JMIR Aging

Journal Impact Factor (JIF) (2023): 5.0
Volume 9 (2026) ISSN 2561-7605 Editor in Chief: Yun Jiang, PhD, MS, RN, FAMIA; Jinjiao Wang, PhD, RN, MPhil

Contents

Original Papers

A Novel Multisegmental Individual-Based Exercise Approach for Postural Dysfunction and Balance Impairment in Older Adults: Randomized Controlled Trial (e78426)
Filza Shoukat, Syed Ur Rehman. . . . 17

Effects of a Combined Zumba and Video Game–Based Cognitive Intervention on Cognitive Function Among Older Adults in Klang Valley, Malaysia: Semirandomized Trial (e88479)
Yook Chia, Sharifah Syed Annuar, Jactty Chew, Yunli Lee, Wei-Gene Lim, Wei Yip, Kristin-Ann Zhe Mun, Michael Jenkins. . . . 31

Motivations, Barriers, and the Potential of AI in Physical Activity Apps for US Older Adults: Interview Study (e96114)
Abhinav Choudhry, Jonathan Mshvelidze, Sean Mullen, Devorah Kletenik, Rachel Adler. . . . 48

Association Between a Hybrid Video-Guided Multicomponent Exercise Program With Staff Assistance and Free-Living Gait in Older Adults With Knee and Low Back Pain: Secondary Analysis of a Randomized Controlled Trial (e97414)
Shuichi Obuchi, Tomoyuki Arai, Hideaki Ishibashi, Takashi Amari, Hisashi Kawai. . . . 65

Community Exergaming and Maintenance-Related Outcomes in Older Adults: Multi-Wave Observational Evaluation (e100007)
Yuezhong Liu, Bing Chia, Herng Koh, Quoc Nguyen, Jiahui Jin, Kimberly Tan, Samuel Long, Yin-Leng Theng. . . . 79

Domain-Specific Associations of Physical Activity With Phase Angle and Extracellular Water–to–Total Body Water Ratio in Middle-Aged or Older Adults: Cross-Sectional Study (e90158)
Seong-Uk Baek, Jin-Ha Yoon. . . . 91

Tailored Online Physical Activity Coaching for Middle-Aged and Older Adults With Cognitive and Mental Health Concerns: Single-Arm Pre-Post Intervention Study (e80040)
Kathryn Ellis, Rhoda Lai, Eleanor Curran, Jennifer Southam, Rebecca Moorhead, Kay Cox, Serafino Mancuso, Alissa Westphal, Terence Chong, Thomas Rego, Victoria Palmer, Kaarin Anstey, Nicola Lautenschlager. . . . 102

Dose-Related Effects of Different Tai Chi Styles Versus Traditional Community-Based Exercises on Cardiometabolic Health and Physical Function in Middle-Aged and Older Adults: Randomized Controlled Trial (e80125)
Jiadong Qiu, Jian Wang, Xiongying Song, Wanyu Shu, Sung Kim. . . . 123

XSL•FO
RenderX

Interactive Vision–Based 3D Augmented Reality System for In-Home Rehabilitation in Older Adults Following Total Joint Replacement: Cross-Sectional Survey of Patients and Service Providers (e87572)
Adriana Salvia, Siona Phadke, Manuela Kunz, Afolasade Fakolade. . . . 142

The 4 Ss Framework for Designing Age-Friendly Communities of the Future: Focus Group Study (e90923)
Inga Hunter, Caroline Lockhart. . . . 158

Smart Speaker–Based Applications to Support Social Connectedness in Older Adult Residents in Affordable Housing: User-Centered Design Study (e90053)
Jane Chung, Natalie Mansion, Tracey Gendron, Rachel Wood, George Demiris. . . . 172

Evaluation of a Gamification-Based Intervention to Enhance Intergenerational Communication: Pre-Post Mixed Methods Evaluation (e77190)
Eliza Wong, Clement Ng, Dorothy Wang, Carol Wong, Shirley Lui, Annie Cheung, Eng-kiong Yeoh. . . . 191

The Influence of Robot Facial Expression and Inclusion Behavior on Rapport and Trust in Older Adults: Mixed-Factorial Experimental Study (e91235)
Hsiu-Ping Yueh, Kuang-Chi Lin, Weijane Lin. . . . 208

Perceived Use of Web-Based Videoconferencing for Social Connectedness Among Older Adults Living in Long-Term Care: Qualitative Study (e73213)
Anna Garnett, Halyna Yurkiv, Denise Connelly, Richard Booth, Lorie Donelle. . . . 222

Association Between Loneliness, eHealth Literacy, and Quality of Life Among Chinese Older Adults: Cross-Sectional Study (e88600)
Shouchuang Zhang, Jiayi Li, Jia Tang, Weiyan Jian, Qing Liu, Jing Guo. . . . 237

Effect of an Intersectoral Telemedicine Intervention on Hospitalization in Nursing Home Patients: Specification Curve Analysis of a Stepped Wedge Cluster Randomized Trial (e75721)
John Grosser, Sophie Pauge, Birthe Aufenberg, Wolfgang Greiner, Miriam Hertwig, Jenny Unterkofler, David Brücken, Christian Hübel, Optimal@NRW Research Group, Jörg Brokmann. . . . 249

Predicting Care Needs in Community-Dwelling Older Adults Using Explainable Machine Learning and a Multidimensional Approach Integrating Health, Social, and Environmental Factors: Cross-Sectional Study (e93371)
Hyeri Shin. . . . 275

Daily Monitoring of Behavioral and Psychological Symptoms of Dementia in Residential Care: Mixed Methods Pilot Study (e98024)
Johannes Malm, Therése Bielsten, Charlotta Nilsen, Elzana Odzakovic, Ingemar Kåreholt. . . . 295

Providing Tech Support as Care Work Among Care Workers in Assisted Living Facilities: Qualitative Interview Study (e80272)
Yiyi Wu, Nicole Ross, Sunyoung Kim. . . . 314

Acceptance and Adoption of the InspireD Reminiscence App Among Care Home Staff: Qualitative Focus Group Study (e91081)
Aoife Conway, Rosemary Bradley, Assumpta Ryan, Claire McCauley, Brighide Lynch, Deirdre Harkin, Sarah Penney. . . . 328

Large Language Model–Based Identification of Subjective Cognitive Decline in Electronic Health Records Among Older Adults With Normal Cognitive Testing: Retrospective Cohort Study (e93011)
Liqin Wang, Rebecca Amariglio, Sheril Varghese, Jiazi Tian, Diane Seger, Li Zhou, Gad Marshall. . . . 341

XSL•FO
RenderX

Development and Temporal External Validation of a Parsimonious, Interpretable Machine Learning Model for Predicting 6-Month Mortality in Long-Term Care Facilities: A Retrospective Cohort Study (e94567)
Yun-Cheng Tsai, Yi-Lin Wu, Yung-Chen Yu, Yi-Wen Chen, Yi-Wen Chen, Yi-Ching Yang, Yu-Tai Lo. 357

A Smart Wheeled Walker With Integrated Physiological Monitoring to Improve Mobility in Older Adults: Randomized Crossover Trial (e99522)
Phurichaya Werasirirat, Pornpimol Muanjai, Oranat Sukkho, Nongnuch Luangpon, Audrius Snieckus, Juntip Namsawang. 382

Leveraging Digital Health Technologies to Assess Older Adults' Frailty and Nutritional Status: Two Cross-Sectional Studies (e77816)
Nunzio Camerlingo, Madisen Wicker, Dimitrios Psaltos, Nina Shaafi-Kabiri, Hao Zhang, Andrew Messere, Isabelle Messina, Zachary Hayden, Meredith Kelly, Gauri Kamat, Fikret Karahano lu, Charmaine Demanuele, Mar Santamaria, David Caouette, Kevin Thomas. 394

Knowledge Performance, User Experience, and Cybersickness With Immersive Virtual Reality for Tai Chi Movement-Encoding in Older Adults: Randomized Comparison Study (e96867)
XiaCheng Song, Lu Sun, Yaqing Song, Huafeng Qu, Jing Jin, Junfeng Zhu, Wenwen Song, Huirong Huang. 413

Assessing a Digital Platform (KOKU-Nut) to Improve Nutrition in Older Adults: Mixed Methods Feasibility Randomized Controlled Trial (e87351)
Chloe French, Sorrel Burden, Jaheeda Gangannagaripalli, Emma Stanmore. 428

Developing Digital Platforms to Help Older Adults Stay Physically and Socially Active: Co-Design Study to Develop a Toolkit and Evaluate Its Impact (e89785)
Leonie Cooper, Hannah Bradwell, Colin Shore, Simone Tomaz, Rory Baxter, Gemma Ryde, Swen Gaudl, Richard Haynes, Catherine Hennessy, Ray Jones, Anna Whittaker, The GOALD Project. 442

Measuring Implicit Attitudes Toward Digital Health Technologies in Older Adults Using an Influence-Aware Affect Misattribution Procedure: Development and Feasibility Study (e87989)
Anna Vinnikova, Jinyue Zhan, Kailing Jin, Xianfeng Ding, Wei Sang, Weina Xu, Qian Yang. 461

Development and Evaluation of a Transportation Planning e-Tool for Older Adults: Co-Design Study (e88297)
Sara Bahrampoor Givi, Dany Baillargeon, Sahar Tahir, Bessam Abdulrazak, Mireille Gagnon-Roy, Mélanie Levasseur, Nathalie Delli-Colli, Patrick Boissy, Mélisa Audet, Véronique Provencher. 476

Usability of a Digital Platform Supporting a Telehealth Intervention for Physical and Social Deconditioning in Older Adults: Mixed Methods Study (e96315)
Nesrine Koubaa, Mélanie Levasseur, Mylène Aubertin-leheudre, Pasquale Roberge, Céline Verchère, Ruth Ndjaboue, Simon Brière, Cinthia Corbin, Patrick Boissy. 498

Inclusive Design Guidelines for Digital Health Interventions for Older Adults: Fuzzy eDelphi Study (e99657)
Kerstin Denecke, Octavio Rivera-Romero, Lana Cvijic, Beatrice Kaufmann, Emilene Zitkus, Yasemin Afacan, Marina Bruderer-Hofstetter, Ruben Debeuf, Joy Goodman-Deane, Laura Haase, Mustafa Ozkaynak, Valérie Santschi, Emanuele Seminerio, Georgia Vardali, Tiina Vuohijoki, Enrique Dorronzoro Zubiete, Carolyn Petersen. 515

Telemonitoring Adoption Among Older Adults: Sequential Mixed Methods Study of Message Framing and Structural Influence (e89817)
Dandan Wang, Huanyu Bao, Amber Qua, Gaelmaine Hoong, Pei He, Ern Ong, Edmund Lee. 536

Usability of the Coach-Supported Dementia Prevention App ENHANCE (Tailored Intervention for Brain Health and Cognitive Enrichment) in Older Adults: 1-Week Mixed Methods Study (e92800)
Tsz Yu, Gill Livingston, Richard Boczko, Kealan Forristal, James Jamison, Carl Leckstein, Vrushti Mehta, Hee Park, Aneesha Singh, Andrew Sommerlad, Sergi Costafreda. 555

XSL•FO
RenderX

Personalized Predictive Model to Predict Subtask Success of Medication Adherence Technologies for Older Adults With Diverse Capabilities: Development and Internal Validation Study (e84616)
Bincy Baby, Ghada Elba, Minzee Kim, SooMin Park, Imra Hudani, Rishabh Sharma, Halak Patel, Sidharth Bajaj, P Patterson, Annette McKinnon, Sara Guilcher, Feng Chang, Linda Lee, Catherine Burns, Ryan Griffin, Joslin Goh, Tejal Patel. 575

Exploring Information Access in Aging Populations and Those With Dementia and Mild Cognitive Impairment in the United Kingdom: Survey and Focus Group Study (e85626)
Claire Rogers, William McGeown, Yashar Moshfeghi. 593

Smartphone Use and Social Activities Among People With Mild to Moderate Dementia: Multi-Informant Cross-Sectional Study (e81927)
Shuji Tsuda, Satoko Hotta, Hiroshige Matsumoto, Kae Ito. 612

Smart Glasses for Older Adults With Cognitive Impairment: Explanatory Mixed Methods Study (e81840)
Brittany Burch, Nayeon Kim, Rachel McPherson, Debora Anokye, Alice Ryan, Odessa Addison, Nirmalya Roy, Chien-Ming Huang, Elizabeth Galik, Barbara Resnick. 626

Development of a Web-Based Resource to Support Driving Safety in Older Adults: Semistructured Interview and Focus Group Study (e79630)
Mollie Cahill, Ranmalee Eramudugolla, Marwin Tang, Diane Hosking, Stephen Cratchley, Michael Regan, Lesley Ross, Jasmine Price, Kaarin Anstey. 642

Web-Based Gamified Auditory-Cognitive Dual-Task Training for Older Adults With Age-Related Hearing Loss: Pilot Randomized Controlled Trial (e84083)
Ivy Zhao, Angela Leung, Chen Li, Laurence Lloyd Parial, Hongming Ma, Jed Montayre, Justin Golub, Robert Sweetow, Janet Ng, Engle Chan.
6 5 3

Usability and Adoption of Smartwatches by Older Adults in Bangladesh: User Study (e76404)
Fariah Mahzabeen, Fazlea Rabby, Neazmul Mowla, Sadia Ibnat. 672

A Tablet Computer–Based Food Record for the Self-Assessment of Nutritional Intake in Patients Undergoing Geriatric Rehabilitation: Prospective Pilot Feasibility Study (e84653)
Julia Berndt, Patrick Elfert, Marco Eichelberg, Juergen Bauer, Andreas Hein, Rebecca Diekmann. 696

Performance-Based Usability of Medication Adherence Technologies Among Older Adults With Diverse Capabilities: Quantitative Study (e88398)
Bincy Baby, Ghada Elba, SooMin Park, Imra Hudani, Rishabh Sharma, Kirk Patterson, Annette McKinnon, Sara Guilcher, Feng Chang, Linda Lee, Catherine Burns, Dagmar Hajducek, Ryan Griffin, Tejal Patel. 710

Predictive Model of Acupuncture Adherence in Alzheimer Disease: Secondary Analysis of Randomized Controlled Trials (e82787)
Ze-Hao Chen, Ran Li, Yu-Hang Jiang, Jia-Kai He, Shan-Shan Yan, Guan-Hua Zong, Zong-Xi Yi, Xin-Yu Ren, Bao-Hui Jia. 731

The Virtual Kitchen Challenge—Version 2: Validation of a Digital Assessment of Everyday Function in Older Adults (e82092)
Marina Kaplan, Moira McKniff, Stephanie Simone, Molly Tassoni, Katherine Hackett, Sophia Holmqvist, Rachel Mis, Kimberly Halberstadter, Riya Chaturvedi, Melissa Rosahl, Giuliana Vallecorsa, Mijiail Serruya, Deborah Drabick, Takehiko Yamaguchi, Tania Giovannetti. 746

Developing Consumer Consensus on Remote Assessment and Management of Physical Function in Older Adults (RAMP): International Modified Delphi Process (e75791)
Elsa Dent, Christopher Hurst, Jack Dalla Via, Jackson Fyfe, Paul Jansons, Eleanor Hayes, Gary Skinner, Marc Sim, Mylene Aubertin-Leheudre, Sabine Britting, Fanny Buckinx, Gavin Connolly, Ruth Dignam, Lora Giangregorio, Jennifer Jones, Pauline Kelly, Robert Kob, Suzanne Morin, Girish Nandakumar, Lucas Orssatto, Maria Pearson, Daniel Pinto, Esmee Reijnierse, Catherine Said, Mohamed Salem, Vina Tan, Rosanna Tran, Jesse Zanker, Robin Daly, David Scott. 766

XSL•FO
RenderX

Effects of a Virtual Reality–Based Natural Environment Intervention on Attention and Mood in Community-Dwelling Older Adults: Randomized Controlled Trial (e87861)
Cheng-Chen Chou, Jong-Ni Lin, Chia-Pin Yu, Chi-Wen Chen, An-Yun Yeh, Yen-Kuang Lin. . . . 784

Evaluating the Use of Generative AI Videos for Health Self-Management of Older Adults: Mixed Methods Study (e88005)
Ting Liu, Patrick Pang, Yiming Luo, Dana McKay, George Buchanan, Shanton Chang. . . . 797

360°-Based Cognitive-Motor Training System for Older Adults With Cognitive Impairment: User-Centered Design and Evaluation Study (e68032)
Francesco Bigotto, Jonathan Panigada, Francesca Bruni, Valentina Mancuso, Silvia Cavedoni, Marco Stramba-Badiale, Karine Goulene, Elisa Pedroli, Silvia Serino. . . . 814

Adoption and Use of a Smart Home Connected Care System by Older Adults: Mixed Methods Study (e76012)
Laura Pemberton, Michelle Carter, Nikolay Mehandjiev. . . . 833

A Multidimensional Digital Health Platform to Support Intrinsic Capacity and Healthy Aging: Pilot Feasibility Study (e79862)
Marcin Kolakowski, Andrea Lupica, Seif Bader, Jaouhar Ayadi, Luca Gilardi, Angelo Consoli, Irina Mocanu, Oana Cramariuc, Lionello Ferrazzini, Eva Reithner, Magdalena Velciu, Barbara Borgogni, Sofia Rivaira, Luca Antognoli, Sara Leonzi, Elisa Felici, Margherita Rampioni, Giacomo Cucchieri, Vera Stara. . . . 848

A Music Interaction System Based on the Emotional Recognition of Older Adults in Smart Homes: Cross-Sectional Experimental Study (e77218)
Chengmin Zhou, Yuxuan Luo, Wenjing Zhan, Ting Huang, Jake Kaner. . . . 872

A Community-in-the-Loop Approach to Smart Home Monitoring for Aging in Place: Mixed Methods Evaluation of a Co-Designed Prototype (e88290)
Roschelle Fritz, Connie Nguyen-Truong, Jennifer Phipps, Ellen Hinderlie, Yolanda Rodriguez, Thai Nguyen, Gabrielle Barling, Anna Mishuk, Heather Schoonover, Christi Zuber, Marilyn Rantz. . . . 900

Smart Home Technology Integration in Home Modification Programs Serving Older Adults: Focus Group Study With Program Grantees (e89729)
Haomin Hu, Moriah Clarke, Meredith Hughes, Palash Sharma, Jack Fried, Portia Singh, Pamela Toto. . . . 915

Feasibility of a Co-Designed Personalized Cognitive Reablement Program Delivered via Digital Voice Assistants for Older Adults Living With Early-Stage Dementia or Mild Cognitive Impairment: Pilot Randomized Controlled Trial (e96922)
Tracy Yiannis, Helen Macpherson, David Scott, Michele Callisaya, Celeste Chapple, Costas Glavas, Caitlyn Gourlay, Ralph Maddison, Eugene Gvozdenko, Paul Jansons. . . . 926

ASSETs for Aging in Place: 12-Month Descriptive Longitudinal Study of Technology-Supported Interprofessional Care Coordination After Nursing Home Discharge (e91524)
Knoo Lee, Erin Robinson, Amy Vogelsmeier, Blaine Reeder, Suzette Bacon, Elizabeth Conrow, Elizabeth Curtis, Ashley Roberts Dulany, Erica Perryman, Lori Popejoy. . . . 943

Alzheimer Disease and Related Dementias Caregiver Burden on Reddit: Large Language Model–Assisted Measurement Study (e103305)
Yan Zhang, Mingfei Zhang. . . . 959

Barriers and Facilitators to Technology Use Among Caregivers of Homebound Older Adults: Descriptive Qualitative Study (e96226)
Carina Katigbak, Esha Shenoy, Padmavathy Ramaswamy, Jessica Lee, Kathrin Milbury, Katie Scott. . . . 980

XSL•FO
RenderX

Use of Digital Tools and Environmental Practices Among Informal Caregivers of Older Adults in the European Union Embracing the Triple Transition: Cross-Sectional Study (e81556)
Beatrice Ietto, Mirko Di Rosa, Giovanni Lamura, Sabrina Quattrini, Ricardo Rodrigues, Sara Santini, Aviad Tur-Sinai, Marco Socci. 993

Public-Facing Communication of Health and Social Services for Older Adults and Their Family or Friend Caregivers: Environmental Scan of 58 Integrated Care Teams' Websites in Ontario, Canada (e80595)
Hung Nguyen, Kyle Corbett, Deanna Vervaecke, Natalia Ivanovski, Soroush Shirazi, Sara Yadollahi, Hom Shrestha, Irene Slavescu, Behnoush Zarandi Baghini, Stella Medvedyuk, Matthias Hoben. 1015

A Telehealth-Adapted Dementia Caregiver Skills Training Intervention (TeleCARE): Single-Arm Pre-Post Intervention Study (e81256)
Maureen O'Connor, Steven Shirk, Jaye McLaren, Andrew Nguyen, Kendra Pugh, Madeline Sullivan, Emily Metcalf, Samantha Harrington, Lauren Moo. 1029

Developing and Testing a Brief Mindfulness Just-in-Time Adaptive Intervention to Reduce Stress Among Caregivers of People With Dementia: Quasi-Experimental Study (e87316)
Patrick Kor, Alex Tsang, Daphne Cheung, Steven Zarit, Haoran Xie, Min Qian, Kay Tan, Yuanqing Zheng, Amanda Chu, Kee Chou. 1044

Combining Subjective Perceptions and Objective Behavioral Metrics With the Elderly Digital Twin System: Quantitative Usability Study (e91873)
Ziaullah Momand, Pornchai Mongkolnam, Debajyoti Pal, Siam Yamsaengsung. 1060

Multidimensional Digital Literacy and Quality of Life Among Informal Care Dyads in Malaysia: Cross-Sectional Survey (e86561)
Fereshteh Mohammadzadeh Yazd, Hui Foong, Rahimah Ibrahim, Puvaneswaran Kunasekaran, Asmidawati Ashari. 1087

Integrating Perceived Psychosocial Outcomes Into the Assessment of Telecare Acceptability Among Slovenian Informal Caregivers of Older People: Cross-Sectional Survey Study (e85534)
Vesna Dolni ar, Nejc Berzelak, Sakari Taipale, Simona Hvali -Touzery. 1106

Characterizing Digital Mindfulness Intervention Utilization and Weekly Assessments in Caregivers of Persons Living With Dementia: Secondary Analysis of a Randomized Controlled Trial (e83618)
Michael Williams, Darby Simon, Morgan Seward, Raquel Tatar, Jennifer Huberty, Ana-Maria Vranceanu, Evan Plys. 1123

Rationale, Design, and Baseline Characteristics of the TARGET (Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments) Cohort of Older Adults for Assessing Fall and Fracture Risk: Prospective Cohort Study (e89285)
Angelique Chan, Abhijit Visaria, Rahul Malhotra, Navrag Singh, Benedikt Helgason, Victor Schinazi, Kok Tan, David Matchar, Rita Sim, Vanessa Koh, Wei Lai, Lakkhina Troeung, Catherine Lim, Sai Pai, Kai Tan, Anitha Praveen, Dheeraj Jha, Stephen Ferguson, Giorgio Colombo, Karolina Minta, Ecosse Lamoureux, William Taylor. 1164

Development and Validation of Machine Learning Models for Predicting Falls Among Hospitalized Older Adults: Retrospective Cross-Sectional Study (e80602)
Xiyao Yang, Juan Ren, Dan Su, Manzhen Bao, Miao Zhang, Xiaoming Chen, Yanhua Li, Zonggui Wang, Xiujing Dai, Zengzeng Wei, Shuiyu Zhang, Yuxin Zhang, Juan Li, Xiaolin Li, Junjin Xu, Nan Mo. 1190

Global, Regional, and National Burden of Falls Among Older Adults Aged 65 Years and Above: Secondary Data Analysis of the Global Burden of Disease Study 2021 (e73802)
Ze Zhang, Yingying Diao, Mingwang Fu, Wantong Han, Haoran Zhou, Ru Fan, Biyun Xu, Bingwei Chen. 1205

Supervised and Self-Directed Technology-Based Dual-Task Exercise Training Program for Older Adults With a History of Falls: Mixed Methods Feasibility Study (e87577)
Prerna Mathur, Afroditi Stathi, Victoria Goodyear, Taylor Krauss, Helen Thomas, Angela Cooper, Philip Kinghorn, Caroline Miller, Natalie Ives, Magdalena Chechlacz, Daisy Wilson, Laura Magill, Shin-Yi Chiou. 1217

XSL•FO
RenderX

Impact of Sleep Quality on Gait Variability: Pilot Cohort Study (e81630)
Francesca Marino, Minzae Kim, Phillip Hwang, Yorghos Tripodis, Michael Alosco, Jesse Mez, Ann McKee, Rhoda Au, Robert Thomas, Ludy Shih. 1232

Feasibility and Preliminary Effects of a Pressure Sensor–Based Exergame on Cognitive and Sensorimotor Frailty in Low-Income Older Adults Receiving Public Community Care Services: Pilot Randomized Controlled Trial (e93467)
Sunmi Song, Suk-Chan Hahm, Seyun Chang, Jinseung Lee, Heewon Kang, Sang Lee, Junesun Kim. 1245

Associations of Midlife Gait Speed and Muscle Performance With Falls in Later Life: Secondary Analysis of the MIDUS Cohort (e103917)
Roee Hayek, Yaniv Nudelman, Dan Keren, Shmuel Springer. 1265

An Ultra-Brief Informant Questionnaire for Case Finding of Cognitive Impairment Across Diverse Literacy: Diagnostic Accuracy Study (e72963)
Tau Liew, King Yip, Kaavya Narasimhalu, Simon Ting, Weishan Li, Sze Tay, Way Koay. 1277

Data-Driven Phenotyping Reveals Nonuniform Association Between Age and Mortality After Aortic Surgery: Retrospective Cohort Study of UK Biobank Data (e75611)
Maria Leinweber, Fadi Taher, Miriam Kliewer, Afshin Assadian, Amun Hofmann. 1297

Associations of Solid Fuel Use and Circadian Rhythm Syndrome With Physical Function and Muscle Strength in Middle-Aged and Older Adults: Nationwide Cohort Study in China (e78352)
Haochen Wang, Pengsen Mou, Donghao Liu, Ze Zhao, Wenzheng Chen, Yuxin Yao, Haoyan Shi, Jiaxin Guan, Jing Dong, Yingliang Wei. . . .
1 3 1 1

Associations Between Indoor Musty Odors and Sleep Quality and Duration Among Older Adults: Longitudinal Study (e107641)
Jinwei Bian, Pui Chau, Kirsten Koehler, Youngmin Cho, Sofia Liu, Junxin Li. 1327

Latent Class Analysis for Identifying Patterns of Digital Device Use and Adaptation Difficulty and Their Associations With Service Needs Among Older Adults: Secondary Data Analysis of a Cross-Sectional Survey Study (e91770)
SangA Lee, Minhee Yang, Jungwon Cho, Jiyoung Kim, Eunhee Cho. 1340

Gaps Between Self-Reported and Documented Use of a Health Care Service Portal Among Older Adults: Cross-Sectional Survey and Retrospective Electronic Data Study (e93057)
Efrat Neter, Diane Levin-Zamir, Ester Eliyahu, Peter Chang, Tuyen Duong, Orna Baron-Epel. 1364

Fear-Specific Technophobia as a Mediator Between Daily Technical Support and Instrumental Versus Hedonic Technology Use in Later Life: Cross-Sectional Survey Study (e93063)
Yu Yang, Shuo Dai, Siyu Yin, Wanyu Xi, Qi Li. 1377

Patient Satisfaction With Telehealth Visits in Rural Compared With Urban Communities: Single-Center Study (e85018)
Corrin Hepburn, Karolina Krawczyk, Cara Joyce, Jonah Rubin. 1400

Digital Self-Efficacy, Satisfaction With the Daily Life Changes Stemming From Digital Transformation, and the Moderating Effect of Social Capital in Middle-Aged and Older Adults: Cross-Sectional Survey Study (e79845)
Jeongeun Choi, Hyeonmi Cho, Hyangkyu Lee. 1409

Digital Inclusion Pathways for Older Chinese Adults in the Context of Active Aging: Secondary Analysis of 2023 China Longitudinal Aging Social Survey Data (e83078)
Bo Zhao, Xiaoyan Wang, Ren Chen, Eun Nam. 1426

XSL•FO
RenderX

Broadband Access, Internet Use, and General Cognition in Middle-Aged and Older Adults: Longitudinal Study (e75947)
Fangfang Cheng, Li Cao. 1441

Factors Associated With the Use of Digital Health and Well-Being Resources in Non–Memory-Led Dementias: Quantitative Survey Study (e85863)
Emilie Brotherhood, Céline El Baou, Caroline Fearn, Oliver Hayes, Nikki Zimmermann, Anastasia Tsipa, Sebastian Crutch, Joshua Stott. . . 1 4 5 7

Attitudes Toward Digital Meal Assistance Services Among Older Adults in China: Cross-Sectional Survey (e84956)
Zhe Chen, Tinghe Deng, Xiaoli Zhu. 1472

Development of the Windmill Model for Mapping Older Adults' Intrinsic Capacity Using Digital Twin Technology: Descriptive Qualitative Study (e81075)
Yirou Niu, Zehui Xuan, Yanling Wang, Xiaoya Zhang, Yibo Guo, Xin Li, Shuai Jin, Jie Zhao, Qian Xiao, Hong Chang. 1486

Segmentation of Older People's Needs and Readiness for Smart Homes by Residentially Based Lifestyles in Spain: Survey Study (e75110)
Jiyeon Yu, Angélica de Antonio, Elena Villalba-Mora. 1502

Web-Based Application to Support Caregivers in the Use of Learning Optimization Methods: Participatory Action Research Study (e76543)
Chantal Viscogliosi, Hubert Ngankam, Romane Duquette-Laplante, Sarah Rahimaly, Hélène Pigot, Bernard Chassé, Sara Ait Bouziaren, Yves Couturier, Jessica Déry, Nicolas Biard, Jordan Duchesne, Dominique Giroux, Sabrina Jacob, Kim Routhier-Chevrier, Jordan Mino-Roy, Mélanie Caron, Véronique Provencher. 1518

Level of eHealth Literacy and Its Associations With Health Behaviors and Outcomes in Chinese Older Adults: Cross-Sectional Analysis of Baseline Data From a Large-Scale Community Project (e74110)
Siu Chau, Wanjia He, Tzu Luk, Sophia Chan. 1531

From Digital Anxiety to Empowerment in Older Adults: Cross-Sectional Survey Study on Psychosocial Drivers of Digital Literacy (e75245)
Han-Jen Niu, Ming-Hsuan Li, Feng-Yu Hsieh, Chun-Chieh Yu, Chun-Ting Lin. 1548

Smartphone Addiction, Use Preferences, and Depression Among Older Adults in the Digital Context: Machine Learning Analysis of Survey Data (e84703)
Sheng Chen, Yue Song, Chien-chung Huang. 1565

Self-Identified Age Cohorts and Personal Experience Sharing in Pseudonymous Online Spaces: Natural Language Processing Analysis of Reddit (e90103)
Ern Khor, Moon Choi. 1577

Developing Customized Personas to Capture Intrinsic Capacity Profiles and Digital Monitoring Intentions in Older Adults: Mixed Methods Study (e82867)
Zehui Xuan, Yirou Niu, Ruifu Kang, Qian Xiao, Shuai Jin, Jie Zhao, Yanling Wang, Hong Chang. 1587

Aging Like Milk or Fine Wine: Content Analysis of Reactions to One's Future Self on TikTok (e70794)
Reuben Ng, Nicole Indran. 1606

Factors Influencing Older Adults' Sustained Engagement in Digital Psychological Interventions in Resource-Limited Settings: Qualitative Study Guided by the Social Ecological Model (e84350)
Jingying Wang, Yating Luo, Sipei Jia, Wenling Yang, Dan Liu, Jiarong Li, Gang Gan, Yi Zhou, Sha Wang, Jianfei Xie. 1615

XSL•FO
RenderX

Real-Time Associations Between Subjective Cognitive Concerns and Daily Activity Participation in Persons Aging With Stroke: Smartphone-Based Ecological Momentary Assessment Study (e80222)
Jiali He, Quoc Bui, Mandy Fong, David Mohr, Christopher Metts, Alex Wong. 1629

Prediction of 2-Year Incident Clinical Knee Osteoarthritis From Extensor Mechanism Mechanical Properties in Asymptomatic Community-Dwelling Older Adults: Prospective Cohort Study (e84494)
Ui-jae Hwang, Zongpan Li, Peng Xia, Tianxiang Fan, Arnold Wong, Oh-yun Kwon, Siu-ngor Fu. 1642

Subsequent Hypertension's Mediation of the Association Between Sleep Duration Trajectories and New-Onset Cardiovascular Disease: Population-Based Cohort Study (e78914)
Hongwei Liu, Minghui Wu, Minheng Zhang, Jian Pan, Haiyan Wu, Haixia Fan. 1655

Deep Learning–Based Estimated Pulmonary Biological Age From Chest Computed Tomography Images in Healthy Adults: Model Development and Validation Study (e78243)
Liping Zuo, Na Zhu, Bowen Wang, Donglai Li, Jinlei Fan, Zhaolei Fan, Yongsheng Shang, Yongxiang Wang, Lei Xu, Peng Zhou, Wangshu Cai, Dexin Yu. 1675

Digital Literacy and Heart Failure Self-Care in Older Patients and Their Caregivers: Dyadic Analysis Using the Actor-Partner Interdependence Model (e85976)
Hokon Kim, Misook Chung, Yeonsoo Jang, JiYeon Choi, Byung Yoo, Sang Chu. 1688

Dynamic Relationship Between Sleep Patterns and Behavioral and Psychological Symptoms of Dementia: Longitudinal Observational Study (e80422)
Eunhee Cho, Sinwoo Hwang, Minhee Yang, Eunkyo Kim, Jungwon Cho, Chang Park. 1701

Experiences of Older Adults and Caregivers With Home Telemonitoring for Heart Failure in Canada: Qualitative Study (e79797)
Guy Paré, Marie-Pierre Moreault, Philippe Voyer, Alexandre Castonguay, Marie-Soleil Hardy, Mickael Ringeval. 1715

Early-Life Experiences as Moderators of the Relationship Between Extreme Heat and Alcohol Consumption among Older Adults: Quantitative Study (e76904)
Hee Lee, Su Shin, Yeon Choi, Eun Choi, Hyunjung Ji. 1731

Global Trends and Research Hot Spots in Medication Regimen Simplification: Bibliometric Analysis (e82274)
Chuwen Dou, Bei Wu, Leyi Wu, An Gu, An Huang, Chen Zhang, Xichenhui Qiu, Jianlin Lou, Lina Wang. 1743

Determinants of Medical Help–Seeking Behavior Following Case Finding of Early Cognitive Impairment: Semistructured Interview Study of Patients and Caregivers (e79386)
Shu Lim, Xiaohui Xin, Gerald Koh, Julian Thumboo, Tau Liew. 1761

A Roadmap for Implementing Virtual Geriatric Mental Health Services for Rural Veterans: Qualitative Longitudinal Evaluation (e95286)
Christine Gould, Marika Humber, Chalise Carlson, Ranak Trivedi, Marisa-Francesca Lindsey, Althea Lloyd, Lynsay Paiko, Amanda Peeples. .
1 7 8 5

Changes in Patterns of Health Care Use Among Older Adults in Singapore Across a Decade: Repeated Nationwide Cross-Sectional Survey (e80157)
Edimansyah Abdin, AshaRani P V, Bernard Tan, Ke Ning, Peizhi Wang, Saleha Shafie, Anitha Jeyagurunathan, Janhavi Vaingankar, Li Ng, Kelvin Tan, Paul McCrone, Siow Chong, Mythily Subramaniam. 1800

Identifying Factors Associated With Patient Portal and Synchronous Telehealth Use Across Age Groups in the Postpandemic Era: Retrospective Analysis of the Health Information National Trends Survey (e83730)
Jaeyoung Park, Su-I Hou. 1814

XSL•FO
RenderX

Digital-Human Public Community Care Integration for Chronic Pain in Low-Income Older Adults in a 6-Week Living Lab Setting: Quasi-Experimental Feasibility Study (e85611)
Sunmi Song, Seo-Yeon Hwang, Hae-Young Kim, Junesun Kim. 1830

Locomotive Syndrome Digital Therapeutics Provided via a Smartphone App: Proof-of-Concept Single-Group Trial Study (e86174)
Tatsuru Sonobe, Itaru Ogawa, Takahiro Seki, Kosuke Watanabe, Yota Kaneko, Takeru Yokota, Taro Mawatari, Satoru Harada, Yasumichi Kadowaki, Youhei Takenaka, Yoshihiro Matsumoto. 1851

Smartphone-Based Field Assessment of Trunk Stability and Its Relationship With Whole-Body Balance in Older Adults: Cross-Sectional Study (e83546)
Javier De Los Ríos-Calonge, Casto Juan-Recio, Amaya Prat-Luri, Aarón Miralles-Iborra, Jose Elvira, David Barbado, Francisco Vera-Garcia. .
1 8 6 3

Feasibility and Acceptability of a Remote Sleep–Dependent Memory Assessment in Older Adults With Cognitive Concerns: Pilot Cross-Sectional Study (e87926)
Clara Tan, Aaron Lam, David Ireland, Dana Bradford, Nathan Cross, Melvern Kurniawan, Simone Simonetti, Jurgen Fripp, Sharon Naismith. .
1 8 7 6

Perceived Algorithmic Care Intensity and Older Adults' Autonomy in Community Smart Care: Mixed Methods Study (e94559)
Xianghui Huang. 1902

Preparation for Future Care in Community-Dwelling Older Adults With Prefrailty and Frailty: Development and Pilot Test of a Digital Intervention (e92492)
Qiqi Ni, Jing Dong, Weilin Jiang, Jia Yi, Ran Yan, Yuhan Chen, Li Wang, Yiyu Zhuang. 1918

Practice Effects on Repeated Self-Administered Digital Cognitive Assessments in Cognitively Intact Older Adults from the Framingham Heart Study: Longitudinal Observational Study (e87208)
Kieffer Christianson, Chenglin Lyu, Cody Karjadi, Sherral Devine, Ting Ang, Ileana De Anda-Duran, Vijaya Kolachalama, Honghuang Lin, Huitong Ding, Katherine Gifford, Rhoda Au, Phillip Hwang, Preeti Sunderaraman. 1938

Detection of Age-Related Cognitive Decline Over Short Time Frames Using Digital Multiday Learning Assessments: Longitudinal Study (e100917)
Iris Blotenberg, Hairin Kim, Daniel Soberanes, Mark Dubbelman, Lex Minto, Onyinye Udeogu, Vivienne Lubbat, Rachel Buckley, Gad Marshall, Keith Johnson, Dorene Rentz, Reisa Sperling, Kathryn Papp, Rebecca Amariglio. 1952

Experiences of Ageism in mHealth App Usage Among Older Adults: Interview Study Among Older Adults Based on Extended Unified Theory of Acceptance and Use of Technology and Risks of Ageism Models (e79457)
Jiayi Sun, Yawen Liu, Chengrui Zhang, Ying Xing, Wanqiong Zhou, Wei Luan. 1964

Designing an mHealth App to Encourage Uptake of Muscle-Strengthening Exercise in Older Adults: Co-Design Focus Group Study (e87332)
Ethan Berry, Nicholas Sculthorpe, Ashley Warner, Nilihan Sanal-Hayes, Lawrence Hayes. 1977

Remote Asynchronous Smartphone 5 Times Sit-to-Stand Testing in Older Adults: Pragmatic Validity Trial (e108345)
Ethan Berry, Lawrence Hayes, Ashley Warner, Nicholas Sculthorpe. 1996

Egocentric Spatial Memory Deficit in Amnestic Mild Cognitive Impairment Revealed Through Virtual Reality: Cross-Sectional Study (e79224)
Cosimo Tuena, Silvia Serino, Chiara Stramba-Badiale, Diana Biondi, Karine Goulene, Marco Stramba-Badiale, Giuseppe Riva. 2008

XSL•FO
RenderX

Long-Term Effects of Mobile-Based Metamemory Cognitive Training in Older Adults With Mild Cognitive Impairment: 15-Month Prospective Single-Arm Longitudinal Study (e81648)
Jung-In Lim, Yeeun Byeon, Sunyoung Kang, Hyeonjin Kim, Keun Kim, Lukas Stenzel, So Jeon, Jun-Young Lee. 2026

Community Readiness Model Applied in Older Adults for the Promotion of Digital Cognitive Training: Mixed Methods Feasibility Study (e69434)
Cíntia Carvalho, Bruno Poltronieri, Beatriz Lima, Rafaela Ferreira, Maria Reis, Christine FitzGerald, Brian Lawlor, Rogério Panizzutti. 2037

Digital Mahjong-Based Cognitive Engagement Intervention for Older Adults: Randomized Controlled Trial (e85339)
Heng-Hsin Tung, Chen-Yuan Kuo, Pei-Lin Lee, Chih-Wen Chang, Kun-Hsien Chou, Katelyn Chen, Ching-Po Lin, Chih-Kuang Liang, Liang-Kung Chen. 2050

Impact of Chatbot-Based Cognitive Training on Cognitive Function and White Matter Integrity in Cognitively Unimpaired Older Adults: Randomized Controlled Trial (e94127)
Bori Kim, Eunhee Kim, Soo Lim, Jinwoo Kim, Jee Jeong, Geon Kim. 2064

Facilitators and Barriers to Digital Self-Management in Older Adults With Depression: COM-B and Theoretical Domain Framework Qualitative Study (e79253)
Ziping Zhu, Xinying Lin, Zhi Chen, Hong Li, Rong Lin. 2082

Response Time Dynamics From Noncognitive Ordinal Ecological Momentary Assessment as a Proxy for Symptom Change in Geriatric Depression: Longitudinal Observational Study (e83891)
Jooho Lee, Jeehang Lee, Sehwan Park, Gangho Do, Jihye Noh, Sangjoon Moon, Kyungmi Chung, Sang Son, Jin Park. 2101

Emotional Regulation and Physiological Responses to a Cultural Heritage Virtual Reality Experience in Older Adults: Mixed Methods Study (e91772)
Yafeng Lai, Pohsun Wang. 2121

Demoralization and Mental Disorders in Community-Dwelling Older Adults: Cluster Analysis (e84657)
Alice Onofrio, Maria Ferrara, Ilaria Domenicano, Martino Belvederi Murri, Tommaso Toffanin, Maria Nanni, Alessandra Canuto, Kerstin Weber, Jana Volkert, Holger Schulz, Martin Härter, Sylke Andreas, Luigi Grassi. 2144

Delirium Occurrence and Associated Factors Among Patients With Dementia and Mild Cognitive Impairment Admitted to Long-Term Care Hospitals: Cross-Sectional Study (e90434)
Kyoung Moon, Yong-Hwan Hyeon. 2158

Assigning Article-Level Themes in Bibliometric Analysis: Mode-Based Mapping Approach Using JMIR Aging Publications (e79906)
Sam Ho, Kang-Ting Tsai. 2258

Integrated High-Throughput Targeted Metabolomics and Machine Learning for Early Prediction and Prevention of Postoperative Delirium in Older Adult Surgical Patients: Prospective Multicenter Cohort Study (e78495)
Gengrui Zhong, Xiaoli Huang, Congye Li, Deqiang Wang, Dingding Huang, Menghan Sun, Quanhong Zhou, Yong Guo. 2724

Web-Based Personalized Machine Learning Recommendations to Enhance Shared Decision-Making in Prostate-Specific Antigen Screening: Randomized Controlled Trial (e83238)
Yi-Ting Lin, Yen Huang, Chih Liu, Hsiao-Yun Cho, Mingchih Chen. 2745

Personalized Digital Care Program Allocation for Older Adults: Reinforcement Learning-Based Simulation Study (e86710)
Hongsoo Kim, Kyungbok Lee, Jae Yi, Myunghee Paik. 2767

XSL•FO
RenderX

Temporal Trends in the Prevalence of Frailty in England From 2002 to 2022, With Forecasts to 2040: Longitudinal Secondary Analysis of a National Cohort Study (e98665)
Dongfeng Tang, Katie Sheehan, Simon Conroy, Aicha Goubar, Julie Whitney, Yanzhong Wang, Matthew O'Connell. 2787

Predicting Frailty Trajectories Using Interpretable Machine Learning Among Older Adults Following Hip Surgery: Prospective Longitudinal Study (e90705)
Jingying Huang, Qinqin Fan, Huiqin Shi, Jiaojiao Chen, Hongbo Gao, Kailai Gu, Jin Yang. 2799

Using Indoor Movement Complexity in Smart Homes to Detect Frailty in Older Adults: Multiple-Methods Case Series Study (e77322)
Katherine Wuestney, Diane Cook, Catherine Van Son, Roschelle Fritz. 2815

Early Identification of Mobility Limitations in Community-Dwelling Middle-Aged and Older Adults: Development of a Prediction Model Based on a Prospective Cohort (e77187)
Alberto Freniche, Wei Hu, Mo Chen, Dantong Wang, Denis Breuillé, Yu-ming Chen. 2833

A Novel Combined Audiovisual-Semantic Digital Tool for Early-Stage Cognitive Decline Detection: Development and Validation Study (e91165)
Qiwei Wu, Binyu Zhao, Zihao Zheng, Betsy Seah, Xueping Zhang, Jing Shao. 2848

Digital Handwriting Kinematics and Physical Performance According to Pentagon-Copy Performance in Community-Dwelling Older Adults: Cross-Sectional Study (e85074)
Pau Ferrer-Ramos, Marcos Faundez-Zanuy, Noemí Serra-Payá, Manuel Garnacho-Castaño, Montserrat Girabent-Farrés. 2862

Machine Learning–Based Biomechanical and Clinical Assessment of the Five Times Chair Rise Test in Older Adults: Sensor-Based Functional Aging Study (e94597)
Anna Schumacher, Ole Meyer, Vincent Quinten, Andreas Hein, Rebecca Diekmann. 2878

Multimodal Dementia Prediction With Large Language Models: Cross-Attention Over Text, Audio, and Image (e93279)
Felix Agbavor, Hualou Liang. 2899

AI-Enhanced Automatic Life Story Structuring for Reminiscence Therapy in Older Adults: Technical Feasibility Study (e83122)
Fang Gui, Mengchen Yang, Liuqi Jin, Jing Qian. 2914

Machine Learning–Derived Cardiovascular Aging Phenotypes From Cardiac Function and Stroke Risk in the UK Biobank: Cohort Study (e77017)
Kang Yuan, Deyan Kong, Jinghui Zhong, Mengdi Xie, Rui Liu, Wen Sun, Xinfeng Liu. 2934

Exploring Expectations, Ethical Dimensions, and Values for Voice-Enabled AI Assistants That Support Older Adults and Informal Family Caregivers: Participatory Speculative Design Study (e79740)
Shweta Singh, Aqueasha Martin-Hammond. 2948

Potential Application Value and Needs of Using Generative AI Videos for Health Management for Older Adults: Qualitative Study (e80661)
Ting Liu, Yiming Luo, Zhenni He, Patrick Pang, Kin Chan, Ying Lau. 2973

Predicting Adherence to Computer-Based Cognitive Training Programs Among Older Adults Using Source-Free Domain Adaptation: Algorithm Development and Validation (e79123)
Ronast Subedi, Shayok Chakraborty, Zhe He, Yuanying Pang, Shenghao Zhang, Mia Lustria, Neil Charness, Walter Boot. 2987

XSL•FO
RenderX

Integrating Care Context With Skeleton and Depth Information for Older Adult Activity Recognition in a Care Facility Using Care-Assessment-Aware Spatiotemporal Transformer: Method and Validation Study (e80102)
Nazmun Nahid, Iqbal Hassan, Md Ahad, Sozo Inoue. . . . 2999

The Ways in Which Stakeholders Make Decisions About AI and Novel Technologies for the Health Care of Older Adults: Qualitative Interview Study (e86148)
Zhang Zhang, Thomas KM Cudjoe, Sato Ashida, Jacqueline Massare, Kacey Chae, Phillip Phan, Peter Abadir, Alicia I Arbaje, Mathias Unberath, Nancy Schoenborn. . . . 3019

Telehealth Barriers and Digital Ageism Experienced by Older Veterans: Formative Ethnographic Study to Inform a Telepharmacy Randomized Trial (e79409)
Maria Venegas, Megan McCullough, Chelsea Hawley, Judith Beizer, William Hung, Lauren Moo. . . . 3035

It's Raining, It's Pouring, the Old Man Is Snoring: Content Analysis of Age Stereotypes in Nursery Rhymes (e70841)
Reuben Ng, Nicole Indran. . . . 3045

Integrating Community and Digital Support Through Social Prescribing to Improve Mental Health in Rural Older Adults in South Korea: Quasi-Experimental Study (e76201)
Bo Zhao, Joo Lim, Subeen Lim, Hye Yoon, Hae Kim, Seok Hong, Xinyu Zhang, Eun Nam. . . . 3056

Online Communication Attitudes and Video Game Co-Play in Older Adults: Cross-Sectional Mediation Study (e80541)
Jeffrey Tseng. . . . 3074

CareMobi to Improve Communication Between Caregivers and Adult Day Centers for People With Dementia: Mixed Methods Feasibility Study (e85090)
Tina Sadarangani, Laura Peralta, Shahrzad Siamdoust, Jackie Finik. . . . 3089

Social Determinants and Oral Health Among Indigenous and Non-Indigenous Older Adults in La Araucanía, Chile: Cross-Sectional Study (e85075)
Fernanda Muñoz-Sepúlveda, Karen Danke Hausdorf, Víctor Beltrán Varas, Claudia Acevedo Atala, Leonardo López, Alejandro Herrera. . . . 3105

Design Preferences for a Digital Advance Care Planning Tool to Support People Living With Dementia: International Stakeholder Survey and Mixed Methods Analysis (e93323)
Ana Miloši , Urška Smrke, Izidor Mlakar. . . . 3131

Reviews

Exergaming Interventions for Preventing Falls and Injurious Falls in Older People: Systematic Review and Meta-Analysis of Randomized Controlled Trials (e89807)
Charlotte Eost-Telling, Lisa McGarrigle, Chunhu Shi, Annemarie Money, Yang Yang, Kimberly Lazo Green, Saima Ahmed, Rachel Christie, Abodunrin Aminu, Kim Delbaere, Eling de Bruin, Emma Stanmore, Chris Todd. . . . 1136

Prediction Models for Frailty in People Living with HIV: Systematic Review (e110499)
Meilian Xie, Yeye Hu, Junjia Yuan, Zhiyun Zhang. . . . 2171

Global Frameworks Toward Effective Integration of Technology in National Dementia Strategies: Narrative Review and Case Study in Mexico (e83861)
Lili Camacho-Bustamante, Claudia Astudillo-García, Guillermo Velázquez López, Luis Moncayo-Samperio, Alberto Palacios-García, Sergio Navarro-Tuch. . . . 2184

XSL•FO
RenderX

Comparative Effectiveness of Digital Exercise-Based Interventions for Improving Balance in Older Adults: Systematic Review and Network Meta-Analysis With GRADE Certainty of Evidence (e101231)
Fangbo Li, Xuecheng Li, Qingfang Li, Xiaoqin Gong, Kexin Shi, Yongyu Huang, Xiaowei Feng, Dong Zhang. . . . 2204

GPS-Based Approaches to Assessing Psychosocial Health in Community-Dwelling Older Adults: Integrative Review (e103996)
JaeWon Hyun, Bomgyeol Kim, Hun Kang, JiYeon Choi. . . . 2231

AI-Based Social Connection Interventions of Emerging Categories of Support for Older Adults: Scoping Review (e99463)
Hana Kim, Cheng Ren, Heather Horton, Jeesoo Jung, Yanjun Dong, Victoria Rizzo. . . . 2242

Application-Based Interventions for Family Caregivers of Older Adults: Scoping Review (e76115)
Xin Lin, Francesca Falzarano, E-Shien Chang, Kristin Hon, Ruoying Zhang, Andy Hickner, Sara Czaja. . . . 2282

Effects of Multisensory Integration Training on Postural Stability Characteristics and Fall Risk in Older Adults: Systematic Review and Meta-Analysis (e80345)
Chaoyu Guo, Lulu Yin, Peng Chen, Jianglong Zhan, Zhongqi Yu, Teck Tan, Yaru Wei, Yixue Gong, Menghan Xu, Van Le, Lin Wang. . . . 2346

Digital and Remote Health for Older Adults in Rural or Underserved Settings: Systematic Review of Clinical, Behavioral, and Implementation Outcomes (e78913)
Víctor Beltrán, Alain Chaple Gil, Constanza Morales-Gómez, Vanessa Campos-Bijit, Rafael Contador, Leonardo Díaz, Jaime Bustos, Eduardo Fernandez. . . . 2360

The Effectiveness of Digital Interventions to Increase Preventive Care Uptake in Older Adults: Systematic Review (e83446)
Lindsay Burton, Kathy Rush, Mindy Smith, Robert Janke. . . . 2374

Effects of Digital-Based Exercise Interventions on Concerns About Falling, Falls Efficacy, and Physical Performance Among Older Adults: Systematic Review and Meta-Analysis (e87070)
Zhaojun Wang, Bochi Zhu, Meng Zhou, Xiaojie Xie, Xueyan Zhang. . . . 2397

The Associations Between Digital Exclusion and Physical or Cognitive Function in Middle-Aged and Older Adults: Systematic Review and Meta-Analysis (e75920)
Jia Yu, Jundan Huang, Shuhan Zhou, Xiao Wan, Qi Xie, Yinan Zhao, Juan Yang, Hui Feng. . . . 2419

Digital Resources and Social Connectedness Among Ethnic Minority Older Adults: Systematic Review and Meta-Analysis (e84962)
Padmore Amoah, Edward Ameyaw, Miron Bhowmik, Prince Adotey, Andrew Donkor. . . . 2442

Humanoid Robot–Assisted Support for Health Care in Older Adults: Systematic Scoping Review (e83849)
Lei Cui, Yufei Li, Xinyao Yang, Xue Liu, Like Zhang, Lili Hou. . . . 2477

Identifying Facilitators and Obstacles in Piloting Dementia Initiatives Within a Living Lab Approach: Systematic Review (e77752)
Sara Alves, Natália Duarte, Teodora Figueiredo, Luis Midao, Maria Bessa, Olalla Rodriguéz, José Fidalgo, Irene Fernandez, Alba Felpete, Maxi Rodriguéz, Juan Bernardéz, David Facal, Joana Carrilho, Elísio Costa. . . . 2494

Key Challenges and Barriers to Digital Literacy for Older Adults: Scoping Review (e80647)
Salman Khan, Sarah Webster, John Puxty, Madison Robertson. . . . 2508

Identifying and Prioritizing Age-Friendly Design Principles and Guidelines for Developing Transportation Planning E-Tools: Scoping Review (e83668)
Sara Bahrampoor Givi, Mireille Gagnon-Roy, Véronique Provencher. . . . 2521

XSL•FO
RenderX

AI-Enabled Smart Glasses for Active Aging: Scoping Review (e81157)
Claudio Delgado-Morales, Juan García-Iglesias, Ana Duarte-Hueros. . . . 2538

Impact of 4 Weeks or More Immersive Virtual Reality on Quality of Life and Physical Activity in Older Adults: Systematic Review and Meta-Analysis (e80820)
Iria Trillo-Charlín, Javier Bravo-Aparicio, Juan Avendaño-Coy, Héctor Beltrán-Alacreu. . . . 2547

Lessons Learned About Digital Health Tool Acceptability Among Rural Older Adults: Systematic Review Guided by the Technology Acceptance Model (e70012)
Zachary Siegel, Ellie Quinkert, Jiya Pai, Corinne Miller, Marquita Lewis. . . . 2565

Service Design Strategies to Enhance Exercise Adherence in Extended Reality Interventions for Older Adults: Systematic Review (e86595)
Jiangpan Niu, Yuanyuan Yin, Shan Wang. . . . 2590

“Living Inside the Walls”: Systematic Review and Qualitative Meta-Synthesis (e78142)
Weitong Li, Ziqi Mei, WenJing Tu, Yulei Song, Yayi Zhao, Yamei Bai, Guihua Xu. . . . 2608

e-Counseling for Fall Prevention in Older Adults: Scoping Review (e78444)
Lina Safarina, Cecep Kosasih, Iqbal Pramukti, Nucki Nursjamsi. . . . 2628

Use of Wearable Technology for Measuring and Characterizing Sedentary Behavior in People With Mild Cognitive Impairment and Dementia: Systematic Review (e85361)
Jenny Wales, Chloe Hinchliffe, Ryan Falck, Silvia Del Din, Alison Yarnall, Ríona Mc Ardle. . . . 2639

Effects of Pure Cognitive Training on Gait and Balance in Older Adults: Systematic Review and Meta-Analysis (e85070)
Xiaohan Li, Yat Chan, Guoshi Liang, Cheuk Lau, Tsz Lam, Chak Lam, Tsz Chan, Hoi Cheung, Wayne Chan, Chun Hsu, Freddy Lam. . . . 2663

AI-Based Monitoring Devices for the Identification of Circadian Parameters as Digital Predictive Biomarkers of Neurodegenerative Diseases in Older Adults: Systematic Review (e83031)
J Herrera-Zamora, Jesus Monroy-Rodriguez, Jesus Garcia-Mancilla, Pedro Santana-Mancilla. . . . 2678

Research Letters

Intergenerational Digital Learning and Relational Outcomes in Junior-Senior Dyads: Exploratory Dyadic Intervention Study (e92774)
May Lwin, Valerie Yu, Gelean Evangelista, Peter Schulz. . . . 1359

Leveraging Multimodal Large Language Models for Fall Risk Reduction in Older Adults in the Home: Proposed Model Design (e77591)
Justin Do, Vivaswat Suresh, Lily Zhang, Bharvi Chavre, Jeremy Cha, Robert Pugliese. . . . 3030

Viewpoints

Rethinking the Prevention of Cognitive Decline in Healthy Older Adults: The Motor-Cognitive Active Gamified Immersive Cultural (MAGIC) Exergames (e100013)
Jean-Jacques Temprado. . . . 2699

Why Frailty Should Be Assessed in Older Adults Affected by Leprosy and How the Use of Smartphones Could Enhance This Health Monitoring (e55963)
Luzielma Gloria, Josafá Barreto, João Oliveira, Areolino Matos, Bianca Callegari, Givago Souza. . . . 2711

XSL•FO
RenderX

Leveraging AI to Advance Age-Friendly Care in the Veterans Health Administration (e75686)
Elizabeth Fine Smilovich, Megha Kalsy, Kimberly Wozneak, Quratulain Syed, Laurence Solberg. 2716

Corrigenda and Addendas

Correction: Speech Recognition–Based Dietary Assessment Tool for Older Adults: Validation and Usability Study (e104170)
Yoonjee Sung, Soyoung Jung, Hae Kang, So Moon, Jee Jeong, Seong Choi, Eun-Hye Lee, Yoo Park. 2780

Correction: Domain-Specific Associations of Physical Activity With Phase Angle and Extracellular Water–to–Total Body Water Ratio in Middle-Aged or Older Adults: Cross-Sectional Study (e114943)
Seong-Uk Baek, Jin-Ha Yoon. 2781

Correction: Development and Validation of Machine Learning Models for Predicting Falls Among Hospitalized Older Adults: Retrospective Cross-Sectional Study (e101284)
Xiyao Yang, Juan Ren, Dan Su, Manzhen Bao, Miao Zhang, Xiaoming Chen, Yanhua Li, Zonggui Wang, Xiujing Dai, Zengzeng Wei, Shuiyu Zhang, Yuxin Zhang, Juan Li, Xiaolin Li, Junjin Xu, Nan Mo. 2783

Correction: Feasibility and Preliminary Effects of a Pressure Sensor–Based Exergame on Cognitive and Sensorimotor Frailty in Low-Income Older Adults Receiving Public Community Care Services: Pilot Randomized Controlled Trial (e109510)
Sunmi Song, Suk-Chan Hahm, Seyun Chang, Jinseung Lee, Heewon Kang, Sanghun Lee, Junesun Kim. 2785

XSL•FO
RenderX

A Novel Multisegmental Individual-Based Exercise Approach for Postural Dysfunction and Balance Impairment in Older Adults: Randomized Controlled Trial

Filza Shoukat*, PhD; Syed Shakil Ur Rehman*, Prof Dr

Faculty of Rehabilitation and Allied Health Sciences, Riphah International University, 6 M Gulberg III Ferozepur Road Lahore, Lahore, Pakistan

*all authors contributed equally

Corresponding Author:
Filza Shoukat, PhD
Faculty of Rehabilitation and Allied Health Sciences, Riphah International University, 6 M Gulberg III Ferozepur Road Lahore, Lahore, Pakistan

Abstract

Background: Age-related decline in mobility and balance substantially increases the risk of falls in older adults. Evidence-based exercise programs are recommended, but their effectiveness requires further evaluation.

Objective: The study aimed to find out the effectiveness of multisegmental individual-based exercise program (MIBEX) approach for postural dysfunction and balance impairment in older adults.

Methods: A 2-arm, single-blinded randomized controlled trial was conducted with 113 community-dwelling adults aged 65 years or older (exercise group, n=56; control group, n=57). Participants in the intervention arm received MIBEX, while the control group received postural education via video. Outcomes included the Timed Up and Go test, tragus-to-wall test, Mini-Balance Evaluation Systems Test, and postural sway, measured at baseline and 3 postintervention time points. Between-group comparisons were analyzed using the Mann-Whitney *U* test with effect sizes (*r*) and 95% CIs.

Results: Significant between-group differences were observed across all outcomes ($P<.01$). The exercise group showed improvements in the Timed Up and Go test (median reduction from 15, IQR 13 - 19 to 10, IQR 8 - 12 s) compared with controls (15, IQR 13-18 to 12, IQR 10-14 s), with a large effect size ($r=0.65$, 95% CI 0.45 - 0.80). Tragus-to-wall test performance improved markedly in the intervention group (median increase from 41, IQR 36-47 to 49, IQR 44-54 cm) versus controls (40, IQR 35-46 to 37, IQR 31-43 cm; $r=0.52$, 95% CI 0.30 - 0.70). Mini-Balance Evaluation Systems Test scores declined in controls but improved in the intervention group ($r=0.48$, 95% CI 0.25 - 0.66). Postural sway demonstrated moderate improvement ($r=0.40$, 95% CI 0.18 - 0.58).

Conclusions: These findings indicate that the MIBEX produced significantly greater improvements in mobility, balance, and postural stability compared with usual care, highlighting its potential as an effective intervention for older adults.

Trial Registration: Iranian Registry of Clinical Trials IRCT20221005056094N1; https://www.irct.ir/trial/66236

(JMIR Aging 2026;9:e78426) doi:10.2196/78426

KEYWORDS

exercise; old age; posture; balance; physical function

Introduction

The 2019 United Nations World Population Prospects data show that approximately 8% of the global population is currently aged 65 years or older, and this figure is projected to reach almost 16% by 2050 [1]. The fastest-growing demographic population is the older adult population aged 80 years and above. Concerns are increasing regarding the increasing proportion of older individuals compared with the total population. It is estimated that by 2030, the percentage of older adults in the total population will increase by 9.3% [2]. According to the Global Burden of Disease Study, the percentage of years lived with disability has increased by 7.9% in men and by 6.5% in women from 1990 - 2017 [3]. Approximately two-thirds of the older adult population lives in low- to middle-income countries, and like many others, Pakistan also experiences a demographic shift. This shift has significant health, social welfare, and economic development implications [4].

Aging is an ongoing and irreversible process resulting in changes in body composition, which affect physical fitness. Physical fitness decreases steadily with age [5]. Sarcopenia, the loss of muscle strength and mass due to biological aging, is a major cause of functional decline, dependency, and increased risk of falls and fall-related injuries in older adults [6]. Aging also

XSL•FO
RenderX

results in various changes in the human body, including changes in body composition, such as increasing body fat, which is linked to a higher prevalence of chronic diseases such as diabetes, chronic lung disease, cancer, and stroke [6,7].

There is strong scientific evidence that regular physical activity provides extensive health benefits for people aged 65 years and above. In some cases, the evidence of health benefits is greater for older people because the negative outcomes associated with inactivity are more common in this age group than in younger individuals [8]. In 2020, the World Health Organization (WHO) updated its recommendations for sedentary behavior and physical activity for children and adolescents, adults, and older individuals.

According to these recommendations, an individual should engage in 150 to 300 minutes of moderate-intensity physical activity, 75 to 150 minutes of vigorous-intensity physical activity per week, or a combination of both [9]. Additionally, individuals are advised to participate in muscle-strengthening exercises targeting all major muscle groups at least twice a week. These recommendations highlight the importance of developing an effective program to promote physical activity for health-related purposes to reduce the risk of chronic illnesses, falls, injuries, morbidity, and premature mortality [10,11].

Different types of exercise programs have been used in previous studies to increase or maintain physical fitness [12] and body composition [13] in the older adult population. Multicomponent or multimodal exercise, which combines various exercise modalities, such as strength, balance, gait, aerobics, and flexibility, into a single exercise session, has gained increasing attention in recent years for improving the physical function of frail older adults [14-16].

Multicomponent physical exercise programs, particularly strength training, are recognized as the most effective approaches to delay the loss of strength [17,18] and functional disability, and to mitigate adverse effects such as hospitalization. Evidence-based systematic reviews and meta-analyses of randomized controlled trials have shown that multicomponent exercises have a positive impact on muscle mass, strength, balance, and physical performance [19-21]. The integration of multimodal exercise with functional balance and mobility training, moderate- to high-intensity progressive resistance training, and 30- to 45-minute sessions at least twice a week for optimal results is recommended [22].

An additional change that affects older adults, particularly females, is the alteration in postural control, which is influenced by hormonal changes, sarcopenia, osteopenia, and reduced physical activity. Postural imbalances contribute to an increased risk of falls [23]. Multicomponent exercises can help prevent falls by enhancing muscular function and postural control [24]. These exercises are recommended for the older population because they improve postural performance and increase muscle strength and power [25,26].

Although multicomponent exercise programs are known to improve physical function in older adults, most apply standardized protocols, and some studies address postural dysfunctions, such as thoracic kyphosis or forward head posture. Few interventions combine segment-specific training with structured postural education, and individualized progression based on functional assessment is rarely implemented. To address this gap, we developed a multisegmental individual-based exercise program (MIBEX), an equipment-free, cost-effective intervention that targets specific anatomical segments related to postural control, integrates postural correction strategies, and adjusts intensity and complexity weekly according to each participant's needs. The objective was to assess its effects on balance, posture, and functional mobility in community-dwelling older adults. We hypothesized greater improvements in postural alignment and balance with this program compared to postural education alone.

Methods

Study Design

This was a 2-arm parallel design, single-blinded, single-center randomized controlled trial approved by the Ethical Committee of Riphah International University (RCRS-RE-PHD-RS/Fall19/127). The trial was registered at the Iranian Registry of Clinical Trials (trial ID IRCT20221005056094N1). The date of registration was October 25, 2022. All participants signed written informed consent before participating in the study. The participants were also informed that the study intervention had no disadvantages or risks. Careful monitoring of the participants was performed throughout the study, and necessary medical care was provided to the participants when needed. The participants were also informed that they could withdraw freely. During the entire study process, the identities of the participants were not disclosed in any publication of this study. The patient may choose not to participate in the study and withdraw their consent to participate at any time and will not be penalized in any way if they choose not to participate or withdraw from this study. The study was conducted according to CONSORT (Consolidated Standards of Reporting Trials) [27], and included the CONSORT checklist (Checklist 1).

Recruitment and Follow-Up Periods

Participant recruitment and all study activities, including follow-up, were conducted between March 2022 and August 2022. The trial comprised a 12-week intervention phase immediately followed by a 12-week no-intervention follow-up period, with final assessments completed in August 2022. All primary and secondary outcomes, including those related to potential benefits and harms, were measured at baseline, during the intervention, post intervention, and at follow-up.

Participants and Study Design

The data were collected at the outpatient department of the University Teaching Hospital, UOL, and in Lahore, Pakistan. A nonprobability convenience sampling technique was used to collect data from the geriatric population. The inclusion criteria were participants aged 65 to 75 years, including both male and female participants; those who could move independently and climb a flight of stairs independently; those who have thoracic kyphosis of 50 degrees or greater; and those who have cognitive function within normal limits, assessed by Mini-Mental State

XSL•FO
RenderX

Examination ≥24. The exclusion criteria were participants who could not maintain balance when standing unassisted; those with diagnosed neurological or vestibular disorders; and those with serious medical conditions that would limit participation in the exercise program, that is, chest pain, myocardial infarction, or cardiac surgery within 6 months. Participants with uncontrolled hypertension and with a history of hip fracture or total hip replacement within the previous 12 months (due to its potential effect on balance and mobility outcomes) and those who used assistive devices for walking (eg, cane or walker) in daily activities were also excluded.

Intervention

The intervention consisted of a 24-week exercise program. The treatment was applied twice weekly for 12 weeks, after which follow-up readings were taken for the next 12 weeks. Data from both the intervention group and the control group were obtained by the researcher at 6 different times: at baseline (T0) for 3 months (T1, T2, and T3), and then for the next 3 months (T4, T5, and T6) for follow-up visits. The details are shown in Table 1.

Table . Data collection and outcome measures scheduled per visit.

Activity/outcome measure	Baseline (wk 0)	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6
Recruitment							
Participant selection	✓						
Allocation	✓						
Intervention							
Group A	✓	✓	✓	✓	✓	✓	✓
Group B	✓	✓	✓	✓	✓	✓	✓
Outcome assessment							
Postural sway	✓	✓	✓	✓	✓	✓	✓
Timed Up and Go test	✓	✓	✓	✓	✓	✓	✓
Mini-BESTest[a]	✓	✓	✓	✓	✓	✓	✓
Tragus-to-wall test	✓	✓	✓	✓	✓	✓	✓

[a]Mini-BESTest: Mini-Balance Evaluation Systems Test.

Procedure

Overview

Baseline measurements of both groups were taken before treatment via a global postural system (GPS400). The GPS was used to measure postural sway as well as static and dynamic balance. Stabilometric assessment was performed via GPS by asking patients to stand on the platform while the participants stood on the postural analyzer for picture acquisition. The distance of the participant from the specific acquisition device was predetermined, and it was an essential requirement. Measurements were carried out on each acquired image, which was saved for future use. Postural sway was measured while the participants stood on a force platform for a 20-second period. This platform measured the displacement of the center of pressure (COP; sway path). Pictures were then taken first with the eyes open and then with the eyes closed. The participants were then asked to move on the platform to measure their dynamic balance. After the initial testing, the participants participated in 1-hour exercise sessions twice weekly for 12 weeks.

The exercise program was divided into 6 sections. Exercises were focused on improving balance, mobility, flexibility, strength, and endurance, as well as providing postural education. To implement the program, the experimental group was supervised at all times by 2 physical therapists who were blinded to which older individuals were in each of the groups. The intensity and complexity of exercises were initially prescribed based on each participant's baseline assessment. Adjustments were made weekly using the rating of perceived exertion (Borg scale) to ensure exercises remained challenging yet safe for all participants. The participants were divided into 2 groups.

Group A (Control Group)

Participants in the control group received postural education delivered twice weekly by a physiotherapist in small groups, with each session lasting 30 - 45 minutes. The sessions used a postural correction video demonstrating proper sitting, standing, and walking posture, as well as exercises to improve body alignment during daily activities. Participants were encouraged to practice the demonstrated postures at home, and the sessions were reinforced with visual aids and *reminders* to ensure adherence. This component aimed to increase awareness of correct posture and support balance and functional performance, complementing the primary outcomes of the study. The postural education program administered to participants in the control group is described in Textbox 1.

XSL•FO
RenderX

Textbox 1. Description of the control group.

Component

- Postural education (control group)

Description

- Proper sitting, standing, and walking posture; exercises to improve body alignment

Frequency and duration

- Twice weekly; 30 - 45 minutes per session

Delivery method

- Conducted by a trained physiotherapist in small groups of 6 - 8; delivered via postural correction video

Group B (Exercise Group) MIBEX

The participants in this group received an exercise program. In accordance with the recommendations of the American College of Sports Medicine [28], every session consisted of the following exercises:

- Warm-up period (10 min)
- Exercise period (35 min)
- Postural education includes instructions for maintaining the correct posture (5 min)
- Cool-down period of stretching and relaxation exercises (10 min)

Warm-Up Period

It includes range-of-motion exercises for the shoulder, hip, knee, wrist, and ankle.

Exercise Period

It includes the following exercises:

Balance Exercises

Balance exercises for both static and dynamic balance were given to the patients. During the static exercises, the participants were instructed to remain in the same position for 10 seconds. After the samples were allowed to rest, the test was repeated 3 times. In the case of dynamic exercises, participants performed different exercises under different conditions that challenged their balance by demanding postural corrections. The exercises were also repeated 3 times.

Strengthening Exercises

Muscle-strengthening training consists of exercises for both the upper and lower limbs:

- Sets and repetitions: The participants performed from one set of 5 to 8 repetitions in the first training session to three sets of 10 to 12 repetitions at the end of the program.
- Intensity and progression: The intensity increased with increasing weight from 0.5 to 1.0 kg, and weight progression was determined at the individual level by the therapist.

Stretching and Flexibility Exercises

The upper limb (neck circles, shoulder circles, triceps stretch, pectoralis major stretch, and trapezius stretch) and the lower limb (quadriceps stretch, calf stretch, hamstring stretch, and back stretch) were used.

Endurance Exercises

These included brisk walking and running, starting first at a slow pace and then gradually increasing the intensity.

Postural Education

Postural education included instructions for maintaining the correct posture (5 min).

Cooling Down Period

The cool-down period consisted of stretching and relaxation exercises (10 min), and a progressive approach was used in all the multicomponent exercises.

A progressive approach was used in all the multicomponent exercises. On the basis of a gradual increase in training volume, sets, and repetitions, we intended to gradually include participants with a variety of different exercise levels.

A detailed description of the exercise program is given in Table 2.

XSL•FO
RenderX

Table . Detailed description of the multisegmental individual-based exercise program.

Exercise protocol	Description of exercises	Duration (min)	Intensity or progression
Warm-up period	• ROM[a] exercises for hip, knee, ankle, shoulder, and wrist	10	• Continue as such for 24 wk
Stretching exercises	• Upper limb: Neck circle, shoulder circle, triceps stretch, pectoralis major stretch, trapezius stretch • Lower limb: Quadriceps stretch, calf stretch, hamstring stretch, and back stretch	10	• Week 1 - 4: 5 min • Week 5 - 8: Continue above • Week 9 - 12: 10 min • Week 13 - 24: Continue above
Strengthening exercises	• Bicep curl, triceps curl, hip extension, hip flexion, hip abduction, and knee extension	10	• Week 1 - 4: 8 reps×2 sets, intensity 65% of 1 RM[b] • Week 5 - 8: 10 reps×3 sets, intensity 75% of 1 RM • Week 9 - 12: 12 reps×3 sets, intensity 85% - 90% of 1 RM • Week 13 - 24: Maintain above
Endurance exercises	• Brisk walk, running, and jogging (individualized; low-impact alternatives such as slow-paced walking if needed)	10	• Week 1 - 4: 10 min at comfortable speed • Week 5 - 8: 5 min warm-up, then increase speed for remaining duration • Week 9 - 12: 10 min at increased speed • Week 13 - 16: 15 min at maximum tolerated speed or intensity
Balance exercises	• Heel raises, 1-leg stand, toe raises, sit-to-stand, marching (hold each for 10 s)	10	• Week 1 - 4: 2-hand support • Week 5 - 8: 2-hand support • Week 9 - 12: 1-hand support • Week 13 - 24: No support
Postural education	—[c]	5	• Twice weekly
Cool-down period	• Stretching and relaxation exercises	5	• Continue as such for 24 wk

[a]ROM: range of motion.

[b]RM: repetition maximum.

[c]Not applicable.

Outcome Measures

Four primary outcome measures were used to evaluate the effects of the MBIEX intervention.

Postural Sway

Postural sway refers to the continuous, small movements of the body's COP during quiet standing, reflecting postural stability and balance control. In this study, postural sway was assessed using the GPS**,** which measures the quantitative measure of overall body sway [29]. Measurements were conducted under eyes-open and eyes-closed conditions while participants stood in a comfortable upright position. Lower sway values indicate greater postural stability and better balance control. The GPS system provides reliable and objective measurements, enabling accurate assessment of subtle changes in postural control [30]. Recent studies have reported good test-retest reliability for COP measures obtained via force platforms (intraclass correlation coefficient, ICC=0.85 - 0.95), supporting the use of GPS in clinical and research settings. Postural sway is widely used as a sensitive outcome to evaluate the effectiveness of interventions aimed at improving balance and reducing fall risk in older adults [31].

Timed Up and Go Test

This test was developed by Podsidalo and Richardson [32]. The functional mobility of independent people was evaluated via the Timed Up and Go (TUG) test. A stopwatch was used to record the time it took to stand from a chair, walk a 3-meter distance at the usual speed, turn around, and sit down again. Timing began when the instructor gave the command to go and ended when the participant returned to the starting position. The TUG test has excellent test-retest reliability (ICC=0.97) [33]. A TUG performance of less than 20 seconds indicates mobility

XSL•FO
RenderX

independence, whereas a performance of more than 30 s indicates dependence [34].

Mini-Balance Evaluation Systems Test

The biomechanical constraints, stability limits or verticality, anticipatory postural adjustments, postural responses, sensory orientation, and stability in gait subsystems of the 6 subsystems that underlie static and dynamic balance are covered by the 14 items of the Mini-Balance Evaluation Systems Test (Mini-BESTest), a clinical assessment tool for measuring balance. The Mini-BESTest's test-retest reliability for people with a range of neurological conditions and balance issues has also been demonstrated to be outstanding (ICC=0.92-0.97) [35]. The maximum score on the 14-item Mini-BESTest, which is graded from 0 to −2, is 28. Only the lower scores for items 3 (single-leg stance) and 6 (compensatory lateral stepping) were used when the overall score was calculated. For item 14, the score should be reduced if a person's gait slows by more than 10% between the TUG with and without a dual task [36].

Tragus-to-Wall Test

It is a reliable test of forward-flexed posture in healthy younger (20 - 30 y) and older (>60 y) adults. The patient stood with their feet separated by the length of the individual's foot; heels 10.0 cm from the wall, and buttocks and posterior thorax against the wall; head looking straight ahead: level being horizontal between the bottom of the nose and the tragus. The distance between the posterior aspect of the tragus and the wall was measured with a ruler [37].

Sample Size

The study sample size was estimated on the basis of the mean and SD reported by Szturm et al [38] for the balance scale. An open-source online calculator (OpenEpi V.3.00) was then used to estimate the total sample size as 91 patients. To account for potential attrition, an anticipated dropout rate of 20% was incorporated, consistent with rates reported in previous longitudinal exercise intervention studies involving older adults [39]. This adjustment was made to minimize the risk of reduced statistical power due to participant loss from health-related issues, personal circumstances, or loss to follow-up. Accordingly, the final estimated sample size was 113 patients, with 56 allocated to the experimental group and 57 to the control group. A 95% CI (α) and 80% power were used in the calculation.

Randomization and Blinding

Participants were recruited using convenience sampling and randomly assigned to 1 of 2 groups using the simple lottery method. Each participant was given a unique identification number by the principal investigator, which was written on identical slips of paper and sealed in opaque envelopes. These envelopes were thoroughly mixed and placed in a box. A research assistant—independent of the recruitment, treatment, and assessment process—randomly drew envelopes to assign participants to either group, thereby ensuring allocation concealment and maintaining a 1:1 ratio between groups.

Due to the nature of the interventions, blinding of the participants and treating therapists was not feasible. However, to minimize detection bias, the outcome assessor remained blinded to group allocation throughout the study. Additionally, the data provided to the statistician were anonymized and coded (eg, as group A and group B) to ensure blinding during analysis.

Statistical Analysis

All statistical analyses were performed using SPSS version 26 (IBM Corp.). Continuous variables such as age, weight, and height were reported as mean (SD). The normality of the data was checked with the Kolmogorov-Smirnov test. For comparisons between groups at each time point (baseline, 3 mo), we used the Mann-Whitney *U* test. For within-group changes over time (eg, baseline vs 3 mo), we applied the Wilcoxon Signed-Rank Test. When more than 2 time points were compared within a group, we first ran the Friedman test, followed by pairwise Wilcoxon tests with Bonferroni correction to adjust for multiple comparisons. In addition to *P* values, effect sizes (*r*) were calculated for all nonparametric tests using the formula $r=Z/\sqrt{N}$ (n=number of observations in the test), along with 95% CIs for *r* to provide an estimate of the magnitude and precision of effects. Statistical significance was set at $P<.05$.

Ethical Considerations

The study was approved by the Ethical Committee of Riphah International University (RCRS-RE-PHD-RS/Fall19/127). All participants signed written informed consent before participating in the study. The subjects were also informed that the study intervention had no disadvantages or risks. Careful monitoring of the participants was performed throughout the study; necessary medical care was also provided to the participants when needed. During the entire study process, the identities of the participants will not be disclosed in any publication of this study. The patient may choose not to participate in the study and withdraw their consent to participate at any time and will not be penalized in any way if they choose not to participate or withdraw from this study. All the participants participated voluntarily in the study.

Completed personal data or other documents containing protected personal health information were kept in a locked file at the principal investigator's office. The data were entered into an electronic deidentified database by authorized study team members and checked for completeness and accuracy. Access to data with identifiers was restricted to authorized study team members and authorities. The electronic data were kept on a server regulated by the local research institute. Data will be destroyed 10 years after the study's finalization or 5 years after the last publication.

Results

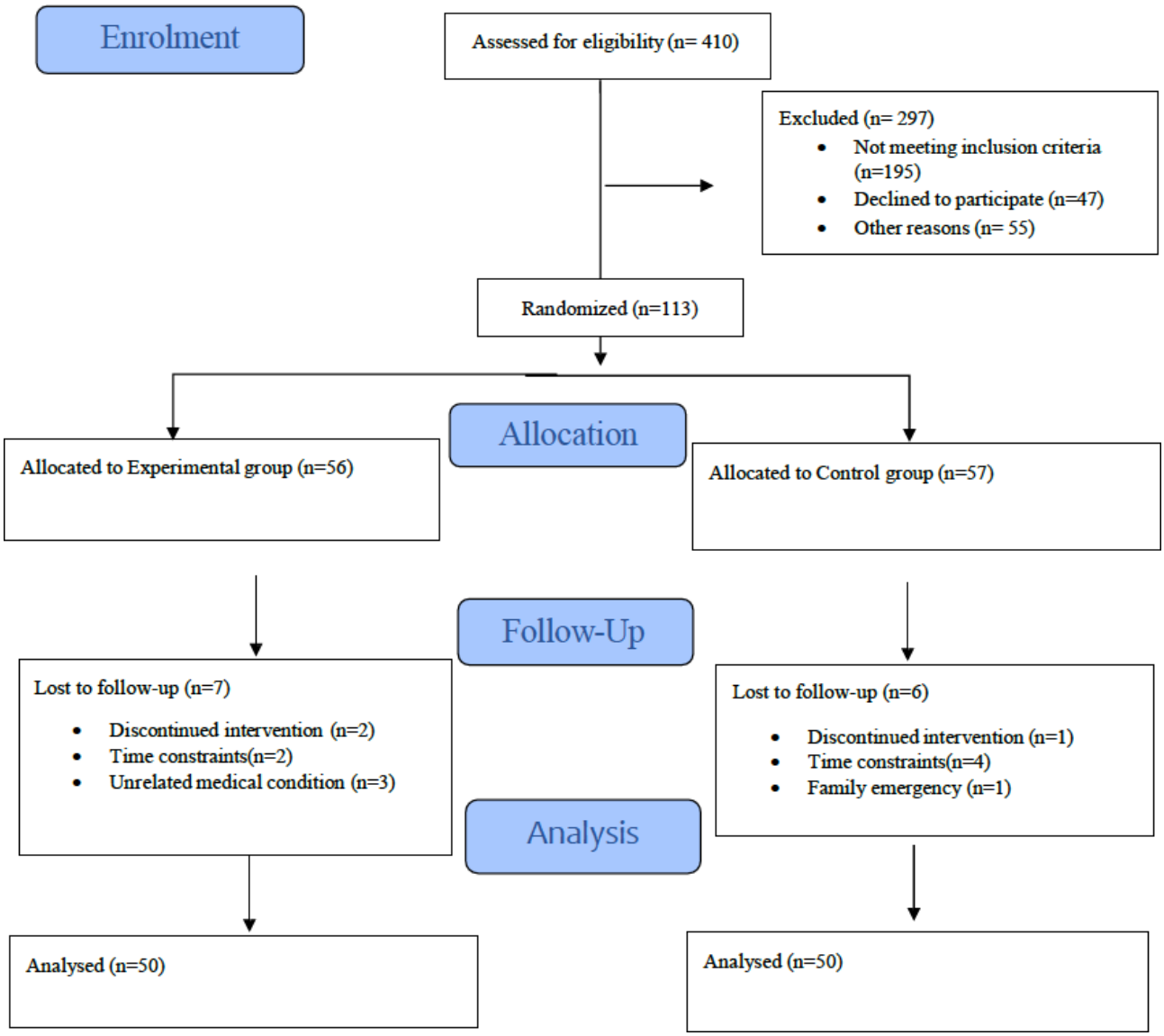
Participant Flow and Baseline Characteristics

Figure 1 provides a detailed overview of participant enrollment, allocation, follow-up, and analysis. A screening of 410 healthy older individuals was performed for participation in the study. After the exclusion of 297 participants, 113 fulfilled the inclusion criteria of the study. Fifty-seven participants were allocated to the control group, and 56 were allocated to the exercise group. During the study, 13 participants dropped out,

XSL•FO
RenderX

and 100 participants completed the study for 24 weeks. The overall schematic flow of the patient sample is given in the CONSORT guidelines. There were no statistically significant differences between the groups in terms of demographic or clinical characteristics at baseline.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.



Basic Demographics

Descriptive statistics of gender, age, sex, weight, height, and BMI were included as the means and SDs and are summarized in Table 3.

Table 3 shows that the groups were comparable in terms of gender distribution, age, height, weight, and kyphosis angle, with no statistically significant differences observed (P>.05). A significant difference was noted for BMI, which was slightly higher in the exercise group compared to the control group (P=.04).

XSL•FO
RenderX

Table . Baseline demographics of study participants.

Variables	Control group	Exercise group	*P* value
Gender, n (%)			
Men	11 (22)	12 (24)	.81
Women	39 (78)	38 (76)	.42
Age (y), mean (SD)	70.36 (3.08)	70.34 (2.99)	.97
Height (cm), mean (SD)	181.0 (14.0)	178.0 (15.0)	.29
Weight (kg), mean (SD)	71.52 (5.36)	72.34 (6.01)	.47
BMI (kg/m²), mean (SD)	25.42 (1.30)	25.99 (1.38)	.04
Kyphosis (°), mean (SD)	51.90 (4.83)	50.78 (5.20)	.27

Table 4 summarizes the between-group comparisons of all primary outcome measures using the Mann-Whitney *U* test. The exercise group demonstrated significantly greater improvements than the control group in TUG (*U*=161.00; *P*<.001; *r*=0.65, 95% CI 0.45 - 0.80), tragus-to-wall test (TWT; *U*=654.00; *P*<.001; *r*=0.52, 95% CI 0.30 - 0.70), Mini-BESTest (*U*=560.00; *P*<.001; *r*=0.48, 95% CI 0.25 - 0.66), and postural sway (*U*=480.00; *P*=.002; *r*=0.40, 95% CI 0.18 - 0.58). The magnitude of effect sizes suggested a moderate-to-large effect of the intervention on mobility and balance outcomes.

Table . Comparison of outcome measures between the groups.

Outcome measure and groups	Baseline, median (IQR)	Postintervention 1, median (IQR)	Postintervention 2, median (IQR)	Postintervention 3, median (IQR)	Mann-Whitney *U*	*P* value	Effect size, *r* (95% CI)
TUG[a] test (s)							
Control	15 (13 - 18)	14 (12 - 16)	13 (11 - 15)	12 (10 - 14)	161.000	<.001	0.65 (0.45 - 0.80)
Exercise	15 (13 - 19)	12 (10 - 14)	11 (9 - 13)	10 (8 - 12)	—[b]	—	—
TWT[c] test (s)							
Control	40 (35 - 46)	39 (34 - 45)	38 (32 - 44)	37 (31 - 43)	654.000	<.001	0.52 (0.30 - 0.70)
Exercise	41 (36 - 47)	45 (40 - 50)	45 (40 - 50)	49 (44 - 54)	—	—	—
Mini-BESTest[d] (score)							
Control	21 (19 - 25)	20 (18 - 23)	19 (17 - 22)	18 (16 - 21)	560.000	<.001	0.48 (0.25 - 0.66)
Exercise	20 (18 - 24)	24 (22 - 28)	26 (23 - 29)	27 (24 - 30)	—	—	—
Postural sway (cm)							
Control	3.2 (2.8 - 3.7)	3.0 (2.6 - 3.5)	2.9 (2.5 - 3.4)	2.8 (2.4 - 3.3)	480.000	.002	0.40 (0.18 - 0.58)
Exercise	3.3 (2.9 - 3.8)	2.6 (2.2 - 3.0)	2.4 (2.0 - 2.8)	2.2 (1.8 - 2.6)	—	.002	0.20 (0.16 - 0.38)

[a]TUG: Timed Up and Go.

[b]Not applicable.

[c]TWT: tragus-to-wall test.

[d]Mini-BESTest: Mini-Balance Evaluation Systems Test.

Table 5 presents the within-group changes in functional performance and postural control outcomes for the exercise and control groups across baseline, month 1, month 2, and month 3 assessments. Both groups demonstrated statistically significant improvements in TUG, TWT, Mini-BESTest, and postural sway (*P*<.001 for most outcomes), with the exercise group showing larger effect sizes (*r*=0.71 - 0.80) compared to the control group (*r*=0.40 - 0.70). The exercise group exhibited the greatest gains in mobility and balance measures. Table 5 shows between-group comparisons of all primary outcome measures using the Mann-Whitney *U* test. At postintervention 3, the exercise group performed significantly better than the control group on the TUG test (median 10, IQR 8-12 vs 12, IQR 10-14 seconds; U=161.00, *P*<.001; *r*=0.65, 95% CI 0.45-0.80), TWT (49, IQR 44-54 vs 37, IQR 31-43 cm; U=654.00, *P*<.001; *r*=0.52, 95% CI 0.30-0.70), Mini-BESTest (27, IQR 24-30 vs 18, IQR 16-21 points; U=560.00, *P*<.001; *r*=0.48, 95% CI 0.25-0.66), and postural sway (2.2, IQR 1.8-2.6 vs 2.8, IQR 2.4-3.3 cm;

XSL•FO
RenderX

U=480.00, *P*=.002; *r*=0.40, 95% CI 0.18-0.58). These findings suggest moderate-to-large intervention effects for mobility and balance outcomes. These results indicated that the MIBEX effectively enhanced postural stability and functional mobility over the intervention period.

Table . Comparison of outcome measures within the group.

Outcome measure and groups	Baseline, median (IQR)	Month 1, median (IQR)	Month 2, median (IQR)	Month 3, median (IQR)	*P* value	Wilcoxon *z*-score	Effect size (*r*)	95% CI for change
Timed Up and Go test (s)								
Control	9.5 (8.6 - 10.3)	6.8 (6.2 - 7.4)	5.1 (4.6 - 5.6)	4.8 (4.3 - 5.3)	<.001	−4.45	0.70 (large)	0.52 - 0.84
Exercise	17.3 (16.4 - 18.3)	15.5 (14.5 - 16.4)	13.9 (13.0 - 14.9)	12.2 (12.0-11.4)	<.001	−5.12	0.80 (large)	0.65 - 0.90
Tragus-to-wall (s)								
Control	3.7 (3.3 - 4.2)	6.6 (5.9 - 7.3)	9.2 (8.7 - 9.7)	10.4 (9.9 - 10.9)	<.001	−4.38	0.66 (large)	0.47 - 0.81
Exercise	9.8 (9.2 - 10.4)	7.4 (6.9 - 7.9)	8.2 (7.7 - 8.7)	6.1 (5.6 - 6.5)	<.001	−4.71	0.71 (large)	0.55 - 0.85
Mini-BESTest[a] (score)								
Control	3.7 (3.3 - 4.2)	6.6 (5.9 - 7.3)	9.2 (8.8 - 9.6)	10.4 (9.6 - 10.2)	<.001	−3.98	0.60 (large)	0.40 - 0.77
Exercise	4.0 (3.5 - 4.5)	3.7 (3.2 - 4.1)	7.9 (7.6 - 8.2)	8.4 (7.1 - 8.3)	<.001	−4.89	0.77 (large)	0.63 - 0.89
Postural sway (cm)								
Control	3.4 (3.0 - 3.8)	2.9 (2.5 - 3.3)	2.6 (2.2 - 3.0)	2.5 (2.1 - 2.9)	<.001	−4.32	0.65 (large)	0.45 - 0.80
Exercise	4.8 (4.4 - 5.2)	3.8 (3.4 - 4.3)	3.5 (3.1 - 4.0)	—[b]	<.001	−4.78	0.72 (large)	0.55 - 0.85

[a]Mini-BESTest: Mini-Balance Evaluation Systems Test.

[b]Not applicable.

Adherence to the Intervention

Adherence to the intervention was tracked using attendance logs maintained by the treating physiotherapist. Participants were encouraged to attend at least 75% of scheduled supervised sessions over the 12-week intervention period, followed by a 12-week follow-up (no adherence data were collected during the follow-up). Adherence rates for both groups during the 12-week intervention period are presented in Table 6. The intervention group maintained high attendance (92% at wk 4% and 87% at wk 12), whereas the control group's adherence declined from 78% to 70% over the same period.

Table . Adherence rates for the intervention and control groups over the 12-week supervised intervention period.

Time point	Intervention group (%)	Control group (%)
T1 (wk 4)	92	78
T2 (wk 8)	89	74
T3 (wk 12)	87	70

Adverse Events

Adverse events were monitored throughout the intervention period by the treating physiotherapist and during each exercise session. Participants were instructed to promptly report any discomfort, injury, or health-related incidents at any time during the study. No adverse events were reported by participants in either the experimental group or the control group during the study period.

Discussion

Principal Findings

This study focused on assessing the impact of a multisystemic exercise program on the physical function of community-dwelling older adults. This study revealed that a multisystemic exercise program, including low-to-moderate intensity resistance, cardiorespiratory, balance, strength, and flexibility training, along with postural education and an increase in moderate-to-vigorous intensity exercise performed twice a week for 24 weeks, significantly improved physical function and reversed the decline in function associated with aging. These

XSL•FO
RenderX

findings support the original hypothesis that an exercise intervention positively and substantially affects physical fitness in community-dwelling older adults.

The participants in this community-based exercise program achieved considerable gains in multiple functional capacities with minimal investment in equipment compared with conventional strength and/or cardiorespiratory training using costly devices [6]. The improvements were similar to those reported in a recent review of 28 studies on resistance training [40]. Consequently, multicomponent exercise programs may be beneficial, as there is a clear connection between high physical activity and low sedentary behavior, as well as improved muscle strength in physically active older adults [41]. Meta-analyses of recent research have shown that multicomponent exercise programs effectively enhance various aspects of health-related physical fitness [14,38].

The multicomponent exercise program significantly improved lower and upper limb strength. This improvement is likely due to the stimulus provided by the exercise program and the adherence to the exercise prescription guidelines [42]. As participants perform strength exercises, they place demands on their muscles, which stimulates muscle fibers and leads to an increase in muscle strength. Progression training increased over the 24 weeks during which participants performed more upper limb and lower limb exercises. These findings align with those of the Beck Jepsen et al study [43], and they support the results of a recent meta-analysis that emphasized the importance of multiple sets in producing benefits in muscle strength [44].

A systematic review of randomized clinical trials with frail older patients revealed conflicting data regarding improvements in the risk of falls, mobility, balance, strength, functional capacity, and body composition [45]. Even among those who showed improvements in functional capacity, attributing the success of these results to a specific element of the intervention protocol is not appropriate.

In addition, multiple systematic reviews have shown that the multicomponent training regimen, comprising sessions that last 30 to 45 minutes and are conducted 3 times per week, is the most commonly recommended protocol for exercise among frail older adults [46]. Specifically, multicomponent exercise programs that incorporate resistance, balance, and flexibility exercises are highly effective for this population [41,42]. Our study is consistent with these findings, as we used 60-minute exercise sessions twice weekly for a 24-week duration. A year ago, a study in which a TWT was used to measure the altered angle was conducted to investigate the effects of exercise programs on age-related kyphosis. A comparison of the preintervention and postintervention values of the occiput-to-wall test revealed a significant difference after the protocol was administered (P<.001) [43], similar to our study's results. Multicomponent exercise programs delivered in primary health care settings have been shown to improve postural control and muscle power in older adults, potentially helping to prevent falls and enhance quality of life, as demonstrated in a study with a similar design to ours conducted on healthy older individuals. The sample size of that study was small, with participants averaging 60 years of age [24].

The main strength of our study is that it uses a multicomponent exercise program (MIBEX), which appears to be a feasible, efficient, and cost-effective option for promoting physical activity among older adults in primary health care settings. This is because the exercises can be performed in groups, require inexpensive equipment, and can be modified to suit local characteristics. Additionally, our study uses a personalized exercise program tailored to individual characteristics and includes postural education.

This study had several limitations. First, it was not possible to blind participants, clinicians, or trainers in this type of exercise program. Although the outcome assessors were blinded to the groupings, some participants may have inadvertently revealed their treatment status, which could have affected the results. This study used a postural teaching program via video for the control group. However, there was no significant difference in the outcome measures since participants did not adhere to the activity, leading to a low rate of participation in the exercise sessions. Despite frequent stimulation and customized work, program adherence fell short of expectations.

Moreover, participation was entirely voluntary, and the main cause for exclusion was when participants refused to participate. As a result, the majority of participants may have established a positive relationship between exercise and health, which could have led to bias. Additionally, we could not determine whether participants who declined to participate would have benefited from physical activity as well.

Future studies should specifically consider the effects of exercise programs on physical performance by gender. Evidence suggests that older women may experience greater declines in physical function and show less pronounced improvements in strength and balance following exercise interventions compared to men [47]. Therefore, future research should account for sex as an important factor when designing and evaluating exercise programs, and may consider including balanced numbers of male and female participants to better assess gender-specific responses.

The current findings can be used to plan and implement future useful public health interventions for community-dwelling older individuals, which could reduce prospective healthcare expenses associated with disabilities among this population. Additionally, primary health care programs should aim to minimize barriers to physical activity promotion. Although recommendations from the literature should be considered in exercise programs for older adults, they may occasionally create barriers to the development of interventions. Nonetheless, socialization and recreational activities seem to be essential components in enhancing adherence among older adults and should be incorporated into multicomponent exercise programs.

Conclusion

The 24-week tailored exercise program had significant positive effects on improving physical function in older adults. The participants exhibited marked improvements in strength, flexibility, balance, and endurance, leading to better daily performance, increased mobility, and a reduced risk of falls. The personalized nature of the program ensured higher

XSL•FO
RenderX

adherence rates and more meaningful outcomes, addressing various aspects of physical function for overall health and well-being.

Acknowledgments

We would like to thank all those who participated in the study. We confirm that generative AI tools were not used. Grammarly was used to correct grammatical mistakes.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets used or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Both authors contributed equally and approved the submitted version.

Conflicts of Interest

None declared.

Checklist 1
CONSORT checklist.
[PDF File, 117 KB - aging_v9i1e78426_app1.pdf]

References

1. Rudnicka E, Napierała P, Podfigurna A, Męczekalski B, Smolarczyk R, Grymowicz M. The World Health Organization (WHO) approach to healthy ageing. Maturitas 2020 Sep;139:6-11. [doi: 10.1016/j.maturitas.2020.05.018] [Medline: 32747042]
2. Fulmer T, Pelton L, Zhang J, Huang W. Age-friendly health systems in an ecosystem perspective. In: Geriatric Medicine: A Person-Centered Evidence-Based Approach: Springer; 2023:1-14. [doi: 10.1007/978-3-030-01782-8_122-1]
3. Haagsma JA, Olij BF, Majdan M, et al. Falls in older aged adults in 22 European countries: incidence, mortality and burden of disease from 1990 to 2017. Inj Prev 2020 Oct;26(Supp 1):i67-i74. [doi: 10.1136/injuryprev-2019-043347] [Medline: 32111726]
4. Noreen N, Abdullah M, Lalani N. Geriatric care in Pakistan: current realities and way forward. Pak J Public Health 2021;11(4):214-219. [doi: 10.32413/pjph.v11i4.872]
5. Dias GNF, Couceiro MS. Active Ageing and Physical Activity: Guidelines, Functional Exercises and Recommendations: Springer; 2017.
6. Rodrigues F, Domingos C, Monteiro D, Morouço P. A review on aging, sarcopenia, falls, and resistance training in community-dwelling older adults. Int J Environ Res Public Health 2022 Jan 13;19(2):874. [doi: 10.3390/ijerph19020874] [Medline: 35055695]
7. Merchant RA, Morley JE, Izquierdo M. Editorial: exercise, aging and frailty: guidelines for increasing function. J Nutr Health Aging 2021;25(4):405-409. [doi: 10.1007/s12603-021-1590-x] [Medline: 33786554]
8. Izquierdo M, Fiatarone Singh M. Promoting resilience in the face of ageing and disease: the central role of exercise and physical activity. Ageing Res Rev 2023 Jul;88:101940. [doi: 10.1016/j.arr.2023.101940] [Medline: 37127094]
9. Okely AD, Kontsevaya A, Ng J, Abdeta C. 2020 WHO guidelines on physical activity and sedentary behavior. Sports Med Health Sci 2021 Jun;3(2):115-118. [doi: 10.1016/j.smhs.2021.05.001] [Medline: 35782159]
10. DiPietro L, Al-Ansari SS, Biddle SJH, et al. Advancing the global physical activity agenda: recommendations for future research by the 2020 WHO physical activity and sedentary behavior guidelines development group. Int J Behav Nutr Phys Act 2020 Nov 26;17(1):143. [doi: 10.1186/s12966-020-01042-2] [Medline: 33239105]
11. Monteiro AM, Rodrigues S, Matos S, Teixeira JE, Barbosa TM, Forte P. The effects of 32 weeks of multicomponent training with different exercises order in elderly women's functional fitness and body composition. Medicina (Kaunas) 2022 Apr 30;58(5):628. [doi: 10.3390/medicina58050628] [Medline: 35630045]
12. Monteiro A, Bartolomeu RF, Forte P, Carvalho J. The effects of three different types of training in functional fitness and body composition in older women. J Sport Health Res 2019;11(3):289-304 [FREE Full text]
13. Marín-Cascales E, Alcaraz PE, Rubio-Arias JA. Effects of 24 weeks of whole body vibration versus multicomponent training on muscle strength and body composition in postmenopausal women: a randomized controlled trial. Rejuvenation Res 2017 Jun;20(3):193-201. [doi: 10.1089/rej.2016.1877] [Medline: 27998241]

XSL•FO
RenderX

14. Li Y, Gao Y, Hu S, et al. Effects of multicomponent exercise on the muscle strength, muscle endurance and balance of frail older adults: a meta-analysis of randomised controlled trials. J Clin Nurs 2023 May;32(9-10):1795-1805. [doi: 10.1111/jocn.16196] [Medline: 34989056]
15. Mahmoudi A, Amirshaghaghi F, Aminzadeh R, Mohamadi Turkmani E. Effect of aerobic, resistance, and combined exercise training on depressive symptoms, quality of life, and muscle strength in healthy older adults: a systematic review and meta-analysis of randomized controlled trials. Biol Res Nurs 2022 Oct;24(4):541-559. [doi: 10.1177/10998004221104850] [Medline: 35619569]
16. Campbell E, Petermann-Rocha F, Welsh P, et al. The effect of exercise on quality of life and activities of daily life in frail older adults: a systematic review of randomised control trials. Exp Gerontol 2021 May;147:111287. [doi: 10.1016/j.exger.2021.111287] [Medline: 33609689]
17. Claudino JG, Afonso J, Sarvestan J, et al. Strength training to prevent falls in older adults: a systematic review with meta-analysis of randomized controlled trials. J Clin Med 2021 Jul 20;10(14):3184. [doi: 10.3390/jcm10143184] [Medline: 34300350]
18. López-López S, Abuín-Porras V, Berlanga LA, et al. Functional mobility and physical fitness are improved through a multicomponent training program in institutionalized older adults. Geroscience 2024 Feb;46(1):1201-1209. [doi: 10.1007/s11357-023-00877-4] [Medline: 37493861]
19. Chen N, He X, Feng Y, Ainsworth BE, Liu Y. Effects of resistance training in healthy older people with sarcopenia: a systematic review and meta-analysis of randomized controlled trials. Eur Rev Aging Phys Act 2021 Nov 11;18(1):23. [doi: 10.1186/s11556-021-00277-7] [Medline: 34763651]
20. de Mello RGB, Dalla Corte RR, Gioscia J, Moriguchi EH. Effects of physical exercise programs on sarcopenia management, dynapenia, and physical performance in the elderly: a systematic review of randomized clinical trials. J Aging Res 2019;2019:1959486. [doi: 10.1155/2019/1959486] [Medline: 31827927]
21. Chen YC, Chen WC, Liu CW, et al. Is moderate resistance training adequate for older adults with sarcopenia? A systematic review and network meta-analysis of RCTs. Eur Rev Aging Phys Act 2023 Nov 29;20(1):22. [doi: 10.1186/s11556-023-00333-4] [Medline: 38030985]
22. McLaughlin EC, El-Kotob R, Chaput JP, et al. Balance and functional training and health in adults: an overview of systematic reviews. Appl Physiol Nutr Metab 2020 Oct;45(10 (Suppl. 2)):S180-S196. [doi: 10.1139/apnm-2020-0279] [Medline: 33054334]
23. Sedaghati P, Goudarzian M, Ahmadabadi S, Tabatabai-Asl SM. The impact of a multicomponent-functional training with postural correction on functional balance in the elderly with a history of falling. J Exp Orthop 2022 Mar 1;9(1):23. [doi: 10.1186/s40634-022-00459-x] [Medline: 35233709]
24. Vieira LA, Cruz JD, Razuk M, Rinaldi NM. Analysis of postural control and muscle function performance of older adults participating in a multicomponent exercise program in primary health care. Geriatr Gerontol Aging 2022;16:e0220007. [doi: 10.53886/gga.e0220007]
25. Tavares DI, Badaró AFV. Perfil da postura corporal em idosos: revisão narrativa [Article in Portuguese]. Revista Kairós-Gerontologia 2017;20(Especial 23):101-118. [doi: 10.23925/2176-901X.2017v20iEspecial23p101-118]
26. Sobrinho ACS, Benjamim CJR, Luciano de Almeida M, et al. Fourteen weeks of multicomponent training associated with flexibility training modifies postural alignment, joint range of motion and modulates blood pressure in physically inactive older women: a randomized clinical trial. Front Physiol 2023;14:1172780. [doi: 10.3389/fphys.2023.1172780] [Medline: 38028788]
27. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg 2012;10(1):28-55. [doi: 10.1016/j.ijsu.2011.10.001] [Medline: 22036893]
28. ACSM's Guidelines for Exercise Testing and Prescription, 9th edition: Lippincott Williams & Wilkins; 2013.
29. Lozano-Quijada C, Poveda-Pagán EJ, Segura-Heras JV, Hernández-Sánchez S, Prieto-Castelló MJ. Changes in postural sway after a single global postural reeducation session in university students: a randomized controlled trial. J Manipulative Physiol Ther 2017 Sep;40(7):467-476. [doi: 10.1016/j.jmpt.2017.06.005] [Medline: 29037786]
30. Prieto TE, Myklebust JB, Hoffmann RG, Lovett EG, Myklebust BM. Measures of postural steadiness: differences between healthy young and elderly adults. IEEE Trans Biomed Eng 1996 Sep;43(9):956-966. [doi: 10.1109/10.532130] [Medline: 9214811]
31. Pajala S, Era P, Koskenvuo M, Kaprio J, Törmäkangas T, Rantanen T. Force platform balance measures as predictors of indoor and outdoor falls in community-dwelling women aged 63-76 years. J Gerontol A Biol Sci Med Sci 2008 Feb;63(2):171-178. [doi: 10.1093/gerona/63.2.171] [Medline: 18314453]
32. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. Is the timed up and go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. BMC Geriatr 2014 Feb 1;14:14. [doi: 10.1186/1471-2318-14-14] [Medline: 24484314]
33. Wall JC, Bell C, Campbell S, Davis J. The timed get-up-and-go test revisited: measurement of the component tasks. J Rehabil Res Dev 2000;37(1):109-113. [Medline: 10847578]
34. Ibrahim A, Singh DKA, Shahar S. "Timed up and go" test: age, gender and cognitive impairment stratified normative values of older adults. PLoS ONE 2017;12(10):e0185641. [doi: 10.1371/journal.pone.0185641] [Medline: 28972994]

XSL•FO
RenderX

35. Anson E, Thompson E, Ma L, Jeka J. Reliability and fall risk detection for the BESTest and Mini-BESTest in older adults. J Geriatr Phys Ther 2019;42(2):81-85. [doi: 10.1519/JPT.0000000000000123] [Medline: 28448278]
36. King L, Horak F. On the Mini-BESTest: scoring and the reporting of total scores. Phys Ther 2013 Apr;93(4):571-575. [doi: 10.2522/ptj.2013.93.4.571] [Medline: 23547173]
37. Tudini FT, Myers BJ, Bohannon RW. Forward flexed posture: reliability and determinants of tragus-to-wall measurement. Physiother Theory Pract 2022 Apr;38(4):579-586. [doi: 10.1080/09593985.2020.1771801] [Medline: 32466740]
38. Szturm T, Betker AL, Moussavi Z, Desai A, Goodman V. Effects of an interactive computer game exercise regimen on balance impairment in frail community-dwelling older adults: a randomized controlled trial. Phys Ther 2011 Oct;91(10):1449-1462. [doi: 10.2522/ptj.20090205] [Medline: 21799138]
39. Theou O, Stathokostas L, Roland KP, et al. The effectiveness of exercise interventions for the management of frailty: a systematic review. J Aging Res 2011 Apr 4;2011:569194. [doi: 10.4061/2011/569194] [Medline: 21584244]
40. Lopez P, Radaelli R, Taaffe DR, et al. Resistance training load effects on muscle hypertrophy and strength gain: systematic review and network meta-analysis. Med Sci Sports Exerc 2021 Jun 1;53(6):1206-1216. [doi: 10.1249/MSS.0000000000002585] [Medline: 33433148]
41. Borges-Machado F, Silva N, Farinatti P, Poton R, Ribeiro Ó, Carvalho J. Effectiveness of multicomponent exercise interventions in older adults with dementia: a meta-analysis. Gerontologist 2021 Nov 15;61(8):e449-e462. [doi: 10.1093/geront/gnaa091] [Medline: 32652005]
42. Marques DL, Neiva HP, Marinho DA, Marques MC. Manipulating the resistance training volume in middle-aged and older adults: a systematic review with meta-analysis of the effects on muscle strength and size, muscle quality, and functional capacity. Sports Med 2023 Feb;53(2):503-518. [doi: 10.1007/s40279-022-01769-x] [Medline: 36307745]
43. Beck Jepsen D, Robinson K, Ogliari G, et al. Predicting falls in older adults: an umbrella review of instruments assessing gait, balance, and functional mobility. BMC Geriatr 2022 Jul 25;22(1):615. [doi: 10.1186/s12877-022-03271-5] [Medline: 35879666]
44. Caldo-Silva A, Furtado GE, Bachi ALL, Barros MP, Zanchi NE, Massart A. Evidence-based effects of multicomponent exercise training on biomarkers related to frailty in older persons: a systematic review. J Phys Educ Sport 2021;21(3):3040-3248. [doi: 10.7752/jpes.2021.s3282]
45. Freire I, Seixas A. Effectiveness of a sensorimotor exercise program on proprioception, balance, muscle strength, functional mobility and risk of falls in older people. Front Physiol 2024;15:1309161. [doi: 10.3389/fphys.2024.1309161] [Medline: 38694207]
46. Lee YC, Chang SF, Kao CY, Tsai HC. Muscle strength, physical fitness, balance, and walking ability at risk of fall for prefrail older people. Biomed Res Int 2022;2022(1):4581126. [doi: 10.1155/2022/4581126] [Medline: 36531652]
47. Fernández-Rodríguez R, Martínez-Vizcaíno V, Reina-Gutiérrez S, et al. Sex differences in effects of exercise on physical function in aging: a systematic review with meta-analysis. World J Mens Health 2024 Oct;42(4):694-711. [doi: 10.5534/wjmh.230257] [Medline: 38449450]

Abbreviations

CONSORT : Consolidated Standards of Reporting Trials
COP: center of pressure
GPS : global postural system
ICC: intraclass correlation coefficient
MIBEX : multisegmental individual-based exercise program
Mini-BESTest: Mini-Balance Evaluation Systems Test
TUG: Timed Up and Go
TWT : tragus-to-wall test
WHO: World Health Organization

Edited by P Jansons; submitted 02.Jun.2025; peer-reviewed by A Hirata, DS Anwar, MC Espada, M Bayattork, PDA Ahmed; revised version received 20.Nov.2025; accepted 20.Nov.2025; published 09.Jul.2026.

Please cite as:
Shoukat F, Ur Rehman SS
A Novel Multisegmental Individual-Based Exercise Approach for Postural Dysfunction and Balance Impairment in Older Adults: Randomized Controlled Trial
JMIR Aging 2026;9:e78426
URL: https://aging.jmir.org/2026/1/e78426
doi:10.2196/78426

XSL•FO
RenderX

© Filza Shoukat, Syed Shakil Ur Rehman. Originally published in JMIR Aging (https://aging.jmir.org), 9.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of a Combined Zumba and Video Game–Based Cognitive Intervention on Cognitive Function Among Older Adults in Klang Valley, Malaysia: Semirandomized Trial

Yook Chin Chia[1,2,3], MBBS FRCP; Sharifah Nadhira Syed Annuar[3], PhD; Jactty Chew[3,4], PhD; Yunli Lee[3,5], PhD; Wei-Gene Lim[3,4], MSc; Wei Zern Yip[6], BSc; Kristin-Ann Leong Zhe Mun[3], BSc; Michael Jenkins[7], PhD

[1]Department of Clinical Medicine and Surgery, Jeffrey Cheah Sunway Medical School, Faculty of Medical and Life Sciences, Sunway University, No. 5, Jalan Universiti, Bandar Sunway, 47500 Selangor Darul Ehsan, Malaysia, Petaling Jaya, Selangor, Malaysia

[2]Department of Primary Care Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Kuala Lumpur, Malaysia

[3]Ageing, Health and Well-Being Research Centre, Jeffrey Cheah Sunway Medical School, Faculty of Medical and Life Sciences, Sunway University, Petaling Jaya, Selangor, Malaysia

[4]Department of Biomedical Sciences, Jeffrey Cheah Sunway Medical School, Faculty of Medical and Life Sciences, Sunway University, Petaling Jaya, Selangor, Malaysia

[5]Research Centre for Human-Machine Collaboration (HUMAC), Faculty of Engineering and Technology, Sunway University, Petaling Jaya, Selangor, Malaysia

[6]School of Psychology, Faculty of Medical and Life Sciences, Sunway University, Petaling Jaya, Selangor, Malaysia

[7]School of Psychology, Faculty of Science and Engineering, University of Nottingham Malaysia, Semenyih, Selangor, Malaysia

Corresponding Author:
Yook Chin Chia, MBBS FRCP
Department of Clinical Medicine and Surgery, Jeffrey Cheah Sunway Medical School, Faculty of Medical and Life Sciences, Sunway University, No. 5, Jalan Universiti, Bandar Sunway, 47500 Selangor Darul Ehsan, Malaysia, Petaling Jaya, Selangor, Malaysia

Abstract

Background: Age-related cognitive decline can threaten independence in older adults, creating an urgent need for effective and practical preventive strategies. Nonpharmacological approaches such as physical activity, cognitive stimulation, and combined programs show promise, but their comparative effectiveness and the specific cognitive domains they influence are not yet clearly established.

Objective: This study evaluated the effects of 3 nonpharmacological interventions. These consisted of a physical activity program using Zumba, a video game–based cognitive stimulation program, and a combined physical and cognitive program. The aim was to determine their effects on global cognition and specific cognitive domains in community-dwelling older adults in the Klang Valley of Malaysia. A secondary objective was to examine the influence of adherence on cognitive outcomes.

Methods: A total of 96 participants (median age 69 y, IQR 65-72 y; n=63, 65.6% female) were assigned to 4 groups: physical (n=12), cognitive (n=23), combined (n=43), and control (n=18). Global cognitive performance was measured using the Montreal Cognitive Assessment at baseline and postintervention. Mixed ANOVA examined time and time × group effects. Additional nonparametric analyses (Kruskal-Wallis, Mann-Whitney *U*, Wilcoxon signed rank tests) were conducted under intention-to-treat and per-protocol (PP) frameworks. Adherence was defined as attending 50% or more of scheduled sessions, with PP analyses reclassifying participants accordingly.

Results: Global cognition improved significantly over time across all groups (time main effect: $F_{1,92}$=7.03, P=.008; time × group interaction: $F_{3,92}$=0.57, P=.64). PP analyses showed moderate gains for participants attending 50% or more of cognitive sessions (r=0.46) and larger gains for those adhering to the combined intervention (r=0.66), with median Montreal Cognitive Assessment scores increasing by approximately 2 points in both adherent groups. Subdomain analyses indicated improvements in language (all groups), visuospatial function (combined group), and orientation (cognitive group).

Conclusions: All intervention types were associated with improvements in cognitive performance over time, with comparable gains across groups. Participants who achieved at least moderate adherence (≥50%) to the cognitive or combined interventions demonstrated larger and more consistent within-group effect sizes, highlighting the importance of sustained engagement. These findings support the feasibility and potential cognitive benefits of structured, multimodal nonpharmacological interventions for community-dwelling older adults.

Trial Registration: ClinicalTrials.gov NCT06376656; http://clinicaltrials.gov/ct2/show/NCT06376656

XSL•FO
RenderX

(JMIR Aging 2026;9:e88479) doi:10.2196/88479

KEYWORDS

cognitive aging; physical activity; cognitive stimulation; combined intervention; Montreal Cognitive Assessment; MoCA; older adults

Introduction

Malaysia is rapidly becoming an aging society. The proportion of the population aged 60 years and older is projected to increase from 11.6% in 2024 to 17.3% by 2040, surpassing the 15% threshold commonly used to define an aging population [1]. This demographic transition reflects a global trend, as many high- and middle-income countries face similar shifts toward older age structures [2]. Based on these projections, Malaysia will officially be classified as an aging society by 2040.

Globally, the prevalence of mild cognitive impairment (MCI) has been estimated at 19.7% among individuals aged 50 years and older [3]. In Malaysia, previous studies have reported even higher estimates, with approximately 21.1% of older adults affected [4,5]. More recently, a study conducted in Kuantan, Malaysia, found the prevalence of MCI among adults aged 60 years and older to be 18.7% [6]. Despite these substantial prevalence rates, most individuals remain undiagnosed until later stages of cognitive decline, with only 8% of those in the early stage of MCI receiving a formal diagnosis [7]. These findings have important implications for independence, health care use, and caregiver burden, contributing to the rising societal impact of dementia [8].

Lifestyle-based interventions are increasingly recognized as promising strategies for delaying cognitive decline, particularly given the limited effectiveness of pharmacological treatments [9]. Aerobic exercise has been associated with improvements in brain structure and function through mechanisms such as enhanced cerebral blood flow and improved regulatory effects on the brain. Sustained aerobic exercise increases heart rate, boosts cerebral perfusion, and promotes glucose metabolism in the brain [10]. Among the various modalities, Zumba represents a culturally engaging form of exercise that combines rhythm, coordination, and social interaction. Participation in Zumba classes has been shown to improve quality of life and visuospatial memory among older adults [11,12]. Zumba has been proposed as a feasible public health strategy due to its enjoyable and socially engaging nature [13], and its increasing popularity within Malaysian community settings further underscores its potential for broader implementation.

Parallel to this, technology-driven cognitive training programs, including video game–based interventions, have emerged as scalable and accessible tools for older adults. Evidence from computerized and video game–based cognitive training indicates that overall effects on cognitive performance in healthy older adults are positive but small. Reported benefits vary substantially across cognitive outcomes and are strongly influenced by intervention design characteristics, including delivery format and training intensity [14]. Improvements in specific domains are inconsistent, and transfer to global cognition remains variable, underscoring the importance of task design and intervention structure. With increasing digital literacy among urban-dwelling Malaysian older adults, including those in the Klang Valley, such approaches are becoming increasingly feasible.

Emerging research suggests that combining physical and cognitive interventions may have synergistic effects, as aerobic exercise provides a neurophysiological foundation for plasticity, while cognitive training directly stimulates neural networks [15]. However, studies directly comparing these approaches, particularly in Asian populations, remain limited.

The present semirandomized controlled trial was designed to address these gaps in the Malaysian context. Older adults living in the Klang Valley were allocated into physical activity (Zumba), cognitive stimulation (video game), combined intervention, or control groups within community settings. This study evaluated the effects of the interventions on global cognitive function using the Montreal Cognitive Assessment (MoCA). We also explored changes across cognitive subdomains, examined whether attendance influenced outcomes, and identified participants who demonstrated meaningful improvements, defined as achieving the minimal clinically important difference of a ≥2-point increase in MoCA scores [16]. Thus, this study aimed to evaluate the effects of Zumba, video game–based cognitive stimulation, and combined interventions on global cognitive function among older adults in Klang Valley, Malaysia, and to examine the influence of adherence on cognitive outcomes, with the hypothesis that the combined intervention would produce the greatest improvements.

Methods

Study Design and Setting

This semirandomized controlled trial was conducted as part of the ongoing MyAgeWell project, which investigates lifestyle-based strategies to promote cognitive health among older adults in the Klang Valley, Malaysia. A total of 400 community-dwelling adults aged 60 years or older completed baseline assessments within the MyAgeWell cohort. Participants were recruited from senior centers and residential communities across the Klang Valley [17,18].

Of these, a subset was invited to participate in the present 6-month intervention trial based on willingness to engage in structured intervention sessions, availability during the intervention period, and feasibility considerations specific to the intervention modalities (eg, access to internet-enabled devices for online cognitive sessions). Individuals who did not meet these requirements or declined participation continued in the observational cohort but were not enrolled in the intervention trial.

XSL•FO
RenderX

Eligible participants were allocated to 1 of 4 groups (Zumba, video game–based cognitive stimulation, combined intervention, or control group) using a semirandomized (quasi-randomized) allocation approach. Participants were primarily assigned in a sequential cycle to maintain approximate balance in group sizes. Minor adjustments were made in 24% (23/96) of participants to accommodate the logistical feasibility of intervention delivery. No allocation concealment was implemented, and neither participants nor outcome assessors were blinded to group assignment. All intervention arms shared identical eligibility criteria, and group assignment differed only by allocation procedure rather than differential enrollment criteria.

This trial evaluated the effects of weekly physical exercise (Zumba), video game–based cognitive stimulation, and a combined intervention over 6 months, with outcomes assessed both between groups (relative to a control group) and within groups (preintervention and postintervention).

Participants

Eligible participants were cognitively normal or without cognitive impairment, including those with well-managed medical conditions such as hypertension. All participants were required to have some mobility, defined as the ability to walk at least 3 m independently, and to be able to communicate in at least 1 of the following languages: English, Malay, Mandarin, or Tamil.

Individuals were excluded if they had a diagnosis of neurodegenerative diseases such as Alzheimer or Parkinson disease, psychiatric disorders, uncorrected hearing or vision impairments, or were currently taking psychiatric medications. Participants who were immobile, required full-time caregiver assistance, or had comprehension difficulties that could interfere with study assessments were also excluded.

Ethical Considerations

All procedures were explained to the participants, and written informed consent was obtained from all participants before study participation in accordance with the Declaration of Helsinki. The study was approved by the Sunway University Research Ethics Committee (SUREC2020/039) and registered at ClinicalTrials.gov (NCT06376656). Participant privacy and confidentiality were protected throughout the study, and all collected data were deidentified prior to analysis. Participants were informed of the study procedures, potential risks, and their right to withdraw at any time without penalty. Participants received cash compensation of RM 50 (US $12.09) for baseline and postassessment completion, RM 30 (US $7.25) per physical intervention session, and RM 10 (US $2.42) per online cognitive intervention session.

Sample Size Considerations

This study included all eligible participants from the ongoing MyAgeWell cohort who consented to join the 6-month intervention. No formal a priori power calculation was performed, and the sample size was determined by feasibility and participant availability within the study timeline.

Intervention Procedures

Adherence Assessment and Per-Protocol Classification

No participants withdrew during the 6-month intervention, and complete MoCA data were available for all participants. Per-protocol (PP) groups were defined post hoc by applying an adherence threshold of ≥50% attendance calculated across the full 6-month intervention period for each intervention component. Participants not meeting this threshold were assigned to the control or low-adherence PP group. No participants were removed, reassigned, or reallocated during the intervention period based on adherence; all reclassifications occurred analytically at the PP stage.

Zumba Exercise

The Zumba intervention was delivered in 19 weekly sessions of approximately 60 minutes each at View Gallery, an activity space at Sunway University. The sessions were low impact, culturally adapted, and designed to emphasize coordination, flexibility, balance, and moderate-intensity aerobic activity. All sessions were led by certified instructors trained in working with older adults. Attendance was carefully recorded to monitor adherence.

Video Game–Based Cognitive Stimulation

The video game–based cognitive intervention consisted of 24 weekly online sessions delivered via Zoom, each lasting 60 minutes. Participants were asked to download and use the video games on their personal devices (Android or iOS). Two commercially available games were used: Tetris, developed by Alexey Pajitnov, targeting visuospatial reasoning and problem-solving, and Aim Lab, developed by Statespace Inc., designed to train attention, processing speed, and hand-eye coordination. Attendance and performance were monitored via Zoom screen sharing, score reporting, and session time logs. The research team did not develop or own any of the software, did not receive funding or sponsorship from the software developers, and had no other financial or personal relationships with the developers. All potential conflicts of interest are disclosed in the *Conflicts of Interest* section.

Combined Intervention

Participants in the combined intervention group received both physical and cognitive training components. Specifically, participants attended one 60-minute Zumba session per week delivered in person and one separate 60-minute video game–based cognitive stimulation session per week delivered online via Zoom. The 2 sessions were scheduled on different days to reduce fatigue and maximize engagement. Thus, while each intervention component was matched in duration to its corresponding single-modality arm, the combined group received a greater total weekly intervention dose (120 min/wk).

Control Group

Participants in the control group continued with their usual daily routines and did not participate in structured exercise or cognitive stimulation.

XSL•FO
RenderX

Outcome Measures

Cognitive function was assessed using the MoCA before and after the intervention. The MoCA, developed and published in 2005 as a brief cognitive screening tool with high sensitivity and specificity [19], assesses global cognitive function across 7 subdomains: visuospatial or executive function, naming, attention, language, abstraction, delayed recall, and orientation [20]. Both total and subdomain scores were analyzed to determine the magnitude and domain specificity of cognitive changes. Attendance was recorded for all sessions, with adherence calculated as the percentage of sessions attended relative to the total scheduled sessions.

Statistical Analysis

Data were analyzed using IBM SPSS Statistics (version 27). Descriptive statistics summarized participants' demographic characteristics. Normality was assessed using the Shapiro-Wilk test, skewness, and kurtosis. Although MoCA before and MoCA after scores showed mild deviations from normality, the MoCA change score was approximately symmetric (skewness=0.13, kurtosis=0.30, Shapiro-Wilk P=.02), and Levene test indicated homogeneity of variances (P=.15). Given the sample size (n=96) and robustness of ANOVA to mild nonnormality, both parametric and nonparametric analyses were conducted.

A 2-way mixed ANOVA (time × group) was performed to examine changes in MoCA scores from baseline to postintervention across the 4 groups. Kruskal-Wallis and Wilcoxon signed rank tests were conducted as nonparametric sensitivity analyses to confirm the robustness of the findings.

Between-group comparisons of change scores were performed using Kruskal-Wallis tests, followed by Mann-Whitney U tests with Bonferroni correction for pairwise comparisons. Effect sizes were reported as $r=|Z|/\sqrt{N}$ for Wilcoxon and Mann-Whitney U tests and as η^2 for Kruskal-Wallis tests, interpreted as small (0.1), medium (0.3), or large (0.5).

Associations between attendance and MoCA change were examined using Spearman rank correlation. To further explore the predictive value of attendance, multiple linear regression analyses were conducted to assess whether attendance variables significantly predicted changes in MoCA scores. In addition, receiver operating characteristic (ROC) analyses were performed to evaluate the discriminative ability of attendance percentage in identifying cognitive responders (defined as ≥2-point MoCA improvement). Statistical significance was set at P<.05.

Results

Participant Flow and Group Allocation (Intention-to-Treat and PP Analyses)

Participant Flow Overview

A subset of 96 participants from the MyAgeWell cohort participated in the 6-month intervention involving physical activity (Zumba), cognitive stimulation (video game–based), or combined physical and cognitive stimulation, following selection based on willingness, availability, and intervention-specific feasibility requirements.

Participants were allocated to physical activity (Zumba; n=12), cognitive stimulation (video game; n=20), combination (n=43), and control (n=21) groups (intention-to-treat [ITT] sample, N=96). There were no dropouts during the intervention period and no missing postintervention MoCA assessments. All 96 participants completed preintervention and postintervention MoCA testing and were included in the ITT analyses.

For the PP analysis, participants were reassigned based on actual adherence, using a ≥50% attendance threshold for each intervention. This threshold was chosen as a pragmatic cutoff, representing moderate engagement while retaining an adequate sample size for analysis. Adherence at the ≥50% threshold corresponded to attending at least 10 of 19 Zumba sessions, 12 of 24 cognitive sessions, or both in the combined intervention, with each session lasting 60 minutes weekly. Although higher thresholds (eg, 70% - 80%) [21] are commonly applied in medication adherence and clinical trials as a conservative definition of "good adherence," previous studies on lifestyle and cognitive interventions have reported wide variability in attendance rates, with thresholds ranging from 50% to 80% [22,23]. Sensitivity analyses at higher thresholds in this study showed similar overall patterns but reduced analytic stability. The ≥50% attendance threshold was applied only after completion of the intervention to define analytic PP groups and did not influence intervention delivery or participant participation. Participants who did not meet the ≥50% attendance threshold, including those originally allocated to the control arm and those with partial but insufficient intervention exposure, were classified into a control or low-adherence PP group. Exposure below this level was considered unlikely to produce a meaningful intervention effect.

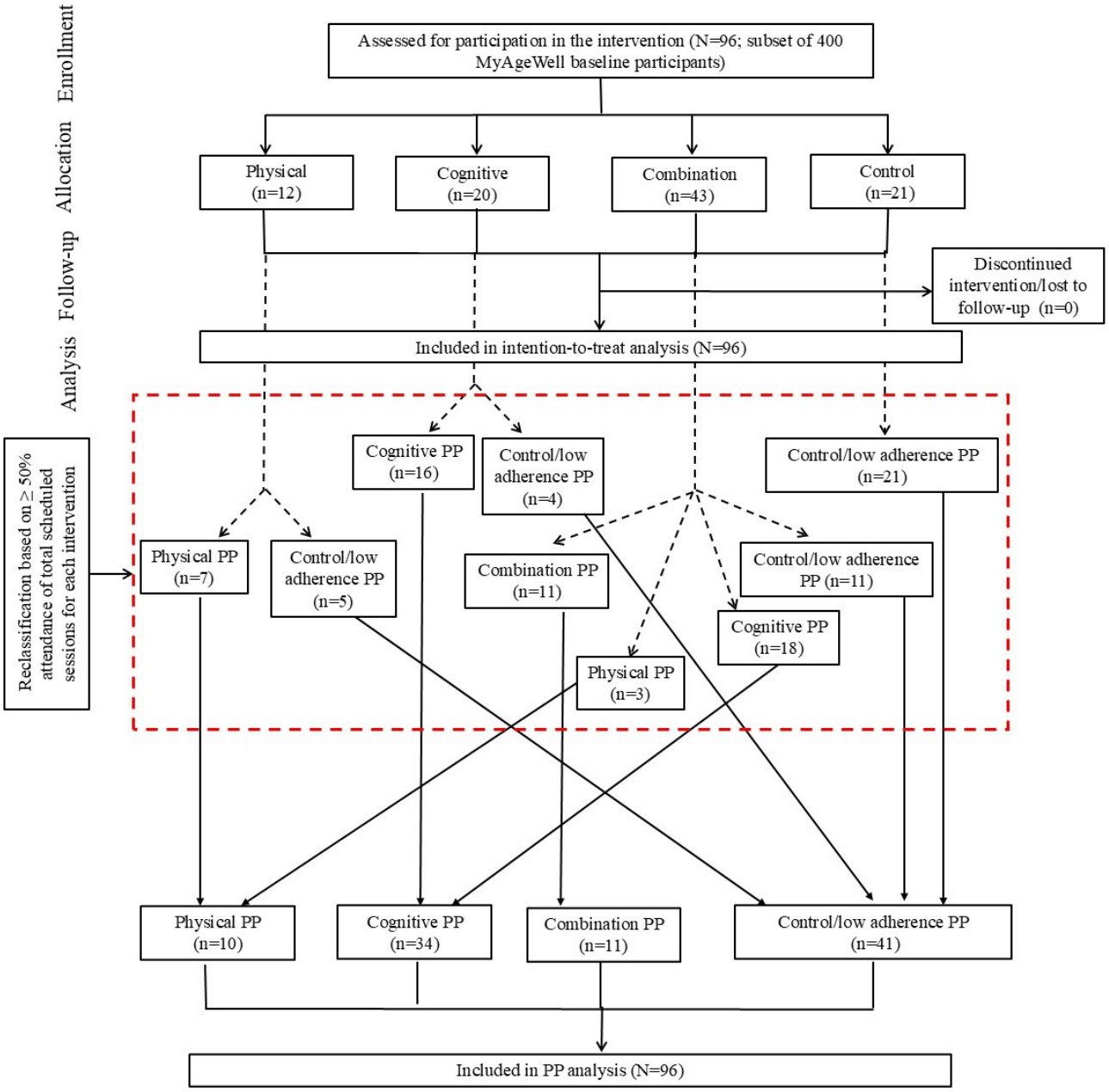
Movement of Participants From ITT Groups to PP Groups

In the ITT physical group (n=12), 7 participants attended at least 50% of the Zumba sessions and were therefore retained in the physical PP group, while the remaining 5 participants, who attended less than 50%, were reclassified into the control or low-adherence PP group. In the ITT cognitive group (n=20), 16 participants achieved at least 50% adherence to the video game sessions and remained in the cognitive PP group, whereas 4 participants with lower adherence were reassigned to the control or low-adherence PP group. In the ITT combination group (n=43), 11 participants met the adherence threshold for both Zumba and video game components and continued in the combination PP group. From the ITT combination group, 3 participants who adhered only to Zumba were reassigned to the physical PP group, 18 participants who adhered only to the video game component were moved to the cognitive PP group, and the remaining 11 participants, who did not achieve at least 50% adherence to either component, were reassigned to the control or low-adherence PP group. All 21 participants in the ITT control group remained in the control or low-adherence PP group. Following reclassification, the PP sample comprised physical (n=10), cognitive (n=34), combination (n=11), and control or low-adherence (n=41) groups. No participants were excluded from ITT analyses. This detailed reporting clarifies how participants transitioned from their original ITT allocation to the PP groups, ensuring transparency in how adherence

XSL•FO
RenderX

influenced the final analytic samples. Figure 1 illustrates the participant flow diagram (adapted CONSORT [Consolidated Standards of Reporting Trials] for quasi-randomized trial; Checklist 1).

Figure 1. Participant flow diagram (adapted CONSORT [Consolidated Standards of Reporting Trials] for quasi-randomized trial). Of 400 participants who completed baseline assessments in the MyAgeWell cohort, a subset of 96 were enrolled in the present 6-month intervention study and allocated to physical (n=12), cognitive (n=20), combination (n=43), and control (n=21). No participants dropped out, and no postintervention Montreal Cognitive Assessment data were missing. Per-protocol reclassification (dashed red box) was based on ≥50% attendance for each intervention component and produced PP groups of physical (n=10), cognitive (n=34), combination (n=11), and control or low adherence (n=41). The 96 participants were a subset of the larger MyAgeWell cohort (N=400). PP: per-protocol.



Baseline Characteristics

Baseline characteristics were assessed according to the original group allocation. Table 1 presents participants' demographic characteristics and baseline MoCA scores by intervention group. Participants' median age was 69 (IQR 65-72) years and did not differ significantly across groups (Kruskal-Wallis H=0.277, P=.96). MoCA scores at baseline were similar across groups (H=1.131, P=.77), with an overall median of 26 (IQR 25-28). Besides, although female participants comprised a larger proportion of the sample (63/96, 65.6%), no statistically significant differences in sex distribution were observed across intervention groups (χ^2_3=1.684, P=.64). Education level differed significantly across groups (χ^2_{18}=55.632, P<.001) and represents a potential confounding factor given its known association with MoCA performance. The cognitive group had the highest proportion of participants with a bachelor's degree (12/20, 60%), whereas the combination group included a higher proportion of participants with a master's degree (8/43, 18.6%).

XSL•FO
RenderX

Table . Demographic characteristics and baseline Montreal Cognitive Assessment scores of participants by intervention group.[a]

Variable	Physical (n=12)	Cognitive (n=20)	Combination (n=43)	Control (n=21)	Total (N=96)	*P* value
Age (y), median (IQR)	68 (66-71)	69 (66-72)	68 (65-71)	69 (65-73)	69 (65-72)	.96[b]
Sex, n (%)						.23[c]
Female	9 (75.0)	13 (65.0)	31 (72.1)	10 (47.6)	63 (65.6)	
Male	3 (25.0)	7 (35.0)	12 (27.9)	11 (52.4)	33 (34.4)	
Education level, n						<.001[d]
No education	3	0	1	0	4	
Primary	3	0	4	1	8	
Secondary	3	1	12	9	25	
Preuniversity/diploma	0	7	7	5	19	
Bachelor's	1	12	7	4	24	
Master's	0	0	8	2	10	
Unknown	2	0	4	0	6	
Hypertension, n (%)	3 (25.0)	5 (25.0)	6 (14.0)	5 (23.8)	19 (19.8)	.64[d]
MoCA[e] before, median (IQR)	26 (25-27)	26 (24-28)	26 (25-27)	26 (23-29)	26 (25-28)	.77[b]

[a]*P*<.05 was considered statistically significant.

[b]Kruskal-Wallis test for continuous variables (age and MoCA before).

[c]Chi-square test for sex distribution (male + female) across groups.

[d]Chi-square test for categorical variables (education level and hypertension) across groups.

[e]MoCA: Montreal Cognitive Assessment.

Descriptive Statistics and Normality

At baseline, participants demonstrated a median MoCA score of 26 (IQR 25-28), with slight negative skew (−0.877) and moderate kurtosis (1.079). The Shapiro-Wilk test indicated significant deviation from normality (*P*<.001), but the distribution was considered suitable for further analysis. Postintervention, the median MoCA score increased to 27 (IQR 26-29), showing moderate negative skew (−1.045) and kurtosis (1.210), with Shapiro-Wilk *P*<.001. The distribution of MoCA change scores (after – before) was approximately normal (mean 0.84, SD 2.30, skewness=0.125, kurtosis=0.304, Shapiro-Wilk *P*=.02), and Levene test confirmed homogeneity of variances across intervention groups (*P*=.15). These results justified the use of both parametric (mixed ANOVA) and nonparametric (Wilcoxon signed rank, Kruskal–Wallis, and Mann-Whitney *U*) analyses.

Mixed ANOVA (Time × Group Effects)

A 2×4 mixed ANOVA was conducted to examine the effects of time (pre vs post) and group (physical, cognitive, combination, and control) on overall MoCA scores (Table 2), with participants analyzed according to their original group allocation (ITT), irrespective of adherence level. The analysis revealed a significant main effect of time ($F_{1,92}$=7.47, *P*=.008, η^2=0.075), indicating that participants demonstrated a significant improvement in cognitive performance from preintervention to postintervention across all groups.

Table . Mixed ANOVA results for global Montreal Cognitive Assessment score.

Effect	Test statistic, *F* (*df*)	*P* value	Effect size (η^2)	Interpretation
Time	7.47 (1, 92)	.008[a]	0.075	Significant main effect
Group	1.28 (3, 92)	.29	0.040	Not significant
Time × group	0.57 (3, 92)	.64	0.018	Not significant

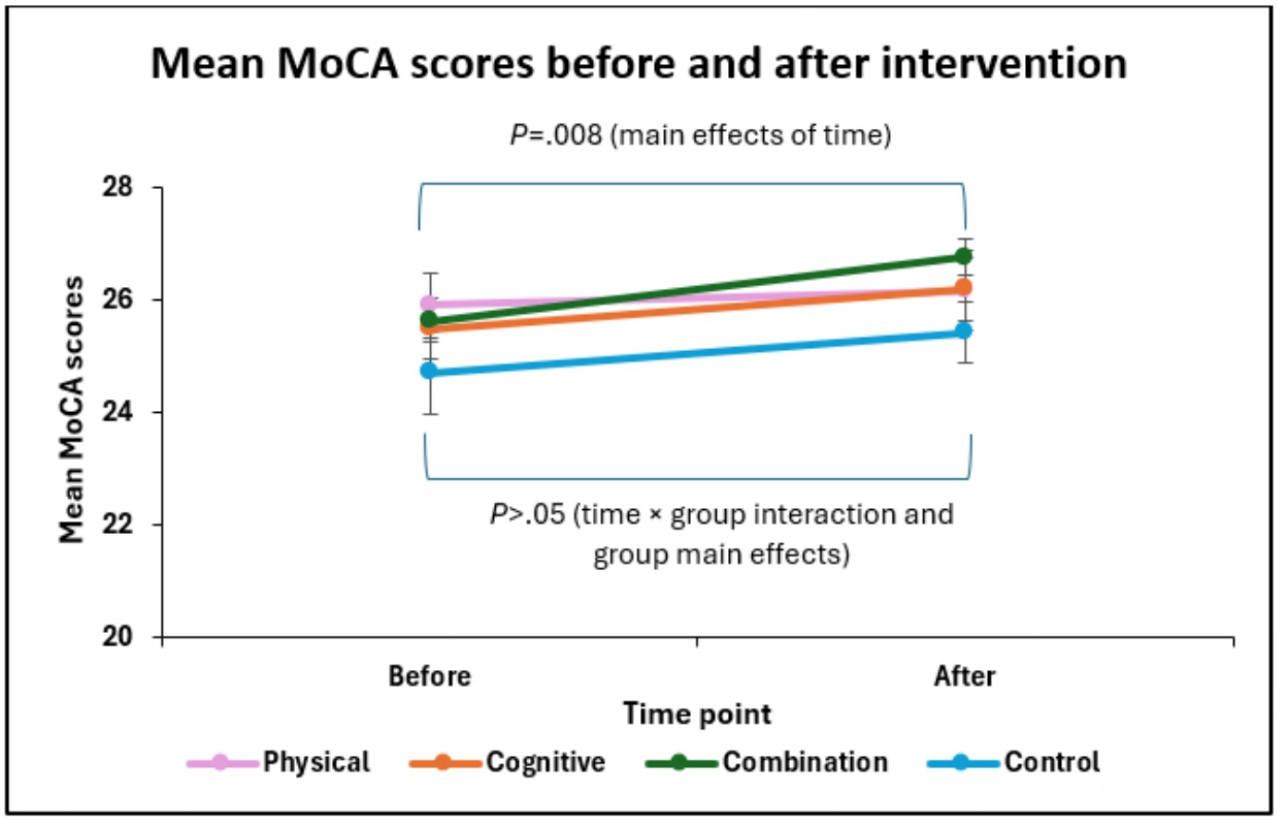
[a] *P*<.05 was considered statistically significant.

However, the time × group interaction was not significant ($F_{3,92}$=0.57, *P*=.64, η^2=0.018), suggesting that the magnitude of cognitive improvement did not differ significantly among the intervention types. Similarly, the between-group main effect

XSL•FO
RenderX

was nonsignificant ($F_{3,92}$=1.28, P=.29, η^2=0.040), indicating no baseline or overall mean differences between the 4 groups. Figure 2 shows the mean MoCA scores for each group before and after the intervention, illustrating a general upward trend across time regardless of intervention condition. Although the combination group showed a numerically larger increase from preassessment to postassessment, the time × group interaction was not significant, indicating comparable improvements across groups.

Figure 2. Mean MoCA scores before and after the intervention across groups. Data represent group mean (SEM). A significant main effect of time was observed ($F_{1,92}$=7.47, P=.008, η^2=0.075), indicating overall improvement in cognitive performance from preintervention to postintervention. The time × group interaction and group main effects were nonsignificant (P>.05), suggesting similar gains across all intervention groups. MoCA: Montreal Cognitive Assessment.



ITT Analysis: Cognitive Outcomes

Within-Group Changes (Wilcoxon Signed Rank Test)

Within-group comparisons of global cognitive performance were assessed using Wilcoxon signed rank tests for each intervention arm. Table 3 presents the preintervention and postintervention MoCA scores, corresponding change values, test statistics, and effect sizes.

Table . Within-group Montreal Cognitive Assessment score changes (intention-to-treat analysis, Wilcoxon signed rank test).

Group	MoCA[a] before, median (IQR)	MoCA after, median (IQR)	Median change	*Z*	*P* value[b]	Effect size (*r*)	Interpretation
Physical (n=12)	26 (25-27)	26 (25-28)	0.00	−0.59	.56	0.17	NS[c]
Cognitive (n=20)	26 (24-28)	26 (24-28)	0.00	−0.94	.35	0.21	NS
Combination (n=43)	26 (25-27)	27 (26-28)	+1.0	−3.55	<.001	0.54	Significant
Control (n=21)	26 (23-29)	26 (25-27)	0.00	−1.22	.22	0.27	NS
Total (N=96)	26 (25-28)	27 (26-29)	+1.0	3.38	<.001	0.34	Significant (overall)

[a]MoCA: Montreal Cognitive Assessment.

[b]P<.05 was considered statistically significant.

[c]NS: nonsignificant.

XSL•FO
RenderX

Across all participants, median MoCA scores increased from 26 (IQR 25-28) at baseline to 27 (IQR 26-29) postintervention, reflecting an overall improvement of +1 points (Z=3.38, P<.001, r=0.34), indicating a moderate effect. Among individual groups, the combination group showed a statistically significant increase in MoCA scores from 26 (IQR 25-27) to 27 (IQR 26-28) (Z=−3.55, P<.001, r=0.54), representing a large effect size. In contrast, the physical (median 26, IQR 25-27 to median 26, IQR 25-28), cognitive (median 26, IQR 24-28 to median 26, IQR 24-28), and control (median 26, IQR 23-29 to median 26, IQR 25-27) groups demonstrated nonsignificant improvements, with small effect sizes ranging from r=0.17 to 0.27. These findings suggest that while general cognitive function improved modestly across the entire sample, the combined physical and cognitive stimulation produced the most meaningful within-group enhancement.

Between-Group Comparisons (Kruskal-Wallis and Mann-Whitney U Tests)

Between-group analyses were conducted to determine whether changes in global cognition (MoCA change scores) differed significantly across intervention groups. The Kruskal-Wallis test showed no significant difference among the 4 groups (H=1.20, P=.75, η^2=0.013), indicating that overall cognitive improvement did not differ meaningfully between intervention types.

Subsequent pairwise Mann-Whitney U tests comparing each intervention group with the control group also yielded nonsignificant results. The physical vs control comparison (Z=−0.208, P=.84, r=0.036), cognitive vs control (Z=−0.026, P=.98, r=0.004), and combination vs control (Z=−0.848, P=.40, r=0.11) all demonstrated small, negligible effect sizes. Taken together, these findings suggest that while participants exhibited within-group improvements over time (Figure 2), the degree of improvement did not significantly differ across intervention conditions (Table 4). Given the small and unbalanced group sizes and the use of nonparametric tests, effect sizes (r and η^2) were reported to indicate the magnitude and direction of effects, formal 95% CIs for these nonparametric effect sizes were not calculated due to limited methodological support and reduced interpretability in small, skewed samples.

Table . Between-group comparisons of Montreal Cognitive Assessment change scores (intention-to-treat analysis).[a]

Comparison	Test statistic	P value	Effect size (r/η^2)	Interpretation
Kruskal-Wallis (4 groups)	H=1.20	.75	η^2=0.013	NS[b]
Physical vs control	Z=−0.208	.84	r=0.036	NS
Cognitive vs control	Z=−0.026	.98	r=0.004	NS
Combination vs control	Z=−0.848	.40	r=0.110	NS

[a]Effect sizes are reported as r or η^2. CIs were not calculated for nonparametric effect sizes due to small and uneven group sizes. P<.05 was considered statistically significant.

[b]NS: nonsignificant.

MoCA Subdomain Analysis (Wilcoxon Within Each Group)

To explore domain-specific cognitive effects, Wilcoxon signed rank tests were conducted for each of the 7 MoCA subdomains within the 4 intervention groups (Table 5). In the physical group, a significant improvement was observed in the language domain (P=.004), whereas other subdomains such as visuospatial, naming, attention, abstraction, memory, and orientation showed no significant changes.

XSL•FO
RenderX

Table . Wilcoxon signed rank test results for Montreal Cognitive Assessment subdomains by group.[a]

MoCA[b] subdomain	Physical (*P* value)	Cognitive (*P* value)	Combination (*P* value)	Control (*P* value)	Interpretation
Visuospatial	.77	.64	.03	.81	Improved only in combination group
Naming	.32	.56	.06	>.99	NS[c] across all groups
Attention	.56	.15	.56	.48	NS across all groups
Memory	.62	.21	.83	.55	NS across all groups
Language	.004	.05	.02	.79	Improved in physical and combination groups; trend in cognitive group
Abstraction	.32	.71	.13	.48	NS across all groups
Orientation	.38	.06	.11	.29	Trend in cognitive group only

[a]*P*<.05 was considered statistically significant.

[b]MoCA: Montreal Cognitive Assessment.

[c]NS: nonsignificant.

In the cognitive group, there were nonsignificant trends toward improvement in language (*P*=.05) and orientation (*P*=.06), suggesting possible subtle gains that did not reach statistical significance. Participants in the combination group demonstrated significant improvements in both visuospatial (*P*=.03) and language (*P*=.02) domains, indicating that combining physical and cognitive activities may yield broader cognitive benefits. In contrast, the control group exhibited no significant changes in any MoCA subdomain across the intervention period.

PP Analysis: Cognitive Outcomes

Overall Group Comparisons

A Kruskal-Wallis test indicated no significant difference in MoCA change scores among the 4 PP groups (H=3.275, df=3, *P*=.35), with a small effect size (η^2=0.034). Although not statistically significant, mean rank trends showed that participants in the cognitive and combination groups tended to exhibit slightly greater improvements compared with the physical and control or low adherence groups.

Within-Group Changes (Wilcoxon Signed Rank Tests)

Within-group comparisons of preintervention and postintervention MoCA scores are presented in Table 6. Significant improvements were observed in the cognitive (Z=−2.693, *P*=.007, r=0.46) and combination (Z=−2.200, *P*=.03, r=0.66) groups, reflecting moderate-to-large effect sizes. An increase of ≥2 points on the MoCA was considered a meaningful improvement. Using this criterion, participants in the cognitive and combination groups demonstrated median gains of +2 points, consistent with the definition of a cognitive responder. The physical (Z=−0.144, *P*=.89, r=0.05) and control or low adherence (Z=−1.468, *P*=.14, r=0.23) groups did not show significant changes. These findings suggest that interventions involving cognitive engagement, either independently or in combination with physical activity, led to more pronounced cognitive benefits.

Table . Within-group changes in Montreal Cognitive Assessment scores by intervention group (per-protocol analysis).

Group	MoCA[a] before, median (IQR)	MoCA after, median (IQR)	Median change	*Z*	*P* value[b]	Effect size (*r*)	Interpretation
Physical (n=10)	26 (25 - 27)	26 (25 - 28)	0.0	−0.144	.89	0.05	NS[c]
Cognitive (n=34)	25 (24 - 26)	27 (26 - 28)	+2.0	−2.693	.007	0.46	Significant
Combination (n=11)	27 (26 - 28)	29 (28 - 30)	+2.0	−2.200	.03	0.66	Significant
Control or low adherence (n=41)	26 (25 - 27)	26 (25 - 27)	+0.0	−1.468	.14	0.23	NS

[a]MoCA: Montreal Cognitive Assessment.

[b]*P*<.05 was considered statistically significant.

[c]NS: nonsignificant.

XSL•FO
RenderX

Between-Group Pairwise Comparisons

Pairwise Mann-Whitney *U* tests comparing each intervention group with the control or low adherence group were nonsignificant (Table 7). Effect sizes were small across all comparisons ($r<0.15$), indicating minimal between-group differences in the magnitude of MoCA change.

Table . Between-group Mann-Whitney *U* comparisons of Montreal Cognitive Assessment change scores (per-protocol analysis).

Comparison	Test statistic (*Z*)	*P* value[a]	Effect size (*r*)	Interpretation
Physical vs control or low adherence	−0.527	.60	0.07	NS[b]
Cognitive vs control or low adherence	−1.291	.20	0.15	NS
Combination vs control or low adherence	−0.894	.37	0.12	NS

[a]*P*<.05 was considered statistically significant.

[b]NS: nonsignificant.

MoCA Subdomain Analysis (Wilcoxon Within Each Group)

Domain-specific analysis using Wilcoxon signed rank tests identified selective improvements across PP intervention groups (Table 8). The physical group demonstrated a significant improvement in the language domain (*P*=.01). The cognitive group showed significant gains in both language (*P*=.03) and orientation (*P*=.03), as well as a near-significant trend in visuospatial (*P*=.07). The combination group also showed a significant improvement in language (*P*=.02). No significant changes were observed in the control or low adherence group across all subdomains. These findings indicate that interventions incorporating cognitive elements, either alone or in combination with physical activity, particularly enhanced language and orientation abilities, which are important for executive and communicative functions among older adults.

Table . Wilcoxon signed rank results for Montreal Cognitive Assessment subdomains (per-protocol analysis).[a]

MoCA[b] subdomain	Physical (*P* value)	Cognitive (*P* value)	Combination (*P* value)	Control/low adherence (*P* value)	Interpretation summary
Visuospatial	.59	.07	.66	.22	Trend in cognitive group
Naming	.32	.32	.32	>.99	NS[c] across all groups
Attention	.78	.13	.66	.51	NS across all groups
Memory	.76	.81	>.99	.89	NS across all groups
Language	.01	.03	.02	.62	Improved in physical, cognitive, and combination groups
Abstraction	.66	.56	.32	.32	NS across all groups
Orientation	.41	.03	.26	.44	Improved in cognitive group

[a]*P*<.05 was considered statistically significant.

[b]MoCA: Montreal Cognitive Assessment.

[c]NS: nonsignificant.

MoCA Responder Analysis

When using a ≥2-point increase as the criterion for meaningful improvement, 37.5 (36/96) participants were classified as responders, while 62.5% (n=60) participants were nonresponders. Responders were most common in the cognitive and combination groups, supporting their greater within-group gains.

Attendance and Cognitive Change

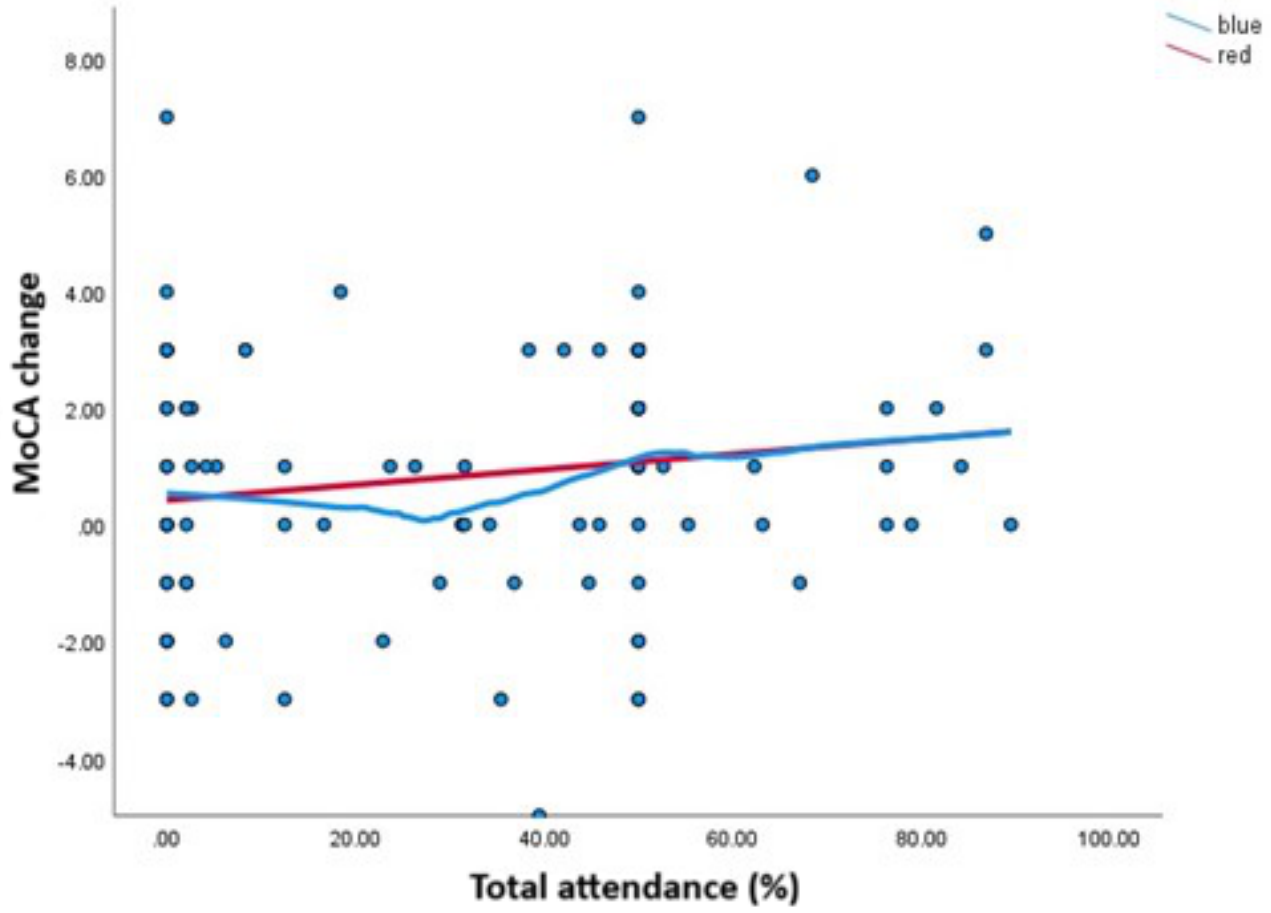
Spearman rank correlation analyses revealed weak and nonsignificant relationships between attendance and MoCA improvement. Video game attendance showed a small positive correlation with MoCA change (*r*=0.192, *P*=.06), while total attendance was similarly weak (*r*=0.159, *P*=.12). Zumba attendance was not significantly associated with cognitive improvement.

Regression analysis further confirmed that attendance variables were not significant predictors of MoCA change (*P*=.18). ROC analysis indicated low discriminative accuracy of attendance measures in identifying cognitive responders, with area under the curve (AUC) values ranging from 0.45 to 0.59, suggesting poor predictive utility. Figure 3 illustrates the relationship between total attendance (%) and MoCA change. Linear (red) and locally weighted scatterplot smoothing (blue) trend lines

XSL•FO
RenderX

show a weak upward trajectory, consistent with the nonsignificant statistical findings.

Figure 3. Relationship between total attendance (%) and MoCA change. Scatter plot illustrating the association between total attendance (%) and change in MoCA scores among participants. Linear (red) and locally weighted scatterplot smoothing (blue) lines are shown. Both lines indicate a weak positive trend, consistent with the nonsignificant correlation and regression analyses. MoCA: Montreal Cognitive Assessment.



Discussion

This semirandomized trial found that physical (Zumba), cognitive (video game), and combined interventions were associated with within-group improvements in global cognitive function among older adults in Klang Valley, Malaysia. The combined intervention demonstrated the largest magnitude of improvement, although between-group differences were not statistically significant. Patterns of improvement suggested that adherence may play an important role in achieving meaningful cognitive benefits.

Cognitive impairment and age-related decline substantially affect daily functioning and independence in older adults. With rapid global aging and increasing dementia prevalence, effective preventive strategies are urgently needed [15]. Pharmacological approaches show limited success and are often accompanied by adverse effects [24-26], prompting growing interest in nonpharmacological interventions targeting cognitive frailty [27-31]. Despite this progress, the comparative effectiveness of different intervention modalities and their combinations remains uncertain [9]. Our study addresses this gap by examining physical, cognitive, and combined interventions on global and domain-specific MoCA performance, providing insights into strategies to preserve cognitive health in later life.

At baseline, there were no significant differences across groups in age (P=.96), sex (P=.23), MoCA scores before intervention (P=.77), or hypertension prevalence (P=.64), indicating a well-balanced distribution among participants. Recently, our research team examined the association between C-reactive protein (CRP) levels and various demographic factors. The analysis demonstrated that CRP concentrations were significantly correlated with gender, ethnicity, and BMI (P<.05), but not with age or economic status [18]. CRP is an inflammatory biomarker that plays a pivotal role in the pathophysiology of age-related conditions, including cognitive decline.

However, a significant between-group difference was observed in educational attainment (P<.001), which is a well-established determinant of cognitive reserve and MoCA performance. The cognitive group had the highest proportion of participants with a bachelor's degree, while the combination group included a relatively higher proportion with postgraduate education. This imbalance may have influenced both baseline cognitive performance and responsiveness to intervention, thereby complicating causal interpretation of between-group comparisons.

Although educational attainment differed across groups at baseline, the modest sample size and small subgroup numbers

XSL•FO
RenderX

limited the feasibility of conducting covariate-adjusted or stratified analyses to account for potential education effects. Consequently, the observed intervention effects may have been influenced by baseline differences in cognitive reserve associated with education. Nevertheless, prior research suggests that cognitively engaging and socially meaningful interventions can benefit older adults across a broad range of educational backgrounds [32]. These considerations highlight the need to interpret our findings cautiously and the importance of larger, well-powered trials to separate intervention effects from educational influences.

Overall, the mixed ANOVA results showed a significant improvement in cognitive performance, as demonstrated in MoCA score from preintervention to postintervention (P=.008), indicating that engagement in structured group activities can enhance cognitive outcomes in older adults, regardless of intervention type [33]. The time × group interaction, based on the mixed ANOVA, was not significant (P=.64), suggesting that the magnitude of cognitive improvement was similar across intervention types. These findings were further supported by Kruskal-Wallis and Mann-Whitney analyses in both ITT and PP approaches, which showed no significant differences in MoCA change scores in between-group comparisons or in pairwise comparisons of each intervention with the control group. Consistent with Barnes et al [34], who reported that 12 weeks of physical, cognitive, or combined activities in inactive older adults did not differ from the control groups, these findings suggest that overall engagement or the total amount of activity may be more influential than the specific type of intervention. One possible explanation is that nonspecific factors, such as social interaction, may contribute to the cognitive benefits observed [33].

The absence of significant between-group differences warrants further consideration. Several methodological and contextual factors may account for this pattern. First, the relatively small and unequal group sizes in the PP analysis, particularly the physical group (n=10), reduced statistical power to detect modest effects. Second, the improvements observed in the control group might be due to practice effects from taking the MoCA test multiple times or from nonspecific benefits such as social interaction, as previously noted by Lam et al [33], as well as incidental contact with the research team during assessments. Third, the sample's relatively high baseline MoCA scores (mean≈25) suggest a potential ceiling effect, limiting the magnitude of observable gains among cognitively healthy participants. Finally, interindividual variability in engagement and adherence may have contributed to heterogeneous treatment responses, attenuating between-group contrasts. These explanations are consistent with prior intervention studies showing that cognitive enrichment benefits can occur across multiple activity types, but group-level differences often emerge only with larger samples or higher training intensity [35-37].

Notably, the combination group showed numerically larger within-group improvements and more consistent effect sizes across analytic approaches, while the cognitive group also demonstrated significant gains in the PP analysis. Importantly, these significant improvements were observed only among participants who achieved at least moderate adherence (≥50% attendance), indicating that sufficient intervention exposure was necessary to elicit measurable gains in global cognition. However, between-group comparisons did not reveal statistically significant differences, indicating that cognitive improvements over time were broadly comparable across intervention conditions. Within this context, the relatively larger effect sizes observed in the combined intervention, reflected by large within-group effects in both the ITT (r=0.54) and PP (r=0.66) analyses, suggest a potential advantage of multimodal engagement when adherence is maintained. These patterns are consistent with previous reports indicating that combined physical and cognitive interventions may be associated with greater improvements in global cognition compared with single-modality approaches, as observed by Yu et al [35] and Han et al [36]. Mechanistically, dual-task or multitask training that integrates physical and cognitive elements has been proposed to support cognitive function through neurophysiological pathways, including reduced bilateral prefrontal cortical oxygenation [36], increased hippocampal volume [38], and improved white matter integrity [39].

The video game-based cognitive intervention demonstrated meaningful benefits in this study, particularly among participants who maintained at least moderate adherence in the PP analysis. This aligns with the meta-analysis by Lampit et al [14], which reported that computerized cognitive training produces small but significant improvements in cognitive performance among healthy older adults. Participants engaging more consistently with the cognitive intervention showed the largest gains, highlighting the importance of adherence in achieving measurable benefits. In contrast, the physical activity group showed minimal cognitive change across both analytic approaches, suggesting that a once-weekly Zumba session may not have provided a sufficient cumulative training dose to elicit substantial cognitive improvements.

Besides, Zhang et al [40] examined the effects of all FITT-VP (frequency, intensity, time [duration], type, volume, and progression) exercise variables on cognition across healthy populations, including older adults, and found that variations in FITT-VP influence cognitive benefits. Older adults showed the greatest improvements in global cognition, executive function, and memory. Yuan et al [31] similarly reported that exercise improved cognition in older adults with cognitive frailty, particularly when conducted 3 or more times per week, with each session lasting at least 45 minutes over 12 weeks or longer. Chia et al [17] also observed that participants with higher physical activity levels exhibited significantly greater median MoCA scores compared with less active participants. Collectively, these findings emphasize that exercise characteristics, including frequency, intensity, and session duration (time), are crucial determinants of cognitive outcomes.

The present findings are consistent with recommendations from the American College of Sports Medicine and prior meta-analyses. Northey et al [41] reported that moderate-intensity exercise lasting at least 45 minutes per session on most days of the week enhances global cognition in adults aged 50 years and older. Xue et al [42] similarly found that moderate-duration, moderate-intensity exercise is associated with significant cognitive improvements among older adults.

XSL•FO
RenderX

Although these studies do not explicitly recommend the FITT-VP framework, their results indirectly support its core principles, particularly frequency, intensity, and time, as key determinants of cognitive benefit.

In this study, Zumba sessions were delivered once weekly for 60 minutes over approximately 19 weeks. While session intensity was appropriate, the low frequency likely limited cumulative training dose, providing a plausible explanation for the minimal cognitive gains observed. Applying FITT-VP principles as an interpretive framework highlights the importance of optimizing exercise frequency, duration, and intensity in future interventions to ensure sufficient training stimulus for measurable cognitive improvements among older adults. While each intervention component was matched in session duration, the combined arm received a higher total weekly exposure, and this difference in intervention "dose" was considered when interpreting between-group comparisons.

The study was further extended with subdomain analyses under both analytic approaches (ITT and PP) to explore domain-specific responsiveness to the different interventions. The results showed that the language domain, which included assessments of verbal, categorical, and phonological fluency [43], demonstrated the most consistent improvement across all intervention groups. This suggests that verbal fluency and communication-related cognitive processes are particularly sensitive to structured and socially interactive activities. These findings are also supported by previous studies, although the results varied depending on the intervention type. Nocera et al [44] reported that aerobic exercise improved verbal fluency, whereas cognitive training was associated with decreased or nonsignificant improvement in this domain. In contrast, Calatayud et al [45] demonstrated that a comprehensive cognitive stimulation program significantly enhanced language proficiency, verbal fluency, and overall cognitive function. These contrasting findings indicate that the type and structure of the intervention may differentially influence specific cognitive domains.

Significant gains in the visuospatial domain were observed in the combination group in the ITT analysis, while improvements in orientation emerged in the cognitive stimulation group and reached significance in the PP analysis. These patterns suggest that distinct cognitive mechanisms may underlie the effects of different stimulation modalities. Interventions that combine physical and cognitive elements may enhance visuospatial and integrative processing, whereas purely cognitive activities may strengthen orientation. Collectively, these subdomain-level findings highlight the importance of examining domain-specific outcomes beyond total MoCA scores, as they may reveal early indicators of intervention efficacy [15].

Exploratory analyses examining adherence showed that attendance rates were not significantly correlated with MoCA improvement. While weak positive associations were observed for video game and total attendance, regression and ROC analyses showed limited predictive value for cognitive improvement. This suggests that mere attendance frequency may not fully capture the quality or engagement level of participation, underscoring the complexity of dose-response relationships in behavioral interventions. Nevertheless, adherence remains clinically relevant, as participants with consistent engagement in cognitive or combined activities demonstrated the most robust improvements in the PP analyses.

From a clinical perspective, these findings highlight the feasibility of combined physical and cognitive interventions as a practical approach for promoting cognitive engagement among older adults. While between-group differences were not statistically significant, the consistent within-group improvements and favorable subdomain patterns observed in participants with higher adherence suggest that sustained participation in structured multimodal activities may be beneficial in real-world community settings.

In relation to previous literature, the present findings are consistent with meta-analyses reporting small-to-moderate improvements in global cognition following cognitive and combined interventions [15,46]. This study adds novel insights by directly comparing physical activity, cognitive stimulation, and combined training within a single cohort, employing both ITT and PP analytic frameworks and mapping MoCA subdomains to identify domain-specific responsiveness. Such findings contribute to growing evidence that multimodal engagement may support cognitive health in aging populations, although definitive conclusions regarding superiority over single-domain training require larger, adequately powered trials.

Several limitations should be acknowledged. The semirandomized allocation and lack of blinding may have introduced expectancy effects, although baseline characteristics were comparable across groups and standardized outcome assessments were used to mitigate potential bias. Additionally, the sample size was modest and uneven across groups, limiting statistical power and precision. The intervention frequency (once weekly) and duration may have been insufficient to elicit large cognitive changes. Cognitive outcomes were based solely on the MoCA, without complementary neuropsychological or neuroimaging measures. Moreover, the study lacked long-term follow-up, preventing evaluation of the persistence of cognitive gains. Furthermore, as participants were community-dwelling older adults from a single urban region, generalizability to rural or clinical populations is limited.

Besides, adherence thresholds for the PP analysis were pragmatically defined at ≥50% attendance. Sensitivity analyses using higher cutoffs (70% - 80%) yielded directionally similar results, but small subgroup sizes reduced statistical stability, highlighting the trade-off between stricter adherence definitions and analytic feasibility in nonpharmacological interventions. Pooling participants with no intervention exposure and those with partial exposure below the ≥50% adherence threshold increased heterogeneity and may have attenuated detectable between-group differences.

Finally, formal 95% CIs for nonparametric between-group effect sizes were not calculated due to methodological constraints with small, skewed samples. Consequently, effect sizes (r and η^2) were used to indicate magnitude and direction rather than statistical precision, and between-group findings should be interpreted cautiously.

XSL•FO
RenderX

In summary, both cognitive and combined interventions produced meaningful improvements in cognitive performance among older adults, with the strongest effects observed when adherence was maintained. Although between-group differences were not statistically significant, the consistent within-group trends and domain-specific patterns suggest practical cognitive benefits of engaging in regular, structured multimodal activities. Future research with larger samples, longer intervention durations, and multidomain cognitive assessments is warranted to confirm these findings and further clarify the mechanisms underlying differential cognitive responsiveness.

This study demonstrates that participation in video game–based cognitive stimulation, physical activity, and combined interventions was associated with within-group improvements in global cognitive performance among community-dwelling older adults. PP analyses indicated that participants who attended at least 50% of the cognitive stimulation sessions showed moderate improvements, while those who adhered to at least 50% of both the physical and cognitive components in the combined intervention demonstrated numerically larger and more consistent gains. Domain-specific analyses suggested that language, visuospatial, and orientation abilities may respond differently to intervention type. Although between-group differences were not statistically significant, consistent within-group trends support the practical value of structured engagement in cognitive and combined activities. Future studies with longer durations, higher frequency or intensity of physical activity, larger sample sizes, and multidomain cognitive assessments are needed to clarify the mechanisms underlying intervention-specific cognitive responsiveness.

Acknowledgments

The authors acknowledge the Ministry of Higher Education Malaysia for funding this work. They are grateful to all postdoctoral research fellows, graduate research assistants, research assistants, and interns for their invaluable contributions to the project. They also extend their sincere thanks to all participants for their time, commitment, and active involvement in the intervention sessions. ChatGPT (OpenAI) was used to support language editing during the preparation of this manuscript. The authors reviewed and edited all content and take full responsibility for the accuracy, originality, and integrity of the manuscript.

Funding

This study was supported by a research grant from the Ministry of Higher Education Malaysia (LRGS/1/2019/SYUC/02/1/1) and Sunway University.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: YCC
Figure preparation: SNSA
Methodology: YCC, SNSA, JC, YL, MJ
Supervision: YCC
Writing – original draft: SNSA
Writing – review & editing: YCC, JC, YL, WGL, WZY, KALZM, MJ
All authors provided critical feedback and helped shape the context of the paper.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 1161 KB - aging_v9i1e88479_app1.pdf]

References

1. Population projections, Malaysia, 2020–2060. : Department of Statistics Malaysia (DOSM); 2024 URL: https://www.dosm.gov.my/uploads/release-content/file_20250710220243.pdf [accessed 2026-05-29]
2. World population ageing 2020 highlights. : United Nations, Department of Economic and Social Affairs; 2020 URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Sep/un_pop_2020_pf_ageing_10_key_messages.pdf [accessed 2026-05-29]
3. Song WX, Wu WW, Zhao YY, et al. Evidence from a meta-analysis and systematic review reveals the global prevalence of mild cognitive impairment. Front Aging Neurosci 2023;15:1227112. [doi: 10.3389/fnagi.2023.1227112]

XSL•FO
RenderX

4. Lee LK, Shahar S, Chin AV, Mohd Yusoff NA, Rajab N, Aziz SA. Prevalence of gender disparities and predictors affecting the occurrence of mild cognitive impairment (MCI). Arch Gerontol Geriatr 2012;54(1):185-191. [doi: 10.1016/j.archger.2011.03.015] [Medline: 21546098]
5. Lai WH, Hon YK, Pang GMH, et al. Dementia of the ageing population in Malaysia: a scoping review of published research. Aging Health Res 2022 Jun;2(2):100077. [doi: 10.1016/j.ahr.2022.100077]
6. Atri Roop Y, Shalihin MSE, Mohammad Aidid E. Mild cognitive impairment and its associated factors amongst the older people attending government health clinics in Kuantan. Med J Malaysia 2025 May;80(3):321-327. [Medline: 40437721]
7. Gaster B. Bringing cognitive assessment to the front lines of primary care. MedPage Today. 2025 Nov 19. URL: https://www.medpagetoday.com/opinion/second-opinions/118602 [accessed 2025-11-20]
8. World Alzheimer report 2015: the global impact of dementia - an analysis of prevalence, incidence, cost and trends. : Alzheimer's Disease International; 2015 URL: https://www.alzint.org/u/WorldAlzheimerReport2015.pdf [accessed 2026-05-29]
9. Mendes AJ, Ribaldi F, Sayin O, et al. Single-domain and multidomain lifestyle interventions for the prevention of cognitive decline in older adults who are cognitively unimpaired: a systematic review and network meta-analysis. Lancet Healthy Longev 2025 Sep;6(9):100762. [doi: 10.1016/j.lanhl.2025.100762] [Medline: 41016407]
10. Zhang D, Lu Y, Zhao X, Zhang Q, Li L. Aerobic exercise attenuates neurodegeneration and promotes functional recovery - why it matters for neurorehabilitation & neural repair. Neurochem Int 2020 Dec;141:104862. [doi: 10.1016/j.neuint.2020.104862] [Medline: 33031857]
11. Parial LL, Kor PPK, Sumile EF, Leung AYM. Dual-task Zumba gold for improving the cognition of people with mild cognitive impairment: a pilot randomized controlled trial. Gerontologist 2023 Aug 24;63(7):1248-1261. [doi: 10.1093/geront/gnac081] [Medline: 35679826]
12. Pawar M, Shinde N. Impact of Zumba on health and its effects on quality of life (QOL): a survey based study. Int J Yoga Physiother Phys Educ 2023;8(3):4-8 [FREE Full text]
13. Stonnington CM, Krell-Roesch J, Locke DEC, et al. Impact of Zumba on cognition and quality of life is independent of APOE4 carrier status in cognitively unimpaired older women: a 6-month randomized controlled pilot study. Am J Alzheimers Dis Other Demen 2020;35:1533317519868370. [doi: 10.1177/1533317519868370] [Medline: 31401841]
14. Lampit A, Hallock H, Valenzuela M. Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. PLoS Med 2014 Nov;11(11):e1001756. [doi: 10.1371/journal.pmed.1001756] [Medline: 25405755]
15. Gheysen F, Poppe L, DeSmet A, et al. Physical activity to improve cognition in older adults: can physical activity programs enriched with cognitive challenges enhance the effects? A systematic review and meta-analysis. Int J Behav Nutr Phys Act 2018 Jul 4;15(1):63. [doi: 10.1186/s12966-018-0697-x] [Medline: 29973193]
16. Lindvall E, Abzhandadze T, Quinn TJ, Sunnerhagen KS, Lundström E. Is the difference real, is the difference relevant: the minimal detectable and clinically important changes in the Montreal Cognitive Assessment. Cereb Circ Cogn Behav 2024;6:100222. [doi: 10.1016/j.cccb.2024.100222] [Medline: 38745691]
17. Chia YC, Low END, Teh JKL, et al. Association between physical activity and cognitive function in a multi-ethnic Asian older adult population. Sci Rep 2025 Feb 28;15(1):7249. [doi: 10.1038/s41598-025-92074-w] [Medline: 40021918]
18. Lim WG, Lim WL, Lau TP, Chia YC, Chew J. Association between salivary C-reactive protein levels and covariates in an older adult population in Malaysia: a cross-sectional study of the MyAgeWell cohort. Prog Microbes Mol Biol 2025;8(1):a0000461. [doi: 10.36877/pmmb.a0000461]
19. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
20. Julayanont P, Nasreddine ZS. Montreal Cognitive Assessment (MoCA): concept and clinical review. In: Larner AJ, editor. Cognitive Screening Instruments: Springer International Publishing; 2017:139-195. [doi: 10.1007/978-3-319-44775-9_7]
21. Baumgartner PC, Haynes RB, Hersberger KE, Arnet I. A systematic review of medication adherence thresholds dependent of clinical outcomes. Front Pharmacol 2018;9:1290. [doi: 10.3389/fphar.2018.01290] [Medline: 30524276]
22. Giovanazzi A, Jones K, Carr RM, Fairhurst CM, Backhouse MR, Adamson JA. Current practice in the measurement and interpretation of intervention adherence in randomised controlled trials: a systematic review. Contemp Clin Trials 2022 Jul;118:106788. [doi: 10.1016/j.cct.2022.106788] [Medline: 35562000]
23. Soldevila-Domenech N, Ayala-Garcia A, Barbera M, et al. Adherence and intensity in multimodal lifestyle-based interventions for cognitive decline prevention: state-of-the-art and future directions. Alzheimers Res Ther 2025 Mar 17;17(1):61. [doi: 10.1186/s13195-025-01691-0] [Medline: 40098201]
24. Lee YY, Wu CY, Teng CH, Hsu WC, Chang KC, Chen P. Evolving methods to combine cognitive and physical training for individuals with mild cognitive impairment: study protocol for a randomized controlled study. Trials 2016 Oct 28;17(1):526. [doi: 10.1186/s13063-016-1650-4] [Medline: 27793183]
25. Petersen RC, Lopez O, Armstrong MJ, et al. Practice guideline update summary: mild cognitive impairment [RETIRED]: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of

XSL•FO
RenderX

Neurology. Neurology (ECronicon) 2018 Jan 16;90(3):126-135. [doi: 10.1212/WNL.0000000000004826] [Medline: 29282327]
26. Campbell EB, Delgadillo M, Lazzeroni LC, et al. Cognitive improvement following physical exercise and cognitive training intervention for older adults with MCI. J Gerontol A Biol Sci Med Sci 2023 Mar 1;78(3):554-560. [doi: 10.1093/gerona/glac189] [Medline: 36099058]
27. Cano A. Cognitive frailty, a new target for healthy ageing. Maturitas 2015 Oct;82(2):139-140. [doi: 10.1016/j.maturitas.2015.07.026] [Medline: 26277253]
28. Ruan Q, Yu Z, Chen M, Bao Z, Li J, He W. Cognitive frailty, a novel target for the prevention of elderly dependency. Ageing Res Rev 2015 Mar;20:1-10. [doi: 10.1016/j.arr.2014.12.004] [Medline: 25555677]
29. Sugimoto T, Arai H, Sakurai T. An update on cognitive frailty: its definition, impact, associated factors and underlying mechanisms, and interventions. Geriatr Gerontol Int 2022 Feb;22(2):99-109. [doi: 10.1111/ggi.14322] [Medline: 34882939]
30. Peng J, Chang R, Wei X, Yin Z, Liu Q. Effect of non-pharmacological interventions in people with cognitive frailty: a systematic review and network meta-analysis. BMC Public Health 2024 Oct 1;24(1):2684. [doi: 10.1186/s12889-024-20079-7] [Medline: 39354435]
31. Yuan Y, Wang S, Zhou C, Zhang A, Zhang S, Wang Y. Effects of exercise interventions on cognition, physical function and quality of life among older adults with cognitive frailty: a systematic review and meta-analysis. Geriatr Nurs 2025;62(Pt A):96-107. [doi: 10.1016/j.gerinurse.2025.01.006] [Medline: 39889512]
32. Vafa S, Thanaraju A, Chan JK, et al. The effectiveness of real-life cognitive and physical interventions on cognitive functioning in healthy older adults: a systematic review of the effects of education and training duration. Appl Psychol Health Well Being 2025 Feb;17(1):e12651. [doi: 10.1111/aphw.12651] [Medline: 39822168]
33. Lam LCW, Chan WC, Leung T, Fung AWT, Leung EMF. Would older adults with mild cognitive impairment adhere to and benefit from a structured lifestyle activity intervention to enhance cognition?: a cluster randomized controlled trial. PLoS ONE 2015;10(3):e0118173. [doi: 10.1371/journal.pone.0118173] [Medline: 25826620]
34. Barnes DE, Santos-Modesitt W, Poelke G, et al. The Mental Activity and eXercise (MAX) trial: a randomized controlled trial to enhance cognitive function in older adults. JAMA Intern Med 2013 May 13;173(9):797-804. [doi: 10.1001/jamainternmed.2013.189] [Medline: 23545598]
35. Yu R, Leung G, Woo J. Randomized controlled trial on the effects of a combined intervention of computerized cognitive training preceded by physical exercise for improving frailty status and cognitive function in older adults. Int J Environ Res Public Health 2021 Feb 3;18(4):1396. [doi: 10.3390/ijerph18041396] [Medline: 33546255]
36. Han K, Tang Z, Bai Z, Su W, Zhang H. Effects of combined cognitive and physical intervention on enhancing cognition in older adults with and without mild cognitive impairment: a systematic review and meta-analysis. Front Aging Neurosci 2022;14:878025. [doi: 10.3389/fnagi.2022.878025] [Medline: 35928994]
37. Chan RCF, Zhou JHS, Cao Y, et al. Nonpharmacological multimodal interventions for cognitive functions in older adults with mild cognitive impairment: scoping review. JMIR Aging 2025 May 12;8:e70291. [doi: 10.2196/70291] [Medline: 40354647]
38. Erickson KI, Voss MW, Prakash RS, et al. Exercise training increases size of hippocampus and improves memory. Proc Natl Acad Sci U S A 2011 Feb 15;108(7):3017-3022. [doi: 10.1073/pnas.1015950108] [Medline: 21282661]
39. Tait JL, Duckham RL, Milte CM, Main LC, Daly RM. Influence of sequential vs. simultaneous dual-task exercise training on cognitive function in older adults. Front Aging Neurosci 2017;9:368. [doi: 10.3389/fnagi.2017.00368] [Medline: 29163146]
40. Zhang M, Jia J, Yang Y, Zhang L, Wang X. Effects of exercise interventions on cognitive functions in healthy populations: a systematic review and meta-analysis. Ageing Res Rev 2023 Dec;92:102116. [doi: 10.1016/j.arr.2023.102116] [Medline: 37924980]
41. Northey JM, Cherbuin N, Pumpa KL, Smee DJ, Rattray B. Exercise interventions for cognitive function in adults older than 50: a systematic review with meta-analysis. Br J Sports Med 2018 Feb;52(3):154-160. [doi: 10.1136/bjsports-2016-096587] [Medline: 28438770]
42. Xue Y, Yang Y, Huang T. Effects of chronic exercise interventions on executive function among children and adolescents: a systematic review with meta-analysis. Br J Sports Med 2019 Nov;53(22):1397-1404. [doi: 10.1136/bjsports-2018-099825] [Medline: 30737201]
43. Rieker JA, Reales JM, Muiños M, Ballesteros S. The effects of combined cognitive-physical interventions on cognitive functioning in healthy older adults: a systematic review and multilevel meta-analysis. Front Hum Neurosci 2022;16:838968. [doi: 10.3389/fnhum.2022.838968] [Medline: 35399365]
44. Nocera JR, Mammino K, Kommula Y, Wharton W, Crosson B, McGregor KM. Effects of combined aerobic exercise and cognitive training on verbal fluency in older adults. Gerontol Geriatr Med 2020;6:2333721419896884. [doi: 10.1177/2333721419896884] [Medline: 31950073]
45. Calatayud E, Subiron-Valera AB, Marcén-Román Y, et al. Effects on language and verbal fluency of a cognitive stimulation program in older adults: randomized controlled trial. Sustainability 2023;15(3):2533. [doi: 10.3390/su15032533]

XSL•FO
RenderX

46. Karssemeijer EGA, Aaronson JA, Bossers WJ, Smits T, Olde Rikkert MGM, Kessels RPC. Positive effects of combined cognitive and physical exercise training on cognitive function in older adults with mild cognitive impairment or dementia: a meta-analysis. Ageing Res Rev 2017 Nov;40:75-83. [doi: 10.1016/j.arr.2017.09.003] [Medline: 28912076]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
FIIT-VP: frequency, intensity, duration time, type, volume, and progression
ITT: intention-to-treat
MCI: mild cognitive impairment
MoCA: Montreal Cognitive Assessment
PP: per-protocol
ROC: receiver operating characteristic

Edited by F Puga; submitted 26.Nov.2025; peer-reviewed by M Mancarella, W Zhu; revised version received 10.Feb.2026; accepted 02.Apr.2026; published 03.Jul.2026.

Please cite as:
Chia YC, Syed Annuar SN, Chew J, Lee Y, Lim WG, Yip WZ, Zhe Mun KAL, Jenkins M
Effects of a Combined Zumba and Video Game–Based Cognitive Intervention on Cognitive Function Among Older Adults in Klang Valley, Malaysia: Semirandomized Trial
JMIR Aging 2026;9:e88479
URL: https://aging.jmir.org/2026/1/e88479
doi:10.2196/88479

© Yook Chin Chia, Sharifah Nadhira Syed Annuar, Jactty Chew, Yunli Lee, Wei-Gene Lim, Wei Zern Yip, Kristin-Ann Leong Zhe Mun, Michael Jenkins. Originally published in JMIR Aging (https://aging.jmir.org), 3.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Motivations, Barriers, and the Potential of AI in Physical Activity Apps for US Older Adults: Interview Study

Abhinav Choudhry[1], BE, MFC, MPA; Jonathan Mshvelidze[2], BS; Sean P Mullen[3], PhD; Devorah Kletenik[4], PhD; Rachel F Adler[1], PhD

[1]School of Information Sciences, University of Illinois Urbana-Champaign, 501 E. Daniel St. MC-493, Champaign, IL, United States

[2]Macaulay Honors College, City University of New York, New York, NY, United States

[3]Department of Health and Kinesiology, College of Applied Health Sciences, University of Illinois Urbana-Champaign, Urbana, IL, United States

[4]Computer and Information Science, Brooklyn College, City University of New York, Brooklyn, NY, United States

Corresponding Author:
Abhinav Choudhry, BE, MFC, MPA
School of Information Sciences, University of Illinois Urbana-Champaign, 501 E. Daniel St. MC-493, Champaign, IL, United States

Abstract

Background: The older adult population is increasing rapidly and experiences a multitude of chronic conditions affecting quality of life. Physical activity and reduced sedentary behavior reduce adverse health conditions and help preserve functional independence and cognitive abilities, but the older adult age group has much lower levels of physical activity than recommended compared with other population demographics.

Objective: The study aimed to understand older adults' motivations toward health and physical activity and the role of technology in achieving their health and physical activity goals and to explore the attitudes toward the use of generative AI in health apps.

Methods: This study was qualitative and based on hour-long interviews with 20 older adults living in the United States. To elicit specific design preferences, participants were presented with visual probes of AI-generated avatars. The interviews were analyzed using thematic analysis, and the coding process drew on reflexive and grounded theory techniques.

Results: Older adults primarily wished to maintain functional independence, both physical and cognitive, prioritizing baseline maintenance rather than continuous optimization. They reported satisfaction with levels of physical activity, adjusting them to chronic conditions and injuries. They were active through everyday functions of living as well as structured exercise programs with many participants using wearable devices and smartphones to help monitor their activity. They reported being motivated by social accountability and gamification, although some found the latter condescending or anxiety-inducing. Accessibility issues were mentioned frequently, including problems with touchscreens, limited viewing, difficulty navigating interfaces with clutter, complex multitouch processes for undertaking actions, and steep learning curves. Computers were preferred for certain tasks over mobile phones due to easier viewability, keyboard access, and greater security features. Personalization was highly desired in health apps. There was a willingness to use more health technology, including generative AI, and they were willing to consider AI in both motivational and informational roles in apps as long as it was personalized. AI embodiment preferences varied widely but with an emphasis on relatability. However, there was marked skepticism toward data privacy, concerns about the loss of independence and self-efficacy, and doubts about AI's accuracy.

Conclusions: Older adults are primarily motivated by sustaining functional independence, although accessibility barriers in current health apps remain. Participants were open to AI for physical activity support, but expectations varied. They viewed AI as a passive "tool," a directive "coach," or a relational "companion." Considerable variation in preferences for gamification and visual design suggested that future apps require greater modularity. As skepticism over data privacy and AI accuracy remains high, greater transparency and explainability might be needed to engender trust.

(JMIR Aging 2026;9:e96114) doi:10.2196/96114

KEYWORDS

older adults; mobile health; mHealth; physical activity; generative AI; GenAI; wearables; human-computer interaction; technology acceptance; qualitative research; conversational agent

XSL•FO
RenderX

Introduction

Older Adults and Physical Activity

The demographic composition of the global population is trending toward a higher proportion of older adults, a shift that presents significant public health challenges across the world [1]. According to the 2020 Census, the US population aged >65 years stood at 55.8 million, or 16.8% of the general population [2]. This demographic grew nearly 5 times faster than the total population over the 100 years from 1920 to 2020 [2]. Older adult populations commonly experience a multitude of chronic conditions that accrue with age and also worsen due to decreasing physical activity (PA) with age [3]. PA and reduced sedentary behavior have proven benefits for older adults, including the prevention of chronic disease, maintenance of functional independence, and preservation of cognitive health [4]. Despite these known benefits, a large portion of the older adult population in the United States does not meet recommended PA guidelines [5], making the development of effective, scalable interventions for sustained PA in older adults increasingly important.

Technological Changes in Digital Health

Recent rapid developments in technologies have resulted in increasing applications to the health domain. Digital health is an umbrella term for all uses of digital technologies to improve health and well-being, and it includes computer-based, remote, mobile, and ubiquitous solutions [6]. On the whole, digital interventions have been found to be effective in PA interventions for older adults across many studies in the past few years [7], primarily by facilitating continuous self-monitoring and reducing logistical barriers to exercise. Such technologies, including smartphone apps, wearables, and AI-powered conversational agents, have shown the potential to be scalable and cost-effective solutions [8]. Real-life applications have also been employing gamification: the use of game design elements in nongame contexts [9]. Gamification has been found to be beneficial in the health domain [10] and shown to help older adults in having higher levels of motivation, enjoyment, and engagement while performing PA [11].

Research Gap Motivating This Study

However, a persistent gap has been observed between the high potential of these tools and their low adoption and sustained engagement by older adults [12]. This suggests a fundamental disconnect between the technological solutions offered and the actual needs, motivations, and practices of this user group. A recent systematic review found that limitations in capability, motivation, and social support were principal in limiting the uptake of mobile health (mHealth) solutions among older adults and suggested blended approaches of face-to-face interactions with technological interventions as a way to improve adoption [13]. Other research has identified 10 factors as determinants of older adults' adoption of mHealth: value, usability, affordability, accessibility, technical support, social support, emotion, independence, experience, and confidence [14]. Tailoring is also associated with much greater adherence rates [15]. Current older adult fitness technologies may also be lacking alignment with the needs and preferences of older adults [16].

To better understand their needs and preferences, we conducted interviews with 20 older adults to answer the following research questions:

1. What are older adults' foundational health goals and primary motivations regarding PA?
2. What are older adults' current practices regarding existing health technologies (eg, smartwatches and fitness apps) and what specific accessibility barriers do they experience?
3. How do older adults perceive the hypothetical use of AI-powered conversational agents for PA, and what specific roles and trust-related concerns do they envision?

Background

The mHealth Technology Landscape

The mHealth landscape for older adults extends beyond the relative ubiquity of smartphones and wearables to include tablets and emerging unobtrusive monitoring technologies. While personal mobile devices are effective for short-term PA promotion, their long-term utility is often limited by high abandonment rates once the novelty fades. There is also a persistent digital divide in active usage. Although hardware ownership is rising, particularly among highly educated older adults, sustained engagement remains a challenge across the demographic [17,18]. To mitigate the visual and dexterity deficits associated with these smaller screens, research demonstrates that tablet-based interventions can offer distinct usability advantages, allowing older adults to engage more effectively with complex health dashboards due to the larger screen real estate [19]. Furthermore, the landscape is shifting toward ambient computing and voice-first interfaces, which offer a lower barrier to entry for users with limited mobility or technical literacy [20].

Accessibility and the Limits of Usability

Historically, low adoption of mHealth technologies among older adults was attributed to physical and cognitive barriers. Systematic reviews confirm that age-related declines—specifically in manual dexterity, visual acuity, and cognitive processing speed—create significant hurdles for standard smartphone interfaces [12]. Wildenbos et al [21] demonstrated that basic interface deficits, such as small touch targets, complex navigation hierarchies, and insufficient contrast, can render an app functionally unusable regardless of the app's intent. However, addressing usability alone does not guarantee engagement. Foundational information systems theory, specifically the senior technology acceptance model (STAM), confirms that for older adults, attitude and utility ultimately mediate the impact of physical limitations [22].

A Look at Engagement: Foundational Techniques and Gamification Strategies

To drive engagement beyond simple utility, gamification is frequently used to foster motivation. While substantial evidence supports its general efficacy in improving health outcomes [23], the evidence specifically for older adults is less robust [10]. Considering self-determination theory (SDT), mechanics

XSL•FO
RenderX

designed to drive extrinsic motivation (eg, points, badges, and leaderboards) may fail to sustain engagement if they distract from the direct insight provided by self-monitoring or fail to support the user's intrinsic health goals. Tabak et al [24] illustrated this risk, noting that specific game elements can sometimes weaken the correlation between the app and the actual health goal if they distract from the direct insight provided by self-monitoring. Consequently, it is argued that gamification for this demographic must be applied with "adaptive engagement," allowing users to customize motivational affordances to their specific needs [11].

Social Connectedness Paradox and Privacy Barriers

Social support is widely recognized as a determinant of PA, but digitizing it presents unique risks; purely cognitive behavior change techniques are often less suitable for older adults compared to meaningful social support [25]. This uncertainty is often driven by a virtual paradox of connectedness: while older adults desire connection, they frequently reject the specific digital mechanisms used to facilitate it. This tension can be understood through socioemotional selectivity theory (SST), which posits that as time horizons shrink, older adults prioritize emotionally meaningful relationships with close friends and family over knowledge building with peripheral acquaintances [26]. Consequently, mHealth features that broadcast data to a wide, impersonal network conflict with this preference for selective intimacy. Orji [27] found that while competition can foster commitment, it introduces significant risks for this demographic, including heightened anxiety regarding self-esteem. Privacy barriers further compound this issue; older adults are significantly more likely to raise concerns about data confidentiality [28]. As a result, forced digital interaction can alienate users, who often prefer offline socialization, such as discussing results privately with family [24], over built-in cooperative elements that appeal only to specific, highly social user types [29].

Social Presence of Conversational AI

Conversational agents (CAs) are evolving rapidly, and those with unconstrained natural language input yield high user satisfaction and improved health outcomes, particularly in chronic disease management; they can foster a relationship similar to human coaching, where the natural language modality itself builds rapport [30]. This is supported by findings that the act of vocalizing health goals to an agent encourages psychological commitment [31,32]. Humans subconsciously apply social rules to computer interfaces, and this tendency toward anthropomorphism is particularly pronounced in older adults, who often use human social scripts to navigate digital interactions [33,34]. Embodiment appears to be helpful in gaining trust and acceptance among older adults [35,36]. An embodied virtual representation, commonly called an "*avatar*," can be a tool for persons for mediating communication and to derive meaning from interactions with the environment [37]. This perception as social entities increases engagement but introduces design risks [38]. Hyperrealistic avatars can elicit "uncanny valley" discomfort, while cartoonish representations may be perceived as patronizing [39]. Concepts of homophily suggest a preference for peer-like avatars to facilitate trust [40], and recent research affirms that properly designed peer-like CAs are effective partners in increasing PA and dietary behaviors [41].

The Research Gap

The existing literature is extensive but fragmented, often examining domains such as usability, gamification, and trust in isolation. Furthermore, the prevalence of interventionist studies creates a bias toward evaluation, where user insights are constrained by the specific features of a prototype rather than revealing the user's broader context. Consequently, there is a lack of comprehensive research that synthesizes the relationship between an older adult's foundational health goals, their existing technology habits, and their trust requirements for future AI interactions. To bridge this gap between current behaviors and future expectations, this study adopts a multitheoretical lens. By integrating the STAM for usability, SDT for motivational goals, and SST for social and privacy preferences, we conducted interviews with older adults about their attitudes toward PA, current PA practices and barriers, and their perceptions of an AI-based tool for supporting PA.

Methods

Participants and Recruitment

To meet study inclusion criteria, participants had to be 65 years of age or older, reside in the United States, and speak conversational English. Recruitment was conducted through university email listservs, flyers, and the newsletter of a prominent lifelong learning institute for older adults. Interested participants completed a demographic questionnaire to determine eligibility, and the research team screened them based on responses. After every fresh outreach effort (listserv, flyer, or newsletter), interviews were scheduled with all eligible participants. Thereafter, a fresh outreach effort was made until the target sample size (n=20) was reached. A total of 30 unique respondents filled out the screening survey, but only 28 individuals were eligible, of whom 20 finally consented to being interviewed.

Data Collection and Procedure

Interviews were conducted between May 2024 and January 2025 via videoconferencing software (Zoom; Zoom Video Communications) by the primary researcher. The sessions followed a semistructured format lasting approximately 1 hour. All interview data regarding PA and technology usage relied on participant self-report. The interview protocol was developed following guidelines for qualitative research questions [42] and covered health technology usage, challenges, and goals (Multimedia Appendix 1).

To elicit specific design preferences regarding AI embodiment, participants were presented with a series of visual probes. These consisted of AI-generated images (created using DALL-E models [OpenAI]; see Multimedia Appendix 2) depicting potential "*health coach*" avatars in fitness-oriented environments. The stimuli included the following:

- Realistic human avatars: young male, young female, older male, and older female.

XSL•FO
RenderX

- Robot avatars: a smiling bipedal robot and a plain bipedal robot.
- Animal avatars: a realistic dog, a cartoonish dog, a realistic cat, and a cartoonish cat.

These were intended to give participants a variety of choices in avatars. The clothing (sportswear) and background (sports field) of the avatars were selected so as to be associated with PA. There were no other selection criteria for the selected avatars, and the first image generated by the prompt was accepted as long as it followed the prompt accurately (eg, a prompt asking for a cartoonish dog got a cartoonish dog as a response and not a real dog image). Prompts used are provided in Multimedia Appendix 2.

Interviews were recorded on Zoom and transcribed using automated transcription services (Zoom and Whisper AI [OpenAI]), then checked for accuracy by the primary researcher.

Data Analysis

Data analysis used a hybrid inductive approach. The overall framework was reflexive thematic analysis as described by Braun and Clarke [43], elected for its flexibility in identifying overarching patterns across a heterogeneous dataset. Within this constructivist framework, we adopted specific coding procedures from the grounded theory tradition [44]—namely open line-by-line coding—strictly as practical tools for systematically building initial codes from the data upward. Acknowledging the epistemological debates between these methods, this pragmatic combination [45] was chosen specifically because the structured coding of grounded theory facilitated deep data familiarization, which then transitioned fully into the interpretive, theme-building processes of reflexive thematic analysis [43,46], without committing to a positivist theory-building paradigm.

The analysis proceeded in 3 stages. First, transcripts were subjected to initial coding to identify initial concepts. Every recorded video was viewed with the transcript in hand during this stage. Second, iterative grouping was applied to organize these codes into broader categories and explore relationships between them. Finally, these categories were synthesized into the final themes reported. To ensure rigor, initial coding was conducted independently by 2 researchers across the entire dataset (n=20). The researchers then brought their independent coding together, and contrasting perspectives were explored through collaborative reflexive dialogue in regular meetings to develop richer, multifaceted conceptual themes. Throughout this process, researchers maintained detailed reflexive notes. The process was conducted and tracked by hand, with pen and paper, whiteboard, sticky notes, and Microsoft Excel to ensure dependability and confirmability.

Thematic sufficiency was assessed informally by monitoring whether new interviews introduced significantly novel concepts. By the final 5 interviews (P16-P20), no new codes were generated during initial coding, suggesting thematic sufficiency within this sample [47]. However, we acknowledge that the demographic homogeneity of our participants may have contributed to earlier convergence and that saturation might not hold if sampling were extended to more diverse populations.

To categorize how participants imagined AI supporting their health and PA, our thematic synthesis organically generated an inductive typology based on participants' primary orientations toward the agent's role. Data interpretation was actively shaped by the coding team's diverse positionalities. The primary researcher had an eclectic multidisciplinary background across computer science engineering, finance, and other disciplines, coupled with personal experience as a caregiver for an older adult. Furthermore, an extensive experience of interviewing individuals in both industry and academic settings, personal engagement with structured exercise routines, nondigital heuristics, fitness apps, and various AI models provided a nuanced understanding of the tools being discussed. This combination had considerable impact on the semistructured interviews by shaping the conversation. The primary researcher's understanding of older adults' heuristics helped elicit deeper insights into technology usage and avoided viewing any struggles as skill deficiencies. Conversely, the secondary coder, a computer science major, brought a younger, digitally native perspective, while a practical background as a lifeguard provided a grounded lens on physical safety and fitness. During the "Four Eyes" evaluation, this contrast was highly productive. It established the credibility of our interpretations by preventing technical biases or overly sympathetic caregiving perspectives from overshadowing the older adults' verbal accounts. Finally, our detailed demographic reporting allows readers to judge the transferability of these insights.

Building on this reflexive approach, our thematic analysis was intentionally theoretically sensitized by the multitheoretical lens introduced earlier. To maintain the inductive integrity of reflexive thematic analysis, initial codes were generated directly from the participants' reported practice and perceptions. However, STAM, SDT, and SST provided the analytical vocabulary necessary to elevate these descriptive codes into our conceptual themes, ultimately shaping the resulting tool, coach, and companion typology.

Ethical Considerations

This study protocol was approved by the University of Illinois Urbana-Champaign's Institutional Review Board (project IRB24-0536). Informed consent was obtained from all participants prior to data collection. Qualitative data and interview transcripts were deidentified and assigned participant codes (eg, P1, P2) to ensure confidentiality. Participants received Amazon gift cards worth US $20 each as compensation for their time.

Results

Demographic Characteristics

Twenty participants (men: n=7, 35%; women: n=13, 65%) were interviewed regarding their experiences with health technology. The sample was highly educated, with all participants holding at least an undergraduate degree. The sample was relatively homogeneous in terms of ethnicity. All collected participant demographic data are presented in Table 1 and are self-reported.

XSL•FO
RenderX

Table . Participant demographics.

Participant code	Sex	Age group (years)	Education	Ethnicity
P1	F[a]	75 - 84	Graduate degree	White
P2	F	65 - 74	Graduate degree	White
P3	F	65 - 74	Graduate degree	White
P4	M[b]	65 - 74	Graduate degree	White
P5	F	65 - 74	Postgraduate degree	White
P6	F	75 - 84	Postgraduate degree	White
P7	F	65 - 74	Graduate degree	White
P8	F	75 - 84	Postgraduate degree	White
P9	M	65 - 74	Bachelor's degree	White
P10	F	≥85	Graduate degree	White
P11	F	65 - 74	Graduate degree	White
P12	M	75 - 84	Postgraduate degree	White
P13	M	75 - 84	Postgraduate degree	White
P14	M	65 - 74	Postgraduate degree	White
P15	F	65 - 74	Graduate degree	White
P16	M	65 - 74	Graduate degree	White
P17	F	65 - 74	Graduate degree	White
P18	F	75 - 84	Graduate degree	White
P19	M	75 - 84	Postgraduate degree	White
P20	F	65 - 74	Graduate degree	White

[a]F: female.

[b]M: male.

Attitudes and Motivations Related to PA

Attitudes Toward PA

Self-Reported Activity and Satisfaction

A majority of participants (12/20, 60%) reported maintaining active lifestyles, although the structure and intensity of this self-reported activity varied widely. (Subjective satisfaction with one's activity level may not necessarily correspond with meeting established PA guidelines.) Their PA regimens included structured exercises (eg, circuit training, resistance training, yoga, swimming, and sports) as well as lifestyle activities, such as gardening, walking the dog, and household maintenance.

An example of a highly structured, multifaceted routine was provided by P1:

> *Right now, I work out with a trainer three mornings a week. I walk every day. I'm always on the go. So, I usually log about 10,000 steps a day.*

Others integrated activity into their necessary daily tasks; P6 said:

> *I'm pretty active in the morning. I walk my dog twice... I do some gardening.*

Participants also reported satisfaction with their current level of PA. Their goal was maintenance, seeking to sustain their current baseline activity, rather than seeking continuous improvement. P8's comment reflected a common sentiment:

> *I'm not training for a marathon, but I get my steps in, I do my gardening... I feel good. At my age, I think I'm doing just fine.*

However, not everyone liked PA. P10 was happy with her low activity levels, listing multiple reasons for her dislike, including profound boredom ("It's so repetitive. I'd rather read a book") and sensory discomfort ("I just hate sweating").

Pragmatic Adaptation to Physical Barriers

This satisfaction was maintained through active negotiation with the physical realities of chronic health conditions and past injuries. These ranged from bone and joint issues to heart issues and past surgery history. P2 described the impact of chronic issues affecting the type of PA:

> *I've had high blood pressure since I was in my early 40 s...I wonder if I should be a little more gentle... I don't want to have surgery on my neck.*

It also manifested in a continual process of adaptation involving finding new activities to replace those no longer safe to perform. P5 swapped hiking with water aerobics:

XSL•FO
RenderX

> *I'd love to go on the long hikes I used to, but with the arthritis in my knee... I just can't. So now, I do my water aerobics. You do what you can.*

Motivations Aiding Health and PA

Functional Independence

Participants by and large were motivated with the aim of living a quality long life and preventing functional immobility. They framed having a certain level of functional fitness as a prerequisite to valued activities, such as being able to engage with family, avoiding emergency health circumstances, preventing cognitive decline, and also looking or feeling good. P13 framed it as a "straightforward" choice:

> *Well, at my age you have to decide, do I want to live or do I want to die?*

PA was similarly linked to cognitive preservation; participants viewed exercise as essential for maintaining mental clarity alongside physical capabilities (especially P13 and P18). Rather than seeking continuous improvement, their long-term objectives focused on sustaining baseline health—such as retaining energy and speed (eg, P6) and avoiding hospitalization (eg, P5).

Familial Motivation

Participants expressed a strong desire to maintain their health to see their descendants grow (eg, P8). P20 vividly described how losing weight would allow her to safely engage with her family:

> *I'll be able to play with my grandson instead of getting down and feeling like I can't get back up because it's hard to get up when you have all this extra weight.*

Sense of Community

Most participants (12/20, 60%) cited social factors and motivations when discussing their PA. Participants liked working with other people such as their spouse, children, friends, or an exercise class.

The sense of community was crucial for some and helped with exercise adherence, providing motivation and even a sense of security. Participants noted that group settings, such as yoga or classes, instilled the discipline required to maintain routines (eg, P11), while also offering peer-monitoring that made them feel safe (eg, P20). P4 was very expressive in describing the value of community:

> *The encouragement usually comes from enjoying the game with other people...[I] make sure I have enough good social activities for my health. I think that's important for the mental side of health. Coming into contact with social groups and through activities that I do... and also with my children to socially interact with them.*

Competition

A healthy sense of competition helped some participants, but this was typically with close family members or even against oneself. P2 mentioned an example of competition between mother and daughter based on achieving the activity levels tracked by Apple Health, which are shown as rings.

> *I found that I was more motivated when I was doing it with my daughter. She was sharing her rings, and I was sharing mine. I liked that competition.* [P2]

While intimate competition within the family proved motivating, broad group-based competition was often rejected. Participants favored collaborative or self-referential tracking (eg, P9), actively disliking generalized leaderboards. As P11 noted regarding university-wide step challenges, comparing oneself against strangers "doesn't motivate me, but against myself, yeah."

One participant was motivated by optimizing physical health, focusing on proactively improving performance in competitive bodybuilding:

> *We are not just messing around. We work out 6 days a week... It's a lot when you're competing.* [P14]

Accountability and Routine

Having someone else to discuss one's health objectives and progress with was listed as a key motivation by a number of participants, and this was often routine based. This accountability took various forms, ranging from formal weekly weigh-ins with nutritionists (eg, P3) to informal peer pressure from a spouse or "buddy" (eg, P14). Some participants even expressed a desire for "nagging reminders."

For some, this accountability was maintained through rigid schedules rather than direct supervision, such as committing specific days of the week to racquetball or fencing (eg, P4). Others relied on self-styled, task-based routines as a proxy for tracking overall activity, measuring their physical output. For example, P7 measured it by daily gardening milestones:

> *How long is the grass? How many weeds are in the garden? How many tomatoes and cucumbers have I picked today?*

Professional accountability from personal trainers or community instructors was highly valued for providing tailored motivation (eg, P17). More importantly, participants prized instructors who understood age-specific physiological issues and could proactively offer modifications for injuries or joint problems, ensuring the exercise remained both safe and effective (eg, P3).

Health Technology Apps: Attitudes and Barriers

Attitudes

Willingness to Adopt

Almost half (9/20, 45%) of the participants were current users of health apps, while others had used them in the past. Of nonusers, several people felt optimistic about using them. For some, potentially life-saving features such as fall detection and continuous heart rate monitoring were the key reason for use, particularly for those living alone (eg, P1). Some participants were more likely to use apps integrated with their current ecosystem, such as Apple's, essentially for the ease of compatibility with existing hardware.

XSL•FO
RenderX

Health apps were widely considered convenient for routine tracking. P6 preferred digital tools that automatically organized their data over manual methods such as writing in notebooks, noting that digital apps provided greater accountability (eg, P9). This convenience was further amplified by age-specific customization and seamless integration with existing services, such as health insurance platforms (eg, P3).

Unwillingness: Threat to Self-Efficacy and Feelings of Condescension

Six (30%) participants mentioned not wanting to track their health at all with reasons ranging from lack of need to it being effortful. A perceived loss of self-efficacy was a major barrier and participants actively resisted technology that felt overly prescriptive. P16 was emphatic:

> *I want to feel that I'm in control of the phone, not the phone in control of me.*

It also felt condescending for P19:

> *I don't need a pat on the back. I know I'm doing it. I don't need to write it down.*

App features were underused partly due to the feeling of loss of independence. P1 expressed this sentiment best:

> *I don't use a lot of the capacities... I don't like that it tells me to breathe. I know to breathe... I was raised to be independent, so I don't need some piece of software telling me that it's time to do something unless I've set an alarm.*

Barriers

Overview

All (20/20, 100%) our participants used technology regularly in their lives with many also possessing smartwatches (18/20, 90%) and voice-based personal assistants (8/20, 40%). Only 2 (10%) participants did not use smartphones, although they used computers. Nevertheless, many barriers and issues with technology were identified. These could be divided into accessibility issues and user preference mismatches. Although older adults aged ≥75 years reported more accessibility concerns, many aged 65 to 74 years also had similar concerns.

Accessibility: Physical Constraints

Accessibility issues proved to be a major barrier to usage. These ranged from physical constraints, such as arthritis affecting touchscreen use, to cognitive difficulties in learning new workflows. Small mobile interfaces actively hindered interaction, with participants struggling against illegible font sizes (eg, P18) and the physical pain of using tiny keyboards (eg, P9). Consequently, a significant number of participants explicitly preferred computers or tablets, seeking the physical affordances of larger screens and tactile keyboards (eg, P15). As P7 stated:

> *I got myself a keyboard with larger keys...My fingers miss a lot, especially if I'm doing texting... [Computer's more comfortable] because the print's bigger... I can see more on the screen.*

While some participants preferred voice interactions to bypass typing difficulties (eg, P4), they often encountered recognition errors, noting that systems struggled to understand their enunciation and frequently transcribed commands incorrectly (eg, P2).

Furthermore, modern devices were often viewed as unforgiving. Heightened touchscreen sensitivity and shifting interaction modes led to hesitation in adopting newer hardware. P13 expressed his exasperation:

> *Sometimes it's annoying that the more features they put on the iPhone or any of these, the harder these buttons are to navigate. I'm always accidentally going somewhere where I don't want to go... this is an iPhone 12. The reason I don't have the 15 is my daughter has it and ... just do that and screens disappear and new stuff comes on. It's too sensitive.* [P13]

Accessibility: Cognitive Overload

Participants frequently indicated a strong preference for visual simplicity, frequently reporting feeling overwhelmed with complex menus (eg, P11) or excessive data visualizations such as multicolored line graphs (eg, P9).

Unfamiliar iconography compounded this frustration, sometimes leading to outright abandonment, as P1 recalled:

> *When I go to an app that's just filled with symbols and I have no idea what they are, I just turn it off and go to something else.*

Extraneous options and constant notifications further contributed to information overload, occasionally driving users to the brink of returning their devices (eg, P13). Participants viewed these interruptions as highly disruptive to their workflow:

> *All these other things are coming up. Do you want to do this? Do you want to upgrade? Very distracting and very annoying... Let me just finish what I'm doing. Don't overload me with too much. People that are older are not used to that level of overload.* [P16]

Accessibility: Learning Difficulty and Complex Processes

Advancing age made learning more difficult for participants, leading to them expressing a marked preference for short and smooth learning curves (eg, P12). Multistep procedures were flagged as a specific barrier to adoption. P3 described this frustration:

> *The only time I think it really gets difficult as if I can't figure out what steps to take to get to what I want to do; if it's too many steps... do this now do that now... there are 10 or 12 steps... The easier it is to access the information I want, the more likely I am to use the technology.* [P3]

Convoluted navigation flows were censured, with some participants finding deeply nested tasks across multiple screens virtually impossible to complete. As P4 stated passionately:

XSL•FO
RenderX

> *Navigation to me is one of the cruelest things if it's not done correctly, especially for people that are a little older.*

Mismatch With User Needs

Apps were also criticized for not putting user needs first but instead prioritizing developer needs or lazily adapting study results:

> *I just think it's so distilled...not even the essence of what you need. It's more like some coder, some developer, who might look at a study or something put down what they think should be in the app, not what you need from the app.* [P5]

Participants strongly desired automated, minimal-effort setup processes for new wearables (eg, P4). When initial configuration proved too complex, users often had to surrender their independence and rely on younger family members for assistance before they could comfortably use the device (eg, P20).

Health App Preferences

Gamification Preferences

The majority (14/20, 70%) of participants showed an affinity for gamification, considering it useful for motivation. Many features of gamification came up, ranging from levels and points to leaderboards and competitive streaks.

Motivated by Gamification

Tangible rewards, such as points tied to health insurance premiums (eg, P2) or daily step targets (eg, P17), provided clear milestones that drove users to finish their daily routines. Participants generally responded well to this positive, automated feedback (eg, P9).

The motivation provided by awards and the ability to share with family added a social dimension. Reference points for distance helped too. This was best expressed by P20:

> *They would send those little award things. You've gone this many steps, walked this many miles...it's funny. In the beginning, I would send it to my kids. I said, "Look what I did." I was all excited, Oh, my gosh! I mean, it was kind of neat because it made you feel like you've gone this far already, and you don't think of it when you're doing it... It was nice. It's silly things, but it was just kind of neat to see.* [P20]

While some users thrived on the interpersonal competition of leaderboards (eg, P16) or the discipline of maintaining multiday streaks (eg, P17), others found deep motivation in personal, visual targets. The visual appeal of closing daily activity rings drove direct behavioral changes, as P19 confessed:

> *I like the [Apple Watch] rings... I want to close my rings. So, I will walk around the kitchen island... to get my steps in to close my rings... It's a bit silly, but it works.*

Unmotivated by Gamification

Conversely, 7 (35%) participants explicitly stated that they were unmotivated by gamification, dismissing such features as childish, distracting, or irrelevant to their intrinsic health goals. Its purported usefulness still failed to make them comfortable with it (eg, P12) and some even called it inherently unsuitable for their demographic, as expressed by P6:

> *I don't think people my age think in those terms [levels, badges, challenges].*

Excessive focus on gamification was also regarded as a source of anxiety. A quote by P8 illustrates this:

> *One of the one of my goals for the last 10 years is not to be too judgmental. I have friends who are really disappointed when they don't get their 10,000 steps... where it becomes worrisome to them... the worrying is worse than not getting the 10,000 steps in.*

Privacy Preferences

A quarter (5/20, 25%) of participants expressed very strong concerns about use of their data and privacy, citing fears of hacking (eg, P4), the commercial exploitation of even ostensibly anonymized personal data (eg, P5), and the financial extortion targeting their age demographic (eg, P13). This often manifested as a visceral and pervasive feeling of unease because the technology was literally "listening in." P13 was cogent:

> *[Siri] gives me the feeling that I'm being watched... I don't know who's all in on that conversation. The same with Amazon.*

Informational Preferences

When seeking health information, participants used a sophisticated, multisource approach characterized by conditional trust. This often involved a rigorous verification process for search engine results, where participants explicitly avoided relying on top hits alone (eg, P8). P20 elaborated that the priority was honesty, even if it resulted in complexity:

> *I always keep reading further... I'll read two, three things past that—sometimes it'll be the same... and sometimes it won't… it's got to be honest. Some things cannot be simplified. It's going to take some time and energy.*

However, this digital searching was balanced by participants who relied on nondigital sources. Participants grounded their trust in established medical authority (eg, P2), curated traditional media such as books and magazines (eg, P11), and their own somatic intuition, prioritizing what their bodies told them over external data (eg, P14).

Attitudes and Preferences Toward AI in Health

The majority of participants (14/20, 70%) demonstrated a pragmatic willingness to try an AI-powered coach. Even in the case of skepticism, outright rejection was rare and participants were curious and wished to give it a chance to prove its utility.

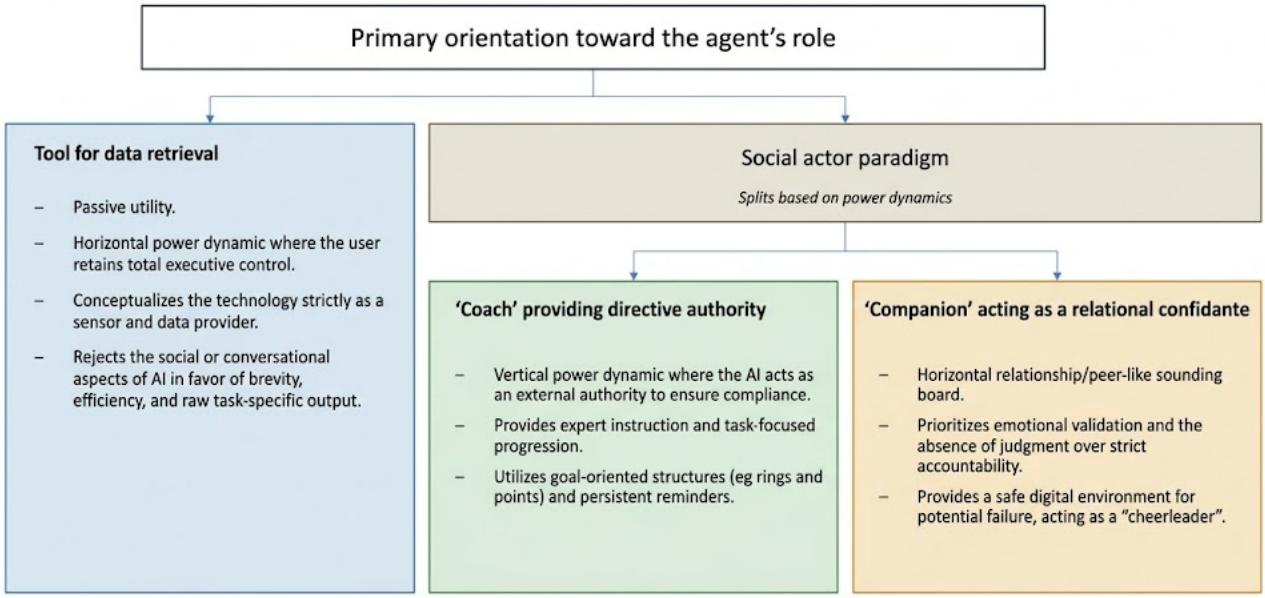
Role of AI

The analysis of participant preferences regarding the AI's role revealed distinct patterns in how users negotiate autonomy and support. Participants viewed AI in different ways based on desired power dynamics: a relatively passive "tool" (n=9, 45%), a directive "coach" (n=8, 40%), or a relational "companion"

XSL•FO
RenderX

(n=3, 15%). Figure 1 maps this visual typology, with individual participant classifications and supporting evidence provided in Multimedia Appendix 3.

Figure 1. Visual typology of AI roles: mapping the tool, coach, and companion.



Many participants (6/20, 30%) were optimistic that an AI coach could provide essential accountability. A substantial number (n=6, 30%) explicitly desired that the AI coach "push" them, tracking specific metrics, such as workout frequency and weightlifting (as exemplified by P2). They preferred this push to be supportive rather than punitive, as P9 envisioned:

> *If you had an AI app and there's some avatar that you talked to who said "hey did you do your walk today? Well, why not?" I think encouragement rather than admonishment would be better.*

The access to exercise data on the move was considered helpful for influencing workout length and intensity. The use was informational but its hypothesized use in session also had the perspective of accountability. P3 felt it could be useful to track distances against targeted norms:

> *If I was walking more than resting... am I going as far? ... Especially if I'm in a new place, like if I'm traveling and I don't know the route as well.*

The facet of accountability was also present in pushing users to do more, as with a real personal trainer. P18 lucidly explained a future scenario of use:

> *Especially if you're wearing an Apple watch, it would know your heartbeat ... exactly how much exertion you're putting into it. It could help keep track of your repetitions and give you advice... Let's say, I said, I'm ready to quit, and if it said, Wait a minute. You need another set of repetitions on certain things that would be great, you know. I think it would help you do it would help make sure you got you got to, because it's easy to say I've done enough when you don't have a personal trainer there.* [P18]

P11 was even more upfront about desiring an AI coach that was blunt rather than soft:

> *As opposed to being nice and saying, "oh, I think you can do better." No, I need someone that says, "come on, you're half-assing it." So, I don't need a drill sergeant, but I need something more than a soft person.*

The personal monitoring could also be in the form of informational suggestions, such as for sleep and enjoyable activities (eg, P13). Participants valued the ability to ask follow-up questions in a "back and forth" dialogue with a CA. For some, this dynamic shifted the AI's role from a strict instructor to a supportive peer. P8 explained:

> *Coach is a little strong... I would like to think it is maybe a companion, an associate... somebody who is sympathetic, empathetic but holds me to it.*

On the other hand, some participants could only perceive AI as a utilitarian tool, viewing it as inherently impersonal. P3 felt that her interactions would be task-specific only:

> *I think I would just be interested in the goal 'tell me how far I've gone.*

She, however, appreciated the convenience of querying health statistics via natural language. Participants such as P10 wholly rejected the concept of an AI health coach for either motivation or information. Ultimately, even those receptive to AI motivation frequently alternated between viewing it as an active coach and a passive informational tool.

Personalization

A strong desire for personalization was expressed by 11 (55%) participants. They wanted an app that would adapt to their specific health conditions and allow them to control their own goals. For instance, P1 wanted any advice to be catered to her situation, schedule, and current weather and not "generic walk 10,000 steps" advice. Similarly, P19 described desired flexibility in setting goals:

XSL•FO
RenderX

> *I would want to be able to set my own goals... be able to go in and adjust it if I'm having a bad day.*

Others expressed interest in getting personalized notifications (eg, P3), while P13 and P17 were enthusiastic about guidance based on specific limitations, such as mobility issues. There was also a desire for a multistage CA that adapted its complexity to each user's needs in "stages" proceeding gradually toward providing detailed research information (eg, P6).

However, some participants felt that current generative AI (GenAI) models lacked the contextual knowledge required to answer highly specific questions. P10 explained that desired specificity makes her prefer trusted websites over GenAI answers:

> *All it could give me would be just generalisations. I've got specific questions. How much does it cost to do this? How many days would I have to wait for this? Google AI would have no way of knowing those answers.*

Some participants wanted very strong referencing of the information that the AI would provide. P6 noted that giving, "a detailed list of even academic references...would be perfect," a sentiment echoed by P17 and P18. P5 wished that the AI be very "science based," share citations carefully, avoid "dumbing down" information, providing very detailed reasons alongside a quick summary:

> *I would want it to give me a reason...the important sources used to come to this conclusion... If you get that type of information, and they don't put their citations with it, then you don't know where the science is coming from.* [P5]

When asked about preferences for a hypothetical assistant, most participants would prefer a conversational assistant integrated into a device they already use, such as a smartphone or smartwatch, rather than a separate, standalone device. A minority (n=2, 10%) explicitly preferred multiple devices, but the rest preferred it to be included within their phone. Regarding notifications, a small number (n=2, 10%) preferred a "pull" model (user-initiated), whereas 4 (20%) participants explicitly valued "push" reminders as a necessary mechanism for accountability.

Visual Presentation

Embodiment was not a strong point of contention for most participants, with several explicitly expressing indifference toward it. However, among those who cared, 8 (40%) participants showed a clear preference for embodiment, while only 1 (5%) preferred a nonembodied CA. Furthermore, every participant showed at least some strong preferences among different avatar choices, displaying a clear preference for relatability. The majority (13/20, 65%) explicitly preferred human-like avatars, rejecting robot avatars for being unrelatable, using words such as "artificial" or "sci-fi" to describe them.

Within the human category, preferences were split by age. Some participants preferred older avatars, perceiving them as peers with shared experience while others preferred younger avatars, associating youth with up-to-date health knowledge, as in the case of P15:

> *I would prefer the younger one... because they might be more up-to-date... and I might feel that they would be more encouraging.*

Beyond age, participants expressed a specific preference for physical attainability, often rejecting avatars with "sculpted" or hyperfit physiques as discouraging (eg, P18). P5 ridiculed the notion:

> *I know zero ladies that have abs like that at her age.*

Similarly, P9 rejected aspirational avatars:

> *If the app were some guy with a six pack... that's just BS... I think more of a peer.*

This sentiment was not universal, though; in contrast, P17 wished to see aspirational physiques in the avatar:

> *I'd like to see what they do if I can do that.*

The embodied version of the AI also had an impact on personalization. Some participants wanted to feel a connection with their AI coach. P13 expressed:

> *I would like for it to more like a confidante... [Pointing to older male coach avatar] This picture feels like they're interested in what I have to say. He's the sort of guy I feel like I could talk with. I could sit down and have a beer with.*

Skepticism and Distrust

AI evoked skepticism among many due to a number of reasons, the most important among them being accuracy and the "black box" nature of the technology. Threat to privacy and perceived loss of independence were other concerns.

Adverse previous experiences often led to concerns about poor accuracy. An illustrative experience was that of P4, who had played around with ChatGPT:

> *The struggle I had when I was trying to get it to do exactly what I wanted it to do. A lot of times when it would fail at doing something, I didn't know its limitations or it couldn't figure out what I was trying to say.*

Similarly, P16 was "wary" for reasons of accuracy and information sources:

> *It pulls up a lot of things from a lot of places. I don't know how accurate everything is.*

P8 stressed the importance of accuracy over response speed:

> *If you want to sacrifice detail, accuracy and completeness for efficiency—for making it quick and easy—you're going to end up with a lot of stupid people, which we already have.*

The gradual creep of AI into daily life also caused discomfort, with some viewing it as an encroachment on independence. This discomfiture was expressed powerfully by P8 who questioned AI creators' underlying motives:

XSL•FO
RenderX

> *I want to question who's putting this together, who is creating. Am I looking at Zuckerberg... Musk? Are these people that I would trust on other things? ... do they have a motive other than just information for me... I know Big Brother's out there. I don't the idea of having a device that would listen to everything I say or do bothers me.* [P8]

Overall, there was a prominent lack of trust for AI among participants. Even those who had used a GenAI-based CA directly exhibited visible unease (eg, P20). Despite acknowledging its potential usefulness, P15 regarded herself as "suspicious" of GenAI and untrusting of its answers:

> *When I ask a question I always go down further, because I don't want to read the AI answer. I want to read the other answer.*

This caution frequently extended to fears regarding the prevalence of digital scams and phishing (eg, P9).

Discussion

Areas of Improvement in App Design for Older Adults

Matching Accessibility and User Needs

In this study, we found most participants to be current users of smart devices, but despite their comfort, there was a discrepancy between their general familiarity with apps and their actual comfort in practice, primarily due to accessibility issues. STAM [22] explains this discrepancy. When the procedural demands and learning curve of an app exceed users' cognitive or sensory thresholds, the perceived utility of the technology collapses. This resonates with previous work done by Vaportzis et al [48] on older adults working with tablet computers. Even older adults who are adept current users of technology would prefer simplicity: their usage alone does not mean that the apps meet accessibility needs.

Participants in our study appeared to be developing adaptive strategies to cope with a fast-changing technological landscape in conjunction with their own cognitive and sensory declines. STAM explains this retreat to familiar desktop computers for substantive tasks as a rational calculation (not a digital deficit) with users actively filtering out complex apps to manage age-related barriers [22]. Excess features contributed to cognitive overload, creating a steep learning curve [48] that acted as a hard filter between basic device usage and mastering new health apps. This burden is compounded by hardware seen as failing to accommodate aging users [12]. While previous literature suggests a potential benefit for voice interfaces [32], our participants expressed mixed preferences, suggesting that giving users more interface options might be best.

Our typology also revealed that for some participants oriented toward the tool archetype, the demand for interface simplicity did not equate to a desire for simplified content. These expert users expected granular, high-level data, demonstrating that accessibility in mHealth should streamline navigation without diluting the scientific complexity of the information provided. This aligns with previous literature cautioning against oversimplification [14].

Social Connectedness

Among our predominantly well-educated sample, socialization and accountability were significant drivers, although whether this generalizes to more diverse populations warrants further investigation. SDT's tenet of relatedness [49] illustrates why, for many, the exercise itself was secondary to the community interaction it provided. Facilitating these meaningful social connections acts as a vital relatedness-supportive mechanism [50]. Solitary app mechanics of most individual-centric mHealth tools fail to replicate this sense of communal security.

While digitizing social connection offers potential, it often conflicted with preferences for privacy and meaningful interaction. SST [26] accounts for this resistance, positing that as older adults prioritize emotionally meaningful relationships, mechanisms that broadcast personal data to peripheral networks conflict with their preference for selective intimacy. This theoretically contextualizes findings on social comparison by Koivisto and Malik [10] and Orji [27]. Thus, successful interventions would do well to shift from individual-centric models to community-centric ones that facilitate this safety net without enforcing high-pressure, privacy-invading competition.

Functional Independence

The older adults' emphasis on maintaining a high degree of functional independence for their full lifespan also resulted in more pragmatic reasons for adoption of health apps or wearables. However, monitoring of step count and heart rate is known to have a contributing effect on increasing levels of PA [51]. It thus might be advisable to consider in mainstream health apps features for monitoring current chronic conditions to attract older users. Moreover, daily routines too warrant inclusion. Many participants got PA done as part of daily activities, such as walking their dog. An app would connect more with people if it discussed PA as part of their daily life rather than as workout routines alone.

Gamification contributed to encouragement for many participants and also added to accountability. This is encouraging given the proven effectiveness of gamification in improving health interventions for older adults [52]. However, extrinsic feedback loops in the form of tracking and gamification were also disliked by some participants despite diligence in performing PA. Here, SDT is explanatory because it predicts that those exercising higher autonomy will reject these as patronizing or irrelevant [49,53]. The taxonomy of motivational behaviors by Ahmadi et al [50] categorizes this exact friction as an autonomy-thwarting mechanism. Conversely, participants adopting the "coach" orientation actively requested strict accountability and persistent prompts. This is also consistent with SDT because actively choosing a directive system to enforce personal discipline remains an autonomous decision and is a delegation. The divergence of participant needs on the complexity of PA tracking measures—optimization vs simplicity—also has implications. PA apps aimed at the older demographic would thus do well to have a simple mode with limited tracking and reduced clutter while allowing for an expanded mode that affords detailed tracking, because complex apps have shown no additional benefit over simple pedometers for PA, as found in a systematic review and meta-analysis [51].

XSL•FO
RenderX

Personalized Interactions

The findings indicate that participants possess distinct and often contradictory requirements that necessitate personalization. The data reveal divergent design preferences that cannot be adequately addressed by a single, static app. A similar divergence appears in gamification preferences, where participants were quite polarized, although slightly more expressed motivation from game elements. Among those motivated by gamification, the underlying drivers varied significantly between those seeking competition and those interacting with features pragmatically to track habits. Those who actively rejected gamification viewed it as patronizing or irrelevant. A standardized strategy, such as a default leaderboard, thus risks actively demotivating the intrinsic and routine-driven groups. Avatar choices too are very divergent, and a wide range of choices are better if embodiment is used because individual preferences are strong.

These conflicting preferences suggest that adoption for this demographic is conditional on the ability to tailor the experience. Under SDT [49], the widespread desire to set one's own goals reflects a fundamental need to preserve autonomy, driving participants to reject generic, prescriptive advice. Providing these meaningful choices operates as a core autonomy-supportive mechanism [50]. Consequently, implementing principles of autonomy support is critical, specifically the provision of choice and the minimization of external pressure [49]. Our findings also indicate that rigid system-defined metrics can further anxiety. Inflexible mandatory targets such as daily step counts can act as competency-thwarting features [50] that undermine user confidence. Therefore, streaks and targets that flexibly adjust to user circumstances may help avoid imposing undue pressure.

Future Implications for AI Integration and the Role of Trust

The study has a number of findings about the perception of AI that would be instrumental to future designers of apps using AI for aiding older adults in the context of PA. We found that participants were willing to try AI-based systems but nursed inherent skepticism. There are 2 major points of discussion concerning the use of AI by older adults. The first is on their hypothesized use cases, and the second is on the prerequisites that massively impact their adoption and perception of AI.

AI in Health Apps: Tool, Coach, and Sometimes Companion

The inductive typology that emerged from the data clearly aligns with functional distinctions in existing human-computer interaction literature. Specifically, it identifies whether participants viewed the AI as a "tool" for data retrieval [54], or following the social actor paradigm [33], as a "coach" providing directive authority [55] or a "companion" acting as a relational confidante [56].

The study shows that older adults are open to increased AI functionality in their apps. A remarkable proportion of the participants wished to feel a sense of being accountable to someone for their PA; even motivation could be attributed to discussions with an intelligent CA. Many use cases were informational but their presumed use was different from search engine behavior because they expected the CA to remember detailed personal context and provide information; this use case, although informational, is closer to that of a personal assistant for health. Previous literature too has found social support to be very helpful in increasing PA among older adults [57]. Their liking for gamification within apps was part of such motivation. However, gamification alone lacks a social component. While the need for social accountability is ideally satisfied by friends and family or even through personal trainers, many older adults lack them; it can be hypothesized that social support CAs could assist in such contexts, which could also help in alleviating loneliness [58]. Research has found that quality goal-setting practices, self-efficacy, and social support predict PA among older adults [59]. These can be facilitated by CAs, and this study too finds older adults to be willing to use them for these purposes. The CA having relatable embodiment and personality might help further.

However, based on our findings, health apps, especially with AI, have to take care not to negatively affect feelings of self-efficacy among older adults. We resonate with Klasnja et al [60] in that the AI should have the provision of modularity features—allowing users to use it as coach, tool, or companion. Providing this modularity ensures the app can dynamically satisfy the differing theoretical drivers of our participants—whether seeking accountability (SDT), relatedness (SST), or pure usability (STAM)—thereby protecting self-efficacy and respecting each user's distinct needs.

Trust and Privacy as Foundational Prerequisites

Trust emerged as a foundational prerequisite for adoption, linking privacy concerns, AI skepticism, and information preferences. Under STAM [22], a lack of trust acts as a hard stop for adoption; if an app feels unsafe, its usability and intended utility are lost. Our participants viewed AI as a tool requiring human oversight rather than as an autonomous authority. This heightened sensitivity aligns with research indicating that older adults with higher educational attainment are significantly more likely to raise data privacy concerns than their peers [28]. The findings on information accuracy and privacy are corroborated by other studies [61].

Trust is also closely related to accuracy and algorithmic transparency. The participants' distrust of AI stems directly from a practical awareness of GenAI's propensity for hallucinations [62]. Inaccurate AI poses severe hazards in health care, and older adults require systems that make their reasoning visible. This demand aligns closely with explainable AI frameworks tailored for older demographics [63,64]. Participants expected the technology to justify its outputs rather than act as an unquestionable authority [64]. Their rejection of dumbed-down summaries serves as a practical test of this requirement. They consistently preferred referenced explanations over simplifications because obscuring the source data is often perceived as a breach of honesty. Designing these explanations to be both transparent and accessible [63] forms the foundation necessary to earn their trust.

XSL•FO
RenderX

Preliminary Design Considerations

On the basis of participants' stated preferences and hypothetical reactions, we propose the following considerations for future designers of PA apps aimed at older adults. As these are derived from interview data and visual probe responses rather than observed usage behavior, they serve as preliminary guideposts rather than definitive rules.

1. Maintenance: Algorithms aimed at older adults may benefit from identifying a steady routine and rewarding consistency rather than aggressively pushing for continuous increases.
2. Social accounting: Designs could explore enabling easy sharing of physical activities and integrating social cues as a proxy for real-life social contact.
3. Relatability: More participants preferred age-appropriate avatars with average body types, suggesting these visual defaults may reduce psychological distance. This preference may be varied by cultural background and familiarity with digital agents.
4. Transparency: Providing direct links to reputable sources allows users to verify recommendations independently, supporting explainable AI principles.
5. Flexibility: Interfaces that support graceful failure states, such as allowing users to modify success criteria or auto-pause streaks during illness, better accommodate changing user physical realities.
6. Screen constraints: Developers might consider offering a companion view on tablets or desktops for complex data review while reserving mobile phones for passive tracking.
7. Passive data collection: Prioritizing automated data collection (eg, via wearables or sensors) can bypass the physical friction of manual entry.
8. Minimal interface changes: Participants expressed frustration with shifting layouts, suggesting that established navigation flows be changed only with substantial care.
9. Lifestyle integration: Rather than framing PA exclusively around structured workouts, apps should recognize and validate daily functional tasks (eg, gardening and housework).

Limitations

This study has several limitations. The participant sample consisted predominantly of individuals in an urban area of the Midwestern United States who are well educated, relatively financially secure, and relatively ethnically homogeneous. Importantly, they possess a high baseline of technological literacy that likely exceeds that of the broader older adult population. As a result, key findings regarding how comfortable participants were with technology, their openness toward AI's roles, and their specific skepticism toward AI may not apply to older adults with different educational, economic, or cultural backgrounds. We caution against overgeneralizing these findings; the nuanced views on AI observed here may represent a "best-case scenario" and cannot be automatically transferred to more diverse or underserved populations who face deeper digital divides. The study also relied on recall and self-reporting with participants reporting their own PA and technology use; an important caveat is that participants' reported satisfaction with their PA levels may be influenced by social desirability, particularly given that they were recruited from an academic community and interviewed by university researchers. Self-report of PA is known to be susceptible to overestimation [65]. While participants expressed confidence in their activity levels, the absence of objective measures (eg, accelerometry) means we cannot confirm whether their activity meets recommended thresholds.

Finally, the discussion regarding AI agents was based on static visual probes and hypothetical scenarios rather than real-world use; participants' trust and engagement might differ if they were using a functioning system over a long period. Visual probes in the future too could feature more diverse human avatars.

Conclusions

In this study, we investigated the current usage of technology for PA among older adults and their attitudes toward AI usage. We found that older adults are primarily motivated by the desire to sustain functional independence and meaningful social connections. They sought maintenance of baseline health rather than chasing continuous improvement. While existing apps and wearables were widely adopted, rigid navigation flows, cognitive overload, and a mismatch with user needs continued to act as significant accessibility barriers. Most participants were willing to try AI-assisted technology but their expectations varied. They viewed it either as a passive user-led "tool," a directive "coach" for strict accountability, or a relational "companion" for empathetic support. Older adults also possessed divergent views on gamification and tracking, and while the majority found it motivating, a significant minority found it condescending or anxiety inducing. Visual design preferences for CAs also varied widely, including split preferences on age, physique, and human versus nonhuman representation. Clearly, future AI apps should consider offering more modularity, personalization, and optionality.

Skepticism among older adults remained high due to pervasive concerns over data privacy, algorithmic hallucinations, and a perceived threat to user independence. AI that is personalized, transparent, and privacy protecting and that provides detailed referencing and explainable outputs when necessary is much more likely to be trusted. Greater use of AI-supported technologies for PA ultimately hinges on gaining user trust.

Acknowledgments

The authors thank the participants for their time and insights. The authors declare the limited use of generative AI in the research and writing process. Generative AI (GenAI; OpenAI's DALL-E 3 used through ChatGPT 4.0) was used to create the visual avatars for the interview probes. According to the GAIDeT (Generative AI Delegation Taxonomy) [66], the following tasks were assisted by GenAI tools under full human supervision: literature search and systematization, proofreading and editing, and

XSL•FO
RenderX

reformatting. The GenAI tool used was Gemini 3.1 (Google). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

The authors declared that no financial support was received for this work.

Data Availability

The qualitative data generated during this study are not publicly available to protect participant anonymity but are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: AC
Data curation: AC
Formal analysis: AC (lead), JM (equal)
Investigation: AC
Methodology: AC (lead), RFA (supporting)
Project administration: AC
Resources: RFA
Software: AC
Supervision: RFA
Validation: AC (lead), RFA (supporting), SPM (supporting)
Visualization: AC
Writing – original draft: AC
Writing – review & editing: AC (lead), DK (supporting), RFA (supporting), SPM (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview protocol.
[DOCX File, 14 KB - aging_v9i1e96114_app1.docx]

Multimedia Appendix 2
AI avatars and prompts used for avatar generation.
[DOCX File, 4705 KB - aging_v9i1e96114_app2.docx]

Multimedia Appendix 3
Typology of participant perceptions of AI in the physical activity context.
[DOCX File, 21 KB - aging_v9i1e96114_app3.docx]

References

1. Decade of healthy ageing: plan of action. World Health Organization. 2020. URL: https://www.who.int/publications/m/item/decade-of-healthy-ageing-plan-of-action [accessed 2026-08-13]
2. Older population and aging. US Census Bureau. 2019. URL: https://www.census.gov/topics/population/older-aging.html [accessed 2026-08-13]
3. Sun F, Norman IJ, While AE. Physical activity in older people: a systematic review. BMC Public Health 2013 May 6;13(1):449. [doi: 10.1186/1471-2458-13-449] [Medline: 23648225]
4. Cunningham C, O' Sullivan R, Caserotti P, Tully MA. Consequences of physical inactivity in older adults: a systematic review of reviews and meta‐analyses. Scandinavian Med Sci Sports 2020 May;30(5):816-827. [doi: 10.1111/sms.13616]
5. Physical activity guidelines for Americans midcourse report: implementation strategies for older adults. : US Department of Health and Human Services; 2023 URL: https://health.gov/sites/default/files/2023-06/PAG_MidcourseReport_508c_final.pdf [accessed 2026-08-13]
6. Mangion A, Piller N. What is digital health, eHealth, uHealth and Healthcare 4.0? Informatics and Health 2025 Sep;2(2):137-142. [doi: 10.1016/j.infoh.2025.06.002]
7. Alley SJ, Waters KM, Parker F, et al. The effectiveness of digital physical activity interventions in older adults: a systematic umbrella review and meta-meta-analysis. Int J Behav Nutr Phys Act 2024 Dec 18;21(1):144. [doi: 10.1186/s12966-024-01694-4] [Medline: 39696583]

XSL•FO
RenderX

8. Gentili A, Failla G, Melnyk A, et al. The cost-effectiveness of digital health interventions: a systematic review of the literature. Front Public Health 2022;10:787135. [doi: 10.3389/fpubh.2022.787135] [Medline: 36033812]
9. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining “gamification". 2011 Presented at: Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek ’11; Sep 28-30, 2011; Tampere, Finland p. 9-15. [doi: 10.1145/2181037.2181040]
10. Koivisto J, Malik A. Gamification for older adults: a systematic literature review. Gerontologist 2021 Sep 13;61(7):e360-e372. [doi: 10.1093/geront/gnaa047] [Medline: 32530026]
11. Kappen DL, Mirza-Babaei P, Nacke LE. Gamification of older adults’ physical activity: an eight-week study. 2018 Presented at: Proceedings of the Annual Hawaii International Conference on System Sciences; Jan 3-6, 2018; Hawaii, USA p. 1207-1216. [doi: 10.24251/HICSS.2018.149]
12. Czaja SJ, Charness N, Fisk AD, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352. [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
13. van Acker J, Maenhout L, Compernolle S. Older adults’ user engagement with mobile health: a systematic review of qualitative and mixed-methods studies. Innovation in Aging 2023 Mar 1;7(2):igad007. [doi: 10.1093/geroni/igad007]
14. Lee C, Coughlin JF. PERSPECTIVE: older adults’ adoption of technology: an integrated approach to identifying determinants and barriers. J of Product Innov Manag 2015 Sep;32(5):747-759. [doi: 10.1111/jpim.12176]
15. Dubbeldam R, Stemplewski R, Pavlova I, et al. Technology-assisted physical activity interventions for older people in their home-based environment: scoping review. JMIR Aging 2025 Sep 15;8:e65746. [doi: 10.2196/65746] [Medline: 40952779]
16. Tacca C, Vazquez Galvez A, Thompson IM, et al. How well do older adult fitness technologies match user needs and preferences? Scoping review of 2014-2024 literature. J Med Internet Res 2025 Sep 24;27:e75667. [doi: 10.2196/75667] [Medline: 40990372]
17. Daniels K, Quadflieg K, Bonnechère B. Mobile health interventions for active aging: a systematic review and meta-analysis on the effectiveness of physical activity promotion. Mhealth 2025;11:4. [doi: 10.21037/mhealth-24-41] [Medline: 39944853]
18. Langford A, Orellana K, Kalinowski J, Aird C, Buderer N. Use of tablets and smartphones to support medical decision making in US adults: cross-sectional study. JMIR Mhealth Uhealth 2020 Aug 12;8(8):e19531. [doi: 10.2196/19531] [Medline: 32784181]
19. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014 Mar 14;2(1):e11. [doi: 10.2196/mhealth.3048] [Medline: 25099993]
20. Sunshine J. Smart speakers: the next frontier in mHealth. JMIR Mhealth Uhealth 2022 Feb 21;10(2):e28686. [doi: 10.2196/28686] [Medline: 35188467]
21. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
22. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
23. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106. [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
24. Tabak M, de Vette F, van Dijk H, Vollenbroek-Hutten M. A game-based, physical activity coaching application for older adults: design approach and user experience in daily life. Games Health J 2020 Jun;9(3):215-226. [doi: 10.1089/g4h.2018.0163] [Medline: 32053023]
25. Zubala A, MacGillivray S, Frost H, et al. Promotion of physical activity interventions for community dwelling older adults: a systematic review of reviews. PLoS One 2017;12(7):e0180902. [doi: 10.1371/journal.pone.0180902] [Medline: 28700754]
26. Carstensen LL, Isaacowitz DM, Charles ST. Taking time seriously: a theory of socioemotional selectivity. American Psychologist 1999;54(3):165-181. [doi: 10.1037/0003-066X.54.3.165]
27. Orji R. Why are persuasive strategies effective? Exploring the strengths and weaknesses of socially-oriented persuasive strategies. In: de Vries P, Oinas-Kukkonen H, Siemons L, Beerlage-de Jong N, van Gemert-Pijnen L, editors. Lecture Notes in Computer Science: Springer, Vol. 2017:253-266. [doi: 10.1007/978-3-319-55134-0_20]
28. Alhammad N, Alajlani M, Abd-Alrazaq A, Epiphaniou G, Arvanitis T. Patients’ perspectives on the data confidentiality, privacy, and security of mHealth apps: systematic review. J Med Internet Res 2024 May 31;26:e50715. [doi: 10.2196/50715] [Medline: 38820572]
29. Kocadere SA, Çağlar S. Gamification from player type perspective: a case study. Educ Technol Soc 2018;21(3):12-22 [FREE Full text]
30. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 1;25(9):1248-1258. [doi: 10.1093/jamia/ocy072] [Medline: 30010941]
31. Wiratunga N, Wijekoon A, Palihawadana C, Cooper K, Mendham V. FitChat: conversational AI for active ageing. 2020 Presented at: Proceedings of the 1st Workshop on Intelligent Information Processing and Natural Language Generation; Sep 7, 2020; Santiago de Compostela, Spain p. 21-31 URL: https://aclanthology.org/2020.intellang-1.3.pdf [accessed 2026-09-30]

XSL•FO
RenderX

32. Kim S. Exploring how older adults use a smart speaker-based voice assistant in their first interactions: qualitative study. JMIR Mhealth Uhealth 2021 Jan 13;9(1):e20427. [doi: 10.2196/20427] [Medline: 33439130]
33. Reeves B, Nass CI. The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places: Cambridge University Press; 1996.
34. Pradhan A, Lazar A, Findlater L. Use of intelligent voice assistants by older adults with low technology use. ACM Trans Comput-Hum Interact 2020 Aug 31;27(4):1-27. [doi: 10.1145/3373759]
35. Loveys K, Sebaratnam G, Sagar M, Broadbent E. The effect of design features on relationship quality with embodied conversational agents: a systematic review. Int J of Soc Robotics 2020 Dec;12(6):1293-1312. [doi: 10.1007/s12369-020-00680-7]
36. Lee YH, Xiao M, Wells RH. The effects of avatars' age on older adults' self-disclosure and trust. Cyberpsychol Behav Soc Netw 2018 Mar;21(3):173-178. [doi: 10.1089/cyber.2017.0451] [Medline: 29638156]
37. Mennecke BE, Triplett JL, Hassall LM, Conde ZJ. Embodied social presence theory. 2010 Presented at: Proceedings of the 43rd Hawaii International Conference on System Sciences; Jan 5-8, 2010; Koloa, Hawaii p. 1-10. [doi: 10.1109/HICSS.2010.179]
38. Biocca F, Harms C, Burgoon JK. Toward a more robust theory and measure of social presence: review and suggested criteria. Presence: Teleoperators & Virtual Environments 2003 Oct;12(5):456-480. [doi: 10.1162/105474603322761270]
39. Mori M, MacDorman KF, Kageki N. The uncanny valley [from the field]. IEEE Robot Automat Mag 2012;19(2):98-100. [doi: 10.1109/MRA.2012.2192811]
40. Baylor AL. The design of motivational agents and avatars. Education Tech Research Dev 2011 Apr;59(2):291-300. [doi: 10.1007/s11423-011-9196-3]
41. Lyzwinski LN, Elgendi M, Menon C. Conversational agents and avatars for cardiometabolic risk factors and lifestyle-related behaviors: scoping review. JMIR Mhealth Uhealth 2023 May 25;11:e39649. [doi: 10.2196/39649] [Medline: 37227765]
42. Roberts RE. Qualitative interview questions: guidance for novice researchers. TQR 2020;25(9):3185-3203. [doi: 10.46743/2160-3715/2020.4640]
43. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
44. Strauss A, Corbin J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory, 2nd edition: Sage Publications; 1998.
45. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006 Mar;5(1):80-92. [doi: 10.1177/160940690600500107]
46. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
47. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
48. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 4;8:1687. [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
49. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist 2000;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68]
50. Ahmadi A, Noetel M, Parker P, et al. A classification system for teachers' motivational behaviors recommended in self-determination theory interventions. J Educ Psychol 2023;115(8):1158-1176. [doi: 10.1037/edu0000783]
51. Chaudhry UAR, Wahlich C, Fortescue R, Cook DG, Knightly R, Harris T. The effects of step-count monitoring interventions on physical activity: systematic review and meta-analysis of community-based randomised controlled trials in adults. Int J Behav Nutr Phys Act 2020 Oct 9;17(1):129. [doi: 10.1186/s12966-020-01020-8] [Medline: 33036635]
52. Chen L, Jang F, Li M, Zong W, Yu H. Effectiveness of mHealth-based gamified interventions on physical activity in older adults: systematic review. JMIR Aging 2025 Oct 31;8:e78686. [doi: 10.2196/78686] [Medline: 41171624]
53. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol Bull 1999 Nov;125(6):627-668. [doi: 10.1037/0033-2909.125.6.627] [Medline: 10589297]
54. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do: Morgan Kaufmann Publishers; 2003.
55. Kamphorst BA. E-coaching systems. Pers Ubiquit Comput 2017 Aug;21(4):625-632. [doi: 10.1007/s00779-017-1020-6]
56. Bickmore TW, Picard RW. Establishing and maintaining long-term human-computer relationships. ACM Trans Comput-Hum Interact 2005 Jun;12(2):293-327. [doi: 10.1145/1067860.1067867]
57. Thummasorn S, Sa-Nguanmoo P, Sirisattayawong P, et al. Social support-based physical activity that exerts beneficial effects for obese older adults with cognitive impairment via increasing participation in leisure-time physical activity. PLoS One 2025;20(6):e0325516. [doi: 10.1371/journal.pone.0325516] [Medline: 40587451]
58. Bickmore TW, Caruso L, Clough-Gorr K, Heeren T. 'It's just like you talk to a friend' relational agents for older adults. Interact Comput 2005 Dec;17(6):711-735. [doi: 10.1016/j.intcom.2005.09.002]
59. Park CH, Elavsky S, Koo KM. Factors influencing physical activity in older adults. J Exerc Rehabil 2014 Feb;10(1):45-52. [doi: 10.12965/jer.140089] [Medline: 24678504]

XSL•FO
RenderX

60. Klasnja P, Consolvo S, Pratt W. How to evaluate technologies for health behavior change in HCI research. 2011 Presented at: CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 7-12, 2011; Vancouver, BC, Canada p. 3063-3072. [doi: 10.1145/1978942.1979396]
61. An R, Ji M, Zhang S, et al. Conversational AI in healthcare: a domain-specific large language model approach. arXiv. Preprint posted online on Oct 12, 2023. [doi: 10.48550/arXiv.2310.08123]
62. Anh-Hoang D, Tran V, Nguyen LM. Survey and analysis of hallucinations in large language models: attribution to prompting strategies or model behavior. Front Artif Intell 2025;8:1622292. [doi: 10.3389/frai.2025.1622292] [Medline: 41098969]
63. Wolf CT, Ringland KE. Designing accessible, explainable AI (XAI) experiences. SIGACCESS Access Comput 2020 Mar 2(125):1-1. [doi: 10.1145/3386296.3386302]
64. van Leersum CM. What should AI explain to you? Visions of older adults regarding explainable AI. Gerontologist 2026 Feb 4;66(2):gnaf308. [doi: 10.1093/geront/gnaf308] [Medline: 41569379]
65. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Int J Behav Nutr Phys Act 2008 Nov 6;5:56. [doi: 10.1186/1479-5868-5-56] [Medline: 18990237]
66. Suchikova Y, Tsybuliak N, Teixeira da Silva JA, Nazarovets S. GAIDeT (Generative AI Delegation Taxonomy): a taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. Account Res 2026 Apr;33(3):2544331. [doi: 10.1080/08989621.2025.2544331] [Medline: 40781729]

Abbreviations

CA: conversational agent
GenAI: generative AI
mHealth: mobile health
PA: physical activity
SDT: self-determination theory
SST: socioemotional selectivity theory
STAM: senior technology acceptance model

Edited by P Jansons; submitted 06.Apr.2026; peer-reviewed by H Tilga, K Waters; revised version received 28.Jul.2026; accepted 14.Aug.2026; published 16.Oct.2026.

Please cite as:
Choudhry A, Mshvelidze J, Mullen SP, Kletenik D, Adler RF
Motivations, Barriers, and the Potential of AI in Physical Activity Apps for US Older Adults: Interview Study
JMIR Aging 2026;9:e96114
URL: https://aging.jmir.org/2026/1/e96114
doi:10.2196/96114

© Abhinav Choudhry, Jonathan Mshvelidze, Sean P Mullen, Devorah Kletenik, Rachel F Adler. Originally published in JMIR Aging (https://aging.jmir.org), 16.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Between a Hybrid Video-Guided Multicomponent Exercise Program With Staff Assistance and Free-Living Gait in Older Adults With Knee and Low Back Pain: Secondary Analysis of a Randomized Controlled Trial

Shuichi P Obuchi[1], PT, PhD; Tomoyuki Arai[2], PT, PhD; Hideaki Ishibashi[3], MD, PhD; Takashi Amari[2], PT, PhD; Hisashi Kawai[1], PhD

[1]Research Team for Human Care, Tokyo Metropolitan Institute for Geriatrics and Gerontology, 35-2 Sakae-cho, Itabashi-Ku, Tokyo, Japan

[2]Faculty of Health Sciences, Saitama Medical University, Saitama, Japan

[3]Ina Hospital, Saitama, Japan

Corresponding Author:
Hisashi Kawai, PhD
Research Team for Human Care, Tokyo Metropolitan Institute for Geriatrics and Gerontology, 35-2 Sakae-cho, Itabashi-Ku, Tokyo, Japan

Abstract

Background: Older adults with chronic joint pain often exhibit decreased mobility. Daily walking measures are being increasingly used in aging research as objective indicators of mobility.

Objective: In this study, we aimed to examine the effects of a 12-week hybrid exercise-based program on daily-life walking characteristics.

Methods: A total of 59 community-dwelling older adults (aged ≥65 years) with mild to moderate knee and/or low back pain were randomized to either a hybrid exercise group or an educational control group. Daily-life walking was monitored using an ankle sensor. Daily observation data were collected within 2 planned 7-day windows. Primary parameters included activity volume and gait characteristics. Statistical analyses were conducted based on a modified intention-to-treat principle using linear mixed-effects models, with a per-protocol set used for sensitivity analyses.

Results: The mixed-effects analysis based on the modified intention-to-treat principle (n=52 participants; 696 observations) suggested an effect of the intervention on improving walking quality, particularly step length (β=2.39 cm; *P*=.04; Hedges *g*=0.493), with a smaller effect with borderline statistical significance observed on improving walking speed (β=0.05 m/s; *P*=.06; Hedges *g*=0.496). Cadence, daily distance, daily steps, active energy expenditure, and accumulated 3 or more metabolic equivalent of task activities showed no clear difference between the groups. In the per-protocol sensitivity analysis (n=44), the same directional pattern of improvement was observed for step length (β=1.95 cm; *P*=.13; Hedges *g*=0.365) and walking speed (β=0.04 m/s; *P*=.14; Hedges *g*=0.368); however, the estimates were imprecise.

Conclusions: A 12-week hybrid exercise program may improve the quality of daily-life walking, with the clearest effect observed on step length. Given the attenuation after multiplicity-aware interpretation and in sensitivity analyses, these findings should be interpreted as suggestive rather than definitive.

Trial Registration: University Hospital Medical Information Network Clinical Trials Registry UMIN000057169; https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000065355

(JMIR Aging 2026;9:e97414) doi:10.2196/97414

KEYWORDS

hybrid video-guided multicomponent exercise program; randomized controlled trial; gait; active energy expenditure; pain

Introduction

Walking function maintenance is a key target for preserving independence in older adults because gait-related measures such as walking speed have been reported as early and strong predictors of subsequent decline in instrumental activities of daily living among community-dwelling older adults [1]. Musculoskeletal pain is a major barrier to maintaining walking function in older adults [2,3]. In Japan, approximately 42.2 million adults (41.2%) experience musculoskeletal pain, with 9.1 million (8.8%) reporting that the pain interferes with their daily activities [4]. The national burden of low back, hip, and

XSL•FO
RenderX

knee pain is projected to increase up to 2055 [5]. These burdens are particularly pronounced in older adults. In Japan's population-based ROAD (Research on Osteoarthritis Against Disability) study of community-dwelling people aged 60 years and older, the prevalences of radiographic knee osteoarthritis (Kellgren-Lawrence grade≥2) were 47.0% and 70.2% in men and women, respectively [6]. In the related LOCOMO (Longitudinal Cohorts of Motor System Organ) study, the prevalence of knee pain was 32.7%, and that of lumbar pain was 37.7%; overall, 12.2% of the participants reported both knee and low back pain concurrently [7]—the very combination targeted by the present study. Scalable approaches to musculoskeletal pain are therefore urgently needed to prevent pain-related functional decline in aging societies.

Numerous randomized controlled trials have demonstrated the effectiveness of multicomponent exercise programs in relieving symptoms and improving function in individuals with chronic knee and low back pain. For knee pain, programs combining stretching, strengthening, and aerobic exercises improve pain and physical function [8]. For low back pain, similar improvements require more than exercise programs alone; interventions integrating exercise with pain coping and behavioral skill training have shown sustained benefits [9-11]. Although many of these interventions were delivered under professional supervision, the high prevalence of knee and back pain in the older adult population highlights the need for more accessible exercise programs [12].

To improve access, audiovisual programs and remote digital interventions have been evaluated for musculoskeletal pain management. These include internet-delivered exercise programs with pain coping skill training for chronic knee pain and digital care programs for chronic low back pain [13,14]. However, adherence to home exercise programs declines over time. Key barriers to sustained participation include symptom flare-ups, fear of worsening pain, limited feedback or support outside of supervised settings, and difficulty integrating exercise into daily routines despite the flexibility of video-based delivery [15-17]. These practical constraints suggest that a hybrid program in which participants perform video-guided exercises during on-site sessions with staff support may be more effective.

Daily walking measures obtained in real-world settings are increasingly used in aging research and complement laboratory-based assessments as objective indicators of mobility [18-20]. Reference values for daily-life walking parameters have been reported in the Japanese population [21], and daily-life walking speed, as assessed using smartphones, has been used to evaluate frailty [22]. Daily-life walking speed is increasingly used as an objective indicator of mobility, disability, and other adverse health outcomes in older adults [23,24].

Thus, in this secondary analysis of data from a randomized controlled trial, we aimed to examine whether a hybrid exercise-based intervention improved objectively assessed daily-life walking outcomes. In the parent randomized controlled trial, participants were allocated to either a hybrid multicomponent exercise program (video-guided exercises with staff assistance during twice-weekly on-site sessions) or a once-a-week face-to-face class. Primary symptom outcomes will be reported separately. We also sought to compare the effects of the two programs on objectively assessed daily-life walking outcomes.

Methods

Study Design and Setting

This study involved a secondary analysis of data from a randomized controlled trial that evaluated a hybrid, exercise-based program targeting knee and low back pain. Participants in the intervention group attended on-site sessions involving video-guided exercise with staff assistance, whereas participants in the control group attended weekly face-to-face educational lectures regarding knee and low back pain. The trial was conducted at a community-based exercise facility in Iruma, Japan, and the study-related procedures took place at Saitama Medical University. The study period was from June 10, 2025, to October 17, 2025.

Participants

The study participants were community-dwelling adults living in the vicinity of the Karada Ugoki Kaifuku Center, Iruma. All participants who were randomized to either the intervention or control group and who contributed any valid daily-life walking data were eligible for inclusion. We included older adults aged 65 years and older who were able to walk independently and access the facility and experienced mild to moderate knee and/or low back pain, with pain intensity rated as less than 8 on a visual analog scale (VAS) from 0 to 10. The exclusion criteria were inability to walk independently, severe joint pain (VAS score ≥8), presence of medical conditions such as neurological or respiratory disease that made participation in the exercise program inappropriate as judged by a physician, and difficulty with decision-making because of dementia or impaired consciousness.

Sample Size

The sample size was determined based on the primary outcomes of the parent trial, specifically the Roland-Morris Disability Questionnaire and the Western Ontario and McMaster Universities Arthritis Index. A sample size of 20 participants per group was calculated to provide 90% power to detect the expected effect based on the mean intergroup difference and SD reported in a similar randomized controlled trial [1,2,25,26]. To allow for attrition and missing outcome data, the total recruitment target was set at 60 participants, comprising 30 participants per group. A separate sample size calculation was not performed for this secondary analysis.

Randomization and Blinding

Participants were randomly assigned to the intervention or control group in a 1:1 ratio using a computer-generated random sequence prepared by the Department of Biostatistics, Saitama Medical University Graduate School of Medicine. Specifically, the sequence was generated using the SAS software (PROC PLAN statement; SAS Institute) with a fixed random seed, and randomization was performed using permuted blocks with a block size of 6. Participant enrollment was conducted by the

XSL•FO
RenderX

second author (T Arai), whereas group assignment was implemented by the Biostatistics Group at Saitama Medical University and was performed independently of the intervention team. Blinding of participants and intervention staff was not feasible due to the nature of the interventions; however, the outcome assessors for performance-based measures were blinded, and questionnaire outcomes were self-reported.

Baseline Measures

Baseline characteristics for which data were collected in the parent trial included demographics (age and sex), anthropometrics and body composition (height, body weight, BMI, and skeletal muscle index), physical performance (maximal grip strength, standing time on one leg, and comfortable and maximal walking speed), pain and symptom measures (the low back and knee pain VAS score, Roland-Morris Disability Questionnaire score, and Western Ontario and McMaster Universities Arthritis Index subscale and total scores), and health-related questionnaires (25-question Geriatric Locomotive Function Scale, Kihon Checklist, Geriatric Depression Scale, and EQ-5D index).

Interventions

Participants in the intervention group received an on-site exercise program delivered through the Curves personalized exercise support system, which aimed to alleviate knee and low back pain through individually prescribed video programs with staff present to support participation and ensure safety. The program consisted of 21 distinct exercises comprising 12 stretching exercises and 9 muscle strengthening exercises. Stretching was performed to the maximal comfortable range of motion, with each of the 12 positions held for 30 seconds. Regarding muscle strengthening, upper-limb and lower-limb exercises used body weight resistance or auxiliary equipment such as elastic bands and a rubber ball, with intensity adjusted to a “somewhat hard” perceived exertion level. Strengthening exercises were performed for 30 seconds each at a speed based on the participants’ subjective intensity. The planned frequency and duration were 2 sessions per week for 12 weeks. The specific content and progression rules were updated based on the participants’ exercise proficiency and visit frequency according to age. Participants initially started with 15 exercises and progressed to performing the full 21-exercise program after 4 weeks. Staff monitored participant responses and implemented safety protocols, including discontinuing exercise for a specific body part if acute pain occurred and adjusting movements to a pain-free range of motion if pain was present. The control group received on-site educational lectures on knee and low back pain once weekly during the intervention period, and these participants were offered access to the intervention program after the completion of follow-up assessments.

Secondary Analysis Outcomes

This secondary analysis focused on daily-life walking outcomes derived from an ankle-worn activity monitor, WAlkX (Sanka). We analyzed daily activity volume outcomes, representing the daily accumulation of distance, steps, active energy expenditure, and activities graded at 3 or more metabolic equivalents of task (METs). In addition, we analyzed gait characteristics, representing the daily average values for walking speed, cadence, and step length. Participants were asked to wear the device continuously unless it caused discomfort.

The device was worn continuously during waking hours throughout the monitoring period, including during the on-site sessions; participants were not asked to remove it. Because the supervised sessions comprised video-guided multicomponent exercises performed in place rather than continuous overground walking, in-session activity did not generate the gait bouts detected by the daily-life walking algorithm and did not contribute to the step count or gait parameters. Therefore, the reported outcomes reflect habitual out-of-session daily-life walking.

The preintervention window was the 7 days starting 1 day after device placement. The postintervention window was the 7 days immediately preceding each participant’s last observed monitoring day, capturing late-program or end-of-program daily-life walking. The timing of the last monitoring day relative to the scheduled program end date varied across participants (in most cases, it clustered near program completion, with earlier cessation of monitoring in some participants), reflecting differences in device return and wear adherence.

Adherence, Protocol Deviations, and Safety

Adherence was quantified as the number of sessions attended and the percentage of prescribed sessions completed during the intervention period, with adequate adherence for the per-protocol analysis defined as completion of 80% to 120% of the prescribed sessions. Protocol deviations and adverse events were monitored throughout the study.

Statistical Analysis

For this secondary analysis, we use the term “modified intention-to-treat” (mITT) analysis because it included all randomized participants who contributed at least one valid daily-life walking observation in both windows rather than all randomized participants. Daily-life walking outcomes were analyzed using repeated-measure linear mixed-effects models (LMMs) incorporating all available daily observations from participants who contributed at least one valid observation in both the pre- and postintervention windows. The LMMs were prespecified based on the repeated-measure data structure rather than selected from a normality test of the outcome distributions. Daily observations were extracted from 2 planned 7-day windows anchored to the monitoring period. The preintervention window was defined as the 7 days starting 1 day after participants began wearing the device (ie, excluding the first day to avoid potential reactivity or partial wear effects). The postintervention window was defined as the 7 days immediately preceding the last observed monitoring day for each participant. Days with no recorded step counts (steps=0) were treated as missing daily walking observations and were excluded from model fitting. As a sensitivity analysis, we conducted a per-protocol set (PPS) analysis based on the trial ledger; we included participants classified as PPS eligible. Intervention effects were estimated using repeated-measure LMMs, with the primary parameter of interest being the group × phase interaction term, interpreted as a difference-in-differences (DID) estimate.

XSL•FO
RenderX

For each outcome, we fitted a repeated-measure LMM including fixed effects for group, phase, day in window, and the group × phase interaction, with a participant-specific random intercept.

Model estimates are reported as regression coefficients with 95% CIs. Models were fitted using restricted maximum likelihood in Python (version 3.14.0; Python Software Foundation) using statsmodels (version 0.14.6) with Pandas (version 3.0.1), NumPy (version 2.4.2), and SciPy (version 1.17.1).

Ethical Considerations

The study was conducted in accordance with the principles of the Declaration of Helsinki and the relevant Japanese ethical guidelines. It is reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) 2010 statement. The protocol was approved by the Saitama Medical University Institutional Review Board (approval 2025-004; May 24, 2025). The trial was registered in the University Hospital Medical Information Network Clinical Trials Registry (UMIN000057169) [27]. All participants provided informed consent prior to enrollment.

Results

Overview

The baseline characteristic data of participants in the parent trial are summarized in Table 1. Overall, 60 participants were assessed for eligibility in the present study; however, 1 withdrew after providing consent. Thus, 59 participants were randomized, with 30 assigned to the intervention group and 29 assigned to the control group. The flow of study participants is summarized in Figure 1 (CONSORT flow diagram).

XSL•FO
RenderX

Table . Parent trial participant characteristics[a].

Variables	Intervention group (n=30)	Control group (n=28)	Intergroup difference	Welch *P* value	SMD[b]
Sex (female), n (%)	13 (43.3)	11 (39.3)	4	.76	0.08
Age (y), mean (SD)	73.7 (5.4)	74.9 (5.2)	−1.2	.39	−0.23
Height (cm), mean (SD)	161.5 (10.1)	161.2 (8.3)	0.3	.90	0.03
Body weight (kg), mean (SD)	59.1 (10.0)	61.6 (11.2)	−2.5	.37	−0.24
BMI (kg/m^2), mean (SD)	22.6 (3.0)	23.6 (3.1)	−1.0	.22	−0.33
Skeletal muscle index (kg/m^2), mean (SD)	7.1 (1.0)	7.3 (1.1)	−0.2	.48	−0.19
Maximal grip strength (kg), mean (SD)	30.3 (7.0)	29.3 (8.0)	1.0	.62	0.13
Standing time on one leg (eyes open; s), mean (SD)	46.2 (41.3)	31.1 (30.8)	15.1	.12	0.41
Comfortable walking speed (m/s), mean (SD)	1.2 (0.2)	1.2 (0.3)	0.0	>.99	0.00
Maximal walking speed (m/s), mean (SD)	1.7 (0.3)	1.7 (0.3)	0.0	>.99	0.00
Low back pain VAS[c] score (0-10), mean (SD)	2.7 (2.4)	3.5 (2.8)	−0.8	.25	−0.31
RDQ[d] total score (range 0-24), mean (SD)	3.0 (3.2)	4.2 (4.8)	−1.2	.27	−0.30
Knee pain VAS score (0-10), mean (SD)	2.7 (2.2)	3.7 (2.7)	−1.0	.13	−0.41
WOMAC[e] pain subscale score (range 0-20), mean (SD)	3.0 (2.3)	4.2 (3.9)	−1.2	.16	−0.38
WOMAC stiffness subscale score (range 0-8), mean (SD)	1.1 (1.4)	2.0 (1.8)	−0.9	.04	−0.56
WOMAC physical function subscale score (range 0-68), mean (SD)	7.2 (8.6)	11.6 (11.4)	−4.4	.11	−0.44
WOMAC total score (range 0-96), mean (SD)	11.3 (10.8)	17.8 (15.5)	−6.5	.07	−0.49
LOCOMO-25[f] total score (range 0-100), mean (SD)	10.0 (6.6)	16.0 (10.9)	−6.0	.02	−0.67
Kihon Checklist total score (range 0-25), mean (SD)	4.5 (3.0)	5.5 (2.9)	−1.0	.20	−0.34
GDS[g] total score (range 0-15), mean (SD)	3.3 (3.2)	3.0 (2.7)	0.3	.70	0.10
EQ-5D index score, mean (SD)	0.9 (0.1)	0.8 (0.2)	0.1	.02	0.64

XSL•FO
RenderX

[a]Intergroup difference = intervention − control. The Pearson chi-square test was used for sex; the Welch *t* test was used for continuous variables. Baseline characteristics are shown for randomized participants with an available baseline assessment (intervention: n=30; control: n=28); 1 control participant lacked baseline data.

[b]SMD: standardized mean difference.

[c]VAS: visual analog scale.

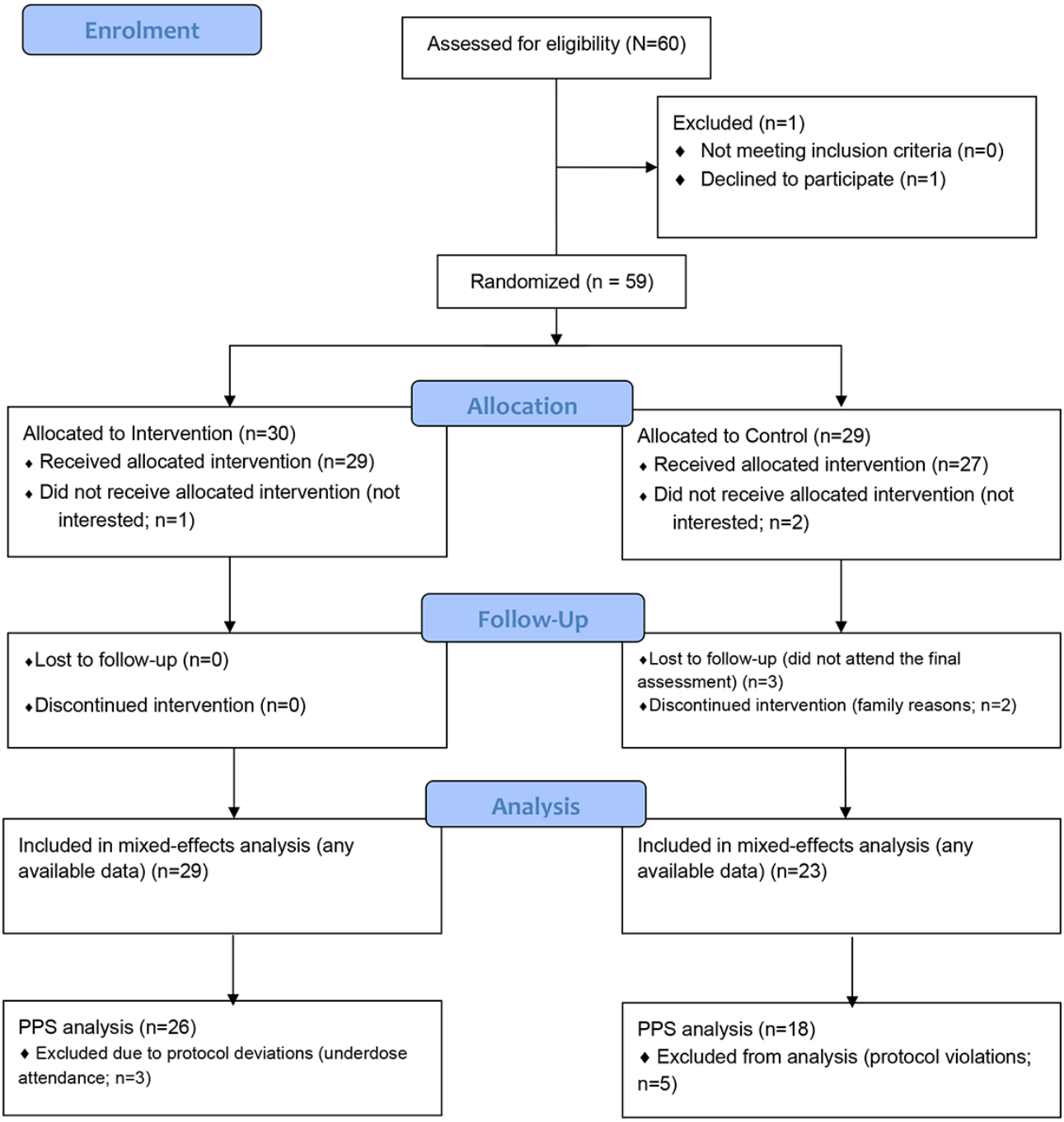
[d]RDQ: Roland-Morris Disability Questionnaire.

[e]WOMAC: Western Ontario and McMaster Universities Arthritis Index.

[f]LOCOMO-25: 25-question Geriatric Locomotive Function Scale.

[g]GDS: Geriatric Depression Scale.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of participant allocation, follow-up, and details of analysis sets, including the final modified intention-to-treat and per-protocol set (PPS) mixed-effects samples after the planned missing data handling for daily walking observations.



A total of 56 participants contributed at least one valid daily observation within the planned 7-day pre- and/or postintervention windows. Of these, 54 participants (intervention: n=29; control: n=25) were mITT eligible (ie, they contributed at least one valid daily observation in both windows). Three randomized participants contributed no valid

XSL•FO
RenderX

activity monitor data within the planned windows, and 2 further control participants had observations from only 1 window. Of the 54 mITT-eligible participants, 52 (intervention: n=29; control: n=23; 696 daily observations) contributed analyzable observations to the fitted mixed-effects models after treating days with no recorded steps (steps=0) as missing. The 2 remaining control participants had valid prewindow days but no analyzable (nonzero steps) postwindow days; therefore, they could not contribute a within-participant pretest-to-posttest contrast and were not estimable in the fitted models.

In the intervention group, the median number of sessions attended was 23 (IQR 21 - 24) over the 12-week period. As a sensitivity analysis, a PPS was defined based on the trial ledger. This yielded 47 PPS-eligible participants (intervention: n=27; control: n=20), of whom 44 contributed at least one valid daily observation in both windows (intervention: n=26; control: n=18).

Mixed-Effects Analysis Using All Available Data

In the mixed-effects models using all available daily observations, 52 participants contributed 696 daily observations across the 2 planned windows (before and after the intervention). The PPS sensitivity analysis contributed 44 participants and 595 observations.

In the mixed-effects models using 696 daily observations from 52 participants, the intervention was associated with a directional effect toward better gait quality, which was most clearly observed for step length. The DID estimate for daily-life step length was 2.39 cm (*P*=.04; Hedges *g*=0.493), whereas walking speed showed a smaller effect with borderline statistical significance (β=0.05 m/s; *P*=.06; Hedges *g*=0.496). Cadence did not differ clearly between the groups (β=1.26 steps per minute; *P*=.50). Daily steps and active energy expenditure also did not significantly differ between the groups (β=–51.0 steps per day and *P*=.92; β=4.84 kcal per day and *P*=.83). The effects of the intervention on daily distance and accumulated 3 or more MET activities were likewise null (β=0.08 km per day and *P*=.75; β=0.08 MET hours per day and *P*=.81; Table 2). As the gait quality outcomes were assessed alongside multiple related outcomes, these findings were interpreted as signals rather than as definitive confirmatory evidence.

Table . Results of the modified intention-to-treat analysis using mixed-effects models with all available daily observations.

Outcome variables	Participants, n	Observations, n	DID[a] coefficient (SE; 95% CI)	*P* value	Hedges *g*[b]
Distance (km/d)	52	696	0.08 (0.26; −0.42 to 0.59)	.75	0.076
Steps (steps per d)	52	696	−51.00 (521.02; −1072.21 to 970.20)	.92	0.000
Acal[c] (kcal per d)	52	696	4.84 (22.05; −38.38 to 48.06)	.83	0.100
MET[d] h (MET h per d; ≥3 METs)	52	696	0.08 (0.32; −0.55 to 0.71)	.81	0.079
Speed (m/s)	52	696	0.05 (0.03; −0.00 to 0.10)	.06	0.496
Step length (cm)	52	696	2.39 (1.18; 0.09 to 4.70)	.04	0.493
Cadence (steps per min)	52	696	1.26 (1.88; −2.43 to 4.95)	.50	0.312

[a]DID: difference in differences. DID indicates the group × phase interaction.

[b]Standardized mean difference in change (after − before) with small-sample correction.

[c]Acal: activity-related calorie expenditure.

[d]MET: metabolic equivalent of task.

PPS Analysis

The PPS sensitivity analysis included 44 participants who provided 595 daily observations. This analysis revealed the same directional pattern as observed with the primary model but with wider CIs. The DID estimate for step length was 1.95 cm (*P*=.13), whereas that for walking speed was 0.04 m/s (*P*=.14) and that for cadence was 1.03 steps per minute (*P*=.62). The effects of the intervention on daily steps and active energy expenditure remained null in the PPS analysis (β=–286.25 steps per day and *P*=.62; β=–8.88 kcal per day and *P*=.72). The effects on daily distance and accumulated 3 or more MET activities were also null (β=–0.01 km per day and *P*=.98; β=–0.09 MET hours per day and *P*=.80; Table 3).

XSL•FO
RenderX

Table . Results of the per-protocol analysis using mixed-effects models.

Outcome variables	Participants, n	Observations, n	DID[a] coefficient (SE; 95% CI)	*P* value	Hedges *g*[b]
Distance (km/d)	44	595	−0.01 (0.28; −0.56 to 0.55)	.98	−0.015
Steps (steps per d)	44	595	−286.25 (576.25; −1415.70 to 843.19)	.62	−0.121
Acal[c] (kcal per d)	44	595	−8.88 (24.67; −57.23 to 39.47)	.72	−0.095
MET[d] h (MET h per d; ≥3 METs)	44	595	−0.09 (0.35; −0.78 to 0.60)	.80	−0.071
Speed (m/s)	44	595	0.04 (0.03; −0.01 to 0.10)	.14	0.368
Step length (cm)	44	595	1.95 (1.29; −0.59 to 4.48)	.13	0.365
Cadence (steps per min)	44	595	1.03 (2.08; −3.04 to 5.10)	.62	0.125

[a]DID: difference in differences. DID indicates the group × phase interaction.

[b]Standardized mean difference in change (after − before) with small-sample correction.

[c]Acal: activity-related calorie expenditure.

[d]MET: metabolic equivalent of task.

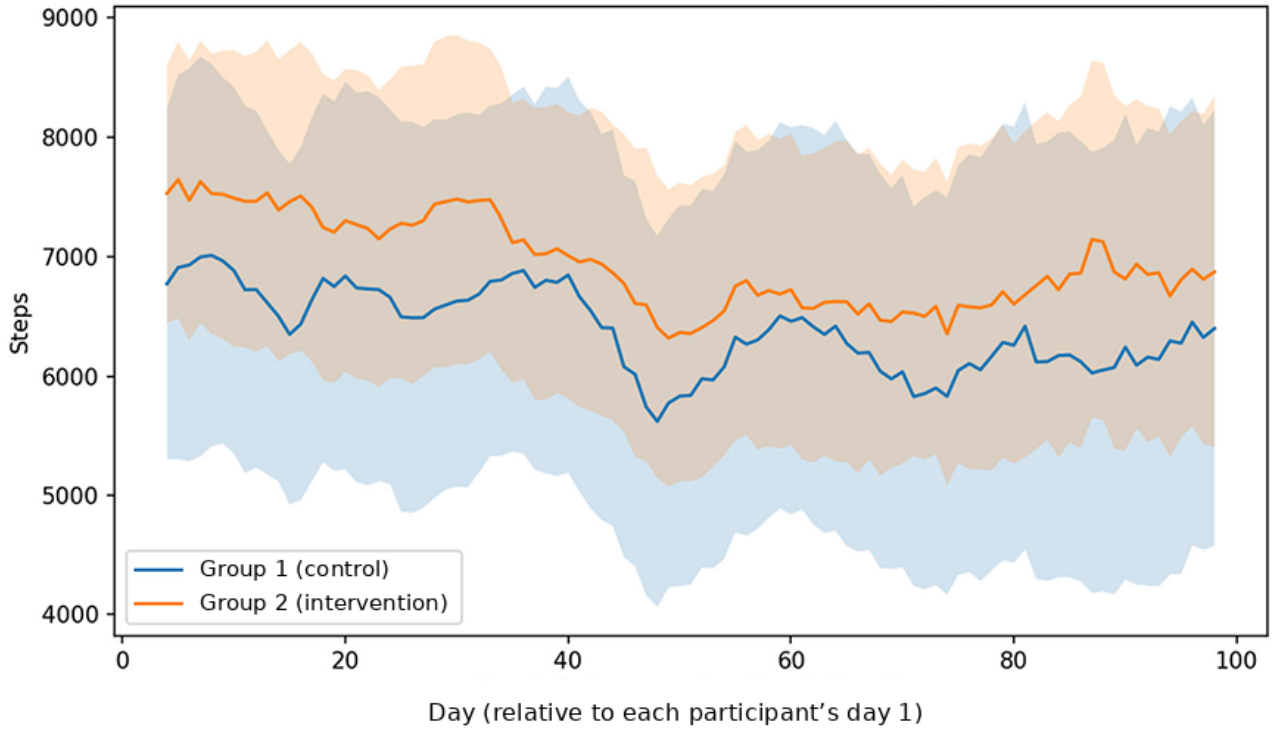
Trajectory of Daily-Life Walking Outcomes

Figure 2 shows descriptive day-by-day trajectories for daily steps (Figure 2A) and daily walking speed (Figure 2B) from days 2 to 98 aligned to each participant's day 1 (first recorded date). The daily walking speed trajectory showed a modest separation between the groups, emerging around the midintervention phase and persisting up to the end of follow-up, whereas the daily step trajectory did not show a similarly clear pattern. According to participation logs, the median program started on day 27 (IQR 26 - 29) in the control group and day 29 (IQR 28 - 30) in the intervention group. These figures are descriptive, and formal inference is provided by the mixed-effects models.

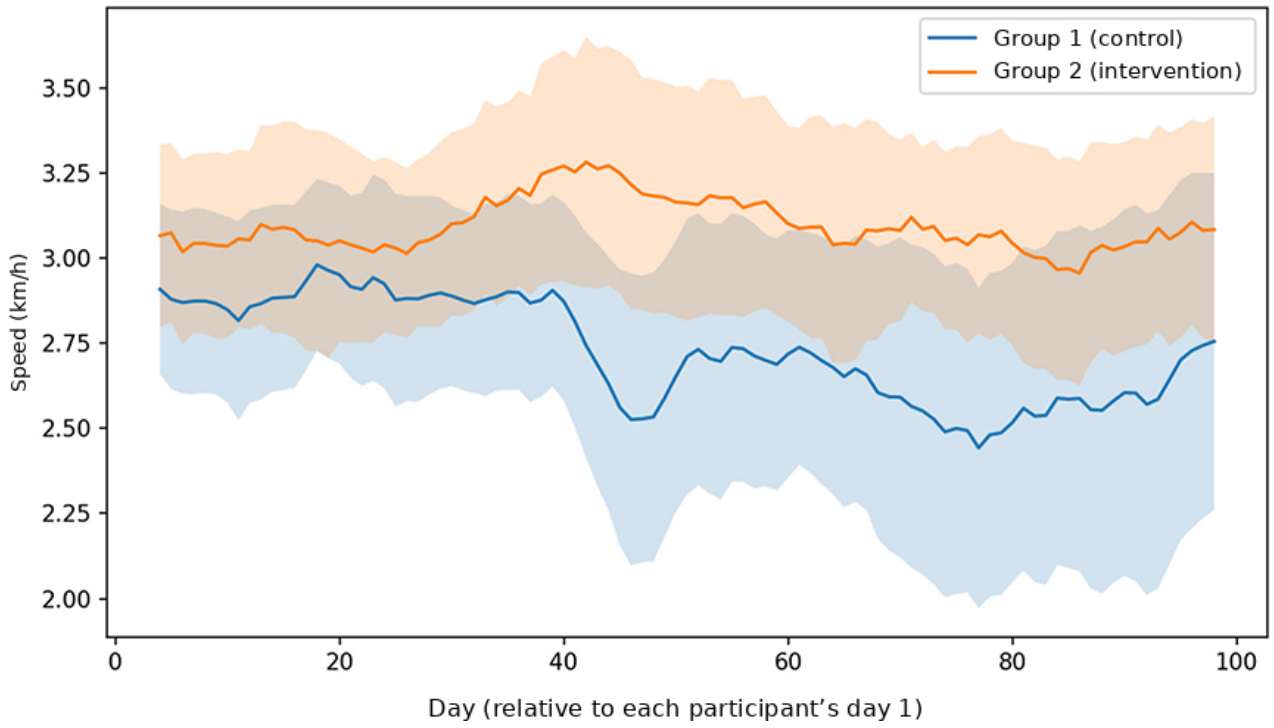
XSL•FO
RenderX

Figure 2. Daily steps and walking speed trajectories. (A)**.** Daily step trajectory (days 2 - 98). Lines depict the 7-day rolling mean of the group's daily mean step count, with each day indexed relative to each participant's day 1 (first recorded date). Shaded bands indicate smoothed 95% CIs for each group's daily mean (95% CI = mean ±1.96 × SE, also smoothed using a 7-day rolling mean). (B) Daily walking speed trajectory (days 2 - 98). Lines depict the 7-day rolling mean of the group's daily mean walking speed, with each day indexed relative to each participant's day 1 (first recorded date). Shaded bands indicate smoothed 95% CIs around each group's daily mean (mean ±1.96 × SE, also smoothed using a 7-day rolling mean). The figure is descriptive as formal inference is provided by the mixed-effects models.

A)



B)



XSL•FO
RenderX

Adverse Events

One adverse event was reported in the control group, and no adverse events were reported in the intervention group.

Discussion

Walking parameters, particularly walking speed and step length, are well-established indicators of adverse health outcomes such as mortality and disability [1,28]. There is accumulating evidence suggesting that these daily-life gait characteristics are also informative for health risk, including the risk of falls and frailty [20,22,29].

Principal Findings

In this study, we examined whether a hybrid multicomponent training program influenced daily-life walking. In mixed-effects analyses using all available daily observations, the clearest effect of the intervention was observed on step length, whereas cadence showed no clear effect of the intervention and walking speed showed a smaller effect, with borderline statistical significance. Although the step length estimate is of a moderate effect size (Hedges *g* of approximately 0.49), the small sample size of this secondary analysis yielded a wide CI; therefore, whether the change reaches a minimal clinically important difference remains uncertain. Activity volume outcomes, including daily steps and active energy expenditure, were not clearly altered by the training program. Overall, the effect of the intervention was more consistently observed on gait quality than on activity volume.

In this study, the most consistent effect of the intervention was observed on step length, whereas changes in cadence were small and imprecise, particularly in the PPS analysis. This pattern suggests that any intervention-related change in daily-life walking quality may have been driven more by step length than by cadence. A previous study demonstrated that daily-life cadence was associated with ambient temperature and may be modulated by autonomic responses [30]. Such influences may attenuate intervention-related changes in cadence compared with more mechanically determined gait parameters such as step length [31]. Although these findings remain exploratory, a between-group difference in step length of approximately 2 cm is not negligible from a gait quality perspective in older adults.

In contrast to the consistent improvements observed in gait characteristics, effects on activity volume outcomes were less clear. The multicomponent program did not incorporate components designed to promote behavior change, such as goal setting, self-monitoring, or structured activity coaching. Moreover, the intervention was intentionally designed as a scalable, video-guided program delivered without continuous specialist supervision. Although this format supports scalability, it may make it more difficult to embed individualized, multifaceted behavior change support. Prior evidence similarly indicates that interventions focused on symptoms and physical function do not necessarily lead to increases in objectively measured physical activity [32]. When increasing activity volume is an explicit goal, exercise-based interventions may need to incorporate behavior change techniques [33,34].

The descriptive walking speed trajectory (Figure 2B) was presented to provide an intuitive representation of how intergroup differences may have evolved over time. According to the participation logs, program initiation clustered around days 26 to 30 (control group: median day 27, IQR 26‐29; intervention group: median day 29, IQR 28‐30), and the speed trajectory suggested modest visual separation between the groups from approximately day 40. This timing corresponds to approximately 2 weeks after initiation and is broadly consistent with that in exercise-based randomized trials for knee osteoarthritis, in which intergroup differences could be detected within approximately 2 weeks for pain and functional outcomes and within approximately 3 weeks for changes in walking speed [35,36]. Notably, the trajectory plot suggested a maintenance pattern in the intervention group rather than a marked improvement. The observed intergroup difference may therefore reflect prevention in the intervention group of the decline observed in the control group.

Limitations

We observed some inconsistencies in statistical significance between the results of mixed-effects analyses using all available daily observations and those of the PPS sensitivity analyses for certain outcomes. Across multiple activity-related measures, including steps per day, daily walking distance, active energy expenditure, and MET-based indexes, no consistent intervention effect was observed. In the PPS set, the smaller sample size reduced statistical precision and widened CIs, making statistical significance less likely. However, the overall interpretation was consistent across analytic sets. Model residuals departed from normality for several outcomes, and in a distribution-robust sensitivity check, the one nominally significant outcome (step length) was no longer statistically significant; therefore, the findings are interpreted as suggestive.

This study has some limitations despite its strengths, such as objective, automatically derived free-living gait outcomes and comparison against an active control condition. The participants were recruited from a single setting and were required to attend sessions, which may limit generalizability. Additionally, this was a secondary analysis of activity monitor–derived outcomes, and the parent trial sample size was determined for symptom outcomes rather than specifically for the present free-living gait outcomes; accordingly, precision may have been limited for modest between-group differences. The cohort was clinically heterogeneous in baseline symptom burden; combined with the modest sample size of this secondary analysis, this may have increased between-participant variability and limited the precision of the between-group estimates. Participants who refused to wear the device or withdrew before outcome collection contributed no activity monitor data within the planned pre- and postintervention windows and, therefore, did not contribute data to the mixed-effects analyses of device-derived outcomes. The postwindow anchor was the last monitoring day rather than a fixed postcompletion assessment date, so the exact interval between intervention completion and the measured window differed between participants. The per-protocol definition may have also influenced PPS estimates, and the PPS findings should be interpreted as sensitivity analyses. Finally, we did not directly measure key mediators of

XSL•FO
RenderX

change (eg, pain trajectories), and longer follow-up will be needed to assess the durability and downstream clinical outcomes.

Practical Implications

This study has practical implications for scalable pain management services. Many exercise-based programs for knee and low back pain demonstrate benefits under professional supervision, yet widespread implementation is constrained by time and workforce. The current hybrid model of video-guided sessions with staff assistance may offer a feasible option that can be delivered at scale while potentially being consistently associated with a signal for daily-life gait quality. From a clinical perspective, the gait outcomes are clinically meaningful because gait speed is a robust indicator of overall health and ability to function and predicts adverse outcomes such as disability and mortality in older adults [28,37-39]. In the present study, the speed signal was modest and imprecise, whereas the step length signal was more consistent. Hence, clinical interpretation may focus primarily on gait quality rather than on activity volume while recognizing that even modest changes in physical performance measures may be clinically meaningful in older adults [40].

Conclusions

A 12-week hybrid multicomponent program combining video-guided exercise with staff assistance may be associated with a suggestive signal for better daily-life walking quality, with the clearest effect observed for step length in older adults with knee and low back pain. However, given the attenuation in sensitivity analyses and the multiplicity of outcomes assessed, these findings should be regarded as exploratory.

Acknowledgments

The authors appreciate the staff who assisted with this study and the study participants. During the preparation and revision of this manuscript, the authors used ChatGPT (OpenAI) to assist with rephrasing, English translation, proofreading, and language polishing of text originally written by the authors. All AI-assisted revisions were critically reviewed and edited by the authors, who take full responsibility for the final content of the manuscript.

Funding

This study was supported by Curves Japan through providing funding and the intervention venue and delivering the intervention services. The funder had no role in the data analysis, interpretation of the findings, or preparation of the manuscript.

Data Availability

The datasets analyzed during the current study are not publicly available due to ethical restrictions and participant confidentiality concerns but are available from the corresponding author to qualified researchers on reasonable request.

Authors' Contributions

SPO contributed to conceptualization, formal analysis, funding acquisition, investigation, and writing—original draft. T Arai, HI, and T Amari contributed to data curation, investigation, project administration, resources, and writing—review and editing. HK contributed to investigation, supervision, writing—original draft, and writing—review and editing.

Conflicts of Interest

None declared.

Checklist 1
CONSORT checklist.
[DOCX File, 29 KB - aging_v9i1e97414_app1.docx]

References

1. Masugi Y, Kawai H, Ejiri M, et al. Early strong predictors of decline in instrumental activities of daily living in community-dwelling older Japanese people. PLoS One 2022;17(4):e0266614. [doi: 10.1371/journal.pone.0266614] [Medline: 35381051]
2. Cruz-Almeida Y, Rosso A, Marcum Z, et al. Associations of musculoskeletal pain with mobility in older adults: potential cerebral mechanisms. J Gerontol A Biol Sci Med Sci 2017 Sep 1;72(9):1270-1276. [doi: 10.1093/gerona/glx084] [Medline: 28505228]
3. Hicks GE, Gaines JM, Shardell M, Simonsick EM. Associations of back and leg pain with health status and functional capacity of older adults: findings from the retirement community back pain study. Arthritis Rheum 2008 Sep 15;59(9):1306-1313. [doi: 10.1002/art.24006] [Medline: 18759261]
4. Suka M, Yoshida K. Musculoskeletal pain in Japan: prevalence and interference with daily activities. Mod Rheumatol 2005;15(1):41-47. [doi: 10.1007/s10165-004-0362-x] [Medline: 17028821]

XSL•FO
RenderX

5. Suka M, Yoshida K. The national burden of musculoskeletal pain in Japan: projections to the year 2055. Clin J Pain 2009 May;25(4):313-319. [doi: 10.1097/AJP.0b013e31818c00c5] [Medline: 19590480]
6. Muraki S, Oka H, Akune T, et al. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthritis Cartilage 2009 Sep;17(9):1137-1143. [doi: 10.1016/j.joca.2009.04.005] [Medline: 19410032]
7. Yoshimura N, Akune T, Fujiwara S, et al. Prevalence of knee pain, lumbar pain and its coexistence in Japanese men and women: the Longitudinal Cohorts of Motor System Organ (LOCOMO) study. J Bone Miner Metab 2014 Sep;32(5):524-532. [doi: 10.1007/s00774-013-0522-1] [Medline: 24213218]
8. Ettinger WH Jr, Burns R, Messier SP, et al. A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis. The Fitness Arthritis and Seniors Trial (FAST). JAMA 1997 Jan 1;277(1):25-31. [Medline: 8980206]
9. Moffett JK, Torgerson D, Bell-Syer S, et al. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences. BMJ 1999 Jul 31;319(7205):279-283. [doi: 10.1136/bmj.319.7205.279] [Medline: 10426734]
10. UK BEAM Trial Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. BMJ 2004 Dec 11;329(7479):1377. [doi: 10.1136/bmj.38282.669225.AE] [Medline: 15556955]
11. Lamb SE, Hansen Z, Lall R, et al. Group cognitive behavioural treatment for low-back pain in primary care: a randomised controlled trial and cost-effectiveness analysis. Lancet 2010 Mar 13;375(9718):916-923. [doi: 10.1016/S0140-6736(09)62164-4] [Medline: 20189241]
12. Yoshimura N, Muraki S, Oka H, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 2009;27(5):620-628. [doi: 10.1007/s00774-009-0080-8] [Medline: 19568689]
13. Bennell KL, Nelligan R, Dobson F, et al. Effectiveness of an internet-delivered exercise and pain-coping skills training intervention for persons with chronic knee pain: a randomized trial. Ann Intern Med 2017 Apr 4;166(7):453-462. [doi: 10.7326/M16-1714] [Medline: 28241215]
14. Cui D, Janela D, Costa F, et al. Randomized-controlled trial assessing a digital care program versus conventional physiotherapy for chronic low back pain. NPJ Digit Med 2023 Jul 7;6(1):121. [doi: 10.1038/s41746-023-00870-3] [Medline: 37420107]
15. Palazzo C, Klinger E, Dorner V, et al. Barriers to home-based exercise program adherence with chronic low back pain: patient expectations regarding new technologies. Ann Phys Rehabil Med 2016 Apr;59(2):107-113. [doi: 10.1016/j.rehab.2016.01.009] [Medline: 27050664]
16. Boutevillain L, Dupeyron A, Rouch C, Richard E, Coudeyre E. Facilitators and barriers to physical activity in people with chronic low back pain: a qualitative study. PLoS One 2017;12(7):e0179826. [doi: 10.1371/journal.pone.0179826] [Medline: 28742820]
17. Skou ST, Pedersen BK, Abbott JH, Patterson B, Barton C. Physical activity and exercise therapy benefit more than just symptoms and impairments in people with hip and knee osteoarthritis. J Orthop Sports Phys Ther 2018 Jun;48(6):439-447. [doi: 10.2519/jospt.2018.7877] [Medline: 29669488]
18. Mueller A, Hoefling HA, Muaremi A, et al. Continuous digital monitoring of walking speed in frail elderly patients: noninterventional validation study and longitudinal clinical trial. JMIR Mhealth Uhealth 2019 Nov 27;7(11):e15191. [doi: 10.2196/15191] [Medline: 31774406]
19. Del Din S, Galna B, Godfrey A, et al. Analysis of free-living gait in older adults with and without Parkinson's disease and with and without a history of falls: identifying generic and disease-specific characteristics. J Gerontol A Biol Sci Med Sci 2019 Mar 14;74(4):500-506. [doi: 10.1093/gerona/glx254] [Medline: 29300849]
20. van Schooten KS, Pijnappels M, Rispens SM, Elders PJ, Lips P, van Dieën JH. Ambulatory fall-risk assessment: amount and quality of daily-life gait predict falls in older adults. J Gerontol A Biol Sci Med Sci 2015 May;70(5):608-615. [doi: 10.1093/gerona/glu225] [Medline: 25568095]
21. Obuchi SP, Kawai H, Murakawa K. Reference value on daily living walking parameters among Japanese adults. Geriatr Gerontol Int 2020 Jul;20(7):664-669. [doi: 10.1111/ggi.13931] [Medline: 32378318]
22. Kawai H, Obuchi S, Ejiri M, Ito K. Association between daily life walking speed and frailty measured by a smartphone application: a cross-sectional study. BMJ Open 2023 Jan 6;13(1):e065098. [doi: 10.1136/bmjopen-2022-065098] [Medline: 36609327]
23. Takayanagi N, Sudo M, Yamashiro Y, et al. Predictivity of daily gait speed using tri-axial accelerometers for two-year incident disability among Japanese older adults. Sci Rep 2022 Jun 16;12(1):10067. [doi: 10.1038/s41598-022-14304-9] [Medline: 35710722]
24. Buchman AS, Yu L, Boyle PA, Shah RC, Bennett DA. Total daily physical activity and longevity in old age. Arch Intern Med 2012 Mar 12;172(5):444-446. [doi: 10.1001/archinternmed.2011.1477] [Medline: 22412115]
25. Rêgo TA, Ferreira AP, Villela DW, et al. Effects of mat Pilates on older adult women with knee osteoarthritis: a randomized controlled trial. J Bodyw Mov Ther 2023 Jan;33:136-141. [doi: 10.1016/j.jbmt.2022.02.007] [Medline: 36775508]

XSL•FO
RenderX

26. Itoh N, Mishima H, Yoshida Y, Yoshida M, Oka H, Matsudaira K. Evaluation of the effect of patient education and strengthening exercise therapy using a mobile messaging app on work productivity in Japanese patients with chronic low back pain: open-label, randomized, parallel-group trial. JMIR Mhealth Uhealth 2022 May 16;10(5):e35867. [doi: 10.2196/35867] [Medline: 35576560]
27. Efficacy of an exercise program for improving physical pain and movement: a randomized controlled trial. UMIN Clinical Trials Registry. URL: https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000065355 [accessed 2026-09-27]
28. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. JAMA 2011 Jan 5;305(1):50-58. [doi: 10.1001/jama.2010.1923] [Medline: 21205966]
29. van Schooten KS, Pijnappels M, Rispens SM, et al. Daily-life gait quality as predictor of falls in older people: a 1-year prospective cohort study. PLoS One 2016;11(7):e0158623. [doi: 10.1371/journal.pone.0158623] [Medline: 27389583]
30. Obuchi SP, Kawai H, Garbalosa JC, Nishida K, Murakawa K. Walking is regulated by environmental temperature. Sci Rep 2021 Jun 9;11(1):12136. [doi: 10.1038/s41598-021-91633-1] [Medline: 34108568]
31. Obuchi SP, Kawai H, Ejiri M, Ito K, Murakawa K. Change in outdoor walking behavior during the coronavirus disease pandemic in Japan: a longitudinal study. Gait Posture 2021 Jul;88:42-46. [doi: 10.1016/j.gaitpost.2021.05.005] [Medline: 33990001]
32. Oliveira CB, Franco MR, Maher CG, et al. Physical activity interventions for increasing objectively measured physical activity levels in patients with chronic musculoskeletal pain: a systematic review. Arthritis Care Res (Hoboken) 2016 Dec;68(12):1832-1842. [doi: 10.1002/acr.22919] [Medline: 27111744]
33. Willett M, Duda J, Fenton S, Gautrey C, Greig C, Rushton A. Effectiveness of behaviour change techniques in physiotherapy interventions to promote physical activity adherence in lower limb osteoarthritis patients: a systematic review. PLoS One 2019;14(7):e0219482. [doi: 10.1371/journal.pone.0219482] [Medline: 31291326]
34. Booth G, Howarth A, Stubbs B, Ussher M. The effectiveness of interventions and intervention components for increasing physical activity and reducing sedentary behaviour in people with persistent musculoskeletal pain: a systematic review and meta-analysis. J Pain 2022 Jun;23(6):929-957. [doi: 10.1016/j.jpain.2021.11.004] [Medline: 34856410]
35. Rosedale R, Rastogi R, May S, et al. Efficacy of exercise intervention as determined by the McKenzie System of Mechanical Diagnosis and Therapy for knee osteoarthritis: a randomized controlled trial. J Orthop Sports Phys Ther 2014 Mar;44(3):173-181. [doi: 10.2519/jospt.2014.4791] [Medline: 24450370]
36. Watanabe S, Someya F. Effect of body weight-supported walking on exercise capacity and walking speed in patients with knee osteoarthritis: a randomized controlled trial. J Jpn Phys Ther Assoc 2013;16(1):28-35. [doi: 10.1298/jjpta.Vol16_004] [Medline: 25792901]
37. Abellan van Kan G, Rolland Y, Andrieu S, et al. Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and Aging (IANA) Task Force. J Nutr Health Aging 2009 Dec;13(10):881-889. [doi: 10.1007/s12603-009-0246-z] [Medline: 19924348]
38. Cesari M, Kritchevsky SB, Penninx B, et al. Prognostic value of usual gait speed in well-functioning older people--results from the Health, Aging and Body Composition Study. J Am Geriatr Soc 2005 Oct;53(10):1675-1680. [doi: 10.1111/j.1532-5415.2005.53501.x] [Medline: 16181165]
39. Perera S, Patel KV, Rosano C, et al. Gait speed predicts incident disability: a pooled analysis. J Gerontol A Biol Sci Med Sci 2016 Jan;71(1):63-71. [doi: 10.1093/gerona/glv126] [Medline: 26297942]
40. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc 2006 May;54(5):743-749. [doi: 10.1111/j.1532-5415.2006.00701.x] [Medline: 16696738]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
DID: difference in differences
LMM: linear mixed-effects model
LOCOMO: Longitudinal Cohorts of Motor System Organ
MET: metabolic equivalent of task
mITT: modified intention to treat
PPS: per-protocol set
ROAD: Research on Osteoarthritis Against Disability
VAS: visual analog scale

XSL•FO
RenderX

Edited by R Lin; submitted 06.Apr.2026; peer-reviewed by O Conde, Y Hu; revised version received 28.Jul.2026; accepted 31.Aug.2026; published 08.Oct.2026.

Please cite as:
Obuchi SP, Arai T, Ishibashi H, Amari T, Kawai H
Association Between a Hybrid Video-Guided Multicomponent Exercise Program With Staff Assistance and Free-Living Gait in Older Adults With Knee and Low Back Pain: Secondary Analysis of a Randomized Controlled Trial
JMIR Aging 2026;9:e97414
URL: https://aging.jmir.org/2026/1/e97414
doi:10.2196/97414

© Shuichi P Obuchi, Tomoyuki Arai, Hideaki Ishibashi, Takashi Amari, Hisashi Kawai. Originally published in JMIR Aging (https://aging.jmir.org), 8.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Community Exergaming and Maintenance-Related Outcomes in Older Adults: Multi-Wave Observational Evaluation

Yuezhong Liu[1], PhD; Bing Xun Chia[1]; Herng Yeow Koh[1]; Quoc Nam Tran Nguyen[1]; Jiahui Jin[2]; Kimberly Miracle Wei Yan Tan[1]; Samuel Qing An Long[3]; Yin-Leng Theng[1,2], PhD

[1]Ageing Research Institute for Society and Education, Nanyang Technological University, 61 Nanyang Drive, Academic Block North (ABN), #01A – 01, Singapore, Singapore, Singapore

[2]Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore

[3]Sports Science & Management, National Institute of Education, Nanyang Technological University, Singapore, Singapore

Corresponding Author:
Yuezhong Liu, PhD
Ageing Research Institute for Society and Education, Nanyang Technological University, 61 Nanyang Drive, Academic Block North (ABN), #01A – 01, Singapore, Singapore, Singapore

Abstract

Background: Community exergaming may help older adults remain physically active by combining movement, enjoyment, and social interaction, but less is known about whether observed changes persist after supervised delivery ends in real-world community settings.

Objective: This study evaluated immediate and maintenance-related outcomes of the International Singapore Intergenerational National Games, a community-delivered exergaming program for older adults in Singapore. We examined positive mental health, intrinsic motivation, habit strength, physical activity, and exploratory implementation determinants aligned with the capability, opportunity, and motivation–behavior (COM-B) framework.

Methods: We conducted a multi-wave observational evaluation from October 2025 to January 2026 among 280 older adults recruited from 14 senior activity centers or active aging centers in Singapore. Assessments were conducted at baseline (T0), immediately after an 8-week supervised training phase (T1), after a 4-week unsupervised free-play phase (T2), and at a later follow-up for selected measures in 6 centers (T3). Immediate within-person changes were examined using 2-tailed paired *t* tests. Maintenance-related trajectories were examined using adjusted linear mixed-effects models. Exploratory analyses examined whether COM-B–aligned capability, opportunity, and motivation composites were associated with physical activity trajectories.

Results: Participants had a mean age of 72.54 (SD 6.15) years, and 77.1% (216/280) were female. From T0 to T1, mean positive mental health increased from 55.19 (SD 9.79) to 57.59 (SD 8.70; Cohen *dz*=0.28; *P*<.001), mean intrinsic motivation increased from 5.94 (SD 1.00) to 6.29 (SD 0.81; Cohen *dz*=0.40; *P*<.001), and mean habit strength increased from 3.34 (SD 1.11) to 3.81 (SD 0.84; Cohen *dz*=0.42; *P*<.001). In adjusted mixed models, habit strength remained higher at T2 than at baseline (β=0.362, 95% CI .113-.611; *P*=.005), and intrinsic motivation remained higher at T2 (β=0.397, 95% CI .156-.638; *P*=.001) and T3 (β=0.305, 95% CI .009-.601; *P*=.04). Positive mental health was higher immediately after supervised training (β=4.140, 95% CI 2.169-6.111; *P*<.001) and at T2 (β=7.390, 95% CI 5.370-9.410; *P*<.001) than at baseline. Among the COM-B domains, only opportunity was associated with greater immediate posttraining physical activity improvement (time × opportunity interaction β=26.464, 95% CI 3.191-49.737; *P*=.03), with no evidence of a sustained interaction at T2.

Conclusions: In this observational evaluation, positive mental health, intrinsic motivation, and habit strength were higher at T2 than at baseline among respondents with available follow-up data, and intrinsic motivation was also higher at the selected-center T3 follow-up. Because this study lacked a control group and attrition increased over time, these findings represent observed longitudinal trajectories rather than causal evidence of program effectiveness. Opportunity-related supports such as access, reminders, routines, staff facilitation, and peer participation may be useful when community exergaming transitions from supervised delivery to reduced-intensity free play, although this interpretation remains exploratory.

(JMIR Aging 2026;9:e100007) doi:10.2196/100007

KEYWORDS

exergaming; older adults; digital health; physical activity; habit formation; motivation; implementation science; intergenerational; community-based intervention

XSL•FO
RenderX

Introduction

Community and group-based exergaming have attracted growing interest as a potentially scalable strategy for promoting later-life activity by combining movement, enjoyment, and social interaction. However, less is known about whether implementation-relevant gains persist once structured supervision is withdrawn in routine community settings. Across older adult studies, the most consistent benefits have been observed for balance and mobility, although effect sizes vary and intervention designs remain heterogeneous [1-5]. Cognitive and psychosocial outcomes are more mixed: reviews suggest possible benefits for mood, cognition, and health-related quality of life, but they also emphasize variability in delivery platforms, dose, comparator conditions, and measurement [2,6-8]. There is recent evidence likewise suggesting that different exergame emphases may yield different cognitive vs balance gains, underscoring the importance of intervention design [9].

Social and implementation questions remain especially important for community translation. Social outcomes are still measured inconsistently, yet there is available evidence suggesting that socially interactive or multiple-player exergaming formats may enhance loneliness-related and depressive outcomes more than solitary modes [10,11]. At the same time, follow-up beyond the immediate posttest remains limited, leaving uncertainty about what happens when staff prompting and structured supervision are reduced [2,5,12]. Qualitative and mixed methods studies further indicate that enjoyable gameplay alone is insufficient for implementation; familiarization, clear instructions, accessible design, confidence building, and social support all shape uptake and continued use [12-14].

The capability, opportunity, and motivation–behavior (COM-B) model provides a useful framework for examining these questions because it conceptualizes behavior as arising from capability, opportunity, and motivation [15]. Within this framework, intrinsic motivation and habit strength can be understood as motivational processes that may support behavioral maintenance, whereas opportunity reflects the social and environmental conditions that make continued activity possible after supervision is reduced. Self-determination theory further suggests that autonomy, competence, and relatedness-supportive delivery practices can foster more self-sustaining forms of motivation; recent consensus work has translated these principles into observable motivational behaviors that may be useful for community program delivery [16]. In a pragmatic community program, opportunity-related conditions such as access, reminders, peer presence, and light-touch support may be especially relevant once formal supervision decreases. The present study evaluated the International Singapore Intergenerational National Games (I-SING), a real-world community exergaming program delivered through senior activity centers and active aging centers in Singapore. The evaluation captured a supervised training phase, a deliberately designed 4-week unsupervised free-play phase, and later follow-up for selected measures.

Guided by the COM-B framework, we examined immediate and maintained changes in positive mental health, intrinsic motivation, habit strength, and physical activity together with descriptive feedback on program acceptability and barriers to continued participation. Because this study used an observational pretest-posttest design, the analyses were intended to describe trajectories and implementation-relevant associations rather than establish causal effectiveness.

This study tested 3 hypotheses:

- Hypothesis 1—positive mental health, intrinsic motivation, and habit strength would improve immediately after the supervised training phase.
- Hypothesis 2—intrinsic motivation and habit strength would remain higher during reduced-intensity follow-up than at baseline, whereas gains in positive mental health were expected to be less stable after structured social contact decreased.
- Hypothesis 3—higher COM-B–aligned capability, opportunity, and motivation scores would be associated with more favorable physical activity trajectories, with opportunity expected to be particularly relevant after supervision was reduced.

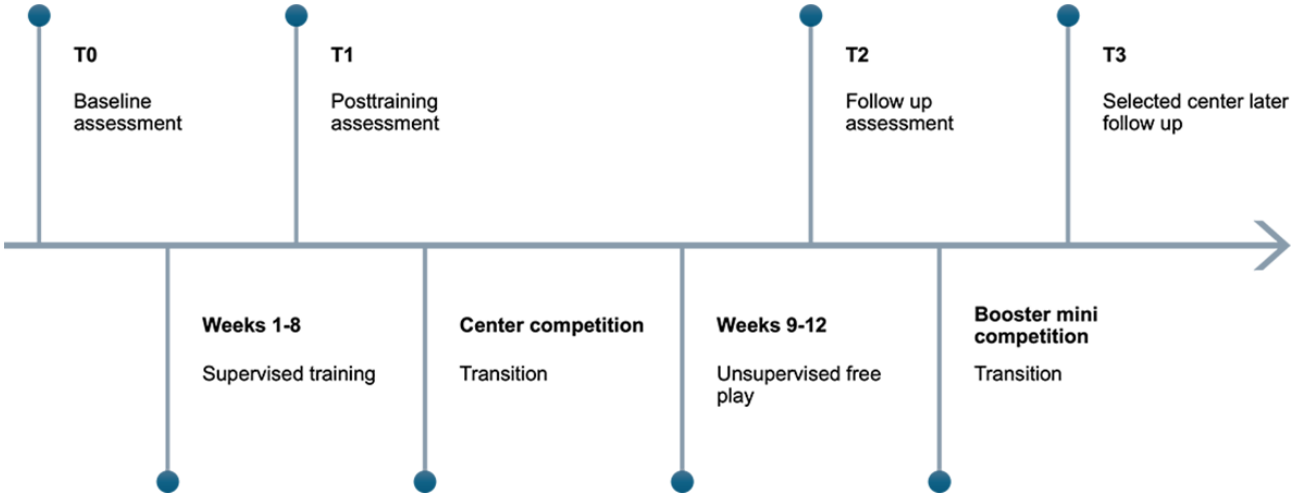
Methods

Study Design

We conducted a multi-wave observational evaluation of the I-SING, a community-delivered intergenerational exergaming program for older adults in Singapore. Assessments were conducted before supervised training (T0), immediately after supervised training (T1), after a 4-week unsupervised free-play period (T2), and at a later follow-up for selected measures (T3). The design was pragmatic and did not include a control or comparison group. Therefore, this study could identify within-person changes and implementation-relevant associations, but it could not rule out secular trends, regression to the mean, center-level contextual effects, or other unmeasured confounders. The study design and assessment timeline are shown in Figure 1.

XSL•FO
RenderX

Figure 1. Study design and assessment timeline. T3 was conducted 1 week after the booster mini competition in 6 selected centers.



Participants and Recruitment

Older adults were recruited from October 2025 to January 2026 through 14 participating senior activity centers and active aging centers using center-based outreach and posters. Eligibility followed the approved study protocol: participants were required to be 60 years of age or older, community ambulant, able to perform light physical activity, and free of a physical impairment that would preclude participation. Serious mental disorders were an exclusion criterion. Of 280 participants with baseline data, 2 (0.7%) withdrew at the program start, leaving 278 (99.3%) who entered the supervised phase. At the end of the 8-week supervised intervention, 243 participants completed the program; 242 (99.6%) provided analyzable posttraining (T1) survey data. After the 4-week unsupervised free-play phase, 230 participants contributed T2 data. Later follow-up (T3) was conducted only in 6 selected centers, where 120 participants were eligible for follow-up and 88 (73.3%) completed the T3 survey. Attrition analyses comparing baseline characteristics between completers and noncompleters at T1, T2, and T3 can be found in Tables S1 to S3 in Multimedia Appendix 1. At T1, completers were more likely than noncompleters to have previously played exergames and participated in I-SING. No clear baseline differences were observed between T2 completers and noncompleters. Among participants selected for T3 follow-up, completers differed from noncompleters in community center attendance frequency and were more likely to have previously played exergames and participated in I-SING. Other baseline demographic characteristics were generally comparable between completers and noncompleters across the assessment points.

Program Procedures

The supervised phase consisted of 8 weekly sessions delivered over 8 weeks. After T1, a center competition was held, followed by a 4-week unsupervised free-play phase intended to observe naturalistic engagement under reduced staff prompting. A booster mini competition was introduced only after T2 data collection to reduce anticipation effects. Surveys required approximately 45 to 60 minutes at each wave.

Outcome Measures

The focal outcomes for the present analyses were positive mental health, intrinsic motivation, and habit strength. Physical activity and COM-B–aligned implementation drivers were examined as exploratory, implementation-oriented outcomes. Detailed descriptions of additional measures collected for broader program evaluation, including healthy lifestyle, self-reported health, and health ownership, can be found in Multimedia Appendix 1.

Ethical Considerations

This study was approved by the Nanyang Technological University Institutional Review Board (IRB-2023-321), including protocol amendments covering revised training formats and the unsupervised free-play period. All participants provided written informed consent before enrollment. Survey data were collected and analyzed in deidentified form to protect participant privacy and confidentiality. Participants received shopping vouchers as compensation for participation.

Demographic Characteristics

A 9-item questionnaire assessed employment status, living arrangement, educational level, marital status, household size, perceived adequacy of money for daily expenses, frequency of community center visits, prior exergaming experience, and prior participation in I-SING. These variables were used to characterize the sample and, where relevant, as covariates.

Positive Mental Health

Positive mental health was measured using the 14-item Mental Health Continuum–Short Form (MHC-SF), which assesses emotional, social, and psychological well-being [17]. Participants rated how often they had experienced each feeling during the previous month on a 6-point scale ranging from 0 (never) to 5 (every day). The 14 harmonized item scores were summed to produce total scores ranging from 0 to 70, with higher scores indicating better positive mental health.

Motivation for Exergaming

Motivation for exergaming was assessed using an adapted version of the Motives for Physical Activity Measure–Revised (MPAM-R) [18]. Adaptation consisted of wording the original physical activity items to refer specifically to exergaming participation; no substantive changes were made to the response scale or intended construct. Items were rated on a 7-point scale ranging from 1 (not at all) to 7 (describes exactly). The present

XSL•FO
RenderX

analyses focused on the intrinsic motivation score, with higher values indicating stronger intrinsically oriented motivation to engage in exergaming.

Habit Strength

Habit strength was assessed using the 12-item Self-Report Habit Index (SRHI) [19]. Items were rated on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). Item responses were averaged, with higher scores indicating stronger habit strength for the target behavior.

Physical Activity

Physical activity was assessed using the Physical Activity Scale for the Elderly (PASE), a measure of leisure-time, household, and work-related activity in older adults [20]. Scores were calculated using the recommended weighted scoring procedure, with higher scores indicating greater physical activity. Because the PASE is a weighted composite of distinct physical activity domains rather than a unidimensional scale, internal consistency was not calculated.

Program Experience and Feedback

At T3, participants from selected centers completed program feedback items on game preferences, enjoyment, recommendations, and suggestions for improvement. These responses were summarized descriptively to characterize program acceptability and practical implications for community delivery.

Measurement Reliability

At T1, internal consistency was good for the 14-item MHC-SF (Cronbach α=0.89; n=242), the 14-item adapted MPAM-R intrinsic motivation score (Cronbach α=0.87; complete n=218), and the 12-item SRHI (Cronbach α=0.90; complete n=201). Designated missing value codes and values outside the prespecified response ranges were treated as missing before the coefficients were calculated. Internal consistency was not estimated for the PASE because its weighted total combines conceptually distinct activity domains rather than parallel items measuring a unidimensional construct.

Operationalization of COM-B–Aligned Drivers

To explore implementation-relevant correlates of physical activity maintenance, we constructed 3 theory-informed composite indexes aligned with the COM-B domains of capability, opportunity, and motivation. These composites were pragmatic summary indicators for exploratory analysis and should not be interpreted as validated latent COM-B scales. Capability comprised health literacy, self-efficacy, information seeking, and health citizenship or control; opportunity comprised supportive accountability, perceived social support, responsibility and social contract, community actor, and collective efficacy; and motivation comprised healthy lifestyle, positive mental health, intrinsic and extrinsic motivation, habit strength, and self-reported health. Because the motivation composite included constructs that overlapped with the primary outcomes, we did not interpret this domain as an independent predictor of those same outcomes. Accordingly, the COM-B analyses were limited to physical activity trajectories and interpreted cautiously to minimize the risk of conceptual circularity.

Because this study focused on maintenance after supervised training, posttraining values were used preferentially to represent participants' readiness for the reduced-intensity phase; baseline values were used only when posttraining values were unavailable. Component measures were standardized before aggregation, and higher composite scores indicated relatively greater capability, opportunity, or motivation. Full details on composite construction, weighting, equations, missing data handling, and a recommended sensitivity analysis excluding overlapping motivation indicators are provided in Multimedia Appendix 1.

Statistical Analysis

Baseline characteristics were summarized using means and SDs for continuous variables and frequencies and percentages for categorical variables. The primary outcomes were positive mental health, intrinsic motivation, and habit strength. Immediate within-person change from T0 to T1 was examined using paired *t* tests based on complete pairs and reported as Cohen *dz* with 95% CIs. Maintenance across waves was examined using linear mixed-effects models with time as a categorical fixed effect and participant as a random intercept, adjusted for age, sex, and cohort and year. These covariates were selected a priori because age and sex are commonly associated with physical activity and psychosocial outcomes in older adults and cohort and year captured implementation differences across program cohorts. Analyses of PASE trajectories and COM-B driver interactions were prespecified as exploratory implementation-oriented analyses. All analyses used available data and were conducted in R (version 4.5.1; R Foundation for Statistical Computing). No imputation was performed; therefore, findings assume that missingness was adequately handled by the available-data mixed-model framework conditional on observed covariates and should be interpreted alongside attrition patterns.

Results

Sample and Baseline Characteristics

At baseline, the sample had a mean age of 72.54 (SD 6.15) years, 77.1% (216/280) of the participants were female, and 92.1% (258/280) identified as Chinese. Most participants were not working (244/277, 88.1%), and many reported frequent visits to community centers. Prior exergaming exposure was present but not dominant: 41.4% (115/278) of the participants had played exergames, and 21.9% (61/278) had participated in I-SING previously. Baseline sample characteristics are shown in Table 1. Participant flow is reported in accordance with STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidance and shown in Figure 2.

XSL•FO
RenderX

Table . Participant demographics (baseline; N=280).

Characteristics	Values
Age (y), mean (SD)	72.54 (6.15)
Age group (y), n (%)	
60-64	22 (7.9)
65-69	72 (25.7)
70-74	84 (30)
75-79	64 (22.9)
≥80	38 (13.6)
Sex, n (%)	
Male	64 (22.9)
Female	216 (77.1)
Ethnicity, n (%)	
Chinese	258 (92.1)
Malay	10 (3.6)
Indian	8 (2.9)
Others	4 (1.4)
Spoken language, n (%)	
English	105 (37.5)
Mandarin	171 (61.1)
Other	4 (1.4)
Occupation status (n=277), n (%)	
Working	33 (11.9)
Not working	244 (88.1)
Living situation (n=278), n (%)	
Alone	84 (30.2)
With spouse	78 (28.1)
With spouse and children	56 (20.1)
With children	36 (12.9)
With friends or relatives	14 (5)
Tenant	1 (0.4)
Others	9 (3.2)
Highest level of education (n=278), n (%)	
No formal school education	38 (13.7)
Primary school	86 (30.9)
Secondary school	101 (36.3)
Diploma or polytechnic or preuniversity education	41 (14.7)
Degree, professional certifications, or higher	12 (4.3)
Marital status (n=278), n (%)	
Single, never married	33 (11.9)
Married	155 (55.8)
Cohabiting	1 (0.4)
Divorced	16 (5.8)
Separated	2 (0.7)

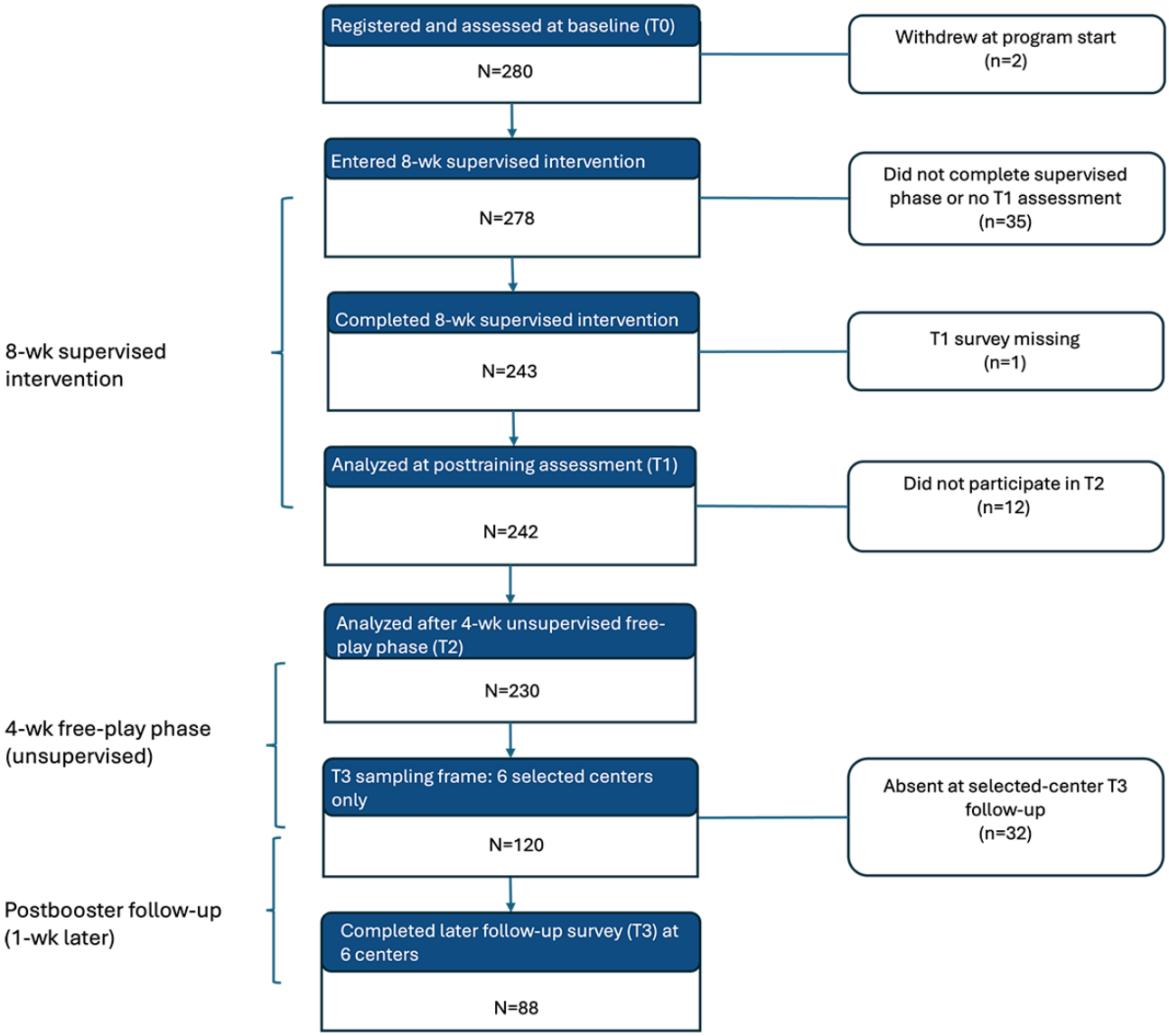
XSL•FO
RenderX

Characteristics	Values
Widowed	71 (25.5)
Household members (n=278), n (%)	
Alone	71 (25.5)
2 people	93 (33.5)
3 people	46 (16.5)
4 people	29 (10.4)
5 people	14 (5)
More than 5 people	25 (9)
Perceived adequacy of daily expenses (n=278), n (%)	
Very insufficient	2 (0.7)
Not enough	15 (5.4)
Just enough	183 (65.8)
More than enough	71 (25.5)
Very abundant	7 (2.5)
Community center visits per week (n=278), n (%)	
1 time	22 (7.9)
2 times	34 (12.2)
3 times	50 (18)
4 times	59 (21.2)
5 times	72 (25.9)
More than 5 times	32 (11.5)
Others	9 (3.2)
Prior exergame experience (n=278), n (%)	
Played exergames before	115 (41.4)
Never played exergames before	163 (58.6)
Prior exergame experience (mo; n=90), mean (SD)	7.13 (13.65)
Prior participation in I-SING[a] participants (n=278), n (%)	61 (21.9)
Previous I-SING program participated in (n=61), n (%)	
I-SING 2024—4 wk	2 (3.3)
I-SING 2024—8 wk	49 (80.3)
I-SING 2023—6 wk	0 (0)
I-SING 2023—12 wk	0 (0)
Others	10 (16.4)

[a]I-SING: International Singapore Intergenerational National Games.

XSL•FO
RenderX

Figure 2. STROBE (Strengthening the Reporting of Observational Studies in Epidemiology)–style participant flow diagram. T0: baseline; T1: after an 8-week supervised training phase; T2: after a 4-week unsupervised free-play phase; T3: at a later follow-up for selected measures in 6 centers.



Immediate Program Effects

From T0 to T1, participants showed within-person increases in the 3 prespecified proximal outcomes. Mean positive mental health increased from 55.19 (SD 9.79) at T0 to 57.59 (SD 8.70) at T1 among 240 complete pairs (Cohen *dz*=0.279, 95% CI 0.149-0.407; *P*<.001). Mean intrinsic motivation increased from 5.94 (SD 1.00) to 6.29 (SD 0.81) among 166 paired observations (Cohen *dz*=0.398, 95% CI 0.239-0.555; *P*<.001). Mean habit strength increased from 3.34 (SD 1.11) to 3.81 (SD 0.84) among 143 paired observations (Cohen *dz*=0.423, 95% CI 0.252-0.594; *P*<.001). These results support hypothesis 1 but should be interpreted as changes observed after program participation rather than controlled intervention effects. Pair-specific sample sizes for the primary outcomes are reported in Table S4 in Multimedia Appendix 1.

Exploratory paired analyses of additional psychosocial variables are summarized in Table S5 in Multimedia Appendix 1. Briefly, small significant increases were observed for extrinsic motivation (*P*=.009) and social responsibility (*P*=.009), whereas collective efficacy decreased slightly (*P*=.03); no reliable immediate change was detected for self-efficacy (*P*=.24), supportive accountability (*P*=.31), perceived social support (*P*=.26), information seeking (*P*=.63), or health literacy (*P*=.73).

Maintenance Under Reduced Intensity

Hypothesis 2 was partially supported. Habit strength was higher immediately after supervised training (β=0.291, 95% CI .047-.535; *P*=.02) and after the free-play phase at T2 (β=0.362, 95% CI .113-.611; *P*=.005) than at baseline. Intrinsic motivation was higher at T2 (β=0.397, 95% CI .156-.638; *P*=.001) and at the selected-center T3 follow-up (β=0.305, 95% CI .009-.601; *P*=.04) than at baseline. Contrary to the expectation that positive mental health gains would be less stable, positive mental health was higher both immediately after supervised training (β=4.140, 95% CI 2.169-6.111; *P*<.001) and at T2 (β=7.390, 95% CI 5.370-9.410; *P*<.001) than at baseline. The paired T0-to-T2 change in positive mental health was moderate (n=228; Cohen *dz*=0.520, 95% CI 0.382-0.658; *P*<.001).

These findings indicate that positive mental health, intrinsic motivation, and habit strength were all higher at T2 than at baseline among respondents with available follow-up data.

XSL•FO
RenderX

Nevertheless, because this study lacked a control group and relied on available observations, the changes cannot be attributed causally to I-SING. Secular changes, contextual influences, and selective retention remain possible alternative explanations (Table 2).

Table . Primary and maintenance outcomes across study waves[a].

Outcome and time point	Participants, n	Score, mean (SD)	Contrast vs baseline, β (95% CI)	*P* value	Paired effect size, Cohen *dz*
Habit strength (range 1-5)					
Before	278	3.319 (1.103)	Reference	—[b]	—
After	242	3.729 (0.836)	0.291 (0.047 to 0.535)	.02	0.423
T2[c]	230	3.729 (0.817)	0.362 (.113 to 0.611)	.005	0.382
Intrinsic motivation (range 1-7)					
Before	278	5.945 (0.982)	Reference	—	—
After	242	6.305 (0.764)	0.184 (–0.054 to 0.421)	.13	0.398
T2	230	6.403 (0.683)	0.397 (0.156 to 0.638)	.001	0.512
T3[d]	88	6.349 (0.988)	0.305 (0.009 to 0.601)	.04	0.291
Positive mental health (MHC-SF[e] total; 0-70)					
Before	278	55.076 (9.726)	Reference	—	—
After	242	57.674 (8.712)	4.140 (2.169 to 6.111)	<.001	0.279
T2	230	60.622 (8.131)	7.390 (5.370 to 9.410)	<.001	0.520

[a]Mixed-effects models included time as a categorical fixed effect and participant as a random intercept and were adjusted for age, sex, and program cohort and year. Positive β values indicate scores higher than baseline. The n values represent the number of available observations at each wave. Paired effect sizes were calculated using complete pairs; pair-specific sample sizes are reported in Table S4 in Multimedia Appendix 1. For the Mental Health Continuum–Short Form, the T0-to-T1 effect size was based on 240 complete pairs, and the T0-to-T2 effect size was based on 228 complete pairs.

[b]Not applicable because baseline was the reference category; therefore, no P value or paired effect size was calculated.

[c]T2: after a 4-week unsupervised free-play phase.

[d]T3: at a later follow-up for selected measures in 6 centers.

[e]MHC-SF: Mental Health Continuum–Short Form.

Taken together, the adjusted models indicated that positive mental health, intrinsic motivation, and habit strength were higher at T2 than at baseline among respondents with available data. Intrinsic motivation was also higher at the selected-center T3 follow-up than at baseline. These results describe observed longitudinal trajectories and should not be interpreted as evidence that the reduced-intensity phase caused the outcomes to be maintained.

COM-B Drivers and Physical Activity Maintenance Over Time

Hypothesis 3 examined whether higher COM-B–aligned capability, opportunity, and motivation were associated with differences in physical activity trajectories across baseline, immediately after training, and the T2 follow-up (Table 3). In mixed-effects models, only opportunity significantly moderated immediate change in physical activity after training (time × opportunity interaction β=26.464, 95% CI 3.191-49.737; *P*=.03). No significant time interactions were observed for capability or motivation, and no COM-B domain significantly moderated T2 physical activity change. In separate end point models, baseline PASE strongly predicted both posttraining and follow-up PASE scores, whereas opportunity showed only borderline evidence of association with posttraining PASE scores (detailed end point regression results can be found in Table S6 in Multimedia Appendix 1). Given the number of exploratory interactions tested, this result should be interpreted as a preliminary hypothesis-generating signal rather than definitive evidence that opportunity is more important than other COM-B domains.

XSL•FO
RenderX

Table . Associations between capability, opportunity, and motivation–behavior (COM-B) drivers and physical activity maintenance over time. Positive estimates indicate stronger increases in Physical Activity Scale for the Elderly among participants with higher standardized driver scores.

COM-B driver and interaction term	Regression coefficient, β (95% CI)	*P* value	Interpretation
Capability			
T1[a] × capability	−9.332 (−32.331 to 13.666)	.42	No evidence of interaction
T2[b] × capability	9.252 (–14.230 to 32.735)	.44	No evidence of interaction
Opportunity			
T1 × opportunity	26.464 (3.191 to 49.737)	.03	Higher opportunity associated with greater immediate increase in PASE[c]
T2 × opportunity	0.164 (–23.606 to 23.934)	.99	No evidence of interaction
Motivation			
T1 × motivation	10.015 (–10.159 to 30.190)	.33	No evidence of interaction
T2 × motivation	−6.623 (–27.145 to 13.899)	.52	No evidence of interaction

[a]T1: immediately after an 8-week supervised training phase.

[b]T2: after a 4-week unsupervised free-play phase.

[c]PASE: Positive estimates indicate greater increases in PASE among participants with higher standardized driver scores.

The COM-B domains were operationalized as follows: capability comprised health literacy, self-efficacy, information seeking, and control; opportunity comprised supportive accountability, perceived social support, responsibility and social contract, community actor, and collective efficacy; and motivation comprised healthy lifestyle, positive mental health, intrinsic and extrinsic motivation, habit, and self-reported health. The standardized COM-B driver scores were centered around 0, with means close to 0 and SDs ranging from 0.605 to 0.638, indicating comparable scaling across the 3 domains. These composites were exploratory and theory informed, and their results should not be interpreted as validated COM-B measurement scores.

Mixed-effects models assessed whether capability, opportunity, and motivation moderated physical activity trajectories from baseline to the posttraining and follow-up periods. Mean PASE scores increased modestly from baseline (mean 143.71, SD 66.42) to the posttraining time point (mean 147.81, SD 66.91) and declined at T2 (mean 137.02, SD 58.53). Among the 3 COM-B domains, only opportunity significantly moderated immediate posttraining change in physical activity. In separate end point models, baseline physical activity strongly predicted both posttraining and follow-up PASE scores, whereas opportunity showed only borderline evidence of association with posttraining PASE scores.

Game Feedback From Selected Centers at T3

Because T3 feedback was available only for selected centers, these findings are presented descriptively. Overall, the exergaming component appeared to be well accepted. Apple Frenzy emerged as the most consistently preferred game and was commonly described as easy to play, enjoyable, relaxing, colorful, and easy to score. Fishing Frenzy was also frequently identified as a favorite because it was perceived as fun, active, and similar to exercise. Punch Punch and Whack-a-Mole were also popular, particularly among participants who valued excitement, reflex training, or challenges. In contrast, more cognitively demanding or faster-paced games such as Chinatown Rush and Red Light Rat Race appeared more polarizing. Participants also most often described Cat Stack, Don’t Forget to Wipe, and Ice Breaker negatively, typically because they were perceived as too difficult, tiring, frustrating, or hard to control. Detailed game-level results are provided in Table S7 in Multimedia Appendix 1.

Open-ended comments suggested that continued engagement during free play depended not only on game enjoyment but also on center-level opportunity conditions, including reminders, scheduling, staff support, and available space. This descriptive pattern was consistent with the quantitative finding that opportunity was the only COM-B domain associated with stronger immediate posttraining physical activity improvement.

Notably, participants often described Fishing Frenzy as enjoyable because it involved substantial movement and “helps me exercise.” This complements the supplementary functional movement mapping, within which Fishing Frenzy is characterized as a repeated sit-to-stand task with likely lower-limb strengthening and balance-related benefits. Thus, participant feedback supports the view that games with clear, functional, exercise-like movement demands may be especially acceptable in community exergaming for older adults.

Discussion

Principal Findings

This multi-wave observational evaluation examined whether changes in psychosocial and behavioral outcomes persisted after a community exergaming program transitioned from supervised training to reduced-intensity free play. Positive mental health, intrinsic motivation, and habit strength increased immediately after supervised training and were higher at T2 than at baseline among respondents with available data; intrinsic motivation was also higher at the selected-center T3 follow-up than at baseline. Opportunity was the only COM-B domain associated

XSL•FO
RenderX

with greater immediate posttraining change in physical activity, and no COM-B domain moderated T2 change. These findings represent observed longitudinal trajectories and exploratory associations rather than causal program effects.

The immediate pattern is broadly consistent with earlier exergaming research. Reviews have repeatedly found the strongest and most consistent benefits in movement-related outcomes such as balance and mobility, whereas psychosocial effects are more heterogeneous across studies and measures [2,3,5,6]. The present study contributes something different: it evaluated a large, community-delivered program operating through routine service infrastructure in Singapore rather than a tightly controlled efficacy trial. This makes the findings especially relevant for implementation decisions in community aging services.

The maintenance-related findings extended beyond motivation and habit [13,14,18]. Positive mental health was also higher at T2 than at baseline, contrary to the expectation that mental health gains would be less stable after structured social contact decreased. The sustained differences in positive mental health, intrinsic motivation, and habit strength suggest that favorable outcomes were observed beyond the supervised phase among participants who remained in follow-up. However, attrition, the relatively short follow-up period, and the absence of direct free-play use data limit conclusions about maintenance in the full cohort and about whether these psychosocial and habit-related scores translated into continued exergaming participation.

Positive mental health was higher immediately after supervised training and at T2 than at baseline. This pattern suggests that the observed improvement was not confined to the supervised phase. However, because this study lacked a control group and the T2 analysis relied on available follow-up data, the higher score cannot be attributed causally to I-SING. Secular changes, center-level context, survey administration conditions, and selective retention may also have contributed to the observed trajectory.

The COM-B results provide a cautious implementation signal rather than a definitive mechanism test. Of the 3 COM-B domains, only opportunity was associated with stronger immediate posttraining physical activity change. In COM-B terms, opportunity refers to the environmental and social conditions that enable behavior, including access, social support, cues, and a facilitating context [15]. This interpretation is plausible in community exergaming because reminders, peer presence, accessible scheduling, and light-touch staff support may help older adults translate interest into actual participation. However, the opportunity result should be considered hypothesis generating because multiple interactions were tested, the composites were not validated COM-B scales, and center-level context may have contributed to the observed association.

The descriptive game feedback supports that same conclusion. Participants preferred games that combined intuitive movement, moderate pace, clear goals, and a rewarding sense of exercise. Fishing Frenzy and Apple Frenzy appeared to work well because they were both accessible and meaningfully active, with Fishing Frenzy often described as helping participants exercise. In contrast, games perceived as confusing, overly fast, poorly tracked, or physically unrewarding were less acceptable. Importantly, free-play barriers were often logistical rather than motivational, reinforcing the inference that access, reminders, scheduling, and space matter as much as game enjoyment once supervision recedes [13,15].

Several limitations should be acknowledged. First, the observational pretest-posttest design lacked a control group, limiting causal inference and preventing attribution of the observed changes to I-SING alone. Second, analyses used available observations, and attrition reduced the sample for paired comparisons and later follow-up. Selective retention of participants who were more engaged, healthier, or more familiar with exergaming may have contributed to the apparent maintenance-related trajectories. Third, T3 was restricted to selected centers, limiting the generalizability of those findings. Fourth, no automated logs or equivalent individual-level records of free-play exposure were available; consequently, higher motivation and habit scores cannot be assumed to represent continued exergaming participation. Fifth, the PASE was self-reported. Sixth, the COM-B composites were theory-informed pragmatic indexes rather than validated measures, and the motivation composite contained constructs that overlapped conceptually with the primary outcomes. Finally, no dedicated validated social outcome battery was administered, so conclusions regarding loneliness, social connectedness, or social participation cannot be drawn. Future studies should include controlled comparisons, direct measures of program exposure and adherence, validated social health outcomes, and sensitivity analyses addressing missing data and center-level clustering.

Conclusions

In this multi-wave observational evaluation of community exergaming for older adults, positive mental health, intrinsic motivation, and habit strength were higher at T2 than at baseline among respondents with available follow-up data, and intrinsic motivation was also higher at the selected-center T3 follow-up. These findings suggest maintenance-related signals across both psychosocial and behavioral outcomes, but the uncontrolled design, attrition, selected-center follow-up, and absence of direct free-play exposure data preclude causal conclusions. Opportunity-related supports such as access, reminders, scheduling, staff facilitation, and peer participation may be useful when programs transition from supervised sessions to reduced-intensity free play, although this interpretation remains exploratory. Future controlled studies with consistent cross-wave scoring, validated social outcomes, and direct tracking of program exposure are needed to clarify maintenance pathways.

XSL•FO
RenderX

Acknowledgments

The authors sincerely thank all older adult participants and the partnering active aging centers involved in the International Singapore Intergenerational National Games program. They are particularly grateful to the senior activity center and active aging center management teams and staff for their generous support in organizing sessions, coordinating participants, and facilitating implementation across sites. The authors also thank the student interns from Nanyang Technological University for their assistance with program delivery, logistics, and administrative tasks. Finally, they acknowledge the support of the Ageing Research Institute for Society and Education, Nanyang Technological University, for this study. The authors used ChatGPT (OpenAI) to assist with language editing and improving the clarity of the manuscript. It was not used to generate study data or perform the statistical analyses. All AI-assisted content was critically reviewed and verified by the authors. The authors take full responsibility for the accuracy, integrity, and final content of the manuscript.

Funding

This work was supported by the Ageing Research Institute for Society and Education, Nanyang Technological University. The funder had no role in the study design; data collection, analysis, and interpretation; manuscript preparation; or decision to submit the manuscript for publication.

Data Availability

The datasets generated and analyzed during this study are not publicly available because they contain potentially identifiable participant information and are subject to institutional data governance requirements of the Ageing Research Institute for Society and Education, Nanyang Technological University. Deidentified data may be made available from the corresponding author on reasonable request, subject to institutional review and permission from the data custodian.

Authors' Contributions

YL conceived the study, led the analysis, and drafted the manuscript. BXC, HYK, and QNTN contributed to program implementation and data acquisition. JJ, KMWYT, and SQAL contributed to data analysis and manuscript revision. YLT contributed to study design, interpretation of the findings, and critical revision of the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods and results.
[DOCX File, 68 KB - aging_v9i1e100007_app1.docx]

References

1. Agmon M, Perry CK, Phelan E, Demiris G, Nguyen HQ. A pilot study of Wii Fit exergames to improve balance in older adults. J Geriatr Phys Ther 2011;34(4):161-167. [doi: 10.1519/JPT.0b013e3182191d98] [Medline: 22124415]
2. Cacciata M, Stromberg A, Lee JA, et al. Effect of exergaming on health-related quality of life in older adults: a systematic review. Int J Nurs Stud 2019 May;93:30-40. [doi: 10.1016/j.ijnurstu.2019.01.010] [Medline: 30861452]
3. Fang Q, Ghanouni P, Anderson SE, et al. Effects of exergaming on balance of healthy older adults: a systematic review and meta-analysis of randomized controlled trials. Games Health J 2020 Feb;9(1):11-23. [doi: 10.1089/g4h.2019.0016] [Medline: 31800322]
4. Kwok BC, Pua YH. Effects of WiiActive exercises on fear of falling and functional outcomes in community-dwelling older adults: a randomised control trial. Age Ageing 2016 Sep;45(5):621-627. [doi: 10.1093/ageing/afw108] [Medline: 27496921]
5. Pacheco TB, de Medeiros CS, de Oliveira VH, Vieira ER, de Cavalcanti FA. Effectiveness of exergames for improving mobility and balance in older adults: a systematic review and meta-analysis. Syst Rev 2020 Jul 18;9(1):163. [doi: 10.1186/s13643-020-01421-7] [Medline: 32682439]
6. Chen X, Wu L, Feng H, et al. Comparison of exergames versus conventional exercises on the health benefits of older adults: systematic review with meta-analysis of randomized controlled trials. JMIR Serious Games 2023 Jun 22;11:e42374. [doi: 10.2196/42374] [Medline: 37347534]
7. Drazich BF, LaFave S, Crane BM, et al. Exergames and depressive symptoms in older adults: a systematic review. Games Health J 2020 Oct;9(5):339-345. [doi: 10.1089/g4h.2019.0165] [Medline: 32551982]
8. Stojan R, Voelcker-Rehage C. A systematic review on the cognitive benefits and neurophysiological correlates of exergaming in healthy older adults. J Clin Med 2019 May 23;8(5):734. [doi: 10.3390/jcm8050734] [Medline: 31126052]

XSL•FO
RenderX

9. Begde A, Alqurafi A, Pain MT, Blenkinsop G, Wilcockson TD, Hogervorst E. The effectiveness of home-based exergames training on cognition and balance in older adults: a comparative quasi-randomized study of two exergame interventions. Innov Aging 2023;7(8):igad102. [doi: 10.1093/geroni/igad102] [Medline: 37941829]
10. Li J, Erdt M, Chen L, Cao Y, Lee SQ, Theng YL. The social effects of exergames on older adults: systematic review and metric analysis. J Med Internet Res 2018 Jun 28;20(6):e10486. [doi: 10.2196/10486] [Medline: 29954727]
11. Li J, Theng YL, Foo S. Play mode effect of exergames on subthreshold depression older adults: a randomized pilot trial. Front Psychol 2020 Oct 26;11:552416. [doi: 10.3389/fpsyg.2020.552416] [Medline: 33192801]
12. Rytterström P, Strömberg A, Jaarsma T, Klompstra L. Exergaming to increase physical activity in older adults: feasibility and practical implications. Curr Heart Fail Rep 2024 Aug;21(4):439-459. [doi: 10.1007/s11897-024-00675-9] [Medline: 39023808]
13. Meekes W, Stanmore EK. Motivational determinants of exergame participation for older people in assisted living facilities: mixed-methods study. J Med Internet Res 2017 Jul 6;19(7):e238. [doi: 10.2196/jmir.6841] [Medline: 28684385]
14. Ning H, Jiang D, Du Y, et al. Older adults' experiences of implementing exergaming programs: a systematic review and qualitative meta-synthesis. Age Ageing 2022 Dec 5;51(12):afac251. [doi: 10.1093/ageing/afac251] [Medline: 36571772]
15. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
16. Ahmadi A, Noetel M, Parker P, et al. A classification system for teachers' motivational behaviors recommended in self-determination theory interventions. J Educ Psychol 2023;115(8):1158-1176. [doi: 10.1037/edu0000783]
17. Lamers SM, Westerhof GJ, Bohlmeijer ET, ten Klooster PM, Keyes CL. Evaluating the psychometric properties of the Mental Health Continuum-Short Form (MHC-SF). J Clin Psychol 2011 Jan;67(1):99-110. [doi: 10.1002/jclp.20741] [Medline: 20973032]
18. Ryan RM, Frederick CM, Lepes D, Rubio N, Sheldon KM. Intrinsic motivation and exercise adherence. Int J Sport Psychol 1997;28(4):335-354 [FREE Full text]
19. Verplanken B, Orbell S. Reflections on past behavior: a self-report index of habit strength. J Appl Soc Psychol 2003 Jun;33(6):1313-1330. [doi: 10.1111/j.1559-1816.2003.tb01951.x]
20. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. J Clin Epidemiol 1993 Feb;46(2):153-162. [doi: 10.1016/0895-4356(93)90053-4] [Medline: 8437031]

Abbreviations

COM-B: capability, opportunity, and motivation–behavior
I-SING: International Singapore Intergenerational National Games
MHC-SF: Mental Health Continuum–Short Form
MPAM-R: Motives for Physical Activity Measure–Revised
PASE: Physical Activity Scale for the Elderly
SRHI: Self-Report Habit Index
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by P Jansons; submitted 01.May.2026; peer-reviewed by H Tilga, MC Espada; revised version received 28.Jul.2026; accepted 31.Jul.2026; published 09.Sep.2026.

Please cite as:
Liu Y, Chia BX, Koh HY, Nguyen QNT, Jin J, Tan KMWY, Long SQA, Theng YL
Community Exergaming and Maintenance-Related Outcomes in Older Adults: Multi-Wave Observational Evaluation
JMIR Aging 2026;9:e100007
URL: https://aging.jmir.org/2026/1/e100007
doi:10.2196/100007

© Yuezhong Liu, Bing Xun Chia, Herng Yeow Koh, Quoc Nam Tran Nguyen, Jiahui Jin, Kimberly Miracle Wei Yan Tan, Samuel Qing An Long, Yin-Leng Theng. Originally published in JMIR Aging (https://aging.jmir.org), 9.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Domain-Specific Associations of Physical Activity With Phase Angle and Extracellular Water–to–Total Body Water Ratio in Middle-Aged or Older Adults: Cross-Sectional Study

Seong-Uk Baek[1], MD; Jin-Ha Yoon[2,3,4], PhD

[1]Graduate School, Yonsei University College of Medicine, Seoul, Seoul, Republic of Korea

[2]Department of Preventive Medicine, Yonsei University College of Medicine, 50-1, Yonsei-ro, Seodaemungu, Seoul, Republic of Korea

[3]The Institute for Occupational Health, Yonsei University College of Medicine, Seoul, Republic of Korea

[4]Institute for Innovation in Digital Healthcare, Yonsei University Health System, Seoul, Republic of Korea

Corresponding Author:
Jin-Ha Yoon, PhD
Department of Preventive Medicine, Yonsei University College of Medicine, 50-1, Yonsei-ro, Seodaemungu, Seoul, Republic of Korea

Related Article:

This is a corrected version. See correction statement: https://aging.jmir.org/2026/1/e114943

Abstract

Background: The phase angle (PhA) and extracellular water (ECW)–to–total body water (TBW) ratio serve as indicators of cell membrane integrity and body fluid balance in older adults.

Objective: This study examined the domain-specific associations of physical activity (PA) with PhA and the ECW-to-TBW ratio.

Methods: A nationally representative sample of 5211 adults aged 50 years and older was analyzed. Weekly amounts of moderate-to-vigorous PA (MVPA) in leisure-time, occupational, and transport-related domains were assessed using the Global Physical Activity Questionnaire. A bioelectrical impedance analyzer was used to measure PhA (°) and the ECW-to-TBW ratio (%). Linear regression analyses were conducted.

Results: For leisure-time PA, compared with 0 min/week, participating in 150 to 299 min/week and ≥300 min/week of MVPA was associated with 0.08 (95% CI 0.02-0.14) and 0.13 (95% CI 0.06-0.20) degrees higher PhA, respectively. Additionally, compared with 0 min/week, engaging in 150 to 299 min/week and ≥300 min/week of MVPA was associated with –0.11 (95% CI –0.19 to –0.03) and –0.15 (95% CI –0.22 to –0.07) percentage points lower ECW-to-TBW ratio, respectively. For transport-related PA, compared with 0 min/week, engaging in ≥300 min/week of MVPA was negatively associated with the ECW-to-TBW ratio (β=–0.07, 95% CI –0.14 to –0.01). Occupational PA showed no clear associations with PhA or the ECW-to-TBW ratio.

Conclusions: While the observed associations were modest, the associations of PA with both PhA and the ECW-to-TBW ratio differed across PA domains. Specifically, leisure-time PA was positively associated with PhA and inversely associated with the ECW-to-TBW ratio. These findings highlight the potential role of leisure-time PA in improving bioelectrical markers of cellular health and fluid distribution.

(JMIR Aging 2026;9:e90158) doi:10.2196/90158

KEYWORDS

bioimpedance analysis; body fluid; exercise; physical activity; epidemiology

Introduction

Physical activity (PA) is recognized as a major priority in promoting public health in older adults. While the World Health Organization (WHO) recommends that adults participate in ≥150 min/week of moderate-to-vigorous physical activity (MVPA), a concerning proportion of the global population exhibits low levels of PA [1]. For example, a recent study reported that insufficient PA was observed in 31.3% of the global population in 2022, representing a gradual increase over the past decades [1]. In South Korea, sedentary lifestyles remain a significant public health concern. Despite various

XSL•FO
RenderX

governmental policies and campaigns to promote PA, 54.5% of adults did not meet the WHO PA guidelines in 2020 [2], a prevalence that has continued to rise over recent decades [3]. Furthermore, older adults in South Korea were significantly less likely to engage in PA, and the decreasing trend was particularly steep among the population aged 50 years and older [3].

The health benefits of PA for older adults are well established. PA is associated with lower risks of overall and cardiovascular disease mortality in older adults [4]. The mechanisms underlying these health benefits can be attributed to multiple factors, including improvements in metabolic profiles, muscle strength and function, and psychological well-being [5]. However, accumulating evidence suggests that the effects of PA vary depending on the domain and context in which it is performed—a phenomenon referred to as the "PA paradox" [6,7]. For instance, studies have shown that reductions in the risks of all-cause mortality and cardiovascular disease are most pronounced for leisure-time physical activity (LTPA) compared with occupational physical activity (OPA) or transport-related physical activity (TRPA) [8-10]. Furthermore, OPA was linked to an elevated risk of mortality [10]. Despite the existing evidence regarding the "PA paradox," differential associations between PA domains and body composition or fluid status have rarely been investigated.

The phase angle (PhA) and extracellular water (ECW)–to–total body water (TBW) ratio, both determined using bioelectrical impedance analysis (BIA), serve as biomarkers of body fluid distribution and cellular function [11]. While PhA reflects cellular membrane integrity, the ECW-to-TBW ratio primarily reflects electrolyte balance and edema status [11,12]. Adverse health conditions, such as malnutrition, edema, systemic inflammation, and low skeletal muscle mass, are associated with lower PhA and higher ECW-to-TBW ratios [11,13,14]. Moreover, low PhA and high ECW-to-TBW ratios have been linked to increased mortality risk among older adults [15-17]. Therefore, PhA and the ECW-to-TBW ratio are regarded as important biomarkers of health status in recent literature.

PA substantially affects both body fluid distribution and body composition by reducing body fat and improving muscle quality in older adults [18-20]. These physiological adaptations can influence BIA-derived parameters, including PhA and the ECW-to-TBW ratio. Prior studies, for instance, reported a positive association between total weekly PA and PhA, along with a negative association with the ECW-to-TBW ratio [18-20]. However, whether the associations of PA with PhA or the ECW-to-TBW ratio differ across PA domains in older adults has been rarely examined. Specifically, no previous studies have investigated how LTPA, OPA, and TRPA are differentially associated with PhA and the ECW-to-TBW ratio. To address this research gap, this study investigated the associations of LTPA, OPA, and TRPA with PhA and the ECW-to-TBW ratio in a nationwide sample of older adults in South Korea. Specifically, we hypothesized that (1) LTPA and TRPA would be positively associated with PhA and inversely associated with the ECW-to-TBW ratio, and (2) OPA would be inversely associated with PhA and positively associated with the ECW-to-TBW ratio.

Methods

Study Population

This study used a cross-sectional design, with the study sample drawn from the Korea National Health and Nutrition Examination Survey (KNHANES). The Korea Disease Control and Prevention Agency (KDCA) administers the KNHANES annually, providing a nationally representative sample of the Korean population. A multistage clustered probability sampling methodology was used to include participants, ensuring representation of the entire Korean population [21]. In this study, datasets from the 2022 and 2023 cycles of the KNHANES were used, during which BIA was performed on participants. Note that, owing to the cross-sectional study design, the temporal relationship between the exposure variables (PA domains) and the outcomes (PhA and the ECW-to-TBW ratio) could not be established, which limits the ability to draw causal inferences from the study findings. A total of 6804 adults aged 50 years and older participated in the 2022 and 2023 survey years of the KNHANES. Individuals meeting any of the following exclusion criteria were excluded: (1) those with pacemakers or implantable cardioverter-defibrillators, (2) pregnant women, (3) individuals with prosthetic limbs, and (4) those who refused or were unable to undergo BIA. In total, 5742 (84.4%) adults participated in BIA. After further excluding participants with missing data, a final sample of 5211 (76.6%) adults was included in this analysis.

Ethical Considerations

All participants provided written informed consent. The protocols of the 2022 and 2023 KNHANES were approved by the institutional review board of the KDCA (2018-01-03-4C-A, 2022-11-16-R-A).

Physical Activity

The Korean version of the Global Physical Activity Questionnaire (K-GPAQ) was used to measure the weekly amounts of total PA, LTPA, OPA, and TRPA. The K-GPAQ has demonstrated good reliability and validity for assessing PA among the Korean population [22]. The questionnaire evaluates PA in 3 domains: LTPA, OPA, and TRPA. LTPA was assessed using the following questions: "Do you regularly participate in vigorous-intensity sports, exercise, or leisure activities that cause you to become significantly out of breath or make your heart beat very fast for at least 10 consecutive minutes?" (vigorous PA) and "Do you regularly participate in moderate-intensity sports, exercise, or leisure activities that cause you to become slightly out of breath or make your heart beat somewhat faster for at least 10 consecutive minutes?" (moderate PA). OPA was assessed using the following questions: "Does your job involve vigorous-intensity physical activity that causes you to become significantly out of breath or make your heart beat very fast for at least 10 consecutive minutes?" (vigorous PA) and "Does your job involve moderate-intensity physical activity that causes you to become slightly out of breath or make your heart beat somewhat faster for at least 10 consecutive minutes?" (moderate PA). TRPA was assessed using the following question: "Do you walk or cycle for transportation (eg, commuting to work, shopping,

XSL•FO
RenderX

running errands, attending religious services, or going to or from school or private academies) for at least 10 consecutive minutes?" (moderate PA). For each question, the duration of each session and the number of sessions per week were recorded. The weekly amount of MVPA for LTPA, OPA, and TRPA was then calculated using the following equation:

MVPA(per week)=2×Vigorous PA(per week)+Moderate PA(per week)

Total PA was calculated as the sum of MVPA across the 3 domains—LTPA, OPA, and TRPA. Finally, for each PA domain, PA levels were divided into 4 groups: 0, 1 - 149, 150 - 299, and ≥300 min/week, in accordance with the WHO PA guidelines [23].

PhA and ECW-to-TBW Ratio

PhA and the ECW-to-TBW ratio were measured using a multifrequency bioelectrical impedance analyzer (InBody 970; Biospace). The InBody 970 has demonstrated good validity for assessing body composition [24]. Moreover, measurements obtained from the InBody 970, including PhA, show high reliability and are comparable with results from other BIA devices [25]. Before the assessment, participants fasted for at least 8 hours. After removing all accessories and digital devices, they stood barefoot on the InBody 970 while holding the hand electrodes. All BIA procedures were supervised by trained health professionals from the KDCA. Whole-body PhA was recorded to the nearest 0.1° at 50 kHz, and ECW, intracellular water (ICW), and TBW were recorded to the nearest 0.1 L. Both PhA (°) and the ECW-to-TBW ratio (%) were analyzed as continuous variables.

Covariates

The confounders included the following sociodemographic characteristics: sex, age, educational attainment (middle school or below, high school, or college or above), income (lowest, low, high, or highest), marital status (married, unmarried, or others), smoking status (no or yes), alcohol use (no or yes), hypertension (no or yes), diabetes (no or yes), chronic kidney disease (no or yes), and BMI (kg/m^2). However, because of the lack of available data, several important confounders, including dietary intake, inflammatory markers, and medication use, which may influence body composition and fluid status, could not be included in the analyses. This limitation may have resulted in residual confounding.

Statistical Analysis

Sample characteristics were summarized according to total PA levels. Linear regression models were applied to evaluate the associations of PA levels with both PhA and the ECW-to-TBW ratio, using β coefficients and 95% CIs. The normality of PhA and the ECW-to-TBW ratio was confirmed through visual inspection of histograms and Q-Q plots. First, the associations of total PA with PhA and the ECW-to-TBW ratio were evaluated. Second, the associations of LTPA, OPA, and TRPA with PhA and the ECW-to-TBW ratio were examined. Both crude and fully adjusted models were fitted. For further analyses, the PA variables were treated as continuous variables in the model. Additionally, the nonlinear associations between total PA and each domain of PA with PhA and ECW-to-TBW were examined using the restricted cubic spline function with 4 knots [26]. Statistical analyses were performed using R (version 4.5.0; R Foundation for Statistical Computing). The *survey* package and its function *svyglm* were used to account for the sampling design of the KNHANES [27].

Compared with the included participants, those with missing values were more likely to be older, have lower levels of education and income, have lower PhA values, and have a higher ECW-to-TBW ratio (Table S1 in Multimedia Appendix 1). To mitigate this selection bias, for sensitivity analyses, multiple imputation was performed to handle missing data using the R package *mice* [28]. Specifically, 20 complete datasets were generated using the multivariate imputation by chained equations method, and estimates from each dataset were combined using Rubin rules.

Results

Among the 5211 participants, 31.7% (n=1650), 28.7% (n=1496), 19.2% (n=1000), and 20.4% (n=1065) engaged in 0, 1 - 149, 150 - 299, and ≥300 min/week of MVPA, respectively (Table 1). Participants with higher levels of total PA tended to be younger, have higher education and income levels, and not have hypertension, diabetes, or chronic kidney disease. The mean PhA values were 5.0° (SD 0.7), 5.0° (SD 0.7), 5.1° (SD 0.7), and 5.2° (SD 0.7) for those reporting 0, 1 - 149, 150-299, and ≥300 min/week of MVPA, respectively. The corresponding mean ECW-to-TBW ratios were 38.7% (SD 0.8), 38.4% (SD 0.8), 38.5% (0.7), and 38.4% (SD 0.7), respectively.

In the fully adjusted model (Table 2), compared with 0 min/week of MVPA, ≥300 min/week of MVPA was positively associated with PhA (β=0.08, 95% CI 0.04-0.13). Compared with 0 min/week, the ECW-to-TBW ratio was negatively associated with 150 - 299 min/week of MVPA (β=–0.08, 95% CI –0.14 to −0.03), and ≥300 min/week of MVPA (β=–0.11, 95% CI –0.15 to –0.06). Therefore, total PA showed a linear positive association with PhA and a linear negative association with the ECW-to-TBW ratio, even after adjustment for covariates.

XSL•FO
RenderX

Table . Characteristics of the participants.

	Overall (n=5211)	Total PA[a] level (MVPA[b] per wk)			
		0 min (n=1650)	1-149 min (n=1496)	150-299 min (n=1000)	≥300 min (n=1065)
Sex, n (%)					
Male	2256 (43.3)	743 (45)	572 (38.2)	411 (41.1)	530 (49.8)
Female	2955 (56.7)	907 (55)	924 (61.8)	589 (58.9)	535 (50.2)
Age (years), mean (SD)	64.5 (8.7)	65.8 (8.9)	65.1 (8.8)	63.2 (8.2)	62.7 (8.4)
Education level, n (%)					
Middle school or below	2111 (40.5)	818 (49.6)	629 (42)	332 (33.2)	332 (31.2)
High school	1770 (34)	531 (32.2)	501 (33.5)	360 (36)	378 (35.5)
College or above	1330 (25.5)	301 (18.2)	366 (24.5)	308 (30.8)	355 (33.3)
Income level, n (%)					
Lowest	1362 (26.1)	516 (31.3)	424 (28.3)	210 (21)	212 (19.9)
Low	1344 (25.8)	423 (25.6)	391 (26.1)	257 (25.7)	273 (25.6)
High	1175 (22.5)	347 (21)	327 (21.9)	256 (25.6)	245 (23)
Highest	1330 (25.5)	364 (22.1)	354 (23.7)	277 (27.7)	335 (31.5)
Marital status, n (%)					
Married	3979 (76.4)	1211 (73.4)	1126 (75.3)	797 (79.7)	845 (79.3)
Unmarried	149 (2.9)	37 (2.2)	47 (3.1)	32 (3.2)	33 (3.1)
Others	1083 (20.8)	402 (24.4)	323 (21.6)	171 (17.1)	187 (17.6)
Smoking, n (%)					
No	4503 (86.4)	1404 (85.1)	1295 (86.6)	880 (88)	924 (86.8)
Yes	708 (13.6)	246 (14.9)	201 (13.4)	120 (12)	141 (13.2)
Alcohol consumption, n (%)					
No	2920 (56)	983 (59.6)	883 (59)	540 (54)	514 (48.3)
Yes	2291 (44)	667 (40.4)	613 (41)	460 (46)	551 (51.7)
Hypertension, n (%)					
No	3176 (60.9)	931 (56.4)	911 (60.9)	665 (66.5)	669 (62.8)
Yes	2035 (39.1)	719 (43.6)	585 (39.1)	335 (33.5)	396 (37.2)
Diabetes, n (%)					
No	4346 (83.4)	1320 (80)	1245 (83.2)	862 (86.2)	919 (86.3)
Yes	865 (16.6)	330 (20)	251 (16.8)	138 (13.8)	146 (13.7)
Chronic kidney disease, n (%)					
No	5049 (96.9)	1604 (97.2)	1446 (96.7)	969 (96.9)	1030 (96.7)
Yes	162 (3.1)	46 (2.8)	50 (3.3)	31 (3.1)	35 (3.3)
BMI (kg/m^2), mean (SD)	24.1 (3.3)	24.3 (3.4)	24.1 (3.2)	23.9 (3.2)	24.2 (3.3)
Phase angle (°), mean (SD)	5.1 (0.7)	5.0 (0.7)	5.0 (0.7)	5.1 (0.7)	5.2 (0.7)
ECW[c]-to-TBW[d] ratio (%), mean (SD)	38.6 (0.8)	38.7 (0.8)	38.7 (0.8)	38.5 (0.7)	38.4 (0.7)

[a]PA: physical activity.

[b]MVPA: moderate-to-vigorous physical activity.

XSL•FO
RenderX

[c]ECW: extracellular water.

[d]TBW: total body water.

Table . Associations of total physical activity (PA) with the phase angle and extracellular water (ECW)–to–total body water (TBW) ratio.

Total PA (MVPA[a])	Phase angle (°)		ECW-to-TBW ratio (%)	
	Crude model, β (95% CI)	Adjusted model, β (95% CI)	Crude model, β (95% CI)	Adjusted model, β (95% CI)
0 min/wk	Reference	Reference	Reference	Reference
1-149 min/wk	−0.02 (−0.09 to 0.04)	0.02 (−0.01 to 0.06)	−0.03 (−0.09 to 0.03)	−0.04 (−0.08 to 0.01)
150-299 min/wk	0.09 (0.02 to 0.17)	0.04 (−0.01 to 0.09)	−0.19 (−0.27 to −0.12)	−0.08 (−0.14 to −0.03)
≥300 min/wk	0.23 (0.16 to 0.30)	0.08 (0.04 to 0.13)	−0.28 (−0.35 to −0.21)	−0.11 (−0.15 to −0.06)

[a]MVPA: moderate-to-vigorous physical activity.

Table 3 presents the associations of the levels of LTPA, OPA, and TRPA with both PhA and the ECW-to-TBW ratio. For LTPA, compared with 0 min/week, engaging in 1-149, 150-299, and ≥300 min/week of MVPA was associated with 0.02 (95% CI –0.03 to 0.06), 0.08 (95% CI 0.02-0.14), and 0.13 (95% CI 0.06-0.20) degrees higher PhA, respectively. In addition, compared with 0 min/week, participating in 1-149, 150-299, and ≥300 min/week of MVPA was associated with –0.06 (95% CI –0.12 to −0.00), –0.11 (95% CI –0.19 to −0.03), and –0.15 (95% CI –0.22 to −0.07) percentage points lower ECW-to-TBW ratio, respectively. For TRPA, compared with 0 min/week, engaging in ≥300 min/week of MVPA was negatively associated with the ECW-to-TBW ratio (β=–0.07, 95% CI –0.14 to –0.01). However, no clear relations were observed between TRPA and the PhA. Finally, OPA showed no significant associations with either PhA or the ECW-to-TBW ratio. Therefore, the associations with both PhA and the ECW-to-TBW ratio were most consistent for LTPA, whereas TRPA was associated only with a lower ECW-to-TBW ratio, and OPA was not significantly associated with either outcome.

Table . Associations of leisure-time, occupational, and transport-related physical activity with the phase angle and extracellular water (ECW)–to–total body water (TBW)' ratio

	Phase angle (°)		ECW-to-TBW ratio (%)	
	Crude model, β (95% CI)	Adjusted model, β (95% CI)	Crude model, β (95% CI)	Adjusted model, β (95% CI)
0 min/wk	Reference	Reference	Reference	Reference
LTPA[a] (MVPA)[b]				
1-149 min/wk	0.21 (0.13 to 0.28)	0.02 (−0.03 to 0.06)	−0.28 (−0.35 to −0.22)	−0.06 (−0.12 to –0.00)
150-299 min/wk	0.25 (0.17 to 0.34)	0.08 (0.02 to 0.14)	−0.33 (−0.43 to −0.23)	−0.11 (−0.19 to −0.03)
≥300 min/wk	0.43 (0.33 to 0.53)	0.13 (0.06 to 0.20)	−0.44 (−0.54 to −0.35)	−0.15 (−0.22 to −0.07)
OPA[c] (MVPA)				
1-149 min/wk	0.13 (−0.00 to 0.26)	0.04 (−0.05 to 0.14)	−0.01 (−0.14 to 0.11)	0.02 (−0.08 to 0.12)
150-299 min/wk	0.02 (−0.20 to 0.24)	−0.09 (−0.21 to 0.03)	0.04 (−0.19 to 0.28)	0.13 (−0.02 to 0.29)
≥300 min/wk	0.31 (0.20 to 0.42)	0.03 (−0.04 to 0.10)	−0.21 (−0.31 to −0.10)	0.05 (−0.03 to 0.14)
TRPA[d] (MVPA)				
1-149 min/wk	−0.14 (−0.19 to −0.09)	−0.00 (−0.04 to 0.03)	0.07 (0.02 to 0.13)	−0.02 (−0.07 to 0.02)
150-299 min/wk	−0.06 (−0.13 to 0.01)	0.02 (−0.03 to 0.08)	−0.02 (−0.09 to 0.05)	−0.06 (−0.11 to 0.00)
≥300 min/wk	−0.04 (−0.12 to 0.05)	0.03 (−0.03 to 0.08)	−0.02 (−0.10 to 0.07)	−0.07 (−0.14 to −0.01)

[a]LTPA: leisure-time physical activity.

[b]MVPA: moderate-to-vigorous physical activity.

[c]OPA: occupational physical activity.

[d]TRPA: transport-related physical activity.

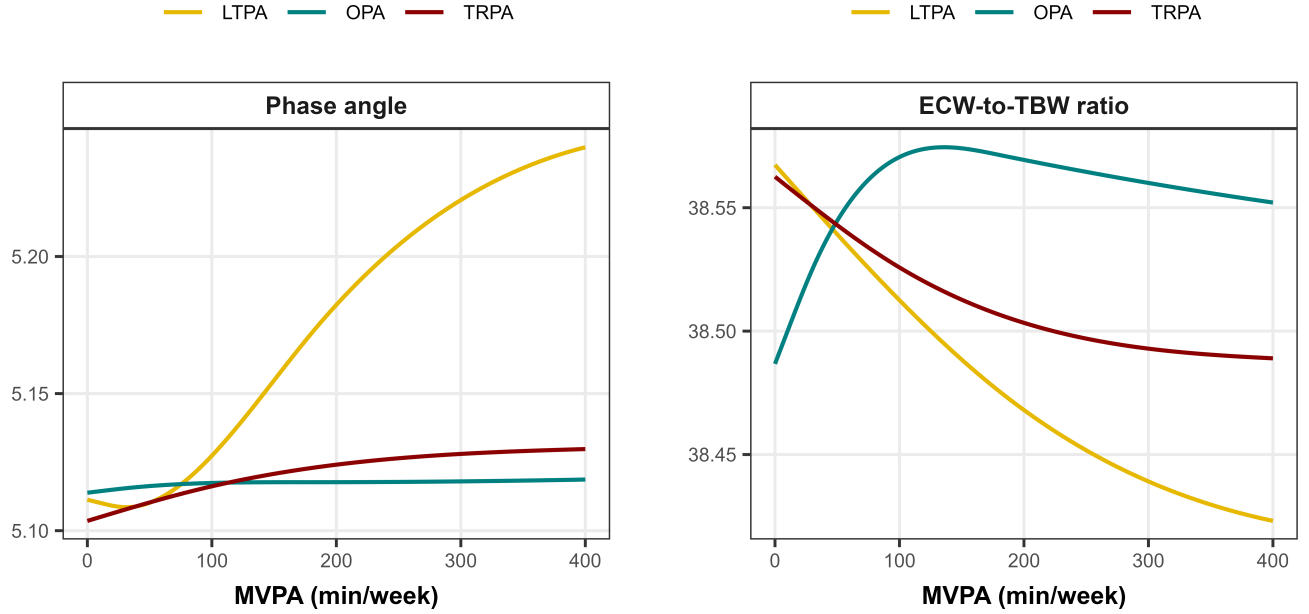
Figure 1 shows the nonlinear associations of LTPA, OPA, and TRPA with PhA and the ECW-to-TBW ratio based on restricted cubic spline models. For PhA, a pronounced positive association was observed with LTPA. TRPA showed a weak positive association, whereas OPA exhibited a flat pattern, indicating no clear association with PhA. For the ECW-to-TBW ratio, LTPA demonstrated the strongest inverse association, whereas TRPA was also negatively associated with the ECW-to-TBW ratio. In contrast, OPA showed a positive association up to

XSL•FO
RenderX

approximately 100 min/week of MVPA, beyond which the association was no longer apparent.

Sensitivity analyses based on the imputed datasets confirmed that total PA was directly associated with PhA and inversely associated with the ECW-to-TBW ratio (Table S2 in Multimedia Appendix 1). The slight differences between the estimates in Table 2 and Table S2 in Multimedia Appendix 1 reflect the differences in analytic approaches between complete-case analysis and multiple imputation, while the estimates are comparable. Moreover, the associations of LTPA, OPA, and TRPA with both PhA and the ECW-to-TBW ratio remained consistent in the sensitivity analyses (Table S3 in Multimedia Appendix 1). When PA was analyzed as a continuous variable (Table S4 in Multimedia Appendix 1), each 150-min/week increment in MVPA from LTPA, OPA, and TRPA was associated with changes in PhA (°) of 0.03 (95% CI 0.01-0.05), 0.00 (95% CI −0.01 to 0.01), and 0.01 (95% CI −0.01 to 0.02), respectively. The corresponding changes in the ECW-to-TBW ratio (%) were −0.03 (95% CI −0.05 to −0.01), 0.01 (95% CI −0.01 to 0.02), and −0.02 (95% CI −0.04 to −0.01), respectively, after adjustment for covariates (model 2).

Figure 1. Nonlinear associations of leisure-time, occupational, and transport-related physical activity with phase angle and the extracellular water (ECW)–to–total body water (TBW) ratio, estimated using restricted cubic spline functions. The models were adjusted for sex, age, education, income, marital status, smoking status, alcohol consumption, hypertension, diabetes, chronic kidney disease, and BMI. LTPA: leisure-time physical activity; MVPA: moderate-to-vigorous physical activity; OPA: occupational physical activity; TRPA: transport-related physical activity.



Discussion

Principal Findings

This study investigated the domain-specific associations of PA with PhA and the ECW-to-TBW ratio in older adults. Total PA was positively associated with PhA and negatively associated with the ECW-to-TBW ratio in older adults. Among the PA domains, LTPA showed the most consistent associations with both PhA and the ECW-to-TBW ratio, whereas high levels of TRPA (≥300 min/week) were negatively associated only with the ECW-to-TBW ratio. Specifically, the nonlinear models demonstrated that the positive association between LTPA and PhA, as well as the negative association between LTPA and the ECW-to-TBW ratio, followed a largely linear pattern without a clear threshold. In contrast, OPA showed no clear associations with either PhA or the ECW-to-TBW ratio.

The results of this study are consistent with prior research demonstrating that PA is positively associated with PhA and negatively associated with the ECW-to-TBW ratio. Studies conducted in South Korea and Poland showed that self-reported PA was associated with a higher PhA among adults [19,29]. Furthermore, previous studies have consistently confirmed the positive association between objectively measured PA and PhA [20,30,31]. Evidence from a meta-analysis also supports the finding that PA is positively associated with PhA [32]. Conversely, previous research reported negative associations between PA levels and the ECW-to-TBW or ECW-to-ICW ratios [18,33]. A recent study using the doubly labeled water method found that PA levels were directly associated with PhA and inversely associated with the ECW-to-ICW ratio [34]. Building on this evidence, the present study provides novel insights by elucidating that the associations between PA and PhA and the ECW-to-TBW ratio vary substantially across PA domains.

It is hypothesized that complex physiological and biochemical mechanisms may underlie the associations between PA and both PhA and the ECW-to-TBW ratio. One of the biological pathways through which PA is linked to BIA-measured outcomes may be associated with the mitigation of systemic inflammation. Long-term engagement in PA was linked to lower levels of circulating inflammatory markers, including C-reactive protein, interleukin-6, and tumor necrosis factor α [35,36]. Chronic inflammation and oxidative stress can induce cellular damage, such as structural injury, membrane disruption, and apoptosis, thereby lowering PhA. Consequently, regular PA may be associated with higher PhA values [12]. Furthermore, PA promotes favorable changes in body composition by

XSL•FO
RenderX

lowering body fat percentage and increasing skeletal muscle mass [37]. Such physiological adaptations to PA are also associated with a lower ECW-to-TBW ratio, reflecting improved intracellular-to-extracellular water balance [38].

Although this study found substantial associations between PA domains and PhA and the ECW-to-TBW ratio, the magnitudes of these associations were modest; for example, engaging in ≥300 min/week of LTPA was associated with a 0.13° higher PhA and a 0.2% lower ECW-to-TBW ratio compared with the 0 min/week. The modest effect sizes of the observed associations may be because BIA-derived profiles are influenced by a wide range of underlying demographic and health-related factors, including sex, age, and chronic conditions. Although these associations may appear modest at the individual level (ie, a 0.13° higher PhA and a 0.2% lower ECW-to-TBW ratio), they correspond to approximately 2 SDs of the distributions of PhA and the ECW-to-TBW ratio in our study population. Therefore, these differences may have meaningful implications at the population level.

The findings that the associations between PA and both PhA and the ECW-to-TBW ratio differed across PA domains warrant attention. These results align with the existing "PA paradox" hypothesis, which posits that the beneficial health effects of PA are primarily derived from LTPA [39]. A study of Korean adults reported that while LTPA was linked to lower high-sensitivity C-reactive protein levels, OPA was associated with higher levels [40]. Furthermore, a recent study of older adults in South Korea found that LTPA was positively associated with the skeletal muscle mass index, whereas OPA showed no significant association [41]. Although multiple factors may contribute to the differential associations of PA domains with health outcomes, one possible explanation is that LTPA generally involves the free and coordinated use of muscle groups throughout the body, whereas OPA often requires repetitive use of specific muscle groups [42]. Evidence from these studies suggests that LTPA is associated with beneficial physiological adaptations, which may explain why LTPA exhibited the most consistent associations with both PhA and the ECW-to-TBW ratio. Particularly for LTPA, individuals can freely determine the frequency and intensity of their activity. Therefore, those with better underlying health status may be more likely to engage in more frequent and vigorous LTPA, which may increase the potential for reverse causation and contribute to the stronger associations observed in the present study. Additionally, individuals with high levels of OPA may be more likely to have lower socioeconomic status and engage in occupations involving substantial physical demands and work-related stress, factors that may be associated with lower PhA values and higher ECW-to-TBW ratios. Such coexposure may have obscured the independent association between OPA and these BIA parameters. Furthermore, OPA was assessed using self-reported questionnaires rather than objective measures, which may not accurately capture the intensity, posture, and frequency of OPA. This limitation may have introduced measurement error and further attenuated the observed associations. Future longitudinal studies are needed to confirm the causal relationship.

The present study has several limitations. First, because of its cross-sectional design, the present study cannot establish the causal effects of PA on PhA and the ECW-to-TBW ratio. For instance, the possibility of reverse causation should be acknowledged; individuals in better physical condition, indicated by higher PhA and lower ECW-to-TBW ratios, may be more likely to engage in PA. This may be particularly relevant for LTPA, as individuals can freely determine the levels and intensity of PA. Therefore, the possibility of reverse causation may be greater for LTPA than for other PA domains and may have contributed to the stronger observed associations. Therefore, the temporal relationship between PA and both PhA and the ECW-to-TBW ratio should be clarified in future longitudinal studies. Second, while the K-GPAQ is a validated tool for evaluating PA among the Korean population and offers the advantage of differentiating LTPA, OPA, and TRPA, potential measurement errors arising from recall or social desirability bias should be recognized. A previous study reported that older adults tend to overreport their levels of PA when assessed using the GPAQ compared with accelerometer-based measurements [22]. Although self-reported and objectively measured PA were significantly correlated, the strength of this correlation was only modest, highlighting the inherent limitations and potential measurement error associated with questionnaire-based PA assessment [22]. Furthermore, the degree of measurement error may vary across PA domains, potentially resulting in differential misclassification and influencing the observed domain-specific associations. It should also be noted that, among older adults, the distinction between TRPA and LTPA may not be clear. Future studies would benefit from using objective measures, such as accelerometers, to improve accuracy. Third, this study did not account for several unmeasured confounders due to the lack of available data. For example, factors such as dietary intake, inflammatory biomarkers, and the use of medications that influence body composition and fluid status were not adjusted for in the regression models. Future in-depth analyses should incorporate these variables to elucidate the physiological mechanisms underlying the associations between PA and both PhA and the ECW-to-TBW ratio.

Despite these limitations, the study has some strengths. To the best of our knowledge, this is the first study to demonstrate domain-specific associations between PA and both PhA and the ECW-to-TBW ratio. Furthermore, the results were derived from a nationally representative sample of middle-aged and older adults, thereby ensuring the generalizability of the results.

Conclusions

This study revealed that PA is differentially associated with PhA and the ECW-to-TBW ratio according to its domain in older adults in South Korea. LTPA demonstrated a notable positive association with PhA and a negative association with the ECW-to-TBW ratio, even at lower levels of activity. The negative association between TRPA and the ECW-to-TBW ratio was observed at higher levels (≥300 min/week), whereas OPA showed no clear associations with either PhA or the ECW-to-TBW ratio. Future research should use longitudinal study designs and objective measures of PA to establish the

XSL•FO
RenderX

temporal sequence of the observed associations and improve the accuracy of PA assessment.

Acknowledgments

The authors thank the researchers at the Korea Disease Control and Prevention Agency. During the preparation of this work, the authors used ChatGPT from OpenAI to enhance grammar and proofread the manuscript under full human supervision. Responsibility for the final manuscript lies entirely with the authors.

Funding

This research was supported by a grant from the National Bio Bigdata Project, funded by four ministries (Ministry of Health and Welfare; Ministry of Science and ICT; Ministry of Trade, Industry and Resources; and Korea Disease Control and Prevention Agency) of Korea (grant RS-2024 - 00462006).

Data Availability

The raw data of the KNHANES were anonymized by the researchers at the Korea Disease Control and Prevention Agency and are available at the agency website [43].

Authors' Contributions

Conceptualization: SUB, JHY
Formal analysis: SUB
Investigation: SUB
Methodology: SUB
Software: SUB
Supervision: JHY
Writing—original draft: SUB
Writing—review and editing: JHY

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample characteristics, sensitivity analyses using imputed datasets, and domain-specific and continuous physical activity associations with phase angle and extracellular water–to–total body water ratio.
[DOCX File, 28 KB - aging_v9i1e90158_app1.docx]

References

1. Strain T, Flaxman S, Guthold R, et al. National, regional, and global trends in insufficient physical activity among adults from 2000 to 2022: a pooled analysis of 507 population-based surveys with 5.7 million participants. Lancet Glob Health 2024 Aug;12(8):e1232-e1243. [doi: 10.1016/S2214-109X(24)00150-5] [Medline: 38942042]
2. Seo YB, Oh YH, Yang YJ. Current status of physical activity in South Korea. Korean J Fam Med 2022 Jul;43(4):209-219. [doi: 10.4082/kjfm.22.0099] [Medline: 35903044]
3. Kim S, Choi S, Kim J, et al. Trends in health behaviors over 20 years: findings from the 1998-2018 Korea National Health and Nutrition Examination Survey. Epidemiol Health 2021;43:e2021026. [doi: 10.4178/epih.e2021026] [Medline: 33872483]
4. Fukushima N, Kikuchi H, Sato H, et al. Dose-response relationship of physical activity with all-cause mortality among older adults: an umbrella review. J Am Med Dir Assoc 2024 Mar;25(3):417-430. [doi: 10.1016/j.jamda.2023.09.028] [Medline: 37925162]
5. Neufer PD, Bamman MM, Muoio DM, et al. Understanding the cellular and molecular mechanisms of physical activity-induced health benefits. Cell Metab 2015 Jul 7;22(1):4-11. [doi: 10.1016/j.cmet.2015.05.011] [Medline: 26073496]
6. Holtermann A, Hansen JV, Burr H, Søgaard K, Sjøgaard G. The health paradox of occupational and leisure-time physical activity. Br J Sports Med 2012 Mar;46(4):291-295. [doi: 10.1136/bjsm.2010.079582] [Medline: 21459873]
7. Pronk N. Physical activity paradox: providing evidence-based guidance while closing research gaps. Br J Sports Med 2024 Jul 31;58(16):932-933. [doi: 10.1136/bjsports-2024-108294] [Medline: 38760157]
8. Bernabe-Ortiz A, Carrillo-Larco RM, Gilman RH, Smeeth L, Checkley W, Miranda JJ. Leisure-time and transport-related physical activity and the risk of mortality: the CRONICAS cohort study. J Phys Act Health 2022 Feb 1;19(2):118-124. [doi: 10.1123/jpah.2021-0672] [Medline: 35051901]

XSL•FO
RenderX

9. Kazemi A, Soltani S, Aune D, et al. Leisure-time and occupational physical activity and risk of cardiovascular disease incidence: a systematic-review and dose-response meta-analysis of prospective cohort studies. Int J Behav Nutr Phys Act 2024 Apr 24;21(1):45. [doi: 10.1186/s12966-024-01593-8] [Medline: 38659024]
10. Coenen P, Huysmans MA, Holtermann A, et al. Associations of occupational and leisure-time physical activity with all-cause mortality: an individual participant data meta-analysis. Br J Sports Med 2024 Dec 18;58(24):1527-1538. [doi: 10.1136/bjsports-2024-108117] [Medline: 39255999]
11. Martins PC, Alves Junior CA, Silva AM, Silva DA. Phase angle and body composition: a scoping review. Clin Nutr ESPEN 2023 Aug;56:237-250. [doi: 10.1016/j.clnesp.2023.05.015] [Medline: 37344079]
12. da Silva BR, Orsso CE, Gonzalez MC, et al. Phase angle and cellular health: inflammation and oxidative damage. Rev Endocr Metab Disord 2023 Jun;24(3):543-562. [doi: 10.1007/s11154-022-09775-0] [Medline: 36474107]
13. Lai TF, Park JH, Jang M, et al. Elevated extracellular water to total body water ratio and low phase angle in relation to muscle function in middle-aged and older adults. J Int Soc Sports Nutr 2025 Dec;22(1):2536693. [doi: 10.1080/15502783.2025.2536693] [Medline: 40771175]
14. Ceolin J, de Borba EL, Mundstock E, de Oliveira JR, Mattiello R, Bodanese LC. Phase angle of bioimpedance as a marker of inflammation in cardiovascular diseases: a systematic review. Nutrition 2023 Aug;112:112064. [doi: 10.1016/j.nut.2023.112064] [Medline: 37263162]
15. Wilhelm-Leen ER, Hall YN, Horwitz RI, Chertow GM. Phase angle, frailty and mortality in older adults. J Gen Intern Med 2014 Jan;29(1):147-154. [doi: 10.1007/s11606-013-2585-z] [Medline: 24002625]
16. Garlini LM, Alves FD, Ceretta LB, Perry IS, Souza GC, Clausell NO. Phase angle and mortality: a systematic review. Eur J Clin Nutr 2019 Apr;73(4):495-508. [doi: 10.1038/s41430-018-0159-1] [Medline: 29695763]
17. Wang Y, Liu J, Hao H, et al. Non-linear association between extracellular water/total body water ratio and all-cause mortality: a population-based cohort study. Sci Rep 2025;15(1):19032. [doi: 10.1038/s41598-025-04202-1] [Medline: 40447756]
18. Jaremków A, Markiewicz-Górka I, Hajdusianek W, Czerwińska K, Gać P. The relationship between body composition and physical activity level in students of medical faculties. J Clin Med 2023 Dec 21;13(1):50. [doi: 10.3390/jcm13010050] [Medline: 38202057]
19. Yang J, Yu J, Kim J, Park E. Association between physical activity and phase angle obtained via bioelectrical impedance analysis in South Korean adults stratified by sex. Nutrients 2024 Jul 4;16(13):2136. [doi: 10.3390/nu16132136] [Medline: 38999883]
20. Nakashima D, Fujii K, Tsubouchi Y, et al. Relationship between phase angle and physical activity intensity among community-dwelling older adults in Japan: a cross-sectional study. Healthcare (Basel) 2024 Jan 10;12(2):167. [doi: 10.3390/healthcare12020167] [Medline: 38255056]
21. Kweon S, Kim Y, Jang MJ, et al. Data resource profile: the Korea National Health and Nutrition Examination Survey (KNHANES). Int J Epidemiol 2014 Feb;43(1):69-77. [doi: 10.1093/ije/dyt228] [Medline: 24585853]
22. Lee J, Lee C, Min J, et al. Development of the Korean Global Physical Activity Questionnaire: reliability and validity study. Glob Health Promot 2020 Sep;27(3):44-55. [doi: 10.1177/1757975919854301] [Medline: 31375056]
23. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462. [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
24. Yi Y, Baek JY, Lee E, Jung HW, Jang IY. A comparative study of high-frequency bioelectrical impedance analysis and dual-energy X-ray absorptiometry for estimating body composition. Life (Basel) 2022 Jul 4;12(7):994. [doi: 10.3390/life12070994] [Medline: 35888083]
25. Yang J, Kim J, Chun BC, Lee JM. Cook with different pots, but similar taste? Comparison of phase angle using bioelectrical impedance analysis according to device type and examination posture. Life (Basel) 2023 Apr 30;13(5):1119. [doi: 10.3390/life13051119] [Medline: 37240764]
26. Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. Stat Med 2010 Apr 30;29(9):1037-1057. [doi: 10.1002/sim.3841] [Medline: 20087875]
27. Lumley T. Analysis of complex survey samples. J Stat Softw 2004;9(8):1-19. [doi: 10.18637/jss.v009.i08]
28. Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. Ann Transl Med 2016 Jan;4(2):30. [doi: 10.3978/j.issn.2305-5839.2015.12.63] [Medline: 26889483]
29. Musijowska M, Kwilosz E. Association between physical activity level, body composition, and phase angle in university students from bioelectrical impedance analysis (BIA). J Clin Med 2024 May 7;13(10):2743. [doi: 10.3390/jcm13102743] [Medline: 38792285]
30. Chen J, Lai TF, Lin CY, Hsueh MC, Park JH, Liao Y. Associations between objectively measured overall and intensity-specific physical activity and phase angle in older adults. Sci Rep 2024;14(1):7309. [doi: 10.1038/s41598-024-57544-7] [Medline: 38538678]
31. Yamada Y, Yoshida T, Murakami H, et al. Phase angle obtained via bioelectrical impedance analysis and objectively measured physical activity or exercise habits. Sci Rep 2022;12(1):17274. [doi: 10.1038/s41598-022-21095-6] [Medline: 36241873]

XSL•FO
RenderX

32. Mundstock E, Amaral MA, Baptista RR, et al. Association between phase angle from bioelectrical impedance analysis and level of physical activity: systematic review and meta-analysis. Clin Nutr 2019 Aug;38(4):1504-1510. [doi: 10.1016/j.clnu.2018.08.031] [Medline: 30224304]
33. Iwasaka C, Yamada Y, Nishida Y, et al. Association of appendicular extracellular-to-intracellular water ratio with age, muscle strength, and physical activity in 8,018 community-dwelling middle-aged and older adults. Arch Gerontol Geriatr 2023 May;108:104931. [doi: 10.1016/j.archger.2023.104931] [Medline: 36701946]
34. Oishi K, Yoshida T, Kim HK, et al. Association of physical activity and hydration markers with cellular health in older adults: a doubly labeled water study. Exp Gerontol 2025 Oct 1;209:112859. [doi: 10.1016/j.exger.2025.112859] [Medline: 40803496]
35. Kasapis C, Thompson PD. The effects of physical activity on serum C-reactive protein and inflammatory markers: a systematic review. J Am Coll Cardiol 2005 May 17;45(10):1563-1569. [doi: 10.1016/j.jacc.2004.12.077] [Medline: 15893167]
36. Zheng G, Qiu P, Xia R, et al. Effect of aerobic exercise on inflammatory markers in healthy middle-aged and older adults: a systematic review and meta-analysis of randomized controlled trials. Front Aging Neurosci 2019;11:98. [doi: 10.3389/fnagi.2019.00098] [Medline: 31080412]
37. Bradbury KE, Guo W, Cairns BJ, Armstrong ME, Key TJ. Association between physical activity and body fat percentage, with adjustment for BMI: a large cross-sectional analysis of UK Biobank. BMJ Open 2017 Mar 24;7(3):e011843. [doi: 10.1136/bmjopen-2016-011843] [Medline: 28341684]
38. Harimawan AI, Prabandari A, Wihandani DM, et al. Association between phase angle and ECW/TBW ratio with body composition in individuals with central obesity: a cross-sectional study. Front Nutr 2025;12:1638075. [doi: 10.3389/fnut.2025.1638075] [Medline: 41098783]
39. Holtermann A, Krause N, van der Beek AJ, Straker L. The physical activity paradox: six reasons why occupational physical activity (OPA) does not confer the cardiovascular health benefits that leisure time physical activity does. Br J Sports Med 2018 Feb;52(3):149-150. [doi: 10.1136/bjsports-2017-097965] [Medline: 28798040]
40. Lee J, Kim HR, Jang TW, Lee DW, Lee YM, Kang MY. Occupational physical activity, not leisure-time physical activity, is associated with increased high-sensitivity C reactive protein levels. Occup Environ Med 2021 Feb;78(2):86-91. [doi: 10.1136/oemed-2020-106753] [Medline: 32912859]
41. Kim B, Osuka Y, Okubo Y, Zhao X, Kim GM, Oh S. The physical activity paradox in low muscle mass in middle-aged and older adults. Am J Prev Med 2025 Feb;68(2):348-356. [doi: 10.1016/j.amepre.2024.10.015] [Medline: 39477133]
42. Quinn TD, Prince SA, Pronk NP, Barone Gibbs B. The physical activity health paradox: what is it, why might it happen, and where to go from here? Int J Behav Nutr Phys Act 2026 May 8;23(1):67. [doi: 10.1186/s12966-026-01922-z] [Medline: 42104355]
43. Health and Nutrition Survey and Analysis Division, Korea Disease Control and Prevention Agency. URL: https://knhanes.kdca.go.kr/knhanes/main.do [accessed 2026-09-17]

Abbreviations

ECW: extracellular water
ICW: intracellular water
K-GPAQ: Korean version of the Global Physical Activity Questionnaire
KDCA: Korea Disease Control and Prevention Agency
KNHANES: Korea National Health and Nutrition Examination Survey
LTPA: leisure-time physical activity
MVPA: moderate to vigorous physical activity
OPA: occupational physical activity
PA: physical activity
PhA: phase angle
TBW: total body water
TRPA: transport-related physical activity
WHO: World Health Organization

XSL•FO
RenderX

Edited by Y Jiang; submitted 22.Dec.2025; peer-reviewed by L Vorobeľová, L Ardigò; revised version received 09.Aug.2026; accepted 10.Aug.2026; published 28.Sep.2026.

Please cite as:
Baek SU, Yoon JH
Domain-Specific Associations of Physical Activity With Phase Angle and Extracellular Water–to–Total Body Water Ratio in Middle-Aged or Older Adults: Cross-Sectional Study
JMIR Aging 2026;9:e90158
URL: https://aging.jmir.org/2026/1/e90158
doi:10.2196/90158

© Seong-Uk Baek, Jin-Ha Yoon. Originally published in JMIR Aging (https://aging.jmir.org), 28.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Tailored Online Physical Activity Coaching for Middle-Aged and Older Adults With Cognitive and Mental Health Concerns: Single-Arm Pre-Post Intervention Study

Kathryn A Ellis[1,2,3], MPsych (Clin), PhD; Rhoda Lai[1], BA (Hons); Eleanor Curran[1,4], MBBS(Hons), BMedSci (Hons), MPM; Jennifer Southam[1], BBehavSci-Psychology(Hons), Grad Dip Health Promotion, Grad Dip Education; Rebecca Moorhead[5], MSW; Kay L Cox[6], PhD; Serafino G Mancuso[1], PhD, MBiostat; Alissa Westphal[1,3], BAppSc (OccTher)(Hons), GDip AppSc(Stats); Terence W H Chong[1,3], MBBS, MP, MBA; Thomas Rego[1,4], MBBS, BMedSci, MPM; Victoria J Palmer[7,8], PhD; Kaarin J Anstey[9,10], PhD; Nicola T Lautenschlager[1,4], MD

[1]Department of Psychiatry, The University of Melbourne, Level 3, Alan Gilbert Building, 161 Barry Street, Parkville, Victoria, Australia

[10]Neuroscience Research Australia, Randwick, New South Wales, Australia

[2]Melbourne School of Psychological Sciences, The University of Melbourne, Parkville, Victoria, Australia

[3]St Vincent's Hospital Melbourne, Kew, Victoria, Australia

[4]Older Adult Mental Health Program, The Royal Melbourne Hospital, Melbourne, Victoria, Australia

[5]Murdoch Children's Research Institute, Parkville, Victoria, Australia

[6]Medical School, The University of Western Australia, Perth, Western Australia, Australia

[7]The ALIVE National Centre for Mental Health Research Translation, The University of Melbourne, Parkville, Victoria, Australia

[8]The Department of General Practice and Primary Care, The University of Melbourne, Parkville, Victoria, Australia

[9]School of Psychology, University of New South Wales, Sydney, New South Wales, Australia

Corresponding Author:
Kathryn A Ellis, MPsych (Clin), PhD
Department of Psychiatry, The University of Melbourne, Level 3, Alan Gilbert Building, 161 Barry Street, Parkville, Victoria, Australia

Abstract

Background: Low levels of physical activity (PA) increase dementia risk, and for middle-aged and older adults with co-occurring cognitive concerns and mental health symptoms, dementia risk increases further. Despite clinical trials showing lower adherence to PA interventions in high-risk groups, there is a sparsity of interventions tailored to support unique behavior change needs. In EXCEL (Exercise for Cognitive Health) phase 1, we developed a model to understand the needs of this population and identified tailoring requirements to enhance engagement. Here we report the findings of a pilot online intervention designed to support middle-aged and older adults with subjective cognitive decline or mild cognitive impairment and mild to moderate symptoms of depression or anxiety to meet PA guidelines.

Objective: We aimed to measure (1) efficacy of the intervention in promoting adoption of PA in line with national guidelines, (2) acceptability, feasibility, and safety of the intervention, (3) changes to dementia risk, PA levels, mental health symptoms, and stages of change, and (4) changes in potential cognitive mechanisms of change.

Methods: A pilot individual 12-week online home-based PA intervention. Participants aged 45-80 years, experiencing both cognitive and mild-to-moderate anxiety and depression symptoms, were prescribed individually tailored PA programs combining aerobic and strength PA, plus balance training as indicated, with fortnightly online coaching.

Results: A total of 55 participants were enrolled (46 females/9 males; mean age 62.2, SD 7.6 y). The intervention was effective; at baseline, only 3 of 55 (6%) participants met all applicable PA guidelines, compared to 24 of 55 (44%) participants post intervention (P<.001). Participants were at least 8 times more likely to meet the age-appropriate guidelines for aerobic, strength, and balance activities after the intervention. Retention rates were high (95% completion rate), and feedback indicated 98% found the program useful. Safety was successfully monitored by a clinical panel via email, using a system of alerts. Dementia risk was reduced (d=−0.32, P=.008), and reductions in depression, anxiety, and stress scores were large and clinically meaningful (d=−1.31, d=−0.89, and d=−1.18, respectively; all P<.001). Over half of the cohort (27 participants) transitioned to a higher stage of change at postintervention. Improvements in action planning (d=0.66, P<.001) and positive outcome expectancies (d=0.33, P=.01) indicate potential cognitive mechanisms of change.

XSL•FO
RenderX

Conclusions: Our findings showed that this tailored, home-based online intervention successfully supports an at-risk cohort of aging adults to adopt PA, in line with national guidelines. The intervention was acceptable, feasible, and safe and led to significant reductions in dementia risk and mental health symptoms and was associated with progression through stage of change. Reflecting the Capability, Opportunity, and Motivation Behavior model underlying the intervention design, improvements in action planning and outcome expectancy appeared to play a role in bridging the PA intention-behavior gap.

Trial Registration: Anzctr.org.au ACTRN12621001152819; https://tinyurl.com/bdeehn6f

(JMIR Aging 2026;9:e80040) doi:10.2196/80040

KEYWORDS

physical activity; exercise; older adults; aging; depression; anxiety; cognitive impairment; subjective cognitive decline; online intervention; coaching

Introduction

Low physical activity (PA) levels are an established modifiable risk factor for dementia [1-3]. Depression and anxiety are also linked to dementia risk [4,5] and are associated with low levels of PA and other dementia risk factors [6]. The presence of cognitive concerns further adds to dementia risk, beyond the contribution of mood symptoms [7-9]. Accordingly, aging individuals who are concerned about their cognition and experience depressive or anxiety symptoms are considered an at-risk group for dementia [1,3,10]. The World Health Organization has emphasized the critical importance of directing dementia risk reduction (DRR) implementation research toward such at-risk populations, recognizing their heightened vulnerability and the urgent need for tailored and targeted strategies [11].

In 2018, our group developed PA guidelines for older adults living with mild cognitive impairment (MCI) or subjective cognitive impairment [12], complementing existing adults' and older adults' PA guidelines [13,14]. These guidelines recommend 150 minutes of moderate aerobic PA plus 2 strength sessions (of approximately 30 min each) per week, along with balance exercises as often as possible for older adults [12,13]. However, research shows that consumer awareness of PA guidelines is very low; none of the middle-aged to older adults interviewed by our team were aware of the guidelines [15], consistent with similarly low awareness internationally [16,17]. Implementation research is now vital to ensure individuals in middle to later life with mental and cognitive health symptoms can benefit from these guidelines, and from DRR interventions more broadly.

Symptoms of depression and anxiety, however, are associated with lower adherence and poorer outcomes in both PA and DRR intervention trials, as well as being a barrier to participation [6,18-20]. Tailored interventions for this population are essential to reduce dementia risk [11,21]. In a prior study [14], our team conducted a comprehensive analysis integrating data from semistructured interviews with 21 participants, a review of existing literature, and the Capability, Opportunity, and Motivation Behavior (COM-B) model to identify three behavior-change areas critical for this high-risk group: (1) emotional regulation, (2) overcoming barriers to turning intention into action, and (3) building confidence in one's abilities [14]. These findings informed a model outlining strategies for tailoring PA interventions to the emotional and cognitive needs of this cohort. We additionally identified engagement preferences with information and communications technology [15,22], which was seen as essential to overcome access barriers [23], and technology was frequently cited as a useful tool by our target group for the accountability and connection they sought from PA interventions [15]. Indeed, older people appear to be well placed to benefit from PA interventions, given that we have seen relatively high and rising information and communications technology use in this age group [24]. We used this model [14] as the foundation for developing the EXCEL (Exercise for Cognitive Health) intervention. The PA programs themselves were based on our established work, which has successfully prescribed PA plans to older adults [25]. The EXCEL intervention was finalized at the beginning of a period of COVID-19–related lockdown in Victoria, Australia. As social isolation is linked to heightened experiences of symptoms of depression and anxiety [26,27] and a higher prevalence of these conditions [28], interventions for individuals with poorer mental health became more crucial than ever, and PA engagement and delivery were essential.

In this paper, we present the findings of EXCEL, a tailored 12-week home-based PA-delivered intervention designed to help middle-aged and older adults living with mild to moderate symptoms of depression, anxiety, or stress and cognitive concerns in adopting PA guidelines for DRR. Our primary hypothesis was that EXCEL would be associated with improved adherence to PA guidelines. We further hypothesized that the program would be acceptable, feasible, and safe, and associated with reductions in dementia risk and mood symptoms, and promote progression through stages of change. We further aimed to explore whether the intervention impacted potential cognitive mechanisms of behavior change.

Methods

Study Design

A single-arm pilot study of a 12-week online intervention combining PA prescription and individualized coaching. The current study measured the efficacy of the intervention for helping participants to meet age-appropriate PA guidelines [12,14], with secondary outcomes exploring acceptability, feasibility, and safety, and whether the intervention modified dementia risk, overall PA levels, mental health symptoms, and stages of change. We examined pre-post changes in measures of cognitive mechanisms of behavior change.

XSL•FO
RenderX

Ethical Considerations

This study received ethical approval from the University of Melbourne Human Research Ethics Committee (2021-20479-16470-3). All published data is deidentified. Participants provided written informed consent to take part in this study before screening. Participants kept all programs and equipment provided during this study but did not receive other compensation for participating.

Participants

Participants were eligible for inclusion if they were aged between 45 and 80 years; English-speaking; living in the community; had access to a phone and/or internet; reported concerns about their cognition (with or without objective cognitive decline) but not dementia; and experienced mild to moderate symptoms of depression, anxiety, or stress (measured using the short-form Depression Anxiety Stress Scales [DASS-21] [29]). Concerns about cognition were reported by a "yes" response to either of the questions: "Do you think your memory or thinking has changed?" or "Do you think your memory or thinking is worse than for other people your age?" Exclusion criteria were lack of capacity to provide informed consent; inability to read or write in English; diagnosis of dementia or a score indicative of dementia (27 or lower) on the Telephone Interview for Cognitive Status [30]; scores in the severe category or higher on either the depression, anxiety, or stress dimensions of the DASS-21 [29]; acute or past history of significant mental illness; current suicidality or history of suicidal attempts within the past 2 years; significant uncorrected sensory impairment; severe mobility impairment (either an inability to walk unaided or a score of 12 or more, indicative of medium to high risk, on the Falls Risk Assessment Tool [31]); any other significant medical condition that would prevent safely undertaking moderate-intensity PA (gauged using the Adult Pre-Exercise Screening System [32] and the Physical Activity Readiness Questionnaire [33]); or a BMI of less than 18.5 or 38 or higher (indicating that they were underweight or very obese). Where applicable, ethnically adjusted BMI indicators were used [34]. Participants were recruited from various sources across Australia, including state-wide seniors' newsletters, university media, general practitioners, online self-registration services (StepUp platform) [35], and a database of volunteers who previously consented to be approached for research. Our target sample size was informed by a 1-sample proportion test [36] based on the AIBL Active Study [25]. To detect a 15% increase in participants meeting moderate-intensity PA guidelines with α=.05 and an expected attrition rate of 30% [37], we calculated a required sample size of 70, which would yield power of 0.80 to detect this effect.

Intervention

EXCEL was a 12-week home-based PA intervention comprising the following resource and support components.

PA Equipment, Guide, and Programs

We developed a PA guide incorporating aerobic, strength, and balance programs tailored to 2 age streams: 45 - 59 years (based on the standard national PA guidelines [14]) and 60 - 80 years (based on national guidelines for older adults with MCI or subjective cognitive decline [12]). While warm-up and cool-down programs were identical for both streams, aerobic and strength programs were customized. Within each stream, participants received either a starter or intermediate program based on their health, abilities, and current activity level. Strength programs differed slightly between males and females. All older age stream participants received a balance program with options for progression. Core program content was developed from successful past studies, including the INDIGO (Individual Goal-Setting and Volunteer Mentors) study [38] and work by Yang et al [39]. The PA guide (Figure 1) provided structured guidance on increasing duration, intensity, number of sessions, and resistance or weights to meet the PA guidelines in 12 weeks. The program incorporated flexibility, with progression guides suggesting rather than prescribing exercises. Each warm-up and cool-down, strength, and balance program included summary pages for easy reference (Figure 2), plus exercise-specific pages detailing performance instructions, pictures, and links to YouTube (Google LLC) demonstration videos (Figure 3). Along with the PA guide and program, participants received a wearable wrist activity tracker, dumbbells, and resistance bands via post if they did not already own this equipment. Participants had access to an online PA diary hosted on the REDCap (Research Electronic Data Capture; Vanderbilt University) platform [40], which could be printed in hard copy or PDF format upon request.

XSL•FO
RenderX

Figure 1. Example of an aerobic program used in the EXCEL intervention. EXCEL: Exercise for Cognitive Health; mod: moderate; MOD: moderate.

EXCEL Study Part Two

OLDER ADULTS AEROBIC PROGRAM

STARTER LEVEL WALKING (5 DAYS)

WEEKS 1-12

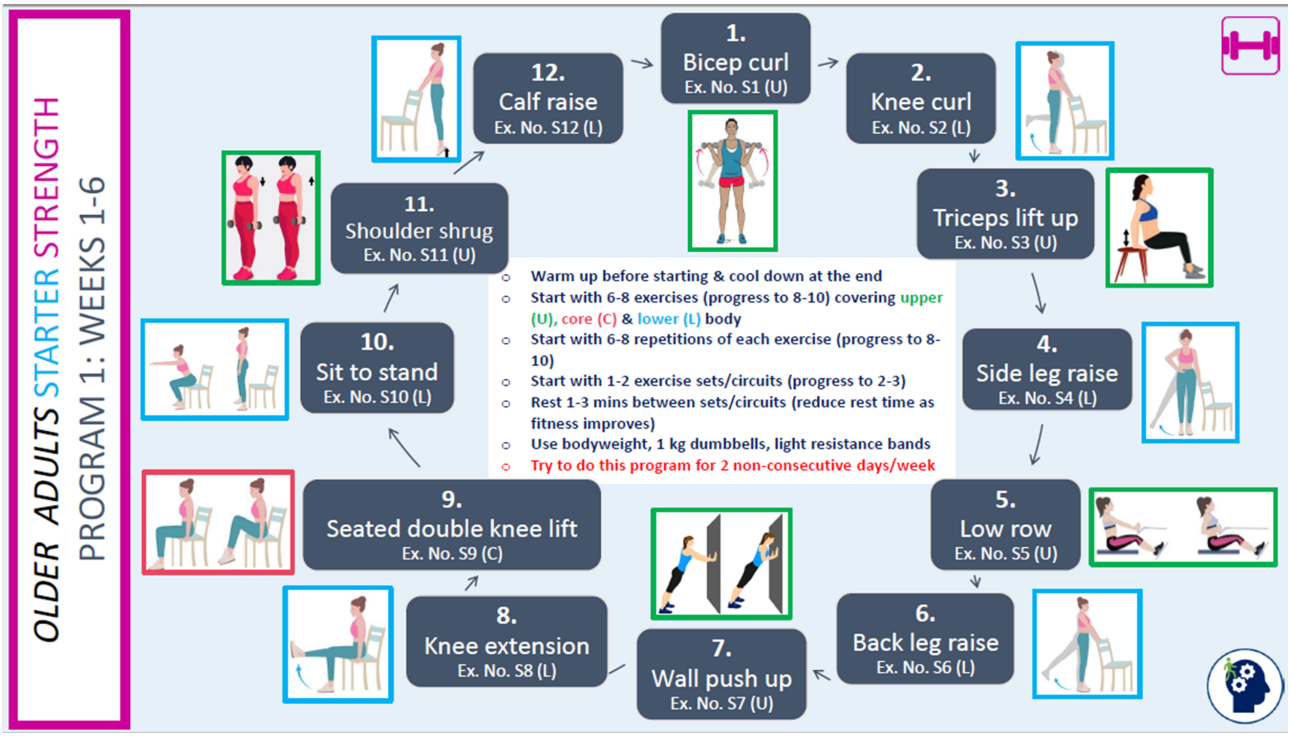
Aerobic

Try to walk most days. This program is for walking on 5 DAYS EACH WEEK
> In the early weeks you can walk in shorter blocks (10 or 15-minutes) with 1-2 min rest if needed
> You can swap WATER WALKING for walking if you have arthritis or joint issues

WEEKS		WALK PACE	TIME (mins)	EFFORT
1 & 2		Slow pace WALK	25	**LOW:** Easy & not puffing
3 & 4	Warm up:	Slow pace WALK	5	**LOW:** Easy & not puffing
		Slow-mod pace WALK	20	**LOW-MOD:** Easy but breathing a bit more
	Cool down:	Slow pace WALK	5	**LOW:** Easy & not puffing
5 & 6	Warm up:	Slow pace WALK	5	**LOW:** Easy & not puffing
		Moderate pace WALK	20	**MODERATE:** Puffing a little, can easily talk
	Cool down:	Slow pace WALK	5	**LOW:** Easy & not puffing
7 & 8	Warm up:	Slow-mod pace WALK	5	**LOW-MOD:** Easy but breathing a bit more
		Moderate pace WALK	20	**MODERATE:** Puffing a little, can easily talk
	Cool down:	Slow-mod pace WALK	5	**LOW-MOD:** Easy but breathing a bit more
9, 10, 11 & 12	Warm up:	Slow-mod pace WALK	5	**LOW-MOD:** Easy but breathing a bit more
		Moderate-brisk pace WALK	20	**MOD-BRISK:** Puffing a little more still able to talk
	Cool down:	Slow-mod pace WALK	5	**LOW-MOD:** Easy but breathing a bit more

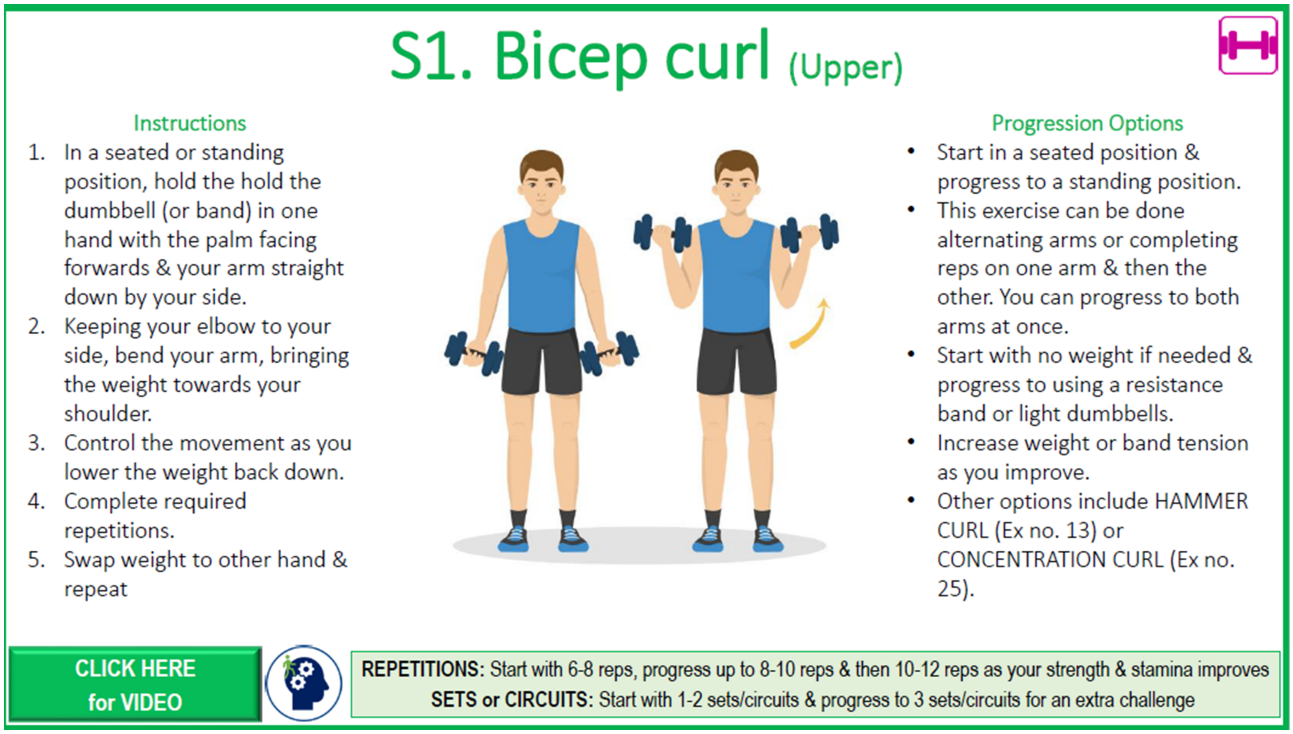
Remember to do your exercises for STRENGTH (at least twice a week) & BALANCE (as often as you can)

Figure 2. Example of part of a strength program used in the EXCEL intervention. C: core; Ex. no.: exercise number; EXCEL: Exercise for Cognitive Health; L: lower; U: upper.



XSL•FO
RenderX

Figure 3. Example of a detailed exercise page in an EXCEL study strength program. Ex. no.: exercise number; EXCEL: Exercise for Cognitive Health; reps: repetitions.



Coaching or Support

The online support structure consisted of 3 types of sessions as outlined below.

The introduction session was an approximately 70-minute online videoconference session that familiarized participants with key sections of the PA guide, including guidelines and safety information. Participants received personalized directions from an exercise specialist (JS) and were able to ask questions.

Of fortnightly catch-up sessions, there were online sessions (20 min or less) with EXCEL coaches, who checked participants' diaries, discussed difficulties, and used behavior change strategies tailored to individual needs [22]. These strategies included education, motivation, incentivization, reinforcement, training, environmental restructuring, modeling, and enablement. The 4 EXCEL coaches came from various backgrounds, including psychology, PA, social work, and occupational therapy.

Of exercise specialist sessions, there were as-needed online sessions with the team exercise science specialist (JS) provided extra assistance on exercise elements of the program for participants who requested or required additional support.

Procedure

Figure 4 summarizes the participant procedure. Participants completed a consent form and prescreening survey and proceeded to a screening call (50 min), completing additional measures to confirm eligibility. Eligible participants were assigned to a PA program and scheduled for an introductory session before completing self-report baseline measures (≈40 min). Their general practitioner was notified of their involvement in the PA intervention. Participants completed the DASS-21 again at weeks 4 and 8 (each ≈5 min). Safety was prioritized, with participants scoring in the severe range or higher on any DASS-21 subscale contacted using a safety protocol script. Fortnightly consultation sessions were held between EXCEL coaches and the research team, which consisted of psychiatrists (NTL, TWHC, TR, and EC), a clinical psychologist (KAE), and exercise physiologists (KLC and JS), to monitor participant safety. Following the 12-week program, participants completed self-report postintervention measures (≈40 min) and an evaluation survey (≈8 min) via Zoom.

XSL•FO
RenderX

Figure 4. The EXCEL study participant flowchart. DASS-21: Depression Anxiety Stress Scale-21; EXCEL: Exercise for Cognitive Health; PA: physical activity.

XSL•FO
RenderX

Outcome Measures

Outcome measures are summarized in Table 1. The primary outcomes reflected successful uptake of age-appropriate PA guidelines, specifically meeting the aerobic recommendation of at least 150 minutes per week of moderate or vigorous intensity activity at follow-up and achieving the full set of PA guidelines encompassing aerobic, strength, and balance activities. Secondary outcomes included acceptability, feasibility, and safety, changes in general dementia risk levels, changes in total moderate-to-vigorous PA that incorporates everyday activities, changes in mental health symptoms, and progression through stage of change. Finally, 2 categories of mechanism-focused outcomes were evaluated. The first relates to potential cognitive mechanisms underpinning exercise behavior change, including shifts in knowledge, motivation, and self-efficacy from pre- to postintervention. The second examined potential mechanisms associated with general dementia risk-reduction behavior change over the same period.

XSL•FO
RenderX

Table . Outcome variables.

Outcome measures	Details
Primary outcomes	• Data for primary outcomes were taken from the CHAMPS[a] questionnaire [41]. For assessing guideline uptake, only purposeful PA[b] was included, that is, PA done with the purpose of doing exercise. This excluded incidental or everyday PAs, such as gardening, housework, etc.
Uptake of aerobic PA guidelines	• Measured by proportion of participants completing 150 minutes or more per week of at least moderate intensity aerobic PA (purposeful PA not including incidental or everyday activities) pre- and postintervention.
Uptake of full PA guidelines	• Measured by the proportion of participants completing 150 minutes or more of at least moderate intensity aerobic PA (purposeful PA not including incidental or everyday activities) as well as 2 or more strength training sessions and at least 1 balance session (older adults only) per week pre- and postintervention.
Secondary outcomes	• Data for secondary outcomes were taken from a range of tools listed below.
Acceptability, feasibility, and safety	• Measured by retention (intervention withdrawals or completions) using a postintervention evaluation questionnaire (developed by the research team to evaluate enjoyment, usefulness, and usability of various elements of the intervention). • Safety was assessed by successful monitoring by a clinical panel using a system of email alerts and a safety protocol.
General dementia risk levels	• Measured by dementia risk and cognitive health scores pre- and postintervention. This tool is a risk assessment calculator used to understand an individual's risk of developing dementia (CogDRisk) [42].
Change in PA	• Measured by overall minutes of moderate to vigorous intensity PA (including everyday relevant activities) pre- and postintervention using CHAMPS [41]. • Measured by overall minutes of any intensity PA (including everyday relevant activities) pre- and postintervention using CHAMPS [41]. • Measured by adherence to PA guidelines in 3 domains (aerobic, strength, and balance) over 12 weeks using self-report PA diary data.
Change in mental health symptoms	• Measured by depression, anxiety, and stress scores pre- and postintervention. This instrument rates these 3 dimensions of mental health (DASS-21[c]) [29].
Changes in change readiness	• Measured by stages of change scores pre- and postintervention including any shift in change stage. This tool assesses an individual's readiness to change a specific behavior (PA) by evaluating their agreement with statements corresponding to different stages of change: precontemplation, contemplation, preparation, action, and maintenance (modified SOC[d] [43]).

XSL•FO
RenderX

Outcome measures	Details
Change in potential exercise-related cognitive mechanisms of exercise behavior change	• Measured by self-efficacy for exercise scores pre- and postintervention. This 9-item scale measures an individual's confidence in their ability to exercise 3 times per week for 20 minutes in the context of 9 different situations that may interfere with exercise. Items are rated from 0 to 10; 0 represents not confident and 10 represents very confident (modified SEES[e] [44]). • Measured by outcome expectations for exercise scores pre- and postintervention. This 13-item scale includes positive and negative subscales and measures the strength of outcome expectations for exercise. It uses a Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree; modified OEES-2[f] [45]). • Measured by exercise benefits and barriers for scores pre- and postintervention. This 43-item scale measures perceptions about benefits or barriers to exercising and uses a Likert-type response ranging from 4 (strongly agree) to 1 (strongly disagree; modified EBBS[g] [46]). • Measured by action planning for exercise scores pre- and postintervention. This scale is used to measure the extent to which individuals plan and commit to actions related to their goals and uses a 6-point Likert scale. Participants respond to items that question specific aspects of planning (ie, "I have a clear plan for how to achieve my goal"; APS[h] [47]).
Change in potential cognitive mechanisms of DRR[i] related behavior change	• Measured by dementia knowledge scores pre- and postintervention (DKAS-subscale D[j] [48]). • Measured by motivation to change dementia risk behavior scores pre- and postintervention. This 27-item scale assesses beliefs and attitudes toward DRR and uses a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree; MCLHB-DRR[k] [49]).

[a]CHAMPS: Community Health Activities Model Program for Seniors.

[b]PA: physical activity.

[c]DASS-21: Depression Anxiety Stress Scale short form.

[d]SOC: Stages of Change questionnaire.

[e]SEES: Self-Efficacy for Exercise Scale.

[f]OEES-2: Outcome Expectations for Exercise Scale.

[g]EBBS: Exercise Benefits and Barriers Scale.

[h]APS: Action Planning Scale.

[i]DRR: dementia risk reduction.

[j]DKAS-subscale D: Dementia Knowledge Assessment Scale.

[k]MCLHB-DRR: Motivation to Change Lifestyle and Health Behaviors for Dementia Risk Reduction Scale.

Statistical Analysis

All quantitative data analyses were conducted using SPSS (version 29; IBM Corp) and R (version 4.1; R Foundation). The *lme4* package was used for the mixed-effects modeling [50], and the *emmeans* package was used to calculate the effect sizes [51], while the *coin* package was used for the asymptotic marginal homogeneity test [52]. Standard assumption checks were undertaken for the linear mixed-effects analyses covering residuals, random effects, variance homogeneity, multicollinearity, influence, and missing data patterns and found no departures that would likely affect the findings. Given the single-group pre-post design, key assumptions required for valid causal mediation analysis, including the ability to rule out unmeasured confounders and separate intervention effects from natural temporal change, could not be met. Therefore, we examined pre-post changes in hypothesized mediators descriptively rather than conducting formal mediation analysis.

Descriptive statistics (means, SDs, frequencies, and proportions) were used to describe demographic and outcome measures. The Marginal Homogeneity test, an extension of the McNemar test, was used to compare the proportion of participants who met PA guidelines at baseline to those who met the PA guidelines at postintervention, with the sample restricted to those participants with both baseline and follow-up data. For continuous variables, linear mixed-effects models were used to compare changes in outcome scores from baseline to follow-up, accounting for missing data at follow-up. The models included a fixed effect of time and a random intercept for participants. The Kenward-Roger degrees of freedom approximation was used to calculate *P* values, CIs, and effect sizes. For these models, Cohen *d* effect size was calculated as the mean difference between baseline and follow-up scores divided by the square root of the sum of the intercept (participant) and residual random effect variances [53]. Cohen *d* effect sizes were

XSL•FO
RenderX

interpreted as small (d=0.2), medium (d=0.5), and large (d≥0.8). The Benjamini and Yekutieli [54] approach was used to correct for multiple testing of the primary outcomes and secondary outcomes, respectively. For these analyses, an α of .05 was used for statistical inference.

Adherence to PA Guidelines

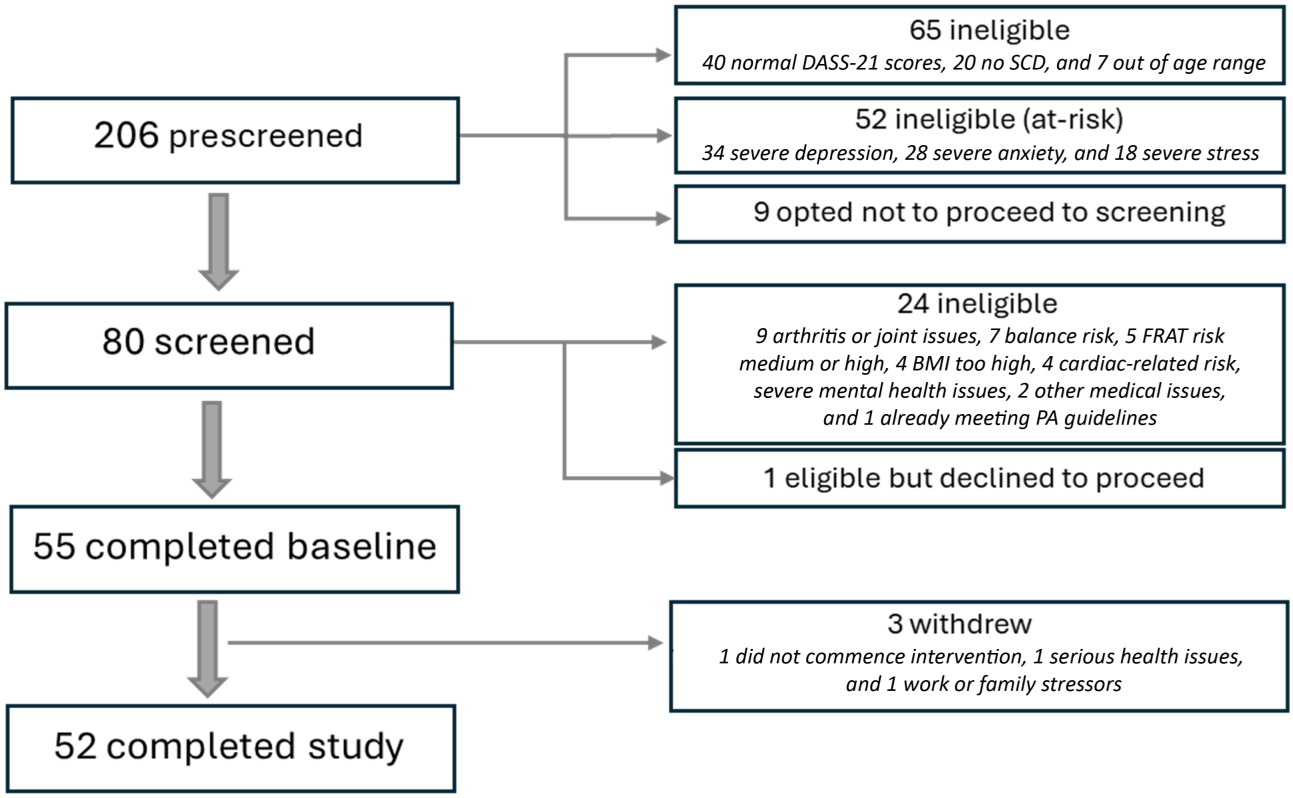
Participants' self-reported diaries collected over the 12-week intervention period were used to assess adherence to PA guidelines. No consensus exists on defining PA adherence; similar studies commonly use meeting 2/3 or more of the prescribed PA [55]. We will apply two thresholds: (1) full adherence, defined as meeting the complete guideline each week (>150 min/wk moderate-plus aerobic PA, >2 strength sessions/wk, and for older adults >1 balance session/wk); and (2) 2/3 adherence, defined as >100 minutes per week of aerobic PA, ≥1 strength session per week, and ≥1 balance session per week for older adults. Aggregate total adherence over the 12-week intervention period will use a realistic expectation of 8 or more out of a possible 12 weeks' meeting in full or with 2/3 guidelines.

Results

Participant Demographics

Figure 5 summarizes this study recruitment and screening process; 55 participants were enrolled and completed baseline assessment (46 female/9 male; mean age at screening 62.2, SD 7.6 y). Demographic characteristics are presented in Table 2.

Figure 5. Study recruitment and screening flow diagram. DASS-21: Depression Anxiety Stress Scale short form; FRAT: falls risk assessment tool; PA: physical activity; SCD: subjective cognitive decline.



XSL•FO
RenderX

Table . Participant demographic characteristics at baseline.

	Baseline (n=55)
Gender, n (%)	
Man	9 (16)
Woman	46 (84)
Age (y)	
Mean (SD)	62.2 (7.6)
Older age group (60 - 80), n (%)	43 (78)
Middle-aged group (45 - 59), n (%)	12 (22)
Country of birth, n (%)	
Australia	33 (60)
Other English-speaking (eg, New Zealand, the United Kingdom, and the United States)	10 (18)
Other	12 (22)
Speaks a language other than English at home, n (%)	
No	52 (95)
Yes	3 (5)
Education, n (%)	
Bachelor's degree or higher	43 (78)
Work status, n (%)	
Retired	16 (29)

Preliminary Analysis

We performed sensitivity analyses adding age, gender, baseline depression score, and baseline total minutes of PA as covariates; their inclusion had a negligible effect on the estimated mean difference and effect size and did not meaningfully change the CIs, indicating the main results are robust to these potential confounders, and the full results are reported in Multimedia Appendix 1.

Primary Outcomes

Uptake of Aerobic PA Guidelines

We used the more rigorous measure of aerobic PA (ie, purposeful aerobic PA which did not include everyday activities, eg, housework, gardening, etc) to gauge the uptake of aerobic PA guidelines via the provided aerobic PA programs. For changes in broader measures of aerobic PA, see Table 3. At baseline, 44% (24/55) of participants met the purposeful aerobic PA guidelines (150 min of moderate or 90 min of vigorous) aerobic PA per week. At postintervention, this figure increased by 11% to 55% (30/55) participants meeting purposeful aerobic PA guidelines (Table 4). Of those participants who did not meet full purposeful aerobic guidelines at baseline, at postintervention, 37% (11/52) of participants met these guidelines (P=.06; Table 5).

XSL•FO
RenderX

Table . Baseline and follow-up for secondary outcome measures. There were 55 participants at baseline and 52 participants at postintervention. *P* values were adjusted for multiple testing.

Secondary outcomes	Baseline, mean (SD)	Postintervention, mean (SD)	Mean differences (SE; 95% CI)	*P* value	Cohen *d* (95% CI)
CogDRisk total risk score	4.80 (4.88)	3.10 (4.60)	−1.52 (0.43; −2.39 to −0.65)	.008	−0.32 (−0.50 to −0.13)
Overall min mod/ vigorous PA[a]	380.45 (347.27)	495.29 (370.14)	114.40 (37.39; 39.37 to 189.44)	.02	0.32 (0.10 to 0.54)
Overall minutes of all PA[b]	32.55 (449.07)	936.06 (588.71)	205.25 (54.93; 95.02 to 315.48)	.005	0.39 (0.17 to 0.62)
DASS-depression[c]	11.02 (4.68)	4.92 (4.61)	−6.07 (0.68; −7.44 to −4.70)	<.001	−1.31 (−1.70 to −0.92)
DASS-anxiety	5.35 (3.98)	2.15 (3.01)	−3.15 (0.54; −4.22 to −2.07)	<.001	−0.89 (−1.23 to −0.54)
DASS-stress	14.18 (4.86)	7.69 (6.15)	−6.51 (0.84; −8.20 to −4.82)	<.001	−1.18 (−1.56 to −0.80)
Stages of change	3.73 (1.18)	4.27 (0.82)	0.54 (0.18; 0.19 to 0.89)	.02	0.53 (0.17 to 0.88)
SEES[d] total self-efficacy score	56.78 (17.71)	55.08 (18.03)	−1.88 (2.22; −6.33 to 2.56)	≥.99	−0.10 (−0.35 to 0.14)
OEES[e]-POESS[f]	4.20 (0.60)	4.40 (0.64)	0.20 (0.06; 0.33 to 0.07)	.02	0.33 (0.54 to 0.11)
OEES-NOESS[g]	4.34 (0.69)	4.49 (0.60)	−0.12 (0.08; −0.28 to 0.03)	.47	−0.19 (−0.42 to 0.05)
EBBS benefit[h] average score	3.19 (0.40)	3.23 (0.48)	0.04 (0.05; −0.06 to 0.14)	≥.99	0.09 (−0.13 to 0.30)
EBBS barrier[i] average score	1.94 (0.39)	1.85 (0.43)	−0.09 (0.05; −0.18 to 0.00)	.30	−0.22 (−0.45 to 0.01)
Action Planning Scale average score	2.88 (1.58)	3.96 (1.64)	1.07 (0.24; 0.58 to 1.56)	.001	0.66 (0.34 to 0.99)
DKAS[j] average subscale D score	0.98 (0.45)	1.07 (0.48)	0.09 (0.05; −0.02 to 0.20)	.47	0.19 (−0.04 to 0.43)
MCLHB[k] average total score	3.40 (0.39)	3.36 (0.34)	−0.05 (0.04; −0.13 to 0.03)	.92	−0.13 (−0.34 to 0.09)

[a]Min mod/ vigorous PA: minimum moderate-vigorous physical activity.

[b]PA: physical activity.

[c]DASS: Depression Anxiety Stress Scale short form.

[d]SEES: Self-Efficacy for Exercise.

[e]OEES: Modified Outcome Expectancy for Exercise Scale.

[f]POESS: Positive Outcome Expectancy for Exercise.

[g]NOESS: Negative Outcome Expectancy for Exercise.

[h]EBBS benefit: Exercise Benefits/Barriers Scale – Benefits subscale.

[i]EBBS barrier: Exercise Benefits/Barriers Scale – Barrier subscale.

[j]DKAS: Dementia Knowledge Assessment Scale.

[k]MCLHB: Motivation to Change Lifestyle and Health Behaviors for Dementia Risk Reduction.

XSL•FO
RenderX

Table . Proportion of cohort meeting aerobic and purposeful guidelines at baseline and 12 weeks.

	Baseline (n=55)	12-weeks (n=55)
Meeting full purposeful aerobic guidelines, n (%)		
No	31 (56)	22 (40)
Yes	24 (44)	30 (55)
Withdrew	0 (0)	3 (6)
Meeting full purposeful total guidelines, n (%)		
No	52 (94)	28 (51)
Yes	3 (6)	24 (44)
Withdrew	0 (0)	3 (6)

Table . Physical activity guidelines transitions from baseline to postintervention (n=52).

Baseline	Postintervention		
	No	Yes	Total
Meeting full purposeful aerobic guidelines, n (%)			
No	19 (64)	11 (37)	30 (100)
Yes	3 (14)	19 (86)	22 (100)
Meeting full purposeful total guidelines, n (%)			
No	28 (56)	22 (44)	50 (100)
Yes	0 (0)	2 (100)	2 (100)

Uptake of Full PA Guidelines

Meeting the full PA guidelines includes aerobic PA as described above, as well as 2 or more strength sessions per week, and for older adults, balance exercises as often as possible (at least once per week). At baseline, 52 of 55 (94%) participants were not meeting the full purposeful PA guidelines, and postintervention, this figure had dropped to just over half (28/55) of participants who were not meeting these guidelines (Table 4). Likewise, of those participants who did not meet full purposeful total guidelines at baseline, 44% (22/52) of participants met these guidelines at postintervention (P<.001; Table 5).

Secondary Outcomes

Acceptability, Feasibility, and Safety

Postintervention feedback was received from 94% (49/52) of intervention completers. Further, 51 of 52 (98%) participants found the program to be useful, and all 52 participants stated they would recommend such a program to a friend, demonstrating acceptability. At week 12, a total of 96% (50/52) of participants expressed an intention to continue with PA beyond the EXCEL program. Fortnightly catch-up sessions were deemed the most helpful component; all 52 participants rated this support as helpful to very helpful, with 79% (41/52) of participants rating it as very helpful. Catch-up call attendance was very high; 89% (46/52) of participants attended all 6 calls. The remaining 6 participants missed only 1 catch-up session. Most liked that the program was provided online and was completed at home (50/52, 96% of participants, and 48/52, 92% of participants, respectively). Retention rates were notably high; 52 of 55 (95%) participants completed this study, providing support for program feasibility. Zoom technology was used for 91% (189/208) of all catch-up sessions, with telephone a practical back-up option for the remaining 9% (19/208) of sessions. Technology issues were uncommon; 76% (40/52) of participants reported seldom or never encountering problems. The intervention was supported by a comprehensive monitoring system that included 44 email alerts from coaches to the medical team, 35 queries or requests from coaches to the PA team, 3 individual participant sessions via Zoom with an exercise specialist, and several case conferences, ensuring participant safety and supporting adherence.

Dementia Risk Levels

The intervention was associated with a reduction in dementia risk, with an average decrease of 1.52 (SE 0.43) points on the CogDRisk (95% CI 0.65 to 2.39, P=.008; see Table 3). This is consistent with a small to moderate effect (d=0.32, 95% CI 0.13 to 0.50).

Moderate to Vigorous Intensity PA and All PA

Moderate to vigorous intensity PA levels, measured by the CHAMPS, increased by an average of 114.40 (SE 37.39; 95% CI 39.37 to 189.44) minutes per week from baseline to postintervention. The effect size was small to moderate (d=0.32; Table 3). This included any PA performed at moderate or higher intensity, encompassing all types (eg, aerobic or strength-based PA) and any purpose, whether purposeful (undertaken with the intention of being active) or incidental or everyday activity (eg, gardening). Likewise, all PAs increased by an average of 202.25 (SE 54.93; 95% CI 95.02 to 315.48) minutes from baseline to

XSL•FO
RenderX

postintervention; the effect size was small to moderate (d=0.39; Table 3). This included any intensity (low to very high), any type (aerobic, strength, and balance), and any purpose (purposeful and incidental).

PA Guideline Adherence

Weekly adherence data is summarized in Table 6. All component guidelines adherence was highest during weeks 3 to 6 (>70%; at least 36 of the 52 participants), and all component guidelines were met by 31% (16/52) of participants for 8 or more weeks. Average weekly volume of moderate intensity aerobic PA fluctuated over the 12 weeks but did not drop below the recommended 150 minutes per week. Weekly strength session numbers were stable across the intervention and consistently exceeded the minimum of 2 sessions per week, with an average of 2.2 (SD) sessions per week. Weekly adherence to the balance guidelines was only applicable for older adults and was high throughout the intervention. Weekly numbers of balance PA sessions were relatively stable over the 12 weeks, with an average of 4.1 (SD) sessions per week, ranging from 3.8 to 4.3 sessions per week.

Table . Weekly adherence to PA[a] guideline components over the 12-week intervention.

PA component and weekly adherence	Full guideline	≥2/3 guideline
All components[b]		
Highest, n (%)	>36 (>70) – Weeks 3 - 6	41 (79) – Week 4
Lowest, n (%)	20 (39) – Week 12	28 (54) – Week 8
Mean (SD)	23 (44)	34 (65)
Aerobic PA[b]		
Highest, n (%)	33 (64) – Week 3	46 (89) – Week 4
Lowest, n (%)	23 (44) – Week 9	32 (62) – Week 9
Mean (SD)	28 (54)	37 (71)
Strength PA[b]		
Highest, n (%)	≈42 (≈80) – Weeks 1, 3 - 5	≥47 (>90) – Weeks 1, 3 - 5
Lowest, n (%)	33 (64) – Week 12	41 (79) – Week 12
Mean (SD)	38 (74)	45 (87)
Balance PA[c]		
Highest, n (%)	43 (100) – Weeks 7, 10	—[d]
Mean (SD)	40 (93)	—

[a]PA: physical activity.

[b]n=52.

[c]n=43 (balance component for older adults only).

[d]Not available.

Mental Health Symptoms

Analysis of mental health outcomes measured using the DASS-21 showed improvements across all dimensions. Effect sizes for changes in depression, anxiety, and stress scores were large at 1.31, 0.89, and 1.18, respectively (Table 3). Depression decreased by 6.1 points (SE 0.7; 95% CI 4.7 to 7.4; P<.001), anxiety decreased by 3.2 points (SE 0.5; 95% CI 2.1 to 4.2; P<.001), and stress decreased by 6.5 points (SE 0.8; 95% CI 4.8 to 8.2; P<.001). DASS-21 subscale scores can be multiplied by 2 to allow comparison with established DASS-42 severity norms. Using this interpretive framework, mean depression and stress scores shifted from the severe to mild range, and anxiety scores from the moderate to normal range.

Exercise-Related Cognitive Mechanisms (Change Readiness, Self-Efficacy, Outcome Expectations, Benefits and Barriers, and Planning)

There was strong evidence of an intervention effect on change readiness (χ^2_4=16.1, P<.003). As shown in Table 7, a total of 27 of 52 (52%) participants transitioned from a lower stage at baseline of change to a higher stage of change at postintervention, and 29 of 52 (56%) participants remained in the same stage of change at postintervention. For the latter group, however, there was a ceiling effect whereby 15 of the 29 participants who started in the highest stage of change at baseline could not transition to a higher stage of change at postintervention. Notably, all participants who started in stage 2 or stage 3 at baseline progressed to a later stage of change postintervention. There were 8 of 52 (15%) of participants who transitioned from a higher stage of change at baseline to a lower stage of change post intervention.

XSL•FO
RenderX

Table . Stages of change transition from baseline to postintervention (n=52).

Baseline	Postintervention					
	1	2	3	4	5	Total
Do not intend, n (%)	0 (0)	0 (0)	0 (0)	1 (100)	0 (0)	1 (100)
Intend to start, n (%)	0 (0)	0 (0)	1 (14)	6 (86)	0 (0)	8 (100)
Started but not regular, n (%)	0 (0)	0 (0)	0 (0)	13 (77)	4 (24)	17 (100)
Have started regular, n (%)	0 (0)	0 (0)	2 (33)	1 (17)	3 (50)	6 (100)
Currently exercise regularly, n (%)	1 (5)	1 (5)	0 (0)	4 (19)	15 (71)	21 (100)

There were no significant changes in self-efficacy or perceived barriers and benefits to exercise (Table 3). However, our data demonstrated an improvement in both planning and positive outcome expectation postintervention. There was a moderate effect of the intervention for average planning (*d*=0.66, 95% CI 0.34 to 0.99), with an average increase of 1.07 points (SE 0.24; 95% CI 0.58 to 1.56, *P*=.001) from baseline to postintervention. Furthermore, there was a small to moderate intervention effect on positive outcome expectations for exercise (*d*=0.33, 95% CI 0.11 to 0.54), with an average increase of 0.20 points (SE 0.06; 95% CI 0.07 to 0.33, *P*=.02).

DRR Cognitive Mechanisms (Dementia Knowledge and Motivation)

There was no change in dementia knowledge (mean difference=0.09, SE 0.5; 95% CI −0.20 to 0.02, *P*=.47) and attitudes and beliefs toward lifestyle adaptations for DRR (mean difference=0.05, SD, 95% CI –0.03 to 0.13, *P*=.92) from baseline to postintervention (Table 3).

Discussion

Principal Findings

These findings support our hypothesis that the EXCEL program, a 12-week digitally delivered, tailored PA intervention, could effectively facilitate middle-aged and older Australians experiencing mild to moderate mental health symptoms alongside cognitive concerns to meet PA guidelines. The program was acceptable, feasible, and safe and was associated with reductions in both dementia risk and mental health symptoms and progression through stages of change. Changes to action planning and positive outcome expectancies were observed and present potential mechanisms of action, reflecting components of the behavior change model that informed the intervention's design [14]. Given the well-documented challenges in engaging this population in sustained PA and the critical importance of PA for DRR, these results are particularly encouraging.

Comparison With Prior Work

EXCEL was highly efficacious; nearly half of the cohort (24/55, 44% participants) met PA guidelines at the end of the program, up from 6% (3/55) of participants at baseline. The increases in PA are particularly encouraging given the low proportion meeting PA guidelines at baseline (<6%), compared to the Australian population levels of 22.4% of adults and 33.4% of older adults [56]. Low baseline PA levels in our cohort fit with evidence that people with mental health disorders are significantly less likely to meet PA guidelines compared to controls [57]. This further underlines the critical importance of interventions that are tailored to support this cohort at high risk for dementia. Retention rates (95%) and positive feedback from participants (ie, 98% found it useful and 100% would recommend it to others) demonstrate that the program was well-received. High levels of acceptability might be attributed to the individually tailored nature of the intervention, which was designed based on comprehensive formative research incorporating the preferences and needs of the target demographic [15]. It is also relevant to note that EXCEL was conducted in part during the COVID-19 lockdowns, which may have increased participants' involvement due to a reduction in alternative activities and distractions. Evidence suggests that while COVID-19 had a mixed impact on research participation, research that used PA was associated with increased involvement [58].

Safety was a central focus of our design, particularly given the online delivery. While the safety, scalability, and effectiveness of digital PA interventions have been well documented in younger adults (see umbrella review by Russell et al [59]), there remains a gap in the literature regarding older adults. We used a robust system of support and monitoring that included email alerts and professional consultations that managed physical and mental health risk successfully. Our current results, taken together with data that show the majority of middle-aged and older adults are online (98% of Australians aged 65 - 74 y and 94% of those aged 75 y + [60]), suggest that age should not be seen as a barrier to accessing safe and effective online PA interventions. Another advantage suggesting future scalability was that coaches were not required to have mental health credentials or exercise physiology expertise themselves to be highly effective. The safety of our trial and the training and support of our coaches did, however, require qualified mental health and exercise physiology clinicians in an oversight capacity, which has implications for real-world implementation.

XSL•FO
RenderX

A significant reduction in dementia risk was observed over the 12 weeks, as measured by the CogDRisk, a validated online measure that assesses individualized risk factors of dementia in various population settings [42,61]. These results support a developing literature showing that PA reduces dementia risk in other high-risk groups, including those with new-onset type II diabetes [62]. Baseline dementia knowledge in our cohort was low and was not modified by our intervention, which is to be expected given that this was not an aim of the current study. This does demonstrate, however, that it is not essential to increase dementia knowledge to reduce dementia risk, which may be particularly pertinent for individuals with cognitive difficulties. Nonetheless, integrating additional psychological or social support structure components could further enhance the intervention's impact on self-efficacy and sustained behavior change.

We evaluated whether the intervention impacted mental health symptoms and found a significant and clinically meaningful improvement across all DASS-21 dimensions. After 12 weeks, our cohort's anxiety levels had returned to within the normal range, while depression and stress symptoms were now mild. This is a meaningful change from the extremely severe (stress), severe (depression), and moderate (anxiety) average symptom scores observed 12 weeks earlier, at baseline. Our findings add to the established literature showing the important positive effect that exercise can play in improving mood; a finding reflected in clinical practice guidelines in the United States, United Kingdom, and Australia [63-65]. A recent meta-analysis by Noetel et al [66] showed exercise to be equally effective for people with and without comorbidities, and with different baseline levels of depression. Their findings reflected over 200 studies [66] and suggest that walking and strength training are among the most effective exercises for depression reduction, in line with PA guidelines that underlie EXCEL. While evidence for the anxiety-reducing effect of PA is currently less consistent (see Chong et al [67] for a meta-analysis), we were encouraged to see that PA reduced anxiety levels in our cohort.

Potential Mechanisms of Action

EXCEL was designed to specifically target proposed mechanisms of action for optimizing engagement in this cohort, including emotional regulation, self-efficacy in existing skills, and the capacities to act on intentions despite barriers, and we measured these potential cognitive mechanisms pre- and postintervention. Our findings showed a small to moderate, but significant, increase in positive outcome expectations for exercise. Positive outcome expectancy refers to an individual's belief that PA will lead to certain (positive) outcomes and is 1 of 2 key constructs in social cognitive theory, alongside self-efficacy. Our findings are consistent with previous research showing that (1) positive, rather than negative, outcome expectancy is associated with PA levels, and (2) there is an apparent age-related effect whereby outcome expectancy is more predictive of PA in older adults [68]. Positive outcome expectancies have previously been observed to facilitate the formation of PA intentions [68] and are associated with progression through the stages of change, as seen in this study. Considered within the COM-B framework, positive outcome expectancies may enhance motivation and support the development of strong intentions to engage in PA.

However, transitioning from intention to behavior often relies on effective action planning—that is, setting concrete, detailed plans that outline the when, where, and how of intended PAs [69]. Our findings support this notion. The mental stimulations involved in planning have also been shown to have a self-regulatory effect that may help individuals connect their behavioral responses to future scenarios and account for potential barriers [70]. In this context, our intervention may have supported participants in mentally rehearsing the steps needed to translate PA intentions into action, reducing the intention-behavior gap. This aligns with our findings on the stages of change measure: at baseline, participants primarily scored in the "preparation" stage, characterized by intentions to begin regular exercise or making sporadic attempts to be active. Following the intervention, averages shifted to the "action" stage, reflecting active and consistent engagement in structured exercise. These results suggest that aging adults with cognitive difficulties related to subjective or MCI and co-occurring mental health symptoms may particularly benefit from pragmatic, structured action planning that bridges the gap between intention and behavior.

Limitations and Future Directions

While our results are promising, they should be interpreted with caution due to the pilot nature of this study, modest sample size, and pre-post design. This study's design prevented a valid causal mediation analysis, meaning that changes in hypothesized mediators can only be described rather than inferred as causal. Further research using a larger sample and a randomized controlled trial design with mediation analyses is needed to determine whether the proposed mechanisms contribute causally to the observed outcomes. Sensitivity analyses indicated that age, gender, baseline depression score, and overall minutes of PA did not confound the findings. However, technological or digital literacy was not measured, and engagement may have been higher during COVID-19 lockdowns; these unmeasured confounders may limit generalizability to nonpandemic conditions. Despite these limitations, the findings exemplify the utility of tailored interventions and the value of a behavior-change framework. It is recommended that future studies include larger cohorts and longer follow-up periods to replicate findings and evaluate the durability of effects. Given the complexity of factors influencing PA adherence and mental health among aging adults, multicomponent interventions that address both physical and cognitive health holistically are recommended.

Conclusions

This pilot study demonstrated EXCEL to be effective, acceptable, feasible, and safe. The tailored, PA-delivered PA intervention supported middle-aged and older adults experiencing mild to moderate mental health symptoms alongside cognitive concerns to meet PA guidelines. Important secondary gains were a reduction of dementia risk and clinically meaningful improvements in mental health symptoms, emphasizing the advantages of boosting PA in high-risk groups. Improvements in action planning and outcome expectancy may

XSL•FO
RenderX

have contributed to bridging the PA intention-behavior gap, in line with the COM-B model underlying the intervention design. While findings require confirmation in larger controlled trials, the results suggest that EXCEL offers a scalable, evidence-informed model for supporting dementia-risk reduction in adults who face well-documented barriers to engaging in sustained PA.

Acknowledgments

The authors would like to thank this study's participants of both phase 1 and phase 2 of the EXCEL study, Dr Francine Moss and Dr Brett Coulson for their contributions to earlier versions of the EXCEL design and the StepUp for Dementia Research service for their assistance with participant recruitment. Study data were collected and managed using REDCap (Research Electronic Data Capture) tools hosted at The University of Melbourne. Figures 2 and 3 include images licensed by the authors via Shutterstock; these images are not covered by the paper’s Creative Commons Attribution (CC BY) license, and reuse requires permission from the copyright holder.

Funding

This study was funded by the National Health and Medical Research Council (NHMRC) through a Medical Research Future Fund (MRFF) grant. The funder had no involvement in this study's design, data collection, analysis, interpretation, or the writing of this paper.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to ethics requirements, but are available from the first (corresponding) author on reasonable request.

Authors' Contributions

Conceptualization: NTL (lead), KAE (supporting), EC, TWHC, KLC, VJP, KJA
Data curation: KAE, RM, RL, JS, AW
Formal analysis: SGM (lead), KAE, JS (supporting)
Funding acquisition: NTL
Investigation: KAE, RM, RL, EC, JS, AW, TWHC, TR, NTL
Methodology: NTL, KAE, EC, KLC, JS
Project administration: KAE (lead), RM (supporting), JS, RL, AW
Resources: JS (lead), KLC, RM, RL, EC
Supervision: NTL, KAE
Visualization: SGM, KAE, JS, RL
Writing – original draft: KAE (lead), RL, EC (supporting)
Writing – review & editing: NTL (lead), all authors

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results of sensitivity analysis examining age, gender, baseline physical activity, and baseline depression severity as potential confounders of primary and secondary outcome measures.
[PDF File, 23 KB - aging_v9i1e80040_app1.pdf]

References

1. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing commission. Lancet 2024 Aug 10;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
2. Chong TWH, Curran E, Ellis KA, et al. Physical activity for older Australians with mild cognitive impairment or subjective cognitive decline - a narrative review to support guideline development. J Sci Med Sport 2020 Oct;23(10):913-920. [doi: 10.1016/j.jsams.2020.03.003] [Medline: 32389610]
3. Iso-Markku P, Kujala UM, Knittle K, Polet J, Vuoksimaa E, Waller K. Physical activity as a protective factor for dementia and Alzheimer’s disease: systematic review, meta-analysis and quality assessment of cohort and case-control studies. Br J Sports Med 2022 Jun;56(12):701-709. [doi: 10.1136/bjsports-2021-104981] [Medline: 35301183]
4. Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds CF. Late-life depression and risk of vascular dementia and Alzheimer’s disease: systematic review and meta-analysis of community-based cohort studies. Br J Psychiatry 2013 May;202(5):329-335. [doi: 10.1192/bjp.bp.112.118307] [Medline: 23637108]

XSL•FO
RenderX

5. Santabárbara J, Lipnicki DM, Olaya B, et al. Does anxiety increase the risk of all-cause dementia? An updated meta-analysis of prospective cohort studies. JCM 2020;9(6):1791. [doi: 10.3390/jcm9061791]
6. Firth J, Siddiqi N, Koyanagi A, et al. The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. Lancet Psychiatry 2019 Aug;6(8):675-712. [doi: 10.1016/S2215-0366(19)30132-4] [Medline: 31324560]
7. Wang SM, Han KD, Kim NY, et al. Late-life depression, subjective cognitive decline, and their additive risk in incidence of dementia: a nationwide longitudinal study. PLoS One 2021;16(7):e0254639. [doi: 10.1371/journal.pone.0254639] [Medline: 34260630]
8. Liew TM. Depression, subjective cognitive decline, and the risk of neurocognitive disorders. Alz Res Therapy 2019 Dec;11(1):70. [doi: 10.1186/s13195-019-0527-7]
9. Liew TM. Subjective cognitive decline, anxiety symptoms, and the risk of mild cognitive impairment and dementia. Alz Res Therapy 2020 Dec;12(1):32917264. [doi: 10.1186/s13195-020-00673-8]
10. Chong TW, Macpherson H, Schaumberg MA, et al. Dementia prevention: the time to act is now. Med J Aust 2021 Apr;214(7):302-304. [doi: 10.5694/mja2.50972] [Medline: 33715184]
11. Risk reduction of cognitive decline and dementia: WHO guidelines. : World Health Organisation; 2019 URL: https://www.who.int/publications/i/item/risk-reduction-of-cognitive-decline-and-dementia [accessed 2026-02-19]
12. Lautenschlager NL, Cox K, Hill KD, Pond D, Ellis KA, Dow B, et al. Physical Activity Guidelines for Older Australians with Mild Cognitive Impairment or Subjective Cognitive Decline: Melbourne, Australia: Dementia Collaborative Research Centres; 2018.
13. Brown W, Moorhead G, Marshall A. Choose health: be active: a physical activity guide for older australians. : Commonwealth of Australia and the Repatriation Commission; 2005 URL: https://www.health.gov.au/resources/publications/choose-health-be-active-a-physical-guide-for-older-australians?language=en [accessed 2026-02-19]
14. Physical activity and sedentary behaviour guidelines - adults (18 to 64 years) - fact sheet. : Australian Government Department of Health; 2014 URL: https://www.health.gov.au/resources/publications/physical-activity-and-sedentary-behaviour-guidelines-adults-18-to-64-years-fact-sheet?language=en [accessed 2026-02-19]
15. Curran E, Palmer VJ, Ellis KA, et al. Physical activity for cognitive health: a model for intervention design for people experiencing cognitive concerns and symptoms of depression or anxiety. JAD 2023;94(2):781-799. [doi: 10.3233/JAD-221216]
16. Piercy KL, Bevington F, Vaux-Bjerke A, Hilfiker SW, Arayasirikul S, Barnett EY. Understanding contemplators' knowledge and awareness of the physical activity guidelines. J Phys Act Health 2020 Apr 1;17(4):404-411. [doi: 10.1123/jpah.2019-0393] [Medline: 32087601]
17. Knox ECL, Esliger DW, Biddle SJH, Sherar LB. Lack of knowledge of physical activity guidelines: can physical activity promotion campaigns do better? BMJ Open 2013 Dec 5;3(12):e003633. [doi: 10.1136/bmjopen-2013-003633] [Medline: 24319271]
18. Coley N, Ngandu T, Lehtisalo J, et al. Adherence to multidomain interventions for dementia prevention: data from the FINGER and MAPT trials. Alzheimer's Dementia 2019 Jun;15(6):729-741. [doi: 10.1016/j.jalz.2019.03.005]
19. Coley N, Coniasse-Brioude D, Igier V, et al. Disparities in the participation and adherence of older adults in lifestyle-based multidomain dementia prevention and the motivational role of perceived disease risk and intervention benefits: an observational ancillary study to a randomised controlled trial. Alz Res Therapy 2021 Dec;13(1):157. [doi: 10.1186/s13195-021-00904-6]
20. Mason JE, Faller YN, LeBouthillier DM, Asmundson GJG. Exercise anxiety: a qualitative analysis of the barriers, facilitators, and psychological processes underlying exercise participation for people with anxiety-related disorders. Ment Health Phys Act 2019 Mar;16:128-139. [doi: 10.1016/j.mhpa.2018.11.003]
21. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655. [doi: 10.1136/bmj.a1655] [Medline: 18824488]
22. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
23. Salisbury C, O'Cathain A, Edwards L, et al. Effectiveness of an integrated telehealth service for patients with depression: a pragmatic randomised controlled trial of a complex intervention. Lancet Psychiatry 2016 Jun;3(6):515-525. [doi: 10.1016/S2215-0366(16)00083-3] [Medline: 27132075]
24. Communications and media in Australia: the digital lives of older Australians. : Commonwealth of Australia (Australian Communications and Media Authority); 2021 May URL: https://www.acma.gov.au/sites/default/files/2021-05/The%20digital%20lives%20of%20older%20Australians.pdf [accessed 2026-02-19]
25. Cox KL, Cyarto EV, Ellis KA, et al. A randomized controlled trial of adherence to a 24-month home-based physical activity program and the health benefits for older adults at risk of Alzheimer's disease: the AIBL active-study. J Alzheimer's Dis 2019 Aug 13;70(s1):S187-S205. [doi: 10.3233/JAD-180521]
26. Robb CE, de Jager CA, Ahmadi-Abhari S, et al. Associations of social isolation with anxiety and depression during the early COVID-19 pandemic: a survey of older adults in London, UK. Front Psychiatry 2020;11:591120. [doi: 10.3389/fpsyt.2020.591120] [Medline: 33132942]

XSL•FO
RenderX

27. Curran E, Nalder L, Koye D, et al. COVID-19 and mental health: Impact on symptom burden in older people living with mental illness in residential aged care. Australas J Ageing 2022 Dec;41(4):522-529. [doi: 10.1111/ajag.13042] [Medline: 35129267]
28. Mann F, Wang J, Pearce E, et al. Loneliness and the onset of new mental health problems in the general population. Soc Psychiatry Psychiatr Epidemiol 2022 Nov;57(11):2161-2178. [doi: 10.1007/s00127-022-02261-7]
29. Lovibond SH, Lovibond PF. Manual for the Depression Anxiety Stress Scales, 2nd edition: Psychology Foundation; 1995.
30. Knopman DS, Roberts RO, Geda YE, et al. Validation of the telephone interview for cognitive status-modified in subjects with normal cognition, mild cognitive impairment, or dementia. Neuroepidemiology 2010;34(1):34-42. [doi: 10.1159/000255464] [Medline: 19893327]
31. Stapleton C, Hough P, Oldmeadow L, Bull K, Hill K, Greenwood K. Four-item fall risk screening tool for subacute and residential aged care: the first step in fall prevention. Australas J Ageing 2009 Sep;28(3):139-143. [doi: 10.1111/j.1741-6612.2009.00375.x] [Medline: 19845654]
32. Adult pre-exercise screening system (APSS) user guide 2019. : Exercise and Sports Science Australia; 2019 URL: https://ausactive.org.au/wp-content/uploads/2022/02/Adult-Pre-Exercise-Screening-System-Guide.pdf [accessed 2026-02-19]
33. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Can J Sport Sci 1992 Dec;17(4):338-345. [Medline: 1330274]
34. Wang S, Shen J, Koh WP, Yuan JM. Comparison of race- and ethnicity-specific BMI cutoffs for categorizing obesity severity: a multicountry prospective cohort study. Obesity (Silver Spring) 2024;32(10):1958-1966. [doi: 10.1002/oby.24129]
35. Jeon YH, Shin M, Smith A, et al. Early implementation and evaluation of StepUp for dementia research: an australia-wide dementia research participation and public engagement platform. IJERPH 2021;18(21):11353. [doi: 10.3390/ijerph182111353]
36. Aberson CL. Applied Power Analysis for the Behavioral Sciences, 2nd edition: New York: Routledge; 2019.
37. Schmidt JA, Gruman C, King MB, Wolfson LI. Attrition in an exercise intervention: a comparison of early and later dropouts. J Am Geriatr Soc 2000 Aug;48(8):952-960. [doi: 10.1111/j.1532-5415.2000.tb06894.x] [Medline: 10968301]
38. Cox KL, Cyarto EV, Etherton-Beer C, et al. A randomized controlled trial of physical activity with individual goal-setting and volunteer mentors to overcome sedentary lifestyle in older adults at risk of cognitive decline: the INDIGO trial protocol. BMC Geriatr 2017 Sep 13;17(1):215. [doi: 10.1186/s12877-017-0617-y] [Medline: 28903738]
39. Yang XJ, Hill K, Moore K, et al. Effectiveness of a targeted exercise intervention in reversing older people's mild balance dysfunction: a randomized controlled trial. Phys Ther 2012 Jan;92(1):24-37. [doi: 10.2522/ptj.20100289] [Medline: 21979272]
40. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
41. Stewart AL, Mills KM, King AC, Haskell WL, Gillis D, Ritter PL. CHAMPS physical activity questionnaire for older adults: outcomes for interventions. Med Sci Sports Exercise 2001;33(7):1126-1141. [doi: 10.1097/00005768-200107000-00010]
42. Anstey KJ, Kootar S, Huque MH, Eramudugolla R, Peters R. Development of the CogDrisk tool to assess risk factors for dementia. Alzheimer's Dementia 2022 Jan;14(1):e12336. [doi: 10.1002/dad2.12336]
43. Marcus BH, Selby VC, Niaura RS, Rossi JS. Self-efficacy and the stages of exercise behavior change. Res Q Exerc Sport 1992 Mar;63(1):60-66. [doi: 10.1080/02701367.1992.10607557] [Medline: 1574662]
44. Resnick B, Jenkins LS. Testing the reliability and validity of the self-efficacy for exercise scale. Nurs Res 2000;49(3):154-159. [doi: 10.1097/00006199-200005000-00007] [Medline: 10882320]
45. Resnick B. Reliability and validity of the Outcome Expectations for Exercise Scale-2. J Aging Phys Act 2005 Oct;13(4):382-394. [doi: 10.1123/japa.13.4.382] [Medline: 16301750]
46. Sechrist KR, Walker SN, Pender NJ. Development and psychometric evaluation of the exercise benefits/barriers scale. Res Nurs Health 1987 Dec;10(6):357-365. [doi: 10.1002/nur.4770100603] [Medline: 3423307]
47. Sniehotta FF, Schwarzer R, Scholz U, Schüz B. Action planning and coping planning for long-term lifestyle change: theory and assessment. Eur J Soc Psychol 2005 Jul;35(4):565-576. [doi: 10.1002/ejsp.258]
48. Annear MJ, Toye CM, Eccleston CE, et al. Dementia knowledge assessment scale: development and preliminary psychometric properties. J Am Geriatr Soc 2015 Nov;63(11):2375-2381. [doi: 10.1111/jgs.13707] [Medline: 26503020]
49. Kim S, Sargent-Cox K, Cherbuin N, Anstey KJ. Development of the motivation to change lifestyle and health behaviours for dementia risk reduction scale. Dementia Geriatr Cogn Disord Extra 2014 May 22;4(2):172-183. [doi: 10.1159/000362228]
50. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
51. Lenth R. Package "emmeans": estimated marginal means, aka least-squares means, version 1.10.3. : R Documentation; 2024 URL: https://search.r-project.org/CRAN/refmans/emmeans/html/emmeans-package.html [accessed 2026-02-19]
52. Hothorn T, Hornik K, Wiel MAVD, Zeileis A. Implementing a class of permutation tests: the coin package. J Stat Soft 2008;28(8):1-23. [doi: 10.18637/jss.v028.i08]
53. Westfall J, Kenny DA, Judd CM. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. J Exp Psychol Gen 2014 Oct;143(5):2020-2045. [doi: 10.1037/xge0000014] [Medline: 25111580]

XSL•FO
RenderX

54. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Statist 2001;29(4):1165-1188. [doi: 10.1214/aos/1013699998]
55. King AC, Kiernan M, Oman RF, Kraemer HC, et al. Can we identify who will adhere to long-term physical activity? Signal detection methodology as a potential aid to clinical decision making. Health Psychol 1997;16(4):380-389. [doi: 10.1037//0278-6133.16.4.380]
56. Australian Bureau of Statistics. : National Health Survey; 2022 URL: https://www.abs.gov.au/statistics/health/health-conditions-and-risks/national-health-survey/latest-release [accessed 2026-02-19]
57. Vancampfort D, Firth J, Schuch FB, et al. Sedentary behavior and physical activity levels in people with schizophrenia, bipolar disorder and major depressive disorder: a global systematic review and meta-analysis. World Psychiatry 2017 Oct;16(3):308-315. [doi: 10.1002/wps.20458] [Medline: 28941119]
58. Abdulhussein D, Yap TE, Manzar H, Miodragovic S, Cordeiro F. Factors impacting participation in research during the COVID-19 pandemic: results from a survey of patients in the ophthalmology outpatient department. Trials 2022 Sep 30;23(1):823. [doi: 10.1186/s13063-022-06748-1] [Medline: 36175935]
59. Russell E, Kirk A, Dunlop MD, Hodgson W, Patience M, Egan K. Digital physical activity and sedentary behavior interventions for community-living adults: umbrella review. J Med Internet Res 2025 Mar 18;27:e66294. [doi: 10.2196/66294] [Medline: 40100250]
60. How we use the internet: consumers and communication services in australia 2022–23. : Australian Communication and Media Authority (ACMA); 2023 URL: https://www.acma.gov.au/publications/2026-02/report/communications-and-media-australia-how-we-use-internet [accessed 2026-02-25]
61. Kootar S, Huque MH, Eramudugolla R, et al. Validation of the CogDrisk instrument as predictive of dementia in four general community-dwelling populations. J Prev Alzheimers Dis 2023 Sep;10(3):478-487. [doi: 10.14283/jpad.2023.38]
62. Yoo JE, Han K, Kim B, et al. Changes in physical activity and the risk of dementia in patients with new-onset type 2 diabetes: a nationwide cohort study. Diabetes Care 2022 May 1;45(5):1091-1098. [doi: 10.2337/dc21-1597]
63. Malhi GS, Bell E, Bassett D, et al. The 2020 Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. Aust N Z J Psychiatry 2021 Jan;55(1):7-117. [doi: 10.1177/0004867420979353]
64. Gelenberg AJ, Freeman MP, Markowitz JC, Rosenbaum JF, Thase ME, Trivedi MH, et al. Practice guideline for the treatment of patients with major depressive disorder. : American Psychiatric Association; 2010 URL: https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf [accessed 2026-02-19]
65. Depression in adults: treatment and management. : National Institute for Health and Care Excellence; 2022 URL: https://www.nice.org.uk/guidance/ng222 [accessed 2026-02-19]
66. Noetel M, Sanders T, Gallardo-Gómez D, et al. Effect of exercise for depression: systematic review and network meta-analysis of randomised controlled trials. BMJ 2024 Feb 14;384:e075847. [doi: 10.1136/bmj-2023-075847] [Medline: 38355154]
67. Chong TWH, Kootar S, Wilding H, et al. Exercise interventions to reduce anxiety in mid-life and late-life anxiety disorders and subthreshold anxiety disorder: a systematic review. Ther Adv Psychopharmacol 2022;12:20451253221104958. [doi: 10.1177/20451253221104958] [Medline: 35833057]
68. Williams DM, Anderson ES, Winett RA. A review of the outcome expectancy construct in physical activity research. Ann Behav Med 2005 Feb;29(1):70-79. [doi: 10.1207/s15324796abm2901_10] [Medline: 15677303]
69. Wee ZQC, Dillon D. Increasing physical exercise through action and coping planning. Int J Environ Res Public Health 2022 Mar 24;19(7):3883. [doi: 10.3390/ijerph19073883] [Medline: 35409564]
70. Gollwitzer PM. Implementation intentions: strong effects of simple plans. American Psychologist 1999;54(7):493-503. [doi: 10.1037/0003-066X.54.7.493]

Abbreviations

COM-B: Capability, Opportunity, and Motivation Behavior
DASS-21: Depression Anxiety Stress Scale short form
DRR: dementia risk reduction
EXCEL: Exercise for Cognitive Health
INDIGO: Individual Goal-Setting and Volunteer Mentors
MCI: mild cognitive impairment
PA: physical activity
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

Edited by M Antoniou; submitted 03.Jul.2025; peer-reviewed by M Irfan, R Stemplewski; revised version received 20.Dec.2025; accepted 03.Feb.2026; published 06.Mar.2026.

Please cite as:
Ellis KA, Lai R, Curran E, Southam J, Moorhead R, Cox KL, Mancuso SG, Westphal A, Chong TWH, Rego T, Palmer VJ, Anstey KJ, Lautenschlager NT
Tailored Online Physical Activity Coaching for Middle-Aged and Older Adults With Cognitive and Mental Health Concerns: Single-Arm Pre-Post Intervention Study
JMIR Aging 2026;9:e80040
URL: https://aging.jmir.org/2026/1/e80040
doi:10.2196/80040

© Kathryn A Ellis, Rhoda Lai, Eleanor Curran, Jennifer Southam, Rebecca Moorhead, Kay L Cox, Serafino G Mancuso, Alissa Westphal, Terence W H Chong, Thomas Rego, Victoria J Palmer, Kaarin J Anstey, Nicola T Lautenschlager. Originally published in JMIR Aging (https://aging.jmir.org), 6.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Dose-Related Effects of Different Tai Chi Styles Versus Traditional Community-Based Exercises on Cardiometabolic Health and Physical Function in Middle-Aged and Older Adults: Randomized Controlled Trial

Jiadong Qiu[1], PhD; Jian Wang[2], MM; Xiongying Song[2], BM; Wanyu Shu[1], PhD; Sung Min Kim[1,3,4,5], PhD

[1]Department of Sports Science, Hanyang University, Seoul, Republic of Korea

[2]Department of Orthopedics, Beijing Fengtai Hospital, Beijing, China

[3]Department of Active Aging Industry, Hanyang University, Seoul, Republic of Korea

[4]BK21 FOUR Human-Tech Convergence Program, Hanyang University, Seoul, Republic of Korea

[5]Center for Artificial Intelligence Muscle, Hanyang University, Seoul, Republic of Korea

Corresponding Author:
Sung Min Kim, PhD
Department of Sports Science
Hanyang University
222, Seong-dong, Seong-gu, Seoul
Seoul, 04762
Republic of Korea
Phone: 82 2 2220 2880
Email: minarthur@hanyang.ac.kr

Abstract

Background: Age-related declines in metabolic, cardiovascular, and physical function contribute to reduced quality of life in older adults. Although structured exercise is central to healthy aging, the optimal modality remains unclear. Community-based exercise programs in China are heterogeneous, and their comparative effects on health outcomes and cardiovascular safety have not been systematically evaluated.

Objective: We aimed to compare the effects of 4 common community-based exercise modalities on cardiometabolic health, physical function, and quality of life in middle-aged and older adults, and to assess their relative efficacy and safety across intervention exposure.

Methods: This single-blind randomized controlled trial included 113 middle-aged and older adults (mean age 62.3, SD 4.25 years). Participants were assigned to one of the 5 groups: 12-form Chen-style Tai Chi (CTC12), 24-form Tai Chi, square dance, walking, or a control group. The 12-week intervention comprised 2 supervised sessions per week, each lasting 90 minutes. Pre- and postintervention assessments included blood pressure, lipid profiles, fasting glucose, interleukin-6 levels, Short Physical Performance Battery (SPPB) scores, activities of daily living, and the World Health Organization Quality of Life (WHOQOL) index.

Results: The CTC12 and square dance groups showed significant improvements in several outcomes. In the CTC12 group, triglyceride and low-density lipoprotein cholesterol levels significantly decreased (P=.008 and P=.002, respectively), whereas SPPB and WHOQOL scores significantly improved (P=.02 and P=.002, respectively). In the square dance group, total cholesterol, triglyceride, and low-density lipoprotein cholesterol levels significantly decreased (P=.01, P=.002, and P=.002, respectively), whereas SPPB, activities of daily living, and WHOQOL scores significantly improved (P=.04, P=.04, and P=.002, respectively). The walking group showed a significant improvement only in WHOQOL scores (P=.02). No significant changes were observed in the control group for any outcome.

Conclusions: CTC12 and square dance were associated with significant improvements in cardiometabolic, physical, and psychological health outcomes in middle-aged and older adults. CTC12 demonstrated broader multidimensional benefits, potentially reflecting the integration of physical, respiratory, and cognitive components. These modalities may represent safe and scalable strategies for promoting healthy aging in community settings.

XSL•FO
RenderX

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400092473; https://www.chictr.org.cn/showprojEN.html?proj=249330

(JMIR Aging 2026;9:e80125) doi:10.2196/80125

KEYWORDS

Tai Chi; cardiometabolic health; functional rehabilitation; community-dwelling population; middle-aged and older adults

Introduction

The global population is aging at an unprecedented rate, with the number of individuals aged >65 years worldwide projected to increase from 703 million in 2019 to 1.5 billion by 2050 [1,2]. As life expectancy rises, older adults are increasingly susceptible to chronic conditions and their associated health burdens [3], a shift that has contributed to higher rates of chronic disease, poor health status, and functional decline among middle-aged and older populations [4,5]. Multimorbidity further compounds this burden, affecting approximately 46% of older adults worldwide [6-8], thereby underscoring the need to prioritize preventive strategies. Among these, exercise remains a key nonpharmacological intervention with well-established benefits for cardiometabolic health in aging populations [9,10].

Advancing age is associated with an increased risk of cardiovascular diseases, including hypertension and hyperlipidemia, as well as metabolic disorders, such as metabolic syndrome and diabetes mellitus (DM) [11,12]. Hypertension, affecting approximately 30% of the population and disproportionately prevalent in older adults, is a major contributor to global mortality [13-15]. Elevated systolic blood pressure (SBP) and diastolic blood pressure (DBP) are well-established predictors of cardiovascular disease and all-cause mortality [16], while pulse pressure (PP), reflecting the interplay between SBP and DBP, has gained recognition as an important indicator of vascular aging and mortality risk [17]. Dyslipidemia—characterized by elevated total cholesterol (TC) and low-density lipoprotein cholesterol (LDL-C), alongside reduced high-density lipoprotein cholesterol (HDL-C)—is likewise strongly associated with adverse outcomes in older populations [18]. Additionally, composite indices integrating glucose and lipid metabolism, including fasting plasma glucose (FPG) and the triglyceride/HDL-C ratio, have emerged as important predictors of cardiometabolic risk [19].

DM, a chronic metabolic noncommunicable disease characterized by persistent hyperglycemia, remains a significant contributor to global mortality [20-22], with diagnosis defined by FPG levels≥7.0 mmol/L (126 mg/dL) [23]. In 2015, approximately 415 million adults aged 20-79 years had diabetes, and this number is projected to increase by an additional 200 million by 2040 [20], placing a growing strain on health care systems worldwide, particularly among older populations [12]. In parallel, age-related chronic low-grade inflammation, commonly termed “inflammaging” [24], is characterized by maintained elevations in proinflammatory markers, such as interleukin-6 (IL-6), and is closely linked to cardiovascular disease, diabetes, and metabolic syndrome [25,26]. Accordingly, IL-6 was used in this study as a biomarker to capture the systemic physiological effects of exercise. Beyond physical health, older adults face an increased risk of mental health challenges [27,28], necessitating a comprehensive evaluation of both functional and psychological domains. Thus, we used validated instruments, including the Short Physical Performance Battery (SPPB), activities of daily living (ADLs), and the World Health Organization Quality of Life (WHOQOL) index to assess multidimensional outcomes [29-31]. Midlife represents a critical period during which risk factors for functional decline emerge, while also offering substantial potential for recovery [32]. Thus, we focused on middle-aged and older adults to examine the role of exercise in disease prevention and functional rehabilitation.

Tai Chi, originating from traditional Chinese medicine and guided by movement-based theoretical frameworks, integrates principles, such as yin-yang philosophy and meridian theory [33]. It has evolved into a comprehensive practice promoting both physical and psychological well-being. Empirical evidence indicates that Tai Chi confers benefits across a range of conditions, including hypertension, hyperlipidemia, DM, metabolic syndrome, inflammation, impaired physical function, limitations in ADLs, and psychological disorders [34-36]. However, most interventions have relied on the simplified 24-form Tai Chi (TC24), derived from Yang-style Tai Chi [37], which emphasizes slow, low-intensity movements and simplified structures that facilitate safety and widespread dissemination.

Despite these advantages, its low intensity and repetitive movements may limit energy expenditure and neuromuscular engagement, potentially constraining its effectiveness in improving cardiometabolic health and higher-level physical function [38,39].

These limitations highlight the necessity for more dynamic and integrative exercise modalities [40]. Chen-style Tai Chi incorporates both slow, controlled movements and rapid, forceful actions, including stomping and explosive movements, thereby enhancing neuromuscular activation and energy metabolism while maintaining safety and fluidity of motion [41,42]. The 12-form Chen-style Tai Chi (CTC12) preserves the core elements of traditional Chen-style practice while integrating 12-meridian regulation theory from traditional Chinese exercise, resulting in a rhythmically diverse and functionally comprehensive intervention. CTC12 is hypothesized to provide greater benefits than TC24 in improving neuromuscular function, metabolic regulation, and multidimensional health outcomes in middle-aged and older adults [43].

In parallel, square dance represents a widely adopted form of aerobic exercise among middle-aged and older populations, offering both physical and psychosocial benefits in community settings [44]. Similarly, walking, another accessible and low-risk group-based activity, has gained popularity. Both modalities

XSL•FO
RenderX

are associated with lower cardiovascular demands and risk of injury, and have been proven efficient in improving blood pressure, lipid metabolism, glucose regulation, and psychological well-being, including reductions in symptoms of depression and anxiety [33,45].

Despite evidence supporting the benefits of Tai Chi and aerobic exercise for cardiovascular and metabolic health among older adults, several gaps remain. First, most studies have focused on TC24, with limited investigation of more dynamic forms, such as Chen-style Tai Chi, and variability in intervention characteristics—including duration, intensity, and form—has contributed to heterogeneity in reported outcomes, underscoring the need for more standardized approaches [46]. Second, previous research has often examined single health outcomes, with limited evaluation of multidimensional physiological effects, particularly inflammatory markers [47]. Third, comparative studies of community-based aerobic activities, including Tai Chi, square dance, and walking, remain limited, and direct comparisons with structured mind-body interventions are scarce. Addressing these gaps requires comprehensive, well-designed randomized controlled trials [48,49].

Accordingly, this study was conducted to systematically evaluate the effects of CTC12, TC24, square dance, and walking on cardiometabolic, functional, and psychological outcomes in community-dwelling middle-aged and older adults, while assessing the safety and dose-related efficacy of each modality. Moreover, the study sought to provide evidence to inform the development of safe, effective, and sustainable individualized exercise strategies for aging populations.

Methods

Participants

In this study, we enrolled community-dwelling middle-aged and older adults who were capable of participating in physical exercise. Participants were recruited from local community centers and health clinics. Of 155 individuals screened, 120 were enrolled (mean age 63.0, SD 4.5 years; 97/120, 80.8% female). Eligibility criteria included age 50-70 years, absence of clinically diagnosed cardiovascular, musculoskeletal, metabolic, or cognitive disorders, and no regular participation in structured exercise within the preceding 6 months (defined as <1 session/week). Exclusion criteria comprised current use of medications affecting blood pressure, lipid metabolism, or glucose regulation (eg, antihypertensive agents, statins, or antidiabetic drugs), as well as current smoking or smoking cessation within the past 6 months.

Study Design

This study was a single-blind randomized controlled trial conducted among community-dwelling middle-aged and older adults. Following baseline assessments, participants were randomly assigned in a 1:1:1:1:1 ratio to one of the 5 groups—CTC12, TC24, square dance, walking, or a control group. Randomization was performed using a computer-generated sequence prepared by an independent researcher, with allocation concealed in sequentially numbered, opaque, sealed envelopes, which were opened after baseline assessments. Outcome assessors and laboratory personnel were blinded to group allocation, and participants were informed only of their assigned intervention. Each participant received a unique identification code, and all assessments were conducted according to standardized procedures by trained evaluators. No sex-based restrictions were applied, and all interventions and assessments were implemented uniformly across sexes. Before the intervention, participants completed a 1-week familiarization period to facilitate adaptation, enhance adherence, and ensure correct execution of the assigned exercises.

Ethical Considerations

The study protocol was approved by the Ethics Committee of Beijing Fengtai Hospital (approval P2024022) and complied with the Declaration of Helsinki. All participants provided written informed consent after receiving detailed information regarding study objectives, procedures, potential risks, and participant rights. Assessments were conducted by licensed medical professionals from Beijing Fengtai Hospital and Nankang District Hospital of Traditional Chinese Medicine (Ganzhou, Jiangxi Province). Adverse events were evaluated and managed on-site. Personal data were anonymized and coded to ensure confidentiality and were used exclusively for research purposes.

After completing the study, participants were provided with full electronic instructional videos for CTC12, TC24, and square dancing.

Intervention Protocol

Participants in each intervention group attended supervised sessions twice weekly for 90 minutes over a 12-week period, conducted from 8 to 9:30 AM. The interventions were defined as follows: (1) CTC12, which emphasized explosive techniques (fa jin), jumps, and complex transitions to enhance neuromuscular coordination and cardiopulmonary function; (2) TC24, a standardized and simplified form of Yang-style Tai Chi, focusing on breathing rhythm, flexibility, and balance [37]; (3) square dance, a moderate-to-high intensity aerobic dance performed at 60%-75% of maximum heart rate (HR) [50]; (4) walking, performed as brisk walking at 3.5-4.0 km/h on flat terrain, targeting 50%-60% of maximum HR, and monitored using HR devices [50,51]; and (5) control, continued daily activities with no additional structured exercise.

All sessions were conducted outdoors under temperatures ranging from 24 °C to 28 °C. Although humidity and air circulation were not controlled, participant safety and comfort were ensured throughout. Each session began with a standardized 10-minute warm-up consisting of light aerobic activity and joint mobility exercises. Tai Chi interventions (CTC12 and TC24) were led by Shaoqing Zeng, with more than 10 years of teaching experience and recognized competitive achievements at the city level. Square dance and walking interventions were designed and supervised by WS, with a doctoral degree in Sports Science and extensive experience in physical education and exercise interventions, who had participated in the design and implementation of several community-based exercise programs (Multimedia Appendix 1).

XSL•FO
RenderX

Outcome Measures

Primary Outcomes

Fasting venous blood samples were collected between 7 AM and 9 AM before and after the 12-week intervention, following ≥10 hours of fasting, and were processed within 2 hours. Blood pressure was measured twice in a seated position using an automated sphygmomanometer (Omron HEM-7320 [Omron Healthcare]), and PP was calculated as SBP−DBP [52]. Lipid profiles, including TC, triglycerides, HDL-C, and LDL-C, were measured using enzymatic colorimetric assays (Cobas c 501 [Roche]) [52,53]. FPG was measured using the glucose oxidase method (LabX) [54]. Serum IL-6 concentrations were quantified using a high-sensitivity ELISA kit (Human IL-6 Quantikine HS ELISA [R&D Systems]; catalog HS600C) [55], with a detection limit of 0.09 pg/mL. According to the manufacturer's specifications, the intra-assay coefficient of variation ranged from 3.6% to 4.7% and interassay coefficient of variation from 3.9% to 10.8% across 3 concentration levels. Serum samples were aliquoted and stored at −80 °C until analysis [56]. Samples were assayed in duplicate, and the mean of the 2 measurements was used for analysis. The assay detection limit was 0.09 pg/mL. Laboratory technicians were blinded to group allocation, and all samples were analyzed in a single batch to minimize variability.

Basal metabolic rate (BMR) was estimated using bioelectrical impedance analysis (InBody H30NWi [InBody Co]) [57]. HR data, including resting HR and average exercise HR, were recorded using photoplethysmography-based monitors (Polar H10 [Polar Electro]) [58]. The change in heart rate (ΔHR) was defined as the difference between average exercise HR and resting HR during each 90-minute session. Energy expenditure (kcal), including calories at rest and additional calories expended during exercise sessions, was estimated from continuous HR monitoring (Polar H10) [59].

Physical function was assessed using the SPPB (score range 0-12), which includes static balance, 4-meter gait speed, and the 5-time chair stand test [60]. Functional independence was evaluated using ADLs, measured as the ability to perform basic daily activities [61]. WHOQOL was assessed using total scores across 6 domains from the WHOQOL-100 instrument [62].

Secondary Outcomes

Secondary outcomes included the six WHOQOL domain scores: (1) Q1: physical health, (2) Q2: psychological health, (3) Q3: social relationships, (4) Q4: environment, (5) Q5: self-perceived health, and (6) Q6: self-rated overall quality of life [62,63]. All evaluations were conducted by trained personnel blinded to group assignment.

Sample Size Estimation

An a priori power analysis was performed using G*Power (version 3.1.9.7) for a mixed repeated-measures ANOVA (within-between interaction; group×time interaction). The SPPB was selected as the primary outcome, reflecting the central role of physical function in this study. The expected effect size was set at η^2=0.065 based on previous literature [64], corresponding to Cohen $f=\sqrt{(\eta^2/(1-\eta^2))}$=0.264. Assuming α=.05, power (1−β)=0.95, 5 groups, 2 measurements (pre- and postintervention), correlation among repeated measures of 0.50, and ε=1, the required sample size was 75 participants (15 per group). To account for an anticipated attrition rate of 20%, a minimum of 94 participants (required final sample size=75, expected retention rate=0.80) were targeted (Multimedia Appendix 2).

Statistical Analysis

All statistical analyses were performed using R (version 4.2.1; R Core Team). Continuous variables are presented as mean (SD). Within-group comparisons were assessed using paired *t* tests, and between-group comparisons were conducted using 1-way ANOVA with Tukey post hoc tests. Normality and homogeneity of variance were evaluated using the Shapiro-Wilk and Levene tests, respectively. To address potential inflation of type I error due to multiple comparisons across outcomes (eg, blood pressure, fasting glucose, lipids, IL-6, SPPB, and WHOQOL subdomains), the false discovery rate correction was applied. All tests were 2-tailed, with statistical significance defined as $P<.05$. Adjusted *P* values, effect sizes, and 95% CIs were reported to complement dichotomous significance testing.

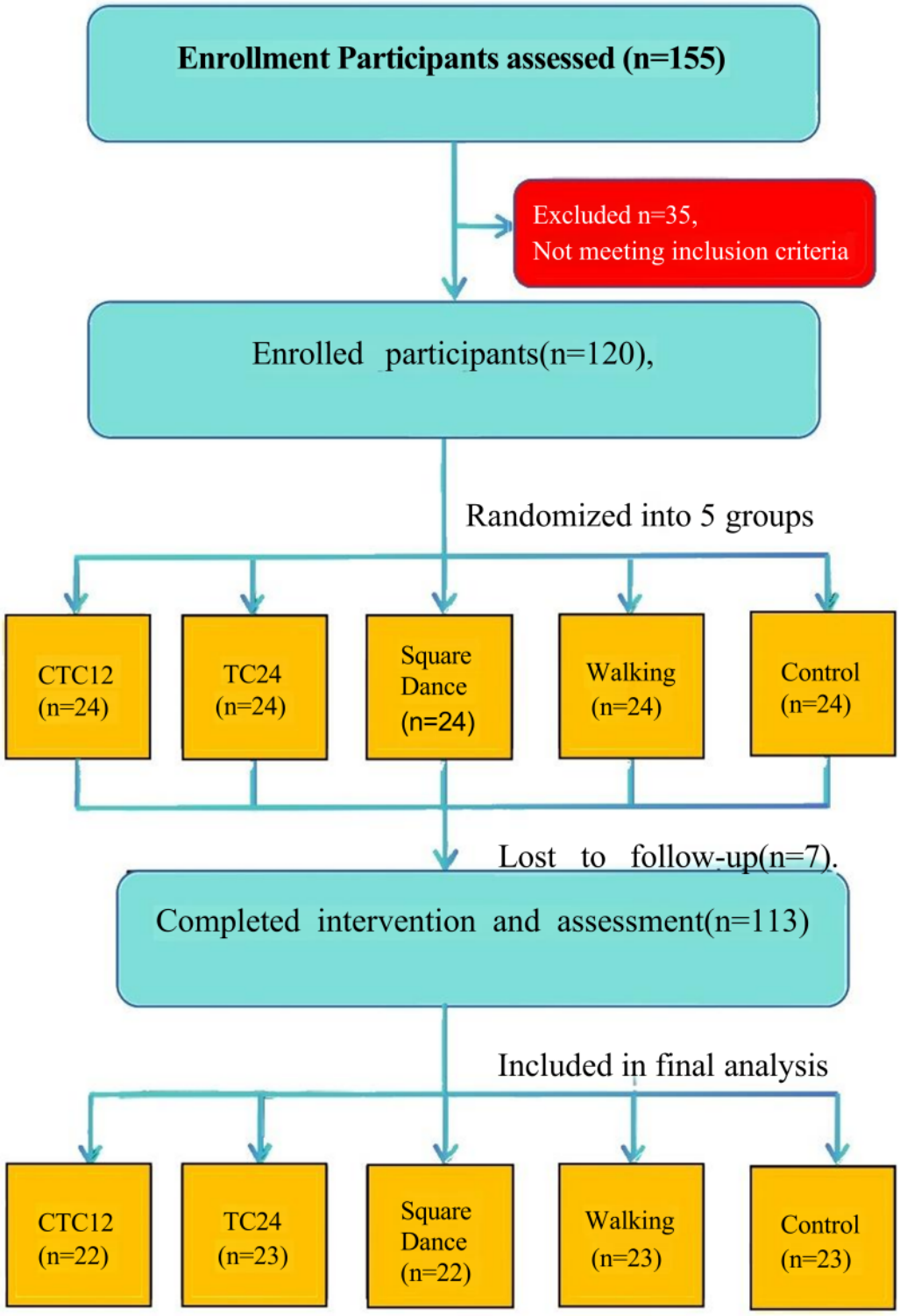
Results

Participant Flow and Baseline Characteristics

As illustrated in the CONSORT (Consolidated Standards of Reporting Trials) diagram, 155 individuals were screened, of whom 120 (77.4%) eligible community-dwelling middle-aged and older adults were enrolled. During the 12-week intervention, 7 participants withdrew due to personal reasons or insufficient attendance (<90%) and were subsequently classified as dropouts. Ultimately, 113 participants completed the intervention and all assessments. The final group sizes were as follows: (1) CTC12: n=22, (2) TC24: n=23, (3) square dance: n=22, (4) walking: n=23, and (5) control: n=23 (Figure 1, Table 1).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of participant recruitment, group allocation, intervention, and analysis. Control: control group; CTC12: Chen-style Tai Chi 12 forms group; Square Dance: square dance group; TC24: Simplified Tai Chi 24 forms group; Walking: walking exercise group. All analyzed participants attended ≥90% of sessions and met protocol compliance.



XSL•FO
RenderX

Table 1. Baseline characteristics of participants across intervention groups.

Variable	CTC12[a] group (n=22), mean (SD)	TC24[b] group (n=23), mean (SD)	Square dance group (n=22), mean (SD)	Walking group (n=23), mean (SD)	Control group (n=23), mean (SD)	*P* value[c]
Age (y)	62.45 (5.26)	61.65 (2.99)	62.64 (4.11)	61.61 (4.29)	63.17 (4.48)	.69
Sex, n						
Male	5	5	4	5	4	.99
Female	17	18	18	18	19	.99
Height (cm)	159.14 (6.07)	158.22 (5.00)	158.32 (5.11)	158.83 (5.04)	160.13 (4.86)	.74
Weight (kg)	60.42 (8.39)	59.09 (6.96)	59.11 (6.38)	59.38 (6.73)	61.47 (6.82)	.75
BMI (kg/m²)	23.87 (3.20)	23.59 (2.35)	23.62 (2.80)	23.51 (2.00)	23.98 (2.54)	.97
DBP[d] (mm Hg)	80.59 (6.81)	81.87 (5.52)	81.09 (6.24)	80.43 (7.04)	80.87 (6.06)	.95
SBP[e] (mm Hg)	126.18 (7.69)	126.04 (7.74)	126.05 (6.42)	126.43 (9.85)	125.70 (7.78)	.99
PP[f] (mm Hg)	45.59 (9.05)	44.17 (8.89)	44.96 (7.51)	46.00 (7.67)	44.83 (7.77)	.95
TC[g] (mmol/L)	5.40 (0.97)	5.44 (1.05)	5.43 (1.06)	5.39 (0.82)	5.51 (1.01)	.99
TG[h] (mmol/L)	1.62 (1.03)	1.70 (0.82)	1.62 (0.52)	1.66 (0.66)	1.70 (0.61)	.99
HDL-C[i] (mmol/L)	1.46 (0.30)	1.37 (0.35)	1.30 (0.30)	1.34 (0.33)	1.30 (0.18)	.35
LDL-C[j] (mmol/L)	2.67 (0.57)	2.71 (0.64)	2.60 (0.54)	2.62 (0.59)	2.74 (0.73)	.94
Fasting glucose (mmol/L)	5.28 (0.67)	5.32 (0.85)	5.26 (0.81)	5.30 (0.81)	5.38 (0.81)	.99
IL-6[k] (pg/ml)	2.91 (1.88)	2.90 (1.75)	2.99 (1.84)	2.93 (2.50)	2.99 (1.50)	.52
BMR[l] (kcal)	1273.32 (108.56)	1322.24 (114.76)	1271.36 (60.90)	1261.61 (103.29)	1264.78 (99.30)	.23
SPPB[m] (scores)	9.59 (0.67)	9.52 (0.99)	9.50 (1.10)	9.57 (0.59)	9.35 (0.83)	.89
ADL[n] (scores)	98.64 (2.28)	98.70 (2.24)	98.18 (2.46)	98.70 (2.24)	98.91 (2.11)	.87
WHOQOL[o] total (scores)	102.36 (7.49)	102.26 (5.77)	102.14 (6.22)	101.87 (5.41)	102.13 (5.27)	.10
Resting heart rate (bpm)	67.77 (3.64)	67.83 (3.88)	67.59 (3.98)	67.70 (4.76)	67.22 (3.58)	.99
Calories at rest (kcal)	84.38 (5.82)	83.20 (6.67)	83.68 (4.57)	83.01 (6.54)	84.03 (6.74)	.94

[a]CTC12: 12-form Chen-style Tai Chi.

[b]TC24: 24-form simplified Tai Chi.

[c]No significant differences were observed between groups at baseline (*P*>.05 for all comparisons).

[d]DPB: diastolic blood pressure.

[e]SPB: systolic blood pressure.

[f]PP: pulse pressure.

[g]TC: total cholesterol.

[h]TG: triglycerides.

[i]HDL-C: high-density lipoprotein cholesterol.

[i]LDL-C: low-density lipoprotein cholesterol.

[k]IL-6: interleukin-6.

[l]BMR: basal metabolic rate.

[m]SPPB: Short Physical Performance Battery.

[n]ADL: activity of daily living.

[o]WHOQOL: World Health Organization Quality of Life questionnaire.

XSL•FO
RenderX

Primary Outcomes

Within-Group Comparisons

No significant changes were observed in DBP, SBP, and PP across all groups (all *P*>.05). Regarding lipid profiles, TC significantly decreased only in the square dance group (*P*=.01). Triglyceride and LDL-C significantly decreased in both the CTC12 (*P*=.008 and *P*=.002, respectively) and square dance (*P*=.002 for both) groups. HDL-C remained unchanged in all groups (all *P*>.05). In terms of fasting glucose and IL-6, no significant changes were noted in any group (all *P*>.05). For physical function and quality of life, SPPB scores significantly improved in the CTC12 (*P*=.02) and square dance (*P*=.04) groups, whereas ADL scores significantly improved only in the square dance group (*P*=.04). WHOQOL scores significantly improved in the CTC12 (*P*=.002), square dance (*P*=.002), and walking (*P*=.02) groups (Table 2).

XSL•FO
RenderX

Table 2. Within-group comparisons of key health indicators before and after intervention (paired *t* tests).

Variable and group	Preintervention, mean (SD)	Postintervention, mean (SD)	*t* test (*df*)	Change[a], marginal mean (95% CI)	*P* value	Adjusted *P* value
DBP[b] (mm Hg)						
CTC12[c]	80.59 (6.81)	80.36 (6.92)	–0.6918 (21)	–0.23 (–0.91 to 0.46)	.50	.62
TC24[d]	81.87 (5.52)	81.65 (5.73)	–0.7740 (22)	1.26 (–0.80 to 0.37)	.45	.62
Square dance	81.09 (6.24)	81.36 (6.18)	1.2396 (21)	0.27 (–0.18 to 0.73)	.23	.48
Walking	80.43 (7.04)	79.61 (6.85)	–2.7018 (22)	–0.82 (–1.46 to –0.19)	.01	.13
Control	80.87 (6.06)	80.83 (5.67)	–0.2041 (22)	–0.04 (–0.49 to 0.40)	.84	.84
SBP[e] (mm Hg)						
CTC12	126.18 (7.69)	125.45 (7.25)	–2.3470 (21)	–0.73 (–1.37 to –0.08)	.03	.14
TC24	126.04 (7.74)	125.57 (7.50)	–0.7359 (22)	–0.47 (–1.83 to 0.87)	.47	.62
Square dance	126.05 (6.42)	125.59 (6.53)	–1.2076 (21)	–0.46 (–1.24 to 0.33)	.24	.48
Walking	126.43 (9.85)	126.30 (9.78)	–0.3708 (22)	–0.13 (–0.86 to 0.40)	.71	.79
Control	125.70 (7.78)	125.96 (7.57)	1.3667 (22)	0.26 (–0.13 to 0.66)	.19	.48
PP[f] (mm Hg)						
CTC12	45.59 (9.05)	45.09 (9.01)	–1.7100 (21)	–0.50 (–1.12 to 0.12)	.10	.15
TC24	44.17 (8.89)	43.91 (8.54)	–0.3800 (22)	–0.26 (–1.68 to 1.16)	.70	.75
Square dance	44.96 (7.51)	44.23 (7.78)	–2.5238 (21)	–0.73 (–1.33 to –0.13)	.02	.07
Walking	46.00 (7.67)	46.70 (7.61)	1.8085 (22)	0.70 (–0.10 to 1.49)	.08	.13
Control	44.83 (7.77)	45.09 (7.65)	2.0205 (22)	0.26 (–0.01 to 0.53)	.06	.11
TC[g] (mmol/L)						
CTC12	5.40 (0.97)	5.03 (0.63)	–2.7694 (21)	–0.37 (–0.65 to –0.09)	.01	.05
TC24	5.44 (1.05)	5.28 (0.99)	–1.2498 (22)	–0.16 (–0.42 to 0.10)	.23	.30
Square dance	5.43 (1.06)	4.79 (1.16)	–3.5620 (21)	–0.64 (–1.00 to –0.26)	.002	.01
Walking	5.39 (0.82)	5.24 (0.71)	–0.9954 (22)	–0.15 (–0.46 to 0.16)	.33	.39
Control	5.51 (1.01)	5.72 (0.84)	1.9356 (22)	0.21 (–0.01 to 0.42)	.07	.12
TG[h] (mmol/L)						
CTC12	1.62 (1.03)	1.18 (0.60)	–3.9299 (21)	–0.44 (–0.68 to –0.21)	.001	.008
TC24	1.69 (0.82)	1.49 (0.77)	–2.7039 (22)	–0.20 (–0.37 to –0.05)	.01	.05
Square dance	1.62 (0.52)	1.27 (0.34)	–4.6725 (21)	–0.35 (–0.50 to –0.19)	<.001	.002
Walking	1.66 (0.66)	1.48 (0.67)	–2.4388 (22)	–0.18 (–0.34 to –0.03)	.023	.067
Control	1.70 (0.61)	1.93 (0.68)	1.8613 (22)	0.23 (–0.03 to 0.47)	.09	.13
HDL-C[i] (mmol/L)						
CTC12	1.47 (0.30)	1.53 (0.30)	2.1168 (21)	0.06 (0.00 to 0.13)	.05	.11
TC24	1.37 (0.35)	1.38 (0.37)	0.2423 (22)	0.01 (–0.076 to 0.10)	.81	.81
Square dance	1.30 (0.30)	1.37 (0.37)	2.0256 (21)	0.07 (–0.00 to 0.13)	.06	.11
Walking	1.34 (0.33)	1.37 (0.26)	1.1281 (22)	0.03 (–0.03 to 0.09)	.27	.34
Control	1.30 (0.18)	1.27 (0.19)	–0.9172 (22)	–0.03 (–0.12 to 0.05)	.37	.41
LDL-C[j] (mmol/L)						
CTC12	2.67 (0.57)	2.23 (0.70)	–4.6953 (21)	–0.44 (–0.64 to –0.25)	<.001	.002
TC24	2.71 (0.64)	2.57 (0.54)	–2.3375 (22)	–0.14 (–0.26 to –0.02)	.03	.07

XSL•FO
RenderX

Variable and group	Preintervention, mean (SD)	Postintervention, mean (SD)	*t* test (*df*)	Change[a], marginal mean (95% CI)	*P* value	Adjusted *P* value
Square dance	2.60 (0.54)	2.19 (0.48)	–4.5876 (21)	–0.41 (–0.60 to –0.23)	<.001	.002
Walking	2.62 (0.59)	2.50 (0.56)	–2.2419 (22)	–0.12 (–0.25 to –0.01)	.04	.08
Control	2.74 (0.73)	2.91 (0.77)	1.7128 (22)	0.17 (–0.035 to 0.37)	.10	.15
Fasting glucose (mmol/L)						
CTC12	5.35 (0.80)	5.06 (0.56)	–2.4815 (21)	–0.29 (–0.53 to –0.05)	.02	.07
TC24	5.32 (0.85)	5.28 (0.77)	–0.4939 (22)	–0.04 (–0.23 to 0.14)	.63	.67
Square dance	5.26 (0.81)	4.89 (0.82)	–2.1352 (21)	–0.37 (–0.72 to –0.01)	.05	.09
Walking	5.30 (0.81)	5.26 (0.59)	–0.3960 (22)	–0.04 (–0.24 to 0.16)	.70	.72
Control	5.38 (0.81)	5.50 (0.75)	1.4302 (22)	0.12 (–0.05 to 0.29)	.17	.21
IL-6[k] (pg/ml)						
CTC12	2.91 (1.88)	2.58 (1.46)	–2.3568 (21)	–0.33 (–0.62 to –0.04)	.03	.07
TC24	2.90 (1.75)	2.83 (1.69)	–0.5833 (22)	–0.07 (–0.34 to 0.19)	.57	.63
Square dance	2.99 (1.84)	2.88 (1.71)	–1.8769 (21)	–0.11 (–0.20 to 0.01)	.08	.14
Walking	2.93 (2.50)	2.87 (2.48)	–1.3773 (22)	–0.06 (–0.17 to 0.03)	.18	.22
Control	2.99 (1.50)	3.09 (1.42)	1.7851 (22)	0.10 (–0.02 to 0.22)	.09	.14
SPPB[l] (scores)						
CTC12	9.59 (0.67)	10.00 (0.82)	3.2504 (21)	0.41 (0.15 to 0.67)	.004	.02
TC24	9.52 (0.99)	9.65 (0.88)	1.8166 (22)	0.13 (–0.02 to 0.28)	.08	.14
Square dance	9.50 (1.10)	10.00 (0.93)	2.9250 (21)	0.50 (0.14 to 0.86)	.008	.04
Walking	9.57 (0.59)	9.61 (0.84)	0.2955 (22)	0.04 (–0.26 to 0.35)	.77	.77
Control	9.35 (0.83)	9.22 (0.74)	–1.3667 (22)	–0.13 (–0.33 to 0.07)	.19	.22
ADL[m] (scores)						
CTC12	98.64 (2.28)	99.55 (1.47)	2.1602 (21)	0.91 (0.03 to 1.78)	.04	.09
TC24	98.26 (2.43)	98.48 (2.35)	1.0000 (22)	0.22 (–0.23 to 0.67)	.33	.38
Square dance	98.18 (2.46)	99.55 (1.47)	2.8062 (21)	1.37 (0.35 to 2.37)	.01	.04
Walking	98.70 (2.24)	99.35 (1.72)	1.8166 (22)	0.65 (–0.09 to 1.40)	.08	.14
Control	98.91 (2.11)	98.48 (2.35)	–1.4475 (22)	–0.43 (–1.06 to 0.19)	.16	.22
WHOQOL[n] total (scores)						
CTC12	102.36 (7.49)	105.23 (5.76)	4.4585 (21)	2.87 (1.53 to 4.20)	<.001	.002
TC24	102.26 (5.77)	104.65 (4.24)	2.3246 (22)	2.39 (0.26 to 4.52)	.03	.07
Square dance	102.14 (6.22)	105.18 (5.33)	4.6643 (21)	3.04 (1.69 to 4.40)	<.001	.002
Walking	101.87 (5.41)	103.48 (5.24)	3.2565 (22)	1.61 (0.58 to 2.63)	.004	.02

XSL•FO
RenderX

Variable and group	Preintervention, mean (SD)	Postintervention, mean (SD)	*t* test (*df*)	Change[a], marginal mean (95% CI)	*P* value	Adjusted *P* value
Control	102.13 (5.27)	101.43 (5.21)	–1.6599 (22)	–0.70 (–1.56 to 0.17)	.11	.16

[a]Change indicates the mean difference between pre- and postintervention values.

[b]DPB: diastolic blood pressure.

[c]CTC12: 12-form Chen-style Tai Chi.

[d]TC24: 24-form simplified Tai Chi.

[e]SPB: systolic blood pressure.

[f]PP: pulse pressure.

[g]TC: total cholesterol.

[h]TG: triglycerides.

[i]HDL-C: high-density lipoprotein cholesterol.

[i]LDL-C: low-density lipoprotein cholesterol.

[k]IL-6: interleukin-6.

[l]SPPB: Short Physical Performance Battery.

[m]ADL: activity of daily living.

[n]WHOQOL: World Health Organization Quality of Life questionnaire.

Between-Group Comparisons

Postintervention differences were assessed using 1-way ANOVA with Tukey post hoc tests (Table 3; Figure S1 in Multimedia Appendix 3).

XSL•FO
RenderX

Table 3. Between-group comparisons of postintervention outcomes among 5 intervention arms (1-way ANOVA).

Variable	CTC12[a] (n=22), mean (SD)	TC24[b] (n=23), mean (SD)	Square dance (n=22), mean (SD)	Walking (n=23), mean (SD)	Control (n=23), mean (SD)	*F* test[c] (*df*)	η^{2}[d]	*P* value[e]	Post hoc[f]
DBP[g] (mm Hg)	80.36 (6.92)	81.65 (5.73)	81.36 (6.18)	79.61 (6.85)	80.83 (5.67)	0.38 (4,108)	0.014	.82	Ns[h]
SBP[i] (mm Hg)	125.45 (7.25)	125.57 (7.50)	125.59 (6.53)	126.30 (9.78)	125.91 (7.75)	0.04 (4,108)	0.002	.99	Ns
PP[j] (mm Hg)	45.09 (9.01)	43.91 (8.54)	44.23 (7.78)	46.70 (7.61)	45.09 (7.65)	0.40 (4,108)	0.015	.81	Ns
TC[k] (mmol/L)	5.03 (0.63)	5.28 (0.99)	4.79 (1.16)	5.24 (0.71)	5.79 (0.90)	3.82 (4,108)	0.124	.006	Control>CTC12, square dance
TG[l] (mmol/L)	1.18 (0.60)	1.49 (0.77)	1.27 (0.34)	1.48 (0.67)	1.93 (0.68)	4.75 (4,108)	0.149	.001	Control>CTC12, square dance
HDL-C[m] (mmol/L)	1.53 (0.30)	1.38 (0.37)	1.37 (0.37)	1.37 (0.27)	1.27 (0.19)	2.14 (4,108)	0.073	.08	Control<CTC12; others ns
LDL-C[n] (mmol/L)	2.23 (0.70)	2.57 (0.54)	2.18 (0.47)	2.50 (0.56)	2.91 (0.77)	4.98 (4,108)	0.156	.001	Control>CTC12, square dance
Fasting glucose (mmol/L)	4.90 (0.58)	5.28 (0.77)	4.89 (0.82)	5.26 (0.59)	5.50 (0.75)	3.07 (4,108)	0.102	.02	Control>CTC12, square dance
IL-6[o] (pg/ml)	2.58 (1.46)	2.83 (1.69)	2.88 (1.71)	2.87 (2.48)	3.09 (1.42)	0.23 (4,108)	0.009	.92	Ns
SPPB[p] (scores)	10.00 (0.82)	9.61 (0.94)	10.00 (0.93)	9.61 (0.84)	9.22 (0.74)	3.31 (4,108)	0.109	.01	CTC12, square dance>Control
ADL[q] (scores)	99.55 (1.47)	98.48 (2.35)	99.55 (1.47)	99.35 (1.72)	98.48 (2.35)	1.88 (4,108)	0.065	.12	Ns
WHOQOL[r] total (scores)	105.23 (5.76)	104.65 (4.24)	105.18 (5.33)	103.48 (5.24)	100.91 (5.09)	2.80 (4,108)	0.094	.03	CTC12, square dance>control
BMR[s] (kcal)	1391.50 (70.78)	1314.96 (91.49)	1396.55 (74.45)	1296.74 (102.82)	1257.22 (92.10)	10.90 (4,108)	0.2877	<.001	CTC12, square dance>control, TC24, walking
ΔHR[t] (bpm)	29.50 (4.61)	22.48 (6.67)	46.68 (7.48)	19.48 (2.19)	0.00 (0.00)	105.35 (4,108)	0.7861	<.001	Square dance>CTC12> TC24 > Walking
1.5 hours additional calories burned (kcal)	363.75 (18.54)	207.80 (15.76)	386.65 (37.29)	166.36 (12.60)	0.00 (0.00)	518.87 (4,108)	0.948	<.001	Square dance>CTC12>TC24, Walking

[a]CTC12: 12-form Chen-style Tai Chi.

[b]TC24: 24-form simplified Tai Chi.

[c]*F* test: analysis of variance test statistic.

[d]η^{2}: representing effect size.

[e]*P* values represent the significance of overall group differences.

[f]Pairwise comparisons were conducted using Tukey's post hoc tests where appropriate.

[g]DPB: diastolic blood pressure.

[h]Ns: not significant.

[i]SPB: systolic blood pressure.

[j]PP: pulse pressure.

[k]TC: total cholesterol.

[l]TG: triglycerides.

[m]HDL-C: high-density lipoprotein cholesterol.

[n]LDL-C: low-density lipoprotein cholesterol.

[o]IL-6: interleukin-6.

[p]SPPB: Short Physical Performance Battery.

XSL•FO
RenderX

[q]ADL: activity of daily living.

[r]WHOQOL: World Health Organization Quality of Life questionnaire.

[s]BMR: basal metabolic rate.

[t]ΔHR: heart rate increase.

No significant between-group differences were observed for SBP (*P*=.99), DBP (*P*=.82), or PP (*P*=.81).

TC, triglyceride, and LDL-C were significantly lower in the CTC12 (mean 5.03, SD 0.63 mmol/L; mean 1.18, SD 0.60 mmol/L; and mean 2.23, SD 0.70 mmol/L, respectively) and square dance (mean 4.79, SD 1.16 mmol/L; mean 1.27, SD 0.34 mmol/L; and mean 2.18, SD 0.47 mmol/L, respectively) groups than in the control group (mean 5.79, SD 0.90 mmol/L; mean 1.93, SD 0.68 mmol/L; and mean 2.91, SD 0.77 mmol/L; and *P*=.006, *P*=.001, and *P*=.001, respectively). HDL-C was higher in the CTC12 group (mean 1.53, SD 0.30 mmol/L) than in the control group (mean 1.27, SD 0.19 mmol/L), although this did not reach statistical significance (*P*=.08).

Fasting glucose was significantly lower in the CTC12 and square dance groups than in the control group (*P*=.02), whereas IL-6 did not differ significantly among groups (*P*=.92).

SPPB and WHOQOL scores were significantly higher in the CTC12 and square dance groups than in the control group (*P*=.01 and *P*=.03, respectively), while ADL did not differ between groups (*P*=.12).

BMR, ΔHR, and additional energy expenditure differed significantly among groups (all *P*<.001). BMR was higher in the CTC12 (mean 1391.50, SD 70.78 kcal) and square dance (mean 1396.55, SD 74.45 kcal) groups than in the TC24, walking, and control groups. ΔHR was greatest in the square dance group (mean 46.68, SD 7.48 bpm), followed by the CTC12 group, which showed an increase of mean 29.50 (SD 4.61) bpm, indicating significant differences between groups (*P*<.001).

Additional energy expenditure during 1.5-h sessions was also higher in the square dance (mean 386.65, SD 37.29 kcal) and CTC12 (mean 363.75, SD 18.54 kcal) groups than in the other groups, indicating greater exercise intensity (Table 3; Figure S1 in Multimedia Appendix 3).

Secondary Outcomes: WHOQOL Subdomains

Within-Group Comparisons

Following the intervention, improvements were observed across multiple WHOQOL subdomains. Q1 and Q2 scores significantly improved in all intervention groups (CTC12, square dance, TC24, and walking; all *P*<.05), whereas no significant changes were observed in the control group. In contrast, Q3 improved significantly only in the square dance group (*P*=.003), with no significant changes in the other groups. The environmental quality of life (Q4) remained unchanged across all groups (*P*>.05), although the square dance group showed a borderline decline (*P*=.05). Self-perceived health status (Q5) significantly improved in all intervention groups (*P*<.01). Similarly, self-rated overall quality of life (Q6) increased significantly in all intervention groups (*P*<.01), with no change observed in the control group (Table S1 in Multimedia Appendix 3).

Between-Group Comparisons

Postintervention comparisons showed varying patterns across subdomains. Q1 scores were highest in the square dance group and lowest in the control group, with a marginal group effect (*P*=.06). Q2 scores were significantly higher in the CTC12 and square dance groups than in the control group (*P*=.005), indicating improved psychological health. No significant differences were observed for Q3 (*P*=.13), although the square dance group had the highest mean score. Similarly, no significant between-group differences were observed for Q4 (*P*=.78). For Q5, a near-significant group effect was observed (*P*=.09), with higher scores in the CTC12 and square dance groups than in the control group. Q6 did not differ significantly among groups (*P*=.13), although the CTC12 group showed the highest mean overall quality of life score (Table S2 and Figure S2 in Multimedia Appendix 3).

Discussion

Overview of Intervention Effects

This 12-week exercise intervention demonstrated significant benefits in the CTC12 and square dance groups. Both interventions improved lipid profiles, with TC reduced in the square dance group and triglyceride and LDL-C reduced in both the CTC12 and square dance groups, while no significant changes were observed in FPG or IL-6. Functional outcomes also improved, as reflected by increased SPPB scores in both groups and improved ADL scores in the square dance group. WHOQOL scores significantly improved in the CTC12, square dance, and walking groups, indicating enhanced quality of life. Collectively, these findings support the effectiveness of CTC12 and square dance in improving lipid metabolism, physical function, and quality of life in middle-aged and older adults.

Improvements in Cardiometabolic Indicators and Underlying Mechanisms

Regarding blood pressure, although no significant changes were observed in DBP, SBP, or PP across the groups, small trends in blood pressure changes were noted. From a clinical perspective, such small blood pressure changes are unlikely to represent meaningful short-term cardiovascular risk reduction, although minor population-level shifts have been associated with long-term risk trends in epidemiological studies [65]. Mechanistically, CTC12 may influence SBP through coordinated breathing and postural control, potentially modulating autonomic balance and arterial compliance [48]. The square dance intervention, characterized by rhythmic and moderate-to-high intensity movements, may contribute to improvements in large-artery elasticity, which could partly explain the observed PP reduction [66]. Walking, as a low-intensity aerobic activity, may enhance endothelial function and peripheral vasodilation, contributing to modest DBP changes [67]. These mechanisms provide plausible explanations for the observed directional

XSL•FO
RenderX

trends but do not imply clinically significant blood pressure reductions within the intervention period of this study.

The CTC12 and square dance groups demonstrated significant reductions in TC, triglyceride, and LDL-C, while HDL-C was significantly increased in the CTC12 group, indicating beneficial effects on lipid metabolism. The CTC12 regimen, which incorporates strength, coordination, and breathing control, may enhance lipid homeostasis and contribute to increases in HDL-C [68]. In contrast, square dance may improve lipid profiles through enhanced lipoprotein lipase activity, facilitating very-low-density lipoprotein clearance and cholesterol transport [68]. Although the CTC12 group exhibited greater increases in HDL-C, and the square dance group showed more pronounced reductions in TC, both groups demonstrated complementary effects on lipid regulation, which may have beneficial implications for cardiovascular prevention [69]. The square dance group exhibited a more pronounced reduction in TC, whereas the CTC12 group was more effective in elevating HDL-C, indicating distinct yet complementary lipid-regulatory mechanisms. Both groups achieved clinically relevant lipid thresholds (eg, triglyceride<1.7 mmol/L and LDL-C<3.4 mmol/L), underscoring their potential role in primary cardiovascular prevention. This suggests that the observed lipid improvements were not only statistically significant but also potentially relevant for practical cardiovascular risk management in community-dwelling older adults [70].

Although not statistically significant, the control group exhibited a worsening trend in TC, triglyceride, and other cardiometabolic markers, which may reflect the absence of structured physical activity [71].

Modulation of Inflammatory Markers

No significant changes in IL-6 levels were observed across intervention groups. This may reflect the relatively healthy baseline status of the participants and the limited duration of the intervention. In such populations, longer intervention periods may be required to detect measurable changes in inflammatory markers. Although mind–body interventions, such as CTC12 and Square Dance, may influence parasympathetic activity and low-grade inflammation, these effects did not translate into detectable changes in IL-6 within the short intervention period. Future studies with extended follow-up are needed to clarify these effects [72].

Improvements in Physical Function and BMR

The CTC12 group demonstrated the greatest improvements in SPPB scores, followed by the square dance group, whereas no significant changes were observed in the walking group. The coordinated activation of core and lower-limb musculature in CTC12 likely contributes to improvements in balance, coordination, and postural control, thereby reducing the risk of falls and improving gait performance. These improvements may be clinically meaningful, as better SPPB performance is closely associated with mobility, reduced fall risk, and preserved functional independence in older adults [73].

Square dance, characterized by rhythmic movement and group participation, was associated with increases in BMR, suggesting enhanced metabolic activity and energy efficiency, which may help delay age-related metabolic decline [74].

Improvements in Quality of Life and Mental Health

WHOQOL scores improved in the CTC12, square dance, and TC24 groups, with the most pronounced improvements in psychological health gains (Q2) observed in the CTC12 and square dance groups. The underlying mechanisms appear distinct. CTC12 may enhance emotional regulation and self-efficacy through long-term cognitive engagement and controlled breathing, facilitating parasympathetic activation [75]. Conversely, the square dance may improve mood, social cohesion, and subjective well-being through social interaction, rhythmic movement, and music [76]. These findings suggest that CTC12 emphasizes internal self-regulation, whereas square dance operates through external social reinforcement, highlighting the importance of tailoring interventions to individual psychological and social profiles [63]. A slight worsening trend in psychological scores was observed in the control group, which may reflect the effects of prolonged inactivity [77].

Safety and Feasibility of the Interventions

Mean exercise HR across all intervention groups remained within 50%-70% of HR_max, consistent with American College of Sports Medicine guidelines for middle-aged and older adults [78]. No falls or cardiovascular events were observed in the CTC12 group, despite the inclusion of explosive movements, likely due to structured warm-up protocols, progressive difficulty, and professional supervision. Similarly, no injuries were reported in the square dance group, despite engaging in moderate-to-high intensity activities. Attendance exceeded 90% across all groups, indicating high adherence and feasibility. These findings support the safety and applicability of these interventions in community-based settings.

Dose-Related Interpretation and Practical Considerations

A descriptive dose-related interpretation was used to examine how differences in exercise intensity and physiological load relate to the breadth of observed health benefits. CTC12 and square dance improved a greater number of outcomes and were therefore compared in terms of exercise dose characteristics and physiological demand. Square dance demonstrated higher energy expenditure (386.7 kcal) and ΔHR (46.7 bpm; approximately 72%-75% HR_max), reflecting moderate-to-high intensity activity. This higher intensity may explain the more rapid improvements in physical function but may also exceed optimal safety thresholds for some older adults, particularly those with cardiovascular risk factors, potentially limiting long-term sustainability. Nevertheless, square dance led to more immediate improvements in physical function, particularly in balance and muscular endurance, possibly due to its higher energy expenditure [79]. CTC12 achieved relatively high energy expenditure (363.8 kcal) and metabolic activation while maintaining a lower cardiovascular load (ΔHR=29.5 bpm; approximately 60%-65% HR_max), consistent with recommended safety ranges for older adults [80]. Although improvements appeared more gradual, CTC12 was associated

XSL•FO
RenderX

with long-term benefits across physical, psychological, and quality-of-life domains with a lower risk of excessive fatigue or cardiovascular strain. These findings suggest distinct practical roles for each intervention. Square dance may be more suitable for individuals seeking higher-intensity activity and faster functional gains, whereas CTC12 may provide a more sustainable, lower-risk option for long-term health maintenance [81-83]. From an implementation perspective, both interventions are feasible in community settings; however, CTC12 may be more appropriate for broader populations due to its moderate intensity and lower cardiovascular demand, whereas square dance may require closer monitoring in higher-risk individuals [84,85].

Limitations

This study has some limitations. The relatively small sample size (n=22 per group) may have limited statistical power, and the 12-week intervention period may not have been adequate to capture long-term effects. Although outcome assessors were blinded to group allocation, participants were necessarily aware of their assigned intervention due to the distinct characteristics of each exercise modality. This awareness may have introduced expectancy-related influences, particularly for self-reported outcomes such as quality of life (eg, WHOQOL). However, such effects are less likely to have substantially influenced objective measures, including blood pressure, biochemical indicators, and physical function. Laboratory personnel responsible for biochemical analyses were blinded to group assignment, which helped mitigate detection bias.

Although the final sample size was adequate for the primary group×time analysis, the study may have been underpowered to detect differences in some secondary outcomes. Accordingly, these findings should be interpreted with caution. In addition, advanced cardiovascular assessments were not included, which limits the ability to characterize underlying physiological changes in greater detail. Future studies should include larger and more diverse samples, extend intervention duration, and incorporate participants at elevated clinical risk. The application of advanced analytical approaches, such as structural equation modeling, may facilitate a more comprehensive understanding of the interrelated metabolic, neural, emotional, and social pathways influenced by exercise. Building on the findings of this study in relatively healthy middle-aged and older adults, further research in populations with cardiovascular disease, metabolic disorders, and other chronic conditions is warranted to clarify clinical benefits and underlying mechanisms across different disease contexts.

Conclusion

In summary, we demonstrated that both CTC12 and square dance were associated with improvements in blood lipid profiles, physical function, and quality of life in middle-aged and older adults. CTC12 was associated with improvements across multiple health indicators, whereas square dance showed more pronounced effects on physical function, metabolic activity, and emotional well-being. In contrast, the effects observed for TC24 and walking were more limited and were confined to selected outcomes. These findings support the role of structured, moderate-to-high-intensity exercise in promoting healthy aging and provide preliminary evidence to inform the development of targeted exercise interventions for age-related health conditions.

Acknowledgments

The authors gratefully acknowledge all participants in this study. We also thank the medical staff of Beijing Fengtai Hospital and Nankang District Hospital of Traditional Chinese Medicine, Ganzhou, Jiangxi Province, for conducting professional pre- and postintervention assessments. Special thanks are extended to Ms Shaoqing Zeng, Ms Xia Wang, and the Nankang District Tai Chi Association for the Elderly in Ganzhou, Jiangxi, for their support in recruiting community-dwelling older adults.

Funding

This study was supported by the research fund of Hanyang University (HY-202400000001631).

Authors' Contributions

JQ conceptualized and designed the study, oversaw intervention procedures, and drafted the manuscript. JW contributed to intervention implementation, clinical data acquisition, and data analysis. XS performed statistical analyses and interpreted the results. WS assisted with data analysis. SMK provided critical input in study conceptualization, design refinement, and data interpretation. All authors reviewed and approved the final manuscript and agreed on the authorship order.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional figures.
[PDF File (Adobe PDF File), 1286 KB - aging_v9i1e80125_app1.pdf]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).

XSL•FO
RenderX

[PDF File (Adobe PDF File), 1335 KB - aging_v9i1e80125_app2.pdf]

Multimedia Appendix 3
Additional Figures and Tables.
[DOCX File , 7090 KB - aging_v9i1e80125_app3.docx]

References

1. GBD 2021 Demographics Collaborators. Global age-sex-specific mortality, life expectancy, and population estimates in 204 countries and territories and 811 subnational locations, 1950-2021, and the impact of the COVID-19 pandemic: a comprehensive demographic analysis for the Global Burden of Disease Study 2021. Lancet 2024;403(10440):1989-2056 [FREE Full text] [doi: 10.1016/S0140-6736(24)00476-8] [Medline: 38484753]
2. Chen C, Lim J, Koh J, Beard J, Rowe JW. A global analysis of adaptation to societal aging across low-, middle- and high-income countries using the Global Aging Society Index. Nat Aging 2025;5(1):113-121. [doi: 10.1038/s43587-024-00772-3] [Medline: 39730823]
3. Wong MYC, Ou KI, Chung PK, Chui KYK, Zhang CQ. The relationship between physical activity, physical health, and mental health among older Chinese adults: a scoping review. Front Public Health 2022;10:914548 [FREE Full text] [doi: 10.3389/fpubh.2022.914548] [Medline: 36684983]
4. Hickey A, Barker M, McGee H, O'Boyle C. Measuring health-related quality of life in older patient populations: a review of current approaches. Pharmacoeconomics 2005;23(10):971-993. [doi: 10.2165/00019053-200523100-00002] [Medline: 16235972]
5. Agostini D, Gervasi M, Ferrini F, Bartolacci A, Stranieri A, Piccoli G, et al. An integrated approach to skeletal muscle health in aging. Nutrients 2023;15(8):1802 [FREE Full text] [doi: 10.3390/nu15081802] [Medline: 37111021]
6. Dhana K, Agarwal P, James BD, Leurgans SE, Rajan KB, Aggarwal NT, et al. Healthy lifestyle and cognition in older adults with common neuropathologies of dementia. JAMA Neurol 2024;81(3):233-239. [doi: 10.1001/jamaneurol.2023.5491] [Medline: 38315471]
7. Chen J, Zhang F, Zhang Y, Lin Z, Deng K, Hou Q, et al. Trajectories network analysis of chronic diseases among middle-aged and older adults: evidence from the China Health and Retirement Longitudinal Study (CHARLS). BMC Public Health 2024;24(1):559 [FREE Full text] [doi: 10.1186/s12889-024-17890-7] [Medline: 38389048]
8. Zhu X, Wang Z, Yang X, Ning Z. About half of older adults have two or more chronic conditions at the same time: a systematic review and meta-analysis. Front Public Health 2025;13:1680745 [FREE Full text] [doi: 10.3389/fpubh.2025.1680745] [Medline: 41458407]
9. Solis-Navarro L, Gismero A, Fernández-Jané C, Torres-Castro R, Solá-Madurell M, Bergé C. Effectiveness of home-based exercise delivered by digital health in older adults: a systematic review and meta-analysis. Meta-Analysis Age Ageing 2022;51(11):afac243. [doi: 10.1093/ageing/afac243] [Medline: 36346736]
10. Radtke T, Smith S, Nevitt SJ, Hebestreit H, Kriemler S. Physical activity and exercise training in cystic fibrosis. Paediatr Respir Rev 2022;44:47-52 [FREE Full text] [doi: 10.1016/j.prrv.2022.08.001] [Medline: 36402662]
11. Campbell P, Rutten FH, Lee MM, Hawkins NM, Petrie MC. Heart failure with preserved ejection fraction: everything the clinician needs to know. Lancet 2024;403(10431):1083-1092. [doi: 10.1016/S0140-6736(23)02756-3] [Medline: 38367642]
12. Shi Y, Sun J. Integrated bagging-RF learning model for diabetes diagnosis in middle-aged and elderly population. PeerJ Comput Sci 2024;10:e2436. [doi: 10.7717/peerj-cs.2436] [Medline: 39650520]
13. Yuan M, He J, Hu X, Yao L, Chen P, Wang Z, et al. Hypertension and NAFLD risk: insights from the NHANES 2017-2018 and Mendelian randomization analyses. Chin Med J (Engl) 2024;137(4):457-464 [FREE Full text] [doi: 10.1097/CM9.0000000000002753] [Medline: 37455323]
14. Kim H, Lee EM, Ahn SY, Kim KI, Kim HC, Kim JH, et al. The 2022 focused update of the 2018 Korean Hypertension Society Guidelines for the management of hypertension. Clin Hypertens 2023;29(1):11 [FREE Full text] [doi: 10.1186/s40885-023-00234-9] [Medline: 36788612]
15. Chen J, Chen X, Mao R, Fu Y, Chen Q, Zhang C, et al. Hypertension, sleep quality, depression, and cognitive function in elderly: a cross-sectional study. Front Aging Neurosci 2023;15:1051298 [FREE Full text] [doi: 10.3389/fnagi.2023.1051298] [Medline: 36824262]
16. Seidu S, Hambling CE, Kunutsor SK, Topsever PJPCD. Associations of blood pressure with cardiovascular and mortality outcomes in over 2 million older persons with or without diabetes mellitus: a systematic review and meta-analysis of 45 cohort studies. Prim Care Diabetes 2023;17(6):554-567 [FREE Full text] [doi: 10.1016/j.pcd.2023.09.007] [Medline: 37806800]
17. Chen Y, Liu H, Zhang M, Liu Z, Liu Q, Liu S, et al. Predictive value of different long-term (10-year) blood pressure parameters variability in cardiovascular disease: a prospective cohort study. Am J Prev Cardiol 2025;24:101336 [FREE Full text] [doi: 10.1016/j.ajpc.2025.101336] [Medline: 41281726]

XSL•FO
RenderX

18. Wang S, Yang S, Jia W, Cao W, Han K, Liu M, et al. Relationships of lipids profile with health-related quality of life in Chinese centenarians. J Nutr Health Aging 2020;24(4):404-411 [FREE Full text] [doi: 10.1007/s12603-020-1340-5] [Medline: 32242208]
19. Yang Y, Liu Z, Lin B, Huang Y, Deng H, Yang X, et al. The association between new insulin resistance indices and all-cause mortality in elderly patients with diabetes: a prospective cohort study. Diabetol Metab Syndr 2025;17(1):181 [FREE Full text] [doi: 10.1186/s13098-025-01732-6] [Medline: 40442829]
20. Gupta S, Sharma N, Arora S, Verma S. Diabetes: A review of its pathophysiology, and advanced methods of mitigation. Curr Med Res Opin 2024;40(5):773-780. [doi: 10.1080/03007995.2024.2333440] [Medline: 38512073]
21. Hossain MJ, Al-Mamun M, Islam MR. Diabetes mellitus, the fastest growing global public health concern: early detection should be focused. Health Sci Rep 2024;7(3):e2004 [FREE Full text] [doi: 10.1002/hsr2.2004] [Medline: 38524769]
22. Liu P, Zhang Z, Cai Y, Li Z, Zhou Q, Chen Q. Ferroptosis: mechanisms and role in diabetes mellitus and its complications. Ageing Res Rev 2024;94:102201. [doi: 10.1016/j.arr.2024.102201] [Medline: 38242213]
23. Kuzuya T, Nakagawa S, Satoh J, Kanazawa Y, Iwamoto Y, Kobayashi M, et al. Report of the committee on the classification and diagnostic criteria of diabetes mellitus. Diabetes Res Clin Pract 2002;55(1):65-85. [doi: 10.1016/s0168-8227(01)00365-5] [Medline: 11755481]
24. Singh P, Gollapalli K, Mangiola S, Schranner D, Yusuf MA, Chamoli M, et al. Taurine deficiency as a driver of aging. Science 2023;380(6649):eabn9257 [FREE Full text] [doi: 10.1126/science.abn9257] [Medline: 37289866]
25. Thadathil N, Selvarani R, Mohammed S, Nicklas EH, Tran AL, Kamal M, et al. Senolytic treatment reduces cell senescence and necroptosis in sod1 knockout mice that is associated with reduced inflammation and hepatocellular carcinoma. Aging Cell 2022;21(8):e13676 [FREE Full text] [doi: 10.1111/acel.13676] [Medline: 35869934]
26. Pan Y, Ma L. Inflammatory markers associated with physical frailty and cognitive impairment. Aging Dis 2024 Apr 17;16(2):859-875. [doi: 10.14336/AD.2024.0258] [Medline: 38739942]
27. Golbach RD, Kleinenberg-Talsma N, van der Lucht F, Hobbelen JS, Jager-Wittenaar H, Finnema EJ. Understanding frailty and its opposites from community-dwelling older peoples' perspectives: a phenomenological qualitative study. Int J Nurs Stud Adv 2024;7:100238 [FREE Full text] [doi: 10.1016/j.ijnsa.2024.100238] [Medline: 39310677]
28. Song Y, Qian L, Sui J, Greiner R, Li X, Greenshaw AJ, et al. Prediction of depression onset risk among middle-aged and elderly adults using machine learning and Canadian longitudinal study on aging cohort. J Affect Disord 2023;339:52-57. [doi: 10.1016/j.jad.2023.06.031] [Medline: 37380110]
29. Liu D, Wang S, Liu S, Wang Q, Che X, Wu G. Frontiers in sarcopenia: advancements in diagnostics, molecular mechanisms, and therapeutic strategies. Mol Aspects Med 2024;97:101270 [FREE Full text] [doi: 10.1016/j.mam.2024.101270] [Medline: 38583268]
30. de Souto Barreto P, Gonzalez-Bautista E, Bischoff-Ferrari HA, Pelegrim de Oliveira V, Gorga Bandeira de Mello R, Andrieu S, et al. Real-life intrinsic capacity screening data from the ICOPE-Care program. Nat Aging 2024;4(9):1279-1289. [doi: 10.1038/s43587-024-00684-2] [Medline: 39122839]
31. Ugwu LE, Idemudia ES, Onyedibe MCCJSR. Decoding the impact of night/day shiftwork on well-being among healthcare workers. Sci Rep 2024;14(1):10246 [FREE Full text] [doi: 10.1038/s41598-024-60882-1] [Medline: 38702344]
32. Brown RT, Covinsky KE. Moving prevention of functional impairment upstream: is middle age an ideal time for intervention? Womens Midlife Health 2020;6:4 [FREE Full text] [doi: 10.1186/s40695-020-00054-z] [Medline: 32695430]
33. Lyu L, Li Y, Fan X, Seo J, Eunna O. The effect of Tai Chi exercise intervention on sleep quality in community residing older adults: a systematic review and meta-analysis. Eur Geriatr Med 2024;15(2):381-396. [doi: 10.1007/s41999-023-00923-6] [Medline: 38227110]
34. Zhang P, Zhang D, Lu D. The efficacy of Tai Chi for essential hypertension: a systematic review and meta-analysis. Int J Nurs Pract 2024;30(2):e13211. [doi: 10.1111/ijn.13211] [Medline: 37798948]
35. Yin J, Yue C, Song Z, Sun X, Wen X. The comparative effects of Tai Chi versus non-mindful exercise on measures of anxiety, depression and general mental health: a systematic review and meta-analysis. J Affect Disord 2023;337:202-214. [doi: 10.1016/j.jad.2023.05.037] [Medline: 37244543]
36. Sun F, Li L, Wen X, Xue Y, Yin J. The effect of Tai Chi/Qigong on depression and anxiety symptoms in adults with cancer: a systematic review and meta-regression. Complement Ther Clin Pract 2024;56:101850. [doi: 10.1016/j.ctcp.2024.101850] [Medline: 38626582]
37. Chen Y, Qin J, Tao L, Liu Z, Huang J, Liu W, et al. Effects of Tai Chi Chuan on cognitive function in adults 60 years or older with type 2 diabetes and mild cognitive impairment in china: a randomized clinical trial. JAMA Netw Open 2023;6(4):e237004 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.7004] [Medline: 37022680]
38. Wang Y, Luo B, Wu X, Li X, Liao S. Comparison of the effects of Tai Chi and general aerobic exercise on weight, blood pressure and glycemic control among older persons with depressive symptoms: a randomized trial. BMC Geriatr 2022;22(1):401 [FREE Full text] [doi: 10.1186/s12877-022-03084-6] [Medline: 35525971]
39. Wen J, Su M. A randomized trial of Tai Chi on preventing hypertension and hyperlipidemia in middle-aged and elderly patients. Int J Environ Res Public Health 2021;18(10):5480 [FREE Full text] [doi: 10.3390/ijerph18105480] [Medline: 34065454]

XSL•FO
RenderX

40. Bradshaw A, Walker L, Borgstrom E, Burke S. Group-based Tai Chi as therapy for alleviating experiences of social death in people with advanced, incurable disease: an ethnographic study. Qualitative Research in Sport, Exercise and Health 2021;14(1):84-100. [doi: 10.1080/2159676x.2021.1879918]
41. Liu H, Gong H, Chen P, Zhang L, Cen H, Fan YJM. Biomechanical effects of typical lower limb movements of Chen-style Tai Chi on knee joint. Med Biol Eng Comput 2023;61(11):3087-3101. [doi: 10.1007/s11517-023-02906-y] [Medline: 37624535]
42. Jin W, Tao Y, Wang C, Wang L, Ao X, Su M, et al. Infrared imageries of human body activated by tea match the hypothesis of meridian system. Phenomics 2023;3(5):502-518. [doi: 10.1007/s43657-022-00090-x] [Medline: 37881315]
43. Huang J, Wang D, Wang JJEBC. Clinical evidence of Tai Chi exercise prescriptions: a systematic review. Evid Based Complement Alternat Med 2021;2021:5558805 [FREE Full text] [doi: 10.1155/2021/5558805] [Medline: 33777155]
44. Wang H, Pei Z, Liu Y. Effects of square dance exercise on cognitive function in elderly individuals with mild cognitive impairment: the mediating role of balance ability and executive function. BMC Geriatr 2024;24(1):156 [FREE Full text] [doi: 10.1186/s12877-024-04714-x] [Medline: 38360628]
45. Li Y, Chen S, Li F, Li B, Li T, Deng S, et al. The cross-sectional and longitudinal associations between fruit intake and cognitive performance: the Chinese healthy dance study. Maturitas 2025;191:108146. [doi: 10.1016/j.maturitas.2024.108146] [Medline: 39536657]
46. Xu J, Huang Y, Li L, Zhang J, Li C, Liu M, et al. An RCT META analysis based on the efficacy of Tai Chi exercise therapy on blood pressure and blood lipids in patients with essential hypertension. Front Cardiovasc Med 2025;12:1506912 [FREE Full text] [doi: 10.3389/fcvm.2025.1506912] [Medline: 40873614]
47. Teng Y, Luo W, Zhou J, Wang M, Wu J, Nie X. Clinical research on Tai Chi: a review of health benefits. Acupuncture and Herbal Medicine 2024;4(3):320-345 [FREE Full text] [doi: 10.1097/HM9.0000000000000097]
48. Li X, Chang P, Wu M, Jiang Y, Gao Y, Chen H, et al. Effect of Tai Chi vs aerobic exercise on blood pressure in patients with prehypertension: a randomized clinical trial. JAMA Netw Open 2024;7(2):e2354937 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.54937] [Medline: 38335001]
49. Ouyang T, Qu Y, Du X, Fan M, Wang Y. The relationship between square dance exercise and quality of life in middle-aged and older adults: chain mediated effects of negative emotions and attitudes toward aging. Front Public Health 2024;12:1470624 [FREE Full text] [doi: 10.3389/fpubh.2024.1470624] [Medline: 39502812]
50. Wollesen B, Herden M, Lamberti N, Giannaki CD. Defining and reporting exercise intensity in interventions for older adults: a modified Delphi process. Eur Rev Aging Phys Act 2024;21(1):3 [FREE Full text] [doi: 10.1186/s11556-024-00337-8] [Medline: 38302886]
51. Blodgett JM, Ahmadi MN, Atkin AJ, Pulsford RM, Rangul V, Chastin S, et al. Device-Measured 24-hour movement behaviors and blood pressure: a 6-part compositional individual participant data analysis in the proPASS consortium. Circulation 2025;151(2):159-170. [doi: 10.1161/circulationaha.124.069820]
52. Liu W, Yang C, Lei F, Huang X, Cai J, Chen S, et al. Major lipids and lipoprotein levels and risk of blood pressure elevation: a Mendelian randomisation study. EBioMedicine 2024;100:104964 [FREE Full text] [doi: 10.1016/j.ebiom.2023.104964] [Medline: 38181703]
53. Bolat S, Ertürk Zararsız G, Doğan K, Kochan N, Yerlitaş SI, Cephe A, et al. Concordance of LDL-C estimating equations with direct enzymatic measurement in diabetic and prediabetic subjects. J Clin Med 2023;12(10):3570 [FREE Full text] [doi: 10.3390/jcm12103570] [Medline: 37240676]
54. Chen J, Liu X, Zheng G, Feng W, Wang P, Gao J, et al. Detection of glucose based on noble metal nanozymes: mechanism, activity regulation, and enantioselective recognition. Small 2023;19(8):e2205924. [doi: 10.1002/smll.202205924] [Medline: 36509680]
55. Qin Y, Cai ML, Jin HZ, Huang W, Zhu C, Bozec A, et al. Age-associated B cells contribute to the pathogenesis of rheumatoid arthritis by inducing activation of fibroblast-like synoviocytes via TNF-α-mediated ERK1/2 and JAK-STAT1 pathways. Ann Rheum Dis 2022;81(11):1504-1514. [doi: 10.1136/ard-2022-222605] [Medline: 35760450]
56. Latham RM, Kieling C, Arseneault L, Kohrt BA, Moffitt TE, Rasmussen LJ, et al. Longitudinal associations between adolescents' individualised risk for depression and inflammation in a UK cohort study. Brain Behav Immun 2022;101:78-83 [FREE Full text] [doi: 10.1016/j.bbi.2021.12.027] [Medline: 34990745]
57. Branco MG, Mateus C, Capelas ML, Pimenta N, Santos T, Mäkitie A, et al. Bioelectrical impedance analysis (BIA) for the assessment of body composition in oncology: a scoping review. Nutrients 2023;15(22):4792 [FREE Full text] [doi: 10.3390/nu15224792] [Medline: 38004186]
58. Charlton PH, Allen J, Bailón R, Baker S, Behar JA, Chen F, et al. The 2023 wearable photoplethysmography roadmap. Physiol Meas 2023;44(11):111001 [FREE Full text] [doi: 10.1088/1361-6579/acead2] [Medline: 37494945]
59. Rennie KL, Hennings SJ, Mitchell J, Wareham NJJM. Estimating energy expenditure by heart-rate monitoring without individual calibration. Med Sci Sports Exerc 2001;33(6):939-945. [doi: 10.1097/00005768-200106000-00013] [Medline: 11404659]
60. Pandey A, Kitzman DW, Nelson MB, Pastva AM, Duncan P, Whellan DJ, et al. Frailty and effects of a multidomain physical rehabilitation intervention among older patients hospitalized for acute heart failure: a secondary analysis of a randomized clinical trial. JAMA Cardiol 2023;8(2):167-176 [FREE Full text] [doi: 10.1001/jamacardio.2022.4903] [Medline: 36598761]

XSL•FO
RenderX

61. Sperling RA, Donohue MC, Rissman RA, Johnson KA, Rentz DM, Grill JD, et al. Amyloid and Tau prediction of cognitive and functional decline in unimpaired older individuals: longitudinal data from the A4 and LEARN studies. J Prev Alzheimers Dis 2024;11(4):802-813 [FREE Full text] [doi: 10.14283/jpad.2024.122] [Medline: 39044488]
62. Mehmet B, Gillard S, Jayasena C, Llahana S. Association between domains of quality of life and patients with Klinefelter syndrome: a systematic review. Eur J Endocrinol 2022;187(2):S21-S34 [FREE Full text] [doi: 10.1530/EJE-21-1239] [Medline: 35639859]
63. Skevington SM, Schick-Makaroff K, Rowland C, Molzahn A, WHOQOL Group T. Women's environmental quality of life is key to their overall quality of life and health: global evidence from the WHOQOL-100. PLoS One 2024;19(10):e0310445 [FREE Full text] [doi: 10.1371/journal.pone.0310445] [Medline: 39356697]
64. Granet J, Peyrusqué E, Ruiz F, Buckinx F, Abdelkader LB, Dang-Vu T, et al. Online physical exercise intervention in older adults during lockdown: can we improve the recipe? Aging Clin Exp Res 2023;35(3):551-560 [FREE Full text] [doi: 10.1007/s40520-022-02329-z] [Medline: 36635450]
65. Celler B, Yong A, Rubenis I, Butlin M, Argha A, Rehan R, et al. Accurate detection of Korotkoff sounds reveals large discrepancy between intra-arterial systolic pressure and simultaneous noninvasive measurement of blood pressure with brachial cuff sphygmomanometry. J Hypertens 2024;42(5):873-882. [doi: 10.1097/HJH.0000000000003651] [Medline: 38230626]
66. Liu W, Lin Y, Mündel T, Chou C, Liao Y. Effects of acute interval exercise on arterial stiffness and cardiovascular autonomic regulatory responses: a narrative review of potential impacts of aging. Front Cardiovasc Med 2022;9:864173 [FREE Full text] [doi: 10.3389/fcvm.2022.864173] [Medline: 35620510]
67. Tsuda K, Hayashi E, Kamiya K, Kudo A, Kakihana H, Nakayama S, et al. Effects of interval-walking training on blood pressure in community-dwelling Japanese older adults. J Sports Med Phys Fitness 2023;63(3):492-502. [doi: 10.23736/S0022-4707.22.14263-5] [Medline: 36239289]
68. Xu S, Tu S, Hao X, Chen X, Pan D, Liao W, et al. Exercise, nutrition, and neuromuscular electrical stimulation for sarcopenic obesity: a systematic review and meta-analysis of management in middle-aged and older adults. Nutrients 2025;17(9):1504 [FREE Full text] [doi: 10.3390/nu17091504] [Medline: 40362811]
69. Mann S, Beedie C, Jimenez A. Differential effects of aerobic exercise, resistance training and combined exercise modalities on cholesterol and the lipid profile: review, synthesis and recommendations. Sports Med 2014;44(2):211-221 [FREE Full text] [doi: 10.1007/s40279-013-0110-5] [Medline: 24174305]
70. Li B, Zhou X, Liu Y, Zhang Y, Mu Y. Remnant cholesterol is more strongly associated with arterial stiffness than traditional lipids and lipid ratios in the general Chinese population. J Atheroscler Thromb 2024;31(5):587-602 [FREE Full text] [doi: 10.5551/jat.64146] [Medline: 38171806]
71. Li Y, Zhai Q, Li G, Peng W. Effects of different aerobic exercises on blood lipid levels in middle-aged and elderly people: a systematic review and Bayesian network meta-analysis based on randomized controlled trials. Healthcare (Basel) 2024 Jun 30;12(13):1309 [FREE Full text] [doi: 10.3390/healthcare12131309] [Medline: 38998844]
72. Kurdziel LBF, McDevitt L, Hardway C. Acute effects of slow-paced breathing on emotion regulation: a pilot study. Psychol Rep 2025:332941251329856. [doi: 10.1177/00332941251329856] [Medline: 40179113]
73. Freiberger E, Fabbietti P, Corsonello A, Lattanzio F, Sieber C, Tap L, SCOPE investigators. Short physical performance battery is not associated with falls and injurious falls in older persons: longitudinal data of the SCOPE project. Eur Geriatr Med 2024;15(3):831-842 [FREE Full text] [doi: 10.1007/s41999-024-00941-y] [Medline: 38416398]
74. Zhai Y, Li F, Lin C, Wu F, Qiu H, Li J, et al. The mediating role of body surface area-adjusted basal metabolic rate: effects of low muscle mass and central obesity on cognitive impairment in Chinese patients with type 2 diabetes mellitus. Front Endocrinol (Lausanne) 2024;15:1513035 [FREE Full text] [doi: 10.3389/fendo.2024.1513035] [Medline: 39926391]
75. Wang Y, Tian J, Yang Q. Tai Chi exercise improves working memory capacity and emotion regulation ability. Front Psychol 2023;14:1047544 [FREE Full text] [doi: 10.3389/fpsyg.2023.1047544] [Medline: 36874821]
76. Li R, Yan Q, Qu Y, Wang Y. Square dance, loneliness, quality of life, and attitude toward aging in middle-aged and older women in China. Front Public Health 2025;13:1508556 [FREE Full text] [doi: 10.3389/fpubh.2025.1508556] [Medline: 40061464]
77. Li Q, Zhao Z, Yang C, Chen Z, Yin W. Sedentary behavior and risk of depression in older adults: a systematic meta-analysis. Iran J Public Health 2024;53(2):293-304 [FREE Full text] [doi: 10.18502/ijph.v53i2.14914] [Medline: 38894832]
78. Riebe D, Ehrman J, Liguori G, Magal M. ACSM's Guidelines for Exercise Testing and Prescription. United Statesn: American College of Sports Medicine; 2018.
79. Boulay P, Ghachem A, Poirier P, Sigal R, Kenny GP. Assessment of maximum heart rate prediction equations in adults at low and high risk of cardiovascular disease. Med Sci Sports Exerc 2025;57(1):60-69. [doi: 10.1249/MSS.0000000000003540] [Medline: 39160700]
80. Nelson M, Rejeski W, Blair S, Duncan P, Judge J, King A, et al. Physical activity and public health in older adults: recommendation from the American college of sports medicine and the American heart association. Med Sci Sports Exerc 2007;39(8):1435-1445. [doi: 10.1249/mss.0b013e3180616aa2] [Medline: 17762378]

XSL•FO
RenderX

81. Bower JE, Lacchetti C, Alici Y, Barton DL, Bruner D, Canin BE, et al. Management of fatigue in adult survivors of cancer: ASCO-society for integrative oncology guideline update. J Clin Oncol 2024;42(20):2456-2487. [doi: 10.1200/JCO.24.00541] [Medline: 38754041]
82. Liang I, Perkin O, Williams S, McGuigan P, Thompson D, Western MJ. The efficacy of 12-week progressive home-based strength and Tai-Chi exercise snacking in older adults: a mixed-method exploratory randomised control trial. J Frailty Aging 2024;13(4):572-581. [doi: 10.14283/jfa.2024.32] [Medline: 39574284]
83. Chen X, Han H, Jiang T, Cai G. Effects of Tai Chi on executive function, single-leg dynamic balance, and brain functional connectivity in older adults. Sci Rep 2025;15(1):11838 [FREE Full text] [doi: 10.1038/s41598-025-93321-w] [Medline: 40195410]
84. Xu F, Soh KG, Chan YM, Bai XR, Qi F, Deng N. Effects of Tai Chi on postural balance and quality of life among the elderly with gait disorders: a systematic review. PLoS One 2023;18(9):e0287035 [FREE Full text] [doi: 10.1371/journal.pone.0287035] [Medline: 37768953]
85. Ou KL, Wong MYC, Chung PK, Chui KYK. Effect of square dance interventions on physical and mental health among Chinese older adults: a systematic review. Int J Environ Res Public Health 2022;19(10):6181 [FREE Full text] [doi: 10.3390/ijerph19106181] [Medline: 35627716]

Abbreviations

ΔHR: change in heart rate
ADL: activity of daily living
BMR: Basal metabolic rate
CONSORT: Consolidated Standards of Reporting Trials
CTC12: 12-form Chen-style Tai Chi
DBP: diastolic blood pressure
DM: diabetes mellitus
FPG: fasting plasma glucose
HDL-C: high-density lipoprotein cholesterol
HR: heart rate
IL-6: interleukin-6
LDL-C: low-density lipoprotein cholesterol
PP: pulse pressure
SBP: systolic blood pressure
SPPB: Short Physical Performance Battery
TC: total cholesterol
TC24: 24-form Tai Chi
WHOQOL: World Health Organization Quality of Life

Edited by Q Fan; submitted 05.Jul.2025; peer-reviewed by S Vengadassalapathy, Y Fan, MC Espada; comments to author 15.Dec.2025; revised version received 31.Mar.2026; accepted 02.Apr.2026; published 23.Apr.2026.

Please cite as:
Qiu J, Wang J, Song X, Shu W, Kim SM
Dose-Related Effects of Different Tai Chi Styles Versus Traditional Community-Based Exercises on Cardiometabolic Health and Physical Function in Middle-Aged and Older Adults: Randomized Controlled Trial
JMIR Aging 2026;9:e80125
URL: https://aging.jmir.org/2026/1/e80125
doi:10.2196/80125

©Jiadong Qiu, Jian Wang, Xiongying Song, Wanyu Shu, Sung Min Kim. Originally published in JMIR Aging (https://aging.jmir.org), 23.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Interactive Vision–Based 3D Augmented Reality System for In-Home Rehabilitation in Older Adults Following Total Joint Replacement: Cross-Sectional Survey of Patients and Service Providers

Adriana C Salvia[1], MEd; Siona Phadke[1], BSc; Manuela Kunz[2], PhD; Afolasade Fakolade[1], PhD

[1]School of Rehabilitation Therapy, Queen's University, Louise D Acton Building, 31 George Street, Kingston, ON, Canada

[2]Digital Technologies Research Center, National Research Council Canada, Ottawa, ON, Canada

Corresponding Author:
Afolasade Fakolade, PhD
School of Rehabilitation Therapy, Queen's University, Louise D Acton Building, 31 George Street, Kingston, ON, Canada

Abstract

Background: Completing home exercise programs remains a significant challenge for older adults after total joint replacement (TJR), often resulting in suboptimal rehabilitation outcomes. Augmented reality (AR) technologies offer a promising approach to addressing limitations associated with traditional paper-based home exercise programs.

Objective: This study aimed to identify priorities for design features and practical or logistical considerations to inform the future design of a 3D AR system (3D ARS) to support home-based rehabilitation for older adults following TJR.

Methods: A cross-sectional online survey was administered via Qualtrics. The survey included items across four domains: sociodemographic and clinical characteristics, proposed design features, practical considerations, and logistical details, and potential barriers to access and use of the 3D ARS. Descriptive statistics (frequencies and percentages) were computed, and the Mann-Whitney *U* test was used to examine differences in feature preferences between older adults and service providers.

Results: A total of 145 participants (56 older adults with TJR and 89 service providers) completed the survey (response rate 88.4%). Older adults and service providers shared similar priorities for several design features of the 3D ARS, including embedding individualized exercise programs, providing customized feedback on exercise progress, and including background music during exercise (all $P>.05$). However, significant differences emerged for other features. Older adults ranked the ability to receive positive, encouraging messages during exercises ($U=1525$, $P<.001$), voice activation prompts ($U=1145$, $P<.001$), and the ability to pause and resume exercises at any time ($U=1843.50$, $P=.01$) as higher priorities than service providers. Most older adults preferred real-time visual prompts to guide joint range of motion during exercise (42/56, 75%), whereas most service providers preferred chart-based progress reports (63/89, 71%). Participants further recommended accessing the system through physiotherapy clinics, initiating use within the first month following surgery, and maintaining access for 6 - 12 months.

Conclusions: Findings identify key user-informed priorities for the design of AR-based rehabilitation technologies for older adults following TJR, including personalized exercise programming, accessible system interfaces, and timely feedback mechanisms. Incorporating these priorities into future system development may support scalable models of home-based rehabilitation that complement conventional physiotherapy care.

(JMIR Aging 2026;9:e87572) doi:10.2196/87572

KEYWORDS

arthroplasty; older adults; exercise prescription; post-exercise recovery techniques; telerehabilitation; vision-based interactive tools; gamification; stakeholder participation

Introduction

Total joint replacement (TJR) is a well-established surgical intervention for older adults with advanced osteoarthritis when conservative management strategies are no longer effective [1]. Advances in surgical techniques, perioperative care, and enhanced recovery pathways have shortened hospital stays, and most patients (74% - 96%) are now discharged directly home after surgery [2-4]. After discharge, patients are typically prescribed home exercise programs to restore joint range of motion, muscular strength, and functional mobility [5]. However, maintaining adherence to home exercise programs remains challenging for many older adults after TJR, often resulting in delayed functional recovery and reduced quality of life [6,7]. Commonly reported barriers include limited access to supervision and feedback once patients return home [8,9],

XSL•FO
RenderX

difficulties recalling or correctly performing prescribed exercises [10], low motivation [11,12], and challenges integrating structured exercise into daily routines [9,11]. Additional contextual factors, such as variability in home environments [13,14], limited social support [15], and geographic or transportation barriers to outpatient physiotherapy services [16], may further constrain engagement in rehabilitation for some older adults after TJR. Conventional paper-based home exercise programs, while inexpensive and widely used, often provide static instructions that lack personalization, interactivity, and real-time performance feedback, which may further limit motivation and adherence [17,18].

Emerging augmented reality (AR) technologies present a promising avenue for addressing some of these challenges. AR systems overlay virtual images and information onto the real-world environment, allowing users to interact with digital content while remaining grounded in their physical surroundings [19,20]. In rehabilitation contexts, AR has been explored as a means of delivering guided exercise, enhancing motor learning through visual and auditory feedback, and supporting task-specific practice within meaningful environments [21-24]. This approach contrasts with fully immersive virtual reality (VR), which replaces the real-world environment entirely and may not be suitable or acceptable for all older adults due to issues, such as disorientation or "cybersickness" and physical discomfort (eg, neck pain) sometimes caused by VR headsets [25,26].

While many contemporary AR applications are delivered using smartphones or head-mounted displays, including optical see-through devices, these platforms may pose usability, cost, and comfort barriers for some older adults, particularly in unsupervised home settings [27,28]. In contrast, AR delivered through familiar nonwearable displays, such as tablets or televisions equipped with cameras, depth sensors, or computer-vision capabilities, may represent a more accessible and acceptable configuration for home-based rehabilitation [21,22,29,30]. In these systems, digital overlays, avatars, or visual cues are superimposed onto a live or screen-based representation of the user's environment, preserving core characteristics of AR while avoiding the need for head-mounted equipment. Although the use of AR glasses as an alternative to the TV or monitor offers a more immersive experience, they can also trigger severe symptoms of visually induced motion sickness and dizziness [31] and might therefore not be suitable for safe use for older adults in a home environment. For older adults, leveraging familiar household devices may reduce cognitive and technical burden, enhance perceived ease of use, and support sustained engagement with rehabilitation technologies [27,32].

One configuration that may support accessible home deployment involves plug-and-play camera-based AR systems that connect directly to household televisions or monitors. Similar systems are already widely used in motion-based gaming. For example, NEX Team Inc (San Jose, CA, USA) offers a small camera-based console (Nex Playground) that connects to a television via HDMI, allowing one to 4 players to interact with virtual games using body movements. Similarly, the MetFutAR Motion Gaming Console (MetaFuture Technologies Inc) connects to televisions via HDMI and uses built-in motion-sensing cameras to track body movements during gameplay. These systems overlay digital elements onto the real-world environment, effectively transforming living rooms into interactive play or activity spaces. Easy installation and relatively low cost are commonly promoted as key advantages of these technologies. These characteristics suggest that similar plug-and-play configurations may offer a practical and scalable platform for delivering AR-supported home-based rehabilitation.

AR systems may enhance the home exercise experience by providing interactive visual cues, gamified tasks, or virtual coaching, thereby increasing engagement and enjoyment. These systems can incorporate motion tracking and sensor-based feedback to provide immediate information on exercise performance, such as identifying errors in joint range of motion or movement quality, capabilities that paper exercise booklets cannot offer [17,33]. AR systems also create opportunities for remote or asynchronous monitoring by clinicians, enabling exercise feedback and progression without requiring frequent in-person visits. By supporting individualized exercise prescription and reducing reliance on in-person visits, AR-enabled home rehabilitation may help address geographic and transportation barriers that can limit access to post-TJR physiotherapy services [34].

Existing rehabilitation research on AR systems includes those targeting improved balance and muscle strengthening, falls prevention, upper limb functionality, and enhanced cognition [35]. Qualitative research has also begun to examine stakeholder perspectives on AR tools designed for home-based use. For example, Fields and colleagues explored the acceptability and appropriateness of an AR home assessment tool, highlighting the importance of user-centered design and stakeholder engagement when developing AR-supported home technologies for older adults [36]. A recent pilot trial of an interactive AR exercise system for knee and hip arthroplasty found that patients and therapists preferred the system over standard paper-based home exercise instructions [17]. Study participants highlighted the system's potential to improve exercise accuracy and adherence to home exercise programs. While AR tools show considerable promise for enhancing TJR rehabilitation, limited engagement of key stakeholders in co-designing, implementing, and evaluating these technologies has hindered their translation into routine rehabilitation care [37,38].

The successful integration of any AR tool into routine rehabilitation care depends on careful alignment with end user needs and real-world clinical workflows [39]. Both older adults and their health care service providers (eg, physiotherapists) are key stakeholders whose perspectives could guide the design of an AR tool to support TJR rehabilitation. Engaging patients and health care service providers can illuminate desirable features, acceptable interfaces, and potential barriers, ensuring that the AR tool is developed to meet users' needs and increases the potential for successful implementation in routine rehabilitation practice. This responds to a gap identified in the ARS literature: that heightened patient understanding of the features and functionality of an AR system is vital to creating beneficial, personalized rehabilitation tools [35].

XSL•FO
RenderX

Therefore, the objective of this study was to identify priorities for design features and practical or logistical considerations to inform the future design of a 3D AR system (3D ARS) to support home-based rehabilitation for older adults following TJR.

Methods

Overview

We began this research program by conducting a qualitative exploration of older adults' perspectives, with particular focus on their pre- and postoperative experiences with TJR [40]. In-depth interviews and focus groups were conducted with 42 older adults with TJR and 4 unrelated family caregivers to explore their everyday use of technology, along with the potential role of AR tools to support home-based rehabilitation post-TJR [19]. Two key themes were identified: (1) dreaming up possibilities related to features, considerations, and logistics that could be incorporated into an ARS, such as providing customized feedback, including motivating messages while performing the exercises, and ensuring timely access to the ARS post-TJR surgery; and (2) being pragmatic about possible barriers to access and use of an ARS system, ranging from cost, confidence, and motivational challenges to concerns about accessing suitable home exercise equipment to maximize recovery while using the ARS [19].

Building on the qualitative findings, the themes informed the development of the survey in this study, which focused on prioritizing system features and outlining practical and logistical considerations to guide the design of a proposed 3D ARS to support home-based exercise programming for older adults after TJR.

As part of the survey development process, an expert-by-experience advisory panel (2 older adults with lived experience and 2 service providers) provided feedback on the survey items. The panel members were sent a draft of the survey and asked to review and complete it independently. After the draft surveys were completed and returned to the research office, a virtual meeting was convened to discuss item clarity, wording, readability, ease of access, and, importantly, overall relevance to older adults. Based on this feedback, revisions were made to refine the wording of survey items and improve clarity. The order of items was also revised based on the feedback to improve readability and flow. The updated survey was subsequently shared with panel members. No additional revisions were suggested at this stage.

Survey participants were recruited through flyers distributed by professional membership organizations (eg, the Canadian Physiotherapy Association, the Canadian Society for Exercise Physiology), organizations that offer support to people with arthritis (eg, the Arthritis Society of Canada), local libraries, seniors' centers, and via emails to participants in our previous research who had consented to be re-contacted for future studies.

The inclusion criteria comprised individuals who self-identified as: (1) an older adult aged ≥65 years with a TJR within the last 6 months—to capture those with relevant post-TJR rehabilitation experience, or (2) a service provider (defined as a health care professional or community service provider with paid or volunteer experience working with older adults with TJR in formal care or community-based settings). All participants were further required to be residents of Canada.

Interested individuals scanned a QR code on the recruitment flyer to access the survey. After an online screening via Qualtrics, interested individuals were provided with detailed information about the study. All individuals were required to provide informed consent before completing the survey online. Each question enabled voluntary responses. Individuals who completed a survey could opt in to receive a $40 CAD (US $30) e-gift card. All survey data were collected using an institutional Qualtrics account (Silver Lake, Seattle, Washington, 2024). The survey took between 30 - 45 minutes to complete.

Data Collection

The survey consisted of 4 sections. Section 1 collected information on general sociodemographic and/or clinical characteristics. Sections 2 to 4 included items derived from our previous qualitative study [40] and were designed to gather information to guide the design and implementation of the 3D ARS.

In Section 2, participants were presented with a list of potential design features and asked to prioritize them for inclusion in the 3D ARS. Participants completed a forced rank-ordering task by dragging and dropping each feature into a ranked list from extremely important (rank=1) to least important (rank=14). Each feature was assigned a unique rank, and ties were not permitted. This approach produces relative, ordinal ranking data rather than independent Likert-type ratings.

Sections 3 and 4 consisted of structured questions with predefined response options using a "select all that apply" format, allowing participants to identify multiple relevant responses. Section 3 asked participants to indicate practical considerations and logistical preferences for the proposed AR system (eg, preferred feedback format, timing of motivational messages, and access pathways). Section 4 presented a list of potential barriers to AR system use and asked participants to select all options that applied to their experience or perspective.

Participants were also provided with a narrative description and a graphic depiction of the proposed 3D ARS, along with a glossary of key terms (eg, AR), to provide context and support informed responses.

Data Analysis

Data were downloaded from Qualtrics and imported into SPSS version 29 (IBM SPSS Statistics) for analysis. Descriptive statistics, including frequencies and percentages, were used to summarize participant characteristics. Items using a "select all that apply" format were also analyzed descriptively using frequencies and percentages.

For survey items involving rankings of desired system features, differences between older adults with TJR and service providers were examined using Mann-Whitney *U* tests. This nonparametric test was selected because the ranking data were ordinal and the 2 groups were independent. Statistical significance was set at $P<.05$.

XSL•FO
RenderX

Since the survey allowed participants to skip individual questions, analyses were conducted using available-case analysis for each item. As a result, the number of responses varies slightly across analyses depending on the number of participants who completed each question.

Ethical Considerations

Ethical approval of the online survey was obtained from the Queen's University Health Sciences and Affiliated Teaching Hospitals Research Board (REH-858 - 23). All participants provided informed e-consent before completing the survey. The survey was anonymous and no identifiable patient-related information was collected. Participants had the right to withdraw from the study by not completing the survey or by contacting the research team after submission to request that their data be withdrawn, without any adverse consequences. All study procedures were in accordance with the relevant guidelines/regulations, including data privacy and the Declaration of Helsinki. Participants who completed the survey and who opted in received a $40 CAD (US $30) e-gift card as a token of appreciation.

Results

Sample Characteristics

Table 1 summarizes the characteristics of the 145 participants who provided demographic information (out of 164 who activated the survey; 88.4% response rate). The sample consisted of 56 older adults with TJR and 89 service providers. Across the sample, most participants were male (68/145, 46.9%). Approximately one-third of the sample lived or worked in Ontario (47/145, 32.4%). Among older adults, left knee replacement was the most reported type of TJR (22/56, 39%). Older adults had a high level of education, with about one-third (22/56, 39%) holding a master's degree. About one-third of the service providers were kinesiologists or exercise specialists (27/89, 30%). Most service providers were practicing in community-based settings (52/89, 58%).

XSL•FO
RenderX

Table . Participants' characteristics (n=145)[a].

Variable	Older adult (n=56)	Service providers (n=89)
Gender, n (%)		
Male	28 (50)	40 (45)
Female	28 (50)	18 (20)
Nonbinary	—[b]	11 (12)
Transgender man	—	12 (14)
Transgender woman	—	7 (8)
Gender fluid	—	1 (1)
Age (years), n (%)		
60 - 65	10 (18)	—
66 - 70	19 (34)	—
71 - 75	14 (25)	—
76 - 80	9 (16)	—
81 - 85	3 (5)	—
86 - 90	1 (2)	—
Province of residence, n (%)		
Alberta	3 (5)	10 (11)
British Columbia	11 (20)	9 (10)
Manitoba	2 (4)	20 (23)
New Brunswick	1 (2)	20 (23)
Newfoundland and Labrador	—	4 (5)
Nova Scotia	3 (5)	5 (6)
Ontario	36 (64)	11 (12)
Prince Edward Island	—	3 (3)
Quebec	—	6 (7)
Saskatchewan	—	1 (1)
Type of community, n (%)		
Rural area (population less than 1000)	2 (4)	5 (6)
Small center (population 1000 - 29,000)	11 (20)	29 (33)
Medium center (population 30,000 - 99,000)	9 (16)	45 (51)
Large urban center (population 100,000 or more)	34 (60)	4 (5)
Prefer not to disclose	—	3 (3)
Highest education, n (%)		
High school or general educational development	5 (9)	—
Technical or trade school	2 (4)	—
College	7 (13)	—
Bachelor's degree	14 (25)	—
Master's degree	22 (39)	—
Doctoral degree	6 (11)	—
Family income (US $)		
20,000 - 49,999	13 (32)	—

XSL•FO
RenderX

Variable	Older adult (n=56)	Service providers (n=89)
50,000 - 99,999	15 (27)	—
100,000 - 149,999	8 (14)	—
150,000 or more	11 (20)	—
Prefer not to disclose	9 (16)	—
Type of joint replacement within the last 5 years, n (%)		
Left hip primary replacement surgery	7 (13)	—
Left hip revision surgery	2 (4)	—
Right hip primary replacement surgery	12 (21)	—
Right hip revision surgery	7 (13)	—
Left knee primary replacement	22 (39)	—
Left knee revision surgery	2 (4)	—
Right knee primary replacement surgery	19 (34)	—
Right knee revision surgery	1 (2)	—
Left ankle primary replacement surgery	4 (7)	—
Left ankle revision surgery	4 (7)	—
Right ankle primary replacement surgery	8 (14)	—
Right ankle revision surgery	2 (4)	—
Family caregiver during postoperative care, n (%)		
No one	1 (2)	—
Spouse	30 (54)	—
Friend	2 (4)	—
Offspring	10 (18)	—
Extended family members	10 (18)	—
Other	3 (5)	—
Service provider profession, n (%)		
Physiotherapist	—	10 (11)
Occupational therapist	—	21 (24)
Kinesiology or exercise specialist	—	27 (30)
Recreational therapist	—	8 (9)
Physician or physiatrist	—	1 (1)
Nurse	—	6 (7)
Social worker	—	6 (7)
Personal support worker	—	9 (10)
Prefer not to disclose	—	1 (1)
Types of technologies regularly used, n (%)		
Smartphone	51 (91)	75 (84)
SmartTV	35 (63)	82 (92)
Conventional TV	15 (27)	10 (11)
Smartwatch	21 (38)	17 (19)
Tablet	35 (63)	29 (33)
Wireless earphones	25 (45)	37 (42)
Laptop computer	35 (63)	31 (35)
Desktop computer	26 (46)	7 (8)

XSL•FO
RenderX

Variable	Older adult (n=56)	Service providers (n=89)
Other	4 (7)	—

[a]Percentages do not add up to 100% because participants could select all options that apply.

[b]Not applicable.

Key Features

The Mann-Whitney *U* test revealed no significant differences between groups in rankings of "capability to embed individualized/tailored recommended exercise options in the 3D ARS" (*U*=2134.50, *Z*=−1.18; *P*=.24), "capability of the 3D ARS to deliver customized feedback on progress with recommended exercises" (*U*=2045, *Z*=−1.60; *P*=.11), or "capability to include background music during exercise" (*U*=2118.50, *Z*=−1.20; *P*=.23). However, older adults and service providers differed in their rankings of "capability of the 3D ARS to deliver positive encouraging messages while performing recommended exercises'" (*U*=1525, *Z*=−3.93; *P*<.001), "capability to activate the 3D ARS using voice activation prompts" (*U*=1145, *Z*=−5.39; *P*<.001), and "ability to pause and resume 3D ARS use at any time" (*U*=1843.50, *Z*=−2.51; *P*=.01). Descriptive statistics for the overall rankings of desired features are presented in Table 2.

XSL•FO
RenderX

Table . Differences between older adults and service providers in ranking of desired features for the 3D augmented reality system.

Design features (rank ordered based on older adults' and service providers' responses)[a] (n=142)	Older adults (n=55), n (%)	Median (IQR)	Service providers (n=87)	Median (IQR)	Mann-Whitney *U* test	Z statistic	*P* value
Individualized or tailored exercise options	40 (71)	1 (1-3)	80 (90)	2 (2-2)	2134.50	−1.18	.24
Positive encouragement delivered through the system while performing recommended exercises	30 (54)	2 (1-3)	77 (87)	2 (1-3)	1525	−3.93	<.001
Customized feedback on progress with recommended exercises	29 (52)	2 (1-3)	13 (15)	3 (0-3)	2045	−1.60	.11
Ability to pause and continue system use at any time	28 (50)	2 (1-3)	19 (21)	3 (3-3)	1843.50	−2.51	.01
System availability 24/7	21 (38)	3 (1-5)	10 (11)	4 (4-4)	1971	−1.87	.06
Embedded technical support to problem-solve issues or difficulties with system use	19 (34)	3 (2-4)	13 (15)	5 (4-5)	1346.50	−4.59	<.001
Background music while performing recommended exercises	12 (21)	4 (3-6)	16 (18)	4 (3-4)	2118.50	−1.20	.23
Interactive gaming features	9 (16)	7 (5-8)	1 (1)	8 (8-8)	1109	−6.84	<.001
System connection to other personal digital devices	8 (14)	5 (4-7)	3 (3)	6 (6-6)	1778.50	−3.02	.003
Availability of various language options	7 (13)	6 (4-6)	7 (8)	6 (6-6)	1655.50	−4.04	<.001
System connection and interactivity with other sources of support	6 (11)	6 (4-7)	1 (1)	7 (7-7)	1610	−3.82	<.001
System activation using voice activation prompts	5 (9)	5 (4-5)	48 (54)	2 (1-5)	1145	−5.39	<.001

XSL•FO
RenderX

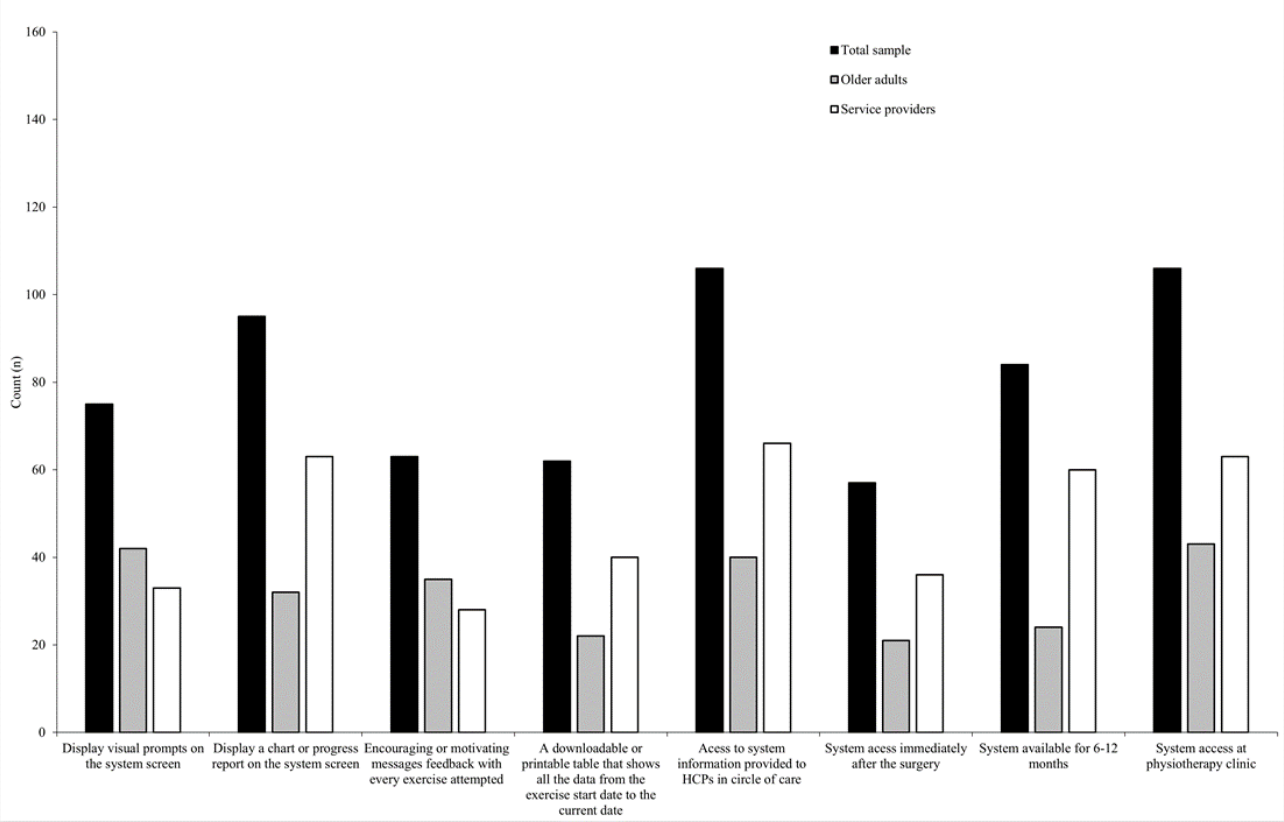
Design features (rank ordered based on older adults' and service providers' responses)[a] (n=142)	Older adults (n=55), n (%)	Median (IQR)	Service providers (n=87)	Median (IQR)	Mann-Whitney *U* test	Z statistic	*P* value
Choice and control over on-screen images presented	3 (5)	4 (3-5)	26 (29)	4 (2-5)	1734.50	−2.85	.004
Choice and control over the virtual environment	3 (5)	5 (4-6)	34 (38)	5 (3-7)	1194.50	−5.26	<.001

[a]Total sample of 145 with 3 values missing.

Practical Logistics for the 3D ARS

Figure 1 summarizes practical considerations and logistical details for the 3D ARS. When asked how customized feedback should be delivered, most older adults with TJR (42/56, 75%) preferred visual prompts displayed on the 3D ARS screen to help guide and adjust joint range of motion. Conversely, most service providers (63/89, 71%) preferred a chart report of joint range-of-motion exercises displayed on the 3D ARS screen.

Figure 1. Practical considerations and logistical details for using the 3D augmented reality system. HCP: health care professional.



Regarding how encouraging or motivational messages should be delivered, most older adults with TJR (35/56, 63%) and about one-third of service providers (28/89, 31%) preferred motivational messages delivered after every exercise attempt. For tracking and documentation, about one-third of older adults (22/56, 39%) and nearly half of the service providers (40/89, 45%) favored a downloadable or printable table that shows all the data from the exercise start date to the current date.

Regarding where and when end users should obtain access, and who should have data access, most older adults with TJR (43/56, 77%) and service providers (63/89, 71%) recommended that users obtain access to the 3D ARS through a physiotherapy clinic. In terms of timing, over one-third of older adults (21/56, 38%) and nearly half of the service providers (42/89, 47%) indicated that access should begin immediately after surgery or within 4 weeks post-surgery. Nearly half of older adults (24/56, 43%) and more than two-thirds of service providers (60/89, 67%) recommended maintaining access for 6 - 12 months. Additionally, most older adults (40/56, 71%) and service providers (66/89, 74%) recommended that health and social

XSL•FO
RenderX

care providers within the user's circle of care should have access to information generated through the 3D ARS.

Potential Barriers to Access and Usage of the 3D ARS

Table 3 summarizes the potential barriers to access and use of the 3D ARS. As shown in the table, over half of the older adults (31/56, 55%) and service providers (48/89, 54%) reported that cost would be a potential barrier. Nearly half of the older adults (25/56, 45%) and service providers (42/89, 47%) indicated that lack of access to suitable exercise equipment to perform recommended exercises at home would also be a top barrier. Among older adults, additional barriers included a lack of confidence in the 3D ARS setup and use (31/56, 55%) and a lack of motivation to use the 3D ARS (26/56, 46%). Service providers did not identify additional top barriers.

Table . Barriers to access and use of 3D augmented reality system (3D ARS; n=145).

Barriers	Older adult (n=56), n (%)	Service providers (n=89), n (%)
Cost of accessing the 3D ARS	31 (55)	48 (54)
Lack of confidence to set up and use 3D ARS	31 (55)	28 (31)
Lack of motivation to use 3D ARS	26 (46)	20 (23)
Lack of access to appropriate equipment to support exercising at home	25 (45)	42 (47)
Lack of adequate space in residence	21 (38)	30 (34)
Overdependence on 3D ARS for recovery	20 (36)	25 (28)
Inability to complete exercises unsupervised at home	20 (36)	6 (7)
Privacy and safety concerns	15 (27)	33 (37)

Discussion

Principal Findings

This study examined priorities for a 3D ARS to support home-based physical rehabilitation for older adults recovering from TJR. Overall, our findings indicate that older adults were open to the potential use of a 3D ARS to support postoperative rehabilitation in the home. Participants emphasized the importance of individualized exercise programming, timely and meaningful feedback during exercises, and user-friendly system features that support ease of interaction. While older adults and service providers shared several priorities for system design, differences emerged in preferences for certain features, including the format and timing of feedback, motivational messaging, and options that support user autonomy, such as control over on-screen images and virtual environments. Together, these findings highlight key user-informed considerations that can guide the development of AR-based rehabilitation technologies for older adults following TJR.

User Priorities and Personalization in AR-Based Rehabilitation

Older adults and service providers shared several priorities for the design of the 3D ARS, including individualized exercise programs, customized feedback on exercise progress, continuous system availability, and optional features such as background music. At the same time, differences emerged in preferences for certain features, particularly those related to motivational messaging and user autonomy. Participants emphasized the importance of the system's ability to deliver positive encouragement during exercise and to provide options that allow users greater control over on-screen images, virtual environments, and language settings. These findings reflect the heterogeneous needs and recovery experiences of older adults following TJR [41] and underscore the importance of tailoring rehabilitation programming to individuals' functional capacities. Personalization has been widely identified as a core design principle in geriatric rehabilitation technologies, supporting engagement and sustained participation in home-based exercise programs [42].

Real-Time Feedback and the Role of Automated Support in Home-Based Rehabilitation

A key finding was the difference in preferred feedback formats during exercises. Service providers favored simple charts and progress reports, whereas older adults preferred immediate, on-screen visual prompts. This distinction is consistent with previous research showing that therapists often prefer concise, actionable summaries rather than navigating extensive patient metrics generated by digital rehabilitation systems [43]. Prioritizing older adults' perspectives as end users highlights the importance of integrating real-time feedback into AR-based rehabilitation technologies. Although previous studies have demonstrated positive responses to synchronous feedback from remote therapists during exercise sessions [17], the shortage of therapists, logistical challenges of scheduling virtual supervision, and associated health care costs may limit the feasibility of such approaches in routine rehabilitation care [44]. While synchronous supervision may enhance safety for certain clinical populations requiring hands-on support [45], AR-based systems that provide automated feedback may offer a more scalable solution for supporting home-based rehabilitation. Another important finding related to the timing of feedback was that participants preferred motivational messages to be delivered after each exercise attempt. This finding aligns with existing evidence suggesting that frequent positive feedback can reinforce desired behaviors and improve exercise adherence [46]. This finding emphasizes the need to incorporate multimodal feedback mechanisms that accommodate both

XSL•FO
RenderX

patient needs and provider oversight requirements in the design of the 3D ARS.

Usability and User Experience Considerations for 3D ARS for Older Adults

Beyond feedback mechanisms, participants emphasized the importance of system usability and ease of interaction when engaging with AR-based rehabilitation technologies. In particular, features such as voice activation and the ability to pause and resume exercises were identified as important for supporting interaction with the 3D ARS. Significant differences between older adults and service providers also emerged for these usability-related features, suggesting that flexibility in how users interact with and control the system may be an important consideration in the design of home-based rehabilitation technologies. Usability is a critical determinant of engagement with digital health technologies, particularly among older adults [47]. Technologies that are difficult to navigate can create confusion and frustration, which may reduce sustained engagement and limit the effectiveness of digital rehabilitation interventions [48,49]. Designing AR-based systems that prioritize intuitive interaction and flexible system control may therefore support greater engagement and usability in home-based postoperative rehabilitation.

Privacy-Preserving Motion Tracking and Data Sharing Considerations in AR-Based Rehabilitation

In addition to usability considerations, participants and system designers must also consider privacy and data management issues associated with camera-based rehabilitation technologies. We propose the use of artificial intelligence–based pose estimation as a promising approach for balancing the need for meaningful movement analysis with concerns related to privacy and data management in home-based rehabilitation. By converting video input into a simplified 3D skeletal representation of the body, the system will capture key movement information while removing identifiable visual features such as facial characteristics, clothing, and background details. This approach may help address privacy concerns that often accompany camera-based monitoring technologies, which can be a barrier to adoption in home environments. In addition, the system focuses on the individual closest to the camera, thereby minimizing the unintended capture of other individuals in the home.

Beyond privacy considerations, the skeletal representation substantially reduces the amount of data generated during exercise sessions compared with full video streams. This reduction in data volume facilitates efficient storage and transmission of exercise information and supports the feasibility of local data storage directly on the console. Such efficiency may be particularly important for in-home rehabilitation contexts where internet connectivity or bandwidth may be limited.

Importantly, this approach also enables flexible models of information sharing with health care providers. At the most basic level, individuals can retain full control over what information is shared by verbally communicating their progress or challenges during follow-up consultations. Alternatively, the system could generate automated summaries of exercise performance, including indicators such as repetitions completed, trends in performance, or potential anomalies. These summaries could be shared manually or electronically through secure communication channels. When more detailed monitoring is required, anonymized skeletal data could also be transmitted to clinicians, enabling more precise review of exercise performance while maintaining a high level of privacy protection.

Taken together, these features suggest that artificial intelligence–based pose estimation may offer a viable pathway for integrating motion tracking into AR-based home exercise systems in a manner that supports both patient privacy and clinical oversight [50]. Such approaches may help address key implementation challenges associated with remote rehabilitation technologies and support more scalable models of home-based care [51,52].

Clinical Integration and Access to AR-Based Rehabilitation

Our findings raise important considerations regarding health equity in the implementation of technology-based rehabilitation. Physiotherapy clinics may vary in their capacity to integrate and support technology-enabled interventions over extended periods, and not all patients may have equal access to such resources due to geographic, financial, or systemic barriers [53]. Addressing these challenges during system development and implementation will be critical to ensuring that the benefits of the 3D ARS are scalable and equitably distributed.

Within this context, participants recommended that access to the 3D ARS be facilitated through physiotherapy clinics, initiated early in the postoperative period, and maintained for at least 6 months. This finding aligns with previous work highlighting physiotherapists' central role and expertise in exercise prescription for older adults with TJR [54]. Situating access within clinical settings may also ensure that the use of the 3D ARS is guided by professional oversight, reducing the risk of misuse while positioning the system as a complement to conventional rehabilitation care.

Cost, Digital Equity, and Access to Rehabilitation Technologies

Other issues identified include the potential cost of the 3D ARS and a lack of access to suitable exercise equipment to complete the recommended exercises at home. Overall, participants' perspectives on costs highlight issues of digital equity and the importance of efforts to address the existing divide in access to digital infrastructure, which gives some individuals advantages over others [55]. To address these concerns, the proposed 3D ARS will leverage a relatively simple hardware configuration similar to existing motion-based gaming consoles, in which a camera-based console connects directly to a television or monitor. Such an approach could facilitate easy integration into home environments while minimizing installation complexity. By using commonly available household electronics, such as a TV or monitor, this configuration may also help reduce system costs and improve the feasibility of implementing AR-based rehabilitation technologies in the home.

XSL•FO
RenderX

Supporting Independent Use of the 3D ARS Through Accessible System Design

Finally, the lack of confidence in the 3D ARS setup identified by older adults highlights the importance of incorporating robust training supports and designing intuitive, user-friendly setup procedures [56]. Embedding step-by-step guidance, interactive tutorials, and a clear user interface design within the system itself could increase user confidence while reducing cognitive load and reliance on external assistance. Such design choices not only empower users to engage more independently but may also help minimize system development and implementation costs by reducing the need for dedicated technical support staff to resolve setup issues.

Overall, these findings highlight several key considerations for the development of AR-based rehabilitation technologies for older adults following TJR. Participants emphasized the importance of systems that combine individualized exercise programming, accessible and intuitive user interfaces, and timely feedback mechanisms that support motivation and adherence. At the same time, our results underscore the need to balance technological capabilities with practical implementation considerations, including privacy protection, clinical oversight, and equitable access to digital rehabilitation resources. By incorporating these user-informed priorities into the design of the 3D ARS, future systems may better align with the needs of both patients and rehabilitation professionals, supporting scalable models of home-based rehabilitation that complement conventional care.

Limitations

Several limitations should be considered when interpreting the findings of this study. First, participants were asked to evaluate proposed features of a 3D ARS rather than interacting with a fully functional system. As a result, the preferences expressed may not fully reflect experiences with actual system use during postoperative TJR rehabilitation. Future studies should examine user experiences with prototype systems to better understand usability, engagement, and clinical effectiveness in real-world contexts.

Second, the sample included a relatively highly educated group of older adults, which may limit the generalizability of findings to individuals with lower levels of education or digital literacy. Because comfort with technology can influence perceptions of digital health tools, the perspectives captured in this study may not fully represent the broader population of older adults undergoing TJR.

Third, the cross-sectional design captures participant perspectives at a single point in time and does not account for how preferences may evolve during different stages of postoperative recovery or with increased familiarity with digital rehabilitation technologies. Longitudinal research may provide additional insight into how user needs and expectations change throughout the rehabilitation process.

Finally, although participants were recruited from multiple regions, a substantial proportion of the sample was located in Ontario. Health system structures, access to rehabilitation services, and availability of digital infrastructure may vary across regions, which could influence the transferability of these findings to other settings.

Conclusions

This study identified key design priorities for a 3D ARS to support home-based physical rehabilitation for older adults recovering from TJR. These findings provide an important foundation for the development of AR-based rehabilitation technologies that respond to the needs of end users and rehabilitation professionals. Future work will focus on user-centered system development, prototype testing, and usability evaluation to support the implementation of a practical and scalable AR-based rehabilitation solution.

Acknowledgments

The authors would like to thank Poornima Kamath and Helen Rosseau for their contributions to the survey design.

Funding

This research was supported by the National Research Council Canada’s Aging in Place Program.

Data Availability

The data generated and analyzed during this study to support the findings are included in this published article and its supplementary files.

Authors' Contributions

Data curation, formal analysis, and writing – original draft: AS
Data curation, formal analysis, and writing – review and editing: SP
Conceptualization, funding acquisition, and writing – review and editing: MK
Conceptualization, methodology, data curation, funding acquisition, formal analysis, supervision, and writing – original draft: AF
All authors reviewed the draft, provided feedback, and approved the final manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Neuprez A, Neuprez AH, Kaux JF, et al. Total joint replacement improves pain, functional quality of life, and health utilities in patients with late-stage knee and hip osteoarthritis for up to 5 years. Clin Rheumatol 2020 Mar;39(3):861-871. [doi: 10.1007/s10067-019-04811-y] [Medline: 31720892]
2. Fleischman AN, Austin MS, Purtill JJ, Parvizi J, Hozack WJ. Patients living alone can be safely discharged directly home after total joint arthroplasty. J Bone Joint Surg Am 2018 Jan 17;100(2):99-106. [doi: 10.2106/JBJS.17.00067] [Medline: 29342059]
3. Fu MC, Samuel AM, Sculco PK, MacLean CH, Padgett DE, McLawhorn AS. Discharge to inpatient facilities after total hip arthroplasty is associated with increased postdischarge morbidity. J Arthroplasty 2017 Sep;32(9S):S144-S149. [doi: 10.1016/j.arth.2017.03.044] [Medline: 28455181]
4. Debbi EM, Mosich GM, Bendich I, Kapadia M, Ast MP, Westrich GH. Same-day discharge total hip and knee arthroplasty: trends, complications, and readmission rates. J Arthroplasty 2022 Mar;37(3):444-448. [doi: 10.1016/j.arth.2021.11.023] [Medline: 34808278]
5. Cofano E, Familiari F, Mori T, et al. Home-based vs. clinic-based rehabilitation after joint arthroplasty: a prospective matched cohort study. Appl Sci (Basel) 2025;15(9):4937. [doi: 10.3390/app15094937]
6. Bravi M, Longo UG, Laurito A, et al. Supervised versus unsupervised rehabilitation following total knee arthroplasty: a systematic review and meta-analysis. Knee 2023 Jan;40:71-89. [doi: 10.1016/j.knee.2022.11.013] [Medline: 36410253]
7. Chen H, Wang Y, Liu C, et al. Benefits of a transtheoretical model-based program on exercise adherence in older adults with knee osteoarthritis: a cluster randomized controlled trial. J Adv Nurs 2020 Jul;76(7):1765-1779. [doi: 10.1111/jan.14363] [Medline: 32202313]
8. Bakaa N, Chen LH, Carlesso L, Richardson J, Shanthanna H, Macedo L. Understanding barriers and facilitators of exercise adherence after total-knee arthroplasty. Disabil Rehabil 2022 Oct;44(21):6348-6355. [doi: 10.1080/09638288.2021.1965232] [Medline: 34397309]
9. Poulsen AG, Gravesen JD, Madsen MN, Mikkelsen LR, Bandholm T, Rossen CB. Patient perspectives on home-based rehabilitation exercise and general physical activity after total hip arthroplasty: a qualitative study (PHETHAS-2). F1000Res 2021;10:382. [doi: 10.12688/f1000research.51684.5] [Medline: 37359253]
10. Husted RS, Bandholm T, Rathleff MS, Troelsen A, Kirk J. Perceived facilitators and barriers among physical therapists and orthopedic surgeons to pre-operative home-based exercise with one exercise-only in patients eligible for knee replacement: a qualitative interview study nested in the QUADX-1 trial. PLoS ONE 2020;15(10):e0241175. [doi: 10.1371/journal.pone.0241175] [Medline: 33095777]
11. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. Int J Environ Res Public Health 2021 Feb 19;18(4):2023. [doi: 10.3390/ijerph18042023] [Medline: 33669679]
12. Daly RM, Gianoudis J, Hall T, Mundell NL, Maddison R. Feasibility, usability, and enjoyment of a home-based exercise program delivered via an exercise app for musculoskeletal health in community-dwelling older adults: short-term prospective pilot study. JMIR Mhealth Uhealth 2021 Jan 13;9(1):e21094. [doi: 10.2196/21094] [Medline: 33439147]
13. Hanson HM, Friesen J, Beaupre L, Jasper L, Millington J, Jones CA. Supporting rehabilitation of rural patients receiving total knee arthroplasty through physical activity: perceptions of stakeholder groups. ACR Open Rheumatol 2022 Oct;4(10):863-871. [doi: 10.1002/acr2.11489] [Medline: 35862257]
14. Stisen MB, Mechlenburg I, Bearne LM, Godfrey E, Pedersen AB, Sørensen D. Exploring needs, barriers to, and facilitators of rehabilitation exercise following revision hip replacement - a grounded theory study. Disabil Rehabil 2024 Oct;46(21):5012-5020. [doi: 10.1080/09638288.2023.2295008] [Medline: 38117004]
15. Nuevo M, Modrego A, Rodríguez-Rodríguez D, et al. The influence of sociodemographic and health factors on adherence to home-based rehabilitation after fast-track total knee arthroplasty: secondary analysis of a randomized controlled trial. Disabil Rehabil 2025 Sep;47(18):4728-4737. [doi: 10.1080/09638288.2025.2458191] [Medline: 39879639]
16. Smith TO, Dainty JR, Clark EM, Whitehouse MR, Price AJ, MacGregor AJ. Demographic and geographical variability in physiotherapy provision following hip and knee replacement. An analysis from the National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. Physiotherapy 2020 Mar;106:1-11. [doi: 10.1016/j.physio.2019.11.003] [Medline: 32026841]
17. Conroy CL, Brunetti GM, Barmpoutis A, Fox EJ. Integrated telehealth and extended reality to enhance home exercise adherence following total hip and knee arthroplasty. Presented at: 2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); Mar 8-12, 2025; Saint Malo, France p. 959-964. [doi: 10.1109/VRW66409.2025.00195]

XSL•FO
RenderX

18. Zhao B, Liu H, Du K, Zhou W, Li Y. Effectiveness and safety of outpatient rehabilitation versus home-based rehabilitation after knee arthroplasty: a systematic review and meta-analysis. J Orthop Surg Res 2023 Sep 19;18(1):704. [doi: 10.1186/s13018-023-04160-2] [Medline: 37726800]
19. Chan ZYS, MacPhail AJC, Au IPH, et al. Walking with head-mounted virtual and augmented reality devices: effects on position control and gait biomechanics. PLoS ONE 2019;14(12):e0225972. [doi: 10.1371/journal.pone.0225972] [Medline: 31800637]
20. Chang E, Kim HT, Yoo B. Virtual reality sickness: a review of causes and measurements. International Journal of Human–Computer Interaction 2020 Oct 20;36(17):1658-1682. [doi: 10.1080/10447318.2020.1778351]
21. Liu Y, Zhang Q, Li W. Enhancing lower-limb rehabilitation: a scoping review of augmented reality environment. J Neuroeng Rehabil 2025 May 20;22(1):114. [doi: 10.1186/s12984-025-01643-7] [Medline: 40394647]
22. Farsi A, Cerone GL, Falla D, Gazzoni M. Emerging applications of augmented and mixed reality technologies in motor rehabilitation: a scoping review. Sensors (Basel) 2025 Mar 25;25(7):2042. [doi: 10.3390/s25072042] [Medline: 40218554]
23. Butz B, Jussen A, Rafi A, Lux G, Gerken J. A taxonomy for augmented and mixed reality applications to support physical exercises in medical rehabilitation—a literature review. Healthcare (Basel) 2022 Mar 30;10(4):646. [doi: 10.3390/healthcare10040646] [Medline: 35455824]
24. De Cecco M, Luchetti A, Butaslac I III, et al. Sharing augmented reality between a patient and a clinician for assessment and rehabilitation in daily living activities. Information 2023;14(4):204. [doi: 10.3390/info14040204]
25. Drazich BF, McPherson R, Gorman EF, et al. In too deep? A systematic literature review of fully-immersive virtual reality and cybersickness among older adults. J Am Geriatr Soc 2023 Dec;71(12):3906-3915. [doi: 10.1111/jgs.18553] [Medline: 37560978]
26. Modlin D, Kuo YT. The effects of using augmented reality in rehabilitation and recovery exercise on patients' outcomes and experiences: a systematic review. Front Virtual Real 2025;6:1641316. [doi: 10.3389/frvir.2025.1641316]
27. Kakoutopoulos K, Drakakis E, Papadopoulou A, Goumopoulos C. Feasibility of augmented reality-based cognitive training for older adults: the MarketMind AR approach. Sensors (Basel) 2025 Mar 26;25(7):2081. [doi: 10.3390/s25072081] [Medline: 40218594]
28. Jin X, Tong W, Wei X, et al. Exploring the opportunity of augmented reality (AR) in supporting older adults to explore and learn smartphone applications. Presented at: CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems; May 11-16, 2024. [doi: 10.1145/3613904.3641901]
29. Arntz A, Weber F, Handgraaf M, et al. Technologies in home-based digital rehabilitation: scoping review. JMIR Rehabil Assist Technol 2023 Jul 27;10:e43615. [doi: 10.2196/43615] [Medline: 37253381]
30. Escalona F, Martinez-Martin E, Cruz E, Cazorla M, Gomez-Donoso F. EVA: EVAluating at-home rehabilitation exercises using augmented reality and low-cost sensors. Virtual Real 2020 Dec;24(4):567-581. [doi: 10.1007/s10055-019-00419-4]
31. Kaufeld M, Mundt M, Forst S, Hecht H. Optical see-through augmented reality can induce severe motion sickness. Displays 2022 Sep;74:102283. [doi: 10.1016/j.displa.2022.102283]
32. Mendoza-Ramírez CE, Tudon-Martinez JC, Félix-Herrán LC, Lozoya-Santos JDJ, Vargas-Martínez A. Augmented reality: survey. Appl Sci (Basel) 2023;13(18):10491. [doi: 10.3390/app131810491]
33. Bell KM, Onyeukwu C, McClincy MP, et al. Verification of a portable motion tracking system for remote management of physical rehabilitation of the knee. Sensors (Basel) 2019 Feb 28;19(5):1021. [doi: 10.3390/s19051021] [Medline: 30823373]
34. Andriollo L, Picchi A, Iademarco G, et al. The role of artificial intelligence and emerging technologies in advancing total hip arthroplasty. J Pers Med 2025 Jan 9;15(1):21. [doi: 10.3390/jpm15010021] [Medline: 39852213]
35. Modlin D, Kuo YT. The effects of using augmented reality in rehabilitation and recovery exercise on patients' outcomes and experiences: a systematic review. Front Virtual Real 2025 Oct 14;6:2025. [doi: 10.3389/frvir.2025.1641316]
36. Fields B, Fitzpatrick M, Kinney L, et al. Evaluating the acceptability and appropriateness of the Augmented Reality Home Assessment Tool (ARHAT): qualitative descriptive study. JMIR Aging 2023 Sep 27;6:e44525. [doi: 10.2196/44525] [Medline: 37787657]
37. Mitchell J, Shirota C, Clanchy K. Factors that influence the adoption of rehabilitation technologies: a multi-disciplinary qualitative exploration. J Neuroeng Rehabil 2023 Jun 20;20(1):80. [doi: 10.1186/s12984-023-01194-9] [Medline: 37340496]
38. Phoenix M, Moll S. Realizing the potential of co-design to build innovation in rehabilitation services. Physiother Can 2021 Nov 1;73(4):299-300. [doi: 10.3138/ptc-2021-0060-gee] [Medline: 34880532]
39. Gorman C, Gustafsson L, Gomura C. The perspectives of stroke survivors and health professionals on the use of augmented reality for inpatient stroke rehabilitation: an anticipatory exploration. Brain Impair 2022 Sep;23(2):206-215. [doi: 10.1017/BrImp.2021.11]
40. Fakolade A, Salvia AC, Phadke S, Kunz M. An interactive vision-based 3D augmented reality system for in-home physical rehabilitation: a qualitative inquiry to inform system development. Health Expect 2024 Oct;27(5):e70020. [doi: 10.1111/hex.70020] [Medline: 39440453]
41. Lim CV, Harithasan D, Ahmad MAB, Mesbah N, Kumar S, Singh DKA. Biopsychosocial determinants for total knee replacement decisions among older adults with chronic knee osteoarthritis: a scoping review. Geriatr Orthop Surg Rehabil 2025;16:21514593251335882. [doi: 10.1177/21514593251335882] [Medline: 40297342]

XSL•FO
RenderX

42. Li G, Li X, Chen L. Personally tailored exercises for improving physical outcomes for older adults in the community: a systematic review. Arch Gerontol Geriatr 2022;101:104707. [doi: 10.1016/j.archger.2022.104707] [Medline: 35462312]
43. Pearce LMN, Howell M, Yamato TP, et al. What attributes of digital devices are important to clinicians in rehabilitation? A cross-cultural best-worst scaling study. Int J Med Inform 2024 Nov;191:105589. [doi: 10.1016/j.ijmedinf.2024.105589] [Medline: 39137666]
44. Hamilton C, Lovarini M, McCluskey A, Folly de Campos T, Hassett L. Experiences of therapists using feedback-based technology to improve physical function in rehabilitation settings: a qualitative systematic review. Disabil Rehabil 2019 Jul;41(15):1739-1750. [doi: 10.1080/09638288.2018.1446187] [Medline: 29513052]
45. Jarbandhan A, Toelsie J, Veeger D, Bipat R, Vanhees L, Buys R. Feasibility of a home-based physiotherapy intervention to promote post-stroke mobility: a randomized controlled pilot study. PLoS ONE 2022;17(3):e0256455. [doi: 10.1371/journal.pone.0256455] [Medline: 35255091]
46. Meade LB, Bearne LM, Sweeney LH, Alageel SH, Godfrey EL. Behaviour change techniques associated with adherence to prescribed exercise in patients with persistent musculoskeletal pain: systematic review. Br J Health Psychol 2019 Feb;24(1):10-30. [doi: 10.1111/bjhp.12324] [Medline: 29911311]
47. Berry ECJ, Sculthorpe NF, Warner A, Mather JD, Sanal-Hayes NEM, Hayes LD. A scoping review of the feasibility, usability, and efficacy of digital interventions in older adults concerning physical activity and/or exercise. Front Aging 2025;6:1516481. [doi: 10.3389/fragi.2025.1516481] [Medline: 40290578]
48. Zanatta F, Giardini A, Pierobon A, D'Addario M, Steca P. A systematic review on the usability of robotic and virtual reality devices in neuromotor rehabilitation: patients' and healthcare professionals' perspective. BMC Health Serv Res 2022 Apr 20;22(1):523. [doi: 10.1186/s12913-022-07821-w] [Medline: 35443710]
49. Contreras-Somoza LM, Irazoki E, Toribio-Guzmán JM, et al. Usability and user experience of cognitive intervention technologies for elderly people with MCI or dementia: a systematic review. Front Psychol 2021;12:636116. [doi: 10.3389/fpsyg.2021.636116] [Medline: 33967901]
50. Yan X, Xu Y, Chen C, Zhang S. Privacy preserving for AI-based 3D human pose recovery and retargeting. ISA Trans 2023 Oct;141:132-142. [doi: 10.1016/j.isatra.2023.04.022] [Medline: 37121822]
51. Abedi A, Colella TJF, Pakosh M, Khan SS. Artificial intelligence-driven virtual rehabilitation for people living in the community: a scoping review. NPJ Digit Med 2024 Feb 3;7(1):25. [doi: 10.1038/s41746-024-00998-w] [Medline: 38310158]
52. Mennella C, Maniscalco U, De Pietro G, Esposito M. The role of artificial intelligence in future rehabilitation services: a systematic literature review. IEEE Access 2023;11:11024-11043. [doi: 10.1109/ACCESS.2023.3236084]
53. Pearce LMN, Hassett L, Sherrington C, Pryor J. Human interactions remain at the heart of rehabilitation with advanced technology: a practice-embedded longitudinal qualitative study with allied health clinicians. J Neuroeng Rehabil 2025 Mar 7;22(1):52. [doi: 10.1186/s12984-025-01576-1] [Medline: 40050924]
54. Aftab S, Ali HN, Saeed B, Sarwar S, Dawood MH, Pervez M. Early physiotherapy for post-total knee arthroplasty recovery: a systematic review of randomized controlled trials on quality of life, pain, and range of motion outcomes. Musculoskeletal Care 2025 Sep;23(3):e70158. [doi: 10.1002/msc.70158] [Medline: 40635148]
55. Hall AK, Bernhardt JM, Dodd V, Vollrath MW. The digital health divide: evaluating online health information access and use among older adults. Health Educ Behav 2015 Apr;42(2):202-209. [doi: 10.1177/1090198114547815] [Medline: 25156311]
56. Wang J, Du Y, Coleman D, et al. Mobile and connected health technology needs for older adults aging in place: cross-sectional survey study. JMIR Aging 2019 May 15;2(1):e13864. [doi: 10.2196/13864] [Medline: 31518283]

Abbreviations

3D ARS: 3D augmented reality system
AR: augmented reality
TJR: total joint replacement
VR: virtual reality

Edited by M Gray; submitted 11.Nov.2025; peer-reviewed by M Ashour, R Tredinnick; revised version received 05.Mar.2026; accepted 25.Mar.2026; published 20.Apr.2026.

Please cite as:
Salvia AC, Phadke S, Kunz M, Fakolade A
Interactive Vision–Based 3D Augmented Reality System for In-Home Rehabilitation in Older Adults Following Total Joint Replacement: Cross-Sectional Survey of Patients and Service Providers
JMIR Aging 2026;9:e87572
URL: https://aging.jmir.org/2026/1/e87572
doi:10.2196/87572

XSL•FO
RenderX

© Adriana C Salvia, Siona Phadke, Manuela Kunz, Afolasade Fakolade. Originally published in JMIR Aging (https://aging.jmir.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The 4 Ss Framework for Designing Age-Friendly Communities of the Future: Focus Group Study

Inga Hunter, MA, MBChB, MPhil; Caroline Lockhart, RN, MBS

School of Management and Marketing, Massey Business School, Massey University, Business Studies Complex, 3rd floor, Tennent Drive, Palmerston North, Manawatu-Wanganui, New Zealand

Corresponding Author:
Inga Hunter, MA, MBChB, MPhil
School of Management and Marketing, Massey Business School, Massey University, Business Studies Complex, 3rd floor, Tennent Drive, Palmerston North, Manawatu-Wanganui, New Zealand

Abstract

Background: Digital technology is increasingly being used to deliver interventions and initiatives to support the well-being of older adults. However, few studies have conducted needs assessments to identify the future well-being service requirements of an older adult population and their preferred modes of delivery, whether via digital technology or, in- person, or a combination of both (ie, a hybrid model).

Objective: This study aims to investigate the requirements of a rural region in New Zealand to inform planning to meet the future well-being needs of its older adult population over the next 30 years.

Methods: In total, 2 focus group discussions and 13 interviews were held with participants using a combination of phone and video calls. A total of 31 adults aged 61 years or older participated. The participants were asked how they saw the future well-being needs of the older adult population evolving, the role of digital technology and/or in-person interactions in delivering well-being services, and perceived barriers to, and enablers of, digital technology for providing services. Focus group and interview transcripts were thematically analyzed.

Results: A total of 4 key well-being themes were identified across both focus group discussions and interviews with participants: —"services," "support," "spaces," and "social connection," —with an overarching finding that future age-friendly planning must prioritize social and community well-being outcomes, ensuring that investments in services, digital inclusion, and public spaces strengthen older adults' ability to connect with others and participate fully in community life.

Conclusions: Digital technology is developing at a rapid pace, and as a result, we need to consider how to plan for the transition and bridge the gap identified between the current use of digital technology and its potential future use if technology is to support the older adults of the future. The findings indicate that older adults prefer to engage in- person, while trust is a barrier to digital technology use for some participants. The future offers many opportunities to support the well-being of older adults and communities through the application of the proposed 4 Ss Framework.

(JMIR Aging 2026;9:e90923) doi:10.2196/90923

KEYWORDS

older adults; digital literacy; digital health; social connection; social isolation; well-being; trust; technology adoption; rural health services; health equity

Introduction

Background

The World Health Organization (WHO) frames health as a dynamic resource that enables people to meet their needs and aspirations [1]. This broad understanding recognizes that health and well-being are shaped by the wider environment in which people live, and many determinants of health exist outside of the traditional scope of the health sector [2]. The WHO applied these determinants to the health and well-being of older adults to promote the concept of "active aging" through its Active Aging Framework [3]. The Healthy Ageing Strategy in New Zealand [4] considers that older people should live well and age well in "age-friendly" communities, thus linking active aging with Age-Friendly Cities and Communities (AFCC).

Van Hoof et al [5] developed their model of the age-friendly city with 8 domains that align with the WHO dimensions of active aging: outdoor spaces and buildings;, transportation;, housing;, civic participation and employment;, respect and social inclusion;, social participation;, communication and information;, and community support and health services. This is very similar to the 5 key areas identified by the Office of the Seniors "Better Later Life–He Oranga Kaumātua 2019 to 2034" [6].

XSL•FO
RenderX

Much of the research around AFCC is at a conceptual or pilot stage and focuses on only specific components for AFCC, and the metrics used to assess AFCC interventions are limited [7,8]. Interventions vary and may include the improvement ofin health risk factors for health, positive behavioral lifestyle changes, changes in organizations, and changes in the domains of AFCC, with the majority having a combination of environmental and psychosocial interventions [7]. Examples include safe and friendly housing, transportation, social inclusion and participation, outdoor spaces and buildings, communication, and information [7].

Smart Cities and Communities, AFCC, and Active Aging

The WHO emphasizes promoting healthy aging in a digital world [9] and identifies loneliness and isolation as critical public health issues requiring urgent intervention [10,11]. The WHO's extension of the Global Digital Health Strategy to 2027 further underscores the relevance of integrating digital solutions for aging populations worldwide [12].

At its most basic, "smart cities and communities" are cities and communities that are interconnected by digital technology and are data-driven, using AI, to achieve better lives (well-being) for all residents. However, they are not without their own problems [13,14], as the predominant technologies used are the Internet of Things (sensors) and AI [15,16], each of which comes with design, implementation, and sustainability issues. Pedell and Borda [17] combined smart cities, AFCC, and active aging, with 3 of the WHO domains to bring in the digital aspect (Smart) in their conceptual model: a simplified version of the AFCC model by Van Hoof et al [5]. When tested, similar to the authors' previous research [13,14], they found that digital technology can reduce social isolation, increase social participation, and increase independence. Pedell and Borda [17] suggest that digital technologies may have a wider role to play in AFCC.

New Zealand Context

Against this international backdrop, this study explores how one New Zealand community is preparing to meet the future well-being needs of its older population. In New Zealand, it is critical that health and well-being are considered under the obligations of Te Tiriti o Waitangi [18] and through a bicultural lens, reflecting Ma¯ori perspectives of health. This approach is embodied in the Te Whare Tapa Wha¯ model [19]. Te Whare Tapa Wha¯ defines well-being across 4 interconnected dimensions: physical health (taha tinana), spiritual health (taha wairua), mental/emotional health (taha hinengaro), and family/social health (taha whānau) [19]. These perspectives extend the concept of well-being beyond individual health to include social and community well-being, which are fundamental to Ma¯ori, who operate within a collective-based society [20]. This study applies a holistic interpretation of well-being that is consistent with Te Whare Tapa Wha¯ and New Zealand's bicultural health context.

Methods

Study Design

This qualitative exploratory study sought to examine the future well-being needs of older adults residing in a small town and its surrounding rural community, with the aim of informing planning for future service delivery.

Setting and Context

This study was conducted in a developing rural area in New Zealand's North Island, historically settled by retired farmers, and characterized by dispersed settlements, limited transport options, and variable internet connectivity. As of 2024, the region had a population of 33,700 people of whom 19.6% were aged ≥65 years [21]. Travel across the region can take approximately 90 minutes. These contextual features create both challenges and opportunities to inform future planning for well-being-related services for an aging population, making older adults a central consideration for this study during the design and analysis stages.

Participants and Eligibility

The eligibility criteria were community-dwelling older adults born between 1940 and 1960 and residing in the region at the time of data collection. This cohort requirement was set by the project funder.

Recruitment

Participants were recruited through purposive sampling [22] to ensure diverse perspectives by gender and ethnicity including: male and female participants, Māori, and other non—New Zealand European participants. Recruitment strategies included posters placed in community venues, notices disseminated via local organizations, posts on community Facebook pages, and the researchers' existing contact networks. These traditional methods of in-person recruiting and using paper-based marketing material were adapted during the recruitment process because the national COVID-19 pandemic lockdown settings prevented in-person contact at the time of data collection. Participation was voluntary. Participants could bring a support person, decline to answer questions, withdraw, and access translation services if required. All participants provided either written or verbal informed consent prior to participation. Verbal consent was an added option for participants due to the restrictions imposed by COVID lockdowns, and returning a scanned informed consent sheet was difficult for some. These recruitment strategies were designed to capture a diverse range of perspectives and ensure cultural inclusivity.

Data Collection

Data collection occurred between April 5 and May 11, 2022. The project was originally designed to include up to 6 in-person focus groups (a maximum of 40 participants) with older adults born between 1940 and 1960. National COVID-19 pandemic lockdown settings prevented in-person focus groups and necessitated a mixed approach, combining 2 semistructured online focus groups (n=18), three online interviews (n=3), and telephone interviews (n=13). This adaptation of the study's design also addressed some of the challenges experienced by

XSL•FO
RenderX

participants when using digital platforms (such as videoconferencing via Zoom), technical difficulties such as inexperience in using these digital platforms, rural internet connectivity challenges, and preferences for individual interviews. Scanning consent forms was a particular challenge and was addressed using recorded verbal consent.

The adaptive approach taken regarding the initial methodology is consistent with previous rural telehealth research, which shows that flexibility is often required when engaging older populations in digitally constrained contexts [23].

Participants self-rated their confidence in using digital technologies and provided demographic details on age, ethnicity, and gender. Semistructured questions [24] were provided in advance to encourage reflection, a strategy that has proven effective in earlier research [23]. The same questions were used across phone interviews and focus groups for standardization. Four questions, with prompts, structured the sessions (Table 1).

Table . Questions for focus groups and phone interviews.

Question number	Summary of questions covered
1	How do you see the well-being needs of your community evolving over the next 30 years? What types of community services will be required?
2	Should these services be delivered in- person, digitally, or in a hybrid form?
3	What concerns or barriers might you have in using digital technology to meet the community's well-being needs? What mitigating factors are required to address these?
4	What benefits could digital technology offer (eg, savings on time, travel, or costs)?

Data Analysis

An inductive thematic analysis was undertaken [25,26]. All focus groups and phone interviews were audio-recorded, transcribed verbatim, and anonymized, and the data were managed in NVivo (Release 1.6.1; Lumivero). Initial codes were generated, grouped into categories, and refined into overarching themes. Coding decisions and analytic notes were documented to ensure transparency, and themes were reviewed against the full dataset for accuracy. Two researchers coded the data independently and resolved any differences through discussion. The final analysis identified 4 key themes, summarized as the "4 Ss Framework": services, support, spaces, and social Connection.

Ethical Considerations

The project followed the Massey University Code of Ethical Conduct for Research, Teaching and Evaluations Involving Human Participants [27] and was judged by peer review as low risk, including suitability from a Māori research perspective. A low-risk notification was recorded with Massey University's Human Ethics Committee [28]. Māori research ethics considerations were reviewed to ensure cultural appropriateness and responsiveness to Te Tiriti o Waitangi, and the findings were reviewed from a Māori perspective to ensure appropriate interpretation. All participants gave informed consent, and participant involvement was entirely voluntary; privacy was preserved by anonymizing the data, and data were kept secure and confidential as per the university data management guidelines. Participants did not receive financial compensation.

Results

Participant Characteristics

A total of 31 older adults aged 61 to 82 years participated, and 45% (n=14) identified as male and 54% (n=17) as female. To allow for birthdays later in the year, the eligibility criteria allowed for 61-year-old participants on the date of data collection.

The majority identified as New Zealand European (n=26, 86%), Māori (n=4, 13%), and "Other" (n=1, 3%), based on the Statistics New Zealand ethnicity classification system [29]. Māori participants were slightly underrepresented when compared to local population data for both the region (18%) and the national population (17%) [30,31]. Findings indicated that 18 (58% (n=18) of participants reported that they were somewhat, or very confident using technology (Figure 1), including the oldest participants, and 63% (n=19) of Manawatū resident participants self-reported as very/ confident or somewhat confident using technology. Table 2 shows the demographic characteristics.

In the findings section, where a participant could be identified with a participant ID, this has been provided. Some quotes are from focus group participants, in which case the relevant focus group participant IDs are provided.

XSL•FO
RenderX

Table . Demographic characteristics.

Participant ID	Type of contact	Ethnic group	GenderSex	Age (y)	Self-reported level of confidence
P1	Focus group	New Zealand European	Female	71	Very unconfident
P2	Focus group	New Zealand European	Male	73	Somewhat confident
P3	Phone interview	Māori	Female	72	Very confident
P4	Phone interview	New Zealand European	Male	81	Somewhat confident
P5	Focus group	New Zealand European	Male	65	Very confident
P6	Focus group	New Zealand European	Male	70	Very confident
P7	Focus group	Other	Female	76	Very confident
P8	Focus group	New Zealand European	Female	67	Neither
P9	Focus group	New Zealand European	Female	61	Somewhat confident
P10	Focus group	New Zealand European	Male	73	Very confident
P11	Focus group	New Zealand European	Male	80	Neither
P12	Focus group	New Zealand European	Female	82	Somewhat confident
P13	Focus group	NZENew Zealand European	Male	70	Somewhat confident
P14	Focus group	New Zealand European	Female	66	Somewhat confident
P15	Focus group	New Zealand	Female	67	Neither
P16	Focus group	New Zealand	Male	70	Neither
P17	Focus group	New Zealand	Male	69	Neither
P18	Focus group	New Zealand	Female	66	Very unconfident
P19	Focus group	New Zealand	Female	77	Very unconfident
P20	Focus group	New Zealand	Female	82	Somewhat confident
P21	Phone interview	New Zealand	Male	73	Somewhat confident
P22	Phone interview	New Zealand	Female	70	Somewhat confident
P23	Phone interview	New Zealand	Female	70	Neither
P24	Phone interview	Māori	Female	61	Very confident
P25	Phone interview	New Zealand	Male	64	Neither
P26	Phone interview	New Zealand	Female	74	Neither
P27	Phone interview	New Zealand	Male	61	Very confident
P28	Phone interview	New Zealand	Male	74	Somewhat unconfident
P29	Phone interview	New Zealand	Female	65	Somewhat confident
P30	Phone interview	Māori	Male	70	Neither
P31	Phone interview	Māori	Female	65	Somewhat confident

Thematic Findings: The 4 Ss Framework

Thematic analysis identified 4 interrelated domains of well-being—services, support, spaces, and social connection—which together form the 4 Ss Framework. Table 3 presents these overarching themes and their associated subthemes. These themes reflect priorities identified by participants when considering what older adults will need to age well over the next 30 years, as required by the project funder.

XSL•FO
RenderX

Table . Themes and subthemes (4 Ss Framework) identified by participants.

Theme	Subtheme
Services	Transport, housing, health care, and communication
Support	Digital technology, connectivity, safety and security, information
Spaces	Existing facilities, purpose-built social spaces, technology hubs, and hybrid (digital and physical spaces)
Social connection	In-person interactions and socializing, loneliness and isolation, feeling part of the community, Te Ao Māori, and connection

Theme 1: Services

Participants consistently emphasized the importance of services that sustain independence and well-being, particularly transport, health care, housing, and communication.

Transport

Transport emerged as the most pressing concern. Rural participants described high fuel costs, limited public transport, and mobility barriers to reaching bus stops. Some worried about dependence on family: "I can get lifts, but I don't want to be dependent on others" (P1). Alternatives suggested included subsidized taxis, shuttles, and mobility scooters. Participants considered autonomous (self-driving) cars as viable future transport options, where one said, "I'm looking forward to autonomous cars when I stop driving" (P6). Taxi services were considered better for social interaction and more personal than taking the bus.

The prospect of losing a driver's license generated considerable anxiety: "I am dreading what happens if I can't drive" (P2). Participants linked this to reduced social interaction and an increased risk of isolation and loneliness: "I hate the idea that we can't go out at night" (P3).

Housing

Housing affordability and limited options constrained the ability to relocate closer to services or family,: "Housing developments are currently targeted to those who are well off financially" (P21). Mixed developments that combine public, private, intergenerational, and sheltered housing were widely supported. "The actual village has young people and children and families and workers and retirees and unemployed, that is the sort of caring community I would like to see for the ageing population" (P3).

Smart home technology in housing was generally regarded positively. Different levels of familiarity with smart home technology were recognized by participants,; "some people in NZ already have semi-intelligent houses, already know how to use Alexa and technological gadgets" (P5), while another acknowledged that "people that haven't grown up with [it] are going to need the reassurance that they can use these things in the next 30 years" (P6).

Health Care

Health care was another key issue. Concerns were raised about limited local providers, long wait times, and travel distances for rural residents. These concerns ran the full gamut of health care, including emergency care, acute care, long-term care, disability support, including support for low vision and deaf communities, mobility, and dementia care. Changes discussed included walk-in clinics, different opening times, and reduced waiting times, with one participant likening the current health service to "a very slow train crash" (P3). Some participants recollected the district nurses' home visits and expressed concern about their decline. Cost was raised as a general issue, ("Living off the pension now is prohibitive, power, phone, insurance, …." ([P27]) and specifically around private health insurance ("By the time you need your health insurance, you can't afford it, it is too expensive" ([P3]).

Telehealth (Tthe interpretation of Telehealth applied here is "the use of digital technology to deliver health care when participants are separated by distance and/or time." [32]) was valued because it gives "the ability for medical practitioners to see into your life where you are more relaxed" (P7). However, downsides raised about telehealth included: confidence, digital skills, and access to technology or connectivity. Participants expressed a clear preference to have the option of in-person consultations and acknowledged that telehealth could not replace all aspects of health care. "You can see a dead cow out there on a drone, or a cow lying down, you don't know what's wrong with it until you go up and see it" (P30).

Delivery of medications by drone (P9, and P24) was of particular interest to rural participants because accessing the local pharmacist required lengthy road trips. Unlike other rural areas, where the local rural bus can deliver medications (and shopping), many rural residents in the study area have no option but to drive to the local pharmacy in town.

One participant with low vision commented: "For our mental health, one of the main things is we need someone who we can connect with, who understands us as an individual, and we feel is on our side, that we can say, 'this is what I am doing, how do I go about it'" (P29). Health and loneliness can be intertwined, and for example, "Isolated people feel pain more than those socially connected" (P3).

Communication Gaps

Communication gaps were seen as a significant issue, especially as both rural and urban populations increase,: "the bigger the place gets, the worse the communication gets" (P5). Many participants lacked timely information about services, events, or smaller community groups, constraining their ability to access services, which would improve their social connection. This was exacerbated for people who were isolated, new to the community, or did not have English as their first language. While community groups exist, "there is no overarching

XSL•FO
RenderX

facilitation of these groups spreading the word" (P7), illustrating the opportunity for community navigators to share information to improve social connection.

Digital tools provided opportunities to communicate, but barriers such as cost, skills, and access excluded some older adults: "A large proportion of the population can't communicate using tech, or don't want to" (P2), and one participant commented that "There was an online survey for the rural bus service, I got it when it was sent out, but couldn't save it on the computer–not like pinning it on the fridge with a magnet" (P1). On the other hand, digital technology assisted with communication,: "Technology allows people to be in contact with friends who live far away" and "translation devices help people connect despite the language barrier" (P22).

Theme 2: Support

Discussions about support centered on the importance of digital literacy, reliable connectivity, support for safety both online and personal safety, and the developing skills around information management.

Digital Technology

Participants recognized the increasing integration of technology in daily life (including as a tool for sharing information) and acknowledged its advantages, including online banking, virtual travel, and telehealth. However, participants also voiced strong concerns about being excluded due to limited access, affordability, or digital literacy, noting that "as time goes on people will have to accept technology" (P11-P20). The rapid rate at which digital technology is changing was considered a barrier,*:* "This is why some people give up, and then they become isolated" (P5).

Some participants reported using devices to connect with family overseas or for entertainment. Barriers to the uptake of digital technology by older adults included low confidence and digital skills,: "Older people are feeling cut out because they aren't technologically literate" (P11-P20), fear of scams, and associated costs with "trust is a huge barrier to people using technology" (P21).

Support groups are one way of addressing low confidence and digital skills: "with the Blind Low Vision [support group], we have been learning how to use Alexa and Siri, but once you get to people age 65 - 70, anything digital is hard work if you don't understand the workings" (P28) and "I also have hearing problems, which a lot of people probably have both problems with vision and deafness" (P29). P29 added that, "the Blind Low Vision person came over and showed me how to use Siri on my phone, I now ask the phone for location directions."

Connectivity bBarriers

Connectivity barriers included unreliable rural internet and phone coverage, unclear pricing, and restrictive data caps, which exacerbated digital exclusion for older adults. Several participants stressed that knowing about and using relevant apps was as important as the infrastructure itself,: "the online communication will only work if people know about the apps, have the apps, can use these and have access to the technology" (P9). Unreliable coverage was perceived as a barrier to engagement: "You can't give your mobile number out, because no one can contact you when you get home" (P8) (due to variable coverage).

Safety and Security Concern

Safety and security concerns were a broad topic, including housing security, especially when living alone, feeling safe when out and about or at home (not just from physical attacks but from falling or injuring themselves), trusting people who came to the house such as trades people, and so on. Participants suggested that pedestrian crossings require level access and "Cross Now" audio signals that are loud enough for people who are deaf or have hearing loss to improve their safety.

Some participants did not know how to improve their internet connectivity and often worried about whether they had completed their online banking or email transactions successfully. This led to frustration and an increased fear around using the internet for certain activities. Concerns heightened when phishing was discussed, because participants worried about online fraud and their ability to recognize it,: "I cancelled my card and started from scratch again, my children helped me with that. What was scary was that the person on the phone [from the bank] said, 'you would be surprised how often that happens.' I don't know what we can do about tightening up security on devices" (P23). In addition, there can be a level of discomfort seeking help with sensitive digital tasks such as banking,: "Older people may not want strangers to help with certain aspects of technology" (P22).

Information

Information, and in particular, digital literacy, emerged as a recurring theme. As more services move online, participants are concerned about their ability to search, verify, and store information effectively,: "if older adults are active in the community, they will get information but if [they are] at home they will be isolated and not get the information" (P27). Concern about managing the "social media afterlife" (P5) following death highlighted the need for better guidance. Practical, community-based training to improve digital skills was strongly supported.

Theme 3: Spaces

Spaces for social connection and well-being were considered essential, with a preference for adapting existing facilities over constructing new ones to enable in-person interactions.

Existing Facilities

Community venues, such as halls, cafés, and churches, were described as underused yet valuable resources: "We don't need new buildings, but we should utilisze what we have" (P10, P27).

Purpose-Built Social Spaces

On the other hand, building new social spaces designed to accommodate the needs of the community by being age-friendly and accessible was considered necessary–"We need more community buildings in place" (P4, P27). Barriers identified included cost, compliance requirements, and physical accessibility challenges. Participants expressed a preference for

XSL•FO
RenderX

casual, spontaneous use of spaces, particularly cafés as community hubs.

Technology hHubs

Technology hubs co-located in libraries or health care centers were recognized as preferred spaces for supporting informal learning and scheduled training. One participant identified that "The library could encourage people to come together–they tend to do this with other groups, but not necessarily older people" (P4). These shared spaces could also support tasks such as logging into patient portals by providing technological support.

Hybrid Spaces

Hybrid spaces that allow both in-person and online participation were considered particularly valuable for individuals with low vision, mobility, or transport challenges. It was proposed that delivering introductory in-person training (eg, teaching videoconferencing skills) would increase older adults' confidence and willingness to engage online. As one participant said, "We could have an in-person session initially to show us all how to use Zoom, so we would feel more comfortable" (P2).

Theme 4: Social Connection

Social connection (also known as social connectedness) was consistently identified as a cornerstone of well-being.

In-Person Interactions and Socializing

In-person interactions and socializing had the most impact on well-being in all the focus groups and interviews. Most conversations concluded with a discussion highlighting the importance of socializing in person. In-person connection was perceived as irreplaceable, with digital tools valued only as supplementary supports. As one participant explained, "We are lucky in that we all belong to each other, we are whakapapa, and we can talk to each other, but in other communities that is not the case where they are diverse" (P24).

Everyday connections through clubs, volunteering, and informal gatherings provided purpose and belonging, with some attending clubs "for the interaction" (P10) rather than for the activity itself. Intergenerational contact, whether with family or through community programs, was described as enriching and valuable for participants,: "Mixing with the local youth is a win" (P10).

While digital platforms supported social ties during the COVID-19 lockdowns, participants agreed that physical spaces and in-person connections were essential to their well-being. Participants were clear that virtual options should complement, not replace, in-person contact.

Loneliness and iIsolation

Loneliness and isolation were attributed to a lack of social connection in the community, limited inclusive spaces, or inadequate access to digital technology, and were further linked to limited mobility or a recent relocation to the region.

As one participant stated, "without a Tablet or Smart phone, older adults miss out on the connection, and struggle to get information" (P27). New residents, in particular, described difficulties integrating and, at times, feeling unwelcome in established communities. Participants were very keen to explore social prescribing and community well-being navigators, both strategies deployed in New Zealand and overseas to increase social contact and connection, "if there was a connector that would help" (P7).

Another participant commented, "For a lot of us, it is family, I have quite a few friends who are on their own, and it is hard. I see that digital can help, but there is nothing like a person who doesn't look like they are in a hurry to get to the next job – they have time" (P29).

Not Part of a Community

Participants were concerned that newcomers from urban areas often preferred to keep to themselves and often worked from home or commuted to larger urban centers, coming home to relax. They felt that these new residents did not always recognize the importance of social and community connection and were therefore less likely to engage in community activities.

Even existing residents can feel disconnected from the wider community, with their only contact being through social services: "I deliver meals on wheels, and it's obvious that a lot of people have no contact other than the person bringing their meal" (P23).

Te Ao Māori and Connection

Māori participants emphasized the value of cultural connection—"I would be more comfortable talking to you face to face, as opposed to this (Zoom)" (P30)—including the central role of marae (Māori meeting places), which operates as multifunctional hubs that provide cultural, social, and increasingly digital engagement. A number of marae now offer fast internet, enabling them to serve as hybrid spaces, as one participant noted: "*Y*es, there is Wi-Fi on the marae…but not in our houses" (P24). This can also provide support in using technology, with one participant observing, "Technology for our community, we have old people out here (The Valley), they don't even know how to operate a mobile phone" (P31).

Discussion

Principal Findings

The 4 Ss Framework (Table 3) provides a practical lens for understanding future well-being priorities in a regional older adult population by highlighting 4 priority areas for future planning for age-friendly, inclusive communities: services, support, spaces, and social connection. Notably, participants described well-being as relational and community-based, reflecting Māori perspectives that emphasize collective rather than solely individual determinants of health. Social and community well-being emerged not only as one of the 4 Ss but also as a determinant that influences access to services, digital engagement, and the use of physical spaces.

The approach proposed aligns with internationally recognized social determinants of health and well-being [2] and the WHO dimensions of active aging [3]. The thematic findings presented here also reflect the models of Van Hoof et al [5] and Pedell and Borda [17], which emphasize the interplay between environments (spaces), service access, and social participation

XSL•FO
RenderX

in shaping aging experiences. Despite reporting confidence with using digital technologies, participants' difficulties in participating online reflected limited skills and low trust in digital systems, illustrating a gap between perceived and actual digital literacy.

These findings confirm and extend prior research demonstrating links between social isolation and poorer health outcomes [33] that telehealth faces persistent equity barriers [34] and the need to co-design housing and transport to support aging in place [13]. Alignment with global and national policies on active and healthy aging, the WHO's Global Network for AFCC [35] and New Zealand's Healthy Aging Strategy [36], is also evident due to their emphasis on transport, housing, communication, and participation as the foundation of age-friendly environments. Similarly, New Zealand's Better Later Life–He Oranga Kaumātua 2019 to 2034 strategy [6] sets a vision to enable participation and value contributions as people age. The alignment of this study's findings within a wider policy context highlights their relevance for future planning.

Most participants commented that younger generations are more digitally literate than older generations, suggesting that communities are experiencing a transition phase, with any perceived lack of digital skills primarily affecting today's older adults and possibly the subsequent generation.

Services

Participants highlighted concerns related to financial insecurity, isolation and loneliness, safety and security, stress and anxiety, and limited access to social services. Challenges in accessing social services can disproportionately affect people who are deaf or have hearing loss, as well as those with low vision, especially those without standard forms of identification, such as a driving license. Looking to the future, essential foundations for aging well in place included the importance of reliable access to health care, housing, transport, and information.

Enabling access to reliable and safe transport supports social participation and inclusion because the older adults' ability to engage in- person is contingent on accessible mobility options, an essential principle in universal design and age-friendly environments [37,38]. Autonomous (self-driving) cars were considered viable and realistic future options, either as owner-user or taxi-based models, and the deployment of these technologies is already emerging internationally [39-41]. Shu and Woo [42] anticipate that autonomous mobility may help reduce older adults' social and geographic isolation by overcoming transportation barriers. These access challenges ultimately constrained opportunities for social participation and inclusion, demonstrating the importance of treating social well-being as a core service outcome.

Health care concerns focused on access, long waits, and rural service gaps. Participants viewed telehealth and wearable technologies as useful in supporting in-person care within a hybrid delivery model. However, they emphasized that telehealth could not replace all aspects of health care and that retaining the option for face-to-face consultations remains essential for many older adults [23]. This finding reflects global evidence that the effectiveness of telehealth depends on confidence, digital skills, technology availability, and reliable connectivity [34,43-45]. Social isolation, and its impact on health, emerged as a key concern for participants, reinforcing international findings that social isolation can exacerbate pain and reduce participation in daily activities [33].

While participants in this study did not raise cost as a specific barrier to health care, this may be due to the population sampled, as several reported holding private health insurance, although many found it too costly to maintain. Despite universal health coverage being the predominant funding model in New Zealand, around 33% of the population buy additional private health insurance at some point [46]. However, in New Zealand's *Better Later Life–He Oranga Kaumātua 2019 to 2034* strategy, cost was clearly identified by older adults as a barrier to health care and well-being, for example, in relation to visiting the doctor and dentist [6].

Housing affordability and diversity of types of housing were equally pressing needs, with participants supporting smart home technologies and universal design features to enable aging in place. Smart home sensors can support safety and daily activities, enabling more older adults to age in place, which is associated with better health outcomes than relocating to assisted living [13,14]. International evidence identifies that early integration of universal design reduces costs and supports disability-inclusive housing [13,14]. Jacques [47] has argued that any new house has a 60% probability of accommodating someone with a disability during the lifespan of that house. Transport and housing are interlinked, with the location and design of new developments requiring integration with accessible public and active transport options to avoid environments that limit older adults' mobility, social participation, and inclusion and thereby creatinge islands of isolation [48].

The transition from print to digital was seen as a significant communication challenge that risks excluding those lacking digital access or skills, thus increasing social isolation [49]. This reflects broader international concerns that digital transitions can exacerbate inequalities in information access, leading to digital exclusion for some communities [50,51]. Seifert et al [52] found that older adults often feel excluded when essential communication moves online without alternative nondigital alternatives. Vaca-Benavides et al [53] also caution that digital transitions in health and social care services can exclude older adults who struggle with digital confidence, access, connectivity, or support.

Support

Trust in digital technology, in the systems that support it, and in participants' own ability to use it emerged as a significant concern and presented issues with information access. Barriers identified included cost, the rapid pace of technological change, and limited training opportunities. One option to address this is to provide community well-being navigators (through local libraries) to support older adults with digital literacy and increase their confidence in the use of digital technologies [54]. Vercruyssen et al [55] observed that for older adults with minimal or no prior experience with digital technology,

XSL•FO
RenderX

acquiring even the most basic or entry-level digital literacy skills is difficult.

Internet and phone connectivity issues, experienced particularly by participants living in rural areas, compounded these barriers to equitable digital access. Even in areas with theoretical internet coverage, participants experienced unreliable connections, consistent with national and global studies highlighting the last mile of the digital divide [23,56].

The participants also raised safety and security concerns about the use of digital technologies, including online risks (eg, fraud and identity theft) and physical vulnerabilities. Technologies such as alarms and GPS trackers were predominantly welcomed, especially for dementia care; however, participants' low levels of trust presented barriers to digital engagement. Kebede et al [57] reported that safety and digital confidence issues represent key barriers to digital engagement, particularly in relation to web-based digital technologies. The ability to access digital information and community services was considered essential to well-being [43], because unreliable access increased the risk of social isolation.

Digital literacy is essential for civic participation and safety. Older adults are at risk of digital exclusion without targeted support [55]. Digital skills therefore support more than information access; they enable older adults' civic participation and social connection, which is critical to community well-being. Seifert et al [52] described a "double burden of exclusion" affecting older adults who face both digital and social exclusion. Other research illustrates that affordability, trust, and skills gaps continue to drive digital exclusion among older adults [58,59], mirroring the findings of this study.

Spaces

Accessible, inclusive, and multifunctional community spaces that support social and recreational activities were considered essential for well-being. One key finding was participants' clear preference for reusing existing facilities as social spaces, such as community halls, cafés, and churches, rather than building new ones, which is consistent with international research. Liddle et al [60] recognized the importance of making better use of existing spaces to improve social connections and considered how to repurpose them. Barriers identified to repurposing shared spaces included cost, compliance requirements, and accessibility challenges. Integrating age-friendly universal design into both housing and public infrastructure could reduce these barriers and promote independence, because urban design influences opportunities for social engagement and inclusion [61]. Physical places that enable social connection were regarded as essential infrastructure for community well-being, not merely a venue for activities. With older adults' strong preference for in-person interaction, it is essential that we retain and repurpose existing spaces to support their well-being. While digital technology can complement these physical spaces, it cannot substitute for them, particularly for those at risk of digital exclusion.

Social Connection

Social and community well-being emerged as a key determinant of healthy aging, influencing participation, belonging, and independence across all aspects of daily life. Social interaction, particularly in-person connection, was consistently described as essential to well-being, especially for those with low vision and hearing loss. A lack of social connection was perceived to contribute to loneliness and isolation, negatively impacting mental health. International research has similarly identified that increasing opportunities for in-person social contact can reduce loneliness and improve well-being in older adults, consistent with the findings of this study [61].

Participants emphasized the value of social prescribing [62] and community well-being navigators [54], both of which are emerging as strategies in New Zealand and internationally to increase social contact and reduce isolation. Māori participants emphasized the importance of marae as multifunctional hubs, reinforcing indigenous models of well-being where holistic dimensions are interconnected. The exploration of marae-based approaches, and similar collective spaces, is recommended to strengthen belonging and intergenerational ties to support social connection. Participants expressed a clear preference for approaches that combine digital access with ongoing in-person contact, such as hybrid models, reflecting findings internationally that found face-to-face interaction crucial for older adults' well-being [63-65].

In summary, these findings suggest that future age-friendly planning must prioritize social and community well-being outcomes, ensuring that investments in services, digital inclusion, and public spaces strengthen older adults' ability to connect with others and participate fully in community life.

Recommendations

The 4 Ss Framework is proposed as a planning framework for future age-friendly, inclusive communities: services, support, spaces, and social connection. It can be used to assess current opportunities, identify gaps in existing opportunities, and thus inform future planning. Using the framework, the authors propose a set of actionable recommendations to guide planning for age-friendly regional communities over the next 30 years. Table 4 summarizes these recommendations, offering actionable guidance for service providers, policymakers, and community organizations. This is also an example of how a community could use the framework for its future age-friendly planning.

These recommendations, drawing on the 4 Ss Framework, are designed to inform future planning for age-friendly regional communities over the next 30 years, as well as corresponding policy and practice implications. The findings highlight the importance of hybrid service models, digital inclusion, universal design, and social connection and emphasize the importance of embedding aging perspectives in policy and planning. Finally, most participants commented that younger generations are more digitally literate than they are, suggesting that communities are experiencing a transition phase , with any perceived lack of digital skills primarily affecting today's older adults, and possibly, the subsequent generation.

The future offers many opportunities for supporting the well-being of individuals and communities, and the findings of this project are supported by existing models of wellness. Using the models and strategies in the literature, highlighted by this thematic analysis and the Te Whare Tapa Whā model, it is

XSL•FO
RenderX

possible to create a matrix to direct planning and evaluation of interventions and initiatives to improve wellness in age-friendly communities that meet Te Tiriti o Waitangi obligations and align with New Zealand healthy aging strategies. Table 5 shows an example of such a matrix, with potential initiatives available in New Zealand.

Table . Recommendations for supporting older adults' well-being in regional communities.

Theme (4 Ss Framework)	Recommendation	Rationale/evidence
Services	Prioritize hybrid service models that combine digital and in-person delivery	Digital tools are useful but cannot replace in-person interaction; hybrid models improve access and support well-being
Support	Reenvisage community spaces as multifunctional hubs for digital training, service access, and social connection	Lack of digital skills and trust creates exclusion, training in familiar settings, builds confidence and inclusion
Spaces	Apply universal design to housing, public infrastructure, and transport	Universal design ensures accessibility; improved transport access enables participation, independence, and aging well in place
Social connection	Establish well-being navigator roles to support older adults in accessing services	Limited trust in technology is a barrier; navigators build trust, reduce isolation, and strengthen community networks

Table . Proposed matrix for wellness interventions, with some examples.

Examples of wellness interventions	Services	Skills	Spaces	Social connectedness
Physical health/te taha tinana	Meals on wheels	Fitbits, waka ama	Gyms, safe pavements	Social prescribing
Spiritual health/te taha wairua (sense of meaning and purpose)	Older adults working with youth	Virtual reality journeys	Marae, museum	Kai Hub, Kapa haka
Mental or emotional health/te taha hinengaro	Well-being navigators	Mobile banking	Safe public gardens	Repair cafes and Menzshed
Family/te taha whānau (social well-being)	Free hearing aids	SeniorNet	Library	Skinny Jump, social media

Limitations

This study was based on a small regional population in New Zealand and relied on self-reported experiences, both of which are common limitations in focus group studies. The COVID-19 public health restrictions and national lockdowns at the time shaped recruitment and influenced preferences for individual interviews over group discussions, as well as limitations to completing documentation such as consent forms. As the number of individual interviews grew, fewer new ideas emerged; however, determining data saturation remains contentious, with varying definitions across qualitative research. Therefore, caution is required when applying the findings beyond the study location.

Future Research

Opportunities for further studies include testing the 4 Ss Framework in other settings, for example, with older adults with low vision examining postpandemic digital adoption and evaluating the effectiveness of social prescribing and well-being navigator roles in facilitating digital engagement. The authors are not qualified to speak on Te Ao Māori but suggest a closer exploration of how marae support, and can be supported, to enhance the well-being of Māori, and indeed other Indigenous peoples, recognizing not only the importance of maintaining tradition and belonging, but also as to how it can be applied to non-Māori and other non-Indigenous populations.

Conclusions

The future presents significant opportunities to enhance the well-being of older adults and communities through targeted, inclusive, and innovative approaches. Building on the thematic analysis and existing strategies, this study proposes a practical framework (the 4 Ss Framework) to guide future planning for age-friendly communities. As a methodology, the 4 Ss Framework provides a structured approach to anticipating the future well-being needs of older adults to support them to age well and age in place. By translating the lived experiences of older adults into an actionable framework, this study contributes to local and international efforts to create inclusive, resilient, and age-friendly communities. The 4 Ss Framework reflects the priorities of the WHO Global Network for AFCC, the UN Decade of Healthy Ageing, and OECD (Organisation for Economic Co-operation and Development) well-being agendas, offering a transferable approach designed as a tool to support and strengthen the future well-being of communities.

XSL•FO
RenderX

Acknowledgments

The authors thank the Manawatū Community Trust, a not-for-profit organization established by the Manawatū District Council in 2008, for their collaboration and support in this research. The authors are especially grateful to all participants for their enthusiasm and willingness to contribute to the focus groups and interviews. The authors also extend special thanks to Olivia Spencer for her invaluable assistance with drafting this paper. No generative AI was used at any stage.

Funding

This study was funded by the Manawatū Community Trust (Manawatū Community Trust is a not-for-profit organisationorganization established in 2008 by the Manawatū District Council, with a mission to “promote and provide housing and well-being services that are relevant and meet the long-term community needs” [https://manawatucommunitytrust.nz]).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: CL, IH
Data curation: CL, IH
Formal analysis: CL, IH
Funding acquisition: CL, IH
Investigation: CL, IH
Methodology: CL, IH
Project administration: CL
Supervision: IH
Writing – original draft: IH
Writing – review and editing: CL, IH

Conflicts of Interest

None declared.

References

1. Ottawa charter for health promotion. : World Health Organization; 2012 URL: https://www.who.int/publications/i/item/ottawa-charter-for-health-promotion [accessed 2025-10-30]
2. Office of Disease Prevention and Health Promotion. Social determinants of health. Healthy People 2030. URL: https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health [accessed 2025-08-17]
3. Active ageing: a policy framework. : World Health Organization; 2002 URL: https://iris.who.int/server/api/core/bitstreams/0418705f-1c82-4fe2-82d6-d4faee89dfa0/content [accessed 2025-10-30]
4. Publications. Ministry of Health NZ (Manatū Hauora). URL: https://www.health.govt.nz/publications [accessed 2026-09-03]
5. Van Hoof J, Kazak JK, Perek-Białas JM, Peek STM. The challenges of urban ageing: making cities age-friendly in Europe. Int J Environ Res Public Health 2018 Nov 5;15(11):2473. [doi: 10.3390/ijerph15112473] [Medline: 30714577]
6. Better later life–He Oranga Kaumātua 2019–2034. : Office for Seniors (Te Tari Kaumātua); 2022 URL: https://tinyurl.com/2exu487z [accessed 2026-09-03]
7. Sánchez-González D, Rojo-Pérez F, Rodríguez-Rodríguez V, Fernández-Mayoralas G. Environmental and psychosocial interventions in age-friendly communities and active ageing: a systematic review. Int J Environ Res Public Health 2020 Nov 10;17(22):8305. [doi: 10.3390/ijerph17228305] [Medline: 33182710]
8. Torku A, Chan APC, Yung EHK. Age-friendly cities and communities: a review and future directions. Ageing Soc 2021 Oct;41(10):2242-2279. [doi: 10.1017/S0144686X20000239]
9. Promoting healthy ageing in a digital world. World Health Organization Regional Office for Europe. URL: https://www.who.int/europe/news-room/events/item/2024/09/17/default-calendar/promoting-healthy-ageing-in-a-digital-world [accessed 2025-08-17]
10. Loneliness and isolation – the hidden threat to global health we can no longer ignore. World Health Organization. URL: https://www.who.int/news-room/commentaries/detail/loneliness-and-isolation-the-hidden-threat-to-global-health-we-can-no-longer-ignore [accessed 2025-08-17]
11. Social isolation and loneliness. World Health Organization. URL: https://www.who.int/teams/social-determinants-of-health/demographic-change-and-healthy-ageing/social-isolation-and-loneliness [accessed 2025-08-17]
12. The world health assembly endorses the extension of the global strategy on digital health to 2027 and approves the next phase for 2028–2033. World Health Organization. URL: https://www.who.int/news/item/23-05-2025-world-health-assembly-endorses-extension-of-the-global-digital-health-strategy-to-2027 [accessed 2025-08-17]

XSL•FO
RenderX

13. Elers P, Hunter I, Whiddett D, Lockhart C, Guesgen H, Singh A. User requirements for technology to assist aging in place: qualitative study of older people and their informal support networks. JMIR mHealth uHealth 2018 Jun 6;6(6):e10741. [doi: 10.2196/10741] [Medline: 29875083]
14. Hunter I, Elers P, Lockhart C, Guesgen H, Singh A, Whiddett D. Issues associated with the management and governance of sensor data and information to assist aging in place: focus group study with health care professionals. JMIR mHealth uHealth 2020 Dec 2;8(12):e24157. [doi: 10.2196/24157] [Medline: 33263551]
15. Hunter IM, Lockhart CA. Use of sensor data to support interRAI assessments of older adults living at home in New Zealand. : Massey University; 2020 URL: https://ehealthforum.nz/t/the-aim-of-this-report-is-to-present-findings-from-research-undertaken-into-the-use-of-sensor-data-available/24005 [accessed 2026-09-03]
16. Sharif RA, Pokharel S. Smart city dimensions and associated risks: review of literature. Sustainable Cities Soc 2022 Feb;77:103542. [doi: 10.1016/j.scs.2021.103542]
17. Pedell S, Borda A. This is how we create the age-friendly smart city. The Conversation. 2021 Mar 3. URL: https://theconversation.com/this-is-how-we-create-the-age-friendly-smart-city-152973 [accessed 2026-09-03]
18. Hauora: report on stage one of the health services and outcomes kaupapa inquiry. : Waitangi Tribunal; 2023 URL: https://forms.justice.govt.nz/search/Documents/WT/wt_DOC_195476216/Hauora%202023%20W.pdf [accessed 2026-09-03]
19. Te Whare Tapa Whā: Māori health model. : Ministry of Health – Manatū Hauora; 2025 URL: https://www.health.govt.nz/system/files/2024-05/maori_health_model_tewhare.pdf [accessed 2025-08-17]
20. Boulton AF, Gifford HH, Potaka-Osborne M. Realising whānau ora through community action: the role of Māori community health workers. Educ Health (Abingdon) 2009 Aug;22(2):188. [Medline: 20029746]
21. Age composition. Regional Economic Profile.: Infometrics URL: https://rep.infometrics.co.nz/manawatu-district/population/age-composition?compare=new-zealand [accessed 2025-09-15]
22. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544. [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
23. Hunter I, Lockhart C, Rao V, Tootell B, Wong S. Enabling rural telehealth for older adults in underserved rural communities: focus group study. JMIR Form Res 2022 Nov 4;6(11):e35864. [doi: 10.2196/35864] [Medline: 36331533]
24. Corbin J, Morse JM. The unstructured interactive interview: issues of reciprocity and risks when dealing with sensitive topics. Qual Inq 2003 Jun;9(3):335-354. [doi: 10.1177/1077800403009003001]
25. Braun V, Clarke V, Terry G. Thematic analysis. In: Rohleder P, Lyons AC, editors. Qualitative Research in Clinical and Health Psychology: Palgrave Macmillan; 2015:95-113.
26. Braun V, Clarke V. What can “thematic analysis” offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152. [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
27. Code of ethical conduct for research, teaching and evaluations involving human participants. : Massey University; 2026 URL: https://www.massey.ac.nz/documents/1590/Code_Ethical_Conduct_Research_Teaching_Evaluations_Involving_Human_Participants.pdf [accessed 2025-08-18]
28. Human ethics. Massey University. URL: https://www.massey.ac.nz/research/ethics/human-ethics [accessed 2025-08-18]
29. Ethnicity classifications levels list (XLSX dataset). : Te Whatu Ora – Health New Zealand; 2024 URL: https://tinyurl.com/ymu4c6aj [accessed 2025-08-28]
30. Place and ethnic group summaries: Manawatu district. Stats NZ. 2018. URL: https://www.stats.govt.nz/tools/2018-census-place-summaries/manawatu-district [accessed 2025-08-18]
31. Māori population estimates: At 30 June 2020. Stats NZ. 2020 Jun. URL: https://www.stats.govt.nz/information-releases/maori-population-estimates-at-30-june-2020 [accessed 2025-06-18]
32. NZ Telehealth–2022 Funder Report. : New Zealand Telehealth Forum; 2022 URL: https://www.telehealth.org.nz/assets/2022/Reports/Telehealth-Funder-Report-2022-FINAL.pdf [accessed 2025-09-12]
33. Karayannis NV, Baumann I, Sturgeon JA, Melloh M, Mackey SC. The impact of social isolation on pain interference: a longitudinal study. Ann Behav Med 2019 Jan 1;53(1):65-74. [doi: 10.1093/abm/kay017] [Medline: 29668841]
34. Kichloo A, Albosta M, Dettloff K, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. Fam Med Community Health 2020 Aug;8(3):32816942. [doi: 10.1136/fmch-2020-000530] [Medline: 32816942]
35. About the global network for age-friendly cities and communities. WHO Age-friendly World. URL: https://extranet.who.int/agefriendlyworld/who-network/ [accessed 2025-08-18]
36. Healthy ageing strategy. : Ministry of Health–Manatū Hauora; 2016 URL: https://www.health.govt.nz/system/files/2016-12/healthy-ageing-strategy_june_2017.pdf [accessed 2025-08-17]
37. Maisel J, Steinfeld E. Universal Design: Creating Inclusive Environments: John Wiley & Sons; 2012.
38. Global age-friendly cities: a guide. : World Health Organization; 2007 URL: https://iris.who.int/server/api/core/bitstreams/08caa73b-5934-4b6b-bffe-d27fbe92e6e7/content [accessed 2025-10-23]
39. Fagnant DJ, Kockelman K. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transp Res A Policy Pract 2015 Jul;77:167-181. [doi: 10.1016/j.tra.2015.04.003]

XSL•FO
RenderX

40. McLoughlin S, Prendergast D, Donnellan B. Autonomous vehicles for independent living of older adults—insights and directions for a cross-European qualitative study. 7th International Conference on Smart Cities and Green ICT Systems 2018;1:294-303. [doi: 10.5220/0006777402940303]
41. Kim P. Adoption of autonomous vehicles among the aging population. Issues Inf Syst 2024;25(2):24-35. [doi: 10.48009/2_iis_2024_103]
42. Shu S, Woo BKP. Applications of self-driving vehicles in an aging population. JMIR Form Res 2025 Apr 28;9(1):e66180. [doi: 10.2196/66180] [Medline: 40294433]
43. Pedell S, Borda A, Keirnan A, Aimers N. Combining the digital, social and physical layer to create age-friendly cities and communities. Int J Environ Res Public Health 2021 Jan 5;18(1):325. [doi: 10.3390/ijerph18010325] [Medline: 33466259]
44. Fernandes LG, Devan H, Fioratti I, Kamper SJ, Williams CM, Saragiotto BT. At my own pace, space, and place: a systematic review of qualitative studies of enablers and barriers to telehealth interventions for people with chronic pain. Pain 2022 Feb 1;163(2):e165-e181. [doi: 10.1097/j.pain.0000000000002364] [Medline: 34433776]
45. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
46. International health care system profiles: New Zealand. The Commonwealth Fund. 2020. URL: https://www.commonwealthfund.org/international-health-policy-center/countries/new-zealand [accessed 2025-08-17]
47. Jaques R. Rethinking universal design. In: BRANZ Build: BRANZ (Building Research Association of New Zealand); 2013, Vol. 138:79-80 URL: https://www.buildmagazine.org.nz/articles/show/rethinking-universaldesign [accessed 2025-08-18]
48. Mendly-Zambo Z, Power L, Khan A, Bryant T, Raphael D. Islands of isolation in a modern metropolis: social structures and the geography of social exclusion in Toronto, Ontario, Canada. CJUR 2021;30:101-118. [doi: 10.36939/cjur/vol30no--/art353]
49. Stansberry K, Anderson J, Rainie L. 5. Leading concerns about the future of digital life. : Pew Research Center; 2019 URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2019/10/PI_2019.10.28_The-Next-50-Years-of-Digital-Life_FINAL.pdf [accessed 2025-08-18]
50. United Nations E-Government Survey 2022: The Future of Digital Government. : United Nations; 2022 URL: https://desapublications.un.org/sites/default/files/publications/2022-09/Web%20version%20E-Government%202022.pdf [accessed 2025-10-23]
51. van Dijk J. The Digital Divide: Polity Press; 2020.
52. Seifert A, Cotten SR, Xie B. A double burden of exclusion? Digital and social exclusion of older adults in times of COVID-19. J Gerontol B Psychol Sci Soc Sci 2021 Feb 17;76(3):e99-e103. [doi: 10.1093/geronb/gbaa098] [Medline: 32672332]
53. Vaca-Benavides DA, Uthayakumar S, McSwiggan E, et al. Experiences of digital transitions in health and social care services in later life: findings and reflections from a collaborative, interdisciplinary exploratory review. Soc Sci 2024 Oct 1;13(10):526. [doi: 10.3390/socsci13100526]
54. Lockhart C, Hunter I. Using ‘health’ to promote older adults’ digital health literacy. Med Res Arch 2024;12(6). [doi: 10.18103/mra.v12i6.5579]
55. Vercruyssen A, Schirmer W, Geerts N, Mortelmans D. How “basic” is basic digital literacy for older adults? Insights from digital skills instructors. Front Educ 2023;8. [doi: 10.3389/feduc.2023.1231701]
56. Lourens M. Explainer: why is mobile reception so bad in parts of our biggest cities? Stuff. 2022. URL: https://www.stuff.co.nz/national/explained/128767213/explainer-why-is-mobile-reception-so-bad-in-parts-of-our-biggest-cities [accessed 2025-08-18]
57. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: acoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
58. Mohan R, Saleem F, Voderhobli K, Sheikh-Akbari A. Ensuring sustainable digital inclusion among the elderly: a comprehensive analysis. Sustainability 2024;16(17):7485. [doi: 10.3390/su16177485]
59. Wale A, Everitt J, Ayres T, et al. Effectiveness of interventions for addressing digital exclusion in older adults in the social care domain: rapid review. JMIR Aging. 2025 Dec 30 p. e70377. [doi: 10.2196/70377] [Medline: 41468583]
60. Liddle J, Pitcher N, Montague K, Hanratty B, Standing H, Scharf T. Connecting at local level: exploring opportunities for future design of technology to support social connections in age-friendly communities. Int J Environ Res Public Health 2020 Jul 31;17(15):5544. [doi: 10.3390/ijerph17155544] [Medline: 32751898]
61. Cowan M, Dupre K, Fernando R. Exploring the relationship between urban design and social capital: a systematic quantitative literature review. Architecture 2024 Sep;4(3):493-514. [doi: 10.3390/architecture4030027]
62. Oster C, Skelton C, Leibbrandt R, Hines S, Bonevski B. Models of social prescribing to address non-medical needs in adults: a scoping review. BMC Health Serv Res 2023 Jun 15;23(1):642. [doi: 10.1186/s12913-023-09650-x] [Medline: 37316920]
63. Cone N, Lee JE. Older adult communication types and emotional well-being outcomes during COVID-19 pandemic. BMC Geriatr 2023 Mar 27;23(1):178. [doi: 10.1186/s12877-023-03856-8] [Medline: 36973663]
64. Khanassov V, Ilali M, Ruiz AS, Rojas-Rozo L, Sourial R. Telemedicine in primary care of older adults: a qualitative study. BMC Prim Care 2024 Jul 17;25(1):259. [doi: 10.1186/s12875-024-02518-x] [Medline: 39020277]

XSL•FO
RenderX

65. Laukka E, Lakoma S, Harjumaa M, et al. Older adults' preferences in the utilization of digital health and social services: a qualitative analysis of responses to open-ended questions. BMC Health Serv Res 2024 Oct 4;24(1):1184. [doi: 10.1186/s12913-024-11564-1] [Medline: 39367429]

Abbreviations

AFCC: Age-Friendly Cities and Communities
OECD: Organisation for Economic Co-operation and Development
WHO: World Health Organization

Edited by A Garcia; submitted 06.Jan.2026; peer-reviewed by J Waycott, K Maki, TAK Manne; revised version received 10.Aug.2026; accepted 13.Aug.2026; published 25.Sep.2026.

Please cite as:
Hunter I, Lockhart C
The 4 Ss Framework for Designing Age-Friendly Communities of the Future: Focus Group Study
JMIR Aging 2026;9:e90923
URL: https://aging.jmir.org/2026/1/e90923
doi:10.2196/90923

© Inga Hunter, Caroline Lockhart. Originally published in JMIR Aging (https://aging.jmir.org), 25.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smart Speaker–Based Applications to Support Social Connectedness in Older Adult Residents in Affordable Housing: User-Centered Design Study

Jane Chung[1], PhD, RN; Natalie Mansion[2], MS; Tracey Gendron[3], MS, PhD; Rachel E Wood[4], PhD, RN; George Demiris[5], PhD

[1]Nell Hodgson Woodruff School of Nursing, Emory University, 57 Executive Park Dr NE, Atlanta, GA, United States

[2]School of Nursing, Virginia Commonwealth University, Richmond, VA, United States

[3]Department of Gerontology, College of Health Professions, Virginia Commonwealth University, Richmond, VA, United States

[4]Department of Nursing, Longwood University, Farmville, VA, United States

[5]School of Nursing, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Jane Chung, PhD, RN
Nell Hodgson Woodruff School of Nursing, Emory University, 57 Executive Park Dr NE, Atlanta, GA, United States

Abstract

Background: Older adults in affordable housing face heightened risks of social isolation and loneliness due to limited social networks, transportation barriers, chronic conditions, and inadequate technology access. Smart speakers offer potential for enhancing social connectedness in this underserved population, yet technology interventions are rarely designed with meaningful input from older adults themselves. User-centered design (UCD) approaches can address this gap by engaging end users throughout the development process to ensure technology solutions align with their needs and living contexts.

Objective: This study aimed to engage older adults in affordable housing in an iterative UCD process to develop prototype scenarios for smart speaker–based applications that promote social connectedness while addressing safety, community-building, and wellness needs.

Methods: We conducted a 3-stage UCD study with 29 older adults (mean age 70, SD 6.8 years; 23/29, 79% African American; 20/29, 69% high school education or less) living alone in affordable housing between April 2021 and April 2022. Stage 1 included 5 focus groups (n=25) combining needs assessment discussions with rapid brainstorming activities. Stage 2 involved research team synthesis of focus group transcripts and brainstorming data to create a Design Strategies Map, development of initial prototype scenarios, 5 evaluation focus groups (n=18) to gather feedback, and iterative scenario refinement. Stage 3 comprised 4 validation focus groups (n=17) to assess refined scenarios and identify implementation recommendations. Participants included both smart speaker users (n=13) and nonusers (n=16). Data were analyzed using thematic analysis for needs assessment, content analysis for brainstorming ideas and feedback, and matrix analysis for systematic comparison across scenarios.

Results: Participants generated 153 ideas for smart speaker use, with Health and Safety and Daily Assistance being the most frequent categories. Analysis revealed that social connection needs were inseparable from safety concerns related to living alone. Through iterative co-design, we developed 7 prototype scenarios across 4 functional categories: Checking-In (peer and management safety verification with privacy controls), Social Companion (conversational artificial intelligence–based companionship and emotional support), Community Involvement (virtual bulletin boards and activity coordination), and Wellness Check (system-initiated monitoring of activity and behavioral patterns as health indicators with user-controlled interventions). Participants emphasized requirements for personalization, opt-in/opt-out controls, “Do-Not-Disturb” functionality, and safeguards preventing replacement of human connection.

Conclusions: Older adults in affordable housing engaged in technology design and provided valuable insights that challenge assumptions about their needs and preferences. The prototype scenarios addressed the dual imperatives of social connection and safety while living alone, offering a foundation for developing technology-based applications tailored to underserved populations. Implementation should prioritize user control, privacy protection, and human-in-the-loop design ensuring that technology facilitates rather than replaces human connection and community programming, alongside consideration of user characteristics to build trust and ensure effective, sustained use of the intended technology platform.

(JMIR Aging 2026;9:e90053) doi:10.2196/90053

XSL•FO
RenderX

KEYWORDS

social connectedness; loneliness; social isolation; smart speakers; disparities; user-centered design; community-dwelling older adults; focus groups

Introduction

Background

Research has repeatedly demonstrated the critical importance of social connectedness in older age. Social isolation and loneliness, the main constructs of social disconnectedness, are prevalent in Americans aged 65 years and older, with prevalence rates approaching nearly 25% for the former [1] and ranging from 17% to 57% for the latter [2-4]. Social isolation is objectively defined as the lack of social contact, communications, and relationships [1,5], while loneliness refers to a subjective feeling resulting from a perceived deficit in the quantity or quality of meaningful social relationships [6]. Older adults who are socially isolated or feel lonely exhibit a wide range of negative health outcomes, including depression, cardiovascular diseases, poor cognitive function, reduced quality of life, and increased mortality [2,5,7,8].

Social isolation and loneliness are pressing concerns for older adults living in affordable housing (hereafter referred to as "residents") because they have limited social networks, transportation options, and access to information and communication technology (ICT) [9]. Chronic health conditions and disabilities are also common in this population [10,11], which affect their ability to establish and maintain meaningful social interactions. Also, affordable housing, particularly in high-rise structures, may lack communal areas [12] or discourage residents from inviting others over due to suboptimal building conditions [13]. The complex interplay of health, housing conditions, and social and financial factors may contribute to elevated levels of social isolation and loneliness among residents [14]. Innovative approaches are necessary to increase meaningful social connections and reduce loneliness in this population. It is recognized that there is no "one-size-fits-all" approach to address this issue and there is a need to design interventions customized to meet specific groups or the varying levels of social isolation and loneliness they may experience [15].

A notable gap exists in technology-based interventions tailored to enhance social connectedness among residents. While face-to-face programs and services for improving social connections are available for older adults (eg, community senior centers and senior cafés), these may be inaccessible to older adults residing alone in affordable housing for multiple reasons mentioned previously. The growing interest in voice-operated smart speakers (eg, Amazon Echo and Google Home) has increased an exploration into how this technology can enhance social connectedness in community-dwelling older adults [16-18]. The hands-free voice control may improve accessibility and usability for people with vision impairment, mobility limitation, or poor dexterity. Artificial intelligence (AI)–based voice agents (eg, Amazon Alexa) on smart speakers have the potential to help older adults in limited-resource settings feel less lonely and engage more in social activities, for example, by providing easy access to communicate with others. Studies provide preliminary evidence that older adults enjoy social health benefits from interactions with an AI voice agent, such as a source of companionship reducing loneliness and improving mental health [18,19].

Despite the potential benefits of smart speakers, unique challenges to their adoption and use may arise, particularly when the needs of target end users are overlooked or inadequately addressed. Moreover, historically marginalized communities, including low-income older adults and racial or ethnic minority groups, still face inequity in ICT access as well as challenges in its adoption and use [20-22]. Low ICT adoption and utilization are attributed to various barriers, including technology costs, internet accessibility, functional impairments, low literacy, lack of self-efficacy, privacy concerns, and mistrust of the technology [23-28]. Recognizing the unique needs of users and the factors influencing acceptance and usability within the user's living context is critical for implementing technology-based interventions. This can be accomplished through user-centered design (UCD), which involves exploring user needs, preferences, and technology requirements; engaging users throughout the design process; and iteratively developing technology prototypes [29].

Aims

The purpose of the study was to engage older adult residents living alone in affordable housing in an iterative UCD process to develop prototype ideas of innovative solutions leveraging smart speaker technology for social connectedness. Through a 3-stage co-design process, we sought to (1) understand the social connection needs and technology preferences of residents, (2) collaboratively develop and refine prototype scenarios that address these needs, and (3) identify design requirements and implementation considerations for future deployment. Our work focused on *early* prototyping to develop technology solutions and build a shared understanding of the social connection needs in this population. Our long-term goal is to develop smart speaker–based applications called *Voice2Connect* for deployment in affordable housing to enhance social connectedness among their residents.

Methods

Study Design

We conducted a UCD study to engage residents in developing smart speaker–based applications for social connectedness. Our process consisted of three major stages conducted between April 2021 and April 2022: (1) needs assessment and idea generation, (2) iterative prototype development, and (3) validation and implementation planning (Table 1).

XSL•FO
RenderX

Table . Study design overview.

Stage and phases	Activities	Participants	Key outcomes
Stage 1: Needs assessment			
Phase 1	• Focus group discussions • Video introduction to smart speakers • 6-8-5 brainstorming exercise • Pictorial sheet ideation	25 older adults (5 groups; 12 users and 13 nonusers)	• Social connection needs • 153 design ideas • User requirements
Stage 2: Iterative development			
Phase 2	• Team synthesis of phase 1 data • Design Strategies Map creation • Initial scenario development	Research team	• 4 functional categories • 7 initial scenarios
Phase 3	• Scenario presentations via role-play • Feedback focus groups	18 older adults (5 focus groups; 8 users and 10 nonusers)	• Design preferences • Concerns identified • Improvement suggestions
Phase 4	• Matrix analysis of feedback • Scenario refinement	Research team	• 7 refined scenarios
Stage 3: Validation			
Phase 5	• Refined scenario presentations • Validation focus groups • Implementation planning	17 older adults (4 groups; 5 users and 12 nonusers)	• Validated prototypes • Deployment recommendations

Setting and Participants

We recruited participants from 3 housing buildings in Richmond, Virginia, United States (referred to as housing A, B, and C) that serve older adults with limited income and individuals with disabilities. Our research team had established relationships with these communities through prior collaborative research and wellness programs. Eligible participants were older than 55 years, lived alone, could speak and understand English, and were able to participate in all study activities. We excluded individuals with severe sensory impairments preventing engagement with audiovisual materials, terminal illness, or a score ≤8 on a short version of the Montreal Cognitive Assessment indicating moderate to severe cognitive impairment. Participants were divided into 2 groups based on smart speaker experience: users (n=13) who had been using Amazon Echo Dot devices deployed by housing A since January 2020, and nonusers (n=16) from all 3 housing sites.

We recruited participants through multiple channels including housing management referrals, community wellness program participant lists, on-site information sessions, and flyers. Due to pandemic-related visitor restrictions, eligibility screening was conducted by phone, during which we also collected demographic and technology experience information. Detailed descriptions of recruitment procedures can be found elsewhere [30].

Ethical Considerations

The Institutional Review Board of Virginia Commonwealth University determined that this project posed minimal risk and waived written informed consent (HM20020008).

Procedures

We conducted focus groups with older adults in phases 1, 3, and 5 to perform an iterative co-design process and gather data. The groups were stratified by housing building and current smart speaker use (Table 2).

XSL•FO
RenderX

Table . Focus group stratification by phase and participant group (older adult residents).

Phase and building	Smart speaker user	Values, n	Total values
1			25
A	Yes	6	
A	Yes	5	
B[a]	No	6	
A	No	5	
C	No	3	
3			18
A	Yes	4	
A	Yes	4	
A	No	4	
B	No	4	
C	No	2	
5			17
A	Yes	5	
A	No	2	
B	No	6	
C	No	4	

[a]One nonuser resident from building A attended this meeting.

Stage 1: Needs Assessment and Idea Generation (Phase 1)

We conducted 5 focus groups (n=25) between April and July 2021. Due to pandemic restrictions, we held the first 3 meetings in a campus lecture room, providing van transportation since most participants lacked reliable transportation. Safety protocols included temperature checks and mask requirements.

Each 1 - to 1.5-hour session consisted of 2 parts. First, we explored participants' needs for establishing and maintaining meaningful social contacts and their desired outcomes from smart speakers. To help nonusers understand the technology, we presented a 5-minute video demonstrating basic smart speaker tasks (weather checking, drop-in calls, and dictionary functions). Second, we facilitated rapid brainstorming using the 6-8-5 design method [31], where participants generated 6 - 8 ideas in 5 minutes. Participants worked in mini groups of 2 - 3 with a facilitator, sharing ideas and asking clarifying questions. We used pictorial sheets (Multimedia Appendix 1) where participants articulated specific social connectedness tasks they wanted to accomplish using smart speakers. After 2 brainstorming rounds, mini groups presented key ideas to the larger group. Audio-recorded sessions were transcribed verbatim and analyzed as described in the Data Analysis section.

Stage 2: Iterative Prototype Development (Phases 2-4)

Phase 2: Initial Scenario Development

The research team synthesized phase 1 data through multiple team meetings, creating a Design Strategies Map that linked user needs with potential smart speaker–based solutions. Through triangulation of focus group data and pictorial sheet ideas, we identified 4 main functional categories: checking-in, social companion, community involvement, and wellness check. We developed 1 - 2 scenarios within each category—short stories depicting user characteristics, context of use, and specific technology-based tasks to achieve goals [32,33]. Scenarios incorporated voice interactions with "Alexa" and were written at a third-grade reading level to ensure readability and accessibility.

Phase 3: Evaluation Focus Groups

We conducted 5 focus groups (n=18) in October 2021, presenting initial scenarios through role-play demonstrations by 2 team members. After each scenario, we solicited feedback on design perceptions, willingness to use, improvement suggestions, and concerns.

Phase 4: Scenario Refinement

Phase 3 focus group recordings were transcribed verbatim and analyzed as described in the Data Analysis section. We identified key suggestions for each scenario and revised prototypes accordingly.

Stage 3: Validation and Implementation Planning (Phase 5)

We conducted 4 focus groups (n=17) in April 2022 to validate refined scenarios. Team members again used role-play to present scenarios, explicitly noting incorporation of previous feedback. We solicited perceptions, desired additional features, and implementation recommendations. Sessions were transcribed verbatim and analyzed as described in the Data Analysis section. Throughout the study, participants received US $20 gift cards at each session. Focus groups were limited to 6 participants to

XSL•FO
RenderX

encourage initiated conversations while minimizing groupthink [34].

Measures

Before phase 1, we conducted a phone-based survey to collect data on demographics, technology experience, social well-being (social isolation, loneliness, and satisfaction with current social interaction levels; Multimedia Appendix 2), and other aspects of health such as functional status, depression, sleep, and physical activity. This manuscript reports only on demographics, technology experience, and social well-being.

Data Analysis

We used a combination of inductive and deductive qualitative approaches across the 3 stages of the UCD process, selected to match the analytic goals of each phase (Table 3).

Table . Focus group data analytic approach.

Stage	Phase	Analytic approach
Stage 1	Phase 1	Inductive thematic analysis; content analysis of pictorial sheets
Stage 2	Phases 3 - 4	Deductive content analysis; matrix display
Stage 3	Phase 5	Deductive content analysis

Phase 1 (Needs Assessment)

Three team members conducted inductive thematic analysis [35] of focus group transcripts to explore participants' social connection needs and identify user preferences for low-fidelity prototype development. Transcripts were imported into Atlas.ti (Lumivero, LLC) software for coding. They independently reviewed transcripts multiple times and met to discuss how participants wanted to use smart speakers for social connectedness and general well-being. These preliminary ideas became codes (eg, how to stay in contact, feeling lonely or isolated, and user needs and preferences), forming the basis for an initial codebook that included code definitions. The team had regular meetings to discuss coding and listed all design ideas mentioned by participants and relevant quotes. We used Miro for affinity mapping to cluster ideas into emergent categories. Rigor strategies included collaborative coding with consensus discussion, peer debriefing with external stakeholders, an audit trail of analytic decisions, and rich description with illustrative quotes. Detailed user challenges with voice user interface on smart speakers and social isolation and loneliness experiences among these participants are published elsewhere [36].

We conducted a content analysis [37,38] of pictorial sheet data using a deductively developed coding framework of 8 categories: daily assistance, health and safety, socializing with others, learning, companionship with a voice agent, entertainment, religion, and miscellaneous. One researcher (NM) assigned category labels to all ideas, which were subsequently reviewed and validated by another researcher (JC). They reconciled any discrepancies through a team discussion, after which the total number of smart speaker–based tasks for each category was calculated.

Phases 3 - 4 (Prototype Evaluation and Refinement)

Two researchers (JC and NM) independently conducted deductive content analysis [37] of phase 3 focus group transcripts using the 4 functional categories established in phase 2 as an organizing framework. A matrix display approach [39] was used to systematically arrange findings by domain, enabling efficient comparison of design preferences and recommendations across focus groups [40]. They (JC and NM) read transcripts multiple times and made notes independently in a pre-agreed template. Then they compared their coding results and reconciled any discrepancies. Based on this discussion, they identified key suggestions for each scenario, presented these to the whole team to get feedback, and then made refinements to the prototype scenarios.

Phase 5 (Validation)

Two team members analyzed transcripts following the phase 4 approach, using deductive content analysis to identify themes relevant to system deployment and implementation planning.

Results

Participant Characteristics

Twenty-nine older adults participated in 1 or more focus group sessions (Table 4). The sample was predominantly African American (n=23, 79%), with 55% (n=16) male. Out of 29 participants, 20 (69%) participants had completed high school education or less. While 21 (72%) participants owned smartphones, only 15 (52%) participants had computer or laptop access, and 12 (41%) participants had never used the internet. Thirteen (45%) participants were smart speaker users at study inception, 11 (85%) of whom reported prior internet use and 2 (15.4%) did not. Among the 11 participants with prior internet use, the majority (n=8, 73%) were daily internet users.

XSL•FO
RenderX

Table . Participant characteristics (N=29).

Characteristics	Participants
Age (years), mean (SD); range	70 (6.8); 60 - 87
60 - 69	12 (41.4)
70 - 79	14 (48.3)
≥80	3 (10.3)
Sex, n (%)	
Female	13 (44.8)
Male	16 (55.2)
Race, n (%)	
African American	23 (79.3)
White/Caucasian	5 (17.2)
Other	1 (3.4)
Hispanic/Latino, no	29 (100)
Education, n (%)	
Less than high school	7 (24.1)
High school diploma/general education development	13 (44.8)
Some college	6 (20.7)
Bachelor's degree	1 (3.4)
Graduate or professional degree	2 (6.8)
Smart speaker ownership, yes, n (%)	13 (44.8)
Smartphone ownership, yes, n (%)	21 (72.4)
Laptop/computer ownership, yes, n (%)	15 (51.7)
Ever used internet, n (%)	
Yes	17 (58.6)
Never	12 (41.4)
Daily internet users, n (%)	13 (44.8)
Social isolation measured by LSNS-6[a], mean (SD); range	15 (SD 6.7); 3 - 24
Social isolation (LSNS-6 <12), yes, n (%)	8 (27.6)
Loneliness measured by UCLA[b] Loneliness Scale Short Form, mean (SD); range	4.6 (SD 1.9); 3 - 9
Loneliness (UCLA-3 ≥ 6), yes, n (%)	9 (31.0)
Participation in any organizations, religious groups, or committees; yes, n (%)	12 (41.4)
Satisfaction with the current social relationships, n (%)	
Very dissatisfied	10 (34.4)
Somewhat dissatisfied	3 (10.3)
Satisfied	16 (55.2)

[a]LSNS-6: Lubben Social Network Scale-6 items.

[b]UCLA: University of California, Los Angeles.

At baseline, participants had a mean Lubben Social Network Scale-6 score of 15 (SD 6.7), with 8 (28%) out of 29 participants meeting the threshold for social isolation (Lubben Social Network Scale-6 score of <12). Mean University of California, Los Angeles (UCLA) Loneliness Scale-Short Form score was 4.6 (SD 1.9), with 9 (31%) out of 29 participants meeting the threshold for loneliness (UCLA Loneliness Scale-3 ≥6). While just over half of the participants (16/29, 55%) reported being satisfied with their current social relationships, 34% (10/29) reported being very dissatisfied and 10% (3/29) somewhat

XSL•FO
RenderX

dissatisfied, and fewer than half (12/29, 41%) reported participating in any organizations, religious groups, or committees.

Stage 1: Understanding Social Connection Needs and Generating Ideas

Social Connection Context and Needs

Focus group discussions revealed multiple factors driving the need for enhanced social connections among residents. Participants described how the pandemic disrupted existing social ties, leaving many feeling isolated. They expressed willingness to help fellow residents experiencing isolation but lacked mechanisms to do so. Many had restricted social networks in general due to limited transportation, mobility constraints, and inadequate access to information about community activities. A critical finding was that participants' social connection needs were inseparable from safety and survival concerns related to living alone. As one participant explained:

> *That would be the best thing we need for Alexa to do. To let everybody—where everybody can say I'm okay, or they check in. Somebody from the office or somebody calls and checks on everybody.* [Group 1, smart speaker user]

Participants emphasized that unexpected health events were a constant worry, creating a need for check-in systems that went beyond traditional socialization support.

Brainstormed Design Ideas

Twenty-four participants generated 153 ideas on pictorial sheets (Multimedia Appendices 3 and 4), with individual contributions ranging from 1 to 16 ideas. The mean number of ideas developed by smart speaker users and nonusers was similar (users: mean 6.6, SD 3.8; nonusers: mean 6.2, SD 5.2). The most frequent categories were Daily Assistance (n=41, 26.8%) and Health and Safety (n=34, 22.2%). Social connectedness generated 35 total ideas across 2 categories: Socializing With Others (n=26, 17.0%) and Companionship With Virtual Agent Alexa (n=9, 5.9%). Additional categories included Learning (n=28, 18.3%), Entertainment (n=7, 4.6%), Religion (n=3, 2.0%), and Miscellaneous (n=5, 3.3%) purposes. Examples of social connection ideas included: "Talk to other residents" and "Call my daughter," "Alexa could call my friends and tell my friends I said hi, and tell them to call me," "Help me check on my health...If I'm feeling a certain way, tell me what you think is wrong," and "What if there was a way to send out a message saying, 'I'm going to go walk. Anyone want to go walk with me?'"

Stage 2: Co-Designing Smart Speaker–Based Applications for Social Connectedness

From User Needs to Design Principles

Through synthesis of focus group transcripts and pictorial sheet data, we developed a Design Strategies Map (Table 5). This analysis revealed overarching design principles: (1) the system should support both social connections and safety while living independently, (2) the system should help build stronger community cohesion, and (3) the design should recognize differing needs across individuals and housing contexts. These principles guided development of 4 functional categories with specific use cases designed to address the social-living context of affordable housing residents aging alone (Multimedia Appendix 5).

XSL•FO
RenderX

Table . Design Strategies Map linking older adults' needs to smart speaker–based application[a].

User need theme	Illustrative quotes	Design response
Safety and Emergency Support While Living Alone	• "That would be the best thing we need for Alexa to do...where everybody can say I'm okay, or they check in. Somebody from the office or somebody calls and checks on everybody." [User, Group 1] • "If we had Alexa, (we could say) 'Alexa, I feel like I'm gonna fall out. Send 911.' A lot of people here have nobody they can talk to." [Nonuser, Group 5]	Checking-in function • Peer-to-peer wellness checks • Management check-in programs • Emergency contact system • Reminders to connect with others
Companionship and Emotional Support	• "It'd be nice if I could have something to talk back to me." [User, Group 2] • "Sometimes I felt down, and I talked to Alexa, and she gave me some good answers... when I'm lonely, that's for sure." (Nonuser, Group 3) • "It could help you—if you're a person with depression and anxiety, you could talk to Alexa to have some company. That would help you have company—have somebody to talk to. Especially if you don't come out and get no fresh air or walk nowhere, where you take walks and see people and speak to them." [Nonuser, Group 4] • "You would need somebody to talk to. Can Alexa read books?" [Nonuser, Group 5]	Social companion function • Conversational AI[b] for daily interaction • Positive affirmations and greetings • Conversational AI reading books aloud to user and encouraging discussions • Recall of past conversations
Community Connection and Social Participation	• "What if there was a way to send out a message saying, 'I'm going to go walk. Anyone want to go walk with me?...You might decide to go, right?" [User, Group 1] • "Alexa could call my friends and tell my friends I said hi, and tell them to call me." [Nonuser, Group 4] • "Keeps me in touch with my daughter." [Nonuser, Group 5] • "That would be helpful to me if it will remind me to call my mom." [Nonuser, Group 5] • "I don't want to just think this [building] is where I'm gonna die at...[We] want something lively." [Nonuser, Group 5]	Community involvement function • Virtual bulletin board for events • Coordination of group activities • Information about local resources • Facilitation of meetups and outings
Health Monitoring and Wellness Support	• "Help me check on my health, such as my temperature, blood pressure....If I'm feeling a certain way, tell me what you think is wrong." [Nonuser, Group 3]	Wellness check function • Integration with wearables • Activity pattern monitoring • Mood and loneliness detection • Prompts for intervention

[a]These ideas were identified through phase 1 focus groups and brainstorming activities. Design responses represent the 4 functional categories developed for prototype scenarios.

[b]AI: artificial intelligence.

Four Functional Categories of Smart Speaker–Based Applications

Category 1: Checking-In (2 Scenarios)

Participants identified checking-in as the most important feature. This included both peer-to-peer and management-to-resident check-ins to address living-alone concerns and emergency support needs, focusing on confirming that a resident is present and safe. As one participant noted: "A lot of people here have nobody they can talk to" [Group 5, nonuser]. The scenarios enabled residents to connect with neighbors through voice commands and established opt-in check-in programs with housing management.

Category 2: Social Companion (2 Scenarios)

Conversational AI could offer companionship, emotional support, and continuity by recalling past interactions. Participants recognized that this could help residents feel less lonely: "Sometimes I felt down, and I talked to Alexa, and she gave me some good answers...when I'm lonely, that's for sure"

XSL•FO
RenderX

[Group 3, nonuser]. Scenarios depicted personalized interactions where Alexa engaged in book discussions and daily conversations with residents.

Category 3: Community Involvement (2 Scenarios)

The technology could facilitate participation by sharing local information (eg, bus routes and community events), enabling virtual gatherings, or arranging meetups. Participants wanted to build "something lively" rather than viewing their housing as purely a place to age in isolation. Scenarios included virtual bulletin boards for community activities and coordinated group trips to essential destinations. Housing providers' role was assumed to coordinate and manage these programs while Alexa's role was access facilitation.

Category 4: Wellness Check (1 Scenario)

Integration with activity tracking devices such as fitness bands capable of monitoring movement, sleep, and inactivity could enable system-initiated health monitoring that detects changes in activity levels as potential signals of physical and emotional health status, prompting the user toward individual action or connection with others (eg, health care providers and significant others). This scenario assumed that participants would have access to such devices, which would communicate activity data to the smart speaker. We developed 7 initial scenarios across these categories based on iterative refinements based on team review (Multimedia Appendix 6), written at accessible reading levels and featuring familiar names to enhance relatability.

Refining Prototypes Through User Feedback

Phase 3 evaluation revealed both enthusiasm and concerns. Participants appreciated applications that could foster connection and help them stay informed about housing events. They strongly preferred technology facilitating human connection over purely automated interactions. The social companion features were seen as useful for those experiencing loneliness, although some worried about replacing rather than supplementing human contact.

Privacy and obtrusiveness emerged as primary concerns. Participants worried about other residents or management checking-in at any time, unintentional sharing of location and activity information, and technology exerting excessive control (described as "too Orwellian" by one group). Responses to the wellness check scenario indicated that participants were receptive to the use of wearable-based data for monitoring physical inactivity and detecting potential falls. However, they expressed a clear preference for systems that prompt rather than diagnose, flagging activity changes and facilitating human connection, rather than inferring emotional states from sensor data.

Participants provided concrete improvement suggestions for each scenario (Multimedia Appendix 7). In summary, these include (1) user control—add "do-not-disturb" features and allow selection of check-in frequencies, (2) privacy protection—create close contact lists rather than building-wide access and ensure that health information is not shared during check-ins, (3) natural interaction—make voice interfaces sound conversational rather than mechanical, (4) minimal management oversight—limit housing management role to essential check-ins only, and (5) personalization options—allow users to set speaking times, choose conversation perspectives, and toggle features on/off.

The depth and quality of participant contributions did not differ by prior smart speaker experience. Analysis of focus group interviews across all 7 design scenarios revealed comparable levels of engagement between smart speaker users and nonusers. Per focus group, both groups generated similar numbers of positive themes (users: 11; nonusers: 11) and design recommendations (users: 14; nonusers: 11.3). The major design themes, including privacy and user control, perceived usefulness for socially isolated residents, trusted contact network preferences, and voice-based security, emerged consistently across both groups. While users occasionally drew on their lived device experience, nonusers contributed equally substantive and actionable recommendations, including novel suggestions such as video interfaces for hearing-impaired residents, information integration, and peer activity coordination via Alexa.

Based on this feedback, we revised all 7 scenarios (Textbox 1). Key changes included adding opt-in or opt-out capabilities, do-not-disturb functions, contact list controls, and explicit privacy protections. For example, in the management check-in scenario, we specified that Alexa would confirm resident status without sharing additional details the resident mentioned (eg, having a cold). Figure 1 shows a graphic representation of a scenario in the checking-in category.

XSL•FO
RenderX

Textbox 1. Refined scenarios of use for the Voice2Connect system prototypes. Italicized text shows the changes made based on the feedback from phase 3 focus groups.

Category 1: Checking-in

Scenario 1: Using Alexa to talk to your neighbor or friend

Tom, 78-year-old man, uses the smart speaker to check in on others in the building. Tom has not heard from or seen his neighbor Jeanne in 2 weeks, but he knows that Jeanne also has an Alexa Dot, and *he and Jeanne have added one another to their contact lists*. So he speaks to Alexa to see whether Jeanne is doing well: "Alexa, connect me to Jeanne" and then he is connected to Jeanne. Tom asks Jeanne, "Hey, how are you doing? I just want to make sure you are well." Tom and Jeanne chat briefly through the speaker to share what is going on with them. Two days later, Tom is watching TV and *he doesn't want to get bothered, so he turns on the Do Not Disturb Feature. When Jeanne tries to contact him, there is an auto-message from Alexa:* "Tom has turned on 'Do Not Disturb.'"

Scenario 2: Check-ins from the management

The building management started a resident check-in program through Alexa. Cathy opted into the wellness check-in service, and she chose management daily check-ins. Tom also opted into the check-in service and he chose weekly check-ins. The office manager put an announcement on the Alexa of check-in program participants in the building: "Hi Cathy, how are you? This is management. Please tell Alexa if you are ok. If we do not hear from you in 24 hours, we will come to your apartment to check on you." Cathy said, "Alexa, tell management I am ok. I have had a bad cold recently." Alexa informs the management that Cathy is okay and *doesn't say anything about Cathy's bad cold*.

Category 2: Social companion

Scenario 1: Being a friend with Alexa

George is a widowed 85-year-old man and lives alone. *He wants to have an individualized interaction with Alexa, so he chooses the highest level of personalization*. George's children come and visit him on the weekends, but during the week he has little social interaction and often feels lonely. Alexa detects when George gets out of bed and greets him by saying "Good Morning, George, how did you sleep?" George replies saying, "I slept well. Alexa, can you read me the next chapter of my Steven King book at 10 am?" At 10 AM, Alexa says, "Here is where we left off reading yesterday" and continues to read the book aloud to George. After some time, George says, "Alexa, stop reading. This book is getting very good." Alexa replies, "It is. What do you think is going to happen next?" George replies, and Alexa says, "That is possible, we should read again tomorrow."

Scenario 2: Conversation between you and Alexa

John has arthritis in his knees and hips, making walking and standing for long amounts of time painful for him, so he likes to stay at home most of the time. Getting out in the community and connecting with other people has been difficult because of John's limited mobility. John likes to talk to Alexa. In the mornings he will say "Good morning, Alexa," and Alexa will reply, "Good morning, John. Today is going to be a blessed day." John is able to reply, "Every day is a blessed day."

John likes to talk out loud about what is happening in his favorite television show and movies. Alexa remembers what John says is happening in the movie and is able to talk with him about it. "John, that is a popular movie. Online it has hundreds of 5-star reviews. One person left a review and said 'this is one of my favorite movies. I was not a big fan of the ending, but overall the plot was good,'" "Alexa, I agree, but it was still a good movie."

Category 3: Social participation

Scenario 1: Getting involved in the community

Alice, a 75-year-old woman, recently moved into a new apartment building to be closer to her children. *Alice wants to make new friends in the building, so she has opted into the community events bulletin board. When Alice comes home from the grocery store, she sees a light on her Alexa, telling her there is an announcement. Alice asks Alexa,* "Alexa, what is the announcement?" *Alexa says,* "There is a Spanish class today at 3 pm. Would you like to hear other activities happening today?*" Alice says,* "yes, please," *and Alexa reads Alice the list of community activities. Alice says,* "Alexa I think I want to try that Spanish class. Remind me when it is time." *Later at 2:50 PM, Alexa says,* "Don't forget, Spanish class starts in 10 minutes." On Sunday, Alexa says to Alice "Good morning Alice, there is the weekly church service going on in the community room downstairs. Everyone in the building is invited."

Scenario 2: Meetups

David is a new resident of affordable senior housing and recently got an Alexa. After settling into his new apartment, *he opted in to a virtual bulletin board to see what events are happening in the building. David asks Alexa,* "Is there any activity going on today?*"* Alexa says, "John and Martha will walk to a park at 3 pm today for exercise. Do you want to join?" David says, "Yes, put me on the list." *Alexa reminds David 30 minutes before,* "David, you will need to meet them at the entrance at 3 pm."

A couple of days later, David made a list of grocery items he needed. Being new to Richmond, David did not feel comfortable going alone. He said, "Alexa, can you ask if anyone needs to go to Kroger today? We can all take the bus together. *Also, add this to the virtual bulletin board so other people know about this*." Alexa replies, "Martha will go to Kroger with you today at 2 pm. You can take the #3 bus at 1:30. It will pick you up in front of the building."

Category 4: Wellness check and mood detection

Scenario 1: Mood detection

XSL•FO
RenderX

Paula, a 69-year-old woman, has been living in her apartment building for 4 years. She would describe herself as fairly active and social with a few good friends living in the same apartment building. Recently, her only child and grandchildren moved further away from Paula, making it more difficult to see each other. *Paula is worried about becoming lonely, so she opts into the loneliness detection feature*. A few weeks later, Paula is seeing her friends less frequently, sleeping in later, taking naps, and is spending more time watching TV. Alexa says "Paula, the data from the Fitbit tell that you have not been as active as you normally are. How are you doing? Would you like me to connect you with one of your friends or your doctor?" *Paula says*, "Well, Alexa, I don't think I need the doctor, but I haven't been very active recently and am feeling a little lonely." Three days later, Paula still has not left her apartment. *Alexa says*, "Paula, you still seem to be less active than usual. Would you like me to connect to your friend or read the list of activities on the bulletin board?" *Paula agrees, and Alexa connects Paula to her friend to make plans for a walk*.

Figure 1. A graphic representation of a Voice2Connect prototype in the checking-in category. This figure shows verbal interactions between user and voice agent.



Stage 3: Validating Final Prototypes and Planning Implementation

Overview

Participants expressed varied but generally positive opinions about refined scenarios. Overall, they appreciated the vision of using smart speakers to foster a more connected and livable community, although enthusiasm varied by scenario. Features requiring movement or mood detection to identify emotional health risks were viewed most skeptically, with concerns about excessive technological control over daily life.

XSL•FO
RenderX

Crosscutting Themes From Validation Sessions

Personalization as Essential

Participants emphasized the importance of controlling features rather than having a one-size-fits-all system. Suggestions included setting specific times for Alexa to speak, choosing conversation styles or perspectives (eg, "Christian perspective"), and selecting which physiological metrics to monitor.

Safety Remains Paramount

Across nearly all scenarios, participants referenced potential health emergencies, particularly for residents in poorer health or facing adverse outcomes. The checking-in and wellness check functions were especially valued. As one participant noted, "That's better than somebody knocking on your door to see if you're okay" [Group 4].

Preventing Replacement of Human Connection

Participants stressed the need for safeguards against AI technology replacing human interaction or enabling abusive behaviors. They wanted technology that facilitated human connection rather than substituted for it.

Discussion

Principal Findings

This study used an iterative UCD process to engage older adult residents in affordable housing in designing prototype scenarios for smart speaker–based applications to promote social connectedness. Throughout 3 stages of co-design, we developed 7 scenarios across four functional categories: (1) checking-in for safety verification, (2) conversational AI as social companions, (3) community involvement, and (4) wellness monitoring. These scenarios provide a foundation for developing a technology platform tailored to address the specific social isolation and loneliness experienced by underserved older adult populations, who often face distinct challenges in locating and accessing the social and financial resources necessary for aging in place. Baseline social well-being data underscore this need. While more than half reported satisfaction with their current social relationships, a meaningful subgroup was experiencing clinically significant loneliness and social isolation, lending ecological validity to their strong interest in features supporting social connectedness in their living communities.

The findings of this study can be interpreted through the lens of Self-Determination Theory [41], which posits that human motivation and well-being are grounded in 3 psychological needs: autonomy, competence, and relatedness. Self-Determination Theory maps meaningfully onto the design preferences and concerns expressed by participants across all phases. The demand for customization, opt-in controls, do-not-disturb features, and close contact lists reflects a fundamental need for *autonomy*, the desire to maintain personal agency over one's social environment and information sharing. Participants' requests for interfaces that are simple, conversational, and responsive to their voice reflect the need for *competence*, the ability to engage effectively with technology without feeling overwhelmed or excluded. Finally, participants' desire for building a livable community for themselves and peers and for preserving human relationships in the loop reflects the need for *relatedness*, belongingness, and meaningful connection with others. Designing smart speaker–based applications that satisfy these needs may be essential not only for usability but for sustaining intrinsic motivation for continued technology engagement and well-being outcomes in this population.

Social Connectedness Inseparable From Safety and Survival

Throughout the design process, it became clear that the study's focus on social connectedness strongly resonated with participants due in part to the isolation they experienced during the COVID-19 pandemic, a period during which residents faced significant restrictions on in-person social interaction. The strong participant interest in peer check-ins and community engagement features may partly reflect this context, and it is plausible that the relative prioritization of these features could differ in a postpandemic sample. Importantly, however, a critical insight emerged: for affordable housing residents living alone, social connection needs cannot be separated from safety and survival concerns. Participants emphasized that any technology system must address both relationship-building and emergency support, reflecting the reality that all residents live alone and some face immediate safety vulnerabilities, including declining health and fall risk, that may arise at times when human assistance is not immediately available.

This dual emphasis shaped our design approach fundamentally. Rather than creating purely social applications, we prioritized checking-in functionality that serves both social and safety purposes. This finding aligns with research suggesting that older adults' technology adoption is strongly influenced by whether innovations address their most pressing daily concerns [42]. This study shows that for affordable housing residents with limited support networks, safety vulnerabilities are not abstract future concerns but immediate realities shaping their daily experiences of living alone [43].

Building Social Infrastructure in Affordable Housing

Our findings indicate that residents were genuinely interested in making a livable community within their housing by building social infrastructure [44], not just to address social needs but also to promote safe, independent living. Participants suggested various ideas, including virtually connecting with individuals in need, conducting remote safety check-ins, advertising community events, and facilitating coordinated trips to essential destinations (eg, grocery stores). Participants were interested in expanding and enhancing social networks within and beyond their housing community. Research suggests that digital tools connecting older adults to existing social networks may not significantly increase social activity for individuals lacking networks and at risk of social isolation [45]. However, tools designed to facilitate new social connections or broaden access to community resources can play a critical role in addressing these gaps [46]. Our scenarios address this through community-building features (eg, virtual bulletin boards, coordinated meetups, and shared activity planning) that help residents discover and connect with neighbors who share similar needs and interests.

XSL•FO
RenderX

The emphasis on community involvement also reflects participants' recognition that their housing buildings could become more than just places to live in isolation. They wanted technology that would help them build "something lively," a community with regular interaction, mutual support, and shared activities. This vision extends beyond typical conceptualizations of smart speakers as individual assistive devices and positions them as one component of a broader, community support ecosystem that can facilitate communication about opportunities and access to them, while the core work of building and sustaining community remains a fundamentally human endeavor. Importantly, technology alone is insufficient to establish such infrastructure. The foundation must be built through human-centered programming, dedicated staffing, and sustained relational investment. Smart speaker–based applications, therefore, should be understood not as substitutes for this work but as complementary tools that facilitate awareness of and access to existing and emerging community resources, functioning as bridges to human connection.

Engaging Underserved Communities in Technology Design

To our knowledge, this is one of few studies using UCD with older adults of lower socioeconomic status. Such populations are often excluded from design and testing processes for new technologies conceptualized to improve quality of life and health [47], probably due to concerns about their ability to engage meaningfully and provide valuable input. Limited access to ICT and low digital literacy also may have contributed to the underrepresentation of these communities in technology research and design initiatives. The limited participation of marginalized communities perpetuates the exclusion of their perspectives [48] and contributes to disparities in digital health access and outcomes. Given that widespread digital technology access has created new opportunities for self-management, quality of life, and social well-being in older adults [49-51], recognizing the importance of engaging community members of underrepresented populations in a technology development and testing process is essential.

Our study demonstrates that older adults in underserved communities, including those with limited prior technology experience, can actively participate in and meaningfully contribute to technology design processes. Our participants across age groups, technology experience levels, and housing communities demonstrated enthusiasm and active engagement and expressed a sense of ownership in the design sessions, contributing valuable insights that shaped prototype development. The 153 ideas generated during brainstorming activities, the nuanced feedback on privacy and control features, and the concrete implementation recommendations all demonstrate that concerns about this population's ability to engage in design work are unfounded.

Methodological Contributions: Adapting UCD for Underserved Communities

Our research has methodological implications for UCD research with and for underserved communities. The 3-stage iterative approach we used successfully balanced rigorous design principles with the unique contexts and abilities of participants. Our inclusive methods included rapid brainstorming sessions, facilitated mini-group discussions to ensure that everyone had a voice and that discussions remained focused and productive, audiovisual materials to enhance nonusers' understanding of the target technology and stimulate idea generation, scenario creation and role-play demonstrations to reduce uncertainty about the technology and facilitate discussions around privacy and security, and iterative design revision informed by continuous feedback.

These methods were thoughtfully chosen based on the literature [31-33] and the team's extensive experience working with affordable housing communities. Beyond specific UCD techniques, the team prioritized creating a comfortable, nonjudgmental environment where participants, especially those not familiar with smart speakers, could freely propose ideas, suggest changes to scenarios, and identify concerning features. We found that using pictorial sheets allowed participants with varying literacy levels to express ideas, while mini-group formats prevented more vocal participants from dominating discussions.

This study also demonstrates the need for flexibility in balancing rigorous design principles with the unique contexts and abilities of participants, as previous research demonstrated [32]. For example, during phase 1, we opted not to conduct mini-group sessions with focus group 5 (nonusers) after observing their difficulty engaging in technology-related discussions. Instead, we adapted our approach to facilitate a large group discussion focused on preferred functionalities for social connectedness, emphasizing their role as co-designers. This adaptive method adds important considerations specific to a UCD process with underserved populations, suggesting that methodological fidelity should sometimes yield to participant comfort and authentic engagement.

The established trust relationships our research team had developed through prior community-engaged work [52] proved essential to recruitment and retention. This suggests that successful UCD with marginalized communities requires sustained institutional commitment beyond individual research projects, and researchers cannot parachute in for design sessions without prior relationship-building and expect meaningful engagement.

Privacy, Control, and Trust as Essential Design Requirements

Participants' privacy concerns emerged consistently across all 3 stages and shaped significant design revisions. Their worries about unintentional information sharing, management surveillance, and loss of autonomy led to specific requirements: opt-in or opt-out controls, do-not-disturb features, close contact lists rather than building-wide access, and explicit privacy protections ensuring that sensitive information would not be shared with management or other residents. These concerns reflect broader issues of power, trust, and control that are particularly salient for minority, low-income older adults who may have experienced systemic marginalization and surveillance [53]. The description of mood detection features as "too Orwellian" signals deep discomfort with technologies that could exert control over their lives without consent. Our design

XSL•FO
RenderX

iterations directly addressed these concerns by centering user control and transparency.

Importantly, participants distinguished between features they found helpful (checking-in on neighbors and being reminded to connect with others) and those they found intrusive (automatic mood detection and unrestricted management access). This nuanced feedback demonstrates why participatory design is essential because researchers and technologists might assume that more automated, comprehensive monitoring would be beneficial, while users clearly articulated the need for boundaries and personal agency.

Balancing Human Connection and Technological Companionship

Participants expressed interest in voice agent companionship while simultaneously voicing concerns about replacing human interaction. This tension reflects broader societal questions about the role of AI in addressing loneliness and social isolation. Participants recognized that for those experiencing severe isolation, having "somebody to talk to" through Alexa might be "better than nothing," while also insisting on the primacy of human relationships.

Our scenarios attempted to navigate this tension by positioning voice agents as facilitators of human connection rather than replacements for it. For example, the social companion scenarios included Alexa suggesting the user connect with friends for walks rather than only offering its own companionship. The wellness check scenario prompted connection to friends or health care providers rather than only providing automated support. This design approach aligns with emerging ethical frameworks for AI companionship that emphasize augmentation rather than substitution [18]. However, participants' concerns suggest that implementation will require careful monitoring to prevent vulnerable, isolated individuals from becoming overly reliant on AI interaction at the expense of human relationships. Future development should include safeguards that encourage and facilitate human connection rather than allowing AI to become the primary or sole social contact.

Implementation Considerations and Future Directions

The varying levels of enthusiasm across scenarios suggest that a modular, customizable system would be most appropriate. Rather than implementing all 7 scenarios uniformly, housing communities should be able to select and configure features based on resident preferences, building infrastructure, and management capacity. This aligns with participants' emphasis on personalization and control. Moreover, given participants' limited experience with ICT and struggles with voice user interface interactions shared in phase 1 focus groups [11], resident training sessions and ongoing technical support are required for successful implementation. These supports are inherently service-oriented, requiring dedicated staff, sustained resourcing, and the cultivation of trust with residents, and should be considered alongside, rather than assumed to be fully addressed by the technology itself.

In the future, research should examine how smart speaker–based social connection tools interact with other community resources and interventions. Our participants identified transportation barriers, lack of information about community activities, and limited social networks as interconnected challenges. Technology solutions are most likely to succeed when integrated into broader community-based approaches that address structural barriers to social connection.

Learning emerged as a prominent brainstorming category in phase 1, comparable in volume to Socializing With Others. Participants—all nonusers—generated ideas spanning news and current events, language acquisition, practical skills, and digital literacy. Although this domain falls outside the social connectedness focus of this study, these observations point to an interest in intellectual engagement within this population, consistent with a growing literature on lifelong learning in older adults [54,55]. Future work might more explicitly examine learning-oriented smart speaker–based applications as a complementary pathway for supporting cognitive engagement and quality of life.

Limitations

There are several limitations of this study. First, due to the COVID-19 pandemic, we were unfortunately unable to follow up with some participants throughout the process, resulting in their loss from the study. The pandemic also necessitated off-site meetings with transportation provision, which may have excluded some potential participants with mobility limitations or discomfort with van travel. Second, established trust relationships that facilitated recruitment may have introduced response bias, as participants with prior positive experiences with the research team may have been more inclined to provide favorable feedback. Given that real-world implementation will inevitably encounter residents who are skeptical, mistrustful of institutional technology, or uninterested in smart speaker–based applications, purposive recruitment of such individuals would elicit insights that a self-selected, engaged sample cannot provide. Third, incorporating diverse and sometimes contrasting ideas into the design was challenging. While we used multiple data synthesis methods, a more systematic approach to identifying and prioritizing user demands (eg, conjoint analysis or formal prioritization exercises) would have been beneficial for making design trade-off decisions explicit. Fourth, hearing-related concerns, such as hearing loss that hinders meaningful engagement with smart speakers, did not emerge spontaneously in focus group discussions. This absence may reflect sample characteristics or participants' limited awareness of the relationship between hearing ability and voice interface usability. Although smart speaker designs incorporate features such as adjustable volume and visual display companions that may support more equitable access, their adequacy in practice remains unclear. Future work should more explicitly examine how hearing loss shapes engagement with smart speakers. Fifth, while we adapted facilitation approaches to encourage broad participation, differential comfort with technology discussions likely shaped whose voices were most prominent in certain design decisions. Furthermore, differences in age, education, technology access and usage, and prior smart speaker experience among participants may have influenced the outcomes at each phase of the study. Finally, this study focused on early-stage prototyping and scenario development. The scenarios represent

XSL•FO
RenderX

participant aspirations and preferences but have not been validated through implementation.

Conclusions

Our study identified the needs for meaningful connections and safe, independent living among older adults living alone in affordable housing and showed the potential of smart speakers and their voice agents to address social isolation and reduce loneliness. Through a 3-stage iterative UCD process, we successfully engaged residents, including those with limited technology experience, in co-designing 7 prototype scenarios across 4 functional domains. The UCD methods used in this study effectively incorporated end user feedback to develop a technology platform for reducing social isolation and loneliness in this population [42]. Findings emphasize the importance of accessible, thoughtfully designed technology tools for social connectedness and more broadly for health and well-being tailored to older adults with limited resources and digital literacy. Further research is needed to fully develop these prototypes into functional applications, pilot-test them in real-world settings, and evaluate their effectiveness in addressing the unique social health needs of underserved older adults. Future evaluation should also extend beyond self-selected participants to deliberately engage skeptical residents, individuals with institutional mistrust, and those for whom this technology is not an intuitive fit. Rather than assuming universal applicability, research should iteratively test and refine these prototypes across diverse residential contexts, trust environments, and user dispositions to generate a more grounded and transparent evidence base to implementation in affordable housing communities.

Acknowledgments

The authors thank all the participants in the study who made outdoor trips during the pandemic to participate in focus group sessions. The authors express their gratitude to Greg Ford, Director of Program Development and Implementation at Beacon Communities; Normal Dorsett, Senior Community Coordinator at a Winn Residential Managed Community; and Kate Noon, Director of Senior Services at Better Housing Coalition at the time of participant recruitment and data collection. The authors are also grateful to Katherine Falls, Director of Clinical Operations at Mobile Health and Wellness Program at Virginia Commonwealth University. Also, this work is dedicated to the memory of Dr Pamela Parsons, Associate Dean for Practice and Community Engagement at Virginia Commonwealth University School of Nursing, at the time of study design and execution. Without their support, this research would not have been possible. During the preparation of this manuscript, generative artificial intelligence tools were used solely as an editorial assistance tool to improve clarity, conciseness, organization, and grammar of draft responses. The authors have thoroughly reviewed and edited the content and remain fully responsible for the accuracy, originality, and integrity of the final manuscript.

Funding

This work was supported by the National Institutes of Health (R03AG069816 to JC).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

JC conceptualized the project, secured funding, administered and coordinated the project, secured recruitment sites and resources, led data collection and analysis, and led the process of manuscript writing and revision. NM and REW assisted the team in recruiting participants, collecting data, and analyzing data. TG participated in data analysis. GD assisted JC in study conceptualization and data interpretation. All authors participated in manuscript writing, revised the manuscript, and reviewed the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Pictorial sheet for the 6-8-5 exercise for brainstorming in phase 1.
[DOCX File, 223 KB - aging_v9i1e90053_app1.docx]

Multimedia Appendix 2
Social well-being questionnaire—social isolation, loneliness, and satisfaction.
[DOCX File, 17 KB - aging_v9i1e90053_app2.docx]

Multimedia Appendix 3

XSL•FO
RenderX

Categories of ideas written or drawn on pictorial sheets.
[DOCX File, 18 KB - aging_v9i1e90053_app3.docx]

Multimedia Appendix 4
Example ideas shared by a smart speaker user and nonuser.
[DOCX File, 1227 KB - aging_v9i1e90053_app4.docx]

Multimedia Appendix 5
Voice2Connect functional categories, description, and potential technology functions.
[DOCX File, 17 KB - aging_v9i1e90053_app5.docx]

Multimedia Appendix 6
Initial use case scenarios.
[DOCX File, 20 KB - aging_v9i1e90053_app6.docx]

Multimedia Appendix 7
A matrix of content analysis findings in phase 3.
[DOCX File, 20 KB - aging_v9i1e90053_app7.docx]

References

1. National Academies of Sciences, Engineering, and Medicine. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System: National Academies Press (US); 2020. URL: http://www.ncbi.nlm.nih.gov/books/NBK557974 [accessed 2023-01-26]
2. Donovan NJ, Wu Q, Rentz DM, Sperling RA, Marshall GA, Glymour MM. Loneliness, depression and cognitive function in older U.S. adults. Int J Geriatr Psychiatry 2017 May;32(5):564-573. [doi: 10.1002/gps.4495] [Medline: 27162047]
3. Perissinotto CM, Stijacic Cenzer I, Covinsky KE. Loneliness in older persons: a predictor of functional decline and death. Arch Intern Med 2012 Jul 23;172(14):1078-1083. [doi: 10.1001/archinternmed.2012.1993] [Medline: 22710744]
4. Musich S, Wang SS, Hawkins K, Yeh CS. The impact of loneliness on quality of life and patient satisfaction among older, sicker adults. Gerontol Geriatr Med 2015;1:2333721415582119. [doi: 10.1177/2333721415582119] [Medline: 28138454]
5. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
6. Gierveld JDJ. A review of loneliness: concept and definitions, determinants and consequences. Rev Clin Gerontol 1998 Feb;8(1):73-80. [doi: 10.1017/S0959259898008090]
7. Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. Heart 2016 Jul 1;102(13):1009-1016. [doi: 10.1136/heartjnl-2015-308790] [Medline: 27091846]
8. Hu J, Fitzgerald SM, Owen AJ, et al. Social isolation, social support, loneliness and cardiovascular disease risk factors: a cross-sectional study among older adults. Int J Geriatr Psychiatry 2021 Nov;36(11):1795-1809. [doi: 10.1002/gps.5601] [Medline: 34231940]
9. Sargent L, Zanjani F, Winship J, et al. Socioeconomic effects on psychosocial factors among low-income older adults. Gerontol Geriatr Med 2022;8:23337214221084866. [doi: 10.1177/23337214221084866] [Medline: 35299880]
10. Parsons PL, Mezuk B, Ratliff S, Lapane KL. Subsidized housing not subsidized health: health status and fatigue among elders in public housing and other community settings. Ethn Dis 2011;21(1):85-90. [Medline: 21462736]
11. Chung J, Gendron T, Winship J, et al. Smart speaker and ICT use in relationship with social connectedness during the pandemic: loneliness and social isolation found in older adults in low-income housing. Gerontologist 2024 May 1;64(5):gnad145. [doi: 10.1093/geront/gnad145] [Medline: 37880825]
12. Kearns A, Whitley E, Mason P, Bond L. 'Living the High Life'? Residential, social and psychosocial outcomes for high-rise occupants in a deprived context. Housing Stud 2012 Jan;27(1):97-126. [doi: 10.1080/02673037.2012.632080]
13. Krieger J, Higgins DL. Housing and health: time again for public health action. Am J Public Health 2002 May;92(5):758-768. [doi: 10.2105/ajph.92.5.758] [Medline: 11988443]
14. Carbone JT, Clift J, Wyllie T, Smyth A. Housing unit type and perceived social isolation among senior housing community residents. Gerontologist 2022 Jul 15;62(6):889-899. [doi: 10.1093/geront/gnab184]
15. Fakoya OA, McCorry NK, Donnelly M. Loneliness and social isolation interventions for older adults: a scoping review of reviews. BMC Public Health 2020 Feb 14;20(1):129. [doi: 10.1186/s12889-020-8251-6] [Medline: 32054474]
16. Jones VK, Hanus M, Yan C, Shade MY, Blaskewicz Boron J, Maschieri Bicudo R. Reducing loneliness among aging adults: the roles of personal voice assistants and anthropomorphic interactions. Front Public Health 2021;9:750736. [doi: 10.3389/fpubh.2021.750736] [Medline: 34957013]

XSL•FO
RenderX

17. O'Brien K, Light SW, Bradley S, Lindquist L. Optimizing voice-controlled intelligent personal assistants for use by home-bound older adults. J Am Geriatr Soc 2022 May;70(5):1504-1509. [doi: 10.1111/jgs.17625] [Medline: 35029296]
18. Park S, Kim B. The impact of everyday AI-based smart speaker use on the well-being of older adults living alone. Technol Soc 2022 Nov;71:102133. [doi: 10.1016/j.techsoc.2022.102133]
19. Corbett CF, Wright PJ, Jones K, Parmer M. Voice-activated virtual home assistant use and social isolation and loneliness among older adults: mini review. Front Public Health 2021;9:742012. [doi: 10.3389/fpubh.2021.742012] [Medline: 34708017]
20. Perrin SA. Home broadband adoption, computer ownership vary by race, ethnicity in the US. Pew Research Center. 2021 Jul 16. URL: https://www.pewresearch.org/short-reads/2021/07/16/home-broadband-adoption-computer-ownership-vary-by-race-ethnicity-in-the-u-s/ [accessed 2024-05-02]
21. Petretto DR, Carrogu GP, Gaviano L, et al. Telemedicine, e-Health, and digital health equity: a scoping review. Clin Pract Epidemiol Ment Health 2024;20(1):e17450179279732. [doi: 10.2174/0117450179279732231211110248] [Medline: 38660571]
22. Howard E, Gillispie-Bell V, Olet S, Glenn B, Ammar N, Price-Haywood EG. Evaluating racial disparities in implementation and monitoring of a remote blood pressure program in a pregnant population-a retrospective cohort study. Ochsner J 2024;24(1):22-30. [doi: 10.31486/toj.23.0111] [Medline: 38510223]
23. Jensen JD, King AJ, Davis LA, Guntzviller LM. Utilization of internet technology by low-income adults: the role of health literacy, health numeracy, and computer assistance. J Aging Health 2010 Sep;22(6):804-826. [doi: 10.1177/0898264310366161] [Medline: 20495159]
24. Schmeida M, McNeal RS. The telehealth divide: disparities in searching public health information online. J Health Care Poor Underserved 2007 Aug;18(3):637-647. [doi: 10.1353/hpu.2007.0068] [Medline: 17675719]
25. Filippi MK, McCloskey C, Williams C, et al. Perceptions, barriers, and suggestions for creation of a tobacco and health website among American Indian/Alaska Native college students. J Community Health 2013 Jun;38(3):486-491. [doi: 10.1007/s10900-012-9634-0] [Medline: 23149568]
26. McCloud RF, Okechukwu CA, Sorensen G, Viswanath K. Beyond access: barriers to internet health information seeking among the urban poor. J Am Med Inform Assoc 2016 Nov;23(6):1053-1059. [doi: 10.1093/jamia/ocv204] [Medline: 27206459]
27. Cashen MS, Dykes P, Gerber B. eHealth technology and Internet resources: barriers for vulnerable populations. J Cardiovasc Nurs 2004;19(3):209-214. [doi: 10.1097/00005082-200405000-00010] [Medline: 15191264]
28. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013 May 2;15(5):e93. [doi: 10.2196/jmir.2645] [Medline: 23639979]
29. Gulliksen J, Göransson B, Boivie I, Blomkvist S, Persson J, Cajander Å. Key principles for user-centred systems design. Behav Inf Technol 2003 Nov;22(6):397-409. [doi: 10.1080/01449290310001624329]
30. Demiris G, Chung J, Turner AM. Using informatics to engage vulnerable populations in research: lessons learned from the COVID-19 pandemic. Stud Health Technol Inform 2022 Jun 6;290:858-861. [doi: 10.3233/SHTI220201] [Medline: 35673140]
31. Gray D, Brown S, Macanufo J. Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers: O'Reilly and Associates; 2010.
32. Hoffman AS, Bateman DR, Ganoe C, et al. Development and field testing of a long-term care decision aid website for older adults: engaging patients and caregivers in user-centered design. Gerontologist 2020 Jul 15;60(5):935-946. [doi: 10.1093/geront/gnz141]
33. Turner AM, Reeder B, Ramey J. Scenarios, personas and user stories: user-centered evidence-based design representations of communicable disease investigations. J Biomed Inform 2013 Aug;46(4):575-584. [doi: 10.1016/j.jbi.2013.04.006] [Medline: 23618996]
34. Halcomb EJ, Gholizadeh L, DiGiacomo M, Phillips J, Davidson PM. Literature review: considerations in undertaking focus group research with culturally and linguistically diverse groups. J Clin Nurs 2007 Jun;16(6):1000-1011. [doi: 10.1111/j.1365-2702.2006.01760.x] [Medline: 17518876]
35. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development: SAGE Publications; 1998.
36. Chung J, Viceconte M, Syros R, Wood R, Mansion N, Jang H. Low-income older adults' user challenges and preferences related to voice interface design on smart speakers. J Gerontol Nurs 2025 Jul;51(7):7-14. [doi: 10.3928/00989134-20250603-03] [Medline: 40626782]
37. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
38. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
39. Miles MB, Huberman AM, Saldana J. Qualitative Data Analysis: A Methods Sourcebook, 4th edition: SAGE Publications; 2019.

XSL•FO
RenderX

40. Caunca MR, Simonetto M, Hartley G, Wright CB, Czaja SJ. Design and usability testing of the stroke caregiver support system: a mobile-friendly website to reduce stroke caregiver burden. Rehabil Nurs 2020;45(3):166-177. [doi: 10.1097/RNJ.0000000000000196] [Medline: 30418319]
41. Ryan RM, Deci EL. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness: Guilford Press; 2017. [doi: 10.1521/978.14625/28806]
42. Essén A, Östlund B. Laggards as innovators? Old users as designers of new services & service systems. Int J Des 2011;5(3) [FREE Full text]
43. Melchiorre MG, D'Amen B, Quattrini S, Lamura G, Socci M. Health emergencies, falls, and use of communication technologies by older people with functional and social frailty: ageing in place in deprived areas of Italy. Int J Environ Res Public Health 2022 Nov 10;19(22):14775. [doi: 10.3390/ijerph192214775] [Medline: 36429499]
44. Heatwole Shank KS, Cutchin MP. Processes of developing "community livability" in older age. J Aging Stud 2016 Dec;39:66-72. [doi: 10.1016/j.jaging.2016.11.001] [Medline: 27912856]
45. Ngiam NHW, Yee WQ, Teo N, et al. Building digital literacy in older adults of low socioeconomic status in Singapore (Project Wire Up): nonrandomized controlled trial. J Med Internet Res 2022 Dec 2;24(12):e40341. [doi: 10.2196/40341] [Medline: 36459398]
46. Chen YRR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18. [doi: 10.2196/jmir.4596] [Medline: 26822073]
47. Wilson S, Tolley C, Mc Ardle R, et al. Recommendations to advance digital health equity: a systematic review of qualitative studies. NPJ Digit Med 2024 Jun 29;7(1):173. [doi: 10.1038/s41746-024-01177-7] [Medline: 38951666]
48. Koehle H, Kronk C, Lee YJ. Digital health equity: addressing power, usability, and trust to strengthen health systems. Yearb Med Inform 2022 Aug;31(1):20-32. [doi: 10.1055/s-0042-1742512] [Medline: 36463865]
49. Stellefson M, Chaney B, Barry AE, et al. Web 2.0 chronic disease self-management for older adults: a systematic review. J Med Internet Res 2013 Feb 14;15(2):e35. [doi: 10.2196/jmir.2439] [Medline: 23410671]
50. Bendixen RM, Horn K, Levy C. Using telerehabilitation to support elders with chronic illness in their homes. Top Geriatr Rehabil 2007 Jan;23(1):47-51. [doi: 10.1097/00013614-200701000-00007]
51. Wijsman CA, Westendorp RG, Verhagen EA, et al. Effects of a web-based intervention on physical activity and metabolism in older adults: randomized controlled trial. J Med Internet Res 2013 Nov 6;15(11):e233. [doi: 10.2196/jmir.2843] [Medline: 24195965]
52. Diallo AF, Mackiewicz M, Sargent L, et al. Cultivating relationships as a community-based recruitment strategy in transdisciplinary aging research: lessons from an academic-community partnership. Fam Community Health 2024;47(1):32-40. [doi: 10.1097/FCH.0000000000000383] [Medline: 37831622]
53. Liang CA, Munson SA, Kientz JA. Embracing four tensions in human-computer interaction research with marginalized people. ACM Trans Comput Hum Interact 2021 Apr 30;28(2):1-47. [doi: 10.1145/3443686]
54. Wang N, Xu H, Dhingra R, et al. The impact of later-life learning on trajectories of cognitive function among U.S. older adults. Innov Aging 2025;9(5):igaf023. [doi: 10.1093/geroni/igaf023] [Medline: 40386023]
55. Park DC, Bischof GN. The aging mind: neuroplasticity in response to cognitive training. Dialogues Clin Neurosci 2013 Mar;15(1):109-119. [doi: 10.31887/DCNS.2013.15.1/dpark] [Medline: 23576894]

Abbreviations

AI: artificial intelligence
ICT: information and communication technology
UCD: user-centered design

Edited by C Lee; submitted 20.Dec.2025; peer-reviewed by D David, L D'Ambrosio; revised version received 30.Mar.2026; accepted 28.May.2026; published 07.Jul.2026.

Please cite as:
Chung J, Mansion N, Gendron T, Wood RE, Demiris G
Smart Speaker–Based Applications to Support Social Connectedness in Older Adult Residents in Affordable Housing: User-Centered Design Study
JMIR Aging 2026;9:e90053
URL: https://aging.jmir.org/2026/1/e90053
doi:10.2196/90053

© Jane Chung, Natalie Mansion, Tracey Gendron, Rachel E Wood, George Demiris. Originally published in JMIR Aging (https://aging.jmir.org), 7.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution

XSL•FO
RenderX

License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of a Gamification-Based Intervention to Enhance Intergenerational Communication: Pre-Post Mixed Methods Evaluation

Eliza Lai-Yi Wong[1,2], PhD; Clement Cheuk Wai Ng[2], PhD; Dorothy Yingxuan Wang[1,2], PhD; Carol Ka Po Wong[2], MPH; Shirley Shuk Kuen Lui[2], BSc; Annie Wai-Ling Cheung[1,2], MPhil; Eng-kiong Yeoh[1,2], MBBS

[1]Centre for Health Systems and Policy Research, The Jockey Club School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Room 418, School of Public Health Building, Prince of Wales Hospital, Ngan Shing St, Sha Tin, China (Hong Kong)

[2]The Jockey Club School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Eliza Lai-Yi Wong, PhD
Centre for Health Systems and Policy Research, The Jockey Club School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Room 418, School of Public Health Building, Prince of Wales Hospital, Ngan Shing St, Sha Tin, China (Hong Kong)

Abstract

Background: Intergenerational communication has been recognized as a key factor in enhancing social connections among older adults by improving the quality of interactions, which is a crucial component of healthy aging. However, effective interventions to foster intergenerational communication and engagement remain scarce at this stage.

Objective: This study examined the effects of a gamification intervention for intergenerational dialogue on intergenerational communication and health-related capacity across 3 age cohorts and further analyzed the differential impacts of intervention characteristics on the game experience, in terms of satisfaction and engagement.

Methods: A 2-stage exploratory method was adopted with a cocreation approach to develop a theoretical framework of intergenerational dialogue for the gamification-based intervention, the Healthy Aging in Dialogue Tabletop Game, followed by a mixed methods study involving pre- and postintervention surveys and qualitative interviews to evaluate the effectiveness of the intervention in the young (aged 18 - 39 years), middle-aged (aged 40 - 64 years), and older adult (aged ≥65 years) groups. Intergenerational communication was assessed using tools measuring satisfaction and competence, whereas health-related capacity was evaluated using health literacy and patient enablement tools.

Results: A total of 68 respondents completed the intervention and evaluation, with 23, 14, and 31 respondents in the young, middle-aged, and older adult groups, respectively. This intergenerational intervention yielded notable benefits in intergenerational communication satisfaction (Cohen *d*=0.745), communication competence (Cohen *d*=0.377), and health literacy (Cohen *d*=0.409), especially among older adults. Patient enablement was significantly higher in older adults (mean score 6.10, SD 2.18) than in young adults (mean score 4.22) (*P*=.04). In addition, this study highlighted the importance of leveraging gamification to enhance intergenerational communication and engagement, ultimately combating social isolation among older adults.

Conclusions: The intergenerational dialogue tabletop game provided an enjoyable platform for cross-generational dialogue and enhanced each age group's communication competence, allowing participants to contribute and learn from one another. The findings provide additional evidence for aging research and shed light on approaches to foster social interaction and connection, thereby promoting health and active aging.

(JMIR Aging 2026;9:e77190) doi:10.2196/77190

KEYWORDS

intergenerational communication; health; gamification intervention; social isolation; social connection; aging

Introduction

Rapid technological advancement, urbanization, and aging have significantly reshaped social dynamics, family structures, and community engagement, leading to fewer face-to-face interactions. Evidence shows that digital communication increasingly replaces face-to-face interactions [1], with individuals often withdrawing from in-person contact due to technology use and fast-paced lifestyles [2]. These societal shifts, combined with changing family structures linked to population aging, contribute to reduced intergenerational engagement and weakened community ties [3]. Among these

XSL•FO
RenderX

transformations, social isolation has emerged as a critical concern since the 1990 s due to its profound implications across levels of society, including individuals, families, workplaces, disease groups, and broader society [4,5]. Social isolation is defined as the objective absence or insufficiency of meaningful and sustained communication [6], reflected in having too few interactions with others. It has become a priority global health issue recognized by the World Health Organization (WHO) and the United Nations, with an estimated 25% of older people and 15% of adolescents experiencing social isolation [7] because it has been linked to numerous adverse outcomes documented by a growing body of evidence, such as mental health decline, depression, dementia, cardiovascular disease, stroke, falls, reduced quality of life, hindered active aging, and increased mortality rates [8,9]. Our preliminary work shows that prevailing circumstances, such as pandemics, climate change, policies, family crises, and health deterioration, further intensify social isolation [10,11]. Loneliness, by contrast, represents a distressing feeling arising from a perceived and experienced lack of interaction [6]. Although social isolation is an objective form of social disconnection, whereas loneliness is a subjective form, they are strongly related [12], and both forms of disconnection carry significant health risks [7,13].

Many intergenerational programs have been shown to mitigate social isolation and loneliness by strengthening cross-age communication and expanding older adults' social networks [14-16]. Intergenerational communication, defined as the exchange of ideas, values, and experiences between different age groups [17], has been recognized as a key factor in addressing social isolation by fostering mutual understanding and improving the quality of intergenerational interactions within an intergenerational program [15]. Existing intergenerational programs mainly focus on structured connections and individual-based interactions (eg, one-way storytelling, companionship for social and health support), and community-based activities (eg, group-based activities related to arts, gardening, physical activity, cooking, digital games, volunteering, and neighborhood support) [14-16]. The conversation-based program remains limited, despite being the central component of social connection. Without engaging in dialogue topics or accommodating communication practices, social isolation can be perpetuated. Key components of communication include listening, observing verbal and nonverbal cues, and expressing thoughts and emotions through both verbal and nonverbal channels [18], all of which are essential for achieving mutual understanding, building a trusting relationship, and fostering an empathetic and caring environment [19]. When individuals lack adequate communication skills or have limited exposure to diverse conversational partners, they may struggle to accurately interpret others' cues or express themselves effectively. This is particularly evident in intergenerational contexts. Giles and Gasiorek [20] found that younger adults often overaccommodate older individuals through overprotection, whereas older adults tend to underaccommodate younger individuals by voicing harsh criticisms of their generation. A lack of intergenerational understanding may deepen generational divides, leading to age stereotypes that may contribute to nonaccommodation [21]. Without engaging in dialogue topics or practicing accommodating communication strategies, social isolation can be reinforced. Furthermore, most programs are time-limited, and older adults often lose access to the platforms and opportunities that enable ongoing interaction with younger people, leading to a decline in sustained engagement.

Given that the intervention was explicitly designed to structure and enhance intergenerational dialogue, this study focuses on proximal communication outcomes, specifically communication competence and communication satisfaction, as direct indicators of interaction quality. Although social isolation and loneliness motivate this work as broader public health concerns, these outcomes are conceptualized as distal and likely to emerge through sustained improvements in the communication process over time. Accordingly, this study evaluates the effects of the intervention on communication-related outcomes, providing a necessary foundation for future research to examine downstream impacts on social isolation and loneliness. This paper addresses this gap by developing a structured intergenerational tabletop game as a repeatable, low-cost platform that provides guided conversation topics and supports the four key communication components as intervention mechanisms that facilitate meaningful interaction beyond formal programs: guided conversational exchange; shared problem-solving through cooperative tasks; perspective-taking and mutual learning across generations; and opportunities to practice and build confidence in communication competence. Additionally, the intervention embedded health content to enhance health literacy and enablement, supporting participants' capacity to sustain and apply communication gains, as poor health literacy has been recognized by the WHO as a key driver of social isolation and a diminished capacity for social participation [3].

Methods

Overview

A 2-stage exploratory method was used to develop intergenerational dialogue prototypes for the gamification-based intervention and to evaluate its effectiveness on the communication process. Gamification was adopted to incorporate game-based intervention elements such as points, leaderboards, and badges into a nongame context to enhance the engagement experience [22]. This study investigated the effectiveness of the "Healthy Aging in Dialogue Card Game" on intergenerational communication, health, and game experience in three age cohorts: 18 - 39 years, 40 - 64 years, and ≥65 years. The objectives were to evaluate communication skills (communication competence and satisfaction), health literacy, self-efficacy, and engagement in the game (fun, encouragement to communicate, health knowledge building, and satisfaction) across age groups in the pre- and postsurveys.

Study Population

Overview

Participants who were Hong Kong residents and spoke Cantonese as their native language were eligible for the study, while those without the cognitive capacity to provide consent were excluded. Recruitment of young adults aged 18 - 39 years and middle-aged adults aged 40 - 64 years was conducted via

XSL•FO
RenderX

mass mailings at universities and social media, including Facebook (Meta Platforms, Inc) and Instagram (Meta Platforms, Inc). Older adults aged 65 years or older were recruited by nongovernmental organizations via flyers and a multicommunity-center approach.

Stage 1: Cocreation and Prototype Development for Intergenerational Dialogue

The development of the intergenerational dialogue intervention followed the revised Medical Research Council guidance for complex interventions [23], with an emphasis on theory-driven design, stakeholder input, and iterative refinement. A 4-member cocreation team, comprising an academic with training in nursing and public health (ELYW), a social worker, a public health practitioner (CKPW), and 1 potential end user (SSKL), led the multistage development process.

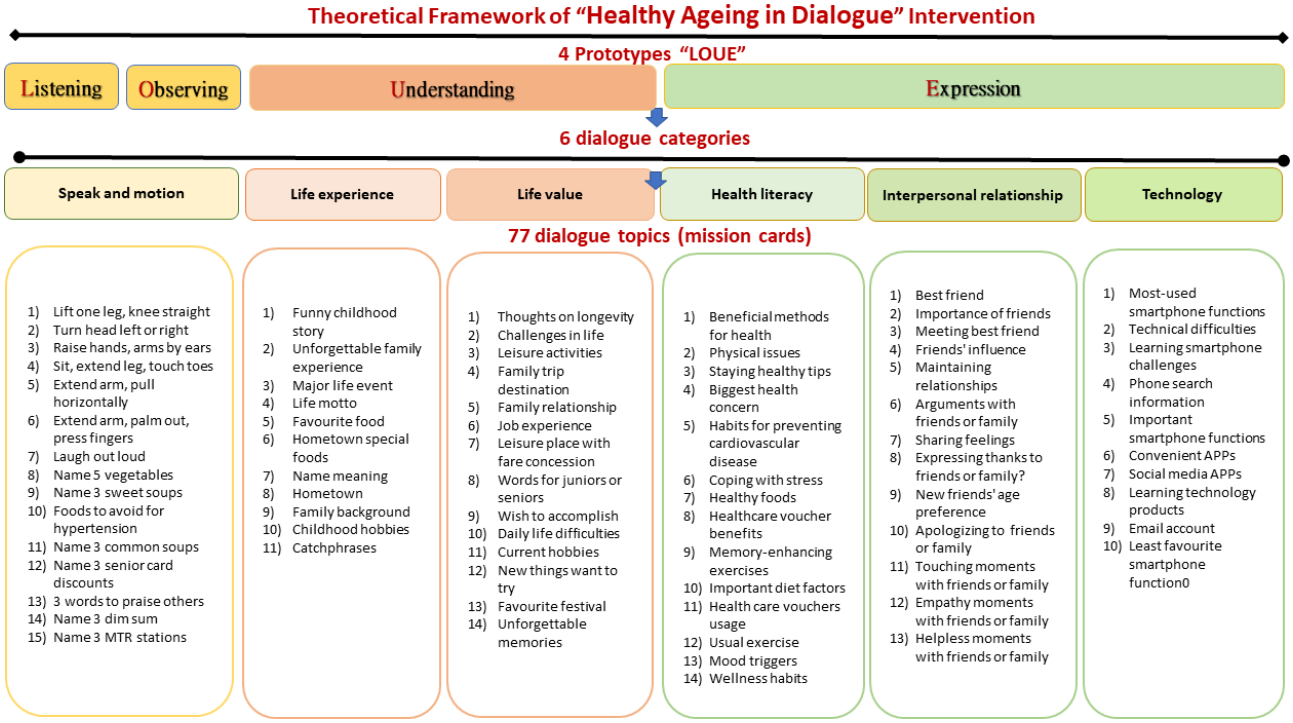
Theoretical Framework: Listening, Observing, Understanding, and Expression

The intervention was theoretically informed by the Listening, Observing, Understanding, and Expression (LOUE) framework, which integrates core verbal, nonverbal, and emotional communication processes [18]. LOUE was selected because it aligns closely with the goals of intergenerational dialogue, where effective communication depends not only on verbal exchange but also on emotional attunement, reciprocal self-disclosure, and perspective-taking. Specifically, active listening and empathic engagement are fundamental to building understanding and trust in interpersonal communication [24], while emotional attunement enables individuals to perceive and respond sensitively to others' affective states [25]. Perspective-taking further supports interpersonal connection by allowing individuals to understand others' thoughts and viewpoints, and reciprocal self-disclosure fosters relational closeness and social bonding [26]. In this study, LOUE served as a design framework to guide the generation, classification, and structuring of dialogue prompts, ensuring that gameplay systematically exercises these 4 communication skills.

Identification of Dialogue Categories and Topic Generation

At the identification stage, the cocreation team reviewed evidence from intergenerational communication and health promotion programs [27,28] and collaboratively identified recurring thematic domains relevant to everyday intergenerational interactions. Through iterative discussion and clustering of draft prompts, six dialogue categories, defined as thematic groupings of communication prompts, were agreed upon: (1) Speak and Motion, (2) Life Experience, (3) Life Values, (4) Health Literacy, (5) Interpersonal Relationship, and (6) Digital Literacy (Figure 1). Building on these categories, the team generated an initial pool of conversation prompts and refined them based on relevance, clarity, and suitability for group dialogue. This process resulted in 77 dialogue topics, which were further refined through a design review to ensure balanced representation across categories and suitability for delivery in a game-based format.

Figure 1. Theoretical framework of the Healthy Aging in Dialogue intervention. MTR: metro.



Operationalization as Mission Cards and Intervention Structure

The 77 dialogue topics were operationalized as mission cards, each containing a structured prompt designed to elicit turn-taking dialogue and reflection. Mission cards were evenly distributed across the 6 dialogue categories. Each card primarily targeted 1 LOUE dimension (eg, listening or expression), while still engaging all 4 communication skills through interaction and response. This structure ensured that each game turn focused on a specific topic while cumulatively strengthening intergenerational communication competencies.

XSL•FO
RenderX

Gamified Tabletop Intervention Design

These mission cards were embedded within a gamified tabletop board game titled the "Healthy Aging in Dialogue Tabletop Game." The intervention consists of players taking turns drawing mission cards, completing dialogue tasks, and accumulating points through participation rather than correct performance. Each player selects a chess piece at the start of the game to represent equal participation and balanced turn-taking, highlighting intergenerational communication as a reciprocal process.

Gamification elements were incorporated to enhance engagement, self-expression, and motivation [29]. A point-based system (1‐3 points, reflecting increasing depth of conversational engagement, ranging from factual exchanges to meaning-making and reflective discussion) was used. Evidence suggests that such gamification strategies can promote interaction, maintain autonomy, and sustain engagement [29,30]. Importantly, 58 of 77 (75%) mission cards require the active involvement of multiple players, such as inviting adjacent or all players to respond, thereby encouraging listening, observing others' perspectives, and shared reflection [31-33].

Game Format and Play Procedure

The intervention is designed for 3‐6 players aged 11 years or above, with each group comprising at least 1 young adult (18‐39 years), 1 middle-aged adult (40‐64 years), and 1 older adult (≥65 years). A typical game session lasts 20‐30 minutes, or until a preset endpoint is reached, to maintain attention and prevent conversational dominance. Players take turns clockwise, draw mission cards, and complete tasks within approximately 1‐2 minutes per turn. There are no correct answers, and judgmental feedback is discouraged to maintain a respectful and inclusive dialogue environment. Details of the game setup and rules are provided in Figure 2.

XSL•FO
RenderX

Figure 2. Healthy Aging in Dialogue Tabletop Game.



- This game helps players from different generations talk, listen, and learn from each other. By complementing mission cards and moving across the game board, players explore six areas of everyday life – actions, life experience, value, health, relationships, and technology – while strengthening communication skills and understanding across ages.
- Setup includes three parts: 1 Score Board, 77 Mission Cards, and 6 Chess. Recommend having 3-6 participants from different generational groups.

Game rule

1. Each player chooses one chess piece and places it at the starting point.
2. Shuffle the mission cards and place them facing down on the table.
3. Players' orders are decided by group consensus. For example, the youngest player starts, or the player with the latest birth month starts.
4. The first player flips over the top card from the shuffled deck, then answers the question or performs the task shown on the card. There are no right or wrong answers, only moments to connect, share, and enjoy others' company.
5. The card may invite different players to participate:



   i. All players participate.
   ii. You and the player on your left participate.
   iii. You and the player on your right participate.
   iv. You and the players on both your left and right participate.
   v. Only you participate.
6. After fulfilling the required task, the player(s) gain points and move their chess forward with the number of steps (ranging from 1 to 3) indicated on the mission card.
7. Players take turns in a clockwise manner by revealing the next top mission card.
8. Recommended play time is about 30 minutes; the first player to reach the end or the player with the highest score within the time limit wins the game.

Stage 2: Mixed Methods Evaluation of the Effectiveness of the Gamification-Based Intervention

A mixed methods study method consisting of a 2-phase longitudinal survey and follow-up qualitative sessions involving focus group or individual interviews was conducted on-site to evaluate the effectiveness of the gamification-based intervention—the Healthy Aging in Dialogue Tabletop Game. Before the intervention, all participants completed the baseline questionnaires independently or with assistance from a trained

XSL•FO
RenderX

research assistant. Immediately after the intervention, they completed the follow-up questionnaire. A minimum sample size of 63 was recommended to detect a medium effect size of 0.5 at a significance level of .05 and statistical power of 80%, with the assumed SD of 1 [34]. The questionnaire consisted of the following four sections: (1) intergenerational communication, (2) health-related capacity, (3) game experience (postintervention only), and (4) demographics (baseline only).

Primary Outcome Measure

Intergenerational communication was assessed in terms of communication satisfaction and communication competence. Communication satisfaction was assessed using four conceptually relevant items selected from Hecht Interpersonal Communication Satisfaction Inventory [35]. The item "I felt I could talk about topic with the other person" was modified from "anything" to "topic" to fit the structured dialogue context, while the remaining three items—"I am generally satisfied with the conversations," "I enjoy the conversations," and "These conversations flow smoothly"—were directly adopted. All items were rated on a 6-point Likert scale ranging from "1=extremely disagree" to "6=extremely agree." Communication competence was evaluated using 6 selected items from the Interpersonal Communication Competence Scale (ICCS) [36], reflecting perspective-taking, emotional awareness, conversational flow, topic management, and sensitivity to nonverbal cues: (1) I can put myself in others' shoes, (2) I do not know exactly what others are feeling, (3) other people think that I understand them, (4) my conversations are characterized by smooth shifts from one topic to the next, (5) I take charge of conversations I am in by negotiating what topics we talk about, and (6) in conversations, I perceive not only what others say but also what they do not say. Responses were rated on a 4-point Likert scale ranging from "1=extremely disagree" to "4=extremely agree."

Secondary Outcome Measures

Health-Related Capacity

Health-related capacity was evaluated using measures of health literacy and enablement. Health literacy was assessed with three items from the European Health Literacy Survey Questionnaire (HLS-EU-Q47) [37], including (1) finding information on treatments of illness that concern you, (2) finding information on how to manage mental health problems such as stress or depression, and (3) judging which everyday behavior (such as drinking and eating habits and exercise) is related to your health. Responses were rated on a 4-point Likert-scale ranging from "1=very difficult" to "4=very easy." A Chinese version of the HLS-EU-Q47 was previously validated to assess health literacy in Chinese populations [38]. Self-care competence was examined by adapting 6 items from the Patient Enablement Instrument (PEI) to the intervention context: (1) able to cope with life, (2) able to understand your health condition, (3) able to cope with your health condition, (4) able to keep yourself healthy, (5) confident about your health, and (6) able to help yourself. Responses were rated on a 3-point Likert scale: "much better," "better," and "same or less." The original PEI was previously translated and validated in the Chinese population [39].

Game Experience

Game experience was evaluated after the intervention using the following five questions about the board game: (1) fun to play, (2) encourage intergenerational communication, (3) enhancement of health knowledge, (4) satisfaction with the mission card design, and (5) overall satisfaction with the experience. Responses were rated on a 4-point Likert scale ranging from "extremely disagree" to "extremely agree." Participants were also asked about their perceived game engagement on a 10-point scale ranging from very disengaged (0) to very engaged (10).

Demographic data were collected before the intervention, including age, gender, living arrangement, education attained, marital status, religion, and self-reported noncommunicable diseases.

All participants were invited to focus group discussions immediately after the game intervention and post survey to collect immediate narrative feedback. The focus groups, consisting primarily of participants from the same age group, were conducted to explore the dialogue experience and views on the board game design. For participants who could not join the focus groups but wished to elaborate on their opinions of the board game, in-depth individual interviews were scheduled within 10 days of their board game experience to minimize recall bias. The focus group discussions and individual interviews lasted approximately 45 - 60 minutes, were audiotaped, and participants were invited to freely express relevant ideas on the discussion topics. The qualitative interviews were led by researchers experienced in qualitative data collection using a discussion guide to discuss [40] with three major components, including (1) attitudes toward intergenerational communication, (2) healthy aging, and (3) views on the design and application of the card game.

Data Analysis

A quantitative analysis of the survey responses was conducted using IBM SPSS (version 29.0). Sample characteristics were reported as descriptive statistics. Items related to intergenerational communication satisfaction, communication competence, health literacy, and the patient enablement index were summed to yield total scores, whereas paired *t* tests were performed to detect significant changes in target outcomes with 95% CIs. Subgroup analyses by age group were further conducted to examine the impact of the card game across different generations. Cohen *d* of the three target outcomes (intergenerational communication satisfaction, communication competence, and health literacy) was computed using the mean difference before and after the intervention to evaluate the effect sizes, and Hedges correction was applied for subgroups with n<20 [41]. Effect sizes were considered small, medium, and large with cutoffs of 0.2, 0.5, and 0.8, respectively [41,42]. Findings related to key game experience components, including "fun to play," "encourage intergenerational communication," "enhance health knowledge," and "satisfied with board game question and mission design," were first presented descriptively. To further examine the associations between the key game experience components and the likelihood of reporting higher levels of overall satisfaction and game engagement, multivariate

XSL•FO
RenderX

binary logistic regression models were conducted. For these analyses, the original Likert-scale or continuous scale outcome variables were dichotomized into "higher" vs "lower" satisfaction or engagement. A higher level of satisfaction was defined as a response of "strongly agree" on the 4-point Likert satisfaction scale, while a higher level of engagement was defined as a score of 9 - 10 on the 0 - 10 engagement scale. The reported odds ratios (ORs) and corresponding 95% CIs were adjusted for age group in the models. A *P* value of <.05 was considered statistically significant.

Audiotapes recorded during the qualitative sessions were transcribed verbatim, and the transcripts were managed with Microsoft Excel. Experiences related to the game in three aspects, including intergenerational communication, health-related capacity, and overall playing experience, were identified by thematic analysis. The themes and quotes were identified with reference to the discussion guide, and the transcripts were coded twice by 2 independent reviewers to develop the study codebook. With the emergent and recurrent themes identified from the transcripts, these codes were checked and discussed among the independent reviewers to enhance consistency throughout the analysis. Codes were discussed among the research team in case of disagreements. Excerpts were quoted to illustrate the findings uncovered from the qualitative phase.

Ethical Considerations

The tabletop game intervention is a component of a social isolation project approved by the Chinese University of Hong Kong Survey and Behavioral Research Ethics Committee (SBRE-21 - 0870). Before the game intervention, informed consent was obtained from participants for participation in the game, surveys, qualitative interviews, and audiotaping.

Results

Demographics

A total of 81 participants were enrolled in the tabletop game, and 68 completed the intervention and evaluations, yielding a response rate of 84%. A total of 19 game groups were arranged, each involving 3 - 5 participants with at least 1 participant from each of 3 age groups. Participants were distributed across the 3 age groups (Table 1). The majority of participants were female (82.4%), living with others (61.8%), had a tertiary education level or above (55.4%), were single (47.1%), and had no religious belief (52.9%). About half of the participants reported having a communicable disease (48.5%). After the intervention and post survey, 52 of 68 (76.5%) participants were interviewed in 8 focus groups and 3 individual interviews. Only 1 focus group was mixed-aged, and 3 individual interviews were arranged due to the unavailability to join the focus group. There was no difference in the direction or richness of sharing; therefore, all focus groups and individual interviews were combined for data analysis. Regarding the characteristics of the interviewees across the 3 age groups, a higher proportion with tertiary education, who were single, who had no religious belief, and who had no chronic disease was observed among the youth, while a higher proportion living alone and being widowed or divorced was observed in the middle-aged and older adult groups. The majority of participants were female (80.8%) across all 3 age groups.

XSL•FO
RenderX

Table . Demographic characteristics of respondents.

Demographic	Total (n=68), n (%)	Young (n=23), n (%)	Middle-aged (n=14), n (%)	Older adults (n=31), n (%)
Age range (years)	19 - 92	19 - 33	46 - 64	66 - 92
Sex				
Male	12 (17.6)	7 (30.4)	1 (7.1)	4 (12.9)
Female	56 (82.4)	16 (69.6)	13 (92.9)	27 (87.1)
Live status				
Living alone	26 (38.2)	1 (4.4)	4 (28.6)	21 (67.7)
Living with others	42 (61.8)	22 (95.6)	10 (71.4)	10 (32.3)
Education level				
Primary or below	7 (10.3)	0 (0.0)	0 (0.0)	7 (23.3)
Secondary	23 (33.8)	1 (4.3)	5 (35.7)	17 (56.7)
Tertiary or above	37 (54.4)	22 (95.7)	9 (64.3)	6 (20.0)
Relationship status				
Single	32 (47.1)	19 (82.6)	3 (21.4)	10 (32.3)
Married or stable relationship	18 (26.5)	3 (13.0)	7 (50.0)	8 (25.8)
Widowed or divorced	18 (26.5)	1 (4.3)	4 (28.6)	13 (41.9)
Religion				
No	36 (52.9)	17 (73.9)	3 (21.4)	16 (51.6)
Yes or traditional worship	32 (47.1)	6 (26.1)	11 (78.6)	15 (48.4)
Self-reported noncommunicable disease				
No	35 (51.5)	21 (91.3)	8 (57.1)	6 (19.4)
Yes	33 (48.5)	2 (8.7)	6 (42.9)	25 (81.6)

Intergenerational Communication

There was a significant improvement in the participants' self-ratings of communication satisfaction, with a mean difference of −2.76 (95% CI −3.66 to −1.87; Table 2). Participants across all 3 age groups showed significant improvement after the intervention, with mean differences of −2.52 in young adults, −3.93 in middle-aged adults, and −2.42 in older adults. The largest effect size was observed in the middle-aged group (Cohen *d*=1.13). For participants' ratings of communication competence, significant improvement was also observed after the intervention, with a mean difference of −1.04 (95% CI −1.71 to −0.37). Improvement was shown across all 3 age groups, but statistically significant improvement was observed only among older adults, with a mean difference of −1.32 (95% CI −2.52 to 0.12). Comparing the 2 communication dimensions, a larger effect size was observed for communication satisfaction (medium effect size of 0.75) than in communication competence (small effect size of 0.38) after the intervention.

XSL•FO
RenderX

Table . Effectiveness of the intervention on intergenerational communication.

Intergenerational communication and measurement	Total (N=68)	Young (n=23)	Middle-aged (n=14)	Older adults (n=31)
Communication				
Baseline, mean (SD)	18.57 (3.31)	18.48 (2.57)	18.14 (3.53)	18.84 (3.74)
Satisfaction				
Intervention, mean (SD)	21.34 (1.87)	21.00 (2.20)	22.07 (1.73)	21.26 (1.61)
Baseline intervention, mean (95% CI)[a]	−2.76 (−3.66 to −1.87)[b]	−2.52 (−4.02; −1.02)[b]	−3.93 (−5.82; −2.04)[b]	−2.42 (−3.91; −0.93)[b]
Cohen *d*[a]	0.745	0.728	1.131	0.595
Communication				
Baseline, mean (SD)	17.62 (2.73)	18.26 (2.40)	18.43 (2.44)	16.77 (2.91)
Competence				
Intervention, mean (SD)	18.66 (2.33)	19.34 (2.21)	18.79 (2.36)	18.10 (2.34)
Baseline intervention, mean (95% CI)[a]	−1.04 (−1.71 to −0.37)[c]	−1.09 (−2.23 to 0.06)	−0.36 (−1.22 to 0.51)	−1.32 (−2.52 to −0.12)[c]
Cohen *d*[a]	0.377	0.411	0.224	0.404

[a]Cohen *d* was computed with the mean difference by Hedges correction for n<20.

[b]Paired *t* test.

[c]*P*<.05

Qualitative interviews supported these quantitative findings, showing that the intervention facilitated easier and more relaxed communication.

A young participant remarked:

> *Sometimes, when talking to middle-aged and older individuals, you may not always know what to say. Not everyone is skilled in discussing various topics with others. On the contrary, this approach raises topics for conversation on which everyone can share their opinions. It's not too tense, and it makes the conversation with them much easier.* [BG22 (E)]

A middle-aged participant further added:

> *… as players often raise different answers, the contrast between them may stimulate even the shy ones to stay engaged with the tasks and responses. This promotes communication, or even builds the trust for any physical assistance with the action tasks when necessary.* [CUBG35 (B)]

An older adult participant also shared:

> *The gamification provides a welcoming space to share personal matters (anonymously), which helps players to soothe and relax, especially on sensitive topics that might be difficult to bring up with friends, as you may be uncertain if the friend may spread the word unintentionally.* [BG26]

Health-Related Capacity

Health literacy significantly improved after the intervention (Table 3). Across the 3 age groups, significant improvements in health literacy with a significant result were observed in the middle-aged (mean difference −0.86, 95% CI −1.60 to −0.11) and older adult groups (mean difference −1.19, 95% CI −1.96 to −0.43) but not in the youth group (mean difference −0.13, 95% CI −0.88 to 0.62). Focusing on the patient enablement effect of the tabletop game intervention, PEI scores significantly increased (Table 3). A comparison of results across the 3 age groups showed that older adults' PEI (mean 6.10, SD 2.18) was significantly higher than that of the young adult group (mean 4.22, SD 2.59). However, no significant difference was observed in the other comparisons. Details of the comparisons are tabulated in Table 2.

XSL•FO
RenderX

Table . Effectiveness of the intervention on health enhancement.

Health enhancement and measurement	Total (n=68)	Young (n=23), mean (SD)	Middle-aged (n=14), mean (SD)	Older adults (n=31), mean (SD)
Health literacy				
Baseline, mean (SD)	8.59 (2.29)	10.00 (1.38)	8.93 (2.06)	7.39 (2.33)
Intervention, mean (SD)	9.35 (1.74)	10.13 (1.46)	9.79 (1.48)	8.58 (1.77)
Baseline intervention, mean (95% CI)[a]	−0.76 (−1.22 to −0.31)[b]	−0.13 (−0.88 to 0.62)	−0.86 (−1.60 to −0.11)[b]	−1.19 (−1.96 to −0.43)[b]
Cohen *d*[c]	0.409	0.075	0.624	0.572
Patient				
Intervention, mean (SD)	5.41 (2.78)	4.22 (2.59)	5.86 (3.68)	6.10 (2.18)
Enablement				
One-way ANOVA (*P* value)	.04[b]	—[d]	—	—
Index (PEI[e])				
Post hoc test, group comparison[f]	1=2; 2=3; 3>1	—	—	—

[a]Paired *t* test.

[b]*P*<.05.

[c]Cohen *d* was computed with the mean difference by Hedges correction for n<20.

[d]Not applicable.

[e]PEI: Patient Enablement Instrument.

[f]For the post hoc test (Bonferroni): 1=young, 2=middle aged, and 3=older adults.

Qualitative data corroborated these results, and a young participant indicated the benefits of the intervention in enhancing health literacy, particularly regarding mental health:

> *There were several cards that addressed topics like how to cope when you're under a lot of stress, or what might make you feel down emotionally. When discussing these topics, everyone can reflect more on their own experiences. It makes us think about times when we ourselves may have felt a lot of emotional pressure or faced emotional issues. Some of the cards also ask how you would cope with such situations, and everyone can share their experiences. This allows for the exchange of skills and strategies for dealing with various topics.* [BG22 (E)]

A middle-aged participant highlighted:

> *The youth and I (middle-aged) had no idea of the welfare offered by the healthcare voucher or the senior citizen card; we had to check online immediately to learn the benefits… it is like acquiring new knowledge. It is not that we don't care, but what we knew was limited, and only then did we realise there are things experienced by the older adults that we lack understanding.* [BG20 (5)]

An older adult participant shared the impact of health literacy:

> *… when you live to a hundred, you would need the help from a carer (social support), and you can't take good care of yourself. This feels powerless, and the other player agrees, sharing the same concern (about Health-related Quality of Life in ageing). We are worried that we become incapable of self-care, and this is not the ideal way (of community ageing).* [BG26]

In terms of patient enablement, a young participant noted:

> *The discussion on the HA Go (mobile application of the Hospital Authority) is important, as it provides information for carers like me, who monitored the health conditions of my family with the HA Go… it is still a means for us to manage health regardless of physical distance...* [CUBG13 (M)]

A middle-aged participant also mentioned:

> *The tabletop game may empower the participants to cope with their future, such as guiding them to think and plan ahead… the positive planning should be for everyone, and everyone should foster a positive mindset.* [BG18 (4)]

An older adult participant expressed:

> *Expressing my thoughts with the game play made me feel more at ease… and by drawing the tasks randomly, we may not dodge from the problem, but to face it directly. It feels like a therapy session, and it's therapeutic.* [BG15 (1)]

Game Experience

A strong positive experience with the tabletop game was revealed, with responses heavily inclined toward "agree" and "extremely agree" in all 4 aspects of game experience. The

XSL•FO
RenderX

highest proportion of participants found that the tabletop game could encourage intergenerational communication, with 51.5% strongly agreeing, followed by "fun to play" (44.1%), "satisfied with mission card design" (30.9%), and "enhanced our health knowledge" (26.5%). The 4 aspects of game experience are reported in Table 4. The association between game experience and engagement, and overall satisfaction is reported in Table 5. It was found that "encourage intergenerational communication" (OR 4.49, 95% CI 1.09 to 18.58) was significantly associated with participants' engagement in the tabletop game, whereas "encourage intergenerational communication" (OR 6.22, 95% CI 1.03 to 37.37) and "satisfied with mission card design" (OR 10.20, 95% CI 1.46 to 71.37) were significantly associated with the overall satisfaction of the tabletop game.

Table . Experience of the gamification-based intervention.

Gamification-based intervention component	Higher level of satisfaction, n (%)	Higher level of satisfaction[a], n (%)	*P* value[b]	Higher level of engagement[c], n (%)	*P* value[b]
Fun to play					
Agree and disagree	38 (55.9)	5 (13.2)	<.001	18 (47.4)	.001
Strongly agree	30 (44.1)	23 (46.7)	<.001	26 (86.7)	.001
Encourage intergenerational communication					
Agree and disagree	33 (48.5)	4 (12.1)	<.001	14 (42.4)	<.001
Strongly agree	35 (51.5)	24 (68.6)	<.001	30 (85.7)	<.001
Enhance health knowledge					
Agree and disagree	50 (73.5)	13 (26)	<.001	28 (56)	.020
Strongly agree	18 (26.5)	15 (83.3)	<.001	16 (88.9)	.026
Satisfaction with board game questions and mission design					
Agree and disagree	47 (69.1)	9 (19.2)	<.001	26 (55.3)	.026
Strongly agree	21 (30.9)	19 (90.5)	<.001	18 (64.7)	<.05

[a]A higher level of satisfaction is defined as "strongly agree" satisfaction.

[b]Chi-square and Fisher exact test were conducted.

[c]A higher level of engagement is defined as scoring 9 - 10 on the engagement question.

XSL•FO
RenderX

Table . Relationship between gamification-based intervention components and board game experience.

Board game experience	Higher level of satisfaction[a], adjusted OR[b] (95% CI)	Higher level of game engagement[c], adjusted OR[b] (95% CI)
Intercept	0.07 (0.012 - 0.39)[d]	0.57 (0.19-1.75)
Fun to play		
Agree and disagree	Ref[e]	Ref
Strongly agree	4.63 (0.71 - 30.02)	2.73 (0.48-15.56)
Encourage intergenerational communication		
Agree and disagree	Ref	Ref
Strongly agree	6.22 (1.03 - 37.37)[d]	4.49 (1.09-18.58)[d]
Enhance health knowledge		
Agree and disagree	Ref	Ref
Strongly agree	3.67 (0.24 - 56.60)	1.47 (0.18-12.11)
Satisfied with mission card design		
Agree and disagree	Ref	Ref
Strongly agree	10.20 (1.46 - 71.37)[d]	1.07 (0.16-7.12)

[a]A higher level of satisfaction is defined as “strongly agree” satisfaction.

[b]Adjusted by 3 age groups: 18 - 39 years, 40 - 64 years, and >65 years.

[c]A higher level of engagement is defined as scoring 9 - 10 on the engagement question.

[d]$P<.05$.

[e]Ref: reference group.

Regarding the qualitative findings, participants emphasized the opportunities this game offered for intergenerational interaction. For example, one young participant mentioned:

> *The tabletop game provides a chance, or a platform, for each generation to interact and communicate. You know how busy the locals are, not to mention having meals together, but it was rare for different generations just to sit together. The card game would offer motivation for different generations to come together.* [BG27]

An older adult participant further added:

> *The card game encourages activities between the youth, middle-aged, and older adults. With these opportunities, we may interact and get to know each other in a friendly setting. It is nice that we can engage with real friends.* [CUBG30]

In addition, a young participant expressed enjoyment of the game:

> *I find the game enjoyable. The tasks are simple and easy to complete. As other students shared, there are tasks that are relatively exciting and playful, especially those with a time limit or an action-based format. These game play strategies are challenging yet fun...* [BG25 (C)]

A middle-aged participant also commented on the game’s enjoyable nature:

> *… the older adults seem to enjoy the action tasks, and it’s such a pleasant scene when everyone engages with the same actions on an intergenerational level.* [CUBG34 (H)]

An older adult participant reflected on the intergenerational engagement:

> *I think the card game experience is nice, and it’s my first attendance at such a game. The intergenerational engagement seems to be drawing everyone closer, and I feel involved as it includes topics related to my childhood.* [CUBG29]

The game’s design was recognized as a key factor in its success. A young participant mentioned:

> *I would say, first of all, what impressed me the most was that the rules of the game were very simple… I think the key is to make sure they understand the rules of the game so they are willing to engage. If the rules were too complicated, they would likely spend too much time figuring out what to do and what each element meant. So, I believe ensuring a smooth game operation is crucial for the success of the game experience.* [BG24 (D)]

A middle-aged participant also expressed satisfaction with the game design:

> *It’s obvious that the card game encourages expansion of social network by providing common topics and activities… By involving adjacent players, this is a great design as there will be no dominating players, but they will share the experience together.* [BG20]

An older adult participant also acknowledged the game design:

XSL•FO
RenderX

Introducing oneself through playing cards makes it more interesting for everyone listening. By using this relaxed atmosphere, it allows everyone to get to know each other better. [CUBG25]

Discussion

Principal Findings

This intergenerational intervention yielded notable benefits in intergenerational communication, health literacy, and patient empowerment, especially for older adults. While the majority of gamification interventions for older adults have focused on evaluating the impact on health outcomes [43], the current evaluation provides a comprehensive review of communication and health-related outcomes. In addition, this study highlighted the importance of leveraging gamification to enhance intergenerational communication and ultimately combat social isolation among older adults.

Across all age groups, participants reported significantly higher communication satisfaction post intervention, indicating that the game fostered enjoyable and fulfilling dialogue. While short-term increases in communication satisfaction following an engaging activity are, to some extent, expected and may partly reflect novelty or positive study participation effects, communication satisfaction reflects perceived interaction quality rather than transient mood alone. In this intervention, satisfaction was fostered through structured dialogue, guided turn-taking, and role rotation, which targeted communication processes that can plausibly generalize beyond the immediate gameplay context. Nevertheless, sustained effects cannot be inferred from the present design and require confirmation through longitudinal follow-up. Notably, only older adults showed a significant improvement in their communication competence scores, whereas younger and middle-aged adults did not. One explanation lies in baseline differences: older participants may have had fewer opportunities for rich social interaction before the study due to the prevalence of social isolation among Hong Kong's older population. Indeed, as of 2021, nearly 190,000 adults aged 65 years and older were living alone in Hong Kong, reflecting a 60% increase since 2011 [44]. Such isolation is linked to declines in communication skills and cognitive function [45]. By facilitating regular, meaningful conversations, the game likely helped older participants practice and regain communication competencies. This aligns with evidence that increasing social interaction can improve older adults' performance on verbal communication tasks [46]. In contrast, younger adults—who typically have more daily social communication—may have been closer to a ceiling in communication ability, limiting observable gains.

Differential effects on health literacy were also observed. Middle-aged and older participants showed significant improvements in health literacy after gameplay, whereas the youth group did not. Younger adults likely entered the study with higher baseline health literacy, given their greater exposure to education and digital information, and thus had less room for measurable improvement. Older adults, on the other hand, often start with lower health literacy and greater health knowledge gaps. The game's health literacy module provided an engaging, accessible means of learning health information, resulting in tangible knowledge gains for these groups. This finding is consistent with prior research showing that game-based interventions can substantially increase health-related knowledge [47]. Moreover, intergenerational learning theory suggests that structured dialogue between generations creates opportunities for mutual education and attitude change [48]. During the health literacy tasks, younger players might have shared digital health information or taught health concepts, benefiting their older counterparts. Reciprocally, older adults may have shared personal health experiences, which could enrich understanding for young and middle-aged adults. The middle-aged group, interestingly, also improved in health literacy, suggesting that those in midlife, often caring for both children and older adult parents, found the health knowledge and cross-generational communication practice directly relevant and empowering for their caregiving and self-care roles.

Importantly, the older group showed a significantly higher PEI score post intervention than the youth group, with no significant differences between the other groups. The high PEI score in older adults suggests that the game may have enhanced their confidence and sense of "enablement" in dealing with health issues. Several mechanisms could explain this. First, older participants likely felt a sense of accomplishment and agency by contributing actively to group tasks, such as sharing life experiences or completing challenges. Being listened to by younger peers may have bolstered their self-esteem and sense of purpose, which, according to Socioemotional Selectivity Theory, is highly valued in late life [49]. The theory posits that as people age and perceive a shorter future time horizon, they prioritize meaningful social interactions and emotional fulfillment. Importantly, this does not imply that brief interactions with unfamiliar partners function as strong-tie relationships; rather, the intervention may have uncovered interaction contexts characterized by emotional affirmation, mutual respect, and psychological safety, which align with the socioemotional qualities prioritized in later life. In our study, the older adults' disproportionate gains in communication ability and health literacy could be a good start toward confidence-building in intergenerational interaction and a strong sense of enablement. Moreover, the opportunity to impart wisdom and bond with younger people likely fulfilled older adults' socioemotional goals, thereby maximizing their engagement and leading to greater gains from the program. In contrast, younger participants generally have a more future-oriented outlook and myriad social connections; while they enjoyed the game, as evidenced by improved satisfaction, they may not have experienced the same degree of program gains or personal empowerment.

The outcome was a unanimous perception across all age groups that the game strengthened intergenerational communication and was well-designed, fun, and educational. These positive subjective evaluations are noteworthy given the challenge of designing content appealing to diverse ages. By using gamification—points, challenges, and playful interactions—the intervention kept participants intrinsically motivated. Gamification theory holds that introducing game elements into nongame contexts (such as health education) can enhance user

XSL•FO
RenderX

engagement and motivation [50]. Consistently, our findings show high engagement and satisfaction scores, indicating that the game format was indeed engaging. Prior studies in older adults have also found that appropriately designed card games can acutely boost mood and well-being during play [51]. Our intergenerational game seems to have capitalized on these effects, creating an enjoyable learning environment that felt less like a health workshop and more like a social gathering.

The association analyses provide further insight into what drove the intervention's success. We found that participants who felt the game "encouraged intergenerational communication" were significantly more likely to report higher engagement in the activities and greater overall satisfaction, reinforcing the importance of perceived interaction quality as a proximal outcome. In other words, when players perceived that the core purpose of the game, bridging generations, was being achieved, their involvement and satisfaction were maximized. This underscores the importance of making the intergenerational benefits salient—participants who recognized the value in connecting with other ages likely invested more effort and attention, potentially creating a virtuous cycle of engagement and positive experience. We also found that satisfaction with the game design was strongly correlated with overall satisfaction. This highlights the role of user-centered design, encompassing elements such as clear rules, culturally relevant content, and balanced difficulty across modules, which likely enhanced enjoyment. From a gamification perspective, well-designed game mechanics (eg, rotating tasks, creative challenges, and immediate feedback through group discussion) can fulfill psychological needs such as autonomy, competence, and relatedness, leading to higher satisfaction [52].

Limitations

First, although the inclusion criteria were broad, participants voluntarily enrolled in the intervention, and the majority were female, which may not be representative of all eligible individuals; selection bias may therefore limit generalizability. However, the qualitative interviews provided in-depth insights that may help inform the design of a larger-scale randomized controlled trial. Second, intergenerational communication outcomes were assessed using selected items from validated instruments rather than the complete scales. While this may limit psychometric completeness, the selected items were purposefully chosen to reflect communicative processes relevant to brief, structured intergenerational interactions within a gamified setting. Accordingly, findings should be interpreted with appropriate caution. Full-length scales include items less applicable to this context, and their use may compromise relevance and participant engagement. Third, some outcomes were dichotomized for analysis due to skewed response distributions, which may have reduced variability and statistical power and potentially obscured differences in participants' responses. Fourth, the sample size for the middle-aged group was comparatively smaller than that of the other 2 groups, although the minimum sample size requirement was met. This imbalance may have limited statistical power for 3-group comparisons. Finally, a 30-minute game duration was recommended, but this was not implemented strictly to sustain positive game dynamics. Variations in actual game time may influence comparisons across groups. However, these variations may also reflect natural engagement patterns in intergenerational communication settings.

In conclusion, the intergenerational card game intervention in Hong Kong demonstrated an effective approach to enhancing communication satisfaction, health literacy, and health enablement, particularly benefiting older adults in an aging, isolated population. By providing an enjoyable and structured platform for cross-generational dialogue, the game enabled meaningful engagement and mutual learning between age groups, fostering high levels of participant engagement. These findings underscore the potential of playful, socially embedded interventions to complement traditional health care services by addressing psychosocial determinants of health. Given its low cost, scalability, and high acceptability, this approach holds promise as a public health tool in aging societies worldwide. Future research should further examine the tool's psychometric properties and extend this work through randomized controlled trials or longitudinal studies to assess longer-term impacts on health behaviors, clinical outcomes, and broader dimensions of social connectedness.

Acknowledgments

The authors are grateful to the Tung Wah Group of Hospitals for their joint efforts in cocreating the card game, recruiting older adults, and arranging the intervention's logistics. Tad Chun-kwan Ho was a major contributor to the card game development as a social worker, while Leopard Chi-keung Lam and Danielle Man-yu Chan provided tremendous support in the tabletop game development. Last but not least, the authors thank all participants for joining the game and sharing their valuable views. No generative artificial intelligence tool was used.

Funding

We thank the Hong Kong Jockey Club Charities Trust for its funding support, which made this intervention possible.

Data Availability

The dataset used or analyzed is not publicly available due to confidentiality considerations and the ongoing project progress. A reasonable request may be made to the corresponding author for further consideration.

XSL•FO
RenderX

Authors' Contributions

Conceptualization and funding acquisition were undertaken by ELYW. Data curation was performed by CCWN, CKPW, and SSKL. Formal analysis was conducted by DYW, CCWN, and AWLC. An investigation was carried out by ELYW, CCWN, and DYW. Methodology, supervision, and validation were the responsibility of ELYW. Visualization was completed by CCWN and AWLC. The original draft of the paper was prepared by ELYW, CCWN, and DYW. Review and editing of the paper were undertaken by CKPW, SSKL, AWLC, and EKY.

Conflicts of Interest

None declared.

References

1. Dwyer RJ, Zhuo AX, Dunn EW. Why do people turn to smartphones during social interactions? J Exp Soc Psychol 2023 Nov;109:104506. [doi: 10.1016/j.jesp.2023.104506]
2. Verduyn P, Schulte-Strathaus JCC, Kross E, Hülsheger UR. When do smartphones displace face-to-face interactions and what to do about it? Comput Human Behav 2021 Jan;114:106550. [doi: 10.1016/j.chb.2020.106550]
3. From loneliness to social connection: charting a path to healthier societies. World Health Organization. 2025. URL: https://www.who.int/publications/i/item/978240112360 [accessed 2026-07-07]
4. House JS, Landis KR, Umberson D. Social relationships and health. Science 1988 Jul 29;241(4865):540-545. [doi: 10.1126/science.3399889] [Medline: 3399889]
5. Holt-Lunstad J, Perissinotto C. Social isolation and loneliness as medical issues. N Engl J Med 2023 Jan 19;388(3):193-195. [doi: 10.1056/NEJMp2208029] [Medline: 36648080]
6. Poscia A, Stojanovic J, La Milia DI, et al. Interventions targeting loneliness and social isolation among the older people: an update systematic review. Exp Gerontol 2018 Feb;102:133-144. [doi: 10.1016/j.exger.2017.11.017] [Medline: 29199121]
7. Decade of healthy ageing: baseline report. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/9789240017900 [accessed 2026-07-07]
8. Lyu C, Siu K, Xu I, Osman I, Zhong J. Social isolation changes and long-term outcomes among older adults. JAMA Netw Open 2024 Jul 1;7(7):e2424519. [doi: 10.1001/jamanetworkopen.2024.24519] [Medline: 39046736]
9. Liang YY, Chen Y, Feng H, et al. Social isolation, loneliness and subsequent risk of major adverse cardiovascular events among individuals with type 2 diabetes mellitus. Gen Psychiatr 2023;36(6):e101153. [doi: 10.1136/gpsych-2023-101153] [Medline: 38170087]
10. Wong ELY, Qiu H, Cheung AWL, Leung HHW, Chen FY, Yeoh EK. Association of social isolation with health status among community-dwelling Chinese older adults living with homecare services: a cross-sectional survey in Hong Kong. Front Public Health 2023;11(1099734):1099734. [doi: 10.3389/fpubh.2023.1099734] [Medline: 37333567]
11. Lai-Yi Wong E, Qiu H, Ho KF, et al. Association of ambient temperature with social isolation among the community-dwelling Chinese older adults: a cross-sectional study in Hong Kong. Heliyon 2025 Jan 15;11(1):e41721. [doi: 10.1016/j.heliyon.2025.e41721] [Medline: 39866493]
12. Hao HJ, Du MQ, Yue JY. Social isolation, loneliness, and subjective wellbeing among Chinese older adults during the COVID-19 pandemic. Front Public Health 2024;12:1425575. [doi: 10.3389/fpubh.2024.1425575] [Medline: 39440171]
13. Wong ELY, Li J, Yuen S, et al. Vulnerable populations during COVID-19 response: health-related quality of life among Chinese population and its influence due to socio-demographic factors and loneliness. Front Public Health 2022;10:857033. [doi: 10.3389/fpubh.2022.857033] [Medline: 36081475]
14. Phang JK, Kwan YH, Yoon S, et al. Digital intergenerational program to reduce loneliness and social isolation among older adults: realist review. JMIR Aging 2023 Jan 4;6:e39848. [doi: 10.2196/39848] [Medline: 36598801]
15. Holloway J, James DL, Robillard A, Hermer J, Hawley N, Sayeed O. Needs of social isolation, loneliness, and intergenerational interventions in the United States: a scoping review. Front Public Health 2024;12:1386651. [doi: 10.3389/fpubh.2024.1386651] [Medline: 39185115]
16. Welch V, Ghogomu ET, Barbeau VI, et al. Digital interventions to reduce social isolation and loneliness in older adults: an evidence and gap map. Campbell Syst Rev 2023 Dec;19(4):e1369. [doi: 10.1002/cl2.1369] [Medline: 38024780]
17. Lin MC. Intergenerational communication. In: Allen M, editor. The SAGE Encyclopedia of Communication Research Methods: SAGE Publications, Inc; 2017:762-766. [doi: 10.4135/9781483381411.n264]
18. Addimando F. Effective communication strategies. In: Trade Show Psychology: Springer Nature Switzerland; 2024:53-72. [doi: 10.1007/978-3-031-53606-9_4]
19. Arnold E, Boggs KU. Interpersonal Relationships: Professional Communication Skills for Nurses: Elsevier; 2020.
20. Giles H, Gasiorek J. Intergenerational communication practices. In: Schaie KW, Willis SL, editors. Handbook of the Psychology of Aging, 7th edition: Elsevier Academic Press; 2011:233-247. [doi: 10.1016/B978-0-12-380882-0.00015-2]
21. Söllner M, Dürnberger M, Keller J, Florack A. The impact of age stereotypes on well-being: strategies of selection, optimization, and compensation as mediator and regulatory focus as moderator: findings from a cross-sectional and a longitudinal study. J Happiness Stud 2022 Feb;23(2):635-665. [doi: 10.1007/s10902-021-00417-x]

XSL•FO
RenderX

22. Landers RN, Bauer KN, Callan RC. Gamification of task performance with leaderboards: a goal setting experiment. Comput Human Behav 2017 Jun;71:508-515. [doi: 10.1016/j.chb.2015.08.008]
23. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30:n2061. [doi: 10.1136/bmj.n2061]
24. Tustonja M, Topić Stipić D, Skoko I, Čuljak A, Vegar A. Active listening – a model of empathetic communication in the helping professions. Med Acad Integr 2024 Mar;1(1):42-47. [doi: 10.47960/3029-3316.2024.1.1.42]
25. Abargil M, Schenkolewski A, Tishby O. Therapists' emotional responses and their relation to patients' experience of attunement and responsiveness. Psychother Res 2025 Jan;35(1):54-66. [doi: 10.1080/10503307.2024.2403422] [Medline: 39508263]
26. Reniers R, Abu-Akel A, Seara-Cardoso A. Editorial: cognitive empathy and perspective taking: understanding the mechanisms of normal and abnormal experiences and abilities. Front Psychiatry 2022;13:945258. [doi: 10.3389/fpsyt.2022.945258] [Medline: 35722573]
27. McCann RM, Giles H, Ota H. Aging and communication across cultures. In: Chen L, editor. Intercultural Communication: De Gruyter Mouton; 2017:289-307 URL: https://www.researchgate.net/publication/316067070_13_Aging_and_communication_across_cultures [accessed 2026-07-07]
28. Giles H, Dorjee T. Communicative climates and prospects in cross-cultural gerontology. J Cross Cult Gerontol 2004 Dec;19(4):261-274. [doi: 10.1023/B:JCCG.0000044684.82186.e3] [Medline: 15604644]
29. Shortt M, Tilak S, Kuznetcova I, Martens B, Akinkuolie B. Gamification in mobile-assisted language learning: a systematic review of Duolingo literature from public release of 2012 to early 2020. Comput Assist Lang Learn 2023 Mar 4;36(3):517-554. [doi: 10.1080/09588221.2021.1933540]
30. Malamed C. The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education: Pfeiffer; 2012. [doi: 10.1145/2207270.2211316]
31. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". Presented at: MindTrek '11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; Sep 28-30, 2011. [doi: 10.1145/2181037.2181040]
32. Talley MH, Ogle N, Wingo N, Roche C, Willig J. Kaizen: interactive gaming for diabetes patient education. Games Health J 2019 Dec;8(6):423-431. [doi: 10.1089/g4h.2018.0107] [Medline: 31769723]
33. Bicen H, Kocakoyun S. Perceptions of students for gamification approach: Kahoot as a case study. Int J Emerg Technol Learn 2018;13(2):72. [doi: 10.3991/ijet.v13i02.7467]
34. Kadam P, Bhalerao S. Sample size calculation. Int J Ayurveda Res 2010 Jan;1(1):55-57. [doi: 10.4103/0974-7788.59946] [Medline: 20532100]
35. Hecht ML. The conceptualization and measurement of interpersonal communication satisfaction. Human Comm Res 1978 Mar;4(3):253-264. [doi: 10.1111/j.1468-2958.1978.tb00614.x]
36. Rubin RB, Martin MM. Development of a measure of interpersonal communication competence. Commun Res Rep 1994 Jun;11(1):33-44. [doi: 10.1080/08824099409359938]
37. Sørensen K, Van den Broucke S, Pelikan JM, et al. Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). BMC Public Health 2013 Oct 10;13(1):948. [doi: 10.1186/1471-2458-13-948] [Medline: 24112855]
38. Zhang M, Lu Z, Song C, Ling Y, Li W, Zuo C. Translation and cross-cultural adaptation of the Chinese version of the HLS-EU-Q47. Health Promot Int 2024 Aug 1;39(4):daae083. [doi: 10.1093/heapro/daae083] [Medline: 39058422]
39. Lam CLK, Yuen NYK, Mercer SW, Wong W. A pilot study on the validity and reliability of the Patient Enablement Instrument (PEI) in a Chinese population. Fam Pract 2010 Aug;27(4):395-403. [doi: 10.1093/fampra/cmq021] [Medline: 20435665]
40. Caserman P, Hoffmann K, Müller P, et al. Quality criteria for serious games: serious part, game part, and balance. JMIR Serious Games 2020 Jul 24;8(3):e19037. [doi: 10.2196/19037] [Medline: 32706669]
41. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 2013 Nov 26;4:863. [doi: 10.3389/fpsyg.2013.00863] [Medline: 24324449]
42. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Routledge; 2013:567 URL: https://utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf [accessed 2026-07-07]
43. Koivisto J, Malik A. Gamification for older adults: a systematic literature review. Gerontologist 2021 Sep 13;61(7):e360-e372. [doi: 10.1093/geront/gnaa047] [Medline: 32530026]
44. Information technology usage and penetration. : Census and Statistics Department, The Government of the Hong Kong Special Administrative Region; 2025 URL: https://www.censtatd.gov.hk/en/EIndexbySubject.html?scode=590&pcode=C0000031 [accessed 2026-07-07]
45. Peters R, Ee N, Ward SA, et al. Intergenerational programmes bringing together community dwelling non-familial older adults and children: a systematic review. Arch Gerontol Geriatr 2021;94:104356. [doi: 10.1016/j.archger.2021.104356] [Medline: 33567363]
46. Miura H, Kariyasu M, Yamasaki K, Sumi Y. Physical, mental and social factors affecting self‐rated verbal communication among elderly individuals. Geriatrics Gerontology Int 2004 Jun;4(2):100-104. [doi: 10.1111/j.1447-0594.2004.00125.x]

XSL•FO
RenderX

47. Gauthier A, Kato PM, Bul KCM, Dunwell I, Walker-Clarke A, Lameras P. Board games for health: a systematic literature review and meta-analysis. Games Health J 2019 Apr;8(2):85-100. [doi: 10.1089/g4h.2018.0017] [Medline: 30256159]
48. Corrigan T, Mcnamara G, O'Hara J. Intergenerational learning: a valuable learning experience for higher education students. Eurasian J Educ Res 2013;52:117-136 [FREE Full text]
49. Carstensen LL, Hershfield HE. Beyond stereotypes: using socioemotional selectivity theory to improve messaging to older adults. Curr Dir Psychol Sci 2021 Aug 1;30(4):327-334. [doi: 10.1177/09637214211011468] [Medline: 34366582]
50. White BK, Martin A, White J. Gamification and older adults: opportunities for gamification to support health promotion initiatives for older adults in the context of COVID-19. Lancet Reg Health West Pac 2023 Jun;35:100528. [doi: 10.1016/j.lanwpc.2022.100528] [Medline: 35815240]
51. Guardabassi V, Manoni E, Di Massimo M, Cirilli E, Maranesi A, Nicolini P. Aging with board games: fostering well-being in the older population. Front Psychol 2024;15(1501111):1501111. [doi: 10.3389/fpsyg.2024.1501111] [Medline: 39624370]
52. Sailer M, Hense JU, Mayr SK, Mandl H. How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction. Comput Human Behav 2017 Apr;69:371-380. [doi: 10.1016/j.chb.2016.12.033]

Abbreviations

HLS-EU-Q47: European Health Literacy Survey Questionnaire
ICCS: Interpersonal Communication Competence Scale
LOUE: Listening, Observing, Understanding, and Expression
OR: odds ratio
PEI: Patient Enablement Instrument
WHO: World Health Organization

Edited by Y Jiang; submitted 09.May.2025; peer-reviewed by J Janssen, M Jordan-Marsh, T Lima; revised version received 29.May.2026; accepted 04.Jun.2026; published 31.Jul.2026.

Please cite as:
Wong ELY, Ng CCW, Wang DY, Wong CKP, Lui SSK, Cheung AWL, Yeoh EK
Evaluation of a Gamification-Based Intervention to Enhance Intergenerational Communication: Pre-Post Mixed Methods Evaluation
JMIR Aging 2026;9:e77190
URL: https://aging.jmir.org/2026/1/e77190
doi:10.2196/77190

© Eliza Lai-Yi Wong, Clement Cheuk Wai Ng, Dorothy Yingxuan Wang, Carol Ka Po Wong, Shirley Shuk Kuen Lui, Annie Wai-Ling Cheung, Eng-kiong Yeoh. Originally published in JMIR Aging (https://aging.jmir.org), 31.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Influence of Robot Facial Expression and Inclusion Behavior on Rapport and Trust in Older Adults: Mixed-Factorial Experimental Study

Hsiu-Ping Yueh[1*], PhD; Kuang-Chi Lin[2*], MS; Weijane Lin[3*], PhD

[1]Department of Psychology/Bio-Industry Communication & Development/Mechanical Engineering, Research Center for Digital Humanities, National Taiwan University, Taipei, Taiwan

[2]Research Center for Digital Humanities, National Taiwan University, Taipei, Taiwan

[3]Department of Library and Information Science, Research Center for Digital Humanities, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Daan Dist., Taipei, Taiwan, Taiwan

[*]all authors contributed equally

Corresponding Author:
Weijane Lin, PhD
Department of Library and Information Science, Research Center for Digital Humanities, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Daan Dist., Taipei, Taiwan, Taiwan

Abstract

Background: The World Health Organization framework for healthy aging emphasizes that the capacity to establish and sustain relationships is a vital component of functional ability. Social robots offer valuable support for this relational aspect, although their effectiveness is contingent upon the quality of the interactions. While it is established that facial expressions and inclusive behaviors can influence rapport building, the combined effects of these elements on older adults remain unexamined.

Objective: This study aimed to investigate the main and interaction effects of robot facial expression (happy vs unhappy) and interaction style (inclusive vs exclusive) on the development of rapport between social robots and older adults residing in the community.

Methods: A 2 × 2 mixed-factorial experimental design was used in this study, involving 53 community-dwelling older adults (mean age 70.57, SD 4.27 years; n=45, 84.9% female). Participants engaged in a gamified interaction known as Radish Squat featuring 2 humanoid robots that exhibited various combinations of facial expressions and interaction styles. The inclusive robot designated participants within the game 75% of the time, whereas the exclusive robot did so approximately 25% of the time. Dependent measures included robot acceptance, rapport, rapport expectation, and trust. Behavioral designation frequencies were recorded throughout the gameplay. A mixed-design repeated-measure ANOVA was conducted to examine both main and interaction effects.

Results: Facial expression exhibited significant main effects on acceptance ($F_{1,51}$=14.27; P<.001; η^2=0.22), rapport ($F_{1,51}$=6.69; P=.01; η^2=0.12), conversation partner expectations ($F_{1,51}$=11.34; P=.001; η^2=0.18), togetherness expectations ($F_{1,51}$=10.28; P=.002; η^2=0.17), and trust ($F_{1,51}$=4.36; P=.04; η^2=0.08). Interaction style significantly affected rapport ($F_{1,51}$=6.22; P=.02; η^2=0.11), as well as both rapport expectation subscales. Significant interaction effects were also observed for rapport ($F_{1,51}$=8.44; P=.005; η^2=0.14) and trust ($F_{1,51}$=16.62; P=.002; η^2=0.25) such that the positive effect of happy facial expressions was substantially larger under exclusive interaction conditions. Older adults designated the exclusive robot within the game more frequently than the inclusive robot (mean 6.42, SD 1.84 vs 5.32, SD 2.12; t_{52}=2.18; P=.03), and qualitative data pointed to the use of fairness-based turn-taking strategies.

Conclusions: Positive facial expressions and inclusive interaction behaviors significantly enhanced rapport and trust, with synergistic interaction effects primarily emerging under exclusive interaction conditions. The observed patterns, whereby older adults compensated for exclusive robot behavior, suggest that they perceive robots as social entities deserving of equitable treatment. Collectively, these findings offer evidence-based guidelines for designing social robots that foster the relational dimensions of healthy aging.

(JMIR Aging 2026;9:e91235) doi:10.2196/91235

KEYWORDS

social robots; human-robot interaction; rapport; facial expression; healthy aging

XSL•FO
RenderX

Introduction

Background

The aging global population necessitates a fundamental shift in our approach to supporting the well-being of older adults. By 2030, it is projected that 1 in 6 individuals worldwide will be aged 60 years or older, and this number is projected to reach 2.1 billion by 2050. In most high-income nations, the age of 65 years is widely accepted as the conventional threshold for retirement and is commonly adopted as a criterion for defining the older adult demographic [1-3]. Accordingly, the World Health Organization has defined healthy aging not simply as the absence of disease but as "the process of developing and maintaining the functional ability that enables well-being in older age" [4-6]. This functional ability, comprising what individuals value in terms of being and doing, emerges from the dynamic interplay between their intrinsic capacity and environmental factors [4,5]. Of the 5 essential domains of functional ability, the capacity to build and maintain relationships is particularly important as social connections profoundly impact stress regulation, cognitive function, and overall well-being [1,6]. However, aging often precipitates social network contraction due to factors such as spousal loss, family dispersion, and mobility limitations. Notably, social isolation has been identified as a mortality risk comparable to smoking [7]. This relational deficit underscores the urgent need for environmental interventions that offer accessible and consistent opportunities for positive social interaction.

Social robots represent a promising environmental resource for addressing this relational dimension of healthy aging. Meta-analytic evidence demonstrates that robot interventions can effectively reduce loneliness and improve psychological well-being among older adults [8,9]. However, the successful integration of social robots into older adults' lives depends critically on interaction quality rather than mere functional utility [10,11]. The computers are social actors (CASA) paradigm establishes that humans unconsciously apply social rules to technological agents, expecting adherence to norms of politeness, turn taking, and emotional congruence [12,13]. For older adults with deeply established social schemas, these expectations are particularly pronounced [11,14]. When robots fail to meet social expectations, the resulting dissonance undermines potential benefits. Therefore, social robots must be conceptualized not as passive tools but as active environmental components whose design features determine whether they enable or hinder the relational aspects of functional ability [14,15].

Rapport, a harmonious relationship characterized by mutual attentiveness, positivity, and coordination [16], emerges as the critical mechanism linking robot design to healthy aging outcomes. When rapport is established, older adults experience enhanced engagement, increased motivation for continued interaction, and greater trust in robot capabilities [17,18]. Research has identified nonverbal cues as primary drivers of rapport formation, with facial expressions serving as powerful signals of social intent [19,20]. Additionally, social inclusion behaviors constitute a fundamental dimension: studies adapting the Cyberball paradigm demonstrate that humans respond to robot-mediated inclusion and exclusion similarly to human interactions [21,22]. However, a critical gap remains. While recent research with university students revealed complex patterns when facial expressions and interaction styles were manipulated simultaneously [23], it remains uncertain how older adults process these potentially conflicting cues. An incongruent robot, one that displays positive expressions while behaving in an exclusive manner, may generate mistrust rather than rapport. Given the distinct generational experiences and varying intrinsic capabilities of older adults [24], understanding the combined effects of these rapport-related characteristics is essential for designing robots that genuinely support healthy aging.

Study Objectives

Driven by the previously mentioned challenges, this study examined the impact of robot facial expressions (happy vs unhappy) and interaction styles (inclusive vs exclusive) on rapport building with community-dwelling older adults (aged 65 years or older). Using a gamified interaction paradigm, we focused on (1) the main effects of facial expressions and interaction styles on acceptance, rapport, rapport expectations, and trust; (2) the interaction effects of these 2 characteristics; and (3) the behavioral patterns exhibited by older adults during gamified human-robot interaction (HRI). By analyzing how older adults respond to both congruent and incongruent combinations of robot social cues, this research offers evidence-based guidelines for designing social robots that effectively support healthy aging.

Methods

Study Design

This study used a mixed-factorial experimental design with a 2 (facial expression: happy vs unhappy) × 2 (interaction style: inclusive vs exclusive) structure. Experimental condition (congruent vs incongruent assignment) served as the between-subject factor, whereas facial expression and interaction style were manipulated within participants. During task 2, each participant interacted with both robots, thereby enabling direct within-subject comparisons. This design allowed for the examination of both main effects and interaction effects. Participants were randomly assigned to 1 of 2 experimental conditions.

In the congruent condition (group 1), robot 1 exhibited a happy facial expression and inclusive interaction behavior, whereas robot 2 exhibited an unhappy facial expression and exclusive interaction behavior. In the incongruent condition (group 2), robot 1 exhibited an unhappy facial expression with inclusive interaction behavior, whereas robot 2 exhibited a happy facial expression with exclusive interaction behavior.

Ethical Considerations

This study was reviewed and approved by the Research Ethics Committee of National Taiwan University (202305HS096). All participants provided informed consent before taking part. All collected data were deidentified and stored securely. Participants received NTD 200 (NTD $1=US $0.03 as of July 6, 2026) as compensation for their participation. Upon completion of all

XSL•FO
RenderX

measures and the postexperiment interview, participants received a full debriefing that explained the experimental manipulation and clarified that the exclusive robot's behavior was experimentally scripted rather than contingent on the participants.

Participants

Community-dwelling older adults were recruited through volunteer networks at the university and through snowball sampling. The following inclusion criteria were established for participants: (1) age of 65 years or above, (2) normal cognitive function and hearing ability, (3) fluency in Mandarin to facilitate verbal interactions with robots, (4) normal or corrected vision adequate for perceiving robot facial expressions at a conversational distance, and (5) sufficient upper-body mobility to engage in game actions.

The sample comprised 53 participants, whose characteristics and experiences with robot interaction are detailed in Table 1. The average age of the participants was 70.57 (SD 4.27; range 65-85) years. The substantial majority (n=45, 84.9% of the participants) identified as female. Educational attainment varied among the group, with the largest proportion holding university degrees (n=22, 41.5%). Furthermore, most participants reported limited experience with robots, interacting with them less than once in the previous 6 months (n=44, 83%). Independent-sample 2-tailed *t* tests and chi-square tests indicated no significant differences between group 1 and group 2 on any baseline characteristics ($P>.10$ in all cases), supporting the success of the randomization procedure.

XSL•FO
RenderX

Table . Demographics and experiences of the participants (N=53).

Characteristic	Total	Group 1[a] (n=27)	Group 2[b] (n=26)
Age (years), mean (SD)	70.57 (4.27)	70.67 (4.79)	70.46 (3.76)
Age range (years)	65-85	65-85	65-79
Age group (years), n (%)			
65-69	24 (45.3)	13 (48.1)	11 (42.3)
≥70	29 (54.7)	14 (51.9)	15 (57.7)
Sex, n (%)			
Female	45 (84.9)	23 (85.2)	22 (84.6)
Male	8 (15.1)	4 (14.8)	4 (15.4)
Educational level, n (%)			
Junior high school	3 (5.7)	2 (7.4)	1 (3.8)
High school	18 (34)	8 (29.6)	10 (38.5)
University or college	22 (41.5)	13 (48.1)	9 (34.6)
Graduate school	10 (18.9)	4 (14.8)	6 (23.1)
Personality, mean (SD)			
Extraversion (TIPI[c] [2-14])	8.42 (2.26)	8.48 (2.31)	8.35 (2.24)
Empathy, mean (SD)			
Cognitive empathy (12-29)	21.04 (7.10)	21.07 (9.07)	21 (4.41)
Emotional contagion (22-40)	28.96 (3.31)	28.22 (2.93)	29.73 (3.55)
Robot contact frequency			
Less than once in the previous 6 mo, n (%)	44 (83)	25 (92.6)	19 (73.1)
Robot acceptance score[d], mean (SD)	35.06 (3.93)	34.04 (4.31)	36.12 (3.24)
Robot attitude, mean (SD)			
NARS[e] score	34.09 (9.80)	35.26 (9.88)	32.88 (9.77)
HRI[f] belief, mean (SD)			
SE-HRI[g] score	46.42 (7.53)	45.85 (6.60)	47 (8.48)

[a]Congruent condition (expression matched behavior).

[b]Incongruent condition (expression was mismatched with behavior).

[c]TIPI: Ten-Item Personality Inventory.

[d]Robot acceptance measured as total score (range 8-40).

[e]NARS: Negative Attitudes Toward Robots Scale (range 14-78; higher scores indicate more negative attitudes).

[f]HRI: human-robot interaction.

[g]SE-HRI: Self-Efficacy in Human-Robot Interaction Scale (range 10-60; higher scores indicate greater self-efficacy).

Apparatus

The study used Kebbi humanoid robots produced by NUWA Robotics (Figure 1). Each Kebbi robot stands at approximately 35 cm tall and is equipped with a wheeled base for mobility, articulated arms and elbows, and a touch screen face capable of conveying a wide range of emotions. In total, 3 Kebbi robots were used: 1 served as the neutral expression model for practice (designated as R0), whereas the other 2 participated in the main experimental task (referred to as R1 and R2). The robots were strategically positioned on a tabletop to ensure they were at eye level with the seated participants. Their behaviors were manipulated through the Wizard of Oz technique combined with the Message Queuing Telemetry Transport protocol to enable semiautonomous communication between the human users and the 2 robots.

XSL•FO
RenderX

Figure 1. Kebbi robot used in this study.



Experimental Task: Radish Squat Game

The experimental task was an adapted version of the Radish Squat game, which closely resembles the Cyberball paradigm used in research on social exclusion [22,23]. In this game, players take turns executing specific actions while verbally stating their identity and designating the next player. Actions are categorized into 2 groups: basic actions (raising a hand, nodding, and shaking the head) and advanced actions (flapping wings, hugging, and drumming). This format facilitates natural, repeated interactions while systematically controlling the dynamics of inclusion and exclusion.

Manipulation of Independent Variables

Facial Expression

The robot's facial expressions were developed based on the elements associated with rapport as outlined by Wang and Gratch [20]. As depicted in Figure 2, the happy expression featured upturned corners of the mouth and a lifted chin, whereas the unhappy expression was characterized by furrowed brows, a wrinkled nose, and a raised upper lip. In the practice session, a neutral expression was used.

XSL•FO
RenderX

Figure 2. Robot facial expression designs.



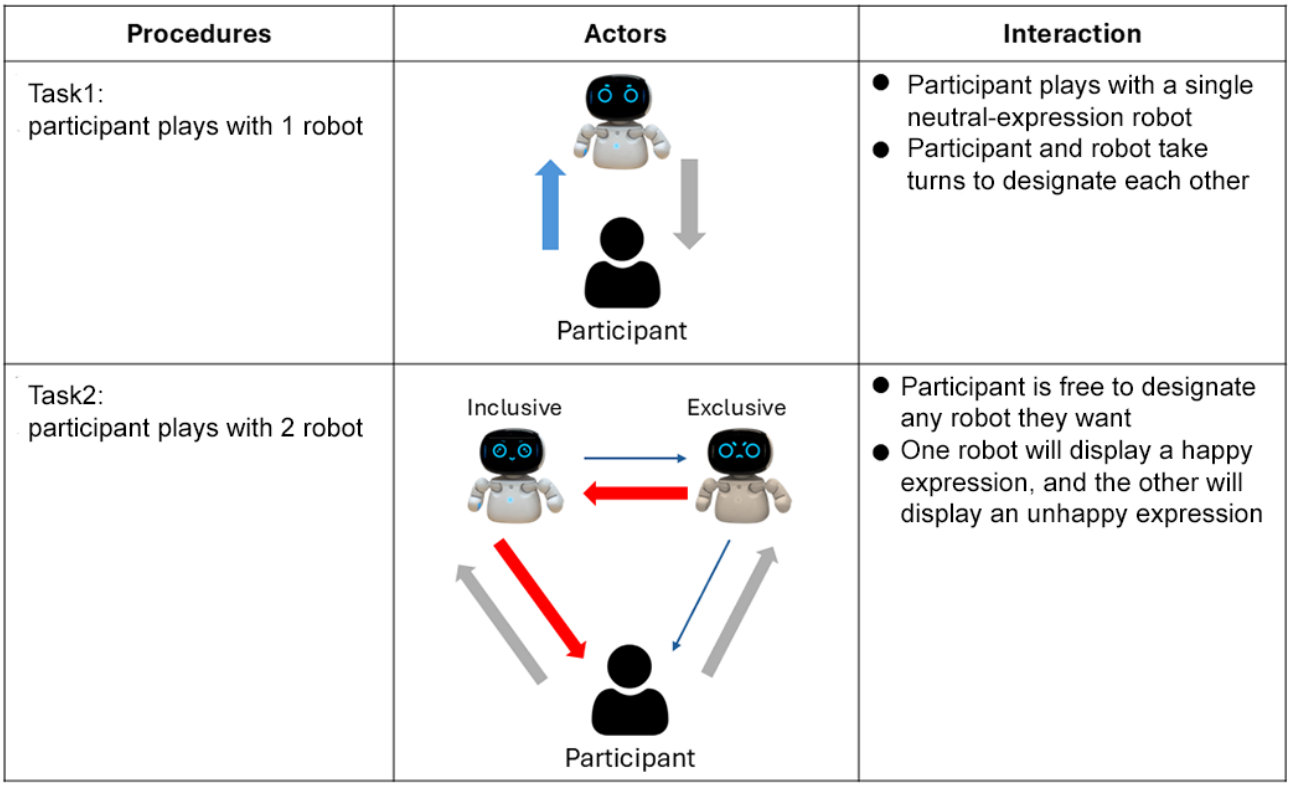
Interaction Style

In accordance with established protocols [21,23], the inclusive robot designated participants 75% of the time while allocating 25% of the time to interaction with the other robot. Conversely, the exclusive robot designated participants only 25% of the time, dedicating the remaining 75% of the time to interaction with the other robot.

Procedure

The experiment consisted of 2 tasks (Figure 3). In task 1 (practice), participants engaged in the Radish Squat game with a single robot (R0) displaying a neutral facial expression for 12 turns to familiarize themselves with the task rules. In task 2 (main session), participants interacted simultaneously with 2 robots (R1 and R2), each exhibiting a distinct combination of facial expression and interaction style, over a total of 30 turns. The task 2 Radish Squat session lasted for an average of 12.4 (SD 2.1; range 9-17) minutes. Following task 2, participants completed postexperimental questionnaires and took part in brief semistructured interviews. The entire experimental session lasted approximately 45 minutes.

Figure 3. Experimental procedure flowchart.



Measures

Four dependent variables were assessed following interaction with each robot: robot acceptance, rapport, rapport expectation, and trust. Robot acceptance was measured using 8 items based on the technology acceptance model ([23,25]; item score range 1-5; higher scores indicate greater acceptance). Rapport was assessed using a 9-item rapport scale ([17]; Cronbach α=0.85; item score range 1-5; higher scores indicate greater rapport). Rapport expectation was measured using the 18-item Rapport-Expectation With a Robot Scale [26], comprising 2 subscales: expectation as a conversation partner (11 items; Cronbach α=0.85) and expectation for togetherness (7 items; Cronbach α=0.82; item score range 1-7; higher scores indicate greater expectation). Trust was assessed using the 15-item Trust Perception Scale–HRI ([27]; item score range 1-6; higher scores indicate greater trust). Behavioral data were collected as the number of times participants designated each robot during gameplay. Baseline measures included extraversion (Ten-Item

XSL•FO
RenderX

Personality Inventory [28]), empathy (Basic Empathy Scale [29]), negative attitudes toward robots (Negative Attitudes Toward Robots Scale [30]), and self-efficacy in HRI (Self-Efficacy in Human-Robot Interaction Scale [31]). Brief semistructured interviews were conducted after the experiment to gather qualitative insights regarding participants' interaction experiences and observations of robot behaviors. Table 2 summarizes all measures used in this study.

Table . Summary of measures.

Category and measure	Items, n	Scale	Reference
Dependent			
Robot acceptance	8	5-point	[23,25]
Rapport	9	5-point	[17]
Rapport expectation—conversation partner	11	7-point	[26]
Rapport expectation—togetherness	7	7-point	[26]
Trust	15	6-point	[27]
Designation frequency	—[a]	—	[22,23]
Baseline			
Extraversion (TIPI[b])	2	7-point	[28]
Basic empathy	20	5-point	[29]
NARS[c]	14	5-point	[30]
SE-HRI[d]	10	6-point	[31]
Qualitative			
Postinteraction interview	—	—	[22,23]

[a]Not applicable.

[b]TIPI: Ten-Item Personality Inventory.

[c]NARS: Negative Attitudes Toward Robots Scale.

[d]SE-HRI: Self-Efficacy in Human-Robot Interaction Scale.

Statistical Analysis

A mixed-design repeated-measure ANOVA was conducted with facial expression (happy vs unhappy) and interaction style (inclusive vs exclusive) as within-subject factors and experimental condition (congruent vs incongruent) as the between-subject factor. Both main effects and interaction effects were analyzed for each dependent variable. Effect sizes were reported as partial η^2 and interpreted using the benchmarks by Cohen ([32]; small: $\eta^2 \geq 0.01$; medium: $\eta^2 \geq 0.06$; large: $\eta^2 \geq 0.1$). Post hoc power analyses ($\alpha = .05$) indicated adequate power for the detected interaction effects: for trust ($\eta^2 = 0.25$), achieved power was approximately 0.97; for rapport ($\eta^2 = 0.14$), power was approximately 0.83. Post hoc comparisons were performed using the Bonferroni correction. Additionally, paired 2-tailed *t* tests were used to compare the frequencies of behavioral designations between the robots ($t_{52} = 2.18$; $P = .03$).

Results

Manipulation Check

To conduct the postexperimental manipulation checks, each participant was interviewed to verify the successful implementation of both independent variables. With respect to facial expressions, chi-square tests indicated that participants correctly identified the happy-expression robot as displaying a happy expression at rates significantly above chance (N=53, $\chi^2_1 = 38.4$; $P < .001$) and the unhappy-expression robot as displaying an unhappy expression at rates significantly above chance (N=53, $\chi^2_1 = 31.7$; $P < .001$). Additionally, as shown in Table 3, the inclusive robot engaged participants significantly more frequently than the exclusive robot, further validating the behavioral manipulation.

XSL•FO
RenderX

Table . Frequency of robot designation by interaction style: manipulation check[a].

Robot type	Designations[b], mean (SD)
Inclusive	8.77 (1.73)
Exclusive	2.79 (1.79)

[a]t_{52}=24.31; P<.001.

[b]Designations refer to the number of times each robot designated the participants during task 2.

Descriptive Statistics

Table 4 presents descriptive statistics for all dependent variables across the 4 within-subject experimental conditions. The patterns were more nuanced than a simple summary of main effects would suggest. For acceptance, robots with happy expressions received higher ratings than those with unhappy expressions across both interaction styles, aligned with their high robot acceptance and HRI belief (Table 1). For rapport and trust, however, the advantage of happy expressions was concentrated primarily in the exclusive interaction condition, with only minimal expression-related differences under inclusive conditions. These patterns are elaborated on in the interaction analysis reported below.

Table . Mean scores of dependent variables by experimental condition.

Variable	Inclusive, mean (SD)	Exclusive, mean (SD)	Overall, mean (SD)
Acceptance (1-5)			
Happy expression	4.25 (0.54)	4.65 (0.41)	4.45 (0.52)
Unhappy expression	4.17 (0.65)	4.08 (0.74)	4.12 (0.689)
Rapport (1-5)			
Happy expression	3.74 (0.62)	4.44 (0.49)	4.08 (0.66)
Unhappy expression	3.88 (0.79)	3.77 (0.85)	3.82 (0.81)
Expectations as a conversation partner (1-7)			
Happy expression	4.33 (0.56)	4.85 (0.77)	4.58 (0.72)
Unhappy expression	4.41 (0.66)	4.01 (0.85)	4.20 (0.79)
Expectations for togetherness (1-7)			
Happy expression	4.81 (0.67)	5.65 (0.78)	5.22 (0.83)
Unhappy expression	5.07 (0.90)	4.54 (1.20)	4.80 (1.09)
Trust (1-6)			
Happy expression	4.48 (0.65)	5.13 (0.47)	4.79 (0.65)
Unhappy expression	4.55 (0.72)	4.66 (0.77)	4.61 (0.74)

Main Effects and Interaction Effects

Overview

Table 5 presents the results of the repeated-measure ANOVA. Facial expression demonstrated significant main effects across all 5 outcome measures, with effect sizes ranging from small to large. Interaction style significantly influenced rapport and both rapport expectation subscales. Significant interaction effects were observed for rapport and trust (both large effects), indicating that the combined influence of facial expression and interaction style was more complex than a simple additive pattern.

XSL•FO
RenderX

Table . Results of repeated-measure ANOVA: main effects and interaction effects[a].

Variable	Facial expression			Interaction style			F × I[b]	
	F test (*df*)	*P* value	η^2	*F* test (*df*)	*P* value	η^2	*F* test (*df*)	η^2
Acceptance	14.27 (1, 51)	<.001	0.22	3.18 (1, 51)	.08	0.06	3.41 (1, 51)	0.06
Rapport	6.69 (1, 51)	.01	0.12	6.22 (1, 51)	.02	0.11	8.44 (1, 51)[c]	0.14
Expectations as a conversation partner	11.34 (1, 51)	.001	0.18	7.99 (1, 51)	.007	0.14	0.29 (1, 51)	0.01
Expectations for togetherness	10.28 (1, 51)	.002	0.17	10.47 (1, 51)	.002	0.16	0.41 (1, 51)	0.01
Trust	4.36 (1, 51)	.04	0.08	2.91 (1, 51)	.10	0.05	16.62 (1, 51)[c]	0.25

[a]Effect size benchmarks [32]: $\eta^2 \geq 0.01$ for small, $\eta^2 \geq 0.06$ for medium, and $\eta^2 \geq 0.1$ for large.

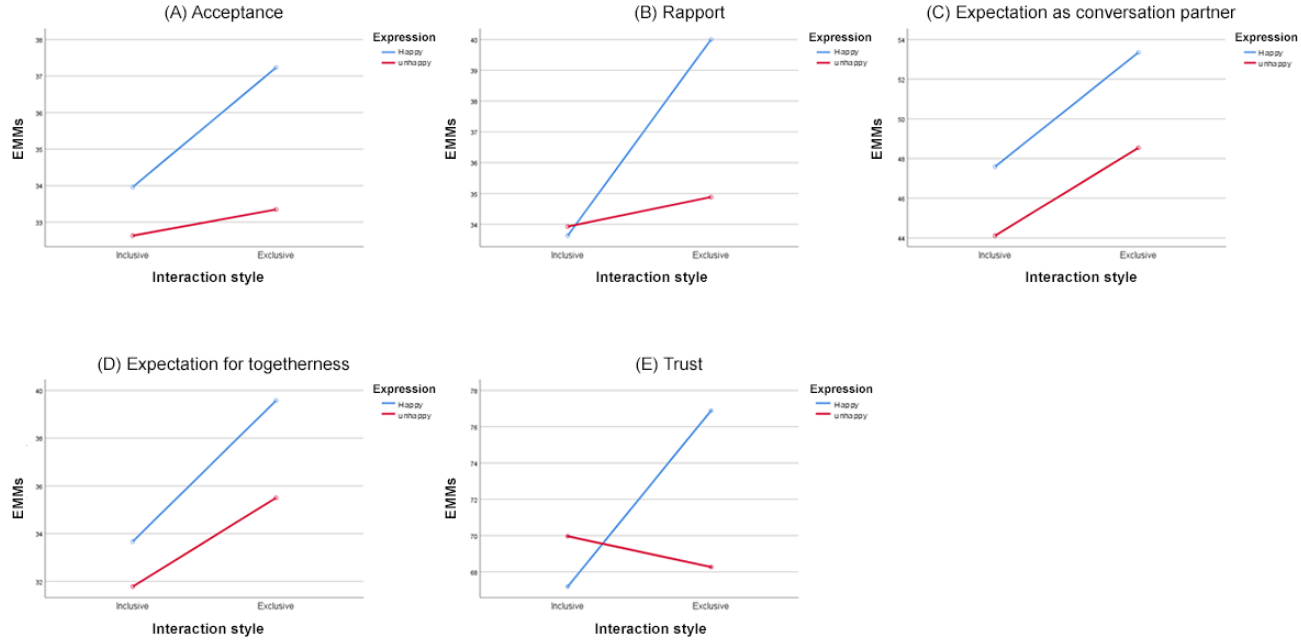
[b]Facial expression × interaction style.

[c]$P<.01$.

Acceptance

Facial expression had a significant main effect on acceptance ($F_{1,51}$=14.27; P<.001; η^2=0.22), with robots displaying happy expressions receiving higher ratings (mean 4.45, SD 0.52) than those displaying unhappy expressions (mean 4.12, SD 0.689). As shown in Figure 4A, the main effect of interaction style was not significant ($F_{1,51}$=3.18; P=.08; η^2=0.06), nor was the interaction between facial expression and interaction style ($F_{1,51}$=3.41; P=.07; η^2=0.06). In the absence of a significant interaction, the main effect of facial expression can therefore be interpreted directly.

Figure 4. Interaction effects of facial expression and interaction style. EMM: estimated marginal mean.



Rapport

For rapport, a significant interaction effect was observed ($F_{1,51}$=8.44; P=.005; η^2=0.14). In light of this interaction, the main effects were qualified and interpreted via simple effects. As shown in Figure 4A, 4B, and 4C, under exclusive interaction, the happy-expression robot generated substantially higher rapport (mean 4.44, SD 0.49) than the unhappy-expression robot (mean 3.77, SD 0.85). In contrast, under inclusive interaction, there was minimal difference between happy (mean 3.74, SD 0.62) and unhappy (mean 3.88, SD 0.79) expressions, and this difference was not significant following Bonferroni correction (P=.65). Thus, the advantage of happy facial expressions was concentrated in the exclusive condition. Accordingly, the overall main effect of facial expression ($F_{1,51}$=6.69; P=.01; η^2=0.12) and the main effect of interaction style ($F_{1,51}$=6.22; P=.02; η^2=0.11) should be understood as being driven by this asymmetry rather than uniform effects across interaction conditions.

XSL•FO
RenderX

Rapport Expectation

Regarding expectations for a conversation partner, both facial expression ($F_{1,51}$=11.34; P=.001; η^2=0.18) and interaction style ($F_{1,51}$=7.99; P=.007; η^2=0.14) exhibited significant main effects, whereas their interaction was not significant (Figure 4C; $F_{1,51}$=0.29; P=.59; η^2=0.01). As shown in Table 4, robots with happy expressions elicited higher conversation partner expectations (overall mean 4.58, SD 0.72) than those with unhappy expressions (mean 4.20, SD 0.79), and exclusive robots elicited higher expectations (marginal mean 4.43) than inclusive robots (marginal mean 4.37). In the absence of a significant interaction, these main effects can be interpreted directly.

For togetherness expectations, both facial expression ($F_{1,51}$=10.28; P=.002; η^2=0.17) and interaction style ($F_{1,51}$=10.47; P=.002; η^2=0.16) yielded significant main effects, with no significant interaction ($F_{1,51}$=0.41; P=.52; η^2=0.01). Robots with happy expressions elicited higher togetherness expectations (overall mean 5.22, SD 0.83) than those with unhappy expressions (mean 4.80, SD 1.09). As shown in Figure 4D, although inclusive robots tended to elicit lower togetherness expectations than exclusive robots within the happy-expression condition, this pattern did not reach the level of significant interaction.

Trust

For trust, a significant interaction effect was observed ($F_{1,51}$=16.62; P=.002; η^2=0.25). As shown in Figure 4E, under exclusive interaction, the happy-expression robot received substantially higher trust ratings (mean 5.13, SD 0.47). Under inclusive interaction, the pattern was slightly reversed: trust ratings for the happy-expression (mean 4.48, SD 0.65) and unhappy-expression (mean 4.55, SD 0.72) robots were essentially equivalent, and this difference was not significant following Bonferroni correction. The significant main effect of facial expression ($F_{1,51}$=4.36; P=.04; η^2=0.08), therefore, reflects the strong expression effect in the exclusive condition. The main effect of interaction style was not significant overall ($F_{1,51}$=2.91; P=.10; η^2=0.05).

Behavioral Patterns: Designation Frequency

The analysis of behavioral data unveiled an unexpected pattern in the designation behavior of participants (Table 6). Contrary to established expectations, participants designated the exclusive robot significantly more frequently than the inclusive robot. This finding stands in contrast to earlier research with university students, who tended to favor the inclusive robot [23]. Qualitative interview data suggested that participants' designation behaviors were guided primarily by considerations of fairness and turn taking.

Table . Participants' designation frequency during task 2[a].

Target robot	Designations, mean (SD)	Total designations, n
Inclusive interaction	5.32 (2.12)	282
Exclusive interaction	6.42 (1.84)	340
Happy expression	6.09 (2.06)	323
Unhappy expression	5.64 (2.04)	299

[a]Inclusive vs exclusive: t_{52}=2.18 (P<.05); happy vs unhappy: t_{52}=0.87.

Representative interview quotes illustrating fairness-based designation strategies included the following:

> *I wanted to ensure fair turn-taking so everyone could participate.* [G1-17-F]
>
> *Since the blue robot kept choosing the red robot, I designated blue to make it fair.* [G2-16-F]
>
> *I simply picked them in turns to give everyone a chance to play.* [G1-4-F]
>
> *I assigned both robots because I believe the game should involve taking turns.* [G2-1-M]
>
> *I chose randomly since we were playing together and wanted everyone to have their turn.* [G1-26-M]

These responses suggest that older adults actively sought to balance the exclusive robot's behavior. They consistently used turn-taking strategies to maintain fairness and equitable participation, including for the robots themselves.

Discussion

Principal Findings

This study examined how a robot's facial expressions and interaction style affect the development of rapport between older adults and social robots. Our findings revealed that (1) positive facial expressions significantly enhanced acceptance, rapport, rapport expectations, and trust; (2) an inclusive interaction style contributed to improved rapport and heightened rapport expectations; (3) there were notable interaction effects between facial expression and interaction style on both rapport and trust; and (4) older adults demonstrated a strong tendency toward inclusivity regarding robots, actively compensating for any behaviors that may seem exclusive.

Comparison With Prior Work

Our findings regarding the impact of facial expressions are consistent with those of prior research indicating that positive robot expressions enhance the quality of interactions [19,33]. The observed effect sizes (η^2=0.08-0.22) point to medium to

XSL•FO
RenderX

large effects, underscoring the importance of facial expression as a significant cue for older adults in rapport building.

The interaction effects between facial expression and interaction style represent a novel theoretical contribution. Under inclusive conditions, the robot's behavior reliably signaled social engagement and satisfied CASA-derived expectations for turn taking and politeness norms [12,13]. This behavioral sufficiency may attenuate reliance on facial expression as a diagnostic cue. Under exclusive conditions, however, the behavioral signal was ambiguous or negative, creating a cue vacuum that heightened the diagnostic value of facial expressions as participants sought to infer the robot's social intentionality. The large interaction effect for trust (η^2=0.25) is particularly noteworthy: trust in the exclusive robot was substantially modulated by facial expression, whereas trust in the inclusive robot remained relatively stable regardless of expression. This pattern aligns with the extension of CASA by Gambino et al [12], which posits that, when behavioral signals deviate from social norms, people rely more heavily on secondary cues such as appearance. Incongruent combinations (positive expression+exclusive behavior) appeared to intensify attention to facial cues as participants attempted to interpret the robot's intentions, a process that, in this sample, ultimately yielded higher rapport with and trust in the happy-faced exclusive robot, suggesting that positive expression may partially compensate for exclusive behavior [23]. Future research using mediation designs could formally test the proposed attentional pathway.

Notably, older adults designated the exclusive robot more frequently than the inclusive robot, in stark contrast to university students in earlier research [23], who favored the inclusive robot. This generational difference may reflect divergent values related to fairness and group harmony. In the socioemotional selectivity theory by Carstensen [34], older adults prioritize emotionally meaningful relationships and prosocial goals. This orientation may lead them to perceive the exclusive robot as a social agent in need of remediation and actively compensate through equitable turn assignment. This pattern is also consistent with prosocial aging research documenting heightened concern for fairness among older adults [24,35]. Future studies using more gender-balanced samples would be valuable for determining whether this inclusivity tendency is particularly characteristic of older women, who constituted the majority of our sample.

Implications for Healthy Aging Interventions

These findings carry important implications for the design of social robots that support healthy aging. First, positive facial expression capabilities appear essential as they produced medium to large enhancements across a range of rapport-related outcomes. Second, designers should prioritize inclusive interaction behaviors that actively engage older users. Third, the observed interaction effects indicate that optimizing individual features in isolation may not yield optimal combined outcomes. The mutual amplification of facial expression and interaction style, particularly in exclusive contexts, needs to be considered from a holistic design perspective.

From a public health standpoint, the demonstrated capacity of older adults to establish rapport with robots through brief, gamified interactions underscores the feasibility of scalable robot-mediated interventions. Moreover, the finding that older adults viewed robots as entities deserving of equitable treatment highlights the potential for meaningful human-robot relationships to complement human care in addressing loneliness and social isolation [36].

Limitations

This study offers valuable observational insights; however, several limitations warrant careful consideration. First, our sample was predominantly female (45/53, 84.9%), likely reflecting recruitment through university-affiliated volunteer networks in Taiwan, which tend to overrepresent women among older adults. Because gender can influence social cue processing and trust building in HRI [30,37], the present findings may be more directly applicable to older women. Future studies with more gender-balanced or male-majority samples are needed to assess generalizability. Second, the gamified Radish Squat interaction enabled precise experimental control but may differ from naturalistic caregiving contexts in duration, emotional stakes, and task demands. The rapport captured in this study reflects initial formation in a brief, structured setting; whether these patterns persist and predict rapport stability in long-term caregiving relationships requires longitudinal investigation [38]. Nonetheless, gamified formats are increasingly adopted in health-promoting robot interventions for older adults, lending direct relevance to the present findings. Third, the single-session design precludes conclusions about long-term rapport development or the cumulative effects of repeated exposure. Additionally, our focus on healthy, community-dwelling adults suggests that the findings may not be applicable to individuals with cognitive impairments. Future research could benefit from systematically examining gender differences, using longitudinal designs, and investigating the cognitive and neurobiological mechanisms that may underpin older adults' strong inclusivity responses toward robots.

Conclusions

This study demonstrates that gamified interactions can effectively foster rapport between older adults and social robots. Positive facial expressions and inclusive interaction behaviors enhanced acceptance, rapport, rapport expectation, and trust. Crucially, these elements interacted synergically: the positive impact of happy facial expressions was substantially amplified when the robot behaved exclusively, suggesting a context-dependent cue-weighting process grounded in the CASA paradigm [12]. The unexpected finding that older adults compensated for exclusive robot behavior through fairness-based turn taking further suggests that they perceived these robots as social entities worthy of equitable treatment. Collectively, these insights offer evidence-based design guidelines for social robots intended to foster meaningful relationships with older adults and contribute to the development of technology-based interventions that promote healthy aging.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank NUWA Robotics Corp for their technical support with the Kebbi robot platform.

Funding

This work is partly funded by the Taiwan National Science and Technology Council (grant NSTC114-2622-H-002-002) and National Taiwan University (114L89470).

Data Availability

The datasets are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: HPY (lead), WL (supporting)
Data curation: KCL
Formal analysis: HPY (lead), KCL (supporting)
Funding acquisition: HPY
Investigation: KCL, WL (supporting)
Methodology: HPY
Project administration: HPY (lead), WL (supporting)
Resources: HPY
Supervision: HPY
Validation: WL
Visualization: KCL (lead), WL (supporting)
Writing—original draft: KCL (lead), HPY (supporting)
Writing—review and editing: HPY (lead), WL (supporting)

Conflicts of Interest

None declared.

References

1. Progress report on the United Nations Decade of Healthy Ageing, 2021-2023. : World Health Organization; 2023 URL: https://www.who.int/publications/i/item/9789240079694 [accessed 2026-07-03]
2. The ageing of populations and its economic and social implications. : United Nations Department of Economic and Social Affairs; 1956 URL: https://digitallibrary.un.org/record/3975472?ln=en&v=pdf [accessed 2026-07-03]
3. Senior Citizens Welfare Act. Laws & Regulations Database of the Republic of China. 2025. URL: https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=D0050037 [accessed 2026-07-03]
4. Michel JP, Sadana R. "Healthy aging" concepts and measures. J Am Med Dir Assoc 2017 Jun 1;18(6):460-464. [doi: 10.1016/j.jamda.2017.03.008] [Medline: 28479271]
5. World report on ageing and health. : World Health Organization; 2015 URL: https://www.who.int/publications/i/item/9789241565042 [accessed 2026-07-03]
6. Ageing and health. World Health Organization. 2025 Oct 1. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2026-07-03]
7. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
8. Chen SC, Moyle W, Jones C, Petsky H. A social robot intervention on depression, loneliness, and quality of life for Taiwanese older adults in long-term care. Int Psychogeriatr 2020 Aug;32(8):981-991. [doi: 10.1017/S1041610220000459] [Medline: 32284080]
9. Pu L, Moyle W, Jones C, Todorovic M. The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies. Gerontologist 2019 Jan 9;59(1):e37-e51. [doi: 10.1093/geront/gny046] [Medline: 29897445]
10. Broadbent E, Stafford R, MacDonald B. Acceptance of healthcare robots for the older population: review and future directions. Int J Soc Robotics 2009 Nov;1:319-330. [doi: 10.1007/s12369-009-0030-6]
11. Breazeal CL, Ostrowski AK, Singh N, Park HW. Designing social robots for older adults. Bridge 2019;49(1):22-31 [FREE Full text]
12. Gambino A, Fox J, Ratan RA. Building a stronger CASA: extending the computers are social actors paradigm. Hum Mach Commun 2020 Feb;1:71-86. [doi: 10.30658/hmc.1.5]

XSL•FO
RenderX

13. Nass C, Moon Y. Machines and mindlessness: social responses to computers. J Soc Issues 2000;56(1):81-103. [doi: 10.1111/0022-4537.00153]
14. Carros F, Meurer J, Löffler D, Unbehaun D, Matthies S, Koch I, et al. Exploring human-robot interaction with the elderly: results from a ten-week case study in a care home. In: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems: Association for Computing Machinery; 2020:1-12. [doi: 10.1145/3313831.3376402]
15. Yueh HP, Lin W, Chen CH, Chang FH, Fu LC. From automation to cognition: a contextual design approach for enhancing elderly patron services with proactive and memory-aware library robots. Electron Libr 2026;44(1):204-226. [doi: 10.1108/EL-06-2025-0244]
16. Tickle-Degnen L, Rosenthal R. The nature of rapport and its nonverbal correlates. Psychol Inq 1990;1(4):285-293. [doi: 10.1207/s15327965pli0104_1]
17. Lubold N, Walker E, Pon-Barry H. Effects of adapting to user pitch on rapport perception, behavior, and state with a social robotic learning companion. User Model User Adapt Interact 2021 Mar;31(1):35-73. [doi: 10.1007/s11257-020-09267-3]
18. Seo SH, Griffin K, Young JE, Bunt A, Prentice S, Loureiro-Rodríguez V. Investigating people's rapport building and hindering behaviors when working with a collaborative robot. Int J Soc Robotics 2018 Jan;10:147-161. [doi: 10.1007/s12369-017-0441-8]
19. Rawal N, Stock-Homburg RM. Facial emotion expressions in human–robot interaction: a survey. Int J Soc Robotics 2022 Sep;14:1583-1604. [doi: 10.1007/s12369-022-00867-0]
20. Wang N, Gratch J. Rapport and facial expression. In: 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops: IEEE; 2009. [doi: 10.1109/ACII.2009.5349514]
21. Erel H, Cohen Y, Shafrir K, Levy SD, Vidra ID, Tov TS, et al. Excluded by robots: can robot-robot-human interaction lead to ostracism? In: HRI '21: Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction: Association for Computing Machinery; 2021:312-321. [doi: 10.1145/3434073.3444648]
22. Williams KD, Cheung CK, Choi W. Cyberostracism: effects of being ignored over the internet. J Pers Soc Psychol 2000 Nov;79(5):748-762. [doi: 10.1037//0022-3514.79.5.748] [Medline: 11079239]
23. Yueh HP, Wang YR, Huang TR. Higher trust toward exclusive robots in a multiplayer human-robot interaction game. Discov Psychol 2025;6:12. [doi: 10.1007/s44202-025-00558-7]
24. Carstensen LL, Turan B, Scheibe S, et al. Emotional experience improves with age: evidence based on over 10 years of experience sampling. Psychol Aging 2011 Mar;26(1):21-33. [doi: 10.1037/a0021285] [Medline: 20973600]
25. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
26. Nomura T, Kanda T. Rapport–expectation with a robot scale. Int J Soc Robotics 2016 Jan;8(1):21-30. [doi: 10.1007/s12369-015-0293-z]
27. Schaefer KE. Measuring trust in human robot interactions: development of the "Trust Perception Scale-HRI". In: Mittu R, Sofge D, Wagner A, Lawless W, editors. Robust Intelligence and Trust in Autonomous Systems: Springer; 2016:191-218. [doi: 10.1007/978-1-4899-7668-0_10]
28. Gosling SD, Rentfrow PJ, Swann WJ. A very brief measure of the big-five personality domains. J Res Pers 2003 Dec;37(6):504-528. [doi: 10.1016/S0092-6566(03)00046-1]
29. Carré A, Stefaniak N, D'Ambrosio F, Bensalah L, Besche-Richard C. The Basic Empathy Scale in adults (BES-A): factor structure of a revised form. Psychol Assess 2013 Sep;25(3):679-691. [doi: 10.1037/a0032297] [Medline: 23815121]
30. Nomura T, Kanda T, Suzuki T, Kato K. Psychology in human-robot communication: an attempt through investigation of negative attitudes and anxiety toward robots. In: RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication: IEEE; 2004. [doi: 10.1109/ROMAN.2004.1374726]
31. Pütten AR, Bock N. Development and validation of the Self-Efficacy in Human-Robot-Interaction Scale (SE-HRI). ACM Trans Hum Robot Interact 2018;7(3):1-30. [doi: 10.1145/3139352]
32. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Routledge; 2013. [doi: 10.4324/9780203771587]
33. Lubold N, Walker E, Pon-Barry H, Ogan A, editors. Automated pitch convergence improves learning in a social, teachable robot for middle school mathematics. In: Artificial Intelligence in Education - 19th International Conference, AIED 2018, Proceedings: Springer; 2018. [doi: 10.1007/978-3-319-93843-1_21]
34. Carstensen LL. Social and emotional patterns in adulthood: support for socioemotional selectivity theory. Psychol Aging 1992;7(3):331-338. [doi: 10.1037/0882-7974.7.3.331] [Medline: 1388852]
35. Carstensen LL. Socioemotional selectivity theory: the role of perceived endings in human motivation. Gerontologist 2021 Nov 15;61(8):1188-1196. [doi: 10.1093/geront/gnab116] [Medline: 34718558]
36. Murayama H, Takase M. Evaluating the effectiveness of digital social robots in reducing loneliness among community-dwelling older adults in Japan: randomized controlled trial and qualitative analysis. JMIR Aging 2025 Oct 22;8:e74422. [doi: 10.2196/74422] [Medline: 41124645]
37. Tay B, Jung Y, Park T. When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. Comput Hum Behav 2014 Sep;38:75-84. [doi: 10.1016/j.chb.2014.05.014]
38. Henschel A, Laban G, Cross ES. What makes a robot social? A review of social robots from science fiction to a home or hospital near you. Curr Robot Rep 2021;2(1):9-19. [doi: 10.1007/s43154-020-00035-0] [Medline: 34977592]

XSL•FO
RenderX

Abbreviations

CASA: computers are social actors
HRI: human-robot interaction

Edited by N Schütz; submitted 11.Jan.2026; peer-reviewed by A Tan, E Erzen, W Qiaofei; revised version received 31.May.2026; accepted 25.Jun.2026; published 15.Jul.2026.

Please cite as:
Yueh HP, Lin KC, Lin W
The Influence of Robot Facial Expression and Inclusion Behavior on Rapport and Trust in Older Adults: Mixed-Factorial Experimental Study
JMIR Aging 2026;9:e91235
URL: https://aging.jmir.org/2026/1/e91235
doi:10.2196/91235

© Hsiu-Ping Yueh, Kuang-Chi Lin, Weijane Lin. Originally published in JMIR Aging (https://aging.jmir.org), 15.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perceived Use of Web-Based Videoconferencing for Social Connectedness Among Older Adults Living in Long-Term Care: Qualitative Study

Anna Garnett[1], RN, PhD; Halyna Yurkiv[1], BScN, MN; Denise M Connelly[2], PhD; Richard Booth[1], RN, PhD; Lorie Donelle[3], RN, PhD

[1]Arthur Labatt Family School of Nursing, Western University, 1151 Richmond Street, London, ON, Canada

[2]School of Physiotherapy, Western University, London, ON, Canada

[3]Biobehavioral Health & Nursing Science, College of Nursing, University of South Carolina, South Carolina, SC, United States

Corresponding Author:
Anna Garnett, RN, PhD
Arthur Labatt Family School of Nursing, Western University, 1151 Richmond Street, London, ON, Canada

Abstract

Background: The COVID-19 pandemic highlighted how restrictions on in-person interactions within long-term care homes (LTCHs) severely compromised social connectedness among older adults and their families. Post pandemic, despite policy changes supporting greater in-person family engagement, frequent outbreaks continue to disrupt face-to-face interactions, and factors such as geography, life circumstances, and health can constrain family members' ability to make regular in-person visits. Research suggests that web-based videoconferencing technology (WVT) may be a practical solution to help older adults within LTCHs to maintain social connection in the absence of physical gathering. However, increased understanding of end user experience is lacking, and more information on LTCHs' readiness to support and sustain WVT will be needed if this modality is to be successfully and widely implemented.

Objective: This study aimed to understand how older adults living in LTCHs, their families, and LTCH staff members perceived the use and ease of use of WVT devices for facilitating social connectedness.

Methods: Using a qualitative description approach, in-depth semistructured interviews were conducted with 7 older adults, 22 family members, and 10 staff across 3 LTCHs via Zoom (Zoom Communications, Inc), Microsoft Teams, or phone calls. Data were analyzed using a directed content analysis informed by the technology acceptance model.

Results: Findings were structured into 3 main themes: actual system use, perceived usefulness of WVT, and perceived ease of use of WVT. Participants described using a range of WVT hardware and software to promote social connection between older adults and family members. Videoconferencing had a crucial role in supporting older adults and their family members' positive emotional state while also enabling them to maintain life and social roles such as participating in family functions. Despite the perceived use of these tools, participants were concerned about the decline in offering videoconferencing services across LTCHs post pandemic. Some participants noted shifting funding priorities toward supporting in-person recreational activities rather than diversifying web-based social connection options. In addition, factors pertaining to WVT ease of use and integration included limited staff to support older adults with different physical and cognitive needs, variability in digital literacy including knowledge about accessibility features to enhance the ease of use, and families' lack of awareness about the availability of WVT for social connectedness.

Conclusions: Web-based videoconferencing technology has the potential to be a meaningful tool to reduce social isolation and promote a sense of social connectedness among older adults and their families and friends. Future research should explore how WVT could be integrated into care planning for this population, particularly in situations where older adults may be at heightened risk for social isolation. Resource allocation toward equipment, infrastructure, and family and staff training would be well-placed to increase engagement with WVT within LTCHs.

(JMIR Aging 2026;9:e73213) doi:10.2196/73213

KEYWORDS

web-based; videoconferencing; virtual technology; long-term care; social connectedness; older adults

XSL•FO
RenderX

Introduction

Background

Globally, the COVID-19 pandemic impacts were profound, causing widespread loss of life and morbidity [1,2]. During the earliest phases of the pandemic, significant restrictions were placed on many social activities, including normal activities of daily living (eg, shopping and leisure activities) [3-8]. One specific aspect of daily living that was significantly affected included the notion of *social connectedness* [9,10]. Social connectedness can be described as the subjective sense of being in close-knit relationships with others [11,12]. Having interpersonal relationships and being able to gather with others are integral to individuals' health and well-being [13,14]. These avenues for social connectedness can prevent people from feeling lonely and reduce the risk of impaired mental health [14-21]. Canadian older adults residing in long-term care homes (LTCHs) were disproportionally affected by the lack of in-person interactions or access to social spaces over the course of the pandemic [22], resulting in their compromised social and emotional well-being [22-24].

During the height of the COVID-19 pandemic, various web-based videoconferencing technologies (WVTs) were introduced by LTCH staff and family members to help maintain social connectivity between older adults residing in LTCHs and their social circles [25-29]. Leveraging WVTs to facilitate social connectivity among older adults was particularly beneficial as these individuals were at greater risk of isolation and loneliness due to factors such as pre-existing isolation (eg, living alone), being reliant on family for support, and impaired physical health such as frailty [30]. Web-based videoconferencing, the use of networked digital telecommunications technology, is a form of telepresence that "simulates the experience of being physically present in a remote environment" [31]. FaceTime (Apple, Inc), Skype (Microsoft Corp), or Zoom (Zoom Communications, Inc) are among the most popular videoconferencing applications used to continue communication among older adults and family or friends when in-person gatherings are restricted [26,32]. Using WVTs enabled people to visually interact with their family members and, in addition to hearing their voice, offered a greater reassurance of their well-being through a visual confirmation [29]. In addition, seeing older adults within LTCHs using WVTs can help foster family members' belief that their family member was well-cared for [33].

Research suggests that web-based videoconferencing is a viable form of connection for people living in long-term care who may be physically separated from their families and familiar social settings (eg, religious mass services and health care consultations) [34-39]. However, there remains a lack of insight pertaining to the experiences of the individuals directly using this technology. Moreover, additional research is required to inform sustainable and consistent use of WVTs as a reliable option to help older adults maintain social connectedness [40]. Although WVTs offer clear benefits, some literature suggests that this technology may cater more to family members than older adults living in LTCHs, especially if the older adult lives with some degree of cognitive challenge (eg, dementia) [28,33,41]. The types of benefits offered by WVTs across a range of users for the purposes of fostering social connection necessitate further exploration, particularly among older adults with varied cognitive abilities residing in LTCHs. As such, greater knowledge is required to understand how older adults with both physical and cognitive impairments may benefit from or experience challenges in using WVTs [33]. Furthermore, there is a need to understand if and how LTCHs are positioned in terms of infrastructure and staffing to support successful widespread WVT use. For example, arthritis, vision impairments, or neurodegenerative conditions such as Parkinson disease could necessitate the need for adaptive technology or additional assistance as well as greater staff requirements and ultimately cost for the LTCH [42,43]. The potential benefits of WVTs as a tool to support social connection have been shown, but within LTCHs, there is a need to better understand end user experiences if it is to be successfully integrated and sustained as a widespread tool to support social connectedness.

Objective

The purpose of this study was to understand how older adults living in LTCHs, their families, and LTCH staff members perceived the use and ease of use of WVTs as viable modalities for facilitating social connectedness.

Methods

Project Registration

This study presents the findings of the first stage of a multistage qualitative research study described in the published protocol [44].

Study Design

Using a qualitative description methodology [45], this study sought to explore how older adults in LTCHs, their family members, and LTCH staff experience and understand the use of WVTs to support the social connectedness of older adults in LTCHs. Qualitative description methodology enables researchers to capture and describe a wide range of participants' experiences and perceptions in a comprehensive, yet clear and practical way using in-depth interviews with study participants [45,46]. In addition, qualitative description is a low-inference approach used in social sciences and health care to provide researchers with a practical way to derive rich, close to data descriptions of participant experiences without the need to interpret the meaning of participant experiences [46]. In addition to in-depth interviews, observational field notes were kept throughout this study, thereby enhancing the dataset and the overall study rigor [47]. Studies using qualitative description methodology seek to find answers that are of high relevance to practitioners or policymakers to improve the quality of a delivered service, including participants' thoughts or attitudes toward a service; why do participants use certain services more so than others; and how participants use a service in specific instances [45,48].

Theoretical Framework

The technology acceptance model (TAM) [49] was used to shape interview questions and to organize and report qualitative

XSL•FO
RenderX

findings in this study. TAM is a widely used model that helps researchers uncover how end users adopt the use of technological systems and how this adoption is influenced by constructs such as usefulness and perceived ease of use [50]. According to TAM, the actual system use construct concerns the observable use of technology [49] and, in the context of this study, refers to the use of WVTs by the study participants (older adults, family members, and staff in LTCHs) for the purpose of social connection. The perceived usefulness construct is the extent to which technology end users perceive technology as enhancing their tasks or helping them achieve a particular goal [49]. In this study, perceived usefulness is the degree to which participants believe that WVT (devices and software) facilitates social connectedness or makes it more meaningful. The perceived ease of use construct pertains to the amount of effort required to operate the technology by end users [49] and, in this study, reflects participants' perceptions of how simple it is to operate or navigate WVTs to engage in video calls with friends and family members.

A range of open-ended interview questions was developed using the TAM constructs to guide the in-depth interviews with study participants. For a full list of interview questions, please see Multimedia Appendix 1 in our protocol [44]. In line with qualitative descriptive methodology supporting the use of deductive approaches (ie, TAM) to guide data analysis [51,52], this study presents descriptions of participants' experiences organized under the following TAM constructs: the (1) *actual system use*, (2) *perceived usefulness*, and (3) *perceived ease of use*.

The Use of TAM With Qualitative Description Approach

Findings grounded in qualitative description methodology result from staying close to the words and events in the research data capturing straightforward descriptive accounts of participant experiences [45,46]. In some qualitative descriptive studies, findings are categorized under constructs of a chosen theoretical framework [51]. This study incorporated TAM as a framework for organizing data-driven coding and practical and identifiable constructs of the TAM that may be of use to policy makers. The process of incorporating TAM into the data analysis phase ensured that participant data were organized and described in a structured way without imposing subjective interpretations or other theoretical assumptions on the data. Such an approach aligns with previously published research that successfully integrated TAM with qualitative description methodology to describe older adults' adoption of digital modalities [51,53,54].

Sample

The desired sample size for qualitative description studies is influenced by factors such as the subject or phenomenon of interest, its representation within the broader population, and the potential variability of experiences with the phenomenon [55]. As such, the recruitment goal of this study was 45 individuals comprised of 3 groups: older adults residing in LTCHs, their family members, and associated staff from 3 participating LTCHs located in southwestern Ontario, Canada. Eligible participants were English-speaking individuals with mild to moderate cognitive disability and able to participate in an interview using the Zoom or Microsoft Teams videoconferencing platforms or a phone. Potential older adult participants with mild cognitive decline living in LTCHs were assessed [56] for suitability to participate in this study by a member of their circle of care within the LTCHs before any study information was shared with them.

Ethical Considerations

This study was reviewed and approved by the Western University Research Ethics Board (2024-121993-91059). Participants were informed about the research study, and their questions were answered by the study research assistant before they consented to participate. Collected data was anonymized and participants who completed an interview were given a $10 (CAD or $7.20 USD) gift card as an honorarium.

Recruitment and Data Collection

Participant recruitment and data collection took place between July 2023 and January 2024 using purposeful and maximum variation sampling to help ascertain a sample with a variety of experiences with WVTs, especially among the older adult and family member participants. Research information was distributed to prospective participants via on-premises physical posters, emails, LTCH newsletters, and word of mouth. Individual in-depth interviews with older adults, their family members, and LTCH staff were offered in person, via phone, or by Zoom videoconferencing to accommodate individual preferences and took place at an agreed-upon date and time. Recreational therapy department staff or a family care partner assisted older adult participants in joining the Zoom or interview if necessary. Interviews with participants were audio recorded and transcribed verbatim prior to analysis.

Data Analysis

The collected data were anonymized and analyzed using NVivo (version 13; Lumivero) qualitative analysis software [57] and directed content analysis, a qualitative analytic technique characterized by a structured stepwise process of coding the data and organizing it into categories and themes [58,59]. The first step of the data analysis entailed deidentification of the cases. Interviews collected from family members were assigned identification such as FM followed by the number of the participant (FM01, FM02, and so on). Similarly, older adults were assigned identification of OA (OA01, OA02, and so on), and staff participants were assigned identification S (S01, S02, and so on). The next step involved reading the interview transcripts, developing memos, and noting down compelling quotes. A formative coding framework was then developed based on the TAM constructs with consensus reached through discussion of the research team. The key constructs of the TAM, perceived usefulness and ease of use of technology, served as initial codes or units of analysis. Data that did not lie within one of the key codes were labeled with a new code and organized into new categories or as subcategories under the initial codes [58]. The preliminary coding framework was tested against the same 5 interview transcripts. Researchers then met on a biweekly basis to refine the coding framework by establishing links between generic and main conceptual categories. The newly identified codes were compared and

XSL•FO
RenderX

discussed for similarities and differences until consensus was reached regarding the emergent patterns as the concepts became more densely packed with meaning and evidence. Study rigor was ensured through conducting member checking [60], acknowledging biases researchers may introduce to the research and practicing reflexivity throughout the research processes, and ensuring confirmability of data through triangulation between interviews and field notes [61].

Results

Demographic Characteristics Overview

The total sample comprised 39 participants, which is congruent with sample sizes in other studies using qualitative descriptive methods [51]. The sample included older adults (n=7, 18%), family members (n=22, 56%), and staff (n=10, 26%) from 3 participating LTCH facilities, offering diverse perspectives on WVT use in these settings. The 3 participating homes were licensed, not-for-profit LTCHs spanning 120 kilometers and providing comprehensive care to between 160 to 394 individuals, including an emphasis on person-centered care.

All 7 (18%) older participants were female and aged between 64 and 95 years. Older participants were widowed (28/39, 71%) or divorced (11/39, 29%). All 39 older adults either had a college diploma (n=22, 57%) or a university degree (n=16, 42%), providing some context that may be related to the ease of using WVTs. The majority of older adults (n=28, 71%) reported living with cognitive or physical challenges (eg, poststroke impairment, dementia, and arthritis), which provides essential context for understanding the impact of cognitive or physical limitations on WVTs' usefulness, ease of use, and intention to continue using WVTs. Moreover, 31 out of 39 (nearly 80%) participants in this study identified that physical (eg, arthritis) or mental disability (eg, dementia) affected individuals' ability to use WVTs.

The 22 family members (FMs) consisted of 18 female and 4 male participants, and were aged between 40 and 69 years. Many family members held a university bachelor's degree (16/39, 41%), with a notable portion being retired (14/39, 36%).

All 10 staff participants were female, with just under half (10/39, 40%) working in their profession between 6 and 10 years. Staff participants included a range of health professionals such as recreational therapists (12/39, 30%), social workers (8/39, 20%), personal support workers (8/39, 20%), a registered practical nurse (4/39, 10%), and a spiritual care practitioner (4/39, 10%).

Thematic Presentation of the Findings

Overview

Study findings on participants' roles (older adults, FMs, and staff) and experiences in using WVTs for the purpose of social connectedness are presented in themes informed by the TAM and include the actual system use, with 2 subthemes—(1) WVTs' hardware and software and (2) trajectory of using WVTs from pre- to postpandemic; perceived usefulness, with 4 subthemes—(1) enabling remote connection, (2) providing emotional and psychological benefits, (3) fostering continuation of life and social roles, and (4) the enriching effect of video presence; and perceived ease of use, with 3 subthemes–(1) design and practicality of the devices and platforms used, (2) usability and accessibility of the devices, and (3) impact on the workload (Textbox 1).

Textbox 1. Summary of technology acceptance model–informed themes and subthemes.

Theme 1: actual system use

- Web-based videoconferencing technologies' hardware and software
- Trajectory of using web-based videoconferencing technologies from pre- to postpandemic

Theme 2: perceived usefulness of web-based videoconferencing technologies

- Enabling remote connection
- Providing emotional and psychological benefits
- Fostering continuation of life and social roles
- Enriching effect of video presence

Theme 3: perceived ease of use of web-based videoconferencing technologies

- Design and practicality of the devices and platforms used
- Usability and accessibility of the devices
- Impact on the workload

Theme 1: Actual System Use

Overview

The actual system (WVTs) use theme consisted of 2 subthemes: (1) the hardware and software used by older adults, FMs, and staff, and (2) WVTs use trajectory from pre- to postpandemic timeframes.

WVTs' Hardware and Software

During the pandemic, LTCH staff introduced iPads and tablets as a virtual means to facilitate social connectedness between families. Older adults and families also used a range of personal

XSL•FO
RenderX

devices with videoconferencing capabilities such as iPads or tablets, laptops, computers, smartboards, and smartphones. Moreover, some families independently researched and purchased nontraditional devices that FMs suggest have more merit for people with physical or cognitive conditions (eg, dementia, arthritis, and paralysis) to access the WVTs they were using, namely, Portal, Alexa Echo Show, and ViewClix hardware. Participants also used various software platforms for WVTs, including Zoom, Teams, WhatsApp, Skype, Facebook Messenger, Signal, and FaceTime (Table 1).

Table . Web-based videoconferencing technologies used across study participants (N=39).

Web-based videoconferencing technology	Older adults, n	Family members, n	Staff, n	Total, n
Hardware types				
iPad or tablet	6	15	10	31
Computer	3	6	1	10
Laptop	0	4	2	6
Smartboard	0	0	2	2
Smartphone	2	8	3	13
Portal	0	2	0	2
Alexa Echo Show	0	1	0	1
ViewClix	0	1	0	1
Software types				
Zoom	0	10	7	17
Teams	6	1	1	8
WhatsApp	0	1	0	1
Skype	0	3	2	5
Facebook messenger	1	2	2	5
Signal	0	2	0	2
FaceTime	1	12	3	16

Trajectory of WVT Use From Pre- to Postpandemic

The iPads or tablets were predominantly leveraged by LTCHs' recreational department teams during the pandemic-related lockdowns and resultant restrictions on in-person visitations as an alternative to socially connect FMs and older adults residing in LTCHs. However, post pandemic, this service was not unanimously integrated into LTCHs' care delivery. Some LTCHs have embraced the technology, integrating it into daily routines or care planning, while others have shifted funding priorities toward in-person recreational activities, which some FMs shared were not always as meaningful as the ability to see and interact with their loved ones on the screen.

Participants reported varied knowledge of acceptance and normalization of WVT use across different LTCHs from pre- to postpandemic. Most participants who used WVTs to socially connect with people outside LTCHs prepandemic continued to do so during and post pandemic. For these participants, it was the only way to connect with family who lived abroad or far away from LTCHs.

Among the study participants, 14 (36%) participants stated they began using WVTs to support social connectedness due to the pandemic (7 FMs, 5 staff, and 1 older adult). Eleven (28%) participants shared that WVTs continue to be offered by the LTCHs or that they still use it post pandemic (6 FMs, 3 staff, and 1 older adult), while 10 (26%) participants believe the service has stopped being offered post pandemic by the LTCHs (5 FMs, 3 staff, and 1 older adult). A few reasons were provided by the participants to explain the decline in use post pandemic. First, a shift of priorities to do more in-person activities and visitations resulted in a lack of recreational department staff support to help with videoconferencing service. As 1 FM whose mother was admitted to LTCH in March 2023 shared:

> *I've never had it offered to me with my mom. I don't think the emphasis is on it and part of it is because the reason they had time to do more of that [videoconferencing] was because so many other programs were cancelled, right?* [FM21]

A second reason provided was a lack of awareness among older adults and their family members that connecting socially using WVTs was still being offered in the LTCHs:

> *If they have anything to offer, they're certainly not advertising it, I think they don't have it. I think that it was simply cut.* [FM19]

A similar concern was echoed by a staff member participant:

> *I don't even know how much recreation [department] even does help with Zoom calls anymore because all family – like anybody can come in now, right? Or they just call too, right? I don't really see much Zoom,*

XSL•FO
RenderX

> *FaceTime, or anything anymore with residents really.* [S07]

The third identified contributor to the decline of WVT use post pandemic shared by the participants was related to the workload challenges imposed by supporting the use of WVTs to socially connect people:

> *We just simply don't have that manpower to be engaging in this environment as well as supporting you to have a conversation with someone. Even though we understand the value in that, and how important that is for their emotional health, you know, when they are feeling positive and good and they know that their family is safe, they don't have responsive behaviours.* [S02]

In summary, the use of WVTs in LTCHs demonstrated benefits for maintaining social connectedness during the pandemic. However, the transition to postpandemic society revealed a decline in its continued use and only partial normalization of videoconferencing across the 3 LTCHs in this study. The trajectory of WVT use from pre- to postpandemic provides insights into the fluid role that technology can have in supporting social connectedness of older adults in long-term care settings.

Theme 2: Perceived Usefulness of WVTs

Overview

The perceived usefulness of WVTs for social connectedness is represented by 4 subthemes: enabling remote connection, providing emotional and psychological benefits, fostering continuation of life and social roles, and the enriching effects of video presence.

Enabling Remote Connection

WVTs enabled residents to connect remotely with family and friends, thereby bridging physical distances that in some cases were difficult for FMs to overcome. This opportunity was particularly important for those families who lived at a great geographical distance from the LTCHs or were unable to visit frequently due to various constraints:

> *We'll FaceTime with family members that are far away. I mean, last year, we were basically in four different time zones. My sister was Central, I'm Eastern, my daughter was in the U.K. and one of my nephews was in Japan.* [FM09]

Conversely, another participant noted a robust example of how WVTs could connect people in the same building:

> *At one point [my parents] had to be separated and put on separate floors, so we used a lot of FaceTime and – is it Facebook Messenger? Yes, we were using that between the two of them. So, two of us would have to go in every day and one to my father, one to my mother, so that the two of them could talk.* [FM03]

In addition, videoconferencing was a useful tool that enabled family to note health care–related details in older adult's presentation, which they later communicated to the staff, as shared by this FM:

> *I recognized when she had an eye infection, or I could see that her hands were really dry. I couldn't have done that if I was on the phone. We were able to see things and it informed us more about how my mother really was, and it improved her care.* [FM18]

WVTs were also used for accessing care as an accessible alternative in place of the traditional in-person meetings:

> *For the annual team meetings, we had the option to do that virtually or in person. And my sister doesn't drive, and the care home is a long way from her place, so she chose to meet with the care staff virtually, which she really appreciated.* [FM02]

Those who have not used WVTs for health care access expressed receptiveness to this possibility (FM12). However, those FMs who lived within proximity to the LTCH preferred in-person meetings for such matters. For instance, 1 FM shared that*:*

> *there is an annual health care review, and they might have offered that as an option, but I'm just ten minutes from where my mum is, so you know, I would just come in for that meeting.* [FM16]

Another FM raised an interesting point regarding the nonverbal communication that takes place during the in-person meeting that can be missed or hard to sense over the video call:

> *In-person you sense if there's anything going on. You sense it more because when you are videoconferencing, you may see just one person at a time. But when you're sitting in the room, we're all sitting around, and I see everybody, and you feel something if there's something wrong, if they're not telling you everything. On videoconferencing you can't feel it. You just see it, but you don't feel like there's any issues.* [FM22]

These quotes exemplify how WVTs ensured continuous social engagement, helping many residents feel less isolated and more connected to the outside world during lockdown or when they were in isolation.

Providing Emotional and Psychological Benefits

Regular videoconferencing interactions helped support older adults' mental health by reducing feelings of loneliness and isolation. The web-based presence connection with family and friends was a vital source of emotional support. Older adults and their FMs experienced substantial emotional benefits from connecting through the video call via WVTs. It provided them with comfort, joy, and a sense of connectedness, which heightened their positive emotional and psychological well-being:

> *The biggest thing it helps with is their mental health, especially if they're just making that change into long-term care. Or if they've been in there, to help lift them out of some bad times or bad moments, because I know that depression in long-term care is huge. There's 24 hours in a day, and if they're sleeping 10 hours and they only get four hours devoted time from someone, there's a lot of time in between there. So, if there could be some way that*

XSL•FO
RenderX

> *they could speak to friends or family or anyone else more, then I think it would decrease depression in long-term care. And that would be big.* [FM07]

Other examples of the ability of WVTs to provide emotional and psychological benefits include its potential to anecdotally delay the decline associated with a chronic disease such as dementia:

> *I think, for the dementia and Alzheimer's, for her [mom] to visually see the relatives on whatever we use, whether it's Zoom or FaceTime or Portal, and connect the face with the voice has sort of delayed her [mom's] decline, because you can actually see them.* [FM09]

In addition, WVTs were used for prevention of responsive behaviors associated with dementia:

> *They [staff] proactively call, because they figure it'll help settle her [mom] down. So now they do FaceTime because they obviously realize that if she [mom] can see somebody then it calms her down even more than just a telephone call.* [FM09]

Fostering Continuation of Life and Social Roles

Fostering continuation of life and social roles was another subtheme associated with WVT use. For example, WVTs provided an avenue to maintain a sense of normalcy by allowing older adults to continue participating in daily life activities and social roles beyond LTCH walls. For instance, older adults kept in touch with FMs and things that were meaningful to them (eg, their pets), remained updated on family events, and engaged in routine conversations that they would typically have in-person prior to being admitted to LTCHs:

> *I was trying to keep [mom] updated, how her pet was doing... show her visually–I would take her for a walk, through the gardens.* [FM07]

Participants found WVTs useful because they allowed them to fulfill their roles as parents, grandparents, and friends. Even if through the screen, they could still offer advice, participate in family decisions, and stay involved in the lives of their loved ones.

In addition to continuing to fulfill their social roles within the family and household, WVTs allowed older adults to participate in other social events. For example, 1 staff member described how WVTs were used for facilitating music concerts for older adults*:*

> *We would get music entertainers at home through video chat and each home area would have on their smart board to get the video chat going.* [S03]

Another example was using videoconferencing to stream mass in churches facilitated by a spiritual care practitioner (1 staff). As a result, 1 older adult requested to attend a virtual mass daily, as it allowed this resident to feel connected to a church community they belonged to prior to the pandemic. These examples demonstrate the expanded role of WVTs in fostering a sense of connectedness with other community members.

Enriching Effects of Video Presence

Compared to other means of communication (eg, landline phone), WVTs were described as having an enriching effect on social connectedness. Participants shared that being able to see faces, share visual experiences, and express emotions more vividly via the screen contributed to a deeper sense of connection:

> *How many times you read an email where you take it wrong. How many times you get on a phone, and you don't understand a person may be sounding angry, but they're visually upset. I think the video adds a lot in those situations, particularly since in a nursing home it is so emotional.* [FM01]

In addition to the interactions between family and older adults, participants described how videoconferencing allowed them to record interactions and watch them later or send them to others (FM01) or observe the resident-staff interactions:

> *A side benefit is, every so often I'll catch one of the staff members in the room with him [my husband], so I have a nice chat with the nurse. I live next door, so I'm there all the time anyway, but it gives me a chance to see the interactions between him and the staff.* [FM17]

In summary, these findings underscore the multifaceted usefulness of WVTs in maintaining social connectedness and well-being not just for older adults, but also for FMs. Usefulness of WVTs may be hindered or enhanced based on how intuitive and user-friendly these technologies are for older adults, FMs, and LTCH staff.

Theme 3: Perceived Ease of Use of Web-Based Videoconferencing Technologies

Overview

The perceived ease of use was informed by 3 subthemes: design and practicality of the hardware and software, usability and accessibility of the WVTs, and ease of use in relation to the impact on the workload for staff members supporting WVT use.

Design and Practicality

Factors impacting the external design of the devices, such as size of the screen, weight, and the availability of supportive accessories like stands, iPad covers, stylus, external speakers, and headphones influenced the practicality and ease of use of WVTs for older adults.

Hardware devices with larger screens, clear audio, and simplified navigation were reported as easier to use. Accessories like headphones or speakers also improved the ease of use. Participants also benefited from using devices with software interfaces requiring minimum to no effort to operate:

> *With our mom, she didn't have to know anything, she just would see faces pop up.* [FM04]

Portable hardware devices were preferred by some, as they could be conveniently moved and adjusted according to the residents' needs and preferences and provided more privacy (4 staff). For instance, 1 FM described how they enjoyed going

XSL•FO
RenderX

for a walk into the garden and video calling their family from that location. At the same time, portability potentially breached privacy when residents were situated in an open public space like the dining area due to a lack of private space or convenience for staff (1 FM and 2 staff). Eight (21%) participants also expressed their concern for device misplacement, theft, or a risk of accessing passwords and private information through older adults' personal devices (1 older adult, 5 FMs, and 2 staff), which could be amplified by the devices' portability.

Usability and Accessibility

The usability and accessibility of the devices were key factors impacting participants' use of WVTs. In this study, 31 out of 39 (nearly 80%) participants identified physical (eg, arthritis) or mental disability (eg, dementia) as the primary reason impeding their ease of using WVTs for social connectedness. Both staff and FMs described older adults' difficulties in comprehending the concept of videoconferencing, which often resulted in disengagement during the video call. Despite this, FMs still appreciated the ability to see their loved ones on screen, even if for a brief period. However, this stimulated FMs to seek other easier-to-use WVTs requiring only the FM to start and end a video call.

In addition to physical or cognitive abilities, digital literacy also varied among participants. Staff noted that the pandemic catalyzed WVT use, and as a result, they had to quickly adapt and engage in learning to use the technologies introduced by their workplaces. Families also engaged in mobile learning either individually or with the help of their technologically literate friends or family. These FMs often assumed responsibility for teaching older adults living in LTCHs how to use WVTs. For those older adults who were particularly isolated or for FMs who lacked support, staff members became responsible for teaching, setting up, or troubleshooting the WVTs for them. One staff member shared a story about teaching a husband outside the LTCH how to download and set up Skype so that he could speak to his isolated wife in the LTCH during the pandemic. Some reported that the assessment for the family and older adult's capacity to use technology and planning to support the ease of use would usually take place during the resident admission into LTCH (2 staff), although this was not always the case. Findings also suggest that some families were unaware they could set up video calls upon admission (3 family members).

Surprisingly, only 1 FM explicitly received special training on accessibility features embedded into the device through their workplace to enhance its ease of use for people with varied cognitive and physical abilities. Moreover, only a few (3 FMs and 1 staff) were aware of the accessibility features and how to enable them for easier device use.

Impact on the Workload

The ease of WVT use impacted the workload among staff within LTCHs. Setting up and supporting older adults with videoconferencing was seen as an additional task to do and required coordination between more than 1 health care team, which led to some resistance to use WVTs among staff. One staff member deployed to assist residents with setting up video calls shared:

> *I wouldn't get the residents ready myself; I would ask a personal support worker to get the resident ready. A lot of times I would go to the floor and some of the residents would still be in bed, so I'd be like: "Someone needs to get this person ready for me because it's out of my scope.* [S05]

Another frontline staff member shared that not all staff were willing to facilitate the video calls*:*

> *There's one resident who has her own tablet, who will ask us: "Can you call my daughter?" And we will go and do it for her. Now, when I say we, it's mostly me and one other person. A lot of people don't want to stop and facilitate that because a lot of times there's troubleshooting with a tablet. And people aren't always comfortable doing that kind of thing. Or they don't feel like it's their job to do.* [S09]

Interestingly, staff reticence to assist with web-based videoconferencing technologies use was received with understanding among family and older adults (1 older adult and 1 FM) with some noting that staff already had many work demands and were too overworked to be helping them with videoconferencing. However, 1 FM also added that contextualizing the reason for a video call (eg, social connection) could overcome the staff's resistance to help with setting it up and staying with the person for the duration of a video call if needed:

> *[…] if you had a palliative resident in the middle of the night that wanted to say last good-byes to somebody, wouldn't it be wonderful to be able to connect them virtually with somebody that couldn't be there. I think if it were put in the context like that, it could be woven into some of the other training that you give your nursing staff.* [FM21]

The perceived ease of use of WVTs in this study was influenced by a range of factors, from choosing the ergonomically suitable device to calibrating it based on unique needs and integrating its use into a care plan given the realities of staff workload.

Discussion

Principal Findings

This qualitative study examined perceptions of usefulness and ease of using web-based videoconferencing hardware and software to facilitate social connectedness between older adults living in LTCHs and FMs. In-depth interviews conducted with 3 participant groups—older adults living in LTCHs, FMs, and LTCH staff—suggest that many older adults and FMs benefited from WVTs, and families were creative in finding workable solutions to facilitate older adults' use of technology despite their potential physical and cognitive limitations. Findings informed by the TAM model were collated in 3 overarching themes: actual system use, perceived usefulness, and perceived ease of use to help inform understanding of WVT use in LTCHs. Key findings in this study include introduction of the WVTs during the COVID-19 pandemic positively impacted emotional

XSL•FO
RenderX

well-being of socially isolated older adults in LTCHs and their FMs when in-person visitations were restricted; design and ergonomics greatly influenced the use of WVTs, and staff along with FMs were instrumental in finding ways to make WVTs useful and easy to use, especially for older adults with cognitive or physical impairments; and there was a notable decline in WVT use in LTCHs for social connectedness post COVID-19 pandemic despite heavy reliance on WVT programs to maintain a sense of connectedness during the pandemic.

Actual System Use

Besides commonly used WVTs like iPads or Zoom platform, FMs and staff participants in this study who provided care to the older adults living with dementia used several devices such as Portal, ViewClix, Alexa Echo Show, and Signal, which are understudied in the literature. For instance, a literature search yielded only 1 study exploring Alexa Echo Show to maintain social connection among older adults without impaired cognitive abilities during the COVID-19 pandemic [62]. Smartphones, iPads, or laptop devices, as well as FaceTime, Skype, or Zoom digital software are the most commonly used WVTs reported in the literature to facilitate social engagement and residents' emotional well-being [25,29,36,41,63-66]. These hardware and software are useful for many people; however, they might not have embedded accessibility features required for people with special needs, which was noted in this study. In addition, WVT users may require additional assistance, either provided by staff (which may not be realistic to integrate in the current workflow) or due to the variable availability of informal support by friends or family.

Participant experiences highlight the limited number of formal roles (eg, recreational activity specialists), resources (eg, tailored technologies and their maintenance), and programs (eg, digital literacy workshops) to support ongoing WVT use. The lack of WVT use for social connectedness, particularly post pandemic, signals an emphasis on in-person activities and, concomitantly, ongoing staff shortages, which are also reported in other research [67]. The sustained use of WVTs for social connectedness in this study was found among those who were most familiar and comfortable with technology and who were separated by great geographical distances, a finding corroborating other research [66]. However, there were FMs of older adults in this study who were unaware of the ongoing availability of WVTs despite expressed interest in using this modality. This indicates a growing preference for multiple modes of fostering social connection with those residing in LTCHs.

Perceived Usefulness

Study findings demonstrated that LTCH staff valued using WVTs with families as a means to enhance social presence and for its positive effects on older adults' well-being, even if they had advanced dementia and were unable to fully comprehend how videoconferencing functioned, which is echoed in other research [36,63,67,68]. Participants in this study and other research [69] reported increased usefulness of WVTs for social connectedness facilitated by LTCH staff or family members, as it allowed older adults to remain remotely connected with their families despite restrictions on in-person visits. Although some of the older adults with more pervasive cognitive shortcomings found it challenging to engage in conversation, videoconferencing experiences still elicited positive reactions and emotions, in accord with findings reported in the literature [36,63,67]. Despite the potential challenges of WVT use, they were perceived as a valuable tool to support well-being, foster engagement between older adults and their family, and enhance patient-centered care [67].

Notable was the lack of clarity voiced by some participants about the availability of and ability to use WVTs, particularly once the acuity of the pandemic had subsided. Although WVT use declined post pandemic in LTCHs, older adults and their FMs described it as a useful and desired service for ongoing use. To achieve optimum quality of life while residing in LTCHs, older adults and their FMs expect greater opportunities for social interactions and higher quality and quantity of family engagement [70], which can be addressed by using WVTs. For example, participants in this study shared that videoconferencing allowed them to stay connected to their life and maintain habits established before moving to LTCHs (eg, virtually visiting the garden they used to walk through and interacting with their pet). This propensity of WVTs to provide access to other contexts may foster a sense of aging in place [71] among older adults who are living in LTCHs.

In addition to supporting social well-being, WVTs may also have an important clinical role, in the form of telemedicine, as these modalities reduce time needed to travel, allow timely access to specialized health care personnel [37] and are a cost-effective solution that could alleviate the need for emergent visits [72]. The added examples of usefulness make it worthwhile not only to sustain but also expand WVT service [73].

Perceived Ease of Use

Participants in this study independently researched and shared their knowledge of Portal, ViewClix, or Alexa Echo Show. These WVTs were perceived as easier to use for people with cognitive impairment (eg, dementia) than iPads. Although iPads have been commonly leveraged for their convenience and wide societal acceptance, these types of mobile devices often presented challenges for those older adults who had cognitive or physical disabilities or who lacked experience in using WVTs. In line with this finding, higher tablet or iPad use is predicted by factors such as younger age, higher cognitive functioning, and absence of hearing impairment [66]. The use of devices such as Portal, ViewClix, and Alexa Echo Show highlighted in this study warrants further exploration. These devices could be well-situated to address some of the identified challenges of WVT use among older adult populations and in those with increasing health complexity, such as is often found within LTCH settings [74]. This is especially relevant because a lack of social contact was identified as a risk factor for developing dementia, and preventative action is best initiated early [75].

Findings in this study also suggest that staff members noted digital literacy challenges among some FMs as well as older adults living in LTCH, which is echoed across other studies [63,67]. Low digital literacy among older adults may amplify social isolation by excluding these individuals from participating in social processes or spaces [76,77], many of which have

XSL•FO
RenderX

shifted their operation to digital spaces in recent years. Moreover, since in many instances the FMs were pivotal for initiating the video calls [36], low digital literacy among them resulted in additional responsibility being assumed by LTCH staff who were required to provide education and manual support [36,67]. These findings highlight the need for broad approaches to foster widespread WVT uses among older adults that consider cost, digital literacy, individual capabilities, usability of these tools, and training and use requirements.

Although WVTs are challenging to use for some older adults [64], looking to the future, more people will likely expect to use and be supported in using WVTs in LTCHs as younger generations of society use these devices regularly and consistently within their daily lives. The importance of digital literacy training and accessibility features to facilitate independent use of WVTs by older adults is documented in the literature [78]. This study also identified limited awareness about, and knowledge of, accessibility features embedded in the WVTs by participants which, in turn, could enhance the ease of their use. Although previous works have identified low digital literacy among older adults [64] and a lack of training available within LTCHs to guide their use of WVTs [36,66], there remains a gap in research regarding implementing digital literacy training programs with older adults in LTCHs and their families. In fact, research suggests that in some cases, FMs have been the ones to take initiative and provide the technology necessary for WVTs within LTCHs [79]. Moving forward, additional funding and education to increase technology availability, enhance digital literacy, and increase staffing support roles within LTCH are warranted.

Staff participants in this study and in others [36,63,67] expressed concern regarding their workload and challenges to integrate videoconferencing with their other roles and responsibilities. Interestingly, a few FM participants in this study were accepting of limited opportunities to connect with their older adult member using WVTs due to staffing shortfalls. It is plausible that widespread awareness of the workload and employment challenges common within LTCHs across Canada contributed to their acceptance [80]. This finding demonstrates families' awareness and compassion for those employed within LTCHs, but it also highlights the ongoing paucity of resources directed toward supporting the psychosocial well-being of older adults within LTCHs. In addition, it highlights the lack of policy to support widespread use of WVTs in LTCHs. For instance, burdened staff being too busy to provide assistance with WVT use is documented in the literature [66,79]. Institutional supports are paramount to sustain the WVT service for social connectedness and include additional staffing and redefining roles and responsibilities among LTCH staff to foster digital literacy among older adults in LTCHs or, alternatively, to assist older adults and their families in using WVTs [81]. Addressing these factors simultaneously may facilitate smoother integration of this service into LTCHs while limiting the burden on older adults, their families, and staff.

Implications

Findings from this study highlight the importance of a breadth and depth of approaches to effectively implement and sustain the integration of WVT programs for social connectedness within LTCHs. Government and the private sectors must allocate funding to allow for technology procurement and infrastructure improvements to support wide-scale WVT use within LTCHs. This could be achieved by targeted funding grants or through capital funding initiatives. To maintain WVT use over time, policies and funding to guide its use are also required. Recognition that personnel is required for successful long-term use is paramount if older adults and families are going to benefit. There could be opportunities for expanded roles or even volunteer positions to help older adults use WVTs. In terms of choosing the WVTs, guidelines that ensure accessibility and user-centeredness for individuals with cognitive or physical impairment are required to ensure equitable use of WVTs. In anticipation of the increasingly digitalized health care and potential future increase in using WVTs by individuals in our society, additional roles for staff should be created to be able to support, sustain, and expand WVT programs for social connectedness and health care in LCTHs.

Interdisciplinary care teams in LTCHs can leverage WVTs in their daily routines or activities to improve older adults' physical and psychosocial well-being and create opportunities for more family engagement. Care teams should create opportunities for older adults and their families to engage in digital literacy workshops to optimize benefits from using WVTs by a greater number of older adults and their families. Staff members should also engage in digital literacy training with a particular focus on troubleshooting WVTs and innovating ways they can incorporate WVTs into their daily activities to promote a sense of meaningful social connectedness among older adults residing in LTCHs.

This study highlights the need to explore long-term impacts on health and well-being of using WVTs with residents and their families. Future studies should focus on exploration and comparison of WVTs to determine the most appropriate WVTs to promote social connectedness, especially for individuals living with cognitive or physical impairments. In addition, metrics on the impact of using WVTs for older adults' well-being should be collected to determine different aspects of WVTs' usefulness for older adults in LTCHs. Future research studies should also explore or develop sustainable models to enhance digital literacy training for older adults and their families.

Strengths and Limitations

A strength of this study is the robust participant sample representing multiple perspectives and experiences in using WVTs for social connectedness. Although staff participants were all female, this parallels research suggesting that 90% of personal support workers employed in Ontario are female and 75% of care workers across Canada are female [82,83]. In addition, collaborating with 3 LTCH sites for participant recruitment supported collection of a diverse range of experiences across the different settings. This mitigated a risk for biased results that would be imposed by the nature and operations of a single facility.

While efforts were made to recruit a range of older participants living in LTCHs, we recognize that the inclusion criteria of mild

XSL•FO
RenderX

or no cognitive impairment constrained the participation of older adults. Those who had advanced dementia or other severe neurocognitive deficits were not included, which likely represented a notable portion of those residing in LTCHs who live with dementia [84]. Those older adults who had mild deficits or other physical limitations that impacted their ability to participate were offered assistance in setting up the technology so they could still participate. The findings revealed that frontline staff, particularly nurses and personal support workers, were instrumental in facilitating WVT use among older adults and their families, although this responsibility posed potential burdens on staff workload. Future work could address targeted personnel support for WVT use, including the potential for volunteer technology facilitation or expanding frontline provider roles to include assisting older adults to use technology such as WVTs. Experiences of staff in various LTCH facilities may differ in terms of digital literacy training provided to staff or role expectations. Moving forward, it would be helpful to have clear guidelines to inform staff roles in providing technology support, particularly in the case of supporting older adults' well-being, including social connectedness. In addition, this study did not delve into other factors that may contribute to the decline in using WVTs post pandemic, like for instance, the complex psychosocial factors that influence technology adoption.

Conclusions

The experiences of older adults residing in LTCHs, FMs, and staff demonstrate that using WVTs for social connectedness positively impacts older adults and their FMs' emotional and social well-being. However, WVTs service needs to be tailored to the needs of the families, including the choice of the device, digital literacy training, and provision of human resources to support connections. Moving forward, LTCHs should develop formal programs that allow for integration of WVTs service to expand the opportunities for older adults in the LTCHs to connect socially with their families or use WVTs in other ways that create a sense of social connectedness. To achieve this, funding initiatives such as capital improvement funds and clearer policies on the roles of personnel support will be required. Prospective studies should implement WVTs in collaboration with other actors participating in the process of socially connecting people with WVTs, such as technology industry partners, companies allowing people to participate in or attend leisure activities virtually, and public health organizations to explore additional impacts of WVTs on social and clinical well-being of older adults in LTCHs and their families.

Acknowledgments

The authors acknowledge Jessica Burford and Christina Oleynikov for their assistance with this project.

Funding

This study was supported by a Social Sciences and Humanities Research Council Insight Development Grant from the Government of Canada. The funders had no role in the study's design; collection, analyses, or interpretation of data; writing of the manuscript; or decision to publish the results.

Authors' Contributions

Conceptualization: AG
Data curation: AG (lead), HY (supporting)
Formal analysis: AG (equal), HY (equal), RB (supporting), LD (supporting), DMC (supporting)
Funding acquisition: AG (lead), RB (supporting), LD (supporting), DMC (supporting)
Investigation: AG (lead), HY (supporting)
Methodology: AG
Project administration: AG (lead), HY (supporting), RB (supporting), LD (supporting), DMC (supporting)
Resources: AG
Supervision: AG
Validation: AG
Visualization: AG (lead), HY (supporting)
Writing–original draft: AG (equal), HY (equal)
Writing–review & editing: AG (lead), HY (supporting), RB (supporting), LD (supporting), DMC (supporting)

Conflicts of Interest

None declared.

References

1. Aburto JM, Schöley J, Kashnitsky I, et al. Quantifying impacts of the COVID-19 pandemic through life-expectancy losses: a population-level study of 29 countries. Int J Epidemiol 2022 Feb 18;51(1):63-74. [doi: 10.1093/ije/dyab207] [Medline: 34564730]

XSL•FO
RenderX

2. Wang H, Paulson KR, Pease SA. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21. Lancet 2022 Apr 16;399(10334):1513-1536. [doi: 10.1016/S0140-6736(21)02796-3] [Medline: 35279232]
3. Felipe SGB, Parreira Batista P, da Silva CCR, de Melo RC, de Assumpção D, Perracini MR. Impact of COVID-19 pandemic on mobility of older adults: a scoping review. Int J Older People Nurs 2023 Jan;18(1):e12496. [doi: 10.1111/opn.12496] [Medline: 35941278]
4. Trevissón-Redondo B, López-López D, Pérez-Boal E, et al. Use of the Barthel index to assess activities of daily living before and after SARS-COVID 19 Infection of institutionalized nursing home patients. Int J Environ Res Public Health 2021 Jul 7;18(14):7258. [doi: 10.3390/ijerph18147258] [Medline: 34299709]
5. Şahiner TA, Şahiner M, Cinar N, Beynex Research Group. In Covid‐19 pandemic, quarantine measures affect daily living activities in normal individuals, but do not change the cognitive state. Alzheimer's Dementia 2021 Dec;17(S10):1-4. [doi: 10.1002/alz.055639] [Medline: 34971285]
6. Champlin C, Sirenko M, Comes T. Measuring social resilience in cities: an exploratory spatio-temporal analysis of activity routines in urban spaces during Covid-19. Cities 2023 Apr;135:104220. [doi: 10.1016/j.cities.2023.104220] [Medline: 36743889]
7. Erfani G, Bahrami B. COVID and the home: the emergence of new urban home life practised under pandemic-imposed restrictions. Cities & Health 2023 Jul 4;7(4):548-555. [doi: 10.1080/23748834.2022.2029241]
8. Breaux R, Cash AR, Lewis J, Garcia KM, Dvorsky MR, Becker SP. Impacts of COVID-19 quarantine and isolation on adolescent social functioning. Curr Opin Psychol 2023 Aug;52:101613. [doi: 10.1016/j.copsyc.2023.101613] [Medline: 37364468]
9. Hugelius K, Harada N, Marutani M. Consequences of visiting restrictions during the COVID-19 pandemic: an integrative review. Int J Nurs Stud 2021 Sep;121:104000. [doi: 10.1016/j.ijnurstu.2021.104000] [Medline: 34242976]
10. Killgore WDS, Cloonan SA, Taylor EC, Lucas DA, Dailey NS. Loneliness during the first half-year of COVID-19 lockdowns. Psychiatry Res 2020 Dec;294:113551. [doi: 10.1016/j.psychres.2020.113551] [Medline: 33171416]
11. Seppala E, Rossomando T, Doty JR. Social connection and compassion: important predictors of health and well-being. Soc Res 2013 Jun;80(2):411-430. [doi: 10.1353/sor.2013.0027]
12. Lee RM, Robbins SB. Measuring belongingness: the social connectedness and the social assurance scales. J Couns Psychol 1995;42(2):232-241. [doi: 10.1037/0022-0167.42.2.232]
13. Townsend KC, McWhirter BT. Connectedness: a review of the literature with implications for counseling, assessment, and research. J Couns Dev 2005 Apr;83(2):191-201. [doi: 10.1002/j.1556-6678.2005.tb00596.x]
14. Parsons MA. "I miss my friends, but I also miss strangers": pandemic loneliness and the importance of place and practice. SSM Ment Health 2022 Dec;2:100121. [doi: 10.1016/j.ssmmh.2022.100121]
15. Wickramaratne PJ, Yangchen T, Lepow L, et al. Social connectedness as a determinant of mental health: a scoping review. PLoS ONE 2022;17(10):e0275004. [doi: 10.1371/journal.pone.0275004] [Medline: 36228007]
16. Holt-Lunstad J. Social connection as a critical factor for mental and physical health: evidence, trends, challenges, and future implications. World Psychiatry 2024 Oct;23(3):312-332. [doi: 10.1002/wps.21224] [Medline: 39279411]
17. Neves BB, Colón Cabrera D, Sanders A, Warren N. Pandemic diaries: lived experiences of loneliness, loss, and hope among older adults during COVID-19. Gerontologist 2023 Jan 24;63(1):120-130. [doi: 10.1093/geront/gnac104] [Medline: 35868621]
18. Lorber M, Černe Kolarič J, Kmetec S, Kegl B. Association between loneliness, well-Being, and life satisfaction before and during the COVID-19 pandemic: a cross-sectional study. Sustainability 2023;15(3):2825. [doi: 10.3390/su15032825]
19. Smith CB, Wong KLY, To-Miles F, et al. Exploring experiences of loneliness among Canadian long-term care residents during the COVID-19 pandemic: a qualitative study. Int J Older People Nurs 2023 Jan;18(1):e12509. [doi: 10.1111/opn.12509] [Medline: 36347829]
20. Kadowaki L, Wister A. Older adults and social isolation and loneliness during the COVID-19 pandemic: an integrated review of patterns, effects, and interventions. Can J Aging 2023 Jun;42(2):199-216. [doi: 10.1017/S0714980822000459] [Medline: 36345649]
21. Prommas P, Lwin KS, Chen YC, et al. The impact of social isolation from COVID-19-related public health measures on cognitive function and mental health among older adults: a systematic review and meta-analysis. Ageing Res Rev 2023 Mar;85:101839. [doi: 10.1016/j.arr.2022.101839] [Medline: 36596396]
22. The impact of COVID-19 on long-term care in canada: focus on the first 6 months. Canadian Institute for Health Information. URL: https://www.cihi.ca/en/long-term-care-and-covid-19-the-first-6-months [accessed 2025-12-01]
23. Paananen J, Rannikko J, Harju M, Pirhonen J. The impact of Covid-19-related distancing on the well-being of nursing home residents and their family members: a qualitative study. Int J Nurs Stud Adv 2021 Nov;3(May):100031. [doi: 10.1016/j.ijnsa.2021.100031] [Medline: 34095858]
24. Stall NM, Brown KA, Maltsev A, et al. COVID‐19 and Ontario's long‐term care homes. J Elder Policy 2021 Sep;1(3):65-110. [doi: 10.18278/jep.1.3.3]

XSL•FO
RenderX

25. Gallistl V, Seifert A, Kolland F. COVID-19 as a “digital push?” research experiences from long-term care and recommendations for the post-pandemic era. Front Public Health 2021;9:660064. [doi: 10.3389/fpubh.2021.660064] [Medline: 34041216]
26. Banskota S, Healy M, Goldberg EM. 15 smartphone apps for older adults to use while in isolation during the COVID-19 pandemic. West J Emerg Med 2020 Apr 14;21(3):514-525. [doi: 10.5811/westjem.2020.4.47372] [Medline: 32302279]
27. Lemaire C, Humbert C, Sueur C, Racin C. Use of digital technologies to maintain older adults' social ties during visitation restrictions in long-term care facilities: scoping review. JMIR Aging 2023 Feb 10;6:e38593. [doi: 10.2196/38593] [Medline: 36599164]
28. Mitchell LL, Albers EA, Birkeland RW, et al. Caring for a relative with dementia in long-term care during COVID-19. J Am Med Dir Assoc 2022 Mar;23(3):428-433. [doi: 10.1016/j.jamda.2021.11.026] [Medline: 34929196]
29. Kelly RM, Xing Y, Baker S, Waycott J. Video calls as a replacement for family visits during lockdowns in aged care: interview study with family members. JMIR Aging 2023 Jun 12;6:e40953. [doi: 10.2196/40953] [Medline: 37191951]
30. Hwang TJ, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. Int Psychogeriatr 2020 Oct;32(10):1217-1220. [doi: 10.1017/S1041610220000988] [Medline: 32450943]
31. Kuksa I, Childs M. Remediating technology, translating experience, immersing in spaces. In: Kuksa I, Childs M, editors. Making Sense of Space: Chandos Publishing; 2014:3-22. [doi: 10.1533/9781780634067.1.3]
32. Khosravi P, Rezvani A, Wiewiora A. The impact of technology on older adults' social isolation. Comput Human Behav 2016 Oct;63:594-603. [doi: 10.1016/j.chb.2016.05.092]
33. Garnett A, Pollock H, Floriancic N, et al. Social connectedness between family caregivers and older adults living in long-term care homes in the context of COVID-19. Can J Aging 2024 Mar;43(1):33-44. [doi: 10.1017/S0714980823000351] [Medline: 37727879]
34. Hart JL, Turnbull AE, Oppenheim IM, Courtright KR. Family-centered care during the COVID-19 era. J Pain Symptom Manage 2020 Aug;60(2):e93-e97. [doi: 10.1016/j.jpainsymman.2020.04.017] [Medline: 32333961]
35. Hardy MS, Fanaki C, Savoie C, et al. Acceptability of videoconferencing to preserve the contact between cognitively impaired long-term care residents and their family caregivers: a mixed-methods study. Geriatr Nurs 2022;48:65-73. [doi: 10.1016/j.gerinurse.2022.09.006] [Medline: 36155311]
36. Naudé B, Rigaud AS, Pino M. Video calls for older adults: a narrative review of experiments involving older adults in elderly care institutions. Front Public Health 2021;9:751150. [doi: 10.3389/fpubh.2021.751150] [Medline: 35096731]
37. Tan AJQ, Rusli KDB, McKenna L, Tan LLC, Liaw SY. Telemedicine experiences and perspectives of healthcare providers in long-term care: a scoping review. J Telemed Telecare 2024 Feb;30(2):230-249. [doi: 10.1177/1357633X211049206] [Medline: 34666535]
38. Sixsmith A, Horst BR, Simeonov D, Mihailidis A. Older people's use of digital technology during the COVID-19 pandemic. Bull Sci Technol Soc 2022 Jun;42(1-2):19-24. [doi: 10.1177/02704676221094731] [Medline: 38603230]
39. Ang S, Lim E, Malhotra R. Health-related difficulty in internet use among older adults: correlates and mediation of its association with quality of life through social support networks. Gerontologist 2021 Jul 13;61(5):693-702. [doi: 10.1093/geront/gnaa096] [Medline: 32744310]
40. Hacker J, vom Brocke J, Handali J, Otto M, Schneider J. Virtually in this together—how web-conferencing systems enabled a new virtual togetherness during the COVID-19 crisis. Eur J Inf Syst 2020 Sep 2;29(5):563-584. [doi: 10.1080/0960085X.2020.1814680]
41. Hung L, Chow B, Shadarevian J, et al. Using touchscreen tablets to support social connections and reduce responsive behaviours among people with dementia in care settings: a scoping review. Dementia (London) 2021 Apr;20(3):1124-1143. [doi: 10.1177/1471301220922745] [Medline: 32380856]
42. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 4;8:1687. [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
43. Gell NM, Rosenberg DE, Demiris G, LaCroix AZ, Patel KV. Patterns of technology use among older adults with and without disabilities. Gerontologist 2015 Jun;55(3):412-421. [doi: 10.1093/geront/gnt166] [Medline: 24379019]
44. Garnett A, Yurkiv H, Booth R, Connelly D, Donelle L. Web-based presence for social connectedness in long-term care: protocol for a qualitative multimethods study. JMIR Res Protoc 2023 Oct 27;12(1):e50137. [doi: 10.2196/50137] [Medline: 37889518]
45. Sandelowski M. Whatever happened to qualitative description? Res Nurs Heal 2000;23(4):334-340. [doi: 10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G] [Medline: 10940958]
46. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
47. Phillippi J, Lauderdale J. A guide to field notes for qualitative research: context and conversation. Qual Health Res 2018 Feb;28(3):381-388. [doi: 10.1177/1049732317697102] [Medline: 29298584]
48. Colorafi KJ, Evans B. Qualitative descriptive methods in health science research. HERD 2016 Jul;9(4):16-25. [doi: 10.1177/1937586715614171] [Medline: 26791375]
49. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]

XSL•FO
RenderX

50. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Jul;9(3):604-634. [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
51. Kim H, Sefcik JS, Bradway C. Characteristics of qualitative descriptive studies: a systematic review. Res Nurs Health 2017 Feb;40(1):23-42. [doi: 10.1002/nur.21768] [Medline: 27686751]
52. Doyle L, McCabe C, Keogh B, Brady A, McCann M. An overview of the qualitative descriptive design within nursing research. J Res Nurs 2020 Aug;25(5):443-455. [doi: 10.1177/1744987119880234] [Medline: 34394658]
53. Rourke S. How does virtual reality simulation compare to simulated practice in the acquisition of clinical psychomotor skills for pre-registration student nurses? A systematic review. Int J Nurs Stud 2020 Feb;102:103466. [doi: 10.1016/j.ijnurstu.2019.103466] [Medline: 31783192]
54. Perdomo Delgato C, Ripat J, Mallory-Hill S, Bohunicky S. Using the technology acceptance model to explore the user experience of smart kitchen use among older adult women. In: Marcus A, Rosenzweig E, Soares MM, editors. Design, User Experience, and Usability: 13th International Conference, DUXU 2024, Held as Part of the 26th HCI International Conference, HCII 2024, Washington, DC, USA, June 29 – July 4, 2024, Proceedings, Part III: Springer:3-12. [doi: 10.1007/978-3-031-61356-2_1]
55. Bradshaw C, Atkinson S, Doody O. Employing a qualitative description approach in health care research. Glob Qual Nurs Res 2017;4:2333393617742282. [doi: 10.1177/2333393617742282] [Medline: 29204457]
56. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
57. NVivo 13. Lumivero. 2023. URL: https://lumivero.com/products/nvivo/ [accessed 2023-06-01]
58. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
59. Assarroudi A, Heshmati Nabavi F, Armat MR, Ebadi A, Vaismoradi M. Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. J Res Nurs 2018 Feb;23(1):42-55. [doi: 10.1177/1744987117741667] [Medline: 34394406]
60. Johnson JL, Adkins D, Chauvin S. A review of the quality indicators of rigor in qualitative research. Am J Pharm Educ 2020 Jan;84(1):7120. [doi: 10.5688/ajpe7120] [Medline: 32292186]
61. Creswell JW, Miller DL. Determining validity in qualitative inquiry. Theory Pract 2000 Aug;39(3):124-130. [doi: 10.1207/s15430421tip3903_2]
62. Soubutts E, Ayobi A, Eardley R, McNaney R, Cater K, O'Kane AA. Amazon Echo Show as a multimodal human-to-human care support tool within self-isolating older UK households. Proc ACM Hum-Comput Interact 2022 Nov 7;6(CSCW2):1-31. [doi: 10.1145/3555193]
63. Zamir S, Hennessy C, Taylor A, Jones R. Intergroup "Skype" quiz sessions in care homes to reduce loneliness and social isolation in older people. Geriatrics (Basel) 2020 Nov 11;5(4):90. [doi: 10.3390/geriatrics5040090] [Medline: 33187242]
64. Moyle W, Jones C, Murfield J, Liu F. 'For me at 90, it's going to be difficult': feasibility of using iPad video-conferencing with older adults in long-term aged care. Aging Ment Health 2020 Feb 1;24(2):349-352. [doi: 10.1080/13607863.2018.1525605] [Medline: 30621436]
65. Tsai HH, Cheng CY, Shieh WY. Effectiveness of laptop-based versus smartphone-based videoconferencing interaction on loneliness, depression and social support in nursing home residents: a secondary data analysis. J Telemed Telecare 2023 Apr;29(3):177-186. [doi: 10.1177/1357633X20972004] [Medline: 33197365]
66. Astell A, Dosanjh S, D'Elia T, et al. Personalized tablets for residents in long-term care to support recreation and mitigate isolation. J Am Med Dir Assoc 2024 Jul;25(7):105022. [doi: 10.1016/j.jamda.2024.105022] [Medline: 38763162]
67. Hardy MS, Fanaki C, Savoie C, et al. Long-term care staffs' experience in facilitating the use of videoconferencing by cognitively impaired long-term care residents during the COVID-19 pandemic: a mixed-methods study. BMC Health Serv Res 2024 May 21;24(1):646. [doi: 10.1186/s12913-024-11095-9] [Medline: 38769512]
68. Schmidt SB, Isbel S, John B, Subramanian R, D'Cunha NM. Examining technology perspectives of older adults with mild cognitive impairment: scoping review. JMIR Aging 2025 Oct 30;8(1):e78229. [doi: 10.2196/78229] [Medline: 41166548]
69. Nguyen M, Fujioka J, Wentlandt K, et al. Using the technology acceptance model to explore health provider and administrator perceptions of the usefulness and ease of using technology in palliative care. BMC Palliat Care 2020 Sep 7;19(1):138. [doi: 10.1186/s12904-020-00644-8] [Medline: 32895060]
70. Liougas MP, Sommerlad A, O'Rourke HM, et al. Assessing social connection for long‐term care home residents: a scoping review of measure content. Alzheimers Dement (N Y) 2024;10(3):e12488. [doi: 10.1002/trc2.12488] [Medline: 39315314]
71. Ratnayake M, Brathwaite S, Lukas S, et al. Aging in place: are we prepared? Dela J Public Health 2022 Aug;8(3):28-31. [doi: 10.32481/djph.2022.08.007] [Medline: 36177171]
72. De Guzman KR, Snoswell CL, Caffery LJ, Smith AC. Economic evaluations of videoconference and telephone consultations in primary care: a systematic review. J Telemed Telecare 2024 Jan;30(1):3-17. [doi: 10.1177/1357633X211043380] [Medline: 34617819]
73. Grant KL, Lee DD, Cheng I, Baker GR. Reducing preventable patient transfers from long-term care facilities to emergency departments: a scoping review. CJEM 2020 Nov;22(6):844-856. [doi: 10.1017/cem.2020.416] [Medline: 32741417]

XSL•FO
RenderX

74. Ng R, Lane N, Tanuseputro P, et al. Increasing complexity of new nursing home residents in Ontario, Canada: a serial cross-sectional study. J Am Geriatr Soc 2020 Jun;68(6):1293-1300. [doi: 10.1111/jgs.16394] [Medline: 32119121]
75. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet Standing Commission. Lancet 2024 Aug 10;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
76. Zapletal A, Wells T, Russell E, Skinner MW. On the triple exclusion of older adults during COVID-19: technology, digital literacy and social isolation. Soc Sci Humanit Open 2023;8(1):100511. [doi: 10.1016/j.ssaho.2023.100511] [Medline: 37021073]
77. Reneland-Forsman L. 'Borrowed access'—the struggle of older persons for digital participation. In J Lifelong Educ 2018 May 4;37(3):333-344. [doi: 10.1080/02601370.2018.1473516]
78. Kelly RM, Xing Y, Baker S, Waycott J. "More than just holding the iPad": family Members' perspectives on the work of video calling in aged care. Proc ACM Hum-Comput Interact 2024 Nov 7;8(CSCW2):1-26. [doi: 10.1145/3687007]
79. Chu CH, Yee A, Stamatopoulos V. Poor and lost connections: essential family caregivers' experiences using technology with family living in long-term care homes during COVID-19. J Appl Gerontol 2022 Jun;41(6):1547-1556. [doi: 10.1177/07334648221081850] [Medline: 35416076]
80. Long-term care and COVID-19: report of a special task force prepared for the Chief Science Advisor of Canada. : Government of Canada; 2020 URL: https://science.gc.ca/site/science/en/office-chief-science-advisor/initiatives-covid-19/long-term-care-and-covid-19 [accessed 2020-10-29]
81. Newbould L, Ariss S, Mountain G, Hawley MS. Exploring factors that affect the uptake and sustainability of videoconferencing for healthcare provision for older adults in care homes: a realist evaluation. BMC Med Inform Decis Mak 2021 Jan 6;21(1):13. [doi: 10.1186/s12911-020-01372-y] [Medline: 33407437]
82. Long-term care staffing study. : Ontario Ministry of Long-Term Care; 2020 URL: https://files.ontario.ca/mltc-long-term-care-staffing-study-en-2020-07-31.pdf [accessed 2025-11-15]
83. Khanam F, Langevin M, Savage K, Uppal S. Women working in paid care occupations. : Statistics Canada; 2022 URL: https://www150.statcan.gc.ca/n1/pub/75-006-x/2022001/article/00001-eng.pdf [accessed 2025-11-15]
84. Dementia in long-term care. Canadian Institute for Health Information. 2025. URL: https://www.cihi.ca/en/dementia-in-canada/dementia-care-across-the-health-system/dementia-in-long-term-care [accessed 2025-11-15]

Abbreviations

FM: family member
LTCH: long-term care home
TAM: technology acceptance model
WVT: web-based videoconferencing technology

Edited by Y Luo; submitted 27.Feb.2025; peer-reviewed by CK Tan, R Kelly; revised version received 04.Nov.2025; accepted 19.Nov.2025; published 09.Jan.2026.

Please cite as:
Garnett A, Yurkiv H, Connelly DM, Booth R, Donelle L
Perceived Use of Web-Based Videoconferencing for Social Connectedness Among Older Adults Living in Long-Term Care: Qualitative Study
JMIR Aging 2026;9:e73213
URL: https://aging.jmir.org/2026/1/e73213
doi:10.2196/73213

© Anna Garnett, Halyna Yurkiv, Denise M Connelly, Richard Booth, Lorie Donelle. Originally published in JMIR Aging (https://aging.jmir.org), 9.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Between Loneliness, eHealth Literacy, and Quality of Life Among Chinese Older Adults: Cross-Sectional Study

Shouchuang Zhang[1], PhD; Jiayi Li[1], BS; Jia Tang[2,3], PhD; Weiyan Jian[1,4], PhD; Qing Liu[5], PhD; Jing Guo[1], PhD

[1]Department of Health Policy and Management, School of Public Health, Peking University, 38 Xue Yuan Rd, Haidian District, Beijing, Beijing, China

[2]School of Population and Health, Renmin University of China, Beijing, Beijing, China

[3]Bigdata and Responsible Artificial Intelligence for National Governance, Renmin University of China, Beijing, Beijing, China

[4]Key Laboratory of Health System Reform and Governance, National Health Commission of China, Beijing, Beijing, China

[5]General Practice Department, Second Outpatient Section, Peking University Third Hospital, Beijing, Beijing, China

Corresponding Author:
Jing Guo, PhD
Department of Health Policy and Management, School of Public Health, Peking University, 38 Xue Yuan Rd, Haidian District, Beijing, Beijing, China

Abstract

Background: Loneliness is a growing public health challenge among older adults and is associated with a wide range of adverse health outcomes. The role of eHealth literacy in shaping the relationship between loneliness and quality of life remains unclear.

Objective: This study aimed to examine the contributing role of eHealth literacy in moderating the associations between loneliness and quality of life among older adults.

Methods: A community-based survey was conducted in 2025 among older adults aged 60 years or older living in northwestern China. A total of 2110 participants were included. Multiple linear regression and interaction models were used to assess associations and moderating effects. Sensitivity analyses were conducted by replacing the outcome variable with depressive symptoms and by performing stratified analyses to assess the robustness of the primary interaction findings.

Results: Loneliness showed a consistent negative association with quality of life (β=−0.83, 95% CI −1.18 to −0.49; *P*<.001). Higher overall eHealth literacy was associated with better quality of life (β=0.41, 95% CI 0.35-0.47; *P*<.001). Interaction models indicated that higher eHealth literacy was associated with a steeper negative association between loneliness and quality of life (β=−0.04, 95% CI −0.08 to −0.002; *P*=.04). Sensitivity analyses produced similar results across alternative outcomes and subgroups.

Conclusions: Higher loneliness was related to a poorer quality of life. Higher eHealth literacy was associated with a steeper negative association between loneliness and quality of life. These findings suggest that eHealth literacy may function as a double-edged sword in later life. Future research is needed to clarify the underlying mechanisms and to examine how digital health competencies interact with psychosocial vulnerability in shaping older adults' well-being.

(JMIR Aging 2026;9:e88600) doi:10.2196/88600

KEYWORDS

loneliness; quality of life; eHealth literacy; health of older adults; China; cross-sectional study

Introduction

Quality of life refers to an overall appraisal, both subjective and objective, based on an individual's physical health, psychological state, social relationships, and living environment [1]. It encompasses both objective resources, such as health and economic conditions, and subjective experiences, including life satisfaction and well-being [2]. With accelerating population aging, the growing burden of chronic diseases and cognitive decline constrains social participation among older adults and increases their risk of reduced quality of life [3,4]. In China, rapid urbanization and changes in family structure are reshaping intergenerational relationships [5], with adult children increasingly living apart from their parents [6]. At the same time, family care and emotional support have weakened, and the proportion of older adults living alone has risen, together compounding the risk of poorer quality of life in later life [7]. Therefore, identifying the psychological and social determinants of quality of life in older adults is essential for designing targeted interventions and shaping effective pathways to healthy aging. In addition to quality of life, depressive symptoms represent a closely related psychological outcome that reflects emotional well-being and mental health status among older adults [8]. Previous studies have shown that loneliness is strongly associated with depressive symptoms and that both constructs capture important but complementary aspects of psychological

XSL•FO
RenderX

vulnerability in later life [9,10]. Given this close conceptual linkage, depressive symptoms can serve as an alternative outcome to examine whether the observed associations are robust across related mental health indicators.

Loneliness is a subjective social emotional state that arises when individuals perceive a gap between the quantity or quality of their actual social relationships and their desired level of connection, leading to feelings of distress and loss [11]. Loneliness is now widely regarded as an important public health concern. A recent global systematic review estimated that the pooled prevalence of loneliness among community-dwelling older adults is about 32%, with substantial variation across regions and settings [12]. In China, survey data indicate that 36.6% of adults aged 60 years or older report some degree of loneliness, with higher levels observed among women, those living alone, individuals in poorer health, and those with lower socioeconomic status [13]. Previous studies have shown that loneliness is an important risk factor for a range of adverse health outcomes, including coronary heart disease and stroke [14], and is associated with a 26% to 50% increase in mortality risk [15,16]. The excess mortality attributable to loneliness is comparable in magnitude to that associated with physical inactivity or smoking [15]. Moreover, loneliness reduces subjective well-being and life satisfaction and further undermines quality of life in later life [17]. These findings highlight the need to treat loneliness as a key psychosocial determinant when examining quality of life in older adults and to clarify its pathways of influence to inform the design of targeted interventions.

eHealth literacy is commonly defined as the combined ability to locate, understand, appraise, and apply health information in internet-based and digital environments in support of health-related decision-making [18,19]. From the theoretical perspective, eHealth literacy may exert a bidirectional moderating influence on the relationship between loneliness and quality of life. According to the stress-buffering model, access to informational and emotional resources can mitigate the adverse impact of psychosocial stressors on health outcomes [20]. Within this framework, higher eHealth literacy could enable older adults experiencing loneliness to obtain trustworthy health information, access online social and community resources, and engage in self-management, thereby attenuating the detrimental effect of loneliness on quality of life. In contrast, frequent and intensive searching for health information online has been linked to heightened health anxiety and cyberchondria. Exposure to large volumes of risk-focused or conflicting information may generate additional stress and confusion [21]. For older adults who are already lonely and emotionally vulnerable, such patterns of information use may be associated with increased concerns about illness and a stronger negative association between loneliness and quality of life.

Therefore, the aims of this study are to (1) investigate the relationship between loneliness, eHealth literacy, and quality of life and (2) further explore the potential moderating role of eHealth literacy between loneliness and quality of life.

Methods

Participants and Procedure

This study was based on a community-based cross-sectional survey conducted among older adults aged 60 years and older in the Lanzhou New Area of Gansu Province, China. Lanzhou New Area is a national-level administrative district that oversees 7 community health service centers and 80 family physician teams. All 7 centers participated in the survey. Each family physician team was required to complete at least 20 household-based interviews. Among more than 30,000 older adults registered in the catchment areas of these centers, approximately 10% were randomly selected, yielding 2486 recruited participants in March 2025.

A standardized data collection procedure was implemented to maintain consistency across centers. Family physicians received centralized training before fieldwork. Questionnaires completed in less than 600 seconds were excluded, removing 290 responses. Another 86 individuals submitted duplicate entries and were therefore excluded. The final analytic sample consisted of 2110 participants..

Measures

Loneliness

Loneliness was assessed using the Three-Item Loneliness Scale [22]. The scale consists of 3 items, each rated on a 3-point response format with “1= hardly ever,” “2= some of the time,” and “3= often.” Scores from the 3 items were summed to generate a total loneliness score ranging from 3 to 9. Higher scores indicate greater perceived loneliness. The internal consistency of the scale was high, with a Cronbach α of 0.872. The full-scale content is provided in Table S1 in Multimedia Appendix 1.

Quality of Life

Quality of life was measured using the brief Older People’s Quality of Life Questionnaire, which includes 13 statements covering core domains of well-being [23]. Participants indicated their level of agreement with each statement using a 5-point response scale ranging from strong disagreement to strong agreement. Item scores were summed to generate a total score ranging from 13 to 65, with higher values reflecting better quality of life. The scale demonstrated excellent internal consistency, with a Cronbach α of 0.975. The full list of items is available in Table S2 in Multimedia Appendix 1.

eHealth Literacy

The eHealth literacy was evaluated using an 8-item instrument [24]. Each item was rated on a 5-point Likert scale ranging from very inconsistent to very consistent. Scores were summed to produce a total ranging from 8 to 40, with higher values indicating stronger eHealth literacy. In this study, the eHealth literacy scale showed excellent internal reliability, with a Cronbach α of 0.980. Following prior studies, the total eHealth literacy score was treated as a continuous variable in the primary analyses [24,25]. In addition, the exploratory analyses were conducted to examine potential heterogeneity across groups of items within the scale. For this purpose, the 8 items of the

XSL•FO
RenderX

eHealth literacy scale were grouped by the authors into 3 sets based on their content, reflecting applying literacy items (items 1‐5), critical literacy items (items 6‐7), and a decision-making literacy item (item 8). These item groupings were defined for exploratory purposes only and do not represent validated subscales of the instrument. Details of the instrument are presented in Table S3 in Multimedia Appendix 1.

Depressive Symptoms

Depressive symptoms were assessed using the 9-item Patient Health Questionnaire. This instrument contains 9 items that capture the frequency of core depressive experiences [26]. Each item is rated on a 4-level scale, ranging from 0 (not at all) to 3 (nearly every day), with response options of 1 (several days) and 2 (more than half of days). Scores were added to create a total ranging from 0 to 27, with higher values reflecting more pronounced depressive symptoms. The internal reliability of the scale was strong, with a Cronbach α of 0.921. Depressive symptoms were included as an alternative outcome variable in sensitivity analyses, replacing quality of life to assess the robustness of the primary interaction findings. Additional details are provided in Table S4 in Multimedia Appendix 1.

Potential Confounders

Demographic and socioeconomic variables controlled for in the regression models included sex (male or female), age in years (60‐69 years, 70‐79 years, or ≥80 years), marital status (single, married, or partnered), education (less than high school or high school or above), employment status (not working or working), self-reported health status (bad, general, or good), chronic diseases (no or yes), smoking (no or yes), and alcohol consumption (no or yes).

Statistical Analysis

Descriptive analyses were conducted to characterize the study sample. Continuous variables were presented as medians together with IQRs. Categorical variables were summarized using counts (n) and proportions (%).

Multiple linear regression models were fitted to examine the association between loneliness and quality of life. To evaluate the moderating role of eHealth literacy, interaction terms between loneliness and eHealth literacy were introduced. The primary analyses were conducted using the total eHealth literacy score as a continuous variable. Analyses based on item groupings of eHealth literacy were exploratory in nature and are presented in Multimedia Appendix 1. Prior to constructing the interaction terms, continuous variables involved in the interaction analyses were mean-centered to improve interpretability and reduce potential multicollinearity. To facilitate interpretation of the moderation effect, predicted values of quality of life were estimated at +1 or −1 SD of eHealth literacy, and the results were visualized. In addition, simple slope analyses were conducted to estimate the association between loneliness and quality of life at different levels of the continuous moderator. Model linearity assumptions were assessed by visually inspecting residual plots of continuous variables against regression residuals. The plots did not indicate substantial deviations from linearity. All analyses were performed using Stata (version 17.0; StataCorp). A 2-sided *P* value of <.05 was regarded as statistically significant.

To assess the robustness of the primary interaction findings based on the total eHealth literacy score, 2 sets of sensitivity analyses were performed. First, the outcome variable was replaced with depressive symptoms. Interaction models were reestimated with depressive symptoms as the outcome to examine the consistency of the moderating relationship. Second, stratified analyses were carried out with sex as the grouping variable. Robustness was evaluated based on whether the direction or statistical significance of the associations remained consistent after conducting stratified analyses [27].

Ethical Considerations

Formal ethics approval was granted by the institutional review board of Peking University Third Hospital (20250224-01025-0129). Electronic informed consent was obtained from all study participants and assisting family physicians before participation. All data were collected and stored in a deidentified and anonymized form to protect participants' privacy and confidentiality. Access to the data was restricted to authorized members of the research team only. Participation in this study was voluntary, and participants were informed of their right to withdraw at any time without consequences. No financial compensation was provided to participants.

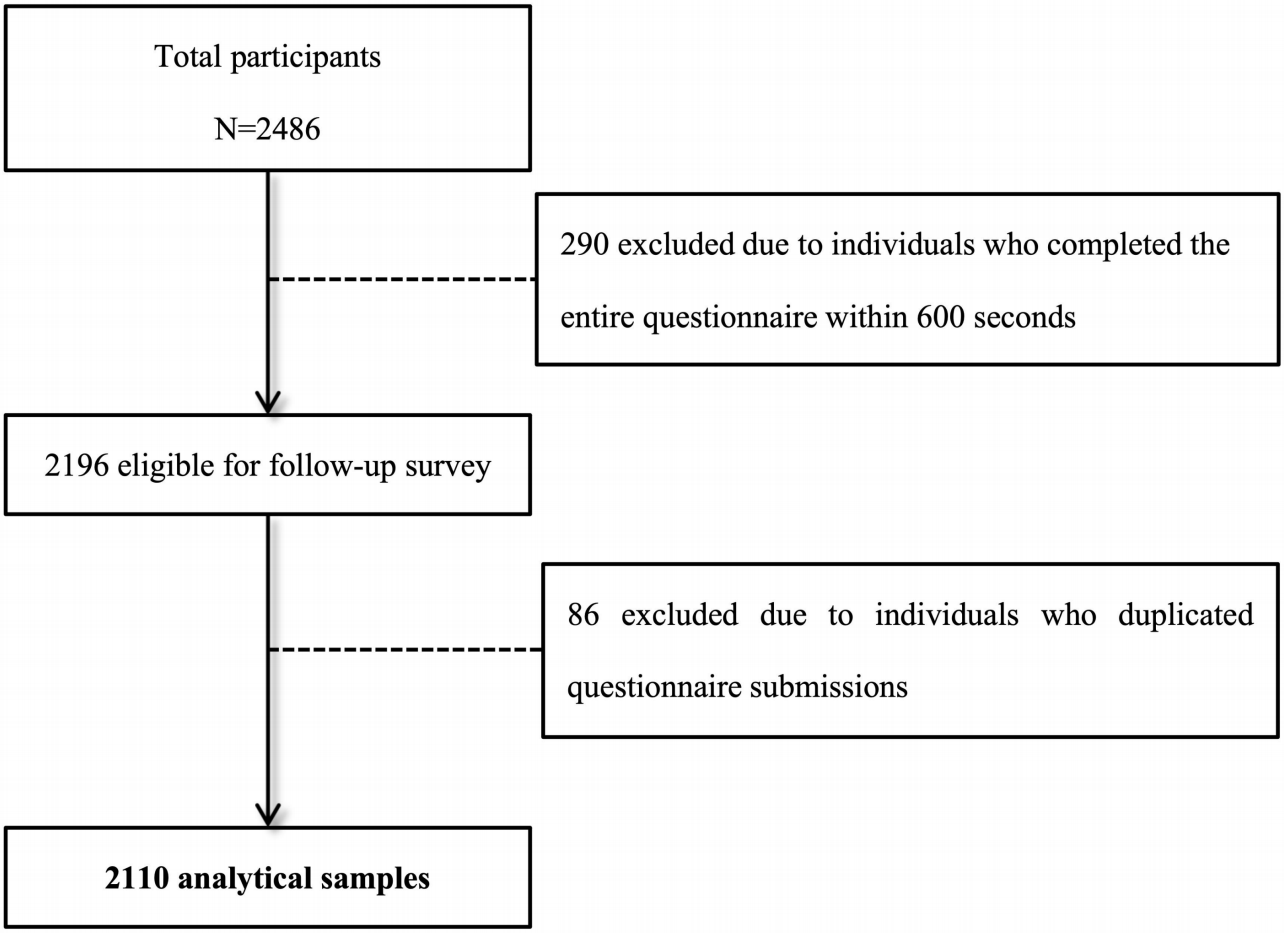
Results

Descriptive Statistical Analysis

Figure 1 presents a flowchart of participant inclusion. Table 1 presents the baseline characteristics of the 2110 participants. Overall, loneliness, quality of life, eHealth literacy, and depressive symptoms were distributed at low to moderate levels. Of the 2110 participants, 1033 (48.9%) were female and 1077 (51.1%) were male. Most participants were aged 60 to 69 years (n=1132, 53.7%), followed by those aged 70 to 79 years (n=798, 37.8%). A large proportion were married or partnered (n=1778, 84.3%) and had education lower than high school (n=1802, 85.4%). The majority were not working (n=1919, 90.9%). Self-reported health was most commonly rated as general (n=1247, 59.1%). Chronic diseases were reported by 1232 (58.4%) participants. Most participants did not smoke (n=1536, 72.8%) and did not consume alcohol (n=1677, 79.5%).

XSL•FO
RenderX

Figure 1. Flow diagram of participant screening and selection.



XSL•FO
RenderX

Table . Descriptive statistics (N=2110).

Variables		Values
Loneliness, median (IQR); range		3.91 (3-5); 3 - 9
Quality of life, median (IQR); range		46.35 (42-52); 13 - 65
eHealth literacy, median (IQR); range		19.22 (8-26); 8 - 40
Depressive symptoms, median (IQR); range		3.68 (0 - 6); 0 - 26
Sex, n (%)		
	Female	1033 (48.96)
	Male	1077 (51.04)
Age (y), n (%)		
	60 - 69	1132 (53.65)
	70 - 79	798 (37.82)
	≥80	180 (8.53)
Marital status, n (%)		
	Single	332 (15.73)
	Married or partnered	1778 (84.27)
Education, n (%)		
	Less than high school	1802 (85.4)
	High school or above	308 (14.6)
Employment status, n (%)		
	Not working	1919 (90.95)
	Working	191 (9.05)
Self-reported health status, n (%)		
	Bad	363 (17.2)
	General	1247 (59.1)
	Good	500 (23.7)
Chronic diseases, n (%)		
	No	878 (41.61)
	Yes	1232 (58.3)
Smoking, n (%)		
	No	1536 (72.8)
	Yes	574 (27.2)
Alcohol consumption, n (%)		
	No	1677 (79.48)
	Yes	433 (20.52)

Relationship Between Loneliness, eHealth Literacy, and Quality of Life

Table 2 summarizes the results of the multiple linear regression models. In model 1, loneliness was significantly associated with lower quality of life (β=−0.66, 95% CI −1.03 to −0.29; *P*<.001). After adding eHealth literacy in model 2, this association remained significant and became slightly stronger (β=−0.83, 95% CI −1.18 to −0.49; *P*<.001). eHealth literacy was positively associated with quality of life (β=0.41, 95% CI 0.35-0.47; *P*<.001).

XSL•FO
RenderX

Table . Multiple linear regression analysis of the relationship between loneliness, eHealth literacy, and quality of life (N=2110).

Variables	Model 1: quality of life, β (robust SE[a]; 95% CI)	Model 2: quality of life, β (robust SE[b]; 95% CI)
Loneliness	−0.66 (0.19; −1.03 to −0.29)[c]	−0.83 (0.18; −1.18 to −0.49)[c]
eHealth literacy[d]	—[e]	0.41 (0.03; 0.35 to 0.47)[c]
Sex (reference: female)		
Male	−0.33 (0.69; −1.67 to 1.02)	−0.43 (0.65; −1.70 to 0.85)
Age (years; ref: 60-69)		
70-79	−1.62 (0.58; −2.76 to −0.47)[f]	−0.64 (0.56; −1.75 to 0.47)
≥80	−1.03 (0.95; −2.89 to 0.84)	0.71 (0.96; −1.17 to 2.59)
Marital status (ref: single)		
Married or partnered	2.04 (0.79; 0.49 to 3.59)[g]	1.10 (0.77; −0.41 to 2.61)
Education (ref: less than high school)		
High school or above	0.50 (0.76; −0.98 to 1.98)	−1.09 (0.70; −2.46 to 0.28)
Self-reported health status (ref: bad)		
General	3.13 (0.72; 1.71 to 4.55)[c]	2.80 (0.70; 1.44 to 4.17)[c]
Good	7.07 (0.87; 5.36 to 8.78)[c]	6.30 (0.84; 4.65 to 7.95)[c]
Employment status (ref: not working)		
Working	−0.39 (0.96; −2.27 to 1.49)	−1.02 (0.86; −2.71 to 0.67)
Chronic diseases (ref: no)		
Yes	−0.55 (0.54; −0.50 to 1.61)	0.55 (0.51; −0.46 to 1.55)
Smoking (ref: no)		
Yes	0.11 (0.75; −1.36 to 1.58)	0.29 (0.72; −1.12 to 1.70)
Alcohol consumption (ref: no)		
Yes	0.98 (0.75; −0.53 to 2.46)	0.46 (0.71; −0.94 to 1.86)
Constant	43.97 (1.50; 41.03 to 46.90)[c]	45.69 (1.41; 42.93 to 48.45)[c]

[a]R^2=0.0658.

[b]R^2=0.1521.

[c]P<.001.

[d]The continuous variable of eHealth literacy has been mean-centered.

[e]Variables were not included in the model.

[f]P<.01.

[g]P<.05.

Exploratory analyses based on author-defined item groupings of eHealth literacy were conducted to examine their associations with quality of life. The results showed broadly consistent patterns with the primary analyses based on the total eHealth literacy score. Detailed results are presented in Table S5 in Multimedia Appendix 1.

Results of Multiple Linear Regression With Interaction Terms

Table 3 shows the results of the interaction models. In model 6, the interaction between loneliness and eHealth literacy was significant (β=−0.04, 95% CI −0.08 to −0.002; P=.04), indicating that higher eHealth literacy was associated with a steeper negative association between loneliness and quality of life. The inclusion of eHealth literacy and its interaction with loneliness increased the explained variance in quality of life from 6.58% (model 1, R^2=0.0658) to 15.37% (model 6, R^2=0.1537), indicating improved model explanatory power (ΔR^2=0.0879). When comparing the model with main effects to the interaction model, the incremental contribution of the interaction term was 0.0016 (model 6, R^2=0.1537; model 2, R^2=0.1521; ΔR^2=0.0016), and the interaction effect was statistically significant. Additional exploratory analyses examined interaction effects between loneliness and the author-defined item groupings of eHealth literacy. The patterns were consistent with the primary interaction findings based on the total score. Detailed results are presented in Table S6 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Results of multiple linear regression with interaction terms examining the relationship between loneliness, eHealth literacy, and quality of life (N=2110).

Variables	Model 6: quality of life, β (robust SE; 95% CI)[a]
Loneliness	−0.83 (0.18; −1.18 to −0.49)[b]
eHealth literacy[c]	0.56 (0.08; 0.40 to 0.73)[b]
Interaction	
Loneliness×eHealth literacy[c]	−0.04 (0.02; −0.08 to −0.002)[d]
Covariates[e]	
Constant	45.64 (1.42; 42.86 to 48.41)[b]

[a]R^2=0.1537.

[b]P<.001.

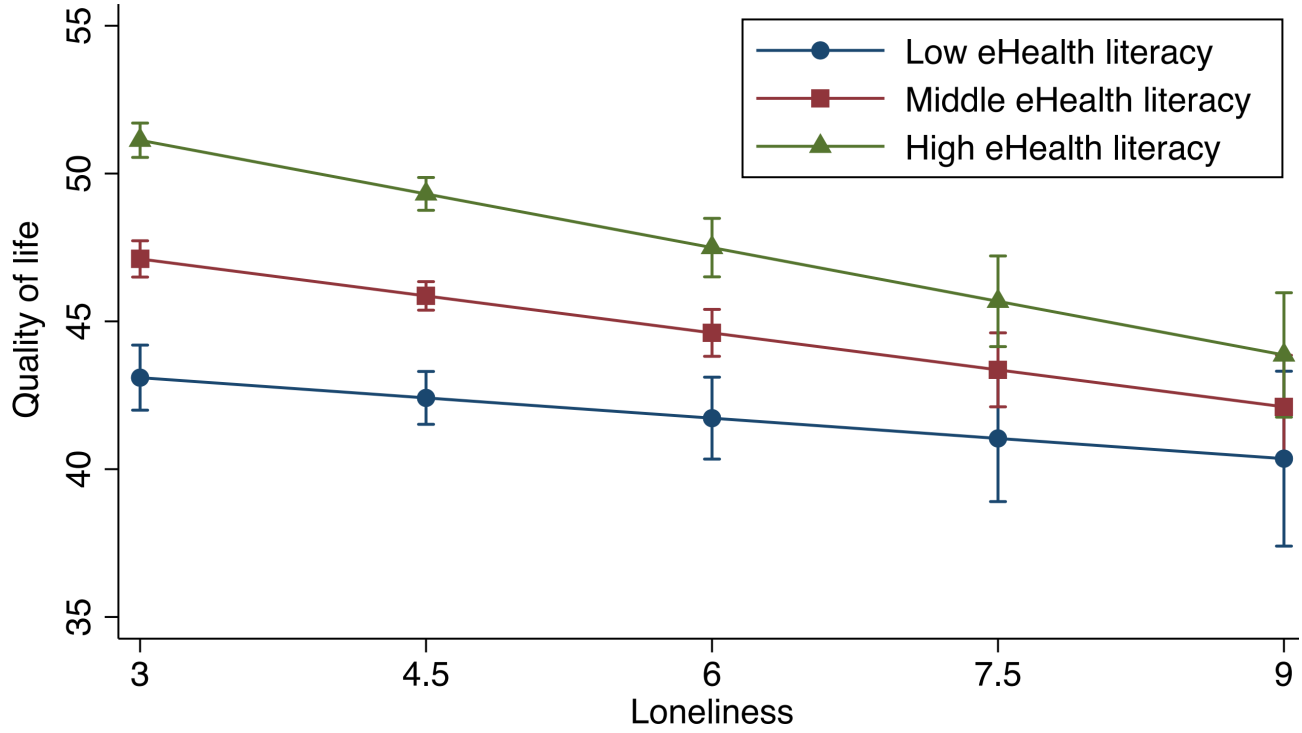
[c]The continuous variable of eHealth literacy has been mean-centered.

[d]P<.05.

[e]Adjusted for all covariates.

To further interpret the interaction between loneliness and eHealth literacy, predicted values of quality of life were estimated at different levels of the continuous moderator. Figure 2 presents the interaction plot based on mean-centered eHealth literacy, with predicted values shown at low (mean−1 SD), middle, and high (mean+1 SD) levels of eHealth literacy. The results indicate that quality of life declined as loneliness increased across all levels of eHealth literacy. Simple slope analyses showed that the association between loneliness and quality of life was not statistically significant at low levels of eHealth literacy (simple slope=−0.46; P=.13), but became significantly stronger at the middle levels (simple slope=−0.83; P<.001) and high levels of eHealth literacy (simple slope=−1.21; P<.001). These findings suggest that the negative association between loneliness and quality of life becomes more pronounced as eHealth literacy increases.

Figure 2. Moderating effect of eHealth literacy on the relationship between loneliness and quality of life.



Sensitivity Analysis

Table S7 in Multimedia Appendix 1 shows the results of the interaction models when depressive symptoms were used as the outcome. Loneliness remained strongly associated with higher depressive symptoms across all models. The interaction between loneliness and eHealth literacy was significant and positive, suggesting that higher eHealth literacy was associated with a steeper positive association between loneliness and depressive symptoms (model 10). Figure S1 in Multimedia Appendix 1 illustrates the simple slopes of loneliness on depressive symptoms at different levels of the continuous eHealth literacy moderator. Depressive symptoms increased with higher levels of loneliness across all levels of eHealth literacy (low eHealth

XSL•FO
RenderX

literacy, simple slope=1.62; *P*<.001 vs middle eHealth literacy, simple slope=1.84; *P*<.001 vs high eHealth literacy, simple slope=2.06; *P*<.001). The slope was steeper at higher levels of eHealth literacy, indicating that the positive association between loneliness and depressive symptoms became stronger as eHealth literacy increased. These results indicate that the moderating role of eHealth literacy remained stable when the outcome variable was replaced, supporting the robustness of the study findings.

Table S8 in Multimedia Appendix 1 presents the stratified results. In both male and female participants, eHealth literacy continued to show a positive association with quality of life. The interaction between loneliness and eHealth literacy was significant among female participants (β=−0.07, 95% CI −0.13 to −0.01; *P*=.02). Overall, the stratified models showed that the interaction results were consistent with the primary interaction findings based on the total eHealth literacy score, indicating the robustness of the findings.

Discussion

Principal Findings

This study examined loneliness, eHealth literacy, and quality of life among older adults in China and assessed the moderating role of eHealth literacy. Loneliness showed a clear negative association with quality of life. Higher eHealth literacy was positively associated with quality of life. Interaction analyses indicated that higher eHealth literacy was associated with a steeper negative association between loneliness and quality of life among older adults.

This study found a consistent negative association between loneliness and quality of life among older adults. This result aligns with evidence from cohort studies and systematic reviews showing that loneliness is linked to depression, anxiety, physical decline, cardiovascular events, and higher all-cause mortality, as well as sustained reductions in subjective well-being and health-related quality of life [10,15,17,28]. Persistent loneliness may impair psychological and social functioning through pathways involving functional deterioration [29], negative emotional accumulation, and reduced social participation, thereby undermining quality of life. These findings highlight the importance of prioritizing loneliness reduction in healthy aging policies and community interventions aimed at improving quality of life in later life.

We observed a robust positive association between eHealth literacy and quality of life in older adults. Existing evidence indicates that higher eHealth literacy enhances the ability of older adults to obtain and understand online health information [30], supports the use of digital health services, and promotes self-management [31], which can improve both psychological and physical health and ultimately enhance quality of life [32]. In China, where health care resources are unevenly distributed and accessing offline services can be burdensome, older adults with higher eHealth literacy are more likely to secure timely information and support through digital channels, translating digital access advantages into better quality of life [30,33].

Our study also showed that higher eHealth literacy was associated with a steeper negative association between loneliness and quality of life. This pattern differs from prior evidence, particularly from studies conducted during the COVID-19 pandemic among younger or student populations, where higher eHealth literacy was more commonly found to attenuate the negative associations between fear, loneliness, depressive symptoms, and health-related quality of life [34-37]. These findings suggest that eHealth literacy may have dual effects in digital health environments. Higher eHealth literacy may be accompanied by more intensive health information seeking and greater reliance on self-directed health management, which can increase cognitive load and decision burden [18,31]. Under conditions of loneliness and limited emotional support, greater engagement with online health information may translate into increased tendencies toward cyberchondria and psychological strain, thereby strengthening the association between loneliness and reduced quality of life [38,39]. In addition, in the absence of adequate family support or professional guidance, older adults with high eHealth literacy may also need to interpret complex health information on their own, which can intensify confusion, stress, and uncertainty and create a cumulative negative effect between loneliness and digital information exposure [40]. Furthermore, the cross-sectional design precludes conclusions regarding temporal ordering. It is possible that older adults with poorer well-being or heightened health concerns are more motivated to engage with online health information and to develop or report higher eHealth literacy, while loneliness simultaneously contributes to lower quality of life. This possibility highlights the need for longitudinal and experimental research to clarify directionality and underlying mechanisms. These findings highlight the need to further investigate the context-dependent functions of eHealth literacy and to examine potential mediating mechanisms, including information overload, health anxiety, cyberchondria, and decision burden, through longitudinal and experimental designs that can clarify temporal ordering and causal pathways. Additionally, strategies aimed at enhancing digital competence among older adults may require complementary forms of structured support, such as assistance in interpreting complex online health information and facilitated linkage to trusted community or primary care resources, particularly for individuals experiencing social vulnerability.

Strengths

This study has several strengths. It draws on a large community-based sample of older adults from multiple health service centers, providing robust evidence from a population that is often underrepresented in digital health research. The study simultaneously examined loneliness, eHealth literacy, and quality of life, allowing for a nuanced understanding of how digital competencies relate to quality of life in later life. The use of validated instruments and rigorous quality control procedures strengthened the reliability of the measures. In addition, the inclusion of interaction models and sensitivity analyses enhanced the robustness of the findings. By identifying a pattern in which higher eHealth literacy was associated with a steeper negative association between loneliness and quality of life, this study provides new insights into the complex role

XSL•FO
RenderX

of eHealth literacy and contributes to the understanding of digital vulnerability among older adults.

Limitations

This study has several limitations. First, the cross-sectional design precludes establishing temporal ordering. Future studies would benefit from longitudinal designs to clarify how loneliness and eHealth literacy interact over time and to identify potential bidirectional pathways. Second, all variables were measured through self-report, which may introduce reporting bias. Future research could incorporate clinical indicators, digital trace data, or performance-based assessments to provide more objective measures of health status and online information use. Third, the sample was drawn from one region in northwestern China, and findings may not generalize to areas with different digital access or social support structures, underscoring the value of multisite or nationally representative studies.

Conclusions

This study found that loneliness was consistently associated with lower quality of life among older adults and that higher eHealth literacy was associated with a steeper negative association between loneliness and quality of life. These results suggest that eHealth literacy functions as a dual-edged resource in later life. Although higher literacy is linked to better quality of life overall, it may intensify the challenges faced by lonely older adults when navigating complex or risk-oriented online health information. Given the cross-sectional and observational nature of the study, these interpretations should be considered exploratory. Future research is needed to examine the underlying psychological mechanisms, including potential roles of health anxiety, information overload, or emotional regulation processes, and to determine whether complementary support strategies could mitigate unintended negative consequences in digitally engaged but socially vulnerable populations.

Acknowledgments

The authors thank Chen Chen and Ban Hu (Department of Health Policy and Management, School of Public Health, Peking University) for their support in data collection.

Funding

This work was supported by the Major Program of the National Social Science Foundation of China (grant 22&ZD143) to Weiyan Jian and the Beijing Natural Science Foundation (grant L242145) to Jing Guo. The funders had no role in the study design; data collection, analysis, or interpretation; manuscript preparation; or the decision to publish the results.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

SZ: conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, writing—original draft, and writing–review and editing.
JL: data curation, formal analysis, methodology, validation, visualization, and writing–review and editing.
JT: validation, visualization, and writing–review and editing.
WJ: resources, project administration, supervision, funding acquisition, and writing–review and editing.
QL: resources, project administration, supervision, and writing–review and editing.
JG: conceptualization, methodology, resources, validation, project administration, supervision, funding acquisition, and writing–review and editing.

Weiyan Jian, Qing Liu , and Jing Guo are co-corresponding authors.
Weiyan Jian, PhD
Department of Health Policy and Management
School of Public Health
Peking University
No.38 Xue yuan Rd., Haidian District
Beijing, 100191
China
Phone: 86 13683354599
Email: jianweiyan@bjmu.edu.cn
Qing Liu, PhD
General Practice Department
Second Outpatient Section
Peking University Third Hospital
No. 49 Huayuan North Rd., Haidian District
Beijing, 100083
China

XSL•FO
RenderX

Phone: 86 13520203689
Email: qingfoxii@163.com

Conflicts of Interest

None declared.

Multimedia Appendix 1
Scales used in study and additional statistical results.
[DOCX File, 251 KB - aging_v9i1e88600_app1.docx]

References

1. Netuveli G, Blane D. Quality of life in older ages. Br Med Bull 2008;85:113-126. [doi: 10.1093/bmb/ldn003] [Medline: 18281376]
2. Felce D, Perry J. Quality of life: its definition and measurement. Res Dev Disabil 1995;16(1):51-74. [doi: 10.1016/0891-4222(94)00028-8] [Medline: 7701092]
3. Renne I, Gobbens RJ. Effects of frailty and chronic diseases on quality of life in Dutch community-dwelling older adults: a cross-sectional study. Clin Interv Aging 2018;13:325-334. [doi: 10.2147/CIA.S156116] [Medline: 29520132]
4. Lawson RA, Yarnall AJ, Duncan GW, et al. Cognitive decline and quality of life in incident Parkinson's disease: the role of attention. Parkinsonism Relat Disord 2016 Jun;27:47-53. [doi: 10.1016/j.parkreldis.2016.04.009] [Medline: 27094482]
5. Zhao L. China's aging population: a review of living arrangement, intergenerational support, and wellbeing. Health Care Sci 2023 Oct;2(5):317-327. [doi: 10.1002/hcs2.64] [Medline: 38938584]
6. Song Q. Aging and separation from children: the health implications of adult migration for elderly parents in rural China. Demogr Res 2017;37:1761-1792. [doi: 10.4054/DemRes.2017.37.55] [Medline: 30581322]
7. He L, Wang J, Wang F, Zhang L, Liu Y, Xu F. Depression symptoms and quality of life in empty-nest elderly among Chengdu: a cross-sectional study. Front Psychiatry 2022;13:1003261. [doi: 10.3389/fpsyt.2022.1003261] [Medline: 36424995]
8. Van As BA, Imbimbo E, Franceschi A, Menesini E, Nocentini A. The longitudinal association between loneliness and depressive symptoms in the elderly: a systematic review. Int Psychogeriatr 2022 Jul;34(7):657-669. [doi: 10.1017/S1041610221000399] [Medline: 33849675]
9. Wang Y, Liu M, Yang F, Chen H, Wang Y, Liu J. The associations of socioeconomic status, social activities, and loneliness with depressive symptoms in adults aged 50 years and older across 24 countries: findings from five prospective cohort studies. Lancet Healthy Longev 2024 Sep;5(9):100618. [doi: 10.1016/j.lanhl.2024.07.001] [Medline: 39208829]
10. Lee SL, Pearce E, Ajnakina O, et al. The association between loneliness and depressive symptoms among adults aged 50 years and older: a 12-year population-based cohort study. Lancet Psychiatry 2021 Jan;8(1):48-57. [doi: 10.1016/S2215-0366(20)30383-7] [Medline: 33181096]
11. Barreto M, Victor C, Hammond C, Eccles A, Richins MT, Qualter P. Loneliness around the world: age, gender, and cultural differences in loneliness. Pers Individ Dif 2021 Feb 1;169:110066. [doi: 10.1016/j.paid.2020.110066] [Medline: 33536694]
12. Stegen H, Duppen D, Savieri P, et al. Loneliness prevalence of community-dwelling older adults and the impact of the mode of measurement, data collection, and country: a systematic review and meta-analysis. Int Psychogeriatr 2024 Sep;36(9):747-761. [doi: 10.1017/S1041610224000425] [Medline: 38525677]
13. Wang J, Zhou Y, Zhang Q, et al. Correction to: loneliness among older Chinese individuals: the status quo and relationships with activity-related factors. BMC Geriatr 2024;24(1). [doi: 10.1186/s12877-024-04821-9]
14. Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. Heart 2016 Jul 1;102(13):1009-1016. [doi: 10.1136/heartjnl-2015-308790] [Medline: 27091846]
15. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
16. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. PLoS Med 2010 Jul 27;7(7):e1000316. [doi: 10.1371/journal.pmed.1000316] [Medline: 20668659]
17. VanderWeele TJ, Hawkley LC, Cacioppo JT. On the reciprocal association between loneliness and subjective well-being. Am J Epidemiol 2012 Nov 1;176(9):777-784. [doi: 10.1093/aje/kws173] [Medline: 23077285]
18. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9. [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
19. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan 27;14(1):e19. [doi: 10.2196/jmir.1619] [Medline: 22357448]
20. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985 Sep;98(2):310-357. [doi: 10.1037/0033-2909.98.2.310] [Medline: 3901065]

XSL•FO
RenderX

21. Starcevic V. Keeping Dr. Google under control: how to prevent and manage cyberchondria. World Psychiatry 2023 Jun;22(2):233-234. [doi: 10.1002/wps.21076] [Medline: 37159368]
22. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based studies. Res Aging 2004;26(6):655-672. [doi: 10.1177/0164027504268574] [Medline: 18504506]
23. Bowling A, Hankins M, Windle G, Bilotta C, Grant R. A short measure of quality of life in older age: the performance of the brief Older People's Quality of Life questionnaire (OPQOL-brief). Arch Gerontol Geriatr 2013;56(1):181-187. [doi: 10.1016/j.archger.2012.08.012] [Medline: 22999305]
24. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
25. Richtering SS, Hyun K, Neubeck L, et al. eHealth literacy: predictors in a population with moderate-to-high cardiovascular risk. JMIR Hum Factors 2017 Jan 27;4(1):e4. [doi: 10.2196/humanfactors.6217] [Medline: 28130203]
26. Levis B, Benedetti A, Thombs BD, DEPRESsion Screening Data (DEPRESSD) Collaboration. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. BMJ 2019 Apr 9;365:l1476. [doi: 10.1136/bmj.l1476] [Medline: 30967483]
27. Sterne JA, Smith GD. Sifting the evidence-what's wrong with significance tests? Phys Ther 2001 Aug 1;81(8):1464-1469. [doi: 10.1093/ptj/81.8.1464] [Medline: 28206639]
28. Perissinotto CM, Stijacic Cenzer I, Covinsky KE. Loneliness in older persons: a predictor of functional decline and death. Arch Intern Med 2012 Jul 23;172(14):1078-1083. [doi: 10.1001/archinternmed.2012.1993] [Medline: 22710744]
29. Santini ZI, Jose PE, York Cornwell E, et al. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. Lancet Public Health 2020 Jan;5(1):e62-e70. [doi: 10.1016/S2468-2667(19)30230-0] [Medline: 31910981]
30. Jung SO, Son YH, Choi E. eHealth literacy in older adults: an evolutionary concept analysis. BMC Med Inform Decis Mak 2022 Jan 31;22(1):28. [doi: 10.1186/s12911-022-01761-5] [Medline: 35101005]
31. Watkins I, Xie B. eHealth literacy interventions for older adults: a systematic review of the literature. J Med Internet Res 2014 Nov 10;16(11):e225. [doi: 10.2196/jmir.3318] [Medline: 25386719]
32. Xie L, Zhang S, Xin M, Zhu M, Lu W, Mo PH. Electronic health literacy and health-related outcomes among older adults: a systematic review. Prev Med 2022 Apr;157:106997. [doi: 10.1016/j.ypmed.2022.106997] [Medline: 35189203]
33. Li S, Cui G, Yin Y, Wang S, Liu X, Chen L. Health-promoting behaviors mediate the relationship between eHealth literacy and health-related quality of life among Chinese older adults: a cross-sectional study. Qual Life Res 2021 Aug;30(8):2235-2243. [doi: 10.1007/s11136-021-02797-2] [Medline: 33661455]
34. Xu RH, Chan HH, Shi L, Li T, Wang D. Moderating effect of eHealth literacy on the associations of coronaphobia with loneliness, irritability, depression, and stigma in Chinese young adults: Bayesian structural equation model study. JMIR Public Health Surveill 2023 Sep 29;9:e47556. [doi: 10.2196/47556] [Medline: 37773621]
35. Nguyen MH, Pham TTM, Nguyen KT, et al. Negative impact of fear of COVID-19 on health-related quality of life was modified by health literacy, eHealth literacy, and digital healthy diet literacy: a multi-hospital survey. Int J Environ Res Public Health 2021 May 6;18(9):4929. [doi: 10.3390/ijerph18094929] [Medline: 34066312]
36. Tran HT, Nguyen MH, Pham TT, et al. Predictors of eHealth literacy and its associations with preventive behaviors, fear of COVID-19, anxiety, and depression among undergraduate nursing students: a cross-sectional survey. Int J Environ Res Public Health 2022 Mar 22;19(7):3766. [doi: 10.3390/ijerph19073766] [Medline: 35409448]
37. Zhao B, Kong F, Nam EW. Relationship between eHealth, perceived risk, and phobia of COVID-19 among Chinese university students in Korea and China. Health Soc Care Commun 2023 Feb 24;2023:1-12. [doi: 10.1155/2023/2755354]
38. Nguyen A, Mosadeghi S, Almario CV. Persistent digital divide in access to and use of the internet as a resource for health information: results from a California population-based study. Int J Med Inform 2017 Jul;103:49-54. [doi: 10.1016/j.ijmedinf.2017.04.008] [Medline: 28551001]
39. Crocco AG, Villasis-Keever M, Jadad AR. Analysis of cases of harm associated with use of health information on the internet. JAMA 2002 Jun 5;287(21):2869-2871. [doi: 10.1001/jama.287.21.2869] [Medline: 12038937]
40. McInnes N, Haglund BJ. Readability of online health information: implications for health literacy. Inform Health Soc Care 2011 Dec;36(4):173-189. [doi: 10.3109/17538157.2010.542529] [Medline: 21332302]

Edited by Y Jiang; submitted 28.Nov.2025; peer-reviewed by M Bujnowska-Fedak, S Matsuda; revised version received 22.Mar.2026; accepted 29.Mar.2026; published 11.May.2026.

Please cite as:
Zhang S, Li J, Tang J, Jian W, Liu Q, Guo J
Association Between Loneliness, eHealth Literacy, and Quality of Life Among Chinese Older Adults: Cross-Sectional Study
JMIR Aging 2026;9:e88600
URL: https://aging.jmir.org/2026/1/e88600
doi:10.2196/88600

XSL•FO
RenderX

© Shouchuang Zhang, Jiayi Li, Jia Tang, Weiyan Jian, Qing Liu, Jing Guo. Originally published in JMIR Aging (https://aging.jmir.org), 11.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effect of an Intersectoral Telemedicine Intervention on Hospitalization in Nursing Home Patients: Specification Curve Analysis of a Stepped Wedge Cluster Randomized Trial

John Grosser[1], MA, MSc; Sophie Pauge[1], MSc; Birthe Aufenberg[1], MSc; Wolfgang Greiner[1], Prof Dr; Miriam Hertwig[2], MSc; Jenny Unterkofler[2], Dr med; David Brücken[3], Dr med; Christian Hübel[2], Dr med; Optimal@NRW Research Group[4,5]; Jörg Christian Brokmann[2], Prof Dr med

[1]Department of Health Economics and Health Care Management, School of Public Health, Bielefeld University, Universitätsstraße 25, Bielefeld, Germany

[2]Department for Acute and Emergency Medicine, University Hospital RWTH Aachen, Aachen, Germany

[3]Department for Acute and Emergency Medicine, Rhine-Meuse Hospital Würselen, Wüselen, Germany

[4]Optimal@NRW Research Group, Aachen, Germany

[5]

Corresponding Author:
John Grosser, MA, MSc
Department of Health Economics and Health Care Management, School of Public Health, Bielefeld University, Universitätsstraße 25, Bielefeld, Germany

Abstract

Background: The use of emergency medical services in Germany has significantly increased, particularly for nonemergency cases. This issue becomes especially relevant for nursing homes, where staff is confronted with resource shortages and organizational challenges, making emergency medical services involvement a common course of action. This leads to potentially avoidable hospitalizations for nursing home residents and exposes them to severe health risks, such as nosocomial infections. To address this, we developed the Optimal@NRW project, implementing an innovative intersectoral telemedicine intervention that includes an early warning system and mobile nonphysician medical assistants in order to enable outpatient treatment in acute medical cases for nursing home residents. We hypothesize that hospitalizations could be significantly reduced, as patients could be treated on-site in the absence of a life-threatening emergency.

Objective: This study aims to determine the effect of an intersectoral telemedicine-based intervention for acute care situations on hospitalization among residents in nursing homes.

Methods: We conducted a prospective multicenter cluster-randomized stepped wedge trial from May 2021 to April 2023 across 24 nursing homes in western Germany. Nursing homes were randomly assigned to one of four sequences and followed over eight 3-month periods. All sequences crossed stepwise from the control to the intervention phase. We collected primary data in an electronic patient record and claims data from statutory health insurance funds. In a specification curve analysis, we evaluated the intervention effect using multiple mixed-effects regression model specifications with varying modeling approaches, datasets, dependent and independent variables, and operationalizations.

Results: The number of participants ranged from 158 to 285 nursing home residents per sequence and period, totaling 1151 participants in the control phase and 1271 participants in the intervention phase. A total of 1260 regression models were analyzed. Approximately two-thirds of all models exhibited a negative (ie, protective) intervention effect. However, no models exhibited a statistically significant (ie, $P<.05$) intervention effect. The distribution of effect direction and size differed by modeling approach, choice of outcome, and dataset used.

Conclusions: Our results do not support the hypothesis of a significant reduction of hospitalizations among nursing home residents. The effectiveness of the intervention may have been compromised by external factors such as a flood disaster in the study region and the COVID-19 pandemic, which introduced unforeseen challenges. Future studies should prioritize targeted outcomes (potentially avoidable hospitalizations) that align with the specific intervention mechanism, while also allowing for extended implementation periods to ensure that interventions reach their full potential.

Trial Registration: ClinicalTrials.gov NCT04879537; https://clinicaltrials.gov/study/NCT04879537

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-022-06613-1

(JMIR Aging 2026;9:e75721) doi:10.2196/75721

XSL•FO
RenderX

KEYWORDS

Telemedicine; nursing home; acute care; emergency medical services; specification curve analysis; nursing care; assisted living; early warning system; potentially avoidable hospitalization; hospitalization

Introduction

The demand for nursing and medical care for nursing home residents (NHR) is steadily increasing due to demographic change and the rising prevalence of chronic diseases, multimorbidity, and care dependency [1]. In Germany, approximately 4.9 million people currently require nursing care, with over half being older than 80 years [2]. Geriatric and multimorbid NHR are significantly sicker compared to older adults living outside such facilities and have higher rates of contact with health care providers [3]. Over the past decade, the use of emergency medical services in Germany has increased substantially [4]. Staffing shortages and organizational challenges in nursing homes, as well as inadequate care capacities in outpatient settings, often result in potentially avoidable hospitalizations for NHR in nonemergency situations, exposing residents to risks like nosocomial infections [5,6], delirium [7,8], and reduced functional capacity [9,10]. This trend is reflected in concepts such as ambulatory care-sensitive conditions (ACSC), core ambulatory care-sensitive conditions (CACSC) [11], and nursing home-sensitive conditions (NHSC) [12], which identify conditions where appropriate on-site or outpatient care could have avoided hospital admission.

These potentially avoidable hospitalizations not only place a financial burden on solidarity-based health care systems [13,14], but also increase costs for municipalities to provide emergency medical services such as ambulances and emergency response vehicles. Furthermore, the growing pressure on nursing homes, driven by societal changes and staff shortages, highlights the need for innovative care models. An intersectoral approach involving all relevant stakeholders is crucial for improving acute care for NHR, as hospital admissions affect not only the nursing homes but also emergency departments, ambulance services, and general practitioners (GPs).

Telemedicine has significant potential to transform acute care by enabling efficient, needs-based medical services across spatial and temporal barriers, optimizing resource use and access, particularly in rural or underserved areas [15,16]. During the COVID-19 pandemic, there was a significant increase in the use of telemedicine in Germany [17]. However, its widespread implementation faces challenges in German nursing homes such as regulatory constraints, reimbursement issues, and limited information technology infrastructure with problems regarding interoperability [18,19]. For instance, it was not compulsory for nursing homes to be connected to the telematics infrastructure until July 2025 (§ 341 (8) of Book V of the German Social Code) [20], which is intended to facilitate secure and efficient data exchange within the health care system. This leads to a fragmented landscape of digitalized medical documentation in nursing homes.

To address these challenges, we designed and implemented an innovative intersectoral telemedicine intervention aimed to enable outpatient treatment for NHR for acute medical conditions. Its effectiveness was evaluated through a prospective multicenter cluster-randomized trial with an open cohort stepped wedge design. The intervention was carried out in 24 nursing homes across the Aachen, Düren, and Heinsberg regions in western Germany, next to the Dutch and Belgian borders, covering both urban and rural areas. In this paper, we analyze the effect of the Optimal@NRW intervention on the hospitalization of NHR through the use of specification curve analysis. In particular, we report the intervention effects from a total of 1260 regression models, resulting from different modeling decisions (eg, choice of dataset, operationalizations, or independent variables).

Methods

Reporting

Our methodology follows the CONSORT (Consolidated Standards of Reporting Trials) statement for stepped wedge cluster randomized trials [21]. The checklist can be found in Checklist 1.

Trial Design

This paper is based on data from a prospective multicenter cluster-randomized trial with an open cohort stepped wedge design [22]. The stepped wedge design was chosen so that all participating nursing homes would eventually receive the intervention, thus fostering commitment to the project. A staggered rollout facilitated the practical and logistical implementation of the intervention. Our open cohort design aimed to recruit a substantial number of participants at the beginning of the trial, while adding more participants during the course of the trial to balance the high expected mortality rates among NHR.

Participants were clustered by nursing home; these clusters were summarized into 4 sequences. All nursing homes started in the control phase, in which participants received standard outpatient care for NHR. Sequences then transitioned stepwise from control to intervention phase, with the timing of the transitions randomized (Figure 1). Each step covered a 3-month interval, and all nursing homes participated in at least 2 intervals (6 mo) in both the control and intervention phases to ensure adequate data collection for each phase. The study took place over 2 years (May 2021 to April 2023), with 8 steps and a total of 24 nursing homes (clusters) across the 4 sequences. The transition phase between the control and intervention phases allowed for staff training and technical setup; data collected during this phase were excluded from our analysis.

XSL•FO
RenderX

Figure 1. Stepped wedge design of the Optimal@NRW trial. White: control phase, gray: intervention phase, blue: transition phase.

Month	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24
Sequence 1 7 Clusters								
Sequence 2 5 Clusters								
Sequence 3 6 Clusters								
Sequence 4 6 Clusters								

Participants and Data Collection

Potential study participants were included based on the predefined eligibility criteria described in Textbox 1. Recruitment was carried out by the study team within the participating nursing homes (clusters). Following the stepped wedge open cohort study design, participants were enrolled both at the beginning of the study and continuously throughout the study period. All participants gave their written consent before data collection.

Textbox 1. Inclusion and exclusion criteria of the Optimal@NRW trial.

Inclusion criteria:

- ≥ 18 years of age
- Residence in a participating nursing home
- Signed consent form for study participation and data transfer or consent form signed by carers (for residents who are unable to give consent)

Exclusion criteria:

- <18 years of age
- No residence in a participating nursing home
- Nursing home residents in a dependent or employment relationship with a member of the study team
- Lack of clarity about nursing home residents' capacity to give consent
- Nursing home residents placed in a nursing home by order of the authorities or the courts

We received data on the participants from two sources. First, we used primary data from a study database. The study database contained demographic information on the participants, such as year of birth, sex, and date of death, as well as insurance status and hospitalization data. Primary data were collected by study nurses employed by each nursing home. Second, we received claims data from five statutory health insurance (SHI) funds. The claims data contained demographic information as well as morbidity and hospitalization data. For all study participants who were insured with one of the participating SHI funds, the claims data were provided for the study period as well as an additional 2-year preobservation period. In addition to the information from the study database and claims data, the study nurses also measured the Barthel Index (BI), a patient-reported endpoint to describe the ability for activities of daily living [23], at two time points: at study enrollment (t0) and at the transition to the intervention phase (t1).

However, these two datasets were incompatible with each other for several reasons, with different advantages and disadvantages. First, the primary dataset covered all study participants, while the claims dataset included data only from those participants who were members of one of the participating SHI funds. Second, some information (such as routinely collected morbidity data and data on hospitalizations before the study period) was available only in the claims dataset. Therefore, instead of merging the datasets, we performed the statistical analysis (described below) separately for the primary data and claims data.

Intervention

The Optimal@NRW project used an innovative, intersectoral intervention to improve acute medical care for residents of nursing homes and residential care facilities (hereafter collectively referred to as "nursing homes"). This intervention consisted of three levels.

First, a telemedical center was initialized at the University Hospital Aachen (UK Aachen) and was available 24/7 for telemedical consultations concerning the NHR. These consultations enabled 3-way communication between patients, nurses, and (tele)physicians. If necessary, mobile nonphysician assistants (nonphysician medical assistants [Nicht-ärztliche Praxisassistent*innen mit Zusatzaufgaben]; NäPa[Z]) were dispatched to perform delegable actions on-site. An electronic health record centralized patient information to ensure informed decision-making and coordination among all care providers.

Second, Optimal@NRW introduced a telemedicine approach via a "virtual hub," where nursing home staff could call the medical on-call service (116 117) in the event of an acute

XSL•FO
RenderX

medical situation involving an NHR. The medical on-call service then assessed the urgency of the situation using a structured initial medical assessment software [24] and determined the next steps. Alternatively, nursing home staff could contact the telephysicians of the UK Aachen directly.

Third, an early warning system (Frühwarnsystem) was implemented, involving routine monitoring of vital signs by nursing home staff, with data transmitted to the telemedicine center for continuous software-based assessment. Further intervention details can be found in the study protocol [25].

During the control phase, care in the participating nursing homes was provided as usual. In acute medical situations, nursing staff were responsible for making decisions independently. This typically involved either contacting a GP or calling the medical on-call service or an ambulance directly.

Sample Size

For the sample size calculation, we assumed an average of 2.8 days in hospital per patient-quarter based on hospitalization data for German nursing home patients from 2019 [26]. Based on estimates from medical experts at UK Aachen, we hypothesized that the number of days in hospital could be reduced by 30%. Informed by a prior study [27], the intracluster correlation coefficient was set at ρ=0.12. For 24 participating nursing homes, a significance level of α=.05, a power of 1-β=.90, and a SD of 4.5 resulted in a required sample size of 43 residents per nursing home per time point. Considering a 12% quarterly dropout due to high mortality [28] and other factors, the final adjusted sample size was 49 residents per nursing home per time point.

Randomization

We used stratified randomization at the nursing home level, taking into account bed capacity (small entities with <80 beds and large entities with ≥80 beds) and specialization in dementia care (yes or no) to ensure balanced group allocation. Additionally, block randomization was applied within these stratified groups to avoid imbalances in cluster distribution. The selection of eligible nursing homes was carried out in advance by the study team, while health economists from Bielefeld University, who were not involved in recruitment, data collection, or the design and implementation of the intervention, designed the allocation plan and performed the randomization process to minimize allocation bias. Nursing homes were informed of their assigned sequence by the study team before the start of the trial.

Objective and Outcomes

The primary objective of our analysis in this paper was to evaluate hospitalization in the intervention versus control phases. It was hypothesized that the intervention would lead to a significant reduction in hospital admissions and, therefore, a significant reduction in days in hospital (it is important to note that the intervention only affects the probability of hospital admission for patients in the nursing home and is not hypothesized to directly affect the length of stay for patients in the hospital).

Hospitalization was measured as a binary variable for each day spent in the nursing home by each patient, defined as 1 if the patient was admitted to the hospital on that day, and 0 otherwise. In addition to a dependent variable considering all hospitalizations, we also defined four further dependent variables, which operationalized potentially avoidable hospitalizations as ACSC hospitalizations, CACSC hospitalizations, NHSC hospitalizations, or same-day discharges (SDD). ACSC refers to a defined set of conditions for the general population in Germany that identifies potentially avoidable hospital admissions. The list is based on *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) and includes 40 *ICD-10* diagnoses (ACSC) and a core list of 22 *ICD-10* diagnoses (CACSC) [11]. Additionally, a list of NHSC, comprising 58 *ICD-10* diagnoses that frequently lead to avoidable hospitalizations in this patient cohort, has recently been finalized for Germany [12]. For each hospitalization, we also measured the duration in days (starting at 1 for SDD). All dependent variables available for analysis are shown in Textbox 2

Participant-level characteristics at baseline were summarized by study phase and by allocated sequence. The participant-level characteristics included gender, age, morbidity, and prior hospitalization. For morbidity and prior hospitalization, multiple operationalizations were available. All independent variables available for analysis in the primary and claims datasets are shown in Textbox 3.

XSL•FO
RenderX

Textbox 2. Available dependent variables. All variables are available at the patient-day level. Variables in the primary dataset were based on information from the study database; variables in the claims dataset were based on information from the statutory health insurance claims data.

Available dependent variables

- Hospitalization: binary variable that equals 1 if there was a hospitalization on the patient-day in question. Equals 0 otherwise.
- Ambulatory care-sensitive conditions (ACSC) hospitalization: binary variable that equals 1 if there was a hospitalization on the patient-day in question and the primary discharge diagnosis of that hospitalization is contained in the ACSC list of *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) codes. Equals 0 otherwise.
- Core ambulatory care-sensitive conditions (CACSC) hospitalization: binary variable that equals 1 if there was a hospitalization on the patient-day in question and the primary discharge diagnosis of that hospitalization is contained in the CACSC list of *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) codes. Equals 0 otherwise.
- Nursing home-sensitive conditions (NHSC) hospitalization: binary variable that equals 1 if there was a hospitalization on the patient-day in question and the primary discharge diagnosis of that hospitalization is contained in the NHSC list of *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) codes. Equals 0 otherwise.
- Same-day discharge (SDD) hospitalization: binary variable, equals 1 if there was a hospitalization on the patient-day in question and the patient in question was discharged on the same day. Equals 0 otherwise.
- Days in hospital: count variable measuring the number of days spent in hospital as the result of hospitalization on the patient-day in question. Equals 0 if there was no hospitalization on the patient-day in question.
- ACSC days in hospital: count variable measuring the number of days spent in hospital as the result of ACSC hospitalization on the patient-day in question. Equals 0 if there was no ACSC hospitalization on the patient-day in question.
- CACSC days in hospital: count variable measuring the number of days spent in hospital as the result of CACSC hospitalization on the patient-day in question. Equals 0 if there was no CACSC hospitalization on the patient-day in question.
- NHSC days in hospital: count variable measuring the number of days spent in hospital as the result of NHSC hospitalization on the patient-day in question. Equals 0 if there was no NHSC hospitalization on the patient-day in question.
- SDD days in hospital: count variable measuring the number of days spent in hospital as the result of SDD hospitalization on the patient-day in question. Equals 0 if there was no SDD hospitalization on the patient-day in question.

XSL•FO
RenderX

Textbox 3. Available independent variables. All variables are available at the patient-day level. Variables in the primary dataset were based on information from the study database; variables in the claims dataset were based on information from the statutory health insurance claims data.

Available independent variables with only one operationalization

- ID: factor variable, a pseudonym indicating the patient in question; nursing home: factor variable indicating the nursing home of the patient in question.
- Time: factor variable indicating the study quarter (period) of the patient-day in question.
- Study group: binary variable indicating the study phase (control or intervention) of the patient in question on the day in question.
- Gender: factor variable indicating the gender of the patient in question.
- Age: continuous variable obtained by subtracting the birth year of the patient in question from the year of the day in question and scaling around the mean.

Morbidity (five operationalizations of the morbidity of the patient in question)

- Factor variable obtained by splitting the most recent observation of the Barthel Index of the patient in question into three categories [29]: no disability (100), moderate disability (56-99), and severe disability (0 - 55).
- Factor variable obtained by splitting the most recent observation of the Barthel Index of the patient in question into five categories [30]: independence (100), slight dependence (91-99), moderate dependence (61-90), severe dependence (21-60), and total dependence (0 - 20).
- Continuous variable (available only in claims dataset) measuring the Charlson Comorbidity Index of the patient in question, calculated daily based on diagnoses during the 365 days prior to the day in question using the comorbidity package [31].
- Factor variable (available only in claims dataset) obtained by splitting the Charlson Comorbidity Index of the patient in question into 3 categories [29]: no comorbidities (0), moderate comorbidities (1-2), and severe comorbidities (3+).
- Factor variable (available only in claims dataset) measuring the care grade of the patient in question (integer between 0 and 5, assigned by the Health Insurance Medical Service according to § 15 and § 18 of the German Social Code, Book XI (SGB XI)) [32].

Prior hospitalization (two operationalizations of prior hospitalizations of the patient in question)

- Binary variable (available only in claims dataset) that equals 1 if the patient in question was admitted to the hospital at least once in the 365 days prior to the day in question. Equals 0 if otherwise.
- Count variable (available only in claims dataset) measuring the number of times the patient in question was admitted to the hospital in the 365 days prior to the day in question.

Statistical Methods

To estimate the effect of the intervention, we performed mixed-effects regression analysis. However, in specifying our regression model, we faced several modeling decisions that offered no clear correct choice. Instead of fitting only a single regression model, we addressed these issues by performing a specification curve analysis. In such an analysis, researchers report the results of all "reasonable specifications," defined as the set of model specifications that are "(1) sensible tests of the research question, (2) expected to be statistically valid and (3) not redundant with other specifications in the set" [33]. The estimated effect sizes from each model are then ordered by size and visualized in a diagram–the specification curve. Specification curve analysis can be considered a formal extension of the common practice of considering the robustness of results to model specification decisions [33]. For our analysis, we ran a total of 1260 separate regression models with different combinations of modeling approaches, datasets, dependent and independent variables, and operationalizations.

First, we needed to choose which modeling approach to use. For this, we considered three options: (1) a negative binomial count model for the number of hospital days at the patient-quarter level, (2) a negative binomial count model for the number of hospitalizations at the patient-quarter level, and (3) a binomial model for hospital admission at the patient-day level. For the count models, we used a log link function. For the binomial regression, we used a complementary log-log (*cloglog*) link function. Unlike logit or probit link functions, *cloglog* link functions are asymmetrical and are therefore most appropriate when the probability of an event is very high or very low [34-36]. This was the case for our data, as there was no hospitalization on the vast majority of person-days in both datasets.

Second, as the primary data collected by the study team and the claims data provided by the SHI funds were incompatible, we needed to choose on which data to fit the regression model. Third, after choosing a dataset, we needed to choose how to operationalize the outcome: as measuring all, ACSC, CACSC, NHSC, or SDD hospitalizations/days in hospital. For ease of reference, we refer to models using ACSC hospitalizations as the outcome as "ACSC models" (and so on for the remaining operationalizations).

In each model specification, we included the variables study group (the treatment variable), time (to model secular trends), nursing home, and id (to model cluster- and patient-level measurement using random intercepts). For the first and second modeling approaches (count models for the number of hospital days and the number of hospitalizations at the patient-quarter level), we also included a factor variable controlling for the time spent in the nursing home by each patient in a given quarter. However, we still needed to decide which combination of the

XSL•FO
RenderX

remaining available independent variables (gender, age, morbidity, and prior hospitalization) to use as control variables.

Finally, multiple operationalizations were available for two independent variables in one or both datasets, with no clear theoretical basis for choosing between these operationalizations. In particular, the primary dataset included two possible operationalizations of morbidity, while the claims dataset included five possible operationalizations of morbidity and two possible operationalizations of prior hospitalization (recall Textbox 3). For each modeling approach, all possible combinations of outcome and control variables amounted to a total of 5*2*2*3=60 model specifications on the primary dataset and 5*2*2*6*3=360 model specifications on the claims dataset. This yields a total of 420 model specifications per modeling approach and 1260 model specifications across all 3 modeling approaches. For each of these model specifications, we removed all patients who had missing values in at least one variable that was included in the model specification. The model formulas for each modeling approach on both datasets are illustrated in Textbox 4.

Textbox 4. Model formulas on the primary and claims datasets.

Primary dataset

1. negative binomial regression (patient-quarter level): hospital days (5 options)~study group+time+gender (2 options)+age (2 options)+morbidity (3 options)+nursing home (random intercept)+id (random intercept)
2. negative binomial regression (patient-quarter level): hospital admissions (5 options)~ study group+ time+gender (2 options)+age (2 options)+morbidity (3 options)+nursing home (random intercept)+id (random intercept)
3. cloglog regression (patient-day level): any hospitalization (5 options)~study group+time+gender (2 options)+age (2 options)+morbidity (3 options)+nursing home (random intercept)+id (random intercept)

For each modeling approach on the primary dataset: 5*2*2*3=60 possible formulas.(Note: for covariates, noninclusion is an option).

Claims dataset

1. negative binomial regression (patient-quarter level): hospital days (5 options)~study group+time+gender (2 options)+age (2 options)+morbidity (6 options)+prior hospitalization (3 options)+nursing home (random intercept)+id (random intercept)
2. negative binomial regression (patient-quarter level): hospital admissions (5 options)~study group+time+gender (2 options)+age (2 options)+morbidity (6 options)+prior hospitalization (3 options)+nursing home (random intercept)+id (random intercept)
3. cloglog regression (patient-day level): any hospitalization (5 options)~study group+time+gender (2 options)+age (2 options)+morbidity (6 options)+prior hospitalization (3 options)+nursing home (random intercept)+id (random intercept)

For each modeling approach on the claims dataset: 5*2*2*6*3=360 possible formulas. (Note: for covariates, noninclusion is an option).

For the first and second modeling approach (count models for the number of hospital days and the number of hospitalizations at the patient-quarter level), we calculated rate ratios by exponentiating the regression coefficients for the study group variable. Note that the rate ratios in the first modeling approach refer to the number of hospital days attributable to hospitalizations in a given quarter, not the number of hospital days that took place during the quarter.

The regression coefficients of the *cloglog* models in the third modeling approach do not allow for a straightforward interpretation in terms of log odds or odds ratios. Therefore, for the third modeling approach, we calculated the intervention effect on the average participant in terms of the yearly probability of hospitalization. This was determined by calculating the predicted daily (total, ACSC, CACSC, NHSC, or SDD) hospitalization probability in the control and intervention phases for a hypothetical participant with modal or mean values for all variables and subtracting the former from the latter (random effects were disregarded for this calculation). We then multiplied the resulting value by 365 to yield the predicted difference in additional (total, ACSC, CACSC, NHSC, or SDD) hospitalizations per 365 days spent in the nursing home between the intervention and control phases.

We processed the data and ran the regression models in the RStudio (Posit Software, PBC) environment using the tidyverse [37], comorbidity [31], and glmmTMB [38] packages and visualized the model configurations (choice of dataset, control variables, and variable operationalizations) and resulting intervention effects for each modeling approach in a specification curve using the specr package [39]. For each modeling approach, we also calculated the number and share of intervention effects by direction, the mean, minimum, and maximum size of the intervention effect, and the median, minimum, and maximum *P* value across all models.

Ethical Considerations

This study was approved by the Ethics Committee of the UK Aachen (CTC-A No. 19-019; EK 463/20). All participants provided informed consent before data collection, which covered both the primary and secondary datasets. The study adhered to the ethical principles outlined in the Declaration of Helsinki.

Primary data for the study were collected by study nurses in the participating nursing homes and stored in a study database, in accordance with the ethics committee approval. Claims data were received from participating SHI funds after authorization by the relevant bodies, including the Federal Office for Social Security (Bundesamt für Soziale Sicherung) and state-level supervisory offices, in accordance with German law. Both datasets were transmitted and stored in accordance with the European General Data Protection Regulation. Detailed descriptions of data security, storage, and transfer mechanisms,

XSL•FO
RenderX

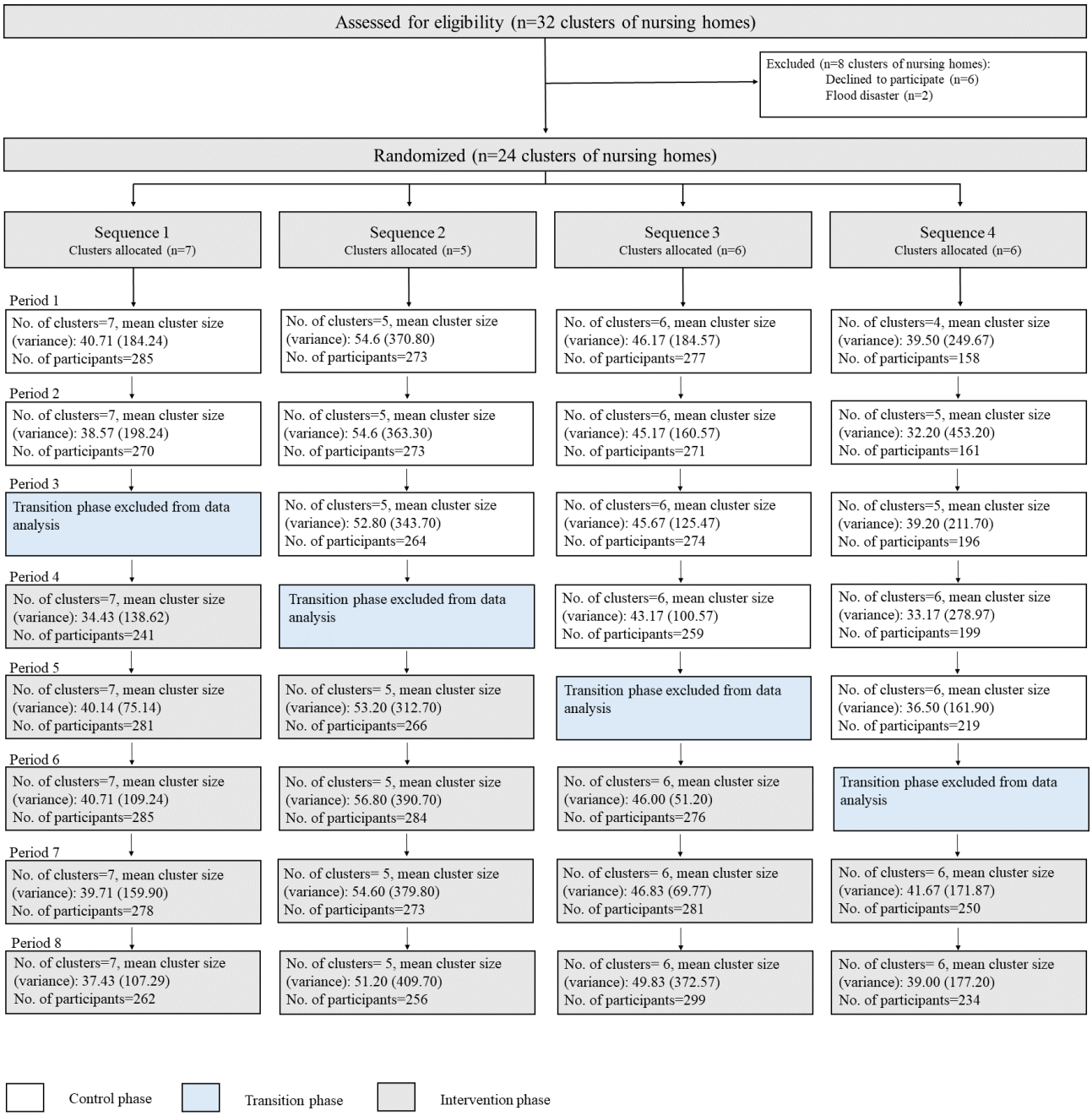
as well as further information on the legal foundation of data collection and use, can be found in the study protocol [25]. The study was registered May 10, 2021, under ClinicalTrials.gov (NCT04879537). A study protocol was published September 27, 2022, in Trials following the SPIRIT checklist [25].

Results

Recruitment and Participant Flow

Figure 2 shows the participant flowchart, which illustrates the recruitment, randomization, and progression of clusters and participants through the study phases of our stepped wedge cluster randomized trial. Initially, 32 clusters of nursing homes were assessed for eligibility. Six nursing homes declined to participate, while 2 nursing homes had to cancel their planned participation after randomization had taken place due to a flood disaster in their region, which forced those homes to shut down operations temporarily. Therefore, 2 additional nursing homes were recruited for sequence 4 in periods 2 and 3. In total, 24 nursing homes participated (in four sequences containing between 4 and 7 nursing homes each). The number of participants assessed varied across periods, ranging from 158 to 285 individuals per sequence per period. Data from the transition phase was not used in our analysis and is therefore excluded from the flowchart.

Figure 2. Study flowchart of participants and nursing homes.



XSL•FO
RenderX

Baseline Characteristics

The baseline characteristics of participants in the control and intervention phases are shown in Table 1. In both the primary and claims datasets, there were more patients, more patient-days, and a slightly lower percentage of female participants in the intervention compared to the control phase. The phases were well balanced in terms of mean age of participants and mean BI (in the BI subgroup). In the claims dataset, participants in both phases also had a similar mean Charlson Comorbidity Index (CCI), median care grade, and number of prior hospitalizations (in the past year). The BI scores indicate that the participants are highly dependent on assistance in their daily lives, while the CCI scores indicate a high level of comorbidity.

As our study followed an open cohort stepped wedge design, some patients were enrolled in both the control and intervention phases, while others were enrolled only in the control phase or only in the intervention phase. As recommended by the CONSORT extension for stepped wedge trials [21], we therefore also summarized participants' baseline characteristics by allocated sequence in Table 2.

Table . Baseline characteristics by study phase in the primary and claims datasets.

Variable	Control	Intervention
Primary dataset		
Total patients, N	1151	1271
Total patient-days, N	281,484	306,947
Age (years[a]), mean (95% CI)	81.02 (80.35-81.68)	81.32 (80.7-81.94)
Female, n (%)	769 (66.81)	826 (64.99)
BI[a,b,c] mean (95% CI)	50.37 (48.49-52.26)	50.85 (49.11-52.58)
Claims dataset		
Total patients, N	762	858
Total patient-days, N	185,329	211,213
Age (years[a]), mean (95% CI)	80.14 (79.3 - 80.98)	80.66 (79.9 - 81.43)
Female, n (%)	510 (66.93)	552 (64.34)
BI[a,c,d] mean (95% CI)	51.88 (49.53-54.22)	52.19 (50.08-54.29)
CCI[a,e] mean (95% CI)	3.10 (2.95-3.26)	2.97 (2.83-3.11)
PG[a,f,g] median	4	4
Hospitalizations in past year[a], mean (95% CI)	1.28 (1.16-1.4)	1.25 (1.13-1.37)

[a]At enrollment and/or the start of the intervention phase.

[b]Data available for subgroup only! (n=941 control phase, n=1157 intervention phase).

[c]BI: Barthel Index.

[d]Data available for subgroup only! (n=632 control phase, n=780 intervention phase).

[e]CCI: Charlson Comorbidity Index.

[f]Data available for subgroup only! (n=725 control phase, n=769 intervention phase).

[g]PG: care grade (German: Pflegegrad).

XSL•FO
RenderX

Table . Baseline characteristics by allocated sequence in the primary and claims datasets.

Variable	SEQ[a] 1	SEQ 2	SEQ 3	SEQ 4	Total
Primary dataset					
Total patients, N	421	379	431	323	1554
Total patient-days, N	155,088	158,191	159,145	116,007	588,431
Age (years[b]), mean (95% CI)	82.94 (81.99-83.88)	80.47 (79.22-81.73)	82.55 (81.6-83.48)	78.51 (77.17-79.85)	81.31 (80.75-81.86)
Female, n (%)	278 (66.03)	256 (67.55)	276 (64.04)	206 (63.78)	1016 (65.38)
BI[b,c,d], mean (95% CI)	52.43 (49.36-55.51)	51.20 (48.05-54.36)	47.14 (44.18-50.1)	54.60 (50.91 - 58.3)	51.07 (49.47 - 52.67)
Claims dataset					
Total patients, N	294	257	278	202	1031
Total patient-days, N	111,280	108,591	100,796	75,875	396,542
Age (years[b]), mean (95% CI)	82.62 (81.48 - 83.76)	79.32 (77.75 - 80.88)	81.60 (80.37 - 82.83)	77.78 (76.09 - 79.46)	80.57 (79.88 - 81.27)
Female, n (%)	199 (67.69)	165 (64.20)	175 (62.95)	131 (64.85)	670 (64.99)
BI[b,d,e], mean (95% CI)	52.94 (49.32 - 56.56)	52.33 (48.38 - 56.28)	47.93 (44.23 - 51.63)	56.56 (51.9 - 61.22)	52.09 (50.12 - 54.06)
CCI[b,g], mean (95% CI)	2.93 (2.7 - 3.16)	3.17 (2.89 - 3.44)	2.93 (2.69 - 3.17)	2.97 (2.66 - 3.28)	3.00 (2.87 - 3.13)
PG[b,f,h], median	4	3	4	4	4
Hospitalizations in past year[b], mean (95% CI)	1.34 (1.17 - 1.52)	1.50 (1.24 - 1.75)	1.26 (1.05 - 1.48)	1.59 (1.31 - 1.87)	1.41 (1.3 - 1.52)

[a]SEQ: Sequence.

[b]At enrollment.

[c]Data available for subgroup only! (n=941 control phase, n=1157 intervention phase).

[d]BI: Barthel Index.

[e]Data available subgroup only! (n=632 control phase, n=780 intervention phase).

[f]Data available subgroup only! (n=725 control phase, n=769 intervention phase).

[g]CCI: Charlson Comorbidity Index.

[h]PG: care grade (German: Pflegegrad).

Outcome Measures

The primary outcomes of our analysis were the number of hospitalizations and days spent in hospital. A descriptive comparison of outcomes by study phase is provided in Table 3. As for the baseline characteristics, we also present the number of hospitalizations and days in hospital by allocated sequence in Table 4. Notably, although the claims dataset covers fewer patients than the primary dataset, the number of ACSC and CACSC hospitalizations (both total and per patient-year in the nursing home) is higher in the claims than in the primary dataset for each of the four sequences. Furthermore, the number of NHSC hospitalizations per patient-year is also larger in the claims dataset than in the primary dataset.

XSL•FO
RenderX

Table . Hospitalizations and days in hospital by study phase in the primary and claims datasets.

Variable	Control	Intervention
Primary dataset		
Total hospitalizations, n	893	1005
Total ACSC[a] hospitalizations, n	215	213
Total CACSC[b] hospitalizations, n	162	164
Total NHSC[c] hospitalizations, n	310	302
Total SDD[d] hospitalizations, n	306	368
Hospitalizations per patient-year in NH[e], mean	1.16	1.20
ACSC hospitalizations per patient-year in NH, mean	0.28	0.25
CACSC hospitalizations per patient-year in NH, mean	0.21	0.20
NHSC hospitalizations per patient-year in NH, mean	0.40	0.36
SDD hospitalizations per patient-year in NH, mean	0.40	0.44
Total days in hospital, n	6070	5838
Total ACSC days in hospital, n	1831	1382
Total CACSC days in hospital, n	1491	1229
Total NHSC days in hospital, n	1960	1430
Total SDD days in hospital, n	306	368
Days in hospital per patient, mean	5.27	4.59
ACSC days in hospital per patient, mean	1.59	1.09
CACSC days in hospital per patient, mean	1.30	0.97
NHSC days in hospital per patient, mean	1.70	1.13
SDD days in hospital per patient, mean	0.27	0.29
Claims dataset		
Total hospitalizations, n	615	641
Total ACSC hospitalizations, n	267	278
Total CACSC hospitalizations, n	256	263
Total NHSC hospitalizations, n	248	260
Total SDD hospitalizations, n	126	105
Hospitalizations per patient-year in NH, mean	1.21	1.11
ACSC hospitalizations per patient-year in NH, mean	0.53	0.48
CACSC hospitalizations per patient-year in NH, mean	0.50	0.45
NHSC hospitalizations per patient-year in NH, mean	0.49	0.45
SDD hospitalizations per patient-year in NH, mean	0.25	0.18
Total days in hospital, n	6106	5775
Total ACSC days in hospital, n	1570	1917
Total CACSC days in hospital, n	1512	1848
Total NHSC days in hospital, n	1414	1766

XSL•FO
RenderX

Variable	Control	Intervention
Total SDD days in hospital, n	126	105
Days in hospital per patient, mean	8.01	6.73
ACSC days in hospital per patient, mean	2.06	2.23
CACSC days in hospital per patient, mean	1.98	2.15
NHSC days in hospital per patient, mean	1.86	2.06
SDD days in hospital per patient, mean	0.17	0.12

[a]ACSC: Ambulatory care-sensitive condition.

[b]CACSC: Core ambulatory care-sensitive condition.

[c]NHSC: Nursing home-sensitive condition.

[d]SDD: Same-day discharges.

[e]NH: Nursing home.

XSL•FO
RenderX

Table . Hospitalizations and days in hospital by allocated sequence in the primary and claims datasets.

Variable	SEQ[a] 1	SEQ 2	SEQ 3	SEQ 4	Total
Primary dataset					
Total hospitalizations, n	498	481	496	423	1898
Total ACSC[b] hospitalizations, n	131	124	83	90	428
Total CACSC[c] hospitalizations, n	95	96	63	72	326
Total NHSC[d] hospitalizations, n	171	152	176	113	612
Total SDD[e] hospitalizations, n	177	136	204	157	674
Hospitalizations per patient-year in NH[f], mean	1.17	1.11	1.14	1.33	1.18
ACSC hospitalizations per patient-year in NH, mean	0.31	0.29	0.19	0.28	0.27
CACSC hospitalizations per patient-year in NH, mean	0.22	0.22	0.14	0.23	0.20
NHSC hospitalizations per patient-year in NH, mean	0.40	0.35	0.40	0.36	0.38
SDD hospitalizations per patient-year in NH, mean	0.42	0.31	0.47	0.49	0.42
Total days in hospital, n	3198	3093	2812	2802	11,905
Total ACSC days in hospital, n	896	839	623	855	3213
Total CACSC days in hospital, n	764	669	495	792	2720
Total NHSC days in hospital, n	912	939	809	730	3390
Total SDD days in hospital, n	177	136	204	157	674
Days in hospital per patient, mean	7.60	8.16	6.52	8.67	7.66
ACSC days in hospital per patient, mean	2.13	2.21	1.45	2.65	2.07
CACSC days in hospital per patient, mean	1.81	1.77	1.15	2.45	1.75
NHSC days in hospital per patient, mean	2.17	2.48	1.88	2.26	2.18
SDD days in hospital per patient, mean	0.42	0.36	0.47	0.49	0.43
Claims dataset					
Total hospitalizations, n	338	380	264	274	1256
Total ACSC hospitalizations, n	154	144	118	129	545

XSL•FO
RenderX

Variable	SEQ[a] 1	SEQ 2	SEQ 3	SEQ 4	Total
Total CACSC hospitalizations, n	147	134	113	125	519
Total NHSC hospitalizations, n	147	128	109	124	508
Total SDD hospitalizations, n	67	57	36	71	231
Hospitalizations per patient-year in NH, mean	1.11	1.28	0.96	1.32	1.16
ACSC hospitalizations per patient-year in NH, mean	0.51	0.48	0.43	0.62	0.50
CACSC hospitalizations per patient-year in NH, mean	0.48	0.45	0.41	0.60	0.48
NHSC hospitalizations per patient-year in NH, mean	0.48	0.43	0.39	0.60	0.47
SDD hospitalizations per patient-year in NH, mean	0.22	0.19	0.13	0.34	0.21
Total days in hospital, n	2889	3896	2553	2570	11,881
Total ACSC days in hospital, n	1065	774	795	853	3487
Total CACSC days in hospital, n	1032	723	766	839	3360
Total NHSC days in hospital, n	986	674	701	819	3180
Total SDD days in hospital, n	67	57	36	71	231
Days in hospital per patient, mean	9.83	15.05	9.18	12.72	11.52
ACSC days in hospital per patient, mean	3.62	3.01	2.86	4.22	3.38
CACSC days in hospital per patient, mean	3.51	2.81	2.76	4.15	3.26
NHSC days in hospital per patient, mean	3.35	2.62	2.52	4.05	3.08
SDD days in hospital per patient, mean	0.23	0.22	0.13	0.35	0.22

[a]SEQ: sequence.

[b]ACSC: Ambulatory care-sensitive condition.

[c]CACSC: Core ambulatory care-sensitive condition.

[d]NHSC: Nursing home-sensitive condition.

[e]SDD: Same-day discharges.

[f]NH: nursing home.

The same holds for days in hospital: the number of ACSC and CACSC days in hospital (both total and per patient) is higher in the claims than in the primary dataset, as is the number of NHSC days in hospital per patient. Notably, despite the claims dataset containing 643 fewer total hospitalizations than the primary dataset, it contained only 24 fewer days in hospital. Overall, these results indicate substantial differences in the quantity and quality of hospitalization data between the primary and claims datasets (ie, incomplete hospitalization information in the primary dataset), necessitating a separate analysis for each dataset.

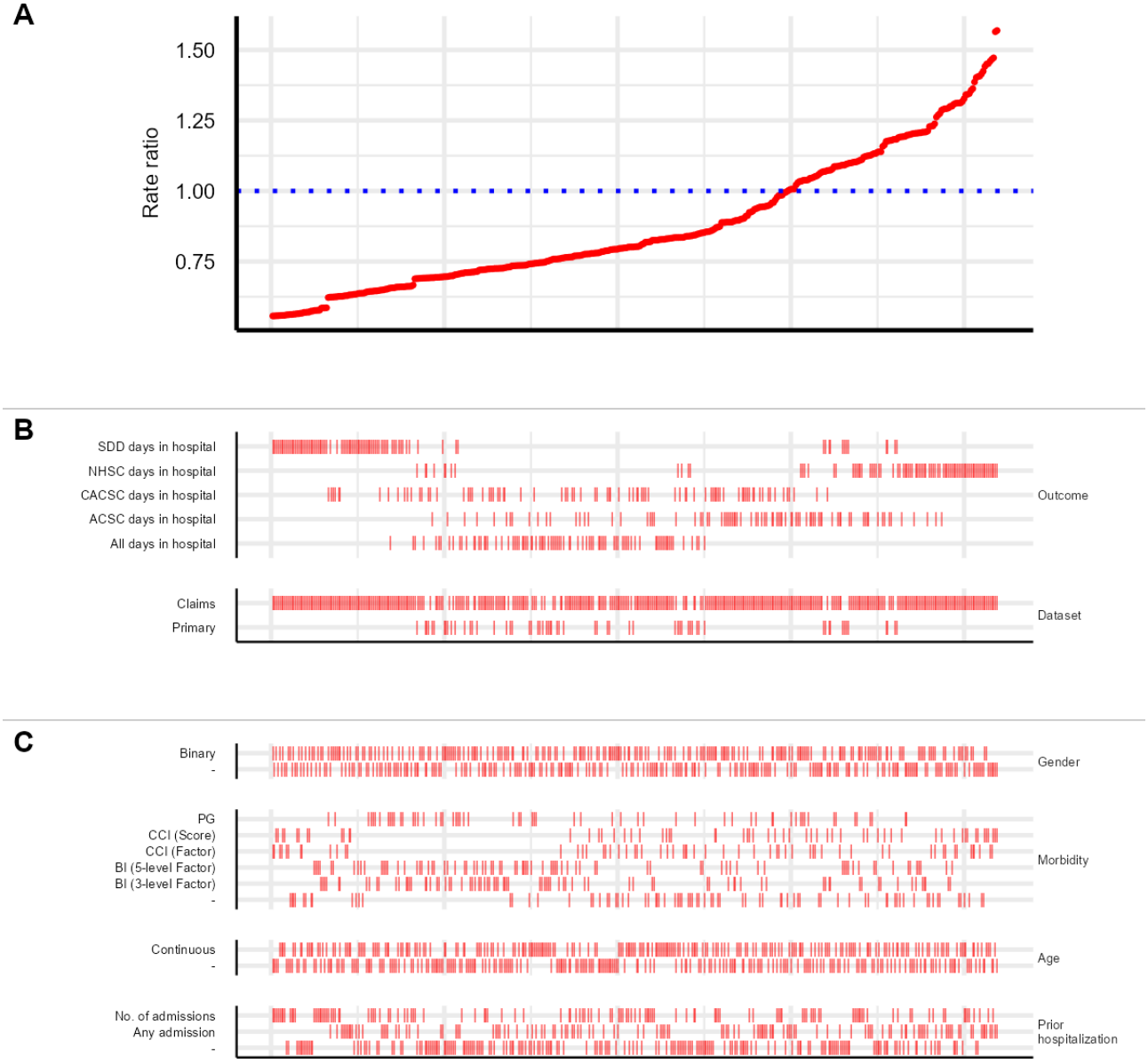
XSL•FO
RenderX

Specification Curve Analysis

Modeling Approach 1: Count Models for Days in Hospital

For modeling approach 1, we ran a total of 420 count regression models with days in hospital as the dependent variable. Of these, one model failed to converge, leaving 419 models for analysis. To illustrate the effect of our modeling choices on the size and direction of the intervention effect, Figure 3 shows the specification curve for the 419 model specifications of modeling approach 1.

Figure 3. Specification curve for modeling approach 1. SDD: same-day discharges; NHSC: nursing home-sensitive conditions; CACSC: core ambulatory care-sensitive conditions; ACSC: ambulatory care-sensitive conditions; PG: care grade (German: Pflegegrad); CCI: Charlson Comorbidity Index; BI: Barthel Index.



The specification curve reveals that the operationalization of the outcome variable has the most noticeable influence on the size of the intervention effect. In particular, SDD models had the lowest rate ratios, while NHSC models had the largest rate ratios. The remaining model specification choices did not have as clear a systematic effect on the size of the intervention effect.

Table 5 shows the proportion of negative intervention effects, the mean, minimum, and maximum effect sizes as well as the median, minimum, and maximum *P* values for the 419 regression models: overall, by outcome variable, and by dataset. Over 71% (298/419) of regression models found a negative (ie, protective) intervention effect, with a mean rate ratio of 0.88. This intervention effect was nonsignificant for all 419 regression models, with a median *P* value of .54 (IQR .34-.75) (although some SDD models exhibited *P* values as low as .08).

NHSC models had the lowest proportion of negative effects (12/84, or less than 15%). Effect sizes (rate ratios) ranged from 0.56 to 1.57, with a mean rate ratio of 0.88. Models using the primary dataset had a higher proportion of negative effects (and a slightly lower mean rate ratio) than those using the claims dataset.

XSL•FO
RenderX

Table . Effect direction, size and significance for modeling approach 1 by outcome variable.

Outcome variable (days in hospital)	Negative effects	RR[a], mean (range)	*P* value, median (range)
Overall (N=419), n (%)	298, 71.12	0.88 (0.56-1.57)	.54 (.08->.99)
By outcome variable (days in hospital), n (%)			
All (n=84)	84 (100.00)	0.76 (0.66-0.85)	.37 (.19-.56)
ACSC[b] (n=84)	51 (60.71)	0.95 (0.69-1.29)	.82 (.48-.99)
CACSC[c] (n=83)	79 (95.18)	0.80 (0.62-1.07)	.66 (.38->.99)
NHSC[d] (n=84)	12 (14.29)	1.18 (0.69-1.57)	.65 (.36-.95)
SDD[e] (n=84)	72 (85.71)	0.68 (0.56-1.18)	.20 (.08-.78)
By dataset, n (%)			
Primary (n=59)	47 (79.66)	0.83 (0.69-1.18)	.53 (.26-.78)
Claims (n=360)	251 (69.72)	0.88 (0.56-1.57)	.55 (.08->.99)

[a]RR: Relative risk.

[b]ACSC: Ambulatory care-sensitive condition.

[c]CACSC: Core ambulatory care-sensitive condition.

[d]NHSC: Nursing home-sensitive condition.

[e]SDD: Same-day discharges.

Modeling Approach 2: Count Models for Hospitalizations

For modeling approach 2, we ran a total of 420 count regression models with the number of hospitalizations as the response. Of these models, none failed to converge, leaving 420 models for analysis. Figure 4 shows the specification curve for these 420 model specifications.

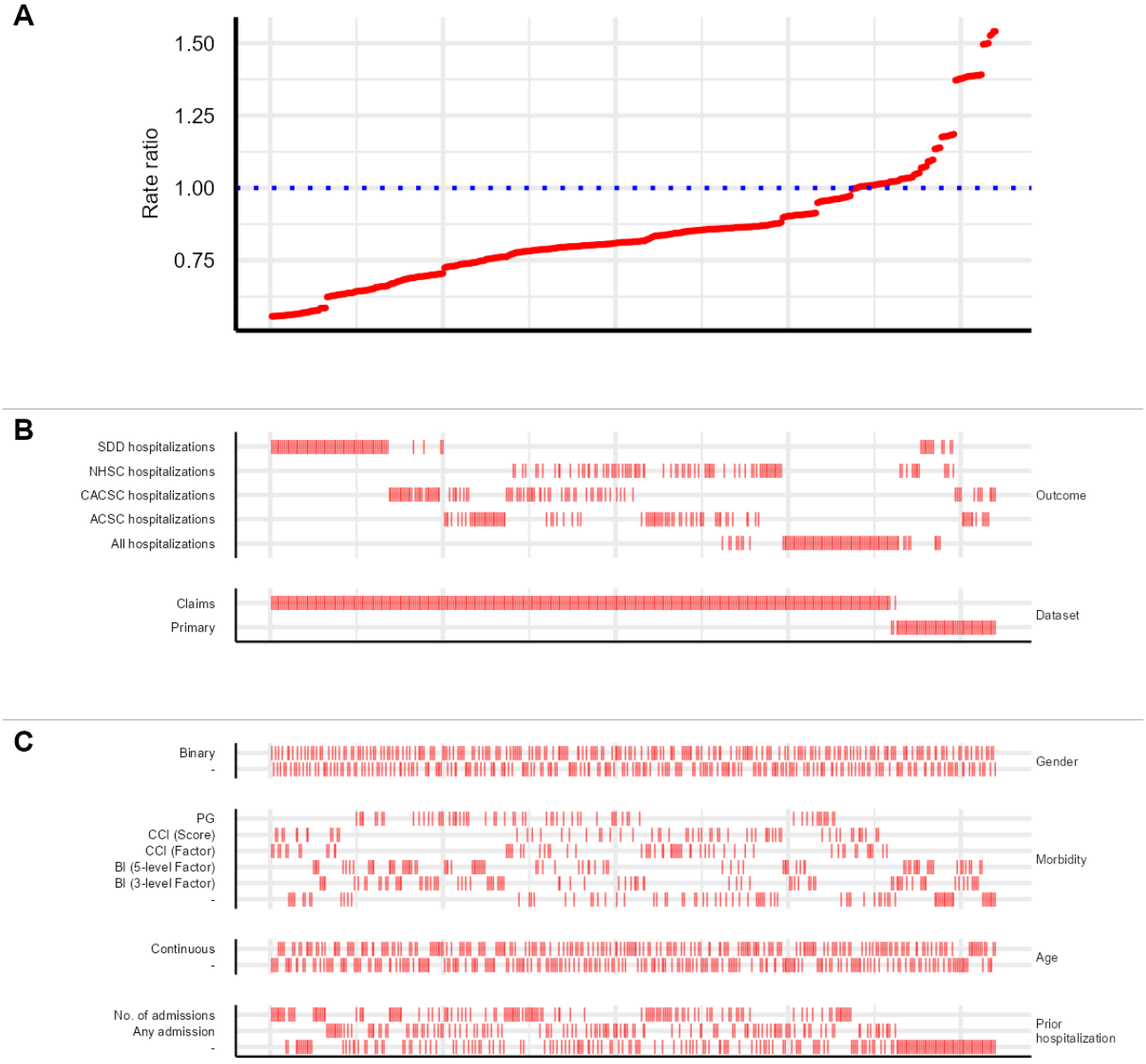
The specification curve reveals that the choice of dataset and the operationalization of the outcome variable have the most noticeable influence on the size of the intervention effect. In particular, models using the primary dataset and SDD models had larger rate ratios than those using the claims dataset or other outcome operationalizations. Within the claims dataset, models using all hospitalizations as the outcome had larger rate ratios than those using other outcome operationalizations. The remaining model specification choices did not have as clear a systematic effect on the size of the intervention effect.

Table 6 shows the proportion of negative intervention effects, the mean, minimum, and maximum effect sizes as well as the median, minimum, and maximum *P* values for the 420 regression models: overall, by outcome variable, and by dataset. Over 80% (339/420) of regression models found a negative (ie, protective) intervention effect, with a mean rate ratio of 0.85. As for modeling approach 1, the intervention effect was nonsignificant for all 420 regression models, with a median *P* value of .42 (IQR .26-.57; although some SDD models exhibited *P* values as low as .08).

Models using all hospitalizations as the outcome had the lowest proportion of negative effects at 60.71% (51/84). Effect sizes (rate ratios) ranged from 0.56 to 1.54, with a mean rate ratio of 0.85. Table 6 also confirms the notable difference in the proportion of negative effects and mean rate ratio between the primary and the claims datasets.

XSL•FO
RenderX

Figure 4. Specification curve for modeling approach 2. SDD: same-day discharges; NHSC: nursing home-sensitive conditions; CACSC: core ambulatory care-sensitive conditions; ACSC: ambulatory care-sensitive conditions; PG: care grade (German: Pflegegrad); CCI: Charlson Comorbidity Index; BI: Barthel Index.



XSL•FO
RenderX

Table . Effect direction, size, and significance for modeling approach 2 by outcome and dataset.

Outcome variable (days in hospital)	Negative effects	RR[a], mean (range)	*P* value, median (range)
Overall (N=420), n (%)	339 (80.71)	0.85 (0.56-1.54)	.42 (.08-.99)
By outcome (days in hospital) (n=84), n (%)			
All	51 (60.71)	0.97 (0.86-1.14)	.82 (.36-.99)
ACSC[b]	72 (85.71)	0.89 (0.72-1.50)	.40 (.12-.57)
CACSC[c]	72 (85.71)	0.84 (0.67-1.54)	.28 (.13-.42)
NHSC[d]	72 (85.71)	0.87 (0.77-1.19)	.51 (.39-.90)
SDD[e]	72 (85.71)	0.68 (0.56-1.18)	.20 (.08-.78)
By dataset, n (%)			
Primary (n=60)	0 (0.00)	1.22 (1.02-1.54)	.45 (.12-.90)
Claims (n=360)	339 (94.17)	0.79 (0.56-1.02)	.41 (.08-.99)

[a]RR: Relative risk.

[b]ACSC: Ambulatory care-sensitive condition.

[c]CACSC: Core ambulatory care-sensitive condition.

[d]NHSC: Nursing home-sensitive condition.

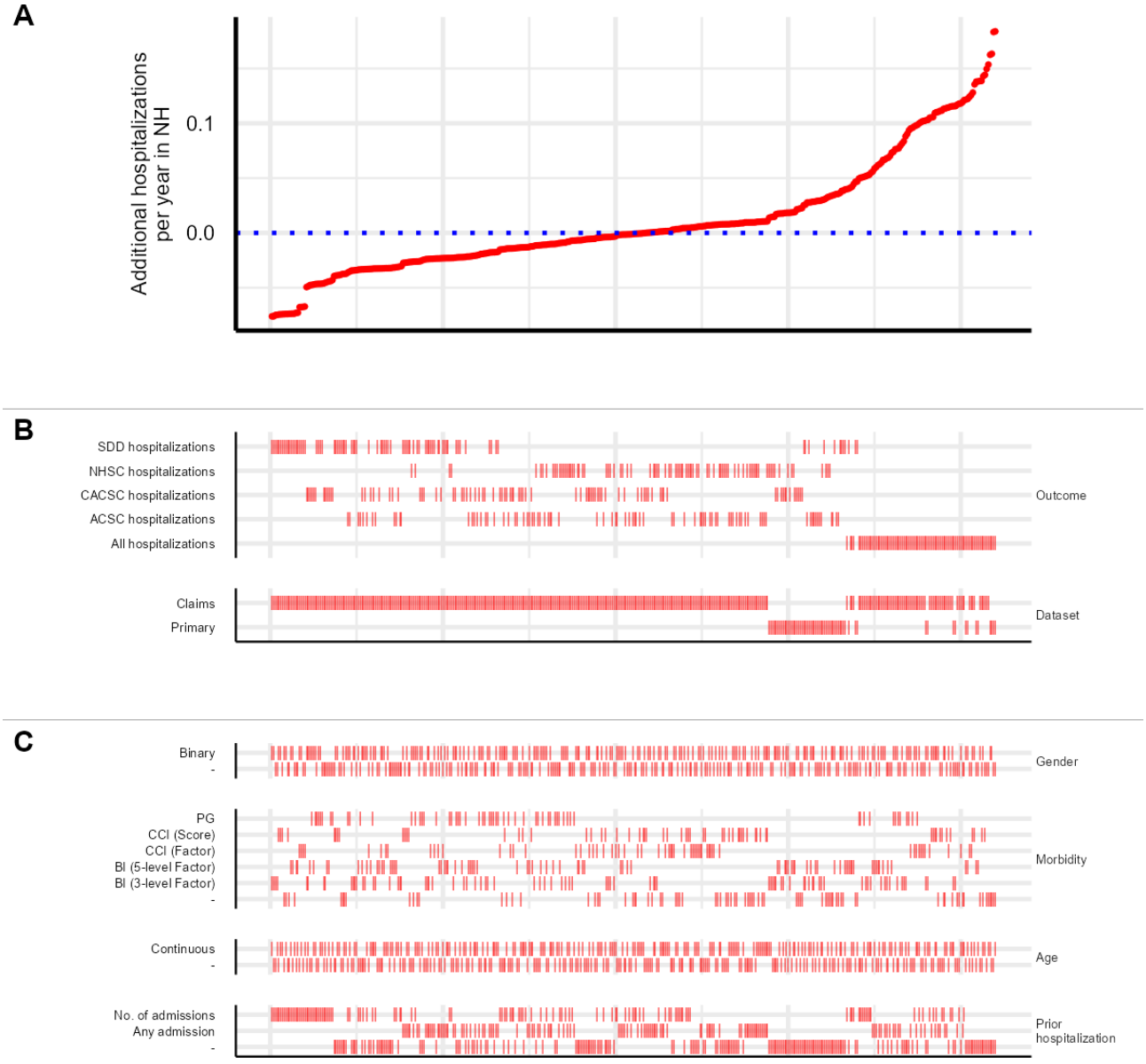
[e]SDD: Same-day discharges.

Modeling Approach 3: Binary Models for Hospitalization

For modeling approach 3, we ran a total of 420 binomial regression models with daily hospitalization as the (binary) response. Of these models, none failed to converge, leaving 420 models for analysis. Figure 5 shows the specification curve for these 420 model specifications.

XSL•FO
RenderX

Figure 5. Specification curve for modeling approach 3. SDD: same-day discharges; NHSC: nursing home-sensitive conditions; CACSC: core ambulatory care-sensitive conditions; ACSC: ambulatory care-sensitive conditions; PG: care grade (German: Pflegegrad); CCI: Charlson Comorbidity Index; BI: Barthel Index



As for modeling approach 2, the specification curve reveals that the choice of dataset and the operationalization of the outcome variable have the most noticeable influence on the size of the intervention effect. In particular, models using all hospitalizations had a larger number of additional hospitalizations per 365 days spent in the nursing home than those using other outcome operationalizations. Furthermore, among those models using outcome operationalizations other than all hospitalizations, those using the primary dataset had a larger number of additional hospitalizations per 365 days spent in the nursing home than those using the claims dataset. The remaining model specification choices did not have as clear a systematic effect on the size of the intervention effect.

Table 7 shows the proportion of negative intervention effects, the mean, minimum, and maximum effect sizes as well as the median, minimum, and maximum *P* values for the 420 regression models. Over half of the regression models (217/420, 51.67%) found a negative (ie, protective) intervention effect. Once again, the intervention effect was nonsignificant for all 420 regression models, with a median *P* value of .56 (IQR .28-.81; although some models exhibited *P* values as low as .07).

XSL•FO
RenderX

Table . Effect direction, size, and significance for modeling approach 3 by outcome and dataset.

Outcome variable (days in hospital)	Negative effects	Effect, mean (range)	*P* value, median (range)
Overall (N=420), n (%)	217 (51.67)	0.01 (–0.08 to 0.18)	.56 (.07->.99)
By outcome (days in hospital) (n=84), n (%)			
All	0 (0.00)	0.10 (0.04 to 0.18)	.37 (.07-.76)
ACSC[a]	48 (57.14)	0.00 (–0.04 to 0.04)	.78 (.09->.99)
CACSC[b]	65 (77.38)	–0.01 (–0.05 to 0.02)	.58 (.10->.99)
NHSC[c]	32 (38.10)	0.00 (–0.03 to 0.03)	.84 (.30-.99)
SDD[d]	72 (85.71)	–0.03 (–0.08 to 0.05)	.22 (.12-.62)
By dataset, n (%)			
Primary (n=60)	0 (0.00)	0.05 (0.01 to 0.18)	.26 (.07-.72)
Claims (n=360)	217 (60.28)	0.00 (–0.08 to 0.15)	.62 (.12->.99)

[a]ACSC: Ambulatory care-sensitive condition.

[b]CACSC: Core ambulatory care-sensitive condition.

[c]NHSC: Nursing home-sensitive condition.

[d]SDD: Same-day discharges.

Models with all hospitalizations as the outcome had the lowest proportion of negative effects, as all such models found a positive (ie, nonprotective) intervention effect. Effect sizes (the number of additional hospitalizations per 365 days spent in the nursing home) ranged from −0.08 to 0.18, with a mean of 0.01. This indicates that the absolute size of most intervention effects was not clinically relevant. As for modeling approach 2, zero of the 60 models on the primary dataset found a negative intervention effect, although the proportion of negative effects in the claims dataset was lower than for modeling approach 2.

Discussion

Principal Findings

We investigated the effect of an intersectoral telemedicine intervention on hospitalization probability for residents in nursing homes. Our analysis was based on data from a multicenter, prospective, cluster randomized controlled trial with an open cohort stepped wedge design. We conducted a specification curve analysis, for which we considered 1260 different regression model specifications, of which 1259 converged. Of these, none found a statistically significant (ie, $P<.05$) intervention effect.

While 854 of 1259 converged models found a (nonsignificant) negative intervention effect (67.83%), the distribution of these effects differed by modeling approach, choice of outcome, and dataset used. For modeling approaches 1 and 2 (count models for the number of hospital d and the number of hospitalizations at the patient-quarter level), more than 70% (298/419) and more than 80% (339/420) of the models found a negative (ie, protective) intervention effect, respectively. For modeling approach 3 (binary models for hospitalization probability at the patient-d level), only just over 50% (217/420) of models found a negative intervention effect. The differences in effect direction by choice of outcome and dataset were even more pronounced. While only 135 of 252 models using all hospitalizations (53.57%) and 116 of 252 NHSC models (46.03%) found a negative intervention effect, these proportions rose to 216 of 252 SDD models (85.71%) and 216 of 251 CACSC models (86.06%). And while almost three-quarters of models (807/1080) on the claims dataset found a negative intervention effect, only 26.26% (47/179) of models on the primary dataset did so.

Based on these results, our analysis does not support the hypothesis of a significant negative intervention effect. This may be due to either (1) issues with our data or methodology, (2) external factors such as the flood disaster and the COVID-19 pandemic, or (3) the actual noneffectiveness of the intervention.

Concerning (1), it could be argued that operationalizing the outcome using SDD or CACSC hospitalizations (where the proportion of negative intervention effects is also higher) may be more appropriate than using all hospitalizations, since the Optimal@NRW intervention specifically targeted potentially avoidable hospitalizations and included no mechanism to prevent necessary hospitalizations. Furthermore, the claims dataset contains more total ACSC and CACSC hospitalizations than the primary dataset, even though the primary dataset contains about 1.5 times as many patients and patient-days. This indicates that there may have been a data quality issue in the primary dataset (see limitations). Therefore, one could argue that using the claims dataset (where the proportion of negative intervention effects is also higher) is more appropriate than using the primary dataset.

Restricting the space of model specifications according to this line of argument can only increase the proportion of negative intervention effects. It does not change the fact that the intervention effect was nonsignificant in all models. However, since the sample size calculation was conducted for all hospitalizations (rather than one of the operationalizations of potentially avoidable hospitalizations), our study might not have

XSL•FO
RenderX

had a sufficient sample size to reliably detect intervention effects for potentially avoidable hospitalizations. This indicates that future studies explicitly focused on, and powered for the analysis of, potentially avoidable hospitalizations (and using high-quality datasets) may be better equipped to identify a (true) significant intervention effect.

Our results were also affected by the flood disaster and the COVID-19 pandemic (see limitations). These external factors directly impacted operations in the participating nursing homes, which might have affected the number of hospitalizations (and potentially avoidable hospitalizations, in particular). For instance, studies showed that there were fewer hospital admissions in general during the COVID-19 pandemic [13,14] in Germany. The potential impact of COVID-19 restrictions at the outset of our study may have led to more avoided hospitalizations compared to later phases, complicating the distinction between external influences and the intervention's true effect. This indicates that future studies conducted without the influence of these external factors may reach a different result than our analysis.

However, the possibility remains that our results reflect the true noneffectiveness of the intervention. This would be surprising, as several international studies have demonstrated an effect of telemedicine in nursing homes reducing hospital admissions (particularly emergency admissions) [40,41]. For instance, Joseph et al (2020) [41] found that NHR receiving telemedicine in acute care situations were less likely to be transferred to emergency departments than a control group.

In Germany, a number of pilot projects [42,43] have emerged that align with our objectives by integrating telemedicine solutions to enhance acute care management in order to reduce potentially avoidable hospitalizations. A previous study, conducted in the same region as the Optimal@NRW project, evaluated a tele-emergency medical service, in which a remote emergency physician supported paramedics to reduce the need for on-site physician involvement [42]. In a randomized controlled trial, the researchers demonstrated that 58% of patients in the intervention group could be dispatched solely by paramedics, compared to only 6% in the control group. This finding underlines the feasibility and suitability of telemedicine approaches in acute care situations and highlights that the successful implementation of telemedicine approaches requires a significant adaptation period before reaching full operational efficiency [42]. The nursing homes in our study exhibited an increasing use of teleconsultations over time. It therefore remains questionable for our study whether an intervention phase of 6 to 15 months was sufficient to fully realize the intervention's potential and achieve sustained effectiveness. Moreover, our intervention did not focus exclusively on a telemedicine approach for acute care situations but also integrated mobile nonphysician medical assistants, whom telephysicians could dispatch to expand the range of on-site treatments in nursing homes. Previous pilot studies have applied similar approaches, such as the community paramedic system (Gemeindenotfallsanitäter), which aimed to provide on-site medical assessment and care for low-acuity emergencies. Studying the Gemeindenotfallsanitäter, Seeger et al [43] demonstrated that 60% of NHR were treated on-site. While they concluded that Gemeindenotfallsanitäter saves resources of emergency services and hospitals, it is essential to note that their study primarily focused on process outcomes rather than quantifying patient-relevant clinical endpoints, as we did.

Successful implementation of telemedicine and complex interventions also strongly depends on the motivation of affected staff. Alongside our study, we systematically monitored the acceptance of participating nursing staff to gain deeper insights into the dynamics, facilitators, and barriers influencing the effective implementation of our intervention. While the results, which have been published elsewhere [44], suggested increased telemedicine acceptance during the intervention phase, these differences did not reach statistical significance. This underlines the need for an extended implementation period to foster greater acceptance over time. The numbers of teleconsultations also differed greatly between nursing homes, in that 15 nursing homes used 75% of all teleconsultations. This cannot be purely attributed to less acute care situations, but also to nursing home specific reasons, which require further research.

Rather than focusing purely on telemedicine to enhance acute care situations, we advocate for a holistic, intersectoral approach that extends beyond merely addressing acute medical needs at the point of care. Therefore, for the Optimal@NRW intervention, we chose a comprehensive approach that integrates preventive measures via our Frühwarnsystem to identify potential health risks in advance and incorporates a follow-up strategy through mobile nonphysician medical assistants to enhance on-site treatment opportunities. However, implementing multiple inherently complex components at once may overwhelm practical application due to staff shortages and workload consolidation in the German health care system. For instance, the Frühwarnsystem measurement was intended to be conducted daily by the nursing staff, but proved to be impractical. Instead, simpler and more manageable solutions are needed, such as implementing wearable technology for NHR, which does not require additional staff resources [45].

We also anticipated incorporating an intersectoral approach by engaging various stakeholders involved in acute care situations, particularly GPs. Recent qualitative and quantitative research in the German nursing sector indicates that including GPs in telemedical and digital interventions for NHR should increase access to GPs outside regular consultation hours, save time, and ease the workload of communicating with GPs for nursing home staff [46,47]. Meanwhile, initial evidence from pilot and pre-post intervention studies in Germany suggests that GP-based telemedicine interventions for NHR can reduce avoidable hospitalizations [48,49], although more evidence from high-quality randomized control trials is needed. Telemedicine interventions for NHR have also been found to reduce avoidable hospitalizations in other contexts, including France [50,51] and the United States [52].

However, during our study, we realized that including GPs proved particularly challenging, with only a very limited number ultimately participating. We invested significant time and effort into personal outreach (such as phone calls and meetings) and provided financial incentives; however, these efforts proved ineffective at recruiting GPs. Given the critical role of GPs in

XSL•FO
RenderX

coordinating and providing medical care for NHR, their inclusion in future intersectoral interventions is essential. This underscores the need for a more effective outreach strategy to enhance GP engagement and facilitate their integration into telemedicine-based acute care models.

Strengths and Limitations

To our knowledge, this is the first study evaluating a comprehensive telemedicine-based approach to treat acute care situations of NHR in Germany regarding patient-relevant clinical outcomes. The main strength of our research is the use of specification curve analysis, which allowed us to investigate the impact of different modeling choices on the direction, size, and significance of the intervention effect. It ensures the robustness of the presented results.

A methodological limitation of our approach is that our specification curve analysis is mainly descriptive. Inferential specification curve analysis is also possible and even recommended in the literature, with Simonsohn et al (2020) suggesting test statistics that include the median effect across all specifications (in our case, 0.11 additional hospitalizations per y for an average participant) and the share of specifications with significant effects in the predicted direction (in our case, none) [33]. To obtain the distributions of these test statistics, they propose modifying the data by shuffling the columns containing randomly assigned variables. For this shuffled dataset, the null hypothesis is now known to be true. Calculating the test statistic of interest for hundreds of random samples from the modified dataset then delivers the estimated distribution of the test statistic under the null hypothesis [33].

In this paper, we did not undertake an inferential specification curve analysis. This was mainly due to the structure of our data. Each row in our datasets represented a patient-day. The assignment of these patient-days to a treatment phase (intervention or control) depends on the date of the patient-day and the nursing home in which the patient was recruited. Therefore, shuffling the treatment column as described above would be inappropriate. Furthermore, even if a suitably shuffled dataset could be created, fitting 1260 mixed-effects regression models for hundreds of samples would have exceeded available computational resources. However, the lack of an inferential specification curve analysis does not represent an overly significant limitation of our results, as all intervention effects were nonsignificant.

Although the primary data was continuously checked for plausibility by a data monitoring team throughout the study, it can be assumed that there are inconsistencies in the primary dataset due to the self-documentation process and an initially paper-based procedure. This became particularly evident when compared with the SHI claims data where, for example, the number of ACSC hospitalizations per patient-year was approximately twice as high in the claims dataset as in the primary dataset. Such inconsistencies raise concerns about the internal validity of the primary dataset. Moreover, models using the (quality-assured) claims data exhibited a notably higher proportion of negative intervention effects. However, this finding must be interpreted with caution, as the smaller sample size in the claims dataset may introduce bias, potentially overestimating the observed intervention effects even though no model showed significant results. It must also be noted that relevant confounders of morbidity included in our regression models were either not captured in the primary dataset (eg, care grade) or could only be applied for a subsample of included NHR due to missing data (eg, BI).

Another limitation of this study is the assumed effect size of a 30% reduction in hospital days during the intervention phase, which served as the basis for the sample size calculation. In retrospect, this estimate appears overly optimistic, potentially leading to an underpowered study design. Furthermore, the heterogeneous utilization of intervention components across different nursing homes may have contributed to variations in the intervention's effectiveness, further complicating the ability to detect a statistically significant effect. This variability, combined with the potential underestimation of the required sample size, likely impacted the study's ability to demonstrate a significant effect, particularly when considering general hospitalization outcomes rather than specific, targeted endpoints.

Accordingly, future studies should base their sample size calculations on small-to-moderate effect sizes. Although this would increase the sample size needed, larger sample sizes may be more feasible in future projects using only claims data, since researchers would not need to collect primary data for each participant. This may also reduce dropout rates, keeping the total sample size needed manageable. Furthermore, future studies may be able to shorten the transition phase (in which participants are recruited and data is collected, but not analyzed), which would further reduce the total sample size needed. Finally, using only claims data, rather than both primary and claims data, may alleviate the concerns regarding primary data quality outlined above.

Beyond these methodological concerns, our study faced several external limitations that may have affected our data and results. A significant portion of the data collection period was impacted by the COVID-19 pandemic and the June 2021 flood disaster. These events affected participant numbers, the feasibility of the study, and data collection processes, while also influencing the analysis, as they could not be directly accounted for. Such unforeseen events may limit the generalizability of our findings.

Conclusion

Based on the results of our specification curve analysis, we conclude that the Optimal@NRW intervention had no significant effect on hospitalization probability for NHR. This could be attributed either to (1) methodological limitations of our study and analysis, (2) external factors, such as the flood disaster and COVID-19 pandemic, or (3) the actual noneffectiveness of the intervention. To address (1), future research should choose more targeted outcomes measuring potentially avoidable hospitalizations, such as SDD or CACSC, which align more closely with the intended underlying mechanisms of the intervention. To address (3), holistic, intersectoral, telemedicine-based interventions should prioritize optimizing usability for intervention components to ensure seamless integration into existing workflows, while minimizing additional workload for nursing staff. Moreover, future studies could benefit from the progressive scalability of telemedicine-based

XSL•FO
RenderX

treatments, leveraging potential upscaling scenarios that enhance both feasibility and efficiency. As familiarity and confidence in telemedicine approaches continue to grow, particularly in acute care settings, increasing trust in digital health care solutions may further strengthen their impact on patient-relevant endpoints, and long-term sustainability.

Acknowledgments

The members of the *Optimal@NRW Research Group* are as follows:

Jörg Christian Brokmann, Priv.-Doz. Dr. med.; David Brücken, Dr. med.; Christian Hübel, Dr. med.; Jenny Unterkofler, Dr. med.; Miriam Hertwig, M.Sc.; Kerstin Voncken, M.Sc.; Susanne Rademacher, M.Sc.; Wolfgang Greiner, Prof. Dr.; Sophie Pauge, M.Sc.; Birthe Aufenberg, M.Sc.; Johannes Pollmanns, Dr. rer. medic.; Sarah Oslislo, Dr. rer. medic.; Thomas Krafft, Prof. Dr.; Eva Pilot; Lahari Yaddanapudi; Maresa Neuerer; Simone Böbel; Simone Doreleijers; Dominik Groß, Prof. Dr. med. Dr. med. dent. Dr. phil.; Saskia Wilhelmy, Dr. rer. medic., M.A.; Nataliya Sira, M.A.; Martina Ziefle, Prof. Dr.; Albrecht Eisert, Priv.-Doz. Dr. rer. nat.; Pia Erdmann.

No generative AI tools were used in the preparation of this manuscript.

Funding

This work was supported by the Innovation Committee of the Federal Joint Committee (G-BA) under grant agreement number 01NVF19015. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The datasets generated and analyzed for the current study are of particular sensibility and high need for protection. Due to project-specific data protection agreements, they will not be made publicly available.

Authors' Contributions

All authors contributed to the conceptualization of the research. JG, BA and SP performed data cleaning and conducted the statistical analysis. JG and BA drafted the manuscript, which was critically reviewed by all authors.

Conflicts of Interest

None declared.

Checklist 1
CONSORT Checklist for Stepped Wedge Trials.
[DOC File, 68 KB - aging_v9i1e75721_app1.doc]

References

1. Schaeffer D, Hämel K. Cooperative care models: an international comparative perspective [Chapter in german]. In: Kriwy P, Jungbauer-Gans M, editors. Handbook of Medical Sociology: Springer VS (Wiesbaden); 2020:463-480. [doi: 10.1007/978-3-658-06392-4_26]
2. Schwinger A, Kuhlmey A, Greß S, Klauber J, Jacobs K. In: Behrendt S, editor. Nursing Report 2024: Arrival of the Baby Boomers – Challenges for Care [Book in German]: Springer; 2024. [doi: 10.1007/978-3-662-70189-8]
3. Sachverständigenrat Zur Begutachtung Der Entwicklung Im Gesundheitswesen Und In Der Pflege. Nachhaltiger Einsatz einer knappen Ressource - Gutachten 2024 - Fachkräfte im Gesundheitswesen. PUBLISSO 2024. [doi: 10.4126/FRL01-006473488]
4. Schehadat MS, Scherer G, Groneberg DA, Kaps M, Bendels MHK. Outpatient care in acute and prehospital emergency medicine by emergency medical and patient transport service over a 10-year period: a retrospective study based on dispatch data from a German emergency medical dispatch centre (OFF-RESCUE). BMC Emerg Med 2021 Mar 9;21(1):29. [doi: 10.1186/s12873-021-00424-4] [Medline: 33750317]
5. Saviteer SM, Samsa GP, Rutala WA. Nosocomial infections in the elderly. Am J Med 1988 Apr;84(4):661-666. [doi: 10.1016/0002-9343(88)90101-5]
6. Isigi SS, Parsa AD, Alasqah I, Mahmud I, Kabir R. Predisposing factors of nosocomial infections in hospitalized patients in the United Kingdom: systematic review. JMIR Public Health Surveill 2023 Dec 19;9:e43743. [doi: 10.2196/43743] [Medline: 38113098]
7. Siddiqi N, House AO, Holmes JD. Occurrence and outcome of delirium in medical in-patients: a systematic literature review. Age Ageing 2006 Jul;35(4):350-364. [doi: 10.1093/ageing/afl005] [Medline: 16648149]

XSL•FO
RenderX

8. Ormseth CH, LaHue SC, Oldham MA, Josephson SA, Whitaker E, Douglas VC. Predisposing and precipitating factors associated with delirium: a systematic review. JAMA Netw Open 2023 Jan 3;6(1):e2249950. [doi: 10.1001/jamanetworkopen.2022.49950] [Medline: 36607634]
9. Gill TM, Allore HG, Holford TR, Guo Z. Hospitalization, restricted activity, and the development of disability among older persons. JAMA 2004 Nov 3;292(17):2115-2124. [doi: 10.1001/jama.292.17.2115] [Medline: 15523072]
10. Loyd C, Markland AD, Zhang Y, et al. Prevalence of hospital-associated disability in older adults: a meta-analysis. J Am Med Dir Assoc 2020 Apr;21(4):455-461. [doi: 10.1016/j.jamda.2019.09.015] [Medline: 31734122]
11. Sundmacher L, Fischbach D, Schuettig W, Naumann C, Augustin U, Faisst C. Which hospitalisations are ambulatory care-sensitive, to what degree, and how could the rates be reduced? Results of a group consensus study in Germany. Health Policy 2015 Nov;119(11):1415-1423. [doi: 10.1016/j.healthpol.2015.08.007] [Medline: 26428441]
12. Valk-Draad MP, Gröne O, Schult T, Stahl K, Bohnet-Joschko S. Needs-based care for nursing home residents through nursing home-sensitive hospital cases: final report [Report in German]. : German Federal Joint Committee (Innovationsfonds G-BA); 2023 URL: https://innovationsfonds.g-ba.de/downloads/beschluss-dokumente/434/2023-08-16_PSK_Ergebnisbericht.pdf [accessed 2026-07-28]
13. Weeks WB, Ventelou B, Paraponaris A. Rates of admission for ambulatory care sensitive conditions in France in 2009-2010: trends, geographic variation, costs, and an international comparison. Eur J Health Econ 2016 May;17(4):453-470. [doi: 10.1007/s10198-015-0692-y] [Medline: 25951924]
14. Shinotsuka M, Matsumura S, Okada T. Emergency admissions of ambulatory care sensitive conditions at a Japanese local hospital: an observational study. J of Gen and Family Med 2020 Nov;21(6):235-241. [doi: 10.1002/jgf2.352]
15. Selvaraj S. Enhancing healthcare access in rural communities: assessing the influence of telehealth services and information technology. Int J Sci Res 2024 Jul 5;13(7):1141-1145. [doi: 10.21275/SR24723160704]
16. Beckers R, Telepflege SV. In: Marx G, Rossaint R, Marx N, editors. Telemedicine: Foundations and Practical Application in Inpatient and Outpatient Facilities [Book in German]: Springer; 2021. [doi: 10.1007/978-3-662-60611-7_23]
17. Knörr V, Dini L, Gunkel S, et al. Use of telemedicine in the outpatient sector during the COVID-19 pandemic: a cross-sectional survey of German physicians. BMC Prim Care 2022 Apr 23;23(1):92. [doi: 10.1186/s12875-022-01699-7] [Medline: 35461212]
18. Kitschke L, Traulsen P, Waschkau A, Steinhäuser J. Determinants of the implementation of telemedicine in nursing homes: a qualitative analysis from Schleswig-Holstein [Article in German]. Z Evid Fortbild Qual Gesundhwes 2024 Jun;187:1-7. [doi: 10.1016/j.zefq.2024.05.005] [Medline: 38811297]
19. Dassel K, Busch A, Lutze M. Reducing workload through nursing software: which quality criteria matter? [Report in German]. : Bertelsmann Stiftung; 2024. [doi: 10.11586/2024004]
20. Sozialgesetzbuch (SGB) Fünftes Buch (V) – Gesetzliche Krankenversicherung: Federal Ministry of Justice, Federal Office of Justice; 1988. URL: https://www.gesetze-im-internet.de/englisch_sgb_5/englisch_sgb_5.html [accessed 2026-08-04]
21. Hemming K, Taljaard M, McKenzie JE, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. BMJ 2018 Nov 9;363:k1614. [doi: 10.1136/bmj.k1614] [Medline: 30413417]
22. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ 2015 Feb 6;350(feb06 1):h391. [doi: 10.1136/bmj.h391] [Medline: 25662947]
23. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965 Feb;14(14):61-65. [Medline: 14258950]
24. Koech L, Ströhl S, Lauerer M, et al. Redirection of patients from the emergency department to ambulatory care: a feasibility study [Article in German]. Gesundheitswesen 2024 May;86(5):339-345. [doi: 10.1055/a-2206-1738] [Medline: 38354744]
25. Brücken D, Unterkofler J, Pauge S, et al. Optimal@NRW: optimized acute care of nursing home residents using an intersectoral telemedical cooperation network - study protocol for a stepped-wedge trial. Trials 2022 Sep 27;23(1):814. [doi: 10.1186/s13063-022-06613-1] [Medline: 36167557]
26. Jacobs K, Kuhlmey A, Greß S, Klauber J, Schwinger A, editors. Care Report 2019: More Personnel in Long-Term Care – But From Where?: Springer; 2019. [doi: 10.1007/978-3-662-58935-9]
27. Hoffmann F, Schmiemann G. Influence of age and sex on hospitalization of nursing home residents: a cross-sectional study from Germany. BMC Health Serv Res 2017 Jan 19;17(1):55. [doi: 10.1186/s12913-017-2008-7] [Medline: 28103927]
28. Vossius C, Selbæk G, Šaltytė Benth J, Bergh S. Mortality in nursing home residents: a longitudinal study over three years. PLoS One 2018;13(9):e0203480. [doi: 10.1371/journal.pone.0203480] [Medline: 30226850]
29. Morishima T, Sato A, Nakata K, et al. Barthel Index-based functional status as a prognostic factor in young and middle-aged adults with newly diagnosed gastric, colorectal and lung cancer: a multicentre retrospective cohort study. BMJ Open 2021 Apr 14;11(4):e046681. [doi: 10.1136/bmjopen-2020-046681] [Medline: 33853804]
30. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989;42(8):703-709. [doi: 10.1016/0895-4356(89)90065-6] [Medline: 2760661]
31. Gasparini A. comorbidity: an R package for computing comorbidity scores. J Open Source Softw 2018 Mar 29;3(23):648. [doi: 10.21105/joss.00648]

XSL•FO
RenderX

32. Elftes Buch Sozialgesetzbuch (SGB XI) – Soziale Pflegeversicherung: Federal Ministry of Justice; 1994. URL: https://www.gesetze-im-internet.de/sgb_11/BJNR101500994.html [accessed 2026-08-04]
33. Simonsohn U, Simmons JP, Nelson LD. Specification curve analysis. Nat Hum Behav 2020;4(11):1208-1214. [doi: 10.1038/s41562-020-0912-z] [Medline: 32719546]
34. Prasetyo RB, Kuswanto H, Iriawan N, Ulama BSS. A comparison of some link functions for binomial regression models with application to school drop-out rates in East Java. Presented at: The 2nd International Conference on Science, Mathematics, Environment, and Education (ICoSMEE 2019); Jul 26-28, 2019; Surakarta, Indonesia p. 020083. [doi: 10.1063/1.5139815]
35. Alves JSB, Bazán JL, Arellano-Valle RB. Flexible cloglog links for binomial regression models as an alternative for imbalanced medical data. Biom J 2023 Mar;65(3):2100325. [doi: 10.1002/bimj.202100325] [Medline: 36529694]
36. Kitaw TA, Tilahun BD, Abate BB, Haile RN. Minimum acceptable diet and its predictors among children aged 6-23 months in Ethiopia. A multilevel cloglog regression analysis. Matern Child Nutr 2024 Jul;20(3):e13647. [doi: 10.1111/mcn.13647] [Medline: 38530126]
37. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. J Open Source Softw 2019 Nov 21;4(43):1686. [doi: 10.21105/joss.01686]
38. Brooks M, Kristensen K, Benthem K, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J 2017;9(2):378. [doi: 10.32614/RJ-2017-066]
39. Masur PK, Scharkow M. Conducting and visualizing specification curve analyses. 2023. URL: https://cran.r-project.org/web/packages/specr/index.html [accessed 2026-07-28]
40. Groom LL, McCarthy MM, Stimpfel AW, Brody AA. Telemedicine and telehealth in nursing homes: an integrative review. J Am Med Dir Assoc 2021 Sep;22(9):1784-1801. [doi: 10.1016/j.jamda.2021.02.037] [Medline: 33819450]
41. Joseph JW, Kennedy M, Nathanson LA, Wardlow L, Crowley C, Stuck A. Reducing emergency department transfers from skilled nursing facilities through an emergency physician telemedicine service. West J Emerg Med 2020 Oct 8;21(6):205-209. [doi: 10.5811/westjem.2020.7.46295] [Medline: 33207167]
42. Schröder H, Beckers SK, Borgs C, et al. Long-term effects of a prehospital telemedicine system on structural and process quality indicators of an emergency medical service. Sci Rep 2024 Jan 3;14(1):310. [doi: 10.1038/s41598-023-50924-5] [Medline: 38172217]
43. Seeger I, Günther U, Schmiemann G, Hoffmann F. Care of older patients by community emergency paramedics: comparison of community-dwellers and nursing home residents [Article in German]. Med Klin Intensivmed Notfmed 2022 Oct;117(7):542-548. [doi: 10.1007/s00063-021-00863-8] [Medline: 34529128]
44. Offermann J, Ziefle M, Optimal@NRW Research Group. You are not alone! Care professionals' acceptance of telemedicine in nursing homes comparing pre- and post-implementation evaluations. Electronics (Basel) 2024 Jul 31;13(15):3022. [doi: 10.3390/electronics13153022]
45. Frech J, Jackson D, Naber J, Masukevitch A, King T. Nursing home patient monitoring system for improved healthcare and outcomes. Presented at: 2024 IEEE International Conference on Consumer Electronics (ICCE); Jan 6-8, 2024; Las Vegas, NV p. 1-4. [doi: 10.1109/ICCE59016.2024.10444209]
46. Martin T, Veldeman S, Großmann H, Fuchs-Frohnhofen P, Czaplik M, Follmann A. Long-term adoption of televisits in nursing homes during the COVID-19 crisis and following up Into the postpandemic setting: mixed methods study. JMIR Aging 2024 Jun 6;7:e55471. [doi: 10.2196/55471] [Medline: 38842915]
47. Denny K, Gundlack J, Bay J, et al. The GP's offices are far away and the gain would be great.'-a qualitative study on nursing home staff's perspectives on digital solutions for communication issues with general practitioners. Scand J Prim Health Care 2026 Dec;44(1):2638515. [doi: 10.1080/02813432.2026.2638515] [Medline: 41943182]
48. Ohligs M, Stocklassa S, Rossaint R, Czaplik M, Follmann A. Employment of telemedicine in nursing homes: clinical requirement analysis, system development and first test results. Clin Interv Aging 2020;15:1427-1437. [doi: 10.2147/CIA.S260098] [Medline: 32884251]
49. Redeker AC, Martin T, Veldeman S, et al. Impact of regular televisits on unplanned hospital admissions of nursing home residents in rural Germany: a pre-post intervention study. BMC Geriatr 2025 Sep 8;25(1):687. [doi: 10.1186/s12877-025-06244-6] [Medline: 40922008]
50. Marjanovic N, Jonchier M, Guenezan J, Delelis-Fanien H, Reuter PG, Mimoz O. Telemedicine in nursing home residents requiring a call to an emergency medical communication center. J Am Med Dir Assoc 2024 Feb;25(2):195-200. [doi: 10.1016/j.jamda.2023.09.019] [Medline: 38623779]
51. Gayot C, Laubarie-Mouret C, Zarca K, et al. Effectiveness and cost-effectiveness of a telemedicine programme for preventing unplanned hospitalisations of older adults living in nursing homes: the GERONTACCESS cluster randomized clinical trial. BMC Geriatr 2022 Dec 22;22(1):991. [doi: 10.1186/s12877-022-03575-6] [Medline: 36550496]
52. Grabowski DC, O'Malley AJ. Use of telemedicine can reduce hospitalizations of nursing home residents and generate savings for medicare. Health Aff (Millwood) 2014 Feb;33(2):244-250. [doi: 10.1377/hlthaff.2013.0922] [Medline: 24493767]

XSL•FO
RenderX

Abbreviations

ACSC: ambulatory care-sensitive condition
BI: Barthel Index
CACSC: core ambulatory care-sensitive condition
CCI: Charlson Comorbidity Index
CONSORT: Consolidated Standards of Reporting Trials
GP: general practitioner
***ICD-10*:** *International Classification of Diseases, 10th Revision*
NHR: nursing home residents
NHSC: nursing home-sensitive condition
SDD: same-day discharge
SHI: statutory health insurance
UK Aachen : University Hospital Aachen

Edited by AF Abdulai; submitted 09.Apr.2025; peer-reviewed by CY Ong, O Pitsun; revised version received 18.May.2026; accepted 12.Jun.2026; published 19.Aug.2026.

Please cite as:
Grosser J, Pauge S, Aufenberg B, Greiner W, Hertwig M, Unterkofler J, Brücken D, Hübel C, Optimal@NRW Research Group, Brokmann JC
Effect of an Intersectoral Telemedicine Intervention on Hospitalization in Nursing Home Patients: Specification Curve Analysis of a Stepped Wedge Cluster Randomized Trial
JMIR Aging 2026;9:e75721
URL: https://aging.jmir.org/2026/1/e75721
doi:10.2196/75721

© John Grosser, Sophie Pauge, Birthe Aufenberg, Wolfgang Greiner, Miriam Hertwig, Jenny Unterkofler, David Brücken, Christian Hübel, Optimal@NRW Research Group, Jörg Christian Brokmann. Originally published in JMIR Aging (https://aging.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predicting Care Needs in Community-Dwelling Older Adults Using Explainable Machine Learning and a Multidimensional Approach Integrating Health, Social, and Environmental Factors: Cross-Sectional Study

Hyeri Shin, PhD

AgeTech-Service Convergence Major, Department of Gerontology, Kyung Hee University, 1732, Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do, Republic of Korea

Corresponding Author:
Hyeri Shin, PhD
AgeTech-Service Convergence Major, Department of Gerontology, Kyung Hee University, 1732, Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do, Republic of Korea

Abstract

Background: Rapid population aging and a worsening shortage of care workers necessitate the identification of older adults who require proactive interventions. Although machine learning (ML) has been increasingly applied in gerontology, existing studies have predominantly focused on social isolation, loneliness, depression, falls, and frailty in isolation rather than on the integrated construct of care needs.

Objective: This study aimed to develop and interpret an explainable ML model that identifies care needs in community-dwelling Korean older adults. Beyond physical health indicators such as disease and functional status, this study adopted a comprehensive approach that included mental health, cognitive function, health behaviors, and socioenvironmental determinants, such as social participation, social support, and the housing environment, to present an integrated model encompassing both health and social care needs.

Methods: Data were obtained from the 2023 Korea Senior Survey, a nationally representative sample of 10,078 community-dwelling adults aged 60 years and older. The data were split 70:30 into training (n=7054) and held-out test (n=3024) sets. Seven algorithms were compared (logistic regression, decision tree, support vector machine, random forest, gradient-boosted decision trees, extreme gradient boosting, and light gradient boosting machine) using stratified 5-fold cross-validation on the training set. Discrimination was assessed by the area under the receiver operating characteristic curve (AUC). Model interpretability used Shapley additive explanations with bootstrap stability assessment across folds.

Results: Model A (excluding activities of daily living or instrumental activities of daily living [IADL]) achieved good discrimination (AUC 0.892, 95% CI 0.866 - 0.918), adequate calibration (calibration slope=0.826), and positive clinical net benefit, demonstrating that upstream factors alone can identify older adults with care needs without relying on functional status. Shapley additive explanations analysis identified age, nutritional risk, employment status, depressive symptoms, self-rated health, cognitive function, household income, and home modification as the leading predictors, with high rank stability across cross-validation folds. Model B (including activities of daily living or IADL) yielded a higher AUC (0.976, 95% CI 0.962 - 0.989), but this reflected the near-tautological relationship between IADL and self-reported care needs rather than genuine upstream predictive value. Using an objective composite outcome yielded equivalent discrimination (AUC 0.892), supporting robustness to the outcome definition.

Conclusions: Explainable ML models offer high predictive accuracy and strong interpretability for identifying care needs among older adults. Care needs in community-dwelling Korean older adults can be identified with good discrimination, calibration, and clinical net benefit using multidimensional nonfunctional factors alone. By highlighting the significant roles of health, social, and environmental factors, this study provides empirical evidence to support evidence-based decision-making for the Long-Term Care Insurance system and integrated community care policies.

(JMIR Aging 2026;9:e93371) doi:10.2196/93371

KEYWORDS

older adults; care needs; long-term care; machine learning; explainable AI

XSL•FO
RenderX

Introduction

The intersection of rapid demographic aging and a shrinking care workforce in South Korea creates an urgent need to identify older adults who currently require care or are at elevated risk. In response to these demographic changes, South Korea introduced the Long-Term Care Insurance (LTCI) system in 2008. As shortages in the care workforce intensify, the Korean government is preparing to implement the Act on Integrated Support for Community Care (Integrated Community Care Act) in 2026. These policy initiatives aim to provide a comprehensive range of services, from mild support to intensive care, enabling older adults to maintain their lives within their communities. With the enactment of this legislation, it is essential to predict care needs across multiple dimensions; however, current eligibility assessments predominantly rely on functional status, such as activities of daily living (ADL). To effectively implement integrated care across the continuum from mild to severe needs, it is necessary to move beyond simple functional assessments and analyze predictive factors from a multidimensional perspective using big data.

Although machine learning (ML) has been increasingly applied in gerontology to address these complexities, existing studies have tended to be thematic rather than holistic. Research in this field has predominantly focused on social isolation and loneliness. For instance, ML has been used to predict social isolation among older adults in China [1] and sleep disorders among older adults with multimorbidity [2]. Similarly, domestic research in Korea has largely focused on specific geriatric syndromes, such as depression [3,4] and fall risk [5]. While valuable, these studies address isolated aspects of aging rather than the comprehensive demand for care. Recent applications of explainable ML to other geriatric outcomes have included cognitive decline [6] and frailty classification [7] in community-dwelling older adults, but none have focused on the integrated construct of care needs.

Research specifically predicting the care needs of older adults remains limited. Few studies have been conducted to predict care needs in Japan [8] and Korea [9] using public insurance claims data. However, these studies are limited by their reliance on administrative data, including medical history, health care, and long-term care service usage records, and do not consider sociodemographic and environmental factors that significantly influence individual care needs. Moreover, some studies have predicted physical disability based on ADL and instrumental activities of daily living (IADL) scores [10], as well as frailty [11]. A conceptual gap exists between functional disability, which is generally defined as a clinical measure, and perceived care needs, which are defined as a subjective demand for assistance [12].

Although the volume of multidimensional ML research in gerontology is expanding, there is a notable absence of studies applying these techniques to predict comprehensive care needs among older Korean adults using survey data. To address this gap, this study has 2 objectives. The first is to develop and interpret an explainable ML model that identifies community-dwelling older adults with care needs using multidimensional upstream factors alone, excluding functional status (ADL and IADL). The second is to characterize the relative contribution of predictors across demographic, health, behavioral, psychological, cognitive, socioeconomic, social support, and environmental domains, thereby examining whether care needs behave as a multidimensional construct. This study aims to provide a robust evidence base for the successful implementation of upcoming integrated care policies.

Methods

Data Source and Participants

This study used data from the 2023 Korea Senior Survey, a nationally representative survey initiated in 2007, to examine the characteristics of older adults in Korea. This survey collects comprehensive information across multiple domains, including health status, economic conditions, and social participation, and enables monitoring of changes over time.

A total of 10,078 community-dwelling adults aged 60 years and older were included in this analysis. Data on demographic, health-related, functional, and psychosocial characteristics were extracted for this study. Outcome prevalence was 8.7% (875/10,078, 95% CI 8.1% - 9.3%), giving an imbalance ratio of approximately 1:10.5.

Missingness was very low across the candidate predictor set (median 0.0%, IQR 0.0% - 0.0% per variable; range 0.0% - 10.6%). Two variables exceeding the prespecified 10% threshold, frequency of contact with children and frequency of contact with friends, were excluded a priori. Continuous variables were imputed using the median and categorical variables using the mode, with imputation parameters fitted only on the training fold and applied to held-out folds to prevent leakage. A complete-case sensitivity analysis yielded substantively identical results.

This study was reported in accordance with the TRIPOD+AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis+AI) statement [13]. The completed checklist is provided in Checklist 1.

Ethical Considerations

The sampling design and survey content of the 2023 National Survey of Older Koreans were officially approved by Statistics Korea (National Statistics Approval Number 117071). The study protocol was reviewed and approved by the Institutional Review Board of the Korea Institute for Health and Social Affairs (2023 - 078).

Measures

In this study, care needs were assessed using a self-reported questionnaire. Participants were asked whether they required assistance or care in their daily lives. Responses were dichotomized to define the outcome variable. Participants who reported requiring care were classified as the care-needs group and assigned a value of 1, whereas those who reported not requiring care were classified as the no–care-needs group and assigned a value of 0.

XSL•FO
RenderX

Based on previous studies and established theoretical frameworks on care needs among older adults, a comprehensive set of multidimensional predictive variables was selected. Demographic characteristics included gender, age, marital status (spouse), educational attainment, current employment status, number of household members, and household income. Subjective health variables comprised self-rated health and spouse's self-rated health. Physical health-related variables included the number of chronic conditions, BMI, sleep quality, and history of falls. Health behavior variables included current smoking status, current drinking status, and regular exercise. Nutritional health was assessed using nutritional risk indicators, including nutritional risk status and the total Nutritional Screening Initiative (NSI) score.

Health care utilization variables included outpatient or hospital visits and health-screening experience within the past 2 years. Mental and cognitive status was measured using depressive symptoms and cognitive function. Functional status was evaluated using ADL and IADL total scores. Binary limitation indicators were excluded from the candidate predictor set to avoid redundancy with the continuous totals. Social participation variables encompassed participation in leisure activities, educational activities, hobby clubs, social groups, political groups, volunteer activities, religious affiliations, and the use of senior welfare centers. Social support was assessed based on the number of available supporters when feeling depressed, when ill, or when needing to borrow money. Finally, living environment factors included housing satisfaction, home modification level, and area of residence.

Composite social-contact indices that overlapped with the outcome construct were excluded a priori as leakage candidates. From the 87 original variables, 6 were excluded a priori: 2 due to high missingness (more than 10%), 2 as redundant binary indicators of functional limitations, and 2 as leakage candidates. This yielded 81 final candidate predictors—79 were used in model A (excluding ADL and IADL totals) and 81 in model B.

Analytical Strategy

Two prespecified models were compared. Model A (primary, upstream) included all 79 candidate predictors except the ADL and IADL totals. This model targets the practical task of identifying older adults at risk before functional dependence is established, the relevant operating point for preventive screening under the Integrated Community Care Act. Model B (benchmark, full) included all 81 candidate predictors and is reported only as a benchmark to characterize the upper bound of predictive performance when functional status is observable.

The analytical process for both models proceeded through descriptive analysis, predictive modeling, feature selection, class-imbalance handling, hyperparameter specification, validation, calibration and decision-curve analysis, model interpretation, and sensitivity analyses.

First, descriptive analyses were conducted to compare the baseline characteristics of older adults with and without care needs. Independent-sample *2-tailed t* tests were used for continuous variables, and chi-square tests were applied for categorical variables.

Second, for predictive modeling, the dataset containing 10,078 samples was split 70:30 (stratified by outcome) into a training set (prevalence, 612/7054, 8.7%) and a held-out test set (prevalence, 263/3024, 8.7%). Seven ML algorithms were compared: logistic regression (LR), decision tree (DT), support vector machine (SVM), random forest (RF), gradient-boosted decision trees (GBDT), extreme gradient boosting (XGBoost), and light gradient boosting machine (LightGBM).

Third, all candidate predictors were retained for model fitting; built-in regularization within each algorithm provided implicit feature weighting (L2 penalty for LR, tree-based split importance for RF and gradient-boosted models, a radial basis function kernel with C=1.0 for SVM, and depth-limited splits for the DT). To address potential redundancy across the candidate set, post hoc Shapley additive explanation (SHAP)-based feature ranking with bootstrap-based stability assessment across cross-validation folds (Jaccard similarity of the top-10 set) was used to characterize the relative contribution of individual predictors (Multimedia Appendix 1).

Fourth, the class imbalance (1:10.5) was handled using 2 complementary approaches. For boosted models (XGBoost, LightGBM, and GBDT), a class-weighted loss function was used, with the positive class weighted by the inverse of its frequency. For nonboosted models, the synthetic minority oversampling technique [14] was applied only within each cross-validation training fold, never on validation or test partitions, to prevent information leakage. LR and SVM models were additionally fitted with class-balanced weights as a robustness check.

Fifth, hyperparameters were prespecified for each algorithm based on values commonly used in the ML literature for moderately sized clinical datasets; no algorithm-specific hyperparameter optimization was performed (XGBoost: 400 estimators, learning rate 0.05, max depth 4, subsample 0.8, colsample_bytree 0.8; LightGBM: 400 estimators, learning rate 0.05, max depth 6, num_leaves 31; GBDT: 300 estimators, learning rate 0.05, max depth 3; RF: 500 trees, min_samples_leaf 2; LR: L2 penalty with C=1.0; decision tree: max depth 6, min_samples_leaf 20; SVM: RBF kernel with C=1.0). Boosted models used positive-class scale weights computed from each training fold; nonboosted models used class-balanced weights (with synthetic minority oversampling technique applied within the cross-validation training fold for nonboosted models). The random seed was set to 42 for reproducibility. All analyses were conducted in Python 3 on Google Colab using *scikit-learn*, *statsmodels*, *SciPy*, *LightGBM*, *XGBoost*, *imbalanced-learn*, and *SHAP*.

Sixth, the validation strategy combined stratified 5-fold cross-validation on the training portion (with the mean, SD across folds reported as the primary performance estimate) with a final evaluation on the held-out test set. Discrimination was assessed using the area under the receiver operating characteristic curve (AUC), with 95% CIs from the DeLong method on the test fold and from 1000 bootstrap resamples. Additional metrics included the precision-recall AUC (PR-AUC), the prevalence-adjusted relative PR-AUC (PR-AUC ÷ outcome prevalence), sensitivity, specificity, the F_1-score,

XSL•FO
RenderX

and the Matthews correlation coefficient (MCC). Performance was reported at 3 operating thresholds, the default of 0.5, prevalence-matched (0.087), and Youden-optimal, to demonstrate the sensitivity-specificity trade-off across the relevant decision range.

Seventh, calibration was assessed on the held-out test set using the Brier score, calibration slope and intercept, and reliability diagrams in which predicted probabilities were grouped into 10 quantile-based bins. Predicted probabilities were post hoc recalibrated using isotonic regression fitted on the training portion via 5-fold cross-validated stacking. Decision-curve analysis was performed across threshold probabilities ranging from 1% to 50% [15] to assess clinical net benefit relative to the "treat all" and "treat none" reference strategies.

Eighth, model interpretability was assessed using SHAP [16]. SHAP values were computed within each of the 5 cross-validation folds, and the mean |SHAP| values were computed with bootstrap-derived 95% CIs, together with the rank stability of the top-10 features quantified by the Jaccard similarity index across folds (Multimedia Appendix 1).

Finally, prespecified sensitivity analyses included leave-one-feature-out for the top-10 predictors of model A; an alternative outcome operationalization using an objective composite (2 or more IADL limitations and/or 1 or more ADL limitations; prevalence, 1411/10,078, 14.0%); stratification of performance by depressive-symptom quartile; and complete-case analysis. Discrimination was preserved across Center for Epidemiologic Studies Depression Scale–Short Form (CES-D-SF) score quartiles (AUC range 0.86 - 0.91), arguing against differential model performance driven by depression-related self-report bias. Complete-case analysis yielded substantively identical results.

Results

Sample Characteristics

Table 1 presents the baseline characteristics of older adults according to the care-needs status. Compared with older adults without care needs, those with care needs were significantly older and more likely to be female older adults. They were also less likely to have a spouse and had substantially lower levels of educational attainment and employment participation. Household income was significantly lower in the care-needs group, whereas the number of household members did not differ significantly between the 2 groups. Regarding health status, older adults with care needs reported poorer self-rated health and poorer spouse's self-rated health, and they had a significantly higher number of chronic conditions. BMI was slightly lower in the care-needs group. Regarding social resources, the care-needs group had fewer available supporters when feeling depressed, when ill, or when needing to borrow money, indicating lower levels of social support.

XSL•FO
RenderX

Table . Respondents' demographic profile (N=10,078).

Variable	Older adults without care needs (n=9203)	Older adults with care needs (n=875)	*P* value
Demographics			
Sex, n (%)			
Female	5620 (61.1)	586 (67.0)	<.001
Male	3583 (38.9)	289 (33.0)	<.001
Age (y), mean (SD)	73.48 (6.43)	80.97 (6.90)	<.001
Spouse, n (%)			
No	3617 (39.3)	510 (58.3)	<.001
Yes	5586 (60.7)	365 (41.7)	<.001
Educational attainment, n (%)			
Elementary school or less	3820 (41.5)	629 (71.9)	<.001
Middle school	2003 (21.8)	124 (14.2)	<.001
High school	2777 (30.2)	99 (11.3)	<.001
College or higher	603 (6.6)	23 (2.6)	<.001
Currently employed, n (%)			
No	5339 (58.0)	791 (90.4)	<.001
Yes	3864 (41.9)	84 (9.6)	<.001
Number of household members, mean (SD)	1.76 (0.66)	1.72 (0.77)	.19
Household income (KR ₩, thousand)[a], mean (SD)	3341.14 (3489.78)	2311.49 (2260.18)	<.001
Subjective health, mean (SD)			
Self-rated health	3.27 (0.92)	2.97 (2.21)	<.001
Spouse self-rated health	5.54 (2.86)	6.31 (3.24)	<.001
Physical health, mean (SD)			
Number of chronic conditions	2.08 (1.51)	3.57 (2.04)	<.001
BMI (kg/m²)	23.68 (2.58)	23.06 (3.09)	<.001
Sleep quality	3.41 (0.88)	3.50 (2.06)	.19
Falls			
History of falls (past year), n (%)			
No	8730 (94.9)	709 (81.0)	<.001
Yes	473 (5.1)	166 (18.9)	<.001
Health behavior			
Current smoking, n (%)			
No	8436 (91.7)	831 (94.9)	<.001
Yes	767 (8.3)	44 (5.0)	<.001
Current drinking, n (%)			
No	5970 (64.9)	748 (85.5)	<.001
Yes	3233 (35.1)	127 (14.5)	<.001
Regular exercise, n (%)			
No	4141 (44.9)	561 (64.1)	<.001
Yes	5062 (55.0)	314 (35.9)	<.001
Nutritional health			

XSL•FO
RenderX

Variable	Older adults without care needs (n=9203)	Older adults with care needs (n=875)	*P* value
Nutritional risk (NSI)[b], n (%)			
Normal	6356 (69.1)	246 (28.1)	<.001
At risk	2847 (30.9)	629 (71.9)	<.001
Nutritional risk total score, mean (SD)	2.02 (2.32)	4.57 (3.27)	<.001
Health care utilization			
Outpatient or hospital visit, n (%)			
No	2952 (32.1)	144 (16.5)	<.001
Yes	6251 (67.9)	731 (83.5)	<.001
Health screening (past 2 years), n (%)			
No	1861 (20.2)	297 (33.9)	<.001
Yes	7342 (79.8)	578 (66.1)	<.001
Mental or cognitive health, mean (SD)			
Depressive symptoms	3.34 (8.78)	19.69 (3.49)	<.001
Cognitive function	24.87 (4.50)	20.28 (5.70)	<.001
Functional status			
ADL[c] limitation, n (%)			
No	8902 (96.7)	356 (40.7)	<.001
Yes	301 (3.3)	519 (59.3)	<.001
ADL (total score), mean (SD)	7.07 (0.54)	8.94 (2.89)	<.001
IADL[d] limitation, n (%)			
No	8317 (90.4)	2 (0.2)	<.001
Yes	886 (9.6)	873 (99.8)	<.001
IADL (total score)	10.32 (1.46)	16.51 (5.65)	<.001
Social participation			
Participation in leisure activities, n (%)			
No	1499 (16.3)	284 (32.5)	<.001
Yes	7704 (83.7)	591 (67.5)	<.001
Participation in educational activities, n (%)			
No	7886 (85.7)	795 (90.9)	<.001
Yes	1317 (14.3)	80 (9.1)	<.001
Participation in hobby clubs, n (%)			
No	8587 (93.3)	857 (97.9)	<.001
Yes	616 (6.7)	18 (2.1)	<.001
Participation in social groups, n (%)			
No	4050 (44.0)	727 (83.1)	<.001
Yes	5153 (55.9)	148 (16.9)	<.001
Participation in political groups, n (%)			
No	9065 (98.5)	869 (99.3)	.07
Yes	138 (1.5)	6 (0.7)	.07
Volunteer experience, n (%)			
No	8937 (97.1)	872 (99.7)	<.001

XSL•FO
RenderX

Variable	Older adults without care needs (n=9203)	Older adults with care needs (n=875)	*P* value
Yes	266 (2.9)	3 (0.3)	<.001
Religious affiliation, n (%)			
No	5438 (59.1)	540 (61.7)	.14
Yes	3765 (40.9)	335 (38.3)	.14
Use of senior center, n (%)			
No	6679 (72.6)	551 (62.9)	<.001
Yes	2524 (27.4)	324 (37.0)	<.001
Use of senior welfare center, n (%)			
No	8328 (90.5)	793 (90.6)	.94
Yes	875 (9.5)	82 (9.4)	.94
Social support, mean (SD)			
Number of available supporters when feeling depressed	2.41 (1.68)	1.77 (1.65)	<.001
Number of available supporters when ill	1.78 (1.32)	1.43 (1.27)	<.001
Number of available supporters to borrow money from	1.15 (1.21)	0.74 (1.15)	<.001
Living environment, mean (SD)			
Housing satisfaction	3.78 (0.71)	4.20 (1.78)	<.001
Home modification level, n (%)			
None	2860 (31.1)	167 (19.1)	<.001
Low	3286 (35.7)	263 (30.1)	<.001
Mid	2447 (26.6)	337 (38.5)	<.001
High	610 (6.6)	108 (12.3)	<.001
Area of residence, n (%)			
Rural	2696 (29.3)	321 (36.7)	<.001
Urban	6507 (70.7)	554 (63.3)	<.001

[a]Monthly household income was measured at KR ₩10,000=US $7.67 as of November 10, 2023.

[b]NSI: Nutritional Screening Initiative.

[c]ADL: activities of daily living.

[d]IADL: instrumental activities of daily living.

Variable Diagnostics

Variance inflation factors (VIFs) were computed for all numeric predictors in model A. The highest VIF was 2.31 (number of available supporters when ill), well below the conventional concerning threshold of 5 and even below the conservative threshold of 2.5. The 7 highest-VIF variables (all VIFs<2.5) reflect expected modest correlations among related health and social-support constructs, but each carries sufficient independent information to be retained. No predictor required removal on multicollinearity grounds.

Multivariable LR

As a confirmatory descriptive complement to the ML analyses, Table 2 presents the results of the multivariable LR analysis identifying independent factors associated with care needs among older adults.

XSL•FO
RenderX

Table . Multivariate stepwise logistic regression analysis of care needs[a].

Variable	β[b] (SE)	OR[c] (95% CI)	*P* value
Age	0.068 (0.008)	1.07 (1.05 - 1.09)	<.001
Number of household members	0.076 (0.092)	1.08 (0.90 - 1.29)	.41
Household income	0.000 (0.000)	1.00 (1.00 - 1.00)	.72
Self-rated health	−0.455 (0.071)	0.63 (0.55 - 0.73)	<.001
Spouse self-rated health	0.046 (0.020)	1.05 (1.01 - 1.09)	.02
Sleep quality	−0.041 (0.066)	0.96 (0.84 - 1.09)	.54
Hypertension	0.050 (0.116)	1.05 (0.84 - 1.32)	.67
Hyperlipidemia	−0.148 (0.111)	0.86 (0.69 - 1.07)	.19
Diabetes mellitus	0.158 (0.108)	1.17 (0.95 - 1.45)	.14
Osteoporosis	0.395 (0.127)	1.48 (1.16 - 1.91)	.002
Dementia	0.571 (0.260)	1.77 (1.06 - 2.94)	.03
BMI	0.019 (0.017)	1.02 (0.99 - 1.05)	.26
Nutritional risk (total score)	0.096 (0.018)	1.10 (1.06 - 1.14)	<.001
Depressive symptoms	0.037 (0.017)	1.04 (1.00 - 1.07)	.03
Cognitive function	−0.027 (0.011)	0.97 (0.95 - 0.99)	.02
Activities of daily living	−0.218 (0.059)	0.80 (0.72 - 0.90)	<.001
Instrumental activities of daily living	0.496 (0.026)	1.64 (1.56 - 1.73)	<.001
Number of available supporters when feeling depressed	0.006 (0.043)	1.01 (0.93 - 1.09)	.89
Number of available supporters when ill	0.056 (0.058)	1.06 (0.94 - 1.18)	.33
Number of available supporters to borrow money from	−0.102 (0.055)	0.90 (0.81 - 1.00)	.06
Housing satisfaction	0.254 (0.079)	1.29 (1.11 - 1.50)	.001
Home modification level	0.056 (0.024)	1.06 (1.01 - 1.11)	.02

[a]Observations=9951; Pseudo-R^2=0.4432; log-likelihood=−1528.1, LL-null=–2744.4; and LLR *P*<.001.

[b]β: regression coefficient.

[c]OR: odds ratio.

Older age was significantly associated with higher odds of care needs, indicating an increased likelihood of requiring care with advancing age. In contrast, better self-rated health was associated with lower odds of care needs, suggesting a protective effect. Regarding health-related factors, spouse's poor self-rated health, osteoporosis, and dementia were significantly associated with increased odds of care needs. Additionally, higher nutritional risk scores and greater depressive symptoms were positively associated with care needs, indicating that nutritional vulnerability and mental health status play important roles.

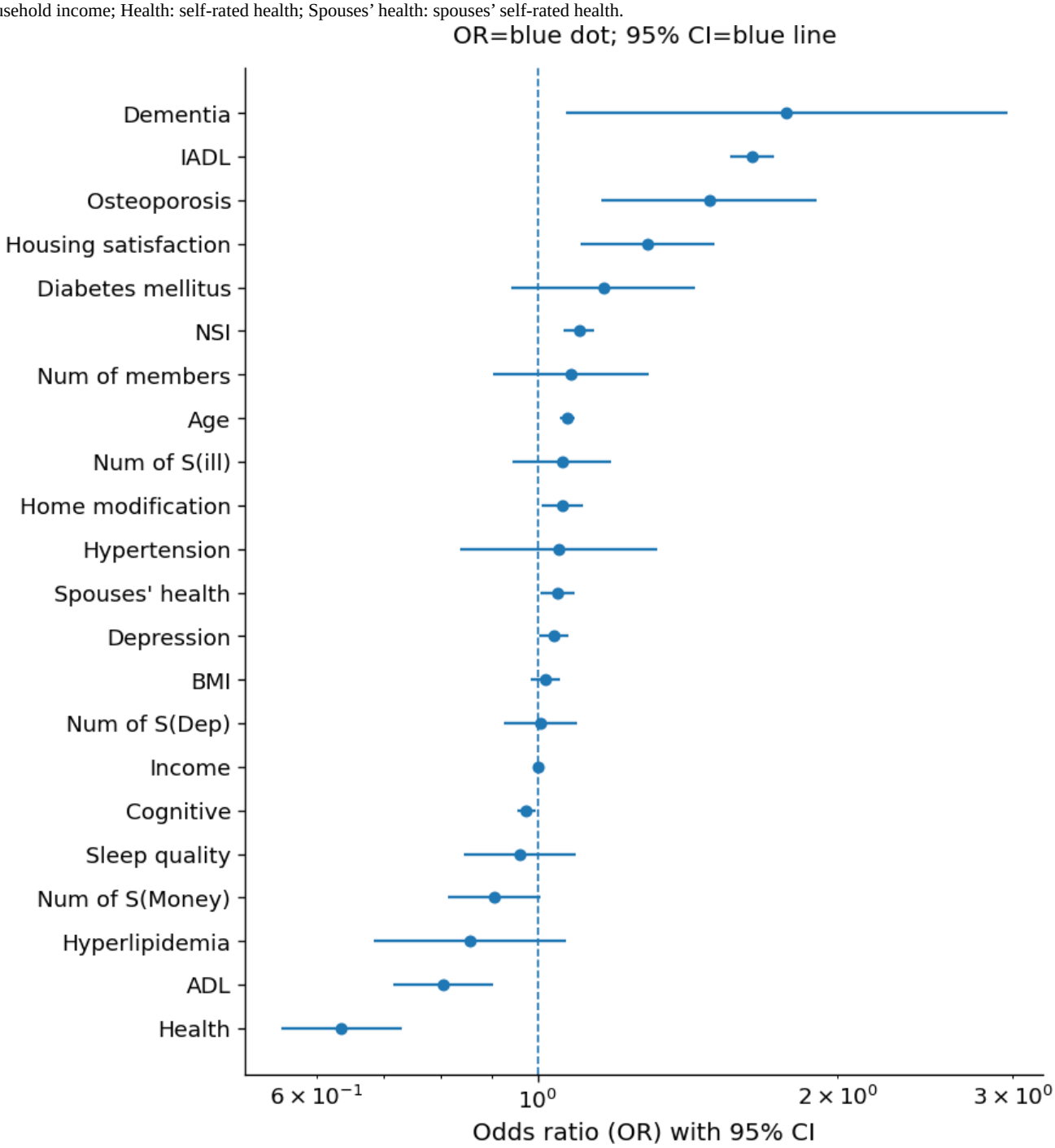
Functional status showed strong and distinct associations. Greater limitations in IADL were associated with higher odds of care needs, whereas better performance in basic ADL was associated with lower odds. Higher cognitive function scores were also inversely associated with care needs. Regarding the residential environment, higher housing satisfaction and home modifications were significantly associated with increased odds of care needs, suggesting that housing conditions and environmental adaptations reflect underlying care demands.

Overall, care needs among older adults were independently associated with advanced age, poor health and nutritional status, mental and cognitive vulnerability, functional impairment, and housing-related factors.

Figure 1 visually summarizes the adjusted odds ratios and 95% CIs from the final multivariable LR model. The vertical dashed line indicates an odds ratio of 1.0, representing a null association. This demonstrates that the determinants of care needs vary substantially in magnitude, underscoring the heterogeneous contributions of functional, health-related, and environmental factors within the multivariable model.

XSL•FO
RenderX

Figure 1. Various factors are associated with the specific care needs of the population. ADL: activities of daily living; IADL: instrumental activities of daily living; NSI: Nutritional Screening Initiative total score; Num of members: number of household members; Num of S(Dep): number of supporters when feeling depressed; Num of S(ill): number of supporters when ill; Num of S(Money): number of supporters to borrow money from; Income: household income; Health: self-rated health; Spouses' health: spouses' self-rated health.



Model Performance

Cross-validated performance (5-fold stratified on the training set) is shown in Table 3. GBDT achieved the highest AUC (mean 0.913, SD 0.007), narrowly ahead of XGBoost (mean 0.909, SD 0.010) and RF (mean 0.907, SD 0.007). All 7 algorithms achieved an AUC of more than 0.85 in model A, indicating that the upstream signal is robust to algorithm choice and not dependent on any single inductive bias.

XSL•FO
RenderX

Table . Performance of 7 ML algorithms on the held-out test set (N=3024) for model A (primary, no activities of daily living [ADL] or instrumental activities of daily living [IADL]) and model B (benchmark, all features).

Feature set and algorithm	AUC[a]	Sensitivity	Specificity	F_1-score	Brier
Model A					
GBDT[b] (best, 95% CI)	0.913 (CV)[d]; 0.892 (0.866 - 0.918)	0.882	0.768	0.409	0.055
XGBoost[c]	0.909 (CV)[d]	0.864	0.815	—[e]	0.076
Random forest	0.907 (CV)[d]	0.863	0.815	—	0.057
LightGBM[f]	0.903 (CV)[d]	0.832	0.831	—	0.066
Logistic	0.901 (CV)[d]	0.859	0.795	—	0.116
SVM[g]	0.869 (CV)[d]	0.856	0.753	—	0.077
Decision tree	0.855 (CV)[d]	0.856	0.723	—	0.101
Model B					
GBDT (best)	0.978 (CV)[d]	0.996	0.914	0.687	0.036

[a]AUC: area under the receiver operating characteristic curve.

[b]GBDT: gradient-boosted decision trees.

[c]XGBoost: extreme gradient boost.

[d]CV=mean across 5-fold cross-validation; 95% CI for held-out test set is shown only for the lead (best) algorithm in each feature set. The full table including operating thresholds and calibration is provided in Multimedia Appendix 2.

[e]Not applicable.

[f]LightGBM: light gradient boosting machine.

[g]SVM: support vector machine.

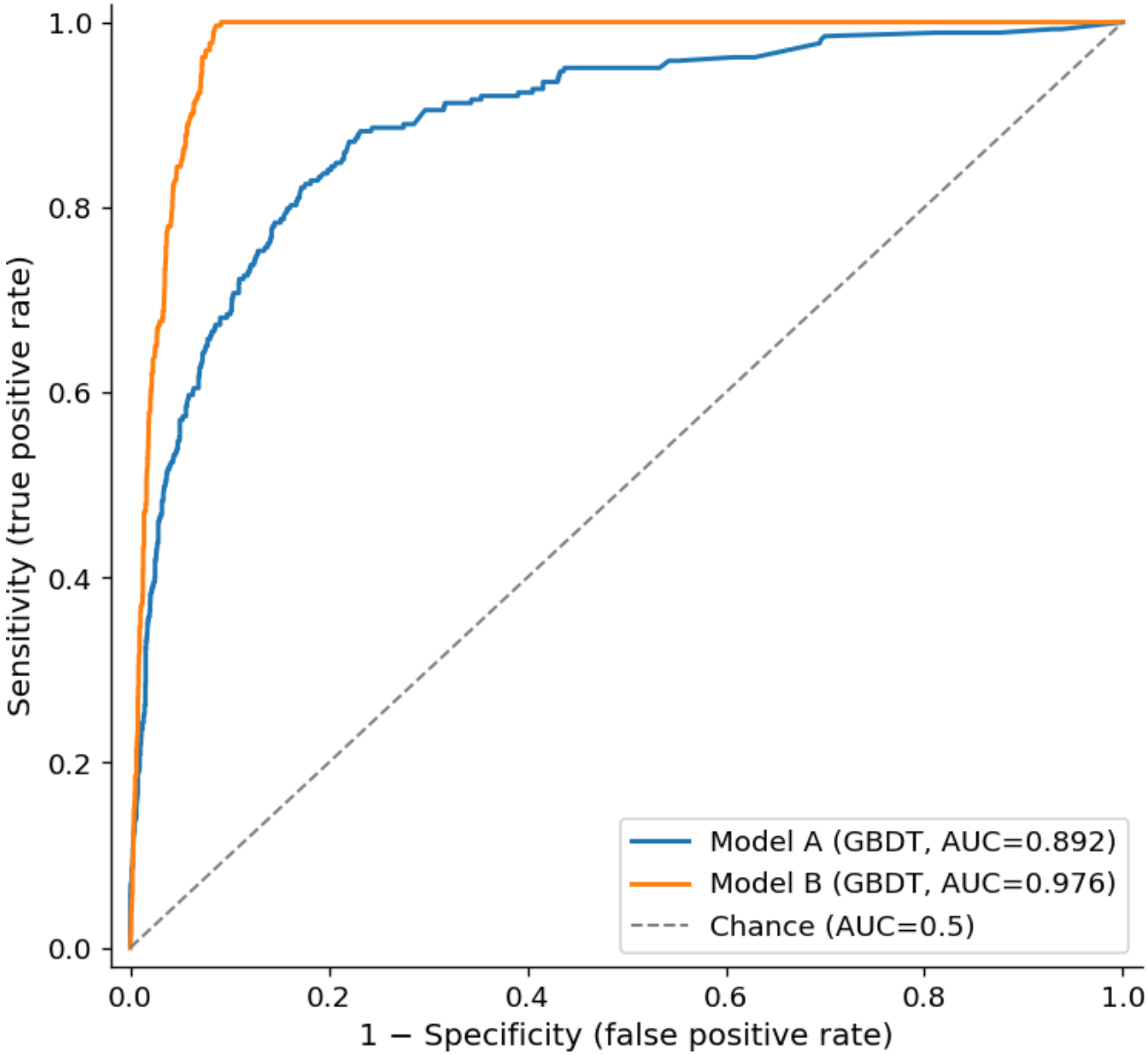
On the held-out test set (n=3024), GBDT (model A) achieved an AUC of 0.892 (95% CI 0.866 - 0.918, DeLong method; bootstrap 95% CI 0.870 - 0.912), with a PR-AUC of 0.530 (6.10 × the no-skill baseline of 0.087). At the Youden-optimal threshold (0.077), sensitivity was 0.882, specificity was 0.768, F_1-score was 0.41, and MCC was 0.41. Calibration was adequate after isotonic recalibration (Brier score=0.055, calibration slope=0.826, and intercept=−0.295); a slope below 1.0 indicates mild overfitting, with predicted probabilities slightly more extreme than observed risks. The cross-validation-to-test AUC drop of 0.021 confirms reliable cross-validated estimates and minimal generalization error. The train-test AUC gap was 0.071, modestly larger than ideal but small enough to rule out severe overfitting given the consistent CV-test gap. Model A thus met the prespecified criterion of an AUC of 0.80 or more without ADL or IADL.

For comparison, model B achieved an AUC of 0.976 (95% CI 0.962 - 0.989) using GBDT on the held-out test set, with a sensitivity of 0.996 at the Youden threshold, a train-test AUC gap of 0.018, and a Brier score of 0.036. The 0.084 AUC gap between models B and A, and the near-perfect sensitivity of model B at the Youden threshold, quantitatively confirm the near-tautological relationship between IADL and the self-reported care-need outcome that motivated the prespecified decision to designate model A as the primary model.

Figure 2 compares the receiver operating characteristic curves of the 7 algorithms on the held-out test set. Boosting-based ensembles (GBDT, XGBoost, and LightGBM) and RF showed similar discrimination, while LR, SVM, and DT achieved slightly lower AUCs.

XSL•FO
RenderX

Figure 2. Receiver operating characteristic (ROC) curves for 7 ML models on the held-out test set using model A (primary, no activities of daily living [ADL] or instrumental activities of daily living [IADL]). AUC: area under the receiver operating characteristic curve; SVM: support vector machine; GBDT: gradient-boosted decision trees.

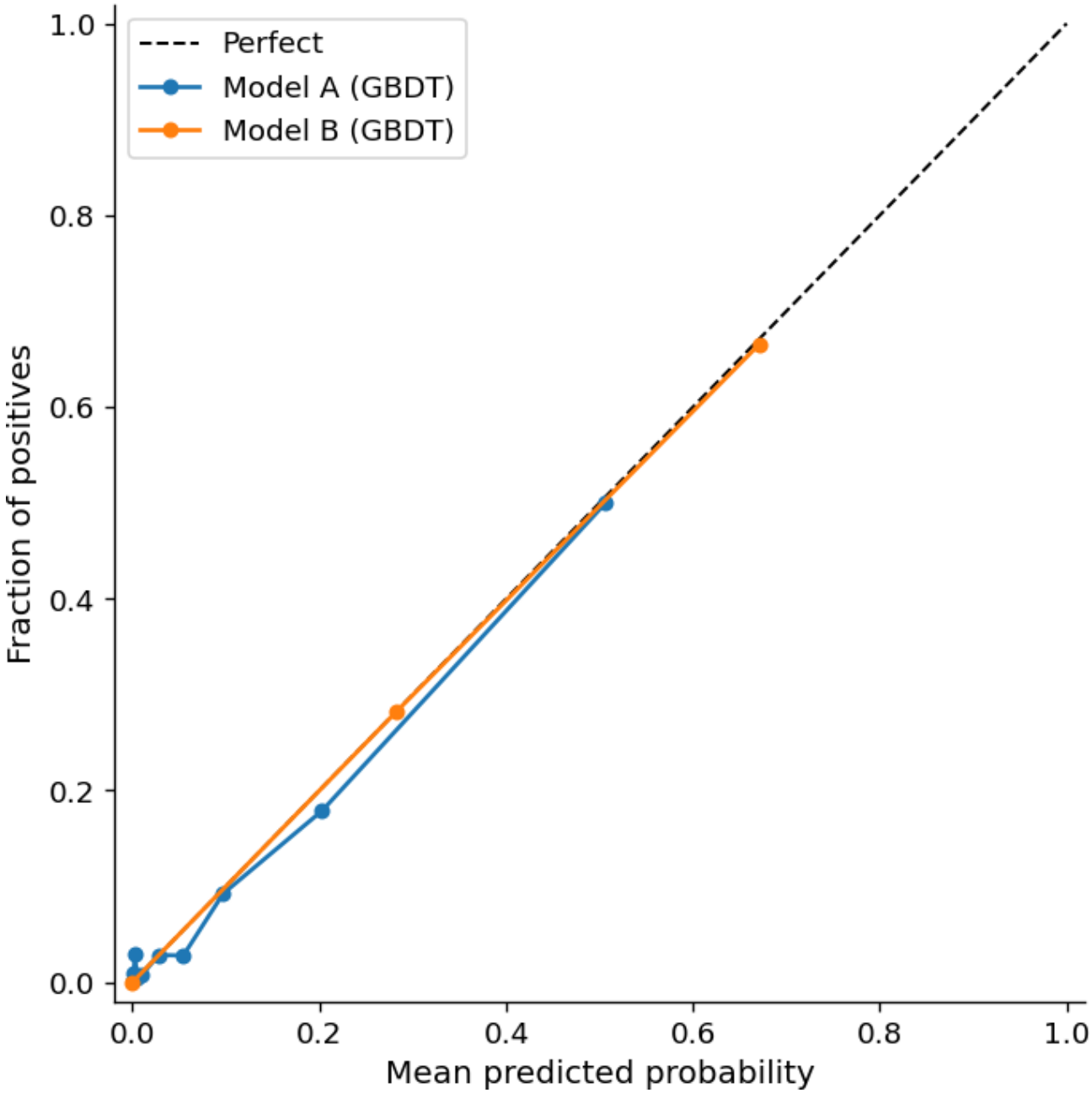


Calibration and Clinical Use

Calibration plots (Figure 3) show that both model A and model B track the perfect-calibration diagonal closely across the entire range of predicted probabilities observed in the held-out test set (0 - 0.7). Decision-curve analysis (Figure 4) demonstrates that model A provides positive net benefit relative to both "treat all" and "treat none" reference strategies across the full clinically plausible range (1% - 50%), suggesting potential use for upstream screening, pending external validation.

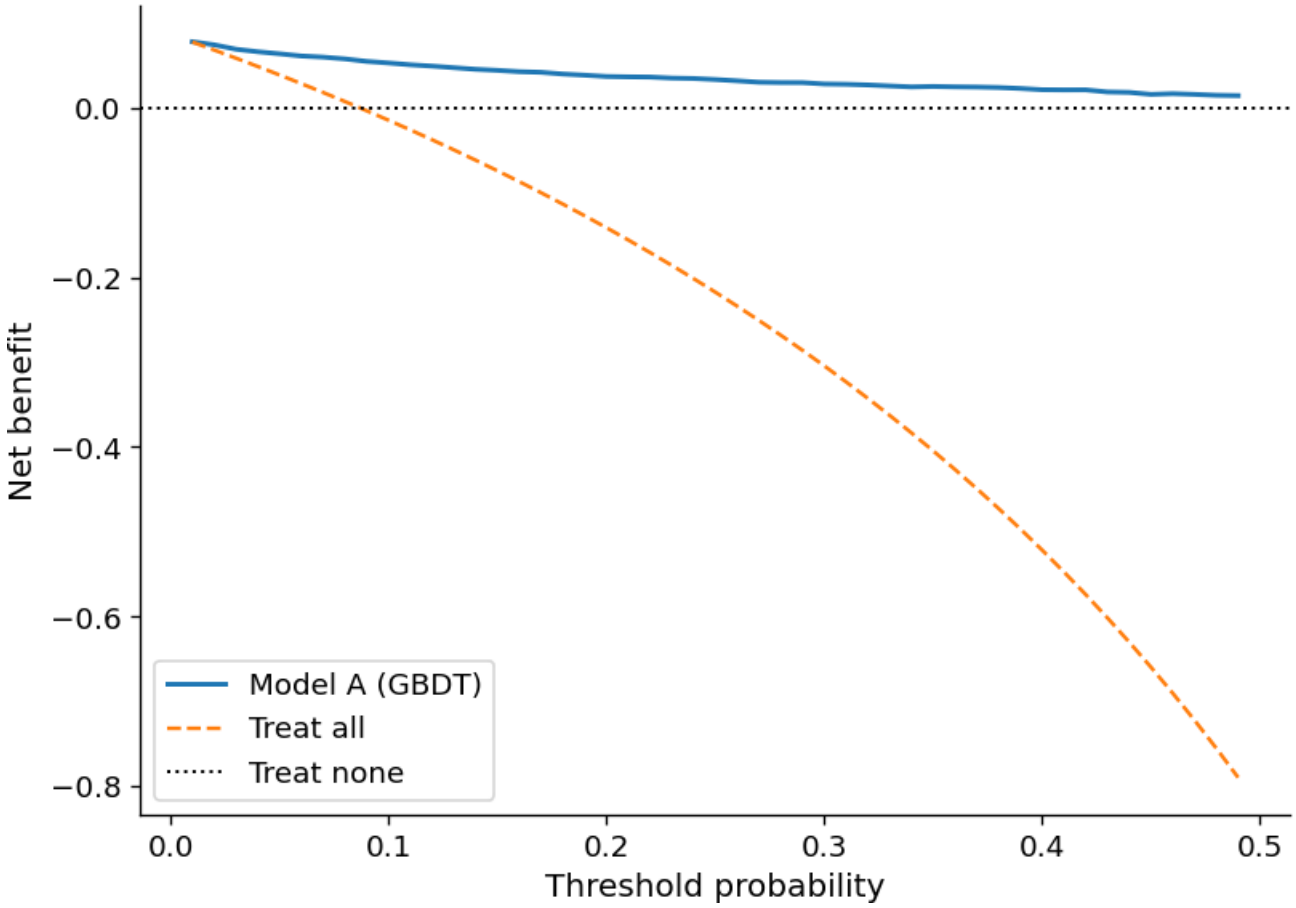
XSL•FO
RenderX

Figure 3. Calibration plot for model A (primary, gradient-boosted decision trees [GBDT]) and model B (benchmark, GBDT) on the held-out test set.



XSL•FO
RenderX

Figure 4. Decision curve analysis for model A on the held-out test set. GBDT: gradient-boosted decision trees.



Model Interpretation

Figures 5 and 6 present the results of the SHAP analysis based on the best-performing model (GBDT, model A), with high rank stability across cross-validation folds (Jaccard=0.806 for the top 10), illustrating how individual predictors contribute to model predictions of care needs. Figure 5 shows the distribution of SHAP values for each predictor across individuals, indicating both the direction and magnitude of each variable's contribution to the predicted probability of care needs. Predictors are ordered according to their overall importance. Figure 6 summarizes the absolute SHAP values, providing a global ranking of feature importance.

XSL•FO
RenderX

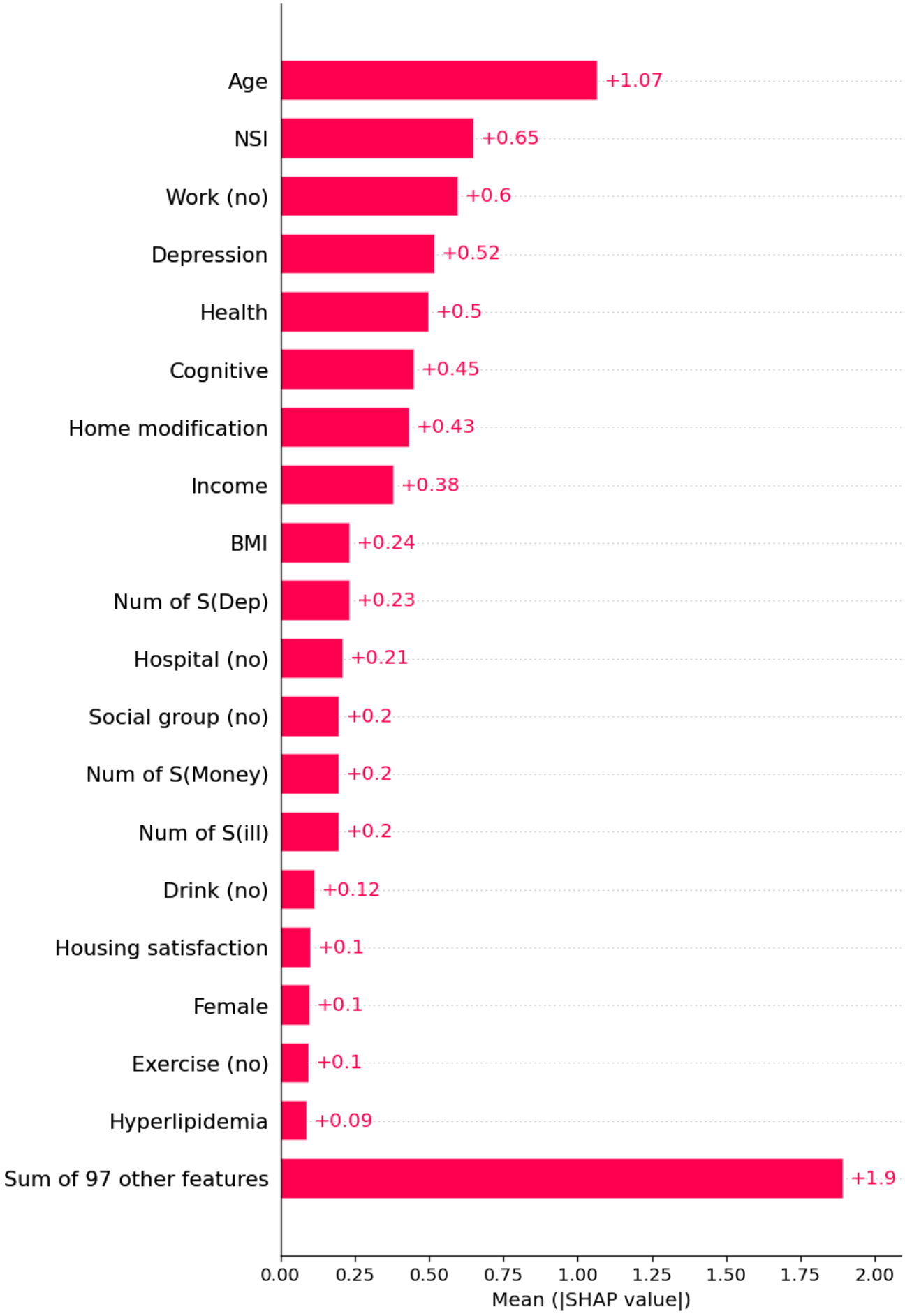
Figure 5. Shapley additive explanations (SHAP) summary plot. All samples and features are displayed, with each row representing a feature and the x-axis representing the SHAP value. Red dots represent higher feature values, while blue and purple dots represent lower feature values. NSI: Nutritional Screening Initiative; Num of S(Dep): number of supporters when feeling depressed; Num of S(ill): number of supporters when ill; Num of S(Money): number of supporters to borrow money from; Income: household income; Health: self-rated health.



XSL•FO
RenderX

Figure 6. Mean absolute Shapley additive explanations (SHAP) values. Variables are ranked by their mean absolute SHAP values; the y-axis ordering reflects each feature's global importance to the model. NSI: Nutritional Screening Initiative; Num of S(Dep): number of supporters when feeling depressed; Num of S(ill): number of supporters when ill; Num of S(Money): number of supporters to borrow money from; Income: household income; Health: self-rated health.

XSL•FO
RenderX



XSL•FO
RenderX

SHAP analyses for model A identified age (mean |SHAP| 1.012, SD 0.073; range 0.910 - 1.093) as the most influential upstream predictor, followed by nutritional risk (NSI; mean 0.686, SD 0.103; range 0.552 - 0.809), employment status (mean 0.628, SD 0.116; range 0.484 - 0.769), depressive symptoms (mean 0.530, SD 0.093; range 0.420 - 0.606), self-rated health (mean 0.515, SD 0.068; range 0.443 - 0.587), cognitive function (mean 0.474, SD 0.069; range 0.398 - 0.567), household income (mean 0.432, SD 0.070; range 0.346 - 0.524), home modification level (mean 0.364, SD 0.061; range 0.270 - 0.407), BMI (mean 0.278, SD 0.032; range 0.244 - 0.316), and number of supporters when feeling depressed (mean 0.251, SD 0.056; range 0.176 - 0.308). All fold-level intervals excluded zero, supporting the stability of the rankings.

The leading predictors span demographic (age), nutritional (NSI), behavioral (employment), mental (depressive symptoms), subjective-health, cognitive, environmental (home modification), socioeconomic (income), and social-support domains, consistent with the study's second objective, indicating that care need behaves as a multidimensional construct. SHAP distributions were smoothly graded with no evident clustering, and all directional patterns were clinically interpretable: higher age, NSI, depressive symptoms, and home modification were associated with higher predicted risk; better self-rated health, higher cognitive function, and higher income were associated with lower predicted risk.

Sensitivity Analyses

For model A, the largest individual ΔAUC drops on removal and retraining were age (−0.005), self-rated health (−0.003), and employment (−0.001). All other top-10 predictors produced |ΔAUC| of 0.002 or less, with 4 (cognitive function, home modification, BMI, and depressive symptoms) showing trivial AUC increases when removed (ΔAUC=0.000 - 0.001), reflecting redundancy within the top-10 set rather than weakness of the included variables. No single nonfunctional variable produced a catastrophic drop, indicating a genuinely multidimensional signal (Multimedia Appendix 3).

When the outcome was redefined as 2 or more IADL limitations and/or 1 or more ADL limitations (1411/10,078, 14.0% prevalence, distinct from the primary self-report 875/10,078, 8.7% prevalence), model A achieved an AUC of 0.892 (95% CI 0.871 - 0.912), identical to its performance against the original self-reported outcome. This congruence between 2 substantively different outcome operationalizations suggests that model A's discrimination is not solely an artifact of self-report bias. However, because the composite outcome itself is based on functional impairment, this analysis primarily demonstrates that the upstream predictors are associated with functional limitation and cannot fully exclude measurement bias in the self-reported outcome.

Discussion

Principal Findings

This study aimed to identify determinants of care needs among community-dwelling older adults in Korea using a comprehensive set of multidimensional variables and explainable ML techniques. The principal finding is that care needs can be identified with good discrimination (AUC 0.892, 95% CI 0.866 - 0.918), adequate calibration (Brier=0.055 and calibration slope=0.826, indicating mild optimism in predicted probabilities), and positive clinical net benefit across the 1% to 50% threshold range using multidimensional nonfunctional features alone (model A). Model A met the prespecified discrimination criterion (AUC≥0.80) without ADL or IADL. Notably, when functional status variables were excluded to identify upstream risk factors, age, nutritional risk, depressive symptoms, self-rated health, employment status, cognitive function, home modifications, and household income emerged as the leading predictors, with high rank stability across folds (Jaccard=0.806). Discrimination is preserved when the outcome is redefined objectively (AUC=0.892), suggesting that the model captures a construct shared between subjective perception and objective functional impairment, though this does not by itself rule out self-report bias in the primary outcome.

The implications of this study are as follows. First, methodologically, the findings demonstrate the strong utility of explainable ML models in gerontological research. By comparing the predictive performance of 7 ML algorithms, the results suggest that ensemble learning methods are highly effective in capturing complex, nonlinear interactions among diverse risk factors inherent in gerontological data. Furthermore, the integration of SHAP values addresses the limitations of ML and provides transparent and interpretable evidence for prioritizing interventions. This study also replicates a recurring lesson in clinical-prediction ML: unusually high AUCs in survey-based studies often reflect outcome-predictor proximity rather than true upstream predictive value. The full-featured model B achieved an AUC of 0.976, but this reflects the near-tautological inclusion of IADL alongside a self-reported care-need outcome [17,18]; reporting model A as the primary result, alongside calibration, decision-curve analysis, and TRIPOD+AI–compliant documentation, provides a more honest characterization of model utility for screening.

Second, this study verified that care needs are multidimensional constructs shaped not only by physical function but also by upstream sociodemographic and environmental determinants across the following domains: demographics; health (eg, functional status, subjective health, nutritional health, mental health, and cognitive health); health behaviors (experience of falls and health-related behaviors); usage (medical use); social domains (participation, support); and the living environment. While previous studies have been limited to medical records or functional status, this study demonstrates that care needs are shaped by a complex interplay of sociodemographic and environmental factors alongside physical function.

Policy Implications

Based on the study's findings, several policy and practical implications can be proposed to support successful aging in place among older adults.

First, it is essential to prioritize functional status for immediate service allocation. Consistent with existing long-term care eligibility criteria, functional limitations (IADL and ADL) were identified as the strongest predictors of care needs in model B.

XSL•FO
RenderX

This finding validates the current assessment system, which prioritizes physical assistance for individuals with established functional dependence.

Second, targeting hidden-risk groups for preventive interventions is a critical secondary strategy. Model A highlighted the need for a preventive pathway. Even among older adults without severe functional impairments, specific multidimensional factors are associated with elevated care needs. The analysis identified advanced age, nutritional instability, depressive symptoms, cognitive decline, and poor housing conditions as key predictors. Therefore, a secondary pathway could be considered for older adults who are physically independent but exhibit these vulnerabilities. These cross-sectional associations are hypothesis-generating and require longitudinal confirmation before informing preventive deployment.

Third, care services must extend beyond physical support to include integrated lifestyle and environmental management. Specifically, services should encompass nutritional management, psychological counseling for depression, and cognitive training programs. Additionally, the strong predictive power of housing satisfaction and home modification levels underscores the importance of the living environment in aging. Housing interventions, such as barrier-free renovations and environmental adjustments, may be integrated into care strategies; whether such interventions delay the onset of physical dependency requires prospective evaluation [19].

Limitations

This study had several limitations. First, as a cross-sectional study, it identified associations but could not establish causal relationships between risk factors and the onset of care needs. Longitudinal studies are needed to track these trajectories over time. Specifically, the model identifies older adults whose contemporaneous profile is consistent with current care needs; it has not been validated as a forward-looking incidence-prediction tool. Longitudinal validation using the panel waves of the Korea Senior Survey is a planned next step.

Second, care need was self-reported and was therefore subject to the cultural, psychological, and gendered patterns of need-perception described in the gerontology literature [20]. This study mitigated this concern through (1) sensitivity analysis using an objective composite outcome (results were substantively unchanged: AUC=0.892), and (2) stratification by depressive symptoms (with discrimination preserved across quartiles). Future work should triangulate self-reported need with caregiver-reported and clinician-assessed need.

Third, the model underwent internal validation only (cross-validation and a held-out test split from the same survey); it has not been externally or temporally validated, and the calibration slope of 0.826 indicates that predicted probabilities would likely require recalibration in new settings. Geographic and cross-national validation in independent cohorts (eg, Japanese LTCI claims; the Survey of Health, Ageing and Retirement in Europe; and the China Health and Retirement Longitudinal Study or Chinese Longitudinal Healthy Longevity Survey) are required before clinical or policy deployment. The present model is framed as an evidence-based step toward, not a substitute for, such validation.

Also, the outcome variable, defined as binary care needs, should be expanded in the context of the LTCI system (eg, predicting grades 1 - 5 or preventive levels). Future research should apply multiclass classification models to predict specific levels of care needs, thereby enabling more granular and tailored resource allocation.

Finally, hyperparameters were prespecified rather than tuned. This avoids the optimistic bias from data-driven tuning, but the relative performance across algorithms should be interpreted as a comparison under common, nonoptimized settings, rather than as each algorithm's best achievable performance.

Conclusions

This study demonstrates that explainable ML models can accurately and transparently predict care needs among older adults using multidimensional survey data. While functional impairment plays a central role, care needs are also shaped by health-related, nutritional, social, and environmental factors. By highlighting the importance of nonfunctional indicators and emphasizing interpretability, this study offers a robust framework for the identification of care needs to inform proactive, community-based aging, and long-term care policies.

Acknowledgments

The authors declare the use of generative AI (GenAI) in the writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: proofreading and editing. The GenAI tool used was Claude (Anthropic). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

The authors declared no financial support was received for this work.

Data Availability

The data used in this study were obtained from the Korea Senior Survey, managed by the Korea Institute for Health and Social Affairs [21]. Data access for analysis purposes was initiated on November 3, 2025, and the latest access date for verification was January 2, 2026.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Shapley additive explanations stability analysis with bootstrap 95% CIs.
[XLSX File, 19 KB - aging_v9i1e93371_app1.xlsx]

Multimedia Appendix 2
Operating-point performance tables at multiple decision thresholds.
[XLSX File, 18 KB - aging_v9i1e93371_app2.xlsx]

Multimedia Appendix 3
Leave-one-feature-out results for the top-10 predictors of model A.
[XLSX File, 11 KB - aging_v9i1e93371_app3.xlsx]

Checklist 1
TRIPOD+AI checklist.
[XLSX File, 20 KB - aging_v9i1e93371_app4.xlsx]

References

1. Jiang M, Li X. Explainable machine learning models predicting the risk of social isolation in older adults: a prospective cohort study. BMC Public Health 2025 May 30;25(1):1999. [doi: 10.1186/s12889-025-23108-1] [Medline: 40448115]
2. Wang X, Zhang D, Lu L, et al. Development and validation of an explainable machine learning model for predicting the risk of sleep disorders in older adults with multimorbidity: a cross-sectional study. Front Public Health 2025;13:1619406. [doi: 10.3389/fpubh.2025.1619406]
3. Lee SM, Ha HJ. Prediction of depression in the elderly using automated machine learning. J Korea Inf Technol Soc 2025 Apr 30;23(4):1-10. [doi: 10.14801/jkiit.2025.23.4.01]
4. Kim JH, Kim HJ, Kim SI, Park BY. Study on creating a depression group discrimination model for the elderly using machine learning and researching factors related to depression. Health Welf 2024 Sep 30;26(3):7-32. [doi: 10.23948/kshw.2024.09.30.3.07]
5. Suh M, Jung H, Kim J. Development of a fall prediction model for community-dwelling older adults in South Korea using machine learning: a secondary data analysis. J Korean Biol Nurs Sci 2024 Nov;26(4):288-299. [doi: 10.7586/jkbns.24.024]
6. Ren H, Zheng Y, Li C, et al. Using machine learning to predict cognitive decline in older adults from the Chinese Longitudinal Healthy Longevity Survey: model development and validation study. JMIR Aging 2025 Apr 30;8:e67437. [doi: 10.2196/67437] [Medline: 40305830]
7. Isaradech N, Sirikul W, Buawangpong N, Siviroj P, Kitro A. Machine learning models for frailty classification of older adults in Northern Thailand: model development and validation study. JMIR Aging 2025 Apr 2;8:e62942. [doi: 10.2196/62942] [Medline: 40262171]
8. Fukunishi H, Kobayashi Y. Care-needs level prediction for elderly long-term care using insurance claims data. Inform Med Unlocked 2023;41:101321. [doi: 10.1016/j.imu.2023.101321]
9. Song MK, Park Y, Han EJ. Development of prediction model identifying high-risk older persons in need of long-term care. Korean J Appl Stat 2022 Aug 31;35(4):457-468. [doi: 10.5351/KJAS.2022.35.4.457]
10. Lu Y, Sato K, Nagai M, Miyatake H, Kondo K, Kondo N. Machine learning–based prediction of functional disability: a cohort study of Japanese older adults in 2013–2019. J Gen Intern Med 2023 Aug;38(11):2486-2493. [doi: 10.1007/s11606-023-08215-2] [Medline: 37127751]
11. Han SY. A study on multidimensional predictors of frailty in community-dwelling older adults: a decision tree-based approach. Health Soc Welf Rev 2025;45(2):4-24. [doi: 10.15709/hswr.2025.45.2.4]
12. Wiles JL, Leibing A, Guberman N, Reeve J, Allen RES. The meaning of “aging in place” to older people. Gerontologist 2012 Jun;52(3):357-366. [doi: 10.1093/geront/gnr098] [Medline: 21983126]
13. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
14. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321-357. [doi: 10.1613/jair.953]
15. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(6):565-574. [doi: 10.1177/0272989X06295361] [Medline: 17099194]

XSL•FO
RenderX

16. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; Dec 4-9, 2017. [doi: 10.5555/3295222.3295230]
17. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019 May;1(5):206-215. [doi: 10.1038/s42256-019-0048-x] [Medline: 35603010]
18. Ramadan B, Liu MC, Burkhart MC, Parker WF, Beaulieu-Jones BK. Diagnostic codes in AI prediction models and label leakage of same-admission clinical outcomes. JAMA Netw Open 2025 Dec 1;8(12):e2550454. [doi: 10.1001/jamanetworkopen.2025.50454] [Medline: 41632159]
19. Cha SM. A systematic review of home modifications for aging in place in older adults. Healthcare (Basel) 2025 Mar 27;13(7):752. [doi: 10.3390/healthcare13070752] [Medline: 40218050]
20. Zhang J, Zhang Y, Bennett MR. Spousal characteristics and unmet care needs: a longitudinal national study of adults aged 50 and over in England. Soc Sci Med 2025 Jan;365:117530. [doi: 10.1016/j.socscimed.2024.117530] [Medline: 39673874]
21. Korea Institute for Health and Social Affairs (KIHASA). URL: https://data.kihasa.re.kr [accessed 2025-11-03]

Abbreviations

ADL: activities of daily living
AUC: area under the receiver operating characteristic curve
CES-D-SF: Center for Epidemiologic Studies Depression Scale–Short Form
DT: decision tree
GBDT: gradient-boosted decision trees
IADL: instrumental activities of daily living
LightGBM: light gradient boosting machine
LR: logistic regression
LTCI: Long-Term Care Insurance
MCC: Matthews correlation coefficient
ML: machine learning
NSI: Nutritional Screening Initiative
PR-AUC: precision-recall area under the curve
RF: random forest
SHAP: Shapley additive explanations
SVM: support vector machine
TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis+AI
VIF: variance inflation factor
XGBoost: extreme gradient boosting

Edited by Y Jiang; submitted 12.Feb.2026; peer-reviewed by AHH Khan, SK Ghosh, SW Park, W Xu; revised version received 18.Jun.2026; accepted 14.Jul.2026; published 31.Jul.2026.

Please cite as:
Shin H
Predicting Care Needs in Community-Dwelling Older Adults Using Explainable Machine Learning and a Multidimensional Approach Integrating Health, Social, and Environmental Factors: Cross-Sectional Study
JMIR Aging 2026;9:e93371
URL: https://aging.jmir.org/2026/1/e93371
doi:10.2196/93371

© Hyeri Shin. Originally published in JMIR Aging (https://aging.jmir.org), 31.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Daily Monitoring of Behavioral and Psychological Symptoms of Dementia in Residential Care: Mixed Methods Pilot Study

Johannes Malm[1,2], MSc; Therése Bielsten[1], PhD; Charlotta Nilsen[1,3,4], PhD; Elzana Odzakovic[5], PhD; Ingemar Kåreholt[1,6], PhD

[1]Institute of Gerontology, School of Health and Welfare, Jönköping University, Barnarpsgatan 39, Box 1026, Jönköping, Sweden

[2]Department of Health, Blekinge Institute of Technology, Karlskrona, Sweden

[3]Aging Research Center, Karolinska Institutet, Stockholm, Sweden

[4]Department of Psychology, The Stress Research Institute, Stockholm University, Stockholm, Sweden

[5]Department of Nursing, School of Health and Welfare, Jönköping University, Jönköping, Sweden

[6]Division of Clinical Geriatrics, Department of Neurobiology, Care Sciences and Society (NVS), Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Johannes Malm, MSc
Institute of Gerontology, School of Health and Welfare, Jönköping University, Barnarpsgatan 39, Box 1026, Jönköping, Sweden

Abstract

Background: Many clinically relevant aspects of health are expressed through behavior, affect, and interaction, and therefore rely on observation rather than instrumental measurement; yet such information is often documented infrequently in routine care. Dementia, and particularly behavioral and psychological symptoms of dementia (BPSD), is a relevant case because symptoms fluctuate across time and context. Little is known about how structured daily symptom assessment can be integrated into routine dementia care practice.

Objective: This pilot study aimed to investigate whether structured daily observation of BPSD, supported by digital documentation, can be incorporated into everyday residential dementia care.

Methods: A convergent mixed methods design was applied across 4 nursing home units in southern Sweden. Feasibility was operationalized through 4 focus areas (implementation, practicality, acceptability, and integration), with adherence to the registration routine as the primary feasibility outcome. Quantitative data comprised structured registrations collected daily by care staff over 90 days for 8 residents with BPSD, analyzed descriptively, and interpreted using selected NASSS (nonadoption, abandonment, scale-up, spread, and sustainability) framework domains to contextualize variation in symptom patterns, adherence, and registration practices. Semistructured dyadic interviews were analyzed using deductive content analysis guided by the same domains, with integration occurring at the analytical and interpretive stages.

Results: A total of 21,993 item-level registrations were completed: 19,123 NPI-NH (Neuropsychiatric Inventory – Nursing Home version) symptom ratings, recorded across 1623 shift registrations (1 per resident per shift), and 2870 individually selected variables. Overall adherence to the registration routine averaged 75% across shifts but varied considerably between residents (range 33% - 90%) and between staff members. Mean time per shift registration decreased from 77.1 seconds in the initial 30-day period to 46.2 seconds in the final 30-day period (P=.008). Symptom prevalence varied widely across NPI domains (7.5% to 38.1% of observation days), and daily symptom counts varied both between and within residents. The routine was experienced as technically straightforward and easy to incorporate into daily work. At the same time, challenges emerged, including time constraints, inconsistent participation, varying levels of interpretive confidence, and difficulties distinguishing among overlapping symptoms. Organizational routines and team stability were central to sustaining staff engagement.

Conclusions: The high number of completed registrations indicates that daily observations can be integrated into daily dementia care. The participants perceived it as meaningful and useful. Feasibility was primarily influenced by organizational and interpretive factors rather than technical constraints. Further research should explore strategies to support shared routines and sustain engagement over time.

(JMIR Aging 2026;9:e98024) doi:10.2196/98024

KEYWORDS

behavioral and psychological symptoms of dementia; BPSD; dementia care; digital health; high-frequency observational data; mixed methods research; Neuropsychiatric Inventory; NPI; pilot study; residential care

XSL•FO
RenderX

Introduction

Care Under Pressure and the Promise of Digitalization

Health care systems in many countries, including Sweden, are under pressure from demographic change: an increasing number of older adults with chronic conditions and complex care needs are met by limited staff and resources [1-4]. At the same time, care is expected to be evidence-based and person-centered [5,6], a standard that is difficult to meet in everyday practice characterized by high workload, time pressure, and gaps in clinical information sharing [7,8]. Digitalization is often proposed as one means to address this gap, through documentation, data collection, and information sharing that can strengthen the informational basis for clinical decision-making and enable more detailed, frequent data on patients' health status [9-11].

Digitalization has expanded the possibility of collecting high-frequency health data that capture patterns and temporal dynamics often missed by episodic assessment and documentation. In several areas of medicine, increasing the frequency of data collection has yielded more detailed insight into health over time, for example, through continuous glucose monitoring in diabetes care [12] and data-intensive cancer diagnostics [13,14]. These advances, however, rely on data generated by sensors or imaging. Many clinically important phenomena, including behavioral and psychological symptoms of dementia (BPSD), are instead captured through staff observation and subjective interpretation, where structured, frequent data are far less available and data-driven methods correspondingly less established. Realizing comparable benefits in such domains therefore depends first on whether observational data can be captured systematically and frequently enough in everyday practice.

The Documentation Gap in Observational Care

In such observation-based aspects of care, professional judgment remains central to detection and interpretation [15], and both the assessment routine itself and the data it generates can provide important insight into a person's situation, needs, and well-being [16]. However, these assessments are typically documented in episodic, narrative formats that lack the standardization and temporal resolution needed to detect gradual or subtle changes over time [17]. This stands in contrast to instrumental measurements, where increasingly continuous and automated data collection enables fine-grained tracking of change [18].

This documentation gap is particularly evident in long-term care settings, where individuals rely on staff to observe and interpret their needs. Through continuous interaction, staff accumulate substantial tacit knowledge about behavioral patterns, functional fluctuations, and early signs of change [19]. While such knowledge is often clinically meaningful, it is unevenly documented and therefore vulnerable to variation [20]. Differences in staff experience and training [21], challenges related to documenting unstructured information [22], and technical constraints in digital systems [23] may all contribute to variation in how observations are recorded. Consequently, clinically relevant information may remain undocumented or circulate primarily through informal communication. Structured, high-frequency observational routines may offer one way to stabilize and render this tacit knowledge more visible and actionable. In dementia care, key data types that benefit from systematic capture include the presence and severity of behavioral and psychological symptoms, daily functional fluctuations, and context cues such as activity, sleep, and interpersonal interactions. Documenting these systematically rather than informally is a prerequisite for using them in care planning.

BPSD as a Case of Rapid Symptom Fluctuation

Dementia, and especially BPSD, provides a particularly relevant case for examining this documentation gap, as symptoms fluctuate episodically across time, environments, and internal states. Dementia is a progressive neurocognitive condition affecting a large and growing proportion of older adults worldwide [24,25], and up to 90% of people living with dementia experience BPSD over the course of the condition [26,27]. Symptoms include agitation, anxiety, hallucinations, delusions, irritability, apathy, and sleep disturbances. BPSD substantially affects quality of life, caregiver burden, rates of institutionalization, and the need for tailored interventions [28,29].

Moreover, several BPSD symptoms share overlapping behavioral expressions, making clinical differentiation challenging even for experienced staff [30]. This interpretive complexity adds a further dimension to the documentation challenge, as structured assessment depends not only on frequency of observation but also on the clarity with which symptoms can be distinguished and categorized.

Because BPSD emerges from dynamic interactions among neurocognitive decline, personal needs, and environmental context, it requires close and repeated observation. Yet current assessment practices rely largely on periodic evaluations, while day-to-day fluctuations are often documented informally or recalled from memory [31]. In Sweden, the national BPSD Registry provides a structured framework for assessment, but registrations typically occur at intervals of 6 months or more [32]. This structural mismatch between rapidly fluctuating symptoms and episodic assessment systems [33] may leave clinically meaningful variation undetected, limiting opportunities for timely and person-centered intervention.

One possible way to capture this variation is through structured high-frequency observational routines carried out as part of everyday care practice. This raises interrelated questions: whether such structured observation is feasible to perform at all and whether it can be integrated as a routine practice in residential dementia care. The latter is shaped by interrelated technical, interpretive, and organizational conditions. Before larger-scale adoption can be considered, it is therefore necessary to examine whether structured daily observational routines are feasible under real-world conditions [34].

Objective

The overall aim of this pilot study was to investigate whether structured daily observation of BPSD, supported by digital documentation, can be incorporated into everyday residential dementia care. Two research questions were addressed: (RQ1)

XSL•FO
RenderX

What is the feasibility of integrating structured daily observations into residential dementia care? (RQ2) What are the day-to-day and between-resident variations in BPSD registrations? Feasibility was examined with adherence to the registration routine as the primary feasibility outcome, and practicality, acceptability, and integration into existing workflows were examined descriptively.

Methods

Study Design

A convergent mixed methods design was used to examine whether structured daily observation of BPSD, supported by digital documentation, can be incorporated into everyday dementia care. Quantitative and qualitative data were collected concurrently, analyzed separately, and subsequently brought together as part of a structured data integration process, in accordance with established principles for convergent mixed methods research [35].

A convergent design was chosen because the 2 strands addressed the same research question (the feasibility of integrating structured daily observational routines into care) from complementary perspectives, with neither strand subordinate to the other. The quantitative data on adherence, registration patterns, and shift-registration time describe what was accomplished; the qualitative data on staff experiences describe how and under what conditions. A sequential design was not appropriate because neither strand needed to inform the other prospectively, and an embedded design was not appropriate because neither was secondary. Concurrent data collection during the same 90-day study window further supported a convergent approach.

Feasibility was conceptualized using the focus areas described by Bowen et al [36]; four areas were targeted in this pilot: implementation, practicality, acceptability, and integration. The primary feasibility outcome was adherence to the registration routine, operationalized as the proportion of expected item-level registrations completed. The remaining 3 dimensions were explored descriptively without prespecified thresholds. The other 4 Bowen focus areas (demand, adaptation, expansion, and limited-efficacy testing) were considered beyond the scope of this pilot.

In this study, two related but distinct concepts are used in their established implementation-science sense: *implementation* refers to the process of introducing the routine into care (operationalized in this study through adherence, the primary feasibility outcome); *integration* refers to the extent to which the routine fits within existing workflows.

Ethical Considerations

This study was approved by the Swedish Ethical Review Authority (Dnr 2022-04509-01) under two specific conditions reflecting the cognitive vulnerability of the resident population. First, because residents could not reasonably be expected to assimilate written study information themselves, written information was directed to a next of kin, legal guardian, or trustee, who provided consent on the resident's behalf and was instructed to consult the resident and convey the resident's opportunity to object. Because residents could not reliably provide active assent, the procedure relied on this proxy consent together with the resident's ongoing opportunity to object, a nonobjection rather than an active-assent model. Second, the structured daily observations imposed no incremental procedural burden: registrations were integrated into the documentation that care staff would otherwise perform informally during routine care, requiring no additional interaction with the resident. To minimize the handling of personal data, signed consent forms were collected and stored locally by participating unit managers and were never transferred to the research team; the team verified that consent had been completed before each resident was included.

Health care professionals were informed of the study by their unit managers but registered their interest in interview participation directly with the research team via an online form, after which a subset were randomly approached; oral informed consent was obtained prior to each interview. Use of the Daily-BPSD (Jönköping University) application required personal staff login to preserve accountability, in line with standard clinical documentation practice; upon export to the research dataset, staff identifiers were replaced with study codes (eg, "Staff Participant 3"), and the anonymized registration data were stored on secure servers at Jönköping University. Audio recordings and verbatim transcripts of interviews were stored on secure university servers, with the key linking interview codes to identifiable information held separately at Jönköping University. The processing of personal data rested on a public-interest legal basis under the EU General Data Protection Regulation, together with the study's ethical approval.

Settings and Participants

The study was conducted in dementia-specific residential care units within municipal long-term care facilities in southern Sweden. The 4 participating units were located in 2 municipalities and distributed across both urban and smaller community settings, with all 4 operated by their respective municipalities. According to Swedish national guidelines for dementia care, such units are designed to provide round-the-clock care for persons with dementia who require extensive support in daily activities and continuous supervision. Care is typically delivered by assistant nurses, with registered nurses available on a scheduled basis. Staff typically worked in rotating shifts, with day shifts (approximately 7 AM to 3 PM), evening shifts (approximately 3 PM to 9 PM), and night shifts (approximately 9 PM to 7 AM). Dementia-specific units are commonly organized into small-scale living environments, or wards, intended to support person-centered care [37]. Although a unit may comprise several wards, the study residents in each unit belonged to a single ward with a capacity of 8 - 10 residents.

Residents were eligible if they lived in a designated dementia unit and had a medically confirmed diagnosis of dementia. No restrictions were applied regarding age or gender. All included residents exhibited BPSD, although no criteria were applied regarding specific symptoms or severity. Because many residents had significant cognitive impairment that limited their

XSL•FO
RenderX

ability to provide informed consent, written proxy consent was obtained from legal representatives or next of kin.

All regular care staff working on the participating units were expected to contribute to the daily observational registrations during the study period as part of their regular work routines. Staff were eligible for qualitative interviews if they had been actively involved in the daily care of included residents.

Unit managers facilitated recruitment by informing staff about the study and coordinating contact with residents' legal representatives or next of kin to obtain written proxy consent. Recruitment followed a convenience sampling approach appropriate for feasibility-oriented research. The sample of 8 residents reflected the practical capacity of the participating units and the research team to support intensive daily observation over the 90-day pilot. Residents were included from all 4 participating units so that the pilot captured daily documentation practices across each care context. Participation in the qualitative component was voluntary, and 10 assistant nurses consented to participate in interviews.

Quantitative Data Collection

The Daily-BPSD Tool

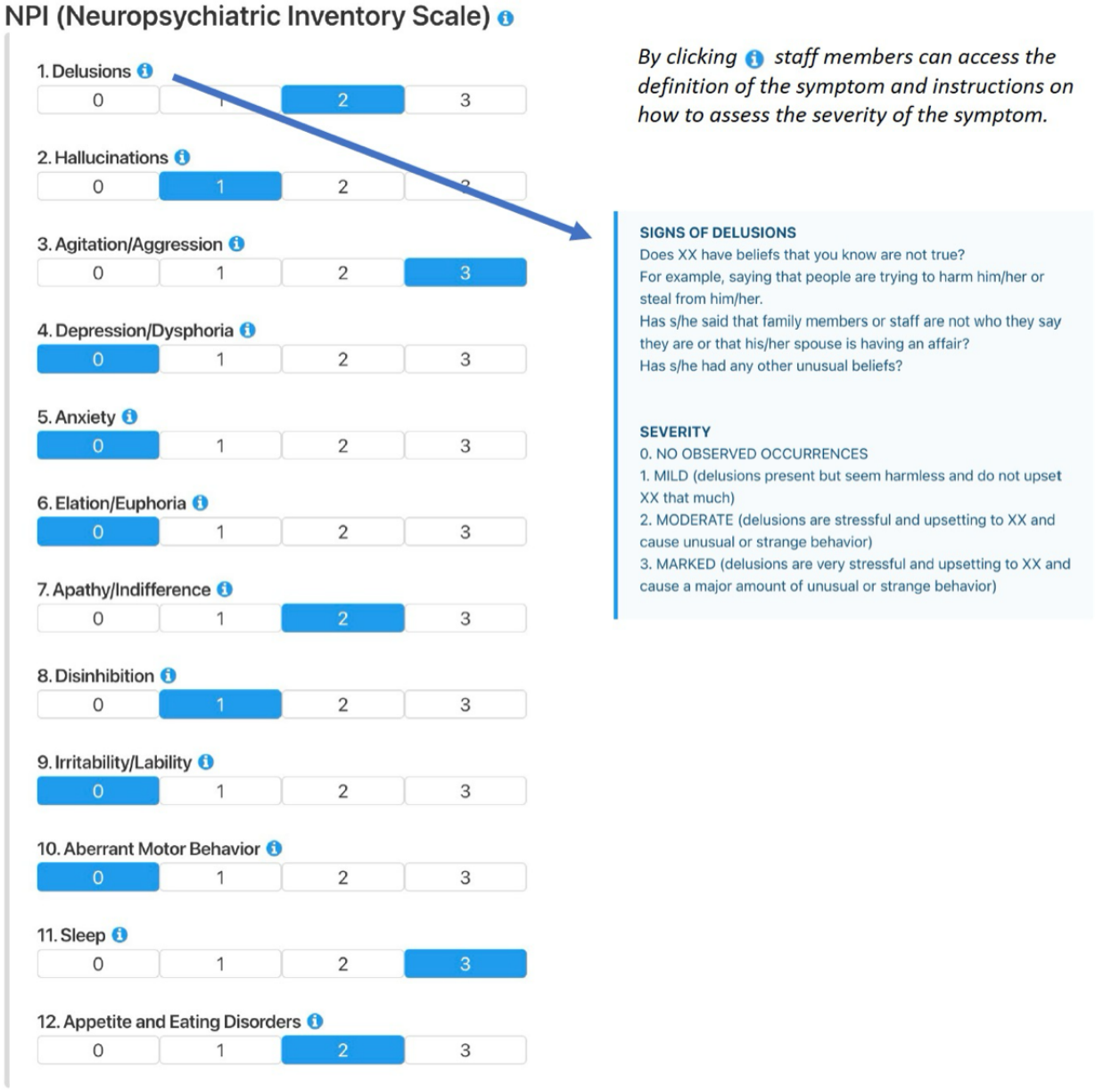
Daily-BPSD is a web-based application designed to support structured daily observation of BPSD as part of everyday care routines. The tool enables staff to register symptom severity up to 3 times per day using smartphones, tablets, or computers.

The interface is based on the 12 domains of the Neuropsychiatric Inventory (NPI) [38], with domain definitions aligned with the Neuropsychiatric Inventory – Nursing Home version (NPI-NH) [39]. Each domain includes a 0 - 3 severity scale, supported by clickable definitions and examples, to promote consistent interpretation among staff (see Figure 1). In addition to the predefined NPI domains, the tool allows optional individualized variables to capture specific care needs or context-sensitive interventions.

Daily-BPSD automatically generates graphical summaries of symptom trends over time directly within the application as new observations are registered, supporting reflection and team communication. The tool was developed in collaboration with care staff and relatives and iteratively adapted to fit existing care workflows, informed by previous findings [31].

XSL•FO
RenderX

Figure 1. Staff interface for registering Neuropsychiatric Inventory (NPI) symptom severity in Daily-BPSD, including clickable definitions and standardized rating guidance. Screenshot from the Daily-BPSD website (2025 Jönköping University, Daily-BPSD project; reproduced with permission).



Quantitative Data Description

Prior to the study period, staff members in participating units received a 2-hour introduction to the Daily-BPSD application, conducted by members of the research team. The introduction covered the rationale for daily symptom registration, navigation of the registration interface, interpretation of NPI severity definitions via the in-app information texts (see Figure 1), and account login.

Using this tool, structured daily observations were recorded for 8 residents over a 90-day period. The system automatically time-stamped each entry, allowing analysis of adherence patterns and temporal distribution of observations. Data were exported in anonymized form and contained no personal identifiers beyond each resident's study code. The quantitative dataset therefore captured both the volume and distribution of observational activity as well as symptom patterns across the study period.

Throughout this study, a *registration* (or *item-level registration*) refers to a single severity rating recorded for one item (either an NPI-NH symptom domain or an individually selected variable) for one resident at one shift; reported registration totals therefore reflect the number of item-level rating events. A *shift registration* refers to one completed NPI form for one resident at one shift (up to 3 per resident per day, one per shift); the 19,123 NPI item-level registrations were recorded across 1623 completed shift registrations. The 2870 individually selected variable registrations are item-level events not tied to a shift form. Registration time is reported per shift registration, not per item-level registration.

XSL•FO
RenderX

Qualitative Data Collection

Qualitative data were collected through semistructured dyadic interviews with 10 assistant nurses conducted in 5 pairs. One pair was recruited from each of the 4 participating units, with one unit contributing an additional pair to include night shift staff. Dyadic interviews were used to encourage interaction between participants and to generate richer accounts of shared experiences within their common care context, allowing participants to validate, contrast, and elaborate on each other's accounts in ways that solo interviews typically do not [40,41]. Participants were eligible if they were regular staff members working on the participating units during the quantitative data collection period and had used the Daily-BPSD tool for daily symptom observation. These 10 assistant nurses were drawn from the care staff who performed the daily observations and contributed to the study by sharing their experiences in the interviews analyzed below. The semistructured interview guide used during the interviews is provided in Multimedia Appendix 1.

The interviews were held at the end of the observation period to allow participants to reflect on their experiences throughout the monitoring period. The interview guide covered topics related to daily observational routines, how behavioral and psychological symptoms were interpreted, factors affecting symptom judgment, and contextual conditions influencing feasibility, such as workload, team coordination, and local documentation practices. Questions were informed by earlier developmental work [31] and by selected domains of the NASSS (nonadoption, abandonment, scale-up, spread, and sustainability) framework [42], which also guided the subsequent deductive analysis.

All interviews were conducted by members of the research team in the participating residential care units, audio recorded with consent, transcribed verbatim, and anonymized. Field notes were taken to capture situational aspects relevant for later analysis.

Data Analysis

Analytic Framework

Whereas the feasibility outcomes specified above identified what was assessed in this study, the NASSS framework [42] was applied to interpret why feasibility-relevant patterns emerged. NASSS is widely used to examine the implementation, scale-up, and sustainability of health technologies in complex care systems [43-45]. Based on the study's aim and familiarization with the empirical material, 5 NASSS domains were identified as analytically relevant: condition, technology, value proposition, adopters, and organization. To facilitate integration of qualitative findings with the quantitative data on adherence and symptom patterns, qualitative findings are reported under the NASSS domains used in the deductive analysis rather than independent categories.

The *organization* domain was included to account for unit-level variation in workflow, leadership support, staffing patterns, and local routines, all of which shaped the feasibility of integrating structured daily observational practices into everyday care.

The *condition* domain was particularly relevant because BPSD are fluctuating, context-dependent, and often unpredictable qualities that fundamentally shape what structured daily observational practices must capture.

The *technology* domain was used to examine the work required to conduct structured daily observations, including usability aspects of the digital interface and how the routine aligned with existing documentation practices.

The *value proposition* domain captured staff perceptions of the usefulness of frequent documentation for reflection, communication, care planning, and understanding residents' needs.

The *adopters* domain addressed how staff members' interpretive skills, motivation, confidence, and perceived workload influenced both the consistency with which observations were conducted and the extent to which the resulting data were used in practice.

These domains informed the deductive coding matrix in the qualitative analysis and provided a structured lens for integrating quantitative and qualitative findings across the convergent mixed methods design. While the NASSS domains provided the primary analytical structure, the analysis remained open to relevant insights not fully captured by the predefined domains. The remaining NASSS domains (wider system and embedding and adaptation over time) were not included because they address system-level influences and long-term adaptation processes that were beyond the scope of this study.

Quantitative Analysis

Quantitative data were analyzed descriptively to examine adherence to daily observational routines, the frequency and distribution of documented symptoms, and temporal variation across the observation period. For each resident, the total number of item-level registrations, the proportion of days with completed entries, and the number of contributing staff members were calculated. Symptom frequencies were summarized for each BPSD domain, and the distribution of daily symptom counts was examined to assess variation across residents. Time-stamped entries enabled analysis of registration frequency across shifts and adherence patterns across units. Because the study aimed to assess feasibility rather than effectiveness, inferential testing was limited to examining changes in the time taken to complete a shift registration. For each resident, shift-registration times (one completed shift-form, measured from form opening to submission) during the first (days 1 - 30) and last (days 61 - 90) 30-day intervals were compared using the Welch independent *t* test, which is robust when the variances are not equal [46]. Durations exceeding 5 minutes were treated as outliers and excluded (43 of 1623 shift registrations, 2.6%), as these likely reflected staff interruptions or forgotten submissions. All quantitative analyses were conducted in Python (version 3.12) using standard scientific computing libraries.

Qualitative Analysis

Qualitative data from dyadic interviews were analyzed using deductive qualitative content analysis following the approach described by Elo and Kyngäs [47]. The analysis proceeded in

XSL•FO
RenderX

3 stages: preparation, organizing, and reporting. In the preparation phase, transcripts were read repeatedly to obtain a comprehensive understanding of staff experiences. In the organizing phase, meaning units related to staff experiences of conducting daily symptom observations were identified and organized into a structured coding matrix informed by selected domains of the NASSS framework (see Multimedia Appendix 2 for an example).

Codes were iteratively compared, refined, and grouped into categories through discussions within the research team. In the reporting phase, these categories were organized within the corresponding NASSS domains, capturing staff perspectives on the feasibility, challenges, and perceived value of conducting daily symptom observations. Microsoft Excel (Microsoft 365 version) and Python (version 3.12) were used for data management, coding, and analysis. Coding was performed by JM. Each meaning unit was assigned a descriptive code that was then mapped to a generic category and to its corresponding NASSS domain; Multimedia Appendix 2 provides a worked example of this quote-to-code-to-category-to-domain chain. EO and TB reviewed a subset of coded segments and provided input on categorization, and points of uncertainty were discussed within the team until a shared interpretation was reached. To support the trustworthiness of the analysis, coding decisions were documented in the structured coding matrix and revisited iteratively in team discussions. We did not compute intercoder agreement statistics, as such measures are of limited interpretive value for deductive content analysis with categories prespecified from a theoretical framework.

Integration of Data

Following separate analyses of the 2 strands, integration was conducted in 3 procedural steps, following the joint-display approach described by Fetters et al [35]. First, findings from each strand were summarized within the 4 feasibility dimensions used in this study (implementation, practicality, acceptability, and integration). Second, quantitative and qualitative evidence pertaining to the same dimension were juxtaposed to identify points of convergence (where strands reinforced each other), complementarity (where strands addressed different facets of the same dimension), and divergence (where strands offered conflicting signals) [48]. Third, integrated meta-inferences were drawn within each NASSS domain to interpret why the observed feasibility patterns emerged, drawing jointly on registration data and staff accounts. A joint display summarizing the integrated evidence per feasibility dimension is provided in Multimedia Appendix 3.

Results

Participant Characteristics

Eight residents living with dementia and exhibiting symptoms of BPSD were included in the pilot study across 4 residential care units. Daily observations were conducted over a 90-day period using the Daily-BPSD tool. In total, 44 staff members contributed registrations to the system.

Table 1 presents residents' age and gender, along with the number of staff members involved in observations for each resident. During the study period, care staff completed 1623 shift registrations (one per resident per shift; 75% of the 2160 possible resident-shifts, that is, 8 residents × 90 days × 3 shifts), comprising 19,123 NPI item-level ratings, together with 2870 individually selected variable registrations (21,993 item-level registrations in total; see Multimedia Appendix 4 for the full list).

Table . Characteristics of people living with BPSD[a] residing in the participating residential care units, as well as the number of care staff using the tool during the data collection period.

Characteristics	Values
Residents, n	8
Age (years), mean (range)	84 (75 - 92)
Gender, n (%)	
Female	5 (63)
Male	3 (37)
Care units, n	4
Unique staff members contributing registrations, n	44
Staff members per resident, median (range)	12 (7 - 15)
Total registrations, n	21,993
NPI[b] domain registrations, n	19,123
Individually selected variable registrations, n	2870

[a]BPSD: behavioral and psychological symptoms of dementia.

[b]NPI: Neuropsychiatric Inventory.

The deductive content analysis yielded 1137 coded segments across the 5 dyadic interviews. As shown in Table 2, the coding was distributed across all 5 NASSS domains, with the *adopters* and *organization* domains most prominently represented, followed by *technology*, *value proposition*, and the *condition*. The following sections present findings within each domain.

XSL•FO
RenderX

Table . Overview of NASSS[a] framework coding across 5 interviews (N=1137 coded segments).

NASSS domain	Segments, n (%)	Generic categories
The Organization	319 (28.1)	Registration routine (n=130), workflow (n=118), team discussion (n=34), decision-making process (n=19), communication (n=9), scalability (n=7), the BPSD[b] registry (n=2)
The Condition	75 (6.6)	Symptom presentation (n=41), symptom assessment (n=18), appropriate symptoms (n=10), dementia (n=6)
The Technology	231 (20.3)	Usability (n=96), registration device (n=42), information boxes (n=32), improvement suggestions (n=30), technical problems (n=25), zero function (n=5), missing data (n=1)
The Value Proposition	179 (15.7)	Perceived value (n=78), changed care practices (n=49), statistics use (n=42), data quality (n=10)
The Adopters	333 (29.3)	Competence (n=72), new employees (n=45), colleagues' views (n=42), knowledge of residents (n=39), engagement (n=37), knowledge development (n=34), expectations (n=29), barriers (n=23), temporary staff (n=12)

[a]NASSS: nonadoption, abandonment, scale-up, spread, and sustainability framework.

[b]BPSD: behavioral and psychological symptoms of dementia.

Domain: The Organization

Organizational routines shaped how Daily-BPSD registrations were carried out in practice. As context for the registration patterns presented below, day shifts typically comprised 3 staff members per unit, with 1 designated to complete the shift registration; these shift registrations were usually completed toward the end of the shift. Staff occasionally consulted colleagues to clarify observations for residents with whom they had less contact.

Units used different strategies to assign this responsibility. In some cases, it was integrated into existing routines by linking the task to other documentation responsibilities. As staff members described, "When we sit down to sign off medications and do our documentation, we added it as part of that routine as well" [Unit B, Staff Participant 3] and "Then we talk it through—the person on the green shift keeps track of the resident we're registering, checks how the past day has been. And the others who worked the day also contribute, reflecting and choosing which ratings fit" [Unit C, Staff Participant 6], illustrating how registration responsibilities were integrated into established practical arrangements within the unit. At night, staff members sometimes worked alone while covering multiple wards, meaning they held overall responsibility for several residents simultaneously.

Staff emphasized the importance of ensuring that all team members were introduced to the tool and felt confident using it, so that the process would not depend on a few individuals. Several also emphasized the importance of initiating new cases directly within the application, without waiting for administrator support.

As shown in Table 3, registration adherence was high, exceeding 70% for all shifts, except for one unit (unit B) where adherence was significantly lower.

XSL•FO
RenderX

Table . Proportion of completed item-level NPI[a] registrations per resident and shift type during the study, calculated as the proportion of expected item-level NPI registrations for each resident–shift-type combination (denominator: 12 NPI domains × 90 study days = 1080 expected item-level registrations per resident per shift type). Residents are coded by unit (A-D) and a sequential resident study number (1-8).

Resident	Morning, %	Evening, %	Night, %	Overall, %	Staff, n
A1	91	90	76	85	12
A2	92	91	80	88	12
B3	58	66	13	46	10
B4	43	47	10	33	8
C5	89	83	98	90	7
D6	87	87	87	87	15
D7	87	86	92	88	15
D8	83	84	87	85	15
Mean	79	77	68	75	12[b]

[a]NPI: Neuropsychiatric Inventory.

[b]Total unique staff members across all residents: 44.

Analysis of registration frequency was conducted to examine changes in the number of completed item-level NPI registrations over the study period. Item-level NPI registration counts were divided into 3 intervals based on the number of days since the first registration for each resident: days 1 - 30 (n=6881), days 31 - 60 (n=6607), and days 61 - 90 (n=5635). The number of completed item-level registrations decreased in the final period compared to the earlier phases. The decline in the final interval coincided with the summer vacation period, when temporary staff without access to the tool were more common (see Strengths and Limitations).

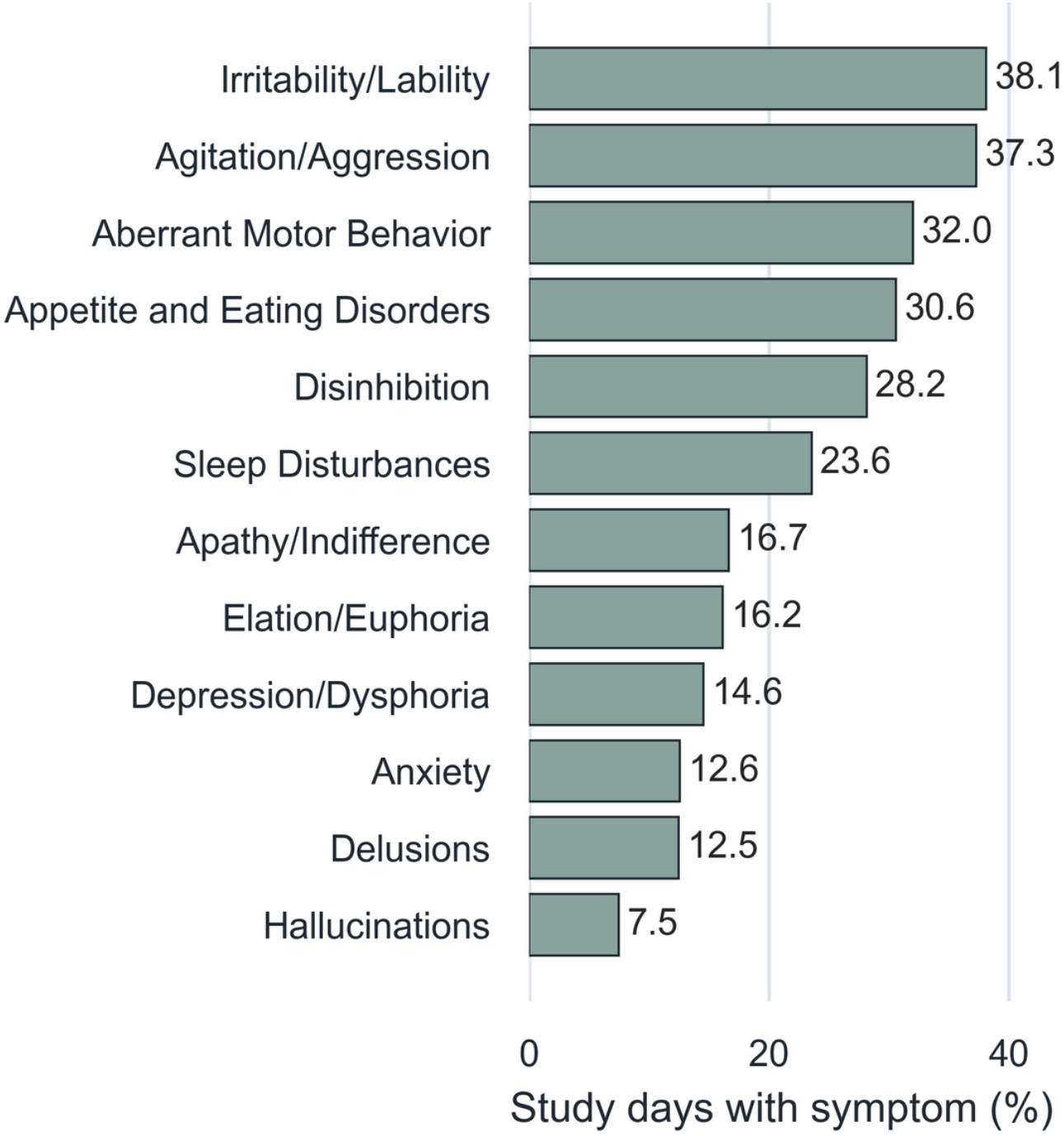
Domain: The Condition or Illness

The occurrence of BPSD symptoms varied across residents, both in terms of overall frequency and in the types of symptoms observed. As shown in Figure 2, the proportion of days with observed NPI symptoms differed between symptom domains. Across the study period, staff recorded symptom ratings (severity 0 - 3) for the 8 included residents, who each contributed 52 - 90 observation days (mean 82.1). Irritability/lability (38.1%) and agitation/aggression (37.3%) were the most frequently observed symptoms, indicating that these behaviors were present on more than one-third of all registered days and represented the most pervasive symptoms. The least frequently reported symptoms included *delusions* (12.5%) and *hallucinations* (7.5%), suggesting that psychotic features were relatively uncommon in daily care observations.

The number of symptoms recorded per day varied across residents during the study period. Median daily symptom counts ranged from 2 to 3 for most residents, although the spread within individuals varied substantially. Some residents showed relatively stable distributions with limited day-to-day variation, whereas others displayed wider distributions and occasional peaks with higher symptom counts (see Figure 3).

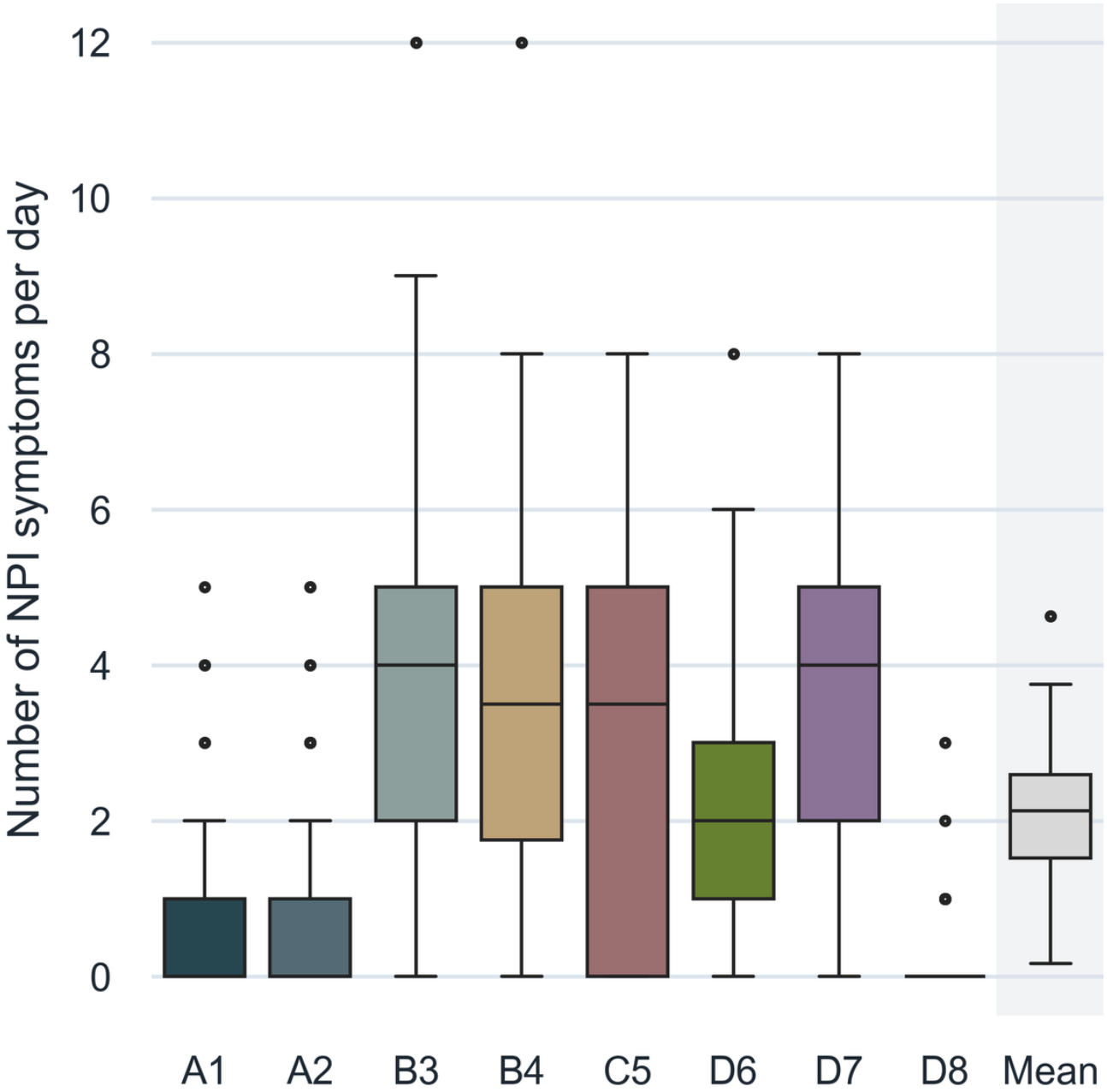
XSL•FO
RenderX

Figure 2. Proportion of observation days with recorded Neuropsychiatric Inventory (NPI) symptoms during the study period. Bars represent, for each symptom, the equal-weight mean across the 8 residents of the proportion of each resident's observation days on which the symptom was recorded at least once (each resident contributing 52 - 90 observation days).



XSL•FO
RenderX

Figure 3. Number of reported Neuropsychiatric Inventory (NPI) symptoms per day across participants. Boxplots illustrate the distribution of daily recorded NPI symptoms for each resident across the 90-day study period. Each box represents IQR (25th-75th percentile), the horizontal line within the box indicates the median, whiskers extend to 1.5 × IQR, and dots represent outliers. The rightmost box "Mean (equal-weight)" shows the distribution of daily equal-weight means across all residents, where each resident contributes equally regardless of the number of registrations. Differences in medians and spread indicate variation in both overall symptom levels and day-to-day fluctuations among residents.



The nature of BPSD symptoms also shaped the interpretive demands placed on staff conducting daily observations. In the interviews, participants perceived certain symptoms, such as anxiety, hallucinations, and delusions, as particularly challenging to identify and assess in terms of severity. As one staff member reflected, "I think we've gotten better at distinguishing whether it's a hallucination or a delusion […] Sometimes it's a bit like, which is which, but then I find it helpful that you can read a little" [Unit D, Staff Participant 9]. The staff described that such situations often required close collaboration within the team to reach consensus and ensure reliable assessments. Knowing the individual resident was viewed as crucial for accurately interpreting symptoms, while temporary or new staff were said to struggle more with recognizing behavioral expressions correctly. Participants noted that the supportive texts (see Figure 1) integrated into the application were helpful when uncertainty arose. Although collective reflections based on group statistics were uncommon, participants reported that registering symptoms daily increased their confidence in assessing symptoms and provided better conditions for adjusting care to residents' needs.

Domain: The Technology

User Experience Feedback

Staff consistently described the Daily-BPSD tool as easy to use and time-efficient. The registration process was generally perceived as quick and straightforward, and few participants reported technical difficulties. As one participant explained, "It has been easy to log in, easy to fill in, and I think it doesn't take

XSL•FO
RenderX

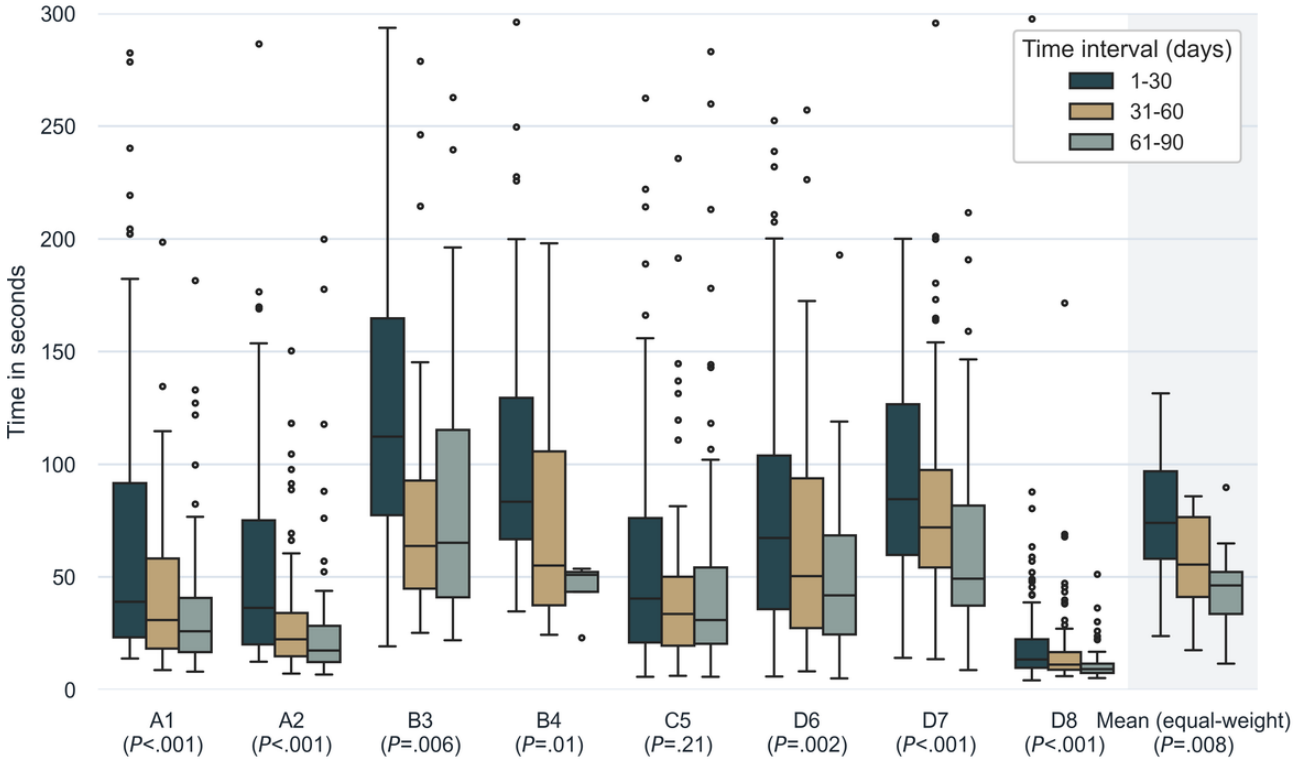
very long" [Unit D, Staff Participant 7], reflecting the general perception that the tool could be integrated smoothly into daily routines. In the interviews, most staff described the smartphone as their most common device for completing registrations, with computers and tablets used occasionally. Night staff emphasized the usefulness of the existing feature, allowing all symptoms to be marked as "no observed occurrences" when a resident was asleep during the shift.

While the overall usability was rated positively, several participants noted areas where functionality could be improved. Some expressed a need for a clearer overview of remaining variables to register, as this occasionally led to omissions—particularly among the individually selected variables. Others highlighted that the information button linked to each symptom was highly appreciated, as it provided helpful definitions and guidance in moments of uncertainty. For future large-scale use, participants suggested adding administrative features to simplify the management of user accounts and the linking of individually selected variables to specific residents.

Registration Time

Consistent with staff perceptions that the tool was quick and easy to use, quantitative analysis of shift-registration times showed a reduction in the time required to complete a shift registration (measured from form opening to submission for one resident at one shift, ie, one complete shift-form) over the study period. Shift-registration times decreased significantly for all residents except C5 (P=.21) (see Figure 4). At the group level, the overall mean shift-registration time, calculated as an equal-weight average across residents (ie, each resident contributing equally regardless of the number of registrations), decreased from 77.1 seconds in the initial period to 46.2 seconds in the final period (P=.008), indicating that completing a full shift registration for all NPI domains together with any individually selected variables required less than 1 minute on average. This quantitative finding aligns with staff perceptions that the tool was quick to use in everyday practice.

Figure 4. Average shift-registration time per participant across three 30-day intervals. Distribution of shift-registration times per resident during the observation period (days 1 - 30, 31 - 60, and 61 - 90). Each box represents IQR (25th-75th percentile), the horizontal line within the box indicates the median, whiskers extend to 1.5 × IQR, and dots represent outliers. Color shading distinguishes the 3 time intervals. *P* values indicate Welch *t* test comparisons between the first and last interval. The category "Mean (equal-weight)" represents the distribution of daily means, with each resident weighted equally regardless of the number of registrations.



Domain: The Value Proposition

Participants described several ways in which daily symptom registration created value in their work. Staff reported that the process of assessing symptom severity on a daily basis increased their understanding of BPSD and strengthened their confidence in recognizing and discussing behavioral changes. As one participant expressed, "Then maybe you've understood that now she's probably going to get angry, because you've seen that it happens every time, so maybe it's washing her hands that triggers it" [Unit C, Staff Participant 6], illustrating how repeated symptom assessment was perceived as a learning process that increased staff awareness of behavioral changes. A reduction in average shift-registration time over the study period (see Figure 4) may also reflect increasing familiarity with both the tool and the symptom domains. Some participants also saw potential in using the aggregated data for team discussions. As one staff member noted, "I think that should have been the goal at team meetings, to bring up those statistics and see that he is much calmer when he has done this" [Unit A, Staff Participant

XSL•FO
RenderX

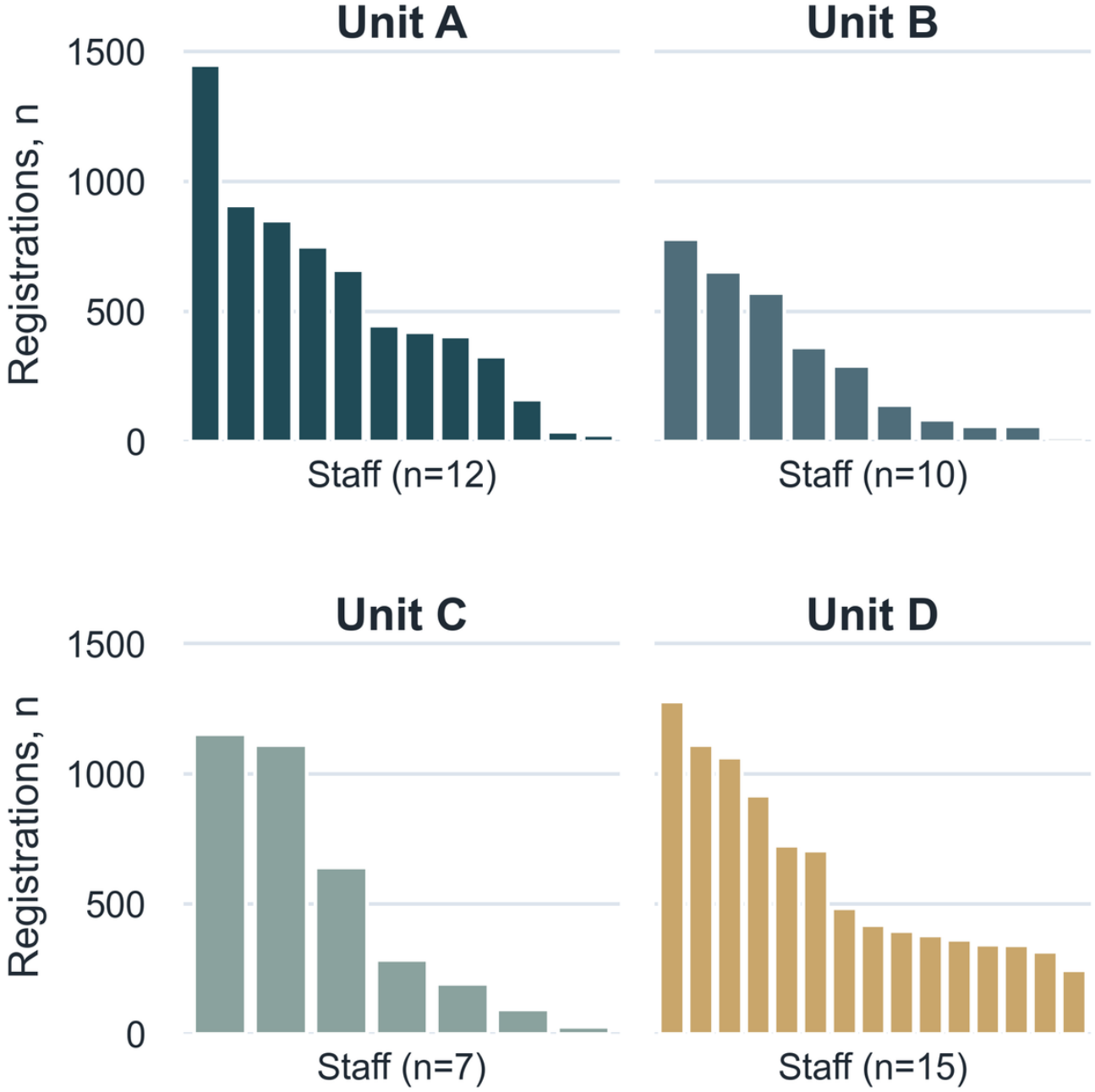
1]. However, only a few participants reported that the collected statistics had been actively used to adjust care for individual residents, although this was not an intended component of the pilot. Most participants nevertheless felt that the structured daily observational routine had clear potential to support both staff and residents if adopted more broadly in practice.

None of the participants raised economic issues, consistent with the study design, as they did not have financial responsibilities within their roles.

Domain: The Adopters

Participants emphasized the importance of engaging all staff members in the use of Daily-BPSD to ensure consistent registrations and shared ownership of the process. On some units, however, the registration effort was unevenly distributed: although all staff were expected to contribute, a few individuals carried out most of the registrations while others contributed comparatively little. As illustrated in Figure 5, the number of item-level registrations varied considerably among staff members within each unit, reflecting participants' descriptions that some individuals took greater responsibility for completing Daily-BPSD registrations than others.

Figure 5. Number of Daily-BPSD item-level registrations per staff member across the 4 participating units. Each bar represents a unique staff member.



Participants also emphasized the importance of involving other health care professionals, such as nurses, occupational therapists, and physiotherapists, to improve the multidisciplinary value of Daily-BPSD. As one participant noted, "Personally, I think it's better when everyone is involved. The more people you have, the more you can bounce ideas around. And then everyone is informed. Everyone has the same goal" [Unit C, Staff Participant 4]. Some noted that nurses currently rely on separate

XSL•FO
RenderX

documentation systems, which limit information sharing and hinder the integration of Daily-BPSD data into broader care planning.

Some participants also expressed that Daily-BPSD might, in the long term, facilitate registrations in the national BPSD Registry by providing more structured, up-to-date information for use in formal assessments.

Integrated Findings

Integration synthesized quantitative registration data with qualitative staff accounts within each of the 4 feasibility dimensions, characterizing each as showing convergence, complementarity, or divergence as defined in the Methods; a joint display summarizing the integrated evidence per dimension is provided in Multimedia Appendix 3.

For *implementation*, overall adherence of 75% varied substantially between units, with resident-level adherence ranging from 33% to 90% (Table 3). The qualitative data complemented this pattern by identifying two contrasting ways of organizing registration responsibility, as a shared task integrated into existing documentation routines, or as a task concentrated in a few individuals and therefore vulnerable to staffing disruptions, that plausibly underlie such between-unit variation.

For *practicality*, the reduction in shift-registration time from approximately 77 to 46 seconds (P=.008) converged with staff descriptions of the routine as quick, easy to log into, and naturally fitted to the closing routines of a shift.

For *acceptability*, quantitative adherence indicated *that* the routine was carried out, while qualitative data described *why* staff found it worth carrying out: reflection prompts, increased symptom awareness, and improved team conversations about residents. Complementarity dominated convergence here. A divergent signal also emerged: although staff perceived the routine as valuable, few reported that aggregated registration data had been used to actively adjust care during the pilot, suggesting that perceived value preceded actual use.

For *integration* into existing workflows, adherence differed markedly between units (Table 3). The qualitative accounts described how registration was fitted into a unit's existing routines, by linking it to documentation or medication tasks and designating a staff member per shift to complete it, mechanisms that plausibly account for how readily the routine was absorbed into established workflows. Across the 4 dimensions, the convergent design surfaced not only points of agreement but also complementary and divergent signals that the quantitative data alone could not have revealed.

Discussion

Principal Findings

This pilot study examined the feasibility of integrating structured daily observational routines, supported by digital documentation, into residential dementia care, and also described the variation in BPSD registrations between residents and from day to day. Across the 4 feasibility dimensions, the primary outcome of *implementation* was supported by an overall adherence of 75% across shifts (range 33% - 90% between residents), indicating that structured daily symptom registration is achievable in routine care but exhibits substantial between-unit variation. *Practicality* was indicated by a reduction in shift-registration time from approximately 77 to 46 seconds over the 90-day period, consistent with staff accounts of the tool as quick and easy to use. *Acceptability* was reflected in staff descriptions of perceived value, increased symptom awareness, and meaningful contribution to their work. *Integration* into existing workflows differed between units, paralleling adherence patterns and reflecting differences in how registration responsibility was incorporated into established routines. Beyond feasibility, the data showed substantial variation in BPSD: NPI symptom prevalence ranged from 7.5% (hallucinations) to 38.1% (irritability/lability), and daily symptom counts varied both between and within residents, supporting the rationale for structured high-frequency observation as a means of capturing meaningful symptom variability.

A fundamental tension exists between the rapid fluctuation of BPSD and the episodic nature of current assessment systems. The strong adherence observed across most units shows that structured, frequent observational documentation does not conflict with routine dementia care, even in settings with workload and staffing challenges. The findings provide preliminary support for the argument that more frequent observational practices may help address the documentation gap between daily symptom dynamics and formal assessment systems [33,49].

At the same time, the decline in registrations during periods with increased temporary staffing highlights an important organizational vulnerability. Structured observational routines may be more sustainable when staff work within shared responsibilities and established workflows. In this study, some units integrated the registration task into collective routines, while others relied more on individual initiative. The considerable variation in individual registration activity within units suggests that integrating high-frequency documentation as a shared professional responsibility, rather than a task carried by a few committed individuals, remains a key implementation challenge in residential care settings.

The variation in completion rates also suggests that local organizational characteristics, rather than shift type alone, influenced adherence. Units with lower adherence, particularly during night shifts where staff sometimes worked alone across multiple wards, may require additional support to ensure consistency. This aligns with implementation research emphasizing that digital innovations become integrated into everyday workflows, through leadership engagement and ongoing reinforcement within teams [50,51].

The results also contribute to the discussion on tacit knowledge in long-term care. Staff described how conducting structured daily symptom assessments enhanced their awareness and reflective engagement with BPSD. This suggests that formalizing observational routines may not only stabilize documentation but also make previously tacit clinical judgments more explicit by encouraging staff to regularly articulate, compare, and reflect on their assessments of residents'

XSL•FO
RenderX

symptoms. At the same time, participants reported difficulties distinguishing between overlapping symptoms such as anxiety and agitation or hallucinations and delusions. Such challenges echo previous research demonstrating the interpretive complexity of BPSD assessment, even among experienced professionals [30,52]. Structuring observation does not eliminate interpretive uncertainty; rather, it makes it more visible and underscores the need for shared definitions and continuous training. Viewed through the NASSS framework, these findings illustrate how the adopter and condition domains interact: the interpretive complexity inherent in BPSD (condition) places demands on staff competence and confidence (adopters) that cannot be resolved by technical solutions alone.

Taken together, the findings indicate that when high-frequency data collection is applied to observation-based, subjectively interpreted phenomena, the conditions for feasibility differ from those of instrumental monitoring: sustainability was shaped less by technical infrastructure than by organizational continuity and shared professional practice. In this sense, structured daily observational documentation functions as a bridge between informal, experience-based knowledge and formalized assessment systems, and may complement existing quality registers by rendering day-to-day fluctuations more systematically visible without replacing professional judgment. Consistent with its feasibility focus, the study characterizes the conditions under which such documentation can be carried out but did not evaluate care outcomes: although staff reported increased symptom awareness and saw potential in using aggregated observations for team discussions, few had actively used the data to adjust care during the study period. Whether these conditions translate into more attentive care, improved symptom management, or resident-level outcomes was not assessed and requires direct evaluation in subsequent studies.

Future research should examine how such routines function across different organizational settings and over longer periods, the role of individualized staff-selected variables in capturing person-specific behavioral patterns, and how observational practices can be scaled and adapted across diverse care contexts.

Strengths and Limitations

A key strength of this pilot study lies in its mixed methods design, which enabled a comprehensive examination of the feasibility and perceived value of structured, high-frequency observational documentation in dementia care. This design combined structured observational data on symptom patterns and adherence to the monitoring routine with qualitative accounts of staff experiences and the contextual conditions shaping its use.

However, several limitations should be acknowledged. The study involved a small, heterogeneous sample from a limited number of units in two municipalities in Sweden, limiting the transferability of the findings to other organizational contexts. Furthermore, the final study period (days 61 - 90) coincided with the summer vacation period, during which a higher proportion of temporary staff, who lacked access to the tool, were on duty, likely contributing to the observed decline in registrations. This seasonal effect should be considered when interpreting adherence patterns. Variation in adoption patterns among staff members also indicates that unit-level feasibility does not necessarily imply uniform individual-level engagement.

The absence of resident-reported outcomes limits conclusions regarding the experiential impact of structured documentation on person-centered care. Furthermore, the relatively short 90-day follow-up period restricts assessment of long-term sustainability and penetration—key implementation outcomes in digital health research. As a pilot study, the findings primarily inform feasibility rather than effectiveness.

Despite these limitations, the study provides preliminary empirical support for the feasibility of integrating structured daily observational routines into residential dementia care and offers a foundation for further refinement and evaluation in larger-scale studies.

Conclusions

This pilot study demonstrates that structured high-frequency observational documentation can be feasibly integrated into residential dementia care when supported by clear routines and team continuity. Integrating such practices into everyday work can establish a structured basis for systematic observation and team communication; whether this translates into earlier detection of behavioral changes or improved care quality requires further evaluation in studies designed to assess clinical and person-level outcomes.

The findings highlight that successful implementation depends less on technical functionality and more on organizational readiness, staff competence, and time allocated for reflection. Although Daily-BPSD served as the empirical platform in this study, the broader contribution concerns the feasibility of structured daily symptom documentation as a complement to existing quality registers. Such approaches may provide a more frequent and detailed informational basis for ongoing care processes; whether this can support more proactive and person-centered dementia care in practice remains an empirical question for future research.

Acknowledgments

We thank the unit managers and residential care staff who participated in this study for sharing their time and insights. Their contributions were essential for this study and for deepening our understanding of its use in everyday dementia care.

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to generative AI tools under full human supervision: code optimization, visualization, proofreading and editing, translation, reformatting, and publication support. The generative AI tool used was Claude Opus 4 (Anthropic). Generative AI was used only to support language editing, translation, formatting, and manuscript preparation. All

XSL•FO
RenderX

study design, data collection, analysis, interpretation, and conclusions were carried out by the authors, who reviewed and verified all AI-assisted text and take full responsibility for the content. Generative AI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

Grants were received from FORTE: the Swedish Research Council for Health, Working Life and Welfare (registration number 2021-01792). Open access funding provided by Jönköping University.

Data Availability

The datasets generated and analyzed during this study are not publicly available and cannot be shared. The data comprise sensitive personal data concerning health, namely daily registrations of behavioral and psychological symptoms for individuals living with dementia, together with verbatim transcripts of interviews with care staff. Under the terms of the ethical approval granted by the Swedish Ethical Review Authority (Dnr 2022-04509-01) and the consent obtained from participants and their legal representatives, and in accordance with the EU General Data Protection Regulation, these data may not be transferred outside the research team. Given the small and specific study population, the risk of reidentification cannot be excluded even from deidentified data; for this reason, the data cannot be made available, including upon individual request. The analysis code is available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JM, TB, CN, IK
Data curation: JM
Formal analysis: JM, IK
Funding acquisition: IK
Investigation: JM, TB, EO
Methodology: JM, TB, EO, IK
Project administration: JM
Software: JM
Supervision: TB, CN, EO, IK
Validation: JM, TB, EO
Visualization: JM
Writing – original draft: JM
Writing – review & editing: TB, CN, EO, IK

Conflicts of Interest

The Daily-BPSD application was developed collaboratively by the author team, with technical implementation by JM. The team, therefore, has a potential intellectual stake in the tool, which is disclosed here. Several steps were taken to limit its influence. The study aims and design were specified in the ethics application before data collection. During the registration period, JM’s contact with staff was limited to technical support and did not extend to influencing what they registered, so the registration data were generated independently of the developers. JM conducted the interviews together with EO and TB, and JM performed the coding, with EO and TB reviewing a subset of coded segments and providing input on categorization. Because the intellectual stake is shared across the whole team, these safeguards mitigate individual rather than team-level bias. No other conflicts of interest have been declared.

Multimedia Appendix 1
Semistructured interview guide used for the dyadic interviews with care staff.
[PDF File, 58 KB - aging_v9i1e98024_app1.pdf]

Multimedia Appendix 2
Example of the NASSS-guided coding matrix, showing coded segments across the 5 selected domains. NASSS: nonadoption, abandonment, scale-up, spread, and sustainability.
[PDF File, 46 KB - aging_v9i1e98024_app2.pdf]

Multimedia Appendix 3
Joint display of integrated quantitative and qualitative evidence per feasibility dimension.
[PDF File, 48 KB - aging_v9i1e98024_app3.pdf]

Multimedia Appendix 4

XSL•FO
RenderX

List of individually selected variables used in addition to the standard Neuropsychiatric Inventory domains in Daily-BPSD. [PDF File, 45 KB - aging_v9i1e98024_app4.pdf]

References

1. Jones CH, Dolsten M. Healthcare on the brink: navigating the challenges of an aging society in the United States. NPJ Aging 2024 Apr 6;10(1):22. [doi: 10.1038/s41514-024-00148-2] [Medline: 38582901]
2. Tynkkynen LK, Pulkki J, Tervonen-Gonçalves L, Schön P, Burström B, Keskimäki I. Health system reforms and the needs of the ageing population – an analysis of recent policy paths and reform trends in Finland and Sweden. Eur J Ageing 2022 Jun;19(2):221-232. [doi: 10.1007/s10433-022-00699-x] [Medline: 35465210]
3. Ogugua JO, Muonde M, Maduka CP, Olorunsogo TO, Omotayo O. Demographic shifts and healthcare: a review of aging populations and systemic challenges. Int J Sci Res Arch 2024;11(1):383-395. [doi: 10.30574/ijsra.2024.11.1.0067]
4. Khan HTA, Addo KM, Findlay H. Public health challenges and responses to the growing ageing populations. Public Health Chall 2024 Sep;3(3):e213. [doi: 10.1002/puh2.213] [Medline: 40496520]
5. Ekman I, Swedberg K, Taft C, et al. Person-centered care – ready for prime time. Eur J Cardiovasc Nurs 2011 Dec;10(4):248-251. [doi: 10.1016/j.ejcnurse.2011.06.008] [Medline: 21764386]
6. Rosengren K, Brannefors P, Carlstrom E. Adoption of the concept of person-centred care into discourse in Europe: a systematic literature review. J Health Organ Manag 2021 Dec 17;35(9):265-280. [doi: 10.1108/JHOM-01-2021-0008]
7. McArthur C, Bai Y, Hewston P, Giangregorio L, Straus SE, Papaioannou A. Barriers and facilitators to implementing evidence-based guidelines in long-term care: a qualitative evidence synthesis. Implementation Sci 2021 Dec;16(1):70. [doi: 10.1186/s13012-021-01140-0]
8. Moore L, Britten N, Lydahl D, Naldemirci Ö, Elam M, Wolf A. Barriers and facilitators to the implementation of person-centred care in different healthcare contexts. Scand J Caring Sci 2017 Dec;31(4):662-673. [doi: 10.1111/scs.12376] [Medline: 27859459]
9. Gopal G, Suter-Crazzolara C, Toldo L, Eberhardt W. Digital transformation in healthcare – architectures of present and future information technologies. Clin Chem Lab Med 2019 Feb 25;57(3):328-335. [doi: 10.1515/cclm-2018-0658]
10. Mitchell M, Kan L. Digital technology and the future of health systems. Health Syst Reform 2019;5(2):113-120. [doi: 10.1080/23288604.2019.1583040] [Medline: 30908111]
11. Pyper E, McKeown S, Hartmann-Boyce J, Powell J. Digital health technology for real-world clinical outcome measurement using patient-generated data: systematic scoping review. J Med Internet Res 2023 Oct 11;25:e46992. [doi: 10.2196/46992] [Medline: 37819698]
12. Maiorino MI, Signoriello S, Maio A, et al. Effects of continuous glucose monitoring on metrics of glycemic control in diabetes: a systematic review with meta-analysis of randomized controlled trials. Diabetes Care 2020 May;43(5):1146-1156. [doi: 10.2337/dc19-1459] [Medline: 32312858]
13. Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med 2019 Aug;25(8):1301-1309. [doi: 10.1038/s41591-019-0508-1] [Medline: 31308507]
14. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. Nature 2020 Jan 2;577(7788):89-94. [doi: 10.1038/s41586-019-1799-6]
15. Connor J, Flenady T, Massey D, Dwyer T. Clinical judgement in nursing – an evolutionary concept analysis. J Clin Nurs 2023 Jul;32(13-14):3328-3340. [doi: 10.1111/jocn.16469] [Medline: 35880251]
16. Casey AN, Low LF, Goodenough B, Fletcher J, Brodaty H. Computer-assisted direct observation of behavioral agitation, engagement, and affect in long-term care residents. J Am Med Dir Assoc 2014 Jul;15(7):514-520. [doi: 10.1016/j.jamda.2014.03.006] [Medline: 24755476]
17. Powers JH, Patrick DL, Walton MK, et al. Clinician-reported outcome assessments of treatment benefit: report of the ISPOR Clinical Outcome Assessment Emerging Good Practices Task Force. Value Health 2017 Jan;20(1):2-14. [doi: 10.1016/j.jval.2016.11.005] [Medline: 28212963]
18. Beattie Z, Miller LM, Almirola C, et al. The collaborative aging research using technology initiative: an open, sharable, technology-agnostic platform for the research community. Digit Biomark 2020;4(Suppl 1):100-118. [doi: 10.1159/000512208] [Medline: 33442584]
19. Laukvik LB, Lyngstad M, Rotegård AK, Slettebø Å, Fossum M. Content and comprehensiveness in the nursing documentation for residents in long-term dementia care: a retrospective chart review. BMC Nurs 2022 Apr 11;21(1):84. [doi: 10.1186/s12912-022-00863-9] [Medline: 35410289]
20. Eikelboom WS, Koch J, Beattie E, et al. Residential aged care staff perceptions and responses towards neuropsychiatric symptoms: a mixed methods analysis of electronic healthcare records. Aging Ment Health 2023 Feb;27(2):243-250. [doi: 10.1080/13607863.2022.2032597] [Medline: 35100918]
21. Tshering G, Troeung L, Walton R, Martini A. Factors impacting clinical data and documentation quality in Australian aged care and disability services: a user-centred perspective. BMC Geriatr 2024 Apr 12;24(1):338. [doi: 10.1186/s12877-024-04899-1] [Medline: 38609868]

XSL•FO
RenderX

22. Larjow E, von Fintel M, Busse A. A mixed-methods study of quality differences between applied documentation approaches in nursing homes. BMC Nurs 2022 Sep 29;21(1):265. [doi: 10.1186/s12912-022-01046-2] [Medline: 36171628]
23. Kariotis TC, Prictor M, Chang S, Gray K. Impact of electronic health records on information practices in mental health contexts: scoping review. J Med Internet Res 2022 May 4;24(5):e30405. [doi: 10.2196/30405] [Medline: 35507393]
24. Nichols E, Steinmetz JD, Vollset SE, et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8]
25. Dementia, fact sheet. World Health Organization. 2025 Mar 31. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-08-04]
26. Cerejeira J, Lagarto L, Mukaetova-Ladinska EB. Behavioral and psychological symptoms of dementia. Front Neurol 2012;3:73. [doi: 10.3389/fneur.2012.00073] [Medline: 22586419]
27. Laganà V, Bruno F, Altomari N, et al. Neuropsychiatric or behavioral and psychological symptoms of dementia (BPSD): focus on prevalence and natural history in Alzheimer's disease and frontotemporal dementia. Front Neurol 2022;13:832199. [doi: 10.3389/fneur.2022.832199] [Medline: 35812082]
28. Górska S, Forsyth K, Maciver D. Living with dementia: a meta-synthesis of qualitative research on the lived experience. Gerontologist 2018 May 8;58(3):e180-e196. [doi: 10.1093/geront/gnw195] [Medline: 28069886]
29. Savva GM, Zaccai J, Matthews FE, et al. Prevalence, correlates and course of behavioural and psychological symptoms of dementia in the population. Br J Psychiatry 2009 Mar;194(3):212-219. [doi: 10.1192/bjp.bp.108.049619] [Medline: 19252147]
30. Melander C, Sävenstedt S, Olsson M, Wälivaara BM. Assessing BPSD with the support of the NPI-NH: a discourse analysis of clinical reasoning. Int Psychogeriatr 2018 Apr;30(4):581-589. [doi: 10.1017/S1041610217002009] [Medline: 28965503]
31. Malm J, Odzakovic E, Kåreholt I, Fristedt S, Bielsten T. Exploring conditions for implementing daily monitoring of behavioral and psychological symptoms of dementia (BPSD) in dementia care: insights from a participatory design study. BMC Geriatr 2026 Mar 11;26(1):457. [doi: 10.1186/s12877-026-07321-0] [Medline: 41814181]
32. Jönsson L, Wibom M, Londos E, Nägga K. Person‐centered care at population scale: the Swedish Registry for Behavioral and Psychological Symptoms of Dementia. Alzheimers Dement Transl Res Clin Interv 2025 Jan;11(1):e70057. [doi: 10.1002/trc2.70057]
33. Fauth EB, Gibbons A. Which behavioral and psychological symptoms of dementia are the most problematic? Variability by prevalence, intensity, distress ratings, and associations with caregiver depressive symptoms. Int J Geriatr Psychiatry 2014 Mar;29(3):263-271. [doi: 10.1002/gps.4002] [Medline: 23846797]
34. Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 6;10(1):1. [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
35. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs – principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
36. Bowen DJ, Kreuter M, Spring B, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457. [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
37. Nationella riktlinjer för vård och omsorg vid demenssjukdom: Stöd för styrning och ledning [in Swedish]. : Swedish National Board of Health and Welfare; 2017 Report No.: 2017-12–2 URL: https://www.socialstyrelsen.se/contentassets/8315c1b65e064fcb92065392c74b942c/2017-12-2.pdf [accessed 2026-7-18]
38. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994 Dec;44(12):2308-2314. [doi: 10.1212/wnl.44.12.2308] [Medline: 7991117]
39. Cummings JL. Neuropsychiatric Inventory – Nursing Home Version (NPI-NH). : npiTEST; 2009 Jun 1 URL: https://www.dementiaresearch.org.au/wp-content/uploads/2016/01/NPI-NH_cr.pdf [accessed 2026-06-30]
40. Eisikovits Z, Koren C. Approaches to and outcomes of dyadic interview analysis. Qual Health Res 2010 Dec;20(12):1642-1655. [doi: 10.1177/1049732310376520] [Medline: 20663940]
41. Morgan DL, Ataie J, Carder P, Hoffman K. Introducing dyadic interviews as a method for collecting qualitative data. Qual Health Res 2013 Sep;23(9):1276-1284. [doi: 10.1177/1049732313501889] [Medline: 23925406]
42. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
43. Abell B, Naicker S, Rodwell D, et al. Identifying barriers and facilitators to successful implementation of computerized clinical decision support systems in hospitals: a NASSS framework-informed scoping review. Implement Sci 2023 Jul 26;18(1):32. [doi: 10.1186/s13012-023-01287-y] [Medline: 37495997]
44. James HM, Papoutsi C, Wherton J, Greenhalgh T, Shaw SE. Spread, scale-up, and sustainability of video consulting in health care: systematic review and synthesis guided by the NASSS framework. J Med Internet Res 2021 Jan 26;23(1):e23775. [doi: 10.2196/23775] [Medline: 33434141]

XSL•FO
RenderX

45. Kadesjö Banck J, Bernhardsson S. Experiences from implementation of internet-delivered cognitive behaviour therapy for insomnia in psychiatric health care: a qualitative study applying the NASSS framework. BMC Health Serv Res 2020 Aug 8;20(1):729. [doi: 10.1186/s12913-020-05596-6] [Medline: 32771024]
46. Derrick B, White P. Why Welch's test is type I error robust. Quant Methods Psychol 2016;12(1):30-38. [doi: 10.20982/tqmp.12.1.p030]
47. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
48. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. Köln Z Soziol 2017 Oct;69(S2):107-131. [doi: 10.1007/s11577-017-0454-1]
49. Lee KH, Lee JY, Kim B. Person-centered care in persons living with dementia: a systematic review and meta-analysis. Gerontologist 2022 Apr 20;62(4):e253-e264. [doi: 10.1093/geront/gnaa207]
50. Cresswell K, Sheikh A. Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review. Int J Med Inform 2013 May;82(5):e73-e86. [doi: 10.1016/j.ijmedinf.2012.10.007]
51. Dugstad J, Eide T, Nilsen ER, Eide H. Towards successful digital transformation through co-creation: a longitudinal study of a four-year implementation of digital monitoring technology in residential care for persons with dementia. BMC Health Serv Res 2019 Jun 10;19(1):366. [doi: 10.1186/s12913-019-4191-1] [Medline: 31182093]
52. Chan DKY, Chan LKM, Kuang YM, Le MNV, Celler B. Digital care technologies in people with dementia living in long-term care facilities to prevent falls and manage behavioural and psychological symptoms of dementia: a systematic review. Eur J Ageing 2022 Sep;19(3):309-323. [doi: 10.1007/s10433-021-00627-5] [Medline: 36052197]

Abbreviations

BPSD: behavioral and psychological symptoms of dementia
NASSS: nonadoption, abandonment, scale-up, spread, and sustainability
NPI: Neuropsychiatric Inventory
NPI-NH: Neuropsychiatric Inventory – Nursing Home version

Edited by H LaMonica, T Leung; submitted 12.Apr.2026; peer-reviewed by LB Lauvik, S Dys, V Rashedi; revised version received 10.Jun.2026; accepted 18.Jun.2026; published 23.Jul.2026.

Please cite as:
Malm J, Bielsten T, Nilsen C, Odzakovic E, Kåreholt I
Daily Monitoring of Behavioral and Psychological Symptoms of Dementia in Residential Care: Mixed Methods Pilot Study
JMIR Aging 2026;9:e98024
URL: https://aging.jmir.org/2026/1/e98024
doi:10.2196/98024

© Johannes Malm, Therése Bielsten, Charlotta Nilsen, Elzana Odzakovic, Ingemar Kåreholt. Originally published in JMIR Aging (https://aging.jmir.org), 23.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Providing Tech Support as Care Work Among Care Workers in Assisted Living Facilities: Qualitative Interview Study

Yiyi Wu[1,2], MA; Nicole Ross[1], MS; Sunyoung Kim[1], PhD

[1]School of Communication and Information, Rutgers, The State University of New Jersey, New Brunswick, NJ, United States

[2]Brookdale Center for Healthy Aging, Hunter College, New York, NY, United States

Corresponding Author:
Yiyi Wu, MA
School of Communication and Information
Rutgers, The State University of New Jersey
4 Huntington St
New Brunswick, NJ, 08901
United States
Phone: 1 848 932 7500
Email: yw1280@scarletmail.rutgers.edu

Abstract

Background: Technology can enhance the quality of life and social engagement for older adults; yet, many require assistance to use it effectively. In assisted living facilities, care workers play a crucial role in supporting residents' use of technology. However, little research has examined the experiences and challenges of care workers in this context.

Objective: In this study, we examine the perspectives and experiences of care workers in supporting older residents with technology.

Methods: We conducted semistructured interviews with 20 professional care workers in the United States, with an average of 4.9 years of experience working in assisted living facilities.

Results: The findings indicate that participants regularly provided technology support to their residents, while their perspectives on and acceptance of this technology support role varied. For many, this responsibility placed additional demands on their existing workload, requiring both explicit and nuanced physical and emotional labor input.

Conclusions: These findings underscore the need for greater recognition and systemic support for care workers as they increasingly take on the responsibility of facilitating technology use in residential care settings.

(JMIR Aging 2026;9:e80272) doi:10.2196/80272

KEYWORDS

assisted living; care workers; long-term care; older adults; technology use; technology support

Introduction

Background

Assisted living facilities, a type of long-term care, serve older adults who can no longer live safely and independently at home. These residential communities provide housing and are staffed with care workers who support residents with a wide range of daily needs and activities [1]. Assisted living emphasizes person-centered care within a home-like environment, catering to a broader spectrum of care and social needs, including those of residents with complex medical conditions [2]. This sector represents a fast-growing segment of long-term care in the United States [3]. However, chronic staffing shortages pose ongoing challenges, which are exacerbated by the rapidly growing aging population [4]. In response, digital and emerging technologies have been explored as a promising means to help bridge this caregiving gap [5,6]. Examples include the use of information and communication technologies (ICTs) to facilitate proactive social engagement and health communication, as well as wearable devices for disease self-management [6-8]. These innovations have demonstrated significant potential to enhance the well-being of older adults, improve their quality of life, and reduce caregiver burden, both within long-term care settings and beyond [6,9].

Yet, the successful applications of technology depend on effective usage, and older adults often face barriers due to lower digital skills, physical and cognitive declines, and a lack of prior exposure to technology [10,11]. While the ownership of digital

XSL•FO
RenderX

technologies among older adults has grown [12], they continue to struggle with independent use, often experiencing anxiety, distrust, and avoidance of technology [13-15].

Consequently, studies consistently highlight that older adults' adoption and effective use of technology heavily depend on their immediate support networks, such as their children and community workers, for assistance with technology [16-18]. In assisted-living settings, care workers often serve as a critical source of technology support for older residents [19]. The term "care workers" refers to individuals who provide essential physical, emotional, and practical support to people requiring assistance with activities of daily living due to age, illness, disability, or other life circumstances [20]. This category encompasses a wide range of occupational roles, including nurses, nursing assistants, and care aides [4,20]. Despite variation in professional training and job titles, care workers share core responsibilities centered on facilitating mobility, administering medication, assisting with daily routines, and supporting the overall well-being of care recipients. Collectively, their labor constitutes a critical component of long-term care systems and plays a vital role in maintaining the quality of life of older adults [4,21].

Previous research on technology use in long-term care indicates that as more older residents adopt digital technologies, their reliance on care workers for technological support may also introduce additional work demand and challenges for care workers [21,22]. However, while emerging studies have begun to attend to those challenges [19,21], the literature has given little attention to care workers' perspectives on and experiences with providing such technology support to older residents. Consequently, there remain limited insights into how to best support care workers within increasingly digitalized caregiving environments.

Understanding the perspectives and experiences of care workers is particularly important given the vulnerabilities they face in the workplace. Caregiving often demands intensive physical and emotional labor, with care workers frequently expected to invest effort beyond their formal job requirements while receiving limited financial compensation and insufficient recognition from employers [21-24]. As the demand for assisted living continues to grow [25], providing better support for care workers is essential for enhancing staff retention and the quality of care.

To explore these issues, we conducted semistructured interviews with 20 care workers employed in assisted living facilities in the United States. Our study aimed to answer the following research questions:

1. How do care workers perceive their involvement in providing technology support?
2. What challenges do care workers face as they navigate the provision of technology support?

Our findings indicate that, although participants held varying views on their engagement in technology support, they consistently recognized it as an integral part of care workers' daily responsibilities. This role adds new dimensions of physical and emotional labor beyond their core job duties. Moreover, the heavy reliance of older residents on care workers for tech support highlights the need for greater recognition of this labor, as well as involving both care workers and older adults in the design of technology for older adults. Additionally, our findings highlight the importance of support programs that empower care workers, enabling them to establish clear work boundaries and navigate their roles more effectively in long-term care settings.

Literature Review

Care Workers' Roles and Workload

Care workers in long-term care broadly refer to individuals who deliver direct care services to people needing daily assistance, typically older adults, in home or residential settings [26]. In residential long-term care facilities, such as assisted living, these professionals are commonly referred to as care assistants, social workers, nursing assistants, and other health care professionals [20,27]. Prior studies indicate that care workers' roles and job responsibilities are often loosely defined, primarily determined by the individualized needs of their residents [28]. This role ambiguity is particularly pronounced in residential care settings, where person-centered care is a key emphasis [29,30]. While fostering meaningful relationships with residents is central to care work, this closeness can blur professional boundaries, making it difficult to delineate responsibilities [29,30]. As a result, care workers frequently feel an overarching sense of duty for residents' well-being, taking on multiple roles beyond their formal job descriptions [28,30]. Qualitative studies on care workers' attitudes toward their responsibility further highlight a persistent tension between care workers' commitment to holistic resident care and the emotional strain of multitasking, burnout, and heavy workloads [23,31].

Commonly recognized care workers' routine responsibilities include, but are not limited to, hygiene assistance, medication management, mobility support, case management, and administrative tasks [32]. While these duties already demonstrate the labor-intensive nature of caregiving, research further reveals that many essential caretaking roles and tasks remain unseen and unpaid [23,28]. These include the emotional investment involved in building trust relationships with clients, managing their clients' emotional responses, and regulating one's own emotions in high-stress environments [33,34]. Additionally, care workers often coordinate to facilitate communication between their clients, health care providers, and other third parties [19,28]. Ongoing studies have continuously advocated for greater recognition of care workers' contribution, particularly the less recognized aspects of their labor, to improve their overall financial and social well-being [23,29].

Technology Use Among Older Adults in Long-Term Care

A growing body of literature has explored the role of technology in enhancing the quality of life and supporting independent engagement in social and health management activities among older adults in long-term care settings [7,9,35]. ICTs, such as computers, laptops, smartphones, and mobile apps, play a crucial role in enabling older residents to maintain connections with family and friends outside of facilities and engage in various

XSL•FO
RenderX

leisure activities [7,36,37]. The use of social media and videoconferencing platforms has been shown to effectively reduce social isolation and feelings of loneliness [37-39]. Beyond social engagement, ICTs are also widely used to deliver essential care services, such as telemedicine and mobile health (mHealth) applications, which facilitate the remote provision of routine care and health information [40,41]. These applications help deliver effective personalized care [40]. Moreover, technologies gaining popularity in long-term care facilities also include wearable devices, such as smartwatches and fitness trackers [6,42], as well as assistive technologies, such as autonomous wheelchairs, designed to enhance functional capacities [43]. These tools support mobility, help monitor health indicators, promote physical activity, and support self-management of chronic diseases [40-43].

The growing use of technologies within long-term care facilities is partly driven by older adults' voluntary adoption, as evidenced by the steady increase in personal device ownership within this demographic [12]. Additionally, as the long-term care sector continues to experience staffing shortages, these technologies are increasingly recognized for their potential to bridge workforce gaps [44]. Consequently, many facilities have expanded their technological resources by providing residents with access to ICT, assistive, and wearable devices, while actively promoting their use [9,17,37,45]. In some cases, technologies have become so embedded in facility operations that residents must engage with them to carry out routine activities and access services and information within the facility [45]. This trend underscores the increasing significance of technology in the daily lives of older residents in long-term care.

Dependency, Technology Support, and Potential Challenges

While most technologies require the direct involvement of older adults, studies have highlighted the heightened challenges that older users face in realizing the intended benefits of these tools [18,46]. Older adults exhibit wide variation in digital literacy, prior experience with technology, and physical or cognitive abilities, all of which affect their capacity to use digital tools independently [13,14,46]. Despite the growing availability of technological resources in long-term care facilities, individuals with higher levels of technological proficiency or prior experience are more likely to engage with these tools [17]. In contrast, those with limited experience or skills often remain disengaged, even when they recognize the potential benefits of technology use [16,17]. Moreover, regardless of their skill levels and motivation, many older adults still struggle with basic interactions due to age-related declines in vision, hearing, fine motor skills, or cognitive processing [21,36]. The continuous diversification of devices and interfaces, each requiring different levels of engagement, from basic device operation to security credential management, can further exacerbate the distrust and anxiety associated with technology use for older adults with cognitive and functional impairments [14,15,46].

Consequently, older adults often rely on close social networks for assistance in using technology effectively [16,18]. These networks typically include family caregivers or other younger family and community members who live in close proximity to older adults [16-18]. In long-term care facilities, several studies have shown that professional care workers often become the primary source of technology support for older residents [19,21]. Nevertheless, providing technology support can be challenging [21,22]. For instance, Hardy et al [22] found that facilitating videoconferencing for cognitively impaired older residents was physically demanding, requiring multiple steps to connect them to virtual platforms. Similarly, Chew et al [21] reported dissatisfaction was found among care workers tasked with setting up and managing virtual visits during the COVID-19 pandemic, a responsibility they found burdensome due to visitor restrictions that placed greater reliance on digital communication.

Despite increasing integration of technology in long-term care facilities, limited research has specifically examined care workers' experiences with providing technology support to their older residents. While prior research emphasizes the importance of care workers in daily resident support, little is known about how technology-related responsibilities intersect with their existing workloads. This gap in the literature risks rendering this emerging aspect of care work unrecognized, thereby negatively affecting their financial and social well-being [22,23]. By examining the direct experiences and perspectives of care workers, this study aims to inform strategies that better support them in long-term care settings.

Methods

Participants

We conducted semistructured interviews with 20 individuals working in assisted living facilities in the Greater New York area. Participants were eligible if they had worked in any care worker role at an assisted living facility for more than 1 year and were at least 18 years old. To recruit participants, we contacted assisted living facilities in the region and provided managers with a detailed explanation of the study's objectives. Upon their approval, recruitment flyers were distributed via email, and managers and supervisors referred potential participants to the research team.

The final sample included 20 participants (9 male and 11 female), with an average of 4.9 (SD 3.08) years of work experience and an age range of 23 to 56 (mean 38.15, SD 8.23) years. All participants were paid professional care workers employed by an assisted-living facility at the time of the interviews. While the resident base in these facilities exhibits varying levels of independence and diverse health needs, many live with chronic conditions, including diabetes, hypertension, or heart disease, and require mobility support, including assistance with walking and using mobility aids. Some also require monitoring for safety or adherence to care routines.

All participants' job titles were self-reported. Nine participants identified as caregivers or care assistants; 6 held medical-care-oriented roles, such as nurse or medication assistant, and 4 worked in social-care-oriented roles, such as case managers and activity assistants. One participant held an administrative role but still reported frequent direct interactions

XSL•FO
RenderX

and support for residents. A detailed breakdown of participant demographics is provided in Table 1. Job descriptions are synthesized from participants' accounts and sources, including professional associations such as the American Nurses Association, government websites like the U.S. Department of Health and Human Services, and career websites like Glassdoor. While they offer a general understanding of each role, they may not fully capture the daily experiences of care workers, a discrepancy explored in the Results section.

Table 1. Participant demographics.

ID	Gender	Age (years)	Work experience (years)	Job title
P1	Female	48	5	Case Manager[a]
P2	Female	40	8	Case Manager
P3	Female	56	10.5	Case Manager
P4	Male	39	3	Caregiver[b]
P5	Male	33	6	Caregiver
P6	Female	39	10	Caregiver
P7	Female	32	3	Care Assistant[b]
P8	Female	35	2	Care Assistant
P9	Female	32	4	Care Assistant
P10	Female	30	3	Care Assistant
P11	Male	23	3	Care Assistant
P12	Female	29	2	Care Assistant
P13	Male	38	7	Nurse[c]
P14	Female	47	1.5	Nurse
P15	Female	51	1	Nurse
P16	Male	40	8	Nurse
P17	Female	30	1	Nursing Assistant[d]
P18	Male	37	5	Medication Assistant[e]
P19	Male	35	7	Activity Assistant[f]
P20	Female	49	8	Administrator[g]

[a]Case managers are health care professionals who help clients with health management, service coordination, and patient advocacy.

[b]Care assistants, caregivers, or personal care aides are care professionals who provide a broad spectrum of services to those who need assistance with daily activities. Their duties include personal care, domestic care, and companionship.

[c]Nurse is a health care professional who attends to patients' biological, physical, and behavioral needs. In the United States, all nurses are required to complete formal training to obtain licenses and certifications.

[d]Nurse assistants provide support in daily activities, health monitoring, and hygiene for clients. Different from nurses, nursing assistants are not always certified.

[e]Medication assistant helps administer medications in a hospital or other care facility.

[f]An activities assistant helps to plan events for socialization and enjoyment in a variety of settings.

[g]Long-term care administrator plans, organizes, directs, and manages the operation and implements policies.

Data Collection

An interview guide was developed to facilitate open-ended, semistructured interviews (Multimedia Appendix 1). The guide consisted of several key topics, with the first few questions concerning participants' job titles and responsibilities, as well as their day-to-day work life. The latter part of the questions focused on participants' perceptions of their residents' technology use and experiences helping their residents with technology-related inquiries. This set of questions is followed by those concerning their use of technology. All interviews were conducted virtually using the videoconferencing software of the participants' choice (eg, Zoom). Each session lasted up to 60 minutes, allowing for an in-depth discussion of their experiences. With participants' consent, all interviews were audio-recorded and transcribed using transcription software. Each transcript was then carefully reviewed and manually corrected to ensure accuracy.

Data Analysis

We used a grounded theory approach to identify recurring patterns within our interview data [47]. The process began with

XSL•FO
RenderX

open coding, where transcripts were systematically examined to label key concepts related to care workers' perspectives and experiences with tech support [48]. We then conducted axial coding, organizing related concepts into broader categories, referred to as phenomena, such as recurring events, actions, or interactions that illustrate how care workers navigate their tech support roles. Lastly, we engaged in selective coding to synthesize the phenomena into a cohesive framework, capturing overarching themes and relationships within the data. This iterative approach continued until data saturation was achieved, ensuring a comprehensive understanding of participants' experiences.

Ethical Considerations

The study received approval from the institutional review board (IRB) at Rutgers, The State University of New Jersey (IRB# Pro2021001731). At the beginning of each interview, participants were informed about the research procedure and their rights as participants, including the voluntary nature of participation and the right to withdraw at any time. All participants provided voluntary, informed, and written consent to participate in the study and had the interview audio recorded. All recordings were deleted upon transcription. The research team ensured that the transcripts contained no identifying information about the participants and that all data was stored on a passcode-protected drive accessible only to the research team. As a token of appreciation, participants received a US $20 e-gift card upon completing the interview.

Results

Overview

Through an analysis of participants' experiences in providing care services to residents, three key themes emerged: (1) the current state of care workers' daily responsibilities, (2) care workers' perspectives on their engagement in technology support, and (3) challenges and labor demands associated with providing technology support. Since this study was conducted in an assisted living setting, we use the term "resident" to refer to the older adults served by the participants. This terminology also aligns with how participants commonly referred to those in their care.

The Current State of Care Workers' Daily Responsibilities

Multifaceted Job Responsibilities

At the beginning of each interview, participants were asked to describe their job titles, primary responsibilities, and a typical day at work. Despite variations in job titles, all participants reported being actively involved in providing a range of direct care services to residents, extending beyond the generally defined duties associated with each job title. Caregivers, for instance, described a wide array of duties addressing residents' individual needs holistically, including assistance with meals, hygiene, medication, mobility, and companionship. Participants in medical-care-oriented positions, such as nurses and medication assistants, whose primary responsibilities involved medication administration, condition monitoring, and care plan development, also frequently assisted with personal care tasks, including hygiene and mobility support. Similarly, the administrator, despite primarily handling administrative duties, also engaged directly with residents and their families to ensure their needs were met. Across participants' narratives, the extensive list of daily tasks highlighted the labor-intensive and multifaceted nature of their work.

> *I help in assisting condition management...administering medications...monitoring them, and ensuring that they carry out the proper exercises...I also help them with grooming and hygiene needs...household tasks like washing the dishes and cleaning up pots.* [Participant 6, caregiver]

> *Medication administration, insulin administration, doctor's visits...Assistance with transfers, ambulation, personal hygiene...On some occasions, nurses would help with showers if we had a shortage of nurses. I am in charge of a lot of responsibilities that change on a daily basis.* [Participant 14, nurse]

There was a strong consensus among the participants that their primary goal was to provide responsive, individualized care tailored to residents' needs. Consequently, their daily tasks were not rigidly defined but instead heavily shaped by evolving demands of those they cared for. The fluid nature of their work demanded flexibility and adaptability to provide timely and effective support, requiring them to juggle multiple responsibilities and roles. However, the continuous demands of addressing age-related frailty and the complexities of care sometimes led to experiences of burnout.

> *In the mornings, people usually call on the system for assistance if they need to. And then they start looking for things to do. We also accompany them if they want us to or if they need supervision. So, it's not like I have a specific role that you carry out every day. I have to switch and perform different roles.* [Participant 19, activities assistant]

> *My typical day at our facility was kind of hectic because I had a lot of roles to take on in the facility. That is the administration of medicine, and also daily monitoring and reporting of vital signs. These residents, most of them are old, so they get sick quite often. On a busy day, on average, my day was kind of tiresome because I had to do this work...it was always tough for me.* [Participant 13, nurse]

Among all these responsibilities, technology support consistently emerged as a significant and recurring task, which we explore in detail in the next section.

Tech Support as a Frequent Need of Residents

When initially outlining their responsibilities, participants rarely mentioned providing technology support. However, upon deeper reflection on their daily interactions with residents, in the context of technology use, it became clear that technology is well-integrated into residents' lives, but often at the condition of their ongoing assistance. Across the board, participants reported varying levels of digital technology use among residents in their facilities, and that many residents frequently

XSL•FO
RenderX

relied on their assistance to use these technologies. The most used devices were residents' personal smartphones and tablets, which they primarily used to access a range of mobile apps. Communication apps, such as WhatsApp, Facebook, and Skype, were frequently used to stay connected with family members and access social networks outside the facilities, as well as to stay informed about news. Additionally, some residents were observed using health and wellness applications, including those for fitness tracking and medication management. A smaller number of participants noted residents using gaming apps and streaming services, such as YouTube. Beyond personal devices, some facilities provided residents with access to shared technological resources, actively encouraging their use. For instance, one facility offered virtual reality equipment to enhance exercise routines, while another maintained a computer lab for residents' personal use. In many cases, participants noted that residents, whether using personal or facility-provided technology, frequently sought their assistance when encountering difficulties or needing additional guidance.

> *They mostly use the apps and platforms I mentioned to reconnect with their family members back home. So, these are the most commonly used in the facility...Some use it on their own, while others seek assistance from me and other staff working in the facility.* [Participant 5, caregiver]
>
> *They like to make video calls with their family, like Skype, and use the laptop. Sometimes, we help them out with that. And there are VR fitness oculars that we sometimes use to engage them in physical activities. And they actually also require assistance with using them.* [Participant 12, primary caregiver]

When explicitly prompted to reflect on the frequency of technology support, participants not only revealed the high occurrence of such tasks but also the high percentage of residents performing them. This suggests that providing technology support becomes an essential part of their daily duties, alongside their other caregiving responsibilities.

> *(Q: Do residents often ask for instructions on using these technologies? [referring to residents' phones]) Yeah, most of the time, 80% of time I would say.* [Participant 18, medication assistant]
>
> *(Q: Do residents often ask for instructions [referring to the in-facility computer lab]?) Yeah, I'd say a big percentage of them. They ask for help whenever they need help and instructions.* [Participant 13, nurse]

Care Workers' Perspectives on Their Engagement in Technology Support

Care Workers' Degrees to Accepting a Technology Support Role

As participants reflected more deeply on residents' technology use and support needs, their narratives revealed varying attitudes toward the boundaries between providing tech support and their assigned job responsibilities. On the one hand, some participants viewed tech support work positively, embracing it as a natural extension of their duties. This perspective was often rooted in the belief that technology assistance represents an emerging and essential need among residents. Given that their responsibilities were already seen as multifaceted, they felt that addressing these needs naturally aligned with their role in providing comprehensive care.

> *(Q: Do residents help each other with technology?) No, they don't. Because it's the staff's job to help them.* [Participant 8, care assistant]
>
> *When residents have difficulties with technology or ask for help, we were all more than happy to help with computers or cell phones...because it was something for them.* [Participant 14, nurse]

In contrast, other participants perceived technology support as outside the scope of their assigned responsibilities. While not entirely opposed to assisting residents, they emphasized the challenges associated with incorporating an additional tech support role into their already multifaceted duties and demanding schedules. This burden was further exacerbated by staffing shortages in long-term care facilities. In response, some participants advocated for alternative solutions, such as hiring dedicated technology support professionals, assigning specific staff for tech-related tasks, or implementing resident training programs to promote digital independence and reduce reliance on care workers for technical assistance. These varying perspectives demonstrated the tensions between evolving care demands in the digital age and traditional understandings of caregiving.

> *Staff is busy. It's not always feasible to help somebody. They have their own work to do.* [Participant 15, nurse]
>
> *Problems arise due to not having enough staff to show residents how to use it...I know a lot of facilities right now experience a lack of staff. Especially with using technology, I believe it needs to be professional, in some way.* [Participant 3, case manager]
>
> *It would be a nice development if all the facilities can give some kind of sensitization or training to residents on how to operate basic technologies themselves, other than leave the tasks solely to staff.* [Participant 4, caregiver]

Dual Recognition of Technology as Essential and Independent Use as Unachievable

Despite different opinions on whether technology support was a part of their core duties or not, all participants continued to assist residents with technology. This was driven, in part, by a shared acknowledgment that technology use is important to their residents and particularly beneficial in providing the familiar emotional support that participants were unable to offer. As identified by existing studies, one of the major challenges older adults residing in long-term care facilities face is experiences of loneliness and isolation due to reduced opportunities to interact with their families and neighbors [49]. Participants' narratives also echoed these findings, underlying the prevalence of this issue, particularly among those whose families did not visit regularly. In this case, ICTs served as the primary tool for maintaining residents' communication with their social networks outside the facilities.

XSL•FO
RenderX

> *They seek out help regularly, because it's not something they can do without. They have to stay in touch with their loved ones, and they rely on technology to receive daily reminders about their medications.* [Participant 16, nurse]

> *A lot of families come in to visit, but some of them don't...Nowadays, Facebook and Skype is probably the best way since technology makes it much easier.* [Participant 15, nurse]

> *They would have a better quality of life if they can actively participate using technology to improve their memory and well-being, because technology has its minuses and pluses. But overall, it keeps people connected in such a disconnected world.* [Participant 2, case manager]

Moreover, there was a collective acknowledgment among participants that the issue related to residents' heavy reliance on care workers for tech support is inevitable. As revealed by participants, it is simply not possible for residents with heightened age-related physical and cognitive disabilities to use technologies independently. While their deep understanding and empathy for their residents are admirable, it is also crucial to attribute the root cause of this issue not to the inevitable age-related constraints, but to the flawed, existing technology systems that consistently overlook the accessibility needs of older adults [50]. In one facility where a voice assistant was used, the participant demonstrated ongoing difficulties; as such, the device consistently failed to recognize and execute the commands of Spanish-speaking residents. She explicitly attributed her residents' reliance on staff for tech support to the lack of inclusive technology design, highlighting how language barriers further complicated their ability to navigate digital tools independently.

> *I think it's normal for someone of their age or physical situation to sometimes require some assistance with the use of technology.* [Participant 12, caregiver]

> *Not everybody's able to do it due to memory issues...so we're always there to assist them.* [Participant 15, nurse]

> *It's very difficult for bilingual residents to use, because in our facility we have Spanish-speaking residents. They have accents and intonation when speaking English. It's very hard for voice assistants to pick up those particular commands from them. So, there are a lot of errors...Those are the reasons related to the interface of those particular technologies that pose difficulties for them...they seek out help regularly.* [Participant 16, nurse]

Technology Support as a Rewarding Experience

While perceptions toward the role boundaries between caregiving and technology support varied, there is a general acknowledgment that the experience of tech support could be rewarding. For instance, participants shared sentimental stories of helping residents reconnect with a long-lost family member through videoconference or experience the outside world through virtual reality devices. Furthermore, participants' kindness and efforts in providing technology support were reciprocated by residents' gratitude and positive emotional feedback, allowing participants to have a sense of meaning and fulfillment in their work. By supporting residents with technology-related issues, participants also strengthened trust and fostered positive relationships, creating a more positive work environment.

> *They do appreciate my assistance with assisting them with their technology.* [Participant 4, caregiver]

> *By integrating them with necessary VR applications, we are able to take them back to their favorite places, like going on tours or climbing mountains...While such interventions can take hours...they responded positively. So it wasn't that stressful.* [Participant 16, nurse]

> *I was helping a resident establish a full connection to a podcast they were listening to. They needed assistance to get the full connection, and they were very happy...We are there to help all of them, so I am always available to help any resident who needs assistance.* [Participant 13, nurse]

Challenges and Labor Demands Associated With Providing Technology Support

The Repetitive Burden of Tech Support

Although not all participants explicitly voiced complaints and concerns about the workload associated with providing tech support, a closer examination of their detailed experiences revealed that this role was both physically and emotionally demanding. When prompted to share their observations on the coping strategies residents used to deal with technology-related difficulties, only a few participants mentioned occasional instances of mutual support among residents. In most cases, participants identified that residents relied primarily on staff for assistance, positioning care workers as the first responders to residents' tech-related requests with the expectation that they would be readily available whenever needed. The high volume and frequency of these demands further exacerbated their already labor-intensive primary responsibilities, contributing to burnout, as outlined in the first section.

> *I feel like their greatest strategy is the fact that they actually have the staff to rely on. So, whenever they have difficulties and all of that, they have the staff at their disposal to help them out.* [Participant 7, care assistant]

> *The only strategy they have is calling for help. So, when they aren't understanding, you have them yelling your name, trying to call you. So sometimes you can have like five people calling you at the same time. It can be really draining and challenging.* [Participant 18, medication assistant]

Moreover, the workload was further complicated by the repetitive nature of technology-related requests, which often involved recurring issues. As previously mentioned, residents required the most support with emerging digital technologies, particularly mobile apps on their phones and tablets. Modern

XSL•FO
RenderX

apps often require users to navigate multistep interactions, manage various settings, and adapt to frequent updates to fully utilize their features. Consequently, assisting residents often involves step-by-step navigation and detailed instructions. However, the initial guidance efforts seldom led to meaningful learning or skill acquisition among residents, leading to repeated interventions for the same issues. As a result, tech support work often became a repetitive task, requiring significant patience, time, and emotional energy, which further strains participants' already demanding roles.

> *You just try to open the apps, see the updates, and then just explain to them what the problem is and show them how they'll go about it. Then the next time they are in the same situation.* [Participant 19, care assistant]

> *Initially, he was unsure of how to use it (an app that tracks residents' movement). I explained that when he wakes up, he needs to open the app, place the phone either on his body or hold it in his hand...He didn't fully understand on the first day, so I had to repeat the instructions multiple times.* [Participant 9, care assistant]

Additional Responsibilities Associated With a Tech Support Role

In addition to the taxing and multifaceted nature of juggling multiple roles and responsibilities, their experiences also revealed that "tech support" encompasses far more than simply addressing technical difficulties. Instead, it often required participants to take on multiple responsibilities, often involving complex coordination tasks. Many participants found themselves acting as coordinators, facilitating communications between residents and third parties, such as families and medical providers. These coordinating tasks included setting up the devices, coordinating schedules with various external parties, and even securing permissions for communication when needed. These responsibilities were particularly pronounced for participants assisting residents experiencing cognitive impairments.

> *I always double-check with the family if it's okay to [call] or no, because you don't know what they're gonna open up and you know viruses and everything else on the internet.* [Participant 15, nurse]

> *We tend to help them do that because some of them don't get to have their telephones with them. So, we tend to call their families for them on a weekly basis or, in special cases, like the woman (with dementia and only remember her child at times) that wanted to see her child at the time. So, we tend to communicate to the families for them.* [Participant 10, care assistant]

Several facilities that offered communal technology resources and devices for residents actively promoted their use to enhance residents' health and social engagement. However, these technologies were rarely used by residents without assistance from participants. In facilities where these technology resources were more established, there were formal guidelines regulating residents' use of technology. Participants' narratives suggest that such regulation stems from the facility's concern about potential damage to the high-cost devices. As a result, participants were directly assigned supervisory responsibilities by their employers to ensure the proper functioning, maintenance, and safe use of these devices. Additionally, while the purpose of technologies such as medicine dispensaries is to facilitate residents' disease self-management, their lack of technology proficiency also led to heightened concerns among participants about the potential health hazards resulting from technology misuse. Consequently, participants voluntarily took on supervisory roles to monitor residents' technology use.

> *In our facility, we have rules because if residents are allowed to use this technology on their own, they may mess things up. So, we guide them to ensure maximum utilization of the technology.* [Participant 13, nurse]

> *This digital medication dispenser, there was a day a resident who is on medication forgot to take her medication. So, she was having doubts since she had missed a dose...I just had to monitor her to ensure that she made use of the dispenser in my presence.* [Participant 20, administrator]

Moreover, intertwined with their coordination role, participants often had to advocate for technology use among residents who avoided it for various reasons, particularly in facilities with communal technology resources. Aligning with prior research on older adults' reduced interest in technology adoption [11], participants observed varying degrees of technology hesitation, especially among the oldest residents. Despite well-intentioned efforts to introduce mobile apps and other emerging technologies to enhance residents' health and social well-being, these efforts were sometimes met with dismissiveness or resistance. Encouraging reluctant residents to engage with technology required patience, empathy, and careful navigation of their moods to foster a supportive and encouraging environment.

> *Most of them are a bit older than I think, and for them to embrace the technology, they need so much convincing.* [Participant 19, activity assistant]

> *I had an encounter with one of the residents who was basically not in the mood because...she didn't want to actually use the app...I tried to convince her that it would be beneficial for her to try it.* [Participant 7, care assistant]

Emotional Dimension of Tech Support

Navigating technology later in life without prior exposure to digital learning experience can be emotionally challenging [51,52], particularly for older adults who wish to use technology but struggle to do so. Participants found these difficulties often triggered a range of negative emotions among residents, including frustrations, confusion, and mood swings, which created challenges in their daily work. As a result, participants were responsible not only for addressing technical problems but also for managing residents' emotional responses while using technologies. To de-escalate these situations, participants used various strategies, such as providing encouragement,

XSL•FO
RenderX

offering empathetic guidance, and maintaining patient and supportive communication.

> *What I've observed is that it's become difficult for some of them; they get depressed, or they feel less motivated because they are trying to access these technological platforms, but it's hard for them. Whenever I find them in these kinds of situations where they feel down or their mood is low, I initiate interactive conversations and some fun activities.* [Participant 5, caregiver]

> *They start with panic because something they really need is not working...So the interaction usually begins in a panic mode. After interacting with them, I assured them that we will take care of it and fix the issue as soon as possible. Then they calm down.* [Participant 6, caregiver]

Defensiveness was a particularly common emotional response among residents who highly valued their independence and believed they possessed sufficient knowledge. This often manifested as stubbornness or a reluctance to seek help. While frequent tech support requests could overwhelm participants' already strained schedules, residents' intentional avoidance of seeking help posed a different challenge. This was especially true for those managing communal technological devices, as they often had to resolve issues caused by residents' mishandling. To manage these situations, participants adopted proactive and sensitive approaches when offering assistance, ensuring that residents felt their independence was respected while still receiving the support they needed.

> *Some of these residents are stubborn...because they think they know, but they don't know. That's why it's important for them to seek help. Some of them are actually stubborn because they always assume they know things, but whenever they're trying to apply, they always mess up.* [Participant 13, nurse]

> *Most of the time when they come around..."I need assistance with this. I need to check this, I need to check that, and can you please help?"...If it's someone who couldn't act fast or you know that this person needs assistance, you just ask the person, Can I help you? Can I help you to do this?* [Participant 11, care assistant]

In turn, managing residents' emotions negatively impacted participants' own emotional well-being and can be mentally draining [53]. Yet, their emotional labor often goes unrecognized by their employers and even by the participants themselves, overshadowed by the deep emotional bonds they share with residents [54]. Correspondingly, in this study, while participants did not explicitly describe technology support work as emotionally burdensome, the constant need to navigate residents' frustrations with patience and sensitivity, alongside physical tasks, suggested that such work takes an emotional toll. Participants' narratives made it clear that to provide effective and satisfying technology support, they frequently needed to regulate their own emotions, even in the face of challenging behaviors.

> *One of the most important things in working in assisted living is that you have to always be calm, even when you are not pleased with some behaviors of the residents; you have to be the one who has to be calm.* [Participant 7, care assistant]

Discussion

Overview

Recent studies on technology use in long-term care facilities revealed the frequent need for care workers to provide technology assistance [21,22]. Nevertheless, research examining care workers' experiences with tech support remains limited. This underscores the need for a deeper understanding of their evolving role in the digital age. This study specifically focuses on the perspectives and experiences of care workers in assisting residents with technology in assisted living settings. In what follows, we discuss implications for better recognizing, supporting, and including care workers in the integration of technology in long-term care.

Increase Recognition of Labor Associated With a Tech Support Role

Our findings highlight that caregivers' technology support extends beyond assisting with devices, encompassing various roles and tasks generated by one seemingly singular tech support role. Echoing existing studies [21], tech support often leads to various coordination activities involving external parties, such as residents' families and health providers. Additionally, our studies uniquely highlight the supervisory responsibilities associated with maintaining and ensuring the safe use of technology devices in care facilities. This burden is particularly pronounced among participants in facilities that implement on-site technological solutions, contributing to workplace stress. As participants indicated, the potential damage to high-cost in-facility devices or residents' misuse of technology can be perceived as a failure in their professional practice. Such pressure, in turn, can compel them to take on even more supervisory and support tasks to mitigate risks, ultimately creating a relentless cycle of labor. As vulnerable workers, even minor perceived failures can overshadow their significant daily contributions, highlighting the urgent need for greater recognition of their labor input.

This study also highlights the emotionally challenging nature of tech support in caregiving. Research increasingly acknowledges the emotional labor inherent in caregiving, defined as the regulation of one's emotions to align with expected professional expressions, even when they conflict with personal feelings [33,53]. Consistent with prior studies [13,51], our participants described how older adults' struggles with technology often trigger frustration, confusion, and emotional outbursts. Consequently, caregivers frequently become the direct target of these frustrations, absorbing residents' emotional distress while striving to offer patient and compassionate support.

Despite being an integral part of resident care, these additional layers of labor associated with a tech support role remain largely unrecognized. Job descriptions and verbal accounts of caregiving

XSL•FO
RenderX

roles often fail to acknowledge the full extent of tech support responsibilities. As previous research noted, care workers' dedication to their professional identity should not justify the normalization of excessive labor [23,28,29]. The continued lack of recognition perpetuates the devaluation of caregiving work [28]. Therefore, given the rapid digitalization of the care sector, this study emphasizes the need for ongoing research and policy interventions to recognize the physical and emotional investments made by care workers in providing tech support and how these responsibilities contribute to their already demanding workloads. Additionally, employers must implement stronger support systems, such as on-site mental health services and training, and recognize this labor through improved compensation and workplace policies.

Technology Use, Design, and Implementation in Long-Term Care Needs Ongoing Careful Evaluation

We found a wide range of technologies are used in assisted living, including residents' personal devices and assistive technologies provided by the facilities. Participants also expressed an overall positive attitude toward residents' use of technology and their willingness to offer support. This willingness stemmed not only from residents' requests but also from participants' genuine belief that technology enhances the quality of life for residents. These findings align with the broader discourse and research on technology in long-term care, where its potential to improve the social and physical well-being of older adults has been widely acknowledged [6,7].

By highlighting how technological innovations introduce new responsibilities for care workers, this study calls for a more nuanced perspective in the ongoing discourse and research on the use of technology and caregiving in long-term care. While technologies can enhance service delivery, support aging independence, and assist care workers, their successful implementation should not be discussed in isolation from the labor of care workers. Too often, the discussions around technology in long-term care focus on the potential for automation to replace routine aspects of caregiving, which risks further devaluing care work [55]. Yet, as this study and others have shown [21], the success of residents' technology use and the implementation of new devices in facilities remain heavily dependent on the support of care workers. Future studies should evaluate the effectiveness of technology use in long-term care by centering care workers' perspectives on both the benefits and shortcomings these innovations introduce.

Moreover, the significant challenges older residents experienced when using these devices and their reliance on others for support, as identified by participants, highlight the persistent digital divide and the lack of age-inclusive design thinking in current technology systems [50,56]. While some may argue that such challenges are unique to the current cohort of older adults, who may have had limited opportunities to engage with technology earlier in life, other challenges, particularly those associated with age-related health conditions such as cognitive impairment, will remain relevant going forward. These findings further demonstrate the lack of inclusivity in existing technology design and the urgent need for improvement.

Existing studies on the technology use of older adults have advocated for a user-centered approach, emphasizing the importance of involving older adults in the design process [57-59]. Building on this, our study recommends the inclusion of care workers in technology design to integrate their expertise and acknowledge their central role in supporting older adults. By involving care workers, we can create more inclusive and practical care-based technologies that address the challenges faced by both older adults and their caregivers. Moreover, this collaborative approach ensures that technological advancements do not come at the expense of frontline care workers, reinforcing the need to value their labor rather than further marginalizing it.

Amplify Care Workers' Agency in Defining and Navigating the Work Boundaries

Consistent with previous research [29,60], this study demonstrates that care workers' job roles are often ambiguously defined and shaped by the evolving needs of residents. Then, the added responsibility of providing tech support further complicated their work [29,30]. While some participants attempted to distinguish their caregiving duties from tech support, their ongoing assistance with residents' technology needs blurred these boundaries over time. This role expansion, driven by their dedication to resident well-being, contributed to an increased workload and emotional exhaustion, aligning with prior research that links caregiving to burnout [61].

However, given the dynamic nature of person-centered care in assisted living, rigidly defining caregiving responsibilities may not be the most effective approach. Indeed, a degree of flexibility is beneficial, as it allows care workers to respond adaptively to residents' needs while reducing workplace strain [60,61]. Therefore, rather than imposing strict role definitions, we emphasize the importance of empowering care workers to define their own work boundaries within a supportive environment [60]. In practice, encouraging open discussions with peers and supervisors will help establish and adjust these boundaries as needed. Lastly, we recommend that employers acknowledge the additional labor associated with less-recognized tasks, such as tech support, and create spaces for care workers to reflect on these responsibilities. Practical steps include offering paid workshops, professional development opportunities, and structured communication channels to foster collaboration and ensure that care workers feel supported in balancing their evolving responsibilities while maintaining their well-being.

Limitations

Our findings must be considered in light of several limitations. First, our sample consisted mainly of younger care workers who did not report experiencing tech-related difficulties while assisting residents. Given that the US caretaking workforce is aging, with a median age of 47 years [62], future research should examine how age-related factors influence care workers' ability to provide tech support. Second, although we did not find differences in experiences across job titles, this may be because our analysis primarily focused on general care-support tasks that remain consistent across caregiving roles. Moreover, the small sample size within each job category might have limited

XSL•FO
RenderX

the feasibility of comparative analysis. Future studies should include a larger sample to explore potential variations in experiences across different caregiving positions. Third, all participants were based in the United States, meaning the findings primarily reflect Western perspectives. Cross-cultural comparisons can provide a more comprehensive understanding of how care workers use technology support. Fourth, the study relied on self-reported experiences, which may be subject to bias. Participants might have underreported difficulties or overemphasized their ability to provide tech support, potentially skewing the findings. Future research could incorporate diverse methods to capture a more objective understanding of the challenges faced by care workers.

Conclusions

Given the rise in technology adoption among older adults and in long-term residential facilities, this study explored the perspectives and experiences of care workers in providing technology support to residents. We identified the nuanced physical and emotional labor involved in technology support, which has become a core yet often unrecognized aspect of caregiving. Additionally, the varying perspectives of care workers on their technology support roles highlight ongoing tensions in work expectations within long-term care. Based on these findings, we advocate for greater recognition of care workers' efforts through recommendations for further research, improved workplace policies, and enhanced program design in long-term care.

Acknowledgments

The research would not have been possible without our participants' contributions. We deeply appreciate the time they took to talk with us and their valuable insights and perspectives.

Data Availability

Data will be made available on reasonable request.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Authors' Contributions

YW led the data analysis and manuscript writing and editing, and assisted with project administration. NR led the data collection and assisted with data analysis. SK oversaw the overall project administration, assisted with data analysis, and participated in manuscript writing and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview guide.
[DOCX File , 2855 KB - aging_v9i1e80272_app1.docx]

References

1. Long-term care facilities: assisted living, nursing homes, and other residential care. National Institute on Aging. 2023 Oct 12. URL: https://www.nia.nih.gov/health/assisted-living-and-nursing-homes/long-term-care-facilities-assisted-living-nursing-homes [accessed 2026-02-04]
2. Trinkoff AM, Yoon JM, Storr CL, Lerner NB, Yang BK, Han K. Comparing residential long-term care regulations between nursing homes and assisted living facilities. Nurs Outlook 2020;68(1):114-122. [doi: 10.1016/j.outlook.2019.06.015] [Medline: 31427078]
3. Grabowski DC, Stevenson DG, Cornell PY. Assisted living expansion and the market for nursing home care. Health Serv Res 2012;47(6):2296-2315 [FREE Full text] [doi: 10.1111/j.1475-6773.2012.01425.x] [Medline: 22578039]
4. Heiks C, Sabine N. Long term care and skilled nursing facilities. Dela J Public Health 2022;8(5):144-149 [FREE Full text] [doi: 10.32481/djph.2022.12.032] [Medline: 36751604]
5. Wang S, Bolling K, Mao W, Reichstadt J, Jeste D, Kim HC, et al. Technology to support aging in place: older adults' perspectives. Healthcare (Basel) 2019;7(2):60 [FREE Full text] [doi: 10.3390/healthcare7020060] [Medline: 30974780]
6. Wang J, Carroll D, Peck M, Myneni S, Gong Y. Mobile and wearable technology needs for aging in place: perspectives from older adults and their caregivers and providers. Stud Health Technol Inform 2016;225:486-490. [doi: 10.3233/978-1-61499-658-3-486]

XSL•FO
RenderX

7. Lemaire C, Humbert C, Sueur C, Racin C. Use of digital technologies to maintain older adults' social ties during visitation restrictions in long-term care facilities: scoping review. JMIR Aging 2023;6:e38593 [FREE Full text] [doi: 10.2196/38593] [Medline: 36599164]
8. Sun Y, Chen J, Ji M, Li X. Wearable technologies for health promotion and disease prevention in older adults: systematic scoping review and evidence map. J Med Internet Res 2025;27:e69077 [FREE Full text] [doi: 10.2196/69077] [Medline: 40553512]
9. Chapman SA, Miller JR, Spetz J. Emerging health technologies in long-term care and suppliers' views on their potential to assist and support the workforce. Med Care Res Rev 2023;80(6):619-630 [FREE Full text] [doi: 10.1177/10775587231190127] [Medline: 37534450]
10. van Deursen AJ. Digital inequality during a pandemic: quantitative study of differences in COVID-19-related internet uses and outcomes among the general population. J Med Internet Res 2020;22(8):e20073 [FREE Full text] [doi: 10.2196/20073] [Medline: 32750005]
11. Francis J, Ball C, Kadylak T, Cotten SR. Aging in the digital age: conceptualizing technology adoption and digital inequalities. In: Neves BB, Vetere F, editors. Ageing and Digital Technology: Designing and Evaluating Emerging Technologies for Older Adults. Singapore: Springer; 2019:35-49.
12. Faverio M. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. 2022 Jan 13. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2026-02-04]
13. Xu L, Lam YY, Ng DTK, Peng P, Suen WL, Lee CSY, et al. Digital literacy education for older adults: a scoping review. Proc Assoc Inf Sci Technol 2024;61(1):691-696. [doi: 10.1002/pra2.1081]
14. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. JMIR Aging 2023;6:e43197 [FREE Full text] [doi: 10.2196/43197] [Medline: 36943333]
15. Quan‐Haase A, Ho D. Online privacy concerns and privacy protection strategies among older adults in East York, Canada. J Assoc Inf Sci Technol 2020;71(9):1089-1102. [doi: 10.1002/asi.24364]
16. Lee S, Ory MG, Vollmer Dahlke D, Smith ML. Technology use among older adults and their caregivers: cross-sectional survey study. JMIR Aging 2024;7:e50759 [FREE Full text] [doi: 10.2196/50759] [Medline: 38717339]
17. Finkelstein R, Wu Y, Brennan-Ing M. Older adults' experiences with using information and communication technology and tech support services in New York City: findings and recommendations for post-pandemic digital pedagogy for older adults. Front Psychol 2023;14:1129512 [FREE Full text] [doi: 10.3389/fpsyg.2023.1129512] [Medline: 37138998]
18. Hunsaker A, Nguyen MH, Fuchs J, Djukaric T, Hugentobler L, Hargittai E. “He explained it to me and I also did it myself”: how older adults get support with their technology uses. Socius 2019;5. [doi: 10.1177/2378023119887866]
19. Kelly RM, Xing Y, Baker S, Waycott J. "More than just holding the iPad": family members' perspectives on the work of video calling in aged care. Proc ACM Hum Comput Interact 2024;8(CSCW2):1-26. [doi: 10.1145/3687007]
20. Scales K, Lepore MJ. Always essential: valuing direct care workers in long-term care. Public Policy Aging Rep 2020;30(4):173-177 [FREE Full text] [doi: 10.1093/ppar/praa022] [Medline: 38626233]
21. Chew HJ, Freeman S, Jackson P, Hemingway D, Klassen-Ross T, Martin-Khan M, et al. Technology use in long-term care during the COVID-19 pandemic: a qualitative study of paid employees' experiences in Western Canada. Digit Health 2024;10 [FREE Full text] [doi: 10.1177/20552076241303195] [Medline: 39711747]
22. Hardy MS, Fanaki C, Savoie C, Wilchesky M, Gagnon MP, Laberge M, et al. Long-term care staffs' experience in facilitating the use of videoconferencing by cognitively impaired long-term care residents during the COVID-19 pandemic: a mixed-methods study. BMC Health Serv Res 2024;24(1):646 [FREE Full text] [doi: 10.1186/s12913-024-11095-9] [Medline: 38769512]
23. Ming J, Kuo E, Go K, Tseng E, Kallas J, Vashistha A, et al. "I go beyond and beyond": examining the invisible work of home health aides. Proc ACM Hum Comput Interact 2023;7(CSCW1):1-21. [doi: 10.1145/3579492]
24. Shaw L, Masood M, Neufeld K, Connelly D, Stanley M, Guitar NA, et al. Work disparities and the health of nurses in long-term care: a scoping review. Healthcare (Basel) 2024;12(20):2065 [FREE Full text] [doi: 10.3390/healthcare12202065] [Medline: 39451482]
25. Harris-Kojetin L, Sengupta M, Park-Lee E, Valverde R, Caffrey C, Rome V, et al. Long-Term Care Providers and Services Users in the United States: Data From the National Study of Long-Term Care Providers, 2013–2014 [Vital and Health Statistics, Series 3, Number 38, DHHS Publication No. 2016–1422]. Hyattsville, MD: National Center for Health Statistics; 2016.
26. Global health and care workers compact: technical guidance compilation. World Health Organization. 2022 Apr 12. URL: https://www.who.int/publications/i/item/9789240073852 [accessed 2026-02-02]
27. Pace D, Betts S, Reda L, Griff M, Stratton L, Shubeck E, et al. Empowering a person-centered long-term care workforce. Alzheimers Dement 2024;20(9):6647-6653. [doi: 10.1002/alz.14111] [Medline: 39023198]
28. Gazzaroli D, D'Angelo C, Corvino C. Home-care workers' representations of their role and competences: a diaphanous profession. Front Psychol 2020;11:581399 [FREE Full text] [doi: 10.3389/fpsyg.2020.581399] [Medline: 33362645]
29. Klostermann J, Funk L. Bounding the boundless: gendered work hierarchies and “boundless work” in Ontario longterm care homes. Stud Polit Econ 2024;105(3):196-214. [doi: 10.1080/07078552.2024.2422684]

XSL•FO
RenderX

30. Johnson E. Dirt, decency and the symbolic boundaries of caregiving in residential homes for older people. Sociol Rev 2022;71(3):679-695. [doi: 10.1177/00380261221140240]
31. Czuba KJ, Kayes NM, McPherson KM. Support workers' experiences of work stress in long-term care settings: a qualitative study. Int J Qual Stud Health Well-being 2019;14(1):1622356 [FREE Full text] [doi: 10.1080/17482631.2019.1622356] [Medline: 31156047]
32. Elices Á, Prieto-Lobato JM, Rodríguez-Sumaza C. The role of social workers in long-term care for older adults: a mixed-methods systematic review in Europe and North America from 2000 to 2022. Int Soc Work 2024;68(1):7-21. [doi: 10.1177/00208728241269670]
33. Franzosa E, Tsui EK, Baron S. "Who's caring for us?": understanding and addressing the effects of emotional labor on home health aides' well-being. Gerontologist 2019;59(6):1055-1064. [doi: 10.1093/geront/gny099] [Medline: 30124808]
34. Duan Y, Thorne T, Iaconi A, Song Y, Saeidzadeh S, Doupe M, et al. Changing trends in job satisfaction and burnout for care aides in long-term care homes: the role of work environment. J Am Med Dir Assoc 2025;26(2):105380 [FREE Full text] [doi: 10.1016/j.jamda.2024.105380] [Medline: 39642910]
35. Hardy MS, Fanaki C, Savoie C. The use of technology by seniors with neurocognitive disorders in long-term care: a scoping review. BMC Geriatr 2024;24(1):573 [FREE Full text] [doi: 10.1186/s12877-024-05174-z] [Medline: 38961323]
36. Seifert A, Cotten SR. Digital distance in times of physical distancing: ICT infrastructure and use in long-term care facilities. Am Behav Sci 2023;69(9):1138-1154. [doi: 10.1177/00027642231155361]
37. Astell A, Dosanjh S, D'Elia T, Kokorelias KM, Stewart S, Grigorovich A, et al. Personalized tablets for residents in long-term care to support recreation and mitigate isolation. J Am Med Dir Assoc 2024;25(7):105022. [doi: 10.1016/j.jamda.2024.105022] [Medline: 38763162]
38. Schuster AM, Cotten SR. COVID-19's influence on information and communication technologies in long-term care: results from a web-based survey with long-term care administrators. JMIR Aging 2022;5(1):e32442 [FREE Full text] [doi: 10.2196/32442] [Medline: 34878989]
39. Baecker RM, Sellen KM, Crosskey S, Boscart V, Neves BB. Technology to reduce social isolation and loneliness. In: ASSETS '14: Proceedings of the 16th International ACM SIGACCESS Conference on Computers and Accessibility. New York: Association for Computing Machinery; 2014:27-34.
40. Adam E, Frielink N, Caljouw M, Bolt S, Embregts P, Zuidema S, et al. Toward successful eHealth adoption in long-term care: insights from a national survey of healthcare professionals. Digit Health 2025;11:20552076251376283 [FREE Full text] [doi: 10.1177/20552076251376283] [Medline: 40979700]
41. Theopilus Y, Davis H, Pedell S. Digital technology for supporting aged care services: a scoping review. In: CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery; 2025:1-12.
42. Auerswald T, Meyer J, von Holdt K, Voelcker-Rehage C. Application of activity trackers among nursing home residents: a pilot and feasibility study on physical activity behavior, usage behavior, acceptance, usability and motivational impact. Int J Environ Res Public Health 2020;17(18):6683 [FREE Full text] [doi: 10.3390/ijerph17186683] [Medline: 32937840]
43. Zheng K, Han F, Yang S, Li N. Building trust in long-term care settings using assistive technology: a systematic review. Front Rehabil Sci 2024;5:1492104. [doi: 10.3389/fresc.2024.1492104] [Medline: 39649373]
44. Tak SH, Benefield LE, Mahoney DF. Technology for long-term care. Res Gerontol Nurs 2010;3(1):61-72 [FREE Full text] [doi: 10.3928/19404921-20091103-01] [Medline: 20128544]
45. Jøranson N, Zechner M, Silva R, Yaylagul NK, Thygesen H. A case study on experiences with integrated technologies in a care home for older adults. Geriatr Nurs 2025;61:121-128 [FREE Full text] [doi: 10.1016/j.gerinurse.2024.10.059] [Medline: 39549444]
46. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024;24(1):506 [FREE Full text] [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
47. Corbin JM, Strauss A. Grounded theory research: procedures, canons, and evaluative criteria. Qual Sociol 1990;13(1):3-21. [doi: 10.1007/BF00988593]
48. Vollstedt M, Rezat S. An introduction to grounded theory with a special focus on axial coding and the coding paradigm. In: Kaiser G, Presmeg N, editors. Compendium for Early Career Researchers in Mathematics Education. Cham: Springer International Publishing; 2019:81-100.
49. Zhang D, Lu Q, Li L, Wang X, Yan H, Sun Z. Loneliness in nursing homes: a qualitative meta-synthesis of older people's experiences. J Clin Nurs 2023;32(19-20):7062-7075. [doi: 10.1111/jocn.16842] [Medline: 37605069]
50. Mannheim I, Wouters EJM, Köttl H, van Boekel LC, Brankaert R, van Zaalen Y. Ageism in the discourse and practice of designing digital technology for older persons: a scoping review. Gerontologist 2023;63(7):1188-1200 [FREE Full text] [doi: 10.1093/geront/gnac144] [Medline: 36130318]
51. An J, Zhu X, Wan K, Xiang Z, Shi Z, An J, et al. Older adults' self-perception, technology anxiety, and intention to use digital public services. BMC Public Health 2024;24(1):3533 [FREE Full text] [doi: 10.1186/s12889-024-21088-2] [Medline: 39702092]

XSL•FO
RenderX

52. Tsai HS, Shillair R, Cotten SR. Social support and "playing around": an examination of how older adults acquire digital literacy with tablet computers. J Appl Gerontol 2017;36(1):29-55 [FREE Full text] [doi: 10.1177/0733464815609440] [Medline: 26491029]
53. Chen CC, Lan YL, Chiou SL, Lin YC. The effect of emotional labor on the physical and mental health of health professionals: emotional exhaustion has a mediating effect. Healthcare (Basel) 2022;11(1):104 [FREE Full text] [doi: 10.3390/healthcare11010104] [Medline: 36611564]
54. Sever M, Tiryaki A. Emotional labor, conflicting caregiving responsibilities and resilience among foreign female caregivers in Japan: a photovoice study. Gen Work Organ 2024;32(3):949-976. [doi: 10.1111/gwao.13195]
55. MacLeavy J. Care work, gender inequality and technological advancement in the age of COVID‐19. Gen Work Organ 2020;28(1):138-154. [doi: 10.1111/gwao.12534]
56. Rosales A, Fernández-Ardèvol M. Ageism in the era of digital platforms. Convergence (Lond) 2020;26(5-6):1074-1087 [FREE Full text] [doi: 10.1177/1354856520930905] [Medline: 33239961]
57. Wegener EK, Bergschöld JM, Whitmore C, Winters M, Kayser L. Involving older people with frailty or impairment in the design process of digital health technologies to enable aging in place: scoping review. JMIR Hum Factors 2023;10:e37785 [FREE Full text] [doi: 10.2196/37785] [Medline: 36705959]
58. Kim S, Trinidad B, Mikesell L, Aakhus M. Improving prognosis communication for patients facing complex medical treatment: a user-centered design approach. Int J Med Inform 2020;141:104147. [doi: 10.1016/j.ijmedinf.2020.104147] [Medline: 32544822]
59. Mannheim I, Schwartz E, Xi W, Buttigieg SC, McDonnell-Naughton M, Wouters EJM, et al. Inclusion of older adults in the research and design of digital technology. Int J Environ Res Public Health 2019;16(19):3718 [FREE Full text] [doi: 10.3390/ijerph16193718] [Medline: 31581632]
60. Abrams R, Vandrevala T, Samsi K, Manthorpe J. The need for flexibility when negotiating professional boundaries in the context of home care, dementia and end of life. Ageing Soc 2018;39(9):1976-1995. [doi: 10.1017/s0144686x18000375]
61. Christianson J, Johnson N, Nelson A, Singh M. Work-related burnout, compassion fatigue, and nurse intention to leave the profession during COVID-19. Nurse Lead 2023;21(2):244-251 [FREE Full text] [doi: 10.1016/j.mnl.2022.06.007] [Medline: 35783544]
62. Margolis H. Report paints picture of care workers in the United States. Center for Retirement Research at Boston College. 2024 Dec 10. URL: https://crr.bc.edu/report-paints-picture-of-care-workers-in-the-united-states/#:~:text=84%20percent%20of%20home%20care,%2B%20(see%20Figure%205) [accessed 2026-02-02]

Abbreviations

ICT: information and communication technology
IRB: institutional review board
mHealth: mobile health

Edited by R Yang; submitted 08.Jul.2025; peer-reviewed by Z Shah, H Heppner, S Fu; comments to author 09.Oct.2025; revised version received 15.Nov.2025; accepted 24.Dec.2025; published 26.Feb.2026.

Please cite as:
Wu Y, Ross N, Kim S
Providing Tech Support as Care Work Among Care Workers in Assisted Living Facilities: Qualitative Interview Study
JMIR Aging 2026;9:e80272
URL: https://aging.jmir.org/2026/1/e80272
doi:10.2196/80272
PMID:

©Yiyi Wu, Nicole Ross, Sunyoung Kim. Originally published in JMIR Aging (https://aging.jmir.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Acceptance and Adoption of the InspireD Reminiscence App Among Care Home Staff: Qualitative Focus Group Study

Aoife Conway[1], MSc; Rosemary Bradley[1], MSc; Assumpta Ryan[1], PhD; Claire McCauley[1], PhD; Brighide Lynch[1], PhD; Deirdre Harkin[1], PhD; Sarah Penney[1], MSc

School of Nursing and Paramedic Science, Faculty of Life & Health Sciences, Ulster University, Derry-Londonderry Campus, Derry, Northern Ireland, United Kingdom

Corresponding Author:
Aoife Conway, MSc
School of Nursing and Paramedic Science, Faculty of Life & Health Sciences
Ulster University, Derry-Londonderry Campus
Northland Road
Derry, Northern Ireland, BT48 7JL
United Kingdom
Phone: 44 2871675151
Email: a.conway@ulster.ac.uk

Abstract

Background: Digital technologies are increasingly used in care homes to support social connection, engagement, and person-centered care for older adults. Digitally facilitated reminiscence is conceptually promising, yet evidence regarding how such interventions are adopted within care home practice remains limited. Existing research has largely focused on feasibility or usability in community settings, with less attention to implementation processes in care homes.

Objective: This study aimed to explore the factors that influence the acceptance and adoption of the InspireD reminiscence app (Ulster University) to support individual reminiscence within everyday care home practice.

Methods: An exploratory qualitative design was used. Twenty-one care home staff, each representing a different care home within a UK region, participated in 3 in-person focus groups. Participants held a range of leadership and nursing roles and were involved in implementing individual reminiscence as part of a leadership program. Data were analyzed using reflexive thematic analysis to generate an inductive, data-led understanding of staff experiences.

Results: Three interrelated themes were identified: (1) enabling possibility—digital infrastructure and organizational readiness for adoption; (2) shaping acceptance—digital confidence, assumptions, and leadership decision-making; and (3) sustaining use—meaningful adoption for residents, staff, and families. Findings highlight how the adoption and sustained use of the InspireD reminiscence app were shaped by organizational constraints and relational care practices. Sustained use of the InspireD reminiscence app depended on how well it was accepted and adopted into everyday care and adapted over time to remain relevant and supportive of meaningful interaction.

Conclusions: Adoption of the InspireD reminiscence app in care homes is not a singular event but an ongoing process shaped by organizational, relational, and individual factors. Meaningful use functions as a key condition for sustainability rather than an outcome in itself. Implementation strategies should address infrastructure, training, leadership support, and relational value to support the sustained use of the InspireD reminiscence app in care practices.

(JMIR Aging 2026;9:e91081) doi:10.2196/91081

KEYWORDS

digital reminiscence; care homes; implementation; technology adoption; person-centered care; apps; digital health

Introduction

Background

Care homes increasingly incorporate digital technologies to support social connection, stimulate engagement, and enhance residents' quality of life [1,2]. Within these settings, digital tools can function as a means of facilitating personalized engagement, strengthening relationships, and supporting identity preservation, particularly for residents who experience communication difficulties, cognitive impairment, or social isolation [3,4]. Interest in using digital technologies as

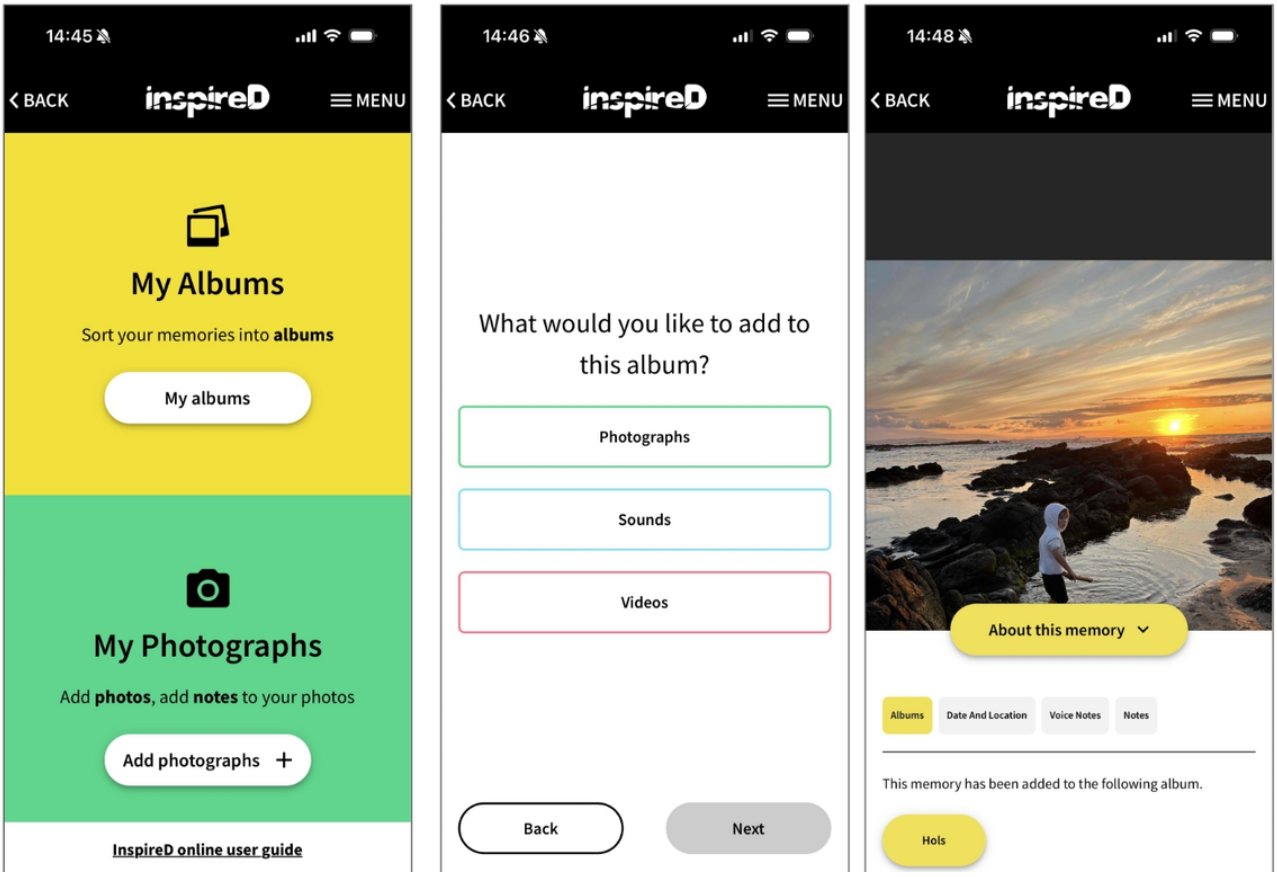
XSL•FO
RenderX

nonpharmacological interventions to enhance the well-being of older adults living in care homes is growing [5-7]. However, evidence indicates that digital practices remain unevenly adopted and differ substantially across services [8-10].

One approach that has gained increasing attention is digitally facilitated individual reminiscence. Reminiscence supports individuals by providing cognitive prompts through images, videos, and music. These cues can evoke past experiences, influence mood, and encourage meaningful conversations with others. Digital reminiscence refers to the use of technology, such as apps, tablets, or touchscreen devices, to support these interactions. The focus is on personalized multimedia content presented in a digital format to support engagement and connection [11]. Previous studies have examined the impact of digital reminiscence [11-15] and, although conceptually promising, robust evidence regarding the implementation of digital reminiscence in care-home contexts remains limited. To date, most research has been conducted with community-dwelling older adults living with dementia, often focusing on feasibility or usability rather than adoption into everyday care. A recent scoping review by Bradley et al [1] examined digital technologies designed to support well-being among people with dementia and family carers, identifying only 2 studies focused on reminiscence, neither of which was conducted in care home settings, which highlights the limited evidence base in this area. This study addresses these gaps by examining the factors that shape the initial adoption and sustained use of the InspireD reminiscence app within everyday care home practice, drawing on the perspectives of care home staff to inform understanding of how digital reminiscence is implemented in care home settings.

This study focuses on the use of the InspireD reminiscence app [16]. The InspireD reminiscence app is a personalized digital reminiscence application originally co-designed with people living with dementia and their carers in partnership with researchers at Ulster University to support the creation and use of individually tailored memorabilia [14]. Its purpose is to support meaningful reminiscence and conversation through personally significant memorabilia. Users can create an individual profile and upload photographs, videos, and sounds or music, organize these materials into albums, and add contextual information, written comments, and voice recordings describing the significance of individual memories (Figure 1). Research on the InspireD reminiscence app has focused on home-dwelling dyads, reporting statistically significant improvements in mutuality, quality of the caregiving relationship, and subjective well-being for people living with dementia, alongside qualitatively reported benefits relating to usability, personal meaning, and relational connection [14,15]. Although initially designed for use in dementia care, the app has since been adapted for broader application within care settings and by family members, friends, and health and social care staff. This shift reflects feedback from users and stakeholders, supporting a dementia-inclusive, rather than dementia-exclusive approach [17]. Within this study, the app was intended to be used collaboratively by residents, staff, and family members, with selected memorabilia acting as prompts for conversation and engagement rather than functioning solely as a digital repository.

Figure 1. InspireD reminiscence app interface.



XSL•FO
RenderX

Aim

To explore the factors that influence the acceptance and adoption of the InspireD reminiscence app to support individual reminiscence within everyday care home practice.

Objectives

The objectives of the study are as follows:

1. To explore how care home staff describe their experiences of introducing and using the InspireD reminiscence app in care home settings.
2. To identify factors that influence the acceptance of the InspireD reminiscence app, from the perspective of care home staff.
3. To examine staff perspectives on factors that support or challenge the adoption of the InspireD reminiscence app.

Methods

Study Design

An exploratory qualitative design was used to examine how the InspireD reminiscence app was adopted and used in care homes (Multimedia Appendix 1). An implementation framework was not applied a priori, as the exploratory aim was to develop an inductive, data-led understanding of staff experiences rather than organize data collection or analysis around predetermined implementation constructs. A reflexive thematic approach [18] was therefore undertaken without coding against an established implementation framework. This allowed for inductive, data-led interpretation of staff experiences without testing a priori theories on technology adoption. Reporting followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidance to enhance completeness and transparency. Ethical approval was obtained from the Ulster University Research Ethics Committee (FCNUR-21-039). All participants provided written informed consent prior to participation. Participants were given an information sheet detailing the purpose of the study, what participation involved, data handling procedures, and their right to withdraw at any time without consequence.

Sampling and Recruitment

A convenience sampling approach was used. Participants were drawn from a preexisting group of staff involved in the My Home Life program [19]; a regional leadership development program delivered across care homes in one UK region. The program targeted staff holding leadership roles within their organizations and typically included senior managers, established leaders, and aspiring leaders seeking to further develop leadership capabilities. As part of the program, participants were involved in a project introducing individualized reminiscence within their care home. The InspireD reminiscence app was offered as one option to support this activity. Participants were not excluded if they chose not to use the app, as understanding reasons for nonadoption or nonuptake was also relevant to the study aims. Twenty-nine staff were involved in implementing the reminiscence activity, of whom 21 provided informed consent to participate in the research study. Of these 21 participants, 19 chose to use the InspireD reminiscence app, while 2 used nondigital approaches to support individual reminiscence.

Intervention and Implementation

The implementation period took place over 4 months and involved the introduction of individualized reminiscence within participating care homes. The wider intervention was not limited to the use of a digital application. Instead, participants were supported to introduce individual reminiscence into routine care using either the InspireD reminiscence app or nondigital approaches, including life story books and memory boxes. The InspireD app was therefore the digital intervention examined within this study, while the workshops, leadership support, peer learning, guidance materials, and provision of tablet devices constituted the implementation strategy used to support its introduction into practice.

Participants who chose to adopt the InspireD app (n=19) were provided with a dedicated Android or Apple tablet device for use within their care home. Ownership of the devices remained with the care homes, which were responsible for setting up the tablet and downloading the app from the Apple App Store or Google Play Store. Photographs, videos, and audio remained stored locally on the device, and app usage data were not accessed or analyzed by the research team. No minimum frequency of app use was prescribed. Instead, participants were encouraged to use reminiscence flexibly in accordance with residents' preferences, local routines, and opportunities for meaningful interaction. App usage was not monitored through backend analytics or usage logs; therefore, information regarding use was derived from participants' accounts during the focus groups.

Implementation of individualized reminiscence was embedded within a structured quality improvement (QI) approach delivered through the wider My Home Life program. participants were supported to introduce, test, and adapt reminiscence approaches within their own care-home contexts over the 4-month implementation period, with opportunities for reflection and peer learning between workshops. A QI approach, informed by Deming [20], the model for improvement [21], and principles of appreciative inquiry [22], guided implementation in practice. To support implementation, participants attended workshops delivered at monthly intervals across the 4-month implementation period (once per month). The workshops were designed to build participants' knowledge of individual reminiscence and support its use in everyday care as well as QI methodologies. Content included principles of individual reminiscence, practical approaches to selecting, curating and using personal memorabilia, an introduction to the InspireD app and available guidance materials, and structured opportunities for participants to reflect with peers on implementation experiences, barriers and emerging learning. Workshops were delivered face-to-face at a location familiar to the participants. Each workshop lasted approximately 7 hours.

The implementation approach was leadership-focused. Participants were not expected to deliver all reminiscence activities themselves; rather, they were encouraged to coordinate implementation within their care home, involve other staff and families where appropriate, and adapt delivery to local routines

XSL•FO
RenderX

and resident preferences. The periods between workshops provided opportunities to trial reminiscence approaches in practice and subsequently discuss experiences, challenges, and adaptations with peers and program facilitators at the following workshop. The number of residents involved and the intensity of app use were determined locally rather than prescribed by the research team. Consequently, exposure was not necessarily uniform across participating homes, reflecting the pragmatic implementation of the intervention within routine care-home practice. This variation was relevant to the study aim, which sought to understand factors influencing adoption and nonadoption under real-world conditions.

Data Collection

Data collection took place following completion of the 4-month implementation period. Three in-person focus groups were conducted in June 2025, after participants had had the opportunity to introduce and use individual reminiscence approaches within their respective care homes. Each focus group comprised 6 to 8 participants and lasted between 60 and 90 minutes. A semistructured interview guide was developed based on the study aims and objectives, relevant literature on reminiscence and implementation in care home settings, and the content of the My Home Life program. Draft questions were reviewed and discussed within the research team to assess their relevance, clarity, and alignment with the study objectives, with amendments made prior to data collection (Multimedia Appendix 2).

All focus groups were audio-recorded and transcribed verbatim by an experienced transcriptionist, with transcripts checked against recordings for accuracy and anonymized prior to analysis. Written informed consent was obtained from all participants, including consent for the use of anonymized quotations. All potential participants were offered the opportunity to take part in the workshops and were free to opt out of any data collection. While all researchers had prior involvement in either the My Home Life program or previous work with the InspireD app, providing valuable contextual insight, reflexive practices were used to critically examine and mitigate potential sources of bias. Rigor was further enhanced through the maintenance of reflexive diaries to document assumptions and analytic decisions throughout the study. Peer validation was undertaken through team discussion of all 3 transcripts to identify potential blind spots, consider alternative interpretations, and support the credibility of the findings and themes.

Data Analysis

Analysis followed Braun and Clarke's [18] reflexive thematic analysis. Familiarization began with repeated reading of the transcripts and review of the audio recordings where necessary to support immersion in the data. Initial coding was undertaken inductively, with codes developed from participants' accounts rather than from a predetermined coding framework. Initial codes were subsequently reviewed across the dataset and brought together into broader patterns of shared meaning (by AC, RB, and CM). These patterns informed the development of initial themes, which were discussed and refined through iterative engagement with the coded data and full transcripts. All themes were examined for internal coherence, distinction from one another, and their ability to capture meaningful patterns relevant to the study aim. This process involved revisiting coded extracts, combining or separating themes where appropriate, and refining theme names and definitions until the final thematic structure was agreed.

Reflexive discussion was embedded throughout the analytic process. Members of the research team had prior involvement in facilitating the wider program and in previous work evaluating the InspireD reminiscence app. This positioning provided contextual insight but also had the potential to shape interpretation. Researchers therefore maintained reflexive diaries to document assumptions, expectations and emerging interpretations, and these reflections were considered during analytic discussions. Team discussions were used not to seek coding consensus, but to challenge assumptions, identify potential blind spots, consider alternative interpretations and examine how researchers' prior knowledge and positioning may have influenced theme development. This reflexive process supported a more critically considered interpretation of the data while remaining consistent with the principles of reflexive thematic analysis.

Results

Overview

Twenty-one care home staff participated in the study. Participants held a range of leadership and nursing roles and brought varied levels of experience to the project. Table 1 outlines participant characteristics. Participants occupied roles that enabled their participation in the program, including managerial, leadership, or senior clinical positions. Consequently, the findings primarily reflect managerial, leadership, and senior clinical perspectives on implementation rather than the direct experiences of frontline care staff. The participants worked across 21 care homes located within one UK region. These homes differed in size, accommodating between 15 and 76 residents, and included both rural and urban settings as well as a mix of nursing and residential provision for people with diverse care needs. Of the 21 participants, 2 used life story books to support individual reminiscence, while the remainder opted to use the InspireD reminiscence app demonstrating initial interest.

XSL•FO
RenderX

Table 1. Participant characteristics.

Categories	Participants, n
Roles	
Care home manager	11
Acting or deputy manager	3
Senior nursing staff (charge nurse, ward sister, and senior staff nurse)	3
Staff nurse	4
Experience (years)	
1-3	2
3-5	1
5-10	5
11-15	4
16-20	3
More than 20	6

Three interrelated themes were identified. The first theme, "enabling possibility—digital infrastructure and organizational readiness for adoption," captures the structural and contextual conditions that shaped early uptake. The second theme, "shaping acceptance—digital confidence, assumptions, and leadership decision-making," reflects how staff perceptions, confidence, and leadership decision-making influenced whether and how the app was introduced. The third theme, "sustaining use—meaningful adoption for residents, staff, and families," explores the conditions that supported or constrained sustained meaningful use and the embedding of the InspireD reminiscence app.

Theme 1: Enabling Possibility—Digital Infrastructure and Organizational Readiness for Adoption

This theme highlights factors that influenced the possibility of adoption of the InspireD reminiscence app, particularly in relation to care home digital infrastructure and organizational readiness. While most participants described minimal technological infrastructure, a few had greater exposure to digital technology and apps, mostly due to changes initiated during the COVID-19 pandemic. Participants recalled that necessity had accelerated digital literacy, particularly through video calls and online communication with families:

> *During Covid we had no face-to-face communication, so we had to learn how to use Teams and video calls. Now it's the norm that families use iPads, saying, "Will you turn that on so I can say goodnight to mummy?" It's become a lot more virtual.* [Participant 4]

> *I'm just thinking back through Covid, when visiting was restricted, we were using tablets and iPads for residents to connect with families. That period was all new to us, to residents living within a care home, and to staff working within a care home, to have to use that technology.* [Participant 7]

This shift meant that, for some homes, digital communication had become part of the routine of care, making the introduction of a reminiscence app less daunting. However, even within these more digitally experienced homes, app use remained largely informal and unstructured, potentially reflecting limited awareness of care-specific or health-related applications:

> *Some of our residents do have iPads, and their families would have brought them in, but it wouldn't be anything structured.* [Participant 9]

> *Sometimes they might use them but it's not through a specific app… I'm not aware of any specific health related apps as such.* [Participant 3]

In other care homes, participants described how the number of devices was limited, with some homes relying on a single shared tablet (provided through this project), and unstable internet connectivity:

> *And technology that doesn't always work.* [Participant 6]

> *It's very hard to share iPads. Technology and internet access are always barriers.* [Participant 11]

> *Currently at the minute we've got one tablet that is dedicated for doing it [using the InspireD reminiscence app] with the one person.* [Participant 8]

Participants anticipated that as future generations enter care settings, residents will be more accustomed to using technology, indicating that access to personal devices is likely to become more common over time:

> *Talking about the generation gap, the younger generation are already used to iPads and picture frames. In another ten years, when they are older, they'll still be familiar with them.* [Participant 17]

> *See, thinking longer term, a lot of people are coming in with their own smart devices, all you have to do is download the app.* [Participant 5]

Participants also described infrastructural assumptions that shaped adoption. The digital infrastructure assumes that memorabilia are already digitized, generating additional tasks

XSL•FO
RenderX

when hard copy materials or materials in a noncompatible format must be converted. This increased task complexity and created additional "hidden" workload, requiring workarounds such as photographing printed photographs rather than uploading files directly:

> *I was struggling to try to upload stuff and navigate my way around it so, you know, even that was time consuming but it was near frustrating when you had to touch this to get into that to do that… you were taking a picture of a picture to get it on as opposed to being able to just upload photographs.* [Participant 12]

Device constraints also added to this, with some iPads lacking ports for external drives, limiting the ability to transfer files from memory sticks:

> *[W]e did have an issue at the beginning was getting the photos uploaded onto the iPad, she… put them on a pen drive and she JPEG'd them all …that was beyond me… So I had to get someone else involved in the care home to transfer them.* [Participant 9]

These examples demonstrate how digital systems that assume predigitized memorabilia create additional work, revealing gaps in readiness for digital approaches. For these participants, the fragility of digital infrastructure, including unstable connectivity and limited device access, constrained the possibility of adoption. This highlights that adequate digital infrastructure is required to make app use workable in day-to-day care.

Theme 2: Shaping Acceptance—Digital Confidence, Assumptions, and Leadership Decision-Making

This theme explores how digital confidence, assumptions, and leadership decision-making shaped acceptance of apps and influenced whether use of the app was taken forward beyond initial introduction. Participants who felt unsure of their own ability to use digital tools were hesitant to delegate or promote their use, which constrained implementation from the outset:

> *I didn't use the Inspire App, I used a scrapbook… plus I don't know technology very well.* [Participant 21]

Perceptions of staff capability also shaped decision-making, with some participants questioning whether their teams were sufficiently prepared or "ready" to use the InspireD reminiscence app, showing how assumptions about competence can influence early adoption behaviors:

> *I never took the app at all in. I just knew my staff would be on the same page as me with technology. No [shakes head and gestures refusal with hands]* [Participant 16]

However, this was not universal. Some managers embraced the challenge and modelled a positive learning stance and, for some, this process of shared learning helped overcome apprehension:

> *Technology wasn't a thing really that I love either, but it was a challenge and through this whole experience and all too you have to challenge yourself. So, I did persevere with it and had some help from my children and all too.* [Participant 4]

> *It probably helped that I'm not technological at all, so they [staff] actually helped me a lot more with it, which made it easier for everyone.* [Participant 9]

Similar reservations were expressed about residents' ability to engage with digital applications, often reflecting preconceived assumptions about cognitive capacity, determining whether the app was taken forward at all:

> *Because I didn't feel that the residents before we even started would be able to cope with an iPad* [Participant 15]

> *And to be quite honest, like myself I wouldn't be very IT literate so it's very hard for me to navigate never mind a dementia resident.* [Participant 13]

Actual differences in digital literacy were a recurring theme. Some staff were apprehensive, particularly when they doubted their own technical skill or worried about doing something wrong, which constrained early adoption:

> *The activity coordinator was not too keen either, she thought she wouldn't be fit for it even after explanation.* [Participant 2]

> *Maybe a bit afraid of it or afraid of doing something wrong…and I think that's something that we did discuss before whenever it comes to new things some members of staff can be apprehensive about it.* [Participant 10]

> *Some staff have accepted it and some have rejected it because of the challenge.* [Participant 7]

In contrast, staff who were already confident with digital applications were more likely to adopt it positively:

> *It was about getting the staff members who were confident in utilising it.* [Participant 20]

> *Now staff members did utilise it and were positive with it, but part of that was in fact that the staff members who were engaging were more confident with technology.* [Participant 13]

Confidence was often associated with age and experience, with younger staff commonly perceived as more digitally comfortable and willing to experiment with new approaches. Participants contrasted this with older colleagues who were sometimes viewed as less confident with technology:

> *The activities girls and myself were all happy enough with the technology side of it. The activities girls are in their early twenties and were quite savvy with it and stuff and we're sort of thinking outside the box.* [Participant 18]

> *She's an older lady, she wouldn't have much with computers and stuff, you know.* [Participant 4]

However, this was not always the case; other participants argued that digital literacy varied within and across age groups; chronological age alone did not predict digital capability:

> *I have got more experienced members of staff who are more tech savvy than some of my school leavers, just because they're young doesn't necessarily mean*

XSL•FO
RenderX

> *they have any real understanding how to do anything much more than make a TikTok video.* [Participant 6]

For staff who felt reasonably comfortable with technology yet lacked confidence in initiating conversations, the app was viewed as a helpful support. They noted that it could enable less experienced staff to begin meaningful, often intergenerational, interactions:

> *It's a good tool because if everything is there it doesn't really take much from the care staff say, if they're a bit apprehensive to initiate the conversations.* [Participant 9]

> *I suppose the app, even those younger staff that don't have the experience, the app giving them tools to be able to initiate that conversation with the individual, and to do that reminiscence* [Participant 14]

Alongside confidence in their own digital ability, some participants also expressed cautious views and uncertainty on the use of digital devices in care home settings, indicating reduced confidence in their value and appropriateness:

> *I can see that yes, it's advancing, and maybe we have to move along with the times, but is it the best? I don't know.* [Participant 7]

Although participants were provided with an introduction to the app and were signposted to guidance materials available via the InspireD webpage, they consistently emphasized the importance of training and clear guidance in building staff confidence for sustained implementation. Participants highlighted that capacity building required more than ad hoc support, noting the need for simple and repeatable guidance to support confidence in using the tool, maintain engagement over time, and manage workload by reducing uncertainty and troubleshooting demands:

> *On the day, X was explaining each section but we didn't have our own iPads set up, so we were just watching and trying to follow. When we went home, we had to figure it out ourselves. A step-by-step guide would have been really helpful.* [Participant 12]

Participants recommended periodic refreshers and regular, brief, targeted learning opportunities, alongside clear step-by-step instructions, to support confidence and consistency in using the app:

> *Updates and refreshers would help. Say every three months, add new features and ideas, otherwise it will never move on.* [Participant 3]

> *More training or clearer step-by-step instructions would help.* [Participant 11]

Theme 3: Sustaining Use—Meaningful Adoption for Residents, Staff, and Families

This theme recognizes that for the InspireD reminiscence app to be adopted and sustained in practice, it must hold meaning for staff, residents, and families. It is clear that continued use is not simply about persistence or frequency, but about perceived value, relational benefit, and fit with care practices. This theme examines the factors that enabled the app to be experienced as meaningful by all involved.

Residents

From the perspective of participants, meaningful use for residents depended on how well the content supported emotional comfort, personal relevance, and interaction. Participants noted that the InspireD reminiscence app became more effective when the material was purposeful rather than excessive. The emphasis was therefore on quality rather than quantity of memorabilia, with carefully selected items seen as more effective. Overpopulating the app with excessive memorabilia could become counterproductive and overcomplicate its use and impact the adoption of the app:

> *I think at the start we maybe tried to put too much in – more photos, more things. So, I think if we'd have stripped it back to more basic it would have been easier.* [Participant 7]

Meaningful use was strengthened by the ability to continually personalize and adapt content. Participants described how ongoing editing allowed staff and families to respond sensitively to residents' emotional reactions, maintaining relevance and protecting well-being. This capacity to update content emerged as a key factor influencing sustained use:

> *Her daughter had left a wee voice note giving me feedback on the app and she said she's actually noticed now that her mum doesn't cope so well with the photos of her mum on the app, so I said to her if you want to take them off, she said there's other ones she absolutely thrives on but she notices that it makes her look for her mum a bit more sad than normal.…. And you can edit it... You can work with it and choose what you want to include.* [Participant 15]

This was essential for participants as residents' fluctuating health, mood and sensory abilities also influenced their capacity to engage meaningfully, indicating that adoption was highly context dependent. Physical impairments and acute illness could temporarily limit interaction, reinforcing the need for flexibility:

> *He has lost his sight now so he was struggling to see the photographs.* [Participant 2]

> *And the resident had an infection. That was a really hard period, about two weeks, and it was quite hard to interact with her, she was just not herself.* [Participant 2]

Residents varied in their engagement with the technology itself. For some, the novelty and ownership of the device enhanced meaningful use and enriched the interaction, while others were less inclined to engage. These diverse responses highlight that adoption was shaped not only by the tool but by individual preference:

> *The resident that we used the app on, she was all biz because she had a piece of technology that was hers and she really enjoyed the technology side of it.* [Participant 11]

Participants consistently emphasized that meaningful use depended on the relational interaction facilitated by the app rather than the digital tool alone. Overemphasis on the

XSL•FO
RenderX

technology risked undermining the core purpose of reminiscence:

> *It's more about the engagement and the way you do it, rather than taking the app.* [Participant 4]
>
> *I found if you overthink it or put too much pressure on the app, it becomes more about the technology than the person.* [Participant 18]

Staff

For staff, meaningful use was dependent on usability and feasibility within daily practice. Where the app was perceived as intuitive and compatible with workload, participants described greater capacity to sustain its use within everyday care practices. Usability and design were frequently highlighted as positive features of the app and a precondition for adoption. Although some staff felt uncertain at the outset, when the interface felt manageable and intuitive perceived usability increased and adoption became more likely:

> *Initially I thought, "I don't know how I'm going to do this." I was thinking it was going to be really complicated and a lot of work on top of everything else. But the app is so user friendly, you can make it as complicated or as simple as you want.* [Participant 9]
>
> *It was great, it was easy to use and it was handy.* [Participant 13]

Participants described how embedding the InspireD reminiscence app within routine care practices created the conditions for meaningful use, by making residents' personal histories more accessible and preventing important stories from being lost. This was perceived as a major factor affecting long-term adoption:

> *I think the app and you can see everything visually as well, it's better than maybe reading it on a care plan or a life story. Like we have the life stories, especially when you've got residents who have been there, you get to know them over the time but then as staff change maybe the life stories don't come out again because they're in their pile, whereas this way if you had the profiles it would sort of keep it real whenever staff can use it more.* [Participant 6]

Family

For families, meaningful use was facilitated by both contributing to the digital content and participating in reminiscence interactions. Their engagement played a crucial role in sustaining adoption, with relatives often recognizing the app as a valuable way to maintain connection and contribute to care. Families were seen as essential partners in sustaining app use, both as contributors of personal material and as co-participants in app use. Several participants noted that relatives valued the presence of technology in the care home and appreciated opportunities to engage in structured digital activities:

> *My family was the same, they found it very interesting, especially that it was like technology in the care home, they never experienced the care home before, they were living thinking once her mum is living in the care home there will be no interaction with her, they found it very difficult at the start to put her in a residential home depending on the case and the family were amazed like wow, we have something to do. And it was giving them a purpose as well.* [Participant 20]

In some cases, relatives became active users of the InspireD reminiscence app, taking the device to be with their loved one and initiate reminiscence interactions:

> *And I suppose from the app perspective, the daughter comes and takes the iPad and goes and sits with her mum and does it a lot of the time.* [Participant 1]

However, meaningful use depended on family members feeling confident and able to engage with the technology. Limited digital familiarity constrained participation and reduced the extent to which families could contribute to sustaining the intervention:

> *If the family could, it would help, but if they're not confident with putting things into the app, that's a challenge.* [Participant 14]
>
> *Her next of kin is her husband who is not too keen as well, afraid to touch the iPad.* [Participant 8]

Family readiness therefore functioned as an enabling or constraining factor in how fully the InspireD reminiscence app could be embedded into everyday routines.

The InspireD reminiscence app also created meaningful opportunities for intergenerational connection and learning. Younger family members often used the technology to initiate and guide conversations, which was particularly valuable when spontaneous interaction felt difficult. For families, the app created a shared point of focus that made visits feel more engaging and emotionally rewarding, encouraged participation across generations and helped them feel more involved in their relative's care:

> *Children or grandchildren sometimes aren't able to have them conversations, so I think from that point of view the app is beneficial as a tool to start the conversations.* [Participant 3]
>
> *And also we were saying for some family members, it's also a tool they can use because people are not really wanting, families come in and sometimes they find this hard, especially if they don't have that two-way conversation, whereas if that's there and they're able to use it I think the app would definitely benefit more than somebody coming in and starting, right let's do some reminiscence.* [Participant 16]

The findings illustrate how a range of organizational, relational, and individual factors shaped the initial adoption of the InspireD reminiscence app and influenced its sustained use within care home practice. Across themes, participants highlighted conditions that either supported or challenged sustained use, demonstrating that adoption was not a singular event, but an ongoing process shaped by context, confidence, and perceived meaning.

XSL•FO
RenderX

Discussion

Principal Findings

This study explored how the InspireD reminiscence app was adopted within care home settings. The findings illustrate the complexity of adoption in these contexts and suggest that it extends beyond commonly emphasized considerations such as access to devices or individual preferences [9,23].

Consistent with earlier literature, participants described increased familiarity with digital technologies following the COVID-19 pandemic, reflecting wider changes reported across care home settings [8,10]. However, the findings indicate that increased exposure to technology and device availability did not necessarily translate into sustained confidence, readiness, or capacity to use digital health apps within everyday practice. Global reports, including the World Health Organization (WHO) dementia research blueprint [24], similarly highlight that despite the potential of digital innovation, progress continues to be constrained by disparities in digital literacy, access to devices, and underlying infrastructure. This study supports those concerns and further shows that the digital divide extends beyond system-level inequity to include variability within individual care home environments. These findings highlight that future considerations with regard to implementation require reliable internet connectivity, sufficient access to suitable devices, straightforward processes for uploading and transferring digital memorabilia, and accessible technical support to reduce implementation burden.

The findings demonstrate that infrastructure alone does not explain digital adoption. Instead, participants' accounts suggested that meaningful use was an important factor shaping whether the InspireD reminiscence app was taken forward in practice. Previous studies have highlighted that digital tools must offer perceived benefit to end users, in this case residents, if they are to be adopted [17]. This study extends this work by suggesting that these conditions were not confined to residents alone but were also shaped by staff and families. In this way, participants' experiences indicated that adoption was shaped by relational and organizational contexts rather than individual perceptions alone, an aspect of technology adoption that remains underexplored in the literature. As Shachak et al [23] reported, technology acceptance models often focus on individual intentions and beliefs, underplaying the complexity of multiuser environments. This framing aligns with broader critiques of technology acceptance models that prioritize perceived usefulness, ease of use, and behavioral intention, but give limited attention to the relational and organizational contexts in which technologies are enacted [25,26]. In care home settings, participants' experiences suggested that sustainability appeared to depend less on the presence of technology itself and more on how it enabled connection, preserved personhood, and fitted within existing practices [27]. Participants' experiences suggested that meaningful use functioned not simply as an outcome of engagement with the InspireD reminiscence app, but as an important condition supporting its continued adoption. This interpretation suggests that the conditions enabling meaningful use may be foundational to adoption rather than representing an endpoint in themselves. While much of the existing literature focuses on whether digital interventions are adopted or accepted, this study extends this perspective by highlighting the conditions through which meaningful use may support adoption.

Digital capability and confidence, whether actual or perceived, also influenced whether the app was adopted, echoing findings reported by Beattie et al [28]. This study adds that assumptions about the capability of staff and/or residents can act as an unrecognized barrier, with decisions about adoption sometimes made before individuals have the opportunity to engage directly with the technology. These findings resonate with wider literature on gatekeeping in dementia research, which demonstrates that assumptions about capability and suitability can shape access and opportunity before individuals are given the chance to engage. Studies have shown that gatekeepers, including managers and staff, may unintentionally exclude people living with dementia from participation based on preconceived notions of vulnerability or unsuitability [29-31]. Prior work suggests that such dynamics can be mitigated through relationship building, trust, and the deliberate challenging of deficit-based assumptions [32,33]. Consistent with this, the present findings suggest that leadership perceptions and behaviors warrant further examination, particularly in how they shape access to, and engagement with, digital interventions in care home settings.

The findings point to the need for sustained workforce investment, including confidence building and digital skills development, echoing wider literature on digital innovation in long-term care [7]. This study indicates that training is required not only for staff introducing the intervention, but also for those supporting its routine use. The findings suggest that implementation of digital apps should consider extending training beyond initial instruction to include brief refreshers, accessible step-by-step guidance, and opportunities for staff to revisit implementation challenges over time. Consideration should also be given to protected time for staff to develop and update personalized content, and to ways of incorporating digital reminiscence flexibly into existing activities, family visits, care planning, or life-story work, rather than positioning it as an additional standalone task. The findings also reinforce the need for training and workload considerations to be co-produced with staff to ensure alignment with everyday routines and capacities. Such approaches may not only support the implementation of the InspireD reminiscence app but also contribute to addressing broader challenges identified in this study and across the literature, including workload pressures [34], the emotional demands of reminiscence [35], and the importance of enabling meaningful family involvement [36].

Strengths and Limitations

This study offers one of the first examinations of the implementation of a digital reminiscence application within care home settings and has several strengths including steps taken to mitigate bias. Credibility was supported through the recruitment of participants across multiple organizations, prolonged engagement over a 4-month implementation period, and the use of rich, illustrative quotations to support analytic

XSL•FO
RenderX

claims. Dependability was enhanced through the maintenance of an auditable record of coding decisions and theme development, alongside regular discussion and sense-checking within the research team and a robust reflexive approach to data analysis. Transferability was strengthened through detailed description of the study context and participant characteristics, as well as the inclusion of a relatively large number of participating care homes. Several limitations are acknowledged. The study used a convenience sample drawn from care homes participating in a leadership and QI program, which may mean that participants were more motivated or open to innovation than staff in other settings. The sample predominantly comprised managers, senior leaders, and clinical staff, with frontline care staff not directly represented in the focus groups. Their perspectives may differ, particularly in relation to workload, usability, and the feasibility of adopting the app into routine day-to-day care. The nature of the My Home Life program and the structured implementation support provided by expert facilitators may also have influenced participants' experiences of introducing the InspireD reminiscence app. Within this supportive context, participants may have been more open to innovation and reflective practice, and may have had access to resources, peer learning, and facilitation that would not necessarily be available in care homes implementing similar technologies independently. Implementation experiences may therefore differ in settings without comparable organizational support, which should be considered when assessing the transferability of the findings. Data were generated from staff perspectives only, without direct input from residents or family members, and without observational data on actual use, which limits insight into how the InspireD reminiscence app is experienced in practice by other stakeholders. Future research should directly include residents' perspectives to explore how they experience the app, what makes its use meaningful or acceptable from their perspective, and how individual preferences, abilities, and responses shape adoption and continued use within everyday care. The study focused on the InspireD reminiscence app within a single UK region, which may constrain transferability to other technologies or health and social care systems. Finally, members of the research team were involved in both the wider program and previous evaluations of the InspireD app, which may have influenced interpretation despite efforts to address this through reflexive practice.

Conclusions

This study examined how the InspireD reminiscence app is adopted within care homes. The findings demonstrate that adoption cannot be understood solely in terms of device availability or individual preference. Instead, it is shaped by the interaction between baseline digital infrastructure, staff confidence and assumptions, and the extent to which the app supports meaningful engagement for residents, staff and families. Planning and implementation need to address infrastructure, training and leadership support alongside coproduced strategies that ensure the technology enhances person-centered relationships.

Acknowledgments

The authors wish to thank the participants who generously shared their time and experiences in this study.

Funding

The My Home Life Program is supported by the Department of Health, Castle Buildings, Stormont, Belfast, Northern Ireland. The funders had no role in the design and conduct of this work.

Data Availability

The datasets generated and analyzed during this study are not publicly available to protect participant confidentiality but are available from the corresponding author on reasonable request.

Generative AI Use

The authors declare that generative AI was not used in the preparation of this manuscript.

Authors' Contributions

Conceptualization: AC, RB, AR, DH, SP
Formal analysis: AC, RB, CM
Investigation: AC, RB, CM, BL
Methodology: AC, RB, CM, BL, DH
Project administration: AC, RB
Supervision: AC, AR
Validation: AR
Writing – original draft: AC
Writing – review & editing: AC, RB, AR, CM, BL, DH, SP

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Visual overview of the study.
[DOCX File , 28 KB - aging_v9i1e91081_app1.docx]

Multimedia Appendix 2
Focus group guide.
[DOCX File , 23 KB - aging_v9i1e91081_app2.docx]

References

1. Bradley L, Shanker S, Murphy J, Fenge L, Heward M. Effectiveness of digital technologies to engage and support the wellbeing of people with dementia and family carers at home and in care homes: a scoping review. Dementia (London) 2023 Aug;22(6):1292-1313 [FREE Full text] [doi: 10.1177/14713012231178445] [Medline: 37235791]
2. Chu CH, Ronquillo C, Khan S, Hung L, Boscart V. Technology recommendations to support person-centered care in long-term care homes during the COVID-19 pandemic and beyond. J Aging Soc Policy 2021;33(4-5):539-554. [doi: 10.1080/08959420.2021.1927620] [Medline: 34278980]
3. Astell AJ, Bouranis N, Hoey J, Lindauer A, Mihailidis A, Nugent C, Technology Dementia Professional Interest Area. Technology and dementia: the future is now. Dement Geriatr Cogn Disord 2019;47(3):131-139 [FREE Full text] [doi: 10.1159/000497800] [Medline: 31247624]
4. Subramaniam P, Woods B. Digital life storybooks for people with dementia living in care homes: an evaluation. Clin Interv Aging 2016;11:1263-1276 [FREE Full text] [doi: 10.2147/CIA.S111097] [Medline: 27698556]
5. Hung L, Liu C, Woldum E, Au-Yeung A, Berndt A, Wallsworth C, et al. The benefits of and barriers to using a social robot PARO in care settings: a scoping review. BMC Geriatr 2019;19(1):232 [FREE Full text] [doi: 10.1186/s12877-019-1244-6] [Medline: 31443636]
6. Wong KLY, Gregorio M, Mann J, Berndt A, Wong L, Wang C, et al. Using videos and films with people with major cognitive disorder living in care settings: a scoping review. Gerontechnology 2023;22(1):1-13. [doi: 10.4017/gt.2023.22.1.823.04]
7. Ren LH, Wong KLY, Wong J, Kleiss S, Berndt A, Mann J, et al. Working with a robot in hospital and long-term care homes: staff experience. BMC Nurs 2024 May 08;23(1):317 [FREE Full text] [doi: 10.1186/s12912-024-01983-0] [Medline: 38720346]
8. Seifert A, Cotten SR. Digital distance in times of physical distancing: ICT infrastructure and use in long-term care facilities. American Behavioral Scientist 2023;69(9):1138-1154. [doi: 10.1177/00027642231155361]
9. Johnston L, Koikkalainen H, Anderson L, Lapok P, Lawson A, Shenkin SD. Foundation level barriers to the widespread adoption of digital solutions by care homes: insights from three Scottish studies. Int J Environ Res Public Health 2022;19(12):7407 [FREE Full text] [doi: 10.3390/ijerph19127407] [Medline: 35742667]
10. Warmoth K, Lynch J, Darlington N, Bunn F, Goodman C. Using video consultation technology between care homes and health and social care professionals: a scoping review and interview study during COVID-19 pandemic. Age and Ageing 2022;51(2):afab279. [doi: 10.1093/ageing/afab279]
11. Boyd K, Bond R, Ryan A, Goode D, Mulvenna M. Digital reminiscence app co-created by people living with dementia and carers: usability and eye gaze analysis. Health Expect 2021;24(4):1207-1219 [FREE Full text] [doi: 10.1111/hex.13251] [Medline: 34128574]
12. Elfrink TR, Ullrich C, Kunz M, Zuidema SU, Westerhof GJ. The online life story book: a randomized controlled trial on the effects of a digital reminiscence intervention for people with (very) mild dementia and their informal caregivers. PLoS One 2021;16(9):e0256251 [FREE Full text] [doi: 10.1371/journal.pone.0256251] [Medline: 34525105]
13. Xu L, Fields NL, Highfill MC, Troutman BA. Remembering the past with today's technology: a scoping review of reminiscence-based digital storytelling with older adults. Behav Sci (Basel) 2023;13(12):998 [FREE Full text] [doi: 10.3390/bs13120998] [Medline: 38131854]
14. Ryan AA, McCauley CO, Laird EA, Gibson A, Mulvenna MD, Bond R, et al. 'There is still so much inside': the impact of personalised reminiscence, facilitated by a tablet device, on people living with mild to moderate dementia and their family carers. Dementia (London) 2020;19(4):1131-1150. [doi: 10.1177/1471301218795242] [Medline: 30193535]
15. Laird EA, Ryan A, McCauley C, Bond RB, Mulvenna MD, Curran KJ, et al. Using mobile technology to provide personalized reminiscence for people living with dementia and their carers: appraisal of outcomes from a quasi-experimental study. JMIR Ment Health 2018;5(3):e57 [FREE Full text] [doi: 10.2196/mental.9684] [Medline: 30206053]
16. The inspireD app. InspireD, Ulster University. 2026. URL: https://www.theinspiredapp.com [accessed 2026-01-08]

XSL•FO
RenderX

17. Conway A, Ryan A, Harkin D, Mc Cauley C. "It's Another Feather in My Hat"-exploring factors influencing the adoption of apps with people living with dementia. Dementia (London) 2023;22(7):1487-1513 [FREE Full text] [doi: 10.1177/14713012231185283] [Medline: 37365816]
18. Braun V, Clarke V. Thematic Analysis: A Practical Guide. London: Sage Publications; 2021.
19. My home life NI. My Home Life Northern Ireland. 2026. URL: https://www.myhomelifeni.co.uk [accessed 2026-01-08]
20. Deming WE. Out of the Crisis. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study; 1986.
21. Langley GL, Moen R, Nolan KM, Nolan TW, Norman CL, Provost LP. The Improvement Guide: A Practical Approach to Enhancing Organizational Performance (2nd edition). San Francisco: Jossey-Bass Publishers; 2009.
22. Reed J. Appreciative inquiry. In: Research for Change. London: Sage; 2007.
23. Shachak A, Kuziemsky C, Petersen C. Beyond TAM and UTAUT: future directions for HIT implementation research. J Biomed Inform 2019;100:103315 [FREE Full text] [doi: 10.1016/j.jbi.2019.103315] [Medline: 31629923]
24. A Blueprint for Dementia Research. Geneva: World Health Organization; 2022.
25. Ajibade P. Technology acceptance model limitations and criticisms: exploring the practical applications and use in technology-related studies, mixed-method, and qualitative researches. Libr Philos Pract 2018:1941 [FREE Full text]
26. Nielsen P, Sahay S. A critical review of the role of technology and context in digital health research. Digit Health 2022;8:20552076221109554 [FREE Full text] [doi: 10.1177/20552076221109554] [Medline: 35769359]
27. McCormack B. Developing Person-centred Cultures in Healthcare Education and Practice: An Essential Guide. New Jersey: John Wiley & Sons; 2024.
28. Beattie M, Muirhead K, Carey N. Exploring care home workers' perceptions of implementing ECHO: a qualitative study. Int J Older People Nurs 2025;20(2):e70017. [doi: 10.1111/opn.70017] [Medline: 39945578]
29. Ries NM, Mansfield E, Sanson-Fisher R. Ethical and legal aspects of research involving older people with cognitive impairment: a survey of dementia researchers in Australia. Int J Law Psychiatry 2020;68:101534. [doi: 10.1016/j.ijlp.2019.101534] [Medline: 32033698]
30. Phenwan T, Sixsmith J, McSwiggan L, Buchanan D. A rapid review of internet mediated research methods with people with dementia: practical, technical and ethical considerations. TQR 2021;26(11):3318-3341. [doi: 10.46743/2160-3715/2021.5089]
31. Götzelmann TG, Strech D, Kahrass H. The full spectrum of ethical issues in dementia research: findings of a systematic qualitative review. BMC Med Ethics 2021;22(1):32 [FREE Full text] [doi: 10.1186/s12910-020-00572-5] [Medline: 33771131]
32. Bryant J, Freund M, Ries N, Garvey G, McGhie A, Zucca A, et al. Volume, scope, and consideration of ethical issues in indigenous cognitive impairment and dementia research: a systematic scoping review of studies published between 2000-2021. Dementia (London) 2022;21(8):2647-2676. [doi: 10.1177/14713012221119594] [Medline: 36054372]
33. Pyer M, Ward A. Developing a dementia friendly approach to consent in dementia research. Aging Ment Health 2024;28(2):294-301 [FREE Full text] [doi: 10.1080/13607863.2023.2264216] [Medline: 37885301]
34. Baumann A, Shaw P, Trotter L, Clinch S, Davies N. Mnemosyne: supporting reminiscence for individuals with dementia in residential care settings. 2024 Presented at: CHI '24: CHI Conference on Human Factors in Computing Systems; May 11-16, 2024; Honolulu HI USA p. 1-17. [doi: 10.1145/3613904.3642783]
35. Midhun CP, Anju P, Mohan AK. Digital storytelling as a medium for continuity, reminiscence, and narrative engagement in later life. IJISRT 2025:894-905. [doi: 10.38124/ijisrt/25nov645]
36. Stasolla F, Di Gioia M, Messina I, Treglia F, Passaro A, Zullo A, et al. Assessing and recovering alzheimer's disease: a comparative analysis of standard neuropsychological approaches and virtual reality interventions with the use of digital storytelling. Front Psychol 2024;15:1406167 [FREE Full text] [doi: 10.3389/fpsyg.2024.1406167] [Medline: 39114597]

Abbreviations

WHO: World Health Organization
COREQ: Consolidated Criteria for Reporting Qualitative Research
QI: quality improvement

XSL•FO
RenderX

Edited by R Bjarnadottir; submitted 15.Jan.2026; peer-reviewed by Z Tasnim, C Nwolise; comments to author 31.Aug.2026; revised version received 11.Sep.2026; accepted 23.Oct.2026; published 23.Nov.2026.

Please cite as:
Conway A, Bradley R, Ryan A, McCauley C, Lynch B, Harkin D, Penney S
Acceptance and Adoption of the InspireD Reminiscence App Among Care Home Staff: Qualitative Focus Group Study
JMIR Aging 2026;9:e91081
URL: https://aging.jmir.org/2026/1/e91081
doi:10.2196/91081
PMID:

©Aoife Conway, Rosemary Bradley, Assumpta Ryan, Claire McCauley, Brighide Lynch, Deirdre Harkin, Sarah Penney. Originally published in JMIR Aging (https://aging.jmir.org), 23.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Large Language Model–Based Identification of Subjective Cognitive Decline in Electronic Health Records Among Older Adults With Normal Cognitive Testing: Retrospective Cohort Study

Liqin Wang[1,2], PhD; Rebecca E Amariglio[2,3], PhD; Sheril Varghese[1,4], BA; Jiazi Tian[5], MS; Diane Seger[6], RPh; Li Zhou[1,2], MD, PhD; Gad A Marshall[2,3], MD

[1]Department of Medicine, Mass General Brigham, 399 Revolution Drive, Somerville, MA, United States

[2]Harvard Medical School, 25 Shattuck Street, Boston, MA, United States

[3]Department of Neurology, Mass General Brigham, Boston, MA, United States

[4]Columbia University Irving Medical Center, New York, NY, United States

[5]Department of Medicine, Massachusetts General Hospital, Boston, MA, United States

[6]Research Information Science & Computing, Mass General Brigham, Boston, MA, United States

Corresponding Author:
Liqin Wang, PhD
Department of Medicine, Mass General Brigham, 399 Revolution Drive, Somerville, MA, United States

Abstract

Background: Subjective cognitive decline (SCD) typically refers to self- or informant-reported decline in cognition despite the absence of objective impairment on standardized testing. Older adults with documented normal cognitive test performance provide a pragmatic anchor cohort for electronic health record (EHR)–based SCD phenotyping. However, cognitive concerns are primarily recorded in unstructured notes and are inconsistently documented, making it unclear how often and in whom concerns are captured in routine care.

Objective: This study aims to operationalize EHR-based SCD phenotyping in an objectively normal-testing cohort using a large language model (LLM)–based natural language processing approach and to examine the frequency and clinical and sociodemographic correlates of cognitive concern documentation.

Methods: We conducted an EHR-based observational study of patients aged 65 years or older with a first normal cognitive test recorded in EHR flowsheets between January 2019 and April 2024 in a large health care system. We developed and iteratively refined a 2-stage LLM-based natural language processing pipeline to identify documented cognitive concerns in unstructured notes during the 12 months prior to the index date, and evaluated performance against manual review. We quantified the frequency of documented concerns within this normal-testing cohort and used multivariable logistic regression to assess associations with sociodemographic factors (age, marital status, insurance, and neighborhood Area Deprivation Index), sequentially adjusting for clinical comorbidities and relevant medications.

Results: Among 15,750 older adults with normal cognitive test scores, 13.8% (n=2175) had at least 1 documented cognitive concern in the prior year, captured in 1.2% (7394/605,177) of notes. On manual validation, the 2-stage pipeline (Med42-v2-8B screening followed by GPT-4o confirmation) achieved a sensitivity of 0.957, positive predictive value of 0.935, specificity of 0.985, and F_1-score of 0.945 for cognitive concern identification. Documentation was more likely in older individuals, those with commercial (vs Medicare) insurance, and those with neurological and psychiatric conditions. In fully adjusted models, Parkinson disease (adjusted odds ratio [aOR] 5.29, 95% CI 3.60 - 7.77), traumatic brain injury (aOR 4.63, 95% CI 3.15 - 6.81), stroke or transient ischemic attack (aOR 3.46, 95% CI 3.15 - 4.18), epilepsy (aOR 2.52, 95% CI 1.83 - 3.49), depression (aOR 1.54, 95% CI 1.36 - 1.75), and excessive alcohol use (aOR 1.48, 95% CI 1.06 - 2.06) were among the strongest correlates of cognitive concern documentation. In contrast, obesity (aOR 0.73, 95% CI 0.65 - 0.82), hyperlipidemia (aOR 0.59, 95% CI 0.51 - 0.67), and residence in more deprived neighborhoods (higher Area Deprivation Index) were associated with lower odds of documented cognitive concerns.

Conclusions: A 2-stage LLM pipeline enabled accurate identification of documented cognitive concerns consistent with SCD among older adults with normal cognitive testing. Documentation was uncommon and selectively captured by clinical and sociodemographic factors, with implications for equity and the validity of EHR-based phenotypes.

(JMIR Aging 2026;9:e93011) doi:10.2196/93011

XSL•FO
RenderX

KEYWORDS

large language model; cognitive concern; subjective cognitive decline; electronic health records; documentation

Introduction

Given emerging therapies for Alzheimer disease (AD), early detection is increasingly critical to ensure interventions are delivered at the most beneficial stage of the disease trajectory. Subjective cognitive decline (SCD) is characterized by self- or partner-reported concerns about memory or other cognitive functions despite normal performance on standardized cognitive tests [1]. SCD represents one of the earliest stages at which cognitive changes become noticeable to patients, family members, or clinicians, often preceding measurable impairment, such as mild cognitive impairment (MCI) [2,3]. Importantly, longitudinal meta-analyses show that SCD is associated with approximately 2-fold higher risk of subsequent dementia and elevated risk of incident MCI, supporting SCD as a clinically meaningful risk marker rather than a universally benign complaint [4,5]. As such, recognizing SCD in clinical settings provides a valuable window of opportunity for early intervention, risk reduction, and enrollment in clinical trials targeting preclinical AD [6].

However, SCD remains underrecognized in routine care and is difficult to identify systematically in electronic health records (EHRs). Unlike MCI or dementia, SCD lacks a standardized diagnostic code, limiting scalable identification for clinical screening and research. In practice, documented cognitive concerns and objective testing are often misaligned: many individuals never undergo formal cognitive testing, while others are tested without any documented concerns [7,8]. Among individuals with normal test scores, the presence vs absence of documented cognitive concerns provides a clinically relevant contrast for examining how early symptoms are recognized and recorded. Characterizing variation in documentation by clinical and sociodemographic characteristics is critical for equitable early detection and for valid EHR-based cognitive aging research. Emerging evidence suggests that recognition, reporting, and clinical documentation of cognitive concerns may vary across socioeconomic, racial or ethnic, and health care access contexts, potentially introducing bias into EHR-based phenotyping and delaying identification of individuals at risk for future cognitive decline [9]. Individuals from disadvantaged communities may also face barriers to early evaluation and specialty memory care, further affecting how cognitive concerns are documented in routine clinical practice [10,11].

Cognitive concerns are typically embedded in free-text clinical notes rather than structured fields, making automated detection challenging, particularly given the wide variation in language describing cognitive concerns among clinicians, patients, and family members [12,13]. Recent advances in large language models (LLMs) present an opportunity to overcome these barriers. LLMs have significantly enhanced the depth, accuracy, and efficiency of natural language processing (NLP) for automated information extraction from EHR data. Unlike traditional rule-based NLP tools that rely on predefined patterns and keywords, LLMs are trained on vast text corpora to learn statistical relationships in language, enabling them to capture context, nuance, and phrasing variation beyond the capabilities of rule-based systems. LLMs can analyze unstructured notes at scale, searching for nuanced mentions of cognitive concerns that might otherwise go unnoticed. Although our prior studies have used deep learning and LLMs to detect broad indicators of cognitive decline in EHRs [14], few studies have focused on SCD, and fewer have compared individuals with and without SCD among those with normal cognitive test performance [12,13]. Moreover, factors that may affect the documentation of subjective concerns, such as comorbidities, medication use, and sociodemographic characteristics, remain incompletely understood.

In the present study, we leveraged LLM and EHR data to systematically identify SCD among individuals with normal cognitive test performance. The aims of this study were to (1) develop and evaluate a scalable LLM-based approach for identifying documented SCD from unstructured clinical notes and (2) compare sociodemographic, clinical, and medication characteristics of individuals with and without documented SCD. We hypothesized that individuals with documented SCD would differ in clinical and sociodemographic profile despite normal cognitive testing.

Methods

Data Sources

We used EHR data from Mass General Brigham (MGB), a large integrated health care system that includes Massachusetts General Hospital, Brigham Women's Hospital, and affiliated community hospitals in Massachusetts. Data were obtained from 2 enterprise EHR data repositories, including the Research Patient Data Registry (RPDR) and the Enterprise Data Warehouse (EDW), spanning from January 1, 2015, to April 10, 2024.

Ethical Considerations

The study was approved by the MGB Institutional Review Board. Cognitive assessment data were collected as part of routine clinical care and obtained retrospectively from the EHR. The study team did not administer any cognitive assessment instruments for research purposes, and licensing information was not available through the EHR repositories.

Study Sample

We defined presumed SCD as the presence of documented cognitive concerns among individuals with normal cognitive test scores. Individuals were initially identified based on cognitive assessments recorded in EHR flowsheets between January 1, 2015, and April 10, 2024. Cognitive assessments included the Mini-Cog, Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), Saint Louis University Mental Status (SLUMS) examination, and the Ascertain Dementia 8-Item Informant Questionnaire (AD8) dementia screening interview. When all MMSE items were documented, but the total score was missing, we calculated the

XSL•FO
RenderX

total by summing item scores. Cognitive assessments without a total score were excluded, including partly administered tests due to time constraints.

To establish a baseline of normal cognition, we defined the index data as each individual's first cognitive assessment occurring between January 1, 2019, and April 10, 2024. The period from January 1, 2015, to December 31, 2018, served as a lookback window to identify prior cognitive assessments. Individuals with any cognitive assessment recorded during this earlier period were excluded to ensure that the index assessment represented the first observed cognitive assessment in the EHR during the study period. We further restricted the cohort to individuals aged ≥65 years at the time of their first assessment to focus on age groups most relevant to AD and related dementias.

We narrowed down the study cohort to include individuals with normal cognitive test scores based on established cutoff criteria: Mini-Cog: 4‐5, MMSE: ≥26, MoCA: ≥26, SLUMS: ≥27, AD8: 0‐1. For the remaining individuals, we extracted all their clinical notes within 1 year prior to the initial cognitive test from MGB's EHR databases. Individuals who did not have notes available in the MGB EHR were excluded. We further excluded individuals who had a prior history of cognitive decline, related diagnosis (eg, MCI or dementia) prior to the index date. Specifically, we excluded individuals with evidence of prior cognitive decline or related diagnoses before the index date. Prior history was identified using structured EHR data extracted from RPDR and EDW, including diagnosis records and problem list entries, based on relevant *International Classification of Diseases* (*ICD*) codes for dementia, MCI, AD, and related cognitive disorders documented before the index cognitive assessment date. In addition, note-derived cognitive concern information (see *LLM-Based Cognitive Concern Extraction* section) was used to identify prior patient-related cognitive decline documentation, including cognitive assessments, MCI, and dementia documentation in the notes. Individuals meeting any of these criteria before the index date were excluded from the final cohort.

Dependent Variable: Cognitive Concerns

The primary dependent variable was the presence of cognitive concerns within one year prior to the first normal cognitive assessment, defined as a binary indicator (yes/no). Given that cognitive concerns in the early stages of cognitive decline were often documented in free-text, we used LLMs to identify and extract cognitive concerns from clinical notes (described in the *LLM-Based Cognitive Concern Extraction* section).

LLM-Based Cognitive Concern Extraction

We used a combination of open-source and commercial LLMs to extract cognitive concerns from clinical notes. Notes were first split into smaller sections of up to 300 words. For each section, the model was prompted to (1) indicate whether a cognitive concern is mentioned (yes/no) and (2) extract the relevant text verbatim (see Table S1 in Multimedia Appendix 1 for prompts). As the open-source model, we selected Med42-v2-8B, a Llama 3-based model fine-tuned on clinical data, which enables efficient processing of large volumes of notes at minimal cost [15]. The prompt of this model was optimized for high sensitivity to exclude cases that clearly lacked cognitive concerns while capturing as many potential positives as possible. As the commercial model, we deployed OpenAI GPT-4o in a Health Insurance Portability and Accountability Act (HIPAA)-compliant Microsoft Azure environment. We adopted a 2-stage pipeline to balance scalability and accuracy. In iterative prompt development, Med42-v2-8B was effective as a high-sensitivity screening model but produced more false positives, including occasional overinterpretation or hallucination of cognitive concerns. More restrictive prompts improved precision but reduced sensitivity and did not fully eliminate these errors. Therefore, Med42-v2-8B was used to screen out notes unlikely to contain cognitive concerns, and GPT-4o was applied only to Med42-v2-8B-positive notes to improve confirmation accuracy while reducing computational cost.

Evaluation of LLMs in Cognitive Concern Extraction

To evaluate cognitive concern identification, we constructed an evaluation set of 500 note chunks using stratified sampling by model-predicted class. Because the prevalence of cognitive concerns in the full corpus was low, we randomly sampled note chunks classified as positive and negative by the Med42-v2-8B model at a 3:1 ratio to enrich positive cases. Two annotators (JT and SV) independently reviewed an initial subset of 100 note chunks to assess interannotator agreement for binary cognitive concern identification using Cohen κ. After substantial agreement was achieved (κ=0.70; overall agreement=90%), the annotators independently reviewed the remaining note chunks, with each annotator reviewing an additional 200 chunks. Discrepancies between annotators were adjudicated by a cognitive neurologist (GAM) to establish the reference standard. In addition, selected discordant cases between LLM predictions and reference labels were reviewed by domain experts to better characterize sources of disagreement and model errors. We then compared the LLM outputs with this reference standard and assessed performance using sensitivity, specificity, positive predictive value (PPV), and F_1 score (the harmonized score of sensitivity and PPV).

Covariates

Individual-level covariates included age (in years), patient-reported sex, race or ethnicity, marital status, education level, language preference, and insurance type. Neighborhood deprivation was assessed using the 2023 Area Deprivation Index (ADI) [16]. For each individual, we derived a 9-digit ZIP code from the residential address and ZIP code, then linked this to the ADI database to obtain the national ADI percentile. Individuals were categorized into three groups based on their national ADI percentile: low (1st-20th percentile), medium (21st-40th percentile), and high (41st-100th percentile).

To account for documentation opportunity, we calculated the number of clinical notes in the year prior to the index date. To capture comorbidity burden, we identified relevant medical conditions across domains like neurological, cardiovascular, cerebrovascular, metabolic, endocrine, psychiatric, autoimmune, infectious diseases, and sensory impairment (eg, hearing and visual loss), as well as lifestyle factors (eg, smoking) [17].

XSL•FO
RenderX

Conditions were identified using rule-based phenotyping algorithms tailored to individual conditions (see Table S2 in Multimedia Appendix 1). We also calculated the Charlson Comorbidity Index using *ICD* diagnosis codes recorded in the 2 years prior to the index date.

In addition, we reviewed the medication history within 1 year prior to the cognitive tests based on prescription records. Medications were grouped into categories based on their potential cognitive effects, including anticholinergic medications, benzodiazepines and sedative-hypnotics, antidepressants, antipsychotics, antiepileptics, opioids, and other medications of concern (eg, beta-blockers, corticosteroids, chemotherapy agents, and H2-receptor antagonists; Table S3 in Multimedia Appendix 1). Each class includes medications known to impair cognition through mechanisms such as anticholinergic burden, GABAergic activity, sedative properties, or blood-brain barrier penetration.

Statistical Analysis

We compared demographic and clinical characteristics between individuals with and without documented cognitive concerns. For continuous variables, we used 2-tailed *t* tests; for categorical variables, we used the chi-square tests or Fisher exact tests. To identify factors associated with documented cognitive concerns, we fitted multivariable logistic regression models with cognitive concern (yes/no) as the dependent variable. For the number of clinical notes in the year prior to the index date, we modeled it as a log2-transformed variable to address right-skew and facilitate interpretation (ie, per doubling of note count). Three nested models were constructed: Model 1 included sociodemographic variables and log2-transformed note count; model 2 additionally included clinical conditions (eg, Charlson Comorbidity Index and selected comorbidities); and model 3 further included medication classes based on medication prescriptions (eg, anticholinergics and antidepressants). Variables with $P<.05$ in univariate analyses were considered for inclusion in multivariable models. We reported crude and adjusted odds ratios (aORs) with 95% CIs. All analyses were conducted in R (version 4.4.1; R Foundation for Statistical Computing) using RStudio.

Results

We identified 57,474 individuals with cognitive assessments documented in EHR flowsheets (Figure 1), of whom 52,862 had total scores available. We then restricted the cohort to 48,146 individuals whose first cognitive test occurred between January 1, 2019, and April 10, 2024, and further to 42,370 individuals aged 65 years or older. Limiting the cohort to those with normal cognitive scores yielded 16,477 individuals. After excluding individuals with prior cognitive decline or earlier cognitive assessments, 15,750 individuals met all inclusion and exclusion criteria.

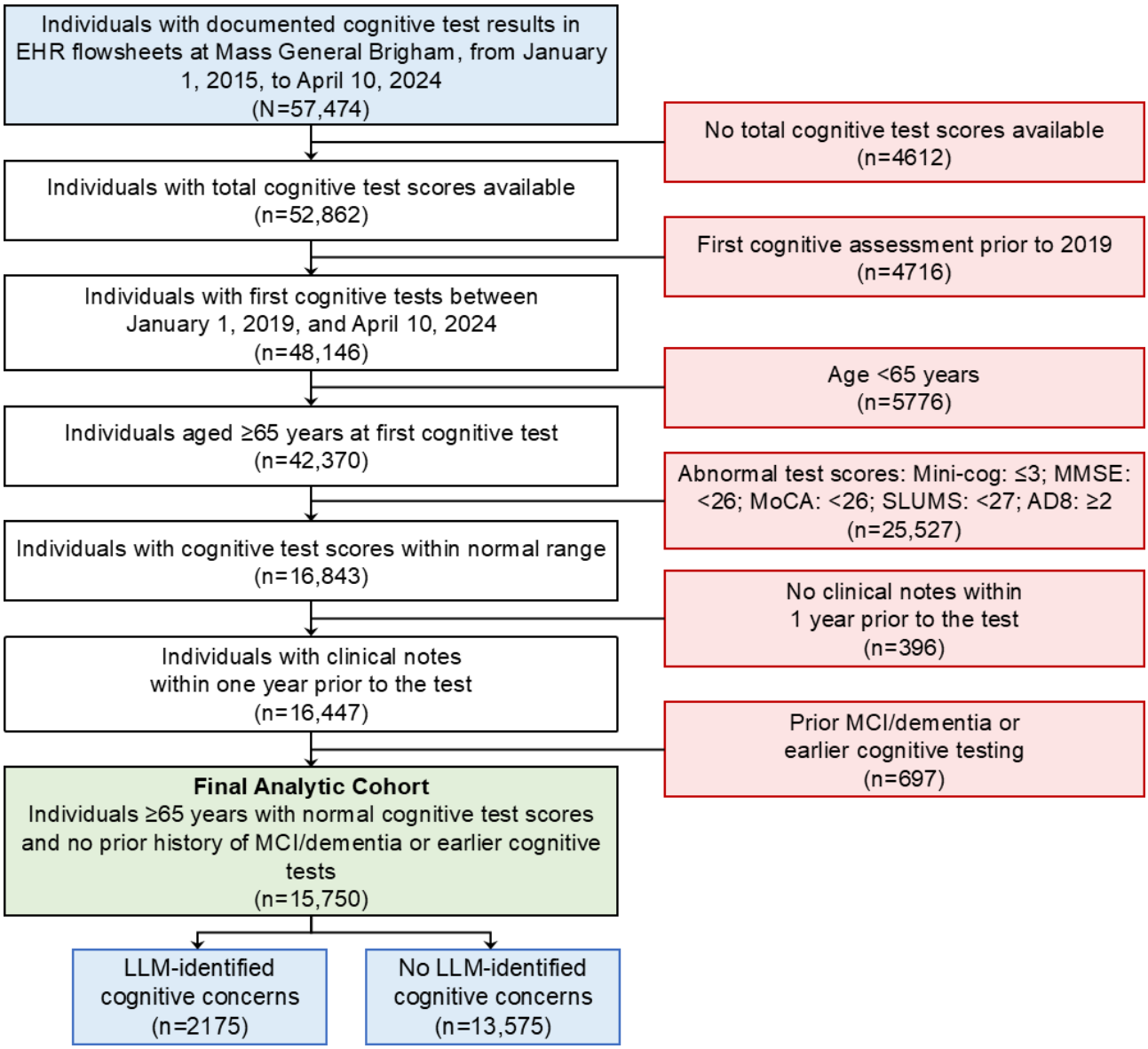
Among the final study cohort, 13.8% (n=2175) had a documented cognitive concern in the prior year. The cognitive concerns were identified from 7395 of 605,177 (1.2%) notes recorded in the year before the first normal cognitive test. The Med42-v2-8B model showed high sensitivity (97.8%) but moderate PPV (64.7%) for identifying cognitive concerns (Table S4 in Multimedia Appendix 1). When GPT-4o was used to reprocess all positive cases identified by the Med42-v2-8B, sensitivity remained high (95.7%) while PPV increased to 93.5% and specificity to 98.5%, improving the F_1-score from 0.779 to 0.945 and yielding more balanced performance across all metrics. We manually reviewed false positives and false negatives from the final pipeline. Most false positives reflected ambiguous or inferential mentions, often related to delirium or other acute conditions, whereas false negatives typically involved very subtle or indirect references to cognitive symptoms (eg, mentions of a “memory unit”) and institutional abbreviations or shorthand.

Most index cognitive assessments occurred in outpatient primary care settings (14,427/15,750, 91.6%), followed by hospital or rehabilitation settings (460/15,750, 2.9%) and outpatient neurology or geriatrics settings (440/15,750, 2.8%; Table S5 in Multimedia Appendix 1). Cognitive concern documentation was distributed across multiple clinical settings, including hospital or rehabilitation settings (574/2175, 26.4% patients), outpatient primary care (1045/2175, 48.0%), neurology or geriatrics (257/2175, 11.8%), and medical specialties (215/2175, 9.9%; Table S6 in Multimedia Appendix 1). These findings suggest that cognitive concerns were captured across diverse clinical contexts rather than solely in primary care or neurology encounters. Detailed score distributions by cognitive test are provided in Table S7 in Multimedia Appendix 1.

Among these 15,750 individuals, 15,038 (95.5%) were White, only 228 (1.4%) identified as Hispanic, and 9651 (61.3%) were female (Table 1). Most were married (n=10,095, 64%). More than half of the cohort (n=8115, 51.5%) had a college education or higher, and 97.9% (n=15,426) reported English as their preferred language. The majority (n=12,202, 77.5%) had Medicare as their insurance, while 3031 (19.2%) had commercial insurance. In terms of neighborhood socioeconomic status, 8278 (52.6%) lived in the least deprived area with national ADI≤20.

XSL•FO
RenderX

Figure 1. Cohort derivation and classification of patients with and without large language model (LLM)-identified cognitive concerns. AD8: Ascertain Dementia 8-Item Informant Questionnaire; EHR: electronic health record; MCI: mild cognitive impairment; MMSE: Mini-Mental State Examination; MoCA: Montreal Cognitive Assessment; SLUMS: Saint Louis University Mental Status Examination.



XSL•FO
RenderX

Table . Demographics and cognitive characteristics of individuals with normal cognitive test scores, stratified by presence of cognitive concerns.

Characteristics	Overall (N=15,750)	With cognitive concerns (n=2175)	Without cognitive concerns (n=13,575)	*P* value[a]
Age (y), mean (SD)	73.5 (6.4)	75 (6.9)	73.2 (6.3)	<.001
Sex (female), n (%)	9651 (61.3)	1298 (59.7)	8353 (61.5)	.10
Race, n (%)				.36
White	15,038 (95.5)	2068 (95.1)	12,970 (95.5)	
Non-White	712 (4.5)	107 (4.9)	605 (4.5)	
Black	191 (1.2)	38 (2)	153 (1.1)	
Asian	174 (1.1)	18 (<1)	156 (1.1)	
Other	347 (2.2)	51 (2)	296 (2.2)	
Ethnicity, n (%)				.23
Non-Hispanic	15,446 (98.1)	2126 (97.7)	13,320 (98.1)	
Hispanic	228 (1.4)	40 (2)	188 (1.4)	
Unknown	76 (<1)	9 (<1)	67 (<1)	
Marital status, n (%)				<.001
Married	10,095 (64.1)	1267 (58.3)	8828 (65.0)	
Nonmarried	5655 (35.9)	908 (41.7)	4747 (35.0)	
Single	1724 (10.9)	268 (12.3)	1456 (10.7)	
Divorced	1534 (9.7)	235 (10.8)	1299 (9.6)	
Widowed	2213 (14.1)	383 (17.6)	1830 (13.5)	
Others or unknown	184 (1.2)	22 (1)	162 (1.2)	
Education, n (%)				.62
Less than high school	325 (2.1)	53 (2)	272 (2.0)	
High school or some college/associates	5372 (34.1)	742 (34.1)	4630 (34.1)	
College and higher	8115 (51.5)	1112 (51.1)	7003 (51.6)	
Unknown	1938 (12.3)	268 (12.3)	1670 (12.3)	
Language, n (%)				
English	15,426 (97.9)	2127 (97.8)	13,299 (98)	.61
Non-English	263 (1.7)	41 (2)	222 (1.6)	
Unknown	61 (<1)	7 (<1)	54 (<1)	
Insurance type, n (%)				<.001
Medicare	12,202 (77.5)	1609 (74.0)	10,593 (78.0)	
Commercial	3031 (19.2)	484 (22.3)	2547 (18.8)	
Other	517 (3.3)	82 (4)	435 (3.2)	
ADI[b], n (%)				<.001
Low (≤20)	8278 (52.6)	1207 (55.5)	7071 (52.1)	
Medium (21-40)	4972 (31.6)	649 (29.8)	4323 (31.8)	
High (≥40)	1821 (11.6)	198 (9.1)	1623 (12.0)	
Unknown	679 (4.3)	121 (5.6)	558 (4.1)	
Cognitive tests, n (%)				<.001
Mini-Cog	14,278 (90.7)	1402 (64.5)	12,876 (94.9)	
MoCA[c]	904 (5.7)	566 (26.0)	338 (2.5)	

XSL•FO
RenderX

Characteristics	Overall (N=15,750)	With cognitive concerns (n=2175)	Without cognitive concerns (n=13,575)	*P* value[a]
MMSE[d]	538 (3.4)	183 (8.4)	355 (2.6)	
SLUMS[e]	11 (<1)	9 (<1)	2 (<1)	
AD8[f]	19 (<1)	15 (<1)	4 (<1)	
Number of visits in prior year per patient, median (IQR)	8 (4-16)	15 (8 - 27)	8 (4-14)	<.001
Total number of notes	605,177	173,475[g]	431,702	—[h]
Number of notes in prior year per patient, median (IQR)	25 (11 - 48)	54 (28 - 102)	22 (10 - 41)	<.001

[a]*P* values were calculated using 2-tailed *t* tests for continuous variables and chi-square tests or Fisher exact tests for categorical variables, as appropriate. For race and marital status, P values were calculated based on the major categories (White vs Non-White and married vs nonmarried, respectively); subcategories are presented for descriptive purposes only.

[b]ADI: Area Deprivation Index.

[c]MoCA: Montreal Cognitive Assessment.

[d]MMSE: Mini-Mental State Examination.

[e]SLUMS: Saint Louis University Mental Status Examination.

[f]AD8: Ascertain Dementia 8-Item Informant Questionnaire

[g]Of these notes, 7394 had mention of cognitive concerns.

[h]Not applicable.

Individuals with cognitive concerns were older (mean 75.0, SD 6.9 y vs 73.2, SD 6.3 y), less likely to be married (1267/2175, 58.3% vs 8828/13,575, 65%), more likely to have commercial insurance (484/2175, 22.3% vs 2547/13,575, 18.8%), and more likely to live in less deprived neighborhoods (low ADI: 1207/2175, 55.5% vs 7071/13,575, 52.1%; high ADI: 198/2175, 9.1% vs 1623/13,575, 12.0%). Sex, race, ethnicity, and education distribution were similar between groups with and without documented cognitive concerns. Consistent with these patterns, unadjusted analyses showed higher odds of concern documentation with older age (crude odds ratio [OR] 1.51 per 10-y increase, 95% CI 1.41 - 1.61) and commercial insurance (vs Medicare; OR 1.25, 95% CI 1.12 - 1.40), and lower odds with being married (OR 0.75, 95% CI 0.68 - 0.82) and with greater neighborhood deprivation (medium vs low ADI: OR 0.88, 95% CI 0.79 - 0.97; high vs low ADI: OR 0.71, 95% CI 0.61 - 0.84; Table S8 in Multimedia Appendix 1).

Cognitive assessment type differed significantly by cognitive concern status (*P*<.001). Individuals without cognitive concerns were more likely to undergo the Mini-Cog (12,876/13,575, 94.9% vs 1402/2175, 64.5%), whereas those with cognitive concerns were more frequently assessed with the MoCA (566/2175, 26% vs 338/13,575, 2.5%) or MMSE (183/2175, 8.4% vs 355/13,575, 2.6%). Mini-Cog scores were marginally lower in the cognitive concern group (4.7 vs 4.8, *P*<.001). Individuals with cognitive concerns also had a substantially higher number of clinical notes per individual (median 54, IQR 28 - 102 vs median 22, IQR 10 - 41; *P*<.001).

Compared with individuals without cognitive concerns, those with cognitive concerns had a markedly higher frequency of several neurological conditions (stroke or transient ischemic attack [TIA], Parkinson disease, traumatic brain injury [TBI], and epilepsy) and cardiovascular conditions (hypertension, coronary artery disease, atrial fibrillation, and heart failure), whereas obesity and hyperlipidemia were slightly less common (Table 2). In crude analysis, the strongest associations with cognitive concern documentation were observed for Parkinson disease (OR 6.13, 95% CI 4.37 - 8.60), TBI (OR 7.75, 95% CI 5.55 - 10.85), stroke or TIA (OR 5.00, 95% CI 4.23 - 5.90), and epilepsy (OR 3.70, 95% CI 2.78 - 4.90), with additional positive associations for depression (OR 2.53, 95% CI 2.29 - 2.79), excessive alcohol use (OR 2.36, 95% CI 1.76 - 3.12), and anxiety (OR 1.82, 95% CI 1.65 - 2.01; Table S5 in Multimedia Appendix 1). Conversely, obesity (OR 0.83, 95% CI 0.75 - 0.92) and hyperlipidemia (OR 0.76, 95% CI 0.68 - 0.84) were associated with lower odds, indicating cognitive concerns were less likely to be documented in patients with these comorbidities.

XSL•FO
RenderX

Table . Clinical characteristics of individuals with normal cognitive test scores, stratified by the presence of cognitive concerns.

Characteristics	Overall (N=15,750)	With cognitive concerns (n=2175)	Without cognitive concerns (n=13,575)	*P* value
Neurological conditions, n (%)				
Stroke or TIA[b]	629 (4.0)	264 (12.1)	365 (2.7)	<.001
Parkinson disease	137 (0.9)	67 (3)	70 (<1)	<.001
TBI[c]	141 (0.9)	77 (4)	64 (<1)	<.001
Epilepsy	213 (1.4)	78 (4)	135 (1.0)	<.001
Cardiovascular conditions, n (%)				
Hypertension	13,001 (82.5)	1936 (89.0)	11,065 (81.5)	<.001
Coronary artery disease	2543 (16.1)	504 (23.2)	2039 (15.0)	<.001
Atrial fibrillation	1907 (12.1)	399 (18.3)	1508 (11.1)	<.001
Heart failure	835 (5.3)	203 (9.3)	632 (4.7)	<.001
Metabolic and endocrine disorders, n (%)				
Type 2 diabetes mellitus	2652 (16.8)	481 (22.1)	2171 (16.0)	<.001
Hypothyroidism	2996 (19.0)	475 (21.8)	2521 (18.6)	<.001
Vitamin B12 deficiency	832 (5.3)	168 (7.7)	664 (4.9)	<.001
Obesity	5084 (32.3)	628 (28.9)	4456 (32.8)	<.001
Hyperlipidemia	12,858 (81.6)	1692 (77.8)	11,166 (82.3)	<.001
Psychiatric conditions, n (%)				
Depression	3150 (20.0)	763 (35.1)	2387 (17.6)	<.001
Anxiety disorders	3624 (23.0)	721 (33.1)	2903 (21.4)	<.001
Excessive alcohol use	244 (1.5)	66 (3)	178 (1.3)	<.001
Autoimmune and inflammatory disorders, n (%)				
Rheumatoid arthritis	364 (2.3)	67 (3)	297 (2.2)	.01
Infectious diseases, n (%)				
Urinary tract infection	1197 (7.6)	262 (12.0)	935 (6.9)	<.001
Upper respiratory infection	823 (5.2)	127 (5.8)	696 (5.1)	.18
Gingivitis	24 (<1)	4 (<1)	20 (<1)	.57
Herpes simplex virus	679 (4.3)	114 (5.2)	565 (4.2)	.03
Epstein-Barr virus	8 (<1)	3 (<1)	5 (<1)	.09
Cytomegalovirus	7 (<1)	3 (<1)	4 (<1)	.06
Streptococcal infection	13 (<1)	4 (<1)	9 (<1)	.09
Staphylococcal infection	26 (<1)	6 (<1)	20 (<1)	.28
COVID-19	3129 (19.9)	510 (23.4)	1619 (19.3)	<.001
Sensory impairment, n (%)				
Hearing impairment	2911 (18.5)	536 (24.6)	2375 (17.5)	<.001
Visual impairment	4428 (28.1)	662 (30.4)	3766 (27.7)	.01
Lifestyle, n (%)				
Smoking	990 (6.3)	141 (6.5)	849 (6.3)	.72
CCI[d], mean (SD)	1.7 (2.4)	2.4 (2.8)	1.6 (2.3)	<.001

[b]TIA: transient ischemic attack.

[c]TBI: traumatic brain injury.

XSL•FO
RenderX

[d]CCI: Charlson Comorbidity Index.

Individuals with subjective cognitive concerns had greater exposure to medications with potential cognitive effects than those without concerns (Table 3). Across most major drug classes, including anticholinergics, benzodiazepines and other sedative-hypnotics, antidepressants, antiepileptics, opioids, and other medications of concern such as H2-receptor antagonists, beta-blockers, corticosteroids, and chemotherapy agents, individuals with cognitive concerns were significantly more likely to have at least one prescription. Subclasses that were rarely prescribed (eg, barbiturates and antiparkinsonian anticholinergics) showed no significant differences between groups.

XSL•FO
RenderX

Table . Medication characteristics of individuals with normal cognitive test scores, stratified by the presence of cognitive concerns.

Characteristics	Overall (N=15,750)	With cognitive concerns (n=2175)	Without cognitive concerns (n=13,575)	*P* value
Anticholinergic medications, n (%)				
Urinary antispasmodics	501 (3.2)	124 (5.7)	377 (2.8)	<.001
Antihistamines (first generation)	1121 (7.1)	284 (13.1)	837 (6.2)	<.001
Antimuscarinics (GI[a])	1750 (11.1)	323 (14.9)	1427 (10.5)	<.001
Skeletal muscle relaxants	1159 (7.4)	256 (11.8)	903 (6.7)	<.001
Benzodiazepines and sedative-hypnotics, n (%)				
Benzodiazepines	2520 (16.0)	553 (25.4)	1967 (14.5)	<.001
Nonbenzodiazepine sedative-hypnotics	802 (5.1)	180 (8.3)	622 (4.6)	<.001
Barbiturates	15 (<1)	9 (<1)	6 (<1)	.44
Antidepressants, n (%)				
TCAs[b]	340 (2.2)	93 (4.3)	247 (1.8)	<.001
SSRIs[c]	2356 (15.0)	525 (24.1)	1831 (13.5)	<.001
SNRIs[d]	635 (4.0)	207 (9.5)	428 (3.2)	<.001
Mirtazapine	287 (1.8)	101 (4.6)	186 (1.4)	<.001
Antipsychotics, n (%)				
Typical antipsychotics (first generation)	908 (5.8)	213 (9.8)	695 (5.1)	<.001
Atypical antipsychotics (second generation)	223 (1.4)	107 (4.9)	116 (0.9)	.55
Antiepileptics, n (%)				
Older generation antiepileptics	107 (0.7)	41 (2)	66 (<1)	.02
Newer generation antiepileptics	1657 (10.5)	452 (20.8)	1205 (8.9)	<.001
Opioids, n (%)	4063 (25.8)	857 (39.4)	3206 (23.6)	<.001
Other medications of concern, n (%)				
Anticholinergic antiparkinsonians	11 (<1)	4 (<1)	7 (<1)	.37
H2[e]-receptor antagonists	1478 (9.4)	338 (15.5)	1140 (8.4)	<.001
Beta-blockers	4066 (25.8)	737 (33.9)	3329 (24.5)	<.001
Corticosteroids	3594 (22.8)	687 (31.6)	2907 (21.4)	<.001
Chemotherapy agents	592 (3.8)	138 (6.3)	454 (3.3)	<.001

[a]GI: gastrointestinal.

[b]TCA: tricyclic antidepressant.

[c]SSRI: selective serotonin reuptake inhibitor.

[d]SNRI: serotonin-norepinephrine reuptake inhibitor.

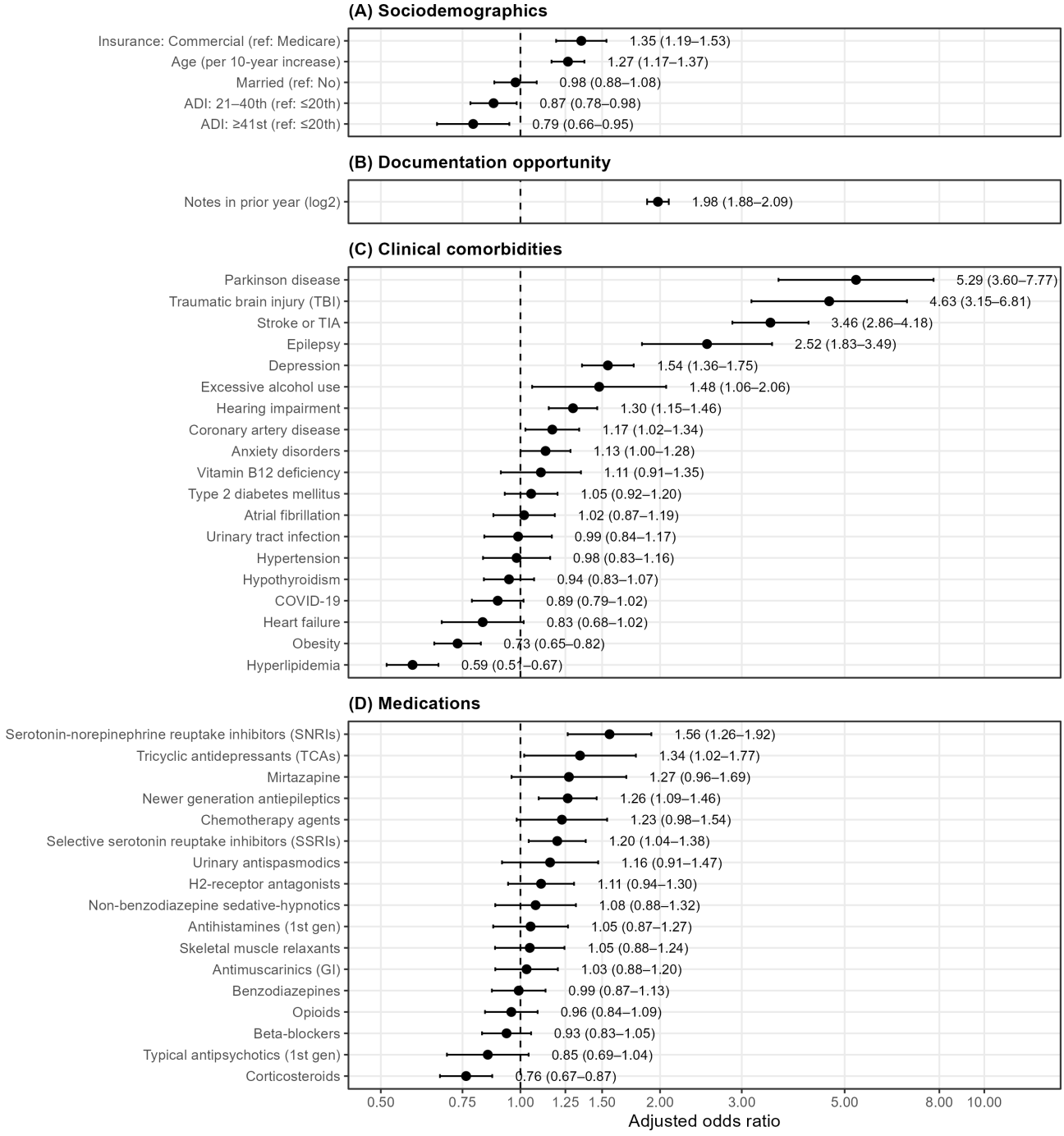
[e]H2: histamine-2.

In multivariable models that are sequentially adjusted for sociodemographic and documentation opportunity factors (model 1), then additionally for clinical comorbidities (model 2), and finally for relevant medication prescriptions (model 3), most associations were attenuated but remained significant (Multimedia Appendix 1 Table S5). Figure 2 shows the aORs and 95% CIs from the fully adjusted model (model 3) across sociodemographics, documentation opportunity, clinical comorbidities, and medication classes. In the fully adjusted model (model 3), among the sociodemographics factors, older age (aOR 1.27 per 10-y increase, 95% CI 1.17 - 1.37) and commercial insurance (vs Medicare; aOR 1.35, 95% CI

XSL•FO
RenderX

1.19 - 1.53) remained strongly associated with cognitive concern documentation, whereas higher neighborhood deprivation was attenuated (medium vs low ADI: aOR 0.87, 95% CI 0.78 - 0.98; high vs low ADI: aOR 0.79, 95% CI 0.66 - 0.95). As expected, the number of notes was strongly associated with the documentation of cognitive concerns across all models. Neurological conditions showed the strongest associations, including Parkinson disease (aOR 5.29, 95% CI 3.60 - 7.77), TBI (aOR 4.63, 95% CI 3.15 - 6.81), stroke or TIA (aOR 3.46, 95% CI 2.86 - 4.18), and epilepsy (aOR 2.52, 95% CI 1.83 - 3.49). Depression (aOR 1.54, 95% CI 1.36 - 1.75) and excessive alcohol use (aOR 1.48, 95% CI 1.06 - 2.06) also remained significant. Obesity (aOR 0.73, 95% CI 0.65 - 0.82) and hyperlipidemia (aOR 0.59, 95% CI 0.51 - 0.67) were associated with lower odds of documentation.

Figure 2. Factors associated with documentation of cognitive concerns in cognitively unimpaired older adults (N=15,750): (A) sociodemographics, (B) documentation opportunity, (C) clinical comorbidities, and (D) medications. GI: gastrointestinal; TIA: transient ischemic attack.



XSL•FO
RenderX

Discussion

Principal Findings

In this study, we examined how cognitive concerns are documented in routine clinical care among older adults with normal cognitive test scores. Using EHR data from a large cohort of individuals aged 65 years or older with an initial normal cognitive assessment, we applied an LLM-based NLP pipeline to identify cognitive concerns from clinical notes in the year prior to testing and used multivariable logistic regression to evaluate the association of cognitive concerns with sociodemographic and clinical factors. We found that cognitive concerns were documented in a minority of individuals and within a very small fraction of notes. Several neurologic (Parkinson disease, TBI, stroke or TIA, and epilepsy) and psychiatric (depression, anxiety, and excessive alcohol use) conditions were associated with higher odds of documentation of cognitive concerns, whereas some metabolic conditions (hyperlipidemia and obesity) and living in more deprived neighborhoods (higher ADI) were associated with lower odds. These associations largely persisted after adjustment for sociodemographic factors, documentation opportunity, comorbidities, and relevant medications. By linking structured cognitive test scores with concerns extracted from unstructured notes at scale, this work provides novel insights into how early cognitive symptoms are selectively recognized and recorded in the EHR, with important implications for both clinical care and EHR-based phenotyping.

The frequency of documented cognitive concerns prior to a normal test was low (13.8% overall). Documentation of cognitive concerns was even less common for individuals assessed with the Mini-Cog, which is a very brief cognitive screening test, than for those assessed with other instruments (MoCA, MMSE, SLUMS, and AD8), most of which are longer cognitive screening tests. This suggests that many cognitive tests are performed without an explicitly documented concern, which may reflect routine screening, clinician-initiated evaluation, or underdocumentation of symptoms. Additionally, it suggests that for individuals with cognitive concerns, a longer cognitive screening test is usually used.

Among sociodemographic factors, we observed clear patterns in documentation of cognitive concerns. Older age was associated with higher odds of documentation, consistent with greater symptom burden and increased clinician vigilance in older adults. Individuals with commercial insurance remained more likely to have cognitive concerns documented than those with Medicare, even in the fully adjusted model. This pattern is consistent with prior work showing that privately insured individuals are more likely to receive preventive services and certain types of outpatient care than publicly insured individuals, which may increase opportunities for patients to report and clinicians to document cognitive concerns [18]. In addition, our finding that higher neighborhood deprivation was associated with lower odds of documented cognitive concerns contrasts with prior work linking more disadvantaged neighborhoods to a higher risk of cognitive decline and dementia [19,20]. However, this discrepancy may reflect differential recognition or documentation of early cognitive symptoms in routine care rather than a truly lower burden of cognitive problems. Individuals from more disadvantaged communities may be less likely to access evaluation at earlier stages and may present for care at later stages of cognitive impairment, when deficits are more apparent [10,11]. Thus, lower documentation of subjective concerns among residents of high-ADI neighborhoods may indicate delayed help-seeking, competing clinical priorities, and fewer opportunities for symptom elicitation and recording [10].

Clinical comorbidities showed a coherent pattern. Neurological conditions, such as Parkinson disease, TBI, stroke or TIA, and epilepsy, had the highest adjusted odds, indicating that clinicians are particularly likely to document cognitive concerns when clear neurologic risk factors are present. Depression and excessive alcohol use were also associated with higher odds, likely reflecting overlapping symptoms (eg, memory complaints and concentration) with cognitive concerns and more frequent discussion of cognition in mental health-focused encounters. Hearing impairment showed a modest but consistent association, which may reflect either shared vulnerability or clinical "diagnostic uncertainty," as hearing-related communication difficulties can prompt concerns about cognition. In contrast, obesity and hyperlipidemia were linked to lower odds of documentation, which may reflect competing clinical priorities, underrecognition of cognitive issues in individuals seen primarily for metabolic disease, and residual confounding, and should be interpreted cautiously pending further study. Other cardiovascular and systemic conditions, including coronary artery disease, showed small positive associations, while several conditions (eg, hypertension and type 2 diabetes) were not independently associated after adjustment. The association for anxiety was attenuated after medication adjustment and became borderline, suggesting the crude comorbidity association may be partly explained by medication use.

Individuals with documented cognitive concerns were also more likely to receive medications with potential cognitive effects, including anticholinergics, benzodiazepines and sedative-hypnotics, antidepressants, antiepileptics, and opioids. These associations may reflect several overlapping mechanisms, including medication-related cognitive symptoms, greater psychiatric and neurological comorbidity burden, and increased clinical attention to cognition among medically complex patients. Because this study focused on documentation of cognitive concerns rather than causal cognitive effects, these findings should not be interpreted as evidence that these medications directly caused cognitive decline. Rather, they highlight the complex clinical context in which cognitive concerns are elicited and documented in routine care.

Strengths

This study is among the first to use LLMs and a large, well-defined cohort of older adults with normal cognitive test scores to examine patterns of subjective cognitive concern documentation before clear cognitive impairment or diagnosed dementia. We applied a 2-tier LLM-based NLP pipeline with high sensitivity and positive predictive value to identify cognitive concerns from unstructured notes, demonstrating the feasibility of scalable, accurate phenotyping beyond structured

XSL•FO
RenderX

data alone. We also integrated rich EHR data, including sociodemographic characteristics, neighborhood deprivation (ADI), detailed comorbidities, and medication prescriptions, enabling a nuanced examination of how clinical and contextual factors jointly affect documentation of SCD. By estimating multiple adjusted models and explicitly focusing on documentation of cognitive concerns rather than test scores or diagnosis, our work offers a novel perspective on how early cognitive symptoms are recognized and recorded in routine care and how this may shape EHR-based research and clinical tools.

Limitations

This study has several limitations. First, we identified the cohort using cognitive test results recorded in flowsheets; individuals whose assessments were documented only in narrative notes may have been missed, introducing selection bias. Second, "normal" cognition was defined using standard screening cutoffs without adjustments. Because these instruments differ in sensitivity and can be biased by educational and sociodemographic factors, subtle impairment may have been misclassified as normal. Third, cognitive concern status may have been misclassified, potentially underestimating differences between groups: we relied on documented concerns, so concerns that were not elicited, reported, or recorded could have been misclassified as "no concern." In addition, we restricted NLP extraction to notes within 1 year before the first qualifying cognitive test, potentially missing concerns documented earlier or outside this window. Because cognitive concerns could be documented repeatedly across multiple encounters and specialties over time, the current study was not designed to evaluate longitudinal care pathways or delays between initial concern documentation and cognitive testing. We also did not explicitly model the temporal ordering of neurological conditions, medication exposure, cognitive concern documentation, and cognitive testing. Some neurological diagnoses and medication exposures may have preceded concern documentation, whereas others may have occurred afterward as part of ongoing clinical evaluation and management. Therefore, observed associations should be interpreted as correlates of cognitive concern documentation rather than evidence regarding causal or prognostic sequencing. Future longitudinal studies should evaluate whether EHR-documented cognitive concerns provide incremental predictive value beyond established neurological and clinical risk factors, particularly when concern documentation precedes the onset or recognition of these conditions. Although our two-tier LLM-based pipeline showed high sensitivity and PPV in cognitive concern extraction, some misclassification is inevitable. Because GPT-4o was applied only to Med42-v2-8B positive note chunks, true positive cases missed at the first screening stage could not be recovered in the second stage. Exploratory analysis identified 2 additional reference-standard positive cases among Med42-v2-8B negative cases, suggesting that direct application of GPT-4o to all note chunks could further improve sensitivity if computational resources permit. However, the staged design substantially reduced computational cost and token usage while maintaining strong overall performance. Finally, this single-health-system, cross-sectional EHR study may have limited generalizability and cannot support causal inference or assessment of subsequent cognitive decline and dementia.

Conclusions

This study used EHR data and LLMs to identify and characterize SCD, showing that cognitive concerns are infrequently documented among older adults with normal cognitive test scores. Documentation was more common in those with neurological, psychiatric, and substance-use conditions, but less common among individuals living in more deprived neighborhoods and those with certain metabolic conditions. These findings suggest inconsistent recognition and documentation of early symptoms, with implications for equity and for EHR-based phenotypes and prediction models. Future work will evaluate whether baseline SCD predicts subsequent cognitive decline, to support more consistent, data-driven early detection and management.

Acknowledgments

The authors thank Taiyang Chen for assistance with editing and formatting the manuscript in accordance with the journal requirements. During manuscript preparation and revision, generative AI tools (ChatGPT) were used for English language editing, text polishing, and code development. The authors have thoroughly reviewed and edited all AI-assisted content and remain fully responsible for the accuracy, originality, and integrity of the final manuscript.

Funding

This work was supported by the National Institute on Aging (grant number: R00AG075190) and the Alzheimer's Association (AARF-22 - 924992). The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

The electronic health record data used in this study contain protected health information and cannot be shared publicly under Mass General Brigham and Health Insurance Portability and Accountability Act (HIPAA) policies. Deidentified aggregate data are available upon reasonable request and with appropriate institutional approvals.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: LW
Data curation: LW, SV, JT, DS
Formal analysis: LW
Funding acquisition: LW, LZ, GAM
Investigation: LW
Methodology: LW, REA, GAM
Project administration: LW
Software: LW
Supervision: LZ, GAM
Validation: LW, REA, GAM
Visualization: LW
Writing – original draft: LW
Writing – review & editing: LW, REA, SV, JT, DS, LZ, GAM

Conflicts of Interest

GAM has received salary support from Eisai Inc and Eli Lilly and Company for serving as a site principal investigator for clinical trials, and has received payments for serving as a consultant for Ono Pharma USA, Inc. All other authors report no conflicts of interest to disclose.

Multimedia Appendix 1
Supplementary tables on large language model–based cognitive concern extraction, model evaluation, clinical and medication-related factors, cognitive assessment characteristics, and factors associated with documentation of cognitive concerns.
[DOCX File, 52 KB - aging_v9i1e93011_app1.docx]

References

1. Cappa SF, Ribaldi F, Chicherio C, Frisoni GB. Subjective cognitive decline: memory complaints, cognitive awareness, and metacognition. Alzheimers Dement 2024 Sep;20(9):6622-6631. [doi: 10.1002/alz.13905] [Medline: 39051174]
2. Rabin LA, Smart CM, Amariglio RE. Subjective cognitive decline in preclinical Alzheimer's disease. Annu Rev Clin Psychol 2017 May 8;13:369-396. [doi: 10.1146/annurev-clinpsy-032816-045136] [Medline: 28482688]
3. Jessen F, Amariglio RE, van Boxtel M, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alzheimers Dement 2014 Nov;10(6):844-852. [doi: 10.1016/j.jalz.2014.01.001] [Medline: 24798886]
4. Wang XT, Wang ZT, Hu HY, et al. Association of subjective cognitive decline with risk of cognitive impairment and dementia: a systematic review and meta-analysis of prospective longitudinal studies. J Prev Alzheimers Dis 2021;8(3):277-285. [doi: 10.14283/jpad.2021.27] [Medline: 34101784]
5. Mitchell AJ, Beaumont H, Ferguson D, Yadegarfar M, Stubbs B. Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: meta-analysis. Acta Psychiatr Scand 2014 Dec;130(6):439-451. [doi: 10.1111/acps.12336] [Medline: 25219393]
6. Molinuevo JL, Rabin LA, Amariglio R, et al. Implementation of subjective cognitive decline criteria in research studies. Alzheimers Dement 2017 Mar;13(3):296-311. [doi: 10.1016/j.jalz.2016.09.012] [Medline: 27825022]
7. Subjective cognitive decline - a public health issue. : Centers for Disease Control and Prevention (U.S); 2019 URL: https://stacks.cdc.gov/view/cdc/85159 [accessed 2026-09-05]
8. Jacobson M, Thunell J, Zissimopoulos J. Cognitive assessment at Medicare's annual wellness visit in fee-for-service and Medicare Advantage plans. Health Aff (Millwood) 2020 Nov;39(11):1935-1942. [doi: 10.1377/hlthaff.2019.01795] [Medline: 33136503]
9. Tomaszewski Farias S, Leng I, Papp K, et al. Subjective cognitive decline among diverse older adults: prevalence and associations with objective cognition. Alzheimers Dement 2025 Jul;21(7):e70432. [doi: 10.1002/alz.70432] [Medline: 40572055]
10. Lewis A, Gupta A, Oh I, et al. Association between socioeconomic factors, race, and use of a specialty memory clinic. Neurology 2023 Oct 3;101(14):e1424-e1433. [doi: 10.1212/WNL.0000000000207674] [Medline: 37532510]
11. Hinton L, Tran D, Peak K, Meyer OL, Quiñones AR. Mapping racial and ethnic healthcare disparities for persons living with dementia: a scoping review. Alzheimers Dement 2024 Apr;20(4):3000-3020. [doi: 10.1002/alz.13612] [Medline: 38265164]
12. Wang L, Laurentiev J, Yang J, et al. Development and validation of a deep learning model for earlier detection of cognitive decline from clinical notes in electronic health records. JAMA Netw Open 2021 Nov 1;4(11):e2135174. [doi: 10.1001/jamanetworkopen.2021.35174] [Medline: 34792589]

XSL•FO
RenderX

13. Du X, Novoa-Laurentiev J, Plasek JM, et al. Enhancing early detection of cognitive decline in the elderly: a comparative study utilizing large language models in clinical notes. EBioMedicine 2024 Nov;109:105401. [doi: 10.1016/j.ebiom.2024.105401] [Medline: 39396423]
14. Shankar R, Bundele A, Mukhopadhyay A. Natural language processing of electronic health records for early detection of cognitive decline: a systematic review. NPJ Digit Med 2025 Mar 1;8(1):133. [doi: 10.1038/s41746-025-01527-z] [Medline: 40025194]
15. Christophe C, Kanithi PK, Raha T, Khan S, Pimentel MAF. Med42-v2: a suite of clinical LLMs. arXiv. Preprint posted online on Aug 12, 2024. [doi: 10.48550/arXiv.2408.06142]
16. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible - the Neighborhood Atlas. N Engl J Med 2018 Jun 28;378(26):2456-2458. [doi: 10.1056/NEJMp1802313] [Medline: 29949490]
17. Taylor CA, Bouldin ED, Greenlund KJ, McGuire LC. Comorbid chronic conditions among older adults with subjective cognitive decline, United States, 2015-2017. Innov Aging 2020;4(1):igz045. [doi: 10.1093/geroni/igz045] [Medline: 31915725]
18. Mahmoudi E, Lin P, Rubenstein D, et al. Use of preventive service and potentially preventable hospitalization among American adults with disability: longitudinal analysis of Traditional Medicare and commercial insurance. Prev Med Rep 2024 Apr;40:102663. [doi: 10.1016/j.pmedr.2024.102663] [Medline: 38464419]
19. Dintica CS, Bahorik A, Xia F, Kind A, Yaffe K. Dementia risk and disadvantaged neighborhoods. JAMA Neurol 2023 Sep 1;80(9):903-909. [doi: 10.1001/jamaneurol.2023.2120] [Medline: 37464954]
20. Vassilaki M, Aakre JA, Castillo A, et al. Association of neighborhood socioeconomic disadvantage and cognitive impairment. Alzheimers Dement 2023 Mar;19(3):761-770. [doi: 10.1002/alz.12702] [Medline: 35666244]

Abbreviations

ADI: Area Deprivation Index
aOR: adjusted odds ratio
EDW: Enterprise Data Warehouse
EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
***ICD*:** *International Classification of Diseases*
LLM: large language model
MCI: mild cognitive impairment
MGB: Mass General Brigham
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
NLP: natural language processing
OR: odds ratio
PPV: positive predictive value
RPDR: Research Patient Data Registry
SCD: subjective cognitive decline
SLUMS: Saint Louis University Mental Status
TBI: traumatic brain injury
TIA: transient ischemic attack

Edited by Q Fan; submitted 06.Feb.2026; peer-reviewed by DMP Jayakody, I Oh; revised version received 03.Jun.2026; accepted 18.Jun.2026; published 07.Oct.2026.

Please cite as:
Wang L, Amariglio RE, Varghese S, Tian J, Seger D, Zhou L, Marshall GA
Large Language Model–Based Identification of Subjective Cognitive Decline in Electronic Health Records Among Older Adults With Normal Cognitive Testing: Retrospective Cohort Study
JMIR Aging 2026;9:e93011
URL: https://aging.jmir.org/2026/1/e93011
doi:10.2196/93011

© Liqin Wang, Rebecca E Amariglio, Sheril Varghese, Jiazi Tian, Diane Seger, Li Zhou, Gad A Marshall. Originally published in JMIR Aging (https://aging.jmir.org), 7.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Temporal External Validation of a Parsimonious, Interpretable Machine Learning Model for Predicting 6-Month Mortality in Long-Term Care Facilities: A Retrospective Cohort Study

Yun-Cheng Tsai[1*], PhD; Yi-Lin Wu[2*], PhD; Yung-Chen Yu[2], MS; Yi-Wen Chen[2,3], PhD; Yi-Wen Chen[4], MS; Yi-Ching Yang[5,6], MS; Yu-Tai Lo[6,7], PhD

[1]PecuLab LLC, Seattle, WA, United States

[2]Department of Nursing, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[3]Department of Nursing, College of Medicine, National Sun Yat-sen University, Kaohsiung, Taiwan

[4]Department of Business Development, Smart Ageing Tech Co., Ltd., New Taipei City, Taiwan

[5]Department of Family Medicine, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[6]Department of Geriatrics and Gerontology, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, No.138, Sheng Li Road, Tainan, Taiwan

[7]Department of Public Health, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[*]these authors contributed equally

Corresponding Author:
Yu-Tai Lo, PhD
Department of Geriatrics and Gerontology, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, No.138, Sheng Li Road, Tainan, Taiwan

Abstract

Background: Early mortality after long-term care facility (LTCF) admission is common; yet, prognostic tools are often derived from Western minimum dataset-based cohorts or require hospital electronic health record linkages that are unavailable at intake in many LTCFs. There is also limited evidence on explainable, admission-feasible machine learning–based prognostication in Asian LTCF settings.

Objective: The aim of the study is to develop and temporally externally validate an interpretable machine learning model for predicting 6-month all-cause mortality among older adults newly admitted to LTCFs in Taiwan using routinely collected LTCF assessment data.

Methods: We conducted a retrospective cohort study using the JUBO Long-Term Care Database, a nationwide private administrative registry covering 636 LTCFs (37.4% of the national LTCFs) in Taiwan. We included residents with first-time LTCF admission and prespecified nonoverlapping cohorts for temporal validation: development (January 1, 2020, to December 31, 2023; n=23,901) and external validation (January 1 to December 31, 2024; n=6216). The outcome measure was death within 180 days of admission. We compared a nonlinear ensemble model (hybrid of extreme gradient boosting and random forest [HybridXGBRF]) with 7 other algorithms, including tree-based and linear benchmarks. Discrimination (area under the receiver operating characteristic curve [AUROC]), classification metrics (accuracy, precision, recall, and F_1), and calibration (Brier score and calibration plots) were assessed. Model interpretability was examined using Shapley Additive Explanations.

Results: In the development cohort, 5272 of 23,901 (22.1%) residents died within 180 days. In the 2024 temporal validation cohort, 1781 of 6216 (28.7%) residents died. In internal cross-validation, HybridXGBRF had the highest AUROC among the evaluated models (0.89, 95% CI 0.88 - 0.89). In temporal validation, HybridXGBRF maintained strong discrimination (AUROC 0.90, 95% CI 0.89 - 0.91), with an accuracy of 0.85 and an F_1-score of 0.68. Calibration plots indicated close agreement between predicted and observed risks across most probability ranges, with mild divergences at higher predicted risks. Shapley Additive Explanations analysis identified frequent hospitalizations within 6 months, activities of daily living impairment, and weight loss as influential predictors. The model showed stable AUROC across sex and age strata (0.89 - 0.90) and maintained high discrimination among residents with improving activities of daily living scores (AUROC 0.91, 95% CI 0.90 - 0.92).

Conclusions: An interpretable machine learning model using routinely collected Taiwanese LTCF assessment data achieved strong discrimination, acceptable calibration, and stable temporal validation performance without requiring hospital-based electronic health record linkage. The HybridXGBRF had the highest AUROC among the evaluated models, but performance

XSL•FO
RenderX

differences from extreme gradient boosting were small. The model may serve as a risk-stratification tool to help identify residents who could benefit from structured care review, advance care planning, or palliative care assessment. Prospective implementation studies could determine its impact on care processes and resident- or family-centered outcomes.

(JMIR Aging 2026;9:e94567) doi:10.2196/94567

KEYWORDS

long-term care facilities; mortality prediction; machine learning; AI; Taiwan

Introduction

The accelerating population aging at a global scale has established long-term care facilities (LTCFs) as cornerstones of geriatric health care systems [1,2]. For many older adults, LTCF transition is a major event in their health trajectory, often marked by precipitous functional decline and limited life expectancy [1,3]. This renders paramount accurate prognostic assessment at admission, not only to align clinical care with patient goals but also to facilitate timely advance care planning (ACP), optimize resource allocation, and ensure equitable access to palliative services [4,5]. Substantial early mortality has been reported after first-time LTCF admission, with approximately 1 in 5 residents dying within the first months of placement, and mortality rates markedly exceeding those in age-matched community populations [2,3]. Therefore, there is importance in promoting the early identification of LTCF residents at high risk for short-term mortality.

Several prognostic tools have been developed to estimate mortality risk among LTCF residents, mostly in the United States and predominantly using a minimum dataset (MDS) or related assessment instruments [5-10]. Examples include the Advanced Dementia Prognostic Tool, designed for residents with advanced dementia [7,9], and the MDS Mortality Risk Index for 6-month risk stratification [5,6,8]. Although several of these models have undergone prospective validation [6,7], most were developed more than a decade ago and may not reflect contemporary care practices, evolving resident acuity, or shifts in LTCF populations following postacute care policy changes. Recent initiatives have attempted to address these limitations, achieving mixed success. The Hospice Eligibility Prediction Index incorporates various clinical indicators of hospital admission history [11], while Jorissen et al [12] developed a comprehensive 51-variable model with individual-, medication-, system-, and health care–related factors known at entry into an LTCF. However, their clinical utility remains hindered by excessive complexity and a lack of external validation in a greater number of populations.

In such contexts, machine learning methods offer potential advantages for risk prediction, including the ability to model nonlinear associations and higher-order interactions within correlated clinical domains [13]. Recent reports have described improved discrimination when using machine learning models that incorporate demographic characteristics, comorbidities, and prior health care use [14,15]. Despite the potential, many of the extant machine learning models for mortality prediction remain dependent on acute care hospitalization data, which are often inaccessible or significantly delayed during the LTCF intake process. Furthermore, evidence of AI-driven prognosis in Asian LTCF contexts remains sparse, leaving a critical gap in health equity and geographical generalizability [16].

To address these gaps, this study aimed to, first, develop an interpretable machine learning prediction model to predict 6-month mortality among older adults newly admitted to LTCFs in Taiwan, using variables that are routinely collected at the time of admission or routine early-stay assessments. Second, to temporally externally validate the model using a later calendar-year cohort.

Methods

Study Design, Dataset, and Population

Taiwan's long-term care system provides a continuum of home-based, community-based, and institutional services for its rapidly aging population [17]. Residential LTCFs represent an institutional care setting within this continuum, serving older adults who require continuous assistance, supervision, nursing support, or help with activities of daily living (ADL) [18]. Accordingly, Taiwanese LTCF residents generally represent a frail and care-dependent population with substantial functional impairment, diverging from the broader community-dwelling older adult population. This study therefore focused on newly admitted LTCF residents, a clinically relevant point at which risk stratification may inform care planning, family communication, and ACP.

As aforementioned, we conducted a retrospective cohort study to develop and validate a machine learning model for predicting 6-month all-cause mortality among older adults newly admitted to LTCFs in Taiwan. This study used deidentified administrative and clinical assessment data routinely collected at the time of LTCF admission. Data were extracted from the JUBO Long-Term Care Database, a nationwide private administrative registry that captures structured admission, discharge, and clinical assessment data from LTCFs in Taiwan. The JUBO LTCFs system plays a significant role in Taiwan's long-term care market, with its services being used by 37.4% of all LTCFs nationwide (ie, a total of 636 institutions). Most (67.7%) of the facilities were in the south, 20% were in the central region, 11% were in the north, and a small number were in the east. This study was reported in accordance with the TRIPOD-AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis-Artificial Intelligence) statement (Checklist 1) [19].

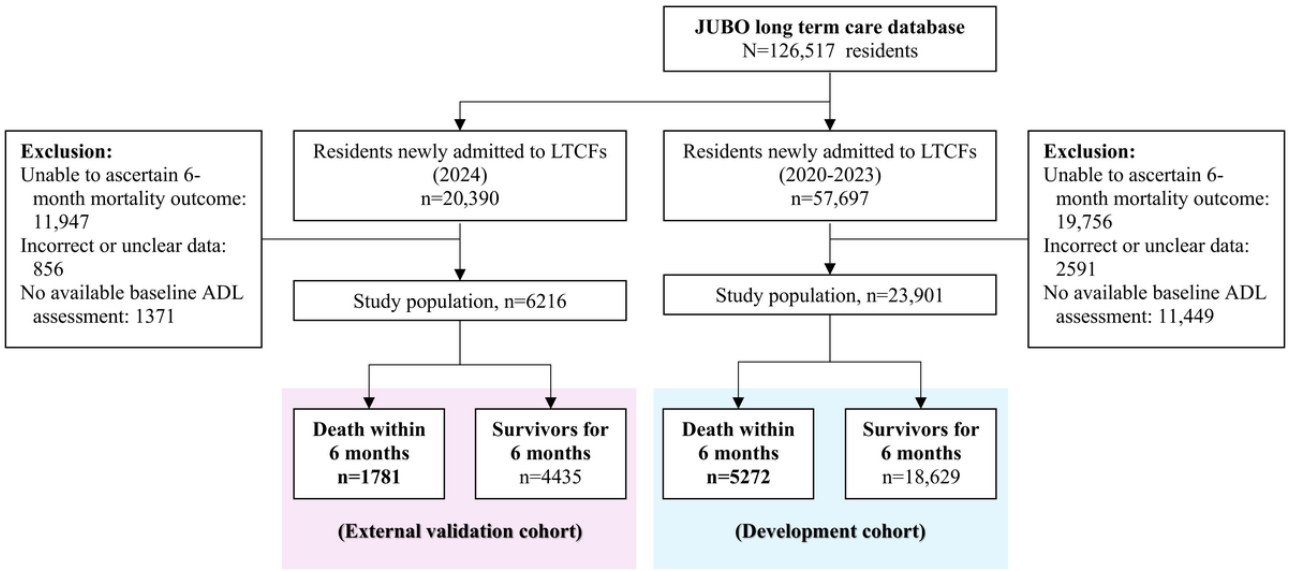
The study sample comprised all residents with a first-time LTCF admission recorded in the JUBO database between January 1, 2020, and December 31, 2024. First-time admission was defined as the first LTCF admission recorded for a resident in the JUBO database during the study period. To enable temporal external

XSL•FO
RenderX

validation, we prespecified 2 nonoverlapping cohorts by admission date, as follows: a model development cohort including admissions from January 1, 2020, to December 31, 2023 (n=23,901); a temporal external validation cohort including admissions from January 1, 2024, to December 31, 2024 (n=6216).

To be included in this study, residents were not required to remain in the LTCF for the entire 6-month outcome window. Residents who died within 6 months after admission while under LTCF follow-up were classified as outcome events. Residents were excluded if their 6-month vital status could not be reliably ascertained, such as discharge or loss to follow-up before completion of the outcome window without subsequent status information in the JUBO database. The exclusion criteria also included implausible admission dates, duplicate records that could not be reconciled, and absence of any available baseline ADL assessment required to define baseline functional status. The cohort selection process, exclusions, and event counts are reported in Figure 1.

Figure 1. Flow diagram for study population inclusion. Unable to ascertain 6-month mortality outcome: residents discharged or lost to follow-up before completion of the 6-month outcome window whose postdischarge vital status was unavailable in the JUBO database. Residents who died within 6 months while under LTCF follow-up were included as outcome events. Incorrect or unclear data: records with implausible admission dates, inconsistent admission or discharge information, or duplicate records that could not be reconciled. No available baseline ADL assessment: residents with no recorded ADL assessment at admission and therefore no baseline ADL information available for feature construction. Residents with a baseline ADL assessment were retained in the analytic cohort even if individual ADL items were partially missing or if subsequent ADL information was insufficient to derive the ADL-change variable. These values were treated as predictor-level missingness and handled using the prespecified imputation procedure. ADL: activities of daily living; LTCF: long-term care facility.



Outcome, Predictors, and Data Preprocessing

Outcome Definition

The primary outcome was all-cause mortality within 6 months after LTCF admission. Mortality status and date of death were ascertained from the discharge or case-closure disposition fields in the JUBO database, which contains the recorded reason for case closure, including death, as documented by facility staff at the time of case closure. The outcome was operationalized as a binary indicator of death occurring within 6 months after LTCF admission. Postdischarge vital status data were not consistently available; thus, residents discharged or lost to follow-up before 6 months without subsequent status information were excluded from the primary analysis.

Candidate Predictors

Candidate predictors were prespecified through a 2-step process. First, we conducted a review of published prognostic models and studies on short-term mortality in LTCF and nursing home populations. Several repeatedly reported prognostic domains were identified, including demographics, functional status, comorbidity burden, recent acute care service use, nutritional risk, psychosocial status, and care-related factors (Multimedia Appendix 1). Second, we mapped these domains to the structured variables routinely collected in the JUBO LTCF assessment system, prioritizing predictors that were clinically interpretable, consistently available in routine LTCF workflows, and feasible to implement without linkage to hospital-based electronic health records. Although prior models have incorporated hospital-derived laboratory measures, such information is not consistently available during LTCF admission. Since this study prioritized real-world feasibility, we excluded laboratory variables from the analyses.

In addition to variables available at or shortly after admission, we also considered predictors capturing clinically meaningful changes during the early LTCF stay. This is because the intended use of the model is risk stratification using routinely collected data available at or before the time of risk estimation, not prediction based solely on variables fixed at the moment of admission. We accordingly included selected longitudinal or look-back variables that can be derived from routine LTCF records. Integrating these considerations, we prespecified candidate predictors, including age, sex, recent hospitalization history, level of consciousness measured by the Glasgow Coma Scale [20], ADL score [21], Geriatric Depression Scale-15 score [22], Study of Osteoporotic Fractures score [23], Mini

XSL•FO
RenderX

Nutritional Assessment—Short Form [24], Cumulative Illness Rating Scale for Geriatrics score [25], do-not-resuscitate documentation, number of medications, body weight, fall count, denture use, tube feeding, and use of respiratory aids. The selected longitudinal or look-back predictors, including hospitalization history, falls, weight change, and ADL change, were defined using prespecified assessment windows in the LTCF record. The operational definition, data source, and measurement window for each predictor, including whether it was available before admission, at admission, shortly after admission, or during routine follow-up before risk estimation, are detailed in Multimedia Appendix 2.

Predictor Measurement and Preprocessing

Predictors were extracted from structured fields recorded at admission and from routine LTCF assessments available before or at the time of risk estimation. Variables recorded after death, or otherwise unavailable at the corresponding risk-estimation time point, were not used.

ADL-related features were handled by distinguishing cohort eligibility from predictor-level missingness. Residents were required to have at least 1 baseline ADL assessment to define baseline functional status; those without such assessment were excluded. Among eligible residents, partially missing ADL items were treated as predictor-level missingness, not exclusion criteria. ADL change was calculated only when at least 2 ADL assessments were available within the prespecified assessment window before risk estimation. Otherwise, ADL change was treated as a missing predictor and handled using the prespecified imputation procedure.

Other variables representing recent clinical instability or longitudinal change, such as hospitalization count, falls, and weight change, were calculated using prespecified look-back or assessment windows before the corresponding risk-estimation time point. Continuous variables were retained on their original scales where appropriate, and categorical variables were encoded using standard binary or categorical representations. To balance data quality with feasibility and interpretability, candidate predictors with high missingness (>30%) were excluded from the model development cohort (Multimedia Appendix 3). The remaining missing values were handled using an imputation strategy.

Missing Data Handling

Guided by the clinical documentation patterns of the source database [26-28], we applied a pragmatic rule-based imputation strategy to handle missing values. For binary or count variables typically documented only when present in routine LTCF workflows (eg, do-not-resuscitate documentation, number of hospitalizations within the previous 6 months, falls, use of respiratory assistive devices, tube feeding, and denture use), missing entries were treated as indicating absence (coded as 0). This assumption reflects the usual documentation practice where the event is actively recorded when it occurs, although we acknowledge that under-documentation in some settings may lead to misclassification.

For continuous measures and scale scores, we standardized each variable using *z*-score normalization, with the mean and SD estimated from the model development cohort only. Then, we imputed missing standardized values as 0 (ie, the cohort mean on the original scale). This approach places missing values at a neutral reference point in the distribution, avoiding the imputation of implausible extremes [29]. All preprocessing parameters estimated in the development cohort were applied unchanged to the external validation cohort to prevent information leakage and mirror real-world deployment.

Model Development, Internal Evaluation, and Calibration

We compared 8 classification algorithms on their predictions of 6-month mortality (Multimedia Appendix 4) to investigate the prognostic advantage of advanced computational techniques over traditional statistical methods. The candidates were categorized into two groups, which are explored hereinafter.

1. Tree-based ensemble learners: This group included an extreme gradient boosting (XGB) classifier (gradient boosting) [30], a random forest classifier (bagging) [31], and our proposed hybrid of extreme gradient boosting and random forest [HybridXGBRF] approach. Ensemble models were selected because of their ability to capture intricate nonlinear feature interactions and robustness against noisy clinical data. The HybridXGBRF model is an ensemble architecture that integrates the predictive strengths of gradient boosting and bagging using a weighted probability blending (soft-voting) mechanism. For each prediction, the model computes a final probability P_{Hybrid}, which is as described below:

$$\text{PHybrid}=\alpha\cdot\text{PXGB}+(1-\alpha)\cdot\text{PRF},\ \alpha\in[0,1]$$

where P_{XGB} and P_{RF} are the predicted probabilities from the base learners, and α controls their relative contribution. The blending weight (α) was tuned via an exhaustive search over a prespecified range ($\alpha\in[0,1]$) in the development cohort. For each candidate α, the Hybrid probability was evaluated under the same cross-validation framework. The α yielding the best cross-validated performance (primary criterion: area under the receiver operating characteristic curve [AUROC]) was selected and then fixed for refitting on the full development cohort, as well as evaluation in the external validation cohort.

2. Linear statistical benchmarks: Five linear models were constructed as performance baselines, including standard logistic regression and regularized versions (ie, Ridge, L2 [32]; Lasso, L1 [33]; and Elastic Net [34]), to address potential multicollinearity among clinical predictors.

An internal model evaluation was conducted in the development cohort using 5-fold stratified cross-validation, preserving the event proportion in each fold. Within each fold, all preprocessing steps, including feature screening rules, imputation, standardization (when applicable), and model fitting, were performed using only the training partition and then applied to the corresponding validation partition. This fold-specific pipeline served to minimize information leakage. For each model and fold, we calculated the discrimination and classification performance metrics, as follows: AUROC, accuracy, precision, recall (sensitivity), F_1-score, and confusion matrix. Multimedia Appendix 5 presents the definitions of the evaluation metrics.

XSL•FO
RenderX

When the accuracy was summarized across folds, we reported the mean (SD), while probability accuracy was assessed using the Brier score whenever applicable. We prioritized the AUROC as the primary metric for model comparison because it is threshold-independent and less sensitive to class imbalance. Meanwhile, threshold-dependent metrics (precision, recall, or F_1) were reported at the classification threshold used across models for comparability.

Calibration was first evaluated using reliability diagrams. For each model, predicted 6-month mortality risks were grouped into deciles of predicted risk based on the predicted probability distributions. Within each bin, we calculated the mean predicted probability and observed event rate (ie, the proportion of residents who died within 6 months). These values were plotted against the 45-degree line representing perfect calibration. For internal evaluation and to reduce optimism, calibration curves were generated using out-of-fold predictions from the cross-validation procedure; for temporal external validation, calibration curves were generated using predictions from the final model applied to the 2024 cohort.

In addition to graphical and decile-based calibration assessments, we calculated calibration intercept, calibration slope, observed-to-expected ratio, and Brier score for the final HybridXGBRF model in both cohorts. Calibration intercept and slope were estimated using logistic recalibration; the observed-to-expected ratio was calculated as observed deaths divided by the sum of predicted probabilities; and the Brier score as the mean square difference between observed outcomes and predicted probabilities. CIs were estimated using resident-level bootstrap resampling. We also summarized calibration for the final HybridXGBRF model using risk-decile calibration tables, reporting the number of residents, observed deaths, observed risk, mean predicted risk, and predicted-risk range within each decile.

Temporal External Validation, Calibration, and Subgroup Analyses

To assess generalizability under a temporal dataset shift, we performed temporal external validation in the 2024 cohort (admissions in 2024), which was not used for model development, feature engineering decisions, or cross-validation. The final models were refitted to the full development cohort using the same predictors and preprocessing rules, and then evaluated in the external validation cohort.

All preprocessing parameters estimated in the development cohort (eg, scaling parameters for continuous variables) were applied to the external validation cohort. Model performance in the external validation was quantified using the same set of metrics as in the internal evaluation (Multimedia Appendix 5).

We additionally evaluated discrimination across sex, age (>85 and ≤85 years), and ADL decline (yes or no) to examine performance stability, reporting subgroup AUROC values using the same prediction pipeline.

Clinical Utility and Threshold Analyses

We conducted 2 complementary analyses for the HybridXGBRF model in the temporal external validation cohort in order to assess potential clinical utility beyond discrimination. First, we performed decision-curve analysis: net benefit was calculated across a range of threshold probabilities and compared with default strategies of classifying all residents as high risk or no residents as high risk. Second, we evaluated threshold-specific performance using clinically interpretable predicted-risk thresholds: for each threshold, we reported sensitivity (recall), specificity, positive predictive value, negative predictive value, F_1-score, and the proportion of residents classified as high risk. This analysis was intended to examine the trade-offs between case detection and alert burden under alternative risk-stratification thresholds.

Model Interpretability and Individual-level Risk Explanation

In an attempt to enhance clinical interpretability, we quantified feature contributions using Shapley Additive Explanations (SHAP) [35,36]. Because HybridXGBRF generates the final predicted probability as a weighted blend of the XGB and random forest component predictions, SHAP values were computed for the final HybridXGBRF blended prediction function. The reported feature attributions correspond to the final predicted probability of 6-month mortality, not to either component model alone. Global interpretability was visualized using SHAP beeswarm, and forest plots were used to rank feature importance across the study population. Local interpretability for individual-level risk explanations was demonstrated using SHAP force plots, illustrating how specific clinical features shifted the final HybridXGBRF-predicted probability from the baseline risk estimate.

Statistical Analysis

The descriptive characteristics of residents in the development cohort were summarized overall and by the 6-month mortality status. Group differences were quantified using standardized mean differences to describe the magnitude of imbalance between residents who died within 6 months and those who survived.

Intending to assess potential selection bias regarding cohort exclusion, we compared baseline characteristics between the development cohort and 2 major excluded groups: residents excluded because their 6-month mortality outcome could not be reliably ascertained and residents excluded because no baseline ADL assessment was available.

Model performance metrics were computed for internal evaluation and temporal external validation, as described earlier. For AUROC, precision, recall, F_1-score, accuracy, Brier score, and subgroup-specific performance metrics, 95% CIs were calculated using resident-level bootstrap resampling of the evaluation set.

For the purpose of comparing the discriminative performance of HybridXGBRF with XGB, we conducted paired bootstrap comparisons of AUROC within each cohort using 2000 resident-level bootstrap resamples. In each bootstrap sample, AUROC was calculated for both models using the same resampled residents, and the paired AUROC difference was computed as AUROC_HybridXGBRF−AUROC_XGB. The

XSL•FO
RenderX

95% CI was estimated from the empirical distribution of the paired AUROC differences, and 2-sided *P* values were calculated to test whether the paired difference differed from 0. Statistical significance was set at $P<.05$. Database management and extraction were conducted using PostgreSQL, and all statistical analyses were performed using Python (version 3.12.7; Python Software Foundation).

Sensitivity Analyses

The primary binary 6-month mortality evaluation was complemented by a time-to-event discrimination analysis using cumulative or dynamic area under the curve (AUC) at monthly time points before the 6-month end point. At each time point, cases were defined as residents who had died by that time, and controls were defined as residents who remained event-free beyond that time. Cumulative or dynamic AUC was calculated separately for the development and temporal external validation cohorts. This analysis evaluated whether model discrimination remained stable across earlier time horizons, while the primary model objective remained 6-month mortality risk prediction.

A concern arises here that missing binary or count variables may represent incomplete documentation, not true clinical absence. Thus, 2 additional post-hoc analyses were conducted to tackle this issue. For one, we examined whether missingness in selected key clinical assessments, including subsequent ADL reassessment, initial body weight, and Glasgow Coma Scale score, was associated with 6-month mortality by using missingness indicators for these variables as exposures in adjusted logistic regression models. For another, facility-level variation in overall documentation missingness was examined while stratified by facility size and geographic region, and heterogeneity across individual facilities was tested using a global chi-square test of independence. This analysis was conducted to assess whether missingness patterns reflected institutional documentation practices in addition to patient-level clinical status.

Ethical Considerations

This study was approved by the institutional review board of National Cheng Kung University Hospital (B-ER-112‐466). All data were anonymized prior to analysis, and the use of AI models was aligned with the ethical principles of transparency, clinical utility, and human oversight.

Results

Study Population and Characteristics

Between 2020 and 2023, 57,697 newly admitted LTCF residents were identified in the development cohort. After excluding residents whose 6-month mortality outcome could not be reliably identified (n=19,756), records with incorrect or unclear data (n=2591), and residents with no available baseline ADL assessment required for feature construction (n=11,449), the final development cohort included 23,901 residents from 493 LTCFs. Among them, 5272 residents died within 6 months after admission and 18,629 survived for at least 6 months. The participant flow diagram is presented in Figure 1.

Residents discharged or lost to follow-up before completion of the 6-month outcome window with unknown subsequent vital status were compared with the analytic cohort (Table S1 in Multimedia Appendix 6). Compared with the analytic cohort, this excluded group was younger and had higher ADL scores, higher body weight, lower prevalence of do-not-resuscitate documentation, and a higher prevalence of denture use. Facility size was similar between groups, although facility region differed; indicators of greater medical dependency, including tube feeding and respiratory support were not significantly different.

Residents excluded due to the absence of a baseline ADL assessment were similarly compared with the analytic cohort (Table S2 in Multimedia Appendix 6). This excluded group (vs analytic cohort) was slightly younger and had lower Modified Cumulative Illness Rating Scale for Geriatric Patients scores, lower prevalence of do-not-resuscitate documentation, fewer hospital admissions within 6 months, and markedly lower documentation of denture use, tube feeding, respiratory support, and fall history.

The characteristics of the development cohort are summarized in Table 1. The mean age was 79.3 (SD 11.8) years, and 12,155 (50.9%) residents were female. Residents who died within 6 months (vs who survived for at least 6 months) had greater functional impairment, lower body weight, more frequent hospital admissions, and higher use of tube feeding and respiratory support. Standardized mean differences greater than 0.1 were observed for Cumulative Illness Rating Scale for Geriatrics, ADL measures, body weight, number of hospitalizations, sex, do-not-resuscitate documentation, denture use, tube feeding, respiratory support, and institutional region.

XSL•FO
RenderX

Table . Characteristics of the development cohort (n=23,901).

Characteristic	Total (n=23,901)	Death group (n=5272)	Survivor group (n=18,629)	SMD[a]
Age (years), mean (SD)	79.3 (11.8)	80.2 (11.5)	79.1 (11.9)	0.09
Age group (years), n (%)				0.06
<65	2542 (10.6)	521 (9.9)	2021 (10.8)	
65 - 69	1926 (8.1)	430 (8.2)	1496 (8)	
70 - 74	2794 (11.7)	549 (10.4)	2245 (12.1)	
75 - 79	3112 (13)	623 (11.8)	2489 (13.4)	
80 - 84	4345 (18.2)	1001 (19)	3344 (18)	
85 - 89	4734 (19.8)	1075 (20.4)	3659 (19.6)	
90+	4448 (18.6)	1073 (20.4)	3375 (18.1)	
Sex, n (%)				0.23
Male	11,746 (49.1)	3069 (58.2)	8677 (46.6)	
Female	12,155 (50.9)	2203 (41.8)	9952 (53.4)	
CIRS-G[b], mean (SD)	1.7 (0.6)	1.8 (0.6)	1.7 (0.6)	0.17
DNR[c] notes (+), n (%)	9777 (40.9)	2472 (46.9)	7305 (39.2)	0.15
ADL[d], mean (SD)				
First score	28.6 (31.7)	15.1 (22.9)	32.3 (32.8)	−0.61
Maximum score in 6 months	31 (32.8)	15.4 (23.3)	35.4 (33.7)	−0.69
Last score	28.5 (31.8)	13.3 (21.7)	30.5 (32.4)	−0.63
GCS[e] (EVM), mean (SD)[f]	7.7 (5.3)	7.0 (4.9)	7.9 (5.4)	−0.16
Dentures, n (%)	3918 (16.4)	720 (13.7)	3198 (17.2)	−0.13
Tube feeding, n (%)	2081 (8.7)	758 (14.4)	1323 (7.1)	0.27
Respiratory support, n (%)	848 (3.5)	410 (7.8)	438 (2.4)	0.26
History of falls, n (%)	1197 (5)	294 (5.6)	903 (4.8)	0.03
BW[g], mean (SD)	51.4 (12.1)	49.1 (13.8)	52 (11.5)	−0.22
Number of admissions in 6 months, mean (SD)	0.7 (1.1)	1.1 (1.1)	0.5 (1)	0.53
Institutional regions, n (%)				0.14
North	2245 (9.4)	378 (7.2)	1867 (10)	
Central	4754 (19.9)	1122 (21.3)	3632 (19.5)	
South	16,603 (69.5)	3745 (71)	12,861 (69)	
East	299 (1.3)	30 (0.6)	269 (1.4)	

[a]SMD: standardized mean difference.

[b]CIRS-G: Modified Cumulative Illness Rating Scale for Geriatric Patients.

[c]DNR: do not resuscitate.

[d]ADL: activities of daily living.

[e]GCS: Glasgow Coma Scale.

[f]GCS scores between 3 and 15. It comprises 3 parameters: best eye response (E), best verbal response (V), and best motor response (M).

[g]BW: body weight.

In 2024, a total of 20,390 newly admitted LTCF residents were identified for temporal external validation. After applying the predefined exclusion criteria, the temporal external validation cohort included 6216 residents from 548 LTCFs. Among them, 1781 residents died within 6 months after admission, and 4435 survived for at least 6 months. The mean age was 78.6 (SD 11.7) years, and 3043 (49%) residents were female. Baseline

XSL•FO
RenderX

characteristics of the temporal external validation cohort are presented in Table 2.

Table . Characteristics of the external validation cohort (n=6216).

Characteristic	Total (n=6216)	Death group (n=1781)	Survivor group (n=4435)
Age (years), mean (SD)	78.6 (11.7)	79.5 (11.4)	78.2 (11.9)
Age group (years), n (%)			
<65	711 (11.4)	183 (10.3)	528 (11.9)
65 - 69	531 (8.5)	143 (8)	388 (8.7)
70 - 74	786 (12.6)	230 (12.9)	556 (12.5)
75 - 79	864 (13.9)	212 (11.9)	652 (14.7)
80 - 84	1116 (18)	325 (18.2)	791 (17.8)
85 - 89	1388 (22.3)	415 (23.3)	973 (21.9)
90+	820 (13.2)	273 (15.3)	547 (12.3)
Sex, n (%)			
Male	3173 (51)	1067 (59.9)	2106 (47.5)
Female	3043 (49)	714 (40.1)	2329 (52.5)
DNR[a] notes (+), n (%)	2832 (45.6)	946 (53.1)	1886 (42.5)
ADL[b], mean (SD)			
First score	26.3 (30.1)	15 (23.2)	30.7 (31.3)
Maximum score in 6 months	28.9 (31.2)	15.3 (23.5)	34.4 (32.2)
Last score	26.6 (30)	12.2 (21)	29.1 (30.7)
GCS[c] (EVM), mean (SD)[d]	7.9 (5.4)	7(5)	8.2 (5.5)
Dentures, n (%)	783 (12.6)	151 (8.5)	632 (14.3)
Tube feeding, n (%)	888 (14.3)	378 (21.2)	510 (11.5)
Respiratory support, n (%)	427 (6.9)	249 (14)	178 (4)
History of falls, n (%)	618 (9.9)	191 (10.7)	427 (9.6)
BW[e], mean (SD)	52.7 (11.8)	50.5 (12)	53.5 (11.7)
Number of admissions in 6 months, mean (SD)	0.7 (1.1)	1 (1.1)	0.6 (1.1)
Institutional regions, n (%)			
North	462 (7.4)	138 (7.7)	324 (7.3)
Central	1328 (21.4)	407 (22.9)	921 (20.8)
South	4333 (69.7)	1217 (68.3)	3116 (70.3)
East	93 (1.5)	19 (1.1)	74 (1.7)

[a]DNR: do not resuscitate.

[b]ADL: activities of daily living.

[c]GCS: Glasgow Coma Scale.

[d]GCS scores between 3 and 15. It comprises 3 parameters: best eye response (E), best verbal response (V), and best motor response (M).

[e]BW: body weight.

Model Development, Internal Evaluation, and Calibration

The performance of the 8 classification algorithms in the internal evaluation is summarized in Table 3 and Figure 2. Overall, the linear models were outperformed by the tree-based models (eg, HybridXGBRF, XGB, and random forest), achieving higher AUROC values (0.87 - 0.89 vs 0.86), higher sensitivity or recall (0.51 - 0.54 vs 0.48 - 0.49), and higher F_1-scores (0.62 - 0.64 vs 0.57 - 0.58).

XSL•FO
RenderX

Table . Performance of the different predictive models.

Model metrics	Model 1[a]	Model 2[b]	Model 3[c]	Model 4[d]	Model 5[e]	Model 6[f]	Model 7[g]	Model 8[h]
AUROC[i] (95% CI)	*0.89 (0.88 - 0.89)* [j]	0.88 (0.88 - 0.89)	0.87 (0.87 - 0.88)	0.86 (0.85 - 0.86)	0.86 (0.85 - 0.86)	0.86 (0.85 - 0.86)	0.86 (0.85 - 0.86)	0.86 (0.85 - 0.86)
Precision (95% CI)	0.79 (0.78 - 0.80)	*0.79 (0.78 - 0.80)*	0.75 (0.74 - 0.76)	0.70 (0.69 - 0.71)	0.70 (0.69 - 0.72)	0.70 (0.69 - 0.71)	0.70 (0.69 - 0.71)	0.70 (0.69 - 0.71)
Sensitivity or recall (95% CI)	0.54 (0.52 - 0.55)	0.51 (0.50 - 0.52)	*0.54 (0.52 - 0.55)*	0.49 (0.48 - 0.51)	0.48 (0.47 - 0.50)	0.48 (0.47 - 0.50)	0.49 (0.47 - 0.50)	0.49 (0.48 - 0.51)
Specificity	0.96	0.96	0.95	0.94	0.94	0.94	0.94	0.94
F_1-score (95% CI)	*0.64 (0.63 - 0.65)*	0.62 (0.61 - 0.63)	0.62 (0.61 - 0.64)	0.58 (0.56 - 0.59)	0.57 (0.56 - 0.59)	0.57 (0.56 - 0.59)	0.57 (0.56 - 0.59)	0.58 (0.56 - 0.59)
Accuracy (95% CI)	*0.87 (0.86 - 0.87)*	0.86 (0.86 - 0.87)	0.86 (0.85 - 0.86)	0.84 (0.84 - 0.84)	0.84 (0.84 - 0.85)	0.84 (0.84 - 0.85)	0.84 (0.84 - 0.85)	0.84 (0.84 - 0.84)
Brier score (95% CI)	*0.10 (0.10 - 0.10)*	0.10 (0.10 - 0.11)	0.11 (0.10 - 0.11)	0.12 (0.11 - 0.12)	0.12 (0.11 - 0.12)	0.12 (0.11 - 0.12)	0.12 (0.11 - 0.12)	0.12 (0.11 - 0.12)

[a]HybridXGBRF (hybrid of extreme gradient boosting and random forest; proposed approach).

[b]Extreme gradient boosting classifier.

[c]Random forest classifier.

[d]Logistic regression (max-iter =1000).

[e]Ridge.

[f]Elastic.

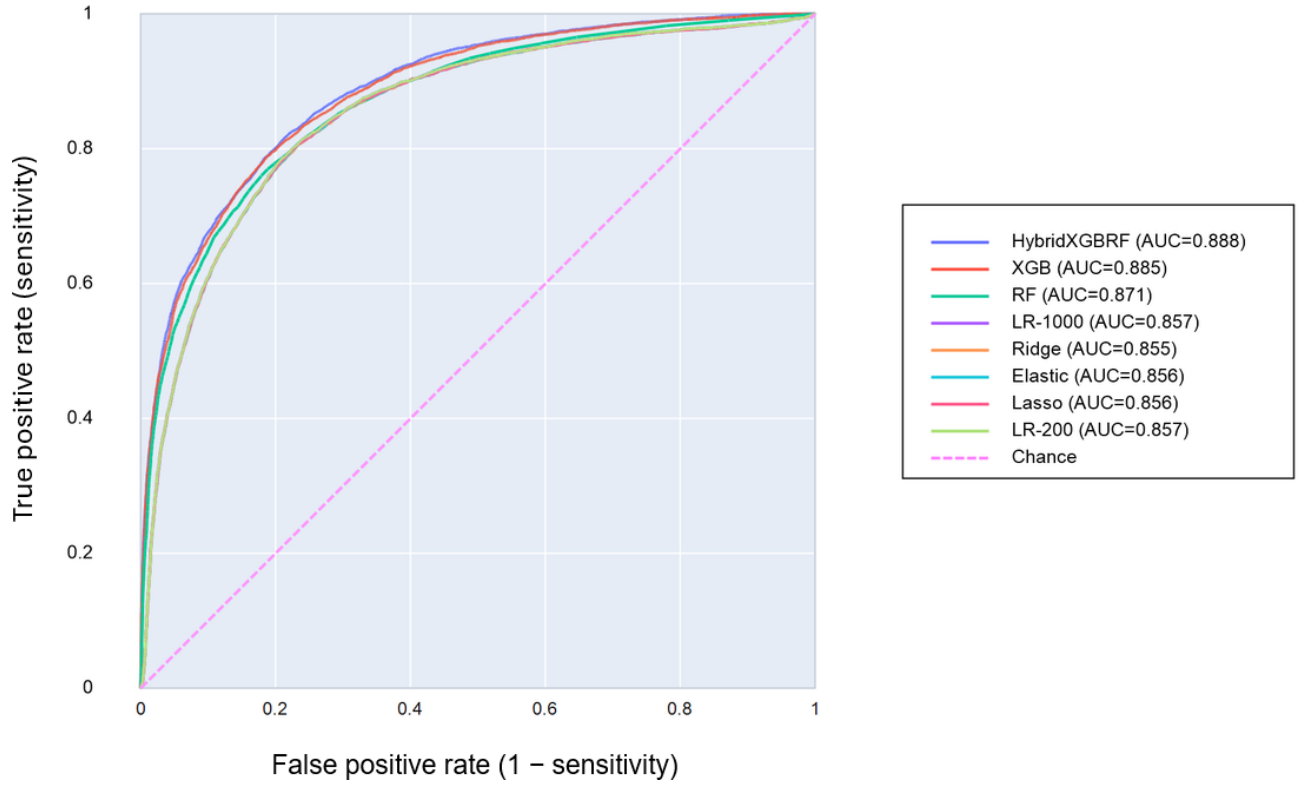
[g]Lasso.

[h]Logistic regression (max-iter =200).

[i]AUROC: area under the receiver operating characteristic curve.

[j]The italics format is used to highlight the best values.

Figure 2. Receiver operating characteristic curve comparison. AUC: area under the curve; HybridXGBRF: hybrid of extreme gradient boosting and random forest; LR: logistic regression; RF: random forest; XGB: extreme gradient boosting.



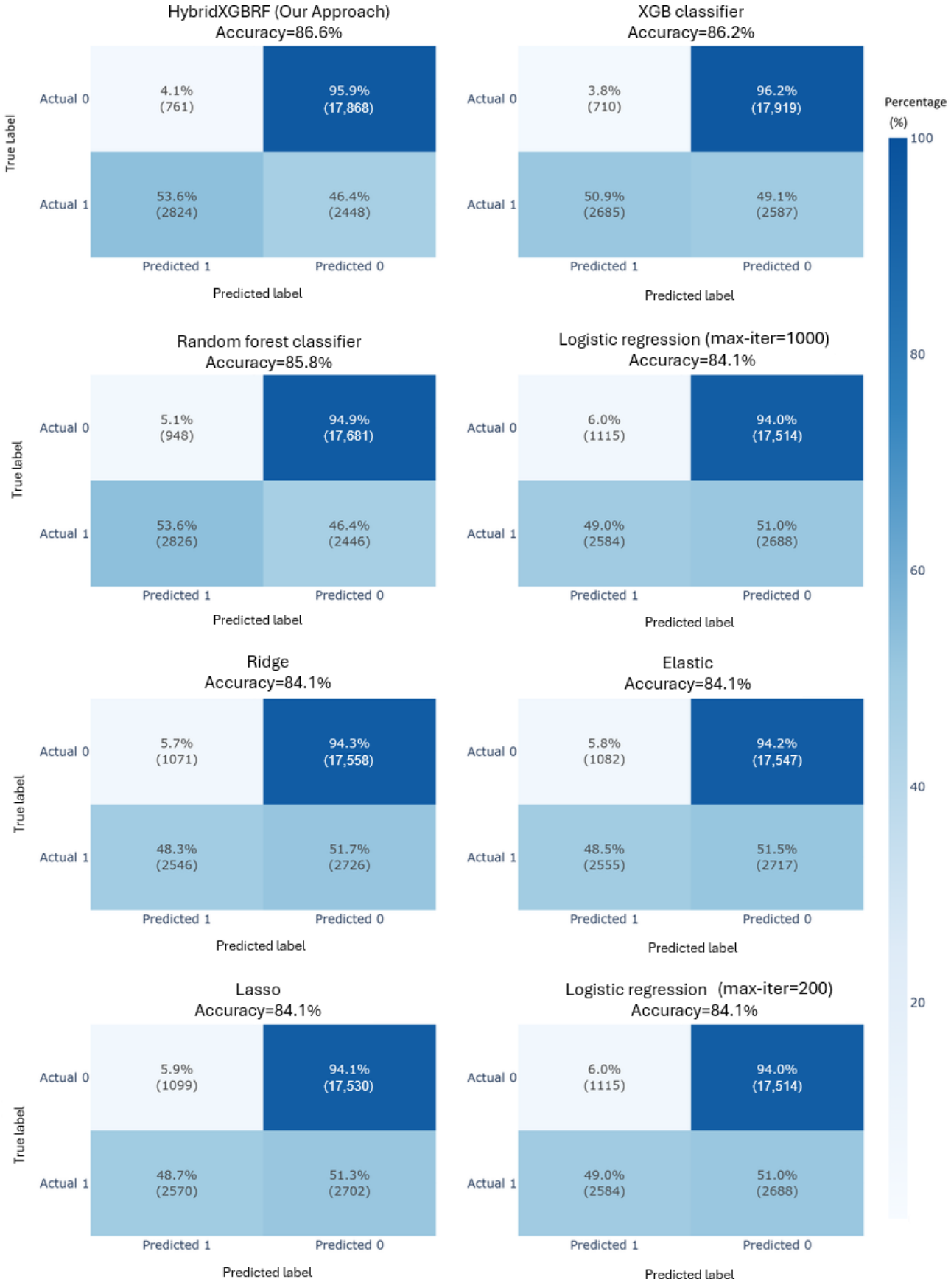
XSL•FO
RenderX

Among all evaluated algorithms, the HybridXGBRF model achieved the best overall internal performance, with the highest AUROC (0.89, 95% CI 0.88 - 0.89), highest F_1-score (0.64, 95% CI 0.63 - 0.65), highest accuracy (0.87, 95% CI 0.86 - 0.87), and lowest Brier score (0.10, 95% CI 0.10 - 0.10). The linear models contrasted this, showing lower classification performance, particularly for sensitivity and F_1-score.

Figure 3 presents the confusion matrices of the evaluated models. Specificity remained consistently high across models (range 0.94 - 0.96), and sensitivity was more modest (range 0.48 - 0.54). Compared with XGB, the HybridXGBRF model identified more positive cases (2824 vs 2685) and achieved higher sensitivity and F_1-score, although XGB generated fewer false-positive predictions (710 vs 761). These findings are aligned with the AUROC, F_1-score, and Brier score results reported in Table 3.

XSL•FO
RenderX

Figure 3. Confusion matrix comparisons between machine learning classifiers. HybridXGBRF: hybrid of extreme gradient boosting and random forest; XGB: extreme gradient boosting.

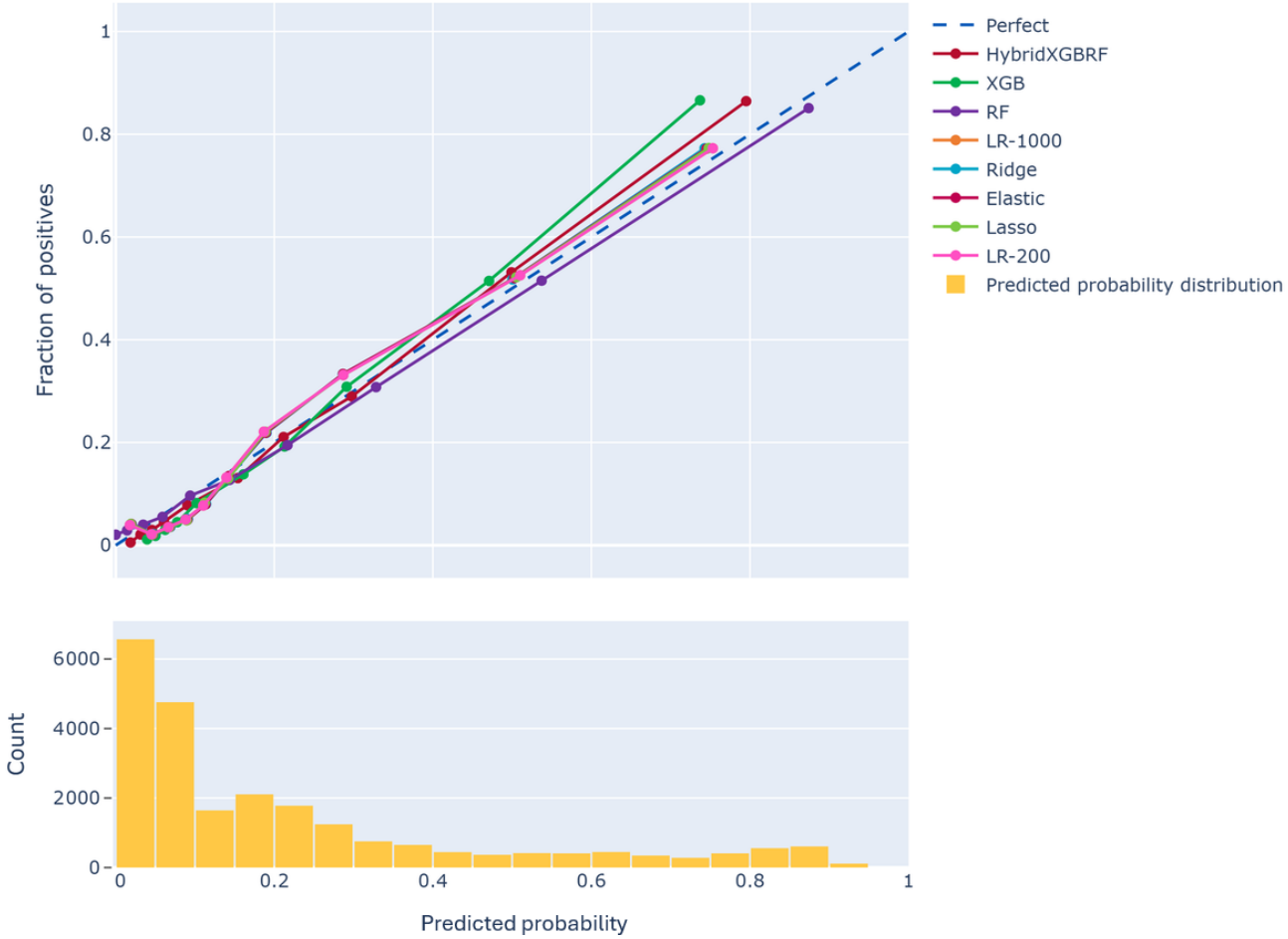


The HybridXGBRF model also demonstrated strong internal calibration. The observed-to-expected (O/E) ratio was 1.00 (Multimedia Appendix 7). Predicted probabilities closely followed the reference 45-degree line across the low-to-moderate

XSL•FO
RenderX

predicted-risk range (Figure 4) based on the risk-decile distribution (Table S1 in Multimedia Appendix 8). A small degree of underestimation was observed in the highest predicted-risk bin (>0.7), while quantitative calibration metrics showed a calibration slope of 1.21 and an intercept of 0.20. Overall, the HybridXGBRF model achieved the lowest internal Brier score (0.10) among the evaluated models (Multimedia Appendix 9), and tree-based models generally showed lower Brier scores than linear models (Table 3). In paired bootstrap comparisons, HybridXGBRF demonstrated a statistically significant but small AUROC improvement over XGB (Multimedia Appendix 10). In the development cohort, the AUROC difference was 0.003 (95% CI 0.002 - 0.005; P<.001).

Figure 4. Calibration plots and distributions of the machine learning models in the development cohort. HybridXGBRF: hybrid of extreme gradient boosting and random forest; LR: logistic regression; RF: random forest; XGB: extreme gradient boosting.



Temporal External Validation, Calibration, and Subgroup Analyses

In the temporal external validation cohort, the models' predictive performance remained stable, indicating good temporal generalizability. The HybridXGBRF model achieved an AUROC of 0.90 (95% CI 0.89 - 0.91), comparable to its internal evaluation performance, as well as achieved an F_1-score of 0.68 (95% CI 0.66 - 0.70), accuracy of 0.85 (95% CI 0.84 - 0.86), and Brier score of 0.11 (95% CI 0.11 - 0.12; Table 4).

XSL•FO
RenderX

Table . Performance metrics of classification algorithms in the external validation dataset[a].

Model metrics	Model 1[b]	Model 2[c]	Model 3[d]	Model 4[e]	Model 5[f]	Model 6[g]	Model 7[h]	Model 8[i]
AUROC[k] (95% CI)	*0.90 (0.89 - 0.91)*	0.89 (0.89 - 0.90)	0.88 (0.87 - 0.89)	0.87 (0.86 - 0.88)	0.87 (0.86 - 0.88)	0.87 (0.86 - 0.88)	0.87 (0.86 - 0.88)	0.87 (0.86 - 0.88)
Precision (95% CI)	*0.86 (0.84 - 0.88)*	*0.86 (0.84 - 0.88)*	0.81 (0.79 - 0.83)	0.77 (0.75 - 0.79)	0.78 (0.75 - 0.80)	0.78 (0.75 - 0.80)	0.77 (0.75 - 0.80)	0.77 (0.75 - 0.79)
Sensitivity or recall (95% CI)	*0.57 (0.55 - 0.59)*	0.54 (0.52 - 0.57)	0.56 (0.54 - 0.59)	0.54 (0.51 - 0.56)	0.53 (0.51 - 0.55)	0.53 (0.51 - 0.55)	0.53 (0.51 - 0.55)	0.54 (0.51 - 0.56)
F_1-score (95% CI)	*0.68 (0.66 - 0.70)*	0.67 (0.65 - 0.68)	0.67 (0.65 - 0.68)	0.63 (0.61 - 0.65)	0.63 (0.61 - 0.65)	0.63 (0.61 - 0.65)	0.63 (0.61 - 0.65)	0.63 (0.61 - 0.65)
Accuracy (95% CI)	*0.85 (0.84 - 0.86)*	0.84 (0.84 - 0.85)	0.84 (0.83 - 0.85)	0.82 (0.81 - 0.83)	0.82 (0.81 - 0.83)	0.82 (0.81 - 0.83)	0.82 (0.81 - 0.83)	0.82 (0.81 - 0.83)
Brier score (95% CI)	*0.11 (0.11 - 0.12)*	0.12 (0.11 - 0.12)	0.12 (0.11 - 0.12)	0.13 (0.12 - 0.13)	0.13 (0.13 - 0.13)	0.13 (0.13 - 0.13)	0.13 (0.12 - 0.13)	0.13 (0.12 - 0.13)

[a]Italicized values are the best ones among the evaluated models.

[b]HybridXGBRF (hybrid of extreme gradient boosting and random forest; proposed approach).

[c]Extreme gradient boosting classifier.

[d]Random forest classifier.

[e]Logistic regression (max-iter=1000).

[f]Ridge.

[g]Elastic.

[h]Lasso.

[i]Logistic regression (max-iter=200).

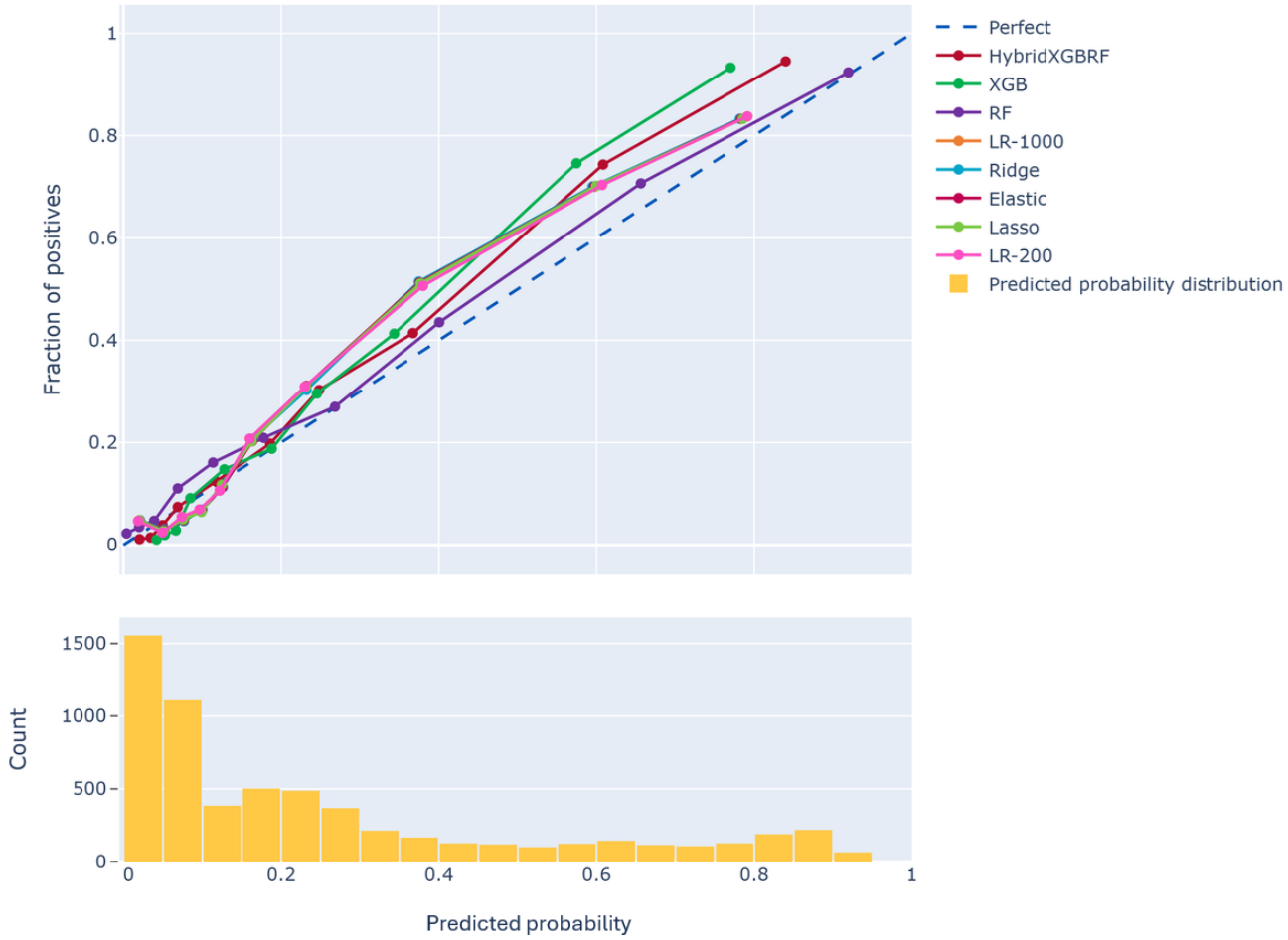
[j]AUROC: area under the receiver operating characteristic curve.

Tree-based models again outperformed linear models in temporal external validation. Tree-based models (vs linear models) showed higher AUROC values (0.88 - 0.90 vs 0.87), higher F_1-scores (0.67 - 0.68 vs 0.63), and higher accuracies (0.84 - 0.85 vs 0.82). In the paired bootstrap comparison, HybridXGBRF showed a small but statistically significant AUROC improvement over XGB in the temporal external validation cohort, with an AUROC difference of 0.004 (95% CI 0.001 - 0.006; *P*=.003; Multimedia Appendix 10).

External validation showed that HybridXGBRF maintained the strongest calibration among the evaluated models, with the lowest Brier score (0.11; Figure 5 and Table 4). Regarding quantitative calibration metrics, the O/E ratio was 1.13, calibration slope of 1.27, and intercept of 0.52 (Multimedia Appendix 7). The detailed risk-decile analysis (Table S2 in Multimedia Appendix 8) showed that the calibration curve closely followed the reference line across most probability ranges, with mild divergence at higher predicted risks. Tree-based models demonstrated better calibration robustness than linear models (Multimedia Appendix 9).

XSL•FO
RenderX

Figure 5. Calibration plots and distributions of the machine learning models in the external validation cohort. HybridXGBRF: hybrid of extreme gradient boosting and random forest; LR: logistic regression; RF: random forest; XGB: extreme gradient boosting.



Subgroup analyses demonstrated that the HybridXGBRF model maintained high and consistent discrimination across sex and age strata, with AUROC values ranging from 0.89 to 0.90 (Table 5). It also showed high discrimination among residents with ADL improvement, with an AUROC of 0.91 (95% CI 0.90 - 0.92), F_1-score of 0.71 (95% CI 0.69 - 0.73), and accuracy of 0.85 (95% CI 0.84 - 0.86). Performance was lower among residents with ADL decline, with an AUROC of 0.71 (95% CI 0.66 - 0.75) and sensitivity of 0.03 (95% CI 0.01 - 0.07).

XSL•FO
RenderX

Table . Subgroup performance of the HybridXGBRF[a] model[b].

Subgroup	AUROC[c] (95% CI)	Precision (95% CI)	Sensitivity or recall (95% CI)	F_1-score (95% CI)	Accuracy (95% CI)	Support, n	Positives, n
ADL[d] improvement	*0.91 (0.90 - 0.92)*	0.86 (0.84 - 0.87)	0.61 (0.58 - 0.63)	0.71 (0.69 - 0.73)	0.85 (0.84 - 0.86)	5524	1658
ADL decline	0.71 (0.66 - 0.75)	0.67 (0.25 - 1.00)	0.03 (0.01 - 0.07)	0.06 (0.02 - 0.12)	0.83 (0.82 - 0.83)	692	123
Age >85 years	*0.90 (0.89 - 0.92)*	0.87 (0.84 - 0.90)	0.60 (0.57 - 0.64)	0.71 (0.68 - 0.74)	0.85 (0.83 - 0.86)	1924	600
Age ≤85 years	0.90 (0.89 - 0.91)	0.85 (0.83 - 0.87)	0.55 (0.52 - 0.58)	0.67 (0.64 - 0.69)	0.85 (0.84 - 0.86)	4292	1181
Female	*0.90 (0.89 - 0.91)*	0.84 (0.80 - 0.87)	0.54 (0.50 - 0.57)	0.65 (0.62 - 0.68)	0.87 (0.86 - 0.88)	3043	714
Male	0.89 (0.88 - 0.91)	0.87 (0.84 - 0.89)	0.59 (0.56 - 0.62)	0.70 (0.68 - 0.72)	0.83 (0.82 - 0.84)	3173	1067

[a]HybridXGBRF: hybrid of extreme gradient boosting and random forest.

[b]Italicized values are the best ones among the evaluated models.

[c]AUROC: area under the receiver operating characteristic curve.

[d]ADL: activities of daily living.

Clinical Utility and Threshold Analyses

Decision-curve analysis showed that the HybridXGBRF model provided higher net benefit than the default strategies. These default strategies involved classifying all residents as high risk or no residents as high risk across a clinically relevant range of threshold probabilities in the temporal external validation cohort (Multimedia Appendix 11).

Alternative threshold analyses demonstrated the expected trade-off between case detection and alert burden (Multimedia Appendix 12). At a predicted-risk threshold of 0.20, the model achieved high sensitivity (0.86) and negative predictive value (0.93), but classified 2659 residents as high risk, corresponding to 42.8% of the temporal external validation cohort. At a threshold of 0.30, the model provided a more balanced profile, with sensitivity of 0.72, specificity of 0.88, positive predictive value of 0.71, and the highest F_1-score among the evaluated thresholds (0.72). At a threshold of 0.50, specificity and positive predictive value increased to 0.961 and 0.86, respectively, but sensitivity decreased to 0.57. The combination of the findings here outlines that lower thresholds prioritized sensitivity at the cost of more high-risk classifications, whereas higher thresholds prioritized specificity and positive predictive value at the cost of reduced sensitivity.

Model Interpretability and Individual-Level Risk Explanation

Figure 6 (ie, showing the distribution of individual SHAP values in a beeswarm plot) and Figure 7 (ie, showing the mean absolute SHAP values with 95% CIs in a forest plot) present SHAP-based explanations for the final HybridXGBRF blended prediction output. Across both visualizations, recent hospitalizations within the past 6 months made the largest overall contribution to the final HybridXGBRF-predicted 6-month mortality risk, with higher hospitalization counts generally contributing positively to predicted risk.

XSL•FO
RenderX

Figure 6. SHAP beeswarm plot of the HybridXGBRF model. ADL: activities of daily living; HybridXGBRF: hybrid of extreme gradient boosting and random forest; SHAP: Shapley Additive Explanations.

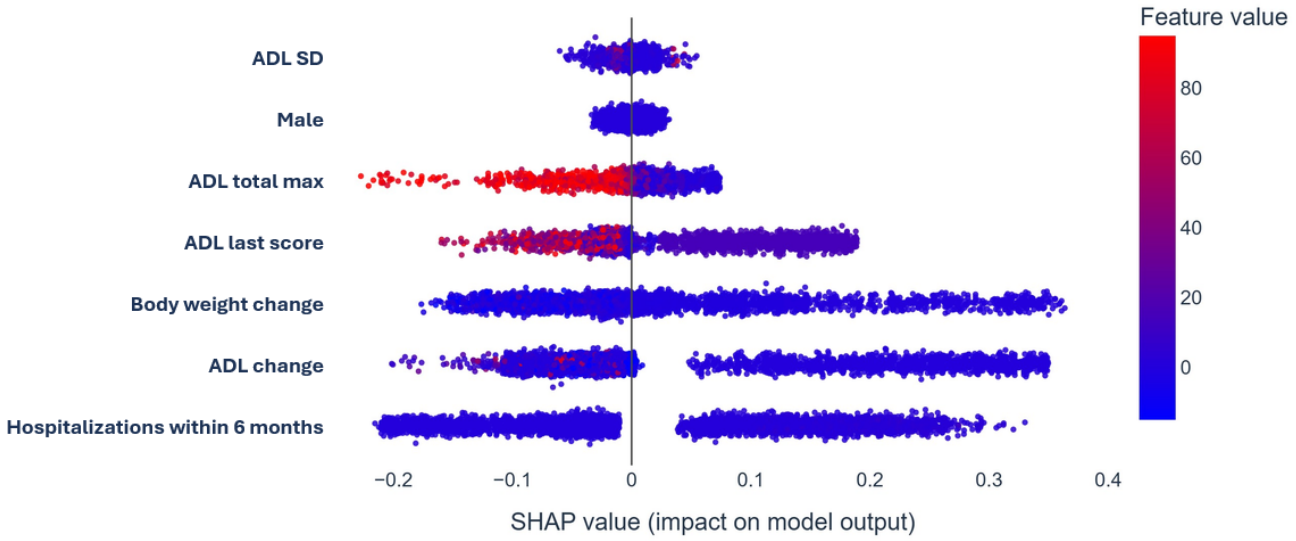
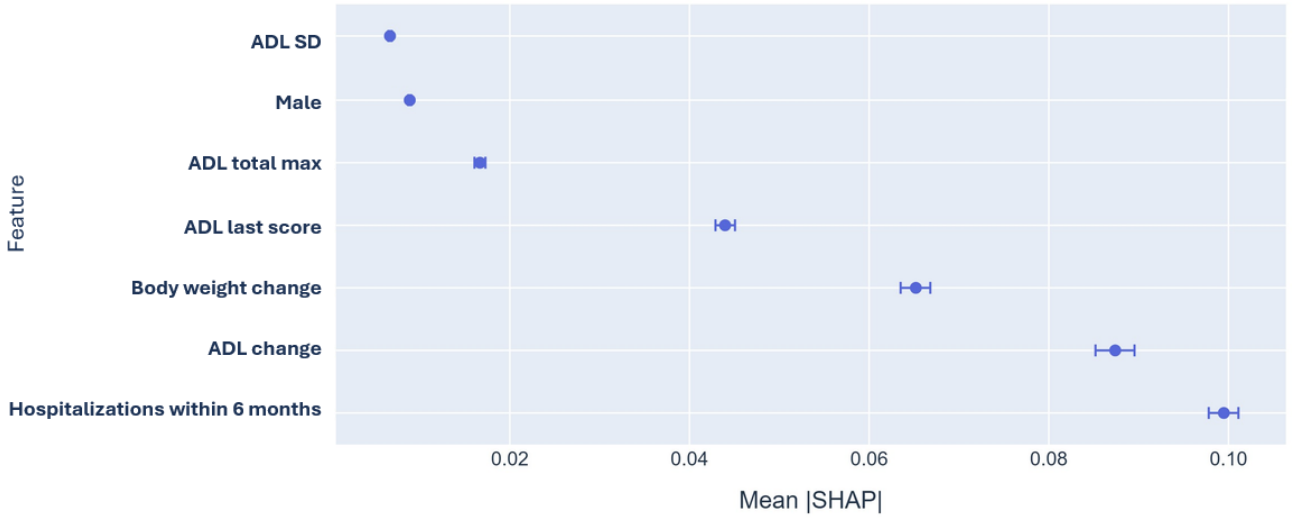


Figure 7. SHAP forest plot of the HybridXGBRF model. ADL: activities of daily living; HybridXGBRF: hybrid of extreme gradient boosting and random forest; SHAP: Shapley Additive Explanations.



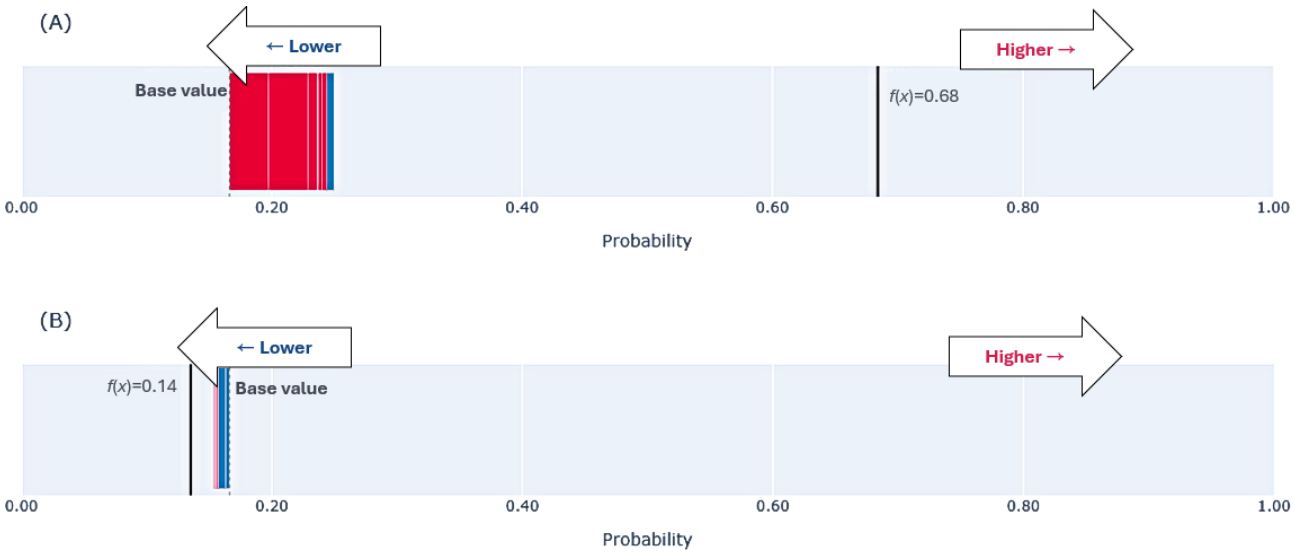
Functional status was also a key determinant of predicted mortality risk. More recent and lower ADL scores, ADL decline, and variability across ADL measurements were associated with increased predicted risk, as reflected by the predominantly positive SHAP values. Conversely, higher maximum ADL scores contributed modestly to lower predicted mortality risk.

Nutritional indicators were similarly influential. Sequential declines in body weight shifted SHAP values upward, indicating heightened predicted risk following nutritional deterioration. Male residents generally showed slightly higher predicted mortality risk than female residents, hence showing that sex contributed to the prediction.

Figure 8 illustrates the SHAP force plots for 2 representative residents. For the positively predicted case (Figure 8A), multiple high-risk features collectively shifted the prediction from the baseline mortality probability (~0.21) to a substantially higher estimated risk (0.68). The contributing factors (ie, pushed the prediction to a higher risk) included impaired functional status, recent functional decline, lower or decreasing body weight, male sex, and recent hospitalization. For the negatively predicted case (Figure 8B), a minimal deviation from the baseline probability (~0.14) was observed, with predominantly protective features (eg, preserved functional ability, stable body weight, and the absence of recent acute care service use) driving the SHAP values toward a lower predicted risk. These examples highlight the manner in which the HybridXGBRF model integrates functional, nutritional, and clinical history features to generate patient-specific risk estimates.

XSL•FO
RenderX

Figure 8. SHAP force plots for 2 representative residents. (A) Positively predicted patient and (B) negatively predicted patient. SHAP: Shapley Additive Explanations.



Sensitivity Analyses

In both cohorts, cumulative or dynamic AUCs remained above 0.89 across months 1 to 5, indicating stable discrimination across the earlier time horizons (Multimedia Appendix 13). In the sensitivity analysis evaluating missingness mechanisms, missingness in subsequent ADL assessment, initial body weight, and Glasgow Coma Scale score was significantly associated with 6-month mortality in both cohorts (adjusted odds ratio ranges 13.7 - 16.6, 3.0 - 3.1, and 1.2 - 1.3, respectively; all $P \leq .01$; Multimedia Appendix 14). Overall documentation missingness varied significantly across the 493 individual facilities (χ^2_{492}=140,786.0; P<.001; Cramer V=0.451), with minimal variation by facility size (7.7% - 8.0%) and greater variation by region (5.5% - 12.1%; Multimedia Appendix 15)[37].

Discussion

Principal Findings

This study developed and temporally externally validated an explainable machine learning model for predicting 6-month all-cause mortality among newly admitted LTCF residents using routinely collected LTCF assessment data. In the temporal external validation cohort, the HybridXGBRF model achieved strong discrimination (ie, AUROC of 0.90) without requiring integration with hospital electronic health records nor data collection beyond standard LTCF assessments. Although HybridXGBRF achieved the highest AUROC among the evaluated models, the absolute improvement over XGB was modest. SHAP analysis pinpointed the most influential predictors as frequent hospitalizations within 6 months, ADL impairment, and weight loss, providing clinically interpretable risk attributions at the individual level. The model also maintained stable discrimination across the sex and age subgroups. Therefore, a parsimonious set of routinely collected LTCF variables, when analyzed through an interpretable machine learning framework, may support risk stratification and structured care planning in Taiwanese long-term care settings.

Interpretation in Context of the Existing Literature

The performance of our model compares favorably with that of established prognostic tools for LTCF populations. The MDS Mortality Risk Index is a validated prognostic tool designed to estimate the risk of 6-month mortality in nursing home residents, and it uses routinely collected clinical and demographic data from the MDS assessment [5,6,8]. The original MDS Mortality Risk Index and its revised versions (MDS Mortality Risk Index-Revised and MDS Mortality Risk Index [version 3]) have demonstrated good discrimination for predicting 6-month mortality (AUROC 0.75 - 0.81) and have been adapted for use in various settings, including nursing homes and community-based care [5,6,8]. Moreover, the Advanced Dementia Prognostic Tool score, specifically designed for residents with advanced dementia, reported an AUROC of 0.67 (95% CI 0.62 - 0.72) [7,9]. More recent approaches include the Hospice Eligibility Prediction Index using hospital diagnoses and age (AUROC 0.838) [11] and Jorissen's comprehensive model using 51 variables (AUROC 0.82), such as individual-level factors, LTCF-level factors, medication history, and health (hospital and medical practitioner) service use [12]. In the 2022 study by Chandra et al [14], an AUROC of 0.84 was reported for predicting mortality among hospitalized patients discharged to LTCFs, albeit this model also required extensive hospitalization data (eg, laboratory results and imaging findings). Importantly, models that depend on detailed hospital-derived diagnoses, medication histories, or longitudinal health service use may not consistently have at their disposal the data (or these data may not be readily linkable) they require to provide their predictions, potentially constraining their scalability and routine implementation in many care settings.

Standing apart from several preceding models, our model achieved strong discrimination using a smaller set of predictors, all of which are routinely documented in standard LTCF admission or early-stay assessments. This parsimony is a deliberate design choice that prioritizes real-world actionability:

XSL•FO
RenderX

by avoiding reliance on hospital-derived diagnoses, medication histories, or granular use data, the model can be executed at, or shortly after, LTCF entry, a point at which timely prognostic awareness may highly meaningfully support ACP and goal-concordant care. This positioning is conceptually aligned with the established Changes in Health, End-Stage Disease, and Signs and Symptoms scale, which uses MDS- or International Resident Assessment Instrument–derived clinical indicators to predict short-term mortality and offers substantial improvement in prognostic precision [38]. Still, while the Changes in Health, End-Stage Disease, and Signs and Symptoms 3.0 [39] scale exhibits moderate discrimination for short-term death (AUROC of 0.72‐0.76 for 30- to 60-day periods), our model achieved a strong AUROC of 0.90 for 6-month mortality in the external validation. What this entails is that by leveraging machine learning to analyze functional trajectories and acute care history (ie, rather than relying primarily on symptomatic clusters), we can provide a resilient and accurate predictive signal supportive of longitudinal care coordination and resource allocation.

Our model was developed and temporally validated in Taiwanese LTCF residents, a population embedded within a long-term care system that includes home-based, community-based, and residential services. Moreover, the baseline characteristics and event rates of our cohort reflect a high-risk institutional population, characterized by substantial functional dependency and 6-month mortality rates of 22.1% in the development cohort and 28.7% in the temporal validation cohort. These mortality rates are broadly consistent with those reported in high-acuity nursing home and long-term care populations internationally [1,2]. Notwithstanding, direct comparisons should be interpreted cautiously, given cross-national differences in admission thresholds, assessment instruments, financing mechanisms, and availability of home- and community-based alternatives.

Most existing mortality prediction tools for LTCF or nursing home residents have been derived from Western cohorts, particularly US MDS-based populations [5-8,11,14]. However, differences in disease prevalence, health care use patterns, long-term care policy, family caregiving availability, cultural attitudes toward end-of-life care, and population health characteristics may limit the generalizability of Western-derived models to Asian LTCF settings [16]. By developing and temporally validating a prediction model using Taiwanese long-term care data, this study contributes region-specific evidence. It outlines that interpretable machine learning can achieve strong discrimination using locally relevant, routinely collected variables. It remains that the greater applicability of the model extends to LTCFs serving functionally dependent and medically complex older adults. External validation and, if necessary, local recalibration would be needed before applying the model to lower-acuity residential care, assisted living settings, or health systems with substantially different care delivery models.

Reliance on machine learning models (vs traditional logistic regression) reflects a deliberate trade-off between interpretability and predictive power. Although logistic regression offers a straightforward coefficient interpretation, it assumes linear relationships and cannot capture complex interactions among predictors without an explicit specification [40]. To address the interpretability challenge inherent in ensemble methods, we used SHAP analysis, which decomposes individual predictions into additive feature contributions grounded in cooperative game theory [35,36]. SHAP analysis revealed that frequent hospitalizations within 6 months, ADL impairment, and weight loss were the most influential predictors; these findings strongly align with established geriatric prognostic frameworks [3,41]. Functional decline (as measured by ADL scores) is a consistent and robust mortality predictor across diverse care settings [3,41]. Weight loss, particularly unintentional loss exceeding 5% of body weight, serves as a clinical marker of frailty, malnutrition, and underlying systemic illnesses [42]. The concordance between the SHAP-identified features and clinically validated risk factors enhances the model's credibility, which may facilitate acceptance among care teams that may otherwise be skeptical of algorithm-driven recommendations.

Additionally, our model was developed and validated exclusively in Taiwanese LTCF residents, specifically and purposefully representing an Asian population. Meanwhile, most existing mortality prediction tools have been derived from Western cohorts, primarily using US MDS registries [5-8,11,14]. The ethnic and geographic differences in disease prevalence, health care use patterns, cultural attitudes toward end-of-life care, and nutritional status may limit the generalizability of Western-derived models to Asian contexts [16]. By developing a prediction model grounded in Taiwanese long-term care data, this study contributes region-specific evidence, emphasizing that interpretable machine learning approaches can achieve strong performance using locally relevant variables.

Subgroup Performance and Algorithmic Fairness

A particularly noteworthy finding was the exceptional performance of the model (accuracy=0.85) among residents who demonstrated ADL improvement after LTCF admission. Functional improvement is typically clinically interpreted as a positive prognostic indicator, being used as a quality metric in institutional care [43,44]. However, a meaningful subset of residents with improving functional status may still face an elevated short-term mortality risk according to our findings, a pattern that may be difficult to identify using conventional clinical assessments alone.

Several mechanisms can explain this paradoxical association. First, some residents may experience transient functional recovery or "terminal lucidity," referring to brief improvements in cognitive or physical function shortly before death, a phenomenon documented in the palliative care literature [45,46]. Second, residents with very low baseline ADL may show relative improvement from institutional rehabilitation while still harboring severe underlying comorbidities or physiological frailty, which may not be fully captured by functional measures alone [47,48]. Similarly, selection effects may operate such that residents admitted to LTCFs in acute decompensation states show functional "improvement" simply by returning toward their prior baseline, yet remain at a high mortality risk due to the severity of the precipitating illness or frailty [47,49].

From an algorithmic fairness perspective, this subgroup represents cases where human clinical judgment is most likely

XSL•FO
RenderX

to underestimate risk, referring to situations sometimes termed "false reassurance" scenarios. The ability of the model to identify elevated mortality risk even in the presence of functional improvement has significant implications for ACP. This holds potential to enable timely goals-of-care conversations and referrals to palliative services for residents whose improving ADL trajectories could otherwise delay such discussions. The potential complementary role of algorithmic decision support is also emphasized here, showing that it does not replace clinical expertise but brings to the surface risk patterns that may be counterintuitive or masked by seemingly positive clinical indicators.

Subgroup analyses demonstrated the model's stable performance across sex and age strata with minimal variation in discrimination (AUROC range 0.89 - 0.90). Therefore, the model does not seem to systematically disadvantage specific demographic groups. This is key if we consider the ongoing concerns regarding algorithmic bias in health care AI [50]. Overall, the evidence from subgroups underlines the model's robustness across key demographic strata, as well as the necessity of a prospective evaluation of whether risk signals in seemingly improving residents translate into better-timed, goal-concordant care when embedded in routine LTCF workflows.

Potential Clinical and Operational Applications

Short-term mortality prediction in LTCFs may support several clinical and operational purposes, including ACP, palliative care assessment, medication review, family meetings, and resource planning [4]. Prognostic information can facilitate serious illness communication and goals-of-care discussions when embedded within structured clinical workflows, although evidence specific to LTCF settings remains lacking [13,51]. In addition, prior work in nursing home residents with advanced dementia has shown that prognostic tools may inform hospice eligibility discussions and care planning, while also underscoring the limitations of relying on prognostic models alone [7]. We put forward that the present model should be interpreted as a risk-stratification tool with potential clinical utility, rather than as evidence that model-guided care will necessarily improve outcomes [19].

Our model has several features that may support future implementation. First, it relies on variables available through LTCF assessment workflows, enabling risk stratification shortly after LTCF entry without requiring hospital record linkage or specialized testing. Second, it generates calibrated probability estimates, not categorical risk labels, allowing facilities to define thresholds according to local resources, resident and family preferences, and palliative care capacity. The threshold and decision-curve analyses also indicate that no single cutoff is optimal for all clinical purposes: lower thresholds may reduce missed high-risk residents but increase the number of residents flagged for review, whereas higher thresholds may reduce alert burden but miss more residents who die within 6 months. Thus, threshold selection should be linked to the intended clinical action and the relative consequences of false-negative versus false-positive classifications [52].

Third, SHAP-based explanations provide interpretable feature contributions that may help clinicians contextualize risk estimates using clinically meaningful domains, such as functional dependency, recent acute-care use, and nutritional decline. Some model features, such as weight loss and frequent hospitalization, should be interpreted as clinically recognizable signals that prompt further assessment, not that provide direction for treatment targets. For example, these signals may lead interdisciplinary teams to review nutritional status, symptom burden, recent care transitions, medication burden, or potentially reversible contributors to functional decline [53,54].

However, clinical utility cannot be assumed from discrimination and calibration alone. These metrics indicate that the model can rank residents by risk and estimate probabilities, while failing to demonstrate that using the model changes clinician behavior, improves care quality, or enhances resident- and family-centered outcomes [17,52]. This brings forth the need for future prospective impact and implementation studies before routine deployment [55]. Such studies should prespecify intended use, risk thresholds, and linked clinical actions; assess workflow integration and implementation outcomes such as acceptability, adoption, appropriateness, feasibility, fidelity, and sustainability; evaluate potential benefits and harms, including ACP uptake, palliative care referral, hospitalization, quality of end-of-life care, family experience, equity, and unintended consequences.

Limitations

This study had several limitations. First, as this was a retrospective analysis of the JUBO database, the findings may be affected by documentation heterogeneity, documentation bias, and unmeasured facility-level effects. Differences in staffing patterns, care protocols, and documentation practices may influence both predictor recording and mortality risk. Facility-level clustering was also not explicitly modeled, and for selected binary or event-based variables, missing values were coded as absence according to the structure of routine LTCF documentation. This later-mentioned approach may conflate true clinical absence with incomplete recording. In sensitivity analyses, missingness in key clinical assessments was significantly associated with 6-month mortality, and facility-level documentation completeness showed significant heterogeneity across facilities and regions (Multimedia Appendices 14 and 15), raising the possibility of informative missingness, not just random documentation gaps. We thus need standardized assessment workflows, routine monitoring of facility-level missingness, and future evaluation of alternative missing-data strategies before prospective implementation of the model.

Second, data availability and temporal measurement limitations may have affected model granularity, benchmark selection, and predictor timing. Although we prioritized routinely available LTCF data to enhance real-world applicability, several potentially high-impact predictors, such as variables related to frailty or specific comorbidities, were excluded because missingness exceeded 30%, which may have reduced the model's clinical granularity. In addition, direct comparison with established MDS- or International Resident Assessment Instrument–based clinical risk scores was not feasible because

XSL•FO
RenderX

the required scoring components are not routinely used in Taiwanese LTCF workflows and were not fully available in the JUBO database. Some predictors also reflected information obtained during routine LTCF assessments after admission rather than strictly at the moment of admission. Although these variables were selected to reflect real-world longitudinal risk monitoring, their use may introduce temporal ambiguity or possible information leakage if the timing of risk estimation is not clearly aligned with predictor availability. Prospective implementation should ensure that each predictor is available before the risk estimate is generated.

Third, post-LTCF discharge vital status data were not consistently available in the JUBO database. Residents discharged or lost to follow-up before completion of the 6-month outcome window without subsequent status information were excluded from the primary analysis. Although the baseline characteristics of included and excluded residents were compared to assess potential selection bias, residual bias cannot be fully excluded because postdischarge mortality could not be directly observed.

Fourth, this study primarily modeled death within 6 months as a binary outcome. Despite adding cumulative or dynamic AUC analyses to evaluate discrimination across earlier time horizons, this approach provides only a complementary time-to-event discrimination assessment, hence not replacing formal survival modeling or fully accounting for censoring and time-to-death. Future studies should compare Cox regression, random survival forests, or gradient-boosted survival models with binary classification approaches to support dynamic risk prediction and updating during LTCF stay.

Fifth, the model's transportability may be affected by cross-national differences in long-term care systems and temporal changes in baseline risk. LTCF admission thresholds vary across countries and may be shaped by financing mechanisms, family caregiving availability, home- and community-based service capacity, and national long-term care policies. In temporal external validation, quantitative calibration analyses showed modest underestimation of observed mortality, particularly at higher predicted-risk levels. This shows that calibration may shift over time or across care settings. External validation, calibration monitoring, and, if necessary, local or temporal recalibration are required before implementation in substantially different care settings.

Finally, the study period encompassed the COVID-19 pandemic, which may have affected LTCF mortality, admission case mix, hospitalization patterns, staffing, infection-control practices, and care delivery. Interpretations of the observed 6-month mortality rates of 22.1% in the development cohort and 28.7% in the temporal validation cohort should thus account for the aforementioned factors. Nevertheless, the model maintained comparable discriminative performance across the development and 2024 temporal validation cohorts, suggesting that the main person-level predictors retained prognostic value despite temporal changes in baseline risk. Validation in future postpandemic cohorts remains warranted.

Future Directions

Future methodological work should go beyond periodic model updating and monitoring of calibration drift, also examining dynamic longitudinal prediction approaches. More recent temporal event prediction methods, including continuous estimation of temporal pattern completion for event prediction, may provide useful frameworks for incorporating repeated ADL assessments, weight trajectories, hospitalization events, falls, resident-reported outcomes, symptom burden, and evolving care needs into updated mortality risk estimates during LTCF stay.

Conclusions

This study developed and temporally externally validated an interpretable machine learning model for predicting 6-month mortality among older LTCF residents in Taiwan using routinely collected LTCF assessment data. The model achieved strong discrimination and acceptable calibration without requiring hospital-based electronic health record linkage, suggesting potential feasibility for risk stratification within LTCF assessment workflows—albeit future prospective impact and implementation studies are recommended before actual deployment in real-world settings. Key predictors, including recent hospitalization, functional dependency and instability, and nutritional decline, were clinically coherent and aligned with geriatric prognostic domains. Prospective implementation studies, formal survival modeling, and external validation in other long-term care systems are needed to determine clinical impact, support dynamic risk updating, and assess broader transportability.

Acknowledgments

The authors thank Smart Ageing Tech Co, Ltd, for their technical support and Ms Chun-Han Tsai for her assistance in this project. The authors also show their appreciation to Professor Sheng-Hsiang Lin and the Biostatistics Consulting Center of National Cheng Kung University Hospital for providing statistical consultation and methodological advice. A portion of the grant funding (NSTC 114-2321-B-006-013) was used to commission the assistance services of Smart Ageing Tech Co, Ltd, for model programming and data processing tasks. The intellectual property of the developed models remains with the research team and the funding agency.

Funding

This study was supported by the National Science and Technology Council of Taiwan (grant NSTC 114-2321-B-006-013).

XSL•FO
RenderX

Data Availability

The datasets generated during and/or analyzed during this study are not publicly available because they contain resident-level administrative and clinical assessment information subject to patient privacy regulations, contractual restrictions, and the data-governance policies of Smart Ageing Tech Co, Ltd, but are available from Smart Ageing Tech Co, Ltd, on reasonable request. Data access is subject to institutional approval, ethics or data-use review, and execution of an appropriate data-use agreement. We provide non–data-containing analysis code, preprocessing specifications, model training and evaluation scripts, random seeds, package versions, and key configuration files in Multimedia Appendix 16 to support future study reproducibility. These materials document the analytic workflow, including data preprocessing, missing-data handling, model training, temporal validation, bootstrap CIs, paired area under the receiver operating characteristic curve comparisons, calibration analyses, decision-curve analysis, threshold analyses, subgroup analyses, and Shapley Additive Explanations–based interpretability.

Authors' Contributions

Conceptualization: YTL, YLW, YC Yu
Methodology: YTL, YLW, YCT
Software: YCT, YWC (Smart Ageing Tech Co, Ltd)
Validation: YCT, YWC (Department of Nursing).
Formal analysis: YCT, YLW, YWC (Smart Ageing Tech Co, Ltd)
Investigation: YLW, YWC (Department of Nursing), YC Yu
Resources: YC Yang, YWC (Smart Ageing Tech Co, Ltd)
Data curation: YLW, YC Yu, YWC (Smart Ageing Tech Co, Ltd)
Visualization: YCT, YLW
Supervision: YC Yang
Project administration: YTL, YLW
Writing—original draft: YLW, YTL
Writing—review and editing: YTL, YC Yang, YLW, YCT, YC Yu, and YWC (Department of Nursing)

Conflicts of Interest

YCT serves as a paid consultant for Jubo Health Technologies, Inc, albeit this role is independent of the current study's funding. YWC (Smart Ageing Tech Co, Ltd) is a full-time employee of Jubo Health Technologies, Inc. The machine learning models evaluated in this study were developed by the research team, while the technical implementation and data extraction services were contracted to Jubo Health Technologies, Inc, using research grant funds. Jubo Health Technologies, Inc, had no role in the study design, analysis, interpretation of data, or decision to publish the results.

Multimedia Appendix 1
Literature review of predictors of mortality in long-term care settings.
[DOCX File, 34 KB - aging_v9i1e94567_app1.docx]

Multimedia Appendix 2
Candidate predictors.
[DOCX File, 40 KB - aging_v9i1e94567_app2.docx]

Multimedia Appendix 3
Assessment of data missingness across study cohorts.
[DOCX File, 34 KB - aging_v9i1e94567_app3.docx]

Multimedia Appendix 4
Model candidates and qualitative comparison with key hyperparameters.
[DOCX File, 34 KB - aging_v9i1e94567_app4.docx]

Multimedia Appendix 5
Definitions of evaluation metrics.
[DOCX File, 33 KB - aging_v9i1e94567_app5.docx]

Multimedia Appendix 6
Characteristics of development cohort and excluded participants.
[DOCX File, 50 KB - aging_v9i1e94567_app6.docx]

XSL•FO
RenderX

Multimedia Appendix 7
Quantitative calibration metrics of the predictive models in internal cross-validation and temporal external validation cohorts.
[DOCX File, 24 KB - aging_v9i1e94567_app7.docx]

Multimedia Appendix 8
Risk-decile calibration tables of the hybrid of extreme gradient boosting and random forest model across validation cohorts.
[DOCX File, 24 KB - aging_v9i1e94567_app8.docx]

Multimedia Appendix 9
Comparison of quantitative calibration metrics across all evaluated models.
[DOCX File, 21 KB - aging_v9i1e94567_app9.docx]

Multimedia Appendix 10
Paired bootstrap test for hybrid of extreme gradient boosting and random forest and extreme gradient boosting.
[DOCX File, 20 KB - aging_v9i1e94567_app10.docx]

Multimedia Appendix 11
Decision curve analysis.
[DOCX File, 220 KB - aging_v9i1e94567_app11.docx]

Multimedia Appendix 12
Evaluation of alternative decision thresholds for the mortality prediction model.
[DOCX File, 23 KB - aging_v9i1e94567_app12.docx]

Multimedia Appendix 13
Cumulative and dynamic area under the curve for time-to-event analysis.
[DOCX File, 21 KB - aging_v9i1e94567_app13.docx]

Multimedia Appendix 14
Association between missing key clinical assessments and 6-month mortality (informative missingness analysis).
[DOCX File, 21 KB - aging_v9i1e94567_app14.docx]

Multimedia Appendix 15
Overall documentation missingness and clinical outcomes stratified by facility size and geographic region.
[DOCX File, 20 KB - aging_v9i1e94567_app15.docx]

Multimedia Appendix 16
Analysis source code and analytic scripts for model reproducibility.
[DOCX File, 24 KB - aging_v9i1e94567_app16.docx]

Checklist 1
TRIPOD-AI checklist.
[DOCX File, 47 KB - aging_v9i1e94567_app17.docx]

References

1. Ferrah N, Ibrahim JE, Kipsaina C, Bugeja L. Death following recent admission into nursing home from community living: a systematic review into the transition process. J Aging Health 2018 Apr;30(4):584-604. [doi: 10.1177/0898264316686575] [Medline: 28553803]
2. Morris JN, Howard EP, Schachter E, Pešić V, Laytham AK, Burney SN. Risk of death among nursing home residents: a cross-national perspective. J Am Med Dir Assoc 2024 Sep;25(9):105142. [doi: 10.1016/j.jamda.2024.105142] [Medline: 38986685]
3. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. JAMA 2012 Jan 11;307(2):182-192. [doi: 10.1001/jama.2011.1966] [Medline: 22235089]
4. Boyd C, Smith CD, Masoudi FA, et al. Decision making for older adults with multiple chronic conditions: executive summary for the American Geriatrics Society guiding principles on the care of older adults with multimorbidity. J Am Geriatr Soc 2019 Apr;67(4):665-673. [doi: 10.1111/jgs.15809] [Medline: 30663782]

XSL•FO
RenderX

5. Niznik JD, Zhang S, Mor MK, et al. Adaptation and initial validation of Minimum Data Set (MDS) Mortality Risk Index to MDS Version 3.0. J Am Geriatr Soc 2018 Dec;66(12):2353-2359. [doi: 10.1111/jgs.15579] [Medline: 30335184]
6. Porock D, Parker-Oliver D, Petroski GF, Rantz M. The MDS Mortality Risk Index: the evolution of a method for predicting 6-month mortality in nursing home residents. BMC Res Notes 2010 Jul 16;3:200. [doi: 10.1186/1756-0500-3-200] [Medline: 20637076]
7. Mitchell SL, Miller SC, Teno JM, Kiely DK, Davis RB, Shaffer ML. Prediction of 6-month survival of nursing home residents with advanced dementia using ADEPT vs hospice eligibility guidelines. JAMA 2010 Nov 3;304(17):1929-1935. [doi: 10.1001/jama.2010.1572] [Medline: 21045099]
8. Porock D, Oliver DP, Zweig S, et al. Predicting death in the nursing home: development and validation of the 6-month Minimum Data Set mortality risk index. J Gerontol A Biol Sci Med Sci 2005 Apr;60(4):491-498. [doi: 10.1093/gerona/60.4.491] [Medline: 15933390]
9. Mitchell SL, Kiely DK, Hamel MB, Park PS, Morris JN, Fries BE. Estimating prognosis for nursing home residents with advanced dementia. JAMA 2004 Jun 9;291(22):2734-2740. [doi: 10.1001/jama.291.22.2734] [Medline: 15187055]
10. Flacker JM, Kiely DK. Mortality-related factors and 1-year survival in nursing home residents. J Am Geriatr Soc 2003 Feb;51(2):213-221. [doi: 10.1046/j.1532-5415.2003.51060.x] [Medline: 12558718]
11. Levy C, Kheirbek R, Alemi F, et al. Predictors of six-month mortality among nursing home residents: diagnoses may be more predictive than functional disability. J Palliat Med 2015 Feb;18(2):100-106. [doi: 10.1089/jpm.2014.0130] [Medline: 25380219]
12. Jorissen RN, Wesselingh SL, Whitehead C, et al. Predictors of mortality shortly after entering a long-term care facility. Age Ageing 2024 May 1;53(5):afae098. [doi: 10.1093/ageing/afae098] [Medline: 38773946]
13. Lee SK, Ahn J, Shin JH, Lee JY. Application of machine learning methods in nursing home research. Int J Environ Res Public Health 2020 Aug 27;17(17):6234. [doi: 10.3390/ijerph17176234] [Medline: 32867250]
14. Chandra A, Takahashi PY, McCoy RG, et al. Risk prediction model for 6-month mortality for patients discharged to skilled nursing facilities. J Am Med Dir Assoc 2022 Aug;23(8):1403-1408. [doi: 10.1016/j.jamda.2022.01.069] [Medline: 35227666]
15. Susnjak T, Griffin E. Towards clinical prediction with transparency: an explainable AI approach to survival modelling in residential aged care. Comput Methods Programs Biomed 2025 May;263:108653. [doi: 10.1016/j.cmpb.2025.108653] [Medline: 39970690]
16. Zhang S, Zhang K, Chen Y, Wu C. Prediction models of all-cause mortality among older adults in nursing home setting: a systematic review and meta-analysis. Health Sci Rep 2023 Jun;6(6):e1309. [doi: 10.1002/hsr2.1309] [Medline: 37275670]
17. Hsu HC, Chen CF. LTC 2.0: the 2017 reform of home- and community-based long-term care in Taiwan. Health Policy 2019 Oct;123(10):912-916. [doi: 10.1016/j.healthpol.2019.08.004] [Medline: 31455563]
18. Huang JJ, Lin KC, Li IC. Service needs of residents in community-based long-term care facilities in northern Taiwan. J Clin Nurs 2008 Jan;17(1):99-108. [doi: 10.1111/j.1365-2702.2007.01974.x] [Medline: 18088262]
19. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
20. Teasdale G, Maas A, Lecky F, Manley G, Stocchetti N, Murray G. The Glasgow Coma Scale at 40 years: standing the test of time. Lancet Neurol 2014 Aug;13(8):844-854. [doi: 10.1016/S1474-4422(14)70120-6] [Medline: 25030516]
21. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
22. Park SH, Kwak MJ. Performance of the Geriatric Depression Scale-15 with older adults aged over 65 years: an updated review 2000-2019. Clin Gerontol 2021;44(2):83-96. [doi: 10.1080/07317115.2020.1839992] [Medline: 33164674]
23. Ensrud KE, Ewing SK, Taylor BC, et al. Comparison of 2 frailty indexes for prediction of falls, disability, fractures, and death in older women. Arch Intern Med 2008 Feb 25;168(4):382-389. [doi: 10.1001/archinternmed.2007.113] [Medline: 18299493]
24. Kaiser MJ, Bauer JM, Ramsch C, et al. Validation of the Mini Nutritional Assessment short-form (MNA-SF): a practical tool for identification of nutritional status. J Nutr Health Aging 2009 Nov;13(9):782-788. [doi: 10.1007/s12603-009-0214-7] [Medline: 19812868]
25. Salvi F, Miller MD, Grilli A, et al. A manual of guidelines to score the modified cumulative illness rating scale and its validation in acute hospitalized elderly patients. J Am Geriatr Soc 2008 Oct;56(10):1926-1931. [doi: 10.1111/j.1532-5415.2008.01935.x] [Medline: 18811613]
26. Digitale J, Franzon D, Pletcher MJ, McCulloch CE, Gennatas ED. Methods for addressing missingness in electronic health record data for clinical prediction models: comparative evaluation. JMIR Med Inform 2025 Nov 14;13:e79307. [doi: 10.2196/79307] [Medline: 41237368]
27. Wells BJ, Chagin KM, Nowacki AS, Kattan MW. Strategies for handling missing data in electronic health record derived data. EGEMS (Wash DC) 2013;1(3):1035. [doi: 10.13063/2327-9214.1035] [Medline: 25848578]
28. Wu K, Dahlem D, Hane C, Halperin E, Zou J. Collecting data when missingness is unknown: a method for improving model performance given under-reporting in patient populations. Presented at: Conference on Health, Inference, and Learning; Jun 22-24, 2023 URL: https://proceedings.mlr.press/v209/wu23b.html [accessed 2026-08-24]

XSL•FO
RenderX

29. Cabello-Solorzano K, de Araujo IO. The impact of data normalization on the accuracy of machine learning algorithms: a comparative analysis. Presented at: 18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023); Sep 5-7, 2023. [doi: 10.1007/978-3-031-42536-3_33]
30. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. Presented at: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 13-17, 2016. [doi: 10.1145/2939672.2939785]
31. Breiman L. Random forests. Mach Learn 2001 Oct;45(1):5-32. [doi: 10.1023/A:1010933404324]
32. Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 1970 Feb;12(1):55-67. [doi: 10.1080/00401706.1970.10488634]
33. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc B 1996 Jan 1;58(1):267-288. [doi: 10.1111/j.2517-6161.1996.tb02080.x]
34. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc B 2005 Apr 1;67(2):301-320. [doi: 10.1111/j.1467-9868.2005.00503.x]
35. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2020 Jan;2(1):56-67. [doi: 10.1038/s42256-019-0138-9] [Medline: 32607472]
36. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Presented at: 31st International Conference on Neural Information Processing Systems; Dec 4-9, 2017 URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf [accessed 2026-08-24]
37. Park EG, Hahn S. An approach to exploring patterns of imbalance and potential missingness in reports of the randomized baseline values for primary outcomes measurable at baseline in randomized controlled trials for meta-analyses. BMC Med Res Methodol 2022 May 28;22(1):154. [doi: 10.1186/s12874-022-01620-x] [Medline: 35643437]
38. Hirdes JP, Frijters DH, Teare GF. The MDS-CHESS scale: a new measure to predict mortality in institutionalized older people. J Am Geriatr Soc 2003 Jan;51(1):96-100. [doi: 10.1034/j.1601-5215.2002.51017.x] [Medline: 12534853]
39. Ogarek JA, McCreedy EM, Thomas KS, Teno JM, Gozalo PL. Minimum Data Set Changes in Health, End-Stage Disease and Symptoms and Signs Scale: a revised measure to predict mortality in nursing home residents. J Am Geriatr Soc 2018 May;66(5):976-981. [doi: 10.1111/jgs.15305] [Medline: 29500822]
40. Stoltzfus JC. Logistic regression: a brief primer. Acad Emerg Med 2011 Oct;18(10):1099-1104. [doi: 10.1111/j.1553-2712.2011.01185.x] [Medline: 21996075]
41. Soh CH, Lim WK, Maier AB. Predictors for the transitions of poor clinical outcomes among geriatric rehabilitation inpatients. J Am Med Dir Assoc 2022 Nov;23(11):1800-1806. [doi: 10.1016/j.jamda.2022.05.019] [Medline: 35760091]
42. Gaddey HL, Holder KK. Unintentional weight loss in older adults. Am Fam Physician 2021 Jul 1;104(1):34-40. [Medline: 34264616]
43. Downer B, Pritchard K, Thomas KS, Ottenbacher K. Improvement in activities of daily living during a nursing home stay and one‐year mortality among older adults with sepsis. J Am Geriatr Soc 2021 Apr;69(4):938-945. [doi: 10.1111/jgs.16915]
44. Wysocki A, Thomas KS, Mor V. Functional improvement among short-stay nursing home residents in the MDS 3.0. J Am Med Dir Assoc 2015 Jun 1;16(6):470-474. [doi: 10.1016/j.jamda.2014.11.018] [Medline: 25659622]
45. Nahm M, Greyson B, Kelly EW, Haraldsson E. Terminal lucidity: a review and a case collection. Arch Gerontol Geriatr 2012;55(1):138-142. [doi: 10.1016/j.archger.2011.06.031] [Medline: 21764150]
46. Silva TO, Levy K, Kerr CW. End-of-life experiences in patients: a scoping review of types, characteristics, and implications for the mind-brain relationship. Int Rev Psychiatry 2025;37(2):142-156. [doi: 10.1080/09540261.2025.2503726] [Medline: 40627515]
47. Carrasco-Ribelles LA, Roso-Llorach A, Cabrera-Bean M, et al. Dynamics of multimorbidity and frailty, and their contribution to mortality, nursing home and home care need: a primary care cohort of 1 456 052 ageing people. EClinicalMedicine 2022 Oct;52:101610. [doi: 10.1016/j.eclinm.2022.101610] [Medline: 36034409]
48. Vetrano DL, Damiano C, Tazzeo C, et al. Multimorbidity patterns and 5-year mortality in institutionalized older adults. J Am Med Dir Assoc 2022 Aug;23(8):1389-1395. [doi: 10.1016/j.jamda.2022.01.067] [Medline: 35218731]
49. Matusik P, Tomaszewski K, Chmielowska K, et al. Severe frailty and cognitive impairment are related to higher mortality in 12-month follow-up of nursing home residents. Arch Gerontol Geriatr 2012;55(1):22-24. [doi: 10.1016/j.archger.2011.06.034] [Medline: 21764151]
50. Ferryman K, Mackintosh M, Ghassemi M. Considering biased data as informative artifacts in AI-assisted health care. N Engl J Med 2023 Aug 31;389(9):833-838. [doi: 10.1056/NEJMra2214964] [Medline: 37646680]
51. Chi S, Kim S, Reuter M, et al. Advanced care planning for hospitalized patients following clinician notification of patient mortality by a machine learning algorithm. JAMA Netw Open 2023 Apr 3;6(4):e238795. [doi: 10.1001/jamanetworkopen.2023.8795] [Medline: 37071421]
52. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(6):565-574. [doi: 10.1177/0272989X06295361] [Medline: 17099194]
53. Tamura BK, Bell CL, Masaki KH, Amella EJ. Factors associated with weight loss, low BMI, and malnutrition among nursing home patients: a systematic review of the literature. J Am Med Dir Assoc 2013 Sep;14(9):649-655. [doi: 10.1016/j.jamda.2013.02.022] [Medline: 23639716]

XSL•FO
RenderX

54. Spector WD, Limcangco R, Williams C, Rhodes W, Hurd D. Potentially avoidable hospitalizations for elderly long-stay residents in nursing homes. Med Care 2013 Aug;51(8):673-681. [doi: 10.1097/MLR.0b013e3182984bff] [Medline: 23703648]
55. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]

Abbreviations

ACP: advance care planning
ADL: activities of daily living
AUC: area under the curve
AUROC: area under the receiver operating characteristic curve
HybridXGBRF: hybrid of extreme gradient boosting and random forest
LTCF: long-term care facility
MDS: minimum dataset
SHAP: Shapley Additive Explanations
TRIPOD-AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis-Artificial Intelligence
XGB: extreme gradient boosting

Edited by A Garcia; submitted 03.Mar.2026; peer-reviewed by C Exmann, N Itzhak; revised version received 13.Jul.2026; accepted 13.Aug.2026; published 16.Sep.2026.

Please cite as:
Tsai YC, Wu YL, Yu YC, Chen YW, Chen YW, Yang YC, Lo YT
Development and Temporal External Validation of a Parsimonious, Interpretable Machine Learning Model for Predicting 6-Month Mortality in Long-Term Care Facilities: A Retrospective Cohort Study
JMIR Aging 2026;9:e94567
URL: https://aging.jmir.org/2026/1/e94567
doi:10.2196/94567

© Yun-Cheng Tsai, Yi-Lin Wu, Yung-Chen Yu, Yi-Wen Chen, Yi-Wen Chen, Yi-Ching Yang, Yu-Tai Lo. Originally published in JMIR Aging (https://aging.jmir.org), 16.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Smart Wheeled Walker With Integrated Physiological Monitoring to Improve Mobility in Older Adults: Randomized Crossover Trial

Phurichaya Werasirirat[1], PhD; Pornpimol Muanjai[1], PhD; Oranat Sukkho[1], PhD; Nongnuch Luangpon[1], PhD; Audrius Snieckus[2], PhD; Juntip Namsawang[1], PhD

[1]Department of Physical Therapy, Faculty of Allied Health Sciences, Burapha University, 169 Longhard Bangsaen Road, Muang District, Saen Suk, Chonburi, Thailand

[2]Institute of Sport Science and Innovations, Lithuanian Sports University, Kaunas, Kaunas, Lithuania

Corresponding Author:
Juntip Namsawang, PhD
Department of Physical Therapy, Faculty of Allied Health Sciences, Burapha University, 169 Longhard Bangsaen Road, Muang District, Saen Suk, Chonburi, Thailand

Abstract

Background: Mobility limitations in older adults are associated with an increased physiological cost of walking, reduced functional independence, and elevated fall risk. Conventional wheeled walkers improve stability but lack real-time physiological monitoring and safety feedback. Digital health–enabled assistive devices integrating physiological monitoring and fall detection may enhance mobility performance and safety.

Objective: This study aimed to evaluate the short-term effects of a smart wheeled walker equipped with real-time physiological monitoring and fall detection compared with a standard wheeled walker in improving walking efficiency, dynamic balance, and fear of falling among community-dwelling older adults.

Methods: A single-blind randomized crossover trial was conducted in 30 community-dwelling older adults aged 65 to 80 years with mild balance impairment (Timed Up and Go >13.5 seconds). Participants completed walking trials using both a standard wheeled walker and a smart wheeled walker in a randomized order, with a 30-minute washout period. The smart walker integrated photoplethysmography-based heart rate (HR) monitoring, pulse oximetry (oxygen saturation [SpO_2]), and a tilt-based fall detection system with mobile alert functionality. The primary outcome was walking efficiency assessed using the Physiological Cost Index (PCI). Secondary outcomes included dynamic balance (Expanded Timed Up and Go [ETUG]) and fear of falling (Falls Efficacy Scale-International). Statistical analysis was performed using paired comparisons with a significance level of $P<.05$.

Results: Thirty (100%) participants completed both intervention conditions and were included in the final analysis. Walking speed was significantly higher with the smart wheeled walker than the standard wheeled walker (mean 28.77, SD 11.94 vs mean 18.14, SD 14.55 m/min; mean difference −10.63 m/min, 95% CI −17.51 to −3.75 m/min; $P<.001$). The PCI was significantly lower with the smart wheeled walker (mean 0.36, SD 0.38 vs mean 0.69, SD 1.23 beats/m; mean difference 0.33 beats/m, 95% CI 0.06 - 0.60 beats/m; $P=.02$), indicating improved walking efficiency. Participants also completed the ETUG test significantly faster with the smart wheeled walker (mean 50.06, SD 24.50 vs mean 64.47, SD 23.04 s; mean difference 14.41 s, 95% CI 2.12 - 26.70 s; $P=.048$), indicating improved dynamic balance. Fear of falling did not differ significantly between walker conditions (mean 31.56, SD 9.58 vs mean 32.91, SD 10.37; mean difference 1.35, 95% CI −3.81 to 6.51; $P=.62$).

Conclusions: A smart wheeled walker integrating real-time physiological monitoring and fall detection significantly improved walking efficiency and dynamic balance during short-term supervised testing in community-dwelling older adults. These findings support the potential clinical value of digital health–enabled mobility aids. Further longitudinal studies conducted in real-world settings are warranted to evaluate long-term mobility, fall prevention, independent use, and comprehensive device safety.

Trial Registration: Thai Clinical Trials Registry TCTR20250302008; https://www.thaiclinicaltrials.org/show/TCTR20250302008

(JMIR Aging 2026;9:e99522) doi:10.2196/99522

KEYWORDS

older adults; smart wheeled walker; assistive technology; Physiological Cost Index; dynamic balance

XSL•FO
RenderX

Introduction

Over the past decade, rapid global demographic aging has posed significant challenges to public health systems. By 2050, the global population aged 60 years and older is projected to reach approximately 2.1 billion, marking a dramatic rise in demand for older adult care and preventive strategies. Mobility impairments and balance disorders are key contributors to fall risk in older adults, with nearly one-third of community-dwelling seniors experiencing at least 1 fall annually—a proportion that increases to approximately 50% for those aged over 80 years [1,2]. Falls among older adults not only precipitate functional decline and loss of independence but are also a leading cause of injury-related mortality in this population [3].

In addition to biomechanical limitations, physiological responses during ambulation, such as elevated heart rate (HR) and increased oxygen consumption, have been shown to rise significantly among older adults with impaired mobility [4]. These responses reflect greater cardiovascular and muscular effort required for walking, which may contribute to fatigue and instability. A higher physiological cost of walking has also been associated with reduced gait efficiency and a greater risk of falls, particularly among frail individuals. Real-time monitoring of HR and oxygen saturation (SpO_2) during the use of a walker may thus serve as a meaningful indicator of exertion and functional capacity, enabling targeted interventions to optimize mobility assistance and reduce physiological strain [5]. Among current clinical strategies, the use of mobility aids such as wheeled walkers has proven effective in enhancing gait stability and reducing musculoskeletal burden [6,7]. Older adults using wheeled walkers exhibited significantly improved backward walking performance, including higher speed, greater step frequency, and fewer steps, indicating better motor control and energy efficiency [8]. The type of walking aid used can influence walking performance and physiological energy cost in older adults. These findings reinforce the importance of wheeled walkers in supporting older adults with compromised mobility [9].

Despite their benefits, traditional wheeled walkers have notable limitations related to maneuverability, brake use, and navigation in challenging situations, such as negotiating obstacles, slopes, and doorways [10,11]. Two-wheeled walkers were associated with a higher mean Physiological Cost Index (PCI) and lower gait speed than four-wheeled walkers, suggesting that walker configuration may influence functional walking performance in older adults [12]. Moreover, lateral falls during turning were more frequent among 2-wheeled walker users, while 4-wheeled walker users experienced more backward falls during weight shifting [13]. These incidents highlight the persistent risk of falls even when assistive devices are used, often due to device limitations such as poor maneuverability and inadequate lateral stability. Advanced walker designs have sought to address these challenges. An instrumented 4-wheeled rollator incorporating load cells, pressure-sensing insoles, and 3D motion capture was developed to objectively assess user stability and device loading during different walking scenarios. These findings underscore the need for assistive technologies that can accommodate the complex mobility demands encountered by older adult users [14].

Although previous studies have reported the development and technical capabilities of smart walker technologies, evidence regarding their effects on walking efficiency, dynamic balance, physiological responses, and fear of falling in community-dwelling older adults remains limited [15,16]. Moreover, few randomized crossover studies have directly compared a smart wheeled walker with a conventional wheeled walker under supervised clinical conditions [16]. To address these knowledge gaps, we developed a novel smart wheeled walker integrating safety features, ergonomic adjustability, and physiological monitoring capabilities. The smart wheeled walker integrates 4 complementary functional components. First, the ergonomic frame, height-adjustable handles, freely rotating wheels, and braking mechanism were designed to improve maneuverability, walking efficiency, and dynamic balance. Second, the integrated HR and SpO_2 monitoring system provides real-time physiological information to support assessment of the physiological cost of walking and cardiopulmonary responses during ambulation. Third, the tilt sensor–based fall detection system and mobile alert function enhance safety monitoring by notifying caregivers when excessive tilt or a fall is detected. Finally, the foldable and portable design was intended to facilitate transportation and storage of the device. These design features may support future implementation of smart mobility technologies, although their real-world usability and acceptability were not evaluated in the present study. HR monitoring is particularly relevant to the present study because the PCI, the primary outcome measure of walking efficiency, is calculated from the difference between walking HR and resting HR divided by walking speed. Consequently, continuous HR monitoring provides the physiological data required for objective estimation of the physiological cost of walking during functional ambulation [17]. In addition, SpO_2 monitoring serves as an adjunct physiological safety indicator for monitoring cardiopulmonary responses during walking. These monitoring functions support physiological assessment and safety monitoring but are not intended to directly improve walking efficiency, dynamic balance, or fear of falling. Rather, improvements in walking efficiency and dynamic balance are more likely attributable to the walker's ergonomic design, height adjustability, maneuverability, and braking mechanism, while the physiological monitoring and fall detection functions provide additional clinical information to support safer mobility assessment and use [15-17]. Previous studies have demonstrated the feasibility of smart walker technologies and physiological monitoring systems in supporting mobility assessment among older adults. However, these studies evaluated different walker designs, sensor systems, walking tasks, and outcome measures, making direct comparisons difficult. Despite these promising technological advances, high-quality comparative clinical evidence evaluating the effectiveness of smart wheeled walkers remains limited, particularly randomized crossover studies assessing both functional and physiological outcomes under standardized supervised conditions.

XSL•FO
RenderX

Therefore, this randomized crossover trial aimed to compare walking efficiency, dynamic balance, physiological responses, and fear of falling between a smart wheeled walker with integrated physiological monitoring and fall detection and a standard wheeled walker during supervised use among community-dwelling older adults.

Methods

Development Process of the Smart Wheeled Walker

Step 1: Structural and Ergonomic Design

Smart Wheeled Walker

The smart wheeled walker was engineered with a lightweight aluminum alloy frame to provide structural stability while minimizing the physical effort required during ambulation. The device weighed approximately 6.9 kg and supported users weighing up to 120 kg. It featured a compact foldable design (51 cm × 59.6 cm × 86.6 cm) and 4 multidirectional wheels (2 front and 2 rear) to enhance maneuverability during walking. The frame was constructed from SUS-304 stainless steel hollow tubing (1-inch diameter, 1 mm thickness) to ensure durability while maintaining a lightweight structure. The handle height was adjustable to three levels (78 - 88 cm), allowing users to maintain approximately 20°-30° of elbow flexion during ambulation, an ergonomic position associated with reduced upper-limb loading and improved walking efficiency. The walker incorporated a foam-cushioned seat with a backrest and underseat storage to enhance user comfort during rest periods. Dual hand-operated brakes provided both speed control during walking and secure wheel locking while seated. These structural and ergonomic features were designed to provide mechanical characteristics comparable to those of the standard wheeled walker used in this study while incorporating additional physiological monitoring and fall detection functions.

Standard Wheeled Walker (Comparator)

The comparator was a commercially available 4-wheeled rollator with a lightweight aluminum alloy frame. The device weighed 7.2 kg and supported users weighing up to 100 kg. Its overall dimensions were 61 cm × 70 cm × 89 cm when unfolded and 61 cm × 40 cm × 104 cm when folded. The walker was equipped with four 8-inch polyvinyl chloride (PVC) wheels, adjustable-height handles (78 - 89 cm), and a foldable frame to facilitate transportation and storage. The seating system consisted of a foam-cushioned seat (35 × 32 cm) with a backrest to provide resting support during ambulation. A dual hand-operated braking system enabled speed control during walking and wheel locking while seated. The standard wheeled walker was selected because its fundamental mechanical characteristics, including the lightweight frame, 4-wheel configuration, adjustable handle height, braking mechanism, foldable design, and seating structure, were comparable to those of the smart wheeled walker. Unlike the smart wheeled walker, the standard device did not incorporate real-time physiological monitoring, fall detection, or mobile alert functions.

Step 2: Integration of Smart Fall Detection and Alert System

The walker is embedded with a high-precision tilt sensor (accuracy ±1.5°) capable of detecting deviations in real time. When tilt exceeds 30°—an angle validated in the fall biomechanics literature as indicative of instability—a 2-tier alert protocol is activated:

1. A 90 dB audible alarm is emitted to draw attention from nearby individuals.
2. A mobile notification system transmits alerts via SMS and the LINE app to preregistered caregivers or family members.

This system was validated in a lab-controlled environment using a standardized mannequin drop test, replicated across multiple angles and surfaces. Sensor reliability was cross-verified with external motion capture systems (Vicon, Oxford Metrics) and achieved 96% detection accuracy across 30 trials.

Step 3: Physiological Monitoring Module

To assess exertional load, the walker integrates a commercially calibrated photoplethysmography (PPG) sensor and a pulse oximeter (certified ISO 80601-2-61). The system continuously monitors:

- HR (beats per minute [bpm])
- Peripheral SpO_2 (%)

Real-time physiological data are displayed on a 2.4-inch OLED (organic light-emitting diode) interface in large, high-contrast fonts, allowing immediate feedback for users and clinicians. Sensor accuracy was benchmarked against a validated fingertip pulse oximeter (Masimo Radical-7) and showed a mean absolute error of less than ±2 bpm (HR) and ±1.8% (SpO_2), consistent with clinical-grade standards.

All instrumentation underwent repeated calibration and testing in both static and dynamic conditions (eg, walking on a treadmill at 1.0 - 1.5 m/s). A pilot usability study (n=10 older adults) confirmed signal stability during typical community-based ambulation.

Step 4: Testing of the Developed Smart Wheel Walker

Participants

This study included 30 community-dwelling older adults aged between 65 and 80 years, residing in Chonburi Province, Thailand. Participants were recruited from local communities in Chonburi Province, Thailand, through community announcements, village health volunteers, and referrals from local health care personnel. Individuals who expressed interest were invited to the Physical Therapy Laboratory, Burapha University, where eligibility screening was conducted according to the predefined inclusion and exclusion criteria. Eligible participants received verbal and written information about the study, provided written informed consent, and were subsequently enrolled in the trial. Eligibility criteria required that participants be capable of walking independently and performing activities of daily living without assistance. To ensure a mild balance deficit, the Timed Up and Go (TUG) test was administered, and only individuals with scores greater than 13.5 seconds were

XSL•FO
RenderX

enrolled. Exclusion criteria included a diagnosis of any neurological condition that impaired gait and orthopedic or musculoskeletal surgery of the upper or lower limbs within the past 3 months. Additional exclusions were uncorrectable visual impairments such as cataracts or diabetic retinopathy and any severe cardiovascular or musculoskeletal disorders that could interfere with safe ambulation. The study adhered to ethical guidelines, including the right to withdraw at any time and the protection of participant confidentiality. Following enrollment, participants were assigned to the intervention sequence in this randomized crossover trial. After completing the first walking condition, participants rested for 30 minutes to minimize carryover effects before using the alternate walker. The order of walker use was counterbalanced across participants to control for potential sequencing bias.

Study Design

This study used a single-blind randomized crossover trial design in which the outcome assessor was blinded to the allocation sequence. Randomization was performed by an independent researcher using a computer-generated random allocation sequence created with the RAND function in Microsoft Excel. Allocation concealment was maintained using sequentially numbered, sealed, opaque envelopes. Outcome assessments were conducted by a separate blinded assessor who was not involved in the randomization process or intervention administration. Because of the visible differences between the smart wheeled walker and the standard wheeled walker, participant blinding was not feasible. Participants were randomly assigned to begin with either the smart wheeled walker or the standard wheeled walker. After completing the first intervention, participants crossed over to the alternate walker following a 30-minute washout period to minimize potential carryover effects. The washout period was selected to allow physiological recovery following the brief 20-m walking task, particularly recovery of HR before the second assessment. Before formal testing, all participants completed a familiarization session with both walkers. Standardized instructions on the correct use of each device were provided, and participants practiced walking with both walkers before outcome measurements were performed. The short walking distance and self-selected comfortable walking speed were considered unlikely to induce substantial fatigue or persistent physiological effects that could influence the subsequent assessment. This randomized crossover trial was designed to compare the short-term effects of a smart wheeled walker and a standard wheeled walker under standardized, supervised clinical conditions. The primary objective was to evaluate walking efficiency using the PCI. Secondary objectives were to assess dynamic balance and fear of falling. HR and walking speed were recorded to support the calculation and interpretation of PCI rather than as independent confirmatory outcomes. This study was not intended to evaluate long-term clinical effectiveness, independent community use, or real-world implementation. The study was reported in accordance with the CONSORT-eHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist (Checklist 1).

Ethical Considerations

This study was approved by the Burapha University Research and Innovation Administration Committee (IRB1-115/2567). The trial was prospectively registered in the Thai Clinical Trials Registry (TCTR20250302008). The study conformed to the ethical standards set forth in the Declaration of Helsinki. All data collection was conducted between April and May 2025 at the Physical Therapy Laboratory of Burapha University, Thailand. Written informed consent was obtained from all participants before enrollment. Participants were informed that participation was voluntary and that they could withdraw from the study at any time without consequences. All data were deidentified before analysis and stored securely to maintain participant confidentiality. Participants received reimbursement for travel expenses but did not receive any additional financial compensation or incentives for participation.

Sample Size Estimation

The sample size was determined using G*Power version 3.1.9.2. The calculation was based on the primary outcome (PCI). An effect size of 0.60 was estimated from preliminary paired PCI data obtained from 10 community-dwelling older adults who completed walking assessments using both the smart wheeled walker and the standard wheeled walker. Assuming a 2-tailed paired *t* test, a significance level (α) of .05, and a statistical power ($1-\beta$) of .80, the estimated minimum sample size was 24 participants. To account for an anticipated 20% attrition rate, the target sample size was increased to 30 participants.

Experimental Procedure

Participants attended an orientation session 1 day prior to testing at the Faculty of Allied Health Sciences, Burapha University. The procedures were explained, and participants practiced adjusting walker height to the level of the greater trochanter, maintaining elbow flexion between 20° and 30°. They were also familiarized with walking using both a standard wheel walker and a smart wheeled walker.

On the testing day, participants were fitted with a Polar FT7 HR monitor (Polar Electro Oy). Resting HR was recorded over 5 minutes before walking trials. Participants then walked 20 m at a normal pace using one of the 2 walker types, with HR recorded every 30 seconds during walking and a 5-minute seated recovery period. Walking speed was also measured. After a 30-minute rest, participants repeated the protocol with the alternate walker. Walker order was randomized.

Outcome Measurements

Primary Outcome: Walking Efficiency

Walking efficiency was quantified using the PCI, calculated as follows: PCI (beats/m) = (HR during walking – Resting HR)/Average walking speed, where HR is expressed in bpm and average walking speed in meters per minute (m/min). A lower PCI value reflects a more efficient gait pattern. This metric was applied to compare energy expenditure during ambulation with both the smart-wheeled walker and the standard walker, thereby reflecting the physiological cost of walking [18]. The PCI was the prespecified primary outcome and served as the principal confirmatory end point. HR and walking speed were

XSL•FO
RenderX

measured to support the calculation and physiological interpretation of PCI and were not considered independent confirmatory outcomes. Dynamic balance (Expanded Timed Up and Go [ETUG]) and fear of falling (Falls Efficacy Scale-International [FES-I]) were evaluated as prespecified secondary outcomes.

Secondary Outcomes

Dynamic Balance

Dynamic balance was evaluated using the ETUG test.

Participants were instructed to rise from a chair, walk a 20-m distance using either walker, circle a designated marker, and return to sit. The time to complete the task was recorded in seconds [19].

Fear of Falling

Fear of falling was assessed using the Thai version of the FES-I, which was adapted and validated by Thiamwong based on the original instrument developed by the Prevention of Falls Network Europe (ProFaNE) [20]. The scale includes 16 items that assess concern about falling during a range of physical and social activities, from basic daily tasks to more demanding functions. Participants rated their concern on a 4-point Likert scale: 1 (not at all concerned), 2 (somewhat concerned), 3 (fairly concerned), and 4 (very concerned). Total scores range from 16 to 64. Scores of 16 to 21 indicate low concern about falling, 22 to 27 reflect moderate concern, and 28 to 64 represent high concern. This scale has been widely used to evaluate fall-related self-efficacy among older adults in community and clinical settings.

Data Analysis

All statistical analyses were conducted using SPSS (SPSS Inc.). Descriptive statistics were used to summarize demographic data. The Shapiro-Wilk test was used to assess the normality of outcome variables. Paired *t* tests were used to compare resting HR, walking HR, average walking speed, PCI, ETUG performance, and fear of falling between the 2 walker conditions for normally distributed data. In addition, a repeated-measures ANOVA was performed as a supplementary analysis for the randomized crossover design. The model included walker condition (smart wheeled walker vs standard wheeled walker) and period (first vs second assessment) as within-participant factors and sequence (smart-standard vs standard-smart) as a between-participant factor, with participants treated as the repeated subject. Sequence and period effects were examined to indirectly assess potential residual carryover effects. Because of the relatively small sample size, these analyses were considered exploratory and were interpreted with caution. A 2-sided *P* value of <.05 was considered statistically significant.

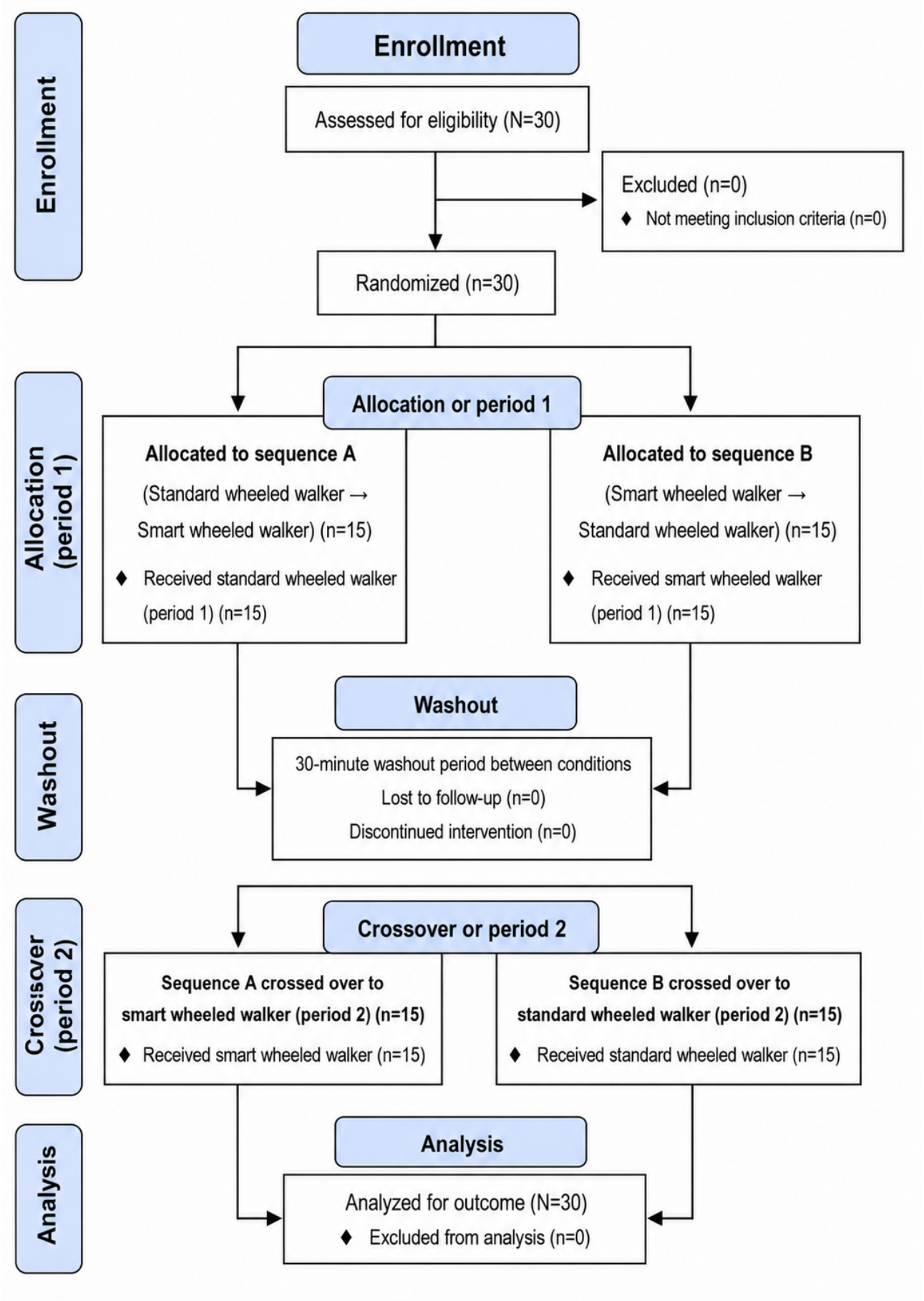
Results

Overview

Figure 1 presents the CONSORT 2010 flow diagram illustrating the flow of participants through enrollment, randomization, allocation to intervention sequence, crossover, follow-up, and analysis. A total of 30 community-dwelling older adults were enrolled and completed both intervention conditions. No participants were lost to follow-up or excluded from the final analysis.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) 2010 flow diagram of participant enrollment, randomization, allocation to intervention sequence, crossover, follow-up, and analysis.



XSL•FO
RenderX

Participant Characteristics

Thirty older adults participated in the study, including 7 (23.33%) men and 23 (76.67%) women. The mean age was 69.83 (SD 3.51) years. The average height and weight were 158.40 (SD 8.29) cm and 60.41(SD 10.99) kg, respectively, yielding a mean BMI of 24.20 (SD 3.56) kg/m². Functional mobility was assessed using the ETUG test, with an average completion time of 13.98 (SD 0.85) seconds (Table 1).

Table . Baseline characteristics of the participants (N=30).

Variable	Values
Sex, n (%)	
Male	7 (23.33)
Female	23 (76.67)
Age (y), mean (SD)	69.83 (3.51)
Weight (kg), mean (SD)	60.41 (10.99)
Height (cm), mean (SD)	158.40 (8.29)
BMI (kg/m²), mean (SD)	24.20 (3.56)
ETUG[a] test (3 m, seconds), mean (SD)	13.98 (0.85)

[a]ETUG: Expanded Timed Up and Go.

Primary Outcome: Walking Efficiency

The analysis showed that HR measurements taken before (*P*=.92), during (*P*=.55), and after walking (*P*>.99) did not differ significantly between the smart wheeled walker and standard wheel walker conditions. However, participants demonstrated a significantly faster average walking speed over the 20-m course when using the smart wheeled walker (*P*<.001). In addition, the PCI, which reflects the energy cost of walking, was significantly lower in the smart wheeled walker condition compared to the standard wheel walker (*P*=.02), indicating more efficient ambulation with the smart device.

Secondary Outcomes

Dynamic Balance

Dynamic balance, evaluated through the ETUG test, showed a statistically significant improvement when participants used the smart wheeled walker. The time required to complete the TUG test was notably shorter with the smart device than with the standard walker (*P*=.048), indicating better balance and mobility performance.

Fear of Falling

The findings of this study indicated that participants who walked with the smart wheeled walker reported lower levels of fear of falling compared to those who used the standard wheel walker. However, the difference between the 2 groups did not reach statistical significance (*P*=.62; Table 2).

Table . Comparison of heart rate (HR), walking efficiency, average speed, and dynamic balance between standard walker and smart wheel walker.

Variables	Standard wheel walker (n=30), mean (SD)	Smart wheel walker (n=30), mean (SD)	Difference among groups (95% CI)	*P*[a] value
Walking efficiency (PCI[b], beats/m)	0.69 (1.23)	0.36 (0.38[c])	0.33 (0.06 to 0.60)	.02
Resting HR (beats/minute)	71.72 (11.36)	71.35 (11.80)	0.37 (−5.61 to 6.35)	.92
Walking HR (beats/minute)	85.47 (11.34)	83.47 (11.89)	2.00 (−4.00 to 8.00)	.55
Postwalking HR (beats/min)	73.23 (12.01)	73.20 (12.62)	0.03 (−6.33 to 6.39)	>.99
Average walking speed (20 m, m/min)	18.14 (14.55)	28.77 (11.94[c])	−10.63 (−17.51 to −3.75)	<.001
Dynamic balance (ETUG[d], seconds)	64.47 (23.04)	50.06 (24.50[c])	14.41 (2.12 to 26.70)	.048
Fear of falling (scores)	32.91 (10.37)	31.56 (9.58)	1.35 (−3.81 to 6.51)	.62

[a]*P*<.05 was considered statistically significant.

[b]PCI: Physiological Cost Index.

[c]Significantly different between walker types.

[d]ETUG: Expanded Time Up and Go.

XSL•FO
RenderX

Additional crossover analyses demonstrated no significant sequence effects for PCI (*P*=.62), ETUG (*P*=.54), walking speed (*P*=.71), or FES-I (*P*=.83). Similarly, no significant period effects were observed for any study outcome (all *P*>.05). These findings indicate that the order of walker administration and testing period did not materially influence the treatment comparisons between the smart wheeled walker and the standard wheeled walker.

Discussion

Principal Findings

This study investigated the comparative impact of a smart wheeled walker and a standard walker on various aspects of mobility and user experience in older adults. The smart wheeled walker, which incorporates advanced functions, including adjustable height, multidirectional wheels, biometric feedback, and real-time fall alert capabilities, was associated with favorable short-term outcomes in several mobility-related measures under supervised testing conditions.

The smart wheeled walker was associated with improved walking efficiency, as reflected by a significantly lower PCI under the supervised conditions of the present study (0.36 beats/m), compared with the standard wheeled walker (0.69 beats/m). This value was lower than that previously reported for conventional wheeled walkers (1.23 beats/m) [12], although direct comparisons between studies should be interpreted with caution because of differences in study populations and experimental protocols. This improvement is likely attributable to multiple design features aimed at optimizing mechanical advantage and reducing muscular workload. The device's adjustable height mechanism allowed users to maintain an elbow flexion angle of approximately 30°, a position that supports optimal force transfer through the upper extremities during assisted gait. This alignment may have minimized unnecessary recruitment of the shoulder depressor muscles, such as the latissimus dorsi and triceps brachii, particularly in taller individuals [2], resulting in reduced fatigue and more efficient propulsion. Additionally, the multidirectional wheels and lightweight aluminum frame likely contributed to smoother gait cycles by decreasing friction and minimizing the need for abrupt deceleration or manual lifting of the device. These factors together may have reduced the cumulative metabolic demand during ambulation. Although the mean HR during walking was not statistically different between the groups, participants using the smart wheeled walker exhibited slightly lower average HRs (83.47, SD 11.89 bpm vs 85.47, SD 11.34 bpm). This trend may reflect a lower cardiovascular load, secondary to the improved mechanical efficiency and reduced physical exertion required to maneuver the device. These physiological findings are in line with prior evidence demonstrating that walker design can influence energy expenditure and cardiopulmonary response during walking [1].

The results of this study suggest that participants demonstrated better dynamic balance while using the smart wheeled walker than the standard wheeled walker, as reflected by ETUG performance under supervised testing conditions. Participants using the smart wheeled walker completed the test in less time (50.06, SD 24.50 s) than those using the standard wheeled walker (64.47, SD 23.04 s), suggesting superior walking efficiency and postural control. These findings align with previous research that demonstrated that wheeled walkers can reduce energy expenditure and support faster straight-line walking in older adults [9]. Previous studies have also reported that wheeled walkers improve balance and stability during ambulation [6]. However, the present study did not evaluate fall incidence; therefore, the potential effect on fall prevention remains to be confirmed in future longitudinal studies [6]. The current study findings are also consistent with previous evidence suggesting that wheeled walkers help prevent physical decline by enabling safe and sustained physical activity in older individuals. Importantly, the smart wheeled walker used in this study featured a height-adjustable frame with 3 levels, allowing users to begin walking from an ergonomically optimal elbow flexion angle of approximately 30° [21]. This positioning reduces arm fatigue and supports more natural and stable movement. Previous research has also demonstrated that height-adjustable walkers are associated with improved balance during ambulation [22]. Additionally, the smart wheeled walker incorporated front and rear wheel locking mechanisms, enhancing safety during use. These braking systems likely provided users with greater control during transitions and when navigating turns or uneven movements. These braking systems may have contributed to improved maneuverability and perceived stability during walking tasks, which may contribute to safer mobility during supervised walking tasks. The potential effect on fall prevention requires confirmation in longitudinal studies. Taken together, these findings suggest that the smart wheeled walker may have the potential to support mobility in older adults under supervised conditions. Because ETUG was a prespecified secondary outcome and no formal adjustment for multiple comparisons was performed, these findings should be interpreted as supportive rather than confirmatory evidence. Therefore, larger randomized controlled trials with long-term follow-up are required to confirm these findings.

In this study, the average score on the fear of falling scale was slightly lower among participants using the smart wheeled walker compared to those using the standard walker, although the difference was not statistically significant. However, the numerical trend observed in the smart wheeled walker condition may warrant further investigation in studies with larger sample sizes and longer follow-up. The smart walker included features such as a lightweight aluminum frame, adjustable handles, 4 rotatable wheels, a built-in seat, and a dual-braking system. Furthermore, it incorporated intelligent capabilities including a tilt sensor that detected deviations greater than 30° and transmitted alerts through mobile messaging apps. The integrated display of real-time HR and SpO_2 may also have provided additional reassurance to users and caregivers. These design features likely enhanced the sense of psychological security, particularly in individuals who ambulate independently or reside alone. Previous studies have suggested that automatic fall detection systems and health-monitoring functions may reduce anxiety related to safety. However, the present study did not evaluate the real-world performance of the fall detection and alert system during actual fall events. Previous research has

XSL•FO
RenderX

highlighted the association between elevated fear of falling and decreased physical and social engagement, which may in turn increase the risk of functional decline and actual falls [23]. Although the difference was not statistically significant, the numerical trend observed in the present study may warrant further investigation in adequately powered studies with longer follow-up to determine whether the smart wheeled walker has a meaningful effect on fear of falling, particularly among older adults at increased risk of falls.

This study has several important limitations. First, the inclusion criteria did not exclude participants with chronic conditions or those taking medications that could influence cardiovascular responses, such as beta-blockers. These factors may have affected HR data and walking performance outcomes. Future research should consider refining the exclusion criteria to control for medical conditions and pharmacological influences that could confound physiological measurements. Second, the number of previous falls among participants was not documented. Since fall history is a known predictor of mobility confidence and functional balance, its absence limits the ability to interpret fear of falling and balance-related outcomes fully. Third, the study was conducted in a controlled indoor setting with uniform flooring. This limits ecological validity, as real-world environments often include uneven surfaces, curbs, or visual distractions that challenge gait and balance. Future studies should assess device performance across diverse environmental conditions to better reflect everyday use. Fourth, no measures of oxygen consumption or electromyographic activity were included. Incorporating such physiological metrics would allow for a more comprehensive understanding of energy expenditure and muscle activation patterns during device use. Fifth, the gender distribution in this sample was notably imbalanced, with a higher number of female participants. Given known sex-related differences in gait characteristics, strength, and perceptions of fall risk, this imbalance may have influenced the study's findings. Future studies should strive for more balanced recruitment or conduct stratified analyses to better understand sex-specific responses to assistive technologies. Sixth, the smart wheeled walker and the standard wheeled walker differed not only in their smart monitoring and fall alert functions but also in several mechanical and ergonomic characteristics, including frame design, wheel configuration, and adjustability. Therefore, the observed differences between the 2 devices cannot be attributed solely to the smart monitoring and alert functions, as the mechanical and ergonomic differences between the devices may also have influenced walking performance and user outcomes. In addition, although the smart wheeled walker incorporates fall detection and mobile alert functions, the present study did not evaluate the real-world performance of these features during actual fall events or independent community use. Although the smart wheeled walker demonstrated promising short-term benefits under supervised conditions, several factors should be considered before its implementation in real-world settings. Older adults may differ in digital literacy, cognitive and sensory abilities, and their willingness to use technology-assisted devices. Caregiver support may be necessary for some users, particularly when responding to remote alerts or assisting with device operation. In addition, privacy and data security should be carefully addressed when transmitting health-related information. Future studies should evaluate the long-term usability, acceptability, safety, and implementation of the smart wheeled walker during independent home and community use in more diverse older adult populations. In addition, the present study did not systematically collect data on usability, user experience, or comprehensive safety outcomes, including adverse events during prolonged use. These aspects should be incorporated into future studies to provide a more comprehensive evaluation of the smart wheeled walker. Seventh, the study assessed only short-term use of each walker type. Longitudinal follow-up would be valuable for evaluating sustained effects on physical performance, fall incidence, and user adherence over time. The randomized crossover design enabled each participant to serve as his or her own control, thereby reducing interindividual variability. Additional crossover analyses demonstrated no significant sequence or period effects, suggesting that the order of walker administration and testing period did not materially influence the study outcomes. However, because of the relatively small sample size, these findings should be interpreted with caution and do not completely exclude the possibility of residual sequence or carryover effects. Moreover, participant familiarization before testing, randomized and counterbalanced allocation, and a 30-minute washout period were implemented to minimize potential carryover, learning, and fatigue effects. Nevertheless, because the study was conducted during a single supervised testing session, residual carryover or learning effects cannot be completely excluded. Future crossover studies with larger sample sizes and longer follow-up should further evaluate these potential influences. In addition, the generalizability of the present findings should be interpreted with caution. The study included only community-dwelling older adults aged 65 to 80 years, with a predominance of female participants, and all assessments were conducted under supervised laboratory conditions. Therefore, the findings may not be generalizable to frailer older adults, institutionalized individuals, people with cognitive impairment, or unsupervised community settings. Future studies should include more diverse populations and evaluate the performance of smart wheeled walkers in real-world environments to improve the external validity of the findings. Finally, this randomized crossover trial involved a relatively small sample size and a single supervised testing session. Therefore, the findings should be interpreted as evidence of short-term performance under supervised conditions rather than definitive evidence of long-term clinical effectiveness or real-world functional benefits. Future adequately powered randomized controlled trials with extended follow-up in community settings are warranted to confirm these findings and evaluate the long-term effectiveness, safety, and user adherence of the smart wheeled walker.

Conclusions

This randomized crossover clinical trial showed that the smart wheeled walker was associated with improved walking efficiency and dynamic balance under supervised testing conditions when compared with a standard wheeled walker. Cardiovascular responses and fear of falling did not differ significantly between walker conditions. Because the smart wheeled walker was evaluated as an integrated device, the

XSL•FO
RenderX

independent contribution of individual design features could not be determined. Further randomized controlled trials with larger and more diverse samples, long-term follow-up, and real-world evaluation are needed to evaluate the individual and combined effects of specific design features and to further establish the clinical effectiveness and practical applicability of smart wheeled walkers for older adults.

Acknowledgments

The authors would like to thank all participants and Mareukin Thongkam, Wariya Kamonsuwan, Uratsaya Loespornrak, and Udomsak Kaewsripot, who worked hard and supported the project.

Generative AI (ChatGPT; OpenAI) was used solely to assist with language editing and improvement of English writing during the preparation of this manuscript. The authors reviewed, verified, and approved all AI-assisted revisions and take full responsibility for the accuracy, integrity, and final content of the manuscript. Generative AI was not used for data analysis, interpretation of results, or generation of scientific conclusions.

Funding

This study was supported by Burapha University (BUU), Thailand Science Research and Innovation (TSRI), and the National Science, Research and Innovation Fund (NSRF) under the Fundamental Fund (grant number 1270/2567).

Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request. The data are not publicly available because they contain information that could compromise the privacy of research participants.

Authors' Contributions

Conceptualization: PW, JN
Data curation: PW, OS, PM
Formal analysis: PW, OS, PM
Investigation: PW, OS, PM, JN
Methodology: PW
Writing – original draft: PW, OS, PM
Writing – review and editing: PW, PM, OS, NL, AS, JN

Conflicts of Interest

None declared.

Checklist 1
CONSORT-eHEALTH checklist.
[PDF File, 429 KB - aging_v9i1e99522_app1.pdf]

References

1. Bergen G, Stevens MR, Burns ER. Falls and fall injuries among adults aged ≥65 years—United States, 2014. MMWR Morb Mortal Wkly Rep 2016 Sep 23;65(37):993-998. [doi: 10.15585/mmwr.mm6537a2] [Medline: 27656914]
2. Salari N, Darvishi N, Ahmadipanah M, Shohaimi S, Mohammadi M. Global prevalence of falls in the older adults: a comprehensive systematic review and meta-analysis. J Orthop Surg Res 2022 Jun 28;17(1):334. [doi: 10.1186/s13018-022-03222-1] [Medline: 35765037]
3. Dionyssiotis Y. Analyzing the problem of falls among older people. Int J Gen Med 2012;5:805-813. [doi: 10.2147/IJGM.S32651] [Medline: 23055770]
4. Mian OS, Thom JM, Ardigò LP, Narici MV, Minetti AE. Metabolic cost, mechanical work, and efficiency during walking in young and older men. Acta Physiol (Oxf) 2006 Feb;186(2):127-139. [doi: 10.1111/j.1748-1716.2006.01522.x] [Medline: 16497190]
5. LaPier J, Chatellier M. Can low cost fingertip pulse oximeters be used to measure oxygen saturation and heart rate during walking? Int J Physiother Res 2016 Oct 11;4(5):1689-1695. [doi: 10.16965/ijpr.2016.166]
6. Bateni H, Maki BE. Assistive devices for balance and mobility: benefits, demands, and adverse consequences. Arch Phys Med Rehabil 2005 Jan;86(1):134-145. [doi: 10.1016/j.apmr.2004.04.023] [Medline: 15641004]
7. Salminen AL, Brandt A, Samuelsson K, Töytäri O, Malmivaara A. Mobility devices to promote activity and participation: a systematic review. J Rehabil Med 2009 Sep;41(9):697-706. [doi: 10.2340/16501977-0427] [Medline: 19774301]

XSL•FO
RenderX

8. Lindemann U, Najafi B, Zijlstra W, et al. Distance to achieve steady state walking speed in frail elderly persons. Gait Posture 2008 Jan;27(1):91-96. [doi: 10.1016/j.gaitpost.2007.02.005] [Medline: 17383185]
9. Cetin E, Muzembo J, Pardessus V, Puisieux F, Thevenon A. Impact of different types of walking aids on the physiological energy cost during gait for elderly individuals with several pathologies and dependent on a technical aid for walking. Ann Phys Rehabil Med 2010;53(6-7):399-405. [doi: 10.1016/j.rehab.2010.06.003] [Medline: 20674533]
10. Lindemann U, Schwenk M, Klenk J, et al. Problems of older persons using a wheeled walker. Aging Clin Exp Res 2016 Apr;28(2):215-220. [doi: 10.1007/s40520-015-0410-8] [Medline: 26226859]
11. Liu HH. Assessment of rolling walkers used by older adults in senior-living communities. Geriatr Gerontol Int 2009 Jun;9(2):124-130. [doi: 10.1111/j.1447-0594.2008.00497.x] [Medline: 19740354]
12. Silva R, Couto F, Almeida M, et al. Use of three types of walkers: impact on the functional performance profile of institutionalized older adults. Rev Enf Ref 2019 Sep;IV Série(22):139-148. [doi: 10.12707/RIV19043]
13. Nickerson KA, Diaz K, Muir BC. Circumstances of falls among older adult walker users in long-term care and the associated walker design deficits. Assist Technol 2024 May 3;36(3):248-253. [doi: 10.1080/10400435.2023.2267623] [Medline: 37796449]
14. Costamagna E, Thies SB, Kenney LPJ, et al. Objective measures of rollator user stability and device loading during different walking scenarios. PLoS One 2019;14(1):e0210960. [doi: 10.1371/journal.pone.0210960] [Medline: 30699170]
15. Strutz N, Meyer-Feil T, Schwesig R, Schulze S, Gissendanner H. Development of a smart walker for clinical settings: a protocol of an exploratory mixed-methods study. BMJ Open 2025 Dec 31;15(12):e105342. [doi: 10.1136/bmjopen-2025-105342] [Medline: 41475804]
16. Sierra M SD, Harris N, Múnera M, Cifuentes CA. Socially assistive walker for daily living assistance in older adults. Front Robot AI 2024;11:1401663. [doi: 10.3389/frobt.2024.1401663] [Medline: 39220205]
17. Sanyod V, Saelim K, Chaarmat K, Kuankid S. IoT-based smart walking assistant for fall detection in the elderly. Int J Onl Eng 2025;21(6):19-35. [doi: 10.3991/ijoe.v21i06.54443]
18. Butler PB, Engelbrecht M, Major RE, Tait JH, Stallard J, Patrick JH. Physiological cost index of walking for normal children and its use as an indicator of physical handicap. Dev Med Child Neurol 1984 Oct;26(5):607-612. [doi: 10.1111/j.1469-8749.1984.tb04499.x] [Medline: 6239799]
19. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
20. Thiamwong L. Psychometric testing of the Falls Efficacy Scale-International (FES-I) in Thai older adults [Article in Thai]. Songkla Med J 2011 Nov;29(6):277-287 [FREE Full text]
21. Hoenig H, Taylor DH Jr, Sloan FA. Does assistive technology substitute for personal assistance among the disabled elderly? Am J Public Health 2003 Feb;93(2):330-337. [doi: 10.2105/ajph.93.2.330] [Medline: 12554595]
22. Choi HJ, Ko CY, Kang S, Ryu J, Mun M, Jeon HS. Effects of balance ability and handgrip height on kinematics of the gait, torso, and pelvis in elderly women using a four-wheeled walker. Geriatr Gerontol Int 2015 Feb;15(2):182-188. [doi: 10.1111/ggi.12246] [Medline: 24612309]
23. Delbaere K, Close JCT, Brodaty H, Sachdev P, Lord SR. Determinants of disparities between perceived and physiological risk of falling among elderly people: cohort study. BMJ 2010 Aug 18;341:c4165. [doi: 10.1136/bmj] [Medline: 20724399]

Abbreviations

bpm: beats per minute
CONSORT-eHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
ETUG: Expanded Timed Up and Go
FES-I: Falls Efficacy Scale-International
HR: heart rate
OLED: organic light-emitting diode
PCI: Physiological Cost Index
PPG: photoplethysmography
ProFaNE: Prevention of Falls Network Europe
PVC: polyvinyl chloride
SpO_2: oxygen saturation
TUG: Timed Up and Go

XSL•FO
RenderX

Edited by A Mavragani; submitted 26.Apr.2026; peer-reviewed by J Qiu, Y Ying; revised version received 31.Jul.2026; accepted 02.Aug.2026; published 28.Aug.2026.

Please cite as:
Werasirirat P, Muanjai P, Sukkho O, Luangpon N, Snieckus A, Namsawang J
A Smart Wheeled Walker With Integrated Physiological Monitoring to Improve Mobility in Older Adults: Randomized Crossover Trial
JMIR Aging 2026;9:e99522
URL: https://aging.jmir.org/2026/1/e99522
doi:10.2196/99522

© Phurichaya Werasirirat, Pornpimol Muanjai, Oranat Sukkho, Nongnuch Luangpon, Audrius Snieckus, Juntip Namsawang. Originally published in JMIR Aging (https://aging.jmir.org), 28.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Leveraging Digital Health Technologies to Assess Older Adults' Frailty and Nutritional Status: Two Cross-Sectional Studies

Nunzio Camerlingo[1], PhD; Madisen Wicker[2], MS; Dimitrios J Psaltos[1], MS; Nina Shaafi-Kabiri[2], PhD; Hao Zhang[1], PhD; Andrew Messere[1], BS; Isabelle Messina[2], MS; Zachary Hayden[2], BS; Meredith Kelly[2], MBA; Gauri Kamat[1], PhD; Fikret Işık Karahanoğlu[1], PhD; Charmaine Demanuele[1], PhD; Mar Santamaria[1], MD; David Caouette[1], MBA; Kevin C Thomas[2], PhD

[1]Pfizer (United States), Cambridge, MA, United States

[2]Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, United States

Corresponding Author:
Nunzio Camerlingo, PhD
Pfizer (United States)
1 Portland Street
Cambridge, MA, 02139
United States
Phone: 1 8573199633
Email: nunzio.camerlingo@pfizer.com

Abstract

Background: With the rising prevalence of aging worldwide, there is a necessity for reliable and frequent assessments of older individuals' health status to manage and potentially prevent age-related complications. Digital health technologies (DHTs), such as wearable devices, provide an opportunity to gather objective, continuous, and unobtrusive measurements, enabling effective health management in everyday life.

Objective: We aimed to evaluate the relationship between digital endpoints and frailty, nutritional status, and patient-reported outcomes (PROs) in older individuals, as well as quantified their compliance and comfort with DHTs.

Methods: In two cross-sectional studies, namely the Geriatric Anorexia Studies 1 and 2, including three in-clinic visits and 2-3 weeks of at-home monitoring, 94 participants (mean age 72.98 years, SD 6.28 years) were stratified based on their frailty status (n=39, 41%, nonfrail; n=45, 48%, prefrail; and n=10, 11%, frail) and nutritional status (n=70, 74%, with normal nutrition and n=24, 26%, at risk for malnutrition), as assessed in a clinical setting using the Fried Frailty Score and the Simplified Nutritional Appetite Questionnaire, respectively. We remotely monitored older adults using different DHTs. At home, participants were monitored with a wrist accelerometer for physical activity, continuous glucose monitoring (CGM) for glucose concentration, a digital body scale for weight and body composition, and a digital nutritional scale for meal tracking. Compliance with devices was assessed via wear time and correct at-home usage, while comfort was evaluated using questionnaires. The association between digital endpoints and frailty/nutritional status was investigated via linear regression, followed by ANOVA, and the relationship between digital endpoints collected at home and PROs was evaluated via Spearman's ρ. Weight and body composition were also assessed during in-clinic visits with a research-grade scale, used to validate the at-home digital body scale measurements, via the intraclass correlation coefficient, Pearson's R, and Bland-Altman plots with mean bias.

Results: Physical activity digital endpoints collected at home were significantly different across frailty and nutritional groups and significantly correlated with self-reported appetite, fatigue, and physical function (eg, for mean daily activity in the maximum 60 minutes of activity, ρ=0.28, –0.23, and 0.47, respectively). More than 80% of participants reported that all devices were mostly to very acceptable to wear/use: 95% (88/93) for the body scale, 86% (43/50) for CGM, 80% (40/50) for the nutrition scale, and 91% (84/93) for the wrist accelerometer. Compliance ranged from 60% of monitoring days for the digital nutrition scale to 90% for the wrist accelerometer.

Conclusions: Physical activity digital endpoints show clear prognostic significance for frailty and malnutrition in older individuals and reflect self-reported measures. DHTs can be reliably deployed at home for older individuals to measure their physical activity, weight and body composition, glucose concentration, and meal intakes, thus enabling patient-centric, data-centric clinical trials.

Trial Registration: ClinicalTrials.gov NCT04858932; https://clinicaltrials.gov/study/NCT04858932; ClinicalTrials.gov NCT05211973; https://clinicaltrials.gov/study/NCT05211973

XSL•FO
RenderX

(JMIR Aging 2026;9:e77816) doi:10.2196/77816

KEYWORDS

digital health technology; older population; frailty; malnutrition; wearable devices

Introduction

Aging is a worldwide phenomenon that is rapidly gaining attention due to the increasing prevalence of older adults and the profound impact it has on individuals and society [1,2]. Aging is associated with a wide variety of comorbidities, such as reduced appetite, undernutrition with associated weight loss and changes in body composition, decreased exercise capacity, and frailty [3]. These conditions are commonly referred to as anorexia of aging (AoA) [4-7].

To diagnose and monitor age-related diseases, typical assessments include, but are not limited to, physical function, weight, and appetite [8]. These assessments are traditionally performed during in-clinic visits using nonscalable devices (eg, research-grade scales to measure weight and body composition) and questionnaires (eg, the Fried Frailty assessment [9]). Notably, in-clinic assessments might not reflect the variability observed in daily life [10,11], because the visits are sporadic and limited in time and space, and participants' performance can be affected by the presence of observers [12]. In addition, self-reported assessments inherently lack objectivity, particularly for individuals with cognitive impairments, making comparisons between individuals difficult [6]. Finally, in-clinic visits require individuals to travel to clinical trial sites, which can be burdensome, especially for those with reduced mobility and/or those living in rural areas [13].

To overcome the limitations of traditional assessments of age-related diseases and address the need for objective, remote, and unobtrusive high-frequency measurements, the use of digital health technologies (DHTs) has been explored in the literature [11,14]. DHTs refer to a broad category of tools, platforms, and systems that integrate software, hardware, and data analytics to monitor different aspects of health and support clinical decision-making, enhance health care management, and improve patient engagement [15]. This work focuses on DHTs to monitor age-related conditions: wearable accelerometers, digital body weight and nutrition scales, continuous glucose monitoring (CGM) devices, and smartphone apps, whose applications are described next.

Wearable accelerometers, such as smartwatches and lumbar belts, allow for objective measurements of the variability of physical activity and gait in daily life. Their use in older individuals has been widely explored in the literature, showing their potential to unlock new insights into older adults' frailty and overall health status [16,17].

Digital body scales, deployed in free-living environments, may afford an opportunity to monitor weight and body composition remotely and with more granularity (eg, daily) and to automatically share data with clinical practitioners. This may enable the early detection of weight loss, weight cycling, or sarcopenia, which are associated with disability and mortality in older adults [18,19].

DHTs can also be used to remotely monitor meal intake and glucose concentration, which can indicate alterations in eating behavior, mitigating the risk for malnutrition and potentially preventing the onset of metabolic disorders [20,21]. Handwritten food diaries are often used to report information about meal timings and amounts, but they rely on users' memory to recall the food consumed, and require the burdensome task of estimating portion size, thus lacking accuracy and reproducibility. In addition, they provide unstructured data, requiring time-consuming and error-prone data harmonization prior to analysis [22-24]. To streamline dietary behavior monitoring, digital nutrition scales can be used to automatically pull meal timings from the network and objectively calculate the serving weight [25,26], resulting in standardized data that can be sent in real time to physicians, dietitians, or nutritionists to accelerate their clinical interpretation. The smartphone apps linked to nutrition scales can provide users with additional features, such as reminders at regular time intervals, which increase compliance and data completeness [27], and the integration of food composition databases, which streamlines the calculation of meals' nutritional content. More advanced tools, such as the one proposed by Ryu et al [28], can perform automatic estimation of meals' macronutrient content by leveraging artificial intelligence models to analyze food pictures taken by individuals.

Furthermore, continuous glucose monitoring (CGM) devices represent state-of-the-art systems to track glucose concentration continuously, with wide applications in the diabetes space across different age ranges [29,30]. Although there is evidence for the use of CGM in different disease areas, including obesity [31], eating disorders [32], or following bariatric surgery [33], its application in healthy older individuals has been poorly explored.

The systematic evaluation and early assessment of frailty and malnutrition are crucial to identify individuals at the highest risk for adverse outcomes, allowing the development of individualized care plans, optimization of physical and mental health, and improvement in overall quality of life for the aging population [34]. In this paper, we investigated the use of DHTs to remotely characterize frailty and malnutrition in older adults. For this purpose, we conducted two cross-sectional studies, namely the Geriatric Anorexia Studies 1 and 2 (GAS-1 and GAS-2), in which older individuals, aged 65-96 years, with different frailty and nutritional statuses were monitored both during in-clinic visits and at home with various DHTs: lumbar and wrist accelerometers, a digital body scale, a digital nutrition scale with a smartphone app, and CGM. Our previous study [35] examined the relationship between older individuals' frailty status and digital endpoints of gait and physical activity using GAS-1 data only. This study integrated GAS-1 and GAS-2 data and broadened the scope by exploring digital endpoints of physical activity, glucose concentration, body composition, and meal composition in relation to (1) the effectiveness of physical

XSL•FO
RenderX

activity digital endpoints in distinguishing between frailty and nutritional status and in correlating with self-reported measures and (2) compliance and comfort with DHTs deployed remotely to measure body weight, glucose concentration, and dietary behavior in order to support the design of more patient-centric clinical trials.

Methods

Dataset: Geriatric Anorexia Studies

The dataset used in this manuscript combined data collected in two studies: GAS-1 (NCT04858932), which monitored community-dwelling older adults, and GAS-2 (NCT05211973), which monitored a cohort of older individuals in long-term care (LTC) settings, such as nursing homes or assisted-living facilities. Both studies included three in-person visits at the participant's facility, where traditional assessments were performed, including questionnaires and weight and body composition data collection via research-grade scales, and 2-3 weeks of at-home monitoring, where DHTs were deployed to collect data remotely. Both studies targeted participants aged 65 years or older, with a Montreal Cognitive Assessment (MoCA) score of ≥19. Additional details on the study design and inclusion/exclusion criteria are reported in Multimedia Appendix 1.

Ethical Considerations

Both studies were reviewed and approved by the Boston University Medical Campus and Boston Medical Center Institutional Review Board (IRB #H-40837 and #H-42048). All participants provided written informed consent prior to enrollment into the study. All study procedures were in accordance with the relevant guidelines/regulations, including data privacy and the Declaration of Helsinki. Data transmission protocols were reviewed by the Boston University Information Technology and the IRB for security and privacy. No identifiable data were used to initialize the devices. Secure File Transfer Protocol (sFTP) was used to transfer data between Boston University and Pfizer, following Good Clinical Practice guidelines, Boston University Information Technology, and Spivack Center for Clinical and Translational Neuroscience policies.

DHTs Used in the Two Studies

Physical activity data were collected using the GENEActiv accelerometer (Activinsights Ltd), which was worn on the nondominant wrist. A lumbar accelerometer was used only in GAS-1, and the related gait endpoints have been discussed elsewhere [35].

Weight and body composition measurements were collected using two scales, the Tanita MC-780U scale (Tanita Corporation of America) and the Renpho ES-BR003 body scale (Renpho Joicom Corporation). The former, used only during in-person visits by study staff, is a Food and Drug Administration (FDA)–cleared scale that uses bioelectrical impedance analysis (BIA) from footplate and hand electrodes [36] and stores data on a wired computer. The latter, used both during in-clinic visits and at home, is a cost-effective, rechargeable, consumer-grade scale that uses lower extremity BIA and stores data through a connection to its own app downloaded on the iPhone XR (Apple Inc) provided. Participants were instructed to use the Renpho body scale every day in the morning upon waking.

The Freestyle Libre Pro CGM (Abbott Laboratories) was used only in GAS-1 to automatically measure the glucose concentration every 15 minutes.

Finally, GAS-1 participants were instructed to use the Renpho ES-SNG01-W smart nutrition scale (Renpho, Joicom Corporation), a digital food scale, paired via Bluetooth with the Gennec app, to capture meals' weight and integrate nutritional insights from a food database, helping keep an automatic and comprehensive food diary. As a backup collection system, participants photographed the weight displayed on the scale using the iPhone provided. Participants were required to perform these operations both before and after consuming each meal, including prepackaged meals provided to them by Epicured.

Assessment of Frailty and Nutritional Status

In this research, the Fried Frailty Score (FFS) was used to evaluate frailty. The FFS involves both questionnaires and objective physical performance tasks (ie, unintentional weight loss, self-reported exhaustion, weakness, slow walking speed, and low physical activity), resulting in a total score of 0-5, which is used to stratify individuals in three frailty groups: nonfrail (FFS=0), prefrail (FFS=1-2), and frail (FFS=3-5) [9]. In some studies [35,37], the FFS has also been used to categorize individuals into two robustness groups, robust (FFS=0, ie, nonfrail), and nonrobust (FFS=1-5, ie, combining prefrail and frail).

In this research, the Mini Nutritional Assessment – Short Form (MNA-SF) and the Simplified Nutritional Appetite Questionnaire (SNAQ) were used to evaluate nutritional status. The MNA-SF includes six questions about food intake, recent weight loss, current mobility, psychological stress or acute disease, neuropsychological problems, and the BMI or calf circumference, resulting in a total score of 0-14, indicating malnutrition (≤7), risk for malnutrition (8-11), and normal nutrition (≥12) [38]. The SNAQ is a four-question tool scored using a 5-point Likert-scale also used to evaluate the nutritional status, resulting in a total score of 4-20, where higher scores indicate better appetite. In particular, scores≤14 have been reported to indicate a significant risk for weight loss of >5% within 6 months [39,40] and to be a marker of malnutrition risk in healthy community-dwelling older adults [41].

Patient-Reported Outcomes

Patient-reported outcomes (PROs) were measured during in-clinic visits and at home on paper by participants. Appetite and fatigue PROs were measured daily. The question on appetite was "How would you rate your appetite during the past 24 hours from 0 to 10 (0=no appetite, 10=very good appetite)?," while the fatigue question was "How would you rate your physical fatigue during the past 24 hours from 0 to 10 (0=no fatigue, 10=worst-possible fatigue)?" These questions were developed by Pfizer based on concept elicitation interviews and cognitive debriefing of questions and response options with the population of interest, following the "FDA Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation

XSL•FO
RenderX

of the Patient's Voice in Medical Product Development and Regulatory Decision Making" [42]. Finally, the well-known PROMIS (Patient-Reported Outcomes Measurement Information System) Physical Function 10a was used during the in-clinic visits [43].

Comfort and wearability questionnaires were used at the last visit to gather feedback on the use of DHTs. The questionnaires included 10 items, with rated questions and free-text comments, reflecting what has been already used in previous studies across different populations [44,45]. In this research, the scores of the questions were not aggregated, but we focused only on the following two questions provided separately per device: "Please rate the ease of wearing and/or using the device," where participants could answer on a 1-7 scale (1=very acceptable, 7=very unacceptable), and "Would you be willing to use the device in the future, and if so, for how long?," where participants could select from "No," "Yes, less than 1 week," "Yes, 1-4 weeks," and "Yes, more than 4 weeks".

Endpoint Generation

Physical activity endpoints were derived from the GENEActiv wrist accelerometer using the vendor-supplied proprietary algorithm. This algorithm aggregates the 50 Hz triaxial raw accelerometry data into 1-minute summaries by extracting the signal vector magnitude (SVMg). Based on SVMg thresholds, the algorithm distinguishes wear from nonwear periods, and during wear periods, it classifies activity as sleep, sedentary, light, moderate, and vigorous. Visit days were excluded from the analysis, as they contained only partial recordings. In addition, only participants with at least 4 compliant days were included in the analysis. For each compliant day, the following physical activity endpoints were derived: time in sedentary activity (hours), time in nonsedentary activity (hours), time in sedentary-to-light physical activity (SLPA; hours), time in moderate-to-vigorous physical activity (MVPA; hours), mean SVMg (arbitrary units [a.u.]), 95th percentile of SVMg (a.u.), mean SVMg during the maximum 6 minutes of activity (M6min; a.u.), and mean SVMg during the maximum 60 minutes of activity (M60min; a.u.). The last two endpoints have been used in the literature to quantify the average activity during the most active 6 and 60 minutes of the day and have shown significant association with traditional measures of functional capacity in older individuals [46]. Finally, physical activity endpoints were averaged across compliant days for each participant.

The following measurements were collected using the Renpho body scale and the Tanita scale: weight (kg), muscle mass (kg), fat mass (kg), fat-free mass (kg), bone mass (kg), and water mass (kg). All body composition measurements were converted to percentages of weight. In-clinic and at-home body measurements were averaged across the three visits and at-home monitoring days per participant, respectively.

To obtain meal-related endpoints, timestamps, which were manually transcribed from each meal photo by two scientists and an independent reviewer, were associated with a meal weight. Weight entries from Gennec and the photos were then paired using their timestamps and meal names. In some cases, weight data were available only from Gennec, if participants did not take the photo correctly (eg, it was too blurry, or the weight was not photographed) or if participants did not take a photo at all. In other cases, weight data were available only from the photo if participants did not connect Gennec to the Renpho nutrition scale, they did not open the app to trigger a new meal recording, or the meal was not left long enough on the scale for the app to record the weight. When the weight was available from both sources, the average weight was computed for both pre- and postmeal entries. The meal net weight (g) was computed as the difference between the post- and premeal weights.

In the case of prepackaged meals, the meal macronutrient content was available, including energy (kcal), fat (g), sodium (g), potassium (g), carbohydrates (g), protein (g), and potassium (g). The net macronutrient content of the meal was computed by applying the percentage of consumed meal weight (ratio of meal net weight and premeal weight) to the macronutrient content of the prepackaged meal, under the assumption that participants consumed the same percentage of macronutrients in the meal. The net macronutrient content was computed only for the prepackaged meals. Meal duration was obtained as the time that elapsed between the pre- and postmeal times. The averages of net weight, energy, macronutrient content, and duration across meals were extracted per participant.

The following endpoints were computed from the Abbott Freestyle Libre Pro CGM data for each at-home monitoring day: mean (SD; mg/dL), minimum (mg/dL), maximum (mg/dL), time in range (TIR, %, defined as the percentage time within 70-180 mg/dL), time below range (TBR, %, defined as the percentage time below 70 mg/dL), and time above range (TAR, %, defined as the percentage time above 180 mg/dL). CGM endpoints were averaged across monitoring days per participant.

PROMIS Physical Function 10a scores were converted to *t*-scores (mean 50, SD 10, representing the US general population) according to standard guidelines [43]. PROMIS Physical Function 10a *t*-scores as well as appetite and fatigue PRO scores were averaged across monitoring days per participant.

Statistical Analysis

The relationship between frailty and nutritional groups was investigated via chi-square tests performed between pairs of categories (ie, FFS vs MNA-SF, FFS vs SNAQ, MNA-SF vs SNAQ), with scatter plots and Spearman's correlation.

The endpoints derived from DHTs at home were tested to assess whether they distinguished the frailty or nutritional groups. For this purpose, a linear regression model was leveraged, with the average endpoint as a dependent variable and the group (ie, nonfrail/prefrail/frail, nonrobust/robust, or normal nutrition/at risk for malnutrition), age, and sex as independent variables. Height was also added as an independent variable for models investigating differences in weight and body composition endpoints.

Model residuals were characterized in terms of normality via the Shapiro-Wilk test and homogeneity of the variance across groups via Levene's test. The presence of an overall statistically significant association between endpoints and groups was investigated via one-way ANOVA in the case of normality and

XSL•FO
RenderX

homogeneity of the residuals, Welch's ANOVA in the case of normality but nonhomogeneity of the residuals, or the Kruskal-Wallis test in the case of nonnormality of the residuals. For $P<0.1$, post hoc pairwise tests were performed to assess differences between pairs of frailty groups via least square comparisons following ANOVA or Wilcoxon tests following the Kruskal-Wallis test. The 0.1 threshold was chosen to flag digital endpoints exhibiting potential trends across groups, thereby reducing type II error. No correction for multiple comparisons was applied [47].

The association between digital endpoints with PRO scores was assessed using Spearman's correlation (ρ) to account for outliers.

Agreement between the Renpho body scale (test device) and the Tanita scale (reference device) endpoints collected during in-clinic visits was assessed using the intraclass correlation coefficient (ICC; two-way random-effects model, absolute agreement) following benchmarks reported by Cicchetti et al [48]: ICC≤0.4, poor agreement; ICC=0.4-0.59, fair agreement; ICC=0.6-0.74, good agreement; and ICC=0.75-1, excellent agreement. The mean difference (bias), limits of agreement (LoA), mean absolute difference, mean percentage error, and Pearson's correlation coefficient (R) were also computed. Scatter plots and Bland-Altman plots with 95% LoA were used to visualize agreement.

All statistical analyses were performed in R version 3.4.2 (R Foundation for Statistical Computing) using the following main packages: *car* for type III ANOVA, *BlandAltmanLeh* for Bland-Altman plots, and *psych* for the ICC.

Results

Demographics

In total, 94 older individuals (n=48, 51%, females) aged 65-96 (mean 72.98, SD 6.28) years with a BMI of 17.3-50.4 (mean 25.81, SD 4.60) kg/m^2 were recruited across GAS-1 (n=50, 53%) and GAS-2 (n=44, 47%) and were monitored with DHTs at home for 8-16 (mean 12.43, SD 1.20) days (excluding visit days). Of the 94 participants, 7 (7%) reported partial use of walking aids (eg, cane, walker, or leg braces), and 6 (6%) reported previous hip or knee replacement.

According to the MNA-SF, nutritional status was classified as "normal" for 67 (71%) participants (mean age 72.39, SD 5.57 years), of which 28 (42%) were females, and "at risk for malnutrition/malnourished" for 27 (29%) participants (mean age 74.44, SD 7.69 years), of which 20 (74%) were females. According to the SNAQ, nutritional status was classified as "normal" for 70 (74%) participants (mean age 73.58, SD 6.17 years), of which 36 (51%) were females, and "at risk for malnutrition/weight loss" for 24 (26%) participants (mean age 71.21, SD 6.39 years), of which 12 (50%) were females. Participants' characteristics are presented in Table 1, stratified by frailty phenotype. Chi-square tests showed expected significant associations between groups (FFS vs MNA-SF: $P<.001$; FFS vs SNAQ: $P<.01$; MNA-SF vs SNAQ: $P<.05$). Scatter plots with Spearman's correlation between the FFS, MNA-SF, and SNAQ are reported in Figure S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Table 1. Participants' demographics characteristics and compliance with digital tools, stratified by frailty group.[a]

Characteristics	Nonfrail group (n=39; FFS[b]=0)	Prefrail group (n=45; FFS=1-2)	Frail group (n=10; FFS≥3)	Total participants (N=94)
Enrolled female participants, n (%)	20 (51)	18 (40)	10 (100)	48 (51)
Age (years), mean (SD, range)	70.79 (3.34, 65.00-81.00	74.60 (6.62, 65.00-93.00	74.20 (10.53, 65.00-96.00	72.98 (6.28, 65.00-96.00
BMI (kg/m^2), mean (SD, range)	25.06 (3.39, 18.60-29.80	26.52 (3.51, 20.80-35.30	25.49 (10.23, 17.30-50.40	25.81 (4.60, 17.30-50.40
Race , n (%)				
Black/African American	5 (13)	11 (24)	2 (20)	18 (19)
White	33 (85)	33 (73)	7 (70)	73 (78)
More than one race	1 (2)	1 (3)	1 (10)	3 (3)
FFS , n (%)				
0	39 (100)	—[c]	—	39 (41)
1	—	33 (73)	—	33 (35)
2	—	12 (27)	—	12 (13)
3	—	—	7 (70)	7 (7.4)
4	—	—	3 (30%)	3 (3.2%)
MNA-SF[d] score				
Normal (score 12-14), n (%)	31 (79)	35 (78)	1 (10)	67 (71)
At risk for malnutrition/malnourished (score 0-11), n (%)	8 (21)	10 (22)	9 (90)	27 (29)
Total (score 0-14), median (IQR, range)	13.00 (2.00, 10.00-14.00)	13.00 (2.00, 8.00-14.00)	9.50 (4.5, 3.00-14.000)	13.00 (3.00, 3.00-14.00)
SNAQ[e] score				
Normal (score 15-20), n (%)	31 (79)	36 (80)	3 (30)	70 (74)
At risk for weight loss/malnutrition (score 4-14), n (%)	8 (21)	9 (20)	7 (70)	24 (26)
Total (score 4-20), median (IQR, range)	16.00 (2.00, 12.00-18.00)	15.00 (2.00, 11.00-18.00)	12.00 (5.50, 6.00-16.00)	15.50 (2.75, 6.00-18.00)
At-home monitoring days (excluding visit days), mean (SD, range)	12.30 (0.67, 11.00-13.00)	12.44 (1.98, 8.00-16.00)	12.44 (1.31, 9.00-15.00)	12.43 (1.20, 8.00-16.00)
Compliance with wrist accelerometer (%), median (IQR, range)	100 (3.50, 42.00-100.00)	100 (9.00, 21.00-100.00)	100 (8.00, 54.00-100.00)	100 (8.00, 21.00-100.00)
Compliance with Renpho body scale (%), median (IQR, range)	92.86 (8.01, 40.00-100.00)	92.30 (9.09, 0.00-100.00)	91.99 (38.78, 45.45-100.00)	92.30 (8.33, 0.00-100.00)
Compliance with CGM[f] (%), median (IQR, range)	100 (11.54, 8.33-100.00)	100 (44.23, 25.00-100.00)	100 (0.00, 75.00-100.00)	100 (24.04, 8.33-100.00)
Compliance with Renpho nutrition scale (%), median (IQR, range)	81.82 (45.64, 0.00-100.00)	78.57 (50.00, 0.00-100.00)	7.14 (16.07, 0.00-71.43)	78.57 (76.23, 0.00-100.00)

[a]Continuous data are reported as the mean (SD, range), categorical data are reported as the median (IQR) or n (%), and compliance data are reported as the median (IQR, minimum-maximum).

[b]FFS: Fried Frailty Score.

[c]Not applicable.

[d]MNA-SF: Mini Nutrition Assessment – Short Form.

[e]SNAQ: Simplified Nutritional Appetite Questionnaire.

[f]CGM: continuous glucose monitoring.

XSL•FO
RenderX

Physical Activity at Home Distinguishes Frailty and Nutritional Status and Correlates With PROs

A compliant day was defined as a day with at least 18 hours of wear time, in agreement with previous literature [49,50], resulting in a mean percentage compliance of 92.04% (SD 17.32%, range 21.00%-100.00%) across participants (Table 1). Data from 3 (3%) participants with less than 4 compliant days were excluded from the analysis, resulting in 91 (97%) compliant participants. During compliant days, participants had a mean wear time of 23.13 (SD 0.82, range 19.91-24.00) hours/day. Questionnaire responses showed that of 93 participants, 84 (91%) found the wrist accelerometer mostly to very acceptable to wear (Figure 1A), and 82 (89%) indicated that they would wear the wrist accelerometer for more than 1 week (Figure 1B).

ANOVA showed a significant effect of frailty status on all physical activity endpoints except SLPA (Table 2). Time spent in sedentary activity, nonsedentary activity, and MVPA (Figure 2A) was significantly different across frailty groups: the nonfrail group spent significantly more time in nonsedentary activity compared to both prefrail and frail groups, significantly more time in MVPA compared to both prefrail and frail groups, and significantly less time in sedentary compared to the frail group. In addition, the mean and 95th percentile of SVMg, as well as M6min and M60min (Figure 2B), were significantly different across frailty groups. The results were concordant when comparisons were performed across the two robustness groups, with the nonrobust group showing significantly higher time in sedentary activity and lower time in nonsedentary activity, MVPA, mean SVMg, 95th percentile of SVMg, M6min, and M60min, compared to the robust group.

Furthermore, a significant age effect was observed for time spent in nonsedentary activity (P<.01), MVPA (P<.01), mean SVMg (P<.05), and 95th percentile of SVMg (P<.01), while a significant effect of sex was observed for M6min (P<.001), with female participants showing a lower M6min (mean 528, SD 135 a.u.) compared to male participants (mean 679.2, SD 270 a.u.).

Physical activity endpoints related to high-intensity activity, that is, M6min and M60min (Figure 2D) were significantly different across SNAQ groups. None of the physical activity endpoints showed a significant difference across MNA-SF groups (Table 3).

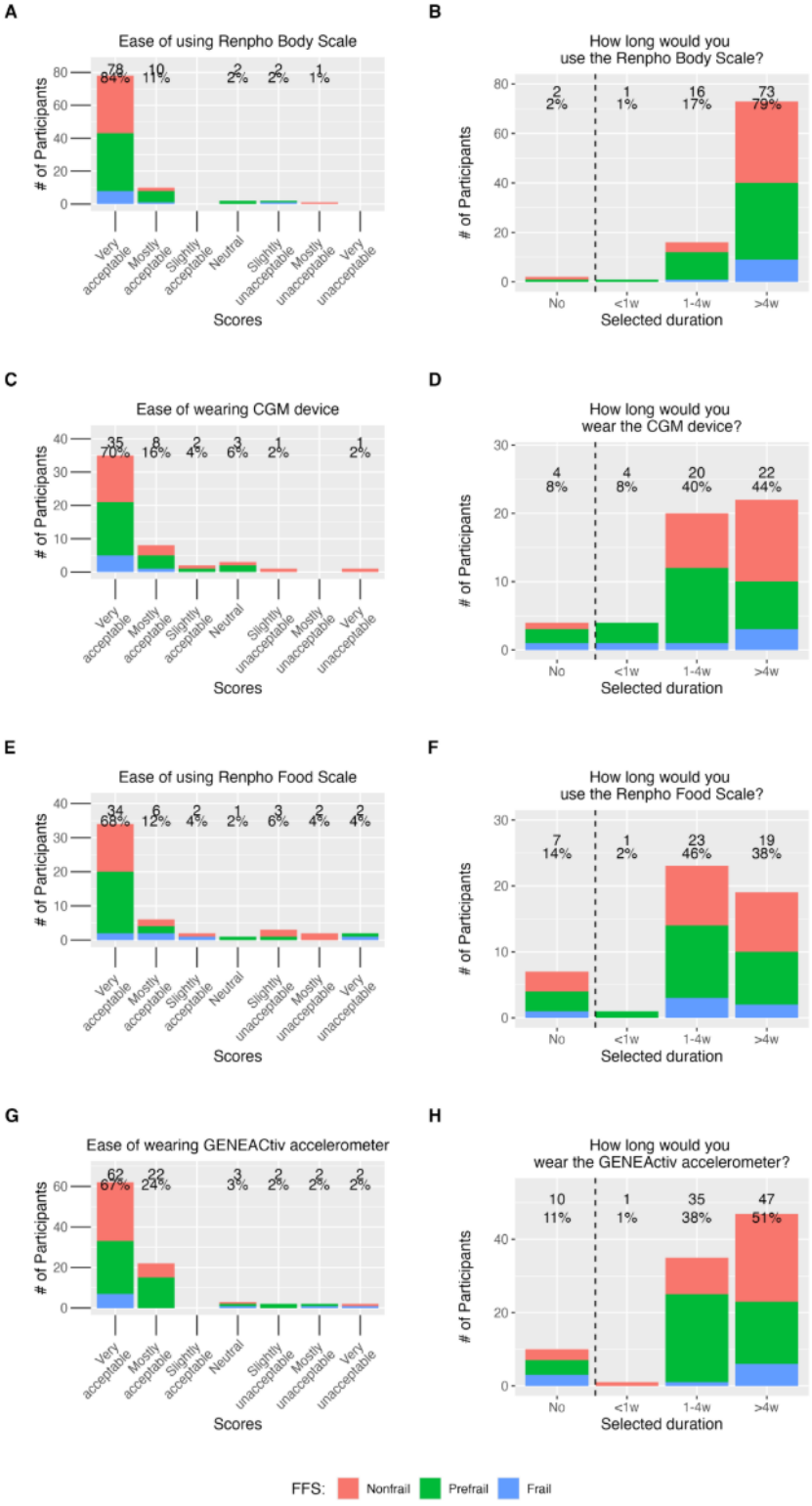
Two participants did not complete the appetite and fatigue questionnaires at home, resulting in a total of 92 (98%) respondents. The question on appetite was answered at home for an average of 13.31 (SD 2.28) days per participant, with an average score of 8.02 (SD 1.87) across participants. The average appetite score differed across frailty groups ($F_{2,87}$=4.53, P=.01), robustness ($F_{1,88}$=5.24, P=.02), and nutritional (MNA-SF: $F_{1,88}$=4.76, P=.03; SNAQ: $F_{1,88}$=18.42, P<.001) groups, as illustrated in Figure S5 in Multimedia Appendix 1. M6min (ρ=0.3, P<.01) and M60min (R=0.28, P<.01) showed a significant correlation with the average appetite score.

The question on fatigue was answered at home for an average of 13.27 (SD 2.34) days per participant, with an average score of 2.95 (SD 2.09) across participants. The average fatigue score differed across frailty groups ($F_{2,87}$=10.01, P<.001), robustness ($F_{1,88}$=14.52, P<.001), and nutritional (MNA-SF: $F_{1,88}$=4.01, P<.05; SNAQ: $F_{1,88}$=6.01, P<0.5) groups, as illustrated in Figure S6 in Multimedia Appendix 1. M60min (ρ=–0.23, P<.05) showed a significant but weak correlation with the average fatigue score.

All participants completed the PROMIS Physical Function 10a questionnaire, resulting in a mean t-score of 52.28 (SD 9.01) across participants. The average t-score differed across frailty ($F_{2,89}$=12.68, P<.001), robustness ($F_{1,90}$=11.6, P<.001), and nutritional (MNA-SF: $F_{1,90}$=4.83, P<.05; SNAQ: $F_{1,90}$=20.2, P<.001) groups, as illustrated in Figure S7 in Multimedia Appendix 1. The time spent in nonsedentary activity (ρ=0.35, P<.001), the time spent in MVPA (ρ=0.42, P<.001; Figure 2E), the mean SVMg (ρ=0.32, P<.01), the 95th percentile of SVMg (ρ=0.4, P<.001), M6min (ρ=0.34, P<.01), and M60min (ρ=0.47, P<.001; Figure 2F) were significantly correlated with the PROMIS Physical Function 10a t-scores.

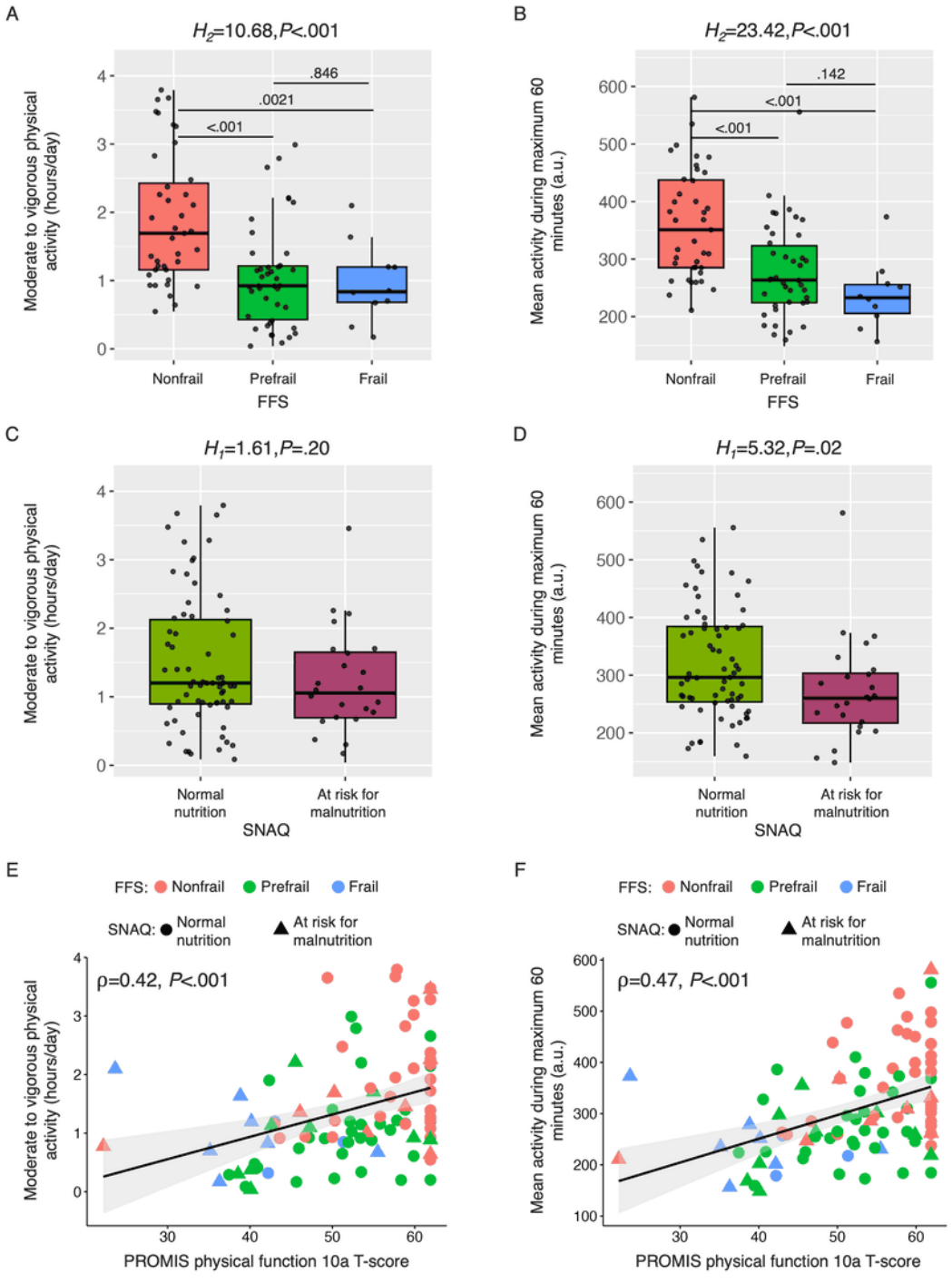
XSL•FO
RenderX

Figure 1. Scores from the comfort and wearability questionnaires, stratified across frailty groups. (A, B) Renpho Body Scale, (C, D), CGM, (E, F), Renpho Food Scale, and (G, H) GENEActiv accelerometer. The color indicates the frailty phenotypes: nonfrail (red), prefrail (green), and frail (blue). Most of the participants found the devices easy to wear or use and indicated their willingness to wear or use them for at least 1 week regardless of their frailty status/score. CGM: continuous glucose monitoring; FFS: Fried Frailty Score.



XSL•FO
RenderX

Figure 2. (A, C, E) MVPA and (B, D, F) and M60min across (A, B) frailty groups, (C, D) across nutritional groups, and (E, F) against PROMIS Physical Function 10a t-scores. (A-D) Data are reported in boxplot representation: central horizontal lines represent the median, boxes mark the IQR, black lines are the whiskers, and black dots indicate outliers. H-scores, degrees of freedom, and *P* value of Kruskal-Wallis checking the association between average endpoints and groups are reported in the title. Colors indicate the frailty group (nonfrail [red], prefrail [green], and frail [blue]) or the nutritional group (normal nutrition [green] and at risk for malnutrition [purple]). Frail participants showed significantly lower MVPA and M60min, and participants at risk for malnutrition showed significantly lower M60min and a smaller MVPA median value. (E, F) Data are reported as scatter plots, where each dot represents the average across participants' daily measurements. Colors indicate the frailty group, while shapes indicate the nutritional group. Spearman's correlation and related *P* values are annotated. Significant correlation with self-reported physical function was observed for both MVPA and M60min. FFS: Fried Frailty Score; M60min: maximum 60 minutes of activity; MVPA: moderate-to-vigorous physical activity; PROMIS: Patient-Reported Outcomes Measurement Information System; SNAQ: Simplified Nutritional Appetite Questionnaire.



XSL•FO
RenderX

Table 2. Physical activity endpoints averaged across all participants and across frailty and robustness groups.

Physical activity endpoint	Total participants (N=91), mean (SD); ANOVA *F*-score, *P* value/Kruskal-Wallis *H* value, *P* value[a]	Nonfrail/robust group (n=39), mean (SD); *P* value vs frail group	Prefrail group (n=42), mean (SD); *P* value vs nonfrail group	Frail group (n=10), mean (SD); *P* value vs prefrail group	Nonrobust group (n=52), mean (SD); ANOVA *F*-score, *P* value/Kruskal-Wallis *H* value, *P* value[a]
Time spent in sedentary activity (hours/day)[b]	11.68 (2.00); $F_{2,86}$=3.50, *P*=.04	11.03 (2.11); .02	12.03 (1.82); .12	12.77 (1.57); .15	12.17 (1.78); $F_{1,87}$=4.78, *P*=.03
Time spent in nonsedentary activity (hours/day)[b]	2.92 (1.48); H_2=13.04, *P*=.002	3.59 (1.38); .02	2.45 (1.41); <.001	2.33 (1.18); .90	2.43 (1.36); H_1=12.97, *P*<.001
Time spent in SLPA[c] (hours/day)	13.20 (1.94); $F_{2,86}$=2.14, *P*=.12	12.72 (2.07); —[d]	13.43 (1.74); —	14.13 (1.86); —	13.56 (1.76); $F_{1,87}$=2.86, *P*=.10
Time spent in MVPA[e] (hours/day)[b]	1.41 (0.93); H_2=20.68, *P*<.001	1.90 (0.96); .002	1.05 (0.75); <.001	0.97 (0.58); .85	1.04 (0.72); H_1=20.60, *P*<.001
SVMg[f] (a.u.[g])[b]	119.9 (29.0); H_2=9.05, *P*=.01	130.3 (27.6); .04	113.0 (26.5); .01	108.1 (23.6); .61	112.1 (25.9); H_1=8.81, *P*=.003
95th Percentile SVMg (a.u.)[b]	348.6 (86.2); $F_{2,86}$=8.19, *P*<.001	394.2 (79.4); .01	316.6 (77.5); <.001	305.3 (66.6); .81	314.4 (75.0); $F_{1,87}$=16.50, *P*<.001
M6min[h] (a.u.)[b]	602.8 (225.0); H_2=16.13, *P*<.001	705.9 (242.0); .002	636.0 (189.0); <.001	480.7 (110.0); .50	525.4 (177.0); H_1=15.57, *P*<.001
M60min[i] (a.u.)[b]	307.8 (96.6); H_2=23.42, *P*<.001	360.2 (93.7); <.001	275.9 (81.9); <.001	238.1 (60.1); .14	268.6 (79.1); H_1=21.34, *P*<.001

[a]*F*-scores and *P* values for ANOVA or *H* values and *P* values for Kruskal-Wallis tests for the effect of frailty and robustness status on physical activity endpoints are reported, along with *P* values of post hoc pairwise comparisons.

[b]Physical activity endpoints significantly different across both frailty and robustness groups.

[c]SLPA: sedentary-to-light physical activity.

[d]Not computed.

[e]MVPA: moderate-to-vigorous physical activity.

[f]SVMg: signal vector magnitude.

[g]a.u.: arbitrary units.

[h]M6min: maximum 6 minutes of activity.

[i]M60min: maximum 60 minutes of activity.

XSL•FO
RenderX

Table 3. Physical activity endpoints averaged across nutritional groups, as assessed using the MNA-SF[a] and the SNAQ[b].

Physical activity endpoint	MNA-SF		SNAQ	
	Normal nutrition (n=65), mean (SD); ANOVA *F*-score, *P* value/Kruskal-Wallis *H* value, *P* value[c]	At risk for malnutrition (n=26), mean (SD); ANOVA *F*-score, *P* value/Kruskal-Wallis *H* value, *P* value[c]	Normal nutrition (n=67), mean (SD); ANOVA *F*-score, *P* value/Kruskal-Wallis *H* value, *P* value[c]	At risk for malnutrition (n=24), mean (SD); ANOVA *F*-score, *P* value/Kruskal-Wallis *H* value, *P* value[c]
Time spent in sedentary activity (hours/day)	11.76 (2.01); $F_{1,87}$=0.60, *P*=.44	11.47 (2.00); $F_{1,87}$=0.60, *P*=.44	11.78 (1.95); $F_{1,87}$=0.18, *P*=.67	11.41 (2.16); $F_{1,87}$=0.18, *P*=.67
Time spent in nonsedentary activity (hours/day)	2.99 (1.48); H_1=0.31, *P*=.58	2.75 (1.49); H_1=0.31, *P*=.58	3.01 (1.51); H_1=0.572, *P*=.45	2.70 (1.39); H_1=0.572, *P*=.45
Time spent in SLPA[d] (hours/day)	13.31 (1.91); $F_{1,87}$=1.16, *P*=.29	12.92 (2.02); $F_{1,87}$=1.16, *P*=.29	13.30 (1.79); $F_{1,87}$=0.36, *P*=.55	12.92 (2.32); $F_{1,87}$=0.36, *P*=.55
Time spent in MVPA[e] (hours/day)	1.44 (0.95); H_1=0.37, *P*=.54	1.31 (0.89); H_1=0.37, *P*=.54	1.48 (0.97); H_1=1.61, *P*=.20	1.19 (0.78); H_1=1.61, *P*=.20
SVMg[f] (a.u.[g])	120.8 (29.3); H_1=0.367, *P*=.54	117.6 (24.7); H_1=0.367, *P*=.54	121.0 (28.5); H_1=0.342, *P*=.56	116.7 (26.7); H_1=0.342, *P*=.56
95th Percentile SVMg (a.u.)	353.2 (87.0); $F_{1,87}$=0.08, *P*=.78	337.0 (84.5); $F_{1,87}$=0.08, *P*=.78	354.4 (86.8); $F_{1,87}$=3.32, *P*=.07	332.3 (83.9); $F_{1,87}$=3.32, *P*=.07
M6min[h] (a.u.)	628.8 (248.0); H_1=1.21, *P*=.27	537.6 (136.7); H_1=1.21, *P*=.27	637.3 (236.0); H_1=7.16, *P*=.007	506.3 (158.0); H_1=7.16, *P*=.007
M60min[i] (a.u.)	316.5 (98.7); H_1=1.30, *P*=.25	286.2 (89.3); H_1=1.30, *P*=.25	320.5 (96.1); H_1=5.32, *P*=.02	272.4 (90.5); H_1=5.32, *P*=.02

[a]MNA-SF: Mini Nutrition Assessment – Short Form.

[b]SNAQ: Simplified Nutritional Appetite Questionnaire.

[c]*F*-scores and *P* values for ANOVA or *H* values and *P* values for Kruskal-Wallis tests for the effect of nutritional status on physical activity endpoints are reported.

[d]SLPA: sedentary-to-light physical activity.

[e]MVPA: moderate-to-vigorous physical activity.

[f]SVMg: signal vector magnitude.

[g]a.u.: arbitrary units.

[h]M6min: maximum 6 minutes of activity.

[i]M60min: maximum 60 minutes of activity.

DHTs Can Enable Patient-Centric Clinical Trials in Older Individuals

The accuracy of the weight and body composition measurements from the Renpho Body Scale and the Tanita scale was compared during in-clinic visits. Perfect agreement (ICC=1) and correlation (R=1) were observed for weight. Mean percentage errors<1% were found for weight and water mass and <4% for muscle mass, fat mass, and fat-free mass. Interestingly, a higher error and nonhomogeneous bias were observed when considering absolute rather than percentage measurements. A comprehensive comparison between the Renpho Body Scale and the Tanita scale is reported in Multimedia Appendix 1.

Compliant days at home using the Renpho Body Scale were defined as those with at least one measurement, resulting in an average percentage compliance of 88.36% (SD 20.49%) across participants. Only 1 (1%) participant did not collect any at-home measurements (0% compliance; this participant felt unsafe using the scale without assistance), while 39 (41%) participants weighed themselves every day (100% compliance). Furthermore, 2 (2%) participants were unable to connect the body scale to the app; thus, only their weight was captured, not the body composition measures.

Questionnaire responses showed that 88 (95%) of 93 participants found the Renpho Body Scale mostly to very acceptable to use, and 89 (97%) of 92 participants indicated that they would use the body scale for more than 1 week.

ANOVA did not show any statistically significant difference across frailty, robustness, or nutritional groups for any of the Renpho Body Scale endpoints. As expected, height showed a significant effect on weight (*P*<.001). In addition, sex showed a significant effect on muscle mass (females: mean 64.96%, SD 8.37%; males: mean 72.85%, SD 4.90%; *P*<.001), fat mass (females: mean 30.92%, SD 8.87%; males: mean 23.30%, SD 5.13%; *P*<.01), water mass (females: mean 47.46%, SD 6.21%; males: mean 55.32%, SD 3.84%; *P*<.001), and fat-free mass (females: mean 69.08%, SD 8.88%; males: mean 76.70%, SD 5.13%; *P*<.01).

Muscle mass (ρ=–0.32, *P*<.01), bone mass (ρ=–0.23, *P*<.05), fat mass (ρ=0.32, *P*<.01), fat-free mass (ρ=–0.32, *P*<.01), and water mass (ρ=–0.31, *P*<.01) showed a significant but weak

XSL•FO
RenderX

correlation with self-reported fatigue. Similarly, muscle mass (ρ=0.32, *P*<.01), bone mass (ρ=0.21, *P*<0.05), fat mass (ρ=–0.32, *P*<.01), fat-free mass (ρ=0.32, *P*<.01), and water mass (ρ=0.31, *P*<.01) showed significant correlations with self-reported physical function.

Of 50 participants, CGM data were available for 43 (86%) participants: For 2 (4%) participants, the devices failed to connect to the reader at the last visit; 1 (2%) participant lost the device during the study; for 3 (6%) participants, the device fell off during the second day of monitoring; and for 1 (2%) participant, the device appeared damaged at the last visit (with the adhesive partially removed and the needle bent sideways).

Excluding visit days, CGM data were recorded for an average of 10.12 (SD 3.37, range 1.00-14.00) days, corresponding to an average of 84.27% (SD 26.68%, range 8.33%-100.00%) of the monitoring period. In total, 25 (50%) participants wore the CGM every day (100% compliance). During days with CGM recordings, the average wear time was 23.23 (SD 1.39, range 18.00-24.00) hours across participants.

Questionnaire responses showed that 43 (86%) of 50 participants found the CGM device mostly to very acceptable to wear, and 42 (84%) of 50 participants indicated that they would use the CGM for more than 1 week.

ANOVA did not show any statistically significant differences across frailty, robustness, or nutritional groups for any CGM endpoints. However, the frail group showed a higher TBR (~7%), and a lower TIR (~8%) compared to nonfrail and prefrail groups (Table S3 in Multimedia Appendix 1).

The minimum glucose concentration showed a significant correlation with self-reported appetite (ρ=0.32, *P*=.04) and self-reported fatigue (ρ=–0.31, *P*=.05).

Participants were considered compliant with using the Renpho Nutrition Scale only if they weighed a meal both before and after consumption. The percentage of days with at least one compliant meal entry was an average of 60.29% (SD 37.00%, range 0.00%-100.00%) across all participants. In addition, 1 (2%) participant was unable to record any meals with the nutrition scale throughout the study, while 4 (8%) participants weighed their meals only once (either before or after consumption) throughout the study. Furthermore, 5 (10%) participants were compliant every day (100% compliance), and 34 (68%) reported at least three meals on an average of 53.40% (SD 24.68%, range 6.66%-92.31%) of the monitoring days. Finally, across participants, a mean total of 25.34 (SD 23.24, range 0.00-121.00) meals were correctly entered during the study using the Renpho Nutrition Scale, resulting in 1267 total meals.

Questionnaire responses showed that 40 (80%) of 50 participants found the nutrition scale mostly to very acceptable to use, and 42 (84%) of 50 participants indicated that they would use the nutrition scale for more than 1 week.

The histograms of meal net weight, meal duration, time between consecutive daily meals, and number of daily meals are reported in Figure S8 in Multimedia Appendix 1. ANOVA showed significant differences between frailty groups only for the net sodium ($F_{2,45}$=4.4, *P*=.02) and net potassium ($F_{2,45}$=3.6, *P*=.03) content, which were lower in the frail group compared to the nonfrail (*P*<.05) and prefrail (*P*<.05) groups. However, no significant differences across robustness groups were observed for any meal endpoints (Table S3 in Multimedia Appendix 1).

Sex showed significant effects on meal net weight (females: mean 263.3, SD 79.8; males: mean 322.2, SD 72.2; *P*<.05) and meal net protein (females: mean 19.71, SD 6.91; males: mean 25.39, SD 5.14; *P*<.05). Age showed a significant effect on meal duration ($F_{1,45}$=9.79, *P*<.01), with older participants being associated with a longer meal duration. When grouping by SNAQ (but not MNA-SF), participants at risk for malnutrition reported a significantly lower number of daily meals ($F_{1,46}$=6.89, *P*<.05) and a significantly lower meal net weight ($F_{1,46}$=6.83, *P*<.05), as summarized in Table S4 in Multimedia Appendix 1. After adding the meal net weight as a model covariate, participants at risk for malnutrition exhibited a significantly longer meal duration ($F_{1,45}$=6.29, *P*<.05).

The meal net weight showed a significant correlation with self-reported appetite (ρ=0.30, *P*<.05) and physical function (ρ=0.48, *P*<.001).

Discussion

Principal Findings

We explored the use of DHTs, including wrist accelerometers, body scales, CGM, and food scales, to remotely monitor the overall well-being of older individuals at risk for frailty and malnutrition. We observed that physical activity endpoints collected at home with a wrist accelerometer and in the lab were significantly different between frailty and nutritional groups. Across all participants, a higher average daily time in SLPA than in MVPA was observed; however, frailty status had a significant effect only on MVPA, with the frail group spending significantly less time in MVPA than the nonfrail group, confirming previous findings [51,52]. A systematic review by Tolley et al [53] reported that time spent in sedentary activity is usually significantly lower in individuals who are frail compared with those who are not. However, the relationship between sedentary time and frailty status is not always consistent [54], and higher sedentary time in nonfrail compared with frail cohorts has also been reported [55]. In this work, sedentary time was significantly higher in the frail compared with the nonfrail group and in the nonrobust compared with the robust group. These findings confirm those reported in our previous study [35], where the frail group spent more time in sedentary activity and less time in MVPA and exhibited significantly a lower 95th percentile of SVMg. In this research, additional physical activity endpoints, including time spent in nonsedentary activity, mean SVMg, and mean activity during M6min and M60min, distinguished frail and nonrobust groups from nonfrail and robust groups. This difference is possibly related to the larger sample size.

A link between nutritional status and physical activity in older adults has already been hypothesized in the literature: Tamamura et al [56] showed that patients with higher MNA-SF scores at admission to rehabilitation wards had greater functional

XSL•FO
RenderX

independence at discharge, and Hsueh et al [57] observed that individuals without a healthy diet were more sedentary compared with those with a healthy diet. In this study, we showed that objectively and remotely measured mean activity during M6min and M60min are significantly lower in participants at risk for malnutrition compared with those with normal nutritional status, as assessed with the SNAQ, confirming the relationship between physical activity and nutritional status in older individuals. Notably, physical activity endpoints were not sensitive to MNA-SF groups, suggesting that the SNAQ might be preferable as an anchor metric in studies assessing minimal clinically important differences in digital endpoints.

Wrist-derived physical activity endpoints have already been associated with self-reported behaviors in older individuals [58]; here, we confirmed the relationship between physical activity endpoints and self-reported measures of appetite [59], fatigue [60], and physical function [61]. These findings suggest that digital endpoints collected in free-living settings are concordant with individuals' perceptions, while unobtrusively providing more objective and sensitive data to complement PROs, as already noted in the literature [37,62]. Indeed, a ceiling effect was observed in self-reported physical function scores but not in physical activity digital endpoints. As noted previously [63], despite the observed correlation, subjective and objective physical activity measures may differ and should potentially be used complementarily [64].

Our findings support the deployment of wrist accelerometers in clinical studies as effective, noninvasive, and highly informative tools for monitoring older adults' physical activity in free-living conditions and for obtaining insights into their frailty and nutritional status. These tools enable vulnerable populations, such as older individuals, to participate in clinical trials with fewer burdensome visits, thereby expanding access to care. In addition, tracking daily levels of physical activity can help health care professionals assess mobility, sleep, and overall well-being, in turn supporting the development and adjustment of coaching programs for older individuals [65].

Furthermore, we showed that other DHTs can be effectively deployed for older individuals to remotely monitor their weight, body composition, glucose concentration, and meal intake, thus enabling more patient-centric and data-centric clinical trials.

In this work, we showed that the Renpho Body Scale, a small and cost-effective device (~US $30 compared with ~US $7000 for the Tanita scale) that offers scalability and ease of use for both participants and health care professionals, provides accurate measures of weight and percentages of muscle mass, fat mass, fat-free mass, and water mass compared to the research-grade Tanita scale. Although these findings should be further corroborated by comparison against other highly accurate measures of body composition (eg, dual X-ray absorptiometry), they provide evidence for the feasibility of frequent monitoring of weight and body composition in older individuals. These measures can support monitoring underlying conditions, such as AoA, fluid retention, or functional decline, and may be indicative of medication effects [66]. For example, daily body composition measurements can reveal reductions in bone mass, which may indicate osteoporosis, or reductions in muscle mass, which may lead to sarcopenia [67].

Data collected in this study using a digital nutrition scale showed that participants at risk for malnutrition consumed a significantly lower number of daily meals, resulting in lower macronutrient intake, confirming previous findings [68,69]. However, as suggested by Madeira et al [68], this relationship is not absolute: eating more meals does not guarantee good nutritional status if the meals are low in nutritional value. Therefore, a digital nutrition scale represents a useful tool for comprehensive assessment of total dietary intake and nutritional quality in clinical trials. These measures are vital for establishing dietary treatment strategies, tailored to meet individuals' needs and to prevent energy and nutrient deficiencies, which are associated with increased length of hospital stay, hospital readmission, decreased survival time, and increased incidence of infection in hospitalized older patients [70,71]. Nevertheless, the lower compliance with this device suggests the need for more automated tools to track dietary behavior. For example, CGM data coupled with suitable artificial intelligence algorithms might offer a solution for automatically detecting meal intakes [72].

The importance of monitoring daily glucose fluctuations, including hypoglycemic periods and postprandial peaks, has been increasingly investigated in recent years, and correlations with various microvascular and macrovascular complications have been identified [73]. Although CGM devices are widely used in type 1 diabetes to capture dynamic glycemic profiles, the applicability in older individuals has been less extensively explored. Uotani et al [30] found that women with eating disorders exhibit higher glucose variability and a greater propensity for hypoglycemia, suggesting CGM as a valuable tool for mitigating the risk for severe morbidity and mortality in this population. Zhong et al [74] used CGM for 2 weeks in 698 healthy older individuals (median age 69.1 years), showing excellent glycemic control with median TIR, TBR, and TAR of 94.7%, 3.29%, and 0.64%, respectively. Compared to these values, we observed a higher average TBR, suggesting that older participants at risk for AoA may have poorer glycemic control and may be more susceptible to hypoglycemia.

Understanding older adults' perceptions of new technologies is crucial for deploying these technologies in clinical trials and everyday life [75,76]. Evaluating the acceptance of DHTs involves measuring both participants' self-reported usability and acceptability as well as their level of continued use. In this study, more than 80% of participants reported the devices to be mostly to very acceptable to wear/use and indicated that they would wear them for more than 1 week. These findings support the deployment of DHTs for older individuals for longer durations, for example, to assess treatment effects in interventional studies or to track physical and physiological changes over time in longitudinal studies, thereby accessing more granular data than those collected sporadically during in-clinic visits [77]. Compliance was also high: all participants were compliant for more than 90% of monitoring days with the wrist accelerometer, more than 85% with the digital body scale, and more than 60% with the digital nutrition scale. CGM data were available for 43 of 50 participants, who were compliant

XSL•FO
RenderX

for more than 80% of monitoring days. These findings support the deployment of DHTs in clinical trials involving older individuals to enhance patient- and data-centricity. Although limited data loss was observed in this study, some issues with CGM and the nutrition scale highlighted the importance of providing clear and detailed training materials to clinical sites for future studies, potentially supported by artificial intelligence [78].

Limitations

This study has several limitations. First, because it was cross-sectional, causality between digital endpoints, frailty, malnutrition, and other PROs cannot be established. In addition, the sample size across frailty groups was unbalanced, with the frail group accounting for only about 10% of the dataset and all being female. To mitigate this limitation, analyses were also performed by grouping nonfrail and prefrail groups into the nonrobust category and by including sex as a covariate in ANOVA models. Finally, because meal assessment occurred in an unsupervised environment, results may have been affected by unreported or misreported meals.

Conclusion

In conclusion, monitoring physical activity, weight, and dietary intake represents a cornerstone of effective health management in daily life. As the population ages worldwide, with the number of people aged 60 years and older expected to double by 2050 to approximately 2 billion, understanding the implications of aging becomes crucial for policymakers, health care professionals, and society. In this context, digitalized data collection represents a reliable and cost-effective approach to reducing in-person visits, while providing health care providers with rich longitudinal data on older individuals living in the community or in assisted-living facilities.

Acknowledgments

The two studies were conducted as a collaboration between Boston University and Pfizer. Boston University was the study sponsor.

Data Availability

Upon request, and subject to review, the study sponsor, Boston University, will provide data that support the findings of this study. Subject to certain criteria, conditions, and exceptions, Boston University may also provide access to the related individual deidentified participant data. Contact Boston University for more information. Further inquiries can be directed to the corresponding author.

Conflicts of Interest

NC, DJP, AM, and FIK are Pfizer employees and might own Pfizer stocks. GK was a paid contractor to Pfizer in connection with the development of this paper. MW, NSK, HZ, IM, ZH, MK, CD, MS, DC and KCT have no conflicts of interest to declare.

Multimedia Appendix 1
Additional information about the Geriatric Anorexia Studies.
[DOCX File , 1681 KB - aging_v9i1e77816_app1.docx]

References

1. McLaughlin S, Connell C, Heeringa S, Li L, Roberts J. Successful aging in the United States: prevalence estimates from a national sample of older adults. J Gerontol B Psychol Sci Soc Sci 2010 Mar;65B(2):216-226 [FREE Full text] [doi: 10.1093/geronb/gbp101] [Medline: 20008481]
2. Maestas N, Mullen KJ, Powell D. The effect of population aging on economic growth, the labor force, and productivity. Am Econ J Macroecon 2023 Apr 01;15(2):306-332 [FREE Full text] [doi: 10.1257/mac.20190196]
3. Landi F, Calvani R, Tosato M, Martone AM, Ortolani E, Savera G, et al. Anorexia of aging: risk factors, consequences, and potential treatments. Nutrients 2016 Jan 27;8(2):69 [FREE Full text] [doi: 10.3390/nu8020069] [Medline: 26828516]
4. Aprahamian I, Coats A, Morley J, Klompenhouwer T, Anker SD, International Advisory Board, and Regional Advisory Boards for North America, Latin America, Europe and Japan. Anorexia of aging: an international assessment of healthcare providers' knowledge and practice gaps. J Cachexia Sarcopenia Muscle 2023 Dec;14(6):2779-2792 [FREE Full text] [doi: 10.1002/jcsm.13355] [Medline: 37897129]
5. Perera LAM, Chopra A, Shaw AL. Approach to patients with unintentional weight loss. Med Clin North Am 2021 Jan;105(1):175-186 [FREE Full text] [doi: 10.1016/j.mcna.2020.08.019] [Medline: 33246517]
6. AlFehaidi A, Khan S, Abdelrahman R, Ahel NT, Shine P, De Ramos MD, et al. Predictors of malnutrition among older residents in Qatari long-term care facilities: a retrospective study. BMC Nutr 2024 Feb 01;10(1):23 [FREE Full text] [doi: 10.1186/s40795-024-00827-z] [Medline: 38303045]
7. Ono S, Sasabuchi Y, Yamana H, Yokota I, Okada A, Matsui H, et al. Weight loss and functional decline in older Japanese people: a cohort study using large-scale claims data. Arch Gerontol Geriatr 2024 May;120:105354 [FREE Full text] [doi: 10.1016/j.archger.2024.105354] [Medline: 38309105]

XSL•FO
RenderX

8. Hoogendijk E, van Kan GA, Guyonnet S, Vellas B, Cesari M. Components of the frailty phenotype in relation to the frailty index: results from the Toulouse frailty platform. J Am Med Dir Assoc 2015 Oct 01;16(10):855-859 [FREE Full text] [doi: 10.1016/j.jamda.2015.04.007] [Medline: 25979777]
9. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar 01;56(3):M146-M156 [FREE Full text] [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
10. Takayanagi N, Sudo M, Yamashiro Y, Lee S, Kobayashi Y, Niki Y, et al. Relationship between daily and in-laboratory gait speed among healthy community-dwelling older adults. Sci Rep 2019 Mar 05;9(1):3496 [FREE Full text] [doi: 10.1038/s41598-019-39695-0] [Medline: 30837520]
11. Czech M, Psaltos D, Zhang H, Adamusiak T, Calicchio M, Kelekar A, et al. Age and environment-related differences in gait in healthy adults using wearables. NPJ Digit Med 2020;3:127 [FREE Full text] [doi: 10.1038/s41746-020-00334-y] [Medline: 33083562]
12. McCambridge J, Witton J, Elbourne D. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol 2014 Mar;67(3):267-277 [FREE Full text] [doi: 10.1016/j.jclinepi.2013.08.015] [Medline: 24275499]
13. Batsis J, DiMilia P, Seo L, Fortuna KL, Kennedy MA, Blunt HB, et al. Effectiveness of ambulatory telemedicine care in older adults: a systematic review. J Am Geriatr Soc 2019 Aug;67(8):1737-1749 [FREE Full text] [doi: 10.1111/jgs.15959] [Medline: 31066916]
14. Ní Scanaill C, Carew S, Barralon P, Noury N, Lyons D, Lyons G. A review of approaches to mobility telemonitoring of the elderly in their living environment. Ann Biomed Eng 2006 Apr;34(4):547-563 [FREE Full text] [doi: 10.1007/s10439-005-9068-2] [Medline: 16550450]
15. Cordeiro JV. Digital technologies and data science as health enablers: an outline of appealing promises and compelling ethical, legal, and social challenges. Front Med (Lausanne) 2021;8(4):647897-647454 [FREE Full text] [doi: 10.3389/fmed.2021.647897] [Medline: 34307394]
16. Kim B, Hunt M, Muscedere J, Maslove DM, Lee J. Using consumer-grade physical activity trackers to measure frailty transitions in older critical care survivors: exploratory observational study. JMIR Aging 2021 Feb 23;4(1):e19859 [FREE Full text] [doi: 10.2196/19859] [Medline: 33620323]
17. Zhang X, Feng L, Munster B, Hobbelen H, Lamoth CJ. Distinguishing (pre)frail from non-frail older adults based on walking pattern: a scoping review on gait parameters derived from inertial sensors. Gait Posture 2024 Sep;113:267. [doi: 10.1016/j.gaitpost.2024.07.287]
18. Murphy RA, Patel KV, Kritchevsky SB, Houston DK, Newman AB, Koster A, et al. Weight change, body composition, and risk of mobility disability and mortality in older adults: a population-based cohort study. J Am Geriatr Soc 2014 Aug;62(8):1476-1483 [FREE Full text] [doi: 10.1111/jgs.12954] [Medline: 25039391]
19. Ulugerger Avci G, Bektan Kanat B, Can G, Suzan V, Unal D, Degirmenci P, et al. The impact of sarcopenia and obesity on mortality of older adults: five years results. Ir J Med Sci 2023 Oct;192(5):2209-2216. [doi: 10.1007/s11845-023-03392-9] [Medline: 37202585]
20. Rudzińska A, Piotrowicz K, Perera I, Gryglewska B, Gąsowski J. Poor appetite in frail older persons-a systematic review. Nutrients 2023 Jun 29;15(13):2966 [FREE Full text] [doi: 10.3390/nu15132966] [Medline: 37447292]
21. Cox N, Morrison L, Ibrahim K, Robinson S, Sayer A, Roberts H. New horizons in appetite and the anorexia of ageing. Age Ageing 2020 Jul 01;49(4):526-534 [FREE Full text] [doi: 10.1093/ageing/afaa014] [Medline: 32043144]
22. Gauthier A, Jaunzarins B, MacDougall SJ, Laurence M, Kabaroff JL, Godwin AA, et al. Evaluating the reliability of assessing home-packed food items using digital photographs and dietary log sheets. J Nutr Educ Behav 2013;45(6):708-712 [FREE Full text] [doi: 10.1016/j.jneb.2013.05.005] [Medline: 23830245]
23. Tanweer A, Khan S, Mustafa F, Imran S, Humayun A, Hussain Z. Improving dietary data collection tools for better nutritional assessment – a systematic review. Comput Methods Programs Biomed Update 2022;2:100067 [FREE Full text] [doi: 10.1016/j.cmpbup.2022.100067]
24. Gariballa S, Forster S. Dietary intake of older patients in hospital and at home: the validity of patient kept food diaries. J Nutr Health Aging 2008 Feb;12(2):102-106 [FREE Full text] [doi: 10.1007/bf02982561]
25. Kalinowska K, Wojnowski W, Tobiszewski M. Smartphones as tools for equitable food quality assessment. Trends Food Sci Technol 2021 May;111:271-279 [FREE Full text] [doi: 10.1016/j.tifs.2021.02.068]
26. Iizuka K, Deguchi K, Ushiroda C, Yanagi K, Seino Y, Suzuki A, et al. A study on the compatibility of a food-recording application with questionnaire-based methods in healthy Japanese individuals. Nutrients 2024 Jun 02;16(11):e14778 [FREE Full text] [doi: 10.3390/nu16111742] [Medline: 38892675]
27. Theodore Armand TP, Kim HC, Kim JI. Digital anti-aging healthcare: an overview of the applications of digital technologies in diet management. J Pers Med 2024 Feb 27;14(3) [FREE Full text] [doi: 10.3390/jpm14030254] [Medline: 38540996]
28. Ryu J, Kim S, Lim Y, Ohn JH, Kim S, Cho JH, et al. Sodium intake estimation in hospital patients using AI-based imaging: prospective pilot study. JMIR Form Res 2024 Feb 16;8:e48690 [FREE Full text] [doi: 10.2196/48690] [Medline: 38363594]
29. Galindo R, Aleppo G. Continuous glucose monitoring: the achievement of 100 years of innovation in diabetes technology. Diabetes Res Clin Pract 2020 Dec;170:108502 [FREE Full text] [doi: 10.1016/j.diabres.2020.108502] [Medline: 33065179]

XSL•FO
RenderX

30. Uotani N, Noma S, Akamine M, Miyawaki T. Continuous glucose monitoring for detection of glycemic variability, hypoglycemia, and hyperglycemia in women with eating disorders. Biopsychosoc Med 2022 Oct 27;16(1):22 [FREE Full text] [doi: 10.1186/s13030-022-00251-4] [Medline: 36303193]
31. Hegedus E, Salvy SJ, Wee CP, Naguib M, Raymond JK, Fox DS, et al. Use of continuous glucose monitoring in obesity research: a scoping review. Obes Res Clin Pract 2021;15(5):431-438 [FREE Full text] [doi: 10.1016/j.orcp.2021.08.006] [Medline: 34481746]
32. Presseller EK, Parker MN, Lin M, Weimer J, Juarascio AS. The application of continuous glucose monitoring technology to eating disorders research: an idea worth researching. Int J Eat Disord 2020 Dec;53(12):1901-1905. [doi: 10.1002/eat.23404] [Medline: 33159708]
33. Lupoli R, Lembo E, Ciciola P, Schiavo L, Pilone V, Capaldo B. Continuous glucose monitoring in subjects undergoing bariatric surgery: diurnal and nocturnal glycemic patterns. Nutr Metab Cardiovasc Dis 2020 Oct 30;30(11):1954-1960 [FREE Full text] [doi: 10.1016/j.numecd.2020.06.029] [Medline: 32807631]
34. Sadıç A, Şahiner Z, Eşme M, Balcı C, Doğu BB, Cankurtaran M, et al. Malnutrition and frailty as independent predictors of adverse outcomes in hospitalized older adults: a prospective single center study. Medicina (Kaunas) 2025 Jul 26;61(8):375-377 [FREE Full text] [doi: 10.3390/medicina61081354] [Medline: 40870398]
35. Camerlingo N, Shaafi Kabiri N, Psaltos DJ, Kelly M, Wicker MK, Messina I, et al. Monitoring gait and physical activity of elderly frail individuals in free-living environment: a feasibility study. Gerontology 2024;70(4):439-454 [FREE Full text] [doi: 10.1159/000535283] [Medline: 37984340]
36. Omran M, Morley J. Assessment of protein energy malnutrition in older persons, part I: history, examination, body composition, and screening tools. Nutrition 2000 Jan;16(1):50-63 [FREE Full text] [doi: 10.1016/S0899-9007(99)00224-5] [Medline: 33159708]
37. McDermott MM, Liu K, O'Brien E, Guralnik JM, Criqui MH, Martin GJ, et al. Measuring physical activity in peripheral arterial disease: a comparison of two physical activity questionnaires with an accelerometer. Angiology 2000 Feb 30;51(2):91-100 [FREE Full text] [doi: 10.1177/000331970005100201] [Medline: 10701716]
38. Winter J, Flanagan D, McNaughton SA, Nowson C. Nutrition screening of older people in a community general practice, using the MNA-SF. J Nutr Health Aging 2013 Apr;17(4):322-325 [FREE Full text] [doi: 10.1007/s12603-013-0020-0] [Medline: 23538653]
39. Wang T, Shen J. Usefulness of Simplified Nutritional Appetite Questionnaire (SNAQ) in appetite assessment in elder patients with liver cirrhosis. J Nutr Health Aging 2018 Feb;22(8):911-915 [FREE Full text] [doi: 10.1007/s12603-018-1086-5] [Medline: 30272092]
40. Acar Tek N, Karaçil-Ermumcu M. Determinants of health related quality of life in home dwelling elderly population: appetite and nutritional status. J Nutr Health Aging 2018 Apr;22(8):996-1002 [FREE Full text] [doi: 10.1007/s12603-018-1066-9] [Medline: 30272105]
41. Rolland Y, Perrin A, Gardette V, Filhol N, Vellas B. Screening older people at risk of malnutrition or malnourished using the Simplified Nutritional Appetite Questionnaire (SNAQ): a comparison with the Mini-Nutritional Assessment (MNA) tool. J Am Med Dir Assoc 2012 Jan;13(1):31-34 [FREE Full text] [doi: 10.1016/j.jamda.2011.05.003] [Medline: 21700503]
42. Perfetto EM, Burke L, Oehrlein EM, Epstein RS. Patient-focused drug development: a new direction for collaboration. Med Care 2015 Jan;53(1):9-17 [FREE Full text] [doi: 10.1097/MLR.0000000000000273] [Medline: 25494232]
43. Terwee C, Crins M, Roorda L, Cook KF, Cella D, Smits N, et al. International application of PROMIS computerized adaptive tests: US versus country-specific item parameters can be consequential for individual patient scores. J Clin Epidemiol 2021 Jun;134:1-13 [FREE Full text] [doi: 10.1016/j.jclinepi.2021.01.011] [Medline: 33524487]
44. Erb MK, Karlin DR, Ho BK, Thomas KC, Parisi F, Vergara-Diaz GP, et al. mHealth and wearable technology should replace motor diaries to track motor fluctuations in Parkinson's disease. NPJ Digit Med 2020 Feb 02;3(2):6-9 [FREE Full text] [doi: 10.1038/s41746-019-0214-x] [Medline: 31970291]
45. Mahadevan N, Christakis Y, Di J, Bruno J, Zhang Y, Dorsey ER, et al. Development of digital measures for nighttime scratch and sleep using wrist-worn wearable devices. NPJ Digit Med 2021 Mar 03;4(1):42-e1010 [FREE Full text] [doi: 10.1038/s41746-021-00402-x] [Medline: 33658610]
46. Howell J, Strong B, Weisenberg J, Kakade A, Gao Q, Cuddihy P, et al. Maximum daily 6 minutes of activity: an index of functional capacity derived from actigraphy and its application to older adults with heart failure. J Am Geriatr Soc 2010 May;58(5):931-936 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.02805.x] [Medline: 20374397]
47. Saville DJ. Multiple comparison procedures: the practical solution. Am Stat 1990 May;44(2):174-180 [FREE Full text] [doi: 10.2307/2684163]
48. Cicchetti D, Bronen R, Spencer S, Haut S, Berg A, Oliver P, et al. Rating scales, scales of measurement, issues of reliability: resolving some critical issues for clinicians and researchers. J Nerv Ment Dis 2006 Aug;194(8):557-564 [FREE Full text] [doi: 10.1097/01.nmd.0000230392.83607.c5] [Medline: 16909062]
49. Troiano R, McClain J, Brychta R, Chen K. Evolution of accelerometer methods for physical activity research. Br J Sports Med 2014 Jul;48(13):1019-1023 [FREE Full text] [doi: 10.1136/bjsports-2014-093546] [Medline: 24782483]

XSL•FO
RenderX

50. Lohne-Seiler H, Hansen B, Kolle E, Anderssen S. Accelerometer-determined physical activity and self-reported health in a population of older adults (65-85 years): a cross-sectional study. BMC Public Health 2014 Mar 27;14:284 [FREE Full text] [doi: 10.1186/1471-2458-14-284] [Medline: 24673834]
51. Kikuchi H, Inoue S, Amagasa S, Fukushima N, Machida M, Murayama H, et al. Associations of older adults' physical activity and bout-specific sedentary time with frailty status: compositional analyses from the NEIGE study. Exp Gerontol 2021 Jan;143:111149 [FREE Full text] [doi: 10.1016/j.exger.2020.111149] [Medline: 33181316]
52. Razjouyan J, Naik AD, Horstman MJ, Kunik ME, Amirmazaheri M, Zhou H, et al. Wearable sensors and the assessment of frailty among vulnerable older adults: an observational cohort study. Sensors (Basel) 2018 Apr 26;18(5):1336 [FREE Full text] [doi: 10.3390/s18051336] [Medline: 29701640]
53. Tolley APL, Ramsey KA, Rojer AGM, Reijnierse EM, Maier AB. Objectively measured physical activity is associated with frailty in community-dwelling older adults: a systematic review. J Clin Epidemiol 2021 Sep;137:218-230 [FREE Full text] [doi: 10.1016/j.jclinepi.2021.04.009] [Medline: 33915264]
54. Park KN, Kim SH. Consumer wearable device-based measures of physical activity and energy expenditure in community-dwelling older adults with different levels of frailty: a STROBE compliant study. Medicine (Baltimore) 2022 Dec 30;101(52):e31863 [FREE Full text] [doi: 10.1097/MD.0000000000031863] [Medline: 36595984]
55. Chen S, Chen T, Kishimoto H, Yatsugi H, Kumagai S. Associations of objectively measured patterns of sedentary behavior and physical activity with frailty status screened by the Frail Scale in Japanese community-dwelling older adults. J Sports Sci Med 2020 Mar;19(1):166-174 [FREE Full text] [Medline: 32132840]
56. Tamamura Y, Hachiuma C, Matsuura M, Shiba S, Nishikimi T. Relationship between improvement in physical activity and three nutritional assessment indicators in patients admitted to a convalescent rehabilitation ward. Nutrients 2024 Aug 02;16(15):275-283 [FREE Full text] [doi: 10.3390/nu16152531] [Medline: 39125410]
57. Hsueh MC, Rutherford R, Huang YH, Chang Chien HY, Chang CH, Park JH, et al. Are older adults without a healthy diet less physically active and more sedentary? Nutrients 2019 May 20;11(5):267 [FREE Full text] [doi: 10.3390/nu11051119] [Medline: 31137459]
58. Maes I, Mertens L, Poppe L, Vetrovsky T, Crombez G, De Backere F, et al. Within-person associations of accelerometer-assessed physical activity with time-varying determinants in older adults: time-based ecological momentary assessment study. JMIR Aging 2023 Nov 23;6:e44425 [FREE Full text] [doi: 10.2196/44425] [Medline: 37995131]
59. Tsai LT, Boyle E, Buhl SF, Kock G, Brønd JC, Visser M, et al. Associations between appetite, physical activity and sedentary behaviour from hip- and wrist-worn accelerometers in community-dwelling older adults. Geriatr Gerontol Int 2023 Jun;23(6):411-417. [doi: 10.1111/ggi.14588] [Medline: 37098733]
60. Zaslavsky O, Su Y, Rillamas-Sun E, Roopsawang I, LaCroix AZ. Accelerometer-measured physical activity levels and fatigue in older women. J Aging Phys Act 2020 Oct 01;28(5):692-698. [doi: 10.1123/japa.2019-0308] [Medline: 32303001]
61. Barone Gibbs B, Sternfeld B, Whitaker KM, Brach JS, Hergenroeder AL, Jacobs DR, et al. Bidirectional associations of accelerometer-derived physical activity and stationary behavior with self-reported mental and physical health during midlife. Int J Behav Nutr Phys Act 2021 Jun 06;18(1):74 [FREE Full text] [doi: 10.1186/s12966-021-01145-4] [Medline: 34090471]
62. Oyeyemi A, Umar M, Oguche F, Aliyu S, Oyeyemi A. Accelerometer-determined physical activity and its comparison with the International Physical Activity Questionnaire in a sample of Nigerian adults. PLoS One 2014;9(1):e87233 [FREE Full text] [doi: 10.1371/journal.pone.0087233] [Medline: 24489876]
63. Vetter V, Özince DD, Kiselev J, Düzel S, Demuth I. Self-reported and accelerometer-based assessment of physical activity in older adults: results from the Berlin Aging Study II. Sci Rep 2023 Jun 21;13(1):10047 [FREE Full text] [doi: 10.1038/s41598-023-36924-5] [Medline: 37344489]
64. Yang Y, Hirdes JP, Dubin JA, Lee J. Fall risk classification in community-dwelling older adults using a smart wrist-worn device and the resident assessment instrument-home care: prospective observational study. JMIR Aging 2019 Jun 07;2(1):e12153 [FREE Full text] [doi: 10.2196/12153] [Medline: 31518278]
65. Foong YC, Chherawala N, Aitken D, Scott D, Winzenberg T, Jones G. Accelerometer-determined physical activity, muscle mass, and leg strength in community-dwelling older adults. J Cachexia Sarcopenia Muscle 2016 Jun;7(3):275-283 [FREE Full text] [doi: 10.1002/jcsm.12065] [Medline: 27239404]
66. Saxon SV, Etten MJ, Perkins EA. Physical Change and Aging: A Guide for Helping Professions. New York: Springer Publishing; 2021.
67. Laurent M, Dedeyne L, Dupont J, Mellaerts B, Dejaeger M, Gielen E. Age-related bone loss and sarcopenia in men. Maturitas 2019 Apr;122:51-56 [FREE Full text] [doi: 10.1016/j.maturitas.2019.01.006] [Medline: 30797530]
68. Madeira T, Severo M, Correia D, Lopes C, Gorjão Clara J. Nutritional intake and malnutrition in institutionalised and non-institutionalised older adults. Br J Nutr 2022 Sep 14;128(5):921-931. [doi: 10.1017/S0007114521003925] [Medline: 34583786]
69. Ozer NT, Akin S, Gunes Sahin G, Sahin S. Prevalence of malnutrition diagnosed by the Global Leadership Initiative on Malnutrition and Mini Nutritional Assessment in older adult outpatients and comparison between the Global Leadership Initiative on Malnutrition and Mini Nutritional Assessment energy-protein intake: a cross-sectional study. JPEN J Parenter Enteral Nutr 2022 Feb;46(2):367-377. [doi: 10.1002/jpen.2123] [Medline: 33893657]

XSL•FO
RenderX

70. Abugroun A, Nayyar A, Abdel-Rahman M, Patel P. Impact of malnutrition on hospitalization outcomes for older adults admitted for sepsis. Am J Med 2021 Feb;134(2):221-226.e1 [FREE Full text] [doi: 10.1016/j.amjmed.2020.06.044] [Medline: 32810466]
71. Chen C, Schilling L, Lyder C. A concept analysis of malnutrition in the elderly. J Adv Nurs 2001 Oct;36(1):131-142 [FREE Full text] [doi: 10.1046/j.1365-2648.2001.01950.x] [Medline: 11555057]
72. van den Brink WJ, van den Broek TJ, Palmisano S, Wopereis S, de Hoogh IM. Digital biomarkers for personalized nutrition: predicting meal moments and interstitial glucose with non-invasive, wearable technologies. Nutrients 2022 Oct 24;14(21):273-282 [FREE Full text] [doi: 10.3390/nu14214465] [Medline: 36364728]
73. Suh S, Kim JH. Glycemic variability: how do we measure it and why is it important? Diabetes Metab J 2015 Aug;39(4):273-282 [FREE Full text] [doi: 10.4093/dmj.2015.39.4.273] [Medline: 26301188]
74. Zhong H, Zhang K, Lin L, Yan Y, Shen L, Chen H, et al. Two-week continuous glucose monitoring-derived metrics and degree of hepatic steatosis: a cross-sectional study among Chinese middle-aged and elderly participants. Cardiovasc Diabetol 2024 Aug 31;23(1):322 [FREE Full text] [doi: 10.1186/s12933-024-02409-6] [Medline: 39217368]
75. Lu J, Wang W, Goh J, Maier A. A practical guide for selecting continuous monitoring wearable devices for community-dwelling adults. Heliyon 2024 Jul 15;10(13):e33488 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e33488] [Medline: 39035501]
76. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. JMIR Aging 2023 Mar 21;6:e43197 [FREE Full text] [doi: 10.2196/43197] [Medline: 36943333]
77. Hulleck A, Menoth Mohan D, Abdallah N, El Rich M, Khalaf K. Present and future of gait assessment in clinical practice: towards the application of novel trends and technologies. Front Med Technol 2022;4:901331 [FREE Full text] [doi: 10.3389/fmedt.2022.901331] [Medline: 36590154]
78. Jayakumar P, Moore MG, Furlough KA, Uhler LM, Andrawis JP, Koenig KM, et al. Comparison of an artificial intelligence–enabled patient decision aid vs educational material on decision quality, shared decision-making, patient experience, and functional outcomes in adults with knee osteoarthritis: a randomized clinical trial. JAMA Netw Open 2021 Feb 01;4(2):e2037107 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.37107] [Medline: 33599773]

Abbreviations

AoA: anorexia of aging
a.u.: arbitrary units
CGM: continuous glucose monitoring
DHT: digital health technology
FFS: Fried Frailty Score
GAS: Geriatric Anorexia Study
HSD: honestly significant difference
ICC: intraclass correlation coefficient
LoA: limits of agreement
M6min: maximum 6 minutes of activity
M60min: maximum 60 minutes of activity
MNA-SF: Mini Nutrition Assessment – Short Form
MVPA: moderate-to-vigorous physical activity
PRO: patient-reported outcome
PROMIS: Patient-Reported Outcomes Measurement Information System
SLPA: sedentary-to-light physical activity
SNAQ: Simplified Nutritional Appetite Questionnaire
SVMg: signal vector magnitude
TAR: time above range
TBR: time below range
TIR: time in range

XSL•FO
RenderX

Edited by J Wang; submitted 23.May.2025; peer-reviewed by Z Zandesh, A Billis; comments to author 05.Nov.2025; accepted 16.Mar.2026; published 17.Apr.2026.

Please cite as:
Camerlingo N, Wicker M, Psaltos DJ, Shaafi-Kabiri N, Zhang H, Messere A, Messina I, Hayden Z, Kelly M, Kamat G, Karahanoğlu FI, Demanuele C, Santamaria M, Caouette D, Thomas KC
Leveraging Digital Health Technologies to Assess Older Adults' Frailty and Nutritional Status: Two Cross-Sectional Studies
JMIR Aging 2026;9:e77816
URL: https://aging.jmir.org/2026/1/e77816
doi:10.2196/77816
PMID:

©Nunzio Camerlingo, Madisen Wicker, Dimitrios J Psaltos, Nina Shaafi-Kabiri, Hao Zhang, Andrew Messere, Isabelle Messina, Zachary Hayden, Meredith Kelly, Gauri Kamat, Fikret Işık Karahanoğlu, Charmaine Demanuele, Mar Santamaria, David Caouette, Kevin C Thomas. Originally published in JMIR Aging (https://aging.jmir.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Knowledge Performance, User Experience, and Cybersickness With Immersive Virtual Reality for Tai Chi Movement-Encoding in Older Adults: Randomized Comparison Study

XiaCheng Song[1], PhD; Lu Sun[2], PhD; Yaqing Song[3], PhD; Huafeng Qu[4], PhD; Jing Jin[5], PhD; Junfeng Zhu[6], PhD; Wenwen Song[7], PhD; Huirong Huang[8], PhD

[1]Guangdong University of Technology, Guangzhou, Guangdong, China
[2]Xi'an Fanyi University, Taiyigong, Chang'an District, Xi'an, Shaanxi, China
[3]Shaanxi Technical College of Finance & Economics, Xi'an, Shaanxi, China
[4]Yunnan College of Business Management, Kunming, Yunnan, China
[5]Guangdong University of Science and Technology, Dongguan, Guangdong, China
[6]Zhaoqing University, Zhaoqing, Guangdong, China
[7]Xidian University, Xi'an, Shaanxi, China
[8]Guangdong Industry Polytechnic, Guangzhou, Guangdong, China

Corresponding Author:
Lu Sun, PhD
Xi'an Fanyi University, Taiyigong, Chang'an District, Xi'an, Shaanxi, China

Abstract

Background: Immersive virtual reality (VR) is increasingly used for embodied training, but its practical value for older adults depends on whether the experiential benefits can be achieved without unacceptable discomfort or overstated claims about motor learning.

Objective: This randomized study compared head-mounted immersive VR with standardized, content-matched 2D video instruction for Tai Chi movement-encoding in community-dwelling older adults.

Methods: A total of 106 older adults were randomized to immersive VR (n=56) or video instruction (n=50). The primary analysis included 92 participants with complete age, pretest, and posttest data (VR: n=45; Video: n=47). The primary outcome was a 5-item movement-encoding knowledge score (range 0‐20), analyzed using analysis of covariance, with posttest score as the dependent variable and group, pretest score, and age as predictors. Secondary outcomes were intrinsic motivation, game experience, and learning engagement; safety, tolerability, and feasibility were assessed using the Intrinsic Motivation Inventory, Game Experience Questionnaire, Utrecht Work Engagement Scale–Short Form, participant flow, and Virtual Reality Sickness Questionnaire. VR system telemetry was analyzed descriptively as an exploratory process outcome.

Results: Adjusted posttest movement-encoding performance did not differ between VR and video instruction (adjusted difference −0.27, 95% CI −2.25 to 1.72; P=.79; partial η^2=0.001). The Game Experience Questionnaire results favored VR for competence (Video: mean 3.36, SD 1.21; VR: mean 4.15, SD 0.92; q=0.009), flow (Video: mean 3.77, SD 1.26; VR: mean 4.46, SD 0.76; q=0.014), and lower tension (Video: mean 1.76, SD 1.04; VR: mean 1.27, SD 0.63; q=0.037), whereas challenge did not remain significant after false discovery rate correction (q=0.073). In the VR Group, recognition accuracy was associated with posttest movement-encoding score (n=44; r=0.317, P=.04), but pose accuracy was not (n=44; r=0.129, P=.41). The Virtual Reality Sickness Questionnaire scores were low among valid VR participants (total score mean 3.77, SD 11.12), and 36 out of 46 (78%) participants reported zero symptoms.

Conclusions: With a strong active control, immersive VR did not produce higher short-term Tai Chi movement-encoding performance than video instruction. Among participants who completed the VR session, the immediate game-experience profile was more favorable. These findings indicate that VR may offer experiential benefits, while also highlighting the need to manage cybersickness and technical exclusions and to distinguish movement knowledge from actual motor execution.

Trial Registration: International Traditional Medicine Clinical Trial Registry ITMCTR2026001760; https://tinyurl.com/3vvunsjm

(JMIR Aging 2026;9:e96867) doi:10.2196/96867

XSL•FO
RenderX

KEYWORDS

human factors; ergonomics; immersive virtual reality; older adults; movement encoding; cybersickness; VRSQ; Virtual Reality Sickness Questionnaire

Introduction

Rationale for the Study

As societies age, maintaining functional independence has become a central public health priority, particularly given the heightened consequences of balance failure and mobility decline in later life [1-3]. Tai Chi is frequently recommended for older adults because it combines controlled postural transitions with sustained attention to movement elements and goal-directed practice that can be adapted to heterogeneous functional levels [4-6]. Unlike purely verbal or screen-based cognitive tasks, Tai Chi learning requires embodied visuomotor mapping and multisensory integration, making it a suitable test bed for studying how immersive virtual reality (VR) shapes action-related memory encoding in later life [7-9]. Yet, the real-world dissemination of Tai Chi in community contexts faces a persistent accessibility paradox: the individuals who could benefit most from structured practice often have the least reliable access to safe, repeatable instruction [10-12]. Shortages of qualified instructors, geographic and scheduling constraints, and the practical challenges of maintaining consistent supervision can all limit participation [11,13,14]. Additionally, complex movement learning may carry a psychological barrier for novice learners, as uncertainty about "doing it correctly" can discourage sustained engagement even when motivation to improve health is high [11,13,15].

Immersive VR has been proposed as a scalable approach to deliver standardized movement learning experiences under controlled conditions, offering a headset-based practice space that can reduce external distraction and support repeated exposure to the same instructional sequence [10,16,17]. However, the evidence base for VR and related immersive training and learning, especially for de novo acquisition of complex movement sequences, remains fragmented and often inconsistent when compared with strong active controls [18-20]. Recent randomized studies and reviews likewise suggest that outcomes depend on training dose, adaptive design, comparator strength, and the distinction between device-recorded performance and transferable motor learning [21-24]. A key methodological limitation is that many studies rely on passive or no-treatment controls, making it difficult to isolate the effect of immersion from instructional content and practice structure. This concern echoes longstanding debates in media-comparison research: learning gains are often driven by instructional design and practice opportunities rather than the delivery medium per se [25,26]. For older adults, these issues are further complicated by immersion-specific feasibility and safety constraints, including VR discomfort and symptoms that may affect adherence and threaten generalizability if not transparently measured and reported [24,27-30].

Beyond these methodological challenges, 2 conceptual gaps remain underaddressed. First, immersive presentation is often implicitly treated as equivalent to improved execution [26,31], yet complex skill learning can be constrained by a cognitive-motor translation bottleneck: forming a representation of "what to do" does not necessarily ensure accurate embodied reproduction of "how to do it" [32]. Movement-encoding knowledge refers to the learner's immediate cognitive representation, recall, and recognition of essential movement elements and sequence-relevant cues. Such encoding is a necessary early step in learning a novel movement sequence and can be assessed comparably across VR and video conditions without treating device-specific tracking metrics as equivalent to motor skill. This distinction is especially salient for Tai Chi, where successful learning requires not only recognizing key movement elements but also coordinating posture, timing, and spatial control. Second, the mechanisms through which VR may confer value are frequently reduced to broad notions of enjoyment or novelty [33,34], while state-level conditions that shape practice quality, such as perceived competence, pressure or tension, and affective experience during learning [15], are less systematically integrated into the evidence chain. We refer to these motivational and experiential states collectively as human factors outcomes. The enclosed nature of a head-mounted display may plausibly create a more private practice context by limiting external social cues, potentially reducing pressure and negative affect while supporting engagement [17,31,35]; however, these claims require disciplined testing alongside learning outcomes, rather than being inferred from immersion alone. At the same time, feasibility depends on clearly defined safety and tolerability boundaries because immersion-related symptoms can emerge even in well-controlled settings [29].

Research Questions and Hypotheses

Guided by the identified gaps in older-adult immersive motor learning, we examined whether the more enclosed practice context created by a head-mounted display was associated with lower pressure or tension and more favorable competence-related and flow-related states. We posed the following research questions (RQs):

1. RQ1 (movement-encoding performance): When the instructional content is matched, does immersive VR produce higher posttest Tai Chi movement-encoding knowledge performance than standardized 2D video instruction after controlling for pretest score and age?
2. RQ2 (human factors outcomes): Does immersive VR systematically reshape older adults' motivational and experiential states during training (eg, interest or enjoyment, perceived competence, pressure or tension, immersion, flow, and engagement) relative to standardized video instruction?
3. RQ3 (safety, tolerability, and feasibility): In real-world community deployment, what safety, tolerability, and feasibility profile does immersive VR show, including VR discomfort or symptoms, discontinuation or exclusion due to discomfort or technical failure, and the proportion of participants reporting zero Virtual Reality Sickness Questionnaire (VRSQ) symptoms?

XSL•FO
RenderX

Based on these questions, we specified the following hypotheses:

1. H1 (movement-encoding performance): Compared with the Video Group, the VR Group will show higher adjusted posttest movement-encoding knowledge scores when controlling for pretest performance and age.
2. H2 (human factors benefit): Compared with the Video Group, the VR Group will report a more favorable motivational and experiential profile, reflected in higher perceived competence and flow or immersion and lower pressure or tension.

Because safety, tolerability, and feasibility were evaluated descriptively, RQ3 was exploratory, and no directional hypothesis was specified.

In addition, for the VR Group, system logs (recognition accuracy, pose accuracy, and movement distance) were used to characterize performance descriptively during task execution, providing process-level signals to complement self-reports and learning outcomes.

Methods

Participants

A total of 106 community-dwelling older adults were recruited and enrolled between November and December 2025. Eligibility criteria were age ≥60 years, a native Chinese language background, and no self-reported major physical or mental disorders that could materially affect task participation or outcome assessment. The sample size was determined based on feasibility considerations and prior community-based studies with comparable designs. Participants were randomly assigned to either the VR Group (n=56) or the Video Group (n=50). Randomization was not stratified by age or prior Tai Chi experience; the observed age imbalance was, therefore, handled analytically by including age as a covariate and by conducting an exploratory Group × Age sensitivity analysis. Following quality control, 95 participants were retained as valid baseline cases for demographic and background comparisons (VR: n=46; Video: n=49).

Experimental Design

We used a 2-arm pretest/posttest control-group design. The independent variable was learning modality (immersive VR learning vs 2D video-based learning). Sessions were conducted in quiet, spacious indoor settings coordinated by local community and village activity rooms. Before each session, the research team confirmed that participants had sufficient, unobstructed space for safe movement, and on-site supervision was provided throughout.

Randomization and Allocation

The study leader generated an approximately 1:1 computer-based randomization sequence before study initiation. Allocation was implemented after eligibility screening, and recruitment and on-site coordination staff followed the prespecified allocation list for scheduling and documentation. The sequence was not stratified by age or prior Tai Chi experience. Participants completed the learning tasks under their assigned condition unless they withdrew or were excluded based on predefined invalid-sample criteria (eg, VR-related discomfort, technical failure, no usable data, or incomplete participation). Study-level exclusions were applied according to feasibility and data-quality criteria before the final analytic datasets were constructed. Participant flow, invalid cases, and exclusions from the primary analysis are reported in the *Results* section.

Apparatus and Materials

The VR intervention was delivered using Meta Quest 3 and Meta Quest 3S head-mounted displays (Meta Platforms, Inc). Both models use the Snapdragon XR2 Gen 2 platform (Qualcomm Technologies, Inc) and provide comparable 6 df inside-out tracking with touch plus controllers, ensuring consistent tracking fidelity and interaction across devices. The Video Group learned the same content via standard 2D instructional videos presented on a conventional display.

Procedure

All participants completed a pretest, the assigned learning session, and a posttest within the same study workflow. The learning content comprised 5 core Tai Chi forms delivered in a standardized order and format across conditions. After the learning session, participants completed postintervention questionnaires assessing intrinsic motivation, game-related experience, and engagement. For VR participants, the system additionally recorded behavioral performance indices during task execution (eg, recognition accuracy, pose accuracy, and movement distance) for descriptive reporting.

Training sessions were supervised on-site. Participants were instructed to stop immediately if they experienced dizziness, nausea, discomfort, or instability. Short breaks were allowed upon request, and the session was terminated if symptoms persisted. Any adverse events (eg, falls, near-falls, musculoskeletal pain, or persistent cybersickness symptoms) and reasons for discontinuation were documented by the staff. These procedures were implemented to minimize physical risks during standing practice among older adults and to ensure that cybersickness and balance-related safety were treated as part of real-world feasibility.

Primary Outcome: Movement-Encoding Knowledge

The primary outcome was short-term movement-encoding knowledge, defined as immediate encoding, recall, and recognition of essential movement elements and sequence-relevant cues. It was assessed using a 5-item scoring rubric developed for the present study and guided by motor-learning frameworks [8,32]. The measure was administered before and after the learning session. It should be interpreted as a cognitive movement-representation outcome, not as a direct assessment of full Tai Chi motor execution, movement quality, or long-term skill acquisition [8,32]. Each item corresponded to one Tai Chi form and was scored from 0 to 4 using a predefined rubric (total range 0 - 20; higher scores indicate more complete encoding and recall of movement elements). Total scores were computed only when all 5 items contained valid responses; otherwise, the total score was treated as missing for the primary analysis [36]. The 5-item scoring components are provided in Multimedia Appendix 1. Diagnostic reliability was acceptable for this rubric-based score (Cronbach

XSL•FO
RenderX

α: pretest=0.910, posttest=0.725; McDonald's omega total: pretest=0.914, posttest=0.743), although these coefficients should be interpreted cautiously because the measure was designed as a scoring rubric rather than a fully validated psychometric scale.

Secondary Outcome: Intrinsic Motivation

Intrinsic motivation refers to participants' immediate interest or enjoyment, perceived competence, and pressure or tension during the learning session. It was measured with 9 items from the Intrinsic Motivation Inventory (IMI) using a 1 - 7 response scale: interest/enjoyment (5 items, including 1 reverse-scored item), perceived competence (2 items), and pressure/tension (2 items, including 1 reverse-scored item). Subscale scores were calculated as item means; higher scores indicate more of the named construct. Prior confirmatory work supports the IMI subscales' factorial validity and internal consistency [37].

Secondary Outcome: Game Experience

Game experience refers to participants' immediate cognitive and affective experience of the learning environment. It was measured with the 14-item Game Experience Questionnaire (GEQ) core module using a 1 to 5 response scale. Seven 2-item subscales were calculated as the item means: competence, immersion, flow, tension, challenge, negative affect, and positive affect; higher scores indicate more of the named experience. Item allocation and scoring followed the GEQ manual [38]. Independent validation has raised concerns about the stability of the original 7-factor structure [39]; in the current sample, Cronbach α values for the seven 2-item subscales ranged from 0.462 to 0.728. GEQ subscale findings were therefore interpreted cautiously and as immediate experience indicators rather than definitive latent-trait measures.

Secondary Outcome: Learning Engagement

Learning engagement refers to the vigor, dedication, and absorption that participants reported during the learning activity. It was measured with the 9-item Utrecht Work Engagement Scale–Short Form (UWES-S) using a 1 to 7 response scale. Vigor, dedication, and absorption scores were calculated as the means of 3 items each, and total engagement as the mean of all 9 items; higher scores indicate stronger engagement. Cross-national validation supports the scale's 3-factor structure and internal consistency [40].

Safety, Tolerability, and Feasibility

Safety and tolerability refer to VR-related symptoms and the ability to complete the session without clinically relevant discomfort or instability. VR symptoms were measured in the VR Group with the 9-item VRSQ (item scores of 0-3). Oculomotor and disorientation scores were normalized to 0 to 100 and averaged to obtain a total score; higher scores indicate more severe symptoms. We also reported the absolute number and proportion with zero symptoms. The VRSQ was developed with evidence of internal consistency and convergent validity for VR-specific symptoms [41]. Feasibility was defined descriptively through participant flow and the number and reasons for discontinuation or exclusion, including discomfort, technical failure/no usable data, and incomplete participation. It was not treated as a separate psychometric construct.

Exploratory VR Behavioral Performance

VR behavioral performance refers to device-recorded process indicators available only in the VR Group. Recognition accuracy was the percentage of correct responses across 5 recognition items; pose accuracy was the percentage of correctly hit pose targets (15 total); and motion distance was the summed movement distance of the head and both hands in meters. These telemetry variables were summarized descriptively and explored using correlations; they were not designated primary or secondary learning outcomes.

Data Quality Control and Analytic Samples

Participant flow and exclusions followed a 2-layer rule consistent with CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth; Checklist 1) reporting guidance [42]. First, study-level invalid cases were removed based on predefined criteria (VR: n=10; Video: n=1), including motion sickness/discomfort, technical failure/no usable data, and incomplete participation/withdrawal, resulting in 95 valid baseline participants. Second, for the primary learning-outcome analysis, participants were included only if age, pretest score, and posttest score were all available, yielding the primary analytic sample of 92 (VR: n=45; Video: n=47). Missingness was limited: in the valid baseline sample, the Video Group had 2 out of 49 participants missing posttest scores, whereas the VR Group had 1 out of 46 missing pretest data and 1 out of 46 missing posttest data. All analyses used complete-case data for the variables involved.

Statistical Analysis

All analyses were conducted in R software (v4.6.0; R Foundation for Statistical Computing). Descriptive statistics are reported as mean (SD) for continuous variables and n (%) for categorical variables. Baseline group differences were examined using independent-samples 2-tailed Welch *t* tests (continuous variables) and Fisher exact tests (categorical variables). The primary intervention effect on posttest performance was evaluated using analysis of covariance (ANCOVA) with posttest score as the dependent variable, group as the fixed factor, and pretest score and age as covariates; adjusted group differences are reported with 95% CIs and partial η^2 [43,44]. For user-experience outcomes, group differences were examined using independent-samples 2-tailed Welch *t* tests with false discovery rate (FDR) adjustment applied across self-report outcomes [45,46]; Cohen *d* values were computed as VR minus Video. Given the baseline age imbalance, we additionally conducted age-adjusted linear models for key self-report outcomes and an exploratory Group × Age model for the primary outcome. Missing data were handled using complete-case analysis for the variables involved; a conservative sensitivity check implemented under the pretest-complete rule did not alter the primary conclusion. VR telemetry correlations with movement-encoding scores and selected self-report measures were examined descriptively using Pearson correlations.

XSL•FO
RenderX

Ethical Considerations

The study protocol was reviewed and approved by the Research Office of Xi'an Translation University on October 16, 2025. The reviewing body did not issue a case or application number. All participants provided written informed consent before participation. Participants were informed of the study procedures and potential risks, including possible VR-related discomfort, such as dizziness or eye strain, and could rest, discontinue a session, or withdraw at any time without penalty. Study data were deidentified for analysis and reporting. Participants received a small commemorative gift valued at approximately RMB $40 to $50 (RMB $1=US $0.14 as of December 8, 2025) for each study visit.

Trial Registration

This report presents first-session findings from a larger pilot randomized trial that was retrospectively registered with the International Traditional Medicine Clinical Trial Registry (registration number ITMCTR2026001760). The registry application was first submitted on April 1, 2026, after participant recruitment had commenced. The registry record includes additional follow-up assessments that are outside the scope of the present report.

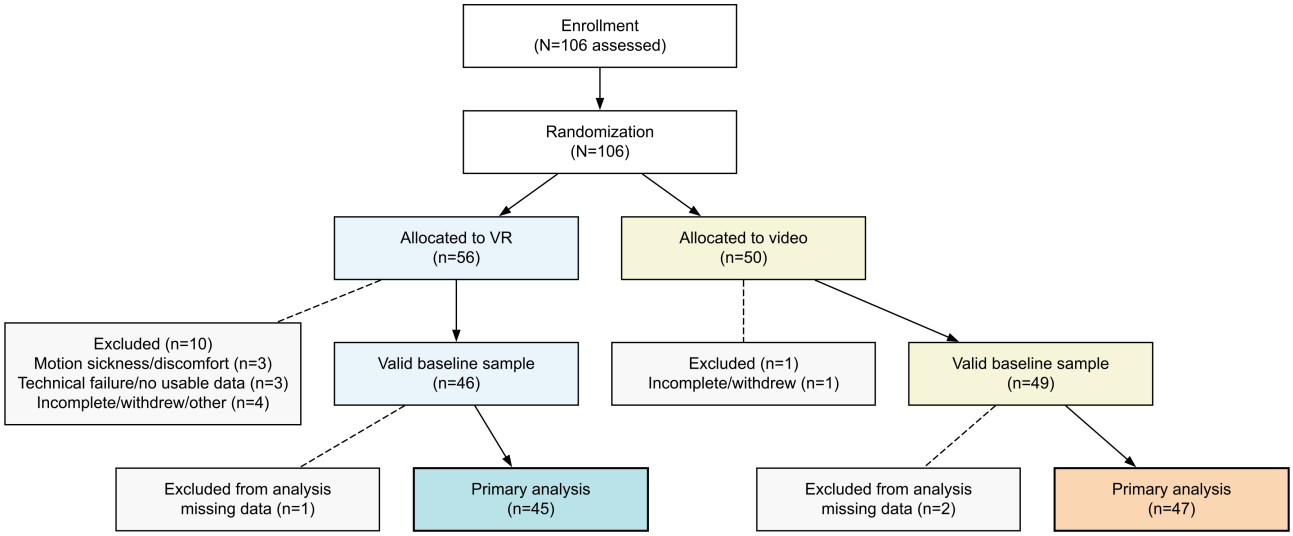
Results

Participant Flow and Baseline Characteristics

As shown in Figure 1, a total of 106 older adults were assessed and randomized (VR: n=56; Video: n=50). Based on predefined invalid-sample criteria, 10 participants in the VR arm were excluded from the study (motion sickness/discomfort, n=3; technical failure/no usable data, n=3; incomplete, withdrew, or other, n=4), and 1 participant in the video arm was excluded (incomplete/withdrew, n=1), yielding 95 valid participants for baseline characterization (VR: n=46; Video: n=49), as shown in Table 1 (the 2 groups did not differ significantly in sex, education, or prior Tai Chi experience, but the Video Group was older on average). For the primary movement-encoding analysis, an additional 1 VR participant and 2 Video participants were excluded under the complete-case rule for age, pretest score, and posttest score, resulting in a final analytic sample of 92 participants (VR: n=45; Video: n=47), which was used for the ANCOVA reported in Table 2.

As shown in Table 1, baseline characteristics were summarized for the valid sample (N=95, Video Group n=49, VR Group n=46). Participants were older adults (overall mean age of 68.67, SD 7.55 y). The Video Group was older (mean 70.69, SD 7.43) than the VR Group (mean 66.52, SD 7.15 y), and this difference was statistically significant (*P*=.006). Sex distribution did not differ between groups (*P*=.16; overall: male 24/95, 25% and female 71/95, 75%). Educational attainment was also comparable across groups (*P*=.41), with the largest categories being primary school or below (31/95, 33%) and junior high school (28/95, 29%). Prior Tai Chi experience was similarly balanced (*P*=.77), with most participants reporting no prior exposure (72/95, 76%).

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram for the randomized comparison of immersive virtual reality (VR) and video-based Tai Chi movement-encoding instruction among community-dwelling older adults, November to December 2025. A total of 106 participants were assessed and randomized (VR Group: n=56; Video Group: n=50). Study-level exclusions were based on prespecified invalid-sample criteria (VR: n=10; Video: n=1), including motion sickness or discomfort (VR: n=3), technical failure or no usable data (VR: n=3), and incomplete participation, withdrawal or other (VR: n=4; Video: n=1). Valid participants contributing to baseline characteristics are reported in Table 1 (VR: n=46; Video: n=49). Primary outcome analyses used the complete-case primary sample (Table 2; VR: n=45; Video: n=47).



XSL•FO
RenderX

Table . Baseline characteristics of community-dwelling older adults randomized to immersive VR[a] or video-based Tai Chi movement-encoding instruction, November to December 2025.[b]

Characteristic	Overall (N=95)	Video Group (n=49)	VR Group (n=46)	*P* value
Age (y), mean (SD)	68.67 (7.55)	70.69 (7.43)	66.52 (7.15)	.006
Sex, n (%)				.16
Male	24 (25)	9 (18)	15 (33)	
Female	71 (75)	40 (82)	31 (67)	
Education, n (%)				.41
Bachelor's degree or above	4 (4)	2 (4)	2 (4)	
Middle school	28 (29)	13 (27)	15 (33)	
Junior college	6 (6)	1 (2.0)	5 (11)	
High school or vocational school	26 (27)	15 (31)	11 (24)	
Primary school or below	31 (33)	18 (37)	13 (28)	
Prior Tai Chi experience, n (%)				.77
Never	72 (76)	37 (76)	35 (76)	
Practice often	5 (5)	2 (4)	3 (7)	
Occasionally watched	9 (10)	6 (12)	3 (7)	
Practiced a little	9 (10)	4 (8)	5 (11)	

[a]VR: virtual reality.

[b]Values are mean (SD) for continuous variables and n (%) for categorical variables. Group differences were tested using independent-samples 2-tailed Welch *t* tests for age and Fisher's exact tests for categorical variables. Missing data were not included in the summaries.

Table . Pretest-posttest Tai Chi movement-encoding knowledge scores and adjusted intervention effect in community-dwelling older adults randomized to immersive VR[a] or video instruction, November to December 2025.[b]

Outcome	Video Group (n=47), mean (SD)	VR Group (n=45), mean (SD)
Pretest score (range 0‐20)	2.70 (5.09)	2.80 (4.98)
Posttest score (range 0‐20)	8.81 (4.97)	8.71 (4.49)
Movement-encoding gain	6.11 (5.88)	5.91 (6.19)

[a]VR: virtual reality.

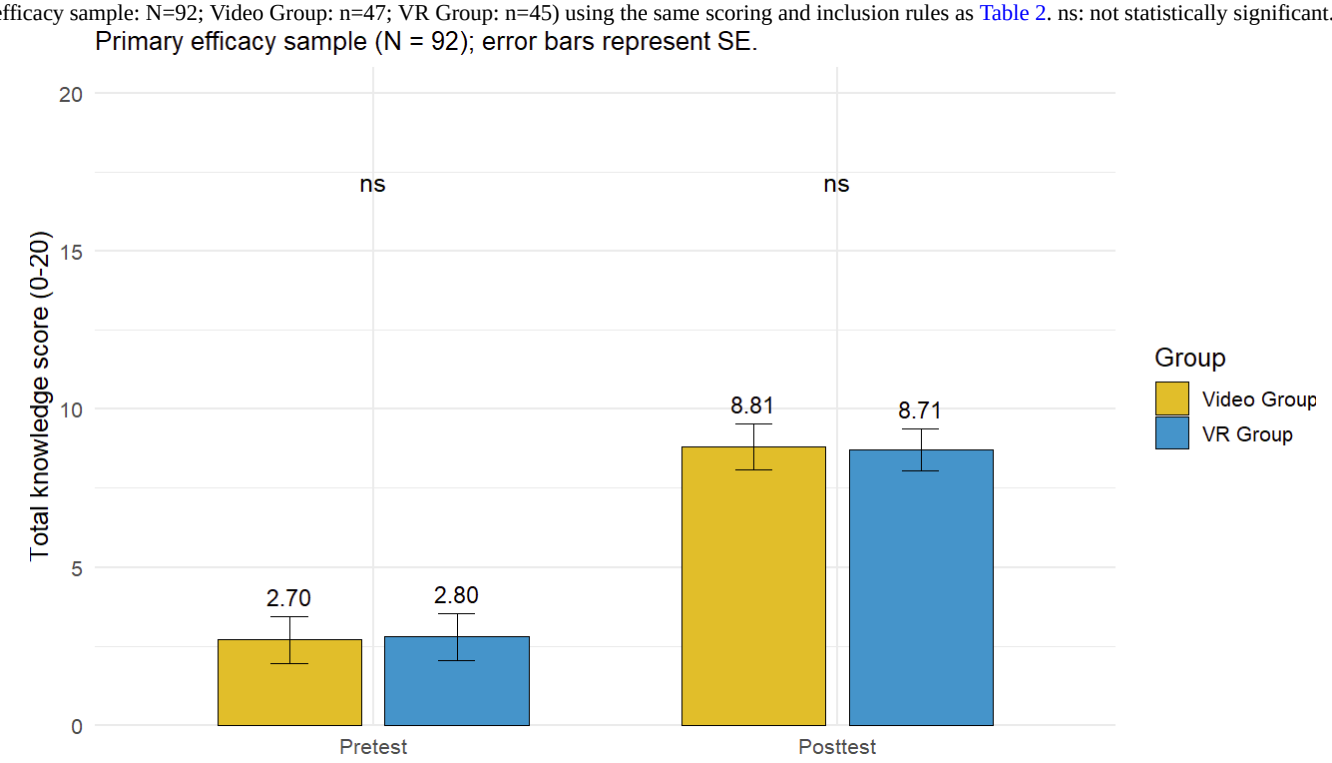
[b]The adjusted mean difference between the VR and Video Groups was −0.269 (95% CI −2.25 to 1.72; *P*=.79), based on an analysis of covariance model with posttest score as the dependent variable, group as the fixed factor, and pretest score and age as covariates (primary analysis sample, N=92). Participants were included in the primary analysis only if age and all 5 pretest and posttest items were complete.

Primary Outcome: Movement-Encoding Knowledge

As shown in Figure 2, total movement-encoding knowledge scores (0‐20) increased from pretest to posttest in both groups within the primary analysis sample (N=92), consistent with the descriptive results summarized in Table 2. The plotted means (SE) indicate comparable patterns of improvement across the Video Group (n=47) and the VR Group (n=45) under the same scoring rule and sample definition used for the main inferential model.

XSL•FO
RenderX

Figure 2. Pretest and posttest Tai Chi movement-encoding knowledge scores among community-dwelling older adults randomized to immersive virtual reality (VR) or video instruction, November to December 2025. Scores represent the sum of 5 movement-encoding knowledge items (each scored 0 to 4), yielding a total score range of 0 to 20. Bars show the mean; error bars represent the SE. Results are based on the primary analysis sample (primary efficacy sample: N=92; Video Group: n=47; VR Group: n=45) using the same scoring and inclusion rules as Table 2. ns: not statistically significant.



As shown in Table 2, the primary analysis sample included 92 participants (Video Group n=47, VR Group n=45) with complete pretest, posttest, and age data. Descriptively, pretest movement-encoding scores were comparable between groups (Video: mean 2.70, SD 5.09; VR: mean 2.80, SD 4.98), as were posttest scores (Video: mean 8.81, SD 4.97; VR: mean 8.71, SD 4.49) and movement-encoding gain (Video: mean 6.11, SD 5.88; VR: mean 5.91, SD 6.19). In the ANCOVA model controlling for pretest score and age, the adjusted intervention effect (VR vs Video) on posttest performance was not significant (adjusted difference −0.269, 95% CI –2.25 to 1.72; P=.79; partial η^2=0.001).

Secondary Outcome: Intrinsic Motivation

As shown in Table 3, intrinsic motivation outcomes were compared between the Video Group (n=47) and the VR Group (n=46) across 3 subscales. Mean scores were similar for interest or enjoyment (Video: mean 5.23, SD 1.33; VR: mean 5.53, SD 1.10), perceived competence (Video: mean 5.47, SD 1.87; VR: mean 5.86, SD 1.27), and pressure/tension (Video: mean 2.27, SD 1.36; VR: mean 2.39, SD 1.65). None of the between-group differences reached statistical significance, and all findings remained nonsignificant after FDR correction (q=0.376, q=0.376, and q=0.879, respectively).

Table . Intrinsic motivation after a single Tai Chi movement-encoding session in community-dwelling older adults randomized to immersive VR[a] or video instruction, November to December 2025.[b]

IMI[c] subscales (range 1-7)	Video Group (n=47), mean (SD)	VR Group (n=46), mean (SD)	*P* value	*q* value
Interest or enjoyment	5.23 (1.33)	5.53 (1.10)	.24	0.376
Perceived competence	5.47 (1.87)	5.86 (1.27)	.24	0.376
Pressure or tension	2.27 (1.36)	2.39 (1.65)	.69	0.879

[a]VR: virtual reality.

[b]Between-group differences were tested using independent-samples 2-tailed Welch *t* tests. False discovery rate adjustment was applied across the self-report outcome set to obtain *q* values.

[c]IMI: Intrinsic Motivation Inventory.

Secondary Outcome: Game Experience

As shown in Figure 3, GEQ ratings showed the clearest experiential separation between groups. The figure displays the GEQ dimensions most relevant to the revised interpretation: competence, flow, tension, and challenge. FDR-adjusted *q* values are shown for visualization. VR participants reported higher competence and flow and lower tension after FDR correction, whereas the lower challenge rating in the VR Group was significant before correction but did not remain significant after FDR adjustment.

XSL•FO
RenderX

Figure 3. Game experience after a single Tai Chi movement-encoding session in community-dwelling older adults randomized to immersive virtual reality (VR) or video instruction, November to December 2025. Ratings are shown for Game Experience Questionnaire (GEQ) competence, flow, tension, and challenge. Box plots display the distribution of participant responses, with jittered points indicating individual observations. False discovery rate–adjusted *q* values are shown for the plotted GEQ subscales; inferential interpretation should follow the *q* values reported in Table 4.

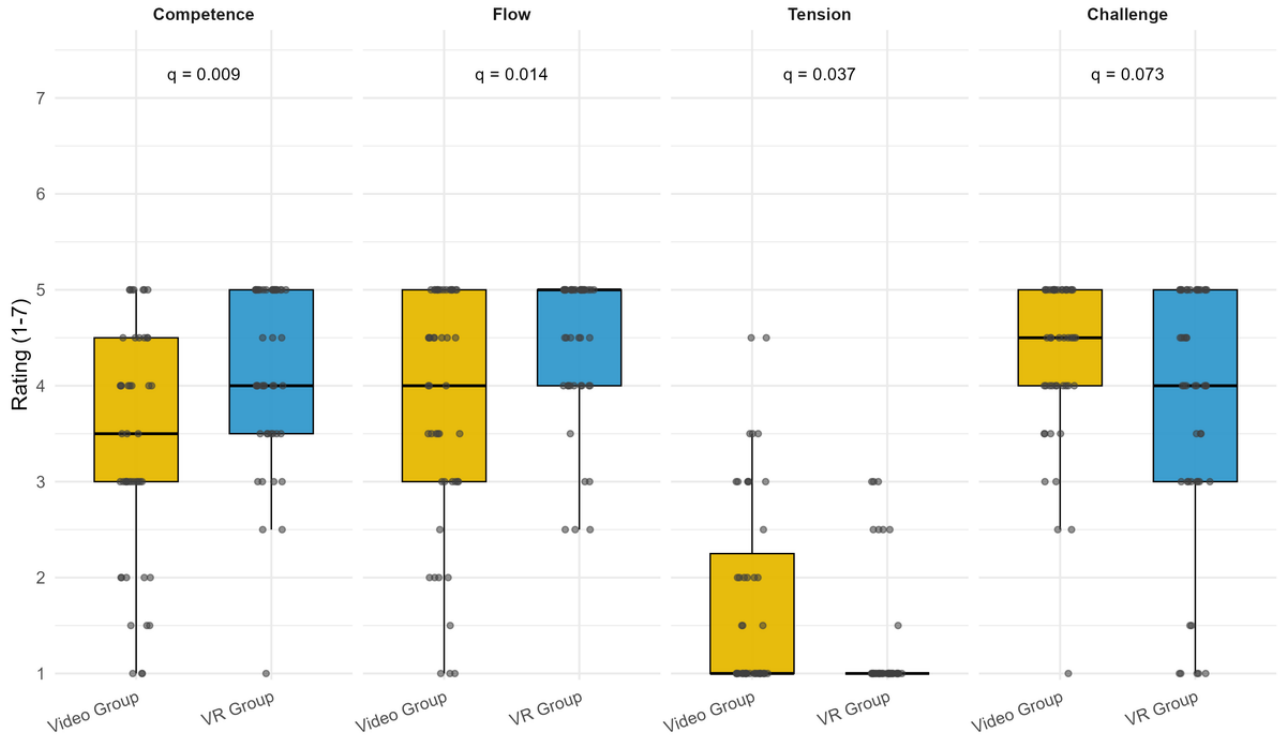


Table . Game experience after a single Tai Chi movement-encoding session in community-dwelling older adults randomized to immersive VR[a] or video instruction, November to December 2025.[b]

GEQ[c] dimensions (range 1-5)	Video Group (n=47), mean (SD)	VR Group (n=46), mean (SD)	*P* value	*q* value
Competence	3.36 (1.21)	4.15 (0.92)	<.001	0.009
Immersion	4.10 (0.95)	4.37 (0.77)	.13	0.302
Flow	3.77 (1.26)	4.46 (0.76)	.002	0.014
Tension	1.76 (1.04)	1.27 (0.63)	.008	0.037
Challenge	4.29 (0.86)	3.74 (1.33)	.02	0.073
Negative affect	1.83 (1.04)	1.83 (1.08)	.99	0.987
Positive affect	4.24 (0.87)	4.49 (0.77)	.16	0.310

[a]VR: virtual reality.

[b]Between-group differences were tested using independent-samples 2-tailed Welch *t* tests. False discovery rate adjustment was applied across the self-report outcome set to obtain *q* values.

[c]GEQ: Game Experience Questionnaire.

As shown in Table 4, game experience (GEQ; 1 - 5) differed on several dimensions between the Video Group (n=47) and the VR Group (n=46). The VR Group reported higher competence (Video: mean 3.36, SD 1.21; VR: mean 4.15, SD 0.92; *P*<.001, *q*=0.009; Cohen *d*=0.733) and higher flow (Video: mean 3.77, SD 1.26; VR: mean 4.46, SD 0.76; *P*=.002, *q*=0.014; Cohen *d*=0.661) after FDR correction. The VR Group also reported lower tension (Video: mean 1.76, SD 1.04; VR: mean 1.27, SD 0.63; *P*=.008, *q*=0.037; Cohen *d*=−0.562). Challenge was lower in the VR Group before correction (Video: mean 4.29, SD 0.86; VR: mean 3.74, SD 1.33; *P*=.02; Cohen *d*=−0.491), but this comparison did not remain significant after FDR adjustment (*q*=0.073). No significant between-group differences were observed for immersion, negative affect, or positive affect after correction.

Secondary Outcome: Learning Engagement

As shown in Table 5, learning engagement (UWES-S; 1 - 7) was broadly comparable between the Video Group (n=47) and the VR Group (n=46). Total engagement did not differ significantly (Video: mean 5.48, SD 1.49; VR: mean 5.70, SD 1.43; *q*=0.669), nor did the vigor (Video: mean 5.68, SD 1.43;

XSL•FO
RenderX

VR: mean 5.65, SD 1.67; q=0.987) or dedication subscales (Video: mean 5.50, SD 1.54; VR: mean 5.55, SD 1.57; q=0.987). Absorption showed a modest, nonsignificant trend favoring the VR Group (Video: mean 5.26, SD 1.85; VR: mean 5.89, SD 1.42; P=.07), which remained nonsignificant after FDR adjustment (q=0.192).

Table . Learning engagement after a single Tai Chi movement-encoding session in community-dwelling older adults randomized to immersive VR[a] or video instruction, November to December 2025.[b]

UWES-S[c] dimensions (range 1-7)	Video Group (n=47), mean (SD)	VR Group (n=46), mean (SD)	*P* value	*q* value
Total engagement	5.48 (1.49)	5.70 (1.43)	.48	0.669
Vigor	5.68 (1.43)	5.65 (1.67)	.93	0.987
Dedication	5.50 (1.54)	5.55 (1.57)	.89	0.987
Absorption	5.26 (1.85)	5.89 (1.42)	.07	0.192

[a]VR: virtual reality.

[b]Between-group differences were tested using independent-samples 2-tailed Welch t tests. False discovery rate adjustment was applied across the self-report outcome set to obtain q values.

[c]UWES-S: Utrecht Work Engagement Scale–Short Form.

Exploratory VR Behavioral Performance

As shown in Table 6, the VR Group's recognition accuracy averaged 75.22% (SD 31.04%), pose accuracy averaged 41.16% (SD 25.15%), and total motion distance was averaged 245.87 (SD 80.23) meters. Among participants with complete paired data (n=44), recognition accuracy was positively associated with posttest movement-encoding knowledge (r=0.317, P=.04), whereas pose accuracy was not (r=0.129, P=.41). These device-recorded indicators are exploratory process measures and should not be interpreted as direct tests of Tai Chi motor execution.

Table . Exploratory device-recorded behavioral performance during immersive virtual reality Tai Chi movement-encoding instruction among community-dwelling older adults, November to December 2025.[a]

Metric	Mean (SD)
Recognition accuracy (%)	75.22 (31.04)
Pose accuracy (%)	41.16 (25.15)
Motion distance (m)	245.87 (80.23)

[a]Recognition accuracy is the percentage of correct responses across 5 recognition items. Pose accuracy is the percentage of correctly hit pose targets (15 total). Motion distance is the sum of movement distances across the head and both hands (m). Correlations with posttest movement-encoding knowledge used 44 complete paired observations.

Safety, Tolerability, and Feasibility

Safety and tolerability outcomes are shown in Table 7. Among VR participants with valid data (n=46), the mean VRSQ total score was 3.77 (SD 11.12) on a 0 to 100 scale; oculomotor and disorientation scores were 2.90 (SD 9.50) and 4.64 (SD 13.25), respectively. A total of 36 out of 46 (78%) participants reported zero symptoms. Among participants with complete paired data (n=45), VRSQ total scores were positively associated with GEQ tension (r=.458, P=.002). Feasibility was reflected in participant flow: of 56 participants assigned to VR, 46 contributed valid baseline data; study-level exclusions included motion sickness or discomfort (n=3), technical failure or no usable data (n=3), and incomplete participation, withdrawal, or other reasons (n=4).

Table . Safety, tolerability, and feasibility of a single immersive virtual reality Tai Chi movement-encoding session among community-dwelling older adults, November to December 2025.[a]

Metric	Values
VRSQ[b] total (range 0 - 100), mean (SD)	3.77 (11.12)
Oculomotor	2.90 (9.50)
Disorientation	4.64 (13.25)
Zero symptom rate, n/N (%)	36/46 (78)

[a]Virtual Reality Sickness Questionnaire (VRSQ) total and subscale scores are scaled from 0 to 100, with higher values indicating more severe symptoms. The zero symptom rate is the absolute number and percentage of virtual reality participants with a total VRSQ symptom sum of 0. The correlation between VRSQ total scores and Game Experience Questionnaire (GEQ) tension scores was based on 45 complete paired observations. Feasibility exclusions are reported in Figure 1 because they occurred before the construction of the valid baseline sample.

[b]VRSQ: Virtual Reality Sickness Questionnaire.

XSL•FO
RenderX

Discussion

Key Findings and Contribution

This randomized study found no evidence that immersive VR improved short-term Tai Chi movement-encoding knowledge relative to content-matched video instruction. After accounting for baseline knowledge and the between-group age difference, posttest performance did not differ between conditions. Among participants who completed the VR session, however, game-related competence and flow were higher, and tension was lower. In this setting, the main observed difference between the delivery modes was, therefore, experiential rather than cognitive performance.

Movement-Encoding Knowledge

Richer sensory presentation does not necessarily produce better learning [19,25,47]. In this study, scores increased descriptively from pretest to posttest in both groups, but VR did not yield higher adjusted posttest performance. This result is consistent with the view that learning gains depend strongly on instructional method and practice structure rather than on the delivery medium alone [19,48]. In Tai Chi, visual representation must also be converted into embodied execution; without haptic guidance, expert correction, or repeated practice, immersive visualization alone may be insufficient to outperform a carefully designed video after a single session [8,48]. Recent evidence indicates that motor benefits are more likely to emerge in repeated, adaptive, or higher-dose VR training than in a single exposure [21-23]. The null between-group result means that the experiential findings should be interpreted separately from claims of improved learning [15,19,32].

Game Experience

The clearest between-group differences were in game experience. VR participants reported higher competence and flow and lower tension, while negative affect was comparable across conditions. These results do not support the assumption that older adults necessarily experience head-mounted displays as more stressful or intimidating [49]. One possible explanation is that VR created a more private, less socially evaluative practice context by reducing external distractions and limiting visible comparisons with others [16,50,51]. This interpretation remains speculative because the study did not directly measure perceived privacy, social-evaluative threat, self-consciousness, or cognitive load. The headset-based setting may have supported concentration and perceived control, but that mechanism requires direct testing [15,16,52].

Intrinsic Motivation and Learning Engagement

Notably, the advantages observed in the game experience did not extend to intrinsic motivation, general engagement, or posttest movement-encoding performance. This dissociation is theoretically informative. Higher perceived competence in VR may reflect a stronger immediate sense of control or confidence in the practice environment, but that feeling did not translate into superior movement-encoding scores within a single session. It is also possible that the immersive presentation induced overconfidence, or that the outcome measure captured cognitive representation more than actual execution quality. Measures such as flow, tension, and perceived competence are sensitive to immediate task conditions, whereas stable motivational orientations, generalized engagement, and measurable motor learning may require repeated exposure, autonomy-supportive scaffolding, and longer time horizons to shift [52-54].

Importantly, differences between GEQ and IMI on conceptually similar dimensions (eg, tension or pressure) may reflect item-level nuance rather than contradiction [37,38,55]. GEQ tension items tend to capture situational frustration or irritation during interaction, whereas IMI pressure or tension items more directly reflect performance-related anxiety or feeling pressured to do well [38]. In this study, immersive VR may have reduced interaction-related tension without materially changing performance-evaluative pressure, which could remain salient in a single-session learning context [54,56]. This dissociation suggests VR operates more as an interface and interaction buffer than a direct remedy for performance anxiety, highlighting design opportunities (eg, supportive feedback or guided practice) to target the latter [16,53,57].

Safety, Tolerability, and Feasibility

The safety profile was encouraging but not uniform. Simulator-sickness symptoms were generally low among VR participants with valid data, and most completers reported no symptoms. At the same time, attrition related to motion sickness and technical problems was not trivial and forms part of the intervention's feasibility and tolerability profile [24,28-30,49]. These exclusions show that acceptability varies across individuals and hardware-software circumstances [24,28-30,49]. Symptom estimates reflect participants who remained in the protocol; therefore, safety conclusions should acknowledge possible survivorship bias and the need for structured screening and monitoring in real-world use [24,28,30,49]. The favorable experiential profile among VR completers may not generalize to older adults who are more susceptible to motion sickness or device-related discomfort.

Practical Implications for Community Implementation

Because the intervention was delivered in community and village activity rooms, the study provides initial evidence of supervised use outside a laboratory [17,24,28,58,59]. For short-term movement-encoding knowledge, no between-group difference was detected; the distinctive findings for VR concerned competence, flow, and tension [16,19,34]. The study did not measure long-term adherence, so implications for sustained participation remain untested. Three practical safeguards are suggested by the feasibility results: (1) brief prescreening for motion sensitivity and comfort with head-mounted displays, (2) gradual exposure with short initial bouts, and (3) reliable on-site support for headset fit, tracking, and interruptions [28,49,58]. VR telemetry may also inform future feedback design and training adjustment [34,58], but these process measures are not interchangeable with learning outcomes. Finally, VR differed from video in immersion, interactivity, embodiment, and sensorimotor coupling; the present design cannot isolate those components.

XSL•FO
RenderX

Limitations and Threats to Interpretation

Several limitations shape interpretation. First, the primary outcome was a brief movement-encoding knowledge test that emphasizes cognitive representation, recall, and recognition of movement elements more than objective movement quality; it cannot establish actual Tai Chi execution, long-term retention, transfer to real-world practice, or expert-rated motor competence [8,32,49]. Second, the analysis focused on immediate posttest performance; comparative benefits of VR may emerge more clearly in delayed retention, transfer, or adherence over time [8,28,49]. Third, the primary analysis required complete pretest and posttest data, while VR-specific attrition due to sickness and technical failure constrains generalizability and should be addressed as an outcome in its own right [28,49]. Fourth, randomization was not stratified by age or prior Tai Chi experience, and the Video Group was older at baseline. Age was controlled in the primary analysis, and exploratory analyses offered no evidence that age modified the intervention effect; however, future trials should consider stratified randomization or minimization procedures.

Future Directions

Future research should move beyond binary comparisons toward mechanism-driven and heterogeneity-driven designs [58,60]. A priority is to test the proposed immersive-shielding pathway by directly measuring social-evaluative threat, perceived privacy, self-consciousness, cognitive load, and self-efficacy as potential mediators linking condition to tension and flow [10,16,50]. Studies should add objective motor metrics (expert ratings, kinematic measures, or pose estimation) to determine whether experiential advantages translate into cleaner movement execution or safer performance under balance demands [16,32]. Longer follow-up is needed to evaluate retention, transfer, and adherence, domains where reduced tension and higher perceived competence may prove consequential but were not tested here [8,15,16]. Finally, safety should be modeled proactively: rather than treating cybersickness as incidental, future trials should identify predictors of intolerance and establish screening and adaptive onboarding protocols so that immersive training can be targeted to those most likely to benefit while minimizing risk for those vulnerable to adverse symptoms [28,49,58].

Conclusions

In this randomized study, immersive VR did not produce higher short-term Tai Chi movement-encoding knowledge than standardized video instruction after adjustment for baseline performance and age [61,62]. Among participants who completed the VR session, perceived competence and flow were higher and tension was lower, suggesting a more favorable immediate game-experience profile [31,62]. These experiential findings do not establish superior learning or motor execution. Discomfort-related withdrawals and technical exclusions also place clear limits on feasibility and generalizability [63-65]. Overall, VR may be useful when experiential engagement is a priority, but the present results do not show an advantage over video for immediate movement-encoding performance.

Acknowledgments

The authors used ChatGPT (OpenAI) and Gemini (Google) during manuscript preparation for text generation and translation. OpenAI Codex was used during the final revision for language revision, editorial formatting, and preparation of the point-by-point response. Generative AI was not used to define the research objectives, questions, or hypotheses; conduct or summarize the literature search; generate figures or other media; collect or analyze data; or interpret the study results. All AI-assisted text was reviewed, verified, and revised by the authors, who take full responsibility for the final manuscript.

Funding

JJ received funding support from the Institutional Research Fund of the School of Art and Design, Guangdong University of Science and Technology (grant GYK-2025BSQDW-70) and the Key Research Platform and Project of Ordinary Higher Education Institutions in Guangdong Province (grant 2025ZDZX4139). Support was received from the Zhejiang Provincial Graduate Education Reform Project (grant JGGC2025743; grant holder: Min Dai).

Data Availability

Individual participant data are not publicly available or shared due to ethical and participant privacy constraints. Analysis code and aggregate nonidentifiable results may be available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: XS
Data curation: XS, LS
Formal analysis: XS
Investigation: LS, YS, HQ, JJ, JZ, WS, HH
Methodology: XS
Project administration: XS
Software: XS
Visualization: XS
Writing – original draft: XS

XSL•FO
RenderX

Writing – review & editing: XS

Conflicts of Interest

None declared.

Multimedia Appendix 1
Five-item movement-encoding scoring rubric for the Tai Chi knowledge test.
[DOCX File, 35 KB - aging_v9i1e96867_app1.docx]

Checklist 1
CONSORT-EHEALTH checklist (V1.6).
[PDF File, 154 KB - aging_v9i1e96867_app2.pdf]

References

1. Anton SD, Cruz-Almeida Y, Singh A, et al. Innovations in geroscience to enhance mobility in older adults. Exp Gerontol 2020;142:111123. [doi: 10.1016/j.exger.2020.111123] [Medline: 33191210]
2. Beard JR, Officer A, de Carvalho IA, et al. The world report on ageing and health: a policy framework for healthy ageing. Lancet 2016 May 21;387(10033):2145-2154. [doi: 10.1016/S0140-6736(15)00516-4] [Medline: 26520231]
3. Montero-Odasso M, Verghese J, Beauchet O, Hausdorff JM. Gait and cognition: a complementary approach to understanding brain function and the risk of falling. J Am Geriatr Soc 2012 Nov;60(11):2127-2136. [doi: 10.1111/j.1532-5415.2012.04209.x] [Medline: 23110433]
4. Li F, Harmer P, Fisher KJ, et al. Tai Chi and fall reductions in older adults: a randomized controlled trial. J Gerontol A Biol Sci Med Sci 2005 Feb;60(2):187-194. [doi: 10.1093/gerona/60.2.187] [Medline: 15814861]
5. Wayne PM, Berkowitz DL, Litrownik DE, Buring JE, Yeh GY. What do we really know about the safety of Tai Chi?: a systematic review of adverse event reports in randomized trials. Arch Phys Med Rehabil 2014;95(12):2470-2483. [doi: 10.1016/j.apmr.2014.05.005] [Medline: 24878398]
6. Wayne PM, Walsh JN, Taylor-Piliae RE, et al. Effect of Tai Chi on cognitive performance in older adults: systematic review and meta-analysis. J Am Geriatr Soc 2014 Jan;62(1):25-39. [doi: 10.1111/jgs.12611] [Medline: 24383523]
7. Fitzroy AB, Kainec KA, Seo J, Spencer RMC. Encoding and consolidation of motor sequence learning in young and older adults. Neurobiol Learn Mem 2021 Nov;185:107508. [doi: 10.1016/j.nlm.2021.107508] [Medline: 34450244]
8. Krakauer JW, Hadjiosif AM, Xu J, Wong AL, Haith AM. Motor Learning. Compr Physiol 2019 Mar 14;9(2):613-663. [doi: 10.1002/cphy.c170043] [Medline: 30873583]
9. Shumway-Cook A, Woollacott MH. Motor Control: Translating Research into Clinical Practice: Lippincott Williams & Wilkins; 2007.
10. Levin MF, Weiss PL, Keshner EA. Emergence of virtual reality as a tool for upper limb rehabilitation: incorporation of motor control and motor learning principles. Phys Ther 2015;95(3):415-425. [doi: 10.2522/ptj.20130579] [Medline: 25212522]
11. Li F, Eckstrom E, Harmer P, Fitzgerald K, Voit J, Cameron KA. Exercise and fall prevention: narrowing the research-to-practice gap and enhancing integration of clinical and community practice. J Am Geriatr Soc 2016 Feb;64(2):425-431. [doi: 10.1111/jgs.13925] [Medline: 26825429]
12. Stevens JA, Voukelatos A, Ehrenreich H. Preventing falls with Tai Ji Quan: a public health perspective. J Sport Health Sci 2014 Mar;3(1):21-26. [doi: 10.1016/j.jshs.2013.10.002] [Medline: 26744633]
13. Franco MR, Tong A, Howard K, et al. Older people's perspectives on participation in physical activity: a systematic review and thematic synthesis of qualitative literature. Br J Sports Med 2015 Oct;49(19):1268-1276. [doi: 10.1136/bjsports-2014-094015] [Medline: 25586911]
14. Spiteri K, Broom D, Bekhet AH, de Caro JX, Laventure B, Grafton K. Barriers and motivators of physical activity participation in middle-aged and older-adults - a systematic review. J Aging Phys Act 2019 Sep 1;27(4):929-944. [doi: 10.1123/japa.2018-0343] [Medline: 31141447]
15. Wulf G, Lewthwaite R. Optimizing performance through intrinsic motivation and attention for learning: the OPTIMAL theory of motor learning. Psychon Bull Rev 2016 Oct;23(5):1382-1414. [doi: 10.3758/s13423-015-0999-9] [Medline: 26833314]
16. Coban M, Bolat YI, Goksu I. The potential of immersive virtual reality to enhance learning: a meta-analysis. Educ Res Rev 2022 Jun;36:100452. [doi: 10.1016/j.edurev.2022.100452]
17. Riva G, Wiederhold BK, Mantovani F. Neuroscience of virtual reality: from virtual exposure to embodied medicine. Cyberpsychol Behav Soc Netw 2019 Jan;22(1):82-96. [doi: 10.1089/cyber.2017.29099.gri] [Medline: 30183347]
18. Chang HY, Binali T, Liang JC, et al. Ten years of augmented reality in education: a meta-analysis of (quasi-) experimental studies to investigate the impact. Comput Educ 2022 Dec;191:104641. [doi: 10.1016/j.compedu.2022.104641]

XSL•FO
RenderX

19. Howard MC, Gutworth MB. A meta-analysis of virtual reality training programs for social skill development. Comput Educ 2020 Jan;144:103707. [doi: 10.1016/j.compedu.2019.103707]
20. Jensen L, Konradsen F. A review of the use of virtual reality head-mounted displays in education and training. Educ Inf Technol 2018;23(4):1515-1529. [doi: 10.1007/s10639-017-9676-0]
21. Everard G, Vermette M, Dumas-Longpré E, et al. Self-adaptive over progressive non-adaptive immersive virtual reality serious game to promote motor learning in older adults - a double blind randomized controlled trial. Neuroscience 2025 Apr 6;571:7-18. [doi: 10.1016/j.neuroscience.2025.02.053] [Medline: 40021077]
22. Lee J, Phu S, Lord SR, Okubo Y. Effects of immersive virtual reality training on balance, gait and mobility in older adults: a systematic review and meta-analysis. Gait Posture 2024 May;110:129-137. [doi: 10.1016/j.gaitpost.2024.03.009] [Medline: 38581933]
23. Kwan RYC, Liu J, Sin OSK, et al. Effects of virtual reality motor-cognitive training for older people with cognitive frailty: multicentered randomized controlled trial. J Med Internet Res 2024 Sep 11;26:e57809. [doi: 10.2196/57809] [Medline: 39259959]
24. Schaumburg M, Imtiaz A, Zhou R, Bernard M, Wolbers T, Segen V. Immersive virtual reality for older adults: challenges and solutions in basic research and clinical applications. Ageing Res Rev 2025 Jul;109:102771. [doi: 10.1016/j.arr.2025.102771] [Medline: 40373984]
25. Clark RE. Reconsidering research on learning from media. Rev Educ Res 1983;53(4):445-459. [doi: 10.3102/00346543053004445]
26. Clark RE. Media will never influence learning. Educ Technol Res Dev 1994;42(2):21-29. [doi: 10.1007/BF02299088]
27. Jochmann E, Jochmann T, Weber M, Weigel K, Klingner C. Impact of sensorimotor mismatch on virtual reality sickness and user experience: age-related differences in a randomized trial. J Neuroeng Rehabil 2025 Jul 3;22(1):143. [doi: 10.1186/s12984-025-01677-x] [Medline: 40611269]
28. Tuena C, Pedroli E, Trimarchi PD, et al. Usability issues of clinical and research applications of virtual reality in older people: a systematic review. Front Hum Neurosci 2020;14:93. [doi: 10.3389/fnhum.2020.00093] [Medline: 32322194]
29. Weech S, Kenny S, Barnett-Cowan M. Presence and cybersickness in virtual reality are negatively related: a review. Front Psychol 2019;10:158. [doi: 10.3389/fpsyg.2019.00158] [Medline: 30778320]
30. Li Y, Shiyanov I, Muschalla B. Assessing cybersickness in a virtual reality intervention among older adults living in nursing homes: a feasibility study. Disabil Rehabil Assist Technol 2025 Oct;20(7):2339-2347. [doi: 10.1080/17483107.2025.2503909] [Medline: 40357987]
31. Makransky G, Petersen GB. The cognitive affective model of immersive learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality. Educ Psychol Rev 2021 Sep;33(3):937-958. [doi: 10.1007/s10648-020-09586-2]
32. Schmidt RA, Lee TD, Winstein C, Wulf G, Zelaznik HN. Motor Control and Learning: A Behavioral Emphasis: Human Kinetics; 2018.
33. Makransky G, Terkildsen TS, Mayer RE. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learn Instr 2019 Apr;60:225-236. [doi: 10.1016/j.learninstruc.2017.12.007]
34. Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I. A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. Comput Educ 2020 Apr;147:103778. [doi: 10.1016/j.compedu.2019.103778]
35. Riva G, Baños RM, Botella C, Mantovani F, Gaggioli A. Transforming experience: the potential of augmented reality and virtual reality for enhancing personal and clinical change. Front Psychiatry 2016;7:164. [doi: 10.3389/fpsyt.2016.00164] [Medline: 27746747]
36. Little RJA, Rubin DB. Statistical Analysis with Missing Data: John Wiley & Sons; 2019. [doi: 10.1002/9781119482260]
37. McAuley E, Duncan T, Tammen VV. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: a confirmatory factor analysis. Res Q Exerc Sport 1989 Mar;60(1):48-58. [doi: 10.1080/02701367.1989.10607413] [Medline: 2489825]
38. IJsselsteijn WA, de Kort YAW, Poels K. The Game Experience Questionnaire: Technische Universiteit Eindhoven; 2013. URL: https://research.tue.nl/en/publications/the-game-experience-questionnaire/ [accessed 2026-09-21]
39. Johnson D, Gardner MJ, Perry R. Validation of two game experience scales: the Player Experience of Need Satisfaction (PENS) and Game Experience Questionnaire (GEQ). Int J Hum Comput Stud 2018 Oct;118:38-46. [doi: 10.1016/j.ijhcs.2018.05.003]
40. Schaufeli WB, Bakker AB, Salanova M. The measurement of work engagement with a short questionnaire: a cross-national study. Educ Psychol Meas 2006 Aug;66(4):701-716. [doi: 10.1177/0013164405282471]
41. Kim HK, Park J, Choi Y, Choe M. Virtual Reality Sickness Questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. Appl Ergon 2018 May;69:66-73. [doi: 10.1016/j.apergo.2017.12.016] [Medline: 29477332]
42. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010;8:18. [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
43. Miller GA, Chapman JP. Misunderstanding analysis of covariance. J Abnorm Psychol 2001;110:40-48. [doi: 10.1037/0021-843X.110.1.40] [Medline: 11261398]

XSL•FO
RenderX

44. Van Breukelen GJP. ANCOVA versus change from baseline: more power in randomized studies, more bias in nonrandomized studies [corrected]. J Clin Epidemiol 2006 Sep;59(9):920-925. [doi: 10.1016/j.jclinepi.2006.02.007] [Medline: 16895814]
45. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc, B: Stat Methodol 1995 Jan;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
46. Pike N. Using false discovery rates for multiple comparisons in ecology and evolution. Methods Ecol Evol 2011 Jun;2(3):278-282. [doi: 10.1111/j.2041-210X.2010.00061.x]
47. Mayer RE. Computer games in education. Annu Rev Psychol 2019 Jan 4;70:531-549. [doi: 10.1146/annurev-psych-010418-102744] [Medline: 30231003]
48. Merchant Z, Goetz ET, Cifuentes L, Keeney-Kennicutt W, Davis TJ. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: a meta-analysis. Comput Educ 2014 Jan;70:29-40. [doi: 10.1016/j.compedu.2013.07.033]
49. Saredakis D, Szpak A, Birckhead B, Keage HAD, Rizzo A, Loetscher T. Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. Front Hum Neurosci 2020;14:96. [doi: 10.3389/fnhum.2020.00096] [Medline: 32300295]
50. Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychol Bull 2004 May;130(3):355-391. [doi: 10.1037/0033-2909.130.3.355] [Medline: 15122924]
51. Riva G, Mantovani F, Capideville CS, et al. Affective interactions using virtual reality: the link between presence and emotions. CyberPsychol Behav 2007 Feb;10(1):45-56. [doi: 10.1089/cpb.2006.9993] [Medline: 17305448]
52. Mihály C. Flow: The Psychology of Optimal Experience: Harper & Row; 1990.
53. Deci EL, Ryan RM. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000 Oct;11(4):227-268. [doi: 10.1207/S15327965PLI1104_01]
54. Howard JL, Bureau J, Guay F, Chong JXY, Ryan RM. Student motivation and associated outcomes: a meta-analysis from self-determination theory. Perspect Psychol Sci 2021 Nov;16(6):1300-1323. [doi: 10.1177/1745691620966789] [Medline: 33593153]
55. Roe RA, Inceoglu I. Measuring states and traits in motivation and emotion: a new model illustrated for the case of work engagement. In: Leong FTL, Bartram D, Cheung FM, Geisinger KF, Iliescu D, editors. The ITC International Handbook of Testing and Assessment: Oxford University Press; 2016:63-88. [doi: 10.1093/med:psych/9780199356942.003.0006]
56. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037//0003-066X.55.1.68] [Medline: 11392867]
57. Hofmann W, Schmeichel BJ, Baddeley AD. Executive functions and self-regulation. Trends Cogn Sci 2012 Mar;16(3):174-180. [doi: 10.1016/j.tics.2012.01.006] [Medline: 22336729]
58. Birckhead B, Khalil C, Liu X, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973. [doi: 10.2196/11973] [Medline: 30702436]
59. Ramalho A, Duarte-Mendes P, Paulo R, Serrano J, Petrica J. Crossing the digital frontier: are older adults ready for virtual reality workouts? Front Public Health 2024;12:1324004. [doi: 10.3389/fpubh.2024.1324004] [Medline: 38389949]
60. Cronbach LJ. The two disciplines of scientific psychology. Am Psychol 1957;12(11):671-684. [doi: 10.1037/h0043943]
61. Makransky G, Andreasen NK, Baceviciute S, Mayer RE. Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. J Educ Psychol 2021;113(4):719-735. [doi: 10.1037/edu0000473]
62. Petersen GB, Petkakis G, Makransky G. A study of how immersion and interactivity drive VR learning. Comput Educ 2022 Apr;179:104429. [doi: 10.1016/j.compedu.2021.104429]
63. Caserman P, Garcia-Agundez A, Gámez Zerban A, Göbel S. Cybersickness in current-generation virtual reality head-mounted displays: systematic review and outlook. Virtual Real 2021 Dec;25(4):1153-1170. [doi: 10.1007/s10055-021-00513-6]
64. Davis R. The feasibility of using virtual reality and eye tracking in research with older adults with and without Alzheimer's disease. Front Aging Neurosci 2021;13:607219. [doi: 10.3389/fnagi.2021.607219] [Medline: 34248596]
65. Drazich BF, McPherson R, Gorman EF, et al. In too deep? A systematic literature review of fully-immersive virtual reality and cybersickness among older adults. J Am Geriatr Soc 2023 Dec;71(12):3906-3915. [doi: 10.1111/jgs.18553] [Medline: 37560978]

Abbreviations

ANCOVA: analysis of covariance
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
FDR: false discovery rate
GEQ: Game Experience Questionnaire
IMI: Intrinsic Motivation Inventory
RQ: research question

XSL•FO
RenderX

UWES-S: Utrecht Work Engagement Scale–Short Form
VR: virtual reality
VRSQ: Virtual Reality Sickness Questionnaire

Edited by Y Jiang; submitted 01.Apr.2026; peer-reviewed by AK Panahi, A Maneuvrier, L Ilodigwe; revised version received 24.Jul.2026; accepted 14.Aug.2026; published 30.Sep.2026.

Please cite as:
Song X, Sun L, Song Y, Qu H, Jin J, Zhu J, Song W, Huang H
Knowledge Performance, User Experience, and Cybersickness With Immersive Virtual Reality for Tai Chi Movement-Encoding in Older Adults: Randomized Comparison Study
JMIR Aging 2026;9:e96867
URL: https://aging.jmir.org/2026/1/e96867
doi:10.2196/96867

© XiaCheng Song, Lu Sun, Yaqing Song, Huafeng Qu, Jing Jin, Junfeng Zhu, Wenwen Song, Huirong Huang. Originally published in JMIR Aging (https://aging.jmir.org), 30.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing a Digital Platform (KOKU-Nut) to Improve Nutrition in Older Adults: Mixed Methods Feasibility Randomized Controlled Trial

Chloe French[1], PhD; Sorrel Burden[1,2], PhD; Jaheeda Gangannagaripalli[1,3,4], PhD; Emma Stanmore[1,3,4], PhD

[1]School of Health Sciences, University of Manchester, Jean McFarlane Building, University of Manchester, Manchester, United Kingdom

[2]Northern Care Alliance NHS Trust, Salford Care Organisation, Salford, United Kingdom

[3]NIHR Applied Research Collaboration Greater Manchester, Manchester, England, United Kingdom

[4]Manchester Academic Health Science Centre, Manchester, United Kingdom

Corresponding Author:
Chloe French, PhD
School of Health Sciences, University of Manchester, Jean McFarlane Building, University of Manchester, Manchester, United Kingdom

Abstract

Background: Keep-on-Keep-up Nutrition (KOKU-Nut) is a free, tablet-based digital platform that focuses on increasing physical activity and improving the dietary intake of older adults.

Objective: The aim of this study was to assess the feasibility of using KOKU-Nut among community-dwelling older adults. Feasibility was assessed by considering recruitment and retention rates, acceptability of the intervention, and study design.

Methods: Participants (community-dwelling adults aged ≥65 years) were randomized 1:1 to either the intervention or control group. The intervention group was asked to engage with KOKU-Nut 3 times a week for 12 weeks. Participants in the control group received a leaflet promoting a healthy lifestyle. All participants completed questionnaires at baseline and 12 weeks. A sample of participants was asked to complete an optional interview. The study collected data on dietary intake (24-hour food diary), physical function (grip strength and 5-time sit-to-stand exercise), usability of the intervention (System Usability Scale), and safety (adverse events).

Results: Of 51 participants assessed for eligibility, 31 were randomized, and 28 completed the 12-week follow-up. Of the 14 participants in the intervention group, 10 (71.4%) reported engagement with KOKU-Nut 3 or more times a week. There was no difference in physical function or dietary intake at 12 weeks between participants in the intervention and control groups after adjusting for age and baseline values.

Conclusions: This feasibility randomized controlled trial indicates that the study design was appropriate and acceptable to older adults, as demonstrated by the recruitment and retention rates. This is promising and demonstrates the potential for conducting a future powered randomized controlled trial to assess the effectiveness of KOKU-Nut.

Trial Registration: ClinicalTrials.gov NCT05943366; https://clinicaltrials.gov/study/NCT05943366

International Registered Report Identifier (IRRID): RR2-10.2196/50922

(JMIR Aging 2026;9:e87351) doi:10.2196/87351

KEYWORDS

feasibility; digital health; malnutrition; older adults; nutrition

Introduction

Data from the Office for National Statistics show that, in 2016, there were 11.8 million UK residents aged 65 years and over, representing 18% of the UK population [1]. Projections for 2066 estimate that this will increase by 8% to 20.4 million people [1]. Despite increases in life expectancy, healthy lifespan is not increasing at the same rate, and these additional years are often spent in poor physical and mental health [2].

Malnutrition is common among older adults and often arises from an interplay of physiological, psychological, and social factors, including dysphagia, chewing difficulties, bereavement, social isolation, and mobility limitations, all of which can influence nutritional intake [3-6]. Estimates suggest that approximately 8.5% of community-dwelling older adults are at risk of or experience malnutrition [7]. Observational and experimental evidence demonstrates that a healthy diet can slow disease progression, reduce the rate of functional decline, and prevent undernutrition and frailty [5,8-11], whereas the presence

XSL•FO
RenderX

of malnutrition leads to adverse health outcomes, premature mortality, and increased costs on health and social care [12,13]. Older adults living in the community are likely to have greater independence and choice regarding their dietary intake than individuals living in care homes or hospital settings, where meals and food provision are more structured and institutionally regulated [14]. A James Lind Alliance Priority Setting Partnership identified that early intervention in vulnerable groups was the number one research and innovation priority for adult malnutrition [15]. This highlights the need and rationale to support the nutritional intake of older adults to help prevent malnutrition

An aging population with high health care demands [1] coupled with technological advances creates an opportunity for innovative digital solutions to support health care [2,16,17]. Keep-on-Keep-up is a National Health Service–approved digital fall prevention program designed with and for older adults [18]. The platform includes progressive strength and balance exercises [19-21], gamification to improve awareness of bone health and safety in the home, and behavior change techniques to support users in building habits and tracking behaviors. Keep-on-Keep-up Nutrition (KOKU-Nut) is the latest iteration [6,22], providing nutrition advice to complement the strength and balance exercises. Multi-domain interventions appear to have greater benefits for functional outcomes in older adults [23] given that adverse health in older age often encompasses multiple deficits across multiple domains [24].

The aim of this study was to assess the feasibility of conducting a randomized controlled trial using the digital platform KOKU-Nut to improve dietary intake in community-dwelling older adults. Feasibility was assessed by considering recruitment and retention rates and acceptability of the study design. This study also considered the usability, preliminary efficacy, and safety of KOKU-Nut.

Methods

Overview

A sequential mixed methods study design was chosen that used quantitative and then qualitative research techniques given the difficulty of evaluating a complex intervention and to maximize the breadth and depth of the research findings. The protocol was previously published [22], and the trial was registered prior to recruitment (NCT05943366). This study was reported in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines [25] to ensure comprehensive and transparent reporting of the qualitative methods and findings (Multimedia Appendix 1).

Ethical Considerations

Ethics approval was obtained from the University of Manchester Research Ethics Committee on August 18, 2023 (reference 2023-17372-30569), and this study complied with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to commencement of the research.

Eligibility Criteria

Participants were included in the study if they were aged 65 years and older, living independently in the community in Greater Manchester, and willing to use an iPad or tablet (their own or one provided by the research team) for the duration of the study. Participants were excluded if they were unable to communicate in English or had a known cognitive impairment.

Recruitment and Sampling

Men and women aged 65 years and older living independently in the community were recruited on a rolling basis through purposive sampling via local Age UK groups, assisted living facilities, and university networks. Initially, a phone call was arranged to determine eligibility. The researcher verbally explained the aims of the study, how data were collected, and what was required from participants and encouraged participants to ask any questions or raise any concerns.

Recruitment to the qualitative study was done purposefully toward the end of the intervention period to recruit a demographically representative group of participants, including variation in sex, age, and treatment group. Researchers continued to invite participants until data saturation was reached, defined a priori as the point at which no new themes emerged from successive interviews during data coding and charting.

Randomization

Upon meeting the eligibility criteria and providing written informed consent, participants were randomized into either the intervention or control group by an independent member of the research team using sealedenvelope [26]. Participants were informed of their group allocation after baseline data had been collected.

Study Intervention

KOKU-Nut [6] is a digital platform that can be downloaded on tablet or iPad devices. The platform includes progressive and personalized strength and balance exercises alongside a suite of educational games, including a nutrition matching pair card game to educate and nudge users to support their dietary intake (see Multimedia Appendix 1 for screenshots). In particular, the educational content focuses on (1) increasing intake of high-quality protein items such as milk, yogurt, and fish; (2) increasing fluid intake to 6 to 8 glasses a day; and (3) educating users on the signs and risks of malnutrition.

Participants in the intervention group were asked to engage with KOKU-Nut in their own time for approximately 30 minutes 3 times a week for 12 weeks. At the baseline visit, participants were supported with downloading the digital platform onto their iPad or tablet. In cases in which participants did not have the necessary device to use the intervention, an iPad (Apple Inc) with KOKU-Nut installed was provided to the participants for the duration of the intervention. The researcher explained and demonstrated the features of the digital platform and assisted with any technical queries. A crib sheet and contact details for the researcher were provided to participants for additional support and to help with technical issues. A family member or friend was invited to attend the baseline visit if the participant

XSL•FO
RenderX

wanted additional support. Participants were advised to continue their usual care and treatment.

Control Group

Participants assigned to the control group received a leaflet developed by Age UK that contained information about the importance of a healthy lifestyle, including the benefits of staying active and eating healthy [27]. Participants in the control group were also advised to continue with usual treatments and care from their health care professionals.

Data Collection

Written informed consent was collected from all participants prior to collecting the baseline case report file or conducting any assessments at the participants' homes. The baseline case report file collected data on the participants' sociodemographic characteristics, as well as the following measures. Health-related quality of life was assessed using the EQ-5D-5L, which considers mobility, self-care, usual activities, pain and discomfort, and anxiety and depression [28]. Each domain is rated across 5 levels of severity, which are combined and converted into a single summary value [29]. The instrument also includes a visual analog scale, where participants rate their overall health from 0 (worst health imaginable) to 100 (best health imaginable). Risk of malnutrition was assessed using the validated Malnutrition Universal Screening Tool [30], and overall risk was assessed as a binary measure to compare those with low vs medium or high risk of malnutrition. Mood was assessed using the short version of the Geriatric Depression Scale [31]. Items were scored dichotomously and summed; overall risk was then assessed as a binary variable so that participants scoring 2 or higher were at risk of depression. These assessments were chosen due to their ease of administration, relevance to the research question, and widespread use across community and clinical settings in the United Kingdom.

During the baseline and 12-week assessments, the researcher also measured grip strength (in kilograms), the time taken to complete 5 sit-to-stand exercises (seconds), and frailty status based on the Clinical Frailty Scale. Further details on these procedures are provided in the study protocol [22].

Participants were asked to complete a 24-hour dietary recall using the online questionnaire Intake24 (version 2) at baseline and 12 weeks [32,33]. Participants were encouraged to follow usual food and drink consumption and then report all food and drink consumed in the online software as accurately as possible. If participants were unable to access Intake24 or experienced technical difficulties, the researcher input this information based on a paper-based version of the participants' food diary.

The researcher arranged the follow-up assessment 12 weeks after the baseline visits. Baseline measurements were repeated, and participants in the intervention group also reported outcome data related to the usability and acceptability of the intervention, assessed using the technology acceptance model (TAM) [34] and System Usability Scale (SUS) [35].

A subsample of participants completed one-to-one semistructured interviews at the end of the intervention period with a female researcher (CF) experienced in qualitative methodology. Interviews were conducted online or face-to-face based on the preference of the participant, audio recorded, and expected to last no more than 60 minutes. Prior to completing the interview, participants provided additional consent regarding the recording, storage, and use of their interview data. A topic guide was informed by the research aims and similar studies and was used to explore barriers to recruitment, the appropriateness of the outcome measures, and the feasibility of implementing the KOKU-Nut intervention.

Outcomes

The main outcome was the feasibility, acceptability, and usability of the KOKU-Nut digital platform (Multimedia Appendix 1). Additional outcomes were the potential benefits of engaging with KOKU-Nut, specifically considering protein intake, risk of malnutrition, physical function, and health-related quality of life, as previously described [22].

Statistical Analysis

Feasibility outcomes and baseline characteristics of the participants were summarized using descriptive statistics, including proportions or means with SDs. Between-group differences in health-related measures at 12 weeks were explored using analysis of covariance adjusting for baseline values of the outcome and participants' age at recruitment. These analyses were conducted on an exploratory basis to provide preliminary estimates of potential intervention effects. Given the feasibility nature of the study, the sample size was not powered to detect statistically significant differences.

Interviews were transcribed verbatim and then checked by a researcher to ensure accuracy. Transcripts were then managed using NVivo (version 12; Lumivero) [36] and analyzed inductively using the 5 stages of framework analysis [37,38]. The initial stage involved repeatedly reading the transcripts, followed by the development of codes and the identification of an initial thematic framework (CF). Subsequently, data were systematically indexed according to this framework. The fourth stage involved organizing and summarizing the data into thematic charts to facilitate comparison and synthesis. Finally, patterns within the data were interpreted and discussed with all coauthors to generate overarching insights. To enhance analytical rigor, all stages of the process were reviewed and discussed by a minimum of 2 researchers.

Results

Sociodemographic Characteristics of the Participants

Table 1 shows the baseline characteristics of all 31 participants enrolled in the study. The mean age was 81.3 (SD 8.3) years, and most participants were female (n=24, 77.4%) and White (n=27, 87.1%). Almost all (n=28, 90.3%) participants were retired, and just under half (n=14, 45.2%) were widowed. Overall, most participants were at a low risk of depression based on the 5-item Geriatric Depression Scale. In total, 12.5% (2/16) of the participants in the intervention group were at risk of depression at baseline, which may have affected motivation to use and engage with KOKU-Nut.

XSL•FO
RenderX

More than half (16/31, 51.6%) of participants were taking 5 or more medications, and there were more participants with polypharmacy in the intervention group (10/16, 62.5%) compared to the control group (6/15, 40%). More than half (16/31, 51.6%) of all participants self-reported one or more falls in the previous 12 months, and fall history was similar between the 2 groups. At baseline, 16.1% (5/31) of the participants were at risk of malnutrition based on a score of 1 or more on the Malnutrition Universal Screening Tool, and there was no difference in risk between the groups. The median frailty score was 4 (IQR 3-5), indicating that the sample was living with very mild frailty. This has been described as “while not dependent on others for daily help, often symptoms limit activities.” A common concern is being “slowed up” and/or being “tired during the day” [39].

XSL•FO
RenderX

Table . Baseline characteristics of the participants (N=31).

	Intervention (n=16)	Control (n=15)	Total
Sex, n (%)			
Male	4 (25)	3 (20)	7 (22.6)
Female	12 (75)	12 (80)	24 (77.4)
Age (y), mean (SD)	79.9 (8.3)	82.7 (8.4)	81.3 (8.3)
Race, n (%)			
White	13 (81.3)	14 (93.3)	27 (87.1)
Other	3 (18.8)	1 (6.7)	4 (12.9)
Marital status, n (%)			
Married	5 (31.3)	5 (33.3)	10 (32.3)
Widowed	7 (43.8)	7 (46.7)	14 (45.2)
Separated or divorced	4 (25)	3 (20)	7 (22.6)
IMD[a] (decile 1-10), median (IQR)	5.5 (4-8.75)	6 (4-7)	6 (4-8)
Number of medications, n (%)			
<4	6 (37.5)	9 (60)	15 (48.4)
≥5	10 (62.5)	6 (40)	16 (51.6)
Fall history, n (%)			
Never	8 (50)	7 (46.7)	15 (48.4)
Once	2 (12.5)	3 (20)	5 (16.1)
Twice or more	6 (37.5)	5 (33.3)	11 (35.5)
Requires support with shopping, n (%)			
Independent	9 (56.3)	11 (73.3)	20 (64.5)
Receives some support	5 (31.3)	4 (26.7)	9 (29.0)
Completely dependent	2 (12.5)	0 (0)	2 (6.5)
Internet use, n (%)			
Daily	9 (56.3)	9 (60)	18 (58.1)
Weekly	4 (25)	3 (20)	7 (22.6)
Never	3 (18.8)	3 (20)	6 (19.4)
Technology use, n (%)			
Owns tablet or iPad	5 (31.3)	4 (26.7)	9 (29.0)
Owns smartphone	14 (87.5)	6 (40)	20 (64.5)
Owns laptop or computer	10 (62.5)	10 (66.7)	20 (64.5)
Depression, n (%)			
Low risk (GDS[b]<2)	14 (87.5)	12 (80)	26 (83.9)
At risk (GDS≥2)	2 (12.5)	3 (20)	5 (16.1)
BMI (kg/m^2), n (%)			
≤24.9	7 (43.8)	8 (53.3)	15 (48.4)
≥25	9 (56.3)	7 (46.7)	16 (51.6)
Risk of malnutrition[c], n (%)			
Low	13 (81.3)	13 (86.7)	26 (83.9)
Medium-high	3 (18.8)	2 (13.3)	5 (16.1)
Frailty score[d], median (IQR)	3.5 (2.3-5)	4 (3-5)	4 (3-5)

XSL•FO
RenderX

[a]IMD: Index of Multiple Deprivation.

[b]GDS: Geriatric Depression Scale.

[c]Assessed using the Malnutrition Universal Screening Tool (MUST; low risk=MUST score of 0; medium-high risk=MUST score of ≥1).

[d]Clinical Frailty Score assessed from 1, representing very fit, to 9, representing terminally ill.

Feasibility Outcomes: Recruitment and Retention

Overview

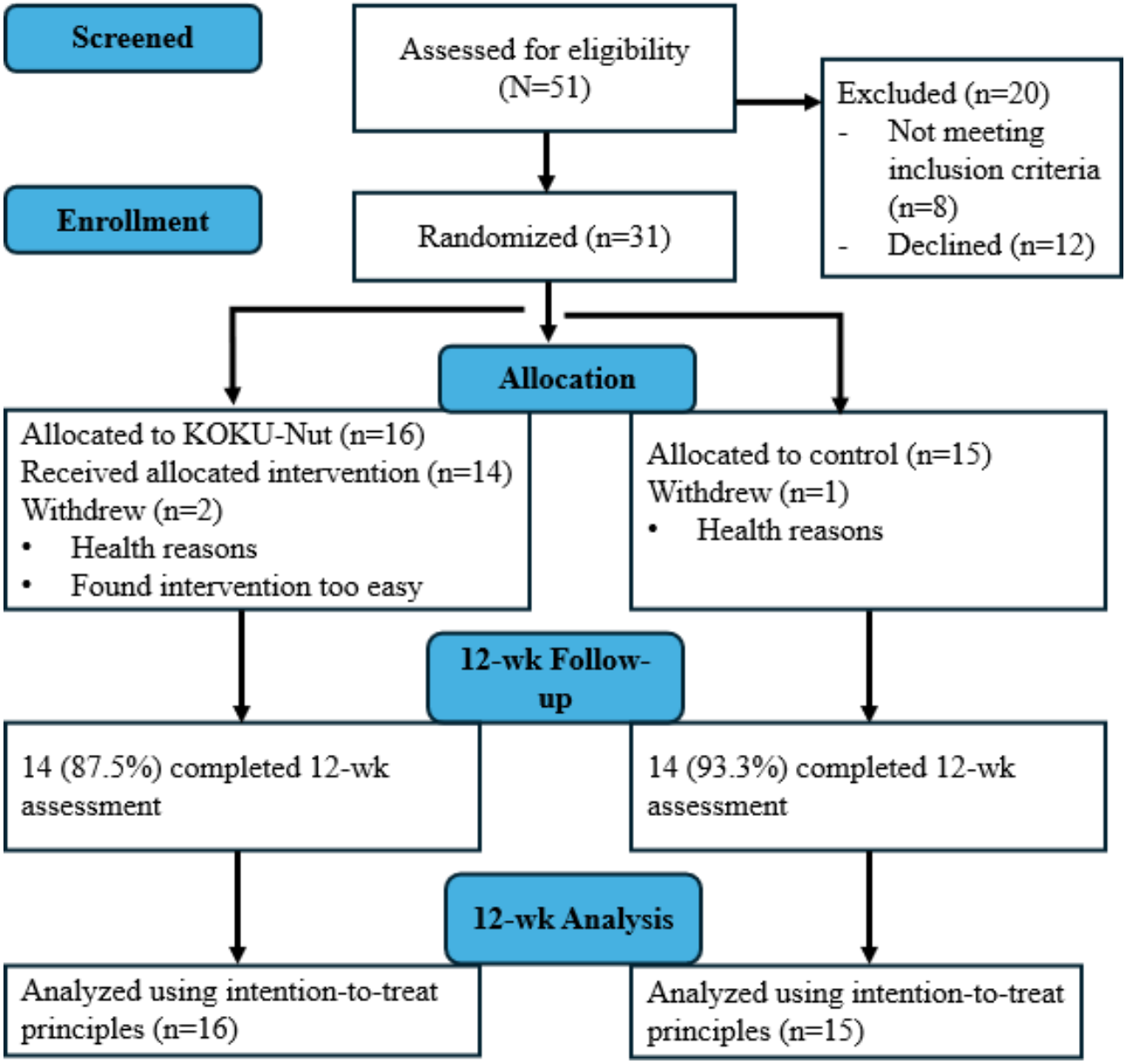
Between March 2024 and August 2024, a total of 51 potential participants were screened to participate in the study, of whom 8 (15.7%) were excluded because they did not meet the eligibility criteria (n=5 were aged <65 years, n=1 lived outside of Greater Manchester, and n=2 had mild cognitive impairment); a further 12 (23.5%) older adults declined to participate.

In total, 31 older adults were enrolled in the study and randomly assigned to the intervention (n=16, 51.6%) or control (n=15, 48.4%) group (Figure 1). Prior to the 12-week follow-up, 12.5% (2/16) of the participants in the intervention group withdrew (one had a knee operation, and the other found the KOKU-Nut program too easy). In total, 6.7% (1/15) of the participants in the control group withdrew because they had a fall and no longer wanted to take part. As a result, the overall retention rate at the 3-month follow-up was 90.3% (28/31). The study was completed with high retention rates and low levels of missing data; these indicators suggest that the study is feasible (Multimedia Appendix 1). It may be beneficial to have more targeted recruitment strategies to increase recruitment rates and support the conversion of eligible to enrolled participants in a future trial.

Ten participants completed an interview, 6 of whom were in the intervention group, and the remaining 4 were in the control group. Interviews were conducted face-to-face at the participants' homes and lasted for a mean of 31.5 (SD 12.1; range 18-51) minutes. Framework analysis of the dataset uncovered 3 key themes.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. KOKU-Nut: Keep-on-Keep-up Nutrition.



Theme 1: Acceptability of the Study

Participants found the questionnaires and physical function tests acceptable; however, there was some resistance regarding the 24-hour dietary recall. The experience of using the online software was variable: some found it interesting to consider the different portion sizes, whereas others were unable or did not want to use the online software and so completed an interviewer-led version. Participants also reported changing their intake to facilitate the recording process or having difficulty entering certain foods:

> *I was very interested in it [Intake24 survey] and the way it was done, particularly with the measurement of quantities and the options you were given which I thought were very helpful...and I actually enjoyed it.* [Female participant; control group]

> *I had a word with a couple of people and described what I had done so far, they found the idea of the meal report off-putting.* [Female participant; intervention group]

> *I omitted some foods as it was tedious to complete, especially when making food from scratch. I didn't know how to enter homemade soup or cake.... It felt like it was tailored towards someone who would buy a lot of pre-packaged foods, also it doesn't ask about salt and pepper.* [Female participant; intervention group]

> *First time I picked the simplest meal that I ever eat.... But the second time I picked a slightly more complicated one.* [Female participant; control group]

In general, the intervention (KOKU-Nut) was acceptable, beneficial, and enjoyable, although there were still some concerns about the digital aspect:

> *Have found that since making changes I have more energy. Used to experience an afternoon slump*

XSL•FO
RenderX

around 3pm but now I feel more energised. [Female participant; intervention group]

For people that are familiar with tablets it would be fantastic.... But I've got friends in their seventies and eighties who don't use smartphones! [Male participant; intervention group]

Theme 2: Healthy Behaviors

Many participants in both the intervention and control groups highlighted that taking part in the study made them reflect on their health and their current dietary choices, which made them want to make positive changes. Participants in the intervention group found that using KOKU-Nut prompted them to consider their health and start a new routine:

Since [starting the study, I] made small changes to diet and exercise routine and app nudged me in the right direction. [Male participant; intervention group]

I want to be active and independent for as long as possible and this provided encouragement to make small changes. [Female participant; control group]

I now try to get protein at each meal.... So yes, I've introduced yoghurt on top of the porridge. [Male participant; intervention group]

Since starting the study have made an effort to increase protein so now, I check the packaging to look for grams of protein to ensure that having protein with all meals. [Female participant; intervention group]

KOKU has made me more conscious of cooking and what we can and should be having for a healthy diet. [Female participant; intervention group]

Participants also recognized the importance of habit formation and that incorporating KOKU-Nut into their routine helped with engagement:

I think it might be good for people to think about where and when they use it and how it can fit in with their routine. [Female participant; intervention group]

Participants in the control group also highlighted that the study prompted them to make more positive lifestyle changes, with one participant saying the following:

And then I was thinking it's not just your physical health, you've got to think of your mental health as well. So, I've signed up to do Tai Chi on a Monday. [Female participant; control group]

However, participants identified the limitations of using a booklet to support behavior change:

The thing with a booklet is, you read it and then you put it to one side and then that's it. And if you need prompting there isn't anything there. [Female participant; control group]

Theme 3: Future Considerations

Participants highlighted additional features that they thought would add value to KOKU-Nut and support them in making positive changes to their diet. This included a myth-busting information section, an example week's menu, or recipe ideas with video and written content:

Due to my hypertension I never know if I can have cheese or yoghurt so it would be great to have information about if I can eat these foods. [Female participant; intervention group]

It would be great to have more nutrition advice and content for example to explain what a high protein diet actually is. [Male participant; intervention group]

I'd find it very useful to have videos or written recipes with an example weeks menu...or even just some recipe inspiration. [Female participant; intervention group]

Additional Outcomes

Usability and Safety of the Intervention

Of 14 participants in the intervention group, 10 (71.4%) self-reported engagement with KOKU-Nut 3 or more times a week as suggested at the start of the intervention. Of 16 participants in the intervention group, 11 (68.8%) received an iPad provided by the research team, and 5 participants (31.2%) had their own devices. Based on the SUS, KOKU-Nut had very good usability, with a mean score of 79.7 (SD 15.2), above the industry standard average of 68 [40]. Participants with their own iPads provided slightly higher and more consistent usability scores (mean 82, SD 6.2) compared to participants who did not have their own iPad (mean 78.6, SD 18.7). There were no adverse events reported over the 12-week period.

User Experience

Fourteen participants considered the acceptability of KOKU-Nut based on the TAM as this was not completed by the 2 participants who withdrew from the study. Participants rated the nutrition game highly on perceived ease of use, perceived usefulness, attitude toward use, and intention to use (Multimedia Appendix 1). There was moderate correlation between perceived usefulness and attitude toward and intention to use (Table 2). This is important as the TAM suggests that there is an association between the user's intention to use the technology and their actual use behavior [41].

XSL•FO
RenderX

Table . Technology acceptance model questionnaire for community-dwelling older adults and their experience of using Keep-on-Keep-up Nutrition (KOKU-Nut). Each aspect was scored on a scale ranging from 0 to 7 such that 7 represented "most favorable" (ie, highest level of perceived usefulness). There were missing data from 2 participants in each case due to withdrawals.

	Perceived ease of use	Perceived usefulness	Attitude toward use	Intention to use
Scores, median (IQR)	6.3 (5.4-7.0)	5.3 (4.9-5.9)	6.9 (5.9-7.0)	5.3 (3.7-6.3)
Correlation with intention to use, Pearson r (*P* value)	0.3 (0.27)	0.5 (0.05)	0.65 (0.01)	—[a]

[a]Correlation between intention to use and intention to use is not reported.

Preliminary Effectiveness of the Intervention

Participants undertook assessments of functional ability, dietary intake, and quality of life at baseline and the 12-week follow-up. An exploratory intention-to-treat analysis was conducted, but as this was a feasibility study, it was not sufficiently powered to detect significant differences between groups.

There was no difference in grip strength or time taken to complete 5 sit-to-stand exercises at the follow-up assessment between groups after adjusting for age and baseline values.

EQ-5D-5L scores were similar between the groups at baseline and reflect the norms for this age group [42]. After the 12-week intervention, there was a difference in EQ-5D-5L scores using analysis of covariance ($F_{1,27}$=7.26; *P*=.01), again adjusting for age and baseline values. Participants in the intervention group had higher mean EQ-5D-5L scores than participants in the control group (at 12 weeks; 0.72, 95% CI 0.67-0.78 vs 0.62, 95% CI 0.57-0.78 in the control group; *P*=.01).

Table 3 shows that baseline mean protein intake was above the reference nutrient intake of 0.75 grams per kilogram of body weight per day [43,44]. There was no difference in energy (kilocalories per day) or protein (grams per kilogram of body weight per day) intake at the follow-up assessment between participants in the intervention and control groups when adjusting for baseline values (Table 3).

Table . Physical function and nutrition outcomes at baseline and 12 weeks by treatment group in community-dwelling older adults (N=31).

Parameter	Baseline, mean (SD)	Week 12, mean (SD)	*P* value
Grip strength (kg)			.77
Intervention	20.19 (11.95)	20.31 (10.07)	
Control	18.33 (6.34)	18.47 (6.39)	
5-STS[a] (s)			.27
Intervention	26.08 (16.98)	24.27 (17.10)	
Control	29.82 (17.07)	30.53 (16.46)	
Energy intake (kcal per d)			.80
Intervention	1459 (468)	1415 (455)	
Control	1548 (468)	1432 (227)	
Protein intake (grams per kilogram of body weight per d)			.87
Intervention	1.00 (0.29)	1.01 (0.32)	
Control	0.96 (0.33)	0.97 (0.24)	
EQ-5D-5L (−0.59 to 1)			.01
Intervention	0.68 (0.27)	0.74 (0.21)	
Control	0.64 (0.22)	0.60 (0.22)	
VAS[b] (0-100)			.61
Intervention	81.10 (18.53)	79.94 (13.96)	
Control	77.73 (17.63)	75.33 (15.41)	

[a]5-STS: 5-time sit-to-stand exercise.

[b]VAS: visual analog scale.

XSL•FO
RenderX

Discussion

Principal Findings

This 12-week intervention informed by the Medical Research Council guidelines for the development and evaluation of a complex intervention [45] in older adults living in the community in Greater Manchester was feasible and acceptable. Participants reported high usability of KOKU-Nut and found the digital platform acceptable and easy to use, as demonstrated by the SUS [35] and TAM [34]. The current study had retention rates above 90%, and framework analysis of the interviews demonstrated that the study design, assessments used, and allocated treatment groups were acceptable to participants.

This study managed to recruit an at-risk population with low attrition at 12 weeks (3/31, 9.7%). Participants had a mean age of 81.3 (SD 8.3) years, ranging from 70 to 96 years. At baseline, more than half (10/16, 51.6%) of the sample had fallen at least once in the previous 12 months, 16.1% (5/31) were at risk of malnutrition, and most of the participants (median score 4, IQR:3,5) were considered vulnerable according to the Clinical Frailty Scale.

The KOKU-Nut intervention had no impact on energy or protein intake and physical function outcomes at 12 weeks; however, this was a feasibility study with a small sample size and, thus, not powered to detect statistical differences. In addition, interpretation of these findings is further limited by the baseline characteristics of the sample, who were well nourished with adequate protein intake of approximately 1 gram per kilogram of body weight per day at baseline [46]. This introduces a potential ceiling effect whereby participants were already meeting dietary requirements and limits the scope for measurable improvement. Future trials should consider additional eligibility criteria to recruit individuals with suboptimal dietary intake at baseline to enhance sensitivity to change and increase the likelihood of detecting a meaningful intervention effect. Increasing protein intake can support the anabolic effects of strength training and minimize age-related muscle loss, and this has previously been demonstrated in older adults [47], including those with frailty [48] and sarcopenia [49]. However, these dietary interventions did not use a digital platform, and very few apps simultaneously improve dietary intake and support strength training in this population [50]. The study by van den Helder et al [51] demonstrates that combining a home-based exercise program delivered via a tablet-based app with dietary protein counseling delivered via face-to-face dietitian sessions can effectively increase protein intake in older adults. Additional work should focus on the development of a digital tool to support dietary intake among older adults. Additional educational games, high-protein recipes, and goal-setting features are likely to support users in changing their behavior, promote protein intake, and increase the effectiveness of KOKU-Nut.

Most participants in the intervention group (10/14, 71.4%) self-reported using KOKU-Nut 3 or more times a week, demonstrating a satisfactory level of adherence to the intervention. This indicates that older adults can regularly and independently engage with a digital platform such as KOKU-Nut even without prior experience of using an iPad or tablet. A feasibility study with 25 older adults found that 80% of participants regularly used a tablet-based self-monitoring nutrition app (Appetitus) [52]; this complements findings from the current study demonstrating the potential to engage older adults with digital nutrition tools, but further research is required on long-term use and effectiveness. Despite high levels of engagement and usability in the current study, future digital interventions may benefit from increased involvement of caregivers or family members, particularly for those with lower digital literacy, to support confidence, engagement, and uptake.

Unlike the interactive KOKU-Nut intervention, the leaflet-based control condition provided a relatively passive form of intervention delivery. Leaflets are often provided to this population for health guidance and information and, thus, reflect treatment as usual; however, an active control condition may provide greater insight into the effectiveness of the KOKU-Nut intervention. Moreover, qualitative findings indicated perceived behavior change in participants across both trial arms, suggesting that study participation itself may have influenced diet and lifestyle behaviors. This is likely influenced by self-monitoring effects associated with completing assessments and heightened motivation due to involvement in a research study.

This study had many strengths. The approach aligns with the Medical Research Council framework for the development and evaluation of complex interventions [45]. The mixed methods design provides a more comprehensive and nuanced understanding of the feasibility and delivery of the study compared to either methodology alone. Attrition bias was low given that this study had a 90.3% (28/31) retention rate, with only 3 participants dropping out during the 12 weeks. This demonstrates that it was possible to recruit and retain a group of older adults who are at a high risk of falls and often not included in research. Moreover, KOKU-Nut is a minimally invasive digital intervention, so it would be suitable for wide-scale implementation, and a small effect in a large population has the potential to have a big overall effect with low cost or risk.

There are also several limitations that must be acknowledged. Participants were unaware of their allocation until after completing baseline assessments; however, randomization had already been performed prior to the assessment. The lack of allocation concealment may introduce selection bias. This study was not blinded due to resource constraints that, again, may introduce detection bias. This study could have compared the use of KOKU-Nut with that of a generic health app that was not designed for this population to blind the intervention. However, due to limited access to iPads, the control group received a leaflet explaining the benefits of healthy lifestyle habits, and this is likely to better represent standard practice or usual care. The nutrition game was embedded into the Keep-on-Keep-up app, and therefore, it was not possible to specifically look at engagement with the nutrition component or identify whether self-reported use and improvements in quality of life were related to the nutrition game or a different part of the digital platform. Engagement with the digital platform was self-reported, so future studies should aim to develop a dashboard or use Google Analytics data to capture use time

XSL•FO
RenderX

objectively. Methodologies to inform the best practice to conduct a gold-standard clinical trial for a digital intervention should be prioritized. Although many assessments were objective, dietary recall is often subject to recall and response bias from underreporting and poor memory. Dietary assessment is prone to misreporting of portion sizes and foods for social approval [53-55]. Moreover, the qualitative work highlighted challenges with completing the 24-hour recall, including burden, difficulties entering complex or homemade foods, and a tendency to simplify intake to facilitate recording. These issues suggest that dietary assessment tools may require greater flexibility and user support to reduce respondent burden and improve data completeness.

Conclusions

Overall, this feasibility randomized controlled trial demonstrates that the study design was appropriate and acceptable to older adults, as indicated by the recruitment and retention rates as well as the themes that emerged from the interviews. This is promising and demonstrates the potential for conducting a larger study in the future that will be powered to detect effectiveness.

KOKU-Nut had good usability, acceptability, and engagement according to the SUS, TAM, and self-report use, indicating that the intervention was designed appropriately to engage this population. However, the exploratory outcomes exhibited no difference in physical function or dietary intake. This demonstrates that further developments should be made to KOKU-Nut to support dietary change before conducting a larger study to see whether KOKU-Nut can lead to clinical and meaningful impact.

Acknowledgments

Generative AI was not used for the production of this manuscript, including writing, refining, and editing.

Funding

CF received funding from the Medical Research Council Co-operative Awards in Science and Engineering (MRC CASE entry MR/R015767/1).

Conflicts of Interest

ES is the director of Keep-on-Keep-up Health, which is a nonprofit company. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Screenshots, study assessment schedule, feasibility progression criteria findings, and technology acceptance model results for the Keep-on-Keep-up Nutrition trial
[DOCX File, 326 KB - aging_v9i1e87351_app1.docx]

Checklist 1
CONSORT checklist.
[PDF File, 269 KB - aging_v9i1e87351_app2.pdf]

References

1. Living longer: how our population is changing and why it matters. : Office for National Statistics; 2018 URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/livinglongerhowourpopulationischangingandwhyitmatters/2018-08-13 [accessed 2026-08-20]
2. Road to recovery: the government's 2025 mandate to NHS England. Department of Health & Social Care, United Kingdom Government. 2025. URL: https://www.gov.uk/government/publications/road-to-recovery-the-governments-2025-mandate-to-nhs-england/road-to-recovery-the-governments-2025-mandate-to-nhs-england [accessed 2026-08-20]
3. O'Keeffe M, Kelly M, O'Herlihy E, et al. Potentially modifiable determinants of malnutrition in older adults: a systematic review. Clin Nutr 2019 Dec;38(6):2477-2498. [doi: 10.1016/j.clnu.2018.12.007] [Medline: 30685297]
4. Volkert D, Beck AM, Cederholm T, et al. ESPEN guideline on clinical nutrition and hydration in geriatrics. Clin Nutr 2019 Feb;38(1):10-47. [doi: 10.1016/j.clnu.2018.05.024] [Medline: 30005900]
5. Norman K, Haß U, Pirlich M. Malnutrition in older adults-recent advances and remaining challenges. Nutrients 2021 Aug 12;13(8):2764. [doi: 10.3390/nu13082764] [Medline: 34444924]
6. French C, Burden S, Tang Y, Stanmore E. A person-centred approach to further develop a digital tool (KOKU-Nut) by developing a nutrition game for older adults living in the community. BMC Geriatr 2025 Nov 18;25(1):925. [doi: 10.1186/s12877-025-06434-2] [Medline: 41254498]
7. Visser M. Undernutrition in community-dwelling older adults: prevalence, causes and consequences. Proc Nutr Soc 2025 Sep 15:1-7. [doi: 10.1017/S0029665125101808] [Medline: 40947981]

XSL•FO
RenderX

8. Kiefte-de Jong JC, Mathers JC, Franco OH. Nutrition and healthy ageing: the key ingredients. Proc Nutr Soc 2014 May;73(2):249-259. [doi: 10.1017/S0029665113003881] [Medline: 24503212]
9. Dorrington N, Fallaize R, Hobbs DA, Weech M, Lovegrove JA. A review of nutritional requirements of adults aged ≥65 years in the UK. J Nutr 2020 Sep 1;150(9):2245-2256. [doi: 10.1093/jn/nxaa153] [Medline: 32510125]
10. Yeung SS, Kwan M, Woo J. Healthy diet for healthy aging. Nutrients 2021 Nov 29;13(12):4310. [doi: 10.3390/nu13124310] [Medline: 34959862]
11. Young C, Argáez C. Interventions for malnutrition in older adults: a review of clinical effectiveness. : Canadian Agency for Drugs and Technologies in Health; 2019 URL: https://www.ncbi.nlm.nih.gov/books/NBK541429/ [accessed 2026-08-21]
12. Gittins M, AlMohaisen N, Todd C, Lal S, Burden S. Overall mortality for community-dwelling adults over 50 years at risk of malnutrition. J Cachexia Sarcopenia Muscle 2024 Dec;15(6):2509-2518. [doi: 10.1002/jcsm.13585] [Medline: 39210532]
13. Elia M. The cost of malnutrition in England and potential cost savings from nutritional interventions. : National Institute for Health Research Southampton Biomedical Research Centre and BAPEN; 2015 URL: https://www.bapen.org.uk/pdfs/economic-report-full.pdf [accessed 2026-08-20]
14. Bamford C, Heaven B, May C, Moynihan P. Implementing nutrition guidelines for older people in residential care homes: a qualitative study using Normalization Process Theory. Implement Sci 2012 Oct 30;7:106. [doi: 10.1186/1748-5908-7-106] [Medline: 23110857]
15. Jones DJ, Baldwin C, Lal S, et al. Priority setting for adult malnutrition and nutritional screening in healthcare: a James Lind Alliance. J Hum Nutr Diet 2020 Apr;33(2):274-283. [doi: 10.1111/jhn.12722] [Medline: 31858685]
16. Jones A, Hedges-Chou J, Bates J, Loyola M, Lear SA, Jarvis-Selinger S. Home telehealth for chronic disease management: selected findings of a narrative synthesis. Telemed J E Health 2014 Apr;20(4):346-380. [doi: 10.1089/tmj.2013.0249] [Medline: 24684478]
17. The Topol Review: preparing the healthcare workforce to deliver the digital future. : National Health Service; 2019 URL: https://topol.digitalacademy.nhs.uk/the-topol-review [accessed 2026-08-20]
18. Choi NG, Stanmore E, Caamano J, Vences K, Gell NM. A feasibility study of multi-component fall prevention for homebound older adults facilitated by lay coaches and using a tablet-based, gamified exercise application. J Appl Gerontol 2021 Nov;40(11):1483-1491. [doi: 10.1177/0733464821991024] [Medline: 33541199]
19. Thomas S, Mackintosh S, Halbert J. Does the "Otago exercise programme" reduce mortality and falls in older adults?: A systematic review and meta-analysis. Age Ageing 2010 Nov;39(6):681-687. [doi: 10.1093/ageing/afq102] [Medline: 20817938]
20. Yeung PY, Chan W, Woo J. A community-based falls management exercise programme (FaME) improves balance, walking speed and reduced fear of falling. Prim Health Care Res Dev 2015 Apr;16(2):138-146. [doi: 10.1017/S1463423614000024] [Medline: 24480023]
21. Campbell A, Robertson MC. Otago Exercise Programme to Prevent Falls in Older Adults: Accident Compensation Corporation (ACC); 2007.
22. French C, Burden S, Stanmore E. Digital intervention (Keep-On-Keep-Up Nutrition) to improve nutrition in older adults: protocol for a feasibility randomized controlled trial. JMIR Res Protoc 2024 Apr 30;13:e50922. [doi: 10.2196/50922] [Medline: 38687981]
23. Dedeyne L, Deschodt M, Verschueren S, Tournoy J, Gielen E. Effects of multi-domain interventions in (pre)frail elderly on frailty, functional, and cognitive status: a systematic review. Clin Interv Aging 2017;12:873-896. [doi: 10.2147/CIA.S130794] [Medline: 28579766]
24. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
25. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
26. Simple randomisation service. Sealed Envelope. 2024. URL: https://www.sealedenvelope.com/simple-randomiser/v1 [accessed 2026-08-21]
27. Healthy living: looking after yourself physically and mentally. Age UK. 2026. URL: https://www.ageuk.org.uk/globalassets/age-uk/documents/information-guides/ageukig24_healthy_living_inf.pdf [accessed 2026-08-21]
28. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
29. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. Health Econ 2018 Jan;27(1):7-22. [doi: 10.1002/hec.3564] [Medline: 28833869]
30. Elia M. The "MUST" report: nutritional screening for adults: a multidisciplinary responsibility. : Malnutrition Advisory Group; 2003 URL: https://www.bapen.org.uk/pdfs/must/must_exec_sum.pdf [accessed 2026-08-21]
31. Hoyl MT, Alessi CA, Harker JO, et al. Development and testing of a five-item version of the Geriatric Depression Scale. J Am Geriatr Soc 1999 Jul;47(7):873-878. [doi: 10.1111/j.1532-5415.1999.tb03848.x] [Medline: 10404935]
32. Bradley J, Simpson E, Poliakov I, et al. Comparison of INTAKE24 (an online 24-h dietary recall tool) with interviewer-led 24-h recall in 11-24 year-old. Nutrients 2016 Jun 9;8(6):358. [doi: 10.3390/nu8060358] [Medline: 27294952]

XSL•FO
RenderX

33. Intake24. URL: https://intake24.org [accessed 2026-08-21]
34. Masrom M. Technology acceptance model and e-learning. Presented at: 12th International Conference on Education, Sultan Hassanal Bolkiah Institute of Education; May 21-24, 2024 URL: https://www.researchgate.net/publication/228851659_Technology_acceptance_model_and_E-learning [accessed 2026-08-26]
35. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry: CRC Press; 1996:189-194. [doi: 10.1201/9781498710411]
36. NVivo. Lumivero. URL: https://lumivero.com/products/nvivo/ [accessed 2026-08-21]
37. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG, editors. Analyzing Qualitative Data: Routledge; 1994:173-194. [doi: 10.4324/9780203413081]
38. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
39. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
40. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text] [doi: 10.5555/2835587.2835589]
41. King WR, He J. A meta-analysis of the technology acceptance model. Inf Manag 2006;43(6):740-755. [doi: 10.1016/j.im.2006.05.003]
42. McNamara S, Schneider PP, Love-Koh J, Doran T, Gutacker N. Quality-adjusted life expectancy norms for the English population. Value Health 2023 Feb;26(2):163-169. [doi: 10.1016/j.jval.2022.07.005] [Medline: 35965226]
43. Dietary Reference Values for Food Energy and Nutrients for the United Kingdom: Report of the Panel on Dietary Reference Values of the Committee on Medical Aspects of Food Policy: H.M. Stationery Office; 1991.
44. Government dietary recommendations: government recommendations for energy and nutrients for males and females aged 1 – 18 years and 19+ years. : Public Health England; 2016 URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/618167/government_dietary_recommendations.pdf/ [accessed 2026-08-21]
45. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
46. Volkert D, Beck AM, Cederholm T, et al. ESPEN practical guideline: clinical nutrition and hydration in geriatrics. Clin Nutr 2022 Apr;41(4):958-989. [doi: 10.1016/j.clnu.2022.01.024] [Medline: 35306388]
47. Campbell WW, Leidy HJ. Dietary protein and resistance training effects on muscle and body composition in older persons. J Am Coll Nutr 2007 Dec;26(6):696S-703S. [doi: 10.1080/07315724.2007.10719650] [Medline: 18187436]
48. Tieland M, Dirks ML, van der Zwaluw N, et al. Protein supplementation increases muscle mass gain during prolonged resistance-type exercise training in frail elderly people: a randomized, double-blind, placebo-controlled trial. J Am Med Dir Assoc 2012 Oct;13(8):713-719. [doi: 10.1016/j.jamda.2012.05.020] [Medline: 22770932]
49. Kim HK, Suzuki T, Saito K, et al. Effects of exercise and amino acid supplementation on body composition and physical function in community-dwelling elderly Japanese sarcopenic women: a randomized controlled trial. J Am Geriatr Soc 2012 Jan;60(1):16-23. [doi: 10.1111/j.1532-5415.2011.03776.x] [Medline: 22142410]
50. French C, McKenzie Y, Stanmore E, Burden S. Are mobile apps effective at targeting physical activity or malnutrition in community-dwelling older adults? A systematic review and meta-analysis. J Aging Phys Act 2024;33(4):313-325. [doi: 10.1123/japa.2023-0449] [Medline: 39708786]
51. van den Helder J, Mehra S, van Dronkelaar C, et al. Blended home-based exercise and dietary protein in community-dwelling older adults: a cluster randomized controlled trial. J Cachexia Sarcopenia Muscle 2020 Dec;11(6):1590-1602. [doi: 10.1002/jcsm.12634] [Medline: 33103379]
52. Aure CF, Kluge A, Moen A. Older adults' engagement in technology-mediated self-monitoring of diet: a mixed-method study. J Nurs Scholarsh 2021 Jan;53(1):25-34. [doi: 10.1111/jnu.12619] [Medline: 33316147]
53. Miles S, Scaife V. Optimistic bias and food. Nutr Res Rev 2003 Jun;16(1):3-19. [doi: 10.1079/NRR200249] [Medline: 19079933]
54. Hebert JR, Ma Y, Clemow L, et al. Gender differences in social desirability and social approval bias in dietary self-report. Am J Epidemiol 1997 Dec 15;146(12):1046-1055. [doi: 10.1093/oxfordjournals.aje.a009233] [Medline: 9420529]
55. Macdiarmid J, Blundell J. Assessing dietary intake: who, what and why of under-reporting. Nutr Res Rev 1998 Dec;11(2):231-253. [doi: 10.1079/NRR19980017] [Medline: 19094249]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
KOKU-Nut: Keep-on-Keep-up Nutrition
SUS: System Usability Scale
TAM: technology acceptance model

XSL•FO
RenderX

Edited by H LaMonica; submitted 07.Nov.2025; peer-reviewed by A Bailer, YY Kristian; revised version received 07.May.2026; accepted 16.Jul.2026; published 28.Aug.2026.

Please cite as:
French C, Burden S, Gangannagaripalli J, Stanmore E
Assessing a Digital Platform (KOKU-Nut) to Improve Nutrition in Older Adults: Mixed Methods Feasibility Randomized Controlled Trial
JMIR Aging 2026;9:e87351
URL: https://aging.jmir.org/2026/1/e87351
doi:10.2196/87351

© Chloe French, Sorrel Burden, Jaheeda Gangannagaripalli, Emma Stanmore. Originally published in JMIR Aging (https://aging.jmir.org), 28.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Developing Digital Platforms to Help Older Adults Stay Physically and Socially Active: Co-Design Study to Develop a Toolkit and Evaluate Its Impact

Leonie Cooper[1], PhD; Hannah Bradwell[1], PhD; Colin B Shore[2], PhD; Simone A Tomaz[3], PhD; Rory Baxter[1], PhD; Gemma C Ryde[4], PhD; Swen E Gaudl[5], PhD; Richard Haynes[6], PhD; Catherine Hennessy[7], PhD; Ray B Jones[1], PhD; Anna C Whittaker[3]; The GOALD Project[8,9]

[1]Faculty of Health, University of Plymouth, Plymouth, England, United Kingdom
[2]School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, United Kingdom
[3]Faculty of Health Sciences and Sport, University of Stirling, Stirling, Scotland, United Kingdom
[4]School of Cardiovascular and Metabolic Health, University of Glasgow, Glasgow, Scotland, United Kingdom
[5]Department of Applied IT, University of Gothenburg, Gothenburg, Västra Götaland, Sweden
[6]Faculty of Arts and Humanities, University of Stirling, Scotland, United Kingdom
[7]Faculty of Social Sciences, University of Stirling, Scotland, United Kingdom
[8]See Acknowlegments
[9]

Corresponding Author:
Anna C Whittaker
Faculty of Health Sciences and Sport, University of Stirling, Stirling, Scotland, United Kingdom

Abstract

Background: Digital technologies have the potential to support physical, cognitive, and social activity among older adults, but many small- and medium-sized enterprises (SMEs) lack the resources to conduct meaningful co-design with older adults. Toolkits derived from rigorous co-design processes may offer a scalable mechanism for translating end-user priorities into real-world product development.

Objective: The Generating Older Active Lives Digitally study aimed primarily to engage older adults in a co-design process to identify priorities for digital technologies that support physical activity and reminiscence. Secondary objectives were to (1) translate these findings into a practical developer-facing toolkit, and (2) evaluate the toolkit's perceived utility and influence among digital technology SMEs.

Methods: A total of 157 participants (121 older, 7 younger, and 29 staff) across 15 community and care settings in England and Scotland engaged in 106 technology interaction sessions, 22 evaluation focus groups, and 10 co-design workshops. These involved more than 20 digital technologies. Thematic analysis and structured card-ranking tasks were used to derive their priorities. Preliminary toolkits were co-designed and used by 10 UK-based SMEs who received small grants to apply the toolkit to active development projects. Their reports and follow-up interviews were analyzed thematically to identify perceived impacts on design decisions, product adaptations, and business outcomes.

Results: Co-design activities generated 7 cross-cutting themes: motivation, content, barriers, design and inclusivity, suitability, acceptability, and motivations to use. These were organized into 3 toolkit sections: general design principles, online physical-activity platforms, and virtual reality. SMEs reported that the toolkits enhanced understanding of older adults' needs, validated design decisions, and inspired some new features. Perceived impacts included improved usability, expanded accessibility options, increased content variety, clearer instructional design, enhanced social components, and reduced operational costs. SMEs also reported some business benefits, including strengthened cases for investment and increased product uptake.

Conclusions: Co-designed toolkits may offer a useful and cost-effective mechanism for translating older adults' priorities into digital product development. SMEs perceived this toolkit as practical, relevant, and impactful for informing design choices. This approach may complement, but does not replace, direct user involvement. It may help accelerate inclusive digital-health innovation for aging populations.

(JMIR Aging 2026;9:e89785) doi:10.2196/89785

KEYWORDS

aging; co-design; digital health; older adults; physical activity; toolkits; user-centered design; virtual reality

XSL•FO
RenderX

Introduction

Background

It is expected that by 2030, a total of 1 in 6 people in the world will be aged older than 60 years, and by 2050 the world's older population will have doubled [1]. This demographic shift is placing increasing pressure on health and social care systems, leading to a growing emphasis on promoting healthy aging across the life course [2]. A key focus within this agenda is promoting physical activity within older populations given the role it can play in maintaining and improving physical function [3]. Another area of focus is that of building older adults' social connectedness and well-being through reminiscence, which has been shown to benefit well-being [4]. Sports-based reminiscence has been particularly successful in men with cognitive impairment, providing, among other mechanisms, a source of specific interest on which to hinge opportunities to increase social participation [5].

Digital technologies have been proposed as a means of supporting physical, cognitive, and social activity in later life [6]. National policies, such as the UK government Consensus on Healthy Aging [7], promote physical, cognitive, and social activity among older adults through exercise, games, reminiscence, and social connection, with digital tools playing a key role. In response, a wide range of digital technologies has emerged, including online exercise platforms, mobile apps, exergames, and immersive virtual reality (VR) experiences. Recent reviews [8,9] covered 16 - 22 studies on technologies to support physical activity, while another study [10] identified 13 apps and 24 websites for promoting strength and balance exercises. Positive effects of online exercise platforms or programs designed for older adults have been found for body strength, flexibility, and balance [11,12]. More recently, VR tools designed to support physical activity and reminiscence in older adults have shown promising benefits. These include improved physical rehabilitation at home [13], positive influences on cognition, memory, and depression [14], and enhanced cognitive and social engagement through reminiscence experiences [11,15]. Despite these advances, older adults often experience challenges when adopting and sustaining engagement with digital technologies. Many report positive attitudes toward the potential usefulness of technology to increase their physical and social activity [16]. However, barriers related to usability, accessibility, confidence, cost, security, and perceived relevance frequently result in poor adoption or disengagement [17]. Evidence suggests that when technologies are well-designed, have clear training or guidance, have adaptations specifically for older adults, highlight the benefits, and focus on an older person's values, conditions, and reducing age-related risks, they can successfully adopt and make use of them [16,18]. However, research has identified the disconnect between what a technology developer may believe an older adult would like or need from the technology and what older adults prioritize [19]. They need to understand the needs and preferences of older adults in using digital products [20-22].

Co-design is widely regarded as a gold-standard approach for addressing this disconnect in digital development. It is the practice of people collaborating and connecting their knowledge, skills, and resources to carry out a design task and to produce new knowledge as people experiment with new ideas [23]. It differs fundamentally from user surveys and market research in that it positions users as active partners in the design process rather than passive respondents. Surveys and market research are valuable for identifying preferences, patterns, and trends, but are typically developer-led and often constrained by predefined questions. In contrast, co-design engages users in generative activities such as ideation, critique, and prototyping. This enables the incorporation of lived experience and tacit knowledge into intervention development. This participatory approach has been associated with improved relevance, acceptability, and implementation of health interventions (eg, other studies [24,25]). As a result, co-design supports successful and relevant design [26] and business growth [27]. However, meaningful co-design can be resource-intensive, requiring time, expertise, and sustained engagement that may be beyond the capacity of small and medium-sized enterprises (SMEs). Further, evidence suggests that poorly implemented co-design offers little advantage over standard design practices [28]. Consequently, alternative mechanisms are needed to support SMEs in integrating end-user insights without requiring them to conduct extensive co-design activities independently.

Despite growing evidence that digital technologies can support physical activity, reminiscence, and social connection among older adults, there remains a gap between older adults' lived experiences and the design priorities of digital technology developers, particularly SMEs. While participatory co-design can generate rich, contextual insights, resource constraints limit SMEs' ability to engage in sustained co-design processes. Consequently, there is a need for mechanisms that can translate rigorous co-design insights into forms that are accessible, practical, and usable within real-world development contexts, alongside empirical evaluation of whether such mechanisms meaningfully support development practice. This study seeks to address this gap.

One mechanism is the use of toolkits. In this context, a toolkit refers to a structured collection of guidance, recommendations, and examples intended to support developers in creating accessible, inclusive, and user-centered digital technologies. Toolkits offer a potential means of translating insights from rigorous co-design research into practical resources that can be applied within real-world development workflows. However, concerns have been raised that toolkits risk becoming tokenistic if they are not grounded in robust evidence or evaluated in applied settings. A review of 83 health-related toolkits found that only 31 (37%) had undergone any form of evaluation, and those that had were often assessed in artificial contexts with designers present, introducing potential bias [29]. The present study sought to address this.

Primary Objective

The primary objective of the Generating Older Active Lives Digitally (GOALD) project was to work with a diverse range of older adults to better understand their preferences and requirements in digital technologies aimed at increasing physical activity.

XSL•FO
RenderX

Secondary Objectives

Secondary objectives were (1) to use the recommendations emerging from working with older adults to co-design a toolkit for SMEs with a focus on digital physical activity, and (2) to evaluate the perceived usefulness of the toolkit with SMEs.

The primary contribution of this study is the development of a co-design–informed toolkit for SMEs, with the participatory process and implementation evaluation presented to support understanding of its development and perceived utility.

Methods

Ethical Considerations

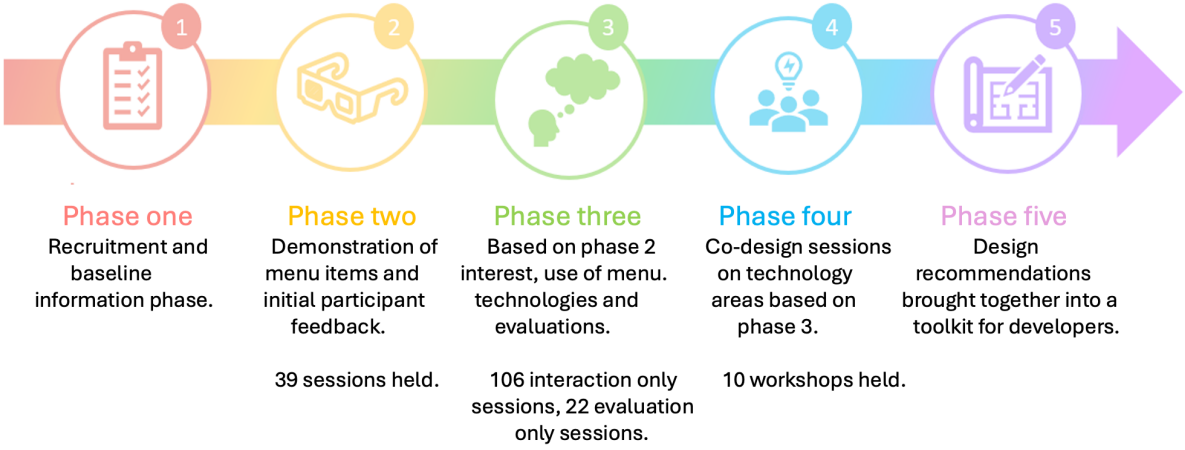
This study received ethical approval from the Faculty of Health Ethics Committee at the University of Plymouth (project ID 3034, on November 29, 2021) and from the University of Stirling General University Ethics Committee (ref: 2021 2080 2106 on October 21, 2021). All participants provided written informed consent.

Study Design

The GOALD project, a collaborative project with the aim of exploring how digital technologies can support older physically and socially active lives, was based in Plymouth (Southwest England) and Stirling (Central Scotland). Older adults were recruited via partner organizations at both sites, plus an advisory group to support the methodology. This study was split into a series of phases, as outlined in Figure 1. Phases 1 - 4 involved participatory in co-design aligning with our primary objective; co-design workshops were held between June 2022 and November 2023. Phase 5 consisted of toolkit development followed by a qualitative implementation evaluation of the toolkit with SMEs, corresponding to our two secondary objectives.

Table 1 presents an overview of the purpose of each phase and details of its impact on the final toolkit.

Figure 1. Flow diagram demonstrating the breakdown of phases of the GOALD (Generating Older Active Lives Digitally) project.



XSL•FO
RenderX

Table . Purpose, number of participants, data collected, analysis, and impact on toolkit formation outlined for each phase and by site.

Phase	Purpose	Participants and sessions	Data collected	Analysis	Output or role in toolkit
Phase 1	Exploratory engagement and recruitment	213 participants recruited across the two sites	N/A[a]	N/A	Informed recruitment and initial engagement
Phase 2	Exploratory engagement and familiarization with digital technologies	England: 14 sessions, 40 OAs[b], 4 YAs[c], 9 staff; Scotland: 14 sessions, 42 OAs, 0 YAs, 11 staff	Informal feedback, baseline discussions	Not formally analyzed	Exploratory: informed selection of technologies for evaluation
Phase 3	Evaluate experiences of interacting with selected technologies	England: 27 technology interaction sessions, 22 formal evaluation sessions 60 OAs, 5 YAs, 12 staff; Scotland: 50 technology interaction sessions (with evaluation[d]), 60 OAs, 2 YAs, 18 staff[d]	Field notes, audio recordings of interactions, focus group and interview transcripts of evaluations	Reflexive thematic analysis	Direct: identified user needs, barriers, and priorities (see the study by Tomaz et al [30])
Phase 4	Co-design and prioritization of features	England: 8 sessions 28 OAs, 1 YAs, 3 staff; Scotland: 4 sessions, 18 OAs, 2 YAs, 6 staff	Ranking cards, workshop artifacts, audio recordings	Thematic synthesis + descriptive ranking	Direct: generated prioritized recommendations for toolkit, additional needs identified
Phase 5	Apply and evaluate toolkit with SMEs[e]	SMEs (n=10)	Reports, interviews	Reflexive thematic analysis	Evaluation: perceived utility and impact of toolkit. Recommended changes from SMEs implemented to produce final toolkit

[a]N/A: not applicable.

[b]OA: older adult.

[c]YA: younger adult.

[d]In Scotland, the care homes needed or wanted a lot more support than community groups or homes or groups in England with some of the more complex technologies, so some sessions in phases 1, 2, and 3 are counted twice as they included an introduction and interaction, or an interaction with a formal evaluation.

[e]SME: small- and medium-sized enterprise.

Participants

Recruitment took place through direct invitation and facilitation from our partners (Active Stirling, Sports Heritage Scotland [now Memories Scotland], iSight Cornwall, Hearing Loss Cornwall, Devon and Cornwall South Asian Society, Nudge, Cornwall Rural Community Charity, Generations Working Together, and UK Active and Sporting Heritage) and local care homes. Consequently, 213 participants reflecting a diversity of age, disability, ethnicity, and location from 18 case sites initially consented to this project; participants are characterized in detail in Table 1 and elsewhere [30]. In brief, older participants' mean age was 78.9 (SD 14.0), but the sample also included care home staff and young people aged 16 - 25 years. Of the participating older adults who provided data, 63% (55/87) were living independently; 75% (65/87) reported a health condition that limited them physically, mentally, or socially, and most reported using some type of corrective or assistive aid—most commonly glasses, hearing aids, or mobility aids; and participants were based in urban and rural areas across Central Scotland and Southwest England. The overall GOALD project examined both physical activity and sports reminiscence, and some participants only engaged in one of these components.

This paper reports data collected from participants who engaged in at least one session of physical activity in phases 2 - 4, resulting in 157 participants from 15 case sites across Central Scotland and Southwest England. Older adult participants were included if they had the capacity to consent, understood English, were aged 65+ years, and were living in a residential care home or attending a community group. Other participants were included if they were staff assisting the older adults in the co-design workshops due to being care staff or running the community group, or if they were members of the research team. Younger people were recruited through the community groups or from local high schools engaging in sport leadership training and intergenerational work (which was a broader goal of the GOALD project reported elsewhere [31]). However, the older adults were the main source of co-design insights, and we report the participant breakdown below simply for transparency as to who was present during the sessions. There were 80 Central Scotland participants and 77 Southwest of England participants.

XSL•FO
RenderX

A total of 56 older adults were from 7 care homes and 65 older adults from 8 independent living community groups. From the staff recruited, 24 of them were care home staff with varying roles and 5 of them were facilitators from community groups. Due to recruitment difficulties, only 7 younger adults from the community were recruited in total (Table 1). Details of sites are available in Multimedia Appendix 1.

Selection of Technology

Initial technologies were identified through the research team's existing academic, clinical, and industry networks to enable rapid access to a diverse range of digital tools relevant to physical activity, reminiscence, and social connection in later life. As this study progressed, additional technologies were purposively identified and introduced. This was in response to older adults' feedback and expressed interests, to ensure that subsequent evaluations and co-design activities aligned with participant priorities. Collectively, the technologies spanned multiple domains and modalities, including online physical activity platforms, immersive VR experiences, and other interactive digital tools. The technologies included in this study were not intended to represent an exhaustive or systematically sampled set of digital interventions.

As a result, older adults could try up to 20 technologies at different stages of development aimed at increasing physical activity and reminiscence. These included online physical activity platforms, VR experiences, social connectivity apps, voice technologies, camera glasses, activity trackers, balance boards, tablet-based aqua-aerobics, and social robots. Due to geographical constraints, not all technologies were available to all participants. Table 2 presents the technologies selected by participants for interaction during this project, which informed the development of the GOALD toolkit.

XSL•FO
RenderX

Table . Technologies participants interacted with throughout this project.

Name	Descriptions
Digital heritage bundle	VR[a] experience created within the University of Plymouth, allowing users to explore heritage remotely. Users are immersed in both audio and visual media, with 3D printed models of artifacts available alongside the VR.
MOTUS VR experiences	Three devices: 1. Standing omnidirectional treadmill with VR headset that allows users to explore both 3D scanned heritage sights and computer-generated environments. 2. Seated treadmill where they can explore computer-generated environments, with incentives to encourage movement. 3. A static seated experience, where 360 videos are uploaded into the headset, allowing users to be immersed in footage. As part of the treadmill experiences, remote users can be connected within virtual environments and explore together regardless of distance.
Oceans Conservation Trust experience	A VR experience created by the National Marine Aquarium (Plymouth, United Kingdom) Ocean Conservation Trust. The experience includes 360-degree videos taken from dives within the aquarium's tanks. Marine artifacts are also included to provide tactile and educational components. This experience was delivered by the Ocean Conservation Trust team.
Oculus Store apps[b]	National Geographic Explore VR Oculus app[b]: this allows physical movements, such as kayaking and walking, to explore hard-to-reach places worldwide. Brink Traveler[b]: users can walk through and learn educational information about locations worldwide.
danceSing/ danceSing Care	Online physical platforms that provide on-demand fitness classes for a range of abilities, including yoga, Pilates, barre, and singing content. There is also a live session component, allowing users to interact with others and instructors.
Mature Movers	Online platform that supports users in finding exercises suited to their needs and provides classes for those aged over 60 years to promote physical activity independence for as long as possible. A mixture of both face-to-face and online classes.
Elder Gym[b]	Online platform that provides exercise classes for older adults, including strength, endurance, balance, and flexibility. Classes are held on a social media platform for live interactions.
Wheel power[b]	Online exercise platform with chair-based exercises suitable for those with reduced mobility or wheelchair users. Live classes, alongside on-demand content.
Goodboost	Waterproof tablet that provides water-based exercises for older adults through an avatar and written instructions.
RetroRehab	Exergaming product created by Goodboost. Online platform with interactive "arcade style" exercise games to encourage movement to improve coordination, balance, and daily activities, for example, sit-to-stand repetitions.
HUR Ltd.	Balance training platform with a display touchscreen that uses the person's movement to control avatars in a range of games to encourage pneumatic strength and balance training.
Body Boosting Bingo[b]	A game of bingo delivered by Age Scotland in which numbers are called, and for each of the possible numbers there is an associated movement to build strength or balance. The movements can be done seated or standing.

[a]VR: virtual reality.

[b]Technologies that were provided in evaluation and co-design workshops only and not the earlier "menu" phase.

Between phases 3 and 4, researchers reviewed participant evaluations and conducted a horizon scan of available technologies based on their requests. Google searches using the phrases "virtual reality to see places for older people" and "online physical activity platforms for older adults" were conducted, with the first 5 pages of results reviewed to identify technologies aligned with participant feedback. Based on availability and location, additional technologies (also shown in Table 2) were provided for participants to trial before phase 4 co-design workshops.

XSL•FO
RenderX

Data Collection

Data were collected across phases 3‐5 using multiple qualitative methods; earlier phases (phases 1‐2) focused on exploratory engagement and familiarization where participants were shown a general menu of technologies related to encouraging physical activity, reminiscence, social connectivity, and access to places and spaces. These phases did not generate formally analyzed data (Table 1).

Regarding phase 3, based on interest expressed in phase 2, researchers (LC, HB, RB, CBS, and SAT) attended community group sessions or visited care homes to support participants in interacting with the technologies. At Plymouth sites, all community participants received logins to access online physical activity platforms at home or during community sessions, and care-home activity leaders received logins to facilitate sessions at their convenience. At Central Scotland sites, sessions were facilitated in-person by researchers; some of these took place in the university laboratory or swimming pool. During phase 3, data included observational field notes recorded by researchers, audio recordings of participant comments, and whiteboard notes captured during group activities. Formal evaluation focus groups and semistructured interviews were conducted in person or online to explore perceived benefits, barriers, usability, and suitability of the technologies. These sessions were audio recorded and transcribed verbatim.

Regarding phase 4, co-design workshops were conducted to identify participant priorities for the technologies. Based on phase 3 findings, 3 co-design themes were addressed: (1) VR for older adults to support physical activity and access to places and spaces; (2) online physical activity platforms for older adults; and (3) general physical activity technologies for older adults. Participants were reminded of the technologies encountered during this project and completed two activities: ranking cards and idea generation.

For the ranking-card activity, researchers synthesized features raised during initial interactions and formal evaluations, categorized them (eg, usability and interaction type), and produced cards for prioritization. Participants ranked cards from low to high priority within assigned categories (one or two per group). Photographs of the ranked cards were taken to generate quantitative priority data. For the idea-generation activity, participants worked in groups to respond to the prompt: “If there were no limitations, what would you want from VR technologies, online physical activity platforms, or physical activity devices to motivate physical activity for older adults or intergenerational groups?” Drawing on Harrington et al [32], participants were encouraged to consider activity type, content, appearance, motivation, and practical considerations. Groups discussed, wrote, and drew ideas using large paper, pens, and sticky notes, then presented them to the wider group. This card-ranking approach was selected because it is widely used in co-design to identify together what is most important and help move beyond discussions to actionable decision-making [33]. It also provides a visible, tangible record of how decisions were reached [34]. During these phase 4 co-design workshops, data comprised audio recordings of group discussions and presentations, written and drawn artifacts produced during idea generation activities, and photographs of ranked cards generated during prioritization exercises. All recordings were transcribed verbatim.

Regarding phase 5, findings from phase 3 interaction and evaluation sessions and phase 4 activities were synthesized into preliminary toolkits to support SMEs designing technologies for older adults, with a focus on physical and cognitive activity. The toolkits were distributed to 22 UK-based software SMEs across 2 calls (July and September 2023). SMEs were invited to submit small projects to apply the toolkit recommendations to improve existing products or develop new ones; 10 SMEs were selected for funding by a panel of GOALD researchers. As a grant condition, funded SMEs provided a short end-of-project report (maximum 6 months) detailing achievements, toolkit use, challenges, observed or anticipated impact, and recommendations for improving the draft toolkits. Following receipt of the reports, researchers (LC, HB, CBS, and John Ritchie, PhD) conducted 30-minute video calls with representatives from the 10 funded SMEs to discuss their feedback in more depth, emphasizing that candid reflections were encouraged given that funding had already been awarded. One interview (Football Memories Scotland) was conducted face-to-face during a researcher visit. Phase 5 data therefore included written end-of-project reports submitted by participating SMEs and audio-recorded follow-up interviews exploring toolkit use, perceived influence on design decisions, and implementation challenges, which were transcribed verbatim.

Data Analysis

Qualitative Analysis

Transcriptions from phases 3 and 4 were analyzed using reflexive thematic analysis following procedures outlined by Braun and Clarke [35]. Researchers (LC, HB, CBS, and SAT) first familiarized themselves with the data through repeated reading of transcripts and field notes. Initial codes were then generated inductively across the dataset, capturing features relevant to older adults’ experiences, preferences, and priorities. Codes were subsequently reviewed and organized into candidate themes, which were refined through iterative discussion within the research team to ensure coherence, internal consistency, and distinction between themes. Themes were then defined and named, and illustrative data extracts were identified to support interpretation. Analyses were initially conducted separately at each site due to differences in data collection timelines. These were then compared across datasets and integrated through team discussion. Where themes were conceptually similar, they were merged into cross-cutting themes. Where differences arose, these were retained and discussed within the research team to preserve contextual variation and ensure that divergent perspectives informed the final interpretation and toolkit recommendations. Analysis was conducted reflexively, with themes treated as interpretative patterns developed through engagement with the data rather than as fixed or exhaustive representations. These themes and codes formed the recommendations and produced 4 preliminary toolkits, which focused on VR for physical activity and accessing places and spaces (Plymouth), online physical activity platforms

XSL•FO
RenderX

(Plymouth), general physical activity platforms including online platforms (Stirling), and a reminiscence-based toolkit (Plymouth initially, updated with input from Stirling data for the second funding call to technology SMEs).

Final reports and interviews from phase 5 challenge-fund software SMEs were thematically analyzed following the same reflexive approach. Researchers (LC, HB, and CBS) familiarized themselves with all texts before independently conducting inductive coding on sections of the data. Texts were then swapped to compare and agree on coding interpretations. Patterns in the codes were identified to create overarching themes, which were reviewed and finalized across all texts. These themes and codes were subsequently merged and refined to produce a single, consolidated toolkit combining data from both sites.

Exploratory Quantitative Analysis

Card-ranking results from phase 4 were calculated after rebalancing scores to account for variation in group size and number of groups per session, and results were banded into tertiles. These 3 bands indicated features ranked as higher or lower priority, with the middle band reflecting neutral importance. Scores were rebalanced by weighting each site's feature score by the number of contributing groups and participants per group, reducing the influence of single participants or small groups. Final feature scores were averaged, excluding zero votes where features were not ranked by a group.

Results

Overview

From the thematic analysis of older adults' data, 7 themes emerged: motivation, content, barriers, design/inclusivity considerations, potential audiences/suitability, acceptability of technology, and motivations to use. Table 3 summarizes the core themes that emerged across the qualitative analysis and provides an overview of the cross-cutting concepts that structure the results presented below.

These themes were separated into 3 sections for the preliminary toolkits, to ensure the information was categorized clearly for SMEs. The 3 sections were general recommendations for technology for older adults, recommendations specific to physical activity platforms (eg, websites, apps, and exergames), and VR recommendations for those working in immersive technologies. Several themes crossed over different subsections; themes and codes (recommendations) are summarized below.

Table . Overview of the major themes that emerged across the qualitative analysis.

Major theme	What it captures
Motivation	Factors encouraging engagement, enjoyment, and sustained use
Barriers	Physical, psychological, and contextual challenges to uptake
Design and inclusivity	Hardware, software, and interaction features enabling accessibility
Content	Preferred types, qualities, and relevance of experiences
Acceptability	Comfort, trust, perceived value, and social acceptability
Suitability and context	Appropriate settings, user groups, and use conditions
Motivations to use	Perceived benefits (eg, well-being, social connection, activity)

General Recommendations for Technology for Older Adults

Themes and related codes that produced the recommendations for general technology for older adults and the evidence (quotes from participants) supporting the development of these themes are shown in Multimedia Appendix 2. Themes within this section included motivation, content, barriers, design, and inclusivity considerations. Recommendations covered suggestions from users on how to make the products motivating to use, from providing social components to incorporating gamification principles. General content suggestions included intensity of exercises, including relatable content, and that having variety and choice were key. Participants discussed barriers such as concerns regarding injury and how accessible the technology was. Finally, the theme of design or inclusivity considerations emerged, which included software considerations, such as using websites or apps, ensuring there are physical adaptations to the technology to make it easier for older adults to use, and seeking to understand what is "easy" for their target audiences to use.

Recommendations Specific to Physical Activity Platforms

Themes and codes that were specific to physical activity platforms and supporting evidence are presented in Multimedia Appendix 3. The main themes that were specific to physical activity platforms included motivation, design or inclusivity considerations, barriers, and potential audiences or suitability. To encourage motivation in users, physical activity platforms should have engaging instructors who are relatable to their audience, such as using older adult instructors. Live sessions were also highlighted as a motivation for users, as these allowed for social interaction components. Barriers included motivating users through a screen, when many preferred going outside or engaging in face-to-face classes, and if there were live classes, a preference was for these to be away from social media platforms. Design and inclusivity considerations included having clear visual follow-along, clear instructions, and having the ability to view on large screens. Potential audiences the older adults felt these platforms would be suited to included care home residents, community groups, people living alongside those less able or independent, and those living without access

XSL•FO
RenderX

to the outdoors. More detail on recommendations specific to physical activity technologies is reported elsewhere [30].

Recommendations Specific to VR Technologies

Of the main themes, 6 were linked to VR-specific recommendations and are presented in Multimedia Appendix 4. Older adults expressed praise for the technology and enjoyed the experiences and felt safe. However, they highlighted the need for consideration of how the technology is introduced to older adults and to consider generational differences with technology use. Barriers included older adults' attitudes to the technology and expectations such as fear of enclosed spaces and the need to consider balance and coordination when using VR technologies, alongside potential mobility challenges that impact older adults. Older adults wanted content not to be repetitive, and to avoid infantilization or child-like games. Content should be smooth, and any movements controlled. They did express concerns around addiction to technology, as well as concerns around confusion, disconnection between senses, and disorientation. Alongside psychological effects, there were concerns of physical effects such as body aches or health and safety concerns around falling or epilepsy. Older adults reported the technology must offer more than what is already available on other technologies.

Design considerations included the hardware used, such as using comfortable headsets but needing to factor in hygiene of the equipment and storage restrictions of locations. Software should include clear tutorials and guidance, have easy navigation between content, have clear and effective audio, simplified setup, and have consistent and reliable technology. Older adults raised having gamification, motivational rewards such as milestones and checkpoints within the virtual environments to encourage continued use. They also wished to be able to interact with the scenery and have movement as an option to explore, such as walking, skiing, seated movements, natural movements, rowing, and more. However, participants reported avoiding using teleportation movement in VR or any movement that results in a disconnect between senses. Participants suggested the ability to project content onto larger screens so that it becomes a group activity for more people to enjoy and facilitate communication.

Content suggestions included nature such as gardens, animals, coastal content, and under the sea. Historical content was also requested, including wonders of the world, archaeology, National Trust sites, museums, and archive footage. A range of activities were raised, such as experience sports, journeys on different modes of transport, shopping, visiting new locations, guided tours, and wanting a mix of passive and active experiences. Alongside this, there was a desire for personal content that could link them to familiar places or past holidays. Overall, older adults wanted longer content, colorful scenery with soothing audio, realistic imagery with details, audio guides, educational content, and most importantly, variety.

Older adults reported various perceived benefits that motivated them to use VR devices. One motivation to use VR was the enhanced experiences it could offer, such as giving access to places that older adults cannot access, whether due to mobility, cost, or mental health challenges, or activities that would otherwise be unsafe. VR offered them the opportunity for new experiences and to see new places, as well as providing an educational angle. As well as new experiences, the devices could be used to escape reality or relive old memories, so it would provide mental and physical stimulation. The physical motivations were the ability for it to facilitate exercise and provide interesting and distracting content while doing so, and identified the ability for it to have therapeutic uses. Older adults expressed the fun and enjoyment they had from using VR, and that it had potential to offer a social experience.

They felt that VR technology that increased physical activity and gave access to places and spaces was suitable for care home residents, people who needed more exercise, those with mobility issues, or those experiencing pain. It was highlighted that the technology could be suited to those who require rehabilitation, or those who are hospitalized or bed bound.

End-User Priorities for Technology SMEs From Rankings

Results from the exploratory ranking analysis in the co-design workshops where participants ranked their priorities for development are presented in Tables 4 and 5. These rankings should be interpreted as indicative outputs of workshop-based prioritization rather than generalizable or statistically robust measures.

XSL•FO
RenderX

Table . End-user priorities ranked for physical activity platforms for older adults.

Priority level	Components
High priority	Follow-along videos, community of users (hosted in the platform), group activity in-person, visual follow-on, variety of content, live sessions, filter function on difficulty level, content for range of abilities, variety of songs to move to, assessment on suitable exercises, voice assistive technology, game scenario—reaching, stretching, moving, armchair activities, warm up and cool down activity, easy to navigate, ability to pause and come back, value for money (affordable), motivation to exercise, easy to use, easy to search, audio instructions (sight impairment), easy initial setup (sight impairment)
Medium priority	Gamified component—collecting badges, easy screen casting options, indoor cycling, varied cultural content, separate from social media platforms, audio description to navigate platform (sight impairment), gradual increase in intensity, individual activity, social forum component, yoga, music content for varied generations, someone to check your exercises, family activity, adjustable pace of sessions, adaptable video and font sizing, older tutors to match audience, big screen or not a mobile app, appropriate length (<45 min), translation options, closed captioning, frequent updates, space needed for activities, clear audio (speaking over music), dancing, group activity, range of video lengths
Low priority	Tutorial to navigate platform, extreme sports (skiing), reminders to take part (notifications), needs to be different from what is already received for free

XSL•FO
RenderX

Table . End user priorities ranked for VR[a] technologies for older adults.

Priority level	Components
High priority	Pedaling, people watching, guided tour, audio from the scenes filmed, different eras, varied content, multisensory (artifacts to hold, breeze to feel, and sand at beach), simple setup, historical information on audio, armchair exercise, information on what you are viewing, relaxing visual content, colorful, bright, interaction with objects, clear instructions on how to use, natural foot movements (walking not shuffling), places with memories, safety features on exercise equipment, comfortable headsets, natural movements, places usually expensive to visit, movement with feet, ability to move around in virtual environment, headset not too heavy, new experiences, not colorless, places no longer physically accessible, realistic, affordable price, detailed visuals (fine details, eg, leaves on trees and petals), cleanable, not simplistic models, high image quality
Medium priority	Rewards (incentives to continue [seeing something new or exciting]), show you locations you may have to go (reduce anxiety hospital appointments and shopping trips), pretravel knowledge (view where you will go), group activity on a big screen displaying VR content, POV[b] adrenaline experiences (zip lining, roller coasters, and horse riding), not cartoonish, physical activity associated with hobbies, that is, gardening, audio, sounds, or music (eg, blue planet), watch animals, sports matches, social element should be optional, seated movements (technology allows seated interaction), passive watching of nature, archeology, inclusive (mobility), related to hobbies, local places, dancing, family interaction, warnings on content (small spaces, heights, sit down, or dizziness), pedals to move around, menu self-orientates in front of you, outdoors, seaside, flowers, heritage sites, movement with hands, haptic feedback technology (gloves), historical events (jubilee or reminiscence), step counter, walking, stable movements (smooth and not teleporting), has to offer more than television, feel control over moving, subtitles (for hearing impairment), headphones for sounds, easy to put on headset, different countries, written cues to sound (for hearing impairment), concerts, interesting content, nature, countryside, educational aspect, paddling or rowing, audio journey or description (for sight impairment)
Low priority	Shopping trips (malls, in stores you no longer go to), treasure hunt (items to find), garden centers, races, social VR world, transferable or portable, peaceful places, busy environments (more hustle and bustle), competitions (including with family or grandchildren), races, audible directions of where to go (sight impairment), avatar customizable, historical information plaques, calm sounds, directional seated movement

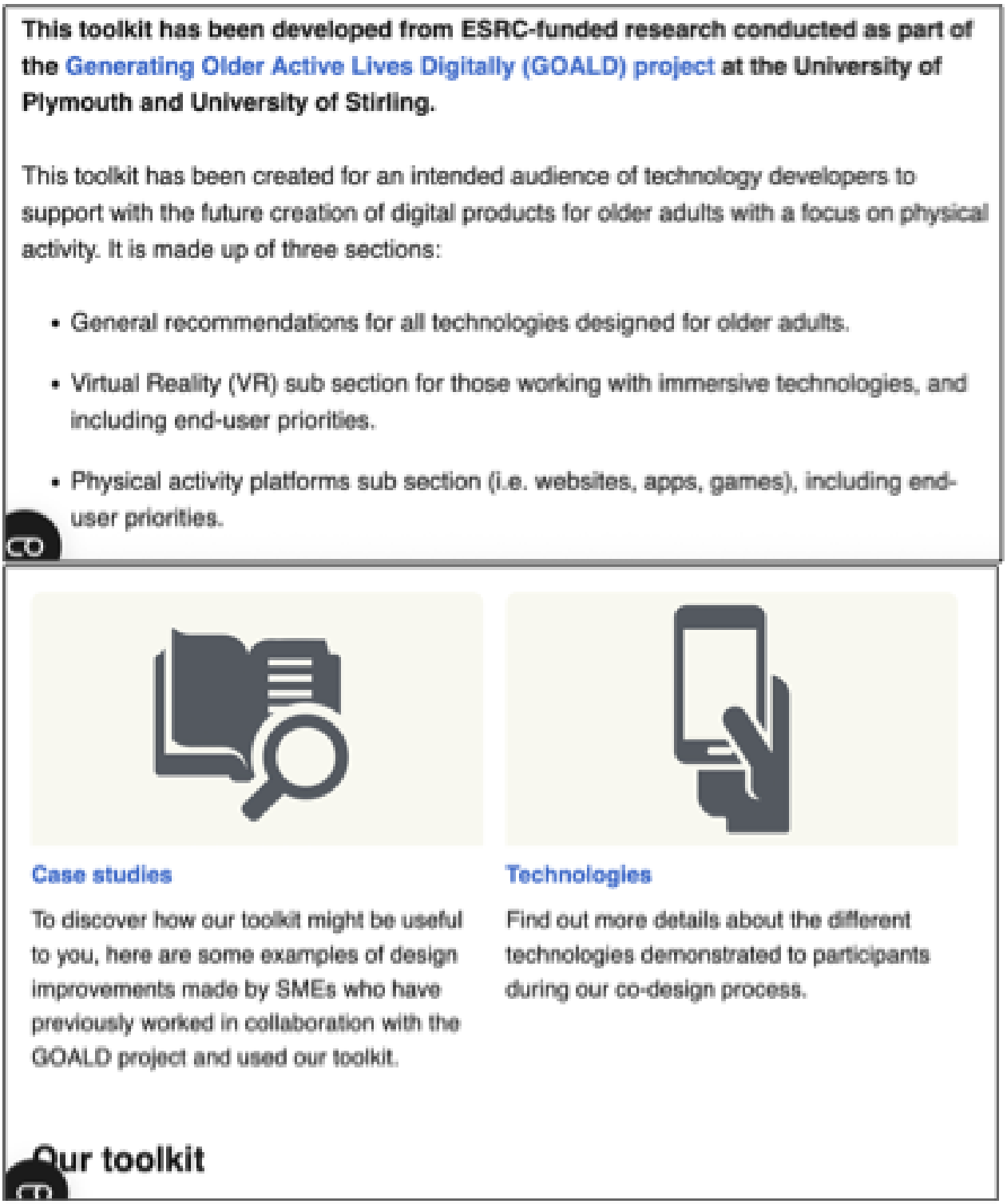
[a]VR: virtual reality.

[b]POV: point of view.

Features were separated into 3 priority categories based on clustering the features according to calculated priorities to offer SMEs guidance on which features to prioritize in their product developments. The priorities were calculated by balancing the ranking across the venues, workshops, and group sizes to equal out scores from individuals over scores from groups. These were provided to the SMEs who won the challenge fund competition alongside the preliminary toolkits. The full GOALD toolkit is available online [36]. Screenshots illustrating the toolkit are shown in Figure 2.

XSL•FO
RenderX

Figure 2. Screenshots of the toolkit to exemplify its contents. ESRC: Economic and Social Research Council; GOALD: Generating Older Active Lives Digitally; SME: small- and medium-sized enterprise; VR: virtual reality.



With the support of the £5000 grants (a currency exchange rate of £1=US $1.23 was applicable) and use of the toolkits, 10 SMEs (Multimedia Appendix 5) were each awarded £5000 based on their ideas to use the preliminary toolkits and the likely impact of the product. They made meaningful adaptations to their products, enhancing accessibility, user experience, and engagement. Changes included merging platforms for better navigation, adding customization and accessibility features such as subtitles, creating instructional guides and videos, integrating physical activity trackers, improving content variety, and developing voice-based and intergenerational features. Several SMEs also introduced new software modules, games, and apps, addressed safety and inclusivity, and tailored content to better suit the needs and preferences of diverse user groups (Table 6). Physical activity-specific adaptations are considered and reported elsewhere [15,16,30].

XSL•FO
RenderX

Table . Adaptations made to products as a result of the challenge fund and toolkits.

Business	Adaptations with challenge fund
Mature Movers	Updated content to have varied workout duration and more variety, took into consideration instructor preferences, added sections and filters, and began the process of exploring recommended exercises.
Triangular Pixels	Added physical activity tracker into virtual reality game for feedback and motivation.
Sentai	Added intergenerational connectivity and reminiscence feature and added voice-based physical activity encouragement.
MOTUS Ltd	Improved set-up procedures and created instructional videos to reduce reliance on the MOTUS staff facilitating sessions.
danceSing	Merged 2 websites into a single platform, added more intuitive navigation, added more user customization options, added accessibility features, for example, subtitles, added instructional guides on using the platform, moved live sessions to the website instead of social media, added a user chat function, added content filters, added analytics on user engagement, used age-relatable instructors.
Age Scotland	Enhanced accessibility features, added subtitles, improved warm-up and cool-down videos, added relatable music and older models, added direction on breathing and posture, explored ways of supporting with safety, added software for alternative languages, added different intensities of exercises.
Good Boost	Included new games into the retro-rehabilitation game that focus on strength, balance, and coordination, developed aquatic resistance exercise equipment to help with strength.
HiCarer	Built a module within the app to support people in finding virtual physical activities and added filters for the content, added activities with music and seated movements based on toolkit recommendations.
Weekday Wow Factor	Created Amazon Alexa voice app to provide people with information on group activities, addressed accessibility concerns, made the app multimodal (audio and visual prompts).
Football Memories Scotland	Created audio and visual content for reminiscence and mental activity.

Themes From Analysis of SME Toolkit Users' Reports and Interviews

From the SME reports and interviews, 6 themes emerged. There were 4 themes related to the impact of using the toolkits on their own product development: adaptations, business benefit, enhanced end-user insight, and challenges to development. The SMEs were also asked for feedback on the toolkits themselves so the final version of the toolkit could integrate their feedback. Themes relating to the toolkit and its further development were aspects of the toolkit positively received, and improvements to the toolkit.

SMEs used the toolkits' suggestions to adapt their products, such as training older adults as instructors, improving exercise delivery, offering varied workout lengths, adding website filters for easier navigation, and simplifying VR setups. They also introduced activity trackers, new delivery platforms (eg, Alexa widgets), and AI-voice features to motivate users. Respondents said the toolkits provided valuable context and helped develop their ideas further. The SMEs reported business benefits, that using the toolkits boosted their credibility, confidence, and speed in making user-focused design decisions. It was especially valuable for early-stage product development, helping expand their audience and market potential. SMEs noted increased enthusiasm, higher product engagement, and reduced support needs due to usability improvements. Economic benefits included easier sales, increased or projected increased sales, lower operational costs, and higher product uptake. The toolkit also supported investment by showing alignment with user needs, and the GOALD challenge fund helped accelerate funding and product changes.

The toolkits provided valuable, enhanced end-user insights, helping SMEs adjust priorities, validate decisions, and inspire new product ideas while avoiding poor design. It increased their audience awareness and highlighted the importance of co-design, encouraging further user involvement. SMEs saw the toolkits as a lasting resource they could keep using and expanding with their own findings. SMEs also reported challenges including development costs, limited funding aligned with older adults' needs, and difficulties in conducting research and measuring impact. They also faced obstacles integrating into certain settings, accessing community facilities, and managing time constraints or technical delays.

Regarding the toolkits themselves, the positive feedback theme highlighted the ranking features for helping prioritize developments, supported by the broad recommendations. Overall, SMEs reported finding the toolkit simple, concise, and easy to follow with clear step-by-step guidance. They suggested improving the toolkit by adding more guidance on co-design research, explaining the recommendations' background, merging

XSL•FO
RenderX

toolkits to reduce repetition, adding visuals, hosting it online, and including interactive features such as checklists and case studies. They also recommended showing participants a wider range of technologies to avoid bias. Some SMEs felt the toolkit's usefulness varied depending on the developer's experience and stage in their journey. To illustrate the utility of the toolkits, quotes illustrating themes and codes are presented in Multimedia Appendices 6 and 7.

Discussion

Principal Findings

This study primarily reports the development of a co-design–informed toolkit for SMEs, supported by a participatory process and qualitative evaluation of its perceived utility.

The GOALD project engaged older and younger adults in intergenerational co-design to produce recommendations for developers, particularly SMEs, designing technologies for older adults, with a focus on increasing physical activity and reminiscence. This process resulted in a final toolkit providing developers with insight into older adults' priorities. Consistent with previous research, older adults perceived benefits of digital health technologies for healthy aging, including physical activity, and identified diverse potential deployment settings [16]. Challenges using some technologies independently or within care settings reinforced the importance of co-designing accessible technologies [18]. Despite these challenges, familiarization enabled older adults to provide usability solutions. Preliminary toolkits and evaluations suggested that this approach could help software developers, particularly SMEs, by reducing the time and cost of considering older adults' needs, while supporting conversations with investors.

Older adults engaged in using the technologies and evaluating them. The benefits of familiarizing themselves with the technologies to guide their suggestions in the final co-design workshops were evident in the abundance of suggestions provided by participants, as others have demonstrated [32]. Interest was, however, uneven across technology types, resulting in the toolkit focusing primarily on VR and online physical activity platforms. To increase applicability, a general technology recommendation section was included to capture codes relevant across multiple technology types. Although exposure to all possible technologies was not feasible, the toolkit is designed to be updated as new co-design research and technologies emerge.

This study suggests that older adults can successfully engage in co-design workshops; however, it was important to adapt to their needs and preferences to ensure they were able to fully participate in the sessions. This included providing technologies and researcher support to connect to online sessions, such as large screens and microphones to assist with hearing. Others have highlighted how not adapting to the older person's needs can result in resistance to participate in co-design and a loss of the credibility surrounding co-design [37]. Therefore, the GOALD project worked with an advisory group to guide the design of the current project and, during workshops, gave longer allocated time for sessions, moved closer to support with hearing, and built rapport with the participants to adapt to their requirements. Many participants became familiar enough with VR headsets to suggest content, often focused on reminiscence, supporting the value of such content for well-being [30]. VR SMEs may particularly benefit from the GOALD toolkit for future older adult projects.

Addressing prior research limitations, which focused on older adults with higher income, education, and digital competence [38], GOALD recruited participants from diverse backgrounds, including care home residents and independently living older adults with physical, visual, and hearing impairments. Recruitment of younger adults for intergenerational co-design was more limited, due to COVID-19 and scheduling conflicts, despite evidence of mutual benefits such as improved attitudes toward older adults and increased community connectedness [39]. Future research should emphasize and communicate these benefits to younger participants.

Dissemination of co-design findings often omits end users, reducing uptake of feedback [40]. GOALD's evaluation of the toolkit with SMEs indicated its perceived value in helping them address older adults' needs, with updated versions available [36]. Unlike previous guidelines for older adult technology design [41], which were generic, this toolkit offers specific, participant-informed guidance. Systematic reviews have noted the merits of older adult involvement in design but often lack evidence on real-world viability or acceptability of resulting technologies [42]. The present study contributes to addressing this gap in knowledge. Further, 4 of the 10 SMEs had not had their software or devices evaluated in our co-design workshops, suggesting transferability of the toolkit ideas beyond the products on which they were based. Many publications or toolkits offering guidance on software design are very "generic" such that software developers might have guessed the advice, for example, "easy navigation" and "good user interface." This research has produced guidance that is more specific, strengthened by examples from older participants and the target market. Previous co-design studies have used structured usability and acceptability evaluations to assess older adults' willingness to interact with specific technologies, such as social or assistive robots [43,44]. These studies typically focus on the feasibility, usability, or acceptability of a single system within a defined evaluation context. In contrast, this study adopted an upstream, design-focused approach aimed at generating transferable insights and recommendations across multiple digital technologies and real-world SME development contexts. These approaches are complementary, with structured usability and acceptability evaluation representing an important subsequent step following toolkit-informed design. This distinction highlights that toolkit-based approaches operate upstream in the design process, whereas usability and feasibility studies typically evaluate downstream implementation and user interaction.

While participating SMEs found the toolkit useful and were able to apply its recommendations, embedding it into routine development practices over the longer term remains challenging. Ongoing use is likely to depend on practical factors such as time, resources, competing priorities, staff changes, and the

XSL•FO
RenderX

stage of product development. These factors help explain why SMEs applied the toolkit selectively in this study and suggest that future evaluations should examine how well such resources fit within everyday SME development practices over time.

Future research should evaluate the toolkit prospectively by embedding it from the earliest stages of product or service development and following its use over longer timeframes. This study suggests that the toolkit informed both adaptations to existing products and the development of new features and services. However, a fully prospective "from inception" evaluation would enable assessment of how toolkit recommendations shape design decisions across the full development lifecycle and their downstream impact on end-user outcomes. Such studies would benefit from longer funding periods and closer integration with SME development cycles to support sustained implementation and evaluation.

Limitations

This study has several limitations. Feature priorities were ranked to indicate their relative importance to older adults; however, differing timelines across sites meant that features could not be merged into a single ranking exercise. This resulted in separate initial toolkits for England and Scotland. In addition, not all technologies were available at every site, and facilitation differed across settings, which may have influenced participants' experiences and priorities. These differences across settings are discussed in detail elsewhere [30].

The card ranking exercise used in phase 4 was intended as a pragmatic prioritization tool to support co-design rather than as a validated quantitative measure. Rankings were used to indicate relative importance of features within the workshop context, and although steps were taken to rebalance scores for variation in group size and participation, the approach remains exploratory and context-specific. As such, ranking outputs should be interpreted as indicative of older adults' priorities to inform toolkit development, rather than as definitive or statistically robust measures of importance across populations or contexts. Despite this, participating SMEs reported that the toolkits provided specific, practical guidance that could be directly implemented to update or adapt their products.

The evaluation of toolkit impact was, of course, based on self-reported feedback from participating SMEs rather than the direct measurement of downstream design processes or end-user outcomes. This approach was taken because this project's funding window and scope did not permit longitudinal follow-up or independent assessment of postimplementation outcomes within participating organizations. As such, findings reflect SMEs' perceived utility and influence of the toolkit on their decision-making, rather than objectively measured changes in product quality or user outcomes. Reported impacts may therefore be subject to social desirability or recall bias, although this risk is reduced by the fact that feedback was collected after funding had been awarded and toolkit use had occurred. In addition, adoption of the toolkit recommendations varied across organizations, reflecting differences in development stage, resources, and priorities. This heterogeneous and partial adoption limits generalizability but reflects realistic patterns of toolkit use within SME development contexts.

The provision of the small challenge funds was likely to have been a catalyst for the use and success of the toolkits, but this was a crucial first step in testing them as part of the GOALD project. Future research could assess the resulting impact on older adults of adaptations to technologies made based on the toolkits; this was beyond the scope of the present research. Finally, the toolkits were produced following the selection of a limited range of digital products focusing on keeping older adults active. Nevertheless, 2 of the 10 SMEs awarded funding were from other areas (HiCarer, a directory type application, and Sentai focusing on connecting people with friends and family), suggesting broader applicability of the recommendations.

Conclusions and Future Research

Involving older adults in the co-design of technologies to support healthy aging is important for developing technologies that are more likely to be acceptable and relevant to older adults.

The GOALD project engaged older adults in co-design for technologies to support healthy aging with a focus on aiming to increase physical activity and connectedness to produce the present toolkit. This study focused on upstream design and implementation processes rather than direct measurement of aging-related health or functional outcomes. However, the toolkit was intended to support age-appropriate design decisions that may influence downstream usability, engagement, and outcomes. Evaluation of health, functional, or behavioral impacts requires longer-term implementation and outcome-focused study designs and is therefore an important next step for future research. The toolkit may act as a knowledge exchange between the research with older adults and with current and future SMEs, to help facilitate continued developments and engagement with co-design. SMEs who tested the toolkit reported useful content which informed some of their product adaptations and was perceived to support some business benefits. Future research could consider evaluating the impact of toolkit-informed adaptations to technology on older adults. It could also incorporate the results of co-design work on a broader range of technologies designed to increase other activity beyond physical activity and reminiscence.

Acknowledgments

We thank the participants of the co-design workshops and the 10 small- and medium-sized enterprises (SMEs) who were funded by the Generating Older Active Lives Digitally (GOALD) "mini-challenge" fund. The support of UK Research and Innovation via the Healthy Ageing Social Behavioural and Design Research Programme (SBDRP; ES/V016113/1) from the Economic and Social Research Council (ESRC) is gratefully acknowledged. Generative AI was not used in this paper.

XSL•FO
RenderX

Funding

UK Research and Innovation provided funding via the Healthy Ageing Social Behavioural and Design Research Programme (SBDRP; ES/V016113/1) from the ESRC. Ten SMEs were funded by the GOALD "mini-challenge" fund. ESRC had no role in the conduct of this study.

Data Availability

Transcripts and researcher notes from focus groups have been archived with this project with the ESRC as per funding requirements and are available via the UK Data Service.

Authors' Contributions

Conceptualization: RH, CH, RBJ, ACW
Data curation: LC, HB, CBS, SAT
Formal analysis: LC, HB, CBS, SAT
Funding acquisition: RH, CH, RBJ, ACW
Project administration: RH, CH, RBJ, ACW
Supervision: RH, CH, RBJ, ACW
Writing – original draft: LC, HB, SAT, RB, RBJ, ACW
Writing – review & editing: LC, HB, CBS, SAT, RB, GCR, SEG, RH, CH, RBJ, ACW

On behalf of the GOALD project: Catherine H. Hennessy, Faculty of Social Sciences, University of Stirling; Ray B. Jones, Faculty of Health, University of Plymouth; Richard Haynes, Faculty of Arts & Humanities, University of Stirling; Anna C. Whittaker, Faculty of Health Sciences & Sport, University of Stirling, (Core Management Team); Sheena Asthana, Plymouth Institute for Health and Care Research (PIHR), University of Plymouth; Rory Baxter, Faculty of Health, University of Plymouth; Hannah Bradwell, Faculty of Health, University of Plymouth; Arunangsu Chatterjee, Faculty of Medicine and Health, University of Leeds; Pete Coffee, School of Social Sciences, Heriot-Watt University; Leonie Cooper, Faculty of Health, University of Plymouth; Alison Dawson, Faculty of Social Sciences, University of Stirling; Katie Edwards, Faculty of Health, University of Plymouth; Swen Gaudl, Department of Applied IT, University of Gothenburg; Tanja Krizaj, Faculty of Health, University of Plymouth; Greg Mannion, Faculty of Social Sciences, University of Stirling; Gemma Ryde, Institute of Cardiovascular and Medical Sciences, University of Glasgow; Alejandro Veliz Reyes, Faculty of Arts, Humanities and Business, University of Plymouth; John Ritchie, Faculty of Arts & Humanities, University of Stirling; Colin B Shore, Faculty of Health Sciences & Sport, University of Stirling; Simone Tomaz, Faculty of Health Sciences & Sport, University of Stirling; Alison Warren, Faculty of Health, University of Plymouth; Karen Watchman, Faculty of Health Sciences & Sport, University of Stirling; Katharine Willis, Faculty of Arts, Humanities and Business, University of Plymouth; and partners: Active Stirling Ltd, Generations Working Together, Hearing Loss Cornwall, Nudge Community Builders, South Asian Society (Devon and Cornwall), Sporting Heritage, Sports Heritage Scotland, St. Breward Community, iSight Cornwall, and UKActive.

Conflicts of Interest

ACW declared an in-kind existing research relationship with danceSing working on a project independent of GOALD. None of the other authors have any competing interests in this study. None of the authors are on the editorial board of the journal or on any committee that may benefit from publication of this paper. There are no legal proceedings relevant to this study. No authors had any financial interest in the 10 companies that participated in this study during this study. ACW has since provided some consultancy to danceSing on a separate project from 2023 to 2024. The authors developed the "toolkits" that were evaluated in this study. These toolkits are freely available to all technology developers.

Multimedia Appendix 1
Descriptor of case sites across England and Scotland (including breakdown of participants at each site).
[DOCX File, 17 KB - aging_v9i1e89785_app1.docx]

Multimedia Appendix 2
Main themes and codes, with evidence, used to form the general recommendations for technology for older adults (case site numbers provided).
[DOCX File, 19 KB - aging_v9i1e89785_app2.docx]

Multimedia Appendix 3
Themes and codes, with evidence, that formed the recommendations directed toward physical activity platforms (case site numbers provided).
[DOCX File, 18 KB - aging_v9i1e89785_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 4
Main themes and codes, with evidence, used to form the recommendations for virtual reality for older adults with the focus on increasing physical activity (case sites provided).
[DOCX File, 20 KB - aging_v9i1e89785_app4.docx]

Multimedia Appendix 5
Organizations awarded funding with project aim and hoped-for outcome.
[DOCX File, 17 KB - aging_v9i1e89785_app5.docx]

Multimedia Appendix 6
Quotes from Mature Movers coded in the thematic analysis showing themes.
[DOCX File, 17 KB - aging_v9i1e89785_app6.docx]

Multimedia Appendix 7
Quotes from developer “4” coded in the thematic analysis showing themes.
[DOCX File, 17 KB - aging_v9i1e89785_app7.docx]

References

1. Ageing and health. World Health Organization. 2025 Oct 1. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2026-08-01]
2. Resnick B, Klinedinst NJ, Yerges-Armstrong L, Choi EY, Dorsey SG. The impact of genetics on physical resilience and successful aging. J Aging Health 2015 Sep;27(6):1084-1104. [doi: 10.1177/0898264315577586]
3. Choi M, Lee M, Lee MJ, Jung D. Physical activity, quality of life and successful ageing among community-dwelling older adults. Int Nurs Rev 2017 Sep;64(3):396-404. [doi: 10.1111/inr.12397] [Medline: 28837231]
4. Shin E, Kim M, Kim S, Sok S. Effects of reminiscence therapy on quality of life and life satisfaction of the elderly in the community: a systematic review. BMC Geriatr 2023 Jul 10;23(1):420. [doi: 10.1186/s12877-023-04001-1] [Medline: 37430198]
5. Tolson D, Schofield I. Football reminiscence for men with dementia: lessons from a realistic evaluation. Nurs Inq 2012 Mar;19(1):63-70. [doi: 10.1111/j.1440-1800.2011.00581.x]
6. Greco G, Poli L, Clemente FM, Francesco F, Cataldi S. The effectiveness of new digital technologies in increasing physical activity levels and promoting active and healthy ageing: a narrative review. Health Soc Care Community 2023 Feb 8;2023:1-11. [doi: 10.1155/2023/2803620]
7. A consensus on healthy ageing. GOV.UK. 2023 Feb 10. URL: https://www.gov.uk/government/publications/healthy-ageing-consensus-statement/a-consensus-on-healthy-ageing [accessed 2026-08-01]
8. Stockwell S, Schofield P, Fisher A, et al. Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults: a systematic review and meta-analysis. Exp Gerontol 2019 Jun;120:68-87. [doi: 10.1016/j.exger.2019.02.020]
9. Ioannou E, Chen HL, Bromley V, et al. The key values and factors identified by older adults to promote physical activity and reduce sedentary behaviour using co-production approaches: a scoping review. BMC Geriatr 2023 Jun 16;23(1):371. [doi: 10.1186/s12877-023-04005-x] [Medline: 37328734]
10. McGarrigle L, Boulton E, Todd C. Map the apps: a rapid review of digital approaches to support the engagement of older adults in strength and balance exercises. BMC Geriatr 2020 Nov 18;20(1):483. [doi: 10.1186/s12877-020-01880-6] [Medline: 33208117]
11. Aksay E. Live online exercise programs during the COVID-19 pandemic – are they useful for elderly adults? J Phys Educ Sport 2021 Jun 1;21(4):1650-1658. [doi: 10.7752/JPES.2021.04209]
12. Dekker-van Weering M, Jansen-Kosterink S, Frazer S, Vollenbroek-Hutten M. User experience, actual use, and effectiveness of an information communication technology-supported home exercise program for pre-frail older adults. Front Med 2017;4:208. [doi: 10.3389/fmed.2017.00208]
13. Piech J, Czernicki K. Virtual reality rehabilitation and exergames—physical and psychological impact on fall prevention among the elderly—a literature review. Appl Sci 2021;11(9):4098. [doi: 10.3390/app11094098]
14. Yen HY, Chiu HL. Virtual reality exergames for improving older adults’ cognition and depression: a systematic review and meta-analysis of randomized control trials. J Am Med Directors Assoc 2021 May;22(5):995-1002. [doi: 10.1016/j.jamda.2021.03.009] [Medline: 33812843]
15. Bradwell HL, Cooper L, Baxter R, et al. Implementation of virtual reality motivated physical activity via omnidirectional treadmill in a supported living facility for older adults: a mixed-methods evaluation.: virtual reality to motivate physical activity for older adults. Presented at: CHI ’24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems; May 11-16, 2024. [doi: 10.1145/3613904.3642281]

XSL•FO
RenderX

16. Campelo AM, Katz L. Older Adults' perceptions of the usefulness of technologies for engaging in physical activity: using focus groups to explore physical literacy. Int J Environ Res Public Health 2020 Feb 11;17(4):1144. [doi: 10.3390/ijerph17041144] [Medline: 32053937]
17. Cook EJ, Randhawa G, Sharp C, et al. Exploring the factors that influence the decision to adopt and engage with an integrated assistive telehealth and telecare service in Cambridgeshire, UK: a nested qualitative study of patient "users" and "non-users". BMC Health Serv Res 2016 Apr 19;16(1):137. [doi: 10.1186/s12913-016-1379-5] [Medline: 27095102]
18. Wandke H, Sengpiel M, Sönksen M. Myths about older people's use of information and communication technology. Gerontology 2012 Oct 29;58(6):564-570. [doi: 10.1159/000339104]
19. Bradwell HL, Edwards KJ, Winnington R, Thill S, Jones RB. Companion robots for older people: importance of user-centred design demonstrated through observations and focus groups comparing preferences of older people and roboticists in South West England. BMJ Open 2019 Sep;9(9):e032468. [doi: 10.1136/bmjopen-2019-032468]
20. LaMonica HM, Davenport TA, Roberts AE, Hickie IB. Understanding technology preferences and requirements for health information technologies designed to improve and maintain the mental health and well-being of older adults: participatory design study. JMIR Aging 2021 Jan 6;4(1):e21461. [doi: 10.2196/21461]
21. Schell R, Hausknecht S, Kaufman D. Barriers and adaptations of a digital game for older adults. Presented at: Proceedings of the 1st International Conference on Information and Communication Technologies for Ageing Well and e-Health (AGEWELL-2015); May 23-25, 2015; Lisbon, Portugal p. 269-275. [doi: 10.5220/0005524002690275]
22. Henson C, Chapman F, Shepherd G, Carlson B, Rambaldini B, Gwynne K. How older indigenous women living in high-income countries use digital health technology: systematic review. J Med Internet Res 2023;25:e41984. [doi: 10.2196/41984]
23. Zamenopoulos T, Alexiou K. Co-design as collaborative research. : University of Bristol/AHRC Connected Communities Programme; 2018 URL: https://oro.open.ac.uk/58301/1/Co-Design_CCFoundationSeries_PUBLISHED.pdf [accessed 2026-08-18]
24. Sanders EBN, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068] [Medline: 18836093]
25. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17. [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
26. Jagtap S. Codesign in resource-limited societies: theoretical perspectives, inputs, outputs and influencing factors. Res Eng Des 2022 Apr;33(2):191-211. [doi: 10.1007/s00163-022-00384-1]
27. Bradwell HL, Cooper L, Edwards KJ, et al. Staff perceptions towards virtual reality-motivated treadmill exercise for care home residents: a qualitative feedback study with key stakeholders and follow-up interview with technology developer. BMJ Open 2023 Nov 23;13(11):e073307. [doi: 10.1136/bmjopen-2023-073307] [Medline: 37996232]
28. Zacharuk A, Ferguson A, Komar C, et al. The effects of co-designed physical activity interventions in older adults: a systematic review and meta-analysis. PLoS One 2024 May 1;19(5):e0297675. [doi: 10.1371/journal.pone.0297675]
29. Barac R, Stein S, Bruce B, Barwick M. Scoping review of toolkits as a knowledge translation strategy in health. BMC Med Inf Decis Making 2014 Dec;14(1):121. [doi: 10.1186/s12911-014-0121-7]
30. Tomaz SA, Shore CB, Ryde GC, et al. Generating Older Active Lives Digitally (GOALD): exploring older adults' views of digital technology for physical activity. J Phys Act Health 2026 May 1;23(5):653-665. [doi: 10.1123/jpah.2024-0799] [Medline: 41569863]
31. Tomaz SA, Taylor L, Ryde GC, et al. Generations active together: an example of using physical activity promotion and digital technology to bring together adolescents and older people in Stirling, Scotland. J Intergenerational Relat 2025 Apr 3;23(2):246-252. [doi: 10.1080/15350770.2024.2322442]
32. Harrington CN, Wilcox L, Connelly K, Rogers W, Sanford J. Designing health and fitness apps with older adults: examining the value of experience-based co-design. Presented at: PervasiveHealth '18: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 21-24, 2018. [doi: 10.1145/3240925.3240929]
33. Knowles S, Sharma V, Fortune S, Wadman R, Churchill R, Hetrick S. Adapting a codesign process with young people to prioritize outcomes for a systematic review of interventions to prevent self-harm and suicide. Health Expectations 2022 Aug;25(4):1393-1404. [doi: 10.1111/hex.13479] [Medline: 35521681]
3. Clanchy K, Mitchell J, Mulholland K, et al. Towards co-design of rehabilitation technologies: a collaborative approach to prioritize usability issues. Front Rehabil Sci 2024;5:1302179. [doi: 10.3389/fresc.2024.1302179] [Medline: 38450206]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. The GOALD toolkit. University of Plymouth. URL: https://www.plymouth.ac.uk/research/centre-for-health-technology/goald/toolkit [accessed 2026-08-01]
37. Sumner J, Chong LS, Bundele A, Wei Lim Y. Co-designing technology for aging in place: a systematic review. Gerontologist 2021 Sep 13;61(7):e395-e409. [doi: 10.1093/geront/gnaa064]
38. Moon S, Park K. The effect of digital reminiscence therapy on people with dementia: a pilot randomized controlled trial. BMC Geriatr 2020 May 6;20(1):166. [doi: 10.1186/s12877-020-01563-2] [Medline: 32375661]

XSL•FO
RenderX

39. De Santis KK, Mergenthal L, Christianson L, Busskamp A, Vonstein C, Zeeb H. Digital technologies for health promotion and disease prevention in older people: scoping review. J Med Internet Res 2023;25:e43542. [doi: 10.2196/43542]
40. Fried LP, Carlson MC, McGill S, et al. Experience corps: a dual trial to promote the health of older adults and children's academic success. Contemp Clin Trials 2013 Sep;36(1):1-13. [doi: 10.1016/j.cct.2013.05.003]
41. Constantin N, Edward H, Ng H, et al. The use of co-design in developing physical activity interventions for older adults: a scoping review. BMC Geriatr 2022 Aug 8;22(1):647. [doi: 10.1186/s12877-022-03345-4] [Medline: 35941570]
42. Duque E, Fonseca G, Vieira H, Gontijo G, Ishitani L. A systematic literature review on user centered design and participatory design with older people. Presented at: IHC 2019 - Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems Association for Computing Machinery, Inc; Oct 22-25, 2019. [doi: 10.1145/3357155.3358471]
43. Amabili G, Maranesi E, Margaritini A, et al. Usability and feasibility assessment of a social assistive robot for the older people: results from the GUARDIAN project. Bioengineering 2024;11(1):20. [doi: 10.3390/bioengineering11010020]
44. Bevilacqua R, Felici E, Cavallo F, Amabili G, Maranesi E. Designing acceptable robots for assisting older adults: a pilot study on the willingness to interact. Int J Environ Res Public Health 2021 Oct 12;18(20):10686. [doi: 10.3390/ijerph182010686] [Medline: 34682433]

Abbreviations

GOALD: Generating Older Active Lives Digitally
SME: small- and medium-sized enterprise
VR: virtual reality

Edited by F Puga; submitted 18.Dec.2025; peer-reviewed by D Zhu, G Amabili; revised version received 17.Jun.2026; accepted 22.Jul.2026; published 27.Aug.2026.

Please cite as:
Cooper L, Bradwell H, Shore CB, Tomaz SA, Baxter R, Ryde GC, Gaudl SE, Haynes R, Hennessy C, Jones RB, Whittaker AC, The GOALD Project
Developing Digital Platforms to Help Older Adults Stay Physically and Socially Active: Co-Design Study to Develop a Toolkit and Evaluate Its Impact
JMIR Aging 2026;9:e89785
URL: https://aging.jmir.org/2026/1/e89785
doi:10.2196/89785

© Leonie Cooper, Hannah Bradwell, Colin B Shore, Simone A Tomaz, Rory Baxter, Gemma C Ryde, Swen E Gaudl, Richard Haynes, Catherine Hennessy, Ray B Jones, Anna C Whittaker, The GOALD Project. Originally published in JMIR Aging (https://aging.jmir.org), 27.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Measuring Implicit Attitudes Toward Digital Health Technologies in Older Adults Using an Influence-Aware Affect Misattribution Procedure: Development and Feasibility Study

Anna Vinnikova[1,2*], BA, MSc, PhD; Jinyue Zhan[3,4*], MB; Kailing Jin[3,5], MB; Xianfeng Ding[6], PhD; Wei Sang[7], PhD; Weina Xu[8], MMed; Qian Yang[1,2], PhD

[1]Department of General Medicine and School of Public Health, The Fourth Affiliated Hospital, and International School of Medicine, Zhejiang University School of Medicine, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, Zhejiang, China

[2]Zhejiang Key Laboratory of Intelligent Preventive Medicine, Zhejiang University, Hangzhou, Zhejiang, China

[3]Department of Geriatrics, School of Public Health, The Fourth Affiliated Hospital, and International School of Medicine, Zhejiang University School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China

[4]Infection Control Department, Lishui City People's Hospital, Lishui, Zhejiang, China

[5]Department of Infection Management and Preventive Health Care, Zhejiang University Affiliated Zhejiang Hospital, Hangzhou, Zhejiang, China

[6]School of Psychology, Central China Normal University, Wuhan, Hubei, China

[7]Faculty of Health and Wellness, City University of Macau, Macau SAR, Macau SAR, China

[8]Department of Geriatric Center for Regeneration and Aging Medicine, The Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, Yiwu, Zhejiang, China

[*]these authors contributed equally

Corresponding Author:
Qian Yang, PhD
Department of General Medicine and School of Public Health, The Fourth Affiliated Hospital, and International School of Medicine, Zhejiang University School of Medicine, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, Zhejiang, China

Abstract

Background: Older adults often express positive attitudes toward digital health technologies in surveys, yet adoption remains low. Self-report measures may not capture automatic affective reactions such as anxiety or distrust. Implicit paradigms such as the affect misattribution procedure (AMP) can reveal these automatic attitudes, but parameters optimized for younger adults may not be suitable for older adults because of age-related slowing and changes in visual processing.

Objective: This study aimed to adapt and evaluate an influence-aware affect misattribution procedure (IA-AMP) for measuring implicit attitudes of older adults toward digital health technologies and to identify an age-appropriate prime duration that balances affect transfer strength with minimal conscious awareness.

Methods: A 2-phase methodological adaptation and feasibility study was conducted among older adults (aged ≥60 y). Phase 1 (n=40) involved the development and validation of age-relevant synthetic images depicting older adults using digital health tools. The images were evaluated based on valence, arousal, thematic relevance, and low-level perceptual features. Phase 2 (n=56) implemented an IA-AMP with 3 prime durations (75, 350, and 425 ms) across 2 sequential cohorts. The first cohort (batch 2A, n=29) used the initial awareness probe, whereas the second cohort (batch 2B, n=27) used a simplified awareness interface. The core IA-AMP target judgment task remained unchanged across batches. The primary inferential outcome was the binary trial-level target judgment, coded as pleasant or unpleasant. Trial-level responses were analyzed using binomial logistic mixed-effects models with prime valence, prime duration, their interaction, and batch as fixed effects, along with random intercepts for participant and prime image. Awareness analyses were restricted to batch 2B.

Results: The primary generalized linear mixed-effects model showed a significant prime valence × prime duration interaction (χ^2_2=38.1; *P*<.001). Positive primes increased the odds of pleasant target judgments relative to negative primes at all durations (odds ratio [OR] 11.2, 95% CI 5.8-21.6, at 75 ms; OR 31.9, 95% CI 16.0-63.4, at 350 ms; and OR 45.6, 95% CI 22.1-94.0, at 425 ms; all Holm-adjusted *P*<.001). The positive-negative contrast was smaller at 75 milliseconds than at 350 and 425 milliseconds, whereas the contrasts at 350 and 425 milliseconds did not differ significantly. Batch sensitivity analyses showed stronger overall priming in batch 2B, but the prime valence × prime duration × batch interaction was not significant (χ^2_2=0.76; *P*=.68).

Conclusions: The IA-AMP offers a promising approach for assessing the affective responses of older adults to digital health technologies beyond self-report. A prime duration of approximately 350 milliseconds appears to be a practical calibration point for AMP studies involving older adults, producing strong affect transfer effects while avoiding the longest exposure duration.

XSL•FO
RenderX

Because reported influence awareness was common, AMP effects should be interpreted alongside awareness measures rather than as awareness-free implicit attitudes. AMP-based affective measures may complement usability and adoption research by identifying emotional responses that users may not readily articulate, thereby supporting more inclusive and evidence-informed development, evaluation, and implementation of digital health technologies for aging populations.

(JMIR Aging 2026;9:e87989) doi:10.2196/87989

KEYWORDS

older adults; digital health; technology adoption; implicit attitudes; affect misattribution procedure; usability; behavioral science

Introduction

Background

As global populations age, maintaining health, independence, and quality of life for older adults has become a public health priority. Digital health technologies, such as telemedicine platforms, wearable sensors, and smart home monitoring systems, are increasingly used to support older adults in managing chronic conditions and maintaining independence while aging in place. These tools can enhance health outcomes, reduce hospitalizations, and promote independent living by extending clinical monitoring into the home environment and supporting self-management [1,2]. However, despite these benefits, adoption remains limited. Many older adults report anxiety, low confidence, and mistrust toward digital tools, reducing engagement even when access and technical support are available [3]. Understanding both conscious and automatic emotional reactions to these technologies is crucial because affective comfort and trust often predict sustained use and adherence to digital health programs.

Most research on technology adoption of older adults relies on self-report measures. However, older adults may provide overly positive responses due to social desirability or limited self-awareness of their emotional responses. As a result, surveys may underestimate discomfort or resistance toward digital health tools [4]. Complementary assessment approaches are therefore needed to uncover emotional barriers that older adults may not express directly. Implicit measures offer one such approach by indexing automatic affective responses that operate outside conscious awareness and sometimes predict behavior beyond explicit reports [5,6]. By capturing intuitive reactions rather than deliberate opinions, these measures can reveal how well a technology truly aligns with users' needs and sense of security.

Implicit Attitudes and Measurement Gaps

Implicit attitudes, defined as automatic positive or negative reactions that occur outside conscious awareness [5], can shape real-world behavior, including technology use. Measuring these implicit responses can reveal how older adults actually feel about digital health technologies, even when they cannot or do not verbalize these attitudes. The affect misattribution procedure (AMP) is a validated tool for capturing such automatic affect transfer [7]. In this task, participants rate neutral stimuli that are briefly preceded by an affective "prime" (eg, an image of a health device), and the prime's emotional valence tends to bias evaluations of the neutral target. This paradigm has been used to study implicit evaluations of political and social stimuli [8,9], but its application to digital health contexts and older populations remains limited.

The AMP has been widely used in studies involving younger adults [7] but has rarely been tested in older populations or with age-relevant digital health imagery. Age-related differences in processing speed and attention suggest that standard AMP settings (approximately 75-ms primes) may be suboptimal for older adults, as slightly longer exposures may improve perceptual registration while also increasing the likelihood of conscious awareness [10,11]. Moreover, ecological validity matters in later life. Older adults may process age-congruent faces differently from younger faces, which highlights the importance of age-relevant materials when assessing affect [12]. Therefore, using stimuli that depict older adults in familiar health technology scenarios may produce more meaningful and representative affective responses.

Current Knowledge Gap

Prior work also indicates that awareness of primes can modulate AMP outcomes, although the direction and interpretation of this modulation remain debated [7]. Some findings suggest that conscious awareness amplifies affect transfer [13], whereas others view awareness reports as post hoc interpretations of implicit responses [14]. For aging research, the practical challenge is to measure automatic affective responses to digital health cues while minimizing demand characteristics and undue reliance on self-reported awareness [7]. To address this gap, there is a need to adapt implicit attitude paradigms so that they remain both sensitive and truly implicit when used with older adults. Determining timing parameters that maximize affect transfer while minimizing awareness may help clarify how best to assess older adults' spontaneous emotional reactions to digital health technologies.

Objective

This study tested whether the AMP can sensitively detect older adults' implicit attitudes toward digital health technologies and examined task parameters that optimize performance for this population. We conducted a 2-phase study with older adults recruited from a hospital setting to develop and validate age-relevant prime images depicting realistic digital health scenarios to ensure ecological validity and administer an AMP with 3 prime durations (75, 350, and 425 ms) and trial-level awareness assessments to evaluate whether affect transfer occurs, whether its magnitude varies by duration, and whether AMP effects persist on trials without reported awareness.

Our goal was to provide an age-appropriate, scalable, and health-relevant method to uncover unspoken comfort or

XSL•FO
RenderX

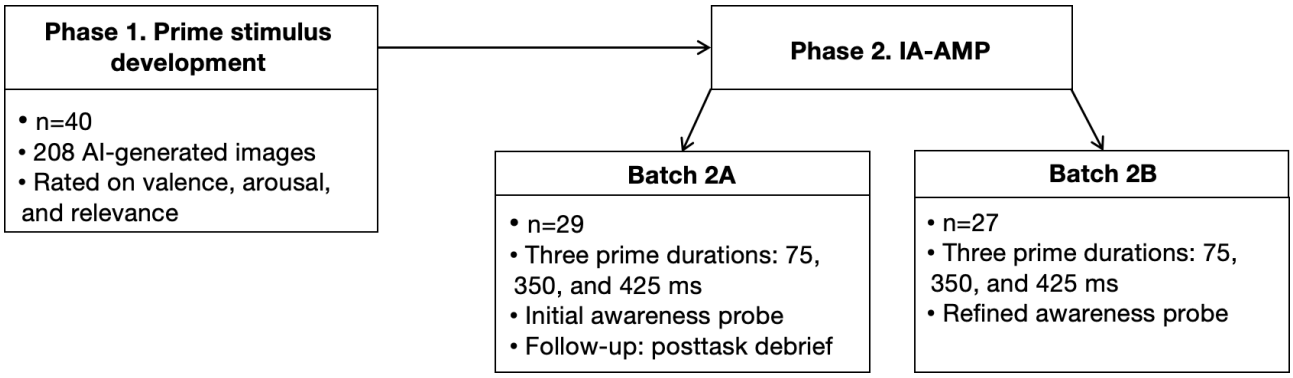
hesitation toward digital health tools so that designers and clinicians can better tailor user-centered interventions that support trust and engagement. By identifying optimal timing parameters for implicit attitude assessment in older adults, this work also contributes a new methodological resource for evaluating emotional readiness and acceptance of digital health technologies.

Methods

Study Design and Overview

We conducted a 2-phase methodological adaptation and feasibility study with older adults to examine whether AMP can capture implicit attitudes toward digital health technologies and to identify timing parameters appropriate for older participants. Phase 1 involved development and validation of age-relevant prime images depicting older adults in digital health contexts, evaluating perceptual balance and affective clarity. Phase 2 was conducted across 2 sequential cohorts: batch 2A and batch 2B. Both cohorts completed the same core influence-aware affect misattribution procedure (IA-AMP) target judgment task, in which participants evaluated ambiguous target images as pleasant or unpleasant after positive or negative digital health primes or gray-square baseline trials. Batch 2A used the initial awareness-probe interface. Because no awareness responses were recorded in batch 2A despite explicit instructions, participant debriefings were used to refine the awareness-probe interface for batch 2B. The batch 2B procedure simplified the awareness response but retained the same core IA-AMP target judgment task. Therefore, both cohorts were included in the primary AMP analysis, whereas awareness-related inference was restricted to batch 2B. All procedures were conducted in Mandarin. Figure 1 illustrates the overall study flow.

Figure 1. The overall study flow. IA-AMP: influence-aware affect misattribution procedure.



Ethical Considerations

Both studies were approved by the Ethics Committee of the School of Public Health, Zhejiang University (approval ID ZGL202411-1). All participants provided written informed consent before participation. Prime stimuli were synthetically generated to avoid any identifiable human participants; no personal data were collected or stored.

Setting, Eligibility, and Recruitment

Participants aged ≥60 years were recruited from an inpatient geriatric department of a hospital in Hangzhou, China. Testing was conducted individually in a quiet, well-lit room using a standard desktop computer with a full keyboard. Inclusion criteria included normal or corrected-to-normal vision and the ability to understand and perform computerized tasks. Exclusion criteria included visual impairment, a history of neurological or psychiatric conditions (eg, stroke, dementia, Parkinson disease, and major depression), use of medications with cognitive side effects, motor impairments preventing keyboard use, or noncompliance during practice (eg, pressing the same response key on all trials).

Phase 1: Stimulus Development and Validation

Objective

The first phase aimed to create a perceptually balanced, age-relevant set of positive and negative digital health primes and verify affective clarity and thematic relevance for older adults.

Stimulus Generation and Screening

A pool of 208 AI-generated images depicting older adults engaging with digital health technologies was created using Imagen 3 (Google). Prompts systematically varied 3 content dimensions: older-adult descriptors (eg, “senior” and “retired individual”), digital health contexts (eg, telemedicine, wearable sensors, and health apps), and emotional expressions (eg, happy, satisfied, frustrated, concerned, and neutral). All images were synthetic and were pilot-screened by 3 independent researchers for realism, clarity of emotional expression, cultural appropriateness, and absence of identifiable or sensitive content before inclusion in the task.

Perceptual Control and Preselection

To minimize low-level visual confounds, all images were processed in Python (OpenCV) to extract spatial frequency (2D Fourier transform on grayscale), chromatic complexity (color entropy in HSV), and luminance (mean intensity of 0 - 255). Feature values were *z*-standardized, and Euclidean distance from the global centroid (0, 0, 0) was calculated to quantify perceptual deviation. We retained 25 images per valence category (positive, neutral, and negative; N=75) with the smallest distances for the behavioral rating phase.

XSL•FO
RenderX

Participants

A total of 40 older adults completed the rating task (2 additional participants were excluded due to noncompliance). Demographic information is summarized in Table 1.

Table . Demographic characteristics (study 1, n=40).

Variable	Value
Age (y), mean (SD)	68 (6.25)
Gender: female, n (%)	19 (47.5)
Education, n (%)	
Primary school	10 (25)
Middle school	14 (35)
High school	9 (22.5)
Junior college	4 (10)
Undergraduate	3 (7.5)
Residence, n (%)	
Urban	27 (67.5)
Rural	13 (32.5)
DHT[a] users, n (%)	30 (75)

[a]DHT: digital health technology.

Procedure

After 10 practice trials, each participant viewed the 75 preselected images once in randomized order (500 × 400 pixels). To prevent cross-dimensional interference, participants were randomly assigned to rate either valence or arousal (7-point scale); all participants also rated relevance (7-point scale) to the theme "older adults using digital health." Standardized instructions in Mandarin were presented onscreen and reiterated verbally. Responses were recorded by mouse click, and the task advanced automatically to the next trial.

Outcome Measures and Data Analysis

For each image, we computed mean ratings of valence, arousal, and relevance. Interrater reliability was estimated using the intraclass correlation coefficient (ICC[2,k]). Images were retained if they met predefined thresholds: valence≥5 (positive), ≤3 (negative), or 3.5 to 4.5 (neutral); relevance≥4; SD<1.5 for valence and relevance; and ICC≥0.70. Because neutral human images were not consistently rated as neutral, the neutral condition for the AMP task in phase 2 was operationalized as uniform gray squares, serving as a perceptual baseline. Normality was assessed using Shapiro-Wilk tests; group comparisons on perceptual features used ANOVAs or Kruskal-Wallis tests, as appropriate. Arousal was compared between final positive and negative sets to confirm matched emotional intensity. The final prime set comprised 5 positive and 5 negative age-relevant images matched in arousal and perceptual complexity.

Phase 2: IA-AMP (Duration Optimization and Awareness)

Objective

The second phase aimed to test whether the IA-AMP can capture older adults' implicit attitudes toward digital health technologies and to assess how variations in prime duration and awareness affect these responses.

Participants and Design

A total of 56 older adults participated across 2 sequential cohorts (Table 2). Batch 2A (n=29; mean age 65.0, SD 3.6 y; range 60 - 71 y) used the initial awareness-probe interface, which prompted participants for 5 seconds to press the "6" key if they felt influenced by the preceding prime; otherwise, they were instructed to press nothing. Following preliminary testing in batch 2A and participant debriefing, the awareness interface was simplified in batch 2B (n=27; mean age 67.3, SD 4.9 y; range 61 - 79 y) using a 4-second prompt requiring a space bar response. Batch 2A and batch 2B did not differ significantly in sex (P=.11) but differed modestly in age (P=.048) and substantially in education, with lower-education participants concentrated in batch 2A (P<.001).

The core AMP judgment task was unchanged across batches. Both batches used a 2 (prime valence: positive and negative) × 3 (prime duration: 75, 350, and 425 ms) within-participant design. Eligibility and ethics procedures mirrored phase 1.

XSL•FO
RenderX

Table . Demographic characteristics (study 2, n=56).

Variable	Batch 2A (n=29)	Batch 2B (n=27)	Total (n=56)
Age (y), mean (SD)	65.0 (3.6)	67.3 (4.9)	66.1 (4.4)
Age range (y)	60 - 71	61 - 79	60 - 79
Sex, n (%)			
Female	21 (72.4)	14 (51.9)	35 (62.5)
Male	8 (27.6)	13 (48.1)	21 (37.5)
Residence, n (%)			
Urban	13 (44.8)	14 (51.9)	27 (48.2)
Rural	16 (55.2)	13 (48.1)	29 (51.8)
Education, n (%)			
Primary school or below	21 (72.4)	3 (11.1)	24 (42.9)
Middle school	8 (27.6)	17 (63.0)	25 (44.6)
High school or vocational	0 (0)	6 (22.2)	6 (10.7)
Undergraduate	0 (0)	1 (3.7)	1 (1.8)
Digital health technology user, n (%)			
Yes	22 (75.9)	22 (81.5)	44 (78.6)

Materials and Apparatus

Five positive and 5 negative images from phase 1 were used as primes. Neutral primes were not carried forward because they were not consistently rated as neutral during phase 1. Instead, a uniform gray square was used as a perceptual baseline. The targets were 40 abstract paintings [15], which were chosen because these images are evaluatively ambiguous [16]. All stimuli were shown in PsychoPy (version 2023.2.1; Open Science Tools Ltd) on a 24-inch monitor (500 × 400 pixels). Participants used a standard keyboard to record their responses, and each trial ended with a blank screen that stayed visible until a key was pressed.

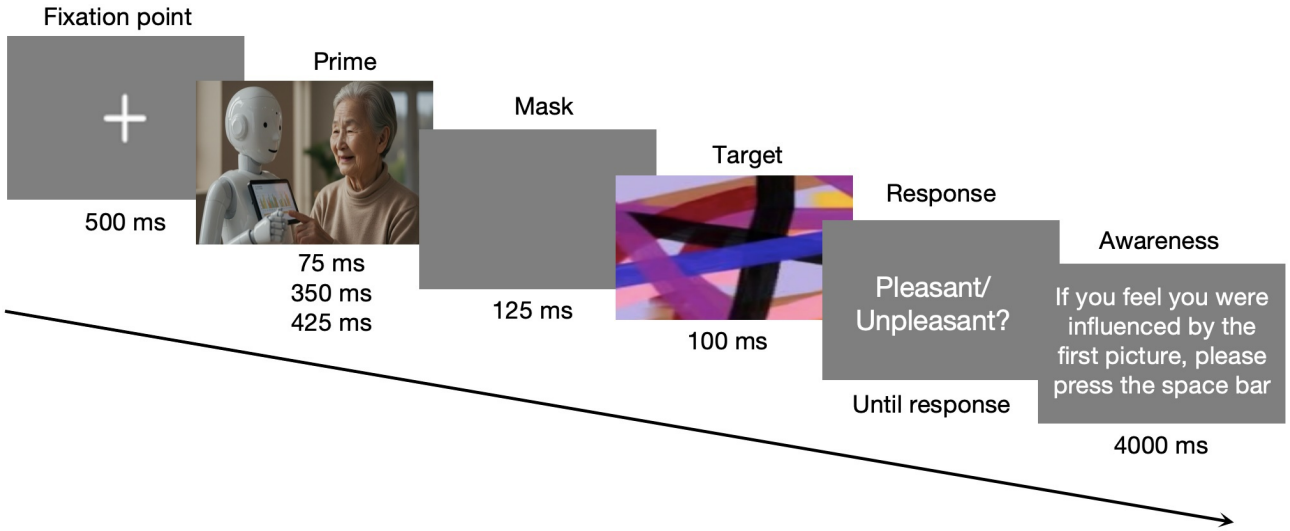
Procedure

Batch 2A

Each participant first completed 10 practice trials to become familiar with the task, followed by 90 critical trials divided into 6 blocks of 15 trials each. The trials included positive, negative, and gray-square baseline trials, and prime duration (75, 350, or 425 ms) was randomized within each block. On every trial, a fixation cross appeared for 500 milliseconds, followed by the prime image for the assigned duration. After a 125-millisecond blank screen, a target abstract painting was shown for 425 milliseconds. A black-and-white mask then remained on the screen until the participant responded. Participants pressed the “E” key if they judged the painting to be less pleasant than average and the “I” key if it was more pleasant than average. Immediately after each target evaluation, an awareness prompt appeared for 5 seconds with the instruction to press the “6” key if they felt that the preceding image had influenced their judgment; otherwise, they waited for the next trial. A 1-second interval separated trials. Participants were free to rest between blocks for as long as needed and could press the space bar to continue when ready. Primes were sampled randomly without replacement, and the order of trials was fully randomized so that the task proceeded in a fixed forward direction. The absence of awareness responses under this initial interface prompted posttask debriefings, which informed a simplified awareness-probe interface for batch 2B (Figure 2). The core IA-AMP target judgment task was unchanged.

XSL•FO
RenderX

Figure 2. Trial structure of the influence-aware affect misattribution procedure used in batch 2B. Batch 2A followed the same sequence but included a 5-second awareness prompt requiring a "6" key press and displayed the awareness screen after all trials, including gray-prime trials.



Posttask Probe

To supplement the trial-level awareness probe, 20 of the 29 participants were asked a brief set of follow-up questions immediately after completing the AMP task. These short probes were conducted in Mandarin by the researcher and focused on 4 domains: (1) task comprehension (eg, whether the instructions were clear), (2) interface usability (eg, ease of key presses and legibility of text), (3) cognitive load and fatigue (eg, whether the task felt confusing or tiring), and (4) perceptions of the awareness key (eg, whether participants understood what it meant).

Responses were grouped into 3 categories: (1) nonconscious processing, (2) procedural and interface-related barriers, and (3) motivational factors. These descriptive data were not analyzed as a full qualitative dataset but were used to contextualize the quantitative awareness results.

Batch 2B

Batch 2B used the same IA-AMP target judgment procedure as batch 2A, with a simplified awareness task based on batch 2A debriefings. Participants pressed the space bar instead of the "6" key to report awareness, and the awareness prompt appeared for 4 seconds. To reduce fatigue, the awareness screen was omitted during gray-square baseline trials. Therefore, awareness responses in batch 2B were collected only for positive-prime and negative-prime trials.

Data Analysis

Outcome Measures and Data Handling

The primary inferential outcome was the trial-level target evaluation in the IA-AMP task, coded as 1 for "pleasant" and 0 for "unpleasant." Consistent with standard AMP scoring, the AMP effect was summarized descriptively as the difference in the proportion of pleasant judgments following positive vs negative primes at each prime duration. "Pleasant judgments" refer to participants' evaluations of the ambiguous target paintings, whereas "positive" and "negative" refer to the valence of the preceding prime image. These proportions are reported only to illustrate the magnitude of the effect; formal statistical inference was performed at the trial level, as described in the *Statistical Analysis* section.

The primary positive-negative AMP analysis included only trials with the validated set of 10 affective primes, comprising 5 positive and 5 negative images. Gray-square baseline trials were excluded from this comparison because they served as perceptual baseline trials and did not contribute to the positive vs negative effect. Trial logs were checked for missing or duplicate responses and for noncompliance (eg, identical responses across all trials), and no missing or duplicate responses or instances of noncompliance were detected.

The secondary outcome was trial-level awareness response frequency among batch 2B trials that included an awareness prompt, indicating whether participants reported that the preceding prime influenced their judgment on a given trial. Awareness analyses were restricted to batch 2B because the awareness-probe interface differed between batches, and gray-square baseline trials were not included because they did not receive awareness prompts. Descriptive feedback from posttask interviews in batch 2A was used to contextualize awareness-probe comprehension and interface-use issues but was not included in inferential analyses.

Statistical Analysis

The binary trial-level outcome described above was analyzed using binomial logistic generalized linear mixed-effects models (GLMMs). The primary AMP model was fitted to the final positive-prime and negative-prime trials and included prime valence, prime duration, their interaction, and batch as fixed effects, with random intercepts for participant and prime image. The focal prime valence × prime duration interaction was evaluated using likelihood-ratio comparisons between models with and without the interaction. Duration-specific positive vs negative contrasts were estimated at each prime duration, with Holm correction applied across the 3 contrasts. Model diagnostics were evaluated using checks for model convergence, overdispersion, and residual patterns.

XSL•FO
RenderX

Positive vs negative prime contrasts were estimated at each prime duration. Because contrasts were examined at all 3 durations, *P* values were adjusted using the Holm method. Odds ratios (ORs) with 95% CIs were reported. Descriptive AMP scores, calculated as positive minus negative differences in pleasant-response proportions, were reported to aid interpretation but were not used as the primary inferential test. To evaluate whether the findings were robust to additional stimulus-level control, an item-control sensitivity model was fitted with crossed random intercepts for participants, prime images, and target images.

Trial-level awareness responses were analyzed in batch 2B only because the awareness-probe interface differed across batches. Awareness was modeled as a binary outcome, indicating whether participants reported that the preceding prime influenced their judgment on a given trial. Gray-square baseline trials were excluded because they did not receive awareness prompts.

To evaluate whether AMP effects differed across the 2 sequential cohorts, batch-interaction sensitivity models were fitted with batch and its interactions with prime valence and prime duration as fixed effects, along with random intercepts for participant and prime image. A targeted likelihood-ratio test compared a model with all 2-way batch interactions to a model that additionally included the prime valence × prime duration × batch interaction. To assess whether differences in cohort composition accounted for the priming effect, the primary model was additionally refitted with participant age, sex, and education as covariates.

All analyses were conducted in R (version 4.5.3; R Foundation for Statistical Computing). Binomial logistic GLMMs were fitted using glmmTMB. Estimated marginal means, model-estimated probabilities, and duration-specific contrasts were obtained using emmeans, with Holm correction applied across the 3 contrasts. The focal prime valence × prime duration interaction was evaluated using a likelihood-ratio test comparing models with and without the interaction term. Model diagnostics were assessed using DHARMa.

Results

Phase 1: Stimulus Validation

Participant Characteristics

A total of 40 older adults (aged 60 - 87 y; mean 68, SD 6.25 y; n=19, 47.5% female participants) completed the image-rating task. Two participants were excluded due to noncompliance (identical responses across trials). All remaining participants met the inclusion criteria and had normal or corrected vision.

Perceptual Feature Analysis

To evaluate whether the AI-generated images were perceptually balanced, we compared low-level visual properties (spatial frequency, chromatic complexity, and luminosity) across the 3 valence groups (Table 3).

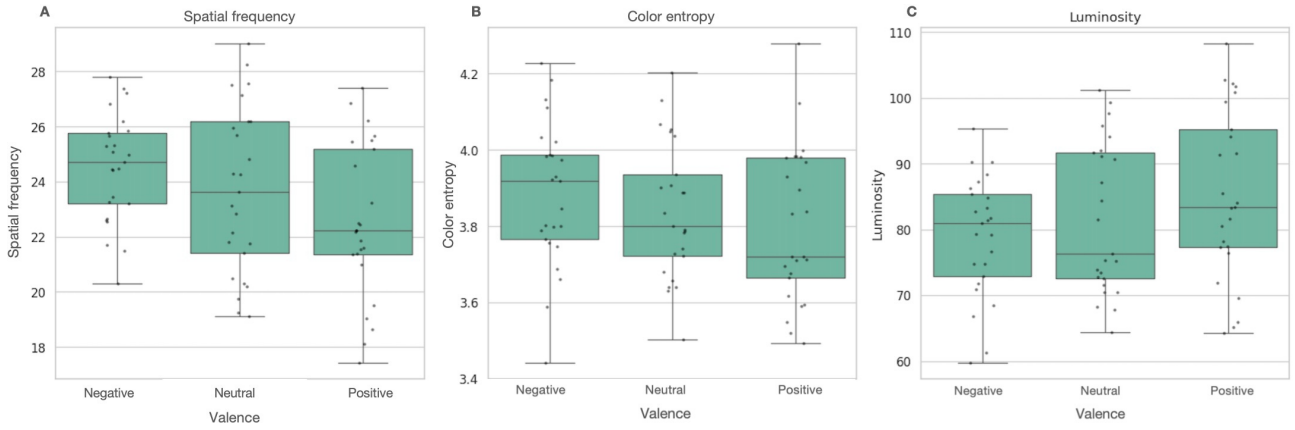
Table . Means and SDs of perceptual features for AI-generated images by valence category (N=75).

Valence	Spatial frequency, mean (SD)	Color entropy, mean (SD)	Luminosity, mean (SD)
Positive	24.51 (1.95)	3.88 (0.19)	78.97 (9.13)
Neutral	23.71 (3.03)	3.84 (0.18)	81.55 (11.46)
Negative	22.52 (2.78)	3.80 (0.20)	85.28 (12.98)

Shapiro-Wilk tests were used to assess normality, and 1-way ANOVAs or Kruskal-Wallis *H* tests were applied as appropriate. Descriptive distributions of the 3 perceptual features across valence categories are presented in Figure 3. Shapiro-Wilk tests indicated that distributions were approximately normal for spatial frequency and color entropy across groups, but not for luminosity (neutral group W=0.913; P=.04). Accordingly, 1-way ANOVAs were used for spatial frequency and color entropy, and a Kruskal-Wallis test was used for luminosity. A significant effect of valence was observed for spatial frequency ($F_{2,72}$=3.61; P=.03) driven primarily by differences between neutral and negative images. No significant differences were found for color entropy ($F_{2,72}$=1.15; P=.32) or for luminosity (H_2=2.95; P=.23). Because neutral images were not carried forward into the AMP task (phase 2), the spatial frequency effect was not considered relevant. Positive and negative image sets, which served as primes in phase 2, were comparable across all 3 perceptual dimensions.

XSL•FO
RenderX

Figure 3. Box plots of perceptual features by valence category: (A) spatial frequency, (B) color entropy, and (C) luminosity.



Behavioral Ratings

Reliability of Ratings

Interrater reliability was assessed for valence, arousal, and relevance ratings using ICC(2,k) (2-way random effects model and absolute agreement [17]; Table 4). Because each participant rated either valence or arousal, but all rated relevance, per-image ICCs could not be estimated; instead, overall reliability estimates are reported. Results indicated excellent agreement for valence (ICC=0.99, 95% CI 0.99-0.99) and good agreement for relevance (ICC=0.81, 95% CI 0.75-0.87) and arousal (ICC=0.77, 95% CI 0.68-0.84; all *P*<.001; Table 2). Split-half correlations yielded similar results (valence *r*=0.986, arousal *r*=0.839, and relevance *r*=0.801), supporting the robustness of these estimates. Together, these findings demonstrate that ratings were highly consistent across participants.

Table . Intraclass correlation coefficients (ICC[2,k]) for interrater reliability of image ratings.

Dimension	ICC(2,k); 95% CI	*F* value (*df*)	*P* value
Valence	0.99; 0.99-0.99	105.96 (72, 1368)	<.001
Arousal	0.77; 0.68-0.84	5.99 (72, 1368)	<.001
Relevance	0.81; 0.75-0.87	7.12 (72, 2808)	<.001

Filtering and Categorization

To ensure that the stimulus set was thematically relevant and affectively unambiguous, images were filtered in several steps. First, we excluded images with mean relevance ratings below 4 on a 7-point scale. Second, images showing high variability across raters (SD≥1.5 for valence or relevance) were removed to eliminate ambiguous stimuli. Third, arousal ratings were standardized, and images outside the 95% CI (*z*-scores ±1.96) were excluded to reduce the influence of atypical extremes. Finally, participant ratings were compared with intended valence categories. Whereas positive and negative sets were retained, most neutral images were inconsistently classified; only 5 neutral images passed the criteria, which was deemed insufficient and unreliable. Consistent with affective norming practices (eg, IAPS [International Affective Picture System] and OASIS [Open Affective Standardized Image Set]) [18,19], neutral primes were therefore operationalized as perceptually simple gray squares in the subsequent AMP task.

From the filtered pool, 5 positive and 5 negative images were retained as primes for the AMP. All selected stimuli were rated as moderately arousing, with mean ratings falling within the 3.5 to 4.5 range. A Welch 2-tailed *t* test confirmed that mean arousal did not differ significantly between the positive (mean 4.15, SD 1.60; n=100) and negative (mean 4.17, SD 1.65; n=100) sets (t_{198}=–0.09; *P*=.93). These results indicate that emotional intensity was successfully balanced across affective conditions.

Phase 2: IA-AMP

AMP Effects

The primary positive-negative AMP analysis used the final validated affective-prime set of 5 positive and 5 negative images (gray-square baseline trials were excluded). Data were analyzed using a trial-level binomial GLMM that included prime valence, prime duration, their interaction, and batch as fixed effects, along with random intercepts for participant and prime image. The prime valence × prime duration interaction significantly improved model fit, as assessed by a likelihood-ratio test (χ^2_2=38.1; *P*<.001), indicating that the positive-negative AMP effect varied across durations.

Descriptively, the raw AMP effect (the positive-negative gap in pleasant judgments) was smallest at 75 milliseconds (0.480) and increased to 0.625 at 350 milliseconds and 0.658 at 425 milliseconds. Full descriptive proportions and model-based estimates for each condition are reported in Table 5 and displayed in Figure 4.

XSL•FO
RenderX

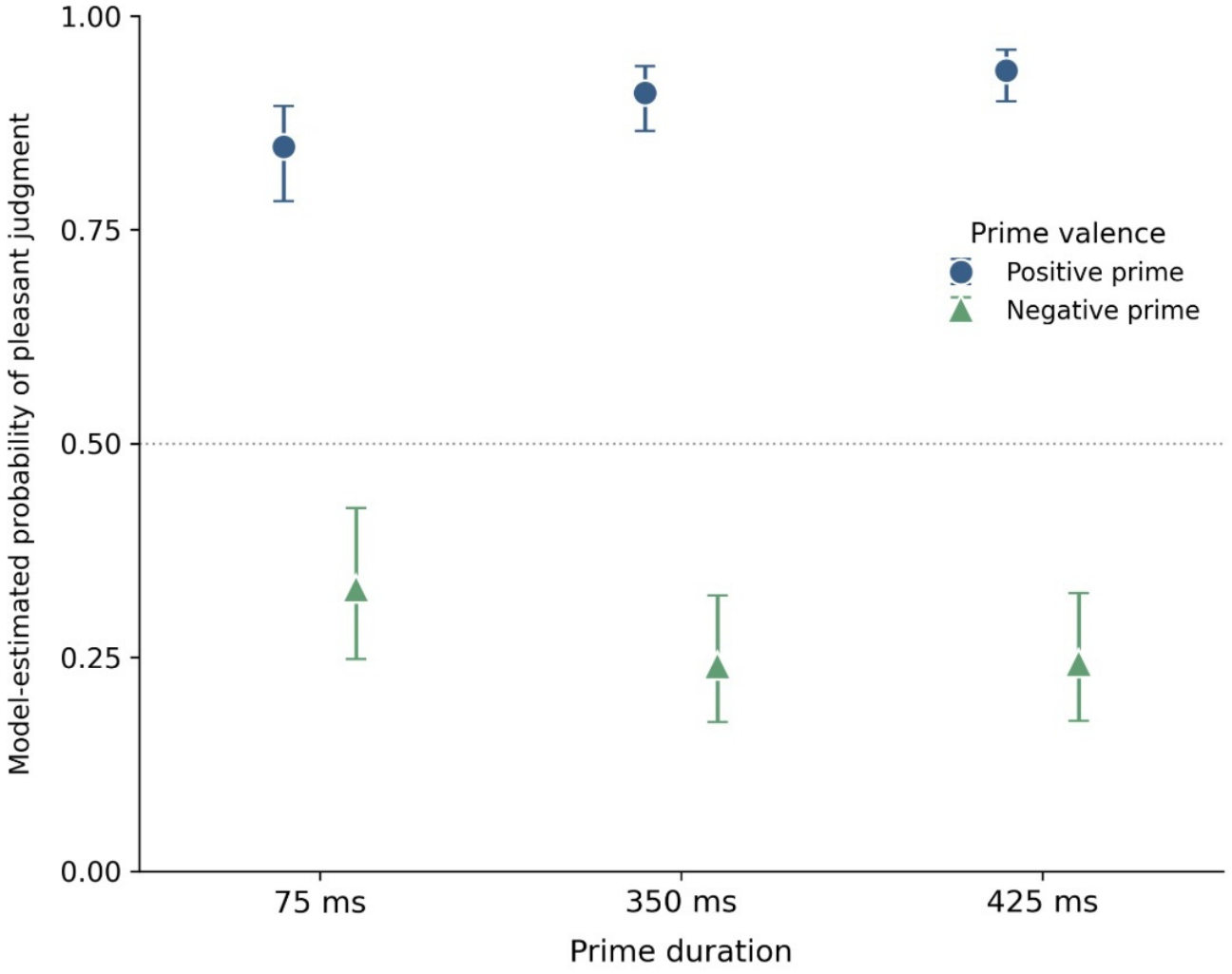
Table . Proportion of pleasant judgments, raw affect misattribution procedure (AMP) effects, and model-based contrasts by prime valence and prime duration[a].

Prime duration (ms)	Positive Pr(pleasant)	Negative Pr(pleasant)	Raw AMP effect	OR[b] (95% CI)	*P* value
75	0.827	0.347	0.480	11.2 (5.8-21.6)	<.001
350	0.889	0.264	0.625	31.9 (16.0-63.4)	<.001
425	0.921	0.263	0.658	45.6 (22.1-94.0)	<.001

[a]Positive Pr(pleasant) and Negative Pr(pleasant) are pooled trial-level proportions. Raw AMP effect = Positive Pr(pleasant) − Negative Pr(pleasant). OR is the odds ratio for positive vs negative prime contrast at each duration, estimated from the primary trial-level binomial GLMM. *P* values are Holm-adjusted across the 3 duration-specific contrasts.

[b]OR: odds ratio.

Figure 4. Model-estimated probability of a pleasant judgment by prime valence and prime duration. Circles represent estimated marginal probabilities for positive primes, and triangles represent negative primes, derived from the trial-level binomial generalized linear mixed-effects model (GLMM), averaged proportionally over batch, with 95% CIs back-transformed from the logit scale. The dotted line marks chance responding (0.5). Follow-up model-based contrasts showed that the positive-negative difference was smaller at 75 ms than at 350 and 425 ms (both Holm-adjusted *P*<.001), whereas the differences at 350 and 425 ms did not differ significantly (*P*=.17). Model-estimated probabilities differ slightly from the raw proportions in Table 5 because the GLMM adjusts for crossed random intercepts and is averaged proportionally over batch.



Holm-corrected positive vs negative contrasts were significant at all durations: positive primes increased the odds of a pleasant judgment relative to negative primes by a factor of 11.2 at 75 milliseconds, 31.9 at 350 milliseconds, and 45.6 at 425 milliseconds (all *P*<.001; 95% CIs given in Table 5).

The magnitude of this effect differed by duration. The positive-negative contrast was significantly smaller at 75 milliseconds than at 350 milliseconds (log-odds difference=−1.046; *P*<.001) and at 425 milliseconds (log-odds difference=−1.404; *P*<.001). The contrast did not differ significantly between 350 and 425 milliseconds (log-odds difference=−0.359; *P*=.17). Affect transfer was, therefore, robust at all durations but weaker at 75 milliseconds than at the 2 longer durations, which were statistically comparable.

An item-control sensitivity model adding target-image random intercepts produced the same substantive fixed-effect pattern,

XSL•FO
RenderX

with the prime valence × prime duration interaction remaining significant (χ^2_2=38.1; *P*<.001). The target-image SD was minimal (0.09 on the logit scale) relative to the participant and prime-image SDs (0.63 and 0.36, respectively), suggesting that the observed AMP effects were not primarily attributable to idiosyncratic features of particular target images.

Batch Sensitivity Analyses

Batch sensitivity analyses were conducted to evaluate whether AMP effects differed across the 2 sequential cohorts. A model allowing batch interactions improved fit relative to the primary model (χ^2_5=21.79; *P*<.001), indicating that batch moderated some aspects of the response pattern. Follow-up contrasts showed significant positive-negative priming at all durations in both batches. In batch 2A, positive vs negative ORs were 7.23 at 75 milliseconds, 25.19 at 350 milliseconds, and 31.86 at 425 milliseconds (all *P*<.001). In batch 2B, the corresponding ORs were 18.52, 43.32, and 72.82 (all *P*<.001).

However, the targeted prime valence × prime duration × batch interaction was not significant (χ^2_2=0.76; *P*=.68). This indicates that although batch 2B showed stronger overall positive-negative priming, the duration-dependent pattern of priming did not differ reliably between batches.

Because the 2 cohorts also differed in age and education (see Participants and Design section), we refitted the primary model adding participant age, sex, and education as covariates to assess whether these differences accounted for the priming effect. The prime valence × prime duration interaction remained significant (χ^2_2=38.0; *P*<.001), and none of the covariates significantly predicted pleasant judgments (for age *P*=.56, for sex *P*=.83, and for education *P*=.48). The duration-dependent priming pattern was therefore not attributable to differences in cohort composition.

Awareness Responses

No awareness responses were recorded in batch 2A under the initial awareness-probe interface. To clarify this zero-response pattern, 20 of the 29 batch 2A participants completed brief posttask probes in Mandarin. Responses were summarized descriptively and suggested that the absence of awareness responses reflected a combination of interface-related barriers, task-load issues, and conservative use of the awareness key, rather than definitive evidence of no prime awareness.

Awareness analyses were, therefore, restricted to batch 2B, where the simplified awareness interface was used. In batch 2B, reported influence awareness was common. Participants reported awareness of the preceding prime's influence on 57.5% of prompted trials at 75 milliseconds, 66.2% at 350 milliseconds, and 68.3% at 425 milliseconds. A binomial mixed-effects model confirmed that awareness increased with prime duration: awareness was significantly lower at 75 milliseconds than at 350 milliseconds (*P*=.004) and 425 milliseconds (*P*<.001), with no significant difference between 350 and 425 milliseconds (*P*=.52). Participant-level summaries indicated that awareness reports were not limited to a small subset of participants. All 27 batch 2B participants reported at least 1 awareness event at each duration.

These findings indicate that the simplified batch 2B awareness-interface successfully elicited awareness reports, but they also show that reported influence awareness was frequent. Therefore, phase 2 should not be interpreted as showing that AMP effects occurred under minimal or absent awareness.

Awareness-Stratified AMP Effects

Exploratory awareness-stratified summaries showed that reported influence awareness substantially shaped AMP effects in batch 2B. On awareness-reported trials, target judgments were highly aligned with prime valence. Pleasant-response rates following positive primes were 92.0%, 94.0%, and 96.1% at 75, 350, and 425 milliseconds, respectively, whereas corresponding rates following negative primes were 4.5%, 3.0%, and 3.7%. This produced raw awareness-reported AMP effects of 0.875, 0.910, and 0.924, respectively.

On trials without reported awareness, positive-negative differences were much smaller. Pleasant-response rates following positive vs negative primes were 78.8% vs 77.3% at 75 milliseconds, 81.6% vs 73.5% at 350 milliseconds, and 85.6% vs 70.6% at 425 milliseconds, corresponding to raw AMP effects of 0.016, 0.081, and 0.150, respectively. These descriptive patterns indicate that the overall AMP effect was strongly associated with reported influence awareness (Table 6).

Table . Awareness-stratified pleasant-response proportions in batch 2B[a].

Prime duration (ms)	Awareness status	Negative Pr(pleasant)	Positive Pr(pleasant)	Raw AMP[b] effect
75	Aware	0.045	0.920	0.875
75	Unaware	0.773	0.788	0.016
350	Aware	0.030	0.940	0.910
350	Unaware	0.735	0.816	0.081
425	Aware	0.037	0.961	0.924
425	Unaware	0.706	0.856	0.150

[a]Proportions are descriptive trial-level values from batch 2B only (n=27). Awareness status reflects whether participants reported that the preceding prime influenced their judgment on a given trial. Raw AMP effect = Positive Pr(pleasant) − Negative Pr(pleasant).

[b]AMP: affect misattribution procedure.

XSL•FO
RenderX

Because awareness-reported trials showed near-complete separation, with highly consistent pleasant responses following positive primes and unpleasant responses following negative primes, ORs from awareness-stratified logistic models were extremely large and are not emphasized in the main text. The cell-level proportions provide a more interpretable summary of the awareness-stratified pattern.

Summary of Timing Pattern

Across both batches, the IA-AMP produced robust positive-negative affect-transfer effects. The effect was significantly smaller at 75 milliseconds than at 350 milliseconds and 425 milliseconds, whereas 350 milliseconds and 425 milliseconds did not differ significantly. Batch 2B showed stronger overall priming than batch 2A, but the duration-dependent pattern was consistent across batches.

The results did not identify a duration that simultaneously maximized affect transfer and minimized awareness. The 75-millisecond duration showed the lowest reported awareness but the weakest overall AMP effect. The 425-millisecond duration showed the largest descriptive AMP effect but also the highest reported awareness and the longest exposure duration. The 350-millisecond duration produced strong affect transfer comparable to 425 milliseconds while avoiding the longest exposure duration. Therefore, 350 milliseconds is best interpreted as a practical calibration point rather than a strictly low-awareness or fully implicit optimum. Reported influence awareness remained common and should be monitored and modeled in future older-adult AMP studies.

Discussion

Principal Findings

This study adapted and evaluated the IA-AMP to assess older adults' affective responses toward digital health technologies. Phase 1 developed and validated age-relevant prime images depicting older adults in digital health contexts. The final positive and negative prime sets were emotionally distinct and comparable in arousal and low-level perceptual features, whereas neutral human images were not consistently rated as neutral and were, therefore, replaced by gray-square baseline stimuli in phase 2, consistent with prior AMP work [9].

Phase 2 evaluated AMP sensitivity across prime durations and examined the reported influence of awareness during affect transfer. Across 2 batches, participants were significantly more likely to judge ambiguous targets as pleasant following positive digital health primes than following negative primes. The magnitude of this effect was duration-dependent: the positive-negative contrast was significantly smaller at the standard 75 milliseconds duration and stronger at both longer durations, with no reliable difference between 350 and 425 milliseconds. This pattern suggests that very brief prime exposure may be insufficient for older adults, whereas moderately extended exposure improves affect transfer. This finding is consistent with broader evidence of age-related slowing in visual and cognitive processing speed [20,21], suggesting that standard AMP timing parameters developed for younger adults may require modification for older adults.

Awareness responses were not observed in the first batch but emerged after interface refinement in batch 2B. Posttask debriefings with batch 2A participants indicated that the absence of responses reflected interface comprehension barriers and conservative key-press behavior rather than a definitive absence of prime awareness. In contrast to the original expectation that awareness would remain relatively low, reported influence awareness in batch 2B was common, exceeding half of prompted trials at all durations. Awareness-stratified summaries further indicated that reported influence awareness substantially shaped the AMP effect, consistent with Hughes et al [13]. On awareness-reported trials, target judgments were highly aligned with prime valence, whereas positive-negative differences were much smaller on trials without reported awareness. These findings show that the IA-AMP can produce robust affect-transfer effects in older adults while also revealing that reported influence awareness plays an important role in those effects. Taken together, the results support 350 milliseconds as a practical timing point for older adults. It produced strong affect transfer comparable to 425 milliseconds while avoiding the longest exposure time, though it should not be interpreted as a strictly low-awareness or fully unconscious parameter.

The contrast between batch 2A and batch 2B is best interpreted as a methodological finding about awareness-probe design. The core IA-AMP target judgment task was unchanged across cohorts, and the awareness prompt appeared only after the pleasant or unpleasant target judgment had already been recorded. Batch-interaction sensitivity analyses showed that the duration-dependent priming pattern did not differ reliably between cohorts, supporting the robustness of the primary AMP finding. The zero-response pattern in batch 2A, therefore, limits awareness-related interpretation but does not compromise the primary AMP target-judgment outcome. Instead, it highlights that awareness probes for older adults must be designed with response accessibility, instruction clarity, and task burden in mind.

Comparison With Prior Work

This work extends the AMP literature by situating the present findings at the intersection of cognitive aging, visual processing, and methodological debates about the role of awareness in AMP effects. First, in relation to cognitive aging, the study applies the AMP to older adults, a population that has rarely been the focus of AMP methodological development. Earlier AMP work showed that evaluative priming can emerge across relatively fast and slower presentation rates [14,22], but much of this evidence was generated in younger adult samples and nonhealth attitude domains. In the present study, the shortest duration produced the weakest positive-negative contrast, whereas the 2 longer durations produced stronger and statistically comparable effects. This pattern is compatible with the idea that older adults may require somewhat longer stimulus exposure for reliable perceptual registration and evaluative carryover [20,21], even when the basic AMP mechanism remains intact. The study also extends the ecological scope of the AMP by using age-relevant digital health imagery, rather than the racial, political, or general social attitude stimuli that dominate much earlier work [7], showing that ambiguous targets can be

XSL•FO
RenderX

systematically biased by digital health primes in a later-life context.

Second, the awareness findings also place the present study within ongoing debates about the role of awareness in AMP effects across the broader implicit-cognition literature. Classic AMP accounts conceptualize the task as an indirect measure of automatically retrieved evaluative responses [9]. More recent work by Kurdi et al [14] suggests that awareness reports may sometimes be outcomes of AMP-consistent evaluative responding rather than causes of AMP effects, cautioning against treating reported awareness as straightforward evidence that awareness drove the AMP effect. In contrast, Hughes et al [13] argue that influence awareness can strongly moderate AMP effects, either by amplifying affect transfer through conscious re-evaluation or by qualifying the task's status as a measure of implicit attitudes. The present findings do not resolve this broader theoretical dispute, but they show that, in an older adult digital health context, reported influence awareness materially shaped the interpretation of AMP scores. Overall, priming was robust, but reported influence awareness was frequent, and awareness-reported trials showed much stronger positive-negative differences than trials without reported awareness. Thus, the influence-aware design was not merely an added procedural feature; it was central to interpreting what the task measured. An additional consideration specific to this population is that older adults characteristically prioritize accuracy over speed and adopt more conservative response thresholds than younger adults [23]. This tendency may extend to awareness reporting such that older adults who detect any prime-related internal reaction are more inclined to endorse an explicit awareness probe when prompted, producing a conservative reporting bias that may have contributed to elevated trial-level awareness rates beyond what strict conscious prime recognition alone would predict. These findings suggest that IA-AMP effects in older adults should be interpreted as affect-transfer responses measured alongside awareness, rather than as straightforward evidence of awareness-free implicit attitudes or pure affect misattribution [13,24].

Implications for Digital Health Research

The present findings have several implications for digital health research and evaluation with older adults when interpreted alongside adoption research showing that older adults' technology use is shaped not only by perceived usefulness or ease of use [25,26], but also by emotional, social, relational, and contextual factors [27,28]. Prior work has consistently documented barriers to older adults' digital health engagement, including limited digital literacy, physical and cognitive challenges, usability problems, privacy concerns, mistrust, and technology anxiety, alongside facilitators such as accessible design, provider endorsement, social support, and co-design [27-29]. Taken together, this body of work suggests that emotional and trust-related responses are not peripheral to digital health engagement—they are part of the adoption process itself.

The IA-AMP contributes a complementary measurement approach for studying affective barriers that conventional self-report methods may not fully capture, particularly given older adults' susceptibility to social desirability bias and limited introspective access to automatic emotional reactions [4]. Innovation Resistance Theory helps situate what those affective barriers may represent by distinguishing functional barriers, such as usage, value, and risk, from psychological barriers, such as tradition and image, and by framing resistance as an active response rather than a passive failure to adopt [30]. In digital health contexts, older adults' resistance may involve emotional discomfort, symbolic misalignment, fears about losing relational care, privacy and data-security concerns, distrust in diagnostic accuracy, and skepticism about care quality [31]. Low self-efficacy and technology anxiety may further create feedback loops that reinforce avoidance behaviors. The present findings are consistent with this perspective: digital health primes elicited rapid affective responses that biased target judgments, suggesting that affective and image-related components of resistance may be measurable at the trial level. Incorporating the IA-AMP into standard technology evaluation pipelines in community or clinical settings would allow developers to assess the "affective usability" of a digital health tool, its capacity to foster emotional comfort and trust, alongside traditional functional usability metrics.

Limitations and Strengths

The primary strength of this work lies in demonstrating that IA-AMP can be adapted for older adults and digital health imagery with a relatively low technical burden. The study was intentionally designed with several simplifications to balance methodological control and participant accessibility. Participants completed the task on standard desktop computers under researcher supervision, suggesting that the procedure is feasible in clinical or applied research settings.

While the sample sizes were modest (phase 1: n=40; phase 2: n=56), they were appropriate for a feasibility study and consistent with prior AMP and implicit-attitude research [32]. Nonetheless, this sample limits generalizability and precludes the examination of subgroup differences (eg, by age or cultural background). Larger and more diverse cohorts will be needed to evaluate the stability of IA-AMP effects, establish psychometric reliability, and test predictive validity against real-world measures of technology adoption, which is necessary before clinical relevance can be established.

The sequential recruitment of the 2 cohorts introduces limitations that shape the interpretation of both the primary and awareness-specific findings. The original batch 2A awareness-interface yielded no recorded responses and was subsequently simplified for batch 2B. Furthermore, the cohorts differed in baseline characteristics, with lower-education participants and a modest age difference concentrated in batch 2A. Although adjusting for age, sex, and education indicated that these demographic differences did not account for the primary AMP effect, the potential influence of temporal context arising from sequential recruitment cannot be fully ruled out. Consequently, while the core duration-dependent priming pattern was robust across both batches, the awareness-related conclusions are tied to the simplified batch 2B interface. Finally, reported influence awareness was common in batch 2B, suggesting that IA-AMP effects in this context reflect

XSL•FO
RenderX

affect-transfer responses measured alongside awareness rather than in its absence.

Taken together, the findings indicate that the IA-AMP is a feasible and scalable method for assessing older adults' affective responses to digital health technologies. Its brevity, low verbal demand, and computerized format make it potentially useful as a complementary method in digital health evaluation, particularly when used alongside self-report, usability testing, and awareness assessment. However, its value should not be understood as providing a simple, awareness-free measure of hidden attitudes. Rather, the IA-AMP may be most useful as part of a multimethod evaluation framework that captures affect-transfer responses while explicitly monitoring reported influence awareness.

Conclusions and Future Directions

The present work demonstrates the feasibility of using the IA-AMP to assess older adults' affective responses to digital health technologies. A prime duration of approximately 350 milliseconds produced strong affect transfer comparable to 425 milliseconds while avoiding the longest exposure duration. These findings suggest that moderately extended prime durations provide a practical timing point for older-adult AMP studies, improving affect-transfer sensitivity relative to very brief exposures. The findings also demonstrate that reported influence awareness must be measured and modeled when using AMP paradigms with older adults. Future studies should expand on this foundation by recruiting larger and more diverse cohorts to evaluate the stability of IA-AMP effects, examining psychometric reliability and predictive validity in relation to real-world digital health engagement, and incorporating richer awareness measures to clarify the boundary between automatic and conscious affect transfer. The next phase of research should also integrate the IA-AMP into participatory design and implementation studies to test whether reducing implicit negativity or enhancing positive affect toward digital health technologies improves adoption and sustained use. Advancing this approach from feasibility to validated measurement will be critical for developing technologies that older adults not only can use but are also willing and affectively prepared to engage with in later life.

Acknowledgments

The authors thank the participants, who generously gave their time to this study, and the medical staff for their assistance with recruitment.

ChatGPT (OpenAI) was used to support language editing during the preparation of this manuscript. The authors reviewed and edited all content and take full responsibility for the accuracy, originality, and integrity of the manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (72474191).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: AV, QY
Data curation: AV, JZ, WS, WX
Formal analysis: AV
Funding acquisition: QY
Investigation: AV, JZ, KJ
Methodology: AV, XD
Project administration: AV, QY
Supervision: QY
Writing – original draft: AV
Writing – review & editing: AV, QY, WS, WX, XD
All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]

XSL•FO
RenderX

2. Facchinetti G, Petrucci G, Albanesi B, De Marinis MG, Piredda M. Can smart home technologies help older adults manage their chronic condition? A systematic literature review. Int J Environ Res Public Health 2023 Jan 10;20(2):1205. [doi: 10.3390/ijerph20021205] [Medline: 36673957]
3. Aranha M, James K, Deasy C, Heavin C. Exploring the barriers and facilitators which influence mHealth adoption among older adults: a literature review. Gerontechnology 2021;20(2):1-16. [doi: 10.4017/gt.2021.20.2.424.06]
4. Soubelet A, Salthouse TA. Influence of social desirability on age differences in self-reports of mood and personality. J Pers 2011 Aug;79(4):741-762. [doi: 10.1111/j.1467-6494.2011.00700.x] [Medline: 21682727]
5. Greenwald AG, Banaji MR. Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychol Rev 1995 Jan;102(1):4-27. [doi: 10.1037/0033-295x.102.1.4] [Medline: 7878162]
6. Kurdi B, Seitchik AE, Axt JR, et al. Relationship between the Implicit Association Test and intergroup behavior: a meta-analysis. Am Psychol 2019;74(5):569-586. [doi: 10.1037/amp0000364] [Medline: 30550298]
7. Payne K, Lundberg K. The affect misattribution procedure: ten years of evidence on reliability, validity, and mechanisms. Soc Personal Psychol Compass 2014 Dec;8(12):672-686. [doi: 10.1111/spc3.12148]
8. Payne BK, Krosnick JA, Pasek J, Lelkes Y, Akhtar O, Tompson T. Implicit and explicit prejudice in the 2008 American presidential election. J Exp Soc Psychol 2010;46(2):367-374. [doi: 10.1016/j.jesp.2009.11.001]
9. Payne BK, Cheng CM, Govorun O, Stewart BD. An inkblot for attitudes: affect misattribution as implicit measurement. J Pers Soc Psychol 2005 Sep;89(3):277-293. [doi: 10.1037/0022-3514.89.3.277] [Medline: 16248714]
10. Vallesi A, Tronelli V, Lomi F, Pezzetta R. Age differences in sustained attention tasks: a meta-analysis. Psychon Bull Rev 2021 Dec;28(6):1755-1775. [doi: 10.3758/s13423-021-01908-x] [Medline: 33772477]
11. Schlaghecken F, Maylor EA. Motor control in old age: evidence of impaired low-level inhibition. J Gerontol B Psychol Sci Soc Sci 2005 May;60(3):158-161. [doi: 10.1093/geronb/60.3.p158] [Medline: 15860786]
12. Fölster M, Hess U, Werheid K. Facial age affects emotional expression decoding. Front Psychol 2014;5:30. [doi: 10.3389/fpsyg.2014.00030] [Medline: 24550859]
13. Hughes S, Cummins J, Hussey I. Effects on the affect misattribution procedure are strongly moderated by influence awareness. Behav Res Methods 2023 Jun;55(4):1558-1586. [doi: 10.3758/s13428-022-01879-4] [Medline: 35689167]
14. Kurdi B, Melnikoff DE, Hannay JW, et al. Testing the automaticity features of the affect misattribution procedure: the roles of awareness and intentionality. Behav Res Methods 2024 Apr;56(4):3161-3194. [doi: 10.3758/s13428-023-02291-2] [Medline: 38030926]
15. Katz JH, Mann TC, Shen X, Goncalo JA, Ferguson MJ. Implicit impressions of creative people: creativity evaluation in a stigmatized domain. Organ Behav Hum Decis Process 2022 Mar;169:104116. [doi: 10.1016/j.obhdp.2021.104116]
16. Mann TC, Cone J, Heggeseth B, Ferguson MJ. Updating implicit impressions: new evidence on intentionality and the affect misattribution procedure. J Pers Soc Psychol 2019 Mar;116(3):349-374. [doi: 10.1037/pspa0000146] [Medline: 30802132]
17. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
18. Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (IAPS): affective ratings of pictures and instruction manual. : University of Florida; 2005 URL: https://csea.phhp.ufl.edu/Media.html [accessed 2026-07-25]
19. Kurdi B, Lozano S, Banaji MR. Introducing the Open Affective Standardized Image Set (OASIS). Behav Res Methods 2017 Apr;49(2):457-470. [doi: 10.3758/s13428-016-0715-3] [Medline: 26907748]
20. Owsley C. Visual processing speed. Vision Res 2013 Sep 20;90:52-56. [doi: 10.1016/j.visres.2012.11.014] [Medline: 23231958]
21. Salthouse TA. The processing-speed theory of adult age differences in cognition. Psychol Rev 1996 Jul;103(3):403-428. [doi: 10.1037/0033-295x.103.3.403] [Medline: 8759042]
22. Hall WJ, Dawes HC, Hannay JW, et al. Development and initial validation of the implicit internalized sexual orientation stigma affect misattribution procedure. Front Psychol 2024;15:1385410. [doi: 10.3389/fpsyg.2024.1385410] [Medline: 39679152]
23. Forstmann BU, Tittgemeyer M, Wagenmakers EJ, Derrfuss J, Imperati D, Brown S. The speed-accuracy tradeoff in the elderly brain: a structural model-based approach. J Neurosci 2011 Nov 23;31(47):17242-17249. [doi: 10.1523/JNEUROSCI.0309-11.2011] [Medline: 22114290]
24. Pillaud N, Ric F. The evaluation-behavior link revisited: it depends on the question you have in mind. Cognition 2025 Jun;259:106097. [doi: 10.1016/j.cognition.2025.106097] [Medline: 40043637]
25. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
26. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view1. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
27. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
28. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]

XSL•FO
RenderX

29. Hepburn J, Williams L, McCann L. Barriers to and facilitators of digital health technology adoption among older adults with chronic diseases: updated systematic review. JMIR Aging 2025 Sep 11;8:e80000. [doi: 10.2196/80000] [Medline: 40934502]
30. Ram S. Successful innovation using strategies to reduce consumer resistance: an empirical test. J Prod Innov Manag 1989 Mar;6(1):20-34. [doi: 10.1111/1540-5885.610020]
31. Birati Y, Tzemah-Shahar R. Barriers to digital health adoption in older adults: scoping review informed by Innovation Resistance Theory. J Med Internet Res 2026 Feb 2;28:e75591. [doi: 10.2196/75591] [Medline: 41627850]
32. Eder AB, Deutsch R. Watch the target! Effects in the affective misattribution procedure become weaker (but not eliminated) when participants are motivated to provide accurate responses to the target. Front Psychol 2015;6:1442. [doi: 10.3389/fpsyg.2015.01442] [Medline: 26441807]

Abbreviations

AMP: affect misattribution procedure
GLMM: generalized linear mixed-effects model
IA-AMP: influence-aware affect misattribution procedure
IAPS: International Affective Picture System
ICC: intraclass correlation coefficient
OASIS: Open Affective Standardized Image Set
OR: odds ratio

Edited by M Antoniou; submitted 18.Nov.2025; peer-reviewed by A Corlateanu, P Watson; revised version received 17.Jun.2026; accepted 03.Jul.2026; published 18.Aug.2026.

Please cite as:
Vinnikova A, Zhan J, Jin K, Ding X, Sang W, Xu W, Yang Q
Measuring Implicit Attitudes Toward Digital Health Technologies in Older Adults Using an Influence-Aware Affect Misattribution Procedure: Development and Feasibility Study
JMIR Aging 2026;9:e87989
URL: https://aging.jmir.org/2026/1/e87989
doi:10.2196/87989

© Anna Vinnikova, Jinyue Zhan, Kailing Jin, Xianfeng Ding, Wei Sang, Weina Xu, Qian Yang. Originally published in JMIR Aging (https://aging.jmir.org), 18.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Evaluation of a Transportation Planning e-Tool for Older Adults: Co-Design Study

Sara Bahrampoor Givi[1,2], PhD; Dany Baillargeon[3,4], PhD; Sahar Tahir[3,5], MSc; Bessam Abdulrazak[3,5], PhD; Mireille Gagnon-Roy[1,2], OTR, PhD; Mélanie Levasseur[2,3], OTR, PhD; Nathalie Delli-Colli[3,6], PhD; Patrick Boissy[3,7], PhD; Mélisa Audet[8,9], PhD; Véronique Provencher[1,2], OTR, PhD

[1]Research Center Charles-Le Moyne, Longueuil, QC, Canada

[2]School of Rehabilitation, Faculty of Medicine and Health Sciences, Université de Sherbrooke, 3001, 12e Avenue Nord, Sherbrooke, QC, Canada

[3]Research Center on Aging, Centre intégré universitaire de santé et de services sociaux de l'Estrie—Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, QC, Canada

[4]Department of Communication, Université de Sherbrooke, Sherbrooke, QC, Canada

[5]Department of Computer Science, AMI-Lab, Faculty of Science, Université de Sherbrooke, Sherbrooke, QC, Canada

[6]School of Social Work, Université de Sherbrooke, Sherbrooke, QC, Canada

[7]Department of Surgery—Orthopedics, Université de Sherbrooke, Sherbrooke, QC, Canada

[8]Centre collégial d'expertise en gérontologie du Cégep de Drummondville, Drummondville, QC, Canada

[9]Faculty of Letters and Human Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada

Corresponding Author:
Sara Bahrampoor Givi, PhD
Research Center Charles-Le Moyne, Longueuil, QC, Canada

Abstract

Background: Older adults often face challenges using digital transportation planning tools to engage in meaningful activities. In Sherbrooke, a Canadian city located in the province of Québec, hilly terrain, bus-based mobility, and varied digital literacy among older adults highlight the need for accessible transportation planning technologies. Despite the growing demand for accessible and inclusive technologies, interfaces are often not tailored to the specific needs and preferences of older adults, making them difficult to use. Addressing this gap is essential to support the mobility and social participation of older adults in their communities.

Objective: This study aimed to enhance older adults' experience with a transportation planning e-tool—called Mobilainés—by co-designing its interface and evaluating its usability and user experience in line with their needs and preferences.

Methods: Using a participatory design approach, the study encompassed both development and evaluation phases, integrating qualitative and quantitative methods. In the development phase, heuristic evaluation, along with the Mobilaînés mock-up, was performed to identify the percentage of compatibility with age-friendly usability design principles and guidelines. Fourteen co-design workshops with the steering committee and research team informed persona development and core functionalities, while 3 workshops with older adults refined the mock-ups. For the evaluation phase, 2 laboratory usability tests (simulated context) and 1 in-field test (real-world context) were conducted with 5, 7, and 4 older adults, respectively. Data were collected using a combination of objective performance metrics recorded in real time by a computer and subjective user feedback, gathered using the System Usability Scale, the Unified Theory of Acceptance and Use of Technology questionnaires, and focus group discussions. For the in-field testing, user experience data were obtained via the User Experience Questionnaire and semistructured interviews. Quantitative data were analyzed descriptively, and qualitative data were analyzed thematically.

Results: The participatory design and iterative evaluation process identified key design decisions for age-friendly transportation planning, including step-by-step navigation, simplified task flow, clear feedback, visual clarity, and personalization. In the final laboratory usability test, the Mobilaînés e-tool showed 84% compatibility with age-friendly guidelines, a 97% task completion rate, and System Usability Scale and Unified Theory of Acceptance and Use of Technology scores corresponding to very good usability. In-field User Experience Questionnaire results and interviews suggested that participants perceived the e-tool as reliable, easy to use, and accurate, despite the lack of real-time data.

Conclusions: Participatory design and iterative usability testing revealed that older adults' experience with Mobilainés transportation planning e-tools is improved. These design learnings contributed to improved usability and confidence in real-world travel planning, offering practical guidance for the development of age-friendly digital transportation tools. Future research should explore collaborative and technological strategies to support its adoption and sustained use, particularly among older adults with limited digital literacy.

XSL•FO
RenderX

(JMIR Aging 2026;9:e88297) doi:10.2196/88297

KEYWORDS

older adults; user experience; participatory design; e-tool; transportation

Introduction

Background

Worldwide, the aging population will increase from 900 million (2015) to 2 billion by 2050 [1]. In Canada, it is expected to reach 15.1 million by 2063 [2]. Age-related changes, such as declining vision [3], reduced mobility [4], and cognitive decline, may hinder traffic signs recognition, decrease flexibility, and slow reaction times while driving [5], which can limit older adults' ability to drive safely and travel independently [6]. Additionally, financial constraints [7], rural living [8], and small social networks [9] further restrict access to transportation options. These challenges increase the risk of sedentary behavior, isolation [10,11], loneliness [11], and anxiety [12]. To enable older adults to maintain their independence and engage in community life—key elements of social participation [13]—it is therefore essential to strengthen older adults' ability to access and use transportation options that reflect their specific needs, including those involving emerging technologies.

Digital transportation planning tools (websites and apps) are intended to support users in accessing information, selecting transportation options, and organizing their travel. While these tools have the potential to improve mobility, they often pose considerable usability challenges for older adults. Age-related changes, such as decreased visual acuity [14], reduced contrast sensitivity [15], and diminished dexterity [16], can make it difficult for them to identify important interface elements, interpret maps, and accurately interact with touch or clickable components [17]. In addition to these sensory and motor declines [4], cognitive changes associated with aging, such as reduced executive functioning, can impair an individual's ability to navigate complex transportation systems or coordinate multistep travel plans [6]. Digital literacy [18] also remains a key barrier; limited familiarity with technology can restrict access to real-time transportation updates or adaptive features within e-tools [19]. Differences in numerical comprehension and screen navigation further emphasize the need for design solutions that reflect the diverse abilities of this population [20].

Beyond usability considerations, digital transportation planning tools may influence older adults' mobility and well-being through psychological and cognitive mechanisms. Accessible and intuitive interfaces can reduce cognitive demands, helping users process information more efficiently and make travel decisions with less effort [21]. This may enhance perceived control over travel and strengthen mobility-related self-efficacy, understood as confidence in one's ability to plan and complete trips, particularly because older adults' transition to public transit use involves acquiring travel-planning skills and confidence through experience [22]. Together, greater perceived control and self-efficacy may encourage more frequent and independent use of transportation options, supporting social participation and well-being. Conversely, poorly designed interfaces may increase cognitive burden, reduce confidence, and discourage engagement with transportation systems [23].

Addressing such specific needs of older adults is vital for creating inclusive tools that enhance their user experience and support independent mobility. A participatory design approach is especially valuable for ensuring that these needs are fully integrated into the development and evaluation of transportation planning interfaces. Participatory design, sometimes known as cocreation or co-design [24], is described as a collaborative process in which various participants engage in shared exploration, understanding, and learning through ongoing reflection and action [25]. Among other benefits, participatory design contributes to a better user experience, as it is conceived with the end users in mind rather than simply tested by them [26]. User experience is defined as "an iterative set of decisions leading to a successful outcome with an interactive tool, as well as a productive and satisfying process to achieve that outcome" [27]. To enhance the user experience of older adults when using a travel planning e-tool, it is important to design an interface that is specifically adapted to their specific needs and allows them to perform their intended tasks easily, intuitively, and without confusion or interruption [28]. Given the notable lack of studies specifically focused on developing and evaluating transport planning interfaces tailored to the needs and preferences of older adults through a participatory design approach [28], addressing this gap is essential for enhancing their user experience with transportation planning e-tools.

Context of Study and Objective

The aim of this study is to improve older adults' experience in using a transportation planning e-tool—called Mobilaînés—by co-designing its interface and evaluating its usability and user experience based on their needs and preferences. Rather than introducing a new mode of transportation, Mobilaînés [29] was designed to integrate existing options, helping older adults make informed travel decisions and supporting their autonomy in everyday life, using a living laboratory methodology [30]. The study took place in Sherbrooke, a city in the province of Quebec, Canada, characterized by its hilly terrain, and its main public transportation is an urban bus service. These environmental characteristics may create mobility challenges for older adults, especially those with reduced endurance, mobility limitations, or unfamiliarity with public transit. In addition, despite increasing digital technology use among older adults in Canada, digital literacy, confidence, and access remain uneven, particularly among those with functional or social vulnerabilities [31]. This context highlights the need for transportation planning tools adapted to older adults' capabilities and preferences. The Mobilaînés living laboratory methodology, as well as its algorithm development, is described elsewhere (see Provencher et al [29] and Tahir et al [32] for additional details). The study is guided by the following research question: How does the co-designed Mobilaînés transportation planning interface improve usability and user experience for older adults?

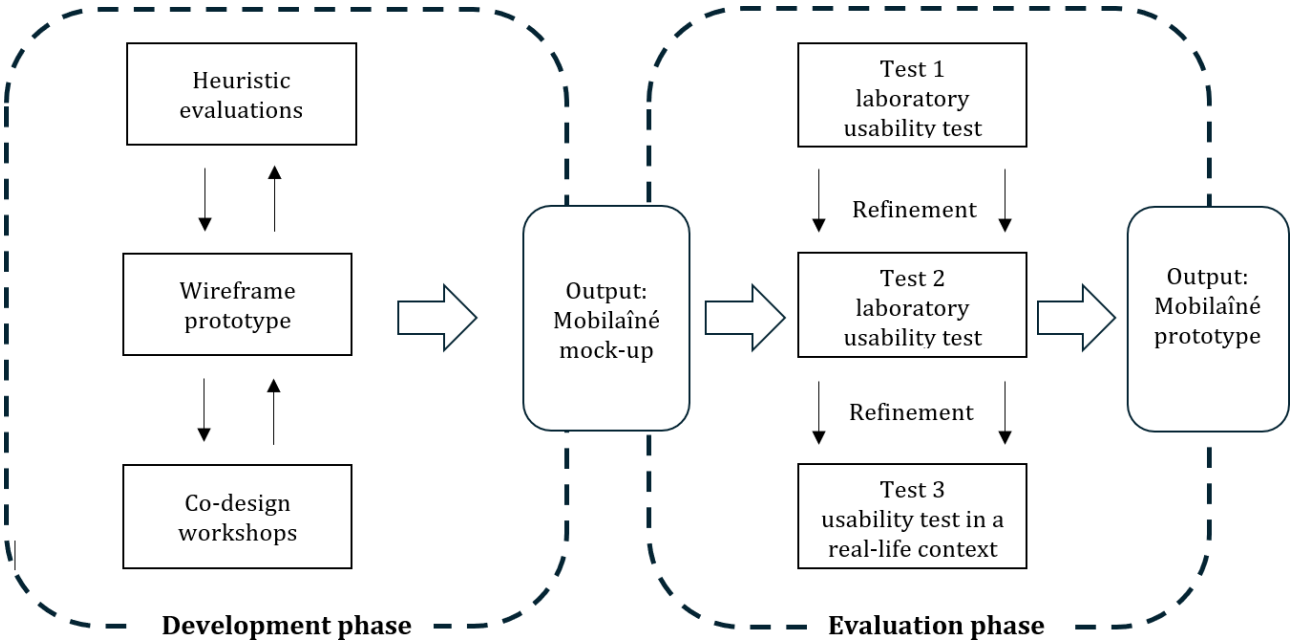
XSL•FO
RenderX

Methods

Study Design

Using a participatory design approach [29], the study had 2 phases using qualitative and quantitative methods: the development phase (May 2020 to July 2022), which focused on the co-design process to ensure that older adults' needs and preferences were considered, and the evaluation phase (August 2022 to August 2024), which examined the usability and user experience of the interface prototype to ensure alignment of the e-tool with the expectations of older adult users (Figure 1).

Figure 1. The development and evaluation process of the Mobilaînés e-tool.



Development Phase

Overview

To co-design the interface of a transportation planning e-tool for older adults, we conducted a heuristic evaluation of existing transportation planning e-tools and the Mobilaînés mock-up, and we carried out a total of 17 co-design workshops with (1) the steering committee, including representatives from municipal and senior advocacy organizations (n=8); (2) the research team, comprising members from communication, computer science, and gerontology (n=10); and (3) older adults (n=5-10).

Heuristic Evaluation

Heuristic evaluation is a widely used expert-based method in which evaluators inspect an interface against a set of usability heuristics to identify design flaws [33], such as supporting *undo* and *redo*. We used the framework of Kumar et al [34], which includes the following phases:

Planning

Transportation planning e-tools that met core functional requirements, such as time autonomy (eg, customizable departure times) and weather integration, were selected in collaboration with the research team [35]. Three websites and 3 mobile apps (Android) were included: the Vermeille App and Website (the transportation planning e-tool provided by Sherbrooke Public Transit at the time of the study), the Google Maps [36] App and Website, the Moovit [37] App, and the OC Transpo [38] Website. The goal was to support the co-design of the Mobilaînés e-tool by documenting the strengths and weaknesses of these transport planning platforms and assessing their alignment with the age-friendly usability design principles and guidelines identified through a scoping review [39]. For additional details on the guidelines, see Multimedia Appendix 1.

Conducting

The heuristic evaluation was initially conducted by an expert evaluator (SBG), who reviewed each e-tool against the age-friendly usability design principles and guidelines identified in the scoping review. Each platform was used repeatedly and reviewed twice to explore its features, functionalities, and interaction pathways. Screenshots of each interface were captured and used during the heuristic evaluation. Findings were documented in structured notes by mapping interface elements to the corresponding heuristics, recording each guideline as compatible or noncompatible, and documenting each issue's location, related heuristic, and potential impact on older adults' ability to understand, navigate, or complete the task. The results were then reviewed and discussed within the research team. Identified usability issues were presented and justified using the heuristic framework, allowing collective reflection and validation of findings and helping refine the interpretation of heuristic violations and compatibility across tools.

Reporting

The identified compatibility and noncompatibility with heuristics were categorized and ranked by severity (minor, moderate, and major). Compatible features were summarized in a table highlighting each tool's strengths, while noncompatible features were presented in a separate table listing usability weaknesses. The evaluation was conducted on both desktop (websites) and

XSL•FO
RenderX

mobile (Android) platforms. Strengths and weaknesses of each e-tool were documented, with particular attention paid to compatibility with age-friendly principles. Recommendations for improvement were then proposed.

Based on the findings, 3 mock-ups for the Mobilaînés website were produced, reviewed with the research team and older adults during co-design activities. Their feedback informed the selection and refinement of the mock-up that best aligned with older adults' needs and preferences, user experience considerations, and the travel-planning algorithm [28,35]. The selected mock-up was then evaluated heuristically by SBG using the same framework [34], with a specific focus on age-friendly usability principles to ensure alignment and identify areas for improvement. The evaluator reviewed the guidelines and assessed whether they were consistently applied across all dimensions of the e-tool.

Co-Design Workshops

Workshops Overview

The development of the Mobilaînés e-tool was grounded in cocreation principles, particularly through workshops that brought together multiple stakeholders (eg, older adults, researchers, municipal representatives, and senior advocacy organizations) using a participatory design [28]. The methods used in the Mobilaînés co-design workshops were informed by Sanders' [40] participatory design framework, and focused on engaging participants in generating, discussing, and refining design ideas. This approach supports the creation of concrete and actionable outputs, such as mock-ups, to inform subsequent development phases [41]. The cocreation workshops involved (1) with the steering committee to develop older adults' personas, (2) with the research team to identify the functionalities of the Mobilaînés mock-up, and (3) with older adults. Lived experience contributors, including older adults and community representatives, played an advisory and co-design role throughout this phase. They contributed by sharing mobility experiences, reviewing personas and mock-ups, identifying usability barriers, and helping prioritize design needs. Their contributions informed the development and refinement of the Mobilaînés interface.

With the Steering Committee to Develop Older Adult Personas

The steering committee guided persona development using the Persona Lifecycle framework [42]. Preliminary personas were developed from semistructured interviews with 6 older adults [29] and a review of scientific and gray literature on aging, mobility, and digital inclusion. Two coconstruction workshops with stakeholders, researchers, municipal representatives, and senior advocacy organizations defined 2 dimensions: mobility profile and planning profile. Mobility profile reflected travel frequency and physical, cognitive, and financial ability to move around, resulting in 4 categories: Mobile Active, Limited Active, Limited Isolated, and Mobile Isolated. Planning profile reflected autonomy in trip planning and access to planning tools, including digital literacy and internet access, resulting in 4 categories: Autonomous Full-Access, Autonomous Limited Access, Dependent Limited Access, and Dependent Full Access. These profiles captured diverse transportation and trip-planning needs and guided recruitment, prototype development, and usage scenarios.

With Research Team to Identify the Functionalities of the Mobilaînés Mock-Up

The scenario-based design technique [43] was used to illustrate common travel-planning issues faced by older adults. This approach incorporated personas and "what if" reasoning to explore both the positive and negative consequences of design choices, helping to assess not only current successes but also the risks of potential changes [43]. In this regard, 3 brainstorming sessions with project team members were conducted using the scenario-based usability engineering technique [44] to (1) identify key functionalities of the e-tool, such as (i) displaying different routes, (ii) providing information on transport options, (iii) presenting travel details (eg, weather, nearby benches, and toilets), and (iv) adapting trip planning to user profiles (eg, walker users); and (2) validating the feasibility of the scenarios. In addition, 9 cocreation workshops were conducted to collaboratively design a preliminary interface for the Mobilaînés e-tool with support from an external agency. Each session involved 8-12 participants, including members of the interdisciplinary Mobilaînés team (from computer science, communication, and gerontology) and an older adult who was a member of the steering committee. Finally, insights from both the co-design sessions and the heuristic evaluation were presented to the research team, informing specific improvements to the selected mock-up ahead of the first usability test.

With Older Adults

Three in-person cocreation workshops were conducted with older adults, with 6-12 participants per session. Eligible participants were older than 65 years, lived in Sherbrooke, and used a computer at least weekly; those unable to travel without human assistance were excluded. Recruitment occurred through municipal and project partners, seniors' residences, posters, and online platforms. Participants were contacted by email and phone to explain the study and assess eligibility and then selected to reflect previously developed personas varying in age, gender, mobility profile, digital competence, travel frequency, and e-tool use. The workshops documented older adults' needs and preferences regarding the Mobilaînés interface and age-friendly features, such as customizable departure times. Feedback on the selected mock-up was documented, thematically analyzed by a research assistant, and validated by 2 principal investigators (VP and DB). Results were integrated into co-design sessions with the research team and design agency to discuss travel-planning challenges and improve the mock-up. The scenarios and mock-up were later evaluated twice by older adults to assess alignment with their expectations. Adjustments were made based on feedback, and perceptions of the interface and functionality were systematically recorded.

All workshops followed a structured but flexible format adapted to session objectives. Steering committee and research team workshops generally lasted 1-2 hours, whereas older adult workshops lasted approximately 3 hours. Sessions typically included (1) presentation of objectives and materials, such as personas, scenarios, or mock-up screens; (2) guided discussion

XSL•FO
RenderX

of transportation planning needs, usability barriers, and expectations; (3) collaborative activities, such as brainstorming, scenario-based discussion, persona review, mock-up walkthroughs, or prioritization of design issues; and (4) synthesis of key decisions and next steps. Outputs included refined personas, prioritized functionalities, user scenarios, annotated mock-ups, usability issues, and design recommendations communicated to the design, programming, and research teams.

Evaluation Phase

Overview

The evaluation comprised 3 experiments: 2 laboratory usability tests and 1 in-field test. According to Nielsen [33], a test with 5 users can reveal 80% of usability issues. Faulkner [45] found that groups of 10 users uncovered 95% of usability issues. Considering possible dropouts, the target sample size was 10, with a minimum of 5 to still capture approximately 80% of practical issues. The same inclusion and exclusion criteria as the co-design workshops were applied, and new participants were recruited for each test to minimize bias. A pretest was conducted with an older adult (aged 65+ years) to ensure that all test materials (eg, scenario tasks, questionnaires, and interview guide) were clear and easy to understand.

Usability Evaluation in a Laboratory Setting

Overview of Usability Evaluation

First, participants (n=7 in test 1; n=5 in test 2) tested the Mobilaînés prototype by completing 3 scenario-based tasks designed to evaluate its usability. These tasks required navigating the interface and entering travel preferences (eg, requesting assistance at the destination). Following guidance on developing usability scenarios and tasks [46], the tasks were designed to reflect realistic travel-planning activities, specify what participants were expected to accomplish, and avoid unnecessary cues about how to use the interface:

1. Task 1: planning a trip while customizing preferences to become familiar with the platform.
2. Task 2: creating and personalizing a user profile, such as saving favorite destinations, to explore additional features.
3. Task 3: planning a future round trip using all the features of the prototype, including both trip directions, save or print options, and personalized settings.

Objective Performance Metrics

The objective performance of participants was assessed using performance metrics to evaluate error tolerance, efficiency, and effectiveness. Observed errors were used to assess error tolerance, task duration measured efficiency (ie, the speed and productivity of task completion), and the number of completed tasks evaluated effectiveness (ie, the success rate of task completion aligned with user expectations). To ensure consistent testing conditions, 3 predefined levels of assistance were provided for participants needing help:

1. Level 1: identifying the problem and clarifying the task.
2. Level 2: providing a clue about the required information and where to find it.
3. Level 3: offering a partial solution by indicating where the information is located on the website.

Assistants (n=4) monitored screens and addressed inquiries according to these guidelines. For each subtask, the outcome was recorded as one of the following: (1) completed without assistance, (2) completed with level 1 assistance, (3) completed with level 2 assistance, (4) completed with level 3 assistance, or (5) not completed. A task or subtask was considered completed when the expected outcome was reached, independently or with predefined assistance. The task completion rate was calculated as completed tasks or subtasks divided by total assigned tasks or subtasks and expressed as a percentage; independent and assisted completion were reported separately. Microsoft Teams recordings of participants' screens and audio were reviewed multiple times to document all observable errors across tasks and subtasks. Errors were summed across participants, tasks, and subtasks for each laboratory usability test. As there was no predefined maximum number of errors per participant or task, participants could make multiple errors and still complete a task independently. All observed errors were documented to capture ease of use, navigation challenges, and alignment with age-friendly usability guidelines. To account for different sample sizes, errors were reported as total counts and observed errors per participant.

Subjective Opinion (User Feedback)

Following task completion, participants completed the validated French version of the System Usability Scale (SUS) [47] to assess perceived usability and satisfaction (Multimedia Appendix 2). This questionnaire demonstrated good convergent construct validity with a single-item measure of user-friendliness (r=0.806) and high internal consistency (Cronbach $\alpha \geq 0.90$) [48]. To evaluate actual usability, participants also completed the Unified Theory of Acceptance and Use of Technology (UTAUT) questionnaire (Multimedia Appendix 3) [49] to measure perceived performance expectancy, effort expectancy, facilitating conditions, and social influence. A customized Mobilaînés UTAUT model was developed to evaluate the effect of these constructs on Behavioral Intention and Behavioral Use (Multimedia Appendix 4), with content tailored to the project.

Subsequently, all participants took part in a focus group to reflect on their satisfaction with sensory navigation, information, and design. Focus group discussions explored participants' overall experience with the Mobilaînés website, elements that enhanced or hindered satisfaction, points of confusion during task completion, comparisons with other planning tools, and their expectations (Multimedia Appendix 5). After analyzing and synthesizing the results of the first laboratory usability test, the e-tool was modified, and all testing procedures were repeated in a second round of usability testing to refine the prototype. The main modifications implemented between the first and second laboratory usability tests included clearer feedback and confirmation messages, improved step indicators and navigation cues, revised icons and colors to enhance visual clarity and consistency, simplified password and account creation guidance, and refined map icons and route-related information to support interpretation of travel options.

XSL•FO
RenderX

Usability Evaluation in a Real-Life Setting

The effects of using the Mobilaînés prototype on older adults' participation in meaningful activities in real-life contexts were explored. Participants (n=4), recruited using the same criteria as in previous usability tests, used the prototype over 6 weeks. Each completed a 1-hour individual introductory session, during which key functions, including account creation, trip planning, saving, and printing, were demonstrated and practiced. Two tutorial videos in French and English were uploaded to the Mobilaînés YouTube channel and emailed to participants to support use and personalization. In weeks 1, 3, and 5, participants were asked by phone to plan trips to 3 Sherbrooke locations; in weeks 2, 4, and 6, the user experience expert followed up by phone to encourage use and address difficulties. Participants were asked to use the tool at least once per week. In weeks 1 and 5, participants completed a paper version of the User Experience Questionnaire (UEQ; Multimedia Appendix 5) to assess satisfaction with the e-tool and changes in user experience over time [50]. The UEQ uses a 7-point Likert scale to assess 6 dimensions: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty [51] (Multimedia Appendix 6). It has acceptable-to-good reliability (Cronbach α=0.77 - 0.92) and construct validity, with correlations between pragmatic and hedonic factors of r=0.42 (P<.001), and attractiveness correlating strongly with pragmatic (r=0.82) and hedonic (r=0.65) user experience aspects (P<.001). After 6 weeks, participants were invited to final interviews about their use of Mobilaînés, its support or lack of support for meaningful activities, and barriers and facilitators in travel planning and daily participation (Multimedia Appendix 7).

Data Analysis

To address the first objective, selected e-tools (Vermeille app or website, Google Maps app or website, Moovit app, and OC Transpo website) were heuristically evaluated for compatibility with age-friendly usability guidelines. Compatibility was assessed descriptively across phases using the same age-friendly usability guideline framework. Applicable guidelines were classified as compatible or noncompatible, and compatibility percentages were calculated as the number of compatible guidelines divided by total applicable guidelines, multiplied by 100. Using the same framework and calculation method allowed for descriptive comparison across phases. However, the evidence used to assess compatibility differed by phase: during development, compatibility was assessed through expert heuristic evaluation, whereas during laboratory and in-field evaluations, observed usability issues, participant feedback, and evaluation findings were mapped to the same framework. An iterative deductive thematic analysis [52] was conducted by SBG on the verbatim transcripts from the cocreation workshops with older adults, using the age-friendly usability design principles and guidelines [39].

To address the second objective, descriptive analyses of participant characteristics, performance metrics, and SUS, UTAUT, and UEQ scores were conducted in SPSS (IBM Corp). SUS scores below 60 were considered unacceptable, and scores of 90 or above were considered excellent [48]. For UTAUT, a score of 3 on the 5-point Likert scale was treated as the acceptable reference value for each construct item. Focus groups and interviews were recorded, transcribed, imported into NVivo (version 20; Lumivero), and analyzed by SBG using constant comparison, deductively for usability tests and inductively for the in-field test, through open, axial, and selective coding [53]. Given the small sample sizes in the laboratory usability tests and in-field test, all quantitative findings were interpreted descriptively, and no inferential comparisons were conducted between tests. Mean and SD values were retained for questionnaire scores to remain consistent with common reporting practices for the SUS, UTAUT, and UEQ; however, these values were interpreted cautiously.

Ethical Considerations

Ethical approval was obtained from the CIUSSS de l'Estrie-CHUS (2020 - 3389: MOBILAÎNÉS). All participants provided written informed consent after being informed about the study aims and procedures. To protect privacy and confidentiality, data were coded and deidentified during collection, analysis, and reporting. Consent forms and the code key were kept securely in a locked location by the researcher in charge. Research data, including audio and video recordings, were stored confidentially in a locked location for 5 years by the principal investigator (VP). Participants in co-design workshops and laboratory usability tests received US $15 per session, with parking reimbursement when applicable; in-field test participants received US $75 each.

Results

Development Phase

Heuristic Evaluation

A comparison of 3 websites (Vermeille, Google Maps, and OC Transpo) showed that Google Maps had the strongest adherence to age-friendly usability design guidelines (see Multimedia Appendix 8 for details). The heuristic evaluation of the Mobilaînés mock-up showed 67% compatibility with age-friendly usability guidelines. While most information-related guidelines were followed, inconsistencies were noted in key principles such as Simplicity, Control, and Feedback. Specific issues included poor prioritization, repetitive actions, an excessive number of options per page, insufficient color contrast, a small font size, the presence of scroll bars, and a lack of step-by-step guidance. Only 1 guideline related to the Ease-of-Use principle (target size) was not met.

Co-Design Workshops

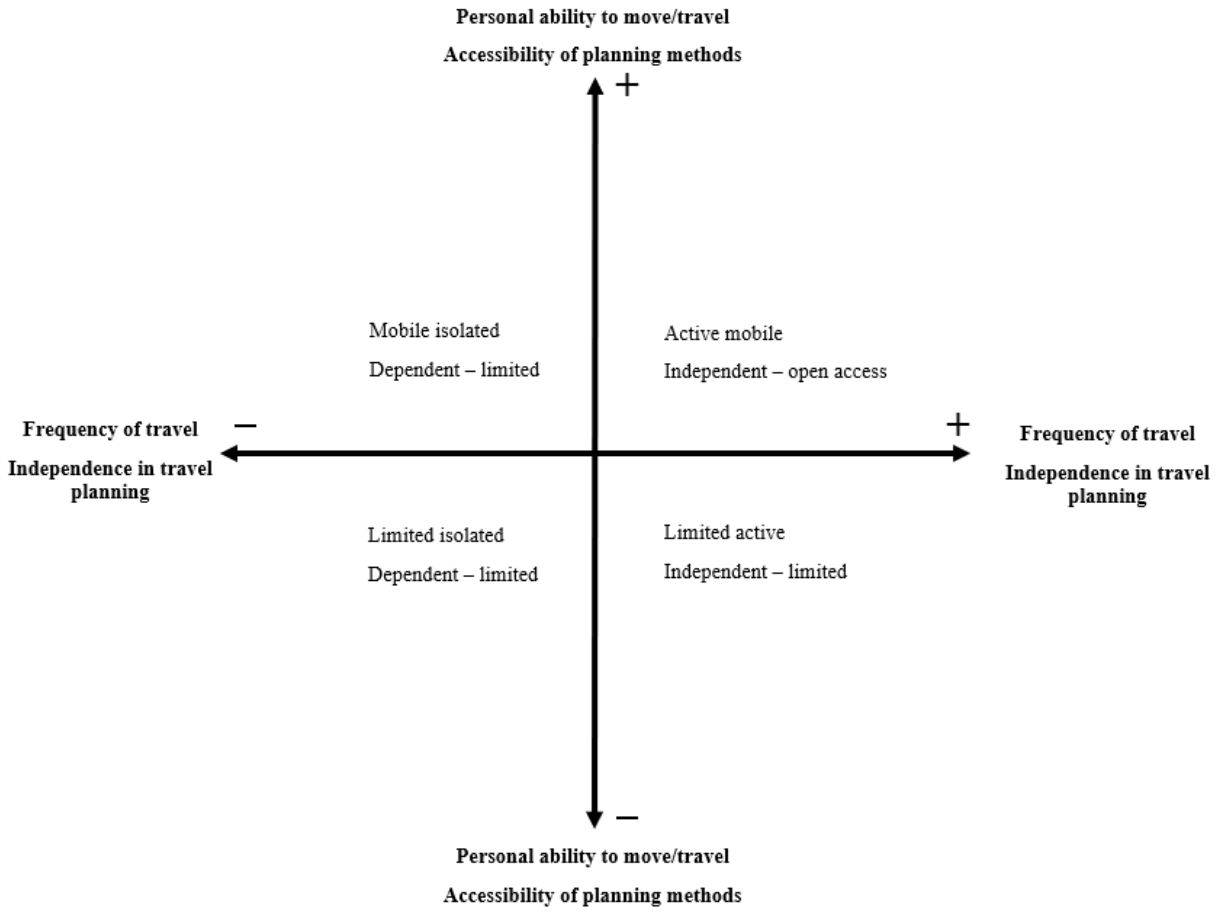
With the Steering Committee to Develop Older Adults Persona

Four personas were created to reflect older adults' diverse mobility needs, abilities, contexts, and levels of digital literacy (Figure 2). Solange, a 72-year-old woman, is independent, physically active, and confident using digital tools for trip planning, representing users with high mobility and digital confidence. Léonard, a 65-year-old man with diabetes-related functional limitations, has reduced driving ability and limited digital literacy despite computer access. René, an 80-year-old man newly using a tablet provided by his family, depends mostly

XSL•FO
RenderX

on others for outings but is motivated to gain independence, representing users with limited experience but willingness to learn. Suzanna, a 78-year-old recent immigrant with limited French proficiency, no social support, financial constraints, pain, and unfamiliarity with public transport, represents users with intersecting linguistic, physical, economic, and digital vulnerabilities. Solange, Léonard, and René were most relevant to Mobilaînés because they could leave home and plan trips with access to technological support. René was specifically used during heuristic evaluations to assess usability for users facing greater challenges. Suzanna could not be represented in the study because she lacked access to digital tools.

Figure 2. Development of personas based on older adults' mobility and trip-planning profiles.



With Research Teams to Identify the Functionalities of the Mobilaînés Mock-Up

The scenario-based design technique guided the definition of the Mobilaînés mock-up features. Based on personas and older adults' co-design feedback, the tool incorporated sheltered bus stops, restrooms, benches, low-light conditions, round-trip and future-trip planning, walking-time adjustment, hill avoidance, profile creation, favorite addresses, and travel history. To improve accessibility and reduce memory load, travel planning was divided into single-task steps, with clear information and feedback.

With Older Adults

Thematic analysis of older adults' co-design feedback highlighted key usability considerations, including personalized travel planning through user profiles and simple, consistent icons, colors, and layouts. Visual clarity was a recurring concern, including font size, color contrast, and distinguishing clickable elements from informational text, such as the Mobilaînés icon and website name. Participants emphasized reducing unnecessary steps in travel planning and profile creation, improving navigation through forward or back options and map zoom, and adding text labels for unclear icons, such as the pencil icon for editing. They also noted that too many route options increased memory load and preferred highlighting 1 selected route in color. Other challenges included understanding path details and using the travel preference personalization page, which had excessive text and small icons. Finally, participants stressed the need for clear feedback after completing a plan and suggested adding a round-trip option at the start.

Synthesis of the Results: Heuristic Evaluation and Co-Design Workshops

Workshop themes from older adults were integrated with mock-up heuristic evaluation results and refined during a Mobilaînés project meeting. Usability issues, including color combinations, font-size buttons, and page layout, were reviewed

XSL•FO
RenderX

using older adults' feedback and age-friendly usability guidelines. SBG documented improvement ideas, and final decisions were based on feasibility and communicated to the interface design, programming, and algorithm teams to prepare the prototype for laboratory usability testing. Some preferences, such as avoiding icy routes, could not be implemented because of limited data quality and development time.

Evaluation

Laboratory Usability Tests

Sociodemographic Overview

In the first usability test, 10 participants were recruited, but 3 were absent due to health reasons, resulting in 7 participants: 5 corresponding to Solange's persona and 2 to Léonard's. Participants were aged between 65 and 78 (mean 71.86, SD 4.85; median 70) years. Four participants identified as female, and 6 held a driver's license. In the second usability test, out of 10 recruited participants, 5 were absent on the testing day, also due to health issues, resulting in 5 participants, all related to Solange's persona. Participants were aged between 66 and 75 (mean 69.4, SD 3.36; median 69) years. In the in-field test, 4 participants aged 67-75 (mean 69.5, SD 3.79; median 68) years were recruited (see Table 1 for more details). Participants were identified using anonymized codes, where "P" refers to the participant number and "T" indicates the test session.

XSL•FO
RenderX

Table . Sociodemographic information of participants in usability tests.

Participant	Age (years)	Sex	Driving license	Means of transport	Social contact[a]	Education[b]	Revenue[c]	Health problems	Walking difficulty
P1 (T1)	78	Female	No	Bus, walking	2	2	2	Yes	No
P2 (T1)	69	Female	Yes	Car, walking	1	3	2	No	No
P3 (T1)	70	Female	Yes	Car, walking	1	2	2	Yes	No
P4 (T1)	75	Male	Yes	Car, bus	1	2	2	No	No
P5 (T1)	77	Female	Yes	Car, walking	2	2	3	No	No
P6 (T1)	69	Male	Yes	Bus, walking	1	1	2	No	No
P7 (T1)	65	Male	Yes	Bus, walking	2	1	1	Yes	Yes
P1 (T2)	66	Male	No	Adapted transport	1	1	4	No	Yes
P2 (T2)	69	Female	No	Bus, walking, taxi	1	1	5	No	No
P3 (T2)	68	Male	Yes	Car	1	1	2	No	Yes
P4 (T2)	75	Female	Yes	Adapted, walking	2	1	5	No	No
P5 (T2)	69	Male	Yes	Car, bus, walking	1	3	2	No	No
P1 (T3)	69	Female	Yes	Car, walking	1	3	2	No	No
P2 (T3)	67	Male	No	Bus, bicycle	1	3	1	No	No
P3 (T3)	67	Male	Yes	Car, car sharing	2	1	3	No	No
P4 (T3)	75	Female	Yes	Car, car sharing	1	2	2	No	No
Mean	• Mean: 70.5 SD: 4.5 Median: 69	• Female: 50% (N=8) • Male: 50% (N=8)	• Yes: 75% (N=12) • No: 25% (N=4)	• Car: 26.7% • Walking: 30% • Bus: 23.3% • Adapted transport: 6.7% • Taxi: 3.3% • Bicycle: 3.3% • Car sharing: 6.7%	• 1: 68.8% (N=11) • 2: 31.2% (N=5)	• 1: 35% (N=7) • 2: 31% (N=5) • 3: 25% (N=4)	• 1: 12.5% (N=2) • 2: 56% (N=9) • 3: 12.5% (N=2) • 4: 6.5% (N=1) • 5: 12.5% (N=2)	• Yes: 19% (N=3) • No: 81 % (N=13)	• Yes: 19% (N=3) • No: 81 % (N=13)

[a]1: more than once a week, 2: at least once a month.

[b]1: university, 2: college, 3: secondary.

[c]1: US $7k-$14k, 2: US $15k-$29k, 3: US $30k-$44k, 4: >US $59k, 5: refused.

Performance Metrics in Laboratory Usability Tests

Navigation, selection, insertion, and task-understanding errors were identified in both tests and used to assess error tolerance. In total, 46 errors were recorded in test 1 (n=7) and 22 in test 2 (n=5), corresponding to 6.6 and 4.4 errors per participant, respectively (Table 2). Most errors occurred during profile

XSL•FO
RenderX

creation or personalization and round-trip planning. Because sample sizes differed between laboratory usability tests, errors were reported as total observed errors and errors per participant. Task performance was reported by assistance level, using counts and percentages for tasks completed independently, completed with assistance, and not completed.

Regarding *effectiveness*, measured by participants' ability to complete tasks with or without assistance, most participants were able to complete the simpler tasks independently, while more support was required for the more complex ones (Table 3).

Table . Observed errors in laboratory usability tests.

Error metric	Test 1 (n=7)	Test 2 (n=5)
Error in task 1: 1-way trip	5	6
Errors in task 2: profile creation or personalization	18	5
Errors in task 3: round-trip planning	23	11
Total observed errors	46	22
Observed errors per participant	6.6	4.4

Table . Task completion by assistance level.

Task and subtask	Test 1 (n=7)	Test 2 (n=5)
Task 1: plan a 1-way trip	• CI[a]: 4 (57%) • CA[b]: 3 (43%) • NC[c]: 0 (0%)	• CI: 2 (40%) • CA: 2 (40%) • NC: 1 (20%)
Task 2a: create profile	• CI: 5 (71%) • CA: 1 (14%) • NC: 1 (14%)	• CI: 3 (60%) • CA: 2 (40%) • NC: 0 (0%)
Task 2b: personalize profile	• CI: 0 (0%) • CA: 4 (57%) • NC: 3 (43%)	• CI: 3 (60%) • CA: 2 (40%) • NC: 0 (0%)
Task 3a: plan a trip	• CI: 2 (29%) • CA: 3 (43%) • NC: 2 (29%)	• CI: 4 (80%) • CA: 1 (20%) • NC: 0 (0%)
Task 3b: use personalized profile information	• CI: 5 (71%) • CA: 0 (0%) • NC: 2 (29%)	• CI: 5 (100%) • CA: 0 (0%) • NC: 0 (0%)
Task 3c: plan a return trip	• CI: 3 (43%) • CA: 3 (43%) • NC: 1 (14%)	• CI: 3 (60%) • CA: 2 (40%) • NC: 0 (0%)
Task 3d: save or print	• CI: 5 (71%) • CA: 1 (14%) • NC: 1 (14%)	• CI: 4 (80%) • CA: 1 (20%) • NC: 0 (0%)

[a]CI: completed independently.

[b]CA: completed with assistance.

[c]NC: not completed.

Efficiency (task duration) remained relatively consistent between the 2 tests, with a mean of 7 minutes and 48 seconds for T1 and 8 minutes and 35 seconds for T2. Overall, test 2 descriptively suggested fewer observed errors per participant and higher task completion, with similar efficiency; however, these findings should be interpreted cautiously, given the small and nonequivalent participant groups.

Questionnaires

The mean SUS score rose from 54.6 (SD 21.2; median 57.5) to 81.5 (SD 12.6; median 82.5) between the first and second tests, shifting the interface rating from "unacceptable" to "very good" [48]. Overall, the SUS results descriptively suggest higher perceived usability after the design revisions, suggesting that these revisions improved the user experience. Mean scores of UTAUT were descriptively higher across all constructs in test

XSL•FO
RenderX

2 (n=5) than in test 1 (n=7), with Social Influence remaining the lowest-rated construct in both tests (Table 4).

Table . UTAUT[a] scores from the first and second laboratory usability tests.

Test	Performance expectancy	Effort expectancy	Facilitating conditions	Behavioral intention	Behavioral use	Social influence
Test 1, mean (SD)	3.9 (0.7)	3.6 (1.2)	3.6 (0.8)	3.1 (1.4)	3.9 (1.4)	1.9 (1.0)
Test 2, mean (SD)	4.6 (1.1)	4.5 (1.4)	4.3 (0.6)	4.1 (0.8)	4.3 (0.7)	3.0 (1.1)

[a]UTAUT: Unified Theory of Acceptance and Use of Technology.

Focus Group Discussion

Following the 2 laboratory usability tests, participants shared valuable feedback during group discussions, which was thematically organized according to age-friendly usability design principles: Ease of Use, Structure and Navigation, Feedback, Control, Information, Simplicity, Visual Clarity, Consistency, and Accessibility (Table 5). Participants' quotes from the laboratory usability tests were translated from French into English.

XSL•FO
RenderX

Table . Themes identified from the focus groups of the first and second usability tests.

Themes and subthemes (tests 1 and 2)	Examples of quotations
Ease of use	
Lack of automatic data entry (T1)	"I spent half the time having the right password then the email, it didn't work. I started again, I started again, I started again" [P6, T1].
Easy to use (T2)	"I thought it [Mobilaînés] was very good because there wasn't too much information; you had to go into it" [P5, T2].
Structure and navigation	
Difficulty in navigation (T1 and T2)	"While I was engaged, I often felt lost on the pages, unsure of how to proceed. I knew what I wanted to do, but I couldn't figure out where to do it" [P2, T1].
Lack of using drop-down menu (T1)	"The drop-down menu was missing; when I searched for Sherbrooke, the result stayed behind the loading wheel. Usually, on the Internet, there is a small menu" [P3, T1].
Lack of visual or textual assistance (T1)	"I actually wish there were arrows" [P6, T1].
Feedback	
Lack of confirmation (T1 and T2)	"It was not easy to know when you finished the task" [P1, T1].
Lack of relevant feedback in error (T1 and T2)	"The little circle that turned, turned, turned, then gave nothing. Meanwhile, the minutes pass. I actually panicked a little" [P1, T1].
Confirmation in performing an action (T1 and T2)	"It is certain that if you use the Sherbrooke transportation system website [...] it does not ask us to confirm the new one. If our direction is correct, we know it's [new address] correct [...]" [P1, T2].
Control	
Lack of control by undoing, redoing, and saving (T1 and T2)	"Once I had one error in the go trip or round trip, I had to go back to the option at the top [...] it took me a while to figure out" [P7, T1].
The possibility of enlarging the maps (T1)	"I was happy to see there were maps. I love that. But I haven't tried to see if we can bring it closer because sometimes, I like to know where the stops [bus stations] are. I like to zoom in a little to see a little" [P6, T1].
Lack of safety in protecting the personal information (T2)	"Personally, I don't go into profiles there, I'm always afraid of having my identity stolen [...] There is no recall [to change the password] [...]" [P4, T2].
Information	
Insufficient relevant information for address searches, age-related travel discounts, purchasing travel tickets, checking route conditions, and comparing different modes of transport (T1 and T2)	"I wouldn't have known about the discount pass without my sister; it would help to integrate practical information like that" [P7, T1].
Lack of real-time travel information (T1 and T2)	"The advantage of Transit is that it provides real-time updates [...] That's the challenge I often face: the lack of real-time awareness" [P7, T1].
Simplicity	
Lack of clarity in terminology (T1)	"About the words used when we talk about Arrive/departure, I didn't know if it meant starting when I was performing the tasks. I was like that the more I progressed" [P3, T1].
Clarity in terminology (T1)	"About the words arrive/departure, I like it, because the departure from or arrival for that is important too" [P6, T1].
Visual clarity	
Lack of accompanying the text with images for better perception (T1)	"There are programs where when you approach the menu like personal information a small explanatory window opens" [P7, T1].
Readable font size (T1)	"Is this the real size of what will appear? Let's say I'm on a computer. Do the writings stay big like that? I think it [font size] is okay" [P3, T1].
Lack of visual clarity in the size of text and icon (T2)	"The little instruction sign was very small" [P4, T2].
Lack of clarity in color contrast (T2)	"You need to put such type of character [color of text] or such color of character on a dark background" [P1, T2].
Consistency	

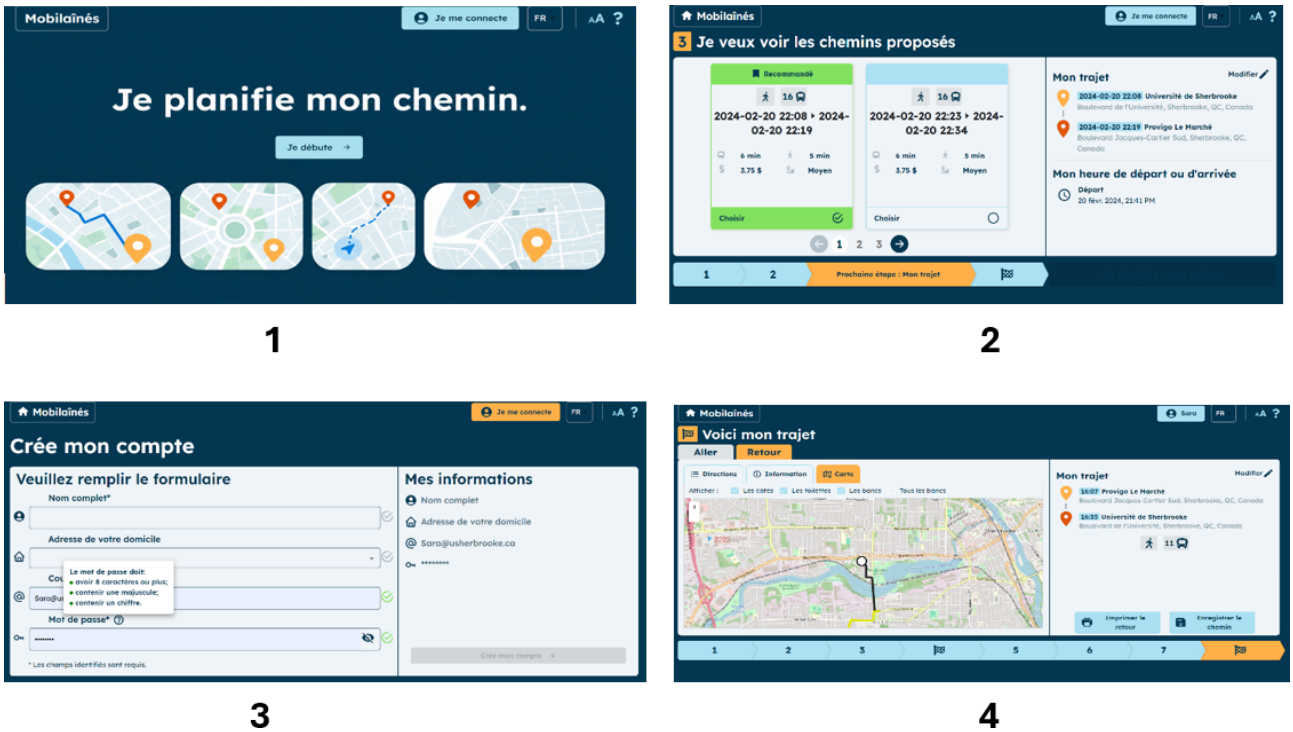
XSL•FO
RenderX

Themes and subthemes (tests 1 and 2)	Examples of quotations
Lack of using familiar buttons, icons, or information (T1 and T2)	"I always wanted to press 'Enter,' but there wasn't there [...]" [P4, T1].
Lack of presenting the information in a familiar framework and structure (T1)	"It was the American 2 am system. In Quebec, we use 14h for the afternoon" [P4, T2].
Inconsistency in navigation structure (T2)	"Maybe it would be better to put a next button instead of it [the buttons at the end of the page]" [P5, T2].
Accessibility	
Possibility of planning a trip for the future (T1)	"If you want to plan a trip for tomorrow afternoon, that's not possible. All it gives you is leave now" [P1, T1].
Lack of adaptation to the preferences of different groups of users (T2)	"Then it [Mobilaînés] didn't want me to take the bus, it forced me to walk. I think that's ageism" [P5, T2].
Adaptation to the preferences of different user groups (T2)	"Totally, I should say there is a word in English, I think it [Mobilaînés] is user-friendly" [P3, T2].
Consistency with other transport planning e-tools (T2)	"I found that it was an application that worked in much the same way as other websites. The fact that we weren't in a completely unknown universe is an advantage of it [Mobilaînés]" [P4, T2].
Possibility of personalizing the account (T2)	"Transit you can't create an account, you always have to fill in what you want, same for Google Maps" [P5, T2].
Accessible and age-friendly transport planning website (T2)	"Totally, the Mobilaînés was good, it was pleasant. There was always something of a discovery there. But it wasn't brilliant, but overall, it was pleasant" [P4, T2].

Across both laboratory sessions, participants encountered similar usability issues, mainly related to navigation, feedback, and data entry. Error analysis informed several refinements before test 2, including clearer confirmation messages and feedback (number 1 in Figure 3); clearer step indicators (number 2 in Figure 3); simplified password guidance (number 3 in Figure 3); renamed tabs, larger navigation arrows, improved map icons, and route controls (number 4 in Figure 3); and refinements to account creation and personalization features. Although real-time updates could not be implemented because of budget constraints, these revisions improved usability and prepared the interface for in-field testing.

In the laboratory usability evaluations, Mobilaînés' compatibility with age-friendly design guidelines increased descriptively from 72% to 84%. Remaining issues involved color contrast, scroll bars, lack of autofill, limited feedback and control features, unfamiliar presentation of new information, and missing real-time travel updates.

Figure 3. Mobilaînés website interface used for the second usability test.



XSL•FO
RenderX

In-Field Test

Overview of the in-Field Test

Over 6 weeks, participants used Mobilaînés for daily trip planning. During week 1, meeting dates and locations were arranged in Sherbrooke, with 1 adjustment for a participant with limited transit experience. Two participants matching the Léonard persona required targeted support due to technical difficulties, including unfamiliarity with Google Chrome and forgetting how to use the tool. Attractiveness, Perspicuity, Efficiency, and Dependability scores were higher in the second UEQ assessment, while Stimulation and Novelty were slightly higher in the first assessment (Table 6).

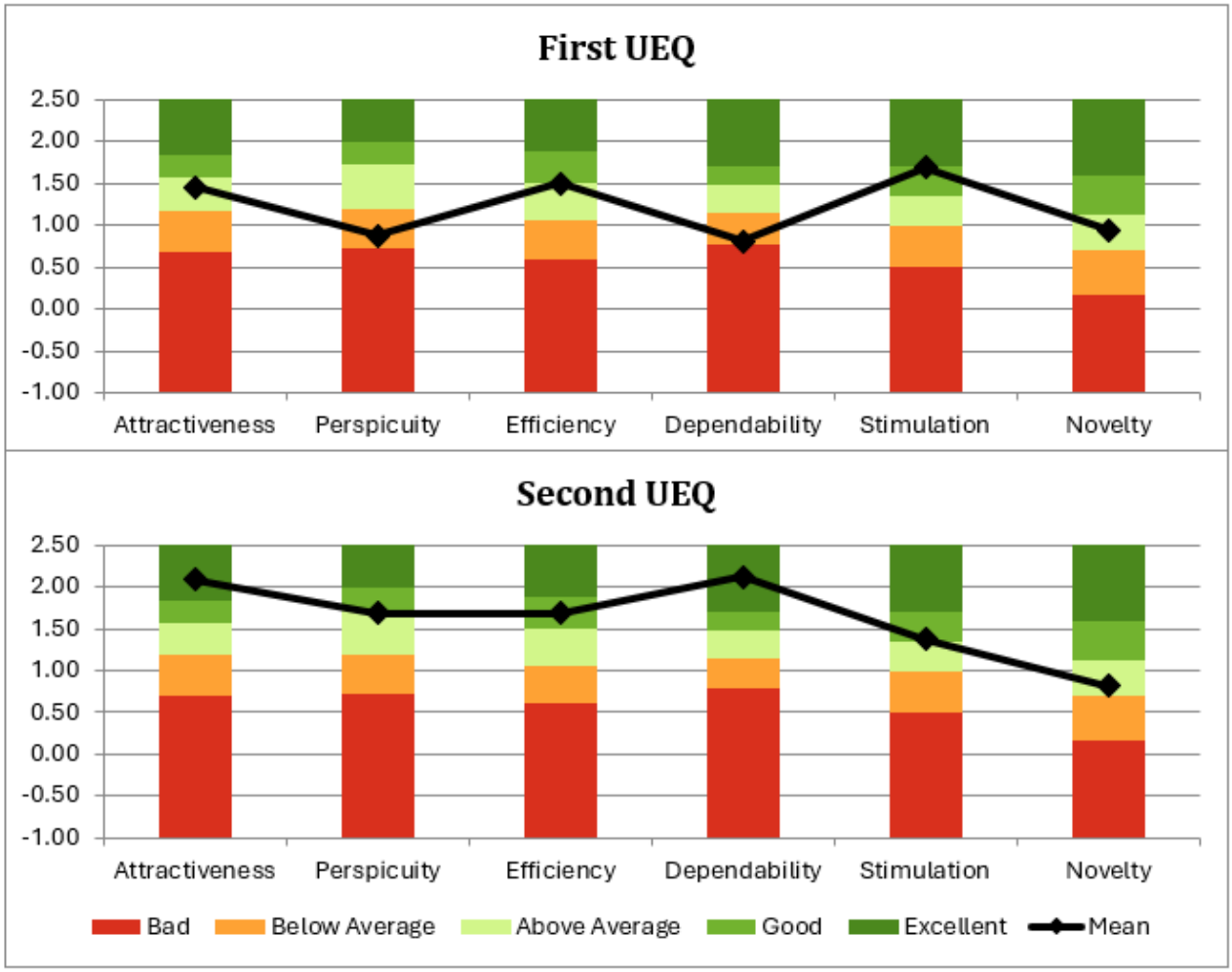
Table . Scale means and SDs of the first and second UEQ[a] assessments.

UEQ scale	First UEQ, mean (SD)	Second UEQ, mean (SD)
Attractiveness	1.46 (0.91)	2.08 (1.13)
Perspicuity	0.88 (0.48)	1.69 (1.13)
Efficiency	1.50 (0.54)	1.69 (1.09)
Dependability	0.81 (0.90)	2.13 (0.85)
Stimulation	1.69 (1.07)	1.38 (1.18)
Novelty	0.94 (1.30)	0.81 (1.84)

[a]UEQ: User Experience Questionnaire.

Comparison of the UEQ scores with the standardized benchmark dataset revealed that Stimulation consistently ranked in the top 25% (good) across both assessments, while Novelty remained in the top 50% (above average). Efficiency and Attractiveness scores increased from the top 50% (above average) to the top 25% (good) and top 10% (excellent), respectively. Dependability and Perspicuity showed the greatest improvement, moving from the bottom 50% (below average) to the top 10% (excellent) and top 50% (above average), respectively (Figure 4).

Figure 4. UEQ benchmark comparison for the Mobilaînés e-tool: first vs second assessment. UEQ: User Experience Questionnaire.



XSL•FO
RenderX

Interviews

A total of 14 subthemes were identified and classified into 9 key themes (Table 7). Five of these themes reflected positive content, including ease of use, an age-friendly travel-planning process, tailored and optimal route options, accurate timing and information, and reliability.

Table . Themes and subthemes identified through the in-field test interview.

Themes and subthemes	Examples of quotations
Ease of use	
• Easy to navigate and functional e-tool for travel planning • Easy-to-use systematic travel-planning process	"I find that although Mobilaînés has more steps, each window focuses me on a specific question [...] so I do give Mobilaînés more stars [...]" [P3].
Absence of real-time and traffic information	
• Lack of real-time travel and traffic information	"[...] for planning, I would use Mobilaînés. For the actual travel time, I would give the advantage to Google Maps because it provides real-time bus stop information, and I would know where I am [...] if Mobilaînés had real-time information, I would still use Mobilaînés first for planning [...]" [P3].
Age-friendly travel-planning process	
• Easy and accessible possibility of planning a round-trip • Detailed information provided on the path, which navigates users in the routing process.	"I used to use Google Maps, but [...] I prefer Mobilaînés because it provides more of what I expect [...]" [P2].
Tailored and optimal route options	
• Route options more tailored to user needs and preferences • Different route options in travel planning	"[...] anytime I've compared Mobilaînés to Google Maps for the proposed bus routes, they've always been the same, or Mobilaînés has been at least as good, if not better. [...] So, it was the best options were provided by Mobilaînés. [...] I would look at Mobilaînés and for going to new locations" [P3].
Accurate timing and information	
• Accurate and precise travel-related information • Accurate timing in travel planning	"It encouraged me to go out more. [...] The mobile app would show me exactly what time to take the bus to Carrefour. [...] So, you could say that the timing information on the website is one of the motivators [...]" [P2].
Reliability	
• Mobilaînés is reliable for going to new places • Mobilaînés is reliable for going out independently	"The first time I was very, very apprehensive [...] But now no problem. Now I feel comfortable, and I trust it" [P4].
Lack of accuracy in navigation	
• Mobilaînés is not that reliable in walking as it sometimes shows the wrong path	"[...] that was the main bug for me [...] it was always having me go in the wrong direction or not in the proper route [...]" [P3].
Absence of pertinent feedback	
• Lack of feedback in displaying the error message	"When you put the time around noon, forget it. [...] At one time, I could not print. [...] It's okay that that [printing error] happened only once" [P4].
Necessity to minimize walking time	
• Lack of possibility of saving the walking time	"[...] For me, it's not a problem, but for an older person [...] it will take a long time to get to the bus. [...] for people who are not used to walking, it's difficult [...]" [P1].

In sum, the in-field usability evaluation suggested that the iterative co-design and testing process led to an interface largely aligned with age-friendly guidelines, reaching 95% compatibility. Key design decisions included maintaining a step-by-step travel-planning process, presenting structured route information, supporting personalization of travel preferences, and improving visual cues and feedback. These design changes appeared to support participants' confidence and perceived reliability of the e-tool in real-world travel planning. Remaining issues included inconsistent presentation of new information, limited simplification of some travel data, lack of real-time travel and traffic information, and technical problems such as poor error handling and incorrect time displays. These issues were forwarded to the development team for correction.

Discussion

Principal Findings

This study sought to enhance the experience of older adults using a transportation planning e-tool by developing an interface

XSL•FO
RenderX

grounded in their needs and preferences through a participatory approach. The process focused on 2 objectives: co-designing an interface with older adults and evaluating its usability and user experience. The compatibility of the Mobilaînés e-tool with age-friendly design guidelines increased descriptively from 67% to 95% through iterative mock-up development and usability tests. These improvements were reflected in user experience scores, moving from unacceptable to very good.

The compatibility of the Mobilaînés e-tool with age-friendly heuristics aligns with findings by Patsoule and Koutsabasis [54], who identified Visibility and Ease of Understanding as the 2 primary heuristics tailored to older adults in website redesign. A clear and intuitive interface can significantly reduce cognitive load for older adults when interacting with digital tools, making tasks more manageable and less stressful [55]. Our results are also in line with those of Hart et al [47], who suggested that improvements to increase design compatibility with guidelines contributed to a more user-friendly experience for older adults. In our study, in-field test participants expressed a high level of satisfaction with the website's interface, which was easy to use and well suited to their travel planning needs. UEQ results showed that after several weeks of use, participants found the Mobilaînés e-tool more interesting, practical, and easy to use for travel planning. With repeated use, they also perceived it as more attractive, efficient, and trustworthy, encouraging more spontaneous and sustained engagement. However, one participant noted the absence of real-time travel and traffic information, which would have allowed tracking of their bus journey and arrival times. This limitation was observed across all 3 usability evaluations and relates to the *Information* principle in usability design guidelines. Although the travel plan printing feature does not offer real-time updates, its visual and informative format ensures accessibility for all older adults, regardless of digital proficiency. For older adults who may experience heightened anxiety from unexpected travel disruptions, Citroen [56] emphasizes that access to up-to-date information can significantly reduce stress and improve decision-making. Incorporating such features not only enhances the tool's practicality but also strengthens user confidence and satisfaction—key factors in promoting adoption and sustained use.

Interpreted through the UTAUT model [49], the design refinements implemented in Mobilaînés may help explain improvements in usability and acceptance. Step-by-step navigation, simplified task flow, familiar terminology, and clearer feedback may have supported effort expectancy by reducing the perceived difficulty of use. Personalized route-planning options, such as saved addresses and walking preferences, may have strengthened performance expectancy by making the tool more useful for users' everyday travel needs. Clear structured guidance may also represent facilitating conditions by helping users complete tasks with greater confidence. From a behavior change perspective [57], these features may support perceived capability and confidence in planning trips independently, which may encourage future use and engagement in community mobility.

However, achieving these improvements required several design and evaluation cycles, highlighting the time and resources needed to develop age-friendly e-tools through participatory and iterative approaches [58]. Qualitative data further highlighted the tool's reliability, accessibility, and effectiveness for transportation planning, despite limitations such as the absence of real-time data. Overall, these findings emphasize the value of involving older adults throughout the design and evaluation process to ensure that the Mobilaînés e-tool is closely attuned to their lived experiences and everyday contexts [58,59].

The in-field test results further revealed that Mobilaînés reduced participants' travel-related concerns by helping them anticipate each step of their journey and boosting their confidence in reaching unfamiliar destinations or adopting new modes of transport, particularly when transitioning away from car use. Although the destinations and usage period of the Mobilaînés tool were predetermined, participants reported that it enhanced trip predictability and served as a reliable resource for leaving their homes independently. In line with behavior change theory [57] and the Technology Acceptance Model [60], generating positive emotions, such as feeling safe and anticipating the pleasure of reaching a destination, may encourage sustained use of Mobilaînés. Greater ease of use enhances the overall experience, which can in turn support older adults' independence, social participation, and overall quality of life [61]. According to Schroeder et al [62], older adults' intention to adopt technology is notably influenced by awareness, emotional needs, and motivational aspects. Understanding these motivators is essential, as personal desire drives activities such as visiting family, whereas medical appointments are often perceived as obligatory [63]. This finding highlights the potential of Mobilaînés to not only facilitate independent mobility but also foster sustained engagement in meaningful daily activities by aligning technological support with older adults' emotional and motivational needs.

The main contribution of this study is not limited to the development of a transportation planning e-tool; rather, it lies in generating practical design lessons for age-friendly transportation planning interfaces. Through an iterative co-design and evaluation process, this study identifies concrete design decisions related to navigation, feedback, visual clarity, personalization, and route information that can inform the development of future digital transportation tools for older adults.

Limitations

This study faced several limitations. First, recruitment challenges resulted in an overrepresentation of tech-savvy participants (Solange personas) and an underrepresentation of those with lower digital literacy (Leonard and René personas). This imbalance may have influenced usability outcomes and should be considered when interpreting the findings. The use of personas also requires caution, as they can simplify the diversity of older adults' experiences if they are not continuously grounded in lived experience. In this study, personas were developed based on interviews, literature, and stakeholder workshops, and were used as design tools to guide recruitment, scenario development, and interface decisions. However, the overrepresentation of participants corresponding to the Solange persona may have influenced design priorities, limiting the

XSL•FO
RenderX

representation of other mobility and digital literacy profiles. Additionally, the small sample size in the laboratory usability tests (n=5 and 7) and particularly in the in-field test (n=4) warrants caution regarding the generalizability and transferability of the findings. The in-field results should therefore be interpreted as exploratory, providing preliminary insight into real-world use rather than definitive evidence of sustained effectiveness. Moreover, comparisons between the 2 laboratory usability tests should be interpreted descriptively, as they involved different participants and participant profiles, with the second test including only those participants corresponding to the Solange persona. Therefore, observed improvements in performance metrics, SUS, UTAUT, and UEQ scores may reflect both interface refinements and differences in participant characteristics, rather than generalizable evidence of effectiveness. Several strategies were used to promote recruitment, such as online advertisements, poster distribution, and the use of local databases. However, given the risks of fraud associated with soliciting older adults by phone or online [64], recruitment through trusted individuals or community partners may be particularly important in future studies [65]. Moreover, flexible scheduling regarding time and place helped to reduce withdrawal due to health reasons, which is common in participatory approaches involving older adults [66,67].

Second, resource constraints affected the feasibility of conducting a broader summative evaluation, such as involving multiple independent evaluators for heuristic analysis. The iterative co-design process required more design and evaluation cycles than anticipated, creating challenges related to time, resources, and project continuity. This highlights the need to plan sufficient resources and procedural continuity from the outset, as the number of refinement cycles can be difficult to predict. To support methodological rigor, usability issues were reviewed with interdisciplinary team members during key development and evaluation phases to help reduce interpretation bias [68]. Design issues were prioritized according to their relevance to age-friendly usability principles, potential impact on task completion and user experience, and feasibility within project constraints.

From a reflexivity and positionality perspective, the analysis and design decisions were shaped by an interdisciplinary team with expertise in user experience, rehabilitation, gerontology, communication, and computer science. While this positioning supported a multidimensional interpretation of the findings, it may also have influenced how usability issues were identified and prioritized. To reduce this risk, interpretations were informed by multiple sources of evidence, including heuristic evaluation results, older adults' feedback, performance metrics, questionnaire results, focus groups, and interviews.

Future Studies

Future studies should address key challenges identified during the development and evaluation of the Mobilaînés application. First, integrating real-time travel and traffic data is essential for effective trip planning [69], reducing anxiety about getting lost [10], protecting against adverse weather and traffic [70,71], and supporting transport choices aligned with individual needs [72]. Features not yet implemented, such as avoiding icy roads, could be enhanced through technical solutions and partnerships, improving adaptation to real-life mobility challenges and addressing safety concerns for pedestrians using mobility devices in winter conditions [73-75].

Second, future research should examine design approaches that support adoption and long-term routinization among older adults with diverse mobility, social participation, and digital literacy profiles. Interview and focus group findings suggested that accessibility features, such as pairing icons with text labels and using familiar visual cues, may benefit all users while supporting age-related visual decline and novice users, such as those represented by the René and Suzana personas. Older adults' technology adoption is shaped by prior experience, available support, perceived usefulness, confidence, and comfort [21]. For Mobilaînés, these factors may involve reliability, accurate real-time and offline information, auditory guidance, and cues that provide ongoing support or training. Universally accessible systems responsive to older adults' needs can improve usability and user experience, thereby encouraging adoption and sustained engagement [76,77]. Finally, future studies should include longitudinal evaluations of spontaneous and sustained real-world use beyond structured, researcher-supported testing to assess long-term impacts on meaningful activities, gradual behavioral changes, and infrequent travel needs, such as medical appointments.

Conclusions

Given the aging population and the critical role of mobility in health outcomes [78], this study addresses a key knowledge gap by advancing age-friendly digital transport planning. Using a participatory approach, the Mobilaînés e-tool was developed and evaluated with older adults to ensure that the interface aligned with their needs and preferences. A key contribution of this study is the identification of concrete design improvements, such as step-by-step task flow, clearer navigation, enhanced visual cues and feedback, structured route information, and personalization of travel preferences that support a more age-friendly planning experience. Findings highlight the need for iterative refinement to reduce cognitive load, support meaningful engagement, and promote digital inclusion, while guiding the development of accessible transport technologies for older adults. Future research should examine human-centered and technological strategies to foster adoption and sustained use, particularly among users with limited digital literacy.

Acknowledgments

The authors gratefully acknowledge the older adults and lived experience contributors, especially Line Coutu, citizen member of the steering committee, and Sylvie Lebrun, director of the Seniors' Consultation Table, for their participation. They also thank

XSL•FO
RenderX

Hélène Pigot for her guidance in conceptualizing phases 1 and 2, Catherine Girard for project coordination, and Véronik Lamoureux for her contribution to data collection. ChatGPT (OpenAI) was used for language refinement during manuscript preparation.

Funding

This study is financially supported by the Fonds de recherche du Québec–Santé (2020-VILL-279242).

Data Availability

The datasets produced or analyzed in this study are not publicly accessible due to ethical and legal restrictions, which limit data availability.

Authors' Contributions

Conceptualization: DB, BA, ML, ND-C, PB, MA, VP (phase 1), SBG, VP, DB (phase 2)
Funding acquisition: VP
Supervision: VP
Methodology: SBG, VP, DB
Investigation: SBG, DB, VP
Formal analysis: SBG
Data curation: SBG
Writing – original draft: SBG
Writing – review & editing: SBG, DB, ST, BA, MG-R, ML, ND-C, PB, MA, VP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Age-friendly usability design principles and guidelines (from Givi et al [39]).
[DOCX File, 20 KB - aging_v9i1e88297_app1.docx]

Multimedia Appendix 2
System Usability Scale (SUS) questionnaire.
[DOCX File, 17 KB - aging_v9i1e88297_app2.docx]

Multimedia Appendix 3
Unified Theory of Acceptance and Use of Technology (UTAUT) questionnaire.
[DOCX File, 18 KB - aging_v9i1e88297_app3.docx]

Multimedia Appendix 4
Mobilaînés Unified Theory of Acceptance and Use of Technology (UTAUT) Model and definition of the core constructs.
[DOCX File, 77 KB - aging_v9i1e88297_app4.docx]

Multimedia Appendix 5
Focus group guide for laboratory usability tests.
[DOCX File, 15 KB - aging_v9i1e88297_app5.docx]

Multimedia Appendix 6
User Experience Questionnaire (UEQ) structure and questionnaire.
[DOCX File, 48 KB - aging_v9i1e88297_app6.docx]

Multimedia Appendix 7
Interview guide for the in-field usability tests.
[DOCX File, 103 KB - aging_v9i1e88297_app7.docx]

Multimedia Appendix 8
Heuristic evaluation.
[DOCX File, 99 KB - aging_v9i1e88297_app8.docx]

XSL•FO
RenderX

References

1. Ageing and health. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-09-27]
2. Section 2—results at the Canada level, 2013 to 2063. Statistics Canada. 2014. URL: https://www150.statcan.gc.ca/n1/pub/91-520-x/2014001/section02-eng.htm [accessed 2014-09-17]
3. Andersen GJ. Aging and vision: changes in function and performance from optics to perception. Wiley Interdiscip Rev Cogn Sci 2012 May;3(3):403-410. [doi: 10.1002/wcs.1167] [Medline: 22919436]
4. Robnett RH, Chop WC. Gerontology for the Health Care Professional Third: Jones & Bartlett Publishers; 2013.
5. Young KL, Koppel S, Charlton JL. Toward best practice in human machine interface design for older drivers: A review of current design guidelines. Accid Anal Prev 2017 Sep;106:460-467. [doi: 10.1016/j.aap.2016.06.010] [Medline: 27374247]
6. Murman DL. The impact of age on cognition. Semin Hear 2015 Aug;36(3):111-121. [doi: 10.1055/s-0035-1555115] [Medline: 27516712]
7. Wong RCP, Szeto WY, Yang L, Li YC, Wong SC. Public transport policy measures for improving elderly mobility. Transp Policy (Oxf) 2018 Apr;63:73-79. [doi: 10.1016/j.tranpol.2017.12.015]
8. Ranković Plazinić B, Jović J. Mobility and transport potential of elderly in differently accessible rural areas. J Transp Geogr 2018 Apr;68:169-180. [doi: 10.1016/j.jtrangeo.2018.03.016]
9. Shimada K, Yamazaki S, Nakano K, Ngoma AM, Takahashi R, Yasumura S. Prevalence of social isolation in community-dwelling elderly by differences in household composition and related factors: from a social network perspective in urban Japan. J Aging Health 2014 Aug;26(5):807-823. [doi: 10.1177/0898264314531616] [Medline: 24812264]
10. Fiedler M. Older people and public transport: challenges and changes of an ageing society. : European Metropolitan Transport Authorities; 2007 URL: https://www.emta.com/wp-content/uploads/2023/08/Final_Report_Older_People_protec.pdf [accessed 2023-06-02]
11. Ring L, Barry B, Totzke K, Bickmore T. Addressing loneliness and isolation in older adults: proactive affective agents provide better support. Presented at: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction; Sep 2-5, 2013. [doi: 10.1109/ACII.2013.17]
12. Kirmizioglu Y, Doğan O, Kuğu N, Akyüz G. Prevalence of anxiety disorders among elderly people. Int J Geriatr Psychiatry 2009 Sep;24(9):1026-1033. [doi: 10.1002/gps.2215] [Medline: 19259977]
13. Levasseur M, Richard L, Gauvin L, Raymond E. Inventory and analysis of definitions of social participation found in the aging literature: Proposed taxonomy of social activities. Soc Sci Med 2010 Dec;71(12):2141-2149. [doi: 10.1016/j.socscimed.2010.09.041] [Medline: 21044812]
14. Whiteside MM, Wallhagen MI, Pettengill E. Sensory impairment in older adults: part 2: Vision loss. Am J Nurs 2006 Nov;106(11):52-61. [doi: 10.1097/00000446-200611000-00019] [Medline: 17068436]
15. Barbur JL, Konstantakopoulou E. Changes in color vision with decreasing light level: separating the effects of normal aging from disease. J Opt Soc Am A 2012;29(2):A27. [doi: 10.1364/JOSAA.29.000A27]
16. Bizier C, Fawcett G, Gilbert S. Dexterity disabilities among Canadians aged 15 years and older, 2012. Government of Canada. 2016. URL: https://www150.statcan.gc.ca/n1/pub/89-654-x/89-654-x2016006-eng.htm [accessed 2026-09-28]
17. Dixon J, Crooks H, Henry K. Breaking the ice: Supporting collaboration and the development of community online. CJLT / RCAT 2006;32(2). [doi: 10.21432/T25S3M]
18. Watering M. The impact of computer technology on the elderly. 2005 URL: https://www.marekvandewatering.com/texts/HCI_Essay_Marek_van_de_Watering.pdf [accessed 2026-10-28]
19. Zhou J, Rau PLP, Gao Q, Qin J, Liao Q. Requirements of transport information service and route guidance service for older adults. 2009 Presented at: 17th World Congress on Ergonomics (IEA 2009; Aug 9-14, 2009 URL: https://www.researchgate.net/publication/230710849_Requirements_of_transport_information_service_and_route_guidance_service_for_older_adults [accessed 2026-09-28]
20. Mansson L, Wiklund M, Öhberg F, Danielsson K, Sandlund M. Co-creation with older adults to improve user-experience of a smartphone self-test application to assess balance function. Int J Environ Res Public Health 2020 May 26;17(11):3768. [doi: 10.3390/ijerph17113768] [Medline: 32466484]
21. Czaja SJ, Lee CC. The impact of aging on access to technology. Univ Access Inf Soc 2007 Mar 19;5(4):341-349. [doi: 10.1007/s10209-006-0060-x]
22. Ravensbergen L, Newbold KB, Ganann R. 'It's overwhelming at the start': transitioning to public transit use as an older adult. Ageing Soc 2024 Jan;44(1):43-60. [doi: 10.1017/S0144686X22000010]
23. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023 Sep 21;11(1):e43186. [doi: 10.2196/43186] [Medline: 37733401]
24. Sanders EBN, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068]
25. Robertson T, Simonsen J. Participatory design: an introduction. In: Routledge International Handbook of Participatory Design: Routledge; 2012:1-17. [doi: 10.4324/9780203108543]

XSL•FO
RenderX

26. Bate P, Robert G. Bringing User Experience to Healthcare Improvement: The Concepts, Methods and Practices of Experience-Based Design: CRC Press; 2007. [doi: 10.1201/9781846197086]
27. Roth R. User interface and user experience (UI/UX) design. GIS&T BoK 2017;2017(Q2). [doi: 10.22224/gistbok/2017.2.5]
28. Bahrampoor Givi S, Gagnon-Roy M, Pigot H, Provencher V. User experience of older adults with an age-friendly transportation planning e-tool: scoping review. JMIR Hum Factors 2025 Jul 28;12(1):e63273. [doi: 10.2196/63273] [Medline: 40720889]
29. Provencher V, Baillargeon D, Abdulrazak B, et al. Developing a one-stop platform transportation planning service to help older adults move around in their community where, when, and how they wish: protocol for a living lab study. JMIR Res Protoc 2022 Jun 9;11(6):e33894. [doi: 10.2196/33894] [Medline: 35679116]
30. Westerlund M, Leminen S. Managing the challenges of becoming an open innovation company: experiences from living labs. Technol Innov Manag Rev 2011;1(1):19-25. [doi: 10.22215/timreview/489]
31. Institut national de la recherche scientifique (INRS). La fracture numérique: contexte québécois, pistes d'action et perspectives internationales [The digital divide: Quebec context, courses of action, and international perspectives]. : Institut national de la recherche scientifique; 2024 URL: https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/emploi-solidarite-sociale/publications-adm/rapport/RA_INRS_fracture_num.pdf [accessed 2026-10-01]
32. Tahir S, Abdulrazak B, Baillargeon D, et al. Towards smart and inclusive cities: MOBILAÎNÉS project—innovating age-friendly mobility solutions for older adults. Presented at: 2024 Mediterranean Smart Cities Conference (MSCC); May 2-4, 2024; Martil, Tetuan, Morocco p. 1-6. [doi: 10.1109/MSCC62288.2024.10697020]
33. Nielsen J. Usability Engineering: Morgan Kaufmann; 1994.
34. Kumar BA, Goundar MS, Chand SS. A framework for heuristic evaluation of mobile learning applications. Educ Inf Technol 2020 Jul;25(4):3189-3204. [doi: 10.1007/s10639-020-10112-8]
35. Tahir S, Abdulrazak B, Baillargeon D, Girard C, Provencher V. Adapting transportation planning e-tools to older adults' needs: scoping review. JMIR Aging 2023 May 16;6:e41938. [doi: 10.2196/41938] [Medline: 37191985]
36. Google Maps. URL: https://www.google.com/maps [accessed 2020-03-24]
37. Moovit. URL: https://moovitapp.com/ [accessed 2020-03-24]
38. OC Transpo. URL: https://www.octranspo.com/en/ [accessed 2020-03-24]
39. Givi SB, Gagnon-Roy M, Provencher V. Identifying and prioritizing age-friendly design principles and guidelines for developing transportation planning e-tools: scoping review. JMIR Aging 2026 Mar 19;9:e83668. [doi: 10.2196/83668] [Medline: 41855421]
40. Sanders EBN. Generative tools for co-designing. In: Scrivener SAR, Ball LJ, Woodcock A, editors. Collaborative Design: Springer; 2000:3-12. [doi: 10.1007/978-1-4471-0779-8_1]
41. Muller MJ, Druin A. Participatory design: the third space in human-computer interaction. In: Jacko JA, editor. The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, th edition: CRC Press; 2012:1125-1154. [doi: 10.1201/b11963-57]
42. Pruitt J, Adlin T. The Persona Lifecycle: Keeping People in Mind Throughout Product Design: Elsevier; 2010.
43. Rosson MB, Carroll JM. Scenario-based design. In: Jacko JA, editor. The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, 3rd edition: CRC Press; 2012:1105-1124. [doi: 10.1201/b11963-ch-48]
44. Bertelsen OW, Bødker S. Activity theory. In: HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science: Morgan Kaufmann, Vol. 2003:291-324.
45. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]
46. Russ AL, Saleem JJ. Ten factors to consider when developing usability scenarios and tasks for health information technology. J Biomed Inform 2018 Feb;78:123-133. [doi: 10.1016/j.jbi.2018.01.001] [Medline: 29329702]
47. Hart TA, Chaparro BS, Halcomb CG. Evaluating websites for older adults: Adherence to 'senior-friendly' guidelines and end-user performance. Behav Inf Technol 2008 May;27(3):191-199. [doi: 10.1080/01449290600802031]
48. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: Adding an adjective rating scale. J Usability Stud 2009;4(3):114-123. [doi: 10.5555/2835587.2835589]
49. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
50. Schrepp M, Hinderks A, Thomaschewski J. Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S). IJIMAI 2017;4(6):103-108. [doi: 10.9781/ijimai.2017.09.001]
51. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. In: Holzinger A, editor. HCI and Usability for Education and Work: Springer; 2008, Vol. 2008:63-76. [doi: 10.1007/978-3-540-89350-9_6]
52. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
53. Onwuegbuzie AJ, Dickinson WB, Leech NL, Zoran AG. A qualitative framework for collecting and analyzing data in focus group research. Int J Qual Methods 2009 Sep;8(3):1-21. [doi: 10.1177/160940690900800301]

XSL•FO
RenderX

54. Patsoule E, Koutsabasis P. Redesigning websites for older adults: A case study. Behav Inf Technol 2014 Jun 3;33(6):561-573. [doi: 10.1080/0144929X.2013.810777]
55. Understanding the cognitive burden of digital interactions for users aged 60 and older. : Harleen; 2025 URL: https://openresearch.ocadu.ca/id/eprint/4732/7/Understanding%20the%20cognitive%20burden%20of%20digital%20interactions%20for%20users%20aged%2060%20and%20older.pdf [accessed 2026-09-29]
56. Citroen CL. The role of information in strategic decision-making. Int J Inf Manage 2011 Dec;31(6):493-501. [doi: 10.1016/j.ijinfomgt.2011.02.005]
57. Sallis JF, Owen N, Fisher EB. Ecological models of health behavior. In: Health Behavior: Theory, Research, and Practice, 4th edition: John Wiley & Sons; 2015. URL: https://webda.medsab.ac.ir/uploads/1_(8)_46103.pdf#page=503 [accessed 2026-09-29]
58. Sumner J, Chong LS, Bundele A, Wei Lim Y. Co-designing technology for aging in place: a systematic review. Gerontologist 2021 Sep 13;61(7):e395-e409. [doi: 10.1093/geront/gnaa064] [Medline: 32506136]
59. Darley A, Carroll Á. Conducting co-design with older people in a digital setting: methodological reflections and recommendations. Int J Integr Care 2022;22(4):18. [doi: 10.5334/ijic.6546] [Medline: 36561633]
60. Venkatesh V, Bala H. Technology Acceptance Model 3 and a research agenda on interventions. Decis Sci 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
61. Jansen-Kosterink SM, Bergsma J, Francissen A, Naafs A. The first evaluation of a mobile application to encourage social participation for community-dwelling older adults. Health Technol 2020 Sep;10(5):1107-1113. [doi: 10.1007/s12553-020-00430-9]
62. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6(1):e44564. [doi: 10.2196/44564] [Medline: 37191976]
63. Hosford K, Pitman B, Brauer M, Lavergne R, Winters M. Characterizing older adults' travel behaviour and unmet needs: findings from the Canadian Longitudinal Study on Aging (CLSA). Can J Aging 2025 Mar;44(1):26-40. [doi: 10.1017/S0714980824000254] [Medline: 39417334]
64. Xiang H, Zhou J, Xie B. Understanding older adults' vulnerability and reactions to telecommunication fraud: the effects of personality and cognition. In: Human Aspects of IT for the Aged Population: Springer; 2020. [doi: 10.1007/978-3-030-50232-4_25]
65. Provencher V, Mortenson WB, Tanguay-Garneau L, Bélanger K, Dagenais M. Challenges and strategies pertaining to recruitment and retention of frail elderly in research studies: a systematic review. Arch Gerontol Geriatr 2014;59(1):18-24. [doi: 10.1016/j.archger.2014.03.006] [Medline: 24745811]
66. Callari TC, Moody L, Saunders J, Ward G, Holliday N, Woodley J. Exploring participation needs and motivational requirements when engaging older adults in an emerging living lab. TIM Rev 2019 Mar 26;9(3):38-49. [doi: 10.22215/timreview/1223]
67. Habibipour A, Georges A, Ståhlbröst A, Schuurman D, Bergvall-Kåreborn B. A taxonomy of factors influencing drop-out behaviour in living lab field tests. TIM Rev 2018 May 29;8(5):5-21. [doi: 10.22215/timreview/1155]
68. McKim C. Meaningful member-checking: a structured approach to member-checking. Am J Qualitative Res 2023;7(2):41-52. [doi: 10.66815/ajqr/12973]
69. Hounsell NB, Shrestha BP, McDonald M, Wong A. Open data and the needs of older people for public transport information. Transp Res Procedia 2016;14:4334-4343. [doi: 10.1016/j.trpro.2016.05.355]
70. Vargas-Acosta RA, Becerra DL, Gurbuz O, Villanueva-Rosales N, Nunez-Mchiri GG, Cheu RL. Smart mobility for seniors through the urban connector. Presented at: 2019 IEEE International Smart Cities Conference (ISC2); Oct 14-17, 2019; Casablanca, Morocco p. 241-246. [doi: 10.1109/ISC246665.2019.9071732]
71. Vechione M, Marrufo C, Vargas-Acosta RA, et al. Smart mobility for seniors: challenges and solutions in el paso, TX, and New York, NY. 2019 Presented at: 2018 IEEE International Smart Cities Conference (ISC2); Sep 16-19, 2018; Kansas City, MO, USA p. 1-8. [doi: 10.1109/ISC2.2018.8656975]
72. Stein M, Meurer J, Boden A, Wulf V. Mobility in later life: appropriation of an integrated transportation platform. Presented at: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver Colorado, USA p. 5716-5729. [doi: 10.1145/3025453.3025672]
73. Amin K, Skyving M, Bonander C, Krafft M, Nilson F. Fall- and collision-related injuries among pedestrians in road traffic environment - A Swedish national register-based study. J Safety Res 2022 Jun;81:153-165. [doi: 10.1016/j.jsr.2022.02.007] [Medline: 35589286]
74. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
75. Ripat J, Sibley KM, Giesbrecht E, et al. Winter mobility and community participation among people who use mobility devices: a scoping review. Arch Rehabil Res Clin Transl 2020 Mar;2(1):100018. [doi: 10.1016/j.arrct.2019.100018] [Medline: 33543060]

XSL•FO
RenderX

76. Subasi Ö, Leitner M, Hoeller N, Geven A, Tscheligi M. Designing accessible experiences for older users: user requirement analysis for a railway ticketing portal. Univ Access Inf Soc 2011 Nov;10(4):391-402. [doi: 10.1007/s10209-011-0223-2]
77. Zhang X, Dolah J. A study of key factors influencing the usability of smartphone graphical user interfaces for older adults. PaperASIA 2024;40(5b):136-151. [doi: 10.59953/paperasia.v40i5b.232]
78. Social determinants of health. World Health Organization. URL: https://www.who.int/health-topics/social-determinants-of-health [accessed 2025-10-01]

Abbreviations

SUS: System Usability Scale
UEQ: User Experience Questionnaire
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by R Bjarnadottir; submitted 22.Nov.2025; peer-reviewed by A Hartanto, I Markoc, J Liddle; revised version received 18.Jul.2026; accepted 02.Sep.2026; published 23.Oct.2026.

Please cite as:
Bahrampoor Givi S, Baillargeon D, Tahir S, Abdulrazak B, Gagnon-Roy M, Levasseur M, Delli-Colli N, Boissy P, Audet M, Provencher V
Development and Evaluation of a Transportation Planning e-Tool for Older Adults: Co-Design Study
JMIR Aging 2026;9:e88297
URL: https://aging.jmir.org/2026/1/e88297
doi:10.2196/88297

© Sara Bahrampoor Givi, Dany Baillargeon, Sahar Tahir, Bessam Abdulrazak, Mireille Gagnon-Roy, Mélanie Levasseur, Nathalie Delli-Colli, Patrick Boissy, Mélisa Audet, Véronique Provencher. Originally published in JMIR Aging (https://aging.jmir.org), 23.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability of a Digital Platform Supporting a Telehealth Intervention for Physical and Social Deconditioning in Older Adults: Mixed Methods Study

Nesrine Koubaa[1,2], Msc; Mélanie Levasseur[1,2,3], PhD; Mylène Aubertin-leheudre[4,5], PhD; Pasquale Roberge[1,6], PhD; Céline Verchère[3], PhD; Ruth Ndjaboue[2,7], PhD; Simon Brière[2], MScA; Cinthia Corbin[2], MA; Patrick Boissy[1,2], PhD

[1]Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, 3001, 12e Avenue Nord, Sherbrooke, QC, Canada

[2]Centre de recherche sur le vieillissement, CIUSSS Estrie CHUS, Sherbrooke, QC, Canada

[3]Institut interdisciplinaire d'innovation technologique (3IT), Université de Sherbrooke, Sherbrooke, QC, Canada

[4]Département des sciences de l'activité physique, Groupe de recherche en Activité Physique Adaptée, Faculté des Sciences, UQAM, Montreal, QC, Canada

[5]Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal (CRIUGM), CIUSSS du Centre-Sud-de-l'Île-de-Montréal, Montréal, QC, Canada

[6]Centre de recherche du CHUS, CIUSSS Estrie CHUS, Sherbrooke, QC, Canada

[7]École de travail social, Université de Sherbrooke, Sherbrooke, QC, Canada

Corresponding Author:
Patrick Boissy, PhD
Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, 3001, 12e Avenue Nord, Sherbrooke, QC, Canada

Abstract

Background: Physical and social deconditioning negatively affect older adults' health and functional independence. Telehealth interventions delivered through digital health platforms may help mitigate these challenges; however, suboptimal usability design can limit user engagement and sustained adoption. To address this issue, the PromoSanté digital platform was iteratively codeveloped and evaluated to support the virtual delivery of a multidisciplinary and multimodal telehealth program targeting deconditioning in older adults.

Objective: The objective of this study was to iteratively design and evaluate the usability of the PromoSanté digital platform with older adults both under controlled conditions and during the full-scale implementation of a 16-week telehealth program. The platform consists of a web-based learning management system with an age-adapted, responsive interface, supporting synchronous videoconferencing and asynchronous engagement through interactive modules, videos, and digital resources.

Methods: An iterative convergent mixed methods design was conducted across 2 sequential phases. In phase 1, preliminary usability testing was performed with 5 older adults who completed a 4-week, task-based virtual PromoSanté scenario. Usability, task performance, and user experience were assessed using the System Usability Scale, direct observation, and focus groups, with findings informing subsequent platform refinements. Phase 2 involved a 16-week real-world implementation of PromoSanté with 30 older adults who participated in synchronous and asynchronous sessions focused on physical exercise, preventive occupational therapy, and well-being. Usability was evaluated through adherence rates, a questionnaire grounded in the unified theory of acceptance and use of technology, and focus groups.

Results: Findings from phase 1 demonstrated high perceived usability, with a median System Usability Scale score of 82.5 (IQR 70-95). Direct observations of synchronous sessions and focus groups indicated strong user satisfaction and informed targeted refinements, including enlarging interactive buttons to improve touch-screen usability. Additional equipment was provided to enhance accessibility, such as tripods and styluses for older adults and wireless headsets for health care professionals. Phase 2 confirmed high adherence rates to synchronous (85.4%) and asynchronous (83.3%) sessions, alongside favorable ratings across unified theory of acceptance and use of technology domains, particularly performance expectancy, effort expectancy, facilitating conditions, and hedonic motivation. Older adults reported positive user experiences, highlighting the accessibility, flexibility, and overall ease of use of the platform, as well as increased confidence in navigating it.

Conclusions: Iterative refinement resulted in targeted usability enhancements. High usability scores, strong adherence rates, and positive user feedback demonstrated the feasibility and acceptability of the PromoSanté digital platform, within a supported implementation context, to deliver tailored, multidisciplinary virtual interventions addressing physical and social deconditioning.

(JMIR Aging 2026;9:e96315) doi:10.2196/96315

XSL•FO
RenderX

KEYWORDS

telehealth; digital health; usability testing; older adults; technology acceptance; mixed methods; physical deconditioning; gerontechnology

Introduction

Background

Although life expectancy has increased, aging populations face growing challenges related to chronic conditions, frailty, and functional decline [1]. Physical and social deconditioning are critical but underrecognized processes within this context [2,3]. Physical deconditioning arises from reduced activity, while social deconditioning reflects declining participation and social engagement [4,5]. Their interaction contributes to falls, loss of autonomy, and social withdrawal, reinforcing a self-perpetuating cycle of physical and social restriction [6,7].

Physical activity and social engagement are essential for preventing deconditioning and promoting healthy aging [8,9]. However, participation in face-to-face programs that deliver these activities is often constrained by logistical, mobility-related, and motivational barriers [10,11]. In response to these limitations, remote interventions delivered through digital platforms have emerged as alternative models to promote physical activity and social participation among older adults. These interventions combine synchronous and asynchronous modalities to deliver structured, interactive activities. By improving accessibility and continuity of participation, this hybrid approach facilitates sustained physical and social engagement and has demonstrated feasibility, satisfactory adherence, and positive user perceptions in older populations [12,13].

Despite their promise, digital platforms are not consistently adopted by older adults, and usability challenges are repeatedly cited as major barriers to sustained engagement [14-16]. The ISO defines usability as the extent to which users can achieve specified goals effectively, efficiently, and satisfactorily in a given context [17]. However, many digital health platforms fail to adequately address older adults' needs and levels of digital literacy during the design process [18,19], leading to complex interfaces and navigation difficulties [20,21]. Age-related cognitive and physical changes, such as visual and dexterity impairments, can further reduce telehealth usability and contribute to disengagement or abandonment [22,23].

Prior Work

To address the need for accessible and sustainable solutions promoting long-term physical activity and social participation, a team at the Research Center on Aging in Sherbrooke, Quebec, co-designed PromoSanté [24,25], a multidisciplinary telehealth program. Developed collaboratively with older adults and community and clinical partners, the program aligns content, delivery methods, and technological features with end users' capacities and preferences. Consistent with best-practice approaches to healthy aging and deconditioning prevention, PromoSanté integrates structured exercise, preventive occupational therapy, and well-being support [26,27]. Its multimodal telehealth model combines face-to-face group sessions with synchronous and asynchronous remote delivery to enhance scalability and continuity beyond traditional settings.

A dedicated digital platform underpins this telehealth approach. Built on a learning management system with custom plugins to simplify authentication and enhance usability, the platform features an age-adapted, responsive interface designed to accommodate sensory, cognitive, and motor changes. It supports synchronous videoconferencing and asynchronous engagement through interactive modules and digital resources. Participants used their own devices or, when necessary, an Android tablet provided during the study period.

Objectives of the Study

The present study examined the usability of the digital platform developed to deliver the PromoSanté telehealth program, addressing two complementary objectives. First, it aimed to validate and iteratively refine the usability of the platform prototype. Second, it evaluated the real-world usability of the optimized platform during the implementation of the full PromoSanté telehealth intervention.

Methods

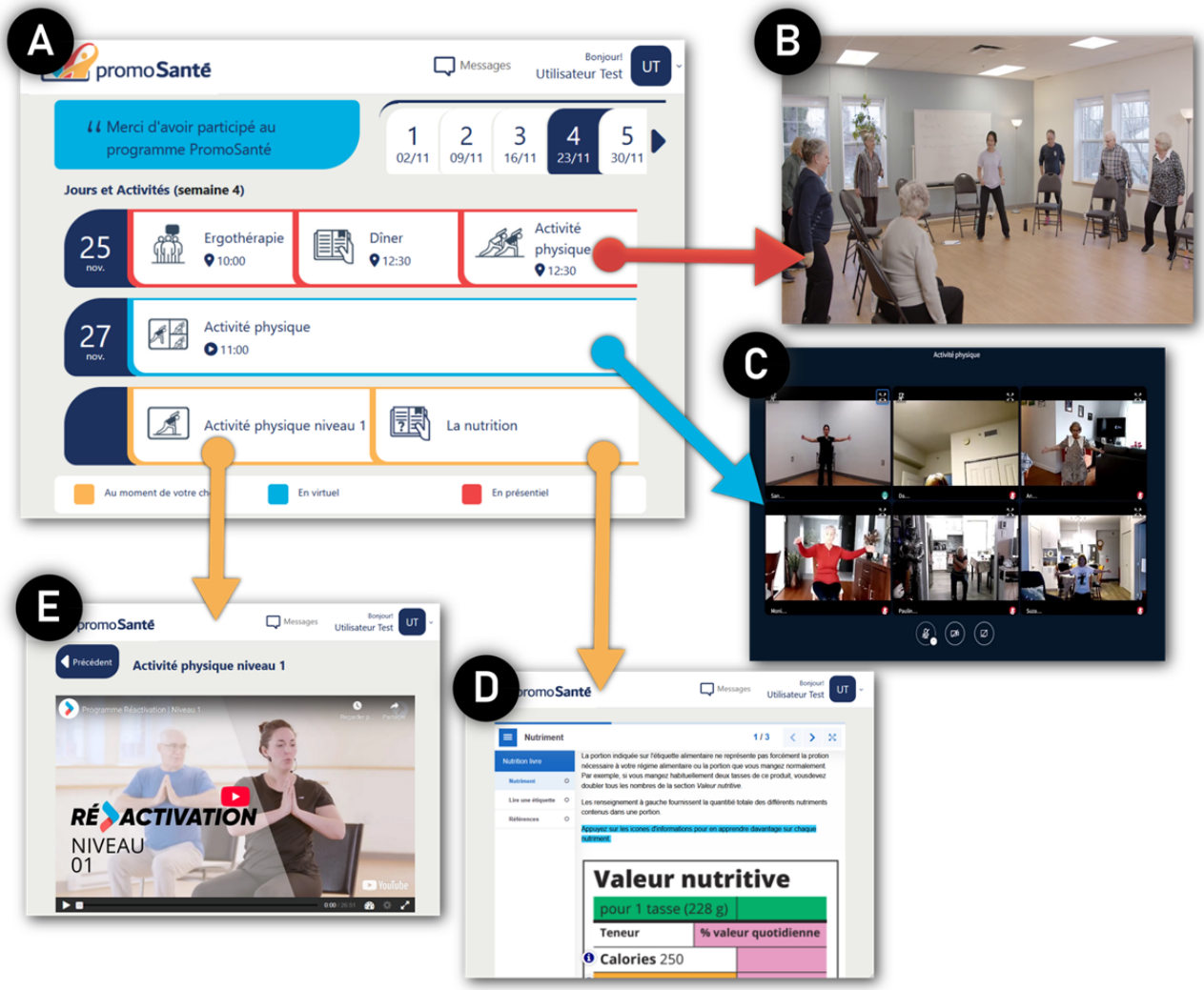
Design of the Digital Platform

Usability was systematically integrated across the platform's development and evaluation phases to align with end users' needs and capabilities [28]. A preliminary qualitative study elicited older adults' preferences and stakeholder perspectives, informing both program structure and platform design [29]. The interface was then developed through an iterative participatory approach involving researchers in aging and digital health alongside a professional UI designer. Design choices followed systematic review recommendations [30], including simplified layouts, larger and well-spaced interactive elements, and avoidance of edge-based controls.

The design of the digital platform interface is presented in Figure 1. Panel A displays the main dashboard, which functions as the central navigation hub and provides a clear overview of scheduled activities for the selected week. As the default landing page upon login, the dashboard facilitates rapid orientation and awareness of upcoming tasks, consistent with the usability principle of visibility of system status. At the top of the interface, the platform logo and direct access buttons for the user profile are prominently displayed, ensuring consistent and predictable access to core functions, in accordance with the principle of consistency and standards.

XSL•FO
RenderX

Figure 1. Interface of the PromoSanté platform. (A) Main dashboard displaying the program's weekly calendar and serving as the central navigation hub; (B) in-person sessions at the community center; (C) synchronous sessions; (D) asynchronous modules of occupational therapy component; (E) recorded physical exercise sessions.



Scheduled activities are presented in a chronological list, with each entry displaying the corresponding date and time. Each activity is accompanied by a distinct icon representing its type, thereby reducing reliance on text and supporting the principle of recognition rather than recall. To further enhance navigation and comprehension, color coding differentiates delivery modalities: red denotes in-person activities (B), blue indicates synchronous remote sessions (C), and yellow represents asynchronous activities. This visual differentiation facilitates the rapid identification of participation requirements and aligns with the heuristic of match between the system and the real world through intuitive visual cues.

Selecting an activity from the dashboard directs users to the corresponding content or action through a predictable navigation pathway. When a synchronous session is selected, users are automatically redirected to a dedicated interface that provides direct access to the videoconferencing link, thereby minimizing navigation steps and reducing cognitive load in accordance with the principle of efficiency of use. For asynchronous activities, selecting opens a structured content page where all relevant materials are organized in a clear, linear format. This layout allows users to progress at their own pace and to revisit previously completed content, supporting user control and freedom.

Across all modules, the platform maintains consistent navigation pathways, terminology, and visual layouts, enhancing learnability and minimizing the risk of user error. To further reduce cognitive load, the interface limits on-screen options and presents only task-relevant information at each stage, consistent with the heuristic of aesthetic and minimalist design [31].

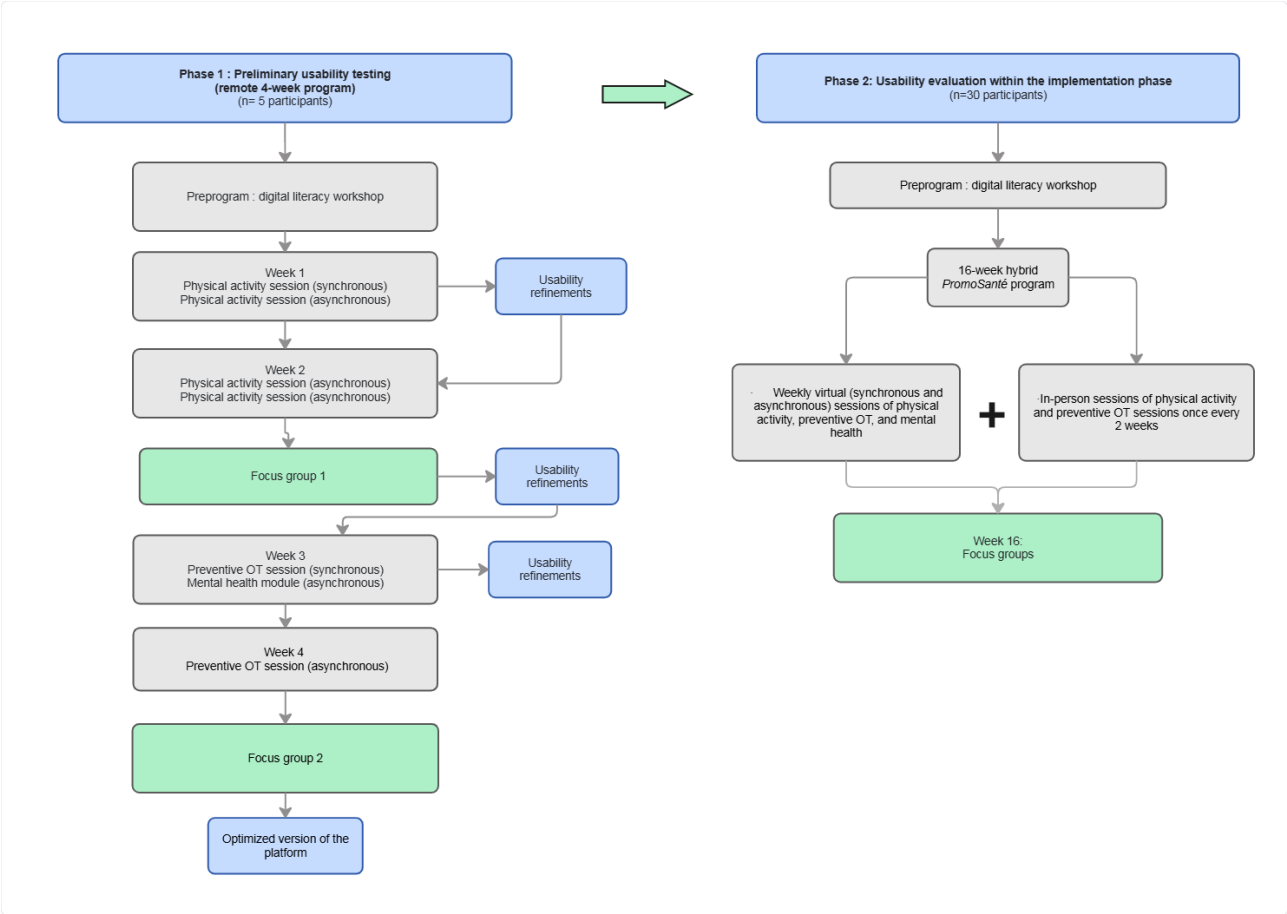
The platform supports several asynchronous activities across the different disciplinary components of the program while preserving these usability principles. The preventive occupational therapy component includes educational capsules and interactive reading materials aimed at promoting functional autonomy and preventive behaviors (D). Within the physical activity components, participants complete prerecorded exercise sessions (E) designed to progressively increase in intensity over the course of the program, with clear visual cues guiding session selection and progression.

XSL•FO
RenderX

Study Design

The usability evaluation of the PromoSanté digital platform was conducted in 2 sequential phases (Figure 2), each corresponding to a specific study objective. Phase 1 focused on validating and iteratively optimizing the usability of the platform prototype. Preliminary usability testing was conducted using task-based scenarios with a small sample of representative older adults to identify navigation, comprehension, and interaction challenges and to inform targeted design refinements.

Figure 2. Study enrollment. OT: occupational therapy.



Phase 2 examined the real-world usability of the finalized platform during full program implementation. This phase was embedded within a broader evaluation of the 16-week *PromoSanté* program, which combined in-person sessions held in a community center with virtual sessions delivered through the digital platform.

Across both phases, an iterative convergent mixed methods design was employed, involving concurrent collection of quantitative and qualitative data across successive cycles of testing and refinement [32]. Data were analyzed separately within each phase and subsequently triangulated to identify usability strengths, barriers, and improvement priorities. This systematic integration of findings guided platform refinement and ensured alignment with the study's objectives of usability optimization and validation.

Phase 1: Preliminary Usability Testing Phase

Participants' Recruitment

This phase aimed to identify usability barriers to task completion and confirm platform functionality. Participants completed typical virtual PromoSanté activities remotely using the developed platform. Five older adults were purposively recruited from a community center in Sherbrooke, Quebec. Inclusion criteria were age ≥65 years, limitation in at least one activity of daily living, residence in a private dwelling, reliable internet access, basic digital literacy, and French proficiency. Individuals with significant cognitive or sensory impairments that could interfere with study procedures were excluded based on an occupational therapist's assessment. Consistent with usability testing standards, five participants were deemed sufficient to identify most usability issues [33]. Two health care professionals (a kinesiologist and an occupational therapist) with at least 2 years of experience working with older adults and/or delivering telehealth services led the synchronous sessions.

Testing Procedure

Participants completed a 90-minute digital literacy workshop prior to testing, during which the platform interface and core functions were introduced through interactive demonstrations. Each participant received a study tablet and practiced key tasks (eg, videoconference access and asynchronous module navigation). In addition, the project coordinator conducted individual technical tests with each participant to verify the reliability of their internet connection. Usability testing consisted of a 4-week task-based scenario (Figure 2) simulating typical

XSL•FO
RenderX

program use. Activities included synchronous and asynchronous exercise sessions, an occupational therapy session on fall prevention with interactive components, a stress management module, and completion of a home safety checklist.

Data Collection

During phase 1, usability outcomes were collected before, during, and after testing. At baseline, sociodemographic characteristics (age, sex, and education level) and information on participants' digital device use were recorded.

Perceived Usability (Satisfaction)

At the end of the program, perceived usability was assessed using the validated French version of the System Usability Scale (SUS) [34]. The SUS includes 10 items rated on a 5-point Likert scale and produces a total score ranging from 0 to 100 [35]. Scores below 68 indicate below-average usability, scores in the 70s and 80s reflect good usability, and scores above 80.3 indicate excellent usability [36]. The French version demonstrates strong reliability and construct validity comparable to the original English scale [34,36-38].

Task Performance and Interaction (Effectiveness and Efficiency)

Direct observation was conducted during synchronous sessions to document task completion, navigation challenges, and real-time usability issues. A research team member observed each session and provided assistance when necessary. Observational notes were reviewed after each session to identify priority issues and guide iterative refinements. Adherence was measured by tracking attendance at synchronous sessions and the completion of asynchronous activities.

User Experience and Interaction Flow (Effectiveness and Satisfaction)

Two focus groups were conducted in weeks 2 and 4 using a semistructured guide informed by Hierarchical Task Analysis. Hierarchical Task Analysis involves decomposing typical tasks into subtasks and operational steps to examine interaction sequences, identify breakdowns, and explore user perceptions [39,40]. Participants were invited to verbally reconstruct the steps required to complete common platform tasks while viewing the interface. Discussion focused on task completion, navigation pathways, and perceptions of ease of use, clarity, and confidence. Focus groups were audio-recorded and transcribed verbatim.

Data Analysis

SUS scores were analyzed using descriptive statistics (median and semi-IQR).

Qualitative data from the focus groups and observations were analyzed to assess usability from a user-experience perspective, with a particular focus on what participants liked and disliked about the platform, as well as any issues that could hinder their motivation to use it and should therefore be considered and addressed before the subsequent testing phase. In this first phase, 2 members of the research team independently coded the qualitative data to identify segments related to user experience, positive and negative aspects of the platform, and issues that could potentially affect participants' motivation to use it. The coding was then compared and discussed to reach agreement on the identified segments and relevant issues. Quantitative and qualitative findings were analyzed in parallel and subsequently triangulated to identify key usability issues and inform iterative refinements to the platform.

Phase 2: Usability Evaluation Within the Implementation Phase

Participants' Recruitment

Following the resolution of the major usability issues identified in phase 1, a second evaluation was conducted during the full implementation of PromoSanté. This phase assessed platform usability under real-world conditions and examined challenges emerging with sustained use. The 16-week program combined in-person sessions with synchronous and asynchronous remote activities, reflecting routine delivery. Inclusion and exclusion criteria remained unchanged to ensure consistency. A larger sample was recruited, comprising 4 groups of up to 8 older adults each, to evaluate the revised platform. The sample size was determined by available resources, and group composition was informed by recommendations derived from phase 1 (see the *Results* section for details).

A recruitment poster was disseminated through the community center's periodic newsletter. Additionally, pivot workers and occupational therapists at the affiliated hospital center supported recruitment by identifying and approaching eligible patients who were receiving home care services.

Data Collection

Sociodemographic Characteristics and Technology Use

As in phase 1, participants attended a digital literacy workshop to familiarize themselves with the platform prior to program initiation. Baseline demographic characteristics were collected at enrollment. In addition, participants were asked which smart devices they had access to and used to connect to the internet, how much time they spent online, and whether they received help from others when accessing the internet. These questions were used to characterize participants' level of digital experience and autonomy in using digital technologies.

Attendance at synchronous sessions and completion of asynchronous activities were recorded throughout the intervention to assess adherence.

Technology Acceptance and Usability (the Unified Theory of Acceptance and Use of Technology–Based Questionnaire)

Technology acceptance and determinants of platform use were assessed using an adapted Canadian French version of a unified theory of acceptance and use of technology (UTAUT2)–based questionnaire [41,42]. The instrument included 23 items across 7 domains: performance expectancy (PE), effort expectancy (EE), facilitating conditions (FC), social influence, hedonic motivation (HM), habit, and behavioral intention. Items were rated on a 7-point Likert scale (1="strongly disagree" to 7="strongly agree").

Focus Groups

At the end of the 16-week program, participants were invited to a focus group to explore their experiences with the platform.

XSL•FO
RenderX

A semistructured interview guide was developed and explicitly structured around the 3 ISO-defined dimensions of usability: effectiveness, efficiency, and satisfaction [43]. Effectiveness referred to participants' ability to achieve their intended goals using the platform; efficiency captured the time, cognitive, and physical effort required; and satisfaction reflected subjective perceptions, comfort, and overall experience [17]. This framework enabled a systematic examination of real-world usability and complemented the quantitative findings.

Data Analysis

Descriptive analyses were conducted to characterize platform use during the implementation phase, including frequency of platform access and adherence to scheduled synchronous and asynchronous activities as objective indicators of engagement. UTAUT2 data were analyzed descriptively at the domain level. For each participant, mean domain scores were calculated by averaging item responses and rounding to the nearest integer on the 7-point Likert scale. Frequency distributions were generated to determine the proportion of participants at each response level (1-7). For interpretability, scores were categorized as below (<4), equal to (=4), or above (>4) the neutral midpoint, representing disagreement, neutrality, and agreement, respectively. This approach enabled the detailed examination of technology acceptance and usability-related perceptions.

Qualitative data from the focus groups were analyzed thematically using the method proposed by Miles and Huberman [41]. An initial coding strategy was developed based on the 3 usability criteria: effectiveness, efficiency, and satisfaction. Two members of the research team independently coded the data. First, each meaningful segment of the verbatim transcripts was assigned an initial code reflecting its content and meaning. Related codes were then grouped and categorized under one of the 3 predefined usability criteria. The 2 coders subsequently met regularly to compare their coding results, harmonize the identified themes, and resolve discrepancies. When consensus could not be reached, a third member of the research team was consulted to facilitate discussion and reach consensus.

Ethical Considerations

The present study was approved by the Ethics Committee of the CIUSSS de l'Estrie-CHUS (approval number 2022 - 4277). Written informed consent was obtained from all individuals prior to their participation. The study was conducted in a manner that maintained confidentiality throughout, and all data were deidentified before analysis.

Results

Preliminary Usability Testing Phase

Participants' Characteristics

For this phase of the usability study, 7 participants were recruited: 5 older adults (n=4 women) and 2 health care professionals (both women; n=1 kinesiologist and n=1 occupational therapist) who delivered the interventions via the digital platform. The median age of older participants was 73 (IQR 69-77) years, and 4 had completed postsecondary education. Three lived with a partner and 2 lived alone. All reported at least 2 chronic conditions, most commonly musculoskeletal and respiratory disorders. Two participants had prior telehealth experience (telephone follow-up with a general practitioner or an online session with a kinesiologist). All participants reported regular use of digital technologies, primarily smartphones; 3 used them for more than 3 hours daily, and 4 indicated they required no assistance when using technology.

Adherence and Perceived Usability

All participants completed the asynchronous PromoSanté sessions. Adherence to synchronous sessions was high (90%), with full attendance across the sessions except for one participant, who missed the first 45 minutes because of selecting an incorrect link within the platform. The digital platform PromoSanté demonstrated high overall usability (a median SUS score of 82.5; the median age of older participants was 73, IQR 70-95 years). Ease of use and confidence in using the platform were particularly well rated. Participants also expressed a strong intention to use the platform frequently in the future.

Task Performance and Interaction

Across 2 synchronous physical activity sessions and 1 occupational therapy session, participants were generally able to access the platform and demonstrated high engagement, supporting overall usability effectiveness. However, several usability issues emerged, primarily related to audiovisual setup, navigation accuracy, and interaction tools. During the first physical activity session, suboptimal camera positioning limited the kinesiologist's ability to observe movements, affecting interaction quality and instructional feedback (efficiency). Introducing tripods and providing clearer preparatory guidance significantly improved participant visibility in subsequent sessions, enhancing both instructional effectiveness and interaction flow. In the second physical activity session, background noise reduced audio clarity and disrupted communication (efficiency). Muting microphones mitigated the issue, and the recommendation to provide wireless headsets to health care professionals aimed to further improve communication quality and reduce cognitive load during facilitation. Participants also demonstrated increasing mastery of platform interaction features (eg, "raise hand"), indicating improved learnability over time. The occupational therapy session highlighted navigation-related usability risks (effectiveness), as one participant misinterpreted calendar information and selected an incorrect session link, requiring telephone assistance. Although content engagement remained high (satisfaction), interaction with shared on-screen tools was uneven, with only 2 participants using them directly. The occupational therapist reported greater cognitive demands compared to in-person delivery, suggesting that virtual facilitation may require adaptation of group size and environmental control to maintain efficiency and reduce facilitator workload.

XSL•FO
RenderX

Qualitative Results: User Perceived Experience and Interaction Flow

Positive Feedback

Overall, participants reported a positive user experience that reflected a user-friendly interface. The interface was described as intuitive and easy to navigate, with clearly displayed information facilitating access to content. Participants found access to synchronous sessions straightforward, noting that clearly labeled links and color-coding effectively distinguished among activity types. Feedback provided within the online physical activity questionnaire—particularly correct answers accompanied by brief explanations—was also appreciated for supporting learning. Synchronous occupational therapy sessions were positively received. Despite limited prior experience with virtual group discussions, participants reported feeling comfortable sharing their perspectives and valued the respectful and engaging atmosphere. Mutual respect was identified as essential to fostering open and dynamic exchanges, reinforcing satisfaction with the virtual group experience.

Issues to Consider

During focus group discussions, some participants reported difficulties with touch-based interactions, particularly typing on virtual keyboards and selecting small interface elements. These challenges, attributed to reduced finger dexterity, occasionally generated frustration and underscored the need for larger interactive targets to improve usability and efficiency. Navigation issues also emerged within asynchronous physical activity modules. Because exercise videos were embedded from the community center's YouTube channel, one participant followed an external link, mistakenly perceiving it as part of the program. Difficulty returning to the PromoSanté platform highlighted the need for clearer navigation cues and stronger visual boundaries when integrating external content.

Final Improvements

As summarized in Table 1, participant feedback from synchronous sessions and focus groups directly informed targeted platform refinements, addressing identified usability barriers and enhancing overall effectiveness, efficiency, and user satisfaction. In addition to technical refinements, a comprehensive user guide was developed to support independent platform use. The guide provides instructions for configuring tablet settings (eg, audio, brightness, and connectivity) and step-by-step directions for accessing the platform, joining synchronous sessions, and navigating asynchronous modules. Key interface sections are visually illustrated and clearly described to reinforce the understanding of their functional roles. This resource was designed to enhance user autonomy, reduce support needs, and promote sustained engagement.

Table . Observations, usability issues, and iterative refinements across synchronous sessions of the PromoSanté telehealth platform.

Session	Observations	Usability issues	Iterative refinements
Session 1–synchronous physical activity	Participants successfully joined and followed exercises; partial body visibility on camera; no adverse events	Limited full-body visibility due to difficulties with camera positioning	Tripods provided; in-person setup support; preparatory instructions added (equipment and environment)
Session 2–synchronous physical activity	Positive feedback on asynchronous exercise module; audio clarity issues; use of the "raise hand" feature	Background noise reduced audibility of instructions	Wireless headset will be provided for next sessions to HCPs[a]; participant microphones muted as needed
Session 3–synchronous occupational therapy	High engagement with discussion on home safety; navigation errors when accessing the session; limited use of shared on-screen tools	Confusion when accessing synchronous sessions; difficulty interacting with shared digital activities	Session access button enlarged and repositioned; font size increased to improve visibility and clickability, providing stylus for future users to facilitate touch-based interactions
Across sessions	High engagement with increased cognitive load for facilitators (occupational therapist) in virtual delivery	Managing interaction and attention in larger virtual groups	Recommendation to limit group size (≤8 participants); reminders to join from quiet environments

[a]HCP: health care professional.

Real-World Usability Evaluation (Phase 2)

Participants' Characteristics

Based on available resources, the platform was implemented across 4 groups over 18 months, each comprising up to 8 participants, consistent with recommendations from preliminary usability testing. A total of 30 participants were recruited for this phase. The mean age was 75.1 (SD 5.4) years, and the sample was predominantly female (n=28, 93.3%). Most participants had completed higher education and reported no prior telehealth experience (Table 2).

XSL•FO
RenderX

Table . Participant characteristics (phase 2; N=30).

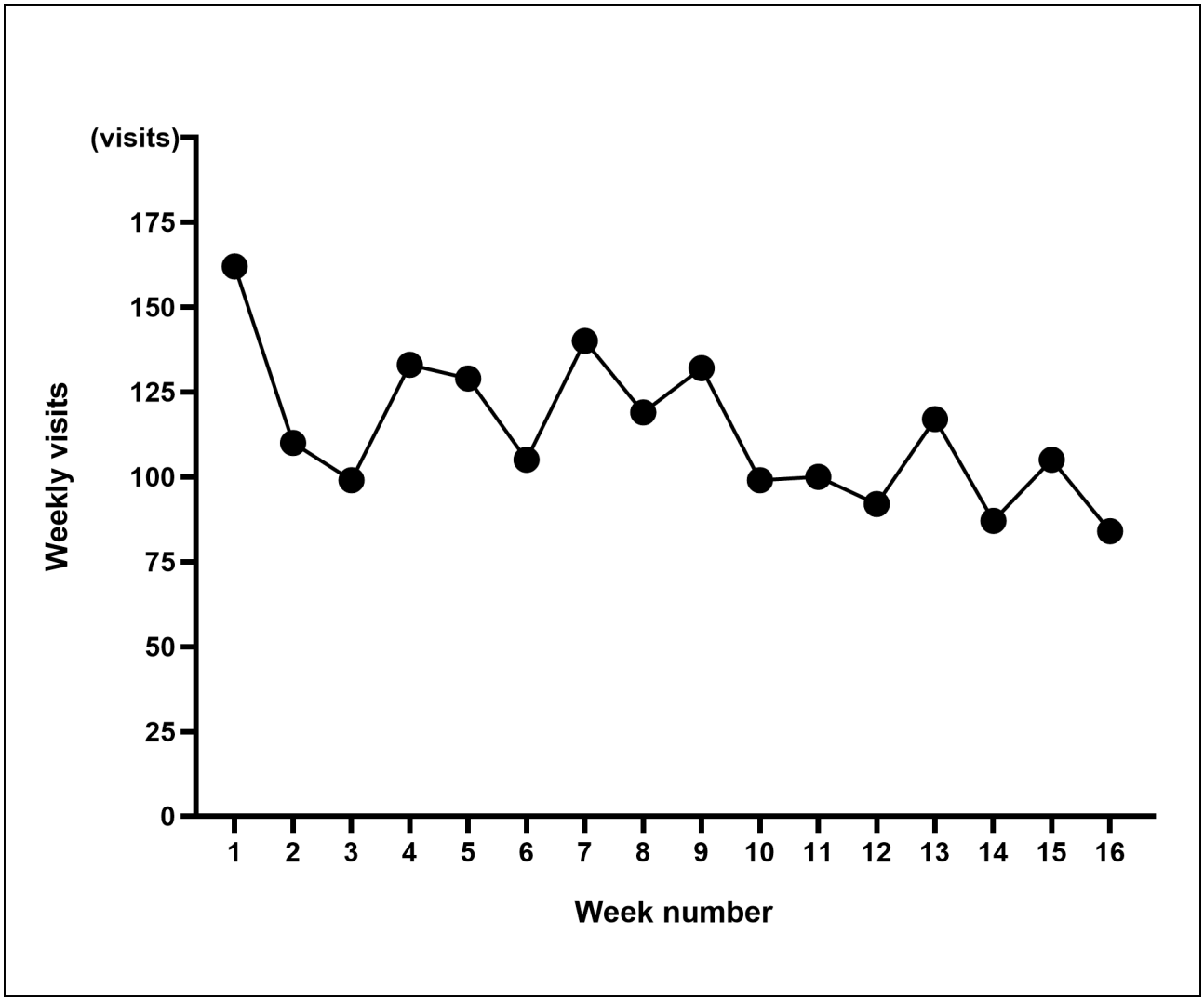
Characteristic	Participant, n (%)
Sex	
Female	28 (93.3)
Male	2 (6.7)
Education level	
Secondary	10 (33.3)
College	15 (50)
University	5 (16.7)
Marital status	
Married	15 (50)
Widow	4 (13.3)
Single	2 (6.7)
Divorced	9 (30)
Living situation	
Alone	14 (46.7)
With a partner	14 (46.7)
Other	2 (6.7)
Prior experience with telehealth services	
Yes	7 (23.3)
No	23 (76.7)
Devices used to connect to internet	
Tablet	15 (50)
Laptop	20 (66.7)
Smartphone	17 (56.7)
Daily use of connected devices (h)	
≤1	5 (16.7)
>1 to 3	13 (43.3)
>3 to 5	10 (33.3)
≥5	2 (6.7)
Need assistance to use digital devices	
Yes	10 (33.3)
No	20 (66.7)

Adherence

Over the 16-week program, the PromoSanté platform recorded 1813 visits, averaging 113 (SD 21) visits per week. Figure 3 shows the weekly number of visits per participant. Use peaked during the first week—likely corresponding to initial onboarding and exploration—before stabilizing across subsequent weeks, suggesting the consistent integration of the platform into participants' routines. A slight decline in week 16 may reflect natural tapering at program completion rather than disengagement. Overall, the curve of visits reflected sustained engagement throughout the intervention. Adherence remained high across modalities, with participation rates of 85.4% for synchronous sessions and 83.3% for asynchronous activities, further supporting the platform's acceptability and usability under real-world conditions. One participant was withdrawn from the study based on the occupational therapist's clinical judgment due to increasing confusion and difficulty remembering scheduled sessions.

XSL•FO
RenderX

Figure 3. The number of weekly visits to the platform throughout the 16-week program period.



UTAUT2 Questionnaire

Descriptive statistics for UTAUT2 domains are presented in Figure 4 and Table 3. More than half of the participants rated all domains above the neutral midpoint (>4). The highest proportions of favorable scores were observed for PE, EE, FC, and HM (80% - 93.3%), indicating strong perceived usefulness, ease of use, and supportive conditions. In contrast, the habit domain showed comparatively lower endorsement, with 60% scoring above 4% and 23.3% below 4, suggesting that routine integration of the platform was still developing. Overall, the findings reflect positive technology acceptance and favorable perceptions of usability.

XSL•FO
RenderX

Figure 4. Unified theory of acceptance and use of technology (UTAUT2)–based questionnaire scores.

XSL•FO
RenderX

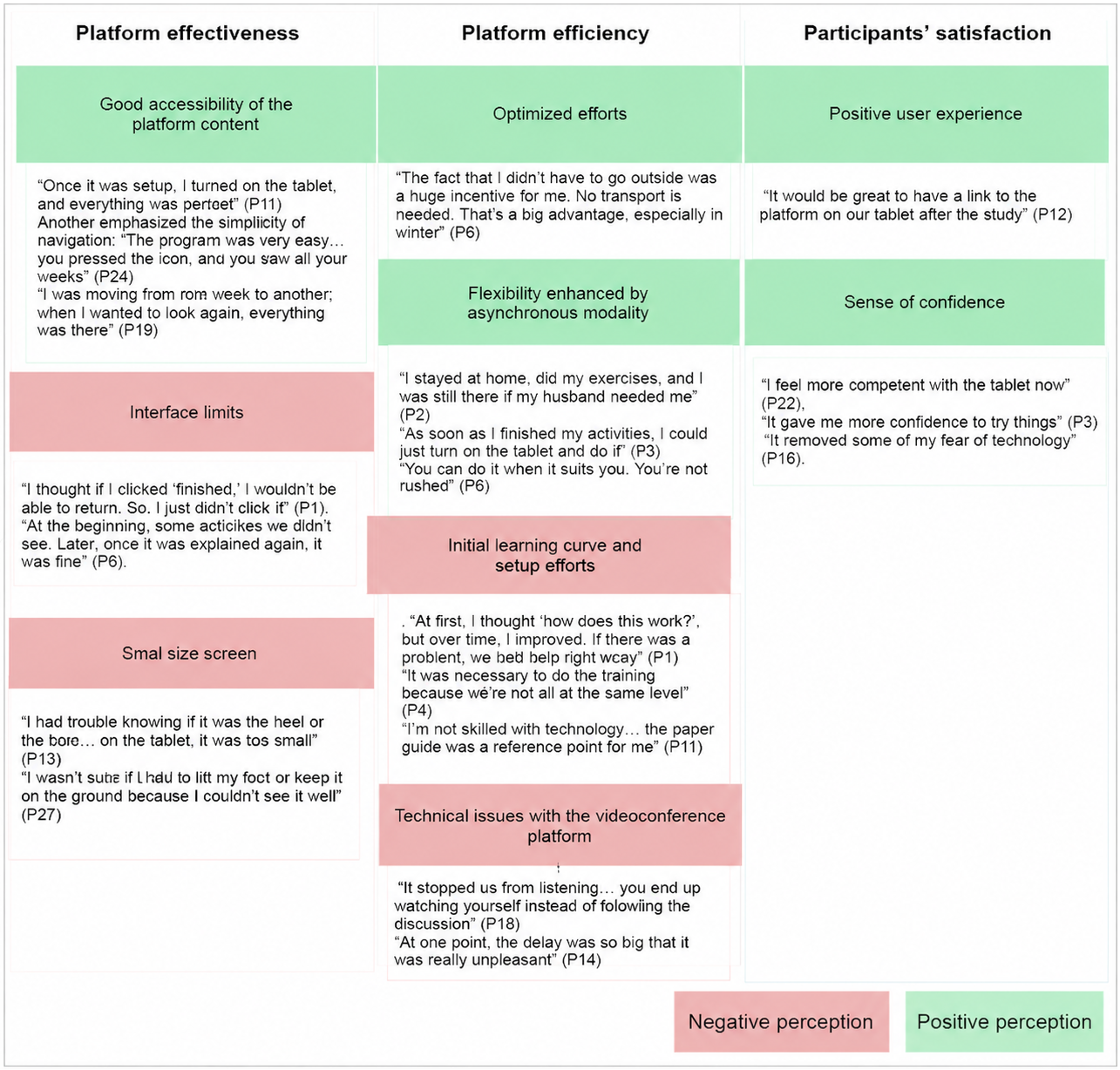
Table . Unified theory of acceptance and use of technology (UTAUT2)–based questionnaire ratings.

UTAUT2 score	Performance expectancy	Effort expectancy	Social influence	Facilitating conditions	Hedonic motivation	Habit	Behavioral intention
Median (IQR)	6.2 (5.1-6.7)	7 (5.8-7.0)	6 (4.5-7.0)	7 (5.5-7.0)	6 (5-7)	5.7 (3.7-6.0)	6 (4.4-7.0)
<4 (%)	6.7	6.7	10.0	3.3	3.3	23.3	13.3
=4 (%)	13.3	0.0	13.3	3.4	6.7	16.7	13.4
>4 (%)	80.0	93.3	70.0	93.3	90.0	60.0	73.3

Qualitative Results

Thematic analysis of focus group data generated several subthemes, which were systematically mapped onto the 3 core usability dimensions—effectiveness, efficiency, and satisfaction (Figure 5). This structured organization enabled interpretation of participants' experiences through an established usability framework and facilitated identification of dimension-specific strengths and challenges.

Figure 5. Participants' perceptions of their experience with the digital platform.



XSL•FO
RenderX

Theme 1: Platform Effectiveness

Good Accessibility of the Platform's Content

Participants reported being able to access the platform and complete scheduled activities without significant barriers, indicating strong usability effectiveness under real-world conditions. After the initial onboarding workshop, the interface was perceived as intuitive, suggesting that the platform's structure supported learnability and reduced cognitive load over time. The ability to revisit materials and repeat activities was particularly valued, as it enabled self-paced engagement and reinforced comprehension. This flexibility not only supported user autonomy but also likely contributed to sustained participation by allowing individuals to adapt platform use to their personal routines and functional capacities.

Interface Limitations

Although most participants were able to access the platform content successfully, a few interface elements caused initial confusion. For instance, one participant misunderstood the "mark as completed" function in asynchronous modules, assuming that it would restrict future access to the content, which led to reluctance to use the feature. In addition, during the first week of use, some participants did not immediately recognize that additional content was available through scrolling. These issues highlight minor learnability challenges and underscore the importance of explicit visual cues to support intuitive navigation.

Small Screen Size

A small number of participants reported that the tablet screen size limited their ability to clearly observe certain physical exercises demonstrated by the kinesiologist. Reduced visual clarity, particularly for foot positioning and movement details, affected the precision of interpretation and may have influenced execution accuracy. These findings highlight the importance of optimizing visual presentation and screen visibility when delivering movement-based content remotely.

Theme 2: Platform Efficiency

Optimized Efforts

Participants identified several benefits associated with the PromoSanté platform, particularly its capacity to reduce logistical and time-related constraints. Virtual delivery minimized transportation demands and mobility-related barriers, thereby enhancing accessibility for individuals with physical limitations. This reduction in practical burdens was perceived as facilitating consistent attendance and supporting adherence to program sessions.

Flexibility Enhanced by Asynchronous Modality

Asynchronous sessions were perceived as an efficient and adaptable modality, enabling participants to integrate program activities into their daily routines. The absence of fixed schedules allowed activities to be completed at times that aligned with individual availability and responsibilities. This flexibility was particularly salient for one participant who served as a caregiver, as it enabled continued engagement in the program while fulfilling caregiving duties. These findings underscore the role of asynchronous delivery in supporting accessibility, autonomy, and sustained participation.

Initial Learning Curve and Setup Efforts

Despite the benefits of virtual delivery, initial engagement with the PromoSanté platform was challenging for some participants, particularly those with limited digital literacy. Over time, however, participants reported increased familiarity and growing confidence in using the platform's features, reflecting improved learnability. Early technical support from the research team was identified as instrumental in facilitating this transition. Participants also emphasized the importance of the initial training session and printed user guide in supporting participants' understanding of the platform's structure and core functionalities. These resources were especially valuable during the onboarding phase and contributed to more efficient and confident interactions as use progressed.

Technical Issues With the Videoconference Platform BigBlueButton

Participants reported technical difficulties during some synchronous occupational therapy sessions, particularly audio and video delays. These disruptions affected conversational flow, limited interaction, and caused overlaps and misunderstandings. Several participants expressed frustration, indicating that unresolved connectivity issues could compromise communication quality and negatively impact overall usability.

Theme 3: Participants' Satisfaction With the Use of the Platform

Positive User Experience

Overall, participants reported high satisfaction with the platform and described the experience as enjoyable. Despite identified challenges, both the content and delivery modalities were perceived as acceptable and engaging throughout the program. Importantly, several participants expressed willingness to continue accessing and using the platform beyond the study period and to recommend it to others, indicating strong acceptability and positive behavioral intention.

Sense of Confidence

Several participants reported that their positive experience extended beyond usability to increased self-confidence in engaging with digital technologies and participating in similar programs. Successfully completing the sessions reinforced their perceived competence and strengthened confidence in their ability to adopt and sustain technology-based health interventions.

Discussion

Principal Findings

This study aimed to iteratively evaluate and refine the usability of the PromoSanté digital platform under both controlled and real-world conditions. Consistent with these objectives, the findings demonstrate that the platform achieved high usability and functional integrity across phases, while iterative refinements addressed barriers identified during preliminary testing. High adherence rates during full implementation indicate

XSL•FO
RenderX

a successful integration of the platform into routine program delivery, and qualitative findings highlight strong user satisfaction alongside minor usability challenges that require optimization.

The first phase demonstrated high perceived usability, with a SUS score of 83 out of 100. This rating exceeds most usability scores reported in a scoping review of digital technologies for older adults, in which values were generally above average but commonly below 80 [44]. Despite this strong overall performance, participants identified challenges with touch-based interactions, particularly related to small interface elements. These findings informed targeted refinements, including enlarging touch targets and optimizing button placement to enhance accessibility and reduce interaction difficulty. Such adjustments are consistent with prior research, emphasizing the importance of larger interactive elements to improve touchscreen usability among older adults [45,46]. Importantly, these usability issues did not impede task completion and were not perceived as major barriers to platform use. Instead, the refinements highlighted the critical role of appropriate supporting equipment—such as tripods for participants and microphones for health care professionals—in optimizing interaction quality during synchronous sessions. This observation aligns with prior findings by Aoun et al [47], who underscored that video-based telehealth modalities require careful technical preparation and ongoing support to ensure an effective and satisfactory user experience. Participants emphasized the importance of providing a user guide—particularly in printed format—to support future users and enhance platform usability. When combined with initial training, these resources were perceived as especially valuable for individuals with limited digital experience, facilitating onboarding and reducing early uncertainty. This finding aligns with prior research demonstrating that structured guidance and instructional materials promote adoption and effective use of new technologies in real-world contexts [48,49].

High adherence rates for both synchronous (85.4%) and asynchronous (83.3%) sessions further support the platform's effectiveness in delivering accessible virtual content. Similar adherence levels have been reported in studies by Granet et al [50,51] evaluating online physical exercise programs using combined modalities. However, in those studies, asynchronous content was delivered through a web platform, whereas synchronous sessions required access via external email links. In contrast, PromoSanté integrates both modalities within a single interface. This unified design likely enhances usability by reducing navigation complexity and supporting seamless participation across program components.

The results from UTAUT2 showed higher ratings on PE, EE, FC, and HM. From a usability standpoint, these domains suggest that the platform was perceived as useful, easy to use, enjoyable, and adequately supported. Consistent with this interpretation, prior studies applying the UTAUT2 framework in older adult populations have identified PE, EE, HM, and FC as key predictors of behavioral intention to adopt digital technologies [52-54].

Qualitative findings further substantiated these results, reflecting positive perceptions of usability across the dimensions of effectiveness, efficiency, and satisfaction. Continuous access to platform content was particularly valued, consistent with evidence that uninterrupted digital access enhances autonomy and accessibility among older adults [55]. The asynchronous modality emerged as a central facilitator, reducing time-related constraints and accommodating individual routines. This flexibility is especially relevant in aging populations marked by heterogeneous schedules and fluctuating energy levels [56]. Collectively, these findings suggest that, when thoughtfully designed, digital interventions can align well with older adults' daily contexts, challenging assumptions that this population is inherently resistant to technology-based solutions. Qualitative findings further reinforced these results, highlighting participants' positive perceptions of platform usability across the dimensions of effectiveness, efficiency, and satisfaction. Continuous access to platform content was particularly valued, aligning with prior research demonstrating that uninterrupted access to digital resources enhances accessibility and supports autonomous engagement among older adults [55]. In addition, the asynchronous modality was identified as a key benefit, as it reduced time-related constraints and accommodated individual routines. This flexibility is especially pertinent in aging populations characterized by heterogeneous daily schedules and varying levels of energy and availability [56]. This suggests that older adults' living contexts may be more supportive of digital interventions than commonly assumed.

Limitations

Despite promising findings, several limitations should be acknowledged. First, most participants reported regular use of digital devices, suggesting relatively high levels of digital literacy. Additionally, findings must be interpreted within the study's implementation context, where digital literacy workshops, technical support, and tablets were provided, and most participants had access to reliable internet connectivity. These resources may have contributed to participants' ability to engage with the platform and should therefore be considered FC for its use. Thus, the positive usability experiences observed in this study may not fully reflect platform use in settings where such support and resources are not readily available, particularly among older adults with lower digital literacy.

In routine practice, inconsistent access to such resources may affect scalability and long-term sustainability. Organizations implementing such a platform should therefore carefully consider and evaluate the resources required to support adoption and sustained use at scale. Second, resource constraints limited implementation to 30 participants, which may not fully represent the social, physical, and cognitive heterogeneity of the broader aging population. Broader deployments across more diverse demographic and geographic contexts are therefore warranted. Although engagement remained high, a slight decline in platform visits was observed over the 16-week period. Longer-term studies are needed to evaluate sustained use, as telehealth engagement may decrease over extended follow-up. Finally, because tablets were provided, future research should assess usability on participants' personal devices to better reflect real-world conditions.

XSL•FO
RenderX

Future Perspectives

Findings from this study indicate that the digital platform supporting virtual delivery of the PromoSanté program is both usable and well accepted by older adults. Building on these results, a research and innovation grant has been secured to further optimize the platform in preparation for large-scale deployment. The next phase will enhance technological maturity through 2 stages. First, the current proof of concept will be transformed into a more secure, modular, and scalable platform capable of supporting multiple programs across diverse user groups, with advanced user engagement functionalities integrated and iteratively validated with end users. Second, the finalized prototype will undergo comprehensive usability testing under real-world conditions to confirm performance, robustness, and scalability.

Conclusion

This study evaluated and refined the usability of a co-designed digital platform supporting the PromoSanté program through an iterative testing process. Initial findings showed high user satisfaction and underscored the importance of a clear, accessible interface and adapted equipment for touch-based interaction. Real-world implementation confirmed the platform's ability to enhance accessibility and flexibility through integrated synchronous and asynchronous delivery, while highlighting the importance of appropriate support and implementation conditions to enable sustained use among older adults. Future large-scale implementation and continued technological optimization will be critical to confirming long-term engagement, scalability, and integration within routine service delivery.

Acknowledgments

The authors would like to acknowledge the support of the research team members, specifically the author ML, who holds a Tier 1 Canada Research Chair in Social Participation and Connection for Older Adults (CRC-2022 - 00331; 2023 - 2030), as well as Constance Calaya, Keren Pelen-Tobar, Antoine Guillerand, and collaborators among the occupational therapy team who contributed to participant recruitment and project coordination. Finally, the authors thank all participants who agreed to give their time to participate in the study and to answer the questionnaires.

Generative AI Use

Generative AI tool (ChatGPT [OpenAI]) was used occasionally for linguistic assistance (activity 1, column 1, according to the AI Activities Table included in the journal guidelines).

Funding

This research was funded by the Fonds de Recherche du Québec—Santé (FRQS; grant number 307980). The study has been funded for a 3-year period, starting in April 2021.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: NK, PB, ML, MA-L, CV, PR, RN
Data curation: PB
Formal analysis: NK, PB, CC, SB
Funding acquisition: PB, ML
Investigation: NK, CC
Methodology: NK, PB, ML, CV, PR, RN
Project administration: PB, CC, SB
Resources: PB
Software: SB
Supervision: PB
Validation: PB
Visualization: NK, CV, CC, PR, RN, SB
Writing – original draft: NK, PB, ML, MA-L, CV, CC, PR, RN, SB
Writing – review & editing: NK, PB, ML, MA-L, CV, CC, PR, RN, SB

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Ageing and health. World Health Organization. 2025 Oct 1. URL: https://www.who.int/en/news-room/fact-sheets/detail/ageing-and-health [accessed 2026-10-28]
2. Booth FW, Roberts CK, Thyfault JP, Ruegsegger GN, Toedebusch RG. Role of inactivity in chronic diseases: evolutionary insight and pathophysiological mechanisms. Physiol Rev 2017 Oct 1;97(4):1351-1402. [doi: 10.1152/physrev.00019.2016] [Medline: 28814614]
3. Lyu C, Siu K, Xu I, Osman I, Zhong J. Social isolation changes and long-term outcomes among older adults. JAMA Netw Open 2024 Jul 1;7(7):e2424519. [doi: 10.1001/jamanetworkopen.2024.24519] [Medline: 39046736]
4. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Soc Sci Med 2000 Sep;51(6):843-857. [doi: 10.1016/s0277-9536(00)00065-4] [Medline: 10972429]
5. Hart DA. Learning from human responses to deconditioning environments: improved understanding of the “use it or lose it” principle. Front Sports Act Living 2021;3:685845. [doi: 10.3389/fspor.2021.685845] [Medline: 34927066]
6. Nicklen B, Delbaere K, Ellmers TJ. Is frailty associated with increased concerns about falling and activity restriction in community-dwelling older adults? A systematic review. J Frailty Aging 2025 Feb;14(1):100002. [doi: 10.1016/j.tjfa.2024.100002] [Medline: 39855888]
7. Hayashi T, Umegaki H, Makino T, et al. Combined impact of physical frailty and social isolation on rate of falls in older adults. J Nutr Health Aging 2020;24(3):312-318. [doi: 10.1007/s12603-020-1316-5] [Medline: 32115613]
8. Sulandari S, Coats RO, Miller A, Hodkinson A, Johnson J. A systematic review and meta-analysis of the association between physical capability, social support, loneliness, depression, anxiety, and life satisfaction in older adults. Gerontologist 2024 Nov 1;64(11):gnae128. [doi: 10.1093/geront/gnae128] [Medline: 39233622]
9. Valenzuela PL, Saco-Ledo G, Morales JS, et al. Effects of physical exercise on physical function in older adults in residential care: a systematic review and network meta-analysis of randomised controlled trials. Lancet Healthy Longev 2023 Jun;4(6):e247-e256. [doi: 10.1016/S2666-7568(23)00057-0] [Medline: 37182530]
10. Meredith SJ, Cox NJ, Ibrahim K, et al. Factors that influence older adults’ participation in physical activity: a systematic review of qualitative studies. Age Ageing 2023 Aug 1;52(8):afad145. [doi: 10.1093/ageing/afad145] [Medline: 37595070]
11. Yang Y, Gao Y, An R, Wan Q. Barriers and facilitators to exercise adherence in community-dwelling older adults: a mixed-methods systematic review using the COM-B model and theoretical domains framework. Int J Nurs Stud 2024 Sep;157:104808. [doi: 10.1016/j.ijnurstu.2024.104808] [Medline: 38823146]
12. Balki E, Hayes N, Holland C. Effectiveness of technology interventions in addressing social isolation, connectedness, and loneliness in older adults: systematic umbrella review. JMIR Aging 2022 Oct 24;5(4):e40125. [doi: 10.2196/40125] [Medline: 36279155]
13. Walha R, Koubaa N, Chagnon M, et al. e-Health interventions for promoting physical activity in aging adults: a scoping review. Telemed J E Health 2025 May;31(5):531-539. [doi: 10.1089/tmj.2024.0414] [Medline: 39757866]
14. Heart T, Kalderon E. Older adults: are they ready to adopt health-related ICT? Int J Med Inform 2013 Nov;82(11):e209-e231. [doi: 10.1016/j.ijmedinf.2011.03.002] [Medline: 21481631]
15. Peek STM, Wouters EJM, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJM. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014 Apr;83(4):235-248. [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
16. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
17. ISO 9241-11:1998: Ergonomic requirements for office work with visual display terminals (VDTs)—Part 11: Guidance on usability. : International Organization for Standardization (ISO); 1998 URL: https://www.iso.org/standard/16883.html [accessed 2026-10-28]
18. Laukka E, Lakoma S, Harjumaa M, et al. Older adults’ preferences in the utilization of digital health and social services: a qualitative analysis of responses to open-ended questions. BMC Health Serv Res 2024 Oct 4;24(1):1184. [doi: 10.1186/s12913-024-11564-1] [Medline: 39367429]
19. Mannheim I, Schwartz E, Xi W, et al. Inclusion of older adults in the research and design of digital technology. Int J Environ Res Public Health 2019 Oct 2;16(19):3718. [doi: 10.3390/ijerph16193718] [Medline: 31581632]
20. Chu RYK, Cheng CHG, Leung KMC, Ezulike JCD, Montayre J. Technology-enabled interventions promoting healthy ageing: an umbrella review. Discov Public Health 2025 Nov 5;22(1):661. [doi: 10.1186/s12982-025-01066-8]
21. Tajudeen FP, Bahar N, Tan MP, Peer Mustafa MB, Saedon NI, Jesudass J. Understanding user requirements for a senior-friendly mobile health application. Geriatrics (Basel) 2022 Oct 1;7(5):110. [doi: 10.3390/geriatrics7050110] [Medline: 36286212]
22. Heponiemi T, Kainiemi E, Virtanen L, et al. Predicting internet use and digital competence among older adults using performance tests of visual, physical, and cognitive functioning: longitudinal population-based study. J Med Internet Res 2023 May 5;25:e42287. [doi: 10.2196/42287] [Medline: 37145836]
23. Sonderegger A, Schmutz S, Sauer J. The influence of age in usability testing. Appl Ergon 2016 Jan;52:291-300. [doi: 10.1016/j.apergo.2015.06.012] [Medline: 26360221]
24. PromoSanté [Article in French]. URL: https://www.promosante.ca/ [accessed 2026-02-15]

XSL•FO
RenderX

25. PROMOSED: co-creation of a telehealth program focused on self-management of the health and social participation of seniors at risk of deconditioning. Dimensions. URL: https://app.dimensions.ai/details/grant/grant.9721602 [accessed 2026-10-28]
26. Lucifora C, Villar E. Multi-dimensional healthy aging interventions: evidence from an age-friendly community program in Italy. Ageing Int 2024 Dec;49(4):749-771. [doi: 10.1007/s12126-024-09567-8]
27. Chi YC, Wu CL, Liu HT. Effect of a multi-disciplinary active aging intervention among community elders. Medicine (Baltimore) 2021 Dec 23;100(51):e28314. [doi: 10.1097/MD.0000000000028314] [Medline: 34941126]
28. Cayola L, Macías JA. Systematic guidance on usability methods in user-centered software development. Inf Softw Technol 2018 May;97:163-175. [doi: 10.1016/j.infsof.2018.01.010]
29. Koubaa N, Levasseur M, Aubertin-Leheudre M, Verchère C, Corbin C, Boissy P. Attitudes and preferences of older adults and healthcare professionals regarding a multidisciplinary hybrid telehealth program for healthy aging. Front Public Health 2026;14:1777557. [doi: 10.3389/fpubh.2026.1777557] [Medline: 41923753]
30. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR mHealth uHealth 2023 Sep 21;11:e43186. [doi: 10.2196/43186] [Medline: 37733401]
31. Nielsen J. Usability Engineering: Morgan Kaufmann; 1993.
32. Alwashmi MF, Hawboldt J, Davis E, Fetters MD. The iterative convergent design for mobile health usability testing: mixed methods approach. JMIR mHealth uHealth 2019 Apr 26;7(4):e11656. [doi: 10.2196/11656] [Medline: 31025951]
33. Nielsen J. Why you only need to test with 5 users. NN/Group (Nielsen Norman Group). 2000 Mar 18. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2026-10-16]
34. Gronier G, Baudet A. Psychometric evaluation of the F-SUS: creation and validation of the French version of the System Usability Scale. Int J Hum Comput Interact 2021 Oct 2;37(16):1571-1582. [doi: 10.1080/10447318.2021.1898828]
35. Brooke J. SUS: a “quick and dirty” usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry: CRC Press; 1996:189-194.
36. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
37. Sauro J. 10 things to know about the System Usability Scale (SUS). MeasuringU. 2013 Jun 18. URL: https://measuringu.com/10-things-sus/ [accessed 2025-11-02]
38. Gao M. Multi-cultural usability assessment with System Usability Scale [Dissertation]. : Rice University; 2019 URL: https://hdl.handle.net/1911/106168 [accessed 2026-10-28]
39. Stanton NA. Hierarchical task analysis: developments, applications, and extensions. Appl Ergon 2006 Jan;37(1):55-79. [doi: 10.1016/j.apergo.2005.06.003] [Medline: 16139236]
40. van den Haak M, De Jong M, Schellens PJ. Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue. Behav Inf Technol 2003 Sep;22(5):339-351. [doi: 10.1080/0044929031000]
41. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012 Mar 1;36(1):157-178. [doi: 10.2307/41410412]
42. Pagé I, Roos M, Collin O, et al. UTAUT2-based questionnaire: cross-cultural adaptation to Canadian French. Disabil Rehabil 2023 Feb;45(4):709-716. [doi: 10.1080/09638288.2022.2037746] [Medline: 35179430]
43. ISO 9241-11:2018: Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts. : International Organization for Standardization (ISO); 2018 URL: https://www.iso.org/standard/63500.html [accessed 2026-10-28]
44. Berry ECJ, Sculthorpe NF, Warner A, Mather JD, Sanal-Hayes NEM, Hayes LD. A scoping review of the feasibility, usability, and efficacy of digital interventions in older adults concerning physical activity and/or exercise. Front Aging 2025;6:1516481. [doi: 10.3389/fragi.2025.1516481] [Medline: 40290578]
45. Yeh P. Exploring touchscreen interface design for home health management products: effects of button size and position on elderly users. Percept Mot Skills 2025 Dec;132(6):1256-1275. [doi: 10.1177/00315125251345594] [Medline: 40440033]
46. Gao Q, Sun Q. Examining the usability of touch screen gestures for older and younger adults. Hum Factors 2015 Aug;57(5):835-863. [doi: 10.1177/0018720815581293] [Medline: 25957042]
47. Aoun B, Ebbert JO, Ramar P, Roellinger DL, Philpot LM. Assessing the ability to use eHealth resources among older adults: cross-sectional survey study. JMIR Form Res 2025 Aug 6;9:e70672. [doi: 10.2196/70672] [Medline: 40768764]
48. Happe L, Sgraja M, Hein A, Quinten V, Diekmann R. Usability and feasibility of a tablet-based e-coach for older adults in rehabilitation units to improve nutrition and physical activity: a prospective pilot study. BMC Geriatr 2023 Sep 19;23(1):578. [doi: 10.1186/s12877-023-04204-6] [Medline: 37726662]
49. Seaton CL, Rush KL, Li EPH, Hasan MK, Fawcus L. Gluu Essentials digital skills training for middle-aged and older adults that makes skills stick: results of a pre-post intervention study. JMIR Aging 2023 Nov 10;6:e50345. [doi: 10.2196/50345] [Medline: 37948115]
50. Granet J, Peyrusqué E, Ruiz F, et al. Online physical exercise intervention in older adults during lockdown: can we improve the recipe? Aging Clin Exp Res 2023 Mar;35(3):551-560. [doi: 10.1007/s40520-022-02329-z] [Medline: 36635450]
51. Granet J, Peyrusqué E, Ruiz F, et al. Web-based physical activity interventions are feasible and beneficial solutions to prevent physical and mental health declines in community-dwelling older adults during isolation periods. J Gerontol A Biol Sci Med Sci 2023 Mar 1;78(3):535-544. [doi: 10.1093/gerona/glac127] [Medline: 35675174]

XSL•FO
RenderX

52. Macedo IM. Predicting the acceptance and use of information and communication technology by older adults: an empirical examination of the revised UTAUT2. Comput Human Behav 2017 Oct;75:935-948. [doi: 10.1016/j.chb.2017.06.013]
53. Phibbs CL, Rahman SSM. A synopsis of "The impact of motivation, price, and habit on intention to use IoT-enabled technology: a correlational study". J Cybersecur Priv 2022 Aug 18;2(3):662-699. [doi: 10.3390/jcp2030034]
54. Wu C, Lim GG. Investigating older adults users' willingness to adopt wearable devices by integrating the technology acceptance model (UTAUT2) and the Technology Readiness Index theory. Front Public Health 2024;12:1449594. [doi: 10.3389/fpubh.2024.1449594] [Medline: 39421816]
55. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
56. Pizarro-Mena R, Rotarou ES, Baracaldo-Campo HA, et al. Implementation of telehealth among older people: a challenge and opportunity for Latin America and the Caribbean—a literature review. Healthcare (Basel) 2025 Oct 23;13(21):2680. [doi: 10.3390/healthcare13212680] [Medline: 41228047]

Abbreviations

EE: effort expectancy
FC: facilitating conditions
HM: hedonic motivation
PE: performance expectancy
SUS: System Usability Scale
UTAUT2: unified theory of acceptance and use of technology

Edited by M Antoniou; submitted 27.Mar.2026; peer-reviewed by L Philpot; revised version received 29.Sep.2026; accepted 13.Oct.2026; published 24.Nov.2026.

Please cite as:
Koubaa N, Levasseur M, Aubertin-leheudre M, Roberge P, Verchère C, Ndjaboue R, Brière S, Corbin C, Boissy P
Usability of a Digital Platform Supporting a Telehealth Intervention for Physical and Social Deconditioning in Older Adults: Mixed Methods Study
JMIR Aging 2026;9:e96315
URL: https://aging.jmir.org/2026/1/e96315
doi:10.2196/96315

© Nesrine Koubaa, Mélanie Levasseur, Mylène Aubertin-leheudre, Pasquale Roberge, Céline Verchère, Ruth Ndjaboue, Simon Brière, Cinthia Corbin, Patrick Boissy. Originally published in JMIR Aging (https://aging.jmir.org), 24.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Inclusive Design Guidelines for Digital Health Interventions for Older Adults: Fuzzy eDelphi Study

Kerstin Denecke[1], Dr rer nat; Octavio Rivera-Romero[2], PhD; Lana Cvijic[1], BSc; Beatrice Kaufmann[1], Dipl.Design; Emilene Zitkus[3], PhD; Yasemin Afacan[4], PhD; Marina Bruderer-Hofstetter[5], Dr sc; Ruben Debeuf[6], MSc; Joy Goodman-Deane[7], PhD; Laura Haase[8,9], M. Eng.; Mustafa Ozkaynak[10], PhD; Valérie Santschi[11], PhD; Emanuele Seminerio[12], MSc; Georgia Vardali[13], MSc; Tiina Vuohijoki[14], PhD; Enrique Dorronzoro Zubiete[2], PhD; Carolyn Petersen[15], MS, MBI, FAMIA, FIAHSI

[1]Institute Patient-centered Digital Health, Bern University of Applied Sciences, Quellgasse 21, Biel, Switzerland

[10]College of Nursing, University of Colorado | Anschutz Medical Campus, Aurora, CO, United States

[11]School of Nursing Sciences, HES-SO University of Applied Sciences and Arts Western Switzerland, Lausanne, Switzerland

[12]IRCCS Azienda Ospedaliera Metropolitana, Genoa, Liguria, Italy

[13]Lab of Medical Physics and Digital Innovation, Aristotle University of Thessaloniki, Thessaloniki, Central Macedonia, Greece

[14]Information Technology and Communication Sciences, Tampere University, Tampere, Pirkanmaa, Finland

[15]College of Medicine and Science, Mayo Clinic, Rochester, MN, United States

[2]Technology Electronic Department, Universidad de Sevilla, Sevilla, Andalusia, Spain

[3]School of Design and Creative Arts, Loughborough University, Loughborough, England, United Kingdom

[4]Departments of Neuroscience and Interior Architecture and Environmental Design, Bilkent University, Ankara, Türkiye

[5]Institute on Ageing and Division of Physiotherapy, Bern University of Applied Sciences, Bern, Switzerland

[6]Rehabilitation Research Group, Department of Physiotherapy, Human Physiology and Anatomy, Vrije Universiteit Brussel, Brussels, Brussels Capital, Belgium

[7]Department of Engineering, University of Cambridge, Cambridge, England, United Kingdom

[8]Institute for Medical Informatics and Biometry, Technische Universität Dresden, Dresden, Saxony, Germany

[9]Department of Cooperative Studies - Computer Science, Berlin School of Economics and Law, Berlin, Berlin, Germany

Corresponding Author:
Kerstin Denecke, Dr rer nat
Institute Patient-centered Digital Health, Bern University of Applied Sciences, Quellgasse 21, Biel, Switzerland

Abstract

Background: People aged 60 years and older represent a rapidly growing population worldwide. Aging is associated with complex health and social needs, and quality of life is influenced by multiple factors, including access to health care services, caregiver support, social participation, and community engagement. Digital health interventions (DHIs), such as mobile apps and wearables, have the potential to support some of these needs. However, their adoption among older adults remains limited, partly due to usability and accessibility barriers, and insufficient consideration of inclusive design principles.

Objective: This study aimed to assess expert consensus on a novel set of inclusive design guidelines for DHIs intended for older adults. An expert panel with expertise in digital health, user experience, design, and geriatrics was convened to evaluate the guidelines.

Methods: The initial guidelines were developed through a multimethod approach in previous work and comprised 13 guidelines. A 3-round eDelphi study was conducted with experts from academia, industry, and public and private health sectors. Participants rated each guideline item in terms of utility, clarity, and ease of use on a 7-point Likert scale, with responses transformed into a fuzzy scale for analysis. Consensus criteria were predefined, and the eDelphi process and reporting followed the Guidance on Conducting and Reporting Delphi Studies (CREDES). Following the eDelphi rounds, experts contributed practical examples to support implementation of the guidelines. These examples underwent an expert review and revision process that was not formally standardized.

Results: A total of 28 experts participated in round 1, 20 in round 2, and 19 in round 3. Across the 42 guideline-by-assessment-aspect evaluations (14 guidelines evaluated for utility, clarity, and ease of use), consensus was reached for 33 (75%). All guideline items achieved consensus regarding their utility and clarity. However, approximately 65% (9/14) of guidelines did not reach consensus with respect to their ease of use. Expert feedback indicated that the absence of concrete examples was an important barrier to applying the guidelines in practice. Practical examples were therefore incorporated into the

XSL•FO
RenderX

final version of the guidelines with the intention to facilitate their interpretation and application; their effect on ease of use was not formally assessed.

Conclusions: The expert panel confirmed that inclusive design of DHIs for older adults should address not only user interface elements but also broader health care, organizational, and contextual factors. Further research is needed to evaluate the guidelines in practice and determine their usability and effectiveness in the development of DHIs for older adults.

(JMIR Aging 2026;9:e99657) doi:10.2196/99657

KEYWORDS

inclusive design; digital health interventions; older adults; inclusive design guidelines; eDelphi

Introduction

Background

Across the globe, individuals aged 60 years and older represent the most rapidly expanding segment of the population [1]. This demographic transition, commonly called population aging, involves a rising proportion and absolute number of older adults alongside a relative reduction in younger age groups [2]. The trend largely results from longer life expectancy; projections estimate that by 2050, the worldwide population aged 60 years and older will reach 2.1 billion [3]. While population aging first became apparent in high-income nations, it is now increasingly evident in low- and middle-income countries [2].

As people age, they might face various health and social considerations, some of which can be quite challenging. The quality of life of older adults is closely linked to factors such as access to appropriate health services, social participation, caregivers' support, and opportunities for meaningful involvement within their communities [4]. Moreover, older individuals face increased risks for multiple health conditions, including cognitive decline, persistent pain, and noncommunicable diseases such as cardiovascular disease, diabetes mellitus, and cancer [3,5]. Consequently, fostering health and well-being in later life is central to healthy aging [6]. Responding effectively to the complex health demands of an aging population is critical not only for personal fulfillment but also for financial stability [7].

Digital health interventions (DHIs) are defined as the application of digital technologies to provide health services, disseminate information, or offer health-related support. They provide one way to address the needs of an aging population [7]. DHIs are being developed to enhance clinical outcomes, to optimize care delivery, and to strengthen overall performance of the health care system. As such, they offer innovative, potentially cost-effective strategies to mitigate the health and social challenges faced by older adults [8]. By facilitating early identification of health conditions, supporting preventive strategies, and enabling individualized care, DHIs may also promote older adults' active participation in managing their own health [9,10]. The World Health Organization categorizes DHIs according to their function within the health system [11]. Among these are interventions designed for personal health monitoring, including mobile health (mHealth) applications, smartphone-based sensors, wearable devices, and web-based platforms. Such technologies allow older adults to track vital signs, record and manage physical activity and medication use, and manage other health-related indicators. Recent findings suggest that interest in digital technologies among older adults is increasing and that DHIs can offer benefits for this population [12,13]. Nevertheless, DHIs' effectiveness is contingent upon ensuring accessibility and usability [14,15]. However, tailoring these tools to meet age-related needs continues to be a significant developmental challenge. Applying principles of inclusive design represents a key strategy for achieving this objective [16].

Inclusive design is an approach used to design products and services that meet the needs of a diverse range of users. Inclusive design meets these needs by accounting for variations in socioeconomic background, gender, age, ethnicity, and language, among other characteristics [17]. Such an approach is especially necessary for older adults, given the absence of a single, representative profile of aging. The aging process differs across individuals and is influenced by both physical surroundings and social contexts that shape health-related behaviors and evolving needs [6]. In accordance, DHIs that incorporate inclusive design principles are more likely to meet users' needs [16]. Tools that overlook the specific capabilities and preferences of older adults risk limited usability, reduced adoption, and diminished impact among this expanding demographic group [18,19].

Previous Work and Scientific Contribution

Previous contributions have proposed several checklists and recommendations that focus on inclusive design of DHIs for older adults. Most of these contributions focus on the design of mobile applications that target older adults as users, and they primarily address interface-related aspects such as navigation, visual design, and interaction mechanisms [20-23]. For instance, Amouzadeh et al [22] provided a set of essential design elements of mHealth applications for older adults, including simplified navigation, enlarged text and touch targets, and incorporation of assistive technologies in the DHIs to improve accessibility. In the area of accessibility, Stephens proposed a comprehensive checklist that covers accessibility needs of older adults related to vision, motor skills, hearing, and cognitive functioning [24]. Similarly, Boot et al [18] developed an extensive set of design recommendations based on the diverse needs of older adults, while Brewer [25] further expanded the field by examining accessibility in nonvisual technologies. Going beyond interface and accessibility aspects, Czaja et al [26] introduced a framework that identifies societal and contextual factors that influence the adoption of DHIs among older adult users.

Despite these valuable contributions, existing guidance remains predominantly focused on usability, user experience (UX), and

XSL•FO
RenderX

accessibility at the level of the user interface. Broader technical, organizational, and ethical aspects such as the integration of DHIs into health care workflows, effective communication of DHI-related information to users and health care professionals, and meaningful involvement of older adults through co-design are often underrepresented, and their integration into a single, comprehensive framework is lacking [27]. Abhari et al [28] reported that current guidelines predominantly emphasize usability and UX of DHIs for older adults, while aspects such as integration and interoperability, data privacy and security, data quality, and long-term adoption receive limited attention in the existing frameworks and guidelines.

In response to these gaps, our work aimed to develop comprehensive inclusive design guidelines for older adults. In this context, inclusive design extends beyond the creation of accessible and usable interfaces. It also encompasses technical, organizational, contextual, and health care–related factors that influence whether older adults can access, understand, and adopt DHIs.

The initial version of the guidelines was developed using a multimethod approach. First, we conducted a scoping literature review that investigated DHIs and their inclusive design elements applied with older adult users in mind [29]. In addition, we conducted interviews with experts to collect their practical experience [30] and a focus group with older adults to identify the needs of the target user group regarding DHIs. By these 3 steps, we collected factors to be considered in the design and development of DHIs for older adults. Based on the factors, we defined the first version of guidelines in a workshop, using a nominal group technique. In this paper, we report on the process of assessing expert consensus on the guideline items through an eDelphi study and present the final version of design guidelines.

Methods

Overview of the Delphi Process

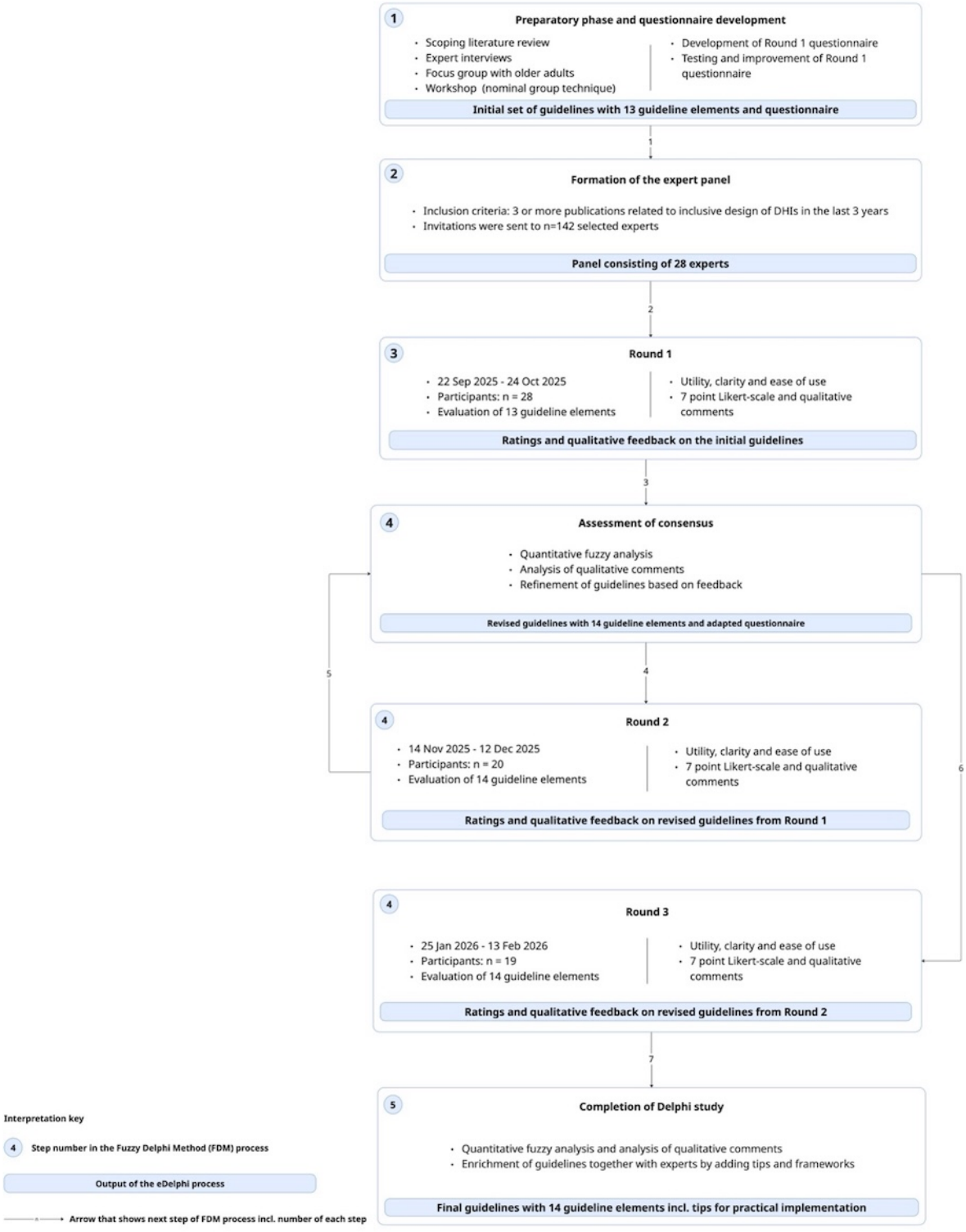
A fuzzy Delphi method (FDM) as defined by Roldán López de Hierro et al [31] was used to reach consensus on the elements of our design guideline. FDM is an adaptation of the traditional Delphi approach that incorporates fuzzy logic to better manage uncertainty and imprecision in expert assessments [31]. The use of defined fuzzy numbers allows for a more detailed and flexible representation of consensus among experts. Due to its advantages, FDM has become widely used in various fields, including health care [32-35].

Compared to the traditional Delphi Method, the FDM offers benefits, especially when it comes to handling subjectivity and uncertainty in expert feedback. While conventional Delphi techniques depend on rigid numerical scales and frequently demand multiple rounds to reach an agreement, FDM integrates fuzzy logic to capture more subtle and realistic perspectives from participants [36]. By using linguistic variables transformed into triangular fuzzy numbers, experts can share their views more flexibly, which effectively captures the natural ambiguity found in complex or delicate subjects. The FDM uses a systematic quantification process of the consensus level that minimizes the subjective interpretation and enhances the reliability of the findings [37].

Our FDM process consisted of 5 steps: (1) preparatory phase, (2) formation of the expert panel, (3) conduction of round 1, (4) assessment of consensus and conducting subsequent rounds, and (5) completion of the Delphi study (Figure 1 contains the visual overview). To organize our study, we followed the recommendations for conducting and reporting of Delphi studies [30,31].

XSL•FO
RenderX

Figure 1. eDelphi process flowchart.



Ethical Considerations

The study design was submitted to the ethics committee of the canton of Bern, which confirmed that ethics approval was not required (Req-2025 - 00387). The participants were informed in each questionnaire that they consent to the anonymous data analysis by submitting their responses to the questionnaire. Email addresses were only collected in rounds 1 and 2 and available to 2 coauthors (KD and ORR), who analyzed the data and shared the questionnaires with the Delphi panel. Data were stored on local machines. We followed principles of data minimization and collected only limited demographic data in round 1.

Preparatory Phase

Development of the Initial Set of Guidelines

In the following, we refer to the term “guideline” as the individual items in our inclusive design guideline, ie, a single

XSL•FO
RenderX

description of recommendations for the design and development of a DHI for older adults, together with concrete implementation guidance.

To develop the initial set of guidelines, findings from a scoping review, expert interviews, and a focus group with older adults were reviewed and organized into candidate inclusive design principles by one author (LC). Then, the initial guidelines were developed during a workshop, using the nominal group technique. Three authors (KD, LC, and BK) participated in the workshop. The workshop had the duration of 3 hours, and it began with an introduction to the workshop objectives and clarification of relevant terminology with the aim to reach the mutual understanding among participants. This was followed by a phase in which each participant independently formulated a set of guidelines. In the next step, each participant presented their proposed guidelines to the group. The candidate guideline items were then jointly compared, discussed, grouped, and mapped according to their thematic similarities. Through this discussion, participants defined a common set of guidelines and prioritized the most relevant guideline items to be included in the initial guidelines. Each candidate guideline item was considered in terms of whether it should be retained, adapted, merged with another item, removed, or reformulated as a new guideline item. In the final stage of the workshop, the selected guidelines were refined and finalized. Through this iterative process, findings from the three evidence sources were synthesized into an initial set that consisted of 13 guideline items.

We decided to split the 13 guidelines into 2 groups. One group focused on the health care ecosystem and aimed to provide a broader perspective on how DHIs can be designed for successful integration into established health care processes. It highlighted the role of health care professionals in recommending suitable DHIs to their patients. It addressed other important aspects, such as costs and support needed to enable older adults to use the DHIs confidently. These guidelines were essential to ensure that older adults have the information and support they need to engage with a DHI and integrate it into their daily routines. The 3 guidelines in this group concerned the integration of the DHI into health care processes, information about the DHI, and support, training, and peer learning.

The second group of guidelines referred to the design process of the DHI itself. This group outlined various aspects that designers and developers should consider when creating a DHI for this population. The 10 guidelines in this group concern co-design and usability testing, onboarding and in-app support, visual design and navigation, accessibility, personalization, motivation and user engagement, inclusive language and trustworthy content, social element, safety, and data security and privacy.

Questionnaire Development

The questionnaire for the first round of the eDelphi study contained the initial set of guidelines. These guidelines formed parts 1 and 2 of the first-round questionnaire. The questionnaire consisted of 5 parts:

- Part 1: Guidelines addressing societal and health care system factors (3 guidelines)
- Part 2: Guidelines addressing the design of a DHI (10 guidelines)
- Part 3: Recommended practical tools, standards, and frameworks for inclusive design (1 open question)
- Part 4: Feedback on the complete set of guidelines (intention, completeness, and missing guidelines)
- Part 5: Demographic data (gender, county of residence, field of expertise, working sector, and years of professional experience)

In parts 1 and 2, participants were shown each guideline, including its description and tips for applying it in a DHI for older adults. Following the methodology of Quiñones et al [38], the participants were asked to judge utility (This guideline is useful for designing and evaluating DHIs for older adults), clarity (This guideline is clearly formulated), and ease of use (This guideline is easy to apply in practice when working with DHIs for older adults) on a 7-item Likert scale (from “strongly disagree” to “strongly agree”). Also, the questions in part 4 followed the methodology of Quiñones et al [38]. It is worth noting that in the first round, we used the term “heuristics” rather than “guidelines” in our Delphi questionnaire. The participants’ feedback led to a change in wording. To avoid confusion, we will use the term “guideline” throughout this paper. As each guideline was evaluated along 3 criteria (clarity, utility, and ease of use), we have 39 items in our consensus calculation in round 1, ie, 3 judgments for 13 guidelines.

Parts 3 and 5 of the questionnaire were used only in the first round. In the last round, we added a question about how to order the guidelines in the final version of the recommendation. A web version of the final round 1 questionnaire was created using Microsoft Forms software (Microsoft Corp). Two researchers (ORR and CP) tested its comprehensibility, usability, and accessibility using several browsers on a laptop. Issues were resolved before deployment. Microsoft Forms was also used for the round 2 and round 3 questionnaires. The participants were asked to provide their email addresses in round 1 and round 2 so they could be invited to the following round. The documentation of changes made to the questionnaire that was used to evaluate the guidelines in 3 rounds is available in Multimedia Appendix 1.

Each guideline was assessed using a 7-point Likert scale. Once the experts had completed the questionnaire, each response was converted to its corresponding fuzzy set as set out in Table 1. For example, if an expert selected a “3” as their response to the first question, the fuzzy set (0.1, 0.3, and 0.5) was used in the analysis of this first question. To this end, fuzzy numbers were used via the transformation set out in Table 1. This transformation has been used in other FDM studies that used a 7-point Likert scale [39]. This fuzzy representation produces 3 values: the minimum, the most plausible, and the maximum. This allowed experts to select the option that best reflects their judgment [39].

XSL•FO
RenderX

Table . Fuzzy scale from a 7-point Likert scale.

Level of agreement	Likert scale	Fuzzy scale		
Strongly agree	7	0.9	1	1
Agree	6	0.7	0.9	1
Somewhat agree	5	0.5	0.7	0.9
Neutral	4	0.3	0.5	0.7
Somewhat disagree	3	0.1	0.3	0.5
Disagree	2	0	0.1	0.3
Strongly disagree	1	0	0	0.1

Participants

The Delphi method used a selected panel for feedback, without needing statistical representativeness. We aimed to recruit 20 - 30 participants from all continents except Antarctica. We considered a person eligible for joining the panel if they had 3 or more publications in inclusive design or DHIs for older adults published in the last 3 years. We aimed to include people who are currently actively contributing to the field. Therefore, we did not include anyone who published on the topic more than 3 years ago.

To identify participants for our expert panel, we considered papers published at domain-specific conferences: Medical Informatics Europe, MedInfo, the American Medical Informatics Association Annual (AMIA) Symposium for experts in digital health and medical informatics, Interact and Association for Computing Machinery Human Computer Interaction for design experts, and Gerontechnology for experts in digital solutions for older adults. We contacted all eligible people once we had identified their email addresses. In addition, people from the authors' professional network were invited to join and share the Delphi information. Participants were encouraged to invite additional participants from their networks. Invitations were emailed to 142 experts.

eDelphi Rounds

We conducted 3 eDelphi rounds, as previous research found that 2 or 3 rounds are preferable [40-42]. Round 1 began with invitations sent on September 22, 2025, and ended on October 24, 2025. Round 2 was conducted from November 14, 2025, to December 12, 2025, and round 3 ran from January 25, 2026, to February 13, 2026. Reminders were sent in rounds 2 and 3 after 1 week and 2 weeks, respectively.

We considered that a consensus had been reached when 3 conditions were met simultaneously: (1) there were no comments from the experts on the open-ended question relating to this item, (2) at least 80% of the experts agreed on the item, and (3) the threshold value (d), determined by calculation of triangular fuzzy numbers, did not exceed 0.2 [43]. The threshold value (d) used in this study represents the average distance between each expert's fuzzy value and the aggregated group fuzzy value. This approach follows the standard FDM protocol [44]. The cutoff point of the d-value is 0.20 ($d \leq 0.2$), so items with a d-value of ≤0.2 were considered to have reached a consensus [45].

Bias Prevention Measures

We decided to take several measures to minimize the potential bias throughout our eDelphi study. First, experts completed the questionnaires independently through an online questionnaire, and they did not directly interact with one another during the evaluation of the guidelines. This measure reduced the risk that dominant participants would influence the ratings of other participants. Second, the same 3 rating dimensions and 7-point Likert scale were applied consistently across all 3 rounds, and consensus criteria were defined before the analysis. Before deployment, the questionnaire was tested by 2 researchers (ORR and CP) for comprehensibility, usability, and accessibility, and identified issues were resolved before the start of the eDelphi study. Between rounds, both quantitative consensus results and qualitative comments were reviewed jointly by 3 authors (KD, LC, and BK), and modifications to the guidelines were agreed upon through discussion and consensus rather than being determined by one researcher alone. In addition, members of the panel were recruited from different disciplinary backgrounds and sectors to incorporate a wide range of perspectives and experience levels. To avoid participants only joining the Delphi study because of coauthorship, we only asked in the last questionnaire whether they wanted to contribute to paper writing and become coauthors.

Experts identified through the authors' professional networks were invited using the same recruitment procedure as other potential panel members and were subject to the same eligibility criteria and study procedures. No preferential follow-up or additional communication regarding the assessment of the guidelines was provided to these individuals. No potential conflicts of interest relevant to the assessment of the guidelines were identified among the panel members or study researchers. The 2 researchers involved in the pretest were independent of the study team and did not participate in the subsequent eDelphi panel.

Analysis and Iterative Improvement of Guidelines

Between rounds 1 and 2 and between rounds 2 and 3, three authors (KD, LC, and BK) held the review meetings to examine the quantitative consensus results and the qualitative comments provided by the expert panel. During these meetings, the authors discussed guideline items that had not reached consensus, as well as items for which participants had suggested clarification, rewording, additional content, or improved practical applicability. The quantitative findings and qualitative feedback

XSL•FO
RenderX

were considered jointly, and revisions were agreed upon through discussion and consensus among the 3 authors. Based on these decisions, the wording, descriptions, and implementation guidance of the guideline items were refined before inclusion in the subsequent eDelphi round. Multimedia Appendix 1 provides a comprehensive log of the modifications made to the guideline items across the 3 eDelphi rounds.

After receiving the experts' scores for each item on the Likert scale, each Likert value was converted to the corresponding fuzzy scale value (Table 1). Thus, for an expert j and item i, we would have 3 values: $m1_{ij}$, $m2_{ij}$, and $m3_{ij}$, representing the minimum, most plausible, and maximum values. Next, the fuzzy mean values for each item were calculated from the expert's fuzzy values ($M1_i$, $M2_i$, and $M3_i$). The threshold for each item was then calculated using the following equation:

We calculated the mean d value for each item (d_i) and the percentage of experts who assigned a d_{ij} value of 0.2 or less. In addition, the fuzzy score for each item was calculated using the following equation:

This study was conducted and reported in accordance with the Conducting and Reporting Delphi Studies (CREDES) guidelines [46]. Checklist 1 presents a summary of the CREDES reporting (items 8 - 16) recommendations, including a reference to sections and pages of this manuscript reporting them.

Finalization of the Guidelines

The experts who assessed the guidelines in the eDelphi study were invited to develop examples for each guideline to enhance their applicability and ease of use. For each guideline, 2 experts developed examples, and an additional expert reviewed them. Experts were selected for these roles based on their professional backgrounds and areas of expertise.

The development and review of the examples were conducted remotely using an online guideline document. This process was not based on a formally standardized review procedure or prespecified review criteria; rather, reviewers drew on their professional experience and judgment to provide comments intended to improve the applicability and clarity of the examples. One author (KD) coordinated the process to facilitate the collection and incorporation of comments.

Experts who had developed the examples revised them in response to the reviewers' comments, with further exchanges taking place where considered necessary by the participating experts. Thus, although examples underwent expert review and revision, the number of review exchanges was not prespecified or systematically documented. Finally, the core team (LC, BK, and KD) reviewed the complete set of guidelines and examples.

Results

Characteristics of the Panel

The initial expert panel consisted of 28 participants who responded in the survey's first round. Of the invited experts, 19.7% (28/142) participated in round 1. Table 2 provides an overview of the participants' characteristics. There were 20/28 participants in the second round (retention rate 71.4%) and 19/20 (retention rate 95%) participants in the third round. We describe the demographics of round 1 in more depth. A total of 17/28 (60.7%) panel participants were female, 10/28 (35.7%) were male, and 1/28 (3.6%) preferred not to say. A significant number of participants (11/28, 39.3%) had 11 - 20 years of professional experience in their field. Several sectors were represented in the panel. Most participants (23/28, 82.1%) were from academia, and 5/28 (17.9%) were from industry. The public and private health sectors were each represented by two participants. The panel was multidisciplinary, with participants from a variety of fields. Participants could select multiple fields of expertise, as these areas often overlap in the field of eHealth. Digital health was the most frequent field of expertise (17/28, 60.7%), followed by UX and usability (13/28, 46.4%), human-computer interaction (11/28, 39.3%), gerontechnology (10/28, 35.7%), health care (10/28, 35.7%), medical informatics (9/28, 32.1%), design research (9/28, 32.1%), accessibility (8/28, 28.6%), geriatrics (7/28, 25%), and health communication (4/28, 14.3%). The panel included participants from Europe (21/28, 75%) and North America (7/28, 25%).

XSL•FO
RenderX

Table . Overview of participants' characteristics in the eDelphi panel[a].

Characteristic and category	Round 1[b], n (%)	Round 2[c], n (%)	Round 3[d], n (%)
Gender			
Female	17 (60.7)	12 (60)	12 (63.2)
Male	10 (35.7)	8 (40)	7 (36.8)
Prefer not to say	1 (3.6)	0 (0)	0 (0)
Experience			
<5 years	4 (14.3)	4 (20)	4 (21)
6 - 10 years	4 (14.3)	2 (10)	2 (11)
11 - 20 years	11 (39.3)	7 (35)	6 (31.6)
20+ years	9 (32.1)	7 (35)	7 (36.8)
Sector[a]			
Academia	23 (82.1)	18 (90)	17 (89.5)
Industry	5 (17.9)	1 (5)	1 (5.26)
Public health sector	2 (7.1)	1 (5)	1 (5.26)
Private health sector	2 (7.1)	1 (5)	1 (5.26)
Other	1 (3.6)	0 (0)	0 (0)
Field of expertise[a]			
Human-computer interaction	11 (39.3)	6 (30)	6 (31.6)
UX[e]/usability	13 (46.4)	8 (40)	8 (42.1)
Medical informatics	9 (32.1)	5 (25)	4 (21)
Health communication	4 (14.3)	2 (10)	2 (10.53)
Geriatrics/age studies	7 (25)	4 (20)	4 (21)
Design research	9 (32.1)	5 (25)	4 (21)
Digital health	17 (60.7)	9 (45)	8 (42)
Gerontechnology	10 (35.7)	7 (35)	6 (31.6)
Accessibility	8 (28.6)	5 (25)	4 (21)
Health care	10 (35.7)	5 (25)	5 (26.3)
Other	3 (10.7)	0 (0)	0 (0)
Continent			
Africa	0 (0)	0 (0)	0 (0)
Asia	0 (0)	0 (0)	0 (0)
Europe	21 (75)	17 (85)	16 (84.21)
Australia and Oceania	0 (0)	0 (0)	0 (0)
North America	7 (25)	3 (15)	3 (15.8)
South America	0 (0)	0 (0)	0 (0)

[a]Multiple options could be selected; percentage was calculated by n/N.

[b]Total=28.

[c]Total=20.

[d]Total=19.

[e]UX: user experience.

XSL•FO
RenderX

Resulting Guidelines

All guidelines remained in the evaluation until the final round. Although some had achieved consensus in previous rounds, the experts' comments affected them, so they were included in the list of questions for the next round. The guidelines from round 3 with a brief explanation are shown in Table 3.

Across the 3 eDelphi rounds, the set of guidelines evolved from 13 original items into 14 final guideline items, with no deletions at any stage. Round 1 framed the set as 13 heuristics (H1-H13). Round 2 reframed, restructured, and renumbered the entire set as 14 guidelines (G1-G14), introducing the standardized description/importance/implementation-tips template for every item. Four guideline titles were revised in round 2 to reflect clearer scope: G1 (Integration into health care processes → Integration into health care workflows), G4 (Co-design and usability testing → Participatory design and user involvement), G6 (Visual design and navigation → Visual clarity and navigation), and G7 (Accessibility → Accessibility and functional inclusion).

The transition from 13 to 14 guidelines is explained by a single substantive structural change: the round 1 item H10 (Inclusive language and trustworthy content) was split into 2 separate guidelines in round 2: G10 (Clarity and comprehensibility of language) and G11 (Trustworthiness). G11 is the only new item that entered the set in round 2; all other items survived as renumbered, reformulated, or retitled versions of their round 1 counterparts. The experts had repeatedly flagged this round 1 item as conflating 2 concerns; the 3 authors (KD, LC, and BK) agreed to separate the linguistic-accessibility content (G10) from the trust/credibility content (G11) before recirculation.

Round 3 retained all 14 guidelines. Thirteen were refined and clarified in wording; one was additionally retitled: G9 changed from "Motivation and user engagement" to "Motivation and sustained use of the DHI" to mark the shift from initial adoption to long-term engagement. No item was added or removed in round 3. Every modification recorded in the change log (restructuring, retitling, splitting, and renumbering) was discussed and agreed upon by the 3 authors (KD, LC, and BK; column K of Multimedia Appendix 1) before each round closed. An overview of the changing process across the 3 rounds is provided in Figure 2.

XSL•FO
RenderX

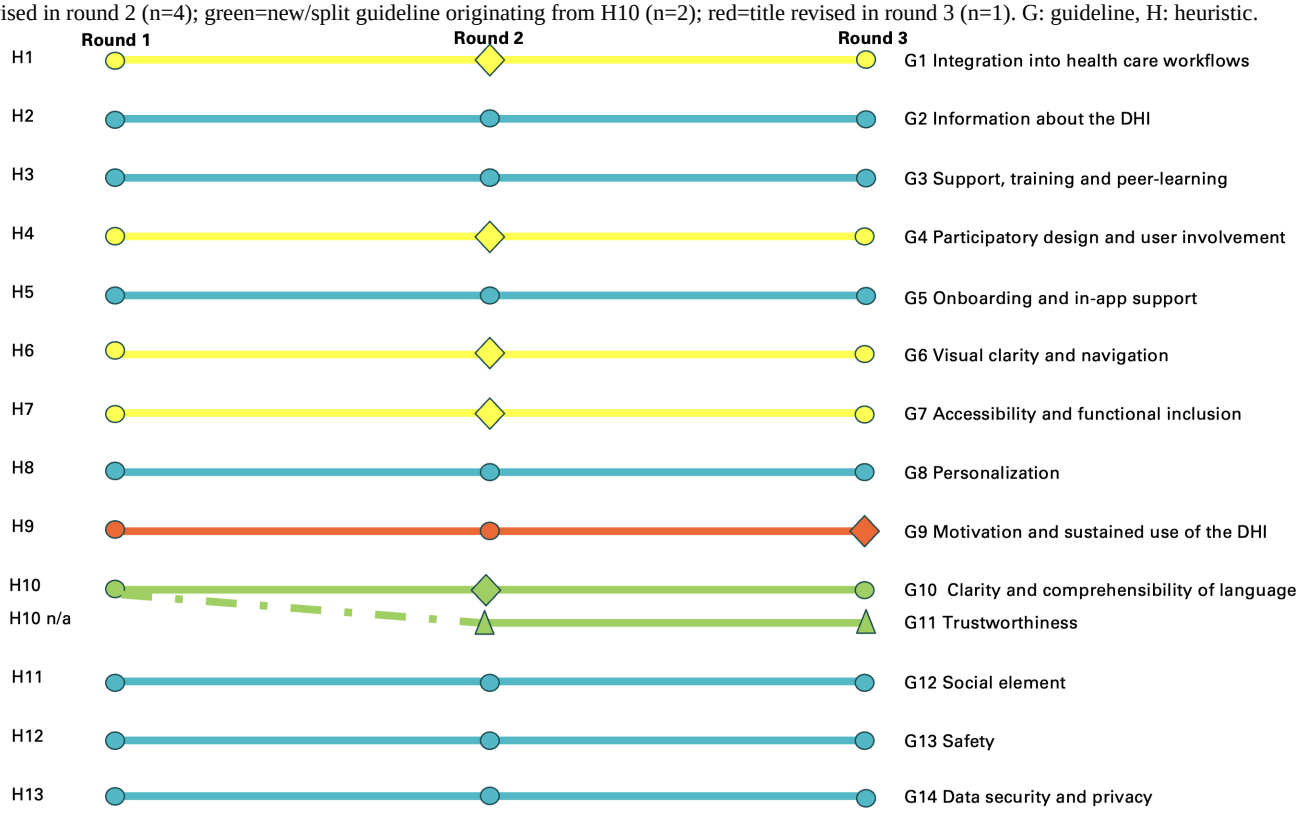
Table . Guidelines in round 3 with brief explanation.

Guideline number	Guideline title	Description
Guidelines addressing the health care system factors		
G1	Integration into health care workflows	Develop a plan for integrating the DHI[a] into health care workflows and identify the added value for each stakeholder involved.
G2	Information about DHI	Provide clear, accessible information about the DHI's purpose, functionality, and limitations for older adults, caregivers, relatives, and health care professionals to promote trustworthy and evidence-based DHIs.
G3	Support, training, and peer-learning	Provide accessible and tailored support, training, and peer learning opportunities that help older adults build confidence in using the DHI.
Guidelines addressing the design of the DHI		
G4	Participatory design and user involvement	Actively involve older adults throughout all stages of the DHI development process to ensure that the solution reflects diverse needs, abilities, and experiences with technology.
G5	Onboarding and in-app support	Provide clear and accessible onboarding, in-app support, and user guides that build confidence and reduce uncertainty.
G6	Visual clarity and navigation	Design the visual interface of the DHI to be clear, consistent, and easy to navigate for older adults.
G7	Accessibility and functional inclusion	Design the DHI to be accessible and usable for older adults with diverse visual, auditory, motor, and cognitive abilities.
G8	Personalization	Design the DHI so it adapts to older adults' abilities, preferences, and goals in a way that supports autonomy and is easy to understand and control.
G9	Motivation and sustained use of the DHI	Design the DHI to include simple and culturally appropriate ways of supporting motivation and sustained use that fit the diverse lifestyles and preferences of older adults.
G10	Clarity and comprehensibility of language	Make sure that all DHI content is easy to understand, respectful, and inclusive for older adults with different levels of health literacy and language proficiency.
G11	Trustworthiness	Build and maintain trust by ensuring reliability, credibility, and transparency in all aspects of the DHI.
G12	Social element	When appropriate for the intervention's goals and context, offer privacy-preserving social features that can help older adults feel connected, supported, and motivated.
G13	Safety	Conduct a systematic risk assessment for older adult DHI users. Implement age-appropriate safety mechanisms with clear escalation pathways to prevent physical and mental harm.
G14	Data security and privacy	Integrate data security and data (processing) transparency into the DHI from the onset.

[a]DHI: digital health intervention.

XSL•FO
RenderX

Figure 2. Evolution of the 14 guideline items across the 3 eDelphi rounds. Each horizontal line represents one guideline (G1–G14). Open circles indicate unchanged titles; filled diamond markers indicate rounds in which the title was revised; the triangle denotes the new G11 (Trustworthiness) guideline created in round 2 by splitting the combined round 1 item H10. Colors track change type: blue=stable title refined each round (n=7); yellow=title revised in round 2 (n=4); green=new/split guideline originating from H10 (n=2); red=title revised in round 3 (n=1). G: guideline, H: heuristic.



The consensus results obtained in round 3 are shown in Table 4. Results from rounds 1 and 2 are provided in Multimedia Appendix 1. This table presents the scores obtained for guideline Gx, considering the 3 evaluated factors: A-utility, B-clarity, and C-ease of use. In the final round, 5 guidelines (G6, G10, G12, G13, and G14) achieved consensus and were fully accepted based on all 3 evaluated factors (utility, clarity, and ease of use).

All the remaining guidelines achieved consensus and were accepted in terms of utility and clarity; however, there was no consensus regarding their ease of use. G5-C was classified as dissent, as one consensus criterion did not meet our threshold (d=0.202>.2). As we followed a strict threshold, this had to be classified as dissent.

XSL•FO
RenderX

Table . Fuzzy scores and agreements of each item obtained in round 3 (N=19)[a].

Items	m1 (SD)	m2 (SD)	m3 (SD)	Fuzzy score	d[b] (SD)	n (%)	Agreement
G1-A	0.795 (0.10)	0.947 (0.05)	1000 (0,0)	0.914	0.064 (0.00)	19 (100.00)	C[c]
G1-B	0.753 (0.09)	0.926 (0.05)	1000 (0.0)	0.893	0.050 (0.03)	19 (100.00)	C
G1-C	0.574 (0.24)	0.758 (0.22)	0.889 (0.17)	0.740	0.172 (0.12)	12 (63.16)	D[d]
G2-A	0.700 (0.27)	0.842 (0.29)	0.911 (0.25)	0.818	0.177 (0.20)	17 (89.47)	C
G2-B	0.684 (0.23)	0.847 (0.22)	0.937 (0.17)	0.823	0.141 (0.15)	17 (89.47)	C
G2-C	0.611 (0.22)	0.789 (0.22)	0.916 (0.17)	0.772	0.148 (0.13)	14 (73.68)	D
G3-A	0.716 (0.25)	0.858 (0.25)	0.926 (0.21)	0.833	0.159 (0.17)	17 (89.47)	C
G3-B	0.700 (0.16)	0.874 (0.13)	0.968 (0.07)	0.847	0.092 (0.09)	18 (94.74)	C
G3-C	0.579 (0.26)	0.753 (0.25)	0.879 (0.20)	0.737	0.188 (0.14)	11 (57.89)	D
G4-A	0.805 (0.12)	0.947 (0.08)	0.995 (0.02)	0.916	0.073 (0.04)	18 (94.74)	C
G4-B	0.784 (0.14)	0.932 (0.09)	0.989 (0.03)	0.902	0.083 (0.05)	17 (89.47)	C
G4-C	0.568 (0.31)	0.737 (0.29)	0.853 (0.23)	0.719	0.232 (0.14)	9 (47.37)	D
G5-A	0.784 (0.14)	0.932 (0.09)	0.989 (0.03)	0.902	0.083 (0.05)	17 (89.47)	C
G5-B	0.784 (0.14)	0.932 (0.09)	0.989 (0.03)	0.902	0.083 (0.05)	17 (89.47)	C
G5-C	0.642 (0.29)	0.800 (0.27)	0.895 (0.21)	0.779	0.202 (0.16)	16 (84.21)	D
G6-A	0.826 (0.12)	0.958 (0.08)	0.995 (0.02)	0.926	0.068 (0.05)	18 (94.74)	C
G6-B	0.774 (0.19)	0.921 (0.16)	0.974 (0.11)	0.889	0.097 (0.12)	18 (94.74)	C
G6-C	0.742 (0.20)	0.900 (0.16)	0.968 (0.12)	0.870	0.100 (0.12)	18 (94.74)	C
G7-A	0.742 (0.22)	0889 (0.18)	0.958 (0.12)	0.863	0.128 (0.12)	18 (94.74)	C
G7-B	0.753 (0.20)	0.900 (0.16)	0.963 (0.10)	0.872	0.117 (0.10)	17 (89.47)	C
G7-C	0.600 (0.33)	0.758 (0.31)	0.853 (0.24)	0.737	0.257 (0.13)	6 (31.58)	D
G8-A	0.711 (0.21)	0.874 (0.17)	0.958 (0.12)	0.847	0.116 (0.12)	18 (94.74)	C
G8-B	0.700 (0.20)	0.868 (0.17)	0.958 (0.12)	0.842	0.112 (0.12)	18 (94.74)	C
G8-C	0.616 (0.26)	0.789 (0.23)	0.900 (0.17)	0.768	0.184 (0.12)	10 (52.63)	D
G9-A	0.711 (0.23)	0.868 (0.19)	0.947 (0.13)	0.842	0.135 (0.13)	17 (89.47)	C
G9-B	0.668 (0.21)	0.842 (0.19)	0.942 (0.13)	0.818	0.132 (0.12)	17 (89.47)	C
G9-C	0.611 (0.27)	0.774 (0.25)	0.889 (0.19)	0.758	0.188 (0.14)	10 (52.63)	D
G10-A	0.753 (0.22)	0.900 (0.19)	0.958 (0.13)	0.870	0.120 (0.13)	17 (89.47)	C
G10-B	0.732 (0.22)	0.884 (0.19)	0.953 (0.13)	0.856	0.128 (0.13)	17 (89.47)	C
G10-C	0.695 (0.23)	0.858 (0.22)	0.942 (0.17)	0.832	0.131 (0.16)	17 (89.47)	C
G11-A	0.805 (0.12)	0.947 (0.08)	0.995 (0.02)	0.916	0.073 (0.04)	18 (94.74)	C
G11-B	0.732 (0.20)	0.889 (0.17)	0.963 (0.12)	0.861	0.110 (0.12)	18 (94.74)	C
G11-C	0.589 (0.32)	0.753 (0.30)	0.858 (0.23)	0.733	0.239 (0.14)	8 (42.11)	D
G12-A	0.742 (0.14)	0.905 (0.10)	0.984 (0.04)	0.877	0.082 (0.06)	19 (100.00)	C
G12-B	0.742 (0.14)	0.905 (0.10)	0.984 (0.04)	0.877	0.082 (0.06)	19 (100.00)	C
G12-C	0.653 (0.23)	0.821 (0.22)	0.926 (0.17)	0.800	0.151 (0.14)	17 (89.47)	C
G13-A	0.753 (0.20)	0.905 (0.16)	0.968 (0.12)	0.875	0.104 (0.12)	18 (94.74)	C
G13-B	0.684 (0.23)	0.847 (0,22)	0.937 (0.17)	0.823	0.141 (0.15)	17 (89.47)	C
G13-C	0.663 (0.27)	0.821 (0.26)	0.911 (0.20)	0.798	0.182 (0.16)	16 (84.21)	C
G14-A	0.795 (0.12)	0.942 (0.08)	0.995 (0.02)	0.911	0.074 (0.04)	18 (94.74)	C
G14-B	0.679 (0.24)	0.847 (0.21)	0.937 (0.16)	0.821	0.143 (0.14)	17 (89.47)	C

XSL•FO
RenderX

Items	m1 (SD)	m2 (SD)	m3 (SD)	Fuzzy score	d[b] (SD)	n (%)	Agreement
G14-C	0.679 (0.25)	0.842 (0.22)	0.932 (0.16)	0.818	0.155 (0.14)	17 (89.47)	C

[a]G (1-14) refers to the number of the guideline (Table 3); G (1-14)-A: utility, G (1-14)-B: clarity, G (1-14)-C: ease of use; m1, m2, and m3 representing the minimum, most plausible, and maximum values.

[b]d: mean threshold value.

[c]C: consent.

[d]D: dissent.

Recommended Practical Tools, Standards, and Frameworks for Inclusive Design

Overview

Experts highlighted a range of established tools, standards, and theoretical frameworks that can support the inclusive design and evaluation of DHIs for older adults. These resources span usability and accessibility standards, human-centered design approaches, behavioral models, and implementation frameworks (Textbox 1).

Textbox 1. Recommended practical tools, standards, and frameworks.

Accessibility standards

- Web Content Accessibility Guidelines, ISO standards (ISO 9241-210, ISO 21801-1:2000)
- Tool: Web Accessibility Evaluation Tool

Design frameworks

- Inclusive, universal design principles, living lab approaches
- Tool: inclusive design toolkit

Health-specific frameworks

- Reach, Effectiveness, Adoption, Implementation, and Maintenance, Xertia mHealth app guidelines, Health Cascade framework, National Institutes of Health Clear Communication principles

Human-centered design

- User- and human-centered design, co-design frameworks

Heuristics and models

- Nielsen heuristics, double diamond model structure design, Garrett 5 Planes of User Experience, design thinking

Accessibility and Usability Standards

Accessibility standards were consistently emphasized as foundational components of inclusive design. In particular, the Web Content Accessibility Guidelines (WCAG 2.1/2.2) [47,48] were frequently recommended, with experts advocating adherence to the highest conformance levels to accommodate age-related sensory, cognitive, and motor impairments. Complementary standards such as ISO 9241‐210 (Human-centered design for interactive systems) [49] and ISO 21801-1:2000, which provides guidance for addressing cognitive and neurodiversity-related accessibility needs [50], were also identified as critical for ensuring usability and accessibility throughout the design lifecycle. Experts further noted the importance of incorporating redundancy in design (eg, multimodal feedback) to support users with diverse and combined impairments.

Practical tools such as the Web Accessibility Evaluation Tools [51] were recommended for assessing accessibility compliance. Experts also pointed to specific guideline sets and systematic reviews (eg, mobile app design guidelines for older adults) as valuable resources for evidence-based design decisions.

Inclusive and Human-Centered Design Frameworks

A range of inclusive design frameworks were recommended to guide the development process. These included the Inclusive Design Toolkit by University of Cambridge [52] and the Microsoft Inclusive Design Toolkit [53], both of which provide practical methods such as personas, exclusion calculators, and capability-based design approaches. The principles of universal design and broader inclusive design principles were also highlighted as essential for addressing variability in users' abilities without stigmatizing older adults.

Human-centered and participatory approaches were strongly endorsed. Experts emphasized user-centered design, human-centered design, and co-design methodologies, including living lab approaches, as critical for ensuring that DHIs reflect the real needs, preferences, and contexts of older adults. The use of persona- and scenario-based design, particularly tailored to diverse older adult populations, was recommended to support empathetic and context-sensitive design decisions.

XSL•FO
RenderX

Heuristics and Design Models

Experts suggested building on established usability heuristics, such as Nielsen heuristics [54], with adaptations to account for age-related changes (eg, enhanced feedback, reduced cognitive load, and support for recognition over recall). In addition, frameworks for persuasive and health technology design were recommended as complementary perspectives.

Design process models such as the double diamond model [55], design thinking [56], and Garrett Five Planes of User Experience [57] were identified as useful for structuring the development of DHIs. Experts also noted the importance of clearly distinguishing between design heuristics (principles) and methodological approaches (processes), as these are often conflated.

Health-Specific and Implementation Frameworks

Several frameworks specific to digital health and implementation were also identified. These included the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework [58], as well as emerging models such as the Health Cascade Framework [59]. Experts referenced domain-specific guidelines such as Xcertia mHealth app guidelines [60] and emphasized the importance of health literacy frameworks, including National Institutes of Health Clear Communication principles [61]. Regulatory and ethical considerations were noted as essential components of inclusive design, with references to the General Data Protection Regulation [62] and the Health Insurance Portability and Accountability Act [63] for data protection and privacy.

Additional Considerations

Experts emphasized that inclusive design should avoid stigmatizing older adults and instead aim to create solutions that are broadly usable across age groups. They also highlighted the importance of aligning DHIs with familiar, mainstream technologies to support acceptance. Finally, the inclusion of practical examples alongside theoretical guidelines was considered essential to enhance usability and real-world applicability for designers and developers.

Discussion

Main Findings

We conducted a 3-round eDelphi study with the objective of collecting expert consensus on a first version of inclusive design guidelines. These guidelines are intended to support designers and developers in creating DHIs for older adults that are both well accepted by older adults and well-integrated into the health care system. A panel of 28 experts participated in round 1, 20 experts in round 2, and 19 experts in round 3 of the Delphi process. Consensus was reached on more than 75% (33/42) of evaluated items related to guidelines' utility, clarity, and ease of use. All guidelines achieved consensus regarding their utility and clarity. However, approximately 65% of guidelines (9/14) did not reach consensus with respect to their ease of use. Participants indicated that the guidelines were not easy to apply in practice due to the lack of examples that illustrate their real-world implementation. To address this, the panel experts enriched the final set of guidelines with examples that show how these guidelines can be implemented during the design and development stages of DHIs for older adults. In addition, existing tools and frameworks can support the implementation of some of the guidelines in practice. We included the recommended tools and frameworks wherever possible in the final version of the guidelines. The final guideline document is available in Checklist 1 and will also be made available on a webpage.

Of note, all the guidelines were considered useful to the DHI design and development processes. Considering the varied backgrounds of the experts and the range of DHIs to which the guidelines will be applied, the complete agreement on usefulness is a significant strength of the guidelines. The experts also regarded all the guidelines to be clear and understandable, which suggests that developers may be more willing to adopt them [64]. In terms of utility, experts rated the following guidelines (in order) as most relevant:

- G6: Visual clarity and navigation
- G4: Active participation
- G11: Trustworthiness
- G1: Integration into health care workflows
- G14: Data security and privacy
- G5: Onboarding and in-app support

The emphasis on navigation and integration into existing workflows is not surprising because ease of use is highly relevant to older adults who use DHIs [65]. These factors become even more important as individuals age and cognitive impairment becomes more likely and perhaps problematic in the completion of activities of daily living. Privacy and data security consistently rank among older adults' concerns related to DHIs [66,67], and failure to account for these concerns in development is likely to result in reduced acceptance and adoption. Providing personalized support at the point of use and with the opportunity for ongoing assistance when needed can help to allay deterrents to DHI use and promote regular DHI use [68].

Although the experts agreed on the usefulness and clarity of the guidelines, they did not necessarily consider all guidelines easy to apply in practice. This discrepancy suggests that a guideline can be perceived as relevant and clearly formulated while still being difficult to operationalize during the design and development of DHIs. One possible barrier is that some guidelines depend on factors that are not fully under the control of individual designers or developers, such as limited access to older adults for participatory activities, the integration of DHIs into the traditional health care workflows, limited resources of the company, institutional data governance requirements, or the availability of staff to provide training and support to older adults. The practical implementation of some guidelines may therefore require additional resources, organizational support, and improved collaboration across different professional groups within an organization. The guidelines that were considered most difficult to use, sorted by perceived difficulty, included:

- G4: Active participation
- G11: Trustworthiness
- G7: Accessibility and functional inclusion

XSL•FO
RenderX

- G3: Support, training, and peer-learning
- G1: Integration into health care workflows

Active participation was ranked second for utility and first for clarity, but it is also considered the most difficult guideline to use. Including older adult users in the development of DHIs is not always possible due to lack of proximity to DHI development facilities, users' cognitive impairment, lack of motivation to be involved, and other reasons [69]. Co-design initiatives may also involve costs (eg, transportation, personal care attendant, and special meals) that are prohibitive in some settings. When co-design is infeasible, use of a participatory design framework such as those developed by Shah et al [70] and Denecke et al [71] could support active participation.

Trustworthiness has been identified as a concern among users of health care technologies. Its abstract nature compared to other guidelines (eg, those for accessibility for users with visual or hearing limitations and DHI training) can make trustworthiness difficult to illustrate. Use of other guidelines to promote ease of use and rapid success may encourage users to trust DHIs even when efforts to demonstrate trustworthiness are not readily apparent. In addition, the trustworthiness guideline could be complex to apply because there is a lack of design frameworks that elucidate how to evaluate this subjective concept.

Support and peer-learning opportunities may be limited or infeasible in some settings, such as skilled nursing facilities, and care team members may reasonably prioritize providing health care services over DHI support. In institutional care settings, such limitations are closely linked to data governance and privacy constraints. Greater institutional control over workflows and integration with medical devices could be implemented through robust privacy-preserving frameworks, and it may enable care teams to better align DHIs with existing practices. Such developments could potentially enhance not only the feasibility of guideline implementation but also the overall quality of care and continuity of support for older adults. However, achieving this would require careful consideration of ethical, legal, and organizational responsibilities. Similarly, care team members may have no control over workflows in use and may be unable to coach older adults in using DHIs along with medical devices. Frameworks such as the normalization process theory may provide guidance in some situations but are unlikely to be a good fit for every use case [72].

Positioning the Guidelines Within Broader Age-Friendly and Inclusive Design Research

Our inclusive design guidelines can also be positioned within the broader literature on digital literacy and inclusive design for older adults. For instance, Shi et al identified that sociodemographic characteristics, device-related factors, and social support are important influences on older adults' digital health literacy [73]. These findings support our broader understanding of inclusive design for older adults, according to which older adults' engagement with DHIs is shaped not only by the usability and accessibility of the user interface but also by their individual capabilities, access to appropriate devices, and social factors (eg, cultural context, social support, peer learning, and training opportunities). Similarly, Dong et al found that interventions aimed at improving digital health literacy can positively affect older adults' health status and health management [74]. These findings indicate that inclusive DHIs should not only reduce the interaction barriers but should also support users in accessing, understanding, and applying information that is distributed by the DHIs. In accordance, our guidelines addressed user support, training, and education, and they highlighted the importance of considering different levels of digital literacy among older adults when creating the DHIs. This might contribute to older adults' meaningful and sustained engagement with DHIs. The relevance of a multidimensional approach is also reflected in inclusive design research beyond the area of digital health. For example, Shi et al [75] developed a multidimensional index for inclusive age-friendly transit design that translates broad age-friendly principles into assessable dimensions and practical scenarios. Although it was developed for public transportation context, this approach illustrates how inclusive design should move beyond general principles toward a structured framework that supports practical implementation. Our guidelines pursue a similar objective in the context of the design and development of DHIs. This was achieved by consolidating user-related, interface-level, technical, organizational, ethical, and contextual factors into guidelines with examples for practical implementation for the design and development of DHIs for older adults. In this way, our work contributes to the translation of inclusive design principles into concrete considerations that can guide researchers, designers, developers, and further health care stakeholders in the creation of DHIs for older adults.

Planned Dissemination of the Guidelines

To support their uptake by academia and industry, the final guidelines will be made openly available through this publication and a dedicated website (milestone 1). The primary target audiences include researchers, designers, and developers of DHIs, health care professionals involved in the implementation of DHI, and organizations working in digital health, human-computer interaction, gerontechnology, and inclusive design. Dissemination will be coordinated by the study team and will initially take place following publication through scientific conferences, academic and professional networks, and relevant professional communities. Where possible, organizations and networks represented by members of the expert panel will be approached to support dissemination within their respective communities.

As an initial real-world application, the guidelines will inform the ongoing design and development of a DHI aimed at supporting the mental well-being of older adults transitioning from working life to retirement (milestone 2). This application will provide the first opportunity to examine the practical usability of the guidelines in a real-world design process, including whether and how the practical examples support their interpretation and ease of use. It will also enable the identification of challenges arising during their implementation. Feedback and experiences from this application will be documented and considered in future revisions of the guidelines.

The reach and uptake of the guidelines will be monitored, where feasible, using indicators such as website access or downloads, citations of the publication, and documented use of the

XSL•FO
RenderX

guidelines in research or development projects. In addition, feedback submitted through the dedicated forum will provide a mechanism for users to report experiences, suggest modifications, or identify emerging evidence relevant to individual guideline items. The guideline set is intended to evolve as new evidence and practical experience emerge. The study team will remain responsible for maintaining the online version and reviewing relevant feedback and new evidence at defined intervals (eg, two years and follow-up milestones). Revisions will consider emerging research, feedback from users, and findings from applications of the guidelines in practice. Substantive updates will undergo review by relevant experts and stakeholders before being incorporated into a revised version, and changes between versions will be documented transparently. Future evaluations should particularly involve older adults, health care professionals, designers and developers, and stakeholders from geographical and cultural contexts that were underrepresented in this study.

Limitations

We acknowledge that the presented eDelphi study has several limitations. First, the geographical representation of the expert panel was limited. All experts were based in Europe or North America, while no experts from Africa, Asia, South America, or Australia and Oceania participated. This imbalance may partly reflect the recruitment strategy we used, which relied primarily on conferences, publications, organizations, and professional networks that are based in Europe and North America. The low response rate could be explained by the recruitment strategy: collecting email addresses from papers may lead to outdated or inactive contact information and may also result in unsolicited invitations being overlooked or filtered as spam. However, we are lacking information on why people were not responding to the request to join the Delphi panel. The survey was conducted in English, which may have introduced language bias by limiting participation to experts with sufficient English proficiency and potentially affecting participants' interpretation of the guideline items and survey questions. The disciplinary and sectoral composition of the panel may also have influenced the findings, as expertise was more strongly represented in digital health, UX/usability, and human-computer interaction than in areas such as geriatrics and health communication, potentially giving greater weight to technology- and design-oriented perspectives. Persons with lived experiences remained unconsidered. Some experts were recruited through the authors' professional networks. Although these experts underwent the same recruitment and study procedures as other panel members, recruitment through existing professional relationships may have introduced selection bias, for example, by increasing the likelihood of including experts with professional perspectives or research interests similar to those of the study team.

Consequently, the consensus that was reached in this eDelphi study should be interpreted primarily within the European and North American contexts represented by the panel. The findings may not fully reflect the differences in healthcare systems, digital infrastructure, socioeconomic conditions, cultural expectations regarding aging and care, language, and levels of digital and health literacy across other world regions. Therefore, the guidelines cannot yet be considered globally validated. Future studies should assess and adapt the guidelines with experts, older adults, health care professionals, designers, and other stakeholders from underrepresented regions and evaluate their applicability across diverse cultural, socioeconomic, and health care contexts.

Furthermore, this study collected only expert consensus and lacks external validation of the guidelines. We did not include an assessment by older adults nor conduct testing of the guidelines in real-world design and development settings. Although older adults' perspectives informed the development of the initial set of guidelines through a focus group, their involvement was limited to this earlier phase. Therefore, future studies should evaluate the guidelines with older adults, caregivers, healthcare professionals, and DHI developers and examine their practical applicability during the design, development, and evaluation of DHIs.

In addition, even after 3 eDelphi rounds, not all guidelines achieved full consensus, particularly regarding their perceived ease of use. In response to this concern, experts from the panel were further involved in refining the guidelines and developing practical examples intended to facilitate their use.

However, although the examples underwent review and revision by participating experts, this process was not formally standardized and relied on their professional experience and judgment. Moreover, the participating experts were also coauthors of this study, and the examples may therefore reflect a limited range of perspectives. Although the examples were developed with the intention of improving ease of use, we did not formally assess whether they achieved this aim. Consequently, the extent to which the examples improve the ease of use of the guidelines remains uncertain. Future research should independently evaluate the guidelines and examples with relevant stakeholder groups, including designers, developers, and health professionals.

Finally, the nature of the eDelphi methodology relies on online data acquisition and limits the direct interaction among participants. The absence of discussion may have restricted opportunities to clarify different perspectives or resolve disagreements in greater depth. While this approach reduces the risk of dominant voices influencing the process, it may also limit the richness of consensus compared to more interactive methods.

Conclusions and Outlook

In this eDelphi study, we assessed expert consensus on a set of inclusive design guidelines for DHIs intended for older adults. The framework extends beyond user interface considerations to address health care, organizational, and contextual factors relevant to DHI design. Consensus was achieved for most items, with all guidelines rated as useful and clear. This consistent agreement across experts with diverse backgrounds underscores the framework's relevance, comprehensibility, and potential for adoption in practice.

Guidelines related to visual clarity and navigation, active participation, trustworthiness, workflow integration, data security, and onboarding were rated as most relevant, reinforcing

XSL•FO
RenderX

the importance of usability, safety, and contextual fit in DHIs for older adults. These priorities align with the known needs of this population, including minimizing cognitive burden, addressing privacy concerns, and ensuring seamless integration into daily routines and care processes.

However, several guidelines did not reach consensus on ease of use, highlighting a gap between perceived value and practical implementation. In particular, active participation and trustworthiness, despite being highly valued, were seen as difficult to operationalize due to barriers such as limited access to users, resource constraints, and the abstract nature of certain ideas. Challenges related to workflow integration and support provision further reflect structural limitations within care settings. These findings indicate that additional methodological and practical guidance is needed, including participatory design approaches, clearer operationalization of complex concepts, and context-sensitive implementation strategies.

Overall, the guidelines bring together interface-level considerations with broader factors such as accessibility, interoperability, sustained use, safety, trustworthiness, and integration into health care processes, providing an expert-informed framework for DHI design and development. Their practical usability, applicability across settings, and contribution to the development of inclusive and accessible DHIs remain to be evaluated empirically. Further studies should evaluate the transferability of the guidelines across diverse health care systems, geographical or cultural contexts, as well as types of DHIs.

Acknowledgments

We acknowledge the participation of the experts who joined the eDelphi panel.

Generative AI Use

Generative AI (ChatGPT OpenAI, models GPT-4o and GPT-5) was used to assist with language editing of the manuscript draft. No AI tool was used to generate original scientific content. AI was also not used to design the study, formulate research questions, extract data, or perform statistical analysis, and AI did not independently determine the interpretation, conclusion, or scientific claims of the study. The authors take full responsibility for the accuracy, integrity, and final content of the manuscript.

Funding

This project was partially funded by the Bern University of Applied Sciences within the funding scheme of Human Digital Transformation. The funding covered parts of the working hour costs of coauthor KD, LC and BK. The funder had no involvement in this study’s design, data collection, analysis, interpretation, or the writing of this paper.

Data Availability

All available data are provided as Multimedia Appendices 1 and 2 and Checklist 1.

Authors' Contributions

Conceptualization: KD, ORR
Methodology: KD, ORR, LC, BK
Data analysis: ORR, KD
Writing (original draft): KD, ORR, LC, CP, EZ
Writing (review & editing): BK, ORR, KD, CP, EZ, YA, LC, JSA, MB-H, RD, JG-D, LH, MO, VS, ES, GV, TV, EDZ
Participation in eDelphi panel: EZ, YA, JSA, MB-H, RD, JG-D, LH, MO, VS, ES, GV, TV, EDZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Inclusive design guidelines and questionnaire modification log, including fuzzy scores and agreements of each item obtained in round 1, round 2, and round 3.
[XLSX File, 48 KB - aging_v9i1e99657_app1.xlsx]

Multimedia Appendix 2
Inclusive design guidelines.
[PDF File, 1201 KB - aging_v9i1e99657_app2.pdf]

Checklist 1
Filled Conducting and Reporting Delphi Studies reporting guidelines.
[DOCX File, 2037 KB - aging_v9i1e99657_app3.docx]

XSL•FO
RenderX

References

1. Active ageing: a policy framework. World Health Organization. URL: https://iris.who.int/items/46476232-b305-47b4-ab28-e3931607c70c [accessed 2026-04-13]
2. Ismail Z, Ahmad WIW, Hamjah SH, Astina IK. The impact of population ageing: a review. Iran J Public Health 2021 Dec;50(12):2451-2460. [doi: 10.18502/ijph.v50i12.7927] [Medline: 36317043]
3. Noto S. Perspectives on aging and quality of life. Health Care (Don Mills) 2023 Jul 26;11(15):2131. [doi: 10.3390/healthcare11152131]
4. Halaweh H, Dahlin-Ivanoff S, Svantesson U, Willén C. Perspectives of older adults on aging well: a focus group study. J Aging Res 2018;2018:9858252. [doi: 10.1155/2018/9858252] [Medline: 30533224]
5. Bland JS. Age as a modifiable risk factor for chronic disease. Integr Med (Encinitas) 2018 Aug;17(4):16-19. [Medline: 31043904]
6. Rudnicka E, Napierała P, Podfigurna A, Męczekalski B, Smolarczyk R, Grymowicz M. The World Health Organization (WHO) approach to healthy ageing. Maturitas 2020 Sep;139:6-11. [doi: 10.1016/j.maturitas.2020.05.018] [Medline: 32747042]
7. Rooper IR, Lewis-Thames MW, Graham AK. Beyond usability: designing digital health interventions for implementation with older adults. J Am Geriatr Soc 2025 May;73(5):1648-1652. [doi: 10.1111/jgs.19286] [Medline: 39607322]
8. Baik D, Allen LA, Coats H, Matlock DD, Jankowski C. Engagement in a digital health physical activity program, TPA4You, among older family care partners of persons with heart failure: a scenario-based design. Geriatr Nurs 2025;66(Pt C):103677. [doi: 10.1016/j.gerinurse.2025.103677] [Medline: 41129906]
9. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023;11:1281923. [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
10. Fothergill L, Latham Y, Hayes N, Holland C. Using a proactive telecare system to support independence, health, and well-being in older adults: feasibility randomized controlled trial. JMIR Form Res 2025 Dec 31;9:e82152. [doi: 10.2196/82152] [Medline: 41474602]
11. Classification of digital health interventions v1.0. : World Health Organization; 2018 URL: https://www.who.int/publications/i/item/WHO-RHR-18.06 [accessed 2026-10-27]
12. Alruwaili MM, Shaban M, Elsayed Ramadan OM. Digital health interventions for promoting healthy aging: a systematic review of adoption patterns, efficacy, and user experience. Sustainability 2023 Dec 2;15(23):16503. [doi: 10.3390/su152316503]
13. Ferraz LT, Frohlich DM, Hodgkins CE, Yuan H, Costa Castro P. Optimising the provision of health information for older adults across paper and screen formats - a requirement study with content producers and consumers. PLOS Digit Health 2025 Nov;4(11):e0001090. [doi: 10.1371/journal.pdig.0001090] [Medline: 41248099]
14. Liu T, Luo YT, Pang PCI, Wong D, Wang Y. Latent obstacles in older adults' digital health participation: a community-based hybrid cluster analysis with natural language processing. BMC Public Health 2025 Dec 18;26(1). [doi: 10.1186/s12889-025-25847-7]
15. Zhou Y, Rees J, Matcham F, Antonelli M, Ourselin S, Liu W. A user-centered interface design framework for the DELONELINESS system in older adults: design indicator development and prioritization. JMIR Hum Factors 2026 Mar 6;13:e88263. [doi: 10.2196/88263] [Medline: 41813191]
16. Berry EC, Sculthorpe NF, Warner A, Sanal-Hayes NE, Hayes LD. Designing an mHealth app to encourage uptake of muscle-strengthening exercise in older adults: co-design focus group study. JMIR Aging 2026 Mar 12;9:e87332. [doi: 10.2196/87332] [Medline: 41818741]
17. Persson H, Åhman H, Yngling AA, Gulliksen J. Universal design, inclusive design, accessible design, design for all: different concepts—one goal? On the concept of accessibility—historical, methodological and philosophical aspects. Univ Access Inf Soc 2015 Nov;14(4):505-526. [doi: 10.1007/s10209-014-0358-z]
18. Boot W, Charness N, Czaja SJ, Rogers WA. Designing for Older Adults: Case Studies, Methods, and Tools: CRC Press; 2020. [doi: 10.1201/b22187]
19. Eklund C, Zander V, Gusdal AK, Åkerlind C, von Heideken Wågert P. Co-designing a digital solution for decreasing loneliness and social isolation among older people in Sweden: explorative study. JMIR Form Res 2025 Nov 21;9:e78213. [doi: 10.2196/78213] [Medline: 41270279]
20. Elguera Paez L, Zapata Del Río C. Elderly users and their main challenges - usability with mobile applications: a systematic review. In: Marcus A, Wang W, editors. Design, User Experience, and Usability Design Philosophy and Theory: 8th International Conference, DUXU 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part I: Springer International Publishing; 2019:423-438. [doi: 10.1007/978-3-030-23570-3_31]
21. Liu N, Yin J, Tan SSL, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. J Am Med Inform Assoc 2021 Oct 12;28(11):2483-2501. [doi: 10.1093/jamia/ocab151] [Medline: 34472601]
22. Amouzadeh E, Dianat I, Faradmal J, Babamiri M. Optimizing mobile app design for older adults: systematic review of age-friendly design. Aging Clin Exp Res 2025 Aug 14;37(1):248. [doi: 10.1007/s40520-025-03157-7] [Medline: 40804492]

XSL•FO
RenderX

23. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023 Sep 21;11:e43186. [doi: 10.2196/43186] [Medline: 37733401]
24. Stephens M. Designing for older audiences: Checklist + best practices.: Medium; 2025. URL: https://uxdesign.cc/designing-for-older-audiences-checklist-best-practices-b6ca3ec5bcbf [accessed 2026-04-13]
25. Brewer R. Understanding voice‐based information uncertainty: a case study of health information seeking with voice assistants. J Assoc Inf Sci Technol 2024 Oct;75(10):1041-1057. [doi: 10.1002/asi.24854]
26. Czaja SJ, Boot WR, Charness N, Rogers WA. Designing for Older Adults: Principles and Creative Human Sactors Approaches, 3rd edition: CRC Press; 2019. [doi: 10.1201/b22189]
27. Stephanidis C, Salvendy G, editors. Designing for Usability, Inclusion and Sustainability in Human-Computer Interaction: CRC Press; 2024. [doi: 10.1201/9781003495147]
28. Abhari S, McMurray J, Randhawa T, et al. Exploring the landscape of standards and guidelines in AgeTech design and development: scoping review and thematic analysis. JMIR Aging 2024 Oct 31;7:e58196. [doi: 10.2196/58196] [Medline: 39481099]
29. Denecke K, Cvijic L, Petersen C. Toward inclusive design heuristics for digital health interventions for the aging population: scoping review. J Med Internet Res 2025 Dec 22;27:e79449. [doi: 10.2196/79449] [Medline: 41428384]
30. Kaufmann B, Cvijic L, Denecke K. Ensuring inclusive design in digital health interventions for the aging population: expert perspectives. Presented at: Congresso Internacional de Design da Informação; Sep 15-19, 2025. [doi: 10.5151/cidi2025-1085923]
31. Roldán López de Hierro AF, Sánchez M, Puente-Fernández D, Montoya-Juárez R, Roldán C. A fuzzy Delphi consensus methodology based on a fuzzy ranking. Mathematics 2021 Sep 19;9(18):2323. [doi: 10.3390/math9182323]
32. Ahmad NA, Abu Kasim NH, Musa S, Othman SA, Naimie Z. Defining the multifaceted roles of dental academics: a consensus approach using the fuzzy Delphi method. PLoS One 2025;20(8):e0330636. [doi: 10.1371/journal.pone.0330636] [Medline: 40845082]
33. Ahmad Basri MAF, Wan Ismail WS, Kamal Nor N, et al. Validation of key components in designing a social skills training content using virtual reality for high functioning autism youth-a fuzzy Delphi method. PLoS One 2024;19(4):e0301517. [doi: 10.1371/journal.pone.0301517] [Medline: 38574084]
34. Abdul Hamid SH, Ellin MR, Mohd Arifin SR, et al. Developing a culturally sensitive framework for sexual and reproductive health education: validating with the fuzzy Delphi method. J Transcult Nurs 2026 Mar;37(2):189-197. [doi: 10.1177/10436596251394358] [Medline: 41355738]
35. Suraya Mohd Dan A, Linoby A, Shahlan Kasim S, et al. Validation of a personalized AI prompt generator (NExGEN-ChatGPT) for obesity management using fuzzy Delphi method. Biol Methods Protoc 2025;10(1):bpaf085. [doi: 10.1093/biomethods/bpaf085] [Medline: 41323029]
36. Jamil MRM, Said S, Azeez MIK. Kompetensi guru terhadap pengurusan pengajaran dan pembelajaran: suatu pendekatan teknik fuzzy Delphi [Teacher Competency in the Management of Teaching and Learning: A Fuzzy Delphi Technique Approach]. J Kep Pendidik 2014;1(3):77-83.
37. Gengatharan K, Beram S, Noor N, Yuet FKC. The fuzzy Delphi techniques in the design and development of health education assessment module constructs for the implementation of classroom based assessment. IJARBSS 2023 Mar 19;13(3):1485-1497. [doi: 10.6007/IJARBSS/v13-i3/16769]
38. Quiñones D, Rusu C, Rusu V. A methodology to develop usability/user experience heuristics. Comput Stand Interfaces 2018 Aug;59:109-129. [doi: 10.1016/j.csi.2018.03.002]
39. Che Jamaludin FI, Abdullah M, Endut M, Saifuddeen SM, Hamimi KA, Harun S. Developing a Shariah-compliant medical services framework in Malaysia: an expert system approach using fuzzy Delphi method and interpretive structural modelling. BMJ Open 2025 Jan 6;15(1):e082263. [doi: 10.1136/bmjopen-2023-082263] [Medline: 39762112]
40. Green B, Jones M, Hughes D, Williams A. Applying the Delphi technique in a study of GPs' information requirements. Health Soc Care Community 1999 May;7(3):198-205. [doi: 10.1046/j.1365-2524.1999.00176.x] [Medline: 11560634]
41. Procter S, Hunt M. Using the Delphi survey technique to develop a professional definition of nursing for analysing nursing workload. J Adv Nurs 1994 May;19(5):1003-1014. [doi: 10.1111/j.1365-2648.1994.tb01180.x] [Medline: 8056906]
42. Beech B. Studying the future: a Delphi survey of how multi-disciplinary clinical staff view the likely development of two community mental health centres over the course of the next two years. J Adv Nurs 1997 Feb;25(2):331-338. [doi: 10.1046/j.1365-2648.1997.1997025331.x] [Medline: 9044008]
43. Cheng CH, Lin Y. Evaluating the best main battle tank using fuzzy decision theory with linguistic criteria evaluation. Eur J Oper Res 2002 Oct;142(1):174-186. [doi: 10.1016/S0377-2217(01)00280-6]
44. Ishikawa A, Amagasa M, Shiga T, Tomizawa G, Tatsuta R, Mieno H. The max-min Delphi method and fuzzy Delphi method via fuzzy integration. Fuzzy Sets Syst 1993 May;55(3):241-253. [doi: 10.1016/0165-0114(93)90251-C]
45. Chang PT, Huang LC, Lin HJ. The fuzzy Delphi method via fuzzy statistics and membership function fitting and an application to the human resources. Fuzzy Sets Syst 2000 Jun;112(3):511-520. [doi: 10.1016/S0165-0114(98)00067-0]
46. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and Reporting Delphi Studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat Med 2017 Sep;31(8):684-706. [doi: 10.1177/0269216317690685] [Medline: 28190381]

XSL•FO
RenderX

47. Web Content Accessibility Guidelines (WCAG) 2.1. World Wide Web Consortium. URL: https://www.w3.org/TR/WCAG21/ [accessed 2026-04-13]
48. Web Content Accessibility Guidelines (WCAG) 2.2. World Wide Web Consortium. URL: https://www.w3.org/TR/WCAG22/ [accessed 2026-04-13]
49. ISO 9241-210:2019: Human-centred design for interactive systems. International Organization for Standardization. URL: https://www.iso.org/standard/77520.html [accessed 2026-04-13]
50. ISO 21801-1:2020: cognitive accessibility. International Organization for Standardization. URL: https://www.iso.org/standard/71711.html [accessed 2026-04-13]
51. WAVE Web Accessibility Evaluation Tools. URL: https://wave.webaim.org/ [accessed 2026-04-13]
52. Inclusive design toolkit. University of Cambridge. URL: https://www.inclusivedesigntoolkit.com/ [accessed 2026-04-13]
53. Microsoft inclusive design. Microsoft. URL: https://inclusive.microsoft.design/ [accessed 2026-04-13]
54. 10 usability heuristics for user interface design. Nielsen Norman Group. URL: https://www.nngroup.com/articles/ten-usability-heuristics/ [accessed 2026-04-13]
55. Kochanowska M, Gagliardi WR. The double diamond model: in pursuit of simplicity and flexibility. In: Perspectives on Design II 2022:19-32. [doi: 10.1007/978-3-030-79879-6_2]
56. Martínez Casanovas M. Exploring design thinking methodologies: a comprehensive analysis of the literature, outstanding practices, and their linkage to sustainable development goals. Sustainability 2025 Aug 6;17(15):7142. [doi: 10.3390/su17157142]
57. Garrett JJ. Meet the elements. In: The Elements of User Experience: User-Centered Design for the Web: New Riders Publishing; 2002. URL: http://www.jjg.net/elements/pdf/elements_ch02.pdf [accessed 2026-10-27] [doi: 10.1145/889692.889709]
58. Standard survey items for RE-AIM. RE-AIM. URL: https://re-aim.org/standard-survey-items-for-re-aim/ [accessed 2026-04-13]
59. Health Cascade. URL: https://healthcascade.eu/ [accessed 2026-04-13]
60. Abts NA, McNicol SA, Branaghan RJ. Guidelines for mobile health applications. Proc Int Symp Hum Factors Ergon Health Care 2019 Sep;8(1):197-200. [doi: 10.1177/2327857919081050]
61. Clear communication. National Institutes of Health (NIH). 2025. URL: https://www.nih.gov/institutes-nih/nih-office-director/office-communications-public-liaison/clear-communication [accessed 2026-04-24]
62. European Parliament and the Council of the European Union. General Data Protection Regulation. URL: https://gdpr-info.eu/ [accessed 2026-04-13]
63. The HIPAA privacy rule. US Department of Health and Human Services. URL: https://www.hhs.gov/hipaa/for-professionals/privacy/index.html [accessed 2026-04-13]
64. Kremer G, Peters I, Bingoel B, Stark R. Better design through shared knowledge via design heuristics. Procedia CIRP 2023;119:957-962. [doi: 10.1016/j.procir.2023.03.140]
65. Rießenberger KA, Povse K, Fischer F. Participatory development of digital innovations for health promotion among older adults: qualitative insights on individual, contextual, and technical factors. Int J Environ Res Public Health 2025 Aug 21;22(8):1311. [doi: 10.3390/ijerph22081311] [Medline: 40869896]
66. Akar C, Koçak O, Solmaz U, Batur Başar T. Exploring older adults' perceptions of online security and privacy: a phenomenological study. J Gerontol Soc Work 2026;69(6):890-914. [doi: 10.1080/01634372.2025.2562403] [Medline: 40974221]
67. Gupta C, Hassan M, Adedayo OM. Mobile privacy and security for older adults: a systematic review of concerns, risks, usability, and pathways for support. J Appl Gerontol 2026 Jan 12:7334648261416184. [doi: 10.1177/07334648261416184] [Medline: 41524149]
68. Choi HS, Song HJ. Key determinants and usage patterns of mobile device proficiency in older adults. Geriatr Nurs 2025;66(Pt C):103645. [doi: 10.1016/j.gerinurse.2025.103645] [Medline: 41014742]
69. Blouin Gagnon L, Turcotte S, DeBroux-Leduc R, Gauthier P, Bier N. Methods for co-designing health innovations with older adults: a rapid review. Australas J Ageing 2025 Sep;44(3):e70065. [doi: 10.1111/ajag.70065] [Medline: 40665831]
70. Shah SGS, Robinson I, AlShawi S. Developing medical device technologies from users' perspectives: a theoretical framework for involving users in the development process. Int J Technol Assess Health Care 2009 Oct;25(4):514-521. [doi: 10.1017/S0266462309990328] [Medline: 19845981]
71. Denecke K, Rivera-Romero O, Giunti G, Van Holten K, Gabarron E. Key components of participatory design workshops for digital health solutions: nominal group technique and feasibility study. J Healthc Inform Res 2025 Sep;9(3):359-379. [doi: 10.1007/s41666-025-00199-4] [Medline: 40726748]
72. Murray E, Treweek S, Pope C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. BMC Med 2010 Oct 20;8(1):63. [doi: 10.1186/1741-7015-8-63] [Medline: 20961442]
73. Shi Z, Du X, Li J, Hou R, Sun J, Marohabutr T. Factors influencing digital health literacy among older adults: a scoping review. Front Public Health 2024;12:1447747. [doi: 10.3389/fpubh.2024.1447747] [Medline: 39555039]
74. Dong Q, Liu T, Liu R, Yang H, Liu C. Effectiveness of digital health literacy interventions in older adults: single-arm meta-analysis. J Med Internet Res 2023 Jun 28;25(1):e48166. [doi: 10.2196/48166] [Medline: 37379077]

XSL•FO
RenderX

75. Shi W, Cai G, Jiang L, Liu Y, Cui YT. A multidimensional index for inclusive age-friendly transit design: development and practical scenarios. Sage Open 2025 Oct;15(4). [doi: 10.1177/21582440251396642]

Abbreviations

AMIA: American Medical Informatics Association
CREDES: Conducting and Reporting Delphi Studies
DHI: digital health intervention
FDM: fuzzy Delphi method
mHealth: mobile health
UX: user experience
WCAG: Web Content Accessibility Guidelines

Edited by PH Chau; submitted 28.Apr.2026; peer-reviewed by SS Funghetto, W Shi; revised version received 23.Sep.2026; accepted 30.Sep.2026; published 12.Nov.2026.

Please cite as:
Denecke K, Rivera-Romero O, Cvijic L, Kaufmann B, Zitkus E, Afacan Y, Bruderer-Hofstetter M, Debeuf R, Goodman-Deane J, Haase L, Ozkaynak M, Santschi V, Seminerio E, Vardali G, Vuohijoki T, Dorronzoro Zubiete E, Petersen C
Inclusive Design Guidelines for Digital Health Interventions for Older Adults: Fuzzy eDelphi Study
JMIR Aging 2026;9:e99657
URL: https://aging.jmir.org/2026/1/e99657
doi:10.2196/99657

© Kerstin Denecke, Octavio Rivera-Romero, Lana Cvijic, Beatrice Kaufmann, Emilene Zitkus, Yasemin Afacan, Marina Bruderer-Hofstetter, Ruben Debeuf, Joy Goodman-Deane, Laura Haase, Mustafa Ozkaynak, Valérie Santschi, Emanuele Seminerio, Georgia Vardali, Tiina Vuohijoki, Enrique Dorronzoro Zubiete, Carolyn Petersen. Originally published in JMIR Aging (https://aging.jmir.org), 12.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Telemonitoring Adoption Among Older Adults: Sequential Mixed Methods Study of Message Framing and Structural Influence

Dandan Wang[1]; Huanyu Bao[1], PhD; Amber Pei Tze Qua[2]; Gaelmaine Yu Heng Hoong[2]; Pei Shan He[2]; Ern Sze Maegan Ong[2]; Edmund Wei Jian Lee[1,3], PhD

[1]Department of Media and Communication, City University of Hong Kong, 5F, Run Run Shaw Creative Media Centre, Hong Kong, China (Hong Kong)

[2]Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore

[3]Institute of Digital Medicine, City University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Huanyu Bao, PhD
Department of Media and Communication, City University of Hong Kong, 5F, Run Run Shaw Creative Media Centre, Hong Kong, China (Hong Kong)

Abstract

Background: Telemonitoring offers solutions for remote health management among older adults; yet, acceptance varies based on individual perceptions and presentation strategies.

Objective: This 2-phased sequential mixed methods study examines the role of message framing on telemonitoring adoption among low-income older adults.

Methods: Phase 1 involved in-depth interviews with 17 older adults, identifying key facilitators (accessibility, affordability, and health benefits) and barriers (preference for in-person consultations, technological limitations, and information processing difficulties) of telemonitoring. Building on these insights, phase 2 used an experimental study with 120 participants assessing message framing effects (episodic vs thematic) and modalities (video vs photo-text) on telemonitoring adoption.

Results: Episodic framing did not produce a significant main effect on attitudes toward telemonitoring; however, its effect was moderated by external locus of control (LOC), with episodic framing associated with more positive attitudes among participants lower in external LOC, and thematic framing more favorable among those higher in external LOC.

Conclusions: Convenience, cost, and health benefits emerged as key adoption drivers, while traditional consultation preferences and cognitive processing challenges posed barriers. Theoretical and practical implications were discussed.

(JMIR Aging 2026;9:e89817) doi:10.2196/89817

KEYWORDS

telemonitoring; older adults; health communication; message framing; digital inequality; aging in place

Introduction

Background

Telemonitoring uses electronic technology to remotely track patients' vital signs and health parameters [1]. By enabling continuous monitoring, telemonitoring facilitates early detection of health changes and supports "aging in place," enabling older adults to live in their communities safely and independently [2,3]. However, telemonitoring services are not equitably accessible due to communication inequalities (disparities in access to and use of ITs), particularly among socioeconomically vulnerable groups [4]. Older adults face greater obstacles in receiving virtual care due to limited use of digital health technologies, insufficient technological resources, and lack of motivation [5]. Those with limited financial resources may lack access to necessary devices [6].

To address these gaps, we conducted a sequential mixed methods study grounded in the Structural Influence Model (SIM) and message framing theory. The SIM posits that health behaviors are shaped by social, economic, and environmental contexts [4], while message framing suggests that information presentation influences attitudes and decisions regarding health behaviors [7,8]. Phase 1 used qualitative methods to examine motivations and barriers for telemonitoring adoption among socioeconomically disadvantaged older adults. Phase 2 evaluated the effectiveness of different message framing approaches (episodic vs thematic) and modalities (video vs photo-text) to evoke motivations identified in phase 1 and therefore enhance telemonitoring acceptance, with attention to how health locus of control (LOC) moderates these effects. By investigating both the factors influencing telemonitoring adoption and testing targeted communication strategies, our research aims to contribute to the development of more inclusive and effective telemonitoring interventions that can help bridge the digital

XSL•FO
RenderX

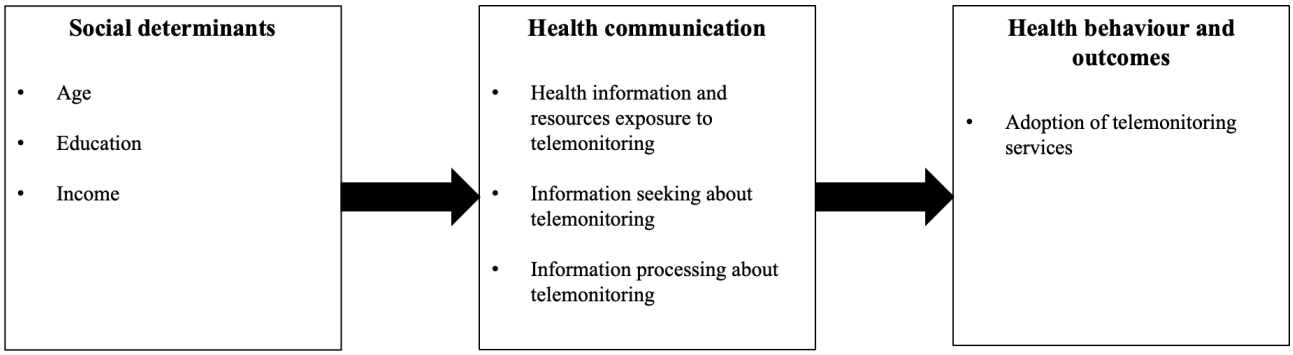
divide and improve health outcomes for socioeconomically disadvantaged older adults. This work responds to calls within social gerontology for research that addresses the intersection of technological innovation, social inequality, and the diverse lived experiences of aging [9,10], ultimately supporting the creation of more equitable models of digital health that enhance older adults' agency and quality of life in their preferred living environments.

Theoretical Framework

Structural Influence Model and Phase 1: Exploring Telemonitoring Adoption Among Older Adults

The SIM provides a theoretical framework for understanding how social determinants interact with communication processes to influence health behaviors and outcomes [11]. This model posits that communication outcomes, including exposure, usage, attention, seeking, and processing of health information, are interlinked with social determinants to influence health outcomes (Figure 1). According to SIM, health information benefits are predominantly accessible to those with the capacity to access, process, and use it, highlighting potential for health technologies to widen but not narrow health disparities.

Figure 1. Modified structural influence model in telemonitoring services.



When applied to older adults' engagement with telemonitoring technologies, the model complements the political economy of aging theory [12,13], which emphasizes how socioeconomic structures create and maintain inequalities in later life. Thus, as shown in Figure 1, we adapted Viswanath et al's [11] 4-component framework into a 3-component framework centered on socioeconomic determinants (age, education, and income) and their relationship to health information exposure, seeking, and processing related to telemonitoring. Consistent with SIM, the digital health care revolution may disproportionately benefit those with greater socioeconomic resources, technological skills, and health literacy [14]. Friemel [15] showed that the digital divide persists among seniors, encompassing not just technology access but also differences in usage patterns, skills, and perceived benefits. Research has identified social determinants such as age, income, and education as significant factors influencing telemonitoring adoption [16,17].

Despite existing literature linking social determinants and communication outcomes with digital health adoption, significant gaps persist. Socioeconomic status (SES) complexity necessitates diverse social deprivation measures, tailored to research objectives and cultural contexts. While scholars commonly use income, occupation, and education levels to delineate social standing when exploring health disparities [18], Darin-Mattsson et al [19] suggest that in old age, educational attainment may have a weaker association with health compared to wealth, income, and tenure. These findings underscore the importance of focusing on low-income older adults when examining telemonitoring adoption motivators and barriers. Consequently, guided by SIM, we aimed to address the following research question (RQ):

RQ1: What are the motivators and barriers associated with the adoption of telemonitoring services among low-income older adults, as understood through the lens of information exposure, usage, attention, seeking, and processing?

Phase 2: Message Framing Strategies for Telemonitoring Acceptance

Building upon Phase 1 findings, which highlighted the importance of affordability, convenience, and health benefits, as well as the challenges posed by preferences for in-person consultations, perceived limitations of remote monitoring, and diminished information processing abilities and health LOC beliefs. Phase 2 aimed to develop and test effective communication strategies to enhance telemonitoring adoption. While Phase 1 identified the key drivers of adoption, altering the technology itself or its accessibility through subsidies, redesign, or infrastructure is better addressed through policy or engineering, and falls out of the scope of health communication. Instead, Phase 2 was designed on the premise that user awareness of its benefits and the perceived relevance of those benefits are essential prerequisites for adoption, regardless of a service's objective advantages [20]. That indicates that the critical question is not only whether these benefits exist, but whether they are being communicated in a way that resonates with older adults, as prior research has shown that older adults often fail to recognize the personal relevance of technologies even when aware of their potential benefits [21]. Accordingly, experimental stimuli of 2 modalities held core service information constant across framing conditions, varying only how this information was presented.

XSL•FO
RenderX

Episodic-Thematic Framing in Message Effects

Framing holds particular promise as a major message format studied across diverse health contexts. It involves strategically presented information where attention is drawn toward specific details [8]. In the context of older adults' technology adoption, framing effects can be particularly significant given the experience gap, the discrepancy between older adults' life experiences and new technologies' demands [22]. This gap creates ambiguity about the value and relevance of technologies, making the framing of messages particularly influential in shaping perceptions and attitudes.

Framing theory posits that media messages shape audience perception and interpretation by highlighting particular aspects of an issue [8]. Iyengar [23] proposed an influential distinction between episodic framing and thematic framing. Episodic framing focuses on concrete exemplars, isolating single events or individuals while offering engaging narratives to depict broader issues via unknown and ordinary individuals [23]. Thematic framing focuses on a more societal context while offering statistical evidence. These framing approaches intersect with a sociocognitive perspective on aging, emphasizing how older adults' perceptions, beliefs, and decision-making are shaped by both individual cognitive processes and social contexts [24]. In the health communication context, episodic messages might describe a specific individual's experience with an illness, while thematic messages might present population-level statistics or attribute the issue to broader social factors [25,26]. In this study, the distinction was operationalized as follows: episodic messages featured a personal account from a named older adult regarding their use of health screening services, while thematic messages presented national survey statistics and referenced community health infrastructure (Experimental Stimuli in Methods section).

Findings from Phase 1 showed that only a small number of individuals believed they were fully responsible for their personal health, while the majority displayed a sense of health responsibility attributions to external parties such as the government or medical professionals. This divergence aligns with research in health communication indicating that responsibility for health outcomes is often attributed to a combination of individual and social or systemic factors [27]. Specifically, older adults' sense of health agency often involves negotiating between personal responsibility and reliance on health care systems, particularly as physical vulnerabilities increase with age, consistent with evidence that older adults' perceptions of aging shape their health and functioning [28]. How health messages frame this balance between individual and systemic responsibility may significantly impact their effectiveness. Iyengar's [23] framework was originally designed to explain how episodic and thematic frames shape audience attributions of responsibility for social issues (ie, attributing responsibility to the individual vs society or institutions); thus, it offers a theoretically grounded perspective for examining whether older adults holding differing views on health responsibility resonate differently with these frames.

The distinction between episodic and thematic framing may be closely related to *exemplification theory* [29]. This theory distinguishes between exemplar-based information (ie, vivid and specific cases) and base-rate statistical information, elucidating the cognitive mechanisms underlying their differing effects. According to the theory, exemplar-based information exerts a disproportionate influence on judgment because it is cognitively more accessible and more readily perceived as representative of broader group characteristics compared to abstract statistical data [29]. Recent research strongly supports this effect: a meta-analysis found that across various communication contexts, exposure to exemplar-based information consistently produced stronger effects than exposure to base-rate information [30]; these effects were most pronounced regarding "first-order reality judgments," such as assessments of prevalence or risk [31]. From this perspective, episodic framing can be viewed as a form of exemplar presentation, whereas thematic framing corresponds to the presentation of base-rate information; and this offers a cognitive-mechanism-level complement to explanations regarding the persuasive advantage of episodic framing [32]. These characteristics may make exemplar-based information particularly relevant for some older adults. Age-related declines in information processing capabilities [33], together with a preference for emotionally meaningful and concrete content [34], suggest that older audiences may be especially sensitive to the availability heuristic and representativeness heuristic underlying the exemplar effect. Therefore, we propose the following hypothesis:

H1: Compared to those exposed to thematic framing, older adults exposed to episodic framing report more positive (1) attitudes toward and (2) intention to use telemonitoring services.

Mediating Role of Attitude

Based on Ajzen's [35] Theory of Planned Behavior (TPB), attitude is one of the predictors demonstrating strong and consistent relationships with intention to perform a behavior. This theoretical relationship has particular relevance for older adults given innovation resistance often observed in later life [36]. This resistance isn't merely about technological difficulty but reflects complex attitudinal factors including concerns about privacy, perceived usefulness, and compatibility with existing routines and values [37]. Understanding how attitudes mediate the relationship between message framing and adoption intentions can therefore provide insights into the psychological mechanisms underlying older adults' technology decisions. Similarly, studies have found evidence supporting significant relationships between attitudes and behavioral intentions in the context of telehealth acceptance and COVID-19 vaccination [38-40]. Thus, we posited:

H2: Attitude toward using telemonitoring services is positively associated with intention to use telemonitoring services.

H3: The relationship between framing and intention to use is mediated by attitude toward using telemonitoring services.

Moderating Role of Video and Photo-Text Modality

The moderating role of modality suggests the delivery method of messages through media platforms significantly influences message reception [41]. Defined modalities, categorized into video and photo-text, are informed by dual-coding theory [42],

XSL•FO
RenderX

which posits that visual and verbal information is processed by distinct cognitive systems. For older adults, message modality may have heightened importance due to age-related changes in sensory processing and cognitive function. Charness and Boot [43] note that age-related changes in visual and auditory processing can affect how older adults engage with different media formats. This theoretical framework supports that videos, combining visual with conversational elements, can enhance comprehension and retention, making them particularly effective for presenting eHealth information to older adults [44].

Evidence suggests varying impacts based on the chosen modality. One study [45] found that both gain- and loss-framed photo-text stimuli significantly influenced health behavior intentions, whereas another [46] demonstrated that loss-framed videos increased perceived severity of health risks more effectively than gain-framed videos. de Graaf et al [47] and Occa and Suggs [48] highlighted the stronger impact of video messages on health behavior attitudes compared to photo-text modalities. Therefore, we proposed:

H4: Modality and framing would interact, such that participants assigned to the video condition would have a more positive attitude toward using telemonitoring services compared to participants assigned to the photo-text condition.

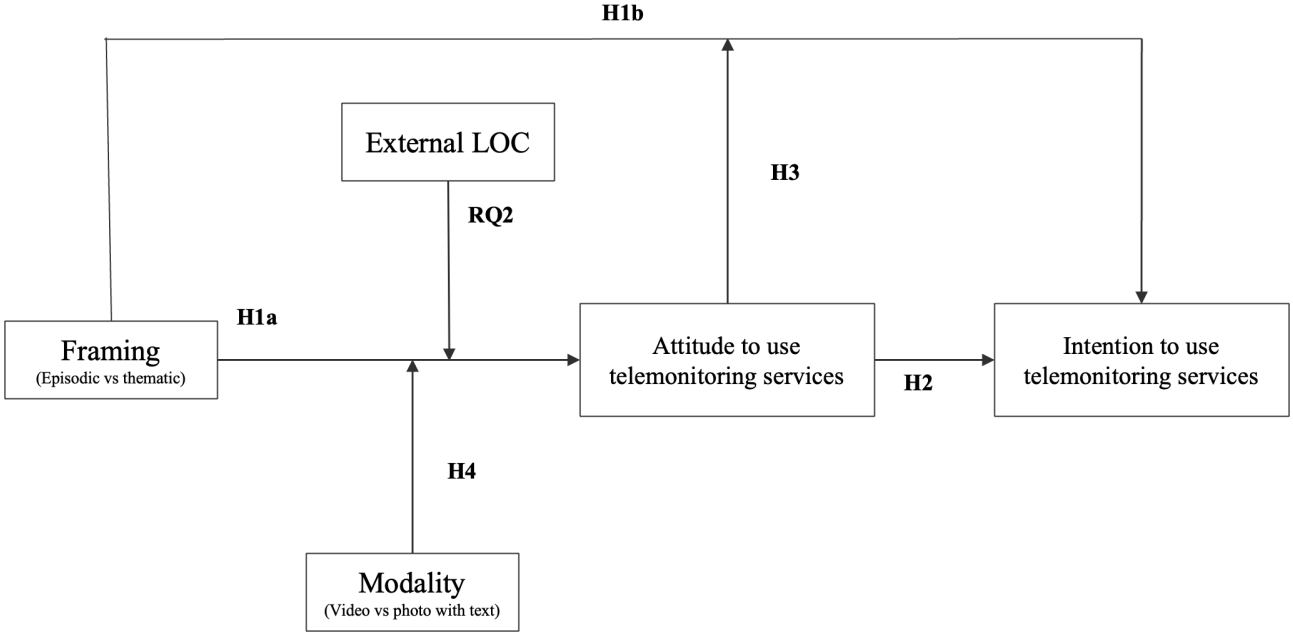
Moderating Role of External LOC

Wallston et al [49] defined health LOC as the degree to which individuals believe their health is controlled by their own behavior versus external forces. People with an internal LOC believe health status is determined by internal causes such as their own behavior. Meanwhile, those with an external LOC believe that one's health depends on external causes or powerful others such as health professionals. Health LOC intersects with gerontological theories about control beliefs in later life. The selective optimization with compensation model [50] suggests successful aging involves adjusting control strategies as resources and capacities change. External LOC might represent an adaptive response to increasing health challenges rather than simply passive orientation [51]. Prior research supports a strong association between numerous sociodemographic variables and external LOC, with studies finding that those of older age, lower income, and lower education were more external in terms of health LOC [52,53].

Current literature found individuals with a stronger external health LOC are more likely to find thematic framing messages more appealing [54], as they emphasize societal or group responsibility. Similarly, Ceylan and Hayran [55] verified that LOC moderated the effects of prosocial and self-interested messages in following COVID guidelines, where individuals with higher external LOC were more influenced by prosocial messages. In the context of interactive health websites, those with higher external LOC are responding more strongly to vivid, exemplar-rich message formats [56]. However, other studies have shown contrasting or inconclusive results on the link between episodic or thematic framing and individual and societal attributions [57,58]. Hence, the following question was asked:

RQ2: To what extent does external LOC moderate the relationship between framing and attitude toward using telemonitoring services? The proposed research model and hypotheses are presented in Figure 2.

Figure 2. Proposed diagram of the pilot health communication campaign on telemonitoring adoption. H1a: hypothesis 1a; H1b: hypothesis 1b; H2: hypothesis 2; H3: hypothesis 3; H4: hypothesis 4; LOC: locus of control; RQ2: research question 2.



Methods

Research Design

This research used a sequential mixed methods design with 2 phases. Phase 1 conducted qualitative interviews and home observations with older adults to identify key motivators and barriers to telemonitoring adoption through the lens of the SIM. Phase 2 experimentally tested how message framing approaches and modalities might influence telemonitoring acceptance among the target population. This sequential design allowed

XSL•FO
RenderX

development of evidence-based communication strategies informed by the lived experiences.

Ethical Considerations

Overview

The research was conducted with FullHealth (provided by FullHealth Asia), a grassroots initiative that provides telemonitoring services through health screening and monitoring booths (Figure 3). This study received approval from the Institutional Review Board (IRB) of Nanyang Technological University (IRB-2022 - 826). All participants provided written informed consent prior to participation in the study, with protocols designed to accommodate potential age-related needs in the consent process. Before each session, participants were briefed on the study's goals and their participation rights, with consent for audio-recording obtained and anonymity assured through the use of pseudonyms. All data were deidentified by using numerical codes, and any identifying information mentioned in the interview transcripts was removed. The data were stored securely and were accessible only to research team members (DW, HB, APTQ, GHYH, PSH, ESMO, and EWJL). All participants received S$30 (US $23.40) compensation.

Figure 3. FullHealth telemonitoring booth at community venues.



Phase 1: Qualitative Study on Motivators and Barriers

Participants and Sampling

Our research focused on older adults in Singapore, a rapidly aging society where telemonitoring services offer significant potential. We targeted Singaporean adults aged 50 years and older to capture current and future older populations. Participants mainly earned ≤S$3000 (US $2340)/person/month, near the national median of S$3287 (US $2563.86) [59], and included English and Mandarin speakers per census data [60]. We recruited participants through convenience sampling facilitated by FullHealth. We included both users and nonusers of the FullHealth telemonitoring service to capture diverse perspectives.

Data Collection

Data collection was conducted in person in November 2022, using semistructured interviews and observation guides in English and Mandarin. We conducted pilot sessions with older adults familiar with telemonitoring but not participating in the main study to refine the guides for effectiveness while minimizing cognitive burden. Main sessions, facilitated by trained FullHealth volunteers, lasted approximately 90 minutes each (N=17), with appropriate breaks.

Data Analysis

We applied thematic analysis, following Braun and Clarke's [61] guidelines. The last (EWJL) and corresponding (HB) authors independently coded a subset of the data (4/17, 23.53% of participants) through iterative open coding. After comparing findings, discrepancies were resolved through discussion. The remaining data were then independently coded. Cohen Kappa coefficient (κ=0.74) indicated substantial intercoder agreement [62].

Phase 2: Experimental Study on Message Framing

Study Design

We implemented a 2×2 between-subjects experimental design with framing conditions (episodic and thematic) and media formats (video and photo-text), engaging 120 participants from January 16 to February 1, 2023. Participants were recruited using convenience and snowball sampling through FullHealth Asia, referrals, and social media. All participants were unfamiliar with the FullHealth booth.

Experimental Stimuli

Framing conditions: thematic framing highlighted societal issues using "Singapore" in titles, statistical data, national

XSL•FO
RenderX

surveys, and references to older adult activity centers to reflect broader societal concerns. Episodic framing focused on the individual, using direct address (eg, "You"), personal stories, and specific experiences. Both framing approaches ensured clarity for older audiences.

Modalities: Stimuli maintained design consistency through uniform color schemes, fonts, and logos. Photo-text stimuli are provided in Multimedia Appendix 1; the 4 video stimuli are provided in Multimedia Appendices 2-5. Differences were marked by imagery, text, and narration changes for video content, accommodating potential age-related sensory and cognitive changes.

Crucially, the core information of service description, emphasizing that FullHealth's booth was free, convenient, and offers comprehensive health checks (blood sugar, blood pressure, height, weight, heart, and eyes), remained consistent across the episodic and thematic conditions in both the English and Chinese versions. The only difference between the presentation styles lay in the narrative approach used to convey this information (personal narratives vs aggregate statistics combined with organizational background), isolating the framing manipulation from message content.

Procedure

Participants were randomly allocated to 4 conditions: (a) episodic video, (b) episodic photo-text, (c) thematic video, or (d) thematic photo-text. After informed consent, participants assessed external LOC, viewed assigned stimulus, then completed attitude and intention measures plus manipulation checks. A debriefing session clarified any misinformation during the study. Participants received S$15 (US $11.70) as compensation.

Measures

External LOC: this moderating variable was measured using a 5-point Likert scale adapted from the Health Locus of Control Scale (HLCS) by Wallston et al [63], using items such as "Good health is largely a matter of good fortune," with mean 3.13 (SD 0.60, Cronbach α=0.72).

Attitudes toward telemonitoring: participants' attitudes were measured using a 5-point semantic differential scale, with items modified from Dabholkar and Bagozzi [64]. Participants rated their feelings toward using the FullHealth booth on dimensions such as "Good to bad" and "Harmful to beneficial," with mean 4.03 (SD 0.47, Cronbach α=0.85).

Intention to use telemonitoring services*:* this was measured on a 5-point Likert scale, adapted from Boontarig et al [65], with items such as "I intend to use this FullHealth booth in the future," mean 4.14 (SD 0.65, Cronbach α=0.90).

Manipulation check: for framing, participants assessed statements reflecting message focus. Items for episodic framing evaluated FullHealth booth experiences from an individual older adult's perspective, while thematic framing assessment involved societal impacts and statistical evidence. For modality, 2 items checked participants' recognition of the stimulus format (video or photo-text).

Data Analysis

We analyzed the data using Hayes' PROCESS macro (Model 9) [66] via SPSS Statistics (IBM Corp), incorporating gender as a covariate. Because PROCESS estimates a series of ordinary least squares regression equations rather than a simultaneously estimated structural equation model, the coefficients reported below reflect independent regression paths but not a jointly estimated path model; accordingly, global model fit indices (such as χ^2, Comparative Fit Index [CFI], and root-mean-square error of approximation [RMSEA]) are not applicable here. Message framing was the independent variable, with attitudes and intentions toward telemonitoring services as dependent variables. External LOC and modality were considered moderating variables. The complete set of model estimation coefficients (including all covariate, interaction paths, and conditional indirect effects at representative values of the moderators) is detailed in Multimedia Appendix 6, together with bootstrap SE and 95% CI.

Statistical Power

A sensitivity power analysis indicated that, based on the sample size (N=120) and the 6-predictor moderated regression model used to test Hypotheses 1 - 4 and RQ2, the study had 80% power (α=.05) to detect an effect of Cohen $f^2 \geq .067$, an effect size falling between the conventional benchmarks for small (.02) and medium (.15) effects [67]. Post hoc power analyses regarding the key effects reported herein confirm that the study possessed sufficient power to detect the significant "framing× locus of control" interaction (f^2=.095, power=.92) and the relationship between attitudes and intentions (f^2=.173, power>.99). The post hoc statistical power for the main effects of framing, presentation modality, and gender, as well as the "framing× presentation modality" interaction, was relatively low (0.06 - 0.10) (Table 1).

XSL•FO
RenderX

Table . Sensitivity and post hoc power analysis results.

Effect	Observed f^2	Post hoc power
Framing × External LOC[a] → Attitude	0.095	0.918
Attitude → Intention	0.173	0.995
External LOC (main) → Attitude	0.018	0.302
Framing × Modality → Attitude	0.004	0.099
Modality (main) → Attitude	0.002	0.082
Framing (main) → Attitude	0.001	0.058

[a]LOC: locus of control.

Results

Phase 1: Motivators and Barriers to Telemonitoring Adoption

The qualitative phase involved 17 participants, 10 telemonitoring users (denoted by “U”) and 7 nonusers (denoted by “N”), aged 56 - 86 years. Most participants (15/17, 88.2%) had attained education up to A-level or Diploma, while 11.8% (2/17) had no formal education. The majority (14/17, 82.4%) reported monthly income below S$2000 (Table 2). Analysis revealed 3 key motivators, 3 barriers, and 2 influential factors affecting telemonitoring adoption emerged (Figure 4).

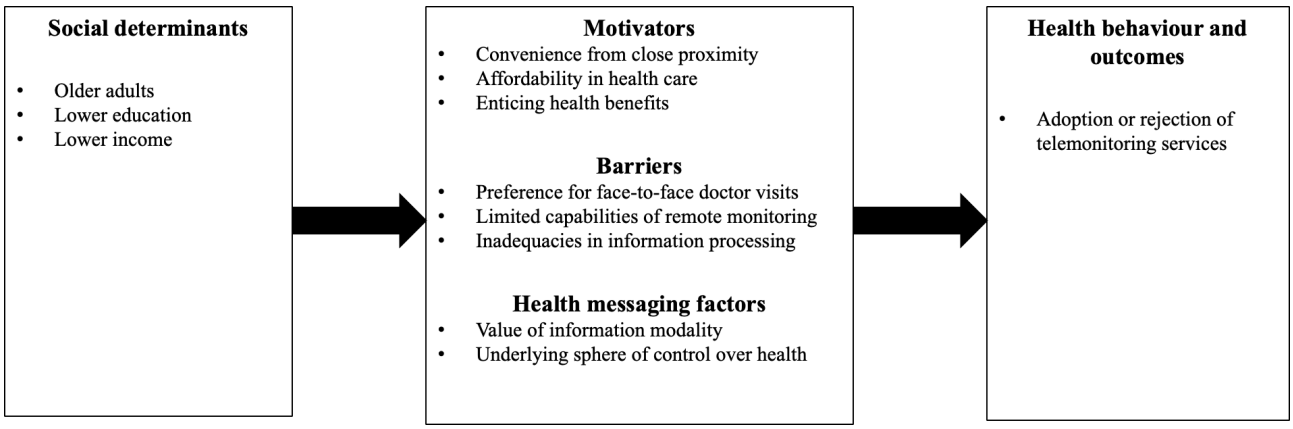
XSL•FO
RenderX

Table . Demographics of participants in the interview and home observation.

Demographics	Total (N=17)	Users (n=10)	Nonusers (n=7)
Sex, n (%)			
Male	12 (70.59)	9 (90.00)	3 (42.86)
Female	5 (29.41)	1 (10.00)	4 (57.14)
Age (in years), n (%)			
50 - 59	4 (23.53)	3 (30.00)	1 (14.29)
60 - 69	6 (35.29)	5 (50.00)	1 (14.29)
70 - 79	6 (35.29)	1 (10.00)	5 (71.43)
80 - 89	1 (5.88)	1 (10.00)	0 (0.00)
Language Spoken, n (%)			
English	4 (23.53)	2 (20.00)	2 (28.57)
Mandarin	12 (70.59)	7 (70.00)	5 (71.43)
English and Chinese	1 (5.88)	1 (10.00)	0 (0.00)
Education, n (%)			
No formal education	2 (11.76)	1 (10.00)	1 (14.29)
Primary 6 or below	5 (29.41)	3 (30.00)	2 (28.57)
Some secondary	3 (17.65)	1 (10.00)	2 (28.57)
N-level or ITE[a]	1 (5.88)	1 (10.00)	0 (0.00)
O-level	4 (23.53)	2 (20.00)	2 (28.57)
A-level or Diploma	2 (11.76)	2 (20.00)	0 (0.00)
Degree	0 (0.00)	0 (0.00)	0 (0.00)
Postgraduate	0 (0.00)	0 (0.00)	0 (0.00)
Monthly Income, n (%)			
Less than S$1,000 (US $780)	6 (35.29)	4 (40.00)	2 (28.57)
S$1,000 to S$1,999 (US $780 to US $1,559.22)	8 (47.06)	6 (60.00)	2 (28.57)
S$2,000 to S$2,999 (US $1,560 to US $2,339.22)	1 (5.88)	0 (0.00)	1 (14.29)
S$3,000 to S$3,999 (US $2,340 to US $3,119.22)	2 (11.76)	0 (0.00)	2 (28.57)

[a]ITE: Institute of Technical Education.

Figure 4. Results of motivators and barriers to telemonitoring services adoption.



XSL•FO
RenderX

Key Motivators

Convenience From Close Proximity

Most FullHealth users cited proximity as the primary reason for using the telemonitoring booth, allowing them to "*save time*" (U7, male, 57 years) and:

> *[not] have to travel all the way to [the] polyclinic.* [U8, male, 63 years]

Nonusers also acknowledged "*convenience*" as a potential motivator. Given the observed mobility challenges, convenience appeared to be particularly significant, reflecting how aging-related changes shape technology preferences.

Affordability in Health Care

The "*FOC [Free-of-Charge]*" (U10, Female, 64 years) nature of the booth services encouraged usage among participants. While some noted that the cost advantage might be less significant compared to other subsidized older adult care services (eg, "*This doctor I visit helps me save money. It's not very expensive because I'm old"* [N1, female, 75 years]), comparing the booth with a regular doctors' visit, cost remained critical. When discussing more expensive health services, participants expressed concerns:

> *Of course, it's money problems. Every time also doesn't allow you to cut [medical fees], sum [up] to 37.50 dollars.* [N7, male, 63 years]

Overall, participants demonstrated high price sensitivity, highlighting how financial considerations become increasingly significant in postretirement life.

Enticing Health Benefits

Many participants cited health benefits as a key motivator. These benefits included addressing specific health issues (eg, "*it helps prevent me from getting a stroke again"* [U3, male, 59 years]) and general health monitoring (eg, "*if we have any illness, we will know"* [U8, male, 63 years]). As one participant summarized:

> *In conclusion, [FullHealth Booth] is good for our health.* [U5, male, 86 years]

Participants noted their increased susceptibility to illnesses and, making health monitoring particularly valuable, reflecting how health vigilance increases with advancing age.

Key Barriers

Preference for Face-to-Face Doctor Visits

A primary barrier was the preference for in-person medical consultations. Participants felt that face-to-face meetings provided better understanding of their health concerns (eg, "*Seeing the doctor. Talk about our illnesses [...] clearer. We can understand"* [U4, male, 76 years]; "*When in front of a doctor, if we have any questions, we can ask, which is better"* [U2, male, 57 years]), and offered a sense of familiarity (eg, "*The doctor if we see often already, [we] will go find that doctor"* [N2, female, 78 years]). These preferences highlight how technology adoption in later life is influenced by established relational patterns and trust.

Limited Capabilities of Remote Monitoring

The booth's restricted monitoring capabilities were considered "*very simple*" (U9, male, 65 years) and "*not comprehensive, it's not full*" (N3, female, 56 years), presenting a barrier to adoption. One participant noted:

> *if you go to poly[clinic], one blood test can cross-check a lot of things.* [N3, female, 56 years]

This perception of limited functionality compared to traditional health care services influenced participants' willingness to adopt telemonitoring.

Inadequacies in Information Processing

Information processing posed a major barrier, especially for lower socioeconomic participants who struggled to understand health advertisements (eg,"*[When shown a health advertisement on the National Steps Challenge] also can't really understand when looking"* [N2, female, 78 years]); *"[When asked what the National Steps Challenge Advertisement is talking about] Don't know*" (U8, male, 63 years). Information processing abilities varied notably based on education and income levels. Participant U4 (male, 76 years) with no formal education and earned less than S$1000 per month, demonstrated poor information processing skills (*"[When asked about the TraceTogether Advertisement] never studied, don't understand"* [U4, male, 76 years]). In contrast, participant N3 (female, 56 years), with an O-level certificate and monthly income between S$3000 and S$3999, showed better comprehension (*"[When asked about the TraceTogether Advertisement] To track people where they, in case of Covid, they know who is where, you know, Yeah, so they can trace the people and then inform them that, you know"* [N3, female, 56 years]). These variations highlight how lifelong socioeconomic advantages and disadvantages accumulate to influence technology engagement.

Factors Influencing Information Processing and Telemonitoring Adoption

Value of Information Modality

Participants cited various media sources including "*posters,*" "*pamphlets,*" "*flyers,*" "*radio,*" "*TV,*" "*newspapers,*" "*Facebook,*" "*WhatsApp,*" "*YouTube,*" "*Websites,*" *and "[digital display screens in] lift.*" Although television and mobile phones were the most commonly owned, neither was frequently used by most participants. Multiple media sources were often mentioned (eg, "*I think from the website, [...] flyers, [...] more news"* [N3, female, 56 years]; "*See the news. [...] I buy newspapers [...] Yes, all [Whatsapp and Facebook] got used, also use Youtube"* [U2, male, 57 years]). The diverse media preferences suggested the importance of considering various modalities in health communication.

Underlying Sphere of Control Over Health

Participants displayed varying beliefs about control. Some demonstrated a sense of personal responsibility (eg, "*This [health screenings] is [my] responsibility. If you don't go, then you have a lot of illnesses, then blame who?"* [U5, male, 86 years]); "*For my blood pressure, if it's high, I have to control [it]. If my blood sugar is high, my food intake has to [be] control[led]. I think it's very useful for elderly, people like us*

XSL•FO
RenderX

to take care of our health" (U7, male, 57 years). However, most participants showed a tendency to be influenced by health care providers and government authorities (eg, "*[When asked if the participant wanted to undergo a cataract surgery] I then refute back and say you're a doctor, you then know what, I'm not a doctor how should I know*" [N4, female, 73 years]; "*over there (i.e Chinese medicine clinic) is the volunteer that comes up and helps me do (pump pressure*)" [U3, male, 59 years]). These varying health LOC orientations suggested potential for further investigation (Figure 4).

Phase 2: Effects of Message Framing on Telemonitoring Acceptance

Building on Phase 1 findings, Phase 2 experimentally tested the effects of different message framing approaches and modalities on telemonitoring acceptance. Participants included 79 (65.8%) women and 41 (34.2%) men, with 98 (81.7%) participants earning below S$3000 monthly. Detailed demographics are presented in Table 3.

XSL•FO
RenderX

Table . Demographics of participants in the pilot experimental study.

Demographics	Participants (N=120), n (%)
Sex	
Male	41 (34.2)
Female	79 (65.8)
Citizenship status	
Singaporean	117 (97.5)
Permanent resident	3 (2.5)
Age (in years)	
50 - 57	19 (15.8)
58 - 65	41 (34.2)
Older than 65	60 (50)
Education	
No formal education	1 (0.8)
Primary 6 or below	6 (5)
Secondary	21 (17.5)
N-level or ITE[a]	1 (0.8)
O-level	20 (16.7)
A-level or Diploma	41 (34.2)
Degree	19 (15.8)
Postgraduate	11 (9.2)
Monthly Income	
Less than S$1,000 (US $780)	45 (37.5)
S$1,000 to S$1,999 (US $780 to US $1,559.22)	38 (31.7)
S$2,000 to S$2,999 (US $1,560 to US $2,339.22)	15 (12.5)
Above S$3,000 (US $2,340)	22 (18.3)
Occupation	
Employed full-time	13 (10.8)
Employed part-time	25 (20.8)
Self-employed	13 (10.8)
Not working	12 (10)
Retired	46 (38.3)
Homemaker	11 (9.2)
Housing type	
Rental block	2 (1.7)
Studio apartment	6 (5)
1-room HDB[b] flat	0 (0)
2-room HDB flat	9 (7.5)
3-room HDB flat	24 (20)
4-room HDB flat	23 (19.2)
5-room HDB flat	30 (25)
Executive condominium	4 (3.3)
Condominium	15 (12.5)

XSL•FO
RenderX

Demographics	Participants (N=120), n (%)
Bungalow or detached house	3 (2.5)
Others	4 (3.3)
Marital status	
Married	77 (64.2)
Single	29 (24.2)
Separated	2 (1.7)
Divorced	6 (5)
Widowed	6 (5)

[a]ITE: Institute of Technical Education.

[b]HDB: Housing and Development Board.

Manipulation Check

T test analysis for framing manipulation revealed significant differences between episodic (mean 3.67, SD 0.58) and thematic (mean 2.11, SD 0.55) conditions (t_{118}= 15.10; *P*<.001). Similarly, the modality manipulation was successful, with participants correctly identifying their assigned condition, with video: mean 4.11, (SD 0.62) versus photo-text: mean 1.80 (SD 0.53, t_{118}= 22.13; *P*<.001).

Effects of Message Framing on Attitudes and Intentions

Analysis revealed that framing did not exert a significant main effect on attitudes toward telemonitoring services (b=−.03, t_{113}=−0.26; *P*=.80). Consistent with this conditional pattern, the direct effect of framing on intentions to use telemonitoring services also did not reach statistical significance (b=−.01, t_{116}=−0.11; *P*=.91). Participants exposed to episodic framing reported attitudes (mean 4.04, SD 0.52) and intentions to use telemonitoring services (mean 4.14, SD 0.61) broadly similar to those exposed to thematic framing ($mean_{attitudes}$ 4.02, $SD_{attitudes}$ 0.41; $mean_{intentions}$ 4.13, $SD_{intentions}$ 0.58). H1 was therefore not supported as an unconditional hypothesis.

Relationship Between Attitudes and Intentions

We found a significant positive relationship between attitudes toward using telemonitoring services and intentions to use these services (b=.56, t_{116}=4.47; *P*<.001), supporting the hypothesis that favorable attitudes positively influence usage intentions (H2). Analysis revealed the unconditional indirect effect of message framing on intention through attitude was not significant, regardless of modality (95% bootstrap CI included zero in both cases). However, this indirect effect was conditional on external LOC (Index=−.28, 95% CI −0.52 to −0.08), while the index for modality was not (Index=.06, 95% CI −0.13 to 0.24). At representative levels of the moderators, framing exerted a significant indirect effect via attitude among participants with low external LOC exposed to video (b=.21, 95% CI 0.02-0.44) and among those with high external LOC exposed to photo-text (b=−.18, 95% CI −0.36 to −0.03). H3 was therefore supported only conditionally, through external LOC.

Moderating Effects of Message Modality

Contrary to expectations, the analysis showed no significant interaction effect between modality (video vs photo-text) and message framing on attitudes toward telemonitoring services (b=.11, t_{113}=0.63; *P*=.53). This suggests the modality did not significantly influence how framing affected attitudes in our experimental conditions, indicating that older adults may process different media formats similarly when evaluating health technology information.

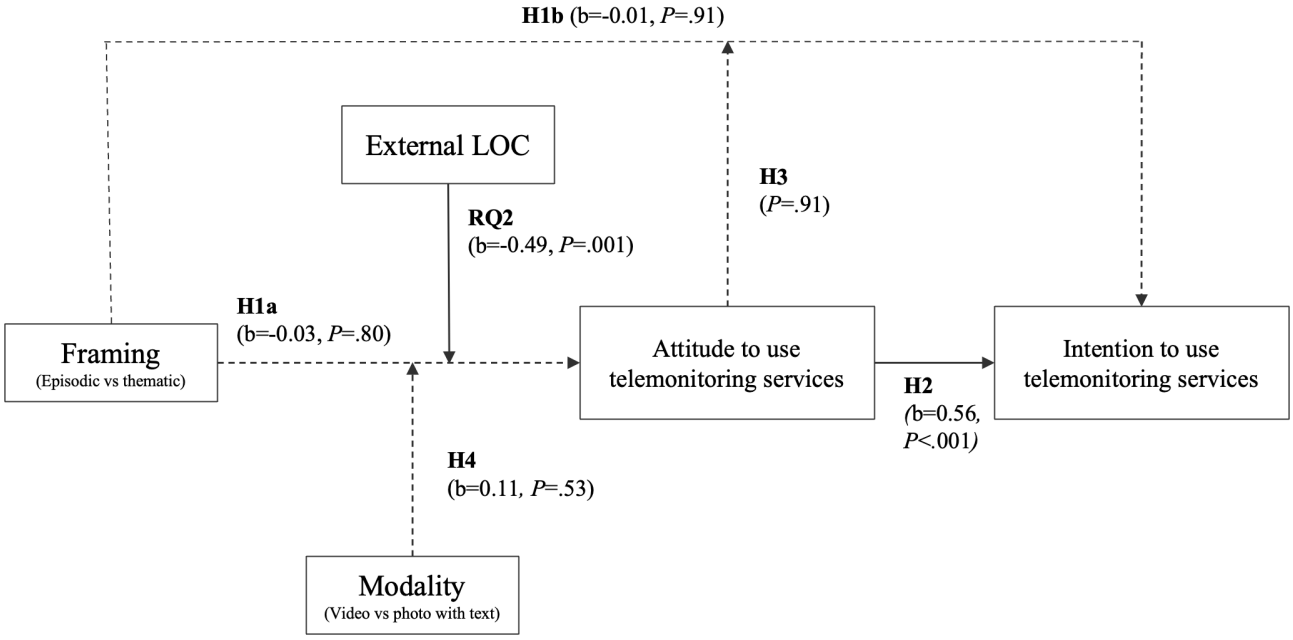
Moderating Effects of External LOC

External LOC significantly moderated the effect of message framing on attitudes toward using telemonitoring services (b=−.49, t_{113}=−3.28; *P*=.001). When exposed to episodic framing, participants with lower external LOC showed more positive attitudes (mean 4.21, SD 0.47) compared to those with higher external LOC (mean 3.99, SD 0.47). Conversely, thematic framing resulted in more positive attitudes among participants with higher external LOC (mean 4.15, SD 0.47) compared to those with lower external LOC (mean 3.98, SD 0.47). This finding suggests that individuals' beliefs about health outcome sources play a crucial role in determining the effectiveness of different message framing approaches, reflecting how psychological orientations influence technology adoption decisions in later life.

Figure 5 provides a conceptual summary of the regression paths estimated using Hayes' PROCESS macro (Model 9). The coefficients shown are unstandardized (b) and are taken from the corresponding ordinary least squares regression equations; as noted in the Data Analysis section, global model fit statistics are not applicable.

XSL•FO
RenderX

Figure 5. Conceptual summary of regression paths estimated using PROCESS Model 9. H1a: hypothesis 1a; H1b: hypothesis 1b; H2: hypothesis 2; H3: hypothesis 3; H4: hypothesis 4; LOC: locus of control; RQ2: research question 2.



Discussion

Principal Findings

Our sequential mixed methods study provides valuable insights into factors influencing telemonitoring adoption among socioeconomically disadvantaged older adults in Singapore. By first identifying key motivators and barriers through qualitative methods then experimentally testing communication strategies, we developed a comprehensive understanding of how to effectively promote telemonitoring acceptance. Our qualitative findings, anchored in the SIM, identified 3 key facilitators of telemonitoring adoption: convenience, affordability, and health benefits. These align with prior research [68,69] and underscore how their absence magnifies existing barriers for older adults with lower socioeconomic status.

Affordability and convenience emerged as essential elements, reflecting the structural theories of aging, wherein socioeconomic constraints often intensify in later life [13]. Prohibitive costs and inconvenience of traditional health care services disproportionately affect those with constrained retirement incomes and limited mobility [70], underscoring how economic inequalities accumulated throughout the life course can restrict technology access and adoption [71]. Perceived health benefits of continuous monitoring served as a powerful motivator, suggesting older adults strategically invest in technologies that help manage age-related health changes and maintain independence [72].

The barriers we identified (preference for in-person consultations, perceived limitations of remote monitoring, and diminished information processing abilities) all relate to SIM components while reflecting age-specific considerations. The preference for face-to-face interactions exemplifies embodied knowledge in health care, showing how accumulated physical and relational experiences shape how older adults interact with health care systems [73]. This reflects established patterns of trust and communication developed throughout the life course [74]. The perceived limitations of telemonitoring relate to information exposure and attention, suggesting a need for clearer communication about the proactive monitoring benefits [1]. Perhaps most critically, the challenge of information processing, exacerbated by cognitive changes and socioeconomic factors [75,76], directly relates to the SIM component of information processing.

Message Framing, Modality, and Health LOC in Later Life Technology Decisions

Building on our qualitative findings, our experimental study revealed important insights about effective communication strategies for promoting telemonitoring adoption among older adults. Most notably, episodic framing's persuasive power was conditional on external LOC: our results showed that individuals with lower external LOC responded more favorably to episodic framing, while those with higher external LOC showed greater receptivity to thematic framing. This message-audience matching pattern aligns with a growing body of evidence that dispositional beliefs about control shape which type of health message works [26,55], and a recent systematic review confirms that health LOC systematically conditions how audiences respond to behavioral health interventions more broadly [77]. Substantively, individuals lower in external LOC likely process episodic narratives as personally self-relevant and congruent with their belief that health outcomes are within their own control; individuals higher in external LOC, who attribute health outcomes more to institutions, systems, or chance, may instead find thematic framing more credible and applicable to their own situation, which explicitly foregrounds statistical evidence and institutional infrastructure [34,78,79]. The prevalence of external health LOC aligns with life-span developmental perspectives suggesting that control beliefs often shift toward greater externality with advancing age, reflecting adaptive responses to increasing dependence on health care systems [51]. Older adults' well-documented preference for emotionally meaningful,

XSL•FO
RenderX

narrative information [33] appears to hold most strongly among those who already see themselves as agents in their own health. However, the nuanced interaction between external LOC levels and message framing calls for further exploration, especially to decipher the effects of different dimensions of LOC, such as chance and powerful others, on health communication strategies in diverse aging populations [80].

Contrary to expectations, we found no significant difference in the effectiveness of video versus photo-text modalities in enhancing framing effects on telemonitoring attitudes. This finding diverges from prior research advocating for video's superior persuasive capacity in health communication [47,48] and challenges simplistic notions about age-related preferences for specific media formats. The lack of observed modality impact suggests that individual differences, such as lifelong media habits, might dilute anticipated modality effects [81,82]. Consequently, the choice of modality for telemonitoring communications should be informed by considerations of content complexity and the cognitive and literacy levels of the target audience [83] while recognizing the diversity of media competencies among older adults.

Taken together, these 2 patterns suggest that what is communicated to older adults, and to whom, matters considerably more than the channel through which it is delivered. Through this lens, external LOC may be more than an individual psychological trait: external LOC may reflect in part how individuals perceive influences beyond their personal agency, from message framing to the behavior of trusted others within one's social network. Though in this study, the adapted SIM mainly focused on socioeconomic determinants, social capital mechanisms may operate alongside the socioeconomic pathway examined here. Individuals with higher external LOC may be more attuned to influences beyond their own agency when forming health-related judgments. This interpretation resonates with social capital and network-based perspectives in technology adoption research [84,85].

Digital Inclusion and Aging in Place: Implications for Policy and Practice

Our findings have important implications for inclusive digital health initiatives supporting older adults' aging in place. First, telemonitoring services should prioritize accessibility and affordability, particularly for socioeconomically disadvantaged older adults through free or subsidized community-based services, within walking distance could enhance adoption, aligning with age-friendly community principles [86]. Singapore's mature estates could serve as ideal locations for community-based services. Second, communications should position telemonitoring as complementing rather than replacing in-person care, emphasizing continuous monitoring benefits [16] while respecting established health care relationships. Messages could highlight how it complements and enhances doctor-patient relationships by providing continuous data that makes consultations more effective [87]. Third, the effectiveness of episodic versus thematic framing depends on external LOC. Practical guidance should center on message–audience matching rather than promoting a single "best" frame. Episodic narratives are likely to resonate more with older adults lower in external LOC, who tend to see health outcomes as within their own agency, while thematic messages emphasizing institutional involvement and statistical evidence may better suit those higher in external LOC. Communicators should therefore tailor framing to audience segment rather than defaulting to one universal approach, using simple language and visual aids to enhance accessibility. Finally, older adults should be viewed as active cocreators rather than passive recipients [88]. Their concerns reflect rational evaluations based on lifelong experiences. Involving them in design and implementation could enhance both solutions and acceptance.

Limitations and Future Research

While our study provides valuable insights, several limitations should be noted. First, this study adopted SIM adaptation, which did not incorporate the model's network-level mediating conditions, such as social capital and peer adoption behavior perception. Future research should directly examine whether social network characteristics or social capital account for this pattern, and integrate validated social network measures (eg, network size, tie strength, and peer adoption status). Second, our qualitative sample size was modest, and the convenience sampling approach may limit generalizability. Future research should expand to include more ethnically diverse participants and explore additional framing approaches. Third, our measure of external LOC would benefit from greater dimensionality in future studies. Longitudinal research examining actual telemonitoring adoption behavior would provide valuable insights. Furthermore, statistical power analysis indicates that the study lacked sufficient power to detect small effects (f^2<.07); therefore, the nonsignificant main effects of presentation modality and gender should be interpreted with caution and not viewed as definitive evidence of the absence of an association.

Conclusion

Our research advances understanding of how socioeconomic factors and communication strategies shape telemonitoring adoption among older adults in Singapore. The identified barriers and facilitators illuminate the interplay between aging experiences, health beliefs, and structural inequalities, rather than simple technological resistance. The message–audience matching pattern that episodic framing benefiting participants lower in external LOC, and thematic framing benefiting those higher in external LOC highlights the importance of psychologically tailored communication approaches that recognize older adults' diverse perspectives. This work contributes to both social gerontology and health communication by demonstrating how targeted interventions can bridge digital divides in aging societies. As Singapore and similar nations develop technological solutions to support aging in place, ensuring these innovations enhance rather than diminish equity, autonomy, and well-being for diverse older populations remains paramount, ultimately helping realize the goal of inclusive, technology-supported aging for all.

XSL•FO
RenderX

Acknowledgments

The authors sincerely thank all older adult participants. The authors also thank Yichi Zhang, Teng Fong Ng, Ren Zhong Neo, Weili Poo, and staff from FullHealth for their collaboration and support in the project. The work described in this paper was partially supported by a fellowship award from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU JRFS2526-9061038). The authors used Claude Sonnet 4.6 (Anthropic) to assist with language (grammar, typographical errors, and editing) and clarity improvement in the initial draft, under full human supervision. The generative AI is not listed as an author and does not bear responsibility for the final outcomes.

Funding

This work was supported by the National Youth Council Singapore; the Ministry of Education (MOE), Singapore, AcRF Tier 1 grant (Grant number: RG38/22); and the Research Grants Council of the Hong Kong Special Administrative Region, China, under the General Research Fund (Project No. 9043936). The funders had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated and analyzed during the current study are not publicly available due to privacy and confidentiality restrictions. The research data contain potentially identifiable personal information from older participants. Reasonable requests for access to anonymized or aggregated data that support the conclusions of this article may be directed to the corresponding author.

Authors' Contributions

Conceptualization: EWJL
Data collection: APTQ, GHYH, PSH, MEZO
Formal analysis: HB
Funding acquisition: EWJL
Investigation: APTQ, GHYH, PSH, MEZO
Supervision: EWJL
Validation: DW
Writing – original draft: APTQ, HB, DW
Writing – review & editing: DW, HB, EWJL.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Stimuli posters.
[DOCX File, 5018 KB - aging_v9i1e89817_app1.docx]

Multimedia Appendix 2
Video stimuli: episodic × video experimental stimuli (English).
[MP4 File, 47799 KB - aging_v9i1e89817_app2.mp4]

Multimedia Appendix 3
Video stimuli: episodic × video experimental stimuli (Chinese).
[MP4 File, 55135 KB - aging_v9i1e89817_app3.mp4]

Multimedia Appendix 4
Video stimuli: thematic × video experimental stimuli (English).
[MP4 File, 44927 KB - aging_v9i1e89817_app4.mp4]

Multimedia Appendix 5
Video stimuli: thematic × video experimental stimuli (Chinese).
[MP4 File, 53905 KB - aging_v9i1e89817_app5.mp4]

Multimedia Appendix 6
Full Estimated Model (Hayes' PROCESS Macro, Model 9).

XSL•FO
RenderX

[DOCX File, 19 KB - aging_v9i1e89817_app6.docx]

References

1. Huygens MWJ, Voogdt-Pruis HR, Wouters M, et al. The uptake and use of telemonitoring in chronic care between 2014 and 2019: nationwide survey among patients and health care professionals in the Netherlands. J Med Internet Res 2021 May 3;23(5):e24908. [doi: 10.2196/24908] [Medline: 33938808]
2. Wiles JL, Leibing A, Guberman N, Reeve J, Allen RES. The meaning of “aging in place” to older people. Gerontologist 2012 Jun;52(3):357-366. [doi: 10.1093/geront/gnr098] [Medline: 21983126]
3. Lindeman DA, Kim KK, Gladstone C, Apesoa-Varano EC. Technology and caregiving: emerging interventions and directions for research. Gerontologist 2020 Feb 14;60(Suppl 1):S41-S49. [doi: 10.1093/geront/gnz178] [Medline: 32057082]
4. Viswanath K, Kreuter MW. Health disparities, communication inequalities, and eHealth. Am J Prev Med 2007 May;32(5 Suppl):S131-S133. [doi: 10.1016/j.amepre.2007.02.012] [Medline: 17466818]
5. Yu E, Hagens S. Socioeconomic disparities in the demand for and use of virtual visits among senior adults during the COVID-19 pandemic: cross-sectional study. JMIR Aging 2022 Mar 22;5(1):e35221. [doi: 10.2196/35221] [Medline: 35134746]
6. Cortelyou-Ward K, Atkins DN, Noblin A, Rotarius T, White P, Carey C. Navigating the digital divide: barriers to telehealth in rural areas. J Health Care Poor Underserved 2020;31(4):1546-1556. [doi: 10.1353/hpu.2020.0116] [Medline: 33416736]
7. Cacciatore MA, Scheufele DA, Iyengar S. The end of framing as we know it … and the future of media effects. Mass Commun Soc 2016 Jan 2;19(1):7-23. [doi: 10.1080/15205436.2015.1068811]
8. Bullock OM, Shulman HC. Utilizing framing theory to design more effective health messages about tanning behavior among college women. Commun Stud 2021 May 4;72(3):319-332. [doi: 10.1080/10510974.2021.1899007]
9. Schulz R, Wahl HW, Matthews JT, De Vito Dabbs A, Beach SR, Czaja SJ. Advancing the aging and technology agenda in Gerontology. Gerontologist 2015 Oct;55(5):724-734. [doi: 10.1093/geront/gnu071] [Medline: 25165042]
10. Peine A, Neven L. From intervention to co-constitution: new directions in theorizing about aging and technology. Gerontologist 2019 Jan 9;59(1):15-21. [doi: 10.1093/geront/gny050] [Medline: 29850812]
11. Viswanath K, Ramanadhan S, Kontos EZ. Mass media. In: Macrosocial Determinants of Population Health: Springer; 2007:275-294. [doi: 10.1007/978-0-387-70812-6_13]
12. Estes CL. Social Policy & Aging: A Critical Perspective: SAGE Publications; 2001. [doi: 10.4135/9781452232676]
13. Phillipson C. The political economy of longevity: developing new forms of solidarity for later life. Sociol Q 2015 Jan;56(1):80-100. [doi: 10.1111/tsq.12082] [Medline: 25678722]
14. Robinson L, Cotten SR, Ono H, et al. Digital inequalities and why they matter. Inf Commun Soc 2015 May 4;18(5):569-582. [doi: 10.1080/1369118X.2015.1012532]
15. Friemel TN. The digital divide has grown old: determinants of a digital divide among seniors. New Media Soc 2016 Feb;18(2):313-331. [doi: 10.1177/1461444814538648]
16. Yang K, Li Y, Qi H. Determinants of and willingness to use and pay for digital health technologies among the urban elderly in Hangzhou, China. Risk Manag Healthc Policy 2023;16:463-478. [doi: 10.2147/RMHP.S393767] [Medline: 37007299]
17. Jung SO, Son YH, Choi E. E-health literacy in older adults: an evolutionary concept analysis. BMC Med Inform Decis Mak 2022 Jan 31;22(1):28. [doi: 10.1186/s12911-022-01761-5] [Medline: 35101005]
18. Barakat C, Konstantinidis T. A review of the relationship between socioeconomic status change and health. Int J Environ Res Public Health 2023 Jun 29;20(13):6249. [doi: 10.3390/ijerph20136249] [Medline: 37444097]
19. Darin-Mattsson A, Fors S, Kåreholt I. Different indicators of socioeconomic status and their relative importance as determinants of health in old age. Int J Equity Health 2017 Sep 26;16(1):173. [doi: 10.1186/s12939-017-0670-3] [Medline: 28950875]
20. Khan S, Webster S, Puxty J, Robertson M. Key challenges and barriers to digital literacy for older adults: scoping review. JMIR Aging 2026 Mar 16;9:e80647. [doi: 10.2196/80647] [Medline: 41838931]
21. Jaana M, Lévesque Ryan M, Tamim H, Riachy E, Paré G. Perspectives of older adults on assistive technology: qualitative study. JMIR Hum Factors 2025 Nov 12;12:e74214. [doi: 10.2196/74214] [Medline: 41222704]
22. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: perceived difficulties, experimentation behaviour and usability. Comput Human Behav 2013 Jul;29(4):1715-1724. [doi: 10.1016/j.chb.2013.02.006]
23. Iyengar S. Is Anyone Responsible? How Television Frames Political Issues: University of Chicago Press; 1991. [doi: 10.7208/chicago/9780226388533.001.0001]
24. Seifert A, Wahl HW. Young at heart and online? Subjective age and internet use in two Swiss survey studies. Educ Gerontol 2018 Mar 4;44(2-3):139-147. [doi: 10.1080/03601277.2018.1427495]
25. Kim SH, Willis LA. Talking about obesity: news framing of who is responsible for causing and fixing the problem. J Health Commun 2007 Jun;12(4):359-376. [doi: 10.1080/10810730701326051] [Medline: 17558788]
26. Koivula A, Marttila E, Koiranen I. Media framing and trust: a randomized survey experiment on attitudes towards the unvaccinated during the COVID-19 pandemic. Soc Sci Humanit Open 2024;10:101157. [doi: 10.1016/j.ssaho.2024.101157]

XSL•FO
RenderX

27. Temmann LJ, Wiedicke A, Schaller S, Scherr S, Reifegerste D. A systematic review of responsibility frames and their effects in the health context. J Health Commun 2021 Dec 2;26(12):828-838. [doi: 10.1080/10810730.2021.2020381] [Medline: 35050839]
28. Warmoth K, Tarrant M, Abraham C, Lang IA. Older adults' perceptions of ageing and their health and functioning: A systematic review of observational studies. Psychol Health Med 2016 Jul;21(5):531-550. [doi: 10.1080/13548506.2015.1096946] [Medline: 26527056]
29. Zillmann D, Brosius HB. Exemplification in Communication: The Influence of Case Reports on the Perception of Issues: Lawrence Erlbaum Associates; 2000. [doi: 10.4324/9781410604743]
30. Bigsby E, Bigman CA, Martinez Gonzalez A. Exemplification theory: A review and meta-analysis of exemplar messages. Ann Int Commun Assoc 2019 Oct 2;43(4):273-296. [doi: 10.1080/23808985.2019.1681903]
31. Krämer B, Peter C. Exemplification effects: a meta-analysis. Hum Commun Res 2020 Jul 1;46(2-3):192-221. [doi: 10.1093/hcr/hqz024]
32. Aarøe L, Petersen MB. Cognitive biases and communication strength in social networks: the case of episodic frames. Brit J Polit Sci 2020 Oct;50(4):1561-1581. [doi: 10.1017/S0007123418000273]
33. Guido G, Pichierri M, Rizzo C, Chieffi V, Moschis G. Information processing by elderly consumers: a five-decade review. JSM 2021 Mar 19;35(1):14-28. [doi: 10.1108/JSM-09-2019-0368]
34. Gross K. Framing persuasive appeals: episodic and thematic framing, emotional response, and policy opinion. Polit Psychol 2008 Apr;29(2):169-192 [FREE Full text] [doi: 10.1111/j.1467-9221.2008.00622.x]
35. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
36. Birati Y, Tzemah-Shahar R. Barriers to digital health adoption in older adults: scoping review informed by innovation resistance theory. J Med Internet Res 2026 Feb 2;28:e75591. [doi: 10.2196/75591] [Medline: 41627850]
37. Lee C, Coughlin JF. PERSPECTIVE: older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J Prod Innov Manag 2015 Sep;32(5):747-759. [doi: 10.1111/jpim.12176]
38. Adams C, Gringart E, Strobel N, Masterman P. Help-seeking for mental health problems among older adults with chronic disease: an application of the theory of planned behaviour. Aust J Psychol 2021 Oct 2;73(4):426-437. [doi: 10.1080/00049530.2021.1952850]
39. Rahman M, Das S, Hossain G, Tajrin T. Teenagers' behavioural intention towards wearable technologies and intention to recommend others: an empirical study in Bangladesh. JSTPM 2022 Feb 2;13(1):110-131. [doi: 10.1108/JSTPM-05-2020-0088]
40. Tarkar P. Predicting intentions to get the COVID-19 vaccine in India: an integration of theory of planned behaviour and health belief model. Int J Health Plan Manag 2023 Jan;38(1):214-238. [doi: 10.1002/hpm.3582] [Medline: 36164891]
41. Bracken CC, Dalessandro A. Modality and medium characteristics. the international encyclopedia of media effects. In: Rössler P, Hoffner CA, Zoonen L, editors. The International Encyclopedia of Media Effects: Wiley; 2017:1-12. [doi: 10.1002/9781118783764]
42. Paivio A. Mental Representations: A Dual Coding Approach: Oxford University Press; 1986. [doi: 10.1093/acprof:oso/9780195066661.001.0001]
43. Charness N, Boot WR. Technology, gaming, and social networking. In: Schaie KW, Willis SL, editors. Handbook of the Psychology of Aging, 8th edition: Academic Press; 2016:389-407. [doi: 10.1016/B978-0-12-411469-2.00020-0]
44. Bol N, van Weert JCM, de Haes HCJM, Loos EF, Smets EMA. The effect of modality and narration style on recall of online health information: results from a Web-based experiment. J Med Internet Res 2015 Apr 24;17(4):e104. [doi: 10.2196/jmir.4164] [Medline: 25910416]
45. Kang J, Lin CA. Effects of message framing and visual-fear appeals on smoker responses to antismoking ads. J Health Commun 2015;20(6):647-655. [doi: 10.1080/10810730.2015.1012242] [Medline: 25868549]
46. Kirkpatrick AW, Park M, Domgaard S, Zhao W, Steinberg C, Hsu YC. Vaccine videos and information sharing: the effects of framing, evidence type, and speaker expertise. J Health Commun 2021 Sep 2;26(9):608-617. [doi: 10.1080/10810730.2021.1983892] [Medline: 34596481]
47. de Graaf A, van den Putte B, Zebregs S, Lammers J, Neijens P. Smoking education for low-educated adolescents: comparing print and audiovisual messages. Health Promot Pract 2016 Nov;17(6):853-861. [doi: 10.1177/1524839916660525] [Medline: 27466267]
48. Occa A, Suggs LS. Communicating breast cancer screening with young women: an experimental test of didactic and narrative messages using video and infographics. J Health Commun 2016;21(1):1-11. [doi: 10.1080/10810730.2015.1018611] [Medline: 26147625]
49. Wallston KA, Wallston BS, DeVellis R. Development of the Multidimensional Health Locus of Control (MHLC) Scales. Health Educ Monogr 1978;6(2):160-170. [doi: 10.1177/109019817800600107] [Medline: 689890]
50. Baltes PB, Baltes MM. Psychological perspectives on successful aging. the model of selective optimization with compensation. In: Baltes PB, Baltes MM, editors. Successful Aging: Perspectives from the Behavioral Sciences: Cambridge University Press; 1990:1-34. [doi: 10.1017/CBO9780511665684.003]

XSL•FO
RenderX

51. Schönenberg A, Kobus C, Günther M, Umfermann L, Prell T. Perceptions of aging and control beliefs: a study on older patients' views of aging. Geriatrics (Basel) 2025 Nov 10;10(6):148. [doi: 10.3390/geriatrics10060148] [Medline: 41283459]
52. Cohen M, Azaiza F. Health-promoting behaviors and health locus of control from a multicultural perspective. Ethn Dis 2007;17(4):636-642. [Medline: 18072372]
53. Jacobs-Lawson JM, Waddell EL, Webb AK. Predictors of health locus of control in older adults. Curr Psychol 2011 Jun;30(2):173-183. [doi: 10.1007/s12144-011-9108-z]
54. Kong Y, Shen F. Impact of locus of control on health message effectiveness. Health Mark Q 2011 Oct;28(4):354-371. [doi: 10.1080/07359683.2011.623114] [Medline: 22054030]
55. Ceylan M, Hayran C. Message framing effects on individuals' social distancing and helping behavior during the COVID-19 pandemic. Front Psychol 2021;12:579164. [doi: 10.3389/fpsyg.2021.579164] [Medline: 33828501]
56. Nah S, Oh J, Atkinson L. Interactive health messages work better for those who feel less in control: the role of external health locus of control and risk perception. J Health Commun 2023 May 4;28(5):321-334. [doi: 10.1080/10810730.2023.2204308] [Medline: 37092292]
57. Boukes M. Episodic and thematic framing effects on the attribution of responsibility: the effects of personalized and contextualized news on perceptions of individual and political responsibility for causing the economic crisis. Int J Press Polit 2022 Apr;27(2):374-395. [doi: 10.1177/1940161220985241]
58. Coleman R, Thorson E, Wilkins L. Testing the effect of framing and sourcing in health news stories. J Health Commun 2011 Oct;16(9):941-954. [doi: 10.1080/10810730.2011.561918] [Medline: 21660828]
59. Key household income trends, 2022. Singapore Department of Statistics. 2023. URL: https://web.archive.org/web/20230210065453/https://www.singstat.gov.sg/-/media/files/news/press09022023.ashx [accessed 2026-06-10]
60. Census of population 2020 statistical release 1: demographic characteristics, education, language and religion. Singapore Department of Statistics. 2021. URL: https://www.singstat.gov.sg/publication-resources/singapore-census-of-population-2020-statistical-release-1-demographic-characteristics-education-language-and-religion [accessed 2026-10-07]
61. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
62. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [doi: 10.2307/2529310] [Medline: 843571]
63. Wallston BS, Wallston KA, Kaplan GD, Maides SA. Development and validation of the Health Locus of Control (HLC) Scale. J Consult Clin Psychol 1976;44(4):580-585. [doi: 10.1037/0022-006X.44.4.580] [Medline: 939841]
64. Dabholkar PA, Bagozzi RP. An attitudinal model of technology-based self-service: moderating effects of consumer traits and situational factors. J Acad Mark Sci 2002 Jul 1;30(3):184-201. [doi: 10.1177/00970302030003001]
65. Boontarig W, Chutimaskul W, Chongsuphajaisiddhi V, Papasratorn B. Factors influencing the Thai elderly intention to use smartphone for e-Health services. 2012 IEEE Symposium on Humanities, Science and Engineering Research (SHUSER) 2012:479-483. [doi: 10.1109/SHUSER.2012.6268881]
66. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach, 2nd edition: Guilford Press; 2018. URL: https://search.worldcat.org/zh-cn/title/1007923866 [accessed 2026-10-07]
67. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Lawrence Erlbaum Associates; 1988. [doi: 10.4324/9780203771587]
68. Ferguson HJ, Brunsdon VEA, Bradford EEF. The developmental trajectories of executive function from adolescence to old age. Sci Rep 2021 Jan 14;11(1):1382. [doi: 10.1038/s41598-020-80866-1] [Medline: 33446798]
69. Ferguson LA. Implementing a video education program to improve health literacy. J Nurse Pract 2012 Sep;8(8):e17-e22. [doi: 10.1016/j.nurpra.2012.07.025]
70. Strazdins L, Griffin AL, Broom DH, et al. Time scarcity: another health inequality? Environ Plan A 2011 Mar;43(3):545-559. [doi: 10.1068/a4360]
71. Dannefer D. Cumulative advantage/disadvantage and the life course: cross-fertilizing age and social science theory. J Gerontol B Psychol Sci Soc Sci 2003 Nov;58(6):S327-S337. [doi: 10.1093/geronb/58.6.s327] [Medline: 14614120]
72. Rowe JW, Kahn RL. Successful aging. Gerontologist 1997 Aug;37(4):433-440. [doi: 10.1093/geront/37.4.433] [Medline: 9279031]
73. Hechinger M, Hentschel D, Aumer C, Rester C. A conceptual model of experiences with digital technologies in aging in place: qualitative systematic review and meta-synthesis. JMIR Aging 2022 Sep 9;5(3):e34872. [doi: 10.2196/34872] [Medline: 36083625]
74. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Dec;21(1):1556. [doi: 10.1186/s12889-021-11623-w]
75. Archer JA, Lee A, Qiu A, Chen SHA. Working memory, age and education: a lifespan fMRI study. PLOS ONE 2018;13(3):e0194878. [doi: 10.1371/journal.pone.0194878] [Medline: 29584781]
76. Sörman DE, Hansson P, Pritschke I, Ljungberg JK. Complexity of primary lifetime occupation and cognitive processing. Front Psychol 2019;10:1861. [doi: 10.3389/fpsyg.2019.01861] [Medline: 31496970]

XSL•FO
RenderX

77. Mozafari S, Yang A, Talaei-Khoei J. Health locus of control and medical behavioral interventions: systematic review and recommendations. Interact J Med Res 2024 Oct 10;13(1):e52287. [doi: 10.2196/52287] [Medline: 39388686]
78. Shen F, Lee SY, Sipes C, Hu F. Effects of media framing of obesity among adolescents. Commun Res Rep 2012 Jan;29(1):26-33. [doi: 10.1080/08824096.2011.639910]
79. Spence LK. Episodic frames, HIV/AIDS, and African American public opinion. Polit Res Q 2010 Jun;63(2):257-268. [doi: 10.1177/1065912908330727]
80. Atilola O, Stevanovic D, Moreira P, et al. External locus-of-control partially mediates the association between cumulative trauma exposure and posttraumatic stress symptoms among adolescents from diverse background. Anxiety Stress Coping 2021 Nov;34(6):626-644. [doi: 10.1080/10615806.2021.1891224] [Medline: 33650438]
81. Zhu Z, Liu S, Zhang R. Examining the persuasive effects of health communication in short videos: systematic review. J Med Internet Res 2023 Oct 13;25:e48508. [doi: 10.2196/48508] [Medline: 37831488]
82. Greenberg K, Zheng R, Gardner M, Orr M. Individual differences in visuospatial working memory capacity influence the modality effect. J Comput Assist Learn 2021 Jun;37(3):735-744. [doi: 10.1111/jcal.12519]
83. Merkt M, Weigand S, Heier A, Schwan S. Learning with videos vs. learning with print: the role of interactive features. Learn Instr 2011 Apr;21(6):687-704. [doi: 10.1016/j.learninstruc.2011.03.004]
84. Choi G, Chung H. Applying the technology acceptance model to social networking sites (SNS): impact of subjective norm and social capital on the acceptance of SNS. Int J Hum Comput Interact 2013 Oct 3;29(10):619-628. [doi: 10.1080/10447318.2012.756333]
85. Hepburn J, Williams L, McCann L. Barriers to and facilitators of digital health technology adoption among older adults with chronic diseases: updated systematic review. JMIR Aging 2025 Sep 11;8:e80000. [doi: 10.2196/80000] [Medline: 40934502]
86. Lui CW, Everingham JA, Warburton J, Cuthill M, Bartlett H. What makes a community age-friendly: a review of international literature. Australas J Ageing 2009 Sep;28(3):116-121. [doi: 10.1111/j.1741-6612.2009.00355.x] [Medline: 19845650]
87. Boyne JJJ, Vrijhoef HJM. Implementing telemonitoring in heart failure care: barriers from the perspectives of patients, healthcare professionals and healthcare organizations. Curr Heart Fail Rep 2013 Sep;10(3):254-261. [doi: 10.1007/s11897-013-0140-1] [Medline: 23666901]
88. Peine A, Marshall B, Martin W, Neven L. Socio-Gerontechnology: Interdisciplinary Critical Studies of Ageing and Technology, 1st edition: Routledge; 2021. [doi: 10.4324/9780429278266]

Abbreviations

CFI: comparative fit index
HLCS: Health Locus of Control Scale
IRB: Institutional Review Board
LOC: locus of control
RMSEA: root-mean-square error of approximation
SES: socioeconomic status
SIM: Structural Influence Model
SPSS: Statistical Package for the Social Sciences
TPB: Theory of Planned Behavior

Edited by M Antoniou; submitted 17.Dec.2025; peer-reviewed by B Pavlenko, F Lu; revised version received 27.Aug.2026; accepted 27.Sep.2026; published 27.Oct.2026.

Please cite as:
Wang D, Bao H, Qua APT, Hoong GYH, He PS, Ong ESM, Lee EWJ
Telemonitoring Adoption Among Older Adults: Sequential Mixed Methods Study of Message Framing and Structural Influence
JMIR Aging 2026;9:e89817
URL: https://aging.jmir.org/2026/1/e89817
doi:10.2196/89817

© Dandan Wang, Huanyu Bao, Amber Pei Tze Qua, Gaelmaine Yu Heng Hoong, Pei Shan He, Ern Sze Maegan Ong, Edmund Wei Jian Lee. Originally published in JMIR Aging (https://aging.jmir.org), 27.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability of the Coach-Supported Dementia Prevention App ENHANCE (Tailored Intervention for Brain Health and Cognitive Enrichment) in Older Adults: 1-Week Mixed Methods Study

Tsz Kiu Clare Yu[1], MSc; Gill Livingston[1,2], MBchB, MD; Richard Boczko[3], PhD; Kealan Forristal[4], MSc; James Jamison[1], PhD; Carl Leckstein[5], MSc; Vrushti Mehta[1], MSc; Hee Kyung Park[1,6], MD, PhD; Aneesha Singh[7], PhD; Andrew Sommerlad[1,2], MBBS, MRCPsych, PhD; Sergi G Costafreda[1,2], MD, PhD

[1]Division of Psychiatry, University College London, Maple House, 149 Tottenham Court Road, London, United Kingdom

[2]North London NHS Foundation Trust, London, United Kingdom

[3]Faculty of Health Sciences, University of Hull, Hull, England, United Kingdom

[4]Social, Genetic & Developmental Psychiatry Centre, King's College London, London, England, United Kingdom

[5]Division of Psychology and Language Sciences, University College London, London, England, United Kingdom

[6]Department of Neurology, Samsung Medical Center, Sungkyunkwan University, Seoul, Republic of Korea

[7]UCL Interaction Centre, University College London, London, England, United Kingdom

Corresponding Author:
Tsz Kiu Clare Yu, MSc
Division of Psychiatry, University College London, Maple House, 149 Tottenham Court Road, London, United Kingdom

Abstract

Background: Multidomain dementia-prevention interventions delivered via apps have the potential to reach large populations. However, existing trials have tended to recruit more socioeconomically advantaged participants, raising concerns that the resulting interventions may be less usable for older adults from minority ethnic, lower educational, or lower socioeconomic backgrounds, who are at higher risk of dementia. The Tailored Intervention for Brain Health and Cognitive Enrichment (ENHANCE) app was designed to address this by prioritizing accessibility and engagement across diverse user groups.

Objective: This study evaluated the usability and user experience of the ENHANCE prototype during a 1-week at-home supported-use test and explored factors influencing use and engagement among older adults.

Methods: We purposively recruited adults aged 60 - 80 years without dementia for a 1-week mixed methods usability evaluation through community settings, including groups underrepresented in dementia-prevention trials. Participants had at least 1 of 10 prespecified dementia risk factors, attended a face-to-face onboarding session with a coach, used the app at home for 7 days with ongoing coach support, and completed a posttest interview and satisfaction survey. We analyzed quantitative data, including app usage metrics and survey responses descriptively, and used reflexive thematic analysis of qualitative data from onboarding sessions, posttest interviews, coaching calls, and in-app messages.

Results: Ten participants participated in the study. The mean age was 68 (SD 6) years, and 7 were females. Participants represented a wide range of deprivation (Index of Multiple Deprivation deciles: mean 4, SD 2, range 1 - 8), with 6 from ethnic minority backgrounds. All met prespecified minimum-use targets (watching a module video, completing a check-in, and playing assigned games at least once). Many demonstrated additional voluntary engagement: 5 rewatched the risk factor video, 7 used the in-app messaging feature, and among the 5 participants with the hypertension module, blood pressure was logged on an average of 5 out of 7 days. Survey responses indicated high satisfaction, perceived usefulness, and ease of use; 9 participants intended to continue using the app and would recommend it to peers. Qualitative analysis identified engagement facilitators, including rewarding game design, familiar interfaces, appropriately challenging gameplay, consistent virtual rewards, trusted expert information combined with peer stories, and coach support. Barriers included unclear visual cues, insufficient accommodation of motor or sensory impairments, and visual discomfort in some games.

Conclusions: Older adults found the ENHANCE prototype usable, acceptable, and engaging over 1 week. Human coaching, inclusive design, and integration of expert and peer narratives were highlighted as key engagement drivers. These findings support further feasibility testing to examine longer-term engagement and provide design insights for more inclusive digital health interventions.

(JMIR Aging 2026;9:e92800) doi:10.2196/92800

XSL•FO
RenderX

KEYWORDS

dementia prevention; older adults; mHealth; underserved populations; usability testing; user experience; cognitive training

Introduction

The number of people with dementia globally is expected to reach 152 million by 2050, posing a major challenge for individuals, families, and health and social care systems [1]. As a result, prevention is a public health priority. The Lancet Commission on dementia prevention, intervention, and care estimates that 45% of cases could be prevented by eliminating 14 potentially modifiable risk factors, including less education, hypertension, diabetes, physical inactivity, smoking, obesity, depression, hearing loss, excessive alcohol, and social isolation [2]. Multidomain interventions addressing several dementia risk factors simultaneously can provide cognitive benefits for older adults at risk, though effects are typically modest [3,4]. For example, a network meta-analysis of over 100 RCTs found small improvements in cognition [3], while the large "Maintain Your Brain" randomized controlled trials (n>6000) reported modest gains over 3 years in the online multidomain lifestyle intervention [4,5].

People with lower socioeconomic status and those from ethnic minority groups are at greater risk of developing dementia likely due to higher risk factor prevalence [6-10]. Despite being at higher risk, these groups are often inadequately represented or reported in prevention trials [11-13]. In the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) cognitive training trial (n=2832), participants were disproportionately highly educated (88.6% high school diploma+ vs 67% in the US population aged 65 y and older), healthier (higher health scores in health survey), and less hospitalized than national norms [14]. Similarly, in the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) trial (a 2-year multidomain intervention; n=1260), participants had higher education levels and were younger or healthier than nonscreened eligible individuals from the same population [15]. These patterns underscore the need for more inclusive dementia prevention strategies that target underserved communities. Given the elevated baseline risk in these populations, even modest improvements in risk factor profiles could yield substantial reductions in dementia incidence at the population level.

In this paper, we use the term "underserved" to refer to groups who may experience higher dementia risk but face barriers to accessing and engaging with prevention interventions and who are often underrepresented in dementia prevention research. This includes (but is not limited to) older adults living in more socioeconomically deprived areas, with lower levels of education, and those from ethnic minority backgrounds.

Digital health interventions, particularly app-based approaches, offer a promising approach to more inclusive dementia prevention as they provide remote access to content, potentially benefiting those with less mobility, in rural areas, or facing financial barriers to in-person programs [16]. Apps can also support multiple languages, improving accessibility for ethnic minority older adults [17].

App use is increasingly common among older adults, including those from low-income groups in high-income countries. In the United Kingdom (2021 - 2022), 81% of adults aged older than 65 years used the internet, with 49% of these older adults accessing it via tablet [18]. Moreover, 67% of individuals from lower socioeconomic backgrounds across all age groups owned a smartphone—a figure likely to have risen due to increased digital adoption during the COVID-19 pandemic [19]. Importantly, a recent study suggests that device access can translate into engagement; individuals from underserved populations use health apps at rates comparable with more advantaged groups and derive similar benefits [20]. However, usability challenges and lower digital confidence remain significant barriers to long-term engagement, especially among this population [21,22]. Addressing these through inclusive design is essential for long-term engagement and use.

To address these gaps, we developed Tailored Intervention for Brain Health and Cognitive Enrichment (ENHANCE), a theory-informed, coach-supported, app-based intervention targeting modifiable dementia risk factors in older adults. A central aim of ENHANCE was to create an intervention that is accessible and usable across the whole older population, including the underserved people. The development process was centered on co-design with community-recruited participants, with a particular focus on individuals underrepresented in dementia prevention research, including those with lower educational attainment, from ethnic minority backgrounds, and with lower socioeconomic status. We also collaborated closely with patient and public involvement representatives and experts in human-computer interaction. Usability testing and iterative refinement were conducted specifically with older adults with lower education and socioeconomic status (refer to our previous work for detailed description) [23].

This study aims to evaluate user experience with the ENHANCE prototype over a 1-week period, identifying key facilitators and barriers to use.

This study addresses the following research questions:

1. What are the experiences of older adults using ENHANCE over 1 week?
2. What factors facilitate or hinder at-home, 1-week use and engagement with the intervention?

Methods

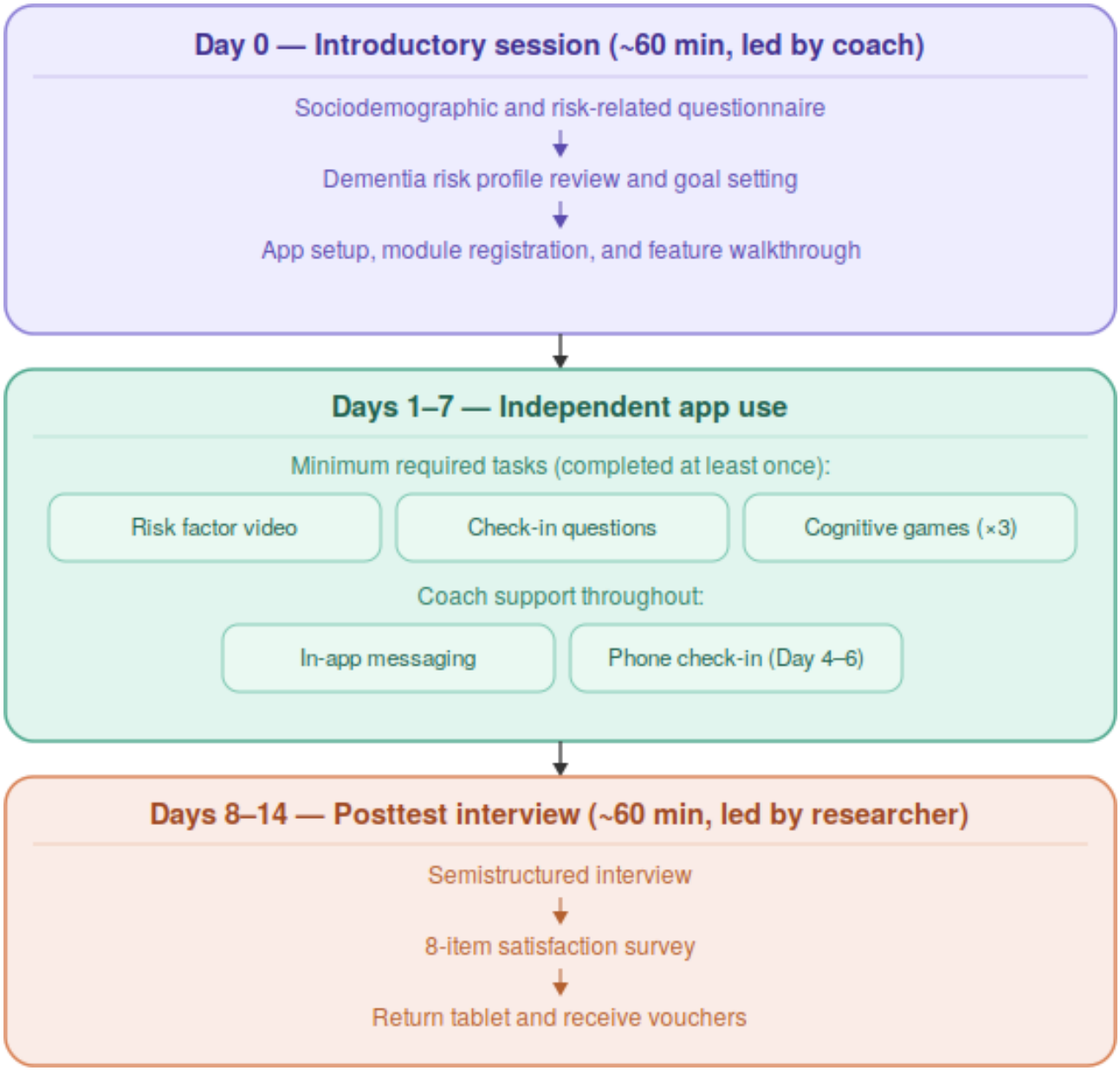
Overview

We conducted a 1-week usability test of the ENHANCE app prototype with older adults, including individuals from underserved backgrounds in the United Kingdom, using a mixed methods approach. Recruitment and data collection took place in London between November 2024 and February 2025. An overview of the study procedure is presented in Figure 1.

XSL•FO
RenderX

This project is part of the ENHANCE study, which will evaluate the feasibility, acceptability, and effectiveness of the ENHANCE intervention in reducing dementia risk among healthy older adults from underserved backgrounds.

Figure 1. Study's flowchart.



Ethical Considerations

The Institutional Review Board of University College London (ID 24235/001) gave ethical approval on June 19, 2023.

Before enrollment, all participants received a participant information sheet detailing the study purpose, investigator names, study procedures, and the expected time commitment (60-min introductory session, 7 days of independent app use, 60-min posttest interview, and a 5-min satisfaction survey). Participants were informed of their right to withdraw at any time without consequence and gave written informed consent before any data collection. All research participation was voluntary.

We pseudonymized collected data using alphanumeric codes. The code list and consent forms were stored separately from the anonymized data in locked cabinets (paper records) and on a password-protected, encrypted institutional drive (electronic records), accessible only to the core research team. Data are retained for 1 year following completion of the relevant analysis.

Eligibility Criteria and Recruitment

Participants were older adults residing in the United Kingdom. Eligibility criteria were as follows:

- Participants were included if they (1) were aged between 60 and 80 years; (2) had at least 1 dementia risk factor targeted by the ENHANCE intervention (<8 y of full-time education, diagnosed diabetes, hypertension, BMI>30kg/m^2, current smoker, >21 units of alcohol/week, <2.5 h of exercise/week, depressive symptoms identified using 2 adapted screening questions as per Patient Health Questionnaire−2 [PHQ-2] [24], social isolation assessed

XSL•FO
RenderX

with 1 adapted question based on the Lubben Social Network Scale [25], and self-perceived hearing impairment); (3) were able to read and communicate in basic English; and (4) had the mental capacity to provide informed consent.
- Participants were excluded if they reported a dementia diagnosis or any disability that would significantly impact tablet use, both assessed via self-reported screening questions (Multimedia Appendix 1).

We predetermined a sample size of 10 participants, based on guidance suggesting that this number can identify approximately 95% of major usability issues [26]. We did not set eligibility criteria based on education, ethnicity, or area deprivation; instead, we used a recruitment strategy designed to reach underserved groups. Participants were drawn from a database of over 100 individuals (aged 60 years or older) recruited through community settings serving underserved populations in London, including food banks, neighborhood shopping malls, religious groups, and community organizations supporting low-income and ethnic minority communities. No centers serving people with dementia were included. Participants had previously taken part in ENHANCE coproduction activities [23] and had provided consent to be contacted for future research.

To inform purposive sampling, we extracted participants' age, ethnicity, and Index of Multiple Deprivation (IMD) scores from self-reported data collected in the previous study [23]. The IMD is a UK-based measure of relative deprivation derived from participants' postcodes and divided into deciles, with lower scores indicating higher levels of deprivation [27]. This information was used to purposefully recruit individuals from more deprived areas and non-White backgrounds.

TKCY, a PhD student with 6 years of dementia research experience, recruited potential participants by phone, email, or text. She informed them that no previous digital skills or home access to a tablet were required. Interested individuals completed a brief telephone screening for eligibility, covering age, dementia diagnosis status, as well as dementia health and lifestyle risk factors (Multimedia Appendix 1).

The ENHANCE Intervention

ENHANCE is a coach-supported, app-based intervention designed to help older adults aged from 60 - 80 years, including those from the underserved background, reduce dementia risk by addressing 10 modifiable risk factors—hypertension, diabetes, physical inactivity, obesity, depression, smoking, excessive alcohol use, social isolation, hearing loss, and low education—through lifestyle behavior change and cognitive engagement.

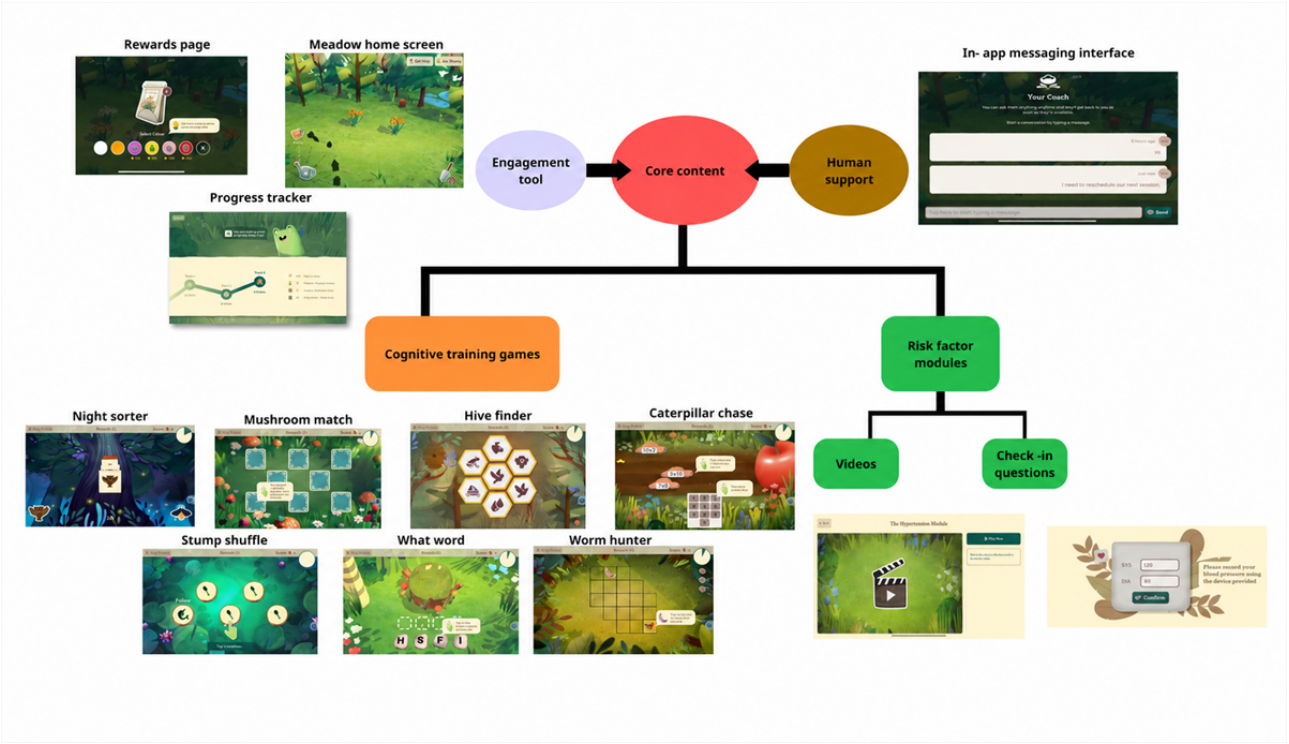
The app's features fall into 3 categories:

- Core content: 7 cognitive training games; modules for 9 other risks which include educational videos and brief weekly check-in questions to support motivation and self-monitoring for lifestyle change tailored to participants' selected risk factors.
- Engagement tools: A unifying "meadow" home screen and rewards page where participants can use virtual seeds to plant flowers in their meadow for playing games, watching videos, and answering check-ins; and a progress tracker showing weekly stars earned from completed activities.
- Human support: Scheduled coaching sessions (45 min in-person onboarding session and fortnightly remote short follow-ups); a coach-only dashboard to monitor participants' activities; in-app messaging with the coach; and optional family involvement.

All participants, regardless of their educational attainment, were given access to the cognitive training games in the ENHANCE intervention as they are rewarding in themselves. For the remaining 9 risk factors, each participant selected 1 risk factor module to focus on during the 1-week testing period. Figure 2 illustrates screenshots from the ENHANCE app. Further details about the app design are reported elsewhere [23].

XSL•FO
RenderX

Figure 2. Screenshots of the ENHANCE app.



Procedures

Overview

The 1-week usability testing consisted of three components: (1) an introductory session on day 0, (2) independent app use from days 1 to 7, and (3) a posttest feedback interview conducted between days 8 and 14.

Introductory Session (Day 0)

Each participant attended a 60-minute introductory session about the app either at the university or at their home, choosing either. A coach (KF, CL, or VM) with an MSc in Psychology who contributed to the coaching manual development led the session, while the researcher (TKCY), a member of the research team who developed the ENHANCE app, observed participants' engagement with both the app and the coach.

The session began with participants completing a brief sociodemographic and risk-related questionnaire (Multimedia Appendix 2), by answering about age, post code (to derive the IMD to measure socioeconomic status), educational level, and key lifestyle and health factors relevant to dementia risk. Based on responses, the coach reviewed each participant's dementia risk profile, discussed the identified risk factors and which one they wished to target during the 1-week testing, and supported the creation of a SMART (specific, measurable, achievable, relevant, time-bound) goal [28] for positive behavior change related to this risk factor. The coach then guided participants through app setup, registration for the selected module, and key features, encouraging questions throughout.

At the end of the session, participants were given a tablet with the app installed and a SIM card and reminded to complete at least 1 video, 1 check-in tool, and each of the 3 cognitive training games at least once during the week; other activities (eg, additional gameplays and video rewatch) were optional.

The session was audio-recorded, and TKCY took notes on participants' interactions with the app, the coach, and any observed difficulties or misunderstandings.

App Testing Week (Day 1-Day 7)

Participants used the app independently for 7 consecutive days. On the app, each participant received:

- 3 randomly assigned cognitive training games from the games on the app. These covered different areas of cognition and could each be played up to 10 times per week after first completion
- 1 video from their selected module (rewatchable without limits)
- Risk factor–specific weekly check-in questions (eg, number of days exercise was completed in the previous wk, with daily logging enabled only for the blood pressure module).

Minimum required use was defined as completion of three core tasks during the testing period: (1) watching the assigned risk factor video at least once, (2) responding to the check-in question at least once, and (3) playing each of the 3 assigned games at least once.

Coach support was available throughout the testing week. Participants could message their coach via the in-app messaging feature. They also received a phone check-in on days 4 - 6. The check-in call was audio-recorded for further analysis.

Posttest Feedback Interview (Day 8-Day 14)

TKCY conducted and audio-recorded a posttest interview (approximately 60 min) at the same venue as the introductory session, within 1 week after app use, to gather feedback on

XSL•FO
RenderX

participants' experiences with the ENHANCE app and coaching. No previous relationship existed between TKCY and the participants before study commencement. Participants were informed that TKCY was a PhD student involved in the project as part of her doctoral study.

TKCY used a semistructured guide co-developed by the core research team and patient and public involvement representatives (Multimedia Appendix 3). The interview covered participants' general impressions, a section-by-section app review for detailed feedback, reflections on the coaching experience, and any potential barriers. TKCY reviewed app usage data in advance to tailor questions, particularly around unused features.

Immediately following the posttest interview, participants completed an 8-item, 1-page satisfaction survey taking approximately 5 minutes on a laptop preloaded with a direct Qualtrics link. To minimize social desirability bias, participants completed the survey independently, with TKCY briefly demonstrating the navigation and submission and then withdrew. All 8 items were mandatory, and Qualtrics was configured to prevent submission unless all items were completed. Participants then saw a review screen summarizing their responses, allowing them to amend answers if needed.

TKCY developed the satisfaction survey from the Client Satisfaction Questionnaire (CSQ-8) [29]—a validated measure of programs and services satisfaction. She adapted it for ENHANCE as no existing tool adequately captured the app-based gameplay, dedicated risk factor modules, and one-to-one health coaching. SGC and a member of the Patient and Public Involvement team provided feedback on wording, clarity, and coverage of the key ENHANCE components, and TKCY iteratively refined the survey. The final survey assessed app satisfaction, game enjoyment, perceived behavior-change potential of the risk factor modules, and coaching satisfaction using a 5-point Likert scale (Multimedia Appendix 4). Participants then returned the tablet and received £80 (£1=US $1.34) vouchers for their time (£25 for onboarding, £25 for postsession feedback, and £30 for app use).

Analyses

All quantitative data were analyzed using SPSS (version 25; IBM), and qualitative data were analyzed using NVivo 14 (Lumivero).

Demographic Data

Demographic data collected via a brief questionnaire administered during the introductory session were analyzed descriptively, with means, SDs, and range reported for continuous variables and counts for categorical variables.

Quantitative Data

App Use (Minimum and Additional App Use)

App use data were extracted from the backend following the app testing week and analyzed to (1) assess whether participants met prespecified minimum-use targets during the testing period and (2) characterize additional patterns of engagement beyond these targets.

Additional app use was assessed using 4 indicators. First, the number of times the risk factor video was rewatched was obtained from backend logs. Second, only participants assigned to the hypertension module had access to respond to the check-in more than once during the study week. Additional responses to the check-in question=(total number of check-in days÷total days of app use)×7. Third, additional gameplay beyond the minimum requirement=(total number of plays÷total days of app use)×7. The values were then averaged across all participants to yield the mean weekly gameplay frequency for each game. Finally, use of the in-app messaging feature at least once via backend data. Participants were coded as having used the feature if they initiated or responded to at least 1 message with the coach. Those with no interaction were coded as nonusers.

Posttest App Satisfaction Survey

Responses to 8 posttest survey questions were analyzed descriptively and summarized as counts for each response option.

Qualitative Data

All audio recordings were transcribed using data regulation-compliant software (Happy Scribe [Happy Scribe Ltd]) and reviewed by TKCY for accuracy and then imported into NVivo 14. TKCY analyzed the data using 6-step reflexive thematic analysis by Braun and Clarke [30,31]. She also maintained a reflexive journal to consider how her background might have influenced the interviews and analysis. She used the COM-B (capability, opportunity, motivation—behavior) model as a guiding framework to identify facilitators and barriers influencing app engagement [32] to provide a comprehensive overview of internal and external factors influencing behavior. It has been widely applied in studies understanding app use and user engagement [32-35] and aligns well with qualitative interview-based data analysis.

- Capability refers to an individual's ability to carry out the behavior, which includes psychological factors (such as knowledge, memory, and cognitive skills) and physical abilities (such as physical strength or motor skills).
- Opportunity encompasses the external circumstances that facilitate or encourage the behavior. These may be physical (eg, time availability and resources) or social (eg, social support or peer influence).
- Motivation involves the internal processes that initiate or maintain behavior. This includes reflective motivation (like conscious goals, intentions, and evaluations) and automatic motivation (such as emotions, desires, and habits).

She analyzed each qualitative data source separately, generating initial codes and developing them into candidate themes with theme refinement carried out through 5 biweekly meetings between TKCY, A Singh, GL, SGC, and HKP. As consistent patterns emerged, findings from each qualitative data source were integrated into a unified set of themes at the end.

Reflexivity Statement

TKCY is a female PhD student in her 30s from Hong Kong with a background in psychology and research experience in aging and dementia research. She is a core member of the team that co-designed the ENHANCE app. Her role in the project

XSL•FO
RenderX

may have influenced how she conducted interviews and interpreted data. Her familiarity with app usability literature may also have shaped how she identified themes during coding. As a young, educated researcher proficient in technology, she acknowledged a risk that she may underestimate usability challenges and overlook barriers faced by older adults. To mitigate this, she remained open to participants' perspectives, actively probed for any concerns raised by participants during interviews, and regularly discussed her reflections with supervisors throughout the process.

Data Integration Between Quantitative and Qualitative Data

TKCY analyzed quantitative and qualitative data separately, then integrated them through collaborative discussions with the research team of A Singh, GL, SGC, and HKP. These discussions involved comparing and contrasting findings across data sources to identify areas of convergence and divergence.

Results

Overview

All 10 participants completed the introductory session, the 1-week app testing period, telephone follow-up, posttest interview, and the satisfaction survey.

Participants used the app in testing for a mean duration of 9 (SD 3, range 7‐16) days. The participant with the longest testing period (16 d) had a delay in follow-up due to illness, which postponed the follow-up interview.

Demographic Characteristics of Participants

Participants' mean age was 68 (SD 6) years; 7 were female; 4 were White and 6 were non-White (South Asian, East Asian, or Mixed); and 7 had English as a first language. Participants lived in areas spanning a wide range of neighborhood deprivation (mean IMD decile 4, SD 2, range 1‐8). Mean years of full-time education was 14 (SD 3). Refer to Table 1 for full demographic details.

Most had previous experience with technology (smartphone: n=9, computer: n=8, and tablet: n=6); and 5 stated that they were comfortable or highly comfortable using it.

Common dementia risk factors included hypertension and hearing impairment (n=6 each), physical inactivity and depression (n=4 each), and diabetes (n=3). Moreover, 6 participants had 1‐2 risk factors, and 4 participants had 4 or more.

Among participants enrolled in the 1-week testing—5 chose hypertension, 2 hearing impairment, 2 diabetes, and 1 depression.

XSL•FO
RenderX

Table . Demographic characteristics of participants (n=10).

Characteristic	Value
Age (y), mean (SD; range)	68 (6; 60 - 77)
Years of full-time education, mean (SD; range)	14 (3; 10 - 18)
Index of Multiple Deprivation Decile (IMD)[a], mean (SD; range)	4 (2; 1-8)
Sex, n	
Male	3
Female	7
Ethnicity, n	
White British	4
White Irish	2
South Asian (Pakistani, Indian, Bangladeshi)	2
East Asian (Chinese)	1
Mixed (White & Black)	1
Highest education qualification, n	
Lower secondary qualification (CSE/GCSE)	1
Higher secondary qualification (A-Levels/BTech)	6
Bachelor's degree	3
English as first language	7
Technology use[b], n	
Used a tablet before	6
Used a smartphone before	9
Used a computer before	8
Owned a tablet with internet access	5
Comfort using technology, n	
Very uncomfortable	0
Somewhat uncomfortable	3
Neutral	2
Somewhat comfortable	4
Very comfortable	1
Dementia risk factors[c], n	
Less education (<8 y full-time education)	0
Diabetes	3
Hypertension	6
Obesity	2
Smoking	1
Excessive alcohol	1
Physically inactive	4
Socially isolated	1
Subjective hearing impairment	6
Depression	4
Number of dementia risk factors, n	
1	3

XSL•FO
RenderX

Characteristic	Value
2	3
3	0
4	3
5	1

[a]The IMD decile (1=most deprived, 10=least deprived) is derived from participant postcodes.

[b]Dichotomous items; only "Yes" responses are shown.

[c]Participants may have multiple risk factors; n do not sum to 10.

Quantitative Findings

App Use (Minimum and Additional App Use)

All 10 participants completed and exceeded the minimum required tasks during the first week: watching at least 1 video, answering a check-in question, and playing at least 3 games once. Moreover, 5 participants watched the risk factor video more than once, and 7 participants used the in-app messaging (Table 2). Of the 3 participants who did not use the in-app messaging, 1 reported he did not need it, and 2 said they were unaware of the feature although they had learned how to use this feature as part of the introductory coaching session.

Of the 5 participants who had chosen the hypertension module, which includes a feature allowing for daily blood pressure logging, entries were made on average 5 out of 7 days during the study period, and all participants entered it at least 4 times.

All participants exceeded the minimum required level of gameplay activity. Table 3 presents gameplay data, showing mean plays per game of approximately 5 - 10 times per week, with Caterpillar Chase most frequently played and Mushroom Match least. Worm Hunter was tested by only 1 participant, who played it 6 times during the week.

Table . Additional app use (videos, in-app messaging, and check-in).

Characteristic	Value
Videos rewatch frequency, n	
0	5
1	3
2	1
>3	1
Use of in-app messaging[a], n	
Yes	7
No	3
Average weekly check-in frequency (days), mean (SD; range)[b]	5 (1; 4–7)

[a]Participants were coded as having used in-app messaging if they initiated or responded to at least one message with the coach. Those with no interaction were coded as non-users.

[b]*Based on data from 5 participants who are on the hypertension module. Average weekly check-in frequency = (total check-in days ÷ total days of app use)×7.*

XSL•FO
RenderX

Table . Weekly gameplay frequency for each game among participants (n=10).

Game	Participants assigned[a], n	Plays per week per participant, mean (SD)	Range
Caterpillar Chase	6	10 (5)	6 - 18
Hive Finder	4	7 (3)	4 - 9
What Word	5	7 (3)	3 - 10
Stump Shuffle	6	7 (2)	3 - 9
Night Sorter	4	6 (6)	2 - 15
Mushroom Match	4	5 (2)	3 - 6
Worm Hunter[b]	1	6 (6)	6 - 6

[a]The number of participants assigned to each game varies because the app randomly allocated three games to each participant per week.

[b]Data for Worm Hunter should be interpreted with caution as only one participant was assigned this game.

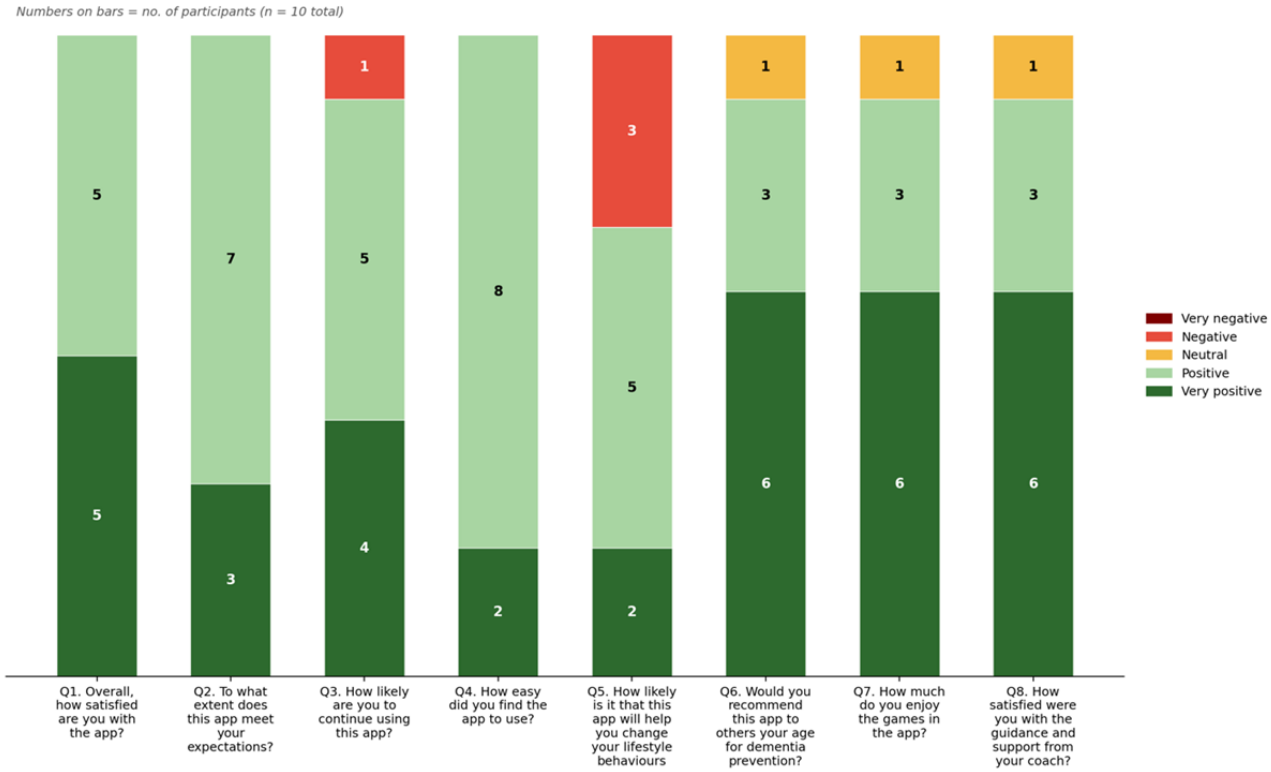
Posttest App Satisfaction Survey

Figure 3 presents posttest survey results summarizing participants' ratings of the app. Response options differed across survey items (eg, satisfaction, likelihood, ease of use, and enjoyment); however, for visualization purposes, responses were harmonized into a standardized 5-point scale ranging from *very negative* to *very positive* based on semantic equivalence. Full survey questions and response labels are provided in Multimedia Appendix 4.

Using the original item-specific response options, all participants reported being either *satisfied* or *very satisfied* with the app and rated it as *easy* or *very easy* to use. All participants felt the app met or exceeded their expectations. Moreover, 9 participants indicated they were *likely* or *very likely* to continue using the app and to recommend it to others their age for dementia prevention. Similarly, 9 participants reported that they *liked* or *liked a lot* the games, and the same number were *satisfied* or *very satisfied* with support from their coach. Regarding the app's potential to support lifestyle behavior change, 7 participants reported that it was *likely* or *very likely* to help, while 3 reported it was *unlikely*.

No adverse events were reported to the coach or research team during the testing period. However, in posttest interviews, some participants described visual discomfort in relation to rapidly moving or flashing visuals in specific games, which are further described in the qualitative analysis result.

Figure 3. User rating of the app.



XSL•FO
RenderX

Qualitative Findings

Overview

Guided by the COM-B model, we identified 6 themes, which are detailed in the further sections. Quotes have age ranges to reduce participant identifiability. The most illustrative quotations are presented in the main text, with additional supporting quotations provided in Multimedia Appendix 5.

Theme 1: Developing Psychological Capability Through Trial-and-Error App Learning

Many participants initially felt uncertain about the unfamiliar interface but gradually gained confidence through trial-and-error app learning. One participant described figuring out *Night Sorter* without initial instructions:

> *I found it [the instructions] a little bit unclear… because it didn't explain what you were trying to do. I had to spend some effort figuring it out myself. But I got a hang of it at last.* [P1, age group 60-69 y, White]

Some app features hindered learning due to competing visual cues. One participant described being confused by the green hand icon in Stump Shuffle until the coach clarified:

> *Because the "hand" is just so overpowering. It was a great big green hand. It's telling you, "Press here." But you're not pressing there—you should be focusing on the odd logo (the target logo; see Multimedia Appendix 6). I didn't get that… I didn't get that at the beginning, but [the coach] explained.* [P4, age group 60-69 y, Asian ethnicity]

App learning also varied depending on participants' previous experience with similar technologies. Familiar interfaces made learning more intuitive. When app visuals mirrored interfaces that participants already knew—such as messaging apps.

> *Yes, for me and I guess most people, the message one [in-app messaging's interface] is intuitive, isn't it? It is like WhatsApp. I thought it pop up the keyboard and everything. That's good.* [P1, age group 60-69 y, White ethnicity]

Theme 2: Drivers of Motivation—Enjoyment and Personal Goal

This theme describes participants' motivation to engage with the ENHANCE app, driven by enjoyment of game challenges, satisfaction from receiving virtual rewards, and desire for self-improvement.

Enjoyment of Game Challenges Boosts Motivation

Many participants described the games as "fun," "exciting," and even said they "got hooked." They especially appreciated games that were mentally stimulating, with just the right level of challenge—difficult enough to engage their thinking, but not so hard as to be discouraging.

> *I think [Hive Finder] is at the right level of challenge because it was pretty challenging for me – I didn't always get it. I'd run out of time sometimes. But it was fun.* [P3, age group 60-69 y, Mixed ethnicity]

Conversely, games perceived as too simple led to disengagement:

> *I don't like this [Night Sorter]. Bosh, bosh, bosh. It's very straightforward. But no, it's not much fun, is it? It's a bit stupid, really, isn't it? Not much thinking needed.* [P1, age group 60-69 y, White]

Participants also valued the adaptive design of some games, where difficulty adjusted based on performance, helping maintain motivation.

> *If you got it wrong, the game [What Word] would give you an easier word to get you more motivated. I like this… Because if it gave me harder ones right after a miss, I'd probably stop and say, "Forget it."* [P3, age group 60-69 y, Mixed ethnicity]

Desire for Self-Improvement Drives Use

Beyond enjoyment, many participants were motivated by a desire to improve their cognitive skills or overall performance, believing that achieving meaningful personal benefits would drive long-term app use.

One participant described the games as a helpful mental exercise:

> *I like games. It's good for me to do it because that's not my strong subject anyway. I regard it as just a good exercise…as an activity which is helpful because it really focuses me on my concentration. So I continued.* [P7, age group 70-79 y, White]

Participants also suggested adding progress visualization features, such as performance trends, believing this would help reinforce personal goals of improving memory and strengthen motivation.

> *Is there anywhere that accumulates those scores to give you, say, a weekly, to show you a weekly, ups and downs or to see whether you've got better during the week or the month? To give you more incentive to keep trying to beat your (game) score.* [P8, age group 70-79 y, White]

Similar progress tracking suggestions extended to other features such as the blood pressure check-in (Multimedia Appendix 5).

In-App Virtual Rewards Reinforce Use

Many participants found the in-game virtual rewards—flowers for completed activities—provided a sense of achievement and encouraged continued participation. One participant described the flower reward as a meaningful achievement:

> *This is a clever idea Yeah. It's accomplishment to me.* [P7, age group 70-79 y, White]

However, when rewards failed to appear due to a bug during testing, motivation declined, leading to frustration and disappointment:

> *Sometimes if I played the game twice in a session, it didn't seem to add, even though it says play for one star, and it didn't, the second time you played it, didn't seem to add the reward to your watering can.…*

XSL•FO
RenderX

> *…This is disappointing. What's the point then?* [P2, age group 60-69 y, White]

Theme 3: Physical Capability Barriers—When Design Overlooks Mobility and Sensory Needs

Some participants with physical or sensory limitations found that certain app features in ENHANCE did not sufficiently support their physical capabilities, creating barriers to engagement. For example, a swipe-based game (*Night Sorter*) was difficult for a participant with reduced movement due to an arm injury. The speed required to play the game interfered with their ability to participate:

> *I've got a problem in my arm because I've had an accident so I can't do fast for this (Night sorter) The swipe thing is kind of like difficult.* [P9, age group 70-79 y, Asian]

Concerns were also raised regarding low-contrast visuals and time constraints (Multimedia Appendix 5).

Theme 4: Discomfort in Gameplay Undermined Motivation

Some participants raised concerns about potential effects from certain games, particularly related to visual discomfort. These experiences undermined motivation and, in some cases, led to disengagement from the games. Some participants reported visual discomfort from specific game elements. For example, *Stump Shuffle*—which involves rapidly moving visuals—was visually demanding:

> *When you had a lot of stumps, and they were moving around… Not just one stump was moving, several of them were moving. I find that…visually quite difficult with flashing numbers, flashing letters, that thing… especially for Epilepsy. That might set you off a bit …* [P2, age group 60-69 y, White]

Similar concerns were raised about Hive Finder due to small, cluttered screen icons (Multimedia Appendix 5).

Theme 5: Enabling Opportunity on App Use—The Role of Coaching and Hands-On Learning

Participants consistently described initial face-to-face coaching as essential, providing real-time support and reassurance.

> *I think it [the introductory session] is very necessary…We can stick to exactly what the coach has shown us. … to have that face-to-face setting up and everything, really.* [P2, age group 60-69 y, White]

Hands-on use during coaching sessions enhanced opportunity to learn by reinforcing understanding. Several participants suggested encouraging users to perform tasks themselves during coaching sessions.

> *[What can be improved in the session?] Yeah, if the coach showed them how to do it once and then had them do it themselves once.* [P3, age group 60-69 y, Mixed ethnicity]

Participants also appreciated ongoing support through phone calls and in-app messaging after the initial session. This helped maintain engagement by providing a sense of personal care and reassurance.

> *I like it [follow-up calls] because you still feel like, oh, somebody really cares.* [P10, age group 60-69 y, White]

Theme 6: Boosting Motivation for Behavior Change: Trusted Experts and Personal Stories

Most participants appreciated the ENHANCE app's educational videos, which combined expert-delivered information with relatable personal stories. This blend of trusted information and real-life experiences encouraged reflection on health behaviors and motivating change.

> *Dr. Sherman [Dr. featured inside the video] is going to be telling the truth because he was going to be real life doctor. It's not like YouTube people… It's going to be true. I was ready and willing to listen carefully and attend.* [P4, age group 60-69 y, Asian]

Practical, actionable advice from professionals also motivated participants to apply recommendations in their daily lives. One participant using the diabetes module appreciated the clear dietary guidance:

> *The video shows me what to do. That's very important…certain foods make the blood sugar go quickly up. If we eat potatoes, they'll make it go up. Glycemic, they call it… That would be really useful and give good tips how to improve it.* [P9, age group 70-79 y, Asian]

Personal stories created emotional resonance that deepened engagement and relevance. One participant expressed a clear preference for human stories over diagrams:

> *I like that it has someone's personal story because it makes it more relatable. personally, I like to see something. I relate more to seeing that human rather than diagrams.* [P1, age group 60-69 y, White]

Furthermore, a participant living with long-term depression found comfort in the simplicity of a story from the "Better Mood" module (Multimedia Appendix 5)

Table 4 summarizes the themes mapped to the COM-B model and their associated design implications, discussed further in the Discussion section.

XSL•FO
RenderX

Table . Design Implications.

COM-B[a] domain and relevant themes	Design implications
Capability	
Theme 1. Developing capability through trial-and-error learning	• Design interfaces that allow users to try, make mistakes, and learn without penalties • Use familiar interface elements and game formats (eg, WhatsApp-style chat, card matching, Sudoku, word games) to support intuitive learning.
Theme 3. When design overlooks mobility and sensory needs	• Use large fonts and high color contrast • Avoid game features that require fine motor skills or fast reactions.
Opportunity	
Theme 5. The role of coaching and hands-on learning	• In-person human coaching is essential to support effective app use. • Introductory sessions should prioritize experiential learning over didactic instruction. • Telephone follow-up coaching is generally preferred over in-app chat, though chat should remain available as an option • Ensure access to a device and reliable internet connectivity. • Family involvement can be beneficial and should remain an optional form of support.
Motivation	
Theme 2. Enjoyment, personal goals, and in-app rewards increase motivation	• Incorporate fun games to boost overall app engagement. • Design games with light problem-solving and adaptive difficulty • Keep games short and update game content regularly • Use consistent virtual rewards linked to personal goals (eg, better brain health) • Use simple, easy-to-understand visual progress tracking linked to users' personal goals.
Theme 4. Visual discomfort in gameplay undermines motivation	• Avoid flashing visuals, rapid motion, or dense layouts. • Offer visually calm alternatives or adjustable display settings. • Test games with users who have sensory sensitivities.
Theme 6. Trusted experts and personal stories increase motivation for behavior change	• Combine trusted expert input with relatable peer narratives • Use clear, actionable health advice • Present information in simple multimedia formats to support reflection and behavior change.

[a]COM-B: capability, opportunity, motivation – behavior.

Discussion

Principal Findings

Our 1-week usability testing suggests that the ENHANCE intervention prototype is acceptable and usable for older adults. All participants met minimum-use targets, with many engaging further, making extensive use of core features (games, videos, and questionnaires). Most also used the in-app messaging feature. All reported high satisfaction with the app and coaching sessions on the postuse survey.

Qualitative interviews identified several facilitators of engagement, such as (1) face-to-face coaching support, (2) appropriately challenging and adaptive games, (3) in-app virtual rewards, (4) visual progress tracking, (5) familiar interface elements, and (6) health information combining expert advice with personal stories. Barriers included unclear interfaces or instructions, design that did not adequately account for users' physical limitations, and visual discomfort linked to specific aspects of game design.

Comparison With Previous Work

Overall, our positive usability and acceptability findings are consistent with the Prevention of Dementia Using Mobile Phone Applications (PRODEMOS) study—the only other UK-based, app-based dementia prevention program targeting underserved populations. The PRODEMOS reported that among its UK participants (n=600), 69% found the app acceptable, 71% considered it feasible, and 77% viewed it as appropriate at 3-month follow-up [36].

While both our study and PRODEMOS included features such as SMART goal setting, educational content, cognitive games, and health tracking, ENHANCE offered more intensive human interaction—including face-to-face onboarding, weekly telephone follow-ups, and in-app messaging with a coach. In contrast, PRODEMOS delivered initial sessions remotely due to COVID-19 restrictions, with follow-up support limited to

XSL•FO
RenderX

asynchronous in-app chat. Participants in PRODEMOS reported difficulties staying engaged, partly attributing this to delays in coach responses [36,37]—a challenge similarly noted in Healthy Ageing Through Internet Counselling in the Elderly (HATICE) internet-based dementia prevention trial [38,39]. .

Our qualitative findings highlighted the value of in-person coaching. Participants described face-to-face onboarding as "necessary" for real-time support and reassurance—echoing findings from the HATICE study [38]. For ongoing support, telephone check-ins were more acceptable than in-app chat: all participants completed phone calls, while only 7 out of 10 used the chat feature. Nonusers tended to be older, less educated, and less familiar with tablets, and were unaware of the chat function despite onboarding demonstrations. This aligns with evidence that human-led approaches—in-person or voice-based coaching—are preferred by older adults and may foster deeper engagement than text-based or AI-driven approaches, such as chatbots [40,41]. In-app chat also posed greater challenges for those with lower digital literacy consistent with evidence that text-based interfaces can present direct barriers when reading or digital skills are limited [42].

While PRODEMOS assessed usability at 3 months, ENHANCE was evaluated after just 1 week, leaving its long-term engagement and feasibility untested. PRODEMOS reported a decline in task completion from 91% in the first 2 weeks to 41% by month 3 [36], reflecting a common pattern of reduced app engagement over time due to loss of novelty and motivation [43,44].

Interpreting Findings Through the COM-B Framework

Using the COM-B framework [32], our qualitative findings can reflect how capability, opportunity, and motivation shaped participants' engagement with the ENHANCE app.

Building Users' Capability Through Effective App Learning

Our findings indicate that trial-and-error learning fostered app engagement by helping users build confidence and skills in using the app (capability). A key facilitator was the use of familiar interface elements—design features similar to those older users already know—which aligns with evidence that recognizable visuals, layouts, and game mechanics reduce cognitive load and enhance usability, particularly for those with limited digital experience [45-48]. In the UK context, games like Scrabble and matching-pair card games are widely recognized by older adults, and several of our games adopted similar mechanics, appearing to lower the cognitive barrier to entry and support a smoother learning curve.

Additionally, we found some app features exceeded users' physical capabilities, reinforcing the need to design in alignment with older adults' motor and visual abilities. Games requiring fine motor skills or fast reactions should be avoided, and interfaces should prioritize large fonts, clear instructions, and limited simultaneous visual cues—consistent with existing recommendations for age-inclusive design that accommodate age-related changes in motor and cognitive function [45-48].

Creating Opportunity Through Coaching

Our study highlights human coaching as a critical factor supporting app engagement, aligning with evidence on blended approaches that combine digital tools with human support, particularly for socioeconomically disadvantaged populations [49]. Personalized coaching can help overcome low digital literacy and complex life challenges (eg, financial barriers to healthy food and poor living conditions) by discussing practical workarounds during sessions [50]. Participants also consistently preferred hands-on, experiential learning during coaching sessions over passive instruction. This supports findings showing that trial-and-error—or "generate-and-correct"—learning enhances recall more effectively than passive study-only approaches [51]. These insights highlight the value of coaching models that emphasize user-led exploration and problem-solving.

Other Opportunities-Related Factors

Our findings partially align with previous research identifying device access, time availability, and family involvement as key opportunity-related factors in app engagement [21,52]. In our study, device and internet barriers were mitigated by providing tablets and connectivity. Time availability, however, did not emerge as a major constraint—possibly reflecting participants' intrinsic motivation or the role of financial incentives in this study.

Although family involvement is often cited as a facilitator [21,34,52], it was notably absent in our findings: despite being asked during onboarding whether they wanted family support, no participant involved family during the testing period, commonly due to living alone or lacking tech-savvy relatives nearby. For older adults from lower socioeconomic or ethnic minority backgrounds, where geographic distance and limited digital literacy within families are common, family support may not be a practical resource [49]. External coaching may therefore serve as a more feasible alternative, with family involvement supplementary where available.

Strengthening Motivation Through App Design

Several app features enhanced both automatic (eg, positive emotional responses) and reflective motivation. These included the games, the combination of professional advice and peer narrative in health information delivery, the desire to improve cognition, and visual feedback and in-app virtual rewards.

The Use of Games

Games were found to be a major source of engagement. Participants enjoyed the cognitive challenge, especially when difficulty was appropriately balanced and adapted to their performance—experiences resembling flow, a state of focused immersion to enhance intrinsic motivation and long-term engagement [53,54]. A meta-analysis found that gamified cognitive training yields greater motivation and engagement than less gamified alternatives [55].

Several design choices likely contributed—games focused on light problem-solving tasks with adaptive difficulty; provided only positive feedback with no penalties, were brief (under 8 min per session); and introduced new games weekly—all

XSL•FO
RenderX

consistent with recommendations for apps or games targeting older adults [46,47]. Competitive elements, such as leaderboards or public performance displays, were deliberately excluded, in line with evidence that these can deter older users [46].

The Use of Expert Guidance and Peer Stories to Deliver Health Information

Our findings suggest that health information was best received when combining expert guidance with peer stories. While older adults tend to trust health care professionals over social media or family and friends [56,57], they also find peer narratives relatable and motivating. Other studies support the value of narrative health communication: personal stories are more emotionally engaging than fact-based messages and are particularly effective at influencing behavioral intentions [58,59]

The Use of Visual Progress Tracking and In-App Virtual Rewards

Participants suggested that visual progress tracking (charts showing improvement in games or behaviors) and in-app virtual rewards (flowers) may support user engagement and health behavior change. Evidence from physical activity apps reinforces this: visual feedback and virtual rewards were the most effective features for predicting actual activity, outperforming elements like leaderboards [60], and rewarding users with virtual currency earned through steps significantly increased activity over 6 months, particularly among sedentary and high-BMI individuals [61]. However, overly complex visualizations of progress tracking can be unclear to older users [62], and extrinsic rewards may undermine long-term motivation once an internally rewarding engagement loop is established [63]. Feedback and rewards should therefore be simple, easy to understand, and aligned with users' personal goals to support long-term motivation.

Strengths and Limitations

This study has several strengths. Recruitment via community settings serving underrepresented groups yielded a sample spanning ethnic minority backgrounds and a wide range of neighborhood deprivation, addressing a key gap in usability evidence for inclusive digital prevention [11]. We used multiple data sources—questionnaires, interviews, and system logs—consistent with the Medical Research Council's guidelines on mixed methods [64] and evidence that combined approaches provide a more comprehensive usability picture than single methods alone [65]. The 1-week testing period captured how user experience evolved across repeated use rather than relying on a single interaction [66], and participants used the app in their natural settings, supporting ecological validity.

Limitations include the 1-week timeframe, which leaves longer-term usability untested, and the small, London-based sample, which limits generalizability to rural areas and other countries. Social desirability bias may have affected interviews, although backend usage data showing use beyond minimum requirements suggests genuine engagement. Despite targeting underserved settings, the sample had relatively high educational attainment; usability barriers may be greater among those with lower literacy. High engagement occurred under supported conditions—device provision, face-to-face onboarding, coach support, and study compensation—which may not reflect real-world implementation. Finally, eligibility relied on not having self-reported dementia without formal clinical assessment. Future studies should incorporate more rigorous eligibility measures.

Conclusion

This study explored how community-recruited older adults, including those underrepresented in dementia prevention trials, engaged with the ENHANCE app over 1 week. The intervention was well-received, with high engagement, positive feedback, and strong usability and acceptability across participants. Key facilitators included face-to-face coaching, familiar interface design, appropriately challenging games, visual progress tracking, and health content that combined expert advice with peer stories. These findings offer practical design insights for developing inclusive digital dementia prevention tools that better accommodate diverse older adults, while future research is needed to evaluate maintained long-term engagement.

Acknowledgments

We thank our research participants and patient and public involvement members for their time, involvement, and valuable feedback.

Funding

This study/project is funded by the NIHR Programme Grant for Applied Research (NIHR203670). The views expressed are those of the authors and not necessarily those of the NIHR. The funder had no role in the study design, data collection, analysis, interpretation of the data, or writing of the manuscript.

Data Availability

The data generated and analyzed during this study are primarily qualitative and include detailed participant accounts. Owing to the risk of re-identification, these data cannot be meaningfully anonymized and are therefore not publicly available. Access to selected, redacted extracts may be considered by the corresponding author on reasonable request, subject to ethical approval and data governance constraints

XSL•FO
RenderX

Authors' Contributions

Conceptualization and methodology: SGC, GL, A Singh, A Sommerlad, HKP, and TKCY formulated the research goals and aims. GL, SGC, and TKCY led the design of the study methodology and selection of testing approaches.
Data collection: TKCY, KF, RB, JJ, CL, and VM
Funding acquisition and supervision: GL and SGC acquired funding. GL, SGC, A Singh, A Sommerlad, and HKP provided oversight and leadership throughout the research activities.
Project administration: JJ
Resources, investigation, and formal analysis: TKCY prepared study materials and performed formal analysis.
Writing: TKCY prepared the initial draft. GL, SGC, A Singh, A Sommerlad, and HKP provided supervisory feedback.
All authors critically reviewed, revised, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Telephone screening questions.
[DOCX File, 202 KB - aging_v9i1e92800_app1.docx]

Multimedia Appendix 2
Sociodemographic and Risk Factor Questionnaire.
[DOCX File, 2666 KB - aging_v9i1e92800_app2.docx]

Multimedia Appendix 3
Guideline for posttest feedback interview.
[DOCX File, 24 KB - aging_v9i1e92800_app3.docx]

Multimedia Appendix 4
Posttest survey.
[DOCX File, 24 KB - aging_v9i1e92800_app4.docx]

Multimedia Appendix 5
Supporting quotes.
[DOCX File, 17 KB - aging_v9i1e92800_app5.docx]

Multimedia Appendix 6
Screenshots of Stump Shuffle.
[DOCX File, 504 KB - aging_v9i1e92800_app6.docx]

References

1. Nichols E, Steinmetz JD, Vollset SE. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
2. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet 2024 Aug 10;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
3. Mendes AJ, Ribaldi F, Sayin O, et al. Single-domain and multidomain lifestyle interventions for the prevention of cognitive decline in older adults who are cognitively unimpaired: a systematic review and network meta-analysis. Lancet Healthy Longev 2025 Sep;6(9):100762. [doi: 10.1016/j.lanhl.2025.100762] [Medline: 41016407]
4. Brodaty H, Chau T, Heffernan M, et al. Maintain your brain: recent findings, future challenges. Alzheimer's & Dementia 2025 Dec;21(S6). [doi: 10.1002/alz70860_103197]
5. Brodaty H, Chau T, Heffernan M, et al. An online multidomain lifestyle intervention to prevent cognitive decline in at-risk older adults: a randomized controlled trial. Nat Med 2025 Feb;31(2):565-573. [doi: 10.1038/s41591-024-03351-6] [Medline: 39875685]
6. Mukadam N, Lewis G, Mueller C, Werbeloff N, Stewart R, Livingston G. Ethnic differences in cognition and age in people diagnosed with dementia: a study of electronic health records in two large mental healthcare providers. Int J Geriatr Psychiatry 2019 Mar;34(3):504-510. [doi: 10.1002/gps.5046] [Medline: 30675737]

XSL•FO
RenderX

7. Cadar D, Lassale C, Davies H, Llewellyn DJ, Batty GD, Steptoe A. Individual and area-based socioeconomic factors associated with dementia incidence in England: evidence from a 12-year follow-up in the english longitudinal study of ageing. JAMA Psychiatry 2018 Jul 1;75(7):723-732. [doi: 10.1001/jamapsychiatry.2018.1012] [Medline: 29799983]
8. Li R, Li R, Xie J, et al. Associations of socioeconomic status and healthy lifestyle with incident early-onset and late-onset dementia: a prospective cohort study. Lancet Healthy Longev 2023 Dec;4(12):e693-e702. [doi: 10.1016/S2666-7568(23)00211-8] [Medline: 38042162]
9. Pham TM, Petersen I, Walters K, et al. Trends in dementia diagnosis rates in UK ethnic groups: analysis of UK primary care data. Clin Epidemiol 2018;10:949-960. [doi: 10.2147/CLEP.S152647] [Medline: 30123007]
10. Ma'u E, Cullum S, Cheung G, Tamatea J, Livingston G, Mukadam N. Variation of the risk associated with potentially modifiable risk factors for dementia between ethnic groups within one country. A retrospective cohort study using routinely collected health data in Aotearoa New Zealand. Int J Geriatr Psychiatry 2025 Jan;40(1):e70042. [doi: 10.1002/gps.70042] [Medline: 39780033]
11. Lazaar N, Van Beek SE, Rirash AF, et al. Diversity in United States dementia prevention trials: an updated systematic review of eligibility criteria and recruitment strategies. Dement Geriatr Cogn Disord 2025;54(5):333-346. [doi: 10.1159/000543905] [Medline: 39947158]
12. Shaw AR, Perales-Puchalt J, Johnson E, et al. Representation of racial and ethnic minority populations in dementia prevention trials: a systematic review. J Prev Alzheimers Dis 2022;9(1):113-118. [doi: 10.14283/jpad.2021.49] [Medline: 35098981]
13. Krishnan A, Pathak A, Nicholas TB, Lee J, Waite L, Stanaway F. Racial and ethnic minority representation in dementia risk factor research: a scoping review of cohort studies. BMJ Open 2024 Sep 25;14(9):e085592. [doi: 10.1136/bmjopen-2024-085592] [Medline: 39322589]
14. Tennstedt SL, Unverzagt FW. The ACTIVE study: study overview and major findings. J Aging Health 2013 Dec;25(8 Suppl):3S-20S. [doi: 10.1177/0898264313518133] [Medline: 24385638]
15. Ngandu T, Lehtisalo J, Levälahti E, et al. Recruitment and baseline characteristics of participants in the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER)-a randomized controlled lifestyle trial. Int J Environ Res Public Health 2014 Sep 10;11(9):9345-9360. [doi: 10.3390/ijerph110909345] [Medline: 25211775]
16. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023;11:1281923. [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
17. Israel‐McLeish K. Development and feasibility of the Community Ageing Research Across Ethnicities (CARE) network app for diverse communities and faith groups. Alzheimer's & Dementia 2025 Dec;21(S4):1002. [doi: 10.1002/alz70858_104972]
18. Briefing: facts and figures about digital inclusion and older people. : Age UK; 2025 Jul URL: https://www.ageuk.org.uk/siteassets/documents/reports-and-publications/reports-and-briefings/active-communities/internet-use-statistics-july-2025-1.pdf [accessed 2026-07-13]
19. Share of smartphone users in the United Kingdom (UK) from 2012 to 2024, by socio-economic group. Statista. 2019. URL: https://www.statista.com/statistics/300421/smartphone-usage-in-the-uk-by-socio-economic-group/ [accessed 2026-07-12]
20. Tang Y, Sinha KK, Moen A, Ertekin N. Towards achieving mental health equity for the underserved population: evaluating the potential of mobile apps. Production and Operations Management 2024 Jan;33(1):87-107. [doi: 10.1177/10591478231224914]
21. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
22. Hepburn J, Williams L, McCann L. Barriers to and facilitators of digital health technology adoption among older adults with chronic diseases: updated systematic review. JMIR Aging 2025 Sep 11;8:e80000. [doi: 10.2196/80000] [Medline: 40934502]
23. Yu CTK, Livingston G, Louise A, et al. Co-producing a coach-supported digital intervention to promote cognitive health in underserved older adults: ENHANCE (tailored intervention for brain health and cognitive enrichment) (preprint). Journal of Medical Internet Research. Preprint posted online on 2026. [doi: 10.2196/preprints.91999]
24. Arroll B, Goodyear-Smith F, Crengle S, et al. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. Ann Fam Med 2010;8(4):348-353. [doi: 10.1370/afm.1139] [Medline: 20644190]
25. Lubben J, Blozik E, Gillmann G, et al. Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. Gerontologist 2006 Aug;46(4):503-513. [doi: 10.1093/geront/46.4.503] [Medline: 16921004]
26. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]
27. Ministry of housing, communities & local government. English Indices of Deprivation 2019 2019.
28. Doran GT. There's a S.M.A.R.T. way to write management's goals and objectives. Manage Rev 1981;70(11):35-36 [FREE Full text]
29. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann 1979;2(3):197-207. [doi: 10.1016/0149-7189(79)90094-6] [Medline: 10245370]

XSL•FO
RenderX

30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
32. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(42):42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
33. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. J Med Internet Res 2020 May 29;22(5):e17572. [doi: 10.2196/17572] [Medline: 32348255]
34. Szinay D, Perski O, Jones A, Chadborn T, Brown J, Naughton F. Perceptions of factors influencing engagement with health and well-being apps in the United Kingdom: qualitative interview study. JMIR Mhealth Uhealth 2021 Dec 16;9(12):e29098. [doi: 10.2196/29098] [Medline: 34927597]
35. Whitehall JM, Cook EJ, Vseteckova J, et al. A systematic review of influences on engagement with remote health interventions targeting weight management for individuals living with excess weight. Int J Obes 2025 Aug;49(8):1427-1468. [doi: 10.1038/s41366-025-01811-8]
36. Moll van Charante EP, Hoevenaar-Blom MP, Song M, et al. Prevention of dementia using mobile phone applications (PRODEMOS): a multinational, randomised, controlled effectiveness-implementation trial. Lancet Healthy Longev 2024 Jun;5(6):e431-e442. [doi: 10.1016/S2666-7568(24)00068-0] [Medline: 38763155]
37. Eggink E, Hafdi M, Hoevenaar-Blom MP, et al. Prevention of dementia using mobile phone applications (PRODEMOS): protocol for an international randomised controlled trial. BMJ Open 2021 Jun 9;11(6):e049762. [doi: 10.1136/bmjopen-2021-049762] [Medline: 34108173]
38. van Middelaar T, Beishuizen CRL, Guillemont J, et al. Engaging older people in an internet platform for cardiovascular risk self-management: a qualitative study among Dutch HATICE participants. BMJ Open 2018 Jan 21;8(1):e019683. [doi: 10.1136/bmjopen-2017-019683] [Medline: 29358447]
39. Richard E, Jongstra S, Soininen H, et al. Healthy ageing through internet counselling in the elderly: the HATICE randomised controlled trial for the prevention of cardiovascular disease and cognitive impairment. BMJ Open 2016 Jun 10;6(6):e010806. [doi: 10.1136/bmjopen-2015-010806] [Medline: 27288376]
40. Loughnane C, Laiti J, O'Donovan R, Dunne PJ. Systematic review exploring human, AI, and hybrid health coaching in digital health interventions: trends, engagement, and lifestyle outcomes. Front Digit Health 2025;7:1536416. [doi: 10.3389/fdgth.2025.1536416] [Medline: 40343215]
41. Bosco C, Shojaei F, Theisz AA, et al. Testing 3 modalities (Voice Assistant, Chatbot, and Mobile App) to assist older African American and Black adults in seeking information on Alzheimer disease and related dementias: wizard of Oz usability study. JMIR Form Res 2024 Dec 9;8:e60650. [doi: 10.2196/60650] [Medline: 39653372]
42. Michelle C, Sharon B. Chatbots in Health Care: Connecting Patients to Information: Emerging Health Technologies 2024. [Medline: 38564540]
43. Giebel GD, Abels C, Plescher F, et al. Problems and barriers related to the use of mHealth apps from the perspective of patients: focus group and interview study. J Med Internet Res 2024 Apr 23;26(1-12):e49982. [doi: 10.2196/49982] [Medline: 38652508]
44. Mustafa AS, Ali N, Dhillon JS, Alkawsi G, Baashar Y. User engagement and abandonment of mHealth: a cross-sectional survey. Healthcare (Basel) 2022 Jan 24;10(2):221. [doi: 10.3390/healthcare10020221] [Medline: 35206837]
45. Lu MH, Lin W, Yueh HP. Development and evaluation of a cognitive training game for older people: a design-based approach. Front Psychol 2017;8:1837. [doi: 10.3389/fpsyg.2017.01837] [Medline: 29089914]
46. Salazar-Cardona J, Vela FLG, Arango-Lopez J, Paderewski P, Moreira F. Older adults and game-based systems: engagement model and player types of characterization based on their motivations. Journal of Engineering Research 2025 Dec;13(4):3107-3120. [doi: 10.1016/j.jer.2024.10.015]
47. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023 Sep 21;11:e43186. [doi: 10.2196/43186] [Medline: 37733401]
48. Liu N, Yin J, Tan SSL, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. J Am Med Inform Assoc 2021 Oct 12;28(11):2483-2501. [doi: 10.1093/jamia/ocab151] [Medline: 34472601]
49. Hengst TM, Lechner L, Dohmen D, Bolman CA. The facilitators and barriers of mHealth adoption and use among people with a low socio-economic position: a scoping review. Digit Health 2023;9(1-21):20552076231198702. [doi: 10.1177/20552076231198702] [Medline: 37691766]
50. Miller LMS, Callegari RA, Abah T, Fann H. Digital literacy training for low-income older adults through undergraduate community-engaged learning: single-group pretest-posttest study. JMIR Aging 2024 May 14;7:e51675. [doi: 10.2196/51675] [Medline: 38599620]
51. Briscoe J, Doherty J, Burgess K, Kent C. Errorful learning improves recognition memory for new vocabulary for people living with memory and dysexecutive impairment following brain injury. Neuropsychol Rehabil 2024 Aug;34(7):974-1004. [doi: 10.1080/09602011.2023.2259017] [Medline: 37733957]

XSL•FO
RenderX

52. Al-Dhahir I, Reijnders T, Faber JS, et al. The barriers and facilitators of eHealth-based lifestyle intervention programs for people with a low socioeconomic status: scoping review. J Med Internet Res 2022 Aug 24;24(8):e34229. [doi: 10.2196/34229] [Medline: 36001380]
53. Nakamura J, Tse DCK, Shankland S. Flow: the experience of intrinsic motivation. In: Ryan RM, editor. The Oxford Handbook of Human Motivation: Oxford University Press; 2019:169-186. [doi: 10.1093/oxfordhb/9780190666453.013.10]
54. Payne BR, Jackson JJ, Noh SR, Stine-Morrow EAL. In the zone: flow state and cognition in older adults. Psychol Aging 2011 Sep;26(3):738-743. [doi: 10.1037/a0022359] [Medline: 21443348]
55. Vermeir JF, White MJ, Johnson D, Crombez G, Van Ryckeghem DML. The effects of gamification on computerized cognitive training: systematic review and meta-analysis. JMIR Serious Games 2020 Aug 10;8(3):e18644. [doi: 10.2196/18644] [Medline: 32773374]
56. Agyemang-Duah W, Arthur-Holmes F, Peprah C, Adei D, Peprah P. Dynamics of health information-seeking behaviour among older adults with very low incomes in Ghana: a qualitative study. BMC Public Health 2020 Jun 15;20(1):928. [doi: 10.1186/s12889-020-08982-1] [Medline: 32539693]
57. Shiroma K, Zimmerman T, Xie B, et al. Older adults' trust and distrust in COVID-19 public health information: qualitative critical incident study. JMIR Aging 2023 Nov 9;6:e42517. [doi: 10.2196/42517] [Medline: 37856774]
58. Dudley MZ, Squires GK, Petroske TM, Dawson S, Brewer J. The use of narrative in science and health communication: a scoping review. Patient Educ Couns 2023 Jul;112:107752. [doi: 10.1016/j.pec.2023.107752] [Medline: 37068426]
59. Xu J. A meta-analysis comparing the effectiveness of narrative vs. statistical evidence: health vs. non-health contexts. Health Commun 2023 Dec;38(14):3113-3123. [doi: 10.1080/10410236.2022.2137750] [Medline: 36278821]
60. Yu T, Parry M, Yu T, et al. Effectiveness of mobile health-based gamification interventions for improving physical activity in individuals with cardiovascular diseases: systematic review and meta-analysis of randomized controlled trials. JMIR Serious Games 2025 Jan 24;13:e64410. [doi: 10.2196/64410] [Medline: 39854099]
61. Elliott M, Eck F, Khmelev E, Derlyatka A, Fomenko O. Physical activity behavior change driven by engagement with an incentive-based app: evaluating the impact of sweatcoin. JMIR Mhealth Uhealth 2019 Jul 8;7(7):e12445. [doi: 10.2196/12445] [Medline: 31287064]
62. While Z, Crouser RJ, Sarvghad A. GerontoVis: data visualization at the confluence of aging. Computer Graphics Forum 2024 Jun;43(3) [FREE Full text] [doi: 10.1111/cgf.15101]
63. Bardach L, Murayama K. The role of rewards in motivation—beyond dichotomies. Learn Instr 2025 Apr;96:102056. [doi: 10.1016/j.learninstruc.2024.102056]
64. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of medical research council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
65. Maqbool B, Herold S. Potential effectiveness and efficiency issues in usability evaluation within digital health: a systematic literature review. Journal of Systems and Software 2024 Feb;208:111881. [doi: 10.1016/j.jss.2023.111881]
66. Bruni F, Mancuso V, Pedroli E. How to evaluate user experience in digital health? A scoping review of questionnaires in virtual reality applications. Front Digit Health 2025;7:1561364. [doi: 10.3389/fdgth.2025.1561364] [Medline: 41031121]

Abbreviations

ACTIVE: Advanced Cognitive Training for Independent and Vital Elderly
COM-B: capability, opportunity, motivation—behavior
CSQ: Client Satisfaction Questionnaire
ENHANCE: Tailored Intervention for Brain Health and Cognitive Enrichment
FINGER: Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability
HATICE: Healthy Ageing Through Internet Counselling in the Elderly
IMD: Index of Multiple Deprivation
PHQ-2: Patient Health Questionnaire−2
PRODEMOS: Prevention of Dementia Using Mobile Phone Applications
SMART: specific, measurable, achievable, relevant, time-bound

Edited by H LaMonica; submitted 03.Feb.2026; peer-reviewed by F Elkourdi, M Neff; revised version received 15.Jun.2026; accepted 23.Jun.2026; published 23.Jul.2026.

Please cite as:
Yu TKC, Livingston G, Boczko R, Forristal K, Jamison J, Leckstein C, Mehta V, Park HK, Singh A, Sommerlad A, Costafreda SG
Usability of the Coach-Supported Dementia Prevention App ENHANCE (Tailored Intervention for Brain Health and Cognitive Enrichment) in Older Adults: 1-Week Mixed Methods Study
JMIR Aging 2026;9:e92800
URL: https://aging.jmir.org/2026/1/e92800
doi:10.2196/92800

XSL•FO
RenderX

© Tsz Kiu Clare Yu, Gill Livingston, Richard Boczko, Kealan Forristal, James Jamison, Carl Leckstein, Vrushti Mehta, Hee Kyung Park, Aneesha Singh, Andrew Sommerlad, Sergi G Costafreda. Originally published in JMIR Aging (https://aging.jmir.org), 23.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Personalized Predictive Model to Predict Subtask Success of Medication Adherence Technologies for Older Adults With Diverse Capabilities: Development and Internal Validation Study

Bincy Baby[1], PharmD, MSc Pharmacy; Ghada Elba[1], PharmD, MSc Pharmacy; Minzee Kim[2], BSc, MMath; SooMin Park[1]; Imra Hudani[1], BSc; Rishabh Sharma[1], PharmD, MSc Pharmacy; Halak Patel[1]; Sidharth Bajaj[2], BMath, MMath; P K Patterson[3], BA; Annette McKinnon[3]; Sara J T Guilcher[4], PT, PhD; Feng Chang[1], PharmD; Linda Lee[5], MD, MCISc(FM); Catherine Burns[6], PEng, PhD; Ryan Griffin[7], P.Eng, PhD; Joslin Goh[2], BSc, MSc, PhD; Tejal Patel[1,8], BScPharm, PharmD

[1]School of Pharmacy, University of Waterloo, 10 Victoria Street South, Waterloo, ON, Canada

[2]Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON, Canada

[3]Patient Advisor's Network, Toronto, ON, Canada

[4]Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada

[5]Department of Family Medicine, McMaster University, Hamilton, ON, Canada

[6]Systems Design Engineering, Faculty of Engineering, University of Waterloo, Waterloo, ON, Canada

[7]National Research Council Canada, Ottawa, ON, Canada

[8]Schlegel-UW Research Institute for Aging, Waterloo, ON, Canada

Corresponding Author:
Tejal Patel, BScPharm, PharmD
School of Pharmacy, University of Waterloo, 10 Victoria Street South, Waterloo, ON, Canada

Abstract

Background: Older adults frequently experience cognitive, physical, sensory, motivational, and environmental barriers that affect medication management. Medication adherence technologies (MATs) can support adherence, but their usability varies widely depending on individual abilities and device features. Prior research has largely focused on overall adherence or user experience, providing limited insight into feature-level usability challenges.

Objective: The aim of the study is to develop and internally validate a personalized predictive model to predict the success of MAT subtasks for older adults with diverse cognitive, physical, sensory, motivational, and environmental capabilities.

Methods: A mixed methods approach was used, incorporating the assessment of impairments using various standardized questionnaires, measurement of usability metrics through cognitive walkthroughs, and one-on-one semistructured interviews. For this study, we used "subtasks" as the representative of features of the devices. A subtask is a discrete, individual action that forms part of a larger task, specifically designed to achieve a step in the overall process. Participants tested between 1 and 7 devices from a selection of 13 devices. The proportion of subtask success was taken as the outcome measure. Predictors included demographic, clinical, cognitive, physical, sensory, motivational, and environmental characteristics. Personalized predictive modeling using cosine similarity and generalized linear models were compared with nonpersonalized and naive models. Model performance was evaluated using mean square error (MSE) through cross-validation and held-out validation.

Results: A total of 117 participants (mean age 74.6, SD 7.9 years) were recruited, including 96 participants for usability testing and 21 for the validation, all varying in cognitive, physical, sensory, motivational, and environmental abilities. Both personalized (m=0.25) and nonpersonalized models (m=1.0) outperformed naive predictions (m=1.21), demonstrating that subtask-level success can be predicted using routinely measurable demographic and functional characteristics. During cross-validation, personalized models achieved optimal performance at a matching proportion of m=0.25, with MSEs lower than those observed at higher matching levels, although differences compared with nonpersonalized models were not statistically significant (Self-Medication Assessment Tool [SMAT]: P=.50; Daily Living Tasks Dependent on Vision [DLTV]: P=.43). In the held-out validation cohort, personalized models achieved MSEs of 0.89 (SMAT-based) and 1.16 (DLTV-based) at m=0.20, whereas nonpersonalized models demonstrated better performance with MSEs of 0.726 (SMAT-based) and 0.815 (DLTV-based). Models incorporating performance-based vision measures (SMAT-based) consistently outperformed those using self-reported vision scores (DLTV-based) across both personalized and nonpersonalized settings.

Conclusions: This study demonstrates the feasibility of predicting subtask success of MATs in older adults. While personalization showed limited added benefit in this dataset, the subtask-focused model provides clinically meaningful insights to support

XSL•FO
RenderX

evidence-informed selection of medication technologies, reduce usability-related medication errors, and improve adherence outcomes.

(JMIR Aging 2026;9:e84616) doi:10.2196/84616

KEYWORDS

medication adherence; technology; older adults; personalized predictive modeling; usability

Introduction

As the global population ages, the proportion of older adults with chronic health conditions is increasing significantly [1]. According to projections by Statistics Canada, approximately 25% of Canadian people will be older than 65 years of age by 2036, highlighting the importance of creating support systems that allow these individuals to live independently and maintain their quality of life [2]. Chronic health conditions, including heart disease, diabetes, arthritis, and other ailments, are more prevalent in this age group, presenting various challenges that impair their ability to manage daily tasks, including the crucial task of medication management [1]. Complex medication regimens and polypharmacy, a common practice due to multimorbid conditions, introduce additional challenges that often lead to various medication-related problems, such as medication errors, potentially inappropriate medications, adverse drug reactions, and nonadherence, subsequently preventing the achievement of therapeutic outcomes [3-5].

Medication management capacity (MMC) is defined as the "cognitive and functional ability to comply with a medication regimen, when it is the individual's wish or desire to follow a medication regimen as prescribed," and is recognized as an important aspect of adherence [6]. In terms of MMC, deviations from adherence result from a lack of ability and are unintentional [6]. Various challenges associated with aging and chronic conditions can affect medication management, including cognitive, physical, sensory, motivational, and environmental barriers [6-9]. Cognitive decline, often exacerbated by chronic illnesses or age-related conditions such as dementia, presents challenges with memory, spatial orientation, and processing speed, which make it difficult for older adults to follow medication schedules accurately, increasing the possibility of errors and adverse events [9-12]. Physical impairments, such as reduced grip strength, fine motor skills, and hand dexterity, further compromise the self-management of medications. For instance, individuals with rheumatoid arthritis often struggle to open medication packaging due to reduced hand function, which can lead to missed doses and unintentional nonadherence [13,14]. Sensory impairments, including decreased visual acuity and hearing, also hinder medication adherence, as patients may be unable to read labels or hear auditory reminders [9,15-17]. Motivational factors, such as confidence in managing medications and understanding the benefits of adherence, play an important role as well [18-20]. Additionally, environmental factors like inadequate lighting at home, cluttered spaces, and a lack of support from caregivers or family can further complicate effective medication management [21-23].

Addressing these barriers is essential, as medication nonadherence can lead to increased health care costs, hospitalizations, and even preventable deaths [7,8,24,25]. Various studies have explored potential solutions through the development and implementation of assistive technologies to support older adults in managing their medications independently [26-29]. Medication adherence technologies (MATs), such as electronic pill dispensers, ingestible sensors, blister pack systems, and smart medication bottles, have emerged as aids, especially for older adults facing cognitive or physical impairments [26-29]. Research indicates that these technologies could improve adherence by providing alerts, reminders, and dosage tracking to simplify the medication-taking process [30,31]. A mobile health–based medication adherence measurement system using electronic blisters in 53 older adults with increased cardiovascular risk significantly improved adherence to diabetes medication (P=.04) during the monitoring phase. Similarly, another 6-month trial in primary care patients with moderate to severe asthma found that inhaler reminders and feedback significantly improved adherence (73% vs 46%; P<.001). However, how easy MATs are to use and how they appear to the user vary depending on individual user characteristics, impairments, and specific device features [32-36]. For example, individuals who are visually impaired may struggle with devices that rely solely on visual cues, and older adult users may have difficulty opening devices that require a lock and key or are too tight to open due to limited hand dexterity, highlighting the importance of selecting technologies with accessible features that match individual abilities.

Usability focuses on the ease, effectiveness, and efficiency with which users can complete tasks, while user experience (UX) encompasses the overall emotional and psychological interaction with a device, including satisfaction and engagement [37-39]. Both are critical for creating successful technologies, especially for older adults, whose cognitive, sensory, and motor impairments require designs that improve accessibility and independence [37-39]. Previous research on MATs has generally assessed UX without focusing on specific device features or the diverse challenges faced by users with different impairments [35]. Given that each older adult may have unique impairments and that MAT features vary widely, it is crucial to evaluate which specific features facilitate or hinder usability. Personalized predictive modeling (PPM) offers a promising approach by accounting for individual heterogeneity and enabling the prediction of feature-level success based on a user's unique functional profile, thereby supporting more targeted and patient-centered MAT selection [40-42].

Therefore, the objective of this study is to develop and internally validate a personalized predictive model to predict the success of MAT subtasks for older adults with diverse cognitive, physical, sensory, motivational, and environmental capabilities.

XSL•FO
RenderX

For this study, we used "subtasks" as the representative of features of the devices. A subtask is a discrete, individual action that forms part of a larger task, specifically designed to achieve a step in the overall process. It involves precise activities, such as inserting a battery or setting a date, which are essential for the successful operation of a device. By developing a personalized predictive model to predict subtask-level success, this study seeks to guide the selection of more user-friendly MATs that meet the diverse needs of older adults. Ultimately, this tailored approach aims to reduce medication nonadherence, minimize health risks, and support the independence of older adults by ensuring that the technologies they use are accessible and aligned with their unique capabilities.

Methods

This paper adheres to the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines for the development of a multivariable prediction model [43]. The TRIPOD checklist was used to ensure complete and transparent reporting of all relevant items (Checklist 1).

Study Design

This study used a prospective mixed methods research design conducted in 2 parts: usability testing and a validation part. Both phases followed the same overall study procedures. This paper focuses specifically on the development and validation of the personalized prediction model. Other outcome variables collected in the study, including additional performance and perception-based measures and qualitative findings from one-on-one interviews, are reported in separate papers.

Eligibility Criteria

Participants were eligible to participate in the study if they were 60 years or older of age and were willing and able to provide informed consent. Participants were excluded from the study if they were unable to speak or read English, were younger than 60 years of age, or were unwilling or unable to provide informed consent. For participants who were unable to provide consent, ones with cognitive impairment, consent was obtained from their caregivers.

Sample and Recruitment

Participants were recruited using purposive and snowball sampling methods to ensure a diverse and representative sample. Initially, we recruited participants and measured their age-related abilities. However, after reaching 50 participants, we encountered challenges in recruiting an adequate number of individuals with diverse impairments. To address this, we modified our recruitment strategy by incorporating screening questions to identify specific symptoms, such as memory problems, hand tremors, limited hand movements, low hand grip strength, low vision, difficulty differentiating colors, blurred vision, low hearing, hand paralysis, and lack of sensation in the hands. Additionally, we included questions assessing medication adherence, such as how often participants skipped their medication due to a lack of motivation or a busy schedule. These modifications allowed us to target participants with specific impairments.

Sample Size

The sample size calculation for the usability study was guided by Green's formula for regression analysis ($n \geq 50+8m$), where m represents the number of independent variables (physical, cognitive, vision, hearing, motivational, and environmental impairments) [44]. Based on this criterion, a minimum sample size of 98 participants was required. Recruitment efforts aimed to ensure that at least 5 participants with each type of impairment tested each of the 13 selected devices, providing a comprehensive and balanced representation across diverse user needs. Recruitment and data collection took place from June 2023 to June 2024.

For internal validation, an increase in sample size by 30% was estimated to be adequate, corresponding to a target of 29 participants. Recruitment and data collection for the validation study took place from May 2025 to August 2025 [45,46].

Study Procedure

An overview of the study procedures followed for the usability and validation phases is provided below. The study procedures remained exactly the same for both phases, with the exception that one-on-one interviews were not conducted during the validation phase.

Step 1: Measurement of Barriers to MMC

The first step involved assessing various barriers to medication self-management, including physical, cognitive, vision, hearing, motivational, and environmental factors. The following tools were used to determine these abilities:

- Self-Medication Assessment Tool (SMAT): Evaluated physical, cognitive, and vision barriers [47,48].
- Daily Living Tasks Dependent on Vision (DLTV) questionnaire: Focused on vision-specific barriers [49,50].
- Whisper test: Measured hearing impairment, categorizing it as 100% impairment if both ears failed, 50% if 1 ear failed, and no impairment if both ears passed [51,52].
- Self-Efficacy for Appropriate Medication Use Scale (SEAMS): Assessed motivational barriers related to confidence in adhering to medication regimens [20,53-55].
- Martin and Park Environmental Demands (MPED) questionnaire: Evaluated environmental challenges that could impact medication management [21].

Step 2: Usability Testing

Overview

After assessing barriers, participants were assigned MATs and performed tasks using these devices following the cognitive walkthrough approach. Two evaluators observed participants' interactions with the devices, systematically documenting task and subtask outcomes, including success, failure, errors, and time taken. The evaluators then reached a consensus regarding the participants' performance on each subtask to ensure consistency and reliability of the observations. More details on the cognitive walkthrough, the devices tested, and the mock medication regimen used in this study are provided below.

XSL•FO
RenderX

Cognitive Walkthrough

The cognitive walkthrough method evaluates how new users interact with a device by simulating their initial experience of exploring its functionality [37-39]. This task-based approach focuses on assessing ease of use and learnability, with the goal of identifying usability challenges that may impede task completion [37-39]. Participants were given a set of predefined tasks outlined in a participant information sheet, designed to evaluate various features and functions of the devices. Tasks ranged from simple actions, such as inserting batteries and unlocking devices, to more complex activities, such as setting alarms, filling medication trays, and retrieving medication after an alert. Each task consisted of smaller, discrete actions called subtasks. For example, in the case of inserting batteries, the subtasks included opening the battery compartment, correctly positioning the batteries, closing the compartment, and ensuring the device was powered on. Across the 13 devices evaluated, 32 subtasks were identified and generalized to ensure broad applicability across devices. The list of subtasks tested in this study is given in Table 1.

Table . Subtasks for using various devices.

Code	Subtask
A1	Locate the battery or cartridge compartment or medication cavity
A2	Put or insert the battery correctly
A3	Lift or close the battery compartment door
A4	Slide in or out the battery compartment door
A5	Slide a tab or button
A6	Check or ensure or verify the device is on or the lock is placed in position or follow instructions or ensure the indicator light flashes for 3 seconds
B1	Flip device
B2	Insert key and rotate
B3	Press and rotate the lid
B4	Open the lid by lifting
C1	Press and hold a button on a device
C2	Press a button on a device
D1	Open the pill box or compartment by rotating the lid
D2	Open the pill box or compartment or tray or door by sliding
D3	Pick up the correct pillbox or pill organizer or open correct compartment
D4	Insert or fill or place medication in the compartment or pillbox or pill organizer
D5	Close lid
D6	Put stickers on pillbox dividers
G1	Remove the medication
G2	Grab or hold the device
G3	Place hand over the open slot
H1	Rotate the carousel 3 days from today’s date
I1	Locate and touch on an icon or button on an app or screen
I2	Enter or type any data in an app or screen
I3	Scroll the screen options
M1	Align and insert the cartridge into the designated slot
P1	Tear the package
R1	Rotate retaining clips at each end of the device in an open or close position
R2	Align connectors to one another and gently push the card into the device
T1	Pierce cavity barrier
T2	Pinch number printed on the card and pull out
U1	Pull the blister packs away from the device

XSL•FO
RenderX

For each device tested, both a participant information sheet and an evaluation sheet were prepared. The information sheet provided instructions for the tasks participants needed to complete, while the evaluation sheet detailed each task and its associated subtasks. The evaluation sheet also recorded key metrics such as time taken for each task, whether each subtask was completed successfully without assistance, with assistance, or failed, and the presence or absence of errors during each subtask.

Device Tested

In this study, 3 smart and 10 electronic MATs were tested for usability and UX. Electronic MATs often include features like multiple compartments and alarms to remind patients to take their medications. In contrast, smart MATs incorporate advanced connectivity features, such as Bluetooth and Wi-Fi, enabling real-time tracking, remote monitoring, and instant feedback between patients, caregivers, and health care providers. These devices frequently include additional capabilities such as audio-visual reminders, adherence data logging, and report generation. The devices to be tested were selected to represent a wide diversity of features, based on a feature-based classification system developed for MATs [56]. Specifically, the selection aimed to include devices with variation across key dimensions of the classification system, including physical features, display, connectivity, system alert, data collection and management, operations, and integration. Devices were purposively chosen to capture differences in compartment design, dispensing and access mechanisms, alert modalities, portability, and levels of automation and connectivity. All selected devices were commercially available and commonly used for medication management. This approach ensured that the sample of MATs reflected a broad range of characteristics and functionalities relevant to older adults and supported a comprehensive testing of usability across diverse MATs. A list of the tested devices is provided in Multimedia Appendix 1. Each participant tested between 1 and 7 devices, with some participants requiring 2 visits to complete their testing. Initially, devices were assigned by the researchers without following specific methods or procedures. However, after testing 50 participants, the assignment strategy was revised to ensure at least 5 participants with each type of impairment tested each device, aiming for balanced representation across 6 types of impairments and devices. No formal training was provided to participants, who relied solely on the instructional materials included with the devices, simulating real-world conditions where users independently learn to use such devices.

Mock Medication Regimen

A simulated medication regimen was designed for this study to replicate the complexity of real-life medication schedules often managed by older adults. Developed in-house, this mock regimen aimed to assess the usability of technologies by mimicking the organization of medications in everyday of MATs while ensuring safety by substituting actual medications with placebo tablets, placebo capsules, and candy.

The mock regimen included placebo equivalents for common medications used to manage chronic conditions, structured as follows:

- Warfarin: 2 mg once daily on Monday, Wednesday, and Friday, and 3 mg once daily on Tuesday, Thursday, Saturday, and Sunday.
- Pantoprazole: 20 mg taken twice daily.
- Phenytoin: 100 mg, with 1 capsule in the morning and 2 capsules in the evening.
- Propranolol: 20 mg, with half a tablet once daily for the first 2 days, followed by 1 full tablet daily.

Step 3: Qualitative One-on-One Interview

Following usability testing, participants engaged in one-on-one interviews to provide feedback on the devices they tested. These interviews explored their experiences, challenges, and perceptions of each device's usability and features, offering qualitative insights to complement the quantitative data.

Outcome Measure—Proportion of Subtask Success

The proportion of subtask success is a measure that shows how often participants successfully completed a specific subtask when using a device. This was calculated using the equation:

$$\text{Proportion of subtask success} = \frac{\text{number of subtasks}}{\text{number of participants}} \times \frac{\text{number of subtasks}}{\text{number of subtasks}}$$

Data Collection and Statistical Analysis

Data Management and Storage

All collected data from this study were securely managed and organized using the REDCap (Research Electronic Data Capture) platform (version 15.1.2; Vanderbilt University). This included participant demographics, outputs from the tools used to measure barriers to MMC, and the detailed data recorded in the evaluation sheets for each device tested. R Studio (version 4.4.1; Posit Software, PBC) was used to conduct data analysis.

Data Analysis Using PPM

The data analysis methodology used PPM to predict the success rates of subtasks for participants based on their unique characteristics. Unlike traditional modeling, where we use large and heterogeneous data to train the model, with personalized prediction modeling, we identify and analyze the data that are similar to the new participant of interest, whose outcome is to be predicted. By training the model on the subset of the data containing only the participants that are similar to the participant of interest, we aim to provide a customized prediction with improved predictive performance [57]. In other words, a new model is fit for each participant to predict their proportion of success.

There are a few key details of interest when implementing the model with PPM. The first is quantifying how similar one participant is to another participant, and the second is identifying the best subpopulation proportion (denoted as m) to train the unique model for each participant—in other words, how many similar participants do we use to train a predictive model for the participant of interest.

Our approach to measuring similarity is using the cosine similarity metric (CSM). CSM is intuitive to interpret, standardizes the similarity between −1 and 1, and can handle mixed data types that contain both categorical and continuous predictors. The data for each participant, such as their demographic and impairment information, are represented as a

XSL•FO
RenderX

Euclidean vector of predictors in a multidimensional feature space, and CSM is used to measure the angle between the vectors for 2 participants. CSM quantifies the similarity between 2 participants as a cosine of an angle with values between −1 and 1, where −1 represents minimum similarity, and 1 represents maximum similarity [53]. For example, if CSM between participant A and B is −0.1, and CSM between participant A and C is 0.6, then C is more similar to A than B is to A.

With PPM, we train a model on a personalized dataset that contains only the most similar participants to the participant of interest. Then, it is crucial to determine what percentage of the most similar participants from the available data we will use to train the model (denoted as m). We find the optimal subproportion of the available data by tuning for this parameter using k-fold cross-validation (CV). The steps are as follows, adapted from previous literature on PPM [57]:

1. Randomly split the available data into 9 folds. The number of folds is 9 to allow the 117 participants to be distributed evenly along all the folds. One of the folds is the test data, and the rest of the data will be the training data.
2. For participant j in the test data, calculate CSM between participant j and every participant in the training data. Fit a generalized linear model (GLM) with a binomial logistic framework using the top m×100% of the most similar participants in the training data, and calculate the predicted proportion of subtask success for participant j using this model. Predictors in this model include demographic and clinical variables such as age, sex, number of medications, cognitive and physical scores, vision metrics (SMAT and DLTV scores), SEAMS scores, and hearing impairment.
3. Repeat step 2 for every participant in the test data. In other words, we are fitting a unique GLM with a logistic model using a tailored dataset containing only the top m×100% of the most similar participants from the training data for each participant in the testing data. Then, calculate the mean square error (MSE) to get the predictive performance for this fold. Predictors for each GLM included demographic and clinical variables such as age, sex, number of medications, cognitive and physical scores, vision metrics (SMAT and DLTV scores), SEAMS scores, and hearing impairment. Two models per participant were compared, one using SMAT vision scores and the other using DLTV vision scores, and success probabilities for subtasks were estimated based on model coefficients.
4. Repeat steps 2 - 3 such that each fold becomes the test data, and the rest of the available data becomes the training data. Average over the MSEs calculated to get the predictive performance of the model for a specific m value.
5. Repeat steps 1 - 4 for values of m ranging from 0.20 to 1.00, in increments of 0.05, to examine the effects of different matching percentages. Smaller subproportions (sub 0.2) were not explored because the resulting dataset would not be heterogeneous enough across all the features to fit the regression model. Note that as m increases, we have larger data to train the predictive model, but we also introduce more noise in the data as a trade-off. We denote m that gives the smallest MSE value as the optimum subpopulation proportion. When m=1, we are using a global (or nonpersonalized) model where we use the exact same data to fit the GLM model for each participant, because we are using all available dataset.

The above CV study can be used to tune for the optimal matching percentage m on the combined dataset of 117 participants. The recommended matching subproportion can be used when predicting the proportion of success at a subtask of a new test participant. However, this simulation cannot be used to estimate the generalization performance on unseen data since all of the data appear in at least 1 of the training folds. This is typically why a separate test set is held out from the training process. This is the data that the models have not trained on; hence, the model's performance on this data is a good estimate of the model's accuracy on unseen data. The dataset of 117 participants is split into the 96 participants collected up to 2024, which would be the training dataset, and the 21 participants collected in 2025, which would be the test data. This would give us an estimate of the performance of PPMs on unseen data and also provide us with benchmarks to outperform for newer models.

In both the above simulations (CV and model validation), model performance is quantified by the squared error of the model in predicting the number of successful attempts at a subtask. This evaluation metric allows us to view the problem in the lens of regression rather than as a classification task where area under the curve or F_1-scores would be used. The choice is made because it is challenging to define the threshold for success at a subtask since multiple attempts are being made at subtasks, and an additional threshold hyperparameter that varies by subtask would be necessary.

The GLMs: Logistic Regression

Personalized subsets of participants were used to train GLMs for predicting subtask success rates. Two separate models were developed for each participant: (1) SMAT-based model using SMAT vision scores and (2) DLTV-based model using DLTV summary vision scores.

For each participant, a personalized GLM was fitted to their subset of training data identified via the CSM methodology. Predictors in the logistic regression model included demographics, clinical history, cognitive, physical, sensory, hearing, motivation, and environmental factor scores. Subtask-specific indicators were incorporated as covariates. Task count weights were applied to adjust for variations in subtask performance across participants.

The GLMs were structured as follows.

SMAT-based model:

DLTV-based model:

where for a participant i, Y_i is the proportion of success of a subtask, age_i is the age, sex_i is the sex, $mednum_i$ is the number of medications, $medhist_i$ is the number of medical history, cog_i is the cognitive score, phy_i is the physical score, $SMATvis_i$ is

XSL•FO
RenderX

the SMAT vision score, $DLTVsumvis_i$ is the DLTV summary vision score, $SEAMS_i$ is the SEAMS score, $MPEDbusy_i$ is the MPED busyness score, $MPEDroutine_i$ is the MPED routine score, $subtask_i$ is a single subtask completed, and $hear_i$ is an indicator variable for hearing impairment.

Ethical Considerations

This study was reviewed by and received approval from the University of Waterloo Office of Research Ethics (ORE##45203). All participants were informed of the study and provided consent before enrolling. Participant confidentiality was maintained by assigning unique participant ID numbers instead of names or identifiable information. All data were deidentified and securely stored. Participants received remuneration for their participation in the study. This study was conducted at multiple locations, including the University of Waterloo School of Pharmacy in Kitchener, Ontario, Canada; various Schlegel Villages across Ontario; and the KW Seniors Day Program in Kitchener, Ontario, Canada.

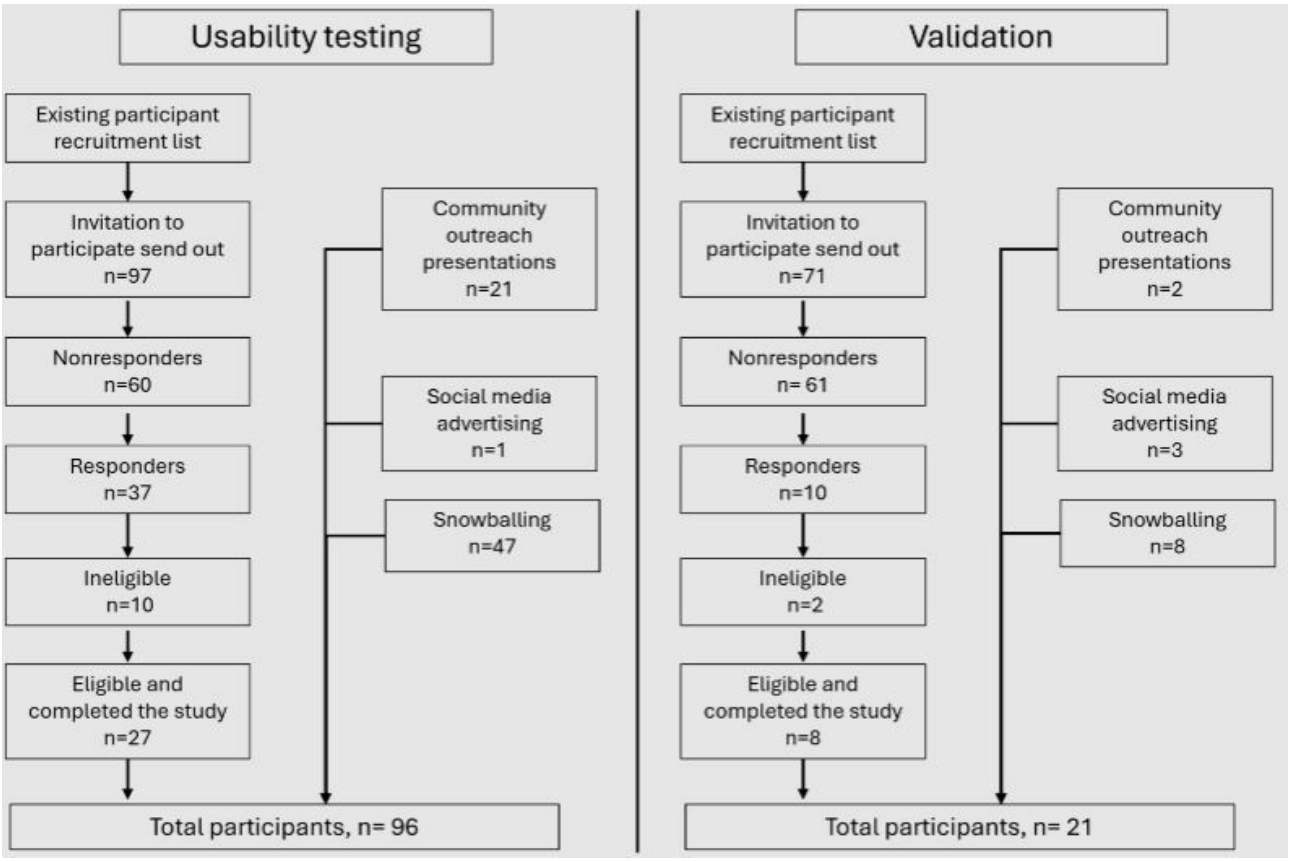
Results

Participant Demographics Characteristics

The study included a total of 117 participants (mean age SD 74.6, SD 7.9 years), including 96 participants for the usability testing and 21 participants for the validation, all with varying cognitive, physical, sensory, motivational, and environmental abilities.

Although we initially planned to recruit 98 participants for the usability testing and 29 participants for the validation, recruitment and inclusion challenges limited enrollment to 96 and 21 participants, respectively. One of our primary goals while recruiting was to ensure that each of the 13 devices was tested by at least 5 participants from each impairment group, which was achieved with the final usability sample and therefore considered sufficient for this component of the study. Additionally, considering that participants tested up to 7 devices, and that each device was treated as a distinct product with its own subtasks, the resulting dataset provided a sufficiently large number of observations. For validation, although the final sample size was smaller than planned, it allowed for an internal CV of the models. Further validation with larger samples is planned in future studies. Figure 1 presents a flow diagram summarizing participant recruitment and inclusion across the usability testing and validation. As the study involved a single assessment session (with a small number of participants completing 2 sessions), there were no participant attribution or dropouts during the study.

Figure 1. Participant recruitment and inclusion across the usability testing and validation.



Barrier Characteristics of Participants

Among the various barriers identified in the usability study, cognitive barriers, defined as low or moderate cognitive scores, were present in 21.87% (21/96) of participants, as determined using the SMAT. Physical barriers, defined as low or moderate to low physical scores, were observed in 38.54% (37/96) of participants. Vision barriers were also notable, with 23.95% (23/96) scoring low or moderate to low on the SMAT vision

XSL•FO
RenderX

scale, and 38.54% (37/96) reporting impairments in DLTV, as assessed using the DLTV. Hearing barriers, evaluated with the whisper test, were present in 57.29% (55/96) of participants. Additionally, motivation and environmental challenges were reported by 35.41% (34/96) of the group (Table 2).

XSL•FO
RenderX

Table . Demographic, clinical, and barrier characteristics of study participants.

Variable	Usability sample (n=96)	Additional validation sample (n=21)
Age (years)		
Mean (SD)	75.11 (7.70)	73.42 (8.56)
Range	61 - 95	60 - 84
Sex, n (%)		
Male	36 (37.50)	14 (66.67)
Female	60(62.50)	7 (33.33)
Medication use (yes), n (%)	86 (89.58)	20 (95.24)
Level of education, n (%)		
Masters or doctoral or professional degree	26 (27.08)	6 (28.57)
Bachelor's degree	27 (28.13)	9 (42.86)
Nonuniversity diploma	17 (17.70)	3 (14.28)
Trade certificate or diploma	1 (1.04)	1 (4.76)
High school	24 (25)	2 (9.52)
Below high school	1 (1.04)	0 (0)
Place of residence, n (%)		
Home	83 (86.46)	19 (90.48)
Retirement home	13 (13.54)	2 (9.52)
Self-reported medical history, n (%)		
Cardiovascular	56 (58.33)	0 (0)
Metabolic and endocrine	37 (38.54)	1 (4.76)
Musculoskeletal	26 (27.08)	11 (52.38)
Eye-related conditions	18 (18.75)	0 (0)
Oral and gastrointestinal	16 (16.67)	0 (0)
Neurological	12 (12.50)	3 (14.28)
Respiratory	10 (10.42)	0 (0)
Mental health	8 (8.33)	2 (9.52)
Renal and urogenital	7 (7.29)	2 (9.52)
Cancer and neoplasms	5 (5.21)	3 (14.28)
Stroke	5 (5.21)	4 (19.05)
None	5 (5.21)	4 (19.05)
Hematological diseases	4 (4.17)	0 (0)
Congenital disorders	2 (2.08)	3 (14.28)
Inflammatory and immune system	2 (2.08)	0 (0)
Injuries and accidents	2 (2.08)	1 (4.76)
Ear-related conditions	1 (1.04)	1 (4.76)
Infection	1 (1.04)	0 (0)
Reproductive health	1 (1.04)	3 (14.28)
Skin	1 (1.04)	4 (19.04)
Other	1 (1.04)	0 (0)
Medication management strategies, n (%)		
Pillbox	43 (44.79)	7 (33.33)
Alarm beeper	11 (11.46)	2 (9.52)

XSL•FO
RenderX

Variable	Usability sample (n=96)	Additional validation sample (n=21)
Blister pack	10 (10.42)	4 (19.05)
Someone else reminds me	8 (8.33)	1 (4.76)
Others	3 (3.12)	6 (28.57)
Medication calendar	2 (2.08)	1 (4.76)
Barriers		
Cognitive barrier—Self-Medication Assessment Tool, n (%)		
High cognitive scores (90% or greater)	70 (72.92)	14 (66.67)
Relatively high cognitive score (80% or greater)	5 (5.21)	1 (4.76)
Moderate cognitive score (70% to 80%)	7 (7.29)	2 (9.52)
Low cognitive score (approximately 69% or less)	14 (14.58)	4 (19.05)
Physical barrier—Self-Medication Assessment Tool, n (%)		
High physical scores (90% or greater)	59 (61.46)	15 (71.43)
Moderate to low physical score (76% to 84%)	11 (11.46)	3 (14.28)
Low physical score (approximately 75% or less)	26 (27.08)	3 (14.28)
Vision barrier—Self-Medication Assessment Tool, n (%)		
High vision scores (90% or greater)	73 (76.04)	18 (85.71)
Moderate to low vision score (76% to 84%)	4 (4.17)	1 (4.76)
Low vision score (approximately 75% or less)	19 (19.79)	2 (9.52)
Vision barrier—Daily Living Tasks Dependent on Vision[a], n (%)		
Impairment present	37 (38.54)	10 (47.62)
Hearing barrier—whisper test, n (%)		
Impairment present (both ears)	55 (57.29)	12 (57.14)
Motivation barrier—Self-Efficacy for Appropriate Medication Use Scale[b], n (%)	34 (35.42)	3 (14.28)
Environmental barrier—Martin and Park Environmental Demands[c] questionnaire, n (%)	34 (35.42)	5 (23.81)

[a]Total score <79: vision barrier present.

[b]Total score <40 low self-efficacy: motivational barrier present.

[c]Busyness subscale score ≥15—greater busyness, routine subscale score <16—less routine: environmental barrier present.

In the additional validation sample, similar patterns were observed. Cognitive barriers were present in 28.57% (6/21) of participants, physical barriers in 28.57% (6/21), and vision barriers in 14.28% (3/21) based on SMAT scores. Impairments in vision-dependent daily living tasks were reported by 47.62% (10/21) of participants, and hearing barriers were present in 57.14% (12/21). Motivational and environmental barriers were less prevalent in the validation sample, reported by 14.28% (3/21) and 23.81% (5/21) of participants, respectively (Table 2).

Table 2 presents the demographic, clinical, and barrier-related characteristics of the study participants.

PPM Outcomes

We begin by illustrating our choice of the tuning parameters and how they impact each of the PPMs. Figure 2 indicates that m=0.25 is the optimal matching percentage in this scenario for both the SMAT-based and DLTV-based models. Note that the averaged MSE decreases from m=0.2 to m=0.25 in Figure 2. Therefore, it is not necessary to modify the matching strategy to accommodate PPMs to be trained on the dataset with lower matching percentages (less than 20%) since the optimal matching subproportion as illustrated by Figure 2 is greater than m=0.2. The fairly volatile nature of the trends in Figure 2 suggests that tuning the matching parameter is beneficial, as there is a large difference between the MSE at m=0.25 compared

XSL•FO
RenderX

to the MSE at m=0.95, where performance is even worse than not using any personalization at all (m=1.0).

Figure 2. Mean square error of the DLTV-based and SMAT-based models averaged over all k-iterations of the cross-validation study as a function of the matching percentage m. DLTV: Daily Living Tasks Dependent on Vision; SMAT: Self-Medication Assessment Tool.



Paired 2-sided Wilcoxon signed rank tests were conducted to determine if there is a statistically significant difference in the median MSE values (along the k-folds) between the personalized models at m=0.25 and their nonpersonalized counterparts (m=1.0). These tests (DLTV: P=.43; SMAT: P=.50) do not show statistical significance. Hence, even though 25% matching is optimal in Figure 2, the MSE populations may not be statistically different enough to necessitate tuning for the matching parameter.

If the datasets become large enough that matching each new participant to m×100% of the training dataset and refitting a new model becomes a computational bottleneck, skipping the personalization step should not statistically degrade model performance. However, when computation is not a constraint and PPMs can be used, the optimal matching percentage with respect to the current data is m=0.25.

Model Validation

The CV procedure as described in the previous section was carried out on the 96 participants but with 8 folds instead of 9 folds, and it was found that the optimal matching percentage was 0.2 for both the DLTV-based and SMAT-based models. On the test set, the SMAT-based model with a matching parameter of 0.20 achieves an MSE of 0.89, and the DLTV-based model with a matching parameter of 0.2 achieves an MSE of 1.16. In other words, the SMAT-based model is off by 0.94 subtasks when predicting the correct number of subtasks a participant has successfully completed, while the DLTV-based model with a matching parameter of 0.20 is off by 1.1 subtasks. Note that these results do not reflect the best possible performance of the models on the test set because the matching parameter, while optimal for the training set, is not necessarily optimal for the test set. The nonpersonalized models (m=1.0) perform much better with an MSE of 0.726 and 0.815 for the SMAT and DLTV-based models, respectively. This does raise the question of whether the optimal matching percentage m transfers over to unseen data that might have a different distribution compared to the training data.

To contextualize these results, note that the average proportion of success in the training data collected in 2024 is 0.87. Consider a naive model that always predicts a fixed proportion of success regardless of the participant's features. Due to the left-skewed nature of the success variable, the naive model predicting a success rate of 100% would do much better than a naive model that predicts a fixed success rate of 0%. To be more precise, the naive model that minimizes MSE on a dataset predicts a constant probability equal to the inner product of the success counts and the corresponding number of attempts across subtasks normalized by the squared l^2-norm of the number of attempts at a subtask. Since the test data are unseen, the naive model can instead compute this ratio on the training set and apply it on unseen data. Numerically, this is a prediction of 0.84 resulting in an MSE of 1.21. The optimal naive model is on average 1.1 subtasks off in predicting the correct number of subtasks.

Overall, the SMAT-based model with m=0.2 is on average more accurate by 0.07 subtasks in prediction, while the DLVT-based model with m=0.2 performs similarly to the naive baseline. In contrast, in the nonpersonalized setting (m=1.0), the SMAT-based model improves accuracy by 0.25 subtasks, and

XSL•FO
RenderX

the DLVT-based model by 0.2 subtasks relative to the naive model. The naive model's MSE of approximately 1.1 subtasks serves as a useful benchmark for future model development and validation studies.

It is possible that the training-testing split is suboptimal if participants recruited during the first phase of data collection (2023 - 2024) systematically differ from those recruited during the second phase (2025). To assess this, statistical tests were conducted on all participant features used in the model to evaluate differences between the 2 cohorts. After adjusting for multiple-comparison error, no statistically significant differences were observed for any feature. This finding supports the chosen training-testing split. However, given the unequal and relatively small sample sizes of the 2 cohorts (96 vs 21), these tests may have been underpowered. Consequently, a random partition of the combined dataset (n=117) may have provided a more statistically robust assessment of both the benefits of personalization and the performance of naive models.

Model Specifications

Unlike conventional prediction models that estimate a single set of regression coefficients applicable to all individuals, PPM generates individualized models tailored to each participant. Specifically, each participant's model is trained on a personalized dataset comprising observations most similar to that participant. Consequently, for the 117 participants included in this study, 117 distinct logistic regression models were estimated, each with its own set of regression coefficients and intercept. Because model parameters are individualized by design, there is no single, global set of regression coefficients, odds ratios, or CIs that can be meaningfully reported. Instead, model specification is defined procedurally rather than through fixed parameter estimates.

To generate predictions for a new individual, we first apply steps 1 - 5 of the PPM procedure (see Data Analysis Using PPM section) to the full available dataset to determine the optimal subpopulation proportion, m. This proportion defines the size of the personalized training set. We then identify the m×100% of observations most similar to the new individual according to the predefined similarity metric and fit a logistic regression model on this subset. The resulting personalized model can be used to estimate the probability of subtask success for that individual and to inform individualized recommendations.

Discussion

Principal Findings

This study was conducted to develop and internally validate a personalized predictive model to estimate subtask success of MATs among older adults with diverse cognitive, physical, sensory, motivational, and environmental abilities. Overall, the findings show that it is feasible to predict subtask success of MATs using measurable functional and demographic characteristics, and that such models can perform better than naive prediction models. These results demonstrate that individual differences meaningfully influence how older adults interact with MATs and that these differences can be quantified and modeled.

Research evaluating the usability of MATs for older adults, while considering their specific impairments, remains limited [28,35]. One study by Woo [58] examined usability through a small-scale device trial involving 2 older adult participants who tested a single device over a 2-day period. The study assessed user perceptions, adherence rates, and usability concerns through observations and interviews. In this study, the first participant was an 82-year-old female with mild cognitive impairment, arthritis, osteoporosis, and mobility challenges, requiring the use of a 4-wheeled walker. The second participant was an 84-year-old male with mild to moderate cognitive impairment, hypertension, and mobility impairments. While both achieved high adherence rates (100% and 90%, respectively), they encountered usability challenges. One participant struggled to grasp pills and had to tip the device over to retrieve medication, while the other faced difficulty closing pill compartments due to manual dexterity impairments. This underscores the importance of examining specific interaction failures related to device features and users' functional abilities.

By focusing on subtasks rather than overall device success, this study addresses an important gap in MAT usability research. As we previously noted, subtasks represent features of the device, the specific actions required to use a device, such as pressing buttons, opening compartments, aligning parts, or reading visual indicators. Evaluating success at this level provides a more detailed and clinically meaningful understanding of usability than global device ratings, which may mask specific interaction challenges. This subtask-based approach allows for a clearer link between user abilities and the demands imposed by individual device features. The decision to model success at the subtask level is supported by assistive technology and human factors literature, which emphasizes that devices should be matched to users based on their functional limitations rather than by device type or overall complexity [59,60]. For example, individuals with mild hand weakness may succeed with simple, low-technology solutions, whereas those with more severe impairments may require devices that reduce physical or cognitive effort by automating certain actions [60]. MATs vary widely in the features, functionalities, effort, and understanding they require, and identifying subtask-level success makes it possible to identify which specific features could create barriers for particular users as well as which features are more suitable. Human factors research further shows that usability failures often occur at discrete interaction points, especially as device complexity increases [61-64]. By treating subtasks as representative device features, this study supports a more user-centered approach to MAT selection that recognizes assistive technologies as essential supports that must align closely with users' abilities for successful use, acceptability, and adoption [61-65].

With respect to modeling strategy, both personalized and nonpersonalized GLMs improved prediction accuracy compared to a naive baseline that assumed uniform subtask success across users. This finding suggests that incorporating individual-level characteristics, such as cognitive function, physical ability, vision, hearing, motivation, and environmental factors, adds

XSL•FO
RenderX

meaningful explanatory value when predicting how older adults interact with specific device subtasks. These results support the broader view that usability arises from the interaction between user characteristics and device features rather than by either factor alone. Although PPM is conceptually appealing, its advantages over nonpersonalized (global or broader) models were not statistically significant in this study. Personalized models showed some improvement during CV but did not significantly outperform nonpersonalized models overall, and in the held-out validation cohort, nonpersonalized models demonstrated better performance. Similar patterns have been reported in prior work. Ng et al [66] showed that personalized predictive models can outperform global models when trained on large, dense datasets with sufficient numbers of clinically similar cases, but also emphasized that reliable personalization depends on stable similarity estimates and adequate neighborhood size. When these conditions are not met, global models that leverage the full dataset may generalize more effectively. Evidence from outside health care further supports this interpretation; for example, in sports performance research, Ahsan et al [67] reported that while personalized modeling can offer targeted insights, broader predictive models are often more robust and easier to maintain when data are limited or heterogeneous. Together, these findings suggest that the lack of statistical significance between the models in this study likely reflects constraints related to sample size and population heterogeneity rather than a lack of conceptual value of personalized modeling itself.

The validation results highlight the importance of evaluating model performance beyond internal training. The stronger performance of nonpersonalized models in the validation cohort suggests that matching parameters optimized during training may not fully transfer to unseen data, particularly when sample sizes are small. Although no statistically significant differences were observed between recruitment phases, the smaller validation cohort may have limited the ability to detect subtle differences. These findings underscore the need for external validation in larger and more diverse populations before using PPM approaches in real-world clinical settings.

An additional important finding was that models incorporating SMAT-based (performance-based) vision measures consistently outperformed those using DLTV-based (self-reported) vision scores. This suggests that task-specific functional assessments may be more informative for predicting technology interaction than broader measures of daily living performance. However, at the same time, it is also important to consider the practical advantages of using task-based measures compared with self-reported measures [68].

From a clinical perspective, this modeling framework has the potential to support more informed MAT selection by predicting which device features an older adult is most likely to use successfully or struggle with. By estimating subtask-level success, the model can help clinicians, caregivers, and users anticipate whether specific device features are well matched to an individual's cognitive, physical, sensory, and environmental abilities. This enables more informed decision-making when selecting MATs, particularly by avoiding devices with features that may be difficult or unsafe for a given user. Because medication management is a safety-critical activity, mismatches between user abilities and device demands can increase the risk of medication errors and nonadherence. Selecting technologies that align with users' capabilities may reduce trial-and-error device adoption, minimize frustration, support correct device use, and ultimately could improve medication adherence, safety, and therapeutic outcomes.

Strengths

This study has several strengths, including a diverse sample of 117 older adults with varying cognitive, physical, sensory, motivational, and environmental barriers. Given the exclusion of older adults with impairments in many studies, this inclusion improves the study's relevance and applicability. Additionally, the study tested both electronic and smart MATs, allowing for a comprehensive comparison of their effectiveness. Another key strength is that it addresses multiple barriers affecting medication self-management, beyond just physical or cognitive impairments. The study's use of validated assessment tools (eg, SMAT and whisper test) strengthens its methodological quality, ensuring reliable and meaningful findings. Finally, the subtask-based success approach provided detailed, feature-level insights into usability, supporting more clinically meaningful interpretations of device-user interactions.

Limitations

Despite its strengths, the study has some limitations. Testing environments may not fully replicate real-world conditions, potentially affecting the generalizability of findings. Additionally, while the sample is diverse, it may not fully capture all cultural and socioeconomic factors influencing MAT use. The variability in participants' technological literacy and health status could also introduce biases, impacting the study's applicability to the broader older adult population. A further limitation is related to sample size planning and model complexity. The initial sample size estimation for the usability phase was guided by Green's formula based on 6 primary impairment domains. However, the final predictive models included a larger set of predictors, including demographic, medication-related variables, multiple functional scores, environmental factors, hearing scores, and subtask indicators. While each personalized model was trained on a subset of similar participants (approximately 20% - 25% of the training sample), the ratio of observations to predictors may still be limited, increasing the risk of overfitting and reducing coefficient stability. This constraint likely contributed to the modest performance gains observed for personalized models compared with nonpersonalized models and highlighted the need for larger datasets to fully support more complex personalized modeling approaches.

Future Recommendations

Future research should address these limitations by conducting real-world studies in home and community settings and including a more representative demographic sample. Additionally, developing a standardized, comprehensive assessment tool for measuring all medication management barriers would improve consistency and applicability. Further studies could explore the longitudinal impacts of MATs on

XSL•FO
RenderX

long-term adherence, hospitalization rates, and health care costs. Research should also focus on developing clinical decision-support tools that integrate usability metrics and patient-specific barriers, helping health care providers tailor recommendations. Examining the learnability and long-term engagement of older adults with these technologies is also important, as older adults may experience technology fatigue or changes in ability that affect long-term use. Finally, external validation of the predictive models in larger, independent, and more diverse populations is essential to assess generalizability and robustness. Future work should evaluate model performance across different health care systems, cultural contexts, and device types and examine whether personalized modeling approaches yield greater benefits as datasets grow, and similarity estimates become more stable. Such efforts can be beneficial for translating predictive usability modeling into routine clinical practice.

Conclusions

This study developed and internally validated a personalized predictive model, demonstrating the feasibility of predicting subtask success of MATs among older adults using demographic, clinical, and functional characteristics. Both personalized and nonpersonalized models outperformed the naive model, highlighting the importance of individual differences in technology usability. Although personalization showed limited added benefit in this dataset, the subtask-focused approach offers clinically meaningful findings to support evidence-informed selection of medication technologies and thereby improving adherence outcomes.

Acknowledgments

The authors would like to extend special thanks to Arshia Chauhan, Nischal Ghookal, Prapti Choudhuri, Hawa Patel, Tanaya Sharma, Victoria Nguyen, Karyman Ahmed Fawzy Ghanem, and Sabiha Lakhi for their valuable support with data collection. This work was presented as a poster at the 2025 Annual Scientific Meeting of the American Geriatrics Society.

Funding

This project was funded by Aging in Place program with National Research Council and Canadian Institutes of Health Research (funding reference-184372) with the Government of Canada.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

BB and TP contributed to all aspects of the study, including study concept and design, acquisition of subjects and/or data, analysis and interpretation of data, and preparation of the manuscript. GE contributed to data acquisition and data entry. MK, SB, and JG contributed to data analysis. SP, IH, and RS contributed to the acquisition of data. PKP, AM, SJTG, FC, LL, CB, and RG contributed to the study concept and design. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

TP received grants from National Research Council, Canadian Institutes of Health Research, Jones Healthcare Group, and Custom Health; travel support from National Research Council and Canadian Institutes of Health Research; and contribution of medical adherence technology from Jones Healthcare Group and Custom Health. The other authors report no conflicts of interest.

Multimedia Appendix 1
List of medication adherence technologies tested.
[DOCX File, 28 KB - aging_v9i1e84616_app1.docx]

Checklist 1
TRIPOD checklist.
[PDF File, 603 KB - aging_v9i1e84616_app2.pdf]

References

1. Griffith LE, Gruneir A, Fisher K, et al. Insights on multimorbidity and associated health service use and costs from three population-based studies of older adults in Ontario with diabetes, dementia and stroke. BMC Health Serv Res 2019 May 16;19(1):313. [doi: 10.1186/s12913-019-4149-3] [Medline: 31096989]
2. Statistics Canada homepage. Government of Canada. URL: https://www.statcan.gc.ca/en/record/2015 [accessed 2022-11-07]
3. Fausset CB, Kelly AJ, Rogers WA, Fisk AD. Challenges to aging in place: understanding home maintenance difficulties. J Hous Elderly 2011;25(2):125-141. [doi: 10.1080/02763893.2011.571105] [Medline: 22072843]

XSL•FO
RenderX

4. Cutchin MP. The process of mediated aging-in-place: a theoretically and empirically based model. Soc Sci Med 2003 Sep;57(6):1077-1090. [doi: 10.1016/s0277-9536(02)00486-0] [Medline: 12878107]
5. Kunstler R. Therapeutic recreation in naturally occurring retirement community (NORC): benefiting “aging in place”. Ther Recreat J 2002;36:186-202 [FREE Full text]
6. Maddigan SL, Farris KB, Keating N, Wiens CA, Johnson JA. Predictors of older adults’ capacity for medication management in a self-medication program: a retrospective chart review. J Aging Health 2003 May;15(2):332-352. [doi: 10.1177/0898264303251893] [Medline: 12795276]
7. Mira JJ, Lorenzo S, Guilabert M, Navarro I, Pérez-Jover V. A systematic review of patient medication error on self-administering medication at home. Expert Opin Drug Saf 2015 Jun;14(6):815-838. [doi: 10.1517/14740338.2015.1026326] [Medline: 25774444]
8. Notenboom K, Beers E, van Riet-Nales DA, et al. Practical problems with medication use that older people experience: a qualitative study. J Am Geriatr Soc 2014 Dec;62(12):2339-2344. [doi: 10.1111/jgs.13126] [Medline: 25516030]
9. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
10. Elliott RA, Goeman D, Beanland C, Koch S. Ability of older people with dementia or cognitive impairment to manage medicine regimens: a narrative review. Curr Clin Pharmacol 2015;10(3):213-221. [doi: 10.2174/1574884710666150812141525] [Medline: 26265487]
11. Edelberg HK, Shallenberger E, Wei JY. Medication management capacity in highly functioning community-living older adults: detection of early deficits. J Am Geriatr Soc 1999 May;47(5):592-596. [doi: 10.1111/j.1532-5415.1999.tb02574.x] [Medline: 10323653]
12. Sino CGM, Sietzema M, Egberts TCG, Schuurmans MJ. Medication management capacity in relation to cognition and self-management skills in older people on polypharmacy. J Nutr Health Aging 2014 Jan;18(1):44-49. [doi: 10.1007/s12603-013-0359-2] [Medline: 24402388]
13. Beckman A, Bernsten C, Parker MG, Thorslund M, Fastbom J. The difficulty of opening medicine containers in old age: a population-based study. Pharm World Sci 2005 Oct;27(5):393-398. [doi: 10.1007/s11096-005-7903-z] [Medline: 16341746]
14. Philbert D, Notenboom K, Bouvy ML, van Geffen ECG. Problems experienced by older people when opening medicine packaging. Int J Pharm Pract 2014 Jun;22(3):200-204. [doi: 10.1111/ijpp.12070] [Medline: 24025086]
15. Sam T. Developing drug administration devices for geriatric use. In: Stegemann S, editor. Developing Drug Products in an Aging Society: Springer; 2016:543-560.
16. Privitera AJ. Sensation and perception. In: Biswas-Diener R, Diener E, editors. Noba Textbook Series: Psychology: DEF Publishers; 2023. URL: http://noba.to/xgk3ajhy [accessed 2026-04-02]
17. Mick P, Foley DM, Lin FR. Hearing loss is associated with poorer ratings of patient-physician communication and healthcare quality. J Am Geriatr Soc 2014 Nov;62(11):2207-2209. [doi: 10.1111/jgs.13113] [Medline: 25413192]
18. Smith SG, Curtis LM, O’Conor R, Federman AD, Wolf MS. ABCs or 123s? The independent contributions of literacy and numeracy skills on health task performance among older adults. Patient Educ Couns 2015 Aug;98(8):991-997. [doi: 10.1016/j.pec.2015.04.007] [Medline: 25936579]
19. Hughes RG, editor. Patient Safety and Quality: An Evidence-Based Handbook for Nurses: Agency for Healthcare Research and Quality; 2008.
20. Son YJ, Won MH. Depression and medication adherence among older Korean patients with hypertension: mediating role of self-efficacy. Int J Nurs Pract 2017 Jun;23(3):e12525. [doi: 10.1111/ijn.12525] [Medline: 28194846]
21. Martin M, Park DC. The Martin and Park Environmental Demands (MPED) questionnaire: psychometric properties of a brief instrument to measure self-reported environmental demands. Aging Clin Exp Res 2003 Feb;15(1):77-82. [doi: 10.1007/BF03324483] [Medline: 12841422]
22. Faisal S, Ivo J, Tennant R, et al. Integration of a smart multidose blister package for medication intake: a mixed method ethnographic informed study of older adults with chronic diseases. PLoS ONE 2022;17(1):e0262012. [doi: 10.1371/journal.pone.0262012] [Medline: 35061773]
23. Somerville E, Massey K, Keglovits M, et al. Scoring, clinical utility, and psychometric properties of the In-HOME medication management performance evaluation (HOME-Rx). Am J Occup Ther 2019;73(2):7302205060p1-7302205060p8. [doi: 10.5014/ajot.2019.029793] [Medline: 30915967]
24. Jindai K, Nielson CM, Vorderstrasse BA, Quiñones AR. Multimorbidity and functional limitations among adults 65 or older, NHANES 2005-2012. Prev Chronic Dis 2016 Nov 3;13:E151. [doi: 10.5888/pcd13.160174] [Medline: 27809419]
25. Sharma P, Maurya P, Muhammad T. Number of chronic conditions and associated functional limitations among older adults: cross-sectional findings from the longitudinal aging study in India. BMC Geriatr 2021 Nov 23;21(1):664. [doi: 10.1186/s12877-021-02620-0] [Medline: 34814856]
26. Patel T, Ivo J, McDougall A, et al. Development of a clinician guide for electronic medication adherence products in older adults. Can Pharm J 2022 Mar;155(2):119-127. [doi: 10.1177/17151635221074977] [Medline: 35300022]
27. Faisal S, Samoth D, Aslam Y, et al. Key features of smart medication adherence products: updated scoping review. JMIR Aging 2023 Dec 19;6:e50990. [doi: 10.2196/50990] [Medline: 38113067]

XSL•FO
RenderX

28. Mason M, Cho Y, Rayo J, Gong Y, Harris M, Jiang Y. Technologies for medication adherence monitoring and technology assessment criteria: narrative review. JMIR Mhealth Uhealth 2022 Mar 10;10(3):e35157. [doi: 10.2196/35157] [Medline: 35266873]
29. Mokhtari M, Abdulrazak B, Aloulou H. Smart homes and health telematics, designing a better future: urban assisted living. Presented at: 16th International Conference, ICOST 2018; Jul 10-12, 2018. [doi: 10.1007/978-3-319-94523-1]
30. Brath H, Morak J, Kästenbauer T, et al. Mobile health (mHealth) based medication adherence measurement—a pilot trial using electronic blisters in diabetes patients. Br J Clin Pharmacol 2013 Sep;76(Suppl 1):47-55. [doi: 10.1111/bcp.12184] [Medline: 24007452]
31. Foster JM, Usherwood T, Smith L, et al. Inhaler reminders improve adherence with controller treatment in primary care patients with asthma. J Allergy Clin Immunol 2014 Dec;134(6):1260-1268. [doi: 10.1016/j.jaci.2014.05.041] [Medline: 25062783]
32. Baltes PB, Smith J. New frontiers in the future of aging: from successful aging of the young old to the dilemmas of the fourth age. Gerontology 2003;49(2):123-135. [doi: 10.1159/000067946] [Medline: 12574672]
33. Czaja SJ, Lee CC. Designing computer systems for older adults. In: Jacko JA, Sears A, editors. The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications: Erlbaum; 2003:413-427.
34. Lindenberger U, Lövdén M, Schellenbach M, Li SC, Krüger A. Psychological principles of successful aging technologies: a mini-review. Gerontology 2008;54(1):59-68. [doi: 10.1159/000116114] [Medline: 18259095]
35. Patel T, Ivo J, Faisal S, et al. A prospective study of usability and workload of electronic medication adherence products by older adults, caregivers, and health care providers. J Med Internet Res 2020 Jun 2;22(6):e18073. [doi: 10.2196/18073] [Medline: 32348292]
36. Ligons FM, Mello-Thoms C, Handler SM, Romagnoli KM, Hochheiser H. Assessing the impact of cognitive impairment on the usability of an electronic medication delivery unit in an assisted living population. Int J Med Inform 2014 Nov;83(11):841-848. [doi: 10.1016/j.ijmedinf.2014.07.004] [Medline: 25153770]
37. Albert B, Tullis T. Measuring the User Experience, 1st edition: Morgan Kaufmann; 2008.
38. Sauro J, Lewis J. Quantifying the User Experience, 2nd edition: Morgan Kaufmann; 2016.
39. Albert B, Tullis T. Measuring the User Experience, 2nd edition: Morgan Kaufmann; 2013.
40. Kasabov N. Global, local and personalised modeling and pattern discovery in bioinformatics: an integrated approach. Pattern Recognit Lett 2007 Apr;28(6):673-685. [doi: 10.1016/j.patrec.2006.08.007]
41. Kasabov N, Hu Y. Integrated optimisation method for personalised modelling and case studies for medical decision support. Int J Funct Inform Pers Med 2010;3(3):236. [doi: 10.1504/IJFIPM.2010.039123]
42. Visweswaran S, Angus DC, Hsieh M, Weissfeld L, Yealy D, Cooper GF. Learning patient-specific predictive models from clinical data. J Biomed Inform 2010 Oct;43(5):669-685. [doi: 10.1016/j.jbi.2010.04.009] [Medline: 20450985]
43. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med 2015 Jan 6;162(1):55-63. [doi: 10.7326/M14-0697] [Medline: 25560714]
44. Green SB. How many subjects does it take to do a regression analysis. Multivariate Behav Res 1991 Jul 1;26(3):499-510. [doi: 10.1207/s15327906mbr2603_7] [Medline: 26776715]
45. Riley RD, Debray TPA, Collins GS, et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Stat Med 2021 Aug 30;40(19):4230-4251. [doi: 10.1002/sim.9025] [Medline: 34031906]
46. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001 Aug;54(8):774-781. [doi: 10.1016/s0895-4356(01)00341-9] [Medline: 11470385]
47. Self-Medication Assessment Tool (SMAT). Mount Allison University. URL: https://mta.ca/self-medication-assessment-tool-smat [accessed 2026-03-20]
48. Irvine-Meek J, Gould ON, Wheaton H, Todd LE. Acceptability and face validity of a geriatric self-medication assessment tool. Can J Hosp Pharm 2010 May;63(3):225-232. [doi: 10.4212/cjhp.v63i3.918] [Medline: 22478982]
49. Schmier JK, Halpern MT, Covert D. Validation of the Daily Living Tasks Dependent on Vision (DLTV) questionnaire in a U.S. population with age-related macular degeneration. Ophthalmic Epidemiol 2006 Apr;13(2):137-143. [doi: 10.1080/09286580600573049] [Medline: 16581618]
50. Grady PA, Gough LL. Self-management: a comprehensive approach to management of chronic conditions. Am J Public Health 2014 Aug;104(8):e25-e31. [doi: 10.2105/AJPH.2014.302041] [Medline: 24922170]
51. Raehl CL, Bond CA, Woods T, Patry RA, Sleeper RB. Individualized drug use assessment in the elderly. Pharmacotherapy 2002 Oct;22(10):1239-1248. [doi: 10.1592/phco.22.15.1239.33473] [Medline: 12389875]
52. Hearing tests. Patient.info. 2022. URL: https://patient.info/doctor/hearing-tests-pro [accessed 2026-03-20]
53. Risser J, Jacobson TA, Kripalani S. Development and psychometric evaluation of the Self-efficacy for Appropriate Medication Use Scale (SEAMS) in low-literacy patients with chronic disease. J Nurs Meas 2007;15(3):203-219. [doi: 10.1891/106137407783095757] [Medline: 18232619]

XSL•FO
RenderX

54. Yang C, Hui Z, Zeng D, et al. A community-based nurse-led medication self-management intervention in the improvement of medication adherence in older patients with multimorbidity: protocol for a randomised controlled trial. BMC Geriatr 2021 Mar 2;21(1):152. [doi: 10.1186/s12877-021-02097-x] [Medline: 33653300]
55. Creech CL. The "Medication Matters…to You!" Educational Intervention Session: A Pilot Study to Improve Medication Management in Community-Living Older Adults [PhD thesis]. : Texas Woman's University; 2016.
56. Baby B, Gill JK, Faisal S, et al. Medication adherence technologies: a classification taxonomy based on features. Mayo Clin Proc Digit Health 2025 Sep;3(3):100237. [doi: 10.1016/j.mcpdig.2025.100237] [Medline: 40881105]
57. Lee J, Maslove DM, Dubin JA. Personalized mortality prediction driven by electronic medical data and a patient similarity metric. PLoS ONE 2015;10(5):e0127428. [doi: 10.1371/journal.pone.0127428]
58. Woo KN. Usability of Medication Adherence Technologies Among Older Adults [Master's thesis]. : University of Alberta; 2012.
59. Assistive technology. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/assistive-technology [accessed 2026-03-20]
60. Scherer MJ. Living in the State of Stuck: How Assistive Technology Impacts the Lives of People with Disabilities, 4th edition: Brookline Books; 2005.
61. Applying human factors and usability engineering to medical devices: guidance for industry and Food and Drug Administration staff. US Food and Drug Administration. 2016. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices [accessed 2026-03-20]
62. International Electrotechnical Commission; Association for the Advancement of Medical Instrumentation; American National Standards Institute. Medical devices—part 1: application of usability engineering to medical devices. ANSI/AAMI/IEC 62366-1:2015/(r)2021+AMD1. : International Organization for Standardization; 2020 URL: https://www.iso.org/standard/63179.html [accessed 2026-04-01]
63. Roma MSG, de Vilhena Garcia E. Medical device usability: literature review, current status, and challenges. Res Biomed Eng 2020 Jun;36(2):163-170. [doi: 10.1007/s42600-019-00037-8]
64. Padmanabhan S. A user-focused approach to medical device development: identifying usability challenges and prioritizing solutions. Decos. 2024. URL: https://healthtech.decos.com/en/resources/blog/user-focused-approach-medical-device-development-identifying-usability-challenges-and [accessed 2026-04-01]
65. Mattern R. Consumer technology has limitations for assistive needs. Macular Degeneration Association. 2025. URL: https://macularhope.org/consumer-technology-has-limitations-for-assistive-needs/ [accessed 2026-03-20]
66. Ng K, Sun J, Hu J, Wang F. Personalized predictive modeling and risk factor identification using patient similarity. AMIA Jt Summits Transl Sci Proc 2015;2015(132–136):132-136. [Medline: 26306255]
67. Ahsan M, Ahmed MD, Azeem K. Role of predictive modeling and personalized modeling in the enhancement of athletic performance. Saudi J Sports Med 2023;23(1):7. [doi: 10.4103/sjsm.sjsm_7_23]
68. Baby B, McKinnon A, Patterson K, et al. Tools to measure barriers to medication management capacity in older adults: a scoping review. BMC Geriatr 2024 Mar 27;24(1):285. [doi: 10.1186/s12877-024-04893-7] [Medline: 38532328]

Abbreviations

CSM: cosine similarity metric
CV: cross-validation
DLTV: Daily Living Tasks Dependent on Vision
GLM: generalized linear model
MAT: medication adherence technology
MMC: medication management capacity
MPED: Martin and Park Environmental Demands
MSE: mean square error
PPM: personalized predictive modeling
REDCap: Research Electronic Data Capture
SEAMS: Self-Efficacy for Appropriate Medication Use Scale
SMAT: Self-Medication Assessment Tool
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
UX: user experience

XSL•FO
RenderX

Edited by AF Abdulai; submitted 22.Sep.2025; peer-reviewed by A Hu, M Gackowski; revised version received 31.Jan.2026; accepted 28.Feb.2026; published 08.Apr.2026.

Please cite as:
Baby B, Elba G, Kim M, Park S, Hudani I, Sharma R, Patel H, Bajaj S, Patterson PK, McKinnon A, Guilcher SJT, Chang F, Lee L, Burns C, Griffin R, Goh J, Patel T
Personalized Predictive Model to Predict Subtask Success of Medication Adherence Technologies for Older Adults With Diverse Capabilities: Development and Internal Validation Study
JMIR Aging 2026;9:e84616
URL: https://aging.jmir.org/2026/1/e84616
doi:10.2196/84616

© Bincy Baby, Ghada Elba, Minzee Kim, SooMin Park, Imra Hudani, Rishabh Sharma, Halak Patel, Sidharth Bajaj, P K Patterson, Annette McKinnon, Sara J T Guilcher, Feng Chang, Linda Lee, Catherine Burns, Ryan Griffin, Joslin Goh, Tejal Patel. Originally published in JMIR Aging (https://aging.jmir.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Information Access in Aging Populations and Those With Dementia and Mild Cognitive Impairment in the United Kingdom: Survey and Focus Group Study

Claire Rogers[1], BSc, MSc; William J McGeown[2], PhD; Yashar Moshfeghi[1], PhD

[1]NeuraSearch Laboratory, Department of Computer and Information Sciences, University of Strathclyde, 26 Richmond Street, Glasgow, United Kingdom
[2]Department of Psychological Sciences and Health, University of Strathclyde, Glasgow, United Kingdom

Corresponding Author:
Yashar Moshfeghi, PhD
NeuraSearch Laboratory, Department of Computer and Information Sciences, University of Strathclyde, 26 Richmond Street, Glasgow, United Kingdom

Abstract

Background: With the growing aging population, technology that supports independent living is increasingly important. Web search systems are well established, whereas generative artificial intelligence (Gen-AI; eg, ChatGPT) represents a newer, adaptive tool that could offer personalized information access. However, little is known about how older adults, particularly those with mild cognitive impairment (MCI) or mild dementia, perceive and engage with these systems.

Objective: This study explored the use of and perspectives on web search and Gen-AI in older adults with and without cognitive impairment (including MCI and early-stage dementia).

Methods: A UK-wide mixed methods study was conducted with older adults, including those with MCI or mild dementia. An online survey captured technology use, Likert-scale ratings of web search and Gen-AI, and reasons for nonuse. Follow-up focus groups provided in-depth qualitative perspectives. Quantitative data were analyzed using descriptive and comparative statistics, while qualitative data were thematically analyzed.

Results: Survey findings showed higher use of web search (275/280, 98.2%) compared to Gen-AI (40/286, 14%) within these groups. Web search was rated positively across participants, although challenges were raised regarding the phrasing of queries and commercialization. Gen-AI use was less common, but more than half of nonusers expressed willingness to adopt it in the future. Combined with focus group responses, themes exploring keyword searching, mistrust, lack of knowledge, and willingness to learn were established. Participants also suggested potential applications of Gen-AI, such as supporting independent living through monitoring and simplifying complex searches.

Conclusions: Web search remains the primary method, and participants highlighted both advantages and frustrations with current systems. Gen-AI was underused but seen as promising, with its adoption mainly limited by mistrust and knowledge gaps. Our findings indicate that structured training, early introduction, and user-centered design could encourage adoption, enhance accessibility, and support independent living among older adults with and without MCI.

(JMIR Aging 2026;9:e85626) doi:10.2196/85626

KEYWORDS

artificial intelligence; dementia; mild cognitive impairment; older adult; generative AI; ChatGPT; search; information retrieval; aging; accessibility

Introduction

Given the rise in the number of people living with dementia worldwide [1], there is an increased demand for technologies that support independent living. In early-stage dementia and mild cognitive impairment (MCI), one potential technology to preserve independence is information retrieval (IR) systems that facilitate access to information. This includes 2 primary areas: web search (eg, Google) and generative artificial intelligence (Gen-AI), such as ChatGPT. In this study, "Gen-AI" refers to text-based conversational artificial intelligence (AI) systems powered by large language models (LLMs), such as ChatGPT. ChatGPT was used illustratively in the survey materials and focus groups, but perceptions of Gen-AI tools were examined more broadly. Much of the existing research on information access via web search has focused on users without impairments [2,3] or younger users [4-6], primarily examining health information seeking rather than broader applications [7-10]. Findings indicate that although health information seeking among older adults is becoming more common, barriers, including digital literacy, stigma, limited support, and information overload, remain [11,12]. Additionally, Gen-AI

XSL•FO
RenderX

systems could serve as a promising tool for information access, enabling independent searches and conversations. Gen-AI's natural language capabilities [13,14] suggest a more adaptable alternative to traditional web search [15], especially beneficial for those with limited digital skills or cognitive impairments.

Recent research suggests that Gen-AI could offer a more personalized and person-centered approach within both new and existing technologies [16], including information search, by simplifying complex information into more accessible formats [17,18]. This indicates that integrating such tools into information access mechanisms could be particularly advantageous in the context of aging [19]. Although many recognize the substantial ethical challenges posed by Gen-AI concerning data management and information sharing, especially given the vulnerability of this population [20,21], it is acknowledged that the potential of Gen-AI, particularly as a means of conversational search, remains significant [20]. Emerging studies have begun to explore older adults' perspectives on LLM-based systems, especially regarding health information seeking and companionship [22], as well as wider AI-driven health technologies [19,23]. These studies suggest that while older adults see potential in these systems, they continue to face barriers such as limited digital literacy, ethical concerns, and persistent ageism in technology design. Yet, despite this growing body of research, little is known about how these tools are experienced by older adults with MCI or dementia, who may have different needs compared to their cognitively healthy peers. Addressing this gap is vital to ensure that future Gen-AI applications support, rather than exclude, those living with cognitive impairments. AI has also been identified as a potential means of enhancing quality of life through informational and social support [24], although studies have noted that responses from Gen-AI can lack the narrative richness and personal experiences often found in peer-to-peer online forums [25]. This highlights the need to balance efficiency with personalization and authenticity valued by older users. Research has begun to investigate their use in this cohort for IR [20], reducing loneliness [26-28], improving mental health [29], and establishing best training practices [30], indicating that the potential for application in the aging population has already been recognized.

Older adults and those with cognitive impairment are often stereotyped as unable or unwilling to use technology [31-33]. This may create a self-fulfilling prophecy [34-37] in which technology is perceived as irrelevant, leading to reluctance, avoidance [38,39], or an underestimation of their abilities. Unfortunately, this stereotype is reflected in some technological developments, which do not attempt to accommodate older populations or are overly simplistic, focusing on skills they have lost rather than retained [37,40]. Additionally, greater emphasis is often placed on young people and their use of technology, which can result in older adults being excluded from its benefits [41]. This "digital ageism" has become an increasingly recognized issue, highlighting how technologies can become misaligned with older adults as age-related changes occur, affecting digital literacy [42] and widening the divide as these groups remain underrepresented in technology development. Beyond equity concerns, reduced digital access has clinical implications, as it may limit autonomy, delay decision-making, or increase dependence on caregivers. More recently, researchers have stressed the importance of addressing this digital divide in relation to Gen-AI specifically, calling for greater attention to accessibility, user-centered design, and inclusivity [43]. Such work demonstrates that while Gen-AI could be transformative for older cohorts, deliberate efforts are necessary to realize these potential benefits. In recent years, the importance of capturing the views of underrepresented groups on AI systems has been increasingly recognized, with older adults identified as a key population whose perspectives should inform the development of these technologies [22]. Therefore, exploring perspectives on information access is vital to understanding how older adults with and without cognitive impairment (including MCI and early-stage dementia) use these technologies and where improvements are needed.

This paper examines usage patterns and perspectives of traditional IR and Gen-AI systems in healthy older adults, those with MCI, and those with early-stage dementia, to characterize current engagement and identify directions for future development. To do so, we ask: How do older adults with and without cognitive impairment (including MCI and early-stage dementia) currently use and perceive traditional web search systems and Gen-AI tools, and what factors shape their engagement with these technologies? To our knowledge, this is the first study to directly compare perspectives on traditional web search and Gen-AI systems across healthy older adults, individuals with MCI, and those with early-stage dementia. In doing so, it provides insights to inform inclusive, clinically relevant design of future information systems.

Methods

Overview

This mixed methods study explored technology use and opinions of older adults with and without MCI or early-stage dementia. The study was conducted in 2 parts: an online survey and follow-up focus groups. The design accounted for heterogeneity in disease and at-risk populations by including participants across the cognitive spectrum, including early-stage dementia, MCI, and cognitively healthy older adults.

Ethical Considerations

The study was given ethical approval by the University of Strathclyde Ethics Committee (UEC24/28). All participants provided informed consent prior to participation. Participants were informed of the nature and purpose of the study and their right to withdraw at any time without penalty. All data were collected and stored in accordance with relevant data protection regulations, and no personally identifiable information was included in the analysis or reporting of results. Participants did not receive compensation for their participation.

Part 1: Online Survey

Participants

The survey gathered a total of 325 responses. As some participants did not go beyond the demographic questions, 39 (12%) responses were excluded before analysis, resulting in

XSL•FO
RenderX

286 (88%) valid responses (162 women, 123 men, and 1 who preferred not to say), with a mean age of 72.87 (SD 5.67) years. Partial responses, where participants completed part of the main survey, were included. Participants were recruited via word of mouth, social networking sites (eg, X and LinkedIn), flyers, and the Join Dementia Research platform, with 272 participants identifying as currently residing in England and 14 in Scotland. Education levels are summarized in Table 1. Participants reported diverse professional and vocational backgrounds, including health care, education, government, engineering, business, and skilled trades. This diversity highlights the varied socioeconomic and professional experiences represented in the sample. They were divided into 3 groups: healthy older adults aged older than 65 years without a neurological diagnosis (n=194, 67.8%, mean 72.38, SD 5.55 y; 119 women and 75 men); adults aged older than 65 years with early-stage or mild dementia or MCI (n=87, 30.4%, mean 74.1, SD 5.62 y; 42 women, 44 men, and 1 who preferred not to say); and caregivers of adults aged older than 65 years with early-stage or mild dementia or MCI, who responded on behalf of the person they cared for, providing their demographics (not their own; n=5, 1.7%, mean 70.4, SD 8.47 y; 1 woman and 4 men). Participants with cognitive impairment included 31 (33.7%) with Alzheimer disease, 44 (47.8%) with MCI, 8 (8.7%) with vascular dementia, 3 (3.3%) with frontotemporal dementia, 2 (2.2%) with Lewy body dementia, and 4 (4.3%) with other or unspecified conditions. Due to the small size of the caregiver group, these responses were combined with those of the dementia and MCI group. Sensitivity analyses excluding caregiver responses showed no meaningful changes in results, justifying their inclusion in the combined cognitive impairment group.

Table . Education level of survey participants[a].

Education level	Response, n (%)
PhD	16 (5.6)
Bachelor's degree	73 (25.5)
Master's degree	31 (10.8)
Some college or further education	88 (30.8)
Secondary school	40 (14)
Other	38 (13.3)

[a]"Other" includes professional, vocational, or partial higher education qualifications not captured by standard degree categories, such as teaching certificates (Postgraduate Certificate in Education), nursing or medical qualifications, engineering or law certifications, postgraduate diplomas, and other professional credentials.

Design

This study was designed as a mixed methods cross-sectional survey using independent samples. The online survey was created and conducted on Qualtrics (Qualtrics International Inc), with questions focusing on respondents' views of online search systems, such as Google, and Gen-AI systems, particularly conversational Gen-AI, such as ChatGPT.

The survey instrument was developed collaboratively by the first author and 2 supervisors to explore awareness, experience, and perceptions of Gen-AI and online search systems in older adults with and without cognitive impairment (including MCI and early-stage dementia). As no existing validated instruments addressed perceptions of conversational Gen-AI in older adults with cognitive impairment, items were created de novo, and established frameworks such as the Technology Acceptance Model (TAM) were not used because they did not capture the specific nuances of these systems in this population. The survey included a combination of Likert-scale and open-ended questions covering prior use, interaction and question phrasing, perceived response relevance, emotional experience, perceived usefulness, and perceived ability for independent use.

For Gen-AI questions, participants were first asked whether they used this technology. Those who said yes were asked about the applications they use, how often, for how long, and through which devices. They responded to questions on a 5-point Likert scale regarding the ease or difficulty of using the applications and the reasons behind their experiences, their usefulness, ease of selecting and writing queries, the relevance of the output, and their experience with the system in terms of boredom, enjoyment, independence, complexity, engagement, and feeling overwhelmed. Those who said no to using Gen-AI tools were asked about their reasons for not using them, what might encourage them to use these tools in the future, and whether they would consider using them in the future. Most Likert-scale questions were rated between "strongly disagree" and "strongly agree." These responses were encoded as integers from 1 through 5 and averaged across groups.

Similarly, participants were asked if they used online search systems. Those who answered yes were also asked about which applications they used (specific to general web search engines), how often, for how long, and through which devices. They were given Likert-scale questions about how easy or difficult the applications were to use and why and questions regarding the overall search process. They were asked about query formulation in terms of phrasing and rephrasing questions and relevance assessment regarding whether the search results were relevant, whether enough relevant information was provided, whether they could decide what information was relevant, and whether results were always relevant to their question. They were also questioned about helpful features, any aspects needing improvement, and their experiences with the system, including memory of their search goals, feelings of being lost, engagement, enjoyment, boredom, independence, complexity,

XSL•FO
RenderX

and feeling overwhelmed. Those who said no to using online search systems were asked about the reasons for not using them, what might encourage them to use these systems, and whether they would consider using them in the future. All participants were asked at the end of the survey what they would like to see in a system they could design themselves and given an opportunity to add any further comments they wished to voice.

Pilot

Five healthy older adults with no cognitive impairments were recruited through the University of Strathclyde Older Adult Research Panel. The pilot sessions were conducted both in person and online. Participants completed the survey with the researcher, providing feedback on the questions, structure, and length of the survey.

Following pilot feedback, minor refinements were made: illustrative images introducing Gen-AI and online search systems were added; it was clarified that "generative AI" referred to conversational, text-based chatbots rather than voice assistants; clarification that questions referred to the search system interface rather than website credibility was included; confusing or redundant response options were removed; and it was more clearly indicated where multiple responses could be selected. Participants reported the survey length and structure to be appropriate, and no major restructuring was required. As the study was exploratory, no formal psychometric validation or reliability testing was conducted.

Procedure

Participants first selected the group they identified with, then filled out the relevant participant information sheet and consent form. If they were interested in joining the focus groups, they had the option to provide their email address for future contact. They received an ID and proceeded to the main questions in the survey, completing the *Generative AI* section followed by the *Online Search System* section. A debrief was then displayed. After the survey recruitment, invitations were sent to those who had expressed willingness to participate in the focus groups. Survey completion times were highly skewed due to a small number of very long durations. The median completion time was 825 (IQR 597.5 - 1223) seconds.

Quantitative Analysis

A series of 2-tailed independent-samples *t* tests were carried out to examine each Likert-scale response between the MCI or dementia and healthy older adult groups. Although Likert-scale data are technically ordinal, independent-samples *t* tests can be considered robust to violations of the interval assumption, particularly for detecting group differences [44]. To account for the small and unequal sample sizes in the Gen-AI comparisons, we conducted supplementary nonparametric analyses using Mann-Whitney *U* tests, which produced results consistent with the *t* tests. Therefore, the reported *t* test results provide a valid representation of group differences.

Part 2: Follow-Up Focus Groups

Focus groups were organized to explore survey results in greater depth.

Participants

Ten survey participants took part in the follow-up focus groups (5 men, 5 women; mean age 73.9, SD 5.67 y). Two (20%) of 10 participants had indicated in the survey that they had used Gen-AI before, and all participants had said yes to having used online search systems. Six (60%) participants identified as healthy older adults, while 4 (40%) had a diagnosis of Alzheimer disease, dysexecutive syndrome due to neurodegeneration, or MCI.

Three focus groups were conducted in total. Group composition was determined based on cognitive status and prior Gen-AI experience: (1) healthy older adults with no prior Gen-AI experience (n=4, 40%), (2) adults with MCI or dementia with no prior Gen-AI experience (n=2, 20%), and (3) a mixed group including both healthy older adults and adults with MCI or dementia, with varying prior Gen-AI experience (n=4, 40%). While group 2 had only 2 participants due to recruitment constraints, their perspectives were included to ensure representation of individuals with cognitive impairment. The characteristics of each group are summarized in Table 2.

Table . Composition of focus groups.

Focus group	Participants, n (%)	Gender (female/male), n	Age, mean (SD)	Diagnosis[a]	Prior Gen-AI[b] use[c]
1	4 (40)	3/1	72.5 (6.56)	No impairments	No
2	2 (20)	0/2	71.5 (6.36)	MCI[d] or dementia	No
3	4 (40)	2/2	76.5 (4.93)	2 MCI or dementia and 2 no impairments	2 yes and 2 no

[a]Diagnosis refers to the self-reported diagnosis provided by participants in the survey demographic questions.

[b]Gen-AI: generative artificial intelligence.

[c]Prior Gen-AI use indicates self-reported experience with Gen-AI tools before completing the survey.

[d]MCI: mild cognitive impairment.

Design

Question topics were informed by initial survey responses, establishing exploration of early patterns, such as trust, which was not questioned in the survey, yet was raised in responses. A semistructured interview guide was used to ensure consistent topic coverage across groups while allowing flexibility for probing and follow-up questions. The full interview guide is provided in Multimedia Appendix 1. To minimize moderator

XSL•FO
RenderX

bias, neutral questioning was used, participation from all group members was encouraged, and both positive and negative perspectives were explored. The presence of a second researcher allowed oversight of adherence to the interview guide and provided a reflexive discussion following each session. All sessions were audio recorded and transcribed verbatim, and data collection continued until thematic saturation was reached, ensuring rigor and reliability of qualitative findings. The focus groups were moderated by the lead researcher, with a senior supervising researcher present during each session.

Procedure

Selected participants were contacted by email for the focus group follow-up. Participants read and completed the participant information sheet and consent form before attending the meeting, and focus groups were conducted online via Zoom (Zoom Communications). Each session lasted approximately 80 minutes, including a short mid-session break. Introductions were made, followed by a short demonstration of Gen-AI, particularly ChatGPT (OpenAI), illustrating the basic functions and capabilities of the system. This demonstration lasted approximately 5 to 10 minutes and followed a scripted sequence of predefined prompts to ensure consistency across groups (Multimedia Appendix 1). The main discussion then centered around Gen-AI, covering overall opinions, experience asking questions, its responses to questions, relevance, trust, and training. The same questions were also posed for online search systems. A brief debrief and a question-and-answer session were provided before closing.

Qualitative Analysis

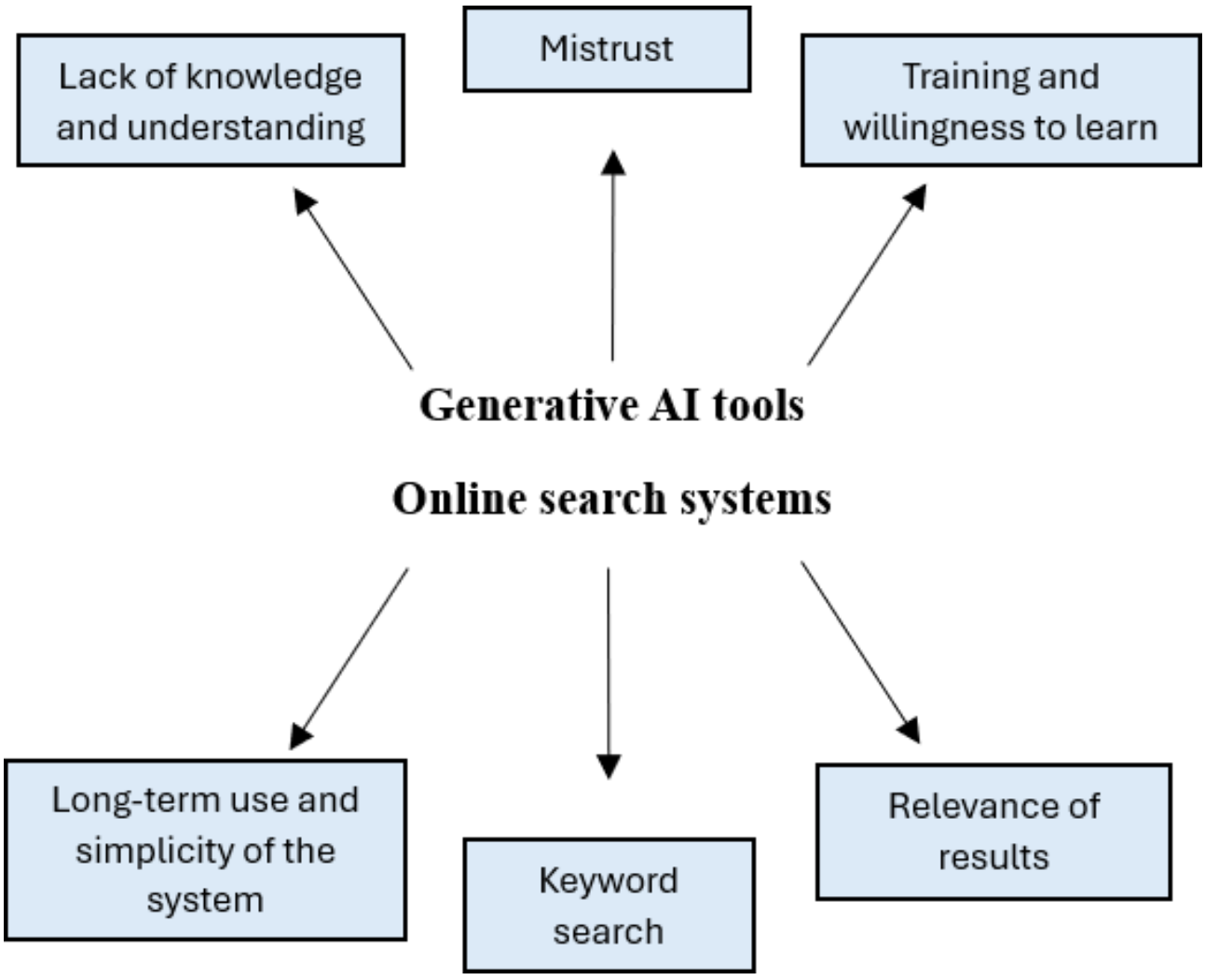
To analyze the qualitative data from both the open-ended survey questions and focus group responses, a reflexive thematic analysis was performed [45]. Themes were iteratively developed and refined across both data sources, allowing the qualitative findings to contextualize and deepen the understanding of the quantitative survey results.

Results

Overview

Thematic analysis of Gen-AI content from the survey open-ended questions and focus groups identified 4 main themes: lack of knowledge about current technologies, mistrust, self-efficacy, and training and willingness to learn. For online search systems, the primary themes were long-term use and simplicity of the system, keyword searching, and relevance of results. Follow-up focus groups explored these themes in more detail, and key findings are presented alongside thematic results (Figure 1).

Figure 1. Final themes developed from qualitative data analysis. AI: artificial intelligence.



XSL•FO
RenderX

Generative AI

Quantitative Data

When asked whether participants currently use a Gen-AI tool, 40 (14%) of 286 participants responded "yes," comprising 27 (13.9%) of 194 older adults and 13 (14.1%) of 92 of those with MCI or dementia, with the most common use being to find information (Table 3). Many participants had heard of this technology, but the reasons given for not having used it varied (Table 4). Despite most participants saying "no" to using Gen-AI, 169 (68.7%) of 246 participants indicated they would consider using Gen-AI tools in the future, highlighting the potential for future adoption. Among those who used these tools, usage was mostly less than monthly, followed by weekly, lasting from a few minutes up to 10 minutes, with ChatGPT being the most frequently used. Thirty-three (94.2%) of 35 Gen-AI users felt that Gen-AI tools could be helpful in their lives. When asked to rate the ease of searching for information using a Gen-AI tool, where 1 was extremely difficult and 5 was extremely easy, the mean rating was 4.15 (SD 0.96).

Table . Uses of generative artificial intelligence (Gen-AI) and web search[a].

Uses	Gen-AI, n (%)	Web search, n (%)
To explore	9 (25.7)	139 (53.7)
To find information	22 (62.9)	257 (99.2)
To learn	20 (57.1)	176 (68)
To socialize	1 (2.9)	20 (7.7)
For entertainment	4 (11.4)	81 (31.3)

[a]Reported use of Gen-AI and web search across participants. N indicates the number of participants who selected each use. Participants could select multiple answers; therefore, totals exceed the number of participants. "Other" and "prefer not to say" were also available options.

Table . Reasons for nonuse[a].

Response option	All groups, n (%)	MCI[b] or dementia, n (%)	Healthy older adults, n (%)
I have heard of it but have never considered using it	113 (45.9)	33 (41.8)	80 (47.9)
I have heard of it but have never had the chance to use it	38 (15.4)	13 (16.5)	25 (15)
I used it and decided to stop using it or didn't like it	8 (3.3)	1 (1.3)	7 (4.2)
I have never heard of it	87 (35.4)	32 (40.5)	55 (32.9)

[a]Participant responses for their reasons behind the nonuse of Gen-AI. Percentages are calculated within participants who reported not using Gen-AI (n=246 total; n=79 MCI or dementia; n=167 healthy older adults).

[b]MCI: mild cognitive impairment.

Independent-samples *t* tests revealed no significant differences between the older adult and MCI or dementia groups, except for independent use of the system, where older adults responded in greater agreement to being able to use Gen-AI independently compared to those with MCI or dementia (Table 5). Nonetheless, there was a noticeable trend of older adults rating items, such as independent use and the ease of use of Gen-AI applications, as more agreeable than the MCI or dementia group. Further investigation through qualitative responses provided additional insights.

XSL•FO
RenderX

Table . Generative artificial intelligence (AI) Likert responses between groups[a].

Question	MCI[b] or dementia, mean (SD)	Healthy older adults, mean (SD)	Significant differences (*P* value)	Effect size (Cohen *d*)
"Generative AI helps me answer my questions". To what extent do you agree with this statement?	3.77 (1.01)	4.18 (0.73)	.17	−0.49
I have to reword my question to get the answer I'm looking for	3.58 (0.9)	3.36 (1.22)	.59	0.2
I find Generative AI tools simple to use	3.69 (1.11)	4.23 (0.87)	.12	−0.56
I can use Generative AI tools independently	4 (1.15)	4.68 (0.65)	.031	−0.79
I feel overwhelmed when using Generative AI tools	1.85 (1.07)	1.5 (0.86)	.30	0.37

[a]Responses from participants who reported that they use Gen-AI applications. Full independent-samples *t* test statistics (*t* values, degrees of freedom, and exact *P* values) for all items are provided in Multimedia Appendix 2. Full Likert-scale response distributions are provided in Multimedia Appendix 3. 1=strongly disagree and 5=strongly agree.

[b]MCI: mild cognitive impairment.

Qualitative Data

Lack of Knowledge and Understanding

When survey participants were asked why they did not use a Gen-AI tool, responses reflected a general lack of knowledge, such as "I don't really understand them," not knowing enough about their potential, such as "don't really know enough about it, its benefits or how to access and use it," and not understanding how to operate it, such as "don't know how to use it." This also extended to the belief that it is solely for education; one survey participant said:

> *...I've read it is used for such things as school/University course work... now I've seen a full explanation... I will investigate further.*

These qualitative responses align with quantitative trends showing that awareness alone does not guarantee confident adoption. For example, the mean ratings for knowing how to phrase their questions were 3.23 (SD 1.3; MCI or dementia) and 3.86 (SD 1.04; healthy older adults), and for understanding when to use Gen-AI, the mean ratings were 3.67 (SD 1.23) and 4.14 (SD 0.89), respectively (Multimedia Appendix 2). This suggests that limited familiarity may restrict effective use, particularly for the MCI or dementia group. Focus groups provided similar insights, with participant 6 (male, aged 69 y, no impairment) stating, "...I didn't know what it was and how to use it," while also expressing interest in using it after hearing about its capabilities.

A minor theme identified in the survey data was the lack of perceived need to use this technology, as reflected by both the MCI or dementia group, "not felt the need," and in the older adult group, "I have no need to generate tracks of commentary or explanation... I feel confident in my ability to read, assess, and formulate written responses." Quantitative data support this: although 68.7% of participants indicated they would consider using Gen-AI tools in the future, current use was only 14%. Older adult survey responses offered more detail, with comments such as "have no need to and don't know enough about it' and "I've not felt a need to use one and I'm not sure how to get started." These insights suggest that perceived lack of need may reflect limited understanding of Gen-AI applications.

Mistrust

A further theme was mistrust. Survey participants who had never used Gen-AI expressed security concerns ("don't trust it. Think it can be used/hacked for nefarious and cruel purposes") and concerns about accuracy ("probably foolishly, I don't feel I can fully trust AI-generated tools or the information they provide"). This aligns with quantitative findings showing lower confidence in independent Gen-AI use (MCI or dementia: mean 4.00, SD 1.15 vs healthy older adults: mean 4.68, SD 0.65; *P*=.03), suggesting that it may limit use. This may relate to limited knowledge about data management.

Survey participants also raised concerns about regulation: "I believe they are dangerous...There is insufficient and inadequate regulation." This aligns with responses from those who had used Gen-AI but stopped or disliked it; when asked what caused this, survey replies included accuracy concerns, ("suspicious of the accuracy of results"), privacy and security ("I need to be absolutely certain that it's secure") and regulation ("better monitoring of these tools by independent third parties"). Quantitative responses showed that although participants generally agreed Gen-AI provides relevant answers (MCI or dementia: mean 3.69, SD 1.32 and healthy older adults: mean 4.09, SD 0.81), lower confidence in phrasing questions (mean 3.23, SD 1.3 vs mean 3.86, SD 1.04) and understanding when to use tools (mean 3.67, SD 1.12, vs mean 4.14, SD 0.89) reflects that mistrust may affect perceived usefulness (Multimedia Appendix 2).

Focus groups discussed trust in more depth. Participants said, "…at the moment I don't trust artificial intelligence," and "…I

XSL•FO
RenderX

actually trust very little of anything." Focus group participants also raised privacy concerns:

> *I think we have enough observation in our lives… from all the CCTVS… a bit of privacy at your own home is quite nice.*
>
> *I'm concerned as well about these algorithms… it gets to know what you like… that worries me too.*

Regulation was also highlighted: "what is being done with social media and how it's abused…we allowed that because nothing was regulated," emphasizing transparency. Participant 3 (female, aged 78 y, no impairment) expressed feeling a lack of control:

> *…I don't know who's feeding it the information…Where's that coming from? Is it coming from one source?*

Survey responses mirrored this cautious use: "if I needed to use AI I would, but I prefer not to increase use of AI until I am sure it is controllable." This mistrust in AI is reflected through sci-fi references by participant 9 (male, aged 74 y, Alzheimer disease), "...these extraordinary machines are not human and all our concerns about sci-fi scenarios in which these machines take over the world," which again mirrors survey responses, "I think they could be dangerous...watch the Terminator films again. If the machines ever become self aware humanity will be in real trouble."

Focus groups also highlighted that trusting Gen-AI depends on the user verifying results. Participant 7 (male, aged 82 y, MCI) said:

> *if I've taken part in history and I quiz the machine... and it responds with a different account... the machine immediately turns around to me and apologizes for providing incorrect facts. So that confuses me a little bit.*

They stressed that their own knowledge is essential:

> *if I hadn't been a witness to the event or taken part in it, I would have believed everything the machine told me...*

Healthy older adults echoed this:

> *…results I get back can be hit and miss and you really have to know what it is you're asking.*
>
> *[Participant 8, female, aged 71 y, no impairment]*
>
> *AI seems to me to be something that just gives you what it thinks you want rather than what you think you want.*
>
> *[Participant 6, male, aged 69 y, no impairment]*

Participant 1 (male, aged 76 y, Alzheimer disease) reinforced this:

> *if you don't have the cognitive ability to be able to discern between fact and fiction, then you're quite likely to be misled, misinformed, or actually harmed.*

These observations help explain why participants reported relevant outputs (MCI or dementia: mean 4.17, SD 0.83), yet lower confidence in independent use, suggesting that mistrust and cognitive concerns influence engagement. Therefore, reliance on the user to identify correct and unbiased information is a significant factor. This was echoed by a participant with MCI or dementia in the survey data, who stated:

> *I'm not a fan of AI. I like to be able to tell the difference between what's real and what's fake... I also don't want to be confused as the disease progresses... I would like all AI generated images and text to have to carry an "AI" generated symbol.*

Thus, mistrust and need for oversight appear linked to lower confidence in independent use shown in the quantitative data, highlighting the importance of distinguishing genuine from AI-generated content.

Perceived Capability and Self-Efficacy

A distinct theme was participants' perceived capability when interacting with Gen-AI, reflecting situational self-efficacy. Some focus group participants reported feeling the need to treat AI as more capable than themselves. For example, participant 7 (male, aged 82 y, MCI) said:

> *...I couldn't treat Gemini or GPI as a machine... For me to relate to them and to use them, I had to treat them as human beings who were more intelligent and more verbal than I was.*

Conversely, focus group participant 8 (female, aged 71 y, no impairment), who uses Gen-AI, offered a contrasting viewpoint, "...I believe I treat it as a machine, as a supercomputer." These perspectives contextualize survey responses reporting discomfort with interactions, "the robotic nature of the exchanges and limitations to interact are uncomfortable to deal with," and may help explain variance in engagement and trust metrics across participants. Although perceived capability was not directly measured, these qualitative insights suggest that participants' situational self-efficacy in using Gen-AI varies depending on how they perceive the AI's abilities relative to their own.

Training and Willingness to Learn

When asked about what would encourage participants to use Gen-AI in the future, survey responses reflected training needs and a general willingness to learn. The MCI or dementia survey group often requested "more knowledge, experience, and training" and "an explanation and teaching of how to do this." Similar responses were given by the older adult survey group, expressing a desire for information on "how it would benefit me followed by training on how to use it," including "a demonstration and practice in using them" and access to "free training sessions possibly at a library." Both survey groups also mentioned the importance of support from others to encourage use, such as "having someone around with the time to explain it and then help me to use it" in the MCI or dementia group, and "recommendation by a friend or family member..." in the healthy older adult group. Quantitative data support this: 68.7% (169/246) of survey respondents indicated they would consider using Gen-AI in the future. Additionally, Likert-scale responses showed moderate agreement that Gen-AI helps answer questions (MCI or dementia: mean 3.77, SD 1.01; and healthy older adults: mean 4.18, SD 0.73) and provides relevant answers (MCI or dementia: mean 4.17, SD 0.83; and healthy older adults: mean

XSL•FO
RenderX

4.18, SD 0.73). This suggests that addressing knowledge gaps and providing transparency could facilitate adoption.

Survey participants also outlined their expectations for training, including desires for "a clear idea of what it can do and be used for," "more information and guidance in how to use and its benefits," and practical examples. Comments also reflected concerns about trust and system performance. The MCI or dementia group mentioned, "only if I could be absolutely assured that what was produced was true and could be collated with other trusted sources of information," while the older adult group expressed concerns about accuracy, stating, "being reassured that they don't make up the answer if they can't find it." These concerns parallel survey Likert items on confidence in independent use, where lower ratings among participants with MCI or dementia (mean 4, SD 1.15 vs 4.68, SD 0.65 in healthy older adults) may reflect caution about trusting outputs and the need for training to build confidence while addressing reliability and accuracy.

Despite encouraging responses, it is important to acknowledge that some reactions showed reluctance to engage. Some participants replied "nothing" when asked what would encourage them to use these systems, while others, such as from the MCI or dementia group, expressed, "I personally don't feel I need anything like this, at the moment," and "Nothing. I do not trust the security of these sites." Older adults also shared perceptions that there is too much technology, with comments such as "I don't want to get involved with anything else," and believed that web search was sufficient, stating "I don't think I need to use such a tool, as I can find most information I need on Google." Willingness to use Gen-AI appears shaped by personal beliefs and experiences, with one noting, "I believe this tool is used to help create content/write something...I have never felt the need to use it." Some responses from the MCI or dementia group cited complexity, "it's too complicated," "I'm not very good with tech stuff," "coping with technology," and "I am a technophobe." These hesitations correspond with quantitative patterns, such as lower confidence in phrasing questions (MCI or dementia: mean 3.23, SD 1.3 vs healthy older adults: mean 3.86, SD 1.04) and understanding when to use Gen-AI (3.67, SD 1.23 vs 4.14, SD 0.89). This indicates that not all barriers can be addressed through training alone.

Both groups also noted ease of use and user-friendliness: "usefulness outweighing frustration." When asked about ease of use, those who used Gen-AI responded positively. The MCI or dementia group said, "...it provides a response that I understand," "it is very responsive to questions posed to it. I often put in further questions...," and "I am just starting with AI as I feel it could be of great help as my condition progresses." Older adults responded similarly, "the info is quite detailed but is easy to query in more detail by using more targeted questions," "no harder than searching in Google," and "a bit better and precise than an internet search." However, some negative points were raised, such as, "with some questions the AI has to be tricked into answering, or will not answer," and concerns about complexity: "I have limited tech" experience and, as I have no family or support network, I can be slow to learn & retain new methods." Overall, these experiences align with Likert-scale ease of use scores (mean 4.15 overall).

The focus group participants expressed varied interest in training. Participant 2 (male, aged 67 y, dysexecutive syndrome) said, "I would go and get that training because I'm curious...the problem is being able to remember... what you've learned," highlighting that training might be most effective early in diagnosis or beforehand. Participant 1 (male, aged 76 y, Alzheimer disease) expressed a preference for being taught by similar individuals, stating:

> *...training would be most effective if- if it was done by people such as myself who were in a MCI state and would have an understanding of what it was that they were talking about…*

Others, who were more skeptical about AI, did not show interest in training, "I just think it's money wasted on me at 80... better given the training to the youngsters so they're prepared for the future rather than me," or simply conveying a lack of interest in attending. Participant 10 (female, aged 79 y, no impairment) expressed a preference to explore it in their own time, stating:

> *I don't think in my case I would need training. I think I would look at it and- and explore it on my own.*

Additionally, it was discussed that targeted training for groups could be most beneficial:

> *I think that training on Gen-AI as a whole for members of the general public would probably possibly not be that useful...*

Despite varied interest, some participants found the focus group demonstration intriguing, saying:

> *when you were givin' the demonstration of ChatGPT, I wanted to start using it myself to, to find some information that I've not been able to get easily from Google.*

Overall, responses highlight that individual experiences, cognitive abilities, and confidence shape willingness to learn, but there is potential for training to help those interested.

Online Search Systems

Quantitative Data

Participants were asked if they currently use an online search system. Of the 286 survey participants, 280 (97.9%) responded to this question, reflecting partial survey completion. Among respondents, 275 (98.2%) of 280 participants responded "yes," with 186 (98.9%) of 188 older adults and 89 (96.7%) of 92 participants in the MCI or dementia group reporting use. Among these users, Google was the most used, and most participants reported using search systems daily and for less than 5 minutes at a time. Participants used these tools primarily to find information and learn (Table 3). When asked how easy they found searching for information using an online search system, where 1 was extremely difficult and 5 was extremely easy, the average response was 4.38 (SD 0.8), indicating generally positive perceptions.

Independent-samples *t* test results reveal significant differences between groups on several survey questions, with the healthy older adult group consistently rating items as more agreeable (Table 6) than the MCI or dementia group. For example, the

XSL•FO
RenderX

MCI or dementia group was more likely to report needing to reword queries, feeling lost, or feeling overwhelmed, whereas healthy older adults more often reported being able to remember what they were searching for, deciding whether information was relevant, and using search systems independently. Survey open-ended questions and focus groups provided additional insight.

Table . Web search Likert responses between groups[a].

Question	MCI[b] or dementia, mean (SD)	Healthy older adults, mean (SD)	Significant differences (*P* value)	Effect size (Cohen *d*)
I have to reword my question to find what I am looking for	3.35 (1.1)	2.94 (1.09)	.005	0.37
I can decide if information is relevant to my question	4.3 (0.74)	4.57 (0.54)	.005	−0.42
I can remember what I was searching for during the search process	4 (1.09)	4.8 (0.44)	<.001	−1.14
I feel lost during the search process	2.39 (1.2)	1.53 (0.85)	<.001	0.88
I can use online search systems independently	4.32 (0.97)	4.67 (0.71)	.003	−0.44
I feel overwhelmed when deciding which information is relevant	2.54 (1.38)	1.7 (1)	<.001	0.74

[a]Responses from participants who reported that they use web search. Full independent-samples *t* test statistics (*t* values, degrees of freedom, and exact *P* values) for all items are provided in Multimedia Appendix 4. Full Likert-scale response distributions are provided in Multimedia Appendix 5. 1=strongly disagree and 5=strongly agree.

[b]MCI: mild cognitive impairment.

Qualitative Data

Long-Term Use and Simplicity of the System

Survey participants were asked about why using these systems was easy or difficult. A prominent response was prior experience, reported by both the MCI or dementia group (“I have a background in IT...”) and the healthy older adult group (“I have used computers in my last job for many years”). Similarly, both survey groups reported personal use, “I have used it for many years, so have been familiar with it for a long time.” These findings align with the quantitative results showing high ease of use (mean 4.38, SD 0.8 overall) and reported high levels of independent use, particularly among healthy older adults (mean 4.67, SD 0.71 vs MCI or dementia mean 4.32, SD 0.97; *P*=.003), suggesting that familiarity underpins perceived ease. Moreover, survey responses often mentioned web search being habitual, “having used various search engines over many years it is almost second nature,” also consistent with quantitative findings of daily use and low feelings of overwhelm when inputting queries (MCI or dementia: mean 2.08, SD 1.24 vs healthy older adults mean 1.34, SD 0.79; *P*<.001). Focus group participants echoed this, noting challenges for learning AI: “I haven’t used it (referring to Gen-AI) because I’m a creature of habit... if I want an answer to something I Google it,” and “one of the necessary keys to people with memory problems in using new technology like AI is that we are people of habit. And that’s going to be the big, the big stumbling block.”

Some survey responses indicated that the enjoyment of discovering information facilitated ease of use, “I enjoy finding things out and have learned how to do this.” Notably, one MCI or dementia survey participant said, “the information is there to be found now that I have learned how to access it. Having family to help me would have been game changing but I have managed. I am acutely aware that as any tech’ ability I have learned… I shall lose these first, so lose this invaluable source of managing my life as Alzheimer’s progresses.” This highlights the impact of IR on independence and daily living and corresponds with survey findings showing lower feelings of being lost among healthy older adults (mean 1.53, SD 0.85) compared with the MCI or dementia group (mean 2.39, SD 1.2; *P*<.001), emphasizing the role of experience in maintaining confidence during searches.

Keyword Searching

Keyword searching was discussed in response to questions about ease of use, with survey participants describing it as both easy and difficult. Some participants with MCI or dementia found it straightforward: “I have always been good at asking questions,” while others found it more challenging: “it depends what I’m looking for. Finding the right words isn’t always easy. The more branches I go down, the more difficult I find it to get back,” and “the most frustrating aspect is when I am unable to select a search term that brings up correct answers... I’m only offered a limited selection of answers and I know there must be more but they don’t appear..." Consequently, some mentioned, “I sometimes have to change my question,” reflecting the quantitative finding that the MCI or dementia group responded

XSL•FO
RenderX

significantly more strongly with needing to reword questions (mean 3.35, SD 1.1) compared to the healthy older adult group (mean 2.94, SD 1.09; *P*=.005; Table 6).

Healthy older adults generally reported less difficulty: "simple to phrase a question," and that rephrasing, if needed, was straightforward: "can usually find what I want quickly but do need to consider how a question is phrased," aligning with quantitative findings that healthy older adults had greater confidence in phrasing questions (mean 4.2, SD 0.73 vs MCI or dementia mean 3.9, SD 1.03; *P*=.02). Both groups highlighted how keywords influence search results, with comments such as "depend on selecting suitable key search words, and the main problem is evaluating the reliability of sources of information..." and "with Google there are a lot of irrelevant web pages returned and it's often necessary to search through the links to find what you are looking for." Overall, there is a contrast between finding keyword searches easy or difficult across both groups, with an acknowledgment that the relevance of results often depends on this.

Focus group participants provided context for these survey patterns. Many participants reported that searching was easy, but occasionally required rephrasing. Participant 5 (female, aged 65 y, no impairment) said:

> *"I think I've always found what I wanted," adding, "...if the results are not relevant... I know I need to rephrase the question..."*

Participant 6 (male, aged 69 y, no impairment) mentioned that their frustration with web search makes Gen-AI appealing:

> *I wanted to start using it myself to- to find some information that I've not been able to get easily from Google...I think it could focus when I want a question answered...*

Overall, survey and focus group data indicate variability in ease of keyword searching, with rephrasing often necessary and keyword selection affecting relevance. Participants with MCI or dementia experienced more difficulty, reflecting quantitative differences.

Relevance of Results

Another identified theme was the relevance of results, particularly concerning advertisements, sponsorships, and location bias. Survey participants highlighted that commercialization reduces relevance, with the MCI or dementia group noting, "it would be easier if the results of searches didn't have such a converging and commercial aspect," and the older adult group stating, "the answers it gives are too random and the paid answers are always crowding out the simple ones." Both groups also reported irrelevant information: "too much irrelevant information given in answers," and "...I am not always convinced by the accuracy or bias in the info I find..." Suggested improvements from survey responses included: "by not showing promoted content first which is irrelevant to my question," "ditch the cookies, by the time I've got through this it's hard to remember what I was thinking about," and "less emphasis on American English terminology, and less bias towards American based info/research." Survey comments also suggested implementing "color-coded hyperlinks" to better distinguish between promoted and relevant results and that "search engines should organize their results so as to assist the searcher rather than according to commercial and financial priorities."

Expanding on the survey concerns, dislike for commercialization was also observed in the focus groups, where participant 5 (female, aged 65 y, no impairment) said, "...the first one that always will come up in the list is sponsored," while participant 3 (female, aged 78 y, no impairment) added, "it disturbs me that...the people who pay most money get up the queue..." This was also linked to trust, with participant 1 (male, aged 76 y, Alzheimer disease) noting that commercial prioritization can facilitate harm:

> *...they will draw revenue from people who are willing to pay to have their entry first in the list... there are a lot of people who get misled.*

Participants also described strategies to reduce risk, such as using privacy-focused browsers, "I use DuckDuckGo because it removes trackers …it just makes me feel more comfortable," although survey responses indicated a knowledge gap: "I wish there was a way of feeling safer when I do the searches."

When survey participants were asked what they would include if they could design their own system, responses addressed relevance filtering, source transparency, and removal of sponsored results, "...A button to click that would identify owners/contributors..," "1. Removal of sponsored responses; 2. Dating of posts to allow disregarding of outdated posts; 3. It may become necessary to identify AI responses derived from old or fake data that should not be relied upon." Interestingly, Gen-AI was suggested as a potential improvement to current systems by the older adult survey group, with comments such as "ChatGPT is useful because you can build a series of questions on a topic to look at different aspects of an issue and then draw on the answers to build a summary," and the MCI or dementia survey group, who said, "I manage my online searches quite easily as long as I am careful with my phrasing. AI is able to help me with that." Others acknowledged that these systems could be improved, stating, "I only know they could be better for me."

Finally, focus group data suggested that the volume of results can affect focus, particularly for participants with MCI or dementia. Participant 7 (male, aged 82 y, MCI) said, "if I ask Google a specific question... I would like a specific answer," whereas participant 3 (female, aged 78 y, no impairment) appreciated the element of choice, saying, "Google gives me choices so I can go in and I consult there." This may be reflected in the survey responses from the MCI or dementia group, which mentioned, "keeping what I am searching for in mind, sometimes I forget mid-search," and "remember what I am doing, short span of concentration, tired easily." The higher volume of results may impose a greater cognitive load or memory pressure when users must determine the most relevant result to select. This may be consistent with quantitative results depicting the MCI or dementia group rating being able to remember what they were searching for as lower (mean 4, SD 1.09) compared to the healthy older adult group (mean 4.8, SD 0.44; *P*<.001).

XSL•FO
RenderX

Participant Perspectives on Future Applications of AI

Building on the survey finding that 68.7% (169/246) of respondents reported openness to using Gen-AI in the future, this theme was explored primarily through qualitative data to understand how participants envisioned AI supporting future daily living needs. Specifically, many envisioned AI systems evolving beyond IR toward proactive monitoring, routine support, and safety. Survey responses acknowledged this:

> *happy with search but still trying to get to grips with AI which I feel has tremendous potential for people with my complaint.*

The MCI or dementia survey group mentioned a monitoring feature:

> *...AI would be ideal to monitor my behaviors...I would be quite happy for my AI to learn all about me and monitor the changes in my behaviors and routine.*

This was taken further in the same response, discussing how their conversational ability could be helpful to them: "...a bit of conversation is helpful when you live on your own!"

These future-oriented applications were not captured in the survey Likert-scale items and were therefore explored in more depth within the focus groups. Participant 4 (female, aged 78 y, no impairment) said, "...I live alone and at the moment I'm hale and hearty but I could have a camera and my daughter could see what I was up to...Things like that I think are good," indicating that monitoring might be accepted as part of a care plan. This was also discussed extensively by participant 1 (male, aged 76 y, Alzheimer disease) who explained:

> *I'm quite happy to share my personal antisocial habits with a computer if it is looking after me...*

This suggests that the potentially adaptable monitoring abilities of AI could be accepted if it improved independent living. However, they also explained that this technology would need to be implemented early in their diagnosis so it could develop alongside them: "...the technology and the help can grow with the disease as it progresses down the step ladder," the step ladder referring to diminishing cognitive decline, "you've got to grow into it... it's got to be led by the individual's needs." Interestingly, this conflicts with concerns about trust and privacy; when asked directly, their perspective on trust was different:

> *by the time my disease develops to the point where I'm going to need AI to look after me, my mind will have degenerated to such a state that I actually don't give two hoots who knows or cares or anything else...I won't have any privacy anyway, so that is really not of a major concern to me.*

A similar stance was taken by participant 5 (female, aged 65 y, no impairment), reflecting on a different meeting about robots used in care:

> *we all went into this meeting saying no way on earth would we have a robot for this. And we all left the meeting saying yes, absolutely, if it meant that we could stay at home...*

Finally, a few notable features were discussed in the focus groups. The first was that AI should learn to forget information, "necessity for AI to learn to forget... when it comes to bank details, jobs in assembly, there are things I think that would be good for AI to recognise it needs to forget…," and proactive autonomy, "AI responds to requests but what AI doesn't do is, is have the autonomous ability to, to, to access prompts ...if AI could be trained to think autonomously about where the needs are and what prompts are—that would be a major step forward in helping people with dementias and brain damage to live independently."

Discussion

Principal Findings

To our knowledge, this is the first study to gather in-depth perspectives of older adults with and without MCI (including MCI and early-stage dementia) on their engagement with and opinions of both traditional IR systems and Gen-AI. This contributes new empirical evidence to the literature on digital engagement in older adults, as most prior studies focus either on younger populations or on health information seeking alone [11,46]. Results show that most participants reported using traditional web search systems, consistent with the understanding that digital literacy in this population is growing [47,48]. Participants were recruited via convenience sampling, which may overrepresent older adults with higher digital literacy or motivation to engage with technology; findings should therefore be interpreted as indicative rather than representative of the broader population. Our findings confirm prior work suggesting that habitual use and familiarity enhance perceived ease of use among older adults [49,50], while also extending this by showing differential experiences in individuals with cognitive impairment. This familiarity likely reflects the long-established nature of web search platforms [12], whereas Gen-AI tools are newer and less commonly adopted, consistent with observations that novelty and lack of exposure limit adoption in older populations [51]. As a result, participants reported greater confidence and skill when using web search compared to Gen-AI. Together, these findings address our aim of examining how these groups use and perceive web search and Gen-AI tools.

Both groups generally found web search helpful, easy to use, and understood when and how to use it. Likert-scale responses showed significant group differences across multiple aspects of the search process, with effect size estimates ranging from moderate to large, indicating meaningful differences in user experience. The MCI or dementia group was more likely to report needing to rephrase questions, feeling lost, and feeling overwhelmed, whereas healthy older adults more often reported knowing how to phrase questions, finding relevant information, remembering search goals, and using systems independently. These findings confirm earlier literature showing that cognitive impairment can increase effort and reduce confidence in information-seeking tasks [52,53] and extend it by quantifying the differences in self-reported behaviors and perceptions. Although overall overwhelm remained low, results suggest that cognitive impairment may increase effort during information

XSL•FO
RenderX

seeking [52]. Despite group differences, quantitative data depicted that web search was generally viewed positively [54], while qualitative data provided deeper insight into these high ratings, highlighting issues such as commercialization, privacy concerns, and keyword selection, which can impact the user experience.

This adds qualitative insight to existing quantitative findings on older adults' web search behavior [52]. Qualitative responses reflect individual struggles based on beliefs and abilities, indicating that cognitive impairment shapes the subjective experience of IR systems [53,55,56]. Participants' proposed solutions indicated that challenges with web search are widespread and may be mitigated through Gen-AI features, with suggestions such as Gen-AI offering more personalized and adaptive responses. These findings complement research on adaptive AI for older adults [57,58] and highlight potential applications supporting independent information access. These findings extend the literature primarily focused on online health information quality [59-61] by highlighting the experiential differences and potential technological improvements for older adults with and without cognitive impairment (including MCI and early-stage dementia). They further suggest that emerging technologies such as Gen-AI may help offset age–related and cognition-related disparities in access [42,62].

Regarding Gen-AI, although reported usage was lower, more than half of nonusers expressed willingness to use these tools in the future, indicating potential uptake if barriers are addressed. Adoption appeared to be influenced by beliefs [51], including mistrust and lack of knowledge, consistent with prior studies showing that trust and perceived knowledge influence technology adoption among older adults [39,63-67]. However, qualitative findings suggest that low engagement cannot be attributed solely to knowledge gaps or mistrust. Some participants reported "no need" for these tools, reflecting perceptions of limited personal utility or preference for existing methods, while others identified practical barriers, such as complexity, navigation difficulties, access limitations, and reliability concerns. These observations extend adoption models [68], suggesting that perceived usefulness and interface design may be equally important in older adult populations.

These findings can be interpreted through the lens of the TAM [69], which proposes perceived usefulness and ease of use as key determinants of adoption. Participants recognized potential usefulness but reported limited adoption due to unfamiliarity and perceived complexity, indicating that both constructs were influential. Applying TAM situates these qualitative insights within an established theoretical framework [19,69]. Together, results suggest that utility, accessibility, design, trust, and perceived relevance interact to shape engagement with Gen-AI among older adults and those with MCI or early-stage dementia. This aligns with established links between technology adoption and trust [39,63-67,70]; however, actual uptake was not measured, and further work is needed to determine whether intentions translate into sustained engagement.

Our findings also highlight an important nuance related to participants' perceived self-efficacy when engaging with Gen-AI. Some focus group participants described feeling less capable than the AI, while others framed their interaction as treating it as a machine. These results extend literature on self-efficacy theory and technology adoption [71-73] by showing that perceived capability may influence engagement, a factor not typically emphasized within TAM. Even when participants recognized usefulness, confidence appeared shaped by perceived competence relative to the AI, particularly among individuals with cognitive impairments, who may already have lower baseline self-efficacy in technology use [72-74]. These perceptions suggest that Gen-AI may unintentionally exacerbate feelings of reduced capability for some users, despite recognized benefits. This highlights the importance of supportive interfaces, gradual exposure, and tailored training to strengthen self-efficacy. Our findings reinforce prior research [19,43] highlighting both barriers and quality of life benefits associated with AI adoption among older adults, while illustrating mechanisms through which Gen-AI may support independent living. Participants' willingness and motivation to learn indicate promise, although prior work shows that intention does not always translate to adoption [51,74,75]; therefore, findings should be interpreted as reflecting potential engagement rather than confirmed behavior change.

Knowledge gaps were common across both groups, indicating that education and training may support adoption. Participants expressed interest in tuition and demonstrations, particularly for ChatGPT [76], aligning with training practices already recognized [30]. However, views were heterogeneous: some participants welcomed structured training, while others viewed it as unnecessary or of limited value. These findings suggest that one-size-fits-all approaches may be ineffective and that training should be flexible and responsive to individual needs and motivation [77,78]. Tailored or optional learning approaches may therefore better support diverse preferences and abilities. Successful interventions may need to balance guidance with autonomy, allowing access to training without assuming universal interest. Structured, transparent education may improve confidence and understanding while complementing web searches and addressing perceived limitations.

Notably, the MCI or dementia group highlighted issues not raised by healthy older adults, including forgetting mid-search, difficulties learning new skills, and challenges in adopting new technologies. They also rated items such as being able to use Gen-AI independently with lower agreement compared to their older adult counterparts. These findings suggest that early introduction of AI and the inclusion of diverse cognitive groups in design processes are crucial. The adaptability and personalization of Gen-AI may help address these challenges [20], potentially supporting independent living outcomes.

Although only one group difference was statistically significant, effect size estimates indicated moderate-to-large differences across measures related to confidence, usability, and independence. This suggests that participants with MCI or dementia may experience greater difficulty when engaging with Gen-AI tools. However, moderate-to-large effect sizes alongside nonsignificant findings likely reflect limited statistical power due to the small number of Gen-AI users, and the observed effect sizes may help inform future power analyses and sample size targets in subsequent studies.

XSL•FO
RenderX

Furthermore, an emerging factor was monitoring, with participants recognizing AI's potential to support daily functioning. This aligns with research showing that chatbots may assist in monitoring and informing care providers about symptom changes [79]. Our findings extend this literature by demonstrating that older adults themselves view monitoring as valuable when independence benefits are clear. Despite privacy concerns, focus group responses expressed willingness to adopt AI if the benefits to independence were substantial. These findings highlight the importance of involving older adults in technology development, especially to address perceptions that AI is unnecessary or irrelevant to them. This is reflected in the theme of "no need," which appears linked to limited awareness of potential benefits [41], yet participants offered clear suggestions for how AI could improve daily life. Therefore, future developments should prioritize early implementation and co-development with older adults and those with cognitive impairment to ensure relevance and utility. Finally, these results complement literature showing that Gen-AI may reduce loneliness, support mental health, and improve person-centered information access, while reinforcing barriers such as ethical concerns, ageism, and digital literacy gaps [17,80]. Our findings demonstrate how these perceptions vary by cognitive status and identify actionable areas for intervention and training.

Limitations

The dementia and MCI groups were smaller than the healthy older adult group, which may limit representativeness. Additionally, one focus group included only 2 participants due to recruitment constraints; although the discussion generated rich data, larger groups may provide broader interactional insights in future research. Furthermore, participants were recruited through convenience sampling, which may overrepresent individuals with higher digital literacy or technology interest. Similarly, the number of current LLM users was small, although this reflects real-world usage patterns. Surveys were conducted online, restricting control over responses and preventing follow-up clarification, although focus groups helped mitigate this.

While demographic information such as age, gender, education level, occupation, and geographic location was collected, ethnicity and detailed socioeconomic indicators were not measured. This limits generalizability, as cultural and socioeconomic factors may influence engagement with web search and Gen-AI tools. Future research should recruit more diverse samples to explore these influences among older adults with and without cognitive impairment (including MCI and early-stage dementia).

Conclusions

Our results provide novel insights into how older adults with and without MCI or early-stage dementia may perceive and engage with web search and Gen-AI tools, addressing an underexplored gap in user experience research [81]. Participants expressed interest in Gen-AI, highlighting areas where training and support may facilitate adoption. While findings reflect perceptions and intentions rather than confirmed behavioral change, they identify factors influencing engagement with Gen-AI. They offer a flexible, personalized approach to information access and conversational search, which may help address limitations identified in web search systems. While web search remains valuable due to its long-standing use, participants highlighted areas for improvement, suggesting that hybrid or complementary approaches with Gen-AI could improve accessibility. The study contributes to an underexplored area of literature where potential applications have been recognized [20,82-84] and aims to inform future technological development grounded in user perspectives.

Future Work

As Gen-AI adoption grows, future research should explore targeted training interventions following best practices [30] to enhance adoption, confidence, and digital literacy among older adults with and without cognitive impairment (including MCI and early-stage dementia). Longitudinal research combining usage tracking and qualitative interviews may clarify evolving perspectives and sustained engagement. Expanding this research to include other underrepresented populations and conditions would be valuable for broadening understanding and assessing the generalizability of the findings. Developers should actively co-design technologies with these groups to combat digital ageism, ensure accessibility, and promote person-centered AI systems. Moreover, effectively evaluating how Gen-AI is integrated into these technologies will be essential to maximize their impact.

Acknowledgments

The authors would like to thank all participants for their time and contributions to this study. The authors also acknowledge the Join Dementia Research platform and the University of Strathclyde Older Adult Research Panel for supporting participant recruitment.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Focus group interview guide and ChatGPT demonstration script.
[DOCX File, 21 KB - aging_v9i1e85626_app1.docx]

Multimedia Appendix 2
Responses from participants who reported that they use generative artificial intelligence applications (1=strongly disagree and 5=strongly agree).
[DOCX File, 30 KB - aging_v9i1e85626_app2.docx]

Multimedia Appendix 3
Generative artificial intelligence Likert-response distributions.
[DOCX File, 21 KB - aging_v9i1e85626_app3.docx]

Multimedia Appendix 4
Responses from participants who reported that they use web search (1=strongly disagree and 5=strongly agree).
[DOCX File, 33 KB - aging_v9i1e85626_app4.docx]

Multimedia Appendix 5
Web search Likert-response distributions (%).
[DOCX File, 23 KB - aging_v9i1e85626_app5.docx]

References

1. Nichols E, Steinmetz JD, Vollset SE. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
2. Pan B, Hembrooke H, Joachims T, Lorigo L, Gay G, Granka L. In Google we trust: users' decisions on rank, position, and relevance. J Comput Mediat Commun 2007 Apr;12(3):801-823. [doi: 10.1111/j.1083-6101.2007.00351.x]
3. Dinet J, Bastien JMC, Kitajima M. What, where and how are young people looking for in a search engine results page? Presented at: IHM '10: Proceedings of the 22nd Conference on l'Interaction Homme-Machine; Sep 20-23, 2010. [doi: 10.1145/1941007.1941022]
4. Gossen T, Höbel J, Nürnberger A. Usability and perception of young users and adults on targeted web search engines. Presented at: IIiX '14: Proceedings of the 5th Information Interaction in Context Symposium; Aug 26-29, 2014. [doi: 10.1145/2637002.2637007]
5. Rowlands I, Nicholas D, Williams P, et al. The Google generation: the information behaviour of the researcher of the future. Aslib Proc 2008 Jul 6;60(4):290-310. [doi: 10.1108/00012530810887953]
6. Buhi ER, Daley EM, Fuhrmann HJ, Smith SA. An observational study of how young people search for online sexual health information. J Am Coll Health 2009;58(2):101-111. [doi: 10.1080/07448480903221236] [Medline: 19892646]
7. Yoon H, Jang Y, Vaughan PW, Garcia M. Older adults' internet use for health information: digital divide by race/ethnicity and socioeconomic status. J Appl Gerontol 2020 Jan;39(1):105-110. [doi: 10.1177/0733464818770772] [Medline: 29661052]
8. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
9. Yang E, Chang SJ, Ryu H, Kim HJ, Jang SJ. Comparing factors associated with eHealth literacy between young and older adults. J Gerontol Nurs 2020 Aug 1;46(8):46-56. [doi: 10.3928/00989134-20200707-02] [Medline: 32936926]
10. Yasin I, Irfan B, Reader JM, Fox‐Fuller JT. ChatGPT‐4 in clinical neurology: an Alzheimer's disease information quality analysis. Alzheimers Dement 2024 Dec;20(S4):e093401. [doi: 10.1002/alz.093401]
11. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022 Feb 16;24(2):e34790. [doi: 10.2196/34790] [Medline: 35171099]
12. Vercruyssen A, Schirmer W, Geerts N, Mortelmans D. How "basic" is basic digital literacy for older adults? Insights from digital skills instructors. Front Educ 2023;8:1231701. [doi: 10.3389/feduc.2023.1231701]
13. Patil R, Gudivada V. A review of current trends, techniques, and challenges in large language models (LLMs). Appl Sci (Basel) 2024;14(5):2074. [doi: 10.3390/app14052074]
14. Huang J, Chang KCC. Towards reasoning in large language models: a survey. Presented at: Findings of the Association for Computational Linguistics; Jul 9-14, 2023; Toronto, Canada p. 1049-1065. [doi: 10.18653/v1/2023.findings-acl.67]
15. Zhu Y, Yuan H, Wang S, et al. Large language models for information retrieval: a survey. ACM Trans Inf Syst 2025 Jan 31;44(1):1-54. [doi: 10.1145/3748304]
16. McCarren L, Kuoppamäki S. A LLM-based chatbot for mindfulness practice with older adults: a development and usability study. Stud Health Technol Inform 2025 May 15;327:888-889. [doi: 10.3233/SHTI250493] [Medline: 40380601]
17. Li P, Wu J, Shang S, Zhan Q. Harnessing large language model agents for healthy aging. Medicine Plus 2025 Jun;2(2):100084. [doi: 10.1016/j.medp.2025.100084]

XSL•FO
RenderX

18. Cheng HY. ChatGPT's attitude, knowledge, and clinical application in geriatrics practice and education: exploratory observational study. JMIR Form Res 2025 Jan 3;9:e63494. [doi: 10.2196/63494] [Medline: 39752214]
19. Wong AKC, Lee JHT, Zhao Y, Lu Q, Yang S, Hui VCC. Exploring older adults' perspectives and acceptance of AI-driven health technologies: qualitative study. JMIR Aging 2025 Feb 12;8:e66778. [doi: 10.2196/66778] [Medline: 39937162]
20. Treder MS, Lee S, Tsvetanov KA. Introduction to large language models (LLMs) for dementia care and research. Front Dement 2024;3:1385303. [doi: 10.3389/frdem.2024.1385303] [Medline: 39081594]
21. Hristidis V, Ruggiano N, Brown EL, Ganta SRR, Stewart S. ChatGPT vs Google for queries related to dementia and other cognitive decline: comparison of results. J Med Internet Res 2023 Jul 25;25:e48966. [doi: 10.2196/48966] [Medline: 37490317]
22. Enam MA, Murmu C, Dixon E. "Artificial intelligence - carrying us into the future": a study of older adults' perceptions of LLM-based chatbots. International Journal of Human–Computer Interaction 2025 Nov 2;41(21):13747-13770. [doi: 10.1080/10447318.2025.2476710]
23. Abdulazeem H, Borges do Nascimento IJ, Weerasekara I, et al. Use of digital health technologies for dementia care: bibliometric analysis and report. JMIR Ment Health 2025 Feb 10;12:e64445. [doi: 10.2196/64445] [Medline: 39928936]
24. Wang J, Liang Y, Cao S, Cai P, Fan Y. Application of artificial intelligence in geriatric care: bibliometric analysis. J Med Internet Res 2023 Jun 23;25:e46014. [doi: 10.2196/46014] [Medline: 37351923]
25. Saha K, Jain Y, Liu C, Kaliappan S, Karkar R. AI vs. humans for online support: comparing the language of responses from llms and online communities of Alzheimer's disease. Presented at: ACM Transactions on Computing for Healthcare; 2025 p. 3709366. [doi: 10.1145/3709366]
26. Alessa A, Al-Khalifa H. Towards designing a chatgpt conversational companion for elderly people. arXiv. Preprint posted online on Apr 18, 2023. [doi: 10.48550/arXiv.2304.09866]
27. Hossain G, Pomare ZS, Prybutok G. ChatGPT: a companion for dementia care. Presented at: 2024 IEEE International Conference on Consumer Electronics (ICCE); 2024; Las Vegas, NV, USA p. 1-6. [doi: 10.1109/ICCE59016.2024.10444253]
28. Pomare ZS, Hossain G, Maguluri DS, Vajpayee P, Prybutok G. ChatGPT in the context of dementia care and cognitive support. Presented at: 2024 IEEE International Conference on Electro Information Technology (eIT); May 30 to Jun 1, 2024. [doi: 10.1109/eIT60633.2024.10609883]
29. Xu J, Wei T, Hou B, Demiris G, Shen L. Revolutionizing dementia care: enhancing talk therapy with fine‐tuned large language models using GPT self‐generated data. Alzheimers Dement 2024 Dec;20(S4):e093496. [doi: 10.1002/alz.093496]
30. Gilman ES, Kot S, Engineer M, Dixon E. Training adults with mild to moderate dementia in ChatGPT: exploring best practices. Presented at: IUI '24 Companion: Companion Proceedings of the 29th International Conference on Intelligent User Interfaces; Mar 18-21, 2024; Greenville SC USA p. 101-106. [doi: 10.1145/3640544.3645230]
31. Ng R, Allore HG, Trentalange M, Monin JK, Levy BR. Increasing negativity of age stereotypes across 200 years: evidence from a database of 400 million words. PLoS ONE 2015;10(2):e0117086. [doi: 10.1371/journal.pone.0117086] [Medline: 25675438]
32. Wangler J, Jansky M. Media portrayal of old age and its effects on attitudes in older people: findings from a series of studies. Humanit Soc Sci Commun 2023;10(1):1-9. [doi: 10.1057/s41599-023-01671-5]
33. Swift HJ, Steeden B. Exploring representations of old age and ageing. : Centre for Ageing Better; 2020 URL: https://ageing-better.org.uk/sites/default/files/2020-03/Exploring-representations-of-old-age.pdf [accessed 2026-02-17]
34. Robertson DA, Savva GM, King-Kallimanis BL, Kenny RA. Negative perceptions of aging and decline in walking speed: a self-fulfilling prophecy. PLoS ONE 2015;10(4):e0123260. [doi: 10.1371/journal.pone.0123260] [Medline: 25923334]
35. Wurm S, Warner LM, Ziegelmann JP, Wolff JK, Schüz B. How do negative self-perceptions of aging become a self-fulfilling prophecy? Psychol Aging 2013 Dec;28(4):1088-1097. [doi: 10.1037/a0032845] [Medline: 24128074]
36. Pikhartova J, Bowling A, Victor C. Is loneliness in later life a self-fulfilling prophecy? Aging Ment Health 2016;20(5):543-549. [doi: 10.1080/13607863.2015.1023767] [Medline: 25806794]
37. Trigg R, Watts S, Jones R, Tod A, Elliman R. Self-reported quality of life ratings of people with dementia: the role of attitudes to aging. Int Psychogeriatr 2012 Jul;24(7):1085-1093. [doi: 10.1017/S1041610212000038] [Medline: 22377451]
38. Ivan L, Cutler SJ. Ageism and technology: the role of internalized stereotypes. University of Toronto Quarterly 2021 Jun;90(2):127-139. [doi: 10.3138/utq.90.2.05]
39. Knowles B, Hanson VL. The wisdom of older technology (non)users. Commun ACM 2018 Feb 21;61(3):72-77. [doi: 10.1145/3179995]
40. Abranches I, Alves LS, Santos LN, et al. Stigma and dementia: an integrative systematic literature review. Alzheimers Dement 2023 Jun;19(S5):e060165. [doi: 10.1002/alz.060165]
41. Knowles B, Hanson VL, Rogers Y, et al. The harm in conflating aging with accessibility. Commun ACM 2021 Jul;64(7):66-71. [doi: 10.1145/3431280]
42. Chu CH, Nyrup R, Leslie K, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. Gerontologist 2022 Aug 12;62(7):947-955. [doi: 10.1093/geront/gnab167] [Medline: 35048111]
43. Kaliappan S, Anand AS, Saha K, Karkar R. Exploring the role of LLMs for supporting older adults: opportunities and concerns. arXiv 2024 Nov 12:2411. [doi: 10.48550/arXiv.2411.08123]

XSL•FO
RenderX

44. Norman G. Likert scales, levels of measurement and the “laws” of statistics. Adv Health Sci Educ Theory Pract 2010 Dec;15(5):625-632. [doi: 10.1007/s10459-010-9222-y] [Medline: 20146096]
45. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
46. Saeidnia HR, Kozak M, Lund BD, Hassanzadeh M. Evaluation of ChatGPT’s responses to information needs and information seeking of dementia patients. Sci Rep 2024 May 4;14(1):10273. [doi: 10.1038/s41598-024-61068-5] [Medline: 38704403]
47. Briefing: facts and figures about digital inclusion and older people. : Age UK; 2024 Jun URL: https://www.ageuk.org.uk/siteassets/documents/reports-and-publications/reports-and-briefings/active-communities/internet-use-statistics-june-2024.pdf [accessed 2026-02-17]
48. Hunsaker A, Hargittai E. A review of Internet use among older adults. New Media & Society 2018 Oct;20(10):3937-3954. [doi: 10.1177/1461444818787348]
49. Yang HJ, Lee JH, Lee W. Factors influencing health care technology acceptance in older adults based on the Technology Acceptance Model and the Unified Theory of Acceptance and Use of Technology: meta-analysis. J Med Internet Res 2025 Mar 28;27:e65269. [doi: 10.2196/65269] [Medline: 40153796]
50. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults’ intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
51. Peek STM, Wouters EJM, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJM. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014 Apr;83(4):235-248. [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
52. Malinowsky C, Almkvist O, Kottorp A, Nygård L. Ability to manage everyday technology: a comparison of persons with dementia or mild cognitive impairment and older adults without cognitive impairment. Disabil Rehabil Assist Technol 2010;5(6):462-469. [doi: 10.3109/17483107.2010.496098] [Medline: 20545577]
53. Dommes A, Chevalier A, Lia S. The role of cognitive flexibility and vocabulary abilities of younger and older users in searching for information on the web. Appl Cogn Psychol 2011 Sep;25(5):717-726. [doi: 10.1002/acp.1743]
54. Hazen TJ, Olteanu A, Kazai G, Diaz F, Golebiewski M. On the social and technical challenges of web search autosuggestion moderation. arXiv. Preprint posted online on Jul 9, 2020. [doi: 10.5210/fm.v27i2.10887]
55. Sanchiz M, Chevalier A, Amadieu F. How do older and young adults start searching for information? Impact of age, domain knowledge and problem complexity on the different steps of information searching. Comput Human Behav 2017 Jul;72:67-78. [doi: 10.1016/j.chb.2017.02.038]
56. Mahmud S, Alvina J, Chilana PK, Bunt A, McGrenere J. Learning through exploration: how children, adults, and older adults interact with a new feature-rich application. Presented at: CHI ’20; Apr 25-30, 2020 URL: https://hci.cs.sfu.ca/MahmudCHI2020.pdf [doi: 10.1145/3313831.3376414]
57. Li HH, Liao YH. Application and effectiveness of adaptive AI in elderly healthcare. Psychogeriatrics 2025 Jan;25(1):e13214. [doi: 10.1111/psyg.13214] [Medline: 39537559]
58. Osei-Asiamah J, Hussain FHT, Chidipothu VK, Ahmed HS, Muhammad AHH, M PV. AI-driven adaptive user interfaces for enhancing interaction and accessibility for elderly users based on cognitive and physical abilities. Presented at: 2025 International Conference on Sustainability, Innovation & Technology (ICSIT); Aug 22-23, 2025; Nagpur, India p. 1-6. [doi: 10.1109/ICSIT65336.2025.11295099]
59. Schnelli A, Hirt J, Zeller A. Persons with dementia as internet users: what are their needs? A qualitative study. J Clin Nurs 2021 Mar;30(5-6):849-860. [doi: 10.1111/jocn.15629] [Medline: 33377238]
60. Zheng J, Gresham M, Phillipson L, et al. Exploring the usability, user experience and usefulness of a supportive website for people with dementia and carers. Disabil Rehabil Assist Technol 2024 May;19(4):1369-1381. [doi: 10.1080/17483107.2023.2180546] [Medline: 37086036]
61. Soong A, Au ST, Kyaw BM, Theng YL, Tudor Car L. Information needs and information seeking behaviour of people with dementia and their non-professional caregivers: a scoping review. BMC Geriatr 2020 Feb 14;20(1):61. [doi: 10.1186/s12877-020-1454-y] [Medline: 32059648]
62. Leukel J, Schehl B, Sugumaran V. Digital inequality among older adults: explaining differences in the breadth of Internet use. Information, Communication & Society 2023 Jan 2;26(1):139-154. [doi: 10.1080/1369118X.2021.1942951]
63. Knowles B, Hanson VL. Older adults’ deployment of ‘distrust’. ACM Trans Comput-Hum Interact 2018 Aug 31;25(4):1-25. [doi: 10.1145/3196490]
64. Altwala N, Bull C, Kirk D. Exploring the potential of AI chatbots for supporting individuals with dementia: thematic insights from user-centered workshops. Presented at: HAI '24: Proceedings of the 12th International Conference on Human-Agent Interaction; Nov 24-27, 2024 URL: https://dl.acm.org/doi/proceedings/10.1145/3687272 [doi: 10.1145/3687272.3690907]
65. Peled-Cohen L, Reichart R. A Systematic Review of NLP for Dementia: Tasks, Datasets, and Opportunities. Trans Assoc Comput Linguist 2024 Sep 29;13:1204-1244. [doi: 10.1162/TACL.a.35]
66. Sun X, Ma R, Zhao X, Li Z, Lindqvist J, Ali AE, et al. Trusting the search: unraveling human trust in health information from google and chatgpt. arXiv. Preprint posted online on Mar 15, 2024. [doi: 10.48550/arXiv.2403.09987]

XSL•FO
RenderX

67. Wolfe BH, Oh YJ, Choung H, et al. Caregiving artificial intelligence chatbot for older adults and their preferences, well-being, and social connectivity: mixed-method study. J Med Internet Res 2025 Mar 13;27:e65776. [doi: 10.2196/65776] [Medline: 40080043]
68. Venkatesh V, Thong J, Xu X, University of Arkansas, Hong Kong University of Science and Technology, The Hong Kong Polytechnic University. Unified Theory of Acceptance and Use of Technology: a synthesis and the road ahead. JAIS 2016;17(5):328-376. [doi: 10.17705/1jais.00428]
69. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
70. Baek C, Tate T, Warschauer M. "ChatGPT seems too good to be true": college students' use and perceptions of generative AI. Computers and Education: Artificial Intelligence 2024 Dec;7:100294. [doi: 10.1016/j.caeai.2024.100294]
71. Schmitter-Edgecombe M, Brown K, Luna C, et al. Partnering a compensatory application with activity-aware prompting to improve use in individuals with amnestic mild cognitive impairment: a randomized controlled pilot clinical trial. J Alzheimers Dis 2022;85(1):73-90. [doi: 10.3233/JAD-215022] [Medline: 34776442]
72. Benge JF, Kiselica AM, Aguirre A, et al. Technology use and subjective cognitive concerns in older adults. Arch Gerontol Geriatr 2023 Mar;106:104877. [doi: 10.1016/j.archger.2022.104877] [Medline: 36459914]
73. Lozoya SVM, Guirado MAZ, Zapata Gonzalez A, Lopez ABM. Use of technologies and self-efficacy in older adults. IEEE R Iberoamericana Tecnologias Aprendizaje 2022;17(2):125-130. [doi: 10.1109/RITA.2022.3166870]
74. Cohen Elimelech O, Rosenblum S, Tsadok-Cohen M, Ferrante S, Demeter N. Understanding older adults' technology use preferences and needs from a triangular perspective: qualitative study. J Med Internet Res 2025 Nov 11;27:e72716. [doi: 10.2196/72716] [Medline: 41218201]
75. Yap YY, Tan SH, Choon SW. Elderly's intention to use technologies: a systematic literature review. Heliyon 2022 Jan;8(1):e08765. [doi: 10.1016/j.heliyon.2022.e08765] [Medline: 35128090]
76. Skjuve M, Brandtzaeg PB, Følstad A. Why do people use ChatGPT? Exploring user motivations for generative conversational AI. FM 2024;29(1). [doi: 10.5210/fm.v29i1.13541]
77. Miller LMS, Callegari RA, Abah T, Fann H. Digital literacy training for low-income older adults through undergraduate community-engaged learning: single-group pretest-posttest study. JMIR Aging 2024 May 14;7:e51675. [doi: 10.2196/51675] [Medline: 38599620]
78. Carrasco-Dajer CM, Vera-Calzaretta AR, Ubillos-Landa S, Oyanedel JC, Díaz-Gorriti V. Impact of a culturally adapted digital literacy intervention on older people and its relationship with health literacy, quality of life, and well-being. Front Psychol 2024;15:1305569. [doi: 10.3389/fpsyg.2024.1305569] [Medline: 38686085]
79. Podoreanu B, Vasilateanu A, Marian CV. Chatbot for patients suffering from mild cognitive impairment. Procedia Comput Sci 2025;256:878-884. [doi: 10.1016/j.procs.2025.02.190]
80. Abadir PM, Battle A, Walston JD, Chellappa R. Enhancing care for older adults and dementia patients with large language models: proceedings of the National Institute on Aging-Artificial Intelligence & Technology Collaboratory for Aging Research Symposium. J Gerontol A Biol Sci Med Sci 2024 Sep 1;79(9):glae176. [doi: 10.1093/gerona/glae176] [Medline: 39001657]
81. Waycott J, Vetere F, Pedell S, Morgans A, Ozanne E, Kulik L. Not for me. Presented at: CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; May 7-12, 2016. [doi: 10.1145/2858036.2858458]
82. Gholizadeh S, Freeman AK, Botticelli MP. Conversational artificial intelligence for people living with dementia and their care partners: a scoping review. Alzheimer's & Dementia 2023 Dec;19(S19):e071965. [doi: 10.1002/alz.071965]
83. Belam G, Nilforooshan R. The use of artificial intelligence and machine learning in the care of people with dementia: a literature review. Eur Psychiatr 2021 Apr;64(S1):S429-S429. [doi: 10.1192/j.eurpsy.2021.1144]
84. Savoia SE, Stern EH, Salardini A, Dorai C. Usability and functional validation of artificial intelligence - enabled technology services in dementia care using a controlled trial in clinical practice. Alzheimers Dement 2021 Dec;17(S7):e056547. [doi: 10.1002/alz.056547]

Abbreviations

AI: artificial intelligence
Gen-AI: generative artificial intelligence
IR: information retrieval
LLM: large language model
MCI: mild cognitive impairment
TAM: Technology Acceptance Model

XSL•FO
RenderX

Edited by A Garcia; submitted 13.Oct.2025; peer-reviewed by A Okolie, R K V, R Boumans; revised version received 25.Feb.2026; accepted 11.Mar.2026; published 21.Apr.2026.

Please cite as:
Rogers C, McGeown WJ, Moshfeghi Y
Exploring Information Access in Aging Populations and Those With Dementia and Mild Cognitive Impairment in the United Kingdom: Survey and Focus Group Study
JMIR Aging 2026;9:e85626
URL: https://aging.jmir.org/2026/1/e85626
doi:10.2196/85626

© Claire Rogers, William J McGeown, Yashar Moshfeghi. Originally published in JMIR Aging (https://aging.jmir.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smartphone Use and Social Activities Among People With Mild to Moderate Dementia: Multi-Informant Cross-Sectional Study

Shuji Tsuda[1], MD, MPH, PhD; Satoko Hotta[2], PhD; Hiroshige Matsumoto[3], PhD; Kae Ito[1], MD, PhD

[1]Research Team for Human Care, Tokyo Metropolitan Institute for Geriatrics and Gerontology, 35-2 Sakaecho, Tokyo, Japan

[2]Graduate School of Health Management, Keio University, Kanagawa, Japan

[3]Department of Community Health Nursing/Public Health Nursing, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Shuji Tsuda, MD, MPH, PhD
Research Team for Human Care, Tokyo Metropolitan Institute for Geriatrics and Gerontology, 35-2 Sakaecho, Tokyo, Japan

Abstract

Background: Smartphones have become deeply embedded in daily life, supporting a range of social and practical activities. Individuals with dementia can potentially use smartphones to compensate for cognitive decline and maintain independence. However, while smartphones are widely studied in controlled research settings, little is known about how individuals with dementia spontaneously use them in everyday life. Understanding usage patterns and their potential link to social participation could inform strategies to support smartphone use in their social and practical daily activities.

Objective: This study aimed to identify factors associated with spontaneous smartphone use, describe usage patterns, and examine whether smartphone use is associated with participation in social activities.

Methods: In 2024, we conducted a cross-sectional survey among community-dwelling individuals with mild to moderate dementia from 17 medical facilities in Tokyo. Structured questionnaires were completed by the participants, their families, and attending physicians. Participation in social activities was assessed using the "Spending Time with Others" subscale of the Social Functioning in Dementia (SF-DEM) scale. Factors associated with smartphone use were analyzed using multinomial logistic regression. Associations with participation in social activities were assessed via hierarchical linear regression.

Results: Among 151 participants with a mean age of 82.9 (SD 6.6) years, 43 (29%) participants were regular smartphone users. Smartphone use was negatively associated with older age (odds ratio [OR] 0.41, 95% CI 0.21 - 0.80) and positively associated with longer education (OR 1.82, 95% CI 1.00 - 3.30), living alone (OR 3.17, 95% CI 1.08 - 9.31), and better cognitive function (OR 2.14, 95% CI 1.24 - 3.69). Common uses included calling, texting, taking photos and videos, and checking the news and weather. Smartphone users reported marginally more frequent participation in social activities than nonusers did (b=1.41, 95% CI –0.01 to 2.83), particularly visiting the homes of friends or family (b=0.43, 95% CI 0.08 - 0.78) and shopping together (b=0.49, 95% CI 0.15 - 0.83).

Conclusions: Despite their limited use at the time of this writing, commercial-based smartphones have the potential to support participation in social activities among individuals with dementia. Targeted support may help bridge the gap between this usage and the broader capabilities of these devices, enhancing their role in sustaining meaningful social participation.

(JMIR Aging 2026;9:e81927) doi:10.2196/81927

KEYWORDS

older adults; dementia; smartphone; everyday information and communication technology; social participation

Introduction

Background

Given the global trends in population aging and growth, an estimated 57.4 million people were living with dementia in 2019, and this number is projected to increase to 152.8 million by 2050 [1]. Dementia is characterized by progressive cognitive decline that gradually interferes with daily functioning and independence [2]. While subjective experiences vary, many people with dementia can maintain a good quality of life with appropriate support [3-5]. Participation in social activities is a key contributor to their well-being, particularly in the mild to moderate stages [5,6]. However, individuals with dementia tend to engage less in or withdraw from social activities before and after their diagnosis [7]. Although this decline may reflect a general tendency toward disengagement, many individuals still want to continue participating in activities that they find meaningful [8]. Supporting the well-being of people with dementia may therefore involve not only sustaining overall levels of participation but also enabling continued participation in personally meaningful activities [6].

XSL•FO
RenderX

Assistive technology, which ranges from purpose-built devices to research-based apps installed on everyday information and communication technologies (EICTs), has been widely explored as a support tool for people with dementia. It varies from memory aids and task-prompting systems to GPS tracking and communication technologies, targeting domains such as cognitive stimulation, assistance with daily activities, and social participation [9-11]. However, to date, most studies have been small-scale feasibility trials, and their findings often reveal only modest [10-12] or inconclusive effectiveness [13]. While modest evidence supports the use of assistive technology for memory support [10], task prompting [12], safety monitoring [9], and promoting social participation [11], assistive technology is rarely implemented in practice [14]. Users and caregivers frequently report barriers such as limited usability, low awareness, stigma, low accessibility, and inadequate technical support [15]. These challenges highlight the need to explore more pervasive alternatives to support the daily lives of individuals with dementia [10].

EICTs, particularly smartphones, have become deeply embedded in daily life. In 2022, smartphone ownership among adults ranged from 76% to 98% across 18 countries with available data; even among older adults aged 50 years and older, ownership rates reached 55% - 96% [16]. People increasingly rely on smartphones to engage in a variety of social and practical activities, such as sending messages, making video calls, searching for information, making digital payments, navigation, and booking restaurants [17]. The diversity of use reflects the fact that everyday social infrastructure is increasingly designed with smartphone users in mind. These integrated services facilitate self-initiated engagement and support the continuity of personally meaningful routines.

Importantly, some community-dwelling individuals with dementia spontaneously use smartphones in their daily lives [18,19]. They consider smartphones nonstigmatizing and enjoy exploring innovative ways to tailor them to their lifestyles [20]. Additionally, recent studies focused on training interventions have revealed positive learning effects and improved task performance in this population [21,22]. These findings suggest that smartphones can be commonly used by individuals with dementia and may serve as accessible tools to support autonomy and engagement. However, little is known about whether and how people with dementia spontaneously use smartphones and whether they leverage commercially available, but not research-based, apps and functions to maintain participation in social activities [10,11]. Gaining such insights could inform strategies to support individuals with dementia in effectively using commercially available EICTs to participate in their preferred social activities, thereby enhancing their well-being.

Objectives

In this study, we aimed to explore the role of smartphones in the lives of older adults with mild to moderate dementia. Specifically, we sought to (1) identify factors associated with spontaneous smartphone use, (2) describe these patterns of smartphone use in this population, and (3) examine whether smartphone users engage in more frequent social activities than nonusers do.

Methods

Study Design

We conducted a cross-sectional observational survey to explore smartphone use among individuals with mild to moderate dementia. The survey was carried out at 17 medical facilities where certified dementia support doctors, who have completed a 2-day governmental training course on community-based dementia care [23], provide outpatient dementia services. The study was conducted in collaboration with these doctors, who engaged with potential participants through outpatient services and provided their clinical information.

Study Settings

Between February and August 2024, data were collected from 17 medical facilities in Tokyo, including 4 hospitals and 13 primary care clinics. All participating hospitals were designated medical centers for dementia that provide memory clinic services through their psychiatric departments for outpatients and patients referred from nearby primary care clinics for diagnostic evaluation or management of severe behavioral and psychological symptoms of dementia (BPSD). The 4 hospitals exhibit varied institutional characteristics. They include an acute care hospital, a chronic care hospital, a psychiatric hospital, and a hospital offering a mix of acute and chronic care services, with bed capacities ranging from 126 to 550 beds. This diversity in hospital type and size reflects the heterogeneous nature of medical centers for dementia in Japan and allows inclusion of a range of clinical roles within regional dementia care.

The 13 primary care clinics were generally small, with 1-6 physicians affiliated with each clinic providing outpatient services on a rotating basis. Five clinics with multiple physicians offered dedicated outpatient sessions for dementia care. The remaining clinics provided care for individuals with dementia during general outpatient visits. In both clinic types, certified dementia support doctors typically collaborate with Medical Centers for Dementia to obtain diagnostic consultations and BPSD management. As this collaborative model between medical centers for dementia and primary care clinics is widely implemented throughout Japan under the national dementia policy, our study setting is likely to be representative of outpatient dementia care in other regions [24].

Participants

At the 17 participating medical facilities, the certified dementia support doctors screened individuals for eligibility based on the following criteria: (1) regularly attended medical appointments at the facility for longer than 6 months, (2) a clinical diagnosis of dementia, (3) a stage of mild cognitive impairment, mild dementia, or moderate dementia as determined using the "Functional Assessment Staging of Alzheimer's Disease" [25], and (4) sufficient cognitive capacity to provide informed consent. We excluded individuals with advanced medical conditions associated with comorbid conditions that suggested an end-of-life stage. The certified dementia support doctors screened their outpatients on a rolling basis until they reached a predetermined number of participants, which ranged from 5 to 30 depending on the volume of their outpatient clinic.

XSL•FO
RenderX

Procedure

Data were collected using 2 self-administered questionnaires, one for individuals with dementia (participants) and the other for their family members. Clinical information was recorded on a data sheet by the certified dementia support doctors. The doctors provided the participants and family members with the questionnaires and a return envelope addressed to the first author; the participants were instructed to return the completed questionnaires by postal mail within 2 weeks. The clinical data sheets were returned separately by the doctors via postal mail. A postal reminder was sent to participants if the completed questionnaires were not received within 2 weeks.

Participants were instructed to complete the questionnaires at home, with assistance from their family members only when necessary. To minimize reporting bias, the family members were instructed not to influence the participants' responses. The family members completed their own questionnaires independently and were requested not to discuss or share their responses with the participants. The doctors were blinded to the participants' and families' responses and were instructed to administer the Mini-Mental State Examination (MMSE) based on standardized written guidance.

Instruments

Participant Questionnaire

The participant questionnaire was used to collect the participants' baseline characteristics, including age, sex, years of education, marital status (married, single, divorced, or widowed), living arrangements (cohabiting or living alone), and perceived economic status (very poor, poor, fair, good, or very good). Factual information such as age and sex was cross-verified using the doctors' clinical data sheets as references. Sex was confirmed in 148 out of 151 (98%) cases, with the remaining 3 (2%) cases unverified owing to missing responses. For age, 131 out of 151 (87%) entries corresponded, while 17 (11%) entries differed by approximately 1 year, likely owing to differences in the time of data entry or recall errors. The 3 participants who did not report their sex also did not report their age. For consistency, we used the clinical sheet data for these variables.

The participant questionnaire also included a section on their mobile phone use, including daily use (smartphones, cellphones, or no use) and years of use. We provided smartphone users with a list of 25 general smartphone apps and functions (eg, calling, texting, maps, and videos) and inquired whether they used them weekly, occasionally, or never (Multimedia Appendix 1). The selection and categorization of apps and functions were informed by governmental surveys of the general public and by previously published studies examining tablet use among people with dementia [26,27]. Multinomial logistic regression was conducted with nonusers as the reference category. All independent variables were categorized as shown in the parentheses, with the first category used as the reference group in the regression analysis.

Family Questionnaire

The family questionnaire was used to collect demographic information, including age, sex, family relationship with the participant, and the frequency of time spent together (categorized as daily, weekly, monthly, or yearly). The remaining sections inquired about the individuals with dementia to assess their daily cognitive functioning, BPSD, and daily social functioning.

To measure daily cognitive functioning, the Dementia Assessment Sheet for Community-based Integrated Care System 8-items (DASC-8) was administered [28]. It is a validated tool for use in community and primary care settings [28] and is widely applied in Japan to assess cognitive and functional decline in older adults. It comprises 8 items covering memory, orientation, and basic and instrumental activities of daily living. Each item was rated by the family member on a scale ranging from 0 (independent) to 3 (fully dependent). In the present study, DASC-8 results demonstrated good internal consistency (Cronbach α=0.83).

To measure BPSD, the 5-item short-form version of the Dementia Behavior Disturbance scale (DBD-5) was used [29]. The DBD-5 is a brief caregiver-rated questionnaire developed to screen for BPSD in community and clinical settings; key behavioral issues such as aggression, agitation, and resistance to care are rated on a 5-point Likert scale ranging from 0 (never) to 4 (always). Moderate internal consistency of DBD-5 results was observed in this study (Cronbach α=0.67).

Daily social functioning was measured using the family-rated version of the Japanese Social Functioning in Dementia scale (SF-DEM-J) [30]. We selected its "Spending Time with Others" subscale for the purpose of this study because it captures clearly differentiated types of social activities that are relevant to those potentially supported or facilitated by smartphone functions [31]. This subscale contains 7 items related to social activities with friends or family, which broadly correspond to levels 3 and 4 of Levasseur taxonomy of social activities, encompassing social contact without performing a specific activity and performing an activity with others [32]. We opted for a carer-reported tool, as it has demonstrated higher reliability than patient-reported measures have [33]. Family members rated the social activities of the individuals with dementia over the previous month on a 4-point scale, ranging from 0 (never) to 3 (very often), with acceptable internal consistency (Cronbach α=0.69). Additionally, one question was derived from a different subscale of the SF-DEM-J to confirm that mobile phone use was associated with communication through phones.

Doctor's Data Sheet

The doctors recorded information on participants' dementia etiologies by selecting the following from a predefined list: Alzheimer disease, Lewy body dementia, frontotemporal dementia, vascular dementia, mixed dementia, others, or unknown. They administered the MMSE in accordance with standardized guidance and documented the resulting scores. Participants' comorbidities were selected from a predefined list on the data collection form, and the total number of comorbidities was calculated for each participant.

XSL•FO
RenderX

Data Analysis

The primary outcome of this study was the frequency of social activities, assessed using the "Spending Time with Others" subscale of the SF-DEM-J. The main exposure variable was the type of mobile phone used. Based on their self-reported usage, the participants were categorized into 3 groups, including smartphone users, cellphone users, and nonusers, with "use" defined as operating the device at least once a week. While the original study objective was to examine the association between smartphone use and social activity among individuals with dementia, a substantial portion of the participants reported using cellphones. Including the data of this group in the analysis provided an opportunity to explore a more nuanced gradient of technology use and its potential relationship with social engagement. Nonusers were used as the reference group in all analyses.

For objective (1), frequencies and percentages were calculated for each group (smartphone users, cellphone users, and nonusers) and compared using chi-square tests. Factors associated with the use of each device were modeled using multinomial logistic regression, with nonusers as the reference category. While exploring the most appropriate model using a list of candidate variables, we first eliminated those with weak associations (P<.20) in the chi-square tests. Spearman correlation coefficients were subsequently calculated among the remaining candidate variables. For each pair of variables with a correlation coefficient >0.40, one was considered for elimination. The variables retained in the final model were selected based on logical considerations and prior research evidence [34].

For objective (2), we assessed the descriptive data of 25 general smartphone apps and functions used by smartphone users.

For objective (3), we compared the total score for the frequency of social activities among smartphone or cellphone users with that among nonusers using linear regression analysis. To confirm the robustness of the association, comparisons were made using 3 different models: Crude, Model 1 that was adjusted for selected baseline characteristics, and Model 2 that was additionally adjusted for cognitive function. An item-level analysis of the frequency of each social activity was performed with adjustment for the Model 2 variables.

Missing values in the family-reported scales (ie, DASC-8, DBD-5, and SF-DEM-J subscale) accounted for less than 2% of the responses for each item; hence, they were imputed using the mean value of the family's available responses on the respective scale. Statistical significance was set at P<.05, and the Bonferroni method was applied to adjust for multiple comparisons in the interpretation of P values in the pairwise comparisons. All statistical analyses were performed using R version 4.4.3 (R Foundation for Statistical Computing).

Ethical Considerations

Written informed consent was obtained from the individuals with dementia and their family members to ensure that all parties were fully informed about the study's purpose, procedures, potential risks, and benefits. They were also assured that their participation was entirely voluntary and that they could withdraw at any time without consequences. Their survey responses would remain confidential and neither be shared with their attending physicians nor linked to any clinical evaluations. In addition, data were anonymized and stored securely, with access limited to the research team. Participants received a small token of appreciation for their participation. All the study procedures were approved by the Tokyo Metropolitan Institute for Geriatrics and Gerontology Research Ethics Committee (approval number: R23-090).

Results

Baseline Characteristics

A total of 165 individuals with dementia were screened, among whom 151 participated in the survey (response rate 92%). Their mean age was 82.9 years (SD 6.6), 95 out of 151 (63%) participants were women, and their average MMSE score was 20.9 (SD 4.6). Dementia etiologies were Alzheimer disease (107/151, 71%), Lewy body dementia (17/151, 11%), mixed dementia (17/151, 11%), vascular dementia (3/151, 2%), frontotemporal dementia (1/151, 1%), and other or unknown (6/151, 4%). Out of 151 individuals, 43 were smartphone users (29%), 30 were cellphone users (20%), and 78 were nonusers (52%). The family members had a mean age of 63.9 years (SD 12.8), and out of 151 individuals, 112 (75%) were women. Their relationships with the participants were spouse (53/151, 35%), child (90/151, 60%), son- or daughter-in-law (4/151, 3%), and others (4/151, 3%). The majority (103/151, 68%) lived with the participants, and they spent time together on a daily (112/151, 74%), weekly (27/151, 18%), or monthly (12/151, 8%) basis.

Differences among participants recruited from hospitals and clinics were observed in several socioeconomic and functional characteristics (Multimedia Appendix 2). Participants recruited from the 4 hospitals tended to be younger, more often male, had higher educational attainment, were married, and reported a better perceived economic status than those recruited from the 13 primary care clinics. Participants recruited from hospitals also demonstrated slightly higher levels of cognitive and physical functioning.

Mobile Phone Users

Smartphone users were significantly younger (P=.002), had more years of education (P=.002), and demonstrated better cognitive function (P=.004 for MMSE score and P<.001 for DASC-8 score) than nonusers did (Table 1). The multivariable predictive model (Table 2) confirmed that an older age was associated with a lower likelihood of smartphone use (odds ratio [OR] 0.41, 95% CI 0.21-0.80; P=.009), while more years of education (OR 1.82, 95% CI 1.00-3.30; P=.05) and a higher MMSE score (OR 2.14, 95% CI 1.24-3.69; P=.007) were positively associated with smartphone use. In addition, living alone was significantly associated with a higher likelihood of smartphone use (OR 3.17, 95% CI 1.08-9.31; P=.04).

XSL•FO
RenderX

Table . Participant characteristics by mobile phone use category.

Participant characteristics	Smartphone use (n=43 28.5%), n (%)	Cellphone use (n=30, 19.9%), n (%)	Nonuse (n=78, 51.7%), n (%)	Overall, *P* value	1 versus 3[a], *P* value	2 versus 3, *P* value
Age (years)				.01	.002	.45
≤74	8 (38.1)	5 (23.8)	8 (38.1)			
75 - 84	28 (37.8)	14 (18.9)	32 (43.2)			
≥85	7 (12.5)	11 (19.6)	38 (67.9)			
Sex				.07	.04	.07
Men	20 (35.7)	14 (25.0)	22 (39.3)			
Women	23 (24.2)	16 (16.8)	56 (58.9)			
Education (years)				.01	.002	.45
≤9	5 (11.9)	8 (19.0)	29 (69.0)			
10 - 12	15 (28.3)	11 (20.8)	27 (50.9)			
≥13	23 (42.6)	11 (20.4)	20 (37.0)			
Marital status				.14	.04	.50
Married	28 (35.0)	16 (20.0)	36 (45.0)			
Single, divorced, or widowed	15 (21.1)	14 (19.7)	42 (59.2)			
Living arrangement				.18	.17	.09
Cohabiting	32 (26.9)	21 (17.6)	66 (55.5)			
Living alone	11 (34.4)	9 (28.1)	12 (37.5)			
Perceived economic status				.16	.04	.77
Good	21 (41.2)	10 (19.6)	20 (39.2)			
Fair	20 (23.5)	17 (20.0)	48 (56.5)			
Poor	2 (15.4)	3 (23.1)	8 (61.5)			
Number of comorbidities				.23	.39	.26
0	5 (26.3)	6 (31.6)	8 (42.1)			
1 - 2	22 (36.1)	8 (13.1)	31 (50.8)			
≥3	16 (22.5)	16 (22.5)	39 (54.9)			
Dementia etiologies				.50	.56	.35
Alzheimer disease	28 (26.2)	25 (23.4)	54 (50.5)			
Lewy body dementia	5 (29.4)	1 (5.9)	11 (64.7)			
Mixed dementia	6 (35.3)	4 (23.5)	7 (41.2)			
Vascular dementia	0 (0.0)	0 (0.0)	3 (100.0)			

XSL•FO
RenderX

Participant characteristics	Smartphone use (n=43 28.5%), n (%)	Cellphone use (n=30, 19.9%), n (%)	Nonuse (n=78, 51.7%), n (%)	Overall, *P* value	1 versus 3[a], *P* value	2 versus 3, *P* value
Frontotemporal dementia	1 (100.0)	0 (0.0)	0 (0.0)			
Others or unknown	3 (50.0)	0 (0.0)	3 (50.0)			
MMSE[b] score				.015	.004	.12
≤19	9 (17.0)	8 (15.1)	36 (67.9)			
20 - 23	15 (28.8)	12 (23.1)	25 (48.1)			
24 - 30	19 (43.2)	10 (22.7)	15 (34.1)			
DASC-8[c] score				<.001	<.001	.006
Intact cognitive function (≤10)	2 (22.2)	5 (55.6)	2 (22.2)			
Mild dementia (11-16)	24 (53.3)	8 (17.8)	13 (28.9)			
Mild to moderate dementia (≥17)	17 (17.7)	16 (16.7)	63 (65.6)			
DBD-5[d] score				.16	.10	.19
5 - 9	21 (36.8)	13 (22.8)	23 (40.4)			
10 - 12	11 (21.6)	11 (21.6)	29 (56.9)			
13 - 25	11 (26.2)	5 (11.9)	26 (61.9)			

[a]Differences across groups tested using chi-square test.

[b]MMSE: Mini-Mental State Examination.

[c]DASC-8: Dementia Assessment Sheet for Community-based Integrated Care System 8-item.

[d]DBD-5: Dementia Behavior Disturbance scale.

XSL•FO
RenderX

Table . Factors associated with mobile phone use compared to nonuse.

Participant factor[a]	Smartphone use		Cellphone use	
	OR[b] (95% CI)	*P* value	OR (95% CI)	*P* value
Age (years), ≤74, 75-84, and ≥85 (years)	0.41 (0.21-0.80)	.009	0.67 (0.33-1.36)	.26
Gender (men and women)	0.68 (0.26-1.74)	.42	0.43 (0.16-1.17)	.10
Education (years), ≤9, 10-12, and ≥13 (years)	1.82 (1.00-3.30)	.05	1.08 (0.59-1.98)	.80
Living arrangement (cohabiting, living alone)	3.17 (1.08-9.31)	.04	3.37 (1.12-10.13)	.03
MMSE[c] score (8-19, 20-23, 24-30)	2.14 (1.24-3.69)	.007	1.67 (0.94-2.98)	.08

[a]Multinomial logistic regression was conducted with nonusers as the reference category. All independent variables were categorized as shown in the parentheses, with the first category used as the reference group in the regression analysis.

[b]OR: odds ratio.

[c]MMSE: Mini-Mental State Examination.

Cellphone users presented less distinctive characteristics. Age, years of education, and MMSE score were not significantly associated with cellphone use (*P*=.45, *P*=.45, and *P*=.12, respectively; Table 1). In the multivariable predictive model, women were marginally less likely to use cellphones (OR 0.43, 95% CI 0.16-1.17; *P*=.10), and a higher MMSE score was marginally associated with cellphone use (OR 1.67, 95% CI 0.94-2.98; *P*=.08). The only factor that was significantly associated with cellphone use was living arrangements: cellphone users were 3.37 times more likely to live alone (OR 3.37, 95% CI 1.12-10.13; *P*=.03).

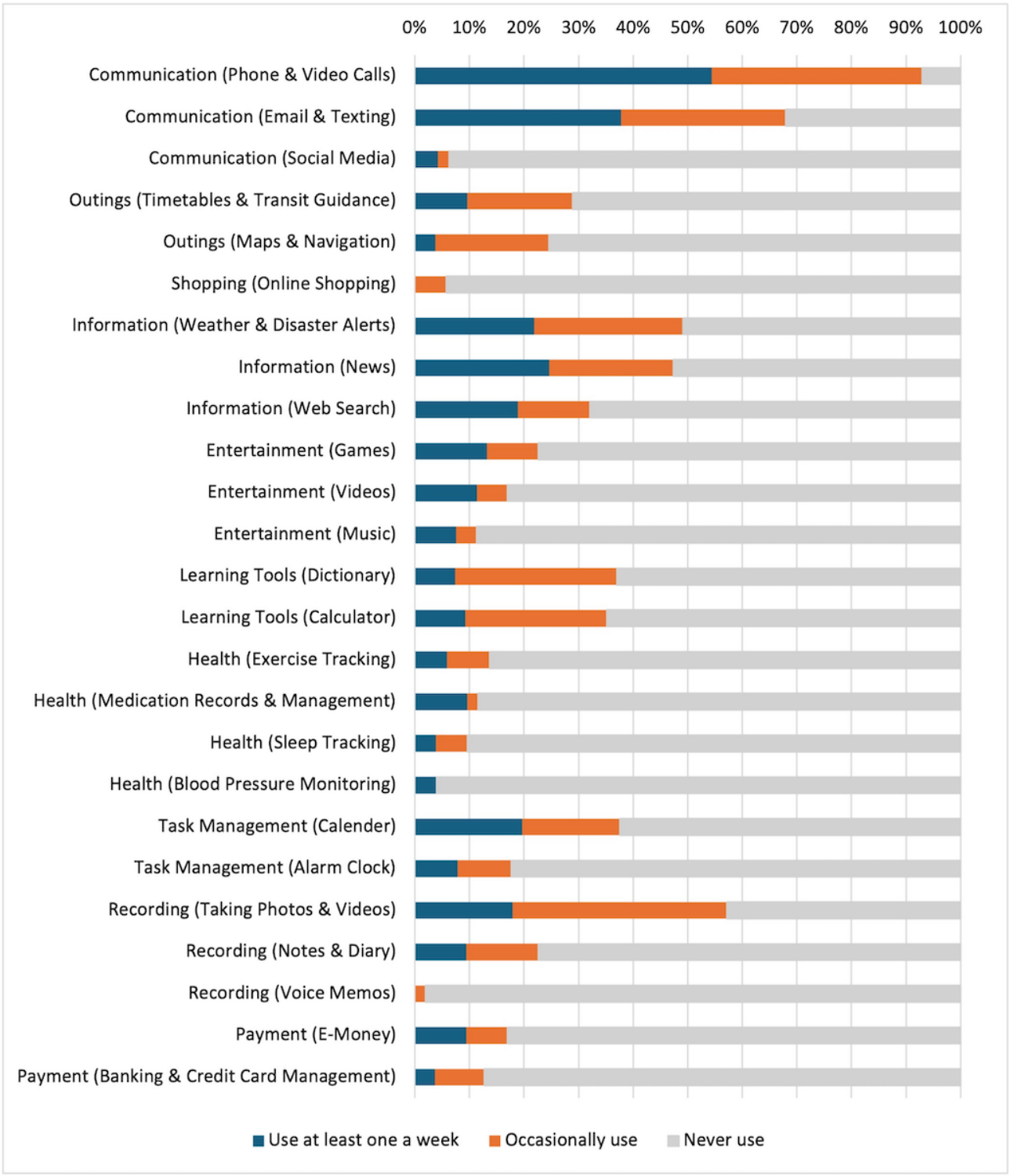
Patterns of Smartphone Use

Among the 43 smartphone users, more than 3-quarters had been using a smartphone for longer than 3 years. Specifically, the duration of smartphone use was up to 1 year for 5 out of 43 (11%) individuals, up to 3 years for 5 out of 43 (11%) individuals, up to 5 years for 14 out of 43 (32%) individuals, and up to 10 years for 20 out of 43 (46%) individuals.

Smartphones were used most often for communication, reading the news, and taking photos and videos (Figure 1). Out of 43 smartphone users, 40 (93%) made phone and video calls more than once a week or occasionally, while 27 (63%) used email and texting. Approximately half of the smartphone users used weather and disaster alerts (20/43, 47%), accessed news (20/43, 47%), and took photos and videos (24/43, 56%). However, apps and functions that could assist out-of-home activities were less likely to be used, as only 11 out of 40 (26%) used transit guidance and navigation systems and 7 out of 40 (16%) made digital payments. Regarding cognitive aid features, 16 out of 40 (37%) used calendars, 9 out of 40 (21%) used notes and diaries, and 7 out of 40 (16%) used alarm clocks. Health monitoring or tracking apps (such as for exercise, medication, and sleep) were used by 5 out of 40 (12%) or fewer smartphone users.

XSL•FO
RenderX

Figure 1. Purposes of smartphone use (n=43).



Mobile Phone Use and Participation in Social Activities

The mean scores on the "Spending Time with Others" subscale of the SF-DEM-J were 6.0 (SD 3.6) for smartphone users, 5.1 (SD 4.1) for cellphone users, and 4.4 (SD 3.2) for nonusers. We first confirmed that mobile phone use was associated with more frequent telephonic communication with friends and family members. After adjusting for age, sex, years of education, living arrangements, and MMSE score, both smartphone use (b=1.07, 95% CI 0.73-1.41; *P*<.001) and cellphone use (b=0.73, 95% CI 0.37-1.10; *P*=.001) showed significant associations.

Table 3 shows the results of the stepwise linear regression analyses examining the association between mobile phone use and the frequency of social activities, as measured using the "Spending Time with Others" subscale of the SF-DEM-J, with nonusers as the reference group. In the crude model, smartphone use was significantly associated with higher social activity scores (b=1.55, 95% CI 0.24-2.86; *P*=.02). This association remained marginally significant after adjusting for age, sex, years of education, and living arrangements in Model 1 (b=1.26, 95% CI −0.13 to 2.64; *P*=.07) and additionally for MMSE score

XSL•FO
RenderX

in Model 2 (b=1.41, 95% CI −0.01 to 2.83; *P*=.05). Cellphone use, by contrast, was not significantly associated with social activity levels in any of the models.

Table . Association between mobile phone use and frequency of social activities.

Participant factor	Crude[a]		Model 1[b]		Model 2[c]	
	b (95% CI)	*P* value	b (95% CI)	*P* value	b (95% CI)	*P* value
Smartphone use	1.55 (0.24 to 2.86)	.02	1.26 (–0.13 to 2.64)	.07	1.41 (–0.01 to 2.83)	.05
Cellphone use	.60 (–0.90 to 2.09)	.44	.58 (–0.92 to 2.09)	.45	.68 (–0.83 to 2.20)	.38
Age (years), 65-74, 75-84, and ≥85 (years)	–1.13 (–1.96 to −0.30)	.008	–.87 (–1.75 to 0.02)	.06	–.80 (–1.70 to 0.09)	.08
Gender (men and women)	–.19 (–1.38 to 1.01)	.76	.62 (–0.67 to 1.91)	.35	.56 (–0.73 to 1.86)	.39
Education (years) ≤9, 10-12, and ≥13 (years)	.59 (–0.13 to 1.30)	.11	.27 (–0.51 to 1.04)	.50	.33 (–0.46 to 1.11)	.41
Living arrangement (cohabiting or living alone)	–.68 (–2.08 to 0.73)	.35	–.77 (–2.22 to 0.69)	.30	–.76 (–2.21 to 0.69)	.30
MMSE[d] score (8-19, 20-23, and 24-30)	–.13 (–0.84 to 0.58)	.72	—[e]	—	–.34 (–1.07 to 0.39)	.36

[a]Stepwise linear regression analysis was conducted to assess the association between mobile phone use and the subscale score, with nonusers as the reference group.

[b]Model 1 adjusted for baseline characteristics (age, sex, education, and living arrangement).

[c]Model 2 additionally adjusted for cognitive function (MMSE).

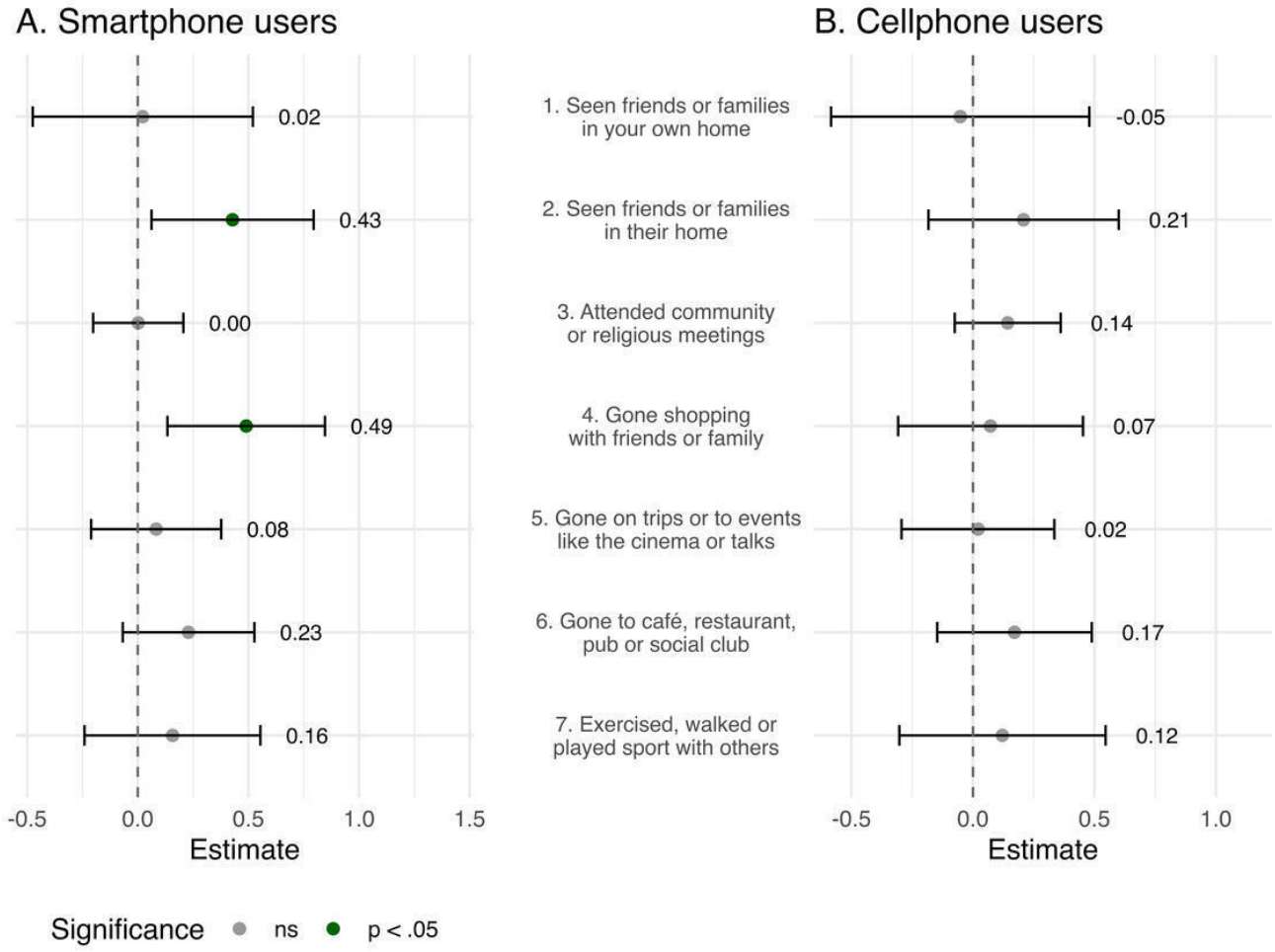
[d]MMSE: Mini-Mental State Examination.

[e]Not applicable.

Figure 2 presents the item-level analysis of the association between mobile phone use and the frequency of specific social activities, adjusted for the covariates in Model 2. Smartphone users reported significantly more frequent visits to the homes of friends or family members (b=0.43, 95% CI 0.08-0.78; *P*=.02) and shopping trips with friends or family members (b=0.49, 95% CI 0.15-0.83; *P*=.005) than did nonusers. Most of the other social activities showed nonsignificant but positive associations with smartphone use. In contrast, among cellphone users, none of the individual items reached statistical significance, although many showed positive trends.

XSL•FO
RenderX

Figure 2. Forest plot of item-level associations.



All estimates were adjusted for baseline characteristics and cognitive function, including age, sex, education, living arrangements, and MMSE scores.

Discussion

Principal Findings

In this cross-sectional study of community-dwelling older adults with mild to moderate dementia, 29% were regular smartphone users, most for longer than 3 years. Smartphone use was independently associated with a younger age, more years of education, living alone, and better cognitive function. The top five smartphone uses were (1) making phone and video calls, (2) sending emails and texts, (3) taking photos and videos, (4) checking the weather, and (5) reading the news. Functions that could more directly aid cognitive function (eg, notes and alarms) and out-of-home social participation (eg, GPS navigation and digital payments) were not often used. Although both smartphone and cellphone users reported more frequent telephonic communication with their friends or family members, only smartphone use was associated with a higher overall score for participation in social activities after adjusting for demographic and cognitive variables. At the activity level, smartphone users reported significantly more frequent visits to the homes of friends and family members and more frequent shopping trips with friends and family members than did nonusers.

Predictors of Smartphone Use

While our participants used smartphones less often than the general older adult population does, their adoption was observed among a substantial proportion of older adults with dementia in their 70s and 80s (mean age 82.9, SD 6.6 years). In 2023, the rates of smartphone use among the general older adult population in Japan were 71% for those aged 65 - 74 years and 38% for those aged 75 years and older [17]. Among our participants with dementia in the same age groups, the rates were lower at 38% and 23%, respectively. Directly comparable data are limited because most studies on smartphone use among individuals with dementia are feasibility or qualitative studies with small sample sizes [10,20,35]; nevertheless, these studies provide some context. In 2016, 75% of 53 individuals (mean age 64.9, SD 13.3 years) who visited a memory clinic for cognitive evaluation in the United States owned a smartphone [36]. In France (2013 - 2015), 33% of 323 older adults (mean age 75.9, SD 7.0 years) who underwent cognitive assessment at a memory clinic reported the regular use of smartphones or tablets [18]. Considering that these rates are from approximately a decade ago, when smartphone adoption among the general older population was 25% - 55% lower [16], the relatively low use among our participants likely reflects their older age and lower cognitive function rather than a persistent trend among individuals with dementia. As smartphone use has become more widespread among older adults in general [16,17], a similar increase can be expected in individuals with dementia.

XSL•FO
RenderX

Consistent with previous findings, we found that a younger age, higher educational level, and better cognitive function were associated with smartphone use [10]. Notably, living alone emerged as a significant predictor of smartphone use. This trend has rarely been explored [10] but may reflect the increasing number of individuals with dementia who live alone [37] and the heightened risk of loneliness in this group [38]. Smartphones may serve as a compensatory tool for maintaining independent living and social connections among these individuals, and this warrants further investigation [39].

Usage Patterns

Commonly reported smartphone uses included communication, searching for information, and taking photos or videos, while the use of features that could more specifically aid cognitive function and out-of-home activities was limited. Similar trends were reported in a US study conducted in 2016, where social and general-purpose functions (eg, email, photos, and news) were favored over cognitive aids (directions, to-do lists, alarms, and reminders) [36]. However, a more recent UK study (2021), in which 29 younger adults with dementia (aged 56 - 78 years) were interviewed, revealed that they actively explored and enjoyed using various smartphone features, including calendars, reminders, and health-tracking apps [20]. These individuals likely developed smartphone proficiency before cognitive decline, which enabled them to integrate EICTs into their daily lives despite their impairments [39]. Given that cognitive aid features have proven both preferred and effective in prior research [10,14], our findings suggest that older individuals such as those who participated in our study may benefit from additional support to effectively use them [34].

Associations With Participation in Social Activities

Although both smartphone and cellphone use were associated with more frequent phone communication with friends and family members, only smartphone use was associated with greater overall social activity. Prior research has revealed that research-based assistive technologies and EICTs, which typically apply cognitive aid features and are tailored to the needs of individuals with dementia, can supplement daily cognitive functions [10,12] and enhance social participation [11]. Our findings suggest that the benefits of these specialized technologies may also be transferable to commercially available smartphones. However, given the limited usage patterns observed among our participants, there remains untapped potential to expand the impact of smartphones through improved design and tailored support [14,20].

Smartphone use was significantly associated with visiting the homes of friends and family members and shopping together, with a marginal significance also observed for going to cafés, restaurants, pubs, or social clubs. In contrast, no significant association was found for attending community or religious meetings, going on trips, or participating in events. These findings align with those of a multinational study of 163 individuals with mild dementia, in which the authors reported that out-of-home participation was most frequent in familiar, informal activities such as visiting friends and local shopping, whereas engagement in more complex or unfamiliar activities was less common [40]. This pattern suggests that individuals with dementia use smartphones to stay engaged in familiar, emotionally rewarding routines that require minimal cognitive planning [8]. Accordingly, support strategies should be focused on enabling them to sustain meaningful activities with familiar social connections rather than on expanding participation into less familiar domains.

Implications

Our findings suggest that commercially available smartphones can serve as meaningful tools for older adults with dementia to maintain social connections and engage in valued social activities. However, the predominant use of basic communication and information functions indicates a missed opportunity to leverage the full potential of smartphone features that can promote independence and social participation. Given the promise of EICT training interventions [21], tailored strategies are needed to help individuals with dementia to use smartphone functions in personally meaningful ways, focusing on preferred activities and offering training in practical tools such as calendars, reminders, and digital payment systems that can directly support these activities.

Strengths and Limitations

This study has several strengths. First, multisite recruitment reflecting collaborative dementia care systems involving medical centers for dementia and primary care clinics enhanced sample diversity, which improved the robustness of the findings across different care settings. Participants recruited from hospitals and primary care clinics differed in sociodemographic and functional characteristics. Second, the inclusion of both patient- and caregiver-reported data allowed for a more comprehensive understanding of smartphone use and social participation, particularly in cases where cognitive limitations might have affected self-reporting. Third, the item-level analysis of social activities provided more nuanced insights into how smartphone use relates to specific aspects of everyday life.

However, several limitations should also be noted. The cross-sectional design prevents any conclusions about causality between smartphone use and social activity. Additionally, reliance on caregiver-reported outcomes may introduce reporting bias, particularly in cases where caregivers had limited knowledge of the participants' daily routines. Finally, the findings may not be generalizable to individuals with more advanced dementia or those not receiving regular medical care, as this study focused on community-dwelling older adults with mild-to-moderate dementia who were recruited through outpatient dementia services.

Future Research Directions

In the future, researchers should examine how spontaneous smartphone use changes over the course of cognitive decline and whether sustained use contributes to better social and psychosocial outcomes. Longitudinal studies are required to clarify these relationships. Moreover, efforts should be focused on identifying effective ways to support and expand the use of commercially available smartphones among older adults with dementia, particularly to enhance their participation in meaningful social activities. Given the limited use of cognitive aid and out-of-home activity features observed in our study,

XSL•FO
RenderX

targeted training or supportive environments would reduce barriers and facilitate more effective use of these functions. Collaborating with individuals with dementia in the design of such training programs and environments may further enhance their relevance and effectiveness.

Conclusions

Our findings highlight that a substantial proportion of community-dwelling older adults with mild to moderate dementia continue to use smartphones primarily for basic communication and information access. While smartphone use was associated with greater telephonic contact and overall social activity, apps and functions that could support independence and social participation were underused. These findings emphasize the potential of commercially available smartphones as meaningful tools for maintaining social engagement in dementia care. Targeted support may help bridge the gap between current usage and the broader capabilities of these devices, enabling individuals with dementia to sustain valued routines and social participation in daily life.

Acknowledgments

We sincerely thank the doctors at the participating medical facilities for their valuable support in data collection and their contributions to the successful completion of this study.

Funding

This work was supported by JSPS KAKENHI Grant Number 22K17531, and the Health Labour Sciences Research Grant by the Japanese Ministry of Health, Labour and Welfare (No. 22GB1003).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey items on smartphone apps and functions.
[DOCX File, 30 KB - aging_v9i1e81927_app1.docx]

Multimedia Appendix 2
Characteristics of participants recruited from hospitals and primary care clinics.
[DOCX File, 30 KB - aging_v9i1e81927_app2.docx]

References

1. Nichols E, Steinmetz JD, Vollset SE. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
2. Gauthier S, Webster C, Servaes S, Morais JA, Rosa-Neto P. World Alzheimer report 2022: life after diagnosis: navigating treatment, care and support. : Alzheimer's Disease International; 2022 URL: https://www.alzint.org/u/World-Alzheimer-Report-2022.pdf [accessed 2026-03-26]
3. Clare L, Gamble LD, Martyr A, et al. Longitudinal trajectories of quality of life among people with mild-to-moderate dementia: a latent growth model approach with IDEAL cohort study data. J Gerontol B Psychol Sci Soc Sci 2022 Jun 1;77(6):1037-1050. [doi: 10.1093/geronb/gbac022] [Medline: 35134935]
4. Wolverson EL, Clarke C, Moniz-Cook ED. Living positively with dementia: a systematic review and synthesis of the qualitative literature. Aging Ment Health 2016 Jul;20(7):676-699. [doi: 10.1080/13607863.2015.1052777] [Medline: 26078084]
5. Martyr A, Nelis SM, Quinn C, et al. Living well with dementia: a systematic review and correlational meta-analysis of factors associated with quality of life, well-being and life satisfaction in people with dementia. Psychol Med 2018 Oct;48(13):2130-2139. [doi: 10.1017/S0033291718000405] [Medline: 29734962]
6. Dröes RM, Chattat R, Diaz A, et al. Social health and dementia: a European consensus on the operationalization of the concept and directions for research and practice. Aging Ment Health 2017 Jan;21(1):4-17. [doi: 10.1080/13607863.2016.1254596] [Medline: 27869503]
7. Hackett RA, Steptoe A, Cadar D, Fancourt D. Social engagement before and after dementia diagnosis in the English longitudinal study of ageing. PLoS ONE 2019;14(8):e0220195. [doi: 10.1371/journal.pone.0220195] [Medline: 31369590]
8. Bohn L, Kwong See ST, Fung HH. Preference for emotionally meaningful activity in Alzheimer's disease. Aging Ment Health 2019 Nov;23(11):1578-1585. [doi: 10.1080/13607863.2018.1506750] [Medline: 30588831]

XSL•FO
RenderX

9. Evans J, Brown M, Coughlan T, Lawson G, Craven MP. A systematic review of dementia focused assistive technology. In: Kurosu M, editor. Human-Computer Interaction: Interaction Technologies: Springer; 2015:406-417. [doi: 10.1007/978-3-319-20916-6_38]
10. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
11. Schepens Niemiec SL, Lee E, Saunders R, Wagas R, Wu S. Technology for activity participation in older people with mild cognitive impairment or dementia: expert perspectives and a scoping review. Disabil Rehabil Assist Technol 2023 Nov;18(8):1555-1576. [doi: 10.1080/17483107.2022.2116114] [Medline: 36067094]
12. Lancioni G, Desideri L, Singh N, O'Reilly M, Sigafoos J. Technology options to help people with dementia or acquired cognitive impairment perform multistep daily tasks: a scoping review. JET 2021 Sep 8;15(3):208-223. [doi: 10.1108/JET-11-2020-0048]
13. Van der Roest HG, Wenborn J, Pastink C, Dröes RM, Orrell M. Assistive technology for memory support in dementia. Cochrane Database Syst Rev 2017 Jun 11;6(6):CD009627. [doi: 10.1002/14651858.CD009627.pub2] [Medline: 28602027]
14. Thorpe JR, Rønn-Andersen KVH, Bień P, Özkil AG, Forchhammer BH, Maier AM. Pervasive assistive technology for people with dementia: a UCD case. Healthc Technol Lett 2016 Dec;3(4):297-302. [doi: 10.1049/htl.2016.0057] [Medline: 28008366]
15. Ye B, Chu CH, Bayat S, Babineau J, How TV, Mihailidis A. Researched apps used in dementia care for people living with dementia and their informal caregivers: systematic review on app features, security, and usability. J Med Internet Res 2023 Oct 12;25:e46188. [doi: 10.2196/46188] [Medline: 37824187]
16. Pew Research Center. 3 internet, smartphone, and social media use. : Pew Research Center; 2022 URL: https://www.pewresearch.org/wp-content/uploads/sites/20/2022/12/PG_2022.12.06_Online-Civic-Engagement_REPORT.pdf [accessed 2026-03-26]
17. Public opinion poll on the use of telecommunications equipment. The Cabinet Office. 2023. URL: https://survey.gov-online.go.jp/hutai/r05/r05-it_kiki [accessed 2026-03-26]
18. Wu YH, Lewis M, Rigaud AS. Cognitive function and digital device use in older adults attending a memory clinic. Gerontol Geriatr Med 2019;5:2333721419844886. [doi: 10.1177/2333721419844886] [Medline: 31080848]
19. Wallcook S, Nygård L, Kottorp A, Malinowsky C. The use of everyday information communication technologies in the lives of older adults living with and without dementia in Sweden. Assist Technol 2021 Nov 2;33(6):333-340. [doi: 10.1080/10400435.2019.1644685] [Medline: 31343958]
20. Wilson SA, Byrne P, Rodgers SE. "I'd be lost without my smartphone": a qualitative analysis of the use of smartphones and tablets by people living with dementia, mild cognitive impairment, and their caregivers. Aging Ment Health 2024 Apr;28(4):595-603. [doi: 10.1080/13607863.2023.2205585] [Medline: 37139944]
21. Kerkhof YJF, Bergsma A, Mangiaracina F, Planting CHM, Graff MJL, Dröes RM. Are people with mild dementia able to (re)learn how to use technology? A literature review. Int Psychogeriatr 2022 Feb;34(2):113-128. [doi: 10.1017/S1041610221000016] [Medline: 33715653]
22. Kwan RYC, Cheung DSK, Kor PPK. The use of smartphones for wayfinding by people with mild dementia. Dementia (London) 2020 Apr;19(3):721-735. [doi: 10.1177/1471301218785461] [Medline: 29973063]
23. Ten-year plan to understand dementia and build community networks. Ministry of Health, Labour and Welfare. 2004. URL: https://www.mhlw.go.jp/topics/kaigo/dementia/c01.html [accessed 2026-03-26]
24. Ministry of Health, Labour and Welfare. Basic plan for the promotion of policies on dementia. 2024 URL: https://www.mhlw.go.jp/content/12300000/001347978.pdf [accessed 2026-04-08]
25. Reisberg B. Functional assessment staging (FAST). Psychopharmacol Bull 1988;24(4):653-659. [Medline: 3249767]
26. White paper information and communications in japan. : Ministry of Internal Affairs and Communications; 2024 URL: https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2024/2024-index.html [accessed 2026-04-08]
27. Neal DP, Kuiper L, Pistone D, et al. FindMyApps eHealth intervention improves quality, not quantity, of home tablet use by people with dementia. Front Med (Lausanne) 2023;10:1152077. [doi: 10.3389/fmed.2023.1152077] [Medline: 37324141]
28. Toyoshima K, Araki A, Tamura Y, et al. Development of the dementia assessment sheet for community-based integrated care system 8-items, a short version of the dementia assessment sheet for community-based integrated care system 21-items, for the assessment of cognitive and daily functions. Geriatr Gerontol Int 2018 Oct;18(10):1458-1462. [doi: 10.1111/ggi.13512] [Medline: 30225857]
29. Ito K, Ogisawa F, Furuta K, Awata S, Toba K. Development of a five-item short-form version of the Dementia Behavior Disturbance Scale. Geriatr Gerontol Int 2021 Sep;21(9):870-871. [doi: 10.1111/ggi.14234] [Medline: 34263527]
30. Umeda S, Kanemoto H, Suzuki M, et al. Validation of the Japanese version of the Social Functioning in Dementia scale and COVID-19 pandemic's impact on social function in mild cognitive impairment and mild dementia. Int Psychogeriatr 2024 Dec;36(12):1205-1218. [doi: 10.1017/S1041610224000401] [Medline: 38462968]
31. Budgett J, Brown A, Daley S, et al. The social functioning in dementia scale (SF-DEM): exploratory factor analysis and psychometric properties in mild, moderate, and severe dementia. Alzheimers Dement (Amst) 2019 Dec;11(1):45-52. [doi: 10.1016/j.dadm.2018.11.001] [Medline: 30623018]

XSL•FO
RenderX

32. Levasseur M, Richard L, Gauvin L, Raymond E. Inventory and analysis of definitions of social participation found in the aging literature: proposed taxonomy of social activities. Soc Sci Med 2010 Dec;71(12):2141-2149. [doi: 10.1016/j.socscimed.2010.09.041] [Medline: 21044812]
33. Sommerlad A, Singleton D, Jones R, Banerjee S, Livingston G. Development of an instrument to assess social functioning in dementia: the Social Functioning in Dementia scale (SF-DEM). Alzheimers Dement (Amst) 2017;7:88-98. [doi: 10.1016/j.dadm.2017.02.001] [Medline: 28317009]
34. Conway A, Ryan A, Harkin D, Mc Cauley C, Goode D. A review of the factors influencing adoption of digital health applications for people living with dementia. Digit Health 2023;9:20552076231162985. [doi: 10.1177/20552076231162985] [Medline: 36937696]
35. Dixon E, Michaels R, Xiao X, et al. Mobile phone use by people with mild to moderate dementia: uncovering challenges and identifying opportunities: mobile phone use by people with mild to moderate dementia. ASSETS 2022 Oct;2022:38. [doi: 10.1145/3517428.3544809] [Medline: 37283282]
36. Benge JF, Dinh KL, Logue E, Phenis R, Dasse MN, Scullin MK. The smartphone in the memory clinic: a study of patient and care partner's utilisation habits. Neuropsychol Rehabil 2020 Jan;30(1):101-115. [doi: 10.1080/09602011.2018.1459307] [Medline: 29661059]
37. Portacolone E, Cohen AB. Living alone with dementia: a reality check. Am J Geriatr Psychiatry 2024 Nov;32(11):1322-1324. [doi: 10.1016/j.jagp.2024.07.006] [Medline: 39181836]
38. Clare L, Martyr A, Henderson C, et al. Living alone with mild-to-moderate dementia: findings from the IDEAL cohort. J Alzheimers Dis 2020;78(3):1207-1216. [doi: 10.3233/JAD-200638] [Medline: 33104029]
39. Talbot CV, Briggs P. The use of digital technologies by people with mild-to-moderate dementia during the COVID-19 pandemic: a positive technology perspective. Dementia (London) 2022 May;21(4):1363-1380. [doi: 10.1177/14713012221079477] [Medline: 35333111]
40. Thalén L, Malinowsky C, Margot-Cattin I, et al. Out-of-home participation among people living with dementia: a study in four countries. Dementia (London) 2022 Jul;21(5):1636-1652. [doi: 10.1177/14713012221084173] [Medline: 35435030]

Abbreviations

BPSD: Behavioral and psychological symptoms of dementia
DASC-8: Dementia Assessment Sheet for Community-based Integrated Care System 8-items
DBD-5: Dementia Behavior Disturbance scale
EICT: Everyday Information and Communication Technology
MMSE: Mini-Mental State Examination
OR: odds ratio
SF-DEM-J: Japanese Social Functioning in Dementia scale

Edited by C Lee; submitted 05.Aug.2025; peer-reviewed by E Dixon, G Feher; revised version received 20.Jan.2026; accepted 06.Mar.2026; published 24.Apr.2026.

Please cite as:
Tsuda S, Hotta S, Matsumoto H, Ito K
Smartphone Use and Social Activities Among People With Mild to Moderate Dementia: Multi-Informant Cross-Sectional Study
JMIR Aging 2026;9:e81927
URL: https://aging.jmir.org/2026/1/e81927
doi:10.2196/81927

© Shuji Tsuda, Satoko Hotta, Hiroshige Matsumoto, Kae Ito. Originally published in JMIR Aging (https://aging.jmir.org), 24.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smart Glasses for Older Adults With Cognitive Impairment: Explanatory Mixed Methods Study

Brittany F Burch[1], RN, MSN, PhD; Nayeon Kim[1], RN; Rachel McPherson[1], PhD; Debora Anokye[1], RN; Alice S Ryan[2,3], PhD; Odessa Addison[2,3], DPT, PhD; Nirmalya Roy[4], PhD; Chien-Ming Huang[5], PhD; Elizabeth Galik[1], CRNP, PhD; Barbara Resnick[1], CRNP, PhD

[1]School of Nursing, University of Maryland Baltimore, 655 W Lombard St, Baltimore, MD, United States

[2]School of Medicine, University of Maryland, Baltimore, MD, United States

[3]Baltimore Geriatric Research, Education, and Clinical Center, VA Maryland Health Care System, Baltimore, MD, United States

[4]Information Systems, University of Maryland, Baltimore, MD, United States

[5]Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Brittany F Burch, RN, MSN, PhD
School of Nursing, University of Maryland Baltimore, 655 W Lombard St, Baltimore, MD, United States

Abstract

Background: Smart glasses might present a promising solution to support older adults with cognitive impairment in maintaining independence. However, there exists a critical gap in smart glasses research that incorporates recently developed models or directly engages older adults with cognitive impairment.

Objective: This study aimed to use survey and interview techniques to explore the acceptability and usability of smart glasses among older adults with cognitive impairment.

Methods: This explanatory mixed methods descriptive study was conducted at an independent living older adult apartment building among residents with memory complaints or cognitive impairment. This study consisted of a quantitative survey (N=26), followed by smart glasses beta testing with qualitative interviews (n=14).

Results: Overall, older participants with cognitive impairment conveyed a generally positive perception of smart glasses and their potential to support memory in daily life. Results suggest that participants prioritized the following smart glasses functions: audio reminders, phone calls, GPS, and distress signals, with audio reminders emerging as the highest-ranked feature. Additionally, participants emphasized the value of an intuitive and quick interface and a preference for audio, rather than visual, information exchange.

Conclusions: This study supports the development and study of smart glasses for older adults with mild cognitive impairment. Smart glasses developers should place great importance on usability, such as intuitive command interfaces, and equip the smart glasses with functions that are relevant to this population, especially audio reminders. Additionally, future work should examine the integration of smart glasses over a longer period and among a larger sample of older adults with cognitive impairment.

(JMIR Aging 2026;9:e81840) doi:10.2196/81840

KEYWORDS

technology; cognitive impairment; smart glasses; dementia; older adults

Introduction

Older Adults With Cognitive Impairment: Prevalence and Burden

Cognitive impairment is a condition associated with a decline in one or more cognitive functions such as memory, attention, reasoning, language, problem-solving, and decision-making [1-3]. The severity of cognitive impairment can range from mild cognitive impairment to more advanced conditions, including severe dementia. In mild cognitive impairment, individuals might struggle with complex tasks such as managing medications and making phone calls, but more basic self-care tasks remain largely preserved [3,4]. The global prevalence of cognitive impairment among community-dwelling older adults is approximately 19%, and this prevalence is expected to rise as the population ages [1,2]. Among older adults who are economically disadvantaged, the onset of cognitive impairment occurs earlier, and the prevalence of cognitive impairment is higher [5,6].

Cognitive impairment is often associated with depression and limitations in activities of daily living, significantly affecting an individual's quality of life [7]. Caregivers of individuals with

XSL•FO
RenderX

cognitive impairment may also experience emotional, physical, or financial strain due to the sequelae of symptoms associated with cognitive impairment [8,9]. While some research indicates that cognitive training, physical activity, and pharmacological interventions can delay the progression of cognitive impairment, no cure currently exists [10-14]. Consequently, innovative solutions are needed to help individuals maintain independence as they live with their cognitive impairment.

Smart Devices to Assist Older Adults With Cognitive Impairment

Smart devices are defined as context-aware electronic tools that are capable of performing autonomous computing and connecting for data exchange [15]. They present innovative solutions to support older adults with cognitive impairment in maintaining independence. For example, researchers have explored smartphones, smart wristwatches, and tablet-based applications to assist with cognitive training or function as an external memory aid [16-19]. Similarly, researchers have investigated virtual assistant smart speakers, such as Google Home or Amazon Alexa, to provide in-home support to older adults living in the community with cognitive impairment [20,21]. While such technologies demonstrate promise in supporting older adults with cognitive impairment in living independently, common acceptability and usability barriers include nonintuitive interfaces, a lack of personalization and accessibility, and technology literacy limitations among older adults [16,20,21]. Over the past decade, a new type of smart device—smart glasses—has emerged as a potential tool to assist individuals with memory-related challenges [22]. Unlike other technological interventions, glasses are a familiar and commonly worn accessory among older adults, potentially making smart glasses a more natural and accepted aid.

The History of Smart Glasses

For the purposes of this paper, smart glasses are defined as an eye-worn device that provides functions similar to a computer to assist and augment the wearer [23]. The release of the Google Glass Explorer Edition from 2013 to 2014 marked the first consumer-grade smart glasses available to the public. However, issues such as unattractive designs, privacy concerns, and high cost led to its limited adoption, and the product was discontinued in 2015 for purchase by general consumers [24-26]. In 2017 and 2019, Google released updated versions, Google Glass Enterprise Edition and Google Glass Enterprise Edition 2, marketed more specifically to businesses; these were discontinued in 2023 [27]. Around this time, the first models of mixed reality headsets were also released, such as the Microsoft HoloLens (2016) [28] and the Magic Leap One (2018) [29]. While at the forefront of innovation in fields such as health care, manufacturing, education, and defense, these devices were still largely considered niche industry products and not marketed to or adopted by general consumers [28-31]. Additionally, their bulky design and lack of portability raise questions about whether the devices should be classified as “smart glasses”—in addition to a headset [22].

In the last 5 years, advancements in features, esthetic, and affordability have led to the growing integration of smart glasses into everyday life, especially among younger individuals [32]. Current models incorporate advanced technologies, including cameras, microphones, GPS, and accelerometers, while maintaining the appearance of regular eyewear [22]. These features enable functionalities such as voice assistance, navigation, reminders, photo live or live stream, and social connection, provided in a more discreet manner. Of note, some commercial smart glasses are becoming powered with advanced artificial intelligence (AI; eg, Ray-Ban Meta) and augmented reality technology (eg, Vuzix V100), defined as digital information overlaid onto the real environment [33]. Given the recent technological and esthetic improvements of commercial smart glasses, alongside the increasing population burden of cognitive impairment, further research is necessary to explore their acceptability and usability among older adults.

Smart Glasses Research and Gap

To contextualize this topic, our research team recently conducted a systematic scoping review on smart glasses for older adults with cognitive impairment [22]. Among the 13 studies identified, only 5 [34-38] involved older adults in hands-on testing of the smart glasses, and only 1 study [37] exclusively recruited older adults living with cognitive impairment, though it was limited by a small sample size (n=6). Additionally, all included studies in the review focused on earlier generations of commercial smart glasses, such as Google Glass and HoloLens, or early-stage rudimentary prototypes (ie, regular glasses with components added to the side arm of the glasses frames). Our review underscores a critical gap in research that incorporates recently developed smart glasses models and directly engages older adults with cognitive impairment. To address this gap, this study aimed to use survey and interview techniques to explore the acceptability and usability of smart glasses among older adults with cognitive impairment. For the purposes of this paper, and consistent with the Sekhon Theoretical Framework of Acceptability, we define acceptability as “the extent to which people delivering or receiving a health care intervention consider it to be appropriate, based on anticipated or experienced cognitive and emotional responses to the intervention” [39]. Next, consistent with the International Organization for Standardization (ISO 9241 - 11), we define usability as “the extent to which a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [40]. The results from this study will guide researchers in selecting appropriate smart glasses and guide engineers in developing smart glasses that are more accessible and relevant for this population across both research studies and real-world settings.

Methods

Study Design

This was an explanatory mixed methods descriptive study to explore the acceptability and usability of smart glasses among older adults with cognitive impairment and consisted of two phases: (1) a quantitative survey and (2) smart glasses beta testing with qualitative interviews.

XSL•FO
RenderX

Participants

Given the high prevalence and burden of cognitive impairment among economically disadvantaged older adults [5], we recruited participants from an income-controlled independent living older adult apartment building in Baltimore. We recruited participants through flyers posted in the apartment building and a brief on-site presentation about the project. Residents were eligible if they met the following criteria: (1) scored 4 or lower on the Mini-Cog [41] or 3 or higher on the Memory Complaint Scale [42], and (2) passed the Evaluation to Sign Consent Measure [43], which assessed whether participants understood the study purpose, risks, and withdrawal rights. As we specifically invited residents who perceived themselves as having memory problems or forgetfulness, nearly all residents (26/27) who expressed interest in the study were eligible based on the cognitive screening criteria. For phase 1 (surveys), we enrolled all interested residents who met the eligibility criteria. For phase 2 (glasses testing and interviews), we randomly selected phase 1 participants for interviews until we achieved data saturation, defined as redundancy of data with no newly emerged themes [44].

Ethical Considerations

The University of Maryland, Baltimore, Institutional Review Board approved the protocol for this study (HP-00109956). We informed participants of the study procedures, risks, benefits, and withdrawal rights prior to participation. We protected confidentiality by assigning unique study IDs to participants and storing all identifiable data on secure, password-protected servers and in a locked office within a locked cabinet. Participants provided informed consent prior to participation. Participants received a US $30 incentive for completing the survey and interview, respectively.

Procedures

During phase 1, we administered a paper-based quantitative survey, which we read aloud and completed for participants upon request. Once phase 1 was completed for all participants, we purchased 3 different pairs of smart glasses that best aligned with participants' preferences and priorities, as informed by the survey feedback: Ray-Ban Meta smart glasses, Alexa Echo Frames, and Vuzix Blade 2 smart glasses. The 3 pairs of selected smart glasses have many similarities, such as including a speaker and a microphone. One primary distinction is that the Alexa Echo Frames do not have a camera, whereas the other 2 devices do. Additionally, Vuzix Blade 2 smart glasses provide augmented reality functionality, which is not available in the other 2 smart glasses.

During phase 2, a random sample of participants from phase 1 beta-tested these 3 pairs of smart glasses and provided qualitative interview feedback. During beta testing, each participant completed a task list for each pair of glasses, guided by the research assistant, as detailed in Textbox 1. Glasses testing took approximately 20 minutes per pair, for a total of 60 minutes. Participants tested the glasses pairs in a rotating order to minimize order effects. For example, participant 1 tested Ray-Ban Meta smart glasses first, Alexa Echo Frames second, and Vuzix Blade 2 smart glasses third. Alternatively, participant 2 tested Alexa Echo Frames first, Vuzix Blade 2 smart glasses second, and Ray-Ban Meta smart glasses third.

XSL•FO
RenderX

Textbox 1. Participants' beta-testing task list.

Ray-Ban Meta smart glasses

- "Hey Meta, how is the weather outside?"
- "Hey Meta, remember for my shopping list I'll need eggs, milk, and bread."
- "Hey Meta, play [insert favorite music]."
- "Hey Meta, remember that I put my purse/wallet by the door."
- "Hey Meta, remember this book."
- "Hey Meta, remind me in 2 minutes to take my medications."
- "Hey Meta, where is the closest Safeway to me?"
- "Hey Meta, what street am I on?"
- "Hey Meta, look and tell me what this says in English?"
- "Hey Meta, what items are on my shopping list?"
- "Hey Meta, look at this snack and tell me if it's healthy."
- "Hey Meta, look at this cooking tool and tell me what it is called."
- "Hey Meta, look at this recipe and summarize it. Hey Meta, remind me what I do after microwaving the butter."
- "Hey Meta, take a photo now and send it to [research assistant]."
- "Hey Meta, call [research assistant]."
- "Hey Meta, take a picture of this flyer so I can refer to it later."
- "Hey Meta, remind me where I put my purse today. What was the book I wanted to remember?"
- "Hey Meta, play 'Mindfulness for Beginners' on the Calm app."

Alexa Echo Frames

- "Hey Alexa, how is the weather outside?"
- "Hey Alexa, play [insert favorite music]."
- "Hey Alexa, remind me in 2 minutes to take my medications."
- "Hey Alexa, where is the closest Safeway to me?"
- "Hey Alexa, what is the metal tool you use to beat eggs?"
- "Hey Alexa, what items are on my shopping list?"
- "Hey Alexa, are raisins healthy?"
- "Hey Alexa, call [research assistant]."
- "Hey Alexa, what's on my 'to do' list?"
- Participant demonstrates the "find my" function within app (GPS tracking).
- "Hey Alexa, turn off lights."

Vuzix Blade 2 smart glasses

- Scroll through apps with voice: "Hey Vuzix. Move right. Move left."
- Scroll through apps with fingers.
- Select notifications with fingers.
- Select notifications with voice: "Hey Vuzix, notifications."
- Determine preferred method (audio or voice).
- Take a picture with preferred method.
- Take a video with preferred method.
- Navigate to Microsoft Teams and check calendar with preferred method.
- Make a call with preferred method.

XSL•FO
RenderX

Measures

Phase 1

The phase 1 quantitative self-report survey consisted of three sections, which assessed (1) demographics, including age, gender, living alone status, comorbidities, and financial strain; (2) frequency of technology use; and (3) preferred smart glasses functions. To assess the frequency of technology usage, we administered the FACETS (Functional Assessment of Currently Employed Technology Scale) [45], a validated and reliable 10-item questionnaire that measures the frequency of technology use across the following areas: sending emails, accessing files, sending text messages, posting on social media, online banking, online purchases, telehealth, communicating with health insurance online, installing computer components, and resetting a modem or router [45]. We removed the last 2 items, which were less pertinent to those who live in the congregate affordable housing setting, many of whom primarily use smartphones and public building Wi-Fi. Item responses ranged from 0 ("never") to 5 ("daily/prefer"). Thus, total scores ranged from 0 to 40, with higher scores indicating greater technology usage.

As this work is exploratory, no validated survey currently exists that examines the preferred functions of smart glasses. Instead, we developed the "preferred functions" items based on our clinical expertise with this population and dementia-friendly survey guidelines from the Alzheimer's Society [46]. This section contained 17 Likert-style questions in which participants rated smart glasses' potential functions as "not helpful," "slightly helpful," "very helpful," and "most helpful." Participants then reported the 3 "best" smart glasses functions from all listed functions. The survey is provided in Multimedia Appendix 1.

Phase 2

Immediately following the testing of each of the three pairs of glasses, the participants answered five open-ended questions: (1) What did you like best about these glasses? (2) What did you like least about these glasses? (3) What function did you like best for these glasses? (4) What function did you like least for these glasses? (5) Is there anything else you would like to share about these glasses? Upon completing beta testing for all 3 pairs, participants responded to five additional overarching questions: (1) What did you think of the smart glasses overall? (2) Describe the glasses you liked the best and why. (3) Describe the smart glasses functions you liked the best and why. (4) Are there any functions you wish smart glasses could have that these did not? (5) Why or why not do you think smart glasses could be used to help with memory? We intended the questions about participants' likes, dislikes, and overall impressions to capture acceptability, while the questions about specific functions, ease of use, and desired features were intended to capture usability. To maintain consistency, a single interviewer (BFB) with qualitative interviewing expertise conducted all interviews. The interviewer provided assistance as needed, while avoiding leading questions or comments. The interviewer transcribed the feedback verbatim in real time.

Analysis

After collecting phase 1 quantitative surveys, we conducted descriptive statistics and applied data visualization approaches to summarize the quantitative results. Specifically, we obtained mean FACETS results and ranked participants' preferred smart glasses functions based on how frequently each was identified as "the best function" or placed among the top three functions.

One study researcher (BFB) conducted a content analysis to organize and interpret the qualitative data. This process consisted of (1) codebook creation, (2) familiarization with the data, (3) data organization, and (4) data interpretation. First, BFB created the codebook in a deductive manner, based on the topics and questions in the interview guide, reflecting the study aims. This codebook contained the codename, when the code should be used, when the code should not be used, and an example corresponding quotation. Second, and consistent with content analysis, BFB engaged in data familiarization by reading each interview twice. Third, BFB used MaxQDA (VERBI Software), a qualitative data management software, to organize and match segments of text with codes that aligned with them. This coding process was iterative in nature. After coding the first 2 interviews, BFB refined the codebook to capture the advantages and disadvantages of each pair of smart glasses with greater granularity. The finalized code names, subcode names, and their frequency of use are listed in Multimedia Appendix 2. Finally, following the methods outlined by Merriam and Tisdell [47], codes were interpreted so that those with similar concepts were merged into categories, and categories were further integrated into themes.

Once the themes were identified and summarized, a second researcher (NK) independently familiarized herself with the qualitative interview data for the purpose of peer-debriefing. During peer debriefing, BFB and NK met to engage in reflexive dialogue, assess for any themes not expressed within the results, consider alternative interpretations of the data (rival explanations), or identify data that did not align with our current conclusions (negative cases). This collaborative process was conducted to enhance the trustworthiness and credibility of our results. While the peer debriefing did not result in the addition of any new themes, it served as the catalyst for a more detailed explanation of participants' preferred smart glasses functions.

The research team then met to compare survey data results with interview findings, identifying convergent, complementary, and divergent conclusions.

Following triangulation, the research team returned to the older adult apartment building to present results and conduct member checking with available participants (n=18). During this session, participants were asked, as a group, to provide feedback on (1) conclusions that reflected their experience and perspective, (2) conclusions that misrepresented their experience or perspective, and (3) any additional insights that were missing or they wanted to share. In response, participants conveyed that the major conclusions drawn by the research team resonated with their experience and perspective. Of note, 2 small changes were made due to the member checking session; namely, language translation was removed from the list of functions described as least helpful, and facial recognition was added within the "Recommendations for Functions Smart Glasses Should Do" section.

XSL•FO
RenderX

Results

Overview

Among the 26 older adult participants, 16 (61.54%) identified as female, 18 (69.23%) identified as Black or African American, and 14 (53.85%) reported some level of financial strain (Table 1). The participants scored an average of 3.50 (SD 1.42) on the Mini-Cog and 7.62 (SD 3.40) on the Memory Complaint Scale, suggesting that participants overall had an average of mild cognitive impairment, although there was a range of cognition from mild to severe cognitive impairment and memory complaints. Participants scored a mean of 27.88 (SD 13.39) on the modified FACETS, demonstrating moderate technology use across most domains [45]. Participants' frequency of technology use was highest for online banking and SMS text messaging and lowest for social media use and accessing files.

Table . Participant characteristics (N=26).

Characteristics	Values
Comorbidities, mean (SD)	2.3 (2.06)
Mini-Cog, mean (SD)	3.50 (1.42)
Memory Complaint Scale, mean (SD)	7.62 (3.40)
FACETS[a], mean (SD)	27.88 (13.39)
Send email	3.35 (2.06)
Find, open, and close files	2.92 (2.19)
SMS text messaging	4.31 (1.98)
Post on social media	2.11 (1.93)
Manage banking online	4.27 (2.13)
Pay bills and make purchases online	4.15 (2.31)
Communicate with doctor online	3.62 (2.26)
Communicate with health insurance company online	3.15 (2.33)
Gender, n (%)	
Female	16 (61.54)
Male	10 (38.46)
Race or ethnicity, n (%)	
Black or African American	18 (69.23)
White	8 (30.77)
Living alone status, n (%)	
Living alone	24 (92.31)
Living with others	2 (7.69)
Age (y), n (%)	
60 - 64	4 (15.38)
65 - 69	5 (19.23)
70 - 74	6 (23.08)
75 - 79	7 (26.92)
≥80	4 (15.38)

[a]FACETS: Functional Assessment of Currently Employed Technology Scale.

Quantitative Survey Results

All 26 participants completed the survey with no missing data. Consistently, participants identified audio reminders as the "best" smart glasses function. Specifically, 8 (30.77%) participants ranked audio reminders as the "best" smart glasses function, and 14 (53.85%) participants ranked audio reminders in the top 3 smart glasses functions. Following audio reminders, the following functions were the most valued by participants: phone calls, GPS, so that their loved ones can locate them, and sending distress signals. Conversely, the following functions were reported as least helpful by participants: playing music, cognitive training, subtitles during conversation, companionship and conversation, and providing video or photos for family. All participants (100%) stated that they would be willing to wear smart glasses that assist with memory; however, 6 (23.08%) reported that the glasses *must* look and feel like normal glasses to warrant wear.

XSL•FO
RenderX

Qualitative Interview Results

Overview

A random sample of 14 participants beta-tested 3 pairs of smart glasses: Ray-Ban Meta smart glasses, Alexa Echo Frames, and Vuzix Blade 2 smart glasses. We selected these 3 devices because of their capacity to provide reminders, make calls, and provide GPS functionality, consistent with preferences from the quantitative survey results. After beta testing the 3 pairs of smart glasses, the participants provided qualitative feedback. This feedback is described below, consisting of six themes: (1) perceived advantages and disadvantages of Ray-Ban Meta smart glasses, (2) perceived advantages and disadvantages of Amazon Alexa Echo Frames, (3) perceived advantages and disadvantages of Vuzix Blade 2 smart glasses, (4) preferred smart glasses and functions, (5) broad smart glasses concerns, and (6) recommendations for functions smart glasses should do. Additional participant quotations reflecting the themes can be found in Multimedia Appendix 3.

Perceived Advantages and Disadvantages of Ray-Ban Meta Smart Glasses

Participants conveyed the most favorable feedback toward the Ray-Ban Meta smart glasses. These favorable views were due to the following: (1) an intuitive and quick interface, (2) functions that they perceived were relevant to them, and (3) comfort and esthetics. First and foremost, participants overwhelmingly appreciated that the Ray-Ban Meta smart glasses were simple to use and responded clearly, quickly, and precisely to their commands. Participant 4 described:

> *They were just real simple. Not a lot involved just speaking the command and it did what it did.*

Second, participants emphasized that the functions provided by the Ray-Ban Meta smart glasses were helpful and relevant to their lives. Features that the participants highlighted as advantages of the Ray-Ban Meta smart glasses included reminder notifications and serving as a memory aid (eg, “Remember where I put my purse”), GPS navigation, phone call functionality, and the inclusion of a portable charging case. Additionally, participants appreciated how the integration of the camera allows the Ray-Ban Meta smart glasses to require less information input than the Amazon Echo Frames. For example, participant 7 described how they “…like how it can take a picture of the book without me saying the name of it…take a picture of the ingredients and it breaking them down.” Finally, participants appreciated the comfort and esthetics of the Ray-Ban Meta smart glasses. They described the glasses as lightweight, attractive, stylish, comfortable to wear, and fitting well. They highlighted the glasses’ resemblance to conventional eyewear.

Although participants generally perceived the Ray-Ban Meta smart glasses as responsive to user commands, compared to the other models tested, 3 participants noted that the interface could still benefit from improvements in accurately detecting user speech. In particular, 1 participant (participant 11), who was edentulous and identified as having an accent, reported needing to repeat the activation command “Hey Meta” multiple times before receiving a response. While the augmented reality of the Vuzix smart glasses was generally viewed unfavorably among the study participants as described below, 2 participants suggested that integrating augmented reality technology into the Ray-Ban Meta smart glasses could produce a pair of glasses with comprehensive functionality.

Perceived Advantages and Disadvantages of Amazon Alexa Echo Frames

The participants provided generally favorable feedback on the Amazon Alexa Echo Frames, citing various features that they would appreciate integrating into their daily lives. Many such features overlapped with those of the Ray-Ban Meta smart glasses. For example, participants highlighted the reminder notifications, the virtual assistant for answering questions, and the phone call functionality. Participant 2 stated:

> *I liked everything about these glasses. They are right up my alley. The [locating] grocery stores, weather, calling someone, [creating and referring to a] to-do list, locating people.*

Unique to the Amazon Alexa Echo Frames, some participants appreciated how the glasses could sync with other household appliances. In terms of comfort and esthetics, participants generally felt positive toward the lightweight and comfortable frames.

The participants described 3 major disadvantages of the Amazon Alexa Echo Frames. First, some participants emphasized the frequent need to repeat the activation command “Hey Alexa” due to inconsistent voice recognition. Participant 19 explained:

> *You have to pause after “Alexa” but then you have to speak fast for what you want. I wish these could do all the things the first one [Meta] could do. You have to repeat.*

Second, participants noted the absence of a camera as a limitation. In addition to the inability to capture photos and videos, the absence of camera restricts the glasses’ capability to perform object recognition or provide information about objects and the environment through visual analysis. Participant 23 stated:

> *I like these. But if they had a camera… I wish they had a camera so she could see a lot of things that I want to know.*

Third, participants were concerned that the glasses had a tendency to slide down their nose.

Perceived Advantages and Disadvantages of Vuzix Blade 2 Smart Glasses

While feedback on the Vuzix Blade 2 smart glasses was generally negative, some participants described the glasses as interesting and helpful for specific, but limited, situations. Participant 8 stated:

> *It was neat having the experience of seeing something inside your glasses [referring to the augmented reality technology]. But that is sort of a novelty thing.*

Participants envisioned the calendar application within Microsoft Teams as a potentially helpful tool for supporting their daily activities. The participants noted that the Vuzix glasses would

XSL•FO
RenderX

require significant training and might be more helpful for specific populations, such as individuals with high technological literacy who have hearing or visual impairment.

Among the 3 pairs of smart glasses tested, the Vuzix Blade 2 smart glasses were the least favorably received by participants. The participants expressed three main disadvantages: (1) operational complexity; (2) the distracting nature of the augmented reality interface; and (3) a bulky, unattractive design. First, participants described the Vuzix smart glasses as "complicated," "strenuous," and "hard to use." While the participants appreciated the choice between navigating the glasses through touch and audio input, they noted that the glasses did not respond easily to either input. The touch input was overly sensitive, leading participants to inadvertently navigate to incorrect apps. The touch input (eg, long 2-finger touch to return home) was also difficult for the participants to remember, especially among those with more significant cognitive impairment. On the other hand, while "…Voice commands are more intuitive for people with memory problems" (participant 1), the participants collectively noted that the Vuzix smart glasses had a restricted vocabulary and often failed to respond to their verbal prompts. Next, as the only pair of smart glasses to be equipped with augmented reality technology, participants found the projection on the lens distracting and confusing. Participant 9 stated:

> *And the clicking and flashbacks…it's overwhelming to me. It bothers my eyes. It breaks your concentration. Especially if you're sitting here talking. It's flashing and stuff and that bothers me. I'd forget what to do.*

Participants also referred to eye strain and instances of double vision. Finally, there was unanimous agreement that the glasses should be smaller and lighter to improve wearability and transportability. For example, Participant 1 stated:

> *I don't like their look… they are bulky. I mean, what do you put these in? I can't even put these in my purse. They wouldn't fit.*

Participants also complained that the glasses fit incorrectly on their faces, such as being heavy on the bridge of the nose, pushing on the head above the ears, sitting on the cheeks, or being tight around the temples.

Preferred Smart Glasses and Functions

Following beta testing of the 3 models, 11 participants stated that their preferred glasses were the Ray-Ban Meta smart glasses, 3 participants stated that their preferred glasses were the Amazon Alexa Echo Frames, and no participants stated that their preferred glasses were the Vuzix Blade 2 smart glasses. When asked why they preferred the Ray-Ban Meta smart glasses, the participants overwhelmingly described how the glasses were the "easiest to use." Participants also mentioned the comfort, esthetics, and camera function in selecting Ray-Ban Meta glasses as their preferred smart glasses. The few participants who preferred the Amazon Alexa Echo Frames smart glasses referred to the lightweight frame.

When asked about the most important functions that smart glasses can have, the participants mainly discussed two features: (1) reminders and (2) instant hands-free information. First, participants found that smart glasses reminders are critical, such as providing alerts to take medications, remembering grocery lists, and remembering where objects are placed. Participant 17 described the most important glasses feature to them:

> *For me, to remember things. What was I supposed to get at the grocery store? Where was I supposed to go today? What time was I supposed to get there? Just yesterday I got things messed up and I was supposed to have a procedure and I called just to be safe and they said "You don't have a procedure today." So something like this could help prevent these embarrassing moments for someone my age and with my condition.*

The second critical smart glasses function highlighted by participants was instant, hands-free information. Participants stated that smart glasses should provide "clear" and "instant" information through the virtual assistant. Additionally, participants appreciated when the information was delivered through audio feedback, rather than a projection on the lens (augmented reality). Participant 18 described:

> *First off you are not holding anything. That is what I like…hands free…to be able to speak out what you need...*

While not endorsed by any other participants, 1 participant (participant 9) emphatically relayed that smart glasses could significantly improve the lives of older adults through offering social companionship for those who experience social isolation. This participant described how older adults could use the glasses for conversation or calling a support group.

Broad Smart Glasses Concerns

The discussion related to the Ray-Ban Meta smart glasses and Vuzix Blade 2 smart glasses both elicited privacy concerns among 3 participants. Two participants emphasized the ethical responsibility of glasses users to be respectful of others' privacy when capturing photos or audio recordings (participants 8 and 23). Participant 3 advocated for more discreet smart glasses designs (compared to Vuzix glasses) to prevent public apprehension about surveillance or invasive behavior. This participant (participant 3) stated that smart glasses could be useful for people with memory problems but said, "You want to make it so, you know, you have the functions, but you don't want everyone knowing you're getting the help." They suggested that a more subtle design would allow for cognitive support without drawing the attention or discomfort of others.

Generally, participants overwhelmingly affirmed that smart glasses could assist with memory by storing information, stirring memory, offering reminders, recalling appointments, aiding with navigation, and recalling where items were placed. Alternatively, 1 among the 14 participants expressed a negative perception of smart glasses' use for memory within the context of a general overreliance on technology. While this participant (participant 11) believed that smart glasses could be "a potentially useful tool," they stated:

> *I am reluctant to let too much electronics in, myself. I enjoy keeping remembering things myself. Keeping*

XSL•FO
RenderX

> *my memory sharp. Whereas with the glasses, it'll tell me something I don't necessarily need to know. I'd rather remember in my head and exercise my memory.*

Recommendations for Functions Smart Glasses Should Do

After testing the 3 models of smart glasses, the participants provided advice regarding what additional functions smart glasses should provide for older adults who have memory problems. First, participants suggested that smart glasses should identify environmental hazards or dangerous situations, such as recognizing fall risks, dangerous traffic, car accidents, or smoke due to a fire. For example, participant 19 stated:

> *They did not point out danger. Like when you got to the corner: "Don't step down." "Wait till the light changes." Help you be more aware of your surroundings.*

Second, participants also stated that smart glasses should provide health assistance, such as step counts or a medical alert in the event of a medical emergency. Third, participants suggested that smart glasses should provide navigation assistance while shopping. For example, smart glasses could group items from a grocery list while navigating the aisles, help locate items in stores, and provide a price check for grocery items. Finally, participants suggested that smart glasses should provide facial recognition for acquaintances. They expressed how a facial recognition functionality would prevent embarrassing moments, such as forgetting the names of other residents in their building.

Triangulation

The integration of the quantitative and qualitative components of this study revealed convergent and complementary results, strengthening the credibility of the findings. Specifically, both survey and interview results suggest that participants prioritized the following smart glasses functions: audio reminders, phone calls, GPS, and distress signals, with audio reminders emerging as the highest-ranked feature (Table 2). Although participants performed other tasks, such as playing music and sending a photo to family or friends, both the quantitative and qualitative results suggest that these functions were perceived as less essential among older adults with memory problems. In addition to confirming functional priorities, complementary insights emerged from the qualitative interviews. First, following hands-on testing, participants emphasized the value of an intuitive and quick interface. Next, the participants highlighted that audio feedback from the glasses, rather than visual feedback, was the preferred method of information delivery. These complementary qualitative findings reveal that even though certain functions might be perceived as critical, the accessibility of the smart glasses interface is key to participant acceptance.

XSL•FO
RenderX

Table . Favored smart glasses functions.

Favored functions	Participants who selected the function as the "best" function, n (%)	Itemized survey (1="least helpful" to 4="most helpful"), mean (SD)	Example representative quote
Audio reminders	8 (30.77)	3.50 (0.71)	• "You know, unless you write it down, you forget it. It can remind you. It's instant information." [Participant 3] • "Whereas with the others (Meta Ray-Ban smart glasses and Alexa Echo Frames, compared to Vuzix Blade 2), they could function 24/7, you just communicate with it. You don't have the visual to distract you. I'm one of those people, I can't multitask, when I'm on my scooter, I have to watch for curbs and people." [Participant 11]
Phone calls	5 (19.23)	3.38 (0.85)	• "Making a call is handy." [Participant 8] • "A lot of the people in this building have family who don't come see them. These would be beneficial for them. They could sit in their apartment. If they feel lonely or depressed, they could tell the glasses to call a support group or something and it would make the world of difference." [Participant 9]
Distress signals	3 (11.54)	3.54 (0.71)	• "They are going in the right direction with them (smart glasses). If they could do a general programming for them and then customize them to the client's needs, it would be wonderful. If a person can't see very well it could analyze for them. It would need sensors to recognize smoke. It should be able to connect to an ambulance like a medical alert…the person's heart rate and all that good stuff. If the diabetics sugar drops low, it should be able to tell it. It should be able to get it off the phone and help." [Participant 24] • "It would be great if you were stranded…and you could still call someone to save you." [Participant 2]

XSL•FO
RenderX

Favored functions	Participants who selected the function as the "best" function, n (%)	Itemized survey (1="least helpful" to 4="most helpful"), mean (SD)	Example representative quote
GPS	3 (11.54)	3.38 (0.80)	• "The question 'Where am I?' is a big one…Or if you get lost. These things are important to know." (Participant 8) • "For example, I used to live on [redacted street]. When I moved down here, I traveled all the way to go to Giant. But I could have asked the glasses where the closest Giant is. I find it complicated to look up things on phone and computers and it's easier with the glasses to get information." (Participant 23)

Discussion

Principal Findings

This study used surveys and interviews to explore the usability and acceptability of smart glasses among older adults with cognitive impairment. Both survey and interview results suggest that participants prioritized the following smart glasses functions: audio reminders, phone calls, GPS, and distress signals, with audio reminders emerging as the highest-ranked feature. Following beta testing of 3 pairs of commercial smart glasses, interview results suggest that participants prefer the Ray-Ban Meta smart glasses due to their intuitive and quick interface, comfort and esthetics, and functionally relevant features such as audio reminders and instant hands-free information. Overall, participants conveyed a generally positive perception of smart glasses and their potential to support memory in daily life.

This study supports the development and exploration of smart glasses for older adults with cognitive impairment. Some clinicians might question the foundation of such work, believing that smart glasses for older adults with cognitive impairment could be problematic due to low technology literacy and the potential for increased confusion in this population [48]. As an alternative perspective, this work contributes to the previous literature [49,50], which demonstrates that older adults are diverse in technology use and ability, and many are indeed comfortable with technology and view it favorably. Furthermore, all 26 older participants who participated in this study, including those with lower levels of technology comfort, stated that they would be willing to wear smart glasses that assist with memory. One potential explanation is that the potential benefits for supporting independence could outweigh the expected initial usability challenges. This work demonstrated the strength of user-centered research in emerging fields, allowing individuals to articulate their own preferences and needs.

Although participants conveyed a generally positive perception of smart glasses and their potential to support memory in daily life, they also identified some concerns. For instance, 2 participants highlighted the importance of smart glasses users being respectful of others' privacy when capturing photo or audio recordings. While enforcing such boundaries could prove practically difficult, most smart glasses have an indicator light that turns on when photo or video is activated to ensure bystanders are aware of such capture. Additionally, 1 participant expressed concern that ongoing use of smart glasses for cognitive assistance could lead to a reliance on the device and a loss of independent cognitive skills. It would indeed be valuable for future longitudinal work to explore whether the cognitive assistance provided by smart glasses in the short term has a "de-training" effect over the long term.

The results of this study suggest that older adults with cognitive impairment strongly value smart glasses that are easy to activate and equipped with a large vocabulary, thus providing interaction that is intuitive. During the past few years, amidst the proliferation of AI, smart glasses have become more user-friendly. As AI technology continues to advance, it is expected that smart glasses will become even more "intelligent," with improvements in connectivity, data collection and processing, pattern recognition, automation, and adaptability [51]. Over time, as smart glasses require less user input to operate correctly, they could become more appropriate for those with low technology literacy. Going forward, even though the participants verbalized a preference for more polished smart glasses models, older adults with cognitive impairment should still be included in early-stage development research. Including end users in early-stage design of technology helps ensure that the development of technology is applicable and helpful for them and improves the future likelihood of use.

Both survey and interview results suggest that older adults with cognitive impairment perceive audio reminders as the most critical smart glasses function. One of the earliest and most common symptoms of cognitive impairment is episodic memory loss, which is characterized by difficulty in recalling specific personal everyday events and experiences [52]. Audio reminders could help with episodic memory by providing alerts for important upcoming events such as doctor's appointments, medication administration, or social activities. Similarly, individuals with deficiencies in episodic memory may have difficulty navigating to a new place after 1 visit (eg, a parking garage). These deficiencies are consistent with the participants' reports of GPS and navigation as prioritized functions.

XSL•FO
RenderX

The participants' additional preferred smart glasses functions—phone calls and distress signals—reflect the challenges and safety concerns often faced by individuals with cognitive impairment who live alone. Such individuals can experience a sense of uncertainty due to the unpredictable course of symptoms and how these symptoms could negatively impact their lives each day [53]. Especially among this population of economically disadvantaged older adults, many lack adequate home-based services [54]. Smart glasses, equipped with accessible communication features such as phone call capabilities and emergency alerts, may provide older adults with cognitive impairment and their caregivers improved security and assurance that assistance can be readily accessible to the older adult when needed.

Implications

This work provides guidance for smart glasses developers regarding the design and features that are most important for older adults with cognitive impairment. First and foremost, smart glasses developers should place great importance on usability, including straightforward activation and intuitive command interfaces. Developers should also equip the smart glasses with functions that are relevant to this population, especially audio reminders, and ensure that the glasses are minimalistic so that older adults can receive help in a discreet manner. Finally, and contrary to much current development research [22], until augmented reality technology has considerably advanced, this work suggests that developers of smart glasses for older adults with cognitive impairment should prioritize audio rather than visual feedback for assistance with memory and daily tasks.

This work also provides information for researchers of smart glasses for older adults with cognitive impairment. A recent review on smart glasses for older adults with cognitive impairment, conducted by Burch et al [22], found that the current literature focuses either on outdated commercial smart glasses or rudimentary prototypes, such as regular glasses with bulky technology components (eg, cameras) added to the side arm. This study suggests that current commercial smart glasses provide a promising platform for assisting older adults with cognitive impairment, but additional studies are needed to examine the integration of smart glasses into long-term daily life. Additionally, if health researchers collaborate with computer science researchers on the development of future smart glasses, this work could serve as a strong launching pad for informing preferred smart glasses design and functionality.

Although this work does not explicitly assess the acceptability of smart glasses compared to other smart devices, some interpretive insights can be considered as to why smart glasses might be particularly advantageous. First, participants consistently appreciated the hands-free nature of the smart glasses, which allow users to request information while using a walker or performing a task with both hands. Many other wearable smart devices (eg, smartwatches) are not as supportive of hands-free interaction. Additionally, similar to "earables," smart glasses are situated near the ear, allowing for more discreet audio assistance than other smart devices. Another reason smart glasses might stand apart compared to other smart devices is that participants valued that the glasses could "see what they see" and provide relevant information about their environment. Finally, participants appreciated that a device they already carry with them could provide such advanced assistance. As smart glasses naturally travel with the user, they provide more possibilities for assistance compared to stationary technologies such as smart speakers. Notably, the aforementioned points should be considered as thought-launching considerations rather than definitive conclusions, as this study was not designed to compare smart glasses with other smart devices.

Limitations and Strengths

Limitations of this study include a small sample size from a single older adult apartment building, short duration of smart glasses use and testing, and a lack of objective measures. First, the sample size of only 26 participants limits the generalizability of the study results, especially to those with more severe cognitive impairment or those who live in more rural settings. Second, participants beta tested the smart glasses for only 20 minutes per pair. During the testing, the participants could have been influenced by branding, such as associating a particular name brand with quality. Thus, the findings might not translate to long-term preferences and do not reveal information about adherence during everyday life. Finally, no validated survey currently exists that examines the preferred functions of smart glasses. Instead, we developed and used a "preferred functions" survey, which has not been tested for reliability and validity. Responses on this survey during phase 1 informed our selection and purchase of smart glasses for phase 2. As such, participants did not have hands-on experience with the glasses prior to completing the survey.

Study strengths include a user-centered focus, a mixed methods design that included rich interviews, and a sample that reflects older adults who are disproportionately affected by cognitive impairment. Within an emerging field of research, especially those involving new technology, obtaining feedback from intended users is critical to ensure that development and technology selection align well with user needs, thereby increasing the likelihood of adoption. The mixed methodology approach used in this study provided rich insights that numbers alone could not reflect. Finally, we recruited our sample from affordable older adult housing to study a group of older adults who are disproportionately affected by cognitive impairment yet underrepresented in technology research.

Conclusions

This study aimed to use surveys and interviews to explore the usability and acceptability of smart glasses among older adults with cognitive impairment. Overall, participants conveyed a generally positive perception of smart glasses and their potential to support memory in daily life. Both survey and interview results suggest that participants prioritized the following smart glasses functions: audio reminders, phone calls, GPS, and distress signals, with audio reminders emerging as the highest-ranked feature. Additionally, participants highlighted that audio feedback from the glasses, rather than visual feedback, was the preferred method of information delivery and emphasized the value of an intuitive and quick interface. Future work should examine the integration of smart glasses over a

XSL•FO
RenderX

longer period, in a larger sample of older adults with cognitive impairment, and use inferential analytic approaches.

Acknowledgments

During the preparation of this work, the first author (BFB) used Microsoft Copilot and ChatGPT solely for the purpose of editing and improving the readability of sentences. The authors used these artificial intelligence tools consistent with the guidelines outlined in “Best Practices for Using AI as an Author, Peer Reviewer, or Editor” [55]. The authors take full responsibility for the content and references within this publication.

Funding

The time allocated for this project was supported by the University of Maryland Claude D. Pepper Older Americans Independence Center (P30-AG028747).

Data Availability

Data from this work can be made available by the authors upon reasonable request.

Authors' Contributions

BFB and BR developed the concept, design, and manuscript draft. BFB, NK, RM, and DA assisted with data collection. BFB and NK performed data organization and analysis. All authors contributed to manuscript editing, revision, and completion.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey: preferred smart glasses’ functions.
[DOCX File, 28 KB - aging_v9i1e81840_app1.docx]

Multimedia Appendix 2
Codebook.
[DOCX File, 20 KB - aging_v9i1e81840_app2.docx]

Multimedia Appendix 3
Additional participant quotations.
[DOCX File, 18 KB - aging_v9i1e81840_app3.docx]

References

1. Dementia statistics. Alzheimer’s Disease International. URL: https://www.alzint.org/about/dementia-facts-figures/dementia-statistics/ [accessed 2025-02-05]
2. Pais R, Ruano L, P Carvalho O, Barros H. Global cognitive impairment prevalence and incidence in community dwelling older adults-a systematic review. Geriatrics (Basel) 2020 Oct 27;5(4):84. [doi: 10.3390/geriatrics5040084] [Medline: 33121002]
3. Hugo J, Ganguli M. Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. Clin Geriatr Med 2014 Aug;30(3):421-442. [doi: 10.1016/j.cger.2014.04.001] [Medline: 25037289]
4. Jongsiriyanyong S, Limpawattana P. Mild cognitive impairment in clinical practice: a review article. Am J Alzheimers Dis Other Demen 2018 Dec;33(8):500-507. [doi: 10.1177/1533317518791401] [Medline: 30068225]
5. Li R, Li R, Xie J, et al. Associations of socioeconomic status and healthy lifestyle with incident early-onset and late-onset dementia: a prospective cohort study. Lancet Healthy Longev 2023 Dec;4(12):e693-e702. [doi: 10.1016/S2666-7568(23)00211-8] [Medline: 38042162]
6. Park S, Amano T, Park S, Kim BR. Prevalence of mild cognitive impairment and Alzheimer’s disease and related dementias among older residents of publicly subsidized housing in the United States: systematic review and meta-analysis. J Gerontol Soc Work 2025 Jun 21:1-16. [doi: 10.1080/01634372.2025.2522986] [Medline: 40542642]
7. Xiang X, An R. The impact of cognitive impairment and comorbid depression on disability, health care utilization, and costs. Psychiatr Serv 2015 Nov;66(11):1245-1248. [doi: 10.1176/appi.ps.201400511] [Medline: 26073413]
8. Shin JY, Habermann B. Caregivers of adults living with Alzheimer’s disease or dementia in 2020: a secondary analysis. J Gerontol Nurs 2022 Sep;48(9):15-25. [doi: 10.3928/00989134-20220805-02] [Medline: 36007215]

XSL•FO
RenderX

9. Watson B, Tatangelo G, McCabe M. Depression and anxiety among partner and offspring carers of people with dementia: a systematic review. Gerontologist 2019 Sep 17;59(5):e597-e610. [doi: 10.1093/geront/gny049] [Medline: 29878117]
10. Buckley JS, Salpeter SR. A risk-benefit assessment of dementia medications: systematic review of the evidence. Drugs Aging 2015 Jun;32(6):453-467. [doi: 10.1007/s40266-015-0266-9] [Medline: 25941104]
11. Qiao Y, Chi Y, Zhang Q, Ma Y. Safety and efficacy of lecanemab for Alzheimer's disease: a systematic review and meta-analysis of randomized clinical trials. Front Aging Neurosci 2023;15:1169499. [doi: 10.3389/fnagi.2023.1169499] [Medline: 37213538]
12. Huang X, Zhao X, Li B, et al. Comparative efficacy of various exercise interventions on cognitive function in patients with mild cognitive impairment or dementia: a systematic review and network meta-analysis. J Sport Health Sci 2022 Mar;11(2):212-223. [doi: 10.1016/j.jshs.2021.05.003] [Medline: 34004389]
13. Butler M, McCreedy E, Nelson VA, et al. Does cognitive training prevent cognitive decline?: A systematic review. Ann Intern Med 2018 Jan 2;168(1):63-68. [doi: 10.7326/M17-1531] [Medline: 29255842]
14. Anderson ND. State of the science on mild cognitive impairment (MCI). CNS Spectr 2019 Feb;24(1):78-87. [doi: 10.1017/S1092852918001347] [Medline: 30651152]
15. Silverio-Fernández M, Renukappa S, Suresh S. What is a smart device? - A conceptualisation within the paradigm of the internet of things. Vis in Eng 2018 Dec;6(1). [doi: 10.1186/s40327-018-0063-8]
16. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
17. Sanders EC, Boot WR. Usability of the smartwatch as a prospective memory aid for cognitively diverse older adults. J Appl Gerontol 2026 Feb;45(2):382-392. [doi: 10.1177/07334648251338876] [Medline: 40432385]
18. Goerss D, Köhler S, Rong E, et al. Smartwatch-based interventions for people with dementia: user-centered design approach. JMIR Aging 2024 Jun 7;7:e50107. [doi: 10.2196/50107] [Medline: 38848116]
19. Salehi W, Gupta G, Bhatia S, Koundal D, Mashat A, Belay A. IoT-based wearable devices for patients suffering from Alzheimer disease. Contrast Media Mol Imaging 2022;2022(1):3224939. [doi: 10.1155/2022/3224939] [Medline: 35542758]
20. Carroll C, Chiodo C, Lin AX, Nidever M, Prathipati J. Robin: enabling independence for individuals with cognitive disabilities using voice assistive technology. Presented at: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17); May 6-11, 2017; Denver, Colorado, United States p. 46-53. [doi: 10.1145/3027063.3049266]
21. Kokorelias KM, Grigorovich A, Harris MT, et al. Longitudinal coadaptation of older adults with wearables and voice-activated virtual assistants: scoping review. J Med Internet Res 2024 Aug 7;26:e57258. [doi: 10.2196/57258] [Medline: 39110963]
22. Burch BF, Gu J, Betz G, et al. Smart glasses for older adults with cognitive impairment: a scoping review. J Am Med Dir Assoc 2025 Nov;26(11):105831. [doi: 10.1016/j.jamda.2025.105831] [Medline: 40882953]
23. Zuidhof N, Ben Allouch S, Peters O, Verbeek PP. Defining smart glasses: a rapid review of state-of-the-art perspectives and future challenges from a social sciences' perspective. Augment Hum Res 2021 Dec;6:15. [doi: 10.1007/s41133-021-00053-3]
24. Wei NJ, Dougherty B, Myers A, Badawy SM. Using Google Glass in surgical settings: systematic review. JMIR mHealth uHealth 2018 Mar 6;6(3):e54. [doi: 10.2196/mhealth.9409] [Medline: 29510969]
25. Yu J, Ferniany W, Guthrie B, Parekh SG, Ponce B. Lessons learned from Google Glass: telemedical spark or unfulfilled promise? Surg Innov 2016 Apr;23(2):156-165. [doi: 10.1177/1553350615597085] [Medline: 26224576]
26. Luckerson V. Google will stop selling glass next week. TIME. 2015. URL: https://time.com/3669927/google-glass-explorer-program-ends/ [accessed 2025-08-05]
27. Garcia A. Google glass lives on in the workplace. The latest pair costs $999. CNN Business. 2019. URL: https://www.cnn.com/2019/05/20/tech/google-glass-enterprise-edition-2 [accessed 2025-07-27]
28. Gsaxner C, Li J, Pepe A, et al. The HoloLens in medicine: A systematic review and taxonomy. Med Image Anal 2023 Apr;85:102757. [doi: 10.1016/j.media.2023.102757] [Medline: 36706637]
29. Metz R. Magic leap raised billions but its headset flopped. Now it's trying again. CNN Business. URL: https://www.cnn.com/2022/03/10/tech/magic-leap-new-headset [accessed 2025-08-05]
30. Eadicicco L. Magic leap's CEO on his 'experiential computer' that blends tech and reality. TIME. URL: https://time.com/5074834/magic-leap-one-rony-abovitz/ [accessed 2025-08-05]
31. Zari G, Condino S, Cutolo F, Ferrari V. Magic Leap 1 versus Microsoft HoloLens 2 for the visualization of 3D content obtained from radiological images. Sensors (Basel) 2023 Mar 11;23(6):3040. [doi: 10.3390/s23063040] [Medline: 36991751]
32. Kaviani F, Lyall B, Koppel S. Exploring social perceptions of everyday smartglass use in Australia. PLoS One 2024;19(11):e0313001. [doi: 10.1371/journal.pone.0313001] [Medline: 39485765]
33. Berryman DR. Augmented reality: a review. Med Ref Serv Q 2012;31(2):212-218. [doi: 10.1080/02763869.2012.670604] [Medline: 22559183]
34. Haesner M, Wolf S, Steinert A, Steinhagen-Thiessen E. Touch interaction with Google Glass – is it suitable for older adults? Int J Hum Comput Stud 2018 Feb;110:12-20. [doi: 10.1016/j.ijhcs.2017.09.006]

XSL•FO
RenderX

35. Essig K, Strenge B, Schack T. ADAMAAS: towards smart glasses for mobile and personalized action assistance. Presented at: Proceedings of the 9th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '16); Jun 29 to Jul 1, 2016. [doi: 10.1145/2910674.2910727]
36. Montuwy A, Cahour B, Dommes A. Using sensory wearable devices to navigate the city: effectiveness and user experience in older pedestrians. Multimodal Technol Interact 2019 Dec;3(1):17. [doi: 10.3390/mti3010017]
37. Wolf D, Besserer D, Sejunaite K, Schuler A, Riepe M, Rukzio E. CARE: an augmented reality support system for geriatric inpatients with mild cognitive impairment. Presented at: MUM '19: Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia; Nov 26-29, 2019. [doi: 10.1145/3365610.3365612]
38. Zhan K, Faux S, Ramos F. Multi-scale Conditional Random Fields for first-person activity recognition on elders and disabled patients. Pervasive Mob Comput 2015 Jan;16:251-267. [doi: 10.1016/j.pmcj.2014.11.004]
39. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
40. Bevan N, Carter J, Earthy J, Geis T, Harker S. New ISO standards for usability, usability reports and usability measures. Presented at: Proceedings of the 18th International Conference on Human-Computer Interaction: Theory, Design, Development and Practice; Jul 17-22, 2016; Toronto, Canada p. 268-278. [doi: 10.1007/978-3-319-39510-4_25]
41. Borson S, Scanlan JM, Chen P, Ganguli M. The Mini-Cog as a screen for dementia: validation in a population-based sample. J Am Geriatr Soc 2003 Oct;51(10):1451-1454. [doi: 10.1046/j.1532-5415.2003.51465.x] [Medline: 14511167]
42. Vale FAC, Balieiro AP Jr, Silva-Filho JH. Memory complaint scale (MCS). Proposed tool for active systematic search. Dement Neuropsychol 2012;6(4):212-218. [doi: 10.1590/S1980-57642012DN06040004] [Medline: 29213800]
43. Resnick B, Gruber-Baldini AL, Pretzer-Aboff I, et al. Reliability and validity of the evaluation to sign consent measure. Gerontologist 2007 Feb;47(1):69-77. [doi: 10.1093/geront/47.1.69] [Medline: 17327542]
44. Walker JL. The use of saturation in qualitative research. Can J Cardiovasc Nurs 2012;22(2):37-46. [Medline: 22803288]
45. Lepkowsky CM, Arndt S. Functional assessment of currently employed technology scale (FACETS): reliability and validity. Int J Med Sci Clin Inven 2018;5(9):4064-4068. [doi: 10.18535/ijmsci/v5i9.07]
46. Tips for dementia-friendly surveys. Alzheimer's Society. URL: https://www.alzheimers.org.uk/dementia-professionals/dementia-experience-toolkit/research-methods/tips-dementia-friendly-surveys [accessed 2026-04-03]
47. Merriam SB, Tisdell EJ. Qualitative Research A Guide to Design and Implementation, 4th edition: Jossey-Bass; 2015.
48. Mace RA, Mattos MK, Vranceanu AM. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022 Dec 30;12(12):1102-1105. [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
49. James CA, Basu T, Nallamothu BK, Kullgren JT. Use of digital health technologies by older US Adults. JAMA Netw Open 2025 Jan 2;8(1):e2454727. [doi: 10.1001/jamanetworkopen.2024.54727] [Medline: 39813034]
50. Drazich BF, Li Q, Perrin NA, et al. The relationship between older adults' technology use, in-person engagement, and pandemic-related mental health. Aging Ment Health 2023 Jan;27(1):156-165. [doi: 10.1080/13607863.2022.2046695] [Medline: 35243945]
51. Wang B, Zheng Y, Han X, et al. A systematic literature review on integrating AI-powered smart glasses into digital health management for proactive healthcare solutions. NPJ Digit Med 2025 Jul 5;8(1):410. [doi: 10.1038/s41746-025-01715-x] [Medline: 40617964]
52. Nordahl CW, Ranganath C, Yonelinas AP, DeCarli C, Reed BR, Jagust WJ. Different mechanisms of episodic memory failure in mild cognitive impairment. Neuropsychologia 2005;43(11):1688-1697. [doi: 10.1016/j.neuropsychologia.2005.01.003] [Medline: 16009250]
53. Portacolone E, Rubinstein RL, Covinsky KE, Halpern J, Johnson JK. The precarity of older adults living alone with cognitive impairment. Gerontologist 2019 Mar 14;59(2):271-280. [doi: 10.1093/geront/gnx193] [Medline: 29373676]
54. Portacolone E, Nguyen TT, Bowers BJ, et al. Perceptions of the role of living alone in providing services to patients with cognitive impairment. JAMA Netw Open 2023 Aug 1;6(8):e2329913. [doi: 10.1001/jamanetworkopen.2023.29913] [Medline: 37594759]
55. Leung TI, de Azevedo Cardoso T, Mavragani A, Eysenbach G. Best practices for using AI tools as an author, peer reviewer, or editor. J Med Internet Res 2023 Aug 31;25:e51584. [doi: 10.2196/51584] [Medline: 37651164]

Abbreviations

AI: artificial intelligence
FACETS: Functional Assessment of Currently Employed Technology Scale
ISO: International Organization for Standardization

XSL•FO
RenderX

Edited by Y Jiang; submitted 06.Aug.2025; peer-reviewed by C Becker, D Bradford, G Spalla; revised version received 27.Feb.2026; accepted 09.Mar.2026; published 27.Apr.2026.

Please cite as:
Burch BF, Kim N, McPherson R, Anokye D, Ryan AS, Addison O, Roy N, Huang CM, Galik E, Resnick B
Smart Glasses for Older Adults With Cognitive Impairment: Explanatory Mixed Methods Study
JMIR Aging 2026;9:e81840
URL: https://aging.jmir.org/2026/1/e81840
doi:10.2196/81840

© Brittany F Burch, Nayeon Kim, Rachel McPherson, Debora Anokye, Alice S Ryan, Odessa Addison, Nirmalya Roy, Chien-Ming Huang, Elizabeth Galik, Barbara Resnick. Originally published in JMIR Aging (https://aging.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of a Web-Based Resource to Support Driving Safety in Older Adults: Semistructured Interview and Focus Group Study

Mollie Cahill[1,2], BSc, MNurs; Ranmalee Eramudugolla[1,2,3], DPsych; Marwin Tang[1,2], BPsych (Hons); Diane Hosking[4], PhD; Stephen Cratchley[5]; Michael Regan[6], PhD; Lesley A Ross[7], PhD; Jasmine Price[2], BHSc; Kaarin J Anstey[1,2,3], PhD

[1]School of Psychology, University of New South Wales, UNSW Sydney, High St, Kensington, Sydney, Australia

[2]Neuroscience Research Australia, Sydney, Australia

[3]University of New South Wales, Ageing Futures Institute, Kensington, Sydney, Australia

[4]National Seniors Australia, Brisbane, Australia

[5]Suncorp Group, Brisbane, Australia

[6]Research Centre for Integrated Transport Innovation (rCITI), University of New South Wales, Sydney, Australia

[7]Department of Psychology, Clemson University, South Carolina, United States

Corresponding Author:
Kaarin J Anstey, PhD
School of Psychology, University of New South Wales, UNSW Sydney, High St, Kensington, Sydney, Australia

Abstract

Background: Older adults face increased crash risk due to age-related declines in cognitive, visual, and physical functioning; yet, many Australians in their 70s are continuing to drive. Web-based platforms are increasingly used to deliver health and mobility information to older adults and may support safer driving; however, existing online resources on driving safety often lack age-specific guidance, have usability limitations, and may not be designed with older adults in mind.

Objective: This study aimed to explore the information needs, preferences, and experiences of older drivers and general practitioners (GPs) to inform the formative, prelaunch development of a user-centered web-based platform, Ageing Well on the Road. Interviews and focus groups were conducted to gather feedback on content, accessibility, and usability, with data collection guided by the Consolidated Framework for Implementation Research and the Technology Acceptance Model frameworks to explore implementation and technology adoption factors.

Methods: Website development was informed by a review of 26 Australian websites and consultations with stakeholders. Semistructured interviews were conducted with 7 older adults (aged >67 years; one additional participant completed only survey data) and a focus group with 10 GPs experienced in fitness to drive assessments for older drivers. Thematic analysis was used to explore participants' experiences with online driving information and feedback on content, usability, and accessibility. Formative feedback from these users directly informed content and design refinements before public launch.

Results: The website addressed priority topics, including age-related health conditions affecting driving, licensing requirements, vehicle safety features, and strategies for maintaining driving confidence. Older adults valued large fonts, plain language, intuitive navigation, and trustworthy, practical advice. GPs highlighted the website's potential to support clinical discussions around driving capacity and safety. Barriers to engagement included digital literacy limitations, accessibility concerns, and uncertainty about online information credibility.

Conclusions: A co-designed, web-based platform tailored to older adults and clinicians can enhance access to relevant, trustworthy information and support driving self-regulation. Involving end users in the development process is critical to optimize usability and ensure the resource aligns with their needs, highlighting the value of theory-informed, user-guided development for digital health resources.

(JMIR Aging 2026;9:e79630) doi:10.2196/79630

KEYWORDS

older drivers; ageing; driving safety; web-based platform; usability; fitness to drive; digital health; technology acceptance; digital literacy; co-design

XSL•FO
RenderX

Introduction

Age-related declines in cognitive, sensory, and motor functions contribute to increased crash risk and driving difficulties among older adults [1,2]. Despite these challenges, over 90% of drivers in their 70s continue to rely on private vehicles, with many expected to drive well into their 80s [3]. Mobility plays a key role in preserving independence and well-being in late life, making it essential to support safe driving practices and provide accessible, evidence-based resources that help older adults make informed decisions about driving safety. However, many existing resources are not tailored to the unique needs of older adults, such as cognitive changes and digital literacy levels, highlighting a gap in accessible, targeted educational tools designed specifically for older drivers.

Web-based platforms have emerged as flexible and cost-effective methods for delivering health interventions, especially as over 75% of older adults now use the internet to seek health and mobility information [4]. While on-road driving assessments have shown effectiveness in evaluating driving ability [5,6], such services are not widely accessible. Older drivers, their family members, and clinicians often report difficulty locating clear, practical guidelines on driving safety, vehicle modifications, and licensing regulations [7,8]. While existing websites about driving safety provide general information, they frequently overlook the unique needs of older users, such as support with age-related decline, planning for driving retirement, and web design accessible to those with reduced physical or cognitive capacity [8].

Although digital health tools can broaden access to information, their effectiveness depends on how well they meet users' expectations regarding usability, accessibility, and content relevance [9]. Older adults may face particular barriers to engaging with digital platforms, such as limited digital literacy, physical impairments, and uncertainty about the credibility of online information [10]. To address these challenges, theoretical frameworks can help guide the design and evaluation of new resources, as we first must understand how older adults search for information and what content they find useful. The Technology Acceptance Model (TAM), which has been applied in studies of older drivers' adoption of advanced driver-assistance systems [11], suggests that perceived ease of use and perceived usefulness are key determinants of whether older adults engage with new technologies. This makes TAM a valuable lens for understanding how older adults evaluate online driving safety tools and identifying features that support or hinder uptake.

This study aimed to explore the information needs, preferences, and experiences of older drivers and general practitioners (GPs) with online driving safety content to inform the development of a user-centered web-based platform. A qualitative approach was used, including a review of existing online resources, consultations with stakeholders, and interviews with older adults (aged ≥67 years) and GPs. GPs play a key role in assessing older drivers' medical fitness to drive (FtD) in Australia, yet often report uncertainty and a lack of appropriate tools to support these discussions [12,13]. These findings underscore the need for clear, accessible resources to support GPs in their decision-making. Given that internet use among older Australians aged >75 years continues to rise, with a sharp increase from 52% in 2019 (before COVID-19–related lockdowns) to 94% in 2022 [14], the development of a tailored web platform, Ageing Well on the Road [15], presents a timely opportunity to improve knowledge, support driving self-regulation, and enhance road safety. Findings from this study contribute formative insights to guide resource refinement and the development of best practice strategies for disseminating online driving safety tools targeted at older drivers.

Methods

Participants and Recruitment

Older adults (aged ≥67 years) were recruited to explore their information-seeking behaviors and perceptions of a newly developed web-based resource on driving safety. Recruitment was conducted via volunteer networks, social media, and partnerships with organizations such as National Seniors Australia. Additionally, participants from the Better Drive Trial and the Community in Health Aging Research Group who had previously consented to research participation were invited. GPs were recruited through the research team's professional network and were eligible to participate if they were currently practicing and had experience conducting FtD assessments with older adults. Eligible participants received a link to an online screening and consent form hosted on REDCap (Research Electronic Data Capture; Vanderbilt University), a secure online survey platform. Those who consented were invited to participate in qualitative interviews or focus groups, where they were shown a prelive version of the website, allowing them to provide feedback on content, design, and usability before public launch, as well as to explore their broader information-seeking behaviors, preferences, and experiences with online driving safety content.

Eight older adults participated in the study, 7 of whom completed semistructured interviews, and 10 GPs participated in a single focus group discussion. Although the total sample size was modest (n=18), this is appropriate for the exploratory and qualitative nature of the study, which aimed to obtain rich, detailed accounts of participants' perspectives rather than statistical generalizability [16]. In usability and user experience research, small samples are widely considered sufficient to identify the majority of issues, with previous studies showing that 5 - 10 participants often reveal key concerns and that additional insights may be gained with up to 20 participants [17]. Recruitment continued until no additional expressions of interest were received, and the research team judged that sufficient feedback had been obtained to inform website development. As a pretesting stage, this sample provided practical insights into user experience, content clarity, and usability. Feedback from participants, together with input from the research team and stakeholders, directly informed iterative refinements to the website. Furthermore, the platform remains open to ongoing feedback via contact details provided on the site, allowing continued updates and improvements post launch.

XSL•FO
RenderX

Website Development

A comprehensive review of 26 Australian websites addressing older driver safety was conducted by a research assistant under the supervision of the project team to assess existing resources, identify content gaps, and inform the development of a new web-based platform. Although not a formal systematic review, this process was guided by content experts, clinicians, stakeholder input, and team discussions, including a meeting with representatives from National Seniors Australia, the Australian Pensioners Insurance Agency (APIA), Suncorp Group, and Transport for NSW to ensure that relevant resources and key perspectives were captured. Websites were identified through targeted Google searches and recommendations from stakeholders.

The review highlighted several consistent patterns and gaps. Most government and national websites provided general information on licensing requirements, legal obligations for drivers with medical conditions, safe driving tips, and alternative transport options. Many included downloadable resources such as fact sheets, PDFs, and videos, and frequently linked to external resources rather than consolidating content in one place. However, few sites offered tailored guidance specifically for older drivers on maintaining driving skills, using vehicle technologies such as advanced driver assistance systems (ADASs), and supporting independence and decision-making around driving retirement. Accessibility features, interactive decision aids, and practical checklists were present in some resources, but not consistently across sites. Table S1 in Multimedia Appendix 1 provides a detailed overview of the content, format, and features of the 26 websites reviewed. Limited information on ADAS, such as lane-keeping assist, adaptive cruise control, and blind spot monitoring, was identified, despite evidence that these technologies can reduce crash risk and improve driving performance among older drivers, particularly when systems are acceptable and usable [18-20].

The resulting website, Ageing Well on the Road [15], was developed through an iterative, user-informed design process that incorporated feedback from older adults, GPs, and key stakeholders. The website design prioritized accessibility features appropriate for older adults, including clear headings, plain language, simple navigation structures, and large easy-to-read fonts. Stakeholder consultations involved representatives from an older adult advocacy group (National Seniors Australia), an older adult's insurance agency (APIA), an insurance group (Suncorp Group), and a government transport authority (Transport for NSW), who emphasized the importance of trustworthy, practical resources that could be easily updated as regulations changed. Interview and focus group data directly informed the structure and content of the website. The resulting platform addressed content gaps identified in the review and reflected user and stakeholder input, with sections focused on maintaining driving skills with age, understanding and using ADAS technologies, planning for driving retirement, navigating licensing and medical requirements, and supporting family and carers. This approach aimed to produce an accessible, evidence-based resource to support safe driving and informed mobility decisions for older adults in Australia. A map of corresponding changes made to the website before launch is available in the Multimedia Appendix 1.

Data Collection

Data were collected using an online questionnaire and through qualitative interviews and focus groups. The questionnaire captured consent, basic demographic information, and participants' driving habits and attitudes toward licensing and safety. It included the participant information statement and consent form, which participants were required to review before proceeding. Consent was recorded electronically, with participants informed of their right to withdraw at any time. Interviews and focus groups were conducted via Microsoft Teams or Zoom (Zoom Communications) and ranged from 30 to 60 minutes in length. A semistructured interview approach was used, combining predetermined open-ended questions with flexibility for participants to elaborate on issues of relevance. The semistructured interview and focus group guides are provided in Multimedia Appendix 1.

Interview guides were informed by the Consolidated Framework for Implementation Research (CFIR), a widely used framework that provides a structured way of identifying the barriers and facilitators influencing the implementation and uptake of health interventions [21,22]. This framework allowed us to systematically design questions that explore contextual factors affecting the adoption of web-based driving safety resources. Specifically, questions about the credibility, usability, and perceived usefulness of the website reflected the domain of intervention characteristics; questions exploring licensing requirements, information sources, and available supports reflected the outer setting; and questions about older adults' motivations, needs, confidence, and comfort with digital tools reflected individual characteristics. By aligning our guide with these domains, we ensured that data captured both personal experiences and broader implementation considerations, providing insights into factors that may support or hinder the uptake of health and safety interventions. This approach also allowed us to capture influences across multiple levels of context. For example, participants' experiences with licensing regulations, access to transport, family involvement, and available community resources reflect the CFIR inner and outer setting domains, highlighting how both organizational and external factors shape engagement with driving safety interventions.

In addition, the TAM [11] guided the design of questions focused on digital engagement. TAM constructs of perceived usefulness and perceived ease of use shaped prompts asking participants to reflect on how straightforward the website was to navigate, whether the information was relevant to their needs, and whether they would recommend the resource to others. For clinicians, TAM principles were reflected in questions about preferred formats for digital screening tools and the likelihood of incorporating these into practice. These links allowed user feedback to be directly interpreted within established theoretical models of technology adoption, particularly for older populations. All sessions were audio-recorded, transcribed, and thematically analyzed. Feedback on user experience and content relevance was reviewed by the research team and stakeholders

XSL•FO
RenderX

and directly informed website revisions to improve clarity, accessibility, and ease of use.

Analysis

Thematic analysis was conducted by 2 researchers using NVivo (QSR International), following the phases described by Braun and Clarke [16,23]. The process involved familiarization with the transcripts, generating initial codes, allocating data segments to the codes, searching for themes, refining themes to ensure they worked in relation to the data, and developing the final thematic framework. Reflexive practices were incorporated throughout, including ongoing consideration of how researcher perspectives shaped coding and interpretation. Two researchers from different disciplinary backgrounds (paramedicine and psychology) independently conducted initial coding of the transcripts. Coding differences and emerging interpretations were discussed collaboratively, and themes were refined iteratively to incorporate multiple perspectives and support reflexivity, producing a more nuanced interpretation of the data [16,23].

Feedback on the website was analyzed primarily deductively, guided by the interview guide and research aim, which were informed by the CFIR and TAM. At the same time, themes related to participants' broader knowledge needs, attitudes toward existing resources, and perceived barriers and facilitators were identified inductively from the data, so unexpected insights could be identified [24]. By combining framework-guided coding with open coding, we were able to use established theory to guide our analysis while also capturing new and unanticipated insights. The coding frameworks are provided in Multimedia Appendix 1.

Participant insights ultimately shaped the content, design, and functionality of the final website. Four main themes from the interviews and focus groups were identified: (1) barriers to accessing information and resources, (2) facilitators of online resource use, (3) driving safety and self-regulation, and (4) motivations and resource needs for information search. Themes were iteratively refined during analysis and further organized into subthemes at the reporting stage to aid clarity. Relevant formative feedback on the prelive website has also been incorporated to illustrate how these insights informed refinements to content, layout, and usability.

Ethical Considerations

This study was approved by the Human Research Ethics Committee at the University of New South Wales (approval no iRECS6396), and participants provided written informed consent. No monetary compensation was provided to participants. Additionally, participants recruited from the Better Drive Trial were originally enrolled under separate approval from the University of New South Wales Human Research Ethics Committee (approval no HC190430) and had previously consented to be contacted for future research. All identifiable information was removed during data collection, and participants were assigned unique codes. The linking file connecting identities to study data was stored separately and was accessible only to the research team. Reidentification was not required, as no participants withdrew their data. Findings are reported in aggregate form to ensure that individual participants cannot be identified.

Results

Characteristics of the Older Adults

Eight older adults consented to participate via the REDCap survey, which collected basic demographics, driving habits, and interview availability. Seven participants completed the interviews. Participants' ages ranged from 67 to 85 years. All were current drivers, with one participant reporting restrictions on their license, requiring spectacles and annual medical reviews. Most participants drove more than half the days of the week, with 3 driving 6 days per week and 3 driving 7 days per week. Weekly driving distances ranged from 31 km to over 150 km. When asked about how many more years they anticipated continuing to drive, the mean number was 17 (SD 6.84) years, with the range spanning from 8 to 30 years. Half of the participants reported having completed a medical review of their FtD by a doctor, while the other half had not.

Confidence in understanding licensing requirements after the age of 75 years was mixed. Three participants expressed high confidence, and 3 expressed low confidence, and the remaining participants reported neutral confidence. Confidence in managing age-related changes affecting driving safety was higher; 4 participants reported high confidence, and 3 reported moderate confidence. Most (n=6) participants also expressed high confidence in managing their own driving safety. Regarding satisfaction with existing information on licensing, aging, and driving safety, participants were generally neutral about both quality (n=4) and quantity (n=5). Common sources of accessing information included family members (n=5), doctors (n=4), and friends (n=2), with online resources being the preferred format for health and driving information (n=6).

Characteristics of GPs

Ten GPs participated in a single focus group. Their ages varied, with 3 participants in their 30s, 2 in their 40s, 2 in their 50s, and 3 in their 60s. Three worked full-time and 7 part-time. Two had joined their practice within the past year, while the others had been with their practice for over 20 years. All GPs held a medical degree and Fellowship of the Royal Australian College of General Practitioners. Two also held PhDs, and 3 held other professional roles, such as being on the board of the Primary Health Network, postgraduate medical education, or running a GP training program.

Barriers to Accessing Information and Resources

Reluctance to Acknowledge Driving Limitations

The older participant group identified a challenge in acknowledging the need for assessment, particularly among long-time drivers. One participant reflected on the difficulty of convincing long-time drivers to recognize when they may no longer be fit to drive, reflecting a broader hesitation in seeking out resources until necessary:

> *The biggest problem all around is persuading people who should be doing something more than an annual test… I think persuading people to use these things*

XSL•FO
RenderX

> *[online resources and tools], is perhaps the key. And perhaps persuading someone who isn't really aware of the difficulties they are having.* [P6, Male, 84]

This reluctance was echoed by the GPs, who noted that older adults typically do not recognize their own limitations when it comes to driving. One GP said:

> *It's pretty rare… when a patient says "maybe I shouldn't be driving." I think that's happened to me once. It may be family members that may say something, but it's rare for the individual to say "no, I'm not good enough."* [GP1]

These observations highlight the challenge both GPs and family members face and suggest the need for resources that help family members recognize signs of potential driving impairment and guide conversations about driving safety.

Reactive Information Seeking

Participants reported that they typically only sought information when it was urgently needed. As one older adult explained, "You're not going to be prompted to go looking for information that you don't need at the moment" (P3, Male, 67). This demonstrates a reactive approach to information-seeking, where older adults may not actively seek driving-related information unless prompted by personal circumstances or external cues such as medical diagnoses or legal requirements.

Skepticism Toward Online Resources

Concerns about the accuracy and reliability of digital tools emerged as another barrier. One older adult expressed doubt about online assessments: "I'm skeptical that a computer questionnaire can actually make an assessment on my ability to drive" (P4, Male, 77). This skepticism was combined with fears regarding privacy and potential misuse of personal data, indicating that trust in online resources remains a significant barrier for older adults, despite the widespread and increasing internet use and online screening tools.

Complexity of Licensing Regulations

Finally, participants highlighted that the complexity of state-by-state licensing regulations can cause confusion, especially for those who regularly travel across Australia. One older adult noted: "The laws are different in every state" (P7, Male, 71), pointing to the need for a centralized, comprehensive resource that consolidates licensing requirements and driving regulations across all states and territories into a single, accessible platform.

Facilitators of Online Use

Accessibility and Convenience

Despite barriers, several factors were identified that could facilitate the use of online resources. Participants highlighted that web-based platforms are appealing because they are quick, easy, and convenient to use. One older adult explained: "It's very quick. It's very easy. It's very comprehensive. You'll get lots more answers" (P4, Male, 77). Similarly, another participant highlighted the convenience of online resources: "Well, it's readily available and easy to access. I don't have to go to the bookshop and look for it or look for a magazine" (P7, Male, 71). These comments underscore that the practical accessibility of online resources is a key facilitator for older adults.

Reducing Anxiety and Stigma

Online resources may also help alleviate the anxiety older adults feel about formal FtD assessments and the potential loss of their license. Participants in the GP focus group noted that many older adults experience significant stress around these assessments: "They get very anxious. Very worried they will fail and their license will be taken away from them, and there's a rapid ball of catastrophising that goes on" (GP). This may discourage older adults from engaging in the process altogether, as one older adult described their personal reaction to medical appointments: "Every time I go to the GP, my blood pressure goes up, because outside the GP's office my blood pressure is normal" (P5, Male, 85). While online resources cannot replace formal assessments, providing clear, transparent, and nonjudgmental information online may reduce the stress associated with seeking information from health care professionals and help individuals feel better prepared before attending in-person assessments. By framing these resources as supportive tools that clarify the process and expectations, they may reduce fear, facilitating greater acceptance and engagement with driving safety information.

Advisory Guidance and Reassurance

Some participants valued online tools that provided guidance rather than a definitive diagnosis, helping them feel more confident about their driving safety. One participant reflected on hesitating to use a potential dementia assessment tool, but noted a different approach for driving:

> *I discovered an online assessment for potential dementia, and I hesitated about doing it. I thought, maybe, I don't want to know. I think that could be less the case in respect to driving. I would want to know, because I'd be putting other people at risk, other than just myself.* [P3, Male, 67]

This highlights how online tools can offer reassurance by helping individuals feel more confident that they are driving safely and not endangering others. Similarly, another participant emphasized the importance of feedback rather than prescriptive outcomes:

> *I wouldn't mind as long as it's a tool that's helping you to determine if you've got some issues or not. But it's not a definitive yes or no. You've got major issues, but something that sort of gives you feedback about how you answer the questions or deal with the online tools would be useful.* [P1, Male, 73]

Participants generally agreed that such tools should be "advisory and not prescriptive" (P2, Male, 74). These findings suggest that online resources can support older adults in managing their driving abilities in a nonthreatening and informative way, providing reassurance and reducing anxiety by framing driving safety as something manageable and shared rather than stigmatizing.

XSL•FO
RenderX

Driving Safety and Self-Regulation

Proactive Approaches to Safety

Some participants took an active role in managing their driving safety by seeking assessments to maintain their driving confidence and reassurance. One participant described:

> *My specialist suggested I have this occupational therapy driving assessment, or something, which I did. I prepared for and did all the reading before that, so that's basically how I got interested and realised that I can still drive properly and safely and confidently.* [P2, Male, 74]

This proactive approach to safety reflects how formal assessments can help older adults make informed decisions and continue driving with confidence, while also acknowledging the need for future adjustments as age-related declines (eg, reduced strength or flexibility) may affect certain driving-related skills, such as towing heavy vehicles.

Social Feedback and Family Influence

Social input was another critical component of driving self-regulation. Participants described relying on their peers or family members for feedback about their driving ability. As P4 (Male, 77) noted, "drivers will always talk to another driver about their ability or a problem." Family concerns also influenced decisions, with one older driver recalling, "At one point, my wife's mum, we were concerned about her driving, and we really didn't want her to drive anymore. It's one of those things—she was mid-80s at that stage" (P4, Male, 77). These examples show that self-regulation often occurs within a broader social context, influenced by both peer discussions and family involvement.

Self-Imposed Driving Restrictions

Self-regulation, including voluntarily limiting driving, was a frequent strategy to maintain safety, shaped by age-related caution or health concerns. For example, one participant explained:

> *My wife says that she isn't going to drive after dark. She didn't go and look and see something somewhere that said "this is how to know whether you can drive after dark or not." She simply just said, "I don't really feel that great driving after dark."* [P6, Male, 84]

This self-imposed restriction was a common behavior among older adults, suggesting that the decision to stop driving may not always be based on formal assessments, but rather personal comfort and awareness of limitations. GPs confirmed this tendency, noting that many older adults consciously restricted their driving to familiar routes or comfortable distances. As one GP stated:

> *I find I have patients who self-restrict their own driving, and say, I only drive where I feel comfortable now…. they've just personally downgraded their driving status within their comfort level.* [GP1]

These findings suggest that self-regulation is a common and pragmatic response, and online tools could help validate these choices by providing reassurance and offering guidance about when to consult with a health care professional.

Access to Alternative Transport

The availability of public transport influenced decisions about driving cessation. Older driver participants living in urban areas reported that giving up driving was more manageable due to better access to amenities and public transport. As P6 (Male, 84) stated, "People in big cities (Melbourne, Sydney, etc) often give up [driving] quite easily as they are near the things they want and the public transport isn't too bad." In contrast, those in rural areas described more difficulty: "Country districts and places like Canberra… it's going to be important and difficult to deal with. If you can't get to the nearest shop" (P6, Male, 84). Limited transport alternatives in these areas make driving cessation more challenging, highlighting the need for tailored resources and support in this area.

Motivations and Resource Needs for Information Search

Health-Related Prompts

Medical diagnoses were a key motivator for seeking information about driving requirements. For example, one older driver participant shared:

> *I've got mild cognitive impairment. And my specialist suggested I complete an occupational therapy assessment…, which I passed with flying colours. And now we're continuing to drive safely and confidently, but it was a good wake-up call.* [P2, Male, 74]

Similarly, GPs noted that new diagnoses such as diabetes, syncope, or stroke often required them to notify licensing authorities, which in turn triggered patients to seek information about driving safety and regulatory requirements.

Family Influence and Forward Planning

Some participants were prompted to search for information because of concerns about their relatives, for example, researching licensing regulations after a relative's dementia diagnosis. One older driver participant explained, "I looked online through the Roads and Traffic Authority, Roads and Maritime Services, and Service NSW websites, and just Googled generally about the issues" (P3, Male, 67). Such examples show how family experiences and concerns can indirectly motivate individuals to seek out resources and prepare for their own future driving needs. In contrast, some participants did not feel an immediate need for information, but acknowledged it would become relevant with age: "The legal requirements for drivers as we get to different ages would be a topic I'd certainly choose to Google" (P1, Male, 73). This reflects a more anticipatory approach, where individuals acknowledge that information will become useful as they age or if their circumstances change.

Preferred Sources of Information

Most participants preferred online resources for accessing driving-related information, with one older driver participant stating, "I prefer to use Google all the time" (P3, Male, 67), while others emphasized the need for these resources to be credible and relevant: "Well, I wouldn't look at anything

XSL•FO
RenderX

American for example, so relevant to my location. But not put there by an insurance company, it's got to be a reputable source" (P7, Male, 71). This emphasis on trustworthiness shows the importance of ensuring that online resources are not only easy to access but also authoritative and specific in their content.

External Prompts and Regulatory Triggers

Despite varied preferences, most participants reported that information-seeking often occurred reactively, in response to external prompts. These included medical advice, legal obligations, family member suggestions, or official reminders. As one GP described:

> *Someone newly diagnosed as a diabetic has to be notified, or a syncope episode or a stroke—those sorts of things—I'll sometimes go and notify the RMS [Roads and Maritime Services] for them, as at least you know that way it's done.* [GP3]

This illustrates how medical diagnoses such as diabetes or syncope can motivate both health care professionals and patients to seek relevant driving safety information and comply with legal requirements. Similarly, one older adult explained, "I got a letter in the mail saying you are due for a check-up… It's a push technique, not a pull" (P6, Male, 84). These examples illustrate how external systems, in addition to personal initiatives, often drive engagement with information about licensing and driving safety.

Preliminary Impressions of the Ageing Well on the Road Website

Participants were invited to review a prerelease version of the Ageing Well on the Road website and provide formative feedback. This input guided refinements to content and design before public launch, as a full usability evaluation was beyond the scope of this study. Older adults described the site as clear, accessible, and easy to navigate. One participant noted: "It wasn't too detailed, and it was in plain English…. I thought it was easy to read, and well presented" (P1, Male, 73), while another commented: "I thought it was easy to navigate, it was broken up into convenient tiles… it included things like technically advanced motor vehicles, and getting cars to assist with lane control etc" (P4, Male, 77). Credibility associated with the university affiliation was also highlighted as supporting trust in the information: "It's got the University of New South Wales symbol... so that gives it some credibility" (P7, Male, 71).

Specific content elements were particularly valued, including state-based licensing information and visual resources. One participant explained: "One aspect I found very good was that you had each state and territory listed, and their licensing requirements" (P4, Male, 77), while another added: "There was a graphic about the blind spots for heavy vehicles, and that was informative to me, because I didn't realize just how blind they are" (P3, Male, 67). Participants indicated they would likely recommend the website to peers, families, and community groups once it is live, noting, "Once it's online, then I'll certainly be promoting it as a good source of information" (P3, Male, 67). Overall, feedback was positive, and participants' broader insights on information needs, preferences, and driving safety concerns were used alongside this input to refine the website before public launch.

GPs also provided positive feedback and highlighted the site's potential use in clinical practice. One GP commented that the site "looks great" and that they would "absolutely refer patients to this." Suggested refinements included adding practical financial guidance for patients retiring from driving, such as demonstrating potential savings from selling a vehicle vs ongoing taxi or ride-share costs, and including information on taxi subsidy schemes. GPs recommended incorporating research-based statistics and evidence on driving frequency and skill decline to support clinical discussions. These suggestions informed adjustments to content and resources before the public launch.

Discussion

Principal Findings

This work represents formative, prelaunch development of the website and aimed to explore how older adults seek information about driving requirements and safety, their motivations and resource needs, and how GPs support older adults in making FtD decisions. Insights on online health-seeking behaviors, information preferences, and driving-related concerns informed the development of the website. Our findings suggest that older adults' information-seeking behaviors are often needs-driven, influenced by personal circumstances, medical conditions, and external prompts, such as family concerns or government notifications, as identified in interviews with older adults and GPs. This aligns with prior research suggesting that older adults often seek information when prompted by a specific need or external factor [25,26].

A key observation was that individuals with existing medical conditions, such as mild cognitive impairment, were more likely to actively seek information. For example, one participant described how a specialist's recommendation for an occupational therapy assessment prompted his search, reinforcing the belief that health-related triggers often lead to increased engagement with driving-related information [3]. Similarly, those with familial experiences of cognitive decline, such as dementia, sought information to better understand the legal and safety implications of continued driving. This finding mirrors studies that emphasize the important role family members and caregivers play in initiating conversations about driving cessation [27]. In contrast, participants who had not yet faced challenges to their driving ability expressed little motivation to seek information. This supports the idea that information-seeking behavior among older adults is often "needs-based" and typically occurs when it becomes personally relevant [26]. However, participants noted that legal requirements would become important over time, indicating they were aware of the information's future relevance even if not immediately needed.

Some participants in the interviews mentioned holding specialized driving licenses, such as for heavy vehicles or trucks. Unsurprisingly, these participants were generally more familiar with regulatory requirements due to the mandatory annual

XSL•FO
RenderX

medical assessments these licenses require. These individuals did not actively seek out additional information, as they felt their licensing processes inherently provided them with updates on medical and cognitive evaluations. This supports the idea that external regulatory frameworks can be effective in disseminating information to older adults, as also found in previous studies focused on policy-driven health behaviors [28].

Regarding preferred information sources, most participants favored online searches, especially Google and government websites. This suggests a need for user-friendly, government-verified online resources tailored to older adults, reflecting other findings from studies that show older adults are increasingly turning to digital platforms for information [29]. Building on these findings, research suggests that internet use can improve the physical and mental health of older adults [30] and that digital literacy is a key factor influencing engagement with online health information. Older adults can be disadvantaged in terms of physiological function, cognitive ability, social status, and economic status, which can limit their use of electronic technology compared with younger individuals [30]. Those with higher digital literacy are more confident navigating websites, evaluating content credibility, and applying information in daily life, whereas those with lower digital literacy may encounter barriers to effectively using digital resources [30,31]. Considering digital literacy and the local sociocultural context when designing a web-based platform could enhance accessibility, usability, and overall engagement, ensuring that the online resource is inclusive for older adults with varying levels of technological skill. When interpreted through the lens of CFIR and TAM, several constructs were particularly relevant in shaping user perceptions. For CFIR, intervention characteristics (such as perceived credibility and usability of the website) and individual characteristics (including confidence and comfort with digital tools) strongly influenced engagement. For TAM, perceived usefulness and perceived ease of use were key determinants of how participants evaluated and were willing to adopt new digital resources [11].

A key theme that emerged was the prevalence of a “push” rather than “pull” approach to information seeking. Many participants reported they did not seek out information unless prompted by an external cue, such as a family intervention or a government letter advising them to go for a medical check-up. This aligns with behavioral models of health communication, which suggest that external cues are often the primary motivators for action [32]. Overall, while some adults proactively sought driving-related information, many reported that they rely on external triggers. To promote safe driving behaviors among aging populations, efforts should leverage these “push” mechanisms to ensure older drivers remain informed and make safe decisions while driving.

In addition to the individual benefits, improving driving safety in older adults has broader societal implications. Older drivers are often overrepresented in crash and injury statistics [11], so supporting safer driving practices through proactive self-regulation and education in later life can contribute to the overall safety of all road users. Therefore, derisking older driver behavior is not only important for maintaining independence and mobility, but also for potentially reducing the public health burden associated with road traffic incidents.

Feedback on the prelive website was constructive and informed content and presentation refinements, providing an initial indication of reach and potential usability. Future research should be conducted to examine longer-term use, gather structured user feedback, and investigate how digital resources can be optimized for this demographic, including strategies to encourage proactive information-seeking. The website does not collect any user data, but visitor metrics, such as the number of views, are monitored monthly to provide ongoing insights into user engagement.

Limitations

This study involved a small, qualitative sample that offers in-depth insights, but limits generalizability to the broader population. The older adult sample was predominantly male and recruited from participants already engaged with research projects, potentially introducing positive bias. Limited demographic information was available for the GP focus group, as the hosting practice did not collect full participant-level data to minimize burden, with only summary information on age ranges, work status, and qualifications available, and sex was not reported. Study findings reflect the Australian licensing and health context and may differ in other jurisdictions. This work was formative and undertaken before the website’s public launch. While participants’ perspectives informed the website design, this study was not a formal usability evaluation. A comprehensive postlaunch evaluation was outside the scope of this project and will require further funding, but would be beneficial, with more diverse users and regions.

Conclusions

This study provides insight into factors influencing older adults’ engagement with driving safety resources. Findings suggest that information-seeking is often reactive, prompted by health issues, family concerns, or legal requirements. Barriers to engagement include skepticism about online resources, the complexity of licensing regulations, and reluctance to acknowledge the need for driving assessments, highlighting the importance of accessible, trustworthy, and user-friendly resources. External cues, such as medical recommendations and family and caregiver input, can encourage engagement with driving safety resources. GPs play a critical role, often serving as the first point of contact for older adults experiencing health issues affecting their driving. Proactive GP involvement, coupled with clear, reliable digital resources, may support older adults in seeking and applying relevant information. Future research should explore how health care professionals, particularly GPs, can be better supported and educated to enhance their role in FtD assessments.

Feedback from older adults, GPs, and stakeholders informed the formative, prelaunch development of the Ageing Well on the Road website [15]. Insights guided the creation of a dedicated clinician section and shaped the site’s plain-language content, intuitive navigation, and additional resources for family members and caregivers, who often play a critical role in initiating conversations about driving safety and helping older

XSL•FO
RenderX

adults navigate relevant information. The multiuser design aims to increase accessibility and promote shared decision-making. Stakeholders involved in the co-design process included representatives from National Seniors Australia, APIA, Suncorp Group, and Transport for NSW, reinforcing the need for accurate, trustworthy, and up-to-date content. Together, these contributions helped create a user-informed, accessible web-based resource tailored to the real-world needs of older adults and clinicians.

Overall, this formative work indicates that a co-designed, web-based platform has the potential to simplify access to essential information, build trust in digital content, and encourage early engagement by older adults and GPs. By addressing key barriers, the platform may contribute to safer driving practices, support independence, and inform broader road safety initiatives. Although developed for an Australian context, the platform's structure could be adapted internationally, with state- or region-specific driving regulations mapped to local legal requirements while maintaining the multiuser design and content approach.

Acknowledgments

The research team would like to thank the research volunteers who contributed to this study.

Funding

We would like to acknowledge the support of the ACT Road Safety Fund (grant reference RSG2023027), which provided funding for this research. We also thank Megan Hiller, Research Assistant, for her contributions to the content development and editing of the website. KJA was funded by an Australian Research Council Laureate Fellowship (FL190100011).

Data Availability

The data are available from the corresponding author, KJA, upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Website review, methodological details, and coding framework.
[DOCX File, 55 KB - aging_v9i1e79630_app1.docx]

References

1. Anstey KJ, Eramudugolla R, Ross LA, Lautenschlager NT, Wood J. Road safety in an aging population: risk factors, assessment, interventions, and future directions. Int Psychogeriatr 2016 Mar;28(3):349-356. [doi: 10.1017/S1041610216000053] [Medline: 26888735]
2. Toups R, Chirles TJ, Ehsani JP, et al. Driving performance in older adults: current measures, findings, and implications for roadway safety. Innov Aging 2022;6(1):igab051. [doi: 10.1093/geroni/igab051] [Medline: 35028434]
3. Anstey KJ, Li X, Hosking DE, Eramudugolla R. The epidemiology of driving in later life: sociodemographic, health and functional characteristics, predictors of incident cessation, and driving expectations. Accid Anal Prev 2017 Oct;107:110-116. [doi: 10.1016/j.aap.2017.07.026] [Medline: 28818682]
4. Tennant B, Stellefson M, Dodd V, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar 17;17(3):e70. [doi: 10.2196/jmir.3992] [Medline: 25783036]
5. Anstey KJ, Eramudugolla R, Kiely KM, Price J. Effect of tailored on-road driving lessons on driving safety in older adults: a randomised controlled trial. Accid Anal Prev 2018 Jun;115:1-10. [doi: 10.1016/j.aap.2018.02.016] [Medline: 29518703]
6. Castellucci HI, Bravo G, Arezes PM, Lavallière M. Are interventions effective at improving driving in older drivers?: A systematic review. BMC Geriatr 2020 Apr 3;20(1):125. [doi: 10.1186/s12877-020-01512-z] [Medline: 32245367]
7. Betz ME, Haukoos JS, Schwartz R, et al. Prospective validation of a screening tool to identify older adults in need of a driving evaluation. J Am Geriatr Soc 2018 Feb;66(2):357-363. [doi: 10.1111/jgs.15222] [Medline: 29231960]
8. Levasseur M, Coallier JC, Gabaude C, et al. Facilitators, barriers and needs in the use of adaptive driving strategies to enhance older drivers' mobility: importance of openness, perceptions, knowledge and support. Transportation Research Part F: Traffic Psychology and Behaviour 2016 Nov;43:56-66. [doi: 10.1016/j.trf.2016.09.014]
9. Lackie ME, Parrilla JS, Lavery BM, et al. Digital health needs of women with postpartum depression: focus group study. J Med Internet Res 2021 Jan 6;23(1):e18934. [doi: 10.2196/18934] [Medline: 33404506]
10. Radovic A, Odenthal K, Flores AT, Miller E, Stein BD. Prescribing technology to increase uptake of depression treatment in primary care: a pre-implementation focus group study of SOVA (Supporting Our Valued Adolescents). J Clin Psychol Med Settings 2020 Dec;27(4):766-782. [doi: 10.1007/s10880-019-09669-5] [Medline: 31630347]

XSL•FO
RenderX

11. Hansen A, Kiely K, Attuquayefio T, et al. Assessment of the application of technology acceptance measures to older drivers' acceptance of advanced driver-assistance systems. Appl Ergon 2025 May;125:104474. [doi: 10.1016/j.apergo.2025.104474] [Medline: 39893764]
12. Jones K, Rouse-Watson S, Beveridge A, Sims J, Schattner P. Fitness to drive - GP perspectives of assessing older and functionally impaired patients. Aust Fam Physician 2012 Apr;41(4):235-239. [Medline: 22472687]
13. Stefanidis KB, Summers MJ, Clark M. The assessment of cognitive function and medical fitness to drive in older adults: a qualitative study with general practitioners. J Transp Health 2025 Jun;42:102020. [doi: 10.1016/j.jth.2025.102020]
14. Communications and media in Australia series: how we use the internet executive summary and key findings. Australian Communications and Media Authority. 2024. URL: https://www.acma.gov.au/sites/default/files/2023-03/HOWWEU~1.PDF [accessed 2026-04-18]
15. Ageing well on the road. Driving, Ageing and Health Research Team. 2024. URL: https://ageingwellontheroad.com.au/ [accessed 2026-04-18]
16. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
17. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]
18. Masello L, Castignani G, Sheehan B, Murphy F, McDonnell K. On the road safety benefits of advanced driver assistance systems in different driving contexts. Transportation Research Interdisciplinary Perspectives 2022 Sep;15:100670. [doi: 10.1016/j.trip.2022.100670]
19. Young KL, Regan MA, Triggs TJ, Tomasevic N, Stephan K, Mitsopoulos E. Impact on car driving performance of a following distance warning system: findings from the Australian Transport Accident Commission SafeCar project. J Intell Transp Syst 2007 Jul 3;11(3):121-131. [doi: 10.1080/15472450701410411]
20. Regan M, Stephan K, Mitsopoulos E, et al. The effect on driver workload, attitudes and acceptability of in-vehicle intelligent transport systems. J Road Safety 2007;18(1):30-36 [FREE Full text]
21. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
22. O'Connor CM, Poulos CJ, Kurrle S, Anstey KJ. Bridging the gap: study protocol for development of an implementation strategy for evidence-informed reablement and rehabilitation for community-dwelling people with dementia. Arch Gerontol Geriatr 2023 May;108:104943. [doi: 10.1016/j.archger.2023.104943] [Medline: 36701945]
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Sharp P, Bottorff JL, Rice S, et al. "People say men don't talk, well that's bullshit": A focus group study exploring challenges and opportunities for men's mental health promotion. PLoS ONE 2022;17(1):e0261997. [doi: 10.1371/journal.pone.0261997] [Medline: 35061764]
25. Fang Z, Liu Y, Peng B. Empowering older adults: bridging the digital divide in online health information seeking. Humanit Soc Sci Commun 2024;11(1). [doi: 10.1057/s41599-024-04312-7]
26. Shen X, Helion C, Smith DV, Murty VP. Motivation as a lens for understanding information-seeking behaviors. J Cogn Neurosci 2024 Feb 1;36(2):362-376. [doi: 10.1162/jocn_a_02083] [Medline: 37944120]
27. Adler G, Rottunda S. Older adults' perspectives on driving cessation. J Aging Stud 2006 Sep;20(3):227-235. [doi: 10.1016/j.jaging.2005.09.003]
28. Hakamies-Blomqvist L, Henriksson P. Cohort effects in older drivers' accident type distribution: are older drivers as old as they used to be? Transp Res Part F Traffic Psychol Behav 1999 Sep;2(3):131-138. [doi: 10.1016/S1369-8478(99)00009-1]
29. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022 Feb 16;24(2):e34790. [doi: 10.2196/34790] [Medline: 35171099]
30. Shi Z, Du X, Li J, Hou R, Sun J, Marohabutr T. Factors influencing digital health literacy among older adults: a scoping review. Front Public Health 2024;12:1447747. [doi: 10.3389/fpubh.2024.1447747] [Medline: 39555039]
31. Yang Y, Yao X, Lu D, et al. Improving the eHealth literacy of older adults: a scoping review. Geriatr Nurs (Lond) 2024 Nov;60:128-136. [doi: 10.1016/j.gerinurse.2024.07.028]
32. Weinstein ND. The precaution adoption process. Health Psychol 1988;7(4):355-386. [doi: 10.1037//0278-6133.7.4.355] [Medline: 3049068]

Abbreviations

ADAS: advanced driver assistance system
APIA: Australian Pensioners Insurance Agency
CFIR: Consolidated Framework for Implementation Research
FtD: fitness to drive
GP: general practitioner

XSL•FO
RenderX

REDCap: Research Electronic Data Capture
TAM: Technology Acceptance Model

Edited by R Bjarnadottir; submitted 25.Jun.2025; peer-reviewed by K Stefanidis, M Lavalliere, NS Aziz, S Mohanadas, X Liang; revised version received 03.Dec.2025; accepted 27.Feb.2026; published 19.May.2026.

Please cite as:
Cahill M, Eramudugolla R, Tang M, Hosking D, Cratchley S, Regan M, Ross LA, Price J, Anstey KJ
Development of a Web-Based Resource to Support Driving Safety in Older Adults: Semistructured Interview and Focus Group Study
JMIR Aging 2026;9:e79630
URL: https://aging.jmir.org/2026/1/e79630
doi:10.2196/79630

© Mollie Cahill, Ranmalee Eramudugolla, Marwin Tang, Diane Hosking, Stephen Cratchley, Michael Regan, Lesley A Ross, Jasmine Price, Kaarin J Anstey. Originally published in JMIR Aging (https://aging.jmir.org), 19.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Web-Based Gamified Auditory-Cognitive Dual-Task Training for Older Adults With Age-Related Hearing Loss: Pilot Randomized Controlled Trial

Ivy Yan Zhao[1,2,3], PhD; Angela Yee Man Leung[1,2,3], PhD; Chen Li[4], PhD; Laurence Lloyd Parial[5], PhD; Hongming Ma[1], MSc; Jed Montayre[1,6], PhD; Justin S Golub[7], MS, MD; Robert Sweetow[8], PhD; Janet Ho-Yee Ng[9], DHSc; Engle Angela Chan[1], PhD

[1]School of Nursing, The Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, Hong Kong, China (Hong Kong)

[2]WHO Collaborating Centre for Community Health Services, School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[3]Research Institute for Smart Ageing (RISA), The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[4]Department of Applied Social Sciences & Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[5]College of Nursing, University of the Philippines Manila, Manila, National Capital Region, Philippines

[6]JBI Hong Kong Centre of Evidence-based Healthcare Excellence, School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[7]Department of Otolaryngology-Head and Neck Surgery, Columbia University Vagelos College of Physicians and Surgeons, NewYork-Presbyterian/Columbia University Irving Medical Center, New York, NY, United States

[8]School of Medicine, University of California, San Francisco, San Francisco, CA, United States

[9]Department of Language Science and Technology, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

Corresponding Author:
Ivy Yan Zhao, PhD
School of Nursing, The Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, Hong Kong, China (Hong Kong)

Abstract

Background: Everyday listening ability is essential for individual health and well-being. Age-related hearing loss (ARHL) is associated with reduced communication engagement, social isolation, loneliness, cognitive decline, and increased dementia risk. Interventions that simultaneously target auditory processing and cognitive function, particularly within engaging and ecologically valid contexts, may offer greater benefits than unimodal approaches. However, culturally adapted, web-based, gamified auditory-cognitive dual-task training (ACDT) tailored for older adults with ARHL remains underexplored. At the time of this writing, few auditory or auditory-cognitive training programs are available in Chinese languages, creating linguistic and cultural barriers for older adults.

Objective: This study aimed to (1) assess the feasibility and acceptability of home-based ACDT among older Chinese adults with ARHL and (2) examine its preliminary effects on global cognition, hearing, social engagement, and loneliness. It was hypothesized that the intervention group would demonstrate greater improvements in global cognition, hearing, and social engagement than the control group.

Methods: Sixty community-dwelling older adults with mild-to-moderate ARHL were randomized 1:1 to either the ACDT group or a waitlist control group in a single-blinded pilot randomized controlled trial. Demographic data and outcome measures were collected at baseline, week 6, and week 12. Postintervention interviews were conducted to assess the feasibility and acceptability of ACDT.

Results: A total of 60 participants were randomized (mean age 67.65, SD 4.78 years; 45/60, 75% male). ACDT demonstrated high feasibility and acceptability. The ACDT group showed significant improvements in focused attention (mean change=–0.15; P=.02; d=–0.46) and divided attention (mean change=–0.21; P=.002; d=–0.63). Significant cognitive improvements on the Hong Kong-Montreal Cognitive Assessment were identified in naming (r=0.37; P=.05) and visual cognition (r=0.44; P=.02) in the intervention group, while no significant improvements were found in the control group. Both groups reported significant decreases in emotional hearing handicap, with slightly greater improvement in the intervention group (r=0.39; P=.03) than in the control group (r=0.37; P=.04). Linear mixed model analysis revealed a small to moderate group effect (Cohen d=0.38) for 5-minute delayed recall on the Auditory Verbal Learning Test, with the fixed effects explaining 69% of the variance (marginal R^2=0.69). A significant time×group interaction was observed for left-ear thresholds (P=.01). Qualitative analysis identified three key themes: (1) intervention coherence and participants' affective attitude toward ACDT; (2) perceived benefits in cognition, information acquisition, and self-awareness from ACDT; and (3) perceiving ACDT as less burdensome with enhanced self-efficacy.

XSL•FO
RenderX

Conclusion: Future iterations should incorporate artificial intelligence–enhanced personalization. Large-scale randomized controlled trials involving diverse samples and active control conditions are needed to confirm sustained effects on auditory and cognitive health, dual-task listening-cognitive abilities, and real-world functioning.

Trial Registration: ClinicalTrials.gov NCT06486285; https://clinicaltrials.gov/study/NCT06486285

(JMIR Aging 2026;9:e84083) doi:10.2196/84083

KEYWORDS

randomized controlled trial; age-related hearing loss; cognitive training; auditory training; gamified; dual-task training

Introduction

Everyday listening ability is essential for individual health and well-being. Age-related hearing loss (ARHL) represents a highly prevalent sensory deficit, affecting approximately one-third of adults aged 60 years or older [1,2]. ARHL is associated with reduced communication engagement, social isolation, and loneliness [3-5]. Beyond these social consequences, ARHL shows consistent associations with cognitive decline and increased risk of dementia [6-8]. Research indicates that even subclinical ARHL may increase the long-term risk of cognitive decline and dementia [9-12], with greater hearing loss severity correlating with poorer performance on memory, executive function, and mental status measures [13,14].

The mechanisms underlying these associations appear multifactorial, involving a complex interplay of sensory deprivation, increased cognitive load during auditory processing, accelerated brain atrophy (particularly in temporal regions), and reduced social-cognitive stimulation leading to neural disuse [15-19]. The sensory deprivation hypothesis suggests that prolonged reductions in sensor input (eg, hearing loss) increase the cognitive effort required for listening, and are subsequently associated with neural atrophy [17-19], while reduced auditory stimulation may cause structural brain changes, such as cortical thinning, reduced gray matter volume, and altered connectivity in auditory and multimodal networks [15,16]. The Cattell-Horn-Carroll theory of cognitive abilities provides a structured framework for describing and classifying the cognitive domains examined in relation to ARHL [20]. For example, increased listening effort may place greater demands on cognitive resources, particularly working memory and processing speed. This chronic reallocation of resources may reduce the capacity available for higher-order cognitive functions, including novel problem-solving, memory retrieval, and the application of acquired knowledge. Thus, hearing loss can cascade into broader cognitive impairment, extending beyond auditory processing alone [20]. Longitudinal evidence indicates that untreated ARHL is associated with a 30% - 50% increased risk of incident dementia and accelerates cognitive decline by 30% - 40% over 6 - 10 years [13,21,22]. Hearing loss treatment in mid-life may reduce dementia risk by 8% [23] with potential benefits in lowering health care use and costs [24,25]. These findings highlight the need for early intervention to mitigate the cognitive risks associated with ARHL.

While hearing aids may help mitigate the risk of cognitive decline [26], adherence challenges persist, and their benefits might be limited primarily to peripheral hearing deficits [27]. However, ARHL also involves central auditory processing difficulties, which appear to be more strongly associated with cognitive impairment than pure tone thresholds. These central deficits place significant demands on cognitive domains that are essential for understanding speech in complex listening environments. Accordingly, interventions that simultaneously target auditory processing and cognitive function [27], particularly in engaging and ecologically valid contexts, may offer greater benefits than unimodal approaches [28,29]. The auditory trainings for older adults with hearing loss mainly focus on speech or word recognition in noise, with auditory processing as the primary outcome [28]. Some studies have expanded these approaches by combining auditory training with cognitive training [30], contextual and visual cues [31], or elements of real-world communication [32]. For example, a randomized controlled trial (RCT) in the United Kingdom used computer-delivered phoneme training with 11 synthesized phoneme continua derived from natural voice recordings and implemented using a 3-interval, 3-alternative forced-choice oddball paradigm. Although the training produced substantial on-task learning effects in older adults with hearing loss, further research is needed to determine which training tasks are most effective, especially given the considerable individual variability in outcomes [30]. Dual-task training approaches, which require participants to perform distinct measurable tasks concurrently or sequentially [33], represent a promising direction for maximizing cognitive benefits [34]. Such training may improve the allocation of limited cognitive resources under challenging listening conditions while also providing repeated stimulation that could support neuroplastic changes in relevant networks [27]. In parallel, web-based cognitive training has shown promise for enhancing cognition [35,36], and gamification may further improve engagement in digital health interventions [37].

However, culturally adapted, web-based, gamified auditory-cognitive dual-task training (ACDT) tailored for older adults with ARHL remains underexplored. At the time of this writing, few auditory or auditory-cognitive trainings are available in Chinese languages, creating linguistic and cultural barriers for older adults. Given evidence that auditory training delivered in one's native language may enhance comprehension, engagement, and listening skill transfer [38], this study aimed to (1) assess the feasibility and acceptability of a web-based, gamified ACDT program performed at home among older Chinese adults with ARHL, and (2) examine its preliminary effects on global cognition, hearing, social engagement, and loneliness. It was hypothesized that the intervention group would demonstrate greater improvements in global cognition, hearing, and social engagement compared with the control group.

XSL•FO
RenderX

Methods

Study Design

A prospective, single-blind, 2-arm parallel pilot RCT, followed by descriptive semistructured individual interviews, was conducted between April 28, 2023, and June 27, 2025. Reporting followed CONSORT (Consolidated Standards of Reporting Trials; Checklist 1) guidelines [39], with the trial protocol available in Multimedia Appendix 1.

Ethical Considerations

The institutional review board of the Hong Kong Polytechnic University (HSEARS20230223006) approved the study. All participants enrolled in this study voluntarily and signed informed consent before the formal start of the assessment after being informed that they could withdraw at any time. Each participant was introduced to the aims and content of the study, potential risks, and benefits. All participants provided written informed consent in accordance with the Declaration of Helsinki. HKD $200 (~US $26; a currency exchange rate on April 28, 2023, is applicable) supermarket vouchers were provided to participants as an incentive to compensate for their time spent on the interview and completion of questionnaires. All data, audio recordings, and written records are stored on a password-protected laptop; only the research team has access to them. In addition, all data were used for research purposes only and could not be disclosed without the participants' consent. To ensure privacy protection, the participants' personal information was anonymized and labeled "P1-P60."

Study Participants

Participants were recruited using convenience sampling through community health centers and advertisements across Hong Kong (eg, social media). Eligibility criteria included (1) aged 60 years or older and living in the community; (2) mild-to-moderate hearing loss, defined as a pure-tone average (PTA) of 20 - 50 dB across octave frequencies from 0.5 to 4 kHz in both ears; (3) no prior use of hearing aids; (4) normal cognitive performance (Montreal Cognitive Assessment score ≥26); and (5) willingness and ability to provide informed consent and comply with study procedures. Exclusion criteria were (1) a history of psychosis, mania, bipolar disorder, substance use disorder, or suicidal ideation; (2) severe or unstable medical illness; (3) significant retrocochlear pathology or an organic lesion causing hearing loss; (4) a diagnosis of probable Alzheimer disease, vascular dementia, frontotemporal dementia, or Parkinson disease; and (5) use of medications, such as antidepressants, sedatives, or antiepileptics, that may affect cognition.

Sample Sizes

This pilot RCT was not designed to definitively test intervention effects, which would require a fully powered trial. Due to the absence of established effect sizes from previous literature, a target sample of 60 participants (30 per arm), allowing for 15% attrition, was determined based on recommendations for pilot RCTs [40,41] and with consideration of the study objectives, timeline, budget, and available tablets.

Interventions

The ACDT training system was co-designed by the research team and older adults with hearing loss, in collaboration with service providers in Hong Kong, and is patented (HK30110757) [42]. The 12-week training intervention consisted of five 60-minute sessions per week [42] and used a sequential dual-task training approach [33]. To enhance engagement, the system incorporates sound effects and interactive gamification elements, including likes, praise, fireworks, and leaderboards (Figure 1). In Figure 1, screenshot 1 illustrates the daily training pathway (only one module is activated per day). Screenshots 2 - 4 show a listening task with background noise, in which the participant listens to a sentence spoken by a male voice, repeats it, and checks whether all words were correctly heard. Screenshots 5 - 8 demonstrate different cognitive training tasks, including indicating a time using hand gestures, identifying locations, counting numbers, and reading Chinese idioms forward and backward. Screenshot 9 shows the scoring system, and Screenshot 10 presents the ranking list. It also offers participants flexibility for individual training while fostering a sense of virtual group participation through the leaderboard.

XSL•FO
RenderX

Figure 1. Screenshot examples of the web-based gamified auditory-cognitive dual-task training (ACDT) system developed by the team.



Each session begins with a 5-minute volume adjustment to establish a comfortable listening level, followed by attention or orientation training exercises where participants answer questions related to person (identifying their name), time (date, month, or year), and place (location at the time of this writing). This was followed by 25 minutes of auditory training based on culturally adapted exercises from the Listening and Communication Enhancement program [43]. The auditory training includes speech-in-noise perception, rapid speech comprehension, and competing speaker discrimination, all contextualized within real-life scenarios such as medical visits or dim sum gatherings. Difficulty hearing a speaker in noisy environments is a common and frustrating consequence of hearing loss. In the training, participants were trained to listen to a medium-length sentence presented in background noise, repeat it aloud, and then see the sentence in writing to confirm whether they understood every word. The difficulty levels of the sentence and background noise are automatically adjusted

XSL•FO
RenderX

based on the participant's performance on the previous sentence. This training enhances the brain's auditory processing by helping participants filter out background noise and focus on the target voice. Understanding rapid speech is another common challenge for older adults with hearing loss. In this training, the system automatically adapts the speed of sentences to help participants focus on fast talkers, aiding in comprehension. Additionally, ARHL makes listening difficult when multiple speakers are talking. During the training, participants listen to one of 3 voices (male, female, or child) as the target voice, while a competing voice is presented at varying volumes to create a challenge. This training teaches the participant's brain to ignore a competing speaker's voice and focus on the conversation they are engaged in. The auditory training was followed by 25 minutes of cognitive training delivered in a game-based format. These activities target executive function (forward/backward serial counting), perceptual-motor ability (arm-clock positioning), memory (forward/backward repetition of numbers/words), and complex attention (forward/backward spelling of Chinese idioms/poetry) using culturally relevant materials. Sessions concluded with a 5-minute wrap-up featuring communication strategy review ("learning classroom") and recall of visuospatial and memory tasks.

Waitlist Control

The waitlist control group received no intervention during the study period but was given access to the ACDT after T2 data collection.

Randomization

Stratified randomization based on key sociodemographic factors (eg, age, gender, and education level) was used to ensure balanced group allocation. An independent statistician assigned participants to either the ACDT or the waitlist control group in a 1:1 ratio using R software (R Core Team). Group assignments were placed in sealed opaque envelopes, and participants were informed of their allocation after completing all baseline assessments.

Study Procedures

After obtaining informed consent, a blinded researcher collected demographic data and administered outcome measures at baseline (T0), week 6 (T1), and week 12 immediately postintervention (T2). The week-6 assessment was included to identify potential early intervention effects, monitor participant attrition patterns, and verify adherence to the intervention dosage. Prior to intervention commencement, participants attended briefing sessions on system operation and received personal accounts with instructional videos. Technical support was available via a hotline.

All the ACDT participants were invited to take part in an individual interview at T2 following a semistructured interview guide (Textbox 1). The interviews were informed by the theoretical framework of acceptability (TFA) [44], which includes 7 constructs, including affective attitude, burden, perceived effectiveness, ethicality, intervention coherence, opportunity costs, and self-efficacy [44]. The TFA was used in the present study to gain a comprehensive understanding of participants' experiences with the ACDT, as well as their perceptions of and acceptability toward the intervention and study design. Interviews were held at community centers or in a university lab and lasted around 45 - 60 minutes. All the interviews were audio recorded to facilitate data analysis. Interview data were reviewed regularly by the study team until no new topics emerged, indicating that data saturation had been reached [45].

Textbox 1. Semistructured interview guide.

Example of questions

- How was your experience of participating in this study?
- Please tell me about your overall experience of attending the auditory-cognitive dual-task training (ACDT) sessions.
- Have you noted any changes in your health/daily life after participating in the ACDT? If so, do you think these changes were related to the training?
- Which ACDT task did you find most helpful, and which did you find least helpful? Please explain your reasons. What made them useful or not useful? How was your experience with the auditory training in this program?
- How was your experience with the cognitive training in this program?
- How was your experience performing the auditory and cognitive trainings simultaneously?
- What difficulties or problems did you experience when using the program?
- Would you like to continue this training at home in the future?
- What did the instructors or facilitators do that was most/least helpful?
- How do you think about the group-based setup of this ACDT program?
- Do you have any suggestions for improving the ACDT program?

XSL•FO
RenderX

Outcome Measures

Primary Outcome

The primary outcome of global cognition was assessed using a neuropsychological battery comprising (1) the Chinese Auditory Verbal Learning Test (AVLT), which assesses multiple cognitive domains including immediate recall, learning efficiency, delayed recall, retention rate, and recognition memory, and has demonstrated high test-retest reliability (intraclass correlation coefficient [ICC]=0.79 - 0.82) and 78% sensitivity; (2) the Chinese Trial Making Test Parts A and B, which measures visual attention, processing speed, and executive function, and has shown strong discriminant validity (Cohen *d*=1.2 for vascular dementia) [46]; and (3) the Hong Kong-Montreal Cognitive Assessment (HK-MoCA), a 30-point cognitive screening tool covering attention, naming, memory, language, visuospatial skills, abstraction, and orientation, with reported sensitivity of 89% and reliability (ICC=0.89 - 0.92) [47]. Alternating versions of HK-MoCA were used to minimize practice effects.

Secondary Outcomes

Secondary outcomes included: (1) hearing, measured by PTA and the Chinese version of the Hearing Handicap Inventory for the Elderly, which has demonstrated strong internal consistency (α=.91) and convergent validity with PTA (*r*=0.74) [48]; (2) social engagement assessed via the validated 6-item Chinese version of the Lubben Social Network Scale, with high test-retest reliability (ICC=0.85) [49]; and (3) loneliness measured by the 6-item Chinese De Jong Gierveld loneliness scale, demonstrating good internal consistency (α=.80) and discriminative validity for social isolation (area under the curve [AUC]=0.79) [50].

Covariates encompassed age, gender, baseline hearing and cognitive function, duration of hearing loss, education level, health status, and psychological well-being.

Feasibility and Acceptability

The feasibility and acceptability of the ACDT were assessed based on qualitative interviews and comprehensive logs documenting recruitment, retention, adherence, participant contacts, implementation challenges, withdrawal reasons, and completion factors (Multimedia Appendix 2). The intervention completion rate was defined as the extent to which participants completed the 60 training modules during the 12-week intervention period. System-generated data (training dates, progress, and performance metrics) were downloaded after the intervention.

Statistical Analysis

Data analyses and plotting were performed using SPSS Statistics (version 29.0.2.0; IBM Corp) and Python (version 3.11; Python Software Foundation). A 2-sided *P*<.05 was set as the predetermined threshold for statistical significance. Missing data (<7% across variables) were imputed using group means to support analyses requiring complete paired observations and to avoid excluding participants due to missing covariates. A natural log transformation was applied to TMT scores at all time points to address skewness. Baseline group equivalence was confirmed via independent *t* tests (continuous variables) and chi-square tests (categorical variables). The primary analysis focused on within-group changes from baseline to T2. Normality of the distribution of outcome change scores was evaluated using the Shapiro–Wilk test. Based on the results, within-group changes for normally distributed variables (*P*≥.05) were analyzed using paired *t* tests, with Cohen *d* calculated to measure effect size. For nonnormally distributed variables (*P*<.05), the nonparametric Wilcoxon signed-rank test was used.

Effect size *r* for the Wilcoxon signed-rank tests was calculated based on the standardized test statistic (*Z*) and the total number of observations. Effect sizes were interpreted using established criteria, with thresholds of 0.1, 0.3, and 0.5 indicating small, medium, and large effects, respectively. Between-group differences across time points (baseline, wk-6, and wk-12) were evaluated using linear mixed effects models. Each model included time (categorical, 3 levels), group, and the time×group interaction as fixed effects, while adjusting for the baseline value of the respective outcome as a covariate. Within-participant correlations across time points were modeled using a first-order autoregressive residual covariance structure. Models were fitted using restricted maximum likelihood. Pairwise comparisons of estimated marginal means for time×group interactions were conducted using the least significant difference adjustment. Given the explorative nature of the study, no adjustment for multiplicity was applied to the multiple comparisons of global cognition scores.

Qualitative data were analyzed using the hybrid thematic analysis approach described by Braun and Clarke [51], combining inductive coding with deductive theme organization guided by the TFA [44]. Two researchers (IYZ and HM) independently analyzed data in Chinese to preserve the original meaning conveyed by participants. Initial codes were generated directly from the transcripts and then mapped onto TFA domains using NVivo (version 14; Lumivero). Codes that did not align with existing TFA domains were used to inform theme refinement or adaptation. All codes and preliminary themes were documented in a spreadsheet, translated into English, and then back-translated into Chinese by HM and verified by IYZ. Any discrepancies were solved with input from the study team.

Results

Participants

A total of 60 participants were randomized. Baseline characteristics were comparable between groups regarding age, gender, education, employment status, marital status, living arrangements, and health status (Table 1). The mean age of the participants was 67.65 years (SD 4.78). Out of 60 participants, the cohort included 45 (75%) men and 15 (25%) women. Most participants (47/60, 78%) were retired and had attained an upper-secondary education or above. Seventy percent were married, while the majority (52/60, 87%) lived with a spouse, family members, or friends. Regarding health status, 42% (25/60) self-reported their health as good to very good, while 52% (31/60) reported their health as normal.

XSL•FO
RenderX

Table . Demographic characteristics of the participants.

Outcomes	Total (n=60)	Intervention (n=30)	Control (n=30)	Test	*P* value
Age (year), mean (SD)	67.65 (4.78)	67.73 (4.38)	67.57 (5.15)	*t* test[a]	.90
Sex, n (%)				χ^2[b]	.60
Female	15 (25)	6 (20)	9 (30)		
Male	45 (75)	24 (80)	21 (70)		
Education, n (%)				Fisher exact test	.90
Primary school	2 (3)	1 (3)	1 (3)		
Lower-secondary	9 (15)	3 (10)	6 (20)		
Upper-secondary	26 (43)	14 (47)	12 (40)		
Bachelor's or equivalent	17 (28)	9 (30)	8 (27)		
Master's or equivalent	6 (10)	3 (10)	3 (10)		
Employment, n (%)				Fisher exact test	.08
Retired	47 (78)	24 (80)	23 (77)		
Full-time employed	5 (8)	0 (0)	5 (17)		
Part-time employed	4 (7)	3 (10)	1 (3)		
Seeking work	1 (2)	1 (3)	0 (0)		
Housewife	3 (5)	2 (7)	1 (3)		
Marital status, n (%)				Fisher exact test	.22
Married	42 (70)	23 (77)	19 (63)		
Single	14 (23)	4 (13)	10 (33)		
Divorced	1 (2)	1 (3)	0 (0)		
Widowed	3 (5)	2 (7)	1 (3)		
Living arrangement, n (%)				Fisher exact test	.60
Live alone	8 (13)	4 (13)	4 (13)		
With spouse	24 (40)	14 (47)	10 (33)		
With family	27 (45)	12 (40)	15 (50)		
With friend	1 (2)	0 (0)	1 (3)		
Health status, n (%)				Fisher exact test	.57
Not good	4 (7)	3 (10)	1 (3)		
Normal	31 (52)	14 (47)	17 (57)		
Good	24 (40)	12 (40)	12 (40)		
Very good	1 (2)	1 (3)	0 (0)		

[a]Continuous variable: independent-samples *t* test,

[b]Categorical variables: chi-square test or Fisher exact test.

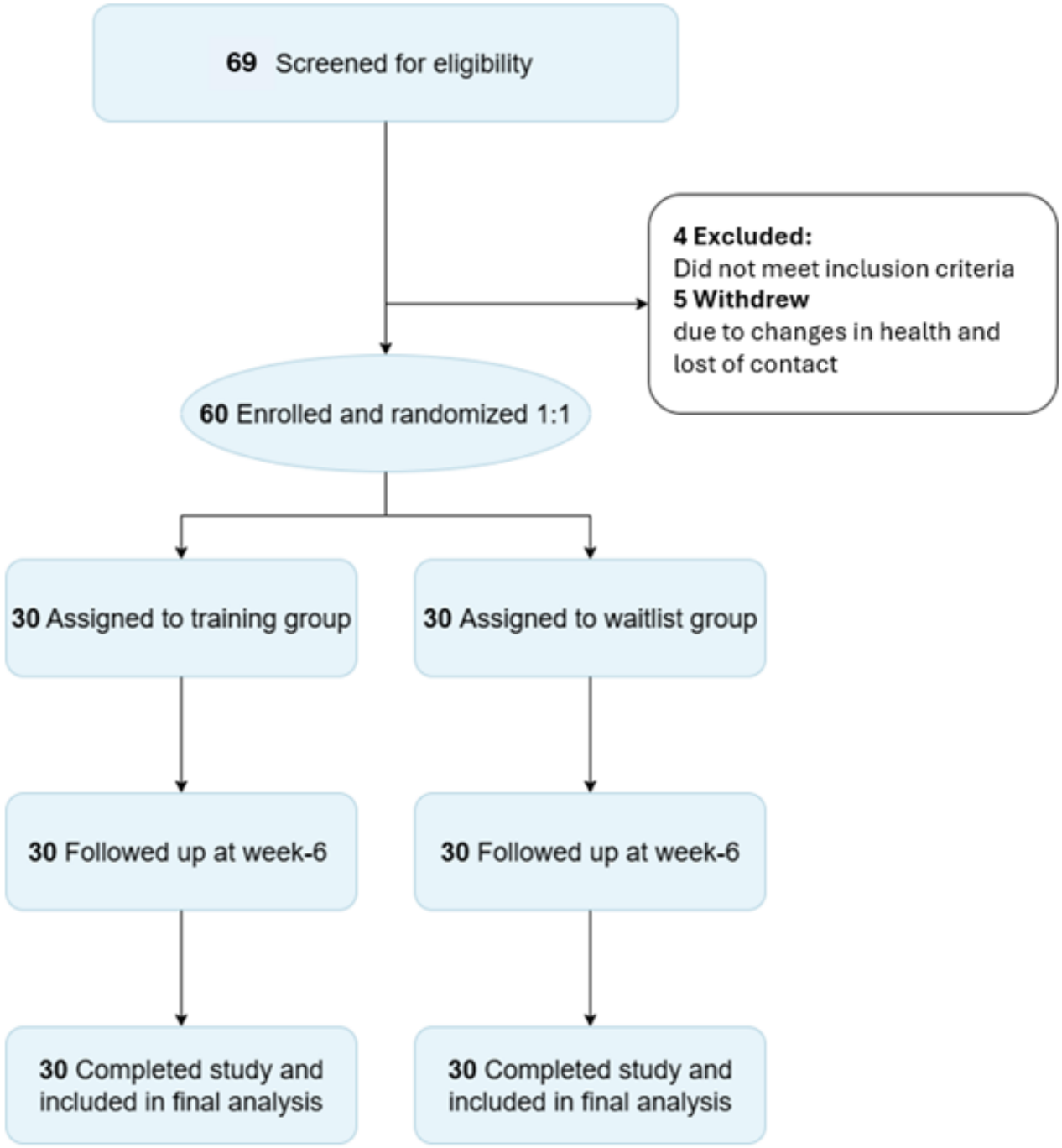
Feasibility of the Intervention

The trial proceeded according to the protocol without early termination. There were 69 candidates screened for eligibility; 4 were excluded for not meeting the eligibility criteria, and 5 withdrew prior to randomization due to health issues or loss of contact (Figure 2). Of the 60 participants who met the eligibility criteria, 30 were randomized to the ACDT group, and 24 of them completed all 60 training modules, while 6 participants completed 55 modules. The overall intervention completion rate was nearly 98% (1770/1800). At the beginning of the study, our hotline received a small number of calls regarding login issues, primarily from participants who had forgotten their passwords or were unsure how to reset them. This issue was resolved after we improved the study briefing sessions. No

XSL•FO
RenderX

adverse events were reported in either the intervention or control group.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) diagram.



Quantitative Results

Within-group analyses showed significant decreases in TMT completion times over 12 weeks, reflecting improved performance in focused attention (intervention: mean change=–0.15; *P*=.02; *d*=–0.46; control: mean change=–0.10; *P*=.027; *d*=–0.46), and divided attention (intervention: mean change=–0.21; *P*=.002; *d*=–0.63; control: mean change=–0.11; *P*=.02; *d*=–0.45). The intervention group demonstrated moderately greater improvements than the control group (Table 2). Both groups reported a significant decrease in overall hearing handicap. No statistically significant within-group changes were observed in HK-MoCA total scores or in social relationship and loneliness scores in either group.

XSL•FO
RenderX

Table . Results of a paired *t* test.

Variable and group	Mean change	*P* value	Effect size (Cohen *d*)
Focused attention (TMT-A)[a]			
Intervention	–0.15	.02	–0.46
Control	–0.10	.02	–0.46
Divided attention (TMT-B)[b]			
Intervention	–0.21	.002	–0.63
Control	–0.11	.02	–0.45
HK-MoCA total score[c]			
Intervention	0.40	.30	0.19
Control	–0.13	.74	–0.06
Hearing handicap (HHIE), overall[d]			
Intervention	–2.43	.05	–0.37
Control	–3.83	.005	–0.56
Hearing handicap (HHIE), social[d]			
Intervention	–1.79	.34	–0.42
Control	–2.50	.14	–0.65
Social engagement (Lubben Social Network Scale)			
Intervention	0.60	.43	0.15
Control	1.00	.11	0.30
PTA (left ear)[e]			
Intervention	–1.64	.16	–0.26
Control	–0.42	.63	–0.09
PTA (right ear)[e]			
Intervention	–1.39	.26	–0.21
Control	–1.79	.08	–0.33

[a]TMT-A: Trial Making Test Part A.

[b]TMT-B: Trial Making Test Part B,

[c]HK-MoCA: Hong Kong-Montreal Cognitive Assessment.

[d]HHIE: Hearing Handicap Inventory for the Elderly.

[e]PTA: pure-tone average.

For outcome variables that violated the normality assumptions, Wilcoxon signed-rank tests were conducted separately within each group (Table 3). In the intervention group, significant cognitive improvements were observed in the HK-MoCA subdomains of naming (Z=2.00; P=.05; r=0.37) and visual cognition (Z=2.39; P=.02; r=0.44), while the control group showed no significant within-group changes in these areas. Both groups reported significant decreases in emotional hearing handicap; however, the intervention group demonstrated a larger effect size (Z=2.15; P=.03; r=0.39) than the control group (Z=2.03; P=.04; r=0.37). Similarly, both groups demonstrated significant improvements in 5-min delayed recall on the AVLT, with the intervention group exhibiting greater improvement (Z=4.16; P<.001; r=0.76) than the control group (Z=4.11; P<.001; r=0.75).

XSL•FO
RenderX

Table . Results of Wilcoxon signed-rank tests.

Outcomes and group	Z	*P* value	Effect size (*R*)	Magnitude
Naming (HK-MoCA)[a]				
Intervention	2.00	.05	0.37	Medium
Control	1.63	.10	0.30	Small
Visual cognition (HK-MoCA)				
Intervention	2.39	.02	0.44	Medium
Control	1.15	.25	0.21	Small
Hearing handicap (emotional), HHIE[b]				
Intervention	2.15	.03	0.39	Medium
Control	2.03	.04	0.37	Medium
5-min delayed recall (AVLT[c])				
Intervention	4.16	<.001	0.76	Large
Control	4.11	<.001	0.75	Large
Recognition (AVLT)[c]				
Intervention	1.81	.07	0.33	Medium
Control	2.08	.04	0.38	Medium
Language (HK-MoCA)				
Intervention	N/A[d]	N/A	N/A	N/A
Control	1.34	.18	0.24	Small
Memory (HK-MoCA)				
Intervention	0.99	>.99	0.18	Small
Control	1.28	.41	0.23	Small
Orientation (HK-MoCA)				
Intervention	1.63	.10	0.30	Small
Control	0.00	>.99	0.00	Trivial
Abstract (HK-MoCA)				
Intervention	0.92	.36	0.17	Small
Control	0.28	.78	0.05	Trivial
Attention (HK-MoCA)				
Intervention	0.83	.41	0.15	Small
Control	0.55	.58	0.10	Small
Loneliness (single question)				
Intervention	0.00	>.99	0.00	Trivial
Control	0.82	.41	0.15	Small
Emotional loneliness				
Intervention	0.68	.50	0.12	Small
Control	0.00	>.99	0.00	Trivial
Social loneliness				
Intervention	0.13	.89	0.02	Trivial
Control	1.22	.22	0.22	Small
Overall loneliness				
Intervention	0.51	.61	0.09	Trivial

XSL•FO
RenderX

Outcomes and group	Z	*P* value	Effect size (*R*)	Magnitude
Control	0.90	.31	0.16	Small

[a]HK-MoCA: Hong Kong-Montreal Cognitive Assessment.

[b]HHIE: Hearing Handicap Inventory for the Elderly.

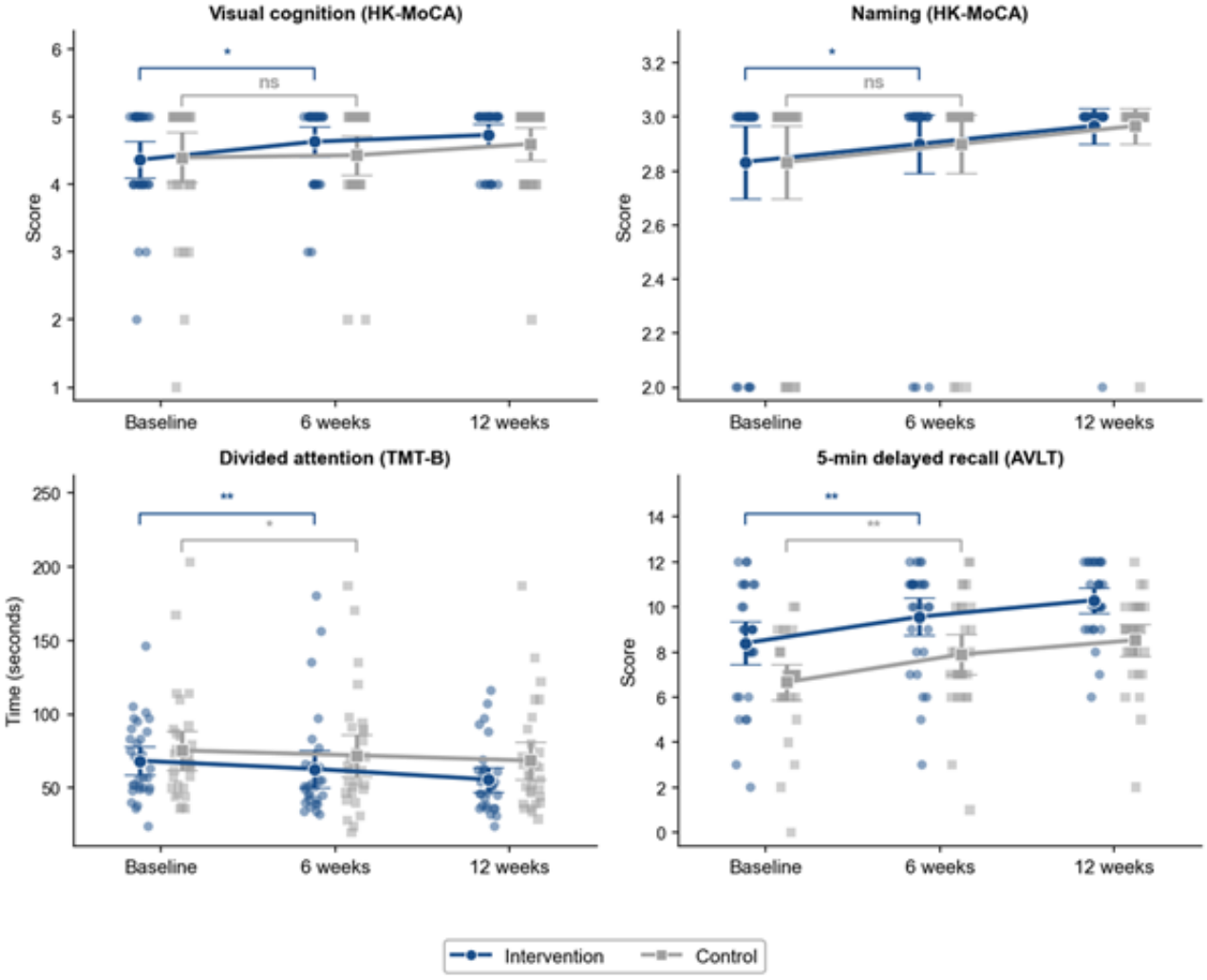
[c]AVLT: Chinese Auditory Verbal Learning Test.

[d]N/A: not applicable due to zero variance in the difference scores.

Changes in visual cognition, naming, divided attention, and 5-minute delayed recall at baseline, 6 weeks, and 12 weeks are shown in Figure 3 for the intervention group (blue circles) and the control group (gray circles). Each data point represents an individual participant's score, with a horizontal offset added for clarity. Lines connect the group means at each time point, and error bars indicate 95% CIs. The intervention group showed significantly greater improvement than the control group in these functions.

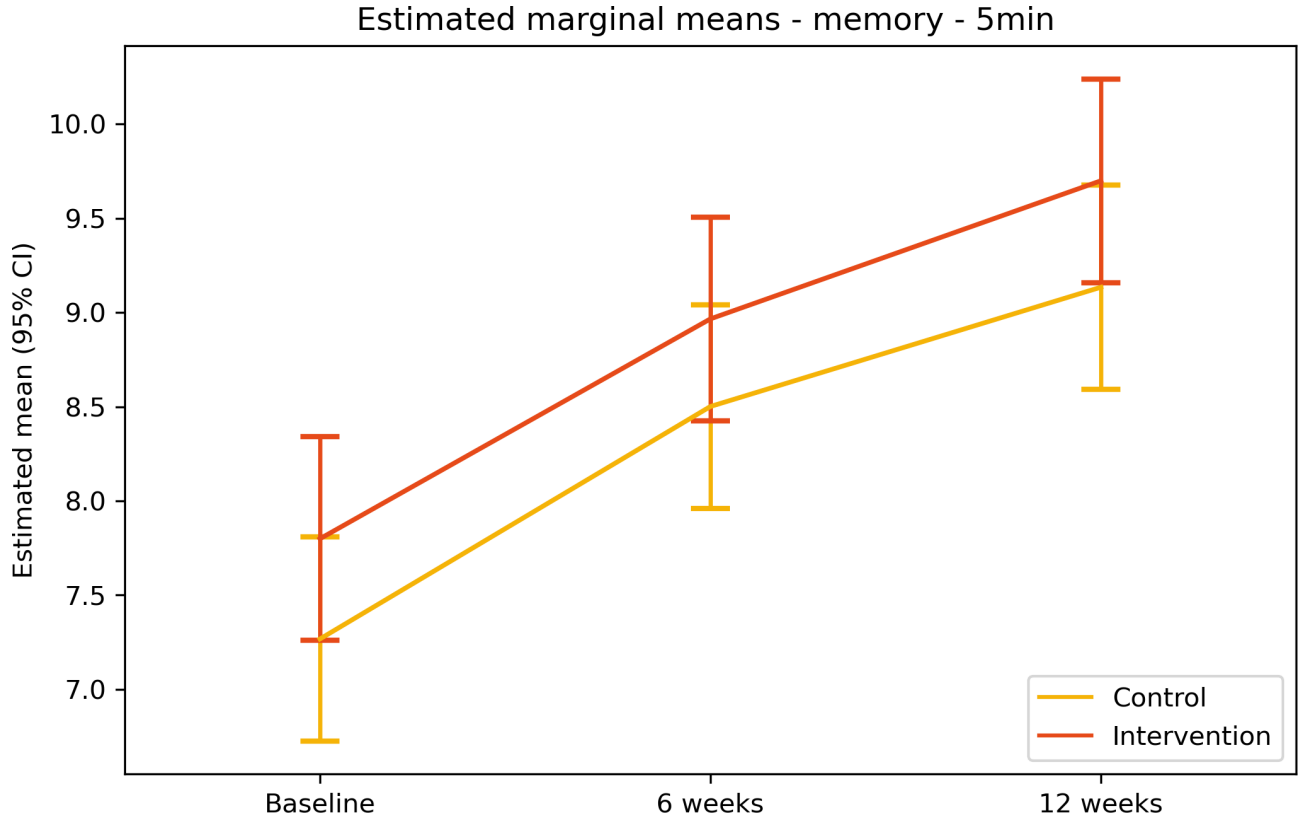
According to the linear mixed effects model analysis, a small to moderate group effect was observed for 5-minute delayed recall on the AVLT (Cohen d=0.38), with the fixed effects explaining 69% of the variance (marginal R^2=0.69). Baseline scores strongly predicted follow-up performance (standardized β=0.78). A significant main effect of time was identified ($F_{2, 129.5}$=25.70; P<.001), along with a significant group difference ($F_{1, 81.2}$=3.997; P=.05), indicating that the intervention group showed better memory scores across visits. The time×group interaction was nonsignificant ($F_{2, 129.5}$=0.022; P=.98; Figure 4).

Figure 3. Trajectories of visual cognition, naming, divided attention, and 5-min delayed recall across the assessment time points. Significance of within-group change from baseline to 6 weeks is indicated above each group's trajectory. AVLT: Auditory Verbal Learning Test; HK-MoCa: Hong Kong-Montreal Cognitive Assessment; TMT-B: Trial Making Test Part. *P<.05, **P<.01, and ***P<.001.



XSL•FO
RenderX

Figure 4. Estimated marginal means of 5-min delayed recall on the auditory verbal learning test (AVLT) for both intervention and control groups from baseline to 6 weeks and 12 weeks.



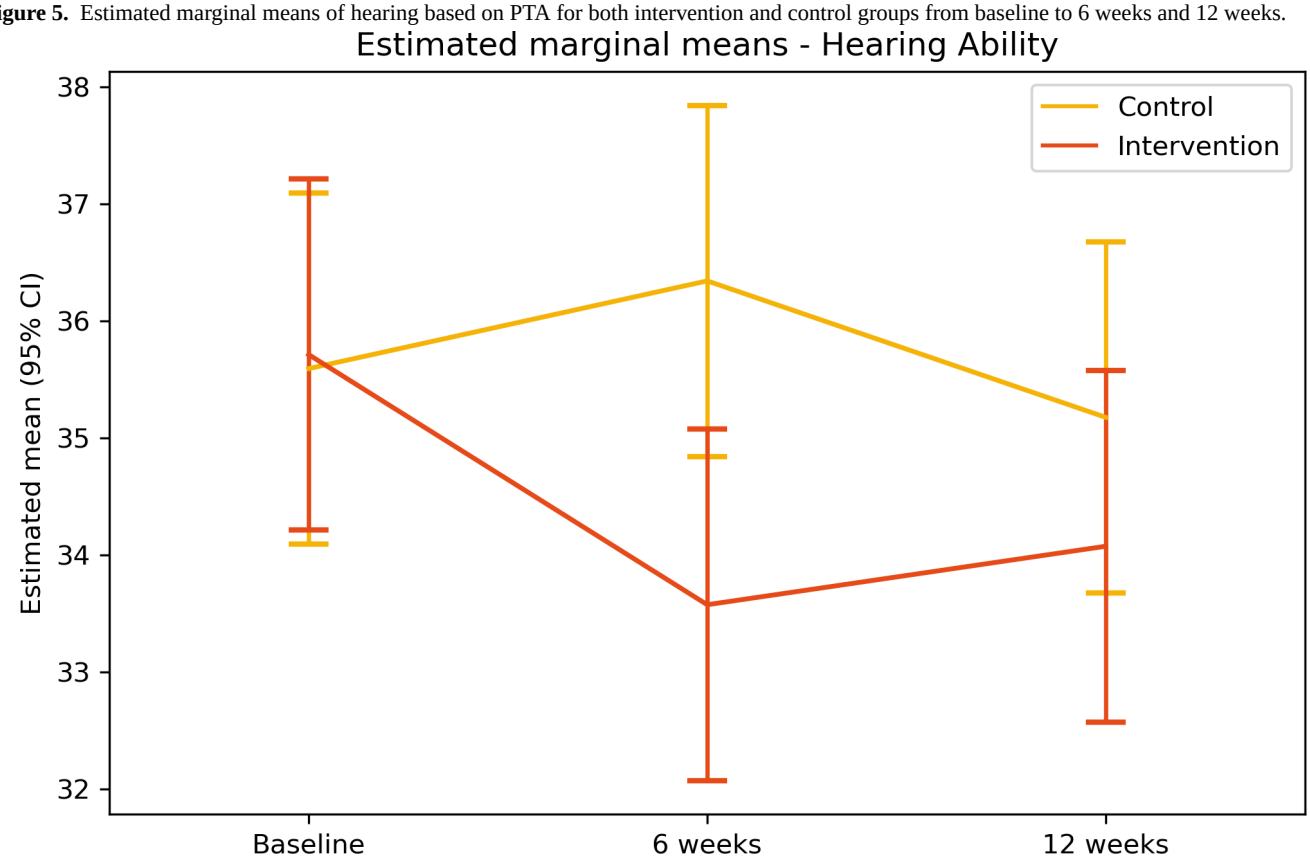
Cognitive performance was similar between the groups at baseline. Post hoc pairwise comparisons showed that cognitive functions on HK-MoCA total scores in the intervention group increased significantly from baseline to week 6 (P=.03) and were maintained at week 12. In contrast, the control group showed a significant decline from week 6 to week 12 (P=.02) with no net change from baseline.

For hearing thresholds (Figure 5), the standardized baseline coefficient was strong (β=0.88), but the group effect remained minimal (Cohen d=0.17), with a marginal R^2 of 0.67. A significant time×group interaction was observed for left-ear thresholds ($F_{2,\ 117.0}$=4.470; P=.01), indicating improved thresholds in the intervention group.

The Hearing Handicap Inventory for the Elderly total score showed a significant main effect of time ($F_{2,\ 112.6}$=14.69; P<.001), with both groups reporting reduced perceived hearing handicap at weeks 6 and 12 compared with baseline. A significant group main effect was observed for the social subscale ($F_{1,\ 62.1}$=4.81; P=.03), indicating that, on average, the intervention group reported lower perceived social hearing handicap than the control group, independent of time. There was no significant time×group interaction for the social subscale (P=.24).

XSL•FO
RenderX

Figure 5. Estimated marginal means of hearing based on PTA for both intervention and control groups from baseline to 6 weeks and 12 weeks.



Qualitative Results

A total of 30 participants in the ACDT group were interviewed, and 3 key themes were generated based on the TFA [44]. No ethical issues were raised by participants. Opportunity costs refer to the trade-offs users may have to make when engaging with the intervention. None of them explicitly expressed such concerns. Instead, most rearranged daily activities to accommodate the training.

Theme 1: Intervention Coherence and Participants' Affective Attitude Toward ACDT

Intervention coherence refers to the extent to which users understand the intervention, its purpose, and how it fits into their existing routines [44]. Participants generally valued the program's alignment with their cognitive health goals. Daily routines involving auditory-cognitive training were embraced as brain-stimulating activities, with Participant 16 emphasizing, "I value daily hearing and cognition training. The system's varied exercises provide consistent training." Participant 38 similarly stated, "I train seriously to improve memory and hearing." Some participants reported feeling nervous and under pressure when they first began the training, mainly because they were unfamiliar with the content and format. As Participant 35 mentioned, "At first, I was nervous because I had never done this before, but after a few sessions, I became more relaxed as I got used to the process." Over time, most participants integrated ACDT into their daily routines, with Participant 5 reporting, "I do it every morning after I get up; it has become like homework," and Participant 60 adding, "I attend the training every day even when traveling."

Affective attitude focuses on how users feel about using the intervention [44]. Participants generally reported neutral to positive attitudes toward ACDT. Most described their training experience as "acceptable," "quite interesting," and "enjoyable" (Participants 18, 22, and 43). However, some participants felt that certain training components were insufficiently engaging. Participant 13 specifically requested more personalized content with interactive features, noting,

> *For hearing training, I hope the content could be more interesting and personalized. I'm not interested in public transportation. The training should be more interactive, and I'd like immediate feedback.* [Participant 13]

This preference for ecological relevance was further reinforced by reports that tasks involving familiar daily scenarios were easier and more interesting than those involving abstract content. As Participant 14 explained, "Daily activities like going to restaurants or supermarkets are more interesting and easier to remember."

Theme 2: Perceived Benefits in Cognition, Information Acquisition, and Self-Awareness From ACDT

Most participants perceived that ACDT enhanced their abilities, particularly when they observed their own progress over time. Participant 60 described this as a source of motivation:

> *I think I do it happily as it is helpful to me. It turns out that when we see ourselves making progress, we can also challenge the difficulties.* [Participant 60]

Several participants reported improvements in attention and concentration during auditory tasks.

XSL•FO
RenderX

> *I notice the issue of concentration. When I want to learn something, I will know that I need to be more attentive and calm down to listen. This is the influence of the training on me.* [Participant 35]

Others described improvements in memory, mental clarity, and processing speed. Participant 5 observed, "I feel that my memory has improved because when doing math by heart, I often have to calculate the subtrahend and the addend."

Participants also found ACDT informative when integrating real-life scenarios. Participant 4 stated, "Yes, in terms of hearing training, the content can help us. There is some information about healthy or daily life, like when you eat dim sum, or you should not eat fried foods," while Participant 56 similarly described it as "very informative."

Beyond functional improvements, participants demonstrated greater awareness of their hearing and attentional limitations, prompting adaptive changes in communication behaviors. As Participant 18 reflected,

> *Sometimes it is not just about relying on hearing. I begin to realize that many times I cannot do certain things because I suddenly think of something else. So in daily life, I start to understand where my problems lie. [Participant 18]*

Sustainability concerns emerged, with Participant 56 expressing her concern about "maintaining gains after completing of this study."

Theme 3: Perceiving ACDT as Less Burdensome With Enhanced Self-Efficacy

Most participants perceived the training as manageable, with several attributing this to prior life experience or visible self-improvement. As Participant 17 noted, "I don't find it difficult as my previous job was very demanding," while Participant 18 connected manageability to perceived benefits, "I feel no burden because it's a training, where I can manage my own time to complete it." Completing sessions often fostered a sense of accomplishment, as illustrated by Participant 34's comment, "I feel confident after finishing tasks." Participants also described progressive mastery of the auditory exercises, with several individuals (34, 42, and 57) noting "improved processing speed and accuracy." Elements like competing voices were perceived as "both challenging and engaging" (Participants 5 and 34). While task difficulty occasionally prompted frustration, as Participant 55 acknowledged, "There is a sense of success when I answer correctly, but difficult tasks can be frustrating." However, they maintained motivation through perceived utility. Participant 49 articulated this resilience, "The hardest parts are likely most useful; backward counting's challenge makes it interesting."

Discussion

Primary Cognitive Outcomes

This pilot RCT provides preliminary evidence for the feasible, acceptable, and potential efficacy of ACDT for older adults with ARHL. Participants in the ACDT group demonstrated significant improvements in focused attention, divided attention, naming, and visual cognition. However, the control group's improvement in some cognitive measures suggests possible practice effects from repeated testing, highlighting the need for active control conditions in future trials. These findings extend evidence that combined auditory-cognitive training selectively enhances executive functions and visual processing, which are critical domains in dementia risk reduction [52]. This pattern converged with qualitative findings of "enhanced attention and concentration," "improved memory and mental clarity," and "increased processing speed" during daily activities. Mechanistically, results align with previous research that auditory training enhances attention in older adults with hearing loss [53], cognitive training improves domain-specific functions, and dual-task trainings strengthen executive control [34]. The gamified design likely amplified these effects through enhanced engagement and motivation, which is supported by participant feedback and studies showing gamification improves adherence and training outcomes in cognitive interventions [54]. Additionally, further research is warranted to examine the effects of ACDT on dual-task listening-cognitive abilities using an active control group, as poor dual-task performance is recognized as an early indicator of dementia [30].

Notably, the ACDT group exhibited improved 5-minute delayed recall with small-to-moderate group effects and maintained higher average scores across assessments than the control group. This finding is clinically significant given delayed recall is an established early marker for cognitive decline and dementia risk [55]. Furthermore, post hoc pairwise comparisons demonstrated that HK-MoCA total scores significantly increased from baseline to week 6 and were sustained at week 12 in the ACDT group, suggesting that ACDT may help stabilize global cognition over time.

Auditory Threshold Findings and Possible Indicators of Training-Induced Neuroplasticity

Hearing threshold analysis revealed minimal overall between-group differences, which means that ACDT did not produce a significant improvement in PTA compared to the control group, and this finding is consistent with the recognized stability of peripheral hearing loss over short time intervals [56]. While the significant time×group interaction observed for left-ear thresholds may indicate training-induced neuroplasticity in central auditory pathways [57], these preliminary findings warrant confirmation in larger trials. Future studies should investigate whether specific training components can be optimized for bilateral benefits and determine if observed neuroplastic changes translate to measurable functional communication improvements.

Perceived Hearing Handicap and Increased Awareness

Interestingly, the control group reported lower perceived overall hearing handicap but higher emotional hearing handicap than the ACDT group. One possible explanation is that the ACDT participants became more aware of their hearing challenges, leading to more accurate self-reporting. In contrast, the control group may have had lower insight into their hearing limitations, potentially resulting in underreporting. This interpretation aligns with qualitative findings and researcher observations during data collection, wherein ACDT participants frequently described

XSL•FO
RenderX

increased awareness of hearing-related challenges in their daily lives. Previous research also supports that targeted interventions can improve auditory awareness, which may initially increase self-reported handicap but ultimately facilitate adaptive behavior change [58].

In our study, participants reported enhanced recognition of auditory and attentional limitations, which subsequently prompted adaptive modifications in communication behaviors. Many participants described integrating training into their daily routines and adopting compensatory strategies to improve attention and listening engagement, which may reflect a meaningful sign of behavior change. However, longitudinal investigation is required to verify sustainability. ACDT also fostered perceived accomplishment and strengthened self-efficacy, potentially encouraging sustained engagement in beneficial behavior changes. Collectively, these findings position ACDT as a catalyst for positive behavioral adaptations, benefiting not only participants but also their families.

In this study, no statistically significant changes were observed in social relationship or loneliness scores for either group. A possible explanation is that participants were comparatively young, most were married and living with a spouse or other family members, and they demonstrated a positive attitude toward adopting this new training to address their hearing problems. Future studies should test ACDT among individuals at higher risk of loneliness and social isolation.

Enhancing Training Content and Interactivity Using Artificial Intelligence

While the intervention was generally well-received, participant feedback identified opportunities for enhancement through more personalized and interactive training content. Integration of artificial intelligence (AI) can dynamically tailor training modules to individual abilities and progression patterns, with real-time adjustments in difficulty and personalized feedback to optimize training efficacy while maintaining engagement. Machine learning algorithms could further enhance ecological validity by generating complex auditory environments that simulate real-world scenarios, improving training transferability to daily communication challenges. Incorporating AI-powered analytics could continuously monitor user performance trajectories, identify areas for improvement, and deliver targeted interventions to maximize auditory and cognitive gains. Future iterations of ACDT should prioritize the development of AI-enhanced features, including personalized content, interactive feedback, and real-time progress tracking. Engaging users and their families in the co-design process is essential to ensure that the platform meets their needs and preferences.

Strengths and Limitations

This study demonstrates several strengths, including (1) multimodal assessment of auditory and cognitive outcomes, (2) ecologically valid gamified design that could enhance participant engagement, and (3) novel documentation of self-awareness as a facilitator for adaptive behavior change. However, several limitations warrant consideration: (1) the modest sample size and absence of longitudinal follow-up limit detection of subtle neuroplasticity effects, such as the observed hearing threshold trend; (2) the waitlist control design cannot exclude attention bias and isolate the specific efficacy of dual-task training; and (3) although ARHL has been reported to be more common, more severe, and to have an earlier onset in men than in women [59], the homogeneous participant characteristics regarding education level and hearing loss severity, along with a predominance of male participants, may restrict generalizability of our findings to broader ARHL populations.

Conclusions

This pilot RCT suggests that ACDT is feasible and acceptable for older adults with ARHL, with preliminary evidence supporting its potential efficacy. The intervention appeared to benefit both auditory and cognitive domains, as well as self-awareness of hearing loss, which may facilitate adaptive behavioral adjustments. Future development could explore AI-enhanced personalization and interactivity to improve ecological relevance. Further trial with larger, more diverse samples, active control conditions, and longitudinal follow-up would help clarify ACDT's sustained effects on auditory-cognitive health, dual-task listening-cognitive abilities, and real-world functioning.

Acknowledgments

We acknowledge the financial support from the School of Nursing and the Faculty of Health and Social Sciences, The Hong Kong Polytechnic University (PolyU). We also thank the Christian Family Service Centre (CFSC) and the PolyU Jockey Club Smart Ageing Hub for their support in participant recruitment.

All authors declare that generative artificial intelligence was used only for language polishing of the manuscript and was not used in any part of the study.

Funding

This research was supported by Start-up Fund for RAPs under the Strategic, The Hong Kong Polytechnic University (PolyU), grant number P0045159, and Faculty Reserve, Faculty of Health and Social Sciences, PolyU, grant number P0048451.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the privacy of participants but are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: IYZ, AYML, CL, LLP, JM, JSG, RS, JHYN, EAC
Data curation: IYZ, HM
Formal analysis: IYZ, HM
Funding acquisition: IYZ
Investigation: IYZ, CL, LLP, JM, JSG, RS, JHYN, EAC
Methodology: IYZ, CL, LLP, JM, JSG, RS, JHYN, EAC
Project administration: IYZ
Resources: IYZ
Software: IYZ
Supervision: IYZ, AYML
Validation: IYZ, CL, LLP, JM, JSG, RS, JHYN, EAC
Visualization: IYZ
Writing – original draft: IYZ, HM
Writing – review & editing: IYZ, AYML, CL, LLP, JM, JSG, RS, JHYN, EAC

Conflicts of Interest

JSG reports Regeneron: 1-time honorarium. Other authors declare no conflicts of interest.

Multimedia Appendix 1
Research protocol.
[DOC File, 581 KB - aging_v9i1e84083_app1.doc]

Multimedia Appendix 2
A log for checking study feasibility.
[PDF File, 440 KB - aging_v9i1e84083_app2.pdf]

Checklist 1
CONSORT checklist.
[PDF File, 147 KB - aging_v9i1e84083_app3.pdf]

References

1. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1211-1259. [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28919117]
2. World report on hearing. World Health Organization. URL: https://www.who.int/publications/i/item/9789240020481 [accessed 2026-06-01]
3. Applebaum J, Hoyer M, Betz J, Lin FR, Goman AM. Long-term subjective loneliness in adults after hearing loss treatment. Int J Audiol 2019 Aug;58(8):464-467. [doi: 10.1080/14992027.2019.1593523] [Medline: 30929531]
4. Shukla A, Harper M, Pedersen E, et al. Hearing loss, loneliness, and social isolation: a systematic review. Otolaryngol Head Neck Surg 2020 May;162(5):622-633. [doi: 10.1177/0194599820910377] [Medline: 32151193]
5. Mick P, Kawachi I, Lin FR. The association between hearing loss and social isolation in older adults. Otolaryngol Head Neck Surg 2014 Mar;150(3):378-384. [doi: 10.1177/0194599813518021] [Medline: 24384545]
6. Uchida Y, Nishita Y, Tange C, et al. The longitudinal impact of hearing impairment on cognition differs according to cognitive domain. Front Aging Neurosci 2016;8:201. [doi: 10.3389/fnagi.2016.00201] [Medline: 27597827]
7. Loughrey DG, Kelly ME, Kelley GA, Brennan S, Lawlor BA. Association of age-related hearing loss with cognitive function, cognitive impairment, and dementia: a systematic review and meta-analysis. JAMA Otolaryngol Head Neck Surg 2018 Feb 1;144(2):115-126. [doi: 10.1001/jamaoto.2017.2513] [Medline: 29222544]
8. A T J, Dias A, Philp I, Beard J, Patel V, Prince M. Identifying common impairments in frail and dependent older people: validation of the COPE assessment for non-specialised health workers in low resource primary health care settings. BMC Geriatr 2015 Oct 14;15(1):123. [doi: 10.1186/s12877-015-0121-1] [Medline: 26467913]
9. Golub JS, Brickman AM, Ciarleglio AJ, Schupf N, Luchsinger JA. Association of subclinical hearing loss with cognitive performance. JAMA Otolaryngol Head Neck Surg 2020 Jan 1;146(1):57-67. [doi: 10.1001/jamaoto.2019.3375] [Medline: 31725853]
10. Irace AL, Armstrong NM, Deal JA, et al. Longitudinal associations of subclinical hearing loss with cognitive decline. J Gerontol A Biol Sci Med Sci 2022 Mar 3;77(3):623-631. [doi: 10.1093/gerona/glab263] [Medline: 34516645]

XSL•FO
RenderX

11. Deal JA, Sharrett AR, Albert MS, et al. Hearing impairment and cognitive decline: a pilot study conducted within the atherosclerosis risk in communities neurocognitive study. Am J Epidemiol 2015 May 1;181(9):680-690. [doi: 10.1093/aje/kwu333] [Medline: 25841870]
12. Amieva H, Ouvrard C, Giulioli C, Meillon C, Rullier L, Dartigues JF. Self-reported hearing loss, hearing aids, and cognitive decline in elderly adults: a 25-year study. J Am Geriatr Soc 2015 Oct;63(10):2099-2104. [doi: 10.1111/jgs.13649] [Medline: 26480972]
13. Lin FR, Ferrucci L, Metter EJ, An Y, Zonderman AB, Resnick SM. Hearing loss and cognition in the Baltimore longitudinal study of aging. Neuropsychology 2011 Nov;25(6):763-770. [doi: 10.1037/a0024238] [Medline: 21728425]
14. Saji N, Makizako H, Suzuki H, et al. Hearing impairment is associated with cognitive function in community-dwelling older adults: a cross-sectional study. Arch Gerontol Geriatr 2021;93:104302. [doi: 10.1016/j.archger.2020.104302] [Medline: 33256998]
15. Lin FR, Ferrucci L, An Y, et al. Association of hearing impairment with brain volume changes in older adults. Neuroimage 2014 Apr 15;90:84-92. [doi: 10.1016/j.neuroimage.2013.12.059] [Medline: 24412398]
16. Dupuis K, Pichora-Fuller MK, Chasteen AL, Marchuk V, Singh G, Smith SL. Effects of hearing and vision impairments on the Montreal Cognitive Assessment. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2015;22(4):413-437. [doi: 10.1080/13825585.2014.968084] [Medline: 25325767]
17. Gopinath B, Wang JJ, Schneider J, et al. Depressive symptoms in older adults with hearing impairments: the blue mountains study. J Am Geriatr Soc 2009 Jul;57(7):1306-1308. [doi: 10.1111/j.1532-5415.2009.02317.x] [Medline: 19570163]
18. Wingfield A, Tun PA, McCoy SL. Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. Curr Dir Psychol Sci 2005;14(3):144-148. [doi: 10.1111/j.0963-7214.2005.00356.x]
19. Uchida Y, Sugiura S, Nishita Y, Saji N, Sone M, Ueda H. Age-related hearing loss and cognitive decline - the potential mechanisms linking the two. Auris Nasus Larynx 2019 Feb;46(1):1-9. [doi: 10.1016/j.anl.2018.08.010] [Medline: 30177417]
20. Schneider WJ, McGrew KS. McGrew KS: the cattell-horn-carroll (CHC) model of human cognitive abilities: past, present, and future. In: Flanagan DP, EM M, editors. Contemporary Intellectual Assessment: Theories, Tests, and Issues, 4th edition 2018:73-163 URL: https://not-equal.org/content/pdf/misc/Kaufman2018.pdf [accessed 2026-06-10]
21. Fritze T, Teipel S, Óvári A, Kilimann I, Witt G, Doblhammer G. Hearing impairment affects dementia incidence. An analysis based on longitudinal health claims data in Germany. PLoS ONE 2016;11(7):e0156876. [doi: 10.1371/journal.pone.0156876] [Medline: 27391486]
22. Kiely KM, Gopinath B, Mitchell P, Luszcz M, Anstey KJ. Cognitive, health, and sociodemographic predictors of longitudinal decline in hearing acuity among older adults. J Gerontol A Biol Sci Med Sci 2012 Sep;67(9):997-1003. [doi: 10.1093/gerona/gls066] [Medline: 22415523]
23. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec 16;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6] [Medline: 28735855]
24. Reed NS, Altan A, Deal JA, et al. Trends in health care costs and utilization associated with untreated hearing loss over 10 years. JAMA Otolaryngol Head Neck Surg 2019 Jan 1;145(1):27-34. [doi: 10.1001/jamaoto.2018.2875] [Medline: 30419131]
25. Ye X, Zhu D, He P. The role of self-reported hearing status in the risk of hospitalisation among Chinese middle-aged and older adults. Int J Audiol 2021 Oct;60(10):754-761. [doi: 10.1080/14992027.2021.1871671] [Medline: 33529120]
26. Jorgensen LE, Messersmith JJ. Impact of aging and cognition on hearing assistive technology use. Semin Hear 2015 Aug;36(3):162-174. [doi: 10.1055/s-0035-1555119] [Medline: 27516716]
27. Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N. A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. Hear Res 2013 Jun;300:18-32. [doi: 10.1016/j.heares.2013.03.006] [Medline: 23541911]
28. Burk MH, Humes LE, Amos NE, Strauser LE. Effect of training on word-recognition performance in noise for young normal-hearing and older hearing-impaired listeners. Ear Hear 2006 Jun;27(3):263-278. [doi: 10.1097/01.aud.0000215980.21158.a2] [Medline: 16672795]
29. Lawrence BJ, Jayakody DMP, Henshaw H, et al. Auditory and cognitive training for cognition in adults with hearing loss: a systematic review and meta-analysis. Trends Hear 2018;22:2331216518792096. [doi: 10.1177/2331216518792096] [Medline: 30092719]
30. Henshaw H, Heinrich A, Tittle A, Ferguson M. Cogmed training does not generalize to real-world benefits for adult hearing aid users: results of a blinded, active-controlled randomized trial. Ear Hear 2022;43(3):741-763. [doi: 10.1097/AUD.0000000000001096] [Medline: 34524150]
31. Wilhelm LA. Using contextual and visual cues to improve sung word recognition in hearing aid users. J Music Ther 2020 Dec 8;57(4):379-405. [doi: 10.1093/jmt/thaa009] [Medline: 32574363]
32. Green T, Faulkner A, Rosen S. Computer-based connected-text training of speech-in-noise perception for cochlear implant users. Trends Hear 2019;23:2331216519843878. [doi: 10.1177/2331216519843878] [Medline: 31010386]
33. Wickens CD. Multiple resources and mental workload. Hum Factors 2008 Jun;50(3):449-455. [doi: 10.1518/001872008X288394] [Medline: 18689052]
34. Parial LL, Kor PPK, Sumile EF, Leung AYM. Dual-task zumba gold for improving the cognition of people with mild cognitive impairment: a pilot randomized controlled trial. Gerontologist 2023 Aug 24;63(7):1248-1261. [doi: 10.1093/geront/gnac081] [Medline: 35679826]

XSL•FO
RenderX

35. Corbett A, Owen A, Hampshire A, et al. The effect of an online cognitive training package in healthy older adults: an online randomized controlled trial. J Am Med Dir Assoc 2015 Nov 1;16(11):990-997. [doi: 10.1016/j.jamda.2015.06.014] [Medline: 26543007]
36. Chae HJ, Lee SH. Effectiveness of the online-based comprehensive cognitive training application, smart brain, for community-dwelling older adults with dementia: a randomized controlled trial. Eur J Phys Rehabil Med 2024 Jun;60(3):423-432. [doi: 10.23736/S1973-9087.24.08043-2] [Medline: 38647533]
37. Cheng C, Ebrahimi OV. Gamification: a novel approach to mental health promotion. Curr Psychiatry Rep 2023 Nov;25(11):577-586. [doi: 10.1007/s11920-023-01453-5] [Medline: 37801212]
38. Kaan E, Wayland R, Bao M, Barkley CM. Effects of native language and training on lexical tone perception: an event-related potential study. Brain Res 2007 May 7;1148:113-122. [doi: 10.1016/j.brainres.2007.02.019] [Medline: 17368579]
39. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trial. Equator Network. URL: https://www.equator-network.org/reporting-guidelines/consort [accessed 2026-06-01]
40. Sim J, Lewis M. The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. J Clin Epidemiol 2012 Mar;65(3):301-308. [doi: 10.1016/j.jclinepi.2011.07.011] [Medline: 22169081]
41. Justifying sample size for a feasibility study. Research Design Service.: NIHR; 2014. URL: https://www.scribd.com/document/657804647/Agriculture-and-Industry-feasibility-study
42. Zhao IY, Lau TW, Li C, et al. Co-designing a web-based, gamified, auditory-cognitive dual-task training system for older adults with hearing loss. Healthcare (Basel) 2025 Nov 15;13(22):2926. [doi: 10.3390/healthcare13222926] [Medline: 41302314]
43. Sweetow RW, Sabes JH. The need for and development of an adaptive Listening and Communication Enhancement (LACE) Program. J Am Acad Audiol 2006 Sep;17(8):538-558. [doi: 10.3766/jaaa.17.8.2] [Medline: 16999250]
44. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
45. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
46. Guo QH, Cao XY, Zhou Y, Zhao QH, Ding D, Hong Z. Application study of quick cognitive screening test in identifying mild cognitive impairment. Neurosci Bull 2010 Feb;26(1):47-54. [doi: 10.1007/s12264-010-0816-4]
47. Lam LCW, Chau RCM, Wong BML, et al. Interim follow-up of a randomized controlled trial comparing Chinese style mind body (Tai Chi) and stretching exercises on cognitive function in subjects at risk of progressive cognitive decline. Int J Geriatr Psychiatry 2011 Jul;26(7):733-740. [doi: 10.1002/gps.2602] [Medline: 21495078]
48. Liu XY, Han Y, Yang SM. A hearing self-reported survey in people over 80 years of age in China by hearing handicap inventory for the elderly–complete version vs screening version. Acta Otolaryngol 2016 Dec;136(12):1242-1247. [doi: 10.3109/00016489.2016.1157729]
49. Chang Q, Sha F, Chan CH, Yip PSF. Validation of an abbreviated version of the Lubben Social Network Scale (“LSNS-6”) and its associations with suicidality among older adults in China. PLoS ONE 2018;13(8):e0201612. [doi: 10.1371/journal.pone.0201612] [Medline: 30071067]
50. Leung GTY, de Jong Gierveld J, Lam LCW. Validation of the Chinese translation of the 6-item De Jong Gierveld Loneliness Scale in elderly Chinese. Int Psychogeriatr 2008 Dec;20(6):1262-1272. [doi: 10.1017/S1041610208007552] [Medline: 18590603]
51. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
52. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 8;396(10248):413-446. [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
53. Henderson Sabes J, Sweetow RW. Variables predicting outcomes on listening and communication enhancement (LACE) training. Int J Audiol 2007 Jul;46(7):374-383. [doi: 10.1080/14992020701297565] [Medline: 17680469]
54. Zając-Lamparska L, Trempała J. Effects of working memory and attentional control training and their transfer onto fluid intelligence in early and late adulthood. Health Psychol Rep 2016;4(1):41-53. [doi: 10.5114/hpr.2016.56846]
55. Papp KV, Snyder PJ, Maruff P, Bartkowiak J, Pietrzak RH. Detecting subtle changes in visuospatial executive function and learning in the amnestic variant of mild cognitive impairment. PLoS ONE 2011;6(7):e21688. [doi: 10.1371/journal.pone.0021688] [Medline: 21779333]
56. Humes LE, Dubno JR, Gordon-Salant S, et al. Central presbycusis: a review and evaluation of the evidence. J Am Acad Audiol 2012 Sep;23(8):635-666. [doi: 10.3766/jaaa.23.8.5] [Medline: 22967738]
57. Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends Cogn Sci (Regul Ed) 2002 Jan;6(1):37-46. [doi: 10.1016/S1364-6613(00)01816-7]
58. Manchaiah VKC. Health behavior change in hearing healthcare: a discussion paper. Audiol Res 2012 Jan 9;2(1):e4. [doi: 10.4081/audiores.2011.e4] [Medline: 26557333]
59. Nolan LS. Age-related hearing loss: why we need to think about sex as a biological variable. J Neurosci Res 2020 Sep;98(9):1705-1720. [doi: 10.1002/jnr.24647] [Medline: 32557661]

XSL•FO
RenderX

Abbreviations

ACDT: auditory-cognitive dual-task training
AI: artificial intelligence
ARHL: age-related hearing loss
AVLT: Auditory Verbal Learning Test
CONSORT: Consolidated Standards of Reporting Trials
HK-MoCA: Hong Kong-Montreal Cognitive Assessment
ICC: intraclass correlation coefficient
PTA: pure-tone average
RCT: randomized controlled trial
TFA: theoretical framework of acceptability

Edited by R Lin; submitted 14.Sep.2025; peer-reviewed by J Fei; revised version received 06.May.2026; accepted 19.May.2026; published 16.Jun.2026.

Please cite as:
Zhao IY, Leung AYM, Li C, Lloyd Parial L, Ma H, Montayre J, Golub JS, Sweetow R, Ng JHY, Chan EA
Web-Based Gamified Auditory-Cognitive Dual-Task Training for Older Adults With Age-Related Hearing Loss: Pilot Randomized Controlled Trial
JMIR Aging 2026;9:e84083
URL: https://aging.jmir.org/2026/1/e84083
doi:10.2196/84083

© Ivy Yan Zhao, Angela Yee Man Leung, Chen Li, Laurence Lloyd Parial, Hongming Ma, Jed Montayre, Justin S Golub, Robert Sweetow, Janet Ho-Yee Ng, Engle Angela Chan. Originally published in JMIR Aging (https://aging.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability and Adoption of Smartwatches by Older Adults in Bangladesh: User Study

Fariah Mahzabeen[1], PhD; Fazlea Rabby[2*], BSc; Neazmul Mowla[2*], BSc; Sadia Afra Ibnat[2], BSc

[1]ECE, North South University, Plot 15, Bashundhara, Dhaka, Bangladesh

[2]ECE, Alumni, Research Assistant, North South University, Dhaka, Bangladesh

[*]these authors contributed equally

Corresponding Author:
Fariah Mahzabeen, PhD
ECE, North South University, Plot 15, Bashundhara, Dhaka, Bangladesh

Abstract

Background: Wearables such as smartwatches can support point-of-care health management for older adults while reducing pressure on health care systems as aging populations grow. Although many studies emphasize technical accuracy, user-centered research on smartwatch adoption among older adults remains limited, particularly in low- and middle-income countries, such as Bangladesh.

Objective: This study evaluated real-world usability, acceptance, and adoption of smartwatches among socioeconomically diverse older adults (aged ≥60 y) in a low- and middle-income country context. It examined how age, education, socioeconomic background, and cultural perceptions influence user experience, and identified barriers and facilitators related to comfort, usability, and long-term engagement. It also explored health-related features that older adults perceived as missing.

Methods: Participants were recruited from older adult communities with diverse socioeconomic and educational backgrounds, and received a commercially available smartwatch. A mixed methods, 3-phase longitudinal design was conducted, including a short-term survey (n=27), 3 long-term survey data points (n=19, 18, and 18), and 2 focus group interviews (n=13 and 12). Quantitative analyses included paired *t* tests across longitudinal data points and chi-square tests for distributional associations, complemented by thematic qualitative analysis.

Results: Comfort improved substantially over time, with participants rating the smartwatch as very comfortable, increasing from 26% (5/19) at baseline to 80% (16/20) at final follow-up. Heart rate monitoring remained the most frequently used feature, rising from 68.42% (13/19) to 75% (15/20). Paired *t* tests showed no significant longitudinal changes in daily usage (t_5=0.00 to 0.22, *P*=.83 to ≥.99), comfort (t_4=−0.07 to 0.30, *P*=.78 to .95), most liked features (t_8=−0.63 to 0.51, *P*=.30 to .62), or clarity of instructions (t_2=0.00 to 0.19, *P*=.87 to ≥.99). Significant increases were observed in feature usage (data point [DP] 2 vs DP3: t_8=−4.16, *P*<.001; DP1 vs DP3: t_8=−2.53, *P*=.03) and perceived difficulty of use (DP2 vs DP3: t_8=−4.29, *P*<.001; DP1 vs DP3: t_8=−2.31, *P*=.05). Chi-square analyses indicated a significant association between study phase and clarity of instructions (χ^2_4=10.3, *P*=.03), with a trend for comfort (χ^2_8=13.8, *P*=.08); other usability dimensions were non-significant (χ^2_{16}=7.1 to 17.5, *P*=.35 to .97). Socioeconomic differences emerged in data-sharing preferences, with higher socioeconomic participants favoring continuous sharing and lower socioeconomic participants preferring emergency-only sharing.

Conclusions: Older adults can adapt to wearable health technologies when given adequate time and support, regardless of education level. Preferences for a limited set of essential health features and faster adoption among participants with lower educational backgrounds highlight the need for simpler smartwatch designs. These findings can guide age-appropriate wearable technologies that prioritize usability, comfort, and familiarity to improve health care access in resource-limited settings.

(JMIR Aging 2026;9:e76404) doi:10.2196/76404

KEYWORDS

smartwatch; wearables; older people; usability; user interface; age-appropriate design; socioeconomic influences; human-computer interaction; low- and middle-income countries

Introduction

The use of wearables for point-of-care health management has the potential to enhance independence and quality of life among older adults while alleviating pressure on health care systems amid an aging population worldwide [1-3]. While numerous studies have examined the technical accuracy and performance of wearable technologies, significant challenges and limitations remain [4]. Validation studies have assessed wearable activity

XSL•FO
RenderX

trackers in older and clinical populations, supporting measurement reliability while highlighting the need for usability-focused research [5]. Furthermore, user-centered investigations into factors influencing the adoption of wearables among older adults, particularly in low- and middle-income countries (LMICs), such as Bangladesh, are scarce [6].

Wearables have demonstrated benefits in chronic disease management, such as diabetes care, through real-time monitoring and personalized feedback [7]. Wearable sensors and systems have also been extensively explored in rehabilitation and long-term monitoring, demonstrating value for supporting functional health needs [8]. Smartwatches and Internet of Things–based applications can also facilitate independent living by enabling safety monitoring and home health supervision for aging populations [9]. Evidence also suggests that wearable-based interventions can improve physical activity and health outcomes among older adults in community settings [10]. Wearables offer diverse applications, including position and posture tracking, activity recognition, and real-time vital sign monitoring, which are particularly useful in home health care [11].

Despite these advantages, adoption among older adults is inconsistent, and mixed methods studies suggest that usability, contextual, and behavioral factors shape long-term engagement [7]. Design aspects may introduce stigma perceptions, affecting willingness to adopt such devices [12]. Acceptance models further indicate that perceived usefulness and health-monitoring value are central determinants of adoption intention [13].

Older adults may also experience technology anxiety and resistance to change, reducing their readiness to adopt wearable health technologies [14]. Predictive studies suggest that adoption is influenced by multiple interacting antecedents, including behavioral, psychological, and contextual factors [15]. Concurrently, wearable health technologies intersect with artificial intelligence–driven markets, where health consciousness and privacy preferences play important roles in user adoption [16].

Community-based research indicates that older adults' acceptance of smartwatches is often driven by health and location tracking functions when these align with personal needs [17]. Adoption of fitness technology encourages physical activity and behavioral change, but adoption and long-term engagement are influenced by both perceived usefulness and social visibility, highlighting the interplay between practical value and social context [18,19]. User-centric research from LMIC settings, such as the Philippines, reinforces that real-world adoption depends on usability, perceived value, and local contextual factors [20,21]. At the same time, wearable technology development faces broader challenges, including reliability, battery life, and interface complexity, which can hinder adoption [21].

Older adults often encounter difficulties using smartwatches, including small displays, discomfort, limited battery life, and complex interfaces, making usability evaluation critical [22]. Qualitative studies emphasize that long-term comfort is shaped by interacting human and design factors and strongly influences continued use [23,24]. Focus group research suggests that perceptions, motivators, and barriers influence tracker use during behavior change maintenance, emphasizing the role of perceived value and ongoing support [24]. Technology adoption decisions are further shaped by psychological, social, and situational factors [25].

It is found that qualitative synthesis methods are effective for exploring lived experiences and uncovering usability barriers that quantitative approaches may overlook [26]. Observational studies show that wearable activity trackers can be acceptable for older adults in real-world settings, although practical challenges remain [27]. Systematic reviews highlight that experiences with wearables vary depending on device type, setting, support systems, and integration into daily routines [28,29]. For example, field testing of fall detection wearables demonstrates that real-world deployment can reveal practical challenges not captured in controlled studies [29]. Wearable systems for measuring gait and physical activity in chronic conditions require usability considerations to ensure adherence and data quality [30,31].

Evidence suggests that continued wearable use among older adults is strongly associated with comfort, alignment with daily routines, and perceived value, underscoring the importance of context-sensitive, user-centered design [23,32-38]. Addressing these factors is essential for designing age-appropriate wearable technologies that maximize adoption, usability, and health impact in resource-limited contexts. Hence, in this work, we use a mixed methods study to examine user adoption and reactions to existing smartwatches across diverse older adult populations, and to evaluate age-centric, socioeconomic, and user-focused adoption barriers and facilitators.

Methods

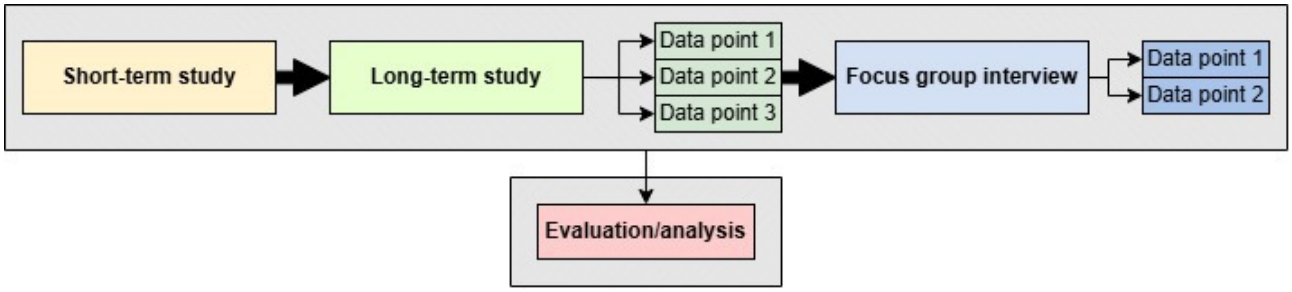
Recruitment

Overview

The core goal of this study is to assess the usability of smartwatches among the older adult communities in an LMIC, like Bangladesh. The first step of the study was to identify interest levels and engage with communities of older adult individuals, as detailed in the following sections. Following that, a survey process of 3 cycles was conducted using purposive sampling techniques by recruiting individuals from different socioeconomic and educational backgrounds, lifestyles, and educational backgrounds. This study cycle includes (1) a short-term survey, (2) a long-term survey with 3 data points (DPs), and (3) a focus group interview with 2 DPs. The final step of the study includes survey analysis and finding key insights from all the cycles for designing a specific older adults–age-centric health monitoring smartwatch. Figure 1 presents a flowchart of the methodological framework of the study.

XSL•FO
RenderX

Figure 1. Overall methodology showing the study design phases: short-term, long-term (with 3 data points), and focus group interviews (with 2 data points), followed by evaluation and analysis.



Participant Demographics and Recruitment

An institutional review board (IRB) was approved by the North South University Scientific Regulatory Committee, and all participants were asked to complete a consent questionnaire. At the start of our study, we surveyed older adult individuals aged 60 years and older from our local community (friends, families, neighbors, and service staff) to understand their interest in using smartwatches for health monitoring. Then we contacted 5 old-age homes in and around Dhaka City and visited them to assess the interest level of participation. Initially, 2 old-age homes and a group of janitors (from residential areas of the authors) agreed and were recruited in the short-term survey study. However, due to internal management issues at one of the old-age homes, the consecutive surveys were conducted with the participants of one old-age home called Probeen Nibash (managed by the Ministry of Social Welfare in Dhaka, Bangladesh). We procured approval from the management of Probeen Nibash to conduct the study at their center and then talked individually to residents to explain the purpose and format of the long-term study. Based on their consent and the inclusion-exclusion criterion, selective participants were recruited from this old-age home. Similarly, several janitors were approached from the authors' local communities, and those who willingly consented were recruited to continue the long-term study. The participants were between 60 and 103 years old, male and female, and purposefully sampled from diverse backgrounds. The participants were divided into 2 groups based on socioeconomic and educational backgrounds.

Inclusion and Exclusion Criteria

Older adults (older than 60 y) who were generally healthy and able to walk >50% time around on their own or with the use of an assistive device (such as a walker) were included. Older adults with restricted mobility (like in a wheelchair >50% time) or conditions that may limit the use of devices, such as blindness and cognitive decline, were excluded.

Survey Method and Participant Demographic Analysis

Overview

Before the first survey, a preliminary questionnaire session was conducted with each participant to collect demographic and baseline health-related information, including age, gender, occupation, educational qualification, and medical history of the past 6 months. The objective of this preliminary questionnaire was to evaluate their eligibility (and health condition) and interest in the study. Older adults (older than 60 y) who are generally healthy and able to walk >50% of the time around on their own or with the use of an assistive device (such as a walker) were included. In contrast, older adults with restricted mobility (like in a wheelchair >50% of the time) or conditions that may limit the use of devices, such as blindness and cognitive decline, were excluded. Aging people with severe cognitive dysfunctions were also excluded. The eligibility criteria included a set of subsectioned questionnaires to assess the participant's health and mental health conditions, technological knowledge, and experience with modern technology to ensure their suitability for our study. From these observations, we found that most participants had no previous experience using technologically advanced devices like any smart wearable, and none had ever used a smartwatch.

Each participant was provided with an Oraimo Watch 4 Plus smartwatch after providing written informed consent to participate in the study. A detailed user guide was also supplied to ensure that participants clearly understood how to operate the device. At the time of the study, the device was commercially available at an approximate cost of US $20.06. The Oraimo Watch 4 Plus was selected due to its affordable price and the availability of standard health-monitoring features commonly found in other commercially available smartwatches.

Short-Term Survey

This survey aimed to determine short-term adherence and compliance to the protocol using the smartwatch device among the participants. Short-term surveys were conducted in 3 places after the participants were distributed the smartwatch for 14 days, with 27 participants from varying socioeconomic and educational backgrounds, as presented in Table 1. Among them, 15 out of 27 (56%) were male participants, and 12 out of 27 (44%) were female, with an average age of 60.55 (SD 7.38) years. In contrast, 13 out of 27 (48%) of participants belonged to a higher socioeconomic background, while 14 out of 27 (52%) belonged to a lower socioeconomic background. This survey involved a comprehensive questionnaire with 19 questions, including 15 multiple-choice and 4 short-answer questions. The following key questions were asked of the participants during the short-term survey:

1. How often do you wear your smartwatch?
2. How comfortable is your smartwatch to wear?
3. Have wearable health technologies helped you in achieving your health goals?
4. Did you find the screen size and text clear and easy to read?
5. Does using a smartwatch present any barriers concerning your religious, social, or cultural beliefs?

XSL•FO
RenderX

6. Did the device provide clear and simple instructions?
7. Did the device feel age-appropriate for you?
8. Have wearable health technologies helped you in achieving your health goals?
9. Which features do you like the best?
10. Which features do you feel uncomfortable with?
11. Which features are important to you in wearable health technology?
12. Would you recommend this device to other older adults?
13. Did you find the device frustrating to use?
14. What challenges do you face when using wearable health technology?
15. What improvements do you expect from this smartwatch?
16. How do you manage and interpret the data collected by your wearable health technology?
17. What kind of support do you expect from smartwatch manufacturers regarding religious use?

Table . Total number of participants, male and female ratio, mean age of participants, percentage of lower socioeconomic class participants, and percentage of higher socioeconomic class participants.

Survey type	Total participants, n	Male, n (%)	Female, n (%)	Age of participants (y), mean (SD)	Participants belonging to lower socioeconomic class, n (%)	Participants belonging to higher socioeconomic class, n (%)
Short-term survey	27	15 (56)	12 (44)	60.55 (7.38)	14 (52)	13 (49)
Long-term survey 1	19	15 (79)	4 (21)	65.00 (8.46)	7 (37)	12 (63)
Long-term survey 2	18	16 (89)	2 (11)	63.11 (5.26)	9 (50)	9 (50)
Long-term survey 3	18	15 (83)	3 (17)	63.16 (5.20)	10 (56)	8 (44)
Focus group interview 1	13	9 (69)	4 (31)	65.08 (9.59)	6 (46)	7 (54)
Focus group interview 2	12	10 (83)	2 (17)	64.75 (8.73)	7 (58)	5 (42)

Long-Term Survey

After completing the short-term survey, the long-term survey had to be proceeded with fewer participants. The participants of one of the old-age homes were excluded from the survey due to internal management issues at the organization, leading us to proceed with just one old-age home's participants and a group of janitors. Also, a few participants from these 2 groups felt disinterested in continuing after the first DP of the long-term survey, primarily due to health issues and inability to adhere to the study guidelines. The first DP, long-term survey-1, was conducted with 19 participants, while long-term survey-2 and long-term survey-3 were conducted with 18 participants due to the sudden death of one of the participants from long-term survey-1. Each long-term survey was conducted at an interval of 10 days. Therefore, this survey section lasted for 30 days. As described in Table 1, the long-term survey-1 was conducted with 15 out of 19 (79%) male and 4 out of 19 (21%) female participants, 12 out of 19 (63%) belonging to a high socioeconomic background, and 7 out of 19 (37%) from a low socioeconomic background. In contrast, long-term survey-2 and long-term survey-3 were conducted with 16 out of 18 (89%) male participants and 2 out of 18 (11%) female participants. Additionally, a minimum of 8 out of 18 (45%) of the participants belonged to high socioeconomic backgrounds, while the other 10 out of 18 (56%) were from low socioeconomic backgrounds. The survey questions of the long-term survey were correlated to the short-term survey, and some variations in question types were used to improve assessment and analysis. The aim of this survey session was to find insights into usability factors of the smartwatch over a long-term period.

The key questions of the long-term survey included:

1. How often do you wear your smartwatch?
2. Which of the following health metrics do you track?
3. Which features do you like the best and are most comfortable with?
4. Which feature of the smartwatch do you find uncomfortable or dislike the most?
5. Rate how comfortable your smartwatch is to wear (out of 5).
6. Does using a smartwatch present any barriers to your religious, social, or cultural beliefs?
7. Did the device provide clear and simple instructions?
8. Has there been any technological challenge in using or understanding the watch?
9. Did the device feel age-appropriate for you?
10. Has the smartwatch helped you in achieving your health goals?
11. What challenges do you face when using the smartwatch?
12. What improvements do you expect from this smartwatch?
13. What health parameters did you collect from the smartwatch?

Focus Group Survey

Based on the socioeconomic background, 2 focus groups were formed among the participants of the long-term survey. However, some participants did not show interest in participating in these sessions due to various reasons. Group A included

XSL•FO
RenderX

participants from higher socioeconomic backgrounds, while group B included participants from lower socioeconomic backgrounds. The first focus group survey was conducted 14 days after completion of the long-term survey (round 3), and the second focus group survey was conducted 14 days after the first focus group survey. These focus groups were repeated in 14-day intervals to investigate if the participants' responses were consistent. More importantly, the focus groups were stratified across socioeconomic backgrounds to understand if and how different backgrounds might influence their experiences with the survey. The first focus group survey was conducted with 13 participants, while group A comprised 7 out of 13 (54%) participants, and group B comprised 6 out of 13 (46%).

In contrast, during the second focus group survey, there was 1 dropout due to illness, consisting of 12 participants, with 5 out of 12 (42%) representing group A and 7 out of 12 (58%) from group B. Each focus group session was conducted for approximately 1 and a half hours at the primary location of the participants to ensure convenience and comfort. Participants were asked structured questions during the sessions, and their responses were collected individually. A casual conversation was facilitated alongside the structured format to explore their thoughts and expectations regarding wearable technologies. The entire session was audio-recorded with participant consent to capture all responses accurately.

Ethical Considerations

This user study or randomized controlled trial to study the adoption behavior of smartwatches among older adults was approved by North South University's IRB and Ethics Review Committee (IRB #2023/OR-NSU/IRB/1231). All the surveys and focus groups described in the study were performed within the guidelines and duration of the approved IRB. A consent form and a baseline health questionnaire were evaluated from each participant before the surveys were conducted. Moreover, participants had the choice to drop out of the study at any time point, which happened over the course of the study. The participants were also provided with contact information of the researcher team for any queries. All data were collected, stored, and analyzed in compliance with HIPAA (Health Insurance Portability and Accountability Act) guidelines and data privacy laws. The participants were recruited by consent on a pro bono basis, although they were gifted the smartwatches at the end of the study.

Results

Short-Term Survey

Overview

The first short-term survey of the study was conducted to determine compliance levels among the participants from varying socioeconomic and educational backgrounds. The survey was designed with a variety of questions mentioned in the short-term survey section. A significant variation in responses was observed.

Responses to the first question, regarding the duration of smartwatch usage, are portrayed in Table 2. Notably, except for 1 individual, all participants used the smartwatch for at least 3 hours daily, with most participants using it for some time (defined as 3 - 4 h), most of the day, and all day, highlighting the interest of the participants in exploring the smartwatch's features and usability.

Table . Daily smartwatch use duration among participants (n=25).

Daily use duration	Participants, n (%)
All the time	7 (28)
Most of the day, except while sleeping	6 (24)
Occasionally	4 (16)
Sometimes	7 (28)
Rarely or never	1 (4)

The second question examined the health features participants tracked. Most of the participants used or tried using multiple features. As illustrated in Table 3, the most frequently tracked features were heart rate monitoring ("Heart Rate"), step counting, and oxygen saturation level monitoring (SpO_2). Less frequently tracked features included sleep quality, calorie burn, respiratory rate monitoring, and stress level monitoring. Participants indicated that "Heart Rate," "Step Count," and "SpO_2" features were easier to access and less complex to use than "Sleep Quality," "Calorie Burn," "Respiratory Rate Monitoring," and "Stress Level Monitoring," which were described as more challenging to understand and monitor effectively.

XSL•FO
RenderX

Table . Participants who reported using each smartwatch feature.

Used feature	Participants, n (%)
Steps	14 (56)
Heart rate	16 (64)
Sleep quality	3 (12)
Calorie, burn	5 (20)
Respiratory rate	3 (12)
Stress	5 (20)
SpO_2[a]	12 (48)
None	4 (16)
Time	4 (16)

[a]SpO_2: oxygen saturation level monitoring.

The third and fourth questions addressed privacy concerns and workplace compatibility. In response to these questions, all participants expressed positive views, indicating comfort in using the smartwatch at their workplaces without privacy concerns.

Questions 5 and 6 explored the features participants liked and disliked. Heart Rate emerged as the most liked feature due to its simplicity and ease of use.

Comfort, addressed in questions 7 and 8, was another key area of investigation. As illustrated in Table 4, most participants reported the smartwatch as "completely comfortable," while a smaller group consisting of 3 (15%) users found it "mostly comfortable," and another group of 4 (20%) users felt "somewhat comfortable," suggesting that participants generally did not encounter significant physiological discomfort while wearing the device.

Table . Number of participants reporting comfort levels during smartwatch use.

Comfortability	Participants, n (%)
Very comfortable	1 (5)
Somewhat comfortable	4 (20)
Mostly comfortable	3 (15)
Completely comfortable	12 (60)

The ninth question examined whether the smartwatch posed religious, social, or cultural barriers for older adult individuals. As illustrated in Table 5, a majority (20 participants) responded "not at all," while 2 participants reported "slight" barriers, and 2 remained neutral. Only 1 participant perceived the watch as a large hindrance while performing religious, social, and cultural activities.

Table . Participant responses regarding barriers.

Response	Participants, n (%)
Very	1 (4)
Slight	2 (8)
Not at all	20 (80)
Neutral	2 (8)

Participants were also asked whether they found the smartwatch age-appropriate and if they would recommend it to others in their age group. As illustrated in Table 6, approximately 52% (n=13) of the participants considered the smartwatch suitable for their age and indicated they would use it for a longer duration in the future. In comparison, 8 (32%) participants expressed neutrality. On the contrary, 2 participants found it very age-inappropriate and declined to use it in the future, citing concerns about advanced technology for older people. Furthermore, 1 female participant (>70 y old) from an old-age home commented:

> *At this age, we cannot use advanced technology like a smartwatch. We are older people, and we believe our instincts suffice to predict physical disorders. I trust natural death will come when it is time. No technology is appropriate to predict diseases early, nor can it serve us with physical complexity.*

XSL•FO
RenderX

Table . Participant responses on whether the smartwatch was compatible with their age.

Response	Participants, n (%)
Yes	13 (52)
No	4 (16)
Neutral	8 (32)

Another question explored whether the smartwatch helped participants achieve health goals. A total of 13 (52%) participants believed it was helpful, while 3 (12%) felt it could not independently establish health goals. Moreover, 9 (36%) participants expressed neutrality. Table 7 presents an analysis of these responses.

Table . Responses from perceptions of whether smartwatch use helped achieve personal health goals.

Response	Participants, n (%)
Yes	13 (52)
No	3 (12)
Neutral	9 (36)

The final question focused on whether participants found the smartwatch frustrating to use. As illustrated in Table 8, 15 (60%) participants responded negatively, 4 (16%) reported frustrations, and 6 (24%) remained neutral. Frustration was reported mostly because of the age-inappropriate user interface.

Table . Responses to whether the device was frustrating to use.

Response	Participants, n (%)
Yes	4 (16)
No	15 (60)
Neutral	6 (24)

Participant Comments and Insights During Short-Term Survey

The survey also captured unique insights from participants, particularly those with lower levels of education. For example, a janitor noted:

> *After using the smartwatch, I often felt hungry and could eat larger portions. This is due to the smartwatch.* [Janitor, 63 y old, primary school graduate]

Similarly, another participant of an old-age home resident noted:

> [I had arthritis, but after using the smartwatch, the pain miraculously healed. I think the device worked like medicine.] [Old-age home resident, 103 y old, with no formal education]

Another female participant (62 y old) from an old-age home reported that she felt nausea, dizziness, and joint pain whenever she used the smartwatch.

The overall responses of this survey phase highlight that both groups of participants were curious about the smartwatch. However, they faced some challenges while using it. According to the responses, Heart Rate, Step Count, and SpO_2 were the easiest features to monitor. Also, the watch was comfortable overall and created no barriers while doing regular activities. However, a few interesting comments were received from participants with lower education and socioeconomic backgrounds. Also, a large portion of the neutrality in the usability of the watch in achieving health goals indicates that a group of participants was confused about the appropriate technical knowledge of using the watch. Many participants shared their complexity in understanding the watch's functionality during the interview. The research team answered that confusion with proper care and ensured clarification about the issues.

Long-Term Study

Overview

The given survey revealed some interesting trends in smartwatch use and responses from participants across 3 phases of the long-term survey. Unlike the former short-term survey, this one included a series of questions categorized into daily usage, comfort level, feature usage patterns, and expected improvements that participants made.

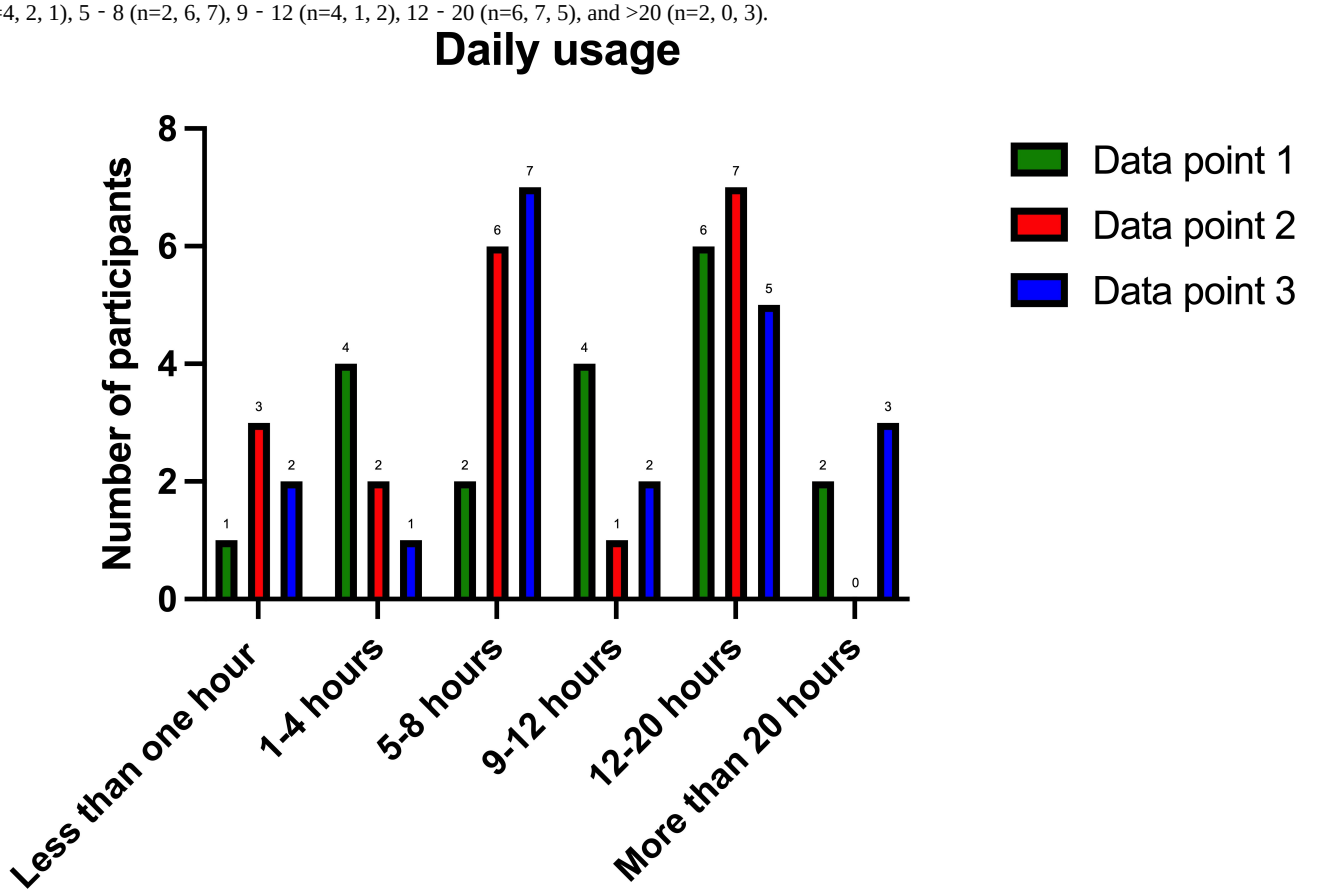
Daily Usage

As Figure 2 illustrates, 31.5% (6/19) of the participants used the smartwatch for 12 - 20 hours daily during the first phase of the survey, thus being the most common usage duration; this was followed by 21% (4/19) of participants who used it from 9 - 12 hours, and exactly the same percentage for those who said they used the smartwatch for 1 - 4 hours daily. Smaller use of the smartwatch (daily periods exceeding 20 h) was reported by 10.5% (2/19), while a single participant reported not using it at all. Usage patterns continued shifting through the second phase, with 31.5% (6/19) of participants using the smartwatch for 5 - 8 hours daily and 36.8% (7/19) for 12 - 20

XSL•FO
RenderX

hours. Only 15.7% (3/19) of participants used the watch less than an hour daily, indicating that almost all participants engaged with the device to some degree. Finally, in the third phase, 20% (5/20) of respondents resumed using the smartwatch in the 12 - 20 hours daily category, and 35% (7/20) used it between 5 - 8 hours.

Figure 2. Grouped bar chart showing the number of participants reporting daily smartwatch use across 3 study time points (hours): <1 (n=1, 3, 2), 1 - 4 (n=4, 2, 1), 5 - 8 (n=2, 6, 7), 9 - 12 (n=4, 1, 2), 12 - 20 (n=6, 7, 5), and >20 (n=2, 0, 3).



The results of the *t* test for daily usage, depicted in Table 2, showed a lack of significant variance in usage patterns across all phases; this is evidenced by *P* values ranging from .78 to ≥.99, which reveals that the overall time spent by the participants on the smartwatch remained relatively stable across the course of the study.

XSL•FO
RenderX

Table . Summary of paired *t* tests across 3 long-term study data points.

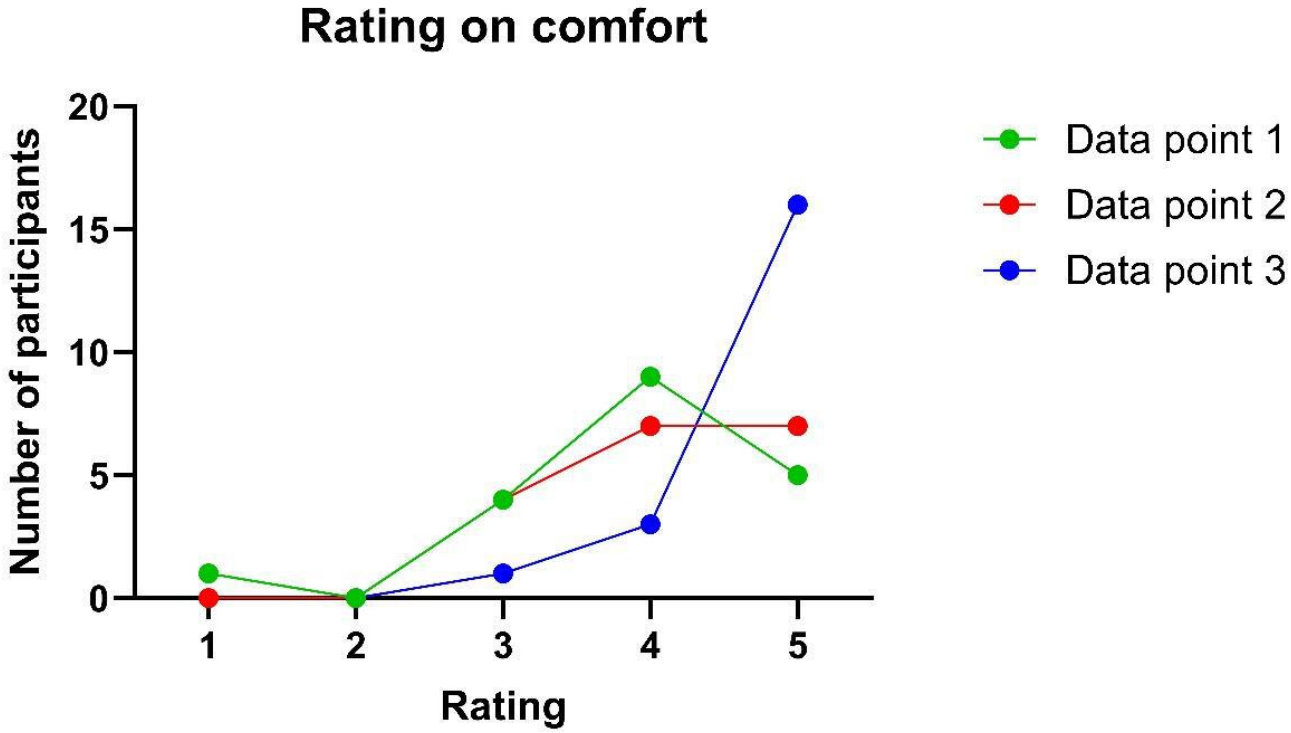
Survey dimension	Comparison	*t* test (*df*)	*P* value	Result
Daily usage	Data point 1 vs data point 2	0.00 (5)	>.99	Not significant
Daily usage	Data point 2 vs data point 3	−0.22 (5)	.83	Not significant
Daily usage	Data point 1 vs data point 3	−0.14 (5)	.89	Not significant
Rating on comfort	Data point 1 vs data point 2	0.30 (4)	.78	Not significant
Rating on comfort	Data point 2 vs data point 3	−0.17 (4)	.87	Not significant
Rating on comfort	Data point 1 vs data point 3	−0.07 (4)	.95	Not significant
Used features	Data point 1 vs data point 2	1.26 (8)	.24	Not significant
Used features	Data point 2 vs data point 3	−4.16 (8)	<.001	Significant
Used features	Data point 1 vs data point 3	−2.53 (8)	.03	Significant
Most liked features	Data point 1 vs data point 2	0.51 (8)	.62	Not significant
Most liked features	Data point 2 vs data point 3	−1.12 (8)	.30	Not significant
Most liked features	Data point 1 vs data point 3	−0.63 (8)	.55	Not significant
Clear instructions	Data point 1 vs data point 2	0.19 (2)	.87	Not significant
Clear instructions	Data point 2 vs data point 3	0.00 (2)	>.99	Not significant
Clear instructions	Data point 1 vs data point 3	0.09 (2)	.94	Not significant
Difficult to use feature	Data point 1 vs data point 2	1.96 (8)	.08	Not significant
Difficult to use feature	Data point 2 vs data point 3	−4.29 (8)	<.001	Significant
Difficult to use feature	Data point 1 vs data point 3	−2.31 (8)	.05	Marginally significant

Comfort Level Investigation

In fact, comfort levels improved a lot in the 3 phases, as portrayed in Figure 3. Initially, 47.3% (9/19) of the participants found the smartwatch comfortable, with only 26.3% (5/19) describing it as very comfortable. A small percentage, 5.3% (1/19), described it as highly uncomfortable since the features were complicated and the users did not have enough knowledge about them. However, in the second phase, no one reported discomfort, while 38.9% (7/18) found the smartwatch very comfortable, which marks the positive shift in user experience. In the third phase, 80% (16/20) rated the smartwatch as very comfortable, while only 1 participant reported it as fairly comfortable, amounting to 5% (1/20).

XSL•FO
RenderX

Figure 3. Line graph showing the distribution of comfort ratings (1=lowest to 5=highest) across 3 data points. The number of participants reporting higher comfort ratings (4-5) increased over time, with rating 5 rising at the final follow-up.



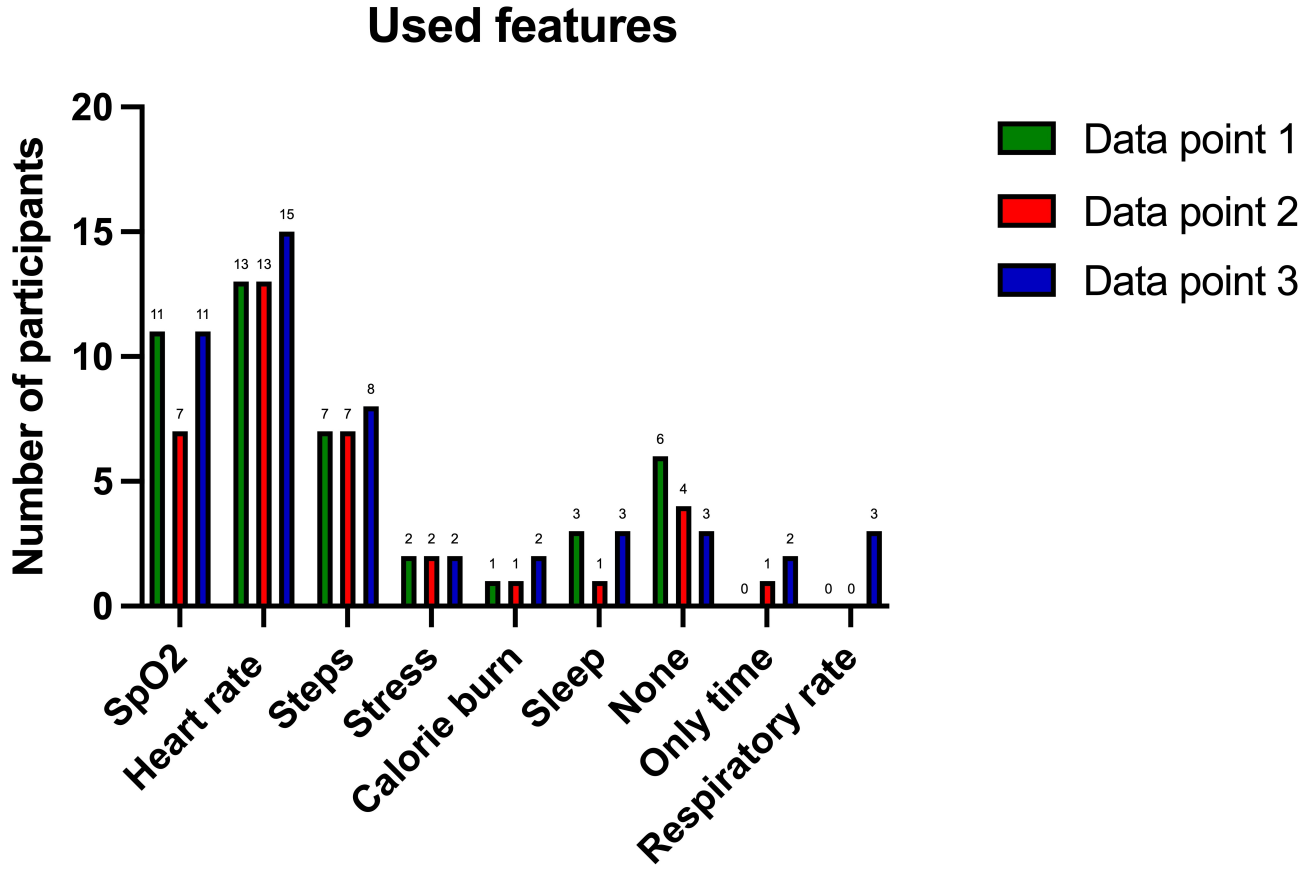
Furthermore, *t* test results of comfort ratings are insignificant among the phases, with *P* values ranging between .78 and .95 (Table 9). However, as revealed by the chi-square test, comfort perception improved dramatically over the survey period (Table 9), with a significant association—namely, *P*=.03 of phases with comfort ratings.

Investigation of Feature Usage Pattern

The usage of the features shows some fixed preferences along with some dynamic behavior, indicated in Figure 4 below. The most used feature in all phases remains heart rate monitoring, which shows an increased usage level from 68.4% (13/19) in phase 1 to 75% (15/20) in phase 3. The participants were primarily using other features such as SpO_2 and Step Count, but with a small fluctuation in them. For example, usage of SpO_2 reduced in phase 2 but increased in phase 3, which shows they were testing different features before making a preference.

XSL•FO
RenderX

Figure 4. Grouped bar chart showing smartwatch feature use across 3 data points. Heart rate and step count were the most used features throughout, while other features were less commonly used and some participants reported using no features.



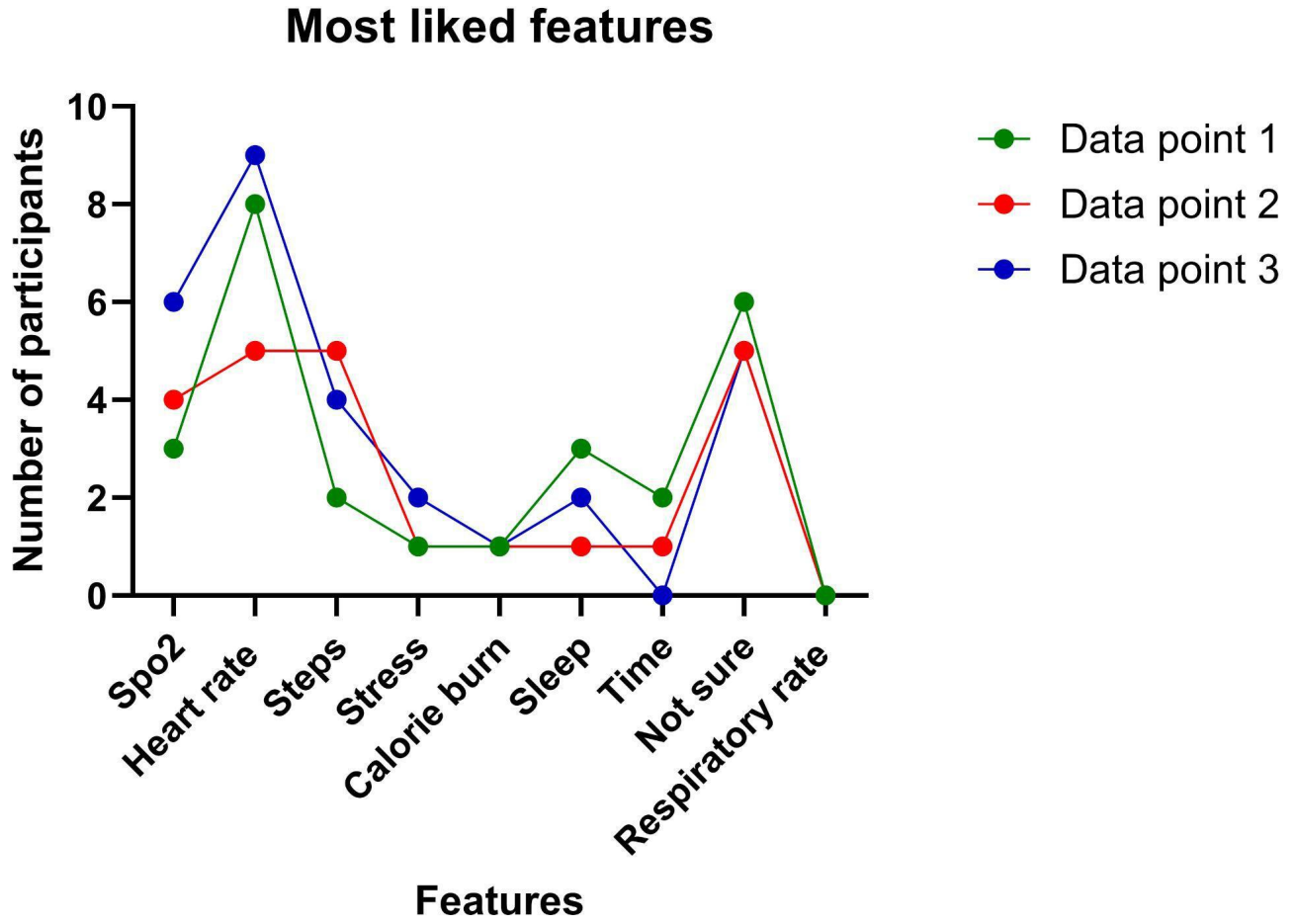
The respiratory rate function, which had not been used in phases 1 and 2, gained acceptance among 15% (3/18) of participants in phase 3, which indicated increased interest in advanced functions with time as participants got accustomed to using the device.

The results of the *t* test indicated a significant change in used features over time, especially when comparing DP2 and DP3 ($P<.001$) and another significant change when comparing DP1 and DP3 ($P=.03$), which indicated that people have gotten used to using the smartwatch features.

Besides usage, another area that we have explored is the most liked features in each stage of our survey, which can be seen in Figure 5. Heart Rate remains among the most liked features in all phases. As seen in Table 2, *t* test analysis for most liked features did not reach significance in the comparisons of DP1 with DP2 with a *P* value of .62, DP2 with DP3 with a *P* value of .30, and DP1 and DP3 with a *P* value of .55 for Heart Rate.

XSL•FO
RenderX

Figure 5. Line graph showing the number of participants selecting comfort ratings from 1 (lowest) to 5 (highest) at 3 measurement points. Comfort ratings shifted upward over time, with the highest rating (5) becoming most common at the final follow-up.



For example, in an attempt to capture features such as SpO_2, the participants recorded a higher preference in the third phase relative to the first phase, as can be noted by the SD recorded in Table 9, which stood at 1.24722 (mean 4.33). Some of these features, such as SpO_2, were thus better appreciated when participants were more conversant with the device. Other functions, such as tracking and respiratory rate measurement, were well-received but demonstrated inconsistency in preference, especially in the early phases.

What is interesting, however, is the fact that respiratory rate tracking was considered hard to use but gained acceptance over time, which is in line with the increased usage of advanced functionality in total. The nonuse of some functions in the first phase, with 31.6% (6/19) of participants not using any functions, highlights the learning curve involved in using smartwatches in general, which had declined substantially in phase 3. The above findings indicate how preferences in features change over time, with Heart Rate being a continually popular choice, given other features, such as SpO_2 and respiratory rate monitoring, becoming more popular with increased familiarity.

Clear Instruction

Several trends were observed based on the feedback obtained from the participants with regard to the clarity of instructions, as presented in Table 10. At the beginning, 78.9% (15/19) found it simple and clear; hence, the reason for the easy adoption. This trend flipped as users continued using the application to explore features that became more complicated. In the second and third phases, the participants began to veer toward the neutral side concerning clarity. This is depicted by the high SD obtained in "Yes" (mean 12, SD 2.44) and "Neutral" (mean 4.33, SD 2.86), indicating that respondents do change with increased usage upon exploring more features. These results stress the need for increasing the clarity of instructions as users proceed to more complex features.

Table . Perceptions on clear and simple instructions across data points.

	Data point 1	Data point 2	Data point 3
Yes	15	12	9
No	3	2	1
Neutral	1	4	8

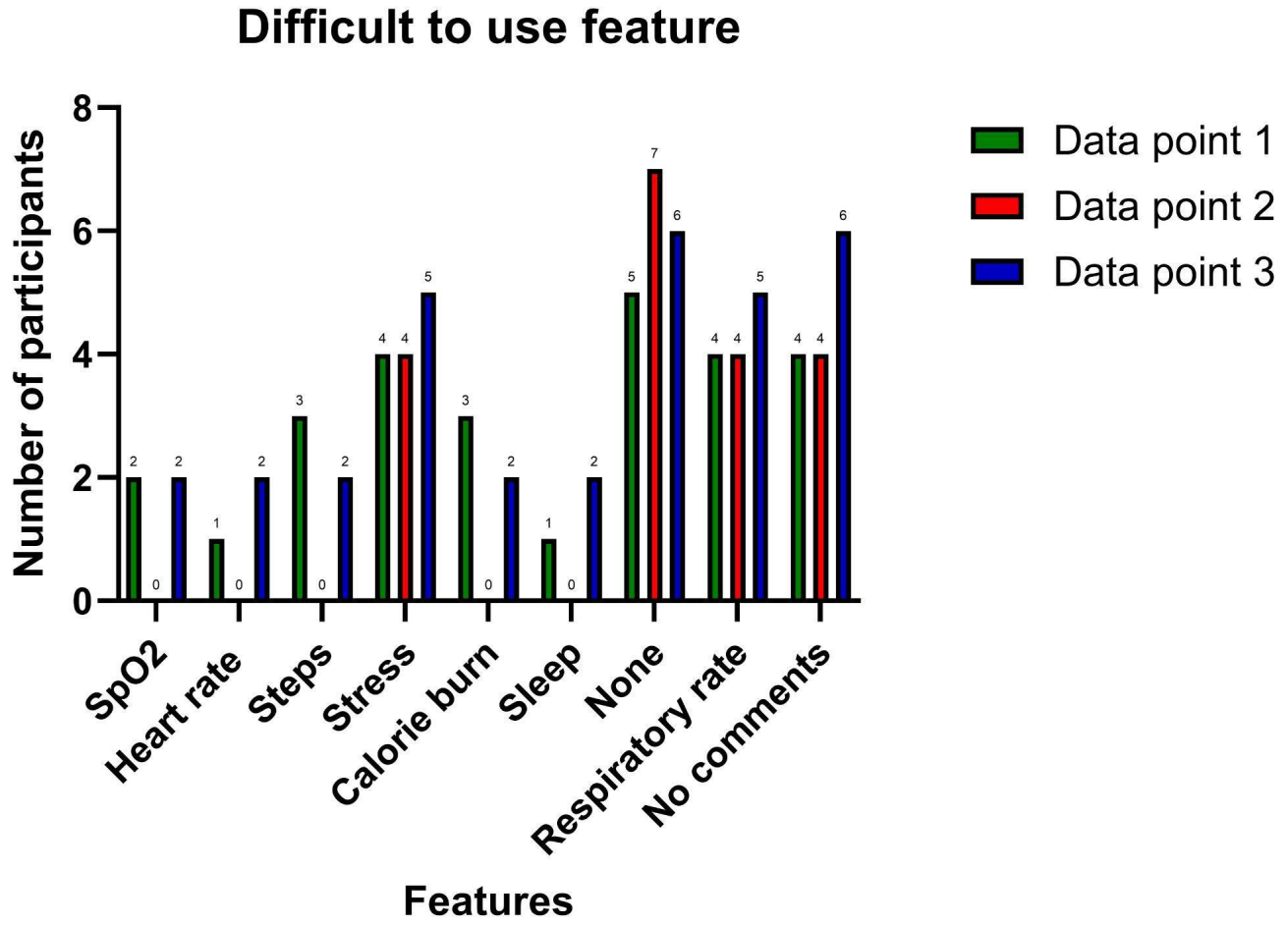
In Table 9, the chi-square test indicates a significant association (P=.04) between survey phase and participants' opinion on instruction clarity, which signals the dynamic user perception as feature complexity increases.

XSL•FO
RenderX

Difficult to Use Features

Figure 6 illustrates that the most difficult feature, especially in the third DP, is stress measurement. This would indicate a possible design or functionality issue with this feature. The same trend occurred for respiratory rate monitoring, which was considered a difficult feature through an increasing number of participants during the course of the experiment. Finally, on the opposite side of the spectrum, Heart Rate was the most favored feature and had no reports of difficulty throughout all points.

Figure 6. Grouped bar chart showing reported difficulty using smartwatch features across 3 data points. Most participants reported no difficulty, while some noted difficulty with specific features.



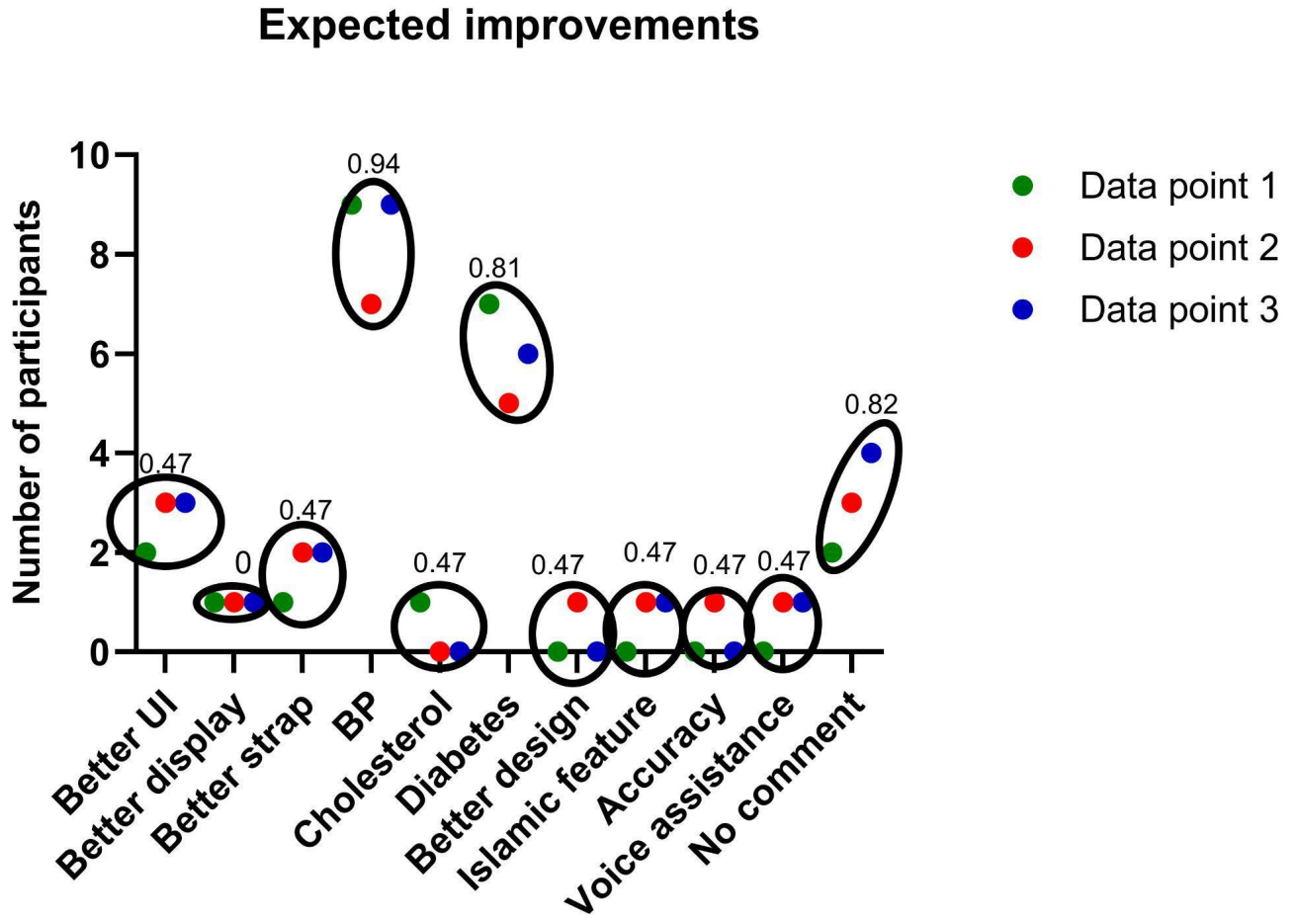
Furthermore, *t* test results indicate significant changes in the perceptions of difficulty regarding difficult to use features, where $P<.001$ for DP2 versus DP3 and $P=.05$ for DP1 versus DP3 support that design improvements are required in some features, such as stress measurement and respiratory rate.

Investigation of Expected Improvement

After long-term use of the smartwatch, adaptability and comfort were observed among the participants, as illustrated in Figure 7. Alongside, a trend was also observed in the expected improvements from the smartwatch (Figure 7). The data show that the most commonly suggested improvements were blood pressure and diabetes monitoring capabilities from the smartwatch, indicating that older people are mostly interested in critical health monitoring features. Interestingly, some participants initially had no comments on improvements but later suggested new features, though this shift was minimal. The overall low SD across all expected improvements confirms that user needs remained stable, reinforcing the necessity of designing a smartwatch specifically for older adults' health monitoring.

XSL•FO
RenderX

Figure 7. Scatter plot showing participant-suggested smartwatch improvements across 3 study time points (data point 1 - 3). Points represent counts for each suggested feature, and labels indicate variability (SD) over time.



Focus Group Survey

Overview

The focus group interviews constituted the final phase of the study and were designed to provide qualitative insights that contextualize and extend the survey findings. Moreover, 2 rounds of focus group interviews were conducted using identical structured interview guides to enable both cross-group and longitudinal comparisons. Participants were stratified by socioeconomic status and educational background to explore how these factors influenced perceptions, adoption, and experiences of smartwatch-based health monitoring.

Group A comprised older adults residing in an assisted living facility, representing a higher socioeconomic and educational profile; all participants had completed at least a high school education, with the majority holding undergraduate or graduate degrees. Group B consisted of individuals from lower socioeconomic backgrounds with less formal education (below high school level), who were employed in custodial and maintenance roles. This stratification enabled a comparative examination of digital health adoption across differing social and educational contexts.

Qualitative responses were analyzed thematically, with dominant patterns reported alongside illustrative quotations. Descriptive response counts are included to enhance transparency, consistent with *Journal of Medical Internet Research* qualitative reporting guidelines.

Findings of Focus Group Interview 1

The focus group interviews took place in separate locations and under the same setting to gather similar perspectives. A total of 13 participants participated in the study, including 9 males and 4 females. Group A included 7 participants, of whom 4 were female and 3 were male. Group B consisted of 6 male participants. Both focus group interviews followed the same structured questionnaire to examine potential differences in responses and variability across the groups.

Perception of Smartwatches

Participants' first impressions of smartwatches revealed clear group-level differences. Group A participants predominantly expressed positive attitudes, whereas group B participants more frequently articulated curiosity rather than immediate endorsement. A small subset of participants (2 from group A and 1 from group B) expressed skepticism or negative perceptions.

Qualitative analysis identified three emergent subthemes:

1. Perceived cognitive and physical burden,
2. Concerns about measurement accuracy, and
3. Skepticism toward replacing clinical expertise.

For instance, concerns about language and accessibility were raised by a group B participant ("I cannot read English texts

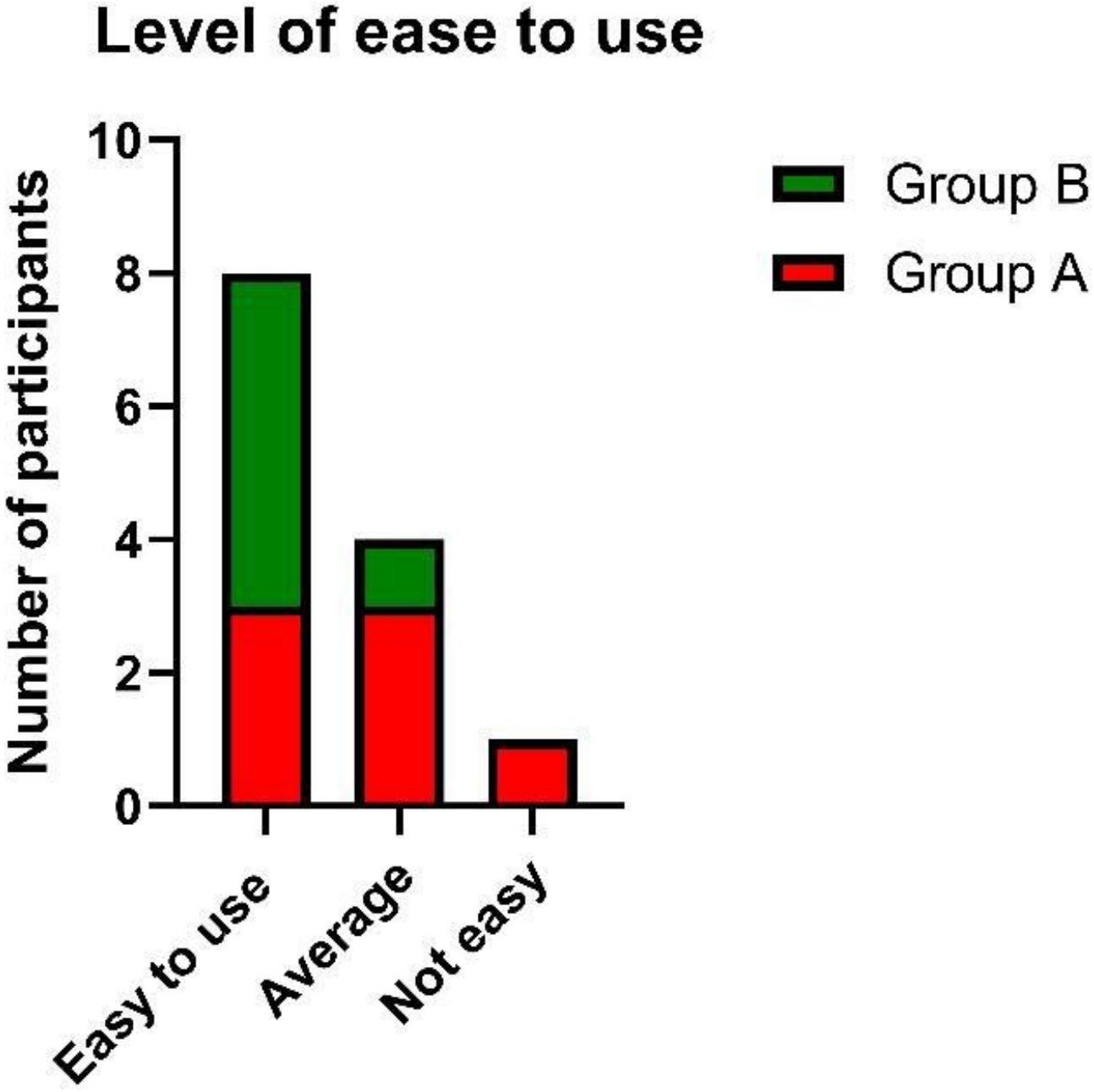
XSL•FO
RenderX

well"), while doubts regarding clinical validity were voiced primarily by group A participants ("Technology cannot replace doctors"). Notably, these reservations were more pronounced among participants from higher socioeconomic backgrounds, despite their higher educational exposure to technology—implying deeper analysis and more cynicism in technology trust among users with higher education.

Usability of Smartwatches

When asked, "Did you find the smartwatch easy to use?" Most participants rated usability as "easy" or "average," with only 1 participant reporting significant difficulty, as represented in Figure 8. Qualitative responses suggest that initial apprehension transitioned into gradual acclimatization, indicating a learning effect rather than persistent usability barriers. Participants emphasized novelty, visual familiarity with traditional watches, and perceived multifunctionality as positive features.

Figure 8. Stacked bar chart showing the number of participants rating the smartwatch as easy to use, average to use, or not easy to use, stratified by group (group A and group B).



In light of this question, a further question was asked, "What did you love about it?" In response to this question, some participants answered,

> *It is a new invention, and we were new users. At first, we were very excited after learning about such devices, and a positive impression of this device came to our heads. Furthermore, it looks quite similar to a normal watch, so we thought it offers more features than a normal watch; it is a good device.*

A participant from group B answered,

> *I was curious when I found out about this. It looked smart, and I thought I would feel smart if I wore this*

XSL•FO
RenderX

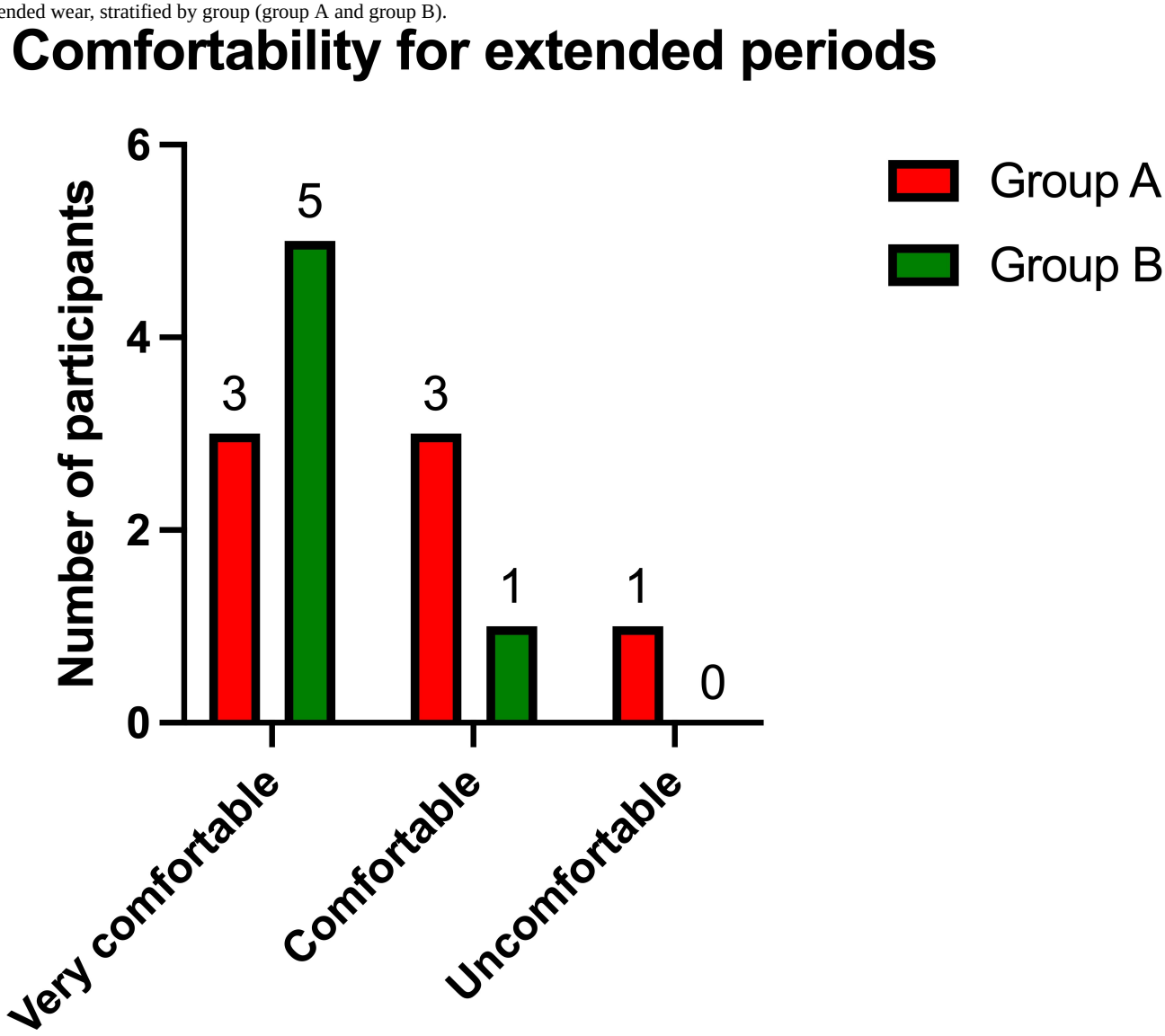
watch. Also, since it offers many health-related topics, it would be a good tool for me.

Therefore, this shows excitement in some older people about using new and "smart" technology, irrespective of their social belonging or understanding of the technology.

Comfort and Design Preferences

While overall comfort ratings were favorable, as illustrated in Figure 9, discomfort related to weight and strap design emerged as a recurring concern, particularly among group A participants. Participants expressed strong preferences for traditional materials (eg, leather or steel straps), suggesting that social acceptability and aesthetic continuity play a role in long-term adoption.

Figure 9. Stacked bar chart showing the number of participants who rated the smartwatch as very comfortable, comfortable, or uncomfortable for extended wear, stratified by group (group A and group B).

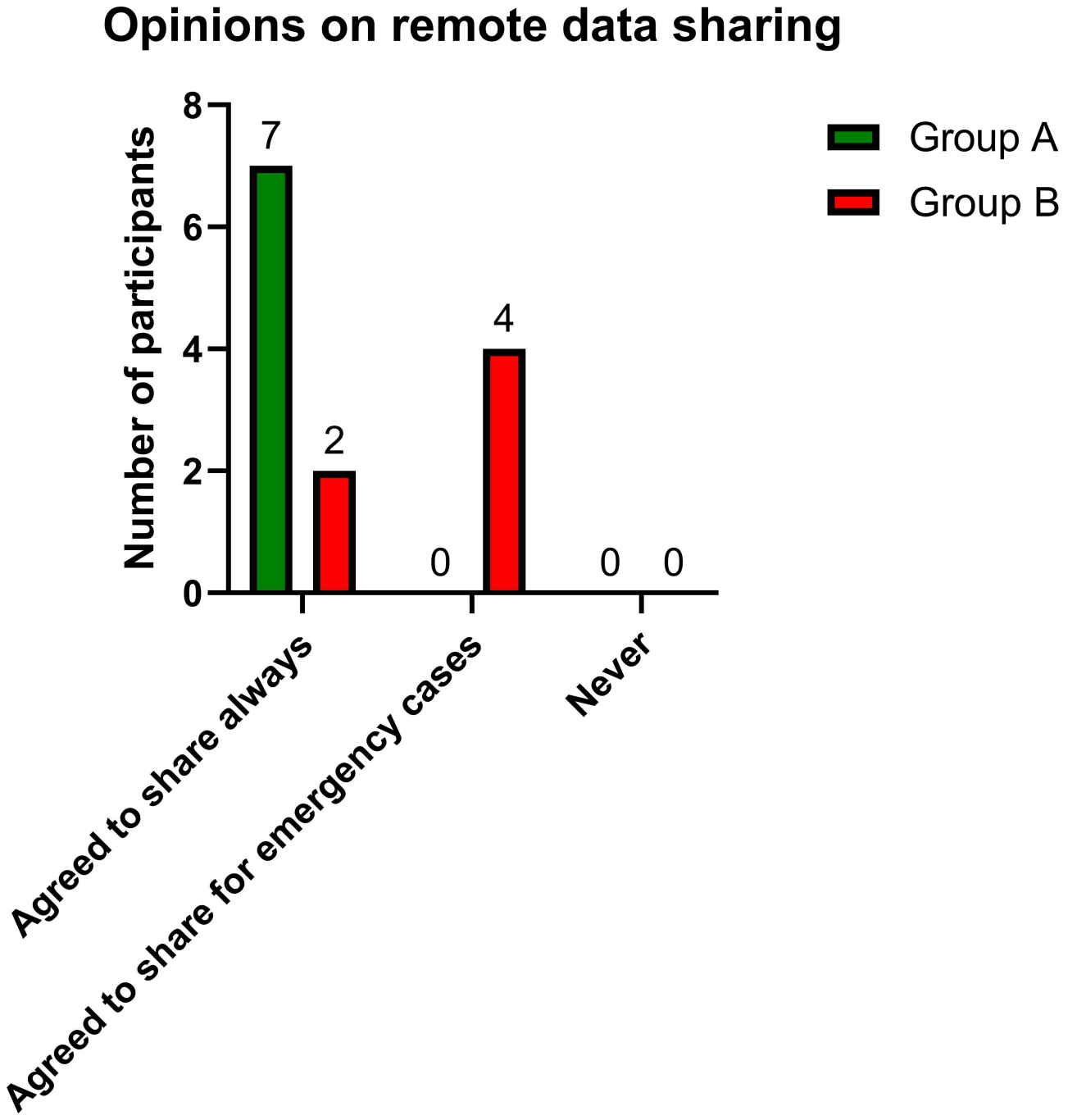


Data Sharing Preferences

The question, "Would you like your data to be continuously and automatically shared with family members and/or designated caregivers?" received mostly positive feedback. As illustrated in Figure 10, all participants in group A responded positively. In contrast, 4 out of 16 participants from group B preferred that data sharing be enabled only during emergencies, while 2 participants expressed full support for continuous sharing. This divergence highlights differing expectations regarding privacy and data oversight between the 2 groups, as the less privileged (and less educated) group identified interpretability concerns and seemed more aware of not burdening their family with their health emergencies.

XSL•FO
RenderX

Figure 10. Bar chart showing the number of participants who agreed to share health data always, only in emergency cases, or never, stratified by group (group A and group B).



Perceived Accuracy of Health Metrics

In response to the question, "How accurate did the health metrics seem to you?" participants generally perceived health metrics as reliable. Limited engagement with data—not distrust—explained neutral responses, indicating that attention and interpretation, rather than sensor credibility, influenced perceived accuracy.

Focus Group Interview 2

The second focus group interview involved 12 participants (10 males and 2 females), with 5 participants in group A and 7 in group B. This interview was conducted 14 days after the first, enabling assessment of temporal stability and experiential adaptation.

Evaluation of Smartwatch Perception

Initial impressions in interview 2 closely mirrored those from interview 1, demonstrating thematic consistency over time. However, qualitative depth increased, with participants expressing more experience-based reflections rather than speculative judgments. Curiosity remained dominant in group B, while group A responses polarized slightly between positive endorsement and skepticism.

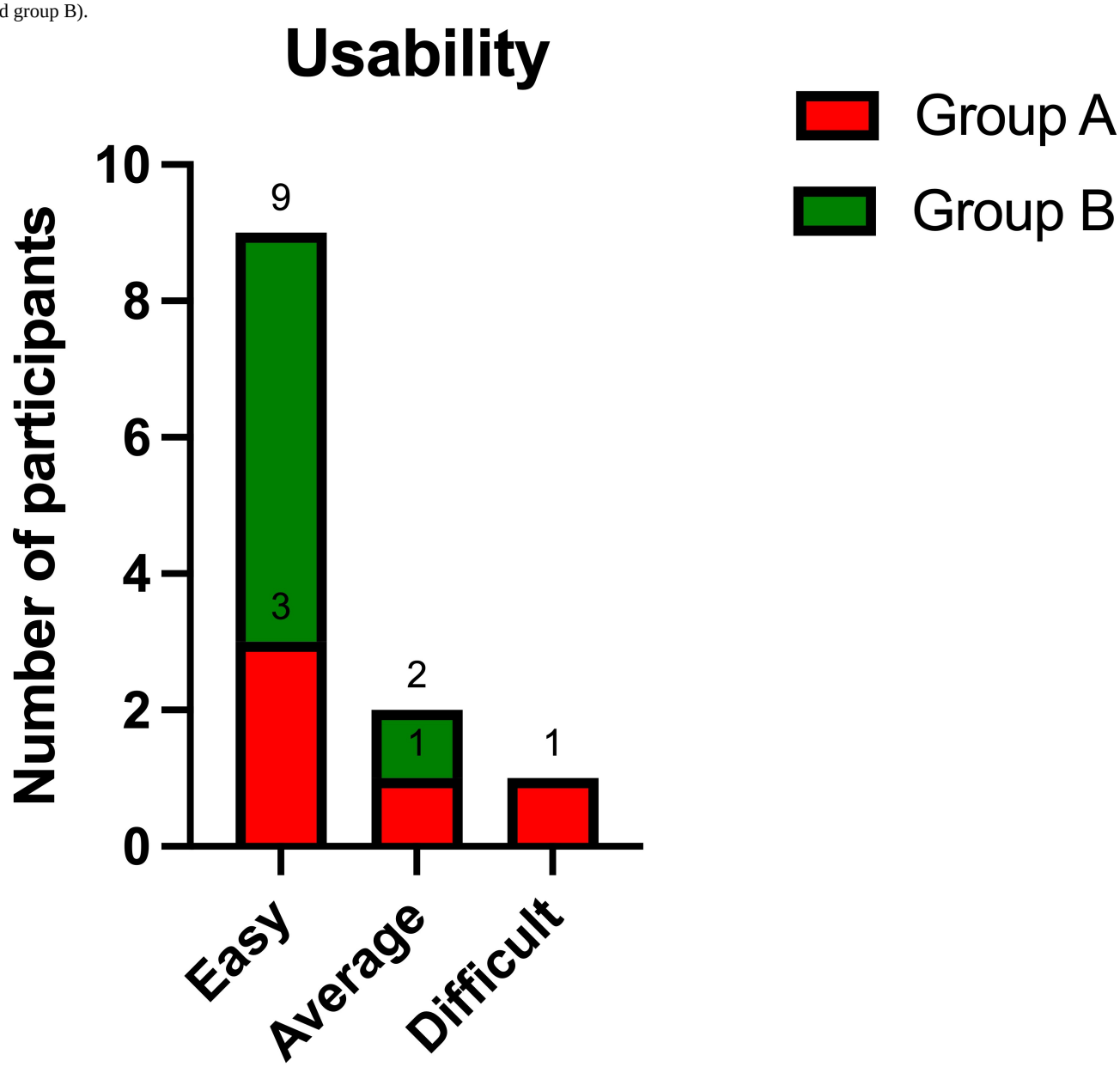
Compared with interview 1, negative perceptions were less frequent and more contextualized, suggesting that prolonged exposure mitigated abstract concerns while amplifying practical considerations.

XSL•FO
RenderX

Ease of Use: Reduced Variability Over Time

Usability ratings improved overall in interview 2, particularly within group B, as depicted in Figure 11. The reduction in "average" or negative usability responses indicates learning consolidation rather than mere tolerance. Difficulty reported by 1 older participant (88 y) underscores the influence of advanced age on interaction with wearable technologies, highlighting the need for age-adaptive interfaces.

Figure 11. Stacked bar chart showing the number of participants who rated smartwatch use as easy, average, or difficult, stratified by group (group A and group B).

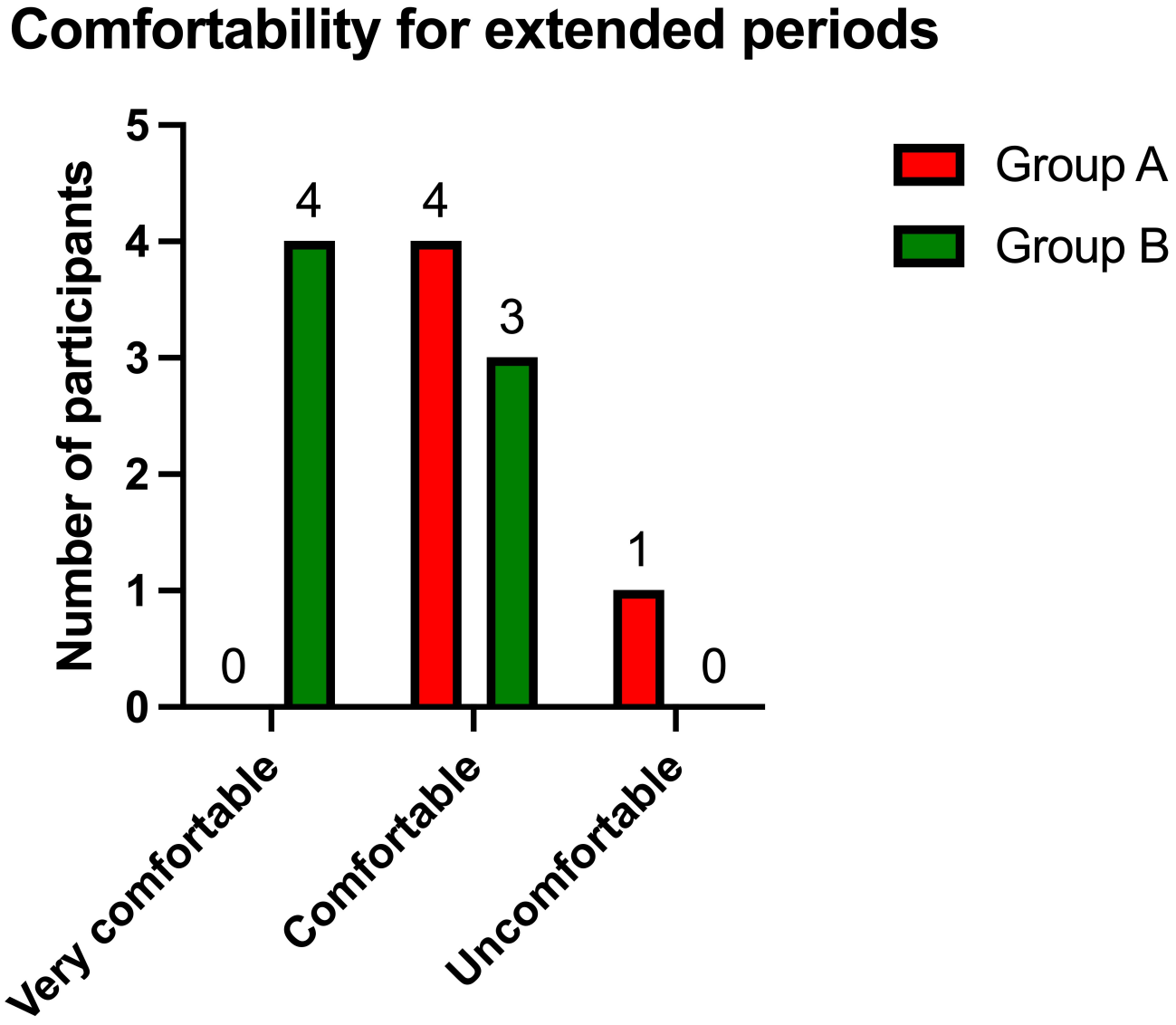


Comfort: Consistency Across Interviews

Figure 12 illustrates that comfort ratings in interview 2 remained stable relative to interview 1, reinforcing the robustness of this finding. Importantly, fewer participants raised new complaints, suggesting that comfort-related perceptions plateaued early and did not deteriorate with extended wear.

XSL•FO
RenderX

Figure 12. Stacked bar chart showing the number of participants who rated the smartwatch as very comfortable, comfortable, or uncomfortable for extended wear, stratified by group (group A and group B).

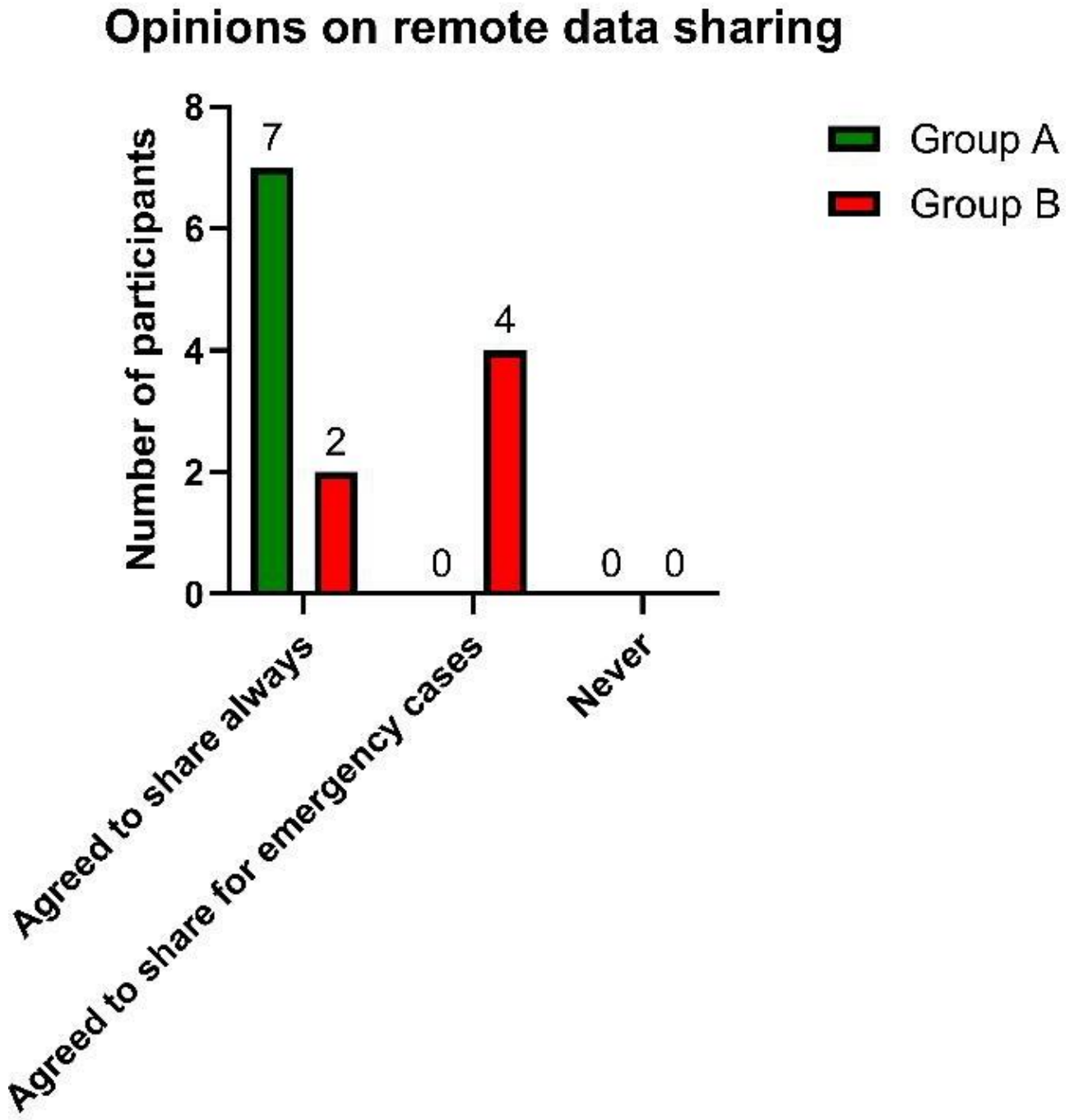


Data Sharing Preferences: Group Divergence Intensifies

A clearer divergence in data sharing preferences emerged in interview 2, as illustrated in Figure 13. All group A participants supported continuous data sharing, whereas the majority of group B participants favored emergency-only sharing. Compared with interview 1, these preferences were articulated with greater certainty, indicating preference crystallization over time rather than ambivalence.

XSL•FO
RenderX

Figure 13. Bar chart showing the number of participants who agreed to share health data always, agreed to share only in emergency cases, or never agreed to share, stratified by group (group A and group B).



Perceived Accuracy of Health Metrics

Confidence in metric accuracy increased slightly in interview 2, particularly among group B participants. Overall, the responses were mostly neutral and a bit vague at this phase, again reflecting lack of attention rather than distrust, reinforcing findings from interview 1.

Comparative Analysis and Analytical Power Considerations

Cross-Interview Thematic Stability

Across both focus group interviews, dominant themes, including usability adaptation, comfort acceptability, data-sharing conditionality, and perceived metric reliability, remained consistent. The repetition of themes across independent samples and time points supports thematic saturation, enhancing analytical power despite modest sample sizes.

Group-Level Comparative Strength

Although formal statistical power calculations are not applicable to qualitative focus group data, analytical power was strengthened through:

1. Replication of findings across 2 interview rounds,
2. Consistent group-level contrasts (higher vs lower socioeconomic status), and
3. Convergence between qualitative insights and descriptive survey trends.

XSL•FO
RenderX

The observed pattern, where participants from lower socioeconomic backgrounds adopted the smartwatch more readily and expressed fewer aesthetic or conceptual reservations, was robust across both interviews. This repeated contrast increases confidence that the finding reflects a genuine behavioral and perceptual trend rather than sampling variability.

Limitations of Quantitative Inference

While descriptive counts are reported with consistency, the study is not powered for inferential statistical comparisons between groups due to the small number of participants. Group differences should therefore be interpreted as qualitative contrasts supported by frequency patterns, not as statistically generalizable effects.

Implications for Mixed Methods Interpretation

The focus group findings provide explanatory depth for survey results, particularly by uncovering skepticism and aesthetic concerns that were not captured quantitatively. The consistency of these findings across 2 interviews enhances credibility and supports their use in informing design recommendations and future hypothesis-driven studies.

Discussion

Principal Findings

This study was guided by 3 overarching hypotheses. First (hypothesis 1), older adults' smartwatch usage patterns would stabilize over time, with selective increases in feature engagement following continuous support, exposure, and training. Second (hypothesis 2), training and ongoing support would reduce perceived difficulty of use and improve long-term adoption. Third (hypothesis 3), socioeconomic and educational background would significantly shape usability experiences and data-sharing preferences.

Overall, the findings provide strong support for all 3 hypotheses.

First, consistent with hypothesis 1, basic health-monitoring features—particularly heart rate tracking—emerged as the most used and preferred functionality across all study phases, increasing from 68% (13/19) in the first phase to 75% (15/20) in the last phase of the final long-term survey. In contrast, more complex features, such as stress and respiratory rate monitoring, showed persistently lower engagement. This stable preference hierarchy underscores the importance of prioritizing core, easily interpretable health metrics in wearable interfaces for older adults, while introducing advanced features progressively as familiarity increases.

Second, longitudinal statistical analyses partially supported hypothesis 2. Across the 3 long-term study DPs, paired *t* test analyses (Table 2) indicated that most user experience dimensions remained stable over time. No significant differences were observed for daily usage, comfort ratings, most liked features, or clarity of instructions across any pairwise comparisons (all $P>.05$). In contrast, significant temporal changes were detected for feature usage, with increased differentiation between DP2 and DP3 ($t_8=-4.16$, $P<.001$) and between DP1 and DP3 ($t_8=-2.53$, $P=.04$). Perceived difficulty of use also showed change over time, reaching significance between DP2 and DP3 ($t_8=-4.29$, $P=.01$) and marginal significance between DP1 and DP3 ($t_8=-2.31$, $P=.05$). Complementary chi-square analyses (Table 9) largely supported these findings, revealing no significant associations between time point and most survey dimensions, except for clarity of instructions, which demonstrated a significant association with study phase ($\chi^2_4=10.3$, $P=.03$), and a trend toward association for comfort ratings ($\chi^2_8=13.8$, $P=.09$). Together, these results suggest overall stability in usage and satisfaction metrics over time, alongside meaningful evolution in feature engagement and perceived usability as participants gained longer-term experience with the system. Hence, our hypothesis was proved that training and support (as provided by the authors) helped the older participants to adopt and familiarize themselves with the smartwatch.

Third, findings strongly supported hypothesis 3, revealing pronounced socioeconomic differences in data-sharing preferences. Participants from higher socioeconomic and educational backgrounds expressed greater willingness to share health data continuously, whereas participants from lower socioeconomic groups predominantly preferred emergency-only sharing. Notably, reluctance to share data was often motivated not by privacy concerns alone, but by a desire to avoid burdening family members perceived as less technologically or medically informed. This highlights the need for flexible, customizable privacy and data-sharing controls that account for diverse social and familial contexts in LMIC settings.

Beyond these hypotheses, qualitative insights revealed a broader trajectory of adoption. Initial skepticism, particularly regarding accuracy and usefulness, gradually gave way to confidence and routine integration of the smartwatch into daily life. Participants consistently emphasized the importance of familiar physical design elements, such as leather or steel straps, suggesting that aesthetic familiarity plays a critical role in lowering psychological barriers to adoption among older adults.

Key Contributions

This study makes several novel contributions to wearable health technology literature. First, it presents one of the few long-term, mixed methods usability evaluations of smartwatches among older adults in an LMIC context, specifically in Bangladesh. Second, the study uniquely incorporates 3 longitudinal survey cycles and 2 focus group interviews, enabling a nuanced understanding of adaptation over time rather than short-term acceptance alone. Third, by stratifying participants according to socioeconomic and educational background, the study reveals how adoption, usability challenges, and privacy preferences diverge across demographic groups—an aspect largely overlooked in previous work. Finally, the findings extend usability discourse beyond daily use to include cultural norms, family dynamics, and aesthetic preferences, offering a more holistic framework for designing inclusive wearable technologies for aging populations.

Conclusion

This study investigated smartwatch usability, adoption, and perception among older adults through short-term surveys,

XSL•FO
RenderX

long-term follow-up surveys, and focus group interviews. Findings demonstrate that while initial engagement was driven by curiosity and exploration, long-term use depended on comfort, usability, and relevance of features. Core health-monitoring functions, such as heart rate tracking, remained central to adoption, whereas complex features faced usability barriers despite training.

Over time, participants became more comfortable and confident with smartwatch use, validating the importance of continuous exposure, structured support, and intuitive design. Comfort-related concerns—particularly device weight and strap material, along with strong preferences for culturally familiar aesthetics, highlight the need for age- and context-sensitive hardware designs. Participants also expressed demand for features directly aligned with prevalent health concerns, including blood pressure monitoring, diabetes tracking, and caregiver-linked data sharing.

Socioeconomic and educational factors significantly shaped user experiences, particularly in relation to data-sharing preferences and perceived responsibility toward family members. These findings emphasize that successful wearable adoption among older adults, especially in the LMIC context, requires not only technical usability but also sensitivity to social values, privacy expectations, and cultural norms.

In conclusion, this study illustrates the evolving relationship between older adults and wearable health technologies, demonstrating both their promise and their limitations. By foregrounding longitudinal adaptation, socioeconomic diversity, and cultural context, the findings offer actionable design and deployment insights to support more equitable, acceptable, and sustainable adoption of smartwatches among aging populations.

Limitations

A potential limitation of this study lies in the limited size and diversity of the participant sample. A significant proportion of the participants came from similar occupational backgrounds. Although the study targeted individuals from lower-income groups, it did not include any participants who were completely illiterate. Furthermore, due to unavoidable circumstances, such as illness or death, several participants either withdrew from the study or missed certain data collection points, resulting in some discrepancies in the dataset. To maintain data balance, some participants had to be excluded from the focus group sessions. Additionally, the sample exhibited minor gender biases. While the initial male-to-female ratio was satisfactory, a number of female participants later withdrew from the study due to health concerns. The absence of representation from more diverse gender identities and from individuals who are enthusiastic about technology but were not recruited may have impacted the inclusivity of the study. Future research should aim to include participants across a broader spectrum of demographics, regardless of gender, and with varying levels of technological literacy, including individuals who are entirely illiterate.

Acknowledgments

This research project was supported by the North South University (NSU), Office of Research, and Department of Electrical and Computer Engineering. We also thank the North South University Institutional Review Board (IRB) for providing the Ethical approval (IRB 2023/OR- NSU/IRB/1231). We have to mention in acknowledgment that large language model tools were used specifically for bringing more clarity and refinement in writing in some sections, although the results and content were entirely the output of our study. We also sincerely acknowledge and appreciate the guidance and mentorship of Dr Sharifa Sultana, Assistant Professor, Computer Science, University of Illinois Urbana-Champaign, and Dr Syed Ishtiaque Ahmed, Associate Professor, Department of Computer Science, University of Toronto. Finally, we are very thankful to the entire community of older adult study participants from old age homes like Probeen Nibash and neighborhoods.

Funding

This research project was supported by the NSU CTRGC research grant fund CTRG-23-SEPS-18 from North South University.

Authors' Contributions

FM conceived the idea, designed and supervised the study, and contributed to data analysis and manuscript preparation. FR and NM contributed equally to recruitment of participants, data collection, analysis, and manuscript preparation. SAI helped with recruitment of participants and data collection.

Conflicts of Interest

None declared.

References

1. Kekade S, Hseieh CH, Islam MM, et al. The usefulness and actual use of wearable devices among the elderly population. Comput Methods Programs Biomed 2018 Jan;153:137-159. [doi: 10.1016/j.cmpb.2017.10.008] [Medline: 29157447]
2. Chandrasekaran R, Katthula V, Moustakas E. Too old for technology? Use of wearable healthcare devices by older adults and their willingness to share health data with providers. Health Informatics J 2021;27(4):14604582211058073. [doi: 10.1177/14604582211058073] [Medline: 34802315]

XSL•FO
RenderX

3. Mather M, Jacobsen A, Pollard KM. Aging in the United States. Popul Bull 2015;70(2) [FREE Full text]
4. Seneviratne S, Hu Y, Nguyen T, et al. A survey of wearable devices and challenges. IEEE Commun Surv Tutorials 2017;19(4):2573-2620. [doi: 10.1109/COMST.2017.2731979]
5. Alharbi M, Bauman A, Neubeck L, Gallagher R. Validation of Fitbit-Flex as a measure of free-living physical activity in a community-based phase III cardiac rehabilitation population. Eur J Prev Cardiol 2016 Sep;23(14):1476-1485. [doi: 10.1177/2047487316634883] [Medline: 26907794]
6. Ahmad A, Rasul T, Yousaf A, Zaman U. Understanding factors influencing elderly diabetic patients' continuance intention to use digital health wearables: extending the technology acceptance model (TAM). Journal of Open Innovation: Technology, Market, and Complexity 2020 Sep;6(3):81. [doi: 10.3390/joitmc6030081]
7. Lim S, Kang SM, Kim KM, et al. Multifactorial intervention in diabetes care using real-time monitoring and tailored feedback in type 2 diabetes. Acta Diabetol 2016 Apr;53(2):189-198. [doi: 10.1007/s00592-015-0754-8] [Medline: 25936739]
8. Patel S, Park H, Bonato P, Chan L, Rodgers M. A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil 2012 Apr 20;9(1):21. [doi: 10.1186/1743-0003-9-21] [Medline: 22520559]
9. Baig MM, Afifi S, GholamHosseini H, Mirza F. A systematic review of wearable sensors and IoT-based monitoring applications for older adults - a focus on ageing population and independent living. J Med Syst 2019 Jun 15;43(8):233. [doi: 10.1007/s10916-019-1365-7] [Medline: 31203472]
10. Jang IY, Kim HR, Lee E, et al. Impact of a wearable device-based walking programs in rural older adults on physical activity and health outcomes: cohort study. JMIR Mhealth Uhealth 2018 Nov 21;6(11):e11335. [doi: 10.2196/11335] [Medline: 30463838]
11. Wang Z, Yang Z, Dong T. A review of wearable technologies for elderly care that can accurately track indoor position, recognize physical activities and monitor vital signs in real time. Sensors (Basel) 2017;17(2):341. [doi: 10.3390/s17020341]
12. Farivar S, Abouzahra M, Ghasemaghaei M. Wearable device adoption among older adults: a mixed-methods study. Int J Inf Manage 2020 Dec;55:102209. [doi: 10.1016/j.ijinfomgt.2020.102209] [Medline: 32834339]
13. Li C, Lee CF, Xu S. Stigma threat in design for older adults: exploring design factors that induce stigma perception. Int J Des 2020;14(1):51-64 [FREE Full text]
14. Li J, Ma Q, Chan AHS, Man SS. Health monitoring through wearable technologies for older adults: smart wearables acceptance model. Appl Ergon 2019 Feb;75:162-169. [doi: 10.1016/j.apergo.2018.10.006] [Medline: 30509522]
15. Tsai TH, Lin WY, Chang YS, Chang PC, Lee MY. Technology anxiety and resistance to change behavioral study of a wearable cardiac warming system using an extended TAM for older adults. PLoS ONE 2020;15(1):e0227270. [doi: 10.1371/journal.pone.0227270] [Medline: 31929560]
16. Talukder M, Sorwar G, Bao Y, Ahmed JU, Palash M. Predicting antecedents of wearable healthcare technology acceptance by elderly: a combined SEM-neural network approach. Technol Forecast Soc Change 2020 Jan;150:119793. [doi: 10.1016/j.techfore.2019.119793] [Medline: 32834137]
17. Zhu Y, Lu Y, Gupta S, Wang J, Hu P. Promoting smart wearable devices in the health-AI market: the role of health consciousness and privacy protection. JRIM 2023 Mar 21;17(2):257-272. [doi: 10.1108/JRIM-10-2021-0246]
18. Chung J, Brakey HR, Reeder B, Myers O, Demiris G. Community-dwelling older adults' acceptance of smartwatches for health and location tracking. Int J Older People Nurs 2023 Jan;18(1):e12490. [doi: 10.1111/opn.12490] [Medline: 35818900]
19. Sullivan AN, Lachman ME. Behavior change with fitness technology in sedentary adults: a review of the evidence for increasing physical activity. Front Public Health 2016;4(289):289. [doi: 10.3389/fpubh.2016.00289] [Medline: 28123997]
20. Chuah SHW, Rauschnabel PA, Krey N, Nguyen B, Ramayah T, Lade S. Wearable technologies: the role of usefulness and visibility in smartwatch adoption. Comput Human Behav 2016 Dec;65:276-284. [doi: 10.1016/j.chb.2016.07.047]
21. Gumasing M, Carrillo G, De Guzman MAA, et al. Investigating user-centric factors influencing smartwatch adoption and user experience in the Philippines. Sustainability 2024;16(13):5401. [doi: 10.3390/su16135401]
22. Ometov A, Shubina V, Klus L, et al. A survey on wearable technology: history, state-of-the-art and current challenges. Computer Networks 2021 Jul;193:108074. [doi: 10.1016/j.comnet.2021.108074]
23. Alshamari MA, Althobaiti MM. Usability evaluation of wearable smartwatches using customized heuristics and system usability scale score. Future Internet 2024;16(6):204. [doi: 10.3390/fi16060204]
24. Xu L, Pan C, Xu B, et al. A qualitative exploration of a user-centered model for smartwatch comfort using grounded theory. International Journal of Human–Computer Interaction 2025 Jan 17;41(2):1091-1106. [doi: 10.1080/10447318.2024.2313273]
25. Kononova A, Li L, Kamp K, et al. The use of wearable activity trackers among older adults: focus group study of tracker perceptions, motivators, and barriers in the maintenance stage of behavior change. JMIR Mhealth Uhealth 2019 Apr 5;7(4):e9832. [doi: 10.2196/mhealth.9832] [Medline: 30950807]
26. Berkowsky RW, Sharit J, Czaja SJ. Factors predicting decisions about technology adoption among older adults. Innovation in Aging 2017 Nov 1;1(3). [doi: 10.1093/geroni/igy002] [Medline: 31517933]
27. Atkins S, Lewin S, Smith H, Engel M, Fretheim A, Volmink J. Conducting a meta-ethnography of qualitative literature: lessons learnt. BMC Med Res Methodol 2008 Apr 16;8:21. [doi: 10.1186/1471-2288-8-21] [Medline: 18416812]
28. Domingos C, Costa P, Santos NC, Pêgo JM. Usability, acceptability, and satisfaction of a wearable activity tracker in older adults: observational study in a real-life context in northern Portugal. J Med Internet Res 2022 Jan 26;24(1):e26652. [doi: 10.2196/26652] [Medline: 35080503]

XSL•FO
RenderX

29. Moore K, O'Shea E, Kenny L, et al. Older adults' experiences with using wearable devices: qualitative systematic review and meta-synthesis. JMIR Mhealth Uhealth 2021 Jun 3;9(6):e23832. [doi: 10.2196/23832] [Medline: 34081020]
30. Thilo FJS, Hahn S, Halfens RJG, Schols J. Usability of a wearable fall detection prototype from the perspective of older people-a real field testing approach. J Clin Nurs 2019 Jan;28(1-2):310-320. [doi: 10.1111/jocn.14599] [Medline: 29964344]
31. Keogh A, Argent R, Anderson A, Caulfield B, Johnston W. Assessing the usability of wearable devices to measure gait and physical activity in chronic conditions: a systematic review. J Neuroeng Rehabil 2021 Sep 15;18(1):138. [doi: 10.1186/s12984-021-00931-2] [Medline: 34526053]
32. Allet L, Knols RH, Shirato K, de Bruin ED. Wearable systems for monitoring mobility-related activities in chronic disease: a systematic review. Sensors (Basel) 2010;10(10):9026-9052. [doi: 10.3390/s101009026] [Medline: 22163393]
33. Muñoz Esquivel K, Gillespie J, Kelly D, et al. Factors influencing continued wearable device use in older adult populations: quantitative study. JMIR Aging 2023 Jan 19;6(1):e36807. [doi: 10.2196/36807] [Medline: 36656636]
34. Duarte RP, Alves VNN, Mota M. Designing inclusive smartwatch interfaces: guidelines for enhancing usability and adoption among older adults. IEEE Access 2025;13:64901-64929. [doi: 10.1109/ACCESS.2025.3559726]
35. Sanders EC, Boot WR. Usability of the smartwatch as a prospective memory aid for cognitively diverse older adults. J Appl Gerontol 2026 Feb;45(2):382-392. [doi: 10.1177/07334648251338876] [Medline: 40432385]
36. Torous J, Linardon J, Goldberg SB, et al. The evolving field of digital mental health: current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality. World Psychiatry 2025 Jun;24(2):156-174. [doi: 10.1002/wps.21299] [Medline: 40371757]
37. Edriss S, Romagnoli C, Caprioli L, et al. The role of emergent technologies in the dynamic and kinematic assessment of human movement in sport and clinical applications. Appl Sci (Basel) 2024;14(3):1012. [doi: 10.3390/app14031012]
38. Vrančić A, Zadravec H, Orehovački T. The role of smart homes in providing care for older adults: a systematic literature review from 2010 to 2023. Smart Cities 2024;7(4):1502-1550. [doi: 10.3390/smartcities7040062]

Abbreviations

DP: data point
HIPAA: Health Insurance Portability and Accountability Act
IRB: institutional review board
LMIC: low- and middle-income country
SpO_2: oxygen saturation level monitoring

Edited by Y Luo; submitted 07.May.2025; peer-reviewed by Q Fang, R Marshall; revised version received 18.Feb.2026; accepted 27.Feb.2026; published 30.Jun.2026.

Please cite as:
Mahzabeen F, Rabby F, Mowla N, Ibnat SA
Usability and Adoption of Smartwatches by Older Adults in Bangladesh: User Study
JMIR Aging 2026;9:e76404
URL: https://aging.jmir.org/2026/1/e76404
doi:10.2196/76404

© Fariah Mahzabeen, Fazlea Rabby, Neazmul Mowla, Sadia Afra Ibnat. Originally published in JMIR Aging (https://aging.jmir.org), 30.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Tablet Computer–Based Food Record for the Self-Assessment of Nutritional Intake in Patients Undergoing Geriatric Rehabilitation: Prospective Pilot Feasibility Study

Julia Berndt[1,2]; Patrick Elfert[3]; Marco Eichelberg[3]; Juergen Martin Bauer[4]; Andreas Hein[1,3]; Rebecca Diekmann[2]

[1]Assistance Systems and Medical Device Technology, Department of Health Services Research, Carl von Ossietzky University, Oldenburg, Germany

[2]Nutrition and Physical Function in Older Adults, Department of Health Services Research, Carl von Ossietzky University, Ammerländer Heerstraße 140, Oldenburg, Germany

[3]OFFIS-Institute for Information Technology, Oldenburg, Germany

[4]Center for Geriatric Medicine, Agaplesion Bethanien Hospital, Heidelberg, Germany

Corresponding Author:
Rebecca Diekmann
Nutrition and Physical Function in Older Adults, Department of Health Services Research, Carl von Ossietzky University, Ammerländer Heerstraße 140, Oldenburg, Germany

Abstract

Background: Nutritional status is an influential factor for functional status and rehabilitation outcomes in patients undergoing geriatric rehabilitation. Although there is evidence for the potential of eHealth interventions in patients undergoing geriatric rehabilitation in general, the evidence for eHealth interventions with a focus on nutrition is scarce. In other target groups with older people, eHealth applications to support nutrition, such as computer-based food records, have been used successfully.

Objective: Therefore, the aim of this study was to verify whether it is feasible for patients undergoing geriatric rehabilitation to independently use a tablet computer−based food record (e-food record) to document their food and beverage intake. The e-food record was developed in advance and tailored to the age- and disease-specific needs of patients undergoing geriatric rehabilitation.

Methods: This prospective pilot study investigated the general feasibility of an e-food record in older adults (≥70 y) in a geriatric rehabilitation center in Germany. It was tested whether the e-food record could be independently used by the participants over 3 days. Furthermore, the usability of the e-food record was assessed by the System Usability Scale (0 - 100 points) after usage. To compare nutritional data, the participants recorded their consumption of food and beverages by the e-food record and by a 24-hour recall for the same time period, and the mean difference was calculated as follows: the value of the 24-hour recall minus the value of the e-food record. As the study was characterized as a pilot, the data analysis was descriptive.

Results: Seventeen out of 25 patients (n=6, 35.3% female, mean age 79.5, SD 3.7 y) maintained the e-food record independently over the study period. The mean System Usability Scale score of the e-food record was 76.0 (SD 11.3) points. Datasets of 9 out of 17 participants (n=5, 55.6% female, mean age 78.2, SD 2.9 y) were analyzed in terms of nutritional data. Mean differences in energy, protein, and fluid intake by the 24-hour recall compared to the e-food record were 4.9 (SD 10.2) kcal/kg body weight (bw), 0.1 (SD 0.3) g/kg bw, and 4.9 (SD 9.4) g/kg bw, respectively.

Conclusions: The use of an e-food record is generally feasible for patients undergoing geriatric rehabilitation characterized by low technical experience, high mean age, and a high rate of functional impairment. Lower intake levels were observed for the e-food record compared to the 24-hour recall with regard to energy, protein, and fluid intake. Aspects for further development of the e-food record were identified to enable evaluation on a larger sample. Following successful evaluation, the e-food record could be used within nutrition therapy in the future to increase the efficiency of the nutritional therapy process.

Trial Registration: German Clinical Trials Register DRKS 00014370; https://drks.de/search/de/trial/DRKS00014370

(JMIR Aging 2026;9:e84653) doi:10.2196/84653

KEYWORDS

eHealth; nutrition; geriatric rehabilitation; tablet computer; food record; older adults; technology; nutrition app; feasibility study

XSL•FO
RenderX

Introduction

Background

Older people represent a growing population in Western countries due to sociodemographic developments. As a consequence, the number of people who have therapeutic or nursing needs is also increasing sharply [1]. Therefore, geriatric rehabilitation is becoming increasingly important in the medical care of older people. Patients undergoing geriatric rehabilitation are characterized by a reduced functional status due to disabling impairments. These are often combined with the presence of multimorbidity and geriatric syndromes such as sarcopenia or frailty [2-4]. The aim of geriatric rehabilitation is to improve functional resources in order to enable patients to return to their home environments [5].

Nutritional status is a well-known factor that is interrelated with physical function, frailty, and sarcopenia, which are associated with negative effects on functional dependence in patients undergoing geriatric rehabilitation [6-8]. However, malnutrition and the risk of malnutrition are widespread conditions among these patients, with a prevalence of 5% to 20% for malnutrition and 40% to 54% for the risk of malnutrition [8]. Because there is profound evidence, malnutrition is treated as part of the structured rehabilitation program within nutritional therapy [9]. The focus here is on optimizing energy and protein intake and on weight stabilization to generate a positive effect on muscle building and muscle maintenance [10]. The amount of energy and protein intake consumed is determined by the nutritional therapist using the method of a food record. Individual nutritional recommendations for the patient are then derived and communicated on this basis. Due to financial and personnel resources, nutritional therapy is often limited in time during rehabilitation treatment, even if the literature indicates that close-meshed nutritional therapy might achieve greater success [11].

eHealth applications, such as computer-based food records, can be used to provide support by intensifying the nutrition therapy process. The time-consuming processes of recording and analyzing dietary intake data are no longer necessary for the nutrition therapist. Maintaining a computer-based food record regularly enables the diet to be monitored more consistently and proactively. Deviations from the nutrition target can be recognized earlier, and reintervention can be initiated by the nutrition therapist. In addition, maintaining the computer-based food record can increase the patient's awareness of their diet and therefore might lead to higher motivation in achieving nutritional goals. In the context of geriatric rehabilitation, advanced age, limited experience with technology, reduced visual acuity, changes in haptics, cognitive limitations, and fears in dealing with technology are characteristic challenges when using eHealth systems in this target group that must be addressed in the technical development process.

eHealth interventions in geriatric rehabilitation have already been shown to have the potential to improve rehabilitation outcomes. In a meta-analysis, the evidence for feasibility, usability, and effectiveness in the treatment of patients undergoing geriatric rehabilitation was demonstrated. However, the eHealth interventions focused predominantly on activity and physical function. No nutritional interventions were reported [12]. Only Happe et al [13,14] demonstrated in 2022 and 2023 the feasibility and usability of a tablet-based e-coach to improve nutrition and physical activity in patients undergoing geriatric rehabilitation. As far as we know, there is no research data on e-food record methods in the group of patients undergoing geriatric rehabilitation published to date.

Objectives

Therefore, the primary objective of this pilot feasibility study was to verify whether it is fundamentally possible for patients undergoing geriatric rehabilitation with low technical experience and simultaneous high postacute or chronic stress to independently use a tablet computer−based food record (e-food record) over 3 days that has been specially adapted to their needs. Secondary and exploratory objectives included the assessment of usability as well as the analysis of the recorded dietary data. In addition, the recorded food items, food portions, and nutritional intake data of the e-food record are compared to the recorded data of a 24-hour dietary recall, a commonly used reference method in clinical practice. Finally, the study aimed to identify potential areas of improvement to inform further development of the e-food record.

Methods

Participants and Study Design

Patients aged 70 years and older who were consecutively admitted to a geriatric rehabilitation ward of a rehabilitation center in the northwest of Germany were eligible for the present prospective pilot feasibility study. Exclusion criteria were (1) immobility in the sense of confinement to bed, (2) the inability to understand the study content and the course of the study, (3) difficulties in communication in the German language, and (4) participation in another study. Inpatient geriatric rehabilitation in Germany is a structured multidisciplinary program with the goal of regaining independence and functionality in older persons with reduced functional conditions due to acute or chronic medical conditions [15]. Daily therapy is administered for a period of usually 20 days [16], including medical therapy, nursing care, physiotherapy, occupational therapy, neuropsychology and social services, nutritional therapy, and speech therapy [15]. Patients undergoing geriatric rehabilitation represent a highly vulnerable and clinically burdened population, often characterized by multimorbidity, functional limitations, and increased physical and psychological stress following acute or chronic health events. Furthermore, familiarity with and affinity for digital technologies are often limited in this age group. Data for the feasibility study were collected between July and December 2018. All participants took part in the study during the geriatric rehabilitation program. Sex and age were requested from the patients and documented. Due to ethical and data protection regulations, no data were collected from patients who declined participation. Therefore, no systematic information on reasons for nonparticipation is available.

XSL•FO
RenderX

Ethical Considerations

The study was approved by the medical ethics committee of the Carl von Ossietzky Universität Oldenburg (ethical approval code 2018-029). We conducted the study in accordance with the Declaration of Helsinki. The study was a priori registered in the German Clinical Trials Register (DRKS-ID: DRKS 00014370).

Informed consent was obtained from all participants prior to participation. Pseudonymized data were collected and stored on secure, password-protected systems accessible only to the research team. Results are reported in aggregated or anonymized form. Participants did not receive any financial compensation, incentives, or other benefits for their participation in this study.

Study Assessment of Food Records

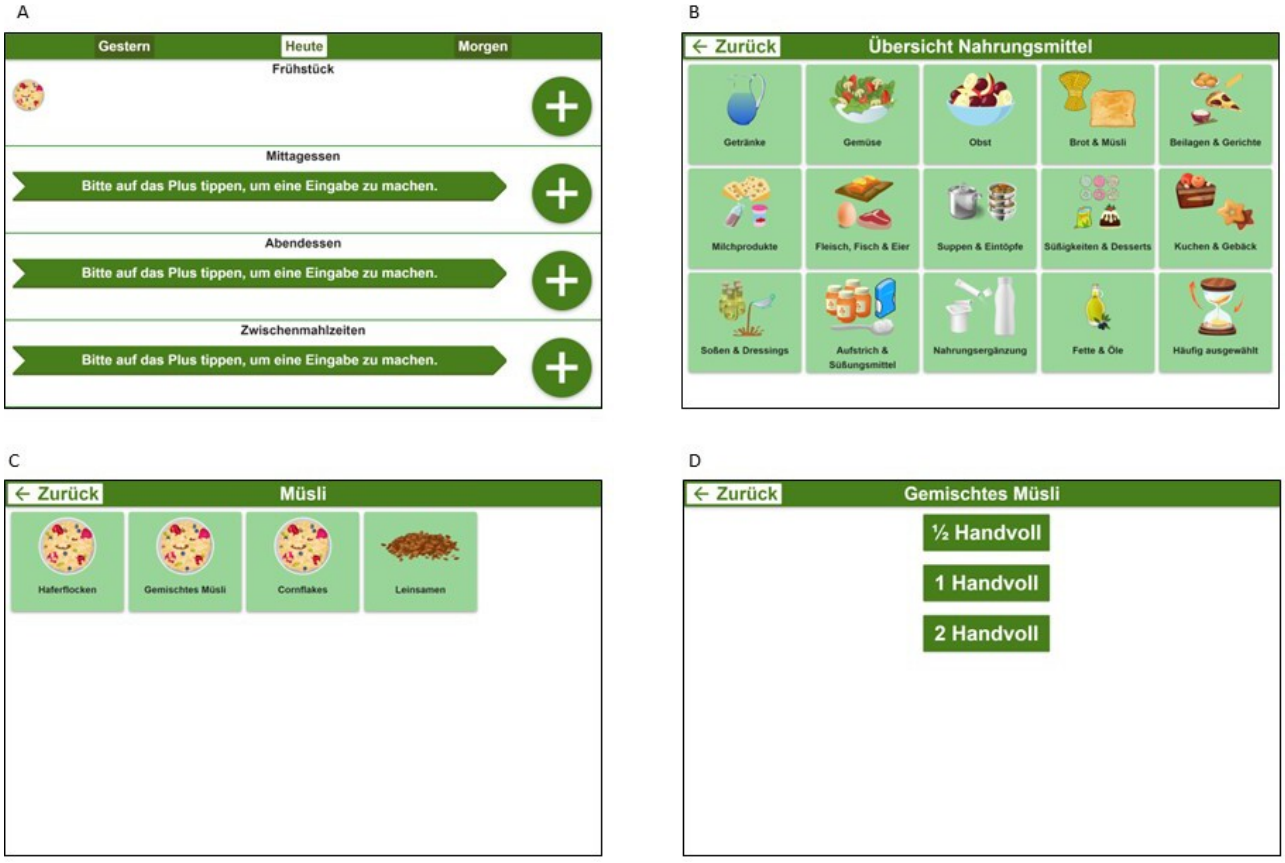
e-Food Record

The patients' food and beverage consumption was recorded by the patients themselves using an e-food record for 3 days including 2 weekdays and 1 weekend day. The e-food record used was an app-based computer program developed in close cooperation between experts in the field of nutritional counseling and app developers. It was installed on a 10.1-inch tablet Lenovo YOGA Tab 3 Plus in landscape mode. With regard to data protection, all data collected via the e-food record were stored locally on the device and transferred via secure local (offline) data transfer to a university computer after study completion. No cloud-based storage or external data transfer was used. No personal identifying information was stored within the application, and the access to study data was restricted to authorized members of the research team only. Technological aspects and the user-centered development of the e-food record were published by Elfert et al [17,18]. Design guidelines for apps for older people were key components in the development of the e-food record. The start screen of the e-food record was present at the current date and prompted the patients to make an entry for either breakfast, lunch, dinner, or a snack by touching the meal-associated plus signs. In total, 195 food symbols out of 14 food categories (eg, "bread and cereals," "vegetables," "meat, fish, and eggs," and "beverages") and 18 subcategories (eg, "bread," "buns," and "cereal" within the upper category "bread and cereals") could be selected by the patients. The view of the start screen and the selection screen of the food upper categories are shown in Figure 1.

Portions suitable for everyday use were predefined in the e-food record based on the food pyramid of the Federal Center for Nutrition [19] and could be selected as half, 1, or 2 portions. Where possible, the portion sizes referred to an easily manageable hand measure. For example, one portion of bread equaled one slice of bread about the size of the hand, and one portion of cereals, fruits, vegetables, and sweets equaled one handful. Beverages were measured in glasses, cups, or mugs. There was no option for free entries. Before the application test, the participants were given a detailed introduction to the e-food record. For this purpose, they entered a defined sample meal into the e-food record under guidance, whereby open questions regarding the handling were addressed and clarified. All data were stored locally on the tablet computer and downloaded from the device via a cable connection upon successful completion of the study.

XSL•FO
RenderX

Figure 1. Screen views of the e-food record in the original German version according to Elfert [20]. (A) Start screen with an entry of muesli under breakfast and the prompt to make an entry for lunch, dinner, or snacks by tapping on the plus symbol. (B) Screen view for the selection of a food upper group (eg, beverages, fruits, milk products, soup and stew, and fat and oil). (C) Screen view for food selection using the example of muesli. (D) Screen view for portion selection using the example of muesli. English translations are as follows: Gestern: yesterday; Heute: today; Morgen: tomorrow; Frühstück: breakfast; Mittagessen: lunch; Abendessen: dinner; Zwischenmahlzeiten: snacks; Bitte auf das Plus tippen, um eine Eingabe zu machen: Please tap the plus sign to enter an entry; Zurück: back; Übersicht Nahrungsmittel: food overview; Getränke: beverages; Gemüse: vegetables; Obst: fruits; Brot & Müsli: bread & cereals; Beilagen & Gerichte: side dishes; Milchprodukte: dairy products; Fleisch, Fisch & Eier: meat, fish & eggs; Suppen & Eintöpfe: soup & stew; Süßigkeiten & Desserts: sweets; Kuchen & Gebäck: cakes; Soßen & Dressings: sauces; Aufstrich & Süßungsmittel: spreads & sweeteners; Nahrungsergänzung: supplements; Fette & Öle: fat & oil; Häufig ausgewählt: frequently selected; Müsli: cereals; Haferflocken: oatmeal; Gemischtes Müsli: mixed granola; Cornflakes: cornflakes; Leinsamen: flaxseeds; Handvoll: handful.



24-Hour Recall—Paper-Based Food Record

In addition to the e-food record, a 24-hour recall was conducted retrospectively for each day the e-food record was also filled out. For this purpose, a prestructured paper-based template of a 24-hour recall was used to enter all consumed food and beverages of the previous day asked by a nutritionist and a member of the study staff. The prestructured 24 hour recall used here consisted of 18 food categories such as “bread,” “vegetables, salad,” “meat, fish, eggs,” “juices, water, other beverages,” etc. Each upper category contained 4 to 11 single foods. Predefined portions were counted as a tally for each of the 3 days. Food not listed was entered as free input at the end of the 24-hour recall.

Feasibility

Feasibility, defined as the independent use of the e-food record over 3 days, was assessed during a 3-day application test in inpatient geriatric rehabilitation. Feasibility was operationalized by the completion of the recording period and quantified using the dropout rate.

Usability and Technology Commitment

As secondary outcomes, usability and technology commitment were assessed. Following the use of the e-food record, the user experience was reflected upon. Usability was measured using the System Usability Scale (SUS, 0 - 100 points) [21] in its German translation [22], with higher scores indicating better usability. The SUS is a standardized questionnaire with 10 items that are rated on a Likert scale from 1 to 5. A value of over 68 is considered above average. The SUS scores determined were grouped together and classified using adjectives according to Bangor et al [23] as follows: best imaginable (100-84.1 points), excellent (84-80.8 points), good (80.7-71.1 points), okay (71-51.7 points), poor (51.6-25.1 points), and worst imaginable (≤25 points).

Prior to the application test, the technology commitment was assessed according to Neyer et al, with a standardized 12-item questionnaire (12 - 60 points). The items were rated on a 5-point Likert scale and reflect the areas of personal contact, interest, and expectations for the use of technologies [24].

XSL•FO
RenderX

Exploratory Analysis: Comparison of E-Food Record and 24-Hour Recall

Nutritional Data

As an exploratory objective, dietary data obtained from the e-food record were compared with those from a 24-hour recall. The nutritional values of energy, protein, and fluid intake of the e-food record dataset and the 24-hour recall were adjusted for body weight (bw) and calculated as grams (g) divided by bw (kg) in a comparable manner for both protocol methods on the basis of the German Federal Food Code. Fluid intake was expressed in g/kg bw and not in mL/kg bw, as provided by the nutritional analysis software DGE Expert (version 1.3.14.1, Deutsche Gesellschaft für Ernährung) used in this study. The mean differences of energy, protein, and fluid intake between the e-food record and 24-hour recall were calculated by subtracting the values of the e-food record from the 24-hour recall, respectively. In addition, agreement in recorded food and beverage entries and portion sizes between both methods was explored. For this purpose, entries were grouped into predefined food categories.

Food and Beverage Portion Entries

After the 3-day application test, the entries of the recorded food and beverages along with the corresponding portion entries from the e-food record and the 24-hour recall were extracted. The recorded data of the 2 methods were grouped into the upper food categories of “beverages,” “vegetables,” “fruits,” “meat, fish, and eggs,” “dairy products,” “soup and stew,” “side dishes,” “sauces,” “sweets,” “cakes,” “bread and cereals,” “spreads,” “fat and oil,” “sweeteners,” and “supplements,” respectively.

Statistical Analysis

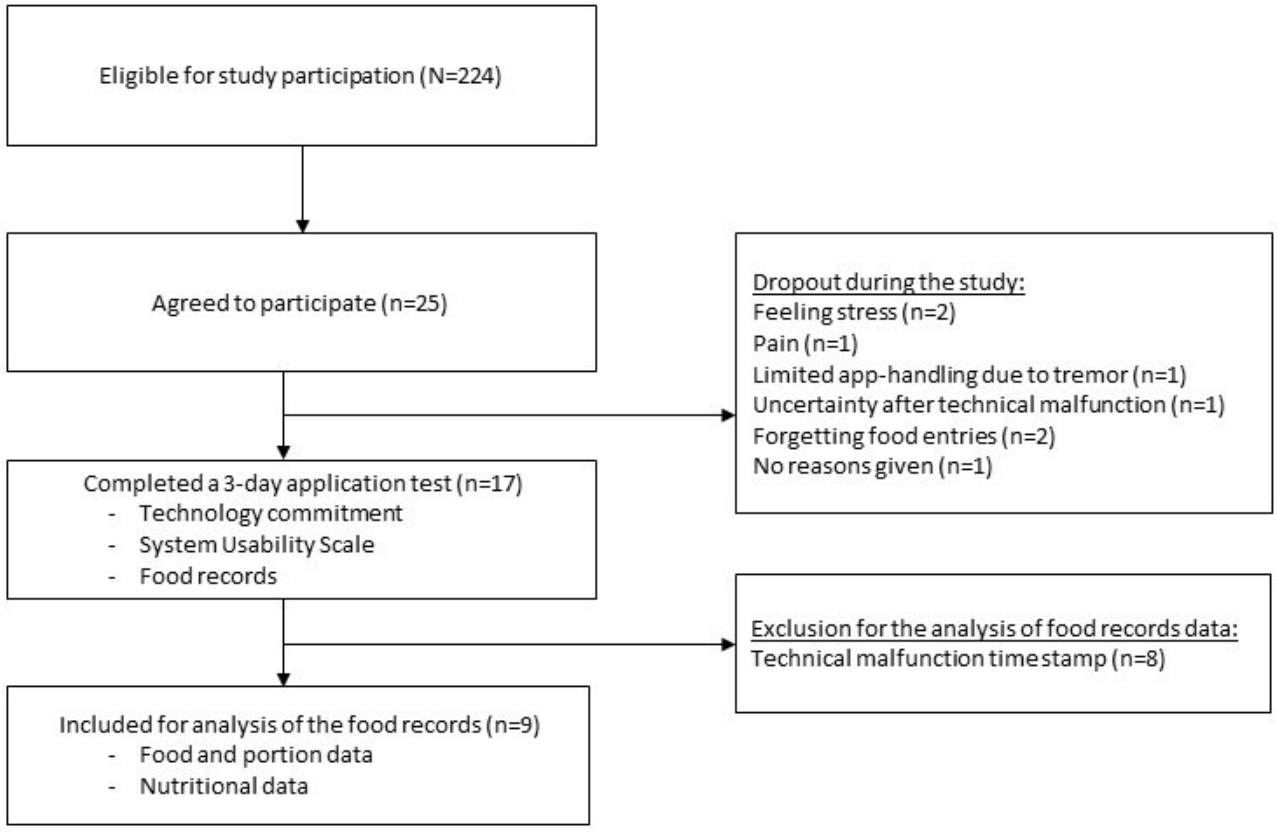
The statistical analysis was conducted using SPSS Statistics version 29 (IBM SPSS Inc). Descriptive statistics are presented as mean values and SD, median and IQR, or as total numbers and percentages. Due to the pilot characteristics of the study, the statistical analysis refers to an explorative analysis, and no interference statistics were performed.

Results

Feasibility

Twenty-five out of 224 (11%) patients who were eligible for inclusion were enrolled in the study. Seventeen out of 25 (68 %) patients completed the study and were included in the analysis of usability and technology commitment, while 8 patients dropped out of the study prematurely. Detailed information about the reasons for dropout is shown in Figure 2. During the study, a technical malfunction occurred in which the day selection did not automatically jump to the current day, contrary to what the study participants were told. As a result, the entries could no longer be validly assigned to the individual protocol days. Due to this, 9 out of 17 datasets were included in the analysis of the nutritional data (group A), while 8 datasets could not be analyzed (group B).

Figure 2. Flowchart for inclusion and dropout of study participants.



XSL•FO
RenderX

Participant Characteristics

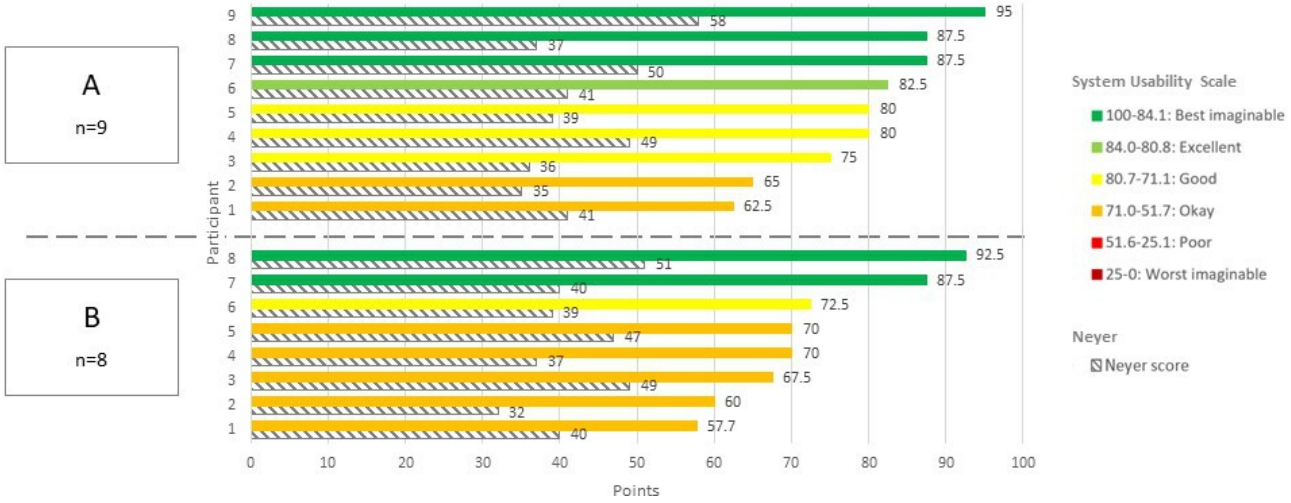
The mean age of all participants (n=17) who completed the study was 79.5 (SD 3.7) years. Six (35.3%) of them were female. The mean age of the participants who completed the 3-day application test (group A, n=9) was 78.2 (SD 2.9) years. Five out of 9 (55.6%) participants were female. The average age of the participants whose data were included for the analysis of the food records (group B, n=8) was 81.0 (SD 4.0) years, and 1 (12.5%) participant was female.

Usability and Technology Commitment

The SUS was rated 76.0 (SD 11.3) points on average by all participants. It was higher in participants of group A with a mean score of 79.4 (SD 10.0) points compared to 72.2 (SD 11.4) points in participants of group B.

The mean technology commitment of all participants was 42.4 (SD 6.8) points. For group A, the mean technology commitment was 42.9 (SD 7.3) points compared to group B with 41.9 (SD 6.1) points. The single-item responses of the Likert scale (1 - 5 points) were rated on average with 3.5 (SD 1.8) points (n=17). Detailed information on the values of the SUS and technology commitment per participant are represented in Figure 3.

Figure 3. Usability and technology commitment per participant. Adjectives were used according to Bangor et al [23].



Comparison of e-Food Record and 24-Hour Recall

Nutritional Data

The data of 9 participants (group A) were included in the following analysis. The mean energy intake determined with the e-food record was 30.3 (SD 10.2) kcal/kg bw compared to 35.0 (SD 11.0) kcal/kg bw determined with the 24-hour recall. The protein intake determined via the e-food record was 1.2 (SD 0.4) g/kg bw, which was lower than the protein intake determined via the recall at 1.3 (SD 0.4) g/kg bw. Figures 4-6 show the individual participant data of the comparison of the nutritional values for energy, protein, and fluid intake, respectively, according to the e-food record and 24-hour recall. For 7 out of 9 participants, the energy and protein intake determined via the 24-hour recall was higher than that determined via the e-food record. For 6 out of 9 participants, the fluid intake determined via the 24-hour recall was higher than via the e-food record. The mean difference in energy intake between the 24-hour recall and the e-food record was 4.9 (SD 10.2) kcal/kg bw and ranged between −17.2 and 16.7 kcal/kg bw, respectively. For protein intake, the mean difference was 0.1 (SD 0.3) g/kg bw and ranged between −0.5 and 0.4 g of protein/kg bw. Fluid intake differed by 4.9 (SD 9.4) g/kg bw on average and ranged between −13.6 and 18.9 g/kg bw.

XSL•FO
RenderX

Figure 4. Mean energy intake per participant (n=9) over 3 days, recorded by the e-food record and 24-hour recall. bw: body weight.

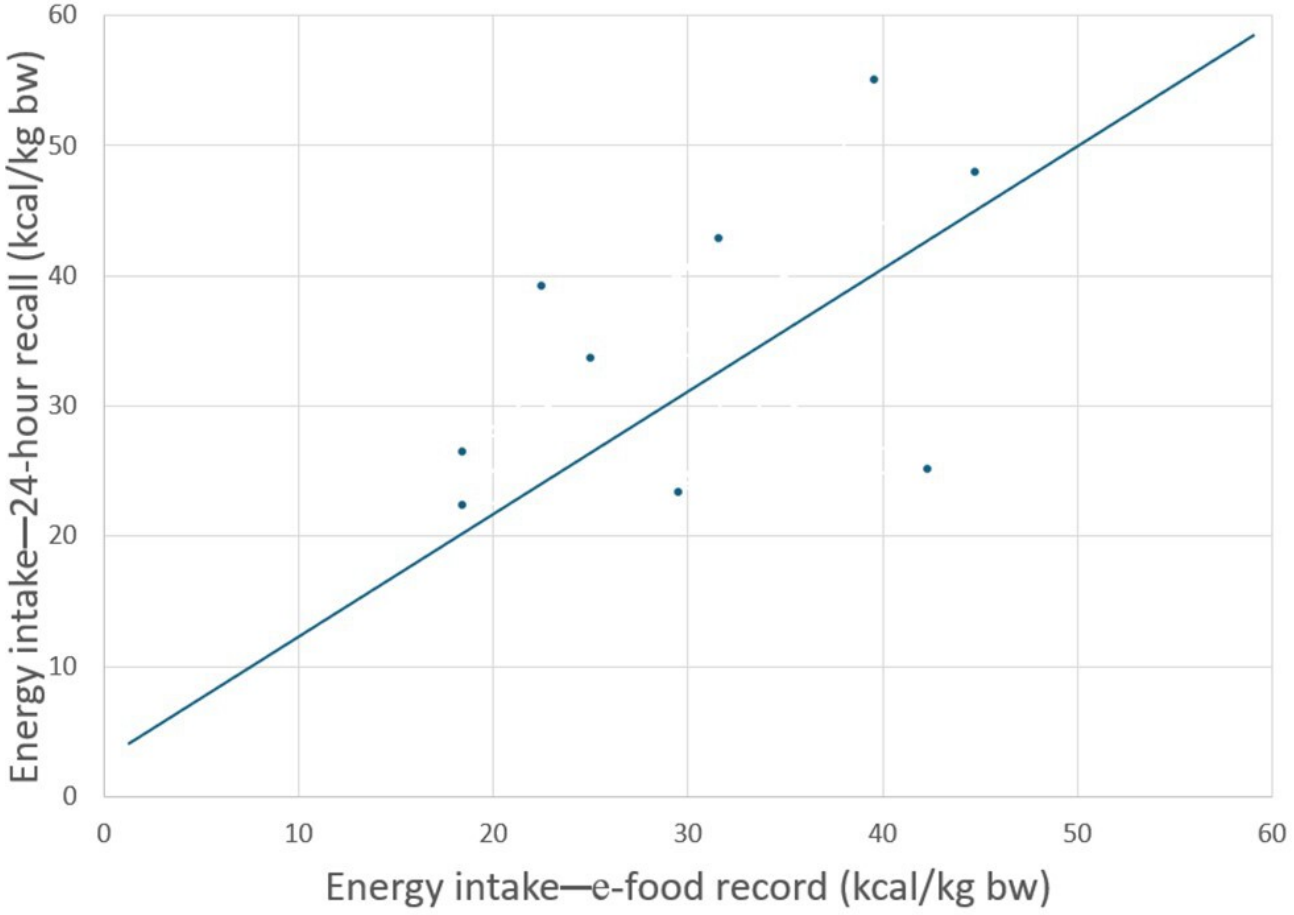
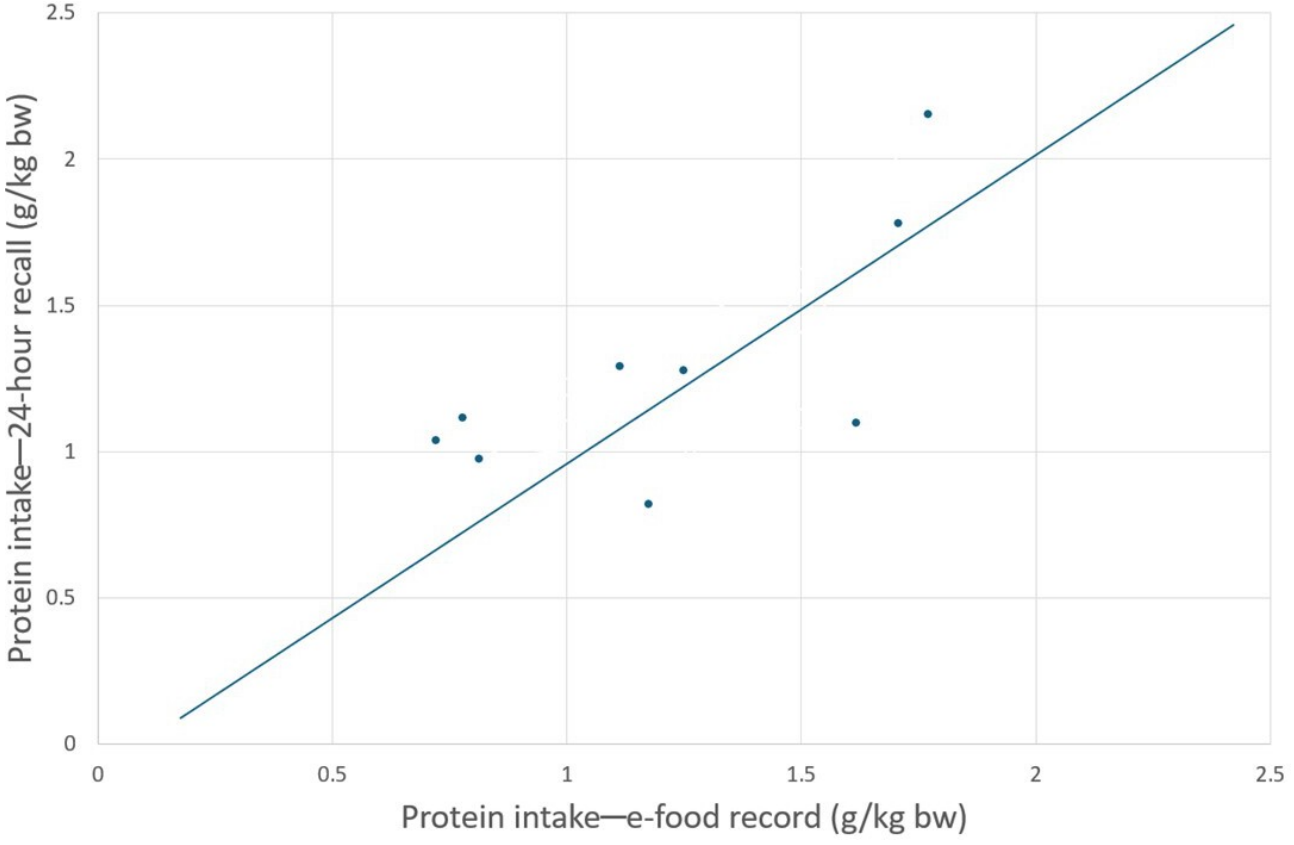
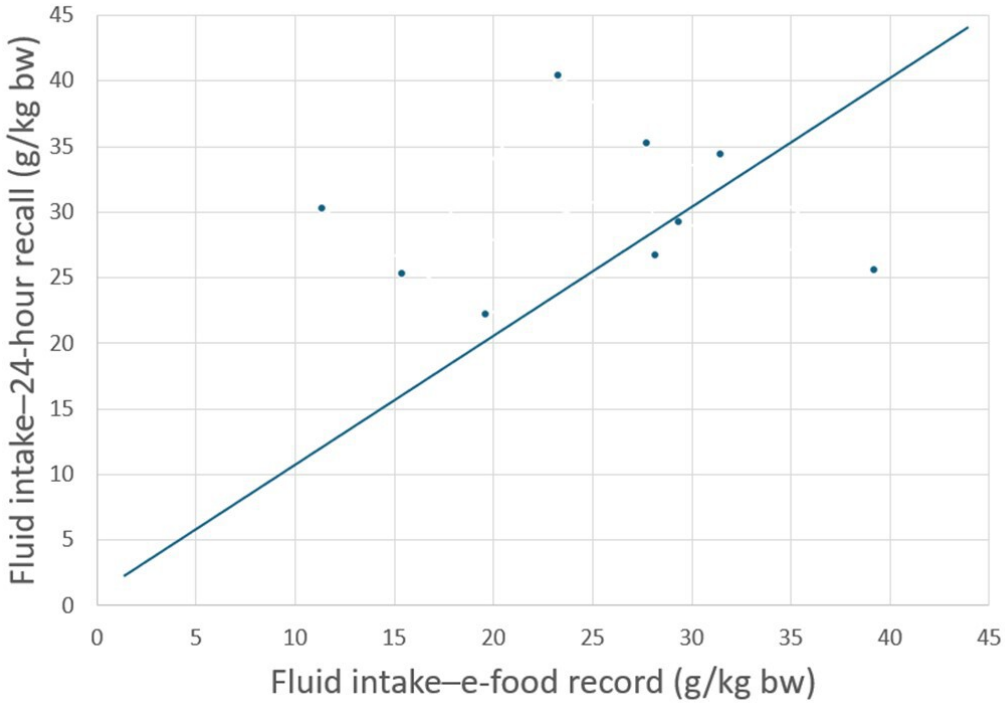


Figure 5. Mean protein intake per participant (n=9) over 3 days, recorded by the e-food record and 24-hour recall. bw: body weight.



XSL•FO
RenderX

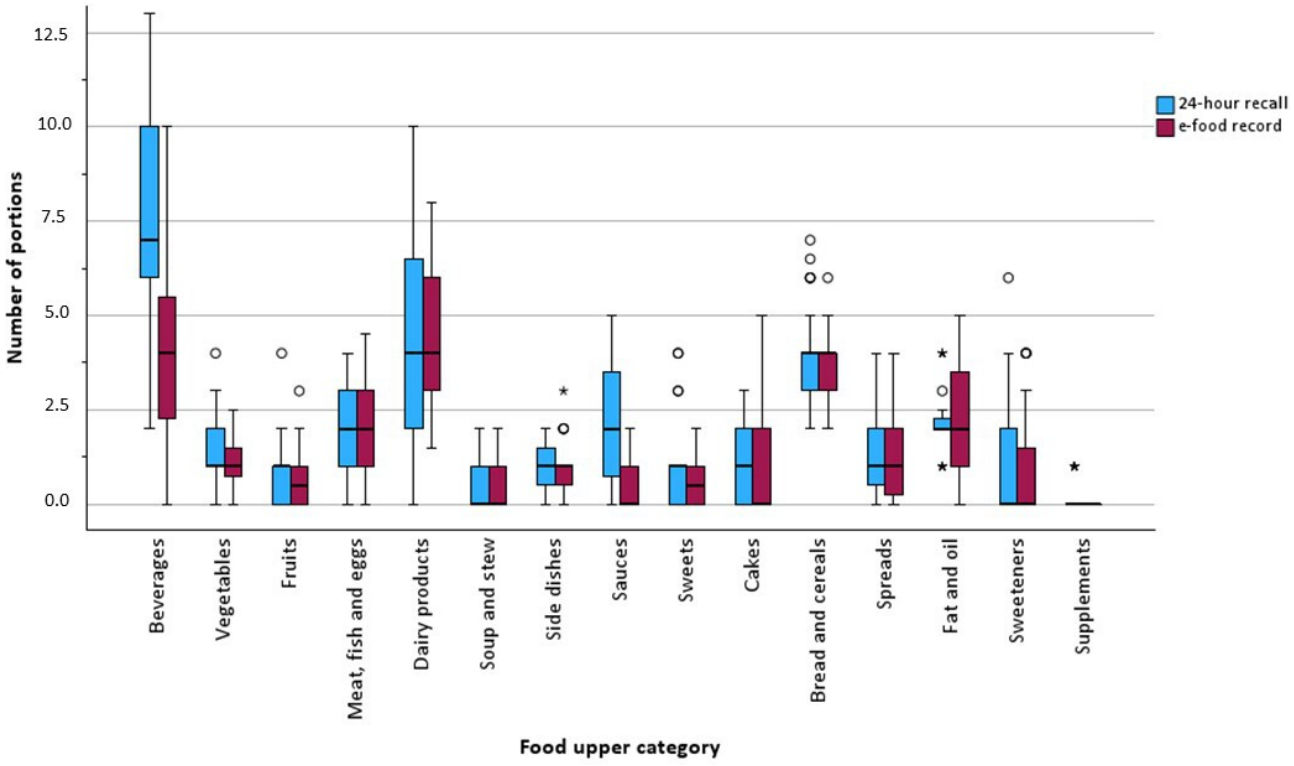
Figure 6. Mean fluid intake per participant (n=9) over 3 days, recorded by the e-food record and 24-hour recall. bw: body weight.



Food and Beverage Portion Entries

Over 3 days and across 9 participants (group A), a total of 625.5 food or beverage portions were recorded by the e-food record and 823.5 food or beverage portions by the 24-hour recall. Per study participant per day, this was an average of 23.9 (SD 5.9) food or beverage portions recorded by the e-food record and 30.8 (SD 6.7) food or beverage portions recorded by the 24-hour recall. On the upper food category level, the median of recorded portions per day was equal for the e-food record and 24-hour recall except for beverages (4, IQR 2-6 vs 7, IQR 6-10.5), fruits (0.5, IQR 0-1 vs 1, IQR 0-1), sauces (0, IQR 0-1 vs 2, IQR 0.5-4), sweets (0.5, IQR 0-1 vs 1, IQR 0-1), and cakes (0, IQR 0-2 vs 1, IQR 0-2), which showed a higher median (IQR) at the 24-hour recall. The results are shown in Figure 7.

Figure 7. Food and beverage portions per protocol day, recorded by the e-food record and by the 24-hour recall.

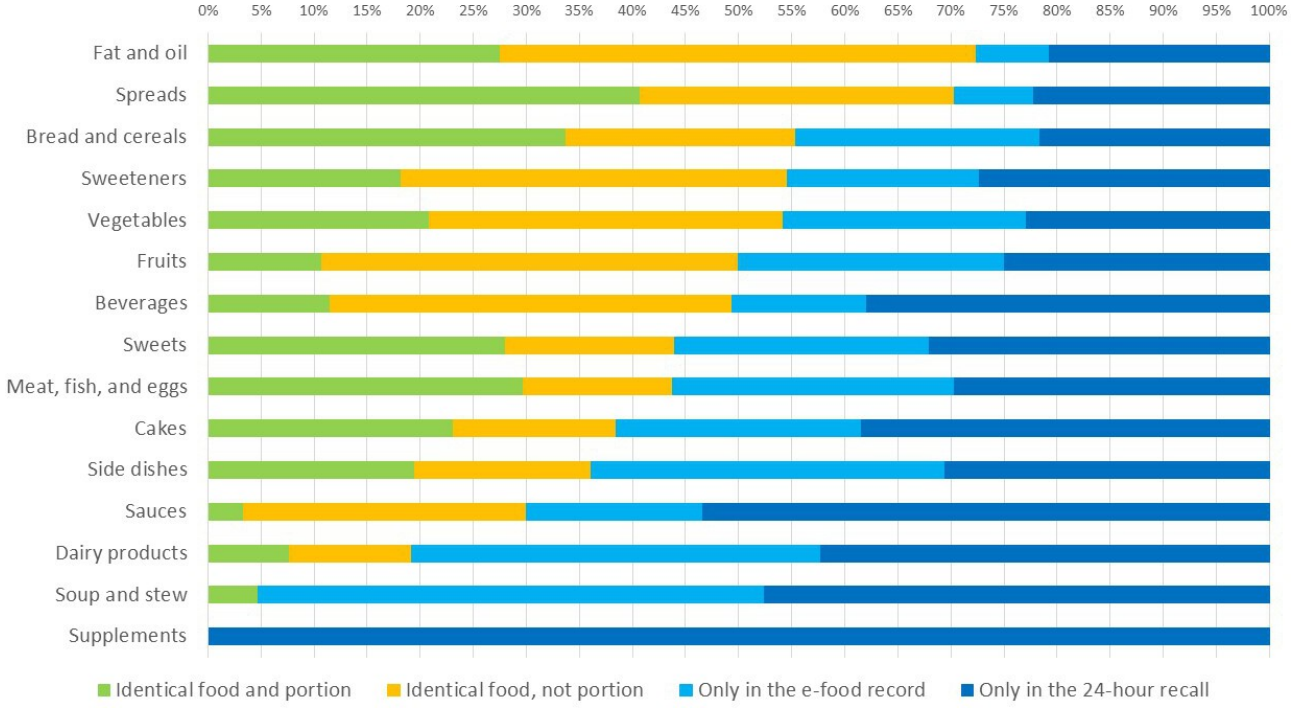


XSL•FO
RenderX

On the single food level, the portion entries for a corresponding food per day were compared between the e-food record and the 24-hour recall. A consistency was achieved when the same food with the same portion was recorded via the e-food record and the 24-hour recall on the same protocol day. The consistency of the identical portion entries between the e-food record and the 24 h recall was below 30% in 13 out of 15 food categories in relation to all entries in the corresponding food category. The highest correspondence was observed for the food category "spread," in which 40.7% (n=27) of the foods were recorded in both the e-food record and in the 24-hour recall with the same portion size, followed by 33.8% (n=74) in the food category "bread and cereals" and 27.6% (n=29) in the food category "fat and oil." In the food category "supplements," no consistency (0%, n=0) was observed for the portion entries of the e-food record and the 24-hour recall. Further consistency was achieved when the same food, but with a different portion size, was recorded via both the e-food record and the 24-hour recall on the same day. The consistency of food entries with different portion sizes between the e-food record and the 24-hour recall was below 30% in 10 out of 15 food categories in relation to all entries in the corresponding food category. The remaining entries in the food categories were related to foods that were either only recorded via the e-food record or only via the 24-hour recall. The results are shown in Figure 8.

According to the results from Figure 8, a range of food or beverage entries were recorded in only one of the 2 food record methods. In total, 65 different foods or beverages were recorded only via the 24-hour recall by at least 1 participant on at least one of the 3 protocol days. In addition, 48 different foods or beverages were recorded only via the e-food record. Condensed milk, water, and black tea were most often recorded only via the 24-hour recall, and salad and cookies were most often recorded only via the e-food-record. Multimedia Appendices 1 and 2 represent data for the respective protocol method starting with 3 cases per food and upward.

Figure 8. Food and portion entries recorded by the e-food record and 24-hour recall. 100% refers to the total entries of all participants over 3 days.



The explorative analysis of the food and beverage entries with regard to fat content, food variety, and preparation type shows that the same food, but with a different fat level (eg, fruit yogurt with 3.5% fat in dry matter vs fruit yogurt with 10% fat in dry matter), was entered for yogurt, milk and milk drinks, curd, as well as for cold cut and fish. A different type of food (eg, multigrain bread vs wholemeal bread) was entered for cheese, milk and milk drinks, bread, cakes, tea, coffee, fat and oil, vegetables, and fruits. The same dish, but different preparation types, were entered for soup and stew, potato dishes, vegetables, and sauce. Data are represented in Multimedia Appendix 3.

Discussion

Principal Results

This pilot study primarily investigated whether an e-food record is a feasible method for documenting nutritional data in older adults undergoing geriatric rehabilitation. The results of this study demonstrate that the e-food record was successfully used by older adults in geriatric rehabilitation to record their food and beverage intake over a period of 3 days. Secondary and exploratory findings are related to usability and the comparison with a 24-hour recall. The e-food record's usability score (mean 76.0, SD 11.3 points) indicates good acceptability and ease of use. However, the e-food record and the 24-hour recall determined different results in terms of recorded nutrient intake and food portion sizes, indicating possible limitations. These differences should be interpreted with caution, as both methods

XSL•FO
RenderX

are subject to measurement errors. While the 24-hour recall may be affected by recall bias and memory limitations, the e-food record may be influenced by input errors or difficulties in the estimation of portion size. These findings suggest that the e-food record is a promising tool for recording food and beverage intake in older adults, but it currently also has some limitations.

Comparison With Prior Work

With respect to the primary objective of feasibility, this study demonstrated that the target group of patients undergoing geriatric rehabilitation was able to use an e-food record independently during their rehabilitation treatment. A systematic review by Kraaijkamp et al [12] showed that eHealth is feasible, usable, and effective in geriatric rehabilitation. However, none of the 40 included studies were related to nutritional content. A nutrition and exercise e-coach was developed iteratively by Happe et al [13,25] and tested successfully in a piloted application study with a follow-up of 9 weeks among older people during and after geriatric rehabilitation [14]. However, this study was designed for usability testing and user behavior analysis, and the data of the e-food record were not analyzed. A study on community-dwelling older adults with a mean age of 72.2 (range 65-89 y) demonstrated that an e-food record is comparable to a paper-based food record method [26]. During the study, 8 out of 25 participants withdrew prematurely from the study. These results are comparable to the study of Happe et al [14] mentioned above. Here, 8 out of 30 participants dropped out of the study prematurely. In this study, we observed a large difference between the number of patients who were in principle eligible for participation and the number of patients who actually participated (25 out of 224 patients undergoing geriatric rehabilitation).

In contrast, findings related to usability and data comparability should be interpreted as secondary and exploratory, as these aspects were not the primary focus of the study. The usability was rated with a mean score of 76.0 (SD 11.3) points, which reflects a high usability of the e-food record. A score of 82 points is described with the adjective “good” in the evaluation by Bangor et al [23] and is near to the score of 85 from which excellent usability begins. The SUS score is slightly lower than that for a tablet computer−based e-coach to improve nutrition and physical activity in patients undergoing geriatric rehabilitation developed by Happe et al [14] with 78.6 (SD 18.1) points on average. This study was conducted as a 3-day application test that did not contain any increased benefits for the participants (eg, feedback on dietary behavior or dietary recommendations). A higher benefit for the participants and thus a higher willingness to participate could be achieved in future studies by providing feedback on nutritional intake, integrating the e-food record into the nutrition therapy, or extending the study period beyond the rehabilitation program.

Implications for Research

The demonstrated feasibility of independent use of the e-food record provides an important basis for future research in the field of nutritional assessment in geriatric rehabilitation. The e-food record shows potential as a tool for documenting nutritional data in geriatric rehabilitation, which should be emphasized in future research. In order to exploit the full potential of the e-food record in the future, it should be integrated into daily treatment, ideally as part of a randomized and controlled intervention study. A comparison between the e-food record and the recorded nutritional values should be conducted using established reference methods and a larger number of participants to further investigate the validity and reliability of the e-food record in adequately powered studies. In addition, subsequent studies should investigate which factors contribute to the successful use of the e-food record and which barriers exist to use an e-food record among patients undergoing geriatric rehabilitation. Relevant aspects for further development and optimization of the e-food record were identified in this study by comparing the e-food record with the practical method of the 24-hour recall. The data in Figure 8 indicate that the minority of food and beverage entries (<30 % in 13 out of 15 food categories) are consistent between the e-food record and the 24-hour recall. To estimate portions more precisely, images of food items with the corresponding portions could be displayed on the user interface of the e-food record. In literature, there are also attempts to estimate portions using photo documentation and image recognition [27-29]. Furthermore, the data of Figure 8 indicate a relatively large discrepancy between the foods recorded via the e-food record and the 24-hour recall. A confirmation prompt could be added to the end of a meal entry, for example, to detect incorrect entries by the patients and give them the opportunity to correct them. The results represented in Multimedia Appendix 2 indicate that certain foods such as condensed milk, water, black tea, coffee, cookies, and gravy were partly reported only via the 24-hour recall. This might be due to the fact that the 24-hour recall was queried using a predefined list of foods in food or meal groups that possibly counteracted the forgetting of foods and beverages consumed. A structured query of frequently forgotten or often related foods or beverages such as coffee and condensed milk, bread and spreadable fat, or salad and dressing could prevent forgetting of foods and beverages. The results of Multimedia Appendix 3 indicate that participants may have rated the fat content of foods such as yogurt milk drinks, curd cheese, cold cuts, and fish differently in the e-food record and 24-hour recall. A barcode scanner function or letter recognition of the packaging may increase input accuracy. Within the use of the e-food record, it was possible to select both individual foods and entire dishes. Entire dishes could be omitted. They could be replaced by individual foods. If entire dishes are used as an input option, ingredient queries may increase accuracy of, for example, a stew. However, it is important to balance the benefits of additional functionality against the potential increase in complexity to ensure ease of use for patients undergoing geriatric rehabilitation. The exploratory findings of this study, particularly regarding usability and discrepancies with the 24-hour recall, provide important directions for further development and hypothesis generation.

Limitations

This study was aimed at serving as a pilot feasibility test and was, therefore, designed as a short-term observation with only a limited number of participants. This could limit the generalizability of the results. In addition, the study was conducted in a single rehabilitation center, which may limit the

XSL•FO
RenderX

transferability of the results to other settings or patient populations.

Only a small number of eligible patients participated in the study, and even a smaller number contributed analyzable dietary data. This may introduce substantial selection bias and limited generalizability, as participants who agreed to participate may differ systematically from nonparticipants, for example, in terms of motivation, health status, or technology affinity. Due to ethical constraints, no data on nonparticipation were collected, although the characteristics of this highly burdened population undergoing geriatric rehabilitation likely contributed to the low participation rate. As a result, the findings are based on a highly selected subgroup, and their generalizability to the broader population of patients undergoing geriatric rehabilitation is strongly limited.

A technical malfunction occurred during the study, as the automatic navigation to the current day's page did not function as intended. This may have affected user interactions and data quality. Feasibility was operationalized primarily based on the dropout rate. However, more detailed indicators of feasibility, such as the level of assistance required during use or the frequency of user errors, were not systematically recorded. This limits a more comprehensive understanding of independent use.

Furthermore, the cognitive status was not systematically recorded within the study. Cognitive impairment may have influenced both the ability to use the e-food record and the accuracy of the 24-hour recall, for example, due to recall bias or difficulties in reporting. However, patients with inability to understand the study content and procedures were excluded from participation, which may partially reduce the impact of severe cognitive impairment in the study sample.

Differences observed between the e-food record and the 24-hour recall could be interpreted with caution, as both methods are subject to different sources of measurement error. The comparison between the two methods was exploratory and not intended as a validation study. Given the small sample size and exploratory design, no firm conclusions regarding the validity or accuracy of either method can be drawn. In addition, analyses related to usability and comparison with the 24-hour recall were exploratory and not powered for definitive conclusions. These results should therefore be interpreted with caution and require confirmation in larger studies.

Implications for Practice

Effective nutrition therapy is crucial for achieving an optimal nutritional status in patients undergoing geriatric rehabilitation with a reduced nutritional status, as malnutrition is directly linked to physical functionality and rehabilitation outcomes [8]. The e-food record method offers several advantages, including direct access to nutritional information for nutrition therapists. This contributes to the implementation of early and targeted interventions. Traditional food records require manual analysis by nutrition therapists and are often limited to 1-day recalls. With the e-food record method, the patients record their nutritional intake themselves. This reduces personnel resources, which in turn can be used for personalized advice.

By utilizing consumption data, the e-food record can improve nutritional status by providing more intensive and targeted therapy. To date, the detailed data on food and beverage intake over several days are rarely available in clinical practice, although it can enhance counseling. Additionally, nutrition therapists can use the e-food record to monitor the progress of nutrition therapy and introduce new interventions when necessary. This was previously more difficult due to time and financial constraints. This can lead to improved patient outcomes, such as enhanced physical function, reduced hospitalization rates, and better quality of life. Furthermore, an e-food record can facilitate communication between health care professionals, patients, and family members, promoting a more collaborative approach to nutrition therapy. In addition, the use of the e-food record is transferable to patient groups with similar requirements for the use of technology, such as patients in acute geriatric, older patients in neurology, or older people in short-term care facilities. However, given that findings beyond feasibility are exploratory, further validation is required before routine implementation in clinical practice.

Conclusions

This pilot feasibility study of an e-food record in the population of inpatients undergoing geriatric rehabilitation is, to our knowledge, the first of its kind and demonstrated, in general, that patients undergoing geriatric rehabilitation were able to independently record their food and beverage intake using the e-food record over 3 days. The results suggest that app-based food recording can be an integral part of nutrition therapy in the future to improve the medical care of patients undergoing geriatric rehabilitation. However, the findings of this study show that the e-food record and the 24-hour recall provide different results in terms of nutrient intake and food portion sizes. Further research with studies involving larger numbers of participants is needed to draw conclusions regarding data accuracy or validity of the e-food record.

Acknowledgments

We would like to thank all study participants for their engagement. Special thanks are due to Andrea Heinks who supported the study by recruiting participants, entering data, and conducting data analysis and to Louisa Dierkes who conducted a first application test as part of her valuable Bachelor's thesis. We would also like to thank the geriatric rehabilitation department of the Rehabilitation Center Oldenburg for allowing us to conduct our study as well as the DiDiER Consortium, within which fruitful and valuable project work for the development of the e-food record was possible. ChatGPT (GPT-5 and OpenAI) was used solely for language editing of the manuscript. No content ideas or analyses were generated by the tool.

XSL•FO
RenderX

Funding

This research was funded by the German Federal Ministry of Education and Research (grant 02K14A155), “Digitized Services in Dietary Counseling for People with Increased Health Risks related to Malnutrition”. The funder had no role in study design, data collection, analysis, interpretation, the decision to publish, or preparation of the manuscript.

Data Availability

The datasets used or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: AH, JMB, RD
Critical review of the manuscript: RD, AH, ME
Formal analysis: JB
Funding acquisition: AH, JMB, RD
Investigation: JB
Project administration: ME
Recruitment: JB
Software: PE
Supervision: RD
Writing – original draft: JB
Writing – review and editing: JB, RD

Conflicts of Interest

JB, PE, ME, AH, and RD declare that they have no competing interests. JMB reports receiving speaker honoraria from Novartis, Pfizer, Fresenius, Nutricia, and Danone outside the submitted work. JMB is an ongoing member of the advisory boards of Nutricia, Danone, Pfizer, Rejuvenate, and Biomed outside the submitted work.

Multimedia Appendix 1
Food recorded only via the 24-hour recall.
[PNG File, 160 KB - aging_v9i1e84653_app1.png]

Multimedia Appendix 2
Food recorded only via the e-food record.
[PNG File, 133 KB - aging_v9i1e84653_app2.png]

Multimedia Appendix 3
Food entries according to fat content, food varieties, and preparation type.
[PNG File, 450 KB - aging_v9i1e84653_app3.png]

References

1. Meyboom-de Jong B, Wynia K, Geluk-Bleumink A. Ageing better in the netherlands. In: D‘Onofrio G, Greco A, Sancarlo D, editors. Gerontology: InTechOpen; 2018:101-111. [doi: 10.5772/intechopen.74748]
2. Low S, Wee E, Dorevitch M. Impact of place of residence, frailty and other factors on rehabilitation outcomes post hip fracture. Age Ageing 2021 Feb 26;50(2):423-430. [doi: 10.1093/ageing/afaa131] [Medline: 32902621]
3. Abbatecola AM, Fumagalli A, Spazzafumo L, et al. Body composition markers in older persons with COPD. Age Ageing 2014 Jul;43(4):548-553. [doi: 10.1093/ageing/aft196] [Medline: 24366838]
4. Churilov I, Churilov L, MacIsaac RJ, Ekinci EI. Systematic review and meta-analysis of prevalence of sarcopenia in post acute inpatient rehabilitation. Osteoporos Int 2018 Apr;29(4):805-812. [doi: 10.1007/s00198-018-4381-4] [Medline: 29455250]
5. Grund S, Gordon AL, van Balen R, et al. European consensus on core principles and future priorities for geriatric rehabilitation: consensus statement. Eur Geriatr Med 2020 Apr;11(2):233-238. [doi: 10.1007/s41999-019-00274-1] [Medline: 32297191]
6. Marshall S, Bauer J, Isenring E. The consequences of malnutrition following discharge from rehabilitation to the community: a systematic review of current evidence in older adults. J Hum Nutr Diet 2014 Apr;27(2):133-141. [doi: 10.1111/jhn.12167] [Medline: 24289811]
7. Mendelson G, Katz Y, Shahar DR, et al. Nutritional status and osteoporotic fracture rehabilitation outcomes in older adults. J Nutr Gerontol Geriatr 2018;37(3-4):231-240. [doi: 10.1080/21551197.2018.1496513] [Medline: 30376425]

XSL•FO
RenderX

8. Wojzischke J, van Wijngaarden J, van den Berg C, et al. Nutritional status and functionality in geriatric rehabilitation patients: a systematic review and meta-analysis. Eur Geriatr Med 2020 Apr;11(2):195-207. [doi: 10.1007/s41999-020-00294-2] [Medline: 32297199]
9. Everink IHJ, Grund S, Benzinger P, et al. Nutritional care practices in geriatric rehabilitation facilities across Europe: a cross-sectional study. J Clin Med 2023 Apr 17;12(8):2918. [doi: 10.3390/jcm12082918] [Medline: 37109255]
10. Volkert D, Beck AM, Cederholm T, et al. ESPEN guideline on clinical nutrition and hydration in geriatrics. Clin Nutr 2019 Feb;38(1):10-47. [doi: 10.1016/j.clnu.2018.05.024] [Medline: 30005900]
11. Beck AM, Dent E, Baldwin C. Nutritional intervention as part of functional rehabilitation in older people with reduced functional ability: a systematic review and meta-analysis of randomised controlled studies. J Hum Nutr Diet 2016 Dec;29(6):733-745. [doi: 10.1111/jhn.12382] [Medline: 27231148]
12. Kraaijkamp JJM, van Dam van Isselt EF, Persoon A, Versluis A, Chavannes NH, Achterberg WP. eHealth in geriatric rehabilitation: systematic review of effectiveness, feasibility, and usability. J Med Internet Res 2021 Aug 19;23(8):e24015. [doi: 10.2196/24015] [Medline: 34420918]
13. Happe L, Sgraja M, Hein A, Diekmann R. Iterative development and applicability of a tablet-based e-coach for older adults in rehabilitation units to improve nutrition and physical activity: usability study. JMIR Hum Factors 2022 Mar 16;9(1):e31823. [doi: 10.2196/31823] [Medline: 35293874]
14. Happe L, Sgraja M, Hein A, Quinten V, Diekmann R. Usability and feasibility of a tablet-based e-coach for older adults in rehabilitation units to improve nutrition and physical activity: a prospective pilot study. BMC Geriatr 2023 Sep 19;23(1):578. [doi: 10.1186/s12877-023-04204-6] [Medline: 37726662]
15. Mitteilungen des Bundesverbands Geriatrie [Article in German]. Z Gerontol Geriat 2016 Apr;49(3):267-268. [doi: 10.1007/s00391-016-1052-y]
16. Grund S, van Wijngaarden JP, Gordon AL, Schols JMGA, Bauer JM. EuGMS survey on structures of geriatric rehabilitation across Europe. Eur Geriatr Med 2020 Apr;11(2):217-232. [doi: 10.1007/s41999-019-00273-2] [Medline: 32297190]
17. Elfert P, Berndt J, Dierkes L, et al. A novel digital nutrition diary for geriatric patients at high risk of frailty syndrome. Nutrients 2022 Jan 18;14(3):400. [doi: 10.3390/nu14030400] [Medline: 35276760]
18. Elfert P, Eichelberg M, Troger J, et al. DiDiER - digitized services in dietary counselling for people with increased health risks related to malnutrition and food allergies. Presented at: 2017 IEEE Symposium on Computers and Communications (ISCC); Jul 3-6, 2017. [doi: 10.1109/ISCC.2017.8024512]
19. Ernährungspyramide: Was esse ich? [Web Page in German]. Bundeszentrum für Ernährung. URL: https://www.bzfe.de/essen-und-gesundheit/ernaehrungspyramide/was-esse-ich [accessed 2025-09-10]
20. Elfert P. Unterstützung des Prozesses der Ernährungsberatung für hochaltrige Personen, die vom geriatrischen Frailty-Syndrom betroffen sind, mit Hilfe von Technik und Digitalisierung [Dissertation] [Article in German]. : University Oldenburg; 2024 URL: https://oops.uni-oldenburg.de/6896/1/ELFUNT24.pdf [accessed 2025-09-17]
21. Brooke J. SUS: a “quick and dirty” usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry: CRC Press; 1996:189-194 URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke [accessed 2026-05-28]
22. Quantitative usablility-analysen mit der system usability scale (SUS) [Web Page in German]. Seibert Group. URL: https://seibert.group/blog/2011/04/11/usablility-analysen-system-usability-scale-sus/ [accessed 2025-09-10]
23. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
24. Neyer FJ, Felber J, Gebhardt C. Entwicklung und Validierung einer Kurzskala zur Erfassung von Technikbereitschaft [Article in German]. Diagnostica 2012 Apr;58(2):87-99. [doi: 10.1026/0012-1924/a000067]
25. Happe L, Hein A, Diekmann R. What do geriatric rehabilitation patients and experts consider relevant? Requirements for a digitalised e-coach for sustainable improvement of nutrition and physical activity in older adults—a qualitative focus group study. BMC Geriatr 2021 Dec 18;21(1):712. [doi: 10.1186/s12877-021-02692-y] [Medline: 34922492]
26. Timon CM, Astell AJ, Hwang F, et al. The validation of a computer-based food record for older adults: the Novel Assessment of Nutrition and Ageing (NANA) method. Br J Nutr 2015 Feb 28;113(4):654-664. [doi: 10.1017/S0007114514003808] [Medline: 25630436]
27. Thompson FE, Subar AF. Dietary assessment methodology. In: Coulston AM, Boushey CJ, Ferruzzi MG, Delahanty LM, editors. Nutrition in the Prevention and Treatment of Disease: Elsevier; 2017:5-48.
28. Martin CK, Nicklas T, Gunturk B, Correa JB, Allen HR, Champagne C. Measuring food intake with digital photography. J Hum Nutr Diet 2014 Jan;27 Suppl 1(1):72-81. [doi: 10.1111/jhn.12014] [Medline: 23848588]
29. Naaman R, Parrett A, Bashawri D, et al. Assessment of dietary intake using food photography and video recording in free-living young adults: a comparative study. J Acad Nutr Diet 2021 Apr;121(4):749-761. [doi: 10.1016/j.jand.2020.09.040] [Medline: 33187931]

Abbreviations

bw: body weight

XSL•FO
RenderX

SUS: System Usability Scale

Edited by R Lin; submitted 30.Sep.2025; peer-reviewed by P Schuss; revised version received 17.Apr.2026; accepted 06.May.2026; published 09.Jul.2026.

Please cite as:
Berndt J, Elfert P, Eichelberg M, Bauer JM, Hein A, Diekmann R
A Tablet Computer–Based Food Record for the Self-Assessment of Nutritional Intake in Patients Undergoing Geriatric Rehabilitation: Prospective Pilot Feasibility Study
JMIR Aging 2026;9:e84653
URL: https://aging.jmir.org/2026/1/e84653
doi:10.2196/84653

© Julia Berndt, Patrick Elfert, Marco Eichelberg, Juergen Martin Bauer, Andreas Hein, Rebecca Diekmann. Originally published in JMIR Aging (https://aging.jmir.org), 9.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Performance-Based Usability of Medication Adherence Technologies Among Older Adults With Diverse Capabilities: Quantitative Study

Bincy Baby[1], MSc, PharmD; Ghada Elba[1], MSc, PharmD; SooMin Park[1], PharmD; Imra Hudani[1], BSc; Rishabh Sharma[1], MSc, PharmD; Kirk Patterson[2], BA; Annette McKinnon[2]; Sara J T Guilcher[3], PT, PhD; Feng Chang[1], PharmD; Linda Lee[4], MD, MCISc(FM); Catherine Burns[5], PEng, PhD; Dagmar Hajducek[1], PhD; Ryan H Griffin[6], PEng, PhD; Tejal Patel[1,7], BScPharm, PharmD

[1]School of Pharmacy, University of Waterloo, 10 Victoria St S A, Waterloo, ON, Canada

[2]Patient Advisors Network, Toronto, ON, Canada

[3]Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada

[4]Department of Family Medicine, McMaster University, Hamilton, ON, Canada

[5]Systems Design Engineering, Faculty of Engineering, University of Waterloo, Waterloo, ON, Canada

[6]National Research Council Canada, Otttawa, ON, Canada

[7]Schlegel-UW Research Institute for Aging, Waterloo, ON, Canada

Corresponding Author:
Tejal Patel, BScPharm, PharmD
School of Pharmacy, University of Waterloo, 10 Victoria St S A, Waterloo, ON, Canada

Abstract

Background: Medication adherence technologies (MATs) offer innovative solutions to support older adults in managing complex medication regimens, yet usability challenges can prevent their successful use. Older adults often face cognitive, physical, sensory, motivational, and environmental barriers, which can influence how they interact with these devices. Therefore, performance-based usability testing is essential for identifying usability issues.

Objective: This study aimed to evaluate the performance-based usability and user experience of 13 MATs among older adults with diverse capabilities.

Methods: A prospective mixed methods design was used, with 96 participants aged 60 years and older testing smart and electronic MATs using cognitive walkthroughs and predefined usability tasks without formal training to simulate first-use conditions. This manuscript focuses specifically on the performance-based usability metrics collected as part of the mixed methods study. Validated tools, including the Self-Medication Assessment Tool (SMAT), Daily Living Tasks Dependent on Vision (DLTV), Whisper Test, Self-Efficacy for Appropriate Medication Use Scale (SEAMS), and Martin and Park Environmental Demands (MPED) Questionnaire, were used to assess individual barriers. Poisson generalized estimating equations (GEE) was used to determine predictors of unassisted task success and error rates while accounting for repeated observations.

Results: A total of 96 participants (mean age 75.1, SD 7.7, range 61 - 95 y) were included, with 37.5% (36/96) male and 62.5% (60/96) female. Poisson GEE models identified significant predictors of unassisted task success rate, including cognitive score (P<.001), physical score (P<.001), vision score (P<.001), motivational score (P<.001), and lower environmental busyness (P=.049). Female sex was associated with a 9% lower success rate (P=.002). Predictors of total error rate included age (P=.02), sex (P≤.001), physical score (P≤.001), vision score (P=.004), DLTV vision score (P=.01), MPED routine score (environmental; P=.05), and interaction between physical score and sex (P=.004).

Conclusions: Both device design and user characteristics strongly influence the usability of MATs. Performance-based evaluations provide actionable findings for user-centered design and selection, guiding manufacturers and health care providers toward developing and choosing accessible technologies that could improve medication adherence among older adults.

(JMIR Aging 2026;9:e88398) doi:10.2196/88398

KEYWORDS

usability; older adults; medication adherence technology; functional abilities; medication management

XSL•FO
RenderX

Introduction

In Canada, the senior population is increasing rapidly and is expected to make up nearly a quarter of the total population by 2040 [1]. As individuals age, their likelihood of developing chronic conditions significantly increases [1-5]. Multimorbidity is associated with polypharmacy, resulting in medication regimen complexity [1-6]. The complexity of such regimens, coupled with age-related cognitive and physical declines, poses significant barriers to medication self-management and adherence [7,8].

For older adults, the ability to manage their own medications is vital to maintaining independence and safety [7-9]. Defined as the ability to organize, take, and adhere to prescribed treatments, self-management requires cognitive functions, such as memory and comprehension, as well as physical capabilities like hand dexterity and visual acuity [8]. Cognitive limitations, such as memory loss and reduced problem-solving abilities, can prevent adherence to prescribed regimens [10,11]. Physical challenges, including arthritis or reduced hand strength, make tasks like opening medication containers difficult [10,12]. Sensory impairments, such as vision loss, can prevent individuals from distinguishing between medications or reading labels accurately [10,13]. Additionally, motivational factors such as low health literacy or reduced confidence, and environmental challenges, such as limited caregiver support or disorganized home environments, can further complicate medication self-management [10,14-16]. These issues often lead to medication nonadherence, resulting in adverse drug events, preventable hospitalizations, and higher health care costs [17,18].

Medication adherence technologies (MATs) offer innovative solutions to support older adults in managing their medications effectively [19,20]. These technologies make use of automation, connectivity, organization, and reminder functions to simplify medication regimens, reduce errors, and improve adherence [19-22]. For example, automated dispensing and reminder systems can reduce cognitive load and support individuals with memory-related challenges, while organized medication compartments and guided workflows may lessen reliance on visual interpretation of labels. While MATs can help address some of the functional challenges encountered by older adults by reducing cognitive load, simplifying medication routines, enabling automated medication dispensing, and facilitating communication between patients, caregivers, and health care providers, they do not eliminate these underlying impairments [21-24]. In fact, these impairments may also influence how effectively individuals interact with these technologies [21,23,25]. Cognitive impairments can limit the ability to understand or operate these devices, while physical difficulties, such as arthritis or tremors, can interfere with interaction with small buttons or complex mechanisms [10,25]. Sensory impairments, including vision and hearing loss, can make it difficult to engage with screens or auditory prompts [10,25]. Moreover, low technology literacy and resistance to adopting unfamiliar tools further restrict MATs' adoption [10,14,25]. The cost of these technologies and limitations in home environments, such as poor internet connectivity, also pose significant barriers [16,26].

The usability of MATs plays an important role in their effectiveness, efficiency, and user satisfaction [27]. Usability is defined as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [28]. For older adults, poor usability can result in medication errors, frustration, device abandonment, and even preventable hospitalizations [27,29]. Usability metrics are vital for evaluating how well users interact with MATs [27]. These metrics can be divided into perception-based and performance-based categories [30-32]. Perception-based metrics focus on users’ subjective satisfaction and feelings about a product’s usability, often captured through standardized tools such as the System Usability Scale (SUS) or the Usefulness, Satisfaction, and Ease of Use (USE) questionnaire [30-32]. Performance-based metrics, in contrast, focus on quantifiable outcomes of user interactions, offering a more objective assessment of usability [30-32]. Key performance metrics include task success, time-on-task, error rates, efficiency, and learnability [30-33]. Unassisted task success evaluates users’ ability to complete specific tasks independently, with lower success rates often reflecting both user-related factors (eg, cognitive, physical, and sensory abilities) as well as potential design-related challenges [30-33]. For example, Ligons et al in a study investigating how cognitive status affects the usability of a medication delivery device (EMMA), found that cognitive impairments reduced success rates for older adults from 69% for individuals with Mini-Mental State Examination (MMSE≥24) to 34.7% for individuals with MMSE<24 [29]. Time-on-task measures the time required to complete tasks, where longer times often indicate inefficiencies in design or difficulties arising from user limitations [30-33]. Error rates, categorized into “slips” and “mistakes,” highlight areas where both interface complexity and user challenges may contribute to confusion [30-33]. Efficiency reflects the effort required to complete tasks, considering both the number of steps and the time spent, and may be influenced by the interaction between device design and user abilities. Learnability assesses how quickly users adapt to a device over time, which can depend on both the intuitiveness of the design and the user’s cognitive and functional capacity [30-33].

By focusing on performance-based metrics, it is possible to pinpoint specific usability problems, such as which features or tasks are most challenging for older adults based on their individual limitations [33]. Barriers to usability, such as confusing interfaces or physically demanding mechanisms, can be identified and addressed to improve functionality, reduce frustration, and improve adoption rates [30-33]. Therefore, the objective of this research is to evaluate the usability and user experience (UX) of 13 different MATs among older adults with diverse physical, cognitive, sensory, motivational, and environmental abilities by examining various performance-based metrics such as unassisted task success rates, total error rates, total task completion time, and efficiency for unassisted tasks. Through comprehensive usability testing, this study aims to identify the design strengths and weaknesses of these devices for older adults with varying limitations, offering actionable

XSL•FO
RenderX

recommendations to improve MATs' usability, functionality, and adoption.

Methods

Study Design

This research used a prospective mixed methods design, integrating both quantitative and qualitative approaches. Quantitative methods included cognitive walkthroughs (refer to Study Procedure section), measurement of various performance-based usability metrics, and assessment of perception-based usability using standardized usability questionnaires. This paper focuses specifically on the performance-based metrics collected in this study.

Recruitment, Eligibility, and Sample Size

Participants aged 60 years and older who were able and willing to provide informed consent were included in the study. Individuals who could not speak or read English were excluded. The sample size for the quantitative analyses was determined by applying Green's formula to ensure sufficient power for regression models, including 6 independent variables (physical, cognitive, vision, hearing, motivational, and environmental limitations). Based on the formula n>50+8 m (where "m" represents the number of independent variables), the recommended sample size was calculated as 98 [34]. In addition, considerations from usability research informed the recruitment strategy. Previous work by Nielsen and Landauer [35] suggests that testing with a small number of users (n=5 approximately) can identify the majority (80% - 85%) of usability issues [35,36]. Accordingly, for each device, we aimed to recruit at least 5 participants representing each type of functional limitation to ensure adequate identification of usability challenges across diverse user groups.

A total of 13 devices were selected for testing. A Microsoft Excel tracking sheet was used to monitor device testing across participants and impairment types, supporting balanced representation across devices. Recruitment targeted a diverse participant pool using convenience, purposive, and snowball sampling methods, with outreach conducted through community organizations, senior fairs, professional networks, and other initiatives. Data collection occurred between June 2023 and June 2024 among older adult residents of Ontario.

MATs Tested

A total of 13 MATs were tested, including three smart devices (SMs) and 10 electronic devices. Electronic MATs typically include features such as multiple compartments, visual or audible alarms, and timers that remind patients when it is time to take their medications [21]. While these devices may operate electronically, they generally do not connect to external networks or systems. Smart MATs incorporate advanced features such as connectivity and automaticity [22]. Connectivity refers to the ability of these devices to connect to other systems or networks using technologies such as Wi-Fi, Bluetooth, or mobile data [22]. This allows smart MATs to synchronize data with health care providers' systems, send reminders to patients via smartphones or other connected devices, and alert caregivers or family members if a dose is missed. Automaticity refers to the device's ability to perform functions without continuous manual input from the user [22]. This may include automatically dispensing medications at scheduled times, recording the timing and quantity of doses taken, and adjusting medication delivery based on real-time data from connected health monitoring devices [22].

These 13 devices were selected for their diversity of features in order to capture the widest range of functionalities and were guided by a taxonomy of MATs based on device features [37]. Participants tested between 1 and 7 devices across multiple sessions, depending on their preference and availability, with some requiring 2 visits to complete the testing. To ensure consistency and minimize potential bias, each device was tested by at least 5 participants representing different types of limitations. Participants were not provided with any formal training; instead, they were instructed to rely solely on the instructional materials provided with the devices, mimicking real-world independent learning scenarios as much as possible in a laboratory setting.

Table 1 provides details of the MATs tested along with their assigned codes, which will be used to refer to the various devices in this paper.

XSL•FO
RenderX

Table . List of MATs[a] tested.

Device code	Device name	Manufacturer
APD[b] 001	MedReady 1700 Automated Medication Dispenser	MedReady Inc
APD 002	GMS Med-e-lert Automatic Pill Dispenser	Group Medical Supply, LLC
PBA[c] 001	MedQ Smart PillBox	Med-Q
PBA 002	MedGlider System 1 with Talking Reminder	Medport
PBA 003	VitaCarry Advanced Pill Case	VitaCarry
PBA 004	e-pill Multi-Alarm Pocket XL	e-pill
PBA 005	100-Hour Pill Reminder	Aidapt
PBA 006	eNNOVEA Weekly Planner with Advanced Auto Reminder	eNNOVEA Medical, LLC
PBA 007	Pill Box with Digital instruction	Shenzhen LongShitong Technology Co LTD
PBA 008	MedCentre System	MedCenter Systems, LLC
SM[d] 001	Spencer Automatic Pill Dispenser	Custom Health
SM 002	CpaX Connected Medication Adherence Packaging	Jones Healthcare Group
SM 003	EllieGrid Smart Pill Organizer	EllieGrid

[a]MAT: medication adherence technology.

[b]APD: automated pill dispenser.

[c]PBA: pill box with alarm.

[d]SM: smart device.

Study Procedure

Upon enrollment, participants provided written informed consent. A structured assessment was then conducted to identify potential barriers to effective medication use, including physical limitations, cognitive challenges, vision and hearing impairments, motivational factors, and environmental influences. Validated tools were used for these assessments: (1) the Self-Medication Assessment Tool (SMAT) [38,39] to measure physical, cognitive, and vision-related barriers; (2) the Daily Living Tasks Dependent on Vision (DLTV) [40,41] for vision-specific challenges; (3) the Whisper Test [42,43] for hearing capabilities; (4) the Self-Efficacy for Appropriate Medication Use Scale (SEAMS) [44-47] to assess motivational factors; and the Martin and Park Environmental Demands (MPED) [48] questionnaire to evaluate environmental factors such as routine complexity and perceived busyness.

Following the barrier assessment, usability testing of the devices was conducted using the cognitive walkthrough method [49]. This is a method used in user interface design to understand how new users interact with a device by simulating the process of exploring its functionality for the first time [49]. It is a task-based approach that focuses on evaluating the device's ease of use and learnability, particularly for users who may not have previous experience or knowledge of the device [49]. The objective is to identify usability issues that could hinder the user's ability to complete tasks efficiently and effectively [30-32]. To conduct a cognitive walkthrough, a participant information sheet and an evaluation sheet were prepared. The participant information sheet outlines specific tasks for users to perform with the devices. These tasks are designed to cover a range of functions and features of various devices, ensuring a comprehensive assessment of their usability. The tasks ranged from basic setup procedures like inserting batteries and unlocking the device to more complex interactions, such as setting up the current time, filling a tray with weekly medication, setting alarms, locking the device, and removing medication after an alarm. The evaluation sheet documented each task's outcome, marking them as unassisted success, assisted success, or failure. Additionally, the sheet recorded task errors, indicating whether multiple attempts were required for a successful completion or if tasks were accomplished on the first try, and the time spent to complete each task. The evaluation sheet was developed based on previous research and expertise in usability testing of similar devices. During the testing sessions, 5 researchers (BB, GE, SP, IH, and RS) were paired in groups of 2 to independently observe participants and simultaneously record task performance. All evaluators received training on the study protocol and data recording procedures to ensure consistency. Training consisted of an initial structured session led by the corresponding author (TP), which lasted 4 hours and introduced all evaluators to the cognitive walkthrough methodology, task procedures, and data recording processes. Mock cognitive walkthroughs were conducted between groups of 3 evaluators to practice observing device testing among participants and recording data. Furthermore, for evaluators who joined the research team after the initial training session, a one-on-one training session was provided by the first author (BB), followed by observation of at least 2 full testing sessions with participants before independently participating in data

XSL•FO
RenderX

collection. Following each session, the researchers reviewed their observations and reached a consensus to ensure accuracy and consistency in the recorded data. A mock medication regimen was developed using placebo medications (eg, candy and inert tablets) to simulate common prescription drugs and dosing schedules. Further details about the regimen are provided in Multimedia Appendix 1.

Outcome Measures

Various outcome measures used in this study, along with their descriptions and the equations used for calculation, are provided in Table 2.

Table . Outcome measures, descriptions, and equations.

Outcome measure	Description and equation
Task success rate (unassisted)	Assesses users' ability to independently complete device-related tasks, reflecting usability and learnability [32]. Task outcomes were categorized as unassisted success, assisted success, or failure based on predefined success criteria [31]. Task success rate unassisted only (%)=Number of tasks completed unassistedTotal number of tasks required to use the device×100
Time on task	The total elapsed time required to complete a task, also referred to as task completion time [32]. Time was recorded from task start to completion using a smartphone clock application.
Efficiency	Measures how effectively users complete tasks by relating unassisted task success to time spent [32]. Efficiency=Number of tasks completed successfully without assistanceTotal number of tasks required to set up the deviceTotal time spent completing the setup steps (minutes)
Error rate	Quantifies unintentional user actions or mistakes made during task performance [31]. Errors were classified as present if additional attempts were required or absent if completed on the first attempt. Error rate (%)=Number of errors encounteredTotal number of tasks required to use the device×100

Data Collection and Statistical Analysis

Data were securely stored and managed using the REDCap platform (REDCap 15.1.2; Vanderbilt University). Descriptive statistics were used to summarize the key variables, and Spearman correlation analyses were performed to examine correlations between outcome measures and predictor variables such as age, number of devices tested, cognitive, physical, vision (SMAT and DLTV), SEAMS, and MPED scores. Product-level outcomes were analyzed using population-averaged generalized estimating equation (GEE) models. The unit of analysis was the individual product tried by a participant (product-level observations), with participants specified as the clustering unit to account for repeated observations within individuals. Models assumed a Poisson distribution with a log link and used an unstructured working correlation structure to account for within-participant correlation. The response variables were the counts of unassisted successes and the counts of errors. To account for variation in the number of subtasks per product, an offset equal to the log of the number of subtasks per product was included, such that model estimates correspond to rates (or percentages) of successes and errors per subtask. Models were fit using the *glmee* function from the *glmtoolbox* R package (version 4.3.3; RStudio). The box plots were created using SAS 9.4 (Enterprise Edition; SAS Institute) on the SAS OnDemand for Academics platform.

Ethical Considerations

This study received ethical approval from the University of Waterloo Office of Research Ethics (ORE #45203). Informed consent was obtained from all participants before study procedures were initiated.

Results

Participant Demographics

A total of 96 individuals participated in the study, with a mean age of 75.1 (SD 7.7; range 61 - 95) years. Among them, 36 (37.5%) were male, and 60 (62.5%) were female. The majority (86/96, 89.6%) reported taking medications regularly. Educational backgrounds varied as follows: 27.1% (26/96) had advanced degrees, 28.1% (27/96) had earned a bachelor's degree, 17.7% (17/96) held diplomas, 17.7% (17/96) had a high school education or lower, and 1% (1/96) had not completed high school or held trade certificates. Most participants (83/96, 86.5%) lived independently in their own homes, while 13.5% (13/96) resided in independent retirement facilities.

Cardiovascular conditions were the most prevalent (56/96, 58.3%), followed by metabolic or endocrine disorders (37/96, 38.5%), musculoskeletal conditions (26/96, 27.1%), and vision-related issues (18/96, 18.8%). Less common health conditions included renal or urogenital disorders (7/96, 7.3%), stroke (5/96, 5.2%), and hematological conditions (4/96, 4.2%). A small proportion (5/96, 5.2%) reported having no significant medical history. Medication management approaches varied: 44.8% (43/96) used pillboxes, 11.5% (11/96) relied on alarm beepers, 10.4% (10/96) used blister packs, 8.3% (8/96) depended on reminders from others, 2.1% (2/96) used calendars, and 3.1% (3/96) reported other methods.

XSL•FO
RenderX

Barriers to Medication Self-Management

Cognitive challenges, defined as low to moderate scores on the SMAT, were observed in 21.9% (21/96) of participants. Physical barriers, also assessed using SMAT, were present in 38.5% (37/96) of participants. Vision-related barriers were also notable, with 24% (23/96) scoring low or moderately low on the SMAT vision scale, and 38.5% (37/96) reporting challenges in vision-dependent tasks as measured by the DLTV scale. Hearing impairments, identified through the Whisper Test, were found in 57.3% (55/96) of participants. Additionally, motivational and environmental barriers, assessed using the SEAMS and the MPED Questionnaire, were each reported by 35.4% (34/96) of participants.

More details on participant demographics and barrier characteristics are provided in Table 3.

XSL•FO
RenderX

Table . Demographics, clinical, and barrier characteristics of study participants (N=96).

Variable and category or description	Statistical value
Age (y), mean (SD; range)	75.1 (7.7; 61-95)
Sex, n (%)	
Male	36 (37.5)
Female	60 (62.5)
Medication use (currently taking medication), n (%)	86 (89.6)
Education level, n (%)	
Below high school	1 (1)
High school	24 (25)
Trade certificate or diploma	1 (1)
Nonuniversity diploma	17 (17.7)
Bachelor's degree	27 (28.1)
Master's, Doctoral, or Professional degree	26 (27.1)
Place of residence, n (%)	
Private home	83 (86.5)
Retirement home	13 (13.5)
Self-reported medical conditions, n (%)	
Cardiovascular	56 (58.3)
Metabolic or endocrine	37 (38.5)
Musculoskeletal	26 (27.1)
Eye-related	18 (18.8)
Oral or gastrointestinal	16 (16.7)
Neurological	12 (12.5)
Respiratory	10 (10.4)
Mental health	8 (8.3)
Renal or urogenital	7 (7.3)
Cancer or neoplasms	5 (5.2)
Stroke	5 (5.2)
Hematological	4 (4.2)
Congenital disorders	2 (2.1)
Inflammatory or immune	2 (2.1)
Injuries or accidents	2 (2.1)
Ear-related	1 (1)
Infection	1 (1)
Reproductive health	1 (1)
Skin	1 (1)
Other	1 (1)
None reported	5 (5.2)
Medication management strategies, n (%)	
Pillbox	43 (44.8)
Alarm or beeper	11 (11.5)
Blister pack	10 (10.4)
Reminder from another person	8 (8.3)

XSL•FO
RenderX

Variable and category or description	Statistical value
Medication calendar	2 (2.1)
Other	3 (3.1)
SMAT[a], n (%)	
Cognitive barrier	
High (≥90%)	70 (72.9)
Relatively high (≥80%)	5 (5.2)
Moderate (70% - 80%)	7 (7.3)
Low (≤69% approximately)	14 (14.6)
Physical barrier	
High (≥90%)	59 (61.5)
Moderate (76% - 84%)	11 (11.5)
Low (≤75%)	26 (27.1)
Vision barrier	
High (≥90%)	73 (76)
Moderate (76% - 84%)	4 (4.2)
Low (≤75%)	19 (19.8)
Vision barrier (DLTV[b]; impairment present), n (%)	37 (38.5)
Hearing barrier (Whisper Test; bilateral impairment), n (%)	55 (57.3)
Motivational barrier (SEAMS[c]; low self-efficacy), n (%)	34 (35.4)
Environmental barrier (MPED[d]; barrier present), n (%)	34 (35.4)

[a]SMAT: Self-Medication Assessment Tool; participants categorized as having moderate or low scores were considered to have impairments in cognitive, physical, and vision domains.

[b]DLTV: Daily Living Tasks Dependent on Vision; total score <79 indicates the presence of a vision-related barrier (lower scores reflect greater impairment).

[c]SEAMS: Self-Efficacy for Appropriate Medication Use Scale; total score <40 indicates low self-efficacy (motivational barrier), with lower scores representing reduced confidence in medication use.

[d]MPED: Martin and Park Environmental Demands; an environmental barrier is indicated by a busyness subscale score ≥15 (higher scores reflect greater busyness) and/or a routine subscale score <16 (lower scores reflect less structured routines).

Distribution of product testing by participant impairment type is presented in Multimedia Appendix 2.

Outcome Measures

Observed Usability Performance Across MATs

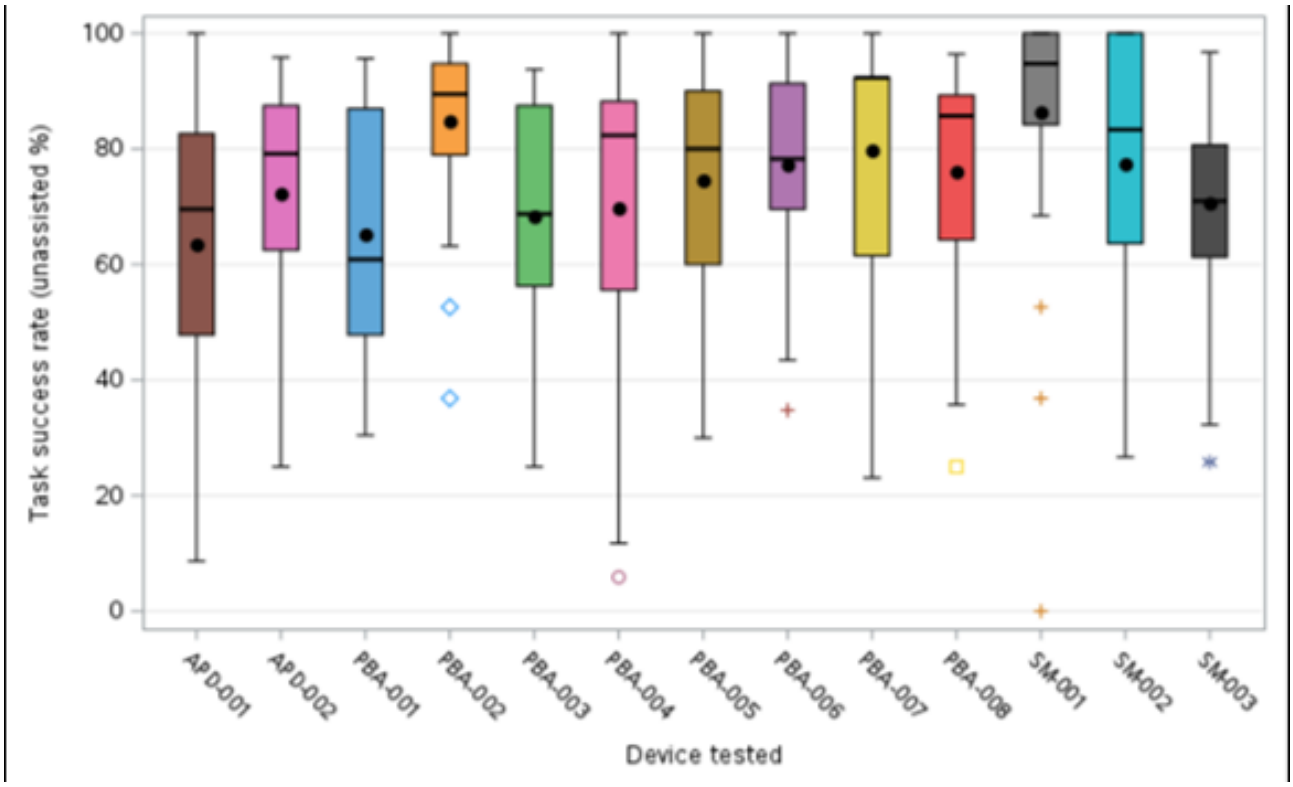
The following section presents observed usability outcomes across the MATs tested in this study, summarizing patterns in average success rate unassisted, total error rate, efficiency for unassisted task success, and total task time.

Average Task Success Rate Unassisted

The overall mean task success rate without assistance across all devices was 74.85% (SD 21.27), with a median of 80% (IQR 61.29-91.30), ranging from 0% to 100%. Smart pill dispensers, including the EllieGrid Smart Pill Organizer (SM-003; EllieGrid), Spencer Automatic Pill Dispenser (SM-001; Custom Health), and Jones Smart Blister Pack (SM-002; Jones Healthcare Group), generally showed higher success rates compared to electronic pill dispensers. Among the SMs, the Spencer Automatic Pill Dispenser (SM-001) had the highest mean success rate of 86.37% (SD 19.98), while the EllieGrid Smart Pill Organizer (SM-003) achieved 70.63% (SD 15.97). Electronic pill dispensers displayed more variability in performance, with success rates ranging from 63.48% to 84.82% (SD 26.42-15.21). The unassisted task success rates of the various tested devices are illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Unassisted task success rates of the various tested devices (boxes represent the IQR, horizontal lines indicate medians, dots indicate mean values, whiskers represent the spread of the data, and points represent outliers). APD: automated pill dispenser; PBA: pill box with alarm; SM: smart device.

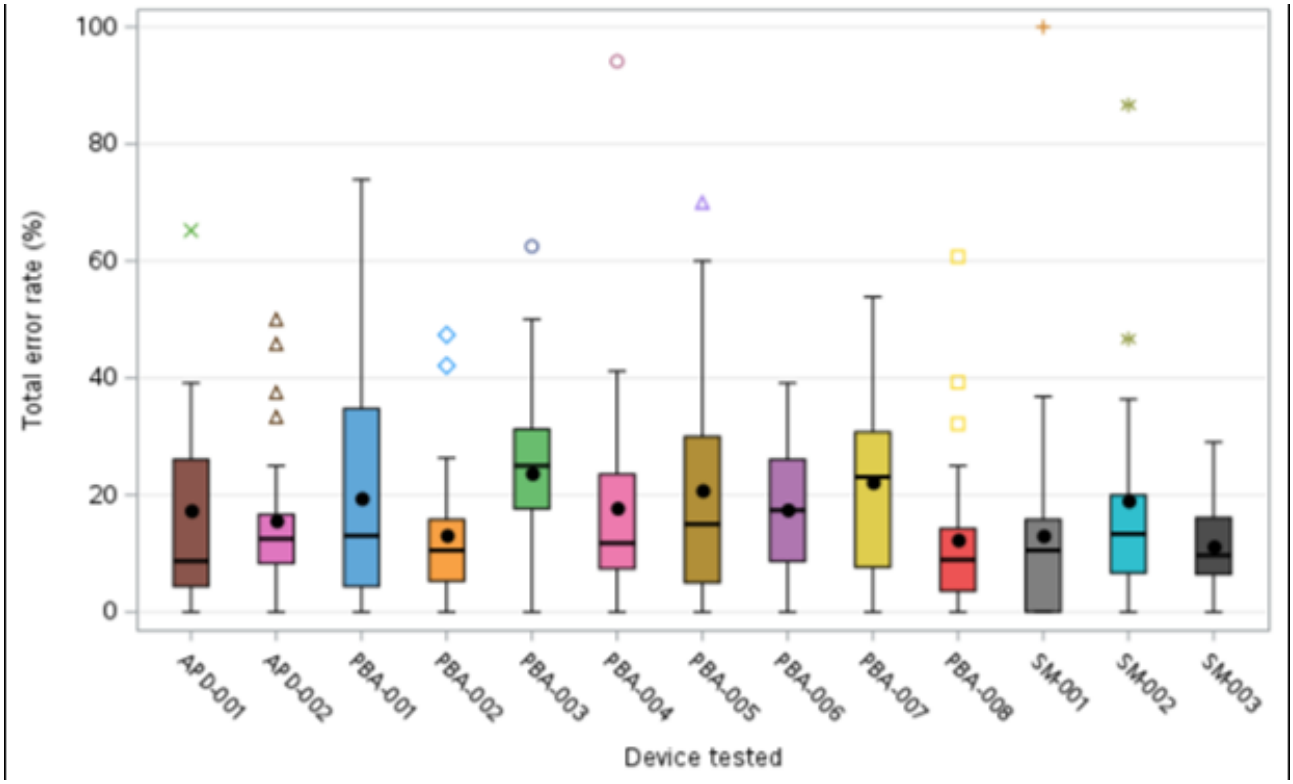


Total Error Rate

The overall mean total error rate across all devices was 16.87% (SD 15.89), with a median of 12.7% (IQR 5.73-23.9), and it ranged from 0% to 100%, indicating considerable variability in user performance. The first quartile (Q1) was 5.73%, while the third quartile (Q3) reached 23.9%, showing that most error rates were clustered within this range. Among the products, the EllieGrid Smart Pill Organizer (SM-003) demonstrated the lowest mean error rate at 11.23% (SD 8.02), with a median of 9.68% (IQR 6.45-16.13), while the highest error rate was observed for the VitaCarry Advanced Pill Case (PBA-003; eNNOVEA Medical, LLC) at 23.76% (SD 14.38). Smart pill dispensers, such as the Spencer Automatic Pill Dispenser (SM-001) and Jones Smart Blister Pack (SM-002), had mixed performance, with Spencer showing a lower mean error rate of 13.09% (SD 18.07) compared with Jones at 19.06% (SD 19.39). The unassisted task error rates of the various tested devices are illustrated in Figure 2.

XSL•FO
RenderX

Figure 2. Total task error rates of the various tested devices (boxes represent the IQRs, horizontal lines indicate medians, dots indicate mean values, whiskers represent the spread of the data, and points represent outliers). APD: automated pill dispenser; PBA: pill box with alarm; SM: smart device.

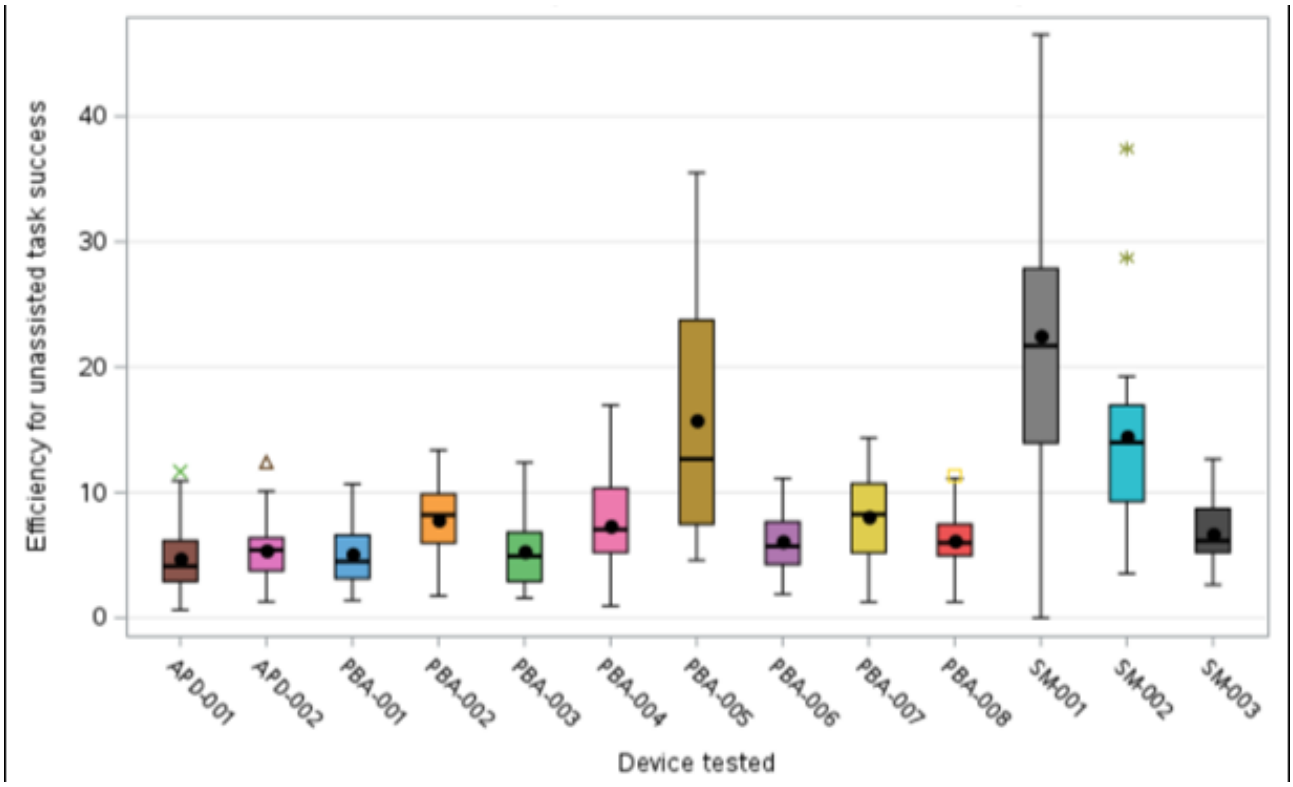


Efficiency for Unassisted Task Success

The overall efficiency for unassisted task success across all devices had a mean of 9.27 (SD 7.87), with a median of 6.83 (IQR 4.67-10.92), and values ranging from 0 to 46.51, indicating significant variability in performance. The highest efficiency unassisted was observed in the Spencer Automatic Pill Dispenser (SM-001), with a mean of 22.51 (SD 11.37) and a median of 21.70 (IQR 14.02-27.15), suggesting that SMs generally outperform electronic dispensers in task efficiency. Among electronic devices, the 100-Hour Pill Reminder (PBA-005; Aidapt) demonstrated relatively high efficiency, with a mean of 15.77 (SD 9.51), while devices such as the MedReady 1700 automated medication dispenser (APD-001; MedReady Inc) had lower efficiency, averaging 4.74 (SD 3.03). Smart pill dispensers, including the Jones Smart Blister Pack (SM-002), also performed well, with a mean efficiency of 14.48 (SD 7.5). The overall efficiency for unassisted task success of the various tested devices is illustrated in Figure 3.

XSL•FO
RenderX

Figure 3. Efficiency for unassisted task success of the various tested devices (boxes represent the IQR, horizontal lines indicate medians, dots indicate mean values, whiskers represent the spread of the data, and points represent outliers). APD: automated pill dispenser; PBA: pill box with alarm; SM: smart device.



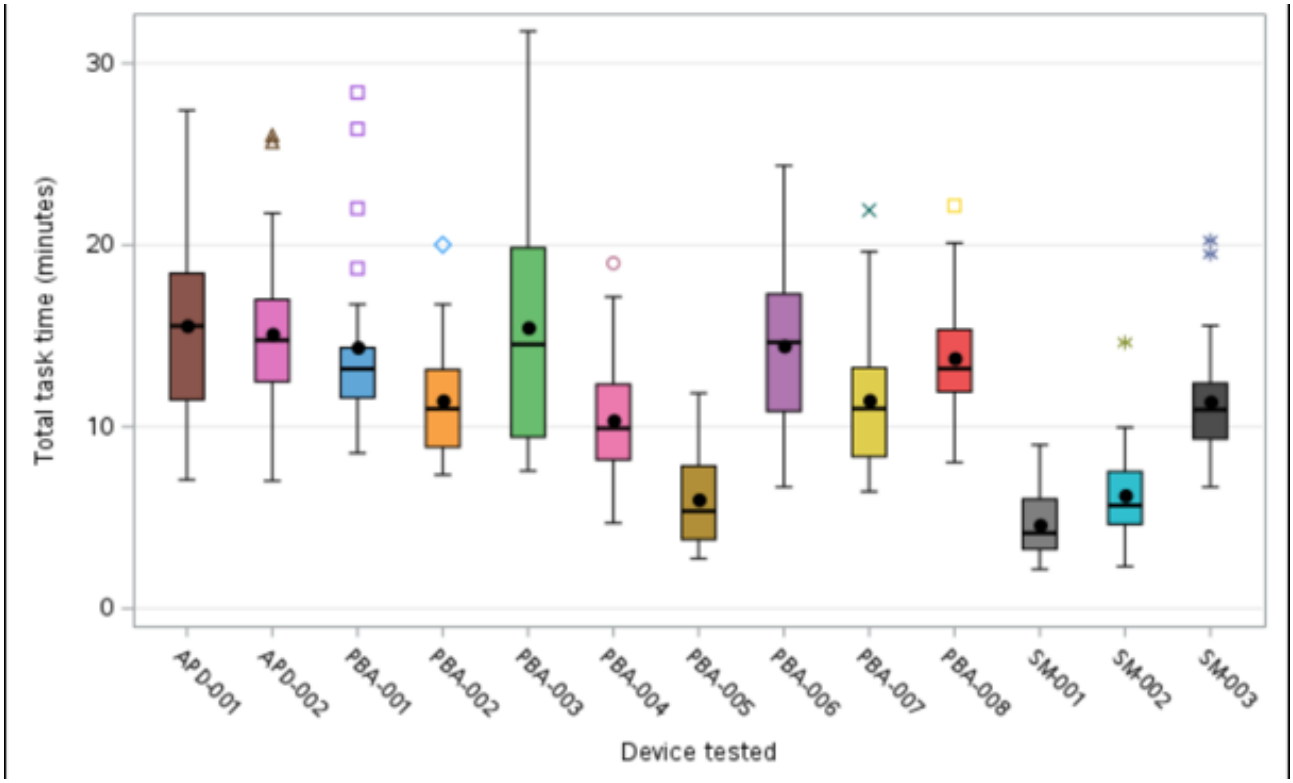
Total Task Time

The overall total task time across all devices had a mean of 11.42 (SD 5.48) minutes, with a median of 10.94 (IQR 7.67-14.67) minutes, and ranged from 2.15 to 31.77 minutes, indicating variation in the time required to complete tasks. Smart pill dispensers, such as the Spencer Automatic Pill Dispenser (SM-001) and the Jones Smart Blister Pack (SM-002), demonstrated the shortest task times, with averages of 4.61 (SD 1.83) minutes and 6.24 (SD 2.65) minutes, respectively, possibly suggesting their ease of use and efficiency. In contrast, electronic pill dispensers, such as the MedReady 1700 Automated Medication Dispenser (APD-001), had longer task times, with an average of 15.58 (SD 5.22) minutes, indicating potentially more timely interactions required for operation. Among electronic dispensers, the 100-Hour Pill Reminder (PBA-005) had the shortest average task time at 6.01 (SD 2.64) minutes, showing relatively quicker use. The overall total task time of the various tested devices is illustrated in Figure 4.

Multimedia Appendix 3 provides more details on various outcome measures based on the devices tested and the barriers assessed. Although the data have been presented according to individual impairment categories, these findings should be interpreted carefully, as most participants had multiple coexisting impairments, making it difficult to isolate the effect of any single impairment category. Therefore, these subgroup comparisons are not fully generalizable.

XSL•FO
RenderX

Figure 4. Total task time of the various tested devices (boxes represent the IQR, horizontal lines indicate medians, dots indicate mean values, whiskers represent the spread of the data, and points represent outliers). APD: automated pill dispenser; PBA: pill box with alarm; SM: smart device.



Significant Predictors of Performance-Based Usability Outcomes

Because participants tested varying numbers of devices and devices were not evaluated an equal number of times across participants, inferential analyses were conducted using Poisson GEE models with a log link, which account for within-participant correlation. This approach accounts for repeated measures within participants, unequal numbers of observations per participant, and variability in device usage, making it well-suited for analyzing correlated outcome data.

The selection of primary outcomes for regression analysis was informed by Spearman correlation analyses examining associations among outcome and predictor variables (Multimedia Appendix 4). Average unassisted task success rate and total error rate were selected as the primary modeled outcomes because they demonstrated meaningful associations with participant characteristics while representing distinct and interpretable dimensions of usability performance. Efficiency for unassisted task success and total task time were not modeled as primary outcomes, as these measures were highly correlated with task success rate and with each other, reflecting overlapping underlying constructs.

Significant Predictors of Average Task Success Rate Unassisted

Poisson GEE model with a log link and within-participant correlation identified several device types and participant characteristics as statistically significant predictors ($P<.05$).

Relative to the reference device (APD-001), success rates were not statistically significantly different for APD-001 ($P=.47$) and PBA-001 ($P=.82$). In contrast, several automated pill dispensers and pill boxes demonstrated higher success rates, ranging from 23.5% (PBA-006; $P=.009$) to 528.5% (PBA-005; $P<.001$). One automated pill dispenser (PBA-008) performed with significantly lower success rates (−28.7%; $P<.001$) in comparison to APD-001. Among SMs, 2 (SM-001 and SM-002) performed with significantly higher success rates (97.2% increase, $P<.001$, and 212.4% increase, $P<.001$, respectively). SM-003 had significantly lower success rates (−39.9%, $P<.001$). Regarding participant characteristics, females had a significantly lower success rate than males (−9.3%, $P=.002$). Higher cognitive (1% increase per unit, $P<.001$), physical (3.8% increase per unit, $P<.001$), and vision scores (3.0% increase per unit, $P<.001$) were all associated with increased success rates. Similarly, higher SEAMS scores were associated with a 1.1% increase in success rate ($P<.001$), while MPED busyness score showed a marginally significant 0.8% increase in success rate ($P=.05$).

Table 4 shows the results of the final Poisson GEE model, showing the estimated effects of device type and participant characteristics on unassisted task success rates.

XSL•FO
RenderX

Table . Fixed effects estimates for the final model predicting average task success rate.

Predictor	Estimate[a], standardized β coefficient (95% CI[b])	SE	*z* score	% change[c] (95% CI[d])	*P* value[e]
APD[f]-001 (reference)	−3.743 (−4.243 to −3.243)	0.255	−14.678	—[g]	—
APD-002	0.056 (−0.096 to 0.208)	0.078	0.720	5.7% (−9.2% to 23.1%)	.47
PBA[h]-001	−0.020 (−0.196 to 0.155)	0.090	−0.228	−2% (−17.8% to 16.8%)	.82
PBA-002	0.663 (0.510 to 0.816)	0.078	8.497	94.1% (66.6% to 126.2%)	<.001
PBA-003	0.784 (0.641 to 0.926)	0.073	10.786	118.9% (89.9% to 152.4%)	<.001
PBA-004	0.715 (0.496 to 0.933)	0.111	6.413	104.4% (64.3% to 154.3%)	<.001
PBA-005	1.838 (1.690 to 1.987)	0.076	24.271	528.5% (441.8% to 629.1%)	<.001
PBA-006	0.211 (0.052 to 0.369)	0.081	2.607	23.5% (5.4% to 44.7%)	.009
PBA-007	1.358 (1.196 to 1.519)	0.082	16.472	288.8% (230.8% to 357%)	<.001
PBA-008	−0.338 (−0.483 to −0.194)	0.074	−4.593	−28.7% (−38.3% to −17.6%)	<.001
SM[i]-001	0.679 (0.529 to 0.829)	0.077	8.866	97.2% (69.7% to 129.2%)	<.001
SM-002	1.139 (0.944 to 1.335)	0.100	11.425	212.4% (156.9% to 279.8%)	<.001
SM-003	−0.509 (−0.678 to −0.340)	0.086	−5.900	−39.9% (−49.2% to −28.8%)	<.001
Sex (female vs male)	−0.098 (−0.161 to −0.035)	0.032	−3.043	−9.3% (−14.9% to −3.4%)	.002
Cognitive score	0.010 (0.005 to 0.015)	0.003	3.763	1% (0.5% to 1.6%)	<.001
Physical score	0.037 (0.021 to 0.053)	0.008	4.420	3.8% (2.1% to 5.5%)	<.001
Vision score	0.030 (0.014 to 0.046)	0.008	3.624	3% (1.4% to 4.7%)	<.001
SEAMS[j] score (motivation)	0.011 (0.004 to 0.017)	0.003	3.314	1.1% (0.4% to 1.7%)	<.001
MPED[k] busyness score (environmental)	0.008 (0.000 to 0.015)	0.004	1.969	0.8% (0% to 1.5%)	.049

[a]Estimates (standardized β coefficient) are presented on the log scale from a generalized estimating equations Poisson model with a log link and an unstructured correlation structure. For categorical variables (product tested), estimates represent comparisons relative to the reference category (APD-001).

[b]95% CI for standardized β coefficient were calculated as β±1.96×SE. Percentage change in the expected error rate was computed as (e^β−1)×100.

[c]For continuous variables (eg, age and physical score), percentage change reflects the expected change in error rate per 1-unit increase in the predictor.

[d]Corresponding 95% CIs for percentage change were derived by exponentiating the confidence limits of β, calculated as [(e^(β−1.96×SE)−1)×100, (e^(β+1.96×SE)−1)×100].

[e]Statistical significance was assessed at $P<.05$.

[f]APD: automated pill dispenser.

[g]Not applicable.

[h]PBA: pill box with alarm.

[i]SM: smart device.

[j]SEAMS: Self-Efficacy for Appropriate Medication Use Scale.

[k]MPED: Martin and Park Environmental Demands.

Model diagnostic checks indicated that the final Poisson GEE model fit the data well. The dispersion parameter (0.289) and deviance-to-degrees-of-freedom ratio (1.06) showed no overdispersion, while the Pearson chi-square statistic ($\chi^2_{328}=328$, $P=.49$) confirmed an adequate fit. Residual diagnostics,

XSL•FO
RenderX

including a normal Q-Q plot and residuals versus fitted values plot, showed no major deviations or heteroscedasticity. The diagnostic plots are given in Multimedia Appendix 5.

Significant Predictors of Total Error Rate

Poisson GEE model with a log link and within-participant correlation identified several device types and participant characteristics as statistically significant predictors of total error rate. Relative to APD-001, most devices did not differ significantly in error rates ($P \geq .24$). However, PBA-008 showed a moderately significant lower error rate (−32.6%, $P=.04$). Among SMs, SM-001 showed a nonsignificant trend toward lower error rates (−35.9%, $P=.08$), while SM-002 and SM-003 did not differ significantly from the reference.

Regarding participant characteristics, increasing age was associated with a small but statistically significant increase in error rate (1.9% per unit increase, $P=.02$). Female participants had a significantly higher error rate compared with males (846.2% increase, $P<.001$). Higher physical scores were associated with increased error rates (46.1% increase per unit, $P<.001$), whereas higher vision scores were associated with reduced error rates (−9.9% per unit, $P=.004$). The DLTV summary score was also positively associated with error rate (2.2% increase per unit, $P=.01$). MPED routine score showed a marginally significant association with decreased error rate (−3.7%, $P=.05$). Additionally, there was a significant interaction between sex and physical score (−16.5%, $P=.004$), indicating that the effect of physical score on error rate differed by sex.

Table 5 presents the results of the final Poisson GEE model, showing the estimated effects of device type and participant characteristics on total error rate.

Model diagnostics indicated no evidence of overdispersion (overdispersion ratio=0.96) and an adequate fit (Pearson $\chi^2_{328}=328$, $P=.49$). Residual diagnostics (normal Q-Q plot and residual vs fitted plot) revealed no major violations of model assumptions. The diagnostic plots are given in Multimedia Appendix 4.

XSL•FO
RenderX

Table . Fixed effects estimates for the final model predicting total error rate.

Predictor	Estimate, standardized β coefficient[a] (95% CI[b])	*z* score	% change[c] (95% CI[d])	*P* value[e]
APD[f]-001 (Ref)	−7.493 (−10.112 to −4.874)	−5.607	—[g]	—
APD-002	−0.015 (−0.418 to 0.387)	−0.075	−1.5% (−34.2% to 47.3%)	.94
PBA[h]-001	0.116 (−0.341 to 0.573)	0.496	12.3% (−28.9% to 77.3%)	.62
PBA-002	−0.282 (−0.751 to 0.187)	−1.178	−24.6% (−52.8% to 20.6%)	.24
PBA-003	0.138 (−0.184 to 0.459)	0.839	14.7% (−16.8% to 58.3%)	.40
PBA-004	0.210 (−0.204 to 0.625)	0.995	23.4% (−18.4% to 86.7%)	.32
PBA-005	0.069 (−0.386 to 0.525)	0.299	7.2% (−32% to 69%)	.76
PBA-006	−0.172 (−0.547 to 0.204)	−0.895	−15.8% (−42.2% to 22.7%)	.37
PBA-007	0.193 (−0.166 to 0.553)	1.053	21.3% (−15.3% to 73.8%)	.29
PBA-008	−0.395 (−0.780 to −0.009)	−2.006	−32.6% (−54.2% to −0.9%)	.04
SM[i]-001	−0.444 (−0.933 to 0.045)	−1.779	−35.9% (−60.7% to 4.6%)	.08
SM-002	0.058 (−0.452 to 0.568)	0.221	5.9% (−36.4% to 76.4%)	.82
SM-003	−0.171 (−0.537 to 0.196)	−0.914	−15.7% (−41.6% to 21.6%)	.36
Age	0.018 (0.003 to 0.034)	2.315	1.9% (0.3% to 3.5%)	.02
Sex (Female)	2.247 (0.926 to 3.568)	3.334	846.2% (152.4% to 3446.3%)	<.001
Physical score	0.379 (0.166 to 0.592)	3.491	46.1% (18.1% to 80.8%)	<.001
Vision score	−0.104 (−0.175 to −0.034)	−2.920	−9.9% (−16% to −3.4%)	.004
DLTV[j] summary score (vision)	0.021 (0.005 to 0.038)	2.536	2.2% (0.5% to 3.9%)	.01
MPED[k] routine score	−0.038 (−0.075 to 0.000)	−1.981	−3.7% (−7.3% to 0%)	.05
Sex×physical score	−0.181 (−0.303 to −0.058)	−2.896	−16.5% (−26.2% to −5.7%)	.004

[a]Estimates (standardized β coefficient) are presented on the log scale from a generalized estimating equations Poisson model with a log link and an unstructured correlation structure. For categorical variables (product tested), estimates represent comparisons relative to the reference category (APD-001).

[b]95% CI for standardized β coefficient were calculated as β±1.96×SE. Percentage change in the expected error rate was computed as (e^β−1)×100.

[c]For continuous variables (eg, age and physical score), percentage change reflects the expected change in error rate per one-unit increase in the predictor.

[d]Corresponding 95% CIs for percentage change were derived by exponentiating the confidence limits of β, calculated as [(e^(β−1.96×SE)−1)×100, (e^(β+1.96×SE)−1)×100].

[e]Statistical significance was assessed at $P<.05$.

[f]APD: automated pill dispenser.

[g]Not applicable.

[h]PBA: pill box with alarm.

[i]SM: smart device.

[j]DLTV: Daily Living Tasks Dependent on Vision.

[k]MPED: Martin and Park Environmental Demands.

Discussion

Principal Findings

This study tested the usability and UX of 13 MATs among older adults with diverse cognitive, physical, sensory, motivational, and environmental capabilities. By focusing on performance-based usability metrics, such as unassisted task success rate, error rate, task completion time, and efficiency, this research extends previous MAT usability literature by examining not only device-level variability but also how individual user characteristics interact with device design during first-time use. These results are intended to provide an overview of how users interacted with a range of commercially available devices under the study conditions, rather than to rank devices or establish definitive performance comparisons. The devices included vary widely in features, functionality, and required task complexity, reflecting real-world options rather than standardized experimental comparators. Additionally, participants tested differing numbers of devices, and devices were not evaluated an equal number of times across participants.

XSL•FO
RenderX

A key finding was the variability in usability across MATs, regardless of device category or intended function. This indicates that usability cannot be inferred solely from whether a device is "smart" or "electronic," but instead reflects specific design features, interaction complexity, and task demands. These findings align with previous work by Patel et al [27], which reported wide variability in unassisted task success, task time, and workload across electronic medication adherence products, even when devices served similar purposes. This study extends this literature by demonstrating that such variability persists among newer-generation smart MATs and users with diverse functional profiles.

The observed variability reinforces the importance of performance-based usability testing. Although subjective measures such as the SUS capture user perceptions, previous research has shown discrepancies between perceived usability and objective task performance. Consistent with this, participants in this study often rated devices as acceptable despite frequent errors or difficulty completing setup tasks. More broadly, usability research has demonstrated weak correlations between subjective ratings and objective performance measures, particularly for complex or safety-critical technologies [50]. Collectively, these findings highlight the value of performance-based metrics for identifying usability risks that may not be captured through self-report alone, especially for MATs where interaction errors may have clinically significant consequences such as medication-related harm.

Regression analyses showed that cognitive, physical, visual, and motivational (SEAMS) scores were significant predictors of unassisted task success, consistent with previous literature on technology use among older adults [27,29]. Cognitive ability has been consistently linked to successful MAT use, with studies of automated dispensers and reminder systems reporting marked reductions in unassisted success among individuals with lower MMSE scores, even when impairment is mild [29]. Vision also emerged as a significant predictor, with both SMAT vision and DLTV scores associated with usability outcomes, echoing evidence that many technologies impose high visual demands and that inadequate accommodation of vision-specific barriers can compromise safe and effective use among older adults [51,52]. Motivational self-efficacy was likewise associated with higher success rates, consistent with findings that individuals with greater confidence are more likely to persist, implement effective strategies, and achieve better outcomes [53,54].

In contrast, most devices did not differ significantly from the reference device in error-rate models, with only 1 device (PBA 008) showing a statistically significant reduction. This apparent inconsistency reflects an important methodological distinction between task failure and task error. Errors were recorded only when tasks were completed with incorrect actions or multiple attempts, whereas failures were not coded as errors. As a result, devices with lower success rates may also show lower error rates simply because users were unable to progress far enough to generate observable errors. Conversely, devices that allowed task completion often did so with frequent missteps, yielding higher error rates despite greater success. From a medication safety perspective, successful task completion accompanied by repeated errors may pose greater real-world risk than early task abandonment. Similar conclusions have been reported in usability and human-factors research, which emphasizes that usability metrics represent distinct but interdependent constructs and should not be interpreted alone [50].

Some predictors of error rate were less intuitive. Higher physical scores were associated with increased error rates, likely because participants with better physical ability progressed further and interacted more extensively with device features, creating more opportunities for observable errors. Similarly, the positive association between DLTV scores and error rates likely reflects greater task exposure rather than poorer vision, as participants with better vision advanced into more complex interactions. These insights underscore the importance of interpreting error rates alongside task success and engagement measures rather than as standalone indicators [50].

Sex differences were evident across usability outcomes, with female participants demonstrating lower unassisted task success, higher error rates, and a significant interaction between sex and physical ability. These results suggest that sex may moderate how older adults interact with MATs, potentially reflecting differences in confidence, experience, and interaction strategies shaped by historical gender roles [55]. Although gender-related factors were not directly assessed, the consistency of sex effects across success and error outcomes highlights the need for sex-sensitive testing and inclusive design approaches to support safe and effective MAT use among older adults.

All devices were evaluated under untrained, first-use laboratory conditions to approximate real-world initial use. While appropriate for assessing early usability and learnability, laboratory testing cannot capture adaptive behaviors, learning over time, or caregiver involvement that may emerge during long-term home use. Evidence from longitudinal studies indicates that some older adults become more proficient and confident with technology with continued exposure [56]. Accordingly, the performance-based outcomes reported here should be interpreted as indicators of initial usability rather than long-term usability or continued adoption in everyday settings.

Study Strengths, Limitations, and Future Research Directions

A key strength of this study is the inclusion of older adults with multiple barriers, including motivational and environmental challenges, and the evaluation of both smart and electronic MATs. While this study provides valuable insights, several limitations should be acknowledged. Although the inclusion of participants with multiple coexisting impairments improves the real-world application, it may limit the ability to draw conclusions about the effects of a single impairment on the usability of a specific device. Future studies using more controlled designs, such as focusing on participants with a single type of impairment and testing a smaller number of devices, may help provide clearer device-specific conclusions. However, such designs may be less feasible in real-world aging populations, as most older adults experience a combination of impairments rather than a single isolated condition. The sample primarily consisted of Canadian English-speaking older adults, which may limit the generalizability of the findings to non-English speakers who might encounter additional

XSL•FO
RenderX

language-related barriers when using MATs. In addition, the initial sample size estimation was guided by Green's formula based on 6 primary impairment domains. However, the final Poisson GEE models included a broader set of predictors, including demographic variables. In addition, variability in the number of observations per device may result in an uneven distribution of data across predictors. While the overall sample size was adequate, these factors may introduce some risk of overfitting and affect the stability of model estimates.

Future research should explore the usability of these technologies in more diverse populations, including individuals with limited health literacy or linguistic diversity, as well as how device design and instructional materials can be optimized to support individuals with different and multiple barriers in safely and successfully completing tasks. Additionally, longitudinal research is needed to examine the long-term impact of MATs on medication adherence and health outcomes. Understanding how long-term use of these technologies influences adherence behaviors over time can provide valuable insights for health care providers and policymakers. Future studies should also explore how older adults interact with these technologies in real-world settings, including their reliance on instructional materials, learning approaches, and help-seeking behaviors. Moreover, this study only collected information on participants' sex. Future research should include both sex and gender measures to better capture sociocultural and biological influences on technology use.

Conclusion

This study highlights the role of performance-based usability testing in identifying the factors for successful use of MATs among older adults with diverse capabilities. By assessing unassisted task success rates, error rates, task completion times, and efficiency, the findings demonstrate substantial variability in usability across MATs, reflecting differences in device design, task complexity, and interaction demands rather than inherent device superiority.

Cognitive, physical, visual, motivational, and environmental factors were significant predictors of performance, highlighting that successful MAT use depends not only on technological capabilities but also on alignment with age-related abilities and confidence in medication management. The presence of higher error rates even among users who successfully completed tasks further highlights that usability and safety cannot be interpreted from task completion alone. Observed sex differences in performance reinforce the need for inclusive, user-centered design and evaluation approaches that account for diverse interaction styles and experiential backgrounds.

Acknowledgments

The authors would like to extend special thanks to Arshia Chauhan, Nischal Ghookal, Prapti Choudhuri, Hawa Patel, Tanaya Sharma, Victoria Nguyen, Karyman Ahmed Fawzy Ghanem, and Sabiha Lakhi for their valuable support with data collection.

This work was presented as a poster at the Canadian Pharmacy Education and Research Conference (CPERC) 2025.

No Generative AI was used for this manuscript.

Funding

This project was funded by Aging in Place program with National Research Council and Canadian Institutes of Health Research (reference 184372) by the Government of Canada.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

BB and TP contributed to all aspects of the study, including study concept and design, acquisition of participants and/or data, analysis and interpretation of data, and preparation of the manuscript. GE contributed for data acquisition and data entry. DH contributed to data analysis. SP, IH, and RS contributed to the acquisition of data. KP, AM, SJTG, FC, LL, CB, and RG contributed to the study concept and design. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

TP received grants from National Research Council (NRC), Canadian Institutes of Health Research, Jones Healthcare Group, and Custom Health; travel support from National Research Council (NRC) and Canadian Institutes of Health Research; and contribution of medical adherence technology from Jones Healthcare Group and Custom Health. The other authors report no conflicts of interests.

Multimedia Appendix 1
Mock medication regimen.
[DOCX File, 16 KB - aging_v9i1e88398_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Distribution of product testing by participant impairment type.
[DOCX File, 24 KB - aging_v9i1e88398_app2.docx]

Multimedia Appendix 3
Descriptive statistics for average task success rate (unassisted), total error rate, efficiency (unassisted), and total task time across participant subgroups and product types.
[DOCX File, 41 KB - aging_v9i1e88398_app3.docx]

Multimedia Appendix 4
Spearman correlation.
[DOCX File, 27 KB - aging_v9i1e88398_app4.docx]

Multimedia Appendix 5
Model diagnostic plots.
[DOCX File, 59 KB - aging_v9i1e88398_app5.docx]

References

1. Aging and chronic diseases: A profile of Canadian seniors. Government of Canada. 2020. URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/aging-chronic-diseases-profile-canadian-seniors-report.html [accessed 2025-01-17]
2. Griffith LE, Gruneir A, Fisher K, et al. Insights on multimorbidity and associated health service use and costs from three population-based studies of older adults in Ontario with diabetes, dementia and stroke. BMC Health Serv Res 2019 May 16;19(1):313. [doi: 10.1186/s12913-019-4149-3] [Medline: 31096989]
3. Wauters M, Elseviers M, Vaes B, et al. Polypharmacy in a Belgian cohort of community-dwelling oldest old (80+). Acta Clin Belg 2016 Jun;71(3):158-166. [doi: 10.1080/17843286.2016.1148298] [Medline: 27105401]
4. Jindai K, Nielson CM, Vorderstrasse BA, Quiñones AR. Multimorbidity and functional limitations among adults 65 or older, NHANES 2005-2012. Prev Chronic Dis 2016 Nov 3;13:E151. [doi: 10.5888/pcd13.160174] [Medline: 27809419]
5. Sharma P, Maurya P, Muhammad T. Number of chronic conditions and associated functional limitations among older adults: cross-sectional findings from the longitudinal aging study in India. BMC Geriatr 2021 Nov 23;21(1):664. [doi: 10.1186/s12877-021-02620-0] [Medline: 34814856]
6. Medication use in older adults. CAMH. 2025. URL: https://www.camh.ca/en/health-info/guides-and-publications/medication-use-in-older-adults#:~:text=Although%20adults%20aged%2065%20and,one%20medication%20at%20a%20time [accessed 2025-01-17]
7. Elliott RA, Goeman D, Beanland C, Koch S. Ability of older people with dementia or cognitive impairment to manage medicine regimens: a narrative review. Curr Clin Pharmacol 2015;10(3):213-221. [doi: 10.2174/1574884710666150812141525] [Medline: 26265487]
8. Maddigan SL, Farris KB, Keating N, Wiens CA, Johnson JA. Predictors of older adults' capacity for medication management in a self-medication program: a retrospective chart review. J Aging Health 2003 May;15(2):332-352. [doi: 10.1177/0898264303251893] [Medline: 12795276]
9. Abed H. Tools to assess functional medication management in older adults living at home: A scoping review. J Am Geriatr Soc 2019;67(Supplement 1).
10. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
11. Smith D, Lovell J, Weller C, et al. A systematic review of medication non-adherence in persons with dementia or cognitive impairment. PLoS ONE 2017;12(2):e0170651. [doi: 10.1371/journal.pone.0170651] [Medline: 28166234]
12. Lisberg RB, Higham C, Jayson MI. Problems for rheumatic patients in opening dispensed drug containers. Br J Rheumatol 1983 May;22(2):95-98. [doi: 10.1093/rheumatology/22.2.95] [Medline: 6850197]
13. McCann RM, Jackson AJ, Stevenson M, Dempster M, McElnay JC, Cupples ME. Help needed in medication self-management for people with visual impairment: case-control study. Br J Gen Pract 2012 Aug;62(601):e530-e537. [doi: 10.3399/bjgp12X653570] [Medline: 22867676]
14. Faisal S, Ivo J, Tennant R, et al. Integration of a smart multidose blister package for medication intake: a mixed method ethnographic informed study of older adults with chronic diseases. PLoS ONE 2022;17(1):e0262012. [doi: 10.1371/journal.pone.0262012] [Medline: 35061773]
15. Kripalani S, Henderson LE, Chiu EY, Robertson R, Kolm P, Jacobson TA. Predictors of medication self-management skill in a low-literacy population. J Gen Intern Med 2006 Aug;21(8):852-856. [doi: 10.1111/j.1525-1497.2006.00536.x] [Medline: 16881946]

XSL•FO
RenderX

16. Somerville E, Massey K, Keglovits M, et al. Scoring, clinical utility, and psychometric properties of the In-HOME Medication Management Performance Evaluation (HOME-Rx). Am J Occup Ther 2019;73(2):7302205060p1-7302205060p8. [doi: 10.5014/ajot.2019.029793] [Medline: 30915967]
17. Kuzuya M, Hirakawa Y, Suzuki Y, et al. Association between unmet needs for medication support and all-cause hospitalization in community-dwelling disabled elderly people. J Am Geriatr Soc 2008 May;56(5):881-886. [doi: 10.1111/j.1532-5415.2008.01676.x] [Medline: 18384585]
18. Field TS, Mazor KM, Briesacher B, Debellis KR, Gurwitz JH. Adverse drug events resulting from patient errors in older adults. J Am Geriatr Soc 2007 Feb;55(2):271-276. [doi: 10.1111/j.1532-5415.2007.01047.x] [Medline: 17302666]
19. Williams A, Manias E, Walker R. Interventions to improve medication adherence in people with multiple chronic conditions: a systematic review. J Adv Nurs 2008 Jul;63(2):132-143. [doi: 10.1111/j.1365-2648.2008.04656.x] [Medline: 18537843]
20. Woo K. Usability of Medication Adherence Technologies among Older Adults 2012.
21. Farooqi M, Carter C, Patel T. Electronic medication adherence technologies – Classification to guide use in older adults. Can Pharm J 2017;150:S26.
22. Faisal S, Samoth D, Aslam Y, et al. Key features of smart medication adherence products: updated scoping review. JMIR Aging 2023 Dec 19;6:e50990. [doi: 10.2196/50990] [Medline: 38113067]
23. Mason M, Cho Y, Rayo J, Gong Y, Harris M, Jiang Y. Technologies for medication adherence monitoring and technology assessment criteria: narrative review. JMIR Mhealth Uhealth 2022 Mar 10;10(3):e35157. [doi: 10.2196/35157] [Medline: 35266873]
24. Center for Technology and Aging. Technologies for Optimizing Medication Use in Older Adults 2009:15-17.
25. Sam T. Developing drug administration devices for geriatric use. In: Stegemann S, editor. Developing Drug Products in an Aging Society: Springer; 2016.
26. Baby B, McKinnon A, Patterson K, et al. Tools to measure barriers to medication management capacity in older adults: a scoping review. BMC Geriatr 2024 Mar 27;24(1):285. [doi: 10.1186/s12877-024-04893-7] [Medline: 38532328]
27. Patel T, Ivo J, Faisal S, et al. A prospective study of usability and workload of electronic medication adherence products by older adults, caregivers, and health care providers. J Med Internet Res 2020 Jun 2;22(6):e18073. [doi: 10.2196/18073] [Medline: 32348292]
28. ISO. Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)—Part 11: Guidance on Usability (ISO 9241-11:1998(E)) 1998.
29. Ligons FM, Mello-Thoms C, Handler SM, Romagnoli KM, Hochheiser H. Assessing the impact of cognitive impairment on the usability of an electronic medication delivery unit in an assisted living population. Int J Med Inform 2014 Nov;83(11):841-848. [doi: 10.1016/j.ijmedinf.2014.07.004] [Medline: 25153770]
30. Albert B, Tullis T. Measuring the User Experience, 1st edition: Morgan Kaufmann Publications; 2008. [doi: 10.1016/B978-0-12-373558-4.00001-7]
31. Sauro J, Lewis J. Quantifying the User Experience, 2nd edition: Morgan Kaufmann Publications; 2016. [doi: 10.1016/B978-0-12-802308-2.00002-3]
32. Albert B, Tullis T. Measuring the User Experience, 2nd edition: Morgan Kaufmann Publications; 2013.
33. Bailey RW, Wolfson CA, Nall J, Koyani S. Performance-based usability testing: metrics that have the greatest impact for improving a system's usability. Presented at: Human-Centered Design Proceedings of the First International Conference; Jul 19-24, 2009; San Diego, CA, USA p. 3-12. [doi: 10.1007/978-3-642-02806-9_1]
34. Green SB. How many subjects does it take to do a regression analysis. Multivariate Behav Res 1991 Jul 1;26(3):499-510. [doi: 10.1207/s15327906mbr2603_7] [Medline: 26776715]
35. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. Presented at: CHI '93: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems; May 1, 1993; Amsterdam, The Netherlands p. 206-213. [doi: 10.1145/169059.169166]
36. Nielsen J. Why you only need to test with 5 users. Neilsen Norman Group. 2000. URL: http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users [accessed 2014-06-22]
37. Baby B, Gill JK, Faisal S, et al. Medication adherence technologies: a classification taxonomy based on features. Mayo Clin Proc Digit Health 2025 Sep;3(3):100237. [doi: 10.1016/j.mcpdig.2025.100237] [Medline: 40881105]
38. Self-medication assessment tool (SMAT). Mount Allison University. URL: https://mta.ca/self-medication-assessment-tool-smat [accessed 2022-09-20]
39. Irvine-Meek J, Gould ON, Wheaton H, Todd LE. Acceptability and face validity of a geriatric self-medication assessment tool. Can J Hosp Pharm 2010 May;63(3):225-232. [doi: 10.4212/cjhp.v63i3.918] [Medline: 22478982]
40. Schmier JK, Halpern MT, Covert D. Validation of the Daily Living Tasks Dependent on Vision (DLTV) questionnaire in a U.S. population with age-related macular degeneration. Ophthalmic Epidemiol 2006 Apr;13(2):137-143. [doi: 10.1080/09286580600573049] [Medline: 16581618]
41. Grady PA, Gough LL. Self-management: a comprehensive approach to management of chronic conditions. Am J Public Health 2014 Aug;104(8):e25-e31. [doi: 10.2105/AJPH.2014.302041] [Medline: 24922170]
42. Raehl CL, Bond CA, Woods T, Patry RA, Sleeper RB. Individualized drug use assessment in the elderly. Pharmacotherapy 2002 Oct;22(10):1239-1248. [doi: 10.1592/phco.22.15.1239.33473] [Medline: 12389875]

XSL•FO
RenderX

43. Hearing tests. Patient. URL: https://patient.info/doctor/hearing-tests-pro [accessed 2022-09-20]
44. Risser J, Jacobson TA, Kripalani S. Development and psychometric evaluation of the Self-efficacy for Appropriate Medication Use Scale (SEAMS) in low-literacy patients with chronic disease. J Nurs Meas 2007;15(3):203-219. [doi: 10.1891/106137407783095757] [Medline: 18232619]
45. Son YJ, Won MH. Depression and medication adherence among older Korean patients with hypertension: mediating role of self-efficacy. Int J Nurs Pract 2017 Jun;23(3):e12525. [doi: 10.1111/ijn.12525] [Medline: 28194846]
46. Yang C, Hui Z, Zeng D, et al. A community-based nurse-led medication self-management intervention in the improvement of medication adherence in older patients with multimorbidity: protocol for a randomised controlled trial. BMC Geriatr 2021 Mar 2;21(1):152. [doi: 10.1186/s12877-021-02097-x] [Medline: 33653300]
47. Creech CL. The “medication matters... to you!” educational intervention session: a pilot study to improve medication management in community-living older adults [Dissertation]. : Texas Woman’s University; 2016 URL: https://twu-ir.tdl.org/items/b3432417-9032-4fff-b0f0-b079885cd276/full [accessed 2026-06-30]
48. Martin M, Park DC. The Martin and Park Environmental Demands (MPED) questionnaire: psychometric properties of a brief instrument to measure self-reported environmental demands. Aging Clin Exp Res 2003 Feb;15(1):77-82. [doi: 10.1007/BF03324483] [Medline: 12841422]
49. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 Feb;37(1):56-76. [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]
50. Hornbæk K. Current practice in measuring usability: challenges to usability studies and research. Int J Hum Comput Stud 2006 Feb;64(2):79-102. [doi: 10.1016/j.ijhcs.2005.06.002]
51. Orrico KB. Caring for visually impaired patients. J Am Pharm Assoc (2003) 2013;53(3):e142-e150. [doi: 10.1331/JAPhA.2013.13514] [Medline: 23699685]
52. Merenda T, Cannella S, Denis J, Patris S. Assistive products in pharmacy practice to optimize medications use for visually impaired patients: focus groups to explore community pharmacists’ opinions and expectations. Explor Res Clin Soc Pharm 2024 Sep;15:100467. [doi: 10.1016/j.rcsop.2024.100467] [Medline: 38983637]
53. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults’ experiences with technologies for health self-management: interview study. JMIR Aging 2023 Mar 21;6:e43197. [doi: 10.2196/43197] [Medline: 36943333]
54. Mekheimer M. Technological self-efficacy, motivation, and contextual factors in advanced EFL e-learning: a mixed-methods study of strategy use and satisfaction. Humanit Soc Sci Commun 2025;12(1):677. [doi: 10.1057/s41599-025-04947-0]
55. Czaja SJ, Charness N, Fisk AD, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352. [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
56. Jakob D, Wilhelm S, Gerl A, Ahrens D, Wahl F. Adapting voice assistant technology for older adults: a comprehensive study on usability, learning patterns, and acceptance. Digital 2025;5(1):4. [doi: 10.3390/digital5010004]

Abbreviations

APD: automated medication dispenser
DLTV: Daily Living Tasks Dependent on Vision
GEE: generalized estimating equation
MAT: medication adherence technology
MMSE: Mini-Mental State Examination
MPED: Martin and Park Environmental Demands
SEAMS: Self-Efficacy for Appropriate Medication Use Scale
SM: smart device
SMAT: Self-Medication Assessment Tool
SUS: System Usability Scale
USE: Usefulness, Satisfaction, and Ease of Use
UX: user experience

XSL•FO
RenderX

Edited by A Garcia; submitted 24.Nov.2025; peer-reviewed by HJ Chen, J Poots, P Santana-Mancilla, R Marshall; revised version received 02.Jun.2026; accepted 04.Jun.2026; published 13.Jul.2026.

Please cite as:
Baby B, Elba G, Park S, Hudani I, Sharma R, Patterson K, McKinnon A, Guilcher SJT, Chang F, Lee L, Burns C, Hajducek D, Griffin RH, Patel T
Performance-Based Usability of Medication Adherence Technologies Among Older Adults With Diverse Capabilities: Quantitative Study
JMIR Aging 2026;9:e88398
URL: https://aging.jmir.org/2026/1/e88398
doi:10.2196/88398

© Bincy Baby, Ghada Elba, SooMin Park, Imra Hudani, Rishabh Sharma, Kirk Patterson, Annette McKinnon, Sara J T Guilcher, Feng Chang, Linda Lee, Catherine Burns, Dagmar Hajducek, Ryan H Griffin, Tejal Patel. Originally published in JMIR Aging (https://aging.jmir.org), 13.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predictive Model of Acupuncture Adherence in Alzheimer Disease: Secondary Analysis of Randomized Controlled Trials

Ze-Hao Chen[1*], MM; Ran Li[2*], MD; Yu-Hang Jiang[1*], MM; Jia-Kai He[3], MD; Shan-Shan Yan[1], MM; Guan-Hua Zong[1], BM; Zong-Xi Yi[4], BM; Xin-Yu Ren[4], BM; Bao-Hui Jia[2*], MD, PhD

[1]Department of Acupuncture and Moxibustion, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, 5 Beixiange, Xicheng District, Beijing, China

[2]Department of Rehabilitation Medicine, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, 5 Beixiange, Xicheng District, Beijing, China

[3]Department of Traditional Chinese Medicine, Peking University People's Hospital, Beijing, China

[4]Department of Acupuncture and Moxibustion, Guang'anmen Hospital, Beijing University of Chinese Medicine, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Bao-Hui Jia, MD, PhD
Department of Rehabilitation Medicine, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, 5 Beixiange, Xicheng District, Beijing, China

Abstract

Background: The therapeutic efficacy of acupuncture in treating Alzheimer disease (AD) largely depends on consistent treatment adherence. Therefore, identifying key factors influencing adherence and developing targeted interventions are crucial for enhancing clinical outcomes.

Objective: This study aims to develop and validate a predictive model for identifying patients with AD who are likely to maintain good adherence to acupuncture treatment.

Methods: This secondary analysis included 108 patients with probable AD, aged 50 to 85 years, from 2 independent randomized controlled trials conducted at Guang'anmen Hospital, China Academy of Chinese Medical Sciences. Of all, 66 patients were assigned to the development cohort and 42 to the external validation cohort. Acupuncture adherence was defined as the proportion of completed sessions relative to scheduled sessions, with good adherence defined as ≥80% completion. Baseline data included demographic, clinical, cognitive, functional, psychological, and caregiving variables. Multivariable logistic regression with backward stepwise selection was used to identify significant predictors, and a nomogram was constructed based on the final model. Model performance was assessed using receiver operating characteristic curves, calibration plots, and decision curve analysis, with external validation performed by receiver operating characteristic analysis. Sensitivity analysis was performed using alternative adherence thresholds of 70% and 90%.

Results: A higher number of treatments during the first month was associated with a significant increase in the odds of good adherence (odds ratio [OR] 3.06, 95% CI 1.68 - 7.01; *P*=.002), while longer disease duration (OR 0.97, 95% CI 0.94 - 1.00; *P*=.049) and receiving care from a part-time caregiver (OR 0.19, 95% CI 0.04 - 0.72; *P*=.022) were associated with lower odds of adherence. Sensitivity analyses further supported the stability and reliability of the model.

Conclusions: This study is the first to develop and validate a predictive model for acupuncture adherence in patients with AD. In clinical research, it can facilitate participant stratification and help identify individuals who may need additional adherence support, thereby reducing bias and enhancing trial quality. In clinical practice, the nomogram enables proactive adherence management by prospectively identifying high-risk patients and guiding targeted strategies to improve adherence and optimize therapeutic outcomes.

(JMIR Aging 2026;9:e82787) doi:10.2196/82787

KEYWORDS

alzheimer disease; acupuncture; adherence; nomogram; predictive model

Introduction

Alzheimer disease (AD) is a progressive neurodegenerative disorder primarily characterized by cognitive decline, functional impairment in activities of daily living, and neuropsychiatric symptoms [1]. Acupuncture, given its favorable safety profile and potential for symptomatic improvement [2-4], has been recommended as a promising nonpharmacological long-term

XSL•FO
RenderX

therapy in clinical practice guidelines for the management of AD [5]. However, the therapeutic efficacy of acupuncture for AD remains a subject of ongoing debate [6,7]. This inconsistency in findings may be partially explained by a critical, yet often underinvestigated factor: treatment adherence.

Consistent adherence is crucial to therapeutic efficacy, particularly in chronic disease management and nonpharmacological interventions. At present, adherence has been studied primarily in the context of pharmacotherapy, with reported rates of long-term pharmacological adherence in patients with AD ranging from 16.5% to 51% [8,9]. Some studies have developed models to predict patients' medication adherence [10,11]. In contrast, adherence to nonpharmacological therapies such as acupuncture remains insufficiently explored. This knowledge gap is compounded by the lack of a standardized definition for what constitutes good or poor adherence. This variability in definitions across studies makes it difficult to compare findings and may be a key confounding factor obscuring the true dose-response relationship of acupuncture [12].

Clinical studies have shown that factors influencing acupuncture adherence are multifaceted, encompassing patient subjective norms, perceived behavioral control, and treatment commitment [13]. Specific factors include perceived effectiveness, family support, emotional status, patient recognition and acceptance of acupuncture, and the availability of medical subsidies [14]. Nonetheless, most of these studies have provided only descriptive insights or general recommendations based on literature reviews [15,16], while few have offered quantitative evidence derived from clinical data. Notably, research specifically focusing on acupuncture adherence in individuals with AD is exceedingly scarce. Given the substantial cognitive decline, behavioral symptoms, and caregiver dependence associated with AD, patients in this population may face unique barriers to maintaining adherence. Thus, findings from studies in other populations may not be directly generalizable to patients with AD. Understanding the specific determinants of adherence in this unique population is crucial for optimizing treatment delivery and improving outcomes.

Given the ongoing debate over acupuncture's efficacy and the critical yet poorly understood role of adherence in treatment outcomes, identifying predictors for treatment engagement in a methodologically robust setting is paramount. In this study, we conducted a secondary analysis of data from 2 independent randomized controlled trials (RCTs) previously conducted by our team, with the aim of identifying key predictors of acupuncture adherence among patients with AD and developing a clinically applicable predictive model. By identifying individuals at high risk of poor adherence and enabling the development of targeted intervention strategies, this study seeks to enhance the adherence and efficacy of acupuncture treatment and provide evidence-based support for adherence management in nonpharmacological interventions for AD.

Methods

Study Design and Data Source

This study was a secondary analysis based on data collected from 2 independent RCTs conducted at Guang'anmen Hospital, China Academy of Chinese Medical Sciences (CACMS). Participants were primarily recruited between December 2021 and June 2024. The first RCT, conducted from December 2021 to June 2024, enrolled 66 patients and served as the model development cohort. The second RCT, conducted from June 2022 to November 2022, included 42 patients and served as an external validation cohort for the predictive model.

Eligibility screening and clinical assessments were conducted by licensed physicians from the departments of encephalopathy and neurology. Inclusion criteria and exclusion criteria are present in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

- **Inclusion criteria:**
  - A diagnosis of probable Alzheimer disease according to the National Institute on Aging–Alzheimer's Association criteria [17]
  - Age between 50 and 85 years
  - A Clinical Dementia Rating score ≥0.5
  - A Mini-Mental State Examination score ≤26
  - A Hachinski Ischemic Scale score ≤4
- **Exclusion criteria:**
  - Other neurological or systemic disorders known to cause progressive cognitive impairment
  - Recent use of medications or exposure to substances known to impair cognition
  - A history of trypanophobia or active skin infections
  - Acupuncture or electroacupuncture treatment within the past 2 weeks
  - Participation in other clinical trials

XSL•FO
RenderX

Ethical Considerations

Both studies were conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. This study is a secondary analysis of data derived from 2 RCTs previously conducted at Guang'anmen Hospital, CACMS. Both original trials received independent ethical approval from the ethics committee of Guang'anmen Hospital, CACMS (approval: 2021-056-KY-01 and 2022-087-KY). Written informed consent was obtained from all participants or their legally authorized representatives prior to data collection. Participants did not receive compensation for participation in the original trials. All treatments and assessments were provided free of charge. To protect participant privacy, all data used in this secondary analysis were anonymized and deidentified prior to analysis.

Treatment and Adherence Assessment

Participants received 20-minute acupuncture sessions 3 times per week (on nonconsecutive days) for a total of 12 weeks. Adherence was assessed by calculating the proportion of completed treatment sessions relative to the total number of scheduled sessions during the intervention period. The proportion of days covered (PDC) was used as the adherence metric. In the absence of an established gold standard for adherence in nonpharmacological trials like acupuncture, we adopted the widely accepted threshold of PDC ≥80% from pharmacotherapy research [9]. This threshold is a well-validated proxy for sufficient exposure to treatment in chronic disease management. Participants with a PDC ≥80% were classified as having good adherence, while those with a PDC <80% were categorized as having poor adherence.

Data Collection

The following clinical data were collected: sex, age, disease duration, disease severity, educational level, occupation, history of acupuncture treatment, Mini-Mental State Examination score, Alzheimer's Disease Assessment Scale–Cognitive Subscale score, basic activities of daily living, instrumental activities of daily living, presence of behavioral and psychological symptoms of dementia, Patient Health Questionnaire-9 score for depressive symptoms, caregiving role, number of treatments in the first month, travel time to the hospital, recruitment method, and adherence outcomes.

Screening of Influencing Factors

A linear regression model was used to assess multicollinearity among independent variables. Multicollinearity was quantified by calculating the variance inflation factor and tolerance values. Variables with a variance inflation factor <10 and tolerance >0.1 were considered to have acceptable levels of multicollinearity [18]. Variables meeting these criteria were subsequently included in a multivariable logistic regression model. Variable selection was performed using a backward stepwise regression approach based on the likelihood ratio test, with a significance threshold of $P<.05$ for retention in the model. This data-driven approach was chosen to build a parsimonious model and reduce the risk of overfitting. We also conducted exploratory analyses by forcing clinically relevant but nonsignificant variables into the model, but this did not lead to a significant improvement in model performance and increased model complexity. Therefore, the final model retained only the statistically significant predictors. Final variables were required to meet criteria for statistical significance, low multicollinearity, and satisfactory predictive performance. To improve model interpretability, Shapley Additive Explanations (SHAP) values were calculated to quantify the relative contribution of each predictor to the outcome [19]. SHAP values were visualized using beeswarm plots.

Construction and Evaluation of the Nomogram

A nomogram was developed based on the final multivariable logistic regression model to predict adherence. The nomogram serves as a graphical tool to visualize the relationships between multiple predictors and the outcome, facilitating individualized risk assessment and clinical decision-making. In the nomogram, each predictor is aligned with its corresponding axis; by drawing a vertical line from the predictor's value to the point scale, a score can be assigned. The total score, obtained by summing the individual scores, corresponds to a predicted probability of adherence on the nomogram's outcome axis.

Model performance was assessed through receiver operating characteristic (ROC) curve analysis, calibration plots, the Hosmer-Lemeshow goodness-of-fit test, and decision curve analysis using the development dataset. Internal validation was conducted using the bootstrap method with 1000 resampling iterations. Agreement between predicted and observed outcomes was evaluated using the κ statistic. An area under the ROC curve (AUC) between 0.5 and 0.7 was interpreted as indicating low discrimination, 0.7 to 0.9 as moderate, and >0.9 as high discrimination [20]. The nomogram's robustness was further evaluated by performing ROC analysis in the external validation cohort.

Statistical Methods

Data completeness for the variables included in the final model was assessed prior to analysis. There were no missing values for the variables included in the final model in either the development or validation cohorts, as complete data collection was a requirement for the per-protocol analysis in the 2 RCTs. All statistical analyses were conducted using R software (version 4.4.2; R Foundation for Statistical Computing). Continuous variables with a normal distribution were presented as mean (SD), while nonnormally distributed variables were expressed as median (IQR), with distribution assessed using the Shapiro-Wilk test. Categorical variables were reported as frequencies and percentages. Group comparisons were performed using independent-samples *t* tests for normally distributed continuous variables and the Mann-Whitney *U* test for nonnormally distributed continuous variables. Categorical variables were compared using the chi-square test or the Fisher exact test, as appropriate. All statistical tests were 2-tailed, and a P value <.05 was considered statistically significant. To assess the robustness of our model to the primary adherence definition (PDC ≥80%), a sensitivity analysis was performed by repeating the multivariable logistic regression using alternative thresholds of 70% and 90%.

XSL•FO
RenderX

Results

Baseline Characteristics of the Training Set

The predictive model was developed using data from the development cohort (cohort 1), which comprised 66 patients with AD (Table 1). Participants were stratified by treatment adherence into a good adherence group (n=43) and a poor adherence group (n=23). Of the total participants, 34 (51.5%) were female participants and 32 (48.5%) were male participants. The mean age was 71.8 (SD 7.9) years, and the mean disease duration was 50.0 (SD 26.0) months. Univariate analysis revealed significant differences between the good and poor adherence groups in caregiving status (P=.025) and the number of treatment sessions during the first month (P=.001).

The P values for testing differences between patients with good and poor adherence to acupuncture treatment were derived from independent samples t tests, Mann-Whitney U tests and chi-square tests or the Fisher exact tests.

XSL•FO
RenderX

Table . Comparison of baseline characteristics between patients with good and poor adherence (n=66).

Variable	Total	Good adherence (n=43)	Poor adherence (n=23)	*t* test (*df*)[a]	Wilcoxon rank-sum test[a]	Chi-square (*df*)	*P* value
Sex, n (%)				—	—[b]	1.88 (1)	.171
Female	34 (52)	19 (44)	15 (65)				
Male	32 (48)	24 (56)	8 (35)				
Age (y), mean (SD)	71.8 (7.9)	70.9 (7.7)	73.5 (8.3)	−1.28 (64)	—	—	.204
Disease duration (mon), mean (SD)	50.0 (26.0)	48.1 (26.0)	53.5 (26.2)	−0.80 (64)	—	—	.427
Disease severity,[c] n (%)				—	—	—	.490
Mild	29 (44)	21 (49)	8 (35)				
Moderate	28 (42)	16 (37)	12 (52)				
Severe	9 (14)	6 (14)	3 (13)				
Education level, n (%)				—	—	0.05 (1)	.816
No higher education	40 (61)	27 (63)	13 (57)				
Higher education	26 (39)	16 (37)	10 (43)				
Occupation, n (%)				—	—	0.00 (1)	.99
Manual work	21 (32)	14 (33)	7 (30)				
Nonmanual work	45 (68)	29 (67)	16 (70)				
MMSE[d], median (IQR)	16.5 (11.0 - 21.0)	18.0 (12.0 - 21.5)	16.0 (9.5 - 20.0)	—	582	—	.241
ADAS-Cog[e], median (IQR)	22.0 (15.0 - 38.8)	22.0 (14.5 - 39.0)	24.0 (16.0 - 38.5)	—	454	—	.590
BADL[f], median (IQR)	10.0 (9.0 - 13.0)	10.0 (9.0 - 12.5)	11.0 (9.5 - 13.5)	—	444.5	—	.500
IADL[g], median (IQR)	26.0 (18.2 - 33.0)	27.0 (16.0 - 32.5)	25.0 (19.5 - 33.5)	—	472.5	—	.772
BPSD[h],[c] n (%)				—	—	—	.99
Present	53 (80)	34 (79)	19 (83)				
Absent	13 (20)	9 (21)	4 (17)				
PHQ-9[i],[c] n (%)				—	—	—	.99
Depressive symptoms	3 (5)	2 (5)	1 (4)				
Normal	63 (95)	41 (95)	22 (96)				
Travel time to the hospital (min), median (IQR)	60.0 (40.0 - 100.0)	50.0 (35.0 - 95.0)	60.0 (50.0 - 135.0)	—	413	—	.272
First-month treatment sessions, median (IQR)	11.0 (11.0 - 12.0)	12.0 (11.0 - 12.0)	10.0 (8.0 - 12.0)	—	736	—	.001
Caregiving role, n (%)				—	—	5.05 (1)	.025

XSL•FO
RenderX

Variable	Total	Good adherence (n=43)	Poor adherence (n=23)	*t* test (*df*)[a]	Wilcoxon rank-sum test[a]	Chi-square (*df*)	*P* value
Full-time caregiver	34 (52)	27 (63)	7 (30)				
Part-time caregiver	32 (48)	16 (37)	16 (70)				
History of acupuncture, n (%)				—	—	0.00 (1)	.99
No	30 (45)	20 (47)	10 (43)				
Yes	36 (55)	23 (53)	13 (57)				
Recruitment method[c], n (%)				—	—	—	.844
Nursing home	4 (6)	3 (7)	1 (4)				
Multimedia	30 (45)	19 (44)	11 (48)				
Study team referral	15 (23)	11 (26)	4 (18)				
Outpatient clinic	17 (26)	10 (23)	7 (30)				

[a]Continuous variables were compared using independent-samples *t* tests when normally distributed and Wilcoxon rank-sum tests otherwise. Categorical variables were compared using the chi-square test or Fisher exact test, as appropriate.

[b]—: not applicable.

[c]Evaluated using the Fisher exact test.

[d]MMSE: Mini-Mental State Examination.

[e]ADAS-Cog: Alzheimer's Disease Assessment Scale–Cognitive Subscale.

[f]BADL: basic activities of daily living.

[g]IADL: instrumental activities of daily living.

[h]BPSD: behavioral and psychological symptoms of dementia.

[i]PHQ-9: Patient Health Questionnaire-9.

Independent Predictors of Adherence to Acupuncture Treatment Among Patients with AD

Initially, 17 potential predictors of adherence to acupuncture treatment in patients with AD were considered. Following multicollinearity analysis and assessment of clinical relevance, 16 variables were ultimately selected for model construction (Table S1 in Multimedia Appendix 1). Multivariable logistic regression analysis was performed to develop a predictive model of adherence to acupuncture treatment in this population. The final model is represented by the following equation (Figure 1; Table S2 in Multimedia Appendix 1):

$$\log(p/(1-p)) = -2.44 + 0.03 \times \text{Disease Duration} + 1.20 \times \text{Number of Treatments in First Month} - 1.63 \times \text{Caregiving Role}$$

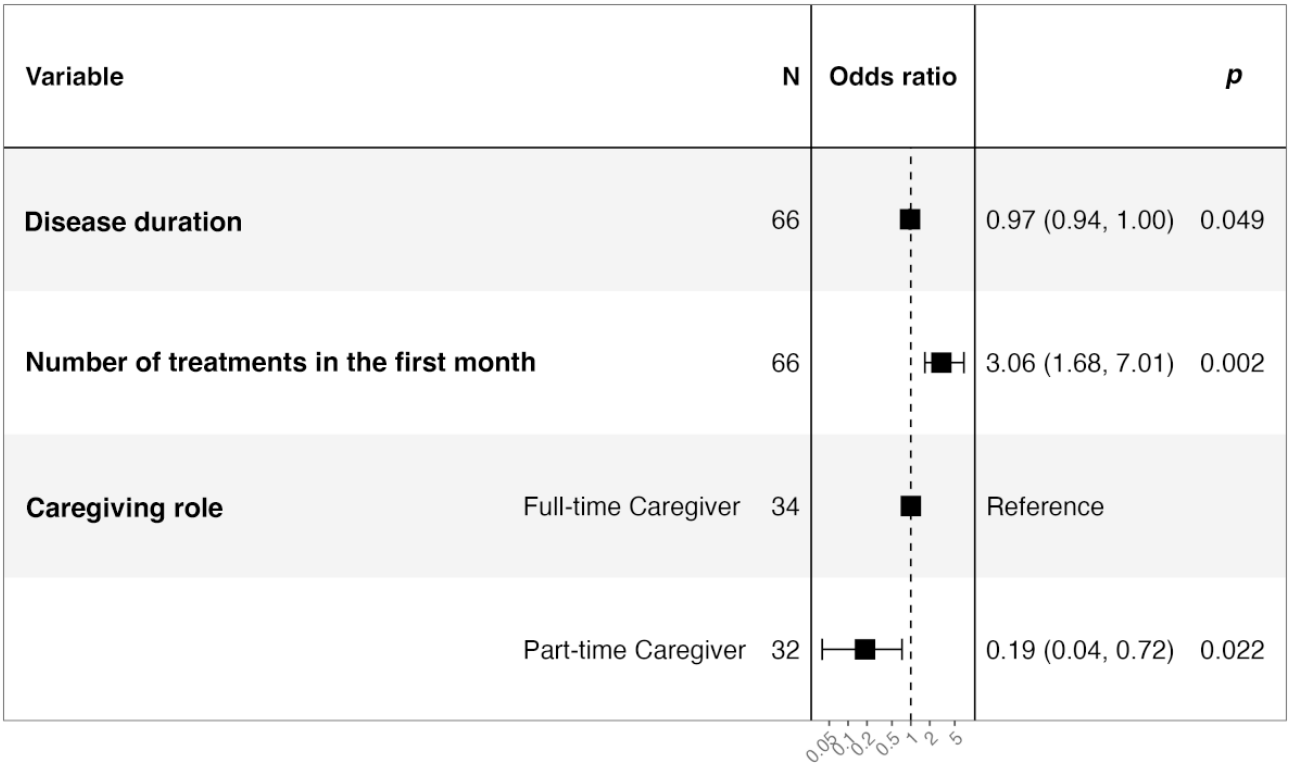
Disease duration, the number of treatments in the first month, and caregiving role were independent predictors of adherence to acupuncture treatment among patients with AD. A higher number of treatments during the first month was associated with a significant increase in the odds of good adherence (odds ratio [OR] 3.06, 95% CI 1.68 - 7.01; *P*=.002), while longer disease duration (OR 0.97, 95% CI 0.94 - 1.00; *P*=.049) and receiving care from a part-time caregiver (OR 0.19, 95% CI 0.04 - 0.72; *P*=.022) were associated with lower odds of adherence. SHAP analysis quantified the contributions of these predictors, confirming their importance in the model (Figure 2). The SHAP results were consistent with the logistic regression findings, enhancing the interpretability of the model.

The results indicated that the number of treatment sessions during the first month, caregiving role, and disease duration were the 3 most important factors influencing adherence. This ranking was highly consistent with the findings of the logistic regression model, demonstrating the substantial contribution of these variables to the model's predictive performance.

The SHAP summary plot visualizes the impact of each variable on the predicted probability of adherence. Each dot represents the SHAP value of an individual observation, indicating the degree and direction of that variable's influence on the model output. Higher SHAP values correspond to a stronger positive contribution to adherence probability. The color gradient from blue to red indicates the relative value of each variable for that observation (blue =lower value; red =higher value). For example, a higher number of treatments in the first month (red) is associated with higher SHAP values, reflecting its positive impact on adherence.

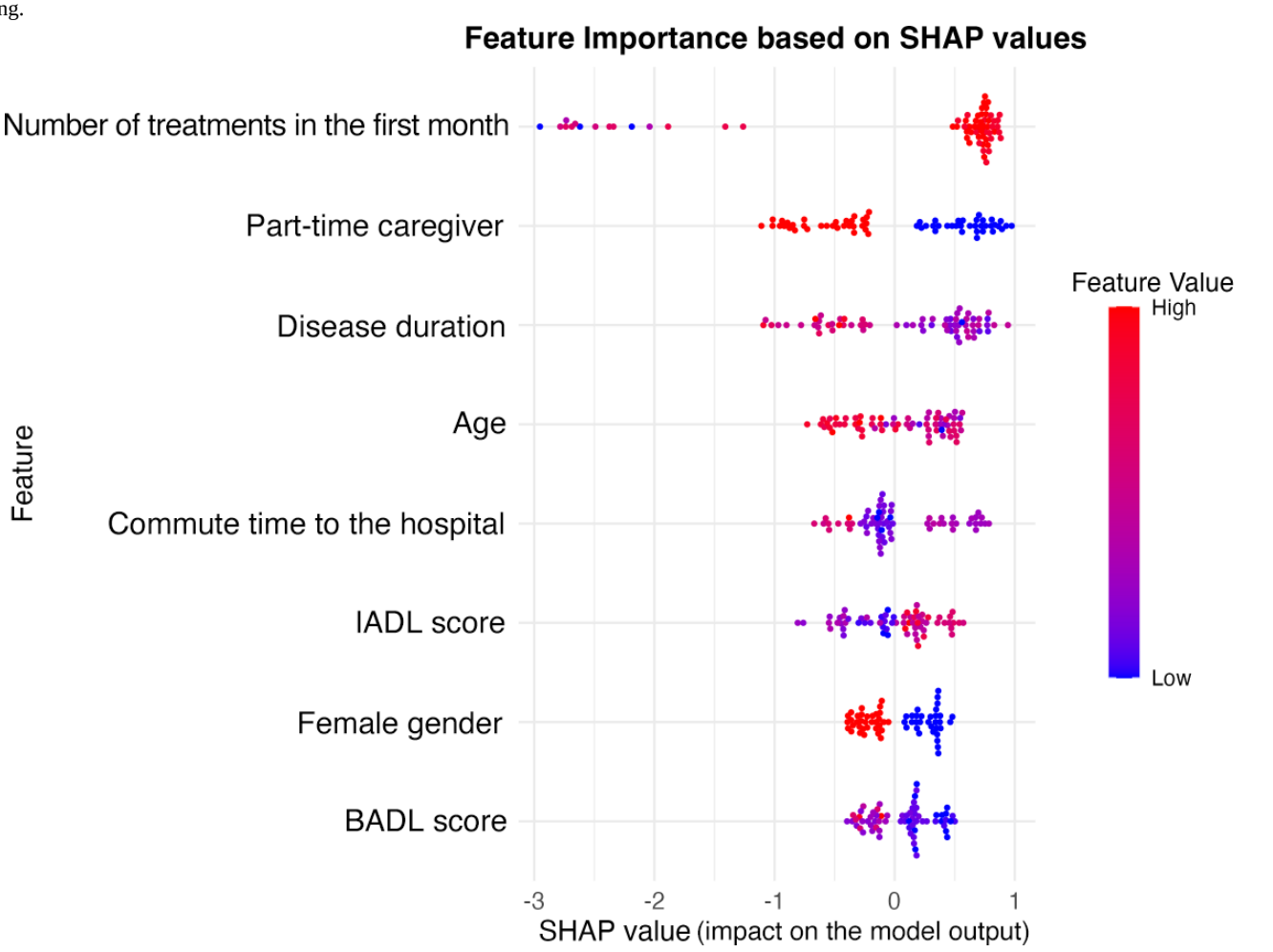
XSL•FO
RenderX

Figure 1. Forest plot of factors influencing adherence to acupuncture treatment among patients with Alzheimer disease based on multivariable logistic regression analysis. Dots represent odds ratios, and horizontal lines indicate 95% CIs. Longer disease duration and care provided by part-time caregivers were associated with lower adherence (odds ratio<1), while a higher number of treatment sessions during the first month significantly increased adherence (odds ratio >1). All variables included in the model were statistically significant (*P*<.05).



XSL•FO
RenderX

Figure 2. Shapely Additive Explanations (SHAP) summary plot of variable importance for predicting adherence to acupuncture treatment in patients with Alzheimer disease based on multivariable logistic regression analysis. BADL: basic activities of daily living; IADL: instrumental activities of daily living.



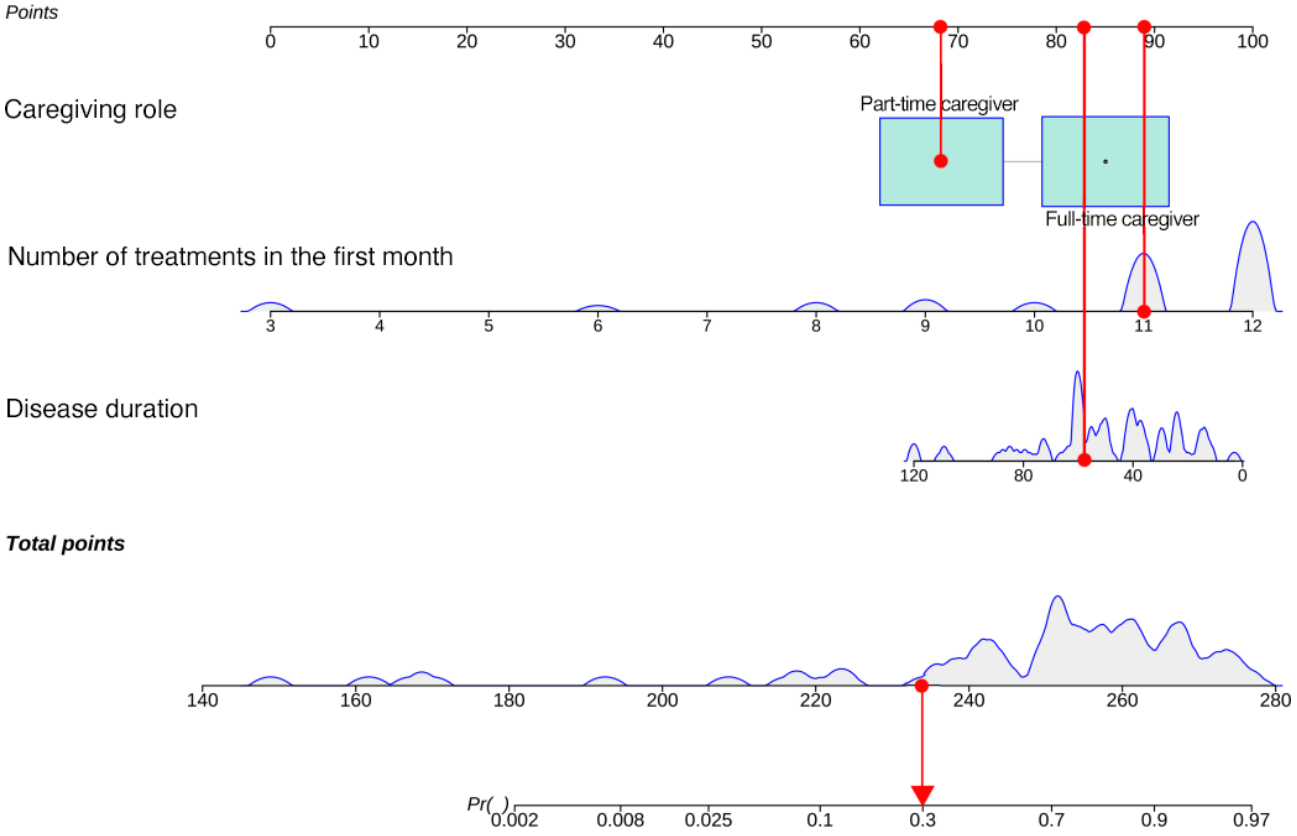
Development and Evaluation of the Nomogram

A nomogram was developed based on the final predictive model, incorporating the number of treatments in the first month, disease duration, and caregiving role (Figure 3). The model demonstrated excellent discrimination, with an area under the ROC curve (AUC) of 0.914 (Figure 4A). Validation was performed using cohort 2 (test set), which included 42 patients with AD (Table S3 in Multimedia Appendix 1). The calibration curve from internal validation indicated good agreement between predicted and observed probabilities (mean absolute error =0.04; mean squared error =0.003; 90th quantile absolute error =0.078) (Figure 4B). The Hosmer-Lemeshow goodness-of-fit test showed that the model predictions were well calibrated (χ^2_8=10.9; *P*=0.21). The decision curve analysis showed that the nomogram provided a net clinical benefit across a wide range of threshold probabilities (Figure 4D). Internal validation yielded an overall predictive accuracy of 89.4%, with a κ statistic of 0.759, indicating good consistency and potential clinical utility. Furthermore, the model performed well in the external validation cohort, achieving an AUC of 0.833 (Figure 4C).

For example, a patient with AD with a disease duration of 60 months, part-time caregiving, and 11 acupuncture sessions during the first month would obtain corresponding point values on each variable axis, indicated by red dots on the upper horizontal scales. The points for the 3 variables are summed to yield a total score of approximately 234. On the bottom axis, a total score of 234 corresponds to a predicted adherence probability of approximately 0.30, as marked by the red arrow. This relatively low predicted probability highlights the need for proactive adherence management and targeted interventions in clinical practice.

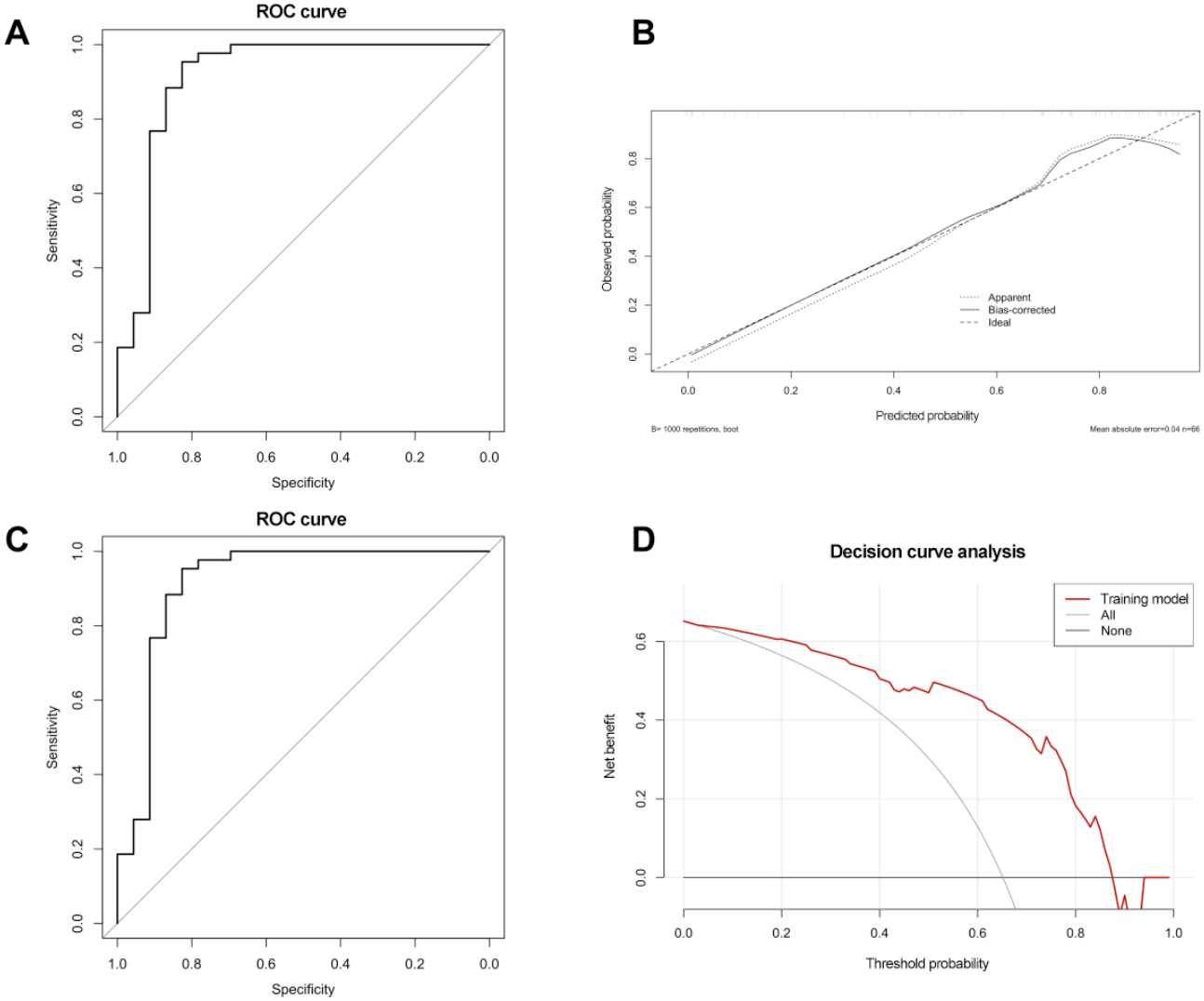
XSL•FO
RenderX

Figure 3. Nomogram for predicting adherence to acupuncture treatment among patients with Alzheimer disease.



XSL•FO
RenderX

Figure 4. Evaluation and validation of the nomogram model for predicting adherence to acupuncture treatment in patients with Alzheimer disease. (A) Receiver operating characteristic (ROC) curve assessing the discriminative ability of the model, with an area under the ROC curve of 0.914, indicating excellent predictive performance. (B) Calibration curve illustrating agreement between predicted and observed adherence probabilities. The x-axis represents predicted probability, while the y-axis shows the observed probability of adherence. The Ideal line indicates perfect concordance between predicted and actual outcomes. The apparent line reflects model performance in the original sample, and the bias-corrected line shows model performance adjusted for overfitting using the bootstrap method (1000 resamples). The close alignment of the calibration curves with the ideal line demonstrates the high accuracy and reliability of the model. (C) ROC curve of the nomogram model in the validation cohort, with an area under the ROC curve of 0.838. (D) Decision curve analysis of the nomogram model in the training cohort. The x-axis represents the threshold probability, and the y-axis indicates the net clinical benefit across different thresholds.



Sensitivity Analysis

To evaluate the robustness of our model, a sensitivity analysis was performed using alternative adherence thresholds of 70% and 90%. The model's predictive performance remained strong across all definitions, with an area under the ROC curve (AUC) of 0.892 for the 70% threshold and 0.860 for the 90% threshold. The statistical significance of the number of treatments in the first month (P=.01 at 70%; P=.05 at 90%), caregiving role (P=.03 at 90%), and disease duration (P=.08 at 70%; P=.09 at 90%) fluctuated near the P=.05 cutoff; this could be caused by the dataset size. Overall, these findings support the stability of our model. Detailed results of the sensitivity analysis are provided in the supplementary materials (Figure S1 in Multimedia Appendix 1).

Adverse Events

Safety monitoring was conducted throughout both RCTs. No serious adverse events related to acupuncture treatment were reported in either the development cohort or the validation cohort. Adverse events unrelated to the intervention are summarized in Table S4 in Multimedia Appendix 1.

Discussion

Predictors of Treatment Adherence in Patients with AD

This study identified the number of treatments in the first month, caregiving role, and disease duration as significant predictors of adherence to acupuncture treatment among patients with AD. The predictive nomogram constructed with these 3 variables demonstrated excellent discrimination in the development cohort (AUC=0.914) and acceptable performance in the external

XSL•FO
RenderX

validation cohort (AUC=0.838). With its simplicity, interpretability, and ease of application, the nomogram may serve as a practical tool to assist clinicians in identifying patients at risk of poor adherence and tailor intervention strategies accordingly.

To understand the clinical implications of these predictors, it is useful to interpret them through the established framework of intentional versus unintentional nonadherence. Unintentional nonadherence typically arises from practical barriers, which can be internal to the patient's condition (eg, forgetfulness and functional limitations) or external and situational (eg, transportation difficulties, inclement weather, and systemic disruptions like a pandemic). In contrast, intentional nonadherence involves a deliberate decision to cancel the treatment plan, often driven by subjective factors such as a perceived lack of efficacy, treatment fatigue, or shifting personal priorities. Our findings suggest that the identified predictors likely influence adherence through mechanisms related to both categories.

Increasing the Number of Treatments in the First Month May Improve Adherence

A key finding of this study is that a higher number of treatments during the first month was a powerful predictor of adherence. This suggests that early and intensive engagement is critical for establishing sustained treatment behaviors. From the perspective of unintentional nonadherence, a structured, frequent schedule in the initial phase may help patients and caregivers with cognitive and organizational deficits to more quickly accept the therapeutic routine, making it a habitual part of their lives [21].

This intensive approach can also mitigate intentional nonadherence through two possible ways. First, given that the therapeutic effect of acupuncture is often cumulative [22], increasing treatment frequency may accelerate the perception of clinical benefits [23,24]. When patients and caregivers observe an improvement early on, their motivation and belief in the treatment's value are naturally reinforced [25]. Second, frequent sessions offer more opportunities for communication among patients, caregivers, and clinicians, enabling the early detection and management of emerging issues. However, this approach requires careful consideration. It is important to acknowledge that a higher treatment frequency does not universally guarantee improved clinical outcomes [26] and may increase the treatment burden on families. The optimal number and timing of sessions may vary depending on disease severity, stage, and individual patient needs [27]. In conclusion, tailoring the intensity of acupuncture interventions to individual profiles remains a key consideration for future research and clinical practice for patients with AD, balancing clinical benefit and treatment burden.

Caregiver Capacity Is a Critical Determinant of Adherence

Our finding that patients supported by part-time caregivers (defined as providing fewer than 41 h of care per week [28]) were significantly less likely to adhere underscores the critical role of caregiver capacity in treatment engagement. This highlights a powerful driver of unintentional nonadherence, as dementia caregiving is a long-term, high-burden undertaking [29-31]. In advanced stages, care demands can often exceed 100 hours per week [32]. This presents formidable logistical barriers for part-time caregivers, who are typically adult children trying to balance employment, their own household duties, and the challenge of not living with the person they care for. These challenges frequently limit their ability to schedule and accompany patients to appointments, provide consistent emotional support, or respond to emergent care needs, thereby compromising adherence despite their best intentions [29].

This relentless demand also takes a significant physical and psychological toll, leading to caregiver burden and fatigue that can further diminish the capacity to support treatment [33-37]. These inherent difficulties are often compounded by unpredictable external factors, such as sudden illness or bad weather, which can disproportionately disrupt the routines of caregivers with less flexibility. A more subtle yet powerful factor is a form of intentional nonadherence driven by altruism, where patients may forgo appointments out of a desire not to burden their children, ultimately leading to treatment discontinuation.

In contrast, spouses often serve as full-time caregivers (≥41 h/wk) [38,39] and are typically more emotionally invested and committed to maintaining treatment routines. Some may even retire early to provide round-the-clock care [40]. Our study also found that professional caregiving within institutional settings can offer stable, structured, and high-quality care, which may facilitate better adherence. To enhance adherence in clinical practice, health care providers should actively involve caregivers in treatment planning and offer tailored education to improve their understanding of disease progression, treatment goals, and the importance of adherence [41,42]. Such interventions can mitigate both unintentional nonadherence (by improving scheduling and problem-solving skills) and intentional nonadherence (by reinforcing the perceived value of the treatment). Strengthening caregivers' motivation and capacity to support treatment is essential. When informal caregiving resources are insufficient, incorporating professional home care services or transitioning to institutional care may help ensure treatment continuity and effectiveness. Notably, countries such as Denmark and the Netherlands have developed comprehensive formal care systems for dementia that integrate medical and social support services [30,43]. These models may provide valuable reference points for improving dementia care infrastructure in China.

Early Detection, Prevention, and Intervention Still Key to Treating AD

Our study found that longer disease duration was associated with poorer adherence to acupuncture treatment. Primarily, this is a form of unintentional nonadherence driven by the patient's own progressive cognitive and functional decline, which impairs their capacity to independently manage appointments [44]. A higher prevalence of neuropsychiatric symptoms can reduce patient cooperation, while accumulating physical comorbidities and mobility limitations create new logistical hurdles. As the disease advances, a cascade of factors converges to further undermine adherence. This decline simultaneously increases

XSL•FO
RenderX

the burden on caregivers, who may experience emotional exhaustion and a deterioration in their own health, diminishing their ability to provide consistent support [45]. The combination of increasing patient dependency, rising caregiver exhaustion, and mounting logistical obstacles creates a formidable barrier to sustained treatment in the later stages of the disease.

This underscores the critical importance of a proactive and early approach to management. Our findings emphasize that the timing of intervention is paramount. Initiating treatment when patients retain greater cognitive and functional capacity offers a crucial window of opportunity. Early initiation of acupuncture may help establish regular treatment routines, foster therapeutic rapport, and enhance patient motivation. Moreover, health care providers should prioritize early education and ongoing support for both patients and caregivers. This includes training in caregiving skills, psychological counseling, and practical strategies to reduce caregiver burden and enhance quality of life [46]. These factors can contribute not only to improved adherence but also to more favorable long-term outcomes by potentially slowing the trajectory of disease progression.

Limitations

This study has limitations. First, although this study suggests that early, intensive treatment may enhance adherence, the optimal frequency and total number of acupuncture sessions for patients with AD remain undetermined. In real-world settings, increasing treatment frequency may impose greater transportation, time, and financial burdens on both patients and caregivers, potentially reducing their motivation and adherence. Second, our analysis was based on data from previous RCTs that were provided free of charge. Consequently, we could not assess the influence of crucial socioeconomic factors, such as treatment costs, or household income, which are known to be factors of health care engagement. Their influence on adherence may have been underestimated. Third, our sample size (N=108) was adequate for the primary analysis; it may be underpowered to detect predictors with more subtle effects. Finally, although we performed external validation, both cohorts were recruited from a single hospital. This shared clinical and demographic context limits the generalizability of our nomogram. Studies are needed to validate our model in multicenter or community-based cohorts to confirm its broader applicability.

Conclusions

In conclusion, this study is the first to develop and validate a predictive model for acupuncture treatment adherence in patients with AD, offering a novel, evidence-based tool for both clinical research and practice. For clinical research, this model provides a method to stratify enrollment or identify participants who may require enhanced adherence support, thereby reducing bias and improving the integrity of future trials. In clinical practice, the nomogram enables a shift from reactive problem-solving to proactive adherence management. By prospectively identifying patients at high risk, clinicians can address specific barriers and implement targeted strategies to improve adherence and, ultimately, enhance therapeutic outcomes.

Acknowledgments

The authors extend their heartfelt gratitude to all patients with Alzheimer disease and their family members for their participation in this study, without whom this research would not have been possible. During the preparation of this work, the authors used ChatGPT to improve the language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the final content of the publication.

Funding

This study was supported by the Clinical Research and Achievement Transformation Capacity Improvement Project of High-Level Traditional Chinese Medicine Hospitals (HLCMHPP2023069); the Scientific and Technological Innovation Project—Innovation Team of the China Academy of Chinese Medical Sciences (CI2021B012); and the Surplus Fund of the Guang'anmen Hospital Scientific Research Project of the China Academy of Chinese Medical Sciences (2023013).

Data Availability

The ethical approval document specifies that access to the data collected in this project is limited to core members of the project team to safeguard patient privacy. Data may, however, be available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: ZHC, BHJ
Data curation: ZHC, RL, YHJ, SSY, GHZ, ZXY, XYR, JKH
Formal analysis: ZHC
Methodology: ZHC, RL, JKH
Supervision: BHJ
Writing – original draft: ZHC
Writing – review & editing: ZHC, RL, YHJ, JKH, BHJ

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
.Supplementary analyses supporting the development and validation of the adherence prediction model.
[DOCX File, 256 KB - aging_v9i1e82787_app1.docx]

References

1. Tatulian SA. Challenges and hopes for Alzheimer's disease. Drug Discov Today 2022 Apr;27(4):1027-1043. [doi: 10.1016/j.drudis.2022.01.016] [Medline: 35121174]
2. Dong XQ, Li XY, Kong XH, et al. Analysis of clinical application patterns in acupuncture-moxibustion treatment of Alzheimer disease. J Acupunct Tuina Sci 2020 Jun;18(3):238-246. [doi: 10.1007/s11726-020-1183-y]
3. Xia KP, Pang J, Li SL, Zhang M, Li HL, Wang YJ. Effect of electroacupuncture at governor vessel on learning-memory ability and serum level of APP, Aβ1-42 in patients with Alzheimer's disease. Zhongguo Zhen Jiu 2020 Apr 12;40(4):375-378. [doi: 10.13703/j.0255-2930.20190728-0003] [Medline: 32275365]
4. Ouyang Q, Mu YY. Clinical observation of electroacupuncture combined with perphenazine in treating psychiatric symptoms of Alzheimer's disease. Shanghai J Acupunct Moxibustion 2000;06:16-17. [doi: 10.13460/j.issn.1005-0957.2000.06.009]
5. Lin L, Ma X, Wang G, Wang HZ, Wang ZQ, Wang ZW, et al. Chinese guidelines for early prevention of Alzheimer's disease (2024). Chin J Alzheimer Dis Relat Disord 2024;7(3):168-175. [doi: 10.3969/j.issn.2096-5516.2024.03.002]
6. Wang YY, Yu SF, Xue HY, Li Y, Zhao C, Jin YH. Effectiveness and safety of acupuncture for the treatment of Alzheimer's disease: a systematic review and meta-analysis. Front Aging Neurosci 2020;12:98. [doi: 10.3389/fnagi.2020.00098] [Medline: 32435187]
7. Ke C, Shan S, Yu J, Wei X, Pan J, Zhang W. Acupuncture for patients with Alzheimer's disease: an evidence map of randomized controlled trials, systematic reviews, and meta-analysis. J Alzheimers Dis 2024 Dec;102(4):924-942. [doi: 10.1177/13872877241295400] [Medline: 39544007]
8. Olchanski N, Daly AT, Zhu Y, et al. Alzheimer's disease medication use and adherence patterns by race and ethnicity. Alzheimers Dement 2023 Apr;19(4):1184-1193. [doi: 10.1002/alz.12753] [Medline: 35939325]
9. Xiong S, Wu J, Li M, Lu K. PDG38 association between satisfaction with quality of care and ANTI-dementia medication adherence among elderly adults with Alzheimer's disease and related dementias. Value Health 2021 Jun;24:S93. [doi: 10.1016/j.jval.2021.04.486]
10. Wu XW, Yang HB, Yuan R, Long EW, Tong RS. Predictive models of medication non-adherence risks of patients with T2D based on multiple machine learning algorithms. BMJ Open Diabetes Res Care 2020 Mar;8(1):e001055. [doi: 10.1136/bmjdrc-2019-001055] [Medline: 32156739]
11. Koesmahargyo V, Abbas A, Zhang L, et al. Accuracy of machine learning-based prediction of medication adherence in clinical research. Psychiatry Res 2020 Dec;294:113558. [doi: 10.1016/j.psychres.2020.113558] [Medline: 33242836]
12. Nagpal TS, Mottola MF, Barakat R, Prapavessis H. Adherence is a key factor for interpreting the results of exercise interventions. Physiotherapy 2021 Dec;113:8-11. [doi: 10.1016/j.physio.2021.05.010] [Medline: 34555674]
13. Wang XL, Hu HQ, Wen Q, Li N. Considerations and suggestions on influencing factors of compliance in clinical trial of acupuncture and moxibustion: experience in case study of knee osteoarthritis treated with acupuncture. Zhongguo Zhen Jiu 2021 Sep 12;41(9):1045-1048. [doi: 10.13703/j.0255-2930.20200910-k0002] [Medline: 34491656]
14. Duan HL, Du ZJ, Zhang WR, Dong XF. Influencing factors of acupuncture adherence based on literature review. World's Latest Medical Information Digest 2018;18(98):146-148. [doi: 10.19613/j.cnki.1671-3141.2018.98.063]
15. Yu MK, Hu RX, Wen LZ, Li X, Zhao ZY, et al. Characteristics of the methodological studies on patient compliance in clinical trials in China. Chin J EvidBased Med 2019;19(6):708-714. [doi: 10.7507/1672-2531.201812053]
16. Wei J, Liu HL, Wang LC, Wang LP, Liu ZS, Yi Y. Discussion on improving compliance of participants in acupuncture-moxibustion clinical trial. Beijing J Tradit Chin Med 2014;33(2):97-99. [doi: 10.16025/j.1674-1307.2014.02.005]
17. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):263-269. [doi: 10.1016/j.jalz.2011.03.005] [Medline: 21514250]
18. Cui Y, Zhang J, Wang Y, et al. Multivariate predictive model of the therapeutic effects of metoprolol in paediatric vasovagal syncope: a multi-centre study. EBioMedicine 2025 Mar;113:105595. [doi: 10.1016/j.ebiom.2025.105595] [Medline: 39946834]
19. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. 2017 Presented at: 31st Conference on Neural Information Processing Systems (NIPS 2017); Dec 4-9, 2017 URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf [accessed 2025-12-30]
20. David W, Lemeshow S, Sturdivant RX. Applied Logistic Regression: John Wiley & Sons; 2013. [doi: 10.1002/9781118548387]
21. Wang YY, Liu Z, Wu Y, et al. Acupuncture for smoking cessation in Hong Kong: a prospective multicenter observational study. Evid Based Complement Alternat Med 2016;2016:2865831. [doi: 10.1155/2016/2865831] [Medline: 28003848]
22. White A, Cummings M, Barlas P, et al. Defining an adequate dose of acupuncture using a neurophysiological approach--a narrative review of the literature. Acupunct Med 2008 Jun;26(2):111-120. [doi: 10.1136/aim.26.2.111] [Medline: 18591910]

XSL•FO
RenderX

23. Xu G, Lei H, Huang L, et al. The dose-effect association between acupuncture sessions and its effects on major depressive disorder: a meta-regression of randomized controlled trials. J Affect Disord 2022 Aug 1;310:318-327. [doi: 10.1016/j.jad.2022.04.155] [Medline: 35504399]
24. Zhang Y, Yang R, Zhang C, Han L. Dose-effect relationship between the number of acupuncture sessions and efficacy for cervical vertigo: a meta-regression analysis based on randomized controlled trials. Zhongguo Zhen Jiu 2025 Aug 12;45(8):1180-1186. [doi: 10.13703/j.0255-2930.20240711-0002] [Medline: 40825705]
25. Li J, Aulakh N, Culum I, Roberts AC. Adherence to non-pharmacological interventions in Parkinson's disease: a rapid evidence assessment of the literature. J Parkinsons Dis 2024;14(s1):S35-S52. [doi: 10.3233/JPD-230266] [Medline: 38640167]
26. Wang TS, Bai P. Clinical study on 37 cases of periarthritis treated with different acupuncture frequencies. Jiangsu J Tradit Chin Med 2019;51(7):62-65 [FREE Full text] [doi: 10.3969/j.issn.1672-397X.2019.07.021]
27. Li Y. Study on the time-effect relationship of acupuncture between needle retaining time and the interval time of acupuncture. Liaoning J Tradit Chin Med 2013;40(3):538-540. [doi: 10.13192/j.ljtcm.2013.03.160.liy.029]
28. Caregiving in the United States 2020. : Health Resources and Services Administration; 2020 URL: https://www.hrsa.gov/sites/default/files/hrsa/advisory-committees/nursing/reports/report-caregiving-us-2020.pdf [accessed 2025-12-30]
29. Committee on Family Caregiving for Older Adults, Board on Health Care Services, Health and Medicine Division, National Academies of Sciences, Engineering, and Medicine. In: Schulz R, Eden J, editors. Families Caring for an Aging America: National Academy of Sciences; 2016. [doi: 10.17226/23606]
30. Chen S, Cao Z, Nandi A, et al. The global macroeconomic burden of Alzheimer's disease and other dementias: estimates and projections for 152 countries or territories. Lancet Glob Health 2024 Sep;12(9):e1534-e1543. [doi: 10.1016/S2214-109X(24)00264-X] [Medline: 39151988]
31. Xu P, Zhong QL. Research progress on long-term care services for disabled elderly in community-based home settings. Chin J Gerontol 2016;36(12):3076-3078 [FREE Full text] [doi: 10.3969/j.issn.1005-9202.2016.12.116]
32. Ma WJ, Wang YH. Influencing factors of informal home care time. Chin J Gerontol 2018;38(14):3560-3563 [FREE Full text] [doi: 10.3969/j.issn.1005-9202.2018.14.088]
33. Berge LI, Angeles RC, Gedde MH, et al. Burden and care time for dementia caregivers in the LIVE@Home.Path trial. Alzheimers Dement 2025 Mar;21(3):e14622. [doi: 10.1002/alz.14622] [Medline: 40042468]
34. Golics CJ, Basra MKA, Finlay AY, Salek S. The impact of disease on family members: a critical aspect of medical care. J R Soc Med 2013 Oct;106(10):399-407. [doi: 10.1177/0141076812472616] [Medline: 23759884]
35. Feng Y, Wu QH, Lin SB, Lin H, Huang XH, Liu BB. Study on the correlation between caregiver burden, coping style, and acceptance behavior in primary caregivers of Alzheimer's disease patients. Chin Gen Pract Nurs 2022;20(32):4593-4596 [FREE Full text] [doi: 10.12104/j.issn.1674-4748.2022.32.035]
36. Du X, Shen J, Wu YF. Correlation between caregiving experience and fatigue among dementia caregivers in old age. Chin J Gerontol 2015;35(9):2504-2506 [FREE Full text] [doi: 10.3969/j.issn.1005-9202.2015.09.093]
37. Pinquart M, Sörensen S. Correlates of physical health of informal caregivers: a meta-analysis. J Gerontol B Psychol Sci Soc Sci 2007 Mar;62(2):126-137. [doi: 10.1093/geronb/62.2.p126] [Medline: 17379673]
38. Wenborn J, O'Keeffe AG, Mountain G, et al. Community occupational therapy for people with dementia and family carers (COTiD-UK) versus treatment as usual (Valuing Active Life in Dementia [VALID]) study: a single-blind, randomised controlled trial. PLoS Med 2021 Jan;18(1):e1003433. [doi: 10.1371/journal.pmed.1003433] [Medline: 33395437]
39. Lamb SE, Sheehan B, Atherton N, et al. Dementia And Physical Activity (DAPA) trial of moderate to high intensity exercise training for people with dementia: randomised controlled trial. BMJ 2018 May 16;361:k1675. [doi: 10.1136/bmj.k1675] [Medline: 29769247]
40. Lilly MB, Laporte A, Coyte PC. Labor market work and home care's unpaid caregivers: a systematic review of labor force participation rates, predictors of labor market withdrawal, and hours of work. Milbank Q 2007 Dec;85(4):641-690. [doi: 10.1111/j.1468-0009.2007.00504.x] [Medline: 18070333]
41. Liu JH, Wang XL, Li WL, Yang L, Wang JG. Effect of synchronous education for children caregivers on hip function recovery after intertrochanteric fracture surgery in elderly patients. Clin Res 2021;29(7):153-154 [FREE Full text]
42. Liu YY, Zhang L. Application of children caregivers' synchronous education in postoperative rehabilitation of elderly patients undergoing heart valve replacement. Contemp Nurse 2022;29(4):150-152. [doi: 10.19791/j.cnki.1006-6411.2022.10.044]
43. Coley N, Gallini A, Garès V, Gardette V, Andrieu S, ICTUS/DSA group. A longitudinal study of transitions between informal and formal care in Alzheimer disease using multistate models in the European ICTUS cohort. J Am Med Dir Assoc 2015 Dec;16(12):1104. [doi: 10.1016/j.jamda.2015.09.010] [Medline: 26593306]
44. MacNeil-Vroomen JL, Thompson M, Leo-Summers L, Marottoli RA, Tai-Seale M, Allore HG. Health-care use and cost for multimorbid persons with dementia in the National Health and Aging Trends Study. Alzheimers Dement 2020 Sep;16(9):1224-1233. [doi: 10.1002/alz.12094] [Medline: 32729984]
45. Balkrishnan R, Housman TS, Carroll C, Feldman SR, Fleischer AB. Disease severity and associated family impact in childhood atopic dermatitis. Arch Dis Child 2003 May;88(5):423-427. [doi: 10.1136/adc.88.5.423] [Medline: 12716715]

XSL•FO
RenderX

46. Zhang DY, Yang JM, Liu H. Correlation between psychological resilience and caregiving burden among caregivers of patients with senile dementia. Chin Gen Pract Nurs 2018;16(32):3969-3973 [FREE Full text] [doi: 10.12104/j.issn.1674-4748.2018.32.001]

Abbreviations

AD: Alzheimer disease
AUC: area under the receiver operating characteristic curve
CACMS: China Academy of Chinese Medical Sciences
OR: odds ratio
PDC: proportion of days covered
RCT: randomized controlled trial
ROC: receiver operating characteristic
SHAP: Shapley Additive Explanations

Edited by Q Fan; submitted 21.Aug.2025; peer-reviewed by H Wu, WX Li, X Wang; accepted 17.Dec.2025; published 21.Jan.2026.

Please cite as:
Chen ZH, Li R, Jiang YH, He JK, Yan SS, Zong GH, Yi ZX, Ren XY, Jia BH
Predictive Model of Acupuncture Adherence in Alzheimer Disease: Secondary Analysis of Randomized Controlled Trials
JMIR Aging 2026;9:e82787
URL: https://aging.jmir.org/2026/1/e82787
doi:10.2196/82787

© Ze-Hao Chen, Ran Li, Yu-Hang Jiang, Jia-Kai He, Shan-Shan Yan, Guan-Hua Zong, Zong-Xi Yi, Xin-Yu Ren, Bao-Hui Jia. Originally published in JMIR Aging (https://aging.jmir.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Virtual Kitchen Challenge—Version 2: Validation of a Digital Assessment of Everyday Function in Older Adults

Marina Kaplan[1], BA; Moira McKniff[1], MA; Stephanie M Simone[1,2], PhD; Molly B Tassoni[1], MA; Katherine Hackett[1,3]; Sophia Holmqvist[1], MA; Rachel E Mis[1,4], PhD; Kimberly Halberstadter[1], MA; Riya Chaturvedi[1], BS; Melissa Rosahl[1], BA; Giuliana Vallecorsa[1,5], MS; Mijiail D Serruya[6], MD, PhD; Deborah A G Drabick[1], PhD; Takehiko Yamaguchi[7], PhD; Tania Giovannetti[1], PhD

[1]Department of Psychology and Neuroscience, Temple University, Philadelphia, PA, United States

[2]Department of Psychiatry, Neuropsychology Section, Michigan Medicine, Ann Arbor, MI, United States

[3]Division of General Internal Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[4]Department of Neurology, Dell Medical School, University of Texas at Austin, Austin, TX, United States

[5]Department of Psychology, LaSalle University, Philadelphia, PA, United States

[6]Raphael Center for Neruorestoration, Thomas Jefferson University, Philadelphia, PA, United States

[7]Department of Applied Information Engineering, Suwa University of Science, Nagano, Japan

Corresponding Author:
Tania Giovannetti, PhD
Department of Psychology and Neuroscience
Temple University
Weiss Hall
1701 N 13th Street
Philadelphia, PA, 19122
United States
Phone: 1 484 843 1321
Email: tgio@temple.edu

Abstract

Background: Conventional methods of functional assessment include subjective self- or informant report, which may be biased by personal characteristics, cognitive abilities, and lack of standardization (eg, influence of idiosyncratic task demands). Traditional performance-based assessments offer some advantages over self- or informant reports but are time-consuming to administer and score.

Objective: This study aims to evaluate the validity and reliability of the Virtual Kitchen Challenge—Version 2 (VKC-2), an objective, standardized, and highly efficient alternative to current functional assessments for older adults across the spectrum of cognitive aging, from preclinical to mild dementia.

Methods: A total of 236 community-dwelling, diverse older adults completed a comprehensive neuropsychological evaluation to classify cognitive status as healthy, mild cognitive impairment, or mild dementia, after adjustment for demographic variables (age, education, sex, and estimated IQ). Participants completed 2 everyday tasks (breakfast and lunch) in a virtual kitchen (VKC-2) using a touchscreen interface to select objects and sequence steps. Automated scoring reflected completion time and performance efficiency (eg, number of screen interactions, percentage of time spent off-screen, interactions with distractor objects). Participants also completed the VKC-2 tasks using real objects (Real Kitchen). All participants and informants for 219 participants completed questionnaires regarding everyday function. A subsample of participants (n=143) performed the VKC-2 again in a second session, 4-6 weeks after the baseline, for retest analyses. Analyses evaluated construct and convergent validity, as well as retest and internal reliability, of VKC-2 automated scores.

Results: A principal component analysis showed that the primary VKC-2 automated scores captured a single dimension and could be combined into a composite score reflecting task efficiency. Construct validity was supported by analyses of covariance results showing that participants with healthy cognition obtained significantly better VKC-2 scores than participants with cognitive impairment (all *Ps*<.001), even after controlling for demographics and general computer visuomotor dexterity. Convergent validity was supported by significant correlations between VKC-2 scores and performance on the Real Kitchen (r=−0.58 to 0.64, *Ps*<.001), conventional cognitive test scores (r=−0.50 to −0.22, *Ps*<.001), and self- and informant report questionnaires evaluating everyday

XSL•FO
RenderX

function (r=0.25 to 0.43, *Ps*<.001). Intraclass correlation coefficients (ICCs) indicated moderate to excellent retest reliability (ICC=0.70-0.90) for VKC-2 scores after 4-6 weeks. Reliability improved in analyses including only participants who reported no change in cognitive status between time 1 and time 2 (n=123). Spearman-Brown correlations showed acceptable to good internal consistency between the VKC-2 tasks (breakfast and lunch) for all scores (0.77-0.84), supporting the use of total scores.

Conclusions: The VKC-2 is an efficient, valid, and sensitive measure of everyday function for diverse older adults and holds promise to improve the status quo of functional assessment in aging, particularly when informants are unavailable or unreliable.

(JMIR Aging 2026;9:e82092) doi:10.2196/82092

KEYWORDS

everyday function; activities of daily living; assessment; dementia; Alzheimer disease; neuropsychology; cognition; mild cognitive impairment; virtual reality; digital assessment

Introduction

As the US population ages and interventions for Alzheimer disease and Alzheimer disease–related dementias become available [1], highly sensitive, objective, and efficient measures of functional abilities are needed for multiple purposes. Mild functional difficulties are among the strongest predictors of future cognitive decline and dementia [2-5]; thus, accurate measurement of functional ability will improve prognostic prediction and help identify the need for early intervention. Given that functional ability level is often the criterion that distinguishes mild cognitive impairment (MCI) from mild dementia, accurate assessment is critical for diagnostic decision-making [6,7]. According to the Food and Drug Administration, the approval of pharmacological treatments for dementia, even at the very early, presymptomatic stage, is contingent on demonstrating gains on meaningful measures of functioning [8]. Recently approved treatments have relied on composite measures such as the Clinical Dementia Rating-Sum of Boxes and the integrated Alzheimer's Disease Rating Scale, but these measures require specialized training, are not readily deployable in typical clinical settings, and lack sensitivity to the earliest functional changes [9,10]. There exists a critical need for sensitive and efficient functional assessment tools that are clinically meaningful, psychometrically sound, and practically implementable across diverse health care settings [11,12]. We developed a nonimmersive virtual reality (VR) measure, the Virtual Kitchen Challenge—Version 2 (VKC-2), an objective, sensitive, efficient, and theoretically based tool for assessing everyday function in older adults to address the gaps in current functional assessments. Here we report results on VKC-2 validity and reliability in racially diverse, community-dwelling older adults with healthy cognition, MCI, or mild dementia.

Self/informant reports of everyday function, which are easy to administer and score, are the current standard method for functional assessment. When used with reliable, observant, and knowledgeable reporters, they generate useful information about how a person is functioning in everyday life [13-15]. In many circumstances, however, the accuracy of self and informant reports is uncertain. Their subjective nature makes them prone to over- or underreporting due to faulty cognitive abilities, psychological factors (eg, denial, depression, burden), or cultural beliefs [16]. Informant reports are often unavailable, as many older adults do not have a living spouse, nearby family members, or close friends. Even when available and willing, informants may have limited opportunities to observe daily functioning and may lack knowledge, particularly when functional difficulties are mild and may be masked by compensatory behaviors [17,18].

Another limitation of questionnaires is that older adults vary widely in the activities they perform and the contexts in which they perform them. For example, informant-reported difficulties with medication management may be profoundly different for an older adult managing a single prescription while residing in a small, highly organized home with her spouse versus an older adult taking dozens of medications while living alone in a large, cluttered house [19]. However, given identical clinical presentations and cognitive test scores suggesting mild cognitive decline, the latter patient would likely be diagnosed with clinical dementia if she were unable to independently manage her medications. Failure to account for context and task complexity confounds the informant report of everyday function and precludes clear comparisons of functional abilities across individuals.

Further, many questionnaires do not distinguish difficulties due to physical versus cognitive limitations [14], and if they do, it may be difficult for an informant to fully understand the nature of the functional difficulties, particularly because physical and cognitive limitations often co-occur [20-22]. Informant and self-reports also do not offer a detailed characterization of types of functional difficulties arising from different underlying cognitive problems (eg, slowing, disorganized actions vs omission of crucial task steps), which could offer insights into interventions for improving function and reducing the risk of future functional disability [23,24].

Performance-based measures of function address many of the limitations of questionnaires; they are objective, standardize task complexity and context, and allow for detailed analysis of behavior and systematic comparison across individuals. The Naturalistic Action Test (NAT), for example, is a performance-based test of everyday function with strong psychometric properties, normative data, and suggested cut scores for healthy cognition versus MCI versus mild dementia [23,25-35]. Scoring NAT performance for subtle inefficient errors, called micro-errors, has increased the sensitivity of NAT tasks for detecting mild difficulties with everyday tasks [35-38]. Results from performance-based tests, such as the NAT, with added sensitive scoring procedures, have demonstrated that (1)

XSL•FO
RenderX

healthy older adults make more errors and require more time to complete everyday tasks than younger adults [36,37,39-43]; (2) people with MCI make more errors than healthy controls but fewer errors than individuals with dementia [32,35,44-46]; and (3) the ability to accurately and efficiently perform everyday tasks is moderately correlated with performance on cognitive tests [27,28,35,37,47] and informant report of everyday function [27,29,31,35,48]. Together, these findings and others [37,49-52] suggest that standardized performance-based assessment of function is valid and reliable.

Despite their objectivity, validity, and potential for rich characterizations of function, current performance-based tests require extraordinary effort, limiting their implementation and scalability. Scoring, particularly scoring for subtle errors and inefficiencies, is time-intensive and requires video recording, detailed scoring instructions, and trained coders. Although some performance-based tests may be scored quickly as pass/fail without video recording [40,53], such gross measures are less sensitive to mild difficulties (ie, MCI) [54], do not advance our understanding of the nature of functional problems [5,55,56], or still require considerable effort to administer. To streamline administration and scoring, a nonimmersive VR task called the Virtual Kitchen, modeled after the NAT, was developed. The original version of the Virtual Kitchen [57] required a mouse to move objects on a computer screen to complete a coffee-making task. Results showed that people with dementia accomplished fewer steps and made more errors than healthy controls on the Virtual Kitchen. Validity was also supported by significant correlations between Virtual Kitchen scores and performance of real tasks, cognitive tests, and informant reports of functioning [57].

Our team revised the original Virtual Kitchen [57] by implementing the following updates: (1) expanding the coffee task to include a more extensive breakfast; (2) adding a lunch task; (3) updating the graphics; and (4) transitioning from a mouse to a computer touchscreen to make interactions more natural [39]. We also added a brief training task to familiarize participants with the touchscreen interface. Automated scores were expanded to include measures computed based on interactions with the touchscreen to increase sensitivity (ie, number of screen interactions). Preliminary results from the revised task, which we called the Virtual Kitchen Challenge (VKC), demonstrated validity and good internal consistency [39]. The VKC automated scores have been validated against conventional cognitive tests in young adults [52] and against neuroimaging markers of cerebral vascular disease (white matter hyperintensities) in a small sample of community-dwelling older adults [48].

In this paper, we present the psychometric properties of the automated scores from the most recent revision of the Virtual Kitchen, the VKC-2. This version includes enhanced graphics and a more extensive basic familiarization task for practice and to obtain a score of participants' digital visuomotor dexterity that may be used as a control measure. We evaluated construct and convergent validity as well as retest and internal reliability in a large, community-based sample of racially diverse older adults with healthy cognition, MCI, or mild dementia. Construct validity of the VKC-2 automated scores was evaluated in a known-group comparison (healthy cognition vs MCI vs mild dementia). Convergent validity was evaluated with correlations between automated VKC-2 measures and performance on the real versions of the VKC-2 tasks (Real Kitchen), demographically adjusted cognitive test scores, and conventional self/informant questionnaires of everyday function. Retest reliability was evaluated over a period of 4-6 weeks. Internal consistency was evaluated for the 2 VKC-2 tasks (breakfast and lunch).

Methods

Participants

Participants were recruited for an observational, longitudinal psychometric study designed to evaluate the psychometric properties of the VKC-2 (n=217; grant R01AG062503) or for a separate, smaller study on activity tracking (n=20; grant F31AG089944). Procedures for the baseline visit of both studies were the same, designed and conducted in accordance with the Helsinki Declaration, and approved by the Institutional Review Board at Temple University (institutional review board protocols 23116 and 29712). All participants and a knowledgeable informant signed informed consent forms, were compensated for their participation (US $50 for participants per session and US $25 for informants per session), and were assigned study numbers to protect their privacy when storing research records. At the end of the study, participants were also offered a research report with their cognitive test scores, if interested.

All participants were recruited from community outreach events, fliers, and referrals from neurology departments in Philadelphia, Pennsylvania, from September 2020 to June 2025. Inclusion and exclusion criteria were screened by phone, with only minor differences between the 2 studies. In both studies, participants were excluded for the following reasons: lifetime history of severe psychiatric disorder (eg, schizophrenia, bipolar disorder); nervous system infections or disorders (eg, epilepsy, brain tumor); current metabolic or systemic disorders (eg, B_{12} deficiency, renal failure, cancer); current moderate-severe depression; current moderate-severe anxiety symptoms; severe sensory deficits that would preclude visual detection or identification of common everyday objects used in the study or the ability to hear task directions (eg, blindness, total hearing loss); severe motor weakness that would preclude the use of everyday objects (eg, severe deformities or paralysis of both upper extremities); intellectual disability; and not being a fluent English speaker. The inclusion criteria for the larger study required participants to be at least 65 years old and have an available informant who could serve as a study partner. Informants were screened by phone for the following eligibility criteria: 18 years of age or older; fluent English speaker; available and willing to complete study questionnaires in person, by phone, or online; has daily contact with the participant; and reports knowledge of the participant's daily functioning. Inclusion criteria for the second, smaller study required participants to be at least 55 years old and did not require a study informant/partner.

XSL•FO
RenderX

Procedures

At the baseline visit (session 1), participants (N=237) completed informed consent, cognitive testing, the VKC-2, the real version of the VKC-2 tasks (ie, Real Kitchen), and questionnaires regarding demographic information, familiarity with the tasks used in the VKC-2, and their ability to perform activities in everyday life. The order of the Real Kitchen and VKC-2 was counterbalanced across participants to control for order effects. At session 1, informants completed questionnaires in person, online, or at home by mail. After reaching our target sample size (n=140) for retest reliability analyses (June 2024), participants were no longer requested to return for a second session 4-6 weeks after session 1 [58]. A total of 143 participants completed session 2, which included a brief interview (for both the participant and informant) regarding changes in cognition or health status (eg, medication changes, falls, illnesses, hospitalizations) since session 1, as well as repeat administration of the VKC-2 and Real Kitchen.

Measures

Conventional Cognitive Tests

Cognitive tests were administered to characterize the sample, classify participants according to their cognitive status, and evaluate the convergent validity of the VKC-2. The cognitive testing protocol is described in Table 1. The protocol included 2 tests from 4 different cognitive domains to classify participants according to Jak/Bondi actuarial criteria [59,60] and clinical criteria originally proposed by Petersen [6] and McKhann [61]. Normative data from the Calibrated Neuropsychological Normative System [62] were used to enable raw score adjustments for sex, age, education, and IQ estimated by a test of reading/vocabulary. Such demographic adjustments are critical for confirming group membership in a diverse sample of older adults [63,64]. Further details on how tests were used for classifying cognitive abilities are provided in Multimedia Appendix 1.

Table 1. Cognitive tests administered at session 1.

Cognitive domain and test	Score(s)	Reference
Premorbid intellectual functioning (IQ)		
Hopkins Reading Test	Estimated IQ	Schretlen et al [65]
Global cognitive status		
Mini-Mental State Examination	Total correct	Folstein et al [66]
Episodic memory		
Hopkins Verbal Learning Test—Revised[a,b]	Delayed free recall total correct	Brandt and Benedict [67]
Brief Visual Memory Test—Revised[a]	Delayed free recall total correct	Benedict et al [68]
Language		
Category (Animal) Fluency[a,b]	Total correct	Schretlen et al [62]
Boston Naming Test—30 item[a]	Total correct	Goodglass and Kaplan [69]
Executive function		
Trail Making Test—Part B[a,b]	Completion time	Reitan [70]
Digit Span Backward[a]	Longest span	Wechsler [71]
Processing speed		
Salthouse Letter Comparison[a,b]	Total correct	Salthouse [72]
Salthouse Pattern Comparison[a,b]	Total correct	Salthouse [72]
Attention		
Digit Span Forward[a]	Longest span	Wechsler [71]
Trail Making Test A[a]	Completion time	Reitan [70]

[a]*t* scores from these tests were used for healthy versus mild cognitive impairment versus dementia classification.

[b]*t* scores from these tests were averaged to compute the modified Knight-Preclinical Alzheimer Cognitive Composite.

For the analysis of VKC-2 convergent validity, composite scores were computed by averaging demographically adjusted *t* scores from tests within each domain (eg, Episodic Memory, Language). A global cognitive composite was modeled after the Knight-Preclinical Alzheimer Cognitive Composite (modified [m]Knight-PACC) [73], which has been validated as a sensitive measure of early cognitive change due to neurodegenerative disease.

Virtual Kitchen Challenge-Version 2

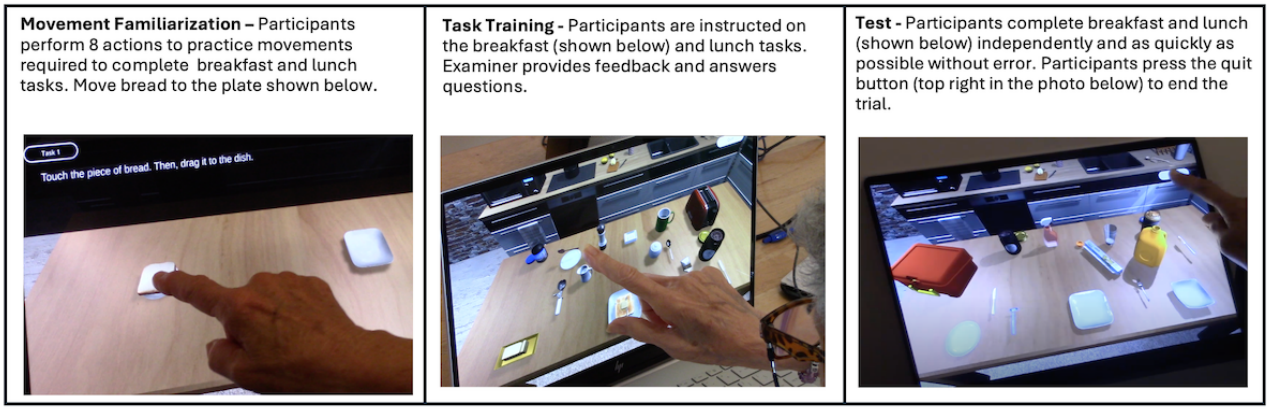
The VKC-2 is a nonimmersive VR test of everyday function that requires participants to complete 2 everyday tasks (breakfast

XSL•FO
RenderX

and lunch) by moving virtual objects using a touchscreen [39,57]. The VKC-2 tasks and objects were modeled after the NAT [25], an extensively studied and theoretically based performance-based test of everyday function that involves completion of familiar everyday tasks using real objects. The VKC-2 breakfast and lunch tasks were designed to be of comparable complexity and difficulty, with each task including 13 target objects and 4 distractor objects. For this study, the VKC-2 was administered on an MSI Creator Z16-A12UET laptop (12th Gen Intel Core i9 Processor) with a 16″ QHD+ (Quad High Definition Plus) (2560 × 1600), 120 Hz, IPS (In-Plane Switching)-level touchscreen display to maximize visibility and portability. Participants were instructed to use the index finger of their dominant hand to move and manipulate objects on the touchscreen.

The VKC-2 included 3 phases: Movement Familiarization, Task Training, and Test. See Figure 1 and the text below for more details.

Figure 1. Photos of participants completing each phase of the VKC-2.



VKC-2 Movement Familiarization

Participants were directed to perform 8 basic touchscreen actions (eg, tap, drag) to complete the following task steps: (1) move bread to dish, (2) stir mug with spoon, (3) pour juice, (4) place thermos in lunch box, (5) spread jelly on bread, (6) wrap cookies in foil, (7) place bread in toaster, and (8) add sugar to mug. Participants first performed all basic touchscreen actions with guidance from the examiner and had the opportunity to ask questions and repeat each action as needed. Next, participants were asked to complete all 8 trials independently as quickly and efficiently as possible. Completion time of the second, independent trial was computed as a measure of basic digital visuomotor dexterity (Digital Dexterity Score).

VKC-2 Task Training

The examiner reviewed the written instructions presented on the computer screen for each task. Participants were asked to point to each of the target objects needed for the task. For example, training for the breakfast task included the direction to “point to all of the objects you will need for the toast” while the examiner named each object out loud (eg, “bread,” “toaster”). Participants were also asked to point to each of the distractor objects and were told that they would not need to touch or use those objects. Participants then proceeded to practice trials, making breakfast and lunch with prompting, cues, and error correction from the examiner. The examiner also answered questions to ensure that participants fully understood each task.

VKC-2 Test

Breakfast and lunch tasks were completed independently without feedback. Instructions regarding the task objectives, which were reviewed during the practice trials, were repeated (eg, “pack a lunch for someone who wants a sandwich, snack, and a drink”). Participants were also instructed to complete test trials as quickly as possible, without making errors, and using clear and precise movements. They were told to touch the quit button at the top right of the screen to end the trial (see Figure 1). Participants were asked to verbally repeat the directions before each task to ensure comprehension; instructions were repeated as often as needed before the participant initiated the task.

VKC-2 Test Automated Scores

Performance on the VKC-2 Test tasks (breakfast and lunch) was scored using data from the touchscreen, as described and validated in our pilot work with the original version of the VKC [39,48,52]:

- Completion time (time) was recorded in seconds from the moment the virtual kitchen screen appeared (after instructions) until the participant pressed the quit button. Results from prior studies of the original VKC indicate that completion time differed significantly between older and younger participants and correlated with completion time on the Real Kitchen, cognitive tests of executive function and episodic memory [39], and neuroimaging markers of cerebrovascular disease [48].
- The number of screen interactions (touches) included the number of discrete instances the participant made contact with the computer touchscreen. Touches were collected as a measure of performance efficiency, with fewer screen interactions reflecting more precise and deliberate actions. Results from the original VKC showed that older adults made significantly more touches than younger adults, with additional touches by older adults including both inefficient correct actions and errors. A higher number of touches was significantly associated with more total errors scored by trained coders who watched video recordings of VKC

XSL•FO
RenderX

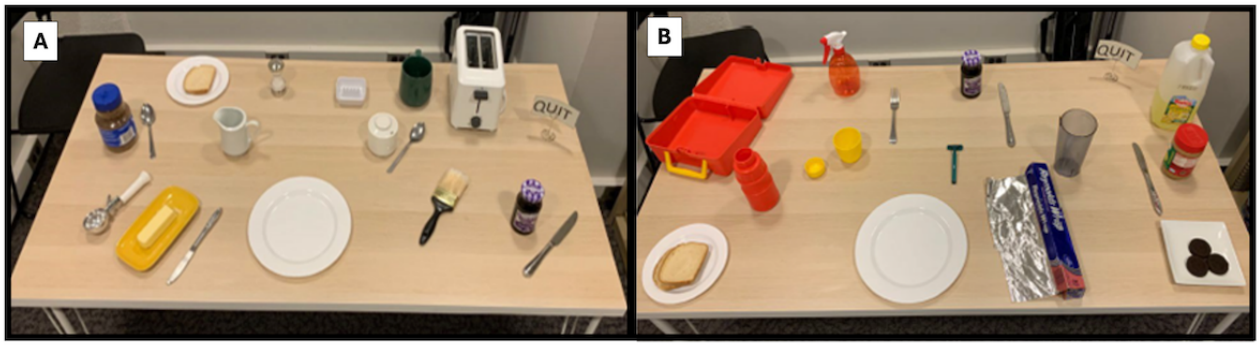
performance. Additionally, screen touches were significantly associated with performance on the Real Kitchen and cognitive tests of executive function and episodic memory [39].

- The percentage of time off-screen (%off-screen) was the percentage of time spent working on the VKC-2 when the participant was not touching the screen. It was computed by subtracting the time spent touching the screen from the completion time, dividing by the completion time, and multiplying by 100. The percentage of time off-screen also reflects performance efficiency. Pilot data from the original VKC [39] indicated that older adults spent a significantly higher percentage of their total time off-screen than younger adults. Correlations between %off-screen and human codes of VKC performance suggested that higher %off-screen times were due to multiple factors, including slower planning, difficulties locating target objects, difficulty resolving competition for object selection, and misreaching toward the computer screen (ie, micro-errors). Higher %off-screen times were significantly associated with more errors on the Real Kitchen, poorer scores on tests of executive function [52] and episodic memory [39], and neuroimaging markers of cerebrovascular disease [48].
- The number of distractor object interactions (distractor interactions) included instances when a distractor object was touched or moved. Our pilot work in a sample of healthy older and younger adults indicated that distractor interactions occurred too infrequently for analysis [39], but they have not been studied in participants with cognitive impairment.

Real Kitchen

The Real Kitchen required participants to complete the breakfast and lunch tasks using real objects placed on a table (Figure 2). Instructions for the Real Kitchen were identical to those for the VKC-2, including the instruction to "press the quit button when finished." In the Real Kitchen, the Quit Button was a piece of paper on the right side of the table labeled "QUIT." Real task objects were similar in appearance (color and shape) to the simulated objects in the VKC-2. Participants repeated the directions before each test trial to ensure comprehension; instructions were repeated as often as needed. Participants were video recorded, and recordings were labeled using a code so that human coders were unaware of participant classification and study session.

Figure 2. Real Kitchen breakfast (A) and lunch (B) tasks.



Real Kitchen performance was scored according to detailed instructions using validated scores and procedures. Real Kitchen scores from a subset of the current sample have been published and show strong interrater reliability, significant differences between participants with healthy cognition versus cognitive impairment, and correlations with cognitive tests and self/informant reports of everyday function [35]. For our current study, the following Real Kitchen scores were used to validate (convergent validity) the VKC-2 automated measures:

- Real Kitchen completion time was recorded in seconds and reliably coded by starting the timer when the first step was initiated and ending when the participant touched the quit button. Prior work shows that participants with greater cognitive impairment demonstrate longer completion times than participants with healthy cognition [35].
- Accomplishment was coded for each completed step and scored from 0 to 13 for the breakfast task and 0 to 20 for the lunch task. A total accomplishment score was computed (0-33), with higher scores reflecting a greater number of task steps accomplished.
- Total errors were coded according to a taxonomy studied in a range of clinical populations [25,74], showing validity and strong interrater reliability in people with stroke [75,76], dementia [27,28,30,47], MCI [26,32,33], and healthy controls [37,38,49], as well as a subset of participants from this sample [35]. The error taxonomy (see Multimedia Appendix 1) includes overt errors (eg, performing task steps in the wrong sequence) and micro-errors (eg, reaching toward a distractor object). In studies of participants with dementia, total overt errors correlate with cognitive tests and informant reports of function. The micro-error category was added to improve detection of subtle, inefficient behaviors in healthy and MCI participants [35,37,38,49]. As overt errors occur with relatively low frequency, they were combined with micro-errors to compute a total error score [35].
- Motor errors were tracked separately from total errors. Motor errors involved instances in which a corrective action was performed with motor or spatial imprecision (eg, spilling coffee grounds, dropping a knife).

XSL•FO
RenderX

Participant Questionnaires

Participants completed a demographic form assessing age, sex, race, ethnicity, income, and education level, as well as the following questionnaires.

The Past Experience Scale [45,77] assessed familiarity with the breakfast and lunch subtasks that comprise the VKC/Real Kitchen. The scale included 4 items (toast, coffee, sandwich, and thermos), each rated from 0 (not at all familiar) to 4 (very familiar). The total familiarity score ranged from 0 to 16, with higher scores reflecting greater familiarity. Participants also rated the frequency with which they had completed each subtask in their day-to-day life over the past 5-10 years, using a scale from 0 (never) to 4 (just about every day), with total scores ranging from 0 (never performed any of the tasks) to 16 (performed each task just about every day).

Functional Activity Questionnaire (FAQ) [14] instructions were modified to reflect only difficulties due to cognitive problems (not physical problems, fatigue, etc) for 10 activities (eg, preparing a balanced meal). Each activity is rated on a scale from 0 (performs normally) to 3 (dependent). Total FAQ scores range from 0 to 30, with higher scores reflecting greater dependence on others in everyday tasks due to cognitive difficulties.

The 12-item Everyday Cognition Scale (ECog-12) [13,78] measures decline over the past 10 years in 12 everyday cognitive abilities (eg, remembering where you have placed objects) on a scale from 1 (better or no change) to 4 (much worse all the time). Total scores reflect an average across all completed items and range from 1 to 4, with higher scores indicating greater decline in everyday cognition.

The Instrumental Activities of Daily Living—Compensation (IADL-C) [15] scale measures the need for assistance and compensatory strategies when performing 27 daily activities (eg, preparing one's own meals). Each activity is rated on a scale from 1 (independent, no aid) to 8 (not able to complete the activity anymore). The total score is the sum of all item responses, with a possible range from 27 (completely independent, no aid needed for any tasks) to 216 (no longer able to perform any task).

Informant Questionnaires

Informants completed questionnaires regarding their demographic information (eg, age, education), their relationship with the participant (eg, cohabitation, years known, hours in contact with the participant), and the participants' everyday function, including the ECog-12 [13,78], FAQ [14], and IADL-C [15]. Instructions and scoring for each questionnaire were the same as those for the participant versions described above.

Analysis Plan

Preliminary Analysis

Analyses were conducted using SPSS version 29.0 (IBM Corp) [79]. VKC-2 automated scores were examined for outliers and Winsorized at the first and ninety-ninth percentiles. The VKC-2 distractor interaction score was dichotomized because a few participants interacted with distractor objects (0=no interactions; 1=at least one interaction with a distractor object during completion of the VKC-2). Relations among VKC-2 scores were evaluated using bivariate correlations. Additionally, a principal component analysis (PCA) was conducted, including the 3 primary VKC-2 variables (time, touches, and %off-screen), to determine whether the dimensional VKC-2 automated scores could be combined into a single composite score. The Digital Dexterity score was not included in the PCA because it is derived from a separate condition intended to be used as a control for basic visuomotor skills. The distractor interaction score was not included because dichotomous variables are not appropriate for PCA. The suitability of the data for PCA was evaluated using the Kaiser-Meyer-Olkin measure of sampling adequacy and the Bartlett test of sphericity.

Construct Validity

VKC-2 automated scores were compared across groups known to differ in functional ability level: healthy cognition, MCI, and mild dementia. As the size of the dementia subgroup was relatively small (n=16), statistical analyses focused on differences between participants with healthy cognition and those with MCI. Participants with dementia were included for descriptive comparisons. One-way analyses of covariance (ANCOVA) were used to test group differences for each VKC-2 automated score (digital dexterity, time, touches, %off-screen, and VKC-2 composite) after controlling for demographics. Group differences were also evaluated in ANCOVA models that controlled for the digital dexterity score to determine whether significant group differences were explained by differences in basic visuomotor or computer abilities. Group differences on the dichotomized VKC-2 distractor interaction score were evaluated using chi-square tests. Significant between-group differences with at least small effect sizes (ie, partial η^2>.01; phi [ϕ] coefficient>.30) were interpreted as supporting the construct validity of the VKC-2 automated scores.

Receiver operating characteristic analyses comparing participant groups (healthy cognition vs impaired cognition [MCI + dementia]; healthy cognition vs MCI) were performed to identify cutoff values for each of the VKC-2 automated scores. Youden indices were used to identify cutoff scores that optimized sensitivity and specificity [80].

Convergent Validity

Correlations between the VKC-2 automated measures and the ability to perform tasks with real objects (Real Kitchen), demographically adjusted cognitive test scores of overall cognition and specific cognitive abilities, and self/informant reports of everyday functioning were performed to evaluate convergent validity. Pearson correlation coefficients were computed using the full sample. Spearman rank-order correlations were also performed and are included in Tables S3-S5 in Multimedia Appendix 1. Significant and moderate-level relationships were interpreted as supporting the convergent validity of the VKC-2 automated scores.

Reliability

Retest reliability was assessed using intraclass correlation coefficients (ICCs), calculated with a 2-way mixed-effects model based on absolute agreement and average measures [81].

XSL•FO
RenderX

ICC values range from 0 to 1, with values above 0.75 generally indicating good reliability and values above 0.90 considered excellent [82]. 95% CIs were computed for each ICC, and significance was determined using *F* tests. Retest reliability for the distractor interaction score (dichotomous variable) was examined using Cohen κ [83]. Retest reliability was evaluated for the full sample who completed session 2 (n=143) and for a subsample that reported no change in cognitive abilities since session 1 (123/143). Internal consistency between the 2 VKC-2 tasks (breakfast and lunch) was tested using the Spearman-Brown formula (*r*), with coefficients>0.70 interpreted as evidence of strong internal consistency [84].

Results

Participant Characteristics

A total of 237 participants were recruited from June 2021 to June 2025 for studies on everyday function. One participant with mild dementia refused to complete the study tasks; thus, the final analytic sample included 236 participants, of whom 172 were classified as having healthy cognition, 48 as having MCI, and 16 as having mild dementia. On average, participants were 72 years old and had completed 15 years of education; of the 236 participants, 156 (66.1%) were women, and nearly equal numbers identified as Black (n=106, 44.9%) and White (n=113, 47.9%). Demographic characteristics of the groups are reported in Table 2. The groups differed in age and education, but post hoc comparisons did not reach statistical significance (*P*>.051 for all). There were no group differences in estimated IQ or in the distributions of sex, Black/African American versus White race, or ethnicity.

Table 2. Demographic and descriptive characteristics by group.

Variable	Healthy (n=172)	Mild cognitive impairment (n=48)	Mild dementia (n=16)	*F* test (*df*) or chi-square (*df*)	*P* value
Age, mean (SD); range	71.95 (6.56); 58-94	74.54 (7.27); 61-98	74.50 (8.70); 55-91	3.30 (2, 235)	.04
Education (years), mean (SD); range	16.06 (2.51); 10-20	15.40 (3.25); 10-20	14.06 (3.04); 10-20	4.64 (2, 235)	.01
Estimated IQ, mean (SD); range	112.44 (11.73); 87-139	112.06 (13.30); 88-138	108.25 (13.19); 88-139	0.87 (2, 235)	.42
Sex: women, n (%)	114 (66.3)	33 (68.8)	9 (56.3)	0.85 (2)	.67
Race				3.53 (2)	.17
Black	72 (41.9)	26 (54.2)	8 (50.0)		
White	89 (51.7)	17 (35.4)	7 (43.8)		
Asian	5 (2.9)	3 (6.3)	1 (6.3)		
Pacific Islander/Hawaiian	2 (1.2)	0 (0)	0 (0)		
American Indian	0 (0)	1 (2.1)	0 (0)		
Multiracial	3 (1.7)	1 (2.1)	0 (0)		
Not reported	1 (0.6)	0 (0)	0 (0)		
Latino/Hispanic Ethnicity	2 (1.2)	0 (0)	0 (0)	0.75 (2)	.69
Past Experience Scale					
Familiarity rating, mean (SD); range	14.20 (2.44); 6-16	13.45 (3.27); 3-16	10.80 (4.44); 4-16	10.69 (2, 235)	<.001
Frequency rating, mean (SD); range	7.87 (2.82); 0-16	7.89 (3.27); 2-13	8.27 (1.62); 5-11	1.11 (2, 235)	.87

Results from the Past Experience Scale showed that task familiarity ratings were generally high, indicating that, on average, the breakfast and lunch tasks were “pretty” to “very” familiar. The groups differed on the familiarity rating, with post hoc tests indicating that the dementia group reported significantly lower task familiarity than the healthy cognition group (*P*<.001) and the MCI group (*P*=.005); however, the healthy cognition group and the MCI group did not differ (*P*=.32). According to the frequency ratings, participants reported that, on average, they had performed the VKC-2 tasks about once per month over the past 5-10 years. Frequency ratings did not differ across groups.

Demographic characteristics of participants who returned for session 2 and were included in the retest reliability analysis are reported in Table S1 in Multimedia Appendix 1. Compared with participants who did not return, the returning participants had completed significantly more years of education, obtained higher estimated IQ scores, and included a greater proportion of White participants.

XSL•FO
RenderX

Informant Characteristics

A total of 219 informants participated in the study. On average, informants were 63.97 years old (SD 13.99 years; range 20-90 years) and had completed 15.73 years of education (SD 2.43 years; range 10-21 years). Informants included spouses (95/219, 43.4%), children (60/219, 27.4%), friends (41/219, 18.7%), and other family members (23/219, 10.5%).

Correlations Among VKC-2 Scores and Principal Component Analysis

Average VKC-2 scores and their bivariate correlations indicate significant, moderate associations among all scores (Table 3). The relationship between VKC-2 time and touches was particularly strong, reflecting nearly overlapping scores, with more touches associated with longer completion times.

Table 3. VKC-2[a] scores and correlation coefficients in the full sample (N=236).

VKC-2 score	Digital dexterity	Time	Touches	%Off-screen
Time	0.67[b]	N/A[c]	N/A	N/A
Touches	0.51[b]	0.83[b]	N/A	N/A
%Off-screen	0.49[b]	0.42[b]	0.34[b]	N/A
Distractor interactions	0.34[b]	0.39[b]	0.42[b]	0.26[b]
Mean	86.76	197.47	67.82	0.48
SD	27.47	111.57	50.59	0.09

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]*P*<.001 (2-tailed). A total of 21 (8.9%) participants interacted with distractor objects; the mean and SD for the distractor interactions score are not reported because it was dichotomized.

[c]N/A: not applicable.

According to the Kaiser-Meyer-Olkin measure of sampling adequacy (0.579) and the significant Bartlett test of sphericity (χ^2_3=311.45, *P*<.001), there was a modest but acceptable level of shared variance among variables and a suitable correlation matrix for factor analysis. PCA results showed that only 1 component was extracted (eigenvalue=2.09), accounting for 69.74% of the total variance. All variables loaded positively on this component (time=0.930, touches=0.905, and %off-screen=0.639), suggesting a single underlying factor representing a common dimension. Thus, a VKC-2 composite score was computed by averaging sample-based *z* scores for time, touches, and %off-screen, with higher scores reflecting worse (ie, more inefficient) performance.

Construct Validity

The construct validity of the VKC-2 automated scores was evaluated by assessing differences among groups known to differ in functional abilities: healthy, MCI, and mild dementia. As shown in Figure 3, average scores on each VKC-2 measure were consistently worse for the dementia group. The same pattern was observed in the VKC-2 composite score (healthy: mean –0.22, SD 0.49; MCI: mean 0.43, SD 1.07; and dementia: mean 1.27, SD 1.36). Statistical analyses focused on differences between the healthy and MCI groups due to the relatively small number of participants with dementia. ANCOVA results comparing healthy versus participants with MCI are reported in Tables 4 and 5 and showed significant group differences (*P*<.001) in all measures after controlling for age. After controlling for the digital dexterity score and age (see Tables 4 and 5), the difference in the time score was no longer significant (*P*=.06), suggesting that the difference in completion time could be explained by low-level visuomotor skill differences between the MCI and healthy groups. By contrast, after controlling for digital dexterity and age, the differences in touches (*P*=.004), %off-screen (*P*=.01), and the VKC-2 (*P*<.001) composite score remained statistically significant, indicating that these between-group differences could not be explained by basic visuomotor skills. Thus, aside from time, the VKC-2 scores—particularly the composite score, which showed the strongest effect size after controlling for digital dexterity and age—likely reflect more than simple visuomotor abilities and capture the cognitive processes required to perform everyday tasks (ie, goal maintenance and control over task goals for the efficient execution of multistep everyday tasks).

XSL•FO
RenderX

Figure 3. Unadjusted VKC-2 mean scores by group.

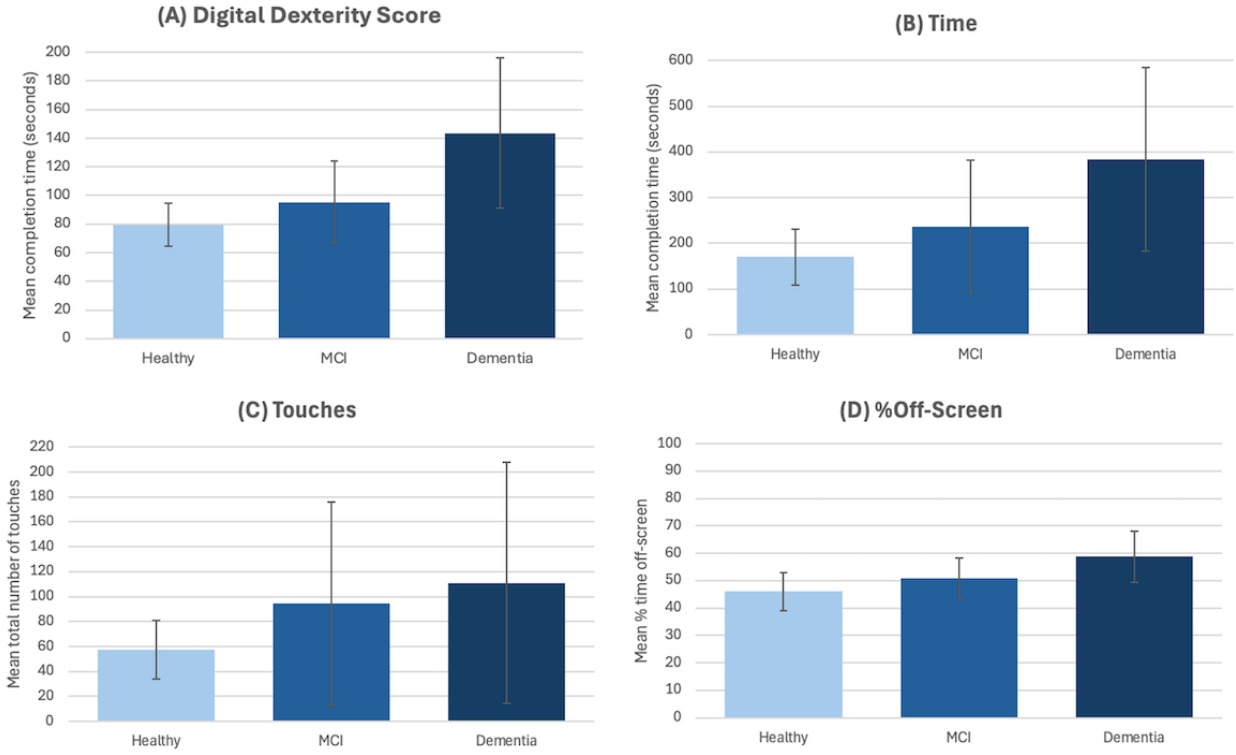


Table 4. Analysis of covariance results comparing participants with healthy cognition (n=172) versus those with mild cognitive impairment (n=48) on all VKC-2[a] automated scores: controlling for age.[b]

VKC-2 score	*F* test (*df*)	*P* value	η^2 (partial eta^2)	Effect size
Digital dexterity	20.68 (2, 219)	<.001	0.087	Medium to large
Time	17.54 (2, 219)	<.001	0.075	Medium
Touches	23.35 (2, 219)	<.001	0.097	Large
%Off-screen	12.71 (2, 219)	<.001	0.055	Medium
VKC-2 composite	29.70 (2, 219)	<.001	0.120	Medium to large

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]Effect sizes (η^2) are interpreted as follows: small=0.01, medium=0.06, and large=0.14.

Table 5. Analysis of covariance results comparing participants with healthy cognition (n=172) versus those with mild cognitive impairment (n=48) on all VKC-2[a] automated scores: controlling for age and digital dexterity.[b]

VKC-2 score	*F* test (*df*)	*P* value	η^2 (partial eta^2)	Effect size
Time	3.47 (3, 219)	.06	0.016	Small
Touches	8.60 (3, 219)	.004	0.038	Small to medium
%Off-screen	6.42 (3, 219)	.01	0.029	Small to medium
VKC-2 composite	11.68 (3, 219)	<.001	0.051	Medium

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]Effect sizes (η^2) are interpreted as follows: small=0.01, medium=0.06, and large=0.14.

The distributions of the distractor interaction score across the 3 groups (not reported in Figure 3) indicated a higher percentage of participants interacting with distractors in the groups with cognitive impairment (dementia: 5/16, 31%; MCI: 8/48, 17%; and healthy: 5/114, 4.3%). The difference in distractor interactions between the MCI and healthy groups was statistically significant (χ^2_1=8.03, *P*=.005; ϕ=0.191, small-to-medium effect size).

Classification Analyses

Classification analyses for distinguishing participants with healthy cognition from those with cognitive impairment (MCI

XSL•FO
RenderX

+ mild dementia combined) are reported in Table 6. All predictors showed statistically significant areas under the curve (AUCs; *P*<.001 for all), indicating that they were better than chance at predicting impaired group status. Time and the VKC-2 composite score were the strongest predictors, as indicated by their high AUCs and sensitivity. Time demonstrated particularly high sensitivity, making it useful for maximizing the identification of participants with impairment for early detection. By contrast, the %off-screen score showed the highest specificity, suggesting it may be more useful for ruling out individuals with healthy cognition during diagnostic confirmation. As expected, scores that increase sensitivity reduce specificity, reflecting the inherent trade-off between identifying true positives and minimizing false positives. Analyses distinguishing participants with healthy cognition versus those with MCI demonstrated similar AUCs (0.68-0.74), cutoff scores, and patterns of sensitivity and specificity, and are reported in Table S2 in Multimedia Appendix 1.

Table 6. Area under the curve values, optimal cutoffs, and specificity/sensitivity for predicting cognitive impairment from VKC-2[a] scores (N=236).

VKC-2 score	AUC[b]	95% CI	SE	*P* value	Optimal cutoff	Sensitivity	Specificity
Digital dexterity	0.73	0.65-0.81	0.41	<.001	87.12	0.67	0.77
Time	0.75	0.68-0.82	0.04	<.001	163.47	0.82	0.58
Touches	0.70	0.62-0.78	0.04	<.001	65.5	0.59	0.78
%Off-screen	0.71	0.64-0.79	0.04	<.001	0.53	0.48	0.87
VKC-2 composite	0.76	0.70-0.84	0.04	<.001	–0.041	0.69	0.72

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]AUC: area under the curve.

Convergent Validity Against Real Kitchen Scores

Bivariate correlations between VKC-2 scores and Real Kitchen scores are reported in Table 7. Correlations with Real Kitchen completion time, accomplishment, and total errors were consistently significant (*P*<.001 for all) and moderate to strong, supporting the convergent validity of the VKC-2 scores against the real versions of the VKC-2 tasks. Relations between VKC-2 measures and motor errors on the Real Kitchen were relatively weaker and not consistently significant (*P* values ranged from <.001 to .08), suggesting that VKC-2 scores correspond more strongly with the cognitive aspects of Real Kitchen performance rather than visuomotor errors made with the real tasks (Table 7). Spearman rank-order correlations showed the same pattern of results and are reported in Table S3 in Multimedia Appendix 1.

Table 7. Correlation coefficients (and P values) between VKC-2[a] scores and Real Kitchen scores (n=201).

VKC-2 score	Completion time	Accomplishment score	Total errors	Motor errors
Digital dexterity	0.59[b]	–0.58[b]	0.50[b]	0.13 (*P*=.08)
Time	0.58[b]	–0.53[b]	0.64[b]	0.22 (*P*=.004)
Touches	0.38[b]	–0.26[b]	0.53[b]	0.30[b]
%Off-screen	0.44[b]	–0.44[b]	0.40[b]	0.15 (*P*=.057)
VKC-2 composite	0.56[b]	–0.48[b]	0.62[b]	0.27[b]
Mean	244.35	32.09	7.42	2.37
SD	93.73	2.35	5.88	2.49

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]*P*<.001 (2-tailed).

Convergent Validity Against Conventional Cognitive Tests

Bivariate correlations between VKC-2 scores and demographically adjusted cognitive test scores are reported in Table 8. The coefficients were statistically significant (*P* values ranged from <.001 to .03), indicating that participants with higher cognitive test scores completed the VKC-2 tasks more quickly and efficiently, supporting the convergent validity of the VKC-2 scores. Spearman rank-order correlations showed the same pattern of results and are reported in Table S4 in Multimedia Appendix 1.

XSL•FO
RenderX

Table 8. Correlation coefficients (and P values) between VKC-2[a] scores and cognitive test scores (N=236).

VKC-2 score	Global Cognition mKnight-PACC[b]	Executive function composite	Episodic memory composite	Processing speed composite	Language composite
Digital dexterity	–0.50[c]	–0.29[c]	–0.42[c]	–0.41[c]	–0.36[c]
Time	–0.42[c]	–0.27[c]	–0.39[c]	–0.30[c]	–0.34[c]
Touches	–0.23[c]	–0.14 (*P*=.03)	–0.29[c]	–0.11 (*P*=.10)	–0.22[c]
%Off-screen	–0.38[c]	–0.27[c]	–0.40[c]	–0.26[c]	–0.27[c]
VKC-2 composite	–0.42[c]	–0.27[c]	–0.42[c]	–0.28[c]	–0.34[c]
Mean	50.50	49.15	45.04	53.08	47.97
SD	7.66	8.77	10.04	9.75	9.55

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]mKnight-PACC: modified Knight-Preclinical Alzheimer Cognitive Composite.

[c]*P*<.001 (2-tailed).

Convergent Validity Against Self/Informant Questionnaires of Everyday Function

Table 9 shows the relationships between VKC-2 scores and questionnaires assessing everyday function completed by participants and informants. Results from participant questionnaires indicated that the associations between VKC-2 scores and the IADL-C and FAQ, which assess current functional abilities, were statistically significant (*P* values ranged from <.001 to .02) and in the expected direction. That is, participants who reported greater current functional difficulties (IADL-C and FAQ) also performed the VKC-2 tasks less quickly and efficiently. The relationship between VKC-2 scores and participants' reports of functional decline (ECog-12) was not significant (*P* values ranged from .21 to .55). By contrast, informant reports of both current functional difficulties (IADL-C and FAQ) and functional decline (ECog-12) were significantly associated with lower VKC-2 scores. Overall, correlations between the VKC-2 and participant/informant questionnaires support the validity of the VKC-2 and are comparable to or stronger than the relationships reported between conventional performance-based tests and questionnaires in the literature [43]. Spearman rank-order correlations showed a similar pattern of results and are reported in Table S5 in Multimedia Appendix 1.

Table 9. Correlation coefficients (and P values) between VKC-2[a] scores and questionnaires.

VKC-2 score	Participant questionnaires (n=236)			Informant questionnaires (n=194)		
	IADL-C[b]	FAQ[c]	ECog-12[d]	IADL-C	FAQ	ECog-12
Digital dexterity	0.26[e]	0.32[e]	0.08 (*P*=.26)	0.43[e]	0.32[e]	0.34[e]
Time	0.28[e]	0.28[e]	0.06 (*P*=.38)	0.41[e]	0.32[e]	0.26[e]
Touches	0.15 (*P*=.03)	0.20 (*P*=.003)	0.04 (*P*=.55)	0.20 (*P*=.005)	0.10 (*P*=.17)	0.07 (*P*=.35)
%Off-screen	0.26[e]	0.25[e]	0.09 (*P*=.21)	0.35[e]	0.29[e]	0.31[e]
VKC-2 composite	0.27[e]	0.30[e]	0.07 (*P*=.30)	0.37[e]	0.28[e]	0.24[e]
Mean	44.76	2.12	1.56	46.75	2.92	1.40
SD	21.41	3.85	.97	30.10	5.98	.54

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]IADL-C: Instrumental Activities of Daily Living—Compensation.

[c]FAQ: Functional Activity Questionnaire.

[d]ECog-12: 12-item Everyday Cognition Scale.

[e]*P*<.001 (2-tailed).

Retest Reliability

ICCs are reported in Table 10 and indicate moderate to excellent reliability for the VKC-2 automated scores. Cohen κ, used to assess agreement between distractor interaction scores at time 1 and time 2, showed only fair agreement (κ=0.27, *P*<.001), indicating limited but statistically significant consistency over time. When ICCs were rerun, including only participants who reported no change in their cognitive status from session 1 (123/143, 86%), results yielded comparable or slightly improved coefficients relative to the full sample (see Table S6 in Multimedia Appendix 1).

XSL•FO
RenderX

Table 10. Intraclass correlation coefficients for VKC-2[a] scores over time (n=143).

VKC-2 score	Intraclass correlation coefficient[b] (average measures)	95% CI	*F* test (*df*)	*P* value
Digital dexterity	0.844	0.766-0.893	6.965 (142, 142)	<.001
Time	0.812	0.736-0.865	5.469 (142, 142)	<.001
Touches	0.849	0.783-0.893	6.943 (140, 140)	<.001
%Off-screen	0.703	0.523-0.806	3.837 (140, 140)	<.001
VKC-2 composite	0.899	0.860-0.923	9.965 (140, 140)	<.001

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]Type A using an absolute agreement definition.

Internal Reliability

Internal consistency between the VKC-2 breakfast and lunch tasks at time 1 was evaluated using Spearman-Brown coefficients in the full sample (N=236). Results indicated acceptable to good internal consistency for all scores (time: 0.81; touches: 0.81; %off-screen: 0.77; and VKC-2 composite score: 0.84).

Discussion

Results of this study support the validity and reliability of the VKC-2 automated scores as measures of everyday function in older adults. As predicted, VKC-2 scores differed significantly between groups known to vary in functional ability (healthy vs MCI vs mild dementia), supporting the construct validity of the VKC-2. Convergent validity was further supported by significant correlations between VKC-2 scores and performance on the real versions of the VKC-2 tasks (Real Kitchen), conventional cognitive test scores, and self/informant questionnaires assessing everyday functioning. Retest reliability analyses showed fair to excellent reliability for the VKC-2 automated scores over 4-6 weeks. Internal consistency between the 2 VKC-2 tasks (breakfast and lunch) was also good. Additionally, participants reported that the tasks included in the VKC-2 were highly familiar (Past Experience Questionnaire). These findings suggest that the VKC-2 automated scores hold strong potential for addressing critical gaps in functional assessment across multiple contexts, including screening older adults at risk for decline in meaningful everyday activities in primary care and serving as a functional end point in clinical trials of Alzheimer disease/Alzheimer disease–related disorder treatments.

To our knowledge, this is the first study to demonstrate significant differences between older adults with healthy cognition and those with MCI on the VKC-2 automated scores. Group differences in all scores except time persisted even after controlling for the digital dexterity score, a novel feature of the updated VKC-2. Thus, differences between MCI and healthy participants on the VKC-2 cannot be attributed solely to differences in digital visuomotor skills or touchscreen accuracy, but rather reflect the additional cognitive demands required to perform everyday tasks accurately and efficiently (eg, accurate object selection, sequencing of task steps, performance monitoring [74,85,86]). This conclusion is further supported by significant correlations with Real Kitchen scores (see also [57,87]) and by the fact that differences between participants with MCI and healthy participants on the VKC-2 mirror those observed on performance-based tasks with real objects in previous studies [26,32,33,35,45]. Significant associations with cognitive tests of episodic memory and language, which do not primarily measure motor skills or processing speed, provide additional evidence that the automated VKC-2 scores reflect cognitive abilities. Collectively, these results strongly support the construct validity of the VKC-2, offering a novel approach to identify everyday task difficulties without the need for video recording or trained coders—a major advantage over traditional performance-based tests—providing a highly efficient, scalable, and sensitive measure of everyday functioning.

It is important to acknowledge that some VKC-2 scores reflect visuomotor skills more than others. For example, the completion time (time) score did not remain significantly different between participants with MCI and those with healthy cognition after controlling for the digital dexterity score. This should not be viewed as a limitation, as mild upper motor dexterity difficulties contribute to functional impairments in people with MCI [20], and mild upper and lower limb difficulties are significantly associated with cognitive challenges in older adults without MCI [21,22,88]. Indeed, the VKC-2 digital dexterity score, as well as VKC-2 measures of efficiency, were associated with a measure of global cognitive abilities (mKnight-PACC) that is sensitive to preclinical Alzheimer disease. Thus, mild motor difficulties may serve as important early indicators of Alzheimer disease/Alzheimer disease–related disorder risk that could be missed by conventional cognitive tests. Additional studies, including longitudinal follow-up, are needed to identify the optimal combination of VKC-2 scores to maximize early detection of functional difficulties and risk.

Correlation analyses with conventional self- and informant-report questionnaires of everyday function provided additional support for the validity of the VKC-2 automated measures as indicators of processes that influence real-world functioning. The strength and pattern of correlations between VKC-2 scores and conventional questionnaires were similar to those reported for validated performance-based tests and questionnaires of everyday function in the existing literature [89,90]. Correlations were stronger and more consistent with informant reports than with self-reports, particularly for the questionnaire assessing cognitive/functional decline (ECog-12 [13]); this pattern has been reported in previous studies [56]

XSL•FO
RenderX

and aligns with our conceptualization of the constructs measured by performance-based tests versus questionnaires. We view the VKC-2, like other performance-based tests, as a measure of everyday functional capacity, making it well-suited for between-participant comparisons, staging, and tracking change over time. Questionnaires assess real-world functioning, which is highly unconstrained, with task demands, motivation, economic resources, social support, and other factors varying widely. Thus, in clinical practice, the VKC-2 could be used alongside questionnaires to provide a comprehensive evaluation of everyday function across contexts.

Significant associations between VKC-2 scores and conventional questionnaires of everyday function support the clinical relevance of the VKC-2 measures. Differences between participants with healthy cognition and those with MCI were small to moderate, with absolute differences amounting to only a few seconds on some scores. Such differences may reflect subtle processing difficulties that lead to inefficiency and increased cognitive load, which could accumulate over the course of a day. We acknowledge, however, that direct evidence that the mild cognitive difficulties captured by the VKC-2 translate to meaningful impacts on everyday tasks is currently lacking. Further validation using ecological momentary assessment or digital phenotyping via wearables (or both) would provide more direct evidence of the VKC-2 as a measure of real-world everyday function.

In addition to validity analyses, the reliability of the VKC-2 represents an important novel contribution of this study. To our knowledge, reliability has not been examined for any prior version of the Virtual Kitchen. Retest reliability estimates (ICC) showed that the automated VKC-2 scores—except for distractor interactions, which occurred very infrequently—were highly stable over a 4-6-week period. ICCs were even stronger when participants who reported notable changes in cognitive abilities were excluded. Strong retest reliability is critical for using the VKC-2 to evaluate meaningful change over time and for clinical trial applications. The VKC-2 tasks (breakfast and lunch) also demonstrated strong internal consistency, supporting the coherence of the combined total VKC-2 scores. Furthermore, correlations and PCA indicate that VKC-2 automated scores reflect a single underlying dimension and can be combined into a composite score representing task efficiency.

Several strengths of the study are worth noting. First, the sample size and inclusion of a substantial proportion of participants (106/236, 44.9%) identifying as Black or African American addresses a critical gap in cognitive assessment research and enhances the generalizability of our findings across the US population. Second, the VKC-2's portability, automated scoring, and standardized administration protocol offer clear advantages over current functional measures and existing regulatory-approved outcome measures for clinical trials, which often require specialized training, lengthy administration times, and access to informants. The VKC-2 does not require Wi-Fi and can be administered on any commercially available, budget-friendly touchscreen computer. Finally, the efficiency of the VKC-2 compared with conventional cognitive test batteries makes it particularly suitable for busy clinical settings where comprehensive neuropsychological assessments are impractical.

Study limitations also warrant consideration. First, although our sample included substantial racial diversity, the predominance of highly educated participants (mean 15.7 years of education) may limit generalizability to populations with lower educational attainment. Additionally, the sample was majority female (156/236, 66.1%), and participants from racial groups other than Black/African American or White or diverse ethnicities were underrepresented. Second, the community-based sample primarily included older adults with healthy cognition, with only 64 of 236 (27.1%) participants meeting criteria for cognitive impairment. The imbalance in subgroup sizes between participants with healthy versus those with impaired cognition limited statistical power for between-group comparisons and AUC/classification analyses. Therefore, additional studies are needed to replicate these findings in samples with larger groups of individuals with MCI or mild dementia. Third, although the virtual task environment is ecologically valid, it may not capture all real-world functional demands, such as physical fatigue, environmental distractions, or competing task requirements. Fourth, the cross-sectional design limits conclusions about the VKC-2's ability to detect meaningful change over time or predict clinical outcomes (predictive validity). Finally, direct validation against regulatory outcome measures is necessary before the VKC-2 can be considered an alternative end point in clinical trials.

As noted, future research on the VKC-2 should include longitudinal studies to determine the predictive validity of its scores. It will be important to evaluate whether the VKC-2 outperforms conventional measures in identifying individuals at risk for cognitive and functional decline. However, even if the VKC-2 performs comparably to traditional cognitive tests or questionnaires, it offers important advantages, including greater efficiency and independence from the need for a reliable informant. Another important future direction is validation of the VKC-2 against biomarkers of neurodegenerative disease. Holmqvist and colleagues [48] demonstrated strong correlations between VKC-2 scores and magnetic resonance imaging–derived measures of cerebral white matter hyperintensities, a biomarker of small vessel disease associated with brain aging and neurodegeneration. Ongoing studies are examining associations between VKC-2 scores and additional biomarkers, including Alzheimer disease–specific positron emission tomography and blood markers. Finally, automated VKC-2 scores that capture task accomplishment are under development, which will further enhance the utility of the VKC-2 by providing a detailed characterization of everyday task performance patterns [85]. Future implementation research should examine and address potential barriers to VKC-2 adoption, including variability in technology literacy, digital skill levels, and computer-related anxiety among diverse older adults [91], as well as strategies for seamless integration into existing clinical workflows.

In conclusion, there is growing interest in the development of digital assessments of cognition, including digitized versions of traditional cognitive tests, smartphone- and tablet-based cognitive assessments, and VR [92]. Digital, performance-based

XSL•FO
RenderX

assessments of everyday tasks, such as the VKC-2, extend this trend to meaningful measures of everyday function. The VKC was designed to address weaknesses of conventional functional measures by providing an objective, standardized, and highly efficient assessment that does not rely on informant reports. The VKC-2 requires approximately 15-20 minutes to administer, is suitable for the full spectrum of cognitive aging—from healthy aging to mild dementia—and includes tasks (breakfast and lunch) that have been extensively studied and shown to be highly familiar to older adults [27,28,32,47,93]. The VKC-2 can be administered on a portable laptop without the need for additional objects or supplies, including a VR headset, avoiding limitations associated with cybersickness and confusion. The touchscreen interface provides a more natural interaction than a mouse or joystick for older adults [94]. Finally, the VKC-2 provides sensitive and detailed performance analysis, including time to completion and measures of performance efficiency derived from the touchscreen, eliminating the need for video recording and human coders. Older adult participants in this study were able to use the touchscreen interface, understood the instructions, did not require extensive training, and performed the tasks consistently with expectations based on data from the Real Kitchen [35]. The VKC-2 shows strong potential as an ecologically valid and scalable tool for capturing everyday functional capabilities in people with healthy cognition, MCI, and mild dementia across various settings, including large longitudinal studies, health clinics, and clinical trials.

Acknowledgments

The authors acknowledge Anna Callahan, Julina Hossfeld, Nicole Lloyd, Julia Manganti, Shrey Patel, Emma Pinsky, and Yuki Tsuchiya for their contributions to the Virtual Kitchen Challenge Project, including participant recruitment, assistance with data collection, and database management. The authors also thank Ms. Sherry Hill, Director of the Community Engagement Fitness Network, for her support with community outreach and participant recruitment. Generative artificial intelligence was not used in the creation of this manuscript.

Funding

This study was funded by grants from the National Institute on Aging (R01AG062503 to TG; F31AG089944 to SH).

Data Availability

The dataset analyzed for this study will be made available in the Virtual Kitchen Challenge Repository. In the interim, data requests may be sent to the corresponding author (TG).

Authors' Contributions

Conceptualization: TG (lead), TY (equal)
Data curation: TG (lead), MK (equal), MM (equal)
Formal analysis: TG (lead), DAGD (equal)
Funding acquisition: TG
Investigation: TG, MK, MM, SMS, MBT, K Hackett, SH, REM, K Halberstadter, RC, MR, GV, MDS
Methodology: TG (lead), TY (equal)
Project administration: TG (lead), MK (equal), MM (equal)
Resources: TG
Supervision: TG
Validation: TG
Visualization: TG (lead), MK (supporting)
Writing – original draft: TG (lead), MK (supporting)
Writing – review & editing: TG (lead), MK (equal), MDS (supporting), SMS (supporting), MBT (supporting), REM (supporting), K Hackett (supporting), K Halberstadter (supporting), DAGD (supporting), MR (supporting), SH (supporting), TY (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional analysis.
[DOCX File , 36 KB - aging_v9i1e82092_app1.docx]

References

1. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Behavioral and Social Sciences and Education , Board on Health Sciences Policy, Board on Behavioral, Cognitive, and Sensory Sciences, Committee on

XSL•FO
RenderX

Research Priorities for Preventing and Treating Alzheimer's Disease and Related Dementias. In: Yost OC, Downey A, Powell T, editors. Preventing and Treating Dementia: Research Priorities to Accelerate Progress. United States: National Academies Press; 2025. URL: https://doi.org/10.17226/28588
2. Pérès K, Chrysostome V, Fabrigoule C, Orgogozo JM, Dartigues JF, Barberger-Gateau P. Restriction in complex activities of daily living in MCI: impact on outcome. Neurology 2006 Aug 08;67(3):461-466. [doi: 10.1212/01.wnl.0000228228.70065.f1] [Medline: 16894108]
3. Pérès K, Helmer C, Amieva H, Orgogozo J, Rouch I, Dartigues J, et al. Natural history of decline in instrumental activities of daily living performance over the 10 years preceding the clinical diagnosis of dementia: a prospective population-based study. J Am Geriatr Soc 2008 Jan 20;56(1):37-44. [doi: 10.1111/j.1532-5415.2007.01499.x] [Medline: 18028344]
4. Lau KM, Parikh M, Harvey DJ, Huang C, Farias ST. Early cognitively based functional limitations predict loss of independence in instrumental activities of daily living in older adults. J Int Neuropsychol Soc 2015 Oct;21(9):688-698 [FREE Full text] [doi: 10.1017/S1355617715000818] [Medline: 26391766]
5. Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C, Davies P, Goldberg TE, Alzheimer's Disease Neuroimaging Initiative. Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's disease neuroimaging initiative. Arch Gen Psychiatry 2011 Sep 05;68(9):961-969. [doi: 10.1001/archgenpsychiatry.2011.96] [Medline: 21893661]
6. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004 Sep;256(3):183-194 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01388.x] [Medline: 15324362]
7. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May 22;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
8. Edgar CJ, Vradenburg G, Hassenstab J. The 2018 revised FDA guidance for early Alzheimer's disease: establishing the meaningfulness of treatment effects. J Prev Alzheimers Dis 2019;6(4):223-227. [doi: 10.14283/jpad.2019.30] [Medline: 31686092]
9. Mintun M, Ritchie CW, Solomon P, Sims JR, Salloway S, Hansson O, et al. Donanemab in early symptomatic Alzheimer's disease: efficacy and safety in TRAILBLAZER‐ALZ 2, a phase 3 randomized clinical trial. Alzheimer's & Dementia 2023 Dec 25;19(S24):e082733. [doi: 10.1002/alz.082733]
10. Rabinovici GD, La Joie R. Amyloid-targeting monoclonal antibodies for Alzheimer disease. JAMA 2023 Aug 08;330(6):507-509. [doi: 10.1001/jama.2023.11703] [Medline: 37459124]
11. Harvey PD, Cosentino S, Curiel R, Goldberg TE, Kaye J, Loewenstein D, et al. Performance-based and observational assessments in clinical trials across the Alzheimer's disease spectrum. Innov Clin Neurosci 2017;14(1-2):30-39 [FREE Full text] [Medline: 28386519]
12. Stoeckel LE, Fazio EM, Hardy KK, Kidwiler N, McLinden KA, Williams B. Clinically meaningful outcomes in Alzheimer's disease and Alzheimer's disease related dementias trials. Alzheimers Dement (N Y) 2025 Feb 19;11(1):e70058. [doi: 10.1002/trc2.70058] [Medline: 39975465]
13. Farias ST, Mungas D, Reed BR, Cahn-Weiner D, Jagust W, Baynes K, et al. The measurement of everyday cognition (ECog): scale development and psychometric properties. Neuropsychology 2008 Jul;22(4):531-544 [FREE Full text] [doi: 10.1037/0894-4105.22.4.531] [Medline: 18590364]
14. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol 1982 May 01;37(3):323-329. [doi: 10.1093/geronj/37.3.323] [Medline: 7069156]
15. Schmitter-Edgecombe M, Parsey C, Lamb R. Development and psychometric properties of the instrumental activities of daily living: compensation scale. Arch Clin Neuropsychol 2014 Dec 25;29(8):776-792 [FREE Full text] [doi: 10.1093/arclin/acu053] [Medline: 25344901]
16. Hackett K, Mis R, Drabick DA, Giovannetti T. Informant reporting in mild cognitive impairment: sources of discrepancy on the Functional Activities Questionnaire. J Int Neuropsychol Soc 2020 Jan 22;26(5):503-514. [doi: 10.1017/s1355617719001449]
17. Wadley VG, Harrell LE, Marson DC. Self- and informant report of financial abilities in patients with Alzheimer's disease: reliable and valid? J Am Geriatr Soc 2003 Nov 24;51(11):1621-1626. [doi: 10.1046/j.1532-5415.2003.51514.x] [Medline: 14687393]
18. Okonkwo OC, Wadley VG, Griffith HR, Belue K, Lanza S, Zamrini EY, et al. Awareness of deficits in financial abilities in patients with mild cognitive impairment: going beyond self-informant discrepancy. The American Journal of Geriatric Psychiatry 2008 Aug;16(8):650-659. [doi: 10.1097/jgp.0b013e31817e8a9d]
19. Tassoni MB, Drabick DAG, Giovannetti T. The frequency of self-reported memory failures is influenced by everyday context across the lifespan: Implications for neuropsychology research and clinical practice. Clin Neuropsychol 2023 Aug;37(6):1115-1135. [doi: 10.1080/13854046.2022.2112297] [Medline: 36000515]
20. Rycroft SS, Quach LT, Ward RE, Pedersen MM, Grande L, Bean JF. The relationship between cognitive impairment and upper extremity function in older primary care patients. J Gerontol A Biol Sci Med Sci 2019 Mar 14;74(4):568-574 [FREE Full text] [doi: 10.1093/gerona/gly246] [Medline: 30358815]

XSL•FO
RenderX

21. Hesseberg K, Tangen G, Pripp A, Bergland A. Associations between cognition and hand function in older people diagnosed with mild cognitive impairment or dementia. Dement Geriatr Cogn Dis Extra 2020 Dec 15;10(3):195-204 [FREE Full text] [doi: 10.1159/000510382] [Medline: 33569075]
22. Teruya SL, Dimino C, Silverman KD, Mielenz T. Poor lower extremity functioning is associated with modest increased incidence of probable dementia. Geriatrics (Basel) 2021 Aug 10;6(3):77 [FREE Full text] [doi: 10.3390/geriatrics6030077] [Medline: 34449642]
23. Giovannetti T, Seligman SC, Britnell P, Brennan L, Libon DJ. Differential effects of goal cues on everyday action errors in Alzheimer's disease versus Parkinson's disease dementia. Neuropsychology 2015 Jul;29(4):592-602. [doi: 10.1037/neu0000167] [Medline: 25495833]
24. Brennan L, Giovannetti T, Libon DJ, Bettcher BM, Duey K. The impact of goal cues on everyday action performance in dementia. Neuropsychol Rehabil 2009 Aug;19(4):562-582. [doi: 10.1080/09602010802405623] [Medline: 18923960]
25. Schwartz MF, Segal M, Veramonti T, Ferraro M, Buxbaum LJ. The Naturalistic Action Test: a standardised assessment for everyday action impairment. Neuropsychological Rehabilitation 2010 Oct 22;12(4):311-339. [doi: 10.1080/09602010244000084]
26. Bruce I, Ntlholang O, Crosby L, Cunningham C, Lawlor B. The clinical utility of naturalistic action test in differentiating mild cognitive impairment from early dementia in memory clinic. Int J Geriatr Psychiatry 2016 Mar 11;31(3):309-315. [doi: 10.1002/gps.4331] [Medline: 26264127]
27. Giovannetti T, Libon DJ, Buxbaum LJ, Schwartz MF. Naturalistic action impairments in dementia. Neuropsychologia 2002 Jan;40(8):1220-1232. [doi: 10.1016/s0028-3932(01)00229-9] [Medline: 11931925]
28. Giovannetti T, Bettcher BM, Brennan L, Libron DJ, Kessler RK, Duey K. Coffee with jelly or unbuttered toast: commissions and omissions are dissociable aspects of everyday action impairment in Alzheimer's disease. Neuropsychology 2008 Mar;22(2):235-245. [doi: 10.1037/0894-4105.22.2.235] [Medline: 18331166]
29. Giovannetti T, Bettcher BM, Libon DJ, Brennan L, Sestito N, Kessler RK. Environmental adaptations improve everyday action performance in Alzheimer's disease: empirical support from performance-based assessment. Neuropsychology 2007 Jul;21(4):448-457. [doi: 10.1037/0894-4105.21.4.448] [Medline: 17605578]
30. Giovannetti T, Libon DJ, Hart T. Awareness of naturalistic action errors in dementia. J Int Neuropsychol Soc 2002 Jul 14;8(5):633-644. [doi: 10.1017/s135561770280131x] [Medline: 12164673]
31. Giovannetti T, Schmidt KS, Gallo JL, Sestito N, Libon DJ. Everyday action in dementia: evidence for differential deficits in Alzheimer's disease versus subcortical vascular dementia. J Int Neuropsychol Soc 2006 Jan 23;12(1):45-53. [doi: 10.1017/S1355617706060012] [Medline: 16433943]
32. Giovannetti T, Bettcher BM, Brennan L, Libon DJ, Burke M, Duey K, et al. Characterization of everyday functioning in mild cognitive impairment: a direct assessment approach. Dement Geriatr Cogn Disord 2008 Mar 14;25(4):359-365. [doi: 10.1159/000121005] [Medline: 18340108]
33. Cornelis E, Gorus E, Van Weverbergh K, Beyer I, De Vriendt P. Convergent and concurrent validity of a report- versus performance-based evaluation of everyday functioning in the diagnosis of cognitive disorders in a geriatric population. Int Psychogeriatr 2018 Dec;30(12):1837-1848 [FREE Full text] [doi: 10.1017/S1041610218000327] [Medline: 29564999]
34. Bettcher BM, Giovannetti T, Libon DJ, Eppig J, Wambach D, Klobusicky E. Improving everyday error detection, one picture at a time: a performance-based study of everyday task training. Neuropsychology 2011 Nov;25(6):771-783. [doi: 10.1037/a0024107] [Medline: 21639639]
35. McKniff M, Holmqvist S, Kaplan M, Simone SM, Tassoni MB, Mis RE, et al. Subtle inefficiencies in everyday tasks indicate early functional difficulties in older adults: implications for clinical practice and research. Clin Neuropsychol 2025 Apr 30:1-20. [doi: 10.1080/13854046.2025.2497381] [Medline: 40304615]
36. Rycroft SS, Giovannetti T, Divers R, Hulswit J. Sensitive performance-based assessment of everyday action in older and younger adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2018 Mar 07;25(2):259-276. [doi: 10.1080/13825585.2017.1287855] [Medline: 28270012]
37. Seligman SC, Giovannetti T, Sestito J, Libon DJ. A new approach to the characterization of subtle errors in everyday action: implications for mild cognitive impairment. Clin Neuropsychol 2014 Nov 05;28(1):97-115. [doi: 10.1080/13854046.2013.852624] [Medline: 24191855]
38. Divers R, Ham L, Matchanova A, Hackett K, Mis R, Howard K, et al. When and how did you go wrong? Characterizing mild functional difficulties in older adults during an everyday task. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2021 Mar 30;28(2):308-326 [FREE Full text] [doi: 10.1080/13825585.2020.1756210] [Medline: 32352347]
39. Giovannetti T, Yamaguchi T, Roll E, Harada T, Rycroft SS, Divers R, et al. The Virtual Kitchen Challenge: preliminary data from a novel virtual reality test of mild difficulties in everyday functioning. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2019 Nov 28;26(6):823-841. [doi: 10.1080/13825585.2018.1536774] [Medline: 30370825]
40. Diehl M, Willis SL, Schaie KW. Everyday problem solving in older adults: observational assessment and cognitive correlates. Psychol Aging 1995 Sep;10(3):478-491. [doi: 10.1037//0882-7974.10.3.478] [Medline: 8527068]
41. McAlister C, Schmitter-Edgecombe M. Naturalistic assessment of executive function and everyday multitasking in healthy older adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2013 Nov;20(6):735-756 [FREE Full text] [doi: 10.1080/13825585.2013.781990] [Medline: 23557096]

XSL•FO
RenderX

42. Sanders C, Schmitter-Edgecombe M. Identifying the nature of impairment in planning ability with normal aging. J Clin Exp Neuropsychol 2012 Aug;34(7):724-737. [doi: 10.1080/13803395.2012.670210] [Medline: 22506736]
43. Schmitter-Edgecombe M, Parsey C, Cook DJ. Cognitive correlates of functional performance in older adults: comparison of self-report, direct observation, and performance-based measures. J Int Neuropsychol Soc 2011 Jun 27;17(5):853-864. [doi: 10.1017/s1355617711000865]
44. Schmitter-Edgecombe M, McAlister C, Weakley A. Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: the day-out task. Neuropsychology 2012 Sep;26(5):631-641 [FREE Full text] [doi: 10.1037/a0029352] [Medline: 22846035]
45. Griffith H, Belue K, Sicola A, Krzywanski S, Zamrini E, Harrell L, et al. Impaired financial abilities in mild cognitive impairment: a direct assessment approach. Neurology 2003 Feb 11;60(3):449-457. [doi: 10.1212/wnl.60.3.449] [Medline: 12578926]
46. Gold DA, Park NW, Murphy KJ, Troyer AK. Naturalistic action performance distinguishes amnestic mild cognitive impairment from healthy aging. J Int Neuropsychol Soc 2015 Jul 06;21(6):419-428. [doi: 10.1017/s135561771500048x]
47. Giovannetti T, Britnell P, Brennan L, Siderowf A, Grossman M, Libon DJ, et al. Everyday action impairment in Parkinson's disease dementia. J Int Neuropsychol Soc 2012 May 24;18(5):787-798. [doi: 10.1017/s135561771200046x]
48. Holmqvist SL, Jobson K, Desalme D, Simone SM, Tassoni M, McKniff M, et al. Preliminary validation of the Virtual Kitchen Challenge as an objective and sensitive measure of everyday function associated with cerebrovascular disease. Alzheimers Dement (Amst) 2024 Feb 05;16(1):e12547 [FREE Full text] [doi: 10.1002/dad2.12547] [Medline: 38318469]
49. Giovannetti T, Schwartz MF, Buxbaum LJ. The Coffee Challenge: a new method for the study of everyday action errors. J Clin Exp Neuropsychol 2007 Oct 24;29(7):690-705. [doi: 10.1080/13803390600932286] [Medline: 17891679]
50. Kessler RK, Giovannetti T, MacMullen LR. Everyday action in schizophrenia: performance patterns and underlying cognitive mechanisms. Neuropsychology 2007 Jul;21(4):439-447. [doi: 10.1037/0894-4105.21.4.439] [Medline: 17605577]
51. Devlin KN, Giovannetti T, Kessler RK, Fanning MJ. Commissions and omissions are dissociable aspects of everyday action impairment in schizophrenia. J Int Neuropsychol Soc 2014 Jul 30;20(8):812-821. [doi: 10.1017/s1355617714000654]
52. Mis RE, Ando T, Yamaguchi T, Brough C, Michalski L, Hoffman LJ, et al. Assessing everyday action in young adult athletes using the Virtual Kitchen Challenge: relations with conventional cognitive tests. J Int Neuropsychol Soc 2025 Sep 01;31(5-6):423-429. [doi: 10.1017/s135561772510101x]
53. Loewenstein DA, Amigo E, Duara R, Guterman A, Hurwitz D, Berkowitz N, et al. A new scale for the assessment of functional status in Alzheimer's disease and related disorders. J Gerontol 1989 Jul 01;44(4):P114-P121. [doi: 10.1093/geronj/44.4.p114] [Medline: 2738312]
54. Weiner MF, Gehrmann HR, Hynan LS, Saine KC, Cullum CM. Comparison of the test of everyday functional abilities with a direct measure of daily function. Dement Geriatr Cogn Disord 2006 Jun 19;22(1):83-86. [doi: 10.1159/000093388] [Medline: 16710087]
55. Binegar D, Hynan L, Lacritz L, Weiner M, Cullum C. Can a direct IADL measure detect deficits in persons with MCI? Curr Alzheimer Res 2009 Feb 01;6(1):48-51 [FREE Full text] [doi: 10.2174/156720509787313880] [Medline: 19199874]
56. Goldberg TE, Koppel J, Keehlisen L, Christen E, Dreses-Werringloer U, Conejero-Goldberg C, et al. Performance-based measures of everyday function in mild cognitive impairment. Am J Psychiatry 2010 Jul 01;167(7):845-853 [FREE Full text] [doi: 10.1176/appi.ajp.2010.09050692] [Medline: 20360320]
57. Allain P, Foloppe DA, Besnard J, Yamaguchi T, Etcharry-Bouyx F, Le Gall D, et al. Detecting everyday action deficits in Alzheimer's disease using a nonimmersive Virtual Reality Kitchen. J Int Neuropsychol Soc 2014 Mar 10;20(5):468-477. [doi: 10.1017/s1355617714000344]
58. Park MS, Kang KJ, Jang SJ, Lee JY, Chang SJ. Evaluating test-retest reliability in patient-reported outcome measures for older people: a systematic review. Int J Nurs Stud 2018 Mar;79:58-69. [doi: 10.1016/j.ijnurstu.2017.11.003] [Medline: 29178977]
59. Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. The American Journal of Geriatric Psychiatry 2009 May;17(5):368-375. [doi: 10.1097/jgp.0b013e31819431d5]
60. Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. JAD 2014 Aug 11;42(1):275-289. [doi: 10.3233/jad-140276]
61. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May 22;7(3):263-269 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.005] [Medline: 21514250]
62. Schretlen D, Testa S, Pearlson GD. Calibrated Neuropsychological Normative System Professional Manual. Lutz, FL: Psychological Assessment Resources; 2010.
63. Manly JJ, Byrd DA, Touradji P, Stern Y. Acculturation, reading level, and neuropsychological test performance among African American elders. Appl Neuropsychol 2004 Mar;11(1):37-46. [doi: 10.1207/s15324826an1101_5] [Medline: 15471745]

XSL•FO
RenderX

64. Manly JJ, Jacobs DM, Touradji P, Small SA, Stern Y. Reading level attenuates differences in neuropsychological test performance between African American and White elders. J Int Neuropsychol Soc 2002 Mar 16;8(3):341-348. [doi: 10.1017/s1355617702813157] [Medline: 11939693]
65. Schretlen DJ, Winicki JM, Meyer SM, Testa SM, Pearlson GD, Gordon B. Development, psychometric properties, and validity of the Hopkins Adult Reading Test (HART). Clin Neuropsychol 2009 Aug;23(6):926-943. [doi: 10.1080/13854040802603684] [Medline: 19191072]
66. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
67. Brandt J, Benedict HRB. Hopkins Verbal Learning Test-Revised Professional Manual. Lutz, FL: Psychologica Assessment Resources, Inc; 2001.
68. Benedict RHB, Schretlen D, Groninger L, Dobraski M, et al. Revision of the Brief Visuospatial Memory Test: studies of normal performance, reliability, and validity. Psychological Assessment 1996;8(2):145-153. [doi: 10.1037//1040-3590.8.2.145]
69. Goodglass H, Kaplan E. Boston Diagnostic Aphasia Examination. Philadelphia, PA: Lea and Febiger; 1983.
70. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. Percept Mot Skills 1958 Dec 01;8(3):271-276. [doi: 10.2466/pms.1958.8.3.271]
71. Wechsler D. Wechsler Memory Scale-Revised Manual. San Antonio, TX: Psychological Corporation; 1987.
72. Salthouse TA. Mediation of adult age differences in cognition by reductions in working memory and speed of processing. Psychol Sci 1991 May 01;2(3):179-183. [doi: 10.1111/j.1467-9280.1991.tb00127.x]
73. McKay NS, Millar PR, Nicosia J, Aschenbrenner AJ, Gordon BA, Benzinger TLS, et al. Pick a PACC: comparing domain-specific and general cognitive composites in Alzheimer disease research. Neuropsychology 2024 Jul;38(5):443-464 [FREE Full text] [doi: 10.1037/neu0000949] [Medline: 38602816]
74. Schwartz MF, Montgomery MW, Buxbaum LJ, Lee SS, Carew TG, Coslett HB, et al. Naturalistic action impairment in closed head injury. Neuropsychology 1998;12(1):13-28. [doi: 10.1037/0894-4105.12.1.13]
75. Schwartz MF, Buxbaum LJ, Montgomery MW, Fitzpatrick-DeSalme E, Hart T, Ferraro M, et al. Naturalistic action production following right hemisphere stroke. Neuropsychologia 1999 Jan;37(1):51-66. [doi: 10.1016/s0028-3932(98)00066-9] [Medline: 9920471]
76. Buxbaum LJ. Ideational apraxia and naturalistic action. Cogn Neuropsychol 1998 Sep 01;15(6-8):617-643. [doi: 10.1080/026432998381032] [Medline: 22448839]
77. Marson D, Sawrie SM, Snyder S, McInturff B, Stalvey T, Boothe A, et al. Assessing financial capacity in patients with Alzheimer disease: a conceptual model and prototype instrument. Arch Neurol 2000 Jun;57(6):877-884. [doi: 10.1001/archneur.57.6.877] [Medline: 10867786]
78. Farias S, Mungas D, Reed B, Harvey D, Cahn-Weiner D, Decarli C. MCI is associated with deficits in everyday functioning. Alzheimer Dis Assoc Disord 2006;20(4):217-223 [FREE Full text] [doi: 10.1097/01.wad.0000213849.51495.d9] [Medline: 17132965]
79. BM SPSS Statistics for Windows (Version 29). Armonk, NY: IBM Corp; 2024.
80. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. Acta Paediatr 2007 May 21;96(5):644-647. [doi: 10.1111/j.1651-2227.2006.00178.x] [Medline: 17376185]
81. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin 1979;86(2):420-428. [doi: 10.1037/0033-2909.86.2.420]
82. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
83. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159. [doi: 10.2307/2529310]
84. Eisinga R, Grotenhuis MT, Pelzer B. The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? Int J Public Health 2013 Aug 23;58(4):637-642. [doi: 10.1007/s00038-012-0416-3] [Medline: 23089674]
85. Giovannetti T, Mis R, Hackett K, Simone SM, Ungrady MB. The goal-control model: An integrated neuropsychological framework to explain impaired performance of everyday activities. Neuropsychology 2021 Jan;35(1):3-18 [FREE Full text] [doi: 10.1037/neu0000714] [Medline: 33393796]
86. Cooper RP, Shallice T. Hierarchical schemas and goals in the control of sequential behavior. Psychological Review 2006;113(4):887-916. [doi: 10.1037/0033-295x.113.4.887]
87. Chirico A, Giovannetti T, Neroni P, Simone S, Gallo L, Galli F, et al. Virtual reality for the assessment of everyday cognitive functions in older adults: an evaluation of the virtual reality action test and two interaction devices in a 91-year-old woman. Front Psychol 2020;11:123 [FREE Full text] [doi: 10.3389/fpsyg.2020.00123] [Medline: 32116928]
88. Dodge H, Mattek N, Austin D, Hayes T, Kaye J. In-home walking speeds and variability trajectories associated with mild cognitive impairment. Neurology 2012 Jun 12;78(24):1946-1952 [FREE Full text] [doi: 10.1212/WNL.0b013e318259e1de] [Medline: 22689734]
89. Mcalister C, Schmitter-Edgecombe M, Lamb R. Examination of variables that may affect the relationship between cognition and functional status in individuals with mild cognitive impairment: a meta-analysis. Arch Clin Neuropsychol 2016 Mar 06;31(2):123-147 [FREE Full text] [doi: 10.1093/arclin/acv089] [Medline: 26743326]

XSL•FO
RenderX

90. Royall DR, Lauterbach EC, Kaufer D, Malloy P, Coburn KL, Black KJ. The cognitive correlates of functional status: a review from the Committee on Research of the American Neuropsychiatric Association. Journal of Neuropsychiatry 2007 Aug 01;19(3):249-265. [doi: 10.1176/appi.neuropsych.19.3.249]
91. Raihan MM, Subroto S, Chowdhury N, Koch K, Ruttan E, Turin TC. Dimensions and barriers for digital (in)equity and digital divide: a systematic integrative review. DTS 2024 Aug 13;4(2):111-127. [doi: 10.1108/dts-04-2024-0054]
92. Öhman F, Hassenstab J, Berron D, Schöll M, Papp K. Current advances in digital cognitive assessment for preclinical Alzheimer's disease. Alzheimers Dement (Amst) 2021;13(1):e12217 [FREE Full text] [doi: 10.1002/dad2.12217] [Medline: 34295959]
93. Roll EE, Giovannetti T, Libon DJ, Eppig J. Everyday task knowledge and everyday function in dementia. J Neuropsychol 2019 Mar 26;13(1):96-120. [doi: 10.1111/jnp.12135] [Medline: 28949080]
94. Holzinger A. Finger instead of mouse: touch screens as a means of enhancing universal access. In: Carbonell C, Stephanidis N, editors. Universal Access: Theoretical Perspectives, Practice and Experience. Berlin/Heidelberg, Germany: Springer; 2003:387-397.

Abbreviations

ANCOVA: analysis of covariance
AUC: area under the curve
ECog-12: 12-item Everyday Cognition Scale
FAQ: Functional Activity Questionnaire
IADL-C: Instrumental Activities of Daily Living—Compensation
ICC: intraclass correlation coefficient
IPS: In-Plane Switching
Knight-PACC: Knight-Preclinical Alzheimer Cognitive Composite
MCI: mild cognitive impairment
mKnight-PACC: modified Knight-Preclinical Alzheimer Cognitive Composite
NAT: Naturalistic Action Test
PCA: principal component analysis
QHD+: Quad High Definition Plus
VKC: Virtual Kitchen Challenge
VKC-2: Virtual Kitchen Challenge—Version 2
VR: virtual reality

Edited by M O'Connell; submitted 12.Aug.2025; peer-reviewed by Z Kunicki, F Akhter; comments to author 19.Oct.2025; revised version received 09.Nov.2025; accepted 05.Dec.2025; published 07.Jan.2026.

Please cite as:
Kaplan M, McKniff M, Simone SM, Tassoni MB, Hackett K, Holmqvist S, Mis RE, Halberstadter K, Chaturvedi R, Rosahl M, Vallecorsa G, Serruya MD, Drabick DAG, Yamaguchi T, Giovannetti T
The Virtual Kitchen Challenge—Version 2: Validation of a Digital Assessment of Everyday Function in Older Adults
JMIR Aging 2026;9:e82092
URL: https://aging.jmir.org/2026/1/e82092
doi:10.2196/82092
PMID:41349042

©Marina Kaplan, Moira McKniff, Stephanie M Simone, Molly B Tassoni, Katherine Hackett, Sophia Holmqvist, Rachel E Mis, Kimberly Halberstadter, Riya Chaturvedi, Melissa Rosahl, Giuliana Vallecorsa, Mijiail D Serruya, Deborah A G Drabick, Takehiko Yamaguchi, Tania Giovannetti. Originally published in JMIR Aging (https://aging.jmir.org), 07.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Developing Consumer Consensus on Remote Assessment and Management of Physical Function in Older Adults (RAMP): International Modified Delphi Process

Elsa Dent[1], PhD; Christopher Hurst[2,3], PhD; Jack Dalla Via[4,5], PhD; Jackson J Fyfe[4], PhD; Paul Jansons[4,6], PhD; Eleanor J Hayes[7], PhD; Gary Skinner[4], MSc; Marc Sim[5,8], PhD; Mylene Aubertin-Leheudre[9,10], PhD; Sabine Britting[11], PhD; Fanny Buckinx[9,10,12], PhD; Gavin Connolly[13], PhD; Ruth Dignam[14†]; Lora Giangregorio[15,16], PhD; Jennifer R A Jones[17,18,19], PhD; Pauline Kelly[14]; Robert Kob[11], PhD; Suzanne N Morin[20], MD; Girish Nandakumar[21], PhD; Lucas B R Orssatto[22], PhD; Maria Pearson[23], MPH; Daniel Pinto[24], PT, PhD; Esmee M Reijnierse[25], PhD; Catherine M Said[17,26,27], PhD; Mohamed Salem[28], MB; Vina PS Tan[29], PhD; Rosanna Tran[30], BAPSC; Jesse Zanker[31], PhD; Robin M Daly[4], PhD; David Scott[4,6], PhD

[1]Division of Health, University of Waikato, Hamilton, New Zealand

[2]AGE Research Group, Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom

[3]NIHR Newcastle Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Cumbria Northumberland Tyne and Wear NHS Foundation Trust and Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom

[4]Institute for Physical Activity and Nutrition (IPAN), School of Exercise and Nutrition Sciences, Deakin University, Burwood, Australia

[5]Nutrition & Health Innovation Research Institute, School of Medical and Health Sciences, Edith Cowan University, Joondalup, Australia

[6]School of Clinical Sciences at Monash Health, Monash University, Clayton, Australia

[7]Department of Sport, Exercise and Rehabilitation, Northumbria University, Newcastle Upon Tyne, United Kingdom

[8]School of Medicine, The University of Western Australia, Perth, Australia

[9]Département des Sciences de l'Activité Physique, Faculté des Sciences, Groupe de Recherche en Activité Physique Adaptée (GRAPA), Université du Québec à Montréal, Montréal, QC, Canada

[10]Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal (CRIUGM), Montréal, QC, Canada

[11]Institute for Biomedicine of Aging, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nuremberg, Germany

[12]Research Unit in Public Health, Epidemiology and Health Economics, University of Liège, Liège, Belgium

[13]Metabolism and Healthy Aging Research Center (MHARC), Division of Endocrinology, Diabetes, and Metabolism, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, United States

[14]IPAN Consumer Network, Deakin University, Geelong, Australia

[15]Schlegel-UW Research Institute for Aging, University of Waterloo, Waterloo, ON, Canada

[16]Department of Kinesiology and Health Sciences, University of Waterloo, Waterloo, ON, Canada

[17]Physiotherapy Department, The University of Melbourne, Parkville, Australia

[18]Physiotherapy Department, Division of Allied Health, Austin Health, Heidelberg, Australia

[19]Institute of Breathing and Sleep, Heidelberg, Australia

[20]Division of General Internal Medicine, Department of Medicine, McGill University, Montréal, QC, Canada

[21]Department of Physiotherapy, Manipal College of Health Professions, Manipal Academy of Higher Education, Manipal, India

[22]Centre for Sensorimotor Performance, School of Human Movement and Nutrition Sciences, Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Brisbane, Australia

[23]Maria Pearson Coaching, Havelock North, New Zealand

[24]Department of Physical Therapy, Marquette University, Milwaukee, WI, United States

[25]Department of Nutrition and Dietetics, Faculty of Health, Sport and Physical Activity, Amsterdam University of Applied Sciences, Amsterdam, The Netherlands

[26]Physiotherapy, Western Health, St Albans, Australia

[27]Australian Institute for Musculoskeletal Science (AIMSS), The University of Melbourne and Western Health, St Albans, Australia

[28]Geriatrics, Active Ageing and Community Care, Ministry for Health and Active Ageing, Valletta, Malta

[29]Exercise and Sports Science Programme, School of Health Sciences, Universiti Sains Malaysia, Kubang Kerian, Malaysia

[30]Rehabilitation and Aged Care Services, Northern Sydney Local Health District, Sydney, Australia

[31]Department of Medicine - Royal Melbourne Hospital, The University of Melbourne, Parkville, Australia

[†]deceased

XSL•FO
RenderX

Corresponding Author:
David Scott, PhD
Institute for Physical Activity and Nutrition (IPAN), School of Exercise and Nutrition Sciences
Deakin University
221 Burwood Highway
Burwood, 3125
Australia
Phone: 61 9246 8438
Email: d.scott@deakin.edu.au

Abstract

Background: Remote health care delivery, including the use of digital health interventions, is emerging as a tool for assessing and managing physical function, but its design and implementation often overlook the needs and preferences of older adult end users.

Objective: The primary aim of this modified Delphi process was to develop consumer consensus on preferences for remote assessment and management of physical function in older adults.

Methods: Research and consumer experts of the Remote Assessment and Management of Physical Function in Older Adults (RAMP) Working Group co-developed the Round 1 Delphi survey, which was advertised to consumers (adults aged ≥60 years) via international clinical and research networks and social media between August and November 2023. The online survey presented 23 Delphi statements for which respondents reported their level of agreement using an 11-point Likert scale (0-10; scores ≥7 indicated agreement). Statements were classified as having "strong agreement" and achieving consensus if ≥80% of participants indicated agreement. Statements classified as having "moderate" (70%-80% of participants indicated agreement) or "low" (<70% of participants indicated agreement) agreement were revised or rejected. Revised statements were presented to participants in Round 2 (January to February 2024), and the final consensus statements were consolidated into recommendations.

Results: A total of 654 consumers (75.7% female) with a mean age of 69.0 (SD 6.0) years from 15 countries (5 continents) were included in analyses in Round 1. Of 23 statements, 13 achieved consensus, with the strongest agreement observed for statements relating to the importance of physical function for quality of life and performing activities of daily living (6 statements; agreement 97.6%-99.5%). Two statements regarding privacy and security concerns when using technology (agreement 20.8%) and the inability to perform physical function assessments or exercise at home (agreement 15.5%) were rejected with low agreement. The remaining 8 statements (agreement 49.5%-79.5%) were modified into 7 new statements for the Round 2 survey, which was completed by 526 (80.4%) respondents from Round 1. Five of seven Round 2 statements were accepted with strong agreement (agreement 80%-82.7%), including the importance of addressing personal preferences for self- versus clinician-led remote interventions, group versus individual exercise, and availability of necessary resources (eg, technology and exercise equipment).

Conclusions: Eighteen statements achieved consensus and were translated into 7 recommendations highlighting that older adults recognize physical function as a health priority, would value more information about it, and are willing to participate in remote assessment and management interventions (including via digital health) to maintain or improve it. These recommendations also reinforce that interventions should be easily accessible and meet individual preferences of consumers.

(JMIR Aging 2026;9:e75791) doi:10.2196/75791

KEYWORDS

aged; Delphi technique; exercise; remote consultation; digital health; physical function

Introduction

Maintaining physical function is crucial for older adults to live independently and maintain a good quality of life [1]. Depending on the task examined, up to 50% of older adults may report difficulty with physical function, and more than 10% use walking aids, with an increasing prevalence of functional limitation in older age (eg, ≥80 years) [2]. Older adults with functional limitations are half as likely to engage with their communities, family, and friends [3] and have significantly increased health care costs compared to those without functional limitations [4]. However, assessment and management of physical function receive limited attention in many clinical settings [5-7]. While multimodal exercise is recommended for improving physical function [8], older adults report that lack of access to interventions, facilities, and relevant health care professionals are barriers to participation [9].

Remote health care services enable patient assessment and monitoring without physical co-location [10] and, especially since the COVID-19 pandemic, have been commonly used by older adults [11]. Emerging evidence suggests digital health interventions using technologies such as telephone calls, videoconferencing, wearable devices, and web applications are feasible and effective for supporting older adults to maintain or

XSL•FO
RenderX

improve their physical function [12-14]. However, both consumers (ie, older adult end users) and health care professionals face barriers to engagement with digital health interventions, such as low digital literacy, lack of equipment, and unsatisfactory social interactions [15,16]. Furthermore, the design and implementation of digital health technologies for remote care have largely failed to consider the perceptions and experiences of consumers in their design and implementation, impeding large-scale sustained adoption [17].

The Delphi method, a rigorous consensus-building approach [18], is effective for developing recommendations for the delivery of care related to maintaining and improving physical function in older adults based on the insights of both experts and consumers (ie, older adults themselves) [19,20]. Incorporating consumer participation in consensus-building processes helps ensure that subsequent recommendations address end users' priorities, maximizing the potential for wide adoption and adherence among consumers in health care and community settings [21]. However, no consumer-focused consensus Delphi process has explored the priorities, acceptability, enablers, and barriers for remote assessment and management of physical function from older adults' perspectives, and there are no consensus guidelines on the delivery of remote care in older adult populations. Thus, this study aimed to develop consumer-informed recommendations for remote assessment and management of physical function in older adults via a modified Delphi process.

Methods

Study Design and Population

The Remote Assessment and Management of Physical Function in Older Adults (RAMP) International Consumer Delphi Process was a modified (2-round) Delphi study. Eligible participants were aged 60 years and older with internet access, residing in any country, and able to complete the survey in the English language.

RAMP was advertised to potential participants via email (consumer mailing lists of the RAMP Working Group), direct invitations (investigator-led conversations with consumers), and social media (Facebook [Meta Platforms, Inc] advertisements; posts on X [X Corp] and LinkedIn [LinkedIn Corp]) between August and November 2023.

Delphi Process

This study's methodology adhered to that of a previous modified Delphi process where 2 rounds, as demonstrated by common practice in previous Delphi studies [22,23], were sufficient to reach consensus on sarcopenia management [19]. Multimedia Appendix 1 summarizes the study timeline. The Round 1 survey (Multimedia Appendix 2) was co-developed by RAMP Working Group research experts (DS, CH, JF, PJ, and RMD) and consumer experts (RD and PK) and included questions on demographics, health status (including the SARC-F questionnaire [24]), and health care experiences, plus 23 Delphi statements related to physical function and its assessment and management. The first 13 statements relate to physical function generally, while the latter 10 statements refer to remote provision of health care related to physical function. Respondents were asked to report their level of agreement with each Delphi statement via an 11-point Likert scale (0=Strongly disagree, 10=Strongly agree) and could optionally provide a free-text comment explaining their response to each statement.

Round 1 survey responses were analyzed by DS, and the results and proposed Round 2 survey were shared with RAMP Working Group members for feedback. In January 2024, an invitation to participate in the Round 2 survey, which included seven new or modified statements, was emailed to respondents who completed Round 1, along with a summary of the Round 1 results and an explanation of decisions taken in developing the Round 2 statements (Multimedia Appendix 3). Further details on reasons for low or moderate agreement to Round 1 statements and decisions taken are provided in Multimedia Appendix 4. Upon completion of the Round 2 survey (Multimedia Appendix 5), DS analyzed the results, and ED and DS initially developed the final recommendations. Finally, this manuscript, including the developed recommendations, was revised (2 rounds of revisions) and approved by RAMP Working Group members.

Statistical Analyses

Survey data were assessed for completeness, and all respondents who completed ≥50% of survey questions were included in analyses. Where duplicate responses were identified, the most complete and/or first response was included. Descriptive characteristics were reported as frequencies or percentages for categorical variables and means and SDs or medians and IQR for continuous variables.

Participants with a response ≥7 out of 10 were considered to have agreed with a given statement, and the level of consensus was determined by the proportion of participants who agreed. Round 1 and 2 Delphi statements were classified as having "strong agreement" if ≥80% of participants responded with a score ≥7 out of 10 [19]. These statements were considered to have achieved consensus and were not further modified. In Round 1 only, statements with "moderate" (70%-80% of participants responded ≥7) or "low" (<70% of participants responded ≥7) agreement were revised or rejected. DS reviewed the associated free-text comments and revised the statements based on common reasons for lack of agreement. The revised statements were shared with the RAMP Working Group for approval. In Round 2, all statements that did not achieve "strong agreement" were rejected. All analyses were performed using SPSS Statistics 28 (IBM Corp).

Ethical Considerations

The study was approved by the Deakin University Human Ethics Advisory Group (Reference number: HEAG-H 111_2023) and complied with the Declaration of Helsinki. Potential participants were directed to an online plain language statement defining physical function (ie, "the ability of a person to perform everyday activities"), as well as the study aims and methods. Potential participants provided an electronic signature confirming their informed consent. Those who consented were subsequently emailed a link to the Round 1 Delphi survey (hosted by Qualtrics XM). Participants who completed the Round 1 survey were invited to participate in the Round 2 survey

XSL•FO
RenderX

between January and February 2024. Participants who consented but had not completed the Round 1 and 2 surveys were sent a reminder 2 weeks prior to the closing date. Participation was voluntary, and respondents did not receive any form of reimbursement.

Results

Round 1

A total of 861 complete consent forms were received, and the Round 1 Delphi survey subsequently received 716 responses during the Round 1 survey period. Of these responses, 50 (7%) completed less than 50% of the survey, and 12 (2%) were duplicate responses. Thus, 654 of 861 (76%) consented consumers were included in the Round 1 survey analyses. Most of the included respondents (n=644, 98%) completed the entire Round 1 survey, while 10 respondents rated their agreement with the first 13 statements only. Respondents' mean age was 69.0 (SD 6.0) years, and three-quarters were female (Table 1). Respondents were residing in 15 countries across 5 continents (Australia, Africa, Asia, Europe, and North America), although the majority (81.5%) were from Australia and the United Kingdom. Three-quarters of respondents had completed higher education and around 70% were retired. Over 87% rated their health as good to excellent, but more than 70% perceived their current physical function was "somewhat worse" or "much worse" than when they were 40 years old. The SARC-F instrument demonstrated that approximately 54% of respondents had some functional limitation, and around one-quarter had experienced at least one fall in the past year (Multimedia Appendix 6).

XSL•FO
RenderX

Table 1. Demographic characteristics of Remote Assessment and Management of Physical (RAMP) participants in Round 1.

Characteristic	All respondents (n=654)
Age (years), mean (SD)	69 (6.0)
Sex (female), n (%)	495 (75.7)
Country of residence, n (%)	
Australia	281 (43.0)
United Kingdom	252 (38.5)
Canada	32 (4.9)
Ireland	22 (3.4)
United States	20 (3.1)
New Zealand	15 (2.3)
Germany	7 (1.1)
Singapore	7 (1.1)
Malta	6 (0.9)
India	5 (0.8)
Netherlands	3 (0.5)
Belgium	1 (0.2)
France	1 (0.2)
Georgia	1 (0.2)
South Africa	1 (0.2)
Highest level of education, n (%)	
Infants or primary school	3 (0.5)
Secondary or high school	154 (23.5)
University, college, or other higher education	497 (76.0)
Employment status, n (%)	
Employed full-time	67 (10.2)
Employed part-time	90 (13.8)
Home duties	3 (0.5)
Pension	18 (2.8)
Retired	462 (70.6)
Student	3 (0.5)
Unemployed	11 (1.7)
General health, n (%)	
Excellent	89 (13.6)
Very good	255 (39.0)
Good	228 (34.9)
Fair	67 (10.2)
Poor	15 (2.3)
Physical function now compared with that at age 40 years, n (%)	
Much better	18 (2.8)
Somewhat better	44 (6.7)
Neither better nor worse	117 (17.9)
Somewhat worse	348 (53.2)
Much worse	127 (19.4)

XSL•FO
RenderX

Characteristic	All respondents (n=654)
SARC-F score, median (IQR)	1 (0-2)

Less than one-third of respondents reported that a health care professional had started a conversation with them about their physical function in the past 5 years, while 38% had initiated a conversation themselves (Table 2). More than 60% had sought information on physical function from sources other than a health care professional. Around 40% of respondents reported having ever completed a physical function assessment with a health care professional, and almost 50% had commenced a supervised exercise program aimed at improving their physical function. The majority of respondents (>60%) had commenced an unsupervised exercise program aimed at improving their physical function.

Table 2. Health care experiences related to physical function for Remote Assessment and Management of Physical (RAMP) participants in Round 1 (n= 654).

Question	Yes, n (%)	No, n (%)	Don't Know, n (%)
In the past five years, have you started a conversation with a health professional (eg, doctor, physiotherapist, and nurse) about your physical function (eg, asking how you can continue to stay independent as you get older, or why you might not be as strong as you were when you were younger)?	245 (37.5)	399 (61)	9 (1.4)
In the past 5 years, has a health professional (eg, doctor, physiotherapist, and nurse) started a conversation with you about your physical function?	209 (32)	436 (66.7)	7 (1.1)
Have you ever tried to find information about physical function from sources other than a health professional (eg, by asking a friend or family member, visiting a website, or reading a book or magazine)?	395 (60.4)	255 (39)	3 (0.5)
Have you ever completed a physical function test (eg, walking speed test, hand grip strength test, and chair stand test) under the supervision of a health professional (eg, where a health professional asked you to perform a specific test while under their supervision to determine whether your physical function was poor)?	270 (41.3)	376 (57.5)	7 (1.1)
Have you ever completed a physical function test (eg, walking speed test, hand grip strength test, and chair stand test) while NOT under the supervision of a health professional (eg, performing a specific test designed to determine whether your physical function is poor after you read or viewed instructions in a document or online)?	101 (15.4)	543 (83)	9 (1.4)
Has a health professional ever prescribed you an exercise program aimed at improving your physical function?	315 (48.2)	333 (50.9)	5 (0.8)
Have you ever commenced an exercise program aimed at improving your physical function while NOT under the supervision of a health professional (eg, an exercise program that you created for yourself, with or without the help of a friend or family member, or using a website/book/magazine)?	402 (61.5)	245 (37.5)	6 (0.9)

Over 58% of respondents had participated in a remote health care service, while less than 5% had participated in a remote physical function test, and less than 20% had engaged in remote physical function management such as an exercise program (Figure 1). For respondents who had participated in any form of remote care, telephone calls were most commonly used (77%), with video calls and emails/text messages each used by over 30% of respondents. For respondents who had participated in a remote physical function assessment, video calls were most commonly used (75%), followed by telephone calls (25%). For respondents who had participated in a remote treatment for their physical function, written documents (eg, flyers, brochures, magazines, and books) were most common (46%), followed by video calls (38%) and telephone calls (31%). Regardless of the type of remote care, respondents generally reported positive experiences; 75%-81% rated their experience as somewhat or very positive.

Only 5% of respondents reported that they would not be willing to participate in remote physical function tests and treatments (Multimedia Appendix 7), with optional comments indicating that this was due to concerns regarding lack of supervision, safety, lack of necessary technologies and technological familiarity, data privacy, hearing and visual problems, potential costs, and insufficient space in the home. Amongst those willing to participate in such interventions, videos (eg, on a website or DVD) were the most preferred delivery method (78%). Approximately 4% were willing to use other approaches; free-text comments indicated that these could include digital voice assistants, wearable devices, webinars, smart televisions, and wall charts. The most common perceived positives of remote tests and treatments, selected by over 80% of respondents, were the convenience of not needing to travel to appointments and the flexibility to perform assessments and exercises when suitable. Some (<4%) respondents nominated other positives in free-text comments, including avoiding body shaming or embarrassment when exercising, self-motivation, empowerment and autonomy for managing exercise, and better use of health care professionals' time. Regarding negatives of remote care, the lack of personalized guidance during exercises was reported by half (51%) of respondents, and concerns regarding the potential ineffectiveness of interventions and a lack of social interaction and motivation were each reported by more than

XSL•FO
RenderX

40%. Approximately 14% of respondents cited other potential negatives, including costs and/or lack of access to necessary equipment and technological devices, lack of confidence in performing exercises correctly, low motivation to exercise, lack of engagement with programs and health care professionals, and insufficient space to perform exercises at home.

Figure 1. Remote health care experiences of RAMP participants in Round 1, related to (A) any type of remote care, (B) remote care to assess physical function, and (C) remote care to manage physical function. Sublabels (1) describe the proportion of respondents who reported having participated in that type of remote care, (2) report the modalities by which respondents engaged with that type of remote care, and (3) report the perceived experience of respondents participating in that type of remote care. Questions 2 and 3 were only asked to participants who responded yes to having participated in that type of remote care. For question 2, participants could choose multiple responses, so the total does not equal 100%.

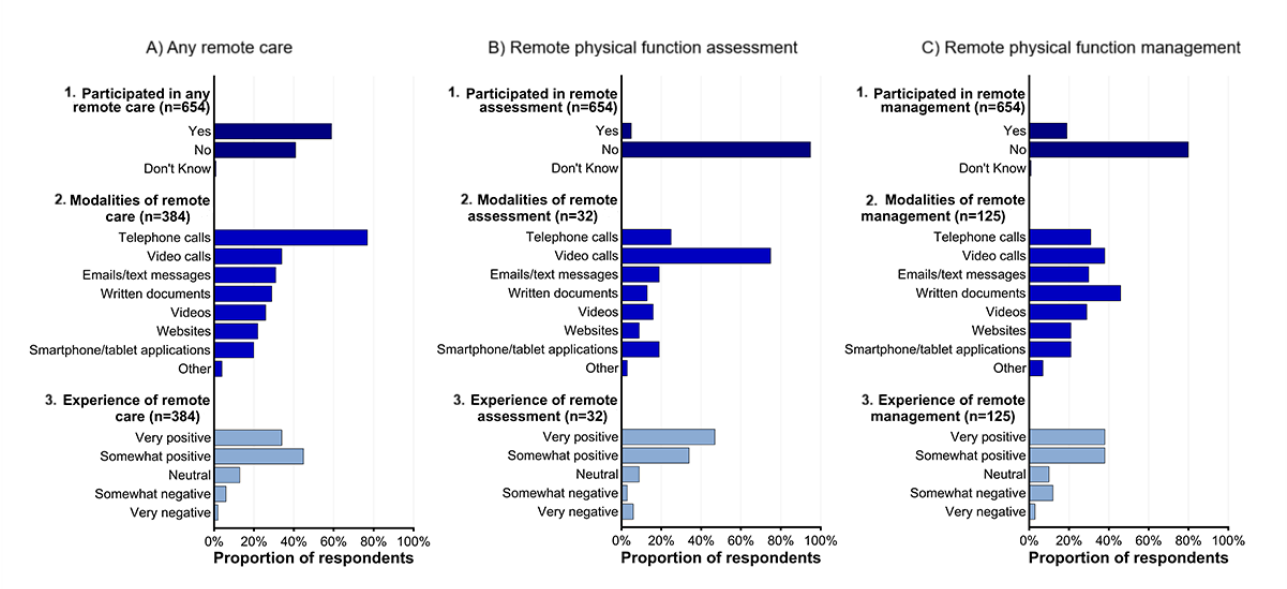
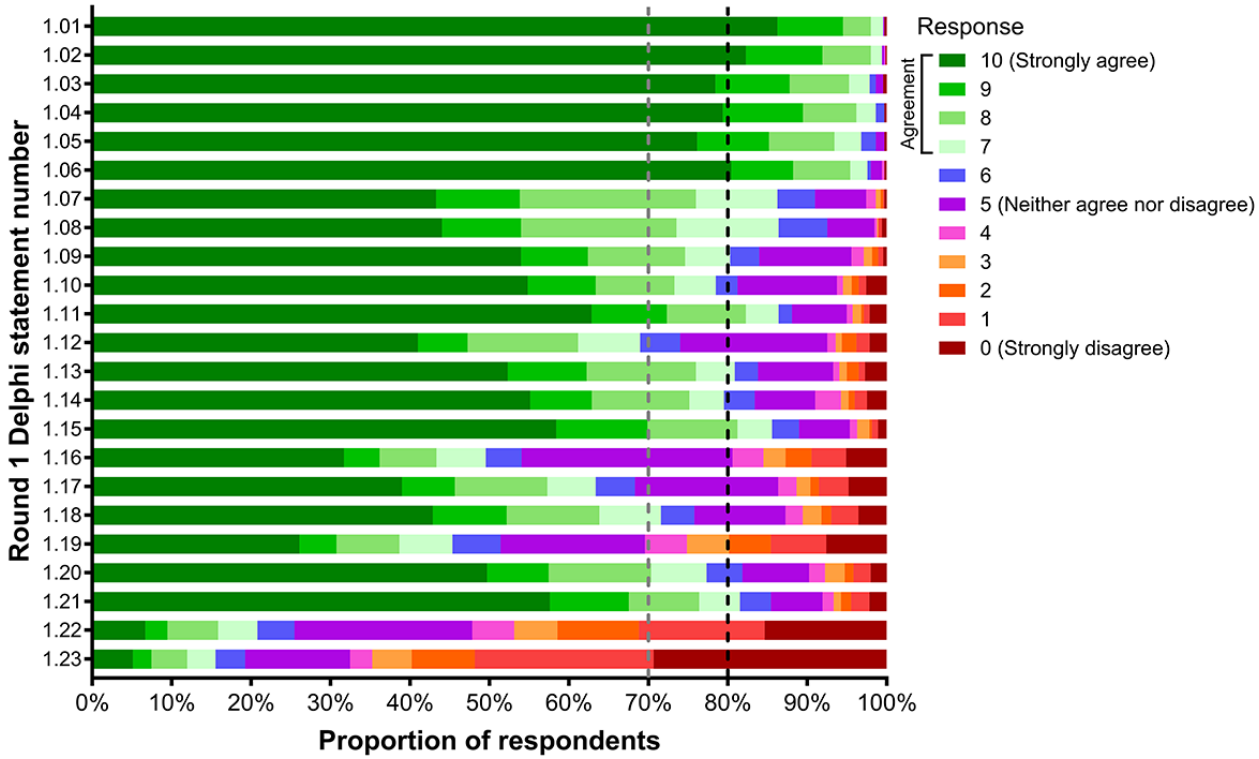


Figure 2 and Table 3 present respondents' reported agreement with the 23 Delphi statements from Round 1. Statements related to the importance of physical function for maintaining overall quality of life, participating in activities in the community and home, and engaging with family and friends (Statements 1.01-1.06) had the highest proportion (97%-100%) of respondents who reported agreement (response ≥7 out of 10) in this round and were accepted. Consensus was also achieved that poor physical function can be prevented and reversed (Statements 1.07 and 1.08; both 86% agreement) and that respondents would discuss with their health care professionals if concerned about their physical function (Statement 1.09; 80% agreement). There was moderate agreement that respondents would like access to information on how to test their physical function (Statement 1.10; 78% agreement) and low agreement that having better access to information would assist in conversations with health care professionals (Statement 1.12; 69% agreement). However, there was strong agreement that respondents would like access to information on how to improve their physical function (Statement 1.11; 86% agreement) and that having better access to information would assist in managing their physical function independently (Statement 1.13; 81% agreement).

Regarding remote provision of health care, almost 80% of respondents agreed that they would be willing to participate in remote physical function tests, but this statement (Statement 1.14) did not meet the criterion for consensus. There was strong agreement that physical function tests would be safe to perform without direct supervision (Statement 1.15; 86% agreement) but only low agreement that respondents would be willing to participate in a remote exercise program that was always (Statement 1.16; 50% agreement) or sometimes (Statement 1.17; 63% agreement) supervised, and moderate agreement that they would be willing to participate in an unsupervised exercise program (Statement 1.18; 72% agreement). There was also low or moderate agreement that respondents would be happy to participate in a remote exercise program in a group setting (Statement 1.19; 45% agreement) or individually (Statement 1.20; 77% agreement). Finally, low proportions of respondents agreed that they would be concerned about their privacy and security when using technology to participate in remote care (Statement 1.22; 21% agreement) or that remote physical function tests or exercise programs would be difficult to perform in their home (Statement 1.23; 16% agreement).

XSL•FO
RenderX

Figure 2. Representation of participants' level of agreement (rated on an 11-point Likert scale; 0-10) with each of 23 Round 1 Delphi statements. n=654 for statements 1.01-1.13, n=644 for statements 1.14-1.23. Participants with a response ≥7 were considered to have agreed with the statement. Statements with strong agreement (≥80%; black dotted line) were accepted. Statements with 70%-<80% agreement (grey dotted line) were considered to have moderate agreement. Statements with <70% agreement were considered to have low agreement.



XSL•FO
RenderX

Table 3. List of Round 1 Delphi statements with levels of agreement and outcome decisions.

Statement number	Statement	Agreement (%)[a]	Agreement rating[b]	Outcome
1.01	Having good physical function is important to the overall quality of life of older adults	99.5	Strong agreement	Accept
1.02	Having good physical function is important for activities involving moving around the community (eg, going shopping or to a restaurant or cafe, visiting your neighbors, friends and family, or the doctor)	99.4	Strong agreement	Accept
1.03	Having good physical function is important for participating in activities with family and friends (eg, playing with grandchildren)	97.9	Strong agreement	Accept
1.04	Having good physical function is important for participating in activities like work, household duties (eg, cooking, cleaning, and gardening), and volunteering	98.6	Strong agreement	Accept
1.05	Having good physical function is important for participating in hygiene activities (eg, showering, dressing, and using the toilet)	96.8	Strong agreement	Accept
1.06	Having good physical function is important for participating in exercise (eg, walking, swimming, dancing, golf, and other types of physical activity)	97.6	Strong agreement	Accept
1.07	It is possible to slow down or prevent poor physical function that occurs as we get older	86.2	Strong agreement	Accept
1.08	If someone already has poor physical function, it is possible to improve it	86.4	Strong agreement	Accept
1.09	If I was concerned about my physical function, I would discuss it with my health professional	80.3	Strong agreement	Accept
1.10	I would like access to information about how to test my physical function myself to determine if it is poor	78.4	Moderate agreement	Modify statement for Round 2
1.11	I would like access to information about things that I can do myself to improve my physical function	86.4	Strong agreement	Accept
1.12	Having better access to information on physical function would help me to have conversations about this with health professionals	69.0	Low agreement	Modify statement for Round 2
1.13	Having better access to information on physical function would help me to take care of my own physical function	80.9	Strong agreement	Accept
1.14	I would be willing to participate in remote tests of my physical function (eg, on a video call with a health professional, or by myself using written instructions and/or video demonstrations provided to me)	79.5	Moderate agreement	Modify statement for Round 2
1.15	I am confident that it would be safe for me to perform physical function tests at home without direct supervision by a health professional if I was provided with instructions (eg, written information or video demonstrations)	85.6	Strong agreement	Accept
1.16	I would be willing to participate in a remote exercise program to improve my physical function if it was ALWAYS supervised (eg, exercising while on a live video call with a health professional for all exercise sessions)	49.5	Low agreement	Modify statement for Round 2
1.17	I would be willing to participate in a remote exercise program if it was SOMETIMES supervised (eg, exercising on a live video call with a health professional for some exercise sessions, but exercising by myself unsupervised using instructions provided by the health professional for other sessions)	63.4	Low agreement	Modify statement for Round 2
1.18	I would be willing to participate in a remote exercise program if it was NOT supervised (eg, exercising by myself unsupervised using instructions provided by a health professional)	71.6	Moderate agreement	Modify statement for Round 2

XSL•FO
RenderX

Statement number	Statement	Agreement (%)[a]	Agreement rating[b]	Outcome
1.19	If I was to participate in a remote exercise program I would be happy to do so with a group (eg, exercising by myself at home but while on a video call with other people like me who are also exercising at home, with or without the supervision of a health professional)	45.3	Low agreement	Modify statement for Round 2
1.20	If I was to participate in a remote exercise program to improve my physical function, I would be happy to do so alone without other people like me involved in the exercise sessions (eg, exercising by myself at home with or without supervision by a health professional)	77.3	Moderate agreement	Modify statement for Round 2
1.21	I would be comfortable using technology (eg, computers, smartphones, and tablets) to participate in remote tests and treatments for my physical function	81.5	Strong agreement	Accept
1.22	I would be concerned about the privacy and security of my personal information when participating in remote tests and treatments for physical function using technology (eg, computer, smartphone, or tablet)	20.8	Low agreement	Reject
1.23	Remote physical function tests or exercise programs would be difficult to perform in my home (eg, because there is limited space)	15.5	Low agreement	Reject

[a]Proportion of participants who rated statement ≥7 out of 10.

[b]Statement classification based on the following criteria: Strong agreement (>80% of respondents rated statement ≥7 out of 10); Moderate agreement (70% to 80% of respondents rated statement ≥7 out of 10); Low agreement (<70% of respondents rated statement ≥7 out of 10).

Round 2

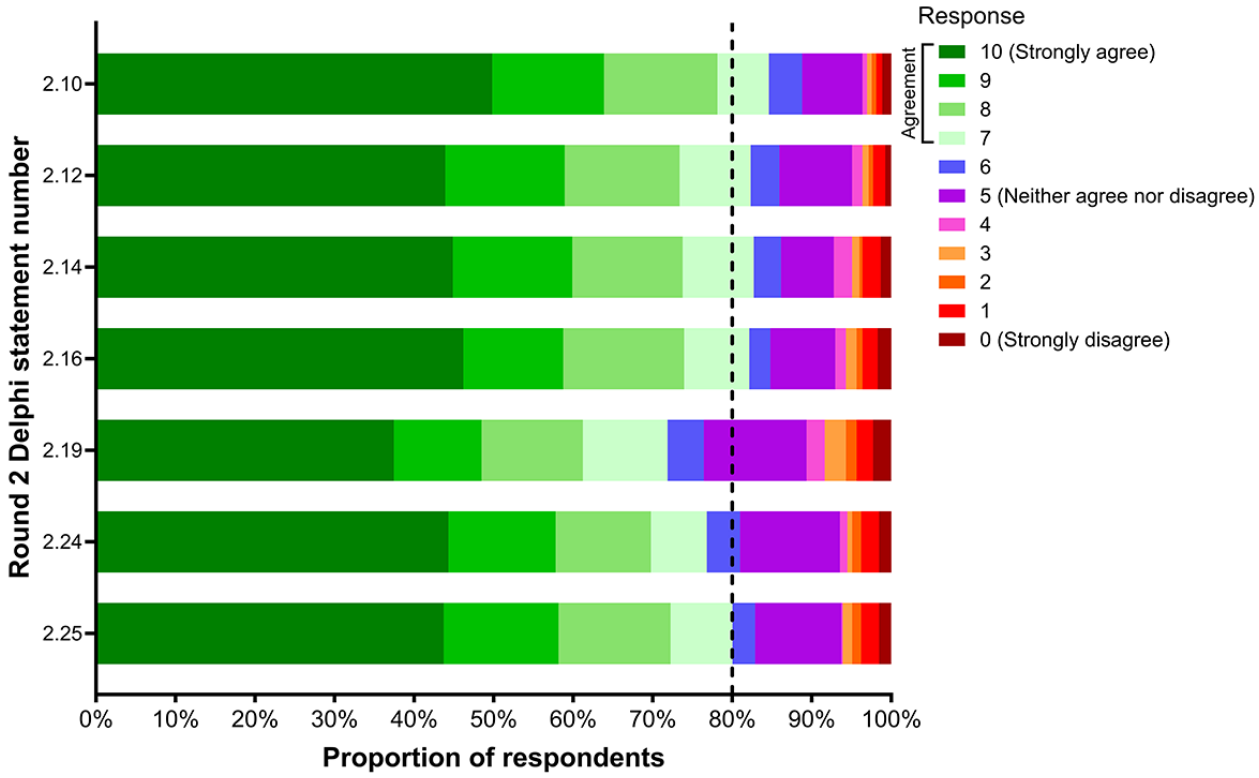
Following Round 1, 13 of 23 statements were accepted to have achieved consensus with strong agreement. It was determined that statements 1.22 and 1.23 had such low agreement that they should be rejected rather than modified and presented again in Round 2. Other statements that achieved moderate or low agreement (Statements 1.10, 1.12, 1.14, 1.16, 1.17, 1.18, 1.19, and 1.20) were modified. In brief, Statements 1.10, 1.12, and 1.14 were revised into Statements 2.10, 2.12, and 2.14, respectively. Statements 1.16, 1.17, and 1.18, which covered similar concepts regarding remote exercise supervision, were merged into a single revised statement (Statement 2.16). Statements 1.19 and 1.20 covered similar concepts regarding preferences for participating in remote exercise individually or in a group setting and were revised into a single statement (Statement 2.19). Finally, based on common themes identified in free-text responses to several statements across Round 1, 2 new statements were introduced for Round 2: Statement 2.24 explored the importance for consumers that remote physical function tests have been demonstrated to be safe and accurate, and Statement 2.25 explored the importance of having access to necessary information and resources to support participation in remote exercise programs. Multimedia Appendices 3 and 4 provide further details on the development of Round 2 Delphi statements.

Among participants who completed Round 1, 526 (80%) completed Round 2. No notable differences were observed between the Round 2 and Round 1 respondents, respectively, for age (mean = 69.3, SD 5.7 vs 69.0, SD 6.0 years) or SARC-F score (median 1.0, IQR 0.0-2.0 vs median 1.0, IQR 0.0-2.0). In Round 2, respondents were asked to rate their physical function compared with when they completed the Round 1 survey (mean interval between survey completions was 117.2, SD 21.5 days). The majority (68%) reported that their physical function was neither better nor worse, while 13% reported their physical function was somewhat or much better, and 19% reported that it was somewhat or much worse.

Figure 3 and Table 4 summarize participant responses to the Round 2 Delphi statements. Respondents strongly agreed that they would like access to instructions on how to test their own physical function and monitor changes (Statement 2.10; 85% agreement) and that having access to information on physical function would allow them to have more informed conversations with health care professionals (Statement 2.12; 82% agreement). Respondents also strongly agreed they would be willing to participate in a remote test of their physical function (Statement 2.14; 83% agreement) and to participate in a remote exercise program suited to their preferences at the time (Statement 2.16; 82% agreement). However, there was only moderate agreement that respondents would be willing to participate in remote exercise programs with a group or individually (Statement 2.19; 72% agreement) and that they would be more likely to participate in a remote physical function test if they were confident that the test was safe and accurate to perform alone (Statement 2.24; 77% agreement). Finally, there was strong agreement that respondents would be more likely to participate in a remote exercise program if they had access to necessary information and resources, including technology and exercise equipment (Statement 2.25; 80% agreement).

XSL•FO
RenderX

Figure 3. Representation of participants' level of agreement (rated on an 11-point Likert scale; 0-10) with each of 7 Round 2 Delphi statements (n=526). Participants with a response ≥7 were considered to have agreed with the statement. Statements with strong agreement (≥80%; black dotted line) were accepted.



XSL•FO
RenderX

Table 4. List of Round 2 Delphi statements with levels of agreement and outcome decisions.

Statement Number	Statement	Agreement (%)[a]	Agreement Rating[b]	Outcome
2.10	I would like access to simple and reliable instructions on how to test my physical function myself so that I can monitor how it changes over time	84.6	Strong agreement	Accept
2.12	If I felt I needed help to improve or maintain my physical function, having access to simple information about this (including advice on appropriate health professionals to discuss it with) would help me to have more informed conversations with health professionals about my physical function	82.3	Strong agreement	Accept
2.14	If I felt I needed help to improve or maintain my physical function, I would be willing to participate in a remote test (eg, supervised on a live video call with a health professional or unsupervised using printed instructions and/or video demonstrations provided to me)	82.7	Strong agreement	Accept
2.16	If I felt I needed help to improve or maintain my physical function, I would be willing to participate in a remote exercise program suited to my preferences at the time which may include exercise supervised by a health professional, and/or exercise led by myself	82.1	Strong agreement	Accept
2.19	If I felt I needed help to maintain or improve my physical function, I would be willing to participate in a remote exercise program suited to my preferences at the time which may include exercise performed by myself, and/or exercise performed with a group of people	71.9	Moderate agreement	Reject
2.24	I would be more likely to participate in a remote test of physical function if I was confident that the test was safe and accurate to perform by myself, and I had access to the necessary information and resources, including technology and equipment, to perform the test myself	76.8	Moderate agreement	Reject
2.25	I would be more likely to participate in a remote exercise program if I was confident that I had access to the necessary information and resources, including technology and exercise equipment, to exercise safely and effectively	80.0	Strong agreement	Accept

[a]Proportion of participants who rated statement ≥7 out of 10.

[b]Statement classification based on the following criteria: Strong agreement (>80% of respondents rated statement ≥7 out of 10); Moderate agreement (70% to 80% of respondents rated statement ≥7 out of 10); Low agreement (<70% of respondents rated statement ≥7 out of 10).

Summary of Delphi Outcomes and Recommendations

Five of seven Round 2 statements achieved consensus (ie, ≥80% agreement), and 2 (2.19 and 2.24) were rejected with moderate agreement. Thus, including the 13 statements accepted in Round 1, a total of 18 statements achieved consensus in this Delphi process and provided a basis for the 7 recommendations relevant to health care professionals, researchers, and policymakers developed by the RAMP Working Group and presented in Table 5.

Table 5. Recommendations for remote assessment and management of physical function in older adults.

Recommendation number	Recommendation	Supporting Delphi statements
1	Recognize that physical function is an important health priority for older adults	1.01, 1.02, 1.03, 1.04, 1.05, 1.06
2	Recognize that older adults are aware that it is possible to prevent, delay, and reverse declines in physical function	1.07, 1.08
3	Regularly engage older adults in discussions about their physical function	1.09, 2.12
4	Provide older adults with accessible and reliable information on how to monitor and maintain their physical function	1.11, 1.13, 2.10
5	For older adults who have concerns about their physical function, facilitate physical function assessments in-person and/or remotely considering feasibility and consumer preferences	2.14
6	Facilitate in-person and/or remote exercise programs for improving older adults' physical function considering feasibility, consumer preferences, and supervision requirements	2.16
7	Ensure that older adults participating in remote assessments or exercise for physical function have access to, or are provided with, appropriate support, technology, and/or equipment to perform assessments and exercises safely and effectively	1.15, 1.21, 2.25

XSL•FO
RenderX

Discussion

Principal Results

The World Health Organization's global strategy on digital health stresses the importance of person-centered approaches where end users are engaged in the design and development phases of digital health approaches [25]. This consumer-focused Delphi process demonstrated that older adults recognize physical function as a health priority and are generally accepting of remote assessment and management, including the use of digital health approaches. Based on these findings, the RAMP Working Group has developed 7 recommendations for researchers, health care professionals, and policymakers to guide remote assessment and management of physical function.

Comparison with Prior Work

Consistent with previous health care professional surveys showing that assessment and management of physical function are infrequently performed in clinical settings [7], our Round 1 survey demonstrated less than half of older adult respondents had received any health care professional-led assessment or intervention for physical function. Even fewer had participated in a remote physical function assessment (<5%) or intervention (<20%). This is despite almost 60% of respondents having participated in some form of remote care, likely due to increased general digital health use during and since the COVID-19 pandemic [11]. For those respondents who had participated in remote physical function care, several methods were reported, likely influenced by their appropriateness for the desired outcome. For example, video calls were most commonly used for remote physical function assessments whereas written documents were most commonly used for remote interventions. Respondents reported receptiveness to a range of different remote care approaches, and less than 5% stated they would not be willing to participate in any remote care for physical function. Key facilitators for remote care included convenience and flexibility in scheduling, while barriers included lack of guidance and motivation as reported previously [26]. Based on the Delphi statements that achieved consensus, seven recommendations for remote care of physical function in older adults were developed. The above barriers and facilitators should be considered when seeking to implement these recommendations in research, clinical care, and policy. Our findings are consistent with a recent position statement on telehealth policy for older adults, which highlights the need for dedicated policies to address common barriers to telehealth among older adults [27].

Recommendations for Researchers, Health Care Professionals, and Policymakers

Recognize That Physical Function Is an Important Health Priority for Older Adults

Over 97% of respondents strongly agreed physical function was integral to their overall quality of life and their ability to participate in activities with family and friends and in the community, to exercise, and to complete self-care tasks. This recommendation is in line with the World Report on Ageing and Health [28], which emphasizes that maximizing functional ability is a priority for older adults.

Recognize That Older Adults are Aware it Is Possible to Prevent, Delay, and Reverse Declines in Physical Function

Over 86% of respondents strongly agreed that it is possible to slow down or prevent poor physical function and that if someone already has poor physical function, it is possible to improve it. This knowledge may encourage older adults to engage in physical function care [29] and health care professionals should leverage it to promote and implement individually tailored strategies to maintain or improve physical function.

Regularly Engage Older Adults in Discussions About Their Physical Function

Respondents strongly agreed that they would raise concerns about their physical function with health care professionals and that access to information about physical function would help them initiate these conversations. Previous research has highlighted the importance that older adults place on mutual goal setting with health care professionals regarding their physical function [29], and that aligning care with patient priorities can lead to better health outcomes [30]. Health care professionals may require upskilling to ensure effective collaboration with consumers on identifying causes and symptoms of poor physical function and promoting benefits and strategies for maintaining physical function [8].

Provide Older Adults With Accessible and Reliable Information on How to Monitor and Maintain Their Physical Function

In addition to conversations with health care professionals, respondents reported a desire for information on how to independently monitor and maintain their physical function. Previous research has highlighted the importance of providing physical function information and advice to patients [30]. It is necessary to develop and promote appropriate resources that empower older adults to monitor and maintain their physical function.

For Older Adults Who Have Concerns About Their Physical Function, Facilitate Physical Function Assessments In-Person and/or Remotely Considering Feasibility and Consumer Preferences

There is a lack of data on the acceptability and appropriateness of remote physical function assessment [31]. In this study, however, respondents agreed they would be willing to participate in a remote physical function assessment if they were concerned about it. Remote assessment of physical function for older adults can be as reliable as face-to-face assessments [32,33] although further research is required to identify the most appropriate physical function tests, protocols, and communication platforms to support reliable remote physical function assessment, as well as facilitators and barriers to implementation in home and community settings.

Facilitate In-Person and/or Remote Exercise Programs for Improving Older Adults' Physical Function

XSL•FO
RenderX

Considering Feasibility, Consumer Preferences, and Supervision Requirements

Round 1 statements regarding remote exercise programs with set levels of supervision (ie, always, sometimes, or not supervised) or a set format (ie, individual or group-based) achieved only low to moderate agreement, reflecting the varied preferences for exercise among older adults [34]. Remotely delivered exercise programs with varying levels of synchronous and asynchronous exercise may be beneficial for community-dwelling older adults who do not have access to exercise facilities or prefer exercising alone and/or at home [35]. Remote programs can support novel and engaging exercise approaches such as integrating exercise into everyday activities [36], "exercise snacking" [12,13] and gamification [37], and may also incorporate other behavioral and educational interventions such as nutrition counseling [38].

Many medical professionals lack knowledge on exercise prescription [39]. Referral of patients to an exercise professional with experience in supporting older adults to exercise via remote care should be a consideration for clinicians who do not feel qualified to prescribe exercise. Upskilling exercise professionals in effectively delivering exercise via remote care and providing access to requisite resources and equipment is also important to build capacity for remote management of physical function.

Ensure That Older Adults Participating in Remote Assessments or Exercise for Physical Function Have Access to, or Are Provided With, Appropriate Support, Technology, and/or Equipment

Respondents believed it was safe and feasible for them to perform remote physical function assessments and exercise programs, and contrary to previous research [40], had relatively low agreement (<21%) that privacy was a concern. Nevertheless, ensuring remote care is administered by secure technologies and adheres to privacy laws is an important consideration for health care professionals to reduce barriers for those who do have concerns. Most of our respondents strongly agreed that they were comfortable using technology to participate in remote tests and treatments for their physical function. Technological literacy can be a barrier to participation in remote health assessments [26] and further research is required to identify approaches to overcoming technological barriers to remote care. Our results demonstrate that older adults agree that access to appropriate instructions, technology, and equipment would increase the likelihood of participation in remote programs. These findings suggest that older adults can successfully participate in remotely delivered exercise programs if appropriate support is provided. This can include standby technical assistance and technology orientation sessions, especially in the early stages of the program [41]. This support can potentially be integrated into existing funding models to support remote care which have increased internationally, particularly following the COVID-19 pandemic, although further development of policy and reimbursement mechanisms is needed for sustainable integration [42].

Recommendations for Future Studies

Based on the current results, it is recommended that future research, informed by collaboration with consumers, caregivers, health care professionals, and policymakers should focus on, but not be limited to, the following outcomes:

- Identifying and addressing barriers to accessibility of remote physical function assessment and management for older populations, particularly in culturally and linguistically diverse populations, those with socioeconomic disadvantage, those from low-middle-income countries, and/or those with low technology literacy or with limited access to technology
- Determining optimal approaches (including protocols and technologies) to delivering remote care for physical function to ensure validity and reliability of assessments and effectiveness and safety of interventions
- Exploring cost-effectiveness and implementation processes to embed remote care for physical function across varying levels of health care internationally
- Developing evidence-based guidelines and health promotion strategies for remote physical function assessment and management in older adults

Strengths and Limitations

Our modified Delphi study was co-developed with health care consumers and incorporated 2 rounds of iterative and anonymous questionnaires and controlled feedback to create consensus. Our study adhered to quality evaluation metrics for Delphi methodology [22] and included an international population of older adults. A high level of agreement (80%) was set a priori for acceptance of statements. There was low attrition of participants between Delphi rounds (<20%), suggesting the respondents were engaged and interested in sharing their views on physical function assessment and management.

Despite these strengths, there were limitations to our study. Given our survey was electronic, the participants required internet access and were likely technologically savvy, and thus a selection bias may be present. Respondents were from 15 generally high-income countries, with the majority residing in Australia and the United Kingdom, a large proportion (76%) were tertiary-educated, and all were English-speaking. It is not known if results are generalizable to those with lower socioeconomic status and/or non–English-speaking individuals. Our respondents also included a large proportion of women (76%), so it may not accurately reflect the views and experiences of older men. Similarly, less than 5% of respondents were aged 80 years or older, and less than 7% had a SARC-F score ≥4 (symptomatic of poor physical function) [24]. However, comparable SARC-F data in our study (median 1, IQR 0-2, proportion with score ≥4=7%) and other similar cohorts (median 0, IQR 1-2, proportion with score ≥4=6%-15%) [24] suggests the level of poor physical function in our sample is generally representative of community-dwelling older adults. Overall, further research investigating remote physical function care preferences in more diverse populations is required to understand the unique and common barriers, enablers, and needs, which would help to enable more widespread adoption of remote

XSL•FO
RenderX

methods. This should include research in men, non–English speakers, the oldest old, those with poorer health and/or digital literacy, those with poorer access to health care and/or technology, and those with poor physical function. Furthermore, effective strategies to address technology access, digital literacy, and support needs for these vulnerable populations need to be explored. This may include reducing access barriers through device and internet connectivity provision, improving digital health literacy through ongoing human coaching and troubleshooting support, and co-designing content, interfaces, and delivery models with underserved communities [43].

The current study captures respondents' preferences and intentions, but not their behaviors or health outcomes. Future research is therefore needed to evaluate engagement with remote physical function interventions among older adults in the real world, as well as the effectiveness of such interventions for relevant health outcomes. The study was also focused on consumer perspectives and does not capture the perspectives of other stakeholders involved in delivering care to older adults (eg, health care professionals and policymakers), which are critical to translating research into practice. To address this, the RAMP working group has recently completed a Delphi process investigating the views of experts involved in the care of older adults (manuscript under review).

Conclusions

This international consumer Delphi process achieved consensus on 18 Delphi statements, which were synthesized into 7 recommendations for health care professionals, researchers, and policymakers to inform remote assessment and management of physical function in older adults. Further research on the feasibility and integration of remote delivery of physical function assessment and exercise programs is required, and this should be co-designed with older adults and other relevant stakeholders. Furthermore, given the recommendations reflect the sample of predominantly highly educated and digitally literate volunteers from high-income countries, further research is also required to explore their generalizability to more diverse older adult populations.

Acknowledgments

The Remote Assessment and Management of Physical Function in Older Adults Working Group gratefully acknowledges the contributions of all consumer participants of this Delphi process.

Data Availability

The study data may be made available upon reasonable request to the corresponding author.

Funding

This research was supported by a Deakin University Institute for Physical Activity and Nutrition (IPAN) Seed Grant (2023). DS is supported by an NHMRC Investigator Grant (GNT1174886). MS is supported by an Emerging Leader Fellowship from the Future Health Research & Innovation Fund, Department of Health, Western Australia. CH is supported by the National Institute for Health and Care Research (NIHR) Newcastle Biomedical Research Centre (reference: NIHR203309). The views expressed in this publication are those of the authors and do not necessarily reflect the views of the National Institute for Health and Care Research or the Department of Health and Social Care. MAL is supported by a T1 Canadian Research Chair.

Conflicts of Interest

DS has received speaker and consulting fees and an educational grant from Abbott, Amgen and Pfizer. RMD has received speaker and consulting fees and an educational grant from Abbott. No other authors have conflicts of interest to declare.

Multimedia Appendix 1
Study design of the RAMP Consumer Delphi Process.
[PDF File (Adobe PDF File), 271 KB - aging_v9i1e75791_app1.pdf]

Multimedia Appendix 2
RAMP Delphi Survey: Round 1. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 364 KB - aging_v9i1e75791_app2.pdf]

Multimedia Appendix 3
RAMP Round 1 results as presented to consumers. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 3165 KB - aging_v9i1e75791_app3.pdf]

Multimedia Appendix 4
Summary of changes to Round 1 Delphi statements with moderate or low agreement.
[PDF File (Adobe PDF File), 152 KB - aging_v9i1e75791_app4.pdf]

XSL•FO
RenderX

Multimedia Appendix 5
RAMP Delphi Survey: Round 2. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 222 KB - aging_v9i1e75791_app5.pdf]

Multimedia Appendix 6
SARC-F scores for RAMP participants in Round 1. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 80 KB - aging_v9i1e75791_app6.pdf]

Multimedia Appendix 7
Delivery preferences and perceived positives and negatives of remote tests and treatments for physical function among RAMP participants in Round 1. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 112 KB - aging_v9i1e75791_app7.pdf]

References

1. Webber SC, Porter MM, Menec VH. Mobility in older adults: a comprehensive framework. Gerontologist 2010;50(4):443-450. [doi: 10.1093/geront/gnq013] [Medline: 20145017]
2. Ostchega Y, Harris TB, Hirsch R, Parsons VL, Kington R. The prevalence of functional limitations and disability in older persons in the US: data from the national health and nutrition examination survey III. J Am Geriatr Soc 2000;48(9):1132-1135. [doi: 10.1111/j.1532-5415.2000.tb04791.x] [Medline: 10983915]
3. Rosso AL, Taylor JA, Tabb LP, Michael YL. Mobility, disability, and social engagement in older adults. J Aging Health 2013;25(4):617-637 [FREE Full text] [doi: 10.1177/0898264313482489] [Medline: 23548944]
4. Hoffman JM, Ciol MA, Huynh M, Chan L. Estimating transition probabilities in mobility and total costs for medicare beneficiaries. Arch Phys Med Rehabil 2010;91(12):1849-1855 [FREE Full text] [doi: 10.1016/j.apmr.2010.08.010] [Medline: 21112425]
5. Offord NJ, Clegg A, Turner G, Dodds RM, Sayer AA, Witham MD. Current practice in the diagnosis and management of sarcopenia and frailty - results from a UK-wide survey. J Frailty Sarcopenia Falls 2019;4(3):71-77 [FREE Full text] [doi: 10.22540/JFSF-04-071] [Medline: 32300721]
6. Reijnierse EM, de van der Schueren MAE, Trappenburg MC, Doves M, Meskers CGM, Maier AB. Lack of knowledge and availability of diagnostic equipment could hinder the diagnosis of sarcopenia and its management. PLoS One 2017;12(10):e0185837 [FREE Full text] [doi: 10.1371/journal.pone.0185837] [Medline: 28968456]
7. Daly RM, Scott D, Kiss N, Tieland M, Baguley B, Fyfe JJ. Knowledge, awareness, behaviours, beliefs, attitudes, and perceptions of older Australians regarding muscle health and sarcopenia: a national survey. Arch Gerontol Geriatr 2025;135:105835 [FREE Full text] [doi: 10.1016/j.archger.2025.105835] [Medline: 40347780]
8. Izquierdo M, Merchant R, Morley J, Anker S, Aprahamian I, Arai H, et al. International exercise recommendations in older adults (ICFSR): expert consensus guidelines. J Nutr Health Aging 2021;25(7):824-853 [FREE Full text] [doi: 10.1007/s12603-021-1665-8] [Medline: 34409961]
9. Hurst C, Dismore L, Granic A, Tullo E, Noble JM, Hillman SJ, et al. Attitudes and barriers to resistance exercise training for older adults living with multiple long-term conditions, frailty, and a recent deterioration in health: qualitative findings from the lifestyle in later life - older people's medicine (LiLL-OPM) study. BMC Geriatr 2023;23(1):772 [FREE Full text] [doi: 10.1186/s12877-023-04461-5] [Medline: 38001414]
10. Nugent CD, Finlay D, Davies R, Donnelly M, Hallberg J, Black ND, et al. Remote healthcare monitoring and assessment. Technol Health Care 2011;19(4):295-306. [doi: 10.3233/THC-2011-0626] [Medline: 21849740]
11. Haimi M, Sergienko R. Adoption and use of telemedicine and digital health services among older adults in light of the COVID-19 pandemic: repeated cross-sectional analysis. JMIR Aging 2024;7:e52317 [FREE Full text] [doi: 10.2196/52317] [Medline: 38656768]
12. Jansons P, Dalla Via J, Daly R, Fyfe J, Gvozdenko E, Scott D. Delivery of home-based exercise interventions in older adults facilitated by amazon alexa: a 12-week feasibility trial. J Nutr Health Aging 2022;26(1):96-102 [FREE Full text] [doi: 10.1007/s12603-021-1717-0] [Medline: 35067710]
13. Fyfe JJ, Dalla Via J, Jansons P, Scott D, Daly RM. Feasibility and acceptability of a remotely delivered, home-based, pragmatic resistance 'exercise snacking' intervention in community-dwelling older adults: a pilot randomised controlled trial. BMC Geriatr 2022;22(1):521 [FREE Full text] [doi: 10.1186/s12877-022-03207-z] [Medline: 35751032]
14. Buckinx F, Aubertin-Leheudre M, Daoust R, Hegg S, Martel D, Martel-Thibault M, et al. Feasibility and acceptability of remote physical exercise programs to prevent mobility loss in pre-disabled older adults during isolation periods such as the COVID-19 pandemic. J Nutr Health Aging 2021;25(9):1106-1111 [FREE Full text] [doi: 10.1007/s12603-021-1688-1] [Medline: 34725669]

XSL•FO
RenderX

15. Brown RCC, Keating SE, Owen PJ, Jansons PS, McVicar J, Askew CD, et al. Client and clinician experiences and perspectives of exercise physiology services during the COVID-19 pandemic: qualitative study. J Med Internet Res 2023;25:e46370 [FREE Full text] [doi: 10.2196/46370] [Medline: 38127430]
16. Saaei F, Klappa SG. Rethinking telerehabilitation: attitudes of physical therapists and patients. J Patient Exp 2021;8:23743735211034335 [FREE Full text] [doi: 10.1177/23743735211034335] [Medline: 34377773]
17. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
18. Hsu CC, Sandford BA. The Delphi technique: making sense of consensus. Practical assessment, research, and evaluation 2007;12(1). [doi: 10.4018/978-1-4666-0074-4.ch011]
19. Zanker J, Sim M, Anderson K, Balogun S, Brennan-Olsen SL, Dent E, et al. Consensus guidelines for sarcopenia prevention, diagnosis and management in Australia and New Zealand. J Cachexia Sarcopenia Muscle 2023;14(1):142-156 [FREE Full text] [doi: 10.1002/jcsm.13115] [Medline: 36349684]
20. Zanker J, Sim M, Anderson K, Balogun S, Brennan-Olsen SL, Dent E, et al. The Australian and New Zealand society for sarcopenia and frailty research (ANZSSFR) sarcopenia diagnosis and management task force: findings from the consumer expert Delphi process. Australas J Ageing 2023;42(1):251-257 [FREE Full text] [doi: 10.1111/ajag.13164] [Medline: 36480154]
21. Barrington H, Young B, Williamson PR. Patient participation in Delphi surveys to develop core outcome sets: systematic review. BMJ Open 2021;11(9):e051066 [FREE Full text] [doi: 10.1136/bmjopen-2021-051066] [Medline: 34475183]
22. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: how to decide its appropriateness. World J Methodol 2021;11(4):116-129 [FREE Full text] [doi: 10.5662/wjm.v11.i4.116] [Medline: 34322364]
23. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 2014;67(4):401-409. [doi: 10.1016/j.jclinepi.2013.12.002] [Medline: 24581294]
24. Malmstrom TK, Miller DK, Simonsick EM, Ferrucci L, Morley JE. SARC-F: a symptom score to predict persons with sarcopenia at risk for poor functional outcomes. J Cachexia Sarcopenia Muscle 2016;7(1):28-36 [FREE Full text] [doi: 10.1002/jcsm.12048] [Medline: 27066316]
25. WHO. Global Strategy on Digital Health 2020-2025. Geneva: World Health Organization; 2021.
26. Jansons P, Fyfe J, Via JD, Daly RM, Gvozdenko E, Scott D. Barriers and enablers for older adults participating in a home-based pragmatic exercise program delivered and monitored by Amazon Alexa: a qualitative study. BMC Geriatr 2022;22(1):248 [FREE Full text] [doi: 10.1186/s12877-022-02963-2] [Medline: 35337284]
27. Holliday AM, Hashmi AZ, Okoli-Umeweni AO, Khan A, Jindal SK, Gaur S, et al. American geriatrics society position statement: telehealth policy for older adults. J Am Geriatr Soc 2025;73(12):3646-3654. [doi: 10.1111/jgs.70004] [Medline: 40650623]
28. Beard JR, Officer A, de Carvalho IA, Sadana R, Pot AM, Michel J, et al. The world report on ageing and health: a policy framework for healthy ageing. Lancet 2016;387(10033):2145-2154 [FREE Full text] [doi: 10.1016/S0140-6736(15)00516-4] [Medline: 26520231]
29. Chan E, Samsudin SA, Lim YJ. Older patients' perception of engagement in functional self-care during hospitalization: a qualitative study. Geriatr Nurs 2020;41(3):297-304. [doi: 10.1016/j.gerinurse.2019.11.009] [Medline: 31787364]
30. Ong M, Soh K, Saimon R, Wai M, Mortell M, Soh K. Fall prevention education to reduce fall risk among community-dwelling older persons: a systematic review. J Nurs Manag 2021;29(8):2674-2688 [FREE Full text] [doi: 10.1111/jonm.13434] [Medline: 34331491]
31. Heslop P, Hurst C, Sayer A, Witham M. Remote collection of physical performance measures for older people: a systematic review. Age Ageing 2023;52(1):2023 [FREE Full text] [doi: 10.1093/ageing/afac327] [Medline: 36721962]
32. Buckinx F, Rezoulat M, Lefranc C, Reginster J, Bruyere O. Comparing remote and face-to-face assessments of physical performance in older adults: a reliability study. Geriatr Nurs 2024;55:71-78. [doi: 10.1016/j.gerinurse.2023.11.004] [Medline: 37976558]
33. Peyrusqué E, Granet J, Pageaux B, Buckinx F, Aubertin-Leheudre M. Assessing physical performance in older adults during isolation or lockdown periods: web-based video conferencing as a solution. J Nutr Health Aging 2022;26(1):52-56 [FREE Full text] [doi: 10.1007/s12603-021-1699-y] [Medline: 35067703]
34. Cohen-Mansfield J, Marx M, Biddison J, Guralnik J. Socio-environmental exercise preferences among older adults. Prev Med 2004;38(6):804-811. [doi: 10.1016/j.ypmed.2004.01.007] [Medline: 15193902]
35. Aaltonen S, Waller K, Vähä-Ypyä H, Rinne J, Sievänen H, Silventoinen K, et al. Motives for physical activity in older men and women: a twin study using accelerometer-measured physical activity. Scand J Med Sci Sports 2020;30(8):1409-1422 [FREE Full text] [doi: 10.1111/sms.13673] [Medline: 32259351]
36. Clemson L, Fiatarone Singh MA, Bundy A, Cumming RG, Manollaras K, O'Loughlin P, et al. Integration of balance and strength training into daily life activity to reduce rate of falls in older people (the LiFE study): randomised parallel trial. BMJ 2012;345:e4547 [FREE Full text] [doi: 10.1136/bmj.e4547] [Medline: 22872695]

XSL•FO
RenderX

37. Martel D, Lauzé M, Agnoux A, Fruteau de Laclos L, Daoust R, Émond M, et al. Comparing the effects of a home-based exercise program using a gerontechnology to a community-based group exercise program on functional capacities in older adults after a minor injury. Exp Gerontol 2018;108:41-47. [doi: 10.1016/j.exger.2018.03.016] [Medline: 29577975]
38. Rodrigues IB, Wang E, Keller H, Thabane L, Ashe MC, Brien S, et al. The moveStrong program for promoting balance and functional strength training and adequate protein intake in pre-frail older adults: a pilot randomized controlled trial. PLoS One 2021;16(9):e0257742 [FREE Full text] [doi: 10.1371/journal.pone.0257742] [Medline: 34559837]
39. Abreu F, Rodrigues A, Baptista F. Low-volume resistance training: a feasible, cost-effective strategy for musculoskeletal frailty in older adults attending daycare centers. Front Sports Act Living 2025;7:1542188 [FREE Full text] [doi: 10.3389/fspor.2025.1542188] [Medline: 40264931]
40. Pool J, Akhlaghpour S, Fatehi F, Gray LC. Data privacy concerns and use of telehealth in the aged care context: an integrative review and research agenda. Int J Med Inform 2022;160:104707. [doi: 10.1016/j.ijmedinf.2022.104707] [Medline: 35131698]
41. Gell N, Hoffman E, Patel K. Technology support challenges and recommendations for adapting an evidence-based exercise program for remote delivery to older adults: exploratory mixed methods study. JMIR Aging 2021;4(4):e27645 [FREE Full text] [doi: 10.2196/27645] [Medline: 34889743]
42. van Kessel R, Srivastava D, Kyriopoulos I, Monti G, Novillo-Ortiz D, Milman R, et al. Digital health reimbursement strategies of 8 European countries and israel: scoping review and policy mapping. JMIR Mhealth Uhealth 2023;11:e49003. [doi: 10.2196/49003] [Medline: 37773610]
43. Oudbier SJ, Souget-Ruff SP, Chen BSJ, Ziesemer KA, Meij HJ, Smets EMA. Implementation barriers and facilitators of remote monitoring, remote consultation and digital care platforms through the eyes of healthcare professionals: a review of reviews. BMJ Open 2024;14(6):e075833 [FREE Full text] [doi: 10.1136/bmjopen-2023-075833] [Medline: 38858155]

Abbreviations

RAMP: Remote Assessment and Management of Physical Function in Older Adults

Edited by D Liu, DRPH, MS, FGSA; submitted 13.Apr.2025; peer-reviewed by N Misser, N Paramita, R Miranda; comments to author 20.Oct.2025; revised version received 02.Nov.2025; accepted 15.Dec.2025; published 06.Feb.2026.

Please cite as:
Dent E, Hurst C, Dalla Via J, Fyfe JJ, Jansons P, Hayes EJ, Skinner G, Sim M, Aubertin-Leheudre M, Britting S, Buckinx F, Connolly G, Dignam R, Giangregorio L, Jones JRA, Kelly P, Kob R, Morin SN, Nandakumar G, Orssatto LBR, Pearson M, Pinto D, Reijnierse EM, Said CM, Salem M, Tan VPS, Tran R, Zanker J, Daly RM, Scott D
Developing Consumer Consensus on Remote Assessment and Management of Physical Function in Older Adults (RAMP): International Modified Delphi Process
JMIR Aging 2026;9:e75791
URL: https://aging.jmir.org/2026/1/e75791
doi:10.2196/75791
PMID:

©Elsa Dent, Christopher Hurst, Jack Dalla Via, Jackson J Fyfe, Paul Jansons, Eleanor J Hayes, Gary Skinner, Marc Sim, Mylene Aubertin-Leheudre, Sabine Britting, Fanny Buckinx, Gavin Connolly, Ruth Dignam, Lora Giangregorio, Jennifer R A Jones, Pauline Kelly, Robert Kob, Suzanne N Morin, Girish Nandakumar, Lucas B R Orssatto, Maria Pearson, Daniel Pinto, Esmee M Reijnierse, Catherine M Said, Mohamed Salem, Vina PS Tan, Rosanna Tran, Jesse Zanker, Robin M Daly, David Scott. Originally published in JMIR Aging (https://aging.jmir.org), 06.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effects of a Virtual Reality–Based Natural Environment Intervention on Attention and Mood in Community-Dwelling Older Adults: Randomized Controlled Trial

Cheng-Chen Chou[1,2], PhD; Jong-Ni Lin[3], PhD; Chia-Pin Yu[4], PhD; Chi-Wen Chen[5], PhD; An-Yun Yeh[6], PhD; Yen-Kuang Lin[7], PhD

[1]Institute of Community Health Care, College of Nursing, National Yang Ming Chiao Tung University, Taipei, Taiwan

[2]Master Program in Transdisciplinary Long-Term Care and Management, National Yang Ming Chiao Tung University, Taipei, Taiwan

[3]Department of Post-Baccalaureate Nursing/ Department of Nursing, College of Nursing and Health Sciences, Da-Yeh University, Changhua, Taiwan

[4]School of Forestry and Resource Conservation, National Taiwan University, Taipei, Taiwan

[5]Department of Nursing, College of Nursing, National Yang Ming Chiao Tung University, Taipei, Taiwan

[6]Department of Nursing, College of Staten Island, The City University of New York, Staten Island, NY, United States

[7]Graduate Institute of Athletics and Coaching Science, National Taiwan Sport University, Taoyuan, Taiwan

Corresponding Author:
Cheng-Chen Chou, PhD
Institute of Community Health Care
College of Nursing
National Yang Ming Chiao Tung University
155, Section 2, Linong St
Beitou District
Taipei, 11221
Taiwan
Phone: 886 0228267000 ext 67176
Email: ccchou@nycu.edu.tw

Abstract

Background: Aging is a phenomenon accompanied by declines in attention and emotional well-being. Exposure to natural environments has been shown to support cognitive and psychological restoration; however, many older adults face barriers to accessing authentic natural environments. Virtual natural environments may offer an accessible alternative, but evidence from community-dwelling older adults remains limited.

Objective: This study examined the immediate effects of a single-session virtual or actual-nature exposure on attention, mood, and physiological indicators of stress among community-dwelling older adults compared with actual-nature and urban environments.

Methods: A randomized controlled trial with a 3-group, pre-post design was conducted among 120 community-dwelling adults aged 65 years and older in northern Taiwan. Participants were randomly assigned to 1 of 3 groups: a virtual-nature group, an actual-nature group, or an actual-urban environment group. Then, they completed a single 20-minute exposure session. Attentional performance, mood, and physiological indicators were assessed before and immediately after the intervention. Intervention effects were analyzed using generalized estimating equations, focusing on group-by-time interaction effects.

Results: Compared with the urban environment group, the virtual-nature group demonstrated a significantly greater improvement in attentional capacity, as measured by the digit span test (B=1.00, 95% CI 0.19-1.80; *P*=.02), and significant improvements in both positive affect (B=2.82, 95% CI 0.02-5.62; *P*=.048) and negative affect (B=–2.22, 95% CI –4.05 to –0.39; *P*=.02). The actual-nature group showed a significant improvement in positive affect compared with the urban group (B=3.92, 95% CI 1.65-6.19; *P*=.001) but did not demonstrate a significant improvement in attention. No significant group-by-time interaction effects were observed for physiological indicators. No adverse events were reported.

Conclusions: A single 20-minute virtual-nature exposure was associated with immediate improvements in attention and mood among community-dwelling older adults, whereas physiological effects were not detected. These findings suggest that virtual nature may serve as an accessible, short-term restorative strategy, particularly when access to real nature is limited. Larger,

XSL•FO
RenderX

multisession trials with longer follow-ups are needed to evaluate the sustained effects of virtual-nature interventions for mental well-being in older populations.

Trial Registration: ClinicalTrials.gov NCT05803460; https://clinicaltrials.gov/study/NCT05803460

(JMIR Aging 2026;9:e87861) doi:10.2196/87861

KEYWORDS

virtual reality; natural environment; community-dwelling older adults; attention; mood

Introduction

Population Aging and Mental Health Challenges

Population aging is a global phenomenon with profound implications for public health. As life expectancy continues to increase worldwide, a growing proportion of older adults experience age-related declines in cognitive functioning and emotional well-being [1,2]. Even among community-dwelling older adults who do not meet diagnostic criteria for cognitive impairment, subtle reductions in attention and mood are common and may precede more pronounced cognitive decline [3,4]. Consequently, identifying accessible and nonpharmacological strategies to support cognitive and emotional well-being has become a key priority for promoting healthy aging [5]. Although aging-related challenges are shared globally, contextual factors, such as urbanization, living environments, and health care infrastructure, substantially shape older adults' daily experiences [6], underscoring the need for feasible interventions that can be integrated into everyday living contexts [7].

Attention Decline and Attention Restoration Theory

Attention is a fundamental cognitive process that supports goal-directed behavior, information processing, and safe engagement in daily activities [8]. Age-related changes in attentional capacity, particularly those involving frontal brain networks, may compromise older adults' ability to sustain attention and manage competing demands [9,10]. Because attention is essential for overall cognitive functioning, attentional impairments can adversely affect daily activities and social behavior, thereby increasing the risk of falls, unsafe driving, and accidental injuries [11,12]. Natural environments are recognized for providing restorative experiences that promote cognitive, psychological, and physiological well-being, particularly by facilitating attentional recovery [13,14]. Attention restoration theory (ART) provides a theoretical framework that explains how exposure to natural environments may facilitate attentional recovery [15,16]. According to ART, natural environments evoke "soft fascination," allowing directed attention resources to rest and replenish [15]. Empirical studies have shown that exposure to natural settings is associated with improved attentional performance; however, findings remain mixed, particularly in studies using brief exposure durations or heterogeneous attentional measures [16,17]. These inconsistencies underscore the need for carefully designed trials that specifically examine attentional outcomes in older populations.

Mood, Stress, and Stress Reduction Theory

Mood, defined as an individual's emotional state, is a core component of psychological well-being and daily functioning [18]. In addition to attentional changes, older adults frequently experience emotional stress, anxiety, and depressive symptoms, which can negatively affect quality of life and may interact with cognitive decline [19,20]. In a community-based health survey, 36.9% of the adults aged more than 65 years exhibited anxiety symptoms, and 12.1% exhibited depressive symptoms [21]. Stress reduction theory suggests that natural environments elicit an affective response that promotes emotional calm and physiological relaxation through the regulation of the autonomic nervous system [22]. Empirical evidence from systematic reviews shows that exposure to natural or simulated natural environments is associated with reduced psychological distress, increased positive affect, and enhanced mood states [23,24].

Barriers to Real-Nature Exposure and Virtual Nature as an Alternative

Despite the potential benefits of real-world natural environments, many older adults face barriers to accessing nature, including physical limitations, mobility constraints, transportation difficulties, urban density, and public health restrictions [25,26]. These barriers have prompted an increasing interest in virtual-nature interventions, which use immersive technologies to simulate natural environments in accessible, indoor settings [23]. Virtual-nature interventions use immersive digital simulations to recreate natural environments and elicit restorative effects comparable to those of real-world nature [27]. Systematic reviews have demonstrated that virtual-nature exposure can improve attention, cognitive performance, mood, and stress regulation [27,28]. Emerging evidence further suggests that even a single session of simulated nature exposure, lasting between 6 and 20 minutes, may reduce stress and enhance positive emotions and attentional restoration [29-32]. However, most previous studies have focused on younger or middle-aged adults, and evidence regarding the effectiveness of virtual-nature interventions among older adults remains limited [27,28]. In addition, the immediate effects of a single-session virtual-nature exposure on attention, mood, and physiological indicators in older populations are not well established. Moreover, few randomized controlled trials have directly compared virtual nature, actual nature, and urban environments within the same experimental framework among community-dwelling older adults.

To address these gaps, this study aimed to examine the immediate effects of a single-session virtual-nature intervention on attention and mood among community-dwelling older adults in Taiwan. Specifically, we investigated whether virtual or actual-nature exposure would produce greater improvements than exposure to an urban environment. We hypothesized that a single session of virtual or actual-nature exposure would lead

XSL•FO
RenderX

to immediate improvements in attention, mood, and physiological outcomes.

Methods

Study Design

This study used a randomized controlled trial with a 3-group, pre-post design to examine the immediate effects of environmental exposure. The trial was registered at ClinicalTrials.gov (NCT05803460). Participants were randomly allocated to 1 of the 3 conditions: virtual nature, actual nature, or actual urban. Each participant completed a single 20-minute exposure session. Outcome measures were collected at baseline and immediately after the intervention to evaluate changes in attention, mood, and physiological responses.

Participants

Participants were recruited from 6 community care centers in Taipei, Taiwan, between April and October 2023. Eligible participants were aged 65 years or older, had intact global cognitive function as assessed by the Short Portable Mental Status Questionnaire (>8), and were able to communicate in Mandarin Chinese and complete study procedures. Individuals were excluded if they had a diagnosed psychiatric disorder or medical condition that could interfere with study outcomes, severe visual or hearing impairments that would impede participation, or active ocular or dermatologic infectious diseases. The required sample size was determined using G*Power statistical software. Given the repeated-measures design and the primary interest in detecting group-by-time interaction effects, a small-to-medium effect size (Cohen f=0.20) was assumed based on an F test for ANOVA, as no previous studies using a comparable virtual-nature intervention design were available [33]. The significance level was set at an α of .05, with statistical power of 0.90. On the basis of these parameters, a total sample size of 84 participants was considered sufficient to detect overall intervention effects. After accounting for an anticipated 20% attrition rate, the minimum required sample size was increased to 105 participants; ultimately, 120 participants were enrolled.

This randomized controlled trial used computer-generated block randomization with a block size of 6. The allocation sequence was generated by a research team member who was not involved in participant recruitment or assessment. As participants were enrolled, the corresponding sequence numbers were placed individually into opaque, sealed envelopes to ensure allocation concealment. A separate research team member opened each envelope sequentially to assign participants to 1 of the 3 groups: the virtual-nature group, the actual-nature group, or the actual-urban group, with 40 participants allocated to each group.

Intervention

Participants assigned to the virtual-nature group viewed a 20-minute immersive video of a natural environment filmed by the research team. This duration was selected based on previous evidence indicating that 10 to 20 minutes of virtual-nature exposure can improve attention and emotional outcomes in older adults [29,31]. The 360° panoramic video was recorded in a natural park in Taipei, a location commonly visited by older adults for its safe walking paths. Filming was conducted under stable weather conditions to minimize visual distractions and ensure ecological validity. The footage was captured using an Insta360 ONE R dual-lens camera, edited using standard audiovisual software, and structured into two 10-minute segments. In line with previous virtual-nature protocols for older adults, no interactive features were included to maintain a passive, restorative experience [34]. The video was delivered using a lightweight head-mounted display (HTC VIVE Flow). Before the session, participants received brief instructions encouraging them to observe in a relaxed manner. Each participant remained seated throughout the 20-minute exposure. For hygiene and safety, all equipment was disinfected after use, and disposable virtual reality (VR) face pads were provided.

Participants in the actual-nature group completed a 20-minute real-world exposure session at the same natural park featured in the virtual intervention. Upon arrival, participants rested for 10 to 15 minutes to acclimate. They were then seated in a safe, shaded location and instructed to observe their surroundings in a calm and relaxed manner.

Participants in the actual-urban group underwent a 20-minute observation session in a nearby urban area selected for safety and accessibility. As with the other groups, participants rested for 10 to 15 minutes before beginning the session and were seated during the exposure period, observing the environment in a relaxed state.

Outcomes

Overview

At baseline, participants completed a questionnaire assessing demographic characteristics (age, sex, years of education, and marital status); health-related variables (BMI, blood pressure, exercise habits, sleep quality, and comorbidities); and time spent in natural environments. Additionally, perceived restorativeness was assessed after the intervention using the Perceived Restorativeness Scale [35]. The primary outcomes were attention and mood. The secondary outcomes were physiological indicators of stress.

Attention

Attention was assessed using the digit span test (DS) and the trail making test (TMT), which capture multiple dimensions of attentional functioning, including attentional span, orienting, and sustained attention. The DS includes 2 subtests: digit span forward (DSF) and digit span backward (DSB). In both tasks, participants were asked to immediately repeat digit sequences read aloud by the assessor, with testing discontinued after 2 consecutive errors. A total DS score was calculated by summing DSF (range 0-16) and DSB (range 0-14), with higher scores indicating better attentional capacity. The Chinese versions of DSF and DSB demonstrate acceptable sensitivity (0.62 and 0.77) and specificity (0.65 and 0.78), respectively [36]. The TMT assessed orienting and sustained attention and consisted of 2 parts. In part A, participants connected numbered circles (1-25) in ascending order, and in part B, they alternated between numbers (1-13) and letters. Performance was scored based on completion time, with longer times indicating poorer attention.

XSL•FO
RenderX

The TMT has demonstrated strong reliability and validity as a measure of attentional processes [37].

Mood

Mood was assessed using the Positive and Negative Affect Schedule (PANAS), which evaluates 2 dimensions of affective states: positive affect (eg, enthusiasm and alertness) and negative affect (eg, distress and fear). The PANAS consists of 20 items, each rated on a 5-point Likert scale ranging from 1 ("very slightly or not at all") to 5 ("extremely"). Higher scores indicated greater intensity of the respective affective states. The PANAS has demonstrated high internal consistency, with Cronbach α coefficients ranging from 0.86 to 0.90 for positive affect and 0.84 to 0.87 for negative affect [38].

Physiological Indicators of Stress

Physiological parameters were assessed using a wrist-worn autonomic monitoring device (ANSWatch, Taiwan Scientific Corporation; approved by the Taiwan Ministry of Health and Welfare, medical device permit number 001525). The device measured heart rate (HR), systolic blood pressure (SBP), diastolic blood pressure (DBP), and indices of sympathetic nervous system and parasympathetic nervous system activity. HR variability (HRV) indices were derived from normal-to-normal RR intervals. High-frequency (HF) power was used as an indicator of parasympathetic activity, whereas low-frequency (LF) power reflects combined autonomic influences. The LF/HF ratio was reported as an index of autonomic balance [31]. To reduce potential confounding, participants were instructed to avoid stimulant beverages (eg, tea or coffee) for at least 2 hours before measurement. All assessments were conducted while participants were seated comfortably in a relaxed position. Each HRV measurement required approximately 6 minutes.

Procedure

Following approval from the institutional review board, community-dwelling older adults were recruited from 6 community health stations in Taipei City from April 2023 to October 2023. Station staff first identified and approached potential participants. Individuals who expressed interest were contacted by the research team for eligibility screening. After written informed consent was obtained, baseline assessments were conducted, including attention tests, mood questionnaires, and physiological measurements (HR, SBP, DBP, and HRV).

Participants were then exposed to 1 of the 3 assigned conditions (virtual nature, actual nature, or actual-urban environment) according to the randomized allocation sequence. A trained research assistant provided standardized instructions and supervised the 20-minute exposure session. Immediately after the intervention, participants completed the posttest assessments, which included physiological measurements, attention tests, and questionnaires. Participants were monitored for discomfort before the posttest assessment was initiated. The entire procedure required approximately 60 minutes per participant. All data were collected by a trained research assistant to ensure standardized administration of assessments and measurement reliability.

Ethical Considerations

The study protocol was reviewed and approved by the institutional review board at National Yang Ming Chiao Tung University (YM111067). Written informed consent was obtained from all participants after they received a full explanation of the study purpose, procedures, potential risks, and expected benefits. Participants were informed that their involvement was voluntary and that they could withdraw from the study at any time without consequence. All members of the research team completed the required human participant protection training and certification before data collection began. A small gift was provided as appreciation for participation.

Data Analysis

Data were analyzed using SPSS (version 25.0; IBM Corp). Descriptive statistics (means, SDs, frequencies, and percentages) were used to summarize participant characteristics and baseline variables. Intervention effects over time were analyzed using generalized estimating equations (GEEs) with an exchangeable working correlation structure and robust SEs. Group-by-time interaction effects were examined. In accordance with the CONSORT (Consolidated Standards of Reporting Trials) 2010 extension for multiarm randomized controlled trials (item 12a) [39], exploratory pairwise comparisons between each of the 2 intervention groups (virtual nature and actual nature) and the control group (urban environment) were conducted without adjustment for multiple comparisons. Baseline group differences and postintervention perceived restorativeness were examined using One-way ANOVA for continuous variables and the chi-square test for categorical variables. One-way ANOVA revealed a significant difference in postintervention perceived restorativeness scores among the 3 groups (P<.05). Post hoc analyses using the Tukey honestly significant difference test indicated significant between-group differences. A 2-tailed P value <.05 was considered statistically significant.

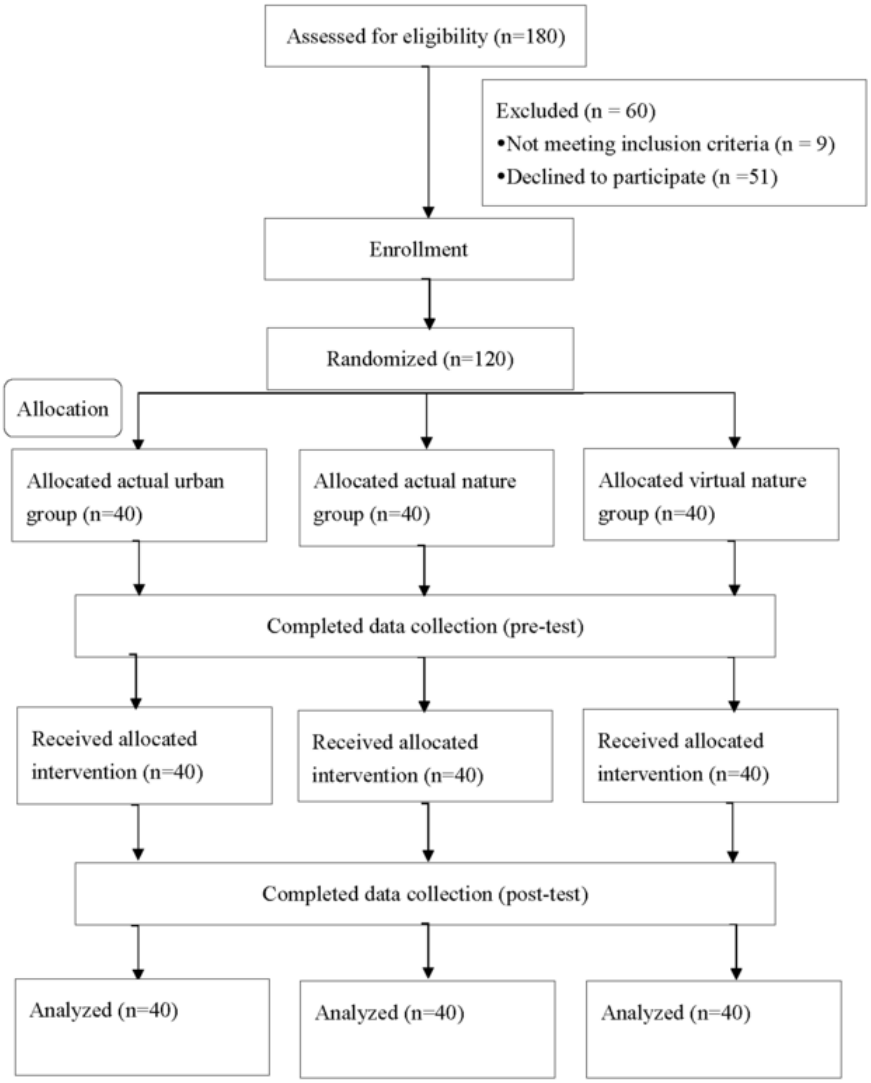
Results

Overview

A total of 180 invitation letters were distributed to community-dwelling older adults. Of these, 9 (5%) individuals did not meet the inclusion criteria, and 51 (28.3%) declined to participate. A total of 120 (66.7%) older adults agreed to participate, completed the study procedures, and were included in the final analysis (Figure 1). No adverse events or intervention-related discomfort was reported during the study.

XSL•FO
RenderX

Figure 1. Flow diagram of the study.



Sample Characteristics

A total of 120 community-dwelling older adults participated in this study. The mean age was 72.6 (SD 4.1) years, and most were female (n=97, 80.8%). Participants had an average of 10.2 (SD 2.8) years of formal education, and 79 (65.8%) were married. The mean BMI was 23.1 (SD 3.4) kg/m^2. Most participants reported regular exercise (n=103, 85.8%), and more than half (n=70, 58.3%) perceived their sleep quality as poor. The most common chronic conditions were hypertension (n=49, 40.8%), hyperlipidemia (n=50, 41.7%), and diabetes (n=33, 27.5%). Regarding exposure to natural environments, 83 (69.2%) participants reported spending less than 1 hour per month in natural settings. Group comparisons showed no significant baseline differences in demographic or health-related characteristics, including age, sex, education, marital status, BMI, blood pressure, exercise habits, number of chronic conditions, or monthly nature exposure. However, a significant between-group difference was observed in perceived restorativeness. Post hoc analyses indicated that the virtual-nature group (P=.01) and the actual-nature group (P=.001) both reported significantly higher restorativeness scores than the actual-urban group (Table 1).

XSL•FO
RenderX

Table 1. Participants' characteristics (N=120).

Characteristic	Total (N=120)	Virtual nature (n=40)	Actual nature (n=40)	Actual urban (n=40)	*P* value
Demographic variables					
Age (y), mean (SD)	72.6 (4.1)	72.2 (3.8)	72.0 (3.8)	73.5 (4.4)	.22
Female, n (%)	97 (80.8)	34 (85.0)	32 (80)	31 (77.5)	.69
Education (y), mean (SD)	10.2 (2.8)	10.6 (2.5)	9.9 (2.9)	10.1 (3)	.52
Marital status, n (%)					.77
Married	79 (65.8)	28 (70)	26 (65)	25 (62.5)	
Single, divorced, or separated	41 (34.2)	12 (30)	14 (35)	15 (37.5)	
Health-related variables, mean (SD)					
BMI (kg/m^2)	23.1 (3.4)	23.5 (3.4)	22.8 (3.2)	23.1 (3.6)	.67
Systolic blood pressure	131.2 (16.9)	130.2 (17.9)	131.3 (14.2)	132.1 (18.5)	.89
Diastolic blood pressure	84.8 (9.6)	83.5 (8.8)	86.7 (9.8)	84.2 (10)	.30
Having an exercise habit, n (%)	103 (85.8)	34 (85)	32 (80)	37 (92.5)	.27
Sleep quality, n (%)					.97
Good	50 (41.7)	17 (42.5)	16 (40)	17 (42.5)	
Poor	70 (58.3)	23 (57.5)	24 (60)	23 (57.5)	
Number of comorbidities, mean (SD)	1.4 (1.1)	1.2 (0.9)	1.6 (1.2)	1.4 (1.1)	.28
Hypertension, n (%)	49 (40.8)	15 (37.5)	14 (35)	20 (50)	.34
Diabetes, n (%)	33 (27.5)	11 (27.5)	9 (22.5)	13 (32.5)	.61
Hyperlipidemia, n (%)	50 (41.7)	14 (35)	16 (40)	20 (50)	.38
Nature-related variables: exposure to nature (monthly), n (%)					.48
≤1 h	83 (69.2)	30 (75)	25 (62.5)	28 (70)	
>1 h	37 (30.8)	10 (25)	15 (37.5)	12 (30)	
Perceived restorativeness after intervention	19.8 (3.1)	20.3 (2.9)	20.8 (2.6)	18.3 (3.3)	.001[a]

[a]Post hoc analysis of between-group comparisons using the Tukey honestly significant difference method: the virtual natural group>actual-urban group (*P*=.01); actual-nature group>actual-urban group (*P*=.001).

Attention

GEE analyses revealed a significant overall group-by-time interaction for total DS scores (Wald χ^2_2=6.28; *P*=.04). Specifically, the pre-post improvement in attentional capacity was significantly greater in the virtual-nature group than in the actual-urban group (B=1.00, 95% CI 0.19-1.80; *P*=.02), indicating a greater increase in DS performance following exposure to virtual nature. The corresponding contrast between the actual-nature and actual-urban groups was not statistically significant. No significant overall group-by-time interaction effects were observed for performance on the TMT Part A (Wald χ^2_2=2.07; *P*=.36) or TMT Part B (Wald χ^2_2=1.90; *P*=.39) between the virtual-nature and actual-urban groups. Additionally, neither the virtual-nature group nor the actual-nature group showed significant improvements in attentional performance compared with the actual-urban group following the intervention (Table 2).

XSL•FO
RenderX

Table 2. Group means and effects of interaction for attention and mood (N=120).

Variables and group	Overall group×time Wald χ^2 (*df*)[a]	Overall group×time *P* value	Before test, mean (SD)	After test, mean (SD)	B (95% CI)	Wald χ^2 (*df*)	*P* value
Digit span test							
Virtual nature	6.29 (2)	.04	24.4 (3.7)	25.8 (4.0)	1.00 (0.19 to 1.80)	5.88 (1)	.02
Actual nature	__[b]	—	25.4 (3.4)	26.5 (3.1)	0.57 (–0.03 to 0.93)	3.30 (1)	.12
Actual urban	—	—	25.6 (3.6)	26.0 (3.4)	—	—	—
TMT[c] part A							
Virtual nature	2.07 (2)	.36	40.8 (20.7)	37.8 (15.4)	1.85 (–2.01 to 5.71)	0.87 (1)	.35
Actual nature	—	—	33.1 (15.6)	30.7 (12.4)	2.35 (–0.91 to 5.61)	1.98 (1)	.16
Actual urban	—	—	39.9 (16.7)	35.1 (14.8)	—	—	—
TMT part B							
Virtual nature	1.90 (2)	.39	83.9 (47.3)	65.9 (30.2)	–4.17 (–13.30 to 4.95)	0.80 (1)	.37
Actual nature	—	—	68.8 (34.3)	57.3 (27.2)	2.42 (–5.13 to 9.98)	0.39 (1)	.53
Actual urban	—	—	79.4 (46.3)	65.5 (37.1)	—	—	—
PANAS[d] (positive)							
Virtual nature	11.46 (2)	.003	35.3 (6.8)	35.6 (7.5)	2.82 (0.02 to 5.62)	3.91 (1)	.048
Actual nature	—	—	33.1 (7.6)	34.2 (7.8)	3.92 (1.65 to 6.19)	11.46 (1)	.001
Actual urban	—	—	33.9 (6.9)	31.5 (8.4)	—	—	—
PANAS (negative)							
Virtual nature	6.05 (2)	.049	13.9 (4.0)	11.6 (2.8)	–2.22 (–4.05 to –0.39)	2.87 (1)	.02
Actual nature	—	—	14.5 (5.8)	12.7 (4.2)	–1.67 (–3.60 to 0.26)	2.87 (1)	.09
Actual urban	—	—	13.7 (4.4)	13.6 (5.7)	—	—	—

[a]The generalized estimating equation analysis was conducted using the actual-urban group as the reference group; an omnibus Wald χ^2_2 test was first conducted for each outcome to evaluate the overall group-by-time interaction. Wald χ^2_1 tests were subsequently used to estimate individual group-by-time contrasts (virtual nature vs actual urban and actual nature vs actual urban).

[b]Not applicable.

[c]TMT: trail making test.

[d]PANAS: Positive and Negative Affect Schedule.

Mood

Significant overall group-by-time interaction effects were observed for both positive affect (Wald χ^2_2=11.46; *P*=.003) and negative affect (Wald χ^2_2=6.05; *P*=.049). Compared with the urban environment group, the virtual-nature group demonstrated significantly greater improvements in both positive affect (B=2.82, 95% CI 0.02-5.62; *P*=.048) and negative affect (B=–2.22, 95% CI –4.05 to –0.39; *P*=.02). These findings indicate that exposure to virtual nature produced more pronounced mood-enhancing effects than the urban condition. In addition, the actual-nature group showed a significantly greater improvement in positive affect compared with the urban environment group (B=3.92, 95% CI 1.65-6.19; *P*=.001), whereas changes in negative affect did not reach statistical significance.

Physiological Variables

GEEs analyses indicated that no significant overall group-by-time interaction effects were observed for any physiological indicators, including HR, SBP, DBP, LF, HF, or the LF/HF ratio (all omnibus Wald χ^2_2 tests; *P*>.05). Consistent with these findings, pairwise group-by-time contrasts showed no significant differences in pre-post changes between the virtual-nature and actual-urban groups or between the actual-nature and actual-urban groups (all Wald χ^2_1 tests; *P*>.05; Table 3).

XSL•FO
RenderX

Table 3. Group means and effects of interaction for physiological indicators of stress (N=120).

Variables and group	Overall group×time Wald χ^2 (*df*)[a]	Overall group×time *P* value	Before test mean (SD)	After the test, mean (SD)	B (95% CI)	Wald χ^2 (*df*)	*P* value
Heart rate							
Virtual nature	3.29 (2)	.19	77.2 (10.5)	73.2 (8.8)	–1.23 (–3.38 to 0.92)	1.24 (1)	.26
Actual nature	—[b]	—	77.1 (9.7)	75.6 (9.2)	1.17 (–1.31 to 3.67)	0.86 (1)	.36
Actual urban	—	—	77.9 (9.4)	75.6 (9.0)	—	—	—
Systolic blood pressure							
Virtual nature	2.82 (2)	.24	130.2 (17.9)	125.6 (16.0)	–0.97 (–8.46 to 6.50)	0.06 (1)	.80
Actual nature	—	—	131.3 (14.2)	132.5 (17.9)	4.92 (–3.29 to 13.14)	1.37 (1)	.24
Actual urban	—	—	132.1 (18.5)	128.2 (19.4)	—	—	—
Diastolic blood pressure							
Virtual nature	0.57 (2)	.75	83.6 (8.8)	81.6 (8.5)	0.20 (–3.25 to 3.65)	0.01 (1)	.91
Actual nature	—	—	86.7 (9.8)	85.6 (8.8)	1.31 (–2.21 to 4.83)	0.52 (1)	.47
Actual urban	—	—	84.2 (10.0)	81.9 (9.4)	—	—	—
LF[c]							
Virtual nature	1.43 (2)	.49	49.8 (20.7)	45.6 (16.6)	–4.76 (–13.06 to 3.52)	1.26 (1)	.26
Actual nature	—	—	46.0 (17.8)	45.5 (18.6)	–0.90 (–9.49 to 7.69)	0.04 (1)	.84
Actual urban	—	—	45.3 (17.4)	45.6 (19.4)	—	—	—
HF[d]							
Virtual nature	1.58 (2)	.45	50.1 (20.7)	54.4 (16.6)	4.75 (–3.53 to 13.05)	1.26 (1)	.26
Actual nature	—	—	54.9 (17.8)	53.8 (18.6)	0.21 (–8.56 to 8.99)	0.002 (1)	.96
Actual urban	—	—	54.6 (17.4)	54.3 (19.4)	—	—	—
LF/HF							
Virtual nature	4.42 (2)	.11	1.4 (1.1)	1.0 (0.7)	–0.45 (–0.93 to 0.03)	3.35 (1)	.07
Actual nature	—	—	1.0 (0.7)	1.1 (0.9)	–0.03 (–0.52 to 0.45)	0.02 (1)	.89
Actual urban	—	—	1.0 (0.8)	1.1 (0.9)	—	—	—

[a]The generalized estimating equation test was conducted using the actual-urban group as a reference; an omnibus Wald χ^2_2 test was first conducted for each outcome to evaluate the overall group-by-time interaction. Wald χ^2_1 tests were subsequently used to estimate individual group-by-time contrasts (virtual nature vs actual urban and actual nature vs actual urban).

[b]Not applicable.

[b]LF: low-frequency power.

[d]HF: high-frequency power.

Discussion

Principal Findings

This randomized controlled trial examined the immediate effects of a single-session virtual-nature exposure on attention, mood, and physiological indicators among community-dwelling older adults. The preliminary findings suggest that, compared to an urban environment, a single 20-minute exposure to virtual nature is associated with greater short-term improvements in attentional capacity and mood. Actual-nature exposure was associated with improved positive affect but did not produce significant attentional benefits. No significant physiological changes were observed across groups. These findings suggest that virtual nature may provide immediate cognitive and emotional benefits in older adults, particularly when access to real-world nature is limited. However, the observed effects were task and domain specific and should be interpreted cautiously within the context of a brief, single-session intervention.

Participants in the virtual-nature group demonstrated a statistically significant improvement in attentional span compared with those in the urban group. This result differed from several previous studies that reported nonsignificant effects of virtual-nature exposure on attention [31,40]. Differences in study design, attention assessments, sample characteristics, intervention duration, and limited statistical power may explain these inconsistent findings. Previous studies often relied on a single attention measure, such as the Sustained Attention to Response Test, which may not adequately capture the

XSL•FO
RenderX

multifaceted nature of attention across environmental contexts [31]. Because distinct environments may engage different attentional components, the use of a single tool may limit sensitivity to change. This study used a multidimensional assessment approach, including the DS and TMT, to evaluate selective attention, sustained attention, and attentional span. This comprehensive measurement strategy may have enhanced the ability to detect subtle yet meaningful changes in attention following environmental exposure [14].

Improvements in attention were observed only in the DS and not in the TMT. This task-specific pattern may reflect differences in cognitive demands and sensitivity to short-term restorative effects. The DS primarily assesses attentional capacity and working memory [36], which may be more responsive to brief environmental restoration. In contrast, the TMT involves additional components, such as visual scanning, psychomotor speed, and executive control [37], which may require longer or repeated interventions to demonstrate measurable change. It is also noteworthy that the virtual-nature group had slightly lower baseline DS scores than the other groups, raising the possibility of a ceiling effect in the comparison groups [41]. Although the analyses focused on group-by-time interaction effects, baseline differences in performance may have influenced responsiveness to the intervention. Practice effects may also contribute to the absence of significant between-group differences in TMT performance [42]. Together, these findings highlight the importance of considering task characteristics and baseline performance when interpreting attention outcomes in environmental intervention studies.

Another explanation for the divergent findings may relate to differences in intervention duration and participant characteristics. This study used a single 20-minute session, whereas previous studies typically administered shorter 10-minute exposures in middle-aged and older adults [31] and younger adult samples [43]. Variations in exposure duration may influence the extent of attentional restoration achieved. In addition, because participants in this study were all aged 65 years or older, they may have been more responsive to restorative stimulation, as age-related declines in attentional capacity could increase susceptibility to the benefits of natural environments [44].

Both virtual and actual-nature exposure were associated with improvements in positive affect, with virtual nature additionally demonstrating reductions in negative affect. These results are consistent with previous research involving mixed middle-aged and older populations [31,45] and extend existing evidence by demonstrating similar effects in community-dwelling adults aged 65 years and older. These findings are broadly consistent with stress reduction theory, which posits that natural environments elicit affective responses that promote emotional calm and psychological well-being [22]. The emotional benefits observed following a single session suggest that even brief exposure to nature-related stimuli may have a positive influence on mood in older adults. The concurrent improvement in attention and mood in the virtual-nature group also aligns with ART [15,16], as emotional relaxation may facilitate the recovery of attentional resources. However, the absence of attentional benefits in the actual-nature group suggests that the relationship between environmental exposure and cognitive outcomes may depend on contextual factors, such as environmental familiarity, attentional engagement, and individual expectations.

A unique strength of this study was the use of locally familiar natural settings, which may have enhanced both the ecological relevance and acceptability of the intervention. Although most participants had no previous experience with VR, previous research suggests that Taiwanese older adults generally perceive VR as useful, enjoyable, and easy to use for supporting active aging [46]. Such positive attitudes may have facilitated greater immersion and engagement with the virtual natural environment.

Contrary to the initial hypothesis, no significant changes in physiological indicators of stress, including HRV and blood pressure, were observed over time between the virtual-nature and urban groups or between the actual-nature and urban groups. Thus, the findings are consistent with previous research reporting similar null effects [31,45]. Several factors may account for these results. First, the single 20-minute intervention may have been insufficient to elicit detectable changes in the autonomic nervous system, particularly in older adults. Second, most participants (102/120, 85%) had no previous experience with VR, and novelty-induced arousal associated with first-time VR exposure may have transiently increased sympathetic activation [47,48], potentially masking relaxation-related physiological responses.

Limitations

Several limitations should be acknowledged. First, the sample comprised predominantly female, community-dwelling older adults recruited from urban areas in northern Taiwan. This demographic profile may limit the generalizability of the findings to male older adults, individuals living in rural settings, those with mobility limitations or limited familiarity with technology, and older adults from different cultural contexts. Future research should explicitly examine the feasibility and effectiveness of the intervention in more diverse and vulnerable older populations. Second, this study was designed as a preliminary, exploratory investigation of immediate effects. Guided by feasibility considerations and real-world conditions, this study used a single-session intervention design and included an urban environment as a comparison group. Future research is recommended to conduct larger-scale trials with multiple intervention sessions and to incorporate additional no-exposure control conditions to further validate the findings. Third, most (102/120, 85%) participants in this study had no previous experience with VR. As a result, the observed effects may partially reflect VR-related factors, such as novelty or immersion, rather than the natural content alone, particularly with respect to physiological outcomes (eg, HRV). In addition, this study included only a single VR condition and did not incorporate formal measures of presence or engagement, which limits the interpretation of the underlying mechanisms of the observed effects. Future research should use multiple VR control conditions and incorporate validated measures of presence or immersion to strengthen the evidential support for the validity of the intervention. Finally, the study sample consisted of older adults who were able to participate in community activities,

XSL•FO
RenderX

excluding those with limited mobility. Future studies should include mobility-impaired older adults to further evaluate the feasibility, accessibility, and effectiveness of virtual-nature interventions in broader older populations.

Conclusions

The findings of this study provide evidence that a single 20-minute virtual-nature exposure can yield immediate improvements in attention and mood among community-dwelling older adults, whereas physiological effects were not detected. Virtual nature may represent a feasible and accessible short-term restorative strategy, particularly for older adults facing barriers to accessing real-world natural environments. Future research using multisession interventions, larger and more diverse samples, additional control conditions, and longitudinal follow-up is needed to clarify sustained effects and underlying mechanisms.

Acknowledgments

The authors thank all the participants who took part in this study.

Funding

This work was supported by the National Science and Technology Council, Taiwan (grant MOST 111-2635-B-A49-002).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

CCC and CPY conceptualized and designed the study. CCC and JNL coordinated and conducted data collection. CCC, CWC, AYY, and YKL performed data analysis and interpretation. CCC and JNL drafted the manuscript. CCC, JNL, CPY, CWC, AYY, and YKL critically reviewed, revised, and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 94 KB - aging_v9i1e87861_app1.pdf]

References

1. Khan HT, Addo KM, Findlay H. Public Health Challenges and Responses to the Growing Ageing Populations. Public Health Chall 2024 Jul;3(3):e213. [doi: 10.1002/puh2.213] [Medline: 40496520]
2. Wang Z, Xu M, Guo X, Zuo Y. Prevalence and associated factors of mental health among older adults in Guangxi, China: insights from depression, anxiety and cognitive function assessments. Gen Psychiatr 2025 Sep 14;38(5):e102220 [FREE Full text] [doi: 10.1136/gpsych-2025-102220] [Medline: 40959770]
3. McDonough IM, Wood MM, Miller WSJ. A review on the trajectory of attentional mechanisms in aging and the Alzheimer's disease continuum through the attention network test. Yale J Biol Med 2019 Mar;92(1):37-51 [FREE Full text] [Medline: 30923472]
4. Hosseini FS, Sharifi N, Jamali S. Correlation anxiety, stress, and depression with perceived social support among the elderly: a cross-sectional study in Iran. Ageing Int 2021 Mar;46(1):108-114. [doi: 10.1007/s12126-020-09376-9]
5. de Sá-Caputo D, Seixas A, Taiar R, Van der Zee EA, Bernardo-Filho M. Editorial: Non-pharmacological interventions in healthy and pathological aging: facts and perspectives. Front Aging Neurosci 2023 Apr 18;15:1191281. [doi: 10.3389/fnagi.2023.1191281] [Medline: 37143692]
6. Zhang S, Wu W, Xiao Z, Wu S, Zhao Q, Ding D, et al. Creating livable cities for healthy ageing: cognitive health in older adults and their 15-minute walkable neighbourhoods. Cities 2023 Jun;137:104312. [doi: 10.1016/j.cities.2023.104312]
7. Horgan S, Prorok J, Ellis K, Mullaly L, Cassidy KL, Seitz D, et al. Optimizing older adult mental health in support of healthy ageing: a pluralistic framework to inform transformative change across community and healthcare domains. Int J Environ Res Public Health 2024 May 23;21(6):664 [FREE Full text] [doi: 10.3390/ijerph21060664] [Medline: 38928911]
8. Posner MI, Rothbart MK, Ghassemzadeh H. Restoring attention networks. Yale J Biol Med 2019 Mar;92(1):139-143 [FREE Full text] [Medline: 30923481]
9. Liu Y, Zeng Z, Huang S, Shang P, Lv Z, Wang Y, et al. Brain activation during working memory task in amnestic mild cognitive impairment patients and its association with memory and attention. J Alzheimers Dis 2023;91(2):863-875. [doi: 10.3233/JAD-220815] [Medline: 36502326]

XSL•FO
RenderX

10. Logue SF, Gould TJ. The neural and genetic basis of executive function: attention, cognitive flexibility, and response inhibition. Pharmacol Biochem Behav 2014 Aug;123:45-54 [FREE Full text] [doi: 10.1016/j.pbb.2013.08.007] [Medline: 23978501]
11. Ramzaoui H, Faure S, Spotorno S. Alzheimer's disease, visual search, and instrumental activities of daily living: a review and a new perspective on attention and eye movements. J Alzheimers Dis 2018;66(3):901-925. [doi: 10.3233/JAD-180043] [Medline: 30400086]
12. Nobre AC, Kastner S. The Oxford Handbook of Attention. Oxford, UK: Oxford University Press; 2014.
13. Gidlow CJ, Jones MV, Hurst G, Masterson D, Clark-Carter D, Tarvainen MP, et al. Where to put your best foot forward: psycho-physiological responses to walking in natural and urban environments. J Environ Psychol 2016 Mar;45:22-29. [doi: 10.1016/j.jenvp.2015.11.003]
14. Jung M, Jonides J, Northouse L, Berman MG, Koelling TM, Pressler SJ. Randomized crossover study of the natural restorative environment intervention to improve attention and mood in heart failure. J Cardiovasc Nurs 2017 Sep;32(5):464-479 [FREE Full text] [doi: 10.1097/JCN.0000000000000368] [Medline: 27755225]
15. Kaplan S. The restorative benefits of nature: toward an integrative framework. J Environ Psychol 1995 Sep;15(3):169-182. [doi: 10.1016/0272-4944(95)90001-2]
16. Ohly H, White MP, Wheeler BW, Bethel A, Ukoumunne OC, Nikolaou V, et al. Attention Restoration Theory: a systematic review of the attention restoration potential of exposure to natural environments. J Toxicol Environ Health B Crit Rev 2016 Sep 26;19(7):305-343. [doi: 10.1080/10937404.2016.1196155] [Medline: 27668460]
17. Bell CN, St George R, Honan C, Bell LJ, Jolly AT, Matthews A. The relationship between nature exposures and attention restoration, as moderated by exposure duration: a systematic review and meta-analysis. J Environ Psychol 2025 Jun;104:102632. [doi: 10.1016/j.jenvp.2025.102632]
18. Brown LJ, Astell AJ. Assessing mood in older adults: a conceptual review of methods and approaches. Int Psychogeriatr 2012 Aug;24(8):1197-1206 [FREE Full text] [doi: 10.1017/S1041610212000075] [Medline: 22340813]
19. Manca R, Venneri A. Minority stress and psychosocial influences on cognitive performance in sexual minority older adults. Innov Aging 2023 Sep;7(8):igad110. [doi: 10.1093/geroni/igad110] [Medline: 37868765]
20. Park HJ, Kim YR, Choi J, Lee SH, Kang HJ, Lee HH. The impact of depression, stress, and self-esteem on quality of life among older adults in South Korea. Pak J Med Sci 2024 Mar;40(4):601-605 [FREE Full text] [doi: 10.12669/pjms.40.4.8912] [Medline: 38545010]
21. Lin YH, Chen HC, Hsu NW, Tsao HM, Lo SS, Chou P. Depression and anxiety correlate differentially with dissatisfaction with healthcare services in older adults: the Yilan Study, Taiwan. J Formos Med Assoc 2022 Aug;121(8):1506-1514 [FREE Full text] [doi: 10.1016/j.jfma.2021.10.021] [Medline: 34782197]
22. Ulrich RS, Simons RF, Losito BD, Fiorito E, Miles MA, Zelson M. Stress recovery during exposure to natural and urban environments. J Environ Psychol 1991 Sep;11(3):201-230. [doi: 10.1016/s0272-4944(05)80184-7]
23. Li H, Ding Y, Zhao B, Xu Y, Wei W. Effects of immersion in a simulated natural environment on stress reduction and emotional arousal: a systematic review and meta-analysis. Front Psychol 2022 Jan 9;13:1058177. [doi: 10.3389/fpsyg.2022.1058177] [Medline: 36698558]
24. Paredes-Céspedes DM, Vélez N, Parada-López A, Toloza-Pérez YG, Téllez EM, Portilla C, et al. The effects of nature exposure therapies on stress, depression, and anxiety levels: a systematic review. Eur J Investig Health Psychol Educ 2024 Mar 7;14(3):609-622 [FREE Full text] [doi: 10.3390/ejihpe14030040] [Medline: 38534901]
25. Patel A, Kolt G, Keogh J, Schofield G. The Green Prescription and older adults: what do general practitioners see as barriers? J Prim Health Care 2012 Dec 1;4(4):320-327. [Medline: 23205382]
26. Batsis JA, Daniel K, Eckstrom E, Goldlist K, Kusz H, Lane D, et al. Promoting healthy aging during COVID-19. J Am Geriatr Soc 2021 Mar;69(3):572-580. [doi: 10.1111/jgs.17035] [Medline: 33470421]
27. Browning MH, Shipley N, McAnirlin O, Becker D, Yu CP, Hartig T, et al. An actual natural setting improves mood better than its virtual counterpart: a meta-analysis of experimental data. Front Psychol 2020 Sep;11:2200 [FREE Full text] [doi: 10.3389/fpsyg.2020.02200] [Medline: 33101104]
28. Browning MH, Saeidi-Rizi F, McAnirlin O, Yoon H, Pei Y. The role of methodological choices in the effects of experimental exposure to simulated natural landscapes on human health and cognitive performance: a systematic review. Environ Behav 2021 Aug;53(7):687-731. [doi: 10.1177/0013916520906481]
29. Appel L, Appel E, Bogler O, Wiseman M, Cohen L, Ein N, et al. Older adults with cognitive and/or physical impairments can benefit from immersive virtual reality experiences: a feasibility study. Front Med (Lausanne) 2020 Jan 15;6:329. [doi: 10.3389/fmed.2019.00329] [Medline: 32010701]
30. Gerber SM, Jeitziner MM, Wyss P, Chesham A, Urwyler P, Müri RM, et al. Visuo-acoustic stimulation that helps you to relax: a virtual reality setup for patients in the intensive care unit. Sci Rep 2017 Oct 16;7(1):13228 [FREE Full text] [doi: 10.1038/s41598-017-13153-1] [Medline: 29038450]
31. Yu CP, Lee HY, Lu WH, Huang YC, Browning MHEM. Restorative effects of virtual natural settings on middle-aged and elderly adults. Urban For Urban Green 2020 Dec;56:126863. [doi: 10.1016/j.ufug.2020.126863]
32. Yu CP, Lee HY, Luo XY. The effect of virtual reality forest and urban environments on physiological and psychological responses. Urban For Urban Green 2018 Oct;35:106-114. [doi: 10.1016/j.ufug.2018.08.013]

XSL•FO
RenderX

33. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
34. Nukarinen T, Rantala J, Korpela K, Browning MHEM, Istance HO, Surakka V, et al. Measures and modalities in restorative virtual natural environments: an integrative narrative review. Comput Hum Behav 2022 Jan;126:107008. [doi: 10.1016/j.chb.2021.107008]
35. Berto R. Exposure to restorative environments helps restore attentional capacity. J Environ Psychol 2005 Sep;25(3):249-259. [doi: 10.1016/j.jenvp.2005.07.001]
36. Leung JL, Lee GT, Lam YH, Chan RC, Wu JY. The use of the Digit Span Test in screening for cognitive impairment in acute medical inpatients. Int Psychogeriatr 2011 Dec;23(10):1569-1574. [doi: 10.1017/s1041610211000792]
37. Wei M, Shi J, Li T, Ni J, Zhang X, Li Y, et al. Diagnostic accuracy of the Chinese version of the trail-making test for screening cognitive impairment. J Am Geriatr Soc 2018 Jan;66(1):92-99. [doi: 10.1111/jgs.15135] [Medline: 29135021]
38. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988 Jun;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
39. Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of multi-arm parallel-group randomized trials: extension of the CONSORT 2010 statement. JAMA 2019 Apr 23;321(16):1610-1620. [doi: 10.1001/jama.2019.3087] [Medline: 31012939]
40. Cassarino M, Tuohy IC, Setti A. Sometimes nature doesn't work: absence of attention restoration in older adults exposed to environmental scenes. Exp Aging Res 2019 Jul;45(4):372-385. [doi: 10.1080/0361073X.2019.1627497] [Medline: 31173559]
41. Andrade C. The ceiling effect, the floor effect, and the importance of active and placebo control arms in randomized controlled trials of an investigational drug. Indian J Psychol Med 2021 Jul;43(4):360-361 [FREE Full text] [doi: 10.1177/02537176211021280] [Medline: 34385732]
42. Kopp B, Rösser N, Wessel K. Psychometric characteristics and practice effects of the Brunswick Trail Making Test. Percept Mot Skills 2008 Dec;107(3):707-733. [doi: 10.2466/pms.107.3.707-733] [Medline: 19235402]
43. Cassarino M, Maisto M, Esposito Y, Guerrero D, Chan JS, Setti A. Testing attention restoration in a virtual reality driving simulator. Front Psychol 2019 Feb 11;10:250 [FREE Full text] [doi: 10.3389/fpsyg.2019.00250] [Medline: 30804862]
44. Zanto TP, Gazzaley A. Aging of the frontal lobe. Handb Clin Neurol 2019;163:369-389. [doi: 10.1016/B978-0-12-804281-6.00020-3] [Medline: 31590742]
45. Yu CP, Lin CM, Tsai MJ, Tsai YC, Chen CY. Effects of short forest bathing program on autonomic nervous system activity and mood states in middle-aged and elderly individuals. Int J Environ Res Public Health 2017 Aug 9;14(8):897 [FREE Full text] [doi: 10.3390/ijerph14080897] [Medline: 28792445]
46. Syed-Abdul S, Malwade S, Nursetyo AA, Sood M, Bhatia M, Barsasella D, et al. Virtual reality among the elderly: a usefulness and acceptance study from Taiwan. BMC Geriatr 2019 Aug 19;19(1):223. [doi: 10.1186/s12877-019-1218-8] [Medline: 31426766]
47. Zimmer HZ, Richter F. Novelty detection and orienting: effects on skin conductance and heart rate. Psychol Res 2023 Jun;87(4):1101-1113 [FREE Full text] [doi: 10.1007/s00426-022-01735-2] [Medline: 36107250]
48. Verastegui-Tena L, van Trijp H, Piqueras-Fiszman B. Heart rate and skin conductance responses to taste, taste novelty, and the (dis)confirmation of expectations. Food Qual Prefer 2018 Apr;65:1-9. [doi: 10.1016/j.foodqual.2017.12.012]

Abbreviations

ART: attention restoration theory
CONSORT: Consolidated Standards of Reporting Trials
DBP: diastolic blood pressure
DS: digit span test
DSB: digit span backward
DSF: digit span forward
GEE: generalized estimating equation
HF: high frequency
HR: heart rate
HRV: heart rate variability
LF: low frequency
PANAS: Positive and Negative Affect Schedule
SBP: systolic blood pressure
TMT: trail making test
VR: virtual reality

XSL•FO
RenderX

Edited by Y Jiang; submitted 15.Nov.2025; peer-reviewed by M Tusconi, TB Xu, L Zhang; comments to author 16.Dec.2025; accepted 16.Jan.2026; published 25.Feb.2026.

Please cite as:
Chou CC, Lin JN, Yu CP, Chen CW, Yeh AY, Lin YK
Effects of a Virtual Reality–Based Natural Environment Intervention on Attention and Mood in Community-Dwelling Older Adults: Randomized Controlled Trial
JMIR Aging 2026;9:e87861
URL: https://aging.jmir.org/2026/1/e87861
doi:10.2196/87861
PMID:

©Cheng-Chen Chou, Jong-Ni Lin, Chia-Pin Yu, Chi-Wen Chen, An-Yun Yeh, Yen-Kuang Lin. Originally published in JMIR Aging (https://aging.jmir.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Use of Generative AI Videos for Health Self-Management of Older Adults: Mixed Methods Study

Ting Liu[1], MFA; Patrick Pang[1], PhD; Yiming Taclis Luo[1], MC; Dana McKay[2], PhD; George Buchanan[2], PhD; Shanton Chang[3], PhD

[1]Faculty of Applied Sciences, Macao Polytechnic University, Macao, China

[2]School of Computing Technologies, Royal Melbourne Institute of Technology University, Melbourne, Australia

[3]School of Computing and Information Systems, The University of Melbourne, Melbourne, Australia

Corresponding Author:
Patrick Pang, PhD
Faculty of Applied Sciences
Macao Polytechnic University
Rua de Luís Gonzaga Gomes, Macao SAR
Macao, 999078
China
Phone: 86 85993815
Email: mail@patrickpang.net

Abstract

Background: Aging is a pressing global issue, and older adults need to build up their knowledge to manage their health. Insufficient self-efficacy and low acceptance of technology hinder their ability to use emerging technologies for self-management.

Objective: This study explored the potential of generative artificial intelligence (GenAI) videos in the health management of older adults.

Methods: We developed a video-based GenAI prototype, AIHealthV. This study used a mixed methods approach, enrolling 20 older adult participants (aged 60 to 80 years) in 3 rounds of iterative workshops. Data collection included pre- and postquestionnaires, in-depth interviews, prompt text records, and the generated video content. Qualitative data were analyzed using the 6-stage reflexivity thematic analysis method by Braun and Clarke, and quantitative data were analyzed using the Wilcoxon signed-rank test.

Results: The findings revealed that GenAI videos can enhance older adults' self-efficacy and technology acceptance, reduce cognitive load, and simultaneously meet their health management needs. The multimodal content generated by GenAI makes health information more comprehensible and thus improves the accessibility of health knowledge. Following the workshop, the interaction between older adults and AIHealthV exhibits a trend of exploration, adaptation, and verification.

Conclusions: Despite ethical, privacy, and usability concerns, AIHealthV has been proven to be beneficial in improving the health management capabilities of older adults. This paper also provides practical insights for developing artificial intelligence health tools tailored to older adults as GenAI tools continue to prevail.

(JMIR Aging 2026;9:e88005) doi:10.2196/88005

KEYWORDS

generative AI; GenAI videos; older adults; health self-management; mobile phone

Introduction

With the acceleration of global aging, the World Health Organization projects that the worldwide population aged 65 years and older will increase to 1.5 billion by 2050 [1]. Physiological declines associated with aging, including deteriorating physical functions, sensory impairments, and memory loss, alongside this increase, have increased older adults' needs for self-health management, particularly in disease prevention, health monitoring, and chronic disease management [1]. Technology offers opportunities to support self-management; however, research has found that older adults encounter significant difficulties in accessing printed or online health information, particularly in technologically complex and information-overloaded environments [2]. Technology-enabled precision health information, for example, on-demand generation

XSL•FO
RenderX

of health educational content, can improve older adults' health management and decision-making abilities [3]. The rapid development of artificial intelligence (AI) sheds light on the generation of tailor-made information for older adults' self-health management.

Older adults are gradually emerging as a significant user group of digital health technologies, which enable them to access health-related information and resources on time, thereby facilitating better health management and chronic disease control [4]. Accessing health knowledge through digital technologies not only enhances older adults' health management capabilities but also strengthens their sense of technological control [5]. However, their persistent reliance on traditional health knowledge acquisition channels (eg, printed materials, doctor visits, or face-to-face classes) also limits the potential of digital technologies in this domain [6]. Studies indicate that some seniors use AI voice assistants (eg, Amazon Echo and Google Home) to access health information [7]. In contrast, others rely on AI conversational agents and robotic systems for health-related queries [8]. Despite their potential, text- or image-based health AI applications often present interaction challenges for older users due to limited comprehension abilities or complex interfaces [9]. Complicated operations and navigation structures further hinder adoption [2]. Additionally, concerns about data privacy and security are prevalent among older adults when using AI [10], with technological distrust potentially reducing acceptance [11].

The capabilities of generative artificial intelligence (GenAI) to create on-demand informational material in a variety of formats, including text, video, or images, offer new opportunities for tailored digital education [12]. The emergence of GenAI has opened new avenues for efficient health knowledge dissemination and precision personalized education [13]. Furthermore, GenAI's capacity for dynamic content modification and optimization enhances the quality and user adaptability of health knowledge outputs through iterative refinement [14]. AI-generated health videos offer high flexibility and adaptability, enabling the creation of personalized video content based on user backgrounds, such as educational level and health needs, such as disease prevention or treatment decision-making [15].

GenAI is revolutionizing the creation and dissemination of health information, particularly in producing high-quality, dynamic content. Personalized health videos generated by GenAI have been preliminarily applied in patient education and health promotion, such as disease-specific tutorials [16]. GenAI technologies can dynamically generate personalized, easily comprehensible health information, particularly through video, a visually rich medium that compensates for the interpretive challenges of text-based content [17]. These videos dynamically tailor information to user needs, outperforming static content in engagement. Videos as a form of educational materials can improve adults' self-efficacy and reflect their critical thinking skills [18], but this is not necessarily true for older adults. Beyond dynamic generation, GenAI's capacity for scenario-specific personalization enhances its utility in health education [19]. Furthermore, generative models integrated with sentiment analysis (eg, detecting user satisfaction or confusion) can swiftly identify older adults' preferences for video design and functionality, optimizing information delivery [20]. GenAI's capacity for dynamic content modification and optimization enhances the quality and user adaptability of health knowledge outputs through iterative refinement [14]. By combining GenAI with user-driven optimization, health communication strategies for older adults can be refined more effectively [21].

While technology offers many opportunities, declines in information processing speed, sensory functions, and memory capacity create numerous barriers for older adults in accessing and comprehending health information via digital technologies [2]. Additionally, research reveals that older adults exhibit deeper fears regarding data security and potential privacy breaches, which further hinder their access to and use of emerging internet-based digital health management tools [22]. Technophobia and heightened concerns about usability and privacy issues represent significant barriers to the adoption of digital health among older adults [12]. These factors contribute to the challenges of promoting the use of AI technologies in general.

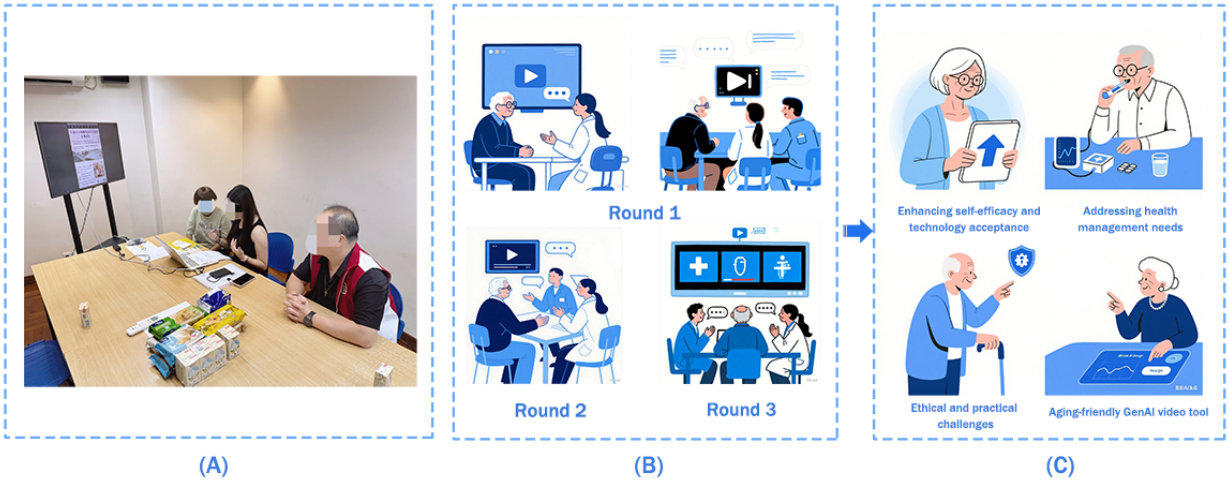
Self-efficacy plays a fundamental role in older adults' use of digital health technologies. According to the social cognitive theory by Bandura [23], an individual's belief in their ability to perform specific tasks influences their willingness to try and persist with new behaviors. In the context of technology adoption, Kim [24] found that older adults with higher technological self-efficacy showed greater persistence when facing initial usage difficulties. The development of technological self-efficacy in older adults follows unique patterns. Unlike younger users who typically acquire digital skills through independent exploration, older adults usually benefit more from organized, guided learning experiences [25]. Schroeder et al [26] emphasized that successful skill-building interventions for this population should include progressive challenge advancement, immediate feedback mechanisms, and opportunities for mastery experiences, which have been proven to enhance self-efficacy [27]. This self-efficacy then becomes a key determinant of technology acceptance. The perceived ease of use (PEOU) influences adoption decisions, especially for older adults who may view technological interfaces as complex [28]. Research by Malureanu et al [29] shows that enhanced self-efficacy directly improves perceived usefulness (PU) and ease of use, the 2 core elements of technology acceptance. This relationship creates a positive feedback loop: as older adults gain confidence through successful experiences, their willingness to adopt more advanced features increases [30]. The interaction between self-efficacy and acceptance has important practical implications. GenAI offers the potential for high self-efficacy, through simple interaction, for older adults, a potential that merits further exploration.

Static output modes in video generation tools neglect user feedback, failing to adapt interactively for maximal information retention. While this approach has theoretical benefits for older adults, these benefits warrant empirical analysis as in this paper. In this paper, we use "GenAI" to refer specifically to its use in creating health promotion videos. The usefulness of this application in supporting health education for older adults is currently underexplored. Figure 1 is a summary of this paper:

XSL•FO
RenderX

Figure 1A shows that workshops were implemented through the following process to obtain requirements from the participants. Prompts and health videos were generated collaboratively between participants and researchers. Figure 1B shows that 3 rounds of videos were generated in a session. Videos were regenerated, discussed, and evaluated in each round. Figure 1C shows key research findings.

Figure 1. Summary of this paper. GenAI: generative artificial intelligence.



To systematically explore the potential of GenAI videos in supporting older adults' health management, this study addresses three key research questions: (1) How can AI-generated videos create an impact on older adults' health management? (2) What is the role of AI-generated videos in the health management processes of older adults? (3) What are the issues and concerns faced by older adults when using GenAI technologies?

Methods

Ethical Considerations

This study was approved by the Human Ethics Committee of Macao Polytechnic University (HEA006-FCA-2025). All procedures followed the guidelines of the principles of the World Medical Association Declaration of Helsinki. All participants provided written explicit informed consent before participation. All data were deidentified. Before the data collection began, the consent form was obtained through an offline paper-based method, and participants were informed about the voluntary nature of this study and their right to withdraw at any time without penalty or loss of benefits. All participant data and visual materials (including figures, video samples, and prototype interface screenshots) in this manuscript have been fully deidentified to eliminate any identifiable personal information, strictly adhering to privacy and confidentiality protection requirements. To ensure the accuracy and safety of generated content, community workers who worked at the community center where this study was conducted were invited to the workshops. They conducted manual reviews of GenAI-generated videos, focusing on verifying the accuracy of health information and mitigating the risk of misleading information. Specifically, these reviews occurred at two critical checkpoints: (1) immediately after the initial script and plot generation to flag potential medical inaccuracies, and (2) after the final video rendering to ensure visual-audio alignment and content safety. All videos required confirmation by community workers to ensure that they did not violate any scientific medical knowledge before they were presented to participants. Additionally, before the workshops, participants were explicitly informed that "AI-generated content should be used in conjunction with professional advice," and the consent form emphasized the limitations of GenAI and the potential of inaccurate information. Participants voluntarily signed up for the workshops and chose a suitable timeslot. They did not receive any additional compensation, but snacks and drinks were provided in the workshop.

Prototype Development

We developed a prototype of a GenAI video tool tailored to older adults, named AIHealthV. This prototype allowed us to customize GenAI for better meeting older adults' needs for health management. For example, AIHealthV integrated voice input and keyword suggestions to reduce the barrier for content creation, as many older adults were reluctant to type; video outputs supported adjustments, including playback speed and voice selection, as well as subtitle fine-tuning with adjustable font sizes. These usability considerations were considered crucial for carrying out our study successfully with the older adult cohort.

For the technical implementation of this prototype, we integrated different software libraries as they have separate functions. The video generation application was built using Python (version 3.12; Python Software Foundation; for the backend) and React (version 18.2; Meta Platforms, Inc; for the frontend), which interconnects with SiliconFlow AI services for image and video generation. The default video length was set to 3-5 minutes, with output parameters optimized for high-contrast mobile viewing to suit the visual needs of older users. The prototype also featured a prompt generator embedded with preloaded health templates (eg, hypertension management and sleep improvement), which leveraged the GPT-4o (OpenAI) model for script generation.

The prompt template followed the structure with content of the role, background, health issue, and constraint, for example, "I

XSL•FO
RenderX

am 70 years old with diabetes and want to learn about a sugar-controlled healthy diet, focusing on [specific constraint].” In this way, the video plots and transcripts generated by the prototype were more health-related and accurate. Furthermore, to make the prototype more trustworthy and encourage older adults to share their health concerns, we used a friendly, human-like voice for feedback with a simplified and conversational interface instead of a technical one. These human-centric design elements help reduce the anxiety that older adults might feel when using the new technology, making them more willing to “talk” to the AI and disclose their real health needs [31].

The user interface design adopts a step-by-step guidance process: (1) a multimodal interactive input area supporting voice and text input, which is used to generate video plots; (2) the output of video plots which are used for video generations, and the content can be modified for fine-tuning video output; (3) a video preview area including multiple alternative videos; (4) features for adjusting elements such as voice, speed, volume, and subtitle size; and (5) generic functions such as settings, saving, and sharing. Figure 2 demonstrates the design of AIHealthV.

Figure 2. The prototype for generating health-related videos with GenAI. GenAI: generative artificial intelligence.



Participant Recruitment

With the assistance of community partners, we recruited 24 older adults aged 60 years and older from nearby communities through offline posters to participate in an iterative workshop. The inclusion criteria used were that participants were 60 years of age or older and had the basic operational ability to use smart devices such as smartphones and tablets. Of the 24 recruited participants, 4 were excluded from the final analysis: 3 withdrew before the pretest due to sudden scheduling conflicts, and 1 failed to complete the full iterative workshop rounds due to physical fatigue. Ultimately, results from 20 participants were included in this study. The purpose of this sample selection is to consciously focus on different population characteristics, such as the different levels of education, living conditions, and health needs (such as high blood pressure, joint health, and sleep disorder) of participants, to maximize the reliability of the information. All participants expressed motivation and a need to use GenAI and health videos for better self-health management. Participants had a range of health conditions (chronic disease, good health, and mild illness) and lived in a variety of residential situations (eg, living alone, with family, or in a nursing facility). In addition, community workers also participated in this study. They worked in community-based health, had a basic understanding of personal health management, and had experience assisting older adults with technology tools. It is the role of community workers to professionally review the content generated by GenAI to effectively circumvent the risk of AI generating erroneous or inaccurate medical information, thereby ensuring the safety and reliability of all discussed information.

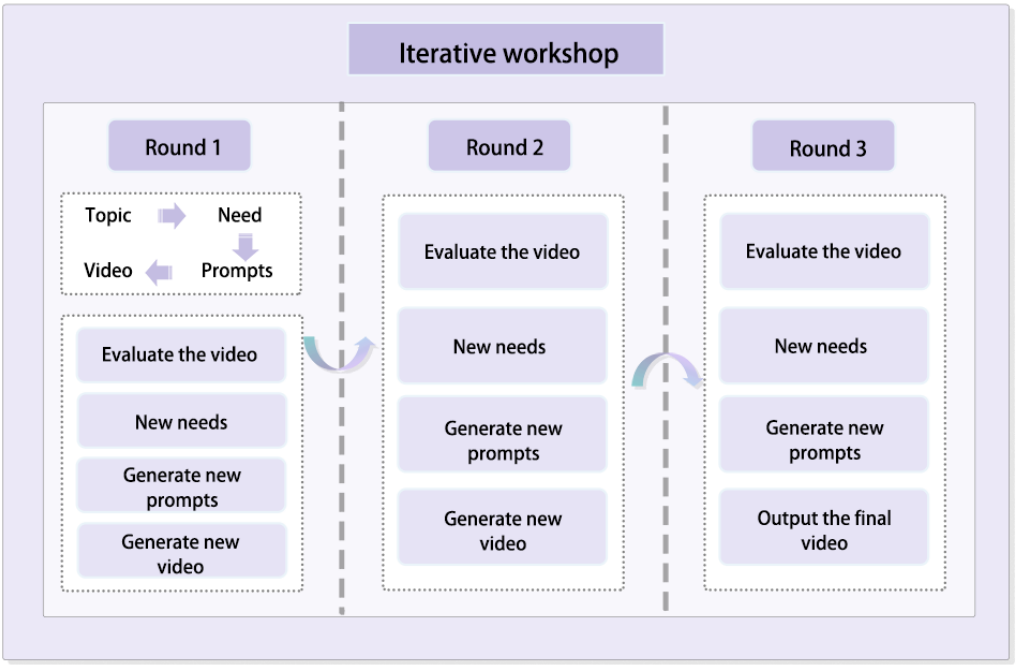
Procedure

Each workshop lasted for about an hour for a single participant and consisted of three phases: (1) pretest, (2) iterative workshop, and (3) posttest. The pretest phase (about 20 minutes) used questionnaires and interviews to collect participants’ subjective evaluations regarding health knowledge needs, technological proficiency, and self-efficacy. In addition, each participant received a 15-minute introduction to the fundamental principles of GenAI and our prototype for video generation. Then, the iterative workshop was carried out with the following procedure

XSL•FO
RenderX

(Figure 3). First, participants identified health topics based on personal interests and articulated their requirements across 5 video dimensions (visual, auditory, narrative, cognitive, and emotional) [32], which were then translated into prompts for video generation by the researcher to produce the first video. Second, after each video generation, community workers reviewed the content and flagged potential issues in the video (eg, incorrect medical advice or misleading claims) if there were any. This step was put in place to ensure discussions were grounded in reliable information. Finally, participants then reviewed the initial video generated by AI, discussed their experiences, provided feedback, and proposed new requirements, which were refined into a revised version of prompts for the generation of the second video.

Figure 3. The procedure of workshop implementation.



This iterative process was repeated for a total of 3 cycles, resulting in a final video incorporating the feedback from all rounds.

The posttest phase (about 20 minutes) involved another series of questionnaires and interviews to evaluate participants' satisfaction with the final video, as well as the changes in self-efficacy and technology acceptance. The entire research process emphasized iterative improvements to the generated videos to ensure progressive alignment between participant needs and the capabilities of the GenAI technology.

Data Collection and Analysis

This study adopts a mixed methods approach, integrating both qualitative and quantitative data for analysis. Data were primarily collected from two sources: (1) pretest and posttest questionnaires, along with interview transcripts from the iterative workshops, and (2) prompts and generated videos recorded in the system. To mitigate potential social-desirability effects stemming from the workshop setting, interviews were conducted by researchers in a neutral community environment, with a strong emphasis to participants that their negative feedback was essential for system improvement.

Interview transcripts were analyzed using the 6-phase reflexive thematic analysis (RTA) by Braun and Clarke [33]. RTA was specifically chosen for this qualitative inquiry due to its emphasis on the researcher's subjective and reflexive involvement in meaning-making, allowing us to generate rich, interpretative themes beyond simple data summation, focusing on "what is said and why" in the context of emerging GenAI technology. The RTA process involved: (1) familiarization (transcribing all interviews and reading the data for initial immersion); (2) generating initial codes (identifying semantic and latent features across the dataset); (3) searching for themes (grouping codes into meaningful broader categories); (4) reviewing themes (checking themes against the coded extracts and the entire dataset to ensure internal consistency and external heterogeneity); (5) defining and naming themes (creating clear, concise, and academically rigorous themes, as presented in Table 1); and (6) producing the report. The coding process was performed collaboratively by two researchers to ensure analytical rigor through constant comparison and to minimize individual bias.

XSL•FO
RenderX

Table 1. Theoretical dimensions and thematic coding.

Theoretical dimension and theme	Sample quotations
Self-efficacy	
Gaining confidence using GenAI[a] videos	• "I believe I can complete the operation with guidance." (E1) • "I am capable of learning as long as someone teaches me." (E6)
Instructions needed for using GenAI videos	• "I need step-by-step instructions to use it." (E4) • "I need specialized guidance on using GenAI." (E7)
Ease of use	• "This video generation process feels simpler than other mobile apps I use." (E3) • "The operation feels as simple as using WeChat." (E2)
Technology acceptance	
Perceived usefulness	• "The generated video effectively explains hypertension symptoms, which is very useful for me." (E1) • "The health videos help address my medical concerns." (E5)
User-friendliness	• "The interface is more user-friendly than expected."(E8)
Intention to use	• "I hold an open attitude toward learning GenAI and creating videos, and I'm willing to try." (E3) • "Gradual adoption is needed." (E9)
Perceived risks	• "I worry about privacy breaches when using GenAI." (E15) • "GenAI outputs require human review." (E19)
Health management	
Ways of understanding information	• "Dynamic visuals are more intuitive than static images or text." (E1) • "Animated video explanations are clearer than doctors' charts." (E3)
Providing personalized information	• "It can provide personalized dietary recommendations." (E4) • "GenAI gives exercise guidance tailored to my current physical condition." (E7)
Needs of management plans	• "I hope GenAI can help me develop fitness plans." (E1) • "I want GenAI to provide sleep improvement plans." (E12)
Requiring further medical or professional advice	• "I'll consult doctors before adopting GenAI suggestions." (E5) • "Family or professionals should review certain content." (E9)

[a]GenAI: generative artificial intelligence.

The coding process was performed collaboratively by 2 researchers who independently coded the data before holding coordination meetings to refine reflexive interpretations and ensure coding consistency. This methodical approach adheres to the standards for reporting qualitative research guidelines. The research team was external to the community center, maintaining a critical distance to minimize potential bias. Thematic saturation was determined when no new codes or subthemes emerged during the analysis of the final 3 participant interviews. We analyzed interview transcripts from all 20 participants, reaching thematic saturation where no new themes related to self-efficacy, acceptance, or concerns emerged in the final interviews. Thematic saturation was determined when no new codes or subthemes emerged during the analysis of the final 3 interviews.

Prompts and videos were analyzed using an inductive content analysis [34] to categorize the evolution of user requests and the content sophistication across the 3 rounds. In addition, comparative analysis was conducted to assess the iterative evolution of prompts and videos.

For quantitative analysis, the pretest and posttest questionnaires were designed to be relatively short and included only a few dimensions, so that older adults could complete them without difficulties. As older adults' needs and preferences for using technologies can be unique, we chose usability dimensions that focus on aging [26]. For this reason, we referred to the literature studying older adults' acceptance of technologies in health care to select appropriate dimensions for our questionnaire. As suggested by recent research and meta-analysis [24,28], we included the following constructs in our study: PU (3 items) refers to the perceived improvement in health management; PEOU (3 items) assesses the effort required; and behavioral intention (BI, 2 items) measures the future usage likelihood. In addition, to gauge the effectiveness and efficacy of older adults in adopting GenAI for health management, we included additional measurements of technology confidence (TechConf,

XSL•FO
RenderX

4 items) [35], task mastery (TaskMastery, 2 items) [25], and the learning of health information (HealthLearn, 3 items) [36]. Internal consistency for these scales was verified, with all constructs demonstrating acceptable reliability (Cronbach α>0.70). The questions were presented to the participants in the pre- and postphases, then Wilcoxon signed-rank tests were used to examine significant differences between pretest and posttest questionnaire results. To evaluate effect sizes, we computed Cohen *d* and additionally reported the Wilcoxon-compatible effect size *r* (calculated as Z divided by the square root of the sample size). In this study, items within core dimensions showed substantial effect sizes; for instance, Cohen *d* values for PEOU items ranged from 0.96 to 1.53 (*r*=0.75-0.86), and the BI items reached a Cohen *d* of 1.05 (*r*=0.81). The full list of questions can be found in Multimedia Appendix 1.

Sample

We recruited 20 older adults residing in nearby communities to participate in iterative workshops of GenAI videos. The participants were aged 60-84 (mean 70.15, SD 6.90) years, possessed basic skills of using smartphones, but had no experience with GenAI. This inclusion criterion ensures that their initial attitudes and user acceptance toward GenAI were not influenced by previous exposure. To enhance the heterogeneity of our sample, the participants reported diverse health statuses (eg, healthy, with chronic or mild conditions), living arrangements (eg, living alone or with family), and educational background (from primary school to university). They were also interested in a wide range of health topics, such as the 3 hyper-conditions, cardiovascular health, sleep disorders, and exercise-related topics. Table 2 lists the demographic information of the participants.

Table 2. Demographic information of our sample.

ID	Age (years)	Gender	Self-reported health status	Living arrangement	Education level	Health topics of concern
E1	60	Woman	Healthy	Alone	High school	Hypertension, hyperlipidemia, and hyperglycemia
E2	80	Woman	Chronic condition	Alone	Primary school	Sleep disorders
E3	78	Woman	Healthy	Living with family	Junior high school	Cardiovascular health
E4	77	Men	Chronic Condition	Living with family	High school	Hypertension, hyperlipidemia, and hyperglycemia
E5	72	Woman	Chronic condition	Living with family	Primary school	Healthy diet
E6	61	Woman	Chronic condition	Living with family	University	Physical wellness
E7	69	Woman	Chronic condition	Living with family	Junior high school	Obesity
E8	73	Woman	Chronic condition	Alone	Primary school	Healthy diet
E9	67	Woman	Healthy	Living with family	High school	Joint health
E10	76	Woman	Healthy	Alone	Primary school	Joint health
E11	74	Men	Mild condition	Living with family	Junior high school	Exercises
E12	69	Men	Chronic condition	Alone	Junior high school	Healthy diet
E13	73	Woman	Chronic condition	Living with family	Primary school	Joint health
E14	62	Woman	Mild condition	Living with family	University	Thyroid issues
E15	84	Woman	Chronic condition	Alone	College diploma	Sleep disorders
E16	67	Woman	Healthy	Alone	Junior high school	Joint health
E17	62	Woman	Healthy	Lives with family	Primary school	Hypertension, hyperlipidemia, and hyperglycemia
E18	72	Woman	Healthy	Lives with family	High school	Healthy diet
E19	62	Woman	Healthy	Alone	Primary school	Joint health
E20	65	Men	Chronic disease	Alone	University	Sleep disorders

Results

Quantitative Impact on Confidence and Learning

Through Wilcoxon signed-rank tests, we analyzed the impact of workshops on multiple dimensions among older adults. The results demonstrated consistently positive mean differences across all dimensions (ranging from 0.375, SD 0.582, to 0.917, SD 0.571) after the workshops, with all reaching statistical significance ($P \leq .01$). Notably, PEOU, PU, BI, technical confidence (TechConf), and the influence on health information learning (HealthLearn) exhibited substantial mean increases. Table 3 shows the results of the Wilcoxon signed-rank tests.

XSL•FO
RenderX

Table 3. Summary of Wilcoxon signed-rank test results.

Item	Mean difference (SD)	Wilcoxon statistic	*P* value
PU[a]	0.867 (0.652)	0	.001[b]
PEOU[c]	0.917 (0.571)	0	.001[b]
BI[d]	0.575 (0.712)	12	.003[b]
TechConf[e]	0.65 (0.455)	0	.001[b]
TaskMastery[f]	0.375 (0.582)	13.5	.01[b]
HealthLearn[g]	0.717 (0.544)	0	.001[b]

[a]PU: perceived usefulness.

[b]Statistically significant.

[c]PEOU: perceived ease of use.

[d]BI: behavioral intention.

[e]TechConf: technical confidence.

[f]TaskMastery: task mastery.

[g]HealthLearn: health information learning.

To evaluate the practical significance of these changes, Cohen *d* was calculated. The results showed large effect sizes for PEOU (*d*=1.61), PU (*d*=1.33), and the influence of health information learning (*d*=1.32). Medium to large effects were also found in technical confidence (*d*=1.43), BI (*d*=0.81), and task mastery (*d*=0.64). These results indicated that these improvements were not only statistically significant but also practically meaningful. Table 4 lists the results of each item and its Cohen *d* tests.

XSL•FO
RenderX

Table 4. Wilcoxon signed-rank test and Cohen *d*.

Item	Mean difference (SD)	Wilcoxon statistic	*P* value	Cohen *d*
PU1[a]	0.75 (0.79)	0	.002[b]	0.95
PU2	1 (1.03)	0	.002[b]	0.97
PU3	0.85 (0.81)	0	.002[b]	1.05
PEOU1[c]	1.1 (0.72)	0	.001[b]	1.53
PEOU2	0.85 (0.81)	0	.002[b]	1.05
PEOU3	0.8 (0.83)	7	.001[b]	0.96
BI1[d]	0.3 (0.92)	37.5	.16	0.32
BI2	0.85 (0.81)	0	.001[b]	1.05
TechConf1[e]	0.5 (0.51)	0	.002[b]	0.97
TechConf2	0.75 (0.79)	0	.002[b]	0.95
TechConf3	0.7 (0.86)	0	.004[b]	0.81
TechConf4	0.65 (0.75)	0	.004[b]	0.87
TaskMastery1[f]	0.35 (0.75)	10	.05	0.47
TaskMastery2	0.4 (0.75)	11	.03[b]	0.53
HealthLearn1[g]	0.8 (1.06)	4.5	.005[b]	0.76
HealthLearn2	0.6 (0.6)	0	.001[b]	1
HealthLearn3	0.75 (0.72)	0	.001[b]	1.05

[a]PU: perceived usefulness.

[b]Statistically significant.

[c]PEOU: perceived ease of use.

[d]BI: behavioral intention.

[e]TechConf: technical confidence.

[f]TaskMastery: task mastery.

[g]HealthLearn: health information learning.

Older Adults' Experiences of Technology

Overview

We qualitatively analyzed the interview transcripts of 20 participants as data saturation was reached, and identified themes around self-efficacy, technology acceptance, and health information comprehension. Data analysis followed the 6-phase method for RTA by Braun and Clarke [33]. RTA emphasizes the researcher's reflexive involvement in the data analysis process, where subjective interpretation is central to understanding and generating meaning from qualitative data [37]. In terms of the changes in self-efficacy, we found that older adults gained confidence using GenAI from the workshops. Meanwhile, they expressed the need for more instructions on using GenAI. The ease of use of the GenAI video technology under simple and clear instructions also contributed to the improvement in self-efficacy.

Another category of emerging themes was related to technology acceptance. Participants perceived that AI-generated videos were useful, and the GenAI tool was user-friendly. The workshops demonstrated the applications of GenAI videos in health management, and participants tended to use them in the future after witnessing the positive side of GenAI. At the same time, they pointed out the potential risks of such AI tools.

From a health management perspective, older adults found that GenAI videos changed their ways of understanding health information and allowed them to retrieve personalized information. However, the AI tool warranted more improvements, such as providing health management plans and integrating validation or reviews from medical and professional experts to ensure the accuracy and reliability of content. Table 1 summarizes the theoretical dimensions, themes, and sample quotations to support the findings. We use the following subsections to offer further insight into each theme.

Improvements in Self-Efficacy

Through content analysis of interview transcripts, it was found that the self-efficacy of participants exhibited a significant upward trend after engaging with the GenAI video tool. Before the workshops, 12 (60%) participants explicitly expressed concerns about the difficulties of GenAI technologies because

XSL•FO
RenderX

of their ages. Many participants mentioned concerns such as, "my memory has deteriorated, and I can't learn new things as quickly as before" (E4), and "I need my children's guidance to operate these new technologies" (E8). This low confidence stemmed from doubts about their digital capabilities and fear of complicated user interfaces. After the workshop, 14 (70%) participants reported that the iterative workshops helped them to use GenAI. For instance, participants noted, "by following your instructions and prompts step by step, I was able to complete the task" (E3) and "I could better understand the content [after the workshop]" (E12).

In the initial workshop stage, many participants demonstrated a strong reliance on external support, which meant that they hoped to complete learning through vocal guidance or assistance from other people. After the workshop, 12 (60%) participants reported that they could independently use our GenAI tool to generate videos or adjust various settings. As 1 participant stated, "[other people's] guidance when making mistakes enables success" (E5).

Before the workshops, 8 (40%) participants perceived GenAI video generation as a highly challenging task. The workshop helped to mitigate cognitive load through template-based prompt designs (eg, preset health-related prompting templates) and video revisions with feedback iterations. The postworkshop interviews showed that 16 (80%) participants expressed the ease of use of GenAI tasks. For example, "after the entire process, I [have] found [that] GenAI video generation simpler than I originally imagined" (E11) and "I now believe I can attempt creating my health videos" (E14).

Increase in Technology Acceptance

After the workshops, 18 (90%) participants emphasized the PU of the generated content was useful for them to manage their health, for example, "the GenAI-generated health video comprehensively explains hypertension knowledge and the content is practical" (E1) and "my legs are weak and I don't know what exercises suit me. GenAI recommended exercises tailored to my condition" (E5).

Many participants commented on the user-friendliness of GenAI features. Among them, 16 (80%) participants mentioned the advantages of adaptable user interfaces (such as adjustable subtitle sizes) and the obvious design for specific functions (such as dedicated buttons for AI interaction). Other participants compared the GenAI with other commonly used mobile phone apps, for instance, "the operations are simple and similar to WeChat" (E2), which further contributed to the intention of using the technology.

While sentiment was generally positive, 5 (25%) participants expressed privacy concerns, including the possible leaks of personal data and the incorrect information provided by GenAI. For example, a participant said, "I worry about privacy breaches when using GenAI" (E15), while others suggested, "GenAI outputs require human review" (E19) or "professionals should review certain content" (E9).

Meeting Health Management Needs

GenAI videos were reported to help older adults better understand health information. A total of 15 (75%) participants described that AI-generated videos were a better way to present health knowledge, for example, "videos explain health knowledge better" (E3), and 11 (55%) participants mentioned that self-caring demonstrations generated by the GenAI were better, for example, "scenario-based demos in videos are more lively than [the demos from] medical experts" (E11).

Furthermore, at least 12 (60%) participants mentioned that GenAI video tools could provide personalized information, for instance, E12 reflected that the AI-generated video "taught how to do daily calorie calculations [for myself]," while many others mentioned that the GenAI could offer information about "healthy diets" and "hypertension control."

Despite the positive feedback, participants highlighted the need for the GenAI to generate health management plans for older adults saying, "I hope GenAI can help me develop fitness plans" (E1), and "I want GenAI to provide sleep improvement plans" (E12). In addition, some participants suggested that they would not rely solely on the information provided by the GenAI. Instead, they will "consult doctors before adopting GenAI's suggestions" (E5).

Interaction Changes With Experience: Prompt Analysis

Overview

With the analysis of a total of 60 prompts collected in all workshops, the changes of prompts across 3 rounds of video generation demonstrated the progressive trajectory of older adult–AI interaction from exploration to decision-making support. The details in each round are shown below.

First Round

In the first round, older adults demonstrated an exploratory attitude toward AI-generated videos. Their understanding of GenAI began at an early stage, and their core concerns centered on exploring the basic capacities of the GenAI tool. Their prompts contained questions about privacy concerns and the accuracy of answers to explore the accuracy and reliability levels of the GenAI tool. Below are examples of this category of prompts:

> *I'm 77 years old. I have Hypertension, Hyperlipidemia, and Hyperglycemia. I want to know how diabetes can be prevented. Will my personal information be disclosed?* [E4]
>
> *I was 62. I lumped my thyroid. I want to know if there is a risk. Is the health video content [that] you provide correct?* [E17]

Second Round

In the second round, older adults focused on functional adaptation, and the prompts were changed accordingly. With a preliminary understanding of the technology, their focus shifted toward whether GenAI could precisely satisfy individual needs and whether the information was simple enough to understand. Prompts emphasizing personalized recommendations and

XSL•FO
RenderX

requiring easy-to-understand information revealed their needs. The following are samples of the prompts:

> *Help me generate a video about joint health care. This video is a bit complicated. Please make it simpler and more understandable.* [E16]

> *I want to know how to lose weight healthily. Please recommend some diet and exercise videos that are suitable for my specific health condition.* [E18]

Third Round

In the final round, after asking the GenAI to provide personalization, older adults escalated to the level of asking for information to support decision-making. For instance, older adults asked for the content to support the discussion during consultations, or the content reviewed by human experts that could be shown to family members. The following shows the excerpts from such prompts:

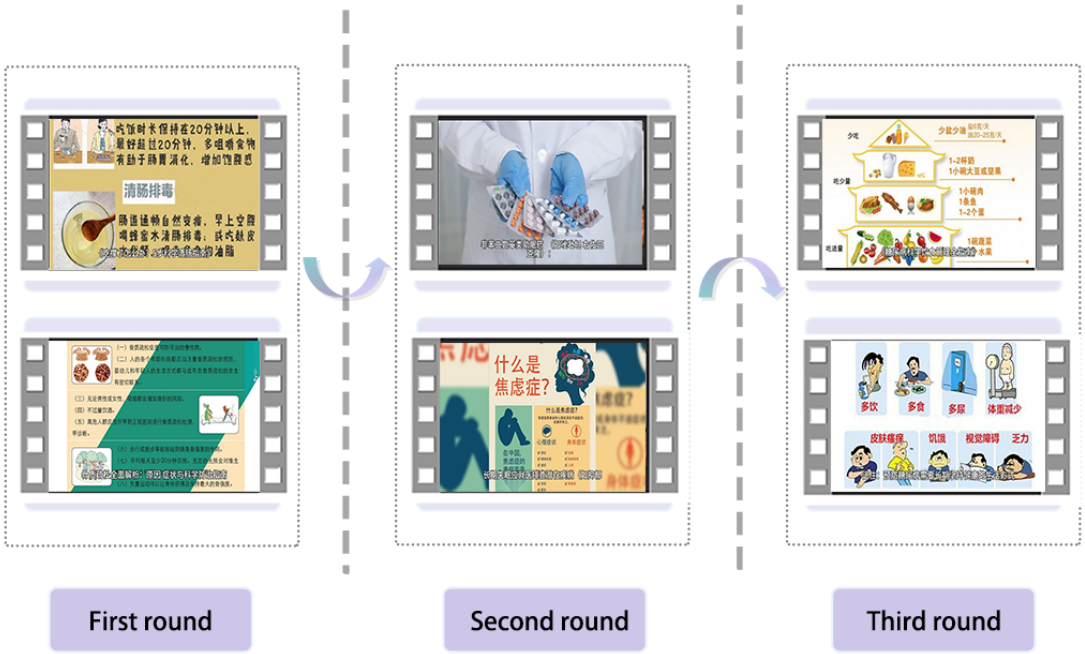
> *I want to know about the issue of obesity among the elderly. Which method is better for losing weight, traditional Chinese medicine or Western medicine? … Are the suggestions you gave me the result of a review by relevant experts?* [E7]

> *I want to know about digestive problems in the elderly. My stomach often feels uncomfortable after eating. … Can dizziness only be caused by low blood sugar? Please help me make a detailed video about a healthy diet and lifestyle. Please provide some health suggestions that experts have reviewed.* [E5]

Results of Generated Videos' Analysis

The analysis of 60 videos produced across 3 rounds of workshops revealed the patterns evolving from vague to concrete. In the first round, participants engaged in exploratory attempts with ambiguous needs. At this stage, most videos consisted solely of text-based captions with static images. Then, the second round saw participants, who had gained preliminary experience with the GenAI tool, begin proposing more concrete demands. The video content at this stage started incorporating scenario-based plots, with a significant portion focusing on specific topics such as health management and medication reminders. However, the explanations about these topics remained vague. In the final iteration, the videos became more sophisticated, with most presenting comprehensive health solutions with the customization of target audiences, content formats, and risk management strategies. Figure 4 shows the video samples in 3 rounds of iterations. Chinese is displayed because the screenshots are captured from the original video clips.

Figure 4. Sample videos in different rounds of the workshops.



Discussion

Principal Findings

This study systematically investigates the application and potential of GenAI video technologies and our prototype AIHealthV in older adults' health management. The research captures the dynamic evolution of older adults' interactions with GenAI and reveals how their needs progress from basic exploration to sophisticated decision-making support. Our findings provide novel insights into the role of AI-generated videos in reducing cognitive load and enhancing health information comprehension among older populations, while highlighting the importance of understanding their concerns in using GenAI technologies.

Interaction and Multimodality: Cognitive Accessibility and Self-Efficacy

The research indicated an improvement in older adults' self-efficacy and confidence, which may be viewed in alignment with the core principles of social cognitive theory. Self-efficacy suggests that individuals' beliefs in their capabilities can be

XSL•FO
RenderX

progressively strengthened [38]. As our GenAI prototype was structured with phased task decomposition, for example, preloaded health templates, and provided immediate feedback mechanisms, for example, step-by-step prompt guidance, these factors likely contributed to older adults building confidence during the experience. Older adults' belief in technological efficacy may be reinforced through gradual, experience-based accumulation when technological interactions are tailored to their cognitive capacities [39].

GenAI health videos offer a personalized visualization of health information that is often complex or difficult to access. This capability has the potential to enhance the PU of the technology, a factor that is widely understood to influence BI. Beyond mere information delivery, our findings suggest that multimodal GenAI content plays a crucial role in satisfying the emotional and cultural needs of older users. This resonance is vital for sustained engagement, as multimodal learning is most effective when it addresses the user's underlying emotional context, a pattern also observed in how younger generations engage with multimodal platforms such as Douyin (ByteDance) for skill learning [40]. Observations suggest that while some participants initially expressed reservations about the complexity of the tools, their acceptance and willingness to use the technology appeared to increase alongside a progressive learning curve [41]. The unified theory of acceptance and use of technology, where experience acts as a critical moderating variable, posits that accumulated familiarity with technologies can substantially improve users' attitudes [42]. Our work proposes that encouraging older adults to adopt GenAI for personal health purposes may necessitate continuous guidance and support, as well as patience, to help them cultivate confidence and practical experience. This further suggests that GenAI functions should be considered as part of a sociotechnical intervention strategy that supports older adults' digital capabilities through both social and technical means. The multimodality of GenAI videos is also thought to contribute to technological confidence and adoption. The media richness theory by Sun and Cheng [43] proposes that high-richness media, such as animation and voice interaction, may be better suited to meet older adults' cognitive needs. While textual data remains the dominant GenAI modality, the maturation of AI-supported interactive technologies, for example, video generation, multimodal fusion, and adaptive learning algorithms, is expected to lead to the increased application of multimodal GenAI in older adults' health education and management.

GenAI Videos as a Potential Virtual Companion and Access to Health Experts

As global aging intensifies, we can witness the potential of GenAI video tools as a potential virtual companion, while the prevalence of older adults living alone has risen markedly [44]. These individuals often face greater challenges in accessing health knowledge through digital technologies due to the absence of familial or societal support networks [45]. GenAI technologies may help mitigate this gap by providing algorithm-driven personalized health guidance, partially substituting traditional interpersonal support functions [46].

However, the sociotechnical implications of GenAI present a paradoxical duality. From a positive perspective, it helps address immediate gaps in health service accessibility for older adults living alone by providing basic health information and guidance when human support is unavailable. While not a complete replacement for professional care or family support, these GenAI tools can serve as a temporary solution that enables older adults to access health knowledge and manage certain aspects of their well-being independently [47]. The research by Niehaves and Plattfaut [48] suggests that digital technologies can compensate for social support deficits. Similarly, the work by Satariano and Scharlach [49] highlights how technology enhances older adults' autonomy. The interaction between older adults and GenAI tools can be further understood through the lens of "reverse generative learning." This approach emphasizes that intergenerational digital inclusion can be reconstructed when technology serves as a bridge, allowing older adults to develop digital readiness through creative engagement with AI [50]. Dynamic video tutorials generated by GenAI tools enable independent completion of health management tasks (such as medication management), reducing reliance on community clinics and the younger generation. On the other hand, prolonged overreliance on virtual interactions may erode real-world social connections and exacerbate emotional isolation risks [51]. Longitudinal studies reveal that while technological interventions improve functional autonomy (eg, self-management capabilities), they cannot replicate the emotional resonance of human care and potentially lead to digital isolation [52]. Recent observations indicate that some older adults living alone have reduced their frequency of visiting community health centers after adapting to GenAI, which further substantiates the contradiction [53].

To this end, technological design should adhere to the complementary principle, for example, integrating AI health videos with community volunteer service networks to provide human-based support, and at the same time, it can sustain emotional bonds with people. Cross-generational digital inclusion strategies (eg, encouraging grandchildren to participate in technology learning) could further supplement this approach, reconstructing social support networks through intergenerational interaction. These findings can underscore the necessity of embedding technological interventions within social ecosystems to avoid the pitfalls.

While older adults appreciate the value of addressing health and lifestyle issues brought by GenAI videos, they have concerns about privacy and highlight the need for human content review. These clearly reflect their uncertainty about the technology while balancing expectations for convenience against fears of misinformation or privacy violations. Finally, GenAI videos present a unique opportunity to bridge the gap between older adults and reliable health information sources. By incorporating verification mechanisms that connect AI-generated content with certified medical expertise, these tools can serve as a gateway to professional health resources. For instance, videos could include prompts encouraging users to discuss the information with their doctors or provide direct links to schedule consultations with health care providers, such as real health experts. This approach not only enhances the credibility of

XSL•FO
RenderX

AI-generated content but also fosters more informed conversations between older adults and medical professionals, ultimately leading to better health outcomes.

Ethical, Privacy, and Usability Concerns for Older Adults to Use GenAI Videos

Older adult users' ethical concerns regarding GenAI health videos reflect profound tensions between digital-era health rights and privacy rights. Despite being aware of risks, participants still tend to accept the use of these technologies, driven by anticipated health benefits. This replicates the view of the risk perception theory by Slovic [54], wherein individuals tolerate higher risks when potential benefits are sufficiently significant. However, the "black box" nature of AI algorithms renders trust establishment that is heavily dependent on external authorities (eg, the guidance from community workers), which exposes the gap in older adults' algorithmic interpretability rights [55]. Intergenerational differences further complicate the ethical issues. Older adults who are living alone exhibit different risk perceptions when compared to those cohabiting with family members, which prompts the necessity of the development of participatory governance frameworks [56].

Our study found that older adults' acceptance of GenAI is influenced by privacy concerns. These concerns mainly stem from worries about personal data security, particularly regarding the storage and potential misuse of sensitive health information, which are consistent with recent findings [57]. Many older users distrust algorithmic systems, fearing accidental data leaks and questioning the reliability of AI-generated health advice. These concerns are further exacerbated by potential algorithmic biases. To address these challenges, GenAI can enhance older adults' trust by establishing transparent data governance frameworks that clearly explain how personal information is collected, processed, and protected.

On the other hand, our findings demonstrate that older adults' use of GenAI video tools follows a process of constructing cognitive perceptions of technological controllability, for example, the needs of adjusting playback speed and the verbal language used in narratives. The interactions between seniors and the GenAI should balance between features and aging-friendly design; for instance, step-by-step guidance, visual support, and accessible user interfaces are crucial for older adults to gain momentum in learning AI tools. In practice, the transition from passive experimentation to active cocreation can be facilitated to inform better design.

Health videos generated by GenAI may contain misleading or inaccurate information, which could disproportionately affect older adults with lower digital literacy [57]. This phenomenon is often associated with GenAI's hallucinations, which may lead to outdated or inaccurate information. Such risks can be mitigated through the implementation of expert validation systems, where health care professionals review AI outputs. In addition, older adults should be educated to understand the limitations of AI-generated advice, and therefore, older adults' AI literacy will be an important topic for reducing the impact of misinformation provided by GenAI.

Limitations and Future Work

This study has several limitations. First, the urban community-focused sample may not capture access disparities of AI among rural or cross-border older adult populations, potentially limiting the generalizability of findings. Second, the short intervention period restricts the observation of long-term dynamics in technological usage. Third, reliance on self-reported data may introduce social desirability bias, possibly overestimating actual technology acceptance levels. Future work should expand to urban-rural and cross-border contexts to enhance sample diversity and use a longitudinal research design, for example, using diary methods and behavioral logs to track long-term usage patterns. The integration of multimodal data (eg, eye-tracking and physiological measurements) can be alternatives to measure the use of GenAI by this cohort.

Conclusions

Our work makes a unique contribution to older adults' health by examining the potential of GenAI videos in supporting older adults' self-management and highlighting important considerations for future implementation. Currently, AI-generated videos show a promising future of empowering older adults to take an active role in managing their health despite several concerns. As AI technologies continue to advance, their thoughtful applications in aging populations can create new possibilities for accessible, personalized, and effective health education while respecting older adults' unique needs and preferences. Future research should explore the longitudinal effects of AI-generated videos on older adults' health outcomes, as well as the feasibility of better integrating these tools with health care systems and social support networks. It is also crucial to reduce hallucinations and misinformation in GenAI, especially for older adults with limited AI literacy. The use of GenAI videos as a learning resource for older adults in other contexts also remains a rich area of exploration for future work in older adults' health, as do the ethicolegal implications of this kind of AI use.

Acknowledgments

We are grateful to the Macao Federation of Trade Unions (FAOM) for their assistance in this research.

Data Availability

The author confirms that all data generated or analyzed during this study are included in this published paper. Furthermore, primary and secondary sources and data supporting the findings of this study were all publicly available at the time of submission.

XSL•FO
RenderX

Funding

This research was funded by the Macao Science and Technology Development Fund (FDCT, 0032/2025/ITP1). The study has been funded for a 2-year period, starting in July 2025. Additionally, it was funded by the Macao Polytechnic University research grant (RP/FCA-24/2025). The study has been funded for a 2-year period, starting in January 2026.

Authors' Contributions

TL and PP proposed and designed this study. YTL helped with data analyses and resolved disagreements in the review process. TL and PP wrote this paper with the assistance of YTL, PP, DM, GB, and SC, and supervised and provided resources for this work. All authors discussed the results and contributed to this final paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
GenAI health video research questionnaire. GenAI: generative artificial intelligence.
[DOCX File , 24 KB - aging_v9i1e88005_app1.docx]

References

1. World Health Organization. Ageing and health. WHO. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2026-02-21]
2. Melenhorst AS, Rogers WA, Bouwhuis DG. Older adults' motivated choice for technological innovation: evidence for benefit-driven selectivity. Psychol Aging 2006;21(1):190-195. [doi: 10.1037/0882-7974.21.1.190] [Medline: 16594804]
3. Zhang S, Wu L, Zhao Z, Massó JRF, Chen M. Artificial intelligence in gerontology: data-driven health management and precision medicine. Adv Gerontol 2024;14(3):97-110. [doi: 10.1134/s2079057024600691]
4. Charness N, Boot WR. A grand challenge for psychology: reducing the age-related digital divide. Curr Dir Psychol Sci 2022;31(2):187-193 [FREE Full text] [doi: 10.1177/09637214211068144] [Medline: 35754678]
5. van der Vaart R, van Deursen AJAM, Drossaert CHC, Taal E, van Dijk JAMG, van de Laar MAFG. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
6. Bickmore TW, Paasche-Orlow MK. The role of information technology in health literacy research. J Health Commun 2012;Suppl 3:23-29 [FREE Full text] [doi: 10.1080/10810730.2012.712626] [Medline: 23030559]
7. Pradhan A, Lazar A, Findlater L. Use of intelligent voice assistants by older adults with low technology use. ACM Trans Comput-Hum Interact 2020;27(4):1-27. [doi: 10.1145/3373759]
8. Shandilya E, Fan M. Understanding older adults' perceptions and challenges in using AI-enabled everyday technologies. Guangzhou: Association for Computing Machinery; 2024 Presented at: Proceedings of the Tenth International Symposium of Chinese CHI; Oct 23, 2022; Guangzhou, China p. 105-116. [doi: 10.1145/3565698.3565774]
9. Bickmore TW, Caruso L, Clough-Gorr K. Acceptance and usability of a relational agent interface by urban older adults. : Association for Computing Machinery; 2005 Presented at: CHI '05 Extended Abstracts on Human Factors in Computing Systems; Apr 02, 2005; Portland, OR p. 1212-1215. [doi: 10.1145/1056808.1056879]
10. Elueze I, Quan-Haase A. Privacy attitudes and concerns in the digital lives of older adults: Westin's privacy attitude typology revisited. Am Behav Sci 2018;62(10):1372-1391. [doi: 10.1177/0002764218787026]
11. Rochmawati E, Kamilah F, Iskandar AC. Acceptance of e-health technology among older people: a qualitative study. Nurs Health Sci 2022;24(2):437-446. [doi: 10.1111/nhs.12939] [Medline: 35297152]
12. Lee OE, Kim DH, Hyenjoo L, Beum KA. Information communication technology use to improve eHealth literacy, technophobia, and social connection among community dwelling older adults. Educ Gerontol 2022;48(10):445-457. [doi: 10.1080/03601277.2022.2045727]
13. Preiksaitis C, Rose C. Opportunities, challenges, and future directions of generative artificial intelligence in medical education: scoping review. JMIR Med Educ 2023;9:e48785 [FREE Full text] [doi: 10.2196/48785] [Medline: 37862079]
14. Ahn S, Yim HJ, Lee Y, Park SI. Dynamic and super-personalized media ecosystem driven by generative AI: unpredictable plays never repeating the same. IEEE Trans Broadcast 2024;70(3):980-994. [doi: 10.1109/tbc.2024.3380474]
15. Uddin J, Feng C, Xu J. Health communication on the internet: promoting public health and exploring disparities in the generative AI Era. J Med Internet Res 2025;27:e66032 [FREE Full text] [doi: 10.2196/66032] [Medline: 40053755]
16. Marey A, Saad AM, Killeen BD, Gomez C, Tregubova M, Unberath M, et al. Generative artificial intelligence: enhancing patient education in cardiovascular imaging. BJR Open 2024;6(1):tzae018 [FREE Full text] [doi: 10.1093/bjro/tzae018] [Medline: 39086557]
17. Musiol M. Generative AI: Navigating the Course to the Artificial General Intelligence Future. New Jersey: Wiley; 2024.

XSL•FO
RenderX

18. Cao HJ, Lee HR, Peng W. Empowering adults with AI Literacy: using short videos to transform understanding and harness fear for critical thinking. : Association for Computing Machinery; 2025 Presented at: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems; April 25, 2025; Yokohama p. 1-8. [doi: 10.1145/3706598.3713254]
19. Cecconi C, Adams R, Cardone A, Declaye J, Silva M, Vanlerberghe T, et al. Generational differences in healthcare: the role of technology in the path forward. Front Public Health 2025;13:1546317 [FREE Full text] [doi: 10.3389/fpubh.2025.1546317] [Medline: 40078753]
20. Fang YM, Chun L, Chu BC. Older adults' usability and emotional reactions toward text, diagram, image, and animation interfaces for displaying health information. Appl Sci 2019;9(6):1058. [doi: 10.3390/app9061058]
21. Pramod D, Patil K. Generative AI for elderly well-being through the computer as social actor paradigm. 2024 Presented at: IC3-2024: Proceedings of the 2024 Sixteenth International Conference on Contemporary Computing; Oct 28, 2024; Noida, India. [doi: 10.1145/3675888.3676033]
22. Wilkowska W, Ziefle M. Privacy and data security in E-health: requirements from the user's perspective. Health Inf J 2012;18(3):191-201 [FREE Full text] [doi: 10.1177/1460458212442933] [Medline: 23011814]
23. Nabavi RT. Bandura's social learning theory & social cognitive learning theory. 2012. URL: https://www.researchgate.net/publication/267750204 [accessed 2026-02-21]
24. Kim SD. Application and challenges of the technology acceptance model in elderly healthcare: insights from ChatGPT. Technologies 2024;12(5):68. [doi: 10.3390/technologies12050068]
25. Zhu W, Mat Jusoh NH, Alan R, Latip M, Md Saad J. Unlocking self-esteem in older adults: a conceptual exploration of technological proficiency and its effects. Behav Sci 2025;15(3):306. [doi: 10.3390/bs15030306]
26. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023;6:e44564 [FREE Full text] [doi: 10.2196/44564] [Medline: 37191976]
27. Bandura A. Self-Efficacy: The Exercise of Control. New York City, NY: W H Freeman/Times Books/ Henry Holt & Co; 1997.
28. Ma QC, Alan HST, Teh L. Insights into older adults' technology acceptance through meta-analysis. Int J Hum–Comput Interact 2021;37(11):1049-1062. [doi: 10.1080/10447318.2020.1865005]
29. Malureanu A, Panisoara G, Lazar I. The relationship between self-confidence, self-efficacy, grit, usefulness, and ease of use of eLearning platforms in corporate training during the COVID-19 pandemic. Sustainability 2021;13(12):6633. [doi: 10.3390/su13126633]
30. Suragarn U, Hain D, Pfaff G. Approaches to enhance social connection in older adults: an integrative review of literature. Aging Health Res 2021;1(3):100029 [FREE Full text] [doi: 10.1016/j.ahr.2021.100029]
31. Zhu X, Liang W, Xv W, Wang Y. The key strategies for increasing users' intention of self-disclosure in human-robot interaction through robotic appearance design. : EDP Sciences; 2023 Presented at: SHS Web of Conferences; May 03, 2023; London, UK. [doi: 10.1051/shsconf/202316501012]
32. Beh J, Pedell S, Doube W. Where is the "I" in iPad? The role of interest in older adults' learning of mobile touch screen technologies. Australia: Association for Computing Machinery; 2015 Presented at: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction; Dec 07, 2025; Parkville, Victoria, Australia p. 437-445. [doi: 10.1145/2838739.2838776]
33. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
34. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x]
35. Yau HK, Cheng ALF. Gender difference of confidence in using technology for learning. J Technol Stud 2012;38(2):74-79 [FREE Full text]
36. Miriovsky BJ, Shulman LN, Abernethy AP. Importance of health information technology, electronic health eecords, and continuously aggregating data to comparative effectiveness research and learning health care. JCO 2012;30(34):4243-4248. [doi: 10.1200/jco.2012.42.8011]
37. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. Qual Quant 2022;56(3):1391-1412. [doi: 10.1007/s11135-021-01182-y]
38. Bandura A. Cultivate self-efficacy for personal and organizational effectiveness. In: Principles of Organizational Behavior. Routledge: The Handbook of Evidence-Based Management; 2023:113-135.
39. Wolfson NE, Cavanagh TM, Kraiger K. Older adults and technology-based instruction: optimizing learning outcomes and transfer. AMLE 2014;13(1):26-44. [doi: 10.5465/amle.2012.0056]
40. Wen J, Tong S, Liang W, He H, Shen H. Is the emotional and cultural need satisfied or not? – A qualitative study based on youth users learning dance through Douyin. Res Dance Educ 2025:1-22. [doi: 10.1080/14647893.2025.2571081]
41. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: perceived difficulties, experimentation behaviour and usability. Comput Hum Behav 2013;29(4):1715-1724. [doi: 10.1016/j.chb.2013.02.006]

XSL•FO
RenderX

42. Dwivedi YK, Rana NP, Jeyaraj A, Clement M, Williams MD. Re-examining the unified theory of acceptance and use of technology (UTAUT): towards a revised theoretical model. Inf Syst Front 2019;21(3):719-734. [doi: 10.1007/s10796-017-9774-y]
43. Sun PC, Cheng HK. The design of instructional multimedia in e-Learning: a media richness theory-based approach. Comput Educ 2007;49(3):662-676. [doi: 10.1016/j.compedu.2005.11.016]
44. Newmyer L, Verdery AM, Wang H, Margolis R. Population aging, demographic metabolism, and the rising tide of late middle age to older adult loneliness around the world. Popul Dev Rev 2022;48(3):829-862 [FREE Full text] [doi: 10.1111/padr.12506] [Medline: 37649654]
45. Goodall K, Newman L, Ward P. Improving access to health information for older migrants by using grounded theory and social network analysis to understand their information behaviour and digital technology use. Eur J Cancer Care (Engl) 2014;23(6):728-738. [doi: 10.1111/ecc.12241] [Medline: 25250535]
46. Li YH, Li YL, Wei MY, Li GY. Innovation and challenges of artificial intelligence technology in personalized healthcare. Sci Rep 2024;14(1):18994 [FREE Full text] [doi: 10.1038/s41598-024-70073-7] [Medline: 39152194]
47. Yang Y, Wang C, Xiang X, An R. AI applications to reduce loneliness among older adults: a systematic review of effectiveness and technologies. Healthcare 2025;13(5):446 [FREE Full text] [doi: 10.3390/healthcare13050446] [Medline: 40077009]
48. Niehaves B, Plattfaut R. Internet adoption by the elderly: employing IS technology acceptance theories for understanding the age-related digital divide. Eur J Inf Syst 2014;23(6):708-726. [doi: 10.1057/ejis.2013.19]
49. Satariano WA, Scharlach AE, Lindeman D. Aging, place, and technology: toward improving access and wellness in older populations. J Aging Health 2014;26(8):1373-1389. [doi: 10.1177/0898264314543470] [Medline: 25502245]
50. Liang W, Shamshudeen RI, Syed MAM, Hui S. Reverse generative learning: intergenerational learning for global readiness in higher education. J Int Stud 2025;15(12):119-140. [doi: 10.32674/hp7sn324]
51. Singh SK, Singh N, Kumari DS. Technology and its effect on social interactions. SSRN 2024. [doi: 10.2139/ssrn.4894688]
52. Peeters J, Wiegers T, Friele R. How technology in care at home affects patient self-care and self-management: a scoping review. Int J Environ Res Public Health 2013;10(11):5541-5564 [FREE Full text] [doi: 10.3390/ijerph10115541] [Medline: 24173139]
53. Ho A. Are we ready for artificial intelligence health monitoring in elder care? BMC Geriatr 2020;20(1):358. [doi: 10.1186/s12877-020-01764-9]
54. Slovic P. Perception of Risk: Reflections on the Psychometric Paradigm. In: Social Theories of Risk. Santa Barbara, CA: Praeger; 1992:117-152.
55. Lythe M, de Cos GM, Mingallon M, Lensen A, Galloway C, Knox D. Explainable AI-building trust through understanding. Open Access Te Herenga Waka-Victoria University of Wellington. 2023. URL: https://openaccess.wgtn.ac.nz/articles/report/Explainable_AI_building_trust_through_understanding/24531433?file=43091017 [accessed 2026-02-21]
56. Rolison JJ, Hanoch Y, Freund AM. Perception of risk for older adults: differences in evaluations for self versus others and across risk domains. Gerontology 2019;65(5):547-559 [FREE Full text] [doi: 10.1159/000494352] [Medline: 30566940]
57. Ko EG, Nanayakkara S, Huff JEW. We need to avail ourselves of [GenAI] to enhance knowledge distribution?: Empowering older adults through GenAI literacy. 2025 Presented at: Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems; Jun 06, 2025; Yokohama, Japan p. 34. [doi: 10.1145/3706599.3720032]

Abbreviations

AI: artificial intelligence
BI: behavioral intention
GenAI: generative artificial intelligence
PEOU: perceived ease of use
PU: perceived usefulness
RTA: reflexive thematic analysis

Edited by N Schütz; submitted 18.Nov.2025; peer-reviewed by T Niu, H Zhang; comments to author 27.Dec.2025; revised version received 01.Jan.2026; accepted 30.Jan.2026; published 04.Mar.2026.

Please cite as:
Liu T, Pang P, Luo YT, McKay D, Buchanan G, Chang S
Evaluating the Use of Generative AI Videos for Health Self-Management of Older Adults: Mixed Methods Study
JMIR Aging 2026;9:e88005
URL: https://aging.jmir.org/2026/1/e88005
doi:10.2196/88005
PMID:

XSL•FO
RenderX

©Ting Liu, Patrick Pang, Yiming Taclis Luo, Dana McKay, George Buchanan, Shanton Chang. Originally published in JMIR Aging (https://aging.jmir.org), 04.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

360°-Based Cognitive-Motor Training System for Older Adults With Cognitive Impairment: User-Centered Design and Evaluation Study

Francesco Bigotto[1], BSc, MSc; Jonathan Panigada[1], BSc, MSc; Francesca Bruni[2], BSc, MSc; Valentina Mancuso[2], PhD; Silvia Cavedoni[1], BSc, MSc; Marco Stramba-Badiale[1,3], MD, PhD; Karine Marie Goulene[1,3], MD, PhD; Elisa Pedroli[1,2], PsyD; Silvia Serino[4], PhD

[1]Older People Rehabilitation and Cerebrovascular Medicine Research Laboratory, IRCCS Istituto Auxologico Italiano, Milano, Lombardia, Italy

[2]Department of Theoretical and Applied Sciences, eCampus University, Novedrate, Italy

[3]Department of Medicine, Neurology and Rehabilitation, IRCCS Istituto Auxologico Italiano, Milano, Italy

[4]Department of Psychology, Università degli Studi di Milano-Bicocca, Milano, Italy

Corresponding Author:
Francesco Bigotto, BSc, MSc
Older People Rehabilitation and Cerebrovascular Medicine Research Laboratory
IRCCS Istituto Auxologico Italiano
via Mosè Bianchi, 90
Milano, Lombardia, 20149
Italy
Phone: 39 3665457767
Email: fra.bigotto@gmail.com

Abstract

Background: Evidence suggests that older adults can improve dual-task (DT) performance through specific motor-cognitive training programs. Recent technological advancements have facilitated the development of novel rehabilitative DT methodologies. In particular, the DUAL-REHAB project exploits 360° technology to develop ecological, cost-effective DT exercises for clinical and home settings.

Objective: This study aimed to comprehensively understand the DUAL-REHAB project's end users (older adults with subjective memory complaints [SMC] and mild cognitive impairment [MCI]) and obtain feedback on an initial DUAL-REHAB mobile app prototype using a user-centered design approach.

Methods: The study used a 2-phase procedure. In the first phase (user requirements), we investigated end users' lifestyles, habits, perceived well-being, technology adoption, and cognitive and cognitive-motor DT training expectations. In the second phase (prototype evaluation), we developed and tested a DUAL-REHAB mobile app prototype to assess its functionality with end users.

Results: Overall, 14 participants were initially recruited for the study. The sample included 7 women with a mean age of 79 (SD 3.79) years and 7 men with a mean age of 82.43 (SD 5.65) years. One female participant was excluded for not meeting the criteria for either group; accordingly, the final sample study included 13 participants, with 8 categorized as MCI and 5 categorized as SMC. Participants reported structured daily routines with high autonomy, although some faced challenges with social connections. Perceived well-being was moderate across physical (mean 2.79, SD 0.97), psychological (mean 3.14, SD 0.86), and cognitive (mean 3.14, SD 0.53) domains. The perceived technology skills were low (mean 2.57, SD 0.94), with low digital literacy (mean 2.21, SD 0.89). Interest in cognitive training was high, with 92.31% (12/13) participants aware of its benefits and demonstrating strong interest (mean 4.15, SD 1.34) and perceived use (mean 4.15, SD 1.28). While only 46.15% (6/13) were aware of the possibility of DT training with technology, 92.31% (12/13) were willing to participate, and 84.61% (11/13) were open to home-based training. The preferred frequency was 2-3 days per week (63.64%, 7/11), with 10-20 minute sessions (8/11, 72.72%). Prototype evaluation revealed specific usability issues related to icon identification, device interaction, and navigation between training days.

Conclusions: We obtained valuable insights into the lifestyles, habits, and technological needs of older adults with MCI and SMC, which will guide DUAL-REHAB training development to align with user needs and capabilities. Our findings emphasize

XSL•FO
RenderX

the importance of simplified technological processes with intuitive interfaces. Additionally, structured interaction opportunities during clinical and at-home training could enhance motivation, facilitate timely problem resolution, and address participants' social needs.

(JMIR Aging 2026;9:e68032) doi:10.2196/68032

KEYWORDS

dual task; user centered; aging; usability; technology; cognitive-motor training

Introduction

Background

In recent decades, a growing focus has been on developing and testing effective strategies to mitigate cognitive decline associated with physiological and pathological aging [1]. The global population of older adults (65 years and older) is projected to reach 1.5 billion by 2050 [2], while Alzheimer disease prevalence may exceed 152 million cases [3]. These conditions significantly impact autonomy through various cognitive, behavioral, and motor symptoms [4-6], including deficits in dual-task (DT) performance, namely the ability to perform 2 tasks simultaneously. The DUAL-REHAB project uses 360° technology to address these challenges and develop ecological training environments for cognitive-motor exercises. Our approach is based on a user-centered design (UCD), prioritizing end user involvement throughout the design and development process. Accordingly, this study explores the needs of older adults with mild cognitive impairment (MCI) and subjective memory complaints (SMC) by investigating their lifestyles, habits, perceived well-being, technology adoption, and expectations regarding cognitive and cognitive-motor DT training, while also gathering their feedback through interaction with a prototype mobile app designed for home-based training.

Cognitive-Motor DT Training: State-of-the-Art and Current Challenges

Deficits in the ability to divide attention, or to perform 2 tasks simultaneously, referred to as DT performance, are early symptoms of dementia that evolve as the disease progresses [7,8]. An interplay exists between motor skills and cognition, with shared brain areas involved in attention, executive functions, and motor control, particularly in the prefrontal and temporal regions [9]. It is suggested that these networks can either process 1 task at a time [10] or that simultaneous engagement in a motor task and a cognitive task could result in interference caused by competing demands for attentional resources [11]. Despite different theoretical accounts, the effect appears to be more pronounced in individuals with cognitive impairments [12], leading to a deterioration in performance. Several studies have highlighted the close relationship between DT impairment alterations and motor dysfunction [13-15], falls [16-18], and a decline in daily activities observed in both normal and pathological aging [13,19]. Additionally, literature indicates that DT impairments could be considered a clinical marker of progression from cognitive impairment to dementia [9,20-22], underscoring the importance of DT skills for both daily functioning and clinical relevance.

There is convincing evidence suggesting that both older adults with Alzheimer disease and MCI can enhance their DT performance, motor skills, and cognitive abilities through specific DT motor-cognitive training programs [23-26]. Moreover, the recent advancements in technological tools have facilitated the development and implementation of novel rehabilitative DT methodologies [27-30]. A recent systematic review [30] indicated that an increasing body of evidence suggested the potential of exploiting technological devices to enhance DT skills across various age-related chronic conditions. Using motion-tracking systems, it is possible to reproduce users' movements within immersive virtual reality (VR) environments or on computer screens while participants are engaged in cognitive tasks. VR platforms offer realistic and immersive cognitive challenges and multisensory stimulation, enriching the interactive experience [31,32], also in home settings. For instance, Downey et al [33] recently exploited the potential of semi-immersive VR for older adults' rehabilitation, specifically targeting executive functions. This protocol used both laboratory and simulated real-world conditions to deliver at-home cognitive training focused on DT performance for both adults with normal-hearing and older adults who use hearing aids.

UCD: A Critical Approach for Effective Technology-Based Interventions

To ensure that technological interventions effectively meet the needs and capabilities of older adults, adopting a UCD approach is essential [34,35]. This methodological framework places end users at the core of the design and development process, systematically incorporating their perspectives, capabilities, and limitations at every stage. The basic premise of UCD is to directly involve end users in shaping the technology-based solutions they will ultimately use, to make them more easily adopted and effective. For instance, Hassandra et al [36] specifically investigated the acceptability, usability, and tolerability of using the VRADA (VR Exercise App for Dementia and Alzheimer's Patients) system compared to standard care physical and cognitive training. In another study by Nayak et al [37], 4 feasibility themes emerged from qualitative analyses: reasons for participation, difficulties using the technologies, engagement, and motivational values. UCD shares principles with related approaches like human-centered design and design thinking, emphasizing active user involvement, iterative design processes, and multidisciplinary collaboration to develop effective technological solutions through different techniques such as participatory design and rapid prototyping [38,39]. The adoption of UCD approaches for developing digital solutions in older adults has been extensively explored [40-43] and is even more crucial for individuals with cognitive decline [44-46]. Standard design

XSL•FO
RenderX

approaches often fall short when developing interventions for these populations because they fail to account for specific cognitive limitations and needs, reduced technology familiarity, and unique user contexts [47-49]. UCD addresses these challenges through structured, iterative cycles that follow 4 key stages: understanding the context of use, specifying user requirements, producing design solutions, and evaluating designs against requirements. When applied to older adults, as suggested by Nimmanterdwong et al [49], this process requires recruiting older participants, identifying their specific needs, incorporating measurable goals, creating rapidly testable prototypes, and adapting usability techniques to accommodate age-related limitations. In particular, UCD becomes particularly relevant for home-based interventions, as these solutions should be integrated into users' daily routines without direct professional support. Without careful consideration of user experience, even evidence-based interventions may fail due to incorrect use or poor adherence. For instance, Mehra et al [43] applied UCD methods to develop a tablet-based exercise app for older adults, revealing that iterative testing with end users was crucial for identifying unexpected interaction barriers and designing appropriate solutions.

The DUAL-REHAB Training: Addressing Research Gaps Through 360° Technology

Despite promising advancements in technology-enhanced DT training [27], several significant research gaps remain unaddressed. First, it remains less investigated whether training gains can be effectively transferred to novel contexts [50], particularly in daily life settings. Most training interventions were conducted in laboratory settings or using commercially available video games, limiting the possibility of generalizing the acquired skills to real-life contexts [29,51]. Second, due to potential mobility and/or physical constraints, transportation limitations, or insufficient facilities in nearby communities, older adults still face significant challenges accessing cognitive training programs. From this perspective, further exploitation of technological innovations could solve the challenges associated with traditional face-to-face interventions. Third, it is crucial to note that controlled randomized trials for older adults with SMCs are rarely included in DT research [21], despite facing a higher risk of developing MCI or dementia compared to those without such disorders [52-54]. To address these research gaps, the DUAL-REHAB project [55] aimed to exploit the potential of 360° technology, which is increasingly being adopted in assessment and cognitive training contexts [56-61]. 360° video represents a particularly accessible technology, offering significant advantages over other VR-based approaches. Since 360° cameras can easily capture real-life environments, they provide a practical solution for generating immersive and realistic scenarios where older adults can actively participate in cognitive-motor exercises, thus providing enhanced opportunities for transferring skills to real-life contexts. Regarding accessibility challenges, 360° content can be easily integrated into high-end applications for in-hospital sessions with therapists. The same format could be delivered through mobile apps and tablets for home sessions, overcoming transportation and facility access barriers. The DUAL-REHAB project will investigate the efficacy of this novel approach with both SMC and MCI populations, specifically including the underrepresented group of older adults with SMCs. The program's structure includes 10 days of training, each featuring approximately 10 DT exercises replicating real-life scenarios, implemented through both an immersive modality in clinical settings (using a head-mounted display and cycle ergometer) and a nonimmersive version for at-home use (via tablet app and portable cycle ergometer).

Study Objectives

To ensure that the DUAL-REHAB training program meets the needs and expectations of end users, as explained before, a UCD approach was adopted throughout the design and development process. This study represents the initial phases of a comprehensive UCD process, and it was designed with two primary objectives: (1) to obtain a comprehensive understanding of the end users of the DUAL-REHAB project, specifically targeting older individuals with SMC and older adults with MCI. This involved exploring various aspects, such as their lifestyles, habits, perceived well-being, technology adoption, and their needs and expectations regarding cognitive training and technologies for cognitive-motor DT training; and (2) to obtain feedback on an initial functional prototype of the DUAL-REHAB mobile app, specifically evaluating its overall functionality from the perspective of the target end users.

By systematically incorporating user perspectives throughout the development process, this study aims to ensure that the resulting training not only addresses clinical objectives but also aligns with the capabilities, preferences, and contexts of its intended users, potentially enhancing engagement, usability, and clinical outcomes.

Methods

Overview

A 2-step procedure was implemented to achieve our objectives. Phase 1 (user requirements) investigated end users' lifestyles, habits, well-being, technology adoption, and expectations regarding cognitive-motor DT training. Phase 2 (prototype evaluation) tested a DUAL-REHAB mobile app prototype with end users. We used semistructured interviews with open-ended and multiple-choice questions, and a think-aloud method to evaluate user-prototype interaction (Multimedia Appendix 1).

Participants and Recruitment

Overall, 14 participants were initially recruited for the study. The sample included 7 women with a mean age of 79 (SD 3.79) years and 7 men with a mean age of 82.43 (SD 5.65) years. Participants were recruited from patients and outpatients of the Department of Medicine, Neurology and Rehabilitation, Istituto di Ricovero e Cura a Carattere Scientifico Istituto Auxologico Italiano, Milan, Italy. In accordance with UCD principles, we recruited participants who represent our intended end users. Researchers contacted potential participants who met the following inclusion criteria: (1) age ≥60 years (no maximum age limit) to target the demographic most affected by age-related cognitive changes, (2) normal or corrected-to-normal vision to ensure participants could adequately view and interact with the visual elements of the technology, and (3) self-reported or

XSL•FO
RenderX

caregiver-reported cognitive concerns or memory complaints. Exclusion criteria were (1) debilitating medical, psychiatric, or neurological conditions that could affect performance; and (2) severe cognitive impairments indicated by a score <24 on the Italian version of the Mini-Mental State Examination [62,63].

To determine participants' cognitive status and establish diagnostic categories, an expert neuropsychologist conducted a comprehensive assessment using a battery of Italian-validated tests, including:

1. The Addenbrooke's Cognitive Examination III [64]
2. The forward version of Digit Span and Corsi Span [65]
3. The phonemic/semantic alternate fluency test [66]
4. The Short-Story Recall and Rey figure task [67]
5. The Frontal Assessment Battery [68]
6. The Trail Making Test [69]

These neuropsychological tests were used only for diagnostic purposes and were not used for analyses in any other part of the study.

MCI participants were identified based on self-reported or caregiver-reported cognitive decline, objective impairment in one or more cognitive domains, preserved functional abilities, and absence of dementia diagnosis [70,71], as documented in hospital medical records. SMC participants were identified based on self-reported memory complaints without objective cognitive decline [52], assessed through an ad hoc qualitative interview using established criteria [72].

Of the 14 initially recruited participants, 13 met the criteria for inclusion: 8 participants (3 women and 5 men) were categorized as MCI, and 5 participants (3 women and 2 men) as SMC. One female participant was excluded because she did not meet the criteria for either group.

Materials and Procedure

Phase 1: User Requirements

A semistructured interview with both open-ended and multiple-choice questions was developed specifically for this study to address the research objectives of the DUAL-REHAB project. Following UCD principles [34,35], both open-ended and multiple-choice questions were developed based on literature on technology acceptance in older adults [47,48], cognitive training preferences [23-26], and UCD approaches for older adults with cognitive impairments [44-46]. The content was subsequently reviewed by the multidisciplinary research team, including neuropsychologists and UCD experts, specifically considering the unique needs of older adults with MCI and SMC [49]. The procedure consisted of a semistructured interview administered by 1 experimenter, with a second experimenter observing and taking notes, conducted in a research room at the Department of Geriatrics and Cardiovascular Medicine of IRCCS – Istituto Auxologico Italiano in Milan. The interview comprised 2 parts: open-ended questions exploring lifestyles, habits, and socio-demographic details, followed by a questionnaire with multiple-choice responses covering perceived well-being, digital habits and skills, cognitive training knowledge, and DT cognitive-motor training expectations. While designed for independent completion, the interviewer assisted when needed or verbally administered questions in cases of fatigue to optimize time and resources. All multiple-choice questions used 5-point Likert scales ranging from "not at all" to "very much," "poor" to "excellent," or "never" to "every day." Each participant was welcomed inside the room with his or her companion, who was dismissed after explaining the procedure. The interview administration lasted between half an hour and 1 hour, depending on participants' reading speed, comprehension, response time, and the extent of their conversational digressions. Experimenters remained flexible regarding digressions as these could provide valuable information and help participants feel comfortable.

Phase 2: Prototype Evaluation

For Phase 2, a prototype of the DUAL-REHAB mobile app designed for DT motor-cognitive training was developed. It included DT exercises planned over 10 days in the final version. The experimenters created a Microsoft PowerPoint presentation displayed on an 11.2-inch iPad (A1566) and administered it to the participants. The presentation slides simulated the user journey from the home screen to the screen displaying DT exercises. The prototype was meticulously designed to ensure an optimal user experience, considering elements such as graphics, colors, user-system dialogue, and text adherence to International Organization for Standardization 9241 guidelines (ISO [International Organization for Standardization] 9241-11:2018, 2018). While the prototype was primarily based on established usability guidelines, some initial considerations from Phase 1 (see Phase 1: user requirements) informed basic design decisions, such as interface simplification and clear visual hierarchy.

Participants were instructed to interact with the slides, following written instructions in a binder. The instructions directed individuals to simulate starting the app and completing exercises scheduled for the first and second training days. The instructions provided were as follows:

Simulation of the first day of rehabilitation

1. Open and enter the app.
2. Launch the first day of exercises.
3. Select the first exercise of the day.
4. Listen to the exercise instructions.
5. Launch the exercise.

Simulation of the second day of rehabilitation

1. Open and enter the app.
2. Launch the second day of exercises.
3. Select the first exercise of the day.
4. Listen to the exercise instructions.
5. Listen to the instructions again.
6. Launch the exercise.

Each exercise was designed with the option to relisten to the instructions if a specific button was selected. Participants were instructed to listen again to the instructions of the exercise, even if they understood them initially, to evaluate whether they could successfully interact with this feature when experiencing comprehension difficulties. According to the think-aloud method

XSL•FO
RenderX

[73], participants were explicitly instructed to verbalize their thoughts and decision-making while interacting with the prototype. This cognitive verbalization technique requires participants to continuously "think aloud" by speaking everything that comes to mind during task completion, without filtering or interpreting their thoughts. This approach provides direct access to participants' cognitive processes and immediate reactions to the interface, capturing usability issues and misunderstandings as they occur rather than relying on retrospective reports that might be affected by memory limitations or post-hoc rationalizations. During the observation, researchers documented participants' verbalizations, hesitations, mistakes, and questions, along with their ability to complete tasks independently or with assistance. This real-time data collection method is particularly valuable for identifying specific interface elements that cause confusion or difficulty.

Ethical Considerations

The study was approved by the Ethics Committee of IRCCS – Istituto Auxologico Italiano, with protocol number approval 2023_01_31_10. All participants provided informed written consent before enrollment in the study. To ensure privacy, all potentially identifying information was removed, and data were kept strictly confidential. Participants received no financial compensation.

Results

Participants Characteristics

Selected participants were the same for this study's Phase 1 and Phase 2. The main socio-demographic information is presented in Table 1.

XSL•FO
RenderX

Table 1. Description of the characteristics of the older adults with mild cognitive impairment and subjective memory complaints (N=13).

Characteristics	Values
Cognitive conditions, n (%)	
MCI[a]	8 (61.54)
SMC[b]	5 (38.46)
Age groups (years), n (%)	
65-75	0 (0)
75-80	7 (53.85)
80-85	1 (7.69)
85-90	5 (38.46)
Sex, n (%)	
Female	6 (46.15)
Male	7 (53.85)
Marital status, n (%)	
Single	2 (15.39)
Married or common-law relationship	5 (38.46)
Divorced	0 (0)
Widowed	6 (46.15)
Education, n (%)	
Primary school diploma	0 (0)
Middle school diploma	4 (30.77)
Professional diploma	1 (7.69)
High school diploma	4 (30.77)
University degree or tertiary education	4 (30.77)
Occupational status, n (%)	
Employed	1 (7.69)
Unemployed	0 (0)
Retired	12 (92.31)
Living situation, n (%)	
Alone	5 (38.46)
With others	8 (61.54)

[a]MCI: mild cognitive impairment.

[b]SMC: subjective memory complaints.

Phase 1: User Requirements

Overview

This section presents findings from our semistructured interviews examining user characteristics, needs, and preferences. We systematically analyze different components of the interview data to understand potential users of the DUAL-REHAB training.

Lifestyle and Habits

In the initial interview phase, participants were presented with an open-ended question to investigate their lifestyle and habits: "Describe a typical day from waking up until going to bed, without unexpected or extraordinary events."

A structured checklist guided the interviewer in systematically collecting information across multiple domains, such as family and social relationships, leisure activities, daily habits, nutritional patterns, sleep quality, and physical activity levels.

Participants typically reported waking up around 7 AM; primary meals (lunch and dinner) were predominantly prepared and consumed at home. Between meals, participants are commonly engaged in activities including walking, reading, television viewing, or resting periods. Household management tasks and fulfillment of daily responsibilities were consistently reported. In total, 5 participants specifically noted attending to

XSL•FO
RenderX

health-related needs, either their own or their spouses', through medication management or attendance at medical appointments for more intensive therapeutic interventions.

Overall, participants demonstrated adequate cognitive and motor autonomy levels for daily functioning, with occasional exceptions due to specific conditions such as temporary limitations following falls or chronic pain. While social relationships were universally identified as fundamental, many participants reported challenges maintaining social connections due to diminished friendship networks or geographical separation from family members with competing obligations. Nevertheless, 7 participants specifically identified interactions with children and grandchildren as significant protective factors for emotional well-being.

These insights informed the development of 2 distinct "personas" (Multimedia Appendices 2 and 3), a UCD technique that synthesizes representative user behaviors and needs [74].

Perceived Well-Being

Participants' perception of well-being appeared moderate, as indicated by mean scores of 2.79 (SD 0.97) for physical health, 3.14 (SD 0.86) for psychological health, and 3.14 (SD 0.53) for cognitive health on a Likert scale ranging from 1 to 5.

Habits and Digital Skills

On a Likert scale ranging from 1 ("very bad") to 5 ("excellent"), participants demonstrated a relatively poor relationship with technology, as evidenced by a mean score of 2.57 (SD 0.94). Moreover, their perception of personal digital literacy (the ability to interact with digital tools) was even lower, with a mean score of 2.21 (SD 0.89).

Tables 2-4 present the frequency distributions for technological assistance received, device use (tablets, smartphones, smart televisions, and PCs), and digital service use (health, shopping, administrative, and entertainment), respectively.

Table 2. Frequency of technological assistance received by participants (N=13).

Technological assistance received	Values (N=13), n (%)
Never	1 (7.69)
Rarely	2 (15.39)
From time to time	5 (38.46)
Often	5 (38.46)
Everyday	0 (0)

XSL•FO
RenderX

Table 3. Frequency of use of technological devices (N=13).

Device use frequency	Values (N=13), n (%)
Tablet	
Never	7 (53.85)
Rarely	1 (7.69)
From time to time	1 (7.69)
Often	2 (15.38)
Everyday	2 (15.38)
Laptop	
Never	5 (38.46)
Rarely	0 (0)
From time to time	3 (23.08)
Often	0 (0)
Everyday	5 (38.46)
Smart-tv	
Never	9 (69.23)
Rarely	0 (0)
From time to time	0 (0)
Often	1 (7.69)
Everyday	3 (23.08)
Smartphone	
Never	5 (38.46)
Rarely	1 (7.69)
From time to time	0 (0)
Often	0 (0)
Everyday	7 (53.85)

XSL•FO
RenderX

Table 4. Frequency of use of digital services (N=13).

Digital service use frequency	Values (N=13), n (%)
Public services and administration	
Never	3 (23.08)
Rarely	3 (23.08)
From time to time	5 (38.46)
Often	2 (15.38)
Everyday	0 (0)
Health services	
Never	10 (76.92)
Rarely	1 (7.69)
From time to time	1 (7.69)
Often	1 (7.69)
Everyday	0 (0)
Shopping services	
Never	11 (84.62)
Rarely	0 (0)
From time to time	0 (0)
Often	2 (15.38)
Everyday	0 (0)
Email and messages	
Never	3 (23.08)
Rarely	3 (23.08)
From time to time	1 (7.69)
Often	3 (23.08)
Everyday	3 (23.08)
Social network	
Never	8 (61.54)
Rarely	0 (0)
From time to time	3 (23-08)
Often	1 (7.69)
Everyday	1 (7.69)
News	
Never	10 (76.92)
Rarely	0 (0)
From time to time	0 (0)
Often	2 (15.38)
Everyday	1 (7.69)
Audio or video streaming	
Never	7 (53.85)
Rarely	1 (7.69)
From time to time	2 (15.38)
Often	2 (15.38)
Everyday	1 (7.69)

XSL•FO
RenderX

Overall, findings suggested that smartphones dominated daily use, with 7 out of 13 participants (53.85%) using them every day. However, other devices show more sporadic patterns, with tablets and laptops being the least frequently used among our participants. On the other hand, it is interesting to note a strong heterogeneity across service categories. Health services showed the lowest penetration (10/13, 76.92% reported never using them), while email services, audio or video streaming, and social networks demonstrate more distributed use.

Two general evaluation parameters were considered in assessing technology: curiosity and use. While participants did not perceive themselves as particularly curious about technology (mean 2.71, SD 1.54, with 3 indicating "neutral"), they did recognize its usefulness in daily life (mean 3.57, SD 1.28).

Cognitive Training

Regarding cognitive training, we aimed to explore the end users' prior knowledge of the feasibility and methodologies involved in training cognitive functions and their level of interest in this area. Twelve participants demonstrated awareness of the potential for cognitive training to improve mental health, encompassing memory, attention, and perception. Eleven individuals also acknowledged the feasibility of incorporating cognitive and motor exercises. At the same time, 12 recognized that the objective of such training could be to enhance cognitive functions or impede their decline.

Interest in improving cognitive abilities was substantial, with a mean rating of 4.15 (SD 1.34), on a Likert scale from 1 to 5. Similarly, participants perceived the personal use of cognitive training, with a mean rating of 4.15 (SD 1.28). Only 1 participant expressed no interest in cognitive training. Furthermore, an interesting finding concerned participants' willingness to recommend cognitive training to others, with 11 individuals expressing readiness to endorse such training. Qualitative analysis of these responses revealed diverse recommendations for peers and younger generations, including children and grandchildren, underscoring the perceived value and broad applicability of cognitive training across age groups.

Cognitive-Motor DT Training With Technology

In implementing integrated DT cognitive-motor training programs, this section explored participants' knowledge and expectations regarding using technology, including tablets, smartphones, and laptops.

Among the participants, 6 individuals (approximately half of the sample) were aware of the possibility of performing cognitive-motor training with the support of technological tools. Despite this limited awareness, 12 participants expressed their willingness to participate in this type of training. However, when asked if they would also perform the training independently at home, the number of willing participants decreased slightly to 11.

Those willing to engage in home-based training were then asked to envision how much time they would dedicate to such a protocol, including weekly frequency, session duration, and overall training period. The results of these 3 questions are reported in Table 5.

Table 5. Time willing to dedicate to cognitive-motor dual task training at home with technology (N=11).

Home training time	Values (N=11), n (%)
Days per week	
1 day	1 (9.09)
2-3 days	7 (63.64)
3-4 days	3 (27.27)
5-6 days	0 (0)
7 days	0 (0)
Time per day	
10-15 minutes	4 (36.36)
15-20 minutes	4 (36.36)
20-30 minutes	1 (9.10)
30-45 minutes	0 (0)
45-60 minutes	2 (18.18)
Overall duration	
1 week	0 (0)
2 weeks	4 (36.36)
3-4 weeks	3 (27.28)
5 weeks	0 (0)
6 weeks	4 (36.36)

XSL•FO
RenderX

Overall, participants showed moderate willingness to dedicate time to home-based DT training. While daily and weekly commitment tended to range from moderate to low, the total duration indicated considerable availability, ranging from medium to high. Further analysis revealed that all 11 older adults would engage in the training for a maximum of 4 days per week. A total of 9 individuals would participate for 30 minutes daily, with the other 2 willing to extend it to 60 minutes. Additionally, 7 participants would undergo training for 2 to 4 weeks, with 4 participants willing to extend it to 6 weeks.

Overall, participants perceived using technology or performing cognitive-motor training with DT exercises as very useful (mean 4.31, SD 0.85) and moderately simple (mean 3.31, SD 1.38). We explored the participants' positive expectations and primary concerns regarding a technology-supported cognitive-motor training protocol. The summarized results are presented in Tables 6 and 7. As participants could provide multiple responses, the total frequencies may exceed the number of interviewees.

Table 6. The most common positive expectations regarding a cognitive-motor training protocol using technological tools at home (N=13).

Positive expectations	Values (N=13), n (%)
Enhanced autonomy	8 (61.54)
Improved self-esteem and self-efficacy	6 (46.15)
Cognitive function enhancement	11 (84.61)
Improved physical health status	3 (23.08)
Enhanced self-care capabilities and awareness	6 (46.15)
Reduced frequency of medical consultations	0 (0)
Improved social interaction	1 (7.69)
Time efficiency benefits	3 (23.08)
Financial benefits	2 (15.39)
Improved daily routines and lifestyle	2 (15.39)
None of the previous answers	0 (0)
Other	0 (0)

Table 7. The most common concerns regarding a cognitive-motor training protocol using technological tools at home (N=13).

Concerns	Values (N=13), n (%)
Privacy concerns	5 (38.46)
Adaptation to new routines	2 (15.39)
User errors in technology operation	8 (61.54)
Technical difficulties with devices	7 (53.85)
Exercise adherence challenges	4 (30.77)
Increased stress levels	3 (23.08)
Training requirements	4 (30.77)
Need for technical assistance	5 (38.46)
None of the previous answers	1 (7.69)
Other	0 (0)

Phase 2: Prototype Evaluation

Using an observation grid during prototype evaluation, all verbalizations were documented, including (1) comments, (2) complaints, or (3) direct requests for help, as well as observable behaviors such as (1) hesitations or (2) errors that indicated the participant's inability to complete the tasks independently. During the data collection, only inputs regarding the system characteristics were documented.

Data collected for each participant were compared to identify tasks causing the most significant difficulties. For each critical task, we determined whether others shared similar challenges. Identifying the most problematic tasks and system characteristics would provide valuable guidance for modifications in the final app version. This comparative analysis is represented in Table 8, which specifies the most critical tasks (and their instructions), synthesizes issues preventing task completion, indicates the number of participants encountering each difficulty, and presents potential solutions.

XSL•FO
RenderX

Table 8. Most critical tasks and respective instructions, problems, and solutions for the think-aloud protocol of the DUAL-REHAB prototype in mild cognitive impairment and subjective memory complaints participants.

Instructions, critical tasks, and problems	Solutions	Participants, n
Open the app		
Device interaction		
Excessive pressure during screen interaction	Implementation of device familiarization sessions for participants without prior touchscreen experience	1
Difficulty identifying the app icon	Elimination of extraneous visual elements from the home screen interface	3
Execution of instructions		
Protocol execution only following researcher's prompting	Development of an interactive tutorial for home screen navigation	1
Launch day 1		
Execution		
Day selection errors due to participants' tendency to search for calendar dates rather than ordinal day numbers in the training sequence	Consideration of replacing numerical day identifiers with calendar dates to align with participants' expectations	1
Proper interaction with the device		
Attentional dispersion due to interface complexity (eg, several buttons on the screen layout)	Reduction of interface elements to minimize cognitive load	1
Launch day 2		
Comprehension of the instructions		
Erroneous day selection	Provision of a tutorial or additional information on the appropriate protocol sequence navigation across different training days	1

Discussion

Principal Findings

The DUAL-REHAB project represents an innovative approach to cognitive-motor rehabilitation that exploits the potential of 360° technology to develop a novel DT training program specifically designed for older adults with SMC and MCI. Following UCD principles, this study revealed important insights about our target users and prototype functionality. Participants demonstrated high autonomy in daily activities and strong motivation to maintain independence despite reduced social interactions. They reported moderate satisfaction with their well-being across physical, psychological, and cognitive domains, with physical health rated lower than cognitive health. Participants expressed strong belief in cognitive training's potential and high interest in enhancing cognitive functions, with expected outcomes primarily focused on autonomy and self-care improvements.

Despite low self-reported technology proficiency, they showed a high willingness to participate in technology-based interventions and perceived such tools as valuable for cognitive-motor training. Participants preferred moderate training frequency (2-3 days/week) and shorter sessions (10-20 minutes). The prototype evaluation revealed some usability issues, with challenges primarily related to basic tablet interaction and navigation understanding. Key design improvements needed include simplifying the interface, redesigning icons, and implementing familiar temporal references for training day selection.

One of the main risks associated with an aging population and increased life expectancy is that older people may lose their autonomy and independence [75], requiring more assistance for longer periods. However, the target population we interviewed demonstrated a high degree of autonomy in daily activities and a strong desire to maintain independence, care for themselves, and improve their well-being. Participants maintained structured daily routines, primarily self-managed their meals, and engaged in various daily activities. This high level of autonomy could benefit the DUAL-REHAB training program, potentially facilitating the implementation of at-home training sessions where participants will access a nonimmersive version of the DT exercises via a tablet app combined with a portable cycle ergometer.

Despite this autonomy, a deeper look into participants' lifestyles showed fewer social interactions than expected, with boredom and loneliness frequently mentioned as negative aspects of their lives. This aligns with previous research identifying social isolation as a significant risk factor for cognitive decline [76]. Interestingly, health maintenance emerged as a meaningful activity for many participants, often becoming time-consuming but purposeful. This finding suggests that framing the DUAL-REHAB program as a health-promoting activity could leverage existing motivational patterns to enhance compliance. Given that 7 participants specifically identified interactions with

XSL•FO
RenderX

children and grandchildren as protective factors for emotional well-being, and 5 participants were already managing health-related needs within their family context, the training program could capitalize on these existing dynamics by incorporating caregivers as "active supporters." This approach would potentially address both training goals and social connection needs simultaneously. For our target populations specifically, family members should serve not merely as technical support but as cognitive scaffolding, a design consideration that might extend beyond the general principle for designing technologies for older adults.

Participants' self-perceptions of well-being deserve careful attention. Despite meeting clinical criteria for MCI or SMC, they reported moderate satisfaction with their cognitive health. Participants reported moderate levels across physical, psychological, and cognitive health domains when assessing their perceived well-being. Even if everyone included in the study is categorized according to MCI or SMC criteria, they feel quite satisfied with their cognitive well-being. It can be interpreted in 2 ways: a lack of self-awareness [77] or psychological acceptance that cognitive decline is inevitable. Notably, participants rated their physical well-being lower than their cognitive health, suggesting potential opportunities for interventions addressing both domains simultaneously, as the DUAL-REHAB approach aimed. Another aspect worth mentioning is their strong belief in the potential of cognitive training. Additionally, participants acknowledged the existence and the potential of motor-cognitive training. Consistently, they also expressed great interest in enhancing their cognitive functions. When discussing expected positive outcomes, our sample most frequently mentioned self-care-related aspects: autonomy, improvement in self-efficacy, cognitive improvements, self-awareness, and self-care.

Technology is the most delicate aspect investigated in this study. While participants reported low technology proficiency, digital literacy, and limited awareness of technology-based cognitive training possibilities, they simultaneously demonstrated high willingness to participate in such interventions (12 out of 13 participants). This apparent contradiction is coupled with their high perceived usefulness of technology for cognitive-motor training, suggesting that older adults' limited technology use likely might derive from motivational barriers rather than a lack of interest, in line with Lee and Coughlin [47], who established perceived value as a key factor in technology adoption. These findings suggest the opportunity to frame technology as a mean to maintain cognitive health could provide sufficient motivation to overcome initial resistance, a strategy supported by Wildenbos et al [48], who found that health-related technologies achieve higher adoption rates when they address specific health concerns valued by older adults.

Finally, it is interesting to highlight that the specific training preferences expressed by participants, favoring moderate frequency (2-3 days/week) and shorter sessions (10-20 minutes), align with typical training guidelines for older adults. These parameters should inform the DUAL-REHAB implementation strategy, as they represent a self-identified sustainable level of engagement that could optimize adherence while minimizing the burden on participants.

Regarding mobile app prototype and tablet use, testing revealed fewer problems than anticipated despite the participants' limited technological skills. Most issues were related to interaction with the tablet touchscreen (eg, pressure calibration), suggesting that the prototype design achieved appropriate simplicity for the target population. A brief training session on tablet interaction fundamentals could effectively address these basic usability barriers.

The 2 key navigation challenges emerged that require design modification. First, participants struggled to recognize the app icon, a finding consistent with research showing that older adults often have difficulty with abstract icons [78]. This necessitates redesigning icons and reducing competing visual elements on the home screen. Second, and more notably, participants had trouble selecting the correct training day, frequently defaulting to "day 1" when attempting subsequent sessions. This conceptual navigation barrier suggests that sequential numbering systems may be cognitively demanding for this population, particularly for individuals with MCI and SMC who may experience episodic memory difficulties and challenges with abstract sequencing. This supports the implementation of familiar temporal references (calendar dates) rather than abstract ordinal numbering. These findings inform some critical design decisions for the DUAL-REHAB: (1) implementing intuitive navigation with familiar conceptual models that accommodate cognitive limitations, and (2) developing a comprehensive initial training phase focusing on basic tablet interaction skills tailored to the cognitive profile of end users.

Finally, the personas developed from our interview data (Multimedia Appendices 2 and 3) will guide future design decisions for the DUAL-REHAB app. These personas reflect key patterns from our study: participants showed technological curiosity alongside dependency concerns, strong motivation for health maintenance, and clear preferences for structured training approaches. The personas will inform design decisions, such as balancing user independence with appropriate support features and involving family members as cognitive support rather than just technical help.

Conclusions

In conclusion, this study successfully addressed its primary objectives by adopting a UCD approach for the development of DUAL-REHAB training. We used 2 methods (a semistructured interview and the think-aloud method) to obtain valuable insights into the lifestyles, habits, and technological needs of older adults with MCI and SMC, as well as crucial feedback on the initial prototype of the DUAL-REHAB mobile app. Our findings will inform the further design and development of the training program, ensuring alignment with user expectations and technological capabilities. A critical need emerged for simplifying technological processes with intuitive and user-friendly interfaces. The importance of a comprehensive pretraining or familiarization phase to prepare participants for using the app was also highlighted. Furthermore, the necessity to consider and monitor participants' engagement and motivation levels throughout the training became evident. Sociality, well-being, and self-care were key areas for implementing and framing the training. Establishing opportunities for interaction

XSL•FO
RenderX

(in particular, with caregivers and family members) during clinical sessions and at-home training could help maintain motivation, promptly solve potential problems, and satisfy participants' need for social engagement. This approach could effectively address the participants' desire for social interactions and feeling supported in managing their well-being.

However, this study has several limitations. First, we did not collect data about the app from experts or caregivers [79]. Future studies should incorporate such feedback to comprehensively evaluate the DUAL-REHAB app's functionality and end users' expectations and lifestyles. This is particularly important considering the social needs that emerged from our findings. Moreover, our sample size represents a methodological limitation. While we followed established usability testing guidelines suggesting that 4-5 participants can identify most major usability issues [80,81], the specific cognitive characteristics of our population warrant consideration. Participants with MCI and SMC may experience varying cognitive difficulties that affect their technology interaction patterns, potentially requiring larger samples to capture the full spectrum of user experiences. Individual usability issues were often identified by small numbers of participants (n=1-3), but these findings align with established usability principles where even single-user observations can reveal critical design flaws. However, we recognize that the cognitive heterogeneity within our target populations limits the generalizability of these specific findings, and larger-scale usability testing would be essential to establish the prevalence and priority of identified issues before final implementation. In addition, it is crucial to note that although the semistructured interview protocol was developed based on existing literature and expert review, the absence of pilot testing represents another limitation. To address potential comprehension issues, interviewers assisted when participants encountered difficulties or fatigue, for example, by providing clarification or reading the multiple-choice questions aloud. Moreover, they allowed for spontaneous digressions that could yield valuable insights. Finally, another limitation of our usability testing regards the use of an initial prototype created with Microsoft PowerPoint rather than a fully functional application. This solution might have affected the ecological validity of our findings, and future testing of DUAL-REHAB solutions should involve more advanced technological prototypes, including, for example, interactive elements, to fully capture potential difficulties in interacting with the training application.

By addressing these limitations and building on the insights gained, future iterations of the DUAL-REHAB project can better address the needs of older adults with MCI and SMC, potentially improving their cognitive function and overall quality of life. Integrating UCD principles throughout development will allow a more effective and engaging cognitive training program tailored to its target users' unique needs and capabilities.

Funding

This work was supported by the Italian Ministry of Health - Ricerca Finalizzata (GR-2021-12374204).

Authors' Contributions

Francesco Bigotto and SS conceptualized the study and prepared the interview procedure. Francesca Bruni, JP, and EP prepared the study prototype of the app. Francesco Bigotto and JP administered the procedures and elaborated the collected data. Francesco Bigotto and SS proceeded with the first draft of the manuscript. Francesco Bigotto, JP, and SS wrote and edited the manuscript. MSB, EP, and SS contributed to the reviewed version of the manuscript. SS and EP supervised the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Microsoft Power Point presentation of the paper prototype of the DUAL-REHAB application, administered to potential end users.
[PPTX File , 7052 KB - aging_v9i1e68032_app1.pptx]

Multimedia Appendix 2
Personas that summarizes male users' prototypical behaviors and needs.
[PNG File , 21 KB - aging_v9i1e68032_app2.png]

Multimedia Appendix 3
Personas that summarizes female users' prototypical behaviors and needs.
[PNG File , 24 KB - aging_v9i1e68032_app3.png]

References

1. Di Nuovo S, De Beni R, Borella E, Marková H, Laczó J, Vyhnálek M. Cognitive impairment in old age: is the shift from healthy to pathological aging responsive to prevention? Eur Psychol 2020;25(3):174-185. [doi: 10.1027/1016-9040/a000391]

XSL•FO
RenderX

2. World Population Prospects 2019. United Nations Department of Economic and Social Affairs. 2019. URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un_2019_wpp_vol1_comprehensive-tables.pdf [accessed 2026-03-08]
3. Guerchet M, Prince M, Prina M. Numbers of people with dementia worldwide: an update to the estimates in the world Alzheimer report 2015. Alzheimer's Disease International. 2020. URL: https://www.alzint.org/resource/numbers-of-people-with-dementia-worldwide/ [accessed 2025-04-11]
4. Castrillo A, Olmos LMG, Rodríguez F, Duarte J. Gait disorder in a cohort of patients with mild and moderate Alzheimer's disease. Am J Alzheimers Dis Other Demen 2016;31(3):257-262 [FREE Full text] [doi: 10.1177/1533317515603113] [Medline: 26395024]
5. Savica R, Wennberg AM, Hagen C, Edwards K, Roberts RO, Hollman JH, et al. Comparison of gait parameters for predicting cognitive decline: the mayo clinic study of aging. J Alzheimers Dis 2016;55(2):559-567. [doi: 10.3233/jad-160697] [Medline: 27662317]
6. Mielke MM, Roberts RO, Savica R, Cha R, Drubach DI, Christianson T, et al. Assessing the temporal relationship between cognition and gait: slow gait predicts cognitive decline in the mayo clinic study of aging. J Gerontol A Biol Sci Med Sci 2013;68(8):929-937 [FREE Full text] [doi: 10.1093/gerona/gls256] [Medline: 23250002]
7. Perry RJ, Hodges JR. Attention and executive deficits in Alzheimer's disease. A critical review. Brain 1999;122(3):383-404. [doi: 10.1093/brain/122.3.383] [Medline: 10094249]
8. Baddeley AD, Baddeley HA, Bucks RS, Wilcock GK. Attentional control in Alzheimer's disease. Brain 2001;124(8):1492-1508. [doi: 10.1093/brain/124.8.1492] [Medline: 11459742]
9. Montero-Odasso M, Verghese J, Beauchet O, Hausdorff JM. Gait and cognition: a complementary approach to understanding brain function and the risk of falling. J Am Geriatr Soc 2012;60(11):2127-2136 [FREE Full text] [doi: 10.1111/j.1532-5415.2012.04209.x] [Medline: 23110433]
10. Pashler H. Dual-task interference in simple tasks: data and theory. Psych Bull 1994;116(2):220-244. [doi: 10.1037//0033-2909.116.2.220] [Medline: 7972591]
11. Kahneman D. Attention and Effort. Englewood Cliffs, NJ: Prentice Hall; 1973.
12. Kahya M, Moon S, Ranchet M, Vukas RR, Lyons KE, Pahwa R, et al. Brain activity during dual task gait and balance in aging and age-related neurodegenerative conditions: a systematic review. Exp Gerontol 2019;128:110756 [FREE Full text] [doi: 10.1016/j.exger.2019.110756] [Medline: 31648005]
13. Sobol NA, Hoffmann K, Vogel A, Lolk A, Gottrup H, Høgh P, et al. Associations between physical function, dual-task performance and cognition in patients with mild Alzheimer's disease. Aging Ment Health 2016;20(11):1139-1146. [doi: 10.1080/13607863.2015.1063108] [Medline: 26161932]
14. Hauer K, Pfisterer M, Weber C, Wezler N, Kliegel M, Oster P. Cognitive impairment decreases postural control during dual tasks in geriatric patients with a history of severe falls. J Am Geriatr Soc 2003;51(11):1638-1644. [doi: 10.1046/j.1532-5415.2003.51517.x] [Medline: 14687396]
15. Smith E, Cusack T, Cunningham C, Blake C. The influence of a cognitive dual task on the gait parameters of healthy older adults: a systematic review and meta-analysis. J Aging Phys Act 2017;25(4):671-686. [doi: 10.1123/japa.2016-0265] [Medline: 28253049]
16. Muir-Hunter SW, Wittwer JE. Dual-task testing to predict falls in community-dwelling older adults: a systematic review. Physiotherapy 2016;102(1):29-40. [doi: 10.1016/j.physio.2015.04.011] [Medline: 26390824]
17. Bayot M, Dujardin K, Dissaux L, Tard C, Defebvre L, Bonnet CT, et al. Can dual-task paradigms predict Falls better than single task? - A systematic literature review. Neurophysiol Clin 2020;50(6):401-440. [doi: 10.1016/j.neucli.2020.10.008] [Medline: 33176988]
18. Peters J, Lauinger A, Mayr M, Ginell K, Abou L. Dual-task assessments for predicting future falls in neurologic conditions: a systematic review. Am J Phys Med Rehabil 2024;103(6):554-560. [doi: 10.1097/phm.0000000000002452] [Medline: 38466165]
19. Brustio PR, Magistro D, Zecca M, Liubicich ME, Rabaglietti E. Fear of falling and activities of daily living function: mediation effect of dual-task ability. Aging Ment Health 2018;22(6):856-861 [FREE Full text] [doi: 10.1080/13607863.2017.1318257] [Medline: 28485621]
20. Montero-Odasso MM, Sarquis-Adamson Y, Speechley M, Borrie MJ, Hachinski VC, Wells J, et al. Association of dual-task gait with incident dementia in mild cognitive impairment: results from the gait and brain study. JAMA Neurol 2017;74(7):857-865 [FREE Full text] [doi: 10.1001/jamaneurol.2017.0643] [Medline: 28505243]
21. Åhman HB, Cedervall Y, Kilander L, Giedraitis V, Berglund L, McKee KJ, et al. Dual-task tests discriminate between dementia, mild cognitive impairment, subjective cognitive impairment, and healthy controls - a cross-sectional cohort study. BMC Geriatr 2020;20(1):258 [FREE Full text] [doi: 10.1186/s12877-020-01645-1] [Medline: 32727472]
22. Verghese J, Robbins M, Holtzer R, Zimmerman M, Wang C, Xue X, et al. Gait dysfunction in mild cognitive impairment syndromes. J Am Geriatr Soc 2008;56(7):1244-1251 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.01758.x] [Medline: 18482293]

XSL•FO
RenderX

23. Gallou-Guyot M, Mandigout S, Combourieu-Donnezan L, Bherer L, Perrochon A. Cognitive and physical impact of cognitive-motor dual-task training in cognitively impaired older adults: an overview. Neurophysiol Clin 2020;50(6):441-453. [doi: 10.1016/j.neucli.2020.10.010] [Medline: 33121880]
24. Fritz NE, Cheek FM, Nichols-Larsen DS. Motor-cognitive dual-task training in persons with neurologic disorders: a systematic review. Journal of Neurologic Physical Therapy 2015;39(3):142-153. [doi: 10.1097/npt.0000000000000090] [Medline: 26079569]
25. Parvin E, Mohammadian F, Amani-Shalamzari S, Bayati M, Tazesh B. Dual-task training affect cognitive and physical performances and brain oscillation ratio of patients with Alzheimer's disease: a randomized controlled trial. Front Aging Neurosci 2020;12:605317 [FREE Full text] [doi: 10.3389/fnagi.2020.605317] [Medline: 33424581]
26. Kuo HT, Yeh NC, Yang YR, Hsu WC, Liao YY, Wang RY. Effects of different dual task training on dual task walking and responding brain activation in older adults with mild cognitive impairment. Sci Rep 2022;12(1):8490 [FREE Full text] [doi: 10.1038/s41598-022-11489-x] [Medline: 35589771]
27. Tuena C, Borghesi F, Bruni F, Cavedoni S, Maestri S, Riva G, et al. Technology-assisted cognitive motor dual-task rehabilitation in chronic age-related conditions: systematic review. J Med Internet Res 2023;25:e44484 [FREE Full text] [doi: 10.2196/44484] [Medline: 37213200]
28. Pedroli E, Cipresso P, Greci L, Arlati S, Boilini L, Stefanelli L, et al. An immersive motor protocol for frailty rehabilitation. Front Neurol 2019;10:1078 [FREE Full text] [doi: 10.3389/fneur.2019.01078] [Medline: 31681149]
29. Gallou-Guyot M, Mandigout S, Bherer L, Perrochon A. Effects of exergames and cognitive-motor dual-task training on cognitive, physical and dual-task functions in cognitively healthy older adults: an overview. Ageing Res Rev 2020;63:101135. [doi: 10.1016/j.arr.2020.101135] [Medline: 32768594]
30. Faria AL, Cameirão MS, Couras JF, Aguiar JRO, Costa GM, Bermúdez i Badia S. Combined cognitive-motor rehabilitation in virtual reality improves motor outcomes in chronic stroke - a pilot study. Front Psychol 2018;9:854 [FREE Full text] [doi: 10.3389/fpsyg.2018.00854] [Medline: 29899719]
31. Gaggioli A, Greci L, Arlati S, Stramba-Badiale M, Pedroli E, Colombo D, et al. "Positive Bike" - An immersive biking experience for combined physical and cognitive training of elderly patients. Annual Review of CyberTherapy and Telemedicine 2017;15:196-199 [FREE Full text]
32. Makmee P, Wongupparaj P. VR cognitive-based intervention for enhancing cognitive functions and well-being in older adults with mild cognitive impairment: behavioral and EEG evidence. Psychosoc Interv 2025;34(1):37-51. [doi: 10.5093/pi2025a4] [Medline: 39781014]
33. Downey R, Gagné N, Mohanathas N, Campos JL, Pichora-Fuller KM, Bherer L, et al. At-home computerized executive-function training to improve cognition and mobility in normal-hearing adults and older hearing aid users: a multi-centre, single-blinded randomized controlled trial. BMC Neurol 2023;23(1):378 [FREE Full text] [doi: 10.1186/s12883-023-03405-1] [Medline: 37864139]
34. Harte R, Glynn L, Rodríguez-Molinero A, Baker PMA, Scharf T, Quinlan LR, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
35. Ergonomics of human-system interaction: human-centred design for interactive systems. Switzerland: International Organization for Standardization (ISO); 2010. URL: https://www.iso.org/standard/77520.html [accessed 2026-03-08]
36. Hassandra M, Galanis E, Hatzigeorgiadis A, Goudas M, Mouzakidis C, Karathanasi EM, et al. A virtual reality app for physical and cognitive training of older people with mild cognitive impairment: mixed methods feasibility study. JMIR Serious Games 2021;9(1):e24170 [FREE Full text] [doi: 10.2196/24170] [Medline: 33759797]
37. Nayak A, Alhasani R, Kanitkar A, Szturm T. Dual-task training program for older adults: blending gait, visuomotor and cognitive training. Front Netw Physiol 2021;1:736232 [FREE Full text] [doi: 10.3389/fnetp.2021.736232] [Medline: 36925571]
38. Göttgens I, Oertelt-Prigione S. The application of human-centered design approaches in health research and innovation: a narrative review of current practices. JMIR Mhealth Uhealth 2021;9(12):e28102 [FREE Full text] [doi: 10.2196/28102] [Medline: 34874893]
39. Altman M, Huang TTK, Breland JY. Design thinking in health care. Prev Chronic Dis 2018;15:E117 [FREE Full text] [doi: 10.5888/pcd15.180128] [Medline: 30264690]
40. Tuena C, Pedroli E, Trimarchi PD, Gallucci A, Chiappini M, Goulene K, et al. Usability issues of clinical and research applications of virtual reality in older people: a systematic review. Front Hum Neurosci 2020;14:93 [FREE Full text] [doi: 10.3389/fnhum.2020.00093] [Medline: 32322194]
41. Chu CH, Biss RK, Cooper L, Quan AML, Matulis H. Exergaming platform for older adults residing in long-term care homes: user-centered design, development, and usability study. JMIR Serious Games 2021;9(1):e22370 [FREE Full text] [doi: 10.2196/22370] [Medline: 33687337]
42. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]
43. Mehra S, Visser B, Cila N, van den Helder J, Engelbert RHH, Weijs PJM, et al. Supporting older adults in exercising with a tablet: a usability study. JMIR Hum Factors 2019;6(1):e11598 [FREE Full text] [doi: 10.2196/11598] [Medline: 30707106]

XSL•FO
RenderX

44. Meiland F, Innes A, Mountain G, Robinson L, van der Roest H, García-Casal JA, et al. Technologies to support community-dwelling persons with dementia: a position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil Assist Technol 2017;4(1):e1 [FREE Full text] [doi: 10.2196/rehab.6376] [Medline: 28582262]
45. Wargnier P, Benveniste S, Jouvelot P, Rigaud AS. Usability assessment of interaction management support in LOUISE, an ECA-based user interface for elders with cognitive impairment. Technol Disabil 2018;30(3):105-126. [doi: 10.3233/tad-180189]
46. Bogza LM, Patry-Lebeau C, Farmanova E, Witteman HO, Elliott J, Stolee P, et al. User-centered design and evaluation of a web-based decision aid for older adults living with mild cognitive impairment and their health care providers: mixed methods study. J Med Internet Res 2020;22(8):e17406 [FREE Full text] [doi: 10.2196/17406] [Medline: 32442151]
47. Lee C, Coughlin JF. Perspective: older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J Prod Innov Manag 2015;32(5):747-759 [FREE Full text] [doi: 10.1111/jpim.12176]
48. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
49. Nimmanterdwong Z, Boonviriya S, Tangkijvanich P. Human-centered design of mobile health apps for older adults: systematic review and narrative synthesis. JMIR Mhealth Uhealth 2022;10(1):e29512. [doi: 10.2196/29512] [Medline: 35029535]
50. Lemke NC, Werner C, Wiloth S, Oster P, Bauer JM, Hauer K. Transferability and sustainability of motor-cognitive dual-task training in patients with dementia: a randomized controlled trial. Gerontology 2019;65(1):68-83. [doi: 10.1159/000490852] [Medline: 30041173]
51. Delbroek T, Vermeylen W, Spildooren J. The effect of cognitive-motor dual task training with the biorescue force platform on cognition, balance and dual task performance in institutionalized older adults: a randomized controlled trial. J Phys Ther Sci 2017;29(7):1137-1143 [FREE Full text] [doi: 10.1589/jpts.29.1137] [Medline: 28744033]
52. Warren SL, Reid E, Whitfield P, Moustafa AA. Subjective memory complaints as a predictor of mild cognitive impairment and Alzheimer's disease. Discov Psychol 2022;2(1):13. [doi: 10.1007/s44202-022-00031-9]
53. Seo EH, Kim H, Choi KY, Lee KH, Choo IH. Association of subjective memory complaint and depressive symptoms with objective cognitive functions in prodromal Alzheimer's disease including pre-mild cognitive impairment. J Affect Disord 2017;217:24-28. [doi: 10.1016/j.jad.2017.03.062] [Medline: 28380342]
54. Jonker C, Geerlings MI, Schmand B. Are memory complaints predictive for dementia? A review of clinical and population-based studies. Int J Geriatr Psychiatry 2000;15(11):983-991. [doi: 10.1002/1099-1166(200011)15:11<983::aid-gps238>3.0.co;2-5] [Medline: 11113976]
55. Pedroli E, Bruni F, Mancuso V, Cavedoni S, Bigotto F, Panigada J, et al. Training cognitive functions using DUAL-REHAB, a new dual-task application in MCI and SMC: a study protocol of a randomized control trial. Technologies 2025;13(3):96. [doi: 10.3390/technologies13030096]
56. Serino S, Repetto C. New trends in episodic memory assessment: immersive 360° ecological videos. Front Psychol 2018;9:1878 [FREE Full text] [doi: 10.3389/fpsyg.2018.01878] [Medline: 30333780]
57. Pieri L, Moro V, Gambina G, Facci E, Amato S, Romano D. The 360° mental screening (MS-360°): a screening test for an ecological assessment of everyday cognitive functioning. 2021 Presented at: Book of Abstracs 25th Annual International CyberPsychology, CyberTherapy & Social Networking Conference (CYPSY25); Sep 13, 2021; Geneva, Switzerland URL: https://www.research.unipd.it/retrieve/e14fb26e-eb43-3de1-e053-1705fe0ac030/BOOK%20OF%20ABSTRACTS_CYPSY25_9Settembre_Definitivo%20%281%29.pdf
58. Jiang Z, Zhang Y, Chiang FK. Meta‐analysis of the effect of 360‐degree videos on students' learning outcomes and non‐cognitive outcomes. Brit J Educational Tech 2024;55(6):2423-2456. [doi: 10.1111/bjet.13464]
59. Negro Cousa E, Brivio E, Serino S, Heboyan V, Riva G, de Leo G. New frontiers for cognitive assessment: an exploratory study of the potentiality of 360° technologies for memory evaluation. Cyberpsychol Behav Soc Netw 2019;22(1):76-81 [FREE Full text] [doi: 10.1089/cyber.2017.0720] [Medline: 30421985]
60. Grewe P, Kohsik A, Flentge D, Dyck E, Botsch M, Winter Y, et al. Learning real-life cognitive abilities in a novel 360°-virtual reality supermarket: a neuropsychological study of healthy participants and patients with epilepsy. J Neuroeng Rehabil 2013;10(1):42 [FREE Full text] [doi: 10.1186/1743-0003-10-42] [Medline: 23618596]
61. Zulkiewicz BA, Boudewyns V, Gupta C, Kirschenbaum A, Lewis MA. Using 360-degree video as a research stimulus in digital health studies: lessons learned. JMIR Serious Games 2020;8(1):e15422 [FREE Full text] [doi: 10.2196/15422] [Medline: 31904577]
62. Magni E, Binetti G, Bianchetti A, Rozzini R, Trabucchi M. Mini-mental state examination: a normative study in Italian elderly population. Eur J Neurol 1996;3(3):198-202. [doi: 10.1111/j.1468-1331.1996.tb00423.x] [Medline: 21284770]
63. Measso G, Cavarzeran F, Zappalà G, Lebowitz BD, Crook TH, Pirozzolo FJ, et al. The mini‐mental state examination: normative study of an Italian random sample. Dev Neuropsychol 1993;9(2):77-85. [doi: 10.1080/87565649109540545]
64. Pigliautile M, Chiesi F, Stablum F, Rossetti S, Primi C, Chiloiro D, et al. Italian version and normative data of Addenbrooke's cognitive examination III. Int Psychogeriatr 2019;31(2):241-249. [doi: 10.1017/s104161021800073x] [Medline: 30021668]

XSL•FO
RenderX

65. Monaco M, Costa A, Caltagirone C, Carlesimo GA. Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. Neurol Sci 2013;34(5):749-754. [doi: 10.1007/s10072-012-1130-x] [Medline: 22689311]
66. Costa A, Bagoj E, Monaco M, Zabberoni S, De Rosa S, Papantonio AM, et al. Standardization and normative data obtained in the Italian population for a new verbal fluency instrument, the phonemic/semantic alternate fluency test. Neurol Sci 2014;35(3):365-372. [doi: 10.1007/s10072-013-1520-8] [Medline: 23963806]
67. Carlesimo GA, Buccione I, Fadda L, Graceffa A, Mauri M, Lorusso S, et al. Normative data of two memory tasks: short-story recall and Rey's figure. Nuova Riv Neurol 2002;12(1):1-13 [FREE Full text]
68. Appollonio I, Leone M, Isella V, Piamarta F, Consoli T, Villa ML, et al. The frontal assessment battery (FAB): normative values in an Italian population sample. Neurol Sci 2005;26(2):108-116. [doi: 10.1007/s10072-005-0443-4] [Medline: 15995827]
69. Siciliano M, Chiorri C, Battini V, Sant'Elia V, Altieri M, Trojano L, et al. Regression-based normative data and equivalent scores for trail making test (TMT): an updated Italian normative study. Neurol Sci 2019;40(3):469-477. [doi: 10.1007/s10072-018-3673-y] [Medline: 30535956]
70. Albert MS. Changes in cognition. Neurobiol Aging 2011;32(1):S58-S63 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2011.09.010] [Medline: 22078174]
71. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the national institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
72. Abdulrab K, Heun R. Subjective memory impairment. A review of its definitions indicates the need for a comprehensive set of standardised and validated criteria. Eur Psychiatry 2008;23(5):321-330. [doi: 10.1016/j.eurpsy.2008.02.004] [Medline: 18434102]
73. Jaspers MWM, Steen T, Van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
74. Nimmanterdwong Z, Boonviriya S, Tangkijvanich P. Human-centered design of mobile health apps for older adults: systematic review and narrative synthesis. JMIR Mhealth Uhealth 2022;10(1):e29512 [FREE Full text] [doi: 10.2196/29512] [Medline: 35029535]
75. Rechel B, Grundy E, Robine JM, Cylus J, Mackenbach JP, Knai C, et al. Ageing in the European Union. Lancet 2013;381(9874):1312-1322. [doi: 10.1016/S0140-6736(12)62087-X] [Medline: 23541057]
76. Courtin E, Knapp M. Social isolation, loneliness and health in old age: a scoping review. Health Soc Care Community 2017;25(3):799-812. [doi: 10.1111/hsc.12311] [Medline: 26712585]
77. Piras F, Piras F, Orfei MD, Caltagirone C, Spalletta G. Self-awareness in mild cognitive impairment: quantitative evidence from systematic review and meta-analysis. Neurosci Biobehav Rev 2016;61:90-107. [doi: 10.1016/j.neubiorev.2015.10.002] [Medline: 26639655]
78. Leung R, McGrenere J, Graf P. Age-related differences in the initial usability of mobile device icons. Behav Inform Technol 2011 Sep;30(5):629-642. [doi: 10.1080/01449290903171308]
79. Cristiano A, Musteata S, De Silvestri S, Bellandi V, Ceravolo P, Cesari M, et al. Older adults' and clinicians' perspectives on a smart health platform for the aging population: design and evaluation study. JMIR Aging 2022;5(1):e29623 [FREE Full text] [doi: 10.2196/29623] [Medline: 35225818]
80. Guillén-Climent S, Garzo A, Muñoz-Alcaraz MN, Casado-Adam P, Arcas-Ruiz-Ruano J, Mejías-Ruiz M, et al. A usability study in patients with stroke using MERLIN, a robotic system based on serious games for upper limb rehabilitation in the home setting. J Neuroeng Rehabil 2021;18(1):41 [FREE Full text] [doi: 10.1186/s12984-021-00837-z] [Medline: 33622344]
81. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? Hum Factors 1992;34(4):457-468. [doi: 10.1177/001872089203400407]

Abbreviations

DT: dual-task
ISO: International Organization for Standardization
MCI: mild cognitive impairment
SMC: subjective memory complaint
UCD: user-centered design
VR: virtual reality

XSL•FO
RenderX

Edited by R Bjarnadottir; submitted 29.Oct.2024; peer-reviewed by Y Li, P Makmee; comments to author 10.Mar.2025; accepted 20.Oct.2025; published 13.Mar.2026.

Please cite as:
Bigotto F, Panigada J, Bruni F, Mancuso V, Cavedoni S, Stramba-Badiale M, Goulene KM, Pedroli E, Serino S
360°-Based Cognitive-Motor Training System for Older Adults With Cognitive Impairment: User-Centered Design and Evaluation Study
JMIR Aging 2026;9:e68032
URL: https://aging.jmir.org/2026/1/e68032
doi:10.2196/68032
PMID:

©Francesco Bigotto, Jonathan Panigada, Francesca Bruni, Valentina Mancuso, Silvia Cavedoni, Marco Stramba-Badiale, Karine Marie Goulene, Elisa Pedroli, Silvia Serino. Originally published in JMIR Aging (https://aging.jmir.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Adoption and Use of a Smart Home Connected Care System by Older Adults: Mixed Methods Study

Laura Pemberton[1], PhD; Michelle Carter[1], PhD; Nikolay Mehandjiev[2], PhD

[1]Alliance Manchester Business School, University of Manchester, Manchester, United Kingdom

[2]School of Social Sciences, University of Birmingham, Dubai, United Arab Emirates

Corresponding Author:
Laura Pemberton, PhD
Alliance Manchester Business School
University of Manchester
Oxford Road
Manchester, M13 9PL
United Kingdom
Phone: 44 0161 306 6000
Email: laura.pemberton@manchester.ac.uk

Abstract

Background: Smart home technologies are often promoted as a means to support independent living and alleviate pressure on health and care systems. However, adoption among older adults remains low, and little is known about how passive, ambient technologies are experienced over time.

Objective: This study aimed to examine the adoption of a Connected Care System among older adults and their caregivers in a real-world deployment, testing the applicability of the Unified Theory of Acceptance and Use of Technology model and identifying design considerations for future systems.

Methods: A mixed methods study was conducted with 91 older adults who used the Connected Care System for approximately 6 months. Quantitative survey data were analyzed using structural equation modeling to assess predictors of behavior intention. Thematic analysis of open-ended survey responses was used to explore participants' experiences and perceptions of the system.

Results: Perceived ease of use (β=0.075; *P*=.02) and facilitating conditions (β=0.232; *P*<.001) were significant predictors of behavior intention, while perceived usefulness and social influence were not. Thematic analysis revealed that older adults often conceptualized usefulness in emotional and relational terms, such as peace of mind and caregiver reassurance. Adoption was frequently described as a shared process involving family members. Key barriers included low confidence in setup and reliability concerns, while valued features included passive monitoring, increased independence, and reduced caregiver burden.

Conclusions: The findings suggest that conventional adoption models require adaptation to account for passive, caregiver-mediated smart home systems. Emotional and relational benefits, rather than task efficiency, drive perceived usefulness. Design recommendations are offered to support user-friendly, personalized, and relationally aware smart home technologies that align with the realities of aging in place.

(JMIR Aging 2026;9:e76012) doi:10.2196/76012

KEYWORDS

smart home technologies; older adults; technology adoption; UTAUT; Unified Theory of Acceptance and Use of Technology; mixed methods; aging in place; caregiver support; passive monitoring

Introduction

Background

Smart home technologies are often positioned as part of the solution to the global challenges of population aging, supporting not only older adults but also family members who provide informal care [1]. As the proportion of older adults continues to grow, there is an urgent need for innovative solutions that promote independent living while reducing the emotional and financial pressures on caregivers, who already bear substantial—yet often underreported—burdens [2-4]. Projections indicate that by 2050, one in four people in the United Kingdom will be aged 65 years or older, compared with one in five in 2019 [5]. This demographic shift is expected to place increasing

XSL•FO
RenderX

strain on health and social care systems, reinforcing the need to investigate technology-enabled models of care. Within this context, smart home technologies such as the Connected Care System (CCS) developed by a major UK telecommunications company offer a practical example of how digital infrastructure could be leveraged to support aging in place.

Despite their potential, the uptake of smart home technologies remains relatively low among older adults [6-8]. Recent reports show that only 54% of adults over the age of 75 years in the United Kingdom use the internet regularly [9], reflecting broader barriers to digital inclusion. This gap between the growing need for scalable care solutions and the slow adoption of digital technologies underscores the importance of examining how integrated systems, such as the CCS, can address barriers while supporting aging in place. In this study, the CCS consisted of a mobile app used by the informal caregiver, an Amazon Alexa device, and several plug-in motion detector sensors placed in the older adult's home. The sensors captured patterns of movement and relayed this information through the mobile app, enabling caregivers to receive notifications as the older adult moved about their house. The inclusion of a voice assistant also provided a familiar interface for interaction, while the mobile app acted as the main point of access for caregivers to monitor activity. Unlike professionally monitored telecare systems, the CCS was designed for informal caregiving contexts, where family members received the notifications directly and used them to support independent living. It also supports daily activity tracking, providing caregivers with insights into routines and well-being over time. These functionalities align with ongoing policy and practice goals to extend independent living while addressing the costs associated with formal care provision [10].

However, existing research on smart home technologies for older adults has significant limitations. Much of the evidence comes from laboratory settings or short-term deployments [11], leaving open questions about sustained use in everyday life. For example, the SmartSenior project evaluated a 36-device smart home system over an average of 45 days, finding that acceptance improved once initial technical issues were resolved [12]. Similarly, a 2-year study in a care home context showed that older adults' attitudes became more positive over time [13]. More recent studies highlight the importance of trust and perceptions of security in influencing adoption [14]. Yet, long-term, real-world evaluations of smart home systems in private households remain rare. This study addresses this gap by examining the adoption and sustained use of the CCS in older adults' own homes over an extended period, offering novel insights into how these systems function beyond pilot or institutional settings.

This study makes several contributions to the field of technology adoption and smart home system design. First, it evaluates theoretical adoption constructs in the context of passive, ambient smart home technologies, identifying potential gaps where traditional models fail to capture the unique dynamics of older adult users. Second, by integrating qualitative insights, the study refines existing adoption frameworks, offering a more holistic understanding of the interplay between behavior intention, perceived usefulness (PU), social influence, and ease of use [15]. A key practical contribution of this research is the development of 6 design requirements based on empirical findings. These requirements provide actionable insights for technology developers, health care providers, and policymakers seeking to design more accessible, personalized, and user-friendly smart home systems. By bridging the gap between technological potential and real-world adoption, this study supports the broader goal of enabling older adults to maintain independence, improve well-being, and enhance quality of life while aging in place.

Theoretical Framework

The Unified Theory of Acceptance and Use of Technology (UTAUT) model [15] is often used to predict technology adoption, but it mainly focuses on task-oriented use rather than on in-home or passive use by older adults [16]. While the UTAUT model has been applied to various groups, studies focusing on older adults are scarce [17,18]. When applied, UTAUT has been found sufficient to explain older adults' technology use [19]. The limited research may be due to the unique needs of older adults, such as physical or cognitive limitations, which can alter the relevance of the model's core constructs [20,21]. For example, older adults may prioritize ease of use and immediate utility [22] more than younger users. Findings regarding the influence of social factors are also inconsistent [11,23], with some research suggesting a strong influence from family and peers, while other studies find it less significant. Empirical findings also point to mixed results when UTAUT is applied to smart technologies with older adults; for example, one study found that performance expectancy strongly predicted intention to use, but effort expectancy, social influence, and facilitating conditions were not significant, despite being core constructs of the model [24]. However, in other cases, social influence has growing importance due to extra reliance on caregivers [25,26]. These inconsistencies highlight the need for more nuanced research on this demographic. This paper also responds to a call for more thorough research using the UTAUT model that also incorporates qualitative data through a longitudinal study [27].

Objective

The objective of this study is to investigate the adoption and use of the CCS among older adults during a 6-month real-world deployment. Specifically, we examine the key factors that shape adoption behavior and assess how well-established constructs from the UTAUT explain technology engagement in this context. In addition, this study seeks to identify the perceived benefits and barriers experienced by both older adults and their caregivers. By integrating quantitative measures with qualitative insights, the research provides a nuanced understanding of how smart home technologies are taken up in everyday life and highlights design implications for systems that aim to support independent living and aging in place.

Research Model and Hypotheses

This study tested the UTAUT model [28] in the context of older adults who had used the CCS for 6 months. The model examines the key determinants influencing older adults' behavior intention to adopt and use smart home technologies.

XSL•FO
RenderX

Behavior intention refers to an individual's readiness to perform a specific behavior—in this case, the intention to use smart home technologies [29]. The model includes 4 independent constructs: perceived ease of use (PEOU), PU, social influence, and facilitating conditions. PEOU is defined as the degree to which an individual believes that using a system is free from effort [11]. PU refers to the extent to which a person believes that using the system will enhance their performance or daily living [28]. Together, these constructs capture both the effort and value components of technology acceptance. Social influence reflects the extent to which individuals perceive that people important to them think they should use a technology [28]. This factor is particularly relevant in the context of older adults, whose technology decisions are often shaped by family members and caregivers. Facilitating conditions refer to the degree to which individuals believe that adequate organizational and technical support exists to enable system use [28]. For older adults, these conditions include access to digital infrastructure, such as Wi-Fi, and availability of guidance or assistance during setup [30].

Based on the UTAUT model and the context of this study, the following hypotheses were proposed:

Hypothesis 1: Perceived ease of use positively influences behavior intention to use a smart home system.

Hypothesis 2: Perceived usefulness positively influences behavior intention to use a smart home system.

Hypothesis 3: Social influence positively influences behavior intention to use a smart home system.

Hypothesis 4: Facilitating Conditions positively influence behavior intention to use a smart home system.

Methods

Research Design

This study used a longitudinal observational design to examine older adults' adoption and use of smart home technologies within the context of a real-world deployment. The trial was designed and conducted by a UK telecommunications company, which managed all aspects of recruitment, technology deployment, and data collection. Recruitment was open to households in the United Kingdom that included an older adult receiving informal care from a family member, friend, or neighbor. No other eligibility restrictions were applied. The study commenced in October 2020 but experienced delays due to the COVID-19 pandemic, which influenced the timing and logistics of technology setup in participants' homes.

A total of 393 participants took part in the trial, with survey completion being entirely voluntary. The smart home technology, the CCS, was installed in participants' homes with only remote support provided by the telecommunications company. The system included an Amazon Alexa device, motion-detecting smart plugs, and a mobile app that enabled informal caregivers to receive notifications and monitor household activity.

All survey data were anonymized prior to being shared with the researchers. The only information available to the researcher was an overview of participants' ages, unlinked to any survey responses.

Ethical Considerations

The data analyzed in this study were collected as part of a trial conducted by a large technology company in the United Kingdom. Ethical approval for the original data collection was obtained through the company's internal ethics and governance review processes, in accordance with national regulations for industry-led research involving noninvasive technologies. All participants in the original trial provided informed consent to participate. Participation in the survey component was voluntary. This study involved secondary analysis of fully anonymized survey data. No personally identifiable information was available to the research team. In line with UK Health Research Authority guidance on Research Ethics Committee review, research limited to the secondary use of previously collected, non-identifiable information is generally excluded from NHS/HSC REC review where participants are not identifiable to the research team [31]. The authors had permission from the telecommunications company to access and analyze the anonymized dataset.

Instrument Development

The researcher involved in this study played a pivotal role in designing the survey questionnaire tailored to meet the trial's specific objectives. Subsequently, the questionnaire was handed over to the organization, the study's industry partner, for further refinement to ensure alignment with their requirements. Following this collaborative effort in questionnaire design, the researcher's involvement with the trial concluded, with no direct participation in the trial's administration. The survey was developed using pre-existing surveys from already recognized models of technology adoption. Constructs selected to be included were primarily based on the UTAUT model [28]. In our instrument development, we opted for the inclusion of open-ended survey questions to provide participants with a platform to articulate their experiences with the CCS. These can also be seen in Multimedia Appendix 1. This design enabled the collection of both quantitative and qualitative data regarding participants' interactions with the CCS, providing a naturalistic view of technology adoption behavior among older adults.

Qualitative Analysis

The results reported in this study were derived from an inductive thematic analysis following the 6-stage reflexive approach proposed by Braun and Clarke [32,33]. The analysis used an open-coding strategy, in which participants' excerpts were first coded line by line and then iteratively grouped into broader categories and overarching themes. This process was conducted using NVivo (version 12, QSR International) software to assist in organizing, coding, and managing the data while enabling the identification of recurring patterns and relationships within the dataset.

The thematic analysis involved the following six stages: (1) becoming familiar with the data through repeated reading and annotation, (2) generating initial codes through open coding of

XSL•FO
RenderX

participant excerpts, (3) collating related codes into potential themes, (4) reviewing and refining these themes against the coded data and entire dataset, (5) defining and naming the final set of themes to ensure clarity and distinctiveness, and (6) producing the analytic narrative and selecting illustrative extracts for reporting. Thematic analysis, as a structured and methodical yet flexible approach, was particularly suited to analyzing the open-ended survey responses, allowing for both depth and breadth in interpreting participants' experiences with the CCS [34].

Quantitative Analysis

Structural equation modeling (SEM) was used as a social science tool to examine complex relationships among latent constructs and observed variables [35]. In this research, SEM was applied to explore factors influencing attitudes and behaviors towards technology adoption. By combining measurement and structural models, SEM validated instruments and tested causal relationships, making it suitable for small sample sizes like this study's 91 participants. SEM allowed for efficient analysis by integrating multiple relationships within one model, enhancing reliability despite limited data. The analysis followed standard operational steps: (1) data screening for missing values and outliers; (2) evaluation of the measurement model using Cronbach α and composite reliability (>0.70), average variance extracted (AVE) (>0.50), and discriminant validity via Fornell–Larcker and Heterotrait-Monotrait ratio criteria; and (3) estimation of the structural model using bootstrapping (5000 resamples) to assess path coefficients, *t* values, and *P* values. Model fit was assessed using SRMR, d_ULS, d_G, and NFI indices, which are appropriate for partial least squares SEM. Predictive accuracy was examined through R^2 values of endogenous constructs. To address limitations, qualitative findings were used to complement SEM results, offering a nuanced view of attitudes towards smart home technologies and balancing potential dataset shortcomings.

Mixed Methods Analysis

This research used a mixed method approach where quantitative and qualitative research methods were used concurrently, taking an exploratory approach. This research compared the findings of the quantitative and qualitative data collected [15,36]. By comparing and contrasting results from different data sources, triangulation enabled researchers to corroborate findings and uncover nuances that may not have been apparent with either method alone [37].

Results

Overview

A total of 91 participants responded to the survey between October 15, 2021, and October 29, 2021. A total of 14% of the participants were younger than 65 years, 16% between the ages of 65 and 75 years, 43% were aged between 75 and 85 years, and finally 27% were older than 85 years across the whole of the study. Additional demographic data were not reported, as only an anonymous dataset was shared.

Quantitative Analysis

The results of the UTAUT model are shown in Figure 1 and Table 1. PU, PEOU, facilitating conditions, and social influence explained 67% of the variance in behavior intention.

XSL•FO
RenderX

Figure 1. Results of Connected Care System data within the Unified Theory of Acceptance and Use of Technology model of technology adoption. BI: behavior intention; FC: facilitating condition; PEOU: perceived ease of use; PU: perceived usefulness; SI: social influence.

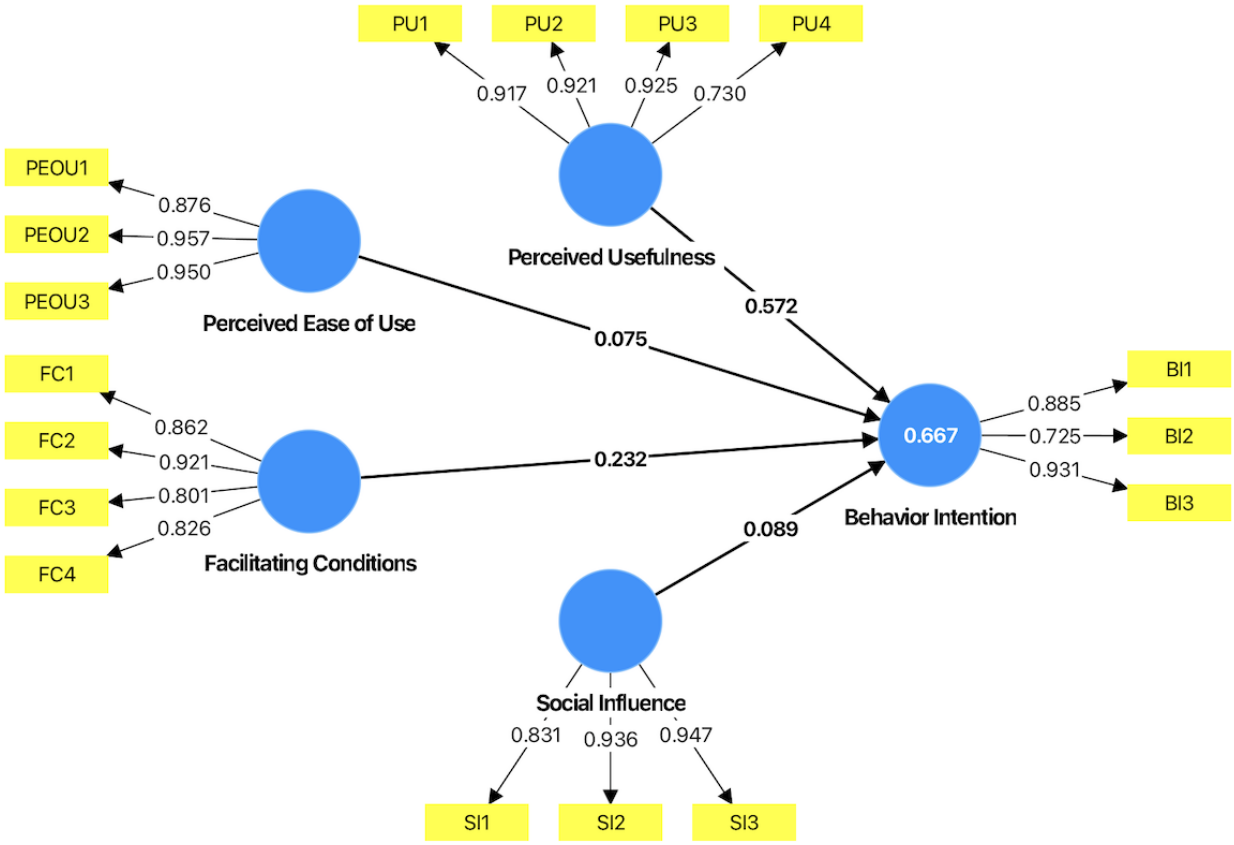


Table 1. Unified Theory of Acceptance and Use of Technology model results (P<.05 and T>1.96).

	Path coefficients	Sample, mean (SD)	Bootstrapped T value	*P* value	Hypothesis
Hypothesis 1	0.075	0.243 (0.095)	2.441	.02	Supported
Hypothesis 2	0.572	0.071 (0.111)	0.673	.50	Not supported
Hypothesis 3	0.089	0.101 (0.075)	1.186	.24	Not supported
Hypothesis 4	0.232	0.561 (0.099)	5.764	<.001	Supported

When measuring for discriminant validity of the model, we completed 2 separate tests of this. First, by considering the latent variable correlations (Table 2), and second, by considering the Heterotrait-Monotrait ratio ratios (Table 3). The recommended threshold for measuring the latent variable correlations is that all values should be below 0.85 [38]. Values above this can indicate that the constructs are too similar and do not measure distinct enough concepts.

Table 2. Unified Theory of Acceptance and Use of Technology latent variable correlations.

	Behavior intention	Facilitating conditions	Perceived ease of use	Perceived usefulness
Facilitating conditions	0.614	__[a]	—	—
Perceived ease of use	0.548	0.697	—	—
Perceived usefulness	0.766	0.502	0.471	—
Social influence	0.507	0.473	0.467	0.477

[a]Not applicable.

XSL•FO
RenderX

Table 3. Unified Theory of Acceptance and Use of Technology model discriminant validity (Heterotrait-Monotrait ratio matrix).

	Behavior intention	Facilitating conditions	Perceived ease of use	Perceived usefulness
Facilitating conditions	0.675	__[a]	—	—
Perceived ease of use	0.622	0.768	—	—
Perceived usefulness	0.890	0.520	0.506	—
Social influence	0.570	0.485	0.503	0.524

[a]Not applicable.

The study also tested for construct reliability measures such as Cronbach α and composite reliability (Table 4). The reliability coefficients are all above the recommended threshold of 0.70 [39], indicating satisfactory reliability. AVE assesses the amount of variance captured by the latent variables relative to measurement error. AVE values above 0.50 are generally considered acceptable [40], indicating that constructs explain more variance than measurement error.

Table 4. Construct reliability.

	Cronbach α	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted
Behavior intention	0.808	0.844	0.887	0.726
Facilitating conditions	0.880	0.923	0.915	0.729
Perceived ease of use	0.919	0.940	0.949	0.861
Perceived usefulness	0.897	0.910	0.930	0.769
Social influence	0.891	0.914	0.932	0.821

Qualitative Results

This results section will describe the most relevant themes found in the thematic analysis, but a full breakdown of themes, including excerpts and frequency counts, is included in Multimedia Appendix 1.

Acquiring Technology

The largest theme emphasized the significant effort required for technology acquisition (n=46), with participants expressing that they would need substantial support, as one participant explained, “Too much/not capable by myself” (ID74).

Followed by participants’ desire for effortless technology acquisition (n=33), highlighting how important it was that the process of obtaining the CCS felt straightforward. Some participants described the experience as smooth and independent, for example, “It would be pretty straightforward to purchase this product” (ID84). A related theme was the desire to improve knowledge and understanding of the technology (n=37). Participants wanted to make more informed and autonomous decisions. For example, one noted simply, “Understanding of devices” (ID35), pointing to the value of accessible information during the acquisition stage. Transparency of cost information also emerged as a recurring concern (n=30). Participants wanted clarity about setup and ongoing costs to support their decision-making. One participant explained, “I need a simple one price for set up and then one monthly price please, nothing for extras” (ID25). Finally, family advice and support played a key role in acquisition processes (n=23). Many participants relied on family members’ expertise and guidance. As one participant put it, “My nephew as he is very good at explaining how to use the technology in a way I understand” (ID3). In several cases, decisions were made jointly, with either the older adult or the caregiver leading the practical acquisition.

Attitude to Technology

A prominent theme to emerge was exposure to the latest technology (n=39), which reflected participants’ curiosity and enthusiasm for innovation. Many were eager to explore and keep up with new technological developments, positioning themselves as proactive users. As one participant explained, “To know what new technologies are available and their potential” (ID72). Another described their involvement in testing initiatives: “I wanted to test the new innovation solution from … and wanted to see the latest and be part of the testing. I really enjoyed it” (ID25). This theme highlights how engagement with emerging technology is not solely the domain of younger groups; some older adults also actively seek new experiences.

Closely linked to this was a positive perception of technology (n=34), with participants frequently describing technology as beneficial, particularly for safety and well-being. For instance, one participant stated, “It won’t be going anywhere. It’s been great and I feel much safer” (ID17), while another simply noted “excellent” (ID15). These accounts illustrate the positive value placed on technology as part of everyday life. By contrast, a theme of neutral perception of technology (n=26) was also evident. Here, participants described the technology as “ok,” reflecting moderate satisfaction or indifference rather than enthusiasm or resistance. This suggests a pragmatic stance, where technology is seen as functional but not transformative.

A further theme concerned the role of technology in everyday life (n=24). Participants emphasized the significance of technology in supporting independence and enabling aging in place. For example, one participant noted, “I think it will play

XSL•FO
RenderX

an important role in allowing older people to live independent lives in their own homes rather than going into care" (ID22), while another commented, "It will play an effective role in keeping me independent but safe" (ID52). These perspectives position technology as an important, sometimes essential, element of daily living for older adults.

Not all participants, however, saw technology as necessary. The theme of limited need for new technology (n=16) captured participants who stated that they did not perceive a need for the system, suggesting varying levels of perceived necessity based on individual circumstances. A smaller but noteworthy theme was updating current technology systems (n=4), where participants discussed adopting new systems as a way to replace or integrate with their existing devices. Finally, some participants expressed uncertainty about the role of technology (n=15), revealing ambivalence and mixed views. For example, one participant commented, "not too keen" (ID10), while another remarked, "It's not really impinged on my life. It's more for the kids" (ID29). These accounts reflect uncertainty about both the relevance and impact of the technology.

Older Adults' Perspective

A central theme concerns mobility decline (n=53), which emerged as one of the most significant challenges for participants. Many described increasing difficulties in moving around their homes, particularly for essential activities such as eating or using the bathroom. As one participant put it, "Getting around the house for food, toilet, etc" (ID52). Another noted, "I need a walking aid as I'm not good on my feet and can't walk far" (ID64). This theme underscores the ways in which declining mobility shapes daily routines and heightens vulnerability. Several participants expressed concerns about falls (n=15), including fears of collapsing alone at home, as one participant put it, "fear of failing" (ID29). Linked to this were concerns around memory decline (n=14), with participants describing difficulties remembering tasks, appointments, or daily routines. For example, one participant said, "My memory isn't so good so it's not so easy for me to remember what I'm supposed to be doing and when" (ID22). Closely related to this was the theme of health decline (n=25), which captured broader concerns about deteriorating physical capacities, including sight and general health. For example, one participant noted, "Further deterioration in my sight" (ID36), reflecting a heightened awareness of age-related change and its implications for independence.

Alongside these physical changes, participants also described challenges with solo tasks and chore difficulty (n=22), highlighting the impact of functional decline on everyday activities. For instance, one participant referred to a "decline in day to day functioning" (ID28), while another noted difficulties with "carrying things around house safely" (ID39). Emotional and social dimensions were also prominent. The theme of loneliness and isolation (n=18) revealed feelings of spending much of the day alone and a desire for companionship. One participant described, "feel lonely at times" (ID41), and another emphasized, "The ability to be able to have human interaction and company" (ID20). Similarly, mental health decline (n=9) captured experiences of boredom, demotivation, and low mood, as in the comment, "Boredom, not seeing anybody" (ID80).

A smaller theme captured older adults' reluctance to rely on assistive systems (n=13). Some participants articulated a preference for maintaining autonomy, for example: "but I want to be able to do things myself and ask for help if needed not rely on it" (ID25). This theme points to the tension between independence and technological support. Another theme, growing need for emergency response (n=8), reflected participants' recognition of the benefits of continuous monitoring and rapid assistance. As one participant explained, "Need more monitoring / access to immediate help" (ID70). These accounts illustrate the desire for reassurance that support is available in emergencies. Finally, the theme of reduced frequency of welfare checks (n=7) highlighted the perceived value of technology in supporting independent living while easing caregiver burden. One participant commented, "Less calls from relatives checking on me as they can see from the app I'm moving" (ID43). This was seen to benefit both older adults, who experienced fewer interruptions, and caregivers, who could monitor well-being remotely.

System Benefits

Participants described multiple benefits of the CCS, particularly in enhancing well-being and quality of life. A prominent theme was the older person's peace of mind (n=45), reflecting reassurance from knowing that assistance is available in case of emergencies. One participant explained, "I am always seeking ways of giving me more peace of mind that if something happens to me, I won't be alone" (ID68), while another noted, "have some relief knowing that my family are able to be notified if something happens" (ID86). Similarly, the theme of caregiver's peace of mind (n=26) highlighted how the system alleviates family concerns: "it stops my family worrying about me, knowing where I am or if I am alright" (ID25).

The CCS also supported enhanced independence (n=26), empowering older adults to live autonomously at home. Participants emphasized this benefit, as one stated, "Being able to live alone longer" (ID45), and another added, "I want to stay in my own home and live independently and these solutions will help me do that" (ID64). Complementing this was the theme of feelings of connectedness (n=13), illustrating how the system fosters familial bonds: "secure I know my daughter is looking after me" (ID10) and "Knowing my family are feeling more connected to me" (ID7).

Practical utility was also emphasized through the theme of usefulness (n=20). Participants described how the system supported daily activities and safety, for instance: "it is like the front door lock, I can't do without, even if I think I am safe!!" (ID25). Relatedly, caregiver support (n=16) reflected how the CCS facilitates informal caregiving by providing monitoring and communication: "It's just used by my family to watch over me" (ID64). Improved communication (n=15) further emphasized connectivity and prompt response in emergencies, as one participant noted, "Confidence there is a way of contacting people and being monitored in case I have issues" (ID28). Several participants valued discreet monitoring and devices (n=34 and n=11, respectively), highlighting unobtrusive

XSL•FO
RenderX

design and passive oversight: “I know that I can be seen to be active without feeling like I’m being watched” (ID60) and “I think the devices all look very smart and fit into my home nicely” (ID68). Finally, ease of use (n=5) and reminders (n=4) were seen as additional benefits, with participants describing the system as simple and supportive: “Very simple” (ID70) and “Little reminders are great” (ID3).

System Issues

The most prevalent theme concerned the reliability of the emergency support system (n=49). Participants consistently expressed worries about their ability to contact someone in case of emergencies, highlighting the need for a dependable and responsive system. One participant stated, “Making sure I can contact someone if I need to” (ID40), while another emphasized, “Fear of not being able to contact someone in the event of emergency” (ID6). These concerns underline that the perceived trustworthiness of emergency support is central to adoption and continued use.

Closely related was the theme of requiring additional and properly functioning devices (n=20). Participants highlighted the importance of reliable call buttons and integrated devices to enhance system effectiveness. As one participant noted, “Our call button never worked and we would definitely want one of those” (ID74), while another commented, “More but needs greater integration with other devices” (ID78). This theme reflects broader concerns about technology reliability and the need for well-functioning, complementary devices. The theme of assistance with setup and maintenance (n=9) captured participants’ need for practical support when installing or maintaining the system. Many reported struggling with new technology or complex setup procedures. For example, one participant explained, “Simple advice as I do not understand new technology” (ID61), and another stated, “I would not understand it or be able to do it without the help of others” (ID37). These accounts highlight that effective user support is critical for ensuring proper use.

Participants also emphasized dependence on additional services or devices (n=7), reflecting a desire for customizable solutions that meet individual needs and preferences. One participant noted, “Video system is likely to be more useful” (ID6), while another stated, “Also if additional types of device become available, they would be considered” (ID74). This theme underscores the value of flexibility in system design to accommodate varying requirements. Smaller themes included financial aid (n=14), with participants recognizing that affordability may influence adoption, and adjustment periods required to trust technology (n=3), exemplified by a participant remarking, “It took a bit of time for me to trust it wasn’t watching everything I do” (ID60). These reflect practical and psychological considerations that can affect engagement with the system.

Discussion

Principal Findings

This study examined how older adults perceive and engage with the CCS, a passive smart home technology designed to support well-being through ambient monitoring. Using a mixed methods approach, we identified several insights that extend existing theoretical models and carry implications for the design, implementation, and policy contexts of care technologies. Notably, our findings suggest that UTAUT [28] requires adaptation when applied to ambient systems. For older adults who engage with the system indirectly, the boundary between PU and behavior intention becomes less meaningful. Consistent with prior research [16,41-44], PU and social influence did not reach statistical significance, highlighting the complexity of measuring technology adoption in this demographic. This finding suggests that PU may require reconceptualization in the context of passive or ambient smart home technologies. For older adults, usefulness may not be derived from direct task performance or system efficiency—as traditionally defined in UTAUT—but from indirect, emotional, and relational benefits, such as reassurance, safety, and reduced caregiver burden. Redefining PU to incorporate these affective and social dimensions may therefore provide a more accurate representation of technology value in aging in place contexts.

A key factor shaping these findings is the passive nature of the CCS. Unlike technologies used actively on a day-to-day basis, participants rarely interacted directly with the system or the data it collected. Instead, benefits were realized indirectly through caregivers or family members, who could monitor the older adult’s well-being without interrupting their daily routines. Consequently, participants’ perceptions of usefulness were primarily emotional rather than functional [14], encompassing peace of mind, reassurance, and ease of access to safety information. As one participant noted, the comfort provided by the system is valuable even if it is not actively used. This underscores the importance of evaluating technology not only in terms of functional outcomes but also in terms of emotional and relational value.

The study also highlights the critical role of social influence and caregiver mediation in technology adoption. Although quantitative analysis found no significant link between social influence and behavior intention, qualitative findings emphasized that family and caregivers play a central role in adoption and ongoing use. This reflects prior research in health settings that integrates social, emotional, and contextual factors into technology acceptance models [45] and suggests that adoption should be conceptualized as a relational process involving shared decision-making and co-dependency within household care arrangements. These dynamics raise broader concerns about digital equity, particularly in how older adults rely on others to access and benefit from smart home technologies.

It is important to acknowledge the potential impact of the small sample size (n=91) on the statistical power of the SEM and the generalizability of the findings. While SEM provides a robust framework for evaluating complex relationships, smaller sample sizes can increase the risk of Type II errors, potentially obscuring significant effects and limiting the reliability of path estimates. In this study, the limited sample may have contributed to the nonsignificant relationships observed between PU, social influence, and behavior intention. Furthermore, the participant pool reflects a specific demographic context—older adults

XSL•FO
RenderX

engaged with a passive smart home system—which may not capture the full diversity of older populations in different care settings or cultural contexts. These considerations highlight that, while the findings provide meaningful insights into the relational and emotional dimensions of technology adoption, caution is warranted in extrapolating the results beyond this sample. Future research should aim to replicate these analyses with larger and more diverse cohorts to confirm the robustness of the observed patterns.

Furthermore, this study contributes to methodological advancements in technology adoption research by demonstrating the value of mixed methods approaches. While UTAUT has been widely applied using survey-based models [46], our research underscores the necessity of integrating qualitative thematic analysis to contextualize adoption behaviors [45]. The discrepancies between quantitative and qualitative findings, particularly regarding PU and social influence, highlight the limitations of relying solely on numerical measures to understand technology adoption in aging populations. By combining statistical modeling with in-depth user perspectives, this study provides a richer, more nuanced understanding of technology adoption, reinforcing the importance of mixed methods research in studying complex user interactions with emerging technologies [23,47].

Ultimately, this research contributes to ongoing debates about how emerging technologies are transforming domestic life and care infrastructures. By integrating qualitative thematic analysis with a widely used adoption model, the study provides a socially embedded account of technology uptake—reflecting the realities of aging, caregiving, and digital participation. The findings highlight the importance of designing technologies not only for individual users but also for the networks and systems that support them, with consideration for emotional reassurance, relational dynamics, and passive engagement.

Comparison With Prior Work

The joint table developed from the mixed method analysis can be seen in Multimedia Appendix 1 [36]. The table presents the qualitative and quantitative findings alongside the construct relationships found in our CC model. The final column of the table presents the meta inference, and the distinction of each inference as being in discordance, expansion or confirmation [48].

Behavior Intention

The quantitative findings of this study provide strong evidence that behavior intention is a significant driver of technology adoption among older adults, as indicated by the high explanatory power in our UTAUT model. This suggests that older adults are not only receptive to smart home technologies like the CCS but also actively recognize their potential to enhance quality of life, increase independence, and provide peace of mind. These findings align with previous studies demonstrating that when technology addresses key psychological and emotional needs, older adults are more willing to adopt and use it [11,49]. The qualitative data further support this conclusion, revealing a strong enthusiasm for continued use of the CCS. Participants particularly valued the peace of mind, security, and independence it provided, not just for themselves but also for their caregivers. However, while the overall attitude toward adoption was positive, qualitative insights also highlight key challenges and barriers that must be addressed to ensure sustained, reliable, and comfortable usage. These include concerns about reliability, usability, and the role of caregivers in facilitating technology adoption. This suggests that technology adoption among older adults is not simply a matter of willingness—rather, it is shaped by contextual factors, such as how well the system integrates into their existing support networks and whether it is perceived as empowering rather than burdensome.

Perceived Ease of Use

Our quantitative findings confirm that PEOU is a significant predictor of behavior intention. This aligns with the technology acceptance model, reinforcing the well-established importance of usability in driving technology adoption among older adults [29]. High outer loadings for PEOU indicators suggest that participants consistently found the CCS to be intuitive and learnable, supporting the need for simplified and user-friendly interfaces. The qualitative findings, however, provide additional depth to this perspective. While participants acknowledged that the system was relatively easy to use, they also emphasized the importance of customizability and flexibility. Many older adults desired the ability to mix and match features based on their individual needs, which were largely influenced by health status, mobility, and caregiving arrangements [50]. This suggests that ease of use is not a one-size-fits-all construct—rather, it should be viewed as a dynamic, individualized factor that varies based on personal circumstances and support systems. A notable insight from the qualitative data is that older adults often rely on caregivers or family members for technology setup, troubleshooting, and maintenance. This highlights an important but underexplored dimension of PEOU: the role of the caregiver in facilitating ease of use. Some participants saw the caregiver as the primary user, while they themselves merely benefited from the system's presence. This raises a critical question: how does ease of use function in systems with shared users, particularly when the technology benefits both the caregiver and the older adult?

Perceived Usefulness

The qualitative data strongly emphasized the usefulness of the CCS, with participants highlighting key benefits such as peace of mind, safety, and independence. These benefits were not just functional but also emotional, reinforcing findings that technology for older adults often plays a psychosocial role beyond simple task efficiency [11,49]. However, the quantitative results indicate that PU did not significantly predict behavior intention. This discrepancy suggests that the traditional conceptualization of PU in technology adoption models may not fully capture its meaning for older adults. Specifically, usefulness in this context extends beyond functionality—it encompasses emotional security, social connectedness, and caregiver support. Furthermore, the qualitative findings suggest that customizability plays a key role in perceptions of usefulness. Older adults highly valued the ability to tailor the system to

XSL•FO
RenderX

their specific needs, which were highly individualized based on health, mobility, and social support structures [50].

Social Influence

The relationship between social influence and behavior intention was the weakest in our model, a finding that aligns with previous studies questioning the role of social influence in older adults' technology adoption [23,47]. This suggests that broader social networks (eg, friends, community) have minimal impact on adoption decisions, while immediate family and caregivers play a much more significant role [49,51]. The qualitative findings reinforce this idea, revealing that older adults primarily rely on caregivers and family members for support, setup, and decision-making. Interestingly, participants often saw the caregiver as the true decision-maker, which may explain the weak correlation between social influence and behavior intention—older adults do not see the choice as their own, but rather as a necessity dictated by their caregivers. This suggests a reconceptualization of social influence is needed in models of technology adoption for older adults, particularly when technology is used collaboratively.

Facilitating Conditions

The quantitative findings confirm that facilitating conditions are a strong predictor of behavior intention, reinforcing existing literature that emphasizes the importance of external support in technology adoption [52]. However, our qualitative findings reveal a more complex picture of facilitating conditions. The moderate positive correlation between facilitating conditions and PEOU (0.697) suggests that the availability of resources and support contributes to users' perceptions of how easy the system is to use. This underscores the importance of providing adequate support structures, clear instructions, and accessible help resources to enhance the PEOU of smart home systems for older adults. Particularly in the case of older adults, who may be starting their adoption process from a point of less knowledge than younger users of smart home systems. A key theme emerging from the data is that older adults perceive facilitating conditions as external rather than internal. Many participants felt that they did not need personal knowledge of the system, as they relied on family members to provide expertise and manage technology-related tasks. This suggests that the traditional framing of facilitating conditions in UTAUT may need revision—rather than measuring an individual's perceived ability to use technology, future models should account for interdependent technology adoption, where one party (the caregiver) facilitates access for another (the older adult). Additionally, the qualitative data highlight the importance of knowledge acquisition—older adults expressed a need for better education and guidance on how to use CCS effectively. This aligns with prior research indicating that older adults require ongoing education and training to improve confidence in new technology [52].

Implications for Design and Practice

The findings from this study inform a set of 6 interrelated design requirements that enhance the adoption, usability, and effectiveness of smart home technologies such as the CCS for older adults and their caregivers. Together, these requirements highlight the importance of designing systems that are not only technically reliable and easy to use but also emotionally supportive, socially connected, and adaptable to users' evolving needs. They address key adoption factors, including behavior intention, facilitating conditions, social influence, PU, and ease of use. Table 5 summarizes how each construct informed specific design requirements, linking empirical insights to actionable design guidance.

Table 5. Design requirements showing the meta inference made across each of the constructs, the system adaptation required to fit in the smart home context, and the requirements.

UTAUT[a] construct	Meta-inference	Requirement based on adaptation
Perceived usefulness	Older adults and caregivers value peace of mind, security, and independence provided by CCS[b]. High behavior intention indicates a strong interest in adoption.	RQ1: Provide clear and tangible benefits emphasizing safety, independence, and emotional reassurance.
Perceived ease of use	Older adults often lack confidence and require structured onboarding to navigate the system effectively.	RQ2: Implement an intuitive, guided onboarding process with tutorials and voice navigation. RQ3: Develop a minimalist, senior-friendly UI with large fonts, voice controls, and simple navigation.
Social influence	Immediate family plays a crucial role in adoption, highlighting the need for real-time shared insights into user well-being.	RQ4: Include a "Shared Dashboard" for real-time well-being monitoring by family members.
Facilitating conditions	Concerns about reliability affect perceived usefulness; trust-building measures can increase confidence in the system.	RQ5: Emphasize trust-building features (fail-safe mechanisms, backup options, transparent logs); provide a dual-interface option with "Essential Mode" and "Full Feature Mode", allowing users to customize their experience based on their comfort level.
Behavior intention	Older adults have varying comfort levels with technology; dual-interface options allow for gradual adaptation and personalized interaction.	RQ6: Provide dual-interface options to allow gradual adaptation and personalized interaction, supporting both mediated and autonomous use.

[a]UTAUT: Unified Theory of Acceptance and Use of Technology.

[b]CCS: Connected Care System.

XSL•FO
RenderX

A central design priority is to ensure that smart home technologies deliver clear and tangible benefits (RQ1). The findings show that both older adults and caregivers highly value peace of mind, safety, and independence, making these factors essential in encouraging adoption. To reinforce behavior intention, systems should communicate emotional and practical benefits early and consistently, for example, through onboarding materials, real-life success stories, and user testimonials that illustrate improvements in daily life and well-being. Integrating real-time feedback, alerts, and status updates can further build confidence and demonstrate reliability, reassuring users that the system is functioning as intended [11].

For many older adults, low digital confidence remains a significant barrier to adoption. A structured onboarding process that includes step-by-step tutorials, voice-assisted navigation, and simplified instructions can help users build comfort and familiarity (RQ2). Importantly, training should not target the older adult alone. Because family members and caregivers often play a central role in setup and troubleshooting, implementing a "Family Support Mode" would allow them to provide remote assistance when needed. To maintain inclusivity, designers should also offer multiple access pathways, such as an on-demand helpline, chatbot, or low-tech phone-based interfaces, ensuring that support is available for users who struggle with app-based interactions [52].

Ease of use also depends on minimizing cognitive load and supporting habitual use. A senior-friendly interface with large fonts, voice commands, and clear navigation can make the system more accessible (RQ3). Compatibility with smart assistants such as Alexa or Google Home would reduce interaction complexity and align with technologies many users already know. Personalized automation features—such as preset routines and scheduled alerts—can further reduce manual adjustments and promote effortless engagement. These design strategies together enhance usability and sustain engagement over time, reinforcing the link between ease of use and adoption [23].

Equally important is recognizing that adoption and use are shared experiences rather than solitary ones. A "Shared Dashboard" can enable caregivers to remotely monitor well-being while preserving older adults' autonomy (RQ4). Multiuser support and customizable permission settings allow older adults to determine what data are visible and to whom, maintaining control and privacy within supportive relationships. These features echo evidence that family involvement—rather than peer influence—is the strongest predictor of adoption among older adults [50,51]. To further enhance trust, fail-safe mechanisms, transparent data logs, and customizable alerts (eg, emergency notifications, wellness check-ins, or reminders) should be integrated. A non-intrusive "Check-in" feature can also allow caregivers to confirm safety in a way that is supportive rather than surveillant [49].

Beyond interpersonal support, the system itself must feel dependable and seamlessly embedded in daily life. Smart home interfaces should be transparent, intuitive, and designed for users with differing levels of digital experience (RQ5). Because many participants preferred minimal engagement, passive monitoring functions should operate quietly in the background, providing meaningful support without creating a sense of constant observation or disruption [45]. This principle extends beyond product design to service delivery, offering guidance for housing providers, health and social care organizations, and policymakers seeking to integrate smart home systems into aging in place strategies [53].

Finally, long-term adoption depends on flexibility and personalization. Older adults' needs and comfort levels change over time, meaning that a single interface or configuration is unlikely to suit all users indefinitely. A dual-interface option, allowing users to switch between a simplified Essential Mode and a more advanced Full Feature Mode, supports gradual learning and adaptation (RQ6). Customization features—such as adjustable settings, notification preferences, and tailored prompts—ensure that technology remains user-centric and empowering. Such personalization not only promotes sustained engagement but also helps mitigate feelings of surveillance or dependency, reinforcing autonomy and dignity in everyday use [30,50]. In this way, personalization and emotional reassurance become mutually reinforcing elements of inclusive smart-home design. This form of personalization enhances usability while helping to reduce perceptions of surveillance or disruption, especially in multiuser households [54].

Taken together, these design implications emphasize that effective smart home technologies for aging populations must integrate emotional reassurance, shared use, and adaptability into their core functionality—not as optional features but as central principles. Doing so will enable systems like the CCS to support aging in place while strengthening trust, confidence, and connectedness between older adults and their caregivers [55].

Limitations and Future Research

Building on the contributions of this study, future researchers should rethink traditional approaches to technology adoption in aging populations by incorporating training interventions and structured support systems that better address older adults' needs. This study highlights the importance of customization and personalization, suggesting that future research should move beyond generic usability considerations and instead investigate which specific customization features are most effective in promoting long-term engagement with smart home and health care technologies.

Additionally, this study challenges the conventional social influence construct in technology adoption models. Future researchers should redefine social influence by considering how shared-use systems operate when one user group (eg, caregivers) has more agency in adoption decisions than the other (older adults). Researchers should explore family dynamics in adoption processes, shifting focus from broad social influence on immediate social networks, particularly how family members shape adoption and continued use. Future studies should explore how caregiver involvement modifies traditional adoption pathways, particularly in health, caregiving, and assistive technologies [52]. A key limitation of this study is that data were collected solely from older adults, without corresponding input from their caregivers. As a result, the findings reflect older

XSL•FO
RenderX

adults' perceptions of caregiver involvement rather than caregivers' own experiences and perspectives.

Furthermore, this work calls for a broader conceptualization of PU in aging populations. Future adaptations of UTAUT should refine the PU construct to better capture indirect, relational, and emotional benefits, rather than relying solely on task efficiency-based measures [11,49]. Last, researchers should integrate care models with technology acceptance frameworks, recognizing that smart home technologies are often collaborative rather than individualistic, requiring a more nuanced understanding of adoption behaviors across multiple stakeholders.

The methodology used in this study was subject to several noteworthy limitations, which significantly impacted the results and constrained the interpretation of the data. The lack of control over participants' access to technology introduced variability in their technological usage throughout the experiment. This may have confounded the results, particularly affecting the PEOU construct. By not having an in-depth understanding of the user's existing experience and self-efficacy with computer use, the results and their interpretation have assumed minimal experience and knowledge regarding computer skills. Despite these limitations, the study's strength lies in its use of real participants who used a real system for a minimum of 6 months. This real-world usage provides valuable practical insights, reflected in the strong relationship between facilitating conditions and behavior intentions. However, the methodological constraints may have led to an incomplete picture of the adoption process, particularly in understanding the temporal aspects of adoption and the evolving perceptions of usefulness over time.

Conclusions

This study sought to examine the adoption and use of the CCS among older adults over a 6-month period, addressing a critical research gap in understanding how this population interacts with smart home technologies in real-world settings. By using both quantitative and qualitative methods, the study offers a more nuanced perspective on the factors influencing the adoption of these technologies. The research contributes to the existing literature by testing the applicability of preexisting adoption constructs, particularly those from the UTAUT model, while also revealing insights that may challenge or extend these frameworks when applied to older adults. Findings indicate that PU in passive smart home technologies must be redefined to include emotional and relational benefits, rather than focusing solely on task efficiency. The role of caregivers as facilitators in adoption decisions emerged as crucial, necessitating a more relational approach to social influence.

Furthermore, personalization and adaptability were identified as key factors in sustained engagement. These insights have practical implications for system design, emphasizing the need for user-friendly, customizable, and caregiver-supported technologies. Future research should further explore caregiver involvement, shared-use adoption models, and long-term engagement strategies to ensure that smart home technologies effectively support aging in place, balancing independence, usability, and support structures.

Acknowledgments

The authors would like to thank all participants for their patience and openness to share opinions on a novel assistive technology. This study was supported by funding from the Engineering and Physical Sciences Research Council (EPSRC) through an Industrial CASE Award, with additional support provided by a large telecommunications organization in the United Kingdom.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Quantitative and qualitative survey instruments, thematic analysis results with code frequencies and participant excerpts, mixed methods meta-inferences, and resulting design requirements.
[DOCX File , 44 KB - aging_v9i1e76012_app1.docx]

References

1. Wass S, Thygesen E, Purao S. Principles to facilitate social inclusion for design-oriented research. J Assoc Inf Syst 2023;24(5):1204-1247 [FREE Full text] [doi: 10.17705/1jais.00814]
2. Lei L, Leggett AN, Maust DT. A national profile of sandwich generation caregivers providing care to both older adults and children. J Am Geriatr Soc 2023;71(3):799-809 [FREE Full text] [doi: 10.1111/jgs.18138] [Medline: 36427297]
3. Xue B, Lacey R, Di Gessa G, McMunn A. Do mental and physical health trajectories change around transitions into sandwich care? results from the UK household longitudinal study. Public Health 2025;239:224-229 [FREE Full text] [doi: 10.1016/j.puhe.2024.12.001] [Medline: 39788842]
4. Bauer JM, Sousa-Poza A. Impacts of informal caregiving on caregiver employment, health, and family. J Popul Ageing 2015;8(3):113-145. [doi: 10.1007/s12062-015-9116-0]
5. Overview of the UK population: January 2021. Office for National Statistics. 2021. URL: https://tinyurl.com/5ymbbtua [accessed 2025-12-20]

XSL•FO
RenderX

6. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. JMIR Aging 2023;6:e43197 [FREE Full text] [doi: 10.2196/43197] [Medline: 36943333]
7. Bergschöld JM, Gunnes M, Eide A, Lassemo E. Characteristics and range of reviews about technologies for aging in place: scoping review of reviews. JMIR Aging 2024;7:e50286 [FREE Full text] [doi: 10.2196/50286] [Medline: 38252472]
8. Shi H, Qi H, Yang K. Enhancing elderly care with smart homes: a comparative study of use and payment willingness. Digit Health 2025;11:20552076251374223 [FREE Full text] [doi: 10.1177/20552076251374223] [Medline: 40933084]
9. Internet users, UK. The Office for National Statistics. 2023. URL: https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2020 [accessed 2025-12-20]
10. State of caring 2019. Carers UK. 2019. URL: https://www.carersuk.org/reports/state-of-caring-2019/ [accessed 2024-07-31]
11. Peek STM, Wouters EJM, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJM. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014;83(4):235-248 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
12. Gövercin M, Meyer S, Schellenbach M, Steinhagen-Thiessen E, Weiss B, Haesner M. SmartSenior@home: acceptance of an integrated ambient assisted living system. results of a clinical field trial in 35 households. Inform Health Soc Care 2016;41(4):430-447. [doi: 10.3109/17538157.2015.1064425] [Medline: 26809357]
13. Galambos C, Rantz M, Craver A, Bongiorno M, Pelts M, Holik AJ, et al. Living with intelligent sensors: older adult and family member perceptions. Comput Inform Nurs 2019;37(12):615-627. [doi: 10.1097/CIN.0000000000000555] [Medline: 31498250]
14. Sorwar G, Aggar C, Penman O, Seton C, Ward A. Factors that predict the acceptance and adoption of smart home technology by seniors in Australia: a structural equation model with longitudinal data. Inform Health Soc Care 2023;48(1):80-94. [doi: 10.1080/17538157.2022.2069028] [Medline: 35473463]
15. Venkatesh V, Brown S, Sullivan Y. Guidelines for conducting mixed-methods research: an extension and illustration. J Assoc Inf Syst 2016;17(7):435-494 [FREE Full text] [doi: 10.17705/1jais.00433]
16. Joa C, Magsamen-Conrad K. Social influence and UTAUT in predicting digital immigrants' technology use. Behav Inf Technol 2021;41(8):1620-1638 [FREE Full text] [doi: 10.1080/0144929x.2021.1892192]
17. Zhang M. Older people's attitudes towards emerging technologies: a systematic literature review. Public Underst Sci 2023;32(8):948-968 [FREE Full text] [doi: 10.1177/09636625231171677] [Medline: 37204075]
18. Wan X, Lighthall N, Roque N. Predictors of technology use: evidence from elite older adults. Act Adapt Aging 2024;48(4):648-669. [doi: 10.1080/01924788.2023.2282276] [Medline: 39744081]
19. Yang C, Liu C, Wang Y. The acceptance and use of smartphones among older adults: differences in UTAUT determinants before and after training. Libr Hi Tech 2022;41(5):1357-1375 [FREE Full text] [doi: 10.1108/lht-12-2021-0432]
20. Carrier L, Cheever N, Rosen L, Benitez S, Chang J. Multitasking across generations: multitasking choices and difficulty ratings in three generations of Americans. Comput Hum Behav 2009;25(2):483-489 [FREE Full text] [doi: 10.1016/j.chb.2008.10.012]
21. Olson KE, O'Brien MA, Rogers WA, Charness N. Diffusion of technology: frequency of use for younger and older adults. Ageing Int 2011;36(1):123-145 [FREE Full text] [doi: 10.1007/s12126-010-9077-9] [Medline: 22685360]
22. Zhang L. Breaking barriers: a dual-factor model unraveling ageism in socially assistive robot adoption among older Chinese adults. Int J Hum Comput Interact 2024;41(2):1397-1413 [FREE Full text] [doi: 10.1080/10447318.2024.2314352]
23. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the almere model. Int J Soc Robotics 2010;2(4):361-375 [FREE Full text] [doi: 10.1007/s12369-010-0068-5]
24. Miguel Cruz A, Lopez Portillo HP, Daum C, Rutledge E, King S, Liu L. Technology acceptance and usability of a mobile app to support the workflow of health care aides who provide services to older adults: pilot mixed methods study. JMIR Aging 2022;5(2):e37521 [FREE Full text] [doi: 10.2196/37521] [Medline: 35583930]
25. Wei W, Gong X, Li J, Tian K, Xing K. A study on community older people's willingness to use smart home-an extended technology acceptance model with intergenerational relationships. Front Public Health 2023;11:1139667 [FREE Full text] [doi: 10.3389/fpubh.2023.1139667] [Medline: 37346111]
26. Wang X, Lee C, Jiang J, Zhang G, Wei Z. Research on the factors affecting the adoption of smart aged-care products by the aged in China: extension based on UTAUT Model. Behav Sci (Basel) 2023;13(3):277 [FREE Full text] [doi: 10.3390/bs13030277] [Medline: 36975302]
27. Blut M, Chong A, Tsigna Z, Venkatesh V. Meta-analysis of the unified theory of acceptance and use of technology (UTAUT): challenging its validity and charting a research agenda in the red ocean. J Assoc Inf Syst 2022;23(1):13-95 [FREE Full text] [doi: 10.17705/1jais.00719]
28. Venkatesh V, Morris MG, Davis G, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27:425-478 [FREE Full text] [doi: 10.2307/30036540]
29. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13:319-339 [FREE Full text] [doi: 10.2307/249008]
30. Peek S, Luijkx K, Rijnaard M, Nieboer ME, van der Voort CS, Aarts S, et al. Older adults' reasons for using technology while aging in place. Gerontology 2016;62(2):226-237 [FREE Full text] [doi: 10.1159/000430949] [Medline: 26044243]

XSL•FO
RenderX

31. NHS Health Research Authority (HRA). ‘Does my project require review by a Research Ethics Committee?’ V 2.0. 2020. URL: https://www.hra-decisiontools.org.uk/ethics/docs/Algorithm%20-%20Does%20my%20project%20require%20REC%20review%20v2.0%2020200304.pdf
32. Braun V, Clarke V. One size fits all? what counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2020;18(3):328-352 [FREE Full text] [doi: 10.1080/14780887.2020.1769238]
33. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
34. Terry G, Hayfield N, Clarke V, Braun V. Thematic analysis. In: The SAGE Handbook of Qualitative Research in Psychology. Los Angeles (USA): SAGE Publications Ltd; 2017.
35. Hair J, Sarstedt M, Pieper T, Ringle C. The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. Long Range Planning 2012;45(5-6):320-340 [FREE Full text] [doi: 10.1016/j.lrp.2012.09.008]
36. Venkatesh V, Brown S, Bala H. Bridging the qualitative-quantitative divide: guidelines for conducting mixed methods research in information systems. MIS Q 2013;37:54. [doi: 10.25300/misq/2013/37.1.02]
37. Bergman M. Advances in Mixed Methods Research. Los Angles (USA): SAGE Publications Ltd; 2008.
38. Henseler J, Ringle C, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J Acad Mark Sci 2014;43(1):115-135 [FREE Full text] [doi: 10.1007/s11747-014-0403-8]
39. Taber K. The use of cronbach’s alpha when developing and reporting research instruments in science education. Res Sci Educ 2017;48(6):1273-1296 [FREE Full text] [doi: 10.1007/s11165-016-9602-2]
40. Hair J, Ringle C, Sarstedt M. PLS-SEM: indeed a silver bullet. J Mark Theory Pract 2014;19(2):139-152 [FREE Full text] [doi: 10.2753/mtp1069-6679190202]
41. Magsamen-Conrad K, Upadhyaya S, Joa C, Dowd J. Bridging the divide: using UTAUT to predict multigenerational tablet adoption practices. Comput Human Behav 2015;50:186-196 [FREE Full text] [doi: 10.1016/j.chb.2015.03.032] [Medline: 25937699]
42. Guner H, Acarturk C. The use and acceptance of ICT by senior citizens: a comparison of technology acceptance model (TAM) for elderly and young adults. Univ Access Inf Soc 2018;19(2):311-330 [FREE Full text] [doi: 10.1007/s10209-018-0642-4]
43. Hong W, Thong J, Chasalow L, Dhillon G. User acceptance of agile information systems: a model and empirical test. J Manag Inf Syst 2014;28(1):235-272 [FREE Full text] [doi: 10.2753/mis0742-1222280108]
44. Lee S, Ory M, Vollmer Dahlke D, Smith M. Technology use among older adults and their caregivers: cross-sectional survey study. JMIR Aging 2024;7:e50759 [FREE Full text] [doi: 10.2196/50759] [Medline: 38717339]
45. Kamal S, Shafiq M, Kakria P. Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). Technol Soc 2020;60:101212 [FREE Full text] [doi: 10.1016/j.techsoc.2019.101212]
46. Venkatesh V, Davis F. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000;46(2):186-204 [FREE Full text] [doi: 10.1287/mnsc.46.2.186.11926]
47. Zhang W, Liu L. Unearthing consumers’ intention to adopt eco-friendly smart home services: an extended version of the theory of planned behavior model. J Environ Plan Manag 2021;65(2):216-239 [FREE Full text] [doi: 10.1080/09640568.2021.1880379]
48. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research. Thousand Oaks: SAGE; 2018.
49. Choi N, Dinitto DM. Internet use among older adults: association with health needs, psychological capital, and social capital. J Med Internet Res 2013;15(5):e97 [FREE Full text] [doi: 10.2196/jmir.2333] [Medline: 23681083]
50. Choi YK, Thompson HJ, Demiris G. Use of an internet-of-things smart home system for healthy aging in older adults in residential settings: pilot feasibility study. JMIR Aging 2020;3(2):e21964 [FREE Full text] [doi: 10.2196/21964] [Medline: 33170128]
51. Luijkx K, Peek S, Wouters E. "Grandma, You Should Do It--It's Cool" older adults and the role of family members in their acceptance of technology. Int J Environ Res Public Health 2015;12(12):15470-15485 [FREE Full text] [doi: 10.3390/ijerph121214999] [Medline: 26690188]
52. Kim S, Choudhury A. Exploring older adults’ perception and use of smart speaker-based voice assistants: a longitudinal study. Comput Hum Behav 2021;124:106914 [FREE Full text] [doi: 10.1016/j.chb.2021.106914]
53. Czaja S, Charness N, Fisk A, Hertzog C, Nair SN, Rogers WA, et al. Factors predicting the use of technology: findings from the center for research and education on aging and technology enhancement. Psychol Aging 2006;21(2):333-352 [FREE Full text] [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
54. Demiris G, Hensel B, Skubic M, Rantz M. Senior residents' perceived need of and preferences for "smart home" sensor technologies. Int J Technol Assess Health Care 2008;24(1):120-124. [doi: 10.1017/S0266462307080154] [Medline: 18218177]
55. Wang X, Wong Y, Sun S, Yuen K. An investigation of self-service technology usage during the COVID-19 pandemic: the changing perceptions of 'self' and technologies. Technol Soc 2022;70:102032 [FREE Full text] [doi: 10.1016/j.techsoc.2022.102032] [Medline: 35757525]

XSL•FO
RenderX

Abbreviations

AVE: average variance extracted
CCS: Connected Care System
PEOU: perceived ease of use
PU: perceived usefulness
SEM: Structural Equation Modeling
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by H Köttl; submitted 15.Apr.2025; peer-reviewed by X Zhang, M Tomita; comments to author 05.Sep.2025; revised version received 02.Nov.2025; accepted 01.Dec.2025; published 15.Jul.2026.

Please cite as:
Pemberton L, Carter M, Mehandjiev N
Adoption and Use of a Smart Home Connected Care System by Older Adults: Mixed Methods Study
JMIR Aging 2026;9:e76012
URL: https://aging.jmir.org/2026/1/e76012
doi:10.2196/76012
PMID:

©Laura Pemberton, Michelle Carter, Nikolay Mehandjiev. Originally published in JMIR Aging (https://aging.jmir.org), 15.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Multidimensional Digital Health Platform to Support Intrinsic Capacity and Healthy Aging: Pilot Feasibility Study

Marcin Kolakowski[1], PhD; Andrea Lupica[2], MSc; Seif Ben Bader[3], PhD; Jaouhar Ayadi[2], PhD; Luca Gilardi[2], BSc; Angelo Consoli[2], Prof Dr; Irina Georgiana Mocanu[4], MSc; Oana Cramariuc[4], MSc; Lionello Ferrazzini[3], MSc; Eva Reithner[5], MSc; Magdalena Velciu[6], PhD; Barbara Borgogni[7], BSc; Sofia Rivaira[7], BSc; Luca Antognoli[8], PhD; Sara Leonzi[8], BSc; Elisa Felici[8], PsyD; Margherita Rampioni[8], PhD; Giacomo Cucchieri[8], BSc; Vera Stara[8], PhD

[1]Institute of Radioelectronics and Multimedia Technology, Warsaw University of Technology, Warsaw, Poland

[2]ECLEXYS Sagl, Riva San Vitale, Switzerland

[3]OCTILIUM Sagl, Chiasso, Switzerland

[4]Centrul IT Pentru Stiinta si Tehnologie (CITST), Bucharest, Romania

[5]EURAG Österreich, Wien, Austria

[6]Ana Aslan International Foundation, Bucharest, Romania

[7]Medea S.r.l., Florence, Italy

[8]IRCCS INRCA Istituto Nazionale di Riposo e Cura per Anziani, Ancona, Italy

Corresponding Author:
Luca Antognoli, PhD
IRCCS INRCA Istituto Nazionale di Riposo e Cura per Anziani
Via Santa Margherita, 5
Ancona, 60124
Italy
Phone: 39 3495970120
Email: l.antognoli@inrca.it

Abstract

Background: By 2050, 22% of the global population will be aged 60 years or older, with Europe experiencing rapid aging. This increase in chronic diseases and functional decline demands a shift from fragmented, hospital-centric care to continuous care models. The World Health Organization's healthy aging framework prioritizes intrinsic capacity (IC)—defined as physical and mental abilities across the locomotion, vitality, sensory, cognition, and psychology domains. Digital health solutions enable remote monitoring but face barriers, such as low digital literacy and usability challenges among older adults.

Objective: This study evaluates the impact of a multidimensional, artificial intelligence–driven digital health platform (CAREUP) on the dynamic monitoring and personalized enhancement of IC in community-dwelling older adults through the integration of real-time physiological, cognitive, and behavioral data streams.

Methods: This pilot feasibility study was conducted from September to December 2024 in Italy, Romania, and Austria. It involved 66 older adults as primary users and 17 caregivers as secondary users. The CAREUP system included Android apps (Careplan and Positive Health), smart devices, and a dashboard. Users followed a 2-month plan. We assessed usability (System Usability Scale [SUS]), user experience (User Experience Questionnaire short version [UEQ-S]), and acceptance (qualitative) at 3 time points (T0, T1, and T2), along with health and quality-of-life measures.

Results: Of the 66 enrolled participants, 61 primary users (37 women and 24 men; mean age 73.9 years) completed the study. The overall SUS score was 68, indicating acceptable but improvable usability, with significant variation across countries. The UEQ-S score of 1.45 indicated a good overall user experience, with high Hedonic Quality (1.66) suggesting strong emotional appeal and Pragmatic Quality (1.23) reflecting above-average usability. Most participants (49/61, 80%) rated the platform as interesting, helpful, and efficient, and 48 out of 61 (79%) were willing to recommend it to friends. Secondary outcomes demonstrated stable physical and mental health and maintained independence in daily activities. Technical connectivity issues and device complexity emerged as the primary barriers to adoption.

Conclusions: CAREUP demonstrated feasibility for IC monitoring, with acceptable usability and a positive user experience. However, technical challenges and cross-country variations indicate a need for adaptations to address cultural differences and

XSL•FO
RenderX

varying levels of digital literacy. Prioritizing reliability, simplifying the system, and providing user support will be essential to enhance broader adoption.

(JMIR Aging 2026;9:e79862) doi:10.2196/79862

KEYWORDS

digital health; healthy aging; intrinsic capacity; older adults; usability; user experience; remote monitoring; integrated care platform; artificial intelligence; multidimensional health assessment

Introduction

Background

By 2050, nearly 22% of the global population will be aged 60 years or older, with Europe experiencing a particularly rapid demographic shift in which those aged over 65 years will outnumber children aged under 15 years [1,2]. This aging trend is accompanied by a rise in chronic diseases, as aging is a primary risk factor for many health conditions [3,4]. Age-related chronic diseases often lead to functional decline and social isolation, making it difficult for older adults to maintain independence and quality of life.

To shift away from a disease-centered model of aging, the World Health Organization introduced the concept of healthy aging, which focuses on preserving functional abilities through the optimization of intrinsic capacity (IC) alongside environmental support. IC encompasses the core physical and mental abilities individuals have at different stages of life, which are influenced by environmental factors and naturally decline with age. It includes 5 key domains: locomotion, vitality, sensory function, cognition, and psychological health [5-7].

Preventing chronic diseases and improving the lives of older adults require care models that focus on maintaining IC. The World Health Organization's Integrated Care for Older People framework recommends an integrative, long-term, multidisciplinary approach centered on person-focused care [8]. However, these complex care needs place significant strain on health care and social systems that are already facing resource shortages, compounded by a limited supply of formal and informal caregivers [9-11].

Digital technologies present promising solutions to support this challenging landscape. Remote monitoring and digital health platforms can streamline and integrate the management of chronic conditions, enhancing older adults' independence while reducing health care costs. Collectively referred to as digital health solutions, these technologies require users to have varying levels of health and digital literacy, which must be considered to ensure accessibility and adoption, especially among older adults with frailty and chronic disability [12-16]. Given the wide variability in digital skills among seniors and their preference for personal contact with health care providers, digital tools must be tailored to be simple, intuitive, and user-friendly—for example, by employing large fonts, minimal button use, and straightforward interactions [17-19].

Knowledge Gap

Many digital platforms targeting healthy aging have been developed in both commercial and research settings, as summarized in Table 1. Most health monitoring platforms collect data using sensors, such as blood pressure monitors, thermometers, oximeters, smartwatches, and activity trackers, as well as digital questionnaires assessing mental well-being and social engagement. These data are typically uploaded via a gateway device, such as a tablet or smartphone, and processed in the cloud. The outcomes are usually accessible through web- or mobile-based interfaces.

XSL•FO
RenderX

Table 1. Exemplary health monitoring platforms.

Platform name	Target	Data sources	Interface	Additional features	Trials	Usability
Comarch HomeHealth 2.0 [20]	Patients and medical professionals	Sensors and questionnaires	Web and mobile	Alerts of parameters outside of predefined scope and virtual visits	Not reported	Not reported
RO-SmartAgeing [21,22]	Patients and medical professionals	Sensors	Web and mobile	Alarms for unusual parameter values	Laboratory tested	High (laboratory tested)
InnoWell [23]	Patients and medical professionals	Sensors, medical records, and questionnaires	Web	No	Not yet reported; protocol available	Low
My AHA [24]	Patients and medical professionals	Sensors and other platforms (acts as middleware)	Web and mobile	Personalized care plan and fall and dementia risk	20 adults (alpha wave)	Not reported directly; platform revised after alpha wave
GER-e-TEC [25,26]	Older adults	Sensors and questionnaires	Web	Alerts of warning signs leading to hospitalization	36 adults monitored for 2 months; specificity could be higher	Not reported
FRAIL [27]	Patients with prefrailty and frailty	Sensors (including custom vest)	Web	Gamification system	40 adults recruited to test the devices	Medium usability of the custom sensors
CAREPATH [28,29]	Older adults	Sensors and questionnaires	Web and mobile	Personalized care plan and physical exercises plan	11 project team members	Good to excellent according to the project team members
HOPES [30,31]	Patients and medical professionals	Sensors and smartphone	Web and mobile	Anomaly detection	100 patients with schizophrenia	Not reported
Vivify +Home [32]	Patients and medical professionals	Sensors and smartphone	Web and mobile	Health tips and virtual visits	Not reported	Not reported
NESTORE [33]	Older adults	Sensors, questionnaires, and location data	Web and mobile	Personalized care plan, social platform, and chatbot	10 older adults	Reported issues with software usability
CAREUP [34]	Older adults	Sensors, questionnaires, and games	Web and mobile	Personalized care plan, intrinsic capacity monitoring, intrinsic capacity prediction, and games	66 older adults and 17 caregivers in 2 months	See the "Usability" section

The platforms differ primarily in how they process data and the services they offer. Medical system–connected platforms often alert caregivers or clinicians to abnormal physiological readings and may provide online consultation features. Home-use platforms tend to focus on personalized care plans that adapt based on the user's current status. More advanced solutions incorporate predictive analytics to estimate the risk of future health problems or declines in capacity.

As maintaining user engagement is crucial for platform success, features such as social networking and gamification are commonly added to enhance adherence. However, usability remains a pervasive and critical challenge. Many platforms incorporate wearables, mobile apps, and surveys to cover all IC domains, but often neglect the linguistic, educational, and social diversity of older adults. This oversight results in low user acceptance and limited engagement [23,35-40]. Reported challenges include difficulties with navigation, lack of personalized experiences, software bugs, and hardware integration issues [27,33]. These barriers are frequently observed in newly developed platforms within research projects.

Caring for an aging population within the healthy aging framework requires supporting residual IC and compensating for its decline. Technology designed specifically for older adults may facilitate integrated and continuous care, thereby addressing physical and psychosocial needs more effectively.

Study Objectives

The CAREUP project introduced an innovative digital platform for the remote monitoring of older individuals' IC to promote inclusive health. It integrates real-time physiological, cognitive, and behavioral data, aiming to slow IC decline and support caregiver-assisted care. This paper reports on a pilot feasibility study detailing the study design and user interactions with the CAREUP platform across Italy, Romania, and Austria. The results focus on usability, acceptance, and the platform's impact on quality of life, cognition, and physical activity. In line with current reporting standards for digital health implementations, this work was structured according to the iCHECK-DH (Guidelines and Checklist for the Reporting on Digital Health Implementations) to enhance completeness, transparency, and

XSL•FO
RenderX

reproducibility of the implementation description. The completed iCHECK-DH checklist is provided as Multimedia Appendix 1.

Methods

Study Design

The pilot feasibility study was one of the main objectives of the CAREUP project, aimed at evaluating the platform in a real-life environment and collecting feedback from actual users. A mixed methods design was used to collect data from primary (older adults) and secondary users (formal or informal caregivers). End users (primary and secondary) were enrolled by project partners in 3 European countries (Austria, Italy, and Romania) from September 2024 to December 2024.

Ethical Considerations

The research protocol was approved by the Ethics Committee of IRCCS INRCA (Istituto Nazionale di Ricovero e Cura per Anziani) on March 30, 2023 (approval number D/GEN 91).

CAREUP Platform Technical Description

The CAREUP platform integrates Android apps installed on a tablet (Careplan and Positive Health [PH]), smart devices, and an online dashboard (CAREUP's Portal Web Interface). The CAREUP platform concept is presented in Figure 1, and a brief description of the platform components is provided in Table 2. The technical aspects of the platform and the implemented algorithms are described in more detail elsewhere [34].

Figure 1. The CAREUP platform concept. AI: artificial intelligence; IC: intrinsic capacity; ICT: information and communications technology.

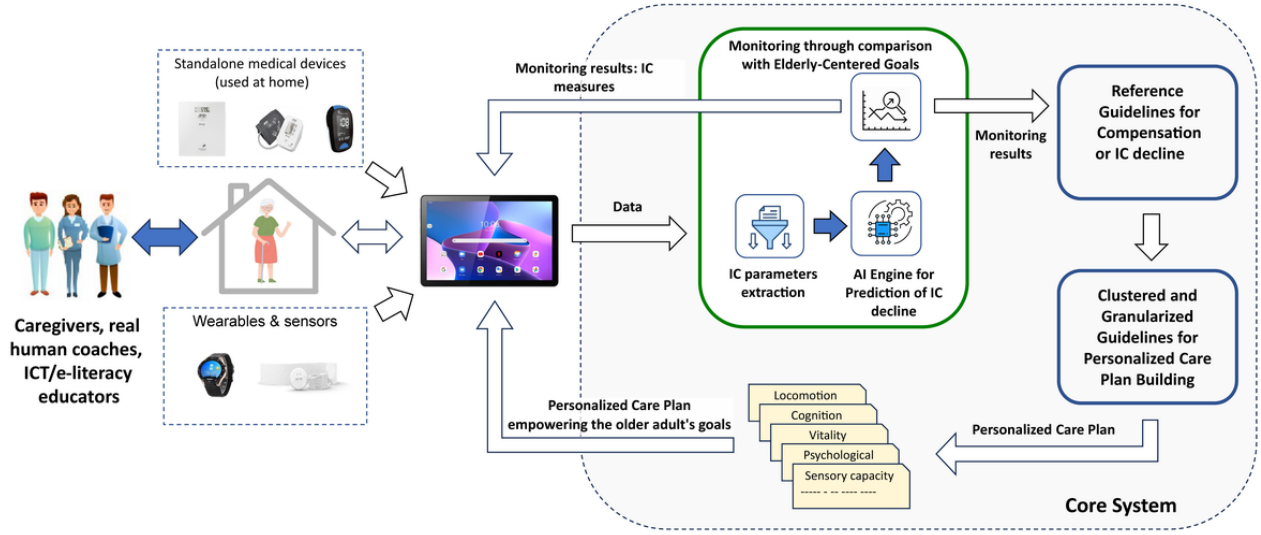


Table 2. CAREUP platform: components and functions.

Component type and the respective component	Function
Android tablet apps	
Careplan app	• Health parameters scan and report • Access to cognitive games apps
Positive Health app	• Health questionnaires and locomotion test
Devices	
Smartwatch (Starmax WS50)	• Step number acquisition
Blood pressure monitor (iHealth track)	• Blood pressure measurement
Weight scale (iHealth weight scale)	• Weight measurements
Oximeter (iHealth PO3M)	• Saturation measurement
Glucometer (iHealth glucometer)	• Blood glucose measurement
Handgrip strength meter (GRIPX Digital Hand Dynamometer Grip Strength Measurement Meter)	• Grip strength measurement
Online dashboard	
CAREUP's Portal Web Interface	• Monitoring of health parameters and intrinsic capacity values

XSL•FO
RenderX

The platform operation follows a common pattern in which primary users, using the platform devices, collect multimodal health data, such as the number of steps taken per day, blood pressure, oxygen saturation, body weight, and sleep quality. The core platform app is the Careplan app, which acts as the entry point to all platform apps. Its main purpose is to collect data from the medical devices. For supported Bluetooth-enabled devices, data are transferred and uploaded automatically using standard, secure protocols; for others, results can be entered manually. The Careplan app presents users with a personalized care plan, including tips for maintaining their IC at the required levels. Users can also access the PH app and the game section through provided links, featuring cognitive games selectable from both commercial and CAREUP-specific options.

The PH app is designed to collect health data that cannot be captured with smart devices. It includes digital health questionnaires to gather information on aspects such as well-being, emotions, and social activities, as well as 2 locomotion tests to assess physical performance. The PH app questionnaires are brief and consist of 3-4 closed-ended questions selected from a pool of 89 items derived from standardized and ad hoc reference questionnaires (Table 3). To reduce respondent burden, the PH app questionnaires do not address family situation, cognition, religion, or instrumental activities of daily living. A questionnaire for a given day is generated according to a predefined 4-week cycle, during which each question is asked at least once. Questions from specific reference questionnaires are distributed throughout the cycle to avoid monotony and maintain user engagement. If a user forgets to complete the PH app on a given day, the prepared questions are presented the following day. The questionnaire generation cycle is described in detail in Multimedia Appendix 2.

Table 3. Reference list of health questionnaires.

Standardized questionnaires	Assessment area
EQ-5D-5L [41]	Quality of life
15-item Geriatric Depression Scale [42]	Depression level
Health Status 12-Item Short Form version 2 [43]	Quality of life
Instrumental Activities of Daily Living scale [44]	Overall well-being
Loneliness Scale version 3 [45]	Perceived social isolation
Lubben Social Network Scale—Revised [46]	Social isolation
Mini Nutritional Assessment [47]	Overall nutritional status
Montreal Cognitive Assessment [48]	Mild cognitive impairment
Patient Health Questionnaire [49]	Depression level
Sensory domain (ad hoc questionnaire)	Sensory system
Short Physical Performance Battery [50]	Functional mobility in older adults
Sleep Quality Scale [51]	Sleep quality
Spiritual Well-Being scale [52]	Perceived spiritual quality of life
User-Centered Goals (ad hoc questionnaire)	User's goals/needs
Locomotion test	
30-second chair stand test [53]	Leg strength and endurance
Balance test [54]	Balance

The app's locomotion tests implemented selected components of the Short Physical Performance Battery, specifically the chair stand and balance tests. The PH app also provides users with appropriate instructions to enable them to perform the tests at home.

Results from device measurements and from the PH app's questionnaires and tests are uploaded to the platform via a Representational State Transfer (REST) application programming interface using the secure HTTPS protocol. In addition, any application programming interface operation requires a valid token to be provided with the request. Tokens are generated only for logged-in users and have an expiration date. After a prolonged break of a few days in platform use, the user must log in again.

The results are stored in a secure database in the CAREUP cloud. For additional data security, they are pseudonymized, and they cannot be directly linked to users' personal information without access to additional data. The collected results are used for weekly assessments of primary users' health status and IC. They also serve as the basis for preparing personalized care plans and IC prognoses.

Summaries and graphs of the analysis results are automatically generated and presented in the CAREUP Portal Web Interface, the online dashboard of the CAREUP platform. This browser-based interface can be accessed by both primary and secondary users using personal credentials to review health status and monitor changes in IC among older adults. In addition, the Web Portal provides access to results from medical

XSL•FO
RenderX

devices and outcomes of game sessions completed by the primary user.

Outcomes Measures

Data collection was conducted at 3 time points: the baseline evaluation (T0), the intermediate evaluation (T1), and the final evaluation (T2).

The primary outcomes, assessed at T1 and T2, aimed to evaluate the system's usability, user experience, and acceptance. Perceived usability of the platform (mobile apps, devices, and online dashboard) was assessed using the System Usability Scale (SUS) [55]; user experience was assessed using the User Experience Questionnaire short version (UEQ-S) [56]; and acceptance was evaluated using an ad hoc qualitative questionnaire consisting of open-ended questions.

The SUS score indicates excellent usability if it is over 80.3 (grade A), good usability from 68 to 80.3 (grade B), and poor or unacceptable usability if it is less than 51 (grade F). The UEQ-S focuses on pragmatic quality (usability-related aspects such as efficiency, clarity, and task support) and hedonic quality (emotional and aesthetic appeal, excitement, and innovation). Its score ranges from –3 to +3 and represents a very bad (–3.00 to –0.81), neutral (–0.80 to 0.80), good (0.81-1.40), excellent (1.41-2.00), or outstanding (2.01-3.00) experience. Acceptance refers to the degree to which users are willing to use a form of technology and was assessed using ad hoc open-ended questions.

The secondary outcomes included various health variables evaluated at T0, T1, and T2. Physical health was assessed using the 12-Item Short Form version 2 (SF-12 v2) survey [43] and the Instrumental Activities of Daily Living (IADL) scale [44]. Perceived loneliness was assessed using the UCLA Loneliness Scale (UCLA version 3 [UCLA v3]) [45], and social interaction was assessed using the Lubben Social Network Scale—Revised (LSNS-R) [46]. Quality of life was assessed using the EQ-5D-5L [41]. Cognitive and mood health were assessed using the Montreal Cognitive Assessment (MOCA) [48] and the 15-item Geriatric Depression Scale (GDS-15) [42], respectively. Spiritual well-being was assessed using the Spiritual Well-Being Scale score [52], and cognitive game experience was assessed using the Game Experience Questionnaire [57]. In addition, care workload among secondary users was assessed as a secondary outcome using ad hoc questions.

The SF-12 [43] provides 2 summary scores: 1 for the physical component summary and 1 for the mental component summary. The UCLA v3 [45] score can range from 20 to 80, with higher scores indicating greater feelings of loneliness and isolation (scores: 20-30, low loneliness; 31-50, moderate loneliness; and 51-80, high loneliness). The LSNS-R [46] focuses on the quality and quantity of social relationships; higher scores indicate a wider and more supportive social network (scores: 12-30, low social support; 31-44, satisfactory network with room for improvement; and 45-60, strong and supportive social network). The GDS-15 [42] score ranges from 0 to 15. Values up to 5 indicate a normal condition, while scores from 5 to 10 indicate moderate depression and scores from 11 to 15 indicate severe depression. The Spiritual Well-Being Scale [52] total score ranges from 20 to 120, with higher scores indicating greater spiritual well-being. The EQ-5D-5L [41] measures health-related quality of life using 5 items, and the Visual Analogue Scale, with values ranging from 0 to 100, measures the person's perceived state of health. The MOCA [48] has a maximum score of 30 points and generally indicates the presence of cognitive impairment when the score is 26 or lower. The Game Experience Questionnaire [57] measures the player's emotional and cognitive status after playing games, and its score follows a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely).

Participants

Overview

During the study, data were collected from both primary and secondary users according to predefined inclusion and exclusion criteria. The primary users were older adults, whereas the secondary users were their formal or informal caregivers.

Inclusion Criteria

Primary users were deemed eligible for the pilot if they (1) were aged 65 years or older; (2) lived independently in their homes in rural or urban areas; (3) had the capacity to provide consent and were available to sign the informed consent form; and (4) were willing to participate voluntarily in the pilot.

Exclusion Criteria

The exclusion criteria were as follows: (1) failure to meet the inclusion criteria; (2) use of active implantable or nonimplantable medical devices; (3) allergy to nickel components; (4) concomitant participation in other studies; (5) lack of written informed consent; (6) acute or untreated medical problems, such as a history of syncopal episodes, epilepsy, or pharmacologically uncontrolled vertigo; serious dysfunction of the autonomic system; severe behavioral syndromes not compensated by medication; concurrent neurological diseases; or severe systemic diseases with a life expectancy of less than 1 year; (7) myocardial infarction or stroke within the previous 6 months; (8) uncontrolled hypertension; (9) pacemaker or implantable cardioverter-defibrillator; (10) advanced Parkinson disease or other neuromuscular disorders; (11) metastatic cancer or immunosuppressive therapy; and (12) significant visual or hearing impairment. No inclusion or exclusion criteria were applied for secondary users.

Recruitment Strategy

Recruitment was multicenter and multinational, involving sites in Italy (n=30), Romania (n=23), and Austria (n=8) from September to December 2024. Participants were recruited through collaborating clinics, local outreach, and direct invitations and were representative of both rural and urban populations. Each site provided local training and support.

Sample Size Rationale

As a pilot study, the sample size (n=66) was determined to enroll a sufficient number of participants to capture usability barriers, acceptance issues, and preliminary outcomes across diverse cultural contexts. No formal power calculation was performed, as the primary aims were formative rather than confirmatory [58]. Although the sample size was limited, the study was explicitly designed as an exploratory feasibility investigation

XSL•FO
RenderX

to evaluate the usability, acceptance, and preliminary impact of the CAREUP platform.

Procedure

At T0, the primary and secondary users were asked to sign the informed consent form. Sociodemographic data were then collected from all participants. The initial evaluation of the primary users also included the MOCA [48] assessment to evaluate cognitive status and the User-Centered Goals questionnaire to define personal health goals and ambitions. Each primary user was provided, on a free-loan basis for the entire duration of the study, with all necessary equipment, including a tablet equipped with a SIM card for internet access and a selection of smart medical devices. Through the tablet, users could access the CAREUP platform apps and the online dashboard. Users were asked to select the smart devices they wished to test, and only the selected devices were provided. Participants received detailed explanations on how to use the system and its role in supporting and monitoring their health and IC. An introductory training phase was also conducted at the start of the pilot at each study site. During this phase, users received a manual in their native language along with recommendations for system use. Participants were instructed to access the tablet apps at least once a week; use the devices at least 2-3 times per week; complete the questionnaires and locomotion test once a week; and play the games as desired. All collected data (measurement values and IC graphs) could be checked daily or weekly by both primary and secondary users through the web portal (online dashboard).

Primary users were invited to use the platform according to these recommendations from T0 to T2, covering the entire pilot period of approximately 2 months. Details of the assessments conducted at T0, T1, and T2 are summarized in Table 4.

Table 4. List of tests and measurements for the pilot.

Dimension	Scale	Type of out-come	Type of end user	Timing of data collection
Socio Demo Data	N/A[a]	N/A	Primary and secondary	T0
User-Centered Goals	N/A	N/A	Primary	T0
Usability and User Experience	SUS[b] and UEQ-S[c]	Primary	Primary and secondary	T1 and T2 (for both)
Acceptance	Qualitative items	Primary	Primary and secondary	T1 and T2 (for both)
Cognitive Status	MOCA[d]	Secondary	Primary user	T0 and T2
Physical Health	SF-12[e] version 2 and IADL[f]	Secondary	Primary user	T0, T1, and T2
Perceived Loneliness	UCLA version 3	Secondary	Primary user	T0, T1, and T2
Social Interaction	LSNS-R[g]	Secondary	Primary user	T0, T1, and T2
Quality of Life	EQ-5D-5L	Secondary	Primary user	T0, T1, and T2
Mental Health	GDS-15[h]	Secondary	Primary user	T0, T1, and T2
Game Experience Questionnaire, part 3 postgame module	Game Experience Questionnaire	Secondary	Primary user	T1 and T2
Privacy and Stigma	Qualitative items	Secondary	Primary and secondary	T1 and T2 (for both)
Care Workload	Qualitative items	Secondary	Secondary user	T1 and T2
Willingness to Pay	Qualitative online questionnaire	Secondary	Primary and secondary	T2

[a]N/A: not applicable.

[b]SUS: System Usability Scale.

[c]UEQ-S: User Experience Questionnaire short version.

[d]MOCA: Montreal Cognitive Assessment.

[e]SF-12: 12-Item Short Form.

[f]IADL: Instrumental Activities of Daily Living.

[g]LSNS-R: Lubben Social Network Scale—Revised.

[h]GDS-15: 15-item Geriatric Depression Scale.

Statistical Analysis

The following section describes the outcomes of the pilot study. The pilot study data were uploaded to the Zenodo repository and are publicly available [59]. The dataset includes the results of the T0, T1, and T2 assessments; physiological measurements performed using the medical devices; and responses to the PH app questionnaires. The data are accompanied by anonymized demographic information for the users.

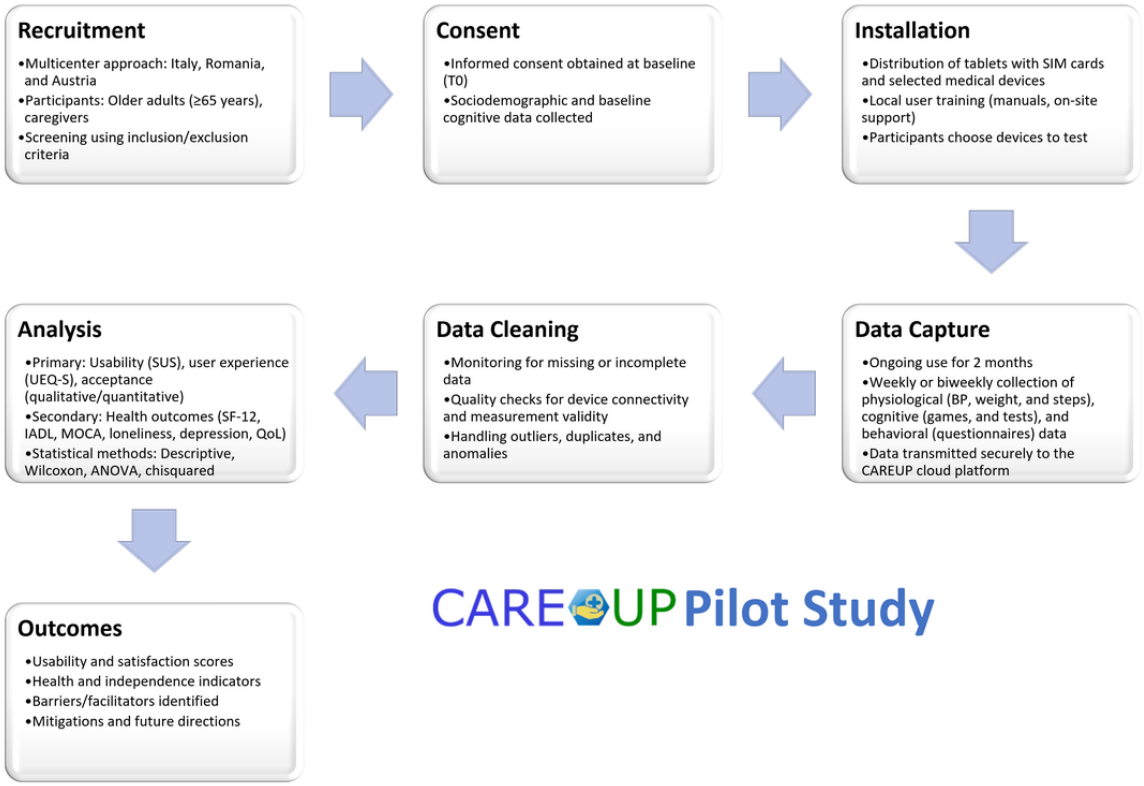
The dataset was curated at the end of the pilot study, and all data were cleaned and encoded using Python (Python Foundation) and the pandas package. The data are stored in standard CSV format. Instructions for dataset use, including a description of the file structure and variable dictionary, are

XSL•FO
RenderX

available in the dataset description PDF file. The repository also includes exemplary Python code for dataset loading and visualization.

In the following analysis, data are reported using mean values (a measure of central tendency representing the typical value of the dataset), frequencies (indicating how often a particular value or category appears in the dataset), and *P* values (used to determine whether observed differences or associations in the data are statistically significant). Statistical analyses included the Wilcoxon test and ANOVA for continuous variables, whereas the chi-square test was used for categorical variables. These analyses were conducted using Python (version 3.11), leveraging libraries such as SciPy for statistical testing, pandas for data handling, and Seaborn for visualization. A descriptive flowchart of all study stages is presented in Figure 2.

Figure 2. Study flowchart. IADL: Instrumental Activities of Daily Living scale; MOCA: Montreal Cognitive Assessment; QoL: quality of life; SF-12: 12-Item Short Form; SUS: System Usability Scale; UEQ-S: User Experience Questionnaire short version;.



Results

User Statistics

A total of 66 older adults participated as primary users: 30 were enrolled in Italy, 25 in Romania, and 11 in Austria. Their caregivers were involved as secondary users (7 in Romania, 5 in Austria, and 5 in Italy).

In Romania, 2 primary users (who enrolled without caregivers) withdrew after the first week of the pilot because they considered the platform too difficult to use and requiring IT knowledge that they did not have. There were also 3 primary user dropouts in Austria: 2 withdrew during the first week after recruitment, and 1 withdrew after the T1 phase. These users did not perceive any benefit from using the system. The pilot was therefore completed by 61 primary users and 17 secondary users. Not every primary user (older adult) was assigned a secondary user (caregiver), as the inclusion criteria for primary users required them to be in good health and not in need of caregiver support for daily activities.

The primary users included 37 women and 24 men, with a mean age of 73.9 years. Most participants were married (n=36), whereas 14 were widowed, 4 divorced, 3 separated, 2 single, and 2 in a full-time relationship. Regarding education level, 27 older adults had a secondary education, 28 had a tertiary education (university or further education), 5 had a primary education, and 1 had no formal education. Most participants (n=54) lived in urban areas, whereas 7 lived in rural areas.

The secondary users included 10 men and 7 women, with a mean age of 47.5 years. All were informal caregivers except for 1 formal caregiver from Italy. Full sociodemographic details are presented in Table 5.

Adherence to platform use was defined as the proportion of days between the first and last app use relative to the individual study period, whereas app usage referred to the total number of distinct days on which the app was used (Figure 3).

XSL•FO
RenderX

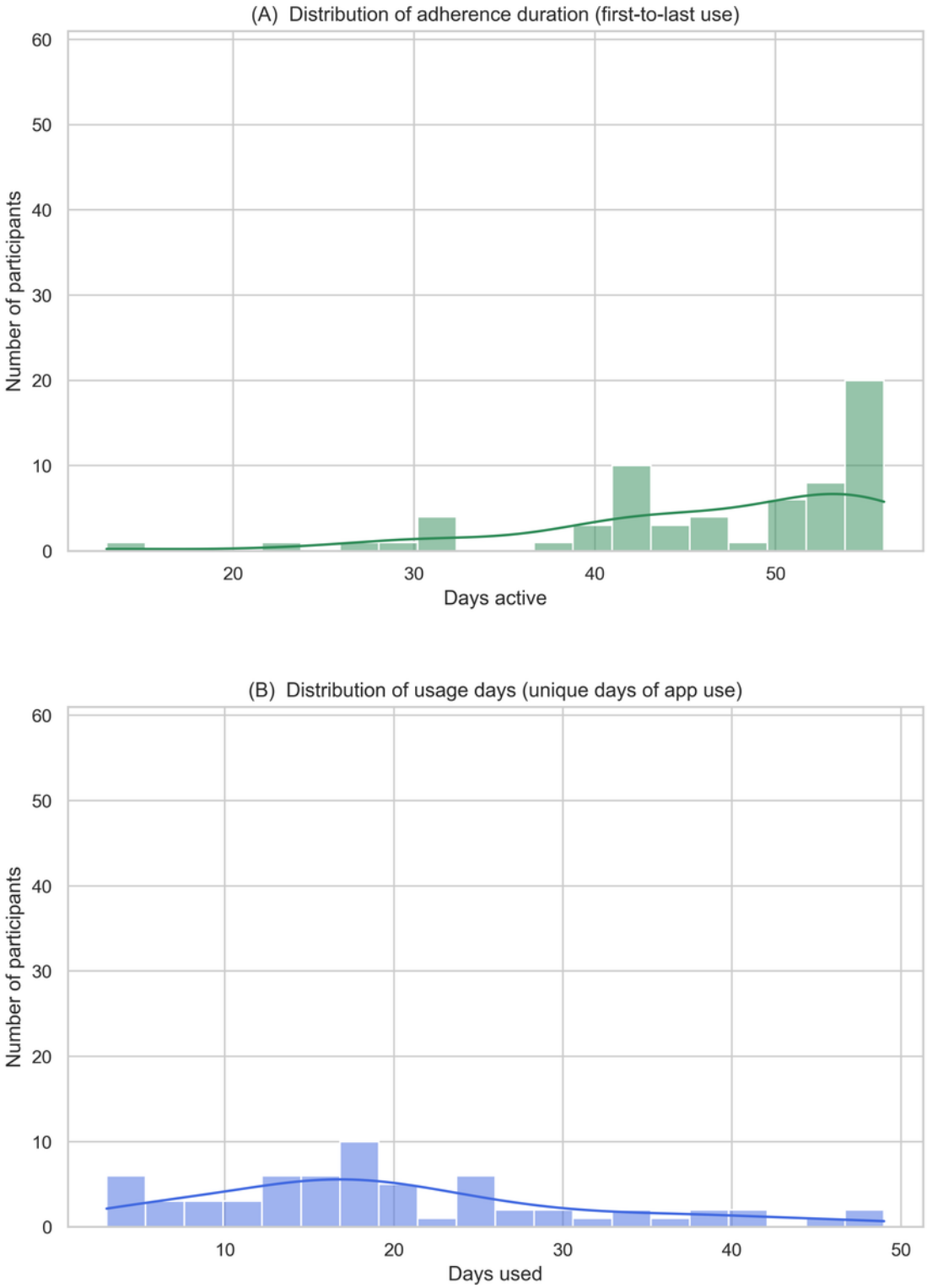
Table 5. Users' sociodemographic characteristics.

Sociodemographic characteristics	Older adults (n=61)	Caregivers (n=17)
Age, mean (SD)	73.9 (6)	47.5 (9.8)
Gender, n (%)		
Male	24 (39)	10 (59)
Female	37 (61)	7 (41)
Marital status, n (%)		
Married	36 (59)	5 (29)
Full-time relationship	2 (3)	4 (24)
Separated	3 (5)	N/A[a]
Divorced	4 (7)	N/A
Single	2 (3)	8 (47)
Widowed	14 (23)	N/A
Education, n (%)		
No education	1 (2)	N/A
Primary	5 (8)	2 (12)
Secondary	27 (44)	6 (35)
Tertiary	28 (46)	9 (53)
Residence, n (%)		
Urban area	54 (89)	16 (94)
Rural area	7 (11)	1 (6)
App use, median (IQR)		
Adherence (%)	85.6 (74.1-96.8)	N/A
Usage (days)	18.0 (13.0-24.2)	N/A

[a]N/A: not applicable.

XSL•FO
RenderX

Figure 3. Platform use. (A) Adherence and (B) usage.



Primary Outcomes

Usability

The average SUS score among all older end users was 68, corresponding to the benchmark 50th percentile. More specifically, the mean SUS score of the primary users was 67.62. According to adjective ratings typically associated with SUS scores, this result would be classified as "OK," suggesting that although the system was generally usable, there remains room for improvement. The lowest SUS scores were observed in Austria (50.83 at T1 and 45.94 at T2), whereas the highest scores were observed in Romania (74.46 at T1 and 75.54 at T2). Among secondary users (caregivers), the average score was 66.3, which was also close to the "OK" range but still indicated room for improvement. Further details are provided in Table 6.

As shown in Figure 4, the SUS scores had median values of 67.50 (IQR 22.50) at T1 and 70.00 (IQR 22.50) at T2. The

XSL•FO
RenderX

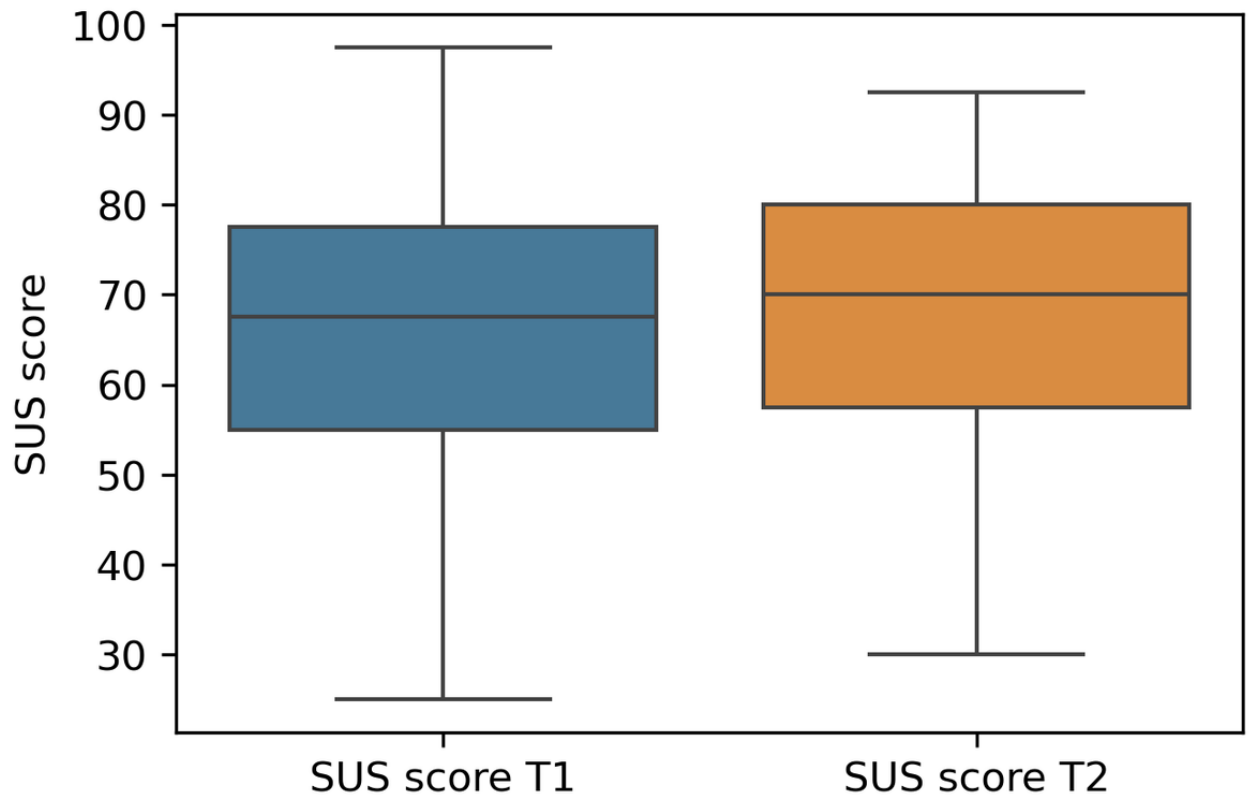
Wilcoxon signed rank test indicated no significant difference between the 2 time points (P=.42).

Table 6. System Usability Scale and User Experience Questionnaire short version scores among older adults and caregivers.

Scales and their scores	Older adults (n=61), mean (SD)	Caregivers (n=17), mean (SD)
System Usability Scale score	67.62 (15.09)	66.2 (21.25)
Nation		
Austria (n=8)	45.94 (9.99)	N/A[a]
Italy (n=30)	67.33 (13.06)	N/A
Romania (n=23)	75.54 (11.28)	N/A
Gender		
Female	70.34 (14.51)	N/A
Male	63.44 (15.30)	N/A
Education level		
Primary or no education	57.08 (14.53)	N/A
Secondary education	66.96 (15.65)	N/A
Tertiary education	70.65 (13.95)	N/A
User Experience Questionnaire short version score		
Hedonic Quality	1.66 (0.90)	1.22 (1.39)
Pragmatic Quality	1.23 (1.02)	0.76 (1.31)
Overall	1.45 (0.83)	0.99 (1.27)

[a]N/A: not applicable.

Figure 4. Boxplot of System Usability Scale (SUS) score at T1 and T2.



XSL•FO
RenderX

User Experience

As reported in Table 6, the overall UEQ score among primary users was 1.45, suggesting a "good" overall user experience in which both functional and emotional needs were reasonably well met.

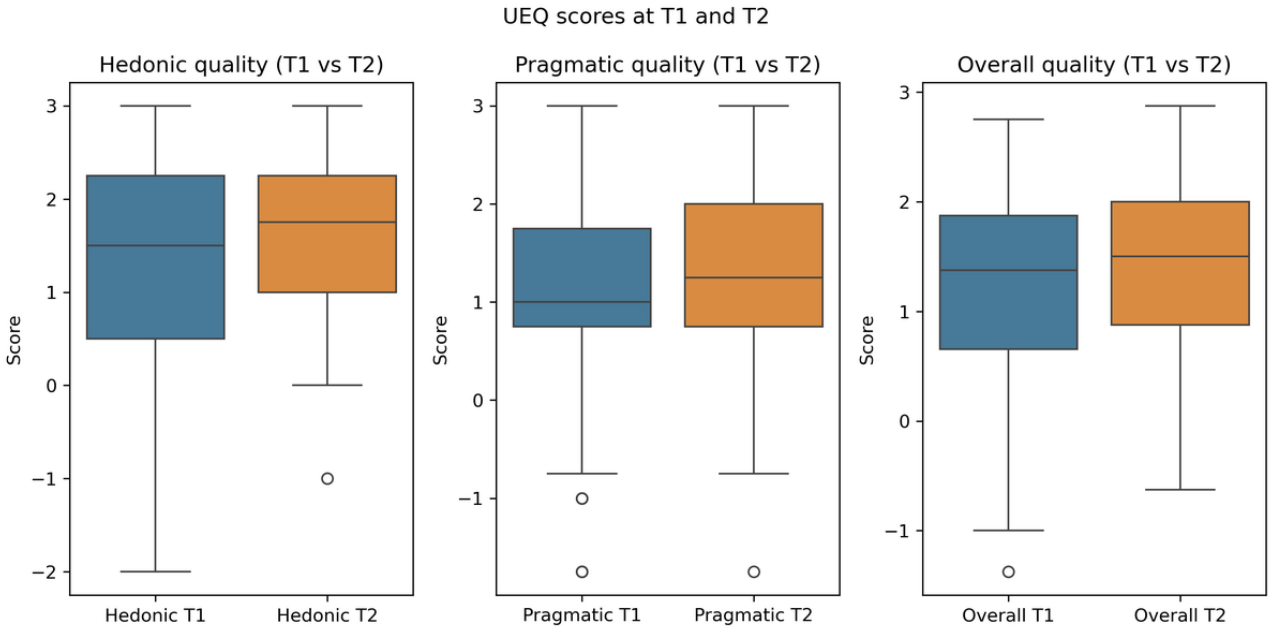
A high Hedonic Quality score of 1.66 indicated that the product was perceived as enjoyable, stimulating, and innovative. This emotional appeal enhanced user satisfaction and could promote greater engagement with the product.

A Pragmatic Quality score of 1.23 reflected above-average usability, indicating that the product was generally effective and supportive for task completion, although there was still room for improvement in areas such as clarity and efficiency.

As reported in Table 6, the overall UEQ score among secondary users was 0.99, suggesting a generally positive experience. A Hedonic Quality score of 1.22 indicated that users found the product somewhat engaging and emotionally appealing, although there was room for improvement to make it more exciting or innovative. Moreover, a Pragmatic Quality score of 0.76 suggested below-average usability, indicating that users experienced challenges related to efficiency, clarity, or task support.

As shown in Figure 5, the UEQ scores demonstrated statistically significant improvements from T1 to T2 across all evaluated dimensions. Hedonic Quality increased, with median values rising from 1.50 (IQR 1.75) at T1 to 1.75 (IQR 1.25) at T2 (*P*=.003). Pragmatic Quality improved from a median of 1.00 (IQR 1.00) to 1.25 (IQR 1.25) (*P*=.04). Overall experience also improved significantly, with median values increasing from 1.38 (IQR 1.25) to 1.50 (IQR 1.12; *P*=.001). These distributions are visually summarized by boxplots in Figure 5, illustrating both the positive trend and the distribution of scores at each time point.

Figure 5. Boxplots comparing the distribution of User Experience Questionnaire (UEQ) scores at 2 time points (T1 and T2). The subplots present hedonic quality, pragmatic quality, and overall experience scores, showing a general increase from T1 to T2. This trend is consistent with statistically significant improvements across all dimensions (hedonic quality: P=.003; pragmatic quality: P=.04; overall experience: P=.001).



Acceptance

Primary users reported a generally positive perception of the CAREUP platform, describing it as interesting, useful, and efficient (n=49). Frequently appreciated features included ease of use and innovative technology (n=25), health monitoring (n=20), games (n=5), the smartwatch for promoting physical activity (n=4), and the mattress sleep-tracking function (n=3).

Key issues raised included the excessive number of devices (n=7), unmet expectations over time (n=5), time-consuming technical and connectivity problems (n=20), game-related issues (n=5), repetitive questionnaires (n=5), overload caused by additional devices (n=3), and concerns regarding the reliability of the results (n=1).

Regarding recommendations, 48 participants stated that they would endorse the platform to a friend, 6 would recommend it only to individuals with strong technical skills, 3 would recommend it if improvements were made, 2 were undecided, and 2 would not recommend it. Regarding willingness to continue using CAREUP, 39 participants agreed, 8 stated they would use it occasionally, 2 would continue if fewer devices were required, 1 was undecided, and 11 declined.

Regarding the system's impact on daily routines, CAREUP improved daily routine management for 34 users, whereas 26 disagreed and 1 was undecided. CAREUP did not interfere with daily routines for 38 users; however, 21 reported interference, and 2 reported occasional interference.

Suggestions for new users included frequent use to become familiar with the system (n=17), active participation because the platform was perceived as simple and useful (n=16), maintaining confidence and patience (n=10), becoming familiar with the technology and reading the manual beforehand (n=7),

XSL•FO
RenderX

seeking help when needed (n=4), and following the provided instructions (n=3). Additional comments included requests for game enhancements (n=6), technical improvements (n=4), fewer devices (n=3), additional training or manuals to address system complexity (n=3), more exercises (n=3), clearer instructions or instructional videos (n=2), and dietary advice (n=1).

Secondary users identified the following as the best aspects of the platform: continuous health monitoring (n=8), data storage (n=10), games (n=4), integration of apps and devices (n=4), the smartwatch for promoting physical activity (n=2), and the sense of providing care and support (n=1).

Key criticisms included system complexity (n=5); difficulties in managing health results (n=2); unclear graphs (n=2); detection errors, technical issues, or a confusing interface (n=9); monitoring being perceived as a form of control (n=1); the short trial duration (n=2); and the need for prior familiarity with technology (n=1).

Regarding recommendations to other caregivers, 8 participants stated that they would endorse CAREUP, 1 would recommend it partially, 3 were undecided, and 4 would not recommend it.

Regarding willingness to continue using CAREUP, 6 participants agreed, 6 were undecided, and 5 declined.

Regarding interference with the care recipients' daily routines, CAREUP was reported not to interfere by 10 participants, to interfere by 3 participants, whereas 4 participants were undecided.

Suggestions for new users included being patient while learning to use the system (n=8), assessing older adults' interest in the platform (n=3), and maintaining real interpersonal relationships alongside technology use (n=1).

Experience-based comments highlighted the need to simplify the system and address technical issues (n=3).

All opinions regarding acceptance are summarized in Multimedia Appendix 3.

Secondary Outcomes

Data Overview

All data related to the secondary outcomes are reported in Table 7.

XSL•FO
RenderX

Table 7. Secondary outcomes.

Test	Respondents, n	Physical component summary at T0, mean (SD)	Physical component summary at T2, mean (SD)	*P* value
Physical component summary	61	46.35 (7.63)	45.88 (7.97)	.75
Nation				
Austria	8	48.67 (5.29)	49.02 (5.95)	.95
Italy	30	49.47 (6.29)	48.50 (7.46)	.76
Romania	23	41.28 (7.62)	41.11 (7.32)	.71
Gender				
Female	37	45.54 (8.01)	44.89 (8.77)	.71
Male	24	47.59 (7.00)	47.39 (6.45)	.96
Mental component summary	61	49.74 (8.56)	51.37 (9.02)	.15
Nation				
Austria	8	49.58 (8.08)	52.19 (8.10)	.57
Italy	30	48.30 (8.84)	48.30 (10.08)	.85
Romania	23	51.70 (8.37)	55.00 (6.43)	.03
Gender				
Female	37	48.76 (9.53)	50.97 (10.51)	.20
Male	24	51.24 (6.75)	51.96 (6.25)	.76
Instrumental Activities of Daily Living	61	7.75 (0.98)	7.84 (0.64)	.64
Nation				
Austria	8	8.00 (0.00)	8.00 (0.00)	>.99
Italy	30	7.93 (0.37)	7.93 (0.37)	.37
Romania	23	7.43 (1.50)	7.65 (0.93)	.23
Gender				
Female	37	7.81 (1.00)	7.92 (0.36)	.37
Male	24	7.67 (0.96)	7.71 (0.91)	.72
UCLA version 3	61	53.51 (4.88)	51.60 (10.36)	.17
Nation				
Austria	8	54.80 (3.26)	43.90 (23.28)	.21
Italy	30	53.53 (4.58)	52.93 (4.14)	.23
Romania	23	52.91 (5.82)	53.22 (5.07)	.81
Gender				
Female	37	53.74 (5.47)	51.89 (9.87)	.19
Male	24	53.16 (3.89)	51.16 (11.25)	.52
Lubben Social Network Scale—Revised	61	34.14 (7.87)	33.62 (10.29)	.50
Nation				
Austria	8	34.10 (7.89)	26.10 (15.18)	.23
Italy	30	32.77 (7.70)	33.60 (9.69)	.32
Romania	23	35.96 (8.07)	36.91 (6.58)	.09
Gender				
Female	37	34.21 (7.81)	34.26 (9.75)	.95
Male	24	34.04 (8.14)	32.64 (11.19)	.36
15-item Geriatric Depression Scale	61	2.51 (2.14)	2.26 (1.89)	.17

XSL•FO
RenderX

Test	Respondents, n	Physical component summary at T0, mean (SD)	Physical component summary at T2, mean (SD)	*P* value
Nation				
Austria	8	2.25 (1.75)	1.88 (1.73)	.17
Italy	30	2.50 (2.00)	2.67 (2.14)	.76
Romania	23	2.61 (2.50)	1.87 (1.52)	.002
Gender				
Female	37	2.65 (2.29)	2.35 (1.89)	.34
Male	24	2.29 (1.92)	2.12 (1.92)	.18
Visual Analogue Scale	61	77.67 (12.55)	79.30 (9.75)	.29
Montreal Cognitive Assessment	61	27.68 (2.29)	27.75 (2.29)	.16

SF-12 v2

The results obtained from the primary users showed that the mean physical component summary scores remained overall stable over time, with values slightly below the standardized population mean of 50. The slight decrease observed at T2 did not appear to be clinically significant. In addition, women showed lower mean scores than men at all time points, indicating poorer physical health status. Men maintained stable scores over time, with values close to the standardized mean of 50, whereas women showed a slight decline at T2. Regarding the mental component summary score, the mean value improved slightly over time, from 49.74 at T0 to 51.37 at T2. The values remained close to the standardized population average of 50, indicating a mental health status consistent with the population average. Whereas the scores for men remained stable across the 3 assessments, women showed a slight improvement in mental well-being from T0 to T2.

IADL

The IADL (Instrumental Activities of Daily Living) scores of the primary users were very close to the maximum value of 8, indicating that participants were largely independent in their daily activities. Although both women and men demonstrated a high level of independence, women showed higher scores across all 3 assessments.

UCLA v3

Scores on the loneliness scale (UCLA Loneliness Scale) showed a slight decrease over time. Overall, the scores remained within the moderate-to-high loneliness range, with a trend toward slight improvement in the social and emotional well-being of the participants (primary users). However, the results showed no significant differences according to gender (*P*=.63) or country (Austria, *P*=.21; Italy, *P*=.23; and Romania, *P*=.81).

LSNS-R

Data analysis showed that the LSNS-R scores of the primary users fell within the 31-44 range on the LSNS scale, indicating moderate social support for the group as a whole. This finding suggested that most participants had a satisfactory social support network, although a slight decline was observed over time. No significant differences were observed between women and men (Austria, *P*=.23; Italy, *P*=.32; and Romania, *P*=.09).

GDS-15

All mean GDS-15 scores (T0 and T2) were below 5, indicating that most participants did not have significant depressive symptoms. No significant differences were observed between women and men (*P*=.79).

EQ-5D-5L

Analysis of the EQ-5D-5L results showed that more than half of the sample reported no problems with mobility, self-care, or performing usual activities. Approximately half of the sample reported slight physical pain, and nearly half reported slight anxiety. The Visual Analogue Scale component of the assessment, which evaluated perceived health status, indicated a generally positive perception of health.

MOCA

The mean MOCA scores of the primary users were 27.68 at T0 and 27.75 at T2, indicating good cognitive function among the participants.

Game Experience Questionnaire, Part 3 Postgame Module

The postgame module's positive experience score of 1.55 suggested that participants were able to play the games without difficulty. No negative experiences, tiredness, or issues related to returning to reality were reported. Further details are presented in Table 8.

XSL•FO
RenderX

Table 8. Postgame module scores.

Game Experience Questionnaire score	T1, mean (SD)	T2, mean (SD)	*P* value
Positive experience	1.55 (1.11)	1.56 (1.08)	.65
Negative experience	0.39 (0.74)	0.35 (0.67)	.55
Tiredness	0.18 (0.41)	0.15 (0.36)	.56
Returning to reality	0.44 (0.54)	0.40 (0.52)	.59

Cares Workload

Overview

The secondary users (caregivers) did not report any notable changes in caregiving workload.

Devices Used During the Pilot

Users utilized the system devices at varying frequencies across different device types and pilot countries. The least-used component was the PH app, which was accessed less than once per week in both Italy and Romania. This finding highlights the importance of tailoring the positive health questionnaires to better address key aspects of older adults' social and psychological well-being. On average, the devices were used 2 times per week. The most frequently used devices were those commonly adopted by older adults, such as the oximeter and blood pressure monitor, followed by the bathroom scale. Several users engaged with these devices very actively, with 1 participant from Romania recording more than 200 weight measurements during the course of the pilot. The glucose meter and hand-grip strength meter were used less frequently. Mattress sleep sensors were used only in Italy. Their average use was approximately 3 times per week, which may have reflected periods during which the devices were unplugged or used for only part of the pilot duration.

Component usage varied across countries. The PH app was used more frequently in Austria, whose population is generally more digitally literate than those of Italy and Romania. Device use was generally similar across countries; however, users in Austria used the devices less actively than users in Italy and Romania. The greatest difference in usage frequency was observed for the hand-grip dynamometer, which was rarely used in Romania. These differences indicate the importance of considering users' digital literacy and preferences when introducing novel technological solutions. Table 9 provides detailed information on weekly system device use across device types and pilot countries.

XSL•FO
RenderX

Table 9. System devices use intensity in times per week.

Device and nation	Mean (SD)	Median	Range
Positive Health app			
Austria	1.47	1.50 (0.90)	0.28-2.92
Italy	0.45	0.66 (0.77)	0.00-3.89
Romania	0.51	0.65 (0.50)	0.08-2.14
Glucose meter			
Austria	N/A[a]	N/A	N/A
Italy	0.80	0.87 (0.49)	0.15-1.69
Romania	0.96	1.29 (1.39)	0.10-4.98
Oximeter			
Austria	1.56	1.82 (0.95)	0.49-3.15
Italy	2.75	2.83 (1.47)	0.60-6.70
Romania	1.69	2.80 (2.66)	0.16-9.48
Bathroom scale			
Austria	1.47	1.73 (1.16)	0.37-3.68
Italy	1.72	2.10 (1.53)	0.22-6.50
Romania	1.37	2.64 (4.03)	0.10-19.49
Hand-grip strength meter			
Austria	2.03	1.77 (0.96)	0.14-2.93
Italy	2.07	2.46 (1.47)	0.74-6.40
Romania	0.16	0.24 (0.21)	0.10-0.60
Blood pressure meter			
Austria	1.30	1.69 (1.38)	0.00-3.41
Italy	2.45	2.50 (1.32)	0.22-5.91
Romania	1.93	2.77 (2.64)	0.10-8.42

[a]N/A: not applicable.

Discussion

Principal Findings

The CAREUP pilot feasibility study evaluated the usability, user experience, and acceptance of an integrated care platform for tracking the IC of older individuals using various monitoring devices. Secondary outcomes included assessments of physical and mental health, quality of life, social interaction, and caregiver workload.

The CAREUP platform demonstrated both successes and challenges during the study. Although participant numbers varied across countries, no direct correlation was found between country and SUS scores. Differences in SUS ratings appeared to be more closely related to factors such as device connectivity, interface clarity, and users' digital literacy levels than to cohort size [60]. These findings suggest that local contextual and technical conditions have a stronger impact on usability perceptions than the number of participants at each site. Digital skills vary among individuals according to their level of autonomy in using technology and are influenced by socioeconomic status [61]. Specifically, Austria's lower scores highlighted issues with device connectivity and interface clarity that remained largely unresolved by the end of the trial. By contrast, Romania's favorable scores suggested that the platform could be effective and well-received when it functions smoothly.

Despite these usability challenges, many older adults (the primary users) viewed CAREUP as engaging, innovative, and empowering. Adopting a technological health solution can initially pose varying degrees of difficulty, depending on an individual's prior experience and available resources [62]. Nonetheless, most people are willing to learn and adopt new digital tools if they perceive tangible benefits for their health and well-being [63]. User experience assessments reflected a predominantly positive outlook, especially regarding the platform's "hedonic qualities"—its emotional appeal. Participants valued access to cognitive games and health-monitoring features, which introduced stimulation and novelty into their daily routines. Nevertheless, some critical feedback was noted.

Acceptance among older adults and caregivers revealed a nuanced picture. Many users expressed enthusiasm for CAREUP, appreciating its capacity to track health parameters

XSL•FO
RenderX

and support proactive health management. However, persistent technical obstacles—such as connectivity issues between devices and software—tested participants' patience. Both groups also found the number of required devices cumbersome. Although health technologies are becoming increasingly accepted in older adults' lives, partly accelerated by the COVID-19 pandemic [64], a digital divide remains, contributing to divergent views on acceptance. Although older adults typically acknowledge the benefits of technology for improving care efficiency and quality, they often struggle with adaptation. Therefore, promoting digital literacy is crucial for fostering their engagement and inclusion in the digital health care landscape [65].

As secondary users, caregivers brought a nuanced perspective to CAREUP. They valued its ability to reduce worry through continuous health monitoring and to provide structured insights into their loved ones' health. However, many also found the system unnecessarily complex. This tension between appreciation and frustration highlighted the need for targeted refinements, particularly simplifying the platform and improving reliability. For digital solutions to succeed, it is vital to involve all key end users—patients, caregivers, and health care professionals—early in development [66]. Caregivers, in particular, play a pivotal role; thus, it is essential not only to engage them but also to provide sufficient digital training. Empowering caregivers in this way enables them to support older users more effectively and substantially increases the accessibility and success of digital solutions [67]. More broadly, the CAREUP pilot exemplified both the challenges and opportunities associated with introducing technology into complex, human-centered settings. Technical reliability emerged as a major challenge, repeatedly undermining the user experience. Frequent Bluetooth issues and software glitches frustrated users and, at times, alienated them from the platform. This underscores the cost of technological shortcomings: regardless of innovation, systems must function seamlessly to earn users' trust and loyalty. Poor reliability can foster frustration, leading to abandonment and an increased risk of digital exclusion [68]. Multiple factors influence the use of digital technology, and disengagement may stem from psychological reasons (such as low motivation, lack of confidence, or perceived lack of usefulness), physical barriers (such as limited access), or practical challenges (including insufficient digital skills) [69].

The issues described are also common in other platforms evaluated in usability studies reported in the literature. Issues such as connectivity problems leading to delayed data availability [22] and difficulties understanding how to use specific components [23,27,33] are particularly prevalent in novel solutions. Interface aesthetics are also a commonly raised concern, with users often requesting personalization options to better tailor the solution to their preferences [70].

Another key takeaway from the pilot was the importance of user education and support. Although digital technologies are a valuable resource in health care, older adults still lag significantly behind younger individuals in their use of these technologies [71]. Their limited digital health literacy is influenced by both subjective factors—such as anxiety, stress related to technology use, and lack of confidence—and objective factors [72]. These include sociodemographic factors (eg, age, education level, and economic status), digital device-related factors (eg, internet access and usability challenges), and social support factors (eg, support provided by family members) [73].

Even though each pilot site organized an instructional phase at the beginning of the pilot and provided a user manual, some participants, especially those less familiar with technology, struggled to fully understand the platform's capabilities. This was particularly evident during the early phase of the trial, when users were still acclimating to the system. User comments also highlighted the potential value of video tutorials or in-person training for enhancing familiarity and confidence with the system. This may serve as a recommendation for future implementation. Given the persistent gap between the availability of digital health solutions and older adults' ability to use them effectively, ensuring adequate digital health literacy is essential [74]. To provide tailored support, it is important to first assess users' level of digital literacy, specifically their ability to search for, navigate, understand, and evaluate digital health information and tools [75].

Limitations, Implications, and Future Directions

Pilot and Implementation Considerations

This study has several limitations inherent to pilot research and the implementation of digital health interventions:

Sample Size

The relatively small sample size (n=66 enrolled; n=61 completed) limits statistical power and the generalizability of the findings. As this was an exploratory pilot study, no formal sample size calculation was performed; therefore, the results should be interpreted as formative and hypothesis-generating rather than confirmatory. The small sample size, coupled with gender imbalance and the specific national contexts of Italy, Romania, and Austria, introduces potential bias and further constrains the generalizability of the findings. These demographic and cultural differences should be considered when interpreting the results.

Selection Bias

Recruitment was conducted at 3 sites using nonrandom convenience sampling. This may have resulted in a study population that was more motivated or technologically skilled than the general older adult population, thereby introducing potential bias.

Missingness

Participant dropout and incomplete app or device use contributed to missing data, particularly for secondary outcomes that relied on self-report and device connectivity.

Device and Data Constraints

Usability challenges, technical difficulties related to device connectivity, and variability in digital literacy affected the consistency and completeness of data collection. The short intervention duration and reliance on self-reported or device-captured metrics further limit the robustness of the clinical inferences.

XSL•FO
RenderX

Technology Readiness Level

Technology Readiness Level (TRL) may have influenced the results. Before the start of the pilots, the platform components (apart from the off-the-shelf sensors) had been tested only in a laboratory or simulated operational environment (TRL 5 and TRL 6). Despite the relatively low TRL, our approach is consistent with current research standards in the field [76], which identified numerous studies in this area that remain in the development and testing phase (TRL 5) or in the demonstration, pilot, and prototypical system validation phase (TRL 6).

Mitigation and Future Directions

To address these limitations, subsequent phases of the CAREUP project will target a larger and more diverse multicenter sample and implement randomized recruitment methods where feasible. Platform development will include expanded modularity and adaptive features to support broader usability and reduce technological barriers. Clinical end points (eg, hospitalization and independent living status) and longer follow-up periods will be incorporated to capture outcomes with greater relevance and reliability. Technical upgrades, enhanced user support, and expanded device interoperability are also planned.

A comprehensive impact evaluation is scheduled within the next 18-24 months, with interim analyses and publication of progress updates at 12 months to guide iterative refinement and reporting.

Sharing both the strengths and limitations of research is fundamental to enabling more effective advancement in the field. Although new technologies increasingly demonstrate substantial utility in supporting health maintenance, self-care, independence, and autonomy—all of which contribute to improvements in quality of life—current approaches remain limited in scope.

Looking ahead, the CAREUP project revealed several valuable lessons for future development. First, robust testing and debugging should be prioritized to ensure the reliability of devices and software. Second, the platform's interface should be simplified, with clearer pathways and visual guidance to make navigation intuitive for a broad demographic. Third, reducing the number of devices or integrating functionalities could address complaints about device overload and improve user-friendliness. Finally, cultural and contextual differences, as reflected in the varied reception across countries, should inform future adaptations to ensure that the platform resonates with diverse user groups. Relevant information was continuously shared with the technical partners to support integration and development of the advanced prototype and final product.

Conclusions

The CAREUP pilot study demonstrates the potential of integrated digital health platforms to support healthy aging through IC monitoring. Despite the challenges associated with introducing technology into human-centered health care, the platform's positive outcomes highlight the promise of digital health solutions for the future. The study also underscores the importance of cultural adaptation in digital health interventions, demonstrating that optimal adoption of multidimensional health-monitoring technologies among older adults requires attention to technical reliability, user interface design, and cultural context. Variations in usability scores across countries further emphasize this need.

Technical reliability is essential for user acceptance, as connectivity issues and software glitches can undermine the user experience and lead to platform abandonment.

However, system complexity and the need for digital literacy support underscore the importance of engaging all stakeholders—patients, caregivers, and health care professionals—early in the development process.

Future implementations of digital health platforms for aging populations must address technical infrastructure, user-friendly interfaces, comprehensive education, cultural adaptation, and multidimensional health monitoring. By meeting these requirements, platforms such as CAREUP can more effectively support healthy aging and enhance the quality of life of older adults and their caregivers.

The pilot also highlighted that technology-based interventions often focus on single domains (physical, cognitive, mental, emotional, etc) rather than adopting a multidimensional approach to health promotion. Expanding research in this direction could substantially strengthen the evidence base and further advance the field.

Funding

This research was cofunded by the Active Assisted Living Programme (reference number AAL-2021-8-95-CP). In Poland, the study was funded by the Polish National Centre for Research and Development (grant AAL/AALCall2021/52/CAREUP/2022). In Italy, the study was partially supported by Ricerca Corrente funding from the Italian Ministry of Health awarded to IRCCS INRCA. In Romania, the study was supported by grants from the Ministry of Research, Innovation, and Digitization (grant CNCS/CCCDI-UEFISCDI) and the AAL Programme, with cofunding from the European Union's Horizon 2020 research and innovation programme (project number AAL-CP-2021-8-95-CAREUP-1 within PNCDI III).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Checklist of iCHECK-DH (Guidelines and Checklist for the Reporting on Digital Health Implementations) guidelines.

XSL•FO
RenderX

[DOCX File , 32 KB - aging_v9i1e79862_app1.docx]

Multimedia Appendix 2
Careup Positive Health app question selection cycle. Questions selected for each 4-week cycle day are marked with an “x.” Question identifiers correspond to the labels used in the CAREUP pilot dataset.
[DOCX File , 26 KB - aging_v9i1e79862_app2.docx]

Multimedia Appendix 3
Acceptance rating among primary users.
[DOCX File , 20 KB - aging_v9i1e79862_app3.docx]

References

1. Ageing and health. World Health Organization (WHO). Geneva, Switzerland: World Health Organization; 2024. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-11-21]
2. World population data sheet 2021: percent of population ages 65 and older. Population Reference Bureau. Washington, DC: Population Reference Bureau; 2021. URL: https://www.prb.org/world-population-data-sheet/ [accessed 2025-05-27]
3. Fiacco S, Walther A, Ehlert U. Steroid secretion in healthy aging. Psychoneuroendocrinology 2019 Jul;105:64-78. [doi: 10.1016/j.psyneuen.2018.09.035] [Medline: 30314729]
4. Zhang ZD, Milman S, Lin J, Wierbowski S, Yu H, Barzilai N, et al. Genetics of extreme human longevity to guide drug discovery for healthy ageing. Nat Metab 2020 Aug 27;2(8):663-672 [FREE Full text] [doi: 10.1038/s42255-020-0247-0] [Medline: 32719537]
5. Shi H, Hu FB, Huang T, Schernhammer ES, Willett WC, Sun Q, et al. Sedentary behaviors, light-intensity physical activity, and healthy aging. JAMA Netw Open 2024 Jun 03;7(6):e2416300 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.16300] [Medline: 38861256]
6. Belloni G, Cesari M. Frailty and intrinsic capacity: two distinct but related constructs. Front Med (Lausanne) 2019 Jun 18;6:133 [FREE Full text] [doi: 10.3389/fmed.2019.00133] [Medline: 31275941]
7. Chhetri JK, Harwood RH, Ma L, Michel J, Chan P. Intrinsic capacity and healthy ageing. Age Ageing 2022 Nov 02;51(11):afac239. [doi: 10.1093/ageing/afac239] [Medline: 36346737]
8. World HO. Integrated Care for Older People: Guidelines on Community-Level Interventions to Manage Declines in Intrinsic Capacity. Geneva, Switzerland: World Health Organization; 2017.
9. Dambha-Miller H, Simpson G, Hobson L, Olaniyan D, Hodgson S, Roderick P, et al. Integrating primary care and social services for older adults with multimorbidity: a qualitative study. Br J Gen Pract 2021 May 17;71(711):e753-e761. [doi: 10.3399/bjgp.2020.1100]
10. Landeiro F, Roberts K, Gray A, Leal J. Delayed hospital discharges of older patients: a systematic review on prevalence and costs. Gerontologist 2019 Mar 14;59(2):e86-e97. [doi: 10.1093/geront/gnx028] [Medline: 28535285]
11. Vos E, de Bruin SR, van der Beek AJ, Proper K. "It's like juggling, constantly trying to keep all balls in the air": a qualitative study of the support needs of working caregivers taking care of older adults. Int J Environ Res Public Health 2021 May 26;18(11):5701 [FREE Full text] [doi: 10.3390/ijerph18115701] [Medline: 34073386]
12. Chan A, Cai J, Qian L, Coutts B, Phan S, Gregson G, et al. In-home positioning for remote home health monitoring in older adults: systematic review. JMIR Aging 2024 Dec 02;7:e57320 [FREE Full text] [doi: 10.2196/57320] [Medline: 39622026]
13. Wang Y, Sun H, Xu S, Xia Q, Ge S, Li M, et al. Smart home technologies for enhancing independence of living and reducing care dependence in older adults: a systematic review. J Adv Nurs 2025 Jun;81(6):2885-2912. [doi: 10.1111/jan.16569] [Medline: 39445693]
14. van de Vijver S, Tensen P, Asiki G, Requena-Méndez A, Heidenrijk M, Stronks K, et al. Digital health for all: how digital health could reduce inequality and increase universal health coverage. Digit Health 2023 Jul 07;9:20552076231185434 [FREE Full text] [doi: 10.1177/20552076231185434] [Medline: 37434727]
15. Smith B, Magnani JW. New technologies, new disparities: the intersection of electronic health and digital health literacy. Int J Cardiol 2019 Oct 01;292:280-282 [FREE Full text] [doi: 10.1016/j.ijcard.2019.05.066] [Medline: 31171391]
16. Maqbool B, Herold S. Potential effectiveness and efficiency issues in usability evaluation within digital health: a systematic literature review. Journal of Systems and Software 2024 Feb;208:111881 [FREE Full text] [doi: 10.1016/j.jss.2023.111881]
17. Oudbier S, Smets E, Nieuwkerk P, Neal DP, Nurmohamed SA, Meij HJ, et al. Patients' experienced usability and satisfaction with digital health solutions in a home setting: instrument validation study. JMIR Med Inform 2025 Jan 08;13:e63703 [FREE Full text] [doi: 10.2196/63703] [Medline: 39800853]
18. Paolillo EW, Lee SY, VandeBunte A, Djukic N, Fonseca C, Kramer JH, et al. Wearable use in an observational study among older adults: adherence, feasibility, and effects of clinicodemographic factors. Front Digit Health 2022 Jun 10;4:884208 [FREE Full text] [doi: 10.3389/fdgth.2022.884208] [Medline: 35754462]
19. Kalicki AV, Moody KA, Franzosa E, Gliatto PM, Ornstein KA. Barriers to telehealth access among homebound older adults. J Am Geriatr Soc 2021 Sep 13;69(9):2404-2411 [FREE Full text] [doi: 10.1111/jgs.17163] [Medline: 33848360]

XSL•FO
RenderX

20. Comarch HomeHealth 2. Comarch. URL: https://www.comarch.com/healthcare/products/comarch-homehealth/ [accessed 2025-10-10]
21. Nicolau D, Alexandru A, Ianculescu M. An IoT, virtual machines and cloud computing-based framework for an optimal management of healthcare data collected from a smart environment. A case study: RO-smart ageing project. Informatica Economica 2019:72-83. [doi: 10.12948/issn14531305/23.3.2019.07]
22. Băjenaru OL, Băjenaru L, Ianculescu M, Constantin V, Gu atu A, Nu ă CR. Geriatric healthcare supported by decision-making tools integrated into digital health solutions. Electronics 2024 Aug 30;13(17):3440. [doi: 10.3390/electronics13173440]
23. LaMonica HM, Roberts AE, Davenport TA, Hickie IB. Evaluation of the usability and acceptability of the InnoWell platform as rated by older adults: survey study. JMIR Aging 2021 Apr 21;4(2):e25928 [FREE Full text] [doi: 10.2196/25928] [Medline: 33881410]
24. Madureira P, Cardoso N, Sousa F, Moreira W, Oliveira-Jr A, Bazzani M, et al. My-AHA: software platform to promote active and healthy ageing. Information 2020 Sep 11;11(9):438. [doi: 10.3390/info11090438]
25. Zulfiqar A, Vaudelle O, Hajjam M, Geny B, Talha S, Letourneau D, et al. Results of the "GER-e-TEC" experiment involving the use of an automated platform to detect the exacerbation of geriatric syndromes. J Clin Med 2020 Nov 26;9(12):3836 [FREE Full text] [doi: 10.3390/jcm9123836] [Medline: 33256080]
26. Zulfiqar A, Hajjam M, Hajjam A, Andres E. GER-e-TEC study: an innovative geriatric risk remote monitoring project. Exon Publications 2022:91-106. [doi: 10.36255/exon-publications-digital-health-remote-monitoring] [Medline: 35605077]
27. Calvillo-Arbizu J, Naranjo-Hernández D, Barbarov-Rostán G, Talaminos-Barroso A, Roa-Romero LM, Reina-Tosina J. A sensor-based mHealth platform for remote monitoring and intervention of frailty patients at home. Int J Environ Res Public Health 2021 Nov 08;18(21):11730 [FREE Full text] [doi: 10.3390/ijerph182111730] [Medline: 34770244]
28. Erturkmen G, Gencturk M, Akpinar AE, Pournik O, Arvanitis T, Steinhoff A, et al. The CAREPATH study for personalized management of multimorbidity in elderly with mild cognitive impairment and mild dementia. MRAJ 2024;12(8):5626. [doi: 10.18103/mra.v12i8.5626]
29. Akpinar A, Gencturk M, Pournik O, Arvanitis T, Laleci EG. User experience and usability evaluation of the CAREPATH adaptive integrated care platform. Studies in health technology and informatics 2025:283-287. [doi: 10.3233/shti250720]
30. Wang X, Vouk N, Heaukulani C, Buddhika T, Martanto W, Lee J, et al. HOPES: an integrative digital phenotyping platform for data collection, monitoring, and machine learning. J Med Internet Res 2021 Mar 15;23(3):e23984 [FREE Full text] [doi: 10.2196/23984] [Medline: 33720028]
31. Yang Z, Heaukulani C, Sim A, Buddhika T, Abdul Rashid NA, Wang X, et al. Utility of digital phenotyping based on wrist wearables and smartphones in psychosis: observational study. JMIR Mhealth Uhealth 2025 Feb 05;13:e56185 [FREE Full text] [doi: 10.2196/56185] [Medline: 39912304]
32. +Home | remote patient monitoring device. Vivify Health. URL: https://www.vivifyhealth.com/products/home-remote-patient-monitoring-device/ [accessed 2022-10-12]
33. Palumbo F, Crivello A, Furfari F, Girolami M, Mastropietro A, Manferdelli G, et al. "Hi this Is NESTORE, your personal assistant": design of an integrated IoT system for a personalized coach for healthy aging. Front Digit Health 2020 Oct 14;2:545949 [FREE Full text] [doi: 10.3389/fdgth.2020.545949] [Medline: 34713033]
34. Kolakowski M, Lupica A, Ben Bader S, Djaja-Josko V, Kolakowski J, Cichocki J, et al. CAREUP: an integrated care platform with intrinsic capacity monitoring and prediction capabilities. Sensors (Basel) 2025 Feb 03;25(3):916 [FREE Full text] [doi: 10.3390/s25030916] [Medline: 39943555]
35. Kouroubali A, Kondylakis H, Logothetidis F, Katehakis DG. Developing an AI-enabled integrated care platform for frailty. Healthcare (Basel) 2022 Feb 26;10(3):443 [FREE Full text] [doi: 10.3390/healthcare10030443] [Medline: 35326921]
36. Wilmink G, Dupey K, Alkire S, Grote J, Zobel G, Fillit HM, et al. Artificial intelligence-powered digital health platform and wearable devices improve outcomes for older adults in assisted living communities: pilot intervention study. JMIR Aging 2020 Sep 10;3(2):e19554 [FREE Full text] [doi: 10.2196/19554] [Medline: 32723711]
37. Alam M, Roy N, Holmes S, Gangopadhyay A, Galik E. Autocognisys: IoT assisted context-aware automatic cognitive health assessment. New York, NY: Association for Computing Machinery; 2020 Presented at: MobiQuitous '20: MobiQuitous 2020 - 17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services; December 7–9, 2020; Darmstadt, Germany p. 184-195. [doi: 10.1145/3448891.3448945]
38. Fan S, Ye J, Xu Q, Peng R, Hu B, Pei Z, et al. Digital health technology combining wearable gait sensors and machine learning improve the accuracy in prediction of frailty. Front Public Health 2023 Jul 20;11:1169083 [FREE Full text] [doi: 10.3389/fpubh.2023.1169083] [Medline: 37546315]
39. Wang Q, Liu J, Zhou L, Tian J, Chen X, Zhang W, et al. Usability evaluation of mHealth apps for elderly individuals: a scoping review. BMC Med Inform Decis Mak 2022 Dec 02;22(1):317 [FREE Full text] [doi: 10.1186/s12911-022-02064-5] [Medline: 36461017]
40. Kim G, Hwang D, Park J, Kim HK, Hwang E. How to design and evaluate mHealth apps? A case study of a mobile personal health record app. Electronics 2024 Jan 03;13(1):213. [doi: 10.3390/electronics13010213]
41. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]

XSL•FO
RenderX

42. Herrmann N, Mittmann N, Silver IL, Shulman KI, Busto UA, Shear NH, et al. A validation study of The Geriatric Depression Scale short form. Int J Geriat Psychiatry 1996 May;11(5):457-460. [doi: 10.1002/(sici)1099-1166(199605)11:5<457::aid-gps325>3.0.co;2-2]
43. Ware J, Kosinski M, Keller S. A 12-Item Short-Form Health Survey. Medical Care 1996 Mar;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
44. Graf C. The Lawton Instrumental Activities of Daily Living (IADL) Scale. Medsurg Nurs 2009;18(5):315-316. [Medline: 19927971]
45. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
46. Lubben J, Blozik E, Gillmann G, Iliffe S, von Renteln Kruse W, Beck JC, et al. Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. Gerontologist 2006 Aug 01;46(4):503-513 [FREE Full text] [doi: 10.1093/geront/46.4.503] [Medline: 16921004]
47. Kaiser MJ, Bauer JM, Ramsch C, Uter W, Guigoz Y, Cederholm T, MNA-International Group. Validation of the Mini Nutritional Assessment short-form (MNA-SF): a practical tool for identification of nutritional status. J Nutr Health Aging 2009 Nov 30;13(9):782-788 [FREE Full text] [doi: 10.1007/s12603-009-0214-7] [Medline: 19812868]
48. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
49. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
50. Pavasini R, Guralnik J, Brown JC, di Bari M, Cesari M, Landi F, et al. Short Physical Performance Battery and all-cause mortality: systematic review and meta-analysis. BMC Med 2016 Dec 22;14(1):215. [doi: 10.1186/s12916-016-0763-7] [Medline: 28003033]
51. Yi H, Shin K, Shin C. Development of the Sleep Quality Scale. J Sleep Res 2006 Sep;15(3):309-316. [doi: 10.1111/j.1365-2869.2006.00544.x] [Medline: 16911033]
52. Ekşi H, Kardaş S. Spiritual Well-Being: scale development and validation. SPC 2017 Feb 5;2(1):73-88. [doi: 10.12738/spc.2017.1.0022]
53. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999 Jun;70(2):113-119. [doi: 10.1080/02701367.1999.10608028] [Medline: 10380242]
54. Lanska DJ. The Romberg sign and early instruments for measuring postural sway. Semin Neurol 2002 Dec;22(4):409-418. [doi: 10.1055/s-2002-36763] [Medline: 12539062]
55. Brooke J. SUS: a quickdirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester BA, editors. Usability Evaluation in Industry. Boca Raton, FL: Taylor & Francis; 1996:4-7.
56. Schrepp M, Hinderks A, Thomaschewski J. Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S). IJIMAI 2017;4(6):103. [doi: 10.9781/ijimai.2017.09.001]
57. Norman K. Geq (game engagement/experience questionnaire): a review of two papers. Interacting with computers 2013;2013(25):278-283. [doi: 10.1093/iwc/iwt009]
58. Montgomery R. Sample size justification in feasibility studies: moving beyond published guidance. Pilot Feasibility Stud 2025 Jun 23;11(1):88 [FREE Full text] [doi: 10.1186/s40814-025-01675-9] [Medline: 40551266]
59. Kolakowski M, Lupica A, Ben BS. CAREUP-IC: dataset of health parameters and questionnaire answers for continuous monitoring of older adults' intrinsic capacity. Zenodo. URL: https://zenodo.org/records/15622154 [accessed 2025-10-10]
60. HLS19 International Report on the methodology, results and recommendations. M-POHL. URL: https://m-pohl.net/Int_Report_methdology_results_recommendations [accessed 2026-05-17]
61. Hargittai E, Piper AM, Morris MR. From internet access to internet skills: digital inequality among older adults. Univ Access Inf Soc 2018 May 3;18(4):881-890. [doi: 10.1007/s10209-018-0617-5]
62. Seifert A. The digital exclusion of older adults during the COVID-19 pandemic. J Gerontol Soc Work 2020 May 13;63(6-7):674-676. [doi: 10.1080/01634372.2020.1764687] [Medline: 32401181]
63. Quan-Haase A, Williams C, Kicevski M, Elueze I, Wellman B. Dividing the grey divide: deconstructing myths about older adults' online activities, skills, and attitudes. American Behavioral Scientist 2018 May 29;62(9):1207-1228. [doi: 10.1177/0002764218777572]
64. Tanhamira L, Randhawa G, Hewson D. Exploring the perceptions and experiences of older people on the use of digital technologies during the COVID-19 pandemic: a qualitative study. Aging Clin Exp Res 2024 Nov 14;36(1):220. [doi: 10.1007/s40520-024-02878-5] [Medline: 39541049]
65. Cabañero-Garcia E, Martinez-Lacoba R, Pardo-Garcia I, Amo-Saus E. Barriers to health, social and long-term care access among older adults: a systematic review of reviews. Int J Equity Health 2025 Mar 12;24(1):72 [FREE Full text] [doi: 10.1186/s12939-025-02429-y] [Medline: 40075452]
66. Goldsack J, Zanetti C. Defining and developing the workforce needed for success in the digital era of medicine. Digit Biomark 2020 Nov 26;4(Suppl 1):136-142 [FREE Full text] [doi: 10.1159/000512382] [Medline: 33442586]

XSL•FO
RenderX

67. Tossaint-Schoenmakers R, Versluis A, Chavannes N, Talboom-Kamp E, Kasteleyn M. The challenge of integrating eHealth into health care: systematic literature review of the Donabedian model of structure, process, and outcome. J Med Internet Res 2021 May 10;23(5):e27180 [FREE Full text] [doi: 10.2196/27180] [Medline: 33970123]
68. Kim H, Kwon IH, Cha WC. Future and development direction of digital healthcare. Healthc Inform Res 2021 Apr;27(2):95-101 [FREE Full text] [doi: 10.4258/hir.2021.27.2.95] [Medline: 34015874]
69. Heponiemi T, Jormanainen V, Leemann L, Manderbacka K, Aalto A, Hyppönen H. Digital divide in perceived benefits of online health care and social welfare services: national cross-sectional survey study. J Med Internet Res 2020 Jul 07;22(7):e17616 [FREE Full text] [doi: 10.2196/17616] [Medline: 32673218]
70. Teh P, Kwok AOJ, Cheong WL, Lee S. Insights into the use of a digital healthy aging coach (AGATHA) for older adults from Malaysia: app engagement, usability, and impact study. JMIR Form Res 2024 May 21;8:e54101 [FREE Full text] [doi: 10.2196/54101] [Medline: 38772022]
71. Lu X, Yao Y, Jin Y. Digital exclusion and functional dependence in older people: findings from five longitudinal cohort studies. EClinicalMedicine 2022 Dec;54:101708 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101708] [Medline: 36353265]
72. Arcury TA, Sandberg JC, Melius KP, Quandt SA, Leng X, Latulipe C, et al. Older adult internet use and eHealth literacy. J Appl Gerontol 2020 Feb;39(2):141-150 [FREE Full text] [doi: 10.1177/0733464818807468] [Medline: 30353776]
73. Shi Z, Du X, Li J, Hou R, Sun J, Marohabutr T. Factors influencing digital health literacy among older adults: a scoping review. Front Public Health 2024 Nov 1;12:1447747 [FREE Full text] [doi: 10.3389/fpubh.2024.1447747] [Medline: 39555039]
74. Vollbrecht H, Arora V, Otero S, Carey K, Meltzer D, Press VG. Evaluating the need to address digital literacy among hospitalized patients: cross-sectional observational study. J Med Internet Res 2020 Jun 04;22(6):e17519 [FREE Full text] [doi: 10.2196/17519] [Medline: 32496196]
75. Xie L, Zhang S, Xin M, Zhu M, Lu W, Mo PK. Electronic health literacy and health-related outcomes among older adults: a systematic review. Prev Med 2022 Apr;157:106997. [doi: 10.1016/j.ypmed.2022.106997] [Medline: 35189203]
76. Liu L, Stroulia E, Nikolaidis I, Miguel-Cruz A, Rios Rincon A. Smart homes and home health monitoring technologies for older adults: a systematic review. Int J Med Inform 2016 Jul;91:44-59. [doi: 10.1016/j.ijmedinf.2016.04.007] [Medline: 27185508]

Abbreviations

GDS-15: 15-item Geriatric Depression Scale
IADL: Instrumental Activities of Daily Living scale
IC: intrinsic capacity
iCHECK-DH: Guidelines and Checklist for the Reporting on Digital Health Implementations
INRCA: Istituto Nazionale di Ricovero e Cura per Anziani
LSNS-R: Lubben Social Network Scale—Revised
MOCA: Montreal Cognitive Assessment
PH app: Positive Health app
REST: Representational State Transfer
SF-12: 12-Item Short Form
SUS: System Usability Scale
TRL: Technology Readiness Level
UEQ-S: User Experience Questionnaire short version

Edited by A Quialheiro; submitted 30.Jun.2025; peer-reviewed by T Huang, R Diekmann; comments to author 28.Sep.2025; revised version received 16.Jan.2026; accepted 26.Mar.2026; published 28.Jul.2026.

Please cite as:
Kolakowski M, Lupica A, Bader SB, Ayadi J, Gilardi L, Consoli A, Mocanu IG, Cramariuc O, Ferrazzini L, Reithner E, Velciu M, Borgogni B, Rivaira S, Antognoli L, Leonzi S, Felici E, Rampioni M, Cucchieri G, Stara V
A Multidimensional Digital Health Platform to Support Intrinsic Capacity and Healthy Aging: Pilot Feasibility Study
JMIR Aging 2026;9:e79862
URL: https://aging.jmir.org/2026/1/e79862
doi:10.2196/79862
PMID:

©Marcin Kolakowski, Andrea Lupica, Seif Ben Bader, Jaouhar Ayadi, Luca Gilardi, Angelo Consoli, Irina Georgiana Mocanu, Oana Cramariuc, Lionello Ferrazzini, Eva Reithner, Magdalena Velciu, Barbara Borgogni, Sofia Rivaira, Luca Antognoli, Sara

XSL•FO
RenderX

Leonzi, Elisa Felici, Margherita Rampioni, Giacomo Cucchieri, Vera Stara. Originally published in JMIR Aging (https://aging.jmir.org), 28.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Music Interaction System Based on the Emotional Recognition of Older Adults in Smart Homes: Cross-Sectional Experimental Study

Chengmin Zhou[1,2], Prof Dr; Yuxuan Luo[1,2], BE; Wenjing Zhan[1,2], MSc; Ting Huang[3], MSc; Jake Kaner[4], Prof Dr

[1]College of Furnishings and Industrial Design, Nanjing Forestry University, No. 159, Longpan Road, Xuanwu District, Nanjing, Jiangsu Province, China

[2]Jiangsu Co-Innovation Center of Efficient Processing and Utilization of Forest, Nanjing, China

[3]Faculty of Natural, Mathematical and Engineering Sciences, King's College London, London, United Kingdom

[4]School of Art and Design, Nottingham Trent University, Nottingham, United Kingdom

Corresponding Author:
Chengmin Zhou, Prof Dr
College of Furnishings and Industrial Design, Nanjing Forestry University, No. 159, Longpan Road, Xuanwu District, Nanjing, Jiangsu Province, China

Abstract

Background: The advent of the Smart Home 3.0 era imposes higher technical requirements for the construction of the home ecosystem. Music is an effective means for humans to regulate their emotions. The emergence of multimodal technology facilitates the use of music for emotional regulation by older adults in their home environment. Therefore, it is necessary to carry out further research on the emotional issues of the music interaction system in smart homes that are targeted at older adult users.

Objective: This study aimed to establish the mapping relationship between music parameters and emotional states based on the Valence-Arousal emotional model and through the analysis of electroencephalogram (EEG) signals. In addition, a novel Multimodal Fusion-Adaptive Wellness Interaction model was constructed to match the most suitable music and meet the emotional needs of older adult users.

Methods: A total of 68 older adult participants aged between 60 and 75 years were recruited. Four groups of music stimulus materials were formulated to conduct an experimental study on the relationship between the emotions of older adult participants and music parameters. Finally, the Multimodal Fusion-Adaptive Wellness Interaction model was constructed. With the recreational sofa as the main body, multimodal, emotional intelligent adaptive interaction was realized.

Results: During the experiment, approximately 68,970,528 EEG signal data were collected. After analyzing the EEG bands and extracting the event-related potential data, the following results were obtained: For P200, the music stimulus materials showed significant differences and activation in the left parietal and occipital lobe regions (P7, P3, O1, O2; $P<.001$). For N400, there were significant differences and activation in the prefrontal lobe region due to the music stimulus materials (O1, CZ, O2; $P<.001$). For P300, the music stimulus materials led to significant differences and activation in the occipital lobe region and the central line region of the temporal area (F7, FPZ, FZ, F8; $P<.001$).

Conclusions: Older adult users respond in a way that the A1 stimulus material, namely the music material with positive valence and high arousal (+V + A), can soothe the emotional state of older adult users. This provides a theoretical basis and evidence for the music interaction regulation module in smart home systems.

(JMIR Aging 2026;9:e77218) doi:10.2196/77218

KEYWORDS

music emotion model; electroencephalogram; geriatric HCI; age-friendly; model building

Introduction

The concept of the Smart Home has continually evolved over the past few decades. It was first introduced in the United States in 1984, when United Technologies Building presented the “Intelligent Building” with integrated and information-based building equipment [1]. With the advancement of Internet of Things technology, smart homes entered the 2.0 era, characterized by interconnectivity and remote control [2]. The advent of the Smart Home 3.0 era imposes higher technical requirements for the construction of home scenarios and user experience, characterized primarily by scenario-based synergy and context awareness [3]. During this evolution, human-computer interaction (HCI) modes have evolved from single-channel physical interfaces, such as buttons in traditional

XSL•FO
RenderX

smart home devices, to multimodal systems that integrate multiple input modalities, including voice, gesture, and eye movement [4]. This low-burden multimodal interaction not only reduces operational difficulty but also facilitates emotional companionship and psychological comfort. As a key potential user group for smart homes, older adults warrant particular attention regarding their growing needs for both physical and emotional support. Their acceptance and use of smart home devices exhibit significant heterogeneity, with increasing engagement particularly in voice-activated and entertainment devices [5]. Emotional interaction is becoming a mainstream adaptive strategy for smart homes in domestic environments, garnering increasing scholarly focus on the emotional integration and psychological adaptation of older adults within their home settings [6]. Consequently, this project aims to explore the integration of multimodal information to achieve adaptive interaction within a smart furniture system. By focusing on behavioral and emotional interaction, it seeks to provide older adult users with an optimized experience that promotes both leisure and emotional well-being.

Among various interaction modalities, music is widely recognized as an effective medium for emotional regulation and for supporting cognitive health for older adults [7]. As a complex auditory input, music plays a crucial role in influencing emotions within the home environment [8,9]. For the older adult population, music serves as a noninvasive means of emotional regulation, helping to alleviate negative emotions such as loneliness, anxiety, and depression, while also enhancing quality of life and overall home-based well-being by evoking positive emotional experiences [10]. Previous studies have demonstrated that music listening can effectively promote social engagement and emotional health among older adults and even improve cognitive function and memory to some extent [11]. Moreover, with the accelerating process of population aging, the dependence and engagement of older adults with music continue to increase. For example, in the United States, more than 50% of older adults report listening to music daily [12]. This indicates that music has gradually become an essential part of daily recreation for older adults and is regarded as an effective approach for emotional regulation and maintaining psychological well-being [13]. Therefore, the design and application of music-based interactions in smart home environments for older adult users hold substantial research significance.

In exploring music's role in emotional regulation mechanisms, studies have found that music intervention can promote hippocampal neurogenesis and improve emotional states [14]. At the same time, music primarily regulates users' emotions by influencing neuronal activity in the brain [15]. Therefore, using physiological measurement devices to examine human emotional changes offers multiple advantages [16], as these signals can continuously and objectively reflect users' behavioral and experiential information. By combining physiological measurements with subjective emotional assessments, researchers can more accurately reveal the mapping relationship between music and emotional states [17].

Compared with subjective evaluations, physiological measurements provide a more objective and comprehensive description of emotional changes [18]. Consequently, neuroscience research, particularly studies using electroencephalography, a noninvasive neurophysiological signal, has been widely applied to emotion recognition and regulation in recent years [19]. From a neural mechanism perspective, emotional processing in the human brain relies on complex neural networks. The amygdala is responsible for evaluating the emotional significance of stimuli, whereas the insular cortex modulates emotional responses [20]. Electroencephalogram (EEG) studies related to emotion have sought to clarify the roles of different frequency bands and brain regions in emotional processing [21,22]. Research has shown that emotional content can amplify certain event-related potentials (ERPs) in EEG signals [23], while hemispheric asymmetry in theta power reflects emotional valence [24]. Moreover, variations in EEG activity in the frontal and parietal lobes, including α, β, θ/β, and σ/β ratios, are associated with individual differences in emotional regulation [25]. The integration of musical stimuli and EEG signals has gradually evolved into the field of music neuroscience [26]. For instance, studies combine analyses of musical elements with EEG data to investigate music perception and training effects [27]. Recording motor-related EEG responses evoked by music has been regarded as a key approach to understanding emotional experiences [28]. By analyzing the neural activity patterns elicited by music through EEG, researchers can gain deeper insights into the mechanisms underlying older adults' emotional experiences induced by music.

The development of multimodal technologies has enabled smart home systems to capture the relationship between users' emotions and music within home environments with greater accuracy. On the one hand, multimodal perception methods such as speech recognition, facial expression analysis, and physiological signal monitoring allow for a more comprehensive understanding of the emotional states of older adults in home settings [29]. Moreover, studies have demonstrated that different musical melodies and rhythms can elicit distinct emotional responses in users [30-33]. Integrating music with smart home technologies can thus serve as an essential multimodal pathway for emotional regulation, providing older adult users with low-burden and continuous emotional regulation support through a closed loop of perception, feedback, and intervention.

With the support of multimodal technologies, smart homes can serve as an important platform for music-assisted emotional regulation. In fact, previous studies have already begun to explore the potential of integrating smart home technologies with music-based emotion regulation. Currently, emotional regulation in smart homes is mainly achieved through 2 approaches. The first involves contact-based acquisition of physiological signals to recognize and regulate emotions. For example, Wagner et al [34] extracted 3 types of features from heart rate, electrodermal activity, electromyography, and respiration signals and used multiple classifiers for emotion recognition with promising results. The second approach relies on computer vision analysis, using cameras to capture users' facial expressions for emotion recognition and feedback. For instance, Yu et al [35] designed a prototype of smart furniture called the "Magic Mirror Table," which can infer users'

XSL•FO
RenderX

emotional states. It captures facial expressions through a camera, performs detailed analysis, and then plays appropriate sentences or music to influence users' emotions. Compared with other emotion recognition techniques, emotion regulation technologies applied in smart home environments have the advantage of passively and continuously collecting multimodal data in natural living contexts [36], thereby providing a more authentic reflection of individual emotional changes.

In the field of multimodal interaction, previous studies have explored various approaches to integrating visual, auditory, motion, and physiological signals to enhance emotion recognition and HCI performance. In the context of smart homes, multimodality enables systems to respond to user needs in a more natural and integrated manner [37]. Some studies have proposed multimodal interaction systems based on 3 different modalities, eye blinking, voice, and touch, to allow individuals with limited mobility to control smart home devices [38]. Other researchers have introduced multimodal interfaces designed to support independent living for older adults in smart home environments [39]. However, in domestic settings, key challenges for multimodal interaction include user diversity and the dynamic nature of interaction contexts [40,41]. Existing studies have primarily focused on improving home interaction by adapting to users and environmental conditions [42-44]. Current mainstream control modes in smart homes include voice and gesture interaction. Related research has also proposed a trajectory recognition interaction method based on vision, acceleration, and dynamic gestures, verifying its feasibility for integration into multimodal and adaptive smart home systems [45]. In terms of emotional regulation, multimodal interaction research has been applied to distraction detection, emotion recognition, and stress-level assessment [46]. Other scholars have investigated the relationship between multimodal fusion of emotional information and interactive behaviors. Their studies used emotional color as the theoretical basis for affective computing in HCI and applied neural network algorithms to datasets, ultimately developing an efficient multimodal affective interaction recognition method [47].

In this study, we expect to explore the relationship between different types of emotional music and the EEG changes of older adult users and derive the degree of stimulation of brain waves by other music genres, which will further perfect the emotion recognition linkage module of the smart home with the sofa as the carrier. The main contributions of this work include the following:

1. Innovative construction of the Multimodal Fusion-Adaptive Wellness Interaction (MF-AWI) model. The MF-AWI model theory can be disassembled into the Multimodal Fusion-Adaptive Wellness model. The main modules can be divided into a posture recognition module, an emotion recognition module, and a wellness HCI module. The theoretical approach guides the realization of the MF-AWI Multimodal Fusion Adaptive Wellness System. This article focuses on the emotion recognition module, relying on the theoretical basis, to achieve the real-time state of the recreation sofa on the older adult users' real-time state recognition, to make the real-time feedback of the recreation adjustment within the scope of the music adjustment, to meet the personalized needs of different older adult users in different scenarios.
2. In the module on emotional state recognition for older adult users, 4 music genres were classified based on the theory of the Valence-Arousal (VA) emotion model, and a control group was set up to explore the role of music genres in the emotional regulation of older adult users. Through the EEG physiological signals, the best music genres for older adult users in different scenarios were identified.

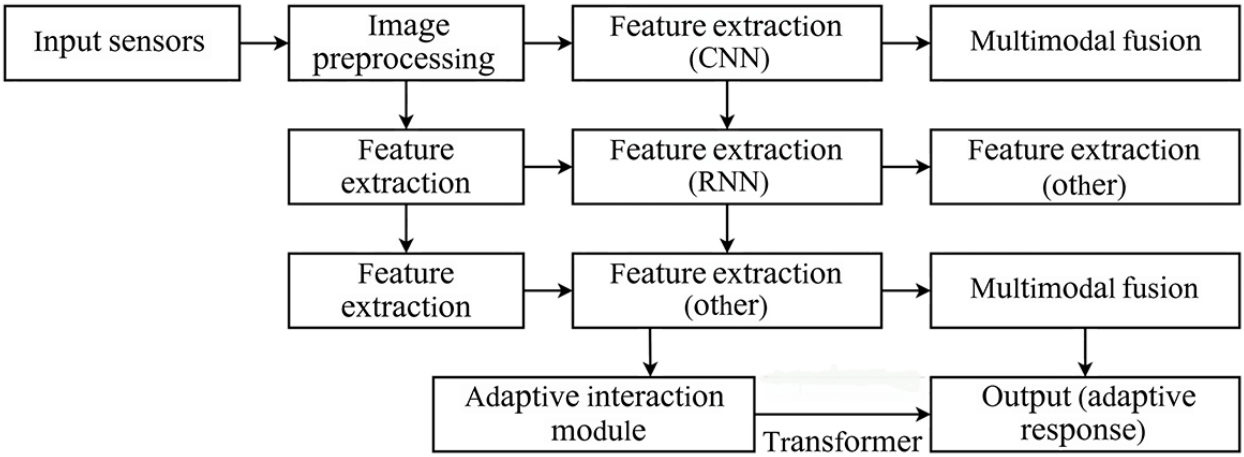
Methods

MF-AWI Model

The MF-AWI model theory can be decomposed into the MF-AWI model, which relies on the foundation of Multimodal Fusion theory to dynamically identify the target participant, locate the target participant, and use the MF-AWI model. The core of the MF-AWI model is based on the concept of multimodal fusion, that is, the integration of information from multiple sources or modalities to gain a more complete understanding of the target participant. In the context of the MF-AWI model, these modalities are poses and emotions that are used to identify the target participant and decompose the target state into its components, as shown in the recognition process in Figure 1.

XSL•FO
RenderX

Figure 1. The Multimodal Fusion-Adaptive Wellness Interaction model responds to process iterations. CNN: convolutional neural network; RNN: recurrent neural network.



The unified representations will be fed into the Wellness HCI module, which will use the pretrained Transformer model to generate adaptive feedback. The Wellness HCI module can recognize different gestalt states and emotional states of the older adult users, enabling the system to provide appropriate adaptive feedback to different older adult users' inputs. For example, if the user indicates that they want to go to a resting state based on their posture and emotional state, the Wellness HCI module can suggest that the wellness sofa adjust the backrest angle to give the older user a more comfortable position, or play soothing music to create a more relaxing environment. Similarly, if the older adult users appear to have negative emotions such as stress or anxiety, the Wellness HCI module can suggest relaxation exercises or provide calming music to help the user relax. The output of the wellness HCI module will be used to control the different module functions of the wellness sofa, including the adjustment of functional modules such as angles and music. This process can be repeated to provide a personalized and adaptive interactive experience for the older adult user on the wellness sofa. In this model, different input data from the wellness sofa is processed using a convolutional neural network (CNN) model, and a multimodal fusion technique is used to combine these output feature vectors. A unified representation of the older adult users' real-time posture and real-time emotion will be used by the wellness HCI module to generate control of the different adaptive modules to produce appropriate adaptive feedback interactions.

The camera captures video signals, which are then preprocessed by resizing and grayscale conversion to enable face detection of older adult users. Using the OpenCV library, once the user's face is detected by applying a pretrained machine learning model, the DeepFace model is used to extract facial features and identify the user's emotional state. Using a CNN, the extracted facial features are used to identify the user's emotion and determine sentiment tuning. Once the real-time emotional state of the older adult users is identified, the type of music required by the recreational sofa can be determined to suit the users' changing emotions. Adjusting the music of the wellness sofa can be done using the emotional state adjustment determined in the previous step to adjust the music of the wellness sofa to suit the emotional state of the user. Finally, the process can be repeated many times to ensure that the Convalescent Sofa adapts to the real-time emotions of the older adult users in real time, as shown in Textbox 1.

XSL•FO
RenderX

Textbox 1. Emotion recognition module framework.

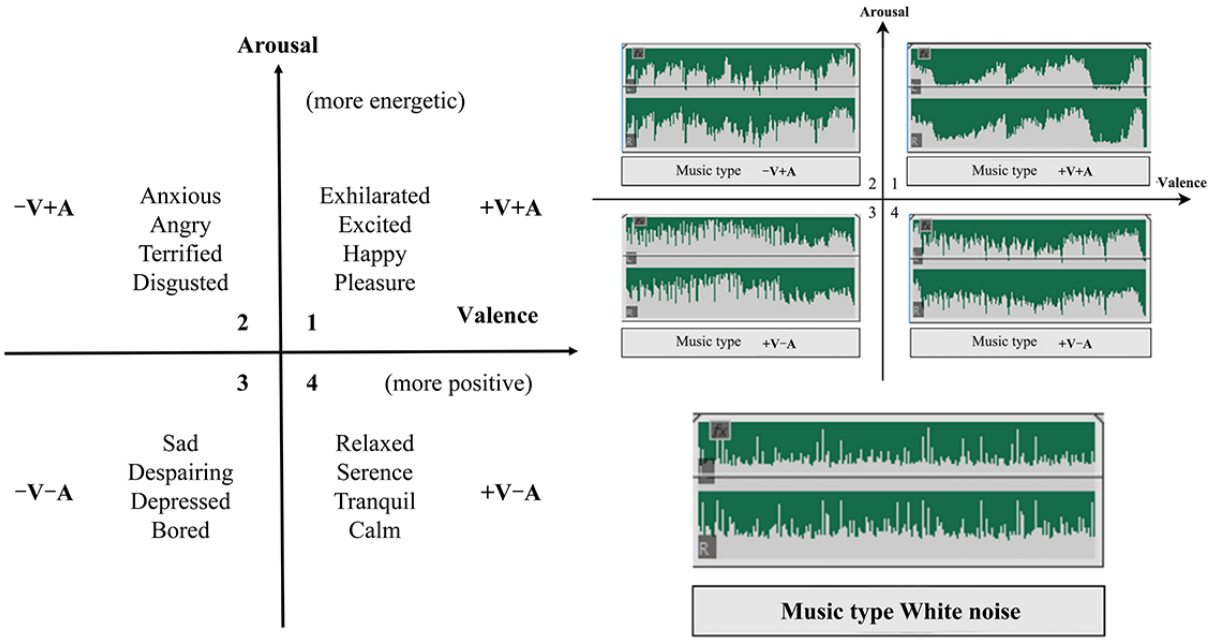
```
# Import necessary libraries
import cv2
import numpy as np
from deepface import DeepFace
# Capture the video feed
cap = cv2.VideoCapture(0)
……
for (x,y,w,h) in faces:
# Extract the facial features using pretrained machine learning models
face_roi = frame[y:y+h, x:x+w]
img = cv2.resize(face_roi, (224, 224))
img = np.expand_dims(img, axis=0)
img = img / 255.0
embeddings = DeepFace.represent(img)
# Recognize the user's emotion using pretrained machine learning models
svm = cv2.ml.SVM_load('emotion_detection_svm.xml')
pred = svm.predict(embeddings)[1][0][0]
# Determine the emotion adjustment needed for the smart sofa
if pred == 0:
emotion_adjustment = 5
elif pred == 1:
emotion_adjustment = -5
```

Materials

The generic continuous dimension model that has been widely adopted in music emotion recognition research is the Circomplex model, also known as the VA model [48]. Research shows that specific musical characteristics are significantly correlated with Valence and Arousal [49]. For example, +V +A refers to a fast tempo, major key, and bright timbre. –V+ A refers to dissonant chords and a fast, irregular rhythm. –V –A refers to a slow tempo, minor key, and dull timbre. +V A refers to slow yet harmonious harmonies. Existing studies have demonstrated that bodily sensations and emotions triggered by music and based on the VA model show consistency across different cultures [50]. In this experiment, 4 pure music genres conforming to the 4 quadrants of the VA model were selected, and the stimulus materials were named A1-A4, that is, A1: +V +A, A2: –V+A, A3: –V–A, A4: +V A, as shown in Figure 2. A control group was set up using white noise to simulate white noise to simulate the daily life environment, and the stimulus material was named A0. The use of white noise provided a constant and semantically neutral auditory background, helping to avoid interference from background sounds on participants' emotional responses [51].

XSL•FO
RenderX

Figure 2. VA emotion models as well as experimental stimulus material. VA: Valence-Arousal.



Participants

Sample size is critical to the reliability and representativeness of research findings. Following the standards proposed by Cohen [52], this study used G*Power software (Heinrich Heine University Düsseldorf, Germany) [53] to estimate the required sample size, selecting the within-factors repeated measures ANOVA design. The parameters included an alpha level of 0.05, a power level of 0.95, a medium effect size of 0.25, and a number of groups of 1. This setup yielded a total sample size requirement of 31 participants. In addition, to ensure greater statistical robustness and account for potential data loss among older adult participants during the experiment, the final sample size was further adjusted. This sample size is also consistent with previous EEG studies investigating music perception and stimulation in similar experimental contexts [54,55]. Therefore, a total of 68 older adult participants aged between 60 and 75 years were recruited for this experiment. All participants were recruited from various regions across China, with culturally diverse backgrounds and different living environments.

Among the participants, there were 34 female participants (n=34) with an average height of 160 cm and a mean BMI of 25.5, and 34 male participants (n=34) with an average height of 171 cm and a mean BMI of 24.86. All participants were required to be free of chronic diseases of the ear, back, and legs and in good physical condition.

Ethical Considerations

This experiment was conducted in strict accordance with ethical guidelines. Prior to the start of the experiment, the experimental procedures were fully explained to all participants, ensuring that the experiment would not pose any risks to them, and written informed consent was obtained from all participants. Also, explicit consent has been granted by a participant to use their images for publication. The experiment has been reviewed and approved by the Ethics Committee of Nanjing Forestry University (Science and Technology Department of Nanjing Forestry University) with the approval number: 2023-03-29. To protect the privacy of participants, all collected data were anonymized before analysis, with no personal identifying information retained. Upon completion of the experiment, each participant received a compensation of 100 RMB (1 RMB= US $0.14) for their time and participation.

Experimental Paradigm Design

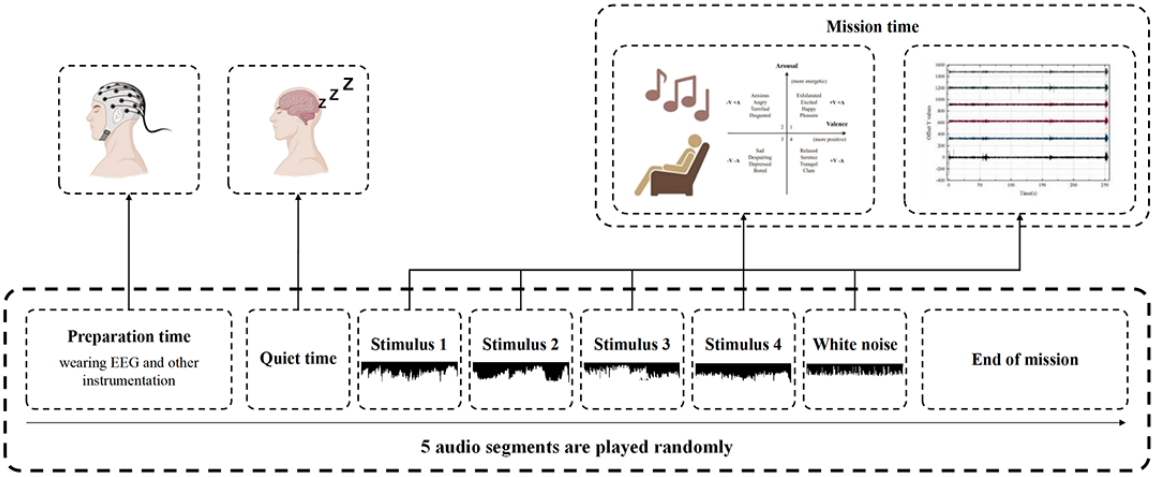
The older adult participants were seated calmly in a comfortable chair. The experimenter prepared 5 preanalyzed audio clips, including 4 audio segments based on the VA model, as the stimulus material. Meanwhile, a white-noise environment was established as the control condition to ensure a consistent auditory ambience. The decibel level of these audio clips was 45 dB. This level was determined with reference to previous studies, which indicate that 45 dB falls within the range of low-noise and quiet environments for older adults, which helps maximize auditory clarity while preventing auditory fatigue or discomfort [56,57]. Although background noise levels in real smart home environments may exceed 45 dB, this value was chosen to obtain stable and reproducible ERP data in a controlled and quiet experimental setting, thereby minimizing the impact of external interference on older adult participants.

Prior to the start of the experiment, the experimental content and precautions were carefully explained to the participants. However, no information regarding the specific types of audio stimuli or their emotional attributes was disclosed to minimize expectation effects. Participants were also instructed to avoid vigorous movements or vocalization. After obtaining the participants' consent, the informed consent form was signed, and preparations for the experiment, including scalp cleaning, began. The experimental equipment was worn by older adult

XSL•FO
RenderX

participants throughout the experiment. The first formal recording was started after the participant indicated full understanding and initial control of the EEG signal. During the experiment, the system continuously monitored electrode connectivity. When head movement or electrode displacement caused channel dropouts or abnormal impedance, the acquisition interface issued a real-time alert. The experimenter immediately paused the task, readjusted the electrodes, and resumed recording once stable contact was restored, ensuring the consistency and comparability of EEG data across all recording stages. The EEG signals were recorded continuously in each session and saved separately. All trial tasks were presented in a randomized order to effectively minimize order and learning effects. The duration of the entire experiment was 15 - 20 minutes, and the experimental paradigm is shown in Figure 3.

Figure 3. Experimental paradigm. The device worn by the participants was the Bitbrain hydroelectric electrode EEG, which had 16 channels, a rated voltage of 3.7 V, and a single-channel sampling rate of 256 Hz. The experiment started after the older adult participants entered the resting state. The experiment was divided into 5 recording stages, which were 5 sets of subexperiments, as a whole. In each stage, EEG signals were continuously recorded and saved separately, while each trial task was randomly ordered. EEG: electroencephalogram.

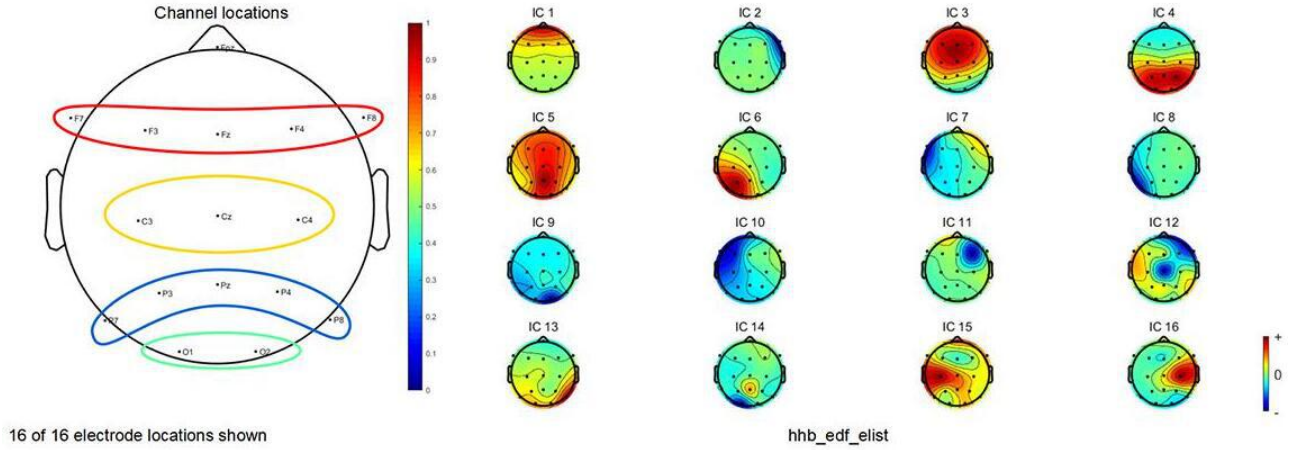


EEG Recording and Preprocessing

The EEG signal was acquired according to the international 10 - 20 system lead setup, with 16 electrodes selected, mainly placed in the frontal, parietal, and central regions, and covering the left and right hemispheres of the brain and the midline region, as shown in Figure 4. This study used a 16-channel EEG acquisition system, a configuration widely used in research on music-induced emotion and cognitive processing [58,59]. Although the number of electrodes was relatively limited, the layout covered key brain regions associated with emotional processing, auditory stimulation, and attentional allocation, effectively capturing the main activity characteristics of typical ERP components such as P200, P300, and N400 [60]. In addition, considering that the participants in this study were older adults, using fewer electrodes helped maintain signal quality while reducing physiological fatigue and discomfort caused by prolonged experiments, thereby improving experimental compliance and data stability. Based on the experimental acquisition parameters, it was estimated that approximately 65,313 data points could be collected per channel for each participant. After excluding 2 invalid samples, the total size of the raw EEG dataset in this study was 65,313×16×66, amounting to approximately 68,970,528 signal data points. This data volume resulted from the normal accumulation determined by the device sampling parameters and experimental design, without any redundant or oversampling processes; therefore, it did not introduce bias into the results.

XSL•FO
RenderX

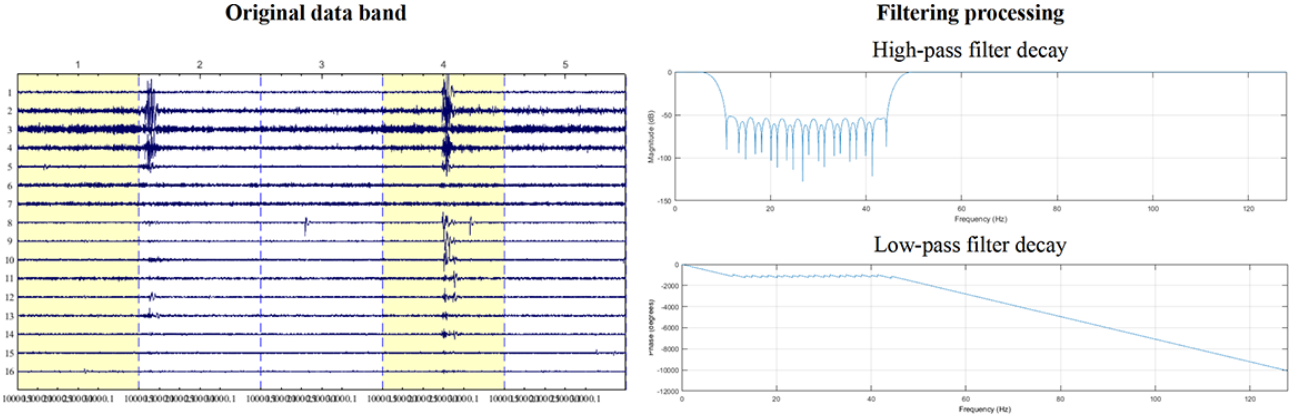
Figure 4. Diagram of the 16 electrode positions selected according to the international 10‐20. EEG electrodes in the 10‐20 system. The area within the red circle represents the frontal lobe (Fpz, **F7, F3, FZ, F4, and F8**), the yellow area represents the central region (**C3, CZ, and C4**), the blue area represents the parietal lobe (**P3, PZ, P4, P7, and P8**), and the green area represents the occipital lobe (**O1 and O2**). EEG: electroencephalogram.



During the acquisition process, the reference and ground electrodes were placed near the Cz electrode and the anterior part of the scalp, respectively. EEG signal preprocessing was conducted using MATLAB with the EEGLAB 2020 toolbox. The acquired .edf files were imported into EEGLAB 2020, and channel localization was performed based on the MNI Coordinate Systems standard brain imaging model. Sixteen channels were identified as Fpz, F7, F3, Fz, F4, F8, C3, Cz, C4, P7, P3, Pz, P4, P8, O1, and O2, as shown in Figure 4.

EEG data preprocessing was performed on the Jinfa EEG Signal Analysis Platform. First, all recordings were visually inspected by researchers to remove segments with obvious artifacts or discontinuities. A zero-phase finite impulse response filter was then applied for band-pass filtering (0.1‐30 Hz), and a notch filter was used to eliminate 50 Hz power-line noise. The original and filtered signals are shown in Figure 5. The data were subsequently rereferenced to the average and interpolated for any bad channels. Artifact removal was based on the standardized pipeline of Infomax independent component analysis and canonical correlation analysis built into the ErgoLAB Man-Machine-Environment Testing Cloud Platform (ErgoLAB 3.0). Identification of artifact components referenced conventional criteria commonly adopted in EEG research. Independent component analysis–derived components were identified and rejected based on their temporal characteristics, scalp topographies, and spectral features corresponding to ocular or muscular artifacts. To further improve signal quality, canonical correlation analysis was applied to remove residual artifacts. For the ERP analysis, the event-related time window was set from −200 to 800 milliseconds, with a baseline period of −200 milliseconds and a measurement window of 200 to 400 milliseconds.

Figure 5. Comparison between the original band of the electroencephalogram and the data processed by filtering.



EEG Bands

The EEG bands in this experiment are divided into 5 bands. Gamma waves play an important role in the cognitive activity of the human brain and in higher-level activities such as the transmission of information in the brain, integrated processing, and feedback [61-63]. Beta waves, with amplitudes of 5-20 µV, are the fastest of the EEG waves and are more pronounced in the frontal and central regions [64]. Alpha waves, with amplitudes of 20-100 µV, are the fastest of the EEG waves and are more pronounced in the occipital and parietal regions [65]. Theta waves have an amplitude of 20 to 150 µV and are the slowest of the EEG waves, generally more pronounced in the parietal and temporal regions [66-69]. Delta waves, which typically exhibit amplitudes of up to approximately 200 µV, are the slowest type of EEG activity and are mainly observed in the frontal and occipital regions and generally occur during deep sleep, making them difficult to capture during the waking state [70-73]. The brain can be divided into parietal, temporal, frontal, and occipital lobes, of which the parietal lobe is mainly

XSL•FO
RenderX

responsible for processing internal feedback, such as skeletal muscles and limbs [74-77].

ERP Data

ERP is a high-level functional potential of the human brain [69]. ERPs are changes in brain potentials associated with cognitive processes such as judgment, attention, perception, decision-making, and working memory content [78]. Recent studies have shown that the amplitude and latency of ERPs are considered to be very important indicators of cognitive load in humans.

There are 3 key parameters of ERP, namely amplitude, latency, and distribution on the scalp [79]. By calculating the average amplitude of the ERP within the corresponding time window, the differences between different conditions can be analyzed. The latency represents the time interval from the onset of the stimulus to the occurrence of the ERP peak, usually expressed in milliseconds. The distribution of the ERP on the scalp can reveal which areas of the brain are functioning when the stimulus appears. The ERP with these 3 key parameters can reveal individuals' cognitive and emotional processing under specific conditions [80,81].

ERPs are typically defined as time-locked neural responses to brief stimuli. Nevertheless, previous studies have shown that ERP components with typical temporal characteristics can also be observed under continuous or natural musical stimulation, particularly when different audio segments exhibit distinct conditional variations [82-84]. In this experiment, each music clip was treated as an independent event type. The onset of each music clip was set as the event trigger point (time=0 milliseconds), and a time window of −200 to 2000 milliseconds was extracted with a 200 milliseconds baseline correction applied. For each participant, the trial waveforms of the same music type were extracted and averaged to obtain the ERP, which reflects the cognitive and emotional processing of different music types in older adults.

Results

EEG Band Analysis

This experiment mainly collected the EEG changes of the older adult participants in the leisure state, so the electrode position of the parietal lobe was mainly selected as the key observation object, and these 3 bands of θ, α, and β were selected as the main discriminating parameters of the brain fatigue in this experiment. The power of the EEG bands under the stimulation materials of A0-A4 is shown in Table 1.

XSL•FO
RenderX

Table . Electroencephalogram band power under A0-A4 stimulation material.

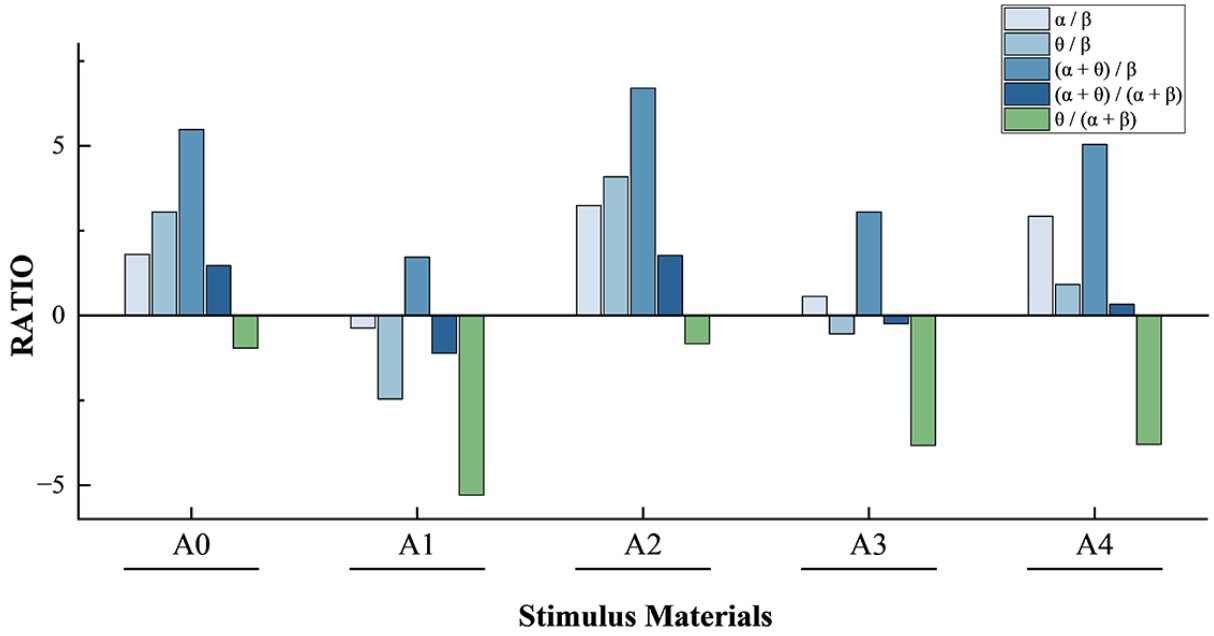
Stimulation materials and project	δ	θ	α	β	γ
A0					
Total power	−4.89	16.43	15.18	13.38	8.29
Average power	−9.66	10.41	7.40	1.34	−4.50
Percentage power (%)	0.31	41.54	31.18	20.6	6.38
A1					
Total power	−25.49	10.18	12.27	12.64	9.31
Average power	−30.27	4.16	4.49	0.60	−3.48
Percentage power (%)	0.01	19.22	31.14	33.89	15.74
A2					
Total power	−2.08	20.35	19.51	16.26	9.28
Average power	−6.85	14.33	11.72	4.22	−3.51
Percentage power (%)	0.25	43.56	35.82	16.97	3.4
A3					
Total power	−18.42	12.18	13.27	12.71	7.91
Average power	−23.19	6.16	5.49	0.67	−4.88
Percentage power (%)	0.02	26.36	33.91	29.83	9.87
A4					
Total power	−18.39	13.29	15.29	12.37	6.15
Average power	−23.16	7.26	7.51	0.33	−6.64
Percentage power (%)	0.02	27.84	44.20	22.56	5.38

The total power value of the theta band was 20. Thirty-five at the highest in stimulus material A2 and occupied 43.56% of the power in this stimulus band, which can be monitored in the sleepy state, especially in the depressed or frustrated state. This indirectly verifies that the stimulus material A2, that is, –V+ A, can induce negative emotions such as sleepiness and frustration. The total power value of the alpha band accounted for the highest percentage of power, which was 44.20% when using stimulus material A4. Clinical research has shown that this is related to spatial attention. Therefore, the stimulus material A4 (ie,+ V A) may facilitate higher levels of creative thinking in older adult participants during cognitive tasks [85]. For stimulus material A1, the beta band's total power value accounted for the highest percentage, 33.89%. Clinically, this band is linked to emotional arousal, indicating that A1 (+ V +A) can easily trigger emotional fluctuations in older adult participants.

The ratio of theta to beta wave energy, F(θ/ β), and the ratio of the sum of theta and alpha wave energy to beta wave energy, F(α + θ) / β, are used as characteristic quantities for EEG fatigue assessment [47,48]. As F(θ / β) rises, the degree of fatigue further increases. The EEG band ratios for the A0 to A4 stimulus material states were derived by further processing of the ratios in Table 1, as shown in Figure 6.

XSL•FO
RenderX

Figure 6. Brain wave band ratio under A0-A4 stimulation material state.



This study found that different musical stimuli have an impact on the mental workload of older adult participants to some extent. The F(α + θ) / β value was much lower than the F(α + θ) / β values for the other 4 musical stimulus states when the stimulus material was A1. So the EEG fatigue level of the older adult participants was lowest when the stimulus material was A1, that is, +V+A, and highest when the stimulus material was A2, that is, –V+A. The F (α + θ) / β value of stimulus material A3 indicates that it can be selected as an auxiliary stimulus in moderate situations. The effect of music stimulation on the brain load of the older adult participants was therefore reflected in some way. It can be seen that music stimulates the brain waves of older adult participants and that different types of music can be used in daily life to induce emotions in older adult participants. +V+A music, such as exhilaration, excitement, and happiness, helps to reduce fatigue, while music of the −V+A type (eg, evoking anxiety, anger, or panic) induces negative emotions and increases brain load [86].

ERP Data Analysis

The frontal zone electrodes (F7, F8, FPZ, and FZ), central zone electrodes (CZ), parietal zone electrodes (P3, P4, P7, P8, and Pz), and occipital zone electrodes (O1 and O2) were selected for the analysis. EEG data from each condition were analyzed using Statistical Product Service Solutions (SPSS; IBM Corp) software. Given that the objective of this study was to explore differences in EEG activity among older adult participants under various music stimulation conditions, one-way ANOVA with least significant difference (LSD) post hoc comparisons was used to maintain sensitivity in detecting potential effects during this exploratory phase. The regions (left, middle, and right) were divided for the evoked P200, P300, and N400 EEG components, as shown in Table 2. The selection of P200, P300, and N400 components as analytical indicators was based on their respective neurophysiological significance in emotional and cognitive processing. The P200 component typically occurs approximately 200 milliseconds after stimulus onset and can be regarded as an initial perceptual response to auditory stimulation [87]. The P300 component usually emerges around 300 milliseconds after stimulus onset, reflecting neural activity associated with attention allocation, working memory, and cognitive evaluation processes [88]. The N400 component generally appears approximately 400 milliseconds after stimulus onset and represents a neural response to semantic incongruity or expectancy violation [89]. By analyzing these 3 components, individual emotional responses can be comprehensively evaluated across different stages of cognitive processing.

Table . Division of evoked potentials into brain regions.

EEG[a] component	Left area	Middle area	Right area
P200	P7, P3, O1	PZ	P4, P8, O2
P300	F7	FPZ, FZ	F8
N400	O1	CZ	O2

[a]EEG: electroencephalogram.

Statistical Analysis of P200 Components

For the P200 EEG component, the latency averages of each electrode (P7, P3, O1, PZ, P4, P8, and O2) were obtained by total superimposed averaging of the participants' waveforms in a total of 5 stimulus material states from A0 to A4, as shown in Table 3.

XSL•FO
RenderX

Table . A0-A4 mean latency for each electrode of the P200 component in the stimulated.

Stimulation materials	P7	P3	O1	PZ	P4	P8	O2
A0	0.094	0.086	0.095	0.088	0.093	0.091	0.088
A1	0.189	0.067	0.165	0.087	0.184	0.086	0.154
A2	0.034	−0.147	−0.099	−0.470	0.027	0.303	0.108
A3	0.072	0.003	0.050	0.080	0.181	0.079	0.110
A4	0.053	0.062	0.680	0.069	0.072	0.060	0.074

Analysis of the mean values of the latency of each electrode for the P200 component in the A0-A4 stimulation material condition shows that there were differences in the mean values of the latency of each electrode for the P200 component. In the A0 stimulation condition, the mean value of the O1 channel was the largest, and the mean value of the P3 channel was the smallest; in the A1 stimulation condition, the mean value of the P7 channel was the largest, and the mean value of the P3 channel was the smallest. In the A2 stimulation condition, the mean value of the P8 channel was the largest, and the mean value of the PZ channel was the smallest. In the A3 stimulation condition, the mean value of the P4 channel was the largest, and the mean value of the P3 channel was the smallest. In the A4 stimulation condition, the mean value of the O2 channel was the largest, and the mean value of the P7 channel was the smallest. The P200 component data obtained were imported into the SPSS mathematical analysis software. One-way ANOVA was used to test the influence of different stimulus materials on the P200 components of each electrode, and the means and SDs were calculated, as shown in Table 4. To further analyze the differences, we used the LSD method for post hoc multiple comparisons, as depicted in Figure 7.

XSL•FO
RenderX

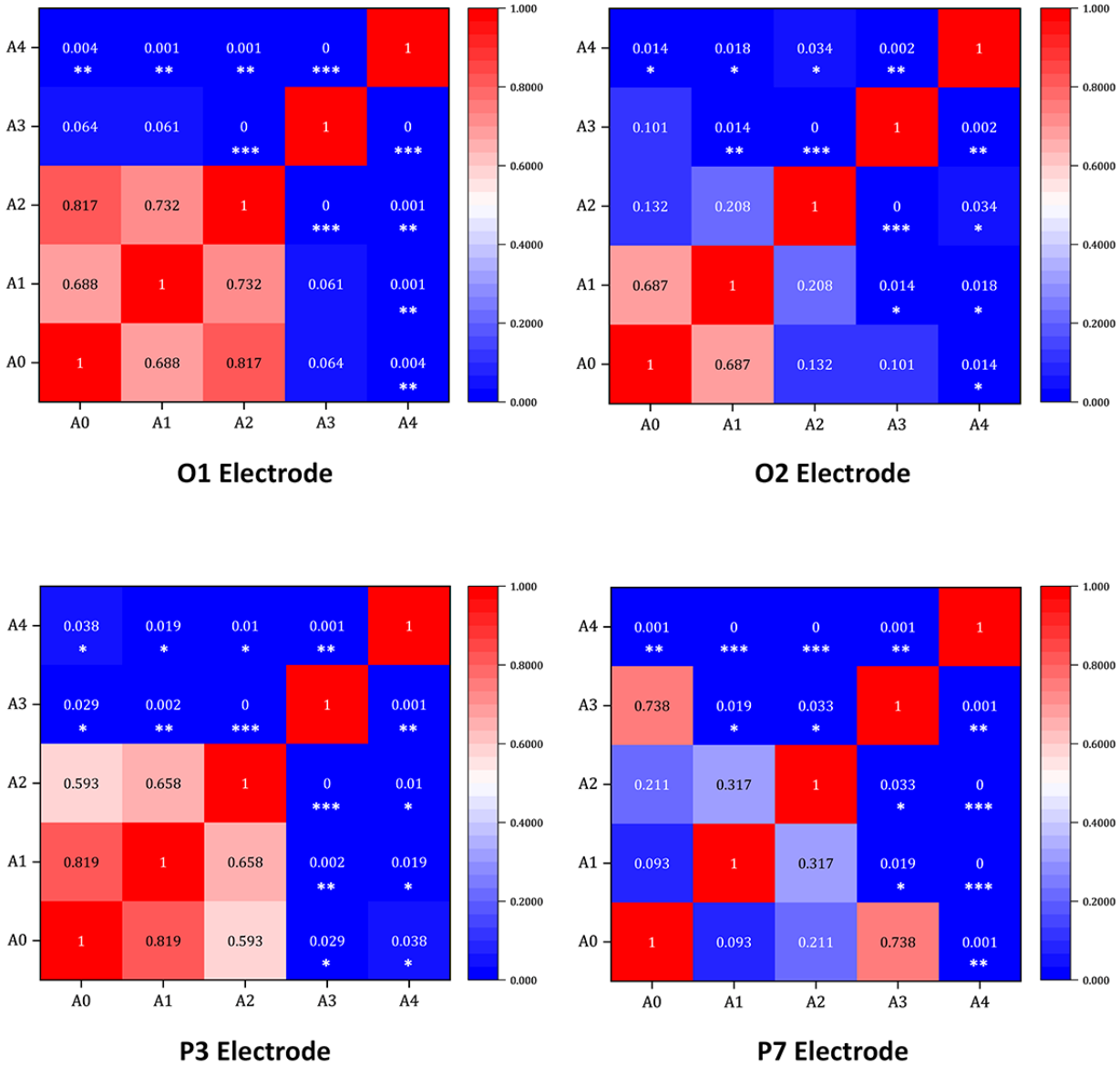
Table . P200 component ANOVA results.

Variable and value	Mean (SD)	F_1-score	*P* value
P7		12.342	<.001
A0	2.382 (8.783)		
A1	−0.276 (6.735)		
A2	0.750 (2.604)		
A3	2.903 (6.550)		
A4	18.507 (32.256)		
Total	4.853 (16.933)		
P3		6.243	<.001
A0	1.638 (9.222)		
A1	1.290 (5.440)		
A2	0.914 (2.482)		
A3	−1.369 (2.579)		
A4	7.605 (17.856)		
Total	2.016 (9.838)		
O1		9.931	<.001
A0	1.002 (7.944)		
A1	0.450 (5.555)		
A2	0.734 (1.840)		
A3	−1.279 (3.282)		
A4	8.307 (15.511)		
Total	1.843 (8.930)		
PZ		7.216	<.001
A0	0.552 (8.025)		
A1	1.358 (3.445)		
A2	1.009 (1.253)		
A3	−2.218 (2.567)		
A4	5.871 (14.469)		
Total	1.314 (8.039)		
P4		10.441	<.001
A0	1.536 (6.285)		
A1	0.954 (3.973)		
A2	1.262 (1.878)		
A3	1.262 (3.524)		
A4	7.185 (14.409)		
Total	1.734 (8.014)		
P8		1.683	.155
A0	0.330 (0.492)		
A1	0.598 (4.628)		
A2	1.224 (1.939)		
A3	−1.699 (3.888)		
A4	2.007 (14.895)		
Total	0.492 (7.590)		

XSL•FO
RenderX

Variable and value	Mean (SD)	F_1-score	*P* value
O2		6.408	<.001
A0	−0.126 (5.572)		
A1	0.280 (4.399)		
A2	1.092 (1.087)		
A3	−1.623 (3.135)		
A4	6.474 (17.694)		
Total	6.474 (9.023)		

Figure 7. Analysis of the P200 heat map under the states of stimulus materials A0-A4. Each cell in the heatmap represents the *P* value of the pairwise comparison between conditions, with deeper colors indicating more significant differences. The diagonal cells represent self-comparisons and are displayed as 1. The numerical values in the figure correspond to the *P* values from the least significant difference post hoc test, and significance levels are denoted as **P*<.05, ***P*<.01, and ****P*<.001.



Further analysis of the potential amplitudes of the P200 component revealed that there were significant differences in different types of musical stimulus materials on the electrodes P7, P3, O1, and O2. For the variable P7 ($F_{4, 268}$=12.342, *P*<.001,

XSL•FO
RenderX

η^2=0.168), the LSD test showed that the amplitude of the A1 stimulation (mean −0.276) was significantly lower than that of the A4 stimulation (mean 18.507, *P*<.001). For the variable P3 ($F_{4, 268}$=6.243, *P*<.001, η^2=0.092), the LSD test showed that the amplitude of the A2 stimulation (mean 0.914) was significantly lower than that of the A4 stimulation (mean 7.605, *P*=.010). For the variable O1 ($F_{4, 268}$=9.931, *P*<.001, η^2=0.140), the LSD test showed that the amplitude of the A1 stimulation (mean 0.450) was significantly lower than that of the A4 stimulation (mean 8.307, *P*=.001). For the variable O2 ($F_{4, 268}$=6.408, *P*<.001, η^2=0.095), the LSD test showed that the amplitude of the A0 stimulation (mean −0.126) was significantly lower than that of the A4 stimulation (mean 6.474, *P*=.014).

A detailed analysis of the heatmap shown in Figure 7 revealed that A4 stimulation, that is, +V A, had the strongest activation, with significantly higher average potential amplitudes across all channels compared with other stimuli and the A0 control. Musical stimuli also differentially activated the left parietal and occipital lobes.

Statistical Analysis of P300 Components

For the P300 EEG component, the voltage-phase EEG data of each electrode in the period of 150 - 250 milliseconds were extracted. The latency averages of each electrode of P300 (F7, FPZ, FZ, and F8) were obtained by averaging the superimpositions of the participants' waveforms in 5 stimulus material states of A0-A4, as shown in Table 5.

Table . A0-A4 mean latency for each electrode of the P300 component in the stimulated.

Stimulation materials	F7	FPZ	FZ	F8
A0	0.037	0.078	0.073	0.112
A1	−0.015	−0.407	0.071	−0.010
A2	−0.527	−0.470	0.274	0.303
A3	−0.340	−0.336	0.141	0.189
A4	0.046	0.167	0.126	0.109

Analysis of the latency averages for each electrode of the P300 component in the A0-A4 stimulus material condition shows that there were differences in the latency averages for each electrode of the P300 component. Under the A0 stimulation condition, the mean value of the F8 channel was the largest, and the mean value of the F7 channel was the smallest. Under the A1 stimulation condition, the mean value of the FZ channel was the largest, and the mean value of the FPZ channel was the smallest. Under the A2 stimulation condition, the mean value of the F8 channel was the largest, and the mean value of the F7 channel was the smallest. Under the A3 stimulation condition, the mean value of the F8 channel was the largest, and the mean value of the F7 channel was the smallest. Under the A4 stimulation conditions, the mean value of the FPZ channels was the largest, and the mean value of the F7 channels was the smallest. The obtained data of the P300 component were calculated for the means and SDs and subjected to one-way ANOVA, as shown in Table 6. The LSD method was used for post hoc multiple comparisons to further analyze the differences, as illustrated in Figure 8.

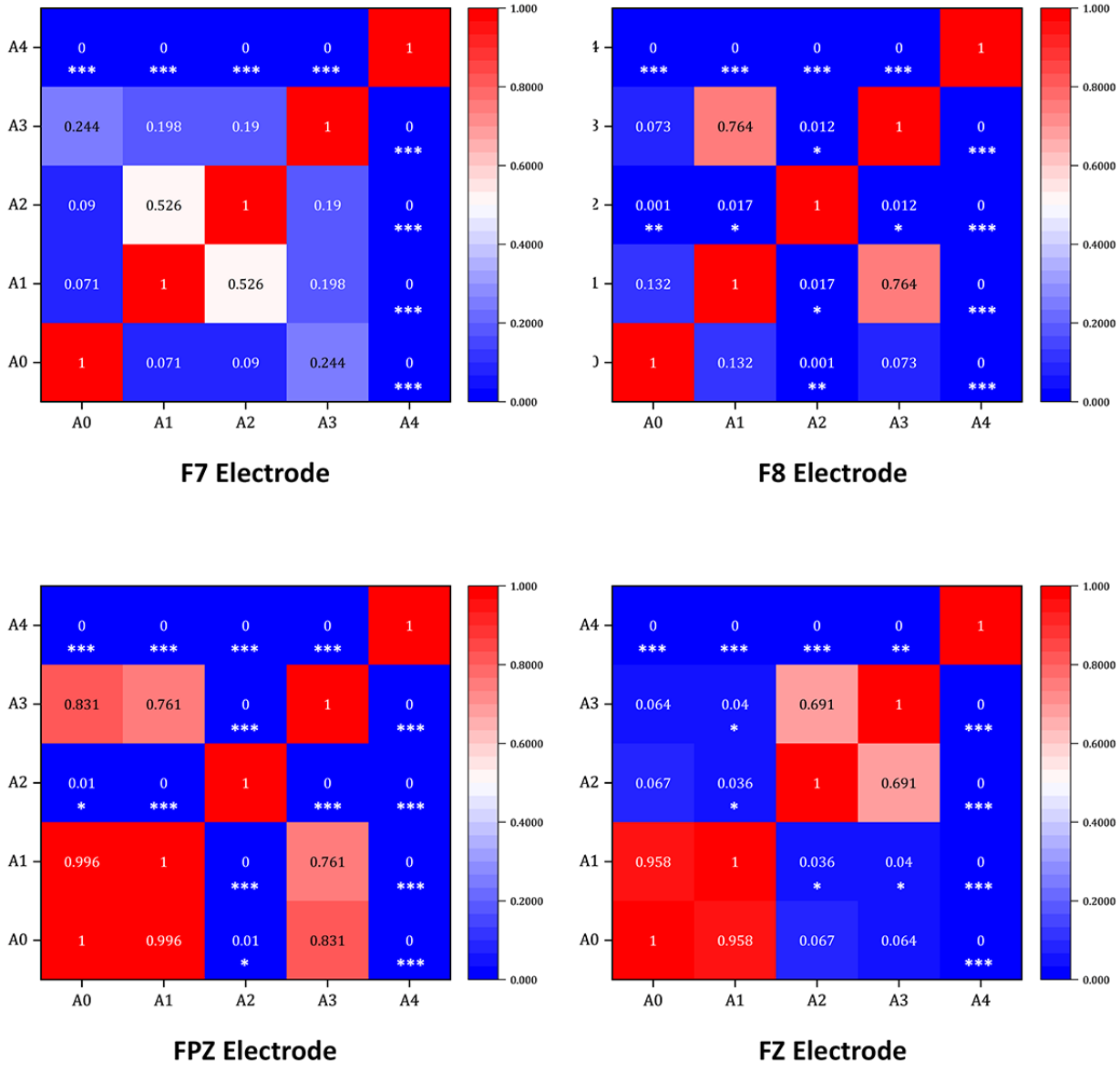
XSL•FO
RenderX

Table . P300 component ANOVA results.

Variable and value	Mean (SD)	F_1-score	*P* value
F7		17.065	<.001
A0	1.608 (6.967)		
A1	−0.482 (4.099)		
A2	−0.101 (1.062)		
A3	0.387 (2.386)		
A4	13.404 (20.879)		
Total	2.963 (11.307)		
FPZ		668.888	<.001
A0	1.044 (4.771)		
A1	−0.152 (2.882)		
A2	−1.249 (1.351)		
A3	−0.307 (2.232)		
A4	44.444 (28.429)		
Total	8.756 (22.062)		
FZ		30.449	<.001
A0	−1.152 (7.616)		
A1	−1.146 (4.626)		
A2	−4.109 (2.355)		
A3	−0.900 (3.342)		
A4	57.083 (12.804)		
Total	9.955 (24.699)		
F8		117.799	<.001
A0	0.990 (7.455)		
A1	0.918 (6.245)		
A2	−0.986 (1.086)		
A3	−1.197 (3.580)		
A4	12.585 (12.766)		
Total	2.462 (8.945)		

XSL•FO
RenderX

Figure 8. Analysis of the heat map of P300 at each electrode under the states of stimulus material A0-A1. Each cell in the heatmap represents the *P* value of the pairwise comparison between conditions, with deeper colors indicating more significant differences. The diagonal cells represent self-comparisons and are displayed as 1. The numerical values in the figure correspond to the *P* values from the least significant difference post hoc test, and significance levels are denoted as **P*<.05, ***P*<.01, and ****P*<.001.



A further analysis of the potential amplitudes of the P300 component revealed that there were significant differences and activation phenomena in the prefrontal lobe region for the musical stimulation materials. Specifically, for the variable F7 ($F_{4, 268}$=17.065, *P*<.001, η^2=0.218), the LSD test showed that the average potential amplitude under the A2 stimulation material (mean −0.101) was much lower than that under the A4 stimulation material (mean 13.404, *P*<.001). For the variable FPZ ($F_{4, 268}$=668.888, *P*<.001, η^2=0.916), the LSD test showed that the average potential amplitude under the A1 stimulation material (mean −1.146) was much lower than that under the A4 stimulation material (mean 57.083, *P*<.001). For the variable FZ ($F_{4, 268}$=30.449, *P*<.001, η^2=0.332), the LSD test showed that the average potential amplitude under the A3 stimulation material (mean −1.197) was much lower than that under the A4 stimulation material (mean 12.585, *P*<.001). For the variable F8 ($F_{4, 268}$=117.799, *P*<.001, η^2=0.658), the LSD test showed that the average potential amplitude under the A1 stimulation material (mean −0.152) was much lower than that under the A4 stimulation material (mean 44.444, *P*<.001).

A detailed analysis of the heatmap shown in Figure 8 reveals that there are significant differences among the different types of musical stimulation materials at the F7, FPZ, FZ, and F8 electrodes. The average potential amplitudes of the variables for each channel under the A4 stimulation material, that is, +V A, were much higher than those under the other 3 types of stimulation materials and the A0 stimulation material of the control group.

XSL•FO
RenderX

Statistical Analysis of N400 Components

For the N400 EEG components, the voltage EEG data of each electrode in the period of 150 - 250 milliseconds were extracted. By averaging the total superposition of the participants' waveforms in the 5 stimulus material states from A0-A4, the latency averages of each electrode of the N400 (O1, CZ, and O2) were obtained, as shown in Table 7.

Table . A0-A4 mean latency for each electrode of the N400 component in the stimulated.

Stimulation materials	O1	CZ	O2
A0	0.095	0.088	0.072
A1	0.165	0.120	0.154
A2	−0.099	0.116	0.108
A3	0.050	0.136	0.110
A4	0.069	0.076	0.074

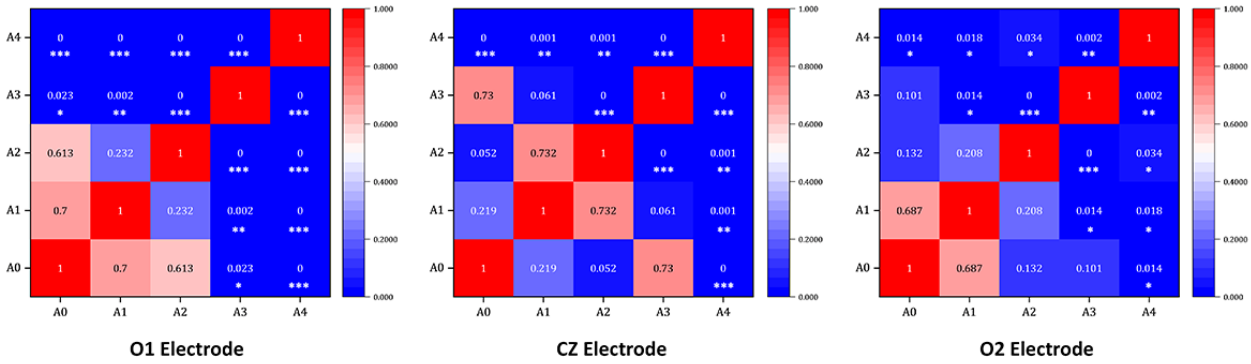
Analysis of the mean latency values of each electrode of the N400 component under the stimulation material conditions of A0-A4 shows that there are differences in the mean latency values of each electrode of the N400 component. Under the A0 stimulation condition, the mean value of the O1 channel was the largest, and the mean value of the O2 channel was the smallest. Under the A1 stimulation condition, the mean value of the O1 channel was the largest, and the mean value of the CZ channel was the smallest. Under the A2 stimulation condition, the mean value of the CZ channel was the largest, and the mean value of the O1 channel was the smallest. Under the A3 stimulation condition, the mean value of the CZ channel was the largest, and the mean value of the O1 channel was the smallest. Under the A4 stimulation condition, the mean value of the CZ channel was the largest, and the mean value of the O1 channel was the smallest. The obtained data of the P300 component were calculated for the means and SDs and subjected to one-way ANOVA, as shown in Table 8. The LSD method was used for post hoc multiple comparisons to further analyze the differences, as illustrated in Figure 9.

Table . N400 component ANOVA results.

Variable and value	Mean (SD)	F_1-score	*P* value
O1		63.016	<.001
A0	1.002 (7.944)		
A1	1.568 (6.642)		
A2	0.430 (0.808)		
A3	−1.806 (3.316		
A4	21.713 (15.859)		
Total	4.581 (12.160)		
CZ		12.095	<.001
A0	−0.954 (5.789)		
A1	0.450 (5.555)		
A2	0.734 (1.840)		
A3	−1.279 (3.282)		
A4	8.307 (15.511)		
Total	1.452 (8.672)		
O2		6.408	<.001
A0	−0.126 (5.572)		
A1	0.280 (4.399)		
A2	1.092 (1.087)		
A3	−1.623 (3.135)		
A4	6.474 (17.694)		
Total	1.219 (9.023)		

XSL•FO
RenderX

Figure 9. Analysis of the heat map of N400 at each electrode under the states of stimulus material A0-A1. Each cell in the heatmap represents the *P* value of the pairwise comparison between conditions, with deeper colors indicating more significant differences. The diagonal cells represent self-comparisons and are displayed as 1. The numerical values in the figure correspond to the *P* values from the least significant difference post hoc test, and significance levels are denoted as **P*<.05, ***P*<.01, and ****P*<.001.



A further analysis of the potential amplitudes of the N400 component reveals that there are obvious differences and degrees of activation in the occipital lobe region and the central line region of the temporal area caused by the musical stimulation materials. The analysis shows that for the variable O1 ($F_{4, 268}$=63.016, *P*<.001, η^2=0.507), the LSD test reveals that the average potential amplitude under the A2 stimulation material (mean 0.430) is much lower than that under the A4 stimulation material (mean 21.713, *P*<.001). For the variable CZ ($F_{4, 268}$=12.095, *P*<.001, η^2=0.165), the LSD test indicates that the average potential amplitude under the A1 stimulation material (mean 0.450) is much lower than that under the A4 stimulation material (mean 8.307, *P*<.001). For the variable O2 ($F_{4, 268}$=6.408, *P*<.001, η^2=0.095), the LSD test shows that the average potential amplitude under the A0 stimulation material (mean −0.126) is much lower than that under the A4 stimulation material (mean 6.474, *P*<.001).

A detailed analysis of the heatmap shown in Figure 9 reveals that there are significant differences among different types of musical stimulation materials at the O1, CZ, and O2 electrodes. In the N400 component, the average potential amplitudes of the variables for each channel under the A4 stimulation material, that is, +V A, are much higher than those under the other 3 types of stimulation materials and the A0 stimulation material of the control group.

Discussion

Principal Results

This study aims to establish a mapping relationship between musical parameters and emotional states based on a music emotion model, and to provide empirical evidence for the design of a music-interaction module for intelligent leisure sofas in home environments. Based on the VA model, 4 sets of musical stimulus materials were formulated, and the state responses of older adult participants while listening to different musical stimulus materials were collected by the Bitbrain hydropolar EEG imaging system. The experimental results indicated that different types of musical stimuli exerted significant effects on the EEG activity of older adults.

First, compared with the white noise control condition, all 4 types of musical stimuli exhibited significant differences in EEG activation levels, indicating that music stimulation can effectively influence the neural workload of older adult participants. This finding is consistent with previous research on the regulatory effects of music on neural activity. For instance, Kučikienė et al [90] reported that preferred music enhanced EEG power across multiple frequency bands, particularly within the α/β bands. Similarly, Yang et al [91] found significant associations between brain activity in different frequency bands and musical rhythm. In the present study, we observed that +V+A pleasurable, excited, and happy types of music elicited higher levels of EEG activation, which may be associated with increased α and β power and decreased θ activity, changes that facilitate positive emotional experiences and reduce fatigue [92]. In contrast, –V+ A anxious, angry, and extremely panicky types of music were characterized by higher proportions of θ activity and peak (α+θ)/β ratios, which may correspond to EEG mechanisms commonly observed during negative emotional stimulation, such as enhanced θ and suppressed α power [93].

Analysis of the ERP data revealed significant differences among the 4 types of musical stimuli for the P200, P300, and N400 components, reflecting the modulatory effects of music on cognitive processing and emotional responses in older adult participants. Regarding the P200 component, our study showed significant differences across electrode sites P7, P3, O1, and O2, primarily localized in the left parietal and occipital regions. This suggests that music can activate the brain's perceptual and attentional networks at an early processing stage, facilitating the initial encoding of auditory information. Similarly, Patel et al [94] reported that musical stimuli could evoke early-stage neural responses for attention and perceptual processing, with the P200 component being closely linked to perceptual and selective attention processes. In addition, Portnova et al [95] found that the perception of emotional stimuli in comatose patients was associated with early P200 responses, whereas healthy individuals exhibited more complex ERP patterns.

Regarding the P300 component, its amplitude is closely associated with attention, working memory, and emotional states

XSL•FO
RenderX

[96]. Evidence suggests that musical stimuli can modulate brain activity during later stages of cognitive processing, significantly influencing P300 amplitude. For instance, Pawlowski et al [97] noted that the P300 component is strongly related to cognitive evaluation and decision-making processes. In the present study, significant differences in P300 amplitudes were observed across electrode sites O1, Cz, and O2, with particularly prominent differences in activation in the occipital and temporal regions. These findings indicate that musical stimuli of different emotional types can significantly affect the allocation of attention and the level of cognitive evaluation in older adult participants, reflecting the specific modulatory role of music in brain cognitive processing.

Regarding the N400 component, although N400 was originally studied in the context of language processing, recent research has extended its application to the field of music to investigate the neural mechanisms underlying musical semantic processing. For example, Calma-Roddin et al [98] found that disharmonious or unexpected changes in familiar melodies elicited an N400-like effect similar to that observed in language, suggesting that music processing also involves semantic-related cognitive mechanisms. In our analysis of the N400 component, significant differences were observed across electrode sites F7, Fpz, Fz, and F8, particularly in the frontal region. This finding is consistent with previous research, indicating that musical stimuli can evoke brain neural responses associated with semantic processing and emotional evaluation. For instance, Portnova et al [95] reported that the N400 component is closely related to emotional perception and semantic processing, and that higher N400 amplitudes have been linked to the perception of emotional expression in sounds.

In summary, different types of musical stimuli exhibit significant differences in the brain activity of older adult participants. These results provide empirical evidence for the neural mechanisms underlying the regulatory effects of music on the emotions and cognition of older adults and offer a basis for the design of music-interaction regulation modules in smart home systems. From a theoretical perspective, the findings of this study further enrich the research framework regarding emotion-driven adaptability of home environments in smart home design [99] and support the necessity of introducing emotional response mechanisms in home environments [100]. Meanwhile, this study also provides empirical supplementation to the affective-rational theory [101,102]. The experimental results indicate an interaction between emotional experience and rational processing in the older adult users under musical stimulation. On this basis, the proposed music feedback mechanism can achieve adaptive regulation of users' emotions by identifying the corresponding relationship between emotions and music, which aligns with the laws of human cognitive processing and people-oriented HCI models [103]. Finally, through a comprehensive analysis of musical stimuli and EEG responses, this study reveals the emotional interaction relationships of the older adult users in the context of smart homes. As Desmet et al [104] noted, emotional artifacts promote well-being and psychological balance by evoking positive emotional experiences. This provides an actionable theoretical basis and experimental support for the future design of emotional products for older adult users and the development of health-promoting interaction systems.

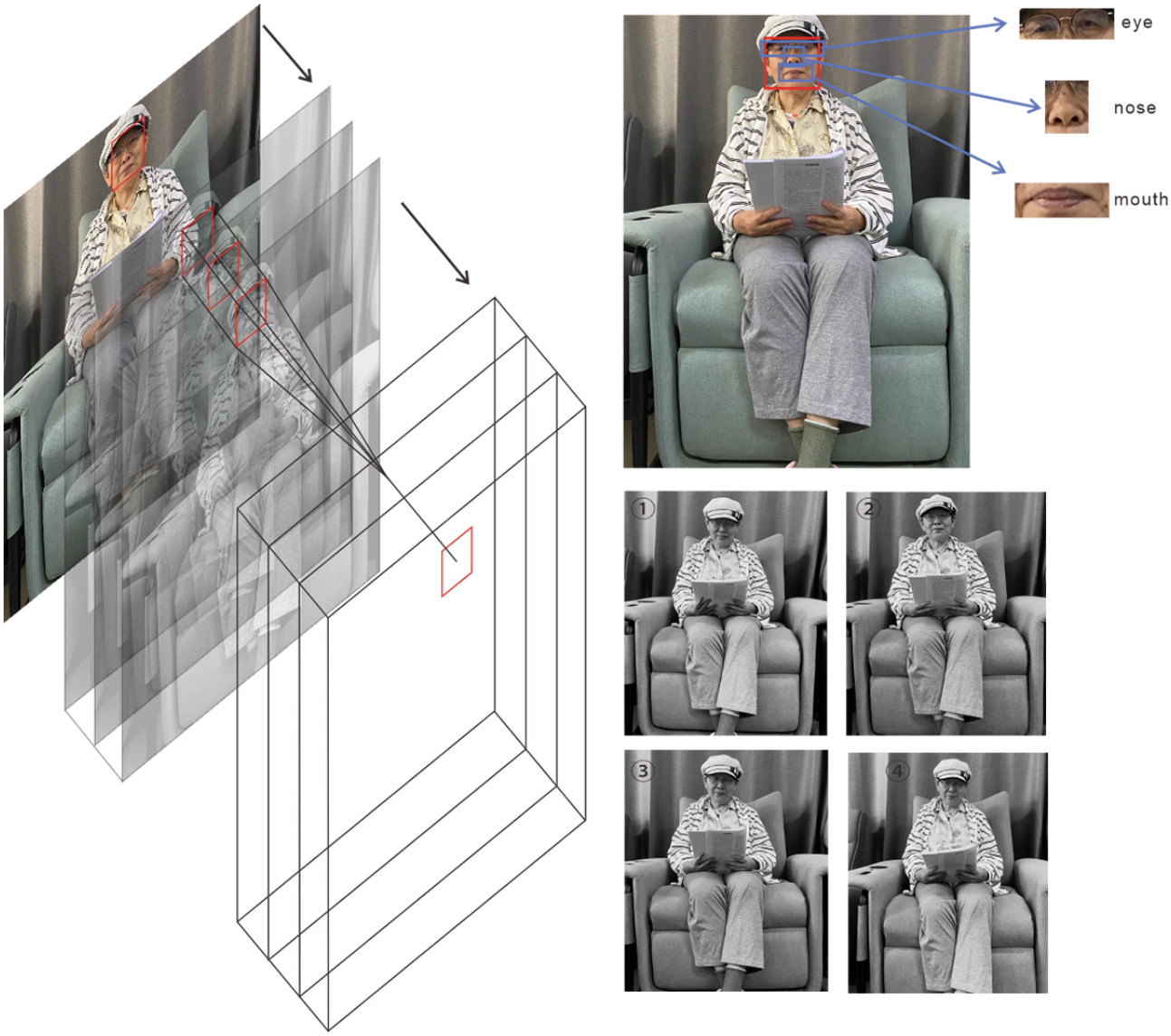
Practice

The emotion recognition module plays a key role in the multimodal HCI system of the Recreation Sofa, which provides an immersive and personalized musical experience for the user. The module is designed to recognize the current behavioral and emotional state of the older adult user and adjust the music and interaction accordingly according to predefined criteria in the recreation module. This involves the integration of multiple sensors and algorithms to detect the emotional, behavioral, and activity levels of the older adult user and select music that matches their preferences and needs. For example, if an older user is feeling stressed or tired, the emotion recognition module can select calming or relaxing music to help them unwind. Or, if the older user is feeling energetic or happy, the music module can select more upbeat or lively music to enhance their emotions. If the older user is in a relaxed state, the emotion recognition module will select calm and soothing music to match the user's emotional and behavioral state. The emotion recognition module also adjusts the volume and tempo of the music to the user's preferences, ensuring a comfortable and enjoyable musical experience when using the lounge sofa. The emotion recognition module thus plays an important role in providing a personalized and adaptive user experience that is tailored to the user's emotional and behavioral state. The emotion recognition module uses the predefined music matching criteria set in the Recreation module to select the appropriate type of music.

The emotion recognition module carries out facial recognition of the older adult users through the Intel RealSense camera. The facial features are divided through the emotion recognition module, which mainly identifies the 3 main parts of the older adult users' face, the eyes, nose, and corners of the mouth, and the recognition process is shown in Figure 10. During the recognition process, the captured image will be grayed out to determine the gender of the target user, establish where the target facial range of the user lies, extract the 3 main facial feature parts of the target user in turn, and determine the actual emotional state of the user through a comprehensive evaluation of the algorithm. At this stage, the module has been developed as a prototype to test the feasibility of facial emotion recognition and its integration effect in the interactive music system. The study adopts the CNN-DeepFace framework to process the facial images of 68 older adult participants in the EEG-music experiment. These images serve as the initial dataset for verifying the effectiveness of the emotion mapping and adaptive playback process. The purpose of this stage is not to train a new deep learning model, but to evaluate the technical feasibility and response accuracy of the system in a controlled environment.

XSL•FO
RenderX

Figure 10. Model foundation framework.



The emotion recognition and music interaction module developed based on the conclusions of this experimental study aims not only to enhance user experience through its technical implementation, but also needs to be examined within the broader context of national, cultural, and ethical frameworks [105]. At the societal level, this study responds to the global trends of intelligent older adults' care and home-based health promotion [106]; it seeks to improve the psychological and emotional well-being of the older adults in home environments through noninvasive emotion perception and matched music interventions. At the cultural level, the randomized controlled experiment conducted using EEG in this study prioritizes cultural universality in music selection and intervention strategies. Music based on the VA model was selected to avoid emotional biases or culturally insensitive interventions caused by cultural differences, thereby ensuring the cross-cultural acceptability of the experiment [107].

At the level of technical ethics and safety practice, the design integrating the interaction module with the leisure sofa adheres to the design guidelines for HCI safety [108]. The main hardware involved includes a speech recognition microphone array, an emotion recognition camera, and an intelligent audio playback device. At the communication level, the system primarily relies on wireless communication protocols such as Wi-Fi, Bluetooth, and LoRa, while realizing multimodal data interaction and scenario linkage through a smart home gateway. However, subsequent research and practice need to consider data security, a critical consideration in the smart home ecosystem, which requires balancing the protection of user privacy with the migration and linkage between smart products [109]. From the perspective of ethics and social well-being, this system aims to provide older adults with an intelligent companionship solution that offers emotional comfort and is oriented toward social care. It enables music to serve not only as feedback for product functions but also as a carrier for emotional care, which aligns with the goal of technology facilitating a better quality of life [110]. Meanwhile, the system should also emphasize that technology should be people-oriented, explainable, and supportive of social dialogue. This requires prioritizing design values over technical implementation, rather than merely pursuing performance optimization [111].

XSL•FO
RenderX

Comparison With Prior Work

The analogies between the findings of this paper and the existing literature related to music emotion recognition or regulation are shown in Table 9.

Table . Correlation of the findings of this paper with the existing literature on musical emotion recognition and regulation.

Comparison	This work	Yang et al [112]	Takashima et al [113]	Wang et al [114]	Sutcliffe et al [115]	Vieillard and Gilet [116]
Research participants	Older adults	General	Datasets	Youth and older adults	Youth and older adults	Youth and older adults
Research methods	Model building, experiments, physiological data analysis	Model building, performance testing	Model building, dataset experiments	Experiments, physiological data analysis	Experiments, physiological data analysis	Experiments, physiological data analysis
Research tools	EEG[a] and ERP[b]	Multimodal research	EMER-CL[c]	Residual convolutional network	Research	Electromyogram (EMG)
Application scenarios	Age-friendly, smart sofa	Multimedia tools	Encoder training	Music generation	Deepen research	Emotion regulation

[a]EEG: electroencephalogram.

[b]ERP: event-related potential.

[c]EMER-CL: embedding-based music emotion recognition using composite loss.

These studies all aim to expand the possibilities of music emotion recognition from different perspectives, thus enabling mutual supplementation and validation. Through comparative and correlational analysis with other literature, this study demonstrates significant differentiation and innovation in several aspects. First, unlike previous studies that mainly targeted the mixed-age groups [112,113], this study specifically focuses on the older adult population, aiming to address the specific emotional regulation needs of the older adults in smart home environments. Methodologically, this study obtained the physiological responses of older adult participants through EEG and ERP analyses and established an emotional state mapping in combination with musical parameters. The MF-AWI model proposed in this study integrates multimodal data and health-promoting HCI, which differs from existing multimodal emotion recognition studies in terms of model framework. Examples include the rhythm prediction model by Wang et al [114], or the experiments on young and mixed-age groups by Sutcliffe et al [115] and Vieillard and Gilet [116]. The proposed model not only considers music and emotion recognition but also emphasizes individual adaptability in home interaction scenarios. Furthermore, numerous studies have demonstrated that multimodal fusion methods outperform single-modality approaches in emotion recognition tasks. For instance, when combining EEG with facial expressions for emotion classification, multimodal fusion consistently achieves higher accuracy than unimodal methods [117]. Therefore, the MF-AWI model incorporates emotional gradients and health-promoting thresholds in feature fusion, enabling personalized emotional feedback, and its preliminary validation in the intelligent home sofa scenario further highlights the potential of multimodal emotion recognition and regulation for practical adaptability and operability.

Limitations

This study has several limitations. First, regarding the participant population, the sample was restricted to older adults aged 60 - 75 years. Future research should further refine user needs and personal habits by examining behavioral habits and individual preferences, as well as potential gender differences in emotional responses and interaction behaviors. Second, all participants in this study were healthy older adults, as the primary aim was to establish a foundational model for music–emotion interaction within smart home systems. However, individuals with cognitive decline, dementia, or other impairments, who represent key potential beneficiaries of such systems, were not included [118]. Since cognitive processing and emotional responses may differ between healthy and cognitively impaired older adult populations, the generalizability of the findings remains to be verified in special older adult groups. Future studies should therefore expand the sample to include older adults with varying levels of cognitive function to comprehensively assess the applicability and effectiveness of the music interaction module across diverse user groups. Comparative studies between groups with different health statuses are also recommended to better understand how cognitive conditions influence emotional and behavioral responses to such interactive systems.

At the model and system level, this study primarily focused on using experimental data to validate the feasibility of the multimodal emotion recognition and adaptive music-matching mechanisms within the MF-AWI model. The study is positioned as a mechanism-level validation. In addition, the CNN–DeepFace emotion recognition component in this study mainly served to verify the feasibility of integrating visual emotion detection into the interactive music system, rather than to develop a fully trained and optimized recognition model. Consequently, systematic quantitative evaluations of the multimodal emotion recognition module and the adaptive feedback module, such as accuracy and F_1-score, have not yet been conducted, and their performance cannot be directly compared with existing EEG-based, multimodal, or single-modality emotion recognition models. Future work will involve large-scale experiments and cross-context validation, complemented by usability testing and subjective satisfaction

XSL•FO
RenderX

assessments, to comprehensively evaluate the model's effectiveness in real-world application scenarios. The model can also be further trained and optimized using larger and more diverse datasets of older adult users to enhance recognition accuracy and generalizability. In addition, standardized performance metrics will be applied to compare multimodal fusion methods with unimodal approaches in home-based environments, thereby further validating the practical benefits of the model in emotion recognition tasks. Finally, regarding statistical analysis, post hoc pairwise comparisons were performed using the LSD method following significant ANOVA results. As this study primarily focused on examining neural activity differences elicited by different musical stimuli, the LSD method provides high sensitivity for detecting effects. However, no further strict multiple comparison corrections were applied, which may increase the risk of type I errors. Future studies will consider using more rigorous correction methods, such as Bonferroni or false discovery rate adjustments, to verify the robustness of the results.

From an application and technical perspective, this study also has certain limitations. The experimental validation was carried out using a leisure sofa as the primary prototype. Future work could integrate multiple types of sensors to collect multichannel physiological signals, thereby improving emotion recognition accuracy and response speed. Moreover, EEG signal acquisition remains susceptible to environmental noise, poor electrode contact, and participant movement. Subsequent studies should focus on improving signal processing algorithms and denoising strategies, enhancing model generalization, and conducting long-term validation in real-world smart home environments to ensure system reliability and sustained usability and user experience.

Finally, from a social and cultural perspective, this study did not deeply address issues related to cross-cultural contexts or the socio-psychological effects of such systems among older adults. Although the system design and experiments were intended to promote user well-being, future interdisciplinary research should consider cultural diversity and social inclusiveness when evaluating and optimizing emotional smart products for the older adult population.

Conclusions

The research object of this paper is explicitly for the older adult users, and it aims to solve the problems related to the emotions of older adult users in their home environment. At the same time, the mapping relationship between music parameters and emotional state is constructed. In conjunction with experimental and physiological data analysis, several aspects are considered, and it is concluded that A1 stimulus materials can soothe the emotions of older adult users. In terms of research methodology, this paper uses EEG and ERP analysis, and the application scene is a smart sofa for older adults. It is compatible with the research object and has a certain degree of practicality. At the same time, based on this, the music emotion interaction module in the multimodal HCI system of the health care sofa is set up, and it provides health care thresholds and theoretical support for the health care sofa.

Acknowledgments

The authors would like to thank the research technical and ErgoLAB Man-Machine-Environment Testing Cloud Platform (ErgoLAB 3.0) related scientific research equipment support of the Kingfar project team. This work was supported by the 2020 Jiangsu Postgraduate International Smart Health Furniture Design and Engineering project (grant number 202002). This work was also supported by the 2022 International Cooperation Joint Laboratory for Production, Education, Research, and Application of Ecological Health Care on Home Furnishing (grant number 20220602) and the Qing Lan Project (grant number 2022QL06).

Conflicts of Interest

None declared.

References

1. Schneider C, Nißen M, Kowatsch T, Vinay R. Impact of digital assistive technologies on the quality of life for people with dementia: a scoping review. BMJ Open 2024 Feb;14(2):e080545. [doi: 10.1136/bmjopen-2023-080545]
2. Tang R, Inoue Y. Services on platform ecosystems in the smart home 2.0 era: elements influencing consumers' value perception for smart home products. Sensors (Basel) 2021 Nov 6;21(21):7391. [doi: 10.3390/s21217391] [Medline: 34770697]
3. Andrés B. Towards the evolution of smart home environments: a survey. JAUSMT 2016;6(3):105-136. [doi: 10.5875/ausmt.v6i3.1039]
4. Dritsas E, Trigka M, Troussas C, Mylonas P. Multimodal interaction, interfaces, and communication: a survey. MTI 2025;9(1):6. [doi: 10.3390/mti9010006]
5. Zhong R, Ma M, Zhou Y, Lin Q, Li L, Zhang N. User acceptance of smart home voice assistant: a comparison among younger, middle-aged, and older adults. Univers Access Inf Soc 2022 Oct 30:1-18. [doi: 10.1007/s10209-022-00936-1] [Medline: 36338377]
6. Xu X, Hu X. The comparative study of emotional interaction design of empty nesters in urban and rural areas in China. HCI Int 2018. [doi: 10.1007/978-3-319-92034-4_42]

XSL•FO
RenderX

7. Zhou M, Wei C, Xie X, Wang Z, Chen L, Zhang X. Effect of music therapy on cognitive function among older adults: a systematic review and meta-analysis of randomized controlled trials. Gerontologist 2025 May 10;65(6):gnaf106. [doi: 10.1093/geront/gnaf106] [Medline: 40084508]
8. Daly I. Neural decoding of music from the EEG. Sci Rep 2023 Jan 12;13(1):624. [doi: 10.1038/s41598-022-27361-x] [Medline: 36635340]
9. Naser DS, Saha G. Influence of music liking on EEG based emotion recognition. Biomed Signal Process Control 2021 Feb;64:102251. [doi: 10.1016/j.bspc.2020.102251]
10. Raja S, Barry C, Upadhyay R, et al. Harmonious ageing: a narrative review of music therapy in the biomedical literature. Eur Geriatr Med 2025 Apr;16(2):461-472. [doi: 10.1007/s41999-024-01146-z] [Medline: 39755779]
11. Dorris JL, Neely S, Terhorst L, VonVille HM, Rodakowski J. Effects of music participation for mild cognitive impairment and dementia: a systematic review and meta-analysis. J Am Geriatr Soc 2021 Sep;69(9):2659-2667. [doi: 10.1111/jgs.17208] [Medline: 34008208]
12. Vidas D, Carrasco R, Kelly RM, et al. Everyday uses of music listening and music technologies by caregivers and people with dementia: survey and focus group study. J Med Internet Res 2024 Aug 27;26:e54186. [doi: 10.2196/54186] [Medline: 39190917]
13. Creech A, Hallam S, Varvarigou M, McQueen H, Gaunt H. Active music making: a route to enhanced subjective well-being among older people. Perspect Public Health 2013 Jan;133(1):36-43. [doi: 10.1177/1757913912466950] [Medline: 23308006]
14. Koelsch S. Brain correlates of music-evoked emotions. Nat Rev Neurosci 2014 Mar;15(3):170-180. [doi: 10.1038/nrn3666] [Medline: 24552785]
15. Daly I, Malik A, Hwang F, et al. Neural correlates of emotional responses to music: an EEG study. Neurosci Lett 2014 Jun 24;573:52-57. [doi: 10.1016/j.neulet.2014.05.003] [Medline: 24820541]
16. Tremmel C, Herff C, Sato T, Rechowicz K, Yamani Y, Krusienski DJ. Estimating cognitive workload in an interactive virtual reality environment using EEG. Front Hum Neurosci 2019;13:401. [doi: 10.3389/fnhum.2019.00401] [Medline: 31803035]
17. Gabriel D, Devos J, Chapuis S, Comte A, Billot PE. Watching live performances enhances subjective and physiological emotional responses compared to viewing the same performance on screen. IBRO Neurosci Rep 2025 Dec;19:381-390. [doi: 10.1016/j.ibneur.2025.08.002] [Medline: 40832113]
18. Yetton BD, Revord J, Margolis S, Lyubomirsky S, Seitz AR. Cognitive and physiological measures in well-being science: limitations and lessons. Front Psychol 2019;10:1630. [doi: 10.3389/fpsyg.2019.01630] [Medline: 31354601]
19. Hamzah HA, Abdalla KK. EEG-based emotion recognition systems; comprehensive study. Heliyon 2024 May;10(10):e31485. [doi: 10.1016/j.heliyon.2024.e31485]
20. Cai Z, Gao H, Wu M, Li J, Liu C. Physiologic network-based brain-heart interaction quantification during visual emotional elicitation. IEEE Trans Neural Syst Rehabil Eng 2024;32:2482-2491. [doi: 10.1109/TNSRE.2024.3424543] [Medline: 38976471]
21. Hofmann SM, Klotzsche F, Mariola A, Nikulin V, Villringer A, Gaebler M. Decoding subjective emotional arousal from EEG during an immersive virtual reality experience. Elife 2021 Oct 28;10:e64812. [doi: 10.7554/eLife.64812] [Medline: 34708689]
22. Costa T, Rognoni E, Galati D. EEG phase synchronization during emotional response to positive and negative film stimuli. Neurosci Lett 2006 Oct 9;406(3):159-164. [doi: 10.1016/j.neulet.2006.06.039] [Medline: 16942838]
23. Schindler S, Straube T. Selective visual attention to emotional pictures: Interactions of task-relevance and emotion are restricted to the late positive potential. Psychophysiology 2020 Sep;57(9):e13585. [doi: 10.1111/psyp.13585] [Medline: 32306463]
24. Tandle A, Jog N, Dharmadhikari A, Jaiswal S. Estimation of valence of emotion from musically stimulated EEG using frontal theta asymmetry. Presented at: 2016 12th International Conference on Natural Computation and 13th Fuzzy Systems and Knowledge Discovery (ICNC-FSKD); Aug 13-15, 2016; Changsha, China p. 63-68. [doi: 10.1109/FSKD.2016.7603152]
25. Tortella-Feliu M, Morillas-Romero A, Balle M, Llabrés J, Bornas X, Putman P. Spontaneous EEG activity and spontaneous emotion regulation. Int J Psychophysiol 2014 Dec;94(3):365-372. [doi: 10.1016/j.ijpsycho.2014.09.003] [Medline: 25219892]
26. Goldman A. Neuroscience in music research: critical challenges and contributions. Music Percept 2024;42(3):179-194. [doi: 10.1525/mp.2024.42.3.179]
27. Nie JC, Bai YR, Zheng Q, et al. EEG studies of the effects of music training on rhythm, music phrases and syntax perception. Presented at: Proceedings of the 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD); Feb 2, 2024; Changsha, China p. 289-295. [doi: 10.1007/978-3-031-51455-5_32]
28. Wang Y, Siu TSC, Cheung H. Effect of music emotion on mu and beta oscillations. Psychol Music 2023 Sep;51(5):1489-1500. [doi: 10.1177/03057356221145960]
29. Exploring emotion expression recognition in older adults interacting with a virtual coach. IEEE Trans Affect Comput 2025;16(3):2303-2320. [doi: 10.1109/TAFFC.2025.1234567]
30. Koelsch S, Skouras S, Fritz T, et al. The roles of superficial amygdala and auditory cortex in music-evoked fear and joy. Neuroimage 2013 Nov 1;81:49-60. [doi: 10.1016/j.neuroimage.2013.05.008] [Medline: 23684870]

XSL•FO
RenderX

31. Salimpoor VN, Benovoy M, Larcher K, Dagher A, Zatorre RJ. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. Nat Neurosci 2011 Feb;14(2):257-262. [doi: 10.1038/nn.2726] [Medline: 21217764]
32. Gagnon L, Peretz I. Mode and tempo relative contributions to “happy-sad” judgements in equitone melodies. Cogn Emot 2003 Jan;17(1):25-40. [doi: 10.1080/02699930302279] [Medline: 29715736]
33. Lehne M, Rohrmeier M, Koelsch S. Tension-related activity in the orbitofrontal cortex and amygdala: an fMRI study with music. Soc Cogn Affect Neurosci 2014 Oct;9(10):1515-1523. [doi: 10.1093/scan/nst141] [Medline: 23974947]
34. Healey J, Picard R. Digital processing of affective signals. Presented at: 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing; May 15, 1998; Seattle, WA, USA p. 3749-3752. [doi: 10.1109/ICASSP.1998.679699]
35. Yu YC, You SD, Tsai DR. Magic mirror table for social-emotion alleviation in the smart home. IEEE Trans Consumer Electron 2012;58(1):126-131. [doi: 10.1109/TCE.2012.6170064]
36. Nelson BW, Allen NB. Extending the passive-sensing toolbox: using smart-home technology in psychological science. Perspect Psychol Sci 2018 Nov;13(6):718-733. [doi: 10.1177/1745691618776008] [Medline: 30217132]
37. Rocha AP, Ketsmur M, Almeida N, Teixeira A. An accessible smart home based on integrated multimodal interaction. Sensors (Basel) 2021 Aug 13;21(16):5464. [doi: 10.3390/s21165464] [Medline: 34450903]
38. Contreras-Castañeda MA, Holgado-Terriza JA, Pomboza-Junez G, Paderewski-Rodríguez P, Gutiérrez-Vela FL. Smart home: multimodal interaction for control of home devices. Presented at: Proc XX Int Conf Human-Computer Interact; Jun 25, 2019. [doi: 10.1145/3335595.3335636]
39. Spinsante S, Cippitelli E, Santis A, et al. Multimodal interaction in an elderly-friendly smart home: a case study. In: Agüero R, Zinner T, Goleva R, Timm-Giel A, Tran-Gia P, editors. Mobile Networks and Management: Springer International Publishing; 2015:373-386. [doi: 10.1007/978-3-319-16292-8_27]
40. Hargreaves T, Wilson C. Smart Homes and Their Users: Springer International Publishing; 2017. [doi: 10.1007/978-3-319-68018-7]
41. Almeida N, Teixeira A, Silva S, Ketsmur M. The AM4I architecture and framework for multimodal interaction and its application to smart environments. Sensors (Basel) 2019 Jun 6;19(11):2587. [doi: 10.3390/s19112587] [Medline: 31174370]
42. Chahuara P, Portet F, Vacher M. Context-aware decision making under uncertainty for voice-based control of smart home. Expert Syst Appl 2017 Jun;75:63-79. [doi: 10.1016/j.eswa.2017.01.014]
43. Gullà F, Ceccacci S, Menghi R, et al. Adaptive interface for smart home: a new design approach. In: Cavallo F, Marletta V, Monteriù A, Siciliano P, editors. Ambient Assisted Living: Springer International Publishing; 2017:107-115. [doi: 10.1007/978-3-319-54283-6_8]
44. Marques B, Dias P, Alves J, et al. Adaptive augmented reality user interfaces using face recognition for smart home control. In: Ahram T, Karwowski W, Pickl S, Taiar R, editors. Human Systems Engineering and Design II: Springer International Publishing; 2020:15-19. [doi: 10.1007/978-3-030-27928-8_3]
45. Wilhelm M, Roscher D, Blumendorf M, Albayrak S. A trajectory-based approach for device independent gesture recognition in multimodal user interfaces. Presented at: Haptic and Audio Interaction Design: 5th International Workshop; Sep 15, 2010; Berlin, Heidelberg p. 197-206. [doi: 10.1007/978-3-642-15841-4_21]
46. Wei T, Shi Y, Sun J. A multimodal interaction-based approach to situational driving anger emotion regulation. Expert Syst Appl 2025 May;274:127051. [doi: 10.1016/j.eswa.2025.127051]
47. Caridakis G, Karpouzis K, Kollias S. User and context adaptive neural networks for emotion recognition. Neurocomputing 2008 Aug;71(13-15):2553-2562. [doi: 10.1016/j.neucom.2007.11.043]
48. Eerola T, Ferrer R, Alluri V. Timbre and affect dimensions: evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds. Music Percept 2012 Sep 1;30(1):49-70. [doi: 10.1525/mp.2012.30.1.49]
49. Gomez P, Danuser B. Relationships between musical structure and psychophysiological measures of emotion. Emotion 2007 May;7(2):377-387. [doi: 10.1037/1528-3542.7.2.377] [Medline: 17516815]
50. Putkinen V, Zhou X, Gan X, et al. Bodily maps of musical sensations across cultures. Proc Natl Acad Sci USA 2024 Jan 30;121(5):e2308859121. [doi: 10.1073/pnas.2308859121]
51. Morgan SD. Comparing emotion recognition and word recognition in background noise. J Speech Lang Hear Res 2021 May 11;64(5):1758-1772. [doi: 10.1044/2021_JSLHR-20-00153]
52. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
53. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
54. Heng GJ, Quek C, Chen SH, et al. Effects of familiarity with musical style on music-evoked emotions: an EEG study. Front Neurosci 2024;18:1123. [doi: 10.3389/fnins.2024.1123]
55. Sarasso P, Perna P, Barbieri P, Neppi-Modona M, Sacco K, Ronga I. Memorisation and implicit perceptual learning are enhanced for preferred musical intervals and chords. Psychon Bull Rev 2021 Oct;28(5):1623-1637. [doi: 10.3758/s13423-021-01922-z] [Medline: 33945127]
56. Sato H, Morimoto M, Ota R. Acceptable range of speech level in noisy sound fields for young adults and elderly persons. J Acoust Soc Am 2011 Sep 1;130(3):1411-1419. [doi: 10.1121/1.3609122]

XSL•FO
RenderX

57. Awada M, Becerik-Gerber B, Lucas G, Roll S. Cognitive performance, creativity and stress levels of neurotypical young adults under different white noise levels. Sci Rep 2022;12(1):14566. [doi: 10.1038/s41598-022-18862-w]
58. Prodhan RA, Akter S, Pias TS, Adnan MA. Optimal EEG electrode set for emotion recognition from brain signals: an empirical quest. arXiv. Preprint posted online on 2023. [doi: 10.1201/9781003371540-5]
59. Burkhard A, Elmer S, Kara D, Brauchli C, Jäncke L. The effect of background music on inhibitory functions: an ERP study. Front Hum Neurosci 2018;12:293. [doi: 10.3389/fnhum.2018.00293] [Medline: 30083099]
60. Lin YP, Wang CH, Jung TP, et al. EEG-based emotion recognition in music listening. IEEE Trans Biomed Eng 2010 Jul;57(7):1798-1806. [doi: 10.1109/TBME.2010.2048568] [Medline: 20442037]
61. Yu Y, Oh Y, Kounios J, Beeman M. Dynamics of hidden brain states when people solve verbal puzzles. Neuroimage 2022 Jul 15;255:119202. [doi: 10.1016/j.neuroimage.2022.119202] [Medline: 35427772]
62. Fang Q, Fang C, Li L, Song Y. Impact of sport training on adaptations in neural functioning and behavioral performance: a scoping review with meta-analysis on EEG research. J Exerc Sci Fit 2022 Jul;20(3):206-215. [doi: 10.1016/j.jesf.2022.04.001]
63. Samaha J, LaRocque JJ, Postle BR. Spontaneous alpha-band amplitude predicts subjective visibility but not discrimination accuracy during high-level perception. Conscious Cogn 2022 Jul;102:103337. [doi: 10.1016/j.concog.2022.103337] [Medline: 35525224]
64. Wang P, Wang M, Zhou Y, Xu Z, Zhang D. Multiband decomposition and spectral discriminative analysis for motor imagery BCI via deep neural network. Front Comput Sci 2022 Oct;16(5):1-13. [doi: 10.1007/s11704-021-0587-2]
65. Sánchez-Moguel SM, Baravalle R, González-Salinas S, Rosso OA, Fernández T, Montani F. Abnormal EEG signal energy in the elderly: a wavelet analysis of event-related potentials during a stroop task. J Neurosci Methods 2022 Jul 1;376:109608. [doi: 10.1016/j.jneumeth.2022.109608] [Medline: 35487316]
66. Sun L, Zhang Y, Li Z, Zhou Z, Zhou Z. inML Kit: empowering the prototyping of ML-enhanced products by involving designers in the ML lifecycle. AIEDAM 2022;36:e8. [doi: 10.1017/S0890060421000391]
67. Qi G, Jing Z. The quantitative research on length and area perception: a guidance on shape encoding in visual interface. Displays 2022 Dec;75:102325. [doi: 10.1016/j.displa.2022.102325]
68. Namayala P, Kondo T, Manyilizu M, Hassan K. Practical applicability of user experience capability/maturity models in the development processes of free and open-source software: a systematic literature review. Behav Inform Technol 2023 Apr 4;42(5):596-623. [doi: 10.1080/0144929X.2022.2032348]
69. Xiong X, Yue X, Wu Z. Current status and development trends of chinese intelligent furniture industry. J Renew Mater 2023;11(3):1353-1366. [doi: 10.32604/jrm.2022.023447]
70. Yu N, Ouyang Z, Wang H. Study on smart home interface design characteristics considering the influence of age difference: focusing on sliders. Front Psychol 2022;13:828545. [doi: 10.3389/fpsyg.2022.828545]
71. Zhao D, Hu X, Xiong S, et al. k-means clustering and kNN classification based on negative databases. Appl Soft Comput 2021 Oct;110:107732. [doi: 10.1016/j.asoc.2021.107732]
72. Tan Y, Geng S, Katsumata S, Xiong X. The effects of ad heuristic and systematic cues on consumer brand awareness and purchase intention: investigating the bias effect of heuristic information processing. J Retail Consum Serv 2021 Nov;63:102696. [doi: 10.1016/j.jretconser.2021.102696]
73. Chao Y, Liu T, Shen LM. Method of recognizing sleep postures based on air pressure sensor and convolutional neural network: for an air spring mattress. Eng Appl Artif Intell 2023 May;121:106009. [doi: 10.1016/j.engappai.2023.106009]
74. Gnanasivam P, T. Bharathy G, Rajendran V, Tamilselvi T. Efficient centralized cooperative spectrum sensing techniques for cognitive networks. Comput Syst Sci Eng 2023;44(1):55-65. [doi: 10.32604/csse.2023.023374]
75. Sun X, Jin W. Elderly's preferences towards rehabilitation robot appearance using electroencephalogram signal. EURASIP J Adv Signal Process 2021 Dec;2021(1). [doi: 10.1186/s13634-021-00757-z]
76. Xia W, Yan L, Yongjian S, et al. Investigation of MAS structure and intelligent + information processing mechanism of hypersonic target detection and recognition system. J Syst Eng Electron 2020;31(6):1105-1115. [doi: 10.23919/JSEE.2020.000083]
77. Al Zaidawi SMK, Prinzler MHU, Lührs J, Maneth S. An extensive study of user identification via eye movements across multiple datasets. Signal Process Image Commun 2022 Oct;108:116804. [doi: 10.1016/j.image.2022.116804]
78. Yang Z, Zhang L, Wu Z. Research on performance evaluation of urban furniture function design based on internet of things digitization. IEEE Access 2022;10:72895-72906. [doi: 10.1109/ACCESS.2022.3188274]
79. Daliri MR, Taghizadeh M, Niksirat KS. EEG signature of object categorization from event-related potentials. J Med Signals Sens 2013 Jan;3(1):37-44. [doi: 10.4103/2228-7477.114318] [Medline: 24083136]
80. Zhang D, He W, Wang T, et al. Three stages of emotional word processing: an ERP study with rapid serial visual presentation. Soc Cogn Affect Neurosci 2014 Dec;9(12):1897-1903. [doi: 10.1093/scan/nst188] [Medline: 24526185]
81. Olofsson JK, Nordin S, Sequeira H, Polich J. Affective picture processing: an integrative review of ERP findings. Biol Psychol 2008 Mar;77(3):247-265. [doi: 10.1016/j.biopsycho.2007.11.006]
82. Malekmohammadi A, Ehrlich SK, Rauschecker JP, Cheng G. Listening to familiar music induces continuous inhibition of alpha and low-beta power. J Neurophysiol 2023 Jun 1;129(6):1344-1358. [doi: 10.1152/jn.00269.2022] [Medline: 37141051]

XSL•FO
RenderX

83. Poikonen H, Alluri V, Brattico E, Lartillot O, Tervaniemi M, Huotilainen M. Event-related brain responses while listening to entire pieces of music. Neuroscience 2016 Jan;312:58-73. [doi: 10.1016/j.neuroscience.2015.10.061]
84. Alday PM, Schlesewsky M, Bornkessel-Schlesewsky I. Electrophysiology reveals the neural dynamics of naturalistic auditory language processing: event-related potentials reflect continuous model updates. eNeuro 2017;4(6). [doi: 10.1523/ENEURO.0311-16.2017] [Medline: 29379867]
85. Ritter SM, Ferguson S. Happy creativity: listening to happy music facilitates divergent thinking. PLoS ONE 2017;12(9):e0182210. [doi: 10.1371/journal.pone.0182210]
86. Qiao Y, Mu J, Xie J, Hu B, Liu G. Music emotion recognition based on temporal convolutional attention network using EEG. Front Hum Neurosci 2024;18:1324897. [doi: 10.3389/fnhum.2024.1324897] [Medline: 38617132]
87. Ferreira-Santos F, Silveira C, Almeida PR, Palha A, Barbosa F, Marques-Teixeira J. The auditory P200 is both increased and reduced in schizophrenia? a meta-analytic dissociation of the effect for standard and target stimuli in the oddball task. Clin Neurophysiol 2012 Jul;123(7):1300-1308. [doi: 10.1016/j.clinph.2011.11.036] [Medline: 22197447]
88. Agrawal D, Timm L, Viola FC, et al. ERP evidence for the recognition of emotional prosody through simulated cochlear implant strategies. BMC Neurosci 2012 Sep 20;13:113. [doi: 10.1186/1471-2202-13-113] [Medline: 22994867]
89. Kutas M, Federmeier KD. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu Rev Psychol 2011;62:621-647. [doi: 10.1146/annurev.psych.093008.131123] [Medline: 20809790]
90. Kučikienė D, Praninskienė R. The impact of music on the bioelectrical oscillations of the brain. Acta Med Litu 2018;25(2):101-106. [doi: 10.6001/actamedica.v25i2.3763] [Medline: 30210244]
91. Yang Z, Su Q, Xie J, et al. Music tempo modulates emotional states as revealed through EEG insights. Sci Rep 2025;15(1):8276. [doi: 10.1038/s41598-025-92679-1]
92. Flores-Gutiérrez EO, Díaz JL, Barrios FA, et al. Metabolic and electric brain patterns during pleasant and unpleasant emotions induced by music masterpieces. Int J Psychophysiol 2007 Jul;65(1):69-84. [doi: 10.1016/j.ijpsycho.2007.03.004] [Medline: 17466401]
93. Zhou Y, Lian J. Identification of emotions evoked by music via spatial-temporal transformer in multi-channel EEG signals. Front Neurosci 2023;17:1188696. [doi: 10.3389/fnins.2023.1188696]
94. Patel SH, Azzam PN. Characterization of N200 and P300: selected studies of the event-related potential. Int J Med Sci 2005;2(4):147-154. [doi: 10.7150/ijms.2.147] [Medline: 16239953]
95. Portnova GV, Proskurnina EV. Acoustic and subjective basis of emotional perception in comatose patients: a comparative study. Clin Psychopharmacol Neurosci 2023 Nov 30;21(4):701-714. [doi: 10.9758/cpn.22.1028] [Medline: 37859443]
96. Öz D, Özbek Y, Kıyı İ, et al. Cognitive evidence on EEG-P300 in healthy individuals with high depression scores. Neuropsychiatr Dis Treat 2022;18:1-9. [doi: 10.2147/NDT.S39020]
97. Pawlowski GM, Ghosh-Hajra S, Fickling SD, et al. Brain vital signs: expanding from the auditory to visual modality. Front Neurosci 2018;12:968. [doi: 10.3389/fnins.2018.00968] [Medline: 30713487]
98. Calma-Roddin N, Drury JE. Music, language, and the N400: ERP interference patterns across cognitive domains. Sci Rep 2020 Jul 8;10(1):11222. [doi: 10.1038/s41598-020-66732-0] [Medline: 32641708]
99. Lee LN, Kim MJ. A critical review of smart residential environments for older adults with a focus on pleasurable experience. Front Psychol 2019;10:3080. [doi: 10.3389/fpsyg.2019.03080] [Medline: 32038424]
100. Kamal AM, Azad M, Chowdhury SJ. Prototype smart home environment with biofeedback. arXiv. Preprint posted online on Jul 28, 2024. [doi: 10.48550/arXiv.2407.19525]
101. Moors A. Theories of emotion causation: a review. Cogn Emot 2009 Jun;23(4):625-662. [doi: 10.1080/02699930802645739]
102. Scherer KR. On the rationality of emotions: or, when are emotions rational? Soc Sci Inf 2011 Sep;50(3-4):330-350. [doi: 10.1177/0539018411411018]
103. Wang CM, Lin CH. Design of an interactive system by combining affective computing technology with music for stress relief. Electronics (Basel) 2025;14(15):3087. [doi: 10.3390/electronics14153087]
104. Desmet PMA, Hekkert P. Framework of product experience. Int J Des 2007;1(1):57-66. [doi: 10.1016/B978-008045089-6.50003-4]
105. Li MG, Olsen KN, Thompson WF. Cross-cultural biases of emotion perception in music. Brain Sci 2025 Apr 29;15(5):477. [doi: 10.3390/brainsci15050477] [Medline: 40426648]
106. Wang H, Chen H, Wang Y. Analysis of hot topics regarding global smart elderly care research - 1997-2021. China CDC Wkly 2024 Mar 1;6(9):157-161. [doi: 10.46234/ccdcw2024.032] [Medline: 38495592]
107. Lee H, Çelen E, Harrison P, et al. GlobalMood: a cross-cultural benchmark for music emotion recognition. arXiv. Preprint posted online on 2025. [doi: 10.48550/arXiv.2505.09539]
108. Paton C, Kushniruk AW, Borycki EM, English M, Warren J. Improving the usability and safety of digital health systems: the role of predictive human-computer interaction modeling. J Med Internet Res 2021 May 27;23(5):e25281. [doi: 10.2196/25281] [Medline: 34042590]
109. Manandhar S, Kafle K, Andow B, Singh K, Nadkarni A. Smart home privacy policies demystified: a study of availability, content, and coverage. Presented at: 31st USENIX Security Symposium (USENIX Security 22); Aug 10-12, 2022; Boston, MA, USA p. 3521-3538 URL: https://www.usenix.org/conference/usenixsecurity22/presentation/manandhar [accessed 2026-02-05]

XSL•FO
RenderX

110. Chong HJ, Kim HJ, Kim B. Scoping review on the use of music for emotion regulation. Behav Sci (Basel) 2024 Sep 9;14(9):793. [doi: 10.3390/bs14090793] [Medline: 39336008]
111. Ho MT, Vuong QH. Five premises to understand human–computer interactions as AI is changing the world. AI Soc 2025 Feb;40(2):1161-1162. [doi: 10.1007/s00146-024-01913-3]
112. Yang L, Shen Z, Zeng J, Luo X, Lin H. COSMIC: music emotion recognition combining structure analysis and modal interaction. Multimed Tools Appl 2023;83(5):12519-12534. [doi: 10.1007/s11042-023-15376-z]
113. Takashima N, Li F, Grzegorzek M, Shirahama K. Embedding-based music emotion recognition using composite loss. IEEE Access 2023;11:36579-36604. [doi: 10.1109/ACCESS.2023.3265807]
114. Wang H, Zhong W, Ma L, Ye L, Zhang Q. Emotional quality evaluation for generated music based on emotion recognition model. Presented at: 2022 IEEE International Conference on Multimedia and Expo Workshops (ICMEW); Jul 18-22, 2022; Taipei City, Taiwan p. 1-6. [doi: 10.1109/ICMEW56448.2022.9859459]
115. Sutcliffe R, Rendell PG, Henry JD, Bailey PE, Ruffman T. Music to my ears: age-related decline in musical and facial emotion recognition. Psychol Aging 2017 Dec;32(8):698-709. [doi: 10.1037/pag0000203] [Medline: 29239655]
116. Vieillard S, Gilet AL. Age-related differences in affective responses to and memory for emotions conveyed by music: a cross-sectional study. Front Psychol 2013;4:711. [doi: 10.3389/fpsyg.2013.00711] [Medline: 24137141]
117. Huang Y, Yang J, Liao P, Pan J. Fusion of facial expressions and EEG for multimodal emotion recognition. Comput Intell Neurosci 2017;2017:2107451. [doi: 10.1155/2017/2107451] [Medline: 29056963]
118. Chimamiwa G, Giaretta A, Alirezaie M, Pecora F, Loutfi A. Are smart homes adequate for older adults with dementia? Sensors (Basel) 2022 Jun 2;22(11):4254. [doi: 10.3390/s22114254] [Medline: 35684874]

Abbreviations

CNN: convolutional neural network
EEG: electroencephalogram
ERP: event-related potential
HCI: human-computer interaction
LSD: least significant difference
MF-AWI: Multimodal Fusion-Adaptive Wellness Interaction
VA: Valence-Arousal

Edited by C Lee; submitted 09.May.2025; peer-reviewed by A Schwerdtfeger, E Yin, HA Aboalsamh, MT Ho; revised version received 13.Dec.2025; accepted 06.Jan.2026; published 10.Feb.2026.

Please cite as:
Zhou C, Luo Y, Zhan W, Huang T, Kaner J
A Music Interaction System Based on the Emotional Recognition of Older Adults in Smart Homes: Cross-Sectional Experimental Study
JMIR Aging 2026;9:e77218
URL: https://aging.jmir.org/2026/1/e77218
doi:10.2196/77218

© Chengmin Zhou, Yuxuan Luo, Wenjing Zhan, Ting Huang, Jake Kaner. Originally published in JMIR Aging (https://aging.jmir.org), 10.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Community-in-the-Loop Approach to Smart Home Monitoring for Aging in Place: Mixed Methods Evaluation of a Co-Designed Prototype

Roschelle L Fritz[1], RN, PhD; Connie Kim Yen Nguyen-Truong[2], RN, PhD; Jennifer Phipps[3], BS, PhD; Ellen E Hinderlie[4], RN, MSN; Yolanda LB Rodriguez[5], ARNP, FNP-BC, DNP; Thai Hien Nguyen[6], RN, BSN; Gabrielle Barling[7], BS; Anna Mishuk[4]; Heather Schoonover[8], MN, ARNP-CNS, PHCNS-BC; Christi Zuber[9], RN, PhD; Marilyn Rantz[10], RN, PhD

[1]Faculty School of Nursing, UC Davis Health System, 2570 48th St., Sacramento, CA, United States

[10]Sinclair School of Nursing, University of Missouri, Columbia, MO, United States

[2]Department of Nursing and Systems Science, Faculty College of Nursing, Washington State University Vancouver, Vancouver, WA, United States

[3]School of Nursing, University of California Davis Health, Sacramento, CA, United States

[4]College of Nursing, Washington State University Vancouver, Vancouver, WA, United States

[5]Advanced Practice & Community-Based Care Department, Faculty College of Nursing, Washington State University Vancouver, Vancouver, WA, United States

[6]Community Health Worker, Vancouver, WA, United States

[7]Amazon (United States), San Diego, CA, United States

[8]Health Catalyst, South Jordan, UT, United States

[9]Aspen Labs, Denver, CO, United States

Corresponding Author:
Roschelle L Fritz, RN, PhD
Faculty School of Nursing, UC Davis Health System, 2570 48th St., Sacramento, CA, United States

Abstract

Background: The population of adults aged 65 and older is rapidly increasing, while the availability of caregivers is declining. Smart homes that provide unobtrusive, continuous monitoring and alerting on clinically relevant changes in daily activity patterns offer a potentially innovative solution for aging in place.

Objective: This study aims to evaluate the barriers and facilitators to the adoption of a low-cost smart home embedded within a community-based approach to health monitoring for older adults with multiple chronic conditions and who are experiencing poverty.

Methods: Using a prospective, mixed methods design and iterative community co-design, 46 older adults from 7 different language groups were continuously monitored for 6 months with ambient sensors installed in their homes. Two older adults were monitored for 4 and 5 months, respectively, resulting in a total sample of 48. The system generated alerts based on movement pattern changes and escalated notifications to participants, support persons, community health workers, and nurses. Sensor data were analyzed descriptively to quantify alert patterns and response rates, while written text-based data from in-the-moment surveys, community health workers' and registered nurses' notes, and semistructured interviews underwent qualitative descriptive analysis and reflexive thematic coding.

Results: The system generated 37 million sensor readings condensed into 1.2 million high-level events and 4719 novel alerts. Qualitative data comprised 34,086 words of text. Participants responded to 1.57% (74) of the initial email alerts and 7.79% (368) of the follow-up SMS text message alerts sent when no email response was received. Community health workers and registered nurses responded to 78.36% (n=3698) of the escalated alerts, resulting in 1060 contacts with participants in response to alerts. Clinical contacts resulted in 72 interventions. Three major qualitative themes emerged: (1) Alone, (2) Trust, and (3) Human Connection. Subthemes included Safety, Personalization, and Digital Distress defined as stress associated with interacting with digital health-monitoring systems. Participants rated the system highly (mean likelihood-to-recommend rating 8.68/10, SD 1.68); however, they expressed a strong preference for phone calls over automated alerts. Cultural expectations influenced adoption, particularly in multigenerational households.

Conclusions: Communities can effectively engage in technology-delivered health care. Future research is needed to improve technical aspects of smart home monitoring systems, including accurate alerting using machine learning, data visualizations for older adults and health care workers, and culturally sensitive features. Additional work should address how and when to

XSL•FO
RenderX

communicate automated messaging, engage older adults with their own data, and integrate sensor-based monitoring into health care workflows. Research should also explore personalization through advanced computational approaches such as machine learning and strategies to reduce digital distress.

(JMIR Aging 2026;9:e88290) doi:10.2196/88290

KEYWORDS

smart homes; ambient sensors; health equity; remote patient monitoring; aging in place; community-based participatory research; co-design; digital distress

Introduction

Background

The global population is aging rapidly, with an estimated 1.6 billion adults aged 65 and over by 2050 [1]. This demographic shift is accompanied by a shrinking caregiver-to-older-adult ratio: in the United States, the ratio for adults aged 80 and older is projected to decline from 7:1 in 2010 to 3:1 by 2050, with fewer caregivers available for the growing aging population [2,3]. This imbalance presents the urgent challenge of how to support older adults in aging safely and independently at home.

Health monitoring smart homes (ie, sensor-based health monitoring at home) offer a promising solution [4,5]. These systems typically consist of ambient sensors (eg, motion, door, humidity, and more) connected to a central hub that transmits data to a secure platform. Algorithms analyze movement patterns to detect anomalies such as prolonged inactivity, disrupted sleep, or changes in kitchen or bathroom use, all potential indicators of emerging health concerns. Alerts can be communicated via email, SMS text message, or phone calls, and escalated to caregivers or clinical teams for follow-up.

However, most smart home health monitoring has been developed and tested in higher-income settings, leaving critical gaps in understanding their feasibility and acceptability among older adults experiencing poverty [6,7]. Older adults in subsidized housing or with limited income face distinctive barriers to technology adoption, including limited internet access, low digital literacy, mistrust of systems, and cultural expectations around caregiving [8-12]. Without deliberate attention to these barriers, smart home innovations may risk widening health disparities rather than closing them [13,14].

This study aims to identify and understand barriers to adopting smart home health monitoring occurring within a co-designed community support system among older adults experiencing poverty. To address these aims, a low-cost, community-in-the-loop “smart health system” prototype was developed, featuring continuous ambient health monitoring capabilities. Using mixed methods, we evaluated how older adults interacted with the system, the challenges they encountered, and the features that supported trust, engagement, and equitable access. Here, the term *smart home* refers broadly to ambient sensing in the home for the purpose of monitoring residents’ health. The prototype deployed for this study is referred to throughout as a *smart health system* because the smart home was embedded in a *community-in-the-loop* support system where older adults received a smart home along with systematic support from self-identified family and friends (ie, a support person) and from appointed community health workers (CHWs) and registered nurses (RNs).

This paper shares the findings from trialing the community-in-the-loop approach to health monitoring in resource-constrained settings that use smart homes and discusses the potential for smart home adoption by older adults living in poverty. The purpose of this paper is to provide insights into how, and if, continuous ambient health monitoring is possible in older adult populations experiencing poverty. The low-cost system is discussed in detail in Multimedia Appendix 1.

Smart Home Technologies for Health Monitoring

Smart home systems leverage ambient sensors, Internet of Things (IoT) devices, and data analytics to monitor daily activities and detect health-related changes [14-16]. Prior studies demonstrate their potential to support aging in place by identifying deviations in mobility, sleep, and self-care routines. Rantz et al [17,18] showed that sensor-based systems can detect early signs of health decline. Pioneering work by Cook and the team at the Center for Advanced Studies in Adaptive Systems (CASAS) developed lightweight, easy-to-install systems for activity recognition and health monitoring, including nurse-in-the-loop models and pain-related behavior detection [5,19-22]. More recent work has advanced IoT-based frameworks for secure remote monitoring [23] and patient-centric agents using blockchain on 5G networks [24]. Furthermore, work has been done that specifically identifies gaps in chronic disease applications [4].

Smart home systems show promise in detecting early signs of clinically relevant conditions such as cardiovascular disease, congestive heart failure, cancer, urinary tract infections, sleep disturbances, and cognitive decline [4,21,25]. These conditions disproportionately affect older adults experiencing poverty who often face delayed diagnoses and limited access to care [26-29]. Sensor-based alerts, such as changes in mobility, sleep patterns, or hygiene routines, can serve as proxies for emerging health issues and prompt timely interventions. These studies confirm the feasibility of unobtrusive monitoring but often remain focused on higher-income populations.

Barriers and Facilitators to Adoption in Low-Income Populations

While smart home technologies are well established, their adoption among diverse older adults remains understudied. Prior work has primarily focused on for-profit senior living settings, including Tiger Place [17,18], Touchmark [21,30,31], and early collaborative aging research using technology (CART)–related studies [32] where adoption has not faced many barriers. In

XSL•FO
RenderX

contrast, adoption barriers among low-income older adults persist, including privacy concerns, perceived intrusiveness, system complexity, digital literacy challenges, limited internet access, and inadequate technical support [4,7,33-35]. Similar inequities appear across digital health interventions, such as mobile health and telehealth [36], for chronic disease management [37-39].

Key facilitators for older adults and the health care team include ease of use, perceived usefulness, and a system's ability to enhance safety and independence [4,9,40-42]. Trust in the technology and the organizations deploying it strongly shapes adoption, particularly amid concerns about data security and reliability [4,9,43]. This study extends prior work by examining barriers and facilitators to adoption among older adults with multiple chronic conditions living in poverty, with particular attention to affordability, usability, cultural relevance, and community co-design as key determinants of trust and engagement.

Community Engagement and Co-Design

Co-design and community-based participatory research methodologies have proven effective in tailoring digital health interventions to the needs of diverse populations [44,45]. Involving end users in design and testing enhances trust, usability, and perceived ownership of technology [46,47]. However, few studies have applied these principles to smart home health monitoring systems serving low-income older adults with multiple chronic conditions.

This study embedded community engagement across all phases of system development and evaluation [48]. We brought together low-income older adult residents in subsidized housing and living in the community, subsidized housing managers, CHWs from diverse cultural and ethnic communities, and nursing and engineering researchers to co-design features including notification preferences, sensor placement, and privacy safeguards. Adapted from the "nurse-in-the-loop" model, this *community-in-the-loop approach* ensured alignment with participants' cultural, social, and socioeconomic contexts. Iterative feedback revealed preferences for low-burden devices, clear feedback loops, and direct human contact during alerts, showing how co-design can make smart home health systems more inclusive, acceptable, and scalable in resource-constrained settings.

Methods

Ethical Considerations

This study was approved by the Washington State University Institutional Review Board (IRB) for human participants' research under IRB#19135. All participants provided informed consent in their preferred language. Seven languages were represented in the participant pool. Each participant received a US $50 gift card and a results summary after study completion.

Study Design and Setting

A community-engaged, prospective, mixed methods design was used to prototype and evaluate a low-cost smart health monitoring system for older adults managing multiple chronic conditions while experiencing poverty. The study was conducted in collaboration with local housing authorities, culturally based health organizations, and a free community health clinic in the US Pacific Northwest.

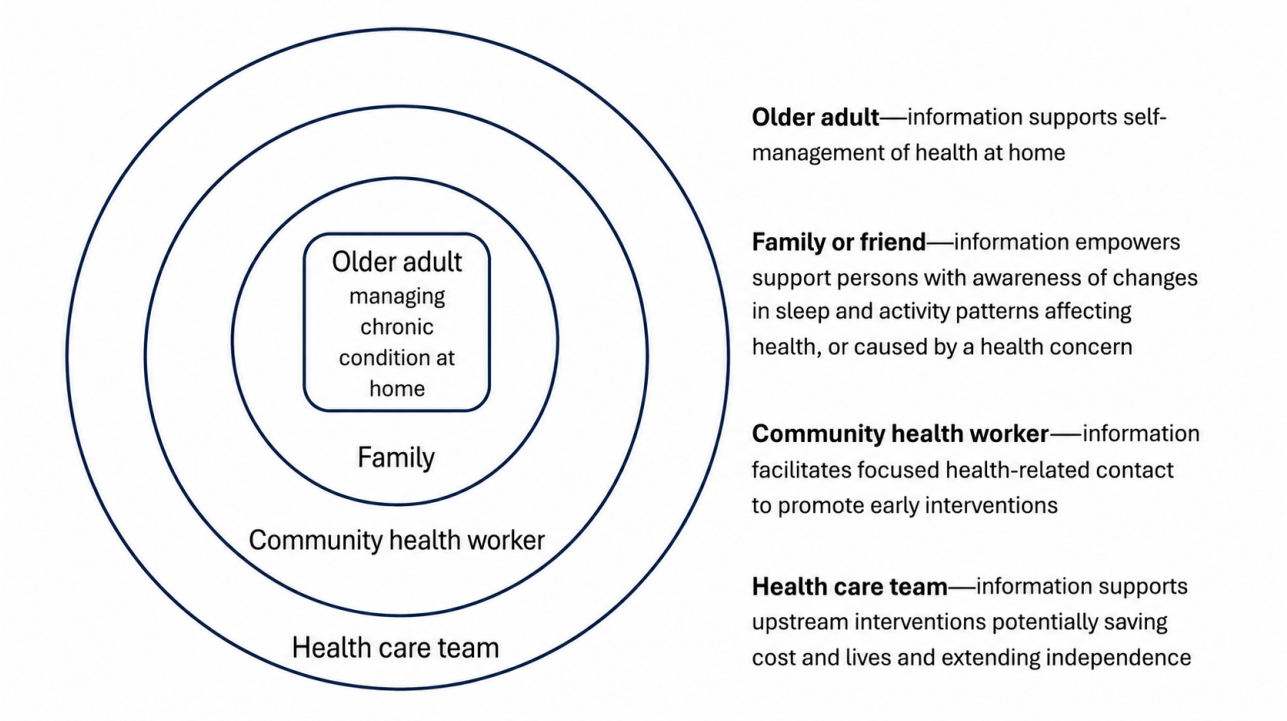
Smart Health System Prototype

The prototype system was designed as a closed-loop model to help families, communities, and health care teams support older adults aging in the place of their choosing (Figure 1). The system operated through three integrated components: (1) ambient sensing, (2) a rules-based algorithm developed for this study that continuously processed sensor data and automatically generated alerts for clinically relevant changes in daily behavior, and (3) community members who worked alongside the technology to respond promptly to alerts, including facilitating engagement with the health care team.

The prototype system included 6 alert types, with alert messaging designed not to imply judgment or provoke stress or fear. The system alerted for social isolation (Homebody), decreased activities (Daytime Resting), kitchen use (Eating Routines), decreased hygiene (Bathing), sleep disturbance (Restless Night), and sleep patterns (Changing Sleep Routine). The participant-facing alert terms associated with clinically relevant behavior changes are italicized. Multimedia Appendix 1 contains detailed descriptions of alerts, including defined thresholds.

XSL•FO
RenderX

Figure 1. The connected smart health system.



Study Sample and Recruiting

A convenience sample was recruited with the assistance of subsidized housing building managers. More than 80 residents across 7 subsidized housing properties attended information presentations about the smart health system, after which interested individuals met privately with a trained research assistant or the lead primary investigator (PI) to review consent forms in their preferred language. Residents were informed that the system was a research prototype, that it was not intended for emergency detection or intervention, and that the absence of alerts would not imply health or clinical stability. They were instructed to use usual health care and emergency services. Adults aged 50 years and above with 2 or more chronic conditions and income at or below the federal poverty level (US $14,580 annually for a single-person household) [49] were included. The lower age threshold was informed by community partner input, which emphasized that chronic conditions often develop earlier among racially and ethnically diverse populations. Older adults living alone were prioritized for recruitment, although homes with fewer than 2 residents met inclusion criteria. A target was set for recruiting >80% single-resident homes. Exclusion criteria included dementia, severe mental illness, or pets that could interfere with sensor readings.

Data Collection

Data collection combined continuous, ambient, home-based sensing with participant-facing and team-documented narratives, and in-the-moment surveys associated with health alerts. Table 1 contains the data elements of the study with descriptions. Alerts were triggered when anomalies exceeded preset thresholds (eg, no door use for 48 hours, reduced kitchen activity, sleep disruption). To minimize alert fatigue, some overlapping alerts were combined into a single alert, yielding 4719 novel alerts. For example, a Restless Night alert and a Changing Sleep Routine alert on the same night would be combined into one alert requiring one survey response. Because it was not possible to have continuous daily CHW coverage for the study duration, the full-time research RN or lead investigator (an RN) assisted in responding to alerts, and alert notifications were delivered only during business hours. Business hours were 9:00 AM to 6:00 PM Pacific Time, Monday to Friday. Nighttime sensing resulting in alerts saw the alerts delivered at 9:00 AM on weekdays, and weekend alerts were delivered at 9:00 AM on Mondays. No alerts were discarded.

XSL•FO
RenderX

Table . Data elements with descriptions.

Data element	Description
Sensor data from 6 multisensors per home	The sensor state (ON) when motion was sensed, or when temperature, humidity, and lights changed.
Alerts	Emails and SMS text messages generated by the system and sent to the participant, a support person, and a CHW[a] (sequentially escalating) when a change in daily routine met algorithm rules aligned with possible health changes.
Alert responses	Discrete and unstructured written text provided by participants, a support person, CHWs, or an RN[b].
Interview transcripts	End-of-study interviews about acceptance perceptions.
Likelihood-to-recommend rating	Participant rating on a 0 - 10 scale indicating how likely they would be to recommend the Smart Health System to family or friends when a system (like the prototype) would be available on the market. Ratings were subsequently categorized as promoters (9-10), passives (7-8), and detractors (0 - 6) to calculate the Net Promoter Score.

[a]CHW: community health worker.

[b]RN: registered nurse.

Each alert included a secure survey link for reporting current health status. Unacknowledged alerts prompted CHW follow-up and escalation to an RN as needed. Narrative data were drawn from three sources: free-text survey responses, CHW/RN field notes after alert-related contacts, and semistructured exit interviews. Although 6 additional languages (Cantonese, Mandarin, Korean, Vietnamese, Spanish, and Russian) were spoken beyond English, multilingual CHWs communicating in preferred languages documented all responses in English.

At a conceptual level, the smart health system generated alerts using 3 components: sensing, rule-based interpretation, and filtering. Ambient motion and door sensors continuously captured movement throughout the home, and these raw signals were aggregated into higher-level behavioral events representing where and when a participant was active (eg, time in the kitchen, nighttime location, frequency of door use). A set of rules, co-designed with community partners and the PIs who were RNs, compared each participant's recent behavior against their own established baseline routine to identify clinically meaningful deviations, such as prolonged inactivity, disrupted sleep, reduced kitchen use, or missed bathing patterns. When a deviation crossed a predefined threshold, an alert was generated. To prevent alert fatigue, repeat alerts of the same type within a 1-week window were suppressed.

Multimedia Appendix 1 describes the full technical specifications of the sensors, processing pipeline, rule definitions, and figures exhibiting the health-monitoring dashboard (deidentified) and the clinical research dashboard, which shows an exemplary behavioral anomaly where a participant did not exit their home for nearly 3 days.

Data Analysis

Quantitative data, including participant characteristics, alert metrics, and response pathways, were summarized descriptively. Response trajectories were prospectively examined to identify alerts that elicited participant, support person, and CHW responses and those that resulted in subsequent RN action.

Qualitative data included free-text survey responses, follow-up narratives documenting health changes, and semistructured exit interviews. The data were analyzed using qualitative descriptive methods [50] and reflexive thematic analysis [51]. Coding was inductive [52] and iterative, with repeated review of written text-based data to refine codes, categories, and themes.

To strengthen rigor, the analytic team included RNs, a family nurse practitioner, bilingual CHWs, and engineering or computing expertise selected to reflect the linguistic, cultural, and clinical heterogeneity of the study population. The team met across 4 synchronous analytic circles and 8 asynchronous review sessions to discuss emerging patterns, compare coding perspectives, and refine thematic structure. After each session, written analytic summaries were synthesized into a shared memo for iterative team review. Related subthemes were consolidated into higher-order concepts through consensus discussion. An audit trail of analytic decisions, supporting written text-based excerpts, and reflexive memos was maintained throughout to enhance transparency and cultural sensitivity.

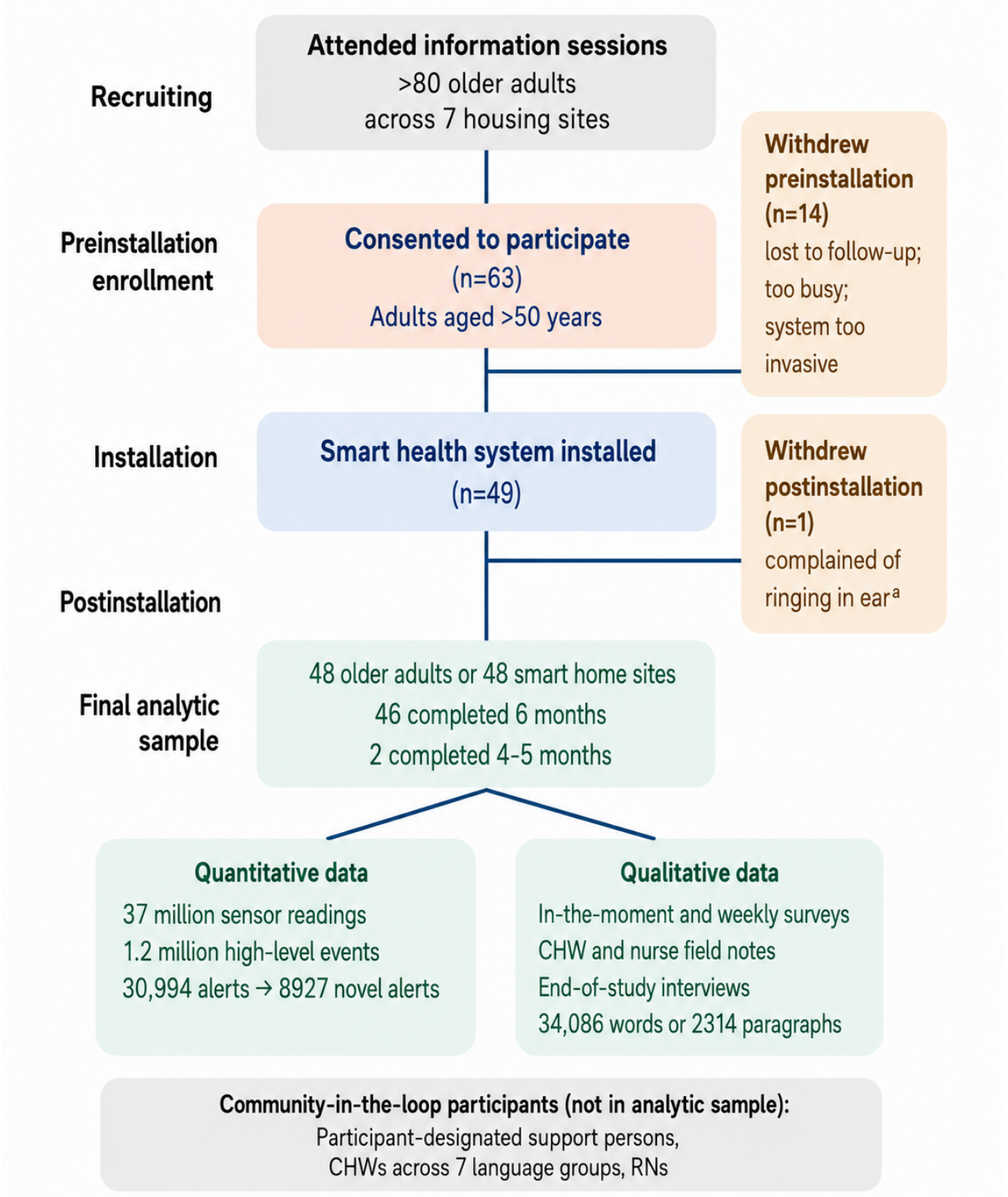
Results

Participant Characteristics

A total of 63 older adults consented to participate in the study, of whom 46 completed 6 months of monitoring and 2 completed 4 and 5 months of monitoring, respectively. These 2 individuals were included in the total sample because they completed two-thirds or more of the study. Fifteen participants withdrew, with reasons including privacy and system invasiveness, discomfort with the perceived invasiveness of questions asked by the Internet Service Provider if the internet was not already connected in the home (free internet was available in subsidized housing), and time constraints. The total sample included 48 older adults (ie, 48 smart home sites). Figure 2 summarizes participant flow from initial information sessions through to the final analytic sample.

XSL•FO
RenderX

Figure 2. Participant flow through the smart health system study. From >80 older adults who attended initial information sessions across 7 subsidized housing sites and individual low-income constituent homes of community partners, 63 consented to participate. Following 14 preinstallation withdrawals and 1 postinstallation withdrawal, the final analytic sample consisted of 48 older adults (48 smart home sites), of whom 46 completed the full 6-month monitoring period and 2 completed 4 to 5 months (≥2/3 of the study duration). [a]The ambient sensors did not have sound-producing capabilities. CHW: community health worker; RN: registered nurse.



Sample demographics are presented in Table 2. Participants ranged in age from 54 to 94 years, with 77.1% (n=37) identifying as female. Educational attainment varied, with 44% (n=18) holding a high school diploma and 17% a college degree. Seven languages were represented across 4 racial groups and 8 ethnicities. Most participants lived alone (n=44, 92%) and had access to the internet before consenting to participate (n=38, 79%). Ten (21%) participants received internet access via study-provided hotspots. The most prevalent conditions among participants were heart (n=23, 47.9%), musculoskeletal (n=12, 25%), and gastrointestinal (n=6, 12.5%) conditions.

XSL•FO
RenderX

Table . Sample demographics (N=48).

Demographic variables	Participants
Age (y), mean (SD; median, IQR; range)	76.1 (8.3; 80.5, 72-83; 54 - 94)
Total (one missing response)	47
Education level, n (%)	
Some middle school or no degree	2 (4.9)
Some high school or no diploma	3 (7.3)
High school or General Educational Development	18 (43.9)
Some college or no degree	11 (26.8)
Associate degree	5 (12.2)
Bachelor's degree	2 (4.9)
Total	41 (100)
Sex, n (%)	
Female	37 (77.1)
Male	11 (22.9)
Total	48 (100)
Race ethnicity (language), n (%)	
Asian Chinese (Cantonese)	5 (10.4)
Asian Chinese (Mandarin)	3 (6.3)
Asian Korean	2 (4.2)
Asian Vietnamese	6 (12.5)
Asian	1 (2.1)
African American	1 (2.1)
Native American—Indian	1 (2.1)
White (Russian)	1 (2.1)
White (English)	28 (58.3)
Total	48 (100.1[a])
Living arrangement, n (%)	
Independent (lives alone)	44 (91.7)
Lives with family	4 (8.3)
Total	48 (100)
Internet, n (%)	
Already had internet	38 (79.2)
Internet provided (hotspot)	10 (20.8)
Total	48 (100)
Chronic conditions, n	
Eyes-ears-nose-throat (sinusitis, vertigo, glaucoma)	3
Brain and mental health (traumatic brain injury, cerebrovascular disease, sleep apnea, panic disorder, bipolar disorder, posttraumatic stress disorder)	5
Metabolic-hormone (type 2 diabetes, hypothyroidism, gout)	14
Heart (congestive heart failure, coronary artery disease, hypertension, high cholesterol)	23
Lung (chronic obstructive pulmonary disorder, asthma)	7
Gastrointestinal (irritable bowel, acid reflux)	6

XSL•FO
RenderX

Demographic variables	Participants
Musculoskeletal (arthritis, osteoarthritis, carpal tunnel, restless leg syndrome, chronic back pain, osteoporosis)	12
Cancer	1

[a]The sum of percentages could exceed 100 due to rounding.

Alert Response Rates and Follow-Up Actions

Participants received 4719 emails and SMS text message alert reminders when baseline movement patterns changed, breaching thresholds of the custom rules-based monitoring algorithm. If a participant did not respond to their email alert within an hour, they received an SMS text message reminder. If they still did not respond, their support person received an email followed by an SMS text message alert. If the support person did not respond, the alert was automatically escalated to a CHW who then contacted the participant as soon as possible by phone. After making contact, the CHW responded to the alert on behalf of the participant and either documented a nonissue or supported the health situation as needed, including escalating a health issue to the research team RN as needed per protocol. CHW and RN responses to alerts included phone outreach and asking targeted questions related to 1 of 6 alert types (Homebody, Daytime Resting, Eating Routines, Bathing, Restless Night, Changing Sleep Routines). For example, for a Homebody alert, the participant would be asked how they were feeling and how long it had been since they last left their house or had visitors over. For a Restless Night alert, the participant would be asked about changes in bedtime, morning wake time, and the number of sleep interruptions that included getting out of bed, as well as the participant's perspective on the cause of sleep interruptions (eg, needing to go to the bathroom to void, experiencing pain, or worrying). If appropriate, the participant was asked to take their vital signs on equipment provided by the study team. Across the 6-month monitoring period, CHWs and nurses completed 1060 participant contacts, resulting in 72 CHW or RN interventions combined (nonurgent and urgent). Commonly, CHW interventions included therapeutic listening and recommendations to avoid dehydration and get enough rest. Common RN interventions included advice to take medications as prescribed, check blood sugar, take vital signs (blood pressure, heart rate, oxygen saturation), or call their support person. For 93% (4389/4719) of the alerts, no action was required. In six instances, the RN suggested the participant call their provider the same day (n=6, <1%). In 5 instances (n=4, <1%), the RN recommended urgent care.

Participant digital response rates were 1.57% (n=74) for initial email alerts and 7.79% (n=368) for follow-up SMS text message alerts. Support persons responded to 2.37% (n=112) of the email alerts and 5.21% (n=246) of the SMS text message alerts. CHWs responded to 78.36% (n=3698) of the alerts, including 4.68% (n=221) within 1 hour, with the remaining responses occurring after 5 hours. In contrast to the very low response rates via email (n=186, 3.94%) and SMS text message (n=614, 13.01%) from participants and support persons, CHWs demonstrated substantially higher engagement, responding to 78.26% (n=3698) of the alerts.

Net Promoter Score

Participants in the study completed a likelihood-to-recommend rating during their end-of-study interview using a 0 - 10 scale (0=would never recommend the health monitoring system and 10=would highly recommend it to family or friends). Fifteen participants did not provide a rating. One response knowingly exceeded the scale maximum and was recoded to 10 prior to analysis in accordance with the instrument's scale limits.

Among respondents (n=33), the mean likelihood-to-recommend rating was 8.68 (SD 1.68), the median was 9.0 (IQR 8.0 - 10.0), and scores ranged from 3 to 10.

Following a standard Net Promoter Score methodology, respondents were then classified as promoters (9-10), passives (7-8), or detractors (0 - 6). Most respondents were promoters (22/33, 67%), while 21% (n=7) were passives and 12% (n=4) were detractors, resulting in a Net Promoter Score of +55.

Qualitative Insights From Participant Narratives

Overview

Three major themes emerged from qualitative analysis—Alone, Trust, and Human Connection—encompassing 11 subthemes and related supporting concepts (Table 3). Narrative data were drawn from in-the-moment surveys with free-text fields, CHW and RN field notes, and end-of-study interviews. The corpus included 34,086 words across 2314 paragraphs. Multimedia Appendix 1 contains an extended table of themes and supporting quotations from participants' own words.

Table . Major themes, subthemes, and supporting concepts emerging from text-based data written by participants, community health workers (CHWs), and registered nurses (RNs).

Theme	Subthemes	Supporting concepts
Alone	Safety, Timing, Reassurance	Watched Over, Comfort, Loneliness, Spoken Language, Promise
Trust	Safe or Secure, Harm, Burden	Fraud, Data Ownership, Reliance, Normalizing Symptoms
Human Connection	Features, Personalization, Digital Distress	Expectations, AI Not Good at Context, Who Assesses Symptoms

XSL•FO
RenderX

Theme 1: Alone

Participants frequently described concerns about being alone while managing multiple chronic conditions, particularly when feeling unwell or lacking nearby family support. Many described the system as increasing feelings of safety, support, and connection. Several participants stated that the system would be especially valuable for those living alone or without family support. One participant shared that at "[my] age and being alone it made [me] feel safe and supported" (participant 25), while another described seeing the sensor light activate at night and thinking, "I'm not alone" (participant 1).

Timing also emerged within this theme, with some participants describing the system as more relevant during periods of worsening illness or future decline. Participants additionally described the comfort associated with the system's presence in the home.

CHW and RN narratives documented many cases in which participants preferred direct phone communication rather than email or SMS text message alerts. Participants were much more responsive to calls than emails or SMS text messages. In addition, they were more responsive when the calls were conducted in their preferred language.

Theme 2: Trust

Trust-related narratives centered on technology, the clinical team introducing the system, data privacy, fraud, and system reliability. Some participants also described broader distrust of the health care system. One participant stated:

> *I've kind of given up on seeing doctors for it [neuropathy pain]; they don't seem to do much or offer any real help.* [Participant 50]

At the same time, participants described trust in nurses and reassurance associated with nurse availability. One participant noted it was comforting to have "a nurse more available than to wait on the line for a nurse" (participant 35). Others described reassurance for distant family members and reduced caregiving pressure, noting that the system allowed loved ones to know they were being "watched over" (participant 23) and the system could "take pressure off [my] daughter" (participant 20).

Three recurring subthemes were evident: Trust in Nurses, Data Security and Identity Protection, and Cultural Values surrounding intergenerational caregiving. Participants frequently reported answering calls only from known phone numbers, and several added the research number to their phone contacts to avoid thinking the number was a scam caller. Among culturally based communities, participants also questioned how the system aligned with traditional family caregiving expectations.

A cross-cutting concept involved participants' own perceptions of health in relation to the system's sensor-based feedback. One participant described ignoring alerts after perceiving that the system had misinterpreted movements (participant 56), whereas another interpreted the absence of an alert as an indication that she was well despite recent symptoms (participant 57).

Theme 3: Human Connection

Participants consistently described preferring direct contact with CHWs or nurses, by phone or in person, over automated alerts and response systems. Many requested CHW or RN phone numbers to save in their contacts and described these interactions as the most valued component of the system.

Access to a health care team member was also described in relation to contextual personalization when system alerts did not align with individual routines. One participant explained:

> *[M]y pattern is different from others, and the system was not getting it. I kept getting alerts that were not relevant to me … but if it worked, I would have a direct nurse to call versus an on-call nurse.* [Participant 47]

Participants also described uncertainty about whether they were using the system correctly, frequently asking questions such as "am I doing it right?" (participant 1). Others reported discomfort with sensor activity, including flashing lights that were "sometimes … bothersome" (participant 11).

Persons withdrawing from the study, both before and after the installation of the sensor, reported concerns about the lack of personalization, invasion of privacy, lack of preferred-language communication, and uncertainty regarding whether a human or the technology was responsible for monitoring health changes.

Discussion

Overview

Central Findings

This mixed methods study extends prior smart home and aging-in-place research by centering older adults managing multiple chronic conditions while experiencing poverty. Consistent with prior work, trust, avoiding aloneness, and human connection emerged as foundational to adoption [12,34,41,43,53-55]. Our findings deepen this literature by showing that adoption barriers in low-income settings extend beyond usability and technical performance to include identity theft concerns, digital distress, cultural caregiving expectations, and mistrust of institutional systems.

A central finding was that trust operated at multiple levels: trust in the technology itself, trust in the people introducing and supporting it, and trust in one's own bodily awareness relative to system feedback. Participants' willingness to engage was strongly shaped by the presence of trusted RNs and CHWs, who functioned as relational anchors and trusted messengers. This aligns with prior research showing that older adults prefer health technologies that augment, rather than replace, human care [56-58]. In this study, RN and CHW involvement increased confidence in the system, reduced uncertainty, and helped participants contextualize alerts within daily life.

The Human Connection theme further reinforces that relational support is not merely an implementation enhancement but may be a core design requirement for aging-in-place technologies. Participants consistently preferred speaking with CHWs or RNs over interacting with automated email or SMS text message

XSL•FO
RenderX

alerts, despite generally favorable overall system ratings. High CHW response rates, alongside low participant digital response rates, suggest that human-mediated communication may be more acceptable than direct self-management of digital alerts, particularly among older adults navigating multimorbidity and poverty. These findings support the integration of human-centered communication pathways into smart home ecosystems, ensuring that technology serves as an extension of trusted care relationships rather than a replacement. Collectively, these findings (the need for human connection alongside low digital response rates) argue against a binary choice between fully automated and fully human-mediated monitoring. Hybrid, tiered models may offer a more pragmatic path forward. Automated sensing can support continuous objective data capture, while human contact, particularly by phone, remains the primary mode of communicating clinically relevant changes. Such a model could reserve intensive CHW and RN involvement for higher-acuity alerts or higher-risk participants. Lower-acuity changes could be routed for brief check-ins by support persons or CHWs or asynchronous follow-up, allowing the relational core of the system to scale without requiring proportional growth in staffing.

Digital Distress

The findings also introduce digital distress as an important extension of the broader technostress literature. The first clues that the concept of digital distress would eventually emerge from the data were noted during the installation phase. Seven people changed their minds about participating after encountering too many stress-inducing personal and financial questions from internet service providers when enrolling in free internet so they could participate in the study. (Note: this led to a design change where we added hot spots to the system.) Additionally, 1 person said it would be too stressful to receive calls from unknown numbers (ie, for the nursing calls), and another said the system was too invasive and would stress them out. Overall, digital distress was noted when technologies felt intrusive, difficult to interpret, poorly personalized, or disconnected from participants' language and cultural context. Digital stress was especially pronounced when users were uncertain whether a human or the system was responsible for monitoring health changes. This role ambiguity, combined with concerns about fraud, flashing sensors, or burdensome connectivity processes, contributed to participant discomfort and study withdrawal. For older adults experiencing poverty, digital distress may be amplified by chronic illness burden, financial precarity, and prior experiences with scams or identity theft, making low-burden and transparent workflows essential. The antidote is the *human connection*, which can mitigate stress, reduce *aloneness*, and promote *trust*.

Safety

A related safety implication concerns the risk of false reassurance through error of omission. Though participants were told that the absence of alerts does not mean the health is stable, some participants still interpreted the absence of alerts as evidence that they were well, even when experiencing symptoms. This suggests that missed detections or insufficient personalization may unintentionally increase reliance on incomplete system feedback. Future smart home systems should therefore prioritize explainable alert logic, transparent thresholds, and explicit communication that the absence of an alert does not equate to clinical stability. These findings have direct implications for digital safety, implementation ethics, and the design of AI-supported aging-in-place tools.

Older Adults' Resources

Poverty-specific barriers were highly influential. The process of accessing free internet felt invasive or financially risky and resulted in several withdrawals from this study. Fear of hidden fees, identity theft, and lack of ongoing support after study completion undermined the willingness to continue with the internet sign-up process. These findings highlight how connectivity itself can become a structural barrier, even when technologies are designed to be low cost.

Though we met our target of less than US $400 per home for hardware, there were other development and implementation costs; primarily personnel time (engineering contractor, nursing, multilingual CHWs, research assistants), training staff, and hotspot connectivity subscriptions for participants without internet. Technical assistance throughout the study was included in the engineering contract, but that may not always be the case. Often, technical assistance is charged separately. Multilingual CHWs received differential pay for each additional language spoken (an additional US $3 per hour). Supply chain issues and inflation impacted hardware availability and costs over a 6-month purchase timeframe during the COVID-19 pandemic (November 2021-May 2022). These pragmatic discoveries suggest that equitable scaling requires budgeting not only for hardware but also for connectivity, technical maintenance, and language access.

Feasibility

Built Environment

The built environment also shaped feasibility. Older housing, limited outlets, and brick-and-steel walls interfered with connectivity and sensor placement. Participants also disliked visible flashing lights and preferred less obtrusive hardware. These findings reinforce that environmental adaptability and unobtrusive sensor design are central to adoption, especially in subsidized or aging housing where infrastructure limitations are common.

Individualization

Our results further suggest that rules-based algorithms are insufficient for diverse older adult populations. Participants frequently reported irrelevant alerts or missed meaningful changes because standardized thresholds failed to reflect individualized routines. One participant responded to an email alert saying they wanted alerts to their support person to stop because "I don't want to alarm her!" and the alerts, in the participant's opinion, were inaccurate. This mismatch undermined trust and increased burden. Context-aware machine learning models that adapt to personal schedules, cultural practices, and evolving health patterns may offer a more appropriate pathway for equitable smart home monitoring.

XSL•FO
RenderX

However, such systems must preserve transparency and human oversight to avoid exacerbating mistrust.

Cultural Sensitivity

Cultural expectations around caregiving were another key determinant of adoption. In some communities, participants expressed concern that technology might undermine family caregiving roles or conflict with intergenerational obligations. Systems that intentionally incorporate family, CHWs, and RN oversight may therefore better align with users' values and improve engagement. Rather than positioning monitoring technologies as substitutes for caregiving, framing them as tools that strengthen family and community support structures may enhance trust and long-term use.

These contextual findings suggest that successful smart home health systems for older adults experiencing poverty require more technical accuracy than was possible within the scope of this study. Such systems require relational trust, cultural responsiveness, personalization, environmental adaptability, and sustainable implementation infrastructure. Human support—particularly through CHWs and nurses—appears central to mitigating digital distress, strengthening trust, and promoting equitable aging in place.

Principal Results

This study identified several barriers to smart home adoption among older adults managing multiple chronic conditions while experiencing poverty. Beyond technical feasibility, adoption was shaped by digital access barriers, identity theft concerns, cultural caregiving expectations, environmental constraints, and structural inequities tied to poverty. Across qualitative themes, Aloneness, Trust, and Human Connection emerged as central determinants of engagement.

A key finding was that participants consistently preferred human-mediated communication, particularly with CHWs and RNs, over automated alerts. Although direct digital response rates were low, CHW engagement was high, suggesting that integrating trusted community and clinical personnel into the health care-technology-patient loop may be critical for sustained use in low-income settings. Future research should test how to reduce the staffing intensity of this model while preserving the human contact participants valued, including through acuity-based tiering, machine learning personalization, and language-concordant outreach.

This study also introduces digital distress, a novel concept describing stress that arises when technology feels intrusive, confusing, insufficiently personalized, or disconnected from linguistic and cultural context. Digital distress was most evident when participants were uncertain how the system worked, whether they were using it correctly, or who was responsible for monitoring health changes.

Collectively, these findings indicate that smart health systems designed around human relationships, cultural responsiveness, and contextual adaptation, and implemented with trusted CHW and nursing support, may better promote equitable aging in place among older adults managing multimorbidity while experiencing poverty.

Limitations and Future Research

This study has several limitations. The sample was relatively small and drawn from a specific population of older adults managing multimorbidity while experiencing poverty. Despite the extant literature defining older adulthood as age 65 years or above, the study sample included immigrant adults aged 50 years or above who were managing multimorbidity, which commonly affects these populations at younger ages due to the impacts of social determinants of health. This may limit transferability across populations, settings, and cultural contexts. In addition, the smart health system was a prototype that underwent iterative refinement during the study, and its rules-based alerting approach may have contributed to irrelevant or missed alerts. The intensive involvement of CHWs and RNs, while central to equitable implementation, may also limit scalability in other care settings. Because digital distress emerged inductively within this context, additional research is needed to examine its relevance across populations and technologies. Future work should evaluate scalable machine learning approaches, multilingual and culturally responsive interfaces, improved alert accuracy and data visualization, integration with clinical workflows, and strategies for communicating automated messages and engaging older adults and families with sensor-derived health data while preserving simplicity and trusted human support.

Conclusions

Smart home health monitoring systems hold promise for supporting aging in place and helping older adults maintain independence, but their success among those experiencing poverty depends on more than technological innovation. If systems are not designed with attention to the unique barriers faced by this population, they risk reinforcing existing health inequities. This study demonstrates that adoption is shaped by a complex interplay of digital literacy, cultural values, trust in institutions, and the physical realities of low-income housing. A community-in-the-loop approach that integrates ambient sensors, automated unobtrusive monitoring, and human-centered care networks may reduce isolation, foster trust, and maintain human connection while offering real-time health insights. To be effective and equitable, future smart health systems must be co-designed with communities and prioritize personalization, cultural sensitivity, and low-burden engagement. In addition, technology-enabled health monitoring for the home must include intentional, explicit strategies to mitigate digital distress. These findings offer a roadmap for researchers, designers, and policymakers seeking to close the digital divide and promote health equity through technology-enabled care.

XSL•FO
RenderX

Acknowledgments

We thank the publicly elected board of the Vancouver Housing Authority (2022 - 2025) for their support as well as the subsidized housing building managers in the Southwest Washington region. We also thank the leaders and community health workers at the Asian Health and Service Center (Oregon) and Battle Ground Healthcare (Washington) for their partnership and assistance in recruiting and data collection. These organizational partnerships and the individuals intimately involved in the study process were critical to the project's success.

Funding

This is funded by the Gordon and Betty Moore Foundation through Grant GBMF9048 to Washington State University Vancouver to support this work led by Roschelle Fritz, PhD, RN, FAAN, now Professor of Nursing at the University of California Davis Health Betty Irene Moore School of Nursing. Generative AI was not used in the preparation of this manuscript. All ideas, analyses, interpretations, and writing are the authors' own.

Data Availability

Deidentified data presented in this study are available for research purposes upon request, due to the small sample size and associated health information.

Authors' Contributions

RLF conceived the study, secured funding, designed the methodology, supervised the project, interpreted the findings, and wrote the manuscript and Multimedia Appendix 1. CKYN-T contributed to community partnerships, participant recruitment, data collection, interpretation of findings, preparation of Table 2, and critical revision of the manuscript. JP critically revised the manuscript and prepared Table A2 in Multimedia Appendix 1. THN contributed to data collection. GB contributed to the engineering, implementation, and technical data collection for the smart home system. HS contributed health economics expertise and critically reviewed the manuscript. EEH and AM contributed to data collection and interpretation of findings. YLBR contributed to interpretation of the findings. CZ contributed to the design and analysis of the Net Promoter Score, interpretation of Net Promoter Score findings, and wrote the associated content. MR provided senior scientific oversight, guided the study design and engineering collaborations, contributed to interpretation of the findings, and critically revised the manuscript. All authors reviewed and approved the final manuscript and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Smart health system prototype description (mHealth Evidence Reporting and Assessment aligned).
[DOCX File, 3258 KB - aging_v9i1e88290_app1.docx]

References

1. World social report 2023: leaving no one behind in an ageing world. : United Nations; 2023 URL: https://www.un-ilibrary.org/content/books/9789210019682 [accessed 2026-07-07]
2. Reyes AM, Schoeni RF, Freedman VA. National estimates of kinship size and composition among adults with activity limitations in the United States. Demogr Res 2021;45:1097-1114. [doi: 10.4054/demres.2021.45.36] [Medline: 36051489]
3. Wolff JL, Cornman JC, Freedman VA. The number of family caregivers helping older US adults increased from 18 million to 24 million, 2011–22. Health Aff (Millwood) 2025 Feb;44(2):187-195. [doi: 10.1377/hlthaff.2024.00978] [Medline: 39899774]
4. Morita PP, Sahu KS, Oetomo A. Health monitoring using smart home technologies: scoping review. JMIR Mhealth Uhealth 2023 Apr 13;11(1):e37347. [doi: 10.2196/37347] [Medline: 37052984]
5. Cook DJ, Schmitter-Edgecombe M. Fusing ambient and mobile sensor features into a behaviorome for predicting clinical health scores. IEEE Access 2021;9:65033-65043. [doi: 10.1109/access.2021.3076362] [Medline: 34017671]
6. Sovacool BK, Martiskainen M, Furszyfer Del Rio DD. Knowledge, energy sustainability, and vulnerability in the demographics of smart home technology diffusion. Energy Policy 2021 Jun;153:112196. [doi: 10.1016/j.enpol.2021.112196]
7. Pirzada P, Wilde A, Doherty GH, Harris-Birtill D. Ethics and acceptance of smart homes for older adults. Inform Health Soc Care 2022 Jan 2;47(1):10-37. [doi: 10.1080/17538157.2021.1923500] [Medline: 34240661]
8. Huang G, Oteng SA. Gerontechnology for better elderly care and life quality: a systematic literature review. Eur J Ageing 2023 Jun 22;20(1):27. [doi: 10.1007/s10433-023-00776-9] [Medline: 37347277]
9. Dermody G, Fritz R, Glass C, Dunham M, Whitehead L. Factors influencing community-dwelling older adults' readiness to adopt smart home technology: a qualitative exploratory study. J Adv Nurs 2021 Dec;77(12):4847-4861. [doi: 10.1111/jan.14996] [Medline: 34477222]

XSL•FO
RenderX

10. MacDonald J, Demiris G, Shevin M, Thadaney-Israni S, Jay Carney T, Cupito A. Health technology for all: an equity-based paradigm shift opportunity. NAM Perspect 2022;2022:10-31478. [doi: 10.31478/202212a] [Medline: 36713773]
11. Demiris G. Information technology to support aging: implications for caregiving. In: Gaugler JE, Kane RL, editors. Family Caregiving in the New Normal: Academic Press; 2015:211-223. [doi: 10.1016/B978-0-12-417046-9.00012-X]
12. Fritz RL, Nguyen-Truong CKY, Leung J, et al. Older Asian immigrants' perceptions of a health-assistive smart home. Gerontechnology 2020 Dec 31;19(4):1-11. [doi: 10.4017/gt.2020.19.04.385]
13. Lakshminarayanan V, Ravikumar A, Sriraman H, Alla S, Chattu VK. Health care equity through intelligent edge computing and augmented reality/virtual reality: a systematic review. J Multidiscip Healthc 2023;16:2839-2859. [doi: 10.2147/JMDH.S419923] [Medline: 37753339]
14. Fritz RL, Dermody G. A nurse-driven method for developing artificial intelligence in "smart" homes for aging-in-place. Nurs Outlook 2019;67(2):140-153. [doi: 10.1016/j.outlook.2018.11.004] [Medline: 30551883]
15. Chien SC, Yen CM, Chang YH, et al. Use of artificial intelligence, internet of things, and edge intelligence in long-term care for older people: comprehensive analysis through bibliometric, Google trends, and content analysis. J Med Internet Res 2025 Mar 4;27(1):e56692. [doi: 10.2196/56692] [Medline: 40053718]
16. Smart home technology for older adults. AgeTech Collaborative from AARP. 2026. URL: https://agetechcollaborative.org/insights/smart-home-tech-for-older-adults/ [accessed 2026-07-07]
17. Rantz M, Phillips LJ, Galambos C, et al. Randomized trial of intelligent sensor system for early illness alerts in senior housing. J Am Med Dir Assoc 2017 Oct 1;18(10):860-870. [doi: 10.1016/j.jamda.2017.05.012] [Medline: 28711423]
18. Rantz MJ, Skubic M, Popescu M, et al. A new paradigm of technology-enabled "Vital Signs" for early detection of health change for older adults. Gerontology 2015;61(3):281-290. [doi: 10.1159/000366518] [Medline: 25428525]
19. Cook DJ, Crandall AS, Thomas BL, Krishnan NC. CASAS: a smart home in a box. Computer 2013 Jul;46(7):62-69. [doi: 10.1109/MC.2012.328] [Medline: 24415794]
20. Cook DJ, Krishnan NC, Rashidi P. Activity discovery and activity recognition: a new partnership. IEEE Trans Cybern 2013 Jun;43(3):820-828. [doi: 10.1109/TSMCB.2012.2216873] [Medline: 23033328]
21. Fritz R, Wuestney K, Dermody G, Cook DJ. Nurse-in-the-loop smart home detection of health events associated with diagnosed chronic conditions: a case-event series. Int J Nurs Stud Adv 2022 Dec;4:100081. [doi: 10.1016/j.ijnsa.2022.100081] [Medline: 35642184]
22. Wilson M, Fritz R, Finlay M, Cook DJ. Piloting smart home sensors to detect overnight respiratory and withdrawal symptoms in adults prescribed opioids. Pain Manag Nurs 2023 Feb;24(1):4-11. [doi: 10.1016/j.pmn.2022.08.011] [Medline: 36175277]
23. Albahri OS, Alamleh A, Al-Quraishi T, Thakkar R. Smart real-time IoT mhealth-based conceptual framework for healthcare services provision during network failures. Appl Data Sci Anal 2023 Nov 2;2023:110-117. [doi: 10.58496/ADSA/2023/010]
24. Rahman A, Wadud MAH, Islam MJ, et al. Internet of medical things and blockchain-enabled patient-centric agent through SDN for remote patient monitoring in 5G network. Sci Rep 2024 Mar 4;14(1):5297. [doi: 10.1038/s41598-024-55662-w]
25. Thomas NW, Beattie Z, Riley T, Hofer S, Kaye J. Home-based assessment of cognition and health measures: the collaborative aging research using technology (CART) initiative and international collaborations. IEEE Instrum Meas Mag 2021 Sep;24(6):68-78. [doi: 10.1109/MIM.2021.9513638]
26. Mode NA, Evans MK, Zonderman AB. Race, neighborhood economic status, income inequality and mortality. PLOS ONE 2016;11(5):e0154535. [doi: 10.1371/journal.pone.0154535] [Medline: 27171406]
27. Beech BM, Ford C, Thorpe RJ, Bruce MA, Norris KC. Poverty, racism, and the public health crisis in America. Front Public Health 2021;9:699049. [doi: 10.3389/fpubh.2021.699049] [Medline: 34552904]
28. Thornton M, Bowers K. Poverty in older adulthood: a health and social crisis. Online J Issues Nurs 2024 Jan 31;29(1). [doi: 10.3912/OJIN.Vol29No01Man03]
29. McIver DL. CMS framework for health equity 2022–2032. : Centers for Medicare & Medicaid Services (CMS); 2022 URL: https://www.cms.gov/files/document/cms-framework-health-equity.pdf [accessed 2026-07-07]
30. Fritz RL, Cook D. Identifying varying health states in smart home sensor data: an expert-guided approach. Presented at: Proceedings of The 21st World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2017); Nov 12-17, 2017 URL: https://www.iiis.org/CDs2017/CD2017Summer/papers/SA395LR.pdf [accessed 2026-07-07]
31. Dahmen J, Cook DJ. Indirectly supervised anomaly detection of clinically meaningful health events from smart home data. ACM Trans Intell Syst Technol 2021 Mar;12(2):1-18. [doi: 10.1145/3439870] [Medline: 34336375]
32. For industry and community partners. Oregon Health & Science University (OHSU). 2026. URL: https://www.ohsu.edu/collaborative-aging-research-using-technology/industry-and-community-partners [accessed 2026-07-07]
33. Tian YJA, Felber NA, Pageau F, Schwab DR, Wangmo T. Benefits and barriers associated with the use of smart home health technologies in the care of older persons: a systematic review. BMC Geriatr 2024 Feb 14;24(1):152. [doi: 10.1186/s12877-024-04702-1] [Medline: 38355464]
34. Dermody G, Shibl R, Wang M, et al. Multi perspective considerations for health smart home: early phase exploratory study. J Adv Nurs 2026 Jan;82(1):849-866. [doi: 10.1111/jan.16945] [Medline: 40172858]
35. Fritz RL, Nguyen-Truong CKY, May T, Wuestney K, Cook DJ. Bioethics principles in machine learning-healthcare application design: achieving health justice and health equity. Harv Public Health Rev (Camb) 2024;79:10-54111. [doi: 10.54111/0001/aaaa1] [Medline: 39850650]

XSL•FO
RenderX

36. Triantafyllidis AK, Tsanas A. Applications of machine learning in real-life digital health interventions: review of the literature. J Med Internet Res 2019 Apr 5;21(4):e12286. [doi: 10.2196/12286] [Medline: 30950797]
37. Szymczak JE, Fiks AG, Craig S, Mendez DD, Ray KN. Access to what for whom? How care delivery innovations impact health equity. J Gen Intern Med 2023 Apr;38(5):1282-1287. [doi: 10.1007/s11606-022-07987-3] [Medline: 36627525]
38. Hadjiat Y. Healthcare inequity and digital health—a bridge for the divide, or further erosion of the chasm? PLOS Digit Health 2023 Jun;2(6):e0000268. [doi: 10.1371/journal.pdig.0000268] [Medline: 37267232]
39. Hoagland A, Kipping S. Challenges in promoting health equity and reducing disparities in access across new and established technologies. Can J Cardiol 2024 Jun;40(6):1154-1167. [doi: 10.1016/j.cjca.2024.02.014] [Medline: 38417572]
40. Whitelaw S, Pellegrini DM, Mamas MA, Cowie M, Van Spall HGC. Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review. Eur Heart J Digit Health 2021 Mar;2(1):62-74. [doi: 10.1093/ehjdh/ztab005] [Medline: 34048508]
41. Whitehead L, Talevski J, Fatehi F, Beauchamp A. Barriers to and facilitators of digital health among culturally and linguistically diverse populations: qualitative systematic review. J Med Internet Res 2023 Feb 28;25:e42719. [doi: 10.2196/42719] [Medline: 36853742]
42. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161. [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
43. Van Hoof J, Demiris G, Wouters EJ. Handbook of Smart Homes, Health Care and Well-Being: Springer International Publishing; 2017. [doi: 10.1007/978-3-319-01583-5]
44. Burns K, Bloom S, Gilbert C, et al. Centering digital health equity during technology innovation: protocol for a comprehensive scoping review of evidence-based tools and approaches. JMIR Res Protoc 2024 Jun 5;13(1):e53855. [doi: 10.2196/53855] [Medline: 38838333]
45. Papoutsi C, Wherton J, Shaw S, Morrison C, Greenhalgh T. Putting the social back into sociotechnical: case studies of co-design in digital health. J Am Med Inform Assoc 2021 Feb 15;28(2):284-293. [doi: 10.1093/jamia/ocaa197] [Medline: 33043359]
46. Weirauch V, Soehnchen C, Burmann A, Meister S. Methods, indicators, and end-user involvement in the evaluation of digital health interventions for the public: scoping review. J Med Internet Res 2024 May 31;26(1):e55714. [doi: 10.2196/55714] [Medline: 38819891]
47. Ravizza A, Conditi N. Human factors engineering methods for engaging end users in the co-design of healthcare 5.0 technologies. In: Bertolaso M, Ilardo ML, Ribera J, editors. Healthcare in the Digital Age: Springer; 2025:249-283. [doi: 10.1007/978-981-96-1437-0_12]
48. Nguyen-Truong CKY, Fritz RL, Xu J, et al. Nurturing belonging, reflection, and responsiveness in cross-sector partnerships: exemplar from community-based smart health system research. J Particip Res Methods 2025 May 23;6(2):242-254. [doi: 10.35844/001c.129279]
49. Annual update of the HHS poverty guidelines. Federal Register. 2023 Aug 18. URL: https://www.federalregister.gov/documents/2023/01/19/2023-00885/annual-update-of-the-hhs-poverty-guidelines [accessed 2026-07-07]
50. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
51. Braun V, Clarke V, Hayfield N, Davey L, Jenkinson E. Doing reflexive thematic analysis. In: Bager-Charleson S, McBeath A, editors. Supporting Research in Counselling and Psychotherapy: Springer International Publishing; 2022:19-38. [doi: 10.1007/978-3-031-13942-0_2]
52. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006 Jun;27(2):237-246. [doi: 10.1177/1098214005283748]
53. Hepburn J, Williams L, McCann L. Barriers to and facilitators of digital health technology adoption among older adults with chronic diseases: updated systematic review. JMIR Aging 2025 Sep 11;8(1):e80000. [doi: 10.2196/80000] [Medline: 40934502]
54. Chen Y, Yuan J, Shi L, et al. Understanding the role of technology anxiety in the adoption of digital health technologies (DHTs) by older adults with chronic diseases in Shanghai: an extension of the unified theory of acceptance and use of technology (UTAUT) model. Health Care (Don Mills) 2024 Jul 15;12(14):1421. [doi: 10.3390/healthcare12141421]
55. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
56. Kuo KT. When voice-based AI becomes a companion: protecting dignity and reimagining human–AI collaboration in aging-in-place. Res Connect 2026 Apr 9;1(2):vmag017. [doi: 10.1093/rescon/vmag017]
57. Genge C, McNeil H, Debergue P, Freeman S. Technology to support aging in place: key messages for policymakers and funders. Front Psychol 2023;14:1287486. [doi: 10.3389/fpsyg.2023.1287486] [Medline: 38078234]
58. Cohen Elimelech O, Rosenblum S, Tsadok-Cohen M, Ferrante S, Demeter N. Understanding older adults' technology use preferences and needs from a triangular perspective: qualitative study. J Med Internet Res 2025 Nov 11;27(1):e72716. [doi: 10.2196/72716] [Medline: 41218201]

XSL•FO
RenderX

Abbreviations

CART: collaborative aging research using technology
CASAS: Center for Advanced Studies in Adaptive Systems
CHW: community health worker
IoT: Internet of Things
IRB: Institutional Review Board
PI: primary investigator
RN: registered nurse

Edited by H Köttl; submitted 22.Nov.2025; peer-reviewed by M Tai-Seale, P Sivaranjani; revised version received 29.Jun.2026; accepted 30.Jun.2026; published 23.Jul.2026.

Please cite as:
Fritz RL, Nguyen-Truong CKY, Phipps J, Hinderlie EE, Rodriguez YLB, Nguyen TH, Barling G, Mishuk A, Schoonover H, Zuber C, Rantz M
A Community-in-the-Loop Approach to Smart Home Monitoring for Aging in Place: Mixed Methods Evaluation of a Co-Designed Prototype
JMIR Aging 2026;9:e88290
URL: https://aging.jmir.org/2026/1/e88290
doi:10.2196/88290

© Roschelle Fritz, Connie Kim Yen Nguyen-Truong, Jennifer Phipps, Ellen Hinderlie, Yolanda Rodriguez, Thai Hien Nguyen, Gabrielle Barling, Anna Mishuk, Heather Schoonover, Christi Zuber, Marilyn Rantz. Originally published in JMIR Aging (https://aging.jmir.org), 23.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smart Home Technology Integration in Home Modification Programs Serving Older Adults: Focus Group Study With Program Grantees

Haomin Hu[1], PhD; Moriah Clarke[2], BS; Meredith Hughes[3], JD, MPH; Palash Sharma[4], MS; Jack Fried[5], MRT; Portia Singh[5], PhD; Pamela Toto[2], PhD, OTR/L

[1]Department of Health Information Management, University of Pittsburgh, 3396 Fifth Ave, Room 4178, Pittsburgh, PA, United States

[2]Department of Occupational Therapy, University of Pittsburgh, Pittsburgh, PA, United States

[3]Department of Health Policy and Management, University of Pittsburgh, Pittsburgh, PA, United States

[4]Health Policy Institute, University of Pittsburgh, Pittsburgh, PA, United States

[5]Department of Rehabilitation Science and Technology, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Haomin Hu, PhD
Department of Health Information Management, University of Pittsburgh, 3396 Fifth Ave, Room 4178, Pittsburgh, PA, United States

Abstract

Background: As people age over the coming decades, demand for in-home support and other interventions, such as home modifications, to help older adults age in place successfully, is also expected to rise. Smart home technologies have the potential to enhance aging in place by complementing traditional home modifications; however, adoption within federally funded home modification programs remains limited.

Objective: This study explored grantees' perspectives on the Older Adults Home Modification Program (OAHMP) to understand current practices, perceived benefits, barriers, and strategies for integrating smart home technologies into home modification services.

Methods: An exploratory qualitative study was conducted with staff, occupational therapists, and home builders from 3 OAHMP grantee organizations. A 1.5-hour virtual focus group was held and thematically analyzed using a deductive approach grounded in the discussion agenda.

Results: A total of 11 participants reported early adoption of select smart devices—most commonly smart speakers, doorbell cameras, motion-activated lights, and smart plugs—to enhance home safety, communication, and independence. Barriers to broader implementation emerged at three levels: (1) older adults' digital literacy, privacy concerns, and device maintenance burden; (2) contextual constraints, such as unreliable internet in rural areas; and (3) organizational limitations, including training needs, staffing capacity, and funding challenges. Participants emphasized the importance of progressive adoption, hands-on training, and low-maintenance technology. In addition, partnerships with university educational programs were well established among selected grantees to provide technology training resources for older adults. To inform a reasonable technology selection for home modifications, criteria were proposed across 5 domains: installation, usability, accessibility, sustainability, and security and privacy.

Conclusions: Integrating smart home technologies into home modification programs provides a scalable, cost-effective opportunity to improve aging in place outcomes. Policy support, workforce training, valid selection criteria, and sustainable funding models are needed to promote equitable adoption across OAHMP and similar federally supported programs.

(JMIR Aging 2026;9:e89729) doi:10.2196/89729

KEYWORDS

smart home technology; aging in place; home modification program; internet of things; technology adoption

Introduction

Background

The share of Americans aged 65 years and older is projected to increase from 17% to 22% by 2040, and the number of Americans aged 85 years and older is projected to double over that same period [1]. The prevalence of disability increases with age, as approximately 14% of adults aged 65 to 74 years and 29% of adults aged 75 years and older have an ambulatory difficulty, defined as a challenge walking or climbing stairs,

XSL•FO
RenderX

and approximately 11% of the population aged 75 years and older has a self-care difficulty, defined as a challenge with bathing or dressing [2]. As the population ages over the coming decades, demand for in-home support and other interventions, such as home modifications, to help older adults age in place successfully, is expected to rise.

"Aging in place" refers to maintaining functional independence and social connections while living in the home or community [3]. A national survey reported that 75% of adults aged 50 years and older wish to stay in their own homes as they age, and 73% hope to stay in their communities [4], underscoring the importance of residential environments in supporting daily living and well-being. While remaining at home is desired by the vast majority of older adults, many American homes lack key features that provide the best person-environment "fit" to support aging in place [5]. As of 2019, about 40% of homes were considered aging-ready, with a bedroom and full bathroom on the first floor and a step-free entryway [2]. Today, only 54% of older adults aged 80 years and older reside in single-floor homes with step-free entryways [6]. Correspondingly, demand for in-home support and residential adaptations is expected to continue, placing a greater emphasis on strategies that enable older adults to live safely and independently.

Home Modifications to Support Aging in Place

One of the most widely recognized approaches to supporting aging in place is home modification. Home modifications are a prominent and essential feature of evidence-based programs, such as Community Aging in Place–Advancing Better Living for Elders (CAPABLE), which seek to reduce falls and improve the safety and independence of older adults aging in place [7]. Home modifications—such as installing grab bars, building ramps, or improving lighting—have been shown to reduce environmental hazards, enhance accessibility, and lower the risk of falls or injury [8]. Evidence indicates that such a solution not only improves safety but also delays institutionalization and reduces health care costs [9]. Despite these benefits, current home modification programs face challenges, including limited funding, a lack of service providers, and difficulties tailoring interventions to the diverse needs of older adults and their caregivers, including considerations of cost and time required to implement home modifications [10-13].

The cost of home modifications varies significantly depending on the nature of the alteration: items such as grab bars typically cost a few hundred dollars to acquire and install, whereas wheelchair ramps or stairlifts can cost several thousand dollars. Nearly one-third of older adult households are cost-burdened, defined as spending at least 30% of their income on housing, utilities, taxes, and insurance [6,14]. To address gaps between the need for and access to home modifications, the Older Adults Home Modification Program (OAHMP), established by the US Department of Housing and Urban Development, focuses on high-impact, low-cost home modification solutions. The OAHMP grant, created in 2021, provides funding to organizations to support low-income older adults in aging in place through home modifications that enhance the safety and accessibility of their home environment [15].

Smart Home Technologies in Home Modification Programs

Over the past decade, smart home technologies have undergone rapid evolution and become increasingly prevalent in everyday life. Devices such as voice-activated intelligent assistants (eg, Google Assistant, Alexa, and Siri), remote monitoring systems, motion-sensor lighting, and video doorbells are now widely available at relatively low cost. These technologies offer new opportunities to extend the benefits of traditional home modifications by enhancing safety, improving communication and monitoring, and supporting independence [16]. Their potential application in home modification programs is particularly promising, given the growing emphasis on technology-enabled care and prevention [17].

Although some OAHMP grantee organizations have reported integrating products such as automatic door openers, security systems, and fall-detection devices into their home modification installations, most OAHMP and similar home modification programs do not include smart technologies in their menu of environmental modifications [15]. Incorporating additional smart home technologies into the OAHMP could support a range of aging in place goals. For example, smart, motion-sensitive lighting can improve visibility and reduce fall risk in areas such as stairwells and bathrooms. Smart speakers can enable automated control of door locks, lighting, window blinds, and other devices, helping individuals with reduced mobility maintain control over their environment. Smart plugs and appliances, such as coffee pots or stoves, can enable caregivers to remotely support and monitor daily routines. Despite the widespread availability of these technologies and the growing inclusion of home modification interventions, their adoption in home modification programs remains limited [18]. Barriers, including a lack of standardization, concerns about cost and maintenance, concerns about technology acceptance by older adults, and uncertainty regarding training and support, have prevented consistent integration into service delivery [16]. Additionally, incorporating smart technologies into traditional home modification approaches requires a readiness for change by both providers and consumers, which may inhibit the uptake and diffusion of this innovation for aging in place [19,20].

To better shape future program design and policy support, it is essential to understand how practitioners view the opportunities and challenges of integrating smart home technologies. This study aimed to explore OAHMP grantees' perspectives on current practices, perceived benefits, barriers, and strategies for integrating smart home technologies into home modification services, and to propose criteria to guide technology selection for program implementation.

Methods

This was an exploratory qualitative study with thematic analysis of focus group feedback.

Recruitment

On the basis of sample size recommendations for focus groups [21], the research team worked with leadership from 3 OAHMP grantee organizations representing rural, suburban, and national

XSL•FO
RenderX

programs to purposively select 9 to 12 (3 - 4 per organization) professionals, consisting of home modification contractors, remodeling contractors, Certified Aging-in-Place Specialists (CAPSs), occupational therapists, other assessment professionals, and home modification program administrators or staff. Recruitment information was emailed to the leaders and subsequently forwarded to English-speaking professionals working for the regional organizations. No professionals were excluded on the basis of gender, race, or ethnicity.

Procedure

A single focus group session was conducted via Zoom and lasted approximately 1.5 hours. An agenda was prepared in advance, with discussion questions organized into three sections: (1) background and ongoing home assessments and modifications; (2) attitudes, current adoption, and barriers related to introducing smart home technology into home modification programs; and (3) future expectations for integrating smart home technology and potential motivators to support adoption (Multimedia Appendix 1). The diffusion of innovations theory, readiness for change, and technology acceptance models were used to guide question content [19,20,22,23]. Two facilitators (HH and MC) from the research team guided the session, alternating leadership across the 3 sections. The primary facilitator in each section posed questions and encouraged participants to share their perspectives, ensuring that all 3 participating organizations responded with comments. While 1 facilitator engaged with participants, the other took detailed notes to capture key points and emerging themes. The entire session was audio-recorded via Zoom to ensure accurate capture of participants' responses.

Data Analysis

The audio recording was transcribed using Zoom's cloud-integrated transcription feature. Data were analyzed using a deductive thematic analysis approach, guided by the topics outlined in the focus group agenda. A 10-theme codebook was developed in advance based on the agenda questions (Table 1). Two independent reviewers (HH and MC) systematically coded the transcript and assigned responses to the predefined themes. Any discrepancies were reviewed by a third, experienced researcher (PT) until an agreement was reached. Beyond categorization, the analytical process involved interpretive synthesis across coded segments to identify patterns and their broader implications within and across themes. Following this process, the initial 10 themes were refined and synthesized by merging overlapping content and eliminating redundancies, yielding the final set of themes used to structure the results.

XSL•FO
RenderX

Table . Focus group agenda and the theme codebook.

Theme	Subthemes or a priori codes	Typical indicators
Clients and referrals	• Primary target population • Referral sources and pathways	• "Mostly low-income seniors" • "Hospital social workers send us referrals"
Assessment practices	• Standard checklist or protocol • Assessor roles (occupational therapist, contractor, and staff) • Data sources (self-report, observation, and performance demonstration) • Scope (entire home vs selected areas)	• "We observe meal prep tasks" • "We skip bedrooms unless requested"
Time and workflow	• Typical timeline (assessment to completion) • Factors that lengthen or shorten the process	• "Funding approval adds 3 weeks"
Common modifications	• Frequently delivered changes • Driven by funding rules or client priorities	• "Grab bars, handrails" • "Ramp only if grant covers it"
Current use of smart home technology and assistive technology	• Technologies already installed • Decision authority (client, assessor, or others) • Staff and contractor training • Client and caregiver receptiveness • Reasons for not using technology (cost, training, and funding compatibility)	• "We add voice-activated lights when caregivers ask"
Perceived benefits	• Promoting independence and safety • Injury prevention and fall reduction • Caregiver peace of mind • Staff workload reduction and service efficiency	• "Sensors alert us before incidents"
Barriers and concerns	• Homeowner resistance or stigma • Digital literacy and technology anxiety • Data privacy and security concerns • Installation complexity and maintenance burden • Cost and funding incompatibility	• "Clients fear being watched" • "Wi-Fi is unreliable"
Proposed solutions and facilitators	• Training and education strategies • Funding or cost-sharing ideas • Technical support models	• "Partner with a local tech club for setup help"
Implementation pathways	• When and where to integrate technology (assessment stage, standard package, or add-on) • Impact on daily practice, coordination, and documentation	• "Add a 'smart tech' checkbox to assessment form"
Organizational capacity and change	• Staff training needs • New roles or expertise required • Workflow restructuring • Leadership buy-in • External partnerships (technology vendors, universities, and aging services)	• "Need a dedicated tech coordinator"

Ethical Considerations

This study was reviewed and granted an exempt determination by the University of Pittsburgh Institutional Review Board (STUDY24120137). Verbal informed consent was obtained from all participants at the beginning of the focus group session, during which they were informed of the study's purpose, the voluntary nature of participation, and that the session would be audio recorded.

Results

Characteristics of Participants

A total of 11 participants from the 3 OAHMP grantee organizations attended the focus group interview. Organizational

XSL•FO
RenderX

roles of the participants are shown in Table 2. Organizational identifiers are not reported at the individual level to protect participant confidentiality, given the small size of the participating organizations and the risk of identification through role-organization combinations.

Table . Professional roles of the participants (N=11).

Roles	Participant, n (%)
President or chief executive officer	1 (9.1)
Program manager or director	3 (27.3)
Grant administrator	1 (9.1)
Occupational therapist	1 (9.1)
Certified Aging-in-Place Specialist	4 (36.4)
Home repair lead	1 (9.1)

Theme 1: Client Characteristics, Referral Pathways, and Technology Readiness for Home Modification

Participants from the 3 OAHMP grantees reported that they primarily served older adults aged 60 years and older, referred by local agencies or partner organizations, or self-referred. Assessments were typically led by occupational therapists, CAPS, or trained occupational therapy students. Notably, participants mentioned that home assessments sometimes involved the use of software, such as Home for Life and REDCap (Vanderbilt University), as well as informal observation or discussion [24,25].

Timelines from assessment to completed modifications varied among grantees, influenced by factors such as funding availability, staffing, contractor schedules, and program structure. Most participants indicated that their projects could be delayed beyond the initial estimates, even if assessments were completed promptly.

> *So you guys at any given time are probably working on 3 to 5 different projects at the same time. So again, sometimes it's just a capacity issue in terms of who is next in line, who has all their paperwork in here, and then which site team is available to go do that whole (renovation).* [Org2]

Regarding the major applications for home modifications, agreed upon by participants from various grantee programs, traditional home modifications—such as grab bars, bathroom safety features, and entry ramps—were still identified as the core service to enhance clients' safety, mobility, and accessibility. This continued reliance on traditional modifications, alongside the limited initial integration of smart technologies, reflects a broader pattern of cautious technology readiness among both clients and grantee organizations—shaped by clients' age profiles, varying levels of digital familiarity, and the constraints of existing program structures.

> *These are things that our staff member can accomplish. So they are comfortable. High toilets, grab bars, handrails, tub cuts, solar powered security lights, levered door handles, levered faucets.* [Org2]

Theme 2: Current Integration of Smart Home Technologies and Perceived Benefits

Several participants representing their OAHMP grantee programs reported that they had incorporated smart home devices into their modification projects, observed smart home technologies used by homeowners, and sometimes collaborated with contractors during installation. In addition, staff from one organization cited their indefinite delivery/indefinite quantity contract as a mechanism for formally recommending smart home devices.

> *The Amazon Echo, the Echo Show, the Echo dot, the Ring doorbell, and then the smart plugs and we also have on our IDIQ, which is what we use...for a lot of you know, items that we recommend a lot of the time. Some of the stuff on there does also have some smart home technology that can be recommended specifically by the occupational therapist, because they don't typically look at the volunteer services as often as the CAPS do so they do have that option to add it in there if they feel a contractor should put it in.* [Org1]

> *But we have had lots of requests for smart technology from homeowners. And so that's I'm glad I'm working with and listening to you all, because that's something that we're hearing more and more of; smart technology. And we need to figure out how to incorporate that better in our programs.* [Org3]

Most installations included Amazon Echo devices, Ring doorbells and cameras, smart plugs, and motion-activated lights. This pattern of selective adoption suggests that grantees prioritized devices that were low cost, easy to install, and directly addressed safety and emergency response needs, reflecting a pragmatic, needs-driven approach to technology integration within the constraints of existing program structures. These devices were valued for enhancing security, facilitating emergency response (eg, voice-activated 911 calls after a fall), and supporting independent living. Interestingly, 1 participant suggested additional applications of smart home technologies, such as detecting pets and wildlife and monitoring air quality for clients in rural or hazard-prone areas, which were not generally considered goals of incorporating those technologies into the programs.

XSL•FO
RenderX

> *One of our cases was really having 8 cats, and like raccoons like, you know, not even knowing that they're up in the top, because our older adults aren't traveling up to their second floor any longer. And so I don't know just thinking creatively to toss it out there...Is there any smart tech to help with that? Particularly when, you know, even if they are animals that are pets, there's that like environmental hazard for the home as well.* [Org2]

Theme 3: Barriers to Adoption and Use of Smart Home Technologies

As OAHMP grantees, participants identified multiple barriers that limited the uptake and sustained use of smart home technologies by older adults. These included older adults' digital literacy and confidence, proficiency in using smartphones and devices, concerns about privacy and data security, and the burden of device maintenance.

> *Do people have the wherewithal to even know what that is, and how it functions, and how to use it? Yes, some people do not have smartphones. They got the old flip phone, or they have their landlines. And okay, we would do this as a modification for them. But is it really gonna work for them?* [Org3]

> *I feel like we need to consider that you know all of us like, what's the maintenance plan and where's the education for when they need to replace that battery? So it doesn't just become something that you know is now a trip hazard, you know, and potentially isn't actually helping them, but could potentially hurt them, since it's not functioning as intended. So there is maintenance associated with this.* [Org2]

Additionally, contextual factors frequently hinder the adoption of smart home technologies. Older adults living in rural areas often face unreliable internet connectivity and limited geographical access to certain services, which restricts the use of specific devices.

> *From where I'm based. It's very rural, and our bandwidth is sometimes shoddy at best, so it's very, um, sometimes that can be kind of a huge obstacle to be able to incorporate the smart home technology.* [Org1]

Theme 4: Adoption and Implementation Strategies

For programs that have introduced smart home technologies, participants from those programs stated that current strategies to improve adoption include providing hands-on technical assistance—often through partnerships with university occupational therapy programs—and developing easy-to-use technology manuals for older adults. Interestingly, participants noted that their partnerships with local universities were viewed as mutually beneficial: older adults received guidance and support during the adoption process, while students gained hands-on learning experiences in aging in place and assistive technology.

> *They are a little bit reluctant because they are older adults and they don't typically like to learn about the technology. But that's kind of why we have the students go in because they go over it with them like very detailed, set up the accounts for them and everything, so that they don't feel overwhelmed. And they actually also leave them a number that they can call whenever they have questions regarding any of the apps that they use.* [Org1]

Participants also recommended choosing devices that require minimal maintenance and can operate with limited internet connectivity.

> *I feel like we need to consider that you know all of us like, what's the maintenance plan and where's the education for when they need to replace that battery? So it doesn't just become something that you know is now a trip hazard, potentially and isn't actually helping them, but could potentially hurt them, since it's not functioning as intended. So there is maintenance associated with this.* [Org2]

On the other hand, participants emphasized that the timing of adopting smart home technologies should be flexible—either during the initial assessment, as part of a standard package, or as optional add-ons later in the home modification process—to address older adults' specific needs. Some older adults initially declined certain devices but became more receptive as their needs evolved, suggesting the value of a progressive adoption approach.

> *We do find that many times any type of recommendation that is made, if declined, is then requested again once a fall or something terrible, unfortunately, happens. So I mean… we do currently implement our smart home technology at each different touch point.* [Org1]

Theme 5: Organizational Capacity and Change Readiness

All participants from various OAHMP grantee organizations agreed that integrating smart home technologies into home modification programs required organizational adaptation. Unlike client-facing adoption strategies, this organizational adaptation focused on building internal capacity—equipping staff, contractors, and volunteers with the knowledge and resources needed to deliver and sustain technology-enhanced services. This included training staff and contractors on device installation and troubleshooting, adding technical support roles, and fostering leadership buy-in. External partnerships with technology vendors, universities, and aging service providers were also seen as essential.

> *We are excited to increase the number of smart home tech devices that we're able to offer. But along with that you know, we would hope that we can coordinate more intensive trainings for both ourselves as staff and our CAPS, but also like to train the trainers, so that we can adequately train our volunteers. Specifically our student volunteers to be able to train the older adults and their caregivers to be able to use the devices appropriately and not cause them more*

XSL•FO
RenderX

> *stress and frustration in trying to use the devices that you know we're providing them.* [Org1]
>
> *Our volunteers perform these modifications. I think we may need some training for the volunteers, just what these smart technologies are, and how they can integrate those in the modifications, and how they can help a homeowner essentially set them up and make them work. So it would be some training that we would need for both ourselves and our and our volunteers who are doing this work.* [Org3]

Policy-wise, 1 participant emphasized that addressing funding limitations, staffing shortages, and technical expertise gaps was crucial for the sustainable implementation of the initiative at scale.

> *And then, you know, we're nonprofit, so more funding to not only support the devices, because, as (one participant) mentioned, they are costly, but also the operations and capacity to roll out and maintain the program, so that we're able to provide that staff support and training opportunities for the volunteers.* [Org1]

Discussion

Principal Results and Policy Implications

Findings from this focus group revealed that, while certain devices—particularly those supporting safety and emergency response—have begun to be integrated into home modification services for older adults, the broader implementation of smart home–equipped modifications remained limited. Barriers emerged across three levels: (1) older adults' perceptions and digital readiness, (2) physical and contextual environmental constraints, and (3) programmatic factors, including policies and budget limitations that shape service delivery.

With an aging population, a lack of accessible, age-friendly US housing stock, and a need for cost-effective solutions to support aging in place, demand for home modifications is likely to increase in the coming years. Incorporating smart home technology into evidence-based home modification programs, such as CAPABLE and OAHMP, can provide a relatively low-cost and high-impact pathway to enhance opportunities for aging in place. Pairing smart home technology with home modifications can provide a variety of benefits, including (1) enhancing mobility, communication, social engagement, safety, independence, and health, (2) compensating for functional limitations, (3) reducing caregiver burden, and (4) preventing injurious falls. Implementing smart-home-enhanced modifications within existing evidence-based home modification programs, such as OAHMP, that support low-income older adults and those in underserved communities can broaden access to impactful, cost-effective services that promote aging in place.

The 3-level barrier structure identified in this study—individual, contextual, and organizational—is consistent with findings from the smart home adoption literature. Digital literacy limitations, privacy concerns, and the burden of device maintenance among older adults have been consistently reported in prior studies [16]. However, our findings extend this literature in 2 important aspects. First, rural internet unreliability emerged as a particularly salient contextual barrier in home modification programs, where clients are often low-income and geographically isolated—a population underrepresented in existing smart home adoption research. Second, organizational-level barriers, including staff training gaps, funding constraints, and leadership buy-in, represent a distinct and underexplored dimension of implementation challenges that go beyond individual user acceptance [26]. These findings suggest that addressing smart home technology adoption in home modification programs requires a multilevel implementation strategy rather than interventions focused solely on end user behavior. Emerging pilot programs are demonstrating the feasibility of such integrated approaches for low-income aging populations [27].

The implementation strategies identified by participants—including university partnerships with reciprocal learning benefits, progressive technology introduction across care touchpoints, and hands-on training—align with emerging evidence on effective support for technology adoption among older adults. Programs pairing students with low-income older adults in community settings have demonstrated strong potential to reduce the digital divide through a reciprocal model in which older adults receive personalized guidance while students gain practical experience. Similarly, introducing technology progressively at multiple touchpoints across the care continuum [28]—rather than as a one-time intervention—aligns with best practices in technology acceptance frameworks that emphasize meeting users at their current readiness level [20-23,29]. These findings reinforce the value of embedding adoption support within existing program structures rather than treating technology training as a separate intervention.

Smart Home Technology Selection Criteria

To guide the selection of commercially available smart home devices for integration into existing home modification services, we derived practical criteria from participant responses during the focus group, organized thematically through the analytical process, and further developed these criteria based on established usability frameworks, including the System Usability Scale [30]. These criteria are organized across 5 perspectives: installation, usability, accessibility, sustainability, and security and privacy (Table 3).

Devices should be easy to install so they do not impose unnecessary burdens on contractors or disrupt home environments. Usability is essential; older adults require straightforward onboarding and intuitive operation without repeated training. Technologies should directly improve safety, mobility, or accessibility rather than complicate daily living. Notably, accessibility matters—interfaces must be clear, with indicators that accommodate sensory needs, and vendor technical support should be available to reduce reliance on program staff. Sustainability is critical for program feasibility, favoring devices that are durable, low maintenance, and free of ongoing costs. Finally, devices must safeguard both physical and digital safety by minimizing injury risks and ensuring strong data privacy protections.

XSL•FO
RenderX

Table . Criteria for selecting smart home devices.

Perspective	Criteria
Installation	• Easy to install with minimum home modifications
Usability	• Ease of onboarding • Ease of use • Easy adoption • No memory interaction • Improves safety, accessibility, or mobility
Accessibility	• Interface design • Clear indicators • Technical support or customer service
Sustainability	• Reasonable cost • Low maintenance requirements • No subscription fees
Security and privacy	• Avoid risks of injury (eg, strangulation and asphyxiation) • Avoid risks of injury (hazardous surfaces) • Protect private information

To our knowledge, standardized criteria for selecting commercially available smart home devices specifically for integration into home modification programs have not been previously proposed in the literature. Prior systematic reviews have called for clearer classification standards and technology selection guidance to align device functions with user needs in aging in place contexts [18,31]. The 5-domain framework proposed here represents an initial step toward filling this gap and warrants further validation through broader stakeholder input and empirical testing.

Limitations

This study has several limitations. First, only 3 OAHMP subawardee organizations participated, limiting the breadth of perspectives and potentially underrepresenting the diversity of all grantees across the program. Second, because participating organizations may differ in size, geographic location, or client characteristics from other OAHMP awardees, the views expressed may not be representative of the program as a whole. Furthermore, detailed organizational characteristics could not be reported at the organizational level to protect the confidentiality of participating organizations; however, purposive sampling ensured representation across rural, suburban, and national program contexts, as described in the Methods section.

From a design perspective, using a single 1.5-hour focus group session may have limited the depth of discussion and introduced dynamics such as unequal participation or social desirability bias. The analysis used a deductive coding approach based on the discussion agenda, which ensured alignment with study aims but may have limited the emergence of unanticipated or novel themes. Additionally, focus group participants represented a variety of roles within each home modification program. It is possible that the professional background of participants (eg, occupational therapist vs grant administrator) influenced perceptions of facilitators and barriers.

Conclusions

Overall, this study underscores the potential of smart home technologies to complement traditional home modifications and enhance aging in place outcomes. Clear technology selection criteria to inform the development of a standard smart home technology package were developed. Barriers to the sustainable implementation of this initiative include older adults' digital literacy, privacy concerns, and organizational capacity. Policy support and resource allocation will be crucial to increasing adoption rates, reducing inequities, and enhancing the impact of federally funded home modification programs. Future research should focus on validating the proposed selection criteria through broader stakeholder input, curating evidence-based device packages for integration into home modification programs, and evaluating the implementation and outcomes of smart home–enhanced modifications across diverse aging populations.

Acknowledgments

The authors thank the staff and program personnel of the participating Older Adults Home Modification Program grantee organizations for their time, expertise, and willingness to share their experiences. Their contributions were essential to this work.

Claude (Anthropic) Sonnet versions 4.5 and 4.6 were used during manuscript preparation for grammar checking and language editing only. The intellectual content, including all study design, data collection, analysis, interpretation, and original writing, was developed by the research team. All AI-assisted edits were reviewed and approved by the authors prior to submission.

XSL•FO
RenderX

Funding

The work that provided the basis for this publication was supported by funding under an award with the US Department of Housing and Urban Development. The substance and findings of the work are dedicated to the public. The authors and publisher are solely responsible for the accuracy of the statements and interpretations contained in this publication. Such interpretations do not necessarily reflect the official views of the government.

Data Availability

The focus group transcript data that support the findings of this study are not publicly available. Participant confidentiality cannot be ensured if the transcript is shared because the combination of organizational context and professional role would risk identifying individual participants, given the small size of the participating organizations. Reasonable requests for further information may be directed to the corresponding author.

Authors' Contributions

HH and PT contributed to the study's conceptualization and methodology. HH and MC conducted the focus group, collected data, and led the formal analysis and data curation. PT provided supervision and served as the third reviewer to resolve coding discrepancies. HH, MH, and PT led the writing of the original draft. All authors contributed to review and editing of the manuscript. PT secured funding and provided resources for the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Focus group discussion guide used in the qualitative focus group study with Older Adults Home Modification Program grantee organizations.
[DOCX File, 23 KB - aging_v9i1e89729_app1.docx]

References

1. 2023 profile of older Americans. U.S. Department of Health & Human Services. URL: https://acl.gov/sites/default/files/Profile%20of%20OA/ACL_ProfileOlderAmericans2023_508.pdf [accessed 2026-07-01]
2. Table S1810: disability characteristics. United States Census Bureau. URL: https://data.census.gov/table?q=S1810 [accessed 2026-07-01]
3. Gustavson AM, Vincenzo J, Miller MJ, et al. Equitable implementation of innovations to promote successful aging in place. J Am Geriatr Soc 2023 Feb;71(2):683-688. [doi: 10.1111/jgs.18177] [Medline: 36522303]
4. New AARP report: majority of adults 50-plus want to age in place, but policies and communities must catch up. : AARP; 2024 URL: https://press.aarp.org/2024-12-10-New-AARP-Report-Majority-Adults-50-plus-Age-Place-Policies-Communities-Catch-Up [accessed 2026-07-01]
5. Lawton MP, Nahemow L. Ecology and the aging process. In: Eisdorfer C, Lawton MP, editors. The Psychology of Adult Development and Aging: American Psychological Association; 1973:619-674. [doi: 10.1037/10044-020]
6. Housing America's older adults. Joint Center for Housing Studies of Harvard University. 2023. URL: https://www.jchs.harvard.edu/sites/default/files/reports/files/Harvard_JCHS_Housing_Americas_Older_Adults_2023_Revised_040424.pdf?__cf_chl_f_tk=s7wVDIuMVXU.wRHJ_vNXjW9fRojEXLPHN_d3.q4w3NU-1782896650-1.0.1.1-JIBgHMNaAoGmMSqkW7WOd89K4BpBsX11HVE5AFoGBVU [accessed 2026-07-01]
7. Szanton SL, Leff B, Li Q, et al. CAPABLE program improves disability in multiple randomized trials. J Am Geriatr Soc 2021 Dec;69(12):3631-3640. [doi: 10.1111/jgs.17383] [Medline: 34314516]
8. Campani D, Caristia S, Amariglio A, et al. Home and environmental hazards modification for fall prevention among the elderly. Public Health Nurs 2021 May;38(3):493-501. [doi: 10.1111/phn.12852] [Medline: 33340382]
9. Abraham J. The cost efficiency of home modifications to reduce health care costs. Innov Aging 2021 Dec 17;5(Supplement_1):558. [doi: 10.1093/geroni/igab046.2144]
10. Home modification and repairs for older adults: challenges and opportunities for State Units on Aging. Advancing States. 2020. URL: https://www.advancingstates.org/hcbs/article/home-modification-and-repairs-older-adults-challenges-and-opportunities-state-units [accessed 2026-07-01]
11. Tenenbaum L, Kimmel S. Overcoming inertia to jumpstart home modifications for aging in place. Generations. 2020 Aug 14. URL: https://generations.asaging.org/home-modifications-aging-place-remodeling/ [accessed 2026-07-01]
12. Older adults and adults with disabilities: federal programs provide support for preventing falls, but program reach is limited. : United States Government Accountability Office; 2022 URL: https://www.gao.gov/assets/730/721979.pdf [accessed 2026-07-01]

XSL•FO
RenderX

13. Wiseman JM, Stamper DS, Sheridan E, Caterino JM, Quatman-Yates CC, Quatman CE. Barriers to the initiation of home modifications for older adults for fall prevention. Geriatr Orthop Surg Rehabil 2021;12:21514593211002161. [doi: 10.1177/21514593211002161] [Medline: 33868764]
14. How much does it cost to make my home wheelchair accessible? Angi. 2025. URL: https://www.angi.com/articles/how-much-does-it-cost-make-my-home-handicap-accessible.htm [accessed 2026-07-01]
15. Reddy AL, Beatley N, Breysse J, Dixon S, Wilson J, Bloom C. Evaluation of the HUD Older Adult Home Modification Grant Program: cohort 1 interim report. : U.S. Department of Housing and Urban Development, Office of Policy Development and Research; 2024 URL: https://www.huduser.gov/portal/publications/Evaluation-of-the-HUD-Older-Adult-Home-Modification-Grant-Program-Cohort-1-Interim-Report.html [accessed 2026-07-01]
16. Tian YJ, Felber NA, Pageau F, Schwab DR, Wangmo T. Benefits and barriers associated with the use of smart home health technologies in the care of older persons: a systematic review. BMC Geriatr 2024 Feb 14;24(1):152. [doi: 10.1186/s12877-024-04702-1] [Medline: 38355464]
17. Tural E, Lu D, Austin Cole D. Safely and actively aging in place: older adults' attitudes and intentions toward smart home technologies. Gerontol Geriatr Med 2021;7:23337214211017340. [doi: 10.1177/23337214211017340] [Medline: 34095352]
18. Cha SM. A systematic review of home modifications for aging in place in older adults. Healthcare (Basel) 2025 Mar 27;13(7):752. [doi: 10.3390/healthcare13070752] [Medline: 40218050]
19. Rogers EM. Diffusion of Innovations, 5th edition: Free Press; 2003.
20. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
21. Gibson WJ, Brown A. Working with Qualitative Data: SAGE Publications; 2009. [doi: 10.4135/9780857029041]
22. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
23. Yang HJ, Lee JH, Lee W. Factors influencing health care technology acceptance in older adults based on the technology acceptance model and the unified theory of acceptance and use of technology: meta-analysis. J Med Internet Res 2025 Mar 28;27:e65269. [doi: 10.2196/65269] [Medline: 40153796]
24. Home for Life Design. URL: https://www.homeforlifedesign.com/ [accessed 2026-07-01]
25. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
26. Dermody G, Wadsworth D, Dunham M, Glass C, Fritz R. Factors affecting clinician readiness to adopt smart home technology for remote health monitoring: systematic review. JMIR Aging 2024 Dec 5;7:e64367. [doi: 10.2196/64367] [Medline: 39012852]
27. Fried J, Hughes M, Singh P, Toto P. Home Safe and Smart: a pilot program integrating smart technology and home modifications for aging in place. Innov Aging 2025 Dec 31;9(Supplement_2):igaf122.1883. [doi: 10.1093/geroni/igaf122.1883]
28. Miller LM, Callegari RA, Abah T, Fann H. Digital literacy training for low-income older adults through undergraduate community-engaged learning: single-group pretest-posttest study. JMIR Aging 2024 May 14;7:e51675. [doi: 10.2196/51675] [Medline: 38599620]
29. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
30. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry: CRC Press; 1996:189-194. [doi: 10.1201/9781498710411]
31. Ma C, Guerra-Santin O, Mohammadi M. Smart home modification design strategies for ageing in place: a systematic review. J Hous Built Environ 2022 Jun;37:625-651. [doi: 10.1007/s10901-021-09888-z]

Abbreviations

CAPABLE: Community Aging in Place–Advancing Better Living for Elders
CAPS: Certified Aging-in-Place Specialist
OAHMP: Older Adults Home Modification Program

XSL•FO
RenderX

Edited by M Mardini; submitted 16.Dec.2025; peer-reviewed by J Kang, M Scott; revised version received 23.May.2026; accepted 18.Jun.2026; published 16.Jul.2026.

Please cite as:
Hu H, Clarke M, Hughes M, Sharma P, Fried J, Singh P, Toto P
Smart Home Technology Integration in Home Modification Programs Serving Older Adults: Focus Group Study With Program Grantees
JMIR Aging 2026;9:e89729
URL: https://aging.jmir.org/2026/1/e89729
doi:10.2196/89729

© Haomin Hu, Moriah Clarke, Meredith Hughes, Palash Sharma, Jack Fried, Portia Singh, Pamela Toto. Originally published in JMIR Aging (https://aging.jmir.org), 16.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Feasibility of a Co-Designed Personalized Cognitive Reablement Program Delivered via Digital Voice Assistants for Older Adults Living With Early-Stage Dementia or Mild Cognitive Impairment: Pilot Randomized Controlled Trial

Tracy Yiannis[1], BSc (Hons), MHN; Helen Macpherson[1], PhD, BA, BSc (Hons); David Scott[1,2], BHumMvt (Hons), PhD; Michele Callisaya[3,4], PhD, BApplSc; Celeste Chapple[1], PhD, GCertHealthResPrac, BFoodSc&Nutr; Costas Glavas[1,2], PhD; Caitlyn Gourlay[1], PhD, BSc (Hons), GCertUniTeach, MClinPsych; Ralph Maddison[1], PhD; Eugene Gvozdenko[5], PhD; Paul Jansons[1,2], PhD, MAppSc, BAppSc

[1]Deakin Institute for Physical Activity and Nutrition, Deakin University, 221 Deakin Hwy, Burwood, Victoria, Australia

[2]School of Clinical Sciences at Monash Health, Monash University, Melbourne, Victoria, Australia

[3]Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia

[4]Peninsula Clinical School, School of Translational Medicine, Monash University, Melbourne, Victoria, Australia

[5]Great Australian Pty Ltd, Keysborough, Australia

Corresponding Author:
Tracy Yiannis, BSc (Hons), MHN
Deakin Institute for Physical Activity and Nutrition, Deakin University, 221 Deakin Hwy, Burwood, Victoria, Australia

Abstract

Background: Home-based reablement programs have improved everyday functioning for people living with early-stage dementia or mild cognitive impairment (MCI), but the substantial time commitment required for delivery by health professionals has created a barrier to widespread implementation. Digital technology could facilitate delivery, but traditional devices such as smartphones and tablets present usability challenges for this population. We have co-designed a cognitive reablement program, VOICE-COG, that can deliver personalized interventions broadcast via voice-controlled digital voice assistants (DVAs).

Objective: This pilot nonblinded randomized controlled trial examined the feasibility and potential effectiveness of VOICE-COG for older adults living with dementia or MCI to guide the design of a full-scale trial.

Methods: Older adults (aged 60 - 85 years) with a self-reported diagnosis of dementia or MCI (N=30) were recruited via Facebook or the StepUp-for Dementia Research database and randomized to either a DVA intervention (n=15) or control (n=15) group for 12 weeks. Each participant randomized to the intervention group and their care partner met with a psychologist via Zoom to select 2 or 3 personalized goals from the VOICE-COG program. The personalized program was input into the software (BuddyLink) by researchers for broadcast via a DVA (Amazon Alexa Show 8). Participants in the control group received enhanced usual care, consisting of treatment as usual, plus weekly informational emails. Primary feasibility outcomes were recruitment, retention, device-recorded adherence, and the System Usability Scale (SUS). Secondary outcomes were self-reported by the participant and care partner via web-based Qualtrics forms and included the Goal Attainment Scale–Light (GAS), changes in participant mood, cognition, and everyday functioning, and carer burden.

Results: The recruitment rate was 86% (30/35) and the retention rate was 97% (29/30), with 1 participant in the intervention group lost to follow-up. Adherence to the VOICE-COG program ranged from 29% to 100% (mean 70%, SD 17%). The VOICE-COG program was rated as marginally acceptable (mean SUS score 56.3, SD 18.1). After 12 weeks, 60% (9/15) of the intervention group reported having met or exceeded their goals (GAS-Light t-score ≥50; mean 57.6, SD 15.4). The intervention group had a significantly higher increase in Multifactorial Memory Questionnaire–Strategy subscale score compared with control, mean difference 5.0 (95% CI 0.03-10.1, P=.049), representing a medium effect size (d=0.55). There were no other significant differences between groups.

Conclusions: This co-designed VOICE-COG program, delivered via DVAs, was feasible for older adults living with early-stage dementia or MCI, resulting in most participants meeting their reablement goals and increases in the use of memory strategies. System usability issues should be addressed in future adequately powered trials.

Trial Registration: ANZCTR ACTRN12623000898651; https://tinyurl.com/3vkzrwmc

(JMIR Aging 2026;9:e96922) doi:10.2196/96922

XSL•FO
RenderX

KEYWORDS

digital voice assistants; reablement; assistive technology; cognitive rehabilitation; dementia; mild cognitive impairment; digital health; feasibility; Amazon Alexa; eHealth; older adults

Introduction

Dementia is a growing public health concern worldwide due to its increasing prevalence and the associated burden of disability and dependency [1]. By 2030, the number of people living with dementia is predicted to reach 85 million, driven by the aging population and higher incidence rates in older adults [2]. As loss of cognitive functioning progresses, people living with dementia become increasingly dependent on informal and formal care, which was estimated to cost US $1.3 trillion globally in 2019 [3,4]. With limited pharmacological options to manage symptoms, studies have focused on nonpharmacological interventions to maintain independence at home and reduce the associated cost of care [5,6].

Previous research in reablement, also referred to as cognitive rehabilitation, has found that people living with dementia experience significant improvements in goal attainment related to daily functioning [7-9], while care partners reported enhanced well-being and coping outcomes [10,11]. Evidence from these studies suggests people living with mild or early-stage dementia experience greater benefits, with dementia severity a barrier to the collaborative goal-setting process [8,12]. However, the substantial time commitment for weekly home-based delivery from a qualified health professional (typically a nurse, occupational therapist, or psychologist), together with travel time and specialized training, has presented significant barriers to widespread implementation [13,14]. Technology could facilitate the remote delivery of reablement programs to overcome these barriers, increasing the availability of goal-oriented reablement programs for people living with dementia [15].

Recent studies have found telehealth delivery of reablement interventions just as effective as face-to-face home-based programs, eliminating the travel time requirement for health professionals [16,17]. This approach, however, still required a substantial time commitment from health professionals, as observed in the adapted Care of People with dementia in their Environments (COPE) [11] trial, which required 2 days training plus eight 1-hour sessions per person [17]. To date, no studies have determined if technology could improve support further by offering reablement self-management and remote monitoring from health professionals. A major issue in digital health research for this population concerns device usability, which is limited by the cognitive burden of remembering passwords and program navigation, in addition to age-related dexterity and vision impairments [18].

Digital voice assistants (DVAs) with embedded speakers and display screens that can broadcast personalized programs and interpret voice commands for navigation may offer a possible solution [19,20]. Existing evidence on DVA use for people living with dementia is confined to exploratory qualitative designs or technical usability evaluations [21-24]. These studies have suffered from methodological limitations such as case study designs, small sample sizes, and laboratory settings, limiting generalization to community-dwelling older adults living with dementia. Furthermore, while these studies provide evidence of daily activity reminders, cognitive stimulation, and social interaction, no studies have evaluated the use of DVAs to deliver a holistic, goal-based reablement program.

To address this need, we developed VOICE-COG, a cognitive reablement program delivered via DVA following a remote goal-setting meeting with a health professional. VOICE-COG is delivered to each participant's DVA in their homes using BuddyLink (Great Australian Pty Ltd) proprietary software [25] that broadcasts personalized health programs via audio, images, text, and video. Previous pilot studies using BuddyLink have demonstrated the feasibility of DVA to deliver personalized exercise and nutrition programs for cognitively healthy older adults [26-28]. The development and testing of VOICE-COG was expanded to include older adults living with mild cognitive impairment (MCI), given its characterization as a predementia stage with similar cognitive symptoms. Co-designed with older adults living with dementia/MCI, their care partners, and health professionals, VOICE-COG is intended to support reablement for regular leisure and ad hoc appointments, emotional regulation exercises, reminiscence activities, self-care adherence, and passive cognitive stimulation [29]. The primary aim of this pilot randomized controlled trial (RCT) was to determine the feasibility and usability of VOICE-COG for older adults living with dementia to inform a full-scale trial. Secondary aims were to assess potential efficacy on goal attainment, cognition, emotional well-being, daily activity functioning, and care partner burden.

Methods

Study Design and Randomization

This 12-week nonblinded pilot RCT examined the feasibility of VOICE-COG, a co-designed cognitive reablement program broadcast via DVA [29], with reporting guided by the Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth (CONSORT-EHEALTH) [30] (Checklist 1). Participants were recruited using convenience sampling, with eligible participants enrolled by the primary project lead. Following completion of baseline assessments, participants were block randomized (1:1), stratified by sex, to either the intervention (VOICE-COG program) or control group. The random allocation sequence was generated in advance using a computer-based randomization method in Microsoft Excel by a research assistant, who was independent of intervention delivery. The research assistant assigned participants according to the pregenerated sequence and notified the project lead of group allocation. Participants were not blinded to group allocation due to the provision of a DVA for the intervention group. Blinding was not possible for the lead researcher, who conducted a midtrial check-in with participants as required for this feasibility pilot study.

XSL•FO
RenderX

Participants and Research Setting

Eligibility criteria included older adults (aged 60 to 85 y) with a self-reported, recent (≤6 mo) clinical diagnosis of dementia and/or MCI, living anywhere in Australia with access to a residential Wi-Fi network. To guarantee an optimal internet connection for the Amazon Alexa DVA, the use of mobile data connections was not supported. Participants could be community-dwelling or in residential aged care, provided they had a care partner with at least 50% close contact during the 12-week trial and had access to a smartphone capable of making and receiving calls on an Australian network. Participants were required to be fluent in English. Digital literacy was not considered an inclusion criterion, as previous research has found prior experience does not influence adaptation for people living with dementia [31], and the voice navigation function of the DVA was expected to reduce this barrier. However, to reduce potential bias, all intervention participants were emailed personalized instructions on how to access their program via the DVA. Participants were excluded if they (1) were deemed to have difficulty communicating with a DVA due to speech or hearing problems, (2) were unwilling to be randomized, (3) planned to be away for ≥4 weeks during the 12-week intervention period, or (4) had any self-reported cognitive or physical impairment that, in the opinion of the research team, would result in the participant having difficulty interacting with the DVA or performing the VOICE-COG program safely. This included a self-reported clinical diagnosis of schizophrenia, bipolar disorder, or other psychotic disorders.

Participants were recruited via 2 methods. First, through paid targeted Facebook (Meta Platforms, Inc) advertisements, potential participants clicked an embedded hyperlink to read the plain language statement and completed an expression of interest form via Qualtrics. Second, eligible volunteers matching the study's inclusion criteria from the StepUp for Dementia Research database were contacted via phone call or email [32]. Individuals in this database have previously provided consent to be contacted for research. Interested participants were screened for eligibility via telephone and then emailed a link to the plain language and consent form on Qualtrics to provide written or electronic informed consent. Participants enrolled in the study provided details of the clinician responsible for their diagnosis, who was notified of the participant's eligibility and was given the opportunity to raise any concerns regarding participation. No further diagnostic criteria or cognitive assessment tools were used to determine eligibility.

Intervention

As detailed elsewhere [29], the development of VOICE-COG was guided by the IDEAS (Integrate, Design, Assess, and Share) framework [33], which combines co-design practice with identification of health behavior change strategies for digital interventions. The initial health professional facilitated goal-setting meeting used a modified, dementia-specific SMART (Specific, Measurable, Achievable, Relevant, and Time-bound) framework [34,35] that establishes current functioning and targeted goal attainment, aligned with previous hybrid digital health interventions [36,37]. The subsequent VOICE-COG program, broadcast via DVA, included several evidence-based behavior change techniques [36,38], such as motivational messages and preferences for social reward [39], individual tailoring with personalized audio and text prompts, and visual cues [40], in addition to instructional videos [41].

Intervention group participants selected goals from the VOICE-COG program that were personalized and broadcast via a DVA in their homes. Participants randomized to the intervention group were mailed a preregistered Amazon Echo Show 8 device ("Alexa") with instructions on how to connect the device to their home Wi-Fi. Each participant and their care partner were emailed a 1-page summary of the VOICE-COG program (Multimedia Appendix 1) in advance of meeting with a psychologist (qualified to a minimum Doctorate in Clinical Psychology with experience in motivational interviewing) via Zoom. Guided by the goal-setting reablement guide for people living with dementia [35], the psychologist assisted the participant in collaboratively setting 2 or 3 modified SMART-goals with their care partner that could be supported by the VOICE-COG program. Examples of participant goals can be found in Multimedia Appendix 2. These included (1) activities and appointments (attending regular weekly community activities and remembering appointments); (2) routine reminders (adherence to daily medication, hydration, and nutrition); (3) family and friends (writing their life story and maintaining relationships); (4) mood monitor (providing coping activities such as music and mindfulness); (5) active minds (providing cognitive stimulation through audiobooks, podcasts, news, and short videos); (6) household manager (remembering when and how to do home chores); and (7) kitchen helper (support for cooking favorite recipes). Further descriptions of how participants interact with the VOICE-COG modules can be found in Multimedia Appendix 3.

VOICE-COG used BuddyLink software to broadcast each participant's personalized program via an Alexa Skill on their Amazon Show 8 DVAs, similar to an app on a smartphone. Personalized goal-related information and preferred program broadcast timings were entered into an online Qualtrics form by the psychologist during the goal-setting meeting. Each participant's goal-related content was uploaded remotely to the BuddyLink software [32] by a research team member, who also configured program broadcast routines using deidentified individual Alexa accounts. A voice-initiated routine was added, "Alexa, Open Mind Buddy," for participants to access their program outside these times. Following the meeting, the intervention group was emailed a personalized 2-page instruction summary for the selected VOICE-COG goals. Participants who elected to write their life story (family and friends goal) received weekly emails with a link to an online Qualtrics form for data collection. Throughout the trial, all intervention participants received fortnightly emails requesting notification of any technical errors via completion of an online Qualtrics form. An optional midtrial check-in by a research team member was offered to participants to address any concerns regarding their program.

Control

Control group participants received usual care from their medical provider or aged care service, plus weekly emails with

XSL•FO
RenderX

Dementia Australia factsheets (available from their website) [42]. These were provided for ethical reasons, to ensure participants received some benefit while minimizing the potential impact on study outcomes.

Primary Outcome Measures

Feasibility

Participant recruitment, retention, and adherence rates were used to examine the feasibility of the VOICE-COG program to inform a future RCT. Recruitment rate (%) was defined as the total number of participants who provided consent at baseline divided by the total number eligible. Retention rate (%) was specified as the number of participants who completed the 12-week postintervention data collection divided by the number of participants randomized. The customized nature of the intervention allowed participants to have substantial variation in the type and number of weekly goal-related prompts. As such, adherence was defined as at least one weekly interaction with a personalized VOICE-COG goal via the DVA, as measured by a voice response to the initial dialog prompt recorded by the BuddyLink system (excluding "no" or "stop" responses). Adherence (%) was calculated as the proportion of weekly goal interactions recorded over the 12 weeks (24 or 36 possible goal interactions for participants selecting 2 or 3 goals, respectively). This measure, together with the predefined targets of ≥70% retention and ≥66% adherence to the VOICE-COG program, was based on previous evidence from DVA health intervention studies [26-28].

System Usability

Perceived usability of the VOICE-COG skill was assessed by DVA participants using the System Usability Scale (SUS) [43] form administered online via Qualtrics. The SUS uses 10 Likert-scale questions to determine subjective usability, providing a score from 0 to 100.

Secondary Outcome Measures

Goal Attainment

All data for baseline and 12-week outcomes were collected remotely via Qualtrics (unless stated otherwise). Demographic data were also collated at baseline. Participants and care partners were emailed a link to the survey including the questionnaires and informed they could complete the survey at any time throughout the week. They were advised that they could save their progress and return to it later if necessary.

DVA participants' goal attainment was measured using Goal Attainment Scale (GAS)–Light, a valid and reliable scoring system recommended for use with people living with dementia/MCI in reablement programs [35,44,45]. The intervention used a co-designed reablement program specifically developed for DVA delivery, thereby enabling administration by health professionals with minimal reablement training. Given the primary aim of this study was to establish feasibility, the control group did not receive a goal-setting-only intervention that would have required a reablement-trained health professional. The baseline goal attainment score was entered by the psychologist into the online Qualtrics goal-setting form during the goal-setting meeting (–2: no function or –1: some function). After 12 weeks, participants were emailed their personal goals and, together with their care partner, reported if they had achieved their goal and to what level: yes (+2 a lot more; +1 a little more; 0 as expected) or no (–0.5 partially achieved; –1 no change; –2 got worse). Results were entered into the GAS-Light spreadsheet [46] to calculate each participant's GAS-Light t-score, where t≥50 is interpreted as met or exceeded expectations.

Participant Questionnaires

Secondary outcome measures included differences in the following participant questionnaires (measuring mood, memory, and cognition) completed via online Qualtrics forms at baseline and at 12 weeks. The Geriatric Depression Scale (GDS) consists of 15 items rated "yes" or "no" to screen for depression in older adults and has been validated for use in older adults with mild to moderate dementia [47]. Scores of >5 suggest depression may be present; >10 is indicative of major depression. The Multifactorial Memory Questionnaire (MMQ) is a memory evaluation tool with 57 items covering 3 subscales—memory satisfaction, ability, and strategy use [48]. Participants' subjective memory experiences from the previous 2 weeks are rated using a 5-point Likert scale from 0 to 4. If there were <4 missing responses for any scale, scores were calculated using the prorated formula; for ≥4 missing items, participant scores were excluded from the analysis [48]. The Illness Cognition Questionnaire (ICQ) is an 18-item questionnaire using a 4-point Likert scale from 1 (disagree) to 4 (strongly agree) to measure the respondents' perception of their own illness with 3 subscales of acceptance, disease benefit, and helplessness [49]. For participants with 1 missing item in a subscale, the missing item was imputed using the mean of the other completed items within that subscale (provided at least 4 of 6 items were present).

Care Partner Questionnaires

At baseline and 12 weeks, care partners completed the following surveys: the Bayer Activities of Daily Living Scale (B-ADL) is a 25-item questionnaire assessing the difficulty that the person living with dementia/MCI has in performing daily activities, using a Likert scale of 0 (never) to 10 (always) [50]. The B-ADL score sums all items, excluding those rated "not applicable" or "unknown" divided by the number of items rated. The Everyday Cognition (ECog) Scale includes 39-items across 6 domains (everyday memory, language, visuospatial abilities, planning, organization, and divided attention) [51]. The Zarit Caregiver Burden 22-item questionnaire asked respondents to rate the level of burden caused using a Likert scale of 1 ("none") to 5 ("nearly always"), giving a score range of no burden (22 to 46), mild burden (47 to 55) or severe burden (56-110) [52].

Data Analysis

We recruited 30 participants (n=15 in each arm), consistent with sample size guidelines for pilot feasibility trials [53] and sufficient to provide preliminary estimates of medium effect sizes (d=0.5) to inform the design of a fully powered RCT [54].

Descriptive data for feasibility and secondary outcomes are presented as either mean and SD or frequency (%). Differences in GAS-Light t-scores for the intervention group were compared using a 2-tailed paired-sample t test. SUS scores were entered

XSL•FO
RenderX

into the web-based analysis toolkit [55] to provide a box plot for visual comparison with the adjective quartile SUS score [56]. Differences in participant and care partner questionnaires were compared between groups using analysis of covariance adjusted for baseline values, and standardized effect sizes calculated using Cohen *d*. Missing data within scales was replaced by the mean value of item responses for that subscale. Effect estimates are reported with 95% CIs and *P* values set at $P < .05$, but as a pilot feasibility trial, this study was not powered to detect any significant changes. Shapiro-Wilk tests assessed assumptions of normality and homogeneity of variance, and natural logarithmic transformations were applied where required before conducting analysis of covariance analysis. All analyses were conducted using Stata Standard Edition (version 18.0; StataCorp).

Ethical Considerations

This study was approved by the Deakin University Human Research Ethics Committee (HREC 2022-043) and registered with the ANZCTR (Australian New Zealand Clinical Trials Registry: ACTRN12623000898651). Intervention participants received an Alexa Show 8 device, which they could retain after the trial to compensate them for their time. Control participants received an Alexa Show 8 device after the 12-week trial was completed.

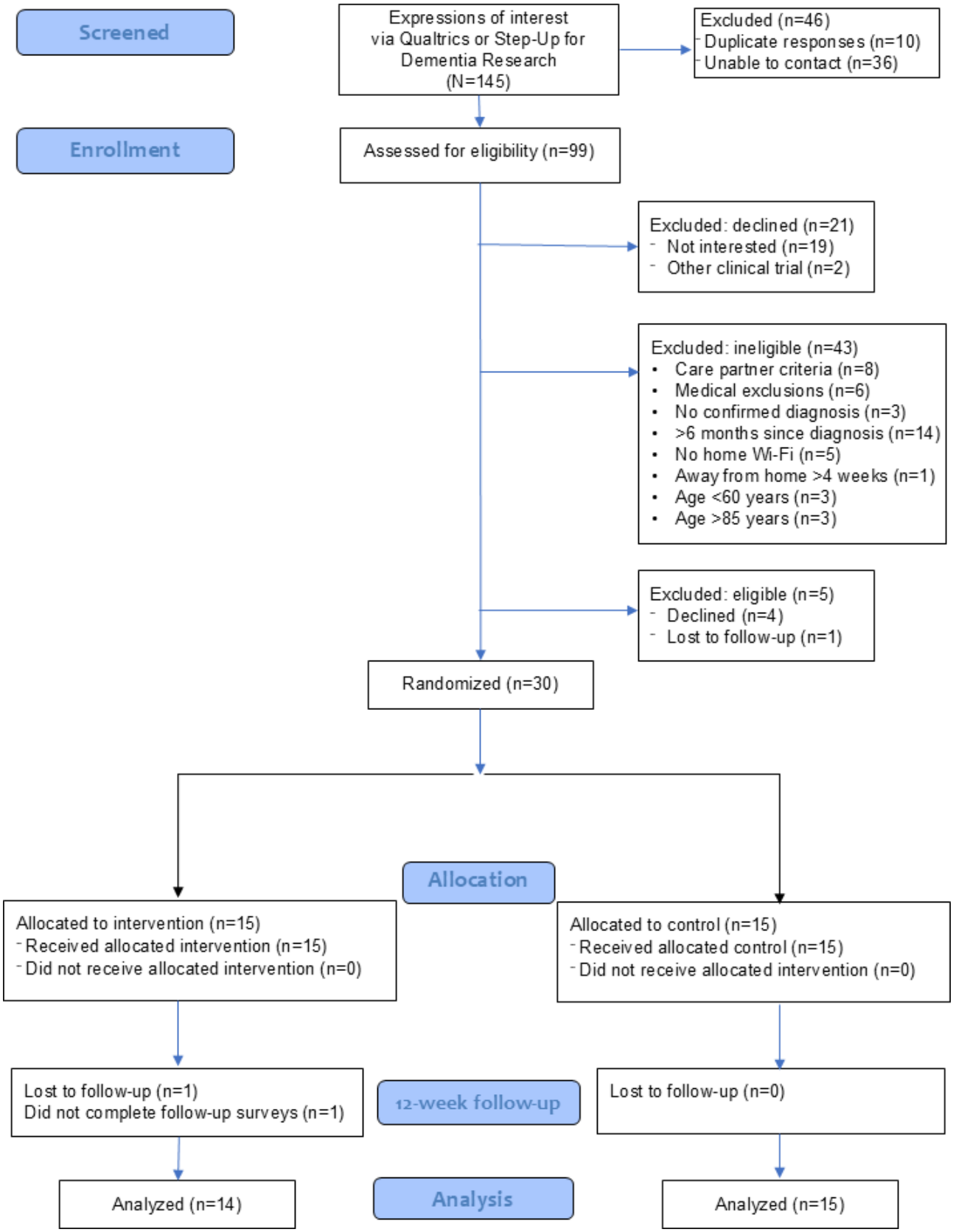
Results

Overview

Participant flow is presented in Figure 1. In total, 145 expressions of interest were received, with 99 potential participants successfully contacted and screened between September 2023 and June 2024. A total of 30 participants completed baseline assessments and were randomly allocated to the DVA group (n=15) or the control group (n=15).

XSL•FO
RenderX

Figure 1. Consolidated Standards of Reporting Trials (CONSORT) diagram showing participant flow through the 12-week VOICE-COG pilot feasibility randomized controlled trial.



Participant Demographics

Table 1 summarizes baseline demographics of the 30 participants. Participants' ages ranged from 60 to 85 years (mean 74.7, SD 5.2 y), and the majority (n=19, 63%) were women. Over half (n=18, 60%) were educated to university level, and while most (n=24, 80%) participants were retired and 13% (n=4) receiving a pension, 7% (n=2) were still employed. The vast majority (n=28, 93%) reported having at least one chronic health

XSL•FO
RenderX

condition in addition to dementia or MCI, with over half (n=17, 57%) having hypertension.

Table . Baseline participant characteristics of the DVA[a] intervention and enhanced usual care control groups (N=30).

Demographics	DVA group (n=15)	Control group (n=15)
Age (y), mean (SD)	72.9 (5.1)	76.5 (4.9)
Sex (female), n (%)	10 (67)	9 (60)
Parent's country of birth, n (%)		
Australia	11 (73)	10 (67)
Other	4 (27)	5 (33)
Highest level of education, n (%)		
Secondary or high school	3 (20)	4 (27)
Technical or further education institution	3 (20)	2 (13)
University or other higher education institution	9 (60)	9 (60)
Marital status, n (%)		
Married or de facto	13 (87)	14 (93)
Divorced	2 (13)	1 (7)
Current employment status, n (%)		
Employed or self-employed part-time	2 (13)	0 (0)
Pension (including disability or sole pension)	2 (13)	2 (13)
Retired	11 (73)	13 (87)
Medical conditions, n (%)		
Any form of cancer	2 (13)	2 (13)
Asthma	2 (13)	1 (7)
Cardiovascular disease[b]	5 (33)	4 (27)
Hypercholesterolemia (high cholesterol)	6 (40)	7 (47)
Hypertension (high blood pressure)	7 (47)	10 (67)
Osteoarthritis	2 (13)	5 (33)
Thrombosis (blood clot)	0 (0)	1 (7)
Rheumatoid arthritis	1 (7)	1 (7)
Diabetes	2 (13)	2 (13)
Other major illness[c]	7 (47)	6 (40)
Reported at least one chronic health condition	13 (87)	15 (100)

[a]DVA: digital voice assistant.

[b]Cardiovascular disease included angina, stroke, peripheral vascular disease, and heart attack.

[c]Parkinson disease, multiple sclerosis, Crohn disease, kidney disease, liver disease, transplant, congenital deafness, vertigo, Sjögren syndrome, dumping syndrome, granulomatous disease, sleep apnea, hypoparathyroidism, hypothyroidism, hypoglycemia, melanoma, dystonic tremor, Marfan syndrome, depression, ankylosing spondylitis, and grand mal epilepsy.

Primary Outcomes

Feasibility

Primary outcome measures are summarized in Table 2. The recruitment rate was 86%; that is, of the 35 eligible participants, 30 people living with dementia/MCI and their care partners provided consent to participate. The overall retention rate was 97% (n=29). Retention was 100% (15/15) for the control group and 93% (14/15) for the intervention group, with 1 participant lost to follow-up (reason unknown). Adherence to the VOICE-COG program ranged from 29% to 100%, with a mean of 70% (SD 17%), exceeding the target adherence of 66%, with no adverse events. As shown in Figure 2, mean participant adherence met or exceeded the target for all VOICE-COG program goals except Activities and Appointments. Participants nominated between 7 and 52 VOICE-COG broadcasts per week, with a mean of 12 (SD 12) weekly voice interactions recorded by the DVA, as detailed in Multimedia Appendix 4.

XSL•FO
RenderX

Table . Primary outcomes of feasibility and usability.

Outcome	Time point	DVA group (n=15)	Control group (n=15)
Recruitment, n/N (%)	N/A[a]	30/35 (86)	30/35 (86)
Retention[b], n/N (%)	12 weeks	14/15 (93)	15/15 (100)
Adherence (%), mean (SD)[c]	Overall	70 (17)	N/A
System Usability Scale (SUS), mean (SD)[d]	12 weeks	56.3 (18.1)	N/A

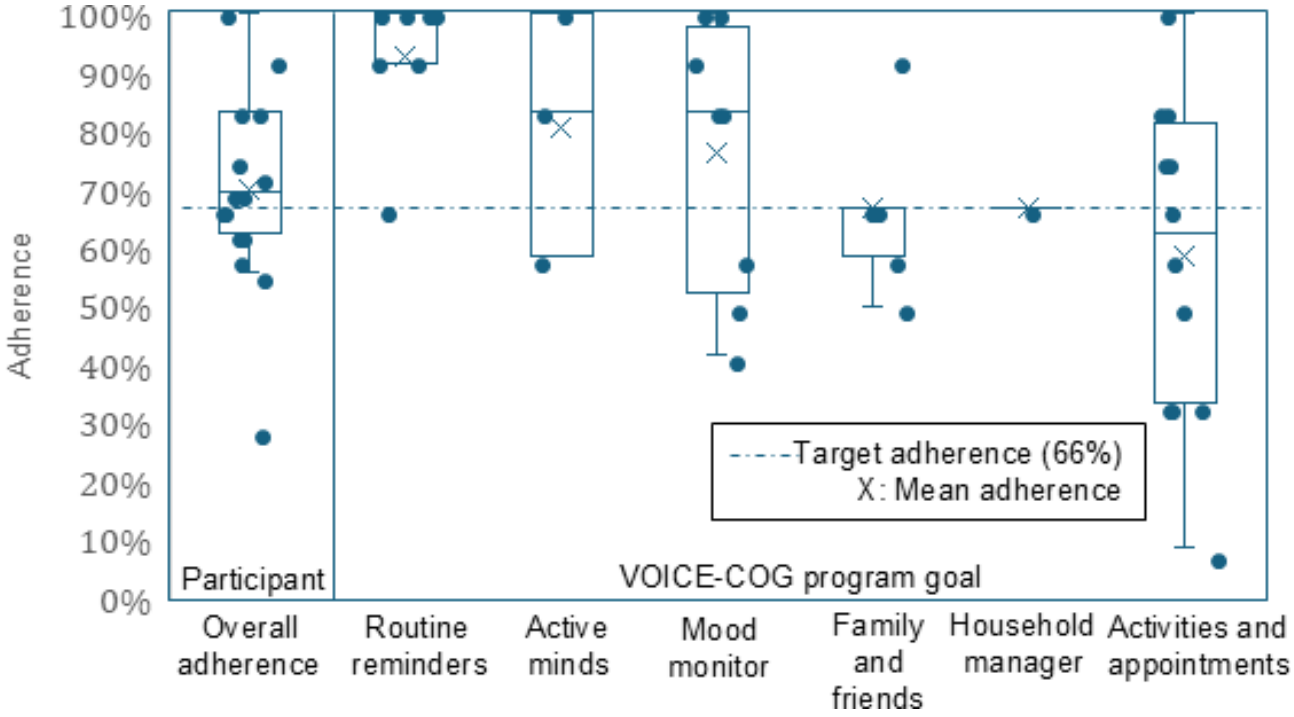
[a]N/A: not applicable.

[b]Retention was defined as participants who completed all postintervention assessments at 12 weeks.

[c]Adherence was defined as ≥1 positive response to the program goal per week (out of the number of weekly goals broadcast over 12 weeks), reported as the percentage completed.

[d]Calculated as a score from 1 to 100.

Figure 2. Box plots showing percentage adherence to the digital voice assistant–delivered VOICE-COG cognitive reablement program over 12 weeks, overall and by program goal.



System Usability

At 12 weeks, the SUS scores for DVA intervention participants indicated marginal acceptability of the VOICE-COG Program (mean 56.3, SD 18.1; range 28 - 79) [56]. The distribution of scores and responses to specific SUS questions can be found in Multimedia Appendix 5. Among the 15 participants, (n=11, 73%) agreed they would like to use the system frequently, and two-thirds (n=10, 67%) agreed the system was easy to use and that most people would learn to use the system very quickly. Just over half (n=8, 53%) believed the various functions in the system were well integrated and disagreed that the system was very cumbersome to use. However, just over half (n=8, 53%) felt they needed to learn a lot of things before they could get going with this system.

Secondary Outcomes

Goal Attainment

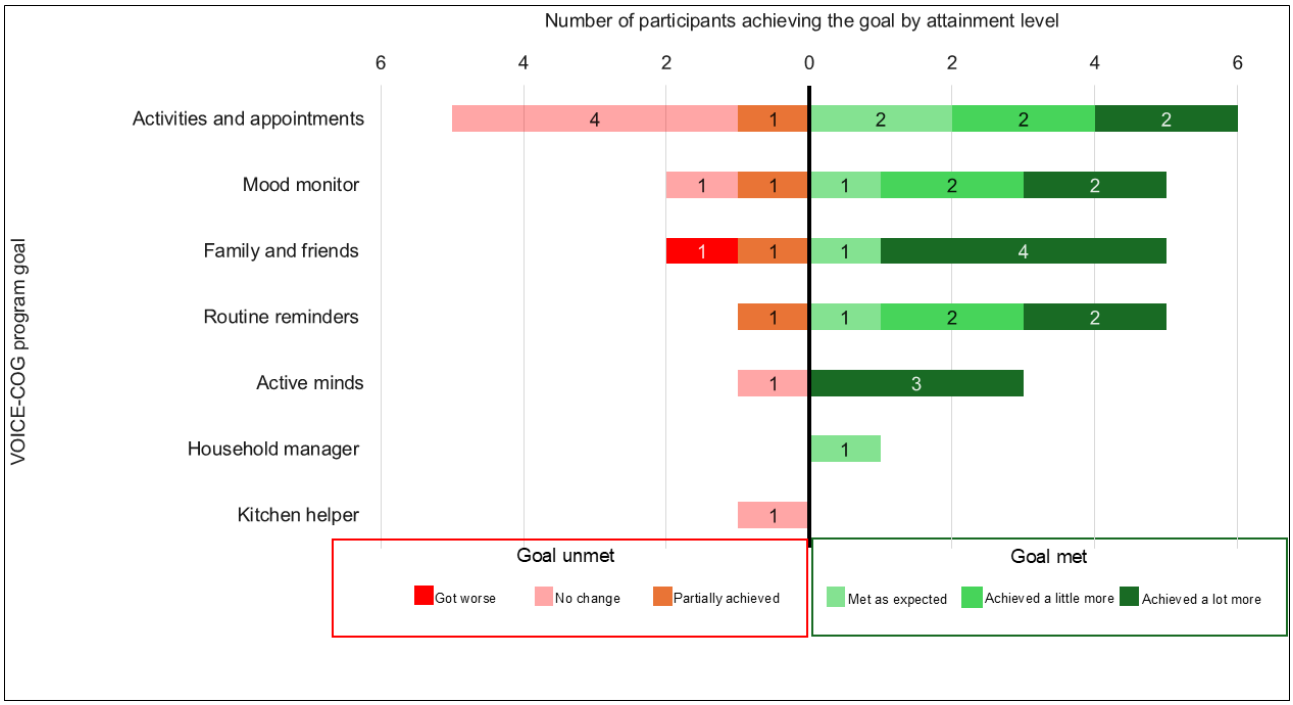
Among the 15 DVA participants, a mean of 2.6 (SD 0.5) goals was selected, with the majority (n=9, 60%) selecting 3 goals. DVA participants reported a significant increase in goal attainment, from a mean GAS-Light t-score of 36.8 (SD 0.7) at baseline to 57.6 (SD 15.4) at 12 weeks (P<.001). Most participants (n=9, 60%) met or exceeded their goals, as determined by a GAS-Light t-score of ≥50.

Thirty-nine VOICE-COG goals were set by participants overall, with 37 assessed at 12 weeks (1 participant who selected 2 goals was lost to follow-up). Further details, including example goals and attainment levels by VOICE-COG program goal, can be found in Multimedia Appendix 2. Of the 15 intervention group participants, 12 (80%) selected the activities and appointments assistant goal, 8 (53%) selected mood monitor, 7 (47%) selected

XSL•FO
RenderX

family and friends, 6 (40%) selected routine reminders, 4 (27%) selected active minds, and just 1 (6.6%) participant selected household manager as a reablement goal. The kitchen helper program goal had to be withdrawn after 2 weeks due to technical issues with device response timing and voice command difficulties with navigating recipes. Over two-thirds of VOICE-COG program goals (25/37, 68%) were attained at the expected level or above, with over a third (13/37, 35%) achieved "a lot more" than expected. Only 1 participant reported "getting worse" on a goal related to writing their life story (family and friends). Dichotomized goal attainment (scores of 0, +1, +2=goal met; scores of –0.5, –1, or –2=goal unmet) by VOICE-COG goal is shown in Figure 3.

Figure 3. Number of digital voice assistant intervention group participants achieving each goal attainment level after 12 weeks by VOICE-COG program goal.



Participant and Care Partner Questionnaires

Table 3 summarizes changes in participant and care partner questionnaire scores for both the intervention and control groups. After 12 weeks, the intervention group participants reported a significant increase in the MMQ-Strategy score (mean difference 5.0, 95% CI 0.03-10.1; *P*=.049) compared to control. This represented a medium size effect (*d*=+0.5). There were no other significant differences between groups after 12 weeks.

XSL•FO
RenderX

Table . Mean scores for participant and care partner questionnaires at baseline and 12 weeks by DVA intervention and enhanced usual care control groups, with estimated between-group effect sizes (95% CI).

Outcome measure	Control[a], mean (SD)		DVA[b], mean (SD)		Estimated difference between groups		Effect size
	Baseline (n=15)	12 weeks (n=15)	Baseline (n=15)	12 weeks (n=14)	Effect size (95% CI)[c]	*P* value	Cohen *d*
Self-reported questionnaires							
GDS-SF[d]	5.5 (3.3)	6.7 (4.4)	3.3 (3.3)	3.5 (2.9)	–1.4 (–3.6 to 0.7)	.19	–0.39
Multifactorial Memory Questionnaire (MMQ)							
Satisfaction	28.9 (16.2)	28.1 (13.5)	29.9 (8.9)	32.2 (9.8)	3.4 (–2.4 to 9.2)	.24	0.29
Ability	34.6 (10.5)	34.1 (14.7)	42.1 (9.3)	42.2 (11.4)	1 (–6.5 to 8.5)	.79	0.08
Strategies	35.1 (8.3)	33.5 (10.1)	37.4 (13.2)	39.9 (8.3)	5 (0.03 to 10.1)	.049[e]	0.55
Illness Cognition Questionnaire (ICQ)							
Helplessness	13.2 (4.4)	12.9 (4.3)	11.1 (2.3)	11 (2.6)	–0.2 (–2 to 1.5)	.77	–0.07
Acceptance	14.4 (3.3)	13.6 (3.9)	15.9 (4)	15.5 (3.2)	0.8 (–1.1 to 2.7)	.41	0.22
Perceived benefits	13.7 (3.4)	13.4 (2.8)	16.1 (4)	14.9 (5.1)	–0.7 (–2.6 to 1.2)	.46	–0.17
Care-partner reported questionnaires							
B-ADL[f]	3.3 (2.3)	3.9 (2.2)	3.3 (2.4)	4.2 (2.3)	0.2 (–1.1 to 1.6)	.71	0.11
ECog[g] questionnaire							
ECog	2.4 (0.7)	2.6 (0.6)	2.2 (0.6)	2.5 (0.8)	–0.0 (–0.4 to 0.4)	.99	0.00
Memory[h]	3 (0.7)	3 (0.6)	2.7 (0.7)	3 (0.8)	0.2 (–0.2 to 0.6)	.27	0.29
Language	2.3 (0.9)	2.5 (0.8)	2.2 (0.8)	2.3 (0.9)	–0.1 (–0.6 to 0.4)	.76	–0.09
Visuospatial	1.8 (0.8)	2.1 (0.8)	1.8 (0.6)	2.1 (0.8)	–0.1 (–0.6 to 0.4)	.61	–0.16
Planning	2.1 (0.9)	2.3 (0.9)	1.9 (0.8)	2.4 (0.9)	0.3 (–0.3 to 0.8)	.31	0.31
Organization	2.5 (0.9)	2.6 (0.8)	1.9 (0.7)	2.1 (0.7)	–0.1 (–0.5 to 0.4)	.79	−0.07
Attention	2.5 (0.8)	2.8 (1)	2.2 (0.9)	2.5 (0.9)	–0.1 (–0.6 to 0.5)	.78	−0.08
Zarit Burden Scale	28.1 (17.7)	28.9 (19.2)	26.2 (14.2)	27 (17.7)	0.3 (–6.6 to 7.3)	.92	0.02

[a]Control=usual care+weekly emails on cognition and memory.

[b]DVA=reablement goal-setting+personalized program via the Amazon Alexa Show 8.

[c]Adjusted for baseline.

[d]GDS: Geriatric Depression Scale.

[e]Indicates a significant difference in change between groups ($P<.05$).

[f]B-ADL: Bayer Activities of Daily Living.

[g]ECog: Everyday Cognition scale.

[h]Analysis was conducted on natural log–transformed data due to nonnormal distribution.

Discussion

Principal Findings

Findings from this pilot RCT demonstrate the feasibility of a co-designed cognitive reablement program broadcast via DVAs to support older adults living with dementia/MCI. The target sample was successfully recruited, and a high retention rate was achieved. Adherence, assessed as weekly interaction with program goals, exceeded target levels (66%), and the mean SUS score indicated marginal acceptability. Secondary outcomes showed the intervention group participants achieved a significant improvement in the attainment of personalized reablement goals. While most outcomes did not differ between groups, memory

XSL•FO
RenderX

strategy use was higher after 12 weeks in the intervention participants compared to the control group. Given that this feasibility study was not designed or powered to evaluate efficacy, a full-scale trial is required to evaluate the impact for people living with dementia/MCI.

Feasibility

Limited evidence on digital health applications to remotely deliver cognitive reablement programs for people living with dementia/MCI restricts comparisons of our study results [57-59]. Previous studies using smartphones or tablet-based apps for cognitive stimulation or training in people living with dementia have reported recruitment challenges related to device incompatibility [60,61]. The provision of the low-cost DVA, currently retailing at Aus $349 (Aus $1=US $0.70 as of Oct 8, 2026), as an additional device may, in part, explain our higher (30/35, 86%) recruitment rate. The intervention group retention rate of 93% (14/15) is higher than in previous telehealth reablement studies that observed retention rates of 60% [16] and 83% [17]. Careful interpretation is required given the significant difference in participant time burden for the virtual face-to-face studies. The 70% adherence rate was lower than 80% [28] and 85% [27] rates recorded in previous BuddyLink studies; however, this result should be interpreted with caution given personalization of VOICE-COG resulted in participants nominating between 7 to 52 weekly prompts compared to 1 or 2 in previous programs. A qualitative study with care partners of people living with dementia (n=5) explored the feasibility of DVA use for individual cognitive stimulation therapy (iCST), finding device accessibility, training provision, and visual stimulation could increase engagement [24]. The moderately high adherence rate in our study may have been due to the co-designed nature of the VOICE-COG program, which improved these functions and features through iterative feedback. Together, our feasibility findings provide support for a future large-scale trial.

System Usability

The VOICE-COG program's usability was rated as marginally acceptable (mean SUS 56.3, SD 18.1). It is possible this result was limited by the SUS outcome measure, which has not been fully validated for use in this population or with the DVA modality. The SUS is commonly used in technology evaluations by people living with dementia/MCI [62]; however, a modified version with simplified language has been proposed [63]. Furthermore, a recent review highlighted that usability assessments designed for graphical user interfaces need to be revised and validated for application with DVA systems [64]. The Voice Usability Scale could be considered as an outcome measure for future trials, having been tested as an alternative to the SUS on the Amazon Alexa [65]. Finally, in alignment with the IDEAS framework, a subsequent qualitative analysis of user experiences will be conducted to inform further development of VOICE-COG for evaluation in large-scale RCT. Future research should consider exploratory analyses examining associations between usability and other outcome measures to determine whether differences in usability for each VOICE-COG program goal influence effectiveness.

Goal Attainment

Goal attainment significantly increased among participants who received the VOICE-COG reablement program via their DVAs, with 60% (9/15) reporting having met or exceeded their goals. It is possible these results were influenced by the goal-setting behavior change process alone. While GAS-Light has recently been recommended for goal setting in reablement programs for people living with dementia [35], the heterogeneity of goal outcome measures used in previous studies limits comparability [66,67]. Furthermore, there is a lack of standardization in the application of the GAS with differences in scales and/or allocation of baseline scores [66]. A study using a touchscreen prompter to support household tasks (only) in people living with dementia (n=8) found 64% (14/22) of goals were met after 4 weeks [68], consistent with our finding of 68% (25/37) of goals achieved. More frequent use of the touchscreen prompter was associated with successful achievement of goals [69], highlighting future research should explore the relationship between adherence and goal attainment. The virtual telehealth study reported a higher percentage of goals were achieved (7/9, 78%) in the telehealth intervention group; however, the smaller sample size (n=3) limits comparison [16].

Cognitive Outcomes

Intervention participants in our study reported significantly higher use of memory strategies after 12 weeks compared with the control group. Despite this promising result, longer duration and adequately powered evaluation trials are required to determine if the increased use of DVA memory strategies could have a positive influence on memory performance and satisfaction over time [70]. It could be argued that this positive result was due to the integrated DVA reminder functions, which include alarms, timers, lists, calendars, and routines, which have been reported to support older adults with memory [71]. Future research should explore the effect of the VOICE-COG program that includes prompts and training support for the integrated functions compared to standalone DVA use. There is limited evidence on the effects of assistive technology for memory support in people living with dementia, with 0 studies included in a 2017 Cochrane review [72]. A recent smartphone study found people living with dementia (n=52) who used a voice recorder or reminder app reported improvements in prospective memory after 4 weeks, with higher levels of app engagement associated with a significant (P=.049) improvement in Instrument of Activities of Daily Living (IADL) [60]. Given the borderline statistical significance in both studies, additional research is needed to determine if there are meaningful effects on cognition and functioning for people living with dementia/MCI longer term.

Implications for Future Research

Feasibility results from this study support consideration of a full-scale trial to evaluate the impact of VOICE-COG, relative to usual care for people living with dementia/MCI. The effect sizes found in this study may be used to inform decision-making regarding adjustments, such as sample size for a future trial [73]. Recruitment via clinically facilitated partnerships, such as memory clinics, should be considered to reduce selection bias from web-based advertisements. VOICE-COG used

XSL•FO
RenderX

co-design to adapt a practical cognitive reablement framework for people living with dementia [74] and a theory-based goal-focused approach [75,76], drawing on various evidence of empirically supported strategies. Future research should examine reasons for lack of goal attainment unrelated to usability and ensure the latest reablement strategies for older adults with dementia are incorporated. This should include increasing accessibility of VOICE-COG to a more diverse population, including minority ethnic groups, or lower socioeconomic status, and those in rural areas. Digital competency measures should be collected at baseline in future trials to investigate potential bias and determine if this influences intervention engagement. Given the high prevalence of depression and anxiety in this population, consideration should be given to the use of positive psychology outcome measures currently in development [77,78]. An important issue for future research is consensus on an objective validated measure of adherence or engagement for digitally delivered reablement support programs and DVA interventions. Finally, the cost-effectiveness of VOICE-COG should be evaluated in comparison to a telehealth virtual reablement program before implementation at scale.

Limitations

Recruitment methods in this study used web-based applications that required a level of digital literacy, which could limit generalization of these findings to all people living with dementia/MCI. This may have led to an unintended bias of a highly educated sample, potentially limiting generalization to all people living with dementia/MCI. Since participants self-reported their diagnosis of dementia/MCI, cognitive impairment stage could not be determined. Therefore, the inclusion of an objective, validated cognitive assessment should be considered for future studies. Goal attainment measures were self-reported and could have been positively influenced by social desirability bias. Furthermore, control group participants did not set reablement goals, so we could not ascertain if goal attainment differences were associated with the intervention or the goal-setting process. Participants had not previously owned an Amazon Alexa Show 8 DVA, so SUS responses may have been influenced by the DVA's integrated features and functionality. Questionnaires were administered online, but several people living with dementia shared email addresses with their care partners, so responses could have been misrepresented due to proxy reporting bias. Secondary outcomes should be interpreted with caution as our sample size was not powered to detect effects. Consistent with the exploratory aims of feasibility studies, we did not correct for multiple comparisons in hypothesis testing, which may have led to false-positive findings. Finally, as a pilot feasibility trial of a minimum viable product, additional human support elements included researcher input during program configuration. Future trials should consider an integrated configuration interface, enabling direct transfer of data from web-based forms into the BuddyLink software for real-world implementation.

Conclusions

This co-designed VOICE-COG cognitive reablement program was feasible to deliver to older adults living with dementia/MCI via DVA, with high levels of retention and adherence. Larger trials with longer-term follow-up and cost-effectiveness evaluation are needed to determine how this remotely delivered and monitored program can support older adults living with early-stage dementia or MCI to meet their reablement goals.

Acknowledgments

StepUp for Dementia Research is funded by the Australian Government Department of Health and is implemented by a dedicated team at the University of Sydney.

TY acknowledges the use of Microsoft Copilot for occasional sentence refinement to improve the clarity and organization of the language while preparing the initial and revised drafts of the manuscript.

Funding

This research was funded by a Dementia Australia Research Foundation grant. TY was supported by an Australian Government Research Training Program (RTP) Scholarship (doi.org/10.82133/C42F-K220).

Data Availability

The datasets generated and analyzed during this study can be made available upon reasonable request and with approval from the Deakin University Human Research Ethics Committee.

Authors' Contributions

PJ, HM, DS, MC, and EG were responsible for the initial conceptualization, funding acquisition, and methodology design. PJ, HM, and DS were responsible for the supervision of the research team. TY was responsible for conducting the investigation, coordinating and planning research activities, creating co-designed modules, programming and testing of the software, data curation and analysis, writing the initial draft of the manuscript, and editing for submission. C Glavas, C Gourlay, and CC supported research activities and data curation. All coauthors, including RM, were responsible for reviewing and editing the manuscript.

XSL•FO
RenderX

Conflicts of Interest

PJ is an editorial board member for *JMIR Aging*. PJ was not involved in the journal's peer-review process or in decisions related to this manuscript.
EG is a director (unpaid position) of Great Australian Pty Ltd. Great Australian Pty Ltd provided in-kind technical support for software and had a nonfinancial interest in the development of a voice-based digital assistant software for telehealth. Great Australian had no input into this study's design, analysis, interpretation of the results, or decision to publish.

Multimedia Appendix 1
VOICE-COG program information for participants.
[PDF File, 169 KB - aging_v9i1e96922_app1.pdf]

Multimedia Appendix 2
Example VOICE-COG program goals.
[PDF File, 99 KB - aging_v9i1e96922_app2.pdf]

Multimedia Appendix 3
VOICE-COG program detail.
[PDF File, 1814 KB - aging_v9i1e96922_app3.pdf]

Multimedia Appendix 4
Weekly VOICE-COG broadcasts and interactions by participant.
[PDF File, 100 KB - aging_v9i1e96922_app4.pdf]

Multimedia Appendix 5
System Usability Scale.
[PDF File, 123 KB - aging_v9i1e96922_app5.pdf]

Checklist 1
CONSORT-EHEALTH (V 1.6.1) checklist.
[PDF File, 1214 KB - aging_v9i1e96922_app6.pdf]

References

1. Chen S, Cao Z, Nandi A, et al. The global macroeconomic burden of Alzheimer's disease and other dementias: estimates and projections for 152 countries or territories. Lancet Glob Health 2024 Sep;12(9):e1534-e1543. [doi: 10.1016/S2214-109X(24)00264-X] [Medline: 39151988]
2. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
3. Chiao CY, Wu HS, Hsiao CY. Caregiver burden for informal caregivers of patients with dementia: a systematic review. Int Nurs Rev 2015 Sep;62(3):340-350. [doi: 10.1111/inr.12194] [Medline: 26058542]
4. Wimo A, Seeher K, Cataldi R, et al. The worldwide costs of dementia in 2019. Alzheimers Dement 2023 Jul;19(7):2865-2873. [doi: 10.1002/alz.12901] [Medline: 36617519]
5. Dyer SM, Harrison SL, Laver K, Whitehead C, Crotty M. An overview of systematic reviews of pharmacological and non-pharmacological interventions for the treatment of behavioral and psychological symptoms of dementia. Int Psychogeriatr 2018 Mar;30(3):295-309. [doi: 10.1017/S1041610217002344] [Medline: 29143695]
6. Paggetti A, Druda Y, Sciancalepore F, et al. The efficacy of cognitive stimulation, cognitive training, and cognitive rehabilitation for people living with dementia: a systematic review and meta-analysis. Geroscience 2025 Feb;47(1):409-444. [doi: 10.1007/s11357-024-01400-z] [Medline: 39485657]
7. Clare L, Kudlicka A, Oyebode JR, et al. Individual goal-oriented cognitive rehabilitation to improve everyday functioning for people with early-stage dementia: a multicentre randomised controlled trial (the GREAT trial). Int J Geriatr Psychiatry 2019 May;34(5):709-721. [doi: 10.1002/gps.5076] [Medline: 30724405]
8. Jeon YH, Simpson J, Fethney J, et al. Effectiveness of the Interdisciplinary Home-based Reablement Programme (I-HARP) on improving functional independence of people living with dementia: a multicentre, pragmatic, randomised, open-label, controlled trial. J Neurol Neurosurg Psychiatry 2025 Jun 12;96(7):705-715. [doi: 10.1136/jnnp-2024-334514] [Medline: 39824599]

XSL•FO
RenderX

9. Regan B, Wells Y, Farrow M, O'Halloran P, Workman B. MAXCOG-maximizing cognition: a randomized controlled trial of the efficacy of goal-oriented cognitive rehabilitation for people with mild cognitive impairment and early Alzheimer disease. Am J Geriatr Psychiatry 2017 Mar;25(3):258-269. [doi: 10.1016/j.jagp.2016.11.008] [Medline: 28034509]
10. Clemson L, Laver K, Jeon YH, et al. Implementation of an evidence-based intervention to improve the wellbeing of people with dementia and their carers: study protocol for 'Care of People with dementia in their Environments (COPE)' in the Australian context. BMC Geriatr 2018 May 9;18(1):108. [doi: 10.1186/s12877-018-0790-7] [Medline: 29739358]
11. Gitlin LN, Winter L, Dennis MP, Hodgson N, Hauck WW. A biobehavioral home-based intervention and the well-being of patients with dementia and their caregivers: the COPE randomized trial. JAMA 2010 Sep 1;304(9):983-991. [doi: 10.1001/jama.2010.1253] [Medline: 20810376]
12. Clare L, Linden DEJ, Woods RT, et al. Goal-oriented cognitive rehabilitation for people with early-stage Alzheimer disease: a single-blind randomized controlled trial of clinical efficacy. Am J Geriatr Psychiatry 2010 Oct;18(10):928-939. [doi: 10.1097/JGP.0b013e3181d5792a] [Medline: 20808145]
13. Clemson L, Laver K, Rahja M, et al. Implementing a reablement intervention, "Care of People with dementia in their Environments (COPE)": a hybrid implementation-effectiveness study. Gerontologist 2021 Aug 13;61(6):965-976. [doi: 10.1093/geront/gnaa105] [Medline: 32803248]
14. Clare L, Kudlicka A, Collins R, et al. Implementing a home-based personalised cognitive rehabilitation intervention for people with mild-to-moderate dementia: GREAT into practice. BMC Geriatr 2023 Feb 13;23(1):93. [doi: 10.1186/s12877-022-03705-0] [Medline: 36782122]
15. da Fonseca MH, Kovaleski F, Picinin CT, Pedroso B, Rubbo P. E-health practices and technologies: a systematic review from 2014 to 2019. Healthcare (Basel) 2021 Sep 10;9(9):1192. [doi: 10.3390/healthcare9091192] [Medline: 34574966]
16. Burton RL, O'Connell ME. Telehealth rehabilitation for cognitive impairment: randomized controlled feasibility trial. JMIR Res Protoc 2018 Feb 8;7(2):e43. [doi: 10.2196/resprot.9420] [Medline: 29422453]
17. Laver K, Liu E, Clemson L, et al. Does telehealth delivery of a dyadic dementia care program provide a noninferior alternative to face-to-face delivery of the same program? A randomized, controlled trial. Am J Geriatr Psychiatry 2020 Jun;28(6):673-682. [doi: 10.1016/j.jagp.2020.02.009] [Medline: 32234275]
18. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
19. Schlomann A, Wahl HW, Zentel P, et al. Potential and pitfalls of digital voice assistants in older adults with and without intellectual disabilities: relevance of participatory design elements and ecologically valid field studies. Front Psychol 2021;12:684012. [doi: 10.3389/fpsyg.2021.684012] [Medline: 34276507]
20. Balasubramanian GV, Beaney P, Chambers R. Digital personal assistants are smart ways for assistive technology to aid the health and wellbeing of patients and carers. BMC Geriatr 2021 Nov 15;21(1):643. [doi: 10.1186/s12877-021-02436-y] [Medline: 34781881]
21. Vinay R, Uetova E, Tommila NC, Biller-Andorno N, Kowatsch T. A hybrid rule- and large language model-based embodied voice assistant (GRACE) for cognitive stimulation in older adults: usability study assessing technical feasibility, technology acceptance, and working alliance. JMIR Aging 2025 Dec 18;8:e76489. [doi: 10.2196/76489] [Medline: 41411649]
22. Striegl J, Gollasch D, Loitsch C, Weber G. Designing VUIs for social assistance robots for people with dementia. MuC '21 2021 Sep 13:145-155. [doi: 10.1145/3473856.3473887]
23. Salai AM, Cook G, Holmquist LE. "The relief is amazing": an in-situ short field evaluation of a personal voice assistive technology for a user living with dementia. In: Nocera JA, Lárusdóttir MK, Petrie H, Piccinno A, Winckler M, editors. Human-Computer Interaction – INTERACT 2023: 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part I: Springer; 2023:165-175. [doi: 10.1007/978-3-031-42280-5_11]
24. Qiu L, Saragih ID, Fick DM, Sundar SS, Abdullah S. Voice assistants to deliver cognitive stimulation therapy for persons living with dementia. Ext Abstr Hum Factors Computing Syst 2025;2025:582. [doi: 10.1145/3706599.3719707] [Medline: 40657447]
25. BuddyLink VIPA therapeutic program. BuddyLink. 2023. URL: https://www.buddylink.au [accessed 2026-10-03]
26. Jansons P, Dalla Via J, Daly RM, Fyfe JJ, Gvozdenko E, Scott D. Delivery of home-based exercise interventions in older adults facilitated by Amazon Alexa: a 12-week feasibility trial. J Nutr Health Aging 2022;26(1):96-102. [doi: 10.1007/s12603-021-1717-0] [Medline: 35067710]
27. Glavas C, Scott D, Sood S, et al. Exploring the feasibility of digital voice assistants for delivery of a home-based exercise intervention in older adults with obesity and type 2 diabetes mellitus: randomized controlled trial. JMIR Aging 2024 Sep 13;7:e53064. [doi: 10.2196/53064] [Medline: 39270212]
28. Hunegnaw MT, Mesinovic J, Jansons P, et al. Feasibility and effectiveness of a digital voice assistant for improving anti-osteoporosis medication adherence, and osteoporosis knowledge and attitudes, in postmenopausal women with osteoporosis: a 12-month randomised controlled trial. Arch Osteoporos 2025 Apr 9;20(1):43. [doi: 10.1007/s11657-025-01529-0] [Medline: 40202535]
29. Yiannis T, Macpherson H, Scott D, et al. Cognitive reablement using digital voice assistants for people living with dementia or mild cognitive impairment: a co-design study. Health Expect 2026 Aug;29(4):e70783. [doi: 10.1111/hex.70783] [Medline: 42542978]

XSL•FO
RenderX

30. Eysenbach G. CONSORT-EHEALTH: implementation of a checklist for authors and editors to improve reporting of web-based and mobile randomized controlled trials. Stud Health Technol Inform 2013;192:657-661. [Medline: 23920638]
31. Øksnebjerg L, Woods B, Ruth K, et al. A tablet app supporting self-management for people with dementia: explorative study of adoption and use patterns. JMIR mHealth uHealth 2020 Jan 17;8(1):e14694. [doi: 10.2196/14694] [Medline: 31951217]
32. Jeon YH, Shin M, Smith A, et al. Early implementation and evaluation of StepUp for Dementia Research: an Australia-wide dementia research participation and public engagement platform. Int J Environ Res Public Health 2021 Oct 28;18(21):11353. [doi: 10.3390/ijerph182111353] [Medline: 34769871]
33. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016 Dec 16;18(12):e317. [doi: 10.2196/jmir.5927] [Medline: 27986647]
34. O'Connor CMC, Rowlands A, Poulos CJ. Development of an assessment guide to evaluate meaningful outcomes for people living with dementia who are engaged in reablement programs. Disabil Rehabil 2022 Oct;44(20):6042-6054. [doi: 10.1080/09638288.2021.1960437] [Medline: 34369257]
35. O'Connor CMC, Poulos CJ. Setting goals and evaluating meaningful outcomes for people living with dementia who are participating in reablement programs: practitioner guide. : HammondCare; 2021 URL: https://www.hammond.com.au/hubfs/HammondCare/HC%20Resource%20Library/Resources/Reablement-Practitioner-Guide.pdf?hsCtaAttrib=183700671346 [accessed 2026-10-05]
36. Mair JL, Salamanca-Sanabria A, Augsburger M, et al. Effective behavior change techniques in digital health interventions for the prevention or management of noncommunicable diseases: an umbrella review. Ann Behav Med 2023 Sep 13;57(10):817-835. [doi: 10.1093/abm/kaad041] [Medline: 37625030]
37. Yang Q, Millette D, Zhou C, Beatty M, Carcioppolo N, Wilson G. The effectiveness of interactivity in improving mediating variables, behaviors and outcomes of web-based health interventions: a meta-analytic review. Health Commun 2020 Oct;35(11):1334-1348. [doi: 10.1080/10410236.2019.1631992] [Medline: 31240958]
38. Liu K, Xie Z, Or CK. Effectiveness of mobile app-assisted self-care interventions for improving patient outcomes in type 2 diabetes and/or hypertension: systematic review and meta-analysis of randomized controlled trials. JMIR mHealth uHealth 2020 Aug 4;8(8):e15779. [doi: 10.2196/15779] [Medline: 32459654]
39. Donovan G, Hall N, Ling J, Smith F, Wilkes S. Influencing medication taking behaviors using automated two-way digital communication: a narrative synthesis systematic review informed by the Behavior Change Wheel. Br J Health Psychol 2022 Sep;27(3):861-890. [doi: 10.1111/bjhp.12580] [Medline: 35080811]
40. Dugas M, Gao GG, Agarwal R. Unpacking mHealth interventions: a systematic review of behavior change techniques used in randomized controlled trials assessing mHealth effectiveness. Digit Health 2020;6:2055207620905411. [doi: 10.1177/2055207620905411] [Medline: 32128233]
41. Pouls BPH, Vriezekolk JE, Bekker CL, et al. Effect of interactive eHealth interventions on improving medication adherence in adults with long-term medication: systematic review. J Med Internet Res 2021 Jan 8;23(1):e18901. [doi: 10.2196/18901] [Medline: 33416501]
42. Help sheets: printable PDFs. Dementia Australia. 2025. URL: https://dementia-org.libguides.com/help-sheets [accessed 2026-10-05]
43. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry: CRC Press; 1996:189-194.
44. Turner-Stokes L. Goal attainment scaling (GAS) in rehabilitation: a practical guide. Clin Rehabil 2009 Apr;23(4):362-370. [doi: 10.1177/0269215508101742] [Medline: 19179355]
45. Rapaport P, Burton A, Palomo M, et al. A mixed-methods feasibility study of a goal-focused manualised intervention to support people with dementia to stay living independently at home with support from family carers: NIDUS (New Interventions for Independence in Dementia Study) Family. Aging Ment Health 2021 Aug;25(8):1463-1474. [doi: 10.1080/13607863.2020.1845299] [Medline: 33222498]
46. Turner-Stokes L. GAS calculation spreadsheet. King's College London. URL: https://www.kcl.ac.uk/cicelysaunders/resources/toolkits/gas-overview [accessed 2025-05-20]
47. Lach HW, Chang YP, Edwards D. Can older adults with dementia accurately report depression using brief forms? Reliability and validity of the Geriatric Depression Scale. J Gerontol Nurs 2010 May;36(5):30-37. [doi: 10.3928/00989134-20100303-01] [Medline: 20349852]
48. Troyer AK, Rich JB. Psychometric properties of a new metamemory questionnaire for older adults. J Gerontol B Psychol Sci Soc Sci 2002 Jan;57(1):19-27. [doi: 10.1093/geronb/57.1.p19] [Medline: 11773220]
49. Evers AW, Kraaimaat FW, van Lankveld W, Jongen PJ, Jacobs JW, Bijlsma JW. Beyond unfavorable thinking: the Illness Cognition Questionnaire for chronic diseases. J Consult Clin Psychol 2001 Dec;69(6):1026-1036. [Medline: 11777106]
50. Nagaratnam N, Nagaratnam K, O'Mara D. Bayer-Activities of Daily Living Scale in mild and moderate dementia of the Alzheimer type. Am J Alzheimers Dis Other Demen 2013 Dec;28(8):784-789. [doi: 10.1177/1533317513500842] [Medline: 24005851]

XSL•FO
RenderX

51. Farias ST, Mungas D, Reed BR, et al. The measurement of Everyday Cognition (ECog): scale development and psychometric properties. Neuropsychology 2008 Jul;22(4):531-544. [doi: 10.1037/0894-4105.22.4.531] [Medline: 18590364]
52. Bédard M, Molloy DW, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. Gerontologist 2001 Oct;41(5):652-657. [doi: 10.1093/geront/41.5.652] [Medline: 11574710]
53. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016 Jun;25(3):1057-1073. [doi: 10.1177/0962280215588241] [Medline: 26092476]
54. Eldridge SM, Lancaster GA, Campbell MJ, et al. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. PLoS ONE 2016;11(3):e0150205. [doi: 10.1371/journal.pone.0150205] [Medline: 26978655]
55. Blattgerste J, Behrends J, Pfeiffer T. A web-based analysis toolkit for the System Usability Scale. 15th Int Conf Pervasive Technol Assist Environ 2022:237-246. [doi: 10.1145/3529190.3529216]
56. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
57. Di Lorito C, Bosco A, Rai H, et al. A systematic literature review and meta-analysis on digital health interventions for people living with dementia and mild cognitive impairment. Int J Geriatr Psychiatry 2022 Jun;37(6). [doi: 10.1002/gps.5730] [Medline: 35588315]
58. Ditton A, Alodan H, Fox C, Evans S, Cross J. Exploring the effectiveness and experiences of people living with dementia interacting with digital interventions: a mixed methods systematic review. Dementia (London) 2025 Apr;24(3):506-551. [doi: 10.1177/14713012241302371] [Medline: 39604136]
59. Neal D, Craven MP, Cross J, et al. Digital assistive technologies for community-dwelling people with dementia: a systematic review of systematic reviews by the INTERDEM AI & assistive technology taskforce. Digit Health 2025;11:20552076251362353. [doi: 10.1177/20552076251362353] [Medline: 40761778]
60. Scullin MK, Jones WE, Phenis R, et al. Using smartphone technology to improve prospective memory functioning: a randomized controlled trial. J Am Geriatr Soc 2022 Feb;70(2):459-469. [doi: 10.1111/jgs.17551] [Medline: 34786698]
61. Rai HK, Schneider J, Orrell M. An individual cognitive stimulation therapy app for people with dementia and carers: results from a feasibility randomized controlled trial (RCT). Clin Interv Aging 2021;16:2079-2094. [doi: 10.2147/CIA.S323994] [Medline: 35221680]
62. Miguel Cruz A, Daum C, Comeau A, et al. Acceptance, adoption, and usability of information and communication technologies for people living with dementia and their care partners: a systematic review. Disabil Rehabil Assist Technol 2023 May;18(4):443-457. [doi: 10.1080/17483107.2020.1864671] [Medline: 33378627]
63. Holden RJ. A simplified System Usability Scale (SUS) for cognitively impaired and older adults. Proc Int Symp Hum Factors Ergon Health Care 2020 Sep;9(1):180-182. [doi: 10.1177/2327857920091021]
64. Deshmukh AM, Chalmeta R. User experience and usability of voice user interfaces: a systematic literature review. Information 2024;15(9):579. [doi: 10.3390/info15090579]
65. Zwakman DS, Pal D, Arpnikanondt C. Usability evaluation of artificial intelligence-based voice assistants: the case of Amazon Alexa. SN Comput Sci 2021;2(1):28. [doi: 10.1007/s42979-020-00424-4] [Medline: 33458698]
66. Budgett J, Sommerlad A, Kupeli N, Zabihi S, Olsen A, Cooper C. Setting individualised goals for people living with dementia and their family carers: a systematic review of goal-setting outcome measures and their psychometric properties. Dementia (London) 2024 Feb;23(2):312-340. [doi: 10.1177/14713012231222309] [Medline: 38105445]
67. Jogie P, Rahja M, van den Berg M, Cations M, Brown S, Laver K. Goal setting for people with mild cognitive impairment or dementia in rehabilitation: a scoping review. Aust Occup Ther J 2021 Dec;68(6):563-592. [doi: 10.1111/1440-1630.12758] [Medline: 34346077]
68. Harris N, Boyd H, Evans N, et al. A preliminary evaluation of a client-centred prompting tool for supporting everyday activities in individuals with mild to moderate levels of cognitive impairment due to dementia. Dementia (London) 2021 Apr;20(3):867-883. [doi: 10.1177/1471301220911322] [Medline: 32249596]
69. Boyd HC, Evans NM, Orpwood RD, Harris ND. Using simple technology to prompt multistep tasks in the home for people with dementia: an exploratory study comparing prompting formats. Dementia (London) 2017 May;16(4):424-442. [doi: 10.1177/1471301215602417] [Medline: 26428634]
70. Kinsella GJ, Ames D, Storey E, et al. Strategies for improving memory: a randomized trial of memory groups for older people, including those with mild cognitive impairment. J Alzheimers Dis 2016;49(1):31-43. [doi: 10.3233/JAD-150378] [Medline: 26444773]
71. O'Brien K, Liggett A, Ramirez-Zohfeld V, Sunkara P, Lindquist LA. Voice-controlled intelligent personal assistants to support aging in place. J Am Geriatr Soc 2020 Jan;68(1):176-179. [doi: 10.1111/jgs.16217] [Medline: 31617581]
72. Ball BH, Peper P, Robison MK. Reminders eliminate age-related declines in prospective memory. Psychol Aging 2025 Feb;40(1):54-65. [doi: 10.1037/pag0000844] [Medline: 39325414]
73. Charlesworth G, Burnell K, Hoe J, Orrell M, Russell I. Acceptance checklist for clinical effectiveness pilot trials: a systematic approach. BMC Med Res Methodol 2013 Jun 13;13:78. [doi: 10.1186/1471-2288-13-78] [Medline: 23758922]

XSL•FO
RenderX

74. Clare L. Rehabilitation for people living with dementia: a practical framework of positive support. PLoS Med 2017 Mar;14(3):e1002245. [doi: 10.1371/journal.pmed.1002245] [Medline: 28267744]
75. O'Connor C, Poulos CJ, Gresham M, Poulos RG. Supporting independence and function in people living with dementia: a technical guide to the evidence supporting reablement interventions. : HammondCare; 2019 URL: https://www.hammond.com.au/hubfs/HammondCare/HC%20Resource%20Library/Resources/Technical-Guide-Second-Edition-Reablement.pdf?hsCtaAttrib=196174735527 [accessed 2026-10-05]
76. O'Connor CMC, Gresham M, Poulos RG, et al. Translating reablement research for dementia practice: development of a handbook using implementation science. Disabil Rehabil 2022 Apr;44(8):1524-1536. [doi: 10.1080/09638288.2020.1797910] [Medline: 32772575]
77. Leung DKY, Chan WC, Spector A, Wong GHY. Prevalence of depression, anxiety, and apathy symptoms across dementia stages: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2021 Sep;36(9):1330-1344. [doi: 10.1002/gps.5556] [Medline: 33905138]
78. Stoner CR, Stansfeld J, Orrell M, Spector A. The development of positive psychology outcome measures and their uses in dementia research: a systematic review. Dementia (London) 2019 Aug;18(6):2085-2106. [doi: 10.1177/1471301217740288] [Medline: 29145739]

Abbreviations

B-ADL: Bayer Activities of Daily Living Scale
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
COPE: Care of People with dementia in their Environments
DVA: digital voice assistant
ECog: Everyday Cognition
GAS: Goal Attainment Scale
GDS: Geriatric Depression Scale
IADL: Instrument of Activities of Daily Living
ICQ: Illness Cognition Questionnaire
iCST: individual cognitive stimulation therapy
IDEAS: Integrate, Design, Assess, and Share
MCI: mild cognitive impairment
MMQ: Multifactorial Memory Questionnaire
RCT: randomized controlled trial
SMART: Specific, Measurable, Achievable, Relevant, and Time-bound
SUS: System Usability Scale

Edited by A Mavragani; submitted 13.Apr.2026; peer-reviewed by A Shenoi, NHM Fadzil; revised version received 27.Aug.2026; accepted 09.Sep.2026; published 23.Oct.2026.

Please cite as:
Yiannis T, Macpherson H, Scott D, Callisaya M, Chapple C, Glavas C, Gourlay C, Maddison R, Gvozdenko E, Jansons P
Feasibility of a Co-Designed Personalized Cognitive Reablement Program Delivered via Digital Voice Assistants for Older Adults Living With Early-Stage Dementia or Mild Cognitive Impairment: Pilot Randomized Controlled Trial
JMIR Aging 2026;9:e96922
URL: https://aging.jmir.org/2026/1/e96922
doi:10.2196/96922

© Tracy Yiannis, Helen Macpherson, David Scott, Michele Callisaya, Celeste Chapple, Costas Glavas, Caitlyn Gourlay, Ralph Maddison, Eugene Gvozdenko, Paul Jansons. Originally published in JMIR Aging (https://aging.jmir.org), 23.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

ASSETs for Aging in Place: 12-Month Descriptive Longitudinal Study of Technology-Supported Interprofessional Care Coordination After Nursing Home Discharge

Knoo Lee[1,2], PhD; Erin L Robinson[3], MPH, MSW, PhD; Amy Vogelsmeier[1], PhD; Blaine Reeder[1,2], PhD; Suzette Bacon[1], MSW; Elizabeth Conrow[1], BSN, MSN; Elizabeth Curtis[1], BSN; Ashley Roberts Dulany[1], MOT, OTD; Erica Perryman[4], PhD; Lori Popejoy[1], PhD

[1]Sinclair School of Nursing, University of Missouri, Columbia, MO, United States

[2]Institute for Data Science and Informatics, University of Missouri, Columbia, MO, United States

[3]College of Health Professions, University of Missouri, Columbia, MO, United States

[4]Fran and Earl Ziegler College of Nursing, University of Oklahoma, Oklahoma City, OK, United States

Corresponding Author:
Knoo Lee, PhD
Sinclair School of Nursing
University of Missouri
915 Hitt St
Columbia, MO, 65211
United States
Phone: 1 5738840421
Email: knoolee@missouri.edu

Abstract

Background: Older and disabled adults face ongoing barriers to aging in place (AIP) after nursing home discharge, particularly following the end of time-limited transition programs. Sustained care coordination and early detection of health and functional changes remain limited, especially in rural and underresourced communities.

Objective: This study aimed to evaluate the feasibility of technology-enabled interdisciplinary care coordination among older and disabled adults enrolled in the ASSETs (Age-Friendly Smart, Sustainable, and Equitable Technologies) for AIP program and to descriptively characterize changes in quality of life (QoL), function, and cognition over 12 months.

Methods: This quantitative, descriptive study used repeated measures over 12 months among older and disabled adults 1 year postdischarge from Missouri's Show-Me Home program. An interdisciplinary team (registered nurse [RN], occupational therapist [OT], and social worker [SW]) provided remote care coordination supported by open-source smart home sensors, consumer-grade wearable devices, and telehealth. Outcomes were assessed quarterly using the interRAI home care (HC) assessment and the Money Follows the Person (MFP) QoL survey.

Results: Nineteen clients were included. Sensor and wearable data were used for remote monitoring, identification of changes in activity patterns, and care coordination. Across repeated assessments, no statistically significant declines were observed in functional status, cognitive performance, instrumental activities of daily living (IADL), or quality-of-life measures. Reports of unmet medical care needs decreased across follow-up.

Conclusions: This study supports the feasibility of a digital health–enabled, interdisciplinary care coordination program using cost-feasible open-source and consumer-grade technologies. Preliminary findings indicate that no statistically significant decline in function or quality-of-life measures was observed during follow-up. Larger-scale implementation is needed to further evaluate effectiveness and refine data-driven care coordination workflows for underserved populations.

(JMIR Aging 2026;9:e91524) doi:10.2196/91524

KEYWORDS

aging in place; quality of life; care coordination; wearable sensors; rural health; interRAI

XSL•FO
RenderX

Introduction

Background

In the United States, one-third of individuals will live in a nursing home at some point in their lives, with nursing home admission often following hospitalization for an acute illness [1,2]. While it is difficult to quantify how many older adults will remain permanently in the nursing home following a hospitalization, Xu et al [3] found that only 29% of individuals admitted to the nursing home for a short stay successfully returned home with few readmissions to the nursing home or hospital. Older and disabled adults face significant barriers to maintaining independence at home following nursing home discharge, including limited access to services and caregiver networks. These barriers are heightened for individuals managing multiple chronic conditions, functional limitations, or limited mobility [4,5]. Successful aging in place (AIP) often depends on early identification of changes in health and timely intervention. Technology-enhanced care coordination can help identify and respond to health status changes, yet few individuals residing in their own homes have access to these services [2,6]. The lack of infrastructure to support remaining at home is particularly challenging in underresourced areas where limited financial, technological, and human resources hinder the adoption of systems that enable early detection and intervention [7,8].

In the United States, federal programs such as Medicare and Medicaid provide health insurance to eligible Americans. Medicare is primarily for older adults aged older than 65 years; however, younger individuals with disabilities may be eligible [9]. Medicaid is a joint federal-state insurance program for low-income Americans, covering about 1 in 5 people, including the older adults and adults with disabilities [9]. The Money Follows the Person (MFP) Initiative, known in Missouri as the Show-Me Home program, provides enhanced Medicaid funding to support individuals for up to 1 year who wish to transition from nursing homes or other institutional settings back into the community. Through this program, clients may access transition coordination, housing support, and home and community-based services designed to reduce reliance on institutional care and improve quality of life (QoL) [10-12]. However, when individuals are discharged from the Show-Me Home program, the risk for nursing home readmission re-emerges as a potential problem; this is particularly concerning for clients who do not meet criteria for home and community care services, as they have fewer community supports. To address these gaps, we launched the ASSETs (Age-Friendly Smart, Sustainable, and Equitable Technologies) for AIP demonstration project [13]. ASSETs for AIP is a digital health program that relies on an interdisciplinary team to provide sensor-enhanced care coordination to support and maintain independence for low-income older adults discharged from nursing homes back to the community. This study examines changes in QoL over 12 months for older adults enrolled in the ASSETs for AIP program and describes how sensor data are used by the care coordination team to support AIP.

AIP

AIP emphasizes supporting older adults' ability to live safely and independently in their homes and communities by leveraging individual strengths, environmental support, and coordinated services. Care coordination is the process of organizing care activities a person receives and sharing information among all clients involved to facilitate appropriate delivery of services [14]. Interdisciplinary care coordination models are central to AIP and integrate nursing, social work, occupational therapy, and community-based resources to address health, functional, and psychosocial needs in a holistic manner [14,15]. Rather than focusing solely on deficits, strengths-based approaches empower older adults to identify personally meaningful goals, actively participate in care planning, and build self-efficacy in managing chronic conditions and functional limitations [16]. When older adults are supported by coordinated, home-based interdisciplinary teams, early identification of health and functional changes can lead to timely interventions that reduce disability, improve QoL, and promote sustained independence [17].

Multicomponent biobehavioral interventions for older adults at home, such as CAPABLE (Community Aging in Place—Advancing Better Living for Elders) [18-20] and STAReD (Smart Technology for Aging and Reducing Disability) [21], have shown improvements in functional outcomes, safety, and health-related QoL, underscoring the value of interdisciplinary, community-based models that align clinical expertise with older adults' goals and lived environments [22]. Digital health technologies (eg, smart home sensor systems, wearable devices including smartwatches and rings, and telehealth [23]) further enhance care coordination by providing objective, longitudinal data on daily activities, mobility, and health-related metrics (eg, heart rate), enabling teams to detect subtle changes before they escalate into adverse outcomes [13,21]. When deployed to support community-based, strengths-focused care coordination frameworks, these technologies support proactive, person-centered decision-making while preserving autonomy and dignity.

Building on these principles, a series of aging-in-place initiatives at the University of Missouri (MU) have explored how interdisciplinary care coordination and technology-enabled monitoring support long-term community living among older adults. These programs provided the conceptual and operational foundation for the development of ASSETs for AIP. Changes in health conditions in older adults, if caught early, can be managed or reversed. In 1996, the AIP team at the MU developed and tested the AIP care coordination model in frail older community-dwelling adults [24,25]. When compared to older adults who received AIP care coordination and those who received traditional home health care, those in AIP care coordination were significantly older, Medicaid recipients, lived alone, had worse cognitive impairment and depression, and had worse activities of daily living (ADL) and instrumental ADL scores [26]. Yet when compared to outcomes of the home health care group, those in the AIP group had lower rates of ADL and instrumental activities of daily living (IADL) decline, fewer rehospitalizations, and lower provider and outpatient costs, but not in overall costs. In 2004, the community-based care model

XSL•FO
RenderX

transitioned to TigerPlace, a public or private partnership with Americare in an independent living with services facility [27]. TigerPlace was the first test site for in-home environmentally embedded sensors used for frail older adults to identify early illness and detect functional decline. Automatically generated sensor-based health alerts identify illness days or weeks earlier than people realize they are ill, allowing them to seek health care services earlier [28-31]. Additionally, living in an environment with sensor technology with health alerts added an average of almost 2 (1.7) years to older adults' length of stay in Independent Senior Housing. When combined with registered nurse (RN) care coordination, sensor technology with health alerts increased length of stay in independent senior housing by an additional 2.5 years, from 1.8 to 4.3 years [32]. A longitudinal analysis of the impact of AIP TigerPlace when compared to a reference group from 2011 to 2019 showed stable functional trajectories, with costs just above residential care costs [33]. While these prior initiatives demonstrated the potential of interdisciplinary care coordination and sensor-enabled monitoring, additional work was needed to develop a scalable, cost-feasible model capable of supporting individuals transitioning from institutional care back to community settings across geographically dispersed and resource-limited environments.

ASSETs for AIP

ASSETs for AIP used the conceptual framework described in the Agency for Healthcare Research and Quality (AHRQ) Care Coordination Atlas to operationalize the care coordination activities used by the interdisciplinary team [14]. The care coordination atlas identifies broad approaches to care coordination, including teamwork-based coordination and health IT-enabled coordination [14]. Care coordination activities provided by ASSETs broadly aligned with the 9 AHRQ care coordination domains, including (1) establishing accountability and negotiating responsibility; (2) communicating; (3) facilitating transitions; (4) assessing needs and goals; (5) creating a proactive plan of care; (6) monitoring, follow-up, and responding to change; (7) supporting self-management goals; (8) linking to community resources; and (9) aligning resources with client needs [14]. ASSETs for AIP uses a team-based care coordination informed by prior aging-in-place, sensor-based monitoring, and community-based intervention research. Services were delivered by an interdisciplinary team consisting of an RN, an occupational therapist (OT), and a social worker (SW). The RN addressed health-related concerns, medication management, and care coordination; the OT focused on home safety, functional independence, and environmental adaptations; and the SW facilitated access to community resources and provided psychosocial support. Team members collaborated closely, shared knowledge of all clients, and worked interchangeably as needed to ensure continuity of care.

ASSETs for AIP provided ongoing interdisciplinary care coordination. The program integrated standardized assessments, telehealth encounters, wearable devices, and ambient smart home monitoring to support early identification of changes in health (eg, heart rate), function (ambulation), and daily activity patterns (time spent in rooms). These data were reviewed by the care coordination team and used to inform client coaching to support client-identified goals for health and function. Additionally, clients were provided simple home safety equipment (eg, bath mats and movement-sensitive light switches) to improve their environment and make their homes more livable. All communication was intended to provide client support to enable them to communicate their needs effectively to service and health care providers so they could continue to live in their community homes and meet their personal goals.

Technology Overview

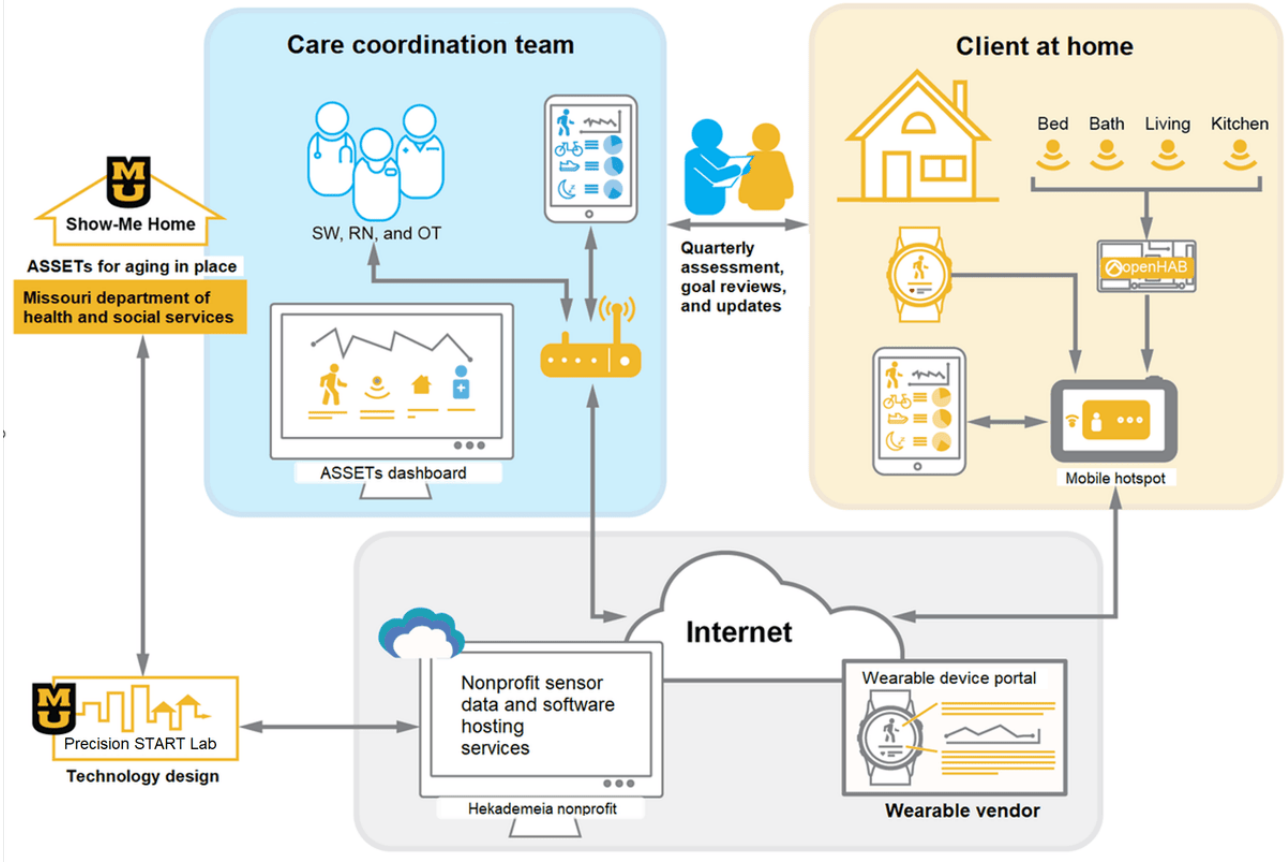
Overview

ASSETs for AIP technologies were selected, developed, configured, and evaluated at the Precision START laboratory [34-37] at the MU Sinclair School of Nursing during and after conceptualization of the program (Figure 1). This selection and test process relied on a *Translation by Design* approach [38] that prioritizes cost-feasibility and adaptability to maximize long-term sustainability and maintenance following implementation. To overcome sustainability issues experienced during previous sensor research to support AIP [39-42], our focus was on identifying free open-source software technologies and widely available consumer-grade devices selected through rapid evaluation processes [34,37]. Low-cost web-based application and data hosting and support for the smart home system are provided by a nonprofit organization created specifically for this purpose [43,44].

System components are a smart home sensor system, a smart home sensor dashboard and data services, a tablet device, and a wearable device.

XSL•FO
RenderX

Figure 1. Workflow of ASSETs for AIP (aging in place). ASSETs: Age-Friendly Smart, Sustainable, and Equitable Technologies; MU: University of Missouri; OT: occupational therapist; RN: registered nurse; SW: social worker.



Smart Home Sensor System

This technology component consisted of a low-cost Raspberry Pi computer running open-source home automation software (openHAB) connected to consumer-grade environmental sensors that capture room-level and light levels.

Smart Home Sensor Dashboard and Data Services

This technology component is a web-based application developed in-house and hosted by the nonprofit service provider [35]. Sensor data are uploaded to individual client accounts and viewed through a dashboard that supports monitoring of system status and home activity patterns. Dashboard information was used as 1 source of information to support follow-up decisions and was interpreted in conjunction with client reports, assessments, and clinical judgment rather than relying on predefined automated alert thresholds.

Tablet Device

Participants were provided with a low-cost tablet to support technology and enable telehealth visits with the remote care coordination team. For troubleshooting after deployment, we used the remote access feature of the device to troubleshoot sensor connection issues.

Wearable Device

Participants were provided with a consumer-grade wearable device (eg, smartwatch or smart ring) to support activity monitoring and self-management.

Internet Connectivity

Broadband internet access was provided through a wireless mobile hotspot when needed for program participation.

The ASSETs for the AIP demonstration project connected older adults at home to interdisciplinary care coordination across geographic distances using the digital health technologies above in the section Technology Overview.

Methods

Study Design

This is a quantitative, descriptive study with repeated measures over 12 months.

Recruitment

Client Eligibility

Eligibility criteria for the ASSETs for the AIP program included older adults and adults with disabilities who were within 1 year of discharge from the Show-Me Home program in Missouri. The Show-Me Home program serves both older adults and younger adults with disabilities transitioning from institutional care back into the community. Accordingly, the study population reflected the underlying population served by the Show-Me Home program. Administrators at the state informed the team about recent discharges and who should be contacted about enrollment into ASSETs for AIP. Those eligible were then contacted via phone to determine interest. No technology experience was required, and all equipment, connectivity, and

XSL•FO
RenderX

training were provided by the program. Participant recruitment and enrollment occurred on a rolling basis between April 2022 and December 2025. Participants were followed quarterly for the duration of their enrollment in the ASSETs for AIP.

Recruitment Strategies

A client list was provided by Show-Me Home to the ASSETs for the AIP team. This list was based on contact information collected at client discharge from the Show-Me Program. Once the team received the list, they began recruitment. Despite diligent efforts, the team encountered challenges including the following: (1) locating individuals who were more than one year postdischarge from a state-funded program; (2) working within a limited-service area; (3) once reached, getting clients to respond about a new, unknown program; and (4) inaccurate or obsolete contact information. Successful strategies included partnering with state and local agencies to retrieve up-to-date contact information, reducing the time from Show-Me Home program discharge to facilitate contact and align better with client needs, extending the service area beyond the mid-Missouri area, including larger urban areas, and sending out postcard mailers to potential clients that provided a clear linkage between the Show-Me Home program, ASSETs for AIP, and MU. Following implementation of the revised recruitment strategies, client enrollment increased by 225% over the subsequent 6 months. Participant flow from screening through final analysis is summarized in Table 1.

Table 1. Participant flow from screening through final analysis.

Participant flow stage	Values, n
Total number of individuals contacted	406
Screened	58
Excluded after screening	13
Enrolled	45
Discharge during study[a]	11
Participant request or withdrawal	5
Limited availability or relocation	1
Diminishing returns	2
Technology loss or noncompliance	2
Incarceration	1
Primary analytic cohort with complete longitudinal assessments	19
Enrolled for MFP[b] quality of outcome item C4 analysis only[c]	1
Enrolled but not included in the analysis (<4 time points at the time of analysis)	25

[a]Discharge during the study did not necessarily result in exclusion from analysis. Participants who were subsequently discharged were included in outcome-specific analyses if they had completed the admission assessment and 4 quarterly follow-up assessments for that outcome.

[b]MFP: Money Follows the Person.

[c]One additional participant outside the primary analytic cohort had complete longitudinal data (admission through quarter 4) for the unmet medical need outcome and was included only in that outcome-specific analysis.

Care Coordination Team Support

The RN, OT, and SW care coordination team collaborated to deliver coordinated care by conducting remote and in-home assessments, setting person-centered goals, and reviewing sensor and wearable data to monitor health status and daily function. The team conducted a full assessment upon client admission using REDCap surveys, an in-home visit for equipment setup, and a home safety checklist. A shorter assessment followed every quarter using selected items. Health data from ambient and wearable sensors, displayed on a cloud-based visual dashboard, were used by the team to monitor clients' activity patterns. The team reviewed the dashboards daily to respond quickly to changes, including reduced activity from ambient sensors or shifts in heart rate and step data from wearable devices. Using information from the assessment and sensor data, the team contacted the client to discuss changes, agreed on a plan of action when needed, and coached the client as to what the next steps could be. This approach let the team monitor the client from a distance and connect by tablet or phone when activity changes suggested the need for further evaluation. The care coordinators did not directly intervene with the clients and their individual health care team but rather coached the client as to what they should be doing next and followed up to support the client as needed.

Data Collection

The interRAI home care (HC) assessment [45-47] is a validated and standardized tool designed to evaluate the health, functional, and psychosocial status of individuals receiving HC services. It provides a comprehensive framework for care planning and monitoring outcomes, using validated scales to assess key domains such as daily functioning and cognition. Participants with sufficient longitudinal data for a given outcome, defined as an admission assessment and 4 quarterly follow-up assessments for that outcome, were included in the

XSL•FO
RenderX

corresponding analysis. Consequently, analytic sample sizes varied (range n=16-20) across outcomes based on outcome-specific data completeness. All interRAI HC and MFP QoL assessments were administered in person during scheduled quarterly client visits. Assessments were completed by members of the interdisciplinary care coordination team, including RNs, OTs, and SWs. Care coordinators received training on assessment administration and program procedures before conducting client assessments to promote consistency in data collection.

ADL Hierarchy Scale

The ADL Hierarchy Scale [46] reflects the disablement process by grouping ADL into stages of functional decline: early (personal hygiene), middle (locomotion, toileting), and late (eating). It uses 4 self-performance items coded over the prior 3 days, with scores ranging from 0 (independent) to 6 (total dependence), where higher scores indicate greater functional loss.

Cognitive Performance Scale

The Cognitive Performance Scale (CPS) [47-49] measures cognitive status using items related to decision-making, short-term memory, communication, and eating performance. It produces a score from 0 (intact) to 6 (very severe impairment). The CPS has been validated against cognitive assessments such as the Mini-Mental State Examination (MMSE) and the Test for Severe Impairment (TSI). The MFP QoL Survey was also administered quarterly.

Analysis

Overview

Given the exploratory nature of this demonstration project and the modest sample size, analyses were conducted to characterize cohort-level longitudinal patterns rather than evaluate program effectiveness or estimate individual-level trajectories. For the interRAI outcomes, client scores were summarized across all clients at each assessment period (admission and quarterly follow-ups), and weighted linear regression was used to examine overall cohort-level trends over time. For the MFP QoL measures, repeated-measures nonparametric tests were used to evaluate changes across assessment periods. The objective was to describe overall trends among clients with complete longitudinal data rather than estimate individual-level trajectories or intervention effects. Analyses were conducted using available data, and no imputation procedures were performed. Differences in outcome-specific sample sizes reflected item-level nonresponse within individual MFP QoL survey items.

interRAI: ADL Hierarchy Scale and CPS

ADL Hierarchy Scale and CPS scores were first summed across clients at each time point (n=19) to produce mean and SD values. To assess overall trends, we fitted weighted linear regression models of mean score on timepoint, with weights calculated as the inverse of the variance ($1/SD^2$) at each timepoint. Slope estimates and CIs were used to evaluate whether ADL or CPS scores showed systematic change over time.

interRAI: IADL Functional Outcomes

The interRAI functional items (C2a-d: managing finances, managing medications, managing use of electronic devices, and managing online shopping or ordering; G4a-e: meal preparation, ordinary housework, stairs, shopping, and transportation) were combined to represent overall functional independence in instrumental and domestic activities. At each assessment, each client received a composite IADL functional score by summing responses across the 9 interRAI functional items. The client-level composite scores were then averaged across all clients to obtain a single mean score for each assessment period (admission and quarters 1-4). Weighted linear regression was performed using these 5 mean scores, with weights calculated as the inverse of the variance ($1/SD^2$) at each timepoint, to characterize overall trends over time.

MFP QoL Survey

Analyses were conducted to examine changes in outcomes across 5 repeated assessments (admission and quarters 1-4). For ordinal and continuous outcomes, including satisfaction with living arrangements, life satisfaction, satisfaction with care received, choice and control of daily activities, and sleep quality, we used the Friedman test for nonparametric repeated measures. To adjust for multiple comparisons, a Bonferroni correction was applied across the 4 admission–quarter contrasts. For binary outcomes, including respect and dignity, unmet need for personal assistance, and unmet need for medical care, the Cochran Q test was used to assess overall change across time points. In addition, McNemar exact tests were performed for pairwise admission-quarter comparisons. All tests were 2-tailed with a significance threshold of $P<.05$.

Ethical Considerations

The study of ASSETs for AIP was approved under MU Health Sciences IRB (#2080782). Research consent was obtained via phone call, followed by an in-person home visit.

Results

Participant Characteristics

Nineteen clients comprised the primary analytic cohort (Table 2). Outcome-specific analytic sample sizes varied according to longitudinal data completeness, as described in the Methods. Most were aged 61 to 80 years, female, divorced, and White. Household income was low, with the majority reporting annual incomes between US $10,000 and US $19,999. Nearly all clients had Medicaid, and over two-thirds also had Medicare. Slightly more than half reported no availability of at-home health care assistance. Falls were uncommon across all assessed periods, with most clients reporting no falls in the prior 30, 31 to 90, or 91 to 180 days. Only 1 client experienced a fall with major consequences within the last 90 days. Comorbidity burden was high. Depression was the most prevalent condition, followed by type 2 diabetes, vascular disease, anxiety, and stroke or cerebrovascular accident. Multiple neurologic, cardiopulmonary, and psychiatric conditions were also reported, indicating substantial multimorbidity within the sample.

XSL•FO
RenderX

Table 2. Participant characteristics (N=19).

Variable	Values, n (%)
Age (years)[a]	
40-60	6 (32)
61-80	13 (68)
Gender	
Male	8 (42)
Female	11 (58)
Marital status	
Divorced	13 (68)
Widowed	4 (21)
Never married	2 (11)
Race	
White	14 (74)
Black or African American	4 (21)
Multiracial[b]	1 (5)
Annual household income (US $)	
<$10,000	2 (11)
$10,000-$19,999	12 (63)
$20,000-$29,999	4 (21)
Preferred not to answer	1 (5)
Health insurance[c]	
Medicaid	18 (95)
Medicare	13 (68)
Availability of at-home help for health care	
Yes	9 (47)
No	10 (53)
Falls	
In the last 30 days	
No falls	15 (79)
One fall	4 (21)
In the last 31-90 days	
No falls	15 (79)
One fall	3 (16)
Two or more falls	1 (5)
In the last 91-180 days	
No falls	15 (79)
One fall	3 (16)
Two or more falls	1 (5)
Any fall with major consequences within the last 90 days	
Yes	1 (5)
No	18 (95)
Morbidity[d]	

XSL•FO
RenderX

Variable	Values, n (%)
Psychiatric diseases	
Depression	11 (58)
Anxiety	7 (37)
Posttraumatic stress disorder	4 (21)
Bipolar disorder	2 (11)
Schizophrenia	1 (5)
Cardiometabolic diseases	
Type 2 diabetes mellitus	9 (47)
Vascular disease	8 (42)
Chronic obstructive pulmonary disorder	5 (26)
Coronary heart disease	3 (16)
Heart failure	3 (16)
CKD[e]	1 (5)
Neurologic diseases	
Stroke or CVA[f]	6 (32)
Hemiplegia	5 (26)
Aphasia	5 (26)
Neuropathy	4 (21)
Seizure	3 (16)
Traumatic brain injury	3 (16)
Multiple sclerosis	1 (5)
Other	
Cancer	3 (16)
Urinary tract infection during the last 30 days	2 (11)
Other fractures during the last 30 days	1 (5)
COVID-19 diagnosis in the last 90 days	1 (5)

[a]Mean 64, SD 10.7.

[b]Multiracial here includes the following races: Black or African American and American Indian or Alaska Native.

[c]Participants may have either Medicare and Medicaid or both.

[d]Participants had been diagnosed with the listed comorbidities; some were undergoing active treatment, while others were being monitored without active treatment.

[e]CKD: chronic kidney disease.

[f]CVA: cerebrovascular accident.

interRAI

ADL Hierarchy Scale

Mean ADL scores were very low at admission and remained low across all 5 time points (Figure 2 and Table 3). The fitted linear trend was essentially flat, with a slope estimate of −0.0185 per time point (SE 0.0738; t_3=−0.25, P=.82). The 95% CI for the slope (−0.253 to 0.216) included 0, indicating no evidence of a systematic increase or decrease in ADL scores over time. The overall conclusion is that ADL scores remained stable. The data are consistent with a small negative trend, no trend, or a small positive trend.

XSL•FO
RenderX

Figure 2. ADL Hierarchy Scale and Cognitive Performance Scale. ADL: activities of daily living; CPS: Cognitive Performance Scale.

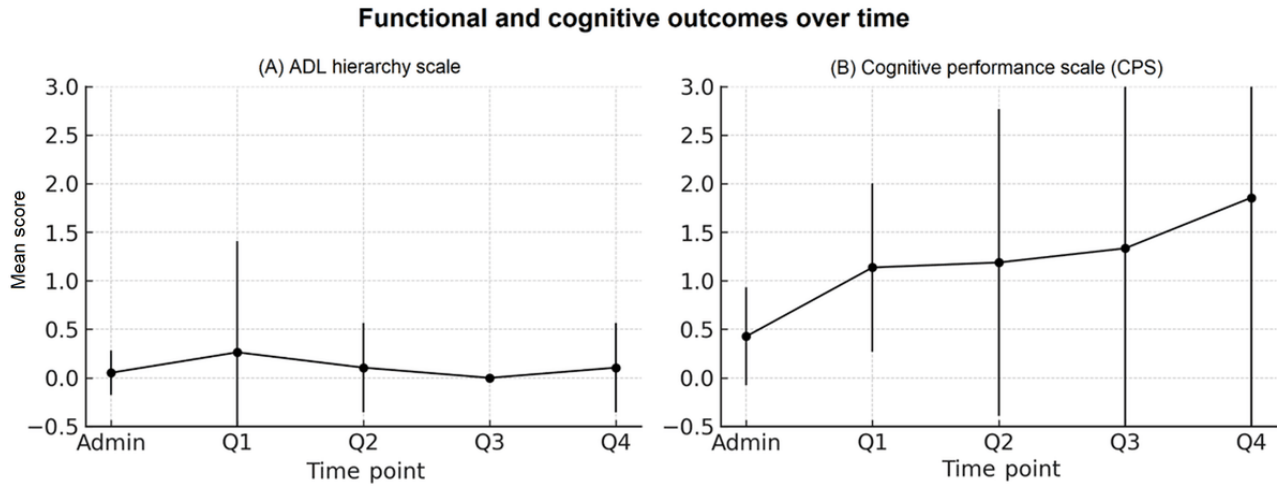


Table 3. Summary of interRAI subscales and IADL[a] outcomes across time points. Means and SDs are aggregated across clients at each time point (admission, quarters 1-4).

Measure	n	Admit	Q1[b]	Q2	Q3	Q4	Slope (per quarter)	
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Slope (95% CI)	*P* value
ADL[c] Hierarchy Scale	19	0.05 (0.23)	0.26 (1.15)	0.11 (0.46)	0.00 (0.00)	0.11 (0.46)	−0.0185 (−0.253 to 0.216)	.82
CPS[d]	19	0.43 (0.50)	1.14 (0.87)	1.19 (1.58)	1.33 (2.31)	1.85 (2.73)	0.390 (−1.03 to 1.81)	.45
interRAI functional items (C2[e], G4[f])	19	17.21 (11.17)	18.63 (11.66)	16.32 (11.76)	16.63 (10.98)	15.11 (11.23)	−0.608 (−11.90 to 10.68)	.875

[a]IADL: instrumental activities of daily living.

[b]Q: quarter.

[c]ADL: activities of daily living.

[d]CPS: Cognitive Performance Scale.

[e]C2: managing finances, medications, using electronic devices, and shopping or ordering online.

[f]G4: assessing instrumental activities of daily living capacity, including meal preparation, routine housework, stair navigation, shopping, and transportation.

CPS

Mean CPS scores increased from 0.43 (SD 0.50) at admission to 1.85 (SD 2.73) at quarter 4, with SDs also increasing across time points (Figure 2 and Table 3). A weighted regression of CPS scores on time indicated a positive slope estimate of 0.39 per time point (SE 0.45 t_3=0.88, *P*=.45). The 95% CI for the slope (−1.03 to 1.81) was wide and included 0, indicating no statistically reliable evidence of an upward trend. Overall, these data are compatible with no change or with small increases or decreases in CPS over time and do not support a conclusion of systematic cognitive decline during follow-up.

IADL Functional Items

The summed score across the 9 interRAI functional items (manage finances, manage medications, manage use of electronic devices, manage online shopping or ordering, meal preparation, housework, stairs, shopping, and transportation) showed mean values of 17.21 (SD 11.17) at admission, 18.63 (SD 11.66) at Q1, 16.32 (SD 11.76) at Q2, 16.63 (SD 10.98) at Q3, and 15.11 (SD 11.23) at Q4 (Table 3). The 9 functional items were summed to create an overall functional independence score, with higher scores indicating greater functional impairment. Weighted linear regression of mean score on time point yielded a slope of −0.61 per quarter (SE 3.55; t_3=−0.17, *P*=.875; 95% CI –11.90 to 10.68). These data show that there is no evidence of systematic improvement or decline in functional outcomes over time.

MFP QoL

Satisfaction With Living Arrangements

At admission, 15 of 19 (78.9%) clients reported being satisfied, and this proportion remained high through follow-up (Q1=84%, 16/19; Q2=84%, 16/19; Q3=90%, 17/19; and Q4=84%, 16/19; Table 4). A Friedman test indicated no significant differences across time (χ^2_4=3.34, *P*=.50). Paired Wilcoxon signed-rank tests comparing admission to each quarter were all nonsignificant after Bonferroni correction. Trends are displayed in Figure 3.

XSL•FO
RenderX

Table 4. Summary of MFP[a] quality of life outcomes across time points[b].

Domain	Admission	Q1[c]	Q2	Q3	Q4	Test	
	n/N (%)	n/N (%)	n/N (%)	n/N (%)	n/N (%)	Values	*P* value
A. Living arrangements	15/19 (79)	16/19 (84)	16/19 (84)	17/19 (90)	16/19 (84)	3.34[d]	.50
B. Life satisfaction	18/19 (95)	15/19 (79)	16/19 (84)	18/19 (95)	15/19 (79)	8.36[d]	.08
C1. Care satisfaction	18/19 (95)	15/19 (79)	17/19 (90)	17/19 (90)	18/19 (95)	3.75[d]	.44
C2. Respect and dignity	15/16 (94)	14/16 (88)	16/16 (100)	15/16 (94)	16/16 (100)	4.00[e]	.41
C3. Unmet personal assistance need	1/19 (5)	2/19 (10)	1/19 (5)	1/19 (5)	1/19 (5)	4.00[e]	.41
C4. Unmet medical need	5/20 (25)	4/20 (20)	1/20 (5)	1/20 (5)	0/20 (0)	10.44[e]	.03
D. Choice and control	5.97 (0.12)[f]	5.85 (0.49)[f]	5.94 (0.24)[f]	6.00 (0.00)[f]	6.00 (0.00)[f]	4.56[d]	.34
E. Sleep quality	16/19 (84)	16/19 (84)	17/19 (90)	18/19 (95)	18/19 (95)	7.56[d]	.11

[a]MFP: Money Follows the Person.

[b]Values are percentages of clients reporting the specified outcome. Analyses were conducted using available data without imputation. Sample sizes vary across outcomes because of outcome-specific missing data. Three participants did not complete the MFP quality of life C2 assessment; therefore, C2-derived outcomes included 16 participants. One additional participant had complete administrative privileges through Q4 data for the unmet medical needs outcome but did not have complete Q4 data for the other MFP quality of life domains; therefore, this outcome included 20 participants.

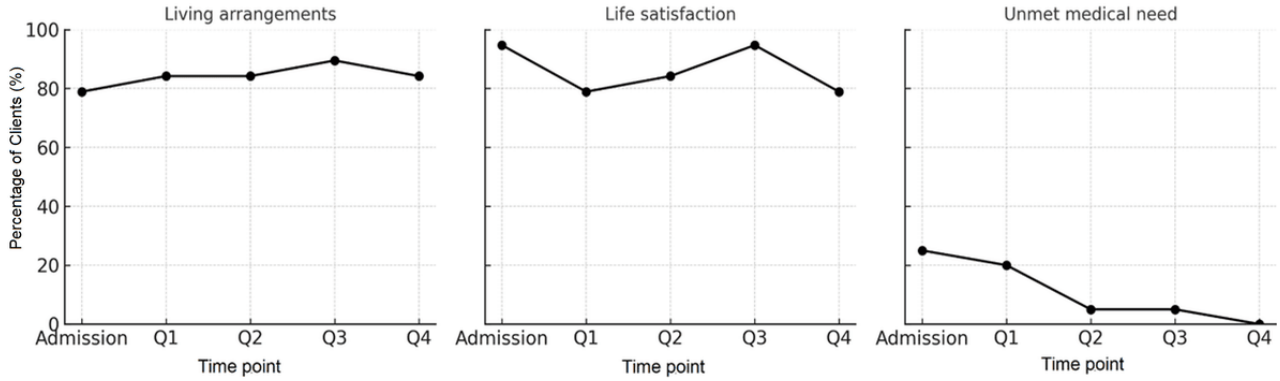
[c]Q: quarter.

[d]Friedman χ^2.

[e]Cochran Q.

[f]Mean (SD).

Figure 3. Key outcomes over time: living arrangements, life satisfaction, and unmet medical need.



Life Satisfaction

A Friedman test was performed but did not reach significance (χ^2_4=8.36, *P*=.08). Pairwise Wilcoxon signed-rank tests showed no significant admission-quarter differences after correction. Full descriptive results appear in Table 4 and Figure 3 for trends.

Satisfaction With Care Received

Proportions reporting happiness with care were consistently high: 95% (18/19) at admission, 79% (15/19) at Q1, 90% (17/19) at Q2, 90% (17/19) at Q3, and 95% (18/19) at Q4. The Friedman test showed no significant differences (χ^2_4=3.75, *P*=.44). Paired Wilcoxon comparisons were all nonsignificant.

Respect and Dignity

Among 16 clients with available responses, 15 (94%) reported being treated with respect and dignity at admission, 14 (88%) at Q1, 16 (100%) at Q2, 15 (94%) at Q3, and 16 (100%) at Q4. The Cochran Q test indicated no significant differences (Q(4)=4.00, *P*=.41). McNemar exact admission-quarter comparisons were all nonsignificant after Bonferroni correction.

Unmet Need for Personal Assistance

Unmet need was rare, with proportions of 5% (1/19) at admission, 10% (2/19) at Q1, 5% (1/19) at Q2, 5% (1/19) at Q3, and 5% (1/19) at Q4. The Cochran Q test revealed no significant differences (Q(4)=4.00, *P*=.41). McNemar tests were nonsignificant.

Unmet Need for Medical Care

At admission, 25% (5/20) of clients reported an unmet medical care need. This proportion declined over time to 20% (4/20) at Q1, 5% (1/20) at Q2, 5% (1/20) at Q3, and 0% (0/20) at Q4. The Cochran Q test indicated a significant decline across time

XSL•FO
RenderX

(Q(4)=10.44, *P*=.03). However, none of the McNemar admission-quarter comparisons reached significance after correction. Trends are displayed in Figure 3.

Choice and Control of Daily Activities

Mean choice-and-control scores were close to the maximum of 6 at all time points (Table 4). A Friedman test was not significant (χ^2_4=4.56, *P*=.34). Pairwise Wilcoxon tests were nonsignificant after correction.

Sleep Quality

At admission, 16 of 19 (84.2%) clients reported good sleep. Proportions remained high, with 84.2% (16/19) at Q1, 89.5% (17/19) at Q2, and 94.7% (18/19) at both Q3 and Q4. The Friedman test showed no significant differences across time (χ^2_4=7.56, *P*=.11). Pairwise Wilcoxon tests were nonsignificant after correction.

Discussion

Principal Findings

This study examined older and disabled adults enrolled in the ASSETs for AIP program following transition from nursing homes to the community in Missouri. We observed no statistically significant decline in measures of physical functions (ie, ADLs and IADLs) and cognition across follow-up, while the QoL indicators were generally high across assessment periods. The only domain showing a statistically detectable change was unmet medical need, which declined from 25% (5/20) at admission to 0% (0/20) by quarter 4, indicating a reduction in reported unmet care needs over time. These findings support the feasibility of technology-enabled care coordination following nursing-home discharge and suggest that clients maintained function, cognition, and QoL over follow-up.

Comparison With Prior Work

Prior research on postdischarge older adults often shows progressive decline in ADL and IADL function [50,51]. In our cohort, no statistically significant decline in functional outcomes was observed over the follow-up periods. Given the small sample and absence of a comparison group, caution in interpretation is required, as it was not possible to establish if observed longitudinal patterns were attributable to real-time monitoring or to care coordination activities. Similarly, no statistically significant decline in cognitive outcomes was observed in this cohort, although the study was not powered to detect small changes over time.

Unlike other established programs, the ASSETs for AIP program combined sensor data with a care coordination team, comprised of a RN, an OT, and a SW, who reviewed individualized data dashboards daily to identify deviations in activity, sleep, or physiological patterns and coached clients as to reasonable next steps. This “human-in-the-loop” approach transformed passive data collection into actionable insights. Deviations from clients’ established daily activity patterns triggered follow-up phone calls from care coordinators to investigate causes for change and coach clients to seek additional services when indicated. Commonly recommended services included health care services, mental health services, physical therapy, social assistance, durable medical equipment resources, and home modification services.

In contrast to studies emphasizing loss of autonomy or satisfaction when transitioning home, no statistically significant decline was observed across all major QoL domains. Participants reported high satisfaction and dignity scores, suggesting that the integration of continuous monitoring and structured care coordination did not compromise person-centered care. Care coordination activities informed by assessment, sensor, and wearable data included patient education, goal setting, reinforcement of positive health behaviors, and self-advocacy training. In this sense, ASSETs for AIP align with the emerging paradigm of digitally augmented care, where continuous monitoring supports, but does not replace, therapeutic human engagement [52].

The integration of sensor data with a comprehensive, interdisciplinary team is an emerging area focused on transforming health care from a visit-centric model to a continuous, preventative care model [53-55]. Since the study’s inception, remote physiological monitoring (RPM), remote therapeutic monitoring, and AI-driven predictive health platforms have emerged from a variety of clinical and commercial offerings [54,56,57]. Overall, the prior studies suggest that technology-enabled monitoring integrated with interdisciplinary care coordination may support care delivery across a range of settings.

Implications

Implications for research, policy, and practice emerge from these findings. Integration of sensor data with care-coordination workflows demonstrates that low-cost, open-source technologies can be effectively deployed in real-world aging contexts. Given the modest sample size (n=19) and the age distribution of participants (40-80 years), these findings should be interpreted as preliminary and may not generalize to broader aging populations, particularly adults over 80 years. Future research should evaluate this approach in larger, more age-diverse cohorts to establish its applicability across different populations and care settings. Future work should use causal designs that fit real-world rollout (eg, cluster randomized trials) to rigorously test effectiveness while allowing phased implementation. There is also a need to define and validate digital biomarkers from sensor data, such as changes in mobility or nocturnal bathroom trips, that can serve as early warning signals of health decline [58]. At the policy level, findings highlight the importance of broadband access as a health intervention, since connectivity is foundational for real-time monitoring and follow-up in rural areas. Equally important is developing reimbursement pathways that align sensor-based monitoring and care coordination with existing remote monitoring and care management payment streams to support sustainability [59,60].

From an implementation perspective, several factors contributed to the practical operation of the ASSETs for the AIP model. First, the program relied on cost-feasible technologies, including open-source software, consumer-grade devices, and low-cost computing infrastructure, which reduced barriers to deployment and long-term maintenance. Second, technology was integrated into existing interdisciplinary care coordination workflows

XSL•FO
RenderX

rather than functioning as a stand-alone monitoring system. Sensor and wearable data served as an additional source of information to support decision-making by nurses, OTs, and SWs. Finally, broadband connectivity and ongoing technical support remained important implementation considerations, particularly for rural and underserved populations. Future work should evaluate implementation outcomes such as adoption, acceptability, scalability, and sustainability to better understand how digitally enabled care coordination programs can be integrated into routine community-based care.

For practice, clear alert management protocols are essential to ensure that sensor signals translate into timely and appropriate action, with thresholds and escalation ladders that distinguish routine checks from urgent interventions. Closed-loop coordination, where every alert generates a task with an assigned owner and documented resolution, can build accountability, reduce missed follow-ups, and provide measurable outcomes for teams supporting AIP.

Limitations

These findings should be interpreted with consideration of limitations. The modest sample size, with outcome-specific analytic samples ranging from 16 to 20 participants, reduces statistical power and may mask subtle changes, particularly when only a few clients shift status between time points. The age distribution, with 6 participants aged 40-60 years and 13 aged 61-80 years, further limits generalizability, particularly to adults aged older than 80 years. Future studies should include larger samples with broader age representation to evaluate whether these findings generalize across older adult populations. The absence of a control group prevents causal inference, underscoring the need for future studies that include a comparison arm using traditional monitoring approaches. Several outcomes (ie, life satisfaction and living arrangements) on QoL showed ceiling effects, with most clients reporting the highest possible levels of satisfaction or autonomy at baseline, leaving little room for improvement. Statistical sensitivity was limited by the small sample and the use of conservative corrections, meaning that some nonsignificant findings should be interpreted as insufficient evidence of change rather than definitive evidence of stability. The aggregation of participant-level data into cohort-level means also does not account for within-person variation over time, further limiting statistical inference from the observed longitudinal patterns. Sensor data completeness was also affected by connectivity gaps, and technology acceptance varied across clients, both of which may have influenced outcomes.

Conclusions

This descriptive study of the ASSETs for AIP project demonstrates the feasibility of an innovative digital health program that relies on a team comprising a nurse, OT, and SW to deliver remote care coordination services enabled by cost-feasible and sustainable open-source and consumer-grade technologies. Specifically, the technology-enabled approach allowed care coordinators to reach clients across geographic distances, while smart home and wearable sensor data provided additional information to support decision-making. In addition, our pilot findings provide preliminary descriptive information regarding quality-of-life, functional, and cognitive outcomes among clients enrolled in the ASSETs for AIP program. Future larger studies are needed to evaluate effectiveness, implementation outcomes, and long-term sustainability while refining data-driven care coordination workflows for underserved populations.

Acknowledgments

We acknowledge the support of Tessa Wells from the University of Missouri Strategic Proposal Development Service for assistance with figure design. We also acknowledge the contribution of Gracey Jerusadan, MSN, for reviewing the results and compiling tables.

Funding

This study was supported by the US Department of Health and Human Services through the ASSETs for Aging in Place demonstration project (Missouri Department of Social Services and Centers for Medicare and Medicaid Services; 1LICMS300153 and 0069078). The funder had no involvement in this study's design, data collection, analysis, interpretation, or the writing of this paper.

Data Availability

The data used in this study were collected under a demonstration project funded by the US Department of Health and Human Services. The project was not designed for data sharing, and the data are subject to legal and regulatory restrictions. Therefore, the data are not publicly available.

Generative AI Use

The authors declare that generative AI was not used in the preparation of this manuscript.

XSL•FO
RenderX

Authors' Contributions

BR and LP conceptualized the project and methodology, led funding acquisition, and administered all aspects of the project. All authors were involved in project implementation and data collection. KL led the data analysis and wrote the original draft. SB, E Conrow, E Curtis, and ARD were involved in project implementation and data collection. All authors participated in writing, review, and approval of the manuscript.

Conflicts of Interest

BR is cofounder and Chair of the Board of Directors for Hekademeia Research Solutions, a nonprofit organization whose mission is to support technology-based academic research. This role is an unpaid volunteer position.

References

1. Gaugler JE, Duval S, Anderson KA, Kane RL. Predicting nursing home admission in the U.S: a meta-analysis. BMC Geriatr 2007;7:13 [FREE Full text] [doi: 10.1186/1471-2318-7-13] [Medline: 17578574]
2. Harris-Kojetin L, Sengupta M, Park-Lee E, Valverde R, Caffrey C, Rome V, et al. Long-term care providers and services users in the United States: data from the national study of long-term care providers, 2013-2014. Vital Health Stat 3 2016;38(x-xii):1-105 [FREE Full text] [Medline: 27023287]
3. Xu D, Garcia E, Arling G. Care trajectories for a cohort of first-time nursing home admissions: a 5-year longitudinal study in Minnesota. J Am Med Dir Assoc 2025;26(12):105885. [doi: 10.1016/j.jamda.2025.105885] [Medline: 41022138]
4. Sehatkarlangrodi S, Zargoush M, Longo CJ, Ghazalbash S. Shaping older adults' care policy: a scoping review of key determinants in post-acute and community reintegration transitions. BMC Health Serv Res 2025;25(1):1236 [FREE Full text] [doi: 10.1186/s12913-025-13433-x] [Medline: 41029370]
5. Sun M, Qian Y, Liu L, Wang J, Zhuansun M, Xu T, et al. Transition of care from hospital to home for older people with chronic diseases: a qualitative study of older patients' and health care providers' perspectives. Front Public Health 2023;11:1128885 [FREE Full text] [doi: 10.3389/fpubh.2023.1128885] [Medline: 37181713]
6. Oh A. Care coordination, social support, home health, and hospice: working towards keeping older adults with dementia in the community.: University of California, San Francisco; 2019. URL: https://escholarship.org/uc/item/5tz2f7d1 [accessed 2026-10-28]
7. Raunaq FF, Islam S, Anam MZ, Bari AM. Assessing the challenges to digital technology adoption in the healthcare sector: implications for sustainability in emerging economies. Inform Health 2025;2(2):194-209. [doi: 10.1016/j.infoh.2025.09.001]
8. Woldemariam MT, Jimma W. Adoption of electronic health record systems to enhance the quality of healthcare in low-income countries: a systematic review. BMJ Health Care Inform 2023;30(1):e100704 [FREE Full text] [doi: 10.1136/bmjhci-2022-100704] [Medline: 37308185]
9. Coleman A, Federman S. What Is Medicaid's value?. 2025 Jan 14. URL: https://www.commonwealthfund.org/publications/explainer/2025/jan/what-is-medicaids-value [accessed 2026-10-28]
10. Report to congress: best practices in money follows the person (MFP) demonstration. Mathematica. United States: United States Department of Health and Human Services and Centers for Medicare & Medicaid Services; 2024 Mar 12. URL: https://www.medicaid.gov/medicaid/long-term-services-supports/downloads/mfp-best-practices-rtc-appx-feb2024.pdf [accessed 2026-11-10]
11. Musumeci M, Chidambaram P, Watts MO. Medicaid's money follows the person program: state progress and uncertainty pending federal funding reauthorization. KFF. 2019 Nov 25. URL: https://www.kff.org/medicaid/issue-brief/medicaids-money-follows-the-person-program-state-progress-and-uncertainty-pending-federal-funding-reauthorization/ [accessed 2026-10-28]
12. Coughlin R, Ward J, Denny-Brown N. Money Follows the Person Demonstration: Overview of State Grantee Progress, January to December 2016. Mathematica. 2017. URL: https://www.mathematica.org/publications/money-follows-the-person-demonstration-overview-of-state-grantee-progress-january-to-december-2016 [accessed 2026-11-10]
13. Popejoy L, Reeder B, Robinson E, Proffitt R, Vogelsmeier A, Lee K. Development of a sensor-enhanced care coordination approach to facilitate aging in place. Innovation Aging 2023;7(Supplement_1):96-97. [doi: 10.1093/geroni/igad104.0312]
14. McDonald K, Schultz E, Albin L. Care coordination measures Atlas update. Agency for Healthcare Research and Quality. 2014 Jun. URL: https://www.ahrq.gov/ncepcr/care/coordination/atlas.html [accessed 2026-10-28]
15. Boult C, Green AF, Boult LB, Pacala JT, Snyder C, Leff B. Successful models of comprehensive care for older adults with chronic conditions: evidence for the institute of medicine's "retooling for an aging America" report. J Am Geriatr Soc 2009;57(12):2328-2337. [doi: 10.1111/j.1532-5415.2009.02571.x] [Medline: 20121991]
16. Gottlieb L, Gottlieb B. Strengths-based nursing: a process for implementing a philosophy into practice. J Fam Nurs (United States) 2017;23(3):319-340. [doi: 10.1177/1074840717717731] [Medline: 28795856]
17. Davey B, Levin E, Iliffe S, Kharicha K. Integrating health and social care: implications for joint working and community care outcomes for older people. J Interprofessional Care 2005;19(1):22-34. [doi: 10.1080/13561820400021734] [Medline: 15842078]

XSL•FO
RenderX

18. Szanton SL, Leff B, Li Q, Breysse J, Spoelstra S, Kell J, et al. CAPABLE program improves disability in multiple randomized trials. J Am Geriatr Soc 2021;69(12):3631-3640. [doi: 10.1111/jgs.17383] [Medline: 34314516]
19. Breysse J, Dixon S, Wilson J, Szanton S. Aging gracefully in place: an evaluation of the capability of the CAPABLE approach. J Appl Gerontol 2022;41(3):718-728. [doi: 10.1177/07334648211042606] [Medline: 34474609]
20. Szanton SL, Wolff JW, Leff B, Thorpe RJ, Tanner EK, Boyd C, et al. CAPABLE trial: a randomized controlled trial of nurse, occupational therapist and handyman to reduce disability among older adults: rationale and design. Contemp Clin Trials 2014;38(1):102-112 [FREE Full text] [doi: 10.1016/j.cct.2014.03.005] [Medline: 24685996]
21. Proffitt R, Robinson EL, Reeder B, Leary E, Botkin L, Marushak S, et al. Smart technology for aging and reducing disability (STAReD): study protocol for a randomized pragmatic clinical trial. Contemp Clin Trials 2024;138:107461. [doi: 10.1016/j.cct.2024.107461] [Medline: 38280484]
22. Flemming J, Armijo-Olivo S, Dennett L, Lapointe P, Robertson D, Wang J, et al. Enhanced home care interventions for community residing adults compared with usual care on health and cost-effectiveness outcomes: a systematic review. Am J Phys Med Rehabil 2021;100(9):906-917. [doi: 10.1097/PHM.0000000000001734] [Medline: 34415887]
23. Reeder B, Meyer E, Lazar A, Chaudhuri S, Thompson HJ, Demiris G. Framing the evidence for health smart homes and home-based consumer health technologies as a public health intervention for independent aging: a systematic review. Int J Med Inform 2013;82(7):565-579 [FREE Full text] [doi: 10.1016/j.ijmedinf.2013.03.007] [Medline: 23639263]
24. Marek KD, Adams SJ, Stetzer F, Popejoy L, Rantz M. The relationship of community-based nurse care coordination to costs in the Medicare and Medicaid programs. Res Nurs Health 2010;33(3):235-242 [FREE Full text] [doi: 10.1002/nur.20378] [Medline: 20499393]
25. Marek KD, Stetzer F, Adams SJ, Popejoy LL, Rantz M. Aging in place versus nursing home care: comparison of costs to Medicare and Medicaid. Res Gerontol Nurs 2012;5(2):123-129. [doi: 10.3928/19404921-20110802-01] [Medline: 21846081]
26. Popejoy LL, Galambos C, Stetzer F, Popescu M, Hicks L, Khalilia MA, et al. Comparing aging in place to home health care: impact of nurse care coordination on utilization and costs. Nurs Econ 2015;33(6):306-313 [FREE Full text] [Medline: 26845818]
27. Rantz M, Popejoy LL, Galambos C, Phillips LJ, Lane KR, Marek KD, et al. The continued success of registered nurse care coordination in a state evaluation of aging in place in senior housing. Nurs Outlook 2014;62(4):237-246. [doi: 10.1016/j.outlook.2014.02.005] [Medline: 24731918]
28. Phillips LJ, DeRoche CB, Rantz M, Alexander GL, Skubic M, Despins L, et al. Using embedded sensors in independent living to predict gait changes and falls. West J Nurs Res 2017;39(1):78-94 [FREE Full text] [doi: 10.1177/0193945916662027] [Medline: 27470677]
29. Rantz MJ, Skubic M, Koopman RJ, Alexander GL, Phillips L, Musterman K, et al. Automated technology to speed recognition of signs of illness in older adults. J Gerontol Nurs 2012;38(4):18-23 [FREE Full text] [doi: 10.3928/00989134-20120307-01] [Medline: 22420519]
30. Rantz MJ, Skubic M, Abbott C, Galambos C, Pak Y, Ho DK, et al. In-home fall risk assessment and detection sensor system. J Gerontol Nurs 2013;39(7):18-22 [FREE Full text] [doi: 10.3928/00989134-20130503-01] [Medline: 23675644]
31. Skubic M, Guevara RD, Rantz M. Automated health alerts using in-home sensor data for embedded health assessment. IEEE J Transl Eng Health Med 2015;3:2700111. [doi: 10.1109/jtehm.2015.2421499]
32. Rantz M, Lane K, Phillips LJ, Despins LA, Galambos C, Alexander GL, et al. Enhanced registered nurse care coordination with sensor technology: impact on length of stay and cost in aging in place housing. Nurs Outlook 2015;63(6):650-655. [doi: 10.1016/j.outlook.2015.08.004] [Medline: 26463735]
33. Popejoy L, Zaniletti I, Lane K, Anderson L, Miller S, Rantz M. Longitudinal analysis of aging in place at TigerPlace: resident function and well-being. Geriatr Nurs 2022;45:47-54. [doi: 10.1016/j.gerinurse.2022.02.030] [Medline: 35305514]
34. Boles KK, Gallimore MR, Howland C, Emezue C, Reeder B. Technical function evaluation of two smart wearables and data analysis methods for step counts. SSRN J 2023 [FREE Full text] [doi: 10.2139/ssrn.4240429]
35. Boles KK, Young L, Emezue C, Lee K, Popejoy L, Reeder B. Discovery of user requirements to support remote health coaching and care coordination in a CMS demonstration project. AMIA Annu Symp Proc 2025 May 22;2024:182-191. [Medline: 40417539]
36. Gallimore MR, Howland C, Chase JD, Grimsley A, Emezue C, Boles K, et al. Digital methodology for mobile clinical decision support development in long-term care. Stud Health Technol Inform 2021;290:479-483. [doi: 10.3233/SHTI220122] [Medline: 35673061]
37. Reeder B, Gallimore M, Howland C. Stepwise evaluation methodology for smart watch sensor function and usability. Cham: Springer International Publishing; 2021 Jul 3 Presented at: Augmented Cognition (HCII 2021); July 24-29, 2021; Virtual p. 221-233. [doi: 10.1007/978-3-030-78114-9_16]
38. Reeder B, Castaneda G. Towards using biometric transcripts for analyzing multimodal sensory information processing data. Cham: International Conference on Human-Computer Interaction. Springer; 2025 May 25 Presented at: Augmented Cognition (HCII 2025); June 22-27, 2025; Gothenburg, Sweden p. 152-165. [doi: 10.1007/978-3-031-93724-8_12]
39. Connelly K, Molchan H, Bidanta R, Siddh S, Lowens B, Caine K, et al. Evaluation framework for selecting wearable activity monitors for research. mHealth 2021 Jan 20;7:6 [FREE Full text] [doi: 10.21037/mhealth-19-253] [Medline: 33634189]

XSL•FO
RenderX

40. Reeder B, Chung J, Lazar A, Joe J, Demiris G, Thompson HJ. Testing a theory-based mobility monitoring protocol using in-home sensors: a feasibility study. Res Gerontol Nurs 2013;6(4):253-263 [FREE Full text] [doi: 10.3928/19404921-20130729-02] [Medline: 23938159]
41. Reeder B, Molchan H, Gutierrez E, Pena E, Caine KE, Demiris G, et al. HomeSHARE: implementing multi-site smart technology infrastructure. 2019 Presented at: AMIA 2019 Annual Symposium; November 16–20, 2019; Washington, DC URL: https://api.semanticscholar.org/CorpusID:218761598
42. Harrington C, Jelen B, Lazar A, Martin-Hammond A, Pradhan A, Reeder B, et al. Taking stock of the present and future of smart technologies for older adults and caregivers. arXiv Preprint posted online on Mar 31, 2021. [doi: 10.48550/arXiv.2104.00096]
43. Anderson B, Cameron J, Jefferson U, Reeder B. Designing a cloud-based system for affordable cyberinfrastructure to support software-based research. Stud Health Technol Inform 2022;290:489-493. [doi: 10.3233/SHTI220124] [Medline: 35673063]
44. Hekademeia Research Solutions. URL: https://hekademeia.org/ [accessed 2026-10-28]
45. Hirdes JP, Ljunggren G, Morris JN, Frijters DH, Finne Soveri H, Gray L, et al. Reliability of the interRAI suite of assessment instruments: a 12-country study of an integrated health information system. BMC Health Serv Res 2008;8(1):277 [FREE Full text] [doi: 10.1186/1472-6963-8-277] [Medline: 19115991]
46. Morris JN, Fries B, Bernabei R, Steel K, Carpenter I, Gilgen R, et al. interRAI home care (HC) assessment form and user's manual, version 9. interRAI 2009 [FREE Full text]
47. Carpenter I, Hirdes JP. Using interRAI assessment systems to measure and maintain quality of long-term care. In: A Good Life in Old Age? Monitoring and Improving Quality in Long-Term Care. Paris, France: OECD/European Commission; 2013:93-139.
48. Morris JN, Fries BE, Mehr DR, Hawes C, Phillips C, Mor V, et al. MDS cognitive performance scale. J Gerontol 1994;49(4):M174-M182. [doi: 10.1093/geronj/49.4.m174] [Medline: 8014392]
49. Morris JN, Howard EP, Steel K, Perlman C, Fries BE, Garms-Homolová V, et al. Updating the cognitive performance scale. J Geriatr Psychiatry Neurol 2016;29(1):47-55. [doi: 10.1177/0891988715598231] [Medline: 26251111]
50. Boyd CM, Landefeld CS, Counsell SR, Palmer RM, Fortinsky RH, Kresevic D, et al. Recovery of activities of daily living in older adults after hospitalization for acute medical illness. J Am Geriatr Soc 2008;56(12):2171-2179 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.02023.x] [Medline: 19093915]
51. Gill TM, Gahbauer EA, Han L, Allore HG. Trajectories of disability in the last year of life. N Engl J Med 2010;362(13):1173-1180 [FREE Full text] [doi: 10.1056/NEJMoa0909087] [Medline: 20357280]
52. Peek STM, Luijkx KG, Rijnaard MD, Nieboer ME, van der Voort CS, Aarts S, et al. Older adults' reasons for using technology while aging in place. Gerontology 2016;62(2):226-237 [FREE Full text] [doi: 10.1159/000430949] [Medline: 26044243]
53. Bodenheimer T, Ghorob A, Willard-Grace R, Grumbach K. The 10 building blocks of high-performing primary care. Ann Fam Med 2014;12(2):166-171 [FREE Full text] [doi: 10.1370/afm.1616] [Medline: 24615313]
54. Totten AM, Womack DM, Eden KB, McDonagh MS, Griffin JC, Grusing S, et al. Telehealth: mapping the evidence for patient outcomes from systematic reviews. 2016. URL: https://www.ncbi.nlm.nih.gov/books/NBK379320/ [accessed 2026-10-28]
55. Recommendations on digital interventions for health system strengthening. World Health Organization. 2019 Jun 6. URL: https://www.who.int/publications/i/item/9789241550505 [accessed 2026-10-28]
56. Kitsiou S, Paré G, Jaana M. Effects of home telemonitoring interventions on patients with chronic heart failure: an overview of systematic reviews. J Med Internet Res 2015;17(3):e63 [FREE Full text] [doi: 10.2196/jmir.4174] [Medline: 25768664]
57. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
58. Majumder S, Aghayi E, Noferesti M, Memarzadeh-Tehran H, Mondal T, Pang Z, et al. Smart homes for elderly healthcare-recent advances and research challenges. Sensors (Basel) 2017;17(11):2496 [FREE Full text] [doi: 10.3390/s17112496] [Medline: 29088123]
59. LaRose R, Gregg JL, Strover S, Straubhaar J, Carpenter S. Closing the rural broadband gap: promoting adoption of the internet in rural America. Telecommun Policy 2007;31(6-7):359-373. [doi: 10.1016/j.telpol.2007.04.004]
60. Prieger JE. The broadband digital divide and the economic benefits of mobile broadband for rural areas. Telecommun Policy 2013;37(6-7):483-502. [doi: 10.1016/j.telpol.2012.11.003]

Abbreviations

ADL: activities of daily living
AHRQ: Agency for Healthcare Research and Quality
AIP: aging in place
ASSETs: Age-Friendly Smart, Sustainable, and Equitable Technologies
CAPABLE: Community Aging in Place—Advancing Better Living for Elders

XSL•FO
RenderX

CPS: Cognitive Performance Scale
HC: home care
IADL: instrumental activities of daily living
MFP: Money Follows the Person
MMSE: Mini-Mental State Examination
MU: University of Missouri
openHAB: open-source home automation software
OT: occupational therapist
QoL: quality of life
RN: registered nurse
RPM: remote physiological monitoring
STAReD: Smart Technology for Aging and Reducing Disability
SW: social worker
TSI: Test for Severe Impairment

Edited by J Wang; submitted 15.Jan.2026; peer-reviewed by S Lin, Y Zhang; comments to author 08.Jun.2026; revised version received 19.Oct.2026; accepted 22.Oct.2026; published 24.Nov.2026.

Please cite as:
Lee K, Robinson EL, Vogelsmeier A, Reeder B, Bacon S, Conrow E, Curtis E, Roberts Dulany A, Perryman E, Popejoy L
ASSETs for Aging in Place: 12-Month Descriptive Longitudinal Study of Technology-Supported Interprofessional Care Coordination After Nursing Home Discharge
JMIR Aging 2026;9:e91524
URL: https://aging.jmir.org/2026/1/e91524
doi:10.2196/91524
PMID:

©Knoo Lee, Erin L Robinson, Amy Vogelsmeier, Blaine Reeder, Suzette Bacon, Elizabeth Conrow, Elizabeth Curtis, Ashley Roberts Dulany, Erica Perryman, Lori Popejoy. Originally published in JMIR Aging (https://aging.jmir.org), 24.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Alzheimer Disease and Related Dementias Caregiver Burden on Reddit: Large Language Model–Assisted Measurement Study

Yan Zhang, PhD; Mingfei Zhang, PhD

[1]Texas State University, 461 Trauth-Huffman Hall, 604 N Guadalupe St, San Marcos, TX, United States

Corresponding Author:
Yan Zhang, PhD
Texas State University, 461 Trauth-Huffman Hall, 604 N Guadalupe St, San Marcos, TX, United States

Abstract

Background: Informal caregivers of older adults with Alzheimer disease and related dementias (ADRD) often face high burden and chronic stress. Reddit provides naturally occurring caregiver narratives to study caregiving burden, but prior studies have often relied on small manual coding or aggregate topic models that may not easily connect caregiving context, burden, and stressors at scale.

Objective: This study aimed to explore what a large language model (LLM)–assisted measurement workflow can recover from ADRD caregiving-related Reddit discourse, such as caregiver burden, caregiving contexts, stressors, temporal patterns, and high-burden subgroups.

Methods: We processed monthly Reddit submission archives from January 2020 through June 2025. Posts were selected from 5 ADRD-specific subreddits and 4 broader caregiving or aging subreddits, yielding a balanced 13,200-post corpus. A locally hosted open-weight LLM inferred structured post-level variables, including caregiver sociodemographic characteristics, caregiver role, ADRD stage, care setting, support need, dominant emotion, sentiment, burden, and primary and secondary stressors. The workflow included schema validation, uncertainty ratings, field-level correction prompts, source matching, a 150-post double-coded human validation benchmark, and pairwise subgroup analyses.

Results: The source-matched analytic sample included 13,179 LLM-coded posts. Human agreement was high for caregiver role, ADRD stage, care setting, dominant emotion, support need type, burden group, and stressor categories (Cohen κ=0.77-1.00), but poor for the initial in-scope screen (κ=0.00). Adult children were the largest inferred caregiver group (8,266/13,179, 62.7%), followed by grandchildren (2,012/13,179, 15.3%). Moderate ADRD was the modal inferred stage (6,554/13,179, 49.7%), and home was the modal care setting (7,145/13,179, 54.2%). Mean burden increased from early-stage ADRD (5.38, SD 2.47) to moderate-stage ADRD (6.96, SD 2.17) and late-stage ADRD (7.77, SD 2.16). Behavioral strain was the dominant primary stressor, while family conflict, financial strain, emotional strain, and social isolation were leading secondary stressors. Pairwise subgroup analysis showed high burden in late-stage ADRD with emotional support needs (mean 8.42, SD 1.46), care transitions with emotional support needs (mean 8.45, SD 1.12), and late-stage care transitions (mean 8.40, SD 1.32).

Conclusions: LLM-assisted measurement can help convert unstructured ADRD caregiving-related Reddit discourse into post-level, theory-guided quantitative measures when paired with validation and cautious interpretation. This workflow extends prior online caregiver research by linking caregiving context, burden, and stressor pathways at scale. The findings highlight care transitions, emotional support needs, financial navigation, and family role negotiation as potential targets for caregiver-support interventions.

(JMIR Aging 2026;9:e103305) doi:10.2196/103305

KEYWORDS

Alzheimer disease; dementia; caregivers; social media; natural language processing; large language models; psychological distress; digital health

Introduction

In the United States, an estimated 7.4 million adults aged 65 years and older were living with clinical Alzheimer dementia in 2026, and 12.7 million family members and other unpaid caregivers provided care for people living with Alzheimer disease or related dementias (ADRD) in 2025 [1]. Caring for people with ADRD is usually long-term and demanding because caregivers must manage progressive cognitive decline, behavioral symptoms, functional dependence, and repeated care transitions [2]. Informal ADRD caregivers help support aging in place and reduce demand for formal long-term care [3,4], but caregiving can also expose them to social isolation, role overload, financial strain, family conflict, and health decline [5,6]. The stress process model developed by Pearlin et al [7]

XSL•FO
RenderX

describes stress as a process in which life strains, social roles, coping resources, and social support shape psychological distress. Its caregiving-specific extension distinguishes primary stressors arising directly from care work, secondary stressors that spill over into the caregiver's broader life, and resources that may buffer burden [8]. This framework is useful for ADRD caregiving because dementia-related symptoms such as wandering, sundowning, hallucinations, aggression, and care refusal can create both immediate care demands and downstream conflicts over money, work, and family.

Despite the importance of ADRD caregivers' experiences, conventional surveys and interviews may only partially capture the complexity of informal caregiver burden. Population surveys are essential for inference but often rely on relatively thin, pre-specified measures [9,10]. Qualitative interviews and focus groups provide depth but are usually small, difficult to replicate, and less suited to tracking change over time [11,12]. ADRD caregiving also involves stigmatized emotions and difficult moral concerns, including guilt, ambivalence, and resentment, that may be under-disclosed in structured interviews because of social desirability pressures [13]. Online communities therefore provide a complementary source of naturally occurring caregiver narratives. Recent studies using online platform data show that caregivers disclose health concerns, relationship strain, emotional problems, financial burden, support-seeking needs, and positive aspects of caregiving in these spaces [14-16]. These studies show the value of online narratives, but traditional manual coding methods remain limited in scale because they require intensive human review. For example, 2 studies by Huang et al [14,15] analyzed >700 Reddit posts and >3000 comments to summarize caregivers' stressors and experiences. These sample sizes are appropriate for close content analysis, but they make it harder to infer caregiving contexts and study subgroup comparisons.

Machine learning, traditional natural language processing (NLP), and broader text-as-data approaches have helped researchers analyze larger collections of online narratives. This literature shows that computational methods can convert large volumes of language into analyzable evidence, which can make large-scale analysis faster, less costly, and more reproducible when paired with careful validation [17]. In dementia caregiving research, sentiment analysis, supervised classification, and topic modeling have been used to study online posts. These studies have identified core themes in large online corpora, including family issues, care facilities, emotional distress, advice seeking, financial struggles, abuse concerns, and sleep problems [18-24]. For example, Cheng and Yin [18] used structural topic modeling to analyze Reddit and ALZConnected posts and found recurring needs related to family issues, disease stages, medical concerns, and care facilities. Ni et al [19] used a 2-stage machine learning framework to identify ADRD caregiving posts on Reddit and showed that caregivers discussed learning challenges, care experiences, and emotional support across ADRD and non-ADRD communities. This work is important because it moves beyond small samples and manual coding. However, many of these approaches are better suited to summarizing broad themes than to preserving more complex post-level information, such as how stressors cluster, how caregiver characteristics relate to stressors, or how burden varies across care contexts.

Large language models (LLMs) and generative AI tools are now being applied in caregiving and dementia research. Some studies have used LLMs to organize caregiving tasks, map digital caregiving functions, or build caregiver-facing support tools. For example, Kim et al [25] used ChatGPT 3.5 (OpenAI) to analyze caregiving education materials and mobile app descriptions across ADRD stages. Hasan et al [26] developed a conversational AI system for Alzheimer disease caregivers and found stronger usability and more relevant support than a baseline ChatGPT model in a user study. He et al [27] used GPT-4o in a 2-stage topic-modeling workflow for more than 230,000 dementia-caregiver tweets and found topics related to emotional burden, cognitive decline, family relationships, caregiving activities, and adaptation or support resources. These studies show the growing role of LLMs in dementia-caregiver research, but much of this work has focused on support tools, educational content, or aggregate topic discovery.

LLMs may also help address a long-standing scale-depth tradeoff in caregiver research. Manual qualitative coding offers depth but is slow and costly, while surveys and fixed-label classifiers offer scale but often reduce narratives to a small set of predefined measures [11,17]. Compared with traditional BERT (Bidirectional Encoder Representations from Transformers)-style classifiers, keyword filters, or bag-of-words topic models, LLMs can use broader context to identify caregiving contexts, burden, and distress even when caregivers use indirect language, euphemisms, or situational descriptions [28-30]. This may reduce reliance on a short list of predefined keywords and allow richer constructs than broad topics alone. LLMs can also provide a relatively fast and inexpensive way to apply structured interpretation to large text corpora while still preserving narrative detail [27]. They can function as measurement tools that convert unstructured text into structured quantitative data for subgroup comparison, such as caregiver role, demographic attributes, burden level, and stressor types [31]. In ADRD caregiving research, these advantages point to a remaining gap: prior studies have identified WHAT questions (main caregiving themes or patterns), but less is known about HOW questions with respect to how caregiver contexts are associated with their needs, burden, and stressors, and how large-scale narrative data can be used to identify the most vulnerable subgroups. LLM-assisted measurement may help describe these patterns more comprehensively and inform policy design, social implications, and caregiver support interventions.

To address these gaps, this study examines what an LLM-assisted measurement workflow can recover from ADRD caregiving-related Reddit discourse. Guided by the stress process model, we sought to move beyond aggregate topic discovery by measuring caregiving contexts, burden, sentiment, and primary and secondary stressors at the post level. Specifically, we asked the following research questions (RQs):

- RQ1: What caregiver contexts, such as roles, ADRD stages, care settings, support needs, sentiment levels, and burden levels, can be identified in ADRD caregiving posts on Reddit?

XSL•FO
RenderX

- RQ2: How do sentiment and burden differ across inferred caregiving contexts?
- RQ3: How do primary stressors and secondary stressors appear and pair together within posts?
- RQ4: How do temporal patterns in caregiving burden and sentiment change over time?
- RQ5: Which subgroups of caregiver contexts and stressor types are associated with especially high burden?

To answer these questions, we first assembled a long-term Reddit corpus from ADRD-specific subreddits and ADRD-keyword-filtered posts in broader caregiving communities. We then used a locally hosted instruction-tuned LLM to infer structured post-level variables, including demographic attributes, caregiver role, ADRD stage, care setting, sentiment, burden, and stressors. We next validated key fields with a human-coded benchmark, summarized confidence coverage, examined post-level patterns, assessed temporal trends, and identified high-burden subgroup intersections relevant to caregiver support. This study contributes a scalable workflow for measuring ADRD caregiving-related Reddit posts and shows how LLM-assisted measurement can support clinical aging research, digital health design, and future caregiver-support interventions.

Methods

Study Design

The workflow had 2 steps. Step 1 built the Reddit post corpus. Step 2 applied LLM coding to the retained posts.

Step 1: Data Collection and Post Extraction

In step 1, we constructed the source corpus from monthly Reddit submission archives provided through Pushshift, a large-scale Reddit data collection infrastructure widely used in social media research [32]. In the completed scrape, we processed 66 monthly archives spanning January 2020 through June 2025. For each monthly archive, we restricted the initial search space to 9 subreddits selected a priori for likely ADRD caregiving content. Five were ADRD-specific communities: r/Alzheimers, r/dementia, r/lewyBody, r/FrontotemporalDementia, and r/vascular_dementia. Posts from these communities were treated as potentially relevant. Four were broader aging or caregiving communities: r/AgingParents, r/CaregiverSupport, r/eldercare, and r/elderlycare. Posts from these broader communities were retained only if the title or body matched an ADRD keyword list covering disease names, cognitive or behavioral symptoms, and dementia-specific care terms. The subreddit selection was informed by prior Reddit caregiver and ADRD caregiver studies that sampled dementia-specific forums as well as broader aging and caregiver forums where dementia caregiving is discussed [14,15,23,33,34].

Across all 66 monthly archives, the pipeline scanned 2,318,164,645 submission records. Subreddit and keyword filtering yielded 41,612 candidate posts, including 33,264 posts from ADRD-specific subreddits and 8348 posts from general caregiving subreddits that satisfied the ADRD keyword filter. Before final ranking, the extraction script excluded posts that were deleted, removed, or empty (n=14,328) and posts with fewer than 50 characters (n=471), leaving 26,813 eligible posts. Candidate posts substantially increased over time, from 207 eligible posts in April 2020 to 1409 in March 2025. Table S1 in the Multimedia Appendix 1 displays the corpus funnel for post extraction.

To prevent later, higher-volume months from dominating the corpus, we retained the top 200 eligible posts by Reddit score within each month after filtering and post-ID deduplication. Because author metadata was not retained in the monthly extracted files, we could not estimate unique contributors or deduplicate by author. A conservative text-level audit identified 112 exact or near-duplicate posts (0.85%); because their removal had minimal impact on the results, they were retained. Accordingly, all reported percentages and means describe posts rather than unique caregivers and may include multiple posts from the same author. This produced a final step-1 corpus of 13,200 posts (200 posts×66 months), including 10,278 (77.9%) from ADRD-specific subreddits and 2922 (22.1%) from broader caregiving communities. The retained source corpus was concentrated in r/dementia (61.1%), followed by r/Alzheimers (16.8%), r/AgingParents (12.0%), and r/CaregiverSupport (8.5%); r/eldercare contributed 1.5%, and r/elderlycare contributed fewer than 0.1%. After LLM output was joined back to the final monthly source corpus by post-ID, 13,179 source-matched LLM-coded posts remained for analysis (99.8% of the retained source corpus). This ID-matching step excluded inference rows that did not correspond to the final retained monthly source files and prevented nonsource records from entering the analytic sample.

Step 2: LLM Inference Procedure

In step 2, we converted each retained Reddit post into a structured measurement record using a locally hosted open-weight instruction-tuned large language model trained by Google DeepMind (Gemma 4 26B A4B) [35,36]. The model was served locally through Ollama on researcher-controlled hardware using the quantized model identifier gemma4:26b-a4b-it-q4_K_M. Local deployment kept post text within the researchers' computing environment, enabled use of a fixed model version, and supported large-scale inference without per-request commercial API charges. The prompting framework treated the LLM as a narrative measurement instrument rather than a generator of free text, following recent work on LLMs as scalable text annotation and computational social science tools [29,30]. This differs from recent LLM dementia-caregiver studies that used LLMs for digital-caregiving task mapping, retrieval-augmented caregiver support, or aggregate topic modeling of caregiver social-media text [25-27]. Rather, each post was supplied with explicit instructions to infer attributes of the caregiver and to return a constrained structured output.

Informed by prior Reddit caregiver studies [14,15,18,20,23], we inferred 15 variables spanning 4 domains. First, caregiver sociodemographic variables included gender, age, race, education, marital status, and income. Second, caregiving-context variables included caregiver role, ADRD stage, and care setting. Third, stress-process variables included primary stressor and secondary stressor. Fourth, socioemotional

XSL•FO
RenderX

variables included dominant emotion, support need type, sentiment, and burden level.

We defined primary stressor and secondary stressor according to the Stress Process Model before running the LLM [8]. The prompt instructed the LLM to read each post and select a short phrase from the narrative that best matched the predefined concept. Primary stressor was defined as the main direct caregiving burden described in the post, such as symptom management, care refusal, supervision demands, or behavioral strain. Secondary stressor was defined as the main spillover burden affecting the caregiver's broader life, such as sibling conflict, financial strain, work disruption, or social isolation. To reduce variation across posts, both stressor fields were constrained to a single phrase of no more than 2 words, or (not available) N/A when no such stressor was clearly present. Dominant emotion was constrained to 1 of 7 categories: anger, sadness, fear, joy, bittersweet, hope, or neutral, following the emotion categories used in prior analysis of supportive communication among Reddit caregivers [15]. Support need type was constrained to 1 of 7 support categories plus N/A: emotional, informational, instrumental, financial, respite, medical, or legal.

Most inferred variables were assigned a 3-level confidence rating (low, mid, or high) intended to capture the degree to which the variable was explicitly stated versus contextually inferred. Confidence ratings were assigned for all sociodemographic variables, caregiving-context variables, support need type, sentiment, and burden level. However, dominant emotion and the 2 stressors were returned without a confidence rating because they were framed as direct summarizations of the narratives. Sentiment and burden level were constrained to a 0‐10 scale, with higher values indicating more positive tone and greater caregiver burden, respectively. The workflow recorded the model's structured label and a low/mid/high confidence category where applicable, but it did not collect explanatory rationales or feature attributions. Confidence therefore reflects the model's requested assessment of textual support; it should not be interpreted as a calibrated probability or an explanation of the internal reasoning that produced a label.

The full prompt template used for LLM inference is provided in Multimedia Appendix 2. An abbreviated prompt logic is summarized below:

> *Analyze each post as a caregiver narrative about ADRD. Infer caregiver attributes only, not care-recipient attributes. Return one structured record per post. Primary_Stressor=main direct caregiving stressor (1‐2 words or N/A) Secondary_Stressor=main spillover stressor (1‐2 words or N/A) Dominant_Emotion=anger/sadness/fear/joy/bittersweet/hope/neutral Support_Need_Type=Emotional/Informational/Instrumental/Financial/Respite/Medical/Legal/N/A Sentiment=0‐10 Burden_Level=0‐10*

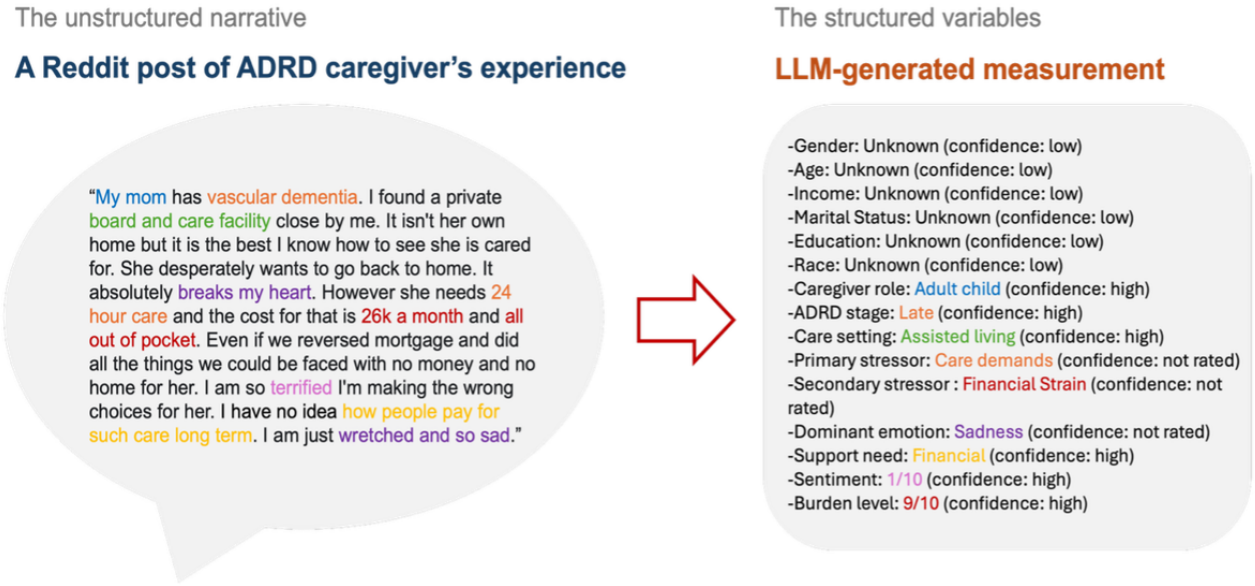
After the model returned a record, we checked whether the output followed the study rules. These checks included whether category labels came from the allowed list, confidence ratings were present when required, sentiment and burden scores stayed within the 0‐10 range, and stressor phrases were no longer than two words. If the whole record was unusable, we asked the model to process the post again. If only one field had a problem, we asked the model to correct only that field. We repeated this field-level correction up to 3 times. This process helped keep correct parts of a record while fixing specific errors.

This design had 2 purposes. First, it made the model output easier to analyze because each post had the same set of variables and the same allowed response options. Second, the confidence ratings helped separate information that was clearly stated in a post from information that required more interpretation. Records with lower confidence or repeated correction needs could then be treated more cautiously in later analyses.

Figure 1 illustrates how the LLM converted a Reddit caregiver narrative into structured inferred variables with confidence ratings.

XSL•FO
RenderX

Figure 1. Example of LLM (large language model)-assisted measurement converting an unstructured narrative into structured inferred variables. The narrative shown in the figure was rephrased to reduce traceability. ADRD: Alzheimer disease and related dementias.



Human Validation Benchmark

To assess the reliability of the LLM-derived measurement fields, we created a 150-post human validation benchmark from the source-matched analytic corpus. The benchmark was selected to include variations in support need, ADRD stage, caregiver role, and burden level. Two coders (YZ and MFZ) independently reviewed the post title and narrative and assigned labels for whether the post was in scope, caregiver role, ADRD stage, care setting, dominant emotion, burden level, support need type, and primary and secondary stressor categories. For each field, we calculated percent agreement and Cohen kappa between the 2 human coders. We then compared LLM labels with human consensus labels for fixed-label fields when both coders agreed, and the LLM output used the same label space. Open-vocabulary primary and secondary stressor phrases were not directly compared with consensus labels because the human benchmark coded broader stressor categories rather than exact LLM phrases.

Analytic Strategy

We conducted post-level descriptive and exploratory analyses after merging the LLM output with monthly metadata. We first examined inference coverage and confidence ratings to show which variables were usually stated clearly in posts. We then evaluated the human validation benchmark to assess whether the LLM-derived fields aligned with independent human coding. After these measurement checks, we described inferred caregiving contexts and burden variables captured in the Reddit corpus. These summaries provided the baseline portrait of who was posting in related subreddits, what care situations they described, what kinds of support they appeared to seek, and how much emotional strain was present.

We then examined whether burden and sentiment varied across nondemographic caregiving groups. This step linked the descriptive fields to caregiver strain by comparing mean burden and sentiment across caregiver contexts. Next, we analyzed primary and secondary stressors. Rather than assigning every post to a fixed stressor checklist, the LLM first summarized the main direct caregiving stressor and the main spillover stressor in each post as short phrases. We then grouped similar phrases after inference using token overlap, shared leading terms, and string similarity. Primary and secondary placements were retained separately so that we could distinguish stressors tied directly to care tasks from stressors that reflected broader effects on the caregiver's family, finances, work, or social life. We also examined quarterly temporal patterns in burden, sentiment, and caregiving contexts, using the post as the unit of analysis. This step was used to assess whether the emotional and practical content of posts changed over the 2020 - 2025 study period.

Finally, we used the structured variables to identify high-burden subgroup intersections that may be most relevant for support intervention. We examined all pairwise combinations of caregiver role, ADRD stage, care setting, and support need type. For each pair, we calculated the number of posts and mean burden for each subgroup and reported the highest-burden cell among those with more than 200 posts. We also examined the 3 most frequent primary stressor topics and the 3 most frequent secondary stressor topics, reporting subgroup patterns only when they had more than 200 posts and mean burden above 8.

Ethical Considerations

This study was reviewed by Texas State University Research Integrity and Compliance and was determined not to meet the criteria for human subjects research Institutional Review Board (IRB-26 - 161). The project involved a secondary analysis of publicly accessible Reddit data and did not involve the recruitment of human participants; therefore, informed consent and participant compensation were not required. To protect confidentiality and anonymity, the paper does not report usernames, URLs, post identifiers, or individual user-level information. Results are reported mainly in aggregate form at the post level, and post excerpts were rephrased and shortened to reduce traceability. These procedures are consistent with ethical approaches used in published Reddit caregiver studies

XSL•FO
RenderX

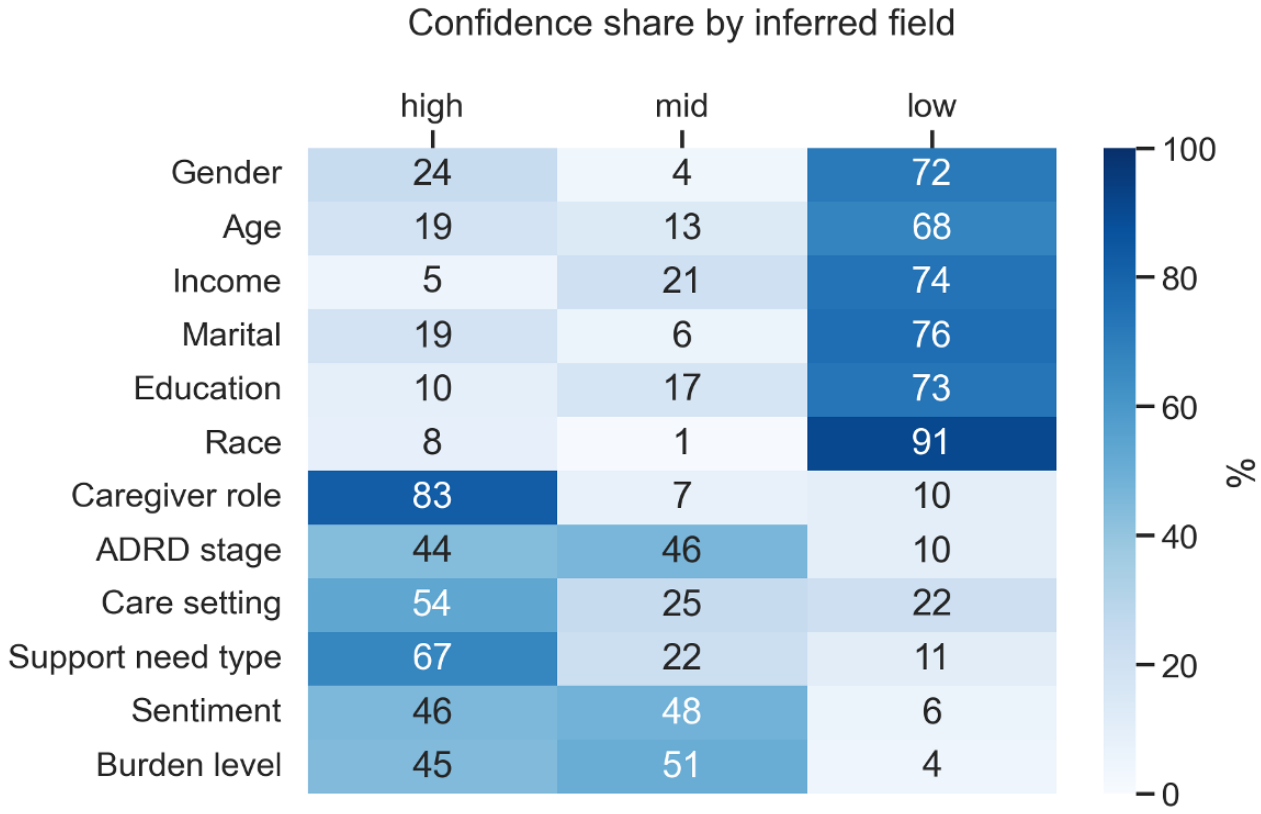
involving public data and IRB-exempt or secondary-data designs [14,20].

Results

Inference Coverage

A high proportion of sociodemographic fields were coded as low confidence, including gender, age, income, marital status, education, and race (8962/13,179, 68% to 11,993/13,179, 91%) (Figure 2). By contrast, several variables were more consistently identified from narratives: caregiver role, ADRD stage, care setting, and support need type had combined mid- and high-confidence shares of 90% (11,861/13,179), 90% (11,861/13,179), 78% (10,280/13,179), and 89% (11,729/13,179), respectively. Sentiment and burden level also had relatively high confidence coverage, with low-confidence ratings in only 6% (791/13,179) and 4% (527/13,179) of posts, respectively. In the following analysis, we skipped reporting analysis by demographic characteristics as they show low confidence ratings.

Figure 2. Confidence coverage across inferred variables. ADRD: Alzheimer disease and related dementias.



Human Validation of LLM-Derived Measures

Human coders had perfect agreement for caregiver role, ADRD stage, care setting, and dominant emotion (all 100.0% agreement; Cohen κ=1.00). Agreement was also high for support need type (99.3%; κ=0.99) and burden level group (97.3%; κ=0.94). Agreement for stressor categories was lower but still substantial: 80.7% agreement for primary stressor category (κ=0.77) and 84.0% for secondary stressor category (κ=0.80). These results support the reliability of the structured constructs used in the main analyses, while also showing that stressor category assignment remains more interpretive than role, stage, setting, emotion, support need, or burden coding.

The main validation weakness was the initial in-scope screen. One human coder (MFZ) applied the broad "from or about an informal caregiver" criterion to all 150 benchmark posts, whereas the second coder applied a stricter authorship threshold and labeled 87 posts "yes," 49 "no," and 14 "unclear." Thus, human coders agreed on only 58.0% of in-scope labels (κ=0.00), with disagreement concentrated in posts that were clinically or dementia-relevant but less clearly first-person informal caregiver narratives. This finding is consistent with the extraction design: posts from ADRD-specific subreddits were retained without a caregiver-perspective keyword requirement, which increased sensitivity but admitted some general discussion, patient-perspective, news, or ambiguous community posts. Accordingly, corpus-level proportions should be interpreted as describing a high-engagement ADRD caregiving corpus rather than a sample of confirmed caregiver-authored posts. The benchmark was diagnostic; no adjudication or post hoc in-scope filtering was used to alter the 13,179-post analytic corpus.

For fixed-label fields with human consensus labels, the LLM labels matched the human consensus labels in all compared cases: caregiver role (150/150), ADRD stage (150/150), care setting (150/150), dominant emotion (150/150), support need type (149/149), and burden level group (146/146). Because primary and secondary stressors were extracted as open-vocabulary phrases and humans coded broader categories, LLM–human agreement for stressors was evaluated indirectly through the human category reliability results rather than exact

XSL•FO
RenderX

phrase matching. Table A2 in the Multimedia Appendix 3 presents agreement statistics for the 150-post double-coded validation benchmark.

Inferred Caregiving Contexts, Emotion, Sentiment, and Burden

Table 1 presents the inferred caregiving contexts, emotional tone, sentiment, and burden level in the source-matched LLM-coded corpus. Among 13,179 posts, caregiver role was dominated by adult children (8266/13,179, 62.7%), followed by grandchildren (2012/13,179, 15.3%) and spouses (523/13,179, 4.0%). Smaller role groups included formal caregivers (202/13,179, 1.5%), friends (75/13,179, 0.6%), and siblings (57/13,179, 0.4%), while caregiver role was unknown in 15.5% (2044/13,179) of posts. ADRD stage was most often coded as moderate (6554/13,179, 49.7%), followed by late stage (2935/13,179, 22.3%) and early stage (1675/13,179, 12.7%); stage was unknown in 15.3% (2015/13,179) of posts. Care recipients were most often described as living at home (7145/13,179, 54.2%). Care transitions were identified in 6.8% (891/13,179) of posts. Institutional settings included memory care (728/13,179, 5.5%), nursing homes (531/13,179, 4.0%), and assisted living (452/13,179, 3.4%), while care setting was unknown in 26.0% (3432/13,179) of posts. The modal combination of adult-child caregivers and moderate-stage ADRD likely suggests that much of the Reddit disclosure captured by this workflow occurs during the period when cognitive and behavioral symptoms are prominent, but care has not yet fully shifted to institutional settings.

XSL•FO
RenderX

Table . Inferred characteristics in the source-matched LLM-coded[a] corpus (N=13,179).

Variable and categories/statistics	Values
Caregiver role, n (%)	
Adult child	8266 (62.7)
Grandchild	2012 (15.3)
Spouse	523 (4.0)
Formal caregiver	202 (1.5)
Friend	75 (0.6)
Sibling	57 (0.4)
Unknown	2044 (15.5)
ADRD[b] stage, n (%)	
Early	1675 (12.7)
Moderate	6554 (49.7)
Late	2935 (22.3)
Unknown	2015 (15.3)
Care setting, n (%)	
Home	7145 (54.2)
Transition	891 (6.8)
Memory care	728 (5.5)
Nursing home	531 (4.0)
Assisted living	452 (3.4)
Unknown	3432 (26)
Support need, n (%)	
Informational	5843 (44.3)
Emotional	4453 (33.8)
Respite	190 (1.4)
Legal	131 (1.0)
Medical	48 (0.4)
Financial	27 (0.2)
Instrumental	18 (0.1)
No explicit support needed	2469 (18.7)
Dominant emotion, n (%)	
Sadness	3839 (29.1)
Fear	3364 (25.5)
Anger	2273 (17.2)
Bittersweet	1655 (12.6)
Hope/joy	722 (5.5)
Neutral	1326 (10.1)
Sentiment[c], mean (SD)	2.61 (1.96)
Burden level[c], mean (SD)	6.53 (2.70)

[a]LLM: large language model.

[b]ADRD: Alzheimer disease and related dementias.

[c]N=13,179.

XSL•FO
RenderX

Support needs were primarily informational or emotional. Informational support was the dominant type (5843/13,179, 44.3%), followed by emotional support (4453/13,179, 33.8%). Posts with no explicit support needed accounted for 18.7% (2469/13,179). Explicit respite (190/13,179, 1.4%), legal (131/13,179, 1.0%), medical (48/13,179, 0.4%), financial (27/13,179, 0.2%), and instrumental (18/13,179, 0.1%) support needs were much less frequent. Table 1 also shows that the corpus was mainly negative in emotional tone. Sadness was the most common dominant emotion (3839/13,179, 29.1%), followed by fear (3364/13,179, 25.5%), anger (2273/13,179, 17.2%), bittersweet emotion (1655/13,179, 12.6%), neutral tone (1326/13,179, 10.1%), and hope or joy (722/13,179, 5.5%). The mean sentiment score was 2.61 (SD 1.96) on a 0 - 10 scale, where higher values indicate more positive tone. The mean burden score was 6.53 (SD 2.7) on a 0 - 10 scale, where higher values indicate greater caregiver burden.

Sentiment and Burden by Nondemographic Groups

Table 2 presents mean sentiment and mean burden scores across all nondemographic inferred variable groups. By caregiver role, adult-child posts had the highest mean burden score of 7.02 (SD 2.29) and the lowest mean sentiment score of 2.40 (SD 1.79), spouse posts had a similar mean burden score of 7.00 (SD 2.31) and mean sentiment of 2.47 (SD 1.84), and grandchild posts had a mean burden score of 6.50 (SD 2.58) and mean sentiment of 2.56 (SD 1.95). Posts coded as formal caregivers, friends, or unknown caregiver role had lower mean burden scores (formal caregivers: 4.01, SD 3.3; friends: 4.71, SD 2.72; and unknown: 4.76, SD 3.39), respectively, while sibling posts had a mean burden score of 6.70 (SD 2.76).

XSL•FO
RenderX

Table . Mean inferred sentiment and burden by nondemographic inferred variable group.

Group	Number of posts	Mean sentiment (SD)	Mean burden (SD)
Caregiver role			
Adult child	8266	2.4 (1.79)	7.02 (2.29)
Grandchild	2012	2.56 (1.95)	6.5 (2.58)
Spouse	523	2.47 (1.84)	7 (2.31)
Formal caregiver	202	4.13 (2.50)	4.01 (3.30)
Friend	75	2.85 (1.64)	4.71 (2.72)
Sibling	57	2.47 (1.88)	6.7 (2.76)
Unknown	2044	3.43 (2.30)	4.76 (3.39)
ADRD[a] stage			
Early	1675	2.88 (1.74)	5.38 (2.47)
Moderate	6554	2.45 (1.79)	6.96 (2.17)
Late	2935	2.05 (1.72)	7.77 (2.16)
Unknown[b]	2015	3.76 (2.45)	4.27 (3.39)
Care setting			
Home	7145	2.36 (1.73)	7.01 (2.29)
Transition	891	2.07 (1.49)	7.82 (1.67)
Memory care	728	2.84 (2.16)	6.59 (2.48)
Nursing home	531	2.23 (1.74)	7.17 (2.31)
Assisted living	452	2.57 (1.84)	6.79 (2.29)
Unknown	3432	3.3 (2.31)	5.05 (3.22)
Support need			
Informational	5843	2.49 (1.37)	6.22 (2.43)
Emotional	4453	1.63 (1.10)	8.02 (1.60)
Respite	190	1.26 (1.28)	9.16 (0.91)
Legal	131	1.67 (1.03)	7.92 (1.67)
Medical	48	1.85 (1.09)	7.58 (1.90)
Financial	27	1.22 (0.75)	8.67 (1.14)
Instrumental	18	1.72 (1.49)	8.22 (1.86)
No explicit support need	2469	4.86 (2.58)	4.24 (3.13)
Dominant emotion			
Sadness	3839	1.49 (0.72)	7.99 (1.59)
Fear	3364	2.07 (0.82)	6.72 (2.10)
Anger	2273	1.48 (0.77)	7.95 (1.63)
Bittersweet	1655	4.18 (1.94)	5.55 (2.60)
Hope	314	6.66 (1.58)	2.9 (2.39)
Joy	408	8.16 (0.96)	2.51 (2.01)
Neutral	1326	4.57 (1.28)	2.69 (2.16)

[a]ADRD: Alzheimer's disease and related dementias.

[b]Unknown denotes insufficient context and should not be read as a clinical category or absence of burden.

According to Table 2, burden increased across inferred ADRD stage. Early-stage posts had a mean burden score of 5.38 (2.47) and mean sentiment of 2.88 (SD 1.74), moderate-stage posts had a mean burden score of 6.96 (SD 2.17) and mean sentiment of 2.45 (SD 1.79), and late-stage posts had the highest mean burden score of 7.77 (SD 2.16) and lowest mean sentiment

XSL•FO
RenderX

among stage groups (2.05, SD 1.72). Unknown-stage posts had lower burden (4.27, SD 3.39) and higher sentiment (3.76, SD 2.45), suggesting that posts without stage cues often contained less specific or less acute caregiving narratives. Burden also varied by care setting. Transition posts had the highest mean burden score (7.82, SD 1.67) and low mean sentiment (2.07, SD 1.49), followed by nursing home (7.17, SD 2.31), home care (7.01, SD 2.29), assisted living (6.79, SD 2.29), and memory care (6.59, SD 2.48). Posts with unknown care setting had a lower mean burden score (5.05) and higher mean sentiment (3.30). Note that posts coded as Unknown for ADRD stage (mean burden 4.27, SD 3.39) or care setting (mean burden 5.05, SD 3.22) lacked sufficient contextual evidence for those classifications. Because stage, setting, and burden were inferred from the same narrative, sparse posts may also have contained fewer explicit burden cues. These lower means may therefore partly reflect information availability or shared-method measurement and should not be interpreted as evidence that an Unknown group is absence of burden.

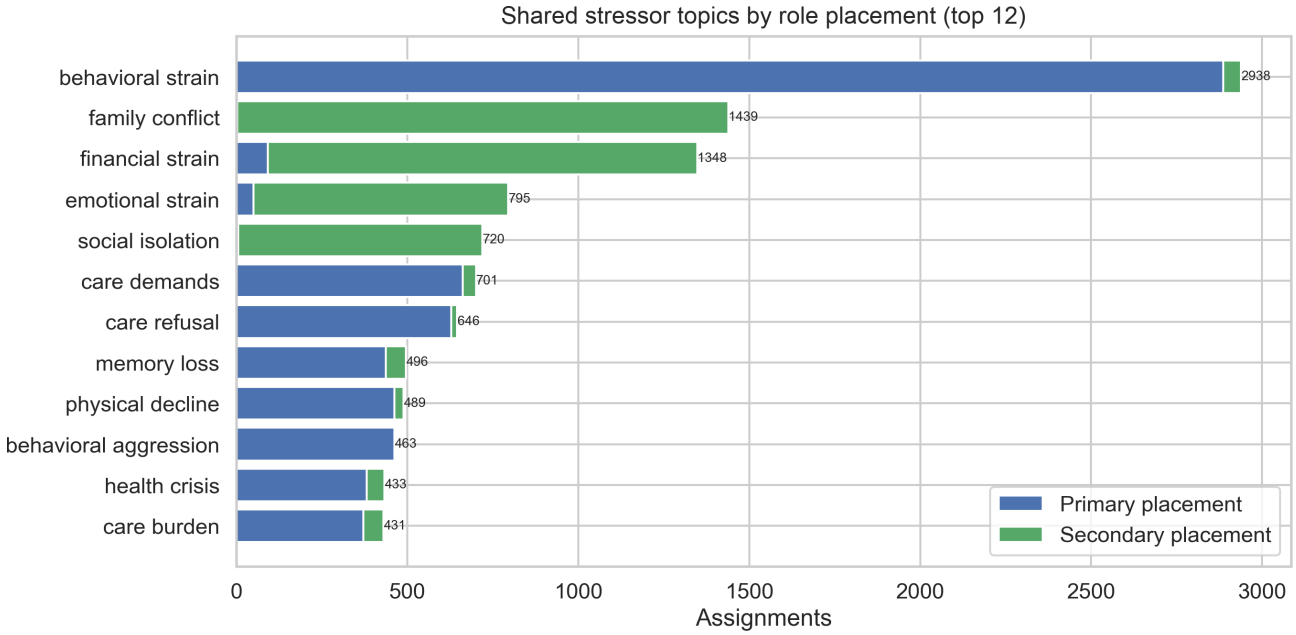
Support need type showed some of the largest differences in Table 2. Posts seeking respite support had the highest mean burden score (9.16, SD 0.91) and low mean sentiment (1.26, SD 1.28), followed by financial support needs (8.67, SD 1.14), instrumental support needs (8.22, SD 1.86), emotional support needs (8.02, SD 1.6), legal support needs (7.92, SD 1.67), and medical support needs (7.58, SD 1.9). Informational support posts had a lower mean burden score (6.22, SD 2.43) and higher mean sentiment (2.49, SD 1.37). Posts with no explicit support need had the lowest mean burden score among support need groups (4.24, SD 3.13) and the highest mean sentiment score (4.86, SD 2.58).

Dominant emotion was also strongly aligned with burden and sentiment. Posts coded as sadness had a mean burden score of 7.99 (SD 1.59) and mean sentiment of 1.49 (SD 0.72), and anger posts had a mean burden score of 7.95 (SD 1.63) and mean sentiment of 1.48 (SD 0.77). Fear posts had moderately high burden (6.72, SD 2.10) and low sentiment (2.07, SD 0.82). In contrast, bittersweet posts had intermediate burden (5.55, SD 2.6) and higher sentiment (4.18, SD 1.94), while hope, joy, and neutral posts had lower burden scores (2.90, SD 2.39, 2.51, SD 2.01, and 2.69, SD 2.16, respectively) and higher sentiment scores (6.66, SD 1.58, 8.16, SD 0.96, and 4.57, SD 1.28, respectively). This pattern supports the internal consistency of the inferred measures: negative emotions were associated with higher burden and lower sentiment, whereas positive or neutral emotions were associated with lower burden and higher sentiment.

Primary and Secondary Stressors

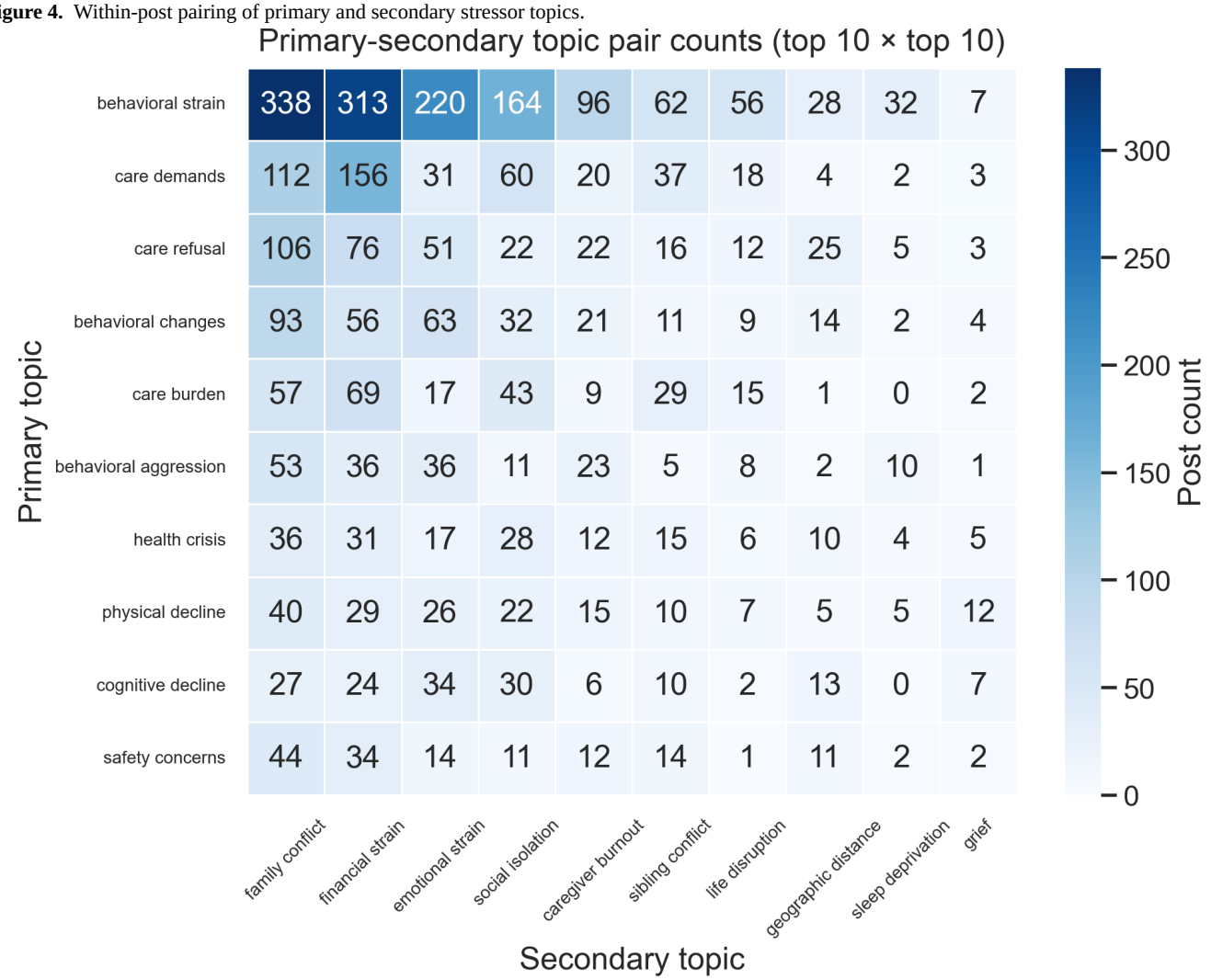
The open-vocabulary stressor extraction produced a broad set of short caregiver-burden phrases. Figure 3 presents clustered stressor topics by primary versus secondary placement, and Figure 4 presents the most common within-post primary-secondary pairings. A primary stressor was identified in 10,961 posts (83.2%), whereas a secondary stressor was identified in 7398 posts (56.1%).

Figure 3. Shared stressor topics by primary versus secondary placement.



XSL•FO
RenderX

Figure 4. Within-post pairing of primary and secondary stressor topics.



Behavioral strain was the dominant clustered primary stressor theme (n=2272), followed by behavioral changes (n=709), care demands (n=663), care refusal (n=629), physical decline (n=462), memory loss (n=418), cognitive decline (n=398), health crisis (n=382), behavioral aggression (n=372), and care burden (n=370). Other clinically recognizable primary stressors included incontinence management, safety concerns, communication difficulty, disease progression, wandering, medical decisions, repetitive questioning, hallucinations, end-of-life concerns, delusional behavior, and placement transition.

Figure 3 also shows that secondary stressors were led by family conflict (n=1436 mentions), financial strain (n=1255), emotional strain (n=744), social isolation (n=714), caregiver burnout (n=354), sibling conflict (n=336), life disruption (n=213), geographic distance (n=209), sleep deprivation (n=122), grief (n=90), mental health concerns (n=73), and legal issues (n=72). The secondary stressor distribution suggests that caregivers' Reddit disclosures often move from symptom-management problems to practical and relational consequences: who pays, who helps, who refuses to help, how caregiving constrains the caregiver's own life, and how distance complicates supervision.

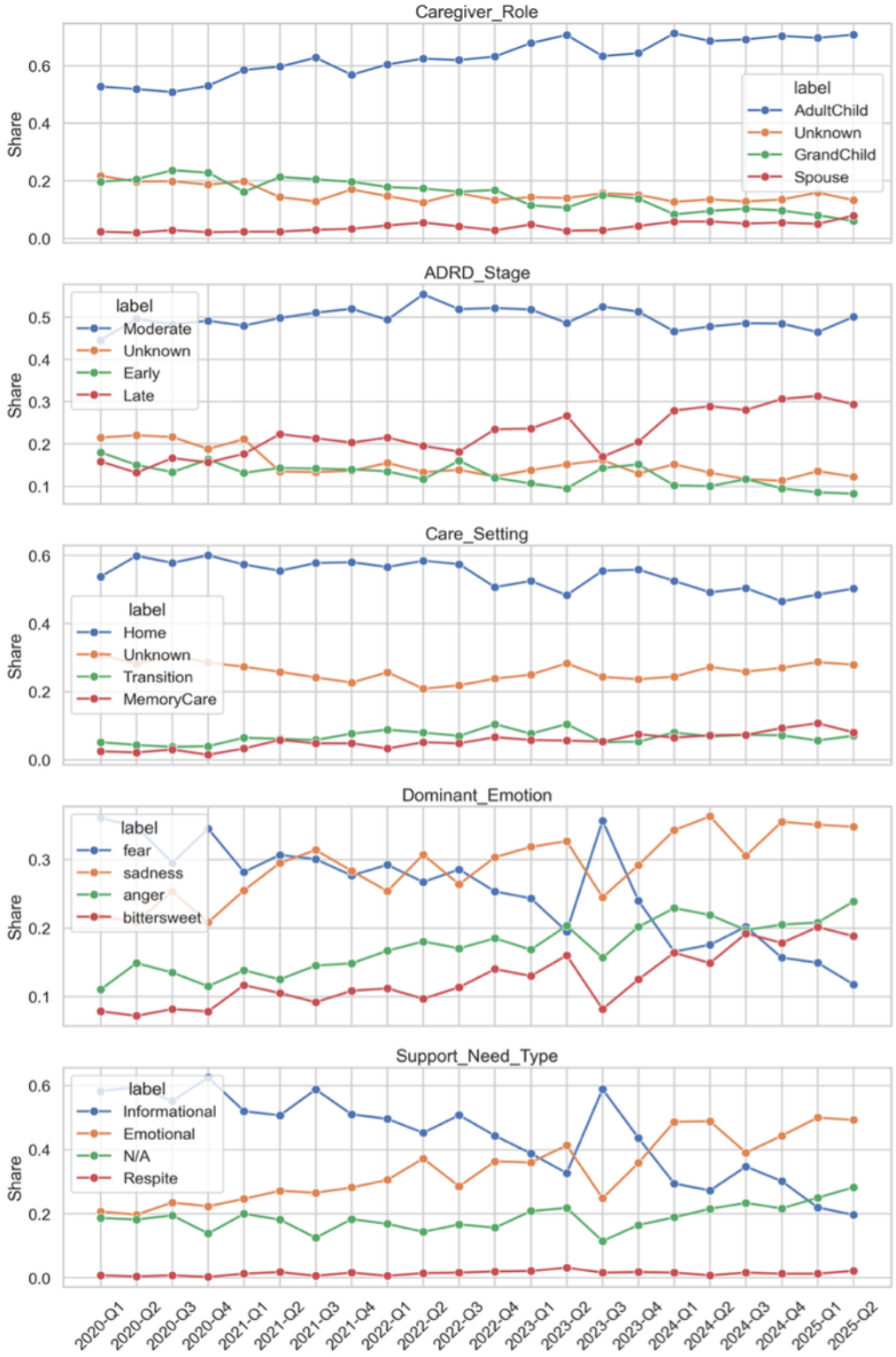
Figure 4 shows that within-post primary-secondary stressor pairings further clarified these pathways. Among posts with both a clustered primary and secondary stressor, the most frequent pairings were behavioral strain with family conflict (n=338), behavioral strain with financial strain (n=313), behavioral strain with emotional strain (n=220), behavioral strain with social isolation (n=164), care demands with financial strain (n=156), care demands with family conflict (n=112), and care refusal with family conflict (n=106). These pairs suggest that behavior-centered and care-demand narratives may function as central stressors linking direct dementia care demands to downstream family, financial, social, and mental-health consequences.

Temporal Patterns

Figures 5 and 6 present quarterly changes in inferred categorical fields, sentiment, and burden. Quarterly percentages represent the distribution of coded posts, and quarterly burden and sentiment values are means across posts. As noted above, these trends represent post-level patterns and should not be interpreted as unique-user prevalence or within-person change.

XSL•FO
RenderX

Figure 5. Quarterly trends in inferred categorical fields. ADRD: Alzheimer disease and related dementias; N/A: not available.



XSL•FO
RenderX

Figure 6. Quarterly trends in inferred sentiment and burden.

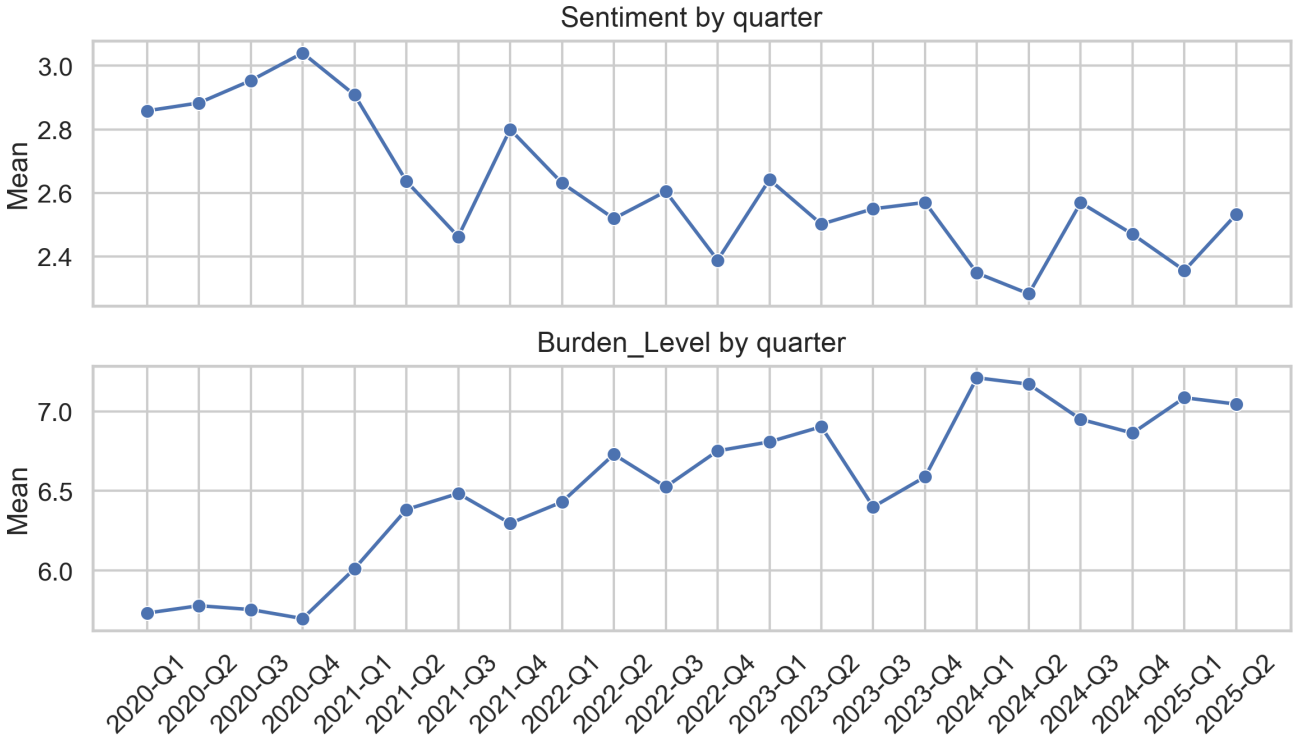


Figure 5 shows that the composition of posts changed over time. Late-stage ADRD posts increased from 15.9% (2095/13,179) in 2020 Q1 to 29.4% (3875/13,179) in 2025 Q2, while emotional-support posts increased from 20.7% (2728/13,179) to 49.2% (6484/13,179) over the same period. Informational-support posts moved in the opposite direction, decreasing from 58.3% (7683/13,179) in 2020 Q1 to 19.7% (2596/13,179) in 2025 Q2. Care-transition posts fluctuated across quarters, ranging from 3.8% (501/13,179) in 2020 Q3 to 10.5% (1384/13,179) in 2022 Q4. Figure 5 also helps interpret the temporary dip in burden observed in 2023 Q3. That quarter retained 600 posts, but it had fewer late-stage posts (2240/13,179, 17.0%) than 2023 Q2 (3519/13,179, 26.7%) and fewer care-transition posts (685/13,179, 5.2%) than 2023 Q2 (1384/13,179, 10.5%). It also had fewer emotional-support posts (3268/13,179, 24.8%) and more informational-support posts (7749/13,179, 58.8%) than adjacent quarters. Dominant emotion also shifted in 2023 Q3: fear increased to 35.7% (4705/13,179), while sadness declined to 24.5% (3229/13,179). These compositional differences are consistent with a case-mix explanation involving more advice-seeking and uncertainty-oriented posts rather than a missing-data artifact. However, because author-level data were unavailable, we cannot determine whether the dip reflected changes in author composition, repeated posting, or stochastic quarter-to-quarter variation.

Figure 6 presents the numeric burden and sentiment trends. Mean burden increased from approximately 5.7 in 2020 to above 7.0 in several later quarters, while sentiment generally declined. These descriptive trends should not be interpreted causally, but they suggest that caregiver burden narratives remained elevated in the later years of the corpus.

Subgroup Intersections and Socioclinical Implications

To identify clinically actionable subgroups, we conducted a targeted pairwise analysis of caregiver role, ADRD stage, care setting, and support need. For each pair of variables, Table 3 reports the subgroup with the highest mean burden among cells with more than 200 posts. This analysis is intended to complement the single-variable summaries in Table 2 by showing where high-burden caregiving contexts overlap.

Table . Highest-burden subgroup in each core variable pairing among cells with more than 200 posts.

Pairing	Highest-burden subgroup	Posts, n	Mean burden (SD)
Caregiver role ADRD[a] stage	Adult child+late stage	1892	7.96 (1.95)
Caregiver role care setting	Adult child+transition	667	7.93 (1.49)
Caregiver role support need	Adult child+emotional support	3096	8.10 (1.48)
ADRD stage care setting	Late stage+transition	299	8.40 (1.32)
ADRD stage support need	Late stage+emotional support	1355	8.42 (1.46)
Care setting support need	Transition+emotional support	359	8.45 (1.12)

[a]ADRD: Alzheimer disease and related dementias.

XSL•FO
RenderX

The pattern in Table 3 is that high burden is concentrated around late-stage ADRD, care transitions, adult-child caregiving, and emotional support needs. Adult-child caregivers appeared repeatedly in the highest-burden pairings. Adult-child posts with emotional support needs had a mean burden score of 8.10 across 3096 posts, and adult-child in care-transition setting posts had a mean burden score of 7.93 (SD 1.49) across 667 posts. Late-stage ADRD combined with emotional support need had a mean burden score of 8.42 (SD 1.46) across 1355 posts, and care-transition posts with emotional support needs had a mean burden score of 8.45 (SD 1.12) across 359 posts. Late-stage care transitions also had high burden (mean 8.40, SD 1.32, across 299 posts).

Table 4 further shows which subgroup categories had the highest burden within the most common primary and secondary stressor topics. To save space and focus on stable, clinically meaningful patterns, Table 4 reports only subgroups with more than 200 posts and mean burden above 8. Behavioral strain was most burdensome in late-stage posts (mean 8.52, SD 1.54) and emotional-support posts (mean 8.49, SD 1.24). Care demands were also most burdensome when paired with emotional support needs (mean 8.33, SD 1.14). Among secondary stressors, financial strain showed the highest burden levels: late-stage financial-strain posts had a mean burden score of 8.76 (SD 1.36), and financial-strain posts with emotional support needs had a mean burden score of 8.72 (SD 1.15). Family conflict and emotional strain also reached high burden when paired with late-stage ADRD or emotional support needs.

Table . High-burden groups within the most common primary and secondary stressor topics[a].

Stressor type	Stressor topic	Highest-burden group	Posts, n	Mean burden (SD)
Primary	Behavioral strain	Late stage	410	8.52 (1.54)
Primary	Behavioral strain	Emotional support need	989	8.49 (1.24)
Primary	Care demands	Emotional support need	279	8.33 (1.14)
Secondary	Family conflict	Late stage	320	8.31 (1.50)
Secondary	Family conflict	Emotional support need	613	8.22 (1.26)
Secondary	Financial strain	Late stage	290	8.76 (1.36)
Secondary	Financial strain	Emotional support need	377	8.72 (1.15)
Secondary	Financial strain	Home care	811	8.25 (1.55)
Secondary	Financial strain	Adult child	890	8.10 (1.57)
Secondary	Emotional strain	Emotional support need	424	8.27 (1.16)
Secondary	Emotional strain	Adult child	514	8.06 (1.37)
Secondary	Emotional strain	Moderate stage	424	8.03 (1.32)

[a]Groups are shown only when the subgroup had more than 200 posts and mean burden above 8.

Discussion

Main Results

In this source-matched analysis of 13,179 LLM-coded Reddit posts from the final 13,200-post scraped corpus, the LLM-assisted workflow identified clinically interpretable ADRD caregiver constructs at scale. Reddit ADRD caregiving narratives were dominated by adult-child and grandchild perspectives; moderate-stage, home-based care was the central disclosure context; sadness, fear, and anger were common; and negative emotions aligned with higher burden. Inferred burden increased across ADRD stages, from early-stage posts to moderate- and late-stage posts. Behavioral strain emerged as the dominant primary stressor, while family conflict, financial strain, emotional strain, social isolation, and burnout were common secondary stressors. Temporal analysis showed that mean burden was relatively stable in 2020, rose through 2021 and 2022, and remained elevated from 2023 through mid-2025; sentiment generally moved in the opposite direction and was lower in later years. These findings suggest that ADRD caregiver burden in Reddit narratives is not only emotional distress but also a set of linked practical, relational, financial, and support-navigation problems.

Several results are especially important for interpreting the measurement workflow. First, caregiver role, ADRD stage, care setting, support need, sentiment, burden, and stressor type were more recoverable from Reddit narratives than sociodemographic characteristics. This pattern is expected because caregivers often describe care tasks, disease stages, emotions, and support needs directly, but rarely provide income, race, education, or marital status. Second, the burden patterns were internally consistent. Late-stage ADRD posts had higher burden and lower sentiment than early-stage posts; transition posts had the highest burden among care-setting groups; and posts expressing sadness or anger had high burden and low sentiment, while hope, joy, and neutral posts had lower burden and higher sentiment. Third, primary and secondary stressors followed the logic of the Stress Process Model: direct dementia-care problems, especially behavioral strain and care demands, often appeared alongside spillover burdens involving family systems, finances, social isolation, and caregiver burnout [8].

XSL•FO
RenderX

The human validation benchmark supports the main structured fields but also clarifies the boundaries of the corpus. Human agreement was excellent for caregiver role, ADRD stage, care setting, dominant emotion, support need, and burden level, and substantial for primary and secondary stressor categories. However, agreement on whether posts were strictly in-scope as informal caregiver narratives was poor; the findings should therefore be interpreted as descriptions of high-engagement ADRD caregiving discourse on Reddit rather than prevalence estimates from a definitively verified caregiver population.

Comparison With Prior Work

This study speaks directly to recent work using Reddit and other online platforms to study caregiver narratives. For example, Huang et al [14] analyzed Reddit posts and comments from r/AgingParents, r/eldercare, and r/elderlycare and found that caregivers often disclosed care recipients' functional problems, caregiving relationship strain, care recipients' emotional problems, and scarcity of health and social resources. Guided by the Stress Process Model, they showed that primary stressors and secondary stressors co-occurred in caregiver posts and that different stressors were associated with different forms of social support received in comments. Our findings echo this stress-process pattern in an ADRD-specific corpus. By applying an LLM-assisted measurement workflow, our study extends this work by analyzing a larger longitudinal corpus and identifying more varied ADRD-specific stressors. The most frequent within-post pairings, such as behavioral strain with family conflict, financial strain, emotional strain, and social isolation, show how dementia-specific primary stressors may connect to downstream caregiver-life stressors within the same post.

The descriptive results are also consistent with prior online caregiver research on poster roles, discussion topics, and visible forms of distress. We found that adult children accounted for 62.7% (8266/13,179) of posts and grandchildren for 15.3% (2012/13,179), which is consistent with prior online-community studies showing that family members, especially adult children and spouses, are prominent caregiver voices in dementia forums [18,20]. Compared with national dementia caregiver estimates, however, the Reddit corpus appears more heavily weighted toward adult-child and grandchild voices and less weighted toward spouse caregivers. The 2026 Alzheimer Disease Facts and Figures report indicates that over half of dementia caregivers provide assistance to a parent or parent-in-law, whereas approximately 10% provide help to a spouse with Alzheimer disease or another dementia [1]. In our post-level corpus, spouse caregivers accounted for only 4.0% of inferred caregiver roles. This difference should not be interpreted as a population difference because Reddit posts are not a representative caregiver sample. Instead, it suggests that Reddit may be especially useful for observing adult-child and younger intergenerational caregiving narratives. The 15.3% (2012/13,179) share of grandchild role inferences is notable because grandchildren are less often examined as a distinct group in traditional caregiver research. Reddit may make younger and intergenerational caregivers more visible because of age-related differences in platform use and the accessibility of anonymous peer support. Our findings also showed that moderate-stage ADRD and home-based care were common, while late-stage posts and transition posts had higher burden and lower sentiment. These findings align with prior dementia-caregiving research showing that care demands change across disease progression and that functional decline, behavioral symptoms, care transitions, and caregiving relationship strain are important sources of burden [37-40].

The emotional and support need findings also fit prior Reddit and online-forum studies. Huang et al [15] supportive-communication study found that Reddit caregivers often used direct support seeking and expressed fear, anger, and sadness. Other ADRD-focused online narrative studies have identified themes such as family issues, care facilities, financial struggles, abuse concerns, symptoms, emotional distress, nursing homes, sleep problems, incontinence, and advice seeking [18,22-24]. Our results echo this literature: informational and emotional support were the 2 dominant support needs, and sadness, fear, and anger together accounted for most posts. Our results extend this literature by adding a burden gradient to these patterns. Respite, financial, instrumental, emotional, legal, and medical support needs all had higher mean burden than informational posts or posts with no explicit support need. Sadness and anger posts had high burden and low sentiment, whereas hope, joy, and neutral posts had lower burden and higher sentiment. This adds a post-level link between emotional expression, support need, sentiment, and inferred burden that is harder to obtain from topic prevalence alone.

This study also extends prior ADRD-specific NLP, machine learning, and LLM studies by changing the unit of measurement. Earlier work used structural topic modeling, machine learning classification, sentiment analysis, or related methods to identify aggregate themes in online posts [18-22,24]. More recent LLM applications have mapped digital caregiving functions, built retrieval-augmented caregiver support systems, or summarized large social-media corpora into aggregate themes [25-27]. These studies are valuable because they scale beyond manual qualitative review and identify clinically meaningful themes. However, topic models, supervised classifiers, and many LLM topic-modeling workflows often produce broad themes or fixed labels, and they may not preserve how several constructs appear together in the same narrative. Our study adds a different use case: structured narrative measurement. The workflow inferred post-level characteristics, caregiving contexts, confidence ratings, and open-vocabulary primary and secondary stressors. This made it possible to compare burden across stage, setting, support need, and emotion; examine stressor clusters; assess uncertainty; and describe temporal changes in both categorical composition and numeric burden. These outputs should not be treated as ground truth, but the findings show how LLMs can help transform naturalistic ADRD caregiving-related Reddit discourse into analyzable, theory-guided measures when paired with source matching, schema validation, field-level correction, confidence ratings, and human validation [29-31].

Social and Clinical Implications

The vulnerable-subgroup findings suggest several practical targets for caregiver support. Care transitions were the clearest pressure point, especially when paired with late-stage ADRD

XSL•FO
RenderX

or emotional support needs. This suggests that transitions should be treated as key intervention windows. Prior work on nursing home admission and home-to-long-term-care transitions shows that institutional transitions can change but do not necessarily eliminate caregiver distress, and that caregivers often have unmet emotional and informational needs during transitions [39,40]. Placement counseling, discharge planning support, anticipatory guidance about memory care or hospice decisions, and follow-up after the move may be useful when caregivers are deciding whether, when, or how to change care settings.

Emotional support needs were consistently present in the highest-burden posts. This suggests that caregiver-support systems should not focus only on informational resources. Peer-support forums, caregiver hotlines, clinical teams, and digital tools may need pathways for validation, distress screening, and referral when caregivers describe guilt, isolation, fear, or unsustainable care. The LLM's role in this study was analytic rather than therapeutic: it classified burden-related signals in posts but did not provide support, assess risk, or make clinical decisions. Automated systems may help summarize needs at scale or direct content to trained human reviewers, but the loneliness, emotional overload, empathy, and validation reflected in high-burden posts are fundamentally relational. Any caregiver-facing application should augment rather than replace reciprocal peer support, clinician and social-worker judgment, and accountable referral pathways. Financial strain also appeared as a high-burden secondary stressor, especially in late-stage, home-care, adult-child, and emotional-support narratives. This finding is consistent with dementia-caregiver studies linking financial strain to role overload and role captivity and with interview research showing that caregivers need integrated financial, legal, functional, educational, and peer-support resources [41,42]. For aging services, this points toward financial-navigation support, benefits literacy, Medicaid and long-term-care planning, and decision aids.

Family conflict and emotional strain also emerged as intervention-relevant secondary stressors. Family conflict reached high burden when paired with late-stage ADRD or emotional support needs, suggesting that conflict often intensifies when care demands become more severe. This is consistent with prior online-community findings that family issues and caregiving relationships are central dimensions of ADRD caregiving discourse [18,20]. Interventions that focus only on symptom management may miss this social context. Caregiver support may also need family mediation, role clarification, and planning tools that help relatives divide responsibility before a crisis or care transition occurs [4]. Grandchildren accounted for 15.3% (2012/13,179) of caregiver role inferences, making them a visible group in this Reddit corpus. This does not imply population prevalence, but it suggests that online ADRD caregiver support environments may include younger and less studied caregivers whose needs may differ from those of spouses or adult children.

The temporal findings add another implication: caregiver-support needs may change in form even when the underlying burden remains high. The later years of the corpus showed higher burden levels and lower sentiment, while emotional support became more common and informational support became less dominant. This suggests that digital health systems and community platforms should be able to respond to both practical advice seeking and emotional distress, rather than treating caregiver support as a single static need. We did not identify a clear external event that could explain the 2023 Q3 pattern by itself. Therefore, the temporal patterns should be interpreted as descriptive signals of changing post composition rather than evidence of a policy or social-event effect.

Looking forward, if LLM-assisted measurement is later embedded in caregiver-facing support systems, algorithmic interpretation may itself become part of the stress process. Opaque or erroneous classifications, perceived surveillance, and mismatched recommendations could create new secondary strains, whereas transparent and well-governed systems could improve navigation and access to support resources. Future research could therefore extend established stress frameworks to examine algorithm-mediated appraisal, trust, perceived control, and support access.

Limitations

This study has several limitations. First, the analytic sample is not representative of the broader population of informal ADRD caregivers. Reddit users constitute a self-selected population, and the monthly top-score sampling strategy further prioritizes posts receiving greater community engagement. Although many users in these communities may be based in the United States, posts from other countries may also be included. Accordingly, differences across subreddits should be interpreted as reflecting variation in community function, audience, and disclosure context rather than differences in the population prevalence of caregiver experiences. In addition, limited agreement in identifying in-scope posts indicates uncertainty in distinguishing general ADRD discussions from narratives clearly authored by informal caregivers. The reported percentages and mean scores therefore characterize this selected corpus of high-engagement Reddit posts and should not be generalized to the prevalence or distribution of caregiver burden in the broader population.

Second, the unit of analysis was the post rather than a verified unique caregiver. Repeated posting by the same caregiver may influence observed distributions. ADRD-specific subreddits may also include posts from patients, clinicians, or general community members. The inference workflow attempted to measure caregiver attributes only when the post supported that framing, but noncaregiver perspectives may remain in the corpus. This limitation was visible in the human validation benchmark, where coders disagreed substantially on the in-scope field despite high agreement for most downstream constructs.

Third, the inferred variables are model-generated measurements rather than ground-truth annotations. The 150-post validation benchmark supported several major fields, but it cannot fully establish performance across all rare subgroups, stressor types, or ambiguous narratives. The source-matched analysis retained only LLM-coded records whose IDs joined to the final scraped corpus. This filtering protects corpus validity but also shows the need for strict ID validation in LLM-assisted pipelines.

Fourth, this study analyzed posts from January 2020 through June 2025 rather than a longer historical period. This time

XSL•FO
RenderX

window was chosen partly to manage computational cost and partly to avoid mixing the main study period with more distant pre-COVID Reddit dynamics. The COVID-19 pandemic may still have shaped caregiver needs, care access, and posting behavior during the study period. Future work should analyze longer time spans and larger post samples to test whether the main burden, stressor, and subgroup patterns are robust across different periods.

Fifth, we did not conduct subgroup analyses by demographic attributes such as race, education, income, marital status, or age. Reddit users often do not disclose these personal characteristics, and the LLM therefore assigned low-confidence ratings to many demographic fields. Analyzing these fields in the current corpus could give a misleading impression of precision. Future studies with larger corpora may still recover enough high-confidence demographic information to support cautious subgroup comparisons and to identify vulnerable caregiver groups for more targeted intervention design.

Finally, although the strict rule- and rubric-based prompt made scoring more standardized, and agreement with human coding is necessary for validation, they may not fully explain how the model reached its judgments. The workflow shows whether model outputs match benchmark labels, not why the model selected a label in an ambiguous narrative. Borderline human cases most often involved whether sparse text justified an “other” stressor rather than “unclear” and whether limited distress cues supported low burden rather than insufficient evidence. Explicit relationship terms, functional-dependence descriptions, behavioral symptoms, named care settings, transition language, and direct support-seeking or emotion language were the observable cues specified in the coding framework, but the study did not estimate their feature importance.

Conclusions

This study demonstrates the potential of LLM-assisted measurement for studying ADRD caregiving-related Reddit discourse at scale. The main innovation is not simply using an LLM to summarize Reddit posts, but using it as a structured measurement tool that links caregiving context, sentiment, burden, support needs, and primary and secondary stressors at the post level. The findings show clinically meaningful patterns: burden increased across inferred ADRD stage, negative emotions aligned with higher burden, behavioral strain was a central primary stressor, and family, financial, emotional, and social strains appeared as common secondary stressors. The subgroup analyses further suggest that late-stage ADRD, care transitions, adult-child caregiving, emotional support needs, and financial strain are key contexts for high burden. These patterns highlight transition navigation, emotional support, financial and benefits counseling, and family role negotiation as potential targets for caregiver-support interventions and digital health tools. Human validation supported the reliability of the main structured fields, while also showing that corpus inclusion and model-generated measures require cautious interpretation. Overall, this workflow offers a scalable and transparent approach for converting naturalistic online ADRD caregiving-related Reddit discourse into theory-guided quantitative measures that can inform clinical aging research, service planning, and future intervention design.

Generative AI Use

The authors acknowledge the use of generative AI tools (ChatGPT) to assist with language editing and manuscript revision. All substantive content, analyses, interpretations, and final wording were reviewed and approved by the authors. The use of large language models (LLMs) as part of the study workflow is described separately in the Methods section.

Funding

This research was supported by grant, P30AG066614, awarded to the Center on Aging and Population Sciences at The University of Texas at Austin by the National Institute on Aging. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The raw Reddit posts analyzed in this study are based on publicly accessible Reddit data. To protect user privacy, raw post text, usernames, and individual-level identifiers are not publicly redistributed. Aggregate tables, figures, and analysis code may be made available by the corresponding author upon reasonable request, subject to ethical and platform-policy considerations.

Authors' Contributions

Conceptualization: YZ (lead), MFZ (equal)
Data curation: MFZ
Formal analysis: MFZ (lead), YZ (supporting)
Funding acquisition: YZ
Investigation: YZ (lead), MFZ (equal)
Methodology: MFZ (lead), YZ (supporting)
Supervision: YZ
Validation: MFZ (lead), YZ (supporting)
Visualization: MFZ (lead), YZ (supporting)

XSL•FO
RenderX

Writing – original draft: YZ (lead), MFZ (supporting)
Writing – review & editing: YZ (lead), MFZ (supporting)
The author MFZ is currently not affiliated with any institution but is an Independent Researcher.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Reddit data corpus funnel for Alzheimer disease and related dementias (ADRD) caregiving post extraction.
[DOCX File, 17 KB - aging_v9i1e103305_app1.docx]

Multimedia Appendix 2
Full Batch-Mode large language model (LLM) Inference Prompt.
[DOCX File, 19 KB - aging_v9i1e103305_app2.docx]

Multimedia Appendix 3
Human validation agreement for the 150-post benchmark.
[DOCX File, 17 KB - aging_v9i1e103305_app3.docx]

References

1. The Alzheimer's Association. 2026 Alzheimer's disease facts and figures. Alzheimers Dement 2026 Apr;22(4):e71345. [doi: 10.1002/alz.71345]
2. Gilhooly KJ, Gilhooly MLM, Sullivan MP, et al. A meta-review of stress, coping and interventions in dementia and dementia caregiving. BMC Geriatr 2016 May 18;16:106. [doi: 10.1186/s12877-016-0280-8] [Medline: 27193287]
3. Wiles JL, Leibing A, Guberman N, Reeve J, Allen RES. The meaning of "aging in place" to older people. Gerontologist 2012 Jun;52(3):357-366. [doi: 10.1093/geront/gnr098] [Medline: 21983126]
4. Zhang Y. Family structure and transitions into and out of homeboundness: evidence from the National Health and Aging Trends Study, 2011-2024. J Gerontol B Psychol Sci Soc Sci 2026 Oct 1;gbag206. [doi: 10.1093/geronb/gbag206] [Medline: 42826503]
5. Caregiving in the United States 2020. : AARP, National Alliance for Caregiving; 2020. [doi: 10.26419/ppi.00103.001]
6. Caregivers of a person with alzheimer's disease or a related dementia. CDC. URL: https://www.cdc.gov/caregiving/about/index.html [accessed 2026-07-30]
7. Pearlin LI, Lieberman MA, Menaghan EG, Mullan JT. The stress process. J Health Soc Behav 1981 Dec;22(4):337-356. [doi: 10.2307/2136676] [Medline: 7320473]
8. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct;30(5):583-594. [doi: 10.1093/geront/30.5.583] [Medline: 2276631]
9. Friedman EM, Shih RA, Langa KM, Hurd MD. US prevalence and predictors of informal caregiving for dementia. Health Aff (Millwood) 2015 Oct;34(10):1637-1641. [doi: 10.1377/hlthaff.2015.0510] [Medline: 26438738]
10. Pristavec T, Luth EA. Informal caregiver burden, benefits, and older adult mortality: a survival analysis. J Gerontol B Psychol Sci Soc Sci 2020 Nov 13;75(10):2193-2206. [doi: 10.1093/geronb/gbaa001] [Medline: 31903481]
11. Vasileiou K, Barnett J, Barreto M, et al. Experiences of loneliness associated with being an informal caregiver: a qualitative investigation. Front Psychol 2017;8:585. [doi: 10.3389/fpsyg.2017.00585] [Medline: 28469589]
12. Akgun-Citak E, Attepe-Ozden S, Vaskelyte A, et al. Challenges and needs of informal caregivers in elderly care: qualitative research in four European countries, the TRACE project. Arch Gerontol Geriatr 2020;87:103971. [doi: 10.1016/j.archger.2019.103971] [Medline: 31756568]
13. Tourangeau R, Yan T. Sensitive questions in surveys. Psychol Bull 2007 Sep;133(5):859-883. [doi: 10.1037/0033-2909.133.5.859] [Medline: 17723033]
14. Huang NM, Timothy B, Ho SS. Stressors disclosed on Reddit by caregivers of older adults and social support received: content analysis. JMIR Aging 2025 Sep 5;8:e71452. [doi: 10.2196/71452] [Medline: 40911350]
15. Huang NM, Timothy B, Ho SS. Supportive communication patterns among informal caregivers of older adults on Reddit: a content analysis. Digit Health 2025;11:20552076251376537. [doi: 10.1177/20552076251376537] [Medline: 41018504]
16. Amorim D, Lam K, Chary AN, Plys E, Shi S. Positive aspects of caregiving: a qualitative analysis of Reddit posts. J Am Geriatr Soc 2026 Jun;74(6):1792-1796. [doi: 10.1111/jgs.70358] [Medline: 41724689]
17. Grimmer J, Stewart BM. Text as data: the promise and pitfalls of automatic content analysis methods for political texts. Polit Anal 2013;21(3):267-297. [doi: 10.1093/pan/mps028]
18. Cheng K, Yin Z. "Please advise": understanding the needs of informal caregivers of people with Alzheimer's disease and related dementia from online communities through a structured topic modeling approach. AMIA Jt Summits Transl Sci Proc 2022;2022:149-158. [Medline: 35854737]

XSL•FO
RenderX

19. Ni C, Malin BA, Song L, Jefferson AL, Commiskey P, Yin Z. “Rough day … need a hug”: learning challenges and experiences of the Alzheimer’s disease and related dementia caregivers on Reddit. ICWSM 2022;16:711-722. [doi: 10.1609/icwsm.v16i1.19328]
20. Ni C, Song Q, Malin B, et al. Examining online behaviors of adult-child and spousal caregivers for people living with Alzheimer disease or related dementias: comparative study in an open online community. J Med Internet Res 2023 Nov 17;25:e48193. [doi: 10.2196/48193] [Medline: 37976095]
21. Shoults CC, Rutherford MW, Kemp AS, et al. Analysis of caregiver burden expressed in social media discussions. Int J Environ Res Public Health 2023 Jan 20;20(3):1933. [doi: 10.3390/ijerph20031933] [Medline: 36767297]
22. Hswen Y, Xiong J, Hurley M, Nguyen TT. Experiences of Alzheimer’s disease and related dementia family caregivers on Reddit communities: a topic modeling and sentiment analysis. Artif Intell Health 2024;1(3):127-135. [doi: 10.36922/aih.3075] [Medline: 39246419]
23. Huang NM, Wong LZ, Ho SS, Timothy B. Understanding challenges and emotions of informal caregivers of general older adults and people with Alzheimer disease and related dementia: comparative study. J Med Internet Res 2025 Feb 28;27:e54847. [doi: 10.2196/54847] [Medline: 40053723]
24. Kim J, Cai ZR, Chen ML, et al. Mental health care needs of caregivers of people with Alzheimer’s disease from online forum analysis. NPJ Ment Health Res 2024 Nov 14;3(1):54. [doi: 10.1038/s44184-024-00100-y] [Medline: 39537826]
25. Kim S, Han DY, Bae J. Transforming Alzheimer’s digital caregiving through large language models. Curr Alzheimer Res 2024;21(7):503-516. [doi: 10.2174/0115672050301740241118044604] [Medline: 39592896]
26. Hasan WU, Zaman KT, Wang X, Li J, Xie B, Tao C. Empowering Alzheimer’s caregivers with conversational AI: a novel approach for enhanced communication and personalized support. NPJ Biomed Innov 2024 Dec 4;1(1):3. [doi: 10.1038/s44385-024-00004-8] [Medline: 42032173]
27. He W, Hou B, Zheng A, et al. Advanced topic modeling with large language models: analyzing social media content from dementia caregivers. Innov Aging 2025 Dec 26;9(Suppl 1):S38-S47. [doi: 10.1093/geroni/igaf120] [Medline: 41458886]
28. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. Presented at: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics; Jun 2-7, 2019; Minneapolis, MN p. 4171-4186. [doi: 10.18653/v1/N19-1423]
29. Gilardi F, Alizadeh M, Kubli M. ChatGPT outperforms crowd workers for text-annotation tasks. Proc Natl Acad Sci U S A 2023 Jul 25;120(30):e2305016120. [doi: 10.1073/pnas.2305016120] [Medline: 37463210]
30. Ziems C, Held W, Shaikh O, Chen J, Zhang Z, Yang D. Can large language models transform computational social science? Comput Linguist 2024 Mar 1;50(1):237-291. [doi: 10.1162/coli_a_00502]
31. Asirvatham H, Mokski E, Shleifer A. GPT as a measurement tool. : National Bureau of Economic Research Working Paper Series, No. 34834; 2026 Feb. [doi: 10.3386/w34834]
32. Baumgartner J, Zannettou S, Keegan B, Squire M, Blackburn J. The pushshift Reddit dataset. ICWSM 2020;14:830-839. [doi: 10.1609/icwsm.v14i1.7347]
33. Tang AY, Kwak J, Xiao L, et al. Online health information wants of caregivers for persons with dementia in social media. SAGE Open 2023 Oct;13(4). [doi: 10.1177/21582440231205367]
34. Garg M, Sohn S. CareD: caregiver’s experience with cognitive decline in Reddit posts. Presented at: 2023 IEEE 11th International Conference on Healthcare Informatics (ICHI); Jun 26-29, 2023; Houston, TX p. 581-587. [doi: 10.1109/ICHI57859.2023.00104]
35. Google. Gemma 4: our most capable open models to date.: Google Blog; 2026. URL: https://blog.google/innovation-and-ai/technology/developers-tools/gemma-4/ [accessed 2026-07-30]
36. Google. Gemma 4 model overview.: Google AI for Developers; 2026. URL: https://ai.google.dev/gemma/docs/core [accessed 2026-07-30]
37. Huang HL, Shyu YIL, Chen MC, et al. Family caregivers’ role implementation at different stages of dementia. Clin Interv Aging 2015;10:135-146. [doi: 10.2147/CIA.S60574] [Medline: 25584022]
38. van den Kieboom R, Snaphaan L, Mark R, Bongers I. The trajectory of caregiver burden and risk factors in dementia progression: a systematic review. J Alzheimers Dis 2020;77(3):1107-1115. [doi: 10.3233/JAD-200647] [Medline: 32804093]
39. Gaugler JE, Mittelman MS, Hepburn K, Newcomer R. Clinically significant changes in burden and depression among dementia caregivers following nursing home admission. BMC Med 2010 Dec 17;8:85. [doi: 10.1186/1741-7015-8-85] [Medline: 21167022]
40. Lee K, Chung J, Meyer KN, Dionne-Odom JN. Unmet needs and health-related quality of life of dementia family caregivers transitioning from home to long-term care: a scoping review. Geriatr Nurs 2022;43:254-264. [doi: 10.1016/j.gerinurse.2021.12.005] [Medline: 34953331]
41. Liu Y, Dokos M, Fauth EB, Lee YG, Zarit SH. Financial strain, employment, and role captivity and overload over time among dementia family caregivers. Gerontologist 2019 Sep 17;59(5):e512-e520. [doi: 10.1093/geront/gnz099] [Medline: 31322654]
42. Fan Q, DuBose L, Ory MG, et al. Financial, legal, and functional challenges of providing care for people living with dementia and needs for a digital platform: interview study among family caregivers. JMIR Aging 2023 Sep 5;6:e47577. [doi: 10.2196/47577] [Medline: 37526513]

XSL•FO
RenderX

Abbreviations

ADRD: Alzheimer disease and related dementias
BERT: Bidirectional Encoder Representations from Transformers
IRB: institutional review board
LLM: large language model
N/A: not available
NLP: natural language processing
RQ: research question

Edited by J Wang; submitted 01.Jun.2026; peer-reviewed by E Yeoh, J Guendouzi; revised version received 17.Sep.2026; accepted 08.Oct.2026; published 24.Nov.2026.

Please cite as:
Zhang Y, Zhang M
Alzheimer Disease and Related Dementias Caregiver Burden on Reddit: Large Language Model–Assisted Measurement Study
JMIR Aging 2026;9:e103305
URL: https://aging.jmir.org/2026/1/e103305
doi:10.2196/103305

© Yan Zhang, Mingfei Zhang. Originally published in JMIR Aging (https://aging.jmir.org), 24.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Barriers and Facilitators to Technology Use Among Caregivers of Homebound Older Adults: Descriptive Qualitative Study

Carina Katigbak[1], PhD; Esha Shenoy[2*], BA; Padmavathy Ramaswamy[3*], PhD; Jessica Lee[4], MD; Kathrin Milbury[5], PhD; Katie Scott[6*], MPH

[1]Department of Research, Cizik School of Nursing, The University of Texas Health Science Center at Houston, 6901 Bertner Ave, Houston, TX, United States

[2]Dell Medical School, The University of Texas at Austin, Austin, TX, United States

[3]Department of Graduate Studies, Cizik School of Nursing, The University of Texas Health Science Center at Houston, Houston, TX, United States

[4]Department of Internal Medicine, McGovern School of Medicine, The University of Texas Health Science Center at Houston, Houston, TX, United States

[5]Department of Behavioral Science, Division of Cancer Prevention and Population Sciences, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[6]CarePartners, Houston, TX, United States

[*]these authors contributed equally

Corresponding Author:
Carina Katigbak, PhD
Department of Research, Cizik School of Nursing, The University of Texas Health Science Center at Houston, 6901 Bertner Ave, Houston, TX, United States

Abstract

Background: Homebound older adults (≥60 y) and their informal family caregivers face multiple challenges in meeting their instrumental, emotional, and informational needs, which result in lower functional status, multiple comorbidities, and poorer mental health. Caregivers are tasked with providing uncompensated care for medically complex older adults, often without formal health care training. This circumstance may translate to high levels of burden and greater vulnerability to social isolation and loneliness, which may precipitate chronic stress and compromise caregiver health. In recent years, numerous health-focused technologies assisting with health-promoting behaviors such as physical activity, stress management, and nutritional tracking have emerged. However, little is known about how these tools are adopted by family caregivers of homebound older adults.

Objective: This descriptive qualitative study aimed to explore how caregivers of homebound older adults perceive technology in the context of supporting caregiving and aging in place through focus groups and one-on-one interviews.

Methods: We conducted a descriptive qualitative study using in-person and online video conference focus group methodology with 40 individuals (32 caregivers, 6 clinicians, and 2 additional key informants) to explore perspectives on caregivers' use of technology to support their caregiving activities. Participants completed a brief sociodemographic questionnaire prior to all interviews. Procedures for thematic analysis were used to analyze the data and identify emerging themes.

Results: Three main themes (and related subthemes) emerged from the data: (1) these technologies are not designed for people like us, (2) balancing the scales, and (3) preserving a sense of self. Caregivers' experiences with technology are marked by a paradox: the same technologies that can ease caregiving activities may also impose new burdens through additional costs, uncertainty, learning demands, and emotional strain.

Conclusions: Caregivers should not navigate these technological tensions alone. Realizing the promise of technology requires more than innovative tools—it demands caregiver-centered design, accessible and equitable implementation strategies, and sustained clinical and health system support to ensure that technology alleviates, rather than amplifies, caregiver burden.

(JMIR Aging 2026;9:e96226) doi:10.2196/96226

KEYWORDS

caregivers; health promotion; technology; qualitative; older adults

Introduction

By 2040, one in five Americans (nearly 78.5 million people) will be aged 65 years or older, a 17% increase from 2022 [1]. While this growth may be attributed to medical advancements and longer life expectancy, a large majority of US older adults also live with multiple comorbidities and functional impairments [2] related to meeting needs for activities of daily living that

XSL•FO
RenderX

require assistance from informal family caregivers [3]. Of this population, more than 7 million US older adults [4] are homebound, which is defined as individuals who never or rarely leave their homes because of their complex care needs [5]. Homebound individuals have a higher chronic disease burden along with reduced access to primary care (due to their inability to leave home). As such, informal family caregivers (herein referred to as caregivers) play a critical role in acting as the primary link between the health care system and homebound older adults requiring comprehensive assistance.

Caregivers often lack formal health care training but are nonetheless tasked with caring for medically complex older adults, a circumstance that may precipitate the harmful effects of chronic stress [6]. The general literature on caregivers of older adults identifies that they face challenges in meeting their tangible, emotional, and informational needs [7,8], which consequently increase their risk for diminishing functional status, multiple comorbidities, and poorer mental health [9]. High burden levels, greater vulnerability to social isolation, and loneliness [10] negatively impact their quality of life [11,12] and may potentially harm the interpersonal dynamic between caregiver and care recipient [13,14]. Similarly, some caregivers of older adults who are homebound shared that rewarding feelings of duty and gratitude in being able to care for their loved ones [15] are frequently overshadowed by providing round-the-clock care and its impact on their other interests and obligations [16]. Over the course of the caregiving journey, caregivers' health can become severely compromised [17-19], underscoring the critical need to support and maintain caregiver health and well-being. Interventions and strategies for caregivers to maintain their own health largely focus on stress management techniques and psychoeducational interventions [20]. A systematic review of interventions targeting caregivers reported greater effectiveness in increasing caregiver knowledge and ability rather than mitigating caregiver burden and depression [21]. Although digital health tools are increasingly marketed to promote caregiver health, and these may be used as promising tools to improve caregiver outcomes, a recent review noted that of the over 200,000 health-focused mobile apps, only 44 were targeted toward caregivers [22]. Remote and digital tools may be especially useful for caregivers of homebound adults, as home-based medical care has proven to alleviate the burden of their fragmented care [15]. These digital apps largely focus on the process of caregiving in 3 categories of support: information and resources, family communication, and caregiver-recipient interactions. Few comprehensively addressed domains of support and seldom addressed specific stress reduction strategies for reducing caregiver burden [22]. Moreover, a scoping review determined that caregiver apps to support care planning and coordination for those with multiple chronic conditions are underdeveloped [23]. Digital solutions (using computer programs, smart-enabled devices, and the internet to improve efficiencies at home) are not limited to mobile apps alone but also extend to telehealth consultations [24], in addition to various devices that those in the caregiving community use for tracking/monitoring [23,25] purposes, such as GPS trackers, wearable technologies, and apps (eg, smart watches, Find My [Apple Inc], and AirTag [Apple Inc]).

Understanding patterns of technology use and device adoption across diverse caregiver populations is increasingly important for enhancing geriatric care and can guide recommendations about appropriate technology interventions [26]. While health-focused technologies exist and continue to evolve rapidly, these technologies must be acceptable, accessible, equitable, and responsive to caregivers' needs [19,27]. Thus, our study sought to describe how family caregivers of homebound older adults use technology to support their caregiving activities, specifically identifying barriers and facilitators to its use; understanding these perspectives is crucial for developing user-centered interventions that optimize the use of technology to support caregivers and encourage high-quality care.

Methods

Study Design

We conducted a qualitative descriptive study to directly capture participants' perspectives on how technology is used to support caregiving activities. Using a community-engaged research approach, we collaborated with Care Partners, a Houston-based community services and advocacy group that delivers high-quality programming for older adults and people with dementia, including family caregivers (such as support groups, educational programming, and in-home visits for caregiver consultation). The organization serves over 2500 older adults in the Houston area. We sought to examine caregivers' experiences using technology to facilitate their caregiving activities and triangulated these data with information shared by clinicians and key informants—specifically around barriers and facilitators to integration and consistent use. The study followed guidelines for the COREQ (Consolidated Criteria for Reporting Qualitative Research).

Participant Recruitment

From July to November 2023, a purposive sample of caregivers and key informants were recruited from CarePartners' roster of volunteer caregivers, clinicians, and community key informants to triangulate data from multiple sources. Consistent with Braun and Clarke reflexive thematic analysis [28], sample adequacy was evaluated based on the richness of the data and the analytic objectives rather than the achievement of data saturation. Recruitment continued until the research team judged that the dataset provided sufficient depth and breadth to address the research question.

Eligibility criteria consisted of the following: informal, unpaid, caregivers of older adults (>60 y old), who were aged over 21 years old, fluent in English, residing in the same home as the older adult, and providing more than 12 hours of care per week; clinicians currently working with CarePartners; and community key informants (eg, community leaders and elder care advocates). Multiple modalities were used to recruit participants: (1) engaging in direct outreach to caregivers participating in online support groups, sharing information with support group moderators, and joining online meetings to provide details about the study; (2) disseminating information about the study through a weekly electronic CarePartners newsletter, which is shared with its membership; and (3) posting physical flyers in common spaces. Prospective participants contacted CarePartners staff to

XSL•FO
RenderX

complete an electronic form so that a member of the research team could follow up and determine study eligibility. A trained research nurse assessed participants for study eligibility, obtained informed consent, and scheduled an interview date and time with the participant.

Researcher Positionality

The study team comprised investigators with expertise in nursing, medicine, psychology, public health, and qualitative methods—all of whom are experienced in working with patients with chronic illness and their families. We acknowledge that our clinical and research backgrounds, as well as our commitment to improving caregiver support and health equity, informed the framing of the research questions and interpretation of the findings. To promote reflexivity, team members engaged in ongoing discussions throughout data collection and analysis to examine assumptions, challenge interpretations, and consider alternative explanations. Coding and thematic development were conducted collaboratively, with analytic decisions documented to enhance transparency and ensure that findings remained grounded in participants' accounts.

Data Collection

Interview Procedures and Participant Consent

Participants were invited to engage in in-person or virtual interviews (via Zoom); caregivers and clinicians participated in 60- to 90-minute focus group interviews (comprising 6 - 8 people per group), while key informants had individual online interviews to glean their unique perspectives on the phenomenon. Prior to initiating data collection, we solicited our community partners' feedback for improving the qualitative interview tool. Participants were informed of the confidential and voluntary nature of their participation in the study and were informed that they could choose to withdraw at any time.

Demographic Survey

Eligible participants first completed a demographic survey (to characterize the caregiver sample and integrate with qualitative interpretation) in REDCap (Vanderbilt University) via a shared email link. In addition to sociodemographic information, we collected data on caregivers' self-reported health (reported as "good," "very good," and "fair") and history of chronic illness. Caregiver burden was assessed with the 4-item Zarit Burden Interview (ZBI-4) [29], an abbreviated version of the original 22-item instrument to screen for caregiver burden. Each item is rated on a 5-point Likert scale ranging from 0 (never) to 4 (nearly always), yielding a total score from 0 to 16, with higher scores indicating greater perceived caregiver burden. Consistent with the original validation study, ZBI-4 scores were analyzed as a continuous measure. Perceived social support from family and friends was measured by the 10-item Lubben Social Network Scale (LSNS) [30] where responses are scored on a 6-point scale (0 - 5) and summed for total scores of 0 to 50. Higher scores indicate larger, supportive networks and lower risk of social isolation. The 14-item senior technology acceptance model [31] assessed attitudinal beliefs, control beliefs, health conditions, and gerontechnology anxiety. Each item is rated on a 10-point scale ranging from 1 (strongly disagree) to 10 (strongly agree). All items are summed to generate a cumulative score ranging from 14 to 140, where higher scores indicate a greater acceptance toward using technology.

Qualitative Interviews

Interviews were conducted by CK and PR, both investigators with prior experience in leading qualitative studies and conducting qualitative interviews. Both interviewers had no prior relationship with the study participants. In-person interviews were conducted in the community collaborator's private meeting rooms, and all interviews were recorded with 2 digital audio recorders to prevent data loss. Virtual interviews were conducted using a secure, HIPAA (Health Insurance Portability and Accountability Act)-compliant version of Zoom (education and health care version); audio and video were captured with Zoom's "record" and transcript functions.

Data Analysis

Quantitative survey data collected in REDCap were exported to SPSS (version 29; IBM Corp) for data management and analysis. Descriptive statistics were performed to characterize the study sample. Qualitative interview data collected in person were professionally transcribed and reviewed by members of the research team, while virtual interviews for which transcript data were automatically transcribed were similarly reviewed to verify transcript accuracy. All data (including digital audio and Zoom meeting recordings and raw interview transcripts) were stored on a secure, institutional drive. Finalized transcripts were deidentified and uploaded to ATLAS.ti (version 5.22.1; Lumivero, LLC) to assist with analysis. Data were analyzed using methods for thematic analysis [28] to assist with addressing the goals of our study. Two members of the research team with prior qualitative research experience first read and reread the entire deidentified, transcribed interview dataset to become familiar with the data. Then, each coder independently developed initial codes (semantic and underlying) to represent interpretive labels for the data that addressed the research question. The codebook contained a comprehensive list of codes with exemplar excerpts from the dataset. Coders met upon completion of the initial coding to discuss diverging assumptions, interpretations, and conceptual insights. Disagreements between coders were treated as opportunities to explore deeper, underlying patterns within the data. After all data were coded, an initial thematic map was developed by grouping together multiple codes sharing similar underlying concepts into higher-order themes and subthemes. The research team then met to review and refine the initial themes, discussing how accurately these represented the coded data (and the entire dataset) and whether to discard or merge themes as needed. Detailed descriptions of the final themes and subthemes were then finalized.

Trustworthiness

To ensure qualitative rigor and trustworthiness as described by Lincoln and Guba [32], we maintained a detailed audit trail, practiced reflexive memoing, held team discussions throughout the analytic process, and used data source triangulation methods to cross-validate the findings.

XSL•FO
RenderX

Ethical Considerations

Participants were offered a US $100 gift card for their time and contribution to the study. Ethics board approval for human subjects research was provided by the UTHealth Houston Institutional Review Board (number HSC-SN-22 - 0575). A trained research nurse (and member of the study team) obtained participants' informed consent using REDCap's e-consent feature. All participants were advised of the voluntary nature of study participation, encouraged to ask questions, and informed that they could withdraw from participation at any time.

Results

A total of 40 individuals participated in the study. The sample consisted of 32 caregivers, 6 clinicians, and 2 key informants to comprise a total of 5 focus group interviews and 2 individual interviews; of these, 4 focus groups were held with caregivers (8 individuals per group), 1 with clinicians, while individual interviews were conducted with key informants.

Demographic Survey

Demographic characteristics for caregiver participants are summarized in Table 1. Caregivers were an average age of 67.6 (SD 10.79) years, female (28/32, 87.5%), a combined category of Black Hispanic (14/32, 43.75%), college degree holders (12/32, 37.5%), and most (20/32, 62.5%) had an average annual income of US $30,000 to US $49,999. Data regarding care recipients were not collected. Caregivers had spent a mean of 7.15 (SD 3.31) years caring for their loved one. They reported having insufficient social connections according to the LSNS-Revised scale (mean 14.38, SD 4.54), while caregiver burden as measured through the Zarit Burden Interview was moderate (mean 7.44, SD 2.90). Finally, technology acceptance measured by the technology acceptance model questionnaire was relatively high (mean 95.00, SD 13.31) among participants, and most respondents (14/32, 43.75%) rated their health as "good."

XSL•FO
RenderX

Table . Caregiver participant characteristics (n=32).

Characteristics	Value
Age (y), mean (SD)	67.63 (10.79)
Social engagement (LSNS[a]), mean (SD)	14.38 (4.54)
Caregiver burden (ZBI-4[b]), mean (SD)	7.44 (2.90)
STAM[c], mean (SD)	95.00 (13.31)
Sex, n (%)	
Male	4 (12.5)
Female	28 (87.5)
Race/ethnicity, n (%)	
White	6 (18.75)
Non-Hispanic Black	6 (18.75)
Hispanic Black	14 (43.75)
Asian	6 (18.75)
Education, n (%)	
High school	2 (6.25)
Some college	2 (6.25)
College degree	12 (37.50)
Graduate degree or higher	6 (18.75)
Prefer not to answer	10 (31.25)
Income (US $), n (%)	
≤30,000	4 (12.5)
30,000-49,999	20 (62.5)
50,000-74,999	2 (6.25)
75,000 and more	4 (12.5)
Prefer not to answer	2 (6.25)
Self-rated health, n (%)	
Very good	16 (50)
Good	14 (43.75)
Fair	2 (6.25)
History of chronic illness, n (%)	
Diabetes	14 (43.75)
Hypertension	8 (25)
Overweight	10 (31.25)[c]

[a]LSNS: Lubben Social Network Scale.

[b]ZBI-4: 4-item Zarit Burden Interview.

[c]STAM: senior technology acceptance model.

Qualitative Interviews

Through our qualitative interviews, participants shared an overarching sentiment that using technology to support their caregiving activities is a balancing act that is largely dictated by their care recipient's needs. Overall, caregivers described a paradoxical relationship with technology, valuing its ability to provide practical assistance, enable emotional connection with others, and provide reassurance, while also recognizing that poorly designed, expensive, or difficult-to-learn technologies could increase the caregiving burden. As such, 3 themes emerged from the data illustrating the competing tensions regarding technology that exist for caregivers of homebound older adults: (1) These Technologies Aren't Designed for People Like Us, Facing These Problems, (2) Balancing the Scales, and

XSL•FO
RenderX

(3) Preserving a Sense of Self. The 8 related subthemes are further described below.

Theme 1: These Technologies Aren't Designed for People Like Us, Facing These Problems

Challenges to Technology Use Among Caregivers

Caregivers shared that they are faced with a technological conundrum; wherein most available technologies are not uniquely tailored to the caregiving experience, leaving caregivers to "pick and choose" what works for them at any given moment, while acknowledging that the investment of time and money to learn these technologies is a barrier to technology use. Three subthemes further illustrate this theme: (1) Cascading Costs of Technology Use, (2) Poor Technological Dexterity, and (3) Technology Unlocks a Black Box of Fears. These subthemes reflect a range of challenges related to accessing technology, required skillsets and abilities, and psychological barriers that are discussed in greater detail below.

Subtheme 1: Cascading Costs of Technology Use

Participants expressed a strong sentiment that using technology bears significant costs. These costs were not only limited to finances but also broadly included time and "mental bandwidth." Caregivers highlighted monetary costs, both up-front and long-term, that hinder technology use. Immediate costs associated with purchasing equipment weighed heavily on participants' decisions to commit to buying and using technology. "I'm willing to spend this much money on this thing and what if I don't [use it]?" (P15; 79-year-old female participant). Although many articulated the direct benefits that certain technologies may contribute to enriching their daily activities, these appeared to be overshadowed by concerns about the time needed to identify suitable, cost-effective technology, the learning curve associated with becoming comfortable with the devices, and setting them up. For instance, a 78-year-old female participant (P9) noted the following:

> *I'm afraid to try something new. And I don't like to talk to the tech people about it, because I don't even know what words to use or what to say. I don't know what [device or app] to pick or where to start.*

Moreover, the belief that the time spent researching, acquiring, and learning how to use a new technology outweighs its potential benefit was shared. One 80-year-old male caregiver (P18) noted the following:

> *…it's just gonna be something new, and it'll be there, and then pretty soon it'll be old and just getting dust on it…. And you don't want to spend it just spinning your wheels and not going any place.*

Insightfully, one 59-year-old male participant (P8) noted that devices lacking system interoperability (Apple iOS and Microsoft, for example) forced him to spend time trying to "make devices (and software) work together," leading to frustration and disappointment with the technology that is marketed to simplify an aspect of one's life. He further stated, "sometimes the technology itself is the lynchpin. It's just it. It gets you on a glitch, and it gets you frustrated." Compounded, these challenges serve as hindrances to technology use despite their potential benefits.

Subtheme 2: Poor Technological Dexterity

Poor technological dexterity, which encompasses challenges in quickly learning, adapting to, and effectively using technology, emerged as a second subtheme. Many caregivers shared that challenges with adopting and adapting to technology were related to lacking a baseline expertise and comfort with technology. These tasks were described as cumbersome, and many caregivers believed they were poorly equipped to engage in these activities. For instance, one 66-year-old female caregiver (P30) stated the following:

> *I'm not always ready for the newest thing, you know, and I know most older adults aren't ready for the newest thing, because it's something else that they have to think about, you know, or they have the process, or it's not that easy, you know, but you think differently when you tend to care for older adults*

Some caregivers stated that they lacked a degree of tech savviness, highlighting a generational digital divide as adult children or grandchildren were often consulted to address these gaps, as one 79-year-old female participant (P6) recounted that she relies on others stating she has "a son-in-law that he's very technology savvy and so…he's the one that put all of these things for me (laughs)." Specific activities that posed challenges included the following: managing passwords for multiple accounts, learning to use new devices, applications, and related platforms. Finally, physical impediments such as issues with eyesight and diminished manual dexterity were noted as significant detractors, with some caregivers emphasizing that these technologies are not designed with these age-related issues in mind. Many participants reported overcoming a steep learning curve to integrate technology into their lives; however, a 59-year-old male participant (P8) suggested that technology should dynamically adjust to its user's unique needs:

> *You look at the font because the older you get; our eyes get worse… I think technology needs to be set up in a way that one of your login things should be like, what's your age? And then the font can adapt itself there. There should be an easy technology to be able to do that.*

For some caregivers, technology emphasized the digital divide, while others saw potential on how it may be used to meet aging adults where they are.

Subtheme 3: Technology Unlocks a Black Box of Fears

The final subtheme centers on a variety of technology-related fears. These fears range from uncertainties that are anticipatory in nature, such as feeling overwhelmed by the prospect of learning something new and viewing that as a stressful activity, to fears of breaking devices and their replacement costs.

> *I've always been afraid of all these devices-I still am. I'm afraid of them. I don't know what I think. I think I'm gonna touch them and they're gonna break, and it's got to cost me more money* [P19; 71-year-old female participant.]

XSL•FO
RenderX

Many respondents reported that despite the potential benefits of technology in their caregiving activities, they were nonetheless doubtful of their own abilities to learn new skills and invest time and energy. A 79-year-old female participant (P16) stated the following:

> *...I think I'm pretty technologically savvy, but I don't want to go down that route, because. I see it as right now as counterproductive timewise and energy wise. It's frustrating to try and deal with it.*

Furthermore, a clinician (P11; 37-year-old female participant) noted that she often encounters caregivers who express the following fear:

> *fear of perhaps messing something up, the fear of the unknown, you know, "maybe I don't have the knowledge to be able to succeed with something that's very technology based"*

In an effort to avert failure, these caregivers will choose to avoid technology altogether.

Theme 2: Balancing the Scales

Facilitators of Technology Use in Caregiving

The theme "Balancing the Scales" illustrates caregivers', clinicians', and key informants' perspectives that technology may serve to temper the complexities of their caregiving experiences. One 66-year-old female caregiver (P20) evoked dramatic imagery in describing their caregiving role as a juggling act where she is "juggling your [care recipient's] life, our life, and my life." The notion of holding multiple balls in the air was a sentiment shared by many caregivers, and technology was described across interviews as a tool to help alleviate the burdens associated with this juggling act. Participants shared their standpoints on factors that facilitate the use of technology to support aging in place, as represented through 3 subthemes described in further detail below: (1) Tailored Teaching Engages and Empowers Caregivers, (2) Technology Streamlines and Organizes Complex Daily Routines, and (3) Promoting Sense of Security.

Subtheme 1: Tailored Teaching Engages and Empowers Caregivers

Participants expressed the desire for tailored teaching to meet their needs. While most caregivers were not opposed to technology, some were hesitant as exemplified by the following quote:

> *I'd be willing to look at them [devices] if someone trained me on how to use them; but to just jump into it without understanding the nuances...I'm not comfortable* [P24; 80-year-old male participant.]

Respondents cited varying modes of instruction they perceived as helpful. One 59-year-old male participant (P8) noted the importance of providing information that is relatable and salient for the end user:

> *They [organization] send[s] out a newsletter, and it's for seniors by seniors, and they do a lot of free technology training for seniors... make sure that you have someone providing that [information] within that person's age range, because it makes the communication easier.*

Suggested strategies for tailored teaching included involving adult children and grandchildren (with whom participants would eventually interact using these devices) as trainers and hosting community events for individuals/caregivers with novice technological abilities. Further, one clinician (P12; female who did not provide her age) suggested that a stepwise teaching approach branching from a basic introduction to smartphones (and similar devices) to guided hands-on training to using various applications would be optimal for caregivers. Thus, multiple training modalities are necessary toward addressing caregivers' needs for technological education and training.

Subtheme 2: Technology Streamlines and Organizes Complex Daily Routines

Caregivers described complex care routines alongside managing a functional household as activities that occupied their days. For many, the COVID-19 pandemic created a need to become technologically literate, as providers moved to telehealth modes of care delivery. This necessary change spurred some caregivers to become more receptive to an array of technological tools such as calendars to organize appointments, logs to document vitals and medication, patient care portals, online grocery delivery, and other online purchasing; these served to aid in streamlining care and organizing tasks of daily living. Telehealth visits were embraced by this group as some care recipients had mobility impairments. A 78-year-old female participant (P9) stated, "when I called to make the appointment with the doctor, I said, Yeah, give me the telehealth because I don't know that I can get him out." The convenience of telehealth was praised by many caregivers; however, some expressed frustration that providers no longer provided telehealth options or that insurance/Medicaid would no longer cover telehealth visits after the COVID-19 pandemic restrictions ended.

Many participants noted that their loved ones required round-the-clock supervision or care, making leaving their homes for simple errands a difficult task; therefore, grocery delivery services and online purchasing options greatly alleviated this burden. One participant remarked, "I don't have to go to the grocery store now. I can have groceries delivered, and how easy you know, and convenient, you know that is" (P10; 52-year-old female participant). Another further reinforced the ease of online purchasing, emphasizing that it's "so much more convenient, and it comes right away" (P8; 59-year-old male participant).

Subtheme 3: Promoting Sense of Security

With care recipients' safety at the forefront of all participants' perspectives, technology was frequently mentioned as a means of maintaining caregivers' peace of mind related to their loved one's safety and security. Many used cellphone-connected video monitoring systems to keep an eye on care recipients. A unique aspect of this technology is that it allows for care coordination between the caregiver and other family members. One caregiver described leaving her husband (care recipient) at her daughter's home, where she could view a live camera-feed and easily check in on him,

XSL•FO
RenderX

> *...so sometimes I go out and I say [to my daughter], " I don't see him in the camera," would you call? She can call because … it's [video camera] connected to her house.* [P6; 79-year-old female participant.]

Others described repurposing technology like AirTags as wearable watches to create inconspicuous trackers to monitor care-recipients with memory loss without infringing on their sense of autonomy and independence. Many caregivers remarked that monitoring care-recipients, especially among those with memory issues and wandering tendencies, maintains safety while avoiding institutional placement. However, despite these generally favorable views toward remote monitoring to promote safety and autonomy, one caregiver expressed a strong belief that "surveilling someone is a denigrating violation of their integrity as a person to be tracking them around" (P5; 71-year-old female participant).

Theme 3: Preserving a Sense of Self

Caregivers described using technology as a means of *Preserving a Sense of Self* amidst their all-consuming caregiving role. Participants shared that they personally used technology to buoy their sense of self through the following 2 subthemes of (1) Maintaining Social Connectedness and (2) Promoting Wellness/Self-Care, enabling them to persevere in their respective caregiving journeys.

Subtheme 1: Maintaining Social Connectedness

Caregivers reported using technology as a strategy to maintain social connections, though they conceded that technological connections are not a replacement for real-life connection. To combat the sense of isolation that has arisen during their caregiving experience, caregivers used video chats like FaceTime (Apple Inc) to connect with family members. Similarly, others used video platforms like Zoom for prayer groups, book clubs, music "jam" sessions, and church services, to maintain active in groups that they enjoyed prior to assuming the caregiving role. One participant praised virtual social-educational opportunities, remarking on a grant-funded opportunity where caregivers could take an online cooking class once a month,

> *The girl would get on there once a month, and we would cook. Everybody... They would have us on camera and we would cook these nutrition meals. […] so, it was fun, because they would give us recipes and, and things like that. That was really good, because we can't get out, you know?* [P3; 67-year-old female participant.]

Overall, technology can be leveraged to maintain a sense of normalcy for caregivers, providing them with the opportunity to connect with family and friends and to maintain personal interests, as they remain largely at home with the care recipient.

Subtheme 2: Promoting Wellness/Self-Care

Lastly, caregivers use technology as a form of self-care. Many remarked on the power of music as a tool to wind down both themselves and their care-recipient. One stated that her grandmother was very responsive to the calming effect of music during episodes of dementia-related agitation, noting:

> *that's an opportunity for sure to say "Alexa" - [when] my grandmother was having a meltdown, and that song started to play. I didn't have to do a lot of effort to play it. Then usually she would stop and listen to that song* [P17; 53-year-old female participant.]

Similarly, another recalled that downloading favorite programs on their smart TV allowed both the caregiver and care recipient to enjoy time together despite limitations in mobility and being immunocompromised:

> *So, I finally said, 'You know what I'm going to download BritBox for you, and I'm going to download the PBS passport' because we love watching the mysteries and stuff.* [P8; 59-year-old male participant.]

Others played digital games to distract themselves and relax, remarking that

> *They [care recipients] are resting... it's a time consumption, but it's also, I feel like, at least I'm using my brain. So sometimes I feel like all I use is my brawn and not my brain.* [P21; 79-year-old female participant.]

Additional examples of technology-mediated self-care included guided meditation apps for relaxation and attending virtual music jam sessions and concerts, as described previously.

Discussion

Principal Findings

Caregivers of older homebound adults describe complicated perspectives on using technology in their caregiving. Nationally, most caregivers of older adults are younger than 64 years old [33]. That most caregivers in our study were older adults themselves presents a unique perspective on technology use in this population, and the tensions they encounter in incorporating technology into their caregiving activities. Our sample was interviewed as COVID-19 restrictions were easing, which reflects the unique experience of those who may have been technology-averse but were forced, due to circumstances created by the pandemic, to integrate modes of virtual interaction into their daily lives. Although most participants found that keeping health appointments with providers was beneficial, allowing them to learn new technologies such as Zoom and other virtual meeting dashboards, this did not translate into a general embrace of technology beyond instrumental activities of daily living such as banking, shopping, or medication management. While participants reported high levels of technology acceptance on the technology acceptance model questionnaire, the qualitative interviews revealed nuanced experiences characterized by frustration, uncertainty, and hesitation when engaging with specific technologies. This divergence may reflect differences between general attitudes toward technology and the practical realities of its use in everyday contexts [34,35]. Previous literature generally supports the notion that research participants are more forthcoming in reporting sensitive topics through anonymous survey questionnaires [36]. It is possible that the context of a group discussion normalized participants' experiences [37] on the barriers they experience when using technology. While participants recognized the potential value

XSL•FO
RenderX

of technology in improving communication, support, and some daily tasks, their narratives highlighted barriers to usability, confidence, technical difficulties, and needs for assistance. These findings suggest that technology acceptance should not be interpreted solely as a willingness to use technology but should also consider the contextual factors that influence sustained engagement. Our first theme aligns with previous work, indicating that physical impairments, in addition to dispositional and technical factors, may limit technology use [38,39]. Nonetheless, the subthemes represent a glaring gap for technologies that are designed expressly for caregivers—there are few that combine their shared concerns (safety, connection to others in their situation, local resources, and clinical expertise). That caregivers expressed a need for resources that combine multimodal types of support aligns with recent guidance to enhance reach and accessibility for technology among older adults [40,41]. Thus, co-designing technologies with caregivers leverages their unique experiential expertise, increasing the chances that digital solutions will not only address explicit caregiving needs but also the subtle, often invisible demands, workarounds, and emotional complexities that shape daily caregiving and are unlikely to be identified through researcher or clinician-driven design alone.

The second theme, Balancing the Scales, illustrated the sentiment echoed across participants that caregivers wear many hats in their role, often feeling like the odds are stacked against them. This theme addressed facilitators to technology use on a broad level, indicating that tailored education and training that meets caregivers where they are in their caregiving journey may address the barriers identified in Theme 1. Its related subthemes further highlighted drivers of technology use as issues of practicality. Using technology to assist with instrumental activities of daily living allows caregivers to simplify their daily routines as they carry a significant mental load in managing their homes and loved ones' condition. The third subtheme, Promoting Sense of Security, centered around balancing caregivers' sense of vigilance—the idea of always being present or "on duty"—with their needs for peace of mind. However, some participants expressed privacy invasion concerns regarding technologies used for surveillance of their loved ones. An important ethical concern for the use of technologies such as home monitoring devices for adults with dementia and cognitive impairment is the appropriate balance between autonomy and safety. A balance between protecting individuals from harm while preserving independence and dignity is essential and should be negotiated carefully [42].

Interestingly, participants did not describe the feeling of a 24/7 responsibility as a burden during their interviews, despite reporting moderate caregiver burden (ZBI-4 scores), but rather as an inherent feature of their situation. It may be that these caregivers have normalized their caregiving demands as an expression of commitment or responsibility due to their familial relation (spouse, adult child, and grandchild) to the care recipient.

The final theme, Preserving a Sense of Self, speaks to the idea that caregivers also have personally driven needs for technology use that may buffer the strains of caregiving. Although it was beyond the scope of this study to examine participants' personal motivations for caregiving, it was evident that a sense of duty and obligation were underlying drivers and that the toll of the caregiving experience called for respite. The subtheme of Maintaining Social Connections and Promoting Wellness further enforces the importance of caregivers maintaining relationships and interests that are critical for their own identities and tempering the role loss experienced by caregivers that has been described by others [43,44]. That caregivers' LSNS scores suggested that participants were experiencing social isolation aligns with this final theme, expressing needs for connection and support. Indeed, loneliness and isolation are pervasive among caregivers of older adults [45,46] and deleteriously affect their mental and physical health [47]. Left unaddressed, declines in caregiver health may negatively impact care recipients [48]. Interestingly, caregiver respondents rated their health as "good," suggesting that caregivers may perceive themselves as physically healthy while nonetheless experiencing substantial psychosocial demands associated with caregiving. This distinction underscores the importance of assessing caregiver well-being across multiple domains, as self-rated health alone may not fully capture the emotional and social consequences of the caregiving role. Finally, despite the robust evidence-base of interventions to reduce loneliness and isolation for older adults, caregivers are less frequently the intended target population [49]. Our findings suggest that while caregivers value technology as a means of accessing information and support, digital tools are likely to be most effective when they foster meaningful human connection rather than replacing it.

Limitations

There are several limitations to our study. First, our sampling frame draws from participants enrolled with one community service agency in Houston, which may not comprehensively represent the diversity of caregiving experiences or those in the general US population. Similarly, geographical differences in access to resources, internet connectivity, and linguistic preferences may have influenced adoption and use of technology among caregivers. As such, the linguistic ability within our research team and our partner organization was limited to English, so our study does not address important issues surrounding digital inclusivity and the application of these approaches across a variety of sociocultural situations.

That the majority of the sample comprised women reflects the larger picture of caregiving in the United States, where females disproportionately (55%) assume caregiving roles [50]. Further, variations in the treatment burden among caregivers according to the type, frequency, and intensity of care provided are difficult to discern, as we did not collect detailed information regarding the care recipients. Finally, since this study was conducted during the post-COVID context, it is difficult to disentangle the lasting consequences of the pandemic from those that reflect broader technological and societal trends in adopting technology in this population. Given the rapidly evolving technological landscape, participants' experiences may reflect practices specific to the post-COVID period, which limits the transferability of the findings as digital tools and technology continue to change.

XSL•FO
RenderX

Conclusions

Caregivers and their care recipients navigate chronic illness in-tandem, using available resources such as technology to alleviate the burdens of care. While technology has considerable potential to support caregiving, it is not a solution in and of itself. As reflected in participants' narratives, technology can also impose additional burdens when it is difficult to use or fails to address their evolving needs. The present technological landscape requires that caregivers and older adults seek out and initiate the process of adopting new technology to integrate it into their lives; its potential is optimized when individuals are not only adequately trained but also fully aware of its capacities and limitations. Our study demonstrates that caregivers of homebound older adults adopt technology tools that reduce their concrete burdens such as time, travel, and monitoring. However, they underuse technologies that introduce financial, cognitive, or emotional costs. These dueling tensions should be recognized by researchers, clinicians, and technology developers as contextual determinants to using technology across the caregiving trajectory. Our findings highlight that targeted, multisectoral efforts toward digital inclusivity among vulnerable groups such as caregivers and their care recipients are sorely needed. Toward that end, it is imperative that health providers, social services clinicians, and tech developers are fully engaged with end users, such as caregivers, in co-design approaches that ensure not only the acceptability and usability of technology but that these technologies are ethically designed while considering caregivers' everyday limitations, financial constraints, emotional workload, and digital confidence.

Acknowledgments

The authors sincerely thank Ms Andrea Williams at CarePartners for facilitating our study recruitment efforts and Harini Kanamarlapudi and Hannah Albano for their assistance in revising this manuscript.

Generative AI Use

Generative AI tools were not used in any stage of the preparation of this manuscript.

Funding

This study was supported by the UTHealth Houston Center for Clinical and Translational Sciences (CTSA) grant (UM1TR004906) and the Research and Extension Experiential Learning for Undergraduate (REEU) Program of the National Institute of Food and Agriculture, USDA, grant number 2022 - 68018 - 36607.

Data Availability

The data that support the findings of this study are available from the corresponding author, CK, upon reasonable request.

Authors' Contributions

Conceptualization: CK (lead), KS (equal)
Formal analysis: CK (lead), ES (supporting)
Funding acquisition: CK
Investigation: CK
Methodology: CK, KM, PR
Project administration: CK (lead), JL (equal), ES (supporting)
Resources: KS
Supervision: CK

Conflicts of Interest

None declared.

References

1. 2023 profile of older Americans. : Administration for Community Living, U.S. Department of Health and Human Services; 2024 URL: https://acl.gov/sites/default/files/Profile%20of%20OA/ACL_ProfileOlderAmericans2023_508.pdf [accessed 2026-10-13]
2. Watson KB, Wiltz JL, Nhim K, Kaufmann RB, Thomas CW, Greenlund KJ. Trends in multiple chronic conditions among US Adults, by life Stage, behavioral risk factor surveillance system, 2013-2023. Prev Chronic Dis 2025 Apr 17;22:E15. [doi: 10.5888/pcd22.240539] [Medline: 40245168]
3. Informal caregiver supply and demographic changes: review of the literature. : Office of the Assistant Secretary for Planning and Evaluation (ASPE), U.S. Department of Health and Human Services; 2021 URL: https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//198051/ICSupplyLR.pdf [accessed 2024-10-02]
4. Leff B, Ornstein K, Pomeroy ML, Ritchie C, Reckrey J, Sheehan O. Prevalence of robust older adults among homebound older adults. Innov Aging 2025 Dec;9(Supplement_2). [doi: 10.1093/geroni/igaf122.1028]

XSL•FO
RenderX

5. Reckrey J, Ritchie C. Home, but homebound: identifying and supporting homebound older adults. Innov Aging 2025 Dec;9(Supplement_2). [doi: 10.1093/geroni/igaf122.1027]
6. Chan KT, Marsack-Topoleswki C, LaFave S, et al. Teaching note-supporting homebound older adults and caregivers through integrative service learning. J Soc Work Educ 2023;59(4):1249-1257. [doi: 10.1080/10437797.2022.2045238] [Medline: 38143590]
7. Reckrey JM, Federman A, Bollens-Lund E, Morrison RS, Ornstein KA. Homebound status and the critical role of caregiving support. J Aging Soc Policy 2020;32(6):590-603. [doi: 10.1080/08959420.2019.1628625] [Medline: 31242823]
8. Gorawara-Bhat R, Graupner J, Molony J, Thompson K. Informal caregiving in a medically underserved community: challenges, construction of meaning, and the caregiver–recipient dyad. SAGE Open Nurs 2019;5:2377960819844670. [doi: 10.1177/2377960819844670] [Medline: 33415239]
9. Kilmer G, Omura JD, Bouldin ED, et al. Changes in health indicators among caregivers - United States, 2015-2016 to 2021-2022. MMWR Morb Mortal Wkly Rep 2024 Aug 29;73(34):740-746. [doi: 10.15585/mmwr.mm7334a2] [Medline: 39207933]
10. Sun AY, Finkelstein E, Ouchida K. Social isolation, loneliness, and caregiver burden among paid and unpaid caregivers of homebound older adults. Innov Aging 2019 Nov 8;3(Supplement_1):S967. [doi: 10.1093/geroni/igz038.3505]
11. Alltag S, Conrad I, Riedel-Heller SG. Caregiver burden among older informal caregivers of patients with dementia and its influence on quality of life: a systematic literature review [Article in German]. Z Gerontol Geriatr 2019 Aug;52(5):477-486. [doi: 10.1007/s00391-018-1424-6] [Medline: 29992357]
12. Kirvalidze M, Mateo-Abad M, Beridze G, et al. Quality of life among older informal caregivers in Sweden: the role of loneliness and social isolation. Qual Life Res 2026 Feb 1;35(3):52. [doi: 10.1007/s11136-025-04156-x] [Medline: 41621026]
13. Gifford A, Marmelat V, Beadle JN. A narrative review examining the utility of interpersonal synchrony for the caregiver-care recipient relationship in Alzheimer's disease and related dementias. Front Psychol 2021;12:595816. [doi: 10.3389/fpsyg.2021.595816] [Medline: 34025493]
14. Penning MJ, Wu Z. Caregiver stress and mental health:impact of caregiving relationship and gender. Gerontologist 2016 Dec;56(6):1102-1113. [doi: 10.1093/geront/gnv038] [Medline: 26035875]
15. Mickler AK, Leff B, Eaton England A, et al. Understanding the daily experiences and perceptions of homebound older adults and their caregivers: a qualitative study. J Appl Gerontol 2021 Dec;40(12):1722-1732. [doi: 10.1177/0733464821990171] [Medline: 33522364]
16. McManus C, Oseroff BH, Zimbroff RM, Riser TJ, Ornstein KA, Sheehan OC. Shedding light on an invisible population in geriatric medicine: a scoping review of older adults who are homebound. J Am Geriatr Soc 2025 Nov;73(11):3562-3576. [doi: 10.1111/jgs.19620] [Medline: 40729452]
17. Musich S, Wang SS, Kraemer S, Hawkins K, Wicker E. Caregivers for older adults: prevalence, characteristics, and health care utilization and expenditures. Geriatr Nurs 2017;38(1):9-16. [doi: 10.1016/j.gerinurse.2016.06.017] [Medline: 27477086]
18. Wilkins VM, Sirey JA, Bruce ML. Caregiver reward and burden: differing constructs in family members providing care for medically ill homebound older adults. J Aging Health 2020;32(5-6):361-370. [doi: 10.1177/0898264319825760] [Medline: 30675813]
19. Lacey RE, Xue B, Di Gessa G, Lu W, McMunn A. Mental and physical health changes around transitions into unpaid caregiving in the UK: a longitudinal, propensity score analysis. Lancet Public Health 2024 Jan;9(1):e16-e25. [doi: 10.1016/S2468-2667(23)00206-2] [Medline: 37977176]
20. Reinhard SC, Given B, Petlick NH, Bemis A. Supporting family caregivers in providing care. In: Hughes RG, editor. Patient Safety and Quality: An Evidence-Based Handbook for Nurses: Agency for Healthcare Research and Quality (US); 2008:149-169 URL: http://www.ncbi.nlm.nih.gov/books/NBK2665/ [accessed 2026-02-09]
21. Zhai S, Chu F, Tan M, Chi NC, Ward T, Yuwen W. Digital health interventions to support family caregivers: an updated systematic review. Digit Health 2023;9:20552076231171967. [doi: 10.1177/20552076231171967] [Medline: 37223775]
22. Park JYE, Tracy CS, Gray CS. Mobile phone apps for family caregivers: a scoping review and qualitative content analysis. Digit Health 2022;8:20552076221076672. [doi: 10.1177/20552076221076672] [Medline: 35154806]
23. Kelley MM, Powell T, Camara D, et al. Mobile health apps, family caregivers, and care planning: scoping review. J Med Internet Res 2024 May 23;26:e46108. [doi: 10.2196/46108] [Medline: 38781588]
24. Chi NC, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. J Telemed Telecare 2015 Jan;21(1):37-44. [doi: 10.1177/1357633X14562734] [Medline: 25475220]
25. McCloskey R, Skerry L, Keeping-Burke L, et al. Caring for self while caring for others: impact of wearable health monitoring devices on self-care of family/friend caregivers of individuals living with a dementia. Int J Adv Nurs Educ Res 2019 Dec 30;4(3):69-80. [doi: 10.21742/ijaner.2019.4.3.10]
26. Lee S, Ory MG, Vollmer Dahlke D, Smith ML. Technology use among older adults and their caregivers: cross-sectional survey study. JMIR Aging 2024 May 1;7:e50759. [doi: 10.2196/50759] [Medline: 38717339]
27. Fernandez-Bueno L, Torres-Enamorado D, Bravo-Vazquez A, Rodriguez-Blanco C, Bernal-Utrera C. Technological innovations to support family caregivers: a scoping review. Healthcare (Basel) 2024 Nov 25;12(23):2350. [doi: 10.3390/healthcare12232350] [Medline: 39684972]

XSL•FO
RenderX

28. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
29. Bédard M, Molloy DW, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. Gerontologist 2001 Oct;41(5):652-657. [doi: 10.1093/geront/41.5.652] [Medline: 11574710]
30. Lubben JE. Assessing social networks among elderly populations. Fam Community Health 1988;11(3):42-52. [doi: 10.1097/00003727-198811000-00008]
31. Chen K, Lou VWQ. Measuring senior technology acceptance: development of a brief, 14-item scale. Innov Aging 2020;4(3):igaa016. [doi: 10.1093/geroni/igaa016] [Medline: 32617418]
32. Lincoln YS. Emerging criteria for quality in qualitative and interpretive research. Qual Inq 1995 Sep;1(3):275-289. [doi: 10.1177/107780049500100301]
33. Wolff JL, Cornman JC, Freedman VA. The number of family caregivers helping older US adults increased from 18 million to 24 million, 2011-22. Health Aff (Millwood) 2025 Feb;44(2):187-195. [doi: 10.1377/hlthaff.2024.00978] [Medline: 39899774]
34. Sun J, Liu Y, Zhang C, Xing Y, Zhou W, Luan W. Experiences of ageism in mHealth app usage among older adults: interview study among older adults based on extended unified theory of acceptance and use of technology and risks of ageism models. JMIR Aging 2026 Jan 7;9:e79457. [doi: 10.2196/79457] [Medline: 41500194]
35. An J, Zhu X, Wan K, et al. Older adults' self-perception, technology anxiety, and intention to use digital public services. BMC Public Health 2024 Dec 19;24(1):3533. [doi: 10.1186/s12889-024-21088-2] [Medline: 39702092]
36. Murdoch M, Simon AB, Polusny MA, et al. Impact of different privacy conditions and incentives on survey response rate, participant representativeness, and disclosure of sensitive information: a randomized controlled trial. BMC Med Res Methodol 2014 Jul 16;14:90. [doi: 10.1186/1471-2288-14-90] [Medline: 25027174]
37. Kornbluh M. Facilitation strategies for conducting focus groups attending to issues of power. Qual Res Psychol 2023;20(1):1-20. [doi: 10.1080/14780887.2022.2066036]
38. Chen K, Chan AHS. Use or non-use of gerontechnology—a qualitative study. Int J Environ Res Public Health 2013 Sep 30;10(10):4645-4666. [doi: 10.3390/ijerph10104645] [Medline: 24084674]
39. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
40. Lindeman DA, Kim KK, Gladstone C, Apesoa-Varano EC. Technology and caregiving: emerging interventions and directions for research. Gerontologist 2020 Feb 14;60(Suppl 1):S41-S49. [doi: 10.1093/geront/gnz178] [Medline: 32057082]
41. National Academies of Sciences, Engineering, and Medicine; Division of Behavioral and Social Sciences and Education; Health and Medicine Division; Board on Health Care Services; Board on Health Sciences Policy; Forum on Aging, Disability, and Independence, Lustig TA, Cilio CM. Artificial Intelligence Applications for Older Adults and People with Disabilities: Balancing Safety and Autonomy: Proceedings of a Workshop—in Brief: National Academies Press (US); 2019. URL: http://www.ncbi.nlm.nih.gov/books/NBK541253/ [accessed 2026-10-08]
42. Wang J, Arthanat S, Opuda E, et al. Ethical considerations in home monitoring technologies for persons living with cognitive impairment: a scoping review. Gerontologist 2026 Feb 12;66(3):gnaf261. [doi: 10.1093/geront/gnaf261] [Medline: 41269110]
43. Skaff MM, Pearlin LI. Caregiving: role engulfment and the loss of self. Gerontologist 1992 Oct;32(5):656-664. [doi: 10.1093/geront/32.5.656] [Medline: 1427278]
44. Montgomery RJ, Kosloski KK. Caregiving as a process of changing identity: implications for caregiver support. Generations 2009;33(1):47-52 [FREE Full text]
45. Qian Y, Pomeroy ML, Petchler CM, Abshire Saylor M, Cudjoe TKM, Ornstein KA. The epidemiology of social isolation and loneliness among family and unpaid caregivers of older adults: findings from the national study of caregiving. J Gerontol B Psychol Sci Soc Sci 2025 Jun 10;80(7):gbaf101. [doi: 10.1093/geronb/gbaf101] [Medline: 40411495]
46. Li L, Wister AV, Mitchell B. Social isolation among spousal and adult-child caregivers: findings from the Canadian Longitudinal Study on Aging. J Gerontol B Psychol Sci Soc Sci 2021 Aug 13;76(7):1415-1429. [doi: 10.1093/geronb/gbaa197] [Medline: 33170276]
47. Perez SG, Nuccio AG, Stripling AM. A rapid review of the detrimental impact of loneliness and social isolation in caregivers of older adults. Am J Geriatr Psychiatry 2021 Apr;29(4):S122-S123. [doi: 10.1016/j.jagp.2021.01.117]
48. Schulz R, Beach SR, Friedman EM. Caregiving factors as predictors of care recipient mortality. Am J Geriatr Psychiatry 2021 Mar;29(3):295-303. [doi: 10.1016/j.jagp.2020.06.025] [Medline: 32718853]
49. Fakoya OA, McCorry NK, Donnelly M. Loneliness and social isolation interventions for older adults: a scoping review of reviews. BMC Public Health 2020 Feb 14;20(1):129. [doi: 10.1186/s12889-020-8251-6] [Medline: 32054474]
50. Unpaid eldercare in the United States-2023-2024: data from the American Time Use Survey. : U.S. Bureau of Labor Statistics, U.S. Department of Labor; 2025 Sep 25 URL: https://www.bls.gov/news.release/pdf/elcare.pdf [accessed 2026-08-02]

XSL•FO
RenderX

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
HIPPA: Health Insurance Portability and Accountability Act
LSNS: Lubben Social Network Scale
ZBI-4: 4-item Zarit Burden Interview

Edited by H Köttl; submitted 26.Mar.2026; peer-reviewed by I Markoc, S Lee; revised version received 18.Aug.2026; accepted 26.Aug.2026; published 06.Nov.2026.

Please cite as:
Katigbak C, Shenoy E, Ramaswamy P, Lee J, Milbury K, Scott K
Barriers and Facilitators to Technology Use Among Caregivers of Homebound Older Adults: Descriptive Qualitative Study
JMIR Aging 2026;9:e96226
URL: https://aging.jmir.org/2026/1/e96226
doi:10.2196/96226

© Carina Katigbak, Esha Shenoy, Padmavathy Ramaswamy, Jessica Lee, Kathrin Milbury, Katie Scott. Originally published in JMIR Aging (https://aging.jmir.org), 6.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Use of Digital Tools and Environmental Practices Among Informal Caregivers of Older Adults in the European Union Embracing the Triple Transition: Cross-Sectional Study

Beatrice Ietto[1], PhD; Mirko Di Rosa[2], PhD; Giovanni Lamura[1], PhD; Sabrina Quattrini[1], MS; Ricardo Rodrigues[3], PhD; Sara Santini[1], PhD; Aviad Tur-Sinai[4], PhD; Marco Socci[1], PhD

[1]Centre for Socio-Economic Research on Ageing, IRCCS INRCA, Ancona, Italy

[2]Centre for Biostatistics and Applied Geriatric Clinical Epidemiology, IRCCS INRCA, C.da Muoio Piccolo, Cosenza, Italy

[3]ISEG Research, Lisbon School of Economics and Management, University of Lisbon, Lisbon, Portugal

[4]School of Public Health, The Faculty of Social Welfare and Health Sciences, University of Haifa, Haifa, Israel

Corresponding Author:
Mirko Di Rosa, PhD
Centre for Biostatistics and Applied Geriatric Clinical Epidemiology, IRCCS INRCA, C.da Muoio Piccolo, Cosenza, Italy

Abstract

Background: The European Union is currently facing a demographic, digital, and environmental transition collectively known as the triple transition. Although these trends significantly impact long-term care services, they have typically been studied in isolation, overlooking potential synergies between the adoption of digital tools and environmental practices among informal caregivers of older adults. This study uniquely bridges this gap by applying an integrative social-ecological-infrastructural systems framework to examine how these transitions interact in everyday caregiving contexts.

Objective: This study examined the relationship between the use of digital tools and eco-friendly practices among informal caregivers of older adults across the European Union and whether the adoption of digital solutions related differently to specific types of eco-friendly behaviors. Unlike previous research focusing on isolated aspects of each transition, we examined their interconnections, providing empirical evidence for developing integrated policy approaches.

Methods: Using data from the Survey of Gender Gaps in Unpaid Care, Individual and Social Activities project (n=5761 informal caregivers of adults aged ≥65 years), we analyzed the association between frequency of use of digital tools for caregiving and environmental practices. We used the social-ecological-infrastructural systems conceptual framework and the theory of planned behavior (TPB) to interpret findings. Analyses included descriptive statistics, Spearman correlations, and multilevel mixed-effects regression models.

Results: A weak positive association was observed between the frequency of use of digital tools and environmentally sustainable behaviors (ρ=0.28; *P*<.001). Among individual behaviors, the largest associations were observed for consumption-related behaviors (fair-trade products: ρ=0.30; used items: ρ=0.30; eco-friendly products: ρ=0.25), whereas the associations were weaker for waste management and transportation choices and were not statistically significant for recycling (ρ=−0.04; *P*=.23) and mindful consumption of natural resources (ρ=−0.02; *P*>.99). Respondents using digital tools daily had scores of environmentally sustainable behavior (SESBs) 22% higher than nonusers. Demographic patterns revealed higher use of digital tools among younger, more educated, and employed caregivers, indicating that any potential digital-environmental co-occurrence may be unevenly distributed across the caregiver population.

Conclusions: The use of digital tools may be related to environmental behavior through improved access to information, peer norms, and a greater perceived ability to act, though these mechanisms were not directly tested. The stronger associations with consumption-related behaviors indicate that digital platforms may be particularly relevant to caregivers' access to eco-friendly products, second-hand goods, fair-trade options, and sustainable local services. By contrast, the lack of statistically significant correlations with recycling and resource conservation suggests that digital solutions alone may be insufficient to address behaviors that depend on structural conditions, physical infrastructure, local service availability, or established household routines. Policies should consider digital inclusion, caregiver support, and green infrastructure as interconnected areas while ensuring that older, less educated, and more burdened caregivers are not excluded from these potential co-benefits.

(JMIR Aging 2026;9:e81556) doi:10.2196/81556

KEYWORDS

informal caregivers; digital tools; environmental behaviors; older adults; triple transition; European Union

XSL•FO
RenderX

Introduction

Setting the Context: The Triple Transition in Long-Term Care

The term "Triple Transition" has been used to describe the concomitant demographic, digital, and green transitions that the European Union (EU) aims to achieve through the integration of social inclusivity, technological advancement, and environmental sustainability [1]. According to the Organisation for Economic Co-operation and Development (OECD) [1], these transitions are reshaping societal structures, presenting unique challenges and opportunities in many sectors, including health care and, more specifically, long-term care (LTC) services.

In terms of demographic transition, the European LTC sector faces growing challenges such as rapid demographic aging and a shrinking working-age population. In Europe, median age increased from 39.0 to 44.5 years (2003 - 2023), life expectancy increased by almost 4 years (2002 - 2023), and the old-age dependency ratio grew by 9.3 percentage points (2001 - 2020) [2,3]. These shifts might have serious implications for care provision, creating an unprecedented demand for care services together with a growing shortage of care staff [1,4].

The digital transition uses technology to enhance care delivery through electronic health records, telehealth, AI-powered diagnostics, and remote monitoring, among others [5,6]. Although some researchers predict digitalization will improve care efficiency [7], others caution that the benefits may be unequally distributed, causing a deeper digital exclusion and discrimination [1,8].

The green transition aims at fostering environmentally friendly practices as well as monitoring climate change [1]. As the health care sector contributes about 4.4% of global greenhouse gas emissions [9], care systems have faced increasing pressure to adopt more sustainable practices and policies [10,11].

While population aging is causing high pressure on the LTC sector and creating new challenges [12,13], digitalization and sustainability have been investigated as potential approaches to reduce resource constraints and improve the delivery of care [7,14-16].

In fact, while digital technologies have been primarily studied for enhancing care efficiency and quality [7,17-20], environmental practices seem to have been mainly investigated for their impact on resource efficiency, waste reduction, and management systems [21,22].

According to the OECD [1], these transitions have typically been studied in isolation, and potential synergies between digitalization and sustainability have been largely overlooked [23,24]. For example, digital solutions might simultaneously improve care efficiency as well as reduce environmental impact [25,26]. Similarly, LTC research has also maintained a siloed approach, with only a few studies exploring the integration and potential synergies between the digital and environmental transition [27-30]. Finally, Eurofound [3] calls for a more comprehensive understanding of these intersecting challenges, as this might inform solutions that jointly address rising care demand, technological change, and sustainability requirements in care systems.

The Social-Ecological-Infrastructural-System Framework

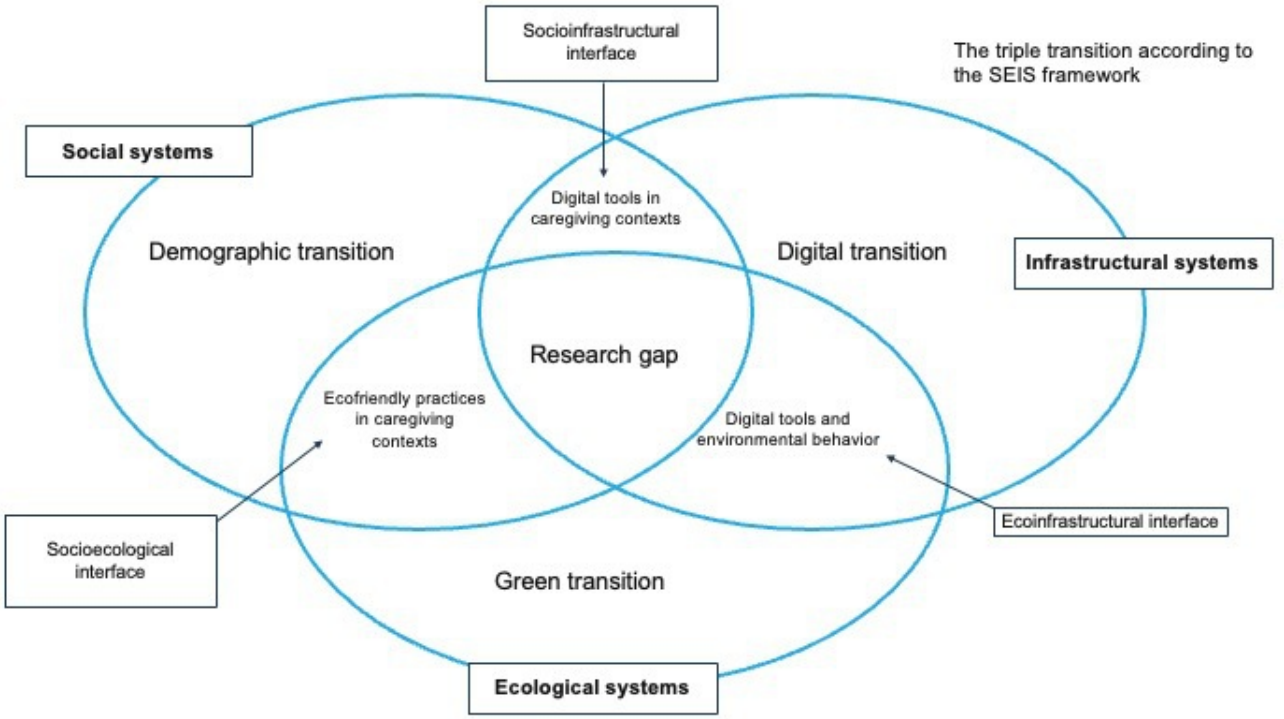
To better understand the interconnections between the demographic, digital, and green transitions, this study uses the social-ecological-infrastructural systems (SEIS) framework developed by Ramaswami et al [31]. This framework belongs to a broader family of conceptual approaches that have been used to examine the relationships between the social, technological, and ecological dimensions of sustainability [32-34]. These approaches have been used to investigate how such dimensions are interconnected and to verify to what extent changes in one domain might affect the other ones [35,36]. Specifically, the SEIS framework merges 3 interconnected dimensions: social systems (people, communities, and institutions), infrastructural systems (transportation networks and engineered solutions), and ecological systems (natural environments and resources). Although the SEIS framework was originally conceived for urban sustainability [31], its application has been extended to other settings, including health care and sustainability [37]. More recent applications have also included digital technologies as a critical component of the infrastructural system [38,39]. Therefore, the SEIS framework is particularly suited to study the Triple Transition in LTC, as it directly aligns with the demographic (social), digital (infrastructural), and environmental (ecological) transitions occurring in the care sector.

The demographic transition primarily affects social systems through changes in population structure (ie, population aging), care needs, and the role of informal caregivers, for example, family or nonfamily members who provide unpaid care, outside of a formal framework, to someone with a chronic illness, disability, or other long-lasting care need [40]. The digital transition impacts the infrastructural system through the introduction of new tools, platforms, and devices for care delivery. Finally, the green transition seeks to develop more sustainable approaches for the use of resources and waste management.

While the Triple Transition has been viewed at these systems as separate pillars, the SEIS framework focuses on the integration within and among them. Specifically, it distinguishes 3 interfaces: the socioinfrastructural (how digital tools are integrated into caregiving contexts), the socioecological (how caregivers engage with eco-friendly practices), and the eco-infrastructural interface (how digital solutions contribute to environmental sustainability). The central intersection of the 3 systems represents the research gap addressed in this paper: the relationship between the use of digital tools and eco-friendly practices among informal caregivers (Figure 1).

XSL•FO
RenderX

Figure 1. The Triple Transition according to the social-ecological-infrastructural systems (SEIS) framework (Ramaswami et al [31]). SEIS: social-ecological-infrastructural systems.



This study investigates how the use of digital tools is interconnected with the adoption of eco-friendly practices among informal caregivers. In doing so, it aims to better understand how the 3 dimensions of the SEIS framework interact with each other within the specific context of LTC. This integrated approach is consistent with more contemporary developments in sustainability theories that emphasize the need for systems thinking rather than siloed approaches [40].

Studying the intersection between informal caregivers' use of digital tools and environmental practices provides a unique opportunity to understand how the Triple Transition manifests in care settings. Informal caregivers are a particularly important population group, increasingly affected by demographic shifts, digitalization, and growing environmental concerns. Situated at this intersection, they are especially relevant for understanding how these transitions interact in practice and how policy might better integrate these interconnected domains rather than addressing them in isolation. Drawing on the SEIS framework, the following 3 subsections examine the relationship between the use of digital tools and environmental behaviors; the adoption of digital tools in caregiving contexts; and eco-friendly practices in LTC.

Digital Tools and Environmental Behaviors

The relationship between the adoption of digital technologies and environmental behaviors has gained increasing attention as both digitalization and environmental awareness have become more widespread [41,42]. This relationship is theoretically grounded in the concept of behavioral spillover effects [43], in which engagement in one domain (use of digital tools) may influence behaviors in another domain (environmental practices) through shifts in values, knowledge, and capabilities. In fact, digital tools may serve as enablers of environmental behaviors in a number of ways, including more efficient access to information, facilitating the creation of support communities, and making eco-friendly consumption easier to perform. According to Hilty and Aebischer [44], digital technology's environmental effects can be categorized into 3 distinct orders.

First-order effects concern the direct environmental impacts of the life cycle of digital devices from production through use to disposal [44]. In the health care context, these impacts include the environmental footprint of digital and telehealth devices [27]. Second-order effects involve how digital technologies enable process optimization, dematerialization, and the substitution of physical resources with digital alternatives [44]. In health care and caregiving, digital services might reduce transportation emissions through telemedicine [45] and lower paper and other resource consumption through the adoption of electronic health records [22,46,47].

Third-order effects, which are considered to have the most profound impact, are the focus of this paper. They involve how digital technologies enable broader behavioral changes [44]. According to Santarius et al [48], digital technologies provide easier access to information about sustainable options, enable remote activities that reduce transportation needs, and foster the formation of online communities, in which peers share and discuss sustainable behaviors. Digital tools can also increase environmental awareness through educational content and real-time feedback [49,50].

Digital platforms make eco-friendly behaviors more convenient by reducing transaction costs of purchasing eco-friendly products [51-53]. Finally, online communities reinforce environmental norms and values through shared experiences and peer influence [54-56]. This explains how digital technologies might relate to environmental behaviors, even

XSL•FO
RenderX

when digital tools themselves are not explicitly used for environmental purposes. While the understanding of digital tools' environmental impact has grown, significant gaps remain in understanding how these effects manifest in specific populations such as caregivers. Research has primarily focused on first- and second-order impacts of digital technologies on environmental sustainability in health care [22,46]. The third-order behavioral effects, specifically how the use of digital tools relates to caregivers' environmental behaviors, remain largely unexplored.

Digital Tools in Caregiving Contexts

Digital tools have become increasingly integrated into caregiving practices, though adoption remains uneven. In the United States, approximately 53% of informal caregivers report using at least one digital tool for caregiving, although this varies significantly by tool type and demographic factors. For example, about 80% report feeling comfortable using basic technologies, compared with only 50% with advanced tools like smart assistants [57,58]. In other countries, where comparable data are scarce, digital solutions seem less widespread but increasingly being used to support informal caregiving [3,59,60]. In fact, they have been used for several purposes, including providing information and educational material [61], platforms for organizing caregiving activities [62], remote monitoring technologies [20], communities for social support [59,63], and administrative tools [64]. Despite variations in the use of digital tools across different caregivers' activities, specialized caregiving apps have a relatively low adoption rate (25%) [65]. Furthermore, a significant digital divide still exists based on the caregivers' demographic characteristics. For example, Shaffer et al [66] found that age and education are strong predictors of technology adoption, with 40.4% of caregivers aged 65 years and older identifying themselves as noninternet users versus just 13.7% of younger caregivers. The likelihood of being offline decreases by 40% among tertiary-educated caregivers, while rural caregivers show higher rates of digital exclusion [67]. The most critical barriers to adoption include financial constraints, digital literacy limitations, and privacy issues [64,67].

Eco-Friendly Behaviors in LTC

The adoption of eco-friendly practices in LTC, particularly among informal caregivers, represents an emerging research area. Recent research has revealed that informal caregivers of older adults engage significantly more in environmentally sustainable behaviors compared to noncaregivers [68]. Their performance is particularly strong in sustainable consumption choices (eco-friendly products, fair-trade items, and second-hand goods), avoiding single-use plastics, prioritizing local food products, and using low-carbon transportation. While caregivers outperform noncaregivers across all measured environmental behaviors, the gap is somewhat smaller for household practices such as recycling and resource conservation [68].

This pattern suggests that caregivers' environmental engagement extends beyond basic resource-saving behaviors to reflect deeper environmental values. The experience of caring for vulnerable individuals may foster a broader sense of responsibility and interconnectedness that extends to environmental concerns [69] and fosters a heightened ethic of care that embraces environmental stewardship [70,71]. Furthermore, because caregivers are typically in charge of households, they make several environmental decisions about consumption, waste disposal, and resource use [72]. This might foster an increased awareness of how environmental factors impact an individual's health, leading to the adoption of a long-term consciousness of the importance of sustainability for future generations [73].

Nonetheless, the distinctive constraints that caregivers face, including time limitations [74,75], financial pressures often associated with longer hours of care [76], and mental and physical exhaustion [76,77], could reduce their engagement in environmental practices. Understanding how digital tools might help overcome these barriers is particularly relevant to supporting sustainable behaviors among this substantial population group, which already shows a predisposition toward eco-friendly practices [75,78,79].

Research Questions

This study offers practical insights into potential interventions that could simultaneously promote digital inclusion and support environmentally sustainable behaviors among informal caregivers, a group that often faces considerable time limitations, financial pressures, and emotional demands. By empirically examining the relationship between the use of digital tools and eco-friendly practices in caregiving contexts, the study addresses a critical gap in the existing literature, which has largely considered digital and environmental transitions in isolation. In doing so, it contributes to a more integrated understanding of how the digital and environmental transitions intersect in everyday care and may inform policies that treat the 2 domains together rather than in isolation. Accordingly, this study addresses 2 key research questions:

1. What is the relationship between the use of digital tools and eco-friendly practices among informal caregivers of older adults with LTC needs?
2. Does the use of digital tools interlace differently with some eco-friendly practices rather than others?

Beyond estimating whether the use of digital tools is associated with environmentally sustainable behaviors, the analysis is intended to identify the behaviors for which digital interventions may be most relevant and those for which they are likely to be of limited use.

The focus on informal caregivers of older adults with LTC needs builds on previous research [68], which suggests that this population shows distinct patterns of environmental engagement compared to noncaregivers. These caregivers face unique challenges at the intersection of aging demographics, digital adoption barriers, and environmental concerns. As the population continues to age in Europe, understanding how they navigate both the digital and environmental transitions is increasingly relevant for the development of sustainable LTC policy.

Theoretical Approach

To examine how the use of digital tools relates to environmental practices among informal caregivers of older adults, this study uses the theory of planned behavior (TPB) [80] as its primary

XSL•FO
RenderX

analytical framework. TPB offers a well-established psychological model of the behavioral determinants, making it particularly suitable for exploring how digital tools might be associated with caregivers' environmental behaviors. According to this theory, behavioral intentions are determined by 3 key factors: attitudes (evaluations of the behavior), subjective norms (perceived social pressure), and perceived behavioral control (perceived ease or difficulty of performing a behavior).

The application of TPB is especially relevant for informal caregivers who face distinctive pressures and motivations compared with those of the general population. Caregivers often operate under significant time pressures, financial constraints, and emotional burdens that may influence their environmental practices. Through the TPB lens, we can examine how digital tools might affect each component in the caregiving context:

- Attitudes: digital tools may influence caregivers' evaluations of eco-friendly practices by providing information and educational content about environmental issues and sustainable practices [49,50]. Online resources can expose caregivers to environmental information that shapes their judgments about these practices
- Subjective norms: digital communities and social networks can strengthen environmental norms among caregivers [55,56], potentially normalizing eco-friendly behaviors or creating social pressure to meet expectations for environmental responsibility
- Perceived behavioral control: digital tools may enhance caregivers' perceived ability to engage in environmental behaviors by making sustainable options more accessible or convenient, for example, through e-commerce platforms [51-53]. This effect is likely to be more pronounced for behaviors that can be changed through information or convenience than for those dependent on external infrastructure

Building on recent extensions of the TPB [80,81], sociodemographic factors that may be associated with these relationships are also considered. Age, education, employment status, urbanization, and country of residence may be systematically related to both the adoption of digital tools and environmental behaviors. Previous research indicates that younger age [66], higher education, and urban residence [67] are associated with more positive attitudes toward digital technologies, potentially contributing to variation in the relationship between digital tools and environmental behaviors across different caregiver subgroups. This approach aligns with the SEIS conceptual framework by providing a psychological dimension for understanding how social (caregiving context), infrastructural (digital tools), and ecological (environmental behaviors) systems interact at the individual level. The analyzed eco-friendly behaviors include consumption choices (eco-friendly products, fair-trade items, second-hand goods, or plant-based options), resource conservation, waste management practices, and transportation decisions [82]. By examining how the use of digital tools relates to these different behavior categories, we can identify which TPB components may best account for the observed associations in the informal caregiving context.

Methods

Main Study Design and Participants

The findings outlined in this study are based on the Survey of Gender Gaps in Unpaid Care, Individual, and Social Activities project (European Institute for Gender Equality; EIGE/2021/OPER/10)—commissioned by the EIGE—conducted between November 2021 and December 2022 by Eurocarers (Belgium), in collaboration with the European Centre for Social Welfare Policy and Research (Austria), the National Institute of Health and Science on Ageing (Istituto di Ricovero e Cura a Carattere Scientifico [IRCCS] Istituto Nazionale Ricovero e Cura Anziani [INRCA] Italy), and IPSOS GmbH (Germany). In total, 60,405 individuals aged 16-74 years participated in the survey, which was carried out across all 27 EU Member States using quantitative methods. The fieldwork took place between August and October 2022.

Sampling and Data Collection

The primary data-collection instrument was a questionnaire designed for a 15-minute survey and structured to comply with the core principles of the General Data Protection Regulation (EU) 2018/1725 (GDPR). Data were collected by IPSOS GmbH using computer-assisted web interviews (CAWIs) across all participating countries except Malta and Luxembourg, where interviews were conducted via computer-assisted telephone interviews (CATIs). For the CAWI mode, the target population was limited to individuals who had registered with an online access panel, who were residents of one of the Member States, and who had access to online services on a digital device at the time of the survey. Panels are pools of registered respondents who are willing to participate in market and opinion research. To register, respondents must reside in the country of the panel they are signing up for. Their language skills must be sufficient to answer the registration questions and surveys in the local language. Robust online access panels do not exist in all EU Member States, particularly in smaller countries, where the overall size of such panels is limited. In Luxembourg and Malta, the available panels could not guarantee a sufficient number of respondents to reach the national sample sizes established on the sampling design. Therefore, in these 2 countries, respondents were interviewed via CATIs [68,83].

The sample size varied by country according to population size, ranging from 500 respondents for smaller countries (eg, Cyprus) to 4000 for more populous countries (eg, France and Germany). Hard quotas were applied for age, gender, and region (Nomenclature of Territorial Units for Statistics–Level 3 [NUTS III] classification), while monitoring quotas covered factors such as household composition, number of children, degree of urbanization, and education level (according to International Standard Classification of Education [ISCED] level of low, medium, and high) [68]. Monitoring education in particular helped account for differing levels of digital literacy among respondents, mitigating potential bias in the CAWI sample.

The survey targeted the general population of individuals aged 16-74 years. In certain countries (Austria, Belgium, Spain, Lithuania, Luxembourg, Latvia, Malta, and Poland), a limited number of respondents aged 75 years and older were included

XSL•FO
RenderX

in the survey to help meet quotas for the 65-74-year age group. In Romania, constraints related to low internet accessibility and limited digital literacy led to a narrower age range, restricting participation to individuals aged 16-64 years. The target quotas were based on official Eurostat statistics and reflected the population stratification within each Member State. The sampling approach for the CATIs differed from that for the CAWIs: no quotas were applied in Luxembourg or Malta, because the sampling was random, thereby reducing the potential for selection bias. Any discrepancies between the net sample and representative figures were subsequently adjusted through weighting. A random iterative proportional fitting procedure was used to weight multiple variables so that, after a predefined maximum number of iterations, the weighted distribution matched the target values of all selected variables [84].

Sample Selection: Exclusion Criteria

The total EIGE project sample numbered 60,405 participants. The exclusion criteria for this study were as follows:

- Participants who did not provide informal care: 47,209
- Participants whose main informal care recipient was younger than 65 years: 6273
- Participants with missing data on the frequency of sustainable behaviors: 1068
- Participants with missing data on the frequency of digital tool use in their caregiving roles: 94.

The resulting subsample selected for this paper, after the above exclusion criteria were applied, comprised 5761 individuals, who were included in the following analysis.

Outcomes and Dimensions

Outcome Variables

The outcome variable of the study was a score of environmentally sustainable behavior (SESB), calculated from participants' answers to 10 questions referring to environmentally friendly behaviors (see Multimedia Appendix 1 for the complete list of these 10 questions). For validity, internal consistency, and further details of the score, see Tur-Sinai et al [68]. To build the score, questions were coded from 1 ("Never") to 5 ("Daily") and summed to obtain a score ranging from 10 to 50 (higher scores indicated more frequent environmentally sustainable behaviors).

Covariates

The demographic covariates included gender and age. Socioeconomic variables included highest educational level according to the ISCED (2011; low=ISCED 0‐2; medium=ISCED 3‐4; high=ISCED 5+), employment status at the time of the survey (employed, unemployed, retired, or other nonworking situation), and perceived adequacy of household income (with great difficulty, with difficulty, with some difficulty, fairly easily, easily, or very easily). Weekly hours of informal care were assessed via 6-hour bands (1‐7 hours, 8-21 hours, 22‐35 hours, 36-39 hours, 50‐70 hours, and ≥71 hours). We also included information about the main form of residence (in your household, in another household, or in a care institution full-time or part-time) and the degree of urbanization of the place of residence (densely populated area, intermediate density area, and sparsely populated area) and country of residence. Finally, we included self-assessed general health (very good, good, fair, bad, or very bad).

Statistical Analysis

For descriptive statistics, the distribution normality of continuous variables was assessed using the Shapiro-Wilk test; means and SDs were reported. Statistical significance of between-group comparisons was assessed using one-way analysis of variance. For categorical variables, absolute frequencies and percentages were reported, and statistical significance was assessed using the Pearson chi-square test. The relationship between the frequency of use of digital tools in the caregiving role and the overall SESB score and each individual environmentally sustainable behavior was evaluated using Spearman rank correlation coefficients, with Bonferroni-adjusted *P* values.

For the multivariate analyses, a linear mixed effects model using maximum likelihood estimation was applied to the SESB score, while 10 mixed effects ordered probit models were estimated for the individual environmentally sustainable behaviors. Because of the hierarchical nature of the data, the statistical approach needed to account for the nesting of individuals within countries [85]. To address this, multilevel mixed effects models using maximum likelihood were employed, enabling the total variability to be partitioned into 2 levels: the primary level (individual-specific variability) and the secondary level (country-level variability) [84].

The validity of the models was evaluated using the Wald chi-square test. Additionally, to identify potential multicollinearity issues among covariates, the standard errors of each coefficient were assessed to mitigate variance inflation.

A 2-tailed *P* value of <.05 was considered statistically significant. All data analyses were conducted using STATA/MP (version 19.0; StataCorp LLC).

The survey microdata are publicly available on the GESIS public repository [86].

Ethical Considerations

The study adhered to the ethical principles outlined in the Declaration of Helsinki, and the survey protocol (including ethical aspects) received approval from IPSOS' Global Data Privacy and Research Compliance Team [87]. Some modifications to the questionnaire were required before approval, including the inclusion of explicit consent for sensitive data, the addition of refusal options, and the use of bracketed/grouped responses when referring to personal details about another adult individual (ie, the partner).

The survey was conducted in full compliance with national and EU data protection regulations, including the GDPR (Regulation [EU] 2016/679), Regulation [EU] 2018/1725, and Regulation [EC] No. 223/2009 regarding the processing of data for statistical aims. An integrated EU-27 dataset was generated from anonymized national datasets, ensuring that respondents could not be identified. A separate informed-consent clause was embedded in the survey's introduction, informing participants of their rights regarding data access, rectification, withdrawal,

XSL•FO
RenderX

and restriction, as well as their right to object to data processing and to file a complaint with the relevant data protection supervisory authority.

Respondents did not receive any compensation for participating in the survey. Informed consent was obtained from all participants, as the survey included questions related to political engagement and health. Respondents were provided with a dedicated section explaining the survey's purpose, content, and data storage policies. Only after giving their consent were they allowed to participate in the survey.

Results

Table 1 presents a bivariate analysis of the relationship between the frequency of use of digital tools for care-related activities and two sets of variables: (1) environmentally sustainable behaviors among informal caregivers—including recycling, avoiding single-use plastics, purchasing eco-friendly and fair-trade products, buying second-hand items, choosing plant-based foods, prioritizing local/seasonal food, mindful resource use, and low-carbon transportation—and (2) sociodemographic covariates. The analysis reveals a consistent positive association between the frequency of use of digital tools and environmentally sustainable behaviors. Specifically, caregivers who use digital tools daily have SESB scores 22% higher than nonusers, with a mean difference of 7 points. This indicates that a more frequent use of digital tools is associated with greater engagement in eco-friendly caregiving practices.

XSL•FO
RenderX

Table . Sample description according to the frequency of use of digital tools and of environmentally sustainable behaviors as a caregiver.

Frequency use of digital tools	Total (n=5761)	Never (n=1616)	Less often (n=1446)	Several times a month (n=864)	Several times a week (n=851)	Daily (n=984)	*P* value
Frequency of sustainable behavior							
SESB[a] (10-50), mean (SD)	33.9 (8.3)	31.4 (8.2)	32.9 (7.3)	33.3 (7.5)	36.0 (7.3)	38.4 (9.0)	<.001
Avoidance of plastic and/or disposable/single-use products, n(%)							<.001
Never	379 (6.6)	184 (11.4)	70 (4.8)	45 (5.2)	23 (2.7)	57 (5.8)	
Less often	1125 (19.5)	392 (24.3)	326 (22.5)	159 (18.4)	122 (14.3)	126 (12.8)	
Several times a month	1263 (21.9)	344 (21.3)	362 (25.0)	236 (27.3)	182 (21.4)	139 (14.1)	
Several times a week	1449 (25.2)	326 (20.2)	362 (25.0)	235 (27.2)	281 (33.0)	245 (24.9)	
Daily	1545 (26.8)	370 (22.9)	326 (22.5)	189 (21.9)	243 (28.6)	417 (42.4)	
Recycling, n (%)							<.001
Never	170 (3)	71 (4.4)	35 (2.4)	22 (2.5)	16 (1.9)	26 (2.6)	
Less often	545 (9.5)	165 (10.2)	142 (9.8)	104 (12.0)	63 (7.4)	71 (7.2)	
Several times a month	773 (13.4)	149 (9.2)	212 (14.7)	155 (17.9)	149 (17.5)	108 (11.0)	
Several times a week	1062 (18.4)	215 (13.3)	246 (17)	184 (21.3)	218 (25.6)	199 (20.2)	
Daily	3211 (55.7)	1016 (62.9)	811 (56.1)	399 (46.2)	405 (47.6)	580 (58.9)	
Avoidance of animal products, n (%)							<.001
Never	1493 (25.9)	622 (38.5)	396 (27.4)	175 (20.3)	132 (15.5)	168 (17.1)	
Less often	1488 (25.8)	424 (26.2)	485 (33.5)	231 (26.7)	177 (20.8)	171 (17.4)	
Several times a month	1056 (18.3)	237 (14.7)	266 (18.4)	202 (23.4)	207 (24.3)	144 (14.6)	
Several times a week	1015 (17.6)	197 (12.2)	189 (13.1)	184 (21.3)	229 (26.9)	216 (22.0)	
Daily	709 (12.3)	136 (8.4)	110 (7.6)	72 (8.3)	106 (12.5)	285 (29.0)	
Low-carbon-emission modes of transport, n (%)							<.001
Never	844 (14.7)	400 (24.8)	180 (12.4)	96 (11.1)	71 (8.3)	97 (9.9)	
Less often	1314 (22.8)	401 (24.8)	406 (28.1)	217 (25.1)	149 (17.5)	141 (14.3)	
Several times a month	1001 (17.4)	223 (13.8)	276 (19.1)	198 (22.9)	170 (20.0)	134 (13.6)	
Several times a week	1229 (21.3)	252 (15.6)	284 (19.6)	216 (25)	261 (30.7)	216 (22.0)	
Daily	1373 (23.8)	340 (21.0)	300 (20.7)	137 (15.9)	200 (23.5)	396 (40.2)	
Buy eco-friendly products and services, n (%)							<.001
Never	302 (5.2)	145 (9.0)	58 (4.0)	35 (4.1)	17 (2.0)	47 (4.8)	
Less often	1278 (22.2)	478 (29.6)	371 (25.7)	173 (20)	126 (14.8)	130 (13.2)	
Several times a month	1617 (28.1)	445 (27.5)	485 (33.5)	286 (33.1)	232 (27.3)	169 (17.2)	
Several times a week	1422 (24.7)	333 (20.6)	313 (21.6)	242 (28.0)	285 (33.5)	249 (25.3)	
Daily	1142 (19.8)	215 (13.3)	219 (15.1)	128 (14.8)	191 (22.4)	389 (39.5)	
Buy fair-trade products and services, n (%)							<.001
Never	496 (8.6)	259 (16.0)	93 (6.4)	56 (6.5)	31 (3.6)	57 (5.8)	

XSL•FO
RenderX

Frequency use of digital tools	Total (n=5761)	Never (n=1616)	Less often (n=1446)	Several times a month (n=864)	Several times a week (n=851)	Daily (n=984)	*P* value
Less often	1503 (26.1)	540 (33.4)	441 (30.5)	204 (23.6)	175 (20.6)	143 (14.5)	
Several times a month	1494 (25.9)	404 (25.0)	443 (30.6)	279 (32.3)	209 (24.6)	159 (16.2)	
Several times a week	1235 (21.4)	235 (14.5)	300 (20.7)	224 (25.9)	242 (28.4)	234 (23.8)	
Daily	1033 (17.9)	178 (11.0)	169 (11.7)	101 (11.7)	194 (22.8)	391 (39.7)	
Buy used instead of new items, n (%)							<.001
Never	1106 (19.2)	474 (29.3)	267 (18.5)	133 (15.4)	95 (11.2)	137 (13.9)	
Less often	2025 (35.2)	671 (41.5)	625 (43.2)	272 (31.5)	236 (27.7)	221 (22.5)	
Several times a month	1170 (20.3)	248 (15.3)	317 (21.9)	244 (28.2)	197 (23.1)	164 (16.7)	
Several times a week	754 (13.1)	104 (6.4)	133 (9.2)	154 (17.8)	205 (24.1)	158 (16.1)	
Daily	706 (12.3)	119 (7.4)	104 (7.2)	61 (7.1)	118 (13.9)	304 (30.9)	
Prioritize seasonal and local food products, n (%)							<.001
Never	199 (3.5)	89 (5.5)	38 (2.6)	25 (2.9)	19 (2.2)	28 (2.8)	
Less often	713 (12.4)	240 (14.9)	206 (14.2)	113 (13.1)	72 (8.5)	82 (8.3)	
Several times a month	1281 (22.2)	362 (22.4)	362 (25.0)	236 (27.3)	193 (22.7)	128 (13.0)	
Several times a week	1649 (28.6)	426 (26.4)	409 (28.3)	286 (33.1)	288 (33.8)	240 (24.4)	
Daily	1919 (33.3)	499 (30.9)	431 (29.8)	204 (23.6)	279 (32.8)	506 (51.4)	
Mindful of recyclability or biodegradability of products and packaging, n (%)							<.001
Never	453 (7.9)	231 (14.3)	95 (6.6)	50 (5.8)	28 (3.3)	49 (5)	
Less often	1046 (18.2)	361 (22.3)	304 (21)	159 (18.4)	113 (13.3)	109 (11.1)	
Several times a month	1160 (20.1)	305 (18.9)	332 (23)	205 (23.7)	169 (19.9)	149 (15.1)	
Several times a week	1416 (24.6)	306 (18.9)	351 (24.3)	255 (29.5)	296 (34.8)	208 (21.1)	
Daily	1686 (29.3)	413 (25.6)	364 (25.2)	195 (22.6)	245 (28.8)	469 (47.7)	
Mindful consumption of natural resources, n (%)							<.001
Never	159 (2.8)	72 (4.5)	28 (1.9)	26 (3.0)	15 (1.8)	18 (1.8)	
Less often	464 (8.1)	126 (7.8)	125 (8.6)	89 (10.3)	58 (6.8)	66 (6.7)	
Several times a month	812 (14.1)	182 (11.3)	219 (15.1)	168 (19.4)	137 (16.1)	106 (10.8)	
Several times a week	1130 (19.6)	222 (13.7)	308 (21.3)	195 (22.6)	229 (26.9)	176 (17.9)	
Daily	3196 (55.5)	1014 (62.7)	766 (53.0)	386 (44.7)	412 (48.4)	618 (62.8)	
Covariates							
Sex, n (%)							<.001
Female	3155 (54.8)	987 (61.1)	763 (52.8)	407 (47.1)	459 (53.9)	539 (54.8)	
Male	2591 (45.0)	624 (38.6)	681 (47.1)	457 (52.9)	388 (45.6)	441 (44.8)	
Other	15 (0.3)	5 (0.3)	2 (0.1)	0 (0)	4 (0.5)	4 (0.4)	
Age, mean (SD)	46.9 (14.0)	50.4 (13.2)	48.1 (13.7)	45.4 (14.4)	43.3 (13.8)	43.9 (14.0)	<.001
Education level, n (%)							<.001

XSL•FO
RenderX

Frequency use of digital tools	Total (n=5761)	Never (n=1616)	Less often (n=1446)	Several times a month (n=864)	Several times a week (n=851)	Daily (n=984)	*P* value
Low	599 (10.4)	227 (14.0)	116 (8.0)	78 (9.0)	79 (9.3)	99 (10.1)	
Medium	2376 (41.2)	676 (41.8)	615 (42.5)	344 (39.8)	332 (39)	409 (41.6)	
High	2786 (48.4)	713 (44.1)	715 (49.4)	442 (51.2)	440 (51.7)	476 (48.4)	
Degree of urbanization							.02
Densely populated area	2685 (46.6)	696 (43.1)	670 (46.3)	423 (49.0)	421 (49.5)	475 (48.3)	
Intermediate density area	1959 (34)	589 (36.4)	512 (35.4)	284 (32.9)	265 (31.1)	309 (31.4)	
Sparsely populated area	1117 (19.4)	331 (20.5)	264 (18.3)	157 (18.2)	165 (19.4)	200 (20.3)	
Perceived household income: household is able to make ends meet, n (%)							<.001
With great difficulty	407 (7.1)	117 (7.2)	71 (4.9)	50 (5.8)	56 (6.6)	113 (11.5)	
With difficulty	692 (12)	177 (11)	165 (11.4)	95 (11)	108 (12.7)	147 (14.9)	
With some difficulty	1719 (29.8)	437 (27)	461 (31.9)	265 (30.7)	270 (31.7)	286 (29.1)	
Fairly easily	1206 (20.9)	340 (21)	288 (19.9)	196 (22.7)	199 (23.4)	183 (18.6)	
Easily	628 (10.9)	167 (10.3)	167 (11.5)	106 (12.3)	87 (10.2)	101 (10.3)	
Very easily	272 (4.7)	68 (4.2)	76 (5.3)	43 (5.0)	39 (4.6)	46 (4.7)	
Prefer not to answer	837 (14.5)	310 (19.2)	218 (15.1)	109 (12.6)	92 (10.8)	108 (11)	
Current situation, n (%)							<.001
Employed	3827 (66.4)	976 (60.4)	968 (66.9)	604 (69.9)	594 (69.8)	685 (69.6)	
Unemployed	351 (6.1)	111 (6.9)	69 (4.8)	44 (5.1)	66 (7.8)	61 (6.2)	
Retired	755 (13.1)	271 (16.8)	222 (15.4)	99 (11.5)	56 (6.6)	107 (10.9)	
Other nonworking situation	828 (14.4)	258 (16.0)	187 (12.9)	117 (13.5)	135 (15.9)	131 (13.3)	
Weekly hours of informal care for main care recipient, mean (SD)	2.0 (1.3)	1.9 (1.2)	1.8 (1.1)	2.1 (1.2)	2.1 (1.2)	2.4 (1.5)	<.001
Residence of main recipient of informal care during period of caring, n (%)							<.001
In your household	1447 (25.1)	406 (25.1)	310 (21.4)	201 (23.3)	215 (25.3)	315 (32)	
In another household	3609 (62.6)	1022 (63.2)	973 (67.3)	552 (63.9)	524 (61.6)	538 (54.7)	
In a care institution	656 (11.4)	172 (10.6)	156 (10.8)	101 (11.7)	104 (12.2)	123 (12.5)	
Prefer not to answer	49 (0.9)	16 (1.0)	7 (0.5)	10 (1.2)	8 (0.9)	8 (0.8)	
General health, n (%)							<.001
Very good	958 (16.6)	219 (13.6)	218 (15.1)	138 (16.0)	162 (19.0)	221 (22.5)	
Good	2544 (44.2)	685 (42.4)	678 (46.9)	411 (47.6)	390 (45.8)	380 (38.6)	
Fair (neither good nor bad)	1814 (31.5)	558 (34.5)	450 (31.1)	261 (30.2)	244 (28.7)	301 (30.6)	
Bad	366 (6.4)	125 (7.7)	91 (6.3)	46 (5.3)	41 (4.8)	63 (6.4)	
Very bad	52 (0.9)	20 (1.2)	6 (0.4)	5 (0.6)	9 (1.1)	12 (1.2)	

XSL•FO
RenderX

Frequency use of digital tools	Total (n=5761)	Never (n=1616)	Less often (n=1446)	Several times a month (n=864)	Several times a week (n=851)	Daily (n=984)	*P* value
Prefer not to answer	27 (0.5)	9 (0.6)	3 (0.2)	3 (0.3)	5 (0.6)	7 (0.7)	
Country, n (%)							<.001
Ireland	85 (1.5)	27 (1.7)	17 (1.2)	12 (1.4)	12 (1.4)	17 (1.7)	
Austria	136 (2.4)	51 (3.2)	29 (2.0)	16 (1.9)	16 (1.9)	24 (2.4)	
Belgium	232 (4.0)	105 (6.5)	53 (3.7)	35 (4.1)	16 (1.9)	23 (2.3)	
Bulgaria	257 (4.5)	51 (3.2)	56 (3.9)	35 (4.1)	44 (5.2)	71 (7.2)	
Croatia	204 (3.5)	41 (2.5)	71 (4.9)	30 (3.5)	35 (4.1)	27 (2.7)	
Cyprus	37 (0.6)	11 (0.7)	14 (1.0)	2 (0.2)	6 (0.7)	4 (0.4)	
Czechia	209 (3.6)	84 (5.2)	54 (3.7)	30 (3.5)	23 (2.7)	18 (1.8)	
Denmark	103 (1.8)	20 (1.2)	18 (1.2)	20 (2.3)	23 (2.7)	22 (2.2)	
Estonia	43 (0.7)	11 (0.7)	18 (1.2)	2 (0.2)	6 (0.7)	6 (0.6)	
Finland	148 (2.6)	40 (2.5)	45 (3.1)	15 (1.7)	30 (3.5)	18 (1.8)	
France	247 (4.3)	89 (5.5)	77 (5.3)	29 (3.4)	28 (3.3)	24 (2.4)	
Germany	306 (5.3)	124 (7.7)	51 (3.5)	45 (5.2)	50 (5.9)	36 (3.7)	
Greece	438 (7.6)	90 (5.6)	112 (7.7)	69 (8.0)	83 (9.8)	84 (8.5)	
Hungary	172 (3.0)	33 (2.0)	36 (2.5)	36 (4.2)	26 (3.1)	41 (4.2)	
Italy	567 (9.8)	204 (12.6)	122 (8.4)	89 (10.3)	74 (8.7)	78 (7.9)	
Latvia	62 (1.1)	15 (0.9)	15 (1.0)	9 (1.0)	10 (1.2)	13 (1.3)	
Lithuania	129 (2.2)	29 (1.8)	36 (2.5)	23 (2.7)	15 (1.8)	26 (2.6)	
Netherlands	286 (5.0)	127 (7.9)	61 (4.2)	37 (4.3)	27 (3.2)	34 (3.5)	
Poland	406 (7.0)	80 (5.0)	108 (7.5)	64 (7.4)	65 (7.6)	89 (9.0)	
Portugal	402 (7.0)	100 (6.2)	108 (7.5)	62 (7.2)	58 (6.8)	74 (7.5)	
Romania	408 (7.1)	68 (4.2)	104 (7.2)	57 (6.6)	62 (7.3)	117 (11.9)	
Slovakia	70 (1.2)	19 (1.2)	30 (2.1)	8 (0.9)	8 (0.9)	5 (0.5)	
Slovenia	106 (1.8)	29 (1.8)	31 (2.1)	14 (1.6)	17 (2.0)	15 (1.5)	
Spain	509 (8.8)	110 (6.8)	117 (8.1)	95 (11.0)	93 (10.9)	94 (9.6)	
Sweden	142 (2.5)	33 (2.0)	49 (3.4)	20 (2.3)	21 (2.5)	19 (1.9)	
Luxembourg	24 (0.4)	6 (0.4)	9 (0.6)	5 (0.6)	2 (0.2)	2 (0.2)	
Malta	33 (0.6)	19 (1.2)	5 (0.3)	5 (0.6)	1 (0.1)	3 (0.3)	

[a]SESB: score of environmentally sustainable behavior.

Interestingly, while all 10 environmental behaviors included in the SESB are statistically significant, they show different patterns. For 8 behaviors, the proportion of caregivers engaging in daily eco-friendly practices is progressively higher from nonusers to daily digital users. Two behaviors, however, deviate from this pattern. For “recycling” and “mindful consumption of natural resources,” engagement does not rise steadily with the frequency of use of digital tools. It remains relatively stable across groups and, in the case of recycling, it declines among the heaviest users. In both cases, moderate digital users show the lowest engagement, producing a U-shaped distribution across the 5 frequency groups.

When considering each eco-friendly behavior separately, Table 1 shows the strongest relationship between the use of digital tools and consumption-related behaviors. Daily users of digital tools are almost 4 times more likely to buy fair-trade products daily (39.7%) than never-users (11%). Conversely, the percentage of caregivers who never buy fair-trade products falls from 16% among never-users of digital tools to 5.8% among daily users.

Daily users of digital tools are over 4 times more likely to purchase used items daily (30.9%) than never-users (7.4%). Conversely, never-users are more than twice as likely to never buy used items (29.3%) than daily users (13.9%).

Daily users of digital tools are 3 times more likely to purchase eco-friendly products daily (39.5%) compared to never-users (13.3%). Never-users are almost twice as likely to never buy eco-friendly products (9.0%) than daily users (4.8%).

XSL•FO
RenderX

Users of digital tools are over 3 times more likely to consume plant-based products (avoid animal products) daily than never-users. Conversely, never-users are more than twice as likely to never avoid animal products (38.5%) than daily users (17.1%).

In terms of resource conservation and waste management, daily users of digital tools are almost twice as likely to avoid single-use plastics daily (42.4%) than never-users (22.9%) and show much higher rates of daily prioritization of seasonal/local food (51.4%) than never-users (30.9%). Daily users of digital tools are much more likely to consider product recyclability daily (47.7%) than never-users (25.6%).

In terms of transportation decisions, daily users of digital tools are almost twice as likely to use low-carbon transportation daily (40.2%) than never-users (21.0%). Conversely, never-users are 2.5 times more likely to never use low-carbon transportation (24.8%) than daily users (9.9%). As for frequency of use of digital tools for care-related activities and sociodemographic covariates, the analysis suggests some specific patterns in the adoption of digital tools (Table 1). In terms of age, never users of digital tools are the oldest group (mean 50.4, SD 13.2 years), while daily users are younger (mean 43.9, SD 14.0 years). Higher education is also associated with more frequent use of digital tools. Daily users have higher rates of high education (48.4%) compared to never users (44.1%). Never users have higher rates of low education (14.0%) compared to daily users (10.1%). Daily users are much more likely to be employed (69.6%) than never users (60.4%), while retired individuals are significantly less likely to use digital tools on a daily basis (10.9%) compared to never users (16.8%). There is a slightly higher use of digital tools in densely populated areas (48.3% of daily users vs 43.1% of never-users). However, the difference in the adoption of digital tools is minimal in scarcely populated areas (20.3% vs 20.5%).

The peculiarity of recycling and mindful consumption of natural resources is further confirmed by the correlation analysis in Table 2. The frequency of use of digital tools shows a weak positive correlation with overall environmental behaviors (ρ=0.28; P<.001), with the largest associations observed for consumption-related behaviors including fair-trade products (ρ=0.30), used items (ρ=0.30), and eco-friendly products (ρ=0.25). In contrast, the weakest relationships are found for recycling (ρ=−0.04; P=.22) and mindful consumption of natural resources (ρ=−0.02; P>.99), neither of which shows a statistically significant correlation with the frequency of use of digital tools among informal caregivers.

Table . Spearman rank correlation coefficients with Bonferroni-adjusted P values between environmentally sustainable behaviors and frequency of use of digital tools for the role of carer.

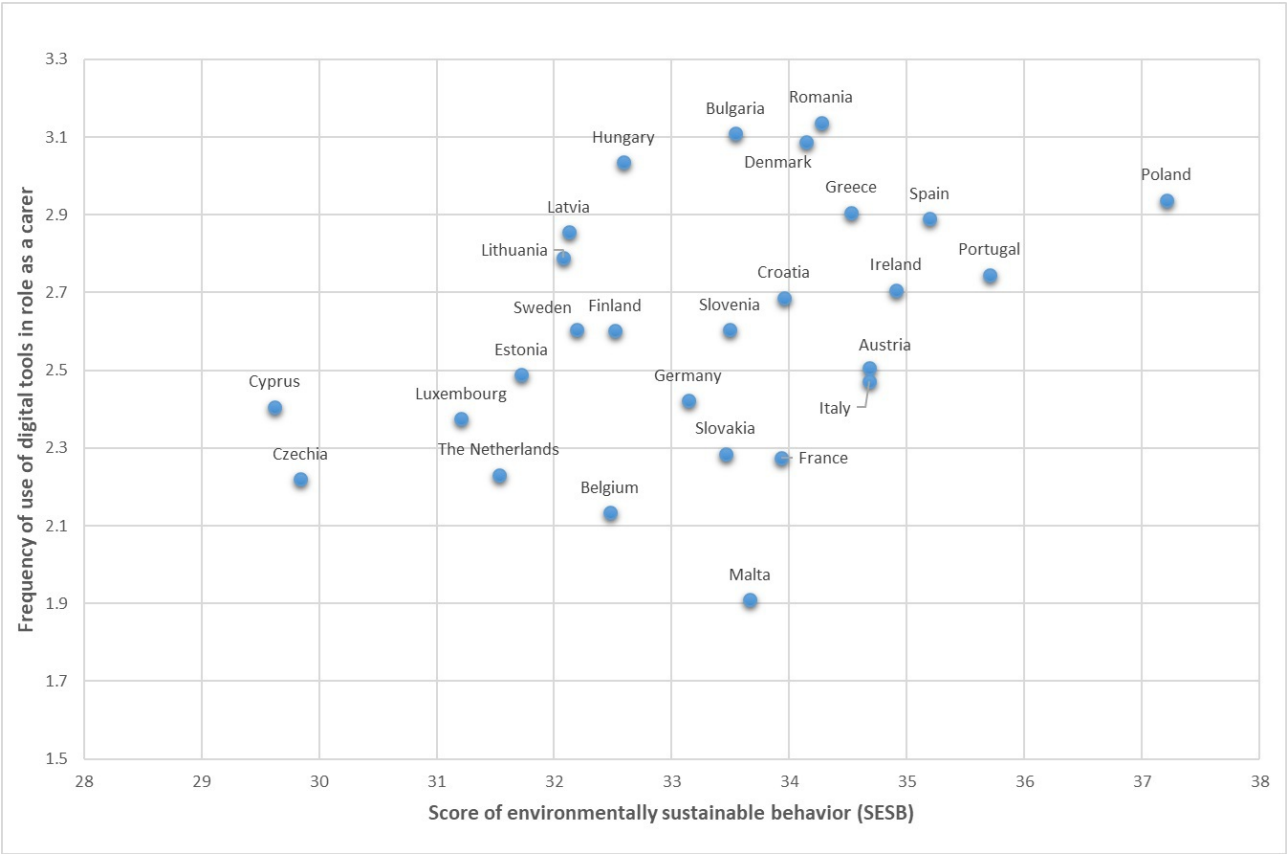
Frequency of sustainable behavior	ρ	P value
SESB[a]	0.28	<.001
Avoidance of plastic and/or disposable/single-use products	0.18	<.001
Recycling	–0.04	.23
Avoidance of animal products	0.27	<.001
Use of low-carbon-emission modes of transport	0.20	<.001
Purchase of eco-friendly products and services	0.25	<.001
Purchase of fair-trade products and services	0.30	<.001
Purchase of used (instead of new) items	0.30	<.001
Prioritization of seasonal and local food products	0.14	<.001
Mindful of recyclability or biodegradability of products and packaging	0.20	<.001
Mindful of consumption of natural resources	–0.02	>.99

[a]SESB: score of environmentally sustainable behavior.

Figure 2 shows how countries score on average for frequency of use of digital tools and of environmentally sustainable behaviors. Interestingly, countries where caregivers report higher scores on both dimensions are almost all located in Eastern Europe (eg, Poland, Romania, and Bulgaria) and Southern Europe (eg, Portugal and Spain). At the opposite end sit the Benelux countries (Belgium, Luxembourg, and the Netherlands), with some outliers such as Malta, Cyprus, and Czechia.

XSL•FO
RenderX

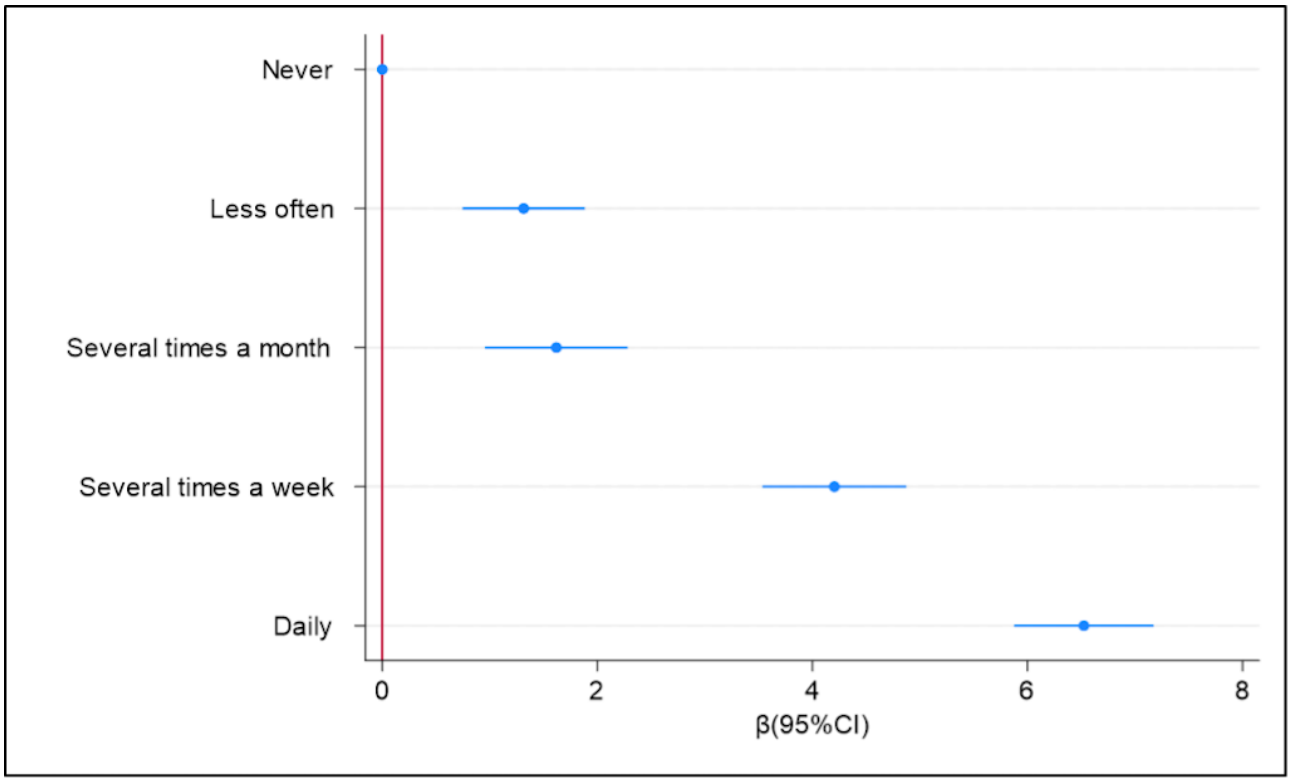
Figure 2. Frequency of use of digital tools and environmentally sustainable behaviors (in terms of score of environmentally sustainable behavior [SESB]) as a caregiver, by country.



The sociodemographic differences between users and never users raise the possibility of composition effects. To address this, a linear mixed effects model (maximum likelihood) was applied, fully adjusted for several potentially confounding variables: age, gender, educational level, weekly hours of informal care, degree of urbanization, perceived household income, working situation, residence of the care recipient, caregiver's general health, and country (as a random effect). This adjustment controls for factors that could be associated with both the use of digital tools and environmentally sustainable behaviors. Figure 3 displays SESB across different frequencies of use of digital tools, showing higher SESB scores at higher frequencies of digital use. This is consistent with the bivariate analyses presented above.

XSL•FO
RenderX

Figure 3. Forest plot: sustainable environmentally behavior score (SESB) according to the frequency of use of digital tools for the role of carer. Linear mixed effects (maximum likelihood) adjusted for age, gender, educational level, weekly hours of informal care for the main care recipient, degree of urbanization, perceived household income, working situation, residence of the main care recipient, caregiver's general health, and country (random effect parameter).



The mixed effects ordered probit analyses were repeated for each of the items included in the SESB and are synthesized in Table 3. Like Figure 3, the models are fully adjusted for the same set of confounding variables and include a random effect for the country. Except for 3 eco-friendly behaviors—recycling, mindful consumption of natural resources, and prioritizing seasonal and local food products (for which statistical significance is present only for some of the frequencies of digital use)—statistical significance holds across all levels of use of digital tools, indicating that more frequent use of digital tools is associated with more frequent sustainable behaviors. This pattern is consistent with the previously observed peculiarity concerning behaviors like recycling and mindful consumption of natural resources.

XSL•FO
RenderX

Table . Environmentally sustainable behaviors according to the frequency of use of digital tools for the role of carer[a].

Frequency of use of digital tools for role as carer	Avoidance of plastic and/or disposable/single-use products (Wald χ^2_{29}=314.00)		Recycling (Wald χ^2_{29}=187.39)		Avoidance of animal products (Wald χ^2_{29}=640.41)		Use of low-carbon-emission modes of transport (Wald χ^2_{29}=449.87)		Purchase of buy eco-friendly products and services (Wald χ^2_{29}=455.35)		Purchase of fair-trade products and services (Wald χ^2_{29}=588.23)		Purchase of used (instead of new) items (Wald χ^2_{29}=897.53)		Prioritize seasonal and local food products (Wald χ^2_{29}=300.84)		Mindful of recyclability or biodegradability of products and packaging (Wald χ^2_{29}=341.03)		Mindful of consumption of natural resources (Wald χ^2_{29}=277.62)	
	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value
Ref. Never																				
Less often	0.16 (0.04)	<.001	–0.03 (0.04)	.55	0.16 (0.04)	<.001	0.19 (0.04)	<.001	0.15 (0.04)	<.001	0.23 (0.04)	<.001	0.23 (0.04)	<.001	0.03 (0.04)	.52	0.14 (0.04)	<.001	–0.08 (0.04)	.07
Several times a month	0.20 (0.05)	<.001	–0.20 (0.05)	<.001	0.37 (0.05)	<.001	0.14 (0.05)	.002	0.21 (0.05)	<.001	0.29 (0.05)	<.001	0.38 (0.05)	<.001	−0.03 (0.05)	.52	0.19 (0.05)	<.001	–0.24 (0.05)	<.001
Several times a week	0.43 (0.05)	<.001	–0.08 (0.05)	.10	0.57 (0.05)	<.001	0.38 (0.05)	<.001	0.47 (0.05)	<.001	0.57 (0.05)	<.001	0.62 (0.05)	<.001	0.21 (0.05)	<.001	0.40 (0.05)	<.001	–0.06 (0.05)	.22
Daily	0.60 (0.05)	<.001	0.16 (0.05)	.002	0.80 (0.05)	<.001	0.62 (0.05)	<.001	0.71 (0.05)	<.001	0.84 (0.05)	<.001	0.82 (0.05)	<.001	0.52 (0.05)	<.001	0.66 (0.05)	<.001	0.20 (0.05)	<.001

[a]Mixed effects probit regression adjusted for age, gender, educational level, weekly hours of informal care for the main care recipient, degree of urbanization, perceived household income, working situation, residence of the main care recipient, and caregiver's general health, with county as a random effect.

From a practical perspective, this pattern suggests that a caregiver's level of digital engagement is more than a technical characteristic; it may reflect a broader capacity to find information, compare alternatives, manage daily tasks, and access services.

Discussion

Principal Findings

This study provides large-scale empirical evidence that the use of digital tools among informal caregivers is positively associated with environmentally sustainable behaviors, expanding current understanding of how digital and ecological transitions intersect in LTC.

The main finding of a weak positive association [88] between the frequency of digital tools use and environmental behaviors (ρ=0.28; *P*<.001) in caregiving contexts provides empirical evidence that demographic, digital, and environmental transitions are interconnected rather than occurring in parallel. Extending the work by Tur-Sinai et al [68] on caregivers' already heightened environmental engagement, our findings are consistent with the possibility that the use of digital tools is associated with caregivers' preexisting environmental responsibility, an interpretation aligned with the ethics of care framework [70,71].

The comparatively larger associations with consumption-related behaviors suggest that improved access to information may be one possible pathway linking the use of digital tools to these behaviors, by helping caregivers identify, compare, and locate sustainable options. Digital platforms offer ways to identify eco-friendly products, compare fair-trade alternatives, locate second-hand goods, arrange delivery, or participate in online exchanges of reusable care-related items; these are the behaviors for which the associations were the largest. By contrast, the weaker associations with behaviors that depend on infrastructure (recycling, mindful resource consumption, and low-carbon transport) suggest that digital solutions alone may be insufficient: a caregiver may recognize the importance of recycling, but, without accessible municipal services, clear local instructions, or convenient collection points, digital engagement has limited practical leverage. The findings thus point to a differentiated intervention model, in which digital tools facilitate informed and convenient choices but must be complemented by physical infrastructure and social support where environmental action depends on local systems.

XSL•FO
RenderX

Furthermore, these behaviors are already widely established: 55% of respondents report engaging daily in recycling and mindful resource consumption, compared with 12% - 33% for other environmental behaviors. This suggests that recycling and resource conservation are long-established habits, likely formed before digital tools became widespread in caregiving. In line with previous research [82], such routine environmental behaviors tend to be less responsive to digital engagement, as they are embedded in habit rather than informed choice. By contrast, consumption-related behaviors—such as buying secondhand or ecolabeled products—are more variable and may be more closely associated with informational and motivational inputs enabled through digital tools [48,51,53]. Consequently, these behaviors show larger associations with digital engagement.

These results support the distinction drawn in the literature between habitual and intentional pro-environmental actions [89], while extending it to the specific context of informal caregiving. In doing so, they offer both theoretical insights and practical guidance for integrated action across the domains of care, digital inclusion, and sustainability. Overall, the observed associations between the use of digital tools and eco-friendly behaviors may operate through attitudes, subjective norms, and perceived behavioral control, although these mechanisms were not directly tested. Digital access to information about sustainable products [48,49] may inform caregivers' evaluative judgments, particularly for consumption choices requiring informed decision-making. Online shopping may support perceived behavioral control by making sustainable choices more accessible [53], especially valuable for caregivers who cannot leave care recipients needing round-the-clock assistance. Online caregiver communities, where sustainable practices are shared and encouraged [56,62], may also relate to subjective norms about what is considered normal or expected. Regarding demographic patterns, consistent with established research on digital divides [66], the adoption of digital tools shows systematic variation across demographic groups, with higher rates among younger, more educated, and employed caregivers. Although the education gap is not substantial and the urban-rural divide is moderate, these disparities suggest that digital-environmental co-occurrence may be unevenly distributed. Older, less educated, and nonemployed caregivers—often those providing the most intensive care—appear less likely to use digital tools frequently. This may leave existing inequities unaddressed unless digital inclusion strategies specifically target these groups [66,67].

From a SEIS framework perspective [31], these findings demonstrate how the infrastructural system (digital tools) might shape behaviors within the ecological system through social system mediators (caregiver decision-making processes).

Policy and Practice Implications

First, policy frameworks addressing demographic, digital, and environmental transitions in LTC should recognize their interconnections rather than treating them as separate domains. Integrated approaches that simultaneously address caregiving needs, digital inclusion, and environmental sustainability could bring collateral benefits across multiple transition domains.

Second, to avoid exacerbating existing inequalities in the care sector [1,8], digital inclusion initiatives should specifically target more vulnerable caregiver groups or those facing greater barriers—particularly older, less educated, and nonemployed caregivers—to ensure that the benefits of digital-environmental synergies are equitably distributed across the caregiver population.

Third, environmental policies and LTC service providers could consider caregivers' existing predisposition toward environmental responsibility [68] when designing digital tools with dual functions, so that the same platforms supporting caregiving tasks could also provide access to sustainable consumption choices, particularly for those providing intensive care who may have limited mobility or greater time constraints. For example, existing caregiver support apps or portals could incorporate features such as mapping locally available second-hand or fair-trade suppliers and low-emission transport options for accompanying care recipients to appointments. Since the findings show a larger association with consumption-related behaviors, integrating sustainable-procurement guidance directly into the tools caregivers already use may be a relevant entry point.

Finally, educational programs for caregivers should integrate digital literacy with environmental awareness, helping caregivers understand how digital tools can facilitate both caregiving responsibilities and environmentally sustainable practices.

These recommendations must also take into account that family caregivers already perform demanding tasks that generate substantial waste. Expecting them to adopt greener practices without structural support risks adding further burden. Since caregivers already lean toward sustainability, the priority is not to demand more of them but to make sustainable action easier through the provision of efficient recyclable products and accessible disposal systems. Digital tools are most valuable here when they reduce, rather than add to, the effort of acting sustainably.

Overall, these results highlight the importance of considering all 3 SEIS dimensions when designing interventions to support the triple transition in LTC. Importantly, digital tools cannot substitute for physical infrastructure. Their associations are the largest for behaviors individuals can act on directly (such as purchasing choices) and weakest where action depends on external provision, such as recycling and resource conservation. The observed associations therefore suggest that the environmental relevance of digital investments may depend on whether enabling infrastructure already exists. Policies should combine digital inclusion measures with complementary local infrastructure, such as low-carbon transport or accessible recycling facilities, so that the options digital platforms recommend are available to caregivers. Digital innovation and physical infrastructure should be considered coordinated investments rather than alternatives.

Limitations

Although our study provides important insights into the intersection of the use of digital tools and environmental behaviors among informal caregivers, 2 primary limitations

XSL•FO
RenderX

should be acknowledged. First, the cross-sectional nature of our data does not allow us to establish causality or temporal ordering in the relationship between the use of digital tools and environmental behaviors. The mechanisms discussed in this paper should therefore be understood as theory-guided and empirically plausible pathways rather than as directly observed causal relationships. Nevertheless, the differentiated pattern across behavior categories offers useful indirect evidence: larger associations for information- and consumption-related behaviors and weaker associations for infrastructure-dependent behaviors are consistent with the proposed TPB and SEIS interpretation. Second, our measurement of digital tools captured frequency rather than the specific type, content, or function of the tools. As a result, the study cannot directly test whether caregivers used digital tools for environmental information, online purchasing, care coordination, peer support, or service navigation, which limits our understanding of which digital applications are most closely associated with eco-friendly behaviors. In light of these limitations, further mixed method studies are encouraged—combining longitudinal data with semistructured interviews—to clarify the temporal ordering and direction of the relationship between digital tools and environmental behaviors, to identify which types of tools caregivers use in relation to eco-friendly practices, and to examine the determinants and motivations underlying environmental attitudes. A further limitation derives from the CAWI recruitment procedure. Because the CAWI sample was drawn from online access panels, participants in these countries had baseline internet access and were registered with an online panel. Consequently, the "never users of digital tools for caregiving" group should not be interpreted as equivalent to noninternet users in the general population. The present dataset therefore cannot provide evidence about caregivers who were offline or otherwise excluded from online access panels, and claims concerning digital exclusion should be interpreted accordingly. Although the sample was weighted by country according to sex, age, region, and educational level, weighting cannot remove this underlying coverage limitation. Despite these limitations, this study provides empirical evidence of the associations between the use of digital tools and environmental behaviors among informal caregivers, contributing to our understanding of how the triple transition manifests in everyday caregiving contexts.

Conclusions

This study provides empirical evidence of an association between the use of digital tools and environmental practices among informal caregivers across the EU, suggesting that demographic, digital, and environmental transitions may co-occur rather than operate in isolation. However, because the design is cross-sectional, these findings should be read as associations rather than as evidence that the use of digital tools changes environmental behavior, and they may not generalize to caregivers without baseline internet access. Overall, the findings support integrated approaches that combine digital inclusion, caregiver support, and environmental infrastructure, while recognizing that the strength of the observed associations varies across behavior types and that recycling and resource conservation were not significantly correlated with the use of digital tools. However, these potential benefits might not be equally accessible to all caregivers, increasing the risk that digital tools may inadvertently widen existing inequalities in the care sector. Realizing the full potential of digital tools to support both caregiving and environmental goals will require inclusive access and appropriate infrastructure support. In practical terms, digitalization in LTC should be valued not only for improving care efficiency but also for its potential to support caregivers' environmental practices by reducing the burden of acting sustainably. This potential is behavior-specific: digital tools are most likely to help with information-, purchasing-, and coordination-related behaviors and least likely to help with practices constrained by physical infrastructure or established routines.

As European countries continue to face the triple transition, the evidence presented in this paper provides a foundation for developing more integrated approaches that combine sustainability-oriented digital platforms, caregiver training in both digital and environmental literacy, and the local infrastructure needed to translate digital information into feasible action.

Acknowledgments

The authors extend their gratitude to the consortium partners who contributed their expertise but are not listed as co-authors: Stecy Yghemonos (Eurocarers, Belgium); Cassandra Simmons and Eszter Zólyomi (European Center for Social Welfare Policy and Research, Austria); Robert Grimm, Carolina Bolz, Alexandra Schoen, and Armgard Zindler (IPSOS GmbH, Germany). Special thanks also go to the interviewers who collaborated to the survey.

Generative AI Use

The authors attest that no generative AI was used in any portion of the manuscript generation, including drafting, editing, revision, or preparation of the manuscript.

Funding

This paper is based on the survey of gender gaps in unpaid care, individual and social activities (CARE), commissioned by the European Institute for Gender Equality (EIGE), an agency of the European Union. Additionally, this work was also partially supported by Ricerca Corrente funding from the Italian Ministry of Health to IRCCS INRCA. The funding entities played no role in the study's design, data collection, analysis, or interpretation, nor in the writing of the manuscript or the decision to publish its findings. Consequently, the perspectives and conclusions expressed in this paper are solely those of the authors and do not

XSL•FO
RenderX

necessarily represent the views of EIGE, the European Union, or the Italian Ministry of Health. None of these institutions can be held accountable for the content.

Data Availability

The survey microdata is publicly available on the GESIS public repository [89].

Authors' Contributions

BI, MS, GL, MDR, AT-S, and RR conceptualized the study. MDR and MS curated the data and conducted the investigation. MDR, MS, and SQ performed the formal analysis and developed the methodology. GL, MS, and RR obtained funding. BI, GL, and MS provided project administration and supervision. MDR and MS created the visualizations. BI, AT-S, MDR, MS, and SQ drafted the original manuscript. All authors participated in writing review and editing, provided critical feedback on intellectual content, validated the results, and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary 1: ten questions list of the score of environmentally sustainable behavior (SESB) score.
[DOCX File, 36 KB - aging_v9i1e81556_app1.docx]

References

1. Towards a Triple Transition: Strategies for Transformational European Development Action. : OECD; 2023 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/10/towards-a-triple-transition_0e9ebd19/094322ba-en.pdf [accessed 2026-11-04] [doi: 10.1787/094322ba-en]
2. Eurostat. Population structure indicators at national level. : European Union URL: https://doi.org/10.2908/DEMO_PJANIND [accessed 2026-11-06] [doi: 10.2908/demo_pjanind]
3. Eurofound. Unpaid care in the EU. : Publications Office of the European Union; 2025 URL: https://assets.eurofound.europa.eu/f/279033/16470d0bab/ef25001en.pdf [doi: 10.2806/5774709]
4. Rocard E, Nozal A. Supporting informal carers of older people: policies to leave no carer behind. : OECD Health Working Papers; 2022 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/05/supporting-informal-carers-of-older-people_447892d0/0f0c0d52-en.pdf [accessed 2026-11-04]
5. Atasoy H, Chen PY, Ganju K. The spillover effects of health IT investments on regional health care costs. Manage Sci 2018 Jun;64(6):2515-2534. [doi: 10.1287/mnsc.2017.2750]
6. Fragão-Marques M, Ozben T. Digital transformation and sustainability in healthcare and clinical laboratories. Clin Chem Lab Med 2023 Mar 28;61(4):627-633. [doi: 10.1515/cclm-2022-1092] [Medline: 36473150]
7. Wanka A, Urbaniak A, Oswald F, Kolland F. Digital transformations in ageing societies. Z Gerontol Geriat 2023 May;56(3):177-180. [doi: 10.1007/s00391-023-02186-z]
8. Choudrie J, Zamani E, Obuekwe C. Bridging the digital divide in ethnic minority older adults: an organisational qualitative study. Inf Syst Front 2022 Aug;24(4):1355-1375. [doi: 10.1007/s10796-021-10126-8]
9. Watts N, Amann M, Arnell N, et al. The 2019 report of The Lancet Countdown on health and climate change: ensuring that the health of a child born today is not defined by a changing climate. The Lancet 2019 Nov;394(10211):1836-1878. [doi: 10.1016/S0140-6736(19)32596-6]
10. WHO. Operational framework for building climate resilient and environmentally sustainable health care facilities. : World Health Organization; 2020 URL: https://iris.who.int/server/api/core/bitstreams/9807d810-5279-4e26-886c-3c98e6b992d0/content [accessed 2026-11-04]
11. Braithwaite J, Ludlow K, Testa L, et al. Built to last? The sustainability of healthcare system improvements, programmes and interventions: a systematic integrative review. BMJ Open 2020 Jun 1;10(6):e036453. [doi: 10.1136/bmjopen-2019-036453] [Medline: 32487579]
12. Spasova S, Baeten R, Vanhercke B, European Commission: Directorate-General for Employment, Social Affairs and Inclusion. Challenges in long-term care in Europe. : Publications Office of the European Union; 2018. [doi: 10.2767/84573]
13. WHO. Global Strategy on Digital Health 2020-2025: World Health Organization; 2021.
14. Kruse CS, Mileski M, Vijaykumar AG, Viswanathan SV, Suskandla U, Chidambaram Y. Impact of electronic health records on long-term care facilities: systematic review. JMIR Med Inform 2017 Sep 29;5(3):e35. [doi: 10.2196/medinform.7958] [Medline: 28963091]
15. Lyons KS, Zarit SH, Sayer AG, Whitlatch CJ. Caregiving as a dyadic process: perspectives from caregiver and receiver. J Gerontol B Psychol Sci Soc Sci 2002 May;57(3):195-204. [doi: 10.1093/geronb/57.3.p195] [Medline: 11983730]
16. Wright J, Hamblin K. Technology and adult social care in England. In: Hamblin K, Lariviere M, editors. Care Technologies for Ageing Societies: Policy Press; 2023:18-48. [doi: 10.51952/9781447364825.ch002]

XSL•FO
RenderX

17. Yeung AWK, Torkamani A, Butte AJ, et al. The promise of digital healthcare technologies. Front Public Health 2023;11:1196596. [doi: 10.3389/fpubh.2023.1196596] [Medline: 37822534]
18. Hamblin KA. Technology in care systems: displacing, reshaping, reinstating or degrading roles? New Technol Work Employ 2022 Mar;37(1):41-58. [doi: 10.1111/ntwe.12229] [Medline: 35911255]
19. Moyle W, Jones C, Dwan T, Ownsworth T, Sung B. Using telepresence for social connection: views of older people with dementia, families, and health professionals from a mixed methods pilot study. Aging Ment Health 2019 Dec;23(12):1643-1650. [doi: 10.1080/13607863.2018.1509297] [Medline: 30450924]
20. Willard S, Cremers G, Man YP, van Rossum E, Spreeuwenberg M, de Witte L. Development and testing of an online community care platform for frail older adults in the Netherlands: a user-centred design. BMC Geriatr 2018 Apr 7;18(1):87. [doi: 10.1186/s12877-018-0774-7] [Medline: 29625562]
21. Pereno A, Eriksson D. A multi-stakeholder perspective on sustainable healthcare: from 2030 onwards. Futures 2020 Sep;122:102605. [doi: 10.1016/j.futures.2020.102605] [Medline: 32834076]
22. Muschol J, Heinrich M, Heiss C, et al. Economic and environmental impact of digital health app video consultations in follow-up care for patients in orthopedic and trauma surgery in Germany: randomized controlled trial. J Med Internet Res 2022 Nov 24;24(11):e42839. [doi: 10.2196/42839] [Medline: 36333935]
23. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
24. Castle-Clarke S. What will new technology mean for the NHS and its patients? : The Health Foundation, the Institute for Fiscal Studies, The King's Fund, and the Nuffield Trust; 2018 URL: https://assets.kingsfund.org.uk/f/256914/x/0158561602/nhs_70_what_will_new_technology_mean_2018.pdf [accessed 2026-11-04]
25. Mosca O, Manunza A, Manca S, Vivanet G, Fornara F. Digital technologies for behavioral change in sustainability domains: a systematic mapping review. Front Psychol 2023;14:1234349. [doi: 10.3389/fpsyg.2023.1234349] [Medline: 38239482]
26. Melkas H, Hennala L, Pekkarinen S, Kyrki V. Impacts of robot implementation on care personnel and clients in elderly-care institutions. Int J Med Inform 2020 Feb;134:104041. [doi: 10.1016/j.ijmedinf.2019.104041] [Medline: 31838291]
27. Tarpani RRZ, Gallego-Schmid A. Environmental impacts of a digital health and well-being service in elderly living schemes. Clean Environ Syst 2024 Mar;12:100161. [doi: 10.1016/j.cesys.2023.100161]
28. Alami H, Rivard L, Lehoux P, Ag Ahmed MA, Fortin JP, Fleet R. Integrating environmental considerations in digital health technology assessment and procurement: stakeholders' perspectives. Digit Health 2023;9:20552076231219113. [doi: 10.1177/20552076231219113] [Medline: 38074337]
29. Pekkarinen S, Melkas H, Hyypiä M. Elderly care and digital services: toward a sustainable sociotechnical transition. In: Toivonen M, Saari E, editors. Human-Centered Digitalization and Services: Springer Nature Singapore; 2019:259-284. [doi: 10.1007/978-981-13-7725-9_14]
30. EAN. Accelerating the digital transition: A stepping stone for sustainable elderly care in Europe. : European Ageing Network; 2024 URL: https://www.ean.care/media/fileman/2024_EAN_Accelerating_the_digital.pdf [accessed 2026-11-04]
31. Ramaswami A, Weible C, Main D, et al. A social - ecological - infrastructural systems framework for interdisciplinary study of sustainable city systems. J Industrial Ecology 2012 Dec;16(6):801-813. [doi: 10.1111/j.1530-9290.2012.00566.x]
32. Ostrom E. A general framework for analyzing sustainability of social-ecological systems. Science 2009 Jul 24;325(5939):419-422. [doi: 10.1126/science.1172133] [Medline: 19628857]
33. McPhearson T, Pickett STA, Grimm NB, et al. Advancing urban ecology toward a science of cities. Bioscience 2016 Mar 1;66(3):198-212. [doi: 10.1093/biosci/biw002]
34. Pickett STA, Cadenasso ML, Grove JM, et al. Urban ecological systems: scientific foundations and a decade of progress. J Environ Manage 2011 Mar;92(3):331-362. [doi: 10.1016/j.jenvman.2010.08.022] [Medline: 20965643]
35. Boone CG, Cook E, Hall SJ, et al. A comparative gradient approach as a tool for understanding and managing urban ecosystems. Urban Ecosyst 2012 Dec;15(4):795-807. [doi: 10.1007/s11252-012-0240-9]
36. Grimm NB, Cook EM, Hale RL, Iwaniec DM. A broader framing of ecosystem services in cities: benefits and challenges of built, natural, or hybrid system function. In: Seto KC, Solecki W, Griffith CA, editors. The Routledge Handbook of Urbanization and Global Environmental Change: Routledge; 2015:203-212. [doi: 10.4324/9781315849256-17]
37. Meerow S, Newell JP. Urban resilience for whom, what, when, where, and why? Urban Geogr 2019 Mar 16;40(3):309-329. [doi: 10.1080/02723638.2016.1206395]
38. Krueger EH, Constantino SM, Centeno MA, Elmqvist T, Weber EU, Levin SA. Governing sustainable transformations of urban social-ecological-technological systems. NPJ Urban Sustain 2022;2(1):10. [doi: 10.1038/s42949-022-00053-1]
39. Alshaikh RZ, Shok ME, Al-Hussaini ZIH, Alkinani AS. The impact of digital technologies on urban life quality and social dynamics in bismayah. Int J Sustain Dev Plan 2024 Sep 30;19(9):3405-3416. [doi: 10.18280/ijsdp.190911]
40. Purvis B, Mao Y, Robinson D. Three pillars of sustainability: in search of conceptual origins. Sustain Sci 2019 May;14(3):681-695. [doi: 10.1007/s11625-018-0627-5]
41. Pricopoaia O, Cristache N, Lupa c A, Iancu D. The implications of digital transformation and environmental innovation for sustainability. J Innov Knowl 2025 May;10(3):100713. [doi: 10.1016/j.jik.2025.100713]

XSL•FO
RenderX

42. Chopra R, Agrawal A, Sharma GD, Kallmuenzer A, Vasa L. Uncovering the organizational, environmental, and socio-economic sustainability of digitization: evidence from existing research. Rev Manag Sci 2024 Feb;18(2):685-709. [doi: 10.1007/s11846-023-00637-w]
43. Thøgersen J. Spillover processes in the development of a sustainable consumption pattern. J Econ Psychol 1999 Feb;20(1):53-81. [doi: 10.1016/S0167-4870(98)00043-9]
44. Hilty LM, Aebischer B. ICT for sustainability: an emerging research field. In: Hilty LM, Aebischer B, editors. ICT Innovations for Sustainability: Springer; 2015, Vol. 310:3-36. [doi: 10.1007/978-3-319-09228-7_1]
45. Degavre F, Kieffer S, Bol D, et al. Searching for sustainability in health systems: toward a multidisciplinary evaluation of mobile health innovations. Sustainability 2022;14(9):5286. [doi: 10.3390/su14095286]
46. Kwon C, Essayei L, Spencer M, et al. The environmental impacts of electronic medical rcords versus paper records at a large eye hospital in India: life cycle assessment study. J Med Internet Res 2024 Feb 6;26:e42140. [doi: 10.2196/42140] [Medline: 38319701]
47. Thompson M. The environmentally impacts of digital health. Digit Health 2021;7:20552076211033421. [doi: 10.1177/20552076211033421] [Medline: 34408902]
48. Santarius T, Pohl J, Lange S. Digitalization and the decoupling debate: can ICT help to reduce environmental impacts while the economy keeps growing? Sustainability 2020;12(18):7496. [doi: 10.3390/su12187496]
49. Johnson D, Horton E, Mulcahy R, Foth M. Gamification and serious games within the domain of domestic energy consumption: a systematic review. Renew Sust Energ Rev 2017 Jun;73:249-264. [doi: 10.1016/j.rser.2017.01.134]
50. Štreimikienė D, Baležentis A. Assessment of willingness to pay for renewables in Lithuanian households. Clean Techn Environ Policy 2015 Feb;17(2):515-531. [doi: 10.1007/s10098-014-0810-z]
51. Islam MS, Proma AM, Wohn C, et al. SEER: sustainable e-commerce with environmental-impact rating. Clean Environ Syst 2023 Mar;8:100104. [doi: 10.1016/j.cesys.2022.100104]
52. Senturk G, Dumludag D. The relationship between consumption of single-use plastic bags, environmental awareness, and socio-demographic factors. J Mater Cycles Waste Manag 2022 Jul;24(4):1494-1507. [doi: 10.1007/s10163-022-01407-8]
53. Eunike IJ, Silalahi ADK, Phuong DTT, Tedjakusuma AP. Towards a holistic model of green consumerism by integrating ethical obligations, collective efficacy, and environmental concerns in sustainable purchase behavior. Sustainable Futures 2025 Jun;9:100746. [doi: 10.1016/j.sftr.2025.100746]
54. Rana S, Shafi F, Rasheed A, Malik MI. Online environmental platforms service and green consumer behavior nexus: a multi-mediator study. Futur Bus J 2024 Dec;10(1):3. [doi: 10.1186/s43093-023-00283-4]
55. Geiger N, Swim JK, Glenna L. Spread the green word: a social community perspective into environmentally sustainable behavior. Environ Behav 2019 Jun;51(5):561-589. [doi: 10.1177/0013916518812925]
56. Hynes N, Wilson J. I do it, but don't tell anyone! Personal values, personal and social norms: Can social media play a role in changing pro-environmental behaviours? Technol Forecast Soc Change 2016 Oct;111:349-359. [doi: 10.1016/j.techfore.2016.06.034]
57. Kakulla B. 2023 tech trends: no end in sight for age 50+ market growth. AARP Research. 2023. URL: https://www.aarp.org/pri/topics/technology/internet-media-devices/2023-technology-trends-older-adults/ [accessed 2026-11-04]
58. Keenan TA. Using technology to assist with caregiving: caregivers are interested. AARP Research. 2022. URL: https://www.aarp.org/pri/topics/ltss/family-caregiving/caregiving-technology-survey/ [accessed 2026-11-04]
59. Hassan AYI. Challenges and recommendations for the deployment of information and communication technology solutions for informal caregivers: scoping review. JMIR Aging 2020 Jul 29;3(2):e20310. [doi: 10.2196/20310] [Medline: 32723720]
60. Soares S, Hoffmeister LV, Fernandes MDF, Henriques A, Costa A. The use of digital technologies in the promotion of health literacy and empowerment of informal caregivers: scoping review. JMIR Aging 2024 Apr 29;7:e54913. [doi: 10.2196/54913] [Medline: 38683655]
61. Chi NC, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. J Telemed Telecare 2015 Jan;21(1):37-44. [doi: 10.1177/1357633X14562734] [Medline: 25475220]
62. Wallcook S, Nygård L, Kottorp A, Malinowsky C. The use of everyday information communication technologies in the lives of older adults living with and without dementia in Sweden. Assist Technol 2021 Nov 2;33(6):333-340. [doi: 10.1080/10400435.2019.1644685] [Medline: 31343958]
63. Yin Z, Stratton L, Song Q, et al. Perceptions and utilization of online peer support among informal dementia caregivers: survey study. JMIR Aging 2024 May 31;7:e55169. [doi: 10.2196/55169] [Medline: 38825836]
64. Guay C, Auger C, Demers L, et al. Components and outcomes of internet-based interventions for caregivers of older adults: systematic review. J Med Internet Res 2017 Sep 19;19(9):e313. [doi: 10.2196/jmir.7896] [Medline: 28928109]
65. Czaja SJ, Perdomo D, Lee CC. The role of technology in supporting family caregivers. In: Human Aspects of IT for the Aged Population Design for Aging: Springer, Cham; 2016, Vol. 9754:178-185. [doi: 10.1007/978-3-319-39943-0_17]
66. Shaffer KM, Chow PI, Cohn WF, Ingersoll KS, Ritterband LM. Informal caregivers' use of internet-based health resources: an analysis of the health information national trends survey. JMIR Aging 2018 Dec 18;1(2):e11051. [doi: 10.2196/11051] [Medline: 31518244]
67. Schütz J. Who benefits from digital transformation in (health) care? Internet use among informal carers and care recipients in Europe. : Bayerisches Zentrum Pflege Digital; 2023. [doi: 10.60785/opus-1752]

XSL•FO
RenderX

68. Tur-Sinai A, Bentur N, Lamura G, Rodrigues R, Di Rosa M, Socci M. The sustainability spillover: uncovering the link between informal elder care and eco-conscious behaviors across the European Union. Innov Aging 2025;9(3):igae108. [doi: 10.1093/geroni/igae108] [Medline: 40061405]
69. Cousins EM. Mobilizing motherhood: the gendered burden of environmental protection. Sociol Compass 2021 May;15(5):e12879. [doi: 10.1111/soc4.12879] [Medline: 35813840]
70. Gilligan C. In a Different Voice: Psychological Theory and Women's Development: Harvard University Press; 1982.
71. Tronto J. Moral Boundaries: A Political Argument for an Ethic of Care: Routledge; 2020. [doi: 10.4324/9781003070672]
72. Xiao C, Hong D. Gender differences in environmental behaviors in China. Popul Environ 2010 Sep;32(1):88-104. [doi: 10.1007/s11111-010-0115-z]
73. Migheli M. Green purchasing: the effect of parenthood and gender. Environ Dev Sustain 2021 Jul;23(7):10576-10600. [doi: 10.1007/s10668-020-01073-6]
74. Pinquart M, Sörensen S. Correlates of physical health of informal caregivers: a meta-analysis. J Gerontol B Psychol Sci Soc Sci 2007 Mar;62(2):126-137. [doi: 10.1093/geronb/62.2.p126] [Medline: 17379673]
75. Schulz R, Beach SR, Matthews JT, Courtney K, De Vito Dabbs A, Mecca LP. Caregivers' willingness to pay for technologies to support caregiving. Gerontologist 2016 Oct;56(5):817-829. [doi: 10.1093/geront/gnv033]
76. Wagle S, Yang S, Osei EA, Katare B, Lalani N. Caregiving intensity, duration, and subjective financial well-being among rural informal caregivers of older adults with chronic illnesses or disabilities. Health Care (Don Mills) 2024;12(22):2260. [doi: 10.3390/healthcare12222260]
77. Spann A, Vicente J, Allard C, Hawley M, Spreeuwenberg M, de Witte L. Challenges of combining work and unpaid care, and solutions: a scoping review. Health Soc Care Community 2020 May;28(3):699-715. [doi: 10.1111/hsc.12912] [Medline: 31845451]
78. Lutz C, Hoffmann CP, Ranzini G. Data capitalism and the user: an exploration of privacy cynicism in Germany. New Media Soc 2020 Jul;22(7):1168-1187. [doi: 10.1177/1461444820912544]
79. Greenfield EA, Oberlink M, Scharlach AE, Neal MB, Stafford PB. Age-friendly community initiatives: conceptual issues and key questions. Gerontologist 2015 Apr;55(2):191-198. [doi: 10.1093/geront/gnv005] [Medline: 26035595]
80. de Leeuw A, Valois P, Ajzen I, Schmidt P. Using the theory of planned behavior to identify key beliefs underlying pro-environmental behavior in high-school students: Implications for educational interventions. J Environ Psychol 2015 Jun;42:128-138. [doi: 10.1016/j.jenvp.2015.03.005]
81. Armitage CJ, Conner M. Efficacy of the theory of planned behaviour: a meta-analytic review. Br J Soc Psychol 2001 Dec;40(Pt 4):471-499. [doi: 10.1348/014466601164939] [Medline: 11795063]
82. Stern PC. New environmental theories: toward a coherent theory of environmentally significant behavior. J Soc Issues 2000 Jan;56(3):407-424. [doi: 10.1111/0022-4537.00175]
83. EIGE. Online panel survey on gender gaps in unpaid care, individual and social activities - technical report. : Publications Office of the European Union; 2023. [doi: 10.2839/075582]
84. Austin PC, Goel V, van Walraven C. An introduction to multilevel regression models. Can J Public Health 2001;92(2):150-154. [doi: 10.1007/BF03404950] [Medline: 11338155]
85. Hox J, Balderjahn J, Mathar R, Schader M, editors. Multilevel modeling: when and why. In: Classification, Data Analysis, and Data Highways Studies in Classification, Data Analysis, and Knowledge Organization: Springer Berlin Heidelberg; 1998:147-154. [doi: 10.1007/978-3-642-72087-1_17]
86. European institute for gender equality (EIGE). survey of gender gaps in unpaid CARE, individual and social activities (CARE) (ZA8752; version 1.0.0). GESIS. 2023. URL: https://search.gesis.org/research_data/ZA8752?doi=10.4232/1.14218 [accessed 2026-10-26]
87. Global Privacy & Data Protection Policy. IPSOS. URL: https://www.ipsos.com/sites/default/files/Ipsos-Global-Privacy-and-Data-Protection-Policy.pdf [accessed 2026-09-23]
88. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Lawrence Erlbaum Associates; 1988.
89. Kurz T, Gardner B, Verplanken B, Abraham C. Habitual behaviors or patterns of practice? Explaining and changing repetitive climate ‐ relevant actions. WIREs Climate Change 2015 Jan;6(1):113-128. [doi: 10.1002/wcc.327]

Abbreviations

CATI: computer-assisted telephone interview
CAWI: computer-assisted web interview
EIGE: European Institute for Gender Equality
EU: European Union
GDPR: General Data Protection Regulation
INRCA: Istituto Nazionale Ricovero e Cura Anziani
IRCCS: Istituto di Ricovero e Cura a Carattere Scientifico
ISCED: International Standard Classification of Education
LTC: long-term care

XSL•FO
RenderX

NUTS III: Nomenclature of Territorial Units for Statistics–Level 3
OECD: Organisation for Economic Co-operation and Development
SEIS: social-ecological-infrastructural systems
SESB: score of environmentally sustainable behavior
TPB: theory of planned behavior

Edited by J Wang; submitted 04.Aug.2025; peer-reviewed by H Chen; revised version received 31.Aug.2026; accepted 06.Oct.2026; published 13.Nov.2026.

Please cite as:
Ietto B, Di Rosa M, Lamura G, Quattrini S, Rodrigues R, Santini S, Tur-Sinai A, Socci M
Use of Digital Tools and Environmental Practices Among Informal Caregivers of Older Adults in the European Union Embracing the Triple Transition: Cross-Sectional Study
JMIR Aging 2026;9:e81556
URL: https://aging.jmir.org/2026/1/e81556
doi:10.2196/81556

© Beatrice Ietto, Mirko Di Rosa, Giovanni Lamura, Sabrina Quattrini, Ricardo Rodrigues, Sara Santini, Aviad Tur-Sinai, Marco Socci. Originally published in JMIR Aging (https://aging.jmir.org), 13.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Public-Facing Communication of Health and Social Services for Older Adults and Their Family or Friend Caregivers: Environmental Scan of 58 Integrated Care Teams' Websites in Ontario, Canada

Hung Nguyen, BSc, MPH, PhD; Kyle Corbett, BA, MA; Deanna Vervaecke, BSc (Hons), MSc; Natalia Ivanovski, BA (Hons); Soroush Shirazi, BSc; Sara Yadollahi; Hom Lal Shrestha, BA (Hons), MA; Irene Slavescu; Behnoush Zarandi Baghini, BSc, MSc; Stella Medvedyuk, BA (Hons), MA, PhD; Matthias Hoben, RN, MSc, Dr rer medic

School of Health Policy and Management, Faculty of Health, York University, 4700 Keele Street, Toronto, ON, Canada

Corresponding Author:
Matthias Hoben, RN, MSc, Dr rer medic
School of Health Policy and Management, Faculty of Health, York University, 4700 Keele Street, Toronto, ON, Canada

Abstract

Background: Family or friend caregivers of older adults are critical in helping older adults navigate fragmented health and social systems, but they face significant challenges in doing so. Their needs for support, information, and resources are often unmet or remain largely invisible to health and social systems and public policy. In Ontario, Canada, Ontario Health Teams (OHTs) were established to integrate and streamline health care services. However, emerging evidence suggests that despite the requirement to integrate patient and caregiver advisors in these activities, caregivers still face substantial navigation barriers.

Objective: This study aimed to systematically evaluate the amount, nature, and accessibility of information provided on each of the 58 OHT websites. Specifically, we focused on information on services and supports for older adults and their caregivers.

Methods: Between November 2024 and May 2025, we conducted an environmental scan of all 58 OHT websites. Using a 5-point Likert scale, 2 team members independently rated how easy or difficult it was to identify services and supports for older adults and their caregivers. They also documented each service and support listed on each website and provided additional details on the experience of navigating the website in an open-text comment. The ratings were discussed in team meetings, and discrepancies were resolved through team consensus. Data analysis included thematic analysis of the services identified and of open-text responses (positive and negative experiences of navigating the websites, rationales for the ratings), as well as descriptive statistics of the ease of access ratings and of the types of services listed on OHT websites.

Results: Almost 60% of the websites were rated as difficult or very difficult to navigate, and 33% provided insufficient information on services and supports. However, information quality and accessibility varied significantly between websites. While some featured clear, well-organized resources, others were poorly designed, lacked a well-functioning search function, or provided vague or incomplete descriptions of services and supports. Design features that improved the accessibility and usefulness of websites included user-friendly, simple navigation and direct links to relevant services. In contrast, poorly designed websites often require multiple steps to access essential information, risking exacerbating caregiver burden.

Conclusions: Our findings highlight significant barriers for caregivers to access and navigate health and social service information, despite the intended goals of OHTs to improve system navigation. Health care system reforms focusing on integrated care need to include older adults and their caregivers as priority populations. Older adults and caregivers need to be engaged systematically and comprehensively, including in the development, design, and evaluation of health care system websites. Further, standards of public reporting need to be developed, and integrated care networks need to be required to follow these standards. This will help to improve transparency and accountability.

(JMIR Aging 2026;9:e80595) doi:10.2196/80595

KEYWORDS

Ontario Health Teams; integrated care; caregivers; older adults; health care system navigation; health care system access; health care system websites

Introduction

Family or friend caregivers (subsequently referred to as caregivers in this paper) play a critical role in the care of older adults [1-4]. Caregivers are relatives, friends, partners, neighbors, or other informal supports who provide care (usually unpaid) to someone who needs help due to physical, intellectual, or developmental disabilities; medical conditions; mental illness; or aging-related needs [1,5]. In 2022, 24.1 million caregivers

XSL•FO
RenderX

in the United States cared for an older adult [6], and around 6.4 million Canadians (approximately 20% of the Canadian adult population) were caregivers to dependent adults [7]. The number of caregivers in Europe was 76.2 million (approximately 10% of the total population) in 2016 [8]. About 80% of ongoing care for older adults in Europe [9], and 90% in the United States [10], is provided by caregivers, and in Canada, 95% of those receiving formal home care also have an unpaid caregiver [11]. The economic value of caregiving (ie, the costs of labor to the health care system if caregivers were to be replaced by formal care) was US $600 billion in the United States in 2021 [12], $9 billion CAD (US $6.6 billion) in Canada in 2019 (estimated to increase to $27 billion CAD [US $36.6 billion] by 2050) [13], and €576.4 billion (US $675.1 billion, 3.6% of the gross domestic product) in Europe in 2016 [8].

Despite their critical role, and despite the fact that caregivers provide large amounts of care to older adults that otherwise would have to be provided by the health care system, caregivers remain largely invisible to public policy—receiving no (or barely any) financial compensation for the labor they provide, and facing a dearth of publicly funded supports to meet their own needs [1-4]. In fact, caregiver supports have been eroding, rather than expanding, in recent years [14-16]. Much has been written about the inherent propensity of capitalist societies to marginalize unpaid care [17-20]. Individuals are conceived as independent, self-interested agents, disregarding issues of dependency. Relationships are seen as reciprocal and contractual, ignoring care-related asymmetries. Care is relegated to the private, domestic sphere, largely considered the responsibility of women, and assigned an idealistic, rather than monetary value. The consequence is a devaluation of care, of those who provide it (primarily women), and of those who receive it (often older adults with frailty and physical and cognitive disabilities) [17-20].

Therefore, unsurprisingly, comprehensive evidence suggests that caregiving increases a person's risk for poor mental health (eg, stress, depression, and anxiety), physical health (eg, pain, hypertension, and diabetes), social relationships (eg, reduced social contacts and conflicts with care recipient), finances (eg, out-of-pocket expenses and loss of income), and quality of life [1,21-25]. These negative impacts of caregiving grow more severe with increasing duration of the caregiver role with higher client care needs, especially if these include responsive behaviors, mental health issues, or sleep disorders [1,21,25-27]. In turn, the health and well-being of caregivers is closely associated with that of the person they care for [1,28-33].

One of caregivers' key roles in the care of older adults is navigating fragmented health and social systems, and they face considerable challenges doing so [34-36]. For example, caregivers often lack information about available services and how to access them. Available services may not be publicly funded, the older adult may not be eligible for publicly funded services, or these services may not be a good fit for the older adult's and caregiver's needs. Caregivers face long waiting times, complicated access procedures, communication gaps between settings and providers that require them to repeat the same information over and over again, and health care providers are poorly informed about the older adult's history or do not have time to listen to or share information with the caregiver [34-36].

When trying to find and navigate services and supports for those they care for or for themselves, the primary source of information for many caregivers is the internet [37]. Therefore, the quality and accessibility of information about these services and supports on the websites of health authorities providing access to these services and supports are a critical issue. In the Canadian province of Ontario, 58 Ontario Health Teams (OHTs) are responsible for delivering and organizing care [38]. In 2019, Ontario introduced a health care system reform, aiming to build "a connected health care system to improve the patient, family, and caregiver experience" [39]. This provincial health care system transformation was informed by international evidence of integrated care models implemented in other countries, such as HealthOne Mount Druitt in Australia, the Massachusetts General Care Management Program in the United States, and the Norrtälje Model in Sweden [40,41]. These integrated care models are similar to the reforms implemented in Ontario in their focus on care coordination and management, multidisciplinary team involvement, and user experience improvement. However, there are variations among OHTs in their geographic scale, governance structures, condition focus, and the specific sectors involved [40,41].

OHTs were established to integrate and connect care among hospitals, primary care practices, and home and community care [38] to "make it easier for [patients and caregivers] to navigate the health care system and transition between health care providers and places where [they] receive care, making sure [they] don't feel lost or unsupported" [39]. Ontario's Ministry of Health decided to leave it to regional networks of health care providers and organizations to connect, decide whether to become an OHT, and undergo a stepwise process to become an approved OHT [42]. The steps to become an OHT were to be repeated until full provincial coverage was achieved and included [42]: (1) a readiness self-assessment; (2) review of this self-assessment by Ontario's Ministry of Health with the possible outcomes of "in discovery" (beginning stages of readiness) or "in development" (higher degree of readiness); (3) postreadiness review activities: for teams in the "discovery" stage, access to supports and continued work toward the "in development" stage and for teams in the "in development" stage, preparing and submitting a full application; (4) review of the full application by Ontario's Ministry of Health with the possible outcomes of "ready for an in-person visit or not yet ready for an in-person visit"; (5) post full application activities: for teams "not yet ready for an in-person visit," continued work on their readiness until they are "ready for an in-person visit" and in-person visit for teams deemed "ready" for that step with the possible outcomes of "full readiness for implementation or still in development"; (6) categorization as "OHT candidates" of teams deemed "fully ready for implementation," who went on to implement the OHT model; and (7) progression toward full designation as an OHT.

By December 2019, the first cohort of 24 OHTs was announced [43], and by November 2022, 54 OHTs had been approved [44]. Full provincial coverage was achieved in January 2024, when the last of the 58 OHTs was approved [45].

XSL•FO
RenderX

One important requirement for OHTs is to ensure meaningful patient and caregiver engagement [46-49]. The government-funded Ontario Caregiver Organization [50], whose mandate is to support caregivers in Ontario [51], actively collaborates with OHTs to support them in these endeavors [52]. While evidence suggests that caregivers have been meaningfully engaged in the development and ongoing activities of some OHTs [53-56], reports from nongovernmental caregiver advocacy organizations point to substantial challenges that caregivers face in navigating health and social supports in Ontario. According to the Ontario Caregiver Coalition, 39% of caregivers experience system navigation challenges [57]. Over half of caregivers with unmet mental health needs [58] and 20% of caregivers with unmet respite needs [59] were unaware of available supports or were never offered supports. According to the Canadian Centre of Caregiving Excellence, 59% of caregivers in Canada spend time each week researching supports and navigating systems, and 50% of caregivers (59% among caregivers aged older than 65 years) report difficulties in finding information about caregiver supports [60]. A snapshot of the Ontario-specific findings suggests that caregivers in that province face similar navigation barriers as those in other provinces [61].

When our team accessed websites of various OHTs for other research projects, we noticed that the information on services and supports for older adults or their caregivers varied substantially between websites. OHTs are required to ensure information about the OHT's "service offerings is readily available and accessible to the public (eg, through a website)" [42]. A document provides guidance to OHTs on public-facing communications [62]. For example, OHTs are required to make "best efforts to include patient and provider benefits in all communications," and to "directly connect people to the right care at the right time and provide 24/7 help in navigating the health care system." Our impression was that the degree to which OHT websites met these requirements—especially related to navigating services and supports for older adults or their caregivers—varied considerably, which is why we decided to conduct this research study. The objective was to systematically evaluate the amount, nature, and accessibility of information provided on each of the 58 OHT websites, related to services and supports available to older adults or their caregivers.

Methods

Research Design

We conducted an environmental scan of the websites of all 58 OHTs [38]. An environmental scan is a structured and adaptable method for identifying and synthesizing publicly available information on existing programs and practices across organizations [63]. In health services research, environmental scans map the current landscape, compare jurisdictions, and identify gaps by integrating diverse sources, ranging from passive reviews of public web content (eg, organizational websites and gray literature) to active approaches such as key informant input [63,64]. Given our aim to appraise public-facing information, we conducted a web-based environmental scan of all 58 OHT websites, an approach consistent with prior OHT scans, to compare structure, accessibility, informational content, currency, and usability at scale. The primary purpose of this study was to evaluate OHT websites from an older adult and caregiver perspective. Therefore, we did not include key informant interviews with OHT representatives. Older adults and caregivers were part of our research team (see below).

Setting and Sample

The Context of Older Adult Care in Ontario

Our setting was ongoing older adult care in the Canadian province of Ontario. Our sample was the websites of all 58 OHTs. With 16.2 million residents (39% of the Canadian population) as of March 2025, Ontario is Canada's most populous province [65]. Over 18% of people in Ontario (2.9 million) are 65 years or older [66], and the estimated number of caregivers to people of all ages is 4 million [51]. Older adult care in Canada is primarily the responsibility of each of the 10 provinces and 3 territories (refer to Deber [67], chapter 4, and Marchildon et al [68], section 5.8 for the following). On a federal level, the Canada Health Act regulates the terms and conditions that provinces and territories must adhere to access federal funding for essential medical services (eg, hospital care, treatment by physicians, and diagnostics). However, federal transfer payments cover only 20% of provincial and territorial health budgets. Provinces and territories individually regulate and organize coverage and delivery of these services. Ongoing care and supports for older adults and their caregivers (eg, home care, respite services such as adult day programs, or long-term care homes) are not insured under the Canada Health Act, meaning they are completely at the discretion of each province and territory. In Ontario, some publicly funded services require out-of-pocket contributions, and those who do not qualify for publicly funded services will have to pay for services themselves [69,70].

Residential Care

The Ontario Ministry of Long-Term Care licenses and oversees 626 long-term care homes with a total of almost 80,000 beds [71,72]. Access is limited to residents with complex care needs who need 24/7 nursing care, and eligibility is assessed by Ontario Health atHome [73]. Services in licensed long-term care homes are publicly funded (with some copayments for board and meals), regardless of whether the organization applies a for-profit or a not-for-profit model. In addition, about 780 licensed, privately owned retirement homes (around 85,000 spaces) rent space and offer services such as meals or assistance with dressing to older adults with less complex care needs, who have the financial means to afford paying for these costs out of pocket (since these services are not publicly funded) [69,74]. Finally, Ontario provides access to a confusing array of publicly subsidized supportive housing options. Each of the 3 ministries (Ministry of Municipal Affairs and Housing; Ministry of Health; and Ministry of Children, Community, and Social Services) provides their own programs (for a total of around 20 programs) with varying target populations, eligibility criteria, access processes, services provided, and copayments required [75]. Target populations often include people with physical or cognitive disabilities, mental health issues, HIV, AIDS, or other conditions, and older adults may be eligible if they meet the

XSL•FO
RenderX

eligibility requirements of the respective program. In contrast to retirement homes, supportive housing settings are owned by not-for-profit entities (eg, municipalities, faith groups, and housing cooperatives), and the rent charged is often either based on a person's ability to pay or publicly subsidized [76].

Home and Community Care

Ontario Health atHome provides various home- and community-based services to well over 650,000 clients, such as care coordination, nursing, physiotherapy, occupational therapy, nutritional counseling, speech therapy, social work, personal support, medical supplies and equipment, meal delivery and dining programs, assistance with daily activities, transportation services, caregiver support, friendly visits, and palliative care [77,78]. For clients who live at home or in a retirement home, and whose care needs are more complex than can be met by home care, assisted living services are available to avoid admissions to long-term care homes or hospitals [79]. Almost 440 publicly funded adult day programs [80] provide respite to caregivers and social and recreational care to clients [81]. On 1 or a few days per week, clients attend these programs for part of the day, returning to their home at night (with some programs providing overnight care and clients returning to their home in the morning) [81]. Publicly funded respite care at home may be available, but availability in general and the funded number of hours vary depending on the client's and caregiver's region and unique situation [82]. Clients' short stays in a long-term care home are funded for up to 60 days at 1 time and a maximum of 90 days per year if caregivers need to take a break, go on vacation, stay in a hospital, or experience health issues preventing them from providing care [82,83]. Short-stay long-term care is also available for up to 90 days per year to older adults who are recovering from a medical condition before returning to their home (convalescent care), or for those waiting for a long-term care admission, but whose care needs cannot be met at home (interim care) [83]. Finally, there are some limited provincial and federal tax credits, employment or leave benefits, and various information and advocacy resources for caregivers [84]. In addition to implementing OHTs, Ontario's home and community care system has also undergone substantial reforms since 2020 [85]. OHTs are not fully integrated in the delivery of home and community care services, and only starting in 2025, an initial group of 12 OHTs was chosen by Ontario's Ministry of Health to deliver home and community services [38].

Measures, Outcomes, and Data Collection

Using a brief electronic information extraction sheet (File S1 in Multimedia Appendix 1), each of 2 team members independently reviewed OHT websites between November 2024 and May 2025. An update to the ratings of all OHT websites was conducted in May 2025 to ensure our analyses reflected any website changes since the initial assessment. Using 5 categories (very difficult, difficult, neither difficult nor easy, easy, and very easy), team members rated, for each website, how easy or difficult they found it to identify services available to older adults or their caregivers. In addition, all services identified were documented, including the type of service, name of the service provider, the weblink, and comments. Additional comments about the website were captured in a final open-text response.

Raters used definitions of each category for their ratings (Table 1), which we developed based on the Web Content Accessibility Guidelines 2.2, developed and published by the World Wide Web Consortium as part of the Web Accessibility Initiative [86]. Team members involved in developing the definitions of each of the 5 rating options were experts in the areas of scale development and validation, and care and support of older adults and their caregivers. Three of our team members have lived experience as caregivers to an older adult, and 1 team member is an older adult with mild cognitive impairment. These team members were involved in the development of rating categories and definitions, website ratings, and reconciliations, and they are coauthors on this paper. We acknowledge the potential of such a setup to introduce bias to the evaluation. On the one hand, older adults and caregivers bring a very specific lived experience that will differ from that of other team members. On the other hand, senior team members who are experienced researchers are at risk of overruling the voices of team members with lived experience. To ensure a diversity of perspectives, we included team members of diverse backgrounds, including variations in ethnocultural background, gender, lived experience, expertise, and seniority. To ensure a safe space and well-rounded discussions, the team agreed on ground rules, such as not interrupting another team member, expressing disagreements in a respectful tone, and if in doubt, giving priority to the perspectives of persons with lived experience. Older adults and caregivers are primary audiences for OHT websites, and their feedback on navigation, accessibility, and usefulness is needed for assessing how public-facing platforms function in practice.

XSL•FO
RenderX

Table . Definitions of Likert scale categories to assess the level of ease or difficulty in identifying information on websites related to services and supports for older adults and caregivers.

Likert scale category	Definition
Very easy	The website is clearly and simply designed, with easily accessible links or dropdown menus directing users to resources relevant to older adults and caregivers. The process for locating nearby services involves straightforward and intuitive steps. The language used throughout the site is clear and easy to comprehend.
Easy	The website is clearly and simply designed, featuring accessible links and dropdown menus for resources relevant to older adults and caregivers. Navigation is generally intuitive, and the steps to locate services are easy to follow. However, users may need to make several clicks before accessing the desired resources. Additionally, the presence of unrelated services may cause some confusion for users specifically seeking support for older adults or caregivers.
Somewhat easy or difficult	While some services are accessible, they are often dispersed across unrelated sections, and the website lacks specific categories for dementia care or for family and friend caregivers. Despite a visually appealing layout, the depth of information provided is limited—often restricted to a phone number or a service provider's name—offering little practical guidance.
Difficult	The website is poorly organized or lacks a coherent structure. Specific categories or groupings such as "older adults" are absent. Although a general resources or support section may exist, it includes minimal content focused on older adults, making it challenging for users to locate relevant information or services.
Very difficult	The website does not offer a dedicated resources page and may simply redirect users to external services. In some cases, users are required to submit extensive personal information to access service details, only to discover that the results may still be irrelevant to their needs.

Team members' expertise and use of an established set of criteria (theory) were measures to ensure the content validity of these categories [87,88]. Using an open-text response option, raters explained their rating in more detail. These open-text responses were used to assess the response process validity [87,88]—that is, the raters' cognitive processes when making the ratings and whether their responses reflected the intended meaning (Table 1) of the respective category. After each dyad had assessed their first website, we met as a team, compared the 2 ratings, discussed and clarified our categories and definitions if needed, discussed discrepancies, and reviewed the website in question as a team until a consensus was reached. Team members then rated the next set of websites. The process continued until all websites were assessed. Because the intended use of the rating categories was purely descriptive and qualitative, and we did not include the ratings in any statistical models, establishing construct validity (ie, association of the category ratings with other outcomes as hypothesized) [87,88] was neither possible nor applicable in the context of this study. Similarly, an assessment of the internal structure of the categories was not conducted because the categories reflected a single item only (ease of navigating the website) [87,88]. Using the prereconciliation ratings (File S2 in Multimedia Appendix 2), we calculated the interrater reliability. The overall weighted kappa value [89] was 0.80 (95% CI 0.69-0.90), indicating strong agreement [90].

Analyses

We reported frequency distributions of difficulty ratings and types of services, as well as the mean (SD) and median (IQR) number of services listed across the websites examined. Two team members independently conducted thematic analysis [91,92] of open-text responses by identifying units of meaning (eg, a phrase, sentence, or paragraph) and assigning a short label to each common theme. The 2 team members compared coding trees and interpretations. Discrepancies were resolved by consensus.

Results

Three (5%) of the OHTs did not have a website. Accessing information was very difficult on 4 OHT websites (7%), difficult on 27 (46%), somewhat easy or difficult on 17 (29%), easy on 4 (7%), and very easy on 3 (5%). File S2 in Multimedia Appendix 2 lists the difficulty rating for each OHT.

Table 2 presents the different types of services listed on the various OHT websites and the number of websites explicitly listing the respective service. Three OHTs had no website, and no services were listed on the websites of 2 OHTs. The websites of each of the other 53 OHTs listed phone numbers or online forms for people to request more information on available services, or provided weblinks to external resources listing services. However, 3 of these 53 websites (6%) provided only this information and listed no other services, 5 (9%) listed only 1 other service, and 6 (11%) listed only 2 other services. Therefore, one-third of the websites (n=19, 33%) were deemed to provide no or insufficient information on services available to older adults or their caregivers, as additional steps were needed to access the required information.

XSL•FO
RenderX

Table . Types of services and supports for older adults and their caregivers listed on Ontario Health Team (OHT) websites and the number (percentage) of websites listing each service or support.

Service or support	Websites, n (%)
Phone numbers, web links, or online forms	53 (91)
Other services or supports	40 (69)
Mental health services	39 (67)
Medical treatments	31 (53)
Caregiver supports	29 (50)
Supports for persons with dementia	22 (38)
Home care	21 (36)
Housing	9 (16)
Meal services	9 (16)
Alzheimer Society supports	7 (12)
Transportation	7 (12)
Adult day programs	3 (5)
No website	3 (5)
No services listed on website	2 (3)

Characteristics of websites that were easy to navigate and that provided easy access to information were described by raters' comments as having "clear buttons or functions for the search." Searches identified a variety of "expected services" for older adults or their caregivers. Some websites even provided information beyond the raters' expectations, such as resources for "self-advocacy, community supports, or handbooks." Simplicity of the online contact form was highlighted as a positive website feature.

In contrast, characteristics that posed challenges in navigating websites and identifying information largely referred to search functions. Often, no search function was available at all. Search functions sometimes did not work properly (eg, dead links or dysfunctional buttons), or search buttons and functions were poorly worded, potentially causing confusion and frustration for users. Raters also commented on websites that required users to follow several steps to identify the expected information. Frequently, the overall website design was cluttered and confusing, making navigation challenging. Websites regularly provided very limited information on the services available to older adults and caregivers, either listing only very few services (if any), poorly describing the service (eg, how to access it, eligibility criteria, and what exactly is offered), or only providing links to external resources for additional information.

Discussion

Our study systematically evaluated the websites of all 58 OHTs, assessing the amount, nature, and accessibility of information on services and supports for older adults or their caregivers. While we found promising best practice examples, the majority of the websites were difficult to navigate, had design issues, or contents were either not available, not helpful, or difficult to access. Overall, website designs and contents varied widely between OHTs.

Numerous studies have identified similar challenges of public health website design and accessibility [93-96]. These problems are exacerbated for caregivers, who often lack the technical skills, time, or capacity to locate the required information [97]. Caregivers for whom English is not the first language, or who have lower levels of education or health literacy, may experience particularly high barriers [98]. Well-designed websites that provide multilingual and culturally safe content and are easily accessible are critical to mitigating these barriers [99]. Otherwise, these navigation challenges may further amplify caregiver burden and distress and negatively affect their mental health [100].

Given that OHTs are required to include caregivers as advisors [46-49], our findings raise the question of why, for many OHTs, caregiver involvement has not resulted in more caregiver-friendly website designs. It is possible that caregivers may not have been involved in the particular activity of website development and that caregiver input has not been obtained on the usefulness and accessibility of these websites. Researchers have pointed out that clear guidance on principles of caregiver involvement is lacking and that the best ways to evaluate the quality of caregiver involvement constitute a critical knowledge gap [101,102]. In fact, research has reported challenges among OHTs in meaningfully engaging caregivers [54]. It is possible that, despite the guidance documents on OHT's public-facing communication [42,62], OHTs may not have identified or prioritized providing information on services and supports on their websites, but rather focused on sharing information on the OHT's development. Also, despite the fact that all 58 OHTs were approved and declared fully functioning by Ontario's Ministry of Health, a fully functioning and searchable website may not have been a major criterion in the approval process, and some of the newer OHTs may not yet have developed such a website. Regardless, the consequences for an older adult or caregiver encountering such a situation are that they will be

XSL•FO
RenderX

unable to locate the required information and will need to identify alternative resources, adding to their burden.

Other reasons for less-than-ideal website design and accessibility may be systemic barriers. Health care systems may lack the technical skills and resources (ie, funding, expertise, and infrastructure) to properly design their websites, may have a poor understanding of their users' design and navigation needs, and the exact audience and goals targeted by the websites may be poorly defined [103,104]. However, best practice guidance for health care systems to support website design is lacking, and there is a critical knowledge gap related to health care system representatives' experiences and challenges in designing these websites [103,104].

In addition to these technological and resource-related challenges, there is reason to critically question whether older adults and their caregivers were given sufficient priority in the process of establishing the OHTs. For example, OHTs were required to select a small number of target populations on which they would initially focus [42]. However, older adults and their caregivers were not mentioned as examples of potential priority populations in the guidance for OHTs [42] or in the full application [105]. As a result, many OHTs may not have selected older adults and their caregivers as priority populations, and we found very limited publicly available documentation on OHTs' selected target populations. Fourteen OHTs explicitly mentioned that older adults were 1 of their priority populations, and 11 OHTs mentioned caregivers. More importantly, while home and community care were always listed, along with other care settings, as part of the care continuum to be integrated by OHTs [38,43], it was not until early 2025 that 12 of the 58 OHTs were selected to integrate the delivery of home and community care more systematically [38]. Not prioritizing the needs of older adults and their caregivers early on, but deferring them until 6 years into the reform process, represents a clear policy decision, suggesting that older adults and their caregivers may not have been at the top of the government's priority list. Systemic marginalization of older adults [106], especially those with dementia [107], and their caregivers [17-20] has been extensively documented in the international literature, including in Ontario [108-110].

We would like to highlight and acknowledge the tremendous work OHTs have done in working toward a more integrated health care system. OHTs are part of a system, and the system's context conditions substantially influence OHTs' possibilities and limits to act. Since the 1980s, every Canadian government has followed international trends of austerity measures [20,111,112], defunding and privatizing public services (including health care services), and dismantling social safety nets [113,114]. The Canadian government has constantly reduced transfer payments to provinces, and provinces, in turn, have implemented similar austerity measures [113,114]. Ontario is no exception [108]. OHTs have struggled with engaging care practitioners who already face heavy workloads [115] and with a lack of external supports and resources [116]. In line with concerns expressed by others [117-119], we strongly believe that without addressing the root causes of health care system issues (predominantly underfunding, privatization, workforce shortages, and burnout), this reform, like others before it, may fall short of its goals. Successful examples of integrated care reforms in Europe highlight the importance of legislation for supporting governance, financial support, and policy support at all levels, from the planning stages to final implementation [120,121].

Despite these struggles, we found various encouraging best practice examples of providing clear, comprehensive, and easily accessible information on services for older adults and their caregivers. Implementing systematic mechanisms for OHTs to connect with one another and to learn from each other's successes and challenges would be an important improvement.

These factors may all contribute to the fact that various OHT websites have seen little (if any) updates since the launch of these teams in 2019 [122], but studies examining the exact reasons for this are lacking. Concerns have been raised about the accountability and public transparency of OHTs [122]. While accountability has improved, a standard method for public reporting has yet to be implemented [122]. Greater accountability and transparency across OHTs could be achieved through a standardized approach to public reporting [122], for example, by leveraging data reporting and sharing by the Canadian Institute for Health Information [123] to provide OHT-level reports on hospital, long-term care, and home and community care performance.

Our study includes several key recommendations for governments or health authorities seeking to create accessible websites to help older adults and their caregivers navigate available services and supports. Providing integrated care networks with sufficient resources, supports, and expertise will be critical in light of population aging, increasingly complex care needs, a declining number of caregivers, and increasing caregiver burden. Clear policies are required that explicitly specify criteria for accessibility and ease of navigation that websites must meet, as well as required contents that websites need to list (eg, details on services and supports, including eligibility criteria, ways to access, funding mechanisms, and so on). Mechanisms of monitoring the availability and quality of websites need to be implemented, along with measures to hold integrated care networks accountable to these criteria. Finally, older adults and caregivers should systematically be integrated into defining criteria for website accessibility and ease of navigation and in assessing whether websites meet these criteria. More research is urgently needed to understand older adults' and caregivers' experiences with accessing and navigating services in more detail, including their interactions with digital platforms. Improving the accessibility and clarity of information should be seen as part of a larger effort to make the health and social care system more navigable, equitable, and responsive to the needs of caregivers and of those they care for.

Major strengths of our study include (1) its focus on an underresearched yet important topic—the usefulness and accessibility of health care system websites from the perspective of caregivers; (2) the comparison of a wide range of health care system websites (those of all 58 OHTs) in the largest Canadian province; (3) our rigorous approach to evaluating these websites (criteria based on best practice standards for website accessibility, independent rating by 2 raters, and reconciliation

XSL•FO
RenderX

of discrepancies); and (4) the inclusion of persons with lived experience (older adults with cognitive impairment and caregivers) in our research team. However, several limitations should be noted. Involving a larger group of older adults and caregivers would have provided more comprehensive and more generalizable insights into their experiences navigating the websites. Despite our approach of independent ratings and subsequent reconciliation, website characteristics remain subjective to some extent, further limiting the generalizability of our findings. While giving persons with lived experience a strong voice was important, we acknowledge that their ratings may differ from those of researchers without such experience. However, the purpose of our study was to evaluate OHT websites from the perspective of potential users. On the other hand, our team members with lived experience are probably not reflective of the typical older adult or caregiver using OHT websites. Our team members (including those with lived experience) were students and study staff with high health literacy. Therefore, the issues we identified with OHT websites are likely underestimated compared with those that a more typical user with lived experience would encounter. We intentionally focused on older adults' and caregivers' perceptions when navigating OHT websites and did not include key informant interviews with OHT representatives. These key informants might have provided important context on OHTs' priorities and processes for website design. However, identifying representatives from each of the 58 OHTs would have been immensely time-consuming, and many of them might not have agreed to participate. This is an important area for future research, but the required efforts would not have been justified given this study's different primary focus. While websites are a critical source of information for caregivers, they may not be the only source. Some caregivers may prefer to call centralized information lines, where the quality of support may vary widely depending on the individual assisting them. Future research should examine caregivers' experiences with such helplines. Finally, websites are dynamic and subject to frequent updates. Some of the websites evaluated in this study may since have improved their accessibility and navigability.

Acknowledgments

This study would not have been possible without the invaluable support of Andrea Ubell, Director of Programs and Client Services at the Alzheimer Society of York Region, whose guidance and introduction to the 58 OHT websites were instrumental to this study's success.

Funding

This study was funded by the Helen Carswell Research Chair in Dementia Care.

Data Availability

All data used to create the results of this study are either publicly available (eg, the OHT websites), presented in the main manuscript, or supplied in supporting files available for download along with the published manuscript (eg, the website ratings or rater comments).

Authors' Contributions

MH led the conceptualization of the study, with support from HN. KC led data curation, with support from HN. MH led the formal analysis, with HN and KC contributing equally. MH acquired the funding and led the methodology. An investigation was conducted equally by HN, KC, DV, NI, SS, SY, HLS, IS, BZB, and SM. Project administration was shared equally by KC and HN. KC led validation, with support from HN. Visualization was carried out equally by HN, KC, and MH. MH led the writing of the original draft and the review and editing process, with supporting contributions from HN, KC, DV, NI, SS, SY, HLS, IS, BZB, and SM.

Conflicts of Interest

DV is a board member of an organization that is a partner of an Ontario Health Team.

Multimedia Appendix 1
Website rating form.
[PDF File, 113 KB - aging_v9i1e80595_app1.pdf]

Multimedia Appendix 2
Ratings and comments for each of the 58 Ontario Health Team websites.
[PDF File, 69 KB - aging_v9i1e80595_app2.pdf]

References

XSL•FO
RenderX

1. Schulz R, Beach SR, Czaja SJ, Martire LM, Monin JK. Family caregiving for older adults. Annu Rev Psychol 2020 Jan 4;71(1):635-659. [doi: 10.1146/annurev-psych-010419-050754] [Medline: 31905111]
2. Parmar J, Torti J, Brémault-Phillips S, Charles L, Chan K, Tian PGJ. Supporting family caregivers of seniors within acute and continuing care systems. Can Geriatr J 2018 Dec;21(4):292-296. [doi: 10.5770/cgj.21.317] [Medline: 30595779]
3. Joseph LM. Aging in place: interprofessional approaches to empower informal caregivers. J Nurse Pract 2023 Jan;19(1):104412. [doi: 10.1016/j.nurpra.2022.07.020]
4. Bangerter LR, Kim Y, Fields B, Wittke MR, Perepezko K. The intersection of successful aging and family caregiving. Gerontologist 2024 Dec 13;65(1):gnae054. [doi: 10.1093/geront/gnae054] [Medline: 38795006]
5. A National Caregiving Strategy for Canada. : Canadian Centre for Caregiving Excellence (CCCE); 2025 URL: https://canadiancaregiving.org/wp-content/uploads/2025/02/National-Care-Giving-Strategy-FINAL-WEB.pdf [accessed 2025-03-20]
6. Wolff JL, Cornman JC, Freedman VA. The number of family caregivers helping older US adults increased from 18 million to 24 million, 2011-22. Health Aff (Millwood) 2025 Feb;44(2):187-195. [doi: 10.1377/hlthaff.2024.00978] [Medline: 39899774]
7. “Sandwiched” between unpaid care for children and care-dependent adults: a gender-based study. Government of Canada. 2024. URL: https://www150.statcan.gc.ca/n1/pub/89-652-x/89-652-x2024002-eng.htm [accessed 2025-03-20]
8. Peña-Longobardo LM, Oliva-Moreno J. The economic value of non-professional care: a Europe-wide analysis. Int J Health Policy Manag 2022 Oct 19;11(10):2272-2286. [doi: 10.34172/ijhpm.2021.149] [Medline: 34814681]
9. Zigante V. Informal care in Europe: exploring formalisation, availability and quality.: European Commission; 2018 URL: https://ec.europa.eu/social/BlobServlet?docId=19681&langId=en [accessed 2026-03-09]
10. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. JAMA 2014 Mar 12;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]
11. 1 in 3 unpaid caregivers in Canada are distressed. Canadian Insitute for Health Information (CIHI). 2020. URL: https://www.cihi.ca/en/1-in-3-unpaid-caregivers-in-canada-are-distressed [accessed 2023-06-05]
12. Reinhard SC, Caldera S, Houser A, Choula RB. Valuing the invaluable: 2023 update - strengthening supports for family caregivers. : American Association of Retired Persons (AARP); 2023 URL: https://www.aarp.org/content/dam/aarp/ppi/2023/3/valuing-the-invaluable-2023-update.doi.10.26419-2Fppi.00082.006.pdf
13. MacDonald BJ, Wolfson M, Hirdes JP. The future co$t of long-term care in Canada. : National Institute on Aging; 2019 URL: https://niageing.ca/reports/the-future-cot-of-long-term-care-in-canada/ [accessed 2026-03-09]
14. Fast J, Duncan KA, Keating NC, Kim C. Valuing the contributions of family caregivers to the care economy. J Fam Econ Iss 2024 Mar;45(1):236-249. [doi: 10.1007/s10834-023-09899-8]
15. Parmar J, Anderson S, Dobbs B, et al. Neglected needs of family caregivers during the COVID-19 pandemic and what they need now: a qualitative study. Diseases 2021 Oct 13;9(4):70. [doi: 10.3390/diseases9040070] [Medline: 34698124]
16. Funk LM, Klostermann J, Symonds-Brown H, Aubrecht K, Giguère L. Producing the public caregiver: the discursive politicization of family caregiving by Canadian caregiver organizations. International Journal of Care and Caring ;9(4):792-809. [doi: 10.1332/23978821Y2024D000000079]
17. Müller B. The careless society—dependency and care work in capitalist societies. Front Sociol 2019;3. [doi: 10.3389/fsoc.2018.00044]
18. Gopnik A. Caregiving in philosophy, biology & political economy. Daedalus 2023 Feb 28;152(1):58-69. [doi: 10.1162/daed_a_01961]
19. Armstrong P. Unpaid Work in Nursing Homes: Bristol University Press; 2023.
20. Fraser N. Cannibal Capitalism: How Our System Is Devouring Democracy, Care, and the Planet – and What We Can Do about It: Verso; 2022.
21. Janson P, Willeke K, Zaibert L, et al. Mortality, morbidity and health-related outcomes in informal caregivers compared to non-caregivers: a systematic review. Int J Environ Res Public Health 2022 May 11;19(10):5864. [doi: 10.3390/ijerph19105864] [Medline: 35627399]
22. Marani H, Peckham A. Unpaid caregiver costs in Canada: a systematic review. Home Health Care Manag Pract 2023 Nov;35(4):277-286. [doi: 10.1177/10848223231169504]
23. Schmidt T, Juday C, Patel P, Karmarkar T, Smith-Howell ER, Fendrick AM. Expanding the catalog of patient and caregiver out-of-pocket costs: a systematic literature review. Popul Health Manag 2024 Feb;27(1):70-83. [doi: 10.1089/pop.2023.0238] [Medline: 38099925]
24. Moody E, Ganann R, Martin-Misener R, et al. Out-of-pocket expenses related to aging in place for frail older people: a scoping review. JBI Evid Synth 2022 Feb;20(2):537-605. [doi: 10.11124/JBIES-20-00413] [Medline: 34738979]
25. Bom J, Bakx P, Schut F, van Doorslaer E. The impact of informal caregiving for older adults on the health of various types of caregivers: a systematic review. Gerontologist 2019 Sep 17;59(5):e629-e642. [doi: 10.1093/geront/gny137] [Medline: 30395200]
26. Lindt N, van Berkel J, Mulder BC. Determinants of overburdening among informal carers: a systematic review. BMC Geriatr 2020 Aug 26;20(1):304. [doi: 10.1186/s12877-020-01708-3] [Medline: 32847493]

XSL•FO
RenderX

27. Choi JY, Lee SH, Yu S. Exploring factors influencing caregiver burden: a systematic review of family caregivers of older adults with chronic illness in local communities. Healthcare (Basel) 2024 May 13;12(10):1002. [doi: 10.3390/healthcare12101002] [Medline: 38786412]
28. Schulz R, Beach SR, Friedman EM. Caregiving factors as predictors of care recipient mortality. Am J Geriatr Psychiatry 2021 Mar;29(3):295-303. [doi: 10.1016/j.jagp.2020.06.025] [Medline: 32718853]
29. Spiers GF, Kunonga TP, Stow D, et al. Factors associated with unmet need for support to maintain independence in later life: a systematic review of quantitative and qualitative evidence. Age Ageing 2022 Oct 6;51(10):afac228. [doi: 10.1093/ageing/afac228] [Medline: 36309974]
30. Nieuwenhuis AV, Beach SR, Schulz R. Care recipient concerns about being a burden and unmet needs for care. Innov Aging 2018 Sep;2(3):igy026. [doi: 10.1093/geroni/igy026] [Medline: 30480145]
31. Beach SR, Schulz R. Family caregiver factors associated with unmet needs for care of older adults. J Am Geriatr Soc 2017 Mar;65(3):560-566. [doi: 10.1111/jgs.14547] [Medline: 27935019]
32. Ejem D, Bauldry S, Bakitas M, Drentea P. Caregiver burden, care recipient depressive symptomology, and social exchange: does race matter? J Palliat Care 2018 Apr;33(2):100-108. [doi: 10.1177/0825859718758120] [Medline: 29514553]
33. Betini RSD, Hirdes JP, Lero DS, Cadell S, Poss J, Heckman G. A longitudinal study looking at and beyond care recipient health as a predictor of long term care home admission. BMC Health Serv Res 2017 Nov 9;17(1):709. [doi: 10.1186/s12913-017-2671-8] [Medline: 29121916]
34. Kim B, Wister A, O'dea E, Mitchell BA, Li L, Kadowaki L. Roles and experiences of informal caregivers of older adults in community and healthcare system navigation: a scoping review. BMJ Open 2023 Dec 9;13(12):e077641. [doi: 10.1136/bmjopen-2023-077641] [Medline: 38070939]
35. Wu J, Zhang H, Shao J, et al. Healthcare for older adults with multimorbidity: a scoping review of reviews. Clin Interv Aging 2023;18:1723-1735. [doi: 10.2147/CIA.S425576] [Medline: 37868094]
36. Barreira LF, Paiva A, Araújo B, Campos MJ. Challenges to systems of long-term care: mapping of the central concepts from an umbrella review. Int J Environ Res Public Health 2023 Jan 17;20(3):1698. [doi: 10.3390/ijerph20031698] [Medline: 36767064]
37. Bangerter LR, Griffin J, Harden K, Rutten LJ. Health information-seeking behaviors of family caregivers: analysis of the Health Information National Trends Survey. JMIR Aging 2019 Jan 14;2(1):e11237. [doi: 10.2196/11237] [Medline: 31518309]
38. Ontario Health Teams. Ontario Ministry of Health. 2024. URL: http://www.ontario.ca/page/ontario-health-teams [accessed 2024-08-28]
39. Improving health care in Ontario. Ontario Health Teams. 2024. URL: http://www.ontario.ca/page/improving-health-care-ontario [accessed 2025-04-03]
40. Kirst M, Im J, Burns T, et al. What works in implementation of integrated care programs for older adults with complex needs? A realist review. Int J Qual Health Care 2017 Oct 1;29(5):612-624. [doi: 10.1093/intqhc/mzx095] [Medline: 28992156]
41. Wodchis WP, Dixon A, Anderson GM, Goodwin N. Integrating care for older people with complex needs: key insights and lessons from a seven-country cross-case analysis. Int J Integr Care 2015;15:e021. [doi: 10.5334/ijic.2249] [Medline: 26528096]
42. Ontario Health Teams: Guidance for Health Care Providers and Organizations. : Ontario Ministry of Health; 2024 URL: https://www.ontario.ca/files/2024-01/moh-oht-hcp-guidance-doc-en-2024-01-22.pdf [accessed 2026-03-02]
43. Ontario introduces 24 Ontario Health Teams across the province to provide better connected care. Ontario Ministry of Health. 2019. URL: https://news.ontario.ca/en/release/55008/ontario-introduces-24-ontario-health-teams-across-the-province-to-provide-better-connected-care [accessed 2025-10-12]
44. Ontario health teams: the path forward. : Ontario Ministry of Health; 2022 URL: https://www.ontario.ca/files/2024-01/moh-oht-path-forward-en-2024-01-22.pdf#:~:text=In%202019%2C%20Ontario%20Health%20Teams%20(OHTs)%20were,and%20engagement%20with%20patients%2C%20families%20and%20caregivers [accessed 2026-03-02]
45. Ontario Health Teams now providing care in every corner of the province. Ontario Ministry of Health. 2024. URL: https://news.ontario.ca/en/release/1004116/ontario-health-teams-now-providing-care-in-every-corner-of-the-province [accessed 2025-10-12]
46. Ontario Health Teams: guidance for health care providers and organizations. : Government of Ontario; 2021 URL: https://www.ontario.ca/files/2024-01/moh-oht-hcp-guidance-doc-en-2024-01-22.pdf [accessed 2026-03-02]
47. Ontario Health Teams: patient, family and caregiver partnership and engagement strategy: guidance document. : Government of Ontario; 2021 URL: https://www.ontario.ca/files/2024-01/moh-oht-pfc-partner-engagement-strat-guidance-en-2024-01-22.pdf [accessed 2026-03-02]
48. Patient, family and caregiver declaration of values for Ontario. Government of Ontario. 2021. URL: http://www.ontario.ca/page/patient-family-caregiver-declaration-values-ontario [accessed 2025-04-18]

XSL•FO
RenderX

49. Abelson J, Tripp L, Abdelhalim R, et al. Building engagement-capable environments for health system transformation: development and early implementation of a capability framework for patient, family and caregiver engagement in Ontario Health Teams. Health Expect 2024 Dec;27(6):e70083. [doi: 10.1111/hex.70083] [Medline: 39492464]
50. The Ontario Caregiver Organization Financial Statements. : The Ontario Caregiver Organization; 2021 URL: https://ontariocaregiver.ca/wp-content/uploads/2024/03/OCO-Financial-Statements-F2021.pdf [accessed 2026-03-02]
51. About us. The Ontario Caregiver Organization. 2024. URL: https://ontariocaregiver.ca/about/about-oco/ [accessed 2025-04-18]
52. Momentum: 2023-24 impact report. : The Ontario Caregiver Organization; 2025 URL: https://ontariocaregiver.ca/wp-content/uploads/2024/06/OCO-Impact-Report-2023-24-English-June21_FINAL-s.pdf [accessed 2026-03-02]
53. Massie J. Deliberating with purpose: deliberative civic engagement for health policy. Can Public Adm 2023 Dec;66(4):556-573 [FREE Full text] [doi: 10.1111/capa.12545]
54. Sibbald SL, Kokorelias KM, Embuldeniya G, Wodchis WP. Engagement of patient and family advisors in health system redesign in Canada. J Health Serv Res Policy 2023 Jan;28(1):25-33. [doi: 10.1177/13558196221109056] [Medline: 35786052]
55. Côté-Boileau É, Rahim A, Vollmer B, Harrilall N, Robertson S. Mapping the newcomer journey for more equitable population health: insights from an Ontario Health Team. Healthc Q 2023 Jan;25(4):23-29. [doi: 10.12927/hcq.2023.27021] [Medline: 36826237]
56. MacNeil M, Tagami A, Sheffield P, et al. Am I attached? A patient-partnered approach to creating infographics about attachment to primary care in Ontario, Canada. Res Involv Engagem 2024 Nov 4;10(1):114. [doi: 10.1186/s40900-024-00652-5] [Medline: 39497223]
57. Policy brief 1: building a better Ontario by listening to the voices of caregivers. Ontario Caregiver Coalition. 2024. URL: https://www.ontariocaregivercoalition.ca/policy-brief-1 [accessed 2025-04-18]
58. Policy brief 3: building a better Ontario for caregivers by providing mental health supports. Ontario Caregiver Coalition. 2024. URL: https://www.ontariocaregivercoalition.ca/policy-brief-3 [accessed 2025-04-18]
59. Policy brief 4: building a better Ontario for caregivers by supporting caregiver respite. Ontario Caregiver Coalition. 2025. URL: https://www.ontariocaregivercoalition.ca/policy-brief-4 [accessed 2025-04-18]
60. Caring in Canada: survey insights from caregivers and care providers across Canada. : Canadian Centre of Caregiving Excellence; 2024 URL: https://canadiancaregiving.org/wp-content/uploads/2024/06/CCCE_Caring-in-Canada.pdf [accessed 2026-03-02]
61. Caring in Canada: snapshot on Ontario. Canadian Centre of Caregiving Excellence. 2024. URL: https://canadiancaregiving.org/wp-content/uploads/2024/12/CIC_Info_ON.jpg [accessed 2025-04-18]
62. Public-facing communications guidance for Ontario Health Teams. Ontario Ministry of Health. 2024. URL: https://www.ontario.ca/files/2024-12/moh-public-facing-communications-guidance-for-oht-en-2024-12-11.pdf [accessed 2025-10-13]
63. Charlton P, Kean T, Liu RH, et al. Use of environmental scans in health services delivery research: a scoping review. BMJ Open 2021 Nov 10;11(11):e050284. [doi: 10.1136/bmjopen-2021-050284] [Medline: 34758992]
64. Wilburn A, Vanderpool RC, Knight JR. Environmental scanning as a public health tool: Kentucky's human papillomavirus vaccination project. Prev Chronic Dis 2016 Aug 18;13:E109. [doi: 10.5888/pcd13.160165] [Medline: 27536901]
65. Population estimates, quarterly. Statistics Canada. 2025. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000901 [accessed 2025-04-21]
66. Population estimates on July 1, by age and gender. Statistics Canada. 2024. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000501 [accessed 2025-04-21]
67. Deber R. Treating Health Care: How the Canadian System Works and How It Could Work Better: University of Toronto Press; 2019.
68. Marchildon G, Allin S, Merkur S. Health Systems in Transition: Canada: Health System Review 2020, 3rd edition: University of Toronto Press; 2021. [doi: 10.3138/9781487537517]
69. Find a retirement home: learn about licensed retirement homes and how to choose a home. Ontario. 2024. URL: http://www.ontario.ca/page/find-retirement-home [accessed 2025-04-21]
70. Home and community care. Government of Ontario. 2024. URL: http://www.ontario.ca/page/home-community-care [accessed 2025-04-21]
71. Ministry of long-term care. Government of Ontario. 2024. URL: http://www.ontario.ca/page/ministry-long-term-care [accessed 2025-04-21]
72. Value-for-money audit: long-term care homes: delivery of resident-centred care. Office of the Auditor General of Ontario. 2023. URL: https://www.auditor.on.ca/en/content/annualreports/arreports/en23/AR_LTCresidential_en23.pdf#:~:text=There%20are%20currently%20626%20long%2Dterm%20care%20homes,than%2020%20beds%20to%20over%20400%20beds [accessed 2026-03-13]
73. Long-term care. Ontario Health atHome. 2022. URL: https://ontariohealthathome.ca/long-term-care/ [accessed 2025-05-22]
74. Retirement homes regulatory authority: follow-up on value-for-money audit, 2020 annual report. Office of the Auditor General of Ontario. 2022. URL: https://www.auditor.on.ca/en/content/annualreports/arreports/en22/1-12retirementhomes_en22.pdf [accessed 2025-04-21]
75. Improving Ontario's supportive housing programs 2020 – 2021. Government of Ontario. 2022. URL: http://www.ontario.ca/page/improving-ontarios-supportive-housing-programs-2020-2021 [accessed 2025-05-22]

XSL•FO
RenderX

76. Supportive housing. Ontario Health atHome. 2022. URL: https://ontariohealthathome.ca/supportive-living/supportive-housing/ [accessed 2025-05-22]
77. Home care. Ontario Health atHome. 2025. URL: https://ontariohealthathome.ca/home-care/ [accessed 2025-04-21]
78. Community care. Ontario Health atHome. 2025. URL: https://ontariohealthathome.ca/community-care/ [accessed 2025-04-21]
79. Assisted living. Ontario Health atHome L. 2023. URL: https://ontariohealthathome.ca/supportive-living/assisted-living/
80. Health services for Ontario. Ontario Health atHome. 2025. URL: https://www.thehealthline.ca/splash.aspx [accessed 2025-05-22]
81. Adult day programs. Ontario Health atHome. 2022. URL: https://ontariohealthathome.ca/supportive-living/adult-day-programs/ [accessed 2024-11-26]
82. Respite care. Government of Ontario. 2024. URL: http://www.ontario.ca/page/respite-care [accessed 2025-05-26]
83. Short-stay. Ontario Health atHome. 2022. URL: https://ontariohealthathome.ca/long-term-care/short-stay/ [accessed 2025-05-26]
84. Resources for caregivers. Ontario Health atHome. 2024. URL: https://ontariohealthathome.ca/above-and-beyond-caregiver-recognition-program/resources-for-caregivers/ [accessed 2025-05-26]
85. New plan to modernize home and community care in Ontario. Ontario Ministry of Health. 2020. URL: https://news.ontario.ca/en/backgrounder/55894/new-plan-to-modernize-home-and-community-care-in-ontario [accessed 2025-10-12]
86. Introduction to understanding WCAG 22. World Wide Web Consortium (W3C). 2025. URL: https://www.w3.org/WAI/WCAG22/Understanding/intro#understanding-the-four-principles-of-accessibility [accessed 2025-06-16]
87. American educational research association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME). Standards for Educational and Psychological Testing: AERA; 2014.
88. Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use, 5th edition: Oxford University Press; 2014.
89. Li M, Gao Q, Yu T. Kappa statistic considerations in evaluating inter-rater reliability between two raters: which, when and context matters. BMC Cancer 2023 Aug 25;23(1):799. [doi: 10.1186/s12885-023-11325-z] [Medline: 37626309]
90. Salmon M, Schumacher D, Höhle M. Monitoring count time series in R: aberration detection in public health surveillance. J Stat Soft 2016;70(10):1-35. [doi: 10.18637/jss.v070.i10]
91. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
92. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1):1609406917733847. [doi: 10.1177/1609406917733847]
93. Zhang Y, Sun Y, Xie B. Quality of health information for consumers on the web: A systematic review of indicators, criteria, tools, and evaluation results. Asso for Info Science & Tech 2015 Oct;66(10):2071-2084. [doi: 10.1002/asi.23311]
94. Zhang Y, Kim Y. Consumers' evaluation of web-based health information quality: meta-analysis. J Med Internet Res 2022 Apr 28;24(4):e36463. [doi: 10.2196/36463] [Medline: 35482390]
95. Sun Y, Zhang Y, Gwizdka J, Trace CB. Consumer evaluation of the quality of online health information: systematic literature review of relevant criteria and indicators. J Med Internet Res 2019 May 2;21(5):e12522. [doi: 10.2196/12522] [Medline: 31045507]
96. Ghalavand H, Nabiolahi A. Exploring online health information quality criteria on social media: a mixed method approach. BMC Health Serv Res 2024 Oct 30;24(1):1311. [doi: 10.1186/s12913-024-11838-8] [Medline: 39478573]
97. Peterson K, Hahn H, Lee AJ, Madison CA, Atri A. In the information age, do dementia caregivers get the information they need? Semi-structured interviews to determine informal caregivers' education needs, barriers, and preferences. BMC Geriatr 2016 Sep 23;16(1):164. [doi: 10.1186/s12877-016-0338-7] [Medline: 27662829]
98. Jia X, Pang Y, Liu LS. Online health information seeking behavior: a systematic review. Healthcare (Basel) 2021 Dec 16;9(12):1740. [doi: 10.3390/healthcare9121740] [Medline: 34946466]
99. Knipping D, Garnett A, Jiang BB. Access and use of services by caregivers of older adults: a scoping review of cultural and linguistic diversity. J Appl Gerontol 2023 Jul;42(7):1672-1686. [doi: 10.1177/07334648231158490] [Medline: 36866817]
100. Ramakrishnan C, Malhotra C. Challenges to accessing community dementia care services: a qualitative study. BMC Health Serv Res 2025 May 23;25(1):747. [doi: 10.1186/s12913-025-12895-3] [Medline: 40410863]
101. Hamilton C, Snow ME, Clark N, et al. Quality of patient, family, caregiver and public engagement in decision-making in healthcare systems: a scoping review protocol. BMJ Open 2019 Nov 6;9(11):e032788. [doi: 10.1136/bmjopen-2019-032788] [Medline: 31699750]
102. Kuluski K, Kokorelias KM, Peckham A, Goldhar J, Petrie J, Alloway CA. Twelve principles to support caregiver engagement in health care systems and health research. Patient Exp J 2019;6(1):141-148. [doi: 10.35680/2372-0247.1338]
103. Sittig DF, Wright A, Coiera E, et al. Current challenges in health information technology-related patient safety. Health Informatics J 2020 Mar;26(1):181-189. [doi: 10.1177/1460458218814893] [Medline: 30537881]
104. Tieman J, Nicholls S. Enhancing the efficacy of healthcare information websites: a case for the development of a best practice framework. BMJ Open 2024 Sep 3;14(9):e088789. [doi: 10.1136/bmjopen-2024-088789] [Medline: 39231553]
105. Ontario Health Team: full application. Ontario Ministry of Health. 2019. URL: https://www.ontario.ca/files/2024-01/moh-oht-full-application-en-2024-01-22.pdf [accessed 2025-10-17]

XSL•FO
RenderX

106. Lagacé M, Mérette M, Navaux J, Rodrigues-Rouleau P. An examination of the social and economic impacts of ageism. : Federal, Provincial and Territorial Forum of Ministers Responsible for Seniors; 2022 URL: https://www.canada.ca/en/employment-social-development/corporate/seniors-forum-federal-provincial-territorial/reports/ageism-social-economic-impacts.html [accessed 2026-03-09]
107. Lévesque M, Oldfield M. A roadmap of noninstitutional living options for people with dementia: “don’t fence me in”. Can J Aging 2025 Mar;44(1):95-104. [doi: 10.1017/S0714980824000308] [Medline: 39564619]
108. Evans B, Fanelli C. Against the People: How Ford Nation Is Dismantling Ontario: Fernwood; 2025.
109. Ageism. Ontario Human Rights Commission. 2024. URL: https://www3.ohrc.on.ca/en/time-action-advancing-human-rights-older-ontarians/ageism [accessed 2025-10-13]
110. The need for social infrastructure to support quality of life and wellbeing for older adults. : Seniors for Social Action (Ontario); 2025 URL: https://www.seniorsactionontario.com/_files/ugd/50033d_e0e77bdd49284ef9878a827c41f4a4a0.pdf [accessed 2026-03-02]
111. Mattei CE. The Capital Order: How Economists Invented Austerity and Paved the Way to Fascism: University of Chicago Press; 2022.
112. Gerstle G. The Rise and Fall of the Neoliberal Order: America and the World in the Free Market Era: Oxford University Press; 2022.
113. Loreto N. The Social Safety Net: Canada in Decline Book One: Dundurn; 2024.
114. Loreto N. Corporate Control: Canada in Decline Book Two: Dundurn; 2025.
115. Grady C, Chan-Nguyen S, Mathies D, Alam N. Family physicians partnering for system change: a multiple-case study of Ontario Health Teams in development. BMC Health Serv Res 2023 Oct 17;23(1):1113. [doi: 10.1186/s12913-023-10070-0] [Medline: 37848926]
116. Ladak Z, Akuamoah-Boateng H, Damba C, et al. Engagement with partners is a leading practice in health workforce planning: what health leaders need to know. Healthc Manage Forum 2024 Sep;37(5):377-383. [doi: 10.1177/08404704241263015] [Medline: 39034474]
117. Majid U. Ontario health teams: innovation or history repeating itself? Longwoods. 2024. URL: https://www.longwoods.com/content/27338/ontario-health-teams-innovation-or-history-repeating-itself- [accessed 2026-03-09]
118. Aggarwal M, Williams AP. Tinkering at the margins: evaluating the pace and direction of primary care reform in Ontario, Canada. BMC Fam Pract 2019 Sep 11;20(1):128. [doi: 10.1186/s12875-019-1014-8] [Medline: 31510942]
119. Jarvis T, Smith RW, Sandhu HS, et al. Promise and peril: how health system reforms impacted public health in three Canadian provinces. Can J Public Health 2023 Oct;114(5):714-725. [doi: 10.17269/s41997-023-00785-2] [Medline: 37410363]
120. Exley J, Glover R, Mccarey M, et al. Governing integrated health and social care: an analysis of experiences in three European countries. Int J Integr Care 2024;24(1):9. [doi: 10.5334/ijic.7610] [Medline: 38344427]
121. Tan S, Farmer J, Wodchis WP, Allin S. How do policy supports enable the implementation, scale, and sustainability of integrated care programs in England, Germany, and The Netherlands? Lessons for Canada. Health Policy 2025 Oct;160:105334. [doi: 10.1016/j.healthpol.2025.105334] [Medline: 40348686]
122. Sethuram C, McCutcheon T, Liddy C. An environmental scan of Ontario Health Teams: a descriptive study. BMC Health Serv Res 2023 Mar 8;23(1):225. [doi: 10.1186/s12913-023-09102-6] [Medline: 36890556]
123. Canadian Institute for Health Information (CIHI). URL: https://www.cihi.ca/en [accessed 2025-12-31]

Abbreviations

OHT: Ontario Health Team

Edited by R Yang; submitted 18.Jul.2025; peer-reviewed by B Agyare-Aggrey, M Northwood; revised version received 31.Dec.2025; accepted 20.Feb.2026; published 16.Mar.2026.

Please cite as:
Nguyen H, Corbett K, Vervaecke D, Ivanovski N, Shirazi S, Yadollahi S, Shrestha HL, Slavescu I, Zarandi Baghini B, Medvedyuk S, Hoben M
Public-Facing Communication of Health and Social Services for Older Adults and Their Family or Friend Caregivers: Environmental Scan of 58 Integrated Care Teams’ Websites in Ontario, Canada
JMIR Aging 2026;9:e80595
URL: https://aging.jmir.org/2026/1/e80595
doi:10.2196/80595

XSL•FO
RenderX

© Hung Nguyen, Kyle Corbett, Deanna Vervaecke, Natalia Ivanovski, Soroush Shirazi, Sara Yadollahi, Hom Lal Shrestha, Irene Slavescu, Behnoush Zarandi Baghini, Stella Medvedyuk, Matthias Hoben. Originally published in JMIR Aging (https://aging.jmir.org), 16.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Telehealth-Adapted Dementia Caregiver Skills Training Intervention (TeleCARE): Single-Arm Pre-Post Intervention Study

Maureen K O'Connor[1,2,3], PsyD; Steven D Shirk[3,4], PhD; Jaye E McLaren[3], OTD; Andrew H Nguyen[3], MA; Kendra Pugh[4,5], PhD; Madeline A Sullivan[3], MA; Emily E Metcalf[3], MA; Samantha Harrington[1], MS; Lauren R Moo[3,6], MD

[1]Department of Neuropsychology, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, MA, United States

[2]Department of Neurology, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, United States

[3]Geriatric Research Education and Clinical Center, Edith Nourse Rogers Memorial Veterans Hospital, 200 Springs Road, Bedford, MA, United States

[4]Department of Psychiatry and Behavioral Sciences, Department of Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, Worcester, MA, United States

[5]Mental Illness Research Education and Clinical Center, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, MA, United States

[6]Department of Neurology, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Maureen K O'Connor, PsyD
Department of Neuropsychology, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, MA, United States

Abstract

Background: Dementia caregivers often want to support aging at home, but as neuropsychiatric symptoms (NPS) become more severe, caregiver challenges increase, often resulting in negative outcomes for both the caregiver and care recipient and institutionalization. Project CARE is a manualized in-person group intervention for dementia caregivers designed to reduce negative caregiver outcomes by teaching skills to manage NPS in care recipients in the home environment. Interventions that occur in person, however, can be difficult for caregivers to attend. Telehealth-based interventions are possible alternatives that reduce barriers to attendance.

Objective: The primary objective of this pilot study was to evaluate the feasibility and acceptability of offering CARE via telehealth (TeleCARE). The secondary objective was to explore quantitative outcome trends and effect sizes postintervention outcomes of TeleCARE for both caregivers and care recipients.

Methods: Rates of recruitment, attendance, and completion were used to assess the feasibility of TeleCARE. Data on technology use and telehealth-based adaptations were also collected. Acceptability was measured using participants' rated satisfaction with the intervention immediately postintervention. Questionnaires were administered at baseline and immediately and 3 months postintervention. Primary outcomes for exploratory analysis included NPS presence, severity, and caregiver NPS-related distress. Secondary outcomes included caregiver depression, anxiety, stress, self-efficacy, positive aspects of caregiving, and meaning and purpose in life.

Results: Of the 109 caregivers contacted for recruitment, 24 (22%) caregivers enrolled in TeleCARE, and 20 (83%) caregivers, predominantly female spouses, completed the study. Feedback from participants in the TeleCARE test group 1 (n=3) was used to modify the intervention to improve the telehealth experience, including adding procedures to improve safety, encourage rapport building, address etiquette, and ensure privacy. The final version of TeleCARE included 7 weekly synchronous video sessions. Ten out of 17 participants (59%) attended all 7 sessions, and all participants attended at least 5 sessions. Satisfaction ratings suggested adequate intervention acceptability. Most participants (11/17, 65%) required technological support, which was needed throughout the intervention. Quantitative trends were observed toward postintervention decreases in care-recipient NPS severity (Cohen *d*=0.16), caregiver depression (*d*=0.15), anxiety (*d*=0.23), and caregiver self-efficacy (*d*=0.21), as well as increases in positive aspects of caregiving (*d*=0.18) and meaning and purpose in life (*d*=0.09). Most improvements were not sustained at the 3-month follow-up.

Conclusions: In this pilot feasibility study, dementia caregivers were successfully recruited and engaged in TeleCARE. Overall, TeleCARE was deemed feasible and acceptable. The current findings suggest that offering interventions via telehealth requires modifications and technological support for older caregiver engagement but is a feasible and acceptable means of offering services.

(JMIR Aging 2026;9:e81256) doi:10.2196/81256

KEYWORDS

Alzheimer disease; caregiver education; dementia; older adults; telehealth; caregivers; telemedicine; veterans; neuropsychiatric symptoms

XSL•FO
RenderX

Introduction

Most individuals with dementia experience neuropsychiatric symptoms (NPS), changes in emotional expression and behavior, during the disease course [1-3]. The presence of NPS is associated with increased caregiver depression and burden and decreased quality of life for both the caregivers and care recipients [4]. Although most dementia caregivers want to keep the care recipient aging at home [5], the presence of NPS often spurs institutionalization [6-9]. The association between NPS and institutionalization has been shown to be mediated by caregiver burden and distress [8]. Interventions that teach dementia caregivers skills to manage NPS have the potential to reduce negative caregiver and care recipient outcomes and institutionalization of the care recipient [6,9,10]. However, due to caregiving demands, attending in-person interventions can be prohibitive. Telehealth offers a solution.

The resources for enhancing Alzheimer's caregiver health (REACH) study [11] is one of the most well-studied, targeted, skills-based interventions aimed at improving caregiver well-being. Participation in REACH decreases dementia caregiver anxiety, depression, and burden [12-14]. The results of REACH suggest that targeted skills-based training may be more effective than broader psychoeducational interventions [11]. CARE was designed as a briefer intervention compared with REACH. CARE is a 5-session group intervention for dementia caregivers that teaches skills to manage NPS in the home environment through behavioral management training, pleasant events scheduling, and emotional regulation techniques [15]. CARE sessions run for 90 minutes weekly and are highly structured, consisting of primarily didactic learning and practice, with some space for content-related open discussion. CARE has been shown to reduce caregiver burden related to NPS. REACH, CARE, and other similar programs demonstrate that providing dementia caregivers with specific skills training to manage NPS can be an effective way to support caregivers and improve caregiver outcomes. Unfortunately, many dementia caregiver skills training programs have historically been offered in person only, limiting accessibility for people with barriers to accessing in-person services.

Many dementia caregivers have unique barriers to in-person care. For example, in 2025, 60% of dementia caregivers were working an average 35 hours per week while caregiving [5], limiting the time available to travel to and participate in interventions. Dementia caregivers also have disproportionate financial stressors, incurring approximately 4 times the average out-of-pocket costs compared with nondementia caregivers, creating a financial barrier to using in-person services (eg, they cannot pay for gas, respite care, or take time off from work) [5]. Other dementia caregiver–specific concerns that limit in-person accessibility of services include feeling uncomfortable leaving the diagnosed individual at home and having difficulty finding respite care [16,17]. For caregivers living in rural areas, there is less awareness of available services, greater difficulties accessing transportation, and barriers related to the sheer distance needed to travel for services compared to urban caregivers [18,19]. We can address these challenges by offering interventions via video telehealth [20].

The acceptance and use of telehealth in dementia care has expanded in recent years, demonstrating positive impacts on depression, anxiety, loneliness, NPS management, and satisfaction and quality of life for both caregivers and care recipients [21]. Rapid expansion of telehealth during the COVID-19 pandemic changed the landscape of the health care system [22]. Older adults and their caregivers expressed satisfaction with telehealth and a desire to keep using telehealth postpandemic [23]. Studies of dementia caregiver telehealth support groups have demonstrated effectiveness in improving caregiver burden, reducing depressive symptoms, and increasing coping skills [24-27]. A systematic review [28] suggested that telehealth supportive interventions were potentially as effective as in-person interventions for reducing dementia caregivers' depression, anxiety, stress, and improving self-efficacy. One recent meta-analysis and systematic review [29] specifically investigating the effectiveness of telehealth skills training for management of NPS in dementia found improvements in caregivers' ability to manage NPS, quality of life, and stress levels.

The primary goal of this study was to evaluate the feasibility and acceptability of offering CARE via a telehealth format (TeleCARE) to reduce many of the current barriers that dementia caregivers face in accessing skills training to manage NPS. Keeping the small sample size of this pilot study in mind, the second objective was to explore quantitative outcomes of TeleCARE with trends and effect sizes.

Methods

Participants

We recruited participants from the Bedford, Massachusetts, and Manchester, New Hampshire, Department of Veterans Affairs Medical Centers and surrounding communities using posted informational flyers. We also reviewed Veterans Affairs (VA) medical records to identify veterans with dementia and mailed recruitment letters to their homes and followed up with phone calls. Criteria for study inclusion were (1) providing informal care for someone with self-reported or physician-confirmed diagnosis of dementia, (2) at least 1 caregiver-reported NPS for the care recipient at the time of enrollment, (3) provision of a minimum of 5 hours per week of direct caregiving, and (4) access to a computer or tablet and internet or Wi-Fi. Participants could be caregivers of veterans with dementia or veterans caring for someone with dementia.

Pilot and feasibility work typically aims to explore recruitment, engagement, and practicality rather than statistical power [30,31]. Therefore, for this pilot study, the focus was on testing feasibility and acceptability rather than testing hypothesis-driven significance with associated power calculations for sample size. To accomplish these goals, we selected a target enrollment of 24 caregivers.

Ethical Considerations

This study was approved and monitored by the VA Bedford Health Care System Institutional Review Board (study no. 1598706). Written informed consent was obtained from all study participants. Privacy and confidentiality of all research

XSL•FO
RenderX

participants' data and identities were maintained using multiple methods, including using a Health Insurance Portability and Accountability Act–compliant web-based platform with end-to-end encryption, coding data so that only the research team had access to identifiable information, storing data securely behind the VA firewall and restricting access to the study team, analyzing coded data and reporting data in aggregate form, and establishing ground rules during group sessions to help participants protect their privacy. Participants received US $25 for completing each of the baseline and postintervention follow-up visits, and US $50 for completing the 3-month follow-up visit, receiving a possible total of US $100 to compensate for their time.

Methodological Adaptations, Procedures, and Intervention Description

Initially, we intended to conduct preintervention, in-person sessions to obtain written consent and complete baseline assessments, as well as to provide in-person instruction about technology use. However, prior to the start of the study, the COVID-19 pandemic forced us to convert all study visits to a virtual format. We added a preintervention live video session to provide technology instruction before the start of groups. Individual assistance to help participants access the initial tutorial was provided as needed by phone. Follow-up data collection visits were conducted by video immediately and 3 months post-treatment. Participants were compensated for their time.

A trained interviewer met with each eligible caregiver via video to explain the study, answer questions, review the Health Insurance Portability and Accountability Act authorization, and sign the informed consent forms. The interviewer also provided initial training on using the videoconferencing platform. During the same visit or a second visit, depending on participant preference, the interviewer administered the baseline questionnaires. Questions were presented visually via PowerPoint (Microsoft Corp) one at a time and read aloud. Participant responses were recorded in VA REDCap (Research Electronic Data Capture), a secure data management software, in real-time. After baseline assessment, a printed TeleCARE manual was mailed to the participant. All caregivers began the intervention within 6 weeks of baseline assessment. The collection of immediate and 3-month follow-up data followed the same procedures as baseline data collection. Additional supplemental visits took place by video or phone as needed to help participants troubleshoot technology issues. All video visits were conducted via Health Insurance Portability and Accountability Act–compliant, encrypted, VA-approved videoconferencing platforms.

TeleCARE was adapted from the original manualized CARE intervention [15]. As in CARE, TeleCARE groups met weekly for 90-minute sessions. At the beginning of each session, a doctoral-level occupational therapist leading the groups went through a series of prompts specific to telehealth to ensure safety and confidentiality during the session. Participants were given time during each session to discuss homework, troubleshoot any problems that may have arisen when attempting to use new skills, and offer each other support and advice.

The intervention was delivered to caregivers in 6 groups that ran sequentially from June 2021 to September 2022. We started with an initial TeleCARE test group (TeleCARE-test) to troubleshoot any issues before the formal pilot. TeleCARE-test participants reported that the 5 sessions did not provide adequate time to cover all the intervention material and allow for discussion. In response, we reviewed all intervention sessions to determine the number of additional sessions needed. The finalized program was extended to 7 sessions. Table 1 provides a summary of each session included in the final manualized intervention. Other intervention-specific modifications to adapt the program to the telehealth setting were made based on the experiences of the study team and feedback from the TeleCARE-test group. At the beginning of each session, we added unique steps to ensure participant safety in the telehealth setting. These steps included gathering information about where the participant was located in the event of a psychological (eg, expressed suicidal intent) or medical (eg, cardiac arrest) emergency. Unlike in-person interventions, telehealth interventions require the provider to know where the participant is located if the need arises to alert emergency services (eg, local police). The group leader also asked for updated contact information at the start of each session for participants joining from a new, unknown location. This information was provided to the group leader through secure chat to maintain privacy. The interventionist encouraged all participants to sign into the platform with their first name to facilitate group rapport. It became clear that video-specific etiquette guidelines should be added to the first session, including muting oneself when not speaking, using the raise hand feature, remaining primarily stationary during the group (eg, not walking around with the video device), and reducing background distractors (eg, television, pets, children). Participants were also asked to locate themselves in a suitable space to reduce privacy concerns (ie, a location where they could not be overheard discussing personal information). Finally, the need for broad technological support for older adults enrolled in TeleCARE. In response, we added a technology tutorial session before the intervention started, as noted above, and access to ongoing support as needed.

XSL•FO
RenderX

Table . Telehealth adaptation of CARE (TeleCARE) group sessions.

Session	Description
1 - 2	*Introduction*: Caregivers introduce themselves to begin building group cohesion. Facilitators review dementia and caregiver burden, with a focus on neuropsychiatric symptoms, the importance of increasing pleasant activities as one method for improving care recipients' behavior, and the caregiver–care recipient relationship. Facilitators present a Pleasant Activity Log.
3	*Increasing Pleasant Activities and Improving Communication*: Discuss the importance of engaging in pleasant activities, suggest common pleasant activities and strategies for generating personally meaningful pleasant activities, and provide education about how to increase pleasant activity engagement. Introduce strategies to improve communication. *Caregiver Homework*: Engage in one new pleasant activity daily and record it, generate additional ideas for pleasant activities, and compliment the care recipient at least once per day.
4 - 5	*Increasing Pleasant Activities and Understanding Behavior*: Review homework and problem-solve any difficulties with homework. Review strategies for increasing pleasant activities. Introduce behavioral management theory and teach participants to identify antecedents, behavior, and consequences (ABCs) to assist with behavior management. Discuss triggers for problem behaviors and how they can be avoided. *Caregiver Homework*: Engage in one pleasant activity daily. Identify and record 3 patient problem behaviors.
6	*Understanding and Changing Difficult Behaviors*: Review homework and problem-solve any difficulties with homework. Practice goal setting and methods for changing problem behaviors in the home. *Caregiver Homework*: Engage in 2 pleasant activities daily. Choose one difficult behavior and attempt to modify it.
7	*Final Review*: Review homework and problem-solve difficulties caregivers had while modifying behaviors at home. Provide guidance on how to continue using these new skills in the future. Caregivers discuss and explore plans for future use of strategies. Celebrate completion of the group, participant contributions, personal growth, challenges, and willingness to participate.

Measures

Following CONSORT (Consolidated Standards of Reporting Trials) guidelines [32] and the proposed guidelines for reporting feasibility study outcomes [33], we collected multiple measures of study feasibility, including recruitment rates, session attendance rates, study completion rates, and a formal rating of the acceptability of the intervention. Given our interest in the telehealth adaptation, we also collected data on technology use. Participants completed outcome measures at baseline, 1 to 2 weeks immediately after the intervention, and 3 months postintervention.

At baseline, caregivers completed a demographic and historical information survey, providing information about themselves and the care recipient. Caregivers also completed [1] the clinical dementia rating sum of boxes [34], a 5-point interview that characterizes 6 domains of the care recipient's cognitive and functional performance and disease severity.

At baseline, immediate follow-up, and 3-month follow-up, caregivers were asked to report on the presence of the care recipient's NPS and functional abilities. The Lawton Instrumental Activities of Daily Living (IADL) Scale [35] assesses the functional capacity of the care recipient in 8 domains of functioning. Scores total from 0 to 8, with higher scores indicating greater functional independence. The NPS inventory questionnaire [36] measures both the presence and severity of the care recipient's NPS and the caregiver's distress related to NPSs. Caregivers were also asked to complete measures related to their own well-being. These included the Beck Depression Inventory-II [37], a 21-item scale to assess the severity of depressive symptoms, and the Beck Anxiety Inventory [38], a 21-item scale to assess the severity of anxiety. The revised Caregiver Self-Efficacy Scale [39] is a 15-item scale exploring caregiver self-efficacy in obtaining respite, responding to patients' disruptive behaviors, and controlling upsetting thoughts. The positive aspects of caregiving (PAC) questionnaire [40] consists of nine items that load onto two factors, self-affirmation and outlook on life, and assesses positive reflections on caregiving. The Relative Stress Scale [41] consists of 15 items and can be divided into 3 subgroups of emotional distress, social distress, and negative feelings. The meaning in life questionnaire [42] is a 10-item self-report assessment designed to measure the presence of and search for meaning in life.

At immediate follow-up only (ie, immediately after completing the TeleCARE intervention), participants also completed the

XSL•FO
RenderX

client satisfaction questionnaire (CSQ-8) [43] to capture the acceptability of the intervention.

Data Analysis

The primary objective of this pilot study was to assess the acceptability and feasibility of TeleCARE. Descriptive statistics were used because the goal was to explore trends rather than test hypotheses, as appropriate for the pilot stage of the current project. As such, we explored numerical trends and effect sizes of outcome variables. Descriptive statistics are reported to characterize the sample and summary scores for each measurement time point (baseline, post-treatment, and a 3-mo follow-up). Cohen *d* effect sizes were calculated to compare changes from baseline to post-treatment and from baseline to 3-month follow-up. All statistical analyses were conducted using SAS software (version 9.4; SAS Institute Inc). Given the early stage of this investigation, data from the TeleCARE test group were excluded; analyses only included participants in groups 2 to 6 who completed the final format of the TeleCARE intervention (n=17). An intent-to-treat approach was not used for this preliminary feasibility investigation.

Results

Participant Characteristics

Participants included 17 White caregivers (100%), who were predominantly female (n=13, 76%) and spouses (n=14, 82%), with a mean age of 71.8 (SD 12.58) years. All caregivers lived with the care recipient (n=17, 100%), and nearly all caregivers (n=16, 94%) reported providing the majority of care. Care recipients had a mean age of 77.7 (SD 4.84) years, were White (n=17, 100%), male (n=15, 88%), and had a mean clinical dementia rating sum of boxes score of 10.8, indicative of moderate dementia severity [44]. See Table 2.

XSL•FO
RenderX

Table . Baseline participant characteristics (N=17).

Characteristic	Participants
CR[a] age (year), mean (SD)	77.7 (4.84)
CR sex (male), n (%)	15 (88)
CR race (White), n (%)	17 (100)
CG[b] age (year), mean (SD)	71.8 (12.58)
CG sex (female), n (%)	13 (76)
CG race (White), n (%)	17 (100)
CG education, n (%)	
High school degree	1 (6)
Some college	8 (47)
College degree	7 (41)
Graduate degree	1 (6)
Household income (US $), n (%)	
0	1 (6)
1-24,999	2 (13)
25,000-49,999	2 (13)
50,000-74,999	7 (44)
75,000-99,999	3 (19)
100,000-149,999	0 (0)
150,000-199,999	1 (6)
Cohabitation, n (%)	
No	0 (0)
Yes	17 (100)
Relationship, n (%)	
Spouse/partner	14 (82)
Adult child	3 (18)
Percentage of caregiving, n (%)	
0%-25%	0 (0)
26%-50%	1 (6)
51%-75%	0 (0)
76%-100%	16 (94)
CDR[c] sum of boxes, mean (SD)	10.8 (3.33)
IADL[d], mean (SD)	1.9 (1.6)
NPI[e], mean (SD)	
Total	6.9 (1.65)
Severity	15.5 (5.08)
Distress	26.8 (10.24)

[a]CR: care recipient.

[b]CG: caregiver.

[c]CDR: clinical dementia rating.

[d]IADL: Instrumental Activities of Daily Living.

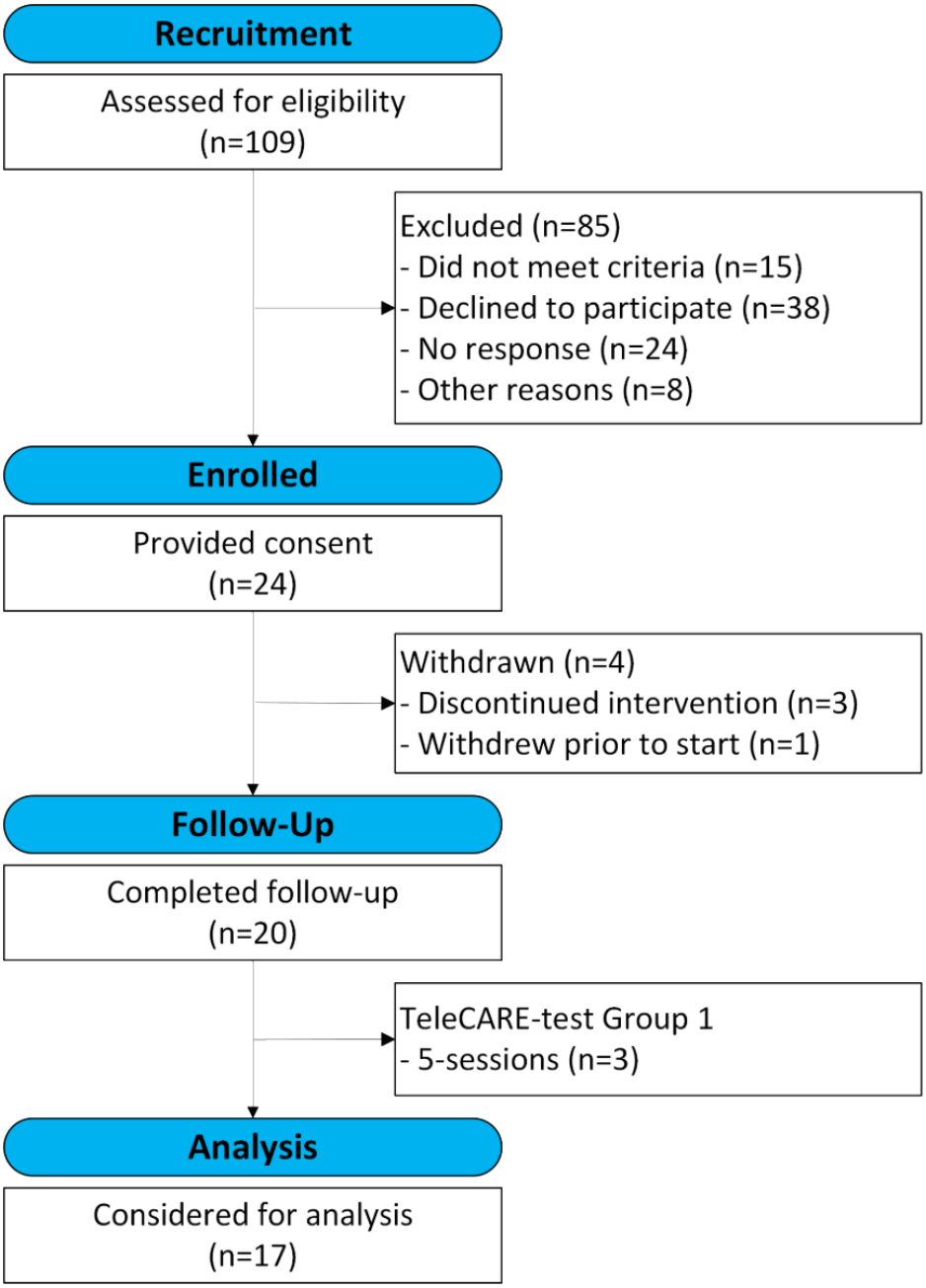
[e]NPI: neuropsychiatric inventory.

XSL•FO
RenderX

Feasibility Results: Recruitment, Enrollment, Attendance, and Retention

Of the 109 caregivers contacted for recruitment, 24 did not respond, 15 did not meet eligibility criteria, 38 declined participation, and 8 had other reasons (Figure 1). Among those who declined (38/109, 35%), the most common reason was that they felt that the intervention was not needed (20/38, 53%). The second most common reason was lack of time (11/38, 29%). The remaining provided no reason (3/38, 8%), said they were not interested (3/38, 8%), or reported a technological barrier (1/38, 3%). Of those that gave other responses (8/109, 7%), 5 said that their loved one was deceased when we contacted them, 2 had been admitted to long-term care, and 1 had recently been diagnosed with cancer. Overall, 24 caregivers (24/109, 22%) were enrolled in the study, with 4 in the TeleCARE-test group 1 and 20 in intervention groups 2 to 6. Across groups 1 to 6, 20 caregivers (20/24, 83%) completed the study. Of the 4 noncompleters, 1 was from the TeleCARE-test group 1, and 3 were from groups 2 to 6. Two of 4 noncompleters withdrew because they disliked the group setting. The remaining 2 noncompleters had medical issues arise that prevented continued participation (1 during the intervention and 1 prior to the start of the intervention but after study enrollment). Because the TeleCARE test group 1 was treated as a test group and subsequent changes were made to the intervention after their feedback, completers from TeleCARE test group 1 (3/4) were not included when reporting intervention attendance or exploratory quantitative analysis presented below. Of those within groups 2 to 6 (n=17), 10 attended all 7 sessions (10/17, 59%), 5 attended 6 of 7 sessions (5/17, 29%), and 2 attended 5 of 7 sessions (2/17, 12%). Occasionally, a participant arrived late, most commonly due to technological difficulties.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.

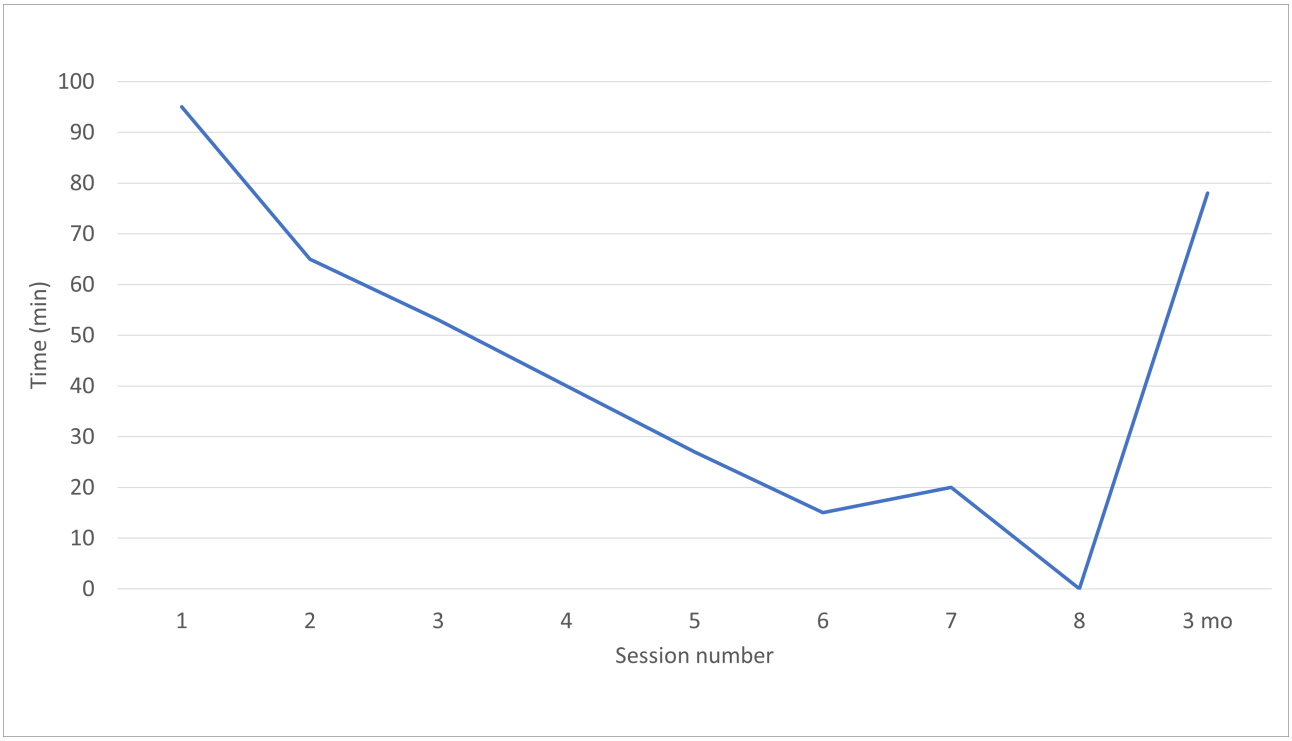


XSL•FO
RenderX

Feasibility Results: Technology Support

To assess the acceptability and feasibility of the telehealth adaptation, data were collected to better understand participant experiences related to the use of technology. Overall, 11 of 17 participants (65%) required some degree of technology support. As mentioned above, a study staff member was assigned to provide personalized technology assistance as needed. Examples of when assistance was needed included helping participants turn on the power if a tablet was provided, access their email and click the video-meeting link, retrieve and/or reset their password, turn on their camera, use the mute and unmute function, and use the raise hand feature. One study participant had no prior experience with technology and needed assistance at all levels, from setting up an email account to using the device. The largest amount of technological assistance was required at session 1, dropping considerably by session 2. Participants needed less support with each passing session, although some continued support remained necessary at almost all time points. Participants reported feeling a sense of accomplishment and increased comfort with technology at the conclusion of the study. At the 3-month follow-up visit, 3 of 17 participants (18%) required assistance with technology, which was an increase compared with session 7 of the intervention. This finding suggests that most study participants were able to retain the taught skills, even after 3 months, while a minority of participants experienced a decay in skills. Figure 2 provides a visual depiction of the amount of time spent providing technological assistance across all sessions for intervention groups 2 to 6.

Figure 2. Total duration of technology assistance given before each session for all participants in groups 2 to 6.



Intervention Acceptability

Smith et al [45] proposed the following guidelines for the interpretation of overall CSQ-8 scores into 4 levels of client satisfaction: “poor” (scores 8 - 13), “fair” (scores 14 - 19), “good” (scores 20 - 25) and “excellent” (scores 26 - 32). The mean CSQ-8 score was 22.3 (SD 1.08), suggesting good intervention acceptance.

Outcome Trends

As mentioned previously, this study was not adequately powered to detect statistical significance, as appropriate for the pilot stage, and the focus was on identifying numerical trends and examining effect sizes. Table 3 provides the mean scores of measures completed at every time point (n=17; also see Table S1 in Multimedia Appendix 1). Between baseline and immediate post-treatment time points, caregivers reported a decline in the care recipient’s ability to complete IADLs. There was no substantial change in the number of NPS reported by the caregiver, though there was a trend towards a decrease in the reported severity of NPS (Cohen d=0.16). Regarding caregiver psychological well-being, trends between baseline and post-treatment time points included a decrease in depression and anxiety symptoms and an increase in PAC and meaning and purpose. In fact, the largest effect between baseline and post-treatment was a decrease in anxiety. Participants reported little change in relative stress. Interestingly, despite improvements in psychological well-being, participants reported a decrease in caregiving self-efficacy between baseline and post-treatment, with the decrease in self-efficacy having the second largest effect size of the outcome measurements. Except for a statistically significant decrease in caregiver self-efficacy (t_{16}=2.17, P=.045, Cohen d=0.21), no outcomes were statistically significant.

XSL•FO
RenderX

Table . Changes in clinical and psychological measures.

Measure	Baseline, mean (min-max)	Posttreatment, mean (min-max)	3-month follow-up, mean (min-max)
IADL[a]	1.9 (0 - 5)	1.6 (0 - 4)	1.5 (0 - 3)
NPI[b]			
Total	6.9 (4-10)	6.8 (2-11)	7.1 (3-10)
Severity	15.5 (6-28)	14.5 (5-28)	15.7 (5-28)
Distress	26.8 (10-46)	25.7 (9-53)	25.8 (9-45)
BDI[c]	8.8 (1-33)	7.6 (0 - 30)	8.5 (0 - 26)
BAI[d]	8.3 (0 - 25)	6.6 (0 - 29)	9.6 (0 - 40)
PAC[e]	34.6 (26-45)	35.8 (26-45)	34.4 (23-45)
RSS[f]	25.9 (9-47)	26.1 (11-42)	25.2 (9-45)
CSES[g]	1347.8 (880 - 1820)	1280 (730 - 1840)	1329.4 (900 - 1840)
MPL[h]	37.9 (17-57)	39.3 (23-62)	39.8 (27-55)

[a]IADL: Instrumental Activities of Daily Living.

[b]NPI: neuropsychiatric inventory.

[c]BDI: Beck Depression Inventory.

[d]BAI: Beck Anxiety Inventory.

[e]PAC: positive aspects of caregiving.

[f]RSS: Relatives' Stress Scale.

[g]CSES: Caregiver Self-Efficacy Scale.

[h]MPL: meaning and purpose in life.

Between baseline and 3-month follow-up, participants reported further decline in IADLs for the care recipient and an increase in NPS. In addition, most improvements reported between baseline and immediate post-treatment were lost by the 3-month follow-up. Participant meaning and purpose in life was the only improvement maintained from post-treatment to the 3-month follow-up. None of the changes between baseline and the 3-month follow-up visits were statistically significant.

Discussion

Feasibility and Acceptability

In this pilot study, we were able to recruit dementia caregivers and engage them in a telehealth adaptation of CARE (TeleCARE). Twenty caregivers completed TeleCARE, including 3 from an initial test group that informed intervention modifications, leaving 17 caregivers in the final intervention group. Caregivers were predominantly older female spouses of the care recipient. Most enrolled caregivers attended all sessions, suggesting that caregivers who engaged in the program stayed engaged throughout. Likewise, satisfaction with the intervention was good. Preliminary quantitative trends suggested immediate postintervention improvements in caregiver outcomes. Overall, TeleCARE was deemed feasible and acceptable, with preliminary quantitative results suggesting that TeleCARE has the potential for a larger-scale study. Future studies of TeleCARE could harness digital recruitment methods for scalability.

Importantly, intervention adaptations had to be made to successfully implement TeleCARE, including both intervention telehealth-specific modifications and the provision of broad technology support. Related to the modifications, previous research emphasizes the importance of safety and risk management, participant privacy, and platform etiquette in telehealth contexts. Telehealth inherently presents unique challenges compared to in-person care, including limited immediate clinical support and barriers to emergency response due to geographic variability and licensure restrictions [46-48]. Presession safety protocols, such as verifying participant location and emergency contacts, are increasingly recommended to mitigate risks during psychological or medical crises [47]. In addition to safety considerations, given evidence that clinical outcomes in telehealth are often comparable to in-person care, researchers emphasize the importance of replicating in-person standards through structured telehealth delivery [49], such as minimizing distractions, limiting movement, and encouraging stationary participation. Privacy concerns also remain central, with recommendations for participants to join from secure, confidential environments to protect sensitive information and foster psychological safety [48,50].

Compared with younger counterparts, older adults tend to have lower digital literacy, which presents a barrier to engaging in telehealth services [51]. Even after adding an initial (before the start of the intervention) technology tutorial session, the majority of study participants (11/17, 65%) required technological support, and though less support was needed after each passing session, some participants required ongoing support throughout.

XSL•FO
RenderX

Digital literacy training programs for older adults have emerged in response to the recognized need to improve access to technology-based services [52-54]. Many studies have found that older adults want to engage with technology and not feel left behind as digital opportunities increase [55]. At study completion, participants reported feeling a sense of accomplishment related to learning to use the necessary technology and increased comfort with technology, highlighting the ability of older adults to engage with technology with individualized and ongoing support.

It is worth noting that half of the caregivers who declined to join the group stated that they did so due to a lack of time, despite the program being less than a 2-month, 90-minute weekly commitment offered via telehealth, thereby minimizing the time needed to travel. Dementia caregivers consistently identify time constraints as a major barrier to participating in interventions [56-58]. Many caregivers report prioritizing the needs of the care recipient over their own and neglecting self-care, thereby reducing opportunities to participate in supportive services [59-61]. Recent research suggests that brief, flexible, telehealth-based interventions may help boost caregiver engagement [62]. The current findings suggest that offering interventions via telehealth may be only one small step toward improving intervention engagement and that brevity and flexibility may play a bigger role in helping to engage busy caregivers. In addition, finding innovative ways to increase respite time for caregivers so they can engage in supportive services offers promise. For example, Iacob and colleagues [63] demonstrated that an online scheduling and planning program with virtual coaching and education helped increase caregiver respite [63].

Outcome Trends

Although the study was not powered to statistically test changes between pre- and posttreatment, quantitative results suggested trends toward a decrease in NPS severity in care recipients, a decrease in caregiver depression and anxiety, and an increase in caregiver-expressed PAC and meaning and purpose in life after intervention completion. However, most improvements were not sustained by the 3-month follow-up. Also of note is the reduction in caregiving self-efficacy following intervention engagement, which may be unexpected. However, when learning new skills, individuals often experience initial failures early in the learning process. Although these failures are part of the growth necessary to learn new skills, failures can lead to temporary reductions in self-efficacy [64]. It is possible that caregivers in TeleCARE experienced reduced self-efficacy related to their initial unsuccessful efforts applying new skills. Caregivers may also have attributed their successes largely to the presence of the interventionist, causing self-efficacy to decline postintervention when transitioning to applying skills independently. In addition, skills may decay between the conclusion of an intervention and longer-term follow-up (in the current study, 3 mo), which can lead to further reductions in self-efficacy [65,66]. Future studies aimed at better understanding the conditions for gains in self-efficacy and sustainment of those gains are warranted. In addition, future trials of TeleCARE could incorporate theoretically driven self-efficacy enhancement techniques based on Bandura's four identified self-efficacy enhancing methods (mastery experience, modeling, social persuasion, altering emotional/somatic states) [67], while retaining the basic intervention elements.

Limitations and Future Directions

The major limitation of the current study is the homogeneous sample, which reduces generalizability. All of the participants in this study were white, mostly female, and primarily spouses of the individuals receiving care, with at least some college education. This fact limits the applicability of our findings to caregivers from diverse ethnic, cultural, and socioeconomic backgrounds.

Studies have found that the perception of caregiver differs for different cultures and, for those with more deprived socioeconomic disadvantage, knowledge about dementia may be low [68-70]. Intervention adaptations that take into account ethnic or cultural considerations and differences related to medical privacy, preferences for keeping the person with dementia at home, the role of religion, and the use of holistic or non-Western medicine approaches to care would be important to consider in future trials of TeleCARE [71-73]. TeleCARE may need to be adapted for caregivers with lower levels of education and lower income associated with limited or inadequate health literacy [74,75]. Of course, adapting the intervention for non-English speaking populations would improve reach. The original CARE intervention was successfully culturally and linguistically tailored for a Latino population and could serve as a basis for a TeleCARE adaptation in future trials [76]. Importantly, and aligned with the focus of telehealth adaptations, racial and ethnic minority caregivers and those with lower educational attainment are less likely to have access to technology (eg, stable internet connectivity, computer, smartphone) and the skills needed to engage with technology, and specific efforts to explore the use of telehealth for diverse cohorts of dementia caregivers is critical [77-80]. Building community partnerships for intervention adaptation, outreach, and recruitment can help strengthen future trials.

The predominantly female participants in the current intervention limit our ability to fully capture male caregivers' experiences. Male caregivers may differ in their use of coping strategies, emotional responses, and support-seeking behaviors, with literature suggesting a greater tendency to use problem-focused strategies, seek help earlier, and prioritize having time for themselves, engaging in personal activities that may buffer caregiving stress [81,82]. As caregiving intensity increases, male caregivers may experience higher burden than their female counterparts, who often develop caregiving-related stress earlier in the process [81-83]. Thus, interventions that have predominantly female participants may differ significantly from a predominantly male group. The current study also included predominantly spousal caregivers. Adult child caregivers face distinct challenges, including the need to balance caregiving responsibilities with employment and parenting obligations [84,85]. Future research should continue exploring these subgroup differences to tailor interventions accordingly.

In summary, we successfully adapted an existing evidence-based, in-person, manualized dementia caregiver intervention (CARE) for synchronous video telehealth

XSL•FO
RenderX

demonstrating good feasibility and acceptability. Successful adaptation required revising the manualized protocol to include procedures to ensure participant safety, etiquette instructions to improve group experience, and technological support to increase accessibility. These findings suggest that when attempting to transition an in-person program to video telehealth, these broad program-nonspecific adaptations should be considered. The important need for ongoing, available technological assistance was particularly salient.

Funding

This study was funded by the Alzheimer's Association Research Grant (AARG; number AARG-20 - 685186; SDS, principal investigator). The study was supported by resources and the use of facilities at the VA Bedford Health Care System and New England Geriatric Research, Education, and Clinical Center. The contents of this paper do not represent the views of the Department of Veterans Affairs or the United States government.

Data Availability

The dataset is available upon reasonable request to the corresponding author (MKO). Due to privacy, confidentiality, and security concerns, data requestors will need to sign a data access agreement to gain access.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Standard deviations and effect sizes of changes in clinical and psychological measures.
[DOCX File, 20 KB - aging_v9i1e81256_app1.docx]

References

1. Kales HC, Gitlin LN, Lyketsos CG, Detroit Expert Panel on Assessment and Management of Neuropsychiatric Symptoms of Dementia. Management of neuropsychiatric symptoms of dementia in clinical settings: recommendations from a multidisciplinary expert panel. J Am Geriatr Soc 2014 Apr;62(4):762-769. [doi: 10.1111/jgs.12730] [Medline: 24635665]
2. Devanand DP, Lee S, Huey ED, Goldberg TE. Associations between neuropsychiatric symptoms and neuropathological diagnoses of Alzheimer disease and related dementias. JAMA Psychiatry 2022 Apr 1;79(4):359-367. [doi: 10.1001/jamapsychiatry.2021.4363] [Medline: 35171235]
3. García-Martín V, de Hoyos-Alonso MC, Ariza-Cardiel G, et al. Neuropsychiatric symptoms and subsyndromes in patients with different stages of dementia in primary care follow-up (NeDEM project): a cross-sectional study. BMC Geriatr 2022 Jan 22;22(1):71. [doi: 10.1186/s12877-022-02762-9] [Medline: 35065615]
4. Majer R, Adeyi O, Bagoly Z, et al. Neuropsychiatric symptoms, quality of life and caregivers' burden in dementia. Open Med (Wars) 2020;15(1):905-914. [doi: 10.1515/med-2020-0124] [Medline: 33336048]
5. 2025 Alzheimer's disease facts and figures. Alzheimers Dement 2025 Apr;21(4):e70235. [doi: 10.1002/alz.70235]
6. Brodaty H, Connors MH, Xu J, Woodward M, Ames D, PRIME study group. Predictors of institutionalization in dementia: a three year longitudinal study. J Alzheimers Dis 2014;40(1):221-226. [doi: 10.3233/JAD-131850] [Medline: 24448780]
7. Gallagher D, Ni Mhaolain A, Crosby L, et al. Determinants of the desire to institutionalize in Alzheimer's caregivers. Am J Alzheimers Dis Other Demen 2011 May;26(3):205-211. [doi: 10.1177/1533317511400307] [Medline: 21378063]
8. Okura T, Plassman BL, Steffens DC, Llewellyn DJ, Potter GG, Langa KM. Neuropsychiatric symptoms and the risk of institutionalization and death: the aging, demographics, and memory study. J Am Geriatr Soc 2011 Mar;59(3):473-481. [doi: 10.1111/j.1532-5415.2011.03314.x] [Medline: 21391937]
9. Spiegl K, Luttenberger K, Graessel E, Becker L, Scheel J, Pendergrass A. Predictors of institutionalization in users of day care facilities with mild cognitive impairment to moderate dementia. BMC Health Serv Res 2021 Sep 24;21(1):1009. [doi: 10.1186/s12913-021-07017-8] [Medline: 34556085]
10. Cerejeira J, Lagarto L, Mukaetova-Ladinska EB. Behavioral and psychological symptoms of dementia. Front Neurol 2012;3:73. [doi: 10.3389/fneur.2012.00073] [Medline: 22586419]
11. Schulz R, Burgio L, Burns R, et al. Resources for Enhancing Alzheimer's Caregiver Health (REACH): overview, site-specific outcomes, and future directions. Gerontologist 2003 Aug;43(4):514-520. [doi: 10.1093/geront/43.4.514] [Medline: 12937330]
12. Basu R, Hochhalter AK, Stevens AB. The impact of the REACH II intervention on caregivers' perceived health. J Appl Gerontol 2015 Aug;34(5):590-608. [doi: 10.1177/0733464813499640] [Medline: 24652899]
13. Belle SH, Burgio L, Burns R, et al. Enhancing the quality of life of dementia caregivers from different ethnic or racial groups: a randomized, controlled trial. Ann Intern Med 2006 Nov 21;145(10):727-738. [doi: 10.7326/0003-4819-145-10-200611210-00005] [Medline: 17116917]

XSL•FO
RenderX

14. Elliott AF, Burgio LD, Decoster J. Enhancing caregiver health: findings from the resources for enhancing Alzheimer's caregiver health II intervention. J Am Geriatr Soc 2010 Jan;58(1):30-37. [doi: 10.1111/j.1532-5415.2009.02631.x] [Medline: 20122038]
15. Gonyea JG, O'Connor MK, Boyle PA. Project CARE: a randomized controlled trial of a behavioral intervention group for Alzheimer's disease caregivers. Gerontologist 2006 Dec;46(6):827-832. [doi: 10.1093/geront/46.6.827] [Medline: 17169938]
16. Potter AJ. Factors associated with caregivers' use of support services and caregivers' nonuse of services sought. J Aging Soc Policy 2018;30(2):155-172. [doi: 10.1080/08959420.2017.1414539] [Medline: 29293072]
17. Lee Y, Choi W, Park MS. Respite service use among dementia and nondementia caregivers: findings from the National Caregiving in the U.S. 2015 survey. J Appl Gerontol 2022 Jun;41(6):1557-1567. [doi: 10.1177/07334648221075620] [Medline: 35303780]
18. Innes A, Morgan D, Kosteniuk J. Dementia care in rural and remote settings: a systematic review of informal/family caregiving. Maturitas 2011 Jan;68(1):34-46. [doi: 10.1016/j.maturitas.2010.10.002] [Medline: 21093996]
19. Appleman ER, O'Connor MK, Boucher SJ, et al. Teleneuropsychology clinic development and patient satisfaction. Clin Neuropsychol 2021 May;35(4):819-837. [doi: 10.1080/13854046.2020.1871515] [Medline: 33504268]
20. Lyu M, Zhao Q, Yang Y, Hao X, Qin Y, Li K. Benefits of and barriers to telehealth for the informal caregivers of elderly individuals in rural areas: a scoping review. Aust J Rural Health 2022 Aug;30(4):442-457. [doi: 10.1111/ajr.12869] [Medline: 35460580]
21. Borges do Nascimento IJ, Abdulazeem HM, Weerasekara I, et al. An overview of reviews on telemedicine and telehealth in dementia care: mixed methods synthesis. JMIR Ment Health 2025 Nov 6;12:e75266. [doi: 10.2196/75266] [Medline: 41194522]
22. Cuffaro L, Di Lorenzo F, Bonavita S, Tedeschi G, Leocani L, Lavorgna L. Dementia care and COVID-19 pandemic: a necessary digital revolution. Neurol Sci 2020 Aug;41(8):1977-1979. [doi: 10.1007/s10072-020-04512-4] [Medline: 32556746]
23. Boudreau JH, Moo LR, Kennedy MA, et al. Needs for successful engagement in telemedicine among rural older US Veterans and their caregivers: qualitative study. JMIR Form Res 2024 May 7;8:e50507. [doi: 10.2196/50507] [Medline: 38713503]
24. Beauchamp N, Irvine AB, Seeley J, Johnson B. Worksite-based internet multimedia program for family caregivers of persons with dementia. Gerontologist 2005 Dec;45(6):793-801. [doi: 10.1093/geront/45.6.793] [Medline: 16326661]
25. Gustafson DH, Gustafson DH, Cody OJ, Chih MY, Johnston DC, Asthana S. Pilot test of a computer-based system to help family caregivers of dementia patients. J Alzheimers Dis 2019;70(2):541-552. [doi: 10.3233/JAD-190052] [Medline: 31256126]
26. Núñez-Naveira L, Alonso-Búa B, de Labra C, et al. UnderstAID, an ICT platform to help informal caregivers of people with dementia: a pilot randomized controlled study. Biomed Res Int 2016;2016(1):5726465. [doi: 10.1155/2016/5726465] [Medline: 28116300]
27. Van Mierlo LD, Meiland FJM, Van de Ven PM, Van Hout HPJ, Dröes RM. Evaluation of DEM-DISC, customized e-advice on health and social support services for informal carers and case managers of people with dementia; a cluster randomized trial. Int Psychogeriatr 2015 Aug;27(8):1365-1378. [doi: 10.1017/S1041610215000423] [Medline: 25872457]
28. Leng M, Zhao Y, Xiao H, Li C, Wang Z. Internet-based supportive interventions for family caregivers of people with dementia: systematic review and meta-analysis. J Med Internet Res 2020 Sep 9;22(9):e19468. [doi: 10.2196/19468] [Medline: 32902388]
29. Liang J, Aranda MP. The use of telehealth among people living with dementia-caregiver dyads during the COVID-19 pandemic: scoping review. J Med Internet Res 2023 May 25;25:e45045. [doi: 10.2196/45045] [Medline: 37227755]
30. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022 Jan 1;60(1):95-103. [doi: 10.1097/MLR.0000000000001664] [Medline: 34812790]
31. Montgomery R. Sample size justification in feasibility studies: moving beyond published guidance. Pilot Feasibility Stud 2025 Jun 23;11(1):88. [doi: 10.1186/s40814-025-01675-9] [Medline: 40551266]
32. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24;355:i5239. [doi: 10.1136/bmj.i5239] [Medline: 27777223]
33. Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. Pilot Feasibility Stud 2019;5(1):114. [doi: 10.1186/s40814-019-0499-1] [Medline: 31608150]
34. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993 Nov;43(11):2412. [doi: 10.1212/WNL.43.11.2412-a] [Medline: 8232972]
35. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [doi: 10.1093/geront/9.3_Part_1.179] [Medline: 5349366]
36. Kaufer DI, Cummings JL, Ketchel P, et al. Validation of the NPI-Q, a brief clinical form of the neuropsychiatric inventory. J Neuropsychiatry Clin Neurosci 2000;12(2):233-239. [doi: 10.1176/jnp.12.2.233] [Medline: 11001602]
37. Beck AT, Steer RA, Brown G. Beck Depression Inventory–II (BDI-II): Psychological Corporation; 1996. URL: https://books.google.co.in/books/about/BDI_II_Beck_Depression_Inventory.html?id=Ka1wAAAACAAJ&redir_esc=y [accessed 2026-02-21]

XSL•FO
RenderX

38. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988;56(6):893-897. [doi: 10.1037/0022-006X.56.6.893] [Medline: 3204199]
39. Steffen AM, McKibbin C, Zeiss AM, Gallagher-Thompson D, Bandura A. The revised scale for caregiving self-efficacy: reliability and validity studies. J Gerontol B Psychol Sci Soc Sci 2002 Jan;57(1):74-86. [doi: 10.1093/geronb/57.1.p74] [Medline: 11773226]
40. Tarlow BJ, Wisniewski SR, Belle SH, et al. Positive aspects of caregiving: contributions of the REACH project to the development of new measures for Alzheimer's caregiving. Res Aging 2004;26(4):429-453. [doi: 10.1177/0164027504264493]
41. Ulstein I, Wyller TB, Engedal K. High score on the Relative Stress Scale, a marker of possible psychiatric disorder in family carers of patients with dementia. Int J Geriatr Psychiatry 2007 Mar;22(3):195-202. [doi: 10.1002/gps.1660] [Medline: 17044130]
42. Steger MF, Frazier P, Oishi S, Kaler M. The meaning in life questionnaire: assessing the presence of and search for meaning in life. J Couns Psychol 2006;53(1):80-93. [doi: 10.1037/0022-0167.53.1.80]
43. Attkisson CC, Greenfield TK. The Client Satisfaction Questionnaire (CSQ) scales and the Service Satisfaction Scale-30 (SSS-30). In: Sederer LL, Dickey B, editors. Outcome Assessment in Clinical Practice: Baltimore; 1996. URL: https://www.researchgate.net/publication/265377848_The_Client_Satisfaction_Questionnaire_CSQ_Scales [accessed 2025-02-12]
44. O'Bryant SE, Waring SC, Cullum CM, et al. Staging dementia using Clinical Dementia Rating Scale Sum of Boxes scores: a Texas Alzheimer's research consortium study. Arch Neurol 2008 Aug;65(8):1091-1095. [doi: 10.1001/archneur.65.8.1091] [Medline: 18695059]
45. Smith D, Roche E, O'Loughlin K, et al. Satisfaction with services following voluntary and involuntary admission. J Ment Health 2014 Feb;23(1):38-45. [doi: 10.3109/09638237.2013.841864] [Medline: 24484191]
46. Kobeissi MM, Hickey JV. An infrastructure to provide safer, higher-quality, and more equitable telehealth. Jt Comm J Qual Patient Saf 2023 Apr;49(4):213-222. [doi: 10.1016/j.jcjq.2023.01.006] [Medline: 36775714]
47. Martiniuk A, Toepfer A, Lane-Brown A. A review of risks, adverse effects and mitigation strategies when delivering mental health services using telehealth. J Ment Health 2024 Jun;33(3):415-438. [doi: 10.1080/09638237.2023.2182422] [Medline: 36866784]
48. Shachar C, Engel J, Elwyn G. Implications for telehealth in a postpandemic future: regulatory and privacy issues. JAMA 2020 Jun 16;323(23):2375-2376. [doi: 10.1001/jama.2020.7943] [Medline: 32421170]
49. Beech EH, Young S, Anderson JK, Belsher BE, Parr NJ. Evidence brief: safety and effectiveness of telehealthdelivered mental health care. : U. S. Department of Veterans Affairs; 2022 Oct URL: https://www.ncbi.nlm.nih.gov/books/NBK586283/pdf/Bookshelf_NBK586283.pdf [accessed 2026-02-12]
50. Houser SH, Flite CA, Foster SL. Privacy and security risk factors related to telehealth services – a systematic review. Perspect Health Inf Manag 2023;20(1):1f. [Medline: 37215337]
51. van Deursen AJ. Digital inequality during a pandemic: quantitative study of differences in COVID-19-related internet uses and outcomes among the general population. J Med Internet Res 2020 Aug 20;22(8):e20073. [doi: 10.2196/20073] [Medline: 32750005]
52. Vidal E. Digital literacy program: reducing the digital gap of the elderly: experiences and lessons learned. 2019 Presented at: 2019 International Conference on Inclusive Technologies and Education (CONTIE); Oct 30 to Nov 1, 2019. [doi: 10.1109/CONTIE49246.2019.00030]
53. Steelman K, Wallace C. Breaking barriers, building understanding: a multigenerational approach to digital literacy instruction for older adults. SIGACCESS Access Comput 2017;118:9-15. [doi: 10.1145/3124144.3124146]
54. Choudhary H, Bansal N. Addressing digital divide through digital literacy training programs: a systematic literature review. Digital Educ Rev 2022(41):224-248. [doi: 10.1344/der.2022.41.224-248]
55. Fields J, Cemballi AG, Michalec C, et al. In-home technology training among socially isolated older adults: findings from the Tech Allies Program. J Appl Gerontol 2021 May;40(5):489-499. [doi: 10.1177/0733464820910028] [Medline: 32141373]
56. Sorrentino M, Fiorilla C, Mercogliano M, et al. Barriers for access and utilization of dementia care services in Europe: a systematic review. BMC Geriatr 2025 Mar 10;25(1):162. [doi: 10.1186/s12877-025-05805-z] [Medline: 40065204]
57. Joshi S, Park T, Brody L, et al. Recruitment of family caregivers of persons with dementia: lessons learned from a pilot randomized controlled trial. Front Pain Res (Lausanne) 2023;4:1125914. [doi: 10.3389/fpain.2023.1125914] [Medline: 37051457]
58. Lech S, O'Sullivan JL, Wellmann L, et al. Recruiting general practitioners and patients with dementia into a cluster randomised controlled trial: strategies, barriers and facilitators. BMC Med Res Methodol 2021 Mar 30;21(1):61. [doi: 10.1186/s12874-021-01253-6] [Medline: 33784967]
59. Lin X, Moxley JH, Czaja SJ. Caring for dementia caregivers: psychosocial factors related to engagement in self-care activities. Behav Sci (Basel) 2023 Oct 18;13(10):851. [doi: 10.3390/bs13100851] [Medline: 37887501]
60. Waligora KJ, Bahouth MN, Han HR. The self-care needs and behaviors of dementia informal caregivers: a systematic review. Gerontologist 2019 Sep 17;59(5):e565-e583. [doi: 10.1093/geront/gny076] [Medline: 29931147]
61. Oliveira D, Zarit SH, Orrell M. Health-promoting self-care in family caregivers of people with dementia: the views of multiple stakeholders. Gerontologist 2019 Sep 17;59(5):e501-e511. [doi: 10.1093/geront/gnz029] [Medline: 30953585]

XSL•FO
RenderX

62. Farnood A, Johnston B, Evans C. The role of digital health interventions in supporting family caregivers for people with dementia to improve quality of life. Curr Opin Support Palliat Care 2024 Dec 1;18(4):224-229. [doi: 10.1097/SPC.0000000000000734] [Medline: 39373310]
63. Iacob E, Caserta M, Donaldson G, et al. Evaluating the efficacy of time for living and caring: an online intervention to support dementia caregivers' use of respite. Innov Aging 2024;8(5):igae043. [doi: 10.1093/geroni/igae043] [Medline: 38803611]
64. Bernacki ML, Nokes-Malach TJ, Aleven V. Examining self-efficacy during learning: variability and relations to behavior, performance, and learning. Metacognition Learn 2015 Apr;10(1):99-117. [doi: 10.1007/s11409-014-9127-x]
65. DiZazzo-Miller R, Pociask FD, Adamo DE. The role of confidence in family caregiving for people with dementia. Phys Occup Ther Geriatr 2020;38(4):355-369. [doi: 10.1080/02703181.2020.1762824]
66. Lewis ML, Hobday JV, Hepburn KW. Internet-based program for dementia caregivers. Am J Alzheimers Dis Other Demen 2010 Dec;25(8):674-679. [doi: 10.1177/1533317510385812] [Medline: 21131674]
67. Bandura A. Social Foundations of Thought and Action: Prentice-Hall; 1986. URL: https://agris.fao.org/search/en/records/65ddef380f3e94b9e5c79ec7 [accessed 2026-02-10]
68. Connell CM, Scott Roberts J, McLaughlin SJ, Akinleye D. Racial differences in knowledge and beliefs about Alzheimer disease. Alzheimer Dis Assoc Disord 2009;23(2):110-116. [doi: 10.1097/WAD.0b013e318192e94d] [Medline: 19474569]
69. Cahill S, Pierce M, Werner P, Darley A, Bobersky A. A systematic review of the public's knowledge and understanding of Alzheimer's disease and dementia. Alzheimer Dis Assoc Disord 2015;29(3):255-275. [doi: 10.1097/WAD.0000000000000102] [Medline: 26207322]
70. Jang Y, Yoon H, Park NS, Rhee MK, Chiriboga DA. Asian Americans' concerns and plans about Alzheimer's disease: the role of exposure, literacy and cultural beliefs. Health Soc Care Community 2018 Mar;26(2):199-206. [doi: 10.1111/hsc.12509] [Medline: 28980423]
71. Cothran FA, Paun O, Strayhorn S, Barnes LL. "Walk a mile in my shoes": African American caregiver perceptions of caregiving and self-care. Ethn Health 2022 Feb;27(2):435-452. [doi: 10.1080/13557858.2020.1734777] [Medline: 32116006]
72. Samson ZB, Parker M, Dye C, Hepburn K. Experiences and learning needs of African American family dementia caregivers. Am J Alzheimers Dis Other Demen 2016 Sep;31(6):492-501. [doi: 10.1177/1533317516628518] [Medline: 26953236]
73. Huggins LKL, Min SH, Dennis CA, Østbye T, Johnson KS, Xu H. Interventions to promote dementia knowledge among racial/ethnic minority groups: a systematic review. J Am Geriatr Soc 2022 Feb;70(2):609-621. [doi: 10.1111/jgs.17495] [Medline: 34628641]
74. Xu H, Kadambi S, Mohile SG, et al. Caregiving burden of informal caregivers of older adults with advanced cancer: the effects of rurality and education. J Geriatr Oncol 2021 Sep;12(7):1015-1021. [doi: 10.1016/j.jgo.2021.04.002] [Medline: 33858803]
75. Lorini C, Buscemi P, Mossello E, et al. Health literacy of informal caregivers of older adults with dementia: results from a cross-sectional study conducted in Florence (Italy). Aging Clin Exp Res 2023 Jan;35(1):61-71. [doi: 10.1007/s40520-022-02271-0] [Medline: 36260214]
76. Gonyea JG, López LM, Velásquez EH. The effectiveness of a culturally sensitive cognitive behavioral group intervention for Latino Alzheimer's caregivers. Gerontologist 2016 Apr;56(2):292-302. [doi: 10.1093/geront/gnu045] [Medline: 24855313]
77. Singh GK, Girmay M, Allender M, Ramey CT. Digital divide: marked disparities in computer and broadband internet use and associated health inequalities in the United States. Int J Transl Med Res Public Health 2020;4:1-16. [doi: 10.21106/ijtmrph.148]
78. Marston HR, Shore L, White PJ. How does a (smart) age-friendly ecosystem look in a post-pandemic society? Int J Environ Res Public Health 2020 Nov 9;17(21):8276. [doi: 10.3390/ijerph17218276] [Medline: 33182413]
79. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: Internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013 May 2;15(5):e93. [doi: 10.2196/jmir.2645] [Medline: 23639979]
80. Mitchell UA, Chebli PG, Ruggiero L, Muramatsu N. The digital divide in health-related technology use: the significance of race/ethnicity. Gerontologist 2019 Jan 9;59(1):6-14. [doi: 10.1093/geront/gny138] [Medline: 30452660]
81. Xiong C, Biscardi M, Astell A, et al. Sex and gender differences in caregiving burden experienced by family caregivers of persons with dementia: a systematic review. PLoS One 2020;15(4):e0231848. [doi: 10.1371/journal.pone.0231848] [Medline: 32310969]
82. Swinkels J, Tilburg TV, Verbakel E, Broese van Groenou M. Explaining the gender gap in the caregiving burden of partner caregivers. J Gerontol B Psychol Sci Soc Sci 2019 Jan 10;74(2):309-317. [doi: 10.1093/geronb/gbx036] [Medline: 28379560]
83. Pöysti MM, Laakkonen ML, Strandberg T, et al. Gender differences in dementia spousal caregiving. Int J Alzheimers Dis 2012;2012(1):162960. [doi: 10.1155/2012/162960] [Medline: 23056990]
84. Rigby T, Ashwill RT, Johnson DK, Galvin JE. Differences in the experience of caregiving between spouse and adult child caregivers in dementia with lewy bodies. Innov Aging 2019;3(3):igz027. [doi: 10.1093/geroni/igz027] [Medline: 31528714]

XSL•FO
RenderX

85. Oldenkamp M, Hagedoorn M, Slaets J, Stolk R, Wittek R, Smidt N. Subjective burden among spousal and adult-child informal caregivers of older adults: results from a longitudinal cohort study. BMC Geriatr 2016 Dec 7;16(1):208. [doi: 10.1186/s12877-016-0387-y] [Medline: 27923347]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
CSQ-8: Client Satisfaction Questionnaire
IADL: Instrumental Activities of Daily Living
NPS: neuropsychiatric symptoms
PAC: positive aspects of caregiving
REACH: resources for enhancing Alzheimer's caregiver health
REDCap: Research Electronic Data Capture
VA: Veterans Affairs

Edited by M O'Connell; submitted 24.Jul.2025; peer-reviewed by C Obasi, N Kaufman; revised version received 16.Dec.2025; accepted 24.Jan.2026; published 12.Mar.2026.

Please cite as:
O'Connor MK, Shirk SD, McLaren JE, Nguyen AH, Pugh K, Sullivan MA, Metcalf EE, Harrington S, Moo LR
A Telehealth-Adapted Dementia Caregiver Skills Training Intervention (TeleCARE): Single-Arm Pre-Post Intervention Study
JMIR Aging 2026;9:e81256
URL: https://aging.jmir.org/2026/1/e81256
doi:10.2196/81256

© Maureen K O'Connor, Steven D Shirk, Jaye E McLaren, Andrew H Nguyen, Kendra Pugh, Madeline A Sullivan, Emily E Metcalf, Samantha Harrington, Lauren R Moo. Originally published in JMIR Aging (https://aging.jmir.org), 12.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Developing and Testing a Brief Mindfulness Just-in-Time Adaptive Intervention to Reduce Stress Among Caregivers of People With Dementia: Quasi-Experimental Study

Patrick Pui Kin Kor[1], PhD; Alex Pak Lik Tsang[1], PhD; Daphne Sze Ki Cheung[2,3], PhD; Steven H Zarit[4], PhD; Haoran Xie[5], PhD; Min Qian[6], PhD; Kay Chen Tan[7], PhD; Yuanqing Zheng[7], PhD; Amanda Man Ying Chu[8], PhD; Kee Lee Chou[8], PhD

[1]School of Nursing, Faculty of Health and Social Sciences, Hong Kong Polytechnic University, Hung Hom, Kowloon, China (Hong Kong)

[2]School of Nursing and Midwifery, Faculty of Health, Deakin University, Melbourne, Victoria, Australia

[3]Centre for Quality and Patient Safety Research/Alfred Health Partnership, Institute for Health Transformation, Deakin University, Melbourne, Victoria, Australia

[4]Department of Human Development and Family Studies, Pennsylvania State University, University Park, PA, United States

[5]Division of Artificial Intelligence, School of Data Science, Lingnan University, Tuen Mun, New Territories, China (Hong Kong)

[6]Mailman School of Public Health, Columbia University, New York, NY, United States

[7]Department of Data Science and Artificial Intelligence, Faculty of Computer and Mathematical Sciences, Hong Kong Polytechnic University, Hung Hom, Kowloon, China (Hong Kong)

[8]Department of Social Sciences and Policy Studies, Faculty of Liberal Arts and Social Sciences, Education University of Hong Kong, 10 Lo Ping Road, Tai Po, New Territories, China (Hong Kong)

Corresponding Author:
Kee Lee Chou, PhD
Department of Social Sciences and Policy Studies, Faculty of Liberal Arts and Social Sciences, Education University of Hong Kong, 10 Lo Ping Road, Tai Po, New Territories, China (Hong Kong)

Abstract

Background: Dementia caregiving entails chronic, fluctuating stress with downstream risks to caregivers' mental health and quality of care. Mindfulness-based interventions can reduce caregiver stress; however, moment-to-moment fluctuations in stress may limit receptivity to practice at any given time. We developed a brief mindfulness just-in-time adaptive intervention (JITAI) that aims to deliver support at the right moment by using machine learning algorithms to optimize notification timing based on receptivity to engage in brief mindfulness practices.

Objective: This study aims to evaluate the feasibility, acceptability, and effectiveness of a brief mindfulness JITAI for caregivers of people with dementia on stress, depressive symptoms, caregiver burden, sleep, quality of life, and trait mindfulness.

Methods: A single-arm, pretest or posttest design was adopted. A total of 120 community-dwelling caregivers were recruited to participate in the 18-day intervention, which included 4 days of psychoeducation delivered via videos and phone coaching, alongside an in-app brief, low-dose mindfulness-based stress reduction component. From days 5 to 11, prompts were delivered either by a static machine learning model trained on prior pilot data or at random times, with equal probability. From days 12 to 18, three delivery models were used with equal probability, namely static, random, and adaptive models, which updated per participant using accumulating receptivity data. Feasibility and acceptability were assessed post intervention. Standardized measures of stress, depressive symptoms, caregiver burden, positive aspects of caregiving, sleep, quality of life, and trait mindfulness were collected via phone interviews at baseline and post intervention.

Results: Retention was 100%. Most participants (111/120, 92.5%) found the app easy to use, 81.7% (98/120) perceived it as helpful for stress management, and 80% (96/120) would recommend it to other caregivers. Pre-post analyses indicated significant reductions in perceived stress (P<.001), depressive symptoms (P<.001), and caregiver burden (P=.003), as well as a significant increase in positive aspects of caregiving (P<.001) and subjective sleep quality (P=.02). Health-related quality of life and trait mindfulness did not change significantly.

Conclusions: A brief, smartphone-delivered mindfulness JITAI for caregivers of people with dementia was feasible and acceptable, with high retention and positive user evaluations. Pre-post findings suggest reductions in perceived stress, depressive symptoms, and caregiving burden, alongside increased positive aspects of caregiving and improved sleep, supporting the potential of adaptive, technology-enabled interventions to provide timely support to caregivers.

XSL•FO
RenderX

Trial Registration: Chinese Clinical Trial Registry ChiCTR2300071361; https://www.chictr.org.cn/hvshowprojectEN.html?id=286697

(JMIR Aging 2026;9:e87316) doi:10.2196/87316

KEYWORDS

dementia caregiving; mindfulness; mobile health; just-in-time adaptive intervention; receptivity; stress; Android app

Introduction

Dementia is a neurodegenerative condition characterized by progressive declines in functional abilities and substantial care needs, with caregiving often provided by family members. With an aging global population, the prevalence of dementia is projected to rise from 55 million in 2019 to 139 million by 2050 [1]. Consequently, the number of family caregivers assuming these demanding roles will increase. The stress associated with dementia caregiving is substantial and can lead to adverse health outcomes for both caregivers and care recipients. For caregivers, prolonged exposure to stress is linked to insomnia, compromised immune function, and psychological distress, including depression and anxiety [2-4]. For care recipients, high levels of caregiver stress can increase the likelihood of premature institutionalization and, in some cases, elder abuse [5,6]. Therefore, developing effective strategies to mitigate caregiver stress is a pressing public health priority.

Since stressors are a normative and unavoidable part of the long-term dementia caregiving trajectory, emotion-focused coping approaches that emphasize the positive reappraisal of caregiving-related emotions, in contrast to problem-focused approaches that target modification of the stressor itself (eg, caregiving skills training), are particularly important for stress management [7]. Mindfulness is one such approach and has been shown to reduce stress among caregivers of people with dementia [8,9]. The mindful coping model emphasizes decentering, defined as the capacity to observe one's thoughts and feelings from a detached perspective [10]. This shift in attentional focus enables caregivers to reframe stressors, engage in positive reappraisal, and ultimately reduce stress. A systematic review has confirmed that mindfulness-based interventions (MBIs) have a moderate effect size in reducing caregiver stress in this population [8].

Despite their effectiveness, MBIs for caregivers of people with dementia are hindered by practical challenges. A primary barrier is the significant time commitment required by traditional program structures. For example, a trial of mindfulness-based cognitive therapy found that about 20% of 113 participants discontinued regular mindfulness practice during follow-up, with lack of time being the most common reason for discontinuing their practice [9]. Caregivers often juggle numerous caregiving tasks and other life commitments, which can disrupt the self-directed practice central to these interventions. The fluctuating nature of caregiving demands and stress can also create episodes of cognitive interference, making it difficult to focus attention during periods of high strain [11]. If an intervention is perceived as too intensive or inflexible, caregivers may be less motivated to engage in self-practice. Therefore, tailoring the delivery of mindfulness practices is crucial to support sustained engagement. To address this challenge, interventions should be delivered at moments of optimal receptivity, defined as states in which a caregiver is able, available, and willing to receive, process, and use support [12].

A just-in-time adaptive intervention (JITAI) can optimize receptivity by delivering support at opportune moments while adapting to a user's changing internal and external contexts [12]. This approach has been applied to a wide range of public health issues, including physical inactivity [13], smoking cessation [14], and obesity [15]. More recently, JITAIs have been integrated with MBIs for individuals with chronic pain, demonstrating reductions in cravings, pain, and stress, along with increases in positive affect [16]. Receptivity in JITAI delivery can be optimized using several models: (1) static, (2) random, and (3) adaptive [17]. A static model delivers prompts at predetermined times, often derived from machine learning analyses of prior user data; however, receptivity patterns from one group may not generalize to others [18]. A random model delivers prompts at random times and often underperforms machine learning–based models [17]. An adaptive model uses machine learning to learn each participant's receptivity patterns as the study progresses. Prior research suggests that although an adaptive model may underperform initially, its performance improves as user-specific data accumulate over time [17].

Taken together, a brief mindfulness-based JITAI may offer a promising approach to the unique challenges of delivering support to this population. By tailoring delivery to moments of high receptivity, a JITAI can facilitate the self-directed practice essential for maximizing the stress-reduction benefits of MBIs [19]. However, there is limited research applying mindfulness-based JITAIs to caregivers of people with dementia. Consequently, their effectiveness for stress reduction, as well as their feasibility and acceptability within this population, remains unclear.

This study aims to address this gap by introducing and evaluating a novel, brief mindfulness JITAI for caregivers of people with dementia using Android phones. The study objectives are (1) to evaluate the feasibility and acceptability of the intervention (eg, retention, adherence or engagement with prompted practice, and user feedback) and (2) to estimate preliminary pre-post changes in caregiver psychosocial outcomes (eg, perceived stress and related outcomes). To operationalize these objectives and manage the initial learning phase of the adaptive model, the program was implemented in the following steps: step 1 (1‐4 d) provided brief psychoeducation and coaching with access to app-based mindfulness content; step 2 (5‐11 d) delivered prompts using either a static model informed by pilot usage data (n=75) or a random model (equal probability) to accumulate initial data;

XSL•FO
RenderX

and step 3 (12‐18 d) introduced an adaptive model (alongside static and random, equal probability) that personalized prompt timing based on continuously collected, participant-specific data.

Methods

Study Design

This is a single-arm, quasi-experimental study. The study is reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) guidelines (Checklist 1). It was conducted in compliance with the Declaration of Helsinki (and its subsequent amendments) and Good Clinical Practice guidelines.

Ethical Considerations

Ethical approval was obtained from the corresponding author's institution (reference: A2021-2022-0276-01). All participants provided written informed consent prior to participation. Participants were informed of the study's purpose, procedures, the voluntary nature of participation, and their right to withdraw at any time without penalty. Privacy and confidentiality were protected by deidentifying participant data and restricting access to research data to authorized members of the research team only. Participants who completed all assessments and attended at least 80% of the intervention sessions received HK $500 (US $64.10) in supermarket coupons as compensation for their time and participation. The trial was registered at the Chinese Clinical Trial Registry (ChiCTR2300071361).

Participants

Participants were family caregivers of people with dementia, recruited through a network of nongovernmental organizations providing community services for older people, caregiver support groups, and the Alzheimer's Association in Hong Kong.

Eligibility criteria included being aged 18 years or older, serving as a family caregiver for a community-dwelling individual with a confirmed medical diagnosis of any type of dementia, providing at least 20 hours per week of assistance with activities of daily living and instrumental activities of daily living for a minimum of 1 year prior to recruitment, having an Android smartphone, and reporting moderate to high stress as measured by the 14-item Perceived Stress Scale (PSS-14) [20]. A cutoff score of 18 was used to identify caregivers with elevated stress (PSS-14>18), consistent with thresholds applied in previous studies [21].

Exclusion criteria included participation in any structured mind-body intervention, cognitive behavioral therapy, or structured psychosocial intervention within the previous 6 months, as well as acute psychiatric or medical comorbidities that were potentially life-threatening (eg, suicidal ideation) or could limit participation or adherence (eg, acute psychosis).

Procedures

Recruitment and Assessment

Interested individuals were approached by a research assistant, who administered a brief screening questionnaire via phone interviews to confirm eligibility. Those who met the inclusion criteria received a standardized study briefing and then provided electronic informed consent via the Qualtrics platform (Qualtrics International Inc). The research team carefully explained the required commitments to ensure that participants were aware of expectations and to minimize attrition during the study period.

Data were collected through structured telephone interviews conducted by a trained research assistant at baseline (T1; the first day of the study period) and post intervention (T2; the last day of the study period).

The Brief Mindfulness JITAI

The brief mindfulness JITAI lasted 18 days. During days 1 to 4, participants received daily 30-minute psychosocial education sessions comprising short videos and phone-based coaching, alongside access to a brief, app-embedded mindfulness program (Figure 1). The study app was distributed to participants as an Android Package Kit for installation and was not publicly available via commercial app stores at the time of the study. Intervention content was developed by a certified mindfulness instructor and drew on "low-dose" mindfulness-based stress reduction (MBSR) elements [22]. MBSR was selected given its demonstrated feasibility, acceptability, and benefits for stress reduction among family caregivers of people with dementia [8,23,24]. Beginning on day 5, participants received push notifications prompting mindfulness practice.

XSL•FO
RenderX

Figure 1. Screenshots of the mobile app. The top-left panel shows the home screen, which greets the user (“您好”) and prompts them to start practice (“開始練習”) for week 0 (“第0週”). The top-right panel displays guided practice for a mindful walking exercise (“靜觀步行練習”), with a progress bar and playback controls. Participants can skip the exercise by pressing “略過.” The bottom-left panel features a check-in function (“每日練習”) that asks users to rate their stress level for the day (“今天你經常感受到的壓力有多大？”). The bottom-right panel showcases the video library (“影片”), which contains modules such as “Communication Skills with Dementia Patients” (“與認知障礙症患者溝通技巧和應對方法”) and “Stress Faced by Caregivers” (“照顧者面對的壓力”).

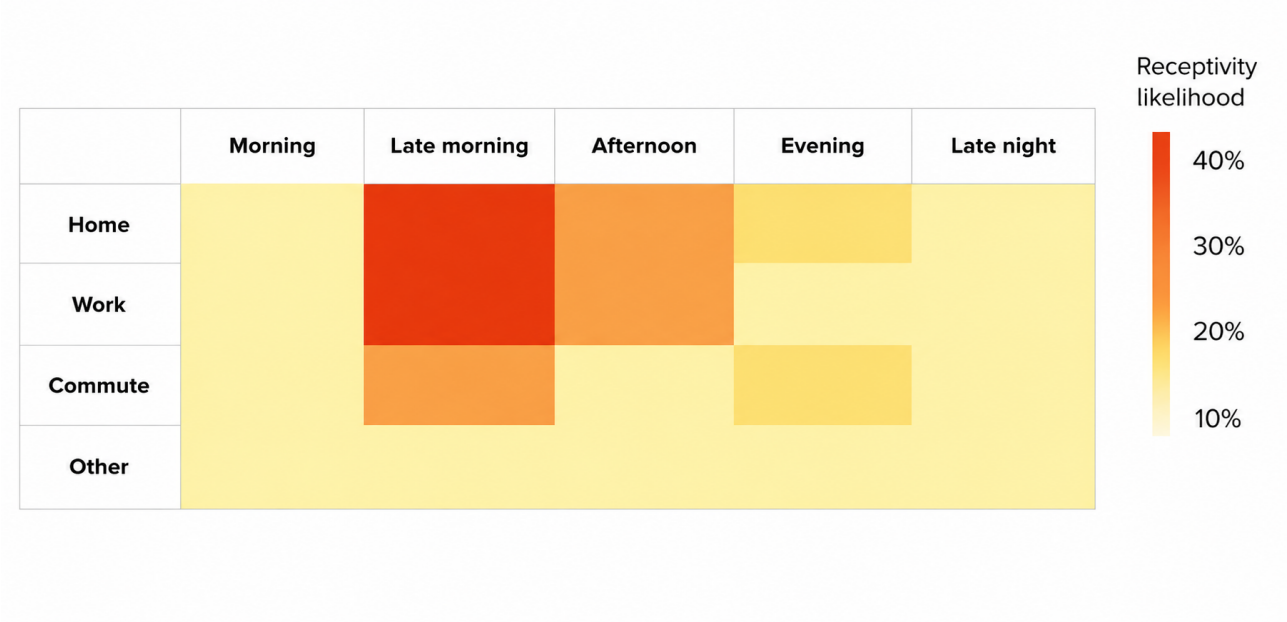


From days 5 to 11, the timing of intervention prompts was determined either by (1) a static machine learning model trained in a pilot study with 75 participants using 4 weeks of passive smartphone sensing data (day of week, time of day, type of day [weekday vs weekend], battery level, and location) and observed receptivity, operationalized as the just-in-time response rate (ie, whether a delivered notification prompted engagement with the mindfulness practice within 10 min of delivery), consistent with definitions used in prior JITAI studies [17] (Figure 2), or (2)

XSL•FO
RenderX

random scheduling with equal probability. The static model remained fixed across all participants and days.

Figure 2. Receptivity likelihood heatmap for training the static model.



From days 12 to 18 (post–warm-up period), an adaptive model was added in addition to the static and random models. The adaptive model updated its predictions as the study progressed using each participant's accumulating observed receptivity data (just-in-time response rate) and the same contextual feature set used in the static model. We introduced the adaptive model after 1 week to ensure that sufficient participant-specific data had been accumulated to inform model predictions.

Decision logic for notification timing was as follows: the static and adaptive models each used a classifier to infer whether the current moment was "receptive" for delivering a prompt; the random scheduler did not screen for receptivity and always approved delivery. If a model judged the current moment as not receptive, the app queried the same model again every 5 minutes. If no receptive moment was inferred within 30 minutes, a notification was sent, and this delivery was recorded under the random model.

Measures

Feasibility Outcomes

Intervention feasibility was assessed using recruitment, attendance, completion, and retention metrics. The recruitment rate was defined as the proportion of eligible individuals who enrolled. The attendance rate reflected the proportion of the 4 psychoeducation sessions attended. The completion rate referred to the proportion of participants who finished all study assessments. The retention rate represented the proportion of enrolled participants who remained throughout the study period.

Acceptability Outcomes

Intervention acceptability was assessed at T2 using a questionnaire with three sections: (1) acceptability of the mobile app, (2) acceptability of notifications, and (3) acceptability of mindfulness activities.

In the app section, participants indicated whether the app was easy to install and use (yes or no), whether they ignored the app's notifications (yes or no), whether the app helped them manage stress (able or helpful or not able or not helpful), and whether they would recommend the app to other caregivers (would or would not).

In the notifications section, participants rated the number of messages (too many, appropriate, or too few) and the helpfulness or appropriateness of their delivery timing (helpful, appropriate, or not helpful). They also reported whether they liked the wording of the reminders (yes or no).

In the mindfulness activities section, participants indicated whether the activities helped them manage stress (able or helpful and not able or not helpful), whether they liked the graphics and characters used (like or do not like), and whether learning mindfulness through these activities was helpful (helpful or not helpful).

Effectiveness Outcomes

The primary outcome was perceived stress, measured using the PSS-14 [20]. Items are rated on a 5-point Likert scale (0="never" to 4="very often"), with higher scores indicating greater perceived stress. Cronbach α for the Chinese PSS-14 was 0.72 at T1 and 0.82 at T2.

Depressive symptoms were assessed using the 9-item Patient Health Questionnaire, which measures the frequency of depressive symptoms over the past week [25]. Items are rated on a 4-point scale (0="not at all" to 3="nearly every day"), with higher scores reflecting greater depression severity. Cronbach α for the Chinese Patient Health Questionnaire-9 was 0.91 at T1 and 0.89 at T2.

Caregiving burden was measured using the 12-item Zarit Burden Interview [26]. Items assess the perceived impacts of caregiving on emotional strain, physical fatigue, role restriction, and social

XSL•FO
RenderX

well-being. Each item is rated on a 5-point scale (0="never" to 4="nearly always"), with higher scores denoting greater burden. Cronbach α for the Chinese 12-item Zarit Burden Interview was 0.86 at T1 and 0.87 at T2.

Positive caregiving appraisal was measured using the Positive Aspects of Caregiving Scale (PAC) [27]. The PAC assesses perceived gains from caregiving, including self-affirmation (a confident, capable self-image derived from the role) and outlook on life (interpersonal relationships and positive life orientation). Items are rated on a 5-point Likert scale (1="strongly disagree" to 5="strongly agree"), with higher scores indicating a more positive appraisal. Cronbach α for the Chinese PAC was 0.91 at both T1 and T2.

Subjective sleep quality and disturbances were assessed using the Pittsburgh Sleep Quality Index [28]. The Pittsburgh Sleep Quality Index aggregates component scores—subjective sleep quality, latency, duration, efficiency, disturbances, and daytime dysfunction—into a global score; higher scores indicate poorer sleep quality. Cronbach α was 0.69 at T1 and 0.67 at T2.

Health-related quality of life was measured using the EQ-5D-5L [29]. The EQ-5D-5L indexes problems across 5 domains (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression), each rated on 5 levels (no, slight, moderate, severe, and extreme problems). A 0 to 100 visual analog scale records self-rated health. Higher scores reflect better health-related quality of life. Cronbach α for the Chinese EQ-5D-5L was 0.71 at T1 and 0.69 at T2.

Trait mindfulness was assessed using the 15-item Mindful Attention Awareness Scale, which measures the frequency of attention lapses and mindlessness in daily life [30]. Items are rated on a 6-point scale (1="almost always" to 6="almost never"), with higher scores indicating greater dispositional mindfulness. Cronbach α for the Chinese Mindful Attention Awareness Scale was 0.91 at T1 and 0.90 at T2.

Sociodemographic and Caregiving Characteristics

Caregivers' sociodemographic characteristics included age, gender (male and female), educational level (primary or below, secondary, and tertiary), economic status (economically inactive and economically active), household size, household income, the use of government or organizational financial assistance (yes or no), and the presence of chronic illness (yes or no) and disability (yes or no). Care recipients' sociodemographic characteristics included age, gender, chronic illness, and disability.

Caregiving characteristics included living arrangement (co-residing with the care recipient: yes or no), sole-caregiver status (yes or no), duration of caregiving (years), and weekly caregiving hours. The quality of the relationship between the caregiver and the care recipient was rated on a 5-point scale from "very good" to "very poor."

Statistical Analysis

Data were analyzed using IBM SPSS version 29. We first examined participants' sociodemographic characteristics, followed by a descriptive assessment of feasibility and acceptability outcomes. Paired-samples *t* tests were then used to compare effectiveness outcomes from T1 to T2. Unless otherwise specified, all tests were 2-tailed with an α level of 5%.

We calculated the sample size using G*Power 3.1 (Heinrich Heine University Düsseldorf). The primary effectiveness outcome was perceived stress, with an effect size of d=0.33 drawn from a prior meta-analysis of MBIs for caregivers of people with dementia [31]. A 2-tailed pre-post analysis with an α level of .05 and 80% power indicated a required sample size of 75 to detect this effect. Allowing for 20% attrition, the target was increased by 15 participants, yielding a total of 90. Recruitment continued as time and resources permitted to further increase statistical power.

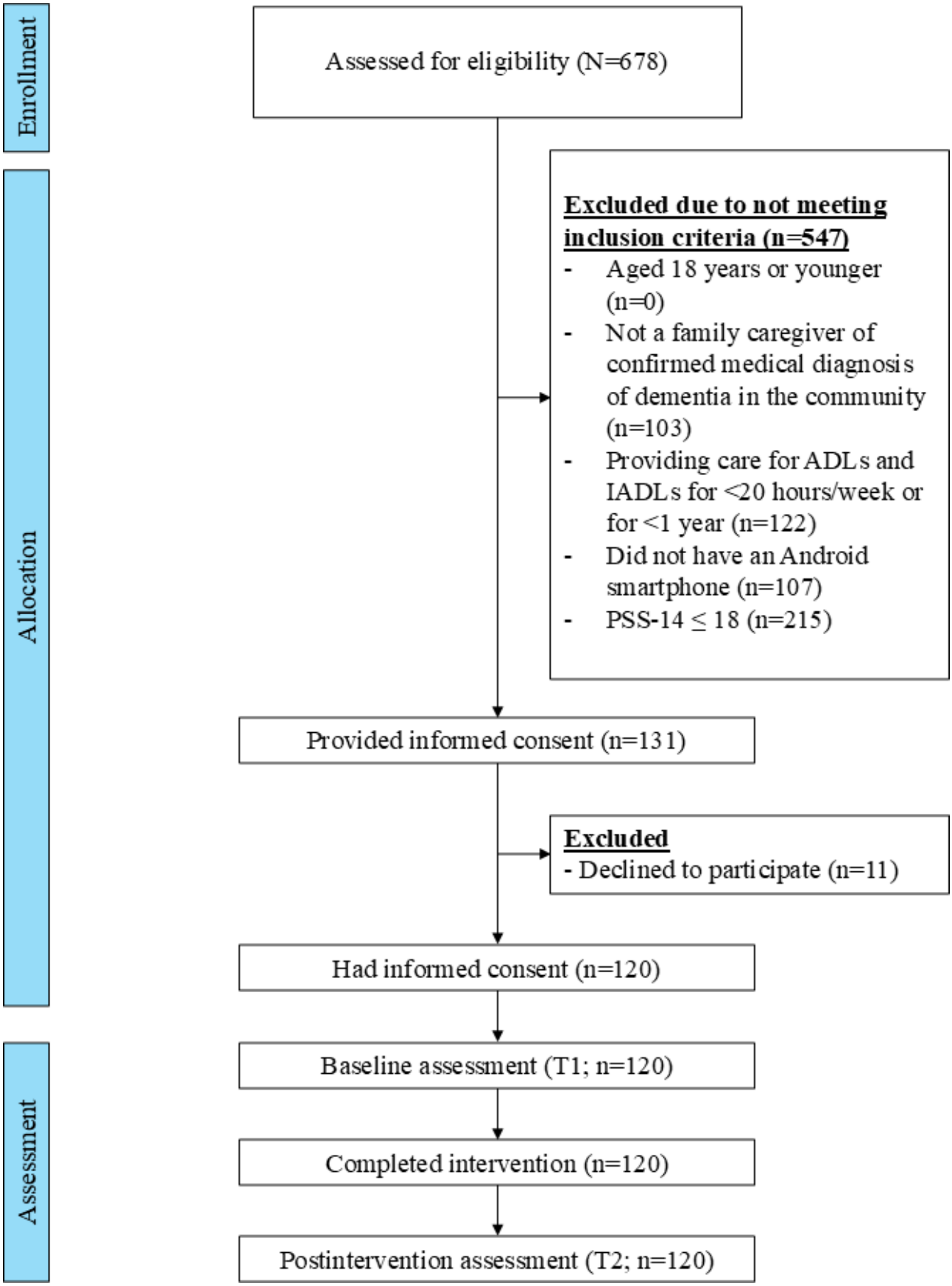
Results

Sociodemographic and Caregiving Characteristics

A total of 120 caregivers participated in the study. A CONSORT flow diagram illustrating the participants' progression is shown in Figure 3. Table 1 presents the participants' sociodemographic and caregiving characteristics.

XSL•FO
RenderX

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. ADL: activities of daily living; IADL: instrumental activities of daily living; PSS-14: 14-item Perceived Stress Scale.



XSL•FO
RenderX

Table . Sociodemographic and caregiving characteristics (N=120).

Characteristics	Values
Caregivers' characteristic	
Age (y), mean (SD)	45.4 (11.8)
Gender, n (%)	
Male	40 (33.3)
Female	80 (66.7)
Educational level[a], n (%)	
Primary or below	13 (10.8)
Secondary	41 (34.2)
Tertiary	66 (55)
Employment status, n (%)	
Economically inactive	26 (21.7)
Economically active	94 (78.3)
Household size, mean (SD)	2.9 (0.9)
Household income (HK $; HK $1=US $0.128), mean (SD)	31,541.70 (20,079.10)
Use of government or organizational financial assistance, n (%)	
Yes	11 (9.2)
No	109 (90.8)
Presence of chronic illness, n (%)	
Yes	12 (10)
No	108 (90)
Presence of disability, n (%)	
Yes	0 (0)
No	120 (100)
Care recipients' characteristics	
Age (y), mean (SD)	76.3 (7.4)
Gender, n (%)	
Male	54 (45)
Female	66 (55)
Presence of chronic illness, n (%)	
Yes	97 (80.8)
No	23 (19.2)
Presence of disability, n (%)	
Yes	20 (16.7)
No	100 (83.3)
Relationship with the caregiver, n (%)	
Mother	55 (45.8)
Father	44 (36.7)
Spouse	9 (7.5)
Paternal grandmother	7 (5.8)
Maternal grandfather	2 (1.7)
Paternal grandfather	1 (0.8)
Mother-in-law	1 (0.8)

XSL•FO
RenderX

Characteristics	Values
Daughter-in-law	1 (0.8)
Caregiving characteristics	
Co-residing with the care recipient, n (%)	
Yes	84 (70)
No	36 (30)
Sole-caregiver status, n (%)	
Yes	43 (35.8)
No	77 (64.2)
Duration of caregiving (y), mean (SD)	5.5 (3.5)
Weekly caregiving hours, mean (SD)	45.9 (20.1)
Relationship with the care recipient, n (%)	
Very bad	0 (0)
Bad	2 (1.7)
Average	39 (32.5)
Good	58 (48.3)
Very good	21 (17.5)

[a]Tertiary education was defined as the completion of any postsecondary program, including subdegree (associate degree or higher diploma), bachelor's degree, or postgraduate qualification.

Caregivers ranged in age from 18 to over 80 years, with a mean age of 45.4 (SD 11.8) years. Most were female (80/120, 66.7%), had completed tertiary education (66/120, 55%), and were economically active (94/120, 78.3%). The mean household size was 2.9 (SD 0.9) persons. The average monthly household income was HK $31,541.70 (SD HK $20,079.10), approximately US $4052.48. A minority of households (11/120, 9.2%) received financial assistance from the government or other organizations. Most caregivers reported having no chronic illness (108/120, 90%) and no disability (120/120, 100%).

Care recipients were older, with a mean age of 76.3 (SD 7.4) years, and the majority were female (66/120, 55%). Most had chronic health conditions (97/120, 80.8%), while a smaller proportion reported disabilities (20/120, 16.7%). The caregiving relationship was predominantly intergenerational. Most caregivers were caring for a parent, most commonly their mother (55/120, 45.8%) or father (44/120, 36.7%). Smaller proportions cared for a spouse (9/120, 7.5%) or other relatives, including grandparents and in-laws (12/120, 10% combined).

Regarding the caregiving characteristics, 70% (84/120) of caregivers lived with their care recipients. Approximately one-third (43/120, 35.8%) were identified as the sole caregiver. The duration of caregiving averaged 5.5 (SD 3.5) years. On average, caregivers devoted 45.9 (SD 20.1) hours per week to caregiving. The caregiver-care recipient relationship was generally positive: 48.3% (58/120) rated it "good" and 17.5% (21/120) "very good."

Feasibility Outcomes

Of the 131 eligible caregivers, 120 (91.6%) enrolled. Attendance, assessment completion, and retention were 120/120 (100%): all participants attended the 4-day psychoeducation sessions and completed both the T1 and T2 assessments, and no dropouts occurred during the study period.

Acceptability Outcomes

Table 2 presents the acceptability outcomes of the brief mindfulness JITAI. Most participants (111/120, 92.5%) found the app easy to install and use. Although a notable proportion of participants reported ignoring notifications at times (68/120, 56.7%), the majority felt the app helped them manage stress (98/120, 81.7%) and would recommend it to other caregivers (96/120, 80%). Regarding notifications, most participants considered the number of messages appropriate (87/120, 72.5%), and a substantial proportion found the timing of the messages appropriate (52/120, 43.3%). The majority of the participants liked the wording of the reminders (107/120, 89.2%).

XSL•FO
RenderX

Table . Acceptability outcomes of the brief mindfulness JITAI[a] (N=120).

Outcomes	Values, n (%)
Acceptability of the mobile app	
The app was easy to install and use	
Yes	111 (92.5)
No	9 (7.5)
Ignored the app's notifications	
Yes	68 (56.7)
No	52 (43.3)
The app helped me manage stress	
Able or helpful	98 (81.7)
Not able or not helpful	22 (18.3)
Would recommend the app to other caregivers	
Would	96 (80)
Would not	24 (20)
Acceptability of notifications	
Adequacy of the number of messages	
Too many	27 (22.5)
Appropriate	87 (72.5)
Too few	6 (5)
Helpfulness or appropriateness of notification timing	
Helpful	47 (39.2)
Appropriate	52 (43.3)
Not helpful	21 (17.5)
Liked the wording of the reminders	
Yes	107 (89.2)
No	13 (10.8)
Acceptability of mindfulness activities	
Mindfulness helped me manage stress	
Able or helpful	102 (85)
Not able or not helpful	18 (15)
Liked the graphics and characters used	
Like	99 (82.5)
Do not like	21 (17.5)
Learning mindfulness through these activities was helpful	
Helpful	102 (85)
Not helpful	18 (15)

[a]JITAI: just-in-time adaptive intervention.

Participants also indicated that the mindfulness activities generally helped them manage stress (102/120, 85%), that they liked the app's graphics and characters (99/120, 82.5%), and that the app was helpful for learning mindfulness (102/120, 85%).

Effectiveness Outcomes

Table 3 presents the effectiveness outcomes of the brief mindfulness JITAI. Figure 4 provides a visual summary of the standardized pre-post changes (effect sizes) across outcomes. The analysis of preintervention and postintervention effectiveness outcomes revealed statistically significant improvements across several psychosocial domains. Participants

XSL•FO
RenderX

showed a significant reduction in perceived stress (P<.001), depressive symptoms (P<.001), and caregiving burden (P=.003). Conversely, there was a significant increase in the positive aspects of caregiving (P<.001). A significant improvement in subjective sleep quality (ie, scores decreased) was also observed (P=.02). However, no statistically significant changes were found for health-related quality of life (P=.08) or for trait mindfulness (P=.14).

Figure 4. Forest plot of standardized outcome changes from T1 to T2. Positive values indicate improvement. For outcomes in which lower scores reflect better status (perceived stress, depressive symptoms, caregiving burden, and subjective sleep quality), effect sizes were sign-reversed so that positive values consistently represent improvement across measures. Error bars indicate 95% CIs.

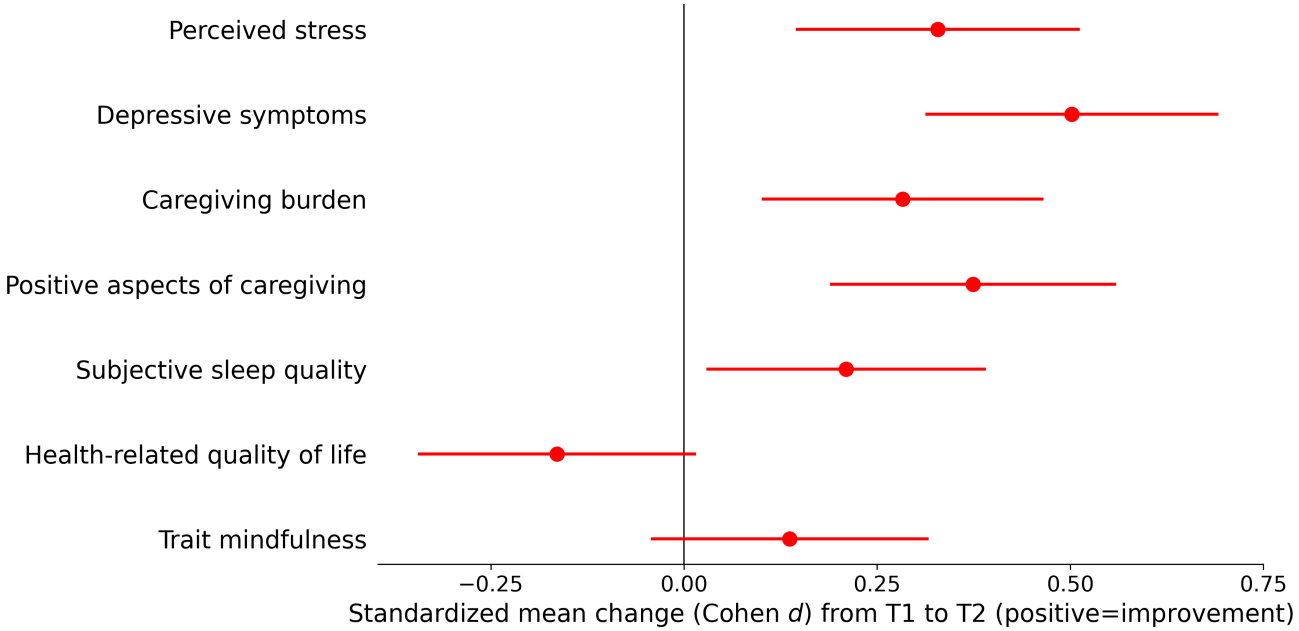


Table . Effectiveness outcomes of the brief mindfulness JITAI[a].

Outcomes	Baseline (T1), mean (SD)	Postintervention (T2), mean (SD)	*t* (*df*)	*P* value	*d*
Perceived stress	30.1 (3.8)	28.6 (4.8)	–3.6 (119)	<.001	–0.33
Depressive symptoms	8.9 (5.1)	6.7 (4.4)	–5.5 (119)	<.001	–0.50
Caregiving burden	22.8 (5.5)	21.1 (5.9)	–3.1 (119)	.003	–0.28
Positive aspects of caregiving	35.7 (6.5)	38.1 (6.3)	4.1 (119)	<.001	0.37
Subjective sleep quality	6.6 (2.8)	6.2 (2.6)	–2.3 (119)	.02	–0.21
Health-related quality of life	0.9 (0.1)	0.8 (0.1)	–1.8 (119)	.08	–0.16
Trait mindfulness	59.2 (9.2)	60.3 (9.1)	1.5 (119)	.14	0.14

[a]JITAI: just-in-time adaptive intervention.

Discussion

Principal Results

This study evaluated a brief mindfulness JITAI for family caregivers of people with dementia and found it to be a feasible, acceptable, and potentially effective approach for improving the mental well-being of caregivers. The intervention demonstrated high feasibility, with a recruitment rate of 91.6% and 100% attendance, completion, and retention. These metrics suggest that the 18-day program was appealing to caregivers and did not place an undue burden on participants.

Acceptability was also high. Most caregivers reported that the application was easy to install and use (111/120, 92.5%), helpful for managing stress (98/120, 81.7%), and something they would recommend to other caregivers (96/120, 80%). Although some participants (68/120, 56.7%) acknowledged ignoring notifications at times, most found the number of messages (87/120, 72.5%) and their delivery timing (99/120, 82.5%) to be helpful or appropriate. The mindfulness activities were positively received, with 85% (102/120) of users reporting that they were helpful for both stress management and learning mindfulness.

Regarding effectiveness, pre-post analyses revealed significant reductions in perceived stress, depressive symptoms, and

XSL•FO
RenderX

caregiving burden, along with a significant increase in positive aspects of caregiving. Sleep quality also improved significantly. However, the intervention did not produce statistically significant changes in health-related quality of life or trait mindfulness.

Comparison With Prior Work

Our findings align with prior JITAI research conducted primarily among noncaregiver populations, including smoking cessation and obesity or dietary adherence interventions, as well as mindfulness-triggering approaches in chronic pain [14,16,32]. However, to our knowledge, this is the first study to apply a mindfulness-based JITAI to family caregivers of people with dementia, a group with distinct constraints and stress dynamics that can undermine adherence to traditional MBIs. Most participants reported that the app was easy to use and helpful for learning and applying mindfulness to reduce stress, consistent with proposed mindfulness-based JITAI mechanisms that target negative affect and automatic responding by increasing present-moment attention in real time [14]. Notably, our results suggest that just-in-time delivery is especially well suited for caregivers of people with dementia, who face substantial time constraints due to heavy caregiving duties and other life commitments. By emphasizing brief windows of receptivity, the intervention may facilitate engagement, as reflected in the high attendance and retention rates. Although mindfulness can help caregivers manage stress, dementia care is a long-term trajectory, and traditional MBIs often struggle to achieve sustained impact because of the required time commitment [9]. By targeting receptive moments for self-practice, the JITAI approach shows promise for supporting longer-term practice across the caregiving trajectory.

Prior mindfulness-based JITAIs, such as those targeting substance use, often trigger intervention prompts at moments of vulnerability (eg, cravings) [14]. For caregivers of people with dementia, however, high-vulnerability states (eg, acute stress) can impede practice because stress taxes attentional control, creating cognitive interference that hinders engagement [11]. As such, our brief mindfulness JITAI focused on prompting during receptive moments. To optimize receptivity, we used three complementary approaches: (1) a machine learning–informed static model derived from pilot data, (2) a random-timing model, and (3) a machine learning–based adaptive model that updated using each participant's receptivity data over time. Overall, most participants rated message frequency and timing favorably, though some notifications were ignored, indicating room for improvement. Because multiple models were deployed concurrently, we cannot isolate the most optimal approach. Future work should directly compare static, random, and adaptive models to maximize receptivity and, in turn, enhance the benefits of mindfulness for this population.

Regarding self-reported outcomes, we observed small to medium reductions in stress and depressive symptoms, consistent with prior MBIs for caregivers of people with dementia [9,24,33]. The mindful coping model posits that decentering disrupts negative thought cycles and supports more positive stress appraisals [10]. Through this mechanism, mindfulness may reduce depressive symptoms by decreasing rumination, which is a repetitive, problem-unresolved focus on negative thoughts [34]. These findings suggest that a mindfulness-based JITAI can yield benefits that align with traditional MBIs while requiring a less intensive and more flexible structure than standard 8-week programs. Such a less intensive approach may reduce burden and improve accessibility by delivering support via a mobile app, thereby enhancing reach for more caregivers. However, because we did not include a traditional MBI control, we cannot directly compare effect sizes on psychosocial outcomes. Future research should compare JITAI-based mindfulness interventions with standard, group-based mindfulness interventions to clarify their relative efficacy in reducing stress and depressive symptoms among caregivers of people with dementia.

We also observed a small reduction in caregiving burden, diverging from findings in previous trials of MBIs for caregivers of people with dementia that often show no effect on caregiving burden [24,33]. One explanation is that self-directed practices and face-to-face attendance in traditional MBIs add to already busy schedules and are frequently affected by low adherence to self-directed practice. If caregivers perceive the intervention as too intensive, they may lack the motivation to engage in self-practice; yet informal mindfulness practice has been linked to improvements in psychological well-being [35]. By delivering prompts when caregivers are most receptive, a just-in-time approach may better motivate self-directed practices. Heightened moment-to-moment attention in real time via a JITAI may also foster more positive appraisals of caregiving experiences, aligning with our observed gains in positive aspects of caregiving, an outcome rarely examined in prior MBI trials in this population. Cross-sectional evidence similarly links mindfulness with more positive caregiving experiences [36]. Consistent with reduced burden, we also found improvements in subjective sleep quality, which contrasts with some prior findings from MBIs in this population [37,38]. A recent scoping review supports an association between caregiving burden and subjective sleep quality [39]. Because behavioral and psychological symptoms of dementia are often a major source of burden [40], delivering mindfulness via a JITAI with increased self-practice may facilitate caregivers' translation of mindfulness into more positive responses to behavioral and psychological symptoms of dementia, potentially reducing burden [9]. Given the absence of a control group, these findings should be interpreted cautiously and not as evidence of causality. Randomized controlled trials (RCTs) with process evaluations are needed to elucidate mechanisms underlying reductions in caregiver burden in mindfulness-based JITAIs for this population.

In contrast, neither trait mindfulness nor health-related quality of life improved after the brief mindfulness–based JITAI. This suggests that the 18-day dosage, including 4 days of psychoeducation, may be insufficient to shift dispositional mindfulness, which typically requires sustained practice over a longer period [41]. For example, longer programs (eg, 16-wk formats) have shown gains in trait mindfulness in caregivers of people with dementia [23]. This may also explain the absence of improvement in health-related quality of life, given that trait mindfulness can mediate the relationship between stress and

XSL•FO
RenderX

health-related quality of life [42]. Future JITAI-based mindfulness interventions should systematically examine dose-response relationships to clarify their effects on trait mindfulness in this population.

Implications

This study demonstrates the feasibility and preliminary effectiveness of a mindfulness-based JITAI to support family caregivers of people with dementia. By prompting brief mindfulness micropractices at moments of higher receptivity, the intervention directly addresses practical barriers that commonly limit caregivers' engagement with traditional MBIs (eg, time constraints, competing care demands, and difficulty sustaining self-directed practice). If these findings are confirmed in controlled trials, this delivery approach could be implemented as a low-intensity, scalable adjunct to the existing community and nongovernmental organization caregiver services by providing between-session support, extending reach to caregivers who are homebound or unable to attend groups regularly, and improving service efficiency through engagement data that can help identify caregivers who may benefit from additional dose, booster support, or stepped-care escalation. In this way, a mindfulness JITAI may strengthen routine caregiver services by enhancing access, personalization, and continuity of support with minimal additional manpower burden to the health and community care system.

Limitations

This study has several limitations. First, the pre-post, quasi-experimental design lacks a control group, precluding causal inference. Although we observed pre-post reductions in perceived stress and other psychosocial outcomes, these changes cannot be attributed solely to the brief mindfulness JITAI. We encourage future studies to evaluate the program using more rigorous research designs, such as RCTs. Second, we did not include process indicators to examine mechanisms of action; thus, we cannot confirm that improvements resulted from enhanced self-directed mindfulness practices. Future studies should incorporate process measures and mediation analyses to elucidate the mechanisms of change in JITAIs and clarify their unique benefits relative to traditional delivery. Third, although our goal was to examine the general effects of a JITAI, the optimal delivery algorithm remains unclear. Future trials may consider using a microrandomized trial design to test whether a static, random, or adaptive model most effectively predicts and acts upon moments of receptivity. Fourth, this study examines only the effects of the brief JITAI on caregivers. However, a recent meta-analysis highlights dyadic psychosocial effects, whereby depressive symptoms in people with dementia can negatively affect caregivers' quality of life [43]. Given the close linkage between caregivers' and care recipients' psychosocial health, future research should evaluate the effects of the brief mindfulness JITAI at the dyadic level. Fifth, the study's generalizability may be affected by the sample characteristics and the technology eligibility criteria. In our sample, participants were relatively young (mean age of 45.4, SD 11.8 y), predominantly female (80/120, 66.7%), and highly educated (66/120, 55% tertiary education), whereas territory-level data on carers of older people overall suggest an older caregiver profile (74.5% female among 966 respondents; mean age of 54.5, SD 18.1 y) [44]. Due to the development and resource constraints in implementing the intervention on both operating systems in parallel, the study was delivered on Android only, and caregivers who primarily used iOS devices were excluded. This platform restriction may have introduced selection bias toward more digitally ready participants and may partly explain the relatively younger and more educated sample. In addition, higher educational attainment may reflect self-selection (eg, greater health literacy or confidence with app-based, self-guided interventions), potentially inflating feasibility, acceptability, and engagement estimates. Future trials should broaden recruitment channels and support multiplatform delivery, alongside enhanced onboarding and technical support, to better capture older, male, and lower-literacy caregiver subgroups. Sixth, the intervention was brief, and there was no long-term follow-up; thus, the sustainability of effects remains unclear. The program lasted 18 days, and outcomes were assessed only immediately post intervention, limiting conclusions about whether observed improvements would be maintained after prompts ended. Given that dementia caregiving is a long-term role, future studies should include longer follow-up (eg, 1-3 - 6 mo) to evaluate durability and to determine whether maintenance supports (eg, tapered prompting, booster sessions, or continued low-intensity prompting) are required to sustain practice and benefits under real-world caregiving conditions.

Conclusions

A brief, mindfulness-based JITAI appears feasible and highly acceptable for supporting caregivers of people with dementia. We observed statistically significant reductions in perceived stress, depressive symptoms, and caregiving burden, alongside increased positive aspects of caregiving and subjective sleep quality, supporting the potential of adaptive, technology-enabled interventions to provide timely, personalized support tailored to caregivers' complex needs. Future work should involve a rigorously designed, adequately powered RCT to confirm efficacy, clarify mechanisms, and inform broader implementation.

Acknowledgments

The authors extend their sincere gratitude to all participants who contributed to this study. No generative artificial intelligence tools were used.

XSL•FO
RenderX

Funding

This study was funded by the Health and Medical Research Fund from the Health Bureau, Hong Kong SAR (project 20210051). The funder had no role in the study design, data collection, analysis, interpretation of results, manuscript preparation, or dissemination of findings.

Authors' Contributions

Conceptualization: KLC, PPKK
Data curation: KLC, PPKK, APLT, MQ
Formal analysis: KLC, PPKK, APLT, MQ
Investigation: KLC, PPKK, APLT, DSKC
Writing – original draft: PPKK, APLT
Writing – review and editing: KLC, PPKK, APLT, DSKC, MQ

Conflicts of Interest

None declared.

Checklist 1
CONSORT checklist.
[DOCX File, 33 KB - aging_v9i1e87316_app1.docx]

References

1. World Alzheimer Report 2024: global changes in attitudes to dementia. : Alzheimer's Disease International; 2024 URL: https://www.alzint.org/u/World-Alzheimer-Report-2024.pdf [accessed 2026-05-19]
2. Whittaker AC, Gallagher S. Caregiving alters immunity and stress hormones: a review of recent research. Curr Opin Behav Sci 2019 Aug;28:93-97. [doi: 10.1016/j.cobeha.2019.02.002]
3. Ferrara M, Langiano E, Di Brango T, De Vito E, Di Cioccio L, Bauco C. Prevalence of stress, anxiety and depression in with Alzheimer caregivers. Health Qual Life Outcomes 2008 Nov 6;6(93):93. [doi: 10.1186/1477-7525-6-93] [Medline: 18990207]
4. McCurry SM, Gibbons LE, Logsdon RG, Vitiello MV, Teri L. Insomnia In caregivers of persons with dementia: who is at risk and what can be done about it? Sleep Med Clin 2009 Dec 1;4(4):519-526. [doi: 10.1016/j.jsmc.2009.07.005] [Medline: 20046806]
5. Gaugler JE, Pot AM, Zarit SH. Long-term adaptation to institutionalization in dementia caregivers. Gerontologist 2007 Dec;47(6):730-740. [doi: 10.1093/geront/47.6.730] [Medline: 18192627]
6. Cooper C, Selwood A, Blanchard M, Livingston G. Abusive behaviour experienced by family carers from people with dementia: the CARD (caring for relatives with dementia) study. J Neurol Neurosurg Psychiatry 2010 Jun;81(6):592-596. [doi: 10.1136/jnnp.2009.190934] [Medline: 19955110]
7. Kor PPK, Chou KL, Tsang APL, et al. Effects of a closed-loop mindfulness-based program for reducing stress in family caregivers of people with dementia: a study protocol of a randomized controlled trial. BMC Psychol 2025 Jul 1;13(1):681. [doi: 10.1186/s40359-025-02919-2] [Medline: 40597419]
8. Kor PPK, Chien WT, Liu JYW, Lai CKY. Mindfulness-based intervention for stress reduction of family caregivers of people with dementia: a systematic review and meta-analysis. Mindfulness 2018 Feb;9(1):7-22. [doi: 10.1007/s12671-017-0751-9]
9. Kor PPK, Liu JYW, Chien WT. Effects of a modified mindfulness-based cognitive therapy for family caregivers of people with dementia: a randomized clinical trial. Gerontologist 2021 Aug 13;61(6):977-990. [doi: 10.1093/geront/gnaa125] [Medline: 32886746]
10. Garland EL, Gaylord SA, Fredrickson BL. Positive reappraisal mediates the stress-reductive effects of mindfulness: an upward spiral process. Mindfulness 2011 Mar;2(1):59-67. [doi: 10.1007/s12671-011-0043-8]
11. Shields GS, Sazma MA, Yonelinas AP. The effects of acute stress on core executive functions: a meta-analysis and comparison with cortisol. Neurosci Biobehav Rev 2016 Sep;68:651-668. [doi: 10.1016/j.neubiorev.2016.06.038] [Medline: 27371161]
12. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
13. Kramer JN, Künzler F, Mishra V, et al. Investigating intervention components and exploring states of receptivity for a smartphone app to promote physical activity: protocol of a microrandomized trial. JMIR Res Protoc 2019 Jan 31;8(1):e11540. [doi: 10.2196/11540] [Medline: 30702430]

XSL•FO
RenderX

14. Yang MJ, Sutton SK, Hernandez LM, et al. A Just-In-Time Adaptive intervention (JITAI) for smoking cessation: Feasibility and acceptability findings. Addict Behav 2023 Jan;136:107467. [doi: 10.1016/j.addbeh.2022.107467] [Medline: 36037610]
15. Goldstein SP, Zhang F, Klasnja P, Hoover A, Wing RR, Thomas JG. Optimizing a just-in-time adaptive intervention to improve dietary adherence in behavioral obesity treatment: protocol for a microrandomized trial. JMIR Res Protoc 2021 Dec 6;10(12):e33568. [doi: 10.2196/33568] [Medline: 34874892]
16. Garland EL, Gullapalli BT, Prince KC, et al. Zoom-based mindfulness-oriented recovery enhancement plus just-in-time mindfulness practice triggered by wearable sensors for opioid craving and chronic pain. Mindfulness (N Y) 2023 May 24;14(6):1-17. [doi: 10.1007/s12671-023-02137-0] [Medline: 37362184]
17. Mishra V, Künzler F, Kramer JN, Fleisch E, Kowatsch T, Kotz D. Detecting receptivity for mhealth interventions in the natural environment. Proc ACM Interact Mob Wearable Ubiquitous Technol 2021 Jun;5(2):74. [doi: 10.1145/3463492] [Medline: 34926979]
18. Goldstein SP, Evans BC, Flack D, et al. Return of the JITAI: applying a just-in-time adaptive intervention framework to the development of m-health solutions for addictive behaviors. Int J Behav Med 2017 Oct;24(5):673-682. [doi: 10.1007/s12529-016-9627-y] [Medline: 28083725]
19. Carmody J, Baer RA. Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. J Behav Med 2008 Feb;31(1):23-33. [doi: 10.1007/s10865-007-9130-7] [Medline: 17899351]
20. Leung DY, Lam TH, Chan SS. Three versions of Perceived Stress Scale: validation in a sample of Chinese cardiac patients who smoke. BMC Public Health 2010 Aug 25;10(1):513. [doi: 10.1186/1471-2458-10-513] [Medline: 20735860]
21. Beneria A, Arnedo M, Contreras S, et al. Impact of simulation-based teamwork training on COVID-19 distress in healthcare professionals. BMC Med Educ 2020 Dec 21;20(1):515. [doi: 10.1186/s12909-020-02427-4] [Medline: 33349248]
22. Klatt MD, Buckworth J, Malarkey WB. Effects of low-dose mindfulness-based stress reduction (MBSR-ld) on working adults. Health Educ Behav 2009 Jun;36(3):601-614. [doi: 10.1177/1090198108317627] [Medline: 18469160]
23. Cheung DSK, Kor PPK, Jones C, et al. The use of modified mindfulness-based stress reduction and mindfulness-based cognitive therapy program for family caregivers of people living with dementia: a feasibility study. Asian Nurs Res (Korean Soc Nurs Sci) 2020 Oct;14(4):221-230. [doi: 10.1016/j.anr.2020.08.009] [Medline: 32931996]
24. Brown KW, Coogle CL, Wegelin J. A pilot randomized controlled trial of mindfulness-based stress reduction for caregivers of family members with dementia. Aging Ment Health 2016 Nov;20(11):1157-1166. [doi: 10.1080/13607863.2015.1065790] [Medline: 26211415]
25. Yu X, Tam WWS, Wong PTK, Lam TH, Stewart SM. The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. Compr Psychiatry 2012 Jan;53(1):95-102. [doi: 10.1016/j.comppsych.2010.11.002] [Medline: 21193179]
26. Tang JYM, Ho AHY, Luo H, et al. Validating a Cantonese short version of the Zarit Burden Interview (CZBI-Short) for dementia caregivers. Aging Ment Health 2016 Sep;20(9):996-1001. [doi: 10.1080/13607863.2015.1047323] [Medline: 26016419]
27. Lou VWQ, Lau BHP, Cheung KSL. Positive Aspects of Caregiving (PAC): scale validation among Chinese dementia caregivers (CG). Arch Gerontol Geriatr 2015;60(2):299-306. [doi: 10.1016/j.archger.2014.10.019] [Medline: 25488014]
28. Chien TW, Hsu SY, Tai C, Guo HR, Su SB. Using Rasch analysis to validate the revised PSQI to assess sleep disorders in Taiwan's hi-tech workers. Community Ment Health J 2008 Dec;44(6):417-425. [doi: 10.1007/s10597-008-9144-9] [Medline: 18437567]
29. Wong ELY, Cheung AWL, Wong AYK, Xu RH, Ramos-Goñi JM, Rivero-Arias O. Normative profile of health-related quality of life for Hong Kong general population using preference-based instrument EQ-5D-5L. Value Health 2019 Aug;22(8):916-924. [doi: 10.1016/j.jval.2019.02.014] [Medline: 31426933]
30. Deng YQ, Li S, Tang YY, Zhu LH, Ryan R, Brown K. Psychometric properties of the Chinese translation of the Mindful Attention Awareness Scale (MAAS). Mindfulness 2012 Mar;3(1):10-14. [doi: 10.1007/s12671-011-0074-1]
31. Liu Z, Chen QL, Sun YY. Mindfulness training for psychological stress in family caregivers of persons with dementia: a systematic review and meta-analysis of randomized controlled trials. Clin Interv Aging 2017;12:1521-1529. [doi: 10.2147/CIA.S146213] [Medline: 29026290]
32. Vinci C, Sutton SK, Yang MJ, Jones SR, Kumar S, Wetter DW. Proximal effects of a just-in-time adaptive intervention for smoking cessation with wearable sensors: Microrandomized Trial. JMIR Mhealth Uhealth 2025 Mar 19;13:e55379. [doi: 10.2196/55379] [Medline: 40106803]
33. Mallya S, Fiocco AJ. The effects of mindfulness training on cognitive and psychosocial well-being among family caregivers of persons with neurodegenerative disease. Mindfulness 2019 Oct;10(10):2026-2037. [doi: 10.1007/s12671-019-01155-1]
34. Deyo M, Wilson KA, Ong J, Koopman C. Mindfulness and rumination: does mindfulness training lead to reductions in the ruminative thinking associated with depression? EXPLORE (NY) 2009;5(5):265-271. [doi: 10.1016/j.explore.2009.06.005] [Medline: 19733812]
35. Birtwell K, Williams K, van Marwijk H, Armitage CJ, Sheffield D. An exploration of formal and informal mindfulness practice and associations with wellbeing. Mindfulness (N Y) 2019;10(1):89-99. [doi: 10.1007/s12671-018-0951-y] [Medline: 30662573]

XSL•FO
RenderX

36. Tolea MI, Camacho S, Cohen IR, Galvin JE. Mindfulness and care experience in family caregivers of persons living with dementia. J Alzheimers Dis Rep 2023;7(1):151-164. [doi: 10.3233/ADR-220069] [Medline: 36891256]
37. Oken BS, Fonareva I, Haas M, et al. Pilot controlled trial of mindfulness meditation and education for dementia caregivers. J Altern Complement Med 2010 Oct;16(10):1031-1038. [doi: 10.1089/acm.2009.0733] [Medline: 20929380]
38. Paller KA, Creery JD, Florczak SM, et al. Benefits of mindfulness training for patients with progressive cognitive decline and their caregivers. Am J Alzheimers Dis Other Demen 2015 May;30(3):257-267. [doi: 10.1177/1533317514545377] [Medline: 25154985]
39. Mattos MK, Bernacchi V, Shaffer KM, et al. Sleep and caregiver burden among caregivers of persons living with dementia: a scoping review. Innov Aging 2024;8(2):igae005. [doi: 10.1093/geroni/igae005] [Medline: 38420182]
40. Huang SS, Lee MC, Liao YC, Wang WF, Lai TJ. Caregiver burden associated with behavioral and psychological symptoms of dementia (BPSD) in Taiwanese elderly. Arch Gerontol Geriatr 2012;55(1):55-59. [doi: 10.1016/j.archger.2011.04.009] [Medline: 21601931]
41. Shapiro S, Siegel R, Neff KD. Paradoxes of mindfulness. Mindfulness 2018 Dec;9(6):1693-1701. [doi: 10.1007/s12671-018-0957-5]
42. de Frias CM, Whyne E. Stress on health-related quality of life in older adults: the protective nature of mindfulness. Aging Ment Health 2015;19(3):201-206. [doi: 10.1080/13607863.2014.924090] [Medline: 24940847]
43. Ning W, Wang S, Xu Y, Cheung D. The association between caregiver psychosocial factors and depressive symptoms in people with dementia: a systematic review and meta-analysis. J Adv Nurs 2026 Apr;82(4):3906-3942. [doi: 10.1111/jan.17093] [Medline: 40556482]
44. Consultancy study on needs and support required of carers of elderly persons and of persons with disabilities in Hong Kong: main report. : Labour and Welfare Bureau; 2022 URL: https://www.lwb.gov.hk/en/other_info/1.%20Carer%20Study-Main%20Report%20(Eng)_2022.05-r.pdf [accessed 2026-05-19]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
JITAI: just-in-time adaptive intervention
MBI: mindfulness-based intervention
MBSR: mindfulness-based stress reduction
PAC: Positive Aspects of Caregiving
PSS-14: 14-item Perceived Stress Scale
RCT: randomized controlled trial

Edited by M O'Connell; submitted 09.Nov.2025; peer-reviewed by A Samraj, S Bartels; revised version received 06.Feb.2026; accepted 19.Mar.2026; published 05.Jun.2026.

Please cite as:
Kor PPK, Tsang APL, Cheung DSK, Zarit SH, Xie H, Qian M, Tan KC, Zheng Y, Chu AMY, Chou KL
Developing and Testing a Brief Mindfulness Just-in-Time Adaptive Intervention to Reduce Stress Among Caregivers of People With Dementia: Quasi-Experimental Study
JMIR Aging 2026;9:e87316
URL: https://aging.jmir.org/2026/1/e87316
doi:10.2196/87316

© Patrick Pui Kin Kor, Alex Pak Lik Tsang, Daphne Sze Ki Cheung, Steven H Zarit, Haoran Xie, Min Qian, Kay Chen Tan, Yuanqing Zheng, Amanda Man Ying Chu, Kee Lee Chou. Originally published in JMIR Aging (https://aging.jmir.org), 5.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Combining Subjective Perceptions and Objective Behavioral Metrics With the Elderly Digital Twin System: Quantitative Usability Study

Ziaullah Momand, PhD; Pornchai Mongkolnam, PhD; Debajyoti Pal, PhD; Siam Yamsaengsung, PhD

School of Information Technology, King Mongkut's University of Technology Thonburi, 126 Pracha Uthit Rd, Bang Mod, Thung Khru, Bangkok, Thailand

Corresponding Author:
Pornchai Mongkolnam, PhD
School of Information Technology, King Mongkut's University of Technology Thonburi, 126 Pracha Uthit Rd, Bang Mod, Thung Khru, Bangkok, Thailand

Abstract

Background: The growing aging population has increased the need for technologies that support informal caregivers in home-based older adult care. Digital twin (DT) systems offer promising capabilities; yet, their effectiveness depends on usability, an aspect still insufficiently evaluated among caregivers.

Objective: This study aimed to assess the usability of an older adult care DT system using a dual-method evaluation that integrates subjective and objective behavioral performance.

Methods: Fifty caregivers participated in a usability assessment combining the System Usability Scale (SUS) and detailed system activity log analytics. Log-based measures included task completion, time on task, errors, and abandonment rate. A composite user engagement score was computed and analyzed for correlation and predictive association with SUS ratings. Engagement clusters were also explored.

Results: Caregivers reported an excellent mean SUS score of 80.45. System logs showed a 94.08% task completion rate, 2.66% abandonment, and an average task duration of 89.16 seconds. User engagement score demonstrated significant correlations with SUS (r=0.626, ρ=0.552; P<.001) and significantly predicted usability in regression analysis (β=52.94, R^2=0.392; P<.001). Engagement-based clustering identified high-, medium-, and low-tier user groups, each exhibiting distinct usability patterns.

Conclusions: Integrating subjective usability ratings with objective behavioral metrics provides a rigorous and comprehensive approach to evaluating DT systems for older adult care. The findings highlight strong usability of the system and offer actionable insights for refining caregiver support technologies.

(JMIR Aging 2026;9:e91873) doi:10.2196/91873

KEYWORDS

elderly digital twin; usability study; System Usability Scale; user engagement score; caregiver support

Introduction

Background

The global demographic shift toward an aging population presents unprecedented challenges for health care systems worldwide. The world's population older than 60 years is projected to double from 1.06 billion (13.5%) in 2020 to 2.13 billion in 2050 (22.0%), raising significant concerns about the impact of aging [1]. This rapid aging process is characterized by an increased prevalence of chronic diseases, complex health care needs, and a growing demand for personalized and proactive health care models [2,3]. In the current technology-driven age, digital twin (DT) has emerged as a transformative approach in health care, offering the potential to revolutionize older adult care through predictive analytics and personalized interventions. DTs are virtual representations of physical entities that enable the dynamic simulation of potential treatment strategies, monitoring and prediction of health trajectories, and early intervention and prevention [4]. In the context of older adult health monitoring, DTs create comprehensive virtual models that integrate multimodal data, including clinical records, genetic information, wearable sensor data, and environmental factors, to provide a holistic view of an individual's health status [5,6].

The application of DT technology in older adult care is particularly compelling because of its ability to monitor vital signs, physiological parameters, and other health-related data in real time, enabling health care providers to detect early signs of deterioration or anomalies and proactively intervene in such scenarios [6]. Recent implementations have demonstrated the potential of this technology in aging-in-place scenarios, where DTs facilitate continuous health monitoring while preserving

XSL•FO
RenderX

the independence and quality of life of seniors [7]. For example, DT systems have been successfully deployed to monitor older adult patients for fall detection, abnormal posture recognition, and health risk assessment, with validation trials showing promising results in real-world nursing home environments [8].

The complexity of older adult care often necessitates significant involvement from both formal and informal caregivers, who play crucial roles in health monitoring, decision-making, and care coordination. Informal caregivers, predominantly family members, face substantial challenges in navigating complex health care systems and making informed decisions [9,10]. Family caregivers' decision support interventions have shown promise in improving care outcomes, with studies demonstrating that well-designed decision support tools can reduce family caregivers' decision uncertainty and improve satisfaction with the quality of care [9]. Such tools leverage trust, cultural humility, strength-based approaches, and effective information sharing to facilitate meaningful conversations between caregivers and health care providers. In the context of DT systems for older adult care, caregiver-facing interfaces have become essential components that must translate complex health data into actionable insights that support informed care–related decisions [10].

The success of DT systems in older adult health care depends critically on their usability, particularly in home care settings, where family members are involved in a variety of experiences [11]. In health care contexts, poor usability can have serious consequences, including medical errors, reduced adoption rates, and compromised patient safety [12]. For DT systems targeting older adult populations and their caregivers, usability considerations have become even more critical because of the potential age-related changes in vision, dexterity, and cognitive processing. Older adults often require caregiver mediation when interacting with digital health technologies, and systems must be designed to accommodate both direct older adult users and caregiver interactions [13].

In health care usability research, evaluation has largely relied on subjective self-report instruments, most prominently the System Usability Scale (SUS), which has been applied across diverse digital health systems such as caregiver monitoring tools HELMA [14], self-management apps for heart failure (Engage) [13], exergame training systems [15], and interdisciplinary hospital information platforms such as therapy and monitoring systems [16]. While SUS provides standardized benchmarking, its reliance on user perception limits insights into actual system interaction. Studies have attempted to complement the SUS with performance or log data: the HELMA project tracked caregiver login frequency and session duration, the Engage system analyzed task completion and observed errors, and exergame interventions logged exercise scores and reaction accuracy. However, these measures often remain surface-level, focusing on frequency, duration, or clinical outcomes rather than granular task behavior analytics such as navigation sequences, error recovery, feature usage, task completion, time on task, click on task ratio, retry rate, session duration, or task abandonment ratio. Moreover, limitations such as older adults' reliance on caregiver mediation [14,17] and testing in controlled settings [13,15] constrain the validity of existing evaluations. Collectively, these findings underscore a critical gap: while subjective usability assessments dominate and objective measures have been trailed, few health care studies [13,14] have integrated log-based behavioral data with standardized questionnaires. To address this gap, our study proposes a dual-method evaluation framework that unites SUS scores with detailed behavioral engagement metrics, providing both perceived and observed dimensions of usability in caregiver-facing DT systems.

To address this gap, this study conducted a usability evaluation of an elderly digital twin (EDT) system designed to support informal caregivers in home care settings. In our previous work [18], we proposed a comprehensive EDT framework to assist caregivers in making informed decisions. To evaluate the feasibility and applicability of the framework, we developed a working EDT system prototype. Unlike prior studies that emphasized technical modeling or data integration, this study explicitly targets the informal caregiver perspective, in which usability is critical for adoption in daily home care. Our evaluation follows a dual-method approach with the following two objectives: (1) subjective evaluation of the usability of our proposed EDT system by assessing caregivers' perceptions using the hugely popular SUS; (2) objective evaluation of usability by proposing a novel behavioral engagement metric that can be used in conjunction with the SUS.

We present a dual-method usability evaluation framework that combines log-based behavioral analytics (objective measurement) with a standardized SUS questionnaire (subjective measurement) [19]. This approach leverages the complementary strengths of objective behavioral data and subjective user feedback to provide a comprehensive view of the system's usability. The log-based component captures detailed user behavior, including interaction patterns, feature usage, task completion pathways, error rates, and temporal usage characteristics. These metrics offer insights into real-world usage that may not be revealed through self-reported measures alone. In parallel, the SUS component provides standardized usability scores that enable benchmarking against established norms in health care applications and allow comparisons with other digital health interventions.

This dual-method approach directly addresses a noticeable gap in existing evaluation practices: the dominance of subjective methods with limited or superficial use of objective interaction data. Previous health care systems' usability studies often relied on the SUS or combined it with a basic usage log [13,14,17], which failed to capture how systems are actually navigated and used in real practice. By integrating behavioral engagement metrics with subjective user perceptions, our study provides an evidence-based, user-centered, and replicable framework for evaluating digital health applications. Applied specifically to caregiver-facing interfaces for DTs of the older adults, this methodology not only fills the gap but also offers actionable insights to guide iterative design and ensure more effective, caregiver-friendly digital health systems. The remainder of this manuscript is organized as follows: Section 2 reviews related work, Section 3 outlines the methods, Section 4 presents the results, Section 5 offers a discussion, and Section 6 concludes the study.

XSL•FO
RenderX

Related Work

The growing complexity of older adult care has driven the development of digital health technologies, including DT systems, to support clinical decision-making and informal caregivers. As these systems advance, ensuring their usability, particularly for nonprofessional caregivers, is critical for adoption and effectiveness. Previous studies have evaluated digital health technologies using subjective measures, such as the SUS or qualitative interviews. This section reviews usability studies and highlights the methodological gap addressed by our dual-method approach.

Usability Evaluation Methods: Subjective Versus Objective

Usability evaluation methods fall into 2 primary categories: subjective and objective. Each category offers unique advantages and faces distinct limitations, making the choice between them or their combination critical for comprehensive usability assessment. SUS remains the most widely adopted usability instrument, valued for its brevity, reliability, and extensive benchmarking data [20,21]. Recent validation studies have reinforced SUS's psychometric properties of the SUS. A validation study of voice user interfaces confirmed SUS's usefulness across different interaction modalities. Additionally, researchers have developed item-level benchmarks for the SUS, allowing practitioners to interpret individual items when specific usability attributes require targeted assessment. These benchmarks enable a more granular evaluation while maintaining the scale's standardized scoring advantage [20,22].

In addition to the SUS, several other standardized questionnaires serve specific evaluation contexts. The usability metric for user experience (UMUX) and its shortened variant, UMUX-LITE, offer alternatives that are more closely aligned with the ISO (International Organization for Standardization) 9241 definitions of usability. Recent psychometric evaluations suggest that the UMUX-LITER provides the closest correspondence to SUS scores when converted to comparable scales [23]. The NASA-TLX (National Aeronautics and Space Administration–Task Load Index) addresses perceived workload in complex, high-consequence environments, making it particularly valuable for health care, aerospace, and military applications. However, its complexity and administration time limit its usefulness for consumer product evaluations [24]. AttrakDiff occupies a unique position by measuring both the pragmatic and hedonic quality dimensions of the user experience. This dual focus allows for the evaluation of traditional usability metrics alongside emotional and aesthetic responses, although recent studies have raised concerns about translation reliability and cultural adaptation [25].

Subjective methods are easy to administer, resource-efficient, and standardized, enabling benchmark comparisons and effective communication among practitioners [26,27]. However, subjective measures have significant limitations. Social desirability bias is a primary concern, with users potentially providing responses they believe evaluators want to hear rather than their genuine perceptions. Cultural factors also influence responses, with different populations rather than actual performance potentially missing critical usability issues that users may not consciously recognize [25,26,28].

Objective evaluation centers on measurable performance indicators, such as task completion, time on task, and error analysis, which together provide insights into system effectiveness and efficiency [29-31]. Error analysis offers detailed insights into the usability problems. Error rates can be calculated globally (total errors divided by total attempts) and task-specific (errors per task opportunity), providing different perspectives on system performance. Modern approaches distinguish between error types and severity levels, enabling targeted improvements [29,31].

Clickstream analysis is a core component of behavioral log methods and has emerged as a powerful objective approach for tracking user navigation patterns, feature usage, and drop-off points. By capturing the sequence of interactions, it highlights problems that are invisible to traditional subjective metrics [32-34]. Advanced behavioral analytics extend this further by combining multiple interaction types, such as clicks, mouse movements, and page transitions, to generate comprehensive user journey maps. These methods are particularly valuable for detecting navigation bottlenecks and optimizing conversion paths in digital systems.

Objective methods generally provide concrete evidence for design decisions by revealing fine-grained interaction patterns that users may not consciously recognize or accurately report [35,36]. However, they also face implementation challenges: data collection and analysis can be complex and resource-intensive, and these methods may not directly capture user satisfaction or emotional responses, requiring supplementary subjective measures. Moreover, standardized interpretation frameworks for many objective metrics remain underdeveloped, complicating cross-study comparisons [26]. Recent health care usability studies have demonstrated effective hybrid implementation [37] and described a hybrid approach that satisfies both pragmatic development needs and academic research requirements. Their framework captures detailed behavioral data for intermediate iterations while enabling deeper qualitative analysis for academic dissemination. Contemporary usability evaluation is shifting toward integrated approaches that combine machine learning with behavioral subjective data analysis. Advanced statistical methods improve the correlation between objective performance and user satisfaction [38]. Recognizing that no single method offers a complete picture, the field increasingly adopts multimethod strategies that balance methodological strengths and limitations while maintaining practical feasibility [37].

Clinical Relevance of Fitbit Sense 2 Data

Consumer wearables are increasingly used in digital health because they enable continuous, low-burden physiological monitoring outside clinical settings. Their clinical relevance, however, is parameter-specific rather than uniform. A systematic review of Fitbit devices found the strongest evidence for step counting under selected conditions, whereas energy expenditure, sleep, and some other measures showed less consistent accuracy and should be interpreted cautiously in health-related decision contexts [39]. A later meta-analysis of wrist-worn Fitbit sleep

XSL•FO
RenderX

models reported that newer sleep-staging devices performed better than earlier motion-only models but still were not substitutes for polysomnography [40]. Because this study used Fitbit Sense 2, newer-generation evidence is also relevant: a prospective multicenter validation study of 11 consumer sleep trackers found that Fitbit Sense 2 showed moderate agreement with polysomnography for sleep-stage classification and competitive performance among wearables, including relatively strong performance in deep-stage detection [41].

Taken together, these findings support a measured role for Fitbit Sense 2 in caregiver-facing systems. Recent free-living validation of a newer Fitbit device against medical-grade references showed moderate to good agreement for daily steps, resting heart rate, respiratory rate, and some heart rate variability (HRV) measures but weaker agreement for oxygen saturation, indicating that some physiological channels are more dependable than others [39,41,42]. Accordingly, Fitbit Sense 2 data in the EDT are best viewed as supporting longitudinal trend monitoring, anomaly awareness, and caregiver-oriented decision support rather than diagnostic inference.

Current Usability Evaluation Methods for Health Care Systems

As shown in Table 1, prior usability studies of digital health systems have relied on subjective methods, particularly the SUS, interviews, and qualitative feedback [13-15,17,43-48]. Only a few studies incorporated partial task-based or observational measures, and these were typically limited and not derived from detailed real-time interaction logs [13,14,45]. Moreover, studies focused on patients, older adults, or clinical professionals in institutional settings rather than informal caregivers in home-based care [16,17,44-48]. Synthesizing these studies reveals three major gaps: (1) heavy reliance on subjective measures without integrating detailed objective interaction logs, (2) limited focus on formal caregivers in home settings, and (3) superficial assessment of task behavior, where satisfaction or completion was reported but navigation patterns, retries, and feature usage were not captured. To address these gaps, this study integrates subjective usability perceptions (SUS) with objective system usage data (user activity logs), providing a more ecologically valid and holistic framework for evaluating caregiver-facing DT systems for older adults.

Table . Comparison of usability methods, target groups, and system features across related studies.

Target users	Method	SUS[a]	Log	Setting	Focus	Behavioral metrics	References
Informal caregivers	SUS, interview	No	No	Home	Stroke support	No	[43]
Older adults + Caregivers	SUS, interview	72.2	Partial	Home	Cognitive monitoring	Partial	[14]
Older adults + Caregivers	SUS, NASA-TLX	82.6	No	Home + Clinical	Heart Failure	Partial (observed)	[13]
Older adults	SUS, acceptability score	59.7	No	Clinical	Gait monitoring	No	[17]
Older adults	SUS, feedback	58.3	No	Lab + Clinical	Fall prevention	No	[15]
Patients	SUS	87.5	No	Clinical	Symptom tracking	No	[45]
Health care professional	SUS	69.2	No	Geriatric wards	Health monitoring	No	[16]
Older adults	SUS + Interview	78.8	No	Home	Home care platform	No	[46]
Patients + Clinicians	SUS + Feedback	86.8	No	Clinical	Trial data capture	No	[48]
Emergency staff	SUS + Interview	53.1	No	Clinical	Clinical information management	No	[47]
Our study (informal caregivers)	SUS + System usage logs	80.45	Activity logs	Home	Decision support in older adult care	Task metrics, user engagement score	—[b]

[a]SUS: System Usability Scale.

[b]Not applicable.

XSL•FO
RenderX

Methods

Study Design

This study evaluated the usability of the EDT system for supporting informal caregivers in home-based care. To provide a comprehensive assessment of user interaction, both subjective and objective measures were integrated. The following sections describe the study design, system features, participant recruitment, testing procedures, and analytical methods used to assess usability.

Overall System Flow

We used a cross-sectional observational design to evaluate the usability of the caregiver-facing EDT prototype. Figure 1 illustrates the system architecture and data flow during the usability sessions. Older adult participants wore a Fitbit smartwatch to collect physiological data, including heart rate, sleep, and activity measures. These data were transmitted to the cloud for secure transfer and preprocessing and then processed and stored by the back end for analysis and insight generation. Caregivers accessed the EDT system through computers, tablets, iPads, or smartphones, where they viewed real-time and historical health information, including digital biomarkers and artificial intelligence (AI)–generated recommendations. During the sessions, informal caregivers completed representative tasks such as exploring dashboards, reviewing AI-based recommendations, and navigating system features. User interactions were recorded through back-end logs, and caregiver feedback was collected using a questionnaire. Sample user interfaces of the system are presented in Figure 2 for cardiac DT and Figure 3 for sleep stage monitoring, while detailed system feature interfaces are provided in the Multimedia Appendix 1.

Figure 1. Illustration of the system usability session setup, showing data flow from a home-based older adult to a digital twin platform accessed by an informal caregiver.

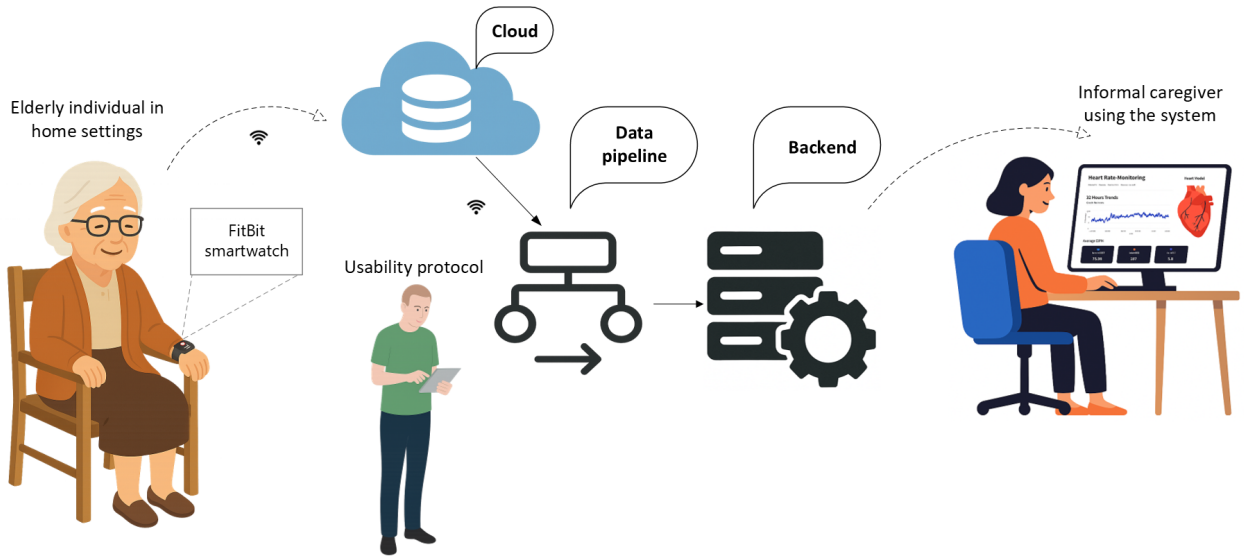
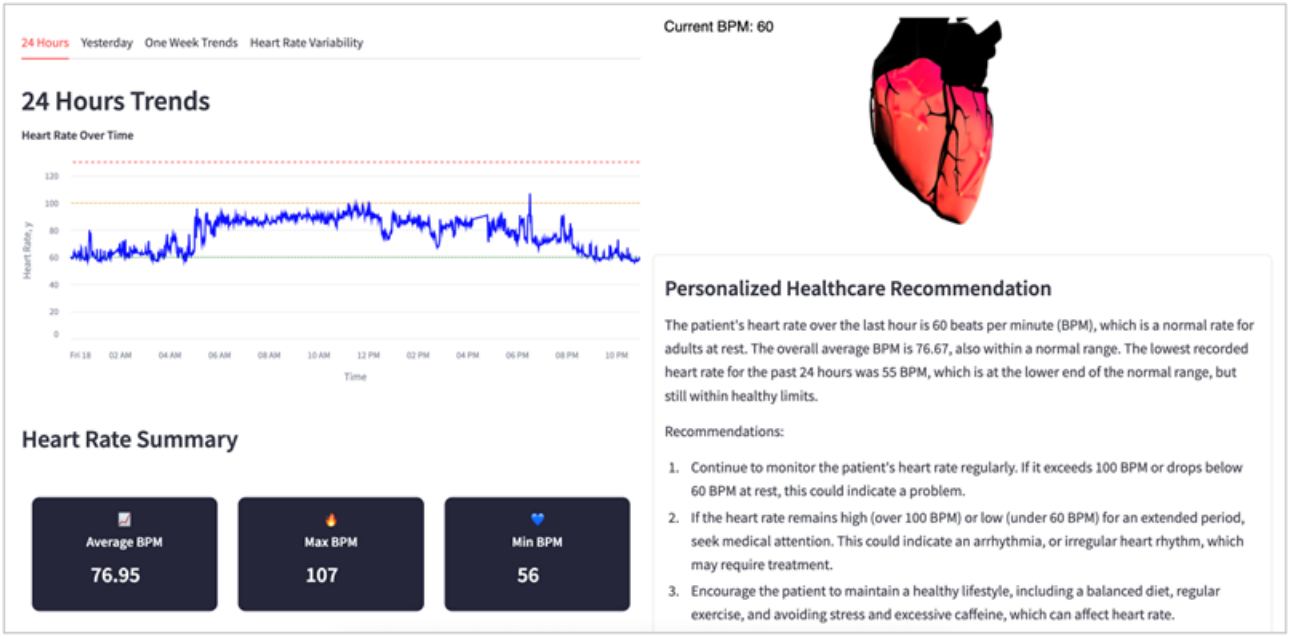


Figure 2. Cardiac Digital Twin model user interface.



XSL•FO
RenderX

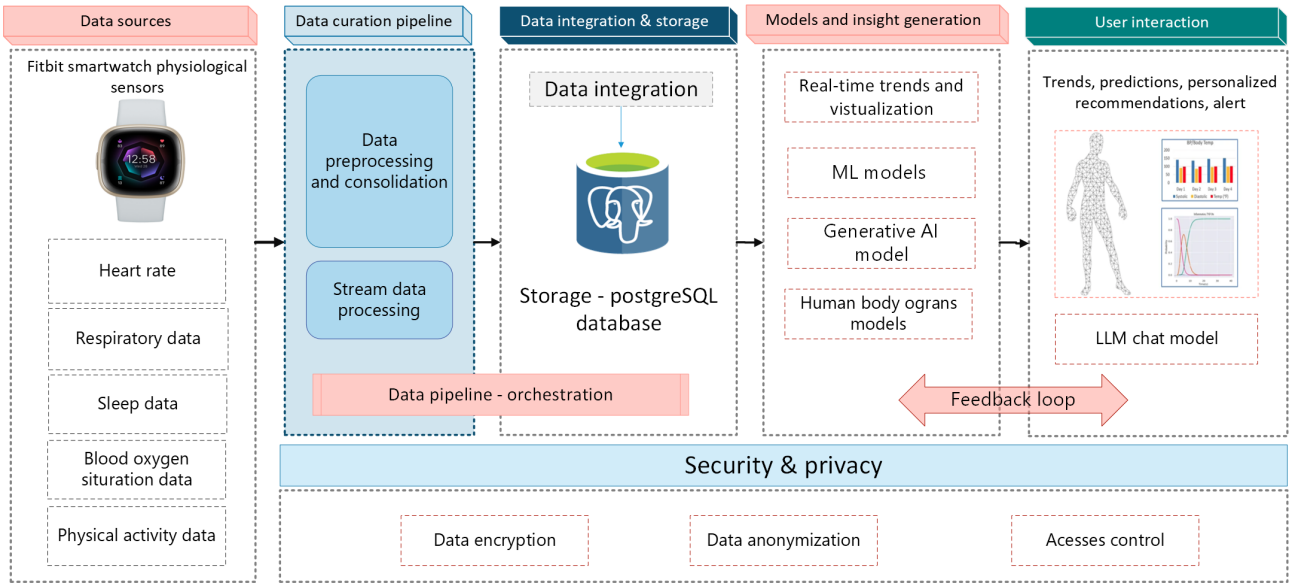
Figure 3. Sleep stages analysis user interface.



System Features

The EDT system evaluated in this study was designed to provide informal caregivers with real-time, personalized insights into the health status of older adults in their homes. The system follows a modular architecture with 5 core components: data sources, data curation pipeline, data integration and storage, models and insight generation, and user interaction, with a dedicated security and privacy layer and a dynamic feedback loop between the user and intelligent models, as shown in Figure 4.

Figure 4. High-level architecture of elderly digital twin system. AI: artificial intelligence; LLM: large language model; ML: machine learning.



Data Sources

The system captures multimodal physiological data using a Fitbit smartwatch worn by older adults. These signals included heart rate, respiratory data, sleep log data, blood oxygen saturation (SpO_2), and physical activity.

Data Curation Pipeline

Incoming sensor data are streamed into an automated preprocessing module that handles data cleaning, synchronization, and consolidation across modalities. A dedicated stream processing engine prepares the data for integration and downstream analysis in real time. This stage ensures quality and consistency before the data are forwarded through an orchestrated pipeline.

Data Integration and Storage

The processed data were stored centrally in a PostgreSQL database. This layer supports structured data storage and efficient querying for both real-time and retrospective analyses. Integration mechanisms ensure temporal alignment of physiological signals for model consumption.

XSL•FO
RenderX

Models and Insights Generation

The EDT system comprises 2 main components: a Cardiac Digital Twin and a Sleep Monitoring Digital Twin, which monitor physiological signals and generate insights for caregivers. The EDT prototype was evaluated as a caregiver-oriented decision support research system rather than as a diagnostic tool, and the sensing device used in this study, Fitbit Sense 2, is not a Food and Drug Administration–approved device for clinical use. Cardiac Digital Twin tracks heart health in older adults through real-time heart rate monitoring, trend visualization, and HRV analysis. A 3D cardiac model simulates a beating heart with animation driven by current beats per minute. For continuous monitoring during data loss, a bidirectional long short-term memory (Bi-LSTM) model predicts beats per minute at 5-minute intervals. Bi-LSTM was selected because of its ability to capture temporal dependencies and improve physiological prediction accuracy. When activated, the system notifies the caregivers that synthetic predictions are being used. The Bi-LSTM model was trained in 579,486 heart rate records, achieving a mean squared error of 0.2944 and a mean absolute error of 0.3410. The reported model performance metrics are presented to characterize the technical behavior of the prototype components and should not be interpreted as evidence of clinical-grade validity.

The Sleep Monitoring Digital Twin processes multisensory physiological data to monitor sleep behavior. It analyzes patterns across sleep stages, including light, deep, rapid eye movement, and awake, and visualizes trends to support long-term observation. The system computes sleep quality metrics, such as efficiency, latency, and awakenings, using probabilistic models to map the stage transitions. These outputs enable the detection of sleep disruptions and behavioral anomalies linked to health decline in the older adult population. A long short-term memory model was developed to infer sleep stages from SpO_2 and heart rate data, chosen for its memory structure suitable for temporal dependencies in physiological time series. The model achieved 92% validation accuracy for 9125 sequences, maintaining sleep-tracking continuity. In this study, however, the sleep-monitoring model was included to enable the functional operation of the caregiver-facing EDT interface and was evaluated as part of the overall usability workflow rather than as a stand-alone clinically validated diagnostic model. A submodule monitored SpO_2 levels to detect hypoxemia, triggering alerts and large language model (LLM) feedback when oxygen saturation fell below the critical threshold. The LLM outputs are intended to support caregiver awareness and follow-up, rather than provide medical diagnosis or treatment advice.

To enhance real-time interpretability and caregiver empowerment, both DT models were supported by a fine-tuned GPT-4o LLM through instruction-based fine-tuning in caregiving. This LLM module delivers context-aware personalized feedback by interpreting physiological indicators, including abnormal heart rate patterns, HRV, disrupted sleep trends, sleep efficiency, latency, awakening, and hypoxemia through SpO_2 monitoring. When the system detects anomalies such as elevated heart rate, low HRV indicating stress, oxygen saturation below thresholds, or deviations from sleep benchmarks, it triggers the LLM to generate clear, caregiver-facing recommendations. These responses are personalized based on recent health trends and are presented in an understandable language. LLM summarizes patterns and provides decision support recommendations for caregiver actions such as monitoring hydration or seeking clinical evaluation. This fine-tuned GPT-4o model translates complex sensor data into actionable guidance, strengthens decision support capabilities, and fosters informed caregiving responses.

In our previous study, the caregiver-facing recommendations were subjected to structured medical experts' evaluation across the dimensions of clinical appropriateness, usefulness, safety, and alignment with clinical guidelines, with overall positive ratings and substantial interrating agreement, supporting their use as decision support guidance in the present system [18]. To reduce the hallucinations, the GPT-4o model was developed using instruction-based fine-tuning in caregiving and grounded in structured physiological indicators from the EDT. The model was activated only when anomalies or threshold deviations were detected. The decision support recommendation generation was constrained through restricted prompting and limited output scope. Its outputs were designed as caregiver-friendly guidance and human-readable summaries of observed sensor patterns and not diagnosis or treatment advice. The system also used escalation language, such as advising clinical evaluation for concerning cases.

User Interaction

The system provides a web-based interface for informal caregivers that is accessible through desktops, smartphones, and tablets. The dashboard enables navigation across modules such as heart rate trends, sleep analysis, SpO_2 monitoring, and alerts. This alert section presents LLM-generated alert messages and recommendations when abnormal physiological patterns or threshold violations are detected. Visualizations use color-coded indicators and summaries to reduce cognitive load. The GPT-4o chatbot further generates natural language feedback and recommendations based on health data and caregiver queries. Caregivers can review alerts, add notes, and interact with the system across devices, thus ensuring flexible and informed older adult care support.

Feedback Loop for Adaptive Learning

The architecture incorporates a dynamic feedback loop based on caregiver responses and system logs. These feedback signals are used to retrain the detection thresholds, improve alert relevance, and fine-tune the language model for more accurate and useful recommendations.

Security and Privacy

A dedicated privacy layer ensures compliance with data protection standards. It includes data encryption, access control mechanisms, and data anonymization to protect sensitive health data and support the deployment of ethical systems.

Participants Recruitment

This study included 2 groups of participants: older adults and their informal caregivers. Six older adult participants were recruited to generate real-world physiological data for the

XSL•FO
RenderX

validation of the monitoring function of the system. These individuals, aged 60 - 85 years, were Thai nationals residing in home care environments and were cognitively and physically capable of using a Fitbit Sense 2 device. Patients with stable chronic health conditions were included in the study, while those with known skin sensitivities, cognitive impairments, or conditions interfering with the use of wearable devices were excluded.

Fifty informal caregivers were recruited for the usability evaluation using a purposive sampling strategy to ensure the representation of individuals actively engaged in home-based older adult care. Table 2 summarizes the demographic characteristics of the informal caregiver participants. Eligible caregivers were adults aged 18 - 55 years with at least 3 months' caregiving experience. To support meaningful interaction with the system, participants were required to demonstrate basic digital literacy (eg, prior use of smartphones or web apps) and the ability to navigate English-language interfaces. Screening interviews confirmed these criteria prior to enrollment. For better recruitment and participant comfort, all initial communications and consent discussions were conducted in the local Thai language with caregivers and older adult individuals. A local Thai assistant was engaged to explain the study procedure and provide step-by-step system guidance in Thai, ensuring clarity and inclusivity during the sessions. Caregivers received a stipend of 100 Thai Baht (average THB to US dollar exchange rate in July 2025: THB 1=US $0.03082; THB 100=US $3.082) to acknowledge their time and contribution. The Fitbit Sense 2 smartwatch used in this study was separately provided by the research team to all older adult participants.

Table . Demographic characteristics of informal caregiver participants.

Characteristic	Caregivers (n=50)
Age (years), mean (SD)	35.6 (7.6)
Age (years), range	18 - 55
Age (years), group distribution, n (%)	
18 - 25	3 (6)
26 - 35	25 (50)
36 - 45	15 (30)
46 - 50	6 (12)
51 - 55	1 (2)
Sex, n (%)	
Male	21 (42)
Female	29 (58)
Caregiving experience (years), n (%)	
<1	11 (22)
1 - 3	21 (42)
4 - 6	12 (24)
>6	6 (12)

Ethical Considerations

The study protocol was reviewed and approved by the institutional review board of King Mongkut's University of Technology Thonburi under reference number KMUTT-IRB-COA-2025 - 048 on July 7, 2025. Written informed consent was obtained from all participants prior to data collection. Specifically, informed consent was secured from both the older adult participants and their informal caregivers before their involvement in the study. Caregiver participants received a small stipend of THB 100, equivalent to approximately US $3.08, as compensation for their time. This compensation was not linked to task performance, usability ratings, or study outcomes and was not considered coercive. All collected data were deidentified prior to analysis, and participant confidentiality was maintained throughout the study.

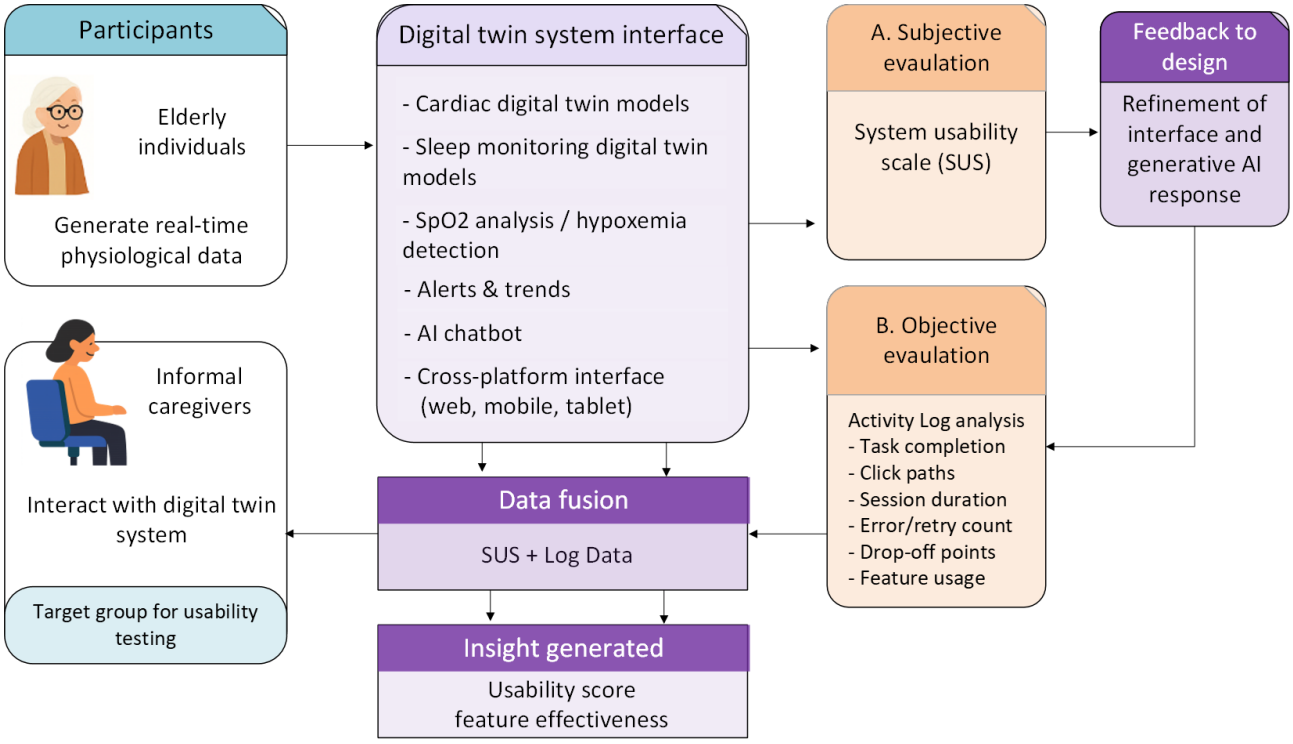
Usability Evaluation Framework

This study used a dual-method usability evaluation framework designed to capture subjective perceptions and objective behavioral interactions with the EDT system. The framework, illustrated in Figure 5, consists of 3 key components: participant roles, system interaction, and evaluation methods, integrated through a feedback-driven design loop. Two participant groups were involved: older adults who contributed real-time physiological data using a Fitbit Sense 2 smartwatch, and informal caregivers who interacted with the system to evaluate its usability. Usability was assessed using both subjective and objective methods. Subjectively, SUS was used to collect caregiver-reported feedback on ease of use, satisfaction, and overall system experience. Objectively, the activity log data captured granular interaction behaviors, including task completion rates, click paths, session duration, error rates, drop-off points, and module usage frequency. These 2 data streams were then fused to generate comprehensive usability

XSL•FO
RenderX

insights. This fusion enabled us to assess both how users perceived the system and how they interacted with it.

Figure 5. Usability evaluation framework for the elderly digital twin system. The framework integrates subjective (System Usability Scale) and objective (log-based) evaluation methods to assess system usability and inform design improvements. AI: artificial intelligence.



Data Collection

Data Collection was conducted over an 8-week period (July 7 to August 31, 2025) and involved 2 participant groups: older adults and informal caregivers. The older adult participants wore the Fitbit Sense 2 smartwatch in their home environments to generate real-time physiological data, including heart rate, respiratory rate, SpO_2, sleep logs, and physical activity. Before commencing data collection, all older adult participants completed a brief health and comfort assessment to confirm their suitability for using a smartwatch. The survey gathered basic demographic information and screened for potential risks, such as allergies, history of skin irritation, relevant medical conditions, or wrist swelling, which could affect comfort or safety. Participants with contraindications were excluded. The results of the health and comfort screening are summarized in Table 3. This prescreening step ensured that the smartwatch could be used safely and comfortably during subsequent data collection.

Table . Health and comfort screening outcomes for older adult participants.

Screening item	Yes/No, n (%)
Allergy to smartwatch materials (silicone, metal, and adhesives)	0 (0)/6 (100)
History of skin irritation from wristband/watches/jewelry	0 (0)/6 (100)
Skin/medical conditions affecting smartwatch use	0 (0)/6 (100)
Swelling in wrist/hands	0 (0)/6 (100)
Physical activity	Daily: 3 (50), occasionally: 1 (16.7), and rarely: 2 (33.3)

The data were streamed to the system back end to validate the predictive analytics and monitoring capabilities of the EDT models under real-life conditions. Informal caregivers evaluated the system usability through guided interaction sessions using their smartphones, tablets, or computers. Each caregiver used the EDT system for 4 consecutive days, engaging with it for 20 - 30 minutes daily. The SUS questionnaire was administered after the fourth day to capture the caregivers' consolidated perception of usability after repeated interactions. This approach balanced ecological validity with participant burden: 4 days ensured sufficient system interaction while remaining feasible for busy caregivers. We also captured users' behavioral data through system-generated activity logs, which were passively recorded through automatic instrumentation. The system obtained informed consent prior to log data collection, disclosing that no personally identifiable information was recorded. All logged data were anonymized and stored per institutional guidelines.

A total of 24 system tasks were predefined for the log data collection phase. These tasks were systematically constructed from the core functional modules of the EDT prototype,

XSL•FO
RenderX

including heart monitoring, sleep tracking, SpO_2 assessment, AI-driven health advice, and the AI chatbot assistant. The task set was designed to ensure structured coverage of the system's principal functions during usability testing, so that all key user interactions could be consistently captured in the log analysis. Thus, the tasks were intended primarily to exercise the major features of the prototype rather than to directly reproduce naturally occurring caregiving workflows derived from interviews, observational studies, or clinical workflow analysis. Each task represented a meaningful user interaction within the prototype (eg, viewing 24-hour heart rate trends or checking sleep stage analysis), and each was assigned a unique code for consistent identification in the analysis, as detailed in Table 4.

XSL•FO
RenderX

Table . System tasks and associated codes used for feature usage analysis during the user activity log data collection.

Model and task	Logged behavior	End role	Associated objective metrics
Cardiac Digital Twin			
Request AI[a] insights for real-time heart rate	Click on AI insights	AI output shown	TC[b], ToT[c], FU[d]
Observe real-time heart rate dashboard	Open live HR[e] screen	Exit/switch page	TC, ToT, RR[f], NF[g], FU
View 24-hour heart rate history	Open 24-hour chart	Exit/switch page	TC, ToT, NF, TAR[h], FU
View 7-day heart rate trends	Open 7-day HR trends	Exit/switch page	TC, ToT, FU
View 24-hour heart rate summary	Open HR summary panel	Exit/switch page	TC, ToT, TAR, FU
Observe heart rate variability dashboard	Open HRV[i] panel	Exit/switch page	TC, RR, ToT, FU
Request AI insights on heart rate variability	Click AI for HRV	AI output shown	TC, FU, CTR[j]
Sleep Monitoring Digital Twin			
View last night's sleep quality	Open sleep summary	Exit/switch page	TC, NF, TAR, ToT
Request AI advice for abnormal deep sleep	Click AI for deep sleep	AI output shown	TC, FU, CTR, ToT
Request AI advice for low light sleep	Click AI for light sleep	AI output shown	TC, FU, CTR, ToT
Request AI advice for REM[k] sleep levels	Click AI for REM sleep	AI output shown	TC, FU, CTR, ToT
View sleep metrics dashboard	Open sleep metrics	Exit/switch page	TC, ToT, TAR
View sleep stage timeline	Open stage timeline	Exit/switch page	TC, NF, CTR, FU
View sleep benchmark reference	Open benchmark panel	Exit/switch page	TC, NF, TAR
Review sleep benchmark interpretation	Open benchmark explanation	Exit/switch page	TC, ToT, RR, FU
Request AI advice on sleep benchmark	Click AI for benchmark	AI output shown	TC, session duration, TAR, FU, ToT, CTR
Open and generate sleep trend report	Open/generate report	Report shown	TC, session duration, TAR, FU
View weekly sleep log	Open weekly sleep log	Exit/switch page	TC, NF, TAR, FU
View sleep fragmentation and stage transitions	Open fragmentation view	Exit/switch page	TC, NF, ToT, FU
Request AI advice on sleep fragmentation	Click AI for fragmentation	AI output shown	TC, FU, session duration, ToT
SpO_2 Monitoring			
View SpO_2 analysis dashboard	Open SpO_2 analysis	Exit/switch page	TC, ToT, NF, FU
Observe real-time SpO_2 for hypoxemia monitoring	Open live SpO_2 screen	Exit/switch page	TC, RR, TAR, FU
Request AI advice on SpO_2 abnormalities	Click AI for SpO_2	AI output shown	TC, FU, CTR, ToT
AI Assistant			
Submit query to AI health chatbot	Enter/send prompt	Response shown	TC, session duration, RR, CTR, TAR, FU

[a]AI: artificial intelligence.

[b]TC: task completion.

[c]ToT: time on task.

[d]FU: feature usage.

XSL•FO
RenderX

[e]HR: heart rate.

[f]RR: retry rate.

[g]NF: navigation frequency.

[h]TAR: task abandonment rate.

[i]HRV: heart rate variability.

[j]CTR: click-to-task ratio.

[k]REM: rapid eye movement

The mapping tasks for objective metrics were derived from the functional characteristics of each task. Metrics such as task completion and time on task were applied universally as core usability indicators. The retry rate was assigned to tasks with likely repeated attempts in monitoring and chatbot interactions. Feature usage was linked to AI-driven features, demonstrating user engagement beyond basic monitoring. Navigation frequency was applied to tasks involving timelines and log exploration. Session duration was associated with tasks requiring extended engagement, such as report generation. The click-to-task ratio was included for tasks with multiple interaction steps. The task abandonment rate was assigned where drop-offs were likely, such as when viewing summaries and reports. This mapping ensures the replicability of linking system interactions to the 8 objective usability metrics.

To ensure that the objective usability measures reflected actual user behavior, each logged task was operationalized as a concrete interface-level action. Tasks such as view and observe were defined as opening and accessing a specific dashboard, chart, or panel until the user exited or switched pages. AI-related tasks were defined as explicit user-triggered actions, such as clicking an AI insight or advice function or submitting a chatbot query, with completion recorded when the system response was displayed. In Table 4, open indicates that the user accessed a dashboard, chart, panel, or report. Click AI indicates an explicit user request for AI-generated insights or advice. Exit/switch page indicates that the user left the target screen or moved to another module. AI output shown indicates successful system generation and display of the requested response.

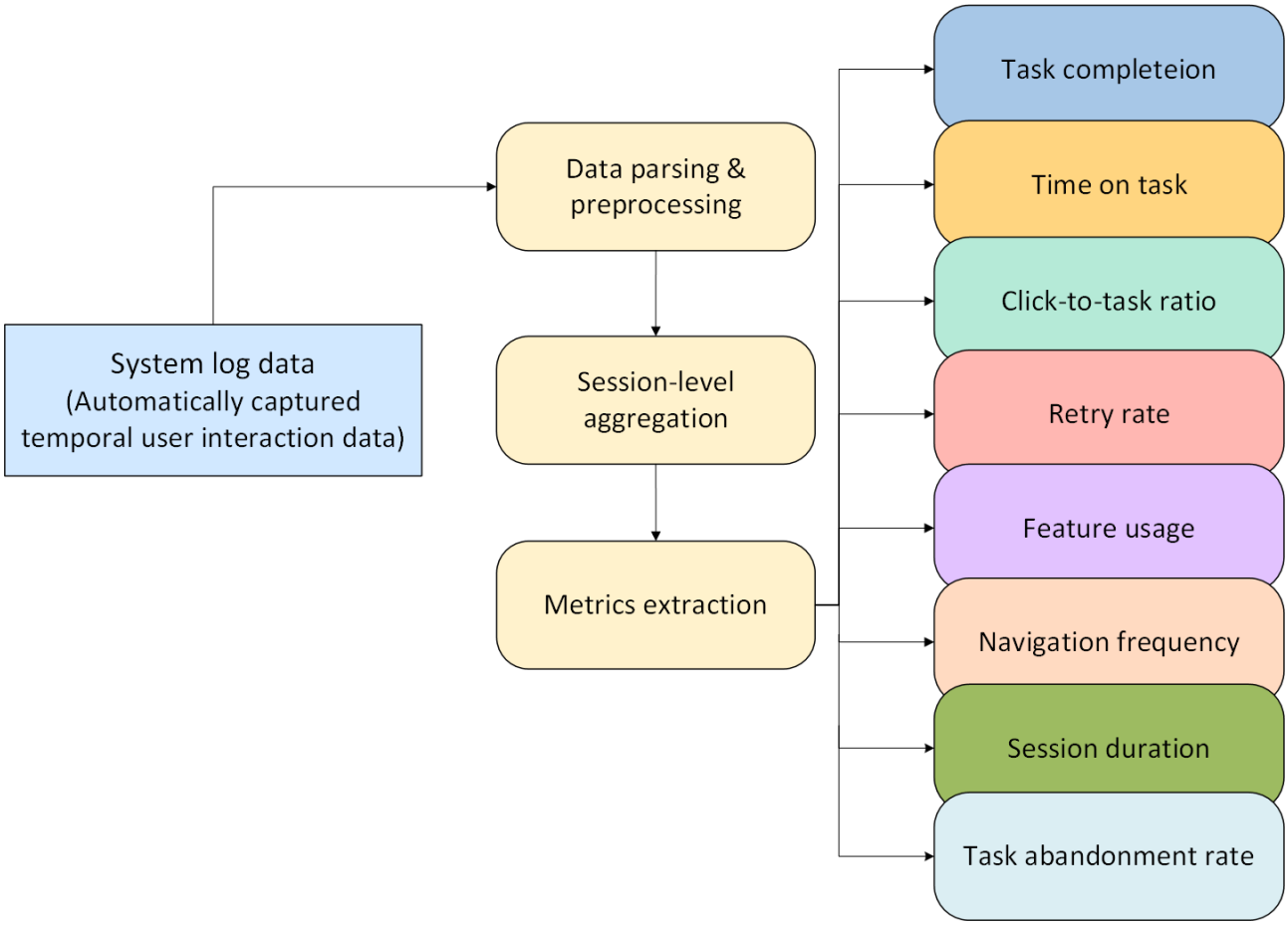
Data Analysis

User Activity Log Analysis

User activity logs were extracted from the system back end for all caregiver interaction sessions that were conducted during the study period. Each log entry included a time stamp, session identifier, user identifier, task name, action type (eg, started, completed, error, retry, and abandoned), task duration (in seconds), navigation events, and click events. As shown in Figure 6, these raw entries were processed at the session-level metrics, aggregated across sessions for each user, and normalized by relevant totals (eg, total tasks started) to enable cross-user comparisons.

XSL•FO
RenderX

Figure 6. Sequential steps for transforming raw log data into quantified usability measures.



Metrics Computation

The raw interaction logs were parsed and preprocessed, aggregated at the session level, and used for metric extraction. From this pipeline, key objective behavioral indicators were computed, including task completion (TC), time on task (ToT), click-to-task ratio (CTR), retry rate (RR), feature usage (FU), navigation frequency (NF), session duration, and task abandonment rate (TAR). TC is the percentage of started tasks that were completed. This metric was calculated using equation 1.

(1)TC(%)=CiSi×100

where C_i is the number of tasks of user i with completed status and S_i indicates the number of tasks with started status of user i. ToT measures how long it takes a user to begin or complete a task once they start interacting with the system. It is defined as:

(2)ToT=∑i=1ndin

where d_i is the duration of the $i - th$ completed task and n is the number of completed tasks. The RR measures how often users had to repeat a task before succeeding. For each task, if a user has multiple STARTED events before success, everything after the first event is a retry, defined as follows:

(3)RRi=riai

where r_i is the number of retries for user i repeated a task after a failed attempt (errors, abandoned, and reclicks) and a_i the total number of tasks the user initiated. FU counts the interaction of users per system feature and was calculated as, $FU_x = N_x$, where, N_x is the count of feature x in the navigation logs. Session duration was calculated as follows:

(4)SD=tend−tstart

where t_{start} and t_{end} are the first and last actions in a session, respectively. The TAR is the percentage of tasks started but never completed and is defined as follows:

(5)TAR(%)=(AS)×100

where A is abandoned tasks. CTR is the number of clicks per completed task, defined as CTR = K/C, where K is the click event. NF is the number of times each navigation step/module is used, defined as $NF_x = N_x$. All metrics were calculated per session and aggregated for user and cross-user analysis. Table 5 presents the descriptive statistics of the 8 objective usability metrics collected from caregivers' interactions with the EDT system.

XSL•FO
RenderX

Table . Descriptive statistics of objective usability metrics.

Metric	Minimum	Maximum	Mean (SD)
TC[a] (%)	82.93	98.02	94.08 (4.10)
RR[b]	0.00	1.00	0.25 (0.28)
ToT[c] (seconds)	39.20	175.67	89.16 (21.97)
TAR[d] (%)	0.00	14.13	2.66 (2.87)
CTR[e]	0.12	1.32	0.48 (0.22)
NF[f]	5.00	12.00	11.02 (1.39)
Session duration (minutes)	36.97	171.90	101.59 (30.17)
FU[g] (event count)	52	98	54 (13.4)

[a]TC: task completion.

[b]RR: retry rate.

[c]ToT: time on task.

[d]TAR: task abandonment rate.

[e]CTR: click-to-task ratio.

[f]NF: navigation frequency.

[g]FU: feature usage.

To establish the relationship between objective usability metrics derived from user activity logs and the dimensions of SUS, we mapped each metric to the SUS factors originally proposed by [49,50]. Their analysis suggested that the SUS can be interpreted as comprising 2 correlated but distinct dimensions: usable (covering items related to effectiveness, efficiency, and overall ease of use) and learnable (covering items reflecting the ease of learning and initial onboarding, specifically items 4 and 10).

The mapping of objective usability metrics to the SUS dimensions is presented in Table 6. In this study, we adopted the 2-factor structure of SUS (usability and learnability) as proposed by Borsci et al [49] and Lewis and Sauro [50], which is the most widely validated and parsimonious model compared with alternative factor structures. In this mapping, performance-related measures such as task completion, efficiency, engagement, and error-related indicators are aligned with the usability factor, as they reflect effectiveness and overall ease of use. In contrast, retry-related measures are aligned with the learnability factor, as they directly capture the effort required for users to acquire proficiency during initial interactions. This mapping provides a conceptual bridge between objective behavioral-based metrics and subjective usability perceptions, ensuring that activity log analysis contributes to the broader usability assessment framework represented by Borsci et al [49] and Lewis and Sauro [50].

XSL•FO
RenderX

Table . The definitions of the usability metrics derived from system activity logs.

Objective metric	Relevant usability aspect (from literature)	Mapped SUS factor (Lewis and Sauro [50])	References
TC[a]	Effectiveness (core usability outcome)	Usable (items 1, 3, 5, 7, 8, and 9)	[49-52]
ToT[b]	Efficiency	Usable (efficiency items contribute to overall ease-of use perceptions)	[49,50,53]
RR[c]	Learnability (ease of initial learning, errors before success)	Learnable (items 4 and 10: technical support and things to learn)	[49,50,52]
FU[d]	Discoverability (linked to usability breadth)	Usable (items 5 and 9: confidence and integration)	[49,50,54]
Session duration	Engagement	Usable (general usability experience, confidence, ease, and satisfaction)	[49,50,52]
TAR[e]	Usability issues/unclear flow	Usable (same rationale as drop-off points)	[49,50,55]
CTR[f]	Efficiency	Usable (efficiency dimension of usability)	[49,50,52]
NF[g]	User flow analysis	Usable (task integration, item 5 “functions well integrated”)	[49,50,56]

[a]TC: task completion.

[b]ToT: time on task.

[c]RR: retry rate.

[d]FU: feature usage.

[e]TAR: task abandonment rate.

[f]CTR: click-to-task ratio.

[g]NF: navigation frequency.

User Engagement Based on Composite Behavioral Metrics

To capture a more holistic measure of user interaction with the EDT system, we computed a user engagement score (UES) by incorporating multiple behavioral indicators that were derived from the system logs. Specifically, 6 objective metrics were included: feature usage, average session duration, session frequency, task completion rate, error rate, and navigation diversity. Each captures a distinct aspect of usability: task efficiency, effectiveness, temporal aspects, and behavioral diversity [51-54,56]. This composite scoring approach is essential for usability studies because it provides a multidimensional perspective on user interaction. These holistic measures are recognized as valid indicators of sustained adoption and user satisfaction [57,58]. To derive the composite engagement score, we initially extracted 8 metrics from the system log. Because not all metrics uniquely represented user engagement, we retained 6 core metrics that directly reflected the breadth, depth, and effectiveness of user interactions. These metrics included FU, ToT (seconds), total session duration (seconds), TC, error count, and NF. Each metric was then normalized to a (0, 1) scale using the maximum value observed in the dataset. The normalized values were subsequently combined using equal weights to produce a composite engagement score, as presented in equation 5.

$$(6)\text{Engi}=\alpha\text{Fi}+\beta\text{Si}+\gamma\text{Ri}+\delta\text{Ci}+\ \text{Ei}+\zeta\text{Di}\alpha+\beta+\gamma+\delta+\ +\zeta$$

where F_i is the normalized feature usage for user i, S_i indicates normalized session duration, R_i is the normalized ToT, C_i shows the normalized values of TC, E_i is normalized error rate, D_i is the normalized navigation diversity, and α, β, γ, δ, ε, and ζ symbols show the weights. All weights were set to be equal ($\alpha = \beta = \gamma = \delta = \varepsilon = \zeta = 1$) so that each engagement dimension contributed uniformly to the composite score. An equal-weight additive formulation was selected as a pragmatic postnormalization approach to combine the engagement metrics while avoiding bias or subjective prioritization in the absence of empirical evidence regarding their relative importance. Prior literature on composite scoring has similarly used unit weights as a baseline when neither theoretical nor empirical justification exists for different weighting [59-61]. Because formal testing of interdependence or dimensional structure among the component metrics was not performed in this study, the resulting engagement score should be interpreted as an operational composite summary rather than a formally validated latent construct. For interpretability, k-means clustering with k=3 was then applied to classify users into low-, medium-, and high-engagement groups.

The SUS

The usability of the EDT system was evaluated using the SUS, a standardized 10-item questionnaire widely used to assess the perceived usability of interactive systems. Each item was rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). SUS data were collected from 50 informal

XSL•FO
RenderX

caregivers. To compute individual SUS scores, responses to odd-numbered items were scored as the participant's rating minus 1, whereas responses to even-numbered items were scored as 5 minus the rating. The adjusted item scores were summed and multiplied by 2.5, yielding total SUS scores ranging from 0 to 100. In addition, item-level descriptive statistics were computed for all 10 items. Internal consistency was evaluated using Cronbach α [62] after reverse-scoring the negatively worded items so that higher scores consistently indicated greater perceived usability, as defined in equation 6.

(7)α=kk−11−∑i=1kσi2σT2

where *k* is the number of items, σi2 is the variance of each individual item, and σT2is the variance of total score (sum across all items).

Correlation and Regression Analysis of Engagement and Perceived Usability

To investigate the relationship between the composite behavioral metric and perceived usability, we first examined the correlation between the UES and the SUS. The normality of both variables (UES and SUS) was tested using the Shapiro-Wilk test. As SUS scores deviated from normality, Spearman rank correlation (ρ) was selected as the primary measure of association, and Pearson correlation (*r*) was also reported for completeness. Bootstrapped 95% CIs were computed to quantify the precision of the correlation estimates. Following the correlation analysis, we conducted a simple linear regression to test whether the UES predicts the SUS. The regression was specified as follows:

(8)SUSi=β0+β1·UESi+ i

where SUS_i is the usability score for user *i*, UES_i is the engagement score, β_0 is the intercept, β_1 is the regression coefficient, and ε_i is the error term. Model fit was evaluated with R^2, F-statistics, and coefficient significance. This 2-step approach enabled us to first establish the strength and direction of the association between user engagement and usability and then the predictive power of the UES for the SUS, thereby linking objective behavioral data with subjective ratings.

Results

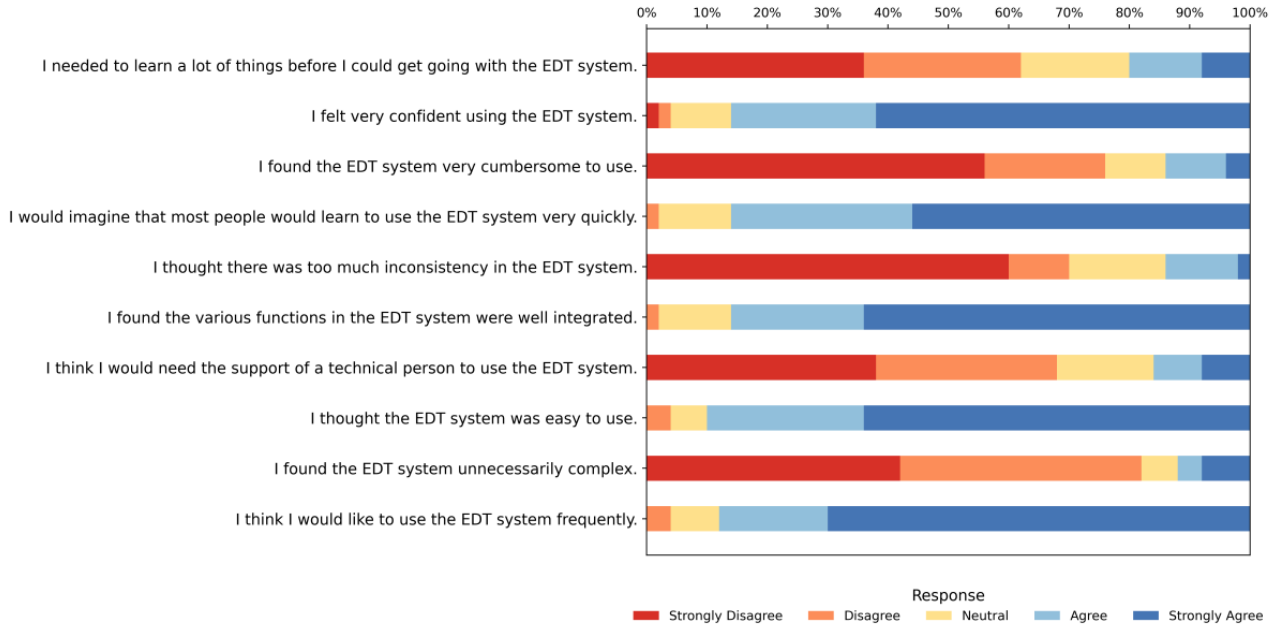
Overview

This study assessed the EDT system usability among informal caregivers by analyzing objective user interactions and subjective perceptions of usability (SUS). Three aspects were analyzed: (1) SUS survey results reflecting perceived usability, (2) patterns of composite UESs from system logs, and (3) the relationship between user engagement and usability using correlation and regression analyses. These findings address the study aim of assessing EDT system usability using objective behavioral metrics and subjective measurements and examining whether behavioral engagement indicates perceived system usability.

SUS Results

Item-level response patterns for the 10 SUS items are presented in Figure 7, with descriptive statistics shown in Table 7. Overall, caregivers responded positively to the positively worded items and showed low agreement with the negatively worded items, indicating favorable usability perceptions of the EDT system. Higher mean scores were observed for confidence in using the system, integration of system functions, and intention to use it frequently, whereas lower engagement on negatively phrased items suggested limited perceived complexity, inconsistency, and need for technical support. The SUS also demonstrated strong internal consistency, with a Cronbach α of 0.87, indicating reliable measurement of perceived usability.

Figure 7. Distribution of responses to the 10 System Usability Scale items among caregivers (n=50). Percentages are shown on the top axis (0% - 100%). EDT: elderly digital twin.



XSL•FO
RenderX

Table . Item-level descriptive statistics for the 10 System Usability Scale questions.

Item	Mean (SD)	Minimum	Maximum
Q1	4.54 (0.8)	2.0	5.0
Q2	1.96 (1.18)	1.0	5.0
Q3	4.50 (0.79)	2.0	5.0
Q4	2.18 (1.26)	1.0	5.0
Q5	4.48 (0.79)	2.0	5.0
Q6	1.86 (1.20)	1.0	5.0
Q7	4.40 (0.78)	2.0	5.0
Q8	1.86 (1.20)	1.0	5.0
Q9	4.42 (0.91)	1.0	5.0
Q10	2.30 (1.30)	1.0	5.0

The overall SUS scores (N=50) ranged from 42.5 to 100, with a mean of 80.45 (SD 17.6), indicating a generally favorable perception of usability by caregivers. As shown in Table 8, the SUS scores varied descriptively across the caregiving experience levels. Caregivers with more than 6 years of experience reported the highest mean score (mean 86.7, SD 21.8), followed by those with 1 - 3 (mean 83.5, SD 14.5) years and 4 - 6 (mean 81.0, SD 16.8) years of experience. Caregivers with less than 1 year of experience reported the lowest mean score (mean 70.9). A 1-way ANOVA was used to test whether these differences were statistically significant. Although the descriptive pattern suggested higher usability perception among more experienced caregivers, the ANOVA results were not statistically significant ($F_{3,46}$=1.58; P=.208, η^2=0.093). The effect size indicates that approximately 9.3% of the variance in the SUS scores was explained by caregiving experience, reflecting a small to moderate effect.

Table . Mean System Usability Scale scores by caregiving experience and age group.

Category (group)	Mean (SD) SUS[a] score[b]
Caregiving experience (years)	
<1	70.9 (20.3)
1 - 3	83.5 (14.5)
4 - 6	81.0 (16.8)
>6	86.7 (21.8)
Age group (years)	
18 - 25	80.8 (17.0)
26 - 35	80.8 (16.1)
36 - 45	77.0 (21.6)
46 - 50	83.3 (18.4)
51 - 55	92.5 (0.0)

[a]SUS: System Usability Scale.

[b]Overall mean score 80.45 (SD 17.65).

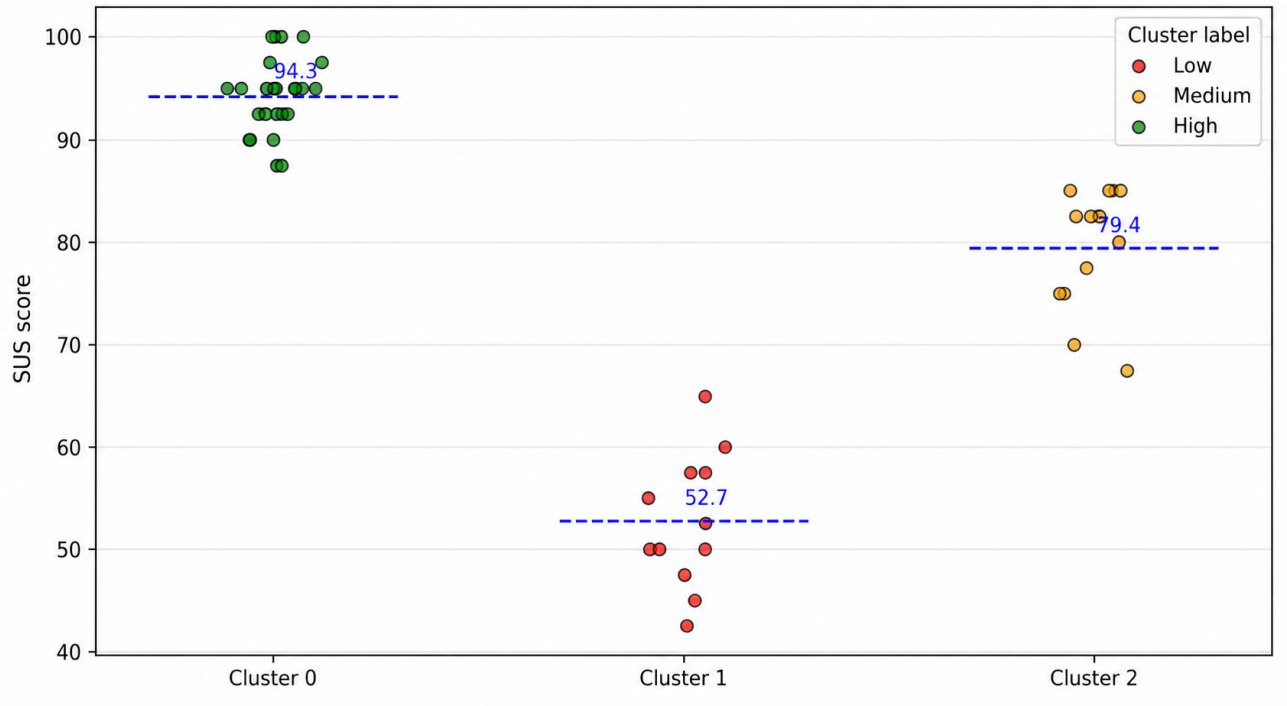
SUS scores also showed descriptive variation across age groups, as presented in Table 8. Caregivers aged 51 - 55 years reported the highest mean score (mean 92.5, SD 0.0), while those aged 36 - 45 years reported the lowest (mean 77.0, SD 21.6). Younger caregivers (aged 18 - 25 and 26 - 35 years) reported similar mean scores (mean 80.8), and caregivers aged 46 - 50 years also rated usability favorably (mean 83.3, SD 18.4). However, 1-way ANOVA found no statistically significant differences across age groups ($F_{4,45}$=0.399; P=.808, η^2=0.034). This small effect size indicates that age accounted for only approximately 3.4% of the variance in the SUS scores. Thus, age did not meaningfully influence caregivers' evaluations of the system. Overall, the mean SUS score of 80.45 exceeded the commonly cited 68-point benchmark and fell within the "excellent" range (≥80), showing that caregivers evaluated the EDT system very favorably.

Beyond demographic differences, we performed k-means clustering to classify caregivers' SUS scores into 3 groups: low (mean 52.7), medium (mean 79.4), and high (mean 94.30). As shown in Figure 8, the clusters clearly distinguished participants with below-average usability perceptions from those who reported excellent usability. Most caregivers fell into the medium to high clusters, reinforcing the overall favorable

XSL•FO
RenderX

evaluation of the system, while a smaller subgroup reflected lower usability perceptions, indicating areas for targeted improvements.

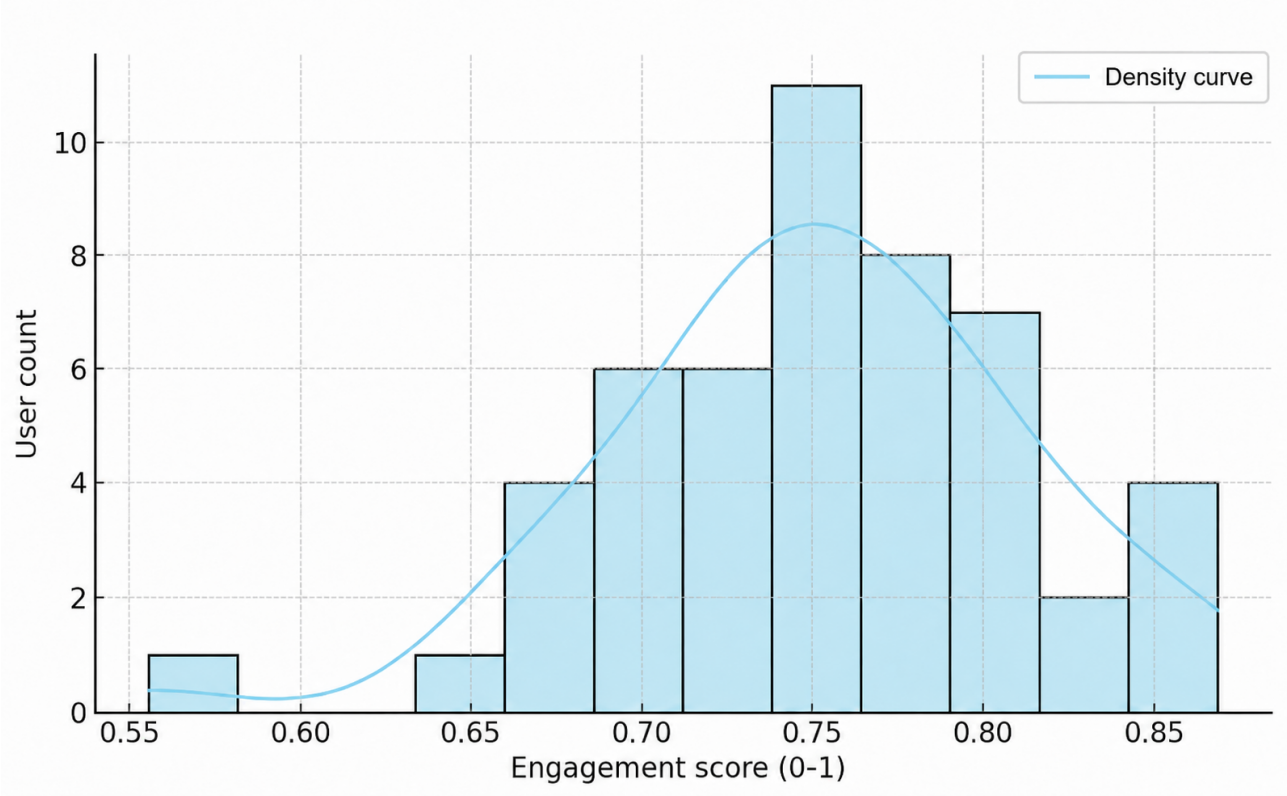
Figure 8. K-means clustering of SUS (N=50) into low, medium, and high groups. Dashed lines indicate cluster centroids. SUS: System Usability Scale.



User Engagement Based on Composite Behavioral Metrics Results

The composite UES was computed for all participants based on normalized behavioral metrics. The scores ranged from 0.55 to 0.86, with the majority of users clustering between 0.70 and 0.80, as shown in Figure 9. The distribution was slightly skewed toward higher engagement, indicating that most participants interacted consistently with the EDT system features, whereas only a few exhibited lower engagement levels.

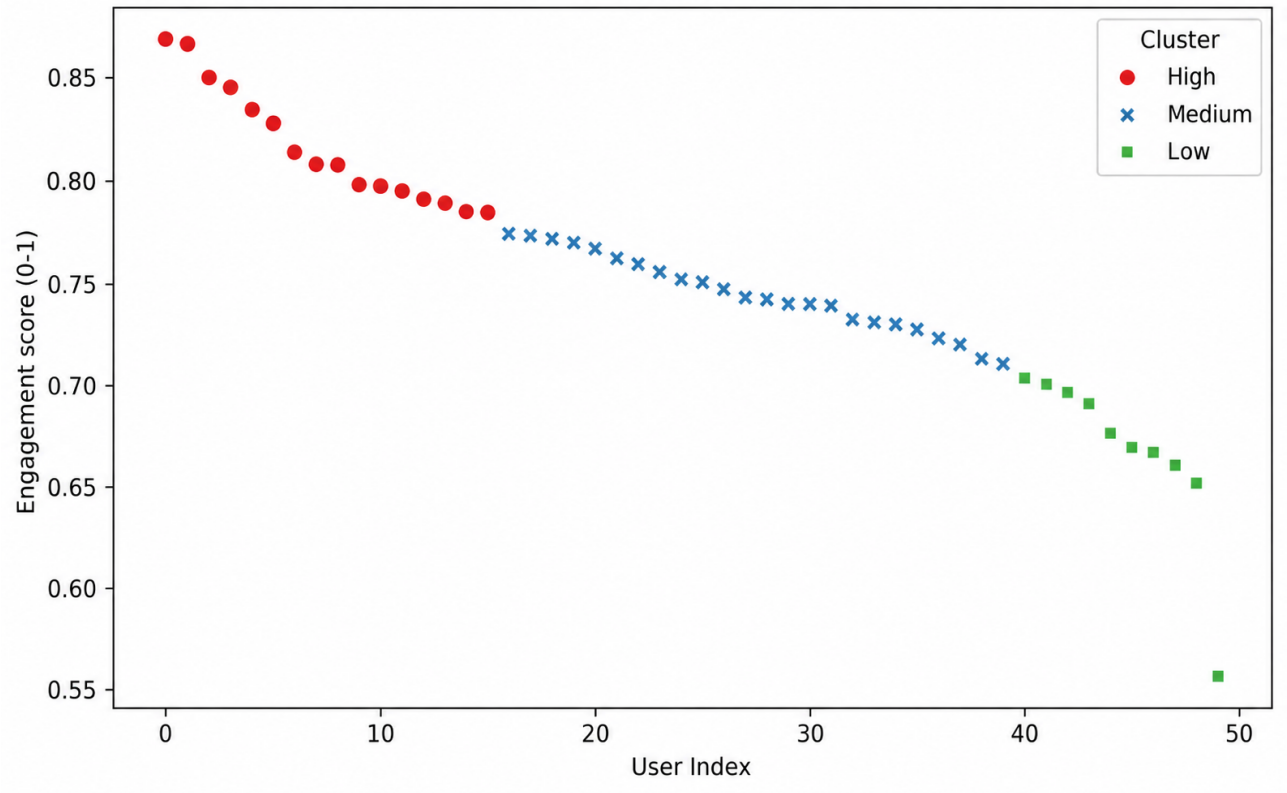
Figure 9. Distribution of composite engagement scores across users.



XSL•FO
RenderX

For interpretability, users were classified into 3 engagement tiers (low, medium, and high) using k-means clustering (k=3). The results showed that most users belonged to the medium engagement group, with smaller proportions in the high- and low-engagement tiers, as shown in Figure 10. This distribution indicates distinct user subgroups, with strong engagement among highly engaged users and lower, yet observable, interaction among the least engaged. Overall, the pattern suggests that the EDT system supported sustained interaction for most users.

Figure 10. Composite user engagement tiers using k-means clustering (k=3).



Correlation and Regression Results for Engagement and Perceived Usability

Building on the relationship between the UES and SUS, we examined how user behavioral measures can predict the SUS of the EDT system. As shown in Table 9, the regression analysis further demonstrated that UES significantly predicted usability scores ($F_{1, 48}$=31.00, P<.001), explaining 39.2% of the variance (R^2=0.392). Engagement was a positive predictor (β=52.94, t_{48}=5.57; P<.001), indicating that higher engagement was associated with higher SUS ratings. Figure 11 illustrates the correlation between SUS and UES.

XSL•FO
RenderX

Figure 11. Scatterplot of engagement and SUS scores with Pearson linear fit (solid line) and Spearman monotonic trend (dashed line). OLS; ordinary least squares SUS: System Usability Scale.

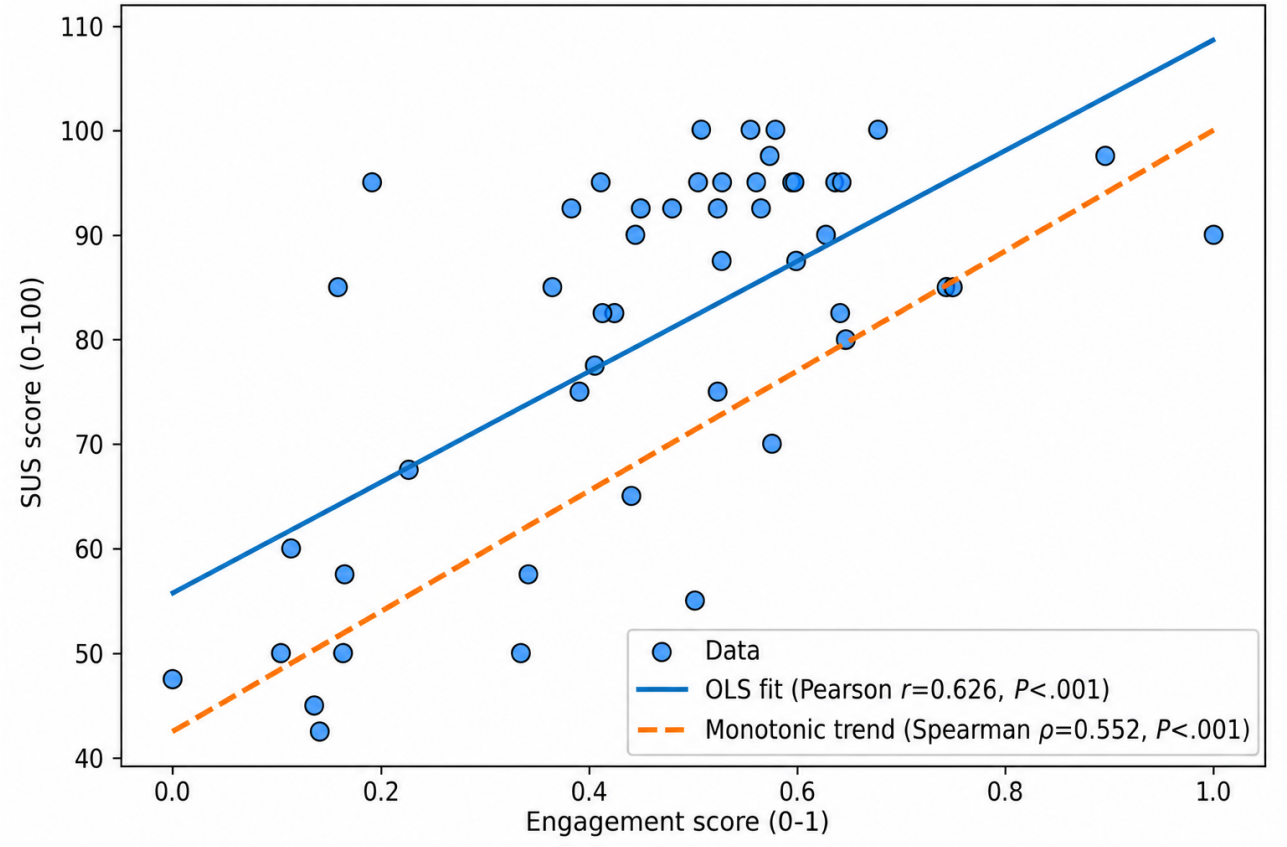


Table . Correlation between System Usability Scale and composite user engagement scores.

Analysis and variable	Statistics (r/p)/β (coefficient)[a]	*P* value	95% CI	R^2	SE	*t* value
Correlation						
Pearson *r*	0.626	<.001	0.417-0.778	—[b]	—	—
Spearman ρ	0.552	<.001	0.311-0.734	—	—	—
Regression						
Intercept	55.73	<.001	45.97-65.49	—	4.86	11.48
UES (β)	52.94	<.001	33.82-72.05	0.392	9.51	5.57

[a]*r* indicates the Pearson correlation coefficient; ρ indicates the Spearman rank correlation coefficient; β indicates the regression coefficient.
[b]Not available.

Discussion

Principal Findings

This study found that the EDT system achieved generally positive usability outcomes among informal caregivers based on both subjective perceptions and objective behavioral metrics. SUS results suggested favorable perceived usability, and user activity logs reflected meaningful engagement with system features. The observed association between SUS and UES further indicates that behavioral metrics may complement self-reported usability measures in evaluating caregiver-facing digital health systems.

Usability Evaluation Outcomes

The EDT system attained a mean SUS score of 80.45, exceeding the widely accepted threshold of 68 for above-average usability [19]. A score above 80 is typically interpreted as excellent usability [63]. In this sample, these results suggest that the participating informal caregivers perceived the EDT interface as relatively easy to learn and efficient to use. The composite UES derived from detailed activity logs was positively correlated with the SUS (Pearson *r*=0.626, Spearman ρ=0.552). The significant correlation between UES and SUS highlights the complementary nature of behavioral and perceptual usability measures. While the SUS captures caregivers' subjective impressions of the system's ease of use and satisfaction, the UES quantifies the depth and quality of their actual interactions.

XSL•FO
RenderX

The strength of the observed correlation between the UES and the SUS can be interpreted as moderate to strong, where the correlation between performance and satisfaction metrics generally ranges from 0.38 to 0.70 and often does not exceed 0.7 [23,64].

Beyond this association, regression analysis demonstrated that UES was a meaningful predictor of the SUS, explaining 39.2% of its variance. In the regression model, the *F*-test result ($F_{1, 48}$=31.00, *P*<.001) indicated that the model with UES as a predictor of SUS was statistically significant overall. The intercept (constant = 55.73) represents the predicted SUS value when UES equals zero; because zero engagement was not observed in the study data, this value should be interpreted as a model constant rather than a substantively meaningful estimate. Engagement score (β=52.94, *P*<.001) indicates that for every 1-point increase in UES (on its 0 - 1 scale), SUS is predicted to increase by about 53 points. The coefficient of determination (R^2=0.392) indicated that 39.2% of the variance in SUS was explained by UES. Because the model included only 1 predictor, this value is also consistent with the squared Pearson correlation between UES and SUS. The adjusted R^2 value of approximately 0.38 suggests a comparable level of explained variance after accounting for sample size. Although much of the remaining variance is likely attributable to other factors influencing perceived usability, these results support a meaningful association between behavioral engagement and perceived usability within this sample.

In human-computer interaction research, behavioral predictors typically explain only a partial variance in subjective satisfaction metrics. Studies have shown that behavioral log data account for 30% - 40% of the variance in perceived usability, with the remainder due to unobserved factors [65,66]. This aligns with strong task performance metrics: task completion 94.08%, time on task 89.16 seconds, abandonment rate 2.66%, and retry rate 0.25%, as shown in Table 5, demonstrating that objective engagement correlates with perceived usability.

The clustering of engagement scores into high-, medium-, and low-tier groups (Figure 10) reveals how caregivers interact with the EDT system. Users in the high-engagement cluster achieved strong task performance and high SUS scores, reflecting meaningful interactions. The medium engagement group showed satisfactory engagement with lower SUS scores, whereas the low-engagement group displayed limited interaction and lower usability perception. This analysis shows that users experience the system differently, highlighting the need to tailor system design and support for different engagement profiles. Demographic checks showed no significant difference in SUS scores by age (ANOVA $F_{4, 45}$=0.399; *P*=.81, η^2=0.034) or caregiving experience (ANOVA $F_{(3, 46}$=1.58; *P*=.21, η^2=0.093), indicating consistent perceived usability across the groups. The convergence of high perceived usability with high task performance provides evidence that the EDT system is usable and effective for caregivers. Higher SUS ratings correlated with more intensive interactions, demonstrating the system's support for meaningful use in caregiving contexts. These behavioral indicators in the composite UES strengthen the predictive relationship between the UES and the SUS, showing that engaged caregivers achieve higher task success and perceive better usability.

Theoretical Implications

The results demonstrate the complementarity between subjective and objective measures. While the SUS provides standardized user perception data with benchmarking values [19,22], it cannot reveal how users succeed or struggle during interactions [64]. By combining the SUS with UES behavioral analytics, our study showed that objective measures strengthen construct validity and highlight engagement heterogeneity that postuse scores may miss. In developing the UES, we refined the engagement score using foundational indicators of breadth, depth, and effectiveness of use, avoiding redundant measures, and ensuring interpretability. Equal weights were applied to the metrics, a defensible choice without prior evidence of differential importance. Future research could explore data-driven weighting strategies, such as principal component analysis, regression models, or SHAP analyses, using larger datasets [67].

Another important implication concerns the demographic sensitivity of the SUS. The ANOVA analysis showed no significant differences in usability scores by age or caregiving experience, suggesting that the SUS may be relatively stable across these demographic groups within this sample. The positive association between SUS and UES indicates that perceived usability and behavioral engagement are related. However, the present findings primarily demonstrate convergence between these measures rather than fully establishing their empirical complementarity. Future work should also examine more directly how both measures may capture distinct usability patterns and subgroup needs.

Practical Implication for Health Care Systems

The findings of this study have important implications for the use of DT systems in health care settings. First, the alignment between high SUS scores and favorable task performance metrics suggests that designers should prioritize core workflows that caregivers value most: real-time vital sign monitoring, health trend visualization, abnormal event alerts, and AI-based recommendations. Maintaining simplicity in these functions is essential for building caregiver trust and encouraging adoption. Second, the clustering analysis of the composite UES showed varied system engagement levels, with a small subgroup in the low engagement cluster. This finding indicates the need for an adaptive user interface. Designers should provide configurable interaction modes, progressive onboarding, and contextual help features to address the diverse needs and digital literacy levels of caregivers. Such adaptations can reduce the engagement barriers among different user groups. Third, given the central role of predictive modeling in EDT systems, such as Bi-LSTM for heart rate prediction and long short-term memory for sleep stage interface, AI-generated insights must be transparent and actionable. Caregivers must distinguish between real-time data and predictions, with alerts that are accompanied by clear rationales. This enhances interpretability and reduces inappropriate decision-making in care settings.

Finally, this study emphasizes the need to support informal caregivers who frequently multitask and experience cognitive

XSL•FO
RenderX

and emotional strain. EDT systems should adopt design strategies for rapid comprehension, including glanceable dashboards, consistent information hierarchies, and conservative notification policies to avoid alarm fatigue in users. This study demonstrates the value of behavioral analytics in pilot deployments, with system-generated logs supporting usability evaluation and providing feedback for refinement. Health care DT developers should integrate privacy-preserving analytics to monitor navigation issues and feature abandonment during real-world applications to guide design improvements and caregiver training. The results show that caregiver-facing DT systems must optimize core workflows while offering adaptability and transparency. These principles enable health care DT systems to support informal caregivers, enhance decision-making, and integrate digital health technologies into caregiving practices.

Comparative Positioning and Methodological Contribution

To situate the usability of the EDT system within the broader digital health literature, we compared its SUS scores with those reported in prior studies on related technologies, as shown in Table 10. The EDT system achieved 80.45, which falls within the "excellent" range and exceeds the values reported for many comparable systems, such as VITAAL (58.3), Pocket Gait (59.7), and ED Health Information System (53.1). Some systems, such as the CHES eDiary (87.5), achieved similarly high SUS scores, although often in more narrowly defined use cases than the others. This benchmark highlights 2 important contributions to literature. First, the EDT system ranks among the higher-performing digital health technologies in terms of perceived usability, reinforcing its readiness for use in caregiver-facing applications. Second, unlike most prior studies that relied exclusively on the SUS or paired it with limited subjective feedback (eg, interviews and acceptability scales), our study integrated objective behavioral metrics derived from detailed user activity logs. This dual-method approach strengthens the robustness of valuation and provides a replicable methodological pathway for future assessments of DT technologies in health care.

Table . Comparative System Usability Scale scores of related studies and the elderly digital twin system.

Study/System	SUS[a] score (reported)	Measures/Metrics used
HELMA [14]	72.2 (caregivers)	SUS + Partial usage logs
Engage [13]	66.3 → 82.6	SUS + NASA-TLX
VITAAL [15]	58.3	SUS + Feedback sessions
Pocket Gait [17]	59.7	SUS + Acceptability scale
CHES eDiary [45]	87.5	SUS + Walkthrough
TMS [16]	69.2	SUS
HeartAround [46]	62.2 → 78.8	SUS + Interview
ED Health Info System [47]	53.1	SUS + Interview
Our Study (EDT System)	80.45	SUS + Activity log metrics (UES[b], tasks)

[a]SUS: System Usability Scale.

[b]UES: user engagement score.

Limitations

Despite these strengths, several limitations must be acknowledged in this study. First, the sample size (n=50) limits the generalizability of the findings, and future research should validate the results with larger and more diverse caregiver populations. Second, the evaluation was conducted in a single-system context (the EDT prototype), which constrains external validity; usability outcomes may differ when deployed in varied care environments or integrated with other health systems. Third, although the SUS is a widely validated tool [63], it remains a subjective self-reported measure that can be influenced by user expectations, prior experience, or social desirability bias [68]. Although we mitigated this limitation by complementing SUS with objective behavioral logs, reliance on a single subjective scale may not fully capture all dimensions of usability. Finally, the cross-sectional study design limits our ability to assess the evaluation of engagement and usability perceptions evolved over time. Longitudinal studies are needed to capture the changes in caregiver interactions and satisfaction with sustained system use.

This study does not establish clinical validity or support the use of EDT as a diagnostic or treatment tool. The prototype was evaluated as a caregiver-oriented decision support and usability research system rather than as a regulated medical device, and the Fitbit Sense 2 used for sensing should be regarded as a consumer-grade wearable rather than a clinical reference device. In addition, the predictive components were evaluated only on the available study dataset, and no independent external sample validation was performed. Accordingly, the reported model performance metrics should be interpreted as preliminary indicators of technical feasibility rather than evidence of clinical-grade validity, external validity, or broader generalizability. Similarly, the LLM-generated outputs were intended to provide caregiver-friendly support, monitoring guidance, and follow-up suggestions rather than medical diagnosis or treatment advice. Furthermore, the composite engagement score was constructed using a pragmatic

XSL•FO
RenderX

equal-weight additive approach, and formal testing of interdependence among its component metrics was not conducted.

System Improvements and Future Work

This usability evaluation provided valuable lessons for refining the EDT. The system demonstrated excellent usability and high effectiveness, with a 94.08% task completion rate and low error rates. However, the findings reveal areas for enhancing system inclusivity and adoption. The small proportion of abandoned tasks (2.66%) and minimal retries (0.25 per task) indicated that certain interactions may challenge some users, particularly those in the low-engagement cluster identified in UES tiering. The average task time indicates overall efficiency but varies across engagement tiers. Medium- and low-engagement users required more time and retries, suggesting opportunities to streamline workflows. Simplifying navigation and reducing cognitive load can improve efficiency across different user profiles. Clustering analysis showed varied usability outcomes across users. Future development should address caregiver experience heterogeneity by tailoring features to different profiles, such as contextual prompts for low and advanced features for high-engagement users. Future work should expand the evaluation to more diverse caregiver populations, examine patterns of engagement and usability, incorporate contextual factors such as digital literacy and caregiving intensity, and evaluate the EDT in more naturalistic home care settings using scenarios grounded in real caregiver workflows to strengthen ecological validity. Future research should also examine intermetric correlations, dimensional structure, and alternative weighting strategies to further validate the composite engagement score. In addition, independent external datasets, prospective evaluation, and real-world deployment testing are needed to establish the robustness and generalizability of the predictive components more rigorously. Together, these efforts will help validate our findings and guide further improvements to ensure that the EDT system remains usable and responsive to caregiving practices.

Conclusions

This study evaluated the usability of the EDT system by combining subjective perceptions with objective behavioral data. The findings demonstrated excellent usability and strong task performance, with high completion, low error, and abandonment rates. Correlation and regression analyses showed that the composite UES was not only positively associated with but also a significant predictor of usability, explaining 39.2% of the variance in SUS. Clustering further revealed differences in engagement tiers, underscoring that usability outcomes are not uniform across users. Taken together, these results highlight the value of integrating subjective and objective measures to capture a comprehensive picture of product usability. This study contributes methodologically by linking behavioral engagement metrics with perceived usability, particularly by identifying the strengths and areas for improvement in EDT systems. Future studies should validate these findings in larger and more diverse caregiver populations, explore longitudinal engagement patterns, and incorporate adaptive features to support users with varying engagement profiles.

Acknowledgments

The authors would like to thank all study participants for their valuable time and insights during the usability testing sessions. The authors also extend special thanks to Miss Sutthikarn Chanthakup for her assistance as a local Thai research assistant, particularly in facilitating communication with Thai participants during data collection. This research was supported by the Petchra Pra Jom Klao Research Scholarship from King Mongkut’s University of Technology Thonburi (KMUTT). The authors confirm that no generative artificial intelligence tools were used for writing, editing, data analysis, figure generation, or other content generation during the preparation of this manuscript. The authors take full responsibility for the accuracy, originality, and integrity of all content in the manuscript.

Funding

No financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The dataset generated and analyzed in this study are available in Zenodo [69] and include anonymized data and supporting materials required to reproduce the study findings. Access to the elderly digital twin system link is available from the authors upon reasonable request.

Authors' Contributions

ZM contributed to conceptualization, methodology, data collection, formal analysis, and writing – original draft. PM participated in supervision, validation, and writing – review & editing. DP participated in supervision, methodology, and writing – review & editing. SY participated in supervision and review & editing.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Screenshots of the elderly digital twin system functions, interfaces, and key caregiver-facing modules.
[PDF File, 1473 KB - aging_v9i1e91873_app1.pdf]

References

1. Promoting health and well-being of older persons: WHO's support to ASEAN member states. World Health Organization. 2025. URL: https://www.who.int/thailand/news/detail/25-02-2025-promoting-health-and-well-being-of-older-persons--who-s-support-to-asean-member-states [accessed 2026-05-13]
2. Hong C, Sun L, Liu G, Guan B, Li C, Luo Y. Response of global health towards the challenges presented by population aging. China CDC Wkly 2023 Sep 29;5(39):884-887. [doi: 10.46234/ccdcw2023.168] [Medline: 37814614]
3. Khan HTA, Addo KM, Findlay H. Public health challenges and responses to the growing ageing populations. Public Health Chall 2024 Sep;3(3):e213. [doi: 10.1002/puh2.213] [Medline: 40496520]
4. Katsoulakis E, Wang Q, Wu H, et al. Digital twins for health: a scoping review. NPJ Digit Med 2024 Mar 22;7(1):77. [doi: 10.1038/s41746-024-01073-0] [Medline: 38519626]
5. Magwa P. Leveraging digital twins for personalized medicine: A framework for predictive therapeutics. Clin Pharmacol Biopharm 2024;13(12) [FREE Full text] [doi: 10.4172/2167-065X.1000527]
6. Vallée A. Digital twin for healthcare systems. Front Digit Health 2023;5:1253050. [doi: 10.3389/fdgth.2023.1253050] [Medline: 37744683]
7. Wang Y, Leung JC, Chen M, Qiu Y, Tan BTH, Zeng Z, et al. An intelligent and privacy-preserving digital twin model for aging-in-place. . Preprint posted online on Apr 4, 2025. [doi: 10.48550/arXiv.2504.03798]
8. Zhang J, Qian H, Zhou H. Application and research of digital twin technology in safety and health monitoring of the elderly in community. Zhongguo Yi Liao Qi Xie Za Zhi 2019 Nov 30;43(6):410-413. [doi: 10.3969/j.issn.1671-7104.2019.06.005] [Medline: 31854524]
9. Harding AJE, Doherty J, Bavelaar L, et al. A family carer decision support intervention for people with advanced dementia residing in a nursing home: a study protocol for an international advance care planning intervention (mySupport study). BMC Geriatr 2022 Oct 26;22(1):822. [doi: 10.1186/s12877-022-03533-2] [Medline: 36289458]
10. Magnusson DM, Shwayder I, Murphy NJ, Ollerenshaw L, Ebendick M, Auer-Bennett E. Creation of a community-driven decision support tool for caregivers of children with developmental concerns. Am J Speech Lang Pathol 2022 May 10;31(3):1084-1094. [doi: 10.1044/2021_AJSLP-21-00072] [Medline: 34731583]
11. Alwashmi MF, Hawboldt J, Davis E, Fetters MD. The iterative convergent design for mobile health usability testing: mixed methods approach. JMIR Mhealth Uhealth 2019 Apr 26;7(4):e11656. [doi: 10.2196/11656] [Medline: 31025951]
12. Ghorayeb A, Darbyshire JL, Wronikowska MW, Watkinson PJ. Design and validation of a new Healthcare Systems Usability Scale (HSUS) for clinical decision support systems: a mixed-methods approach. BMJ Open 2023 Jan 30;13(1):e065323. [doi: 10.1136/bmjopen-2022-065323] [Medline: 36717136]
13. Cornet VP, Daley CN, Srinivas P, Holden RJ. User-centered evaluations with older adults: testing the usability of a mobile health system for heart failure self-management. Proc Hum Factors Ergon Soc Annu Meet 2017 Sep;61(1):6-10. [doi: 10.1177/1541931213601497] [Medline: 30930610]
14. Cossu-Ergecer F, Dekker M, van Beijnum BJF, Tabak M. Usability of a new eHealth monitoring technology that reflects health care needs for older adults with cognitive impairments and their informal and formal caregivers. Presented at: 11th International Conference on Health Informatics; Oct 25-29, 2026; Funchal, Madeira, Portugal p. 197-207. [doi: 10.5220/0006639301970207]
15. Thalmann M, Ringli L, Adcock M, et al. Usability study of a multicomponent exergame training for older adults with mobility limitations. Int J Environ Res Public Health 2021 Dec 20;18(24):13422. [doi: 10.3390/ijerph182413422] [Medline: 34949028]
16. Hochwarter S, Gutheil J, Stampfer P, Truskaller T, Deutsch M, Feichtner F. One system, many professions: a System Usability Scale evaluation of a multi-professional therapy and monitoring system. Stud Health Technol Inform 2025 May 15;327:308-312. [doi: 10.3233/SHTI250334] [Medline: 40380445]
17. Zhong R, Rau PLP. A mobile phone-based gait assessment app for the elderly: development and evaluation. JMIR Mhealth Uhealth 2020 Feb 29;8(2):e14453. [doi: 10.2196/14453] [Medline: 32452821]
18. Momand Z, Mongkolnam P, Chan JH, Charoenkitkarn N, Pal D. Building digital twins for elderly care: an end-to-end framework from data acquisition to modeling. IEEE Access 2025;13:169415-169445. [doi: 10.1109/ACCESS.2025.3607603]
19. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
20. Deshmukh AM, Chalmeta R. Validation of system usability scale as a usability metric to evaluate voice user interfaces. PeerJ Comput Sci 2024;10:e1918. [doi: 10.7717/peerj-cs.1918] [Medline: 38435614]
21. Khan Q, Hickie IB, Loblay V, et al. Psychometric evaluation of the System Usability Scale in the context of a childrearing app co-designed for low- and middle-income countries. Digit Health 2025 May;11:20552076251335413. [doi: 10.1177/20552076251335413]
22. Lewis JR, Sauro J. Item benchmarks for the System Usability Scale. J Usability Stud 2018;13(3):158-167 [FREE Full text]

XSL•FO
RenderX

23. Lewis JR. Measuring perceived usability: the CSUQ, SUS, and UMUX. Int J Hum Comput Interact 2018 Dec 2;34(12):1148-1156. [doi: 10.1080/10447318.2017.1418805]
24. Laubheimer P. Beyond the NPS: measuring perceived usability with the SUS, NASA-TLX, and the single ease question after tasks and usability tests. Nielsen Norman Group. 2018. URL: https://www.nngroup.com/articles/measuring-perceived-usability [accessed 2026-05-13]
25. Takahashi L, Nebe K. Observed differences between lab and online tests using the AttrakDiff Semantic Differential Scale. J Usability Stud 2019;14(2):65-75 [FREE Full text]
26. Robertson IW. Subjective Usability Evaluation: A Comparison of Four Methods: Rice University; 2018. URL: https://hdl.handle.net/1911/105826 [accessed 2026-05-25]
27. Sauro J, Kindlund E. A method to standardize usability metrics into a single score. In: Sauro J, Kindlund E, editors. Presented at: CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 2-7, 2005. [doi: 10.1145/1054972.1055028]
28. Latkin CA, Edwards C, Davey-Rothwell MA, Tobin KE. The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. Addict Behav 2017 Oct;73:133-136. [doi: 10.1016/j.addbeh.2017.05.005] [Medline: 28511097]
29. Strba M. Usability testing metrics. UXtweak. 2024. URL: https://www.uxtweak.com/usability-testing/metrics [accessed 2026-05-13]
30. Hariyanti NKD, Sudhana IFP, Sanjaya IGN, Elfarosa KV. Implementation of usability testing in measuring the effectiveness and efficiency of mobile application. Presented at: Proceedings of the 5th International Conference on Applied Science and Technology on Engineering Science (iCAST-ES 2022); Oct 21-23, 2022. [doi: 10.5220/0011892900003575]
31. Nielsen J, Budiu R. Success rate: the simplest usability metric. Nielsen Norman Group. 2021. URL: https://www.nngroup.com/articles/success-rate-the-simplest-usability-metric [accessed 2026-05-13]
32. Wang G, Zhang X, Tang S, Wilson C, Zheng H, Zhao BY. Clickstream User Behavior Models. ACM Trans Web 2017 Nov 30;11(4):1-37. [doi: 10.1145/3068332]
33. Gerken J, Bak P, Jetter HC, Klinkhammer D, Reiterer H. How to use interaction logs effectively for usability evaluation. 2008 Presented at: CHI 2008 Workshop BELIV ' 08: Beyond time and errors—novel evaluation methods for Information; Apr 2008 URL: http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-52435 [accessed 2025-08-25]
34. Dumais S, Jeffries R, Russell DM, Tang D, Teevan J. Understanding user behavior through log data and analysis. In: Ways of Knowing in HCI: Springer Science & Business; 2014. [doi: 10.1007/978-1-4939-0378-8_14]
35. Roberts MJ. Objective and subjective methods for evaluating the usability of schematic maps: the case against informal expert assessments. Cartogr J 2023 Oct 2;60(4):308-325. [doi: 10.1080/00087041.2023.2246742]
36. Gao M, Kortum P. The relationship between subjective and objective usability metrics for home healthcare devices. Proc Hum Factors Ergon Soc Annu Meet 2015 Sep;59(1):1001-1005. [doi: 10.1177/1541931215591286]
37. Mann DM, Chokshi SK, Kushniruk A. Bridging the gap between academic research and pragmatic needs in usability: a hybrid approach to usability evaluation of health care information systems. JMIR Hum Factors 2018 Nov 28;5(4):e10721. [doi: 10.2196/10721] [Medline: 30487119]
38. Callejo A, Macías JA. Enhancing tree testing analysis to improve the usability evaluation of websites. Behav Inf Technol 2026 Apr 3;45(6):1117-1135. [doi: 10.1080/0144929X.2025.2546971]
39. Feehan LM, Geldman J, Sayre EC, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth 2018 Aug 9;6(8):e10527. [doi: 10.2196/10527] [Medline: 30093371]
40. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of wristband Fitbit models in assessing sleep: systematic review and meta-analysis. J Med Internet Res 2019 Nov 28;21(11):e16273. [doi: 10.2196/16273] [Medline: 31778122]
41. Lee T, Cho Y, Cha KS, et al. Accuracy of 11 wearable, nearable, and airable consumer sleep trackers: prospective multicenter validation study. JMIR Mhealth Uhealth 2023 Nov 2;11(1):e50983. [doi: 10.2196/50983]
42. Hermans F, Arents E, Blondeel A, et al. Validity of a consumer-based wearable to measure clinical parameters in patients with chronic obstructive pulmonary disease and healthy controls: observational study. JMIR Mhealth Uhealth 2024 Nov 6;12(1):e56027-e56027. [doi: 10.2196/56027]
43. Caunca MR, Simonetto M, Hartley G, Wright CB, Czaja SJ. Design and usability testing of the stroke caregiver support system: a mobile-friendly website to reduce stroke caregiver burden. Rehabil Nurs 2020;45(3):166-177. [doi: 10.1097/RNJ.0000000000000196] [Medline: 30418319]
44. Abbate S, Avvenuti M, Light J. Usability study of a wireless monitoring system among Alzheimer's disease elderly population. Int J Telemed Appl 2014;2014:617495. [doi: 10.1155/2014/617495] [Medline: 24963289]
45. Lehmann J, Schreyer I, Riedl D, et al. Usability evaluation of the Computer-Based Health Evaluation System (CHES) eDiary for patients with faecal incontinence: a pilot study. BMC Med Inform Decis Mak 2022 Mar 28;22(1):81. [doi: 10.1186/s12911-022-01818-5] [Medline: 35346170]
46. Panagopoulos C, Menychtas A, Tsanakas P, Maglogiannis I. Increasing usability of homecare applications for older adults: a case study. Designs 2019;3(2):23. [doi: 10.3390/designs3020023]

XSL•FO
RenderX

47. Østervang C, Jensen CM, Coyne E, Dieperink KB, Lassen A. Usability and evaluation of a health information system in the emergency department: mixed methods study. JMIR Hum Factors 2024 Feb 21;11:e48445. [doi: 10.2196/48445] [Medline: 38381502]
48. Paulissen JMJ, Zegers CML, Nijsten IR, et al. Performance and usability evaluation of a mobile health data capture application in clinical cancer trials follow-up. Tech Innov Patient Support Radiat Oncol 2022 Dec;24:107-112. [doi: 10.1016/j.tipsro.2022.10.005] [Medline: 36387779]
49. Borsci S, Federici S, Lauriola M. On the dimensionality of the System Usability Scale: a test of alternative measurement models. Cogn Process 2009 Aug;10(3):193-197. [doi: 10.1007/s10339-009-0268-9] [Medline: 19565283]
50. Lewis JR, Sauro J. The factor structure of the System Usability Scale. : Springer; 2009 Presented at: Proceedings of the 1st International Conference on Human Centered Design: Held as Part of HCI International 2009; Jul 19-24, 2009; San Diego, California, United States. [doi: 10.1007/978-3-642-02806-9_12]
51. Liu Y, Tan H, Cao G, Xu Y. Enhancing user engagement through adaptive UI/UX design: A study on personalized mobile app interfaces. World J Innov Mod Technol 2024;7(5):1-21. [doi: 10.53469/wjimt.2024.07(05).01]
52. Wibawa AP, Manik LP, Pranolo A, Drezewski R, Hernandez L, Ismail AR, et al. Exploring usage-based and usability metrics for user experience for sustainable e-learning systems. 2024 Presented at: International Conference on Computer Science, Electronics, and Information (ICCSEI 2023); Mar 17-19, 2023. [doi: 10.1051/e3sconf/202450102003]
53. Trukenbrod AK, Backhaus N, Thomaschke R. Measuring subjectively experienced time in usability and user experience testing scenarios. Int J Hum Comput Stud 2020 Jun;138:102399. [doi: 10.1016/j.ijhcs.2020.102399]
54. Yang J, Abraham A. Analyzing the features, usability, and performance of deploying a containerized mobile web application on serverless cloud platforms. Future Internet 2024;16(12):475. [doi: 10.3390/fi16120475]
55. Debora D, Kharisma NN, Rifaldi W, Nugraha U. UI/UX transformation of XYZ retail information system through user-centered design approach. Jurnal Sistem Informasi dan Teknologi Informasi [J Inform Syst Inform Technol] 2024;2(2). [doi: 10.33197/justinfo.v2i2.2539]
56. Henriksson E, Lundström M. Navigation systems' impact on usability in mobile applications: a study on mobile newspaper applications. : Jonkoping University; 2021 URL: https://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-53474
57. Weingarden H, Garriga Calleja R, Greenberg JL, et al. Characterizing observed and effective behavioral engagement with smartphone cognitive behavioral therapy for body dysmorphic disorder: a methods roadmap and use case. Internet Interv 2023 Apr;32:100615. [doi: 10.1016/j.invent.2023.100615] [Medline: 36969390]
58. Lalmas M, O'Brien H, Yom-Tov E. Measuring User Engagement: Springer Nature; 2022.
59. Bobko P, Roth PL, Buster MA. The usefulness of unit weights in creating composite scores. Organ Res Methods 2007 Oct;10(4):689-709. [doi: 10.1177/1094428106294734]
60. Wehbe C, Baroud H. Limitations and considerations of using composite indicators to measure vulnerability to natural hazards. Sci Rep 2024 Aug 20;14(1):19333. [doi: 10.1038/s41598-024-68060-z] [Medline: 39164315]
61. Alabbas A, Alomar K. A weighted composite metric for evaluating user experience in educational chatbots: balancing usability, engagement, and effectiveness. Future Internet 2025;17(2):64. [doi: 10.3390/fi17020064]
62. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55. [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
63. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum Comput Interact 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
64. Hornbæk K. Current practice in measuring usability: Challenges to usability studies and research. Int J Hum Comput Stud 2006 Feb;64(2):79-102. [doi: 10.1016/j.ijhcs.2005.06.002]
65. Egorow O, Siegert I, Wendemuth A, Egorow O, Siegert I, Wendemuth A. Prediction of user satisfaction in naturalistic human-computer interaction. Kognitive Systeme 2017(1). [doi: 10.17185/duepublico/44534]
66. Fennell PG, Zuo Z, Lerman K. Predicting and explaining behavioral data with structured feature space decomposition. EPJ Data Sci 2019 Dec;8(1). [doi: 10.1140/epjds/s13688-019-0201-0]
67. Alenazi SA. Predictive modelling of omni-channel customer behavior using big data analytics for retail marketing. Int J Innov Res Sci Stud 2025;8(5):1350-1359. [doi: 10.53894/ijirss.v8i5.9134]
68. Orn A. Pros and cons of the system usability scale (SUS). Research Collective. 2017. URL: https://research-collective.com/pros-and-cons-of-the-system-usability-scale-sus [accessed 2026-05-13]
69. Momand Z, Mongkolnam P, Chan J, Charoenkitkarn N, Pal D. Longitudinal wearable physiological dataset of elderly individuals. Zenodo. URL: https://zenodo.org/records/18745170 [accessed 2026-05-25]

Abbreviations

AI: artificial intelligence
Bi-LSTM: bidirectional long short-term memory
CTR: click-to-task ratio
DT: digital twin
EDT: elderly digital twin

XSL•FO
RenderX

FU: feature usage
HRV: heart rate variability
ISO: International Organization for Standardization
LLM: large language model
NASA-TLX: National Aeronautics and Space Administration—Task Load Index
NF: navigation frequency
RR: retry rate
SpO_2 : peripheral oxygen saturation
SUS: System Usability Scale
TAR: task abandonment rate
TC: task completion
ToT: time on task
UES: user engagement score
UMUX: usability metric for user experience

Edited by PH Chau; submitted 21.Jan.2026; peer-reviewed by N Mor, X Zhang; revised version received 14.Apr.2026; accepted 18.Apr.2026; published 26.May.2026.

Please cite as:
Momand Z, Mongkolnam P, Pal D, Yamsaengsung S
Combining Subjective Perceptions and Objective Behavioral Metrics With the Elderly Digital Twin System: Quantitative Usability Study
JMIR Aging 2026;9:e91873
URL: https://aging.jmir.org/2026/1/e91873
doi:10.2196/91873

© Ziaullah Momand, Pornchai Mongkolnam, Debajyoti Pal, Siam Yamsaengsung. Originally published in JMIR Aging (https://aging.jmir.org), 26.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multidimensional Digital Literacy and Quality of Life Among Informal Care Dyads in Malaysia: Cross-Sectional Survey

Fereshteh Mohammadzadeh Yazd[1], MSc; Hui Foh Foong[1], PhD; Rahimah Ibrahim[1,2], PhD; Puvaneswaran Kunasekaran[1,2], PhD; Asmidawati Ashari[1,2], PhD

[1]Malaysian Research Institute on Ageing (MyAgeing™), Universiti Putra Malaysia, UPM Serdang, Malaysia

[2]Faculty of Human Ecology, Universiti Putra Malaysia, Serdang, Malaysia

Corresponding Author:
Rahimah Ibrahim, PhD
Malaysian Research Institute on Ageing (MyAgeing™), Universiti Putra Malaysia, UPM Serdang, Malaysia

Abstract

Background: Digital literacy (DL) is a key determinant of health and social participation in later life. In Malaysia, where population aging and family caregiving are rising, limited digital engagement among older adults may widen the gray digital divide. As caregivers and care recipients are interdependent, their digital capacities may jointly shape each other's quality of life (QoL).

Objective: This study examined the multidimensional associations between DL and QoL among informal caregivers and care recipients in Malaysia and explored dyadic actor-partner effects.

Methods: A cross-sectional dyadic survey was conducted with 72 caregiver-care recipient dyads (N=144) recruited from 4 sites across Klang Valley, Malaysia, through purposive sampling. Dyads were the unit of analysis. DL was measured using the 22-item Everyday Digital Literacy Questionnaire across 3 dimensions (information and communication, content creation and management [CCM], and safety and security [SS]), while QoL was assessed with the World Health Organization Quality of Life-BREF. Partial least squares structural equation modeling tested dimension-specific links, and the actor-partner interdependence model examined bidirectional effects.

Results: Caregivers reported higher DL and QoL than care recipients. Among caregivers, only "SS" significantly predicted QoL (β=.44, *P*<.001); among care recipients, all 3 dimensions were significant, with "CCM" showing the strongest effect (β=.53, *P*<.001). The actor-partner interdependence model indicated significant actor and partner effects (caregiver actor β=.53, partner β=.34; care recipient actor β=.60, partner β=.45; all *P*<.001), explaining 62% and 60% of variance in QoL.

Conclusions: DL shows a clear actor-partner pattern. For actor effects, caregivers' QoL was linked to their own SS skills, while care recipients' QoL was linked to their own CCM skills. Partner effects showed that care recipients' CCM skills also predicted caregivers' QoL. Both effects were significant, indicating interdependence within the caregiving dyad. These findings can inform dyad-focused and dimension-specific DL interventions to enhance mutual well-being.

(JMIR Aging 2026;9:e86561) doi:10.2196/86561

KEYWORDS

digital literacy; quality of life; older adults; care recipients; caregivers; informal caregiving; caregiving dyads; dyadic research; actor-partner interdependence model; Malaysia

Introduction

Population aging is a global phenomenon. It is transforming social, economic, and health systems worldwide [1]. Increasing longevity and declining fertility rates have led to a larger proportion of older adults in the population. This has caused reliance on informal caregivers, such as spouses or relatives, who provide unpaid care at home [2]. Meanwhile, caregivers have a pivotal role in maintaining the health and well-being of older adults. However, they often experience heavy burdens, such as physical strain, emotional stress, and financial burden, which reduce their quality of life (QoL) [3]. Research has shown that the health and well-being of caregivers and care recipients are interdependent. Emotional distress in one often mirrors the experience of the other [4,5]. This interdependence highlights that we should seek digitally inclusive solutions to improve the well-being of both caregivers and care recipients. In Asia, this need is even more crucial with the rapid growth of the aging population and the spread of digital technologies.

Population aging is progressing rapidly in Asia [6], and the need for long-term and family-based care systems is increasing [7]. Malaysia is an example of this demographic shift, with an estimated 15% of the population projected to be over 60 years of age by 2030 [8]. However, the digital engagement and

XSL•FO
RenderX

participation of older adults are still limited. A recent study in Malaysia highlighted that while device ownership is increasing, actual utilization, specifically for health information, remains low (27%). Furthermore, a majority of older adults (68%) perceive online information as unreliable or lack the necessary confidence and skills to navigate digital platforms effectively [9]. These figures indicate a persistent "digital gray divide." This means that inequalities in access, trust, and digital skills among older adults persist [10]. As family-based caregiving is still the main form of supporting older adults in Malaysia [11], strengthening digital skills within these caregiving relationships is essential.

Digital literacy (DL) is the ability to find, evaluate, and use digital information effectively [12]. It has been identified as a major determinant of health and care outcomes [13]. DL goes beyond technical skills. It helps people access online health care services, communicate with health care professionals, and use digital resources for their daily living. Higher DL is associated with lower caregiver stress and better QoL for both caregivers and care recipients [14,15]. A recent scoping review shows that informal caregivers with stronger digital skills experience greater empowerment. As a result, they feel confident in their caregiving role [16]. Yet, a lack of DL is associated with limited access to reliable information. It is linked to increased vulnerability to misleading information and may exacerbate inequalities in health participation [17,18]. This gap not only affects how caregivers search for information but also how they support care recipients.

Malaysia's aging policies have identified technology as a tool to enhance independence and achieve "supported aging" [19]. However, the National Policy for Older Persons [19], which was developed in the early 2000s, was formulated at a time when concepts such as "digital literacy" and "digital inclusion" were not yet embedded in social policy-making. For this reason, this document does not explicitly address digital participation or the role of informal caregivers in technology use and adoption [20]. Similarly, the Malaysia Digital Economy Blueprint [21] focuses more on innovation, economic growth, and industrial transformation. Although the document emphasizes the need for DL, it does not pay much attention to the specific needs of care recipients and caregivers. Overall, despite the growing attention to technology in existing policies, a comprehensive framework for integrating care recipients and caregivers into the country's digital ecosystem has not yet been formed [16].

Recent studies suggest that digital skills can be associated with increased autonomy, social participation, and health-promoting behaviors among care recipients and caregivers [7,22]. Targeted training programs have also been shown to improve digital skills and self-efficacy in these groups [23,24]. In dyadic caregiving relationships, the caregiver's use of technology is associated with the care recipient's level of digital participation [25], such that higher DL in one person is associated with higher DL in the other. Although several recent studies have applied dyadic approaches to examine health literacy or eHealth literacy in patient-caregiver dyads [26,27], these investigations have generally focused on health behaviors or technology use intentions rather than on everyday DL or QoL. Consequently, little is known about how multidimensional DL shapes actor-partner patterns of QoL within informal caregiving relationships.

Given the distinct roles and digital engagement patterns of caregivers and care recipients, we expected that different DL dimensions might show differential importance for each group. For caregivers, who frequently seek health information online on behalf of care recipients, safety and security (SS) skills may be particularly salient in reducing exposure to misinformation and digital risks [16,28]. For care recipients, communication and content creation skills may be more strongly linked to QoL, as these dimensions enable active social participation and self-expression, key components of autonomy and social connectedness in later life [29,30].

Building on this evidence and the concept of interdependence in caregiving relationships [4,31], this study examines how multidimensional DL, including the dimensions of "information and communication (IC)," "content creation and management (CCM)," and "SS," is related to the QoL of informal caregivers and care recipients in Malaysia. Using the actor-partner interdependence model (APIM), this study seeks to test the hypothesis that the DL of each member of a caregiving dyad is related not only to their own QoL (actor effect) but also to the QoL of their partner (partner effect). It is expected that examining this model will lead to a more comprehensive understanding of the multidimensional role of DL in promoting the mutual well-being of caregivers and care recipients and provide a basis for future research and intervention programs in the field of digital aging.

Methods

Study Design, Setting, and Participants

This study was a cross-sectional, dyadic design conducted in the Klang Valley region of Malaysia (Selangor and the Federal Territory of Kuala Lumpur) between May and October 2025. Eligible participants were informal caregivers aged 18 years and older and community-dwelling older adults aged 60 years and older (hereinafter referred to as "care recipients") residing in the urban Klang Valley region. The caregiver had to be the primary support person nominated by the care recipient, and both were required to complete the questionnaire independently. Informal caregiver-care recipient dyads were the unit of analysis.

A pilot study with 20 dyads was conducted to estimate the population parameters for sample size determination. To address the dyadic nature of the data, a Monte Carlo power analysis was performed using Mplus Version 8.3 [32], following guidelines for sample size estimation in structural equation modeling (SEM) [33]. Based on pilot results and dyadic research recommendations, we hypothesized standardized actor effects of 0.50 and partner effects of 0.37. The simulation parameters were set at α=.05 with 1000 replications using the maximum likelihood robust (MLR) estimator. Results indicated that a sample size of 69 dyads would achieve a statistical power of greater than 0.80 to detect both actor and partner effects. Participants were recruited from 4 sites, including community centers, care service organizations, and public events across Klang Valley, through purposive sampling. A total of 100 dyads

XSL•FO
RenderX

were approached and consented. As expected in dyadic field research, some dyads were incomplete because one member did not return the questionnaire, in which case the entire dyad was excluded. Ultimately, 72 complete dyads were obtained (72% dyad-level participation), exceeding the a priori target and falling within the sample size range often recommended for stable estimation in APIM studies [34].

To ensure reliable self-reports, participants were screened for subjective memory complaints using a single-item cognitive self-assessment. This brief approach is supported by previous research showing that the absence of memory complaints is a useful indicator of adequate cognitive ability for independent questionnaire completion [35]. No dyads were excluded based on this screening item.

Measures

Demographic Characteristics

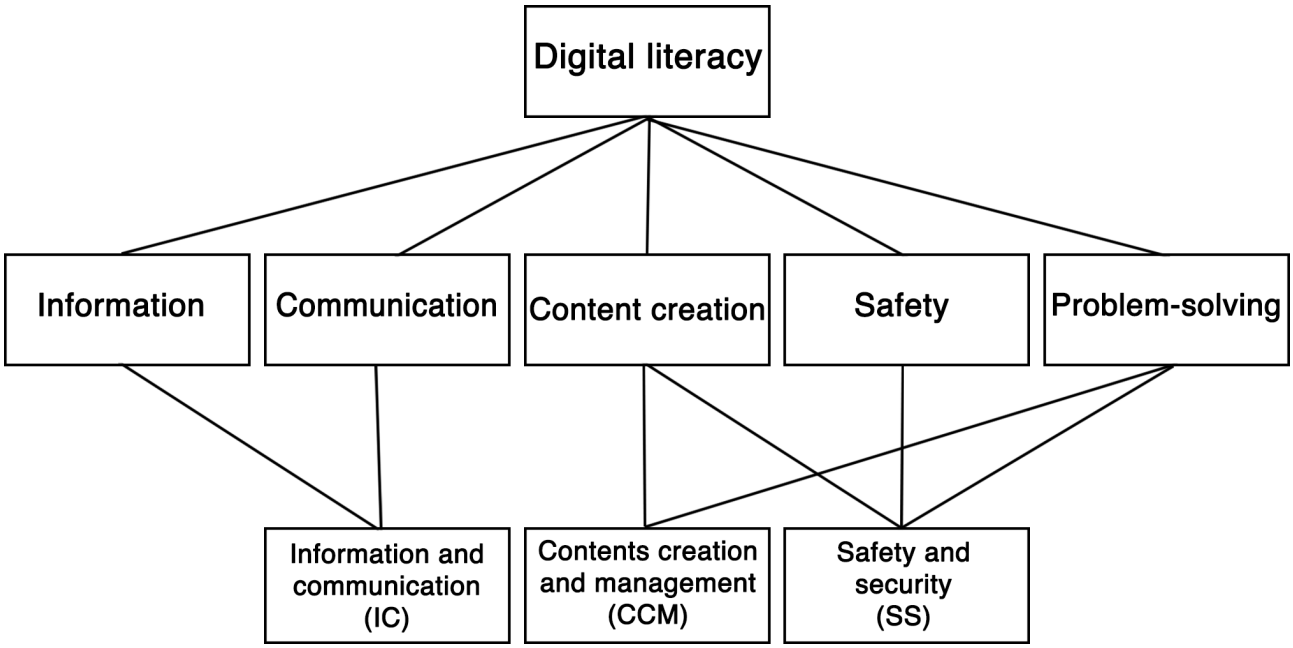
Demographic characteristics were collected for both caregivers and care recipients, including age, gender, marital status, ethnicity, education level, employment status, and monthly household income. Information on internet access (frequency, connection quality, and device use) and self-rated health status was also obtained.

Digital Literacy

DL was assessed using the 22-item Everyday Digital Literacy Questionnaire (EDLQ) [36]. The instrument is conceptually grounded in the European Commission's DigComp framework and broader multidimensional conceptualizations of DL. It was validated with a 3-factor structure. Each item is rated on a 5-point Likert scale ranging from 1 ("Never True") to 5 ("Always True"), with higher scores indicating greater DL.

Following the authors' scoring guidelines, 3 domain scores were calculated: IC (mean of DL01-DL09), CCM (mean of DL10-DL13), and SS (mean of DL14-DL22). In addition, a total score (sum of all 22 items, range 22 - 110) was computed to capture overall DL. Scores were calculated separately for caregivers and care recipients. The EDLQ was originally developed and validated in Korean. English versions of all 22 items were available in the original validation study [36]. The English items were translated into Malay and reviewed by a bilingual academic with expertise in gerontology for linguistic clarity and cultural appropriateness. Due to sample size constraints, measurement invariance testing across languages was not conducted and is acknowledged as a limitation. The 3 domains of the EDLQ are illustrated in Figure 1.

Figure 1. The Everyday Digital Literacy Questionnaire structure, adapted from Choi et al [36].



Quality of Life

QoL was assessed using the World Health Organization Quality of Life-BREF (WHOQOL-BREF), a 26-item instrument developed by the World Health Organization to measure individuals' perceived QoL across 4 domains: physical health, psychological health, social relationships, and environment [37,38]. Each item is rated on a 5-point Likert scale with varying anchors (eg, 1="very poor" to 5="very good," or 1="not at all" to 5="completely"), reflecting the intensity, capacity, frequency, or evaluation of life experiences during the past 2 weeks. Higher scores indicate better QoL. Domain scores were computed following WHO scoring guidelines and transformed to a standardized 0 to 100 scale using the recommended formula: (transformed score=[raw score–minimum possible score]×100/[maximum possible score–minimum possible score]) [37].

An overall QoL index was calculated by averaging the 4 domain scores. In this study, the 4 WHOQOL-BREF domains showed very high intercorrelations (caregivers: r=0.851-0.928; care recipients: r=0.863-0.938; all P<.001), indicating substantial shared variance and supporting the use of a composite measure of overall well-being [38]. Moreover, examining all 4 domains separately would require multiple independent models (raising concerns about statistical power and Type I error inflation with

XSL•FO
RenderX

N=72 dyads). Therefore, this composite score was used for subsequent SEM and APIM analyses. Scores were analyzed separately for caregivers and care recipients.

Ethical Considerations

Data were collected through structured, self-administered questionnaires. The questionnaires were available in both English and Malay, and participants completed the version in their preferred language. Each dyad received 1 set of questionnaires, and an online completion option was also provided. Responses were checked to ensure caregiver-care recipient matching.

Institutional Review Board approval was granted by the Ethics Committee for Research Involving Human Subjects (JKEUPM) of Universiti Putra Malaysia (approval number: JKEUPM-2025-037). Participation was voluntary, and written informed consent was obtained from all respondents. Confidentiality was ensured through anonymized data, and participants could withdraw at any stage without consequence. Participants were not compensated for their participation in this study.

Statistical Analysis

Statistical analyses for this cross-sectional dyadic data were conducted to describe participant characteristics, validate the measurement model, and examine dyadic actor-partner effects of DL on QoL.

Descriptive statistics (means, SDs, frequencies, and percentages) were computed in IBM SPSS Statistics Version 26. Construct validity and reliability of the measurement model were evaluated in SmartPLS Version 4.0.9.6 for both the EDLQ and WHOQOL-BREF. Reliability and convergent validity were confirmed based on item loadings ≥0.70, Cronbach α and composite reliability >0.70, and average variance extracted >0.50. Discriminant validity was established via the heterotrait-monotrait ratio <0.85. Collinearity was assessed using the variance inflation factor (VIF<3.6). Model fit was assessed using standardized root mean square residual (SRMR=0.069 for caregivers; 0.075 for care recipients). Path significance was tested using bootstrapping with 5000 resamples, and predictive relevance (Q^2=0.429 for caregivers; 0.499 for care recipients) was examined via blindfolding.

To test dyadic relationships, the APIM was estimated in Mplus Version 8.3 using MLR, consistent with established applications of dyadic data analysis [32,39,40]. Dyads were treated as distinguishable (caregiver vs care recipient). Because these roles are theoretically and structurally nonexchangeable, an empirical distinguishability test was not required [41]. Missing data were handled using full-information maximum likelihood (FIML), which used all available data under the missing-at-random assumption. In the APIM variables, there were no missing data for the DL predictors, whereas QoL outcomes had 25% (18/72) missing values for caregivers and 18% (13/72) for care recipients.

To assess the robustness of FIML estimates, 3 sensitivity analyses were conducted: Little MCAR test to verify the missing-at-random assumption, expectation-maximization (EM) imputation to assess potential bias from missingness, and APIM re-estimation using listwise deletion (n=47 complete dyads) to compare with FIML results. These analyses were performed in IBM SPSS Statistics 26 (Little MCAR, EM imputation) and Mplus 8.3 (listwise deletion APIM).

It is important to note that while APIM permits the examination of actor and partner effects, the cross-sectional design limits the ability to infer causality or temporal precedence. The APIM framework is used here to model patterns of association and interdependence within dyads at a single time point, consistent with standard applications in cross-sectional dyadic research [34].

The APIM accounts for actor effects (an individual's DL → their own QoL) and partner effects (one member's DL → the other's QoL). Residuals of the 2 QoL outcomes were allowed to covary. Standardized coefficients (β), standard errors, 95% CIs, and explained variance (R^2=0.619 for caregivers; 0.595 for care recipients in the basic model) were reported. Because both APIM specifications (basic and multivariate) were saturated (df=0), global fit indices (Comparative Fit Index, Tucker-Lewis Index, root mean square error of approximation, SRMR) were not interpreted, and modification indices were not applicable. Model evaluation focused on path coefficients and explained variance.

Two APIM specifications were estimated. The basic model used each member's composite DL score to predict their own QoL (actor effect) and their partner's QoL (partner effect). Composite scores were used in the model because latent APIM requires substantially larger samples for stable estimation, and APIM is commonly estimated using observed scores in small to moderate samples [39,41]. Additionally, the high reliability of the measures (care recipient >0.90), as shown in Table 4, supports the use of composite scores as valid proxies for the latent constructs. The multivariate model simultaneously included the 3 EDLQ dimensions (IC, CCM, and SS) as predictors of QoL, allowing estimation of domain-specific actor and partner effects while retaining dyad distinguishability and correlated residuals. All analyses were 2-tailed, and statistical significance was set at P≤.05.

Results

Participant Characteristics

Figure 2 presents the participant recruitment and selection flow. Of the 100 dyads approached, 72 caregiver-care recipient dyads were included in the analysis. The mean age of caregivers was 51 (SD 7.4) years and that of care recipients was 75.3 (SD 4.5) years. Most caregivers were female (43/72, 60%). Among care recipients, 57% (41/72) were female. The sample was predominantly Malay, followed by Chinese and Indian participants. Educational levels were generally higher among caregivers than care recipients. About one-third of caregivers had a bachelor's degree or higher, compared to 11% (8/72) of care recipients. Caregivers reported better internet access than care recipients, more than half of whom had weak access. Most caregivers rated their health as good, whereas only one-third of

XSL•FO
RenderX

care recipients reported good health. Table 1 presents the characteristics of caregivers and care recipients.

Figure 2. Flowchart of participant recruitment and selection process.

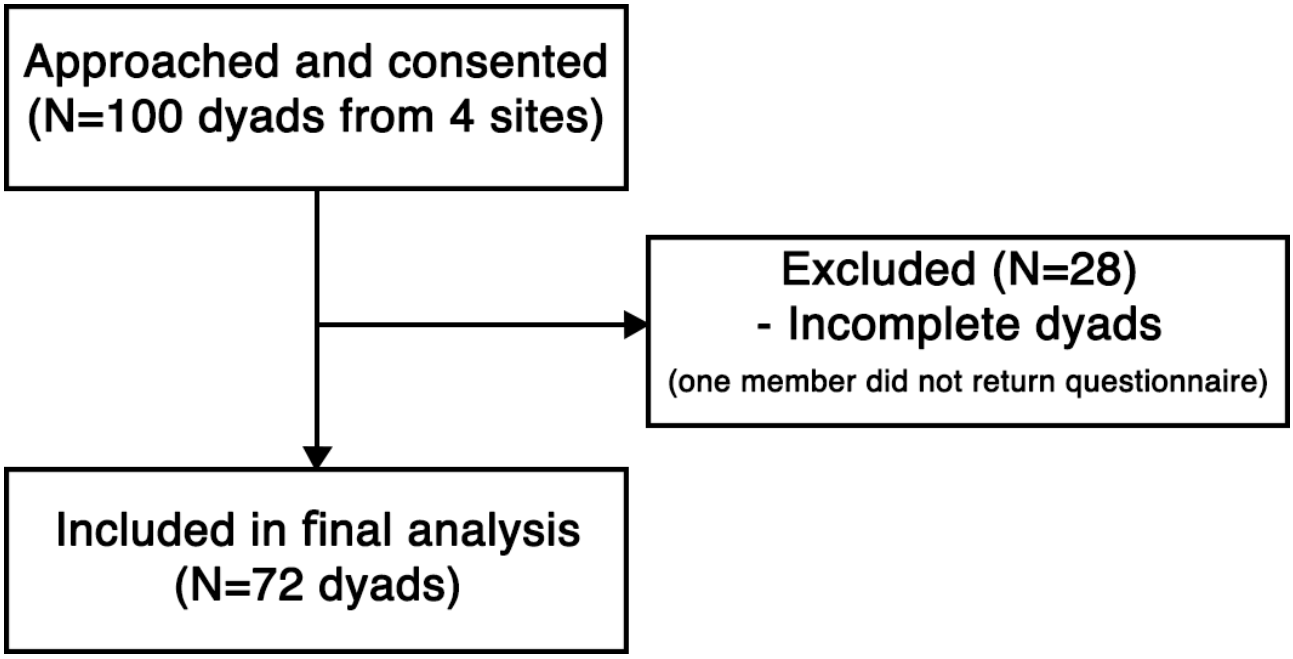


Table . Participant characteristics of caregivers and care recipients (N=72 dyads).

Characteristics	Caregivers	Care recipients
Age (y), mean (SD; range)	51 (7.4; 37 - 65)	75.3 (4.5; 61 - 85)
Gender, n (%[a])		
Female	43 (60)	41 (57)
Marital status, n (%)		
Married	45 (71)	42 (64)
Widowed	13 (21)	24 (36)
Ethnicity, n (%)		
Malay	38 (57)	36 (55)
Chinese	15 (22)	15 (23)
Indian	14 (21)	14 (22)
Education level, n (%)		
High (≥bachelor's degree)	25 (35)	8 (11)
Low (≤secondary)	22 (31)	44 (61)
Internet access, n (%)		
Strong	33 (46)	1 (1)
Moderate	22 (31)	32 (44)
Weak	17 (24)	39 (54)
Good self-rated health, n (%)	42 (58)	25 (35)

[a]Percentages may not sum to 100% due to rounding and missing data.

Descriptive Statistics and Correlations

Table 2 presents the descriptive statistics for the main study variables. Caregivers reported moderate levels of DL across the 3 dimensions (mean range 3.4 - 3.6; Table 2). Care recipients showed slightly lower scores, particularly for CCM. Caregivers had higher overall QoL scores than care recipients, and because care recipient QoL was nonnormally distributed, median (IQR) is also reported.

XSL•FO
RenderX

Table . Descriptive statistics of caregivers and care recipients (N=72 dyads).

Variables	Mean (SD)	Minimum-maximum	Median (IQR)
Caregivers			
Information and Communication	3.47 (0.74)	1.63-5.00	—[a]
Content Creation and Management	3.44 (0.76)	1.25-5.00	—
Safety and Security	3.60 (0.78)	1.75-5.00	—
Overall QoL[b] (0 - 100)	64.46 (14.68)	37.50-97.28	64.36 (20.45)[c]
Care recipients			
Information and Communication	3.23 (0.78)	1.38-4.86	—
Content Creation and Management	2.57 (0.70)	1.00-4.00	—
Safety and Security	3.17 (0.72)	1.43-4.71	—
Overall QoL (0 - 100)	55.27 (16.34)	19.01-84.84	60.26 (21.35)

[a]Not applicable.

[b]QoL: quality of life.

[c]Median and IQR are reported for QoL variables.

The 3 DL dimensions were positively intercorrelated within each group. All DL dimensions were significantly and positively correlated with QoL within each group (all $P<.001$). Cross-dyad correlations also emerged: caregivers' DL dimensions were positively associated with care recipients' QoL, and vice versa. These preliminary results suggest potential interdependence between caregiver and care recipient DL and QoL, supporting the rationale for subsequent APIM analyses. Pearson correlations among study variables (computed using pairwise deletion) are shown in Table 3.

Table . Pearson correlation coefficients among digital literacy dimensions and QoL[a] (N=72 dyads).

Variable	1	2	3	4	5	6	7	8
CG[b] information and communication	—[c]	0.57	0.64	0.18	0.56	0.08	0.63	0.48
CG content creation and management	0.57	—	0.53	0.12	0.46	0.19	0.59	0.56
CG safety and security	0.64	0.53	—	0.15	0.62	0.13	0.68	0.48
CR[d] information and communication	0.18	0.12	0.15	—	0.34	0.62	0.49	0.59
CR content creation and management	0.56	0.46	0.62	0.34	—	0.25	0.96[e]	0.68
CR safety and security	0.08	0.19	0.13	0.62	0.25	—	0.40	0.49
CG QoL	0.63	0.59	0.68	0.49	0.96[e]	0.40	—	0.74
CR QoL	0.48	0.56	0.48	0.59	0.68	0.49	0.74	—

[a]QoL: quality of life.

[b]CG: caregiver.

[c]Not applicable.

[d]CR: care recipient.

[e]The correlation between CR content creation and management and CG QoL (0.96) reflects the subsample with complete data (n=54). Note that the primary actor-partner interdependence model analysis uses FIML (full-information maximum likelihood; n=72) to handle missingness, yielding a robust but lower standardized estimate (β=.83) as reported in Table 7. A scatter plot confirming variable distinctness is provided in Multimedia Appendix 1.

XSL•FO
RenderX

Measurement and Partial Least Squares SEM

The measurement model demonstrated satisfactory reliability and validity. All constructs demonstrated adequate reliability and convergent validity (Table 4). Discriminant validity was established using heterotrait-monotrait ratio criteria (all values <0.85). VIF values indicated no multicollinearity among constructs (all VIF <3.6). Model fit was acceptable for both caregiver and care recipient models (SRMR=0.069 for caregivers; 0.075 for recipients).

Table . Construct reliability and convergent validity of measurement model (N=72 dyads).

Group and construct	Cronbach α	Composite reliability	AVE[a]
Caregivers			
CCM[b]	0.89	0.92	0.75
IC[c]	0.94	0.95	0.66
SS[d]	0.95	0.96	0.71
QoL[e]	0.95	0.97	0.87
Care recipients			
CCM	0.86	0.91	0.71
IC	0.93	0.94	0.63
SS	0.91	0.93	0.58
QoL	0.96	0.97	0.90

[a]AVE: average variance extracted.

[b]CCM: content creation and management.

[c]IC: information and communication.

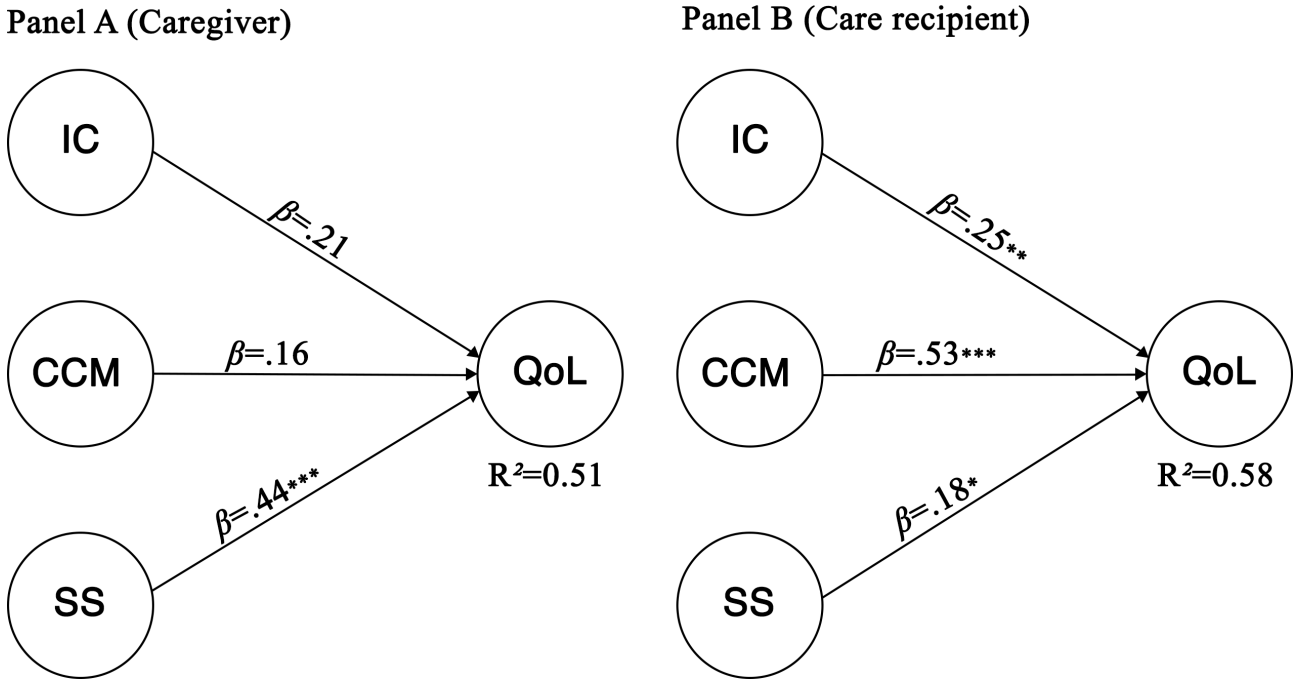
[d]SS: safety and security.

[e]QoL: quality of life.

The structural model explained 51% (37/72) of the variance in caregivers' QoL and 58% (42/72) of the variance in care recipients' QoL. Predictive relevance was supported for both models. Among caregivers, only the SS dimension significantly predicted QoL (β=.44, *P*<.001), whereas other dimensions were not significant. For care recipients, all 3 dimensions significantly predicted QoL, with CCM showing the strongest effect (β=.53, *P*<.001). Effect size analyses supported these patterns, indicating large effects for CCM among care recipients and for SS among caregivers. Figure 3 shows the structural partial least square-SEM models with standardized path coefficients and explained variance (R^2) for caregivers (panel A) and care recipients (panel B). Full results are displayed in Table 5.

XSL•FO
RenderX

Figure 3. Structural partial least squares structural equation modeling models for caregivers (panel A) and care recipients (panel B), showing standardized path coefficients (β) and explained variance (R^2). N=72 dyads. CCM: content creation and management; IC: information and communication; QoL: quality of life; SS: safety and security. *P<.05, **P<.01, ***P<.001.



The partial least squares-structural equation modeling analyses established the validity of the measurement model and identified dimension-specific associations between DL and QoL within each role. These analyses treated caregivers and care recipients as independent groups. To examine bidirectional interdependence between dyad members as a central feature of caregiving relationships, we next employed the APIM, which explicitly accounts for the nonindependence of paired observations and tests whether each member's DL predicts both their own (actor effect) and their partner's (partner effect) QoL.

Table . Structural model results from partial least squares-structural equation modeling (bootstrapping).

Group and path	β	*t*	*P* value	R^2 (QoL[a])	f^2	Q^2
Caregivers						
CCM[b] → QoL	0.16	1.94	.052	0.51	0.03	0.43
IC[c] → QoL	0.21	1.83	.07	—[d]	0.05	—
SS[e] → QoL	0.44	4.36	<.001	—	0.21	—
Care recipients						
CCM → QoL	0.53	7.68	<.001	0.58	0.57	0.50
IC → QoL	0.25	2.74	.006	—	0.08	—
SS → QoL	0.18	2.01	.04	—	0.05	—

[a]QoL: quality of life.

[b]CCM: content creation and management.

[c]IC: information and communication.

[d]Not applicable.

[e]SS: safety and security.

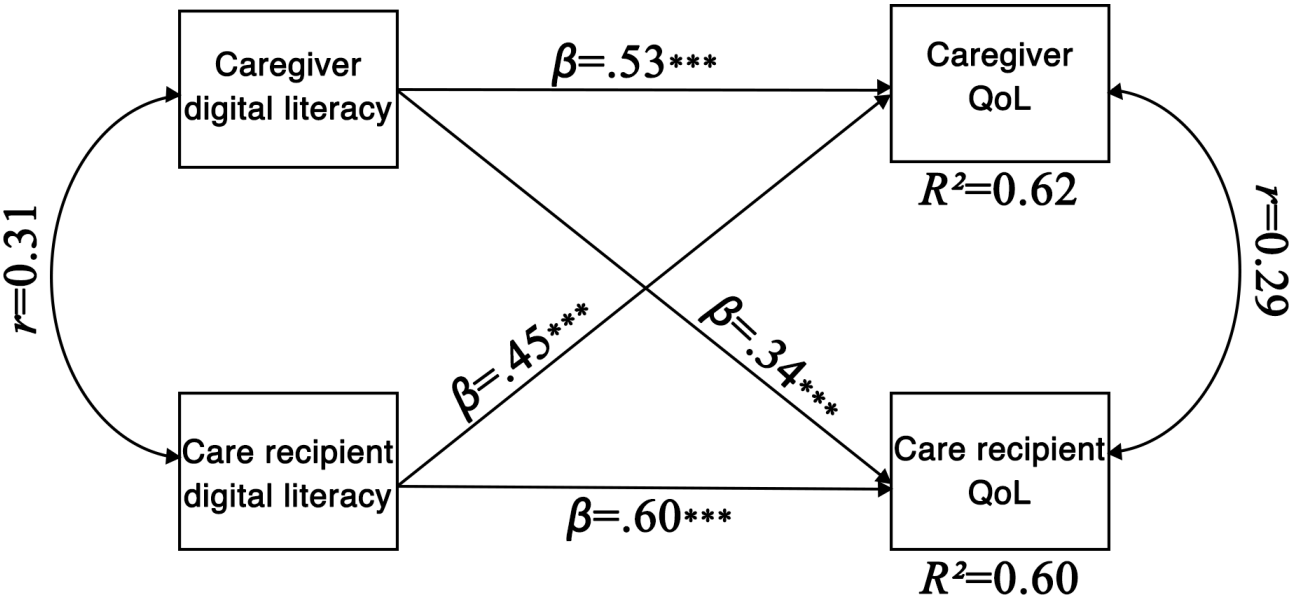
Dyadic Analysis (APIM)

The basic APIM was estimated for 72 dyads using MLR estimation with full-information handling of missing data. Observed composite scores were used for caregiver and care recipient DL and QoL. The model included actor and partner paths for both caregivers and care recipients, allowing for correlated predictors and residuals between dyad members. All 4 APIM paths were statistically significant. Actor effects were positive for both caregivers (β=.53, P<.001) and care recipients (β=.60, P<.001). Partner effects were also significant, from caregiver to care recipient (β=.34, P<.001) and from care recipient to caregiver (β=.45, P<.001). Predictors were moderately correlated (r=0.31, P=.004). The residual correlation between caregiver and care recipient outcomes was small but significant (r=0.29, P=.045). Variance explained was substantial,

XSL•FO
RenderX

with R^2=0.619 for caregivers' QoL and R^2=0.595 for care recipients' QoL. These R^2 values represent the combined variance explained in each outcome by both actor and partner effects. Figure 4 shows the basic APIM with standardized path coefficients and explained variance. Full statistical results are presented in Table 6.

Figure 4. Basic actor-partner interdependence model of digital literacy predicting quality of life, showing standardized path coefficients (β), correlations (r), and explained variance (R^2). N=72 dyads; QoL: quality of life. *P<.05, **P<.01, ***P<.001.



In the multivariate APIM, caregiver and care recipient QoL were simultaneously regressed on 3 dimensions of DL (IC, CCM, and SS) for both members of the dyad. For caregiver QoL, a significant actor effect was observed for caregiver SS (β=.11, P=.022, 95% CI 0.02 - 0.20). Other caregiver DL dimensions were not significant predictors. Among partner effects, care recipient CCM strongly predicted caregiver QoL (β=.83, P<.001, 95% CI 0.75 - 0.91), whereas the other recipient dimensions were nonsignificant. The predictors explained 92% of the variance in caregiver QoL (R^2=0.92). Model diagnostics indicated high intercorrelations among predictors, suggesting that the large R^2 primarily reflects shared variance rather than independent predictive effects. For care recipient QoL, significant actor effects were found for IC (β=.33, P<.001, 95% CI 0.15 - 0.51) and CCM (β=.40, P<.001, 95% CI 0.21 - 0.59). The effect of SS was positive but did not reach significance (β=.16, P=.072). Among partner effects, caregiver CCM significantly predicted care recipient QoL (β=.20, P=.047, 95% CI 0 - 0.40), while other caregiver dimensions were nonsignificant. The predictors explained 67% of the variance in care recipient QoL (R^2=0.67). Figure 5 shows the multivariate APIM with standardized path coefficients and explained variance. Full statistical results are presented in Table 7.

Table . Estimates from the actor-partner interdependence model of digital literacy predicting quality of life (N=72 dyads).

Path[a]	Role	*B* (SE)	*Z*	*P* value	β/*r* (95% CI)
CG[b] DL[c] → CG QoL[d]	Actor (CG)	0.55 (0.09)	6.00	<.001	β=0.53 (0.37-0.68)
CR[e] DL → CG QoL	Partner (CR→CG)	0.56 (0.12)	4.68	<.001	β=0.45 (0.28-0.61)
CR DL → CR QoL	Actor (CR)	0.87 (0.15)	5.75	<.001	β=0.60 (0.45-0.75)
CG DL → CR QoL	Partner (CG→CR)	0.40 (0.11)	3.82	<.001	β=0.34 (0.16-0.51)
CG DL ↔ CR DL	Predictor correlation	45.60 (17.93)	2.54	.004	*r*=0.31 (0.10-0.52)
CG QoL ↔ CR QoL	Residual correlation	25.46 (14.81)	1.72	.045	*r*=0.29 (0.01-0.58)

[a]Estimator: maximum likelihood robust.

[b]CG: caregiver.

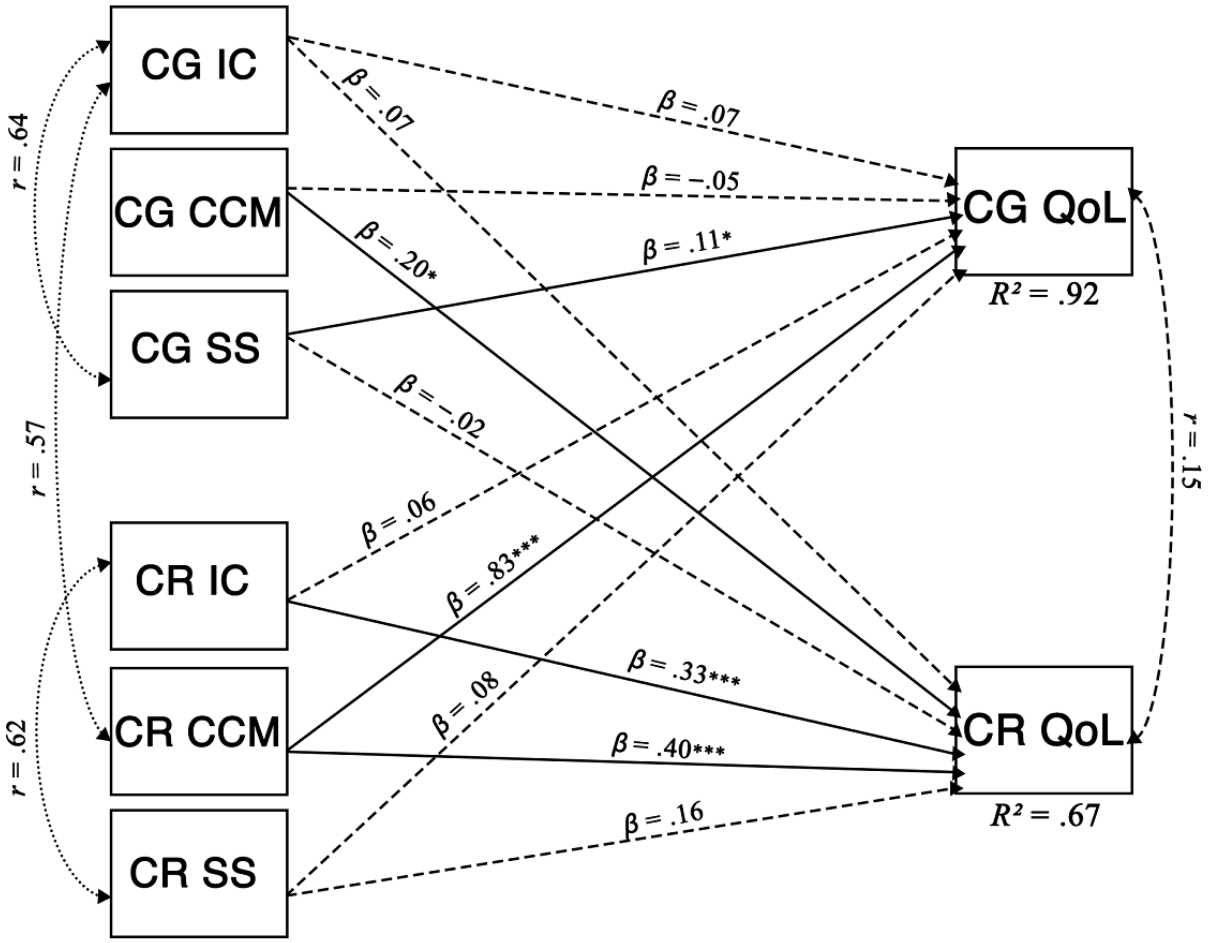
[c]DL: digital literacy.

[d]QoL: quality of life.

[e]CR: care recipient.

XSL•FO
RenderX

Figure 5. Multivariate actor-partner interdependence model of digital literacy dimensions predicting caregiver and care recipient QoL, showing standardized path coefficients (β) and explained variance (R^2). N=72 dyads. Solid arrows indicate significant paths (P<.05); dashed arrows indicate nonsignificant paths; dotted double-headed arrows represent significant correlations among predictors. *P<.05, **P<.01, ***P<.001. CCM: content creation and management; CG: caregiver; CR: care recipient; IC: information and communication; QoL: quality of life; SS: safety and security.



XSL•FO
RenderX

Table . Multivariate actor-partner interdependence model estimates of digital literacy dimensions predicting quality of life (N=72 dyads).

Outcome and path[a]	Role	*B* (SE)	*Z*	*P* value	β (95% CI)
CG[b] QoL[c]					
CG IC[d] → CG QoL	Actor (CG)	1.17 (1.02)	1.15	.25	0.06 (−0.04 to 0.17)
CG CCM[e] → CG QoL	Actor (CG)	−0.88 (1.09)	−0.80	.41	−0.05 (−0.16 to 0.07)
CG SS[f] → CG QoL	Actor (CG)	1.93 (0.86)	2.24	.02	0.11 (0.02 to 0.20)
CR[g] IC → CG QoL	Partner (CR→CG)	1.12 (0.97)	1.16	.25	0.06 (−0.04 to 0.17)
CR CCM → CG QoL	Partner (CR→CG)	16.49 (1.06)	15.63	<.001	0.83 (0.75 to 0.91)
CR SS → CG QoL	Partner (CR→CG)	1.48 (1)	1.48	.14	0.08 (−0.03 to 0.18)
CR QoL					
CR IC → CR QoL	Actor (CR)	6.82 (2.10)	3.25	.001	0.33 (0.15 to 0.51)
CR CCM → CR QoL	Actor (CR)	9.20 (2.30)	4.00	<.001	0.40 (0.21 to 0.59)
CR SS → CR QoL	Actor (CR)	3.66 (1.96)	1.87	.07	0.16 (−0.02 to 0.34)
CG IC → CR QoL	Partner (CG→CR)	1.56 (2.24)	0.70	.49	0.07 (−0.13 to 0.27)
CG CCM → CR QoL	Partner (CG→CR)	4.26 (2.25)	1.98	.05	0.20 (0 to 0.40)
CG SS → CR QoL	Partner (CG→CR)	−0.38 (2.24)	−0.17	.87	−0.02 (−0.23 to 0.19)

[a]Estimator: maximum likelihood robust.

[b]CG: caregiver.

[c]QoL: quality of life.

[d]IC: information and communication.

[e]CCM: content creation and management.

[f]SS: safety and security.

[g]CR: care recipient.

Table 8 shows correlations among predictors in the multivariate APIM. Several strong associations were found, particularly between caregiver IC and caregiver CCM (r=0.57, P<.001), between caregiver IC and SS (r=0.64, P<.001), and between care recipient IC and SS (r=0.62, P<.001). These findings indicate moderate-to-strong associations among DL dimensions.

XSL•FO
RenderX

Table . Correlations among predictors in the multivariate actor-partner interdependence model (APIM; standardized estimates, N=72 dyads).

Path[a]	*r*[b]	SE	Z	*P* value
CG[c] IC[d] ↔ CG CCM[e]	0.57	0.08	7.04	<.001
CG IC ↔ CG SS[f]	0.64	0.07	9.22	<.001
CG IC ↔ CR[g] IC	0.18	0.11	1.72	.09
CG IC ↔ CR CCM	0.56	0.08	6.94	<.001
CG IC ↔ CR SS	0.08	0.12	0.62	.53
CG CCM ↔ CG SS	0.53	0.08	6.53	<.001
CG CCM ↔ CR IC	0.12	0.12	0.98	.33
CG CCM ↔ CR CCM	0.46	0.08	5.67	<.001
CG CCM ↔ CR SS	0.19	0.11	1.68	.09
CG SS ↔ CR IC	0.15	0.10	1.47	.14
CG SS ↔ CR CCM	0.62	0.06	10.71	<.001
CG SS ↔ CR SS	0.13	0.13	1.01	.31
CR IC ↔ CR CCM	0.34	0.09	3.80	<.001
CR IC ↔ CR SS	0.62	0.06	10.99	<.001
CR CCM ↔ CR SS	0.25	0.10	2.47	.01

[a]Estimator: maximum likelihood robust.

[b]Reported values are standardized (STDYX) correlations among predictor variables in the multivariate APIM.

[c]CG: caregiver.

[d]IC: information and communication.

[e]CCM: content creation and management.

[f]SS: safety and security.

[g]CR: care recipient.

Sensitivity Analyses

Three sensitivity analyses confirmed the robustness of FIML-based findings. Among 72 dyads, 47 (65%) dyads had complete data, 12 (17%) dyads had missing care recipient QoL only, 7 (10%) dyads had missing caregiver QoL only, and 6 (8%) dyads had both QoL variables missing. Little MCAR test (χ^2_8=7.82, *P*=.45) indicated that data were missing completely at random, supporting the MAR assumption underlying FIML.

EM imputation showed minimal deviation from observed values. For caregiver QoL, EM-estimated mean was 64.38 (SD 13.96) versus the observed mean of 64.46 (SD 14.68), a difference of 0.08 points. For care recipient QoL, EM mean was 54.69 (SD 16.12) versus the observed mean of 55.27 (SD 16.34), a difference of 0.58 points.

Re-estimating the basic APIM using listwise deletion (n=47 dyads) produced results highly consistent with FIML (N=72 dyads), as shown in Table 9. All 4 APIM paths remained significant (all *P*<.001), effect directions were unchanged, and standardized coefficients differed by an average of only 0.04 units. The consistency across missing data approaches confirms that FIML estimates are robust and that substantive conclusions are unchanged.

XSL•FO
RenderX

Table . Comparison of actor-partner interdependence model parameter estimates using full-information maximum likelihood and listwise deletion (N=72 dyads)[a].

Parameter	FIML[b]	Listwise[c]	Difference[d]
Actor effects			
Caregiver DL[e] → Caregiver QoL[f]	0.53	0.52	0.01
Care recipient DL → Care recipient QoL	0.60	0.54	0.06
Partner effects			
Care recipient DL → Caregiver QoL	0.45	0.41	0.04
Caregiver DL → Care recipient QoL	0.34	0.39	0.05
Explained variance (R^2)			
Caregiver QoL	0.62	0.65	0.03
Care recipient QoL	0.60	0.65	0.05

[a]All path coefficients are standardized (β). All $P<.001$ for both full-information maximum likelihood and listwise deletion models. The average absolute difference across all 6 parameters was 0.04 units, indicating high consistency between methods.

[b]FIML: full-information maximum likelihood using all 72 dyads (18 caregivers and 13 care recipients with missing QoL data).

[c]Listwise deletion: complete-case analysis using 47 dyads with no missing QoL data.

[d]Absolute difference between FIML and listwise deletion estimates.

[e]DL: digital literacy.

[f]QoL: quality of life.

Discussion

DL and QoL: Role-Specific Patterns

The results of this study showed that DL has a multidimensional and role-specific relationship with QoL. The pattern of this association differed between caregivers and care recipients. Among caregivers, only the dimension of "SS" was a significant predictor of QoL. This finding suggests that the ability to navigate online spaces safely and maintain privacy was associated with a greater sense of control [16,28] and higher reported QoL.

Among care recipients, all 3 dimensions were associated with higher QoL. The largest effect was for "CCM," followed by "IC." This pattern suggests that care recipients who actively produce and organize digital content or engage in online communication are not just passive consumers of information but are also more likely to report greater autonomy, social connectedness, and overall satisfaction [29,30]. The findings are consistent with the study by Kyaw et al [7] among Korean older adults, which found that higher DL was associated with better perceived health and more technological self-care behaviors.

The APIM showed that the effects between caregivers and care recipients were reciprocal but asymmetric. Care recipients' CCM skills showed a strong partner effect on caregivers' QoL, while similar skills in caregivers showed a smaller but significant partner effect on care recipients' QoL. These results emphasize that DL is not a uniform skill but rather a multidimensional construct that operates through different yet interdependent pathways between members of the caregiving dyad. The models explained more than 50% of the variance in QoL, indicating strong and stable associations.

In psychological and behavioral terms, safety-related DL can be associated with better perceived outcomes for caregivers. This skill is linked to higher trust in online resources and lower exposure to digital risks. For care recipients, communication and content creation skills were associated with greater self-expression and social engagement, 2 key components of psychological and social well-being. Recent research has also shown that improving digital health literacy increases self-efficacy and health-promoting behaviors in both groups. Taken together, these findings suggest that DL should be viewed as a multidimensional and relational factor within caregiving contexts [12,42].

Dyadic Interdependence Between Caregivers and Care Recipients

The present study showed a clear pattern of bidirectional interdependence between caregivers and care recipients in DL and QoL. In the APIM, both the actor and partner effects were significant; that is, each individual's DL predicted not only their own QoL but also their partner's QoL.

In the basic model, interaction effects were observed in both directions. In the multivariate model, the paths were asymmetrical: CCM skills of care recipients showed a stronger association with caregivers' QoL, while similar skills in caregivers showed a smaller but significant association with care recipients' QoL. In contrast, safety-related DL in caregivers was associated only with their own QoL and was not significantly associated with the other party's QoL. This finding suggests that digital safety may contribute more to caregivers'

XSL•FO
RenderX

sense of competence and personal stability than to care recipients' direct outcomes.

This interdependence can be seen as part of the social and behavioral bonds in caregiving relationships [43]. When care recipients have higher DL, especially in communication and content creation, they become more active in self-care and digital interaction. This was associated with higher caregiver QoL and strengthens mutual reciprocity. On the other hand, caregivers' DL was associated with the perception of a safer and more supportive environment within online environments, which was linked to greater trust and confidence among care recipients.

Therefore, educational interventions should be 2-way and interactive, moving beyond solely individual-focused approaches; they should involve training that focuses on shared learning and collaborative use of technology [44,45]. In Asian societies, where the family is still the center of care, strengthening DL within caregiving dyads may improve mutual well-being and may help reduce the digital divide in older adults. Future programs should consider DL not simply as an individual skill but as a relationship-centered asset that enhances the QoL of both.

Digital Divide and Socio-Contextual Factors

In this study, care recipients had lower levels of DL than caregivers. Although this difference was not statistically tested, it is likely to reflect the "gray digital divide" that has also been reported in both global and regional studies. Recent meta-analyses suggest that eHealth literacy in older adults is still below optimal levels. This issue is particularly evident among women, the very old, single individuals, and those living in lower-income countries [46]. Similarly, in a national study conducted in China, differences in education, economic status, and media attitudes led to significant variations in technology use [47]. These underlying inequalities could explain the DL gap observed in Malaysian caregiving dyads.

As shown in Table 1, a significant portion of care recipients 54% (39/72) reported weak internet access, which may contribute to the digital divide and impact their QoL. At the national level, studies from Malaysia have shown that cognitive impairment and low socioeconomic status reduce the likelihood of using digital tools. Gender, age, and education also play a role in this pattern [48]. Broader research in Southeast Asia has also highlighted specific cultural barriers, such as fear of making mistakes, reliance on family, and limited rural infrastructure, that limit older adults' independence in technology use [49]. Therefore, improving internet access and digital infrastructure for older adults may help alleviate these challenges and reduce the burden on caregivers. Overall, the digital divide in the care context is rooted in cultural and social factors rather than simply individual ones. These inequalities may undermine older adults' digital self-efficacy and reduce their online participation. They may also place a greater educational burden on caregivers. Therefore, to achieve equitable well-being, digital inclusion programs should be designed at the family and community levels.

Cultural-Contextual and Policy Dimensions of DL in Malaysia

Digital inclusion among Malaysian older adults is not limited to individual skills alone. It is also shaped by cultural norms and national policies that determine how technology is used. In our data, care recipients had lower DL than caregivers, consistent with a common pattern in collectivist Asian cultures, where older adults rely on family for digital access and decision-making. Digital learning in this context is often achieved through intergenerational interaction rather than through self-learning [50-52]. This may explain why, in our model, care recipients' digital competence was associated with higher caregivers' QoL.

In policy terms, Malaysia's National Aging Policy recognizes the role of technology in enhancing independence but has yet to consider DL and participation as part of active aging. Similarly, Malaysia's Digital Economy Roadmap [21] focuses on innovation and competitiveness but does not address the specific needs of care recipients and informal caregivers. As a result, despite our findings showing that DL is associated with better outcomes for both groups, existing policies still view technology as an economic and individual phenomenon rather than a social and relational one. The results showed that caregivers' SS skills strongly predicted their QoL. Therefore, national aging programs should incorporate targeted cyber-safety modules for caregivers. Results also showed a strong effect of content-creation and communication skills on care recipient QoL. Hence, community DL initiatives should include training for care recipients to create, manage, and communicate digital content. These targeted areas align more closely with the dyadic and multidimensional patterns observed in this study. Together, these dyad-level recommendations operationalize a "2-way digital inclusion" approach, in which digital skills are developed reciprocally across care partners to enhance shared well-being and participation.

Policy and Practice

The findings of this study emphasize that DL is a multidimensional and bidirectional concept that is related to the QoL of both members of the caregiving couple. The observed effects highlight 2 priority areas for intervention: caregivers' SS skills and care recipients' information, communication, and content-creation abilities. These results indicate that DL interventions should be dimension-specific and relationship-oriented.

Research by Dong et al [24] suggests that theory-based training programs, delivered in brief face-to-face or blended formats, can improve self-efficacy and digital health literacy.

In Malaysia, this approach could be incorporated into the "MyAgeing Digital Literacy Modules" [53], which align with the EDLQ domains. Integrating such modules into Pusat Aktiviti Warga Emas (PAWE) centers [54] and the Digital Rakyat platform [21] may strengthen caregivers' cyber-safety practices and enhance care recipients' content-creation competencies. This integration is likely to be more effective when supported by intergenerational guidance and real-time assistance. These efforts can complement broader national initiatives, including

XSL•FO
RenderX

MyDIGITAL [21] and the Healthy Ageing 2030 plan [8], by promoting equitable and relationship-centered digital participation.

Limitations and Future Directions

The cross-sectional design of this study limits the ability to infer causality and examine behavioral changes over time. Longitudinal dyadic studies are needed to establish how changes in one member's DL may causally influence changes in the other's QoL over time.

Moreover, our use of a composite QoL score, while methodologically necessary for the APIM framework and empirically supported by very high intercorrelations among WHOQOL-BREF domains (r>0.85), precluded examination of domain-specific effects. Future research with larger samples could employ separate APIM models for each QoL domain to explore whether DL dimensions differentially predict physical, psychological, social, or environmental QoL.

In terms of participant screening, our reliance on a single-item subjective memory complaint screen to ensure reliable self-report may not have adequately excluded participants with mild cognitive impairment, particularly among care recipients aged 75 years and older. While this pragmatic approach was necessary for feasibility in a community-based survey, the absence of objective cognitive screening means that some participants with undetected cognitive impairment may have been included, which could affect the reliability of self-reported DL and QoL assessments.

Although the multivariate model explained a high percentage of the variation in caregivers' QoL (R^2=0.92), this should be interpreted with caution. The high correlations between the different dimensions of DL (Table 8) suggest that part of this value may be attributed to the overlap of the constructs rather than the independent effect of each dimension.

As this study used purposive sampling and was conducted only in urban areas of Klang Valley, the generalizability of the findings to care recipients and informal caregiver populations in semiurban or rural areas of Malaysia is limited. Klang Valley differs from many other states in terms of DL, internet access, and socioeconomic conditions; therefore, the patterns obtained are not necessarily representative of the entire country.

Furthermore, the Malay translation of the EDLQ was reviewed by a bilingual academic with expertise in gerontology for linguistic appropriateness, but formal forward-backward translation protocols and comprehensibility pretesting were not conducted. Although both English and Malay versions were offered based on respondents' language preference, the sample size did not permit measurement invariance testing to confirm equivalence across language versions. Future work with larger samples is needed to confirm measurement consistency across languages and respondent groups.

Future research should be conducted with more diverse samples from different regions to increase generalizability. Longitudinal or intervention studies could also shed more light on the causal pathways and lasting effects of digital education. Additionally, designing bidirectional and dimension-based programs with a focus on SS for caregivers and communication and content creation for care recipients can better demonstrate the mechanisms of digital empowerment. Finally, attention to digital equity and cultural adaptation is essential to improve access for less advantaged groups.

Conclusions

In summary, this study demonstrated that DL is a multidimensional and bidirectional factor that is associated with QoL in informal caregiving relationships. Considering both actor and partner effects, the results revealed that caregivers' safety skills and care recipients' communication and management abilities were jointly associated with the well-being of both parties. Integrating bidirectional and dimension-based interventions into Malaysia's national digital health and aging strategies can expand digital inclusion, reduce technological inequality, and may improve the experience of care. Future studies with longitudinal and educational designs can demonstrate how digital empowerment may enhance sustainable well-being in increasingly digital societies.

Acknowledgments

We sincerely thank the older adults and their family caregivers who participated wholeheartedly in this study. We are also grateful to the community centers and service organizations in the Klang Valley for their support in participant recruitment and data collection.

Funding

This study was part of the Transforming Cognitive Frailty to Later-Life Self-sufficiency (AGELESS) study funded by the Ministry of Higher Education Malaysia Long-Term Research Grant Scheme (LRGS/1/2019/UM/01/1/2). The study was also supported by the Research Management Centre (RMC), Universiti Putra Malaysia (UPM).

Data Availability

Data are not publicly available due to institutional policy to protect the privacy of research participants but are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

FMY and RI conceptualized the study. FMY, RI, and PK designed the methodology. FMY conducted data collection and investigation, performed formal analysis, managed the software, administered the project, and led the writing of the original draft. HFF contributed to manuscript writing, reviewing, editing, and revision. RI and PK supervised the project. AA reviewed the manuscript. All authors reviewed and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Scatter plot demonstrating variable distinctness.
[PNG File, 23 KB - aging_v9i1e86561_app1.png]

References

1. Bloom DE, Canning D, Lubet A. Global population aging: facts, challenges, solutions & perspectives. Daedalus 2015 Apr;144(2):80-92. [doi: 10.1162/DAED_a_00332]
2. Roth DL, Fredman L, Haley WE. Informal caregiving and its impact on health: a reappraisal from population-based studies. Gerontologist 2015 Apr;55(2):309-319. [doi: 10.1093/geront/gnu177] [Medline: 26035608]
3. Pinquart M, Sörensen S. Spouses, adult children, and children-in-law as caregivers of older adults: a meta-analytic comparison. Psychol Aging 2011 Mar;26(1):1-14. [doi: 10.1037/a0021863] [Medline: 21417538]
4. Godwin KM, Swank PR, Vaeth P, Ostwald SK. The longitudinal and dyadic effects of mutuality on perceived stress for stroke survivors and their spousal caregivers. Aging Ment Health 2013;17(4):423-431. [doi: 10.1080/13607863.2012.756457] [Medline: 23323629]
5. Chow EO, Fung SF, Singh H. Actor-partner effects of wellbeing, hope and self-esteem on depression in stroke survivor-caregiver dyads: a randomized controlled trial. Clin Rehabil 2023 Mar;37(3):394-406. [doi: 10.1177/02692155221128758] [Medline: 36453001]
6. World population ageing 2019: highlights. : United Nations Department of Economic and Social Affairs, Population Division; 2019 URL: https://digitallibrary.un.org/record/3846855?ln=en&v=pdf [accessed 2025-10-13]
7. Kyaw MY, Aung MN, Koyanagi Y, et al. Sociodigital determinants of eHealth literacy and related impact on health outcomes and eHealth use in Korean older adults: community-based cross-sectional survey. JMIR Aging 2024 Aug 13;7:e56061. [doi: 10.2196/56061] [Medline: 39140239]
8. Abdullah JM, Ismail A, Yusoff MSB. Healthy ageing in Malaysia by 2030: needs, challenges and future directions. Malays J Med Sci 2024 Aug;31(4):1-13. [doi: 10.21315/mjms2024.31.4.1] [Medline: 39247109]
9. Kang JLS, Cheah WL. Digital health literacy among elderly Dayak in southern Sarawak: a cross-sectional study. Malays J Med Health Sci 2025;21(2):29-38. [doi: 10.47836/mjmhs.21.2.5]
10. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. Inquiry 2022;59:469580221096272. [doi: 10.1177/00469580221096272] [Medline: 35471138]
11. Jawahir S, Tan EH, Tan YR, Mohd Noh SN, Ab Rahim I. The impacts of caregiving intensity on informal caregivers in Malaysia: findings from a national survey. BMC Health Serv Res 2021 Apr 27;21(1):391. [doi: 10.1186/s12913-021-06412-5] [Medline: 33906646]
12. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023 Oct;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37824584]
13. Payton S, Hague C. Digital literacy across the curriculum. : Futurelab; 2010 URL: https://www.nfer.ac.uk/media/jnhety2n/digital_literacy_across_the_curriculum.pdf [accessed 2025-04-16]
14. Jang SH, Je NJ. The relationship between digital literacy, loneliness, quality of life, and health-promoting behaviors among the elderly in the age of COVID-19. Int J adv appl sci 2022 Apr;9(4):71-79. [doi: 10.21833/ijaas.2022.04.009]
15. Águas D, Paço M, Henriques A, Arriaga M, Costa A. Digital health literacy in the training of informal caregivers – community intervention. Eur J Public Health 2022 Oct 23;32(Suppl 3). [doi: 10.1093/eurpub/ckac131.338]
16. Soares S, Hoffmeister LV, Fernandes MDF, Henriques A, Costa A. The use of digital technologies in the promotion of health literacy and empowerment of informal caregivers: scoping review. JMIR Aging 2024 Apr 29;7:e54913. [doi: 10.2196/54913] [Medline: 38683655]
17. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9. [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
18. van Deursen AJ, van Dijk JA. The digital divide shifts to differences in usage. Inf Commun Soc 2014 May;16(3):507-526. [doi: 10.1177/1461444813487959]
19. National Policy for Older Persons (Dasar warga emas negara) [Report in Malay]. : Ministry of Women, Family and Community Development, Government of Malaysia; 2011 URL: https://www.kpwkm.gov.my/uploads/content-downloads/file_20250105165538.pdf [accessed 2026-04-04]

XSL•FO
RenderX

20. Tey NP. Population ageing in Malaysia. In: Abeykoon A, Murat N, Rocas G, Naraval AC, editors. Ageing: Thailand, Malaysia, Indonesia and Cambodia: Demographic Transition, Policy and Programmatic Responses: International Council on Management of Population Programmes (ICOMP); 2017:47-56.
21. Malaysia Digital Economy Blueprint (MyDIGITAL). URL: https://www.mydigital.gov.my [accessed 2025-10-14]
22. Bastoni S, Wrede C, da Silva MC, et al. Factors influencing implementation of eHealth technologies to support informal dementia care: umbrella review. JMIR Aging 2021 Oct 8;4(4):e30841. [doi: 10.2196/30841] [Medline: 34623314]
23. Xie B, Watkins I, Golbeck J, Huang M. Understanding and changing older adults' perceptions and learning of social media. Educ Gerontol 2012 Apr 1;38(4):282-296. [doi: 10.1080/03601277.2010.544580] [Medline: 22639483]
24. Dong Q, Liu T, Liu R, Yang H, Liu C. Effectiveness of digital health literacy interventions in older adults: single-arm meta-analysis. J Med Internet Res 2023 Jun 28;25:e48166. [doi: 10.2196/48166] [Medline: 37379077]
25. Lee S, Ory MG, Vollmer Dahlke D, Smith ML. Technology use among older adults and their caregivers: cross-sectional survey study. JMIR Aging 2024 May 1;7:e50759. [doi: 10.2196/50759] [Medline: 38717339]
26. Jin X, Zhang Y, Zhou M, et al. An actor-partner interdependence mediation model for assessing the association between health literacy and mHealth use intention in dyads of patients with chronic heart failure and their caregivers: cross-sectional study. JMIR mHealth uHealth 2025 Mar 6;13:e63805. [doi: 10.2196/63805] [Medline: 40048155]
27. Xu C, Qian L, Qiu Y, et al. Dyadic effects of eHealth literacy on health-promoting behaviors in older adults with stroke and their family caregivers: the mediating role of self-efficacy. J Psychosom Res 2025 Jun;193:112144. [doi: 10.1016/j.jpsychores.2025.112144] [Medline: 40347538]
28. Oh KM, Hong SR, Beran K, Song Y, Lee JA. Digital health literacy and self-efficacy in using digital health resources among caregivers of individuals with Alzheimer's disease and related dementias. Sage Open Aging 2025;11:30495334251398089. [doi: 10.1177/30495334251398089] [Medline: 41356564]
29. Baek EM, Choi B, Kang SH. Exploring the relationship between digital competency and life satisfaction among the elderly. Digit Health 2025;11:20552076251358756. [doi: 10.1177/20552076251358756] [Medline: 40893171]
30. Yu L, Yang HH, Ma D, Liu F, Sun M. Navigating ICT-related factors and life satisfaction in older adults: a social cognitive theory approach. SAGE Open 2025 Apr;15(2):21582440251341173. [doi: 10.1177/21582440251341173]
31. Goh ZZS, Ho MHR, Ng KYY, et al. Using the actor-partner interdependence model to explore the psychological impact of COVID-19 on anxiety in dyads of patients with cancer and caregivers. Int J Behav Med 2024 Feb;31(1):19-30. [doi: 10.1007/s12529-023-10154-5] [Medline: 36788172]
32. Mplus version 8 user's guide. : Muthén & Muthén; 2017 URL: https://www.statmodel.com/download/MplusUserGuideVer_8.pdf [accessed 2026-04-15]
33. Muthén LK, Muthén BO. How to use a Monte Carlo study to decide on sample size and determine power. Struct Equ Modeling 2002 Oct;9(4):599-620. [doi: 10.1207/S15328007SEM0904_8]
34. Kenny DA, Kashy DA, Cook WL. Dyadic Data Analysis: Guilford Press; 2006.
35. Mitchell AJ. The clinical significance of subjective memory complaints in the diagnosis of mild cognitive impairment and dementia: a meta-analysis. Int J Geriatr Psychiatry 2008 Nov;23(11):1191-1202. [doi: 10.1002/gps.2053] [Medline: 18500688]
36. Choi J, Choi S, Song K, et al. Everyday digital literacy questionnaire for older adults: instrument development and validation study. J Med Internet Res 2023 Dec 14;25:e51616. [doi: 10.2196/51616] [Medline: 38095999]
37. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL group. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/s0033291798006667] [Medline: 9626712]
38. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. a report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
39. Cook WL, Kenny DA. The actor–partner interdependence model: a model of bidirectional effects in developmental studies. Int J Behav Dev 2005 Mar;29(2):101-109. [doi: 10.1080/01650250444000405]
40. Ledermann T, Macho S, Kenny DA. Assessing mediation in dyadic data using the actor-partner interdependence model. Struct Equ Modeling 2011 Oct 5;18(4):595-612. [doi: 10.1080/10705511.2011.607099]
41. Kenny DA, Ledermann T. Detecting, measuring, and testing dyadic patterns in the actor-partner interdependence model. J Fam Psychol 2010 Jun;24(3):359-366. [doi: 10.1037/a0019651] [Medline: 20545409]
42. Li Z, Li D, Li Z. Digital literacy, life satisfaction, social capital, and multidimensional health behaviors among middle-aged and older adults in rural China: a nationwide cross-sectional study. Digit Health 2025;11:20552076251374114. [doi: 10.1177/20552076251374114] [Medline: 41018519]
43. Rand S, Forder J, Malley J. A study of dyadic interdependence of control, social participation and occupation of adults who use long-term care services and their carers. Qual Life Res 2017 Dec;26(12):3307-3321. [doi: 10.1007/s11136-017-1669-3] [Medline: 28786019]
44. Bach A, Thiel F. Collaborative online learning in higher education—quality of digital interaction and associations with individual and group-related factors. Front Educ 2024;9. [doi: 10.3389/feduc.2024.1356271]

XSL•FO
RenderX

45. Botas G, Pires S, Fonseca C, Ramos A. Reducing Caregiver Burden Through Dyadic Support in Palliative Care: A Systematic Review Focused on Middle-Aged and Older Adults. J Clin Med 2025 Aug 16;14(16):5804. [doi: 10.3390/jcm14165804] [Medline: 40869629]
46. Jiang X, Wang L, Leng Y, et al. The level of electronic health literacy among older adults: a systematic review and meta-analysis. Arch Public Health 2024 Nov 7;82(1):204. [doi: 10.1186/s13690-024-01428-9] [Medline: 39511667]
47. Zhou Y, He T, Lin F. The digital divide is aging: an intergenerational investigation of social media engagement in China. Int J Environ Res Public Health 2022 Oct 10;19(19):12965. [doi: 10.3390/ijerph191912965] [Medline: 36232259]
48. Md Fadzil NH, Shahar S, Singh DKA, et al. Digital technology usage among older adults with cognitive frailty: a survey during COVID-19 pandemic. Digit Health 2023;9:20552076231207594. [doi: 10.1177/20552076231207594] [Medline: 37868158]
49. Chen IC, Phanumartwiwath A. Bridging the gray digital divide: a cross-cultural qualitative study on digital inclusion and healthy aging in Germany, Japan, and Thailand. Telematics Inform Rep 2025 Jun;18:100218. [doi: 10.1016/j.teler.2025.100218]
50. Jin H, Qu Y. Association between intergenerational support, technology perception and trust, and intention to seek medical care on the internet among Chinese older adults: cross-sectional questionnaire study. J Med Internet Res 2025 Jan 6;27:e65065. [doi: 10.2196/65065] [Medline: 39761564]
51. Li S, Yin Y, Md Yunus M. Elderly People in the Digital Cultural Environment: Social media use, environmental stress, and intergenerational interaction. Environ Behav Proc J 2025;10(31):433-439. [doi: 10.21834/e-bpj.v10i31.6643]
52. Pan S, Jordan-Marsh M. Internet use intention and adoption among Chinese older adults: from the expanded technology acceptance model perspective. Comput Human Behav 2010 Sep;26(5):1111-1119. [doi: 10.1016/j.chb.2010.03.015]
53. Digital literacy program for older person. Malaysian Research Institute on Ageing (MyAgeing®). URL: https://myageing.upm.edu.my/services/training_courses/digital_literacy_program_for_older_person-64866?L=en [accessed 2025-10-23]
54. Pusat Aktiviti Warga Emas (PAWE). Jabatan Kebajikan Masyarakat. URL: https://wargaemas.jkm.gov.my/portal/article?id=77636dd6-4f2f-447f-9098-e783f8c3b398&category=AK01&parent_id=97a9e705-fef3-4287-81b4-441a84915acb [accessed 2026-04-13]

Abbreviations

APIM: actor-partner interdependence model
CCM: content creation and management
DL: digital literacy
EDLQ: Everyday Digital Literacy Questionnaire
EM: expectation-maximization
FIML: full-information maximum likelihood
IC: information and communication
MLR: maximum likelihood robust
QoL: quality of life
SEM: structural equation modeling
SRMR: standardized root mean square residual
SS: safety and security
VIF: variance inflation factor
WHOQOL-BREF: World Health Organization Quality of Life-BREF

Edited by PH Chau; submitted 27.Oct.2025; peer-reviewed by B Ong, L Zwar; revised version received 23.Dec.2025; accepted 28.Dec.2025; published 27.Apr.2026.

Please cite as:
Mohammadzadeh Yazd F, Foong HF, Ibrahim R, Kunasekaran P, Ashari A
Multidimensional Digital Literacy and Quality of Life Among Informal Care Dyads in Malaysia: Cross-Sectional Survey
JMIR Aging 2026;9:e86561
URL: https://aging.jmir.org/2026/1/e86561
doi:10.2196/86561

© Fereshteh Mohammadzadeh Yazd, Hui Foh Foong, Rahimah Ibrahim, Puvaneswaran Kunasekaran, Asmidawati Ashari. Originally published in JMIR Aging (https://aging.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited.

XSL•FO
RenderX

The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Integrating Perceived Psychosocial Outcomes Into the Assessment of Telecare Acceptability Among Slovenian Informal Caregivers of Older People: Cross-Sectional Survey Study

Vesna Dolničar[1], PhD; Nejc Berzelak[2], PhD; Sakari Taipale[3], PhD; Simona Hvalič-Touzery[1], PhD

[1]Faculty of Social Sciences, University of Ljubljana, Kardeljeva pl. 5, Ljubljana, Slovenia

[2]Social Protection Institute of the Republic of Slovenia, Ljubljana, Slovenia

[3]Department of Social Sciences and Philosophy, University of Jyväskylä, Jyväskylä, Finland

Corresponding Author:
Vesna Dolničar, PhD
Faculty of Social Sciences, University of Ljubljana, Kardeljeva pl. 5, Ljubljana, Slovenia

Abstract

Background: While telecare has considerable potential to improve home care services, benefit older adults, and support informal caregivers, its adoption remains limited. There is also a notable gap in empirical research on the factors that influence informal caregivers' behavioral intention to use telecare. In particular, existing theoretical acceptance models are insufficiently comprehensive, as they do not specifically address informal caregivers of older adults and have largely overlooked the psychosocial outcomes of engaging with telecare.

Objective: This study examined the behavioral intention of Slovenian informal caregivers of older adults to use telecare, focusing on the preuse stage of technology acceptance, by integrating the perceived psychosocial outcomes of telecare engagement (peace of mind, reduced anxiety, sense of safety, and perceived connectedness with the care recipient) into the Unified Theory of Acceptance and Use of Technology (UTAUT) model.

Methods: The study applied nested ordinary least squares regression analysis to survey data collected in 2021 in Slovenia from informal caregivers of older adults, all of whom were aged 40 years or older. Of 612 respondents in the dataset, 558 were retained after excluding current or former users of care-related assistive technologies; the regression analyses were based on the 526 participants with complete data on all model variables.

Results: The regression analysis used 3 nested models to examine predictors of the intention of informal caregivers of older adults to use telecare. The first model, which included only background variables (burden of care, financial coping, and caregiver sociodemographic characteristics), had limited explanatory power (R^2=0.03), with financial coping as the only significant predictor (β=−0.14, P=.005). Adding UTAUT predictors significantly improved the model's fit (R^2=0.49). Performance expectancy (β=0.52, P<.001), facilitating conditions (β=0.11, P=.005), and social influence (β=0.21, P<.001) emerged as additional significant predictors. Incorporating perceived psychosocial outcomes further improved the model's fit and increased the explained variance (R^2=0.54); this additional factor also showed the strongest association (β=0.41, P<.001). Patterns for the other predictors remained consistent across models, but including perceived psychosocial outcomes substantially reduced the coefficient for performance expectancy.

Conclusions: In this cross-sectional sample, perceived psychosocial outcomes were the predictors most strongly associated with behavioral intention in the extended UTAUT model, indicating that they may be important factors in understanding telecare acceptability among informal caregivers. The findings suggest that developers, policymakers, and health and social care providers should consider not only functional aspects of telecare but also the psychosocial needs of informal caregivers. As the study measured self-reported intention at the preuse stage rather than actual use, longitudinal and behavioral research is needed to establish whether these associations translate into telecare adoption.

(JMIR Aging 2026;9:e85534) doi:10.2196/85534

KEYWORDS

aging; behavioral intention to use; e-care; informal carers; informal care; older adults; psychosocial factors; psychosocial impacts; technology acceptance; telecare; Unified Theory of Acceptance and Use of Technology; UTAUT

XSL•FO
RenderX

Introduction

Background

Population aging and the increasing prevalence of multiple chronic conditions among older adults are amplifying demand for caregiving services [1,2]. In this context, informal caregivers, who provide unpaid care to someone with a long-term illness, disability, or other long-lasting health or care need outside professional or formal frameworks, are essential [3]. Globally, the number of individuals providing such care is increasing [4,5], with informal caregivers representing a substantial share of the population in many countries, such as 9% in the United Kingdom and 11.5% in the United States [6-8]. In Europe, they provide 80% of all care [3,9,10], making them the largest providers of social care. In Slovenia, informal caregivers form the backbone of the long-term care system [11], with 23.6% of the adult population providing informal care to older adults [12]. This reflects a familialist care model in which the caring responsibilities for older people primarily fall on the family [13,14]. Despite this reliance, strong policies to support and facilitate these responsibilities are lacking [3,11,14].

Caregiving responsibilities can be extensive and demanding. They often place a significant burden on informal caregivers and affect their physical and emotional well-being [11,15-18]. They are likely to experience significant stress when they lack the internal and external resources needed to manage their caregiving responsibilities [19].

In this context, there is growing evidence supporting the potential benefits of telecare solutions in empowering and supporting informal caregivers [9,10,20-24]. In informal care for older adults, the term telecare refers to services that use digital technologies to enable older people to live more independently and safely in their homes for longer [25,26]. Telecare encompasses many services and devices that use digital technology to provide remote health care and social support, including personal alarm systems, trackers, environmental sensors, and reminders. Such services can also help informal caregivers manage their caregiving responsibilities more efficiently while offering peace of mind by providing reassurance about the well-being of the person they care for [25,26]. These technologies can thus represent supportive external resources that can relieve some of the pressure on informal caregivers.

While telecare has considerable promise for improving home care services, benefiting older people, and helping informal caregivers, its adoption remains limited [20,23,26]. Nevertheless, according to a report by Berg Insight [27], the number of telecare users in Europe was 9.0 million at the end of 2023 and is expected to exceed 13.1 million by 2028. This growth, fueled by an aging population and the increasing prevalence of chronic diseases [1,2], indicates the scale of potential telecare use. Notwithstanding the growing need and potential demand, several challenges hinder the wider implementation of telecare solutions [9,28,29]. In particular, their acceptability to informal caregivers is a key barrier [16,25,26,30].

Theoretical Framework and Goal of the Study

Although telecare offers significant potential in informal caregiving, there is a lack of research into the factors that influence informal caregivers' behavioral intention to use it [23,26]. Existing research in this area typically focuses on the classic variables of technology acceptance models: usefulness and ease of use. Several theoretical models have been applied to explore the factors influencing the intention to use telecare [31-34], including the health belief model [35], theory of reasoned action [36], theory of planned behavior [36], diffusion of innovations theory [37], domestication theory [38], technology acceptance model [39], and Unified Theory of Acceptance and Use of Technology (UTAUT) [40]. Among these, UTAUT was selected as the theoretical framework for this study because it integrates key determinants from earlier technology acceptance models and has been widely applied in studies examining the acceptance of digital health and assistive technologies, including research involving informal caregivers [32,41].

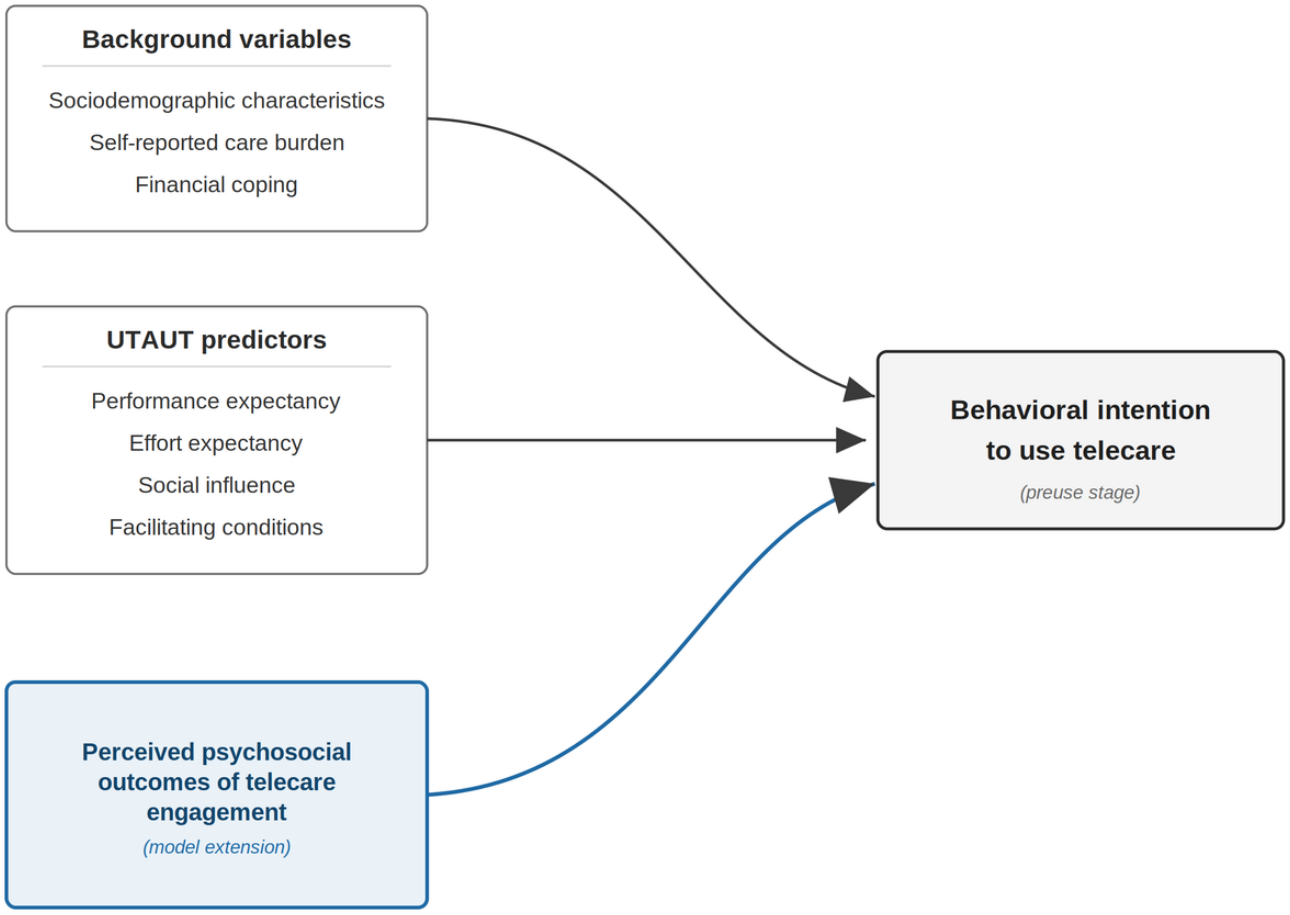
While UTAUT provides a solid foundation for examining factors influencing the intention to use telecare, it was not specifically developed to explain telecare use among informal caregivers of older adults and therefore does not adequately capture factors particularly relevant in health care settings [42] and informal caregiving contexts [43]. Furthermore, it and other telecare acceptance models have largely overlooked the psychosocial outcomes [23,43] associated with telecare engagement, such as peace of mind, reduced anxiety, and a sense of safety. These outcomes are particularly relevant in informal caregiving contexts because telecare technologies are often valued not only for functional support but also for the emotional reassurance they provide to caregivers [23,44-47]. Psychosocial outcomes of telecare engagement refer to the mediating effects of both internal factors (such as psychological and emotional responses) and external sociotechnical environments on an individual's social and psychological well-being [48-50]. In this context, telecare is part of the sociotechnical environment that shapes informal caregivers' responses while delivering unpaid care to older adults within a particular social and structural framework. The psychosocial outcomes of telecare engagement can then affect psychological and social functioning, feelings, behaviors, relationships, interactions, and cognitive processes [51]. A recent scoping study [23] identified peace of mind, reduced anxiety, a sense of safety, and perceived connectedness with the care recipient as the most common psychosocial outcomes of telecare engagement among informal caregivers.

To address the limitations of UTAUT and other acceptance models in this field and to enhance understanding of telecare acceptability among informal caregivers of older people, this study aims to integrate the perceived psychosocial outcomes of telecare engagement into the UTAUT model. The dependent variable in the UTAUT model is the behavioral intention to use telecare, defined as a measure of the strength of one's intention to perform a specified behavior [40,52-54]. Behavioral intention refers to a person's motivation to engage in a specific behavior [52]; the more a person is willing to perform a particular behavior, the more likely it is that the behavior will be performed [53,54].

XSL•FO
RenderX

The aim of this study is, therefore, to examine the behavioral intention of Slovenian informal caregivers of older adults to use telecare by integrating perceived psychosocial outcomes into the UTAUT model. According to the UTAUT model, behavioral intention is influenced by performance expectancy (the degree to which using the system is perceived to improve performance), effort expectancy (the degree of ease associated with system use), social influence (the degree to which important others are perceived to influence its use), and facilitating conditions (the degree to which resources and support are perceived to be available) [40]. We extend this model by including perceived psychosocial outcomes of telecare engagement as an additional predictor (Figure 1). Based on prior research [23,25,55-57], we hypothesize that perceived psychosocial outcomes will positively predict the behavioral intention to use telecare.

Figure 1. Conceptual model of telecare acceptability among informal caregivers, extending the Unified Theory of Acceptance and Use of Technology (UTAUT) framework with perceived psychosocial outcomes. The blue box represents the model extension. Perceived psychosocial outcomes comprise 4 indicators: peace of mind, reduced anxiety, sense of safety, and perceived connectedness with the care recipient.



This study thus addresses the following research question: how are background characteristics (sociodemographics, self-reported care burden, and financial coping), UTAUT constructs (performance expectancy, effort expectancy, social influence, and facilitating conditions), and the perceived psychosocial outcomes of telecare engagement associated with the behavioral intention of Slovenian informal caregivers of older people to use telecare? The study makes a theoretical, methodological, and practical contribution by extending the UTAUT model with psychosocial outcomes, providing empirical evidence on telecare acceptability among informal caregivers of older people in Slovenia, and offering insights relevant to the design and implementation of telecare solutions.

Following the technology acceptance lifecycle [43], which conceptualizes technology acceptance as a staged process, this study focuses on the initial preadoption stage (termed “preuse acceptability” [43]), which occurs prior to the actual first use of the technology, rather than on actual use or postadoption experiences. The research concentrates on 4 distinct telecare functionalities, grouped into 2 main categories: personal alarm systems, comprising the emergency call button and automatic fall detection, and mobility-related devices, comprising home activity sensors and location-tracking devices. Telecare acceptability is thus captured through behavioral intention to use, reflecting caregivers’ motivational readiness to use the technology in the preuse stage.

Methods

Study Design

The study population was Slovenian residents aged 40 years or older who had provided at least 4 hours of unpaid care per week to a family member or relative aged 65 years or over during the previous 6 months. The lower age threshold of 40 years was applied because international evidence indicates a higher prevalence of informal caregiving in this age group [58,59],

XSL•FO
RenderX

while comparable population-level data on informal caregivers were not available for Slovenia. Unpaid care was defined as the provision of activities of daily living, instrumental activities of daily living, or emotional support.

In the absence of appropriate sampling frames for informal caregivers, a survey with a probability sample of the general population was not feasible. Estimates of the prevalence of informal caregivers vary substantially across studies, depending on the definitions and characteristics of care recipients, which further complicates the construction of a reliable sampling frame [59]. Considering estimates of the proportion of informal caregivers in Slovenia in the target age group [60,61], such an approach would require large-scale screening to identify eligible individuals, which would result in prohibitive costs and practical challenges. We therefore used a nonprobability online panel with web survey data collection as a feasible approach to accessing this hard-to-reach population. This approach is increasingly used in social research to survey relatively small and difficult-to-identify population segments. Because nonprobability online panels may vary considerably in quality [62], we paid particular attention to the choice of panel provider, including an assessment against established ESOMAR guidelines for online panels [63]. The study requirements were clearly laid out, and the contract with the selected provider included a minimum sample size, within-panel selection and recruitment guidelines, control over questionnaire content and format, and clear reporting of the sample recruitment and data collection procedures.

Participants and Recruitment

Data were collected between April and May 2021. The panel provider sent invitations to individuals selected from the panel members by quota sampling. Quotas were defined by 10-year age groups, sex, and region to approximate the structure of the general population rather than that of the informal caregiver population due to the lack of reliable population-level data about informal caregivers. In addition, to improve the comparability of the sample with the population of informal caregivers, we incorporated screening questions into the representative Slovenian Public Opinion Survey [64] to identify family caregivers and obtain their sociodemographic characteristics. These data were used to calculate poststratification weights.

Eligible respondents were identified with a screening web questionnaire consisting of questions about the respondent's age, types of care provided to a family member, and the average number of hours spent weekly on care provision. Respondents with answers matching the definition of the target population were immediately routed to the main web questionnaire. Respondents who completed the questionnaire were assigned value points according to the standard incentivization practice of the panel provider. Information about voluntary participation, survey length, and data protection measures was provided, and no personally identifiable information about the participants was obtained. The respondents were allowed to skip individual questions and change their answers until the questionnaire was completed. To reduce nonresponse bias, the panel provider sent at least one email reminder to nonrespondents before attempting to contact additional panelists to meet the preset quotas. Access to the survey was limited to the invited panelists via unique access codes, and a respondent was able to participate in the survey only once.

Of 3284 panelists invited, 1834 completed the screening questionnaire and 781 were eligible. After excluding invalid, duplicate, and low-quality entries and respondents with insufficient completion or ineligible answers, 612 (18.6%) respondents remained in the final dataset. As the study focused on preuse behavioral intentions rather than user experience, 54 respondents were excluded because they were current or former users of care-related assistive technologies, such as emergency call services, fall detectors, sensors, or mobility trackers, or had failed to answer questions on prior use. This distinction is important, as factors influencing preuse behavioral intentions may differ from those shaping actual usage experience. This group was relatively small, reflecting the limited use of telecare services in Slovenia at the time of the study. A total of 558 participants were therefore included in the analysis. In summary, 1834 of the 3284 invited panelists completed the screening questionnaire, 612 were retained after exclusions, and 558 remained after removing current or former users of care-related assistive technologies. This sample of 558 was used for the reliability and factor analyses, while the regression analyses were based on the 526 participants with complete data on all model variables (see *Data Quality and Missing Data* section).

Instruments and Measures

Questionnaire

The main questionnaire had 64 questions, with several of these comprising multiple items (see Multimedia Appendix 1 for key items; full materials are available online [65]). The number of questions presented varied across respondents because inapplicable questions were automatically skipped based on relevant previous answers. The mean questionnaire completion time after screening was 17 minutes.

The questionnaire included items on participants' care provision and personal characteristics, as well as the study constructs. Self-reported care burden was measured using the Self-Rated Burden Scale [66], a single-item self-assessment scale ranging from 0 ("not burdensome at all") to 10 ("extremely burdensome"), with higher scores indicating greater burden. UTAUT constructs and behavioral intention were measured using items adapted from prior studies, while perceived psychosocial outcomes were assessed separately through items capturing peace of mind, reduced anxiety, sense of safety, and perceived connectedness with the care recipient. The questionnaire was developed by the research team and evaluated in a multistage process, including initial pilot testing with informal caregivers (n=93), cognitive interviews (n=9), expert review (n=4), and a final pilot test (n=49), to identify and address potential issues.

Measures of Behavioral Intention

In this study, behavioral intention represents a person's motivational readiness to engage with telecare [40,52,53], that is, how firmly they plan or are willing to use it. We operationalized the construct as the willingness to invest personal resources, specifically, time and money, to reflect the

XSL•FO
RenderX

commitment they are prepared to make to potentially use the technology. This operationalization reflects the typical cost of comparable telecare services in Slovenia and was adopted because pretesting indicated that respondents had difficulty assessing their intention to use such services without information on potential costs, particularly given the low levels of awareness and use of telecare at the time of data collection. Behavioral intention to use telecare was assessed as the average score of 2 items, adapted from the UTAUT behavioral intention items [40], and measured on 5-point agreement scales ranging from 1 (strongly disagree) to 5 (strongly agree): "I would be willing to devote some time to learn how to properly use the services" and "I would be willing to pay 25 EUR (€1=US $1.14 as of September 21, 2026) per month for the services."

UTAUT Measures and Perceived Psychosocial Outcomes

Four types of perceived psychosocial outcomes of telecare engagement identified in a prior scoping study [23] were measured: peace of mind, reduced anxiety, sense of safety, and perceived connectedness with the care recipient. For the exact wording of the questionnaire items, and their assignment to individual constructs, see Multimedia Appendix 1. All items were rated on a 5-point Likert-type scale from 1 (strongly disagree) to 5 (strongly agree). Four UTAUT constructs (performance expectancy, facilitating conditions, effort expectancy, and social influence) were measured with items adapted from prior studies [39,40,67], using the same response scale. Scores were calculated as means of the responses to their respective items.

Measurement Characteristics

The Cronbach α values for the scales with 3 or more items indicated good-to-acceptable reliability, with α=0.87 for perceived psychosocial outcomes, α=0.88 for performance expectancy, and α=0.62 for facilitating conditions. The somewhat lower α for facilitating conditions can be partially attributed to the scale's small item count (3 items). The mean interitem correlation (r=0.35) falls within the recommended range of 0.2 to 0.4 [68]; however, α is mathematically constrained by the number of items for a given level of interitem correlation [69]. Given this, the obtained alpha is generally considered acceptable in exploratory research [70].

A confirmatory factor analysis supported the expected one-factor structure of the measures for the 3 scales comprising 3 or more items. Standardized loadings were all significant (P<.001), ranging between 0.73 and 0.91 for perceived psychosocial outcomes, between 0.62 and 0.86 for performance expectancy, and between 0.52 and 0.70 for facilitating conditions. Model fit was evaluated using the root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root mean square residual (SRMR). Model fit was good for perceived psychosocial outcomes (χ^2_1=0.92, P=.34, RMSEA<0.001, CFI>0.99, TLI>0.99, SRMR=0.004) and for performance expectancy (χ^2_3=3.26, P=.35, RMSEA=0.013, CFI >0.99, TLI>0.99, SRMR=0.004). Because the factor model for facilitating conditions was just identified, fit indices could not be computed.

Common Method Bias

Given that the data were collected using a single self-report survey, we assessed the potential for common method bias using the Harman single-factor test. The first factor accounted for 41% of the variance, which is below the commonly accepted threshold of 50% [71]. These results indicate that a single factor did not account for the majority of the variance, suggesting that common method bias is unlikely to substantially affect the findings. However, the Harman single-factor test has recognized limitations and cannot conclusively rule out the presence of common method variance. To further mitigate this risk, several procedural remedies were applied, including the use of validated measurement scales, careful questionnaire design, comprehensive pretesting, and assurances of respondent confidentiality, thereby reducing the likelihood of systematic response bias and common method bias.

Data Analysis

Methods of Analysis

Univariate distributions of all the included variables were inspected before commencing the data analysis. The measurement properties of the latent constructs were evaluated using correlational analyses, reliability (internal consistency) analyses, and exploratory and confirmatory factor analyses.

We used nested ordinary least squares regression to determine the degree to which adding perceived psychosocial outcomes to the UTAUT model increases its explanatory power. Behavioral intention to use telecare was specified as the dependent variable, while background characteristics, UTAUT constructs, and perceived psychosocial outcomes were treated as independent variables. Since the dependent variable was constructed as the average of two 5-point agreement scale items, the resulting composite measure was treated as approximately continuous, following common practice in the literature. Accordingly, ordinary least squares regression was considered appropriate for the analysis, and the results were robust to alternative specifications using ordered logit models.

First, we fit a model containing only background variables (burden of care and the sociodemographic characteristics of caregivers) as predictors. Second, we added the scores for the 4 UTAUT predictors (performance expectancy, facilitating conditions, effort expectancy, and social influence). Finally, we added the score for perceived psychosocial outcomes.

Multicollinearity was assessed using variance inflation factors (VIFs). The maximum VIF among the independent variables was 3.97, with only 2 variables exceeding a value of 3. These values remain below commonly accepted thresholds, indicating that multicollinearity is not a concern.

Possibilities for creating weights to improve the match with the studied population were explored using estimates from another population survey based on a probability sample. However, the sufficient accuracy of the source estimates could not be reliably confirmed. In addition, evaluation of substantive differences between weighted and unweighted regression models, as well as a nonsignificant F test result [72], revealed no substantive effects of weighting on the regression estimates in any of the 3

XSL•FO
RenderX

models. We therefore used unweighted data for the analyses, while weighted regression results are reported in Multimedia Appendix 2 for comparison.

Data Quality and Missing Data

As noted above, respondents who did not complete at least 50% of the questionnaire were excluded from the analysis. A combination of response quality indicators was analyzed to identify potential low-quality responses, including straightlining indicators (frequency of selecting the same scale value to all items in grid questions), speeding (exceedingly short times on individual survey pages), and other response patterns (consistent tendencies to choose initial answers, midpoints, etc). No such patterns were observed among eligible respondents that would warrant further exclusions.

For the majority of statistical analyses, we used listwise deletion to maintain a consistent sample across the analyses. Of the 558 participants eligible for this study, 526 answered all items and were thus included in the regression analyses. For the reliability and factor analyses assessing measurement characteristics, we used all available data (n=558). Given the relatively low proportion of missing data, the resulting reduction in sample size is unlikely to have biased the results or meaningfully reduced statistical power.

Ethical Considerations

The study was conducted in accordance with Slovenian and international codes of medical ethics, declarations, and conventions and with the General Data Protection Regulation [73], and it was approved by the Committee for Ethics in Research at the Faculty of Social Sciences of the University of Ljubljana on March 12, 2021 (801-2021-020/JG). This study was designed and is reported in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys). The completed checklist is provided in Checklist 1.

Results

Characteristics of the Respondents

Table 1 presents the structure of survey respondents by key socioeconomic characteristics. They were, on average, 55.5 years old; the majority were female and in active employment. The relatively high average age and the proportion of individuals aged 70 years or older are in line with expectations, given the definition of the target population: a lower age bound of 40 years and the requirement of providing at least 4 hours of unpaid care, which becomes more likely with age.

XSL•FO
RenderX

Table . Sociodemographic characteristics, financial coping, and care burden of respondents (N=526).

Characteristic	Participants, n (%)
Sex	
Male	245 (46.6)
Female	281 (53.4)
Age (y)	
40 - 49	180 (34.2)
50 - 59	175 (33.3)
60 - 69	109 (20.7)
70 - 79	62 (11.8)
Education	
Low (ISCED[a] 1-2)	14 (2.7)
Medium (ISCED 3-4)	271 (51.5)
High (ISCED 5+)	241 (45.8)
Employment status	
Active	328 (62.4)
Retired	140 (26.6)
Other	58 (11.0)
Financial coping	
Without problems	202 (38.4)
Just coping	244 (46.4)
Difficult or very difficult	80 (15.2)
Self-reported care burden	
Lower (0 - 3)	267 (50.8)
Medium (4-6)	154 (29.3)
Higher (7-10)	105 (19.9)

[a]ISCED: International Standard Classification of Education.

The survey reached a low proportion of individuals with primary school education or less (International Standard Classification of Education [ISCED] 1 - 2), which is likely related to the structure of the panel from which the sample was obtained.

Slightly below half of the respondents reported being able to just cope financially, while a minority (n=80, 15.2%) found their financial situation difficult or very difficult. The self-reported care burden for the majority of respondents was low, with a higher burden identified among nearly a fifth of respondents.

Correlations Between Constructs

Behavioral intention, perceived psychosocial outcomes, performance expectancy, effort expectancy, and social influence were all positively correlated as expected (Table 2). Behavioral intention was particularly strongly correlated with perceived psychosocial outcomes and performance expectancy, whereas its correlation with facilitating conditions was relatively weak and the weakest among all correlations between behavioral intention and its predictors.

XSL•FO
RenderX

Table . Pearson correlations among Unified Theory of Acceptance and Use of Technology (UTAUT) constructs and perceived psychosocial outcomes (N=526).

Variable	BI	PPO	PE	FC	EE	SI
Behavioral intention (BI)						
r	1	0.68	0.66	0.23	0.32	0.49
P value	—[a]	<.001	<.001	<.001	<.001	<.001
Perceived psychosocial outcomes (PPO)						
r	0.68	1	0.83	0.11	0.36	0.47
P value	<.001	—	<.001	.02	<.001	<.001
Performance expectancy (PE)						
r	0.66	0.83	1	0.12	0.36	0.48
P value	<.001	<.001	—	.008	<.001	<.001
Facilitating conditions (FC)						
r	0.23	0.11	0.12	1	0.42	0.27
P value	<.001	.02	.008	—	<.001	<.001
Effort expectancy (EE)						
r	0.32	0.36	0.36	0.42	1	0.36
P value	<.001	<.001	<.001	<.001	—	<.001
Social influence (SI)						
r	0.49	0.47	0.48	0.27	0.36	1
P value	<.001	<.001	<.001	<.001	<.001	—

[a]Not applicable.

The correlations between the predictors varied considerably, with a strong correlation between performance expectancy and perceived psychosocial outcomes (r=0.83). This strong association is conceptually expected, as both constructs tap closely related evaluations of telecare's value; performance expectancy reflects the functional usefulness of telecare, whereas perceived psychosocial outcomes capture its anticipated emotional benefits (see *Discussion*). Despite the strength of this correlation, it did not produce multicollinearity problems in the regression models: the maximum VIF among the independent variables was 3.97, which is below commonly accepted thresholds (see *Data Analysis*).

Regression Analysis

Table 3 presents the results of 3 nested regression models. The first model contains only background variables (sociodemographics and self-reported care burden), the second model adds UTAUT predictors, and the third model adds perceived psychosocial outcomes.

XSL•FO
RenderX

Table . Nested regression models of factors associated with the behavioral intention to use telecare, including background variables, UTAUT[a] constructs, and perceived psychosocial outcomes (N=526).

Variable	Model 1: background variables[b]			Model 2: UTAUT predictors added[c]			Model 3: perceived psychosocial outcomes added[d]		
	β[e]	*t* test (*df*)	*P* value	β	*t* test (*df*)	*P* value	β	*t* test (*df*)	*P* value
Perceived psychosocial outcomes	—[f]	—	—	—	—	—	0.41	7.19 (512)	<.001
Performance expectancy	—	—	—	0.52	13.23 (513)	<.001	0.19	3.15 (512)	.002
Facilitating conditions	—	—	—	0.11	2.80 (513)	.005	0.12	3.32 (512)	<.001
Effort expectancy	—	—	—	−0.01	−0.18 (513)	.86	−0.02	−0.60 (512)	.55
Social influence	—	—	—	0.21	5.58 (513)	<.001	0.19	5.07 (512)	<.001
Self-reported care burden	0.01	0.25 (517)	.80	−0.04	−1.31 (513)	.19	−0.02	−0.76 (512)	.45
Sex (reference: male)									
Female	−0.02	−0.40 (517)	.69	0.02	0.63 (513)	.53	0.01	0.18 (512)	.86
Age	−0.09	−1.50 (517)	.14	−0.07	−1.46 (513)	.15	−0.05	−1.19 (512)	.24
Education	0.05	1.15 (517)	.25	0.02	0.64 (513)	.52	0.04	1.12 (512)	.26
Employment status (reference: active)									
Retired	0.02	0.38 (517)	.71	0.01	0.11 (513)	.91	0.01	0.21 (512)	.83
Other	−0.05	−1.00 (517)	.32	<0.01	−0.13 (513)	.90	<0.01	0.01 (512)	.99
Financial coping (reference: no problems)									
Just coping	−0.14	−2.79 (517)	.005	−0.10	−2.55 (513)	.01	−0.09	−2.43 (512)	.02
Difficult or very difficult	−0.06	−1.22 (517)	.22	−0.05	−1.28 (513)	.20	−0.04	−1.07 (512)	.29

[a]UTAUT: Unified Theory of Acceptance and Use of Technology.

[b]Model 1: *F* test (residual *df*)=2.07 (517), *P*=.04, R^2 (coefficient of determination)=0.03, adjusted R^2=0.02, Akaike information criterion=1236.30, Bayesian information criterion=1274.69.

[c]Model 2: *F* test (residual *df*)=41.06 (513), *P*<.001, R^2 (coefficient of determination)=0.49, adjusted R^2=0.48, R^2 change from Model 1=0.46; likelihood ratio test vs Model 1, *P*<.001, Akaike information criterion=906.80, Bayesian information criterion=962.25.

[d]Model 3: *F* test (residual *df*)=45.62 (512), *P*<.001, R^2 (coefficient of determination)=0.54, adjusted R^2=0.52, R^2 change from Model 2=0.05; likelihood ratio test vs Model 2, *P*<.001, Akaike information criterion=858.18, Bayesian information criterion=917.89.

[e]Standardized regression coefficient.

[f]Not applicable.

The model containing only background variables explains behavioral intention relatively poorly, with financial coping being the only significant predictor. The proportion of explained variance is low (3%), although the overall model fit is significant. The inclusion of UTAUT predictors significantly improves the model fit and its explanatory power, increasing the proportion of explained variance by 46 percentage points. The associations of performance expectancy, social influence, and facilitating conditions with behavioral intention are significant and positive. Consistent with the initial model, those struggling to cope financially are less willing to invest time and money.

The addition of perceived psychosocial outcomes again significantly improves model fit and also increases the explained variance. While this increase may seem modest, the added variable has the strongest effect size among all predictors, with a positive effect as expected. The pattern of observed relationships between the other predictors and the dependent variable is consistent with the previous model, except for performance expectancy, which has a markedly smaller effect size. The 2 constructs share a high proportion of the explained variance. Perceived psychosocial outcomes are associated with a reduction in the effect size of performance expectancy.

XSL•FO
RenderX

Discussion

Principal Findings and Comparison With Prior Work

In this study, perceived psychosocial outcomes emerged as the strongest predictor of informal caregivers' behavioral intention to use telecare, and adding these outcomes significantly improved the explanatory power of the UTAUT model. Performance expectancy, social influence, and facilitating conditions were also significant predictors, whereas effort expectancy and self-reported care burden were not.

This study extended the UTAUT model by incorporating perceived psychosocial outcomes to explain the behavioral intention of Slovenian informal caregivers of older adults to use telecare. While existing acceptance models provide a useful foundation, they do not consider psychosocial factors that are central to caregiving contexts [23,44-46]. To address this gap, we investigated whether including perceived psychosocial outcomes (such as reduced anxiety and increased peace of mind) improves the UTAUT model's explanatory power beyond conventional sociodemographic variables and standard UTAUT predictors (performance expectancy, effort expectancy, social influence, and facilitating conditions).

The model including only background variables showed limited explanatory power, which is consistent with previous research showing mixed or weak effects of sociodemographic factors on behavioral intention [74-76]. Financial coping (ie, having sufficient income) was the only statistically significant sociodemographic variable associated with behavioral intention. This finding aligns with previous studies in which coping financially was found to be an important factor in informal caregivers' acceptability of telecare [9,75-77]. Self-reported caregiving burden was not a significant predictor, although prior research has indicated that higher burden may increase the need for support [25,47,78]. One possible explanation is that, before caregivers have actually used telecare, their intention to do so may depend more on what they expect it to offer (its practical and psychosocial benefits) than on how much strain they are currently under. Such strain might instead translate into actual uptake only once the technology is available and its relevance to a specific caregiving situation becomes apparent. The single-item measurement of subjective burden may also have limited its predictive value, a point to which we return in the *Strengths and Limitations* section.

Adding the UTAUT factors significantly improved the model's explanatory power. Performance expectancy, social influence, and facilitating conditions were significant positive predictors, in line with previous research [23,25,56,79]. However, the relatively weak effect of facilitating conditions suggests that in the preadoption stage, structural and resource-related aspects may play a less prominent role compared with performance expectancy and social influence. Effort expectancy was not a significant predictor, which is likely due to the respondents' lack of direct experience with telecare [80]. Without hands-on experience, users may find it difficult to assess ease of use [23,80,81], whereas evidence identifying effort expectancy as a significant predictor largely comes from postimplementation research involving participants with direct experience of the technology [82-84]. This suggests that cross-sectional studies (especially those involving nonusers) may need to consider alternative or additional constructs. For instance, assessing caregivers' perceptions of the importance of ease of use may be more germane than asking them to estimate the ease of use of hypothetical or unfamiliar technologies.

Although UTAUT is a robust model for predicting behavioral intention to use telecare, our results show that extending it with perceived psychosocial outcomes is well justified. Their inclusion in the UTAUT model improved its explanatory power. Psychosocial outcomes showed the strongest association among all variables, indicating their key role in explaining informal caregivers' preuse behavioral intention to use telecare. This inclusion substantially reduced the traditional dominance of performance expectancy in explaining telecare acceptability. This may suggest that psychosocial outcomes partially mediate the relationship between performance expectancy and behavioral intention. However, given the study design, mediation cannot be formally established. An alternative interpretation is that psychosocial outcomes represent a distinct but closely related dimension of perceived usefulness that is not fully captured by standard measures of performance expectancy. In line with the original formulation of the technology acceptance model and its extensions, perceived usefulness (and performance expectancy) is typically defined in terms of functional and performance-related benefits, such as efficiency and effectiveness [40]. At the same time, research in telecare and caregiving contexts shows that emotional benefits, such as reassurance, peace of mind, and reduced anxiety, are central to how technologies are evaluated and are often treated as either implicit components of usefulness or as distinct experiential drivers of acceptability [56,85,86]. Notably, while quantitative models tend to emphasize functional utility as the primary predictor of acceptance, qualitative research consistently identifies emotional benefits as key motivations underlying caregivers' decisions [24,25,47,55], suggesting that these dimensions may be insufficiently captured in standard model operationalizations. The strong association between performance expectancy and psychosocial outcomes observed in this study is consistent with this partial conceptual overlap, with performance expectancy reflecting functional usefulness and psychosocial outcomes capturing anticipated emotional benefits, while their independent effects indicate that psychosocial outcomes contribute additional explanatory value. These findings indicate that psychosocial outcomes represent a central, yet still underrepresented, dimension of telecare acceptability in existing technology acceptance models.

Implications for Practice, Policy, and Future Research

These findings have practical implications for how telecare solutions are designed, implemented, and supported in policy. In addition to functional features, telecare services should address the psychosocial needs of informal caregivers by providing reassurance, reducing anxiety, and enhancing a sense of safety. These aspects are often overlooked in favor of technical features or the needs of care recipients, despite informal caregivers often codeciding on the use of telecare [22]. Developers, service providers, and policymakers should

XSL•FO
RenderX

therefore consider these dimensions when designing, communicating, implementing, and evaluating telecare solutions.

Based on our findings, future research should look at how the influence of psychosocial outcomes varies at different stages of the acceptance process. While our study focused on the initial stage of the user journey (ie, preuse acceptability), actual use of the technology and user engagement are likely to shape acceptance during both the initial and sustained use stages [43]. Consequently, the predictive strengths of specific acceptance factors, including psychosocial outcomes, may vary at these stages. Future studies could also further examine the interaction between functional and psychosocial determinants of telecare acceptance, including by testing whether psychosocial outcomes mediate the relationship between performance expectancy and behavioral intention and how these relationships evolve over time. In addition, there is a need to better integrate insights from qualitative research into quantitative models, particularly through the development of more refined measures of psychosocial outcomes.

Strengths and Limitations

A strength of this study is its focus on the theoretical framework for the acceptability of telecare by informal caregivers of older adults, which is often overlooked. The study also introduces a new dimension, namely the psychosocial impact of telecare use, which existing empirical models of telecare acceptability have largely neglected [23,43].

However, it has several limitations. As a cross-sectional survey of self-reported intention, the study describes associations rather than causal relationships, and behavioral intention at the preuse stage is not equivalent to actual telecare use.

First, the sampling relied on a nonprobability online panel due to the absence of a dedicated sampling frame for informal caregivers in Slovenia. This may introduce self-selection bias and limit the representativeness of the results, as individuals who participate in online panels may differ systematically from the broader population of informal caregivers. The lack of a reliable source of population-level data on caregivers in Slovenia limits our ability to assess how representative the sample is of the broader caregiver population. A comparison with a similar target group in an independent, probability-based population survey suggests that the panel sample somewhat overrepresents female (by 3.3 percentage points), highly educated (by 8.3 percentage points), and younger respondents (2.2 y younger on average), with urban and employed respondents also likely overrepresented. Because the 2 surveys differed in question design and coverage, a formal statistical comparison of sample structures was not possible, and these findings should be interpreted as indicative. In addition, respondents with greater familiarity with digital environments may be overrepresented, making the results less representative for less digitally engaged caregiver populations. At the same time, as telecare services typically require access to and basic proficiency with digital technologies, such as a smartphone and an internet connection, individuals without such access would be unlikely to use these services even if they were available. To mitigate this problem, we applied poststratification weights based on data from the representative Slovenian Public Opinion Survey [64], which improved the accuracy of population estimates. However, regression analyses were conducted on unweighted data, as weighting did not substantively affect the results. The findings should therefore be interpreted with caution when generalizing to caregiver populations that are less digitally engaged or to contexts outside Slovenia.

Second, as with all survey-based research, the survey instrument and context may have influenced the results. In particular, social desirability bias may affect responses, especially when reporting perceived psychosocial outcomes, which may result in more favorable assessments. Although self-administration and the assurance of confidentiality generally reduce such effects, their influence cannot be entirely excluded. Furthermore, because all variables were collected using the same survey instrument, the observed relationships may be partially influenced by common method variance, which could inflate correlations between constructs. However, the Harman single-factor test did not indicate that this is a major concern. In addition, caregiving burden was measured using a single-item indicator, which, although commonly applied, may not adequately reflect the multidimensional nature of caregiving experiences [66]. The facilitating conditions scale also showed limited, albeit acceptable, reliability, which indicates the need to improve the measurement characteristics of the scale.

Third, the study surveyed family caregivers aged 40 years and older because international research suggests caregiving prevalence peaks in middle adulthood and remains high into older age [87,88]. This decision is further supported by data from the Slovenian Public Opinion Survey [64], indicating that 77% of informal caregivers of older adults fall within this age range.

Fourth, contrary to expectations, care burden was not a significant factor in our model, which may be due to the use of a single-item subjective burden measure that may not adequately account for the complexity of caregiving experiences. Additionally, self-reported measures are prone to biases, such as social desirability, caregiver guilt, and individual coping mechanisms. Perceptions of burden may also fluctuate over time, further affecting the accuracy of the assessment. Future studies should therefore incorporate both subjective and objective measures of care burden to ensure a more complete assessment.

Fifth, the operationalization of behavioral intention also as willingness to pay for telecare services may introduce bias related to respondents' financial situation, as individuals with fewer financial resources may express lower intention independent of their attitudes toward the technology. To partially account for this, financial coping was included as a control variable in the regression models, although it may not capture all income-related influences on the measure. Relatedly, the 2 items used to operationalize behavioral intention capture conceptually distinct facets of motivational readiness: willingness to invest time reflects an attitudinal commitment to engaging with the technology, whereas willingness to pay reflects an economic commitment that is more strongly conditioned by personal financial circumstances. Treating them as a single measure is therefore a simplification, since it may

XSL•FO
RenderX

blur the line between the attitudinal and economic sides of intention. Future research should measure attitudinal intention and economic willingness separately to better untangle the role each one plays in the acceptability of telecare.

Finally, this study was conducted in a single country, and the relatively low representation of participants with lower educational levels may further limit the generalizability of the findings, as this group may have different levels of digital literacy and attitudes toward telecare. Differences in informal care systems, elder care practices, and the availability of telecare solutions may also limit the transferability of our findings to other contexts.

Despite these limitations, our findings offer a better understanding of informal caregivers' acceptability of telecare. They may help research and policy across different countries and for different technologies.

Conclusions

This study aimed to assess telecare acceptability among informal caregivers of older people and to identify relevant acceptance factors. Our findings show that adding perceived psychosocial outcomes of telecare engagement significantly improved the explanatory power of the UTAUT model and that these outcomes were the predictor most strongly associated with self-reported behavioral intention. This suggests that psychosocial factors are key to telecare acceptability within informal caregiving contexts. Their inclusion goes beyond improving the model empirically by also advancing it theoretically, as it incorporates the emotional aspects of caregiving (peace of mind, reduced anxiety, sense of safety, and perceived connectedness with the care recipient) that traditional technology acceptance models do not capture. Developers, policymakers, and health care providers should consider psychosocial outcomes when designing, testing, and deploying telecare solutions, as addressing informal caregivers' needs for reassurance, safety, and reduced anxiety may be important for their adoption. This can be translated into practice by incorporating caregiver perspectives into user-centered design, addressing psychosocial outcomes in implementation, and reflecting them in evaluation processes. Future research should examine how psychosocial outcomes influence technology acceptance at different stages, such as preadoption, initial adoption, and sustained use, to better understand how this influence changes over time.

Acknowledgments

The authors thank the informal caregivers who took part in pilot testing and cognitive interviews during questionnaire development, as well as the experts who reviewed the questionnaire. We also thank all the informal caregivers who participated in the survey.

The authors declare the use of generative AI (GenAI) in the preparation of this manuscript. According to the GAIDeT (Generative AI Delegation Taxonomy) taxonomy [89], the following tasks were delegated to a GenAI tool under full human supervision: proofreading and editing; and visualization, namely generating Figure 1 and the table of contents image. The tool used was GPT-5.6, including its built-in image-generation capability. No GenAI tool was used to generate original text, ideas, arguments, analyses, or interpretations. The AI-generated images are illustrative; they do not display study data and do not depict identifiable real individuals. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes; responsibility for the final manuscript lies entirely with the authors.

Funding

This work was supported by the Slovenian Research and Innovation Agency (grant numbers J5-1785, J5-4578, J5-70195, and P5-0399). ST's contribution to the research was funded by the Strategic Research Council (SRC) established within the Research Council of Finland (grant numbers 372415 and 372418). The funders had no involvement in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Data Availability

The dataset and codebook used in this study are publicly available at PsychArchives [65]. No restrictions apply to the availability of these data.

Authors' Contributions

VD and SH-T conceptualized and designed the study. VD, SH-T, NB, and ST developed the methodology. NB performed the analysis, and NB, VD, and SH-T interpreted the results. VD, NB, and SH-T drafted the manuscript. VD, SH-T, NB, and ST critically reviewed and revised the manuscript. All authors read and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Variables (constructs) and corresponding questionnaire items used in regression models.
[DOCX File, 21 KB - aging_v9i1e85534_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Comparison of regression models of factors associated with behavioral intention to use telecare based on weighted and unweighted data.
[DOCX File, 28 KB - aging_v9i1e85534_app2.docx]

Checklist 1
CHERRIES checklist.
[PDF File, 62 KB - aging_v9i1e85534_app3.pdf]

References

1. Wang L, Tang Y. Changing trends and the effectiveness of informal care among rural elderly adults in China. SAGE Open 2023 Oct;13(4):0-14. [doi: 10.1177/21582440231202580]
2. Chowdhury SR, Chandra Das D, Sunna TC, Beyene J, Hossain A. Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. EClinicalMedicine 2023 Mar;57:101860. [doi: 10.1016/j.eclinm.2023.101860] [Medline: 36864977]
3. Enabling carers to care: an EU strategy to support and empower informal carers—the 10 steps towards carer-friendly societies across Europe. : Eurocarers; 2020 URL: https://www.omastehooldus.eu/sites/default/files/public/eurocarers_strategy_final.pdf [accessed 2026-09-18]
4. The global carer well-being index: who carers for carers? Perspectives on COVID-19 pressures and lack of support. : Embracing Carers; 2021 URL: https://www.embracingcarers.com/wp-content/uploads/Global-Carer-Well-Being-Index-Report_FINAL.pdf [accessed 2026-09-18]
5. Global state of caring. : International Alliance of Carer Organizations; 2021 URL: https://www.aidants.fr/wp-content/uploads/2022/04/IACO-Global-State-of-Caring-July-13-MONDE_compressed.pdf [accessed 2026-09-18]
6. Egan KJ, Clark P, Deen Z, et al. Understanding current needs and future expectations of informal caregivers for technology to support health and well-being: national survey study. JMIR Aging 2022 Jan 27;5(1):e15413. [doi: 10.2196/15413] [Medline: 35084339]
7. State of caring 2022: a snapshot of unpaid care in the UK. : Carers UK; 2022 Nov URL: https://www.carersuk.org/media/p4kblx5n/cukstateofcaring2022report.pdf [accessed 2026-09-18]
8. Reinhard SC, Caldera S, Houser A, Choula RB. Valuing the invaluable: 2023 update strengthening supports for family caregivers. : AARP Public Policy Institute; 2023 URL: https://www.aarp.org/content/dam/aarp/ppi/2023/3/valuing-the-invaluable-2023-update.doi.10.26419-2Fppi.00082.006.pdf [accessed 2026-09-18]
9. Hassan AYI. Challenges and recommendations for the deployment of information and communication technology solutions for informal caregivers: scoping review. JMIR Aging 2020 Jul 29;3(2):e20310. [doi: 10.2196/20310] [Medline: 32723720]
10. European Commission, Directorate-General for Employment, Social Affairs and Inclusion (DG EMPL), in collaboration with the London School of Economics and Political Science (LSE), Zigante V. Social situation monitor: the role of new technologies in modernising long-term care systems: a scoping review. : Publications Office of the European Union; 2021 URL: https://op.europa.eu/en/publication-detail/-/publication/b2a95cef-61e9-11eb-aeb5-01aa75ed71a1/language-en [accessed 2026-09-22]
11. Hvalič-Touzery S, Trkman M, Dolničar V. Caregiving situation as a predictor of subjective caregiver burden: informal caregivers of older adults during the COVID-19 pandemic. Int J Environ Res Public Health 2022 Nov 4;19(21):14496. [doi: 10.3390/ijerph192114496] [Medline: 36361373]
12. Slovenian Public Opinion 2024/1 – summary of Results and Report on the Implementation of the Survey [Report in Slovenian]. : Center for Public Opinion and Mass Communication Research, University of Ljubljana, Faculty of Social Sciences; 2024 URL: https://www.cjm.si/wp-content/uploads/2025/08/SJM24_1-SUM-L-fin.pdf [accessed 2026-09-18]
13. Hlebec V, Rakar T, Dolničar V, Petrovčič A, Filipovič Hrast M. Intergenerational solidarity in Slovenia: key issues. In: Albert I, editor. Families and Family Values in Society and Culture: Information Age Publishing; 2021:359-380.
14. Filipovič Hrast M, Rakar T. Care policy in Slovenia: divergent trends and convergent attitudes. Revija za socijalnu politiku 2021 Dec 16;28(3):303-320. [doi: 10.3935/rsp.v28i3.1802]
15. Le DD, Ibuka Y. Understanding the effects of informal caregiving on health and well-being: heterogeneity and mechanisms. Soc Sci Med 2023 Jan;317:115630. [doi: 10.1016/j.socscimed.2022.115630] [Medline: 36580861]
16. Xiong C, D'Souza A, El-Khechen-Richandi G, et al. Perceptions of digital technology experiences and development among family caregivers and technology researchers: qualitative study. JMIR Form Res 2022 Jan 28;6(1):e19967. [doi: 10.2196/19967] [Medline: 35089150]
17. Dale J, Nanton V, Day T, et al. Uptake and use of Care Companion, a web-based information resource for supporting informal carers of older people: mixed methods study. JMIR Aging 2023 Sep 21;6:e41185. [doi: 10.2196/41185] [Medline: 37733406]
18. Lindt N, van Berkel J, Mulder BC. Determinants of overburdening among informal carers: a systematic review. BMC Geriatr 2020 Aug 26;20(1):304. [doi: 10.1186/s12877-020-01708-3] [Medline: 32847493]

XSL•FO
RenderX

19. Roth DL, Fredman L, Haley WE. Informal caregiving and its impact on health: a reappraisal from population-based studies. Gerontologist 2015 Apr;55(2):309-319. [doi: 10.1093/geront/gnu177] [Medline: 26035608]
20. Information & Communication Technology (ICT) for informal carers. : Eurocarers; 2021 URL: https://eurocarers.org/wp-content/uploads/2026/05/ICT-for-informal-carers.pdf [accessed 2026-09-18]
21. European Commission, Directorate-General for Employment, Social Affairs and Inclusion. “Long-term care report: trends, challenges and opportunities in an ageing society”. volume II, country profiles. : Publications Office of the European Union; 2021. [doi: 10.2767/183997]
22. Hvalič-Touzery S, Smole-Orehek K, Dolničar V. Exploring reciprocity in perceptions on telecare within the informal carer–care receiver dyad. Teor Praksa 2021 Nov 2;58(3):840-859. [doi: 10.51936/tip.58.3.840-859]
23. Hvalič-Touzery S, Dolničar V, Prevodnik K. Factors influencing informal carers’ acceptance of assistive telecare systems in the pre- and post-implementation phase: a scoping study. Health Soc Care Community 2022 Sep;30(5):e1484-e1504. [doi: 10.1111/hsc.13840] [Medline: 35574935]
24. Hvalič-Touzery S, Smole-Orehek K, Dolničar V. Informal caregivers’ perceptions of self-efficacy and subjective well-being when using telecare in the home environment: a qualitative study. Geriatrics (Basel) 2022 Aug 23;7(5):86. [doi: 10.3390/geriatrics7050086] [Medline: 36136795]
25. Cook EJ, Randhawa G, Guppy A, et al. Exploring factors that impact the decision to use assistive telecare: perspectives of family care-givers of older people in the United Kingdom. Ageing Soc 2018 Sep;38(9):1912-1932. [doi: 10.1017/S0144686X1700037X]
26. Kavčič M, Petrovčič A, Dolničar V. Model of acceptance and use of market-ready home-based e-care services: a qualitative study with care receivers and informal caregivers. Int J Hum Comput Interact 2023 Mar 16;39(5):988-1002. [doi: 10.1080/10447318.2022.2041898]
27. The Telecare Market in Europe and North America. : Berg Insight; 2024 URL: https://media.berginsight.com/2024/03/27192544/bi-telecare1-ps.pdf [accessed 2026-09-18]
28. Lindeman DA, Kim KK, Gladstone C, Apesoa-Varano EC. Technology and caregiving: emerging interventions and directions for research. Gerontologist 2020 Feb 14;60(Suppl 1):S41-S49. [doi: 10.1093/geront/gnz178] [Medline: 32057082]
29. Hopwood J, Walker N, McDonagh L, et al. Internet-based interventions aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018 Jun 12;20(6):e216. [doi: 10.2196/jmir.9548] [Medline: 29895512]
30. Holthe T, Jentoft R, Arntzen C, Thorsen K. Benefits and burdens: family caregivers’ experiences of assistive technology (AT) in everyday life with persons with young-onset dementia (YOD). Disabil Rehabil Assist Technol 2018 Nov;13(8):754-762. [doi: 10.1080/17483107.2017.1373151] [Medline: 28891356]
31. Evans N, Boyd H, Harris N, et al. The experience of using prompting technology from the perspective of people with dementia and their primary carers. Aging Ment Health 2021 Aug;25(8):1433-1441. [doi: 10.1080/13607863.2020.1745145] [Medline: 32223428]
32. Dai B, Larnyo E, Tetteh EA, Aboagye AK, Musah AAI. Factors affecting caregivers’ acceptance of the use of wearable devices by patients with dementia: an extension of the Unified Theory of Acceptance and Use of Technology model. Am J Alzheimers Dis Other Demen 2020;35:1533317519883493. [doi: 10.1177/1533317519883493] [Medline: 31679390]
33. Perez H, Miguel-Cruz A, Daum C, et al. Technology acceptance of a mobile application to support family caregivers in a long-term care facility. Appl Clin Inform 2022 Oct;13(5):1181-1193. [doi: 10.1055/a-1962-5583] [Medline: 36257602]
34. Khaksar SMS, Khosla R, Singaraju S, Slade B. Carer’s perception on social assistive technology acceptance and adoption: moderating effects of perceived risks. Behav Inf Technol 2021 Mar 12;40(4):337-360. [doi: 10.1080/0144929X.2019.1690046]
35. Rosenstock IM. Historical origins of the Health Belief Model. Health Educ Monogr 1974 Dec;2(4):328-335. [doi: 10.1177/109019817400200403]
36. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research: Addison-Wesley Publishing Company; 1975.
37. Rogers EM. Diffusion of Innovations, 5th edition: Free Press; 2003.
38. Silverstone R, Haddon L. Design and the domestication of information and communication technologies: technical change and everyday life. In: Mansell R, Silverton R, editors. Communication by Design: The Politics of Information and Communication Technologies: Oxford University Press; 1996:44-74. [doi: 10.1093/oso/9780198289418.003.0003]
39. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
40. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
41. Bsharat B, Al-Duhoun A, Ghanouni P. The acceptance and attitudes towards using assistive technology for people with stroke in Jordan: caregivers’ perspectives. Assist Technol 2024 Jan 2;36(1):40-50. [doi: 10.1080/10400435.2023.2202723] [Medline: 37083581]
42. Thanthrige A, Lu B, Sako Z, Wickramasinghe N. Determinants of health care technology adoption using an integrated Unified Theory of Acceptance and Use of Technology and Task Technology Fit Model: systematic review and meta-analysis. J Med Internet Res 2025 Dec 30;27:e64524. [doi: 10.2196/64524] [Medline: 41468595]

XSL•FO
RenderX

43. Nadal C, Sas C, Doherty G. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. J Med Internet Res 2020 Jul 6;22(7):e17256. [doi: 10.2196/17256] [Medline: 32628122]
44. Smole-Orehek K, Hvalič-Touzery S, Petrovčič A, Dolničar V, Debevc M, Kožuh I. Psychological outcomes of eCare technologies use for informal carers: a scoping study. Gerontechnology 2019 Apr 15;18(1):15-28. [doi: 10.4017/gt.2019.18.1.002.00]
45. Lebar L, Nagode M, Hvalič-Touzery S, Dolničar V, Natek I. The psychosocial impacts of e-care technology use for long-term care recipients and informal carers. Int J Integr Care 2022;22(S2):72. [doi: 10.5334/ijic.ICIC21299]
46. Lorca-Cabrera J, Grau C, Martí-Arques R, Raigal-Aran L, Falcó-Pegueroles A, Albacar-Riobóo N. Effectiveness of health web-based and mobile app-based interventions designed to improve informal caregiver's well-being and quality of life: a systematic review. Int J Med Inform 2020 Feb;134:104003. [doi: 10.1016/j.ijmedinf.2019.104003] [Medline: 31790857]
47. Smole-Orehek K. Psychosocial outcomes of eCare services use for employed informal carers of older people [Dissertation] [Article in Slovenian]. : Faculty of Social Sciences, University of Ljubljana; 2022 URL: https://repozitorij.uni-lj.si/Dokument.php?id=165224&lang=slv [accessed 2026-09-18]
48. Vizzotto ADB, Oliveira AM, Elkis H, Cordeiro Q, Buchain PC. Psychosocial characteristics. In: Gellman M, editor. Encyclopedia of Behavioral Medicine: Springer; 2020:1-2. [doi: 10.1007/978-1-4614-6439-6_918-2]
49. Martikainen P, Bartley M, Lahelma E. Psychosocial determinants of health in social epidemiology. Int J Epidemiol 2002 Dec;31(6):1091-1093. [doi: 10.1093/ije/31.6.1091] [Medline: 12540696]
50. Jutai J, Day H. Psychosocial Impact of Assistive Devices Scale (PIADS). Technol Disabil 2002;14(3):107-111. [doi: 10.3233/TAD-2002-14305]
51. Abda A, Bolduc ME, Tsimicalis A, Rennick J, Vatcher D, Brossard-Racine M. Psychosocial outcomes of children and adolescents with severe congenital heart defect: a systematic review and meta-analysis. J Pediatr Psychol 2019 May 1;44(4):463-477. [doi: 10.1093/jpepsy/jsy085] [Medline: 30452652]
52. Davis FD, Bagozzi RP, Warshaw PR. Extrinsic and intrinsic motivation to use computers in the workplace. J Appl Soc Psychol 1992 Jul;22(14):1111-1132. [doi: 10.1111/j.1559-1816.1992.tb00945.x]
53. Kim J, Park HA. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 1;14(5):e133. [doi: 10.2196/jmir.2143] [Medline: 23026508]
54. Werner P, Karnieli E. A model of the willingness to use telemedicine for routine and specialized care. J Telemed Telecare 2003;9(5):264-272. [doi: 10.1258/135763303769211274] [Medline: 14599329]
55. Jaschinski C, Ben Allouch S. Listening to the ones who care: exploring the perceptions of informal caregivers towards ambient assisted living applications. J Ambient Intell Human Comput 2019 Feb;10(2):761-778. [doi: 10.1007/s12652-018-0856-6]
56. Boyle LD, Husebo BS, Vislapuu M. Promotors and barriers to the implementation and adoption of assistive technology and telecare for people with dementia and their caregivers: a systematic review of the literature. BMC Health Serv Res 2022 Dec 23;22(1):1573. [doi: 10.1186/s12913-022-08968-2] [Medline: 36550456]
57. Wisniewski P, Linton C, Chokshi A, Perlingieri B, Gurupur V, Gabriel M. We have built it, but they have not come: examining the adoption and use of assistive technologies for informal family caregivers. In: Ahram TZ, Falcão C, editors. Advances in Usability, User Experience and Assistive Technology: Springer; 2019:824-836. [doi: 10.1007/978-3-319-94947-5_81]
58. Di Gessa G, Deindl C. Determinants of trajectories of informal caregiving in later life: evidence from England. Eur J Ageing 2024 Aug 31;21(1):24. [doi: 10.1007/s10433-024-00818-w] [Medline: 39215804]
59. Tur-Sinai A, Teti A, Rommel A, Hlebec V, Lamura G. How many older informal caregivers are there in Europe? Comparison of estimates of their prevalence from three European surveys. Int J Environ Res Public Health 2020 Dec 19;17(24):9531. [doi: 10.3390/ijerph17249531] [Medline: 33352669]
60. European quality of life survey 2016. : Eurofound; 2018 Jan 23 URL: https://www.eurofound.europa.eu/en/publications/all/european-quality-life-survey-2016 [accessed 2026-09-18]
61. Nagode M, Srakar A. Neformalni oskrbovalci: kdo izvaja neformalno oskrbo, v kolikšnem obsegu in za koga. In: Majcen B, editor. Značilnosti Starejšega Prebivalstva v Sloveniji – Prvi Rezultati Raziskave SHARE [Book in Slovenian]: Inštitut za ekonomska raziskovanja (Institute for Economic Research), Ljubljana; 2015:232-243 URL: http://www.share-slovenija.si/files/documents/prvi_rezultati_slovenija/Publikacija_IER_23.pdf [accessed 2026-09-18]
62. Callegaro M, Baker R, Bethlehem J, Göritz AS, Krosnick JA, Lavrakas PJ. Online panel research. In: Online Panel Research: Data Quality Perspective, A: John Wiley & Sons; 2014:1-22. [doi: 10.1002/9781118763520]
63. 37 questions to help buyers of online samples. : ESOMAR; 2023 URL: https://esocorpwebsitestg.blob.core.windows.net/strapi-uploads/uploads/37_questions_updated_version_c4f6e46f0e.pdf [accessed 2026-09-18]
64. Hafner-Fink M, Broder Ž, Doušak M, Falle Zorman R, Gerdina O, Jagodic A, et al. Slovensko javno mnenje 2021/1–Poročilo o izvedbi raziskave in sumarni pregled rezultatov [Report in Slovenian]. : Center za raziskovanje javnega mnenja in množičnih komunikacij, Univerza v Ljubljani, Fakulteta za družbene vede; 2021 URL: https://www.cjm.si/wp-content/uploads/2022/10/SUM-SJM21_1_L.pdf [accessed 2026-09-18]

XSL•FO
RenderX

65. Dolničar V, Berzelak N, Hvalič-Touzery S. Dataset for: factors impacting intention to use smart technology enabled care services among family carers of older people in the context of long-distance care (J5 - 1785) [dataset]. PsychArchives. 2022. URL: https://doi.org/10.23668/psycharchives.8240 [accessed 2026-09-19]
66. van Exel NJA, Scholte op Reimer WJM, Brouwer WBF, van den Berg B, Koopmanschap MA, van den Bos GAM. Instruments for assessing the burden of informal caregiving for stroke patients in clinical practice: a comparison of CSI, CRA, SCQ and self-rated burden. Clin Rehabil 2004 Mar;18(2):203-214. [doi: 10.1191/0269215504cr723oa] [Medline: 15053130]
67. e-Connected Family Caregiver: Bringing Caregiving into the 21st Century. : National Alliance for Caregiving; 2011 URL: https://www.caregiving.org/wp-content/uploads/2020/05/FINAL_eConnected_Family_Caregiver_Study_Jan-2011.pdf [accessed 2026-09-18]
68. Briggs SR, Cheek JM. The role of factor analysis in the development and evaluation of personality scales. J Pers 1986 Mar;54(1):106-148. [doi: 10.1111/j.1467-6494.1986.tb00391.x]
69. Cortina JM. What is coefficient alpha? An examination of theory and applications. J Appl Psychol 1993;78(1):98-104. [doi: 10.1037/0021-9010.78.1.98]
70. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis: Cengage Learning EMEA; 2019.
71. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
72. Dumouchel WH, Duncan GJ. Using sample survey weights in multiple regression analyses of stratified samples. J Am Stat Assoc 1983 Sep;78(383):535-543. [doi: 10.1080/01621459.1983.10478006]
73. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). : European Union; 2016 URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj [accessed 2026-09-22]
74. Vollmer Dahlke D, Lee S, Smith ML, Shubert T, Popovich S, Ory MG. Attitudes toward technology and use of fall alert wearables in caregiving: survey study. JMIR Aging 2021 Jan 27;4(1):e23381. [doi: 10.2196/23381] [Medline: 33502320]
75. Schulz R, Beach SR, Matthews JT, Courtney K, De Vito Dabbs A, Mecca LP. Caregivers' willingness to pay for technologies to support caregiving. Gerontologist 2016 Oct;56(5):817-829. [doi: 10.1093/geront/gnv033] [Medline: 26035899]
76. Lee S, Ory MG, Vollmer Dahlke D, Smith ML. Technology use among older adults and their caregivers: cross-sectional survey study. JMIR Aging 2024 May 1;7:e50759. [doi: 10.2196/50759] [Medline: 38717339]
77. Scott JA, Cepukenas E, Himes M, et al. Dementia caregivers' perspectives on technology's place in care practices: mixed methods survey. JMIR Form Res 2025 Nov 4;9:e69596. [doi: 10.2196/69596] [Medline: 41187277]
78. Chien SC, Chang YH, Yen CM, et al. Understanding and alleviating informal caregiver burden through the development and validation of a caregiver strain index-based model in Taiwan. BMC Geriatr 2024 Jun 26;24(1):558. [doi: 10.1186/s12877-024-05136-5] [Medline: 38918715]
79. Liu L, Miguel Cruz A, Ruptash T, Barnard S, Juzwishin D. Acceptance of global positioning system (GPS) technology among dementia clients and family caregivers. J Technol Hum Serv 2017 Apr 3;35(2):99-119. [doi: 10.1080/15228835.2016.1266724]
80. Felber NA, Lipworth W, Tian YJA, Roulet Schwab D, Wangmo T. Informing existing technology acceptance models: a qualitative study with older persons and caregivers. Eur J Ageing 2024 Mar 29;21(1):12. [doi: 10.1007/s10433-024-00801-5] [Medline: 38551677]
81. Burstein AA, DaDalt O, Kramer B, D'Ambrosio LA, Coughlin JF. Dementia caregivers and technology acceptance: interest outstrips awareness. Gerontechnology 2015;14(1):45-56. [doi: 10.4017/gt.2015.14.1.005.00]
82. Verloo H, Kampel T, Vidal N, Pereira F. Perceptions about technologies that help community-dwelling older adults remain at home: qualitative study. J Med Internet Res 2020 Jun 4;22(6):e17930. [doi: 10.2196/17930] [Medline: 32496197]
83. Thordardottir B, Malmgren Fänge A, Lethin C, Rodriguez Gatta D, Chiatti C. Acceptance and use of innovative assistive technologies among people with cognitive impairment and their caregivers: a systematic review. Biomed Res Int 2019;2019(1):9196729. [doi: 10.1155/2019/9196729] [Medline: 30956989]
84. Mitchell LL, Peterson CM, Rud SR, et al. "It's like a cyber-security blanket": the utility of remote activity monitoring in family dementia care. J Appl Gerontol 2020 Jan;39(1):86-98. [doi: 10.1177/0733464818760238] [Medline: 29504488]
85. Guisado-Fernández E, Giunti G, Mackey LM, Blake C, Caulfield BM. Factors influencing the adoption of smart health technologies for people with dementia and their informal caregivers: scoping review and design framework. JMIR Aging 2019 Apr 30;2(1):e12192. [doi: 10.2196/12192] [Medline: 31518262]
86. Sriram V, Jenkinson C, Peters M. Informal carers' experience of assistive technology use in dementia care at home: a systematic review. BMC Geriatr 2019 Jun 14;19(1):160. [doi: 10.1186/s12877-019-1169-0] [Medline: 31196003]
87. Semere W, Yank V, Lisha NE, Lindquist LA, Huang AJ. Older adults with overlapping caregiving responsibilities and care needs in a U.S. national community-based sample. J Am Geriatr Soc 2024 Jun;72(6):1824-1830. [doi: 10.1111/jgs.18794] [Medline: 38344822]

XSL•FO
RenderX

88. Hopps M, Iadeluca L, McDonald M, Makinson GT. The burden of family caregiving in the United States: work productivity, health care resource utilization, and mental health among employed adults. J Multidiscip Healthc 2017;10:437-444. [doi: 10.2147/JMDH.S135372] [Medline: 29255364]
89. Suchikova Y, Tsybuliak N, Teixeira da Silva JA, Nazarovets S. GAIDeT (Generative AI Delegation Taxonomy): a taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. Account Res 2026 Apr;33(3):2544331. [doi: 10.1080/08989621.2025.2544331] [Medline: 40781729]

Abbreviations

CFI: comparative fit index
ISCED: International Standard Classification of Education
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index
UTAUT: Unified Theory of Acceptance and Use of Technology
VIF: variance inflation factor

Edited by AF Abdulai; submitted 09.Oct.2025; peer-reviewed by GGD Alencar, S Chien-Chen; revised version received 05.Aug.2026; accepted 06.Aug.2026; published 20.Oct.2026.

Please cite as:
Dolničar V, Berzelak N, Taipale S, Hvalič-Touzery S
Integrating Perceived Psychosocial Outcomes Into the Assessment of Telecare Acceptability Among Slovenian Informal Caregivers of Older People: Cross-Sectional Survey Study
JMIR Aging 2026;9:e85534
URL: https://aging.jmir.org/2026/1/e85534
doi:10.2196/85534

© Vesna Dolničar, Nejc Berzelak, Sakari Taipale, Simona Hvalič-Touzery. Originally published in JMIR Aging (https://aging.jmir.org), 20.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Characterizing Digital Mindfulness Intervention Utilization and Weekly Assessments in Caregivers of Persons Living With Dementia: Secondary Analysis of a Randomized Controlled Trial

Michael P Williams[1,2], PhD; Darby M Simon[1,2], PhD; Morgan Seward[1], BA; Raquel G Tatar[3], MA, MSPH, PhD; Jennifer Huberty[4], PhD; Ana-Maria Vranceanu[1,2*], PhD; Evan Plys[1,2*], PhD

[1]Center for Health Outcomes and Interdisciplinary Research, Massachusetts General Hospital, 1 Bowdoin Square, Suite 100, Boston, MA, United States

[2]Department of Psychiatry, Harvard Medical School, Boston, MA, United States

[3]Healthy Minds Innovation, Madison, WI, United States

[4]Department of Medicine, University of Texas Health, San Antonio, TX, United States

[*]these authors contributed equally

Corresponding Author:
Michael P Williams, PhD
Center for Health Outcomes and Interdisciplinary Research, Massachusetts General Hospital, 1 Bowdoin Square, Suite 100, Boston, MA, United States

Abstract

Background: Caregivers of persons living with dementia are at increased risk of reporting high stress. Mindfulness-based interventions (MBIs) teach caregivers mindfulness skills and are effective at reducing stress. Digital MBIs are a feasible way to improve access to MBIs for caregivers of persons living with dementia. Yet, caregiver improvement with digital MBI utilization is less defined in the literature.

Objective: The goal of this secondary data analysis was to characterize weekly mindfulness and stress ratings among caregivers of persons living with dementia and to examine how digital MBI utilization impacted these ratings throughout 12 weeks of a feasibility trial.

Methods: Participants were eligible for this secondary analysis if they were randomized to the digital MBI condition (Healthy Minds Program for Caregivers [HMP-C], n=46) and completed weekly ratings over the 12-week trial. At baseline and at the end of each week of the trial, participants rated their mindfulness and stress in the past week from 0 to 10. Weekly HMP-C utilization was defined as the time spent using HMP-C in the week prior to the weekly ratings. Descriptive statistics and visualizations were used to characterize mindfulness and stress ratings. Generalized linear mixed models were used to estimate the effect of mindfulness on stress and the effect of HMP-C utilization on stress and mindfulness throughout the trial (α=.05).

Results: Baseline mindfulness and stress ratings were 4 (IQR 3-6, range 0 - 8) and 7 (IQR 6.5-8, range 4 - 10), respectively. Stress had the greatest decrease between baseline and week 3 (−2 points on average), whereas mindfulness had the greatest increase between baseline and week 4 (+2.5 points on average). There was a significant fixed effect of baseline mindfulness on baseline stress (β=.5, *P*<.001), with a significant interaction between mindfulness and study week (β=.05, *P*=.001), suggesting that this relationship was attenuated over time. There was variability in baseline stress (τ_{00}=2.04) and the relationship between mindfulness and stress ($\tau_{1\ 1}$ =0.08), with a high correlation (ρ_{01}=0.86), suggesting that those with high baseline stress benefited most from increases in mindfulness. For every 10 minutes of HMP-C utilization between baseline and week 1, mindfulness and stress ratings were 0.14 points higher (*P*<.001) and 0.14 points lower (*P*<.001), respectively. Despite a significant interaction between HMP-C utilization and study week in both models, the effect size was small (mindfulness: β=.02, *P*<.001; stress: β=.01, *P*=.013), suggesting that this relationship was sustained over time.

Conclusions: Mindfulness and stress changed mostly during the first 3 to 4 weeks of the trial. During this time of mindfulness skill acquisition, mindfulness and stress were most significantly negatively associated, especially among those with high baseline stress. The consistent relationship between HMP-C utilization and mindfulness and stress suggests that continued use of HMP-C may be helpful for skill maintenance.

(JMIR Aging 2026;9:e83618) doi:10.2196/83618

KEYWORDS

caregivers; mindfulness; stress; dementia; digital health; utilization

XSL•FO
RenderX

Introduction

There are approximately 53 million informal caregivers of persons living with dementia in the United States [1]. Caregivers of persons living with dementia are at higher risk for poor mental and physical health outcomes compared with age-matched noncaregiver peers [2]. Notably, caregivers of persons living with dementia experience high levels of stress [3], defined as a state of worry or mental tension caused by difficult situations [4]. Reducing stress provides an opportunity to improve caregiver well-being, and approximately half of caregivers report a desire for stress management interventions [5]. Together, there is a need and demand for stress reduction interventions among caregivers of persons living with dementia.

Mindfulness-based interventions (MBIs) are an effective nonpharmacological treatment for caregiver stress [6]. Mindfulness can be conceptualized as both a *trait*, or one's predisposition to be mindful in daily life, and a *state* that is achieved while practicing mindful meditation [7]. A framework proposed by Dahl et al [8] posits that state mindfulness meditation practices consist of skills that can be trained and developed (awareness, connection, insight, and purpose). MBIs primarily teach participants mindfulness skills as a coping strategy to manage everyday stressors. Mindfulness skills center on nonjudgmentally observing thoughts, feelings, and behaviors in the present moment [9]. The continual practice of state mindfulness through these skills increases trait mindfulness over time [7]. Both state and trait mindfulness are associated with lower perceived stress across multiple populations [10-12], including caregivers of persons living with dementia [13-15].

Although MBIs can reduce caregiver stress, caregivers often experience barriers to accessing these interventions in outpatient settings. For example, caregivers may experience limited access to respite care [16], limited free time [17], and resistance to leaving their care recipient due to guilt or fear [18], all of which interfere with help-seeking. Even when interventions are offered remotely, they are often synchronous and do not allow for additional flexibility when acute caregiving responsibilities arise. Digital MBIs provide an opportunity to overcome these barriers and increase access to MBIs for caregivers [19-22]. Digital MBIs are often delivered via mobile phone apps, are free or inexpensive, and can be accessed at any time, promoting flexibility in use, which may be particularly beneficial for those in the caregiving role. Digital MBIs are feasible and acceptable among caregivers of persons living with dementia in pilot trials [23-25] and align with a growing trend of technology acceptance among caregivers [26,27].

Although digital MBIs show promise for reducing stress among caregivers, the trajectory of skill acquisition and perceived stress reduction during participation in digital MBIs is rarely studied. According to Dahl and colleagues' [8] conceptualization of mindfulness as a series of skills to be learned, mindfulness is not expected to increase indefinitely, but rather to be learned (ie, skill acquisition), applied (ie, used as a state in stressful situations), and maintained (ie, developed as a trait across situations). A more granular understanding of the trajectory of skill acquisition and perceived stress reduction could help the scalability of interventions in a real-world context (eg, recommended usage, benchmarks for progress). This information may be particularly beneficial for self-directed interventions, such as digital MBIs, which lack regular monitoring of progress by a trained clinician.

This study aimed to elucidate the relationships among caregiver mindfulness, stress, and utilization of a digital MBI. This study is an exploratory secondary analysis of a feasibility pilot randomized controlled trial (RCT) of Healthy Minds Program for Caregivers (HMP-C), a digital MBI tailored to caregivers of persons living with dementia [25]. The primary pilot RCT reported findings primarily in the domains of feasibility, acceptability, treatment satisfaction, and adherence. The distinct aims of this study were 3-fold: (1) to characterize the trends of caregiver mindfulness and stress across the 12-week duration of the primary study, (2) to characterize the relationships between participants' weekly mindfulness and stress ratings, and (3) to explore how participants' utilization of the digital MBI is associated with their weekly mindfulness and stress ratings across the study duration.

Methods

Study Design

This study is an exploratory secondary data analysis of a single-blind pilot RCT (National Institutes of Health Stage Model, Stage 1b) of HMP-C, a digital MBI tailored for caregivers of persons living with dementia.

Ethical Considerations

This trial is registered with ClinicalTrials.gov (NCT05732038), and all study procedures were approved by the affiliated institutional review board (2022P001601) [25,28]. Informed consent was collected electronically. A member of the study staff was present to answer any questions. Participants received a total of US $50 for completing the postassessment and follow-up assessments. The privacy and confidentiality of all participant data were maintained.

Participants

The primary single-blind feasibility pilot RCT recruited informal adult caregivers of a known person living with Alzheimer disease or a related dementia reporting elevated levels of perceived stress. Recruitment was carried out through flyers and online advertisements. Interested caregivers referred themselves via an online link included in the recruitment materials. Here, they completed a brief eligibility questionnaire (verified by a phone call with study staff) based on the following inclusion criteria: (1) >18 years old, (2) English literacy, (3) self-identify as an informal caregiver, (4) Perceived Stress Scale-10 score of ≥6 [29,30], (5) willing to be randomized, and (6) care recipient rated >1 on the Functional Assessment Staging Tool [31]. Exclusion criteria included (1) any planned change in psychotropic pharmacological treatment for the duration of the study, (2) use of any consumer-based mindfulness meditation app for more than 60 minutes per month in the past 6 months, (3) current participation in a meditation program, (4) major illness anticipated to worsen dramatically or require surgery in the next 20 weeks, (5) active treatment for cancer (eg,

XSL•FO
RenderX

chemotherapy, radiation), (6) placement of the care recipient in a nursing home, (7) involvement in another clinical trial for caregivers, and (8) 4+ errors on the Short Portable Mental Status Questionnaire (ie, for caregivers older than 65 years of age) [32]. The primary feasibility RCT analyzed 90 participants across the HMP-C intervention (n=46) and a control (n=44) condition [25]. The present secondary analysis utilized participants in the intervention condition only, as the control condition did not teach mindfulness skills.

Intervention

HMP-C is a digital MBI tailored for caregivers of persons living with dementia, based on content from HMP, a publicly available mobile app [25,33]. HMP consists of written and guided listening content that aims to help users develop the following mindfulness skills: (1) awareness (eg, mindfulness, attention, self-awareness), (2) connection (eg, appreciation, kindness, compassion), (3) insight (eg, self-inquiry, self-knowledge, transcending the self), and (4) purpose (eg, purpose clarity, embodying values, finding meaning) [8,33]. Based on the Dahl et al [8] framework, these skills are taught to listeners through educational content based on the science of well-being and procedural guided mindfulness practices [25].

Primary Pilot RCT Procedures

Participants were directed to use the app for at least 10 minutes per day, or 70 minutes per week, for 12 weeks. Participants were reminded each week by text message to use their respective phone app for the prescribed amount of time. App utilization data were reviewed weekly by study staff, and participants who did not use at least 75% of their recommended amount in the past week were sent an additional standardized reminder message. Participants who did not complete their recommended utilization for 2 consecutive weeks were called by a study staff member to problem-solve potential barriers to adherence. Throughout the 12-week trial period, participants were emailed a REDCap (Research Electronic Data Capture) link to weekly single-item stress and mindfulness assessments. Participants who completed all assessments were offered a monetary incentive of up to US $50.

Measures

Weekly Mindfulness and Stress Assessments

At the end of each trial week, participants were asked to self-report perceived levels of mindfulness and stress from the prior week on single-item assessments, with scores ranging from 0 to 10. These assessments were also administered at baseline for a total of 13 weekly assessments (ie, baseline and 12 trial weeks). The stress assessment asked, “During the past week, how would you rate your stress on a scale from 0 to 10? (0=not at all; 10=extremely stressed).” The mindfulness assessment asked, “During the past week, how would you rate your mindfulness on a scale from 0 to 10? (0=not at all; 10=extremely mindful).” Mindfulness and stress among caregivers are multifaceted constructs, and several sources may contribute to stress among caregivers of persons living with dementia. The single-item measures utilized in this study represent broad, subjective ratings of mindfulness and stress regardless of perceived source.

Weekly HMP-C Utilization

Weekly HMP-C app utilization was defined as a continuous measure representing the sum of minutes of HMP-C utilization in the 7 days prior to each participant’s completion of their weekly mindfulness and stress assessments. These data were automatically captured through the HMP-C app.

Statistical Analysis

Overview

All statistical analyses were conducted using R (R Foundation for Statistical Computing) [34]. Four participants were removed from the data prior to analysis due to missing data. Specifically, 2 participants completed no weekly stress or mindfulness assessments, and 2 participants completed several weekly assessments in batches (eg, completing 2 or more weekly assessments at the same time). The final analytic dataset used in this study comprised the remaining 42 participants who received HMP-C. Descriptive statistics were used to describe participant characteristics for the 42 eligible participants included in this analysis. Statistical approaches relevant to each aim are detailed below. Consistent with study aims outlined below, we analyzed completers only. An α of .05 was used for all statistical tests.

Aim 1: Characterize Stress and Mindfulness Throughout the Study

Weekly stress and mindfulness assessments were characterized using descriptive statistics, including the median, minimum, maximum, and IQR. Further, we described change scores for the study variables throughout the trial.

Aim 2: Characterize the Relationship Between Weekly Mindfulness and Stress Ratings Throughout the Study Duration

Generalized linear mixed models (GLMMs) were used to assess the relationship between mindfulness and stress throughout the course of the 12-week trial period. Mindfulness was mean-centered to improve interpretation. This model also included an interaction term for mindfulness and time (study week) to assess how the effect of mindfulness on stress changed throughout the study duration. This model included a random intercept and a random slope for mindfulness to account for individual differences in baseline stress and the effect of mindfulness on stress. Random and fixed effects are reported for the effect of weekly mindfulness ratings on weekly stress ratings. Additionally, we visualized the Pearson correlation coefficient between weekly mindfulness and stress ratings. The final reported model was assessed for the following model assumptions: (1) linearity of fixed effects, (2) residual distribution, (3) outliers, (4) autocorrelation, (5) dispersion, (6) link function specification, and (7) zero inflation. No major deviations were detected for any of the aforementioned criteria. A power analysis was conducted using a simulation-based approach (1000 replications) and the observed effect size for the interaction term.

XSL•FO
RenderX

Aim 3: Explore How HMP-C Utilization Is Associated With Weekly Mindfulness and Stress Ratings Throughout the Study Duration

GLMMs were used to test the association between HMP-C utilization and weekly mindfulness and stress ratings. Both models included an interaction term between HMP-C utilization and time (study week) to assess how these effects changed throughout the study duration. Initially, models were specified to include a random slope and random intercept, as in Aim 2. However, both models in Aim 3 exhibited a singular fit, indicating that the variance of the random slope was approximately zero. This is common in datasets with small-to-medium sample sizes [35]. Due to the inability to estimate a random slope, we reported models that only included a random intercept, which aligns with guidelines for building GLMMs [35]. HMP-C utilization was mean-centered for improved interpretation, and study week was centered at week 1 (as opposed to baseline), as no participants utilized HMP-C prior to the baseline assessment. Since we are interested in how past-week HMP-C utilization is associated with stress and mindfulness, centering study week at week 1 allows for a week's worth of HMP-C utilization to transpire before estimating that association. Random and fixed effects are reported for the effect of HMP-C utilization on weekly mindfulness and stress ratings. The final reported model was assessed for the following model assumptions: (1) linearity of fixed effects, (2) residual distribution, (3) outliers, (4) autocorrelation, (5) dispersion, (6) link function specification, and (7) zero inflation. No major deviations were detected for any of the aforementioned criteria. Power analyses were conducted for both models using a simulation-based approach (1000 replications) and the observed effect size for the interaction term.

Results

Participants

Table 1 outlines participant characteristics at baseline for the 42 eligible caregivers included in this study. The average age was 55 (SD 14) years. Most participants were female by sex (n=39, 93%) and gender identity (n=38, 90%). Most participants were White (n=34, 81%). Most participants were married and living with their partner (n=28, 67%). Participants in this study used HMP-C for a median of 70 (IQR 41.7-102.2) minutes per week.

XSL•FO
RenderX

Table . Participant characteristics (N=42).

Characteristic	Value
Age (y), mean (SD)	55 (14)
Sex, n (%)	
Female	39 (93)
Male	3 (7.1)
Gender identity, n (%)	
Man	3 (7.1)
Woman	38 (90)
Prefer not to say	1 (2.4)
Race, n (%)	
Asian	1 (2.4)
Black/African American	4 (9.5)
More than one race	2 (4.8)
White	34 (81)
Choose not to answer	1 (2.4)
Ethnicity, n (%)	
Hispanic or Latino/Latina	3 (7.1)
Not Hispanic or Latino/Latina	38 (90)
Choose not to answer	1 (2.4)
Education, n (%)	
Completed high school or GED[a]	1 (2.4)
Some college	5 (12)
Associate degree	3 (7.1)
Completed 4 years of college	11 (26)
Graduate/professional degree	20 (48)
Trade/vocational training	2 (4.8)
Household income (US $), n (%)	
Less than 15,000	2 (4.8)
15,000 to 29,999	1 (2.4)
30,000 to 49,999	5 (12)
50,000 to 69,999	3 (7.1)
70,000 to 99,999	9 (21)
100,000 or more	17 (40)
Choose not to answer	5 (12)
Marital status, n (%)	
Married/partnered, living together	28 (67)
Separated/divorced	7 (17)
Single, never married	6 (14)
Choose not to answer	1 (2.4)

[a]GED: General Educational Development.

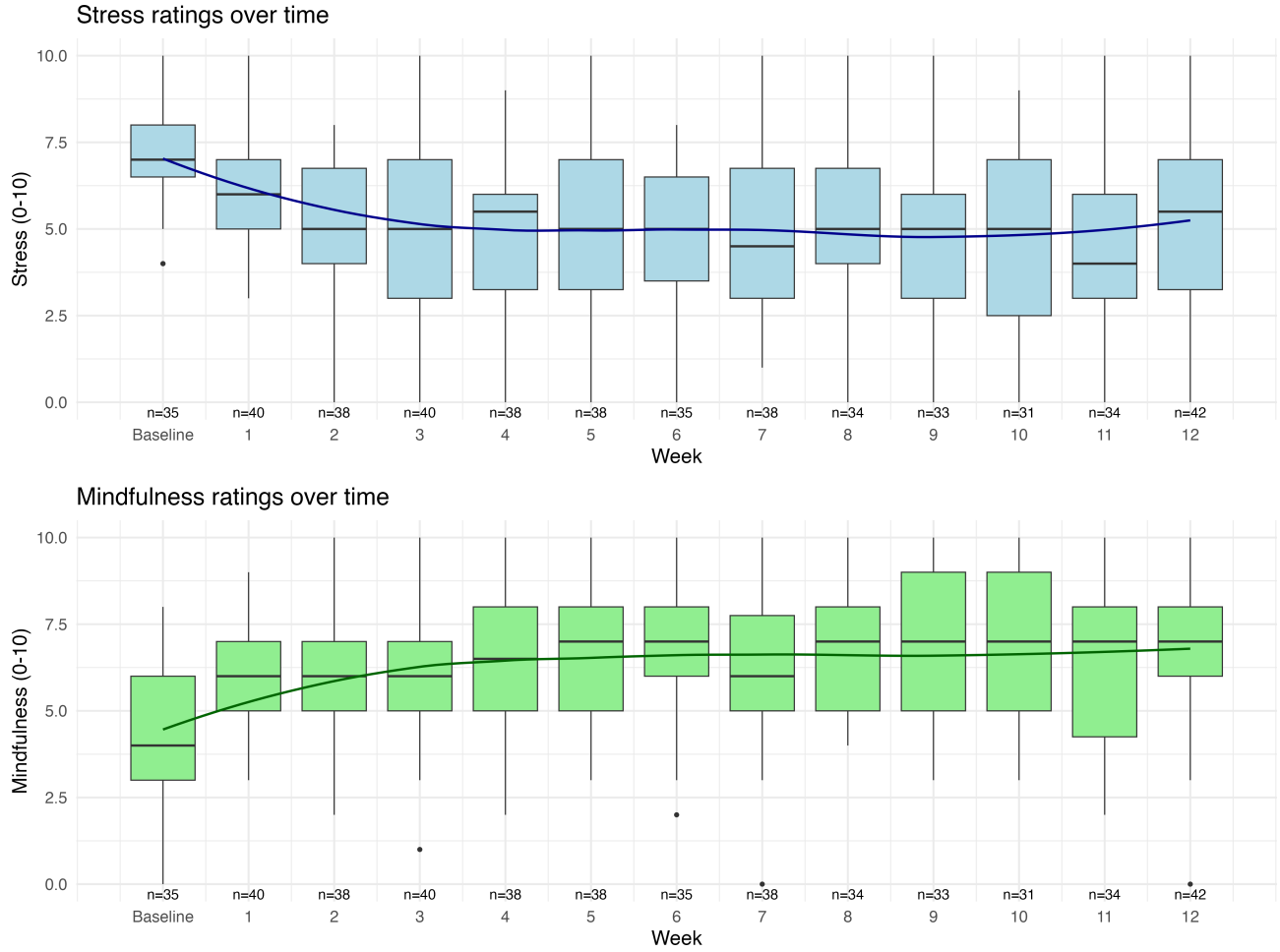
XSL•FO
RenderX

Aim 1

Median mindfulness and stress ratings at baseline were 4 (IQR 3-6, range 0 - 8) and 7 (IQR 6.5-8, range 4 - 10), respectively (Figure 1). By week 6 (ie, midway through the study), median mindfulness ratings increased to 7 (range 2 - 10) and stress ratings decreased to 5 (range 0 - 8). This represents a net increase in median mindfulness of 3 points and a net decrease in median stress of 2 points between baseline and week 6. Mindfulness and stress ratings remained similar throughout the remainder of the trial, with a net increase in median mindfulness and stress of 1 point between week 6 and 12 of the trial. Overall, there was an increase in median mindfulness of 3 points and a decrease in median stress of 1.5 points across the full trial (ie, baseline to week 12).

According to Figure 1, stress ratings appear to have the greatest decrease between baseline and week 3. By week 3, the median stress rating was 5 (IQR 4, range 0 - 10), which corresponds to a 2-point decrease from baseline. Mindfulness appears to experience the greatest increase between baseline and week 4. By week 4, the median mindfulness rating was 6.5 (IQR 3, range 2 - 10), which corresponds to a 2.5-point increase from baseline.

Figure 1. Weekly stress and mindfulness ratings among caregivers of persons living with dementia using the Healthy Minds Program for caregivers.



Aim 2

Figure 2 demonstrates the variability in the correlation between weekly mindfulness and stress ratings. Table 2 reports the results of the GLMM characterizing the relationship between mindfulness and stress. There was a significant fixed effect of mindfulness on stress, such that every additional point of mindfulness above the mean at baseline was correlated with a decrease in participant stress by an average of 0.5 points (P<.001). Similarly, participant stress was reduced by 0.06 (P=.007), on average, for each week throughout the duration of the study. However, the significant interaction between mindfulness and study week (β=.05, P=.001) showed that this association attenuated over time (ie, the stress-reducing effect of mindfulness was most potent during the early weeks of the trial and less potent during the later weeks in the trial). Random effects showed variability in baseline stress (τ_{00}=2.04) and the association between mindfulness and stress ($\tau_{1\ 1}$ =0.08) among participants. Notably, there was a high correlation (ρ_{01}=0.86) between intercepts and slopes, meaning that participants with higher baseline stress had a higher association between mindfulness and stress reduction over the course of the trial. Simulation results for the interaction term in this model show approximately 88.80% (95% CI 86.70-90.70) power to detect effects.

XSL•FO
RenderX

Figure 2. Pearson correlation between mindfulness and stress ratings over time among caregivers of persons living with dementia using Healthy Minds Program for Caregivers.

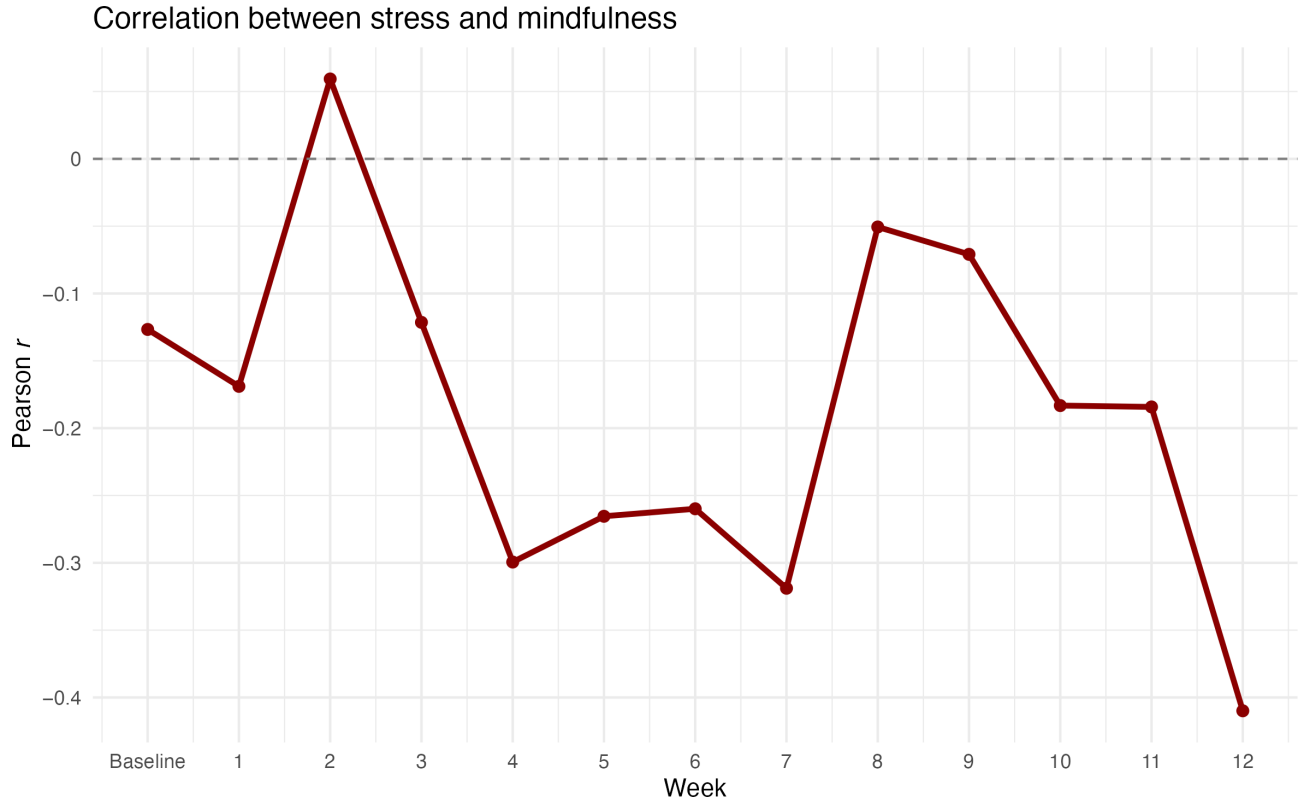


Table . Longitudinal effect of mindfulness ratings on stress ratings among caregivers of persons living with dementia using Healthy Minds Program for Caregivers (N=42).[a]

Predictors	Estimate (95% CI)	*P* value
(Intercept)	5.56 (5.03 to 6.09)	<.001
Mindfulness rating	−0.52 (−0.72 to −0.32)	<.001
Study week	−0.06 (−0.11 to −0.02)	.007
Mindfulness rating × study week	0.05 (0.02 to 0.08)	.001

[a]Random effects: σ^2=2.77; τ =2.04; $\tau_{1\ 1}$ =0.08; ρ_{1} =0.86; intraclass correlation coefficient=0.44.

Aim 3

Figure 3 details the variability in the correlation between HMP-C utilization and stress and mindfulness ratings each week through the 12-week trial. Correlations largely reflect an inverse relationship, where utilization generally correlates with higher mindfulness and lower stress. This is reflected in Table 3, where the fixed effects of utilization on mindfulness and stress were both significant. For every additional 10 minutes of HMP-C utilization between baseline and week 1, participants rated their mindfulness as 0.14 points higher on average (*P*<.001). Similarly, for every additional 10 minutes of HMP-C utilization, participants rated their stress as 0.14 points lower on average (*P*<.001). Both models demonstrated an effect of time, where for every week of utilization, participants' mindfulness increased by 0.10 (*P*<.001) and stress decreased by 0.07 (*P*<.001) on average. Both models had statistically significant interaction terms with study week. However, the effect sizes of these interaction terms were very small for both mindfulness (β=.02, *P*<.001) and stress (β=.01, *P*=.01). This suggests that while there is a statistically significant attenuation of the impact of HMP-C utilization on stress and mindfulness ratings, the size of that impact was very small, meaning that this correlation was mostly stable throughout the trial. Simulation results for the interaction term show approximately 99.50% (95% CI 98.84-99.84) power and 68.90% (95% CI 65.93-71.76) power in the mindfulness and stress models, respectively.

XSL•FO
RenderX

Figure 3. Pearson correlation between Healthy Minds Program for Caregivers (HMP-C) utilization and mindfulness or stress ratings over time among caregivers of persons living with dementia using HMP-C.

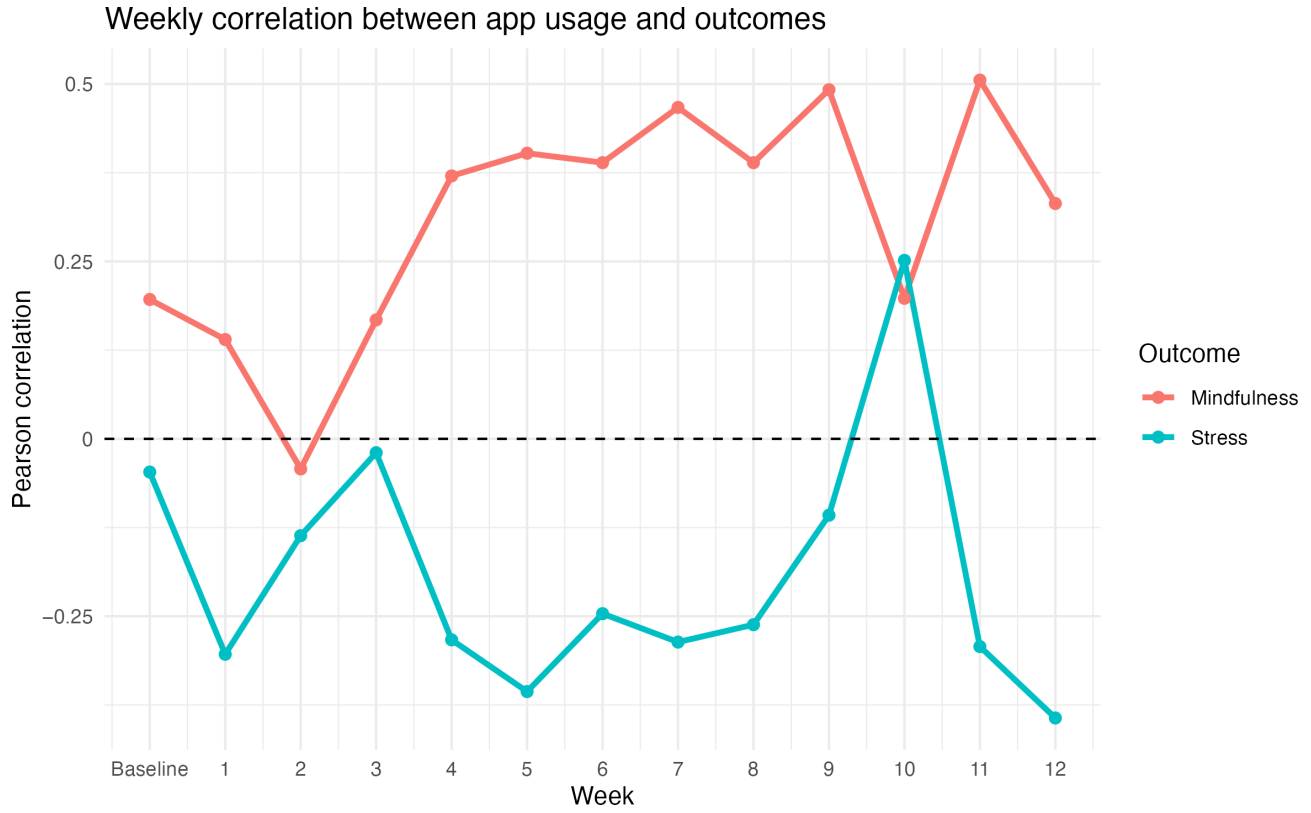


Table . Longitudinal effect of Healthy Minds Program for Caregivers (HMP-C) utilization on weekly mindfulness and stress ratings among caregivers of persons living with dementia using HMP-C (N=42).

Predictors	Mindfulness ratings[a]		Stress ratings[b]	
	Estimates (95% CI)	*P* value	Estimates (95% CI)	*P* value
(Intercept)	5.73 (5.22 to 6.23)	<.001	5.62 (5.11 to 6.13)	<.001
HMP-C utilization[c]	0.14 (0.10 to 0.18)	<.001	−0.14 (−0.20 to −0.08)	<.001
Study week[d]	0.10 (0.07 to 0.13)	<.001	−0.07 (−0.12 to −0.03)	.001
HMP-C utilization × study week	−0.02 (−0.02 to −0.01)	<.001	0.01 (0.00 to 0.02)	.01

[a]For the mindfulness ratings model, random effects were σ^2=1.51, τ =2.39, intraclass correlation coefficient=0.61, and N=42.

[b]For the stress ratings model, random effects were σ^2=3.00, τ =2.03, intraclass correlation coefficient=0.40, and N=42.

[c]HMP-C utilization was mean-centered and scaled to show the effect per 10 additional minutes.

[d]Study week was centered at the first week after baseline because HMP-C utilization captures past week utilization.

Discussion

Principal Results

This secondary data analysis aimed to characterize weekly mindfulness and stress ratings among caregivers of persons living with dementia and explore how digital MBI utilization was associated with those ratings throughout 12 weeks of a feasibility trial. We found that, on average, participants' mindfulness and stress ratings improved throughout the initial weeks of the study and then sustained that improvement throughout the remainder of the study duration. Further, we found that there was a significant association between mindfulness and stress ratings throughout the study duration among our sample of caregivers of persons living with dementia, such that in weeks when mindfulness was high, stress was low. However, this association became weaker as the study progressed, suggesting that perceived stress and mindfulness are less intertwined over time. Finally, HMP-C utilization was associated with changes in mindfulness and stress ratings, suggesting that continued utilization may be helpful for maintaining these improvements over time.

The sample of participants from this study in some ways reflects the broader population of caregivers of persons living with dementia and in other ways potentially deviates from the average caregiver of persons living with dementia. On the one hand, participants overwhelmingly identified as women. Women have historically undertaken the role of caregiving for persons living

XSL•FO
RenderX

with dementia, and this finding is consistent across geographic regions. Further, women often spend more time caregiving than their male caregiver counterparts and often experience a higher level of caregiver burden [36]. Conversely, the present sample of participants is largely White, has attained a graduate degree, and is earning more than US $100,000 per year. Past work has shown that approximately two-thirds of caregivers of persons living with dementia are White (compared to roughly 4 in 5 caregivers in the present sample). Further, past work has shown that the average caregiver of persons living with dementia earns US $50,000 or less per year. Additionally, the present sample represents a cohort of individuals positioned to have greater access to and literacy with technology due to their demographic characteristics (eg, race, income, education, and age) [37]. The complex intersectionality of underresourced identities and technological access and literacy can impact digital intervention engagement. Indeed, past reviews have shown that technological literacy is a significant barrier to implementing digital interventions for caregivers of persons living with dementia, despite their high health literacy [38-44].

Both mindfulness and stress ratings improved throughout the first 6 weeks of the study, with most of the improvement occurring between the first 3 to 4 weeks. From there, mindfulness and stress ratings leveled off throughout the remainder of the study. These findings suggest that participants may have developed some level of mastery in mindfulness skills within the first 4 weeks of HMP-C utilization, with more minor improvements occurring after week 4. Since Dahl et al [8] conceptualize mindfulness practice through a series of developable skills, it would be expected that once those skills are mastered, continued practice would lead to maintenance (ie, not a linear increase). This study reflects this trajectory, where caregivers improved mindfulness and stress early (weeks 1 - 4) and then stabilized by week 6. This may reflect acquisition of new skills followed by subsequent consolidation of those skills. It remains plausible that continued practice would yield additional benefit at longer follow-up. Future work should evaluate this by extending assessments over a longer period of time.

Improvements in weekly mindfulness ratings correlated with stress reductions throughout the study duration. However, there is context to this finding. First, there was variability in baseline stress and the association between mindfulness and stress across participants. Notably, there was a positive correlation between baseline stress and the effect of mindfulness on stress reductions, meaning those who were the most stressed at baseline benefited the most from increases in mindfulness. This suggests that mindfulness skills may have offered these participants a strategy to immediately manage stressful situations in their lives. However, the association between mindfulness and stress decreased over time, regardless of baseline stress levels. This trajectory in mindfulness and stress ratings reflects several previous findings, where an initial effect of mindfulness skills on the primary outcome is observed, followed by a period of attenuation [45-48]. These findings underscore the importance of considering both individual differences at baseline and the temporal dynamics of mindfulness practice. The early pronounced association for those with higher stress highlights mindfulness as a potentially powerful mechanism for alleviating acute distress (ie, state mindfulness for coping). At the same time, the subsequent attenuation of effects suggests that the role of mindfulness may shift from producing large, immediate reductions in stress to supporting the maintenance of stress reduction over time (ie, trait mindfulness). This pattern is consistent with theoretical models, such as Dahl et al [8], that view mindfulness as a set of skills that, once acquired, foster a sense of resiliency and stability, rather than indefinite linear improvement. Taken together, these results suggest that digital MBIs should emphasize intensive early engagement while also providing education and strategies to sustain and deepen benefits during later phases of practice. Future research should aim to assess the mediating role of mindfulness in stress reduction for caregivers of persons living with dementia as compared to an attention-placebo control condition in a fully powered efficacy trial.

HMP-C utilization was broadly associated with increases in mindfulness and reductions in stress ratings throughout the duration of the study. Despite a statistically significant interaction of HMP-C utilization and time in study weeks in both utilization models, the effect sizes of these interactions were very small. While there may be some minor attenuation of the association between HMP-C utilization on mindfulness and stress over time, the association was largely maintained throughout the study duration. This suggests that continued utilization of HMP-C may serve as an important tool for mindfulness skill maintenance, such that utilization may be protective against returning to baseline levels of mindfulness and stress. Therefore, continued use of the digital MBI would be recommended for achieving and sustaining improvements in mindfulness and stress among caregivers of persons living with dementia. These findings suggest several directions for future work. First, to examine if and to what extent a dose-response relationship exists between HMP-C utilization and mindfulness or stress, a future study could randomize participants to incrementally different “dosages” of HMP-C utilization (eg, 45, 60, 75, and 90 min/wk). Further, future research is needed to address whether mindfulness skills and their effect on stress can be maintained without continued practice or the same amount of continued practice. For example, future studies may choose to randomize participants to continued, reduced, or no HMP-C utilization after an initial period of 4 weeks. These would provide the rigorous groundwork needed to tune digital MBI utilization recommendations.

Limitations

There are several limitations to consider when interpreting the findings from this study. First, the participants in this study were a small sample from a stage 1b feasibility RCT. In many ways, the sample was not representative of caregivers of persons living with dementia broadly with respect to race, income, education, and potentially digital literacy. Due to this, the generalizability of these findings to the broader population of caregivers may be limited. Further, analytically, we were not powered to control for possible sources of confounding, such as digital and health-related literacy, which may have led participants to be more adept at improving from mindfulness

XSL•FO
RenderX

skill practice. Second, this study does not include a control or comparator arm in the analysis. This precludes making causal statements about the efficacy of HMP-C or digital MBIs with respect to changes in mindfulness and stress. Possible sources of confounding due to lack of randomization include seasonal changes, regression toward the mean, or simply random fluctuations. Third, single-item weekly ratings are less explanatory than more comprehensive measures. For example, from this study, we cannot ascertain exactly what stressors participants were experiencing or were being mitigated. However, single-item measures have been shown to be comparable to full-scale measures as broad measures of stress or mindfulness, suggesting that the overall trends observed are likely valid [49-51]. Fourth, week-over-week changes in mindfulness and stress may be reflective of the Hawthorne effect or assessment reactivity, as opposed to true changes in the measured constructs.

Conclusions

This study is an exploratory secondary analysis of a feasibility pilot RCT of HMP-C, a digital MBI tailored for caregivers of persons living with dementia. We found that mindfulness and stress ratings improved throughout the first 3 to 4 weeks of the trial before leveling off throughout the remainder of the trial. Particularly, in these first 4 to 6 weeks of skill acquisition, mindfulness and stress were significantly negatively related to one another. This association was higher for those with high baseline stress. Together, these findings highlight the potential benefit of digital MBIs for stress reduction among caregivers of persons living with dementia and suggest that mindfulness skills may be acquired within the first 4 to 6 weeks of the intervention and then maintained over a longer period of time. Future research is needed to move closer to precise dosing recommendations for mobile apps with this population.

Acknowledgments

The authors would like to collectively acknowledge and thank the study team and participants for their contributions to this work.

Funding

This work was supported by the National Institutes of Health (R21 AG075187 and K24AT011760 to AMV; K23AG078410 to EP; and T32 AG081327 to MPW and DMS).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

RGT is the chief science officer for Humin (formerly Healthy Minds Innovations), the nonprofit organization that deploys the Healthy Minds Program used in this study. All other authors declare no conflicts of interest.

References

1. Caregiving in the U.S. 2020. : American Association of Retired Persons (AARP) and National Alliance for Caregiving (NAC); 2020 Mar URL: https://www.aarp.org/content/dam/aarp/ppi/2020/05/full-report-caregiving-in-the-united-states.doi.10.26419-2Fppi.00103.001.pdf [accessed 2026-06-09]
2. Janson P, Willeke K, Zaibert L, et al. Mortality, morbidity and health-related outcomes in informal caregivers compared to non-caregivers: a systematic review. Int J Environ Res Public Health 2022 May 11;19(10):5864. [doi: 10.3390/ijerph19105864] [Medline: 35627399]
3. Gilhooly KJ, Gilhooly MLM, Sullivan MP, et al. A meta-review of stress, coping and interventions in dementia and dementia caregiving. BMC Geriatr 2016 May 18;16(1):106. [doi: 10.1186/s12877-016-0280-8] [Medline: 27193287]
4. Stress. World Health Organization. URL: https://www.who.int/news-room/questions-and-answers/item/stress [accessed 2025-08-19]
5. Caregiving in the U.S. 2015. : American Association of Retired Persons (AARP) and National Alliance for Caregiving (NAC); 2015 URL: https://www.aarp.org/content/dam/aarp/ppi/2015/caregiving-in-the-united-states-2015-report-revised.pdf [accessed 2025-08-09]
6. Li G, Yuan H, Zhang W. The effects of mindfulness-based stress reduction for family caregivers: systematic review. Arch Psychiatr Nurs 2016 Apr;30(2):292-299. [doi: 10.1016/j.apnu.2015.08.014] [Medline: 26992885]
7. Kiken LG, Garland EL, Bluth K, Palsson OS, Gaylord SA. From a state to a trait: trajectories of state mindfulness in meditation during intervention predict changes in trait mindfulness. Pers Individ Dif 2015 Jul 1;81:41-46. [doi: 10.1016/j.paid.2014.12.044] [Medline: 25914434]
8. Dahl CJ, Wilson-Mendenhall CD, Davidson RJ. The plasticity of well-being: a training-based framework for the cultivation of human flourishing. Proc Natl Acad Sci U S A 2020 Dec 22;117(51):32197-32206. [doi: 10.1073/pnas.2014859117] [Medline: 33288719]
9. Siegel RD, Germer CK, Olendzki A. Mindfulness: what is it? Where did it come from? In: Didonna F, editor. Clinical Handbook of Mindfulness: Springer; 2009:17-35. [doi: 10.1007/978-0-387-09593-6_2]

XSL•FO
RenderX

10. Sun S, Goldberg SB, Loucks EB, Brewer JA. Mindfulness-based interventions among people of color: a systematic review and meta-analysis. Psychother Res 2022 Mar;32(3):277-290. [doi: 10.1080/10503307.2021.1937369] [Medline: 34098859]
11. Khoury B, Sharma M, Rush SE, Fournier C. Mindfulness-based stress reduction for healthy individuals: a meta-analysis. J Psychosom Res 2015 Jun;78(6):519-528. [doi: 10.1016/j.jpsychores.2015.03.009] [Medline: 25818837]
12. Kriakous SA, Elliott KA, Lamers C, Owen R. The effectiveness of mindfulness-based stress reduction on the psychological functioning of healthcare professionals: a systematic review. Mindfulness (N Y) 2021;12(1):1-28. [doi: 10.1007/s12671-020-01500-9] [Medline: 32989406]
13. Oken BS, Fonareva I, Haas M, et al. Pilot controlled trial of mindfulness meditation and education for dementia caregivers. J Altern Complement Med 2010 Oct;16(10):1031-1038. [doi: 10.1089/acm.2009.0733] [Medline: 20929380]
14. Hou RJ, Wong SYS, Yip BHK, et al. The effects of mindfulness-based stress reduction program on the mental health of family caregivers: a randomized controlled trial. Psychother Psychosom 2014;83(1):45-53. [doi: 10.1159/000353278] [Medline: 24281411]
15. Whitebird RR, Kreitzer M, Crain AL, Lewis BA, Hanson LR, Enstad CJ. Mindfulness-based stress reduction for family caregivers: a randomized controlled trial. Gerontologist 2013 Aug;53(4):676-686. [doi: 10.1093/geront/gns126] [Medline: 23070934]
16. Lee Y, Choi W, Park MS. Respite service use among dementia and nondementia caregivers: findings from the National Caregiving in the U.S. 2015 survey. J Appl Gerontol 2022 Jun;41(6):1557-1567. [doi: 10.1177/07334648221075620] [Medline: 35303780]
17. Rokicka M, Zajkowska O. Informal elderly caregiving and time spent on leisure: evidence from time use survey. Ageing Int 2020 Dec;45(4):393-410. [doi: 10.1007/s12126-020-09396-5]
18. Leocadie MC, Roy MH, Rothan-Tondeur M. Barriers and enablers in the use of respite interventions by caregivers of people with dementia: an integrative review. Arch Public Health 2018;76(1):72. [doi: 10.1186/s13690-018-0316-y] [Medline: 30479766]
19. Chong SOK, Pedron S, Abdelmalak N, Laxy M, Stephan AJ. An umbrella review of effectiveness and efficacy trials for app-based health interventions. NPJ Digit Med 2023 Dec 16;6(1):233. [doi: 10.1038/s41746-023-00981-x] [Medline: 38104213]
20. Linardon J, Messer M, Goldberg SB, Fuller-Tyszkiewicz M. The efficacy of mindfulness apps on symptoms of depression and anxiety: an updated meta-analysis of randomized controlled trials. Clin Psychol Rev 2024 Feb;107:102370. [doi: 10.1016/j.cpr.2023.102370] [Medline: 38056219]
21. Macrynikola N, Mir Z, Gopal T, et al. The impact of mindfulness apps on psychological processes of change: a systematic review. Npj Ment Health Res 2024 Mar 18;3(1):14. [doi: 10.1038/s44184-023-00048-5] [Medline: 38609511]
22. Plaza I, Demarzo MMP, Herrera-Mercadal P, García-Campayo J. Mindfulness-based mobile applications: literature review and analysis of current features. JMIR mHealth uHealth 2013 Nov 1;1(2):e24. [doi: 10.2196/mhealth.2733] [Medline: 25099314]
23. Goodridge D, Reis N, Neiser J, et al. An app-based mindfulness-based self-compassion program to support caregivers of people with dementia: participatory feasibility study. JMIR Aging 2021 Nov 26;4(4):e28652. [doi: 10.2196/28652] [Medline: 34842530]
24. Kozlov E, McDarby M, Pagano I, Llaneza D, Owen J, Duberstein P. The feasibility, acceptability, and preliminary efficacy of an mHealth mindfulness therapy for caregivers of adults with cognitive impairment. Aging Ment Health 2022 Oct;26(10):1963-1970. [doi: 10.1080/13607863.2021.1963949] [Medline: 34428993]
25. Plys E, Woodworth E, Seward M, et al. Feasibility and mechanism engagement of a mindfulness mobile app for caregivers: findings from an NIH stage 1b trial. Innov Aging 2024 Dec 31;8(Suppl 1):105-105. [doi: 10.1093/geroni/igae098.0339]
26. Faverio M. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. 2022 Jan 13. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2025-08-19]
27. Health and caregiving among the 50+: ownership, use, and interest in mobile technology. : American Association of Retired Persons (AARP); 2011 Jan URL: https://assets.aarp.org/rgcenter/general/health-caregiving-mobile-technology.pdf [accessed 2026-06-09]
28. Woodworth EC, Briskin EA, Plys E, et al. Mindfulness-based app to reduce stress in caregivers of persons with Alzheimer disease and related dementias: protocol for a single-blind feasibility proof-of-concept randomized controlled trial. JMIR Res Protoc 2023 Oct 13;12(1):e50108. [doi: 10.2196/50108] [Medline: 37831492]
29. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
30. Eskildsen A, Dalgaard VL, Nielsen KJ, et al. Cross-cultural adaptation and validation of the Danish consensus version of the 10-item Perceived Stress Scale. Scand J Work Environ Health 2015 Sep 1;41(5):486-490. [doi: 10.5271/sjweh.3510] [Medline: 26111225]
31. Sclan SG, Reisberg B. Functional Assessment Staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality. Int Psychogeriatr 1992;4 Suppl 1(3):55-69. [doi: 10.1017/s1041610292001157] [Medline: 1504288]

XSL•FO
RenderX

32. Roccaforte WH, Burke WJ, Bayer BL, Wengel SP. Reliability and validity of the Short Portable Mental Status Questionnaire administered by telephone. J Geriatr Psychiatry Neurol 1994;7(1):33-38. [Medline: 8192828]
33. The Free Healthy Minds Program. Humin. URL: https://hminnovations.org/meditation-app [accessed 2025-02-10]
34. R: The R Project for Statistical Computing. URL: https://www.r-project.org/ [accessed 2025-02-10]
35. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015 Oct 7;67:1-48. [doi: 10.18637/jss.v067.i01]
36. 2024 Alzheimer's disease facts and figures. Alzheimers Dement 2024 May;20(5):3708-3821. [doi: 10.1002/alz.13809] [Medline: 38689398]
37. Yang R, Gao S, Jiang Y. Digital divide as a determinant of health in the U.S. older adults: prevalence, trends, and risk factors. BMC Geriatr 2024 Dec 21;24(1):1027. [doi: 10.1186/s12877-024-05612-y] [Medline: 39709341]
38. Madeira R, Esteves D, Pinto N, Vercelli A, Pato MV. Acceptance factors and barriers to the implementation of digital interventions in older people with dementia and/or their caregivers: an umbrella review. J Clin Med 2025 Nov 10;14(22):7974. [doi: 10.3390/jcm14227974] [Medline: 41303010]
39. Efthymiou A, Middleton N, Charalambous A, Papastavrou E. Adapting the eHealth Literacy Scale for carers of people with chronic diseases (eHeals-Carer) in a sample of Greek and Cypriot carers of people with dementia: reliability and validation study. J Med Internet Res 2019 Nov 28;21(11):e12504. [doi: 10.2196/12504] [Medline: 31778120]
40. Oh KM, Hong SR, Beran K, Song Y, Lee JA. Digital health literacy and self-efficacy in using digital health resources among caregivers of individuals with Alzheimer's disease and related dementias. Sage Open Aging 2025;11:30495334251398089. [doi: 10.1177/30495334251398089] [Medline: 41356564]
41. Efthymiou A, Middleton N, Charalambous A, Papastavrou E. Health literacy and eHealth literacy and their association with other caring concepts among carers of people with dementia: a descriptive correlational study. Health Soc Care Community 2022 May;30(3):1109-1119. [doi: 10.1111/hsc.13341] [Medline: 33956368]
42. Yuen EYN, Knight T, Ricciardelli LA, Burney S. Health literacy of caregivers of adult care recipients: a systematic scoping review. Health Soc Care Community 2018 Mar;26(2):e191-e206. [doi: 10.1111/hsc.12368] [Medline: 27426731]
43. Cho H, Cho E, You SB, Sefcik JS, Hodgson NA, Demiris G. Understanding digital literacy of persons with dementia and their caregivers: a scoping review and an evolutionary concept analysis of empirical studies. J Appl Gerontol 2026 Apr;45(4):742-766. [doi: 10.1177/07334648251348703] [Medline: 40537959]
44. Kim H. Understanding internet use among dementia caregivers: results of secondary data analysis using the US Caregiver Survey data. Interact J Med Res 2015 Feb 23;4(1):e1. [doi: 10.2196/ijmr.3127] [Medline: 25707033]
45. Ribeiro L, Atchley RM, Oken BS. Adherence to practice of mindfulness in novice meditators: practices chosen, amount of time practiced, and long-term effects following a mindfulness-based intervention. Mindfulness (N Y) 2018 Apr;9(2):401-411. [doi: 10.1007/s12671-017-0781-3] [Medline: 30881517]
46. Fiocco AJ, Hytman L. Preliminary evidence for lasting benefits one year following mindfulness training among older family caregivers of persons with neurodegenerative disease. Mindfulness 2022 Dec;13:2968-2978. [doi: 10.1007/s12671-022-02005-3]
47. Strohmaier S. The relationship between doses of mindfulness-based programs and depression, anxiety, stress, and mindfulness: a dose-response meta-regression of randomized controlled trials. Mindfulness 2020 Jun;11(6):1315-1335. [doi: 10.1007/s12671-020-01319-4]
48. Hsieh CC, Lin ZZ, Ho CC, et al. The short- and long-term causal relationships between self-compassion, trait mindfulness, caregiver stress, and depressive symptoms in family caregivers of patients with lung cancer. Mindfulness (N Y) 2021;12(7):1812-1821. [doi: 10.1007/s12671-021-01642-4] [Medline: 33968270]
49. Moore RC, Depp CA, Wetherell JL, Lenze EJ. Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. J Psychiatr Res 2016 Apr;75:116-123. [doi: 10.1016/j.jpsychires.2016.01.011] [Medline: 26851494]
50. Yang YS, Ryu GW, Choi M. Methodological strategies for ecological momentary assessment to evaluate mood and stress in adult patients using mobile phones: systematic review. JMIR mHealth uHealth 2019 Apr 1;7(4):e11215. [doi: 10.2196/11215] [Medline: 30932866]
51. Murray AL, Xiao Z, Zhu X, et al. Psychometric evaluation of an adapted version of the perceived stress scale for ecological momentary assessment research. Stress Health 2023 Oct;39(4):841-853. [doi: 10.1002/smi.3229] [Medline: 36697362]

Abbreviations

GLMM: generalized linear mixed model
HMP-C: Healthy Minds Program for Caregivers
MBI: mindfulness-based intervention
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

Edited by F Puga; submitted 05.Sep.2025; peer-reviewed by C Carrasco, S Chi Lik Au; revised version received 12.Mar.2026; accepted 12.Apr.2026; published 21.Jul.2026.

Please cite as:
Williams MP, Simon DM, Seward M, Tatar RG, Huberty J, Vranceanu AM, Plys E
Characterizing Digital Mindfulness Intervention Utilization and Weekly Assessments in Caregivers of Persons Living With Dementia: Secondary Analysis of a Randomized Controlled Trial
JMIR Aging 2026;9:e83618
URL: https://aging.jmir.org/2026/1/e83618
doi:10.2196/83618

© Michael P Williams, Darby M Simon, Morgan Seward, Raquel G Tatar, Jennifer Huberty, Ana-Maria Vranceanu, Evan Plys. Originally published in JMIR Aging (https://aging.jmir.org), 21.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exergaming Interventions for Preventing Falls and Injurious Falls in Older People: Systematic Review and Meta-Analysis of Randomized Controlled Trials

Charlotte Eost-Telling[1,2,3], PhD; Lisa McGarrigle[2,3,4], PhD; Chunhu Shi[1,2], PhD; Annemarie Money[1,2], PhD; Yang Yang[1,2], PhD; Kimberly Lazo Green[2,3], PhD; Saima Ahmed[1,2], PhD; Rachel Christie[1,2], MSc; Abodunrin Aminu[2,3], PhD; Kim Delbaere[5,6], PhD; Eling D de Bruin[7,8,9], PhD; Emma Stanmore[1,2,10], PhD; Chris Todd[1,2,3,10], PhD

[1]Division of Nursing Midwifery and Social Work, Faculty of Biology, Medicine and Health, National Institute for Health and Care Research (NIHR) Applied Research Collaboration Greater Manchester, University of Manchester, Manchester, England, United Kingdom

[10]Manchester University NHS Foundation Trust, Manchester, England, United Kingdom

[2]Manchester Academic Health Science Centre, Manchester, England, United Kingdom

[3]Division of Nursing Midwifery and Social Work, Faculty of Biology, Medicine and Health, National Institute for Health and Care Research, Policy Research Unit in Older People and Frailty / Healthy Ageing, University of Manchester, Manchester, England, United Kingdom

[4]School of Nursing and Healthcare, Atlantic Technological University, Castlebar, County Mayo, Ireland

[5]Falls, Balance and Injury Research Centre, Neuroscience Research Australia, Randwick, New South Wales, Australia

[6]School of Health Sciences, University of New South Wales, Kensington, New South Wales, Australia

[7]Department of Health, University of Applied Sciences St. Gallen, St. Gallen, Saint Gallen, Switzerland

[8]Department of Health Sciences and Technology, Motor Control and Learning Group - Institute of Human Movement Sciences and Spor, ETH Zurich, Zurich, Zurich, Switzerland

[9]Department of Neurobiology, Care Sciences, and Society, Division of Physiotherapy, Karolinska Institutet, Huddinge, Stockholm, Sweden

Corresponding Author:
Charlotte Eost-Telling, PhD
Division of Nursing Midwifery and Social Work, Faculty of Biology, Medicine and Health, National Institute for Health and Care Research (NIHR) Applied Research Collaboration Greater Manchester, University of Manchester, Manchester, England, United Kingdom

Abstract

Background: Exergames, which combine physical exercise with interactive gameplay, are increasingly being incorporated into fall prevention programs for older adults. Gamified elements, such as real-time feedback and progress tracking, may enhance motivation, engagement, and adherence. Although several systematic reviews have examined the effects of exergaming on balance and physical function, fewer have focused specifically on clinically meaningful outcomes, such as falls and injurious falls, or on indicators that may influence real-world adoption of exergames.

Objective: This study aimed to evaluate the effectiveness of exergaming interventions for preventing falls and injurious falls in people aged ≥60 years and to synthesize evidence on implementation-related outcomes, including adherence, acceptability, concerns about falling, quality of life, adverse events, and cost-effectiveness.

Methods: MEDLINE, Embase, CINAHL Plus, PsycINFO, and the Cochrane Central Register of Controlled Trials (CENTRAL) were searched from inception to February 2025 for randomized controlled trials evaluating exergaming interventions in older adult populations across all settings. Outcomes included fall rate, number of fallers and injurious falls, and implementation-related secondary outcomes. Risk of bias was assessed using RoB 2.0, and certainty of evidence was assessed using Grading of Recommendations Assessment, Development, and Evaluation (GRADE). Data were synthesized narratively and, where appropriate, pooled using meta-analysis.

Results: Nine studies (N=1385) met the inclusion criteria. Comparator-specific analyses suggested that exergaming may reduce fall rates compared with active intervention comparators, although the magnitude and certainty of effect varied, and substantial heterogeneity was present across analyses. Moderate-certainty evidence also suggested that exergames reduced the number of older adults experiencing one or more falls at 12-month follow-up compared with usual care (risk ratio 0.75, 95% CI 0.61 - 0.92). Evidence for injurious falls, quality of life, concerns about falling, adherence, acceptability, and cost-effectiveness was limited or inconsistent. When pooled across all control groups, exergaming interventions were associated with a lower overall fall rate than comparator interventions (incidence rate ratio 0.53, 95% CI 0.41 - 0.68), although substantial heterogeneity was present (I^2=76%).

XSL•FO
RenderX

Conclusions: Low- to moderate-certainty evidence suggests that exergames may reduce fall rates, particularly in comparisons with active intervention control groups, and may reduce the number of fallers compared with usual care. These findings indicate that exergaming may offer a useful adjunct to established fall prevention strategies for older adults, particularly where sustained engagement with conventional exercise is challenging. However, substantial heterogeneity, modest sample sizes, and limited long-term follow-up reduce confidence in these estimates, and more rigorous, large-scale trials are needed before routine implementation can be recommended. This review extends previous exergaming syntheses by focusing on clinically meaningful outcomes, including falls and injurious falls, while also considering implementation-related factors relevant to real-world uptake.

Trial Registration: PROSPERO CRD42020214721; https://www.crd.york.ac.uk/PROSPERO/view/CRD42020214721

(JMIR Aging 2026;9:e89807) doi:10.2196/89807

KEYWORDS

exergames; active video gaming; video games; virtual reality; exercise; physical activity; aging; fall prevention; accidental falls

Introduction

Rationale

Each year, approximately one-quarter to one-third of adults aged >65 years living in the community experience a fall, a finding consistently reported across international populations [1-3]. A fall is defined as “an unexpected event in which the participant comes to rest on the ground, floor, or lower level” [4]. The consequences of falls range in severity from minor bruising to debilitating fractures or even death [5]. Beyond physical injuries, falls can also have considerable psychosocial effects, as concerns about falling and loss of independence can lead to reduced social engagement and diminished quality of life (QoL) [6].

Falls also place a substantial financial burden on health care systems worldwide due to hospitalizations and the need for long-term care following injurious falls [7]. Many risk factors for falls, including balance and strength deficits, cognitive decline, and reduced muscle mass, are associated with aging and low levels of physical activity [8]. Regular physical activity can mitigate these issues by improving muscle strength, balance, and overall physical function, as well as enhancing cognitive performance, mood, and QoL [9-11]. Strong evidence suggests that physical activity programs, particularly those incorporating balance and functional exercises, can reduce both the rate of falls and the number of older people who experience falls in the community [12]. Nevertheless, adherence to these programs is often modest and tends to decline over time, with systematic reviews reporting full adherence rates as low as 21% for home-based programs [13] and only approximately half of participants remaining adherent at 12 months [14,15].

Exergaming offers a promising alternative by providing an immersive and engaging experience, incorporating gamified elements such as feedback, achievements, and progress tracking to enhance motivation and adherence [16]. Exergames, broadly defined as active video games that integrate gameplay with physical exercise, use real-time motion detection to track players’ movements [17]. These games may be delivered via animated gaming platforms or within 2D or 3D virtual reality (VR) environments and are increasingly being explored as a tool for fall prevention [18,19].

As exergaming can be tailored to an individual’s functional ability, it may make physical activity more accessible, enjoyable, and acceptable for older adults, addressing barriers such as low motivation and physical limitations [16,20]. As a fall prevention tool, exergaming interventions often gamify physical activities targeting balance, strength, 3D movement (eg, dance), endurance [20-22], and cognitive function [23,24]. These interventions can be provided to individuals or groups in a wide range of settings, including home-based settings, community settings, and hospitals, with or without facilitation by a health professional.

Several systematic reviews and meta-analyses have examined exergaming in older adults, reporting improvements in intermediate outcomes, such as balance, physical function, cognitive outcomes, and concerns about falling [25-32]. These findings suggest that exergaming may improve important fall-related risk factors and may enhance exercise engagement through interactive feedback and gamification. However, comparatively less attention has been given to whether these interventions translate into reductions in clinically measured falls and injurious falls, which remain the most meaningful outcomes for patients, clinicians, and health services. In addition, previous reviews have given limited consideration to implementation-relevant factors, such as adherence, acceptability, adverse events, and cost-effectiveness, all of which are important when evaluating the real-world feasibility of integrating exergaming into fall prevention pathways. Furthermore, prior syntheses have rarely distinguished between usual care, evidence-based active exercise controls, and non–evidence-based active comparator interventions, limiting the interpretation of whether exergaming offers benefit beyond established fall prevention practice. A more clinically focused synthesis of the available evidence is therefore needed.

Objectives

To address these gaps, the aim of this systematic review and meta-analysis was to evaluate the effectiveness of exergaming interventions for preventing falls and injurious falls in adults aged ≥60 years. The secondary objectives were to synthesize evidence relating to QoL, concerns about falling, balance confidence, adherence, acceptability, adverse events, and cost-effectiveness and to examine findings according to comparator type.

XSL•FO
RenderX

Methods

Overview

We followed Cochrane systematic review methods [33]. The review was conducted and reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement, using the PRISMA 2020 expanded checklist [34], the PRISMA 2020 for Abstracts checklist [35], and the PRISMA-S extension for reporting literature searches [36]. The study protocol was prospectively registered with PROSPERO (CRD42020214721).

Information Sources and Search Strategy

Searches were completed first in March 2023 and updated using the same search strategy in February 2025. The following databases were searched from inception to this date: MEDLINE (Ovid), Embase (Ovid), CINAHL Plus (EBSCO), PsycINFO (Ovid), and Cochrane Central Register of Controlled Trials (CENTRAL) (Ovid). The PICOS framework was used to develop eligibility criteria around the themes of exergames, falls, and older people (Table 1). The search strategy combined database-specific controlled vocabulary terms (eg, MeSH, Emtree, CINAHL Headings, and APA Thesaurus) with free-text terms. Full search strategies for all databases, including database-specific adaptations, are provided in Supplementary Material S1 in Multimedia Appendix 1. No date restrictions were applied, but searches were limited to studies published in English in accordance with the eligibility criteria. Where published reports lacked sufficient methodological or outcome detail, study authors were contacted for clarification or to request additional data. In addition to database searches, reference lists of included studies and relevant systematic reviews were screened, and forward citation searching was conducted manually by 2 independent reviewers (CE-T and YY) using Google Scholar to identify any additional eligible studies.

Table . Inclusion and exclusion criteria.

Eligibility criterion	Inclusion	Exclusion
Population	• Sample was predominantly aged ≥60 y (operationalized as mean age minus one SD >60 where necessary) • All health conditions	• Sample was predominantly aged <60 y
Interventions	• All types of exergaming interventions • All delivery methods	• Nonexergaming intervention
Control	• No intervention or usual care • Active intervention not believed to reduce falls • Active intervention considered likely to reduce falls	• None
Outcomes	• Primary: falls or injurious falls • Secondary: health-related quality of life, concern about falling or fear of falling, balance confidence, adherence and/or acceptability, adverse events, cost-effectiveness	• Studies where the only outcomes relate to balance, functional mobility, and/or strength
Setting	• All settings	• None
Types of studies	• Randomized controlled trials • Cluster randomized controlled trials	• Nonrandomized studies • Observational, quasi-experimental, qualitative
Aim	• Main purpose is to investigate the effectiveness of exergaming on falls or injurious falls in older people	• Main purpose is not to investigate the effectiveness of exergaming on falls or injurious falls in older people
Type of publication	• Peer-reviewed journal publication	• Not a peer-reviewed journal publication • Conference abstract, theses, letter to editor, reviews, descriptive (editorials, books, and reports), and protocol only
Language	• English language	• Other languages
Dates	• All dates	• None

XSL•FO
RenderX

Eligibility Criteria

We included randomized controlled trials from any setting that evaluated the effects of exergaming interventions on falls or injurious falls in older adults aged ≥60 years (Table 1). This threshold aligns with international definitions of older populations (eg, World Health Organization [2]) and allows inclusion of a broad range of relevant studies. We included trials with generally healthy participants and participants with conditions that may increase the risk of falls (eg, dementia, stroke, Parkinson disease, and frailty).

We included trials in which the main purpose of the study was to investigate the effectiveness of exergaming. We excluded studies that focused solely on the design of exergaming interventions.

Interventions and Comparators

The review considered all types of exergaming interventions and all delivery methods. This included all interactive active games and immersive or nonimmersive VR platforms, and interventions of any length and duration were eligible. Control interventions included usual care, standard exercise programs with an evidence base in fall prevention (eg, strength and balance training), as well as those with no evidence base in fall prevention (eg, an educational leaflet on exercise and low-intensity or low-frequency exercise) [12].

The types of exercise primarily targeted by the intervention were categorized according to the Prevention of Falls Network Europe (ProFaNE) taxonomy [37], as follows: (1) gait, balance, coordination, and functional training; (2) strength or resistance training; (3) flexibility; (4) 3D exercise (eg, Tai Chi or dance); (5) general physical activity; (6) endurance (eg, treadmill walking); and (7) other kinds of exercise.

Outcomes

The primary outcomes of interest were (1) rate of falls (number of falls over a period and falls per person-year), (2) number of fallers (number of people experiencing one or more falls), and (3) number of people who experienced one or more injurious falls [38].

The secondary outcomes were (1) health-related QoL (measured using a validated scale such as the EQ-5D [39]); (2) concern about falling (measured using a validated scale such as the Falls Efficacy Scale-International [40]); (3) balance confidence (measured using a validated scale such as the Activities-Specific Balance Confidence Scale [41]); (4) adherence to the intervention (defined as the extent to which participants completed the prescribed exergaming program, including attendance, exercise frequency, duration, or retention); (5) acceptability of the intervention (defined as participants' perceptions of usability, satisfaction, enjoyment, and willingness to engage with or continue the program, as reported by study authors); (6) adverse events (defined as any negative outcome resulting either directly or indirectly from the assigned treatment [42], measured as the number of people who experienced one or more adverse events); and (7) cost-effectiveness of the intervention. These outcome definitions were based on measures reported in the included studies.

Protocol Registration and Amendments

This review was prospectively registered in PROSPERO (CRD42020214721) and conducted in accordance with the registered protocol. One methodological amendment was made during the review process. We initially planned to combine all active intervention comparators into a single analysis; however, important conceptual and clinical differences were identified between evidence-based fall prevention interventions and non–evidence-based active controls. These comparator groups were therefore analyzed separately to improve the interpretability and clinical relevance of the findings. This decision was made prior to final data synthesis. No other substantive deviations from the registered protocol were made.

Study Selection

Titles and abstracts from the database searches were imported into the Rayyan web app for systematic reviews [43], and duplicate records were removed. Records were then screened for relevance by 2 independent reviewers (selected from a pool of CE-T, LM, EDB, ES, KD, YY, AM, KG, RC, and AA), and full texts of studies considered potentially relevant were assessed for eligibility. The reasons for exclusion at the full-text stage were recorded and are reported in Supplementary Material S2 in Multimedia Appendix 1. The same prespecified eligibility criteria (Table 1) were applied at title or abstract and full-text screening, and the table served as the screening framework for reviewers. No automated decision-making tools were used, and all decisions were made by reviewers. Any disagreements were resolved through discussion with a third independent reviewer from the pool.

Data Extraction

Data were extracted from included studies by 2 independent reviewers (KG and YY) using a standardized data extraction form in an Excel (Microsoft) spreadsheet, and an independent reviewer (CE-T or AA) checked for accuracy. Extracted data included author details, country, year of publication, participant characteristics (eg, age, gender, and health status), study characteristics (eg, design, location, setting, sample size, and length of follow-up), intervention details (eg, exergaming device and dose), outcomes measured, and key results.

We used PROGRESS-Plus criteria (place of residence, race/ethnicity/culture/language, occupation, gender/sex, religion, education, socioeconomic status, and social capital) [44] when extracting participant characteristics.

The TIDieR (Template for Intervention Description and Replication) checklist was used to describe the included interventions and controls [45].

Risk of Bias Assessment

The Excel version of the Cochrane risk of bias (RoB 2.0) tool was used to assess the risk of bias [46,47]. Two independent reviewers assessed the risk of bias for each study, and any disagreements were resolved through discussion with an additional reviewer. Risk of bias was classified as high, low, or some concerns, across the following domains: randomization process, deviations from intended interventions, missing

XSL•FO
RenderX

outcome data, measurement of the outcome, and selection of the reported result.

Reporting Bias Assessment

Reporting bias, including publication bias and selective reporting, was considered during synthesis. Formal statistical assessment (eg, funnel plots or Egger test) was not undertaken because of the small number of studies included in each meta-analysis, making such methods unreliable. Potential reporting bias was therefore considered qualitatively when interpreting the findings.

Assessment of the Certainty of Evidence

We conducted GRADE assessments on the overall evidence for each primary outcome, split by comparator group [48]. Two independent reviewers assessed each outcome (CE-T and AM or LM and SA), and any disagreements were resolved through discussion. Overall bias per outcome was based on the majority rating across domains.

Effect Measures and Data Synthesis

Findings were synthesized narratively and summarized in overview tables to facilitate comparison across interventions. Where 2 or more studies contributed sufficiently clinically comparable data, pooled overall or comparator-specific meta-analyses were conducted in Review Manager (RevMan; version 5.4) [49]. Outcomes informed by single studies are presented for consistency of effect estimation but interpreted narratively. Incidence rate ratios (IRRs) with 95% CIs were calculated for fall rate outcomes, risk ratios (RRs) with 95% CIs for binary outcomes (number of fallers and number of participants experiencing one or more injurious falls), mean differences (MDs) for continuous outcomes reported on the same scale, and standardized MDs (SMDs) where different measurement scales were used. Fixed effects or random effects models were selected according to the clinical comparability between studies. Where pooling was not appropriate because of limited study numbers, heterogeneous outcome measures, or inconsistent reporting, results were synthesized narratively in accordance with Synthesis Without Meta-analysis guidance [50].

Subgroup Analysis

Planned subgroup analysis for the primary outcomes included participant age (older vs younger populations), health condition (eg, healthy, stroke, Parkinson disease, and frailty), and intervention setting (community, residential care home, hospital, assisted care facility, sheltered housing, and retirement community). However, these analyses were not feasible because too few studies contributed data within each subgroup. Comparator-specific analyses were undertaken instead, as these provided the most clinically interpretable basis for exploring variation in effect estimates.

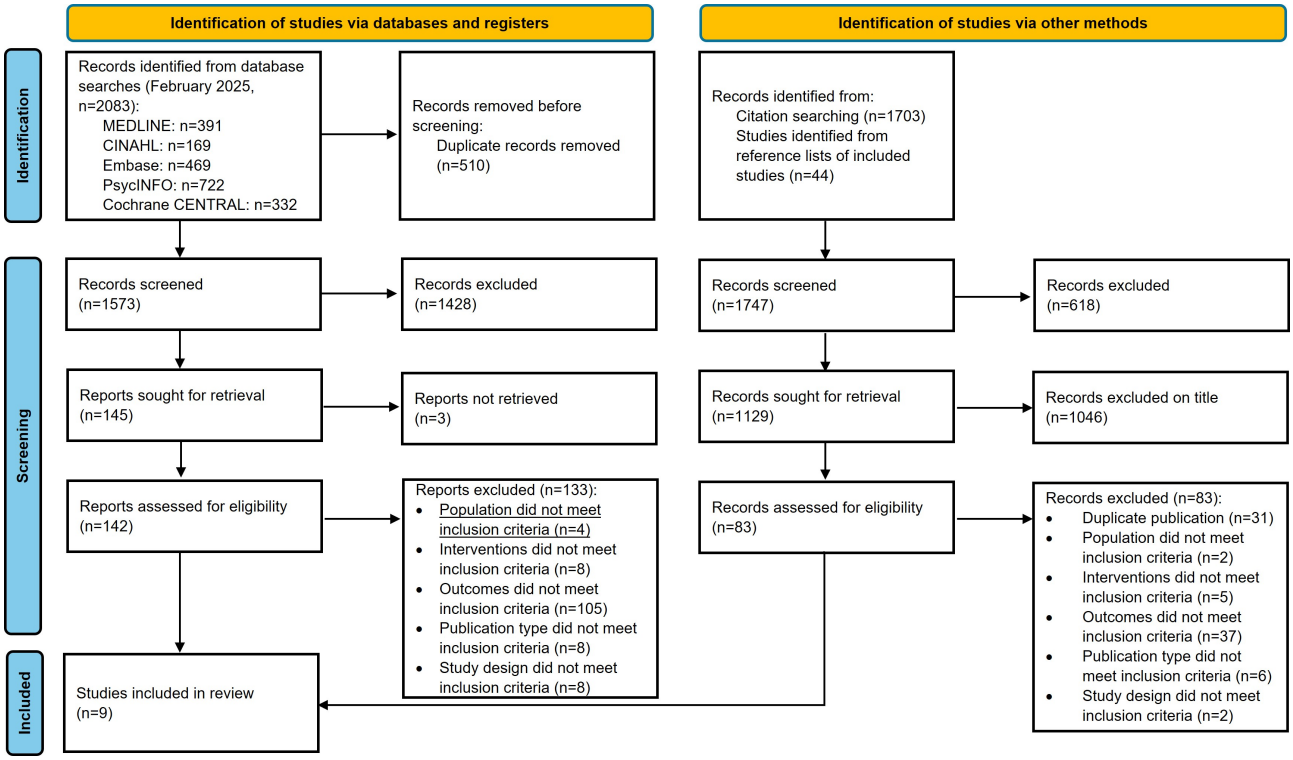
Results

Study Selection

We identified 3830 records, and after removal of duplicates, 3320 (86.7%) were screened on title and abstract, and 225 (5.9%) full-text reports were assessed for eligibility. Of these, 216 (96%) full-text reports were excluded, most commonly because of ineligible study design, inappropriate intervention, ineligible population, or absence of relevant fall outcomes. A full list of studies excluded at the full-text screening, together with reasons for exclusion, is provided in Supplementary Material S2 in Multimedia Appendix 1. Nine randomized controlled trials involving 1385 participants met the inclusion criteria, with 51% (701/1385) in intervention groups and 49% (684/1385) in control groups. Studies varied considerably in participant populations, intervention formats, comparator types, and follow-up duration (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the systematic review process.



Participant Characteristics

Participant characteristics and PROGRESS-Plus aspects of equity are presented in Table 2. In total, 59% (817/1385) were women (intervention groups: 416/817, 51% and control groups: 401/817, 49%). The mean age of participants was 72.8 (SD 7.9; intervention groups: 72.9, SD 7.8, and control groups: 72.6, SD 8.0) years. The mean age of participants in the no contact or usual care control studies was 72.8 (SD 6.9) years; in the active control studies (evidence based), the mean age was 73.1 (SD 9.0) years; and in the active control studies (non–evidence based), the mean age was 72.6 (SD 7.9) years. Four studies recruited participants based on a diagnosis of Parkinson using clinical criteria [51-54]. Participants were recruited in 3 studies on the basis that they had a history of falls [51,53,55], and 1 study specifically recruited participants with mild cognitive impairment [53].

XSL•FO
RenderX

Table . Participant characteristics.

Author, year	Participant characteristics				PROGRESS-Plus[a] measures
	n	Age (y), mean (SD)	Gender (female), n (%)	Clinical condition	
Control group: usual care or no intervention					
Song et al, 2018 [54] (N=53)	I[b]: 28; C[c]: 25	I: 68 (7); C: 65 (7)	I: 16 (52); C: 20 (69)	Idiopathic Parkinson disease	P[d]
Stanmore et al, 2019 [56] (N=92)	I: 49; C: 43	I: 77.9 (8.9); C: 77.8 (10.2)	I: 45 (80.4); C: 38 (76.0)	None reported	P, R[e], O[f], SES[g], SC[h]
Sturnieks et al, 2024 [57] (N=507)	I: 252; C: 255	I: 72.6 (5.7); C: 72.5 (5.5)	I: 178 (70.6); C: 182 (71.4)	None reported (excluded people with neurological, acute psychiatric, or cognitive impairment)	P, E[i], SC
Control group: evidence-based active control					
Fu et al, 2015 [55] (N=60)	I: 30; C: 30	I: 82.4 (3.8); C: 82.3 (4.3)	I: 19 (63.3); C: 20 (66.6)	One or more falls in the past year	P
Gandolfi et al, 2017 [52] (N=70)	I: 36; C: 34	I: 67.5 (7.2); C: 69.8 (9.4)	I: 15 (39.5); C: 10 (26.3)	Parkinson disease (modified Hoehn and Yahr stages 2.5 - 3)	P
Kwok et al, 2016 [58] (N=80)	I: 40; C: 40	I: 70.5 (6.7); C: 69.8 (7.5)	I: 36 (90); C: 30 (80)	43 participants had a history of falls in the past year; mild-to-moderate physical frailty (9 - 5 points) on the short physical performance battery	P, R, L, SC
Control group: non–evidence-based active control					
Alagumoorthi et al, 2022 [51] (N=192)	I: 96; C: 96	I: 69.7 (10); C: 68.5 (9.8)	I: 45 (46.9); C: 33 (34.4)	Idiopathic Parkinson disease (Hoehn-Yahr stage 2.5 - 4). History of at least three falls in previous 3 mo	P
Eggenberger et al, 2015 [59] (N=49)	I: 24; C: 25	I (dance): 77.3 (6.3); C: 80.8 (4.7)	I (dance): 14 (58.3); C: 16 (64.0)	None reported	P
Mirelman et al, 2016 [53] (N=282)	I: 146; C: 136	I: 74.2 (6.9); C: 73.3 (6.4)	I: 48 (33); C: 52 (38)	Self-reported 2 or more falls in the past 6 mo; some with MCI[j] (0.5 on Clinical Dementia Rating scale) or Parkinson disease (Hoehn and Yahr stage II-III)	P, E

[a]PROGRESS-Plus: place of residence, race/ethnicity/culture/language, occupation, gender/sex, religion, education, socioeconomic status, and social capital.

[b]I: intervention.

[c]C: control.

[d]P: place of residence.

[e]R: race, ethnicity, or culture.

[f]O: occupation.

[g]SES: socioeconomic status.

[h]SC: social capital.

[i]E: education.

[j]MCI: mild cognitive impairment

No study included data on the research locale (eg, urban or rural), 2 studies noted race, ethnicity, or culture [56,58], and 1 study each presented data on the participants' language [58], occupation [56], number of years in education [53], and

XSL•FO
RenderX

socioeconomic status [56]. Three studies included some information on aspects, which could contribute to social capital (ie, living arrangements) [56-58].

Study and Intervention Characteristics

Overview

A summary of study and intervention characteristics is presented in Table 3. Studies were conducted across 11 countries, including 1 study each from India [51], Switzerland [59], China [55], Italy [52], Singapore [58], the United Kingdom [56], Australia or New Zealand [54], and Australia [57], and 1 study from a multicountry study (Belgium, Israel, Italy, the Netherlands, and the United Kingdom) [53]. Study settings included 2 clinics: a clinical center [53] and a geriatric clinic [59]. The remaining studies were conducted across hospital and community settings [51], in both community and university laboratory settings [54], in a satellite center, and at home [58], and 4 in community settings, such as the home [52], an assisted living facility [56], a nursing home [55], and 1 unspecified community location [57].

XSL•FO
RenderX

Table . Study and intervention characteristics of included randomized controlled trials.

Author, year	Setting	Intervention	Duration or frequency	Devices	Type of exercise	Type of control	Length of follow-up	Outcomes assessed
Control group: usual care control								
Song et al, 2018 [54]	C[a]	Home-based exergame step training: stepping on central panel to match onscreen directional cues. Accuracy feedback and scores were provided after each round.	15 min per session, 3 times a week for 12 wk	"Dance Dance Revolution" Stepmania	iv[b]	UC[c]: continue with usual health care	3 mo	Falls, CF[d], ADV[e]
Stanmore et al, 2019 [56]	LTCF[f]	Tailored strength and balance exergames: standardized strength and balance exergames individually prescribed and progressed based on ability (eg, increased difficulty, duration, or number of games).	30 min per session, 3 times a week for 12 wk	Microsoft Kinect and tailored software	i[g]	UC: leaflets on fall prevention advice	3 mo	Falls, HQoL[h], CF, ADH[i], ACC[j], ADV, C/CE[k]
Sturnieks et al, 2024 [57]	C	Smart±step program: participants stepped on a Bluetooth-connected mat to match directional targets. Games trained speed, accuracy, motor control, and cognitive skills (working memory, visuospatial skills, dual-tasking, inhibition, and attention).	120 - 150 min per wk for 12 mo	Smart±step computer game system and step mat	i	UC: health information brochure	12 mo	Falls, HQoL, CF, BC[l], ADH, ACC, ADV
Control group: evidence-based active control								
Fu et al, 2015 [55]	LTCF	Wii Fit balance training: balance training games (Soccer Heading, Table Tilt, and Balance Bubble)	1 h per session, 3 times a week for 6 wk	Nintendo Wii Fit balance board	i	AC[m]: conventional balance training, including lower limb muscle strengthening exercises	12 mo	Falls

XSL•FO
RenderX

Author, year	Setting	Intervention	Duration or frequency	Devices	Type of exercise	Type of control	Length of follow-up	Outcomes assessed
Gandolfi et al, 2017 [52]	C	Tele-Wii (home-based virtual reality telerehabilitation): 10 Wii exergames remotely supervised and progressed by a physiotherapist based on each patient's clinical status and improvement	50 min per session, 3 times per week for 7 consecutive weeks (21 sessions)	Nintendo Wii Fit	i and vii[n]	AC: sensory integration balance training of balance and gait exercises	1 mo	Falls, HQoL, BC, ACC, C/CE
Kwok et al, 2016 [58]	C	Nintendo Wii exercise program: WiiActive Balance Board and resistance band training, incorporating cardiovascular, strengthening, calisthenics, and balance training. Sessions included 20 min each of Wii training, individualized exercises, and home exercise	1 h per session, 2 - 3 times a week for 12 wk	Nintendo Wii Active	i, ii[o], iii[p], iv, and vii	AC: traditional gym exercise	3 and 6 mo	Falls, CF, ADH, ACC, ADV
Control group: non–evidence-based active control								
Alagumoorthi et al, 2022 [51]	C	Wii Sports-based strategy training: 8 games, selected based on validated movement analyses to target key balance strategies and movements needed to prevent falls	30 - 40 min per session, 3 times per week for 12 wk	Nintendo Wii console	i	AC: traditional balance training	3 mo	Falls, HQoL

XSL•FO
RenderX

Author, year	Setting	Intervention	Duration or frequency	Devices	Type of exercise	Type of control	Length of follow-up	Outcomes assessed
Eggenberger et al, 2015 [59]	C	VR[q] dance-based aerobic training (DANCE): 1×1 m platform with 4 pressure-sensitive zones that detected steps forward, backward, left, and right. Followed arrow cues on a screen, with feedback on accuracy	1 h per session, twice per week for 6 mo	Impact Dance Platforms, Stepmania	i and iv	AC: treadmill walking (PHYS) and treadmill walking with verbal memory exercise (MEMORY)	12 mo	Falls, CF
Mirelman et al, 2016 [53]	C	Treadmill training with virtual reality: real-time foot-tracking projected onto a virtual environment while walking on a treadmill, with added cognitive challenges (eg, attention, planning, dual-tasking)	45 min per session, 3 times a week for 6 wk	Modified Microsoft Kinect with nonimmersive VR and treadmill	vi	AC: treadmill training	6 mo	Falls, HQoL, ADH, ADV

[a]C: community.

[b]iv: 3D exercise (eg, Tai Chi or dance).

[c]UC: usual care.

[d]CF: concerns about falling.

[e]ADV: adverse event.

[f]LTCF: long-term care facility.

[g]i: gait, balance, coordination, and functional task training.

[h]HQoL: health-related quality of life.

[i]ADH: adherence.

[j]ACC: acceptability.

[k]C/CE: costs/cost-effectiveness.

[l]BC: balance confidence.

[m]AC: active control.

[n]vii: other kinds of exercise, which improve the correct use of ankle and hip strategy during static conditions.

[o]ii: strength or resistance training (alternating biceps curl).

[p]iii: flexibility.

[q]VR: virtual reality.

The interventions in the included studies were described using the TIDieR checklist (see Supplementary Material S3 in Multimedia Appendix 1 for details). A range of exergaming and VR systems were used across studies: 4 used the Nintendo Wii platform, 1 used Microsoft Kinect with tailored software, 1 used a modified Microsoft Kinect combined with nonimmersive VR and treadmill training, 2 used dance-based systems (Konami Dance Dance Revolution or Positive Gaming Impact Dance Platform) paired with a Stepmania mat, and 1 used bespoke stepping software with a pressure-sensitive mat [57].

In terms of the control group, 3 studies included usual care or no intervention [54,56,57], 3 included an evidence-based exercise fall prevention program [52,55,58], and 3 used a non–evidence-based program [51,53,59]. Active controls included balance training (with or without physiotherapy)

XSL•FO
RenderX

[51,52,55], treadmill training [53,59], or gym-based exercises [58].

Intervention Dosage and Delivery Characteristics

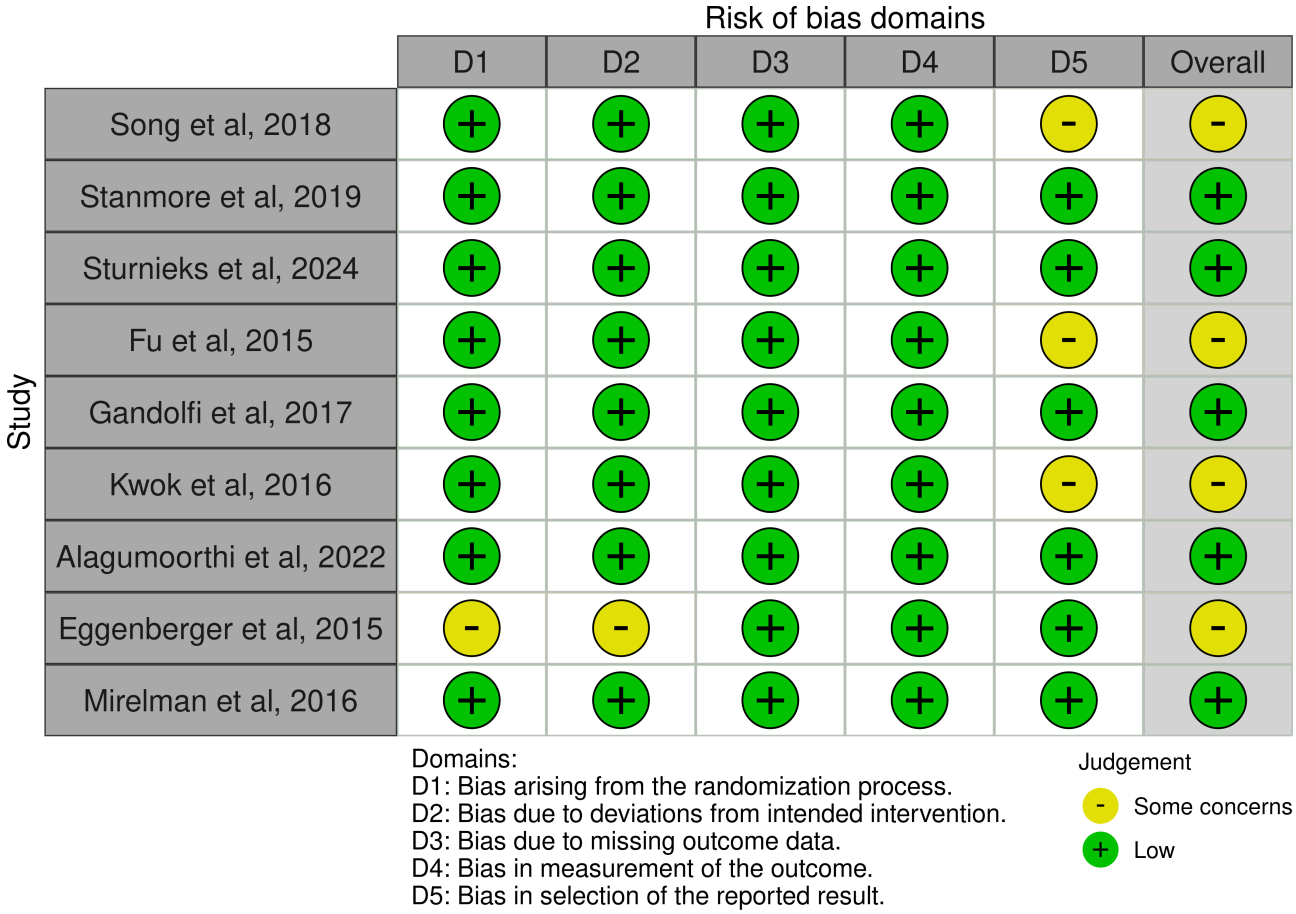
The active exergame dose ranged from 6 to 52 weeks, with a frequency of 2 to 3 sessions per week, where specified. Session duration ranged from 15 minutes to 1 hour. One study gave an overall number of minutes per week to be completed but did not specify how many sessions should be undertaken or how long each should last [57].

All studies involved support for the exergame sessions provided by physiotherapists, health care professionals, or trained postgraduate students. Seven interventions included supervised sessions throughout the study [51-53,55,56,58,59], while the remaining 2 provided supervision during the initial setup and training session, with ongoing technical support and advice available as needed for the remainder of the study [54,57]. Studies did not report on the type and frequency of help needed by participants to undertake the exergame exercises.

Risk of Bias and Certainty of the Evidence

Risk-of-bias (RoB 2) summary assessments are shown in Figure 2.

Figure 2. Risk of bias summary assessments [51-59].



Due to the nature of the intervention, no studies were able to blind participants to their group assignment. Nevertheless, 8 studies were single blinded with assessors blinded to the participant group [51-55,57-59], and the final study was nonblinded because the cluster design revealed participants' allocation due to their assisted living facility [56]. Four studies had some additional concerns over the risk of bias: 3 over the selection of the reported result because there was a lack of information about whether the data that were used to produce the results had been analyzed following a prespecified analysis plan, which had been finalized before the unblinded outcome data were available for analysis [54,55,58]. The fourth study had some concerns over a lack of information on the randomization process and deviations from the intended intervention [59].

Certainty of evidence (GRADE) for each primary outcome is presented in Table 4, with full assessment details provided in Supplementary Material S4 in Multimedia Appendix 1. Key drivers for downgrading of outcome certainty were imprecision and inconsistency across studies.

XSL•FO
RenderX

Table . Overview of Grading of Recommendations Assessment, Development, and Evaluation assessments by primary outcome.

Primary outcome	Comparison 1: usual care or no control	Comparison 2: evidence-based active control	Comparison 3: non–evidence-based active control
Rate of falls	(low)	(low)	(moderate)
Number of fallers	(low)	(low)	(low)
Number of people experiencing injurious falls	(low)	N/A[a]	N/A

[a]N/A: not applicable.

Effectiveness of Interventions

The effectiveness of exergaming interventions is presented according to comparator type: usual care, evidence-based active controls, and non–evidence-based active controls. This approach was used to distinguish the effects of exergaming against usual care or no intervention, against established exercise programs with evidence for fall prevention, and against active control activities not specifically designed to reduce falls.

Comparison 1: Exergames Versus Usual Care or No Intervention Control Group

Three studies, including 652 participants, compared the use of exergames against usual care or no intervention [54,56,57]. Forest plots for this comparison are presented in Figure 3.

XSL•FO
RenderX

Figure 3. Forest plots of exergames versus usual care [54,56,57]. HR: health related; IRR: incidence rate ratio; QoL: quality-adjusted life year.

XSL•FO
RenderX

Primary Outcomes

Rate of Falls: IRR

All 3 studies reported fall rates, which were converted to IRR for comparison (n=652) [54,56,57]. Using a random-effects model, the pooled estimate suggested a possible reduction in fall rate (low certainty evidence). However, the CI crossed the line of no effect, indicating uncertainty in the effect estimate. Substantial heterogeneity was present, and this finding should therefore be interpreted with caution (IRR 0.60, 95% CI 0.35 - 1.01; I^2=86%). Heterogeneity was evident in participants, settings, and intervention components, and one small study, which included participants with Parkinson disease, showed no significant result with wide CIs [54].

Number of Fallers (Number of People Experiencing One or More Falls)

Three studies (n=652), 2 at low risk of bias and 1 with some concerns, presented data on the number of fallers (low certainty evidence): one each at 3-month follow-up [56], 6-month follow-up [54], and a 12-month follow-up [57]. No significant differences were found at 3-month or 6-month follow-up (RR 0.80, 95% CI 0.40 - 1.63 and RR 0.84, 95% CI 0.55 - 1.28, respectively). At 12-month follow-up, exergames may result in a reduction in the number of fallers (RR 0.75, 95% CI 0.61 - 0.92).

Number of People Who Experienced One or More Injurious Falls

One study (n=507) [57] showed exergaming may reduce the relative risk of people experiencing one or more injurious falls by an average of 19% (low certainty evidence). This estimate was imprecise, ranging from a 39% reduction to an increase of 8% (RR 0.81, 95% CI 0.61 - 1.08).

Secondary Outcomes

Quality of Life

One study (n=92) reported on QoL using the EQ-5D-5L VAS scale [56]. The evidence suggests that exergames do not improve QoL at 3-month follow-up (MD −3.40, 95% CI −12.40 to 5.60).

Effects on Concerns About Falling

Three studies (n=652) reported on the effect of exergaming interventions on concerns about falling [54,56,57]. Two studies reported this outcome at 12-week follow-up [54,56] and the third study at 6- and 12-month follow-up [57]. For this study, we focused on the 12-month data, as it represented the longest follow-up period. The evidence from a meta-analysis of the 12-week follow-up data suggests that exergames result in no difference in concerns about falling (SMD −0.28, 95% CI −1.18 to 0.61), and there was high heterogeneity between the studies (I^2=85%). Evidence from the study with the 12-month follow-up period [57] also suggested that exergames did not reduce concerns about falling (SMD −0.10, 95% CI −0.23 to 0.02).

Balance Confidence

No study reported on balance confidence in this comparison.

Acceptability

Two studies (n=599) included acceptability measures [56,57]. The System Usability Scale [60] mean score in the first study was 82.4 (SD 15.5) at 12-week follow-up [56]. The second study measured the System Usability Scale at 6 and 12 months and found a mean score of 80.4 (SD 15.8) at 6-month and 83.3 (SD 13.9) at 12-month follow-up [57]. At all measurement points in both studies, scores were above 80, which is considered excellent and indicates good usability. The Technology Acceptance Model [61] was included in one paper [56] at 12-week follow-up, and overall, all domains were considered high or very high, but with some variation in behavioral intention and perceived usefulness: easy to use, 6.3 (SD 1.4); useful, 5.9 (SD 1.9); favorable attitude, 6.6 (SD 1.2); and intention to use, 5.7 (SD 2.2).

Adherence

All 3 studies reported on retention rates, with a median of 87% and a range from 81% to 91.7% [54,56,57]. The median adherence rate to the minimum prescribed dose of exercise was 69%, ranging from 50.8% to 85%. However, 1 study also noted that only 0.6% of participants achieved the target of 120 minutes of exercise per week over the 12-month intervention period [57].

Adverse Events

No studies reported serious adverse events as a result of the intervention: 2 reported no adverse events from the intervention [56,57] and the third study reported 1 noninjurious fall while completing the exergame step training under the guidance of an experienced physiotherapist [54].

Cost-Effectiveness

One study (n=92) included the cost-effectiveness of the exergame intervention [56]. The authors reported a point estimate incremental cost-effectiveness ratio of £15,209.80 (US $20,239.71) per quality-adjusted life year (QALY). Using 1000 bootstrap replications, they found there was a 61% probability that exergames were cost-effective at the lower National Institute for Health and Care Excellence threshold of £20,000 (US $26,624.08) per QALY, rising to 73% at the upper threshold of £30,000 (US $39,970) per QALY.

Comparison 2: Exergames Versus Evidence-Based Active Control

Three studies (n=210) reported on the use of exergame interventions and an evidence-based active control [52,55,58]. Forest plots for this comparison are presented in Figure 4.

XSL•FO
RenderX

Figure 4. Forest plots of exergames versus evidence-based active controls [52,55,58]. HR: health related; IRR: incidence rate ratio; QoL: quality-adjusted life year.

XSL•FO
RenderX

Primary Outcomes

Rate of Falls: IRR

The IRRs were calculated for all studies, and the evidence from a meta-analysis suggests that exergames result in a reduction in fall rate (IRR 0.41, 95% CI 0.27 - 0.62; $I^{2=}$39%; low certainty evidence).

Number of Fallers (Number of People Experiencing One or More Falls)

A single study (n=80) reported intervention effects on the number of fallers at 12-month follow-up [58]. The evidence, which was of low certainty, showed that exergames result in little to no difference in the number of fallers (RR 1.38, 95% CI 0.62 - 3.06).

Number of People Who Experienced One or More Injurious Falls

No studies reported on this outcome.

Secondary Outcomes

Quality of Life

One study (n=70) examined the effects of an exergame intervention on QoL using the EQ5D scale [52]. Findings indicated that exergames do not increase QoL compared to evidence-based active controls (MD −1.91, 95% CI −8.49 to 4.67).

Concerns About Falling

Concerns about falling were reported in one study (n=80), using the Modified Falls Efficacy Scale [58]. The evidence suggests exergames result in a reduction in concerns about falling at 12-week follow-up (MD 0.80, 95% CI 0.61 - 0.99).

Balance Confidence

Balance confidence was reported in 1 study (n=70), using the Activities-Specific Balance Confidence scale [52]. Exergames may result in little to no difference in balance confidence at 1-month follow-up (MD −4.61, 95% CI −13.10 to 3.88).

Acceptability

Two studies included measures of acceptability (n=150) [52,58]. Evidence from a meta-analysis of these studies suggests that exergames result in little or no difference in acceptability (SMD −0.19, 95% CI −0.53 to 0.16 I^2=0%).

Adherence

One study (n=80) reported on adherence rates [58], assessed through attendance at intervention sessions and the frequency of home exercises performed. Attendance data suggest little to no difference in adherence (MD 0.10, 95% CI −1.02 to 1.22). The home exercise compliance (number of days per week completed over the 1-y follow-up) also suggested no difference in adherence (MD 0.30, 95% CI −0.34 to 0.90).

Adverse Events

Two studies included adverse events [52,58], but both reported no adverse events during the studies or follow-up periods.

Cost Data

Cost-effectiveness was not included in any study; however, one study did consider cost (n=70) [52]. The cost of rehabilitation was calculated as €383.55 (US $439.20) per intervention group participant (€23,299.00 [US $26,679.22] for the whole intervention group) and €602.10 (US $689.45) per control group participant (€28,899.80 [US $33,092.58] for the whole control group).

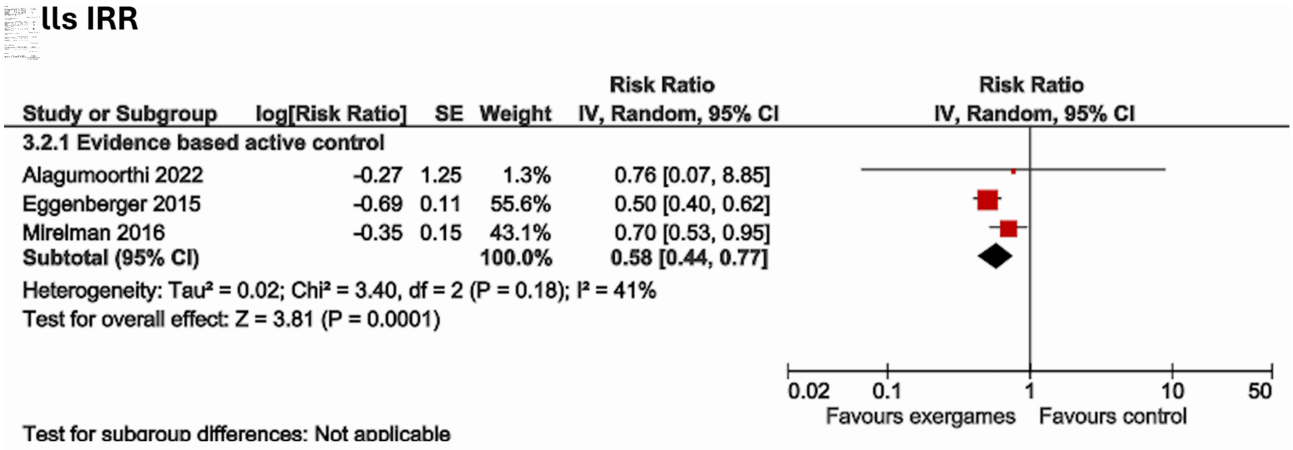
Comparison 3: Exergames Versus Active Intervention (Non–Evidence Based)

Three studies (n=523) reported on the use of exergame interventions and a non–evidence-based active control [51,53,59]. Forest plots for this comparison are presented in Figure 5.

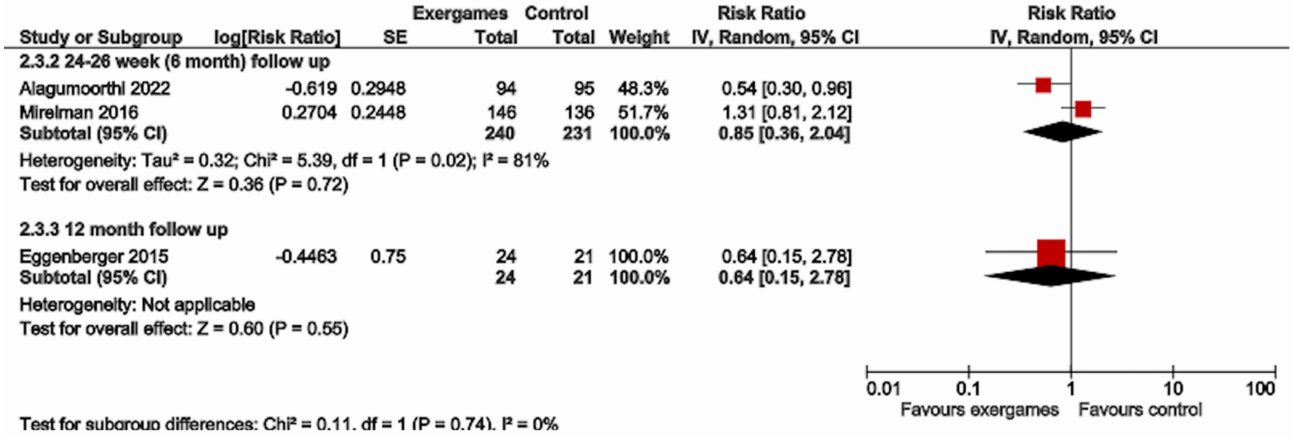
XSL•FO
RenderX

Figure 5. Forest plots of exergames versus non–evidence-based active controls [51,53,59]. HR: health related; IRR: incidence rate ratio; QoL: quality-adjusted life year.

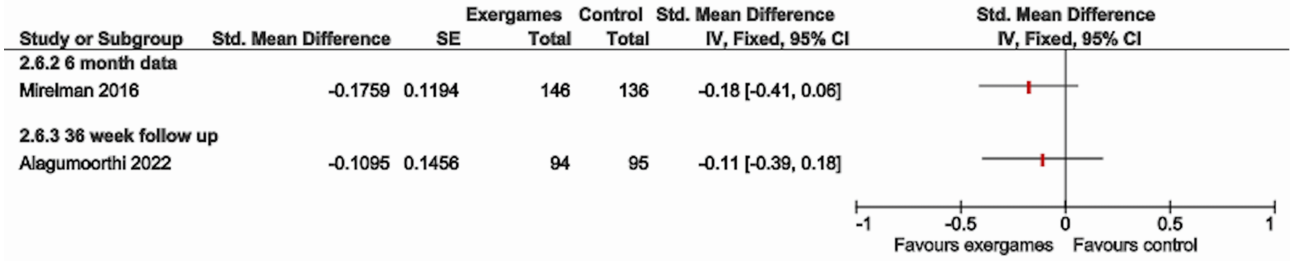
XSL•FO
RenderX



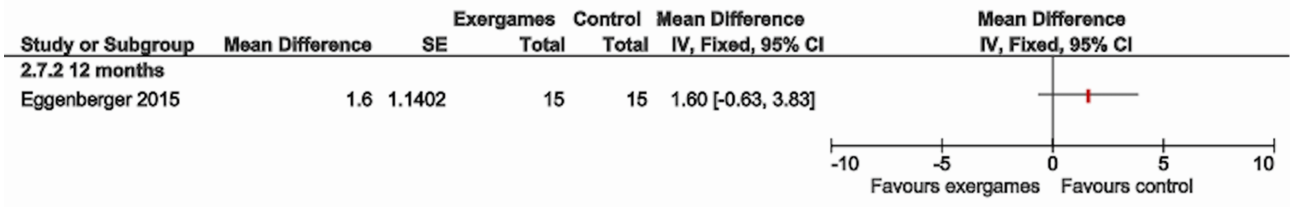
Number of fallers



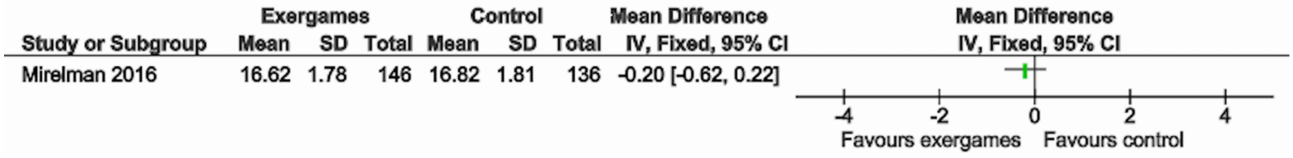
HR QoL



Concerns about falling



Adherence



XSL•FO
RenderX

Primary Outcomes

Rate of Falls: IRR

All 3 studies reported the effects of exergame interventions on the fall rate in this comparison [51,53,59]. Two reported follow-ups at 6 months (n=474) [51,53] and one at 12 months (n=49) [59]. These data were converted to IRR for comparison, and the moderate-certainty evidence from a meta-analysis indicates that exergames likely result in a reduction in fall rate (IRR 0.58, 95% CI 0.44 - 0.77; I^2=41%).

Number of Fallers (Number of People Experiencing One or More Falls)

Three studies (n=523) reported the number of people experiencing one or more falls [51,53,59]. At 6-month follow-up, evidence from a meta-analysis of 2 studies (n=474) suggests no difference in the number of fallers (RR 0.85, 95% CI 0.36 - 2.04) [51,53]. At 12-month follow-up, low-certainty evidence indicated that exergames did not reduce the number of fallers (RR 0.64, 95% CI 0.15 - 2.68) [59].

Number of People Who Experienced 1 or More Injurious Falls

No studies reported on this outcome

Secondary Outcomes

Quality of Life

QoL was measured in 2 studies (n=474) using either the 36-Item Short Form Health Survey (SF-36) [62] at 6-month follow-up [53] or the Parkinson's Disease Questionnaire (PDQ-39) summary index [63] at 36-week follow-up [51]. Evidence from the studies suggests that exergames result in little or no difference in QoL.

Concerns About Falling

One study (n=49) reported on concerns about falling using the FES-I scale [59]. The evidence is very uncertain about the effect of exergames on concerns about falling (MD 1.60, 95% CI −0.63 to 3.83).

Balance Confidence

No studies in this comparison group reported on this outcome.

Acceptability

No studies in this comparison reported on this outcome.

Adherence

Evidence from a single study (n=282) [53] was very uncertain about the effect of exergames on adherence (MD −0.20, 95% CI −0.62 to 0.22).

Adverse Events

Two studies reported on adverse events [51,53]. Although overall 24 events were reported, none occurred because of the study.

Cost-Effectiveness

No studies in this comparison reported on this outcome measure.

Overall Falls Pooled Effect Across Comparator Groups

An exploratory pooled analysis across all comparator groups, with low- to medium-certainty evidence, suggested that exergaming interventions were associated with a lower fall rate than comparator interventions (IRR 0.53, 95% CI 0.41 - 0.68). However, substantial heterogeneity was present (I^2=76%), indicating considerable variation in effect estimates across studies, and therefore, this pooled effect should be interpreted cautiously. Comparator-specific analyses are likely to provide more clinically meaningful insights, as the magnitude and certainty of effect differed across control categories. Differences in participant populations, intervention characteristics, comparator content, and follow-up duration are likely to have contributed to the observed heterogeneity.

Discussion

Summary of Findings

This review synthesized evidence on the effectiveness of exergaming interventions for fall prevention in older adults by examining falls and injurious falls as primary outcomes alongside implementation-relevant secondary outcomes. Comparator-specific analyses suggest that exergaming interventions may reduce fall rates when compared with usual care and some active comparators, although the magnitude and certainty of effect varied across comparisons. When data were pooled across all comparator groups, exergaming was also associated with an overall reduction in fall rate; however, this broad summary estimate should be interpreted cautiously because substantial heterogeneity indicates important variation between intervention contexts, participant groups, and comparator conditions. Moderate-certainty evidence also indicated that exergames may reduce the number of older adults experiencing one or more falls at 12-month follow-up compared with usual care. However, evidence for injurious falls was limited to a single study, and confidence in several pooled estimates was reduced by substantial heterogeneity, small numbers of contributing studies, and imprecision.

These findings suggest that exergaming may have value as an adjunct to established exercise-based fall prevention approaches, particularly where motivation, adherence, and sustained participation in conventional exercise programs are challenging. At the same time, the review highlights that exergaming interventions are not homogeneous, and effectiveness is probably influenced by intervention design, level of supervision, participant health status, comparator type, and follow-up duration.

Evidence for secondary outcomes, including QoL, concerns about falling, balance confidence, adherence, acceptability, and cost-effectiveness, was limited, of low certainty, or inconsistently reported. Adverse events were uncommon and were not clearly attributed to the interventions, but the small number of studies reporting implementation outcomes limits broader conclusions regarding long-term feasibility, sustainability, and economic value. Overall, these findings suggest that exergaming may offer a promising additional approach within fall prevention; however, the current evidence

XSL•FO
RenderX

base is not yet robust enough to support widespread routine adoption.

Comparison With Prior Work

Previous systematic reviews have generally reported beneficial but mixed effects of exergaming on intermediate outcomes, such as balance, mobility, physical function, concerns about falling, and cognitive performance, but have offered limited synthesis of clinically meaningful fall outcomes [64-74]. This review extends this literature by focusing specifically on falls and injurious falls, by examining implementation-related outcomes relevant to real-world delivery, and by considering effects according to comparator type. This provides a more clinically relevant understanding of where exergaming may offer a benefit within fall prevention pathways.

Our findings are broadly consistent with previous reviews in suggesting that exergames can be effective, but they also confirm that these effects are not uniform and may depend on intervention design, comparator type, and duration of follow-up. Previous reviews, including Chen et al [75], have reported short-term improvements were more commonly seen than long-term effects, and this was evident in our included studies. This highlights the importance of longer follow-up periods, in line with ProFaNE recommendations, to establish whether short-term gains translate into sustained reductions in falls over time.

Variation in game design, exercise challenge, alignment with the Systems Framework for Postural Control [76], progression, feedback, and professional supervision may also explain differences in effectiveness across studies. Standardizing key intervention components and clearer reporting of intervention dose and support needs would improve comparability across future studies and improve confidence in findings [45]. This is particularly relevant given that adherence was stronger in studies with structured supervision, consistent with previous work emphasizing declining motivation in unsupported programs [77].

Economic and implementation evidence is also limited. Only 2 included studies reported intervention costs and only one considered cost-effectiveness [56], reflecting the limited economic evaluation seen across the wider exergaming literature [78,79]. Similarly, acceptability and usability were also underreported, despite their importance for long-term use of digital interventions among older adults [16,77,80,81]. Together, these gaps indicate that future exergaming studies should move beyond efficacy measures alone and address whether exergames represent a viable and scalable fall prevention approach for real-world adoption [82,83].

Strengths and Limitations

This review has several strengths, including a comprehensive multi-database search, prospective protocol registration, comparator-stratified analyses, and use of established tools including RoB 2.0 and GRADE. However, several limitations should be considered.

Our search was limited to studies published in English, which may have introduced language bias. There was also considerable clinical and methodological heterogeneity across studies in participant populations, intervention design, comparators, outcome reporting, and follow-up duration, which limited the precision and generalizability of pooled estimates. Exergame interventions varied widely in the types of equipment used, the extent to which games were purpose designed and the degree to which they were based on known effective strength and balance exercises (eg, Otago [84]), the amount of supervision by physiotherapists or health care professionals, and the duration and frequency of delivery. Furthermore, in several studies, exergames were delivered as part of a wider package of support, making it difficult to isolate the individual contribution of the exergame.

Notably, the studies involving support did not report on the amount or types of support needed by participants. Future research should therefore collect and report this information to better understand its impact on outcomes and to help assess the cost of delivery of the intervention in the real world.

Participant health status also varied considerably: 4 studies included people living with Parkinson disease or mild cognitive impairment, and 3 studies specifically recruited people with a history of falls. Such comorbidities, known to affect strength, balance, and falls, may have impacted the study outcomes [85-88]. Although we planned to conduct subgroup analyses, including those based on health status and PROGRESS-Plus, as outlined in the PROSPERO protocol, the relatively small number of studies included in meta-analyses limited the feasibility of more extensive subgroup analyses and formal assessment of publication bias. Many of the studies also had small sample sizes, limiting their ability to detect intervention effects and increasing the risk of type II error. While pooled “overall” estimates are presented, these should be interpreted with caution, and subgroup analyses by comparator type are likely to provide more clinically meaningful insights.

In addition, pooled estimates were calculated using the standard DerSimonian-Laird random effects approach in RevMan, which, although widely used, may provide less conservative interval estimates than alternative methods such as Hartung-Knapp-Sidik-Jonkman when only a small number of studies are available. Follow-up periods ranged from 1 to 12 months but were often relatively short, with 22% of studies including a follow-up of less than 6 months, and 33% at 6 months. Given that the ProFaNE recommends a minimum follow-up of 12 months in fall prevention trials, because these interventions may produce delayed effects, requiring sustained adherence over time to demonstrate efficacy [12], the longer-term sustainability of exergaming effects remains uncertain.

Implications for Policy and Practice

The findings of this review suggest that exergames may play a role as a complementary component of fall prevention programs for older adults, particularly given the evidence that they may reduce fall rates and the number of people experiencing a fall when compared with usual care. However, the current evidence does not support exergames as a stand-alone replacement for established exercise-based fall prevention interventions.

XSL•FO
RenderX

While the included studies varied considerably in intervention design and dosage, several interventions associated with favorable outcomes incorporated structured, progressive exercise delivered multiple times per week over several weeks or months. This is broadly consistent with established fall prevention exercise principles, which emphasize sufficient frequency, intensity, and progression [12,89]. Given the heterogeneity of interventions and limited evidence on optimal dosage, specific recommendations for exergaming interventions cannot yet be made, and further research is needed to determine the most effective training parameters. However, given the variability in intervention design and delivery, standardized guidelines for exergame design and delivery could be developed to ensure greater consistency across interventions. Purpose-designed serious exergames with appropriate theoretical underpinnings should be developed to unlock the full potential of exergame-enhanced fall prevention interventions [90]. This would facilitate the integration of exergaming with established exercise-based fall prevention programs such as Otago [91] or FaME [92], potentially optimizing effectiveness.

Adherence remains a recognized challenge. Most studies reported good adherence rates when structured support, such as supervision by physiotherapists, was provided, suggesting policies promoting supported or remotely monitored delivery may improve engagement and effectiveness [77]. Ensuring equitable access will therefore require investment in user-friendly technology, digital literacy programs, and technical support to reduce barriers to sustained participation among older adults [93-95].

Finally, our review highlights the need for longer follow-up periods and more robust economic evaluations before exergames can be confidently integrated into routine service delivery. Given the preliminary evidence suggesting cost-effectiveness within National Institute for Health and Care Excellence thresholds, further evaluations are needed to confirm these findings. Policymakers should consider further evaluations to guide future resource allocation and determine whether exergames represent a financially sustainable addition to health care and social care fall prevention pathways.

Conclusions

This review suggests that exergaming interventions may help reduce falls among older adults, particularly in comparisons against active intervention control groups, while pooled analyses also indicate potential overall benefit across diverse intervention settings. However, these findings should be interpreted cautiously because effect estimates varied considerably between studies and were limited by substantial heterogeneity, modest sample sizes, and relatively few long-term follow-up assessments.

Evidence for injurious falls and implementation-relevant outcomes, such as adherence, acceptability, QoL, and cost-effectiveness, remains limited or inconsistently reported. As such, although exergaming appears to offer a potentially engaging and clinically relevant adjunct to existing fall prevention programs, the evidence base is not yet sufficiently robust to support widespread routine implementation. Larger, methodologically rigorous trials with standardized fall reporting and longer follow-up are needed to determine which exergaming formats, doses, and delivery models are most effective for sustained fall prevention in older adults.

Acknowledgments

Generative artificial intelligence (ChatGPT-5.3, OpenAI) was used under full human supervision for language editing, proofreading, and formatting support during manuscript revision. No artificial intelligence tools were used for study selection, data extraction, data analysis, or interpretation of findings. All final scientific decisions and manuscript content remain the responsibility of the authors.

Funding

This work is funded in part by the National Institute for Health and Care Research Applied Research Collaboration-Greater Manchester (NIHR ARC-GM; reference NIHR200174) and Professor Chris Todd's NIHR Senior Investigator Awards (references NIHR200299 and NIHR205156). CE-T, AM, CS, YY, SA, RC, and CT are supported by the NIHR ARC-GM. KG, LM, and CT are supported by the National Institute for Health and Care Research, Policy Research Unit in Older People in Healthy Ageing (NIHR206119). YY is additionally supported by the Vivensa Foundation, NIHR, and University of Manchester Healthy Ageing PhD Programme (PDM2202/90). The views expressed are those of the authors and not necessarily those of the National Health Service (NHS), the NIHR, the Department of Health and Social Care or its partner organizations or any other funder organization. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The original findings of this study are encompassed within the article. For additional inquiries, please contact the corresponding authors.

Authors' Contributions

Conceptualization: LM, ES, CT, KD, EDB
Data curation: CE-T, LM, KLG, YY, AM, SA, AA

XSL•FO
RenderX

Formal analysis: CE-T, CS
Funding acquisition: ES, CT
Investigation: CE-T, LM, EDB, ES, KD, YY, AM, KLG, RC, AA
Methodology: CE-T, LM, ES, KD, EDB, CT
Project administration: CE-T, LM
Supervision: CE-T, LM, ES, CT
Validation: CE-T, LM, YY, KLG, AM, SA, AA, RC
Visualization: CE-T, CS
Writing – original draft: CE-T, LM
Writing – review & editing: CE-T, LM, CS, AM, YY, KLG, SA, RC, AA, KD, EDB, ES, CT

Conflicts of Interest

EDB was a cofounder of Dividat [96], an ETH Zurich (Switzerland) spin-off company that developed an exergame platform and was associated to the company as an external advisor. No revenue was paid (or promised to be paid) directly to EDB or his institution over the 36 months prior to this work. ES is coauthor of paper 12. KD is coauthor of paper 15. ES and CT are coauthors of paper 43. EDB is coauthor of paper 45. These 4 authors played no part in data extraction, data analysis, or quality assessment.

Multimedia Appendix 1
Supplementary material, including full search strategy, excluded paper list, TIDieR checklist, and GRADE assessment.
[PDF File, 728 KB - aging_v9i1e89807_app1.pdf]

Checklist 1
PRISMA checklist.
[PDF File, 187 KB - aging_v9i1e89807_app2.pdf]

References

1. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):afac205. [doi: 10.1093/ageing/afac205] [Medline: 36178003]
2. World Health Organization. WHO Global Report on Falls Prevention in Older Age: World Health Organization; 2008.
3. Web-based Injury Statistics Query and Reporting System (WISQARS). Centers for Disease Control and Prevention (CDC). 2025. URL: https://wisqars.cdc.gov/ [accessed 2026-07-15]
4. Lamb SE, Jørstad-Stein EC, Hauer K, Becker C, Prevention of Falls Network Europe and Outcomes Consensus Group. Development of a common outcome data set for fall injury prevention trials: the Prevention of Falls Network Europe consensus. J Am Geriatr Soc 2005 Sep;53(9):1618-1622. [doi: 10.1111/j.1532-5415.2005.53455.x] [Medline: 16137297]
5. Lord SR, Sherrington C, Hicks C. Epidemiology of falls and fall-related injuries. In: Falls in Older People: Risk Factors, Strategies for Prevention and Implications for Practice, 3rd edition 2021:3-22. [doi: 10.1017/9781108594455.002]
6. Schoene D, Heller C, Aung YN, Sieber CC, Kemmler W, Freiberger E. A systematic review on the influence of fear of falling on quality of life in older people: is there a role for falls? Clin Interv Aging 2019;14:701-719. [doi: 10.2147/CIA.S197857] [Medline: 31190764]
7. Falls: key facts. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2026-06-29]
8. Uusi-Rasi K, Patil R, Karinkanta S, et al. A 2-year follow-up after a 2-year RCT with vitamin D and exercise: effects on falls, injurious falls and physical functioning among older women. The Journals of Gerontology 2017 Sep 1;72(9):1239-1245. [doi: 10.1093/gerona/glx044]
9. Bherer L, Erickson KI, Liu-Ambrose T. A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. J Aging Res 2013;2013(1):657508. [doi: 10.1155/2013/657508] [Medline: 24102028]
10. Gheysen F, Poppe L, DeSmet A, et al. Physical activity to improve cognition in older adults: can physical activity programs enriched with cognitive challenges enhance the effects? A systematic review and meta-analysis. Int J Behav Nutr Phys Act 2018 Jul 4;15(1):63. [doi: 10.1186/s12966-018-0697-x] [Medline: 29973193]
11. Bjerk M, Brovold T, Skelton DA, Liu-Ambrose T, Bergland A. Effects of a falls prevention exercise programme on health-related quality of life in older home care recipients: a randomised controlled trial. Age Ageing 2019 Mar 1;48(2):213-219. [doi: 10.1093/ageing/afy192] [Medline: 30615055]
12. Sherrington C, Fairhall NJ, Wallbank GK, et al. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev 2019 Jan 31;1(1):CD012424. [doi: 10.1002/14651858.CD012424.pub2] [Medline: 30703272]
13. Simek EM, McPhate L, Haines TP. Adherence to and efficacy of home exercise programs to prevent falls: a systematic review and meta-analysis of the impact of exercise program characteristics. Prev Med 2012 Oct;55(4):262-275. [doi: 10.1016/j.ypmed.2012.07.007] [Medline: 22813920]

XSL•FO
RenderX

14. Nyman SR, Victor CR. Older people's participation in and engagement with falls prevention interventions in community settings: an augment to the Cochrane systematic review. Age Ageing 2012 Jan;41(1):16-23. [doi: 10.1093/ageing/afr103] [Medline: 21875865]
15. Teng B, Gomersall SR, Hatton AL, Khan A, Brauer SG. Predictors of real-world adherence to prescribed home exercise in older patients with a risk of falling: a prospective observational study. Aging Med (Milton) 2023 Dec;6(4):361-369. [doi: 10.1002/agm2.12270] [Medline: 38239715]
16. Meekes W, Stanmore EK. Motivational determinants of exergame participation for older people in assisted living facilities: mixed-methods study. J Med Internet Res 2017 Jul 6;19(7):e238. [doi: 10.2196/jmir.6841] [Medline: 28684385]
17. Witherspoon L. ACSM information on exergaming. : American College of Sports Medicine; 2013 URL: https://pressbooks.uwf.edu/app/uploads/sites/16/2023/03/Exergaming-1.pdf [accessed 2026-06-29]
18. Chu CH, Quan AML, Souter A, Krisnagopal A, Biss RK. Effects of exergaming on physical and cognitive outcomes of older adults living in long-term care homes: a systematic review. Gerontology 2022;68(9):1044-1060. [doi: 10.1159/000521832] [Medline: 35290983]
19. de Bruin ED, Schoene D, Pichierri G, Smith ST. Use of virtual reality technique for the training of motor control in the elderly. Some theoretical considerations. Z Gerontol Geriatr 2010 Aug;43(4):229-234. [doi: 10.1007/s00391-010-0124-7] [Medline: 20814798]
20. McCaskey MA, Schättin A, Martin-Niedecken AL, de Bruin ED. Making more of IT: enabling intensive motor cognitive rehabilitation exercises in geriatrics using information technology solutions. Biomed Res Int 2018;2018(1):4856146. [doi: 10.1155/2018/4856146] [Medline: 30581853]
21. Choi SD, Guo L, Kang D, Xiong S. Exergame technology and interactive interventions for elderly fall prevention: a systematic literature review. Appl Ergon 2017 Nov;65:570-581. [doi: 10.1016/j.apergo.2016.10.013] [Medline: 27825723]
22. Gonçalves A, Muñoz J, Cameirão MS, Gouveia É, Sousa H, Bermúdez I Badia S. The benefits of custom exergames for fitness, balance, and health-related quality of life: a randomized controlled trial with community-dwelling older adults. Games Health J 2021 Aug;10(4):245-253. [doi: 10.1089/g4h.2020.0092] [Medline: 34370609]
23. Chen X, Chen L, Chen X, Ye Q, Cai G, Zeng Y. Assessing the impact of immersive virtual reality technology on the psychological recovery of patients with Parkinson's disease depression: study protocol of a randomized controlled trial. Trials 2024 Oct 25;25(1):715. [doi: 10.1186/s13063-024-08552-5] [Medline: 39456018]
24. Tseng CN, Gau BS, Lou MF. The effectiveness of exercise on improving cognitive function in older people: a systematic review. J Nurs Res 2011 Jun;19(2):119-131. [doi: 10.1097/JNR.0b013e3182198837] [Medline: 21586989]
25. Alhasan H, Alandijani E, Bahamdan L, et al. Home-based virtual reality training for enhanced balance, strength, and mobility among older adults with frailty: systematic review and meta-analysis. JMIR Serious Games 2025 Jul 18;13:e67146. [doi: 10.2196/67146] [Medline: 40680186]
26. Azimi Sanavi B, Talebpour M, Saatchian V, Oveysi Sani F, Mirakhori F. Effects of virtual reality-based exercise training on risk of falling, walking capacity, and quality of life in healthy older adults: a systematic review and meta-analysis research. J Appl Gerontol 2026 Jul;45(7):1223-1244. [doi: 10.1177/07334648251362512] [Medline: 40738368]
27. Hernandez-Martinez J, Cid-Calfucura I, Vásquez-Carrasco E, Branco BHM, Herrera-Valenzuela T, Valdés-Badilla P. Benefits of exergaming regarding to conventional physical therapies on balance and fall risk in prefrail and frail older people: a meta-analysis of randomized controlled trials. Inquiry 2025;62:469580251372362. [doi: 10.1177/00469580251372362] [Medline: 40956936]
28. Kashi SK, Saatchian V. Effectiveness of exergame intervention on depressive symptoms, daily living activities, and fear of falling in older adults: a systematic review and meta-analysis of randomized controlled trials. Clin Gerontol 2026;49(1):5-19. [doi: 10.1080/07317115.2025.2503250] [Medline: 40372700]
29. Kasicki K, Klimek Piskorz E, Rydzik Ł, et al. Effectiveness of virtual reality-based training versus conventional exercise programs on fall-related functional outcomes in older adults with various health conditions: a systematic review. J Clin Med 2025 Aug 6;14(15):5550. [doi: 10.3390/jcm14155550] [Medline: 40807171]
30. Li JZ, Tang X, Hu W, et al. Efficacy of exergame interventions on muscle strength in older adults: a systematic review and meta-analysis. Exp Gerontol 2025 Dec;212:112958. [doi: 10.1016/j.exger.2025.112958] [Medline: 41187796]
31. Martínez Montilla LA, López Cruces K, Calderón Erazo HS, et al. Effectiveness of virtual reality in balance training for fall prevention in older adults: systematic review. Sports Med Arthrosc Rev 2023 Jun 1;31(2):41-48. [doi: 10.1097/JSA.0000000000000367] [Medline: 37418173]
32. Saragih ID, Chen YM, Suarilah I, Susanto H, Lee BO. Virtual reality intervention for fall prevention in older adults: a meta-analysis. J Nurs Scholarsh 2025 Sep;57(5):759-775. [doi: 10.1111/jnu.70028] [Medline: 40533914]
33. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ. Cochrane Handbook for Systematic Reviews of Interventions Version 65 (Updated August 2024): Cochrane; 2024. URL: www.cochrane.org/handbook [accessed 2026-06-29]
34. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
35. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160. [doi: 10.1136/bmj.n160] [Medline: 33781993]

XSL•FO
RenderX

36. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev 2021 Jan 26;10(1):39. [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
37. Lamb SE, Becker C, Gillespie LD, et al. Reporting of complex interventions in clinical trials: development of a taxonomy to classify and describe fall-prevention interventions. Trials 2011 May 17;12:125. [doi: 10.1186/1745-6215-12-125] [Medline: 21586143]
38. Schwenk M, Lauenroth A, Stock C, et al. Definitions and methods of measuring and reporting on injurious falls in randomised controlled fall prevention trials: a systematic review. BMC Med Res Methodol 2012 Apr 17;12(1):50. [doi: 10.1186/1471-2288-12-50] [Medline: 22510239]
39. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
40. Kempen G, Yardley L, van Haastregt JCM, et al. The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. Age Ageing 2008 Jan;37(1):45-50. [doi: 10.1093/ageing/afm157] [Medline: 18032400]
41. Activities-specific balance confidence scale. AbilityLab SR. 2013. URL: https://www.sralab.org/rehabilitation-measures/activities-specific-balance-confidence-scale [accessed 2022-02-02]
42. Ethgen M, Boutron I, Baron G, Giraudeau B, Sibilia J, Ravaud P. Reporting of harm in randomized, controlled trials of nonpharmacologic treatment for rheumatic disease. Ann Intern Med 2005 Jul 5;143(1):20-25. [doi: 10.7326/0003-4819-143-1-200507050-00006] [Medline: 15998751]
43. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
44. O'Neill J, Tabish H, Welch V, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. J Clin Epidemiol 2014 Jan;67(1):56-64. [doi: 10.1016/j.jclinepi.2013.08.005] [Medline: 24189091]
45. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 7;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
46. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Assessing risk of bias in a randomized trial. In: Cochrane Handbook for Systematic Reviews of Interventions, Second Edition: John Wiley & Sons; 2019:205-228. [doi: 10.1002/9781119536604.ch8]
47. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
48. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 2011 Apr;64(4):380-382. [doi: 10.1016/j.jclinepi.2010.09.011]
49. Review Manager (RevMan). Version 5.4. Cochrane. URL: https://www.cochrane.org/authors/handbooks-and-manuals/style-manual/references/reference-types/software [accessed 2026-07-15]
50. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368:l6890. [doi: 10.1136/bmj.l6890] [Medline: 31948937]
51. Alagumoorthi G, D BJ, Thirunavukarasu S, V R, A K. Effectiveness of Wii sports- based strategy training in reducing risk of falling, falls and improving quality of life in adults with idiopathic Parkinson's disease- a randomized comparative trial. Clin Rehabil 2022 Aug;36(8):1097-1109. [doi: 10.1177/02692155221089030] [Medline: 35315706]
52. Gandolfi M, Geroin C, Dimitrova E, et al. Virtual reality telerehabilitation for postural instability in Parkinson's disease: a multicenter, single-blind, randomized, controlled trial. Biomed Res Int 2017;2017:7962826. [doi: 10.1155/2017/7962826] [Medline: 29333454]
53. Mirelman A, Rochester L, Maidan I, et al. Addition of a non-immersive virtual reality component to treadmill training to reduce fall risk in older adults (V-TIME): a randomised controlled trial. Lancet 2016 Sep 17;388(10050):1170-1182. [doi: 10.1016/S0140-6736(16)31325-3] [Medline: 27524393]
54. Song J, Paul SS, Caetano MJD, et al. Home-based step training using videogame technology in people with Parkinson's disease: a single-blinded randomised controlled trial. Clin Rehabil 2018 Mar;32(3):299-311. [doi: 10.1177/0269215517721593] [Medline: 28745063]
55. Fu AS, Gao KL, Tung AK, Tsang WW, Kwan MM. Effectiveness of exergaming training in reducing risk and incidence of falls in frail older adults with a history of falls. Arch Phys Med Rehabil 2015 Dec;96(12):2096-2102. [doi: 10.1016/j.apmr.2015.08.427] [Medline: 26360975]
56. Stanmore EK, Mavroeidi A, de Jong LD, et al. The effectiveness and cost-effectiveness of strength and balance Exergames to reduce falls risk for people aged 55 years and older in UK assisted living facilities: a multi-centre, cluster randomised controlled trial. BMC Med 2019 Feb 28;17(1):49. [doi: 10.1186/s12916-019-1278-9] [Medline: 30813926]
57. Sturnieks DL, Hicks C, Smith N, et al. Exergame and cognitive training for preventing falls in community-dwelling older people: a randomized controlled trial. Nat Med 2024 Jan;30(1):98-105. [doi: 10.1038/s41591-023-02739-0] [Medline: 38228913]
58. Kwok BC, Pua YH. Effects of WiiActive exercises on fear of falling and functional outcomes in community-dwelling older adults: a randomised control trial. Age Ageing 2016 Sep;45(5):621-627. [doi: 10.1093/ageing/afw108] [Medline: 27496921]

XSL•FO
RenderX

59. Eggenberger P, Theill N, Holenstein S, Schumacher V, de Bruin ED. Multicomponent physical exercise with simultaneous cognitive training to enhance dual-task walking of older adults: a secondary analysis of a 6-month randomized controlled trial with 1-year follow-up. Clin Interv Aging 2015;10(1711-32):1711-1732. [doi: 10.2147/CIA.S91997] [Medline: 26604719]
60. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry: CRC Press; 1996:189-194.
61. Davis FD. A technology acceptance model for empirically testing new end-user information systems: theory and results [PhD thesis]. : Massachusetts Institute of Technology; 1985 URL: http://hdl.handle.net/1721.1/15192 [accessed 2026-06-29]
62. Jenkinson C, Coulter A, Wright L. Short form 36 (SF36) health survey questionnaire: normative data for adults of working age. BMJ 1993 May 29;306(6890):1437-1440. [doi: 10.1136/bmj.306.6890.1437] [Medline: 8518639]
63. Jenkinson C, Fitzpatrick R, Peto V, Greenhall R, Hyman N. The Parkinson's Disease Questionnaire (PDQ-39): development and validation of a Parkinson's disease summary index score. Age Ageing 1997 Sep;26(5):353-357. [doi: 10.1093/ageing/26.5.353] [Medline: 9351479]
64. Alhagbani A, Williams A. Home-based exergames for older adults balance and falls risk: a systematic review. Physical & Occupational Therapy In Geriatrics 2021 Jul 3;39(3):241-257. [doi: 10.1080/02703181.2020.1867286]
65. Fang Q, Ghanouni P, Anderson SE, et al. Effects of exergaming on balance of healthy older adults: a systematic review and meta-analysis of randomized controlled trials. Games Health J 2020 Feb;9(1):11-23. [doi: 10.1089/g4h.2019.0016] [Medline: 31800322]
66. Prosperini L, Tomassini V, Castelli L, et al. Exergames for balance dysfunction in neurological disability: a meta-analysis with meta-regression. J Neurol 2021 Sep;268(9):3223-3237. [doi: 10.1007/s00415-020-09918-w] [Medline: 32447551]
67. Lapierre N, Um Din N, Belmin J, Lafuente-Lafuente C. Exergame-assisted rehabilitation for preventing falls in older adults at risk: a systematic review and meta-analysis. Gerontology 2023;69(6):757-767. [doi: 10.1159/000528854] [Medline: 36580901]
68. Hai L, Hou HY, Zhou C, Li HJ. The effect of exergame training on physical functioning of healthy older adults: a meta-analysis. Games Health J 2022 Aug;11(4):207-224. [doi: 10.1089/g4h.2021.0173] [Medline: 35653720]
69. Manser P, Herold F, de Bruin ED. Components of effective exergame-based training to improve cognitive functioning in middle-aged to older adults - a systematic review and meta-analysis. Ageing Res Rev 2024 Aug;99:102385. [doi: 10.1016/j.arr.2024.102385] [Medline: 38914262]
70. Nishchyk A, Chen W, Pripp AH, Bergland A. The effect of mixed reality technologies for falls prevention among older adults: systematic review and meta-analysis. JMIR Aging 2021 Jun 30;4(2):e27972. [doi: 10.2196/27972] [Medline: 34255643]
71. Pacheco TBF, de Medeiros CSP, de Oliveira VHB, Vieira ER, de Cavalcanti FAC. Effectiveness of exergames for improving mobility and balance in older adults: a systematic review and meta-analysis. Syst Rev 2020 Jul 18;9(1):163. [doi: 10.1186/s13643-020-01421-7] [Medline: 32682439]
72. Schoene D, Valenzuela T, Toson B, et al. Interactive cognitive-motor step training improves cognitive risk factors of falling in older adults - a randomized controlled trial. PLoS ONE 2015;10(12):e0145161. [doi: 10.1371/journal.pone.0145161] [Medline: 26673919]
73. Stojan R, Voelcker-Rehage C. A systematic review on the cognitive benefits and neurophysiological correlates of exergaming in healthy older adults. J Clin Med 2019 May 23;8(5):734. [doi: 10.3390/jcm8050734] [Medline: 31126052]
74. Zheng L, Li G, Wang X, et al. Effect of exergames on physical outcomes in frail elderly: a systematic review. Aging Clin Exp Res 2020 Nov;32(11):2187-2200. [doi: 10.1007/s40520-019-01344-x] [Medline: 31520334]
75. Chen Y, Zhang Y, Guo Z, Bao D, Zhou J. Comparison between the effects of exergame intervention and traditional physical training on improving balance and fall prevention in healthy older adults: a systematic review and meta-analysis. J Neuroeng Rehabil 2021 Nov 24;18(1):164. [doi: 10.1186/s12984-021-00917-0] [Medline: 34819097]
76. Tahmosybayat R, Baker K, Godfrey A, Caplan N, Barry G. Movements of older adults during exergaming interventions that are associated with the Systems Framework for Postural Control: a systematic review. Maturitas 2018 May;111:90-99. [doi: 10.1016/j.maturitas.2018.03.005] [Medline: 29673837]
77. Buyle M, Jung Y, Pavlou M, Gonzalez SC, Bamiou DE. The role of motivation factors in exergame interventions for fall prevention in older adults: a systematic review and meta-analysis. Front Neurol 2022;13:903673. [doi: 10.3389/fneur.2022.903673] [Medline: 35989930]
78. Collado‐Mateo D, Merellano‐Navarro E, Olivares PR, García‐Rubio J, Gusi N. Effect of exergames on musculoskeletal pain: a systematic review and meta‐analysis. Scandinavian Med Sci Sports 2018 Mar;28(3):760-771. [doi: 10.1111/sms.12899]
79. van Santen J, Meiland FJM, Dröes RM, van Straten A, Bosmans JE. Cost-effectiveness of exergaming compared to regular day-care activities in dementia: results of a randomised controlled trial in The Netherlands. Health Soc Care Community 2022 Sep;30(5):e1794-e1804. [doi: 10.1111/hsc.13608] [Medline: 34657346]
80. Herren S, Seebacher B, Mildner S, et al. Exergame (ExerG)-based physical-cognitive training for rehabilitation in adults with motor and balance impairments: usability study. JMIR Serious Games 2025 Feb 14;13:e66515. [doi: 10.2196/66515] [Medline: 39951650]

XSL•FO
RenderX

81. de Medeiros CSP, Farias LBA, Santana M, Pacheco TBF, Dantas RR, Cavalcanti F. A systematic review of exergame usability as home-based balance training tool for older adults usability of exergames as home-based balance training. PLoS ONE 2024 Aug 22;19(8):e0306816. [doi: 10.1371/journal.pone.0306816] [Medline: 39172915]
82. Patel SB, Iqbal FM, Lam K, Acharya A, Ashrafian H, Darzi A. Characterizing behaviors that influence the implementation of digital-based interventions in health care: systematic review. J Med Internet Res 2025 Jun 12;27:e56711. [doi: 10.2196/56711] [Medline: 40505130]
83. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
84. Campbell J, Robertson C. Otago Exercise Programme to Prevent Falls in Older Adults: A Home-Based, Individually Tailored Strength and Balance Retraining Programme: Accident Compensation Corporation (ACC); 2007.
85. Alenazi AM, Alhwoaimel NA, Alqahtani BA, et al. Multiple long-term conditions and disability are independently associated with higher risk of fall among community adults: a cross-sectional study. Risk Manag Healthc Policy 2024;17:1407-1416. [doi: 10.2147/RMHP.S463570] [Medline: 38828104]
86. Huberty S, Freystätter G, Wieczorek M, et al. Association between multimorbidity and rate of falls: a 3-year 5-country prospective study in generally healthy and active community-dwelling adults aged ≥70 years. J Am Med Dir Assoc 2023 Jun;24(6):804-810. [doi: 10.1016/j.jamda.2022.12.011] [Medline: 36657487]
87. Chantanachai T, Taylor ME, Lord SR, et al. Risk factors for falls in community-dwelling older people with mild cognitive impairment: a prospective one-year study. PeerJ 2022;10:e13484. [doi: 10.7717/peerj.13484]
88. Eost-Telling C, McNally L, Yang Y, et al. The association between delirium and falls in older adults in the community: a systematic review and meta-analysis. Age Ageing 2024 Nov 28;53(12):afae270. [doi: 10.1093/ageing/afae270] [Medline: 39686680]
89. Suen J, Dawson R, Kneale D, et al. Qualitative comparative analysis of exercise interventions for fall prevention in residential aged care facilities. BMC Geriatr 2024 Sep 3;24(1):728. [doi: 10.1186/s12877-024-05246-0] [Medline: 39227773]
90. Manser P, de Bruin ED, Temprado JJ, Bherer L, Herold F. Beyond “just” fun: the role of exergames in advancing health promotion and disease prevention. Neuroscience & Biobehavioral Reviews 2025 Sep;176:106260. [doi: 10.1016/j.neubiorev.2025.106260]
91. Shubert TE, Goto LS, Smith ML, Jiang L, Rudman H, Ory MG. The Otago Exercise Program: innovative delivery models to maximize sustained outcomes for high risk, homebound older adults. Front Public Health 2017;5:54. [doi: 10.3389/fpubh.2017.00054] [Medline: 28386536]
92. Skelton D, Dinan S, Campbell M, Rutherford O. Tailored group exercise (Falls Management Exercise -- FaME) reduces falls in community-dwelling older frequent fallers (an RCT). Age Ageing 2005 Nov;34(6):636-639. [doi: 10.1093/ageing/afi174] [Medline: 16267192]
93. Hepburn J, Williams L, McCann L. Barriers to and facilitators of digital health technology adoption among older adults with chronic diseases: updated systematic review. JMIR Aging 2025 Sep 11;8:e80000. [doi: 10.2196/80000] [Medline: 40934502]
94. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
95. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults’ intention to adopt digital technologies. JMIR Aging 2023 May 16;6(1):e44564. [doi: 10.2196/44564] [Medline: 37191976]
96. A unique approach to cognitive-motor training. Dividat. URL: https://dividat.com/en [accessed 2026-07-21]

Abbreviations

GRADE: Grading of Recommendations Assessment, Development, and Evaluation
IRR: incidence rate ratio
MD: mean difference
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROGRESS-Plus: place of residence, race/ethnicity/culture/language, occupation, gender/sex, religion, education, socioeconomic status, and social capital
QALY: quality-adjusted life year
QoL: quality of life
RR: risk ratio
SMD: standardized mean difference
TIDieR: Template for Intervention Description and Replication
VR: virtual reality

XSL•FO
RenderX

Edited by AF Abdulai; submitted 18.Dec.2025; peer-reviewed by L Taylor, R Stemplewski, WKC Leung; revised version received 11.May.2026; accepted 17.Jun.2026; published 31.Jul.2026.

Please cite as:
Eost-Telling C, McGarrigle L, Shi C, Money A, Yang Y, Lazo Green K, Ahmed S, Christie R, Aminu A, Delbaere K, de Bruin ED, Stanmore E, Todd C
Exergaming Interventions for Preventing Falls and Injurious Falls in Older People: Systematic Review and Meta-Analysis of Randomized Controlled Trials
JMIR Aging 2026;9:e89807
URL: https://aging.jmir.org/2026/1/e89807
doi:10.2196/89807

© Charlotte Eost-Telling, Lisa McGarrigle, Chunhu Shi, Annemarie Money, Yang Yang, Kimberly Lazo Green, Saima Ahmed, Rachel Christie, Abodunrin Aminu, Kim Delbaere, Eling D de Bruin, Emma Stanmore, Chris Todd. Originally published in JMIR Aging (https://aging.jmir.org), 31.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Rationale, Design, and Baseline Characteristics of the TARGET (Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments) Cohort of Older Adults for Assessing Fall and Fracture Risk: Prospective Cohort Study

Angelique Chan[1,2], PhD; Abhijit Visaria, PhD; Rahul Malhotra[1,2], MD; Navrag B Singh[3], PhD; Benedikt Helgason[3,4], PhD; Victor R Schinazi[3,5], PhD; Kok Yang Tan[1], MSocSci; David B Matchar[2], MD; Rita Sim, BSc; Vanessa Jean Wen Koh[2], PhD; Wei Xuan Lai[2], MD, PhD; Lakkhina Troeung[1], PhD; Catherine Wen Huey Lim[1], BSc (Hons); Sai G S Pai[3], PhD; Kai Zhe Tan[3], PhD; Anitha D Praveen[3], PhD; Dheeraj Jha[3], PhD; Stephen J Ferguson[4], PhD; Giorgio Colombo[3], PhD; Karolina Minta[3], PhD; Ecosse L Lamoureux[2], PhD; William R Taylor[4], PhD

[1]Centre for Ageing Research & Education (CARE), Duke-NUS Medical School, 8 College Road, Singapore, Singapore

[2]Health Services Research & Population Health (HSRPH), Duke-NUS Medical School, Singapore, Singapore

[3]Future Health Technologies (FHT), Singapore-ETH Centre, Singapore, Singapore

[4]Institute for Biomechanics, ETH Zurich, Zurich, Zurich, Switzerland

[5]Faculty of Society & Design, Bond University, Gold Coast, Queensland, Australia

Corresponding Author:
Angelique Chan, PhD
Centre for Ageing Research & Education (CARE), Duke-NUS Medical School, 8 College Road, Singapore, Singapore

Abstract

Background: Falls and fractures are a major clinical concern for older adults. International clinical guidelines recommend annual fall risk screening for all adults aged 60 years and older. However, existing falls risk–screening algorithms have shown limited sensitivity and specificity in detecting future falls. Emerging health technologies, including wearable sensors, in-silico finite element models (FEMs), and virtual reality (VR) technology, show promise in enhancing fall and fracture risk assessments by providing more personalized objective data to support targeted primary prevention. However, large-scale longitudinal studies are required to evaluate their predictive accuracy and cost-effectiveness for systematic community–based screening.

Objective: The TARGET (Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments) study is an ongoing national prospective cohort study in Singapore designed to develop cost-effective and scalable frameworks for the early detection and prevention of falls and fractures using novel health technologies. This paper describes the rationale, design, and baseline characteristics of TARGET, and provides preliminary insights into how these novel technologies perform in discriminating between older adults with and with no history of falls at study baseline.

Methods: A total of 2291 community-dwelling Singapore residents aged ≥60 years were enrolled between 2022 and 2024. Participants underwent a comprehensive baseline fall risk assessment comprising (1) home-based interview to assess sociodemographic, anthropometric, cognitive, physical, functional, and psychosocial status; (2) gait assessment using wearable inertial measurement units (IMUs) motion sensors; (3) dual-energy X-ray absorptiometry and whole-body 3D scans to construct subject-specific FEMs; and (4) VR-based cognitive assessment to probe spatial navigation deficits. Prospective follow-up for 2 years is ongoing, with linkage to national electronic medical records (EMRs) to obtain detailed clinical, medication, and pathological data.

Results: Of the total cohort, 58% (1327/2291) were female, with a mean age of 74.6 years at baseline. TARGET represents a relatively healthy older adult population, with 83% (1898/2291) remaining fully independent, 44% (1016/2290) with mild cognitive impairment, 50% (440/879) with osteopenia, and 15% (358/2291) reporting falls in the past year. Significant baseline functional, psychosocial, cognitive, and biomechanical differences were identified between fallers and nonfallers. Fallers had poorer physical and psychosocial health, greater gait variability, and a higher risk of osteoporotic fractures. Novel risk prediction models integrating epidemiological, IMUs, FEM, VR, and EMR data are in development and being validated against gold standard clinical assessments, with cost-effectiveness evaluations of the economic feasibility of adopting these screening measures at scale.

Conclusions: Leveraging comprehensive epidemiological, EMR, and health technology data, TARGET is well positioned to identify novel biomarkers and systemic interactions that may predict falls and fracture risk, with potential application to other age-related diseases with shared pathophysiology. In particular, TARGET's focus on community-based screening using technologies

XSL•FO
RenderX

that can be administered by nonclinical personnel has the potential to reduce substantial burden and costs within the health care sector.

(JMIR Aging 2026;9:e89285) doi:10.2196/89285

KEYWORDS

aging; falls; fractures; health technology; predictive algorithms; population health

Introduction

Falls and fall-related injuries are a major public health and clinical challenge worldwide, especially given the ubiquity of population aging [1]. Approximately 30% of community-dwelling adults aged 65 years or older sustain at least 1 fall each year, increasing to 50% in persons aged older than 80 years [2,3]. Falls are a prominent clinical feature in later-life diseases affecting older adults, defined as persons aged 60 years or older. Notably, falls occur frequently in neurological disorders, affecting up to 70% of individuals with Parkinson disease [4], at least 50% of stroke survivors [5], and 20% - 30% of older adults with peripheral neuropathy [6]. The incidence of falls has been estimated to be 50% among older adults with diabetes, particularly when peripheral neuropathy or hypoglycemia is present [7]. Moreover, falls occur in 30% - 50% of older adults with cardiovascular diseases [8] or cancer, especially among those receiving chemotherapy [9]. Together, these data highlight the considerable significance of falls in the context of aging and disease.

Critically, falls are often a sentinel event for older adults and mark a turning point in the trajectory toward the end of life [10]. A single fall can result in hospital admission, major surgical intervention, and a prolonged convalescence [11]. Recurrent falls are a widely accepted clinical indicator of advanced underlying pathology [12]. Notably, severe acute fall-related injuries such as hip fractures often trigger cascading complications, including prolonged hospitalization and secondary infections [13-15]. The 12-month mortality rate following a hip fracture is between 22% and 29% with excess mortality observed up to 10 years postinjury compared with age-matched controls [16]. In the United States and Europe alone, US $50 billion is spent annually on medical costs related to nonfatal fall injuries, while US $754 million is spent related to fatal falls [17].

Age-related physiological decline significantly increases fall susceptibility [18]. Senescent cells accumulate with age, displaying shortened telomeres and altered signaling that impair tissue repair [19], while sarcopenia progressively weakens skeletal muscles, slowing reflexes and postural control [20]. Changes in visual, vestibular, and proprioceptive pathways compromise precise feedback on body positioning and motion [21]. In parallel, age-related neurodegeneration, including brain atrophy and reduced synaptic connectivity, disrupts the central coordination of motor signals [22]. Endothelial dysfunction and reduced mitochondrial capacity also impair cardiovascular reserves, manifesting as orthostatic hypotension or arrhythmias that lead to sudden drops in cerebral perfusion [23]. Collectively, these age-related physiological and cellular changes not only elevate the risk of falls but also contribute to broader age-related diseases through shared pathophysiological pathways [18].

Given the multisystem involvement of falls, the challenge remains how to optimally screen for fall risk in older adults at the population level. Current best practice is focused on secondary and tertiary fall risk prevention and management following a fall event. The seminal American Geriatrics Society and British Geriatrics Society Clinical Practice Guidelines [24] and recent World Falls Guidelines [15] both recommend annual screening of adults aged 65 years or older for fall risk during routine health care encounters. Currently, there is no universally validated biomarker for predicting falls. While inflammatory mediators (eg, interleukin-6 and tumor necrosis factor-α), hormonal factors (eg, insulin-like growth factor-1 and vitamin D metabolites), and markers of neuromuscular integrity (eg, creatine kinase and neurofilament light chain) have been investigated, there remains no definitive gold standard biomarker for fall risk prediction [25,26]. In the absence of validated biomarkers, fall history in the past 12 months is commonly used as a first-line screener for future fall risk, as research shows that a previous fall is the single best predictor for future falls [27]. The relative risk of falling is 3 times greater for older adults who have previously fallen than for those who have not [28,29], and those who present to health care services with a fall or fall-related injury have a 70% risk of falling again in the next year without intervention [15]. Multifactorial investigations are then recommended for older adults with a positive 12-month fall history to identify specific risk factors, including medication review and assessments for gait, balance, strength, neurological function, cardiovascular status, and visual acuity [24,30].

However, this approach is limited as it does not allow for the identification of at-risk individuals prior to their first fall for targeted primary prevention. Indeed, the sensitivity of common falls-screening algorithms based on the American Geriatrics Society and British Geriatrics Society and World Falls Guidelines, such as the STEADI (Stopping Elderly Accidents, Deaths, and Injuries) algorithm [31], has been shown to be around 60% - 70% with a high false-negative rate in the low-risk categorization group [32,33]. Moreover, current screening approaches rely on patient recall of fall history, which may be impaired by age-related cognitive decline [34], as well as clinician judgment of gait patterns, which can vary depending upon experience, training, and underlying knowledge of gait [35]. Among older adults with a negative 12-month fall history, gait and balance deficits have been identified as the most consistent predictors of future falls [29]. Thus, recognition of gait and balance problems is crucial for identifying persons at risk of first-time falls. However, gait patterns ascertained from clinic and laboratory-based settings often differ from the real-world movement patterns that cause falls, posing an added

XSL•FO
RenderX

challenge for clinicians to assess [36]. Furthermore, implementation of physician-led screening in clinical practice has been challenging, with studies finding that over 60% of primary care providers did not routinely ask older adults about fall history or initiate multifactorial falls assessments due to time and resource constraints [31].

In view of these limitations, there has been considerable interest in novel and scalable approaches that can objectively screen and identify individuals at risk of falls, including first-time falls, in real-world settings [36,37]. The ability to predict and prevent falls holds considerable significance in the context of global aging and disease. It is especially critical in Asia, where many regions are experiencing unprecedented population aging. Singapore, in particular, is among the top 5 nations most affected by population aging, with the proportion of older adults in the population expected to reach 40% by 2050 [38]. The standardized incidence rate of fractures among adults aged older than 50 years in Singapore is 314.2 per 100,000, ranking third highest globally and the highest in Asia [39]. By 2050, more than 50% of hip fractures are predicted to occur in Asia [40]. The estimated total economic burden for osteoporotic fractures alone is expected to cost at least SG $289.9 million (SGD $1=US $0.79 as of September 11, 2026) by 2035 [41]. Thus, there is an urgent need for solutions to reduce the health and associated socioeconomic burden of falls and fractures in Singapore and other Asian populations.

While a number of cohort studies have investigated fall or fracture risk factors in Singapore and Asia, the focus of these studies was on general aging [42-44], broad population health [45,46], cancer [46], and eye disease [47], with falls or fractures as a secondary outcome. To bridge this knowledge gap, we implemented a national prospective cohort study of older adults in Singapore to study fall and fracture risk, titled the “Targeted Assessment and Recruitment of Geriatrics for Effective fall prevention Treatments (TARGET)” study. The specific focus of the TARGET study is risk prediction and stratification to support effective therapeutic measures. Its objectives are to (1) develop and validate personalized screening algorithms for assessing the risk of falls (including first-time falls) and fractures among community-dwelling older adults, (2) create risk stratification methods to segment the older adult population into different fall and fracture risk levels with individualized intervention recommendations, (3) conduct cost-effectiveness analyses to support clinical decision-making by identifying the optimal therapeutic window of falls and fractures risk where older adults derive the greatest benefit from interventions, and (4) determine the psychosocial, cognitive, and sociodemographic correlates of the risk of falls and fractures. In this paper, we provide a detailed description of the rationale, design, methodology, and baseline characteristics of the TARGET cohort.

Methods

Study Design

TARGET is an ongoing national prospective cohort study of community-dwelling older adults in Singapore, evaluating 3 novel health technology approaches to identify objective functional, cognitive, and imaging-based biomarkers of fall and fracture risk for population-based screening. Between October 2022 and October 2024, a total of 2291 older adults aged 60 years and older were recruited. After study enrollment, participants underwent a comprehensive baseline fall risk assessment (T_0), followed by 8 follow-up assessments (T_1-T_8) via telephone every 3 months for 2 years to prospectively monitor their fall, fracture, and cognitive status. A 2-year follow-up period was selected to inform the development of clinically meaningful predictive models, which can identify individuals at short- (<1 year) and medium-term (1 - 2 years) risk for falls for targeted preventive intervention. This duration also provides a balance between collecting sufficient fall events for statistical validity and minimizing participant attrition and recall bias. Ongoing follow-up of the TARGET cohort beyond the initial 2 years is planned to ascertain longer-term fall and fracture risk and outcomes.

Ethical Considerations

Ethics approval for the study was granted by the National University of Singapore Institutional Review Board (NUS-IRB-2021 - 168) and ETH Zurich Ethics Commission (EK 2021-N-145). Written informed consent was obtained from all participants at study enrollment. Participant information sheets and informed consent forms were translated into Chinese, Malay, and Tamil languages by a certified translation company. All participants received an SG $50 (SGD $1=US $0.79 as of September 11, 2026) grocery voucher as a token of appreciation upon successful completion of the baseline home-based assessment. In addition, participants received an SG $140 (SGD $1=US $0.79 as of September 11, 2026) grocery voucher, including the reimbursement for transportation costs, for completion of the baseline clinic assessment.

Sampling and Recruitment

Overview

A multistage random sample of 7501 households with at least 1 older adult resident (Singapore citizen or permanent resident) was purchased from the Department of Statistics of Singapore representing approximately 2.2% of all households with an older adult resident in Singapore in 2022. Household addresses were sampled across all 5 geographic regions of Singapore (Central, East, North, Northeast, and West). To increase the participant count and precision of statistical estimates within minority subgroups, we oversampled addresses with predominant (1) Malay and (2) Indian residents, and with (3) at least 1 resident aged 75 years or older by a factor of 2.

Eligibility Criteria

Individuals residing in the random sample of households were eligible for inclusion in the TARGET cohort if they were (1) a Singapore citizen or permanent resident, (2) aged 60 - 105 years, (3) spoke at least 1 of the 4 official languages in Singapore (English, Mandarin [or a Chinese dialect], Malay, and Tamil), (4) ambulatory with or without mobility aids, (5) able to see with or without glasses, and (6) able to hear with or without hearing aids. Individuals with at least one of the following criteria at the time of recruitment were excluded: (1) possible delirium or dementia defined as an Abbreviated Mental

XSL•FO
RenderX

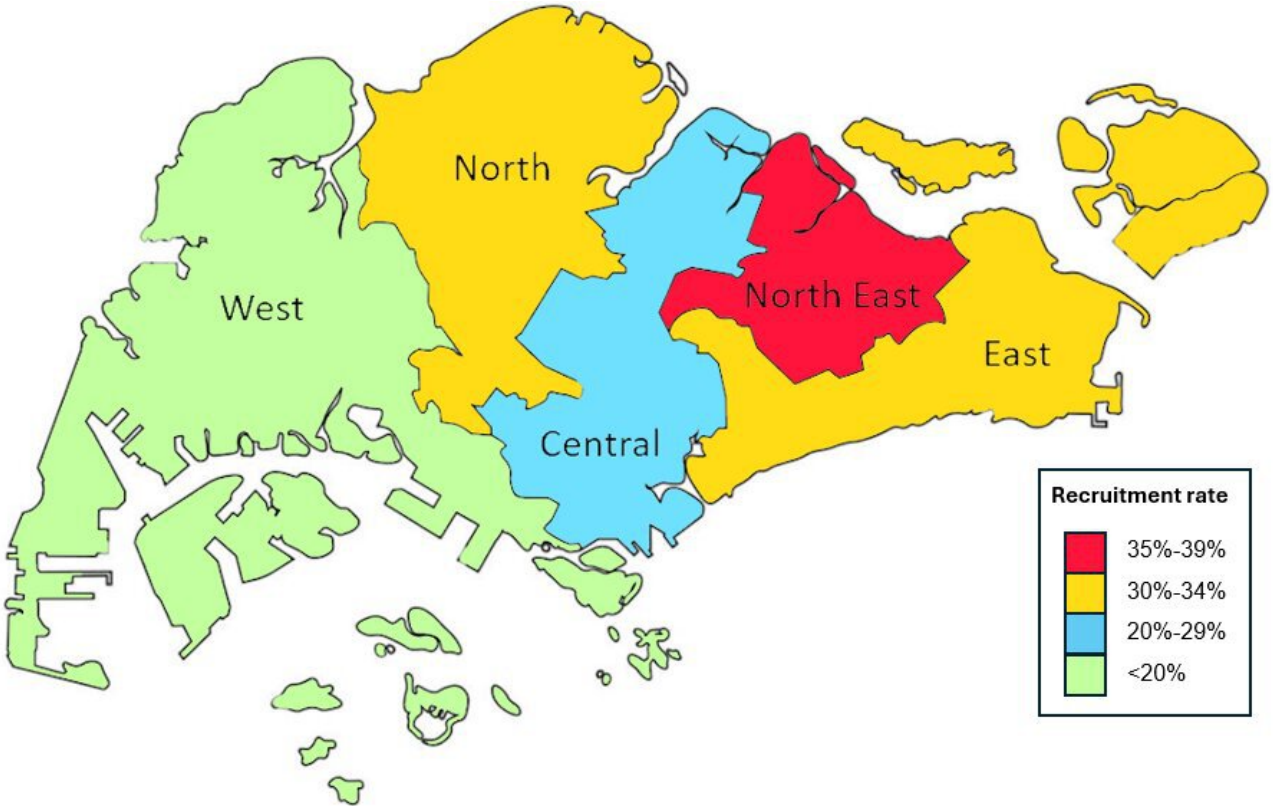
Test score <6 [48]; (2) blood pressure outside the normal range (defined as ≥200/110 mm Hg or ≥180/100 mm Hg and heart rate >120 per minute or ≤90/60 mm Hg); (3) any symptoms of chest discomfort, breathlessness, dizziness, or profuse sweating; or (4) any acute cardiac event in the past 30 days.

Recruitment

An invitation letter was mailed to each household in the random sample stating that participation in the study was completely voluntary, and the study team could be contacted to opt out of the study. Trained field interviewers then conducted home visits 2 - 4 weeks later to households that did not opt out. All willing potential participants within each household were first screened to determine their eligibility against the inclusion and exclusion criteria. Multiple residents from the same household were able to participate if they independently met the eligibility criteria. Once participants consented to the study, they were able to complete the baseline assessment on the same day, if convenient, or reschedule the assessment for a later date. Interviewers made up to 5 visits at differing times of day to each residence to recruit participants. Households, which remained uncontactable after 5 visits, were excluded due to nonresponse.

At least 1 participant was recruited from 1674 households, giving a household recruitment rate of 22.3% (1674/7501). The household participant yield was 1.3 participants per household (range: 1-4). Anecdotally, the field interviewers reported that the households approached were generally reluctant to interact with strangers, coinciding with active national scam prevention campaigns targeting older adults in Singapore during the recruitment period. Demographic characteristics were compared between household responders and nonresponders. There was no significant difference in the age group of the household reference person (primary resident) between household responders and nonresponders. However, differences in housing type and region were identified. The proportion of private households (ie, residential properties that are not subsidized, built, and managed by the Singapore government) was significantly higher among nonresponders. Response rates also significantly differed by geographic region (Figure 1). The highest response was within the Northeast Region (577/1516, 38.1%) while the West Region (215/1549, 13.9%) yielded the lowest response rate.

Figure 1. TARGET (Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments) study household recruitment rate by geographic region in Singapore.



Data Collection

Data Collection Overview

Table 1 summarizes key measures, variables, and outcomes. Both primary and secondary data sources were used. Primary data collected included epidemiological survey data (eg, sociodemographic, anthropometric, cognitive, psychosocial, physical, and environmental data), biomechanical data, clinical imaging, and spatial performance data to provide a holistic understanding of fall and fracture risk factors. Additionally, secondary data in the form of electronic medical records (EMRs) containing data on health status, health care use, prescriptions, procedures, and mortality status are being linked from the Singapore Ministry of Health's TRUST (Trusted Research and Real-World Data Utilisation and Sharing Tech) platform [49] for participants who have consented to this linkage.

XSL•FO
RenderX

Table . Data collected at baseline and follow-up assessments for the TARGET[a] study.

Category	Key measures or variables	Baseline (T_0)	Follow-up (T_1-T_8)
1. Fall risk survey			
1.1. Fall history	Number of falls, number of injurious falls	✓	✓
1.2. Anthropometry	Height, weight, blood pressure, and hand grip strength	✓	—[b]
1.3. Sociodemographics	Age, gender, ethnicity, education, employment, housing, and family	✓	—
1.4. Cognitive impairment	Montreal Cognitive Assessment [50]	✓	—
	Telephone Montreal Cognitive Assessment [51]	—	T_2, T_4, T_6, T_8
1.5. Social network	Modified Lubben Social Network-6 [52]	✓	—
1.6. Quality of life	EQ-5D-5L [53]	✓	—
1.7. Environmental	Modified Life Space Questionnaire [54]	✓	—
1.7. Environmental	Home safety	✓	—
1.8. Functional independence	Modified Barthel Index [55]	✓	—
1.8. Functional independence	Lawton Instrumental Activities of Daily Living [56]	✓	—
1.9. Health	Medical history	✓	—
1.9. Health	Medications	✓	—
1.9. Health	Health and lifestyle behaviors	✓	—
1.9. Health	International Physical Activity Questionnaire-Short Form [57]	✓	—
1.10. Fall risk	Three Key Questions	✓	—
1.10. Fall risk	Fall Risk Questionnaire [58]	✓	—
1.10. Fall risk	Iconographical Falls Efficacy Scale [59,60]	✓	—
1.11. Fracture risk	FRAX tool [61] items including age, gender, weight, height, previous fragility fracture, parental history of hip fracture, smoking, glucocorticoid treatment, rheumatoid arthritis, other secondary causes of osteoporosis, and alcohol	✓	—
1.12. Psychological	Brief Resilience Scale [62]	✓	—
1.12. Psychological	Patient Health Questionnaire [63]	✓	—
2. Gait assessment using wearable sensors	Walking speed, cadence, gait asymmetry, stride time, stride length, stance time, swing time, double-limb support time, heel strike angle, toe-off angle, head stability, postural control, and acceleration patterns	✓	—
3. DXA[c] scan	Bone mineral density of lumbar spine and proximal hip, whole body 3D composition (bone mineral content, fat mass, lean mass, total mass) of trunk, limbs, and head)	✓	—
4. Spatial navigation assessment	Path integration, egocentric pointing, mapping, associative memory, perspective taking, and Santa Barbara Sense-of-Direction Scale [64]	✓	—

XSL•FO
RenderX

Category	Key measures or variables	Baseline (T_0)	Follow-up (T_1-T_8)
5. Electronic medical record linkage	Outpatient (GP[d] and polyclinic) and inpatient medical records (diagnosis, procedures, medications, and pathology)	✓	✓

[a]TARGET: Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments.

[b]Not available.

[c]DXA: dual-energy X-ray absorptiometry.

[d]GP: general practitioner.

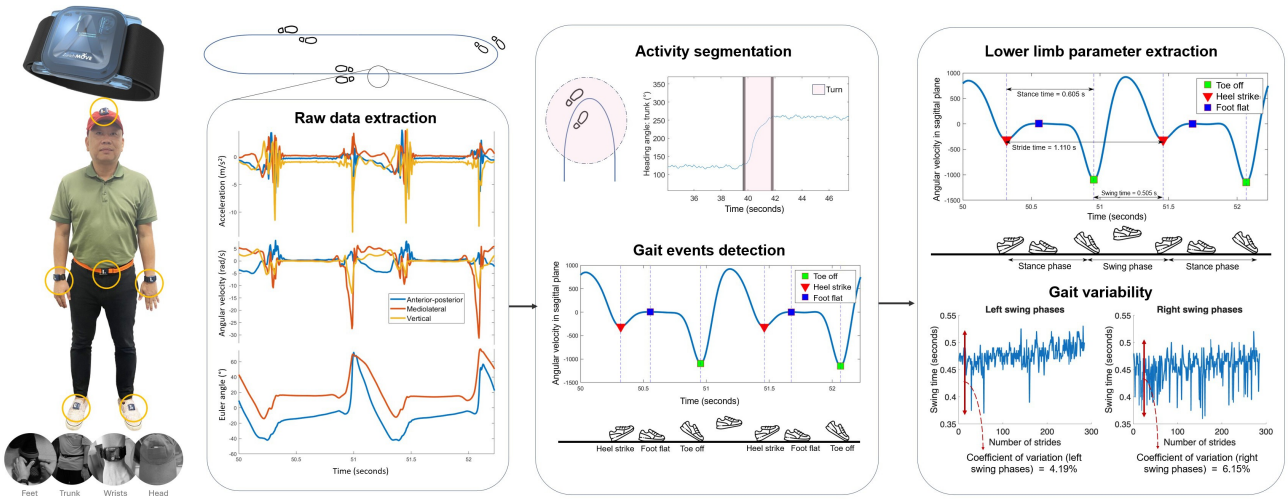
Baseline Interview

A face-to-face computer-assisted personal interview was conducted with each participant at their residence or other preferred place of convenience using a structured questionnaire to capture sociodemographic, environmental, physical, sensory, medical, cognitive, and psychosocial characteristics related to fall and fracture risk at baseline. Anthropometric measurements (height, weight, and bilateral hand grip strength) were also taken. Responses were collected using the survey software Qualtrics (Qualtrics). The survey was designed by TARGET study investigators and administered by trained field interviewers in the participant's preferred language (English, Mandarin or Chinese dialect, Malay, or Tamil). The baseline interview took approximately 2 hours to complete, including the screener, consent-taking, and gait assessment detailed in the following text.

Gait Assessment

As a component of the home-based interview, all participants performed a 5-minute walk inside or outside their home, in an open physical space, with the freedom to walk at their preferred pace and navigate turns, while wearing 6 ZurichMOVE inertial measurement units (IMUs) motion sensors [65] affixed to designated anatomical locations (feet, wrists, trunk, and head; Figure 2).

Figure 2. Participants were requested to wear 6 inertial measurement unit sensors (2 each on the feet and wrists, as well as 1 on the trunk and head) while performing a 5-minute walk test under natural settings. The data from the sensors are processed to provide information on turning, gait events, and variability and asymmetry of foot placement while walking.



There has been growing research and clinical interest in the use of wearable sensors for identifying fall risk through biomechanical data, specifically, dynamic gait signatures [37,66,67]. Gait and balance assessments are a core component of fall risk assessments in clinical practice. However, existing gold standard assessments such as the Timed Up and Go test and Berg Balance Scale have shown poor predictive accuracy (pooled sensitivity ranging from 0.25 to 0.32) in predicting prospective falls [68,69]. As the majority of falls occur while walking, dynamic gait characteristics such as step length, step time, variability, and asymmetry need to be factored into fall risk assessments for accurate risk prediction [66,70]. Wearable sensor systems can objectively measure a wide range of dynamic gait characteristics, which can then be used to train predictive models for fall risk assessment [71]. IMU sensors, which acquire data using accelerometers and gyroscopes, have gained research and clinical interest as a possible alternative to the current gold standard optical motion capture systems for gait analysis [72]. Compared with optical motion capture, IMU sensors are compact, inexpensive, easier to operate, and can be combined with functional walking tests for gait assessment in the home environment [73].

The IMUs embedded within ZurichMOVE are time-synchronized and have a dynamic range of ±16 g for the accelerometer and 2000° per second for the gyroscope at a sampling frequency of 200 Hz and have been shown to reliably assess gait and identify gait differences in healthy and neurological populations across laboratory settings and real-world environments [74,75]. Triaxial acceleration and angular velocity signals recorded during the walk test are first filtered and analyzed with the validated gait event detection

XSL•FO
RenderX

algorithm [70,71], identifying key phases in the gait cycles over the 5 minutes. From these phases, key spatiotemporal parameters such as stride time, length, cadence, and double-limb support time were computed for every participant. Furthermore, intercycle variations are evaluated during the entire walking trial to capture fluctuations and variability in movement patterns, indices associated with deficits in gait, balance, and risk of falling [37,76,77].

Fracture Risk Assessment

After the home-based interview, participants who reported that they did not have bilateral hip replacement were invited to a clinic visit at the Singapore Eye Research Institute for a dual-energy X-ray absorptiometry (DXA) scan. Each participant underwent 3 DXA scans (hip, lumbar spine, and whole-body) within a 20 - to 30-minute session using the Hologic Horizon W scanner (Hologic, Inc). The hip and lumbar spine scans were used to assess areal bone mineral density (aBMD), which was then used to calculate the T-score to evaluate osteoporotic status. Participants with a T-score of ≤−2.5 were classified as osteoporotic, those with a T-score between −2.5 and −1 were classified as osteopenic, and those with a T-score of >−1 were classified as normal, while the whole-body DXA scan captured clinical anthropometric data and body composition, including measurements of fat mass, lean mass, and total mass across the trunk, limbs, and head. These data were used to categorize participants into the low muscle mass group, defined as appendicular lean mass/height^2 of <7 kg/m^2 for males and <5.4 kg/m^2 for females [78]. Additionally, a subset of participants completed 3D-optical body shape scans using the TC2-19M scanner (TC2 Labs LLC), wearing tight-fitting clothing (Spandex top and pants). Participants held an A-line stance for approximately 9 seconds during scanning, which generated a triangular mesh representing their body shape. Data from both the whole-body DXA and 3D optical scans were used to estimate trochanteric soft tissue thickness (TSTT) [79]. TSTT refers to the thickness of the soft tissues, including skin, fat, and muscle, over the greater trochanter of the femur. Clinical risk factors were also gathered, and the risk for major osteoporotic fractures (MOFs) and hip fractures was assessed using the Fracture Risk Assessment Tool (FRAX) without aBMD for the entire cohort. For those who had a DXA scan, FRAX was calculated with the aBMD from the femoral neck. A high risk for MOF was assigned if FRAX-MOF ≥9%, and a high risk for hip fractures was assigned if the FRAX–Hip Fracture (FRAX-HF) score is ≥2% [80]. Age- and sex-adjusted analyses were conducted for all variables, except FRAX-related variables, as age and sex are already incorporated into the FRAX calculation algorithm.

Despite widespread use, studies have shown that aBMD measurements from DXA scans and the 10-year fracture probabilities of hip or MOFs computed using FRAX [61] often fail to detect individuals at risk, as only between 28% and 61% of hip fractures occur in individuals with osteoporosis [62,81,82]. Although FRAX incorporates 12 risk factors and provides country- and ethnicity-specific scores, research suggests that FRAX may be less accurate in Asian populations, compared with Caucasians [63,64]. Finite element models (FEMs) have demonstrated strong potential for improving fracture risk prediction, especially given the limited sensitivity of current standard-of-care methods. Subject-specific FEMs based on computed tomographic scans provide direct estimates of femoral strength and have outperformed aBMD in predicting future fractures, with 1 recent study demonstrating this superiority over a follow-up period of up to 16 years [83-85]. Emerging research from Caucasian cohorts also suggests that FEM-based strength estimates derived from DXA outperform aBMD alone in identifying individuals with hip fractures [86,87]. This body of evidence paves the way for FEM-based fracture risk assessment, particularly when adapted for use with clinical DXA scans for primary screening and with tomographic scans for opportunistic screening, as a more accurate alternative to current methods.

Virtual Reality Spatial Navigation Assessment

Participants who completed the DXA scan were also invited to play a virtual reality (VR) spatial navigation game during the clinic visit to assess their cognitive ability. The SPACE (Spatial Performance Assessment for Cognitive Evaluation) is a novel iPad serious game designed to identify deficits in spatial navigation indicative of cognitive impairment and potential fall risk [88]. SPACE has been validated with participants across community and clinical cohorts in Singapore. Studies revealed that SPACE is well suited for older adults [88], and its performance metrics correlate with scores from the Montreal Cognitive Assessment (MoCA); predict hippocampal atrophy beyond age, education, and standard neuropsychological tests [89]; and distinguish between dementia severity levels with high accuracy [90].

Studies have highlighted the potential application of VR technology for identifying markers of cognitive impairment and fall risk [88]. While analysis of cerebrospinal fluid and positron emission tomography can reliably detect dementia-related pathophysiological changes, these methods are invasive and costly, which limits their use for population-based screening [91]. Blood biomarkers are a promising alternative, but these are still undergoing clinical validation and are not yet widely available for routine diagnostic use [92]. In contrast, neuropsychological tests, such as the Mini Mental State Examination and MoCA, are relatively easier to deploy but have shown limited sensitivity to detect early cognitive impairment [93,94]. Critically, most cognitive assessments do not include tests that assess deficits in spatial navigation ability. This lacuna is surprising since spatial disorientation is one of the earliest signs of declining cognitive ability. Indeed, the hippocampus and the entorhinal cortex, which play an instrumental role in navigation, are also among the earliest brain regions to be affected by Alzheimer disease pathophysiology [95-98]. VR technology now offers an accurate, cost-effective, and safe method to assess navigation deficits indicative of cognitive impairment and fall risk in a quick and portable manner [88,99].

In SPACE, participants assume the role of an astronaut tasked with exploring an unfamiliar planet by completing a series of spatial navigation tasks with the help of a companion robot. Each task is specifically designed to probe different aspects of their spatial navigation abilities. SPACE includes a detailed training phase in which players can familiarize themselves with

XSL•FO
RenderX

the controls while being assessed on basic visuospatial and motor skills. After the initial training phase, participants perform 5 spatial tasks (ie, path integration, egocentric pointing, mapping, associative memory, and perspective taking) designed to recruit critical brain regions involved in spatial navigation. During gameplay, both performance data and metadata (eg, position and orientation, time, coordinates, and screen interaction) are stored and used to produce individual scores for the various tasks [88].

Prospective Follow-Up

Prospective follow-up of the TARGET cohort is currently ongoing. The cohort is being followed up via telephone every 3 months from the baseline interview for 2 years to determine (1) whether they have experienced a fall or fall-related injury over the past quarter, and (2) their cognitive status every 6 months measured using an abbreviated version of the MoCA designed for telephone administration (T-MoCA [51]). Interviewers may also visit participants to conduct the follow-up assessments in person if the participant has difficulties communicating clearly over the phone. As of May 24, 2026, T_6 (18-month follow-up) is complete for all participants, with 80.5% (1844/2291) successfully followed up (48/2291, 2.1% deaths, and 399/2291, 17.4% lost to follow-up), and 55.3% (1267/2291) of participants have completed T_8 (24-month follow-up). The study protocol allows for participants with missing follow-up data to be recontacted at subsequent follow-ups to maximize data completeness and cohort retention. Prospective outcomes will be validated against national EMR data. These data will provide granular diagnostic, procedural, and laboratory data on falls, fractures, and broader age-related diseases for the cohort.

EMR Linkage

Finally, EMRs for each TARGET cohort member who provided consent for record linkage at baseline (1243/2291, 54.3%) are being obtained through Singapore Ministry of Health's TRUST platform [49]. TRUST is a secure data exchange platform that enables the fusion and sharing of anonymized health-related research and real-world data for research purposes between public institutions and the public and private sectors. Linked datasets to be requested for the TARGET cohort will include health records (from acute hospitals, polyclinics, general practice clinics, and medical centers), health (emergency department presentations), and birth and death registry (mortality). EMR data will be used to analyze longitudinal pathology, morbidity, and mortality patterns associated with falls and broader age-related diseases in older adults. Analysis will be performed on the deidentified fused datasets in the secure TRUST Workspace.

Table S1 in Multimedia Appendix 1 compares key characteristics between participants who completed and did not complete the clinic visit components (DXA and 3D body scans and SPACE game), and participants who consented and did not consent to the EMR linkage to evaluate potential selection bias. Overall, participants who did not complete the clinic visit components were more likely to be older and have greater ADL/IADL difficulty.

Quality Assurance and Control

All researchers were required to complete the Collaborative Institutional Training Initiative online program for research ethics and compliance education. Fieldwork interviewers, including university interviewers and interviewers from a survey research company, were required to undergo training and fieldwork shadowing and to demonstrate competency in the relevant procedures and questionnaire administration prior to being approved to perform household recruitment with participants. Study documents, including participant information sheets, informed consent forms, and questionnaires, were translated into Chinese, Malay, and Tamil languages by a certified translation company. The study team validated the translated documents to ensure that the meaning is consistent with the original version. A computer-assisted personal interview questionnaire was designed in Qualtrics using logic and rules to assist fieldwork interviewers to conduct proper inclusion and exclusion criteria eligibility checks. A pilot study (n=18) was undertaken to test recruitment processes and ensure that interviewers were familiar with the recruitment and interviewing procedures. Data collected by all fieldwork interviewers were checked periodically by key study team personnel. Blood pressure machines and weighing scales were checked and certified by the supplier before being issued to the interviewers. All faulty study instruments are either serviced or replaced depending on the availability of spare parts to ensure measurement accuracy.

Study Outcomes and Planned Analysis

Primary Outcomes

The primary outcomes of interest are the incidence of falls and fractures over the 2-year prospective follow-up period. Falls are defined following the World Health Organization's definition as "an event which results in a person coming to rest inadvertently on the ground, floor, or lower level" [100]. Fracture events will include both self-reported and electronically validated incident fractures identified during follow-up and through linkage with EMR data. MOF will include fractures of the hip, clinical spine, wrist, and humerus, consistent with the FRAX definition. Hip fractures will be analyzed separately due to their substantial morbidity, mortality, and health care burden among older adults.

Prospective outcomes collected through quarterly follow-up interviews will be validated against linked EMR using relevant diagnostic and procedural coding systems, including ICD (International Classification of Diseases) codes, available within the TRUST-linked datasets. Time-to-event analyses will be conducted to evaluate predictors of incident falls and fractures over the follow-up period.

Risk Prediction and Stratification

Primary analyses will focus on the development and validation of novel personalized risk prediction models integrating variables from the epidemiological survey, gait assessment, clinical imaging, SPACE, and EMR data to segment the older adult population into different fall and fracture risk levels. Risk prediction models will be built using state-of-the-art machine learning, deep learning, and statistical techniques. The

XSL•FO
RenderX

performance of the novel risk prediction models will be evaluated against existing gold standard assessments (eg, STEADI algorithm, Three Key Questions, and FRAX) and validated against the prospective data. Secondary analyses will focus on (1) developing predictive models of fracture risks among older adults based on holistic body composition parameters obtained from DXA and whole-body 3D shape scans, (2) assessing the performance of SPACE to detect cognitive impairment and its predictive usefulness in detecting falls, and (3) assessing the psychological and sociodemographic correlates of fall risk.

Economic Analysis

Cost-effectiveness analyses of screening strategies will be evaluated using decision tree models. The analyses will use Markov models to simulate the long-term costs and quality-adjusted life years over the lifespan. Prospective epidemiological data collected from TARGET will be used as input parameters for event rates and diagnostic accuracy. Performance of novel risk prediction tools (ie, sensitivity, specificity, and accuracy) will be evaluated against current methods of evaluating falls and fracture risk to determine the incremental cost-effectiveness ratios of different screening strategies. A budget impact analysis will determine the short- and medium-term fiscal implications of a national implementation of these strategies.

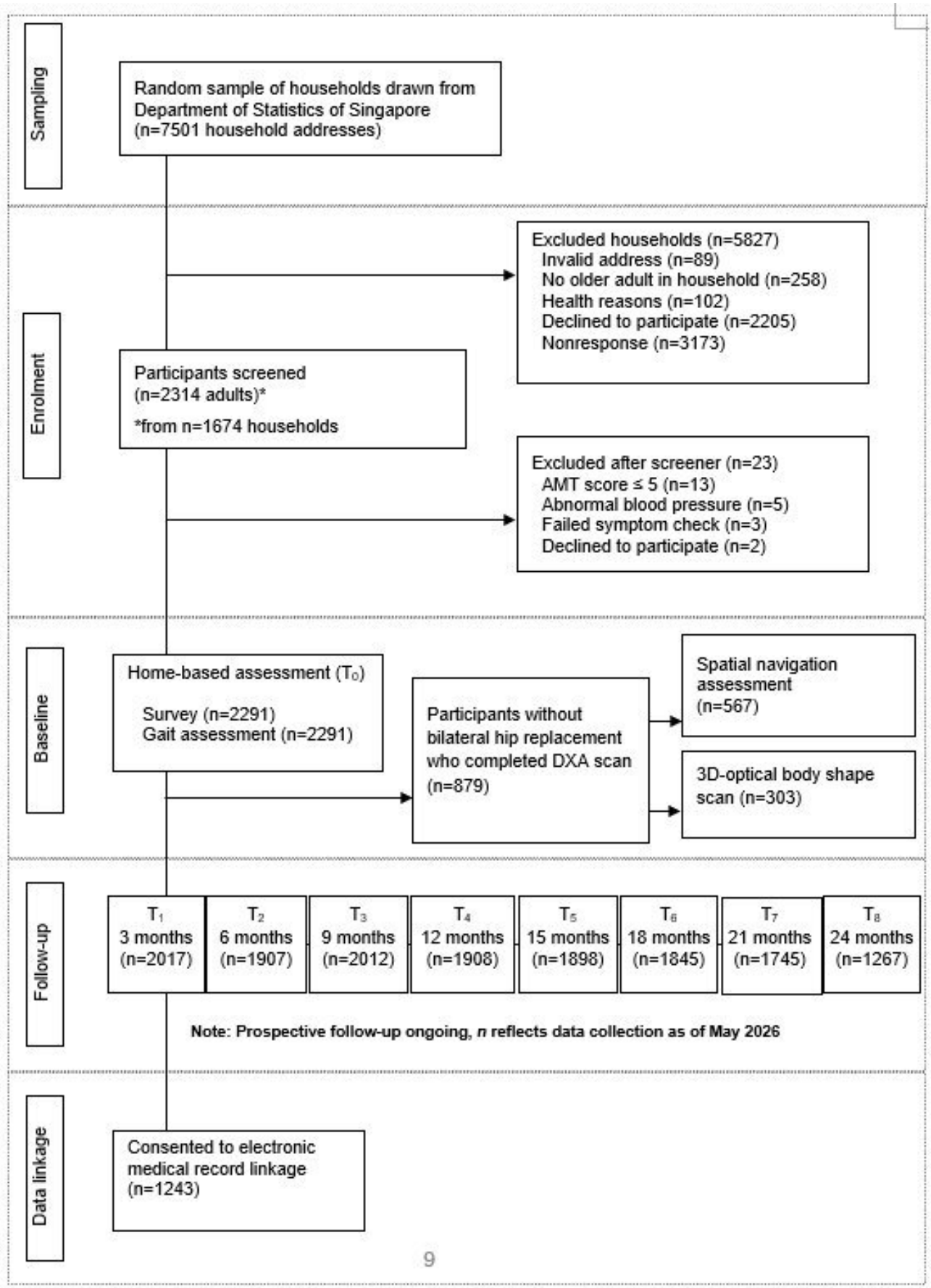
Results

Cohort Characteristics

Baseline data collection was completed in October 2024. Baseline cohort characteristics and key preliminary findings to date are summarized in this section. Figure 3 summarizes the TARGET study design and participant flow. Of the 2291 older adults in the TARGET cohort, the majority were female (327/2291, 57.9%) with a mean age of 74.6 (SD 7.9) years at study enrollment (Table 2). Most participants were Chinese (1521/2291, 66.4%), followed by Malay (433/2291, 18.9%), Indian (290/2291, 12.7%), and Other ethnicity (47/2291, 2.1%). Compared with national population reference data from 2023, representing the midpoint of the baseline data collection, the TARGET cohort had a higher proportion of older adults aged ≥75 years, females, and ethnic minorities, and participants from the Northeast region.

XSL•FO
RenderX

Figure 3. Cohort design and participant flow for the TARGET (Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments) study as of May 2026. AMT: Abbreviated Mental Test; DXA: dual-energy X-ray absorptiometry.



XSL•FO
RenderX

Table . Key sociodemographic characteristics of the TARGET[a] cohort (N=2291). National data for employment status and education status were not available.

Characteristics	Values, n	Proportion	National proportion (2023)[b]
Age (years)			
60 - 64	302	13.2	29.0
65 - 69	388	16.9	25.6
70 - 74	362	15.8	19.6
75 - 79	614	26.8	12.0
80 - 84	380	16.6	7.5
85+	245	10.7	6.3
Sex			
Male	964	42.1	47.2
Female	1327	57.9	52.8
Ethnicity			
Chinese	1521	66.4	81.3
Malay	433	18.9	10.5
Indian	290	12.7	6.5
Others	47	2.1	1.6
Geographic region in Singapore			
Central	616	26.9	26.5
East	484	21.1	18.2
North	284	12.4	11.9
Northeast	602	26.3	21.7
West	305	13.3	21.7
Highest education completed			
No formal education	367	16.0	N/A[c]
Primary school	741	32.3	N/A
Secondary school	757	33.0	N/A
Above secondary school	426	18.6	N/A
Employment status			
Full-time	285	12.4	N/A
Part-time	227	9.9	N/A
Not currently working	1582	69.1	N/A
Never worked	197	8.6	N/A
Housing type			
1- or 2-room government-built flat	258	11.3	7.5
3-room government-built flat	570	24.9	18.4
4- or 5-room government-built flat	1355	59.1	55.0
Private housing	108	4.7	19.1
Marital status			
Married	1464	63.9	70.8
Widowed	531	23.2	15.6

XSL•FO
RenderX

Characteristics	Values, n	Proportion	National proportion (2023)[b]
Divorced or separated	115	5.1	5.7
Never married	180	7.9	7.9

[a]Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments.

[b]From Department of Statistics Singapore 2023 (representing the midpoint of baseline data collection).

[c]N/A: not applicable.

Over 80% (1898/2291) of the cohort were fully independent at baseline (Table 3). Most participants had mild cognitive impairment (1016/2290, 44.4%; based on MoCA score 18‐25, which represents a level of cognitive impairment greater than expected for age without significant functional impairment) [101], or no cognitive impairment (926/2290, 40.4%; MoCA score ≥26), and 70% (1610/2291) had 2 or more chronic health conditions.

XSL•FO
RenderX

Table . Summary of select fall and fracture risk factors at baseline of the TARGET[a] cohort participants (N=2291).

Characteristics	N with data	Values, n	Percentage
Fall history in past 12 months	2291		
At least 1 fall		358	15.6
Recurrent falls (≥2 falls)		104	4.5
Fall-related injury		248	10.8
Hip replacement	2290	48	2.2
Left		22	1.0
Right		22	1.0
Bilateral		4	0.2
Fall risk score (FRQ[b])	2290		
High risk (≥4)		620	27.1
Low risk (0 - 3)		1670	72.9
Cognitive impairment (MoCA[c])	2290		
None (≥26)		926	40.4
Mild (18-25)		1016	44.4
Moderate (10-17)		340	14.8
Severe (0 - 9)		8	0.4
Walking aid used	2288		
Yes		405	17.7
Chronic conditions	2291		
None		302	13.2
≥1		1989	86.8
≥2		1610	70.3
≥3		1096	47.8
≥4		591	25.8
≥5		289	12.6
Functional independence (MBI[d])	2291		
Independent (100)		1898	82.9
Slight dependence (91-99)		102	4.5
Moderate dependence (61-90)		257	11.2
Severe dependence (21-60)		32	1.4
Total dependence (≤20)		2	0.1
Fear of falling (ICONFES[e])	2287		
Low concern (10-18)		1462	63.9
High concern (19-40)		825	36.1
Regular prescription medications	2283		
None		441	19.2
1 - 4		1276	55.9
≥5 (polypharmacy)		566	24.8

[a]Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments.

[b]FRQ: Falls Risk Questionnaire.

[c]MoCA: Montreal Cognitive Assessment.

XSL•FO
RenderX

[d]MBI: Modified Barthel Index.

[e]ICONFES: Iconographical Falls Efficacy Scale.

Falls Epidemiology and Risk Profile at Baseline

Among the total cohort, 15.6% (358/2291) reported at least 1 fall within the prior 12-month period and 4.5% (104/2291) reported recurrent falls. This is comparable with previously reported fall incidence rates in community-dwelling older adult populations in Singapore [102] but lower than reported rates in the international literature (33%) [15]. Regional differences in falls incidence have previously been identified between Asian and Western countries owing to differences in sociodemographic, geographic, environmental, and cultural factors [103].

Psychosocial and Sociodemographic Correlates of 12-Month Fall History

Table 4 compares baseline characteristics for TARGET participants by 12-month fall history. Older adults who had fallen in the past 12 months demonstrated significantly poorer physical and psychosocial health across multiple measures at baseline than nonfallers, reinforcing the need for fall risk screening and early intervention in this population. Notably, fallers demonstrated significantly poorer functional independence and quality of life, greater depressive symptoms, higher number of chronic health conditions, and greater fear of falling than nonfallers. However, it is unclear whether preexisting poorer physical and psychosocial health contributed to falls or was a consequence of the fall sequelae. Structural equation modeling analyses are in progress to better understand the causal pathways between falls, functional independence and quality of life, and the potential mediating and moderating mechanisms that can be targeted through interventions. Predictive analyses using the follow-up data are also underway to determine risk factors of prospective falls and fractures in the cohort.

XSL•FO
RenderX

Table . Comparison of baseline characteristics of TARGET[a] cohort participants by fall status in past 12 months (N=2291).

Characteristic, mean (SD)	N with data	Faller, n=358	Nonfaller, n=1933	*P* value[b]
Age (years)	2291	75.7 (7.9)	74.3 (7.8)	.02[b]
Anthropometrics				
BMI (kg/m^2)	2290	24.3 (4.7)	24.6 (4.7)	.26
Blood pressure: systolic	2291	134.5 (18.2)	133.5 (17.2)	.02[b]
Blood pressure: diastolic	2291	75.6 (9.2)	75.8 (9.0)	.47
Heart rate (beats per minute)	2291	75.5 (11.1)	74.8 (10.1)	.27
Left hand grip strength, kg	2245	16.1 (6.0)	18.1 (6.4)	<.001[b]
Right hand grip strength, kg	2259	17.3 (6.3)	19.9 (18.6)	<.001[b]
Functional independence				
IADL[c]	2290	14.0 (3.2)	14.8 (2.5)	.01[b]
ADL[d]	2291	96.1 (9.9)	97.7 (7.5)	.008[b]
Cognitive impairment				
MoCA[e]	2290	22.9 (5.2)	23.3 (4.9)	.01[b]
Physical health				
Chronic conditions	2291	3.2 (2.0)	2.4 (1.7)	<.001[b]
Medications	2283	3.6 (3.1)	2.9 (2.7)	<.001[b]
MET[f]-minutes per week	2272	1374 (1580)	1742 (1717)	.008[b]
Psychosocial				
EQ5D[g]	2290	85.4 (17.8)	91.7 (14.2)	<.001[b]
PHQ-9[h]	2287	2.5 (2.9)	1.4 (2.4)	<.001[b]
Loneliness_Score	2285	3.9 (1.4)	3.6 (1.2)	.002[b]
LSNS-6[i]	2286	13.1 (5.8)	13.9 (5.6)	.001[b]
BRS[j]	2274	22.4 (3.5)	23.0 (2.9)	<.001[b]
Falls				
ICONFES[k]	2287	19.9 (7.2)	16.5 (6.8)	<.001[b]
FRQ[l]	2290	5.4 (2.8)	1.9 (2.3)	<.001[b]

[a]Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments.

[b]Significance tests evaluated using multilevel mixed-effects regression models adjusted for clustering at household level, $P<.05$.

[c]IADL: Instrumental Activities of Daily Living score ranging from 0 to 16 (fully independent).

[d]ADL: Activities of Daily Living score ranging from 0 to 100 (fully independent).

[e]MoCA: Montreal Cognitive Assessment score ranging from 0 to 30 (normal cognitive function).

[f]MET: metabolic estimates.

[g]EQ5D: EuroQol Index ranging from 1 to 100 (best health).

[h]PHQ-9: Patient Health Questionnaire–9 score ranging from 0 to 27 (severe depression).

[i]LSNS-6: Lubben Social Network Score ranging from 0 to 30 (strong social network).

[j]BRS: Brief Resilience Scale score ranging from 6 to 30 (high resilience).

[k]ICONFES: Iconographical Falls Efficacy Scale score ranging from 10 to 40 (greater fear of falling).

[l]FRQ: Fall Risk Questionnaire score ranging from 0 to 12 (high fall risk).

XSL•FO
RenderX

Gait Assessment

During the home-based interview, 86.4% (1980/2291) of TARGET participants independently completed the full 5-minute walking test, while the remaining 13.6% (311/2291) were able to walk between 3 and 5 minutes. Table 5 compares selected extracted baseline gait metrics between participants with a 12-month history of falls and those without. All participants were analyzed together without adjusting for walking duration as our prior research has shown that 3 minutes of walking covers more than 50 gait cycles (strides) and is sufficient for reliably measuring gait variability [104]. There were no significant differences in the mean and variability of temporal parameters (cadence, stride time, stance time, and double-limb support time) or the spatial measures (step length and gait speed) except for average step length. Among the signal-derived metrics, fallers showed significantly greater variability in the magnitude of foot acceleration (P=.02) and significantly lower angular velocity root-mean-square on the trunk sensor (P=.02). These preliminary findings suggest that higher gait variability and reduced movement intensity are likely to be associated with a history of falls. These gait parameters are being used for fall risk prediction modeling using machine learning and deep learning models, validated by prospective follow-up of fall status.

Table . Age- and sex-adjusted comparisons of gait parameters extracted from inertial measurement units sensors by fall status in past 12 months (N=2149).

Gait parameters, mean (SD), range	Faller (n=339)	Nonfaller (n=1810)	*P* value
Temporal parameters			
Mean			
Cadence (steps per minute)	85.7 (18.6), 24.4 - 139.5	85.6 (17.8), 23.1 - 145.2	.97
Stride time, seconds	1.5 (0.5), 0.9 - 4.9	1.5 (0.5), 0.8 - 5.2	.75
Stance time, seconds	1.0 (0.5), 0.5 - 4.2	1.0 (0.4), 0.5 - 4.5	.70
Double limb support time, seconds	0.2 (0.2), 0.1 - 1.7	0.2 (0.2), 0.0 - 1.9	.64
Variability (coefficient of variation)			
Stride time, %	10.6 (6.1), 2.3 - 49.1	9.9 (5.7), 2.1 - 55.0	.27
Stance time, %	14.8 (8.4), 3.8 - 58.9	13.9 (8.2), 3.0 - 72.5	.37
Step time, %	14.2 (8.4), 3.9 - 71.4	13.2 (7.8), 2.7 - 60.2	.14
Double limb support time, %	37.0 (19.5), 10.0 - 143.9	35.2 (20.3), 9.5 - 211.9	.35
Spatial parameters			
Mean			
Step length, meters	0.3 (0.2), 0.1 - 0.6	0.3 (0.2), 0.0 - 0.9	.01[a]
Gait speed, meter per second	0.5 (0.2), 0.0 - 1.3	0.5 (0.2), 0.0 - 1.3	.30
Variability (coefficient of variation: %RSD[b])			
Step length	25.4 (10.4), 6.5 - 62.9	25.7 (12.2), 4.2 - 175.1	.64
Gait speed	22.9 (8.8), 6.6 - 55.7	22.2 (9.5), 4.1 - 110.2	.43
Signal parameters			
Variability (coefficient of variation: %RSD)			
SD of acceleration magnitude, left foot	29.5 (8.3), 9.9 - 62.7	28.3 (7.7) [9.7 - 83.8]	.02[a]
Root-mean-square of angular velocity magnitude, trunk	36.2 (12.9), 9.4 - 135.0	38.1 (29.0), 9.7 - 1121.0	.01[a]

[a]Values indicate significant differences (P<.05).

[b]RSD: root SD.

BMD, Body Composition, Morphometry, and FRAX Assessment

Among the 38.4% (879/2291) participants (n=6 were removed for scan artifacts) who underwent DXA scans, 54.9% (483/879) were female. Data from 4 participants were excluded due to lumbar spine scans with motion artifacts. Table 6 summarizes the baseline prevalence of osteoporosis, osteopenia, low muscle mass, and high fracture risk by DXA and FRAX, stratified by gender and ethnicity. A higher percentage of females had osteoporosis than males. However, fewer females were classified as osteopenic than males, and fewer females had low muscle mass than males. Chinese participants had lowest TSTT across both genders compared with Indians and Malays [79]. A higher

XSL•FO
RenderX

percentage of females exceeded the FRAX-MOF threshold with aBMD and FRAX-HF threshold with aBMD than males. Similarly, a greater proportion of females surpassed the threshold for FRAX-MOF without aBMD and FRAX-HF without aBMD than males. When comparing ethnicities, the prevalence of osteoporosis, osteopenia, and low muscle mass was the highest in Chinese participants, regardless of gender, compared with their Indian and Malay counterparts [79].

Table . Prevalence of osteoporosis, osteopenia, low muscle mass, and high FRAX[a] with aBMD[b] score (n=879) and FRAX without aBMD score (n=2285), stratified by gender and ethnicity.

Characteristic	Ethnicity			
	Pooled[c]	Chinese	Indian	Malay
Male				
With DXA[d] scan, n	396	275	49	67
Osteoporotic, %	10.6	11.3	8.2	10.4
Osteopenic, %	51.5	55.6	30.6	50.7
Low muscle mass, %	50.3	53.8	49.0	35.8
High FRAX-MOF[e], %[f]	21.9	28.4	6.12	5.9
High FRAX-HF[g], %[f]	67.2	73.8	59.2	44.8
FRAX-related data, n	963	640	118	195
High FRAX-MOF, %[h]	23.9	32.2	10.2	4.6
High FRAX-HF, %[h]	72.6	79.7	65.3	53.3
Female				
With DXA scan, n	483	337	66	78
Osteoporotic, %	13.0	16.0	4.5	7.7
Osteopenic, %	48.9	51.9	36.4	46.2
Low muscle mass, %	39.8	45.7	28.8	24.4
High FRAX-MOF, %[f]	61.7	71.2	50.0	30.8
High FRAX-HF, %[f]	70.4	78.0	59.1	47.4
FRAX-related data, n	1322	880	180	257
High FRAX-MOF, %[h]	71.4	81.1	57.8	47.8
High FRAX-HF, %[h]	79.2	86.4	70.6	60.7

[a]FRAX: Fracture risk assessment tool.

[b]aBMD: areal bone mineral density.

[c]Pooled includes other ethnicities.

[d]DXA: dual-energy X-ray absorptiometry.

[e]MOF: major osteoporotic fractures.

[f]Computed with areal bone mineral density.

[g]HF: hip fractures.

[h]Computed without areal bone mineral density.

Table 7 compares baseline aBMD, body composition, morphometry, and FRAX scores by fall status in the past 12 months. TSTT was calculated for the subcohort of 260 participants who had 3D-optical scans (39 fallers and 221 nonfallers, excluding 43 scans with motion artifacts); no significant difference in TSTT was found between fallers and nonfallers. Following age and sex adjustment, the relationships between fallers and nonfallers remained nonsignificant. The mean FRAX scores for both fallers and nonfallers were significantly above the respective clinical thresholds of 9% for FRAX-MOF and 2% for FRAX-HF. Specifically, fallers had significantly higher FRAX-MOF with aBMD scores (mean 13.16, SD 8.34%) than nonfallers (mean 9.97, SD 6.57%), and higher FRAX-HF with aBMD scores (mean 6.11, SD 5.21% vs mean 4.44, SD 4.43%; $P<.001$ for both). Similar patterns were observed without aBMD, with fallers showing higher FRAX-MOF without aBMD scores (mean 14.27, SD 8.58%) and FRAX-HF without aBMD scores (mean 7.17, SD 6.12%) than nonfallers (mean 10.81, SD 6.72% and mean 5.12, SD 4.61%, respectively; $P<.001$).

XSL•FO
RenderX

Table . Age- and sex-adjusted comparisons of DXA[a] scanner– and 3D-optical scanner–derived variables, FRAX[b] with aBMD[c] scores (n=879), and FRAX without aBMD (n=2285), by fall status in the past 12 months.

Variable	Fallers		Nonfallers		*P* value
	Values, n	Mean (SD)	Values, n	Mean (SD)	
aBMD (total hip), g/cm^2	145	0.8 (0.2)	734	0.8 (0.2)	.76
aBMD (total hip), g/cm^2[d]	145	0.8 (0.1)	734	0.8 (0.1)	.21
aBMD (femoral neck), g/cm^2	145	7 (0.2)	734	0.7 (0.1)	.47
aBMD (femoral neck), g/cm^2[d]	145	7 (0.1)	734	0.7 (0.1)	.44
aBMD (lumbar spine), g/cm^2	145	0.9 (0.2)	730	1.0 (0.2)	.79
aBMD (lumbar spine), g/cm^2[d]	145	1.0 (0.2)	730	1.0 (0.2)	.33
T-score (total hip)	145	-1.2 (1.2)	734	−1.3 (1.1)	.52
T-score (total hip)[d]	145	-1.2 (1.1)	734	−1.3 (1.1)	.28
Appendicular fat mass, kg	145	10.5 (3.5)	734	10.3 (3.8)	.37
Appendicular fat mass, kg[d]	145	10.3 (3.4)	734	10.4 (3.4)	.81
Appendicular lean mass, kg	145	15.6 (3.8)	734	16.2 (3.9)	.07
Appendicular lean mass, kg[d]	145	16.3 (2.4)	734	16.1 (2.4)	.39
Whole-body total fat mass, kg	145	23.5 (7.1)	734	23.3 (7.1)	.95
Whole-body total fat mass, kg[d]	145	23.2 (6.8)	734	23.4 (6.8)	.82
Whole-body total lean mass, kg	145	37.1 (7.6)	734	38.4 (7.8)	.06
Whole-body total lean mass, kg[d]	145	38.5 (4.9)	734	38.1 (4.9)	.47
Muscle mass, kg/m^2	145	6.1 (1.0)	734	6.3 (1.1)	.09
Muscle mass, kg/m^2[d]	145	6.3 (0.9)	734	6.2 (0.9)	.87
TSTT[e], cm	39	3.4 (1.1)	221	3.5 (1.3)	.85
TSTT, cm[d]	39	3.9 (1.1)	221	3.5 (1.1)	.55
FRAX[f]-MOF[g], %[h]	145	13.2 (8.3)	734	10.0 (6.6)	*<.001*[k]
FRAX-HF[i], %[h]	145	6.1 (5.2)	734	4.4 (4.4)	*<.001*[k]
FRAX-MOF, %[j]	357	14.3 (8.58)	1928	10.8 (6.7)	*<.001*[k]
FRAX-HF, %[j]	357	7.2 (6.1)	1928	5.1 (4.6)	*<.001*[k]

[a]DXA: dual energy X-ray absorptiometry.

[b]FRAX: fracture risk assessment tool.

[c]aBMD: areal bone mineral density.

[d]Adjusted for age and sex.

XSL•FO
RenderX

[e]TSTT: trochanteric soft tissue thickness.

[f]FRAX: fracture risk assessment tool.

[g]MOF: major osteoporotic fractures.

[h]Computed with bone mineral density.

[i]HF: hip fractures.

[j]Computed without areal bone mineral density; the Mann-Whitney *U* test was used to compare the distributions of key variables between fallers and nonfallers.

[k]Values in italics indicate significant differences (*P*<.05).

With the body shapes from the 3D-optical scans used to develop a soft tissue database for Asian adults across genders and ethnicities, DXA-derived FEMs will be enhanced by incorporating this soft tissue information to more accurately predict fracture risk in a patient-specific manner. These models will be prospectively validated against fracture outcomes obtained from electronic health records. Through this validation, we aim to determine optimal thresholds for identifying high-risk individuals. Establishing these thresholds will, in turn, inform screening protocols, ensuring that they are cost-effective and scalable for implementation within Singapore's health care system.

SPACE

Of 2291 participants, the subset of 567 (24.7%) participants who completed SPACE had a higher average MoCA score (mean 25.9, SD 3.6), scoring 2.8 points above the remainder of the cohort. All tasks in SPACE significantly correlated with MoCA scores. Robust comparisons using the M-estimator median method revealed no significant differences between fallers and nonfallers for all tasks in SPACE; however, a large group imbalance was noted (Table 8). Planned analyses using the TARGET cohort will focus on evaluating the combined predictive usefulness of SPACE and gait metrics in detecting falls and fractures. SPACE has been shown to predict hippocampal atrophy beyond age, education, and neuropsychological assessment [89], and can accurately distinguish between dementia severity levels (area under the curve=0.94 - 0.95) [90].

Table . Age- and sex-adjusted comparisons of SPACE[a] performance by fall status in the past 12 months (N=567).

SPACE tasks, median (n)	Faller	Nonfaller	*P* value
Training time	271.2 (98)	269.7 (468)	.97
Path integration	207.5 (80)	205.7 (393)	.82
Pointing	79.9 (66)	77.3 (340)	.84
Mapping	0.7 (66)	0.5 (339)	*.03* [b]
Memory	100.0 (66)	100.0 (339)	.98
Perspective taking	45.7 (84)	49.4 (447)	.42

[a]SPACE: Spatial Performance Assessment for Cognitive Evaluation.

[b]Value in italics indicates significant differences (*P*<.05).

Discussion

Principal Findings

To the best of our knowledge, TARGET is the largest prospective cohort study in Asia to jointly investigate fall and fracture risk factors among a national population of community-dwelling older adults. Overall, the TARGET cohort represents a relatively healthy community-dwelling older adult population, with over 80% retaining full functional independence at baseline. However, there is some evidence of early or moderate age-related decline: 70% (1610/2291) of the cohort reported at least 2 chronic conditions, 59% (1356/2290) exhibited mild-moderate cognitive impairment, and among those who underwent DXA scans, 50% (440/879) had osteopenia. Falls were reported by 15% (358/2291) of participants in the past year, whereas 85% (1933/2291) remained fall-free. Thus, TARGET is an ideal cohort for investigating the risk factors that predispose relatively healthy older adults to transition to first-time falls and fractures. Through EMR linkage, this can be extended to study the onset of broader age-related diseases.

By leveraging novel health technology data, TARGET aims to identify novel patterns in gait, cognition, and skeletal fragility as key markers of fall and fracture risk, and potentially other age-related diseases with shared pathophysiology. Data from the wearable IMU sensors capture subtle deficits in balance, stride variability, and reaction times, which cannot be detected through clinical observation alone, providing ecologically valid measurements of real-world performance where traditional clinical thresholds of 0.8 meters per second [15] or 1.0 meter per second [78] may no longer be applicable. These data can be particularly informative for age-related diseases where gait and balance impairments play an early or significant role such as Parkinson disease, stroke, multiple sclerosis, and peripheral neuropathy [105]. SPACE tasks probe spatial orientation that is often overlooked by traditional cognitive tests, with important implications for the early detection of falls as well as dementia [88,89]. The DXA-derived FEM-based models may also identify

XSL•FO
RenderX

more subtle skeletal vulnerabilities that might be missed by standard T-score classifications [83,86,87], offering potential insights into other skeletal conditions such as degenerative joint disease or metabolic bone disorders.

These novel health technology data are strengthened by clinical and laboratory data from EMR linkage, which add essential pathological and biochemical context, allowing for correlations between physiological changes and real-world health outcomes. Finally, robust epidemiological survey data capturing sociodemographic, environmental, psychosocial, and functional parameters enable the identification of synergistic risk factors such as social isolation and psychological resilience that may affect fall and fracture risk and accelerate age-related decline. Collectively, these rich and novel cohort data place TARGET in a strong position to identify emerging biomarkers and systemic interactions that may predict falls and fracture risk and potential other broader age-related pathophysiology.

Critically, insights generated by the TARGET study are likely to make a significant contribution to the development of personalized community-based screening and intervention frameworks for Asian populations, where differences in genetics, lifestyle, health care systems, and environmental factors have been shown to influence both fall and fracture risk profiles and outcomes; yet, where there remains a critical gap in region-specific evidence and scalable screening solutions. While the fracture risk component of TARGET relies on specialized equipment such as DXA scans and 3D body scanning, the fall risk component uses portable, low-cost technologies (eg, IMU sensors) that allow trained nonclinical personnel to conduct on-site assessments. This expanded approach has the potential to reduce substantial burden and costs within the health care sector, with whom the responsibility of health screening traditionally lies, by broadening access in community settings.

Limitations

The TARGET cohort is not a fully representative national sample of community-dwelling older adults in Singapore. Specifically, TARGET contains a higher proportion of individuals aged 75‐84 years and of Malay and Indian backgrounds, reflecting the intentional oversampling of these minority groups to ensure statistical power for subgroup analyses. Furthermore, it excludes certain older adult subgroups, such as those not ambulatory or with possible delirium or dementia. While this approach allows for more nuanced investigation of fall and fracture risk factors within these subpopulations, it may limit the generalizability of the findings to the broader older adult population in Singapore. Additionally, not all participants contributed data for every health technology assessed in the study, particularly for the DXA scan and VR assessment. As a result, some analyses will be based on subcohorts, which could affect the statistical power and introduce potential bias in technology-specific findings. Finally, it is acknowledged that our household response rate of 22.3% (1674/7501) is lower than expected compared with previous national surveys of older adults we have conducted in Singapore (35% response rate) [106]. This lower response rate may introduce selection bias, as individuals who chose to participate could differ systematically from those who declined, potentially in terms of health status, sociodemographic characteristics, or other relevant factors.

Acknowledgments

The research was conducted at the Centre for Ageing Research and Education, Duke-NUS Medical School, in partnership with the Future Health Technologies at the Singapore-ETH Centre, which was established collaboratively between ETH Zurich and the National Research Foundation Singapore. This research is supported by the National Research Foundation Singapore under its Campus for Research Excellence and Technological Enterprise program. The authors wish to thank Christina Chai, Thulase Thamootharam Pillai, Thu Khong, Dr Qiqi Cheng, and Dr Deepani Siriwardhana for their contributions to the baseline data collection and data cleaning. The authors declare the use of generative AI (GenAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: proofreading and editing, and recommendations. The GenAI tool used was ChatGPT 4.5. Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Authors AV and RS are currently not affiliated with any institution, but are Independent Researchers.

Funding

The research was conducted at the Future Health Technologies at the Singapore-ETH Centre, which was established collaboratively between ETH Zurich and the National Research Foundation Singapore. This research is supported by the National Research Foundation Singapore under its Campus for Research Excellence and Technological Enterprise program. AC, RM, KYT, LT, and CWHL acknowledge the funding from Singapore's Ministry of Health (MOH) to the Centre for Ageing Research & Education, Duke-NUS Medical School. Any opinions, findings and conclusions or recommendations expressed in this work are those of the authors and do not reflect the views of MOH.

Data Availability

The TARGET cohort data can be accessed for research collaboration. All collaboration enquiries should be submitted to the cohort lead and principal investigator, associate professor Angelique Chan (angelique.chan@duke-nus.edu.sg). If approved, the

XSL•FO
RenderX

requested data will be provided after the necessary data transfer agreements between institutions have been completed, where applicable.

Authors' Contributions

Conceptualization: AC, AV, RM, DBM, NBS, BH, VRS, WRT
Funding acquisition: AC, AV, RM, DBM, NBS, BH, VRS, WRT
Supervision: AC, AV, RM, KYT, NBS, BH, VRS, SJF
Investigation: KYT, RS, VJWK, WXL, CWHL, SGSP, KZT, DJ, ADP, GC, KM, ELL
Formal Analysis: LT, CWHL, KZT, DJ, and GC
Writing – original draft: LT
Writing – review & editing: All authors

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparison of baseline characteristics of TARGET (Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments) cohort participants by completion of DXA, 3D body scan, and SPACE (Spatial Performance Assessment for Cognitive Evaluation) assessments.
[DOCX File, 23 KB - aging_v9i1e89285_app1.docx]

References

1. James SL, Lucchesi LR, Bisignano C, et al. The global burden of falls: global, regional and national estimates of morbidity and mortality from the Global Burden of Disease Study 2017. Inj Prev 2020 Oct;26(Supp 1):i3-i11. [doi: 10.1136/injuryprev-2019-043286]
2. Hausdorff JM, Rios DA, Edelberg HK. Gait variability and fall risk in community-living older adults: a 1-year prospective study. Arch Phys Med Rehabil 2001 Aug;82(8):1050-1056. [doi: 10.1053/apmr.2001.24893]
3. Ganz DA, Latham NK. Prevention of falls in community-dwelling older adults. N Engl J Med 2020 Feb 20;382(8):734-743. [doi: 10.1056/NEJMcp1903252]
4. Allen NE, Schwarzel AK, Canning CG. Recurrent falls in Parkinson's disease: a systematic review. Parkinsons Dis 2013;2013(1):906274. [doi: 10.1155/2013/906274]
5. Batchelor F, Hill K, Mackintosh S, Said C. What works in falls prevention after stroke?: A systematic review and meta-analysis. Stroke 2010 Aug;41(8):1715-1722. [doi: 10.1161/STROKEAHA.109.570390]
6. Richardson JK, Ashton-Miller JA. Peripheral neuropathy: an often-overlooked cause of falls in the elderly. Postgrad Med 1996 Jun;99(6):161-172.
7. Yang Y, Hu X, Zhang Q, Zou R. Diabetes mellitus and risk of falls in older adults: a systematic review and meta-analysis. Age Ageing 2016 Nov;45(6):761-767. [doi: 10.1093/ageing/afw140]
8. Bourke R, Doody P, Pérez S, Moloney D, Lipsitz LA, Kenny RA. Cardiovascular disorders and falls among older adults: a systematic review and meta-analysis. J Gerontol A Biol Sci Med Sci 2024 Feb 1;79(2):glad221. [doi: 10.1093/gerona/glad221]
9. Overcash JA, Beckstead J. Predicting falls in older patients using components of a comprehensive geriatric assessment. Clin J Oncol Nurs 2008 Dec;12(6):941-949. [doi: 10.1188/08.CJON.941-949]
10. Tinetti ME, Williams CS. Falls, injuries due to falls, and the risk of admission to a nursing home. N Engl J Med 1997 Oct 30;337(18):1279-1284. [doi: 10.1056/NEJM199710303371806]
11. Kannus P, Parkkari J, Koskinen S, et al. Fall-induced injuries and deaths among older adults. JAMA 1999 May 26;281(20):1895-1899. [doi: 10.1001/jama.281.20.1895]
12. Prabhakaran K, Gogna S, Pee S, Samson DJ, Con J, Latifi R. Falling again? Falls in geriatric adults-risk factors and outcomes associated with recidivism. J Surg Res 2020 Mar;247:66-76. [doi: 10.1016/j.jss.2019.10.041]
13. Feng JN, Zhang CG, Li BH, Zhan SY, Wang SF, Song CL. Global burden of hip fracture: the global burden of disease study. Osteoporos Int 2024 Jan;35(1):41-52. [doi: 10.1007/s00198-023-06907-3]
14. Haentjens P, Autier P, Barette M, Boonen S, Belgian Hip Fracture Study Group. The economic cost of hip fractures among elderly women: A one-year, prospective, observational cohort study with matched-pair analysis. J Bone Joint Surg Am 2001 Apr;83(4):493-500. [doi: 10.2106/00004623-200104000-00003] [Medline: 11315777]
15. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):afac205. [doi: 10.1093/ageing/afac205]
16. Abrahamsen B, van Staa T, Ariely R, Olson M, Cooper C. Excess mortality following hip fracture: a systematic epidemiological review. Osteoporos Int 2009 Oct;20(10):1633-1650. [doi: 10.1007/s00198-009-0920-3]

XSL•FO
RenderX

17. Florence CS, Bergen G, Atherly A, Burns E, Stevens J, Drake C. Medical costs of fatal and nonfatal falls in older adults. J Am Geriatr Soc 2018 Apr;66(4):693-698. [doi: 10.1111/jgs.15304]
18. Rubenstein LZ, Josephson KR. Falls and their prevention in elderly people: what does the evidence show? Med Clin North Am 2006 Sep;90(5):807-824. [doi: 10.1016/j.mcna.2006.05.013]
19. Kirkland JL, Tchkonia T. Cellular senescence: a translational perspective. EBioMedicine 2017 Jul;21:21-28. [doi: 10.1016/j.ebiom.2017.04.013]
20. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 1;48(1):16-31. [doi: 10.1093/ageing/afy169]
21. Seidler RD, Bernard JA, Burutolu TB, et al. Motor control and aging: Links to age-related brain structural, functional, and biochemical effects. Neurosci Biobehav Rev 2010 Apr;34(5):721-733. [doi: 10.1016/j.neubiorev.2009.10.005] [Medline: 19850077]
22. Raz N, Lindenberger U, Rodrigue KM, et al. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. Cereb Cortex 2005 Nov;15(11):1676-1689. [doi: 10.1093/cercor/bhi044]
23. Ungvari Z, Tarantini S, Donato AJ, Galvan V, Csiszar A. Mechanisms of vascular aging. Circ Res 2018 Sep 14;123(7):849-867. [doi: 10.1161/CIRCRESAHA.118.311378]
24. Panel on Prevention of Falls in Older Persons, American Geriatrics Society and British Geriatrics Society. Summary of the updated American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls in older persons. J Am Geriatr Soc 2011 Jan;59(1):148-157. [doi: 10.1111/j.1532-5415.2010.03234.x]
25. Schaap LA, Pluijm SMF, Deeg DJH, et al. Higher inflammatory marker levels in older persons: associations with 5-year change in muscle mass and muscle strength. J Gerontol A Biol Sci Med Sci 2009 Nov;64(11):1183-1189. [doi: 10.1093/gerona/glp097]
26. Maggio M, De Vita F, Lauretani F, et al. IGF-1, the cross road of the nutritional, inflammatory and hormonal pathways to frailty. Nutrients 2013 Oct 21;5(10):4184-4205. [doi: 10.3390/nu5104184]
27. Gerdhem P, Ringsberg KAM, Akesson K, Obrant KJ. Clinical history and biologic age predicted falls better than objective functional tests. J Clin Epidemiol 2005 Mar;58(3):226-232. [doi: 10.1016/j.jclinepi.2004.06.013]
28. Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls. A prospective study. JAMA 1989 May 12;261(18):2663-2668. [doi: 10.1001/jama.1989.03420180087036]
29. Ganz DA, Bao Y, Shekelle PG, Rubenstein LZ. Will my patient fall? JAMA 2007 Jan 3;297(1):77-86. [doi: 10.1001/jama.297.1.77]
30. Eckstrom E, Vincenzo JL, Casey CM, et al. American Geriatrics Society response to the World Falls Guidelines. J Am Geriatr Soc 2024 Jun;72(6):1669-1686. [doi: 10.1111/jgs.18734]
31. Stevens JA, Phelan EA. Development of STEADI: a fall prevention resource for health care providers. Health Promot Pract 2013 Sep;14(5):706-714. [doi: 10.1177/1524839912463576] [Medline: 23159993]
32. Nithman RW, Vincenzo JL. How steady is the STEADI? Inferential analysis of the CDC fall risk toolkit. Arch Gerontol Geriatr 2019;83:185-194. [doi: 10.1016/j.archger.2019.02.018] [Medline: 31075677]
33. Montero-Odasso M, Pieruccini-Faria F, Son S, et al. Fall risk stratification in older adults: low and not-at-risk status still associated with falls and injuries. Age Ageing 2025 Mar 3;54(3):afaf064. [doi: 10.1093/ageing/afaf064]
34. Salthouse TA. When does age-related cognitive decline begin? Neurobiol Aging 2009 Apr;30(4):507-514. [doi: 10.1016/j.neurobiolaging.2008.09.023]
35. Toro B, Nester C, Farren P. A review of observational gait assessment in clinical practice. Physiother Theory Pract 2003 Jan;19(3):137-149. [doi: 10.1080/09593980307964]
36. Swanenburg J, de Bruin ED, Uebelhart D, Mulder T. Falls prediction in elderly people: a 1-year prospective study. Gait Posture 2010 Mar;31(3):317-321. [doi: 10.1016/j.gaitpost.2009.11.013]
37. König N, Taylor WR, Armbrecht G, Dietzel R, Singh NB. Identification of functional parameters for the classification of older female fallers and prediction of “first-time” fallers. J R Soc Interface 2014 Aug 6;11(97):20140353. [doi: 10.1098/rsif.2014.0353]
38. Malhotra R, Bautista MAC, Müller AM, et al. The aging of a young nation: population aging in Singapore. Gerontologist 2019 May 17;59(3):401-410. [doi: 10.1093/geront/gny160]
39. Sing CW, Lin TC, Bartholomew S, et al. Global epidemiology of hip fractures: secular trends in incidence rate, post-fracture treatment, and all-cause mortality. J Bone Miner Res 2023 Aug;38(8):1064-1075. [doi: 10.1002/jbmr.4821]
40. Cooper C, Campion G, Melton LJ III. Hip fractures in the elderly: a world-wide projection. Osteoporosis Int 1992 Nov;2(6):285-289. [doi: 10.1007/BF01623184]
41. Chandran M, Lau TC, Gagnon-Arpin I, et al. The health and economic burden of osteoporotic fractures in Singapore and the potential impact of increasing treatment rates through more pharmacological options. Arch Osteoporos 2019 Nov 26;14(1):114. [doi: 10.1007/s11657-019-0664-4]
42. Choo PL, Tou NX, Jun Pang BW, et al. Timed Up and Go (TUG) reference values and predictive cutoffs for fall risk and disability in Singaporean community-dwelling adults: Yishun cross-sectional study and Singapore longitudinal aging study. J Am Med Dir Assoc 2021 Aug;22(8):1640-1645. [doi: 10.1016/j.jamda.2021.03.002]

XSL•FO
RenderX

43. Chu LW, Chi I, Chiu AYY. Incidence and predictors of falls in the Chinese elderly. Ann Acad Med Singap 2005 Jan;34(1):60-72.
44. Makino K, Lee S, Bae S, et al. Prospective associations of physical frailty with future falls and fear of falling: a 48-month cohort study. Phys Ther 2021 Jun 1;101(6):pzab059. [doi: 10.1093/ptj/pzab059]
45. Yip WF, Ge L, Heng BH, Tan WS. Risk factors for incident falls in Singaporean community-dwelling adult men and women: a prospective cohort study. BMJ Open 2023 Mar 3;13(3):e057931. [doi: 10.1136/bmjopen-2021-057931]
46. Dai Z, Ang LW, Yuan JM, Koh WP. Association between change in body weight after midlife and risk of hip fracture-the Singapore Chinese Health Study. Osteoporos Int 2015 Jul;26(7):1939-1947. [doi: 10.1007/s00198-015-3099-9]
47. Dai W, Tham YC, Chee ML, et al. Falls and recurrent falls among adults in a multi-ethnic Asian population: the Singapore epidemiology of eye diseases study. Sci Rep 2018 May 15;8(1):7575. [doi: 10.1038/s41598-018-25894-8]
48. Hodkinson HM. Evaluation of a mental test score for assessment of mental impairment in the elderly. Age Ageing 1972 Nov;1(4):233-238. [doi: 10.1093/ageing/1.4.233]
49. MOH TRUST: improving health outcomes through trusted data exchange. Ministry of Health. 2024. URL: https://www.trustplatform.sg [accessed 2026-09-06]
50. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x]
51. Pendlebury ST, Welch SJV, Cuthbertson FC, Mariz J, Mehta Z, Rothwell PM. Telephone assessment of cognition after transient ischemic attack and stroke: modified telephone interview of cognitive status and telephone Montreal Cognitive Assessment versus face-to-face Montreal Cognitive Assessment and neuropsychological battery. Stroke 2013 Jan;44(1):227-229. [doi: 10.1161/STROKEAHA.112.673384]
52. Lubben J, Blozik E, Gillmann G, et al. Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. Gerontologist 2006 Aug;46(4):503-513. [doi: 10.1093/geront/46.4.503]
53. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x]
54. Stalvey BT, Owsley C, Sloane ME, Ball K. The Life Space Questionnaire: a measure of the extent of mobility of older adults. J Appl Gerontol 1999 Dec;18(4):460-478. [doi: 10.1177/073346489901800404]
55. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index: A simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. Md State Med J 1965;14:61-65. [Medline: 14258950]
56. Graf C. The Lawton Instrumental Activities of Daily Living (IADL) Scale. Medsurg Nurs 2009;18(5):315-316. [doi: 10.1097/01.NAJ.0000314810.46029.74] [Medline: 19927971]
57. Craig CL, Marshall AL, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB]
58. Rubenstein LZ, Vivrette R, Harker JO, Stevens JA, Kramer BJ. Validating an evidence-based, self-rated Fall Risk Questionnaire (FRQ) for older adults. J Safety Res 2011 Dec;42(6):493-499. [doi: 10.1016/j.jsr.2011.08.006]
59. Delbaere K, Smith ST, Lord SR. Development and initial validation of the Iconographical Falls Efficacy Scale. J Gerontol A Biol Sci Med Sci 2011 Jun;66(6):674-680. [doi: 10.1093/gerona/glr019]
60. Lim ML, van Schooten KS, Radford KA, et al. The Iconographical Falls Efficacy Scale (IconFES) in community-dwelling older people-a longitudinal validation study. Age Ageing 2021 May 5;50(3):822-829. [doi: 10.1093/ageing/afaa213]
61. FRAX: fracture risk assessment tool. Osteoporosis Research Ltd. URL: https://www.fraxplus.org [accessed 2026-09-06]
62. Wainwright SA, Marshall LM, Ensrud KE, et al. Hip fracture in women without osteoporosis. J Clin Endocrinol Metab 2005 May;90(5):2787-2793. [doi: 10.1210/jc.2004-1568] [Medline: 15728213]
63. Jha D, Chandran M, Hong N, et al. Discriminatory accuracy of Fracture Risk Assessment Tool in Asian populations: a systematic review and meta-analysis. J Bone Metab 2024 Nov;31(4):296-315. [doi: 10.11005/jbm.24.781]
64. Marques A, Ferreira RJO, Santos E, Loza E, Carmona L, da Silva JAP. The accuracy of osteoporotic fracture risk prediction tools: a systematic review and meta-analysis. Ann Rheum Dis 2015 Nov;74(11):1958-1967. [doi: 10.1136/annrheumdis-2015-207907]
65. Schneider S, Popp WL, Brogioli M, et al. Reliability of wearable-sensor-derived measures of physical activity in wheelchair-dependent spinal cord injured patients. Front Neurol 2018;9:1039. [doi: 10.3389/fneur.2018.01039]
66. Rehman RZU, Zhou Y, Del Din S, et al. Gait analysis with wearables can accurately classify fallers from non-fallers: a step toward better management of neurological disorders. Sensors (Basel) 2020 Dec 7;20(23):6992. [doi: 10.3390/s20236992]
67. Subramaniam S, Faisal AI, Deen MJ. Wearable sensor systems for fall risk assessment: a review. Front Digit Health 2022;4:921506. [doi: 10.3389/fdgth.2022.921506]
68. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. BMC Geriatr 2014 Feb 1;14:1-14. [doi: 10.1186/1471-2318-14-14]
69. Bogle Thorbahn LD, Newton RA. Use of the Berg Balance Test to predict falls in elderly persons. Phys Ther 1996 Jun;76(6):576-583. [doi: 10.1093/ptj/76.6.576]

XSL•FO
RenderX

70. Renggli D, Graf C, Tachatos N, et al. Wearable inertial measurement units for assessing gait in real-world environments. Front Physiol 2020;11:90. [doi: 10.3389/fphys.2020.00090]
71. Kim YK, Pai SGS, Choi JO, et al. Leveraging deep learning and wearables for automatically identifying gait event: effects of age and location of sensors on the assessment of gait events. IEEE Sensors J 2024;25(1):792-802. [doi: 10.1109/JSEN.2024.3487018]
72. Picerno P. 25 years of lower limb joint kinematics by using inertial and magnetic sensors: a review of methodological approaches. Gait Posture 2017 Jan;51:239-246. [doi: 10.1016/j.gaitpost.2016.11.008]
73. Iosa M, Picerno P, Paolucci S, Morone G. Wearable inertial sensors for human movement analysis. Expert Rev Med Devices 2016 Jul;13(7):641-659. [doi: 10.1080/17434440.2016.1198694] [Medline: 27309490]
74. Werner C, Awai Easthope C, Curt A, Demkó L. Towards a mobile gait analysis for patients with a spinal cord injury: a robust algorithm validated for slow walking speeds. Sensors (Basel) 2021 Nov 6;21(21):7381. [doi: 10.3390/s21217381]
75. Willi R, Werner C, Demkó L, et al. Reliability of patient-specific gait profiles with inertial measurement units during the 2-min walk test in incomplete spinal cord injury. Sci Rep 2024 Feb 6;14(1):3049. [doi: 10.1038/s41598-024-53301-y]
76. Hamacher D, Singh NB, Van Dieën JH, Heller MO, Taylor WR. Kinematic measures for assessing gait stability in elderly individuals: a systematic review. J R Soc Interface 2011 Dec 7;8(65):1682-1698. [doi: 10.1098/rsif.2011.0416]
77. Ravi DK, Gwerder M, König Ignasiak N, et al. Revealing the optimal thresholds for movement performance: A systematic review and meta-analysis to benchmark pathological walking behaviour. Neurosci Biobehav Rev 2020 Jan;108:24-33. [doi: 10.1016/j.neubiorev.2019.10.008] [Medline: 31639377]
78. Chen LK, Woo J, Assantachai P, et al. Asian working group for sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020 Mar;21(3):300-307. [doi: 10.1016/j.jamda.2019.12.012]
79. Jha D, Praveen AD, Baker A, et al. Dual-energy X-ray absorptiometry based biofidelic finite element models for simulating falls to the hip: impact of trochanteric soft tissue thickness on fracture risk. Clin Biomech (Bristol) 2025 Nov 21;2025:106720. [doi: 10.1016/j.clinbiomech.2025.106720]
80. Chandran M, Ganesan G, Tan KB, Reginster JY, Hiligsmann M. Using health-economic evidence to support policy-level decision-making in Singapore-sensitivity analysis that provides further confidence in fracture probability-based cost-effective intervention thresholds. Osteoporos Int 2021 Apr;32(4):787-789. [doi: 10.1007/s00198-021-05876-9]
81. Schuit SCE, van der Klift M, Weel A, et al. Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam Study. Bone 2004 Jan;34(1):195-202. [doi: 10.1016/j.bone.2003.10.001]
82. Stone KL, Seeley DG, Lui LY, et al. BMD at multiple sites and risk of fracture of multiple types: long-term results from the Study of Osteoporotic Fractures. J Bone Miner Res 2003 Nov;18(11):1947-1954. [doi: 10.1359/jbmr.2003.18.11.1947]
83. Fleps I, Pálsson H, Baker A, et al. Finite element derived femoral strength is a better predictor of hip fracture risk than aBMD in the AGES Reykjavik study cohort. Bone 2022 Jan;154:116219. [doi: 10.1016/j.bone.2021.116219]
84. Praveen AD, Jha D, Baker A, et al. Comparison of the time-dependent discriminatory accuracy of femoral strength and bone mineral density for predicting future hip and major osteoporotic fractures: a 16-year follow-up of the AGES-Reykjavik cohort. Osteoporos Int 2025 Jul;36(7):1175-1184. [doi: 10.1007/s00198-025-07503-3]
85. Yosibash Z, Trabelsi N, Buchnik I, et al. Hip fracture risk assessment in elderly and diabetic patients: combining autonomous finite element analysis and machine learning. J Bone Miner Res 2020 Dec 1;38(6):876-886. [doi: 10.1002/jbmr.4805]
86. Grassi L, Väänänen SP, Voss A, et al. DXA-based 3D finite element models predict hip fractures better than areal BMD in elderly women. Bone 2025 Jun;195:117457. [doi: 10.1016/j.bone.2025.117457]
87. Grassi L, Väänänen SP, Jehpsson L, et al. 3D Finite element models reconstructed from 2D Dual-energy x-ray absorptiometry (DXA) images improve hip fracture prediction compared to areal BMD in osteoporotic fractures in men (MrOS) Sweden cohort. J Bone Miner Res 2023 Sep;38(9):1258-1267. [doi: 10.1002/jbmr.4878]
88. Colombo G, Minta K, Grübel J, Tai WLE, Hölscher C, Schinazi VR. Detecting cognitive impairment through an age-friendly serious game: the development and usability of the Spatial Performance Assessment for Cognitive Evaluation (SPACE). Comput Human Behav 2024 Nov;160:108349. [doi: 10.1016/j.chb.2024.108349]
89. Minta* K, Colombo* G, Tee* M, et al. SPACE: a novel digital tool for assessing hippocampal structural integrity in older adults. in review. Preprint posted online on 2024. [doi: 10.21203/rs.3.rs-5484040/v1]
90. Colombo G, Minta K, Taylor WR, et al. Spatial navigation as a digital marker for clinically differentiating cognitive impairment severity. Commun Med (Lond) 2026 Mar 5;6(1):228. [doi: 10.1038/s43856-026-01484-y] [Medline: 41787066]
91. Hampel H, Au R, Mattke S, et al. Designing the next-generation clinical care pathway for Alzheimer's disease. Nat Aging 2022 Aug;2(8):692-703. [doi: 10.1038/s43587-022-00269-x]
92. van der Flier WM, de Vugt ME, Smets EMA, Blom M, Teunissen CE. Towards a future where Alzheimer's disease pathology is stopped before the onset of dementia. Nat Aging 2023 May;3(5):494-505. [doi: 10.1038/s43587-023-00404-2]
93. Arevalo-Rodriguez I, Smailagic N, Roqué I Figuls M, et al. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev 2015 Mar 5(3):CD010783. [doi: 10.1002/14651858.CD010783.pub2]
94. Dautzenberg G, Lijmer J, Beekman A. Diagnostic accuracy of the Montreal Cognitive Assessment (MoCA) for cognitive screening in old age psychiatry: determining cutoff scores in clinical practice. Avoiding spectrum bias caused by healthy controls. Int J Geriatr Psychiatry 2020 Mar;35(3):261-269. [doi: 10.1002/gps.5227]

XSL•FO
RenderX

95. Coughlan G, Laczó J, Hort J, Minihane AM, Hornberger M. Spatial navigation deficits - overlooked cognitive marker for preclinical Alzheimer disease? Nat Rev Neurol 2018 Aug;14(8):496-506. [doi: 10.1038/s41582-018-0031-x] [Medline: 29980763]
96. Hafting T, Fyhn M, Molden S, Moser MB, Moser EI. Microstructure of a spatial map in the entorhinal cortex. Nature 2005 Aug 11;436(7052):801-806. [doi: 10.1038/nature03721]
97. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol 1991;82(4):239-259. [doi: 10.1007/BF00308809] [Medline: 1759558]
98. DeTure MA, Dickson DW. The neuropathological diagnosis of Alzheimer's disease. Mol Neurodegener 2019 Aug 2;14(1):32. [doi: 10.1186/s13024-019-0333-5]
99. Liu Q, Song H, Yan M, et al. Virtual reality technology in the detection of mild cognitive impairment: a systematic review and meta-analysis. Ageing Res Rev 2023 Jun;87:101889. [doi: 10.1016/j.arr.2023.101889]
100. Falls. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2026-09-06]
101. Anand S, Schoo C. Mild Cognitive Impairment: StatPearls Publishing; 2024. [Medline: 38261679]
102. Chan KM, Pang WS, Ee CH, Ding YY, Choo P. Epidemiology of falls among the elderly community dwellers in Singapore. Singapore Med J 1997 Oct;38(10):427-431.
103. Romli MH, Tan MP, Mackenzie L, Lovarini M, Suttanon P, Clemson L. Falls amongst older people in Southeast Asia: a scoping review. Public Health 2017 Apr;145:96-112. [doi: 10.1016/j.puhe.2016.12.035]
104. König N, Singh NB, von Beckerath J, Janke L, Taylor WR. Is gait variability reliable? An assessment of spatio-temporal parameters of gait variability during continuous overground walking. Gait Posture 2014 Jan;39(1):615-617. [doi: 10.1016/j.gaitpost.2013.06.014]
105. Godinho C, Domingos J, Cunha G, et al. A systematic review of the characteristics and validity of monitoring technologies to assess Parkinson's disease. J Neuroeng Rehabil 2016 Mar 12;13(1):24. [doi: 10.1186/s12984-016-0136-7]
106. Chan A, Saito Y, Matchar DB, et al. Cohort Profile: Panel on Health and Ageing of Singaporean Elderly (PHASE). Int J Epidemiol 2019 Dec 1;48(6):1750-1751f. [doi: 10.1093/ije/dyz172]

Abbreviations

aBMD: areal bone mineral density
DXA: dual-energy X-ray absorptiometry
EMR: electronic medical record
FEM: finite element model
FRAX: Fracture Risk Assessment Tool
FRAX-HF: Fracture Risk Assessment Tool–Hip Fracture
FRAX-MOF: Fracture Risk Assessment Tool–Major Osteoporotic Fracture
***ICD*:** *International Classification of Diseases*
IMU: inertial measurement unit
MoCA: Montreal Cognitive Assessment
MOF: major osteoporotic fracture
SPACE: Spatial Performance Assessment for Cognitive Evaluation
STEADI: Stopping Elderly Accidents, Deaths, and Injuries
T-MoCA: MoCA designed for telephone administration
TARGET: Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments
TRUST: Trusted Research and Real-World Data Utilisation and Sharing Tech
TSTT: trochanteric soft tissue thickness
VR: virtual reality

Edited by P Jansons; submitted 10.Dec.2025; peer-reviewed by TC Peng, Y Hu; revised version received 01.Jul.2026; accepted 17.Jul.2026; published 25.Sep.2026.

Please cite as:
Chan A, Visaria A, Malhotra R, Singh NB, Helgason B, Schinazi VR, Tan KY, Matchar DB, Sim R, Koh VJW, Lai WX, Troeung L, Lim CWH, Pai SGS, Tan KZ, Praveen AD, Jha D, Ferguson SJ, Colombo G, Minta K, Lamoureux EL, Taylor WR
Rationale, Design, and Baseline Characteristics of the TARGET (Targeted Assessment and Recruitment of Geriatrics for Effective Fall Prevention Treatments) Cohort of Older Adults for Assessing Fall and Fracture Risk: Prospective Cohort Study
JMIR Aging 2026;9:e89285
URL: https://aging.jmir.org/2026/1/e89285
doi:10.2196/89285

XSL•FO
RenderX

© Angelique Chan, Abhijit Visaria, Rahul Malhotra, Navrag B Singh, Benedikt Helgason, Victor R Schinazi, Kok Yang Tan, David B Matchar, Rita Sim, Vanessa Jean Wen Koh, Wei Xuan Lai, Lakkhina Troeung, Catherine Wen Huey Lim, Sai G S Pai, Kai Zhe Tan, Anitha D Praveen, Dheeraj Jha, Stephen J Ferguson, Giorgio Colombo, Karolina Minta, Ecosse L Lamoureux, William R Taylor. Originally published in JMIR Aging (https://aging.jmir.org), 25.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Validation of Machine Learning Models for Predicting Falls Among Hospitalized Older Adults: Retrospective Cross-Sectional Study

Xiyao Yang[1*], MSc; Juan Ren[1*], MSc; Dan Su[2,3*], PhD; Manzhen Bao[1], MSc; Miao Zhang[1], MSc; Xiaoming Chen[1], BSc; Yanhua Li[1], BSc; Zonggui Wang[1], BSc; Xiujing Dai[1], BSc; Zengzeng Wei[1], BSc; Shuiyu Zhang[2], BSc; Yuxin Zhang[1], MSc; Juan Li[1], BSc; Xiaolin Li[1], BSc; Junjin Xu[1], BSc; Nan Mo[1], MSc

[1]Second Hospital of Anhui Medical University, No.678 Furong Road, Shushan District, Hefei, China

[2]School of Nursing, Anhui Medical University, Hefei, China

[3]The Taikang Health and Wellness Industry Research Institute, Anhui Medical University, Hefei, China

[*]these authors contributed equally

Corresponding Author:
Nan Mo, MSc
Second Hospital of Anhui Medical University, No.678 Furong Road, Shushan District, Hefei, China

Related Article:

This is a corrected version. See correction statement: https://aging.jmir.org/2026/1/e101284

Abstract

Background: Falls are one of the leading causes of injury or death among older adults. Falls occurring in individuals during hospitalization, as an adverse event, are a key concern for health care institutions. Identifying older adults at high risk of falls in clinical settings enables early interventions, thereby reducing the incidence of falls.

Objective: This study aims to develop and validate machine learning models to predict the risk of falls among hospitalized older adults.

Methods: This study retrospectively analyzed data from a tertiary general hospital in China, including 342 older adults who experienced falls and 684 randomly matched nonfallers, between January 2018 and December 2024, encompassing demographic information, comorbidities, laboratory parameters, and medication use, among other variables. The dataset was randomly split into training and testing sets in a 7:3 ratio. Predictors were selected from the training set using stepwise regression, least absolute shrinkage and selection operator, and random forest-recursive feature elimination. Seven machine learning algorithms were employed to develop predictive models in the training set, and their performance was compared in the testing set. The optimal model was interpreted using Shapley Additive Explanations (SHAP).

Results: The gradient boosting machine model demonstrated the best predictive performance (C-index 0.744, 95% CI 0.688 - 0.799). The 8 most important variables associated with fall risk were dizziness, epilepsy, fall history within the past 3 months, use of walking assistance, emergency admission, Morse Fall Scale scores, modified Barthel Index scores, and the number of indwelling catheters. The model was interpreted using SHAP to enhance the clinical utility of the predictive model.

Conclusions: The gradient boosting machine model was identified as the optimal predictive model. The SHAP method enhanced its integration into clinical workflows.

(JMIR Aging 2026;9:e80602) doi:10.2196/80602

KEYWORDS

machine learning; risk prediction; older adults; fall prediction; gradient boosting machine; random forest

Introduction

Falls, the second leading cause of global unintentional injury deaths, are a significant public health concern. They are defined as “an event that leads to a person inadvertently coming to rest on the ground, floor, or other lower surface than their original position [1].” Age is one of the main risk factors for falls [2], and statistics indicate that the incidence of falls among older adults is approximately 26.5% [3]. Among individuals aged >60 years globally, falls are one of the most common causes of injury or death, with one out of every 5 falls resulting in a fracture or head injury [2,4]. In addition, falls generate

XSL•FO
RenderX

substantial medical costs, imposing a heavy economic burden worldwide [5].

Notably, falls are adverse events in hospitals, and the prevention of falls is also a priority for improving the quality of nursing care [6]. The incidence of falls in hospitals is typically in the range of 2 to 16 per 1000 bed days [7,8]. Despite a declining incidence of falls among hospitalized older adults, the increasing number of older adults admitted to hospitals, driven by an expanding aging population, suggests that falls prevention will remain a critical concern in hospitals [8,9]. Falls are preventable adverse events in hospitals, and implementing fall prevention programs can avoid costs of US $14,600 per 1000 patient-days [6]. Therefore, identifying individuals at high risk of falls in hospitals to take preventive measures is particularly important, especially among older adults.

The MFS (Morse Fall Scale) and STRATIFY (St. Thomas's Risk Assessment Tool in Falling Elderly Inpatients), widely used in hospitals to identify individuals at high risk of falls, have drawbacks such as low specificity [10,11]. Several studies have developed predictive models for fall risk in older inpatients [12-17]. While some employed traditional regression methods [12-15,17], these conventional approaches often struggle with complex, multidimensional data [18]. Other models exhibit limited applicability, being restricted to specific clinical settings or units [12,13,17]. Additionally, certain models rely solely on clinical texts for prediction, a methodology constrained by single-variable limitations that compromise performance [14].

In recent years, machine learning (ML) algorithms have attracted considerable interest in health care predictive modeling due to their capacity to develop highly accurate prediction models at low cost [19]. The capacity of ML algorithms to process high-dimensional data not only enhances the accuracy and efficiency of predictive models but also enables personalized risk prediction [4,20]. Although existing studies have employed ML algorithms to develop fall prediction models for hospitalized older adults, these models exhibit limitations, including suboptimal performance, applicability restricted to specific geriatric subpopulations, and reliance on environmental detection systems that hinder their widespread clinical adoption [16,21-23]. Critically, limited studies have offered comprehensive explanations or analyses of model predictions, restricting clinical applicability and diminishing the practical value of these models.

Therefore, the objective of our study is to develop and validate multiple ML models utilizing clinically accessible data to predict fall risk of hospitalized older adults. We seek to identify the optimal model while interpreting its predictions through the Shapley Additive Explanations (SHAP) method.

Methods

Data Source and Participants

Using an adverse event reporting system integrated into electronic nursing workstations of a tertiary general hospital, researchers retrieved fall incident records for hospitalized older adults (aged ≥60 y) occurring between January 2018 and December 2024, extracting hospitalization identifiers and fall timestamps. An electronic health record (EHR) system was used to record admission and discharge dates along with hospitalization identifiers for older adults without a history of falls hospitalized between January 2018 and December 2024. The fall timestamp of each case patient was used to anchor the index time for the matched controls. For each case, 2 controls were matched. Specifically, we first preprocessed the data by removing duplicate records from individuals with multiple hospitalizations (retaining only the first admission). From this refined pool of potential controls, we then used a Visual Basic algorithm in Microsoft Excel (version 16.0) to identify patients whose entire hospitalization period (from admission to discharge) encompassed the fall timestamp. This approach ensures that both cases and controls were exposed to similar time-dependent clinical factors at the same specific time point, thereby minimizing potential time-dependent bias. The case and control groups were not matched on demographics such as age or gender in order to maintain the natural distribution found in real-world clinical settings. With the aim of capturing all relevant information, variables with clinical or predictive relevance were included as model features for the ML algorithm to parse their associations with the outcome.

Matched controls identified as day cases were excluded and replaced until a 1:2 case-control ratio was maintained. This ratio was selected based on considerations of statistical power, cost-effectiveness, and practical constraints, as increasing the control-to-case ratio beyond 2:1 yields diminishing returns in power while substantially increasing costs and workload [24,25]. Cases were initially identified from the adverse event reporting system as any patient with a documented fall event occurring within the hospital premises and were excluded if they were aged <60 years old at admission, experienced subsequent falls occurring during the same hospitalization, or were nonhospitalized patients or day cases. Controls were selected from the EHR system as hospitalized patients aged ≥60 years with no record of an in-hospital fall and were excluded for having duplicate admission records (only the first was retained), day-case status, or if they could not be matched to a case. We excluded day case patients because more than 20% of the data were missing in EHR. The sample size was estimated using the "pmsampsize" package in R software (version 4.5.0). According to other researchers, the c-statistic is 0.73, the number of predictor parameters chosen for our study is 17, and the prevalence is 0.33 (1/3), with a required sample size of 992 for the calculation.

Ethical Considerations

The study was approved by the Ethics Committee Board of the Second Affiliated Hospital of Anhui Medical University (approval number: YX2025-162). This study adheres strictly to privacy protection principles. Nonessential identifying information is omitted during data processing, and informed consent is obtained when necessary. Informed consent was waived for patients who died or were disconnected. No financial compensation is provided to participants. This study conforms to the principles outlined in the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) statement.

XSL•FO
RenderX

Data Collection and Processing

All records containing timestamps and hospitalization identifiers were randomly split into 2 datasets. Two uniformly trained data collectors independently extracted variables through the EHR system using these identifiers, followed by cross-verification upon completion. Five categories of variables were collected: demographic characteristics, comorbidities, medications, laboratory indicators, and other variables. Table S1 in Multimedia Appendix 1 provides the list of 64 extracted variables. Demographic characteristics and sleep duration data were extracted from hospital admission records. The absence of BMI values was directly attributable to practical barriers in anthropometric data collection for patients with mobility limitations (bedridden or wheelchair-dependent status). Comorbidities were identified by integrating inpatient diagnoses from admission summaries with discharge diagnoses in corresponding discharge records. Medication administration records were retrieved from both permanent and temporary medical orders to capture all medications administered within the 24-hour period preceding the timestamp. Polypharmacy was defined as taking 5 or more medications daily. Laboratory indicators were collected from laboratory test reports. For indicators with repeated measurements, data within the 7 days before and after the timestamp were selected for analysis. The remaining variables were extracted from nursing records within 1 week before and after the timestamp. Given that at least 2 nursing records are documented weekly, there is no missing data for these variables. A total of 64 variables were initially extracted. With 27.49% missing values, BMI was removed from analysis. For the remaining variables, only albumin and hemoglobin contained missing values (0.03% and 0.04%, respectively). The missing values for albumin and hemoglobin were imputed using the random forest (RF) imputation method, implemented via the "missForest" package in R software (version 4.5.0). This approach offers the advantage of handling mixed data types (continuous and categorical) and effectively capturing nonlinear relationships among variables [26].

Feature Selection

The dataset was randomly split into a training set (70%) and a testing set (30%). A three-step selection strategy was implemented in the training set to identify optimal predictors. First, univariate (LR) was applied for preliminary screening (P<.05) to retain statistically significant variables. Second, 5 feature selection methods were integrated: stepwise regression (SR) comprises 3 variants—forward selection, backward selection, and bidirectional elimination; least absolute shrinkage and selection operator (LASSO); and random forest-recursive feature elimination (RF-RFE). Predictors were determined by the overlap among the results of these methods. This approach aimed to mitigate high correlation among predictors while capturing their complex relationships with the outcome variable [27]. SR iteratively adjusts variables based on statistical significance, LASSO addresses high dimensionality and multicollinearity while preventing overfitting, and RF-RFE captures nonlinear patterns and variable interactions. Both LASSO and RF-RFE incorporated 10-fold cross-validation. Finally, clinical experts validated the selected predictors to ensure clinical applicability.

Models Development and Validation

To comprehensively evaluate predictive performance and ensure robust results, we employed multiple algorithms to construct predictive models in the training set, including seven ML models: LR, support vector machines, RF, gradient boosting machine (GBM), extreme gradient boosting (XGBoost), k-nearest neighbor (KNN), and neural network (NN). Grounded in distinct modeling philosophies, each algorithm offers unique advantages. LR establishes an optimal linear decision boundary, valued for its conceptual simplicity and high interpretability, serving as a reliable performance benchmark [20]. Support vector machines aim to determine a separating hyperplane that maximizes the geometric margin for robust classification. They address nonlinear problems by employing kernel functions to project data into a higher-dimensional feature space where the maximum-margin principle is applied [20]. As a representative bagging ensemble, RF enhances predictive stability and captures complex feature interactions by aggregating numerous decorrelated decision trees, also providing inherent resistance to overfitting and enabling feature importance evaluation [20,28]. GBM employs a sequential modeling strategy that iteratively corrects errors from preceding models, often achieving high predictive accuracy [20]. XGBoost, an optimized implementation of gradient boosting, incorporates regularization and advanced algorithmic techniques to further improve computational efficiency and performance [20]. KNN is an instance-based learning method operating on the principle of local similarity. Predictions are derived from the majority label or average value of a sample's KNNs in the feature space, offering an intuitive perspective on the local data structure [29]. NN, or deep learning models, function as universal approximators by leveraging multiple layers of tunable nonlinear transformations. This architecture enables them to automatically learn hierarchical data representations and extract complex, high-level features through training [30]. This systematic selection of algorithms, encompassing linear models, kernel methods, bagging and boosting ensembles, instance-based learning, and NNs, ensures our evaluation is comprehensive and avoids bias toward any single modeling strategy.

To mitigate class imbalance, we applied random upsampling to the training dataset, which involves duplicating instances from the minority class at random to balance the class distribution. Subsequently, to rigorously tune hyperparameters and guard against overfitting, we performed a grid search with 10-fold cross-validation on this processed training set to identify the optimal parameters. The test set was used to evaluate model performance. The area under the receiver operating characteristic curve (AUROC) in the testing set served as the primary metric for assessing discriminative ability. Model discrimination was primarily assessed using the AUROC. This metric is considered a standard method for evaluating ranking ability, as it provides a threshold-independent assessment of a model's inherent discriminative power [31]. Additionally, model performance was comprehensively evaluated using the area under the precision-recall curve (AUPRC), which is particularly informative for imbalanced datasets, along with sensitivity, specificity, accuracy, recall, F_1-score, positive predictive value, and negative predictive value. Calibration curves were plotted

XSL•FO
RenderX

to assess prediction accuracy. Decision curve analysis (DCA) was performed to quantify clinical utility. SHAP is a model interpretation tool that calculates feature contribution values to provide both global (model-level) and local (individual prediction) explanations, making models more interpretable and applicable [20,32].

Therefore, we employed the SHAP method to elucidate how individual features influence fall risk predictions in hospitalized older adults within the optimal model.

Statistical Analysis

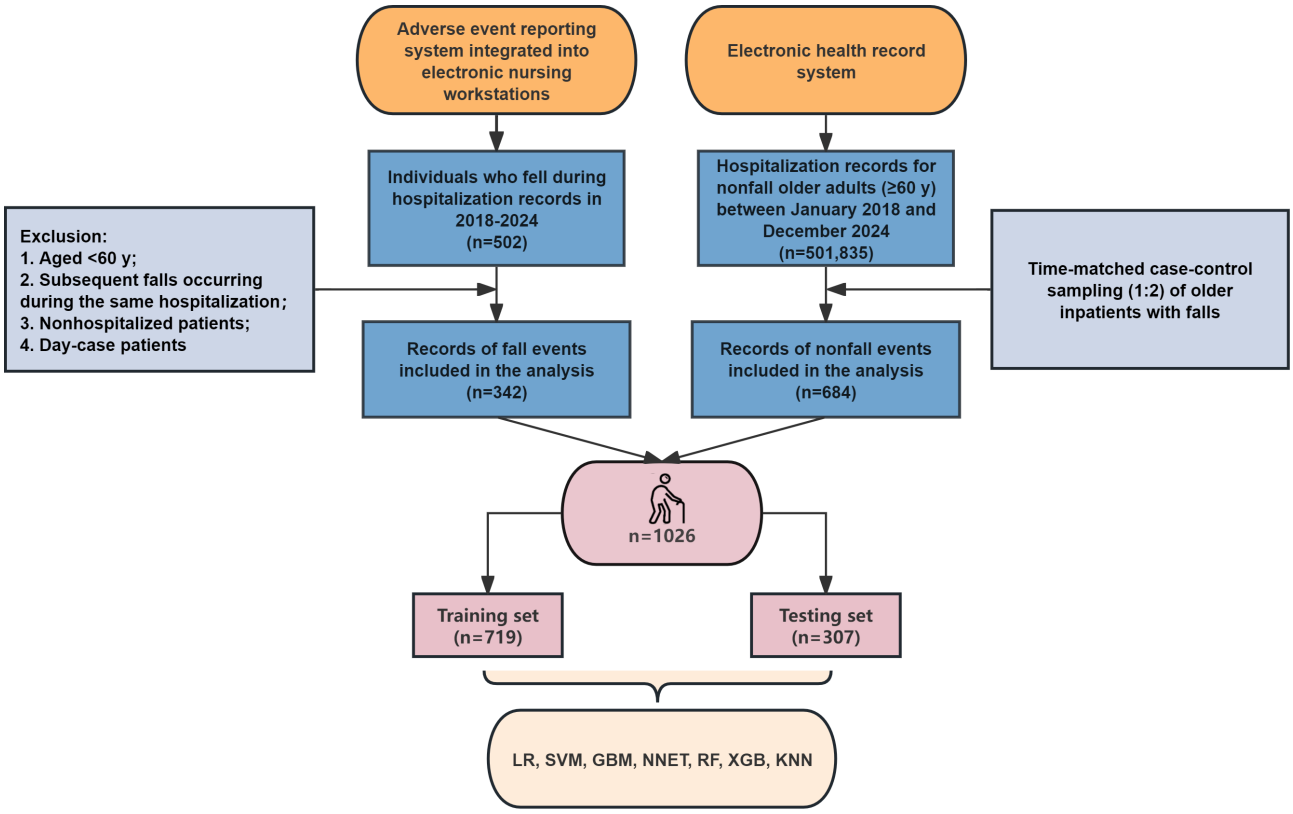
All statistical analyses were performed using R software (version 4.5.0), with categories merged when necessary to address sparse data. Use of walking assistance (UWA) was classified into 4 groups: no assistance, wheelchair or bedridden, support by others or furniture, and walker/crutches/cane. Continuous variables were categorized as follows: age into 60 to 69, 70 to 79, and ≥80 years; serum albumin into ≥34 and <34 g/L [33]; MFS scores into <45 points and ≥45 points [34]; modified Barthel Index (mBI) [35] scores into 0 to 20 points, 21 to 60 points, 61 to 90 points, 91 to 99 points, and 100 points; Nutritional Risk Screening 2002 (NRS 2002) scores into <3 points (no nutritional risk) and ≥3 points (at risk) [36]; and Numeric Pain Rating Scale scores into 0 points (no pain) and ≥1 point (pain) [37]. Continuous variables, none of which followed a normal distribution, were expressed as medians and IQR (M, Q1-Q3). Categorical variables were reported as numbers and percentages (n, %). Differences between groups were analyzed using the Mann-Whitney *U* test for nonnormally distributed continuous variables and the Chi-square test (or Fisher exact test for sparse data) for categorical variables. A 2-sided $P<.05$ was considered statistically significant.

Results

Baseline Characteristics

Ultimately, 1026 older adults were included in the study. Figure 1 illustrates the process of patient screening. The comparison between fallers and nonfallers in the overall dataset is presented in Table S2 in Multimedia Appendix 1. Among the 1026 patients, 40.84% (419/1026) were aged 60 to 69 years and 55.65% (571/1026) were male. Among the 342 fallers, 40.06% (137/342) were aged 70 to 79 years and 52.05% (178/342) were male. Significant differences were observed between fallers and nonfallers in the following variables: age, blood pressure, diabetes, chronic kidney disease, heart failure, hypothyroidism, cancer, Parkinson's disease, dizziness, stroke, gait abnormality, epilepsy, visual impairment, hearing impairment, polypharmacy, antiplatelet drugs, statins, α-blockers, vasodilators, antidiabetic drugs, anti-Parkinson's disease drugs, antiepileptic drugs, benzodiazepines, Z-drugs, albumin levels, fall history in the past 3 months, UWA, emergency admission (EA), MFS scores, NRS 2002 scores, mBI scores, number of indwelling catheters (Indw Cath), and departments.

Figure 1. Flowchart of the patient screening. GBM: gradient boosting machine; KNN: k-nearest neighbor; LR: logistic regression; NNET: neural network; RF: random forest; SVM: support vector machine; XGB: extreme gradient boosting.



Selection of Predictor Variables

Univariate LR identified 26 potential predictors ($P<.05$) in the training set. The mBI did not reach statistical significance in the univariate analysis, but its *P* value was .054, which is very close to the threshold. Given that the mBI is an important clinical indicator associated with fall risk and to avoid omitting

XSL•FO
RenderX

potentially influential factors due to a rigid threshold, we decided after expert discussion to include it as the 27th variable in the subsequent feature selection, as detailed in Table S3 of Multimedia Appendix 1. Table 1 displays the predictors identified by the 5 methods (SR-forward selection, SR-backward selection, SR-bidirectional elimination, LASSO, and RF-RFE). Table S4, Figure S1, and Figure S2 in Multimedia Appendix 1 provide detailed information. Figure 2 visualizes the overlap of predictors selected across 5 methods. The intersecting predictors from these methods formed the final predictor set, comprising dizziness, epilepsy, fall history in the past 3 months, UWA, EA, MFS scores, mBI scores, and Indw Cath. After expert consultation, no predictors were added or removed. The final development model included these 8 predictor variables.

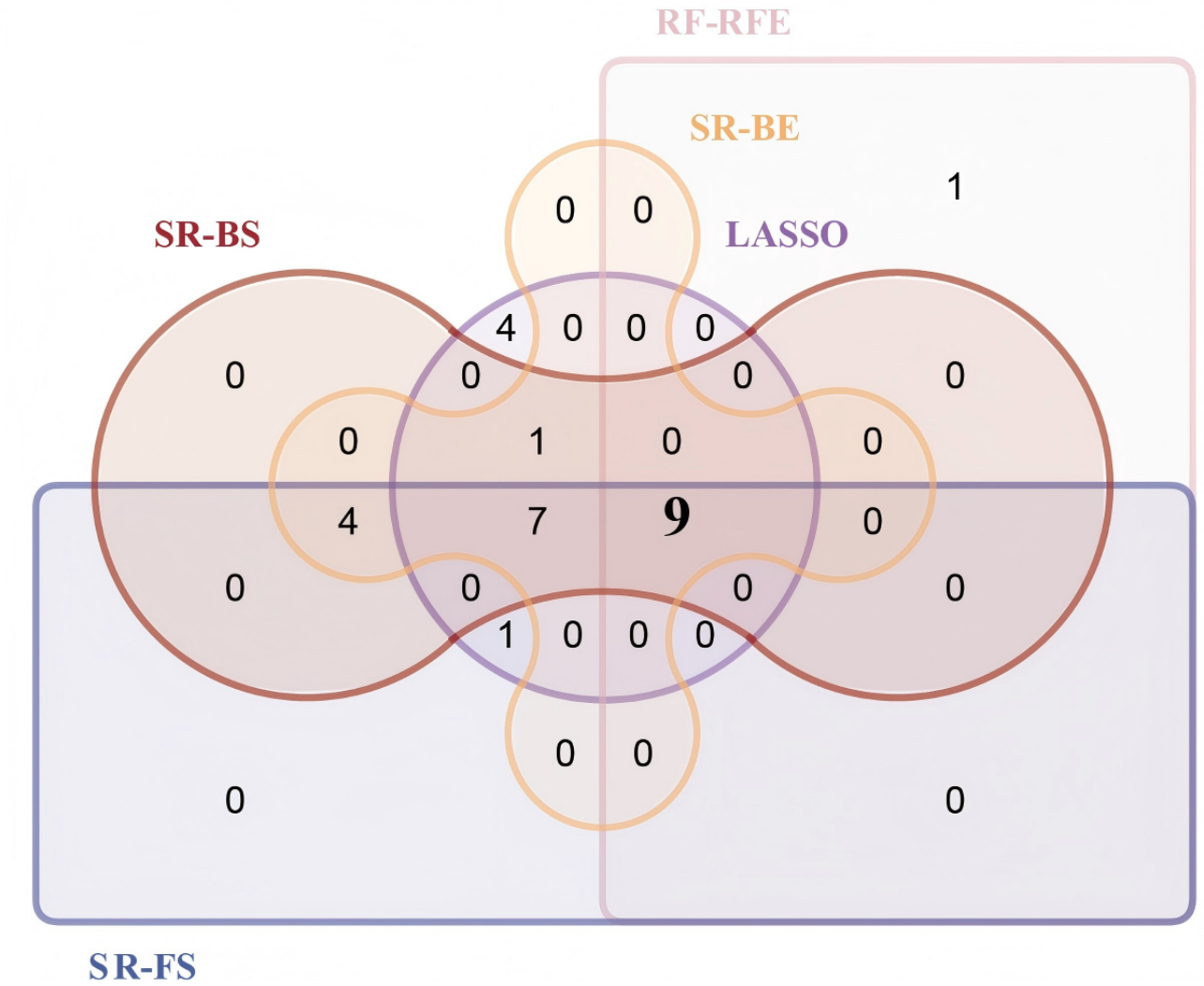
Table . The predictors obtained through 5 selection methods.

Methods	Number of predictors (categories)	Predictor variables
SR-FS	15 (21)	Hypothyroidism, OP[a], dizziness, stroke, epilepsy, polypharmacy, AC[b], BZDs[c], Alb[d], FH-3M[e], UWA.1[f], UWA.2[g], UWA.3[h], EA[i], MFS[j] scores, mBI.1[k], mBI.2[l], mBI.3[m], mBI.4[n], Indw Cath.1[o], Indw Cath.2[p]
SR-BS	15 (21)	Hypothyroidism, OP, dizziness, stroke, epilepsy, AC, BZDs, Zdrugs, Alb, FH-3M, UWA.1, UWA.2, UWA.3, EA, MFS scores, mBI.1, mBI.2, mBI.3, mBI.4, Indw Cath.1, Indw Cath.2
SR-BE	15 (21)	Hypothyroidism, OP, dizziness, stroke, epilepsy, AC, BZDs, Zdrugs, Alb, FH-3M, UWA.1, UWA.2, UWA.3, EA, MFS scores, mBI.1, mBI.2, mBI.3, mBI.4, Indw Cath.1, Indw Cath.2
LASSO	19 (22)	Gender, hypothyroidism, OP, dizziness, stroke, epilepsy, polypharmacy, AC, antidiabetics, BZDs, Zdrugs, Alb, FH-3M, UWA.1, UWA.2, UWA.3, EA, MFS scores, mBI.0[q], mBI.4, Indw Cath.2, Department.2[r]
RF-RFE	9 (10)	CA[s], dizziness, epilepsy, FH-3M, UWA.2, UWA.3, EA, MFS scores, mBI.4, Indw Cath.2

[a]OP: osteoporosis.
[b]AC: anticoagulants.
[c]BZDs: benzodiazepines.
[d]Alb: albumin.
[e]FH-3M: fall history in the past 3 months.
[f]UWA.1: use of walking assistance category 1 (wheelchair or bedridden).
[g]UWA.2: use of walking assistance category 2 (support by others or furniture).
[h]UWA.3: use of walking assistance category 3 (walker/crutches/cane).
[i]EA: emergency admission.
[j]MFS: Morse fall scale.
[k]mBI.1: modified Barthel Index scores category 1 (21-60 points).
[l]mBI.2: modified Barthel Index scores category 2 (61-90 points).
[m]mBI.3: modified Barthel Index scores category 3 (91-99 points).
[n]mBI.4: modified Barthel Index scores category 4 (100 points).
[o]Indw Cath.1: number of indwelling catheters 1 (1).
[p]Indw Cath.2: number of indwelling catheters 2 (≥2).
[q]mBI.0: modified Barthel Index scores category 0 (0-20 points).
[r]Department.2: department category 2 (department of rehabilitation medicine).
[s]CA: cancer.

XSL•FO
RenderX

Figure 2. Upset plot of the overlap of predictors selected across 5 methods. BE: bidirectional elimination; BS: backward selection; FS: forward selection, LASSO: least absolute shrinkage and selection operator; RF-RFE: random forest-recursive feature elimination; SR: stepwise regression.



Models Development and Validation

Table S5 in Multimedia Appendix 1 compares the characteristics of the training and testing sets. The training set comprised 719 (70%) older adults, while the testing set included 307 (30%). All final model predictors and fall status (yes/no) were balanced between the training and testing sets, as shown in Table 2. The AUROC for the 7 models in the testing set is shown in Figure 3B. Among these, the GBM model demonstrated the highest discrimination with an AUROC of 0.744 (95% CI 0.688 - 0.799) compared to the other 6 models. The LR model followed closely with an AUC of 0.742 (95% CI 0.685 - 0.798). The NN and RF models had the lowest AUROCs, at 0.705 (95% CI 0.646 - 0.765) and 0.715 (95% CI 0.657 - 0.772), respectively. Table 3 displays the detailed predictive performance of the 7 ML models. In the testing set, the LR model achieved the highest AUPRC of 0.570 (0.475 - 0.663), while the RF model showed the lowest AUPRC of 0.477 (0.386 - 0.580). Regarding other performance metrics, the NN models had the best sensitivity (0.931), the XGBoost model had the best specificity (0.644), and the LR model showed the highest accuracy (0.687). The calibration curves for the predictive models in the testing set are shown in Figure 3D. The LR model demonstrated the best calibration ability in the testing set. The DCA curves for the predictive models in the testing set are shown in Figure 3F. The DCA curves suggest that the 7 models have certain clinical utility, generating net benefits within the threshold range of 0 to 0.5. Considering AUROC, sensitivity, and specificity, the GBM was determined to be the best-performing model.

XSL•FO
RenderX

Table . Characteristics of the predictors in the training and testing sets.

Predictors	Training set (n=719)	Testing set (n=307)	*P* value
Fall, n (%)			>.99
No	479 (66.62)	205 (66.78)	
Yes	240 (33.38)	102 (33.22)	
Dizziness, n (%)			.93
No	705 (98.05)	302 (98.37)	
Yes	14 (1.95)	5 (1.63)	
Epilepsy, n (%)			>.99
No	706 (98.19)	302 (98.37)	
Yes	13 (1.81)	5 (1.63)	
FH-3M[a], n (%)			.99
No	660 (91.79)	281 (91.53)	
Yes	59 (8.21)	26 (8.47)	
UWA[b], n (%)			.08
No assistance	475 (66.06)	179 (58.31)	
Wheelchair or bedridden	192 (26.7)	100 (32.57)	
Support by others or furniture	10 (1.39)	3 (0.98)	
Walker/crutches/cane	42 (5.84)	25 (8.14)	
EA[c], n (%)			.64
No	701 (97.50)	297 (96.74)	
Yes	18 (2.50)	10 (3.26)	
MFS[d] (points), n (%)			.42
<45	196 (27.26)	92 (29.97)	
≥45	523 (72.74)	215 (70.03)	
mBI[e] (points), n (%)			.24
0-20	36 (5.01)	13 (4.23)	
21-60	172 (23.92)	93 (30.29)	
61-90	322 (44.78)	121 (39.41)	
91 - 99	79 (10.99)	30 (9.77)	
100	110 (15.3)	50 (16.29)	
Indw Cath[f], n (%)			.81
0	566 (78.72)	247 (80.46)	
1	95 (13.21)	38 (12.38)	
≥2	58 (8.07)	22 (7.17)	

[a]FH-3M: fall history in the past 3 months.

[b]UWA: use of walking assistance.

[c]EA: emergency admission.

[d]MFS: Morse fall scale.

[e]mBI: modified Barthel Index.

[f]Indw Cath: number of indwelling catheters.

XSL•FO
RenderX

Figure 3. Receiver operating characteristic (ROC) curves, calibration curves, and decision curve analysis (DCA) curves of different machine learning (ML) models in the training and testing sets. (A) ROC curves and area under the ROC curve (AUC) values of different ML prediction models in the training set. (B) ROC curves and AUC values of different ML prediction models in the testing set.(C) Calibration curves of different ML prediction models in the training set. (D) Calibration curves of different ML prediction models in the testing set. (E) DCA curves of different ML prediction models in the training dataset. (F) DCA curves of different ML prediction models in the testing dataset. GBM: gradient boosting machine; KNN: k-nearest neighbor; LR: logistic regression; NNET: neural network; RF: random forest; SVMs: support vector machines; XGBoost: extreme gradient boosting.

XSL•FO
RenderX

Table . The performance of 7 machine learning models for predicting falls in hospitalized older adults.

Model	AUPRC[a] (95% CI)	Sensitivity	Specificity	F_1-score	Accuracy	PPV[b]	NPV[c]	Recall
Training set								
LR[d]	0.614 (0.555 - 0.672)	0.708	0.656	0.591	0.673	0.508	0.818	0.708
SVMs[e]	0.599 (0.534 - 0.661)	0.767	0.585	0.591	0.645	0.480	0.833	0.767
GBM[f]	0.620 (0.559 - 0.681)	0.654	0.714	0.588	0.694	0.534	0.805	0.654
NN[g]	0.646 (0.590 - 0.705)	0.629	0.741	0.586	0.704	0.549	0.800	0.629
RF[h]	0.580 (0.517 - 0.643)	0.771	0.608	0.604	0.662	0.496	0.841	0.771
XGBoost[i]	0.637 (0.579 - 0.696)	0.654	0.718	0.590	0.697	0.538	0.806	0.654
KNN[j]	0.626 (0.567 - 0.685)	0.783	0.553	0.586	0.630	0.468	0.836	0.783
Testing set								
LR	0.570 (0.475 - 0.663)	0.794	0.634	0.628	0.687	0.519	0.861	0.794
SVMs	0.537 (0.437 - 0.640)	0.873	0.561	0.634	0.665	0.497	0.898	0.873
GBM	0.560 (0.464 - 0.654)	0.873	0.561	0.634	0.665	0.497	0.898	0.873
NN	0.509 (0.409 - 0.610)	0.931	0.424	0.603	0.593	0.446	0.926	0.931
RF	0.477 (0.386 - 0.580)	0.863	0.576	0.635	0.671	0.503	0.894	0.863
XGBoost	0.535 (0.435 - 0.635)	0.745	0.644	0.606	0.678	0.510	0.835	0.745
KNN	0.547 (0.446 - 0.643)	0.794	0.590	0.607	0.658	0.491	0.852	0.794

[a]AUPRC: area under the precision recall curve.

[b]PPV: positive predictive value.

[c]NPV: negative predictive value.

[d]LR: logistic regression.

[e]SVM: support vector machine.

[f]GBM: gradient boosting machine.

[g]NN: neural network.

[h]RF: random forest.

[i]XGBoost: extreme gradient boosting.

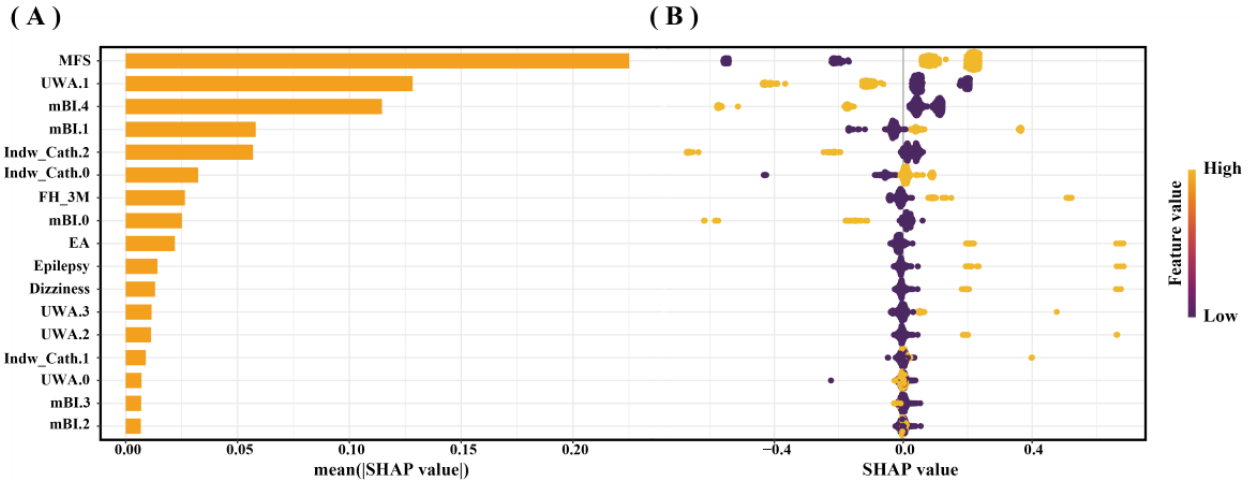
XSL•FO
RenderX

[j]KNN: k-nearest neighbor.

Interpretability Analysis

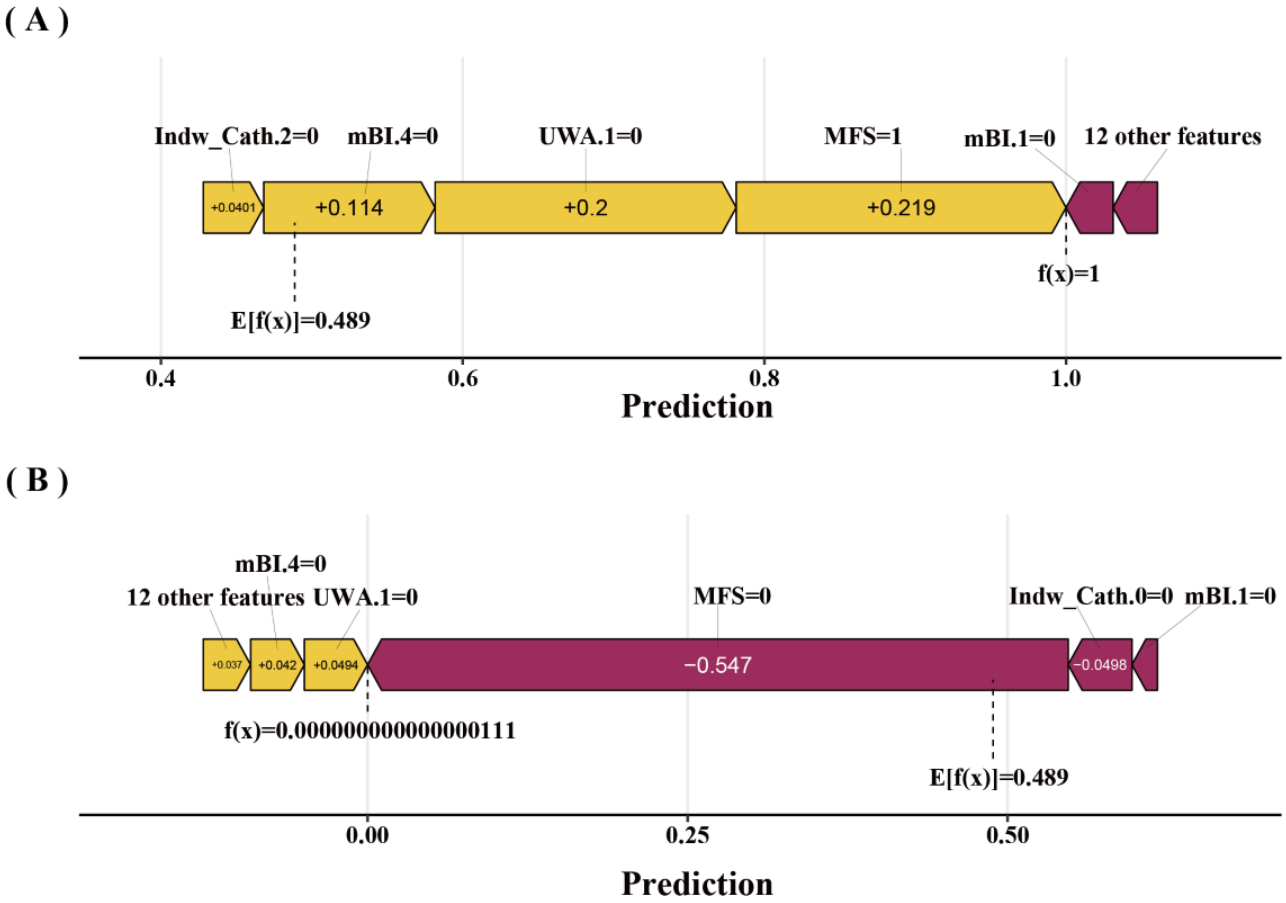
SHAP was utilized to illustrate how the features predict the occurrence of falls in old adults during hospitalization within the GBM model. Figure 4A displays the 17 features sorted by their average absolute SHAP values, and higher absolute SHAP indicates greater contribution to fall risk. Figure 4B shows the impact values and explanations of these features, and yellow dots represent high risk, while purple dots indicate low risk. An MFS score of ≥45, an mBI score that is not 100 points, an mBI score not between 0 and 20 points, having fewer than 2 indwelling tubes, a history of falls in the past 3 months, EA, epilepsy, dizziness, use of a walker/cane/crutch, requiring assistance from others/furniture for walking, and not using a wheelchair or not being bedridden are associated with a higher risk of falls in old adults during hospitalization. Beyond global SHAP interpretations, local interpretability was demonstrated. Figure 5A and B visualizes how the GBA model makes predictions about falls in older adults during hospitalization; yellow arrows indicate risk-increasing features and purple arrows risk-decreasing features. The *f(x)* values inside arrows quantify each feature's contribution. Summing these yields the model's final prediction, which is represented by the *f(x)* value outside arrows.

Figure 4. Interpretation of the gradient boosting machine model by the Shapley Additive Explanations (SHAP) method. (A) A bar summary of the most important features according to the SHAP values. (B) Summary and explanation of the most influential features. Yellow dots indicate high-risk values, and purple dots indicate low-risk values. EA: emergency admission; FH-3M: fall history in the past 3 months; Indw Cath.0: number of indwelling catheters 0 (0); Indw Cath.1: number of indwelling catheters 1 (1); Indw Cath.2: number of indwelling catheters 2 (≥2); mBI.0: modified Barthel Index scores category 0 (0-20 points); mBI.1: modified Barthel Index scores category 1 (21-60 points); mBI.2: modified Barthel Index scores category 2 (61 - 90 points); mBI.3: modified Barthel Index scores category 3 (91 - 99 points); mBI.4: modified Barthel Index scores category 4 (100 points); MFS: Morse Fall Scale; UWA.0: use of walking assistance category 0 (no assistance); UWA.1: use of walking assistance category 1 (wheelchair or bedridden); UWA.2: use of walking assistance category 2 (support by others or furniture); UWA.3: use of walking assistance category 3 (walker/crutches/cane).



XSL•FO
RenderX

Figure 5. Compositional risk of feature contributions for predicting the occurrence of falls in 2 older adults during hospitalization. Purple arrows denote factors that decrease the risk of falls for old adults during hospitalization, while yellow arrows indicate features that increase the risk. (A) An older adult with fewer than 2 indwelling catheters, a modified Barthel Index (mBI) score not equal to 100 and not within the range of 21 - 60 points, not using a wheelchair or bedridden, and a Morse Fall Scale (MFS) score ≥45 points experienced a fall during hospitalization. (B) An older adult with an mBI score not equal to 100 or within the range of 21 - 60 points, not using a wheelchair or bedridden, with the presence of indwelling catheters, and an MFS score <45 points did not experience a fall during hospitalization. Indw Cath.0: number of indwelling catheters 0 (0); Indw cath.2: number of indwelling catheters 2 (≥2); mBI.1: modified Barthel Index scores category 1 (21 - 60 points); mBI.4: modified Barthel Index scores category 4 (100 points); UWA.1: use of walking assistance category 1 (wheelchair or bedridden).



Discussion

Principal Results

In this study, we used 7 ML algorithms to predict in-hospital falls among hospitalized older adults based on clinically accessible data, including demographic characteristics, comorbidities, laboratory parameters, and medications. The GBM algorithm demonstrated the optimal predictive performance. Model interpretability was achieved at both global and individualized patient levels using SHAP [38]. The SHAP approach bridges the gap between ML models and realistic clinical decision-making, enabling health care providers to understand the model's predictive process and trust its predictive power [39].

In our study, multiple ML algorithms based on distinct principles were employed to develop predictive models. However, the predictive performance across these methods showed limited variation, consistent with prior research [20], which suggests that advanced ML algorithms generally perform well on relatively small and low-dimensional datasets. Through a comprehensive evaluation of the AUROC, AUPRC, sensitivity, and specificity, the GBM model was ultimately selected as the optimal model. Notably, LR also demonstrated competitive performance, and DCA curves indicated that the LR model could provide favorable clinical net benefit. Nevertheless, compared to LR, GBM offers distinct advantages in handling nonlinear relationships and complex data patterns [27].

Fall risk factors among hospitalized older adults encompass multiple domains, including demographic characteristics, comorbidities, laboratory parameters, and medications [40-42]. Previous studies have relied on subjective nursing documentation texts, comprehensive geriatric assessments, or environmental detection systems to develop fall prediction models for hospitalized individuals [16,21], which limits their clinical utility. Identifying predictors is a critical step in building predictive models. It is notable that the predictors identified in our study are aligned with routinely collected clinical data, ensuring practical accessibility in health care settings. Conventional approaches often select predictors using a single method, such as regression models, whereas combining multiple feature selection techniques may yield simplified models with higher generalizability [43]. Different from previous studies, we used multiple methods such as SR, RF-RFE, and LASSO to identify predictors, which is one of the advantages of our study.

XSL•FO
RenderX

Eight variables were ultimately identified: dizziness, epilepsy, fall history within the past 3 months, UWA, EA, MFS scores, mBI scores, and the number of Indw Cath. They are also key predictors in other predictive models [11,15,16,44]. Our study identified MFS scores ≥45, nonbedridden and not using a wheelchair, and scores of mBI not 100 as the 3 strongest predictors of falls in hospitalized older adults. These findings align with previous studies [11,15,16]. MFS is widely used for fall risk assessment in hospitals. Previous research has shown that MFS exhibits lower sensitivity than ML models [11,44]. Nevertheless, including it as a predictor in ML models permits the evaluation of its predictive contribution relative to other variables. MFS remains a valid predictor of falls among hospitalized patients [11]. Similarly, in our study, MFS emerged as a strong feature in the ML model.

Moreover, since patients' clinical data often include MFS scores, an integrated model that incorporates MFS can better simulate real-world decision-making, providing a more practical foundation for clinical decision. One of the key strengths of our study lies in integrating a simple, widely used assessment tool with a high-performance ML method, leveraging the advantages of both methodologies to develop and validate a simple, easily generalizable predictive model. This study found that older adults who are not bedridden or not using a wheelchair had a higher fall risk during hospitalization. This may occur as over half of falls happen during daily activities [45], whereas bedridden or wheelchair-bound patients have very low activities of daily living (ADL) ability, limiting activity engagement and thereby reducing fall risk. Similarly, patients with mBI scores >0 had higher fall risk, where higher scores indicate better ADL ability [35]. Notably, those with mBI scores <100 or 21 to 60 also showed increased risk, implying a nonlinear relationship between mBI scores and fall risk. This contradicts findings by Dormosh et al [15] and Chu et al [16] that low ADL ability predicts falls but aligns with Nagarkar et al's [45] longitudinal study linking difficulty with >4 ADL to higher fall odds. Functional decline impairs musculoskeletal integrity and body composition, reducing mobility and increasing fall risk [45-47]. However, the relationship between functional ability and fall risk in elderly patients requires further investigation. Identifying functional states associated with the highest fall risk and implementing dynamic interventions are crucial for preventing falls in this population.

Despite the growing number of ML-based clinical prediction models being developed, most studies lack interpretability of these models, limiting their clinical understanding and practical adoption. The interpretability of ML predictions requires attention from researchers so that physicians can understand, trust, and ultimately apply these predictive models to guide their clinical practice [28,38,39]. SHAP is a model-agnostic interpretation framework grounded in cooperative game theory. Its core lies in computing Shapley values to quantify the marginal contribution of each input feature to individual predictions. This approach provides consistent and locally accurate explanations for every prediction made by the model [38]. In this study, we addressed the "black-box" nature of ML models by implementing SHAP to interpret the GBM model at both global and individualized levels. This means that in a clinical setting, the model can calculate a patient's fall risk in real time and simultaneously provide the primary clinically interpretable factors contributing to that risk, thereby enabling rapid screening and informed decision-making. SHAP improves the clinical utility of prediction models, providing fall risk prediction and interpretable descriptions for older adults during hospitalization, thereby demonstrating its potential to address the "black-box" problem [28,39].

Limitations

This study has several limitations. First, the predictive model was developed using single-center retrospective data, which may introduce potential biases and limit its generalizability to other health care settings. External validation in multicenter cohorts is required to confirm broader applicability. Second, incorporating environmental variables (eg, ward layout, lighting conditions) was challenging due to constraints in single center data collection. Lastly, the exclusion of additional laboratory parameters may have overlooked potential predictors. Future research should prioritize integrating environmental variables, expanding laboratory indicators, and leveraging multicenter datasets for model development and validation.

Conclusions

In this study, multiple ML models were developed and validated using multifaceted clinical data to identify the risk of falls among hospitalized older adults. The GBM model demonstrated the optimal predictive performance. By SHAP, the clinical utility of the predictive model was significantly enhanced. In the future, this GBM fall prediction model could be integrated into the hospital EHR system as an embedded decision support module to dynamically assess fall risk among inpatients and generate real-time alerts. Simultaneously, based on the SHAP values provided by the model, the system could offer evidence to support health care providers in developing personalized intervention measures, thereby translating risk prediction into clinical actions aimed at reducing the incidence of falls in hospitalized older adults.

Funding

This study was supported by Nursing Project of Anhui Institite of Translational Medicine (No. 2024zhyx-hl-A04).

Authors' Contributions

XY, JR, and DS cooperatively designed the study, analyzed the data, and drafted the manuscript. NM conceived and designed the study. MB and MZ contributed to the revision of the manuscript. XC, YL, ZW, XD, ZW, SZ, YZ, JL, XL, and JX contributed to the acquisition of data.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
The variable list, patient characteristics, regression analysis results, and model development process.
[DOCX File, 456 KB - aging_v9i1e80602_app1.docx]

References

1. Falls. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2025-12-20]
2. Colón-Emeric CS, McDermott CL, Lee DS, Berry SD. Risk assessment and prevention of falls in older community-dwelling adults: a review. JAMA 2024 Apr 23;331(16):1397-1406. [doi: 10.1001/jama.2024.1416] [Medline: 38536167]
3. Salari N, Darvishi N, Ahmadipanah M, Shohaimi S, Mohammadi M. Global prevalence of falls in the older adults: a comprehensive systematic review and meta-analysis. J Orthop Surg Res 2022 Jun 28;17(1):334. [doi: 10.1186/s13018-022-03222-1] [Medline: 35765037]
4. Chen X, He L, Shi K, Wu Y, Lin S, Fang Y. Interpretable machine learning for fall prediction among older adults in China. Am J Prev Med 2023 Oct;65(4):579-586. [doi: 10.1016/j.amepre.2023.04.006] [Medline: 37087076]
5. Jung YS, Suh D, Kim E, Park HD, Suh DC, Jung SY. Medications influencing the risk of fall-related injuries in older adults: case-control and case-crossover design studies. BMC Geriatr 2023 Jul 22;23(1):452. [doi: 10.1186/s12877-023-04138-z] [Medline: 37481554]
6. Dykes PC, Curtin-Bowen M, Lipsitz S, et al. Cost of inpatient falls and cost-benefit analysis of implementation of an evidence-based fall prevention program. JAMA Health Forum 2023 Jan 6;4(1):e225125. [doi: 10.1001/jamahealthforum.2022.5125] [Medline: 36662505]
7. Morris ME, Webster K, Jones C, et al. Interventions to reduce falls in hospitals: a systematic review and meta-analysis. Age Ageing 2022 May 1;51(5):afac077. [doi: 10.1093/ageing/afac077] [Medline: 35524748]
8. Cho I, Kwon JM, Choe W, Cho J, Park SH, Bates DW. Under-reporting of falls in hospitals: a multisite study in South Korea. BMJ Qual Saf 2025 Jul 18;34(8):491-498. [doi: 10.1136/bmjqs-2024-017993] [Medline: 40139777]
9. Moreland B, Kakara R, Henry A. Trends in nonfatal falls and fall-related injuries among adults aged ≥65 years - United States, 2012-2018. MMWR Morb Mortal Wkly Rep 2020 Jul 10;69(27):875-881. [doi: 10.15585/mmwr.mm6927a5] [Medline: 32644982]
10. Shim S, Yu JY, Jekal S, et al. Development and validation of interpretable machine learning models for inpatient fall events and electronic medical record integration. Clin Exp Emerg Med 2022 Dec;9(4):345-353. [doi: 10.15441/ceem.22.354] [Medline: 36128798]
11. Lindberg DS, Prosperi M, Bjarnadottir RI, et al. Identification of important factors in an inpatient fall risk prediction model to improve the quality of care using EHR and electronic administrative data: a machine-learning approach. Int J Med Inform 2020 Nov;143:104272. [doi: 10.1016/j.ijmedinf.2020.104272] [Medline: 32980667]
12. Peel NM, Jones LV, Berg K, Gray LC. Validation of a falls risk screening tool derived from InterRAI acute care assessment. J Patient Saf 2021 Dec 1;17(8):e1152-e1156. [doi: 10.1097/PTS.0000000000000462] [Medline: 29360675]
13. Vratsistas-Curto A, Tiedemann A, Treacy D, Lord SR, Sherrington C. External validation of approaches to prediction of falls during hospital rehabilitation stays and development of a new simpler tool. J Rehabil Med 2018 Feb 13;50(2):216-222. [doi: 10.2340/16501977-2290] [Medline: 29260235]
14. Kawazoe Y, Shimamoto K, Shibata D, Shinohara E, Kawaguchi H, Yamamoto T. Impact of a clinical text-based fall prediction model on preventing extended hospital stays for elderly inpatients: model development and performance evaluation. JMIR Med Inform 2022 Jul 27;10(7):e37913. [doi: 10.2196/37913] [Medline: 35896017]
15. Dormosh N, Damoiseaux-Volman BA, van der Velde N, Medlock S, Romijn JA, Abu-Hanna A. Development and internal validation of a prediction model for falls using electronic health records in a hospital setting. J Am Med Dir Assoc 2023 Jul;24(7):964-970. [doi: 10.1016/j.jamda.2023.03.006] [Medline: 37060922]
16. Chu WM, Kristiani E, Wang YC, et al. A model for predicting fall risks of hospitalized elderly in Taiwan-a machine learning approach based on both electronic health records and comprehensive geriatric assessment. Front Med (Lausanne) 2022;9:937216. [doi: 10.3389/fmed.2022.937216] [Medline: 36016999]
17. Satoh M, Miura T, Shimada T. Development and evaluation of a simple predictive model for falls in acute care setting. J Clin Nurs 2023 Sep;32(17-18):6474-6484. [doi: 10.1111/jocn.16680] [Medline: 36899476]
18. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods 2018 Apr;15(4):233-234. [doi: 10.1038/nmeth.4642] [Medline: 30100822]
19. González-Castro A, Leirós-Rodríguez R, Prada-García C, Benítez-Andrades JA. The applications of artificial intelligence for assessing fall risk: systematic review. J Med Internet Res 2024 Apr 29;26:e54934. [doi: 10.2196/54934] [Medline: 38684088]

XSL•FO
RenderX

20. Liu C, Zhang K, Yang X, et al. Development and validation of an explainable machine learning model for predicting myocardial injury after noncardiac surgery in two centers in China: retrospective study. JMIR Aging 2024 Jul 26;7:e54872. [doi: 10.2196/54872] [Medline: 39087583]
21. Adeli V, Korhani N, Sabo A, et al. Ambient monitoring of gait and machine learning models for dynamic and short-term falls risk assessment in people with dementia. IEEE J Biomed Health Inform 2023 Jul;27(7):3599-3609. [doi: 10.1109/JBHI.2023.3267039] [Medline: 37058371]
22. Kim DW, Seo J, Kwon S, et al. Predicting in-hospital fall risk using machine learning with real-time location system and electronic medical records. J Cachexia Sarcopenia Muscle 2025 Feb;16(1):e13713. [doi: 10.1002/jcsm.13713] [Medline: 39994910]
23. Lim ZK, Connie T, Goh MKO. Fall risk prediction using temporal gait features and machine learning approaches. Front Artif Intell 2024;7:1425713. [doi: 10.3389/frai.2024.1425713] [Medline: 39263525]
24. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology, 3rd edition: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008. URL: https://books.google.com/books/about/Modern_Epidemiology.html?id=Z3vjT9ALxHUC [accessed 2025-12-20]
25. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. I. Principles. Am J Epidemiol 1992 May 1;135(9):1019-1028. [doi: 10.1093/oxfordjournals.aje.a116396] [Medline: 1595688]
26. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics 2012 Jan 1;28(1):112-118. [doi: 10.1093/bioinformatics/btr597] [Medline: 22039212]
27. Qi W, Wang Y, Wang Y, et al. Prediction of postpartum depression in women: development and validation of multiple machine learning models. J Transl Med 2025 Mar 7;23(1):291. [doi: 10.1186/s12967-025-06289-6] [Medline: 40055720]
28. Song Y, Yuan Q, Liu H, Gu K, Liu Y. Machine learning algorithms to predict mild cognitive impairment in older adults in China: a cross-sectional study. J Affect Disord 2025 Jan 1;368:117-126. [doi: 10.1016/j.jad.2024.09.059] [Medline: 39271065]
29. Sadat-Hosseini M, Arab MM, Soltani M, Eftekhari M, Soleimani A, Vahdati K. Predictive modeling of Persian walnut (Juglans regia L.) in vitro proliferation media using machine learning approaches: a comparative study of ANN, KNN and GEP models. Plant Methods 2022 Apr 11;18(1):48. [doi: 10.1186/s13007-022-00871-5] [Medline: 35410228]
30. Butler LR, Chen KA, Hsu J, Kapadia MR, Gomez SM, Farrell TM. Predicting readmission after bariatric surgery using machine learning. Surg Obes Relat Dis 2023 Nov;19(11):1236-1244. [doi: 10.1016/j.soard.2023.05.025] [Medline: 37455158]
31. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010 Jan;21(1):128-138. [doi: 10.1097/EDE.0b013e3181c30fb2] [Medline: 20010215]
32. Wang Z, Gu Y, Huang L, et al. Construction of machine learning diagnostic models for cardiovascular pan-disease based on blood routine and biochemical detection data. Cardiovasc Diabetol 2024 Sep 28;23(1):351. [doi: 10.1186/s12933-024-02439-0] [Medline: 39342281]
33. Manolis AA, Manolis TA, Melita H, Mikhailidis DP, Manolis AS. Low serum albumin: a neglected predictor in patients with cardiovascular disease. Eur J Intern Med 2022 Aug;102:24-39. [doi: 10.1016/j.ejim.2022.05.004] [Medline: 35537999]
34. Morse JM, Black C, Oberle K, Donahue P. A prospective study to identify the fall-prone patient. Soc Sci Med 1989;28(1):81-86. [doi: 10.1016/0277-9536(89)90309-2] [Medline: 2928815]
35. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989;42(8):703-709. [doi: 10.1016/0895-4356(89)90065-6] [Medline: 2760661]
36. Jabłońska B, Mrowiec S. Nutritional status and its detection in patients with inflammatory bowel diseases. Nutrients 2023 Apr 20;15(8):1991. [doi: 10.3390/nu15081991] [Medline: 37111210]
37. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. J Clin Nurs 2005 Aug;14(7):798-804. [doi: 10.1111/j.1365-2702.2005.01121.x] [Medline: 16000093]
38. Kang CW, Yan ZK, Tian JL, Pu XB, Wu LX. Constructing a fall risk prediction model for hospitalized patients using machine learning. BMC Public Health 2025 Jan 20;25(1):242. [doi: 10.1186/s12889-025-21284-8] [Medline: 39833780]
39. Shen L, Jin Y, Pan AX, et al. Machine learning-based predictive models for perioperative major adverse cardiovascular events in patients with stable coronary artery disease undergoing noncardiac surgery. Comput Methods Programs Biomed 2025 Mar;260:108561. [doi: 10.1016/j.cmpb.2024.108561] [Medline: 39708562]
40. Machado-Duque ME, Camacho-Arteaga L, Sabaté M, Vidal-Guitart X, Machado-Alba JE. Falls in hospitalized older adults and the use of fall risk-increasing drugs and anticholinergic medications in Colombia: a case control study. Front Pharmacol 2024;15:1369200. [doi: 10.3389/fphar.2024.1369200] [Medline: 39021833]
41. Liang XZ, Chai JL, Li GZ, et al. A fall risk prediction model based on the CHARLS database for older individuals in China. BMC Geriatr 2025 Mar 13;25(1):170. [doi: 10.1186/s12877-025-05814-y] [Medline: 40082807]
42. Mao A, Su J, Ren M, Chen S, Zhang H. Risk prediction models for falls in hospitalized older patients: a systematic review and meta-analysis. BMC Geriatr 2025 Jan 14;25(1):29. [doi: 10.1186/s12877-025-05688-0] [Medline: 39810076]
43. Yang Q, Cheng H, Qin J, et al. A machine learning-based Preclinical Osteoporosis Screening Tool (POST): model development and validation study. JMIR Aging 2023 Nov 8;6:e46791. [doi: 10.2196/46791] [Medline: 37986117]

XSL•FO
RenderX

44. Yang A, Shi M, Lau ESH, et al. A novel electronic-health-record based, machine-learning model to predict 1-year risk of fall hospitalisation in older adults: a Hong Kong territory-wide cohort and modelling study. Age Ageing 2025 Aug 29;54(10):afaf285. [doi: 10.1093/ageing/afaf285] [Medline: 41066674]
45. Nagarkar A, Kulkarni S. Association between daily activities and fall in older adults: an analysis of longitudinal ageing study in India (2017-18). BMC Geriatr 2022 Mar 14;22(1):203. [doi: 10.1186/s12877-022-02879-x] [Medline: 35287596]
46. Jehu DA, Davis JC, Falck RS, et al. Risk factors for recurrent falls in older adults: a systematic review with meta-analysis. Maturitas 2021 Feb;144:23-28. [doi: 10.1016/j.maturitas.2020.10.021] [Medline: 33358204]
47. Dipietro L, Campbell WW, Buchner DM, et al. Physical activity, injurious falls, and physical function in aging: an umbrella review. Med Sci Sports Exerc 2019 Jun;51(6):1303-1313. [doi: 10.1249/MSS.0000000000001942] [Medline: 31095087]

Abbreviations

ADLs: activities of daily living
AUPRC: area under the precision-recall curve
AUROC: area under the receiver operating characteristic curve
DCA: decision curve analysis
EA: emergency admission
EHR: electronic health record
FH-3M: fall history in the past 3-months
GBM: gradient boosting machine
Indw Cath: indwelling catheters
KNN: k-nearest neighbor
LASSO: least absolute shrinkage and selection operator
LR: logistic regression
mBI: modified Barthel Index
MFS: Morse Fall Scale
ML: machine learning
NN: neural network
NRS 2002: Nutritional Risk Screening 2002
RF: random forest
RF-RFE: random forest-recursive feature elimination
SHAP: Shapley Additive Explanations
SR: stepwise regression
STRATIFY: St. Thomas's risk assessment tool in falling elderly inpatients
TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis
UWA: use of walking assistance
XGBoost: extreme gradient boosting

Edited by M Gray; submitted 14.Jul.2025; peer-reviewed by S Sharma, S Shim; revised version received 13.Dec.2025; accepted 16.Dec.2025; published 05.Jan.2026.

Please cite as:
Yang X, Ren J, Su D, Bao M, Zhang M, Chen X, Li Y, Wang Z, Dai X, Wei Z, Zhang S, Zhang Y, Li J, Li X, Xu J, Mo N
Development and Validation of Machine Learning Models for Predicting Falls Among Hospitalized Older Adults: Retrospective Cross-Sectional Study
JMIR Aging 2026;9:e80602
URL: https://aging.jmir.org/2026/1/e80602
doi:10.2196/80602

© Xiyao Yang, Juan Ren, Dan Su, Manzhen Bao, Miao Zhang, Xiaoming Chen, Yanhua Li, Zonggui Wang, Xiujing Dai, Zengzeng Wei, Shuiyu Zhang, Yuxin Zhang, Juan Li, Xiaolin Li, Junjin Xu, Nan Mo. Originally published in JMIR Aging (https://aging.jmir.org), 5.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Global, Regional, and National Burden of Falls Among Older Adults Aged 65 Years and Above: Secondary Data Analysis of the Global Burden of Disease Study 2021

Ze Zhang[1], BM; Yingying Diao[1], BSc; Mingwang Fu[1], BM; Wantong Han[1], BM; Haoran Zhou[1], BM; Ru Fan[2], MPH; Biyun Xu[2], MSc, PhD; Bingwei Chen[1,3], MSc, PhD

[1]Department of Epidemiology and Biostatistics, School of Public Health, Southeast University, Southeast University Dingjiaqiao Campus, No. 87 Dingjiaqiao, Gulou District, Nanjing, China

[2]Medical Statistics and Analysis Center, Nanjing Drum Tower Hospital, Nanjing, China

[3]Key Laboratory of Environmental Medicine Engineering, Ministry of Education, School of Public Health, Southeast University, Nanjing, China

Corresponding Author:
Bingwei Chen, MSc, PhD
Department of Epidemiology and Biostatistics, School of Public Health, Southeast University, Southeast University Dingjiaqiao Campus, No. 87 Dingjiaqiao, Gulou District, Nanjing, China

Abstract

Background: Falls are a leading cause of injury, disability, and death among older adults, posing significant public health challenges. However, comprehensive global analyses of fall-related burdens in older populations remain scarce.

Objective: This study aimed to explore the patterns and distribution of the global, regional, and national burden of falls among adults aged 65 years and older.

Methods: Data from the Global Burden of Disease study 2021 were used to assess the overall, disability, and mortality burden of falls among adults aged 65 years and older. Age-standardized rates of disability-adjusted life years (DALYs), years lived with disability (YLDs), and years of life lost were calculated to compare burdens across countries. Health inequalities were evaluated via the slope index of inequality and the concentration index. Frontier analysis identified optimal burden levels by sociodemographic index (SDI). Bayesian age-period-cohort models projected trends up to 2050.

Results: DALY age-standardized rates showed a U-shaped distribution across SDI regions: lower-SDI countries faced higher mortality burdens, while higher-SDI countries had elevated disability burdens. Despite an increase in absolute overall burden inequality from 1990 to 2021, absolute inequalities in YLDs and years of life lost declined, with DALYs and YLDs exhibiting relatively more balanced distributions. Frontier analysis pinpointed countries with the greatest burden reduction potential. Projections suggest decreasing overall and mortality burdens by 2050 but rising disability burdens.

Conclusions: Higher- and lower-SDI countries face distinct fall-related challenges. Reducing cross-national health inequalities and closing gaps between the observed burden and the optimal burden level achievable at a similar SDI level are critical. Despite projected declines in the overall burden (DALYs), the rising disability burden (YLDs) could present evolving challenges, potentially underscoring the importance of proactive preparedness.

(JMIR Aging 2026;9:e73802) doi:10.2196/73802

KEYWORDS

falls; older adults; global burden of disease; mortality burden; disability burden; health inequality; projection

Introduction

Falls are one of the leading causes of injury, disability, and death among older adults, posing significant challenges to both public health and health care systems worldwide. As global populations age, the burden of falls is expected to increase, particularly among individuals aged 65 years and older. According to the Global Burden of Disease (GBD) study 2021, the age-standardized rate (ASR) of fall incidence was 6198.42 per 100,000 population in 2021, while the ASR of fall-related mortality was 78.36 per 100,000 population among older adults [1,2]. Furthermore, the World Health Organization reports that approximately 28% to 35% of older adults experience falls each year, resulting in outcomes ranging from minor injuries to severe fractures and mortality [3,4]. Beyond the immediate physical consequences, falls are linked to long-term disability, loss of independence, psychological trauma, and increased health care costs, making them a critical concern for both individuals and society at large [5,6].

XSL•FO
RenderX

In light of this growing issue, understanding the global, regional, and national burden of falls has become increasingly important. The GBD study 2021 offers a robust framework for quantifying the impact of diseases and injuries worldwide, including falls. While falls among older adults have been widely studied, many of the existing studies are limited in scope—either focusing on specific regions or countries or addressing only certain aspects of mental and physical effects [7-10]. A study based on the GBD study discussed the mortality rate of falls across 59 high-income and upper-middle–income countries [11]. However, significant gaps persist in comprehensively understanding the global-scale mortality and disability burdens across diverse populations and regions, as well as the underlying socioeconomic, environmental, and health care–related factors driving these disparities.

This study aimed to examine the global, regional, and national burden of falls among individuals aged 65 years and older using the latest data from the GBD study 2021. By analyzing variations in fall burdens across countries or regions, the research sought to identify current and future challenges across different development levels and geographic locations.

Methods

Study Data

This study was a secondary data analysis primarily using the GBD study 2021 database [1,2]. The GBD study 2021, led by the Institute for Health Metrics and Evaluation at the University of Washington, provides comprehensive global health data, covering 369 diseases, injuries, and risk factors. Data on the burden of falls from 1990 to 2021 were retrieved using the GBD Results Tool across 7 age groups: 65 to 69 years, 70 to 74 years, 75 to 79 years, 80 to 84 years, 85 to 89 years, 90 to 94 years, and 95 years and above [12]. Demographic data from 1990 to 2050 were acquired from the GBD study database and a projection of the population [13,14]. Additionally, the sociodemographic index (SDI) was accessed from the GBD study database to evaluate the development status of the included countries and territories [15].

Definition

The disability-adjusted life years (DALYs) represent the total number of years lost due to ill-health, disability, or early death. This study selected DALYs as the measure of overall burden of falls, with years lived with disability (YLDs) measuring burden of disability and years of life lost (YLLs) measuring burden due to premature mortality. ASR is a statistical measure used to compare rates of health metrics (per 100,000 people) across different populations, adjusting for variations in age distribution. By standardizing the measures using the GBD study world population age standard, the rates of DALYs, YLDs, and YLLs become comparable across countries or regions. The SDI is a composite measure used to summarize a country or region's sociodemographic development; it ranges from 0 to 1, where higher values indicate greater sociodemographic development [16,17].

Statistical Analysis

Joinpoint regression analysis was performed using the Windows command-line (batch/callable) version of Joinpoint Regression Software (Surveillance Research Program, National Cancer Institute) to acquire the annual percentage change (APC) and average APC of ASRs [18-20]. This study calculated 2 key indexes of health inequality: the slope index of inequality (SII) and the concentration index, which reflect absolute and relative health inequities, respectively. The SII measures absolute inequality by quantifying the difference in indicator values between the most disadvantaged and most advantaged subgroups considering the overall population distribution. It is derived from a regression model that accounts for the cumulative distribution of the population. Larger SII values indicate greater disparities between these groups. The concentration index is a relative measure that assesses the concentration of a health indicator across subgroups. It ranges from −1 to 1, with 0 indicating no inequality. Positive values suggest concentration in advantaged groups, while negative values indicate concentration among disadvantaged groups. The further the concentration index is from 0, the greater the level of inequality [21]. Frontier analysis was conducted to identify the optimal burden levels that countries could achieve based on their corresponding development statuses [22]. A Spearman rank correlation analysis was conducted to evaluate the relationship between the burden of falls and aging-related indicators among older adults. Bayesian age-period-cohort models and integrated nested Laplace approximations using the R packages *BAPC* (version 0.0.36) and *INLA* (version 24.06.27; R Foundation for Statistical Computing) were used to project the future trends in the burden of falls among older adults [23,24].

Ethical Considerations

This study was a secondary analysis of publicly available, deidentified, aggregated data from the GBD study 2021, used in accordance with the relevant data use terms. No additional ethics approval was sought for this study because it was based exclusively on publicly available, anonymized data and involved no direct contact with human participants. This study involved no collection of, or access to, identifiable personal information; therefore, additional informed consent was not required. All analyses were conducted using anonymized, publicly available data, ensuring privacy and confidentiality. No participants were recruited, and no compensation was provided.

Results

Spatiotemporal Patterns of Falls Among Older Adults

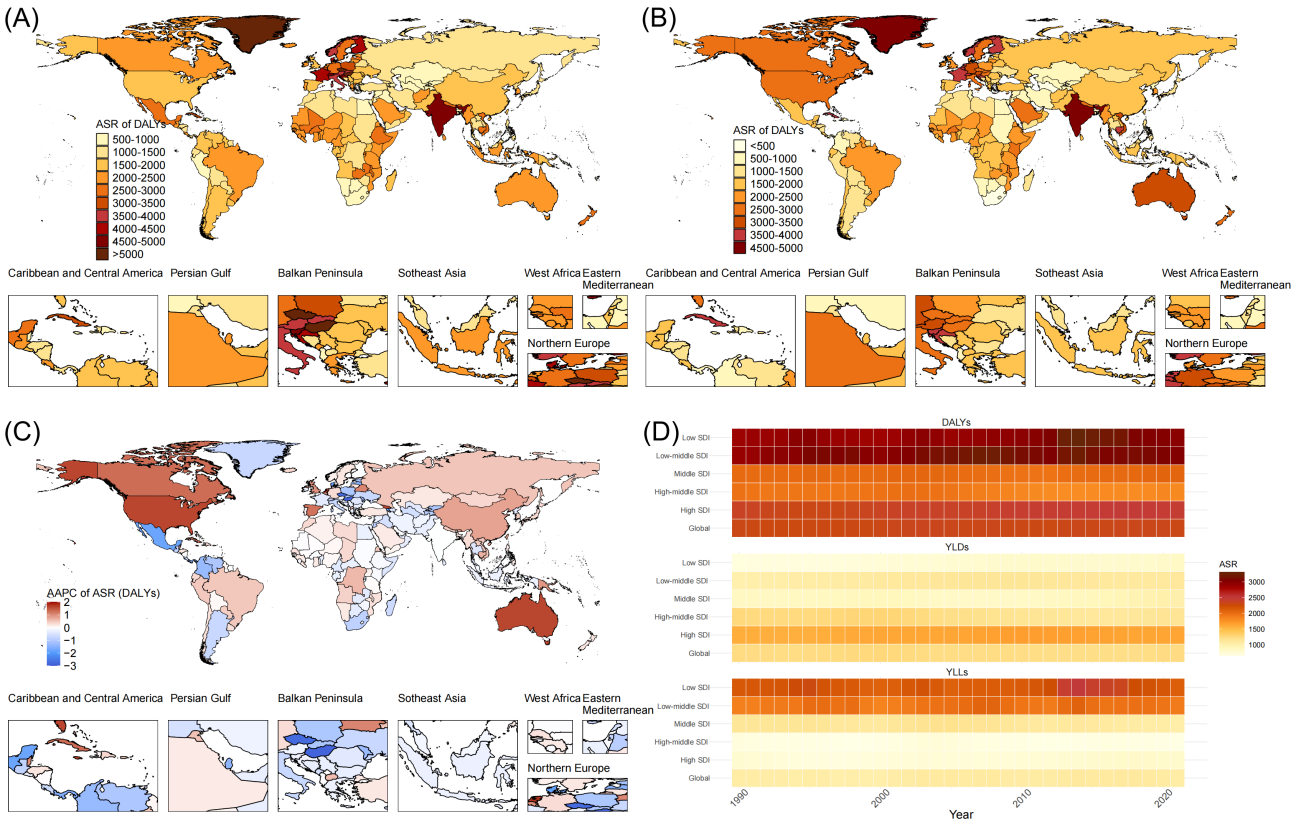
In the initial analysis, it was found that ASRs of mortality and incidence due to falls varied among countries with different SDIs (Figures S1-S3 in Multimedia Appendix 1 and Tables S1-S4 in Multimedia Appendix 2). Higher-SDI countries exhibited lower ASRs for mortality and higher ASRs for fall incidence, whereas lower-SDI countries exhibited more falls and higher mortality rates among older adults. To gain a comprehensive understanding, we used DALYs, YLDs, and YLLs as indicators to assess the overall burden, disability burden, and mortality burden of falls among individuals aged

XSL•FO
RenderX

65 years and above (Tables S5-S10 in Multimedia Appendix 2).

From 1990 to 2021, the overall burden of falls among older adults remained consistently high in India, Australia, Europe, North America, and Africa (Figures 1A and 1B). Specifically, North America and Australia exhibited an increasing trend in their overall burden, while other regions maintained stable levels or experienced only minor increases (Figure 1C). Regarding disability and mortality burdens, regions with low and low-middle SDI exhibited higher mortality burdens, whereas high-SDI regions exhibited greater disability burdens. The overall burden of falls among older adults followed a U-shaped distribution, with middle- and high-middle–SDI regions experiencing the lowest burden and low-, low-middle–, and high-SDI regions facing heavier overall burdens (Figure 1D).

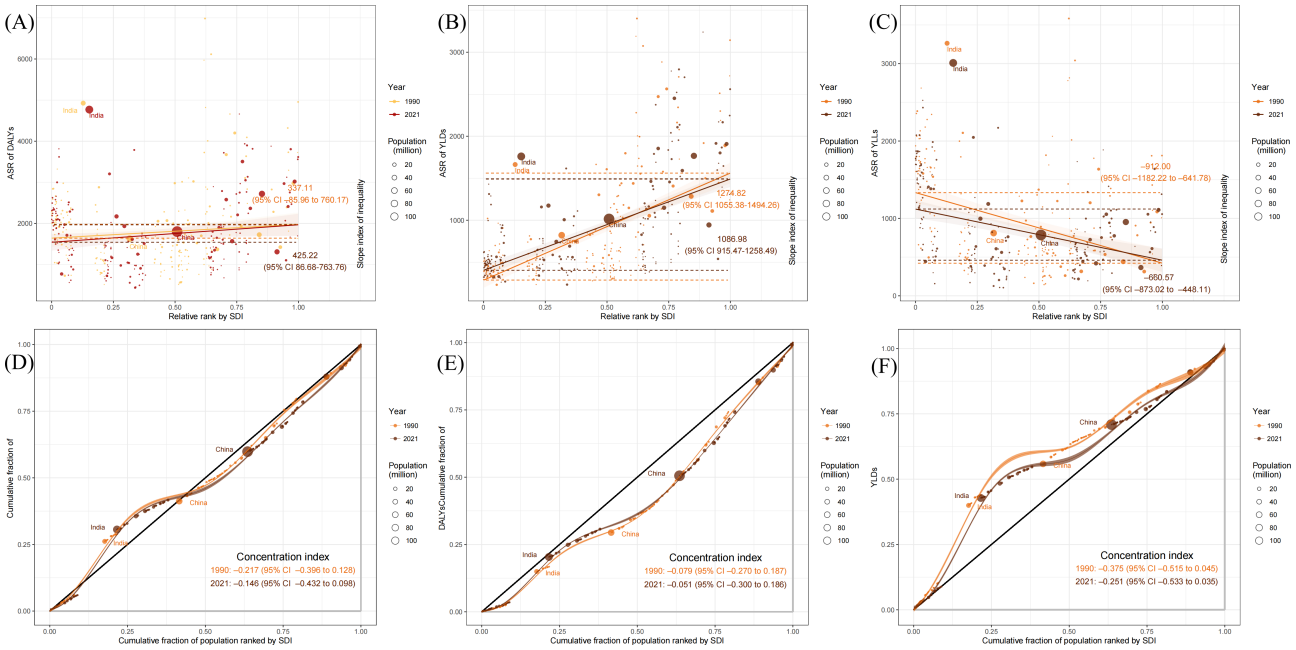
Figure 1. Spatiotemporal patterns of burden of falls among older adults: (A) age-standardized rate (ASR) of disability-adjusted life years (DALYs) in 1990; (B) ASR of DALYs in 2021; (C) average annual percentage change (AAPC) in the ASR of DALYs from 1990 to 2021; and (D) heat map of the ASR of DALYs, years lived with disability (YLDs), and years of life lost (YLLs) at the global and sociodemographic index (SDI) region levels in 2021.



Health Inequality Analysis

The SII was used to quantify absolute differences in ASRs (rates) of burden indicators between the most disadvantaged and most advantaged countries. From 1990 to 2021, the SII for DALYs increased from 337.11 to 425.22, indicating a widening of absolute overall health inequality (Figure 2A). In contrast, the SII for YLDs decreased from 1274.84 to 1086.98, and the SII for YLLs decreased from −912.00 to −660.57, suggesting a reduction in the absolute inequality of both disability and mortality burdens (Figures 2B and 2C). Consistent with the patterns presented in Figure 1D, these metrics reveal distinct underlying drivers for disability and mortality: the positive SII for YLDs confirms that the disability burden was disproportionately concentrated in higher-SDI countries, while the negative SII for YLLs aligns with our finding that the mortality burden remained more pronounced in lower-SDI regions.

XSL•FO
RenderX



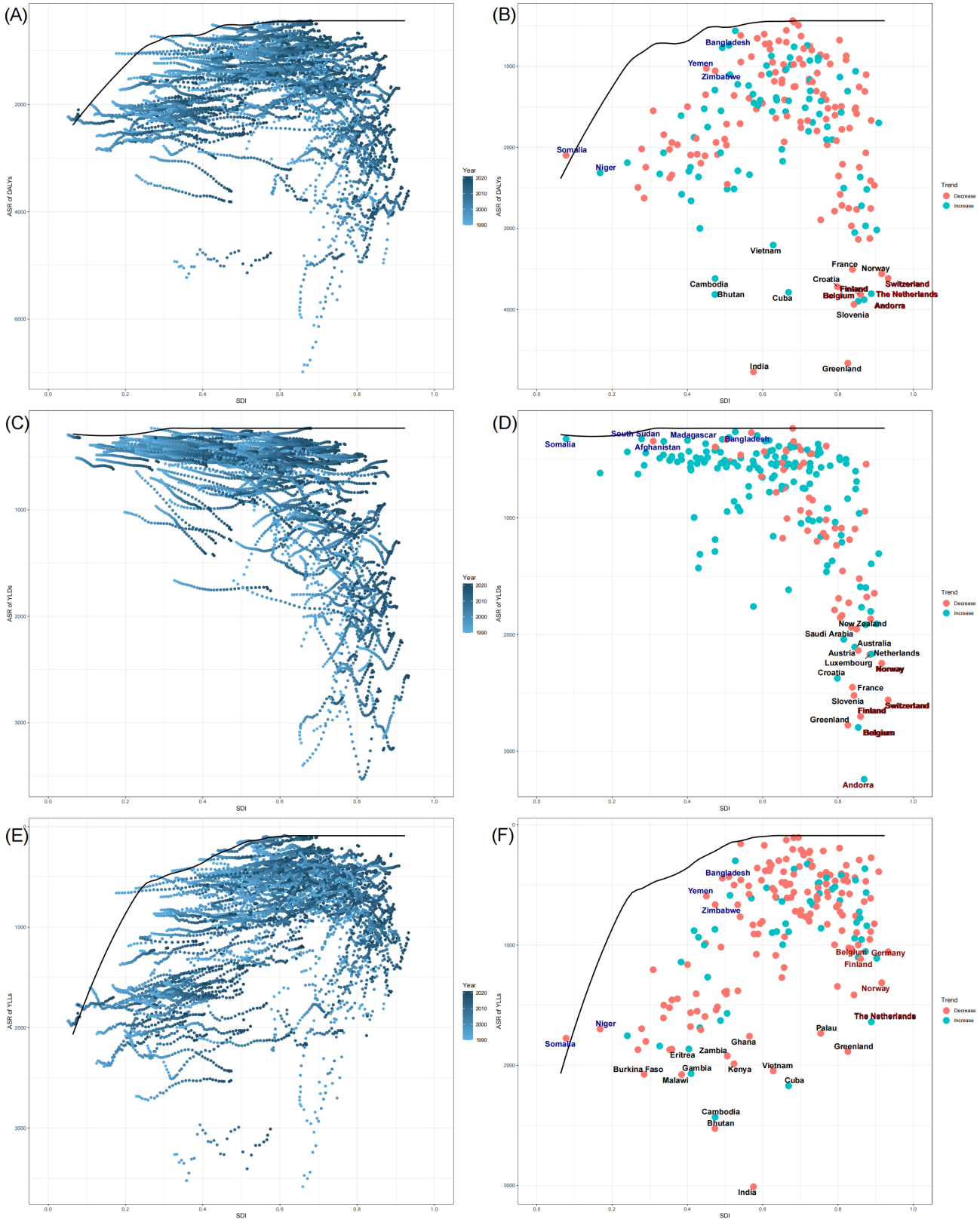
Distinct from the SII, which compares absolute inequality at the country level, the concentration index evaluates relative health inequality by analyzing the distribution of burden indicators (total numbers) across the global population ranked by SDI. Between 1990 and 2021, the concentration index for DALYs shifted from −0.217 to −0.146, indicating a reduction in relative inequality. For YLDs, the concentration index values remained close to 0 (−0.079; *P*=.50; and −0.051; *P*=.68, respectively), suggesting that the disability burden was distributed relatively evenly across the global population [25]. Meanwhile, the concentration index for YLLs increased from −0.375 to −0.251, reflecting a narrowing of relative inequality in mortality. Notably, the negative concentration indexes for both DALYs and YLLs indicate that, despite improvements over time, the overall and mortality burdens remained disproportionately concentrated in lower-SDI populations.

Frontier Analysis

Frontier analysis was conducted to determine the lowest burden that could be potentially achieved based on development status. In Figures 3A, 3C, and 3E, the scatterplot illustrates the relationship between SDI and the ASRs of DALYs, YLDs, YLLs (the color change from light blue [1990] to dark blue [2021]), indicating a shift in their distribution over time. The black frontier line represents the efficiency frontier, which delineates the lowest possible burden of DALYs, YLDs, and YLLs for a given SDI level. The distance from the frontier represents the gap between the observed burden and the potentially achievable burden of falls in a given country or territory considering its SDI. For ASRs of DALYs, countries with both low and high SDIs experienced a higher burden, while those with middle SDI levels had the lowest DALYs. In addition, for ASRs of YLDs, countries with higher SDIs generally exhibited higher disability burdens. High-SDI countries showed a declining or stable trend, while low-SDI countries exhibited an increasing trend in YLDs. In contrast, for ASRs of YLLs, most curves showed a stable or slight decline, and countries with higher SDIs generally exhibited lower mortality burdens.

XSL•FO
RenderX

Figure 3. Frontier analysis based on sociodemographic index (SDI) and age-standardized rates (ASRs) of disability-adjusted life years (DALYs), years lived with disability (YLDs), and years of life lost (YLLs) in 204 countries and territories: (A-B) ideal ASRs of DALYs corresponding to the SDI, (C-D) ideal ASRs of YLDs corresponding to the SDI, and (E-F) ideal ASRs of YLLs corresponding to the SDI.



In Figures 3B, 3D, and 3F, the scatterplots show the relationship between SDI and the ASRs of DALYs, YLDs, and YLLs across different countries in 2021. The colors indicate the trend of ASRs between 1990 and 2021, where red dots represent countries with a decrease in ASRs, while blue dots indicate an increase. High-SDI countries (SDI>0.85) with the top 5 largest distances are marked in red. In contrast, among countries with an SDI below 0.50, those performing best are marked in blue. Additionally, the top 15 countries and territories that are furthest from the frontier are colored in black.

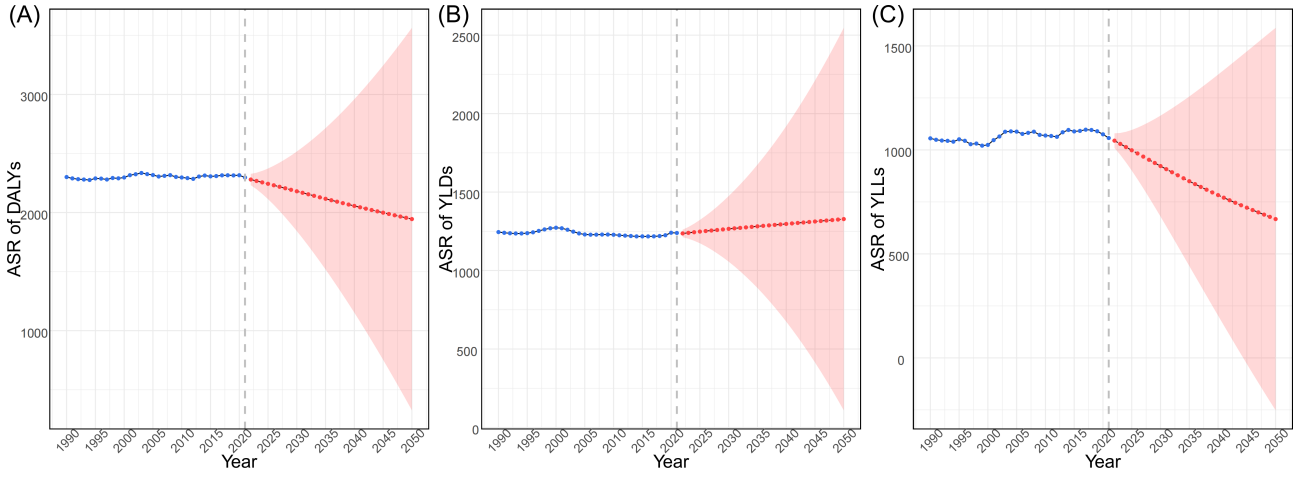
XSL•FO
RenderX

Among the ASRs of DALYs, the countries furthest from the frontier line include India, Greenland, Slovenia, Belgium, and Andorra. For ASRs of YLDs, the countries furthest from the frontier line include Andorra, Belgium, Greenland, Finland, and Switzerland. Regarding ASRs of YLLs, the countries furthest from the frontier line include India, Bhutan, Cambodia, Cuba, and Vietnam. The frontier fit line for ASRs of YLDs is flatter than those of other burden indicators, denoting that the ideal ASRs of YLDs vary little across SDI levels.

Projections of Burden of Falls Among Older Adults

This study projected the global ASRs of DALYs, YLDs, and YLLs over the next 3 decades by integrating demographic forecasts from 2020 to 2050 using the Bayesian age-period-cohort model (Figure 4). The blue dots in the figure illustrate actual data from 1990 to 2021, and the red dots represent forecasts for the period spanning 2022 to 2050. Overall, the global burden exhibits a declining trend. However, the ASR of YLDs shows an upward trend, indicating that the disability burden may become the future challenge in terms of improving falls among older adults.

Figure 4. Projection of age-standardized rates (ASRs) of disability-adjusted life years (DALYs), years lived with disability (YLDs), and years of life lost (YLLs) up to 2050.



Discussion

Principal Findings

In this study, we used DALYs, YLDs, and YLLs to evaluate the overall, disability, and mortality burden of falls among individuals aged 65 years and above. APCs and average APCs of ASRs were calculated to provide insights into trends in burden from 1990 to 2021. Our findings demonstrate that the burden of falls exhibited distinct patterns according to development status. Specifically, YLDs (disability burden) were more prevalent in high-SDI regions, whereas YLLs (mortality burden) were more concentrated in low- and low-middle–SDI regions, reflecting a heavier mortality burden. Consequently, the overall burden (DALYs) followed a U-shaped distribution, with burdens being comparatively more severe at both development extremes.

The observed variations in fall-related burden across geographical regions appear to mirror the broader patterns observed across different levels of socioeconomic development. In high-SDI regions such as Australia, Europe, and North America, the high burdens are primarily driven by older population structures [26,27]. However, beyond demographics, environmental and lifestyle factors play a significant role. Higher rates of urbanization and sedentary lifestyles in these regions may contribute to age-related sarcopenia and impaired balance [28-30]. While infrastructure in high-income settings is generally superior, the urban built environment introduces unique hazards; complex public transportation systems, high-speed escalators, and dense pedestrian traffic can pose challenges for older adults with diminished reaction times [31-33]. Additionally, seasonal climatic factors significantly modulate this burden in high-latitude regions, where winter ice and snow create hazardous conditions that drive seasonal surges in fall-related fractures [34-36]. Furthermore, while health systems in these regions often demonstrate high proficiency in acute trauma surgery, potentially reducing short-term mortality, the elevated YLDs potentially point toward a relative gap in specialized geriatric rehabilitation [37-40]. This indicates that, although life-saving interventions are effective, the current infrastructure for long-term fall prevention and functional recovery might not yet fully keep pace with the needs of an increasing aging population [41-43].

In contrast, the burden observed in countries such as India and parts of the African region appears to reflect a dual pattern of persistent mortality alongside rising disability, potentially shaped by an interplay of systemic, environmental, and cultural factors [44-47]. Structurally, health system capacity may represent a significant challenge; limited emergency medical services and a scarcity of geriatric specialists are potentially associated with suboptimal clinical outcomes and higher case fatality rates following fall injuries [48]. Environmentally, natural topography and infrastructure limitations appear to play a notable role, particularly in rural settings. Older adults often navigate uneven, unpaved terrain and steep slopes, frequently without the aid of stabilized walkways or safety railings [49,50]. These risks may be further compounded by insufficient lighting in both communal and domestic spaces, which can hinder navigation for individuals with age-related visual impairments [51,52]. Additionally, domestic environments in these regions may not

XSL•FO
RenderX

always feature age-friendly modifications such as grab bars or standardized step heights [53]. Culturally, while the traditional reliance on informal family caregiving is significant, access to the specialized medical and physiotherapeutic knowledge required for effective postfall rehabilitation may be limited [54-56]. Without professional guidance on secondary prevention, these informal care structures, despite their supportive intent, might correlate with challenges in preventing recurrent falls and managing long-term disability [57,58].

The observed epidemiological transition in fall-related burdens is further elucidated by the health inequality analysis, which reveals distinct patterns between absolute gaps (country-level rates) and relative distributions (population-level volume). Regarding absolute inequality, the positive SII for DALYs and YLDs indicates that higher-SDI countries bear a greater absolute burden per 100,000 population, particularly in terms of disability. This is likely attributable to the fact that higher-income countries possess more advanced medical and nursing resources, ensuring timely intervention and survival after a fall. While these resources successfully reduce the mortality burden (YLLs), the combination of survival following injury and a higher proportion of older individuals leads to a more pronounced disability burden. Conversely, the negative SII for YLLs confirms that mortality rates remain higher in lower-income regions, where limited medical infrastructure often results in higher fatality rates following falls. From a temporal perspective, the period from 1990 to 2021 witnessed a narrowing of the absolute gaps in both disability (YLDs) and mortality (YLLs) between the most advantaged and least advantaged countries. It was observed that the absolute gap in the overall burden (DALYs) expanded during this time frame; however, this divergence should be interpreted with caution. It appears to be linked to the reduction in age-standardized DALY rates in lower-SDI countries, suggesting that the gap may be widening due to unequal rates of improvement rather than worsening outcomes [59,60].

While absolute gaps indicate disparities in rates between the highest- and lowest-income countries, the analysis of relative inequality (concentration index) provides a complementary perspective by examining how the total burden (absolute number) is distributed across the global population ranked by SDI. Notably, a near-zero concentration index for YLDs suggests a relatively even distribution of disability, and the negative concentration indexes for DALYs and YLLs underscore a persistent health inequity. This implies that, despite higher disability rates in the highest-income countries, the global volume of mortality and overall burden remains predominantly concentrated among populations in lower-income regions. This concentration of burden appears to be driven by 2 primary factors: first, lower-SDI countries often have significantly larger population sizes, meaning that even a moderate incidence rate translates into a much larger absolute number of affected individuals, and second, the dominant contribution of mortality (YLLs) in disadvantaged regions is often exacerbated by inadequate emergency response and a lack of specialized geriatric rehabilitation. Therefore, while high-SDI countries are increasingly challenged by the long-term care of survivors, the most disadvantaged populations continue to carry the heaviest cumulative weight of the burden due to high fatality rates and large population volumes.

However, this global landscape has evolved favorably over the past 3 decades, as indicated by the narrowing concentration indexes for DALYs and YLLs between 1990 and 2021, suggesting a more balanced global distribution of the fall burden. This shift may be partly attributable to the socioeconomic advancements in populous emerging economies, most notably China and India. As these countries experienced steady increases in their SDI, their massive populations gradually shifted toward higher-SDI quintiles. This demographic-economic transition has played a pivotal role in reducing relative inequality. In summary, while absolute metrics (SII) highlight an increasing disability rate in high-SDI countries as a byproduct of aging and survival, relative indexes (concentration index) emphasize that global health efforts must remain focused on the persistent inequities in lower-income areas. Thus, addressing the fall-related burden requires a dual strategy: managing the rising disability in aging, high-income societies while simultaneously tackling the high mortality burden in disadvantaged populations.

The frontier analysis identified the ideal levels of burden that countries could achieve given their stage of development. Thus, countries positioned farther from the frontier have greater potential to enhance fall prevention among older adults. The analysis also highlights which countries excel in controlling the burden of falls and which fail to meet these ideal levels. Although one might expect higher-SDI countries to more effectively manage the burden of falls in older populations, the relatively flat frontier curve for YLDs against SDI suggests little difference in ideal burden of disability between lower- and higher-SDI countries.

Projections of the fall-related burden among older adults highlight potential public health challenges in the coming decades. Although the age-standardized overall burden (DALYs) is expected to decline, the ASR of YLDs is projected to increase, suggesting that the disability burden may pose an increasingly significant future challenge. To mitigate this anticipated shift toward a disability-predominant burden, forward-looking strategies must simultaneously address 3 critical dimensions: preserving intrinsic capacity, minimizing extrinsic hazards, and generalizing system capacity.

Specifically, prevention efforts may benefit from prioritizing the preservation and enhancement of intrinsic capacity, defined as the composite of an individual's physical and mental reserves [61]. The integration of systematic screening approaches such as the World Health Organization's integrated care for older people framework into primary health care systems could facilitate the early identification of deficits in functional ability [62,63]. Crucially, screening must be coupled with targeted interventions; identifying risk is only the first step. Comprehensive management should extend to personalized lifestyle guidance, encompassing nutritional optimization to combat sarcopenia, sleep hygiene to maintain cognitive alertness, and psychological support to alleviate the fear of falling [64-66]. Subsequent individualized exercise prescriptions, such as the Otago Exercise Program or tai chi, should be implemented to improve muscle strength and balance [67,68].

XSL•FO
RenderX

This proactive continuum aims to reverse physical frailty, thereby fortifying the individual's intrinsic resilience against falls.

Complementing these active measures, policy efforts may benefit from prioritizing the reduction in extrinsic hazards through environmental modifications. Implementing universal design principles in urban planning, including barrier-free walkways and adequate street lighting, alongside subsidizing home modifications such as grab bars and nonslip flooring, is critical for supporting "aging-in-place" [69-71]. By systematically mitigating external environmental stressors, such measures can provide a passive safety buffer that operates independently of an individual's fluctuating physical condition, thereby lowering the probability of injury in daily life.

Finally, the adoption of emerging technologies offers a potential solution to generalize and extend system capacity, bridging the workforce gap in geriatric care [72,73]. Wearable fall detection sensors and tele-rehabilitation platforms could provide continuous, unobtrusive monitoring and accessible home-based training [74-77]. These technologies hold the potential to democratize access to specialized resources, supporting the delivery of fall prevention and rehabilitation services and ensuring they are available even in resource-constrained or remote settings.

Limitations

While this study provides a comprehensive global overview of fall-related burdens among older adults, several limitations warrant consideration. First, this research used a macrolevel analysis examining national data from 1990 to 2021; consequently, it did not involve in-depth clinical investigations or granular subnational analyses of individual countries. Localized heterogeneity, such as specific rural-urban disparities within a single country, unique cultural safety practices, or variations in regional infrastructure, may not be fully captured by national-level SDI metrics. Second, the findings are subject to the inherent limitations of the GBD study's primary data sources. In many low-SDI regions, vital registration systems may be fragmented or inconsistent. This can lead to potential underreporting or the misclassification of fall-related injuries. However, to mitigate these disparities, the GBD study framework uses rigorous statistical modeling methods to calibrate data inconsistencies and fill gaps. While these advanced techniques significantly reduce bias and ensure global comparability, the estimates in data-sparse regions should still be interpreted with a degree of caution. Finally, while the frontier and inequality analyses provide robust insights into global patterns, they are based on aggregated statistical models. These models describe observed trends and performance gaps but do not account for sudden shifts in health policy or major global disruptions. Future research using more granular, local-level data would be beneficial to validate these national trends and further explore the specific sociocultural drivers of the fall-related burden.

Conclusions

This study elucidates divergent patterns in the global burden of falls among older adults, consistent with an epidemiological transition: lower-SDI regions bear a predominant mortality burden (YLLs), while higher-SDI regions face an elevated disability burden (YLDs). Regarding health inequality, although the relative distribution of the overall burden (concentration index) showed signs of convergence during the study period, the absolute gap (SII) in ASRs between the most and least advantaged populations appears to persist. Furthermore, the frontier analysis indicated that potential exists for countries across the sociodemographic spectrum to further reduce their burden relative to achievable benchmarks.

Projections of the fall-related burden suggest that, while overall rates may decline, the disability burden (YLDs) is likely to pose an increasingly significant challenge in the coming decades. Collectively, these findings suggest that enhanced proactive preparedness could play a pivotal role in mitigating this anticipated future strain.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Spatiotemporal patterns of incidence and mortality and Joinpoint regression analyses of temporal trends in falls among adults aged 65 years and older.
[DOCX File, 9054 KB - aging_v9i1e73802_app1.docx]

Multimedia Appendix 2
Additional results on cross-national variations in fall-related mortality, incidence, disability-adjusted life years, years lived with disability, and years of life lost among adults aged 65 years and older by sociodemographic index level.
[XLSX File, 478 KB - aging_v9i1e73802_app2.xlsx]

References

XSL•FO
RenderX

1. GBD 2021 Diseases and Injuries Collaborators. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2024 May 18;403(10440):2133-2161. [doi: 10.1016/S0140-6736(24)00757-8] [Medline: 38642570]
2. Murray CJ, GBD 2021 Collaborators. Findings from the Global Burden of Disease Study 2021. Lancet 2024 May;403(10440):2259-2262. [doi: 10.1016/S0140-6736(24)00769-4] [Medline: 38762327]
3. WHO global report on falls prevention in older age. : World Health Organization; 2007 URL: https://extranet.who.int/agefriendlyworld/wp-content/uploads/2014/06/WHo-Global-report-on-falls-prevention-in-older-age.pdf [accessed 2026-04-01]
4. Zangbar B, Rafieezadeh A, Rodriguez G, et al. Geriatric falls: an enormous economic burden compared to firearms. Trauma Surg Acute Care Open 2024;9(1):e001492. [doi: 10.1136/tsaco-2024-001492] [Medline: 39416955]
5. Dykes PC, Curtin-Bowen M, Lipsitz S, et al. Cost of inpatient falls and cost-benefit analysis of implementation of an evidence-based fall prevention program. JAMA Health Forum 2023 Jan 6;4(1):e225125. [doi: 10.1001/jamahealthforum.2022.5125] [Medline: 36662505]
6. Wang J, Li Y, Yang GY, Jin K. Age-related dysfunction in balance: a comprehensive review of causes, consequences, and interventions. Aging Dis 2024 Jan 24;16(2):714-737. [doi: 10.14336/AD.2024.0124-1] [Medline: 38607735]
7. Kohle F, Stark C, Klünter HD, et al. Peripheral neuropathy, an independent risk factor for falls in the elderly, impairs stepping as a postural control mechanism: a case-cohort study. J Peripher Nerv Syst 2024 Dec;29(4):453-463. [doi: 10.1111/jns.12656] [Medline: 39219364]
8. Li Y, Smith RM, Whitney SL, Seemungal BM, Ellmers TJ. Association between dizziness and future falls and fall-related injuries in older adults: a systematic review and meta-analysis. Age Ageing 2024 Sep 1;53(9):afae177. [doi: 10.1093/ageing/afae177] [Medline: 39293812]
9. Yao S, Chen XW. Effects of global cognitive function on the prevalence of fear of falling in older adults. J Psychiatr Res 2025 Jan;181:72-79. [doi: 10.1016/j.jpsychires.2024.11.060] [Medline: 39603164]
10. Nie XY, Dong XX, Lu H, et al. Multimorbidity patterns and the risk of falls among older adults: a community-based study in China. BMC Geriatr 2024 Aug 7;24(1):660. [doi: 10.1186/s12877-024-05245-1] [Medline: 39112944]
11. Kim S, Kim S, Woo S, et al. Temporal trends and patterns in mortality from falls across 59 high-income and upper-middle-income countries, 1990-2021, with projections up to 2040: a global time-series analysis and modelling study. Lancet Healthy Longev 2025 Jan;6(1):100672. [doi: 10.1016/j.lanhl.2024.100672] [Medline: 39848266]
12. GBD results. Institute for Health Metrics and Evaluation. URL: https://vizhub.healthdata.org/gbd-results [accessed 2024-09-11]
13. Vollset SE, Goren E, Yuan CW, et al. Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study. Lancet 2020 Oct 17;396(10258):1285-1306. [doi: 10.1016/S0140-6736(20)30677-2] [Medline: 32679112]
14. Global Burden of Disease Study 2021 (GBD 2021) demographics 1950-2021. Institute for Health Metrics and Evaluation. 2024. URL: https://ghdx.healthdata.org/record/ihme-data/global-burden-disease-study-2021-gbd-2021-demographics-1950-2021 [accessed 2024-09-12]
15. Global Burden of Disease Study 2021 (GBD 2021) socio-demographic index (SDI) 1950–2021. Institute for Health Metrics and Evaluation. 2024. URL: https://ghdx.healthdata.org/record/global-burden-disease-study-2021-gbd-2021-socio-demographic-index-sdi-1950%E2%80%932021 [accessed 2024-09-11]
16. GBD 2015 SDG Collaborators. Measuring the health-related Sustainable Development Goals in 188 countries: a baseline analysis from the Global Burden of Disease Study 2015. Lancet 2016 Oct;388(10053):1813-1850. [doi: 10.1016/S0140-6736(16)31467-2] [Medline: 27665228]
17. GBD 2015 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct;388(10053):1603-1658. [doi: 10.1016/S0140-6736(16)31460-X] [Medline: 27733283]
18. Joinpoint trend analysis software. National Cancer Institute. URL: https://surveillance.cancer.gov/joinpoint/ [accessed 2026-04-06]
19. Kim HJ, Chen HS, Byrne J, Wheeler B, Feuer EJ. Twenty years since Joinpoint 1.0: two major enhancements, their justification, and impact. Stat Med 2022 Jul 20;41(16):3102-3130. [doi: 10.1002/sim.9407] [Medline: 35522060]
20. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000 Feb 15;19(3):335-351. [Medline: 10649300]
21. Schlotheuber A, Hosseinpoor AR. Summary measures of health inequality: a review of existing measures and their application. Int J Environ Res Public Health 2022 Mar 20;19(6):3697. [doi: 10.3390/ijerph19063697] [Medline: 35329383]
22. Xie Y, Bowe B, Mokdad AH, et al. Analysis of the Global Burden of Disease study highlights the global, regional, and national trends of chronic kidney disease epidemiology from 1990 to 2016. Kidney Int 2018 Sep;94(3):567-581. [doi: 10.1016/j.kint.2018.04.011] [Medline: 30078514]
23. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc Series B Stat Methodol 2009 Apr;71(2):319-392. [doi: 10.1111/j.1467-9868.2008.00700.x]

XSL•FO
RenderX

24. Martins TG, Simpson D, Lindgren F, Rue H. Bayesian computing with INLA: new features. Comput Stat Data Anal 2013 Nov;67:68-83. [doi: 10.1016/j.csda.2013.04.014]
25. Handbook on Health Inequality Monitoring: With a Special Focus on Low- and Middle-Income Countries: World Health Organization; 2013. URL: https://www.who.int/docs/default-source/gho-documents/health-equity/handbook-on-health-inequality-monitoring/handbook-on-health-inequality-monitoring.pdf [accessed 2024-12-20]
26. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
27. Haagsma JA, Graetz N, Bolliger I, et al. The global burden of injury: incidence, mortality, disability-adjusted life years and time trends from the Global Burden of Disease study 2013. Inj Prev 2016 Feb;22(1):3-18. [doi: 10.1136/injuryprev-2015-041616] [Medline: 26635210]
28. Ó Breasail M, Zengin A, Pearse C, et al. Rural-urban differences in osteoporosis and sarcopenia prevalence among Gambian older adults: a pilot study. J Bone Miner Res 2025 Sep 22:zjaf130. [doi: 10.1093/jbmr/zjaf130] [Medline: 40981665]
29. Seo Y, Kim M, Shin H, Won C. Perceived neighborhood environment associated with sarcopenia in urban-dwelling older adults: the Korean Frailty and Aging Cohort Study (KFACS). Int J Environ Res Public Health 2021 Jun 10;18(12):6292. [doi: 10.3390/ijerph18126292] [Medline: 34200703]
30. Curl A, Fitt H, Tomintz M. Experiences of the built environment, falls and fear of falling outdoors among older adults: an exploratory study and future directions. Int J Environ Res Public Health 2020 Feb 14;17(4):1224. [doi: 10.3390/ijerph17041224] [Medline: 32074960]
31. Zhang N, Yang Q. Public transport inclusion and active aging: a systematic review on elderly mobility. J Traffic Transp Eng (Engl Ed) 2024 Apr;11(2):312-347. [doi: 10.1016/j.jtte.2024.04.001]
32. Clarke P, Gallagher NA. Optimizing mobility in later life: the role of the urban built environment for older adults aging in place. J Urban Health 2013 Dec;90(6):997-1009. [doi: 10.1007/s11524-013-9800-4] [Medline: 23592019]
33. O'Neil J, Steele GK, Huisingh C, Smith GA. Escalator-related injuries among older adults in the United States, 1991-2005. Accid Anal Prev 2008 Mar;40(2):527-533. [doi: 10.1016/j.aap.2007.08.008] [Medline: 18329403]
34. Al-Azzani W, Adam Maliq Mak D, Hodgson P, Williams R. Epidemic of fractures during a period of snow and ice: has anything changed 33 years on? BMJ Open 2016 Sep 14;6(9):e010582. [doi: 10.1136/bmjopen-2015-010582] [Medline: 27630066]
35. Dahl C, Madsen C, Omsland TK, et al. The association of cold ambient temperature with fracture risk and mortality: national data from Norway-a Norwegian Epidemiologic Osteoporosis Studies (NOREPOS) study. J Bone Miner Res 2022 Aug;37(8):1527-1536. [doi: 10.1002/jbmr.4628] [Medline: 35689442]
36. Odén A, Kanis JA, McCloskey EV, Johansson H. The effect of latitude on the risk and seasonal variation in hip fracture in Sweden. J Bone Miner Res 2014 Oct;29(10):2217-2223. [doi: 10.1002/jbmr.2250] [Medline: 24715585]
37. Gauss T, de Jongh M, Maegele M, Cole E, Bouzat P. Trauma systems in high socioeconomic index countries in 2050. Crit Care 2024 Mar 16;28(1):84. [doi: 10.1186/s13054-024-04863-w] [Medline: 38493142]
38. Li A, Feng Q, Zhao Y, Zhang X, Jiang W. Comprehensive meta-analysis of emergency trauma outcomes: trends, interventions, and survival rates. Front Public Health 2025;13:1696401. [doi: 10.3389/fpubh.2025.1696401] [Medline: 41393036]
39. Arinzon Z, Fidelman Z, Zuta A, Peisakh A, Berner YN. Functional recovery after hip fracture in old-old elderly patients. Arch Gerontol Geriatr 2005;40(3):327-336. [doi: 10.1016/j.archger.2004.10.003] [Medline: 15814166]
40. Zhao L, Zhao X, Dong B, Li X. Effectiveness of home-based exercise for functional rehabilitation in older adults after hip fracture surgery: a systematic review and meta-analysis of randomized controlled trials. PLoS One 2024;19(12):e0315707. [doi: 10.1371/journal.pone.0315707] [Medline: 39700091]
41. Jesus TS, Landry MD. Global need: including rehabilitation in health system strengthening. Lancet 2021 Feb 20;397(10275):665-666. [doi: 10.1016/S0140-6736(21)00207-5] [Medline: 33610203]
42. Chudyk AM, Jutai JW, Petrella RJ, Speechley M. Systematic review of hip fracture rehabilitation practices in the elderly. Arch Phys Med Rehabil 2009 Feb;90(2):246-262. [doi: 10.1016/j.apmr.2008.06.036] [Medline: 19236978]
43. Cieza A, Causey K, Kamenov K, Hanson SW, Chatterji S, Vos T. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2021 Dec 19;396(10267):2006-2017. [doi: 10.1016/S0140-6736(20)32340-0] [Medline: 33275908]
44. Biswas I, Adebusoye B, Chattopadhyay K. Risk factors for falls among older adults in India: a systematic review and meta-analysis. Health Sci Rep 2022 Jul;5(4):e637. [doi: 10.1002/hsr2.637] [Medline: 35774830]
45. James SL, Lucchesi LR, Bisignano C, et al. The global burden of falls: global, regional and national estimates of morbidity and mortality from the Global Burden of Disease Study 2017. Inj Prev 2020 Oct;26(Supp 1):i3-i11. [doi: 10.1136/injuryprev-2019-043286] [Medline: 31941758]
46. Kalula SZ, Scott V, Dowd A, Brodrick K. Falls and fall prevention programmes in developing countries: environmental scan for the adaptation of the Canadian falls prevention curriculum for developing countries. J Safety Res 2011 Dec;42(6):461-472. [doi: 10.1016/j.jsr.2010.12.003] [Medline: 22152264]

XSL•FO
RenderX

47. Okoye EC, Akosile CO, Maruf FA, Onwuakagba IU, Mgbeojedo UG. Falls and fear of falling among older adults in an assisted-living facility: a qualitative and foundational study for intervention development in a developing country. Arch Gerontol Geriatr 2021;94:104375. [doi: 10.1016/j.archger.2021.104375] [Medline: 33571785]
48. Bhattarai HK, Bhusal S, Barone-Adesi F, Hubloue I. Prehospital emergency care in low- and middle-income countries: a systematic review. Prehosp Disaster Med 2023 Aug;38(4):495-512. [doi: 10.1017/S1049023X23006088] [Medline: 37492946]
49. Kitayuguchi J, Abe T, Okuyama K, et al. Association between a hilly neighborhood environment and falls among rural older adults: a cross-sectional study. J Rural Med 2021 Oct;16(4):214-221. [doi: 10.2185/jrm.2021-028] [Medline: 34707730]
50. Li W, Keegan TH, Sternfeld B, Sidney S, Quesenberry CP Jr, Kelsey JL. Outdoor falls among middle-aged and older adults: a neglected public health problem. Am J Public Health 2006 Jul;96(7):1192-1200. [doi: 10.2105/AJPH.2005.083055] [Medline: 16735616]
51. Anderer S. Home hazards increase fall risk in older adults with visual impairment. JAMA 2026;335(5):392. [doi: 10.1001/jama.2025.23017]
52. Schmitz MF, Giunta N, Parikh NS, Chen KK, Fahs MC, Gallo WT. The association between neighbourhood social cohesion and hypertension management strategies in older adults. Age Ageing 2012 May;41(3):388-392. [doi: 10.1093/ageing/afr163] [Medline: 22166684]
53. Keall MD, Pierse N, Howden-Chapman P, et al. Home modifications to reduce injuries from falls in the home injury prevention intervention (HIPI) study: a cluster-randomised controlled trial. Lancet 2015 Jan 17;385(9964):231-238. [doi: 10.1016/S0140-6736(14)61006-0] [Medline: 25255696]
54. Conradie T, Berner K, Louw Q. Describing the rehabilitation workforce capacity in the public sector of three rural provinces in South Africa: a cross-sectional study. Int J Environ Res Public Health 2022 Sep 26;19(19):12176. [doi: 10.3390/ijerph191912176] [Medline: 36231495]
55. Ghassi HT, Muntessu Nsangou DL, Buh FC, Anwar F. Falls in older adults in sub-Saharan Africa: an urgent call for rehabilitation-centred action. BMJ Public Health 2026;4(1):e004402. [doi: 10.1136/bmjph-2025-004402] [Medline: 41626609]
56. Schulz R, Beach SR, Czaja SJ, Martire LM, Monin JK. Family caregiving for older adults. Annu Rev Psychol 2020 Jan 4;71:635-659. [doi: 10.1146/annurev-psych-010419-050754] [Medline: 31905111]
57. Yempabe T, Edusei A, Donkor P, Buunaaim A, Mock C. Factors affecting utilization of traditional bonesetters in the Northern Region of Ghana. Afr J Emerg Med 2021 Mar;11(1):105-110. [doi: 10.1016/j.afjem.2020.09.002] [Medline: 33680729]
58. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):afac205. [doi: 10.1093/ageing/afac205] [Medline: 36178003]
59. Jamison DT, Summers LH, Alleyne G, et al. Global health 2035: a world converging within a generation. Lancet 2013 Dec;382(9908):1898-1955. [doi: 10.1016/S0140-6736(13)62105-4] [Medline: 24309475]
60. GBD 2019 Universal Health Coverage Collaborators. Measuring universal health coverage based on an index of effective coverage of health services in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct;396(10258):1250-1284. [doi: 10.1016/S0140-6736(20)30750-9] [Medline: 32861314]
61. Bautmans I, Knoop V, Amuthavalli Thiyagarajan J, et al. WHO working definition of vitality capacity for healthy longevity monitoring. Lancet Healthy Longev 2022 Nov;3(11):e789-e796. [doi: 10.1016/S2666-7568(22)00200-8] [Medline: 36356628]
62. Sum G, Lau LK, Jabbar KA, et al. The World Health Organization (WHO) Integrated Care for Older People (ICOPE) framework: a narrative review on its adoption worldwide and lessons learnt. Int J Environ Res Public Health 2022 Dec 22;20(1):154. [doi: 10.3390/ijerph20010154] [Medline: 36612480]
63. Tavassoli N, de Souto Barreto P, Berbon C, et al. Implementation of the WHO Integrated Care for Older People (ICOPE) programme in clinical practice: a prospective study. Lancet Healthy Longev 2022 Jun;3(6):e394-e404. [doi: 10.1016/S2666-7568(22)00097-6] [Medline: 36098317]
64. Kuzuya M, Kanno Y, Arai H, et al. Guidelines for nutritional management of sarcopenia and frailty 2025. Guest editor: Masafumi Kuzuya. This publication has been supported by the Japan Geriatrics Society and National Center for Geriatrics and Gerontology (NCGG) (Japan). Geriatr Gerontol Int 2025 Dec;25 Suppl 2:5-53. [doi: 10.1111/ggi.70275] [Medline: 41524691]
65. Zhang Z, Diao Y, Fu M, et al. Self-reported sleep duration and recurrent falls in people aged 50 and above: evidence from two prospective cohorts. J Epidemiol Community Health 2025 Dec 25. [doi: 10.1136/jech-2025-224958] [Medline: 41448939]
66. Drahota A, Udell JE, Mackenzie H, Pugh MT. Psychological and educational interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev 2024 Oct 3;10(10):CD013480. [doi: 10.1002/14651858.CD013480.pub2] [Medline: 39360568]
67. Chen W, Li M, Li H, Lin Y, Feng Z. Tai Chi for fall prevention and balance improvement in older adults: a systematic review and meta-analysis of randomized controlled trials. Front Public Health 2023;11:1236050. [doi: 10.3389/fpubh.2023.1236050] [Medline: 37736087]
68. Yang Y, Wang K, Liu H, et al. The impact of Otago exercise programme on the prevention of falls in older adult: a systematic review. Front Public Health 2022;10:953593. [doi: 10.3389/fpubh.2022.953593] [Medline: 36339194]

XSL•FO
RenderX

69. Keall MD, Tupara H, Pierse N, et al. Home modifications to prevent home fall injuries in houses with Māori occupants (MHIPI): a randomised controlled trial. Lancet Public Health 2021 Sep;6(9):e631-e640. [doi: 10.1016/S2468-2667(21)00135-3] [Medline: 34371005]
70. Stark S, Keglovits M, Somerville E, et al. Home hazard removal to reduce falls among community-dwelling older adults: a randomized clinical trial. JAMA Netw Open 2021 Aug 2;4(8):e2122044. [doi: 10.1001/jamanetworkopen.2021.22044] [Medline: 34463746]
71. Global age-friendly cities: a guide. World Health Organization. 2007. URL: https://www.who.int/publications/i/item/9789241547307 [accessed 2026-04-02]
72. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023 Jul;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8] [Medline: 37464029]
73. Maita KC, Maniaci MJ, Haider CR, et al. The impact of digital health solutions on bridging the health care gap in rural areas: a scoping review. Perm J 2024 Sep 16;28(3):130-143. [doi: 10.7812/TPP/23.134] [Medline: 39135461]
74. Konara Mudiyanselage SP, Yao CT, Maithreepala SD, Lee BO. Emerging digital technologies used for fall detection in older adults in aged care: a scoping review. J Am Med Dir Assoc 2025 Jan;26(1):105330. [doi: 10.1016/j.jamda.2024.105330] [Medline: 39492106]
75. Pech M, Sauzeon H, Yebda T, Benois-Pineau J, Amieva H. Falls detection and prevention systems in home care for older adults: myth or reality? JMIR Aging 2021 Dec 9;4(4):e29744. [doi: 10.2196/29744] [Medline: 34889755]
76. Tsekoura M, Kastrinis A, Nomikou E, Katsoulaki M. Telerehabilitation and fall prevention in older adults. Adv Exp Med Biol 2023;1425:485-489. [doi: 10.1007/978-3-031-31986-0_47] [Medline: 37581822]
77. Xiao H, Zeng W, Lu L, Yuan J, Yan Z, Wang J. Effectiveness of telerehabilitation in postoperative outcomes in patients on hip fracture surgery: a meta-analysis of randomized controlled trials. BMC Sports Sci Med Rehabil 2025 May 26;17(1):130. [doi: 10.1186/s13102-025-01169-4] [Medline: 40420240]

Abbreviations

APC: annual percentage change
ASR: age-standardized rate
DALY: disability-adjusted life year
GBD: Global Burden of Disease
SDI: sociodemographic index
SII: slope index of inequality
YLD: year lived with disability
YLL: year of life lost

Edited by Y Jiang; submitted 12.Mar.2025; peer-reviewed by J Richardson, Y Tanabe; revised version received 07.Feb.2026; accepted 20.Mar.2026; published 10.Apr.2026.

Please cite as:
Zhang Z, Diao Y, Fu M, Han W, Zhou H, Fan R, Xu B, Chen B
Global, Regional, and National Burden of Falls Among Older Adults Aged 65 Years and Above: Secondary Data Analysis of the Global Burden of Disease Study 2021
JMIR Aging 2026;9:e73802
URL: https://aging.jmir.org/2026/1/e73802
doi:10.2196/73802

© Ze Zhang, Yingying Diao, Mingwang Fu, Wantong Han, Haoran Zhou, Ru Fan, Biyun Xu, Bingwei Chen. Originally published in JMIR Aging (https://aging.jmir.org), 10.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Supervised and Self-Directed Technology-Based Dual-Task Exercise Training Program for Older Adults With a History of Falls: Mixed Methods Feasibility Study

Prerna Mathur[1], MSc; Afroditi Stathi[1], PhD; Victoria A Goodyear[1,2], PhD; Taylor Krauss[3], MSc; Helen Thomas[3], BSc; Angela Cooper[1], MSc; Philip Kinghorn[4], PhD; Caroline Miller[5,6], PhD; Natalie Ives[7], MSc; Magdalena Chechlacz[8,9], PhD; Daisy Wilson[6,10], PhD; Laura Magill[11], PhD; Shin-Yi Chiou[1,9], PT, PhD

[1]School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Edgbaston, Birmingham, United Kingdom

[10]Healthcare for Older People Department, Queen Elizabeth Hospital Birmingham, Birmingham, United Kingdom

[11]Birmingham Centre for Observational and Prospective Studies (BiCOPS), School of Health Sciences, University of Birmingham, Birmingham, United Kingdom

[2]Institute for Mental Health, University of Birmingham, Birmingham, United Kingdom

[3]Solihull Community Specialist Falls Service, Solihull Hospital, University Hospitals Birmingham NHS Foundation Trust, Solihull, United Kingdom

[4]Health Economics Unit, Department of Applied Health Sciences, University of Birmingham, Birmingham, United Kingdom

[5]Physiotherapy Department, Queen Elizabeth Hospital Birmingham, University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom

[6]Department of Inflammation and Ageing, University of Birmingham, Birmingham, United Kingdom

[7]Birmingham Clinical Trials Unit, School of Health Sciences, University of Birmingham, Birmingham, United Kingdom

[8]School of Psychology, University of Birmingham, Birmingham, United Kingdom

[9]Centre for Human Brain Health, University of Birmingham, Birmingham, United Kingdom

Corresponding Author:
Shin-Yi Chiou, PT, PhD
School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Edgbaston, Birmingham, United Kingdom

Abstract

Background: Older adults who have fallen are at an increased risk of future falls. Training cognitive and physical functions simultaneously, known as dual-task (DT) training, has been shown to improve mobility and reduce fall risks. With appropriate digital tools, such as smartphones and mobile apps, it is possible to deliver DT training in unsupervised, home-based settings, thereby increasing accessibility beyond the clinical environment.

Objective: This study aimed to evaluate the feasibility and acceptability of a technology-based DT training program delivered through a blended model of supervised and self-directed sessions for older adults with a history of falls. Perspectives of health care professionals working in fall-prevention services were also explored.

Methods: A single-arm, nonrandomized feasibility study was conducted with 45 community-dwelling adults aged 65 years or older with a history of falls. Participants were recruited through primary care practices, secondary care fall-prevention services, and community outreach. The 24-week DT program, which integrated balance and strength exercises with cognitive training using a mobile app, was delivered in two phases: (1) for12 weeks, weekly 50-minute physiotherapist-led group classes in the community, and 2 additional 50-minute self-directed sessions at home, and (2) for 12 weeks, 3 weekly 50-minute self-directed sessions at home. Feasibility and acceptability were assessed through recruitment and retention rates, adherence, app usage, and self-reported satisfaction. Qualitative data were obtained from focus groups with 28 participants who completed the program and 16 health care professionals. Quantitative data were analyzed descriptively, and qualitative data were analyzed thematically.

Results: We recruited 45 of the target 50 participants, with most participants (n=41) recruited through community pathways; 4 were recruited via National Health Service (NHS) pathways. Adherence was 64%, with higher adherence during phase 1 (81%) than phase 2 (50%). App usage was high (95%), and self-reported program satisfaction was moderate to high. Retention at 24 weeks was 76%, and no adverse events occurred. The qualitative findings supported the program's feasibility and acceptability, emphasizing social connection and tailored exercises as key to adherence—especially in home-based sessions. Health care professionals identified community organizations and referral pathways as the most practical routes for implementation.

Conclusions: A blended, technology-based DT training program is both feasible and acceptable for older adults at risk of falling and can be effectively delivered beyond clinical settings. Community-based recruitment outperformed NHS pathways, highlighting

XSL•FO
RenderX

the value of community engagement. These findings support the feasibility and acceptability of a full-scale trial, with targeted refinements to recruitment, support structures, and delivery to maximize scalability and impact.

Trial Registration: ISRCTN ISRCTN15123197; https://www.isrctn.com/ISRCTN15123197

International Registered Report Identifier (IRRID): RR2-10.1371/journal.pone.0314829

(JMIR Aging 2026;9:e87577) doi:10.2196/87577

KEYWORDS

dual-task; aging; falls prevention; exercise; cognition; older adults; technology

Introduction

Falls and fall-related injuries are a major public health concern. One in three adults over 65 years experiences at least 1 fall each year, and individuals who have already fallen are at an increased risk of falling again [1,2]. Globally, falls are the second leading cause of unintentional injury–related deaths [3], and in the United Kingdom, falls and fall-related injuries cost an estimated £4.4 billion annually [1]. It is estimated that by 2050, 1 in 6 people worldwide will be aged 65 years or older. This demographic shift highlights the urgent need for effective prevention strategies [4,5].

Evidence suggests that falls can be prevented using appropriately prescribed exercise programs. Aging affects not only balance but also cognition, and domains such as attention, memory, and executive function are linked to mobility [6]. Dual-task (DT) training, which combines physical and cognitive exercises simultaneously, has been shown to improve walking speed, executive function, and balance and reduce fall risk in older adults [7,8]. Several lines of evidence suggest that training motor and cognitive function simultaneously, that is, DT training, is superior to single-task (balance or resistance) training or no training for improving mobility, global cognition, and reducing fall rates [9-11], leading to a better quality of life [12]. Training cognition and physical function simultaneously rather than sequentially is more beneficial [13,14]. Simultaneous cognitive and physical training can be delivered via technology (eg, mobile apps), enabling professionals to select cognitive exercises that are suitable to be combined with balance exercises, allowing DT training to be delivered unsupervised at home. Mobile apps are interactive, provide instant feedback, and can send reminders to users. These features promote engagement and adherence to exercise [15-17]. Emerging evidence supports the remote delivery of balance and strength exercise training using technology for falls prevention in older adults [18-20]. For example, a smartphone app delivering an unsupervised, home-based DT program to community-dwelling older adults without a history of falls showed feasibility, high adherence, and improved balance [20]. However, it remains unclear whether older adults with a history of falls, who are at greater risk of recurrence, would be willing or able to engage in an unsupervised, technology-supported DT exercise program. Additionally, the potential role of such programs within existing health care services is uncertain. Current services often cater to frailer individuals with multiple comorbidities who typically require close supervision [21], raising questions about the feasibility of recruitment solely through fall-prevention services as well as whether technology-supported DT training is better delivered in clinical or community settings.

The objectives of this study were to evaluate (1) the feasibility of recruiting and retaining community-dwelling older adults with a history of falls in a blended DT program combining supervised and self-directed sessions supported by technology; (2) the acceptability of the program to participants; (3) its potential effects on mobility, cognition, fall risks (eg, number of falls, fear of falling), quality of life, and use of health care resources; and (4) health care professionals' perspectives on integrating a blended DT program into existing care pathways for older adults.

Methods

Study Design

This was a prospective, single-arm, nonrandomized feasibility study. The study protocol has previously been published [22]. The study is reported according to the CONSORT (Consolidated Standards of Reporting Trials) pilot and feasibility trials checklist (Checklist 1).

Ethical Considerations

The study received ethical approval from the Health Research Authority East of England – Cambridge East Research Ethics Committee (Reference: 24/EE/0059) and was registered with ISRCTN (ISRCTN15123197; April 16, 2024). All participants provided written informed consent and were assigned a unique study ID to ensure anonymity. Personal data were stored securely on password-protected University of Birmingham servers, with identifiable information accessible only to authorized members of the research team and managed in accordance with the Data Protection Act 2018, UK General Data Protection Regulation, and relevant research governance frameworks. Participants were reimbursed for travel expenses associated with study participation.

Recruitment

Recruitment took place in 3 periods—May to August 2024, October to November 2024, and March to May 2025—to align with program delivery schedules and public holidays. Participants were recruited through text messages or invitation letters from general practitioners and hospital fall-prevention clinics, and community outreach (eg, retirement villages, charities, interest groups) using open events and study flyers. All participants received the intervention in the gym of the partnering retirement villages and community well-being and leisure centers.

XSL•FO
RenderX

Participants

The sample size (n=50) was chosen based on the recommendations of published literature for feasibility studies [23-25]. If we approach 100 eligible participants, we will be able to estimate a participation rate of 50% (green stop-go criteria; see Table 1) within a 95% CI of ±9.8%. With a sample size of 50 participants, we will be able to estimate an attrition rate of 20% (green stop-go) within a 95% CI of ±11%. Inclusion criteria included community-dwelling adults aged 65 years or older who had experienced at least 2 falls in the past year (including near misses); had access to a smartphone or tablet; could stand with 1-hand support on their usual walking aid for at least 60 seconds; could stand up from a chair and walk 6 m independently with their current walking aid; were self-toileting; and demonstrated sufficient cognition, hearing, and vision to follow the instructions for the assessment and exercise program. Exclusion criteria included medical conditions precluding safe exercise participation and current participation in another research study involving a fall-prevention intervention. Health care professionals involved in services for older adults within the study regions were recruited to participate in qualitative focus groups or one-to-one interviews.

Table . Stop-go criteria.

	Criteria
Green	Recruitment ≥50%; adherence ≥80% (ie, completion of at least 58 out of 72 prescribed sessions over 24 weeks); and retention ≥80%. Qualitative data confirm the intervention being feasible and acceptable, with no safety concerns identified.
Amber	Recruitment: 30%-49%; adherence: 50%-79%; and retention: 50% - 79%. Qualitative data indicate that modifications to program design or delivery are needed, but no safety concerns identified.
Red	Recruitment: <30%; adherence: <50%; and retention: <50%. Qualitative data do not support progression to a larger-scale randomized controlled trial and/or indicate unresolved safety concerns.

Dual Task Exercise Program

All participants underwent the 24-week DT exercise program, consisting of 2 phases:

- Phase 1 (weeks 1 - 12): weekly 50-minute physiotherapist-led group sessions (4 - 8 participants per group) and two 50-minute home-based sessions, which replicated the class exercises but were completed independently and unsupervised.
- Phase 2 (weeks 13 - 24): three weekly 50-minute home-based sessions, guided by a structured exercise plan provided by the research team.

Each session included a 5-minute warm-up, 20-minute DT exercises (static balance or strength exercises while engaging with cognitive games using the commercial *Peak Brain Training* app [26]; Supplementary Material 1 in Multimedia Appendix 1), 20-minute dynamic balance and strength training, and a 5-minute cool-down. The app was used on participants' smartphones or tablets, positioned on adjustable stands for ease of use.

All participants received a handbook with detailed exercise instructions and had access to online videos demonstrating the balance and strength exercises. Participants recorded weekly adherence in an exercise diary. At the end of phase 1, an in-person education session was provided to reinforce fall awareness and explain the program content in phase 2. Participants could contact the research team for technical support with the app as needed. A TIDieR (Template for Intervention Description and Replication) checklist is presented in Checklist 2. The intervention was co-designed with members of the patient and public involvement and engagement (PPIE) group, particularly regarding the selection of cognitive games to ensure they were appropriate and engaging for older adults. The PPIE group also helped refine program components intended to promote social interaction and engagement, including strategies for maintaining contact during the home-based phase.

Outcome Measures

Overview

The primary outcome measures were the feasibility and acceptability of the program, assessed through:

1. Recruitment rate: calculated as (number recruited via National Health Service [NHS] pathways/number of patients approached by primary and secondary care teams)×100. Recruitment via community pathways was excluded due to unquantifiable outreach.
2. Adherence: calculated as (class attendance +self-reported home-based sessions)/(total prescribed sessions)×100. It was reported separately for phase 1, phase 2, and overall.
3. Retention rate: calculated as (number of participants completing the study/number consented) × 100. Completion was defined as attendance at the postintervention assessment.
4. Acceptability: assessed via a mixed methods exit survey administered at the end of phase 1 (wk 12) and phase 2 (wk 24). The survey comprised structured closed-ended items to capture quantitative data on usability, satisfaction, and perceived effectiveness, alongside open-ended free-text fields designed to elicit qualitative insights into participants' experiences and perspectives.

Stop-Go Criteria

Predefined stop-go criteria [22] were used to provide guidance on whether the project should progress to a full trial with an

XSL•FO
RenderX

internal pilot (green), progress to a full trial with an internal pilot with adaptation (amber), or no progression (red; Table 1).

Data were also collected on the following outcomes before and after the intervention. These were possible primary and secondary outcome measures for the main trial, so this enabled the assessment of data collection methods and data completeness of these measures, and exploratory data analysis:

1. Mobility and balance were assessed using the Timed Up and Go (TUG) test and TUG with a DT version of concurrent cognitive task (TUG-Cog), where participants counted backward by 7 from a randomly selected number between 100 and 200. The time taken to complete the test was recorded. All participants first completed 1 familiarization trial, followed by 1 recorded trial for the TUG and 1 recorded trial for the TUG-Cog.
2. Cognitive function was assessed by the total score of the 12-item Everyday Cognition (ECog-12) scales short version [27,28].
3. Fear of falls was assessed by the total score of the Falls Efficacy Scale-International (FES-I) questionnaire [29].
4. Health-related quality of life was assessed using the EQ-5D-5L [30].
5. Usage of health care resources was self-reported using ModRUM [31].
6. Self-reported number of falls in the last 6 months.

TUG and TUG-Cog were assessed in person by an experienced physiotherapist; all other outcomes were collected using a REDCap (Research Electronic Data Capture; v15.4.3) online survey. Survey links were sent to participants 1 week before the intervention and again in week 24, with up to 3 reminders to facilitate response rates. Paper copies were provided to participants upon request.

At baseline, demographic information was collected, including age, sex, ethnicity, level of education, area deprivation (Index of Multiple Deprivation), and type of residence. Frailty was assessed retrospectively using the Clinical Frailty Scale [32] to characterize the participant cohort in this feasibility study and to provide insight into the types of individuals for whom the intervention may be feasible and appropriate. An experienced geriatrician (DW) determined a frailty score following a detailed review of the demographic and functional data alongside a discussion with the team using the Clinical Frailty Scale [33].

Qualitative Methods

Focus Groups

Ten focus group semistructured interviews with 28 participants who completed the DT program (2 - 5 participants per group) were conducted. The focus groups were constructed to provide explanatory data on the study objectives (Supplementary Material 2 in Multimedia Appendix 1). Additionally, 4 online focus groups (n=12) and 4 online one-to-one interviews (n=4) were conducted with health care professionals (12 were physiotherapists, 1 occupational therapist, 1 general practitioner, 1 consultant geriatrician, and 1 research nurse) using a semistructured interview guide (Supplementary Material 2 in Multimedia Appendix 1). Focus groups were led by a member of the research team (PM), took place at the same venues as the exercise classes, were audio-recorded, and later transcribed verbatim. Online interviews were conducted using Microsoft Teams and were transcribed using the Microsoft Teams function. The interviews with the participants and health care professionals lasted between 45 and 80 minutes and 25 and 70 minutes, respectively.

Exit Surveys

Data were also collected from participants at the end of each program phase using an exit survey, which included both quantitative satisfaction ratings (Likert-scale items) and qualitative feedback from 8 open-ended questions. The questions were related to perceived benefits from the program, the experience of using the PEAK app while dual tasking, and recommendations for improving program delivery (Supplementary Material 3 in Multimedia Appendix 1).

Qualitative Data Analysis

Data from the focus groups and exit surveys were organized using NVivo 14 (Lumivero). Framework analysis was applied to the dataset to guide the examination of the study objectives, while also allowing for the identification of unanticipated themes and insights that emerged during the analytical process [34]. The data were initially coded deductively by members of the research team (PM, AS, and SYC) using a set of predefined categories that had been agreed upon in advance. These categories were directly aligned with the study objectives and included engagement and adherence, app usability, and retention. Once the data were organized according to these categories, an inductive approach was used to explore and identify additional factors contributing to each of the main categories. The coding process was led by PM and SYC, and a deliberative strategy [35] was used to enhance rigor. This involved multiple researchers (PM, SYC, AS, VAG) collaboratively reviewing and discussing the codes to refine interpretations and develop overarching themes. PPIE members also contributed to the interpretation of the focus group findings, providing perspectives that helped contextualize participants' experiences and identify practical implications for program delivery.

The explanatory themes were organized within the main categories identified through the framework analysis, which aligned with the study's objectives. These qualitative findings were then integrated with the descriptive quantitative analysis during the reporting phase. To ensure analytical rigor, a relativist approach to quality was applied, guided by key criteria including breadth, credibility, and coherence in the study design [36].

Quantitative Data Analysis

The analysis undertaken within this feasibility study was mainly descriptive. The progression criteria in the stop-go were summarized as proportions or means along with 95% CIs. The number of days of exercise and the number of physical exercises and cognitive games completed and recorded in the exercise diaries during the program were summarized to describe the adherence and usage of the app. To understand any change in the usage of the app between phase 1 and phase 2, we assessed the number of cognitive games completed in phase 1 and phase 2 separately.

XSL•FO
RenderX

The data collected on proposed outcome measures (eg, TUG) were summarized at baseline and follow-up using appropriate summary statistics. Exploratory analysis using a paired *t* test or Wilcoxon signed-rank test (depending on the distribution of the data) was undertaken to provide an exploratory assessment (using CIs; no *P* values are reported) of the effect of the blended intervention. All analyses were conducted in IBM SPSS Statistics (version 29; IBM Corp., 2022).

Results

Participant Recruitment

Forty-five older adults were recruited: 41 (91%) from community sources and 4 (9%) through NHS pathways (Figure 1). Community recruitment involved 10 retirement villages, 19 local activity groups, 8 instructors, and the Birmingham 1000 Elders mailing list, using open events and flyer distribution. As these approaches were open-access and did not involve direct invitations or approaches, it was not possible to determine the number of individuals exposed to the study information or formally approached. Therefore, a conventional recruitment rate (approached vs consented) could not be calculated for the community pathway. In the NHS recruitment pathways, 155 potential participants were identified through general practitioner surgeries (n=105) and secondary care falls clinics (n=50), but only 4 were enrolled (2.6%; 95% CI 0.7% - 6.5%), meeting the red stop-go criteria (Table 2-4). Nonparticipation was mainly due to health concerns, travel distance, or lack of a suitable device (eg, smartphone).

Figure 1. Flowchart. GP: general practitioner; NHS: National Health Service.

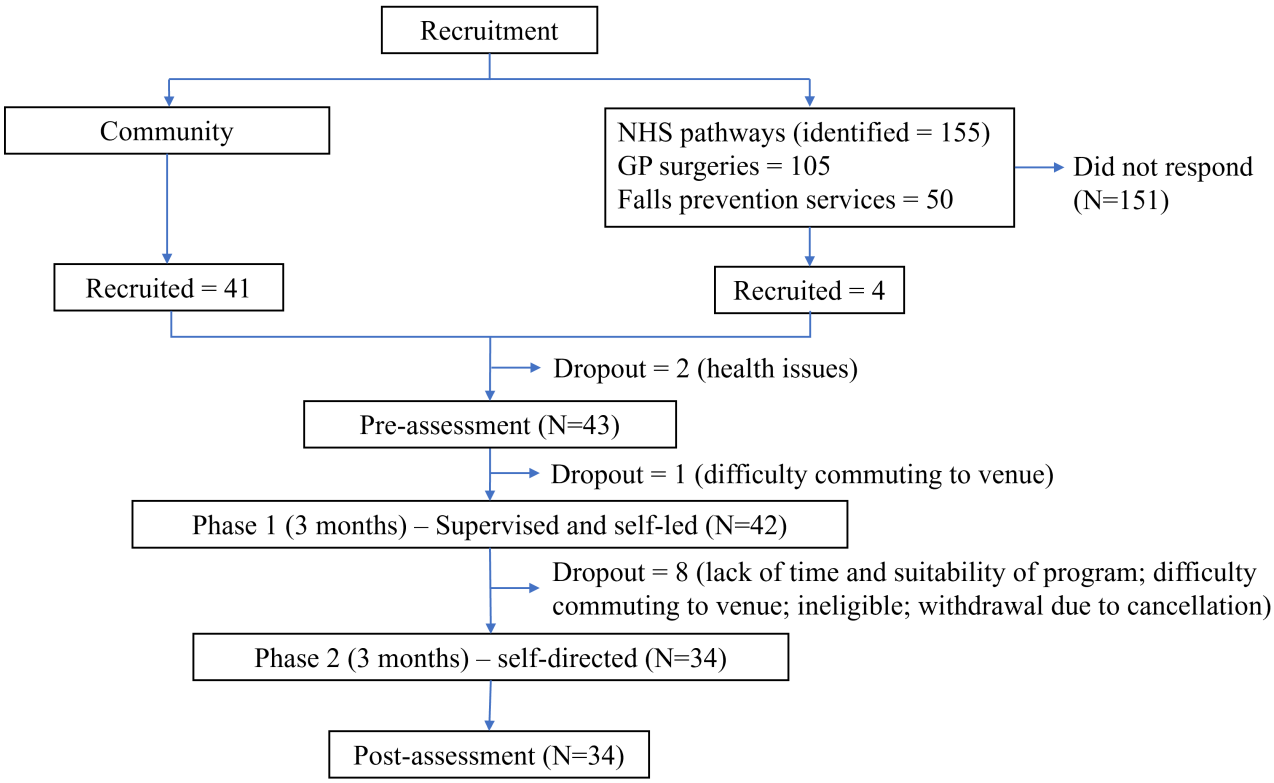


Table . Feasibility study outcomes and the stop-go criteria.

Outcome	Values; stop-go criteria	Full stop-go criteria
Recruitment rate from the NHS[a] pathways	Proportion 3% (95% CI 1% - 7%); red (red to red)	Green: ≥50%; amber: 30% - 49%; red: <30%
Adherence	Median 64% (IQR 46% - 82%; 95% CI 53% - 72%); amber (amber to amber)	Green: ≥80%; amber: 50% - 79%; red: <50%
Retention rate	Proportion 76% (95% CI 60% - 87%); amber (amber to green)	Green: ≥80%; amber: 50% - 79%; red: <50%

[a]NHS: National Health Service.

XSL•FO
RenderX

Table . Median adherence in each phase and overall.

Phase	Number of actual exercise days, median (IQR; 95% CI)	Adherence (%), median (IQR; 95% CI)
Phase 1	27 (24 - 31; 22 - 29)	81 (67 - 87; 64 - 82)
Phase 2	18 (14 - 21; 14 - 23)	50 (17 - 83; 39 - 65)
Phase 1+2	46 (33 - 57; 37 - 51)	64 (46 - 82; 53 - 72)

Table . Number of physical exercises and cognitive games completed (in each phase and overall).

Phase	Number of physical exercises completed, median (IQR; 95% CI)	Number of cognitive games completed, median (IQR; 95% CI)
Phase 1	156 (122 - 148; 123 - 167)	78 (69 - 92; 63 - 85)
Phase 2	137 (73 - 190; 94 - 162)	67 (36 - 96; 44 - 80)
Phase 1+2	275 (194 - 338; 224 - 322)	133 (98 - 174; 111 - 161)

The program consists of 3 exercise days per week in each phase (12 weeks); the total number of prescribed exercise days is 36 in each phase (Table 3). Participants are prescribed 6 physical exercises and 3 cognitive games in a week in each phase; the total number of prescribed physical exercises and cognitive games in each phase is 216 and 108, respectively (Table 4).

Sample Characteristics

The mean age of the participants was 78.3 (SD 7.6) years, and the majority were female participants (n=31, 68.9%). The number of participants recruited from sheltered and private accommodation was 13 (28.9%) and 32 (71.1%), respectively. The median Index of Multiple Deprivation score was 4 (IQR 1-10). The majority of the participants were White British (n=42, 93.3%). All participant characteristics as well as the characteristics of participants contributing to the qualitative study are presented in Table 5.

XSL•FO
RenderX

Table . Participant characteristics for all participants (N=45) at baseline and for participants (n=28) who took part in the focus groups.

Characteristics	All participants	Participants in focus groups
Age (y), mean (SD)	78.3 (7.6)	78.1 (8.9)
Sex, n (%)		
Female	31 (68.9)	20 (71.4)
Sheltered residents, n (%)	13 (28.9)	10 (35.7)
Private residents, n (%)	32 (71.1)	18 (64.3)
IMD[a] scores, median (IQR)	4 (1-10)	4.5 (1-9)
Highest level of education, n (%)		
Secondary education	9 (20)	7 (25)
Further education	6 (13.3)	4 (14.3)
Higher education	29 (64.4)	17 (60.7)
Prefer not to say	1 (2.2)	0 (0)
Ethnicity, n (%)		
White British	42 (93.3)	25 (89.3)
Asian or Asian British	3 (6.7)	3 (10.7)
Participant-reported data		
TUG[b] (s), mean (SD)	11.1 (3.5)	11.1 (4.0)
TUG-Cog[c] (s), mean (SD)	15.1 (7.2)	14.5 (6.0)
ECog-12[d], mean (SD)	1.61 (0.4)	1.6 (0.4)
Clinical frailty score, median (IQR)	3 (2-6)	3 (2-5)
FES-I[e], mean (SD)	31.8 (10.4)	32.9 (9.7)
High concern, n (%)	25 (59.5)	18 (64.3)
Moderate concern, n (%)	14 (33.3)	10 (35.7)
Low concern, n (%)	3 (7.1)	0 (0)
Falls in last 6 months, median (IQR)	2 (0 - 8)	2 (0 - 8)
EQ-5D-5L VAS score (out of 100), mean (SD)	72.1 (18.0)	74.1 (16.8)
EQ-5D-5L score, mean (SD)	0.644 (0.207)	0.655 (0.151)
Number of A&E[f] visits (sum)[g]	19	14

[a]IMD: Index of Multiple Deprivation.

[b]TUG: Timed Up and Go.

[c]TUG-Cog: Timed Up and Go test with a concurrent cognitive task.

[d]ECog-12: The Everyday Cognition scales short version.

[e]FES-I: Falls Efficacy Scale International.

[f]A&E: accident and emergency.

[g]This variable corresponds to the first item of the Modular Resource-Use Measure; the full results are reported in Supplementary Material 5-6 in Multimedia Appendix 1.

Study Retention

Thirty-four of the 45 participants completed the study, giving a retention rate of 76% (95% CI 60% - 87%), which met the amber stop-go criteria (Tables 2-4). A total of 11 participants did not complete the trial. Three withdrew after providing consent but before commencing the intervention due to health issues (n=2) and difficulty commuting to the venue where exercise classes were scheduled to take place (n=1). A further 8 participants withdrew during the intervention phase. Five withdrew due to time constraints or because they felt the program was not suitable for them. One participant was later deemed ineligible due to severe balance issues. Two additional participants were recruited during a third recruitment round; however, one of them withdrew after the first session because the venue was too far from their home. As this left only 1 participant in that group, the planned group sessions could not

XSL•FO
RenderX

proceed, and the final participant was therefore withdrawn from the study (Figure 1).

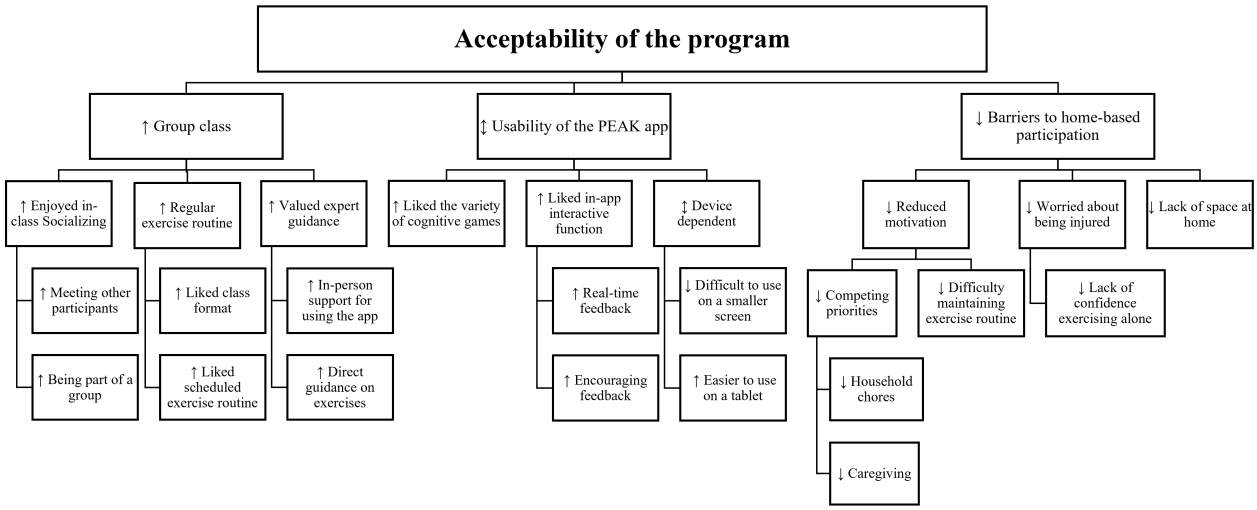
Adherence and Acceptability of the Intervention

Overall adherence to the program was 64% (95% CI 53% - 72%), meeting the amber stop-go criteria (Tables 2-4). On average, participants completed 46 days (95% CI 37 - 51) of combined cognitive and physical exercises out of the 72 days prescribed. Of the 432 prescribed physical exercises and 216 prescribed cognitive games, participants completed 275 physical exercises (64%, 95% CI 224 - 322) and 133 cognitive games (62%, 95% CI 111 - 161), respectively. These results indicate that participants engaged with and completed the intervention largely as intended. Adherence was higher in phase 1 (27 out of 26 days, 81%), which combined supervised and self-directed sessions, compared with phase 2 (18 out of 36 days, 50%), which was entirely self-directed. Within phase 1, supervised class attendance was 81% (9 out of 12 classes), and adherence to home-based sessions was 79% (18/24 home sessions). Difficulty finding time for weekly exercises was noted by 17% (n=5) of the participants in phase 1 and 41% (n=12) in phase 2. Adherence to the PEAK app was 95% (n=32). Participants valued its flexibility (n=23, 69%), although the proportion reporting ease of use declined from 59% (n=17) at the end of phase 1 to 34% (n=10) at the end of phase 2. No adverse events were reported.

The results of the exit survey highlighted that the median perceived program benefit was 7 out of 10 (1=not beneficial; 10=very beneficial). To explore whether perceived benefit influenced adherence, participants were divided into 2 groups based on this median score. Those who rated the program as highly beneficial (score ≥7) demonstrated higher adherence (n=15, 83%; median 33.3, IQR 63.9-97.2) compared to those with lower perceived benefit (score <7; n=13, 36%; median 27.8, IQR 16.7-44.4). Just over half (52%) of the participants agreed that integrating the app into the blended digital therapy program was a good way to improve balance. The majority (n=25, 86%) expressed satisfaction with the support received from both the research team and the physiotherapist. On average, participants reported a 6 out of 10 likelihood of continuing the program after the study concluded. The full results from the exit survey are available in Supplementary Material 3 in Multimedia Appendix 1.

Qualitative findings reinforced the quantitative results, highlighting social interaction as key to participant engagement (Figure 2—thematic tree). Group classes were valued for structure, motivation, enjoyment, and social connection, while the unsupervised phase saw reduced motivation and adherence due to competing priorities. Participants appreciated app features such as age-based rankings and the variety of cognitive games, though some struggled with small smartphone screens. Barriers to home exercise included limited space, low confidence, inappropriate difficulty levels, and concerns about joint pain. Illustrative quotes are provided in Supplementary Material 4 in Multimedia Appendix 1.

Figure 2. Thematic tree from the focus group with participants. Symbols indicate positive (↑), negative (↓), and neutral (↕) feedback.



Participants appreciated the feedback features (eg, ranking within age group, positive reinforcement) of the app and the variety of cognitive games, many of which were perceived as appropriately challenging. Challenges included difficulty using the app on smaller screens of smartphones.

Exploratory Analysis of Proposed Outcome Measures

The TUG is the planned primary outcome for the proposed future trial. Both TUG and TUG-Cog scores decreased after the program (Table 6). The FES-I scores decreased following the program, suggesting a reduction in fear of falling; changes in the category of the FES-I [29] at postassessment are presented in Table 7. There was no change in ECog-12 scores, but there were decreases in EQ-5D-5L tariff and VAS scores (Table 6). The completion rates for TUG and TUG-Cog were 93.94%, and for ECog-12, FES-I, Modular Resource-Use Measure, and EQ-5D-5L, the rates were 98.5%.

XSL•FO
RenderX

Table . Exploratory data analysis of proposed outcomes for the full trial.

Outcomes	Preassessment	Postassessment	95% CI of the difference
TUG[a] (s; n=31), mean (SD)	10.32 (4.17)	8.80 (5.74)	−0.94 to 0.99
TUG-Cog[b] (s; n=31), mean (SD)	12.13 (7.66)	11.51 (8.52)	−1.90 to 0.94
ECog-12[c] (points, n=33), mean (SD)	1.57 (0.36)	1.52 (0.39)	−0.21 to 0.48
FES-I[d] (points, n=33), mean (SD)	29.50 (11.25)	25.00 (9.75)	−5.50 to −1.50
High concern, n (%)	20 (60.61)	14 (42.42)	—[e]
Moderate concern, n (%)	12 (36.36)	13 (39.39)	—
Low concern, n (%)	1 (3.03)	6 (18.18)	—
Number of falls in last 6 months (n=33), median (IQR)	2 (0 - 8)	0 (0 - 10)	—
EQ-5D-5L[f] tariff (n=33), mean (SD)	0.675 (0.191)	0.652 (0.212)	−0.316 to 0.367
EQ-5D-5L VAS (n=33), mean (SD)	73.79 (16.91)	71.30 (17.54)	−0.13 to 0.56
Number of A&E[g] visits (sum, n=33)[h]	16	10	−0.5 to 0.00

[a]TUG: Timed Up and Go test (seconds; lower scores indicate better mobility).

[b]TUG-Cog: Timed Up and Go test with a concurrent cognitive task (seconds; lower scores indicate better dual-task performance).

[c]ECog-12: The Everyday Cognition scales short version (higher scores indicate greater perceived cognitive difficulty).

[d]FES-I: Falls Efficacy Scale International (higher scores indicate greater fear of falling).

[e]Not applicable.

[f]Higher scores indicate better health-related quality of life.

[g]A&E: accident and emergency.

[h]This variable corresponds to the first item of the Modular Resource-Use Measure; the full results are reported in Supplementary Material 5-6 in Multimedia Appendix 1.

Table . Changes in the category of Fall Efficacy Scale–International (FES-I)[a].

FES-I cutoff scores	Postassessment		
	High	Moderate	Low
Preassessment			
High[a]	11	2	7
Moderate	0	1	0
Low	3	3	6

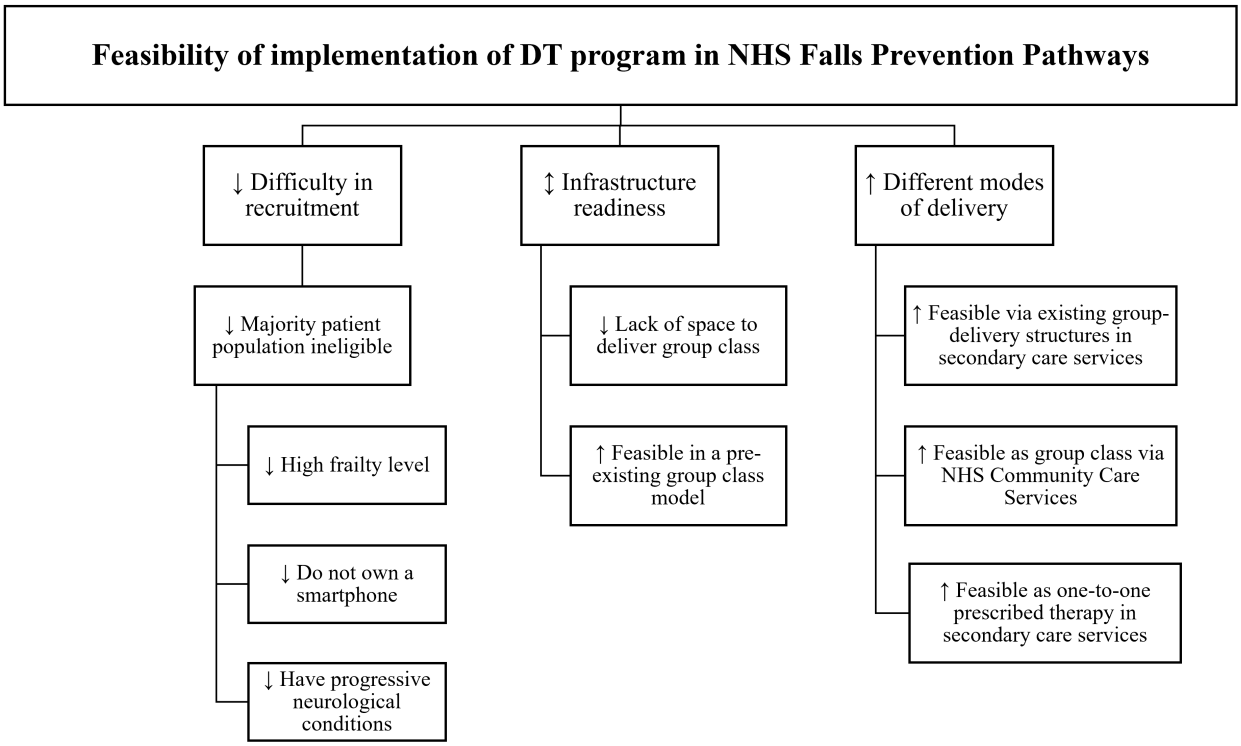
[a]High concern: 28 - 64; moderate concern: 20 - 27; low concern: 16 - 19 [29].

Implementation of the Intervention in Health Care and Community Pathways

A thematic tree illustrating the qualitative results is presented in Figure 3. Health care professionals viewed recruitment from the NHS as challenging, as many patients attending fall-prevention clinics were too frail, had comorbidities (eg, syncope), lacked the capacity for unsupervised balance exercises, or did not have access to suitable devices for cognitive training. Several suggested widening recruitment beyond falls services to include people under 65 years or those beginning to notice balance difficulties, allowing issues to be managed earlier.

XSL•FO
RenderX

Figure 3. Thematic tree from the focus group with health care professionals. Symbols indicate positive (↑), negative (↓), and neutral (↕) feedback. DT: dual-task; NHS: National Health Service.



Implementation was seen as dependent on hospitals. Hospitals already offering group classes were viewed as best suited for integration, while those without such provisions were less likely to consider adopting the program. One-to-one delivery was seen as a possible option if supported by evidence and a wider range of exercise levels to enable personalization. Many professionals also emphasized that delivery through community organizations, accessed via health referral pathways, would reduce travel demands for older adults. Qualitative quotes are in Supplementary Material 4 in Multimedia Appendix 1.

Discussion

Principal Findings

This study demonstrates that a blended, technology-supported DT training program can be both feasible and acceptable for older adults who have experienced at least 2 falls in the past 12 months (including near misses). The combination of supervised group sessions and self-directed home training was well received, with digital engagement enhanced by appropriate support. With design modifications to improve adherence and retention, DT training can be effectively delivered in community settings, representing a shift from clinic-based rehabilitation models that typically rely on one-to-one supervised, therapist-led exercise interventions [37].

The study's key contribution lies in demonstrating that complex interventions, such as DT training, can be adapted for scalable, community-based formats that integrate digital health technologies. This approach has the potential to extend the reach of evidence-based fall-prevention strategies and support their integration into routine community health care pathways.

Participant Recruitment

Recruitment via NHS fall-prevention and primary care pathways—using emails, letters, or texts—was difficult and met the red stop-go criteria. This aligns with prior trials on the physical health of older adults reporting low general practitioner recruitment rates (~6%) [38-40]. Recruiting older adults into exercise studies is widely recognized as challenging. Many older adults experience complex health needs, chronic pain, and multiple comorbidities, all of which may limit their capacity or willingness to participate in research [41]. In our study, health care professionals within NHS Falls Prevention Services also reported a limited pool of eligible participants. Many of their patients were frail, dependent on caregivers, or living with progressive neurological conditions, such as dementia, which restricted their ability to safely participate in group-based or self-directed, unsupervised exercise programs. The literature further highlights several barriers to recruiting older adults into exercise studies. These include social stigma associated with programs targeted at "older people" or " fallers", logistical barriers such as transportation, concerns about the demands or expectations of participation [42-44], and protective attitudes from family members [41].

To address these challenges, we suggest several strategies for future trials. These include working with patient groups to develop more acceptable and inclusive language when describing the study, forming partnerships with community organizations to help address logistical barriers, and framing study aims in ways that appeal to older adults with diverse motivations for participation.

Limited contact (a single invitation) and ethical restrictions on reminders likely contributed to our low response rate. Evidence

XSL•FO
RenderX

supports the use of multiple prompts [41], and Cochrane reviews recommend telephone follow-up reminders in low-response trials [42].

Despite limited success with NHS-based recruitment, the study did enroll 45 of the planned 50 participants (90% of the target), the majority (41/45) through community-based strategies. Face-to-face events and flyers distributed via local organizations were more effective, with in-person engagement supporting informed decision-making—especially important for digital interventions unfamiliar to some older adults [45].

Although NHS-based recruitment met the red stop-go criteria, the study still achieved 90% of its recruitment target, primarily through community routes—demonstrating feasibility for a larger trial. Expanding recruitment to include services such as frailty front door services and fracture liaison services could enhance reach and inclusivity.

Study Retention

Retention met the amber stop-go criteria, indicating a need for trial design modifications. The main reasons for withdrawal were the distance of class venues from participants' homes and perceptions that the program was not sufficiently tailored to individual needs. Classes were scheduled at different times of the day but were restricted to a single day each week due to limitations in staff availability and venue costs. Increasing delivery resources and expanding venue options could improve accessibility and retention [46]. Additionally, incorporating greater flexibility and personalization may boost relevance and continued engagement [47].

Adherence and Acceptability of the Intervention

Adherence met the amber stop-go criteria, indicating the need for trial design modifications—particularly for the unsupervised home-based phase. Consistent with previous evidence, adherence to home-based exercise was lower compared with the group-based programs, particularly when structured supervision and support were reduced [38,48-50]. For comparison, a meta-analysis of community-based group exercise reported an average adherence of 69% (SD 14.6) [51]. Participants valued the motivation, enjoyment, and accountability of supervised group classes but reported reduced motivation, competing priorities, and difficulties sustaining routines once the program became self-directed. Other programs with ongoing support (eg, coaching [52], follow-up calls [19]) have achieved higher adherence rates (75% and 85%, respectively).

Social interaction emerged as a key driver of engagement with participants, suggesting that peer-support groups; virtual group sessions; and reminders via text, phone calls, or app notifications could help support longer-term engagement. Additionally, in-app feedback and a sense of progress also supported adherence [53].

These findings highlight that continuous support, rather than delivery format alone, is key to sustaining engagement among older adults. While supervised sessions help build initial confidence, long-term adherence depends on ongoing social interaction, feedback, and motivational support. Incorporating continuous support mechanisms—delivered in person, virtually, or through structured follow-ups—can help maintain participation and maximize program effectiveness over time.

Implementation of the Intervention in Health Care and Community Pathways

Health care professionals' views on implementing the DT program within NHS services were shaped by patient profiles and infrastructure. According to National Institute for Health and Care Excellence guidelines on falls assessment and prevention, patients referred to fall-prevention clinics are typically frail, present with multiple comorbidities, or have impaired mobility [21]. These individuals often require close supervision (eg, one-to-one delivery) and are therefore not well suited to a program that includes group-based and unsupervised components. The findings from the qualitative data suggest that clinicians in fall-prevention services tended to view this supervised and self-directed technology-based DT program as more appropriate for preventive care in at-risk populations, rather than as a treatment option for those already attending specialist services.

While established programs (eg, Otago and Falls Management Exercise) primarily target physical components of falls risk (eg, strength and balance), the DT program additionally trains executive functions such as attention allocation and task switching during movement—skills that are important for navigating real-world environments where falls often occur during multitasking. This approach may complement existing programs by addressing cognitive-motor demands that are not explicitly targeted in the existing interventions.

In contrast, clinicians working in preventive pathways reported that many of their patients could benefit from the program and were more receptive to its integration into the NHS. Infrastructure was also identified as a key factor influencing implementation. The findings from the qualitative interviews indicated that a lack of space for group classes was commonly cited by health care professionals as a barrier to implementation, whereas services with existing exercise programs expressed greater readiness to adopt the intervention (Figure 3). While the intervention in this feasibility study was delivered by physiotherapists from NHS Falls Prevention Services, health care professionals indicated that, in routine practice, the program could be delivered by appropriately trained exercise professionals, such as therapy assistants or support workers, personal trainers, and qualified community-based exercise instructors. Embedding the program within primary care and community-based infrastructures—such as Primary Care Network referral pathways, social prescribers, local support groups, and leisure providers—offers a scalable, sustainable model aligned with evidence supporting high-dose, balance-challenging exercise [54] and may enhance accessibility and adherence through community engagement and health care referral pathways. These settings often already use qualified instructors and provide accessible venues, which may facilitate wider reach and long-term delivery.

Strengths and Limitations

A key strength of this feasibility study was its mixed methods approach, which added valuable context to the quantitative

XSL•FO
RenderX

outcomes. High completion rates of the exploratory outcome measures among participants who completed the study (>90%) support the suitability of the data collection methods used in this study for older adults. However, the participants recruited lacked ethnic diversity and were skewed toward participants with higher education. While exercise frequency was captured via participant diaries, we did not collect data on exercise intensity (eg, heart rate, perceived exertion) or exercise quality during independent home sessions. This limits our ability to fully assess intervention fidelity and determine whether differences between phases may have been influenced by variations in exercise dose or execution. Future trials should incorporate objective or structured monitoring of home-based exercise intensity and quality to strengthen fidelity assessment. Furthermore, conducting the study exclusively in English—due to funding constraints—also limited inclusivity. Future trials should consider translation services to improve representativeness and reach.

Conclusions

This feasibility study shows that a blended DT training program is acceptable to older adults and can be delivered outside clinical settings. Adherence declined when participants were without direct supervision, highlighting the importance of ongoing support and social connectedness for sustained engagement. Community-based recruitment outperformed NHS pathways, and health care professionals identified community organizations and referral schemes as the most viable implementation channels. Overall, these results support progression to a full trial with adaptations to recruitment strategies, support mechanisms, and delivery pathways to enhance feasibility and scalability.

Acknowledgments

The authors are grateful to all the study participants for their time and generosity. They also thank the Hawkesley Community Centre and Longbridge Retirement Village for kindly providing venues free of charge. Special thanks go to Liz Hensel and Tom Tierney for their valuable contributions on the intervention design and results interpretation as members of the patient and public involvement and engagement group.

Funding

This project is funded by the National Institute for Health and Care Research under its Research for Patient Benefit Programme (grant NIHR205389). The views expressed are those of the author(s) and not necessarily those of the National Institute for Health and Care Research or the Department of Health and Social Care.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: PM, AS, VAG, TK, HT, AC, PK, CM, NI, MC, DW, LM, SYC
Formal analysis: PM, SYC, AS, VAG, PK
Funding acquisition: SYC, MC, AS, VAG, CM, NI, LM, PK, DW
Investigation: PM, TK
Methodology: PM, AS, VAG, TK, HT, AC, PK, CM, NI, MC, DW, LM, SYC
Writing – original draft: PM, SYC
Writing – review & editing: PM, AS, VAG, TK, HT, AC, PK, CM, NI, MC, DW, LM, SYC

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of cognitive games, focus group/interview schedules, exit surveys results, qualitative data with quotes, ModRUM results, and unit costs used in the calculation of ModRUM results.
[DOCX File, 62 KB - aging_v9i1e87577_app1.docx]

Checklist 1
CONSORT 2010 checklist.
[DOCX File, 29 KB - aging_v9i1e87577_app2.docx]

Checklist 2
TIDieR checklist.
[DOCX File, 28 KB - aging_v9i1e87577_app3.docx]

XSL•FO
RenderX

References

1. Falls: applying all our health. GOVUK. 2022. URL: https://www.gov.uk/government/publications/falls-applying-all-our-health/falls-applying-all-our-health#:~:text=medications%2C%20especially%20corticosteroids.-,Focusing%20on%20falls%20in%20your%20professional%20practice,both%20high%20volume%20and%20costly [accessed 2025-10-07]
2. Jehu DA, Davis JC, Falck RS, et al. Risk factors for recurrent falls in older adults: a systematic review with meta-analysis. Maturitas 2021 Feb;144:23-28. [doi: 10.1016/j.maturitas.2020.10.021] [Medline: 33358204]
3. Song K, Portwood C, Jindal J, et al. Preventable deaths involving falls in England and Wales, 2013-22: a systematic case series of coroners' reports. Age Ageing 2023 Oct 2;52(10):10. [doi: 10.1093/ageing/afad191] [Medline: 37847796]
4. Ageing and health. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-10-09]
5. Independent expert on the enjoyment of all human rights by older persons. Office of the High Commissioner for Human Rights. URL: https://www.ohchr.org/en/special-procedures/ie-older-persons/about-human-rights-older-persons [accessed 2026-04-28]
6. Gale CR, Allerhand M, Sayer AA, Cooper C, Deary IJ. The dynamic relationship between cognitive function and walking speed: the English Longitudinal Study of Ageing. Age (Dordr) 2014;36(4):9682. [doi: 10.1007/s11357-014-9682-8] [Medline: 24997019]
7. Rieker JA, Reales JM, Muiños M, Ballesteros S. The effects of combined cognitive-physical interventions on cognitive functioning in healthy older adults: a systematic review and multilevel meta-analysis. Front Hum Neurosci 2022;16:838968. [doi: 10.3389/fnhum.2022.838968] [Medline: 35399365]
8. Wang X, Pi Y, Chen P, Liu Y, Wang R, Chan C. Cognitive motor interference for preventing falls in older adults: a systematic review and meta-analysis of randomised controlled trials. Age Ageing 2015 Mar;44(2):205-212. [doi: 10.1093/ageing/afu175] [Medline: 25377745]
9. Gallou-Guyot M, Mandigout S, Bherer L, Perrochon A. Effects of exergames and cognitive-motor dual-task training on cognitive, physical and dual-task functions in cognitively healthy older adults: an overview. Ageing Res Rev 2020 Nov;63:101135. [doi: 10.1016/j.arr.2020.101135] [Medline: 32768594]
10. Wong PL, Cheng SJ, Yang YR, Wang RY. Effects of dual task training on dual task gait performance and cognitive function in individuals with Parkinson disease: a meta-analysis and meta-regression. Arch Phys Med Rehabil 2023 Jun;104(6):950-964. [doi: 10.1016/j.apmr.2022.11.001] [Medline: 36574531]
11. Weng WH, Yeh NC, Yang YR, Wang RY. Effects of motor-cognitive training on cognitive function and gait performance in older adults with dementia: a systematic review and meta-analysis. Sci Rep 2025 Jul 10;15(1):24915. [doi: 10.1038/s41598-025-09582-y] [Medline: 40640327]
12. Grau-Pellicer M, Chamarro-Lusar A, Medina-Casanovas J, Serdà Ferrer BC. Walking speed as a predictor of community mobility and quality of life after stroke. Top Stroke Rehabil 2019 Jul;26(5):349-358. [doi: 10.1080/10749357.2019.1605751] [Medline: 31063439]
13. Fraser SA, Li KZH, Berryman N, et al. Does combined physical and cognitive training improve dual-task balance and gait outcomes in sedentary older adults? Front Hum Neurosci 2016;10:688. [doi: 10.3389/fnhum.2016.00688] [Medline: 28149274]
14. Sipilä S, Tirkkonen A, Savikangas T, et al. Effects of physical and cognitive training on gait speed and cognition in older adults: a randomized controlled trial. Scand J Med Sci Sports 2021 Jul;31(7):1518-1533. [doi: 10.1111/sms.13960] [Medline: 33772877]
15. Sullivan AN, Lachman ME. Behavior change with fitness technology in sedentary adults: a review of the evidence for increasing physical activity. Front Public Health 2016;4:289. [doi: 10.3389/fpubh.2016.00289] [Medline: 28123997]
16. Valenzuela T, Okubo Y, Woodbury A, Lord SR, Delbaere K. Adherence to technology-based exercise programs in older adults: a systematic review. J Geriatr Phys Ther 2018;41(1):49-61. [doi: 10.1519/JPT.0000000000000095] [Medline: 27362526]
17. Ashe MC, Dos Santos IK, Erome J, Grant J, Mollins J, Soh SE. Systematic review of adherence to technology-based falls prevention programs for community-dwelling older adults: reimagining future interventions. PLOS Digit Health 2024 Sep;3(9):e0000579. [doi: 10.1371/journal.pdig.0000579] [Medline: 39226315]
18. Taylor ME, Ambrens M, Hawley-Hague H, et al. Implementation of a digital exercise programme in health services to prevent falls in older people. Age Ageing 2024 Aug 6;53(8):afae173. [doi: 10.1093/ageing/afae173] [Medline: 39113467]
19. Callisaya ML, Jayakody O, Vaidya A, Srikanth V, Farrow M, Delbaere K. A novel cognitive-motor exercise program delivered via a tablet to improve mobility in older people with cognitive impairment—Standingtall Cognition and Mobility. Exp Gerontol 2021 Sep;152:111434. [doi: 10.1016/j.exger.2021.111434] [Medline: 34098009]
20. Papi E, Chiou SY, McGregor AH. Feasibility and acceptability study on the use of a smartphone application to facilitate balance training in the ageing population. BMJ Open 2020 Dec 2;10(12):e039054. [doi: 10.1136/bmjopen-2020-039054] [Medline: 33268409]
21. Falls: assessment and prevention in older people and in people 50 and over at higher risk. : National Institute for Health and Care Excellence: NICE; 2025 URL: https://www.nice.org.uk/guidance/ng249/resources/

XSL•FO
RenderX

falls-assessment-and-prevention-in-older-people-and-in-people-50-and-over-at-higher-risk-pdf-66143964997573 [accessed 2026-04-28]
22. Mathur P, Thomas H, Cooper A, et al. Supervised and self-directed technology-based dual-task exercise training programme for older adults at risk of falling—protocol for a feasibility study. PLoS ONE 2025;20(3):e0314829. [doi: 10.1371/journal.pone.0314829] [Medline: 40127074]
23. Sim J, Lewis M. The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. J Clin Epidemiol 2012 Mar;65(3):301-308. [doi: 10.1016/j.jclinepi.2011.07.011] [Medline: 22169081]
24. Teare MD, Dimairo M, Shephard N, Hayman A, Whitehead A, Walters SJ. Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. Trials 2014 Jul 3;15:264. [doi: 10.1186/1745-6215-15-264] [Medline: 24993581]
25. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j..2002.384.doc.x] [Medline: 15189396]
26. Peak. 2024. URL: https://www.peak.net [accessed 2026-04-28]
27. Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. Phys Ther 2000 Sep;80(9):896-903. [Medline: 10960937]
28. Tomaszewski Farias S, Mungas D, Harvey DJ, Simmons A, Reed BR, DeCarli C. The measurement of everyday cognition: development and validation of a short form of the everyday cognition scales. Alzheimers Dement 2011 Nov;7(6):593-601. [doi: 10.1016/j.jalz.2011.02.007] [Medline: 22055976]
29. Delbaere K, Close JCT, Mikolaizak AS, Sachdev PS, Brodaty H, Lord SR. The falls efficacy scale international (FES-I). a comprehensive longitudinal validation study. Age Ageing 2010 Mar;39(2):210-216. [doi: 10.1093/ageing/afp225] [Medline: 20061508]
30. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
31. Garfield K, Thorn JC, Noble S, Husbands S, Hollingworth W. Development of a brief, generic, modular resource-use measure (ModRUM): piloting with patients. BMC Health Serv Res 2023 Sep 15;23(1):994. [doi: 10.1186/s12913-023-10011-x] [Medline: 37710265]
32. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005 Aug 30;173(5):489-495. [doi: 10.1503/cmaj.050051] [Medline: 16129869]
33. Rockwood K, Theou O. Using the clinical frailty scale in allocating scarce health care resources. Can Geriatr J 2020 Sep;23(3):210-215. [doi: 10.5770/cgj.23.463] [Medline: 32904824]
34. Furber C. Framework analysis: a method for analysing qualitative data. Afr J Midwifery Womens Health 2010 Apr;4(2):97-100. [doi: 10.12968/ajmw.2010.4.2.47612]
35. Tracy SJ. Qualitative quality: Eight “Big-Tent” criteria for excellent qualitative research. Qual Inq 2010 Dec;16(10):837-851. [doi: 10.1177/1077800410383121]
36. Smith B, McGannon KR. Developing rigor in qualitative research: problems and opportunities within sport and exercise psychology. Int Rev Sport Exerc Psychol 2018 Jan;11(1):101-121. [doi: 10.1080/1750984X.2017.1317357]
37. Finnegan S, Bruce J, Skelton DA, Withers EJ, Lamb SE, PreFIT Study Group. Development and delivery of an exercise programme for falls prevention: the Prevention of Falls Injury Trial (PreFIT). Physiotherapy 2018 Mar;104(1):72-79. [doi: 10.1016/j.physio.2017.06.004] [Medline: 28801033]
38. Iliffe S, Kendrick D, Morris R, et al. Promoting physical activity in older people in general practice: ProAct65+ cluster randomised controlled trial. Br J Gen Pract 2015 Nov;65(640):e731-e738. [doi: 10.3399/bjgp15X687361] [Medline: 26500320]
39. Stathi A, Greaves CJ, Thompson JL, et al. Effect of a physical activity and behaviour maintenance programme on functional mobility decline in older adults: the REACT (Retirement in Action) randomised controlled trial. Lancet Public Health 2022 Apr;7(4):e316-e326. [doi: 10.1016/S2468-2667(22)00004-4] [Medline: 35325627]
40. Brach M, Moschny A, Bücker B, et al. Recruiting hard-to-reach subjects for exercise interventions: a multi-centre and multi-stage approach targeting general practitioners and their community-dwelling and mobility-limited patients. Int J Environ Res Public Health 2013 Dec 2;10(12):6611-6629. [doi: 10.3390/ijerph10126611] [Medline: 24317380]
41. Knechel NA. The challenges of enrolling older adults into intervention studies. Yale J Biol Med 2013 Mar;86(1):41-47. [Medline: 23482244]
42. Cavill NA, Foster CEM. Enablers and barriers to older people’s participation in strength and balance activities: a review of reviews. J Frailty Sarcopenia Falls 2018 Jun;3(2):105-113. [doi: 10.22540/JFSF-03-105] [Medline: 32300698]
43. Cherubini A, Del Signore S, Ouslander J, Semla T, Michel JP. Fighting against age discrimination in clinical trials. J Am Geriatr Soc 2010 Sep;58(9):1791-1796. [doi: 10.1111/j.1532-5415.2010.03032.x] [Medline: 20863340]
44. Ridda I, MacIntyre CR, Lindley RI, Tan TC. Difficulties in recruiting older people in clinical trials: an examination of barriers and solutions. Vaccine (Auckl) 2010 Jan 22;28(4):901-906. [doi: 10.1016/j.vaccine.2009.10.081] [Medline: 19944149]
45. Hung M, Mohajeri A, Almpani K, et al. Successes and challenges in clinical trial recruitment: the experience of a new study team. Med Sci (Basel) 2024 Aug 14;12(3):39. [doi: 10.3390/medsci12030039] [Medline: 39189202]

XSL•FO
RenderX

46. Buttgereit T, Palmowski A, Forsat N, et al. Barriers and potential solutions in the recruitment and retention of older patients in clinical trials-lessons learned from six large multicentre randomized controlled trials. Age Ageing 2021 Nov 10;50(6):1988-1996. [doi: 10.1093/ageing/afab147] [Medline: 34324628]
47. Pitkala KH, Strandberg TE. Clinical trials in older people. Age Ageing 2022 May 1;51(5):afab282. [doi: 10.1093/ageing/afab282] [Medline: 35526845]
48. Gawler S, Skelton DA, Dinan-Young S, et al. Reducing falls among older people in general practice: the ProAct65+ exercise intervention trial. Arch Gerontol Geriatr 2016;67:46-54. [doi: 10.1016/j.archger.2016.06.019] [Medline: 27420150]
49. Liu WT, Wang CH, Lin HC, et al. Efficacy of a cell phone-based exercise programme for COPD. Eur Respir J 2008 Sep;32(3):651-659. [doi: 10.1183/09031936.00104407] [Medline: 18508824]
50. Boongird C, Keesukphan P, Phiphadthakusolkul S, Rattanasiri S, Thakkinstian A. Effects of a simple home-based exercise program on fall prevention in older adults: a 12-month primary care setting, randomized controlled trial. Geriatr Gerontol Int 2017 Nov;17(11):2157-2163. [doi: 10.1111/ggi.13052] [Medline: 28436154]
51. Farrance C, Tsofliou F, Clark C. Adherence to community based group exercise interventions for older people: a mixed-methods systematic review. Prev Med 2016 Jun;87:155-166. [doi: 10.1016/j.ypmed.2016.02.037] [Medline: 26921655]
52. van den Helder J, Mehra S, van Dronkelaar C, et al. Blended home-based exercise and dietary protein in community-dwelling older adults: a cluster randomized controlled trial. J Cachexia Sarcopenia Muscle 2020 Dec;11(6):1590-1602. [doi: 10.1002/jcsm.12634] [Medline: 33103379]
53. Arensman R, Kloek C, Pisters M, Koppenaal T, Ostelo R, Veenhof C. Patient perspectives on using a smartphone app to support home-based exercise during physical therapy treatment: qualitative study. JMIR Hum Factors 2022 Sep 13;9(3):e35316. [doi: 10.2196/35316] [Medline: 36098993]
54. Lesinski M, Hortobágyi T, Muehlbauer T, Gollhofer A, Granacher U. Effects of balance training on balance performance in healthy older adults: a systematic review and meta-analysis. Sports Med 2015 Dec;45(12):1721-1738. [doi: 10.1007/s40279-015-0375-y] [Medline: 26325622]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
DT: dual-task
FES-I: Falls Efficacy Scale-International
NHS: National Health Service
PPIE: patient and public involvement and engagement
REDCap: Research Electronic Data Capture
TIDieR: Template for Intervention Description and Replication
TUG: Timed Up and Go

Edited by A Garcia; submitted 11.Nov.2025; peer-reviewed by SJ Meredith, Z Zeng; revised version received 23.Mar.2026; accepted 17.Apr.2026; published 18.May.2026.

Please cite as:
Mathur P, Stathi A, Goodyear VA, Krauss T, Thomas H, Cooper A, Kinghorn P, Miller C, Ives N, Chechlacz M, Wilson D, Magill L, Chiou SY
Supervised and Self-Directed Technology-Based Dual-Task Exercise Training Program for Older Adults With a History of Falls: Mixed Methods Feasibility Study
JMIR Aging 2026;9:e87577
URL: https://aging.jmir.org/2026/1/e87577
doi:10.2196/87577

© Prerna Mathur, Afroditi Stathi, Victoria A Goodyear, Taylor Krauss, Helen Thomas, Angela Cooper, Philip Kinghorn, Caroline Miller, Natalie Ives, Magdalena Chechlacz, Daisy Wilson, Laura Magill, Shin-Yi Chiou. Originally published in JMIR Aging (https://aging.jmir.org), 18.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Sleep Quality on Gait Variability: Pilot Cohort Study

Francesca R Marino[1], PhD; Minzae Kim[2]; Phillip H Hwang[1,3,4], MPH, PhD; Yorghos Tripodis[5,6], PhD; Michael Alosco[1,5,7], PhD; Jesse Mez[1,4,5,7], MS, MD; Ann McKee[5,7,8,9], MD; Rhoda Au[1,2,3,4,5,7], PhD; Robert Joseph Thomas[10], MD; Ludy Shih[11], MD

[1]Department of Anatomy and Neurobiology, Chobanian and Avedisian School of Medicine, Boston University, 72 E Concord Street, Boston, MA, United States

[10]Department of Medicine, Division of Pulmonary, Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

[11]Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

[2]Department of Medicine, Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, United States

[3]Department of Epidemiology, School of Public Health, Boston University, Boston, MA, United States

[4]The Framingham Heart Study, Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, United States

[5]Alzheimer's Disease Research Center and CTE Center, Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, United States

[6]Department of Biostatistics, School of Public Health, Boston University, Boston, MA, United States

[7]Department of Neurology, Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, United States

[8]VA Bedford Healthcare System, US Department of Veteran Affairs, Bedford, MA, United States

[9]Department of Pathology and Laboratory Medicine, Chobanian and Avedisian School of Medicine, Boston University, Boston, MA, United States

Corresponding Author:
Francesca R Marino, PhD
Department of Anatomy and Neurobiology, Chobanian and Avedisian School of Medicine, Boston University, 72 E Concord Street, Boston, MA, United States

Abstract

Background: Higher step width variability while walking is associated with poor physical function and falls. Sleep is an established modifiable risk factor for both gait and physical function impairments, but it remains to be examined whether sleep is also related to step width variability.

Objective: This study aimed to evaluate the cross-sectional associations between objectively measured sleep quality, using cardiopulmonary coupling spectrograms, and step width variability during a preferred walking condition among middle-aged and older adults.

Methods: This study included 72 adults (mean age 71, SD 8.3 y; n=37, 51% female; n=65, 90% non-Hispanic White) who had ≥2 nights of objectively measured sleep (cardiopulmonary coupling via SleepImage ring) and completed a 10-m walk at preferred speed while wearing inertial sensors (APDM Mobility Lab). Sleep measures included sleep duration, efficiency, fragmentation, stability, apnea-hypopnea index, percentage of time oxygen saturation level <90%, oxygen desaturation index, and respiratory disturbance index. Additional derived sleep variables were explored using least absolute shrinkage and selection operator models. Step width variability was defined by the asymmetry of lateral step variability and categorized as medial (≤−7.5 cm), minimal (within ±7.5 cm; reference), or lateral displacement (≥7.5 cm). Multinomial logistic regression models adjusted for age, sex, race, education, BMI, and usual gait speed evaluated cross-sectional associations between sleep and step width variability categories.

Results: We found that a 1% higher sleep fragmentation was associated with a 6% higher probability of step width variability ≥7.5 cm (95% CI 1.01 - 1.11), while a 1% higher sleep stability was associated with a 5% lower probability of variability ≥7.5 cm (95% CI 0.91 - 0.99), compared to minimal variability. From the least absolute shrinkage and selection operator models, we found that a 1% higher sleep quality index, a 1% higher rapid eye movement sleep, a 1-second shorter apnea duration, and a 1-beat per minute slower mean heart rate were also associated with a lower probability of lateral compared to minimal displacement.

Conclusions: Poor sleep quality was associated with higher step width variability among middle-aged and older adults. This suggests that sleep may be a modifiable risk factor for maintaining postural stability while walking among middle-aged and older adults. Future studies are needed to examine whether intervening in these sleep measures also lowers the risk of falls.

(JMIR Aging 2026;9:e81630) doi:10.2196/81630

KEYWORDS

sleep; gait; variability; wearables; sensors; older adults

XSL•FO
RenderX

Introduction

Ambulating within one's environment is a complex task that requires cognitive coordination of the motor, vestibular, proprioceptive, and sensory systems [1]. This is typically an automatic task that has low variability [1,2]. Variability in gait reflects stride-to-stride fluctuations in parameters such as gait speed, stride length, step time, and other spatiotemporal characteristics [2]. Gait variability tends to increase with age as ambulation becomes less automatic [3,4]. A key variability measure is step width variability, which reflects dynamic postural control and balance [5-7]. Alterations in gait spatiotemporal characteristics, such as step width variability, impact gait biomechanics. Indeed, variations in step width are associated with alterations in stride time, step length, lower extremity muscle activity, as well as rearfoot, ankle, joint (eg, knee, hip), or trunk kinematics, all of which are related to stability and posture during walking [8]. Step width variability is a marker of poor physical function, as higher variability is associated with slower usual gait speed [9], longer time to complete chair stands [9], and a higher risk of falls [6,10]. Given the high prevalence of falls and physical function impairment among older adults [11,12], studies are needed to identify modifiable risk factors to help maintain step width variability in later life.

Sleep is an established risk factor for both gait and physical function impairments. A recent systematic review found that sleep efficiency, sleep duration, and wake after sleep onset were related to motor outcomes, such as gait speed, time to complete chair stands, and postural control [13]. Among older adults, lower self-reported sleep quality is associated with falls and slower gait speed [14], longer sleep latency is associated with an increased risk of mobility difficulty [15], and shorter sleep duration is associated with declines in gait speed [15]. Self-reported insomnia is also linked with differences in gait speed, stride or step length, and variability in single support, double support, stance, stride or step length, and stride or step time [16]. A study of 49 participants also reported associations between apnea-hypopnea index severity and stride time variability [17]. Most prior studies have focused on gait speed, stride or step length or time variability, or falls, and less is known about the relationships between sleep and step width variability. Many epidemiologic studies also rely on self-reported sleep measures, which may misclassify sleep quality [18,19]. As such, there is a need to evaluate associations between objectively measured sleep quality and step width variability.

Methods to assess sleep have evolved from electroencephalographic (EEG) measurements to actigraphy and autonomic or respiratory signals. One such method is cardiopulmonary coupling (CPC), which integrates electrocortical activity, autonomic activity, and respiration to provide profiles of sleep stability [20,21]. With the increasing availability of wearable devices, the easy assessment of free-living sleep quality is possible. By reducing the need for more traditional in-clinic procedures, such as EEG measurements, wearable devices may also lessen participant burden and improve accessibility [22].

This study aimed to establish associations between sleep and step width variability by leveraging digital data from a clinical sample of middle-aged and older adults. We hypothesized that lower objectively measured sleep quality would be cross-sectionally associated with higher step width variability during a preferred walking speed condition. Given that associations between sleep and gait have been found to differ by sex or cognitive function [14,23,24], we also expected that associations between sleep quality and step width variability would be stronger among men or individuals living with cognitive impairment compared to their respective counterparts.

Methods

Participants

This study included participants enrolled in the Boston University (BU) Alzheimer's Disease Research Center (ADRC) Clinical Core who opted into a digital brain health platform study (Digital Study). In brief, the BU ADRC is one of over 30 ADRCs nationwide that aim to promote collaborative research on Alzheimer's disease and related dementias. The BU ADRC enrolled adults aged ≥50 years with and without memory complaints who were English-speaking, community-dwelling, and local to the greater Boston area. The BU ADRC sample is representative of a clinic-based sample, including people with concerns about their cognition either because they have cognitive complaints or because they have certain risk factors such as head trauma or a family history. Full inclusion and exclusion criteria are published elsewhere, but participants were generally eligible unless they had a history of major psychiatric or neurologic illness (eg, schizophrenia, bipolar disorder, epilepsy) or head injury with significant loss of consciousness [25].

Starting in 2021, BU ADRC participants have had the option to use various technologies both at home and in the clinic as part of the Digital Study. Participants are given the choice of which technologies they would like to use [22]. This study included Digital Study participants who chose to use both sleep and gait digital technologies. All participants were eligible to complete the sleep and gait assessments. Participants were excluded from the analysis if sleep and gait data were measured >90 days apart (n=2) or if covariate data were missing (n=3). The flowchart for inclusion in the study sample is shown in the *Results* section.

Ethical Considerations

All study procedures were approved by the Boston University Medical Campus Institutional Review Board (approval number: H-40542). All participants provided written informed consent. The original informed consent and institutional review board approval allowed secondary analysis of the data without requiring additional participant consent. All data analyzed in this study were deidentified. No financial compensation was provided to participants.

XSL•FO
RenderX

Sleep Measures

Recording Process and Automated Analysis

Participants were asked to wear the SleepImage ring (MyCardio LLC) in the free-living environment for at least 3 nights. The ring contains a photoplethysmography sensor that collects heart rate, oxygen saturation (SpO_2), and actigraphy data. Data were automatically uploaded and processed by SleepImage using the CPC algorithm.

Cardiopulmonary Coupling

The CPC method analyzes changes in signals modulated by the autonomic nervous system (eg, heart or pulse rate variability and breathing) to derive sleep stages, states, and a Sleep Quality Index (SQI). The method is based on calculating coupling and coherence between heart or pulse rate variability (eg, time between peaks of consecutive R-waves [RR interval time series]) and fluctuations in R-wave or pulse-wave amplitude induced by respiration to detect changes in breathing (eg, tidal volume variability [TVV]) [20]. These outputs are strongly modulated by sleep-wake state and stage. The cross-spectral power and coherence of the RR time series and corresponding TVV time series are calculated for consecutive windows, and a product of coherence and cross-spectral power is used to obtain the ratio of coherent cross power in the low frequency (0.01 - 0.1 Hz) to that in the high-frequency band (0.1 - 0.4 Hz). The logarithm of the high-to-low-frequency CPC ratio is then computed to yield a continuous and moving average of overlapping CPC windows and output of stable and unstable nonrapid eye movement (NREM) sleep, REM sleep, and wake. Graphing CPC at relevant frequencies versus time provides the SleepImage spectrogram [20,26]. Stable NREM sleep is characterized by stable breathing and stable oxygenation, high vagal tone, non–cyclic alternating pattern on the EEG, continuous occurrence of slow oscillations, high delta power, blood pressure dipping, and stable arousal threshold. Unstable NREM sleep is a marker of sleep instability and has exactly the opposite features of stable NREM sleep, with variability in TVV, cyclic variation in heart rate, cyclic alternating pattern on EEG, low relative delta power, nondipping of blood pressure, and unstable arousal thresholds. The SQI is a proprietary summary index of the CPC biomarkers of sleep quality, sleep stability, fragmentation, and periodicity, which provides a meaningful unit of measure for sleep health. The SQI is displayed on a scale of 0 to 100 with expected values for both children and adults [27-30]. Sleep spectrograms for participants with higher versus lower SQI values are shown in Figure S1 in Multimedia Appendix 1.

Sleep Measures

The following SleepImage measures were identified a priori: (1) sleep duration (hours), defined as the total hours of sleep per night; (2) sleep efficiency (%), defined as the proportion of time asleep out of the total time in bed; (3) sleep fragmentation (%), defined as the probability of transitioning from sleep to wakefulness; (4) stable sleep (%), defined as the proportion of sleep classified as stable (eg, regular breathing, heart rate, etc); (5) apnea-hypopnea index (events/hour), defined as the number of apnea or hypopnea events with ≥4% desaturation per hour; (6) duration (minutes) or percentage (%) of time SpO_2 levels were <90%; (7) oxygen desaturation index (events/hour), defined as the number of events lasting ≥10 seconds with ≥4% desaturation per hour; and (8) respiratory disturbance index (events/hour), defined as the number of times per hour that breathing was disrupted during sleep. The full methods for calculating these sleep measures are described elsewhere [31]. The remaining variables from the SleepImage report were examined using least absolute shrinkage and selection operator (LASSO) models, as described in the *Statistical Analysis* section.

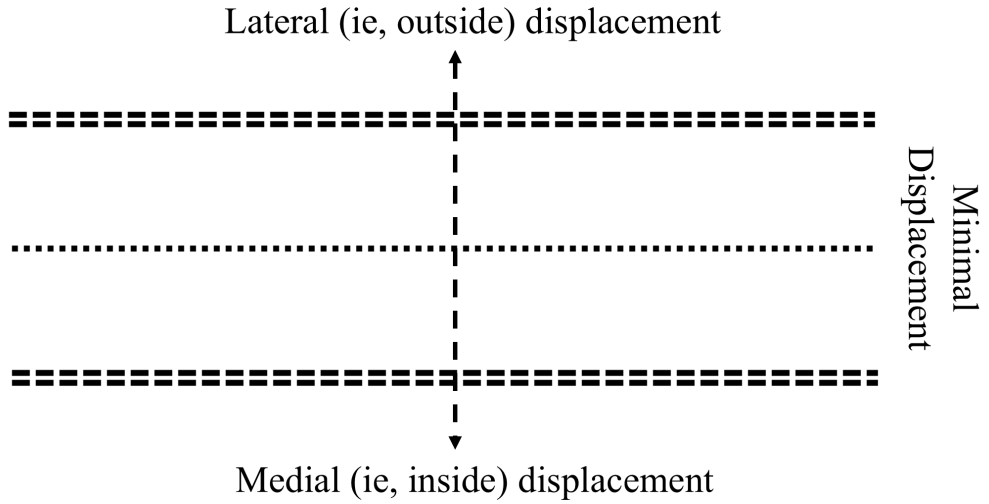
Step Width Variability

As part of the in-clinic visits for the Digital Study, participants completed several walking trials while wearing inertial sensors (Opal; APDM) over the L4-5 vertebra, on each wrist, and on each foot. After the sensors were attached, participants walked back and forth down a 10-m hallway at their preferred comfortable (eg, usual) speed for 1 minute. The number of straight walking bouts varied depending on the participant's usual walking speed. Participants completed this 1-minute walking trial at their usual pace, as well as additional trials at a fast pace, slow pace, and fast-paced dual-task (eg, serial subtractions). Only the usual pace walking trial was included in this analysis. Then, the data were uploaded and processed using Mobility Lab version 2 [32]. Derived variables from these sensors have been validated among healthy adults (mean age 24, SD 6.2 y) [33] and populations with Parkinson disease [34]. These sensors have been used to monitor gait and balance [35-40], and several studies have also evaluated step width variability [36-40].

The outcome of interest was step width variability, as measured by asymmetry in lateral step variability during the preferred comfortable speed walking trial. This was defined as the variability of perpendicular deviations of the middle foot placement from the line connecting the first and third placements, when considering 3 consecutive foot placements made by the same foot [32]. Negative values indicated movement to the inside (eg, medial displacement), whereas positive values indicated movement to the outside (eg, lateral displacement; Figure 1) [32]. Mobility Lab 2 generated 1 summary value of asymmetry in lateral step variability. Mobility Lab 2 generated separate derived variables for information related to turning during the walking trials, which were not used in this analysis. For this study, variability was categorized as displacement within ±7.5 cm (eg, minimal displacement), ≤−7.5 cm (eg, medial displacement), or ≥7.5 cm (eg, lateral displacement). To determine whether greater displacement to either the inside or outside was associated with poor sleep quality, we also calculated the absolute value of asymmetry lateral step variability and analyzed this continuously (eg, higher values reflecting greater displacement in either direction).

XSL•FO
RenderX

Figure 1. Measurement of asymmetry in lateral step variability by Mobility Lab. For an individual walking along the center dashed line, displacement of the foot to the outside is defined as lateral displacement, displacement of the foot to the inside is defined as minimal displacement, and displacement of the foot within the outside and inside limits is defined as minimal displacement.



Covariates

Covariates were measured at each ADRC visit as part of the National Alzheimer's Coordinating Center Uniform Data Set [41]. Age was measured in years. Sex was categorized as male or female. Race was categorized as White, Black, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Asian, other, or unknown. Participants in the current sample identified as White, Black, or Asian. Due to small sample sizes, participants were classified as White or either Black or Asian. Educational attainment (years) was self-reported. BMI was measured as body mass (kg) divided by squared height (m^2). Usual gait speed (m/s) was measured by the APDM sensors [32]. Cognitive status was measured by the Clinical Dementia Rating (CDR) [42]. Participants were classified as having scores of 0 (n=56) or ≥0.5 (n=16; range 0.5 - 2.5).

Statistical Analysis

Sample characteristics were compared overall and by categories of step width variability. Means and SDs or counts and percentages were calculated for continuous or categorical variables, respectively.

Cross-sectional associations between each a priori sleep measure and categories of step width variability were estimated using separate multinomial logistic regression models. These models estimated the relative risk ratio (RRR) and 95% CIs for medial or lateral displacement compared to minimal displacement (eg, reference group). RRR is a measure of association that estimates the probability of being in a particular outcome category (eg, medial or lateral displacement) relative to the reference category (eg, minimal displacement), per 1-unit increase in the exposure variable (eg, sleep measure). All models were adjusted for age, sex, race, education, BMI, and usual gait speed.

LASSO models evaluated the remaining SleepImage features to identify additional sleep variables for differentiating step width variability categories. In brief, this method is a machine learning approach that selects a subset of independent variables (eg, sleep) that best describe a dependent variable (eg, step width variability). A regularization process is applied to shrink the coefficients of variables that are less related to the outcome to zero. Then, variables with a nonzero coefficient are selected. This method is particularly useful for providing higher prediction accuracy and reducing the risk of overfitting in datasets with smaller sample sizes and a larger number of possible variables [43]. The features identified from the LASSO model were examined in separate multinomial regression models adjusted for the same covariates.

We also conducted several sensitivity analyses. First, to determine whether associations differed by certain individual-level characteristics, we stratified each model by sex (men vs women) or cognitive status (CDR of 0 versus ≥0.5). Second, to examine whether poor sleep was associated with higher step width variability irrespective of direction (eg, medial or lateral), we estimated the cross-sectional associations between sleep measures and the absolute value of step width variability using linear regression models. Finally, to evaluate whether associations were driven by extreme values, we repeated the primary analyses after excluding 1 participant with step width variability >57 cm.

The significance level for the multinomial logistic regression models was set to $P<.05$. Due to the number of comparisons, we applied a Benjamini-Hochberg false discovery rate correction (false discovery rate–corrected *P* value [pFDR]) for models estimating associations between either the a priori (9 comparisons) or LASSO-identified (5 comparisons) sleep features with step width variability categories [44]. All analyses were performed using Stata (version 18.0; StataCorp). Figures were generated using R (version 4.2.2).

Results

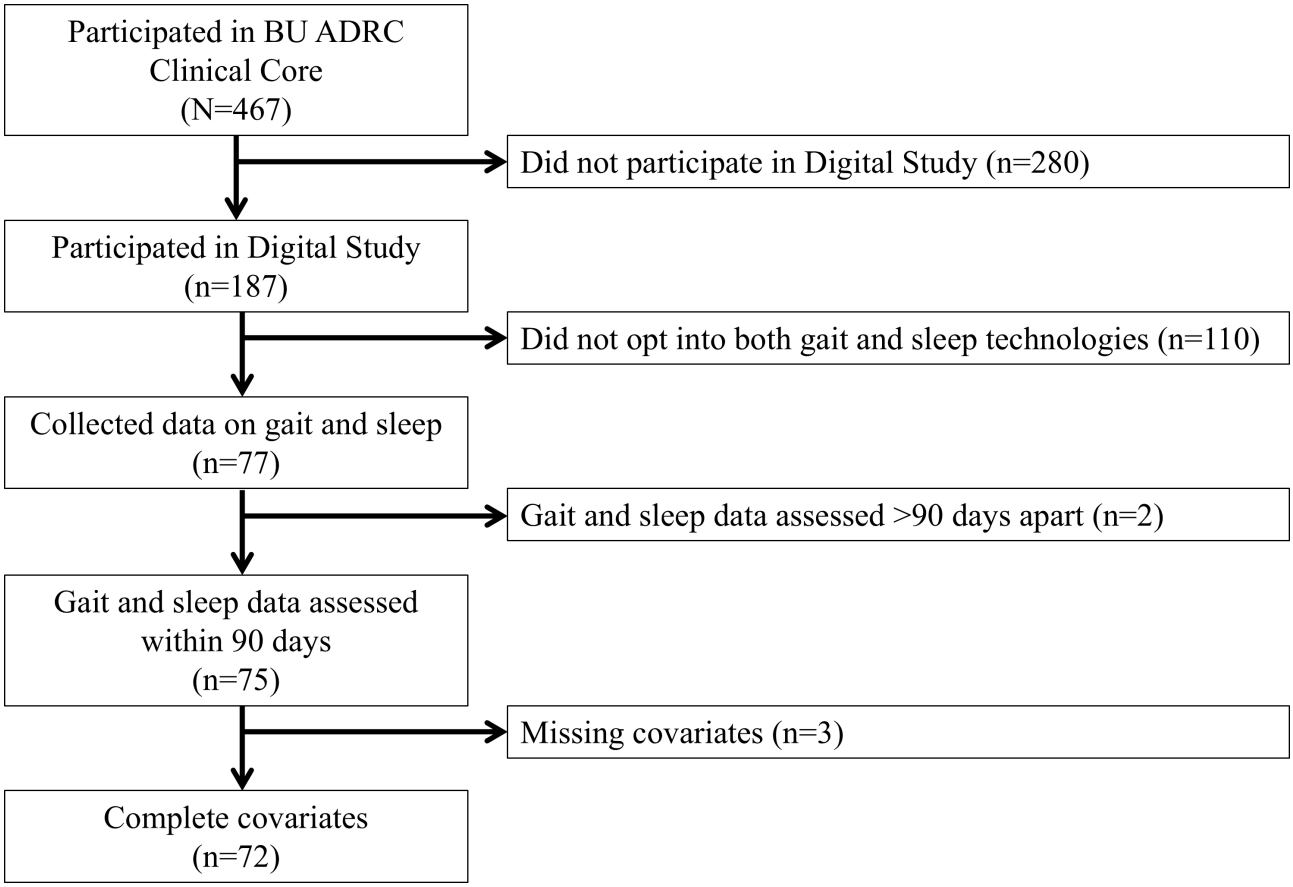
Sample Characteristics

Of the 72 participants in this sample (Figure 2), the mean age was 71 (SD 8.3; range 52 - 88) years, 51% (n=37) were female, 90% (n=65) were non-Hispanic White, and 22% (n=16) had CDR scores ≥0.5. On average, asymmetry in lateral step variability was 1.92 (SD 20.54) cm, with a range of −38.85 to 57.91 cm. For step width variability categories, 31% (n=22)

XSL•FO
RenderX

had minimal, 29% (n=21) had medial, and 40% (n=29) had lateral displacement. Participants with minimal displacement (eg, step width variability within ±7.5 cm) were older; more likely to be White; had higher educational attainment; slower usual gait speed; and lower prevalence of hypertension, diabetes, or cognitive impairment compared to those with medial (eg, ≤−7.5 cm) or lateral (eg, ≥7.5 cm) displacement. However, none of the differences in sample characteristics across step width variability groups were statistically significant (Table 1).

Figure 2. Flowchart for inclusion in the analytic sample. BU ADRC: Boston University Alzheimer's Disease Research Center.



XSL•FO
RenderX

Table . Sample characteristics overall and by step width variability categories.

Characteristics	Overall (N=72)	Minimal[a] (n=22)	Medial[a] (n=21)	Lateral[a] (n=29)	*P* value
Demographics					
Age (y), mean (SD)	71.0 (8.3)	72.2 (9.9)	70.8 (7.4)	70.2 (7.7)	.71
Female, n (%)	37 (51)	11 (50)	9 (43)	17 (59)	.54
Non-Hispanic White, n (%)	65 (90)	22 (100)	19 (90)	24 (83)	.12
Education (y), mean (SD)	16.7 (2.1)	17.5 (1.8)	16.2 (1.8)	16.4 (2.5)	.09
Married, n (%)	53 (74)	16 (73)	16 (76)	21 (72)	.63
Medical conditions, n (%)					
Hypertension	24 (33)	6 (27)	9 (43)	9 (31)	.52
Diabetes	5 (7)	0 (0)	2 (10)	3 (10)	.30
Sleep apnea	22 (31)	7 (32)	7 (33)	8 (28)	.90
Sleep disturbances	10 (14)	4 (18)	3 (14)	3 (10)	.72
CDR[b] ≥0.5	16 (22)	3 (14)	6 (29)	7 (24)	.47
Gait speed (m/s), mean (SD)	1.04 (0.17)	1.01 (0.17)	1.08 (0.14)	1.04 (0.19)	.46
Timing, mean (SD)					
Nights of sleep (n)[c]	6.4 (6.0)	7.2 (7.3)	6.5 (5.9)	5.8 (5.0)	.70
Time lag (days)[d]	12.9 (16.1)	14.0 (19.1)	12.0 (10.7)	12.6 (17.5)	.92

[a]Minimal displacement within ±7.5 cm; medial displacement ≤−7.5 cm; and lateral displacement ≥7.5 cm during the preferred walking speed condition.

[b]CDR: Clinical Dementia Rating scale.

[c]Number of nights of sleep data.

[d]Number of days between sleep and step width variability measurements.

Cross-Sectional Associations Between Sleep and Step Width Variability Categories

Of the sleep features identified a priori, sleep fragmentation and stable sleep percentage were associated with the probability of having high step width variability. After adjustment for age, sex, race, education, BMI, and usual gait speed, a 1% higher sleep fragmentation was associated with a 6% higher probability of lateral displacement, and a 1% higher stable sleep percentage was associated with a 5% lower probability of lateral displacement compared to minimal displacement (fragmentation: RRR 1.06, 95% CI 1.01 - 1.11; stable sleep: RRR 0.95, 95% CI 0.91 - 0.99; Table 2). However, these findings were not statistically significant after FDR correction (Table S1 in Multimedia Appendix 1). There were no associations between sleep duration, sleep efficiency, apnea-hypopnea index, SpO_2 <90%, oxygen desaturation index, or respiratory disturbance index with step width variability categories or between any sleep measure and probability of medial displacement (Table 2). The results were unchanged after excluding 1 participant with step width variability >57 cm (data not shown).

XSL•FO
RenderX

Table . Cross-sectional associations between a priori sleep and step width variability categories (N=72).

Sleep variable	Medial[a,b] (RRR[c], 95% CI)	Lateral[a,b] (RRR, 95% CI)
Sleep duration (h)	1.07 (0.73-1.59)	0.86 (0.59-1.26)
Sleep efficiency (%)	1.02 (0.96-1.09)	0.99 (0.94-1.05)
Sleep fragmentation (%)	1.00 (0.96-1.06)	1.06 (1.01-1.11)
Stable sleep (%)	1.01 (0.96-1.06)	0.95 (0.91-0.99)
AHI[d] (numbers per hour)	1.01 (0.94-1.08)	1.00 (0.94-1.08)
SpO_2[e] <90% (minute)	1.00 (0.98-1.01)	0.99 (0.98-1.01)
SpO_2 <90% (%)	0.99 (0.93-1.05)	0.96 (0.90-1.03)
ODI[f] (numbers per hour)	1.02 (0.95-1.09)	1.01 (0.94-1.08)
RDI[g] (numbers per hour)	1.00 (0.95-1.06)	1.01 (0.96-1.07)

[a]Minimal displacement within ±7.5 cm (reference); medial displacement ≤−7.5 cm; and lateral displacement ≥7.5 cm during the preferred walking speed condition.

[b]Model adjusted for age, sex, race, education, BMI, and usual gait speed.

[c]RRR: relative risk ratio.

[d]AHI: apnea-hypopnea index.

[e]SpO_2: oxygen saturation.

[f]ODI: oxygen desaturation index.

[g]RDI: respiratory disturbance index.

Sleep Measures Identified by LASSO Models

The LASSO model identified SQI, REM sleep percentage, mean SpO_2, minimum apnea duration, and mean heart rate as being related to lateral compared to minimal displacement. After full covariate adjustment, 1% higher SQI and 1% higher REM sleep were associated with lower probability of lateral displacement, whereas a 1-second longer minimum apnea duration and a 1 beat per minute faster mean heart rate were associated with higher probability of lateral displacement compared to minimal displacement (SQI: RRR 0.92, 95% CI 0.87 - 0.98; REM sleep: RRR 0.87, 95% CI 0.77 - 0.99; apnea duration: RRR 2.34, 95% CI 1.06 - 5.16; heart rate: RRR 1.14, 95% CI 1.03 - 1.26; Table 3). The findings related to SQI were robust at the trend-level after FDR correction (P=.09; Table S2 in Multimedia Appendix 1). The LASSO model did not identify any sleep measures associated with medial displacement. However, 1 beat per minute faster mean heart rate was associated with 13% higher probability of medial compared to minimal displacement (95% CI 1.01 - 1.25; Table 3). The results were unchanged after excluding 1 participant with step width variability >57 cm (data not shown).

Table . Cross-sectional associations between the additional sleep measures identified from least absolute shrinkage and selection operator (LASSO) and step width variability categories (N=72).

Sleep variable	Medial[a,b] (RRR[c], 95% CI)	Lateral[a,b] (RRR, 95% CI)
SQI[d] (%)	1.00 (0.95-1.05)	0.92 (0.87-0.98)
REM[e] sleep (%)	0.91 (0.81-1.03)	0.87 (0.77-0.99)
Mean SpO_2[f] (%)	1.01 (0.69-1.48)	1.31 (0.90-1.92)
Minimum apnea duration (s)	1.35 (0.59-3.05)	2.34 (1.06-5.16)
Mean heart rate (BPM[g])	1.13 (1.01-1.25)	1.14 (1.03-1.26)

[a]Minimal displacement within ±7.5 cm (reference); medial displacement ≤−7.5 cm; and lateral displacement ≥7.5 cm during the preferred walking speed condition.

[b]Model adjusted for age, sex, race, education, body mass index, and usual gait speed.

[c]RRR: relative risk ratio.

[d]SQI: sleep quality index.

[e]REM: rapid eye movement.

[f]SpO_2: oxygen saturation.

[g]BPM: beats per minute.

XSL•FO
RenderX

Effect Measure Modification by Sex

Among men (n=35), the findings were robust for stable sleep percentage, SQI, and REM sleep percentage. All other primary findings were attenuated, but the magnitude of the effect sizes was consistent with the primary analyses. Among women (n=37), the findings were robust for mean heart rate. All other primary findings were attenuated, but the magnitude of the effect sizes was consistent with the primary analyses (Table S3 in Multimedia Appendix 1).

Effect Measure Modification by Cognitive Status

The models among participants with CDR ≥0.5 did not converge due to small sample sizes with the addition of usual gait speed as a covariate. Therefore, these models were adjusted for age, sex, race, education, and BMI. Among participants with normal cognition (CDR=0; n=56), the findings were robust for SQI. All other findings were attenuated, although the magnitude of the effect sizes was consistent with the primary analyses. Among participants with cognitive impairment (CDR ≥0.5; n=16), all findings were attenuated. The magnitude of the effect size was similar for sleep fragmentation and mean heart rate (Table S4 in Multimedia Appendix 1).

Cross-Sectional Associations Between Sleep and the Absolute Value of Step Width Variability

After taking the absolute value of step width variability, the mean was 16.63 (SD 12.05) cm, with a range of 0.19 to 57.91 cm. After covariate adjustment, 1% longer REM sleep was associated with 0.81 cm lower step width variability (95% CI −1.32 to −0.30). There were no associations with any other sleep measures (Table S5 in Multimedia Appendix 1).

Discussion

Principal Results

In this clinical sample of middle-aged and older adults, lower objectively measured sleep quality was cross-sectionally associated with higher step width variability during the preferred speed walking condition. Specifically, lower SQI, shorter stable or REM sleep, higher sleep fragmentation, longer apnea duration, and faster mean heart rate were each associated with a higher probability of having lateral (eg, ≥7.5 cm) versus minimal displacement (eg, within ±7.5 cm). The associations between sleep and step width variability were also independent of usual walking pace. Overall, these findings suggest that sleep may be a modifiable risk factor for minimizing step width variability. This has potentially important implications for preventing falls and physical function decline in later life.

This study found that several objectively measured sleep variables related to lower sleep quality were associated with higher step width variability. Most prior studies have focused on associations between sleep and gait speed [14,15,23,24], despite evidence that measures of gait variability may be stronger risk markers for falls [45]. Postural control, which is related to step width variability [2], is impaired among individuals who experience sleep deprivation [46] or sleep-disordered breathing [47]. Additionally, studies have evaluated other measures of gait variability, such as variability in stride length, step length, stride time, or step time. Indeed, lower sleep efficiency, the presence of insomnia, or sleep-related breathing disorders are associated with higher variability in these measures [13,16,17]. This study builds on existing evidence to suggest that sleep quality, even without a sleep disorder, may be an important modifiable risk factor for gait dysfunction, including higher step width variability.

There are several possible mechanisms through which sleep may impact gait and step width variability. Gait is a complex task that requires coordination of several systems [48,49]. Brain regions, including the prefrontal cortex, basal ganglia, brainstem, and cerebellum, are responsible for receiving information about body positioning from the visual, vestibular, and proprioceptive systems. These brain regions must interpret afferent signals and coordinate a motor response through the pyramidal tract, peripheral nerves, and neuromuscular system [48,49]. Further, step width variability is a measure of postural control during locomotion [2], and dynamic postural control is likely mediated through spinal pathways, the cerebellum, the sensorimotor thalamus, the supplementary motor area, and the premotor cortex [50]. Evidence links intermittent hypoxia, which occurs during sleep apnea, with damage to the prefrontal cortex and cerebellum [51], as well as sleep apnea with incident cognitive impairment [52]. Poor sleep quality is also associated with the dysregulation of synaptic homeostasis and the glymphatic system, a higher burden of cerebrovascular disease (eg, white matter hyperintensities) and amyloid-β, and worse cognitive performance [53]. While the role of the basal ganglia in normal sleep mechanisms is less well defined than cortical and brainstem neural circuits, the profound changes associated with disorders of the basal ganglia suggest key roles [54]. These findings may reflect alterations of neural networks from cortical to spinal cord levels, while cortex-to-brainstem areas are more involved in sleep regulation, including the interactions mapped by CPC outputs.

These mechanisms align with prior studies linking poor sleep with gait dysfunction. First, sleep deprivation is associated with poorer postural control among older adults [55]. Sleep deprivation among college students is also related to lower metabolic activity in the prefrontal cortex, basal ganglia, cerebellum, and thalamus, as well as greater difficulty synchronizing gait to an auditory signal [56]. Second, sleep disruption is associated with reduced gait complexity (eg, less adaptability to the environment) [57], indicating greater regularity of stride-to-stride fluctuations [58]. Finally, sleepiness due to poor sleep quality may be related to gait dysfunction. This is supported by several studies linking cumulative exposure to drugs with sedative properties to decreased gait regularity, pace, and complexity [45,59], as well as self-reported sleepiness with a higher risk of falls [51]. Additional research is needed to identify the specific mechanisms through which sleep may impact step width variability.

In this study, associations between sleep quality and step width variability were stronger among men than women. Indeed, several findings were robust for male participants, while only associations with mean heart rate persisted among female participants. However, the magnitude of the effect sizes was similar for both men and women. This aligns with a prior study

XSL•FO
RenderX

that found that slower gait speed was cross-sectionally associated with poor sleep quality and longer sleep duration (eg, >8 h) among men but not women [23]. One potential mechanism explaining the links between sleep and gait among men is that poor sleep quality leads to reductions in muscle mass and strength. This mechanism is supported by a study finding that grip strength mediates associations between self-reported sleep quality and both gait speed and falls among men but not women [14]. Though replication is needed, these preliminary findings suggest that men may be particularly susceptible to higher step width variability related to poor sleep.

Contrary to our hypothesis, the findings for SQI were robust among participants with normal cognition, while no findings were statistically significant among those with cognitive impairment. Given that cognitive decline and dementia are associated with slower gait speed and abnormal gait patterns [4,60], we hypothesized that individuals with cognitive impairment would be more susceptible to gait dysfunction related to poor sleep quality. Populations living with cognitive impairment, such as Alzheimer or Parkinson disease, have higher gait variability [61], lower postural control [61], and a higher risk of falls [62] than cognitively intact controls. As such, we would expect that lower step width variability may be of particular importance for postural control and balance during walking among those with cognitive impairment. Additionally, slower gait speed is linked with lower global cognition among individuals with sleep apnea [24] and lower executive function or attention [49]. The combination of slow gait and poor sleep duration is also associated with higher odds of cognitive impairment [63]. There are a few possible explanations for why we did not observe associations between sleep and step width variability among participants with cognitive impairment in our study. First, we used a global measure of cognitive impairment, but specific cognitive domains, particularly attention and executive function, are linked to both gait and sleep [49,64,65]. Sleep may be more strongly related to step width variability among those with impairment in these domains. Second, the system coordinating gait may have already been disrupted among those with cognitive impairment, potentially attenuating the added harm from poor sleep quality. Finally, we may have been underpowered to detect associations with global cognitive impairment or within specific cognitive domains due to the modest sample size. The replication of the findings is needed in larger, more diverse cohorts.

This study supports the feasibility of remotely monitoring risk factors using digital technologies in the free-living environment. Integrating digital technologies into research has several advantages, including improved feasibility and accessibility, as well as reduced participant burden [22]. Data from these technologies provide additional information about an individual's free-living behaviors and function. Indeed, studies have criticized self-reported questionnaires for misclassifying sleep quality [18,19]. The digital acquisition of gait characteristics in this study also allowed for more nuanced measures of gait characteristics, including step width variability. While this study is limited by a modest sample size and cross-sectional design, these findings support continued research to identify digital sleep biomarkers that relate to functional outcomes. Future studies are also needed to evaluate whether these digital sleep measures also relate to the risk of falls or physical function decline (eg, grip strength, lower extremity strength, balance) and to test whether intervening in these sleep characteristics improves gait and physical function.

Limitations and Strengths

This study has several important limitations. First, there is a lack of consensus on cut points to define minimal, medial, and lateral displacement. Future studies are needed to validate these cut points for asymmetry in lateral step variability. Second, there was a gap in time between measurement of sleep and step width variability. We attempted to address this by excluding participants with measurements >90 days apart. Third, the BU ADRC is a clinical cohort that may be less representative compared to population- or community-based samples. This sample also had limited racial and ethnic diversity, including variations in skin pigmentation, which can bias optoelectronic sensing (eg, plethysmography sensors). Future studies with larger, more diverse samples are needed to evaluate generalizability and sensor performance. Fourth, the sample size was modest, and some analyses may be underpowered. This study is a first step toward establishing associations between sleep and step width variability. Finally, there is a possibility of residual or unmeasured confounding. A major strength of this study is the use of objectively measured sleep quality and step width variability. The use of digital tools provided nuanced measures of both our exposure and outcome. Machine learning methods were also leveraged for sleep feature selection, reducing the risk of overfitting given the modest sample size [43]. Additionally, data were available on cognitive function and important covariates.

Conclusions

This study found that poor objectively measured sleep quality was associated with higher step width variability during preferred speed walking. Future studies are needed to evaluate whether monitoring, or possibly intervening in, these sleep measures could help to minimize the risk of falls or physical function decline in later life.

Acknowledgments

The authors acknowledge the participants of the Boston University Alzheimer's Disease Research Center for their commitment, without which this study would not have been possible. Generative artificial intelligence tools were not used at any stage in the preparation of this manuscript.

XSL•FO
RenderX

Funding

The Boston University Alzheimer's Disease Research Center was supported by the National Institute on Aging, National Institutes of Health (grant number 1P30AG072978), National Institute of Neurological Disorders and Stroke, National Institutes of Health (grant number U54NS115266), and the National Center for Advancing Translational Sciences, National Institutes of Health (grant number 1UL1TR001430). The authors were supported by the National Institute on Aging, National Institutes of Health (grant numbers 5R01AG083735, 1RF1AG072654, and 5K01AG080119), American Heart Association (grant number 20SFRN35360180), Alzheimer's Association (grant number 23AARF-1019048), Parkinson's Foundation, Neuron23, Denali, and University of Rochester. The sponsor was not involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Data Availability

The datasets generated and analyzed for this study can be shared on reasonable request.

Authors' Contributions

FRM, PHH, YT, MA, JM, AM, RA, RJT, and LS made substantial contributions to the conception or design of the work. FRM, MK, RJT, and LS made substantial contributions to the analysis or interpretation of data for the work. FRM, RJT, and LS made substantial contributions to the drafting of the work and reviewing it critically for important intellectual content.
All authors gave final approval of the version to be published. FRM and LS had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

MA received an honorarium from the Michael J Fox Foundation Inc and research support from Life Molecular Imaging Inc. He receives royalties from Oxford University Press Inc. RA is a scientific advisor to Signant Health and Novo Nordisk. RJT developed the cardiopulmonary coupling technology used for this study and licensed to MyCardio, LLC, by Beth Israel Deaconess Medical Center. RJT received royalties per institutional policy. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Results of secondary and sensitivity analyses.
[DOCX File, 585 KB - aging_v9i1e81630_app1.docx]

References

1. Cohen JA, Verghese J, Zwerling JL. Cognition and gait in older people. Maturitas 2016 Nov;93:73-77. [doi: 10.1016/j.maturitas.2016.05.005] [Medline: 27240713]
2. Hausdorff JM. Gait variability: methods, modeling and meaning. J Neuroeng Rehabil 2005 Jul 20;2(1):19. [doi: 10.1186/1743-0003-2-19] [Medline: 16033650]
3. Hollman JH, McDade EM, Petersen RC. Normative spatiotemporal gait parameters in older adults. Gait Posture 2011 May;34(1):111-118. [doi: 10.1016/j.gaitpost.2011.03.024] [Medline: 21531139]
4. Tian Q, Resnick SM, Studenski SA, Ferrucci L. Lap time variability from a 400-m walk is associated with future mild cognitive impairment and Alzheimer's disease. J Am Med Dir Assoc 2019 Dec;20(12):1535-1539. [doi: 10.1016/j.jamda.2019.04.019] [Medline: 31227471]
5. Brach JS, Perera S, Studenski S, Newman AB. The reliability and validity of measures of gait variability in community-dwelling older adults. Arch Phys Med Rehabil 2008 Dec;89(12):2293-2296. [doi: 10.1016/j.apmr.2008.06.010] [Medline: 19061741]
6. Brach JS, Berlin JE, VanSwearingen JM, Newman AB, Studenski SA. Too much or too little step width variability is associated with a fall history in older persons who walk at or near normal gait speed. J Neuroeng Rehabil 2005 Jul 26;2:21. [doi: 10.1186/1743-0003-2-21] [Medline: 16042812]
7. Skiadopoulos A, Moore EE, Sayles HR, Schmid KK, Stergiou N. Step width variability as a discriminator of age-related gait changes. J Neuroeng Rehabil 2020 Mar 5;17(1):41. [doi: 10.1186/s12984-020-00671-9] [Medline: 32138747]
8. Wang Y, Mei Q, Jiang H, et al. The biomechanical influence of step width on typical locomotor activities: a systematic review. Sports Med Open 2024 Jul 27;10(1):83. [doi: 10.1186/s40798-024-00750-4] [Medline: 39068296]
9. Urbanek JK, Harezlak J, Glynn NW, Harris T, Crainiceanu C, Zipunnikov V. Stride variability measures derived from wrist- and hip-worn accelerometers. Gait Posture 2017 Feb;52:217-223. [doi: 10.1016/j.gaitpost.2016.11.045] [Medline: 27936440]
10. Callisaya ML, Blizzard L, Schmidt MD, et al. Gait, gait variability and the risk of multiple incident falls in older people: a population-based study. Age Ageing 2011 Jul;40(4):481-487. [doi: 10.1093/ageing/afr055] [Medline: 21628390]
11. Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults: a review of the literature. Maturitas 2013 May;75(1):51-61. [doi: 10.1016/j.maturitas.2013.02.009] [Medline: 23523272]

XSL•FO
RenderX

12. Landré B, Fayosse A, Ben Hassen C, et al. Terminal decline in objective and self-reported measures of motor function before death: 10 year follow-up of Whitehall II cohort study. BMJ 2021 Aug 4;374:n1743. [doi: 10.1136/bmj.n1743] [Medline: 34348957]
13. Kirshner D, Spiegelhalder K, Shahar RT, Shochat T, Agmon M. The association between objective measurements of sleep quality and postural control in adults: a systematic review. Sleep Med Rev 2022 Jun;63:101633. [doi: 10.1016/j.smrv.2022.101633] [Medline: 35504085]
14. Zhang L, Liu S, Li Y, Li S, Wu Y. Associations of sleep quality with gait speed and falls in older adults: the mediating effect of muscle strength and the gender difference. Gerontology 2022;68(1):1-7. [doi: 10.1159/000514894] [Medline: 33895741]
15. Holingue C, Owusu JT, Tzuang M, et al. Accelerometer-assessed sleep and decline in physical function in older men. Sleep Health 2024 Feb;10(1):129-136. [doi: 10.1016/j.sleh.2023.11.004] [Medline: 38143154]
16. Lee T, Lee M, Youm C, Noh B, Park H. Association between gait variability and gait-ability decline in elderly women with subthreshold insomnia stage. Int J Environ Res Public Health 2020 Jul 17;17(14):5181. [doi: 10.3390/ijerph17145181] [Medline: 32709118]
17. Celle S, Annweiler C, Camicioli R, Barthélémy JC, Roche F, Beauchet O. Sleep-related breathing disorders and gait variability: a cross-sectional preliminary study. BMC Pulm Med 2014 Aug 23;14:140. [doi: 10.1186/1471-2466-14-140] [Medline: 25150985]
18. Jackson CL, Ward JB, Johnson DA, Sims M, Wilson J, Redline S. Concordance between self-reported and actigraphy-assessed sleep duration among African-American adults: findings from the Jackson Heart Sleep Study. Sleep 2020 Mar 12;43(3):zsz246. [doi: 10.1093/sleep/zsz246] [Medline: 31616945]
19. Jackson CL, Patel SR, Jackson WB II, Lutsey PL, Redline S. Agreement between self-reported and objectively measured sleep duration among White, Black, Hispanic, and Chinese adults in the United States: Multi-Ethnic Study of Atherosclerosis. Sleep 2018 Jun 1;41(6):zsy057. [doi: 10.1093/sleep/zsy057] [Medline: 29701831]
20. Thomas RJ, Mietus JE, Peng CK, Goldberger AL. An electrocardiogram-based technique to assess cardiopulmonary coupling during sleep. Sleep 2005 Sep;28(9):1151-1161. [doi: 10.1093/sleep/28.9.1151] [Medline: 16268385]
21. Thomas RJ, Wood C, Bianchi MT. Cardiopulmonary coupling spectrogram as an ambulatory clinical biomarker of sleep stability and quality in health, sleep apnea, and insomnia. Sleep 2018 Feb 1;41(2):zsx196. [doi: 10.1093/sleep/zsx196] [Medline: 29237080]
22. Anda-Duran ID, Hwang PH, Popp ZT, et al. Matching science to reality: how to deploy a participant-driven digital brain health platform. Front Dement 2023;2:1135451. [doi: 10.3389/frdem.2023.1135451] [Medline: 38706716]
23. Wang L, Zou B. The association between gait speed and sleep problems among Chinese adults aged 50 and greater. Front Neurosci 2022;16:855955. [doi: 10.3389/fnins.2022.855955] [Medline: 35557611]
24. Lee S, Shin C. Interaction of obstructive sleep apnoea and cognitive impairment with slow gait speed in middle-aged and older adults. Age Ageing 2017 Jul 1;46(4):653-659. [doi: 10.1093/ageing/afw228] [Medline: 28057621]
25. Jefferson AL, Wong S, Bolen E, Ozonoff A, Green RC, Stern RA. Cognitive correlates of HVOT performance differ between individuals with mild cognitive impairment and normal controls. Arch Clin Neuropsychol 2006 Aug;21(5):405-412. [doi: 10.1016/j.acn.2006.06.001] [Medline: 16893623]
26. Al Ashry HS, Ni Y, Thomas RJ. Cardiopulmonary sleep spectrograms open a novel window into sleep biology-implications for health and disease. Front Neurosci 2021;15:755464. [doi: 10.3389/fnins.2021.755464] [Medline: 34867165]
27. Yuanjie Z, Yunxiao W, Thomas RJ, Yufen T, Zhifei X. Cardiopulmonary coupling estimated sleep quality and memory in children with obstructive sleep-disordered breathing. Sleep Med 2025 May;129:8-13. [doi: 10.1016/j.sleep.2025.01.024] [Medline: 39970701]
28. Hilmisson H, Lange N, Magnusdottir S. Objective sleep quality and metabolic risk in healthy weight children results from the randomized Childhood Adenotonsillectomy Trial (CHAT). Sleep Breath 2019 Dec;23(4):1197-1208. [doi: 10.1007/s11325-019-01802-w] [Medline: 30798410]
29. Magnusdottir S, Witmans M, Hilmisson H. Sleep quality, sleep apnea, and metabolic health in children treated with adenotonsillectomy. Sleep Breath 2023 Aug;27(4):1527-1540. [doi: 10.1007/s11325-022-02747-3] [Medline: 36434376]
30. Magnusdottir S, Hilmisson H, Thomas RJ. Cardiopulmonary coupling-derived sleep quality is associated with improvements in blood pressure in patients with obstructive sleep apnea at high-cardiovascular risk. J Hypertens 2020 Nov;38(11):2287-2294. [doi: 10.1097/HJH.0000000000002553] [Medline: 32649638]
31. Al Ashry HS, Hilmisson H, Ni Y, Thomas RJ, Investigators A. Automated apnea-hypopnea index from oximetry and spectral analysis of cardiopulmonary coupling. Ann Am Thorac Soc 2021 May;18(5):876-883. [doi: 10.1513/AnnalsATS.202005-510OC] [Medline: 33472017]
32. Mobility Lab white paper. : APDM; 2015 URL: https://www.apdm.com/wp-content/uploads/2015/05/02-Mobility-Lab-Whitepaper.pdf [accessed 2026-05-21]
33. Washabaugh EP, Kalyanaraman T, Adamczyk PG, Claflin ES, Krishnan C. Validity and repeatability of inertial measurement units for measuring gait parameters. Gait Posture 2017 Jun;55:87-93. [doi: 10.1016/j.gaitpost.2017.04.013] [Medline: 28433867]

XSL•FO
RenderX

34. Godinho C, Domingos J, Cunha G, et al. A systematic review of the characteristics and validity of monitoring technologies to assess Parkinson's disease. J Neuroeng Rehabil 2016 Mar 12;13(1):24. [doi: 10.1186/s12984-016-0136-7] [Medline: 26969628]
35. Mancini M, Horak FB. Potential of APDM mobility lab for the monitoring of the progression of Parkinson's disease. Expert Rev Med Devices 2016 May;13(5):455-462. [doi: 10.1586/17434440.2016.1153421] [Medline: 26872510]
36. Smith Fine A, Kaufman M, Goodman J, et al. Wearable sensors detect impaired gait and coordination in LBSL during remote assessments. Ann Clin Transl Neurol 2022 Apr;9(4):468-477. [doi: 10.1002/acn3.51509] [Medline: 35257509]
37. Ilg W, Seemann J, Giese M, et al. Real-life gait assessment in degenerative cerebellar ataxia: toward ecologically valid biomarkers. Neurology 2020 Sep;95(9). [doi: 10.1212/WNL.0000000000010176] [Medline: 32611635]
38. Seemann J, Daghsen L, Cazier M, et al. Digital gait measures capture 1-year progression in early-stage spinocerebellar ataxia type 2. Mov Disord 2024 May;39(5):788-797. [doi: 10.1002/mds.29757] [Medline: 38419144]
39. He M, Qi Z, Shao Y, et al. Quantitative evaluation of gait changes using APDM inertial sensors after the external lumbar drain in patients with idiopathic normal pressure hydrocephalus. Front Neurol 2021;12:635044. [doi: 10.3389/fneur.2021.635044] [Medline: 34305775]
40. Kowalski KL, Boolani A, Christie AD. Sex differences in the impact of state and trait fatigue on gait variability. Hum Mov Sci 2021 Dec;80:102884. [doi: 10.1016/j.humov.2021.102884] [Medline: 34600164]
41. Morris JC, Weintraub S, Chui HC, et al. The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. Alzheimer Dis Assoc Disord 2006;20(4):210-216. [doi: 10.1097/01.wad.0000213865.09806.92] [Medline: 17132964]
42. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993 Nov;43(11):2412-2414. [doi: 10.1212/wnl.43.11.2412-a] [Medline: 8232972]
43. Fonti V, Belitser E. Feature selection using LASSO. : Vrije Universiteit Amsterdam; 2017 URL: https://vu-business-analytics.github.io/internship-office/papers/paper-fonti.pdf [accessed 2026-05-21]
44. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Methodol 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
45. Osman A, Kamkar N, Speechley M, Ali S, Montero-Odasso M. Fall risk-increasing drugs and gait performance in community-dwelling older adults: a systematic review. Ageing Res Rev 2022 May;77:101599. [doi: 10.1016/j.arr.2022.101599] [Medline: 35219903]
46. Castor CGM, Santos TRT, Souza TR, et al. Effect of sleep deprivation on postural control and dynamic stability in healthy young adults. Neurosci Lett 2023 Feb 16;797:137055. [doi: 10.1016/j.neulet.2023.137055] [Medline: 36610588]
47. Degache F, Goy Y, Vat S, Haba Rubio J, Contal O, Heinzer R. Sleep-disordered breathing and daytime postural stability. Thorax 2016 Jun;71(6):543-548. [doi: 10.1136/thoraxjnl-2015-207490] [Medline: 26892395]
48. Pirker W, Katzenschlager R. Gait disorders in adults and the elderly: a clinical guide. Wien Klin Wochenschr 2017 Feb;129(3-4):81-95. [doi: 10.1007/s00508-016-1096-4] [Medline: 27770207]
49. Mirelman A, Shema S, Maidan I, Hausdorff JM. Gait. Handb Clin Neurol 2018;159:119-134. [doi: 10.1016/B978-0-444-63916-5.00007-0] [Medline: 30482309]
50. Takakusaki K. Functional neuroanatomy for posture and gait control. J Mov Disord 2017 Jan;10(1):1-17. [doi: 10.14802/jmd.16062] [Medline: 28122432]
51. Stevens D, Jackson B, Carberry J, et al. The impact of obstructive sleep apnea on balance, gait, and falls risk: a narrative review of the literature. J Gerontol A Biol Sci Med Sci 2020 Nov 13;75(12):2450-2460. [doi: 10.1093/gerona/glaa014] [Medline: 32039438]
52. Bubu OM, Andrade AG, Umasabor-Bubu OQ, et al. Obstructive sleep apnea, cognition and Alzheimer's disease: a systematic review integrating three decades of multidisciplinary research. Sleep Med Rev 2020 Apr;50:101250. [doi: 10.1016/j.smrv.2019.101250] [Medline: 31881487]
53. Gottesman RF, Lutsey PL, Benveniste H, et al. Impact of sleep disorders and disturbed sleep on brain health: a scientific statement from the American Heart Association. Stroke 2024 Mar;55(3):e61-e76. [doi: 10.1161/STR.0000000000000453] [Medline: 38235581]
54. Blandini F, Nappi G, Tassorelli C, Martignoni E. Functional changes of the basal ganglia circuitry in Parkinson's disease. Prog Neurobiol 2000 Sep;62(1):63-88. [doi: 10.1016/s0301-0082(99)00067-2] [Medline: 10821982]
55. Robillard R, Prince F, Filipini D, Carrier J. Aging worsens the effects of sleep deprivation on postural control. PLoS One 2011;6(12):e28731. [doi: 10.1371/journal.pone.0028731] [Medline: 22163330]
56. Umemura GS, Pinho JP, Duysens J, Krebs HI, Forner-Cordero A. Sleep deprivation affects gait control. Sci Rep 2021 Oct 26;11(1):21104. [doi: 10.1038/s41598-021-00705-9] [Medline: 34702960]
57. Manor B, Lipsitz LA. Physiologic complexity and aging: implications for physical function and rehabilitation. Prog Neuropsychopharmacol Biol Psychiatry 2013 Aug 1;45:287-293. [doi: 10.1016/j.pnpbp.2012.08.020] [Medline: 22985940]
58. Oliveira JH, Gomes JS, Santos P, Pezarat-Correia P, Vaz JR. Effect of sleep deprivation on gait complexity. J Sleep Res 2025 Dec;34(6):e14478. [doi: 10.1111/jsr.14478] [Medline: 39904824]
59. Wouters H, Van Campen JP, Kuitert MJ, et al. Anticholinergic and sedative medications and dynamic gait parameters in older patients. Drugs Aging 2021 Dec;38(12):1087-1096. [doi: 10.1007/s40266-021-00902-1] [Medline: 34855162]

XSL•FO
RenderX

60. Rosso AL, Verghese J, Metti AL, et al. Slowing gait and risk for cognitive impairment: the hippocampus as a shared neural substrate. Neurology 2017 Jul 25;89(4):336-342. [doi: 10.1212/WNL.0000000000004153] [Medline: 28659421]
61. Pieruccini-Faria F, Black SE, Masellis M, et al. Gait variability across neurodegenerative and cognitive disorders: results from the Canadian Consortium of Neurodegeneration in Aging (CCNA) and the Gait and Brain Study. Alzheimers Dement 2021 Aug;17(8):1317-1328. [doi: 10.1002/alz.12298] [Medline: 33590967]
62. Minta K, Colombo G, Taylor WR, Schinazi VR. Differences in fall-related characteristics across cognitive disorders. Front Aging Neurosci 2023;15:1171306. [doi: 10.3389/fnagi.2023.1171306] [Medline: 37358956]
63. Yoon E, Bae S, Park H. Gait speed and sleep duration is associated with increased risk of MCI in older community-dwelling adults. Int J Environ Res Public Health 2022 Jun 22;19(13):7625. [doi: 10.3390/ijerph19137625] [Medline: 35805289]
64. Rodrigues T, Shigaeff N. Sleep disorders and attention: a systematic review. Arq Neuropsiquiatr 2022 May;80(5):530-538. [doi: 10.1590/0004-282X-ANP-2021-0182] [Medline: 35476076]
65. Tai XY, Chen C, Manohar S, Husain M. Impact of sleep duration on executive function and brain structure. Commun Biol 2022 Mar 3;5(1):201. [doi: 10.1038/s42003-022-03123-3] [Medline: 35241774]

Abbreviations

ADRC: Alzheimer's Disease Research Center
BU: Boston University
CDR: Clinical Dementia Rating
CPC: cardiopulmonary coupling
EEG: electroencephalogram
FDR: false discovery rate
LASSO: least absolute shrinkage and selection operator
NREM: nonrapid eye movement
RRR: relative risk ratio
SpO_2: oxygen saturation
SQI: sleep quality index
TVV: tidal volume variability

Edited by J Wang; submitted 07.Aug.2025; peer-reviewed by E Mico-Amigo, R Choudhury; revised version received 24.Mar.2026; accepted 11.May.2026; published 16.Jun.2026.

Please cite as:
Marino FR, Kim M, Hwang PH, Tripodis Y, Alosco M, Mez J, McKee A, Au R, Thomas RJ, Shih L
Impact of Sleep Quality on Gait Variability: Pilot Cohort Study
JMIR Aging 2026;9:e81630
URL: https://aging.jmir.org/2026/1/e81630
doi:10.2196/81630

© Francesca R Marino, Minzae Kim, Phillip H Hwang, Yorghos Tripodis, Michael Alosco, Jesse Mez, Ann McKee, Rhoda Au, Robert Joseph Thomas, Ludy Shih. Originally published in JMIR Aging (https://aging.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Feasibility and Preliminary Effects of a Pressure Sensor–Based Exergame on Cognitive and Sensorimotor Frailty in Low-Income Older Adults Receiving Public Community Care Services: Pilot Randomized Controlled Trial

Sunmi Song[1,2*], MA, PhD; Suk-Chan Hahm[3*], PT, PhD; Seyun Chang[4], BS; Jinseung Lee[1,2], BS; Heewon Kang[5], MPH, PhD; Sang Hun Lee[4], BS; Junesun Kim[1,6], PT, PhD

[1]Rehabilitation Science Program, Department of Health Science, Graduate School, Korea University, Seoul, Seoul, Republic of Korea

[2]Transdisciplinary Major in Learning Health Systems, Department of Healthcare Sciences, Graduate School, Korea University, Seoul, Seoul, Republic of Korea

[3]Graduate School of Integrative Medicine and Health, Cha University, Pocheon, Gyeonggi-do, Republic of Korea

[4]Geumcheon-gu, 205 Gasan digital 1-ro, MIDAS Health & Technology Co, Seoul, Seoul, Republic of Korea

[5]Department of Public Health Administration, Daejin University, Pocheon, Gyeonggi-do, Republic of Korea

[6]Department of Physical Therapy/Health and Environmental Science, Undergraduate School, College of Health Science, Korea University, Seoul, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Junesun Kim, PT, PhD
Department of Physical Therapy/Health and Environmental Science, Undergraduate School
College of Health Science
Korea University
665 Hana Health Science Building B
145 Anam-ro, Sungbuk-gu
Seoul, Seoul, 02841
Republic of Korea
Phone: 82 232905689
Email: junokim@korea.ac.kr

Related Article:

This is a corrected version. See correction statement: https://aging.jmir.org/2026/1/e109510

Abstract

Background: Age-related cognitive and sensorimotor declines co-occur and reinforce one another, increasing vulnerability to functional dependence and falls. Exergaming offers integrated cognitive-motor training, but most systems lack features that support personalized care for socially vulnerable, community-dwelling older adults. Low-income older adults receiving Public Community Care services bear a disproportionate burden of cognitive and sensorimotor decline, yet remain underrepresented in digital health research. As part of a broader program to develop a digital health monitoring-to-intervention framework, we evaluated a pressure sensor–based exergame designed to deliver targeted cognitive-motor training within the existing Public Community Care infrastructure.

Objective: This study evaluates the feasibility and preliminary effects of the Brain Step exergaming intervention on cognitive and sensorimotor frailty in low-income older adults living alone who were receiving Public Community Care services through a pilot study (study 1) and a randomized controlled trial (RCT; study 2).

Methods: Study 1 (n=15) was a 12-week single-arm pilot study. Study 2 was a parallel-group pilot RCT conducted in Seoul, South Korea. Thirty-one low-income older adults living alone who were receiving Public Community Care services (mean age 81.5 years; 26/31, 84%, female) were randomized to the intervention (n=16) or control (n=15) group. The 16-week intervention used custom hardware with integrated safety bars and arcade-style games that progressed from single- to dual-task challenges.

XSL•FO
RenderX

Primary outcomes were cognitive function (Korean Mini-Mental State Examination [K-MMSE]), physical function (Berg Balance Scale [BBS] and Timed Up and Go [TUG]), and sarcopenia indicators (appendicular skeletal muscle mass [ASM], Appendicular Skeletal Muscle Mass Index, and grip strength). Secondary outcomes were the Korean Montreal Cognitive Assessment, physical fitness (Figure-8 Walk, 30-second Sit-to-Stand, and 2-Minute Step Test), and fall incidence. Generalized estimating equations with Benjamini-Hochberg false discovery rate (FDR) correction for multiple testing were applied.

Results: In the single-arm pilot study (study 1), a pre-post increase in BBS scores and a reduction in TUG time survived FDR correction among the primary outcomes (both FDR-adjusted *P*=.002), as did all 3 secondary physical fitness outcomes (Figure-8 Walk, FDR-adjusted *P*=.006; 30-second Sit-to-Stand, FDR-adjusted *P*=.01; and 2-Minute Step Test, FDR-adjusted *P*=.02). The K-MMSE Visuospatial subscale showed an uncorrected pre-post improvement (*P*=.02), with differential risk-stratified patterns for Recall (*P*=.04) and Attention (*P*=.006). In the pilot RCT (study 2), a nominally significant group × time interaction was observed for ASM (*P*=.04, FDR-adjusted *P*=.11), with the intervention group maintaining muscle mass while the control group declined. The risk-stratified analyses showed an uncorrected significant 3-way interaction for K-MMSE Recall (*P*=.01). No findings from study 2 survived FDR correction.

Conclusions: These sequential studies demonstrate the feasibility of delivering a culturally adapted exergaming intervention through the existing Public Community Care infrastructure. The pilot study (study 1) showed FDR-corrected pre-post improvements in balance and mobility within an uncontrolled single-arm design. In the pilot RCT (study 2), no primary outcomes survived FDR correction. The observed preliminary signals for muscle mass preservation and, in exploratory risk-stratified analyses, cognitive improvement in the high-risk subgroups should be regarded as hypothesis-generating only. Adequately powered trials with risk-enriched samples and active comparators are needed to evaluate efficacy.

Trial Registration: ClinicalTrials.gov NCT06664229; https://clinicaltrials.gov/study/NCT06664229

(JMIR Aging 2026;9:e93467) doi:10.2196/93467

KEYWORDS

exergaming; older adults; cognitive function; sensorimotor function; frailty; fall prevention; dual-task training; community-based intervention; digital health; health equity

Introduction

Aging brings concurrent declines across cognitive, sensory, postural, and muscular systems that together threaten independent living in older adults [1,2]. Frailty is increasingly understood not as disability or comorbidity but as a biologic syndrome of decreased physiologic reserve and reduced resistance to stressors, arising from cumulative declines across multiple systems and conferring vulnerability to falls, disability, and loss of independence [3]. This study addresses 2 interrelated aspects of this syndrome. Cognitive frailty denotes the co-occurrence of physical frailty and cognitive impairment in the absence of Alzheimer disease or other dementia—a potentially reversible, pre-neurodegenerative state distinct from physiologic brain aging and, thus, a target for early intervention [4]. Its sensorimotor counterpart, here termed sensorimotor frailty, comprises age-related deficits in proprioception, balance, postural control, and skeletal muscle function that increase the risk of falls and functional dependence [3,5]. It is closely linked to cognitive decline, sarcopenia, and increased fall risk [2,6]. Critically, these 2 aspects are interrelated rather than independent. This interdependence provides the rationale for integrated interventions that simultaneously engage cognitive, sensory, and motor systems rather than targeting each domain in isolation.

The relationship between sensorimotor decline and cognitive aging is bidirectional. Executive subdomains such as attention, sensory integration, and motor planning are strongly associated with postural instability and gait dysfunction [7], while age-related loss of proprioception and sarcopenia independently compromise both mobility and cognitive function [8,9]. Sarcopenia—the progressive loss of skeletal muscle mass and strength with aging—further compounds these vulnerabilities, as lower limb muscle mass is a critical determinant of balance confidence and functional independence in older adults, and sarcopenia is associated with increased risks of falls, fractures, cognitive impairment, and mortality [9-11]. Integrated dual-task approaches that simultaneously engage cognitive and sensorimotor systems may therefore be more effective than targeting either domain alone [12,13]. Such approaches may be particularly relevant for older adults experiencing preclinical functional decline—those who remain independently ambulatory but show individualized deterioration below their own functional baseline. Network meta-analyses have confirmed that exercise interventions significantly reduce both the number of fallers (risk ratio 0.83, 95% CI 0.77-0.89) and the rate of falls (risk ratio 0.79, 95% CI 0.73-0.86), with emerging evidence favoring combined physical-cognitive training [13]. In South Korea, where adults aged 65 years and older constitute 19.2% of the population [14], preserving functional independence through early detection and recovery from these preclinical declines represents a pressing public health priority. While the number of Koreans with dementia is projected to exceed 1 million by 2026 [14,15], up to 45% of dementia cases may be attributable to modifiable risk factors, including sensory loss, physical inactivity, sarcopenia, and depression [16], underscoring the value of interventions that target the functional precursors of cognitive and sensorimotor decline before irreversible impairment occurs.

Exergaming—physically active video gaming that requires whole-body movements—has emerged as a promising approach for delivering integrated cognitive-motor training. Recent

XSL•FO
RenderX

randomized controlled trials (RCTs) have demonstrated that exergaming can improve executive function [17,18], dual-task performance [19], and global cognition [18,20] in older adults, with a landmark trial by Sturnieks et al [21] showing that home-based step exergaming and cognitive training reduced fall rates by 26% over 12 months. However, most exergaming research has used commercially available gaming systems—primarily Nintendo Wii, Microsoft Xbox Kinect, or stationary cycling platforms [12,18,19]—which may present significant barriers for older adults with limited technology experience, cognitive impairment, or balance deficits that require postural support during exercise [18].

Despite the growing evidence base for exergaming interventions, critical gaps remain. First, socially vulnerable populations, who bear disproportionate burdens of both falls and cognitive decline, are typically underrepresented in clinical trials, limiting the generalizability of existing evidence [22]. Low-income older adults living alone face multiple barriers to accessing conventional rehabilitation services, including transportation difficulties, financial constraints, and social isolation [23]. Second, most prior RCTs have been conducted in individual home-based settings [12] or laboratory settings [18,19]. While home-based approaches can reduce falls, the social isolation inherent in such formats may limit engagement for some populations. Third, the generational appropriateness of gaming content—identified by Begde et al [12] as a critical determinant of usability and adherence—has not been systematically explored. Most commercially available gaming systems were designed for younger, technology-literate users, and their game interfaces and interaction paradigms may not align with the cultural preferences and prior experiences of today's older adults.

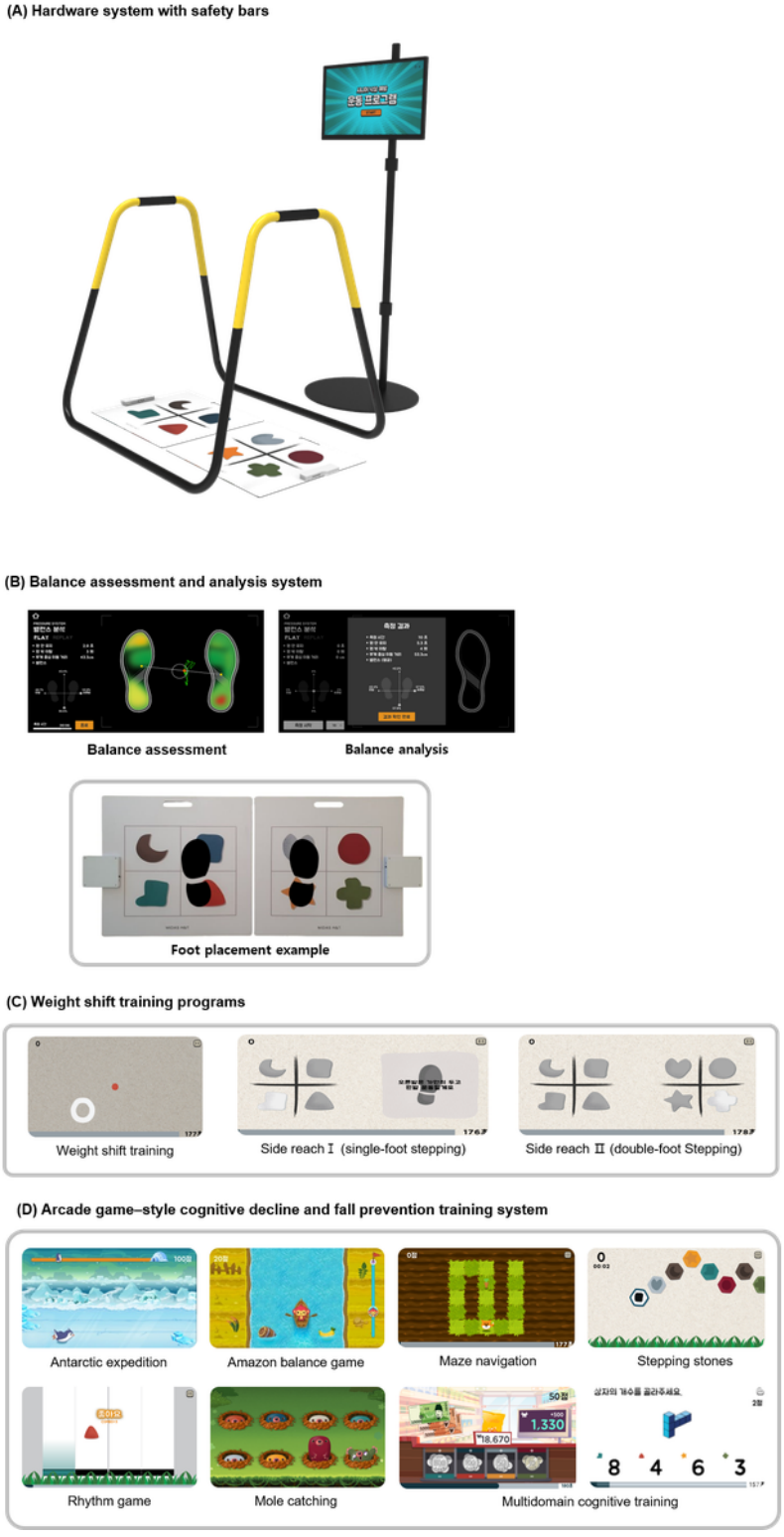
To address these gaps, we developed the Brain Step program, a pressure sensor–based exergaming intervention specifically designed for older adults with elevated cognitive and fall risk (Figure 1). The program features custom hardware with integrated safety bars for postural support, culturally adapted arcade-style games familiar to the target generation, and a progressive structure that advances from single-task to cognitive-motor dual-task challenges [12,18]. The intervention was embedded within the existing service infrastructure of a community welfare center, with on-site staff trained as facilitators to ensure sustainability beyond the research period.

This research is part of a 3-phase program to develop and evaluate a digital health monitoring-to-intervention framework for low-income older adults receiving government-supported Community Care services. Phase 1 established a digital health monitoring platform integrating smartphone chatbots, smartwatch sensors, and motion-sensing technology [24], and phase 2 demonstrated that this platform, combined with caregiver-mediated intervention within the public Community Care infrastructure, could maintain functional status in low-income older adults living alone [25]. This study (phase 3) evaluates Brain Step, a pressure sensor–based exergaming intervention designed to deliver targeted cognitive-motor training when signs of functional decline are detected through the monitoring platform.

By embedding the intervention within the existing Public Community Care service infrastructure, this study aims to address a critical translational gap between digital health monitoring and actionable interventions for populations that face structural barriers to conventional rehabilitation services. This research consisted of 2 sequential studies that used a user-centered, community-participatory approach to develop and evaluate the intervention. Study 1 was a single-arm pilot study designed to evaluate the feasibility, safety, and preliminary efficacy of the Brain Step program and to refine the intervention protocol based on participant and staff feedback. Study 2 was a pilot RCT designed to estimate the preliminary effects of the refined 16-week Brain Step intervention on cognitive and sensorimotor frailty in socially vulnerable older adults receiving Public Community Care services. Secondary objectives were to explore the preliminary effects on self-reported health outcomes (depressive symptoms, sleep quality, and pain-related functional limitations) and to examine whether baseline functional risk status moderated intervention effects. Given the exploratory nature of this pilot RCT, we examined whether the intervention group would show preliminary signals of change in cognitive and sensorimotor frailty compared with the control group and whether individuals with higher baseline risk would show differential intervention responses.

XSL•FO
RenderX

Figure 1. Brain Step system hardware and software components. (A) Hardware system with safety bars. The intervention setup comprises a display screen, dual safety support bars, and a pressure sensor–based response mat with symbol-coded target zones for single- and dual-task training. (B) Balance assessment and analysis system. Real-time visualization of the center of pressure during balance assessment, a postassessment report summarizing balance metrics (duration, sway distance, and weight distribution), and an example of foot placement on the symbol-coded target zones used during training. (C) Weight shift training programs. Exercises targeting lower-limb coordination and weight-shifting control, including Weight Shift Training (directional target tracking), Side Reach I (single-foot stepping), and Side Reach II (double-foot stepping). (D) Arcade game–style cognitive decline and fall prevention training system. Exergames integrating proprioceptive-motor movements with cognitive demands, progressing from motor-focused tasks (Antarctic Expedition, Amazon Balance Game, Maze Navigation, Stepping Stones, Rhythm Game, and Mole Catching) to dual-task cognitive training (multidomain cognitive training targeting orientation, memory, language, visuospatial function, and arithmetic).



XSL•FO
RenderX

Methods

Overview

This research consisted of 3 phases conducted at the Seongbuk Municipal Senior Welfare Center (hereafter, the Welfare Center) in Seoul, South Korea. Study 2 participants were recipients of the Korean government's Public Community Care Service—a nationally standardized, publicly funded program that provides independently ambulatory, low-income older adults living alone with regular caregiver support, including twice-weekly phone calls, weekly home visits, and additional follow-up during episodes of acute illness [26]. Although this study was conducted at a single site, the standardized service protocol—with regulated caregiver work hours (25 hours/week) and caseloads (up to 12 older adults per caregiver)—offers a replicable infrastructure for embedding intervention programs nationwide. A brief prepilot phase tested the initial safety and feasibility of the Brain Step system. Study 1 was a single-arm pilot study designed to evaluate the feasibility and safety of the Brain Step program. Based on the findings of study 1, the intervention protocol was refined and tested in study 2, an RCT. In both studies, participants were predominantly female (38/46, 83%, with 12/15 in the pilot and 26/31 in the RCT), had a mean age of approximately 80 years, and presented with fall risk factors such as a history of falls or gait difficulties. Study 1 recruited community-dwelling older adults from the Welfare Center without income restrictions, whereas study 2 specifically enrolled low-income older adults living alone who were receiving the Public Community Care Service.

Ethics Approval

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board of Korea University (approval number KUIRB-2023-0360). The study was registered at ClinicalTrials.gov (NCT06664229). Written informed consent was obtained from all participants before enrollment.

Pre-Pilot Testing

Before the pilot study, a brief prepilot phase was conducted with 4 older adults to assess the initial safety and feasibility of the Brain Step system. Participants attended sessions 3 times per week for 1 week. This preliminary testing confirmed that the system was safe for use by older adults and that the basic exercises could be completed without adverse events, informing the design of the subsequent pilot study. The overall development approach—progressing from needs assessment through pilot testing to RCT evaluation—was guided by a participatory, user-centered process that engaged older adult participants, community caregivers, and welfare center administrators through qualitative interviews and open-ended feedback collected at posttest. Their input informed device and program design decisions, including the adoption of safety features (integrated support bars), difficulty-level adjustments, and gradual progression from sensorimotor to more complex dual-task training, ensuring that the intervention was contextually appropriate for this population.

Study 1: Pilot Study

Design and Participants

Study 1 was a single-arm pre-post study conducted from March to May 2024. Community-dwelling older adults aged 65 years or older were recruited through bulletin board postings at the community senior welfare center. Inclusion criteria were (1) the ability to communicate and follow instructions and (2) fear of falling. Exclusion criteria were (1) severe sensory impairments (hearing or vision), (2) cognitive or psychiatric impairments (eg, hallucinations), and (3) inability to ambulate independently.

Of 20 older adults screened, 18 met the eligibility criteria and agreed to participate. Three participants withdrew (2 before baseline because of sequelae of back surgery and 1 during the intervention because of scheduling conflicts), resulting in 15 participants (mean age 80.20 years, SD 6.88 years; 12/15, 80%, were women) who completed the study. The characteristics of the pilot study participants are reported in Table S1 in Multimedia Appendix 1.

To examine differential effects according to baseline cognitive function and sensorimotor risk levels, staff at the community senior welfare center assisted with recruitment to ensure balanced enrollment of participants across high- and low-risk categories.

Intervention

Participants received the Brain Step program for 12 weeks, attending 50-minute sessions twice weekly. The program consisted of pressure sensor–based exercises, including center of gravity training, single- and double-foot stepping, and exergames (maze navigation, stepping stones, and mole catching). Sessions were conducted in small groups (≤8 participants per session because of space constraints at the Welfare Center) with trained facilitators, including 2 nonclinical welfare center staff members who served as session facilitators from this pilot phase onward, providing guidance and safety monitoring.

Outcomes and Analysis

Primary outcomes included cognitive function (Korean Mini-Mental State Examination [K-MMSE]), physical function (Berg Balance Scale [BBS] and Timed Up and Go [TUG]), and sarcopenia-related measures (appendicular skeletal muscle mass [ASM], Appendicular Skeletal Muscle Mass Index [ASMI], and handgrip strength).

Secondary outcomes included physical fitness measures (30-second Sit-to-Stand, 2-Minute Step Test, and Figure-8 Walk) and device and program usability. Participants also completed qualitative interviews and satisfaction surveys. Pre-post changes were analyzed using generalized estimating equations (GEEs). Participant feedback was collected through focus group interviews with participants and senior welfare center staff.

XSL•FO
RenderX

Study 2: RCT

Design

Study 2 was a parallel-group RCT. Recruitment began in August 2024, the intervention was delivered from November 2024 to March 2025, and the final assessment was completed in April 2025. Participants were assessed at 3 time points: baseline (preintervention), 8 weeks (mid-intervention), and 16 weeks (postintervention).

Participants

Study 2 targeted low-income older adults living alone who were receiving the Customized Community Care Service described above. Participants were identified through the Community Care service database at the Welfare Center. Community Care providers facilitated recruitment by identifying eligible older adults from their caseloads, arranging initial contact with the research team, and coordinating transportation to the Welfare Center for intervention sessions. Through this involvement, Community Care providers also learned the intervention procedures and progressively assumed facilitator roles, further supporting the program's sustainability within routine welfare center operations. The inclusion and exclusion criteria were identical to those of study 1.

Sample Size

Sample size was calculated using G*Power version 3.1.9.7. Based on a repeated-measures analysis of variance with a within-between interaction, a medium effect size (Cohen *f*=0.25) [27] was assumed. This assumption was based on meta-analytic evidence from exergaming interventions reporting pooled effect sizes of standardized mean difference of 0.35 for executive function [28] and standardized mean difference of 0.36-0.78 for balance outcomes in older adults [29], consistent with prior meta-analytic evidence [17]. Assuming an α level of .05, statistical power of 0.80, 2 groups, and 3 measurement time points, a minimum of 28 participants was required. Accounting for a 15% (5/33) dropout rate, the target sample size was 33 participants. To ensure balance between groups, 34 participants were planned for recruitment.

Randomization and Blinding

The random allocation sequence was generated using block randomization with a fixed block size of 4 by an independent researcher who was not involved in participant recruitment, intervention delivery, or outcome assessment. The allocation sequence was concealed from the other investigators until the time of assignment. Owing to the nature of the intervention, participants were not blinded and were aware of their group assignment, whereas outcome assessors were blinded to group allocation to minimize detection bias.

Intervention

As shown in Figure 1, the Brain Step program consists of a large pressure sensor mat (1100 mm × 1100 mm) with 4 quadrants and 8 response zones, a display screen, and integrated safety bars.

Participants attended 50-minute sessions twice weekly for 16 weeks (32 sessions total). Sessions were conducted in small groups of up to 8 participants (limited by the available space and time at the Welfare Center), with morning and afternoon sessions available. Each session was led by 1 trained facilitator, with 1-2 research assistants providing individual guidance and safety monitoring. Instructional videos were provided to explain the program objectives and demonstrate each game task. Facilitators received training in device operation, safety monitoring, emergency procedures, and participant guidance before program initiation. To prepare for technology-assisted translation into routine community care practice, 2 welfare center staff members—who were not specialists in cognitive or physical interventions—were intentionally included as facilitator trainees, reflecting the typical workforce of community-based older adult care settings. Their operational needs were elicited through qualitative interviews and incorporated into the program design, including adjustable difficulty levels, automated game modules, and instructional videos demonstrating each task to support independent program delivery. Intervention fidelity was monitored through session checklists completed by research assistants and weekly supervision meetings. Adherence was tracked through individual session attendance records and game completion rates documented automatically by the Brain Step system.

As shown in Table 1, the main trial program was structured into 3 progressive phases. Phase 1 (weeks 1-4) focused on single-task training, including weight-shift exercises with real-time center of pressure feedback and side-reaching exercises (single- and double-foot stepping) to strengthen ankle mobility and lower limb muscle function. Phase 2 (weeks 5 and 6) introduced measurement system level 2, combining physical exercises with location memory tasks and learning system games to increase task complexity. Phase 3 (weeks 7-16) emphasized dual-task training through multidomain cognitive training targeting image association, arithmetic, and recognition. Participants selected answers by stepping on designated target zones among the 8 positions on the pressure sensor mat, thereby integrating cognitive decision-making with weight-shifting movements.

The software comprised 3 components (Figure 1): a balance assessment system (Figure 1B) providing real-time center-of-pressure visualization, a weight-shift training system (Figure 1C), and a cognitive-motor training system (Figure 1D) featuring 7 exergames inspired by classic arcade games from the 1980s and 1990s to enhance engagement through nostalgic familiarity. Detailed descriptions of each game are provided in the Methods S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Table 1. Table1. Brain step program structure and progression.

Phase	Weeks	Training focus	Tasks	Task type
1	1-4	Ankle mobility and lower limb strengthening	Center of gravity training and side reach (single/double foot stepping)	Single task
2	5 and 6	Memory and spatial cognition	Maze navigation, stepping stones, and mole catching	Increased complexity
3	7-16	Cognitive-motor integration	Cognitive training quizzes (image association, arithmetic, and recognition) with proprioceptive-motor movements	Dual task

Control Group

Participants in the control group received no intervention during the 16-week study period and attended only the 3 assessment sessions (baseline, 8 weeks, and 16 weeks). A passive (no-contact) control was selected at this feasibility stage to establish whether the intervention could be delivered and whether participants would engage with it in the real-world Public Community Care setting. As the intervention was delivered on-site within the existing infrastructure of the community welfare center, an active or attention-matched comparator was not logistically feasible at this stage. Limited space availability and a shortage of on-site staff precluded implementation of a parallel active-control program. Our priority was to evaluate the real-world applicability of the intervention.

Risk Stratification

Participants were stratified into high-risk (n=17) and low-risk (n=14) groups based on initial screening assessments. High-risk classification required the presence of 2 or more of the following: a history of recurrent falls, gait abnormalities, or cognitive test scores indicating mild cognitive impairment (K-MMSE ≤21 and Montreal Cognitive Assessment [MoCA] <23). This stratification was prespecified during the study design phase. Community stakeholders requested the inclusion of participants at the highest cognitive and sensorimotor risk, and the pilot study findings of differential response patterns informed the decision to classify participants by risk level using pretest screening (K-MMSE and fall history).

Outcome Measures

Comprehensive health outcomes were assessed to provide individualized feedback within the digital-human integrated personalized care framework. Outcomes were organized into 3 tiers: primary outcomes for which intervention effects were hypothesized, secondary outcomes to capture more detailed functional changes, and self-reported health outcomes routinely monitored in Community Care services. Primary outcomes included cognitive function (K-MMSE total and subscale scores), physical function (BBS and TUG), and sarcopenia indicators (ASM, ASMI, and grip strength). The K-MMSE [30] evaluates orientation, recall, attention/calculation, language, and visuospatial function (total score 0-30). Balance was measured using the Korean version of the BBS (14 items; total score 0-56) [31,32]. The TUG measured the time (in seconds) required to rise from a chair, walk 3 m, turn, return, and sit down [33]. Sarcopenia-related measures were assessed according to the Asian Working Group for Sarcopenia 2019 criteria [34]. ASM was measured using bioelectrical impedance analysis (InBody 370), and ASMI was calculated as ASM/height^2 (kg/m^2). Grip strength was measured using a dynamometer (CAMRY EH101) [35,36], with the highest value from 3 trials of the dominant hand used for analysis.

Secondary outcomes were complementary, hypothesis-generating measures, including cognitive function (Korean Montreal Cognitive Assessment [MoCA-K] total and subscale scores), physical fitness (Figure-8 Walk Test, 30-second Sit-to-Stand, and 2-Minute Step Test [36,37]), and fall incidence. The MoCA-K [38,39] was included as a secondary cognitive measure because the K-MMSE has known ceiling effects in detecting subtle cognitive changes in individuals with mild cognitive impairment [40]. The MoCA-K provides greater sensitivity for mild cognitive impairment through more demanding assessments of executive function, visuospatial ability, and attention (total score 0-30) [39].

Fall incidence was assessed at 3 time points. At baseline, participants completed a screening question (Have you experienced a fall in the past year?). At the immediate postintervention assessment, they were asked, “Have you experienced a fall during the past 16 weeks?” At the 1-year postintervention follow-up, fall incidence was determined from community caregiver fall incident reports shared by the municipal welfare center as part of routine Public Community Care service monitoring. Fall incidence was included as a secondary outcome to complement the primary balance and mobility measures.

Other outcomes included sleep quality (Pittsburgh Sleep Quality Index [PSQI]; total score 0-21, with higher scores indicating poorer sleep quality) [41], depressive symptoms (Geriatric Depression Scale-Short Form [GDS-SF]) [42], and pain-related functional limitations (Western Ontario and McMaster Universities Osteoarthritis Index [WOMAC]) [43].

Game Performance Metrics

Game performance data were automatically recorded throughout the 16-week intervention for all participants in the intervention group (n=16). A total of 63 performance indicators were generated across 9 games, including play count, average score, play time, reaction time, and game-specific metrics.

Statistical Analysis

Baseline characteristics were compared between groups to assess randomization balance using *t* tests (unpaired and 2-tailed) for continuous variables and chi-square or Fisher exact tests for categorical variables. The primary analysis examined intervention effects on primary, secondary, and other health

XSL•FO
RenderX

outcomes using GEEs to account for the correlation among repeated measures within participants. GEE models specified an exchangeable working correlation structure with robust SEs. All randomized participants were analyzed according to their assigned groups (intention-to-treat). Missing observations were handled at the observation level using available-case analysis without imputation, with all 31 randomized participants retained in their originally assigned groups and each contributing all available time points to the estimation [44-46]. For grip strength, 1 participant could not be assessed because of hand surgery unrelated to the intervention and was therefore excluded from the grip strength model (n=30). All other outcomes included all 31 randomized participants (see the "Data Exclusion and Missing Data" section).

The full model included group (intervention vs control), time (baseline, 8 weeks, and 16 weeks), and the group × time interaction to examine whether intervention effects were significant. Risk-stratified analyses were also conducted, as prespecified during the study design phase. Years of education were included as a covariate in the models for cognitive function to account for potential confounding. Given the number of outcome variables tested, the Benjamini-Hochberg false discovery rate (FDR) correction was applied within the 6 primary outcomes (K-MMSE total, BBS, TUG, ASM, ASMI, and grip strength; 6 comparisons) at q=0.05. Secondary outcomes (MoCA-K total, Figure-8 Walk, 30-second Sit-to-Stand, 2-Minute Step Test, and fall incidence) and other outcomes (PSQI, GDS-SF, and WOMAC) were reported with unadjusted P values only. Subscale scores were reported descriptively without FDR correction. These analyses were distinguished by both outcome tier and inferential status. Confirmatory inference was restricted to the 6 prespecified primary outcomes with FDR correction. The secondary and other outcomes, the K-MMSE subscale analyses, and the prespecified risk-stratified subgroup analyses were exploratory and hypothesis-generating and were reported with unadjusted P values without confirmatory interpretation. The only post hoc, data-driven analyses were the exploratory correlations between game performance metrics and clinical outcomes (Figure S1 in Multimedia Appendix 1).

An exploratory analysis examined the associations between game performance metrics and significant changes in outcome measures in study 2, with P<.05 applied as the exploratory threshold for identifying significant outcome changes in the intervention group. Pearson correlation coefficients were calculated between game performance metrics and nominally significant changes in primary outcomes (pre-to-post and mid-to-post; for ASM, pre-to-post change, P=.04). For K-MMSE Recall (P=.01) and ASMI (P=.05), analyses were conducted in the high-risk subgroup (n=8) based on significant group × risk × time interactions. For ASM, analyses were conducted in the full intervention group (n=14).

Statistical significance was set at P<.05 (2-tailed). The FDR correction described above was applied to account for multiple outcome comparisons. The study was not powered to detect intervention effects for all outcomes or subgroup effects. All risk-stratified and correlation analyses were exploratory and hypothesis-generating. Effect sizes were calculated as Cohen d based on the between-group differences in pre-to-post change scores divided by the pooled SD of the change scores, with 95% CIs [27]. Analyses were performed using Python 3.12 (Python Foundation) with the statsmodels package for GEE models and SciPy for correlation analyses. Additional statistical details, including the analysis of fall incidence outcomes, are provided in Methods S1 in Multimedia Appendix 1.

Data Exclusion and Missing Data

Of the 31 randomized participants, all completed the baseline assessment except for 1 intervention participant who declined the grip strength measurement because of a history of bilateral hand surgery. This participant subsequently withdrew before the 8-week assessment. By the 8-week midpoint, 5 participants had withdrawn (2 intervention and 3 control), leaving 26 participants with data at that time point. By the 16-week postintervention assessment, 1 additional control participant had withdrawn because of scheduling conflicts, leaving 25 participants with data at that time point. The full longitudinal data set thus comprised 82 participant-by-time point observations across the 3 assessments.

Missing data resulted primarily from participant attrition rather than item-level nonresponse and were not imputed. All GEE analyses used available-case analysis at the observation level. Each participant-by-time point record was included when the outcome value was nonmissing, whereas records missing key covariates (group assignment or baseline risk stratum) were excluded. This approach leverages GEE's ability to accommodate unbalanced repeated-measures data without listwise deletion at the participant level [45,46], such that all 31 randomized participants were retained in their originally assigned groups, with each contributing all available observations to the estimation.

Results

Study 1: Pilot Study Results

The pilot study results are presented in Table 2. Among the primary outcomes, significant pre-post changes that survived FDR correction were observed for the BBS (P<.001, FDR-adjusted P=.002) and TUG (P<.001, FDR-adjusted P=.002). ASM (P=.04) and ASMI (P=.047) showed modest decreases but did not survive FDR correction (FDR-adjusted P=.07). The K-MMSE Visuospatial subscale showed a significant uncorrected pre-post change (P=.02). Among the secondary outcomes (Table S2 in Multimedia Appendix 1), all 3 measures survived FDR correction: Figure-8 Walk (P=.002, FDR-adjusted P=.006), 30-second Sit-to-Stand (P=.007, FDR-adjusted P=.01), and 2-Minute Step Test (P=.02, FDR-adjusted P=.02).

Risk-stratified analyses revealed nominally significant time × risk interactions for the K-MMSE Memory (P=.04) and Attention (P=.006) subscales, with differential patterns according to risk group (Figure 2).

Based on the pilot findings and participant feedback, the intervention protocol was refined for study 2. Instructional videos were added, real-time feedback was enhanced, the

XSL•FO
RenderX

program duration was extended from 12 to 16 weeks, and a midpoint assessment was added at 8 weeks.

Table 2. Comparison of cognitive and sensorimotor frailty over time for pilot testing.[a]

Variable	Intervention group (n=15)		*P* value	FDR[b] *P* value
	Pre	Post		
Cognitive frailty, mean (SD)				
Korean Mini-Mental State Examination, total [c]	24.47 (3.23)	24.53 (2.85)	.83	.83
Orientation	9.47 (0.99)	9.40 (1.12)	.74	N/A[d]
Recall	5.47 (0.99)	5.47 (0.64)	>.99	N/A
Attention	4.07 (1.44)	4.00 (1.41)	.65	N/A
Language	7.87 (0.35)	7.80 (0.41)	.30	N/A
Visuospatial	0.60 (0.51)	0.87 (0.35)	.02[e]	N/A
Sensorimotor frailty, mean (SD)				
Berg Balance Scale (score)	51.10 (3.92)	53.20 (2.70)	<.001[e]	.002[e]
Timed Up and Go (seconds)	8.45 (2.47)	7.41 (2.27)	<.001[e]	.002[e]
Appendicular skeletal muscle mass (kg)	14.65 (2.18)	14.31 (2.44)	.04[e]	.07
Appendicular Skeletal Muscle Mass Index (kg/m^2)	6.14 (0.74)	5.99 (0.74)	.047[e]	.07
Grip strength (kg)	22.01 (4.58)	22.95 (5.36)	.09	.11

[a]Pre-post changes were analyzed using generalized estimating equations with an exchangeable correlation structure. The Benjamini-Hochberg FDR correction (q=0.05) was applied within the 6 primary outcomes.

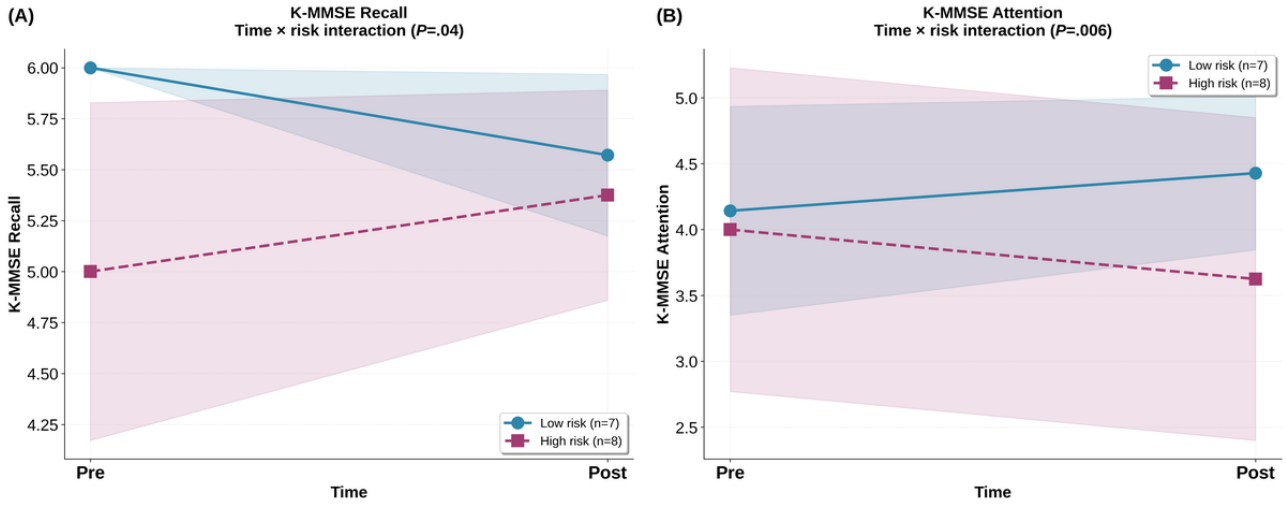
[b]FDR: false discovery rate.

[c]Analyses with Korean Mini-Mental State Examination outcomes were adjusted for years of education.

[d]N/A: not applicable.

[e]P<.05.

Figure 2. Risk-stratified effects on K-MMSE subscale scores in study 1 (n=15). (A) Memory: significant time × risk interaction (P=.04), with improvement observed only in the high-risk group. (B) Attention: significant time × risk interaction (P=.006), with improvement observed only in the low-risk group. Shaded areas indicate 95% CIs around the estimated means. K-MMSE: Korean Mini-Mental State Examination.



Study 2: The Main RCT Results

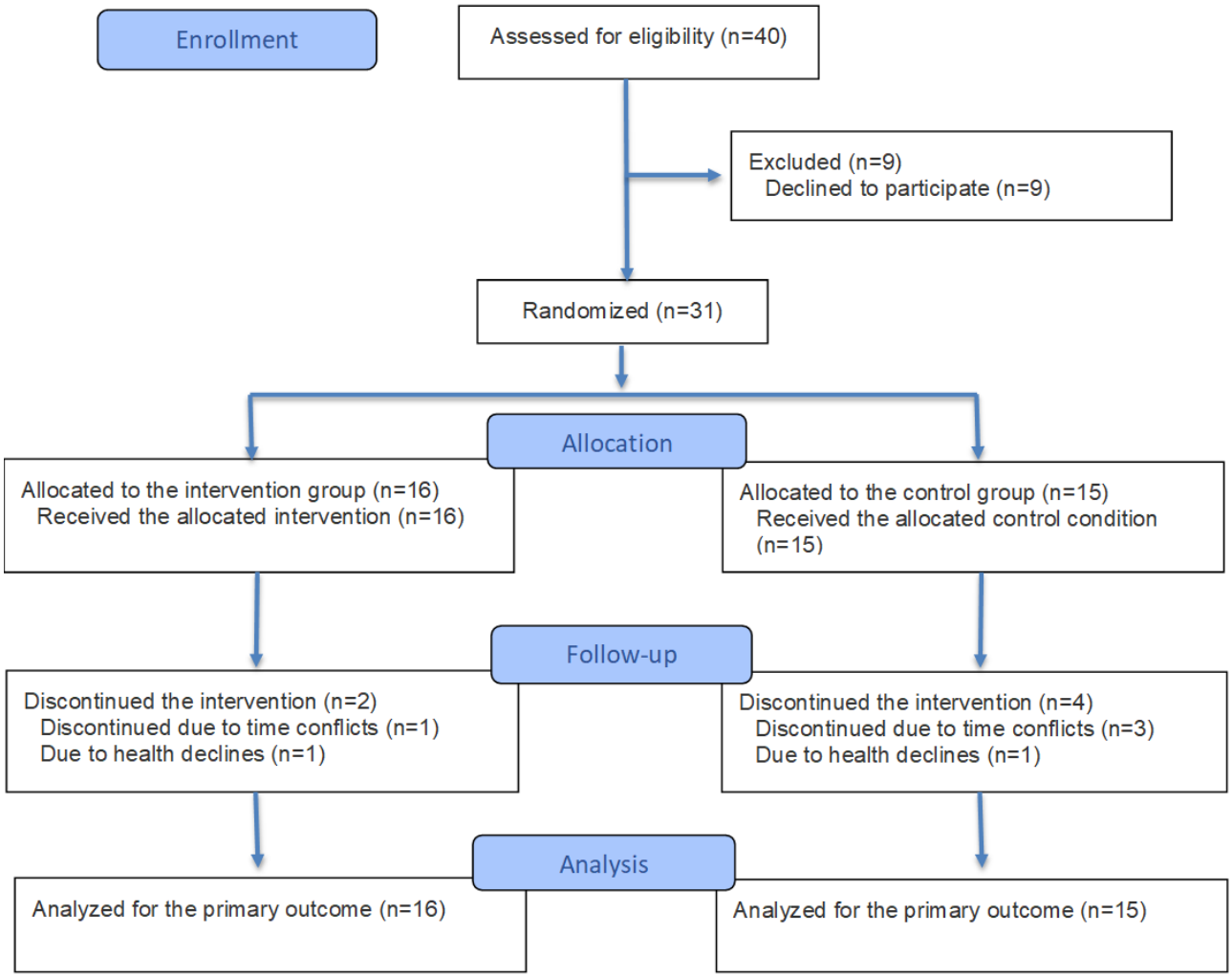
Participant Characteristics

Participant flow is presented in Figure 3. Of the 40 older adults screened, 31 were randomized to the intervention (n=16) or control (n=15) group. Six participants withdrew during the study. All 31 randomized participants were included in the intention-to-treat analysis (82 total observations across 3 time points). See Multimedia Appendix 2 for the study's CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist.

XSL•FO
RenderX

Baseline characteristics are presented in Table 3. Participants (mean age 81.5 years; 26/31, 84%, female) were low-income older adults living alone who were receiving Public Community Care services. The groups were comparable at baseline (see Table 3 for *P* values).

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) flowchart.



XSL•FO
RenderX

Table 3. Baseline characteristics of older adult participants by groups for the randomized controlled trial.

Variable	Values		*P* values
	Intervention (n=16)	Control (n=15)	
Demographic variables			
Age (years), mean (SD)	82.06 (5.42)	80.93 (4.92)	.55
Sex (women), n (%)	13 (81)	13 (87)	>.99
Years of education, mean (SD)	5.88 (4.86)	4.87 (3.23)	.86
Working (yes), n (%)	7 (44)	8 (53)	.86
Monthly income (KRW[a]), n (%)			.23
Less than 500,000	5 (31)	2 (13)	
500,001-1,000,000	10 (63)	11 (73)	
1,000,000 or more	1 (6)	2 (13)	
Charlson Comorbidity Index, n (%)			.47
0	9 (56)	8 (53)	
1-2	7 (44)	5 (33)	
3 or more	0 (0)	2 (13)	
Cognitive functions, mean (SD)			
Montreal Cognitive Assessment	18.81 (4.18)	18.93 (4.08)	.56
Korean Mini-Mental State Examination	24.44 (2.90)	25.73 (1.87)	.28
Sensorimotor function and body composition, mean (SD)			
BMI (kg/m^2)	26.54 (4.01)	24.55 (3.76)	.16
Berg Balance Scale	52.88 (1.76)	52.73 (2.46)	.92
Appendicular skeletal muscle (kg)	15.35 (4.48)	17.03 (6.00)	.28
Appendicular Skeletal Muscle Index (kg/m^2)	6.41 (1.22)	6.96 (2.40)	.71
Grip strength	21.16 (6.88)	20.79 (4.61)	.86
Fall incidence at baseline, n (%)			
Falls in the past year (yes)	5 (31)	6 (40)	.61
Other health outcomes, mean (SD)			
Pittsburg Sleep Quality Index	8.12 (3.90)	9.47 (4.85)	.40
Geriatric Depression Scale-Short Form	3.56 (3.93)	3.33 (2.92)	.86
Western Ontario and McMaster Universities Osteoarthritis Index total	40.81 (15.36)	48.93 (17.73)	.22

[a]US $1=1344 Korean won (KRW).

Intervention Effects on Primary Outcomes

As shown in Table 4, there were no significant group × time interactions for cognitive function. The K-MMSE total score showed no significant interaction (*P*=.27, FDR-adjusted *P*=.38). No significant interactions were observed for the K-MMSE subscales (Orientation, *P*=.36; Recall, *P*=.09; Attention, *P*=.37; Language, *P*=.27; and Visuospatial, *P*=.12).

A nominally significant group × time interaction was observed for the sarcopenia-related index ASM (*P*=.04; Cohen *d*=0.39, 95% CI −0.41 to 1.18). The intervention group maintained ASM over the 16-week period (baseline: 15.35 kg; postintervention: 15.77 kg), whereas the control group showed a decline (baseline: 17.03 kg; postintervention: 15.48 kg; Table 4). The group × time interaction for ASMI did not reach statistical significance (*P*=.05; *d*=0.38, 95% CI −0.41 to 1.18). This pattern is consistent with possible preservation of muscle mass, although the finding did not survive FDR correction. The BBS showed a marginally significant group × time interaction (*P*=.05; *d*=−0.27), likely reflecting a ceiling effect given the high baseline scores in both groups (intervention: 52.88; control: 52.73 of a maximum score of 56). No significant group × time interactions were observed for the other sensorimotor frailty measures (Timed Up and Go, *P*=.65; for grip strength, *P*=.32).

After FDR correction, none of the group × time interactions remained statistically significant (all FDR-adjusted *P*>.16; see Table 4). However, effect sizes are presented to provide context for the observed differences. Before correction, the nominally

XSL•FO
RenderX

significant findings showed small-to-medium effect sizes (ASM *d*=0.39; ASMI *d*=0.38), and the risk-stratified analyses revealed medium-to-large effect sizes among high-risk participants for K-MMSE Recall (*d*=0.73) and ASMI (*d*=0.90), warranting confirmation in larger trials with risk-enriched samples.

Table 4. Comparison of cognitive and sensorimotor frailty outcomes between intervention and control groups over time (n=31).

Variable	Intervention (n=16)		Control (n=15)		Group × time[a]		*d* [b]
	Pre, mean (SD)	Post, mean (SD)	Pre, mean (SD)	Post, mean (SD)	*P* value	FDR[c,d] *P* value	
Cognitive frailty							
Korean Mini-Mental State Examination[e], total	24.44 (2.90)	22.00 (3.01)	25.73 (1.87)	23.27 (2.24)	.27	.38	–0.17
Orientation	9.12 (0.89)	9.36 (1.08)	9.13 (1.30)	9.73 (0.47)	.36	N/A[f]	–0.63
Recall	4.88 (1.09)	4.86 (0.95)	5.00 (0.93)	5.00 (1.10)	.09	N/A	–0.29
Attention	2.38 (1.75)	2.64 (1.69)	3.13 (1.36)	3.00 (1.55)	.37	N/A	0.30
Language	7.56 (0.51)	7.43 (0.76)	7.60 (0.63)	7.73 (0.47)	.27	N/A	–0.08
Visuospatial	0.50 (0.52)	0.71 (0.47)	0.87 (0.35)	0.82 (0.40)	.12	N/A	0.69
Sensorimotor frailty							
Berg Balance Scale	52.88 (1.76)	53.82 (1.15)	52.73 (2.46)	53.82 (2.21)	.05	.11	–0.27
Timed Up and Go (seconds)	7.67 (1.80)	6.86 (0.82)	7.97 (1.99)	7.54 (1.34)	.65	.65	–0.03
Appendicular skeletal muscle mass (kg)	15.35 (4.48)	15.77 (4.56)	17.03 (6.00)	15.48 (3.06)	.04[g]	.11	0.39
Appendicular Skeletal Muscle Mass Index (kg/m^2)	6.41 (1.22)	6.56 (1.16)	6.96 (2.40)	6.36 (1.04)	.05	.11	0.38
Grip strength (kg)	21.16 (6.88)	23.27 (5.37)	20.79 (4.61)	19.71 (3.73)	.32	.38	0.50

[a]Group × time interaction effects were analyzed using generalized estimating equations with exchangeable correlation structure.

[b]Cohen *d* was calculated from between-group differences in change scores divided by pooled SD.

[c]FDR: false discovery rate.

[d]The Benjamini-Hochberg FDR correction (*q*=0.05) was applied within the 6 primary outcomes (Korean Mini-Mental State Examination total, Berg Balance Scale, Timed Up and Go, appendicular skeletal muscle mass, Appendicular Skeletal Muscle Mass Index, and grip strength). No outcomes reached significance after FDR correction.

[e]The Korean Mini-Mental State Examination models were adjusted for years of education.

[f]N/A: not applicable.

[g]*P*<.05.

Intervention Effects on Secondary and Other Health Outcomes

As a secondary outcome, self-reported fall incidence was examined at 3 time points (Table S5 in Multimedia Appendix 1). The group × time interaction for pre-post fall incidence was not statistically significant (Wald χ^2_1=0.66, *P*=.42; odds ratio [OR] 0.47, 95% CI 0.07-2.94). At the postintervention assessment, fall incidence was numerically lower in the intervention group than in the control group (5/14, 35.7% vs 7/11, 63.6%; OR 0.32, 95% CI 0.06-1.64). Although this between-group difference did not reach statistical significance (Fisher exact *P*=.24), the direction and magnitude of the effect favored the intervention group. As a secondary outcome, this finding should be interpreted as hypothesis-generating only.

A 1-year postintervention follow-up assessed fall incidence through community caregiver reports collected as part of routine Public Community Care service monitoring. No fall events were reported for any participant in either the intervention or control group during the 12-month follow-up period (intervention, 0/14; control, 0/14). A separate GEE model comparing baseline with the 12-month follow-up also showed no significant group × time interaction (Wald χ^2_1=0.13, *P*=.71; OR 1.39, 95% CI 0.24-8.05), although this model should be interpreted with caution because of quasi-complete separation resulting from the absence of events at follow-up. Although the absence of falls during follow-up is notable for this population with elevated fall risk, the floor effect precluded between-group comparisons, and the absence of a nonparticipating comparison group limited causal interpretation.

No significant group × time interactions were observed for the secondary outcomes (MoCA-K, *P*=.12; Figure-8 Walk, *P*=.65; 30-second Sit-to-Stand, *P*=.72; and 2-Minute Step Test, *P*=.62; Tables S3 and S4 in Multimedia Appendix 1) or the self-reported health outcomes (PSQI, *P*=.19; GDS-SF, *P*=.55; and WOMAC, *P*=.68).

XSL•FO
RenderX

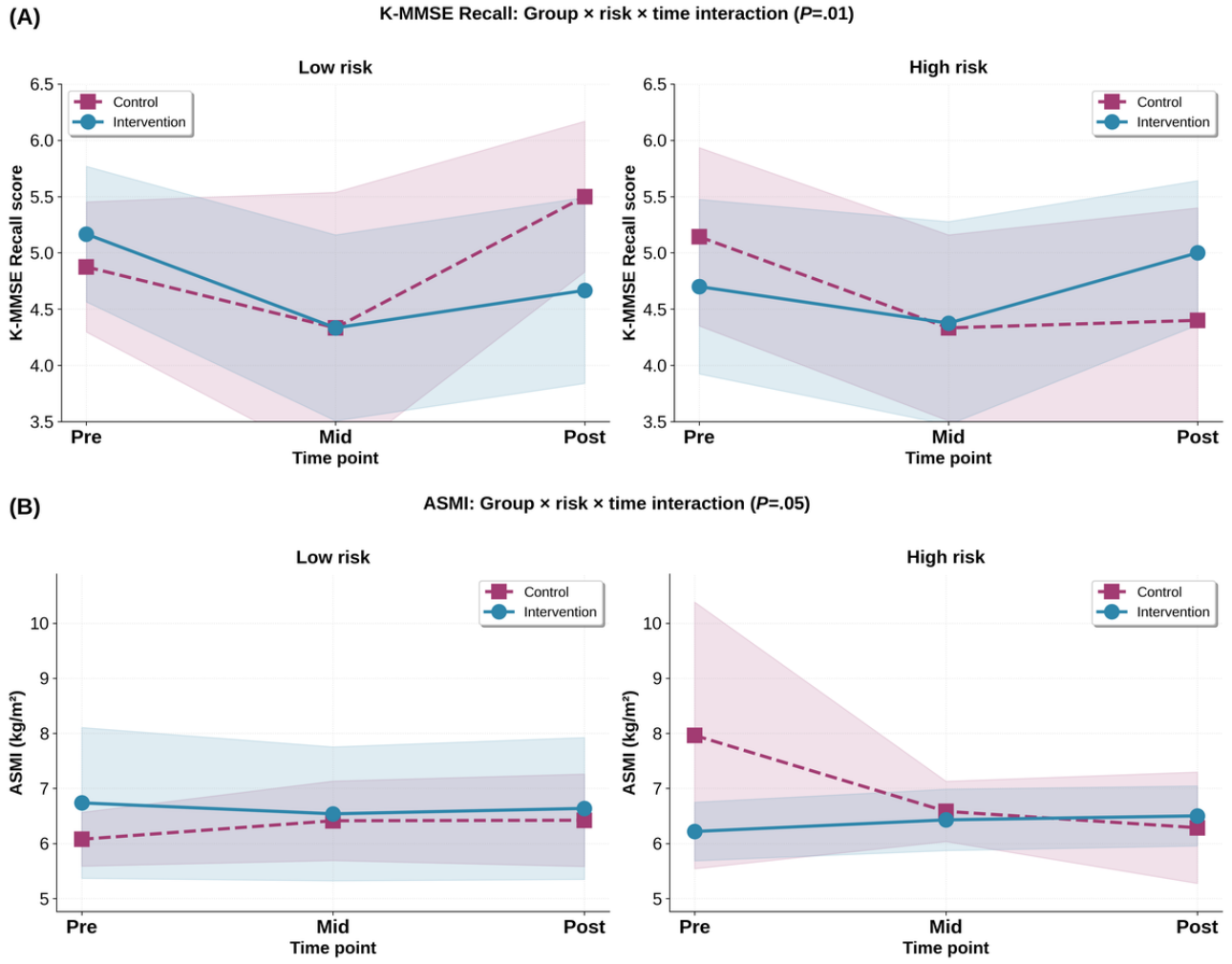
Risk-Stratified Analysis

Figure 4 presents outcomes with risk-level interactions. These analyses are exploratory and hypothesis-generating because of the small subgroup cell sizes (n=5-8 per cell). High risk was defined as meeting the criteria for cognitive impairment (K-MMSE ≤21 or Montreal Cognitive Assessment [MoCA-K] <23) combined with a history of recurrent falls.

A significant uncorrected 3-way interaction (group × risk × time) was observed for K-MMSE Recall (Wald χ^2_2=6.16, *P*=.01; Figure 4A). Among the high-risk participants, the intervention group showed an improvement in Recall (4.70-5.00; *d*=0.73), whereas the control group declined (5.14-4.40). The low-risk participants showed the opposite pattern (*d*=−0.99), suggesting that cognitive effects may be concentrated among individuals with greater baseline impairment.

The 3-way interaction for ASMI reached borderline significance (Wald χ^2_2=6.00, *P*=.05; Figure 4B). Among the high-risk participants, the between-group effect size was large (*d*=0.90): the control group showed a substantial decline (8.0-6.3 kg/m^2), whereas the intervention group maintained ASMI levels (6.2-6.5 kg/m^2). No other outcome variables showed significant uncorrected 3-way interactions (group × risk × time), including the MoCA-K total (*P*=.71) and subscale scores (Memory, *P*=.83; Attention, *P*=.71; Language, *P*=.71; Visuospatial, *P*=.75; Orientation, *P*=.89), K-MMSE total (*P*=.71) and other subscale scores (Recall, *P*=.37; Orientation, *P*=.85; Attention, *P*=.75; Language, *P*=.71; Visuospatial, *P*=.75), and sensorimotor function measures (BBS, *P*=.82; TUG, *P*=.83; Figure-8 Walk, *P*=.75; grip strength, *P*=.81; 30-second Sit-to-Stand, *P*=.70; 2-Minute Step Test, *P*=.70; and ASM, *P*=.70).

Figure 4. Risk-stratified intervention effects on cognitive and sensory-motor frailty outcomes. (A) Korean Mini-Mental State Examination (K-MMSE) Recall scores: a significant uncorrected group × risk × time interaction was observed (*P*=.01). In the high-risk subgroup, the intervention group showed improvement in Recall from mid- to postintervention, whereas the control group continued to decline. (B) Appendicular Skeletal Muscle Mass Index (ASMI): a group × risk × time interaction reached borderline significance (*P*=.05). In the high-risk subgroup, the intervention group maintained relatively stable muscle mass, whereas the control group showed continued decline. Lines represent estimated marginal means, and shaded areas indicate 95% CIs.



Exploratory Game Performance Analyses

Post hoc exploratory correlations between game performance metrics and clinical outcomes (Figure S1 in Multimedia Appendix 1) suggested potential associations between cognitive challenge during gameplay and cognitive outcomes, although the small subgroup sizes require cautious interpretation.

XSL•FO
RenderX

Discussion

Principal Findings

This study examined the effects of a pressure sensor–based exergame intervention on cognitive and sensorimotor frailty in socially vulnerable older adults through 2 sequential studies. In the pilot RCT, no primary outcomes survived FDR correction, indicating that the present findings do not establish efficacy and should be interpreted as preliminary. Within this cautious framework, the pilot study (study 1), an uncontrolled single-arm design, showed pre-post improvements in balance and mobility that survived FDR correction. It also showed an uncorrected pre-post improvement in the K-MMSE Visuospatial subscale. This finding is consistent with evidence that proprioceptive training can enhance visuospatial processing through sensorimotor integration in the posterior parietal cortex [47,48]. In the RCT (study 2), nominally significant uncorrected signals were observed for ASM preservation. In exploratory risk-stratified analyses based on small subgroup cell sizes, comparable uncorrected signals emerged for K-MMSE Recall and ASMI among the high-risk participants. These exploratory patterns are hypothesis-generating only. They suggest that older adults with mild-to-moderate impairment may represent a potentially more responsive subgroup in whom the residual capacity for improvement is greater. Confirmation in adequately powered, risk-enriched trials is required.

Comparison With Prior Work

The uncorrected preservation of ASM, together with trends toward greater grip strength, may reflect the functional loading produced by the repeated weight-shifting and stepping movements in the Brain Step program. A meta-analysis of 24 RCTs reported that exercise interventions produced a significant increase in appendicular muscle mass (+0.4 kg) among community-dwelling older adults [49], and a recent exergame RCT similarly demonstrated significant improvements in ASM following exergame-based training in residents of long-term care facilities [50]. However, these findings should be interpreted cautiously given the lack of statistical significance after FDR correction in our study.

Although the intervention group showed numerically higher BBS scores and shorter TUG times, no significant group × time interaction was observed. Participants were independently ambulatory older adults with high baseline functional levels (BBS approximately 52-53 of 56; TUG approximately 7-8 seconds). At these baseline levels, the BBS and TUG had limited capacity to detect further improvement, indicating ceiling effects. This likely reflects a mismatch between the selected outcome measures and the functional level of the sample rather than definitive evidence of no benefit. Therefore, the lack of significant effects on balance and mobility should not be interpreted as evidence of no intervention benefit because these measures may have been insufficiently sensitive to detect change in this high-functioning sample. Consistent with this interpretation, previous studies have reported limited sensitivity of the BBS for detecting subtle balance changes in high-functioning older adults [51,52] and reduced responsiveness of the TUG in individuals with preserved mobility [53]. Conventional clinical assessments such as the BBS and TUG may not adequately capture subtle changes in dynamic balance or real-world mobility, particularly under complex or dual-task conditions [54,55]. These findings underscore the need to select outcome measures that are both sensitive and ecologically valid and that are appropriately matched to the baseline functional level of the target population.

The K-MMSE Recall pattern among the high-risk participants was consistent across both studies, aligning with exergame RCTs reporting cognitive effects in individuals with MCI [19,20,56] but not in healthy [57,58] or severely impaired [59] populations. The concurrent preservation of muscle mass alongside improved recall in the high-risk participants is consistent with growing evidence that lower limb muscle function and cognitive function are bidirectionally linked [60,61] and that dual-task training engaging both motor and cognitive systems may produce benefits across both domains [12,62]. However, these concurrent patterns were observed in small, underpowered subgroups and should be regarded as hypothesis-generating signals requiring confirmation in adequately powered trials.

Self-reported fall incidence showed a nonsignificant directional reduction favoring the intervention group at the postintervention assessment (intervention, 35.7% vs control, 63.6%; d=−0.42). At the 12-month follow-up, no falls were reported in either group. This likely reflects the shift from retrospective self-report to prospective caregiver monitoring rather than a true treatment effect [63]. Whether proprioceptive-motor training can attenuate fall risk warrants investigation in a larger sample with longer follow-up.

Implications for Digital Health Equity

The target population—low-income older adults living alone who were receiving Public Community Care services—is systematically underrepresented in digital health research. The successful completion of an RCT with this population, facilitated by a participatory research committee that endorsed randomization, demonstrates the feasibility of conducting rigorous evaluations within community care settings. This study completes a 3-phase research program progressing from digital health monitoring [24] through community care–mediated intervention, in which caregivers identify needs and lay workers deliver services [25], to targeted cognitive-motor training, suggesting that the standardized Public Community Care infrastructure offers a scalable platform for embedding technology-based interventions. The high retention rate despite the participants' advanced age (mean 81.5 years), low-income status, and living alone suggests that welfare center–based delivery is a feasible model for reaching underserved populations. From a policy perspective, pressure-sensor mat systems may offer a potentially cost- and resource-efficient option for community care settings.

Limitations

Several limitations should be considered. The small sample size (n=15 for the pilot study; n=31 for the pilot RCT) limited statistical power, and no findings survived FDR correction. The small sample size and the inclusion of outcomes across multiple

XSL•FO
RenderX

domains reflected 2 design priorities: embedding the intervention within the welfare center's existing infrastructure and adopting a multidimensional monitoring framework [24,25]. The risk-stratified subgroup analyses involved very small cell sizes (n=5-8 per subgroup), substantially limiting the reliability of subgroup effect estimates and increasing the risk of inflated type I error and unstable estimates. Although the observed effect sizes were small-to-medium overall (d=0.38-0.50) and medium-to-large in the high-risk subgroups (d=0.73-0.92), these findings should be interpreted as exploratory and hypothesis-generating and require confirmation in adequately powered trials with prespecified subgroup hypotheses. Nonetheless, low-income older adults receiving Public Community Care services constitute a hard-to-reach and underrepresented population that is rarely included in digital health research. We therefore deliberately examined the high-risk subgroup pattern that recurred from the single-arm pilot study (study 1) to the pilot RCT (study 2) and present it strictly as exploratory and hypothesis-generating to inform hypotheses, outcome selection, and sample enrichment for other investigators studying similarly constrained, hard-to-recruit populations.

Relatedly, the selected balance and mobility outcome measures were not well matched to the sample's baseline functional level. Several balance and mobility measures, particularly the BBS and TUG, exhibited ceiling effects because participants were already highly functional at enrollment. These ceiling effects limited the sensitivity of the measures to detect change; accordingly, the null findings for balance and mobility should be interpreted as inconclusive rather than negative. Future trials should select outcome measures that are matched to the target population's functional level and incorporate more sensitive, ecologically valid measures (eg, instrumented gait analysis or dual-task assessments).

Despite randomization, the control group had numerically higher baseline K-MMSE scores than the intervention group (25.73 vs 24.44). This minor imbalance raises the possibility of regression to the mean for cognitive outcomes. The use of a passive (no-contact) control group constitutes a key limitation of internal validity. Without an active or attention-matched comparator, any observed differences cannot be attributed unambiguously to the cognitive-motor training itself because the effects of attention, social contact, novelty, and facilitator engagement remain confounded with the intervention. Although passive control groups are often used in feasibility-stage studies, particularly when the primary objective is to evaluate delivery feasibility within an existing care infrastructure, definitive trials should employ active or sham-exergame controls [28] to strengthen causal inference.

Cognitive outcomes were limited to global screening measures (K-MMSE and MoCA-K) rather than comprehensive neuropsychological batteries, and practice effects cannot be fully excluded. Generalizability is limited by the single-site recruitment context. Multisite replication across populations with diverse socioeconomic characteristics is needed. For independently ambulatory older adults, the 1-year follow-up period may not have been sufficient to detect meaningful differences in fall incidence. The 16-week intervention period also represents a limitation for evaluating durable cognitive and sensorimotor changes.

Conclusion

The 2 sequential studies demonstrate the initial feasibility of delivering a pressure sensor–based exergaming intervention through the existing Public Community Care infrastructure as an engaging, community-based program for socially vulnerable older adults experiencing age-related cognitive and sensorimotor decline. Designed to operate within the limited resources of a community welfare center using nonspecialist staff as facilitators, the Brain Step program offers a scalable template for multisite expansion. However, the current evidence is preliminary. The observed effects did not survive FDR correction, and the preliminary signals for muscle mass preservation and cognitive improvement were concentrated among participants with mild-to-moderate impairments, suggesting that this subgroup may benefit most from movement-mediated cognitive-motor training. Adequately powered multisite trials with longer follow-up are needed to confirm these findings and to evaluate the integration of exergame-based interventions within personalized digital health care frameworks that combine wearable-based monitoring of daily functional changes with adaptive intervention delivery.

Acknowledgments

We are very grateful to the older adults and care service providers for their participation and to the staff and community caregivers at the Seongbuk Municipal Senior Welfare Center for their support. We are also deeply grateful to the late Mr. Jong-hoon Kim, father of Junesun Kim, for generously serving as the first pretester of the Brain Step system; his early contribution helped lay the foundation for this work. During manuscript preparation, the authors used Claude (Anthropic) to assist with English grammar correction, reference formatting and reordering, and general manuscript formatting. All artificial intelligence–assisted content was reviewed and edited by the authors, who take full responsibility for the content of this publication.

Funding

This study was supported by the ANCHOR through the Seoul ANCHOR Center, funded by the Ministry of Education (MOE) and the Seoul Metropolitan Government (2026-ANCHOR-01-003-01), the Pioneer Research Center Program (RS-2022-NR067566) through the National Research Foundation of Korea, and the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science and ICT (MSIT) (RS-2024-00348012).

XSL•FO
RenderX

Data Availability

The data contain sensitive personal and health information from socially vulnerable older adults. In accordance with the conditions of the Institutional Review Board approval (approval number KUIRB-2023-0360) and participant informed consent, the dataset cannot be made publicly available or shared with researchers who are not authorized under the original study protocol.

Authors' Contributions

Conceptualization: JK, SCH, SS
Data curation: JL, SHL
Formal analysis: SS, JL, HK
Funding acquisition: SS, JK
Investigation: JL
Methodology: SS, SCH
Project administration: SS, SCH
Resources: JK
Software: SC, SHL
Supervision: SS, SCH, JK
Validation: SS
Visualization: SC
Writing – original draft: SS, SCH
Writing – review & editing: SS, SCH, HK, JK

Conflicts of Interest

SC and SHL work at the MIDAS Health & Technology Co., which developed the Brain Step system. SS, SCH, JK, and SC are co-inventors on a patent related to the Brain Step technology. SC and SHL contributed to the Brain Step system development including hardware specifications and software design.

Multimedia Appendix 1
Additional baseline characteristics, outcome measures, fall incidence results, and exploratory correlation analyses supporting the findings of this study.
[DOCX File , 318 KB - aging_v9i1e93467_app1.docx]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1150 KB - aging_v9i1e93467_app2.pdf]

References

1. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125 [FREE Full text] [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
2. Zhang X, Zeng R, Zhu A, Xie F, Ye D, Chen L, et al. Association between sensory impairment and cognitive frailty among older people: evidence from four nationwide cohort studies. J Nutr Health Aging 2025 Aug;29(8):100590 [FREE Full text] [doi: 10.1016/j.jnha.2025.100590] [Medline: 40441121]
3. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 2001 Mar 01;56(3):M146-M157. [doi: 10.1093/gerona/56.3.M146]
4. Kelaiditi E, Cesari M, Canevelli M, van Kan GA, Ousset P, Gillette-Guyonnet S, IANA/IAGG. Cognitive frailty: rational and definition from an (I.A.N.A./I.A.G.G.) international consensus group. J Nutr Health Aging 2013 Sep;17(9):726-734 [FREE Full text] [doi: 10.1007/s12603-013-0367-2] [Medline: 24154642]
5. Beier F, Löffler M, Nees F, Hausner L, Frölich L, Flor H. Sensory and motor correlates of frailty: dissociation between frailty phenotype and frailty index. BMC Geriatr 2022 Sep 15;22(1):755 [FREE Full text] [doi: 10.1186/s12877-022-03416-6] [Medline: 36109693]
6. Humes LE, Busey TA, Craig J, Kewley-Port D. Are age-related changes in cognitive function driven by age-related changes in sensory processing? Atten Percept Psychophys 2013 Apr 20;75(3):508-524 [FREE Full text] [doi: 10.3758/s13414-012-0406-9] [Medline: 23254452]
7. Zhang W, Low L, Schwenk M, Mills N, Gwynn J, Clemson L. Review of gait, cognition, and fall risks with implications for fall prevention in older adults with dementia. Dement Geriatr Cogn Disord 2019 Nov 19;48(1-2):17-29. [doi: 10.1159/000504340] [Medline: 31743907]

XSL•FO
RenderX

8. Montero-Odasso M, Speechley M. Falls in cognitively impaired older adults: implications for risk assessment and prevention. J Am Geriatr Soc 2018 Feb 10;66(2):367-375. [doi: 10.1111/jgs.15219] [Medline: 29318592]
9. Sha T, Zhang Y, Wei J, Li C, Zeng C, Lei G, et al. Sarcopenia and risk of cognitive impairment: cohort study and Mendelian randomization analysis. JMIR Aging 2025 Jun 11;8:e66031 [FREE Full text] [doi: 10.2196/66031] [Medline: 40499039]
10. Yuan S, Larsson SC. Epidemiology of sarcopenia: prevalence, risk factors, and consequences. Metabolism 2023 Jul;144:155533 [FREE Full text] [doi: 10.1016/j.metabol.2023.155533] [Medline: 36907247]
11. Sturnieks DL, Hicks C, Smith N, Ratanapongleka M, Menant J, Turner J, et al. Exergame and cognitive training for preventing falls in community-dwelling older people: a randomized controlled trial. Nat Med 2024 Jan;30(1):98-105. [doi: 10.1038/s41591-023-02739-0] [Medline: 38228913]
12. Begde A, Alqurafi A, Pain MTG, Blenkinsop G, Wilcockson TDW, Hogervorst E. The effectiveness of home-based exergames training on cognition and balance in older adults: a comparative quasi-randomized study of two exergame interventions. Innov Aging 2023;7(8):igad102 [FREE Full text] [doi: 10.1093/geroni/igad102] [Medline: 37941829]
13. Sherrington C, Fairhall NJ, Wallbank GK, Tiedemann A, Michaleff ZA, Howard K, et al. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev 2019 Jan 31;1:CD012424 [FREE Full text] [doi: 10.1002/14651858.CD012424.pub2] [Medline: 30703272]
14. Eggleston K, Kim D. Dementia care in a rapidly aging society. Alzheimers Dement 2025 May 23;21(5):e70268. [doi: 10.1002/alz.70268] [Medline: 40407100]
15. Seoul dementia prevalence status by district. National Institute of Dementia. 2023. URL: https://www.nid.or.kr/info/today_list_2023.aspx [accessed 2024-12-31]
16. Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. The Lancet 2024 Aug;404(10452):572-628. [doi: 10.1016/s0140-6736(24)01296-0]
17. Jiang J, Guo W, Wang B. Effects of exergaming on executive function of older adults: a systematic review and meta-analysis. PeerJ 2022;10:e13194 [FREE Full text] [doi: 10.7717/peerj.13194] [Medline: 35433124]
18. Goumopoulos C, Danousis M, Drakakis E. Feasibility and initial impact of a multi-technology exergaming intervention on motor and cognitive function in older adults: a pilot randomized controlled trial. International Journal of Human–Computer Interaction 2025 Jul 24;42(4):2433-2462. [doi: 10.1080/10447318.2025.2528146]
19. Liu C, Cheng F, Wei M, Liao Y. Effects of exergaming-based tai chi on cognitive function and dual-task gait performance in older adults with mild cognitive impairment: a randomized control trial. Front Aging Neurosci 2022 Mar 15;14:761053 [FREE Full text] [doi: 10.3389/fnagi.2022.761053] [Medline: 35370622]
20. Li A, Qiang W, Li J, Geng Y, Qiang Y, Zhao J. Effectiveness of an exergame-based training program on physical and cognitive function in older adults with cognitive impairment: a randomized controlled trial in rural China. BMC Geriatr 2025 Nov 11;25(1):892. [doi: 10.1186/s12877-025-06341-6] [Medline: 41219720]
21. Sturnieks DL, Hicks C, Smith N, Ratanapongleka M, Menant J, Turner J, et al. Exergame and cognitive training for preventing falls in community-dwelling older people: a randomized controlled trial. Nat Med 2024 Jan 16;30(1):98-105. [doi: 10.1038/s41591-023-02739-0] [Medline: 38228913]
22. Herrera AP, Snipes SA, King DW, Torres-Vigil I, Goldberg DS, Weinberg AD. Disparate inclusion of older adults in clinical trials: priorities and opportunities for policy and practice change. Am J Public Health 2010 Apr 01;100 Suppl 1:S105-S112. [doi: 10.2105/AJPH.2009.162982] [Medline: 20147682]
23. Kunonga TP, Spiers GF, Beyer FR, Hanratty B, Boulton E, Hall A, et al. Effects of digital technologies on older people's access to health and social care: umbrella review. J Med Internet Res 2021 Nov 24;23(11):e25887 [FREE Full text] [doi: 10.2196/25887] [Medline: 34821564]
24. Song S, Seo Y, Hwang S, Kim H, Kim J. Digital phenotyping of geriatric depression using a community-based digital mental health monitoring platform for socially vulnerable older adults and their community caregivers: 6-week living lab single-arm pilot study. JMIR Mhealth Uhealth 2024 Jun 17;12:e55842 [FREE Full text] [doi: 10.2196/55842] [Medline: 38885033]
25. Song S, Hwang S, Kim H, Kim J. Digital-human public community care integration for chronic pain in low-income older adults in a 6-week living lab setting: quasi-experimental feasibility study. JMIR Aging 2026 Apr 20;9:e85611 [FREE Full text] [doi: 10.2196/85611] [Medline: 42061838]
26. 2025 customized care service for older adults: program guidelines in Korean. Ministry of Health and Welfare (South Korea). Sejong, Korea: Ministry of Health and Welfare; 2025 Jan 07. URL: https://www.mohw.go.kr/board.es?mid=a10411010100&bid=0019&act=view&list_no=1484224 [accessed 2026-07-17]
27. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
28. Zhao Y, Feng H, Wu X, Du Y, Yang X, Hu M, et al. Effectiveness of exergaming in improving cognitive and physical function in people with mild cognitive impairment or dementia: systematic review. JMIR Serious Games 2020 Jun 30;8(2):e16841 [FREE Full text] [doi: 10.2196/16841] [Medline: 32602841]

XSL•FO
RenderX

29. Fang Q, Ghanouni P, Anderson SE, Touchett H, Shirley R, Fang F, et al. Effects of exergaming on balance of healthy older adults: a systematic review and meta-analysis of randomized controlled trials. Games Health J 2020 Feb 01;9(1):11-23. [doi: 10.1089/g4h.2019.0016] [Medline: 31800322]
30. Kang Y, Na D, Hahn S. A validity study on the Korean Mini-Mental State Examination (K-MMSE) in dementia patients. J Korean Neurol Assoc 1997;15(2):300-308 [FREE Full text]
31. Berg K, Wood-Dauphinee S, Williams J, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83 Suppl 2:S7-11. [Medline: 1468055]
32. Jung H, Park J, Shim J, Kim M, Hwang M, Kim S. Reliability test of Korean version of Berg Balance Scale. J Korean Acad Rehabil Med 2006;30:611-618 [FREE Full text]
33. Podsiadlo D, Richardson S. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
34. Chen L, Woo J, Assantachai P, Auyeung T, Chou M, Iijima K, et al. Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020 Mar;21(3):300-307.e2. [doi: 10.1016/j.jamda.2019.12.012] [Medline: 32033882]
35. Bae J, Li X, Kim T, Bang H, Lee S, Seo DY. Prediction models of grip strength in adults above 65 years using Korean National Physical Fitness Award Data from 2009 to 2019. Eur Geriatr Med 2023 Oct 01;14(5):1059-1064. [doi: 10.1007/s41999-023-00817-7] [Medline: 37392358]
36. Rikli R, Jones C. Development and validation of a functional fitness test for community-residing older adults. J Aging Phys Act 1999;7(2):129-161. [doi: 10.1123/japa.7.2.129]
37. Huang L, Liu Y, Lin T, Hou L, Song Q, Ge N, et al. Reliability and validity of two hand dynamometers when used by community-dwelling adults aged over 50 years. BMC Geriatr 2022 Jul 15;22(1):580 [FREE Full text] [doi: 10.1186/s12877-022-03270-6] [Medline: 35840905]
38. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
39. Lee J, Dong Woo Lee, Cho S, Na DL, Hong Jin Jeon, Kim S, You Ra Lee, Maeng Je Cho. Brief screening for mild cognitive impairment in elderly outpatient clinic: validation of the Korean Version of the Montreal Cognitive Assessment. J Geriatr Psychiatry Neurol 2008 Jun 01;21(2):104-110. [doi: 10.1177/0891988708316855]
40. Kang I, Beom I, Cho J, Son H. Accuracy of Korean-Mini-Mental Status Examination based on Seoul Neuro-Psychological Screening Battery II Results. Korean J Fam Med 2016 May;37(3):177-181 [FREE Full text] [doi: 10.4082/kjfm.2016.37.3.177] [Medline: 27274389]
41. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
42. Yesavage JA, Sheikh JI. Geriatric Depression Scale (GDS). Clinical Gerontologist 2008 Oct 25;5(1-2):165-173. [doi: 10.1300/J018v05n01_09]
43. Bellamy N, Buchanan W, Goldsmith C, Campbell J, Stitt L. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15(12):1833-1840. [Medline: 3068365]
44. LIANG K, ZEGER SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73(1):13-22. [doi: 10.1093/biomet/73.1.13]
45. National RC. The Prevention and Treatment of Missing Data in Clinical Trials. Washington, DC: National Academies Press; 2010.
46. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res 2014 Oct;23(5):440-459. [doi: 10.1177/0962280213476378] [Medline: 23427225]
47. Boyle S, Fothergill M, Metcalfe J, Docherty S, Haskell-Ramsay C. The effects of low-intensity multimodal proprioceptive exercise on cognitive function in older adults. J Phys Act Health 2021 Jan 01;18(1):2-7. [doi: 10.1123/jpah.2020-0134] [Medline: 33260144]
48. Bao R, Chang S, Liu R, Wang Y, Guan Y. Research status of visuospatial dysfunction and spatial navigation. Front Aging Neurosci 2025 May 14;17:1609620 [FREE Full text] [doi: 10.3389/fnagi.2025.1609620] [Medline: 40438501]
49. González-Rocha A, Mendez-Sanchez L, Ortíz-Rodríguez MA, Denova-Gutiérrez E. Effect of exercise on muscle mass, fat mass, bone mass, muscular strength and physical performance in community dwelling older adults: systematic review and meta-analysis. Aging and disease 2022;13(5):1421. [doi: 10.14336/ad.2022.0215]
50. Tuan S, Chang L, Sun S, Li C, Chen G, Tsai Y. Assessing the clinical effectiveness of an exergame-based exercise training program using ring fit adventure to prevent and postpone frailty and sarcopenia among older adults in rural long-term care facilities: randomized controlled trial. J Med Internet Res 2024 Jul 18;26:e59468 [FREE Full text] [doi: 10.2196/59468] [Medline: 39024000]

XSL•FO
RenderX

51. Muir SW, Berg K, Chesworth B, Speechley M. Use of the Berg Balance Scale for predicting multiple falls in community-dwelling elderly people: a prospective study. Phys Ther 2008 Apr;88(4):449-459. [doi: 10.2522/ptj.20070251] [Medline: 18218822]
52. Downs S, Marquez J, Chiarelli P. The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review. Journal of Physiotherapy 2013 Jun;59(2):93-99. [doi: 10.1016/s1836-9553(13)70161-9]
53. Beauchet O, Fantino B, Allali G, Muir SW, Montero-Odasso M, Annweiler C. Timed Up and Go test and risk of falls in older adults: a systematic review. J Nutr Health Aging 2011 Dec 3;15(10):933-938 [FREE Full text] [doi: 10.1007/s12603-011-0062-0] [Medline: 22159785]
54. Verghese J, Holtzer R, Lipton RB, Wang C. Quantitative gait markers and incident fall risk in older adults. J Gerontol A Biol Sci Med Sci 2009 Aug 06;64(8):896-901 [FREE Full text] [doi: 10.1093/gerona/glp033] [Medline: 19349593]
55. Montero-Odasso M, Verghese J, Beauchet O, Hausdorff JM. Gait and cognition: a complementary approach to understanding brain function and the risk of falling. J Am Geriatr Soc 2012 Nov;60(11):2127-2136 [FREE Full text] [doi: 10.1111/j.1532-5415.2012.04209.x] [Medline: 23110433]
56. Zhu Y, Lin C, Yang H, Jin G, Chiu H. Effects of exergaming on cognitive functions and loneliness of older adults with cognitive frailty. Int J Geriatr Psychiatry 2023 Jun;38(6):e5944. [doi: 10.1002/gps.5944] [Medline: 37260087]
57. Ordnung M, Hoff M, Kaminski E, Villringer A, Ragert P. No overt effects of a 6-week exergame training on sensorimotor and cognitive function in older adults. a preliminary investigation. Front Hum Neurosci 2017;11:160 [FREE Full text] [doi: 10.3389/fnhum.2017.00160] [Medline: 28420973]
58. Chen X, Wu L, Feng H, Ning H, Wu S, Hu M, et al. Comparison of exergames versus conventional exercises on the health benefits of older adults: systematic review with meta-analysis of randomized controlled trials. JMIR Serious Games 2023 Jun 22;11:e42374 [FREE Full text] [doi: 10.2196/42374] [Medline: 37347534]
59. van Santen J, Dröes R, Twisk JWR, Blanson Henkemans OA, van Straten A, Meiland FJM. Effects of exergaming on cognitive and social functioning of people with dementia: a randomized controlled trial. J Am Med Dir Assoc 2020 Dec;21(12):1958-1967.e5 [FREE Full text] [doi: 10.1016/j.jamda.2020.04.018] [Medline: 32651132]
60. Kim A, Ko H. Lower limb function in elderly Korean adults is related to cognitive function. J Clin Med 2018 May 01;7(5):99 [FREE Full text] [doi: 10.3390/jcm7050099] [Medline: 29723997]
61. Tessier A, Wing SS, Rahme E, Morais JA, Chevalier S. Association of low muscle mass with cognitive function during a 3-year follow-up among adults aged 65 to 86 years in the Canadian Longitudinal Study on Aging. JAMA Netw Open 2022 Jul 01;5(7):e2219926 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.19926] [Medline: 35796211]
62. Liao Y, Chen I, Hsu W, Tseng H, Wang R. Effect of exergaming versus combined exercise on cognitive function and brain activation in frail older adults: a randomised controlled trial. Ann Phys Rehabil Med 2021 Sep;64(5):101492. [doi: 10.1016/j.rehab.2021.101492] [Medline: 33454398]
63. Romli MH, Mackenzie L, Tan PJ, Chiew RO, Tan SH, Tan MP. Comparison of retrospective and prospective falls reporting among community-dwelling older people: findings from two cohort studies. Front Public Health 2021 Mar 12;9:612663 [FREE Full text] [doi: 10.3389/fpubh.2021.612663] [Medline: 33777881]

Abbreviations

ASM: appendicular skeletal muscle mass
ASMI: Appendicular Skeletal Muscle Mass Index
BBS: Berg Balance Scale
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
FDR: false discovery rate
GDS-SF: Geriatric Depression Scale-Short Form
GEE: generalized estimating equation
K-MMSE: Korean Mini-Mental State Examination
MoCA-K: Korean Montreal Cognitive Assessment
OR: odds ratio
PSQI: Pittsburgh Sleep Quality Index
RCT: randomized controlled trial
TUG: Timed Up and Go
WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index

XSL•FO
RenderX

Edited by P Jansons; submitted 13.Feb.2026; peer-reviewed by K Lee, MC Espada; comments to author 16.Mar.2026; revised version received 30.Jun.2026; accepted 30.Jun.2026; published 24.Jul.2026.

Please cite as:
Song S, Hahm SC, Chang S, Lee J, Kang H, Lee SH, Kim J
Feasibility and Preliminary Effects of a Pressure Sensor–Based Exergame on Cognitive and Sensorimotor Frailty in Low-Income Older Adults Receiving Public Community Care Services: Pilot Randomized Controlled Trial
JMIR Aging 2026;9:e93467
URL: https://aging.jmir.org/2026/1/e93467
doi:10.2196/93467
PMID:

©Sunmi Song, Suk-Chan Hahm, Seyun Chang, Jinseung Lee, Heewon Kang, Sang Hun Lee, Junesun Kim. Originally published in JMIR Aging (https://aging.jmir.org), 24.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations of Midlife Gait Speed and Muscle Performance With Falls in Later Life: Secondary Analysis of the MIDUS Cohort

Roee Hayek[1*], MSci; Yaniv Nudelman[1*], PhD; Dan Keren[1], MSci; Shmuel Springer[1], PhD

Ariel University, Ariel, Israel

*these authors contributed equally

Corresponding Author:
Shmuel Springer, PhD
Ariel University
3 Kiryat Hamada St
Ariel, 40700
Israel
Phone: 972 584572869
Email: shmuels@ariel.ac.il

Abstract

Background: Falls are a major cause of morbidity among older adults, yet fall risk may begin to emerge during midlife. Physical performance measures may identify early functional vulnerability associated with later-life health outcomes.

Objective: This study examined whether midlife gait speed and muscle performance were associated with recurrent falls later in life.

Methods: This study was a secondary analysis of prospectively collected data from the Midlife in the United States (MIDUS) cohort. Independent adults (n=309) aged 45 to 64 years at wave 2 (2004-2009) who completed physical performance assessments were followed to wave 3 (2013-2021). Baseline measures included gait speed from a 50-foot walk test, grip strength, sit-to-stand time, and calculated lower-extremity power. The primary analysis examined associations between baseline physical performance and recurrent falls at follow-up (≥2 falls in the previous 12 months). A sensitivity analysis repeated the primary models but was restricted to participants without a baseline history of recurrent falls.

Results: At follow-up, 20.4% (63/309) of participants reported recurrent falls. Slower midlife gait speed was associated with higher odds of recurrent falls (odds ratio 1.43, 95% CI 1.05-1.95; *P*=.02), whereas muscle performance measures were not associated with falls in the primary analysis. The association between slower gait speed and recurrent falls remained significant in the sensitivity analysis. Over follow-up, muscle performance declined, whereas gait speed remained stable.

Conclusions: Slower midlife gait speed was associated with higher odds of recurrent falls, suggesting that fall vulnerability may be detectable as early as midlife. Given the exploratory nature of this study, further studies are warranted to confirm these findings.

(JMIR Aging 2026;9:e103917) doi:10.2196/103917

KEYWORDS

middle age; mobility; fall risk; subjective age; physical performance

Introduction

Maintaining locomotor ability across the adult lifespan is a central goal of healthy aging. In many cases, increased life expectancy is not accompanied by preserved functional capacity, making mobility-related outcomes in later life a major health concern [1]. At the individual level, mobility limitations in older adulthood are associated with reduced participation, loss of independence, and increased fall risk [2]. Falls represent a substantial public health burden, contributing to substantial health care use, economic costs, and mortality among older adults [3].

Fall prevention efforts focus on identifying risk factors in clinical and research settings. Among these, markers of physical function are particularly informative because they reflect mobility capacity and physiological reserve and have been associated with increased fall risk and adverse health outcomes [4]. Recurrent falls, defined as 2 or more falls within 12 months, are of particular concern because they are more strongly associated with frailty, hospitalization, and mortality than

XSL•FO
RenderX

isolated falls [3]. Despite extensive research and prevention initiatives, population-level progress remains limited, with increasing fall-related mortality among US adults aged 65 years and older [5,6], and annual fall prevalence remaining approximately 1 in 4 [6].

Importantly, falls are not limited to older adulthood, and midlife may represent a critical window for prevention. Recent harmonized cohort analyses across multiple countries report an annual fall prevalence of 17% to 28% among community-dwelling middle-aged adults. Fall rates increase markedly across midlife, nearly tripling from early (45-54 years) to late (55-64 years) midlife [7,8]. Furthermore, approximately 1 in 5 middle-aged adults report mobility limitations despite the absence of overt clinical conditions [9]. These impairments are often unrecognized in routine care, highlighting the importance of early detection of mobility decline to enable timely preventive interventions [8,10]. Screening physical performance during midlife may therefore provide insight into early functional vulnerability.

To better understand functional aging dynamics, several population-based cohorts initiate assessments in midlife to support a life-course perspective on aging [8,11-14]. One such example is the Midlife in the United States (MIDUS) study, a national longitudinal cohort that integrates objective assessments of physical function with psychosocial and health measures [12,13]. These include fall prevalence as well as self-rated health and subjective age, indicators that have been independently associated with subsequent physical decline, frailty, disability, and mortality [15-17]. Together, this multidimensional framework enables the examination of physical function alongside health perceptions in midlife and their relevance to aging-related outcomes.

Among the wide range of physical function variables available in MIDUS, gait speed and muscle performance capture distinct yet complementary aspects of physical function and mobility. These domains have been widely used to elucidate mechanisms and predictors of age-related decline [14,18-22]. Gait speed is an integrative marker reflecting multisystem physiological function, including musculoskeletal, cardiorespiratory, neurological, and cognitive domains, and is widely recognized as a “sixth vital sign” for aging and frailty [23]. Muscle performance specifically quantifies neuromuscular capacity and is a direct indicator of sarcopenia and physical reserve [20,24].

Several studies have demonstrated that slower gait speed and reduced muscle performance are associated with recurrent falls. However, most prior research has focused on older and often more functionally impaired populations [22,25-27]. Studies including middle-aged adults have shown that these measures are associated with mobility difficulties [14,28], whereas a study of community-dwelling adults aged 50 to 65 years reported no association with adverse events, including falls [14]. Evidence specifically examining the relationship between physical performance and falls in middle-aged populations therefore remains limited. Examining this association in independent, community-dwelling, middle-aged adults may help identify risk factors that emerge before overt functional decline.

Using longitudinal MIDUS data, this study primarily examined whether midlife gait speed and muscle performance were associated with recurrent falls later in life. In addition, longitudinal changes in gait speed and muscle performance from midlife onward were examined to characterize their trajectories across adulthood. Secondary analyses assessed whether these same measures were associated with subjective age and self-rated health at follow-up.

Methods

Study Design and Participants

We conducted a longitudinal analysis using data from the MIDUS cohort, focusing on wave 2 (baseline, 2004-2009) [12] and wave 3 (follow-up, 2013-2021) [13]. Although MIDUS is a prospective cohort study, the present investigation represents a secondary analysis linking midlife physical performance at wave 2 with outcomes assessed at wave 3. This study is reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [29], and the completed checklist is available in Table S1 in Multimedia Appendix 1.

The analytic sample was defined based on eligibility at wave 2 and was restricted to relatively healthy, independent adults aged 45 to 64 years who did not use walking aids and were able to complete the 50-foot walking test. To minimize confounding and strengthen internal validity, participants with conditions likely to influence physical performance or fall risk were excluded. Specifically, participants were excluded if they reported a history of stroke or transient ischemic attack, head injury requiring an overnight emergency department or hospital stay, neurological disorder or cancer, blindness in one or both eyes, prior antiepileptic medication use, or opioid use at wave 2.

Physical performance measures (gait speed and muscle performance) obtained at wave 2 were examined in relation to the primary outcome of recurrent falls reported at wave 3. At wave 3, participants were asked to report the number of falls experienced during the previous 12 months. Responses were recorded as a count variable and subsequently categorized as recurrent falls (≥2 falls) versus nonrecurrent (0-1 falls). In addition, we examined changes in physical performance between the 2 waves. Secondary exploratory analyses examined associations between baseline physical performance and subjective age and self-rated health at follow-up. All MIDUS data are fully deidentified and anonymized, ensuring compliance with ethical research standards.

Physical Performance Measures

As part of the MIDUS data collection, participants completed the 50-foot timed walk test twice. They were instructed to walk at their usual comfortable speed from a standing start to a designated turnaround point and back. For the present analysis, gait speed (m/s) was calculated by dividing the total distance (50 feet; 15.24 m) by completion time in seconds. The average of the 2 trials was used in all analyses.

During the MIDUS assessment, handgrip strength was measured using a handheld dynamometer and recorded in kilograms

XSL•FO
RenderX

(equivalent to kilogram-force). Participants performed 3 maximal-effort trials. For the present analysis, the mean value across the 3 trials of the dominant hand was used.

Participants also completed the 5-repetition sit-to-stand test as part of the MIDUS protocol. For the present analysis, 2 indicators of lower-limb muscle performance were derived from this test: the total sit-to-stand completion time (seconds) and lower-extremity muscle power. Sit-to-stand completion time was obtained directly from the MIDUS dataset, whereas lower-extremity muscle power was calculated for the present analysis using an established allometric equation incorporating body mass, height, chair height (approximated as 45 cm), and total sit-to-stand duration [30]:



Self-Perceived Age and Health

Subjective age was measured across MIDUS waves by asking, "Many people feel older or younger than they actually are. What age do you feel most of the time?" Participants reported the age they felt in years as an open-ended numeric response, which was treated as a continuous variable in subsequent analyses. Self-rated health was assessed using 2 ordinal items: general physical health ("In general, would you say your physical health is excellent, very good, good, fair, or poor?"; 1=excellent to 5=poor) and comparative health relative to age peers ("Compared to others your age, would you say your health is much better, somewhat better, about the same, somewhat worse, or much worse?"; 1=much better to 5=much worse).

Statistical Analysis

Descriptive statistics are presented as mean (SD) or median with IQR, as appropriate. Continuous variables were assessed for distribution using visual inspection of histograms and Q-Q plots. Several complementary modeling approaches were used to address the study objectives. For the primary analysis, logistic regression examined whether midlife gait speed and muscle performance (grip strength, sit-to-stand completion time, and lower-limb muscle power) were associated with recurrent falls (≥2 falls in the preceding 12 months) at follow-up. Except for sit-to-stand completion time, all physical measures were directionally inverted prior to modeling so that odds ratios (ORs) reflected the effect of poorer performance. Age-related changes in gait speed and muscle performance between MIDUS waves were examined using linear mixed-effects models with random intercepts for participant ID and adjustment for baseline age and follow-up interval to estimate within-person change over time.

Secondary analyses evaluated associations of physical measures with subjective age and self-perceived health at follow-up, using linear regression for subjective age and ordinal regression for both self-perceived health measures (general physical health and comparative health). Primary and secondary regression models were adjusted for follow-up interval, baseline age, sex, BMI, height, baseline symptom burden (count of self-reported symptoms and chronic conditions), and the baseline value of the respective dependent variable.

Additionally, subgroup analyses were conducted by stratifying participants by age at follow-up (<65 or ≥65 years). The magnitude of associations in linear models was assessed using standardized regression coefficients (β), with values of 0.1, 0.3, and 0.5 denoting small, medium, and large effects, respectively [31,32]. For multiple testing, separate domains were defined in advance. Consistent with the literature, gait speed and muscle performance were treated as distinct domains. Gait speed is an integrative measure of multisystem physiological function [23], whereas grip strength, sit-to-stand time, and sit-to-stand power together index neuromuscular capacity [20,24]. The significance threshold was therefore adjusted within the muscle performance domain (P<.016) but not for gait speed, and unadjusted P values with 95% CIs are reported throughout.

As a sensitivity analysis, each primary adjusted and unadjusted model, for gait speed and each muscle performance measure, was refitted in participants free of recurrent falls at baseline (wave 2). This tested whether midlife performance was associated with newly emerging, rather than persistent, recurrent falls at follow-up. These sensitivity analyses were not adjusted for multiplicity and were interpreted as supportive rather than confirmatory. Statistical analyses were conducted in R (version 4.3.0; R Foundation for Statistical Computing) using the *tidyverse* and *lme4* packages [33-35].

Ethical Considerations

This study is a secondary analysis of deidentified data from the MIDUS Biomarker Project. The original data collection received ethics approval from the institutional review board at the University of Wisconsin-Madison, and informed consent was obtained from all participants by the MIDUS investigators.

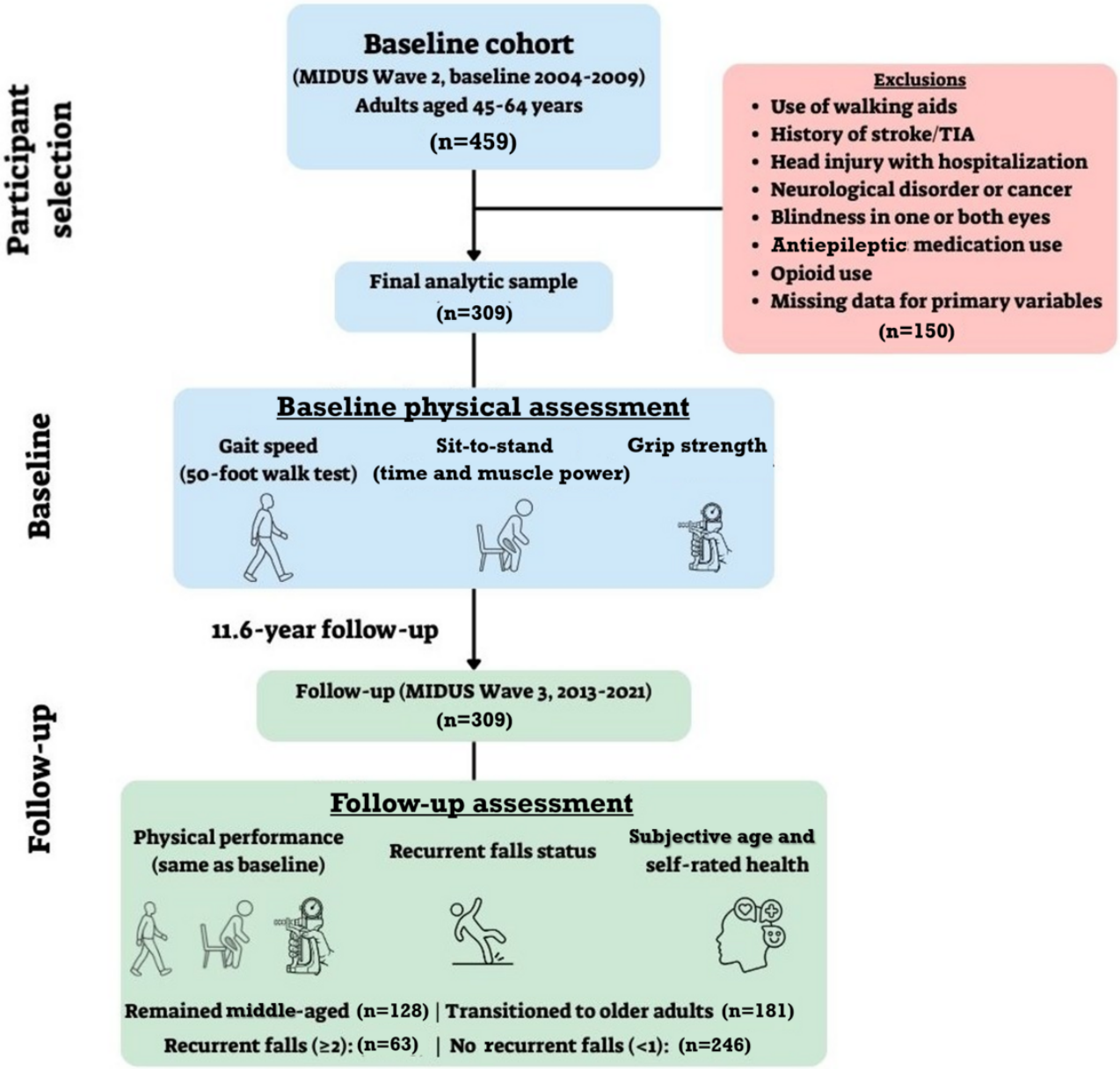
Results

Overview

The participant selection process and the final analytic sample are shown in Figure 1. The initial dataset included 459 middle-aged participants at baseline (MIDUS wave 2). After applying prespecified exclusion criteria and removing cases with missing data, the final analytic sample consisted of 309 participants, 45% (n=141) of whom were male. The mean age was 54.62 (SD 5.60) years at baseline and 66.26 (SD 5.66) years at follow-up. At follow-up, 181 (58.6%) participants were 65 years or older. The mean interval between waves was 11.64 (SD 1.25) years. The median baseline count of self-reported symptoms and chronic conditions was 3 (IQR 2-4). A total of 63 (20.4%) participants reported recurrent falls (≥2 falls) at follow-up; of these, 15 (23.8%) had already been recurrent fallers at wave 2, and 48 (76.2%) were newly recurrent at follow-up. Demographic characteristics, health indicators, and physical performance measures at MIDUS waves 2 and 3 are presented in Table 1.

XSL•FO
RenderX

Figure 1. Flow diagram of participant selection and the analytic sample from the Midlife in the United States (MIDUS) cohort. TIA: transient ischemic attack.



XSL•FO
RenderX

Table 1. Sample characteristics in wave 2 and wave 3.

Variable	MIDUS wave 2 (baseline)	MIDUS wave 3 (follow-up)
General characteristics (N=309)		
Age (y), mean (SD)[a]	54.62 (5.60)	66.26 (5.66)
Female, n (%)	168 (54.4)	168 (54.4)
BMI (kg/m^2), mean (SD)	29.25 (5.85)	29.74 (6.43)
Reported symptoms and chronic conditions, median (IQR)	3 (2-4)	4 (2-6)
Recurrent fallers[b], n (%)	55 (17.8)	63 (20.4)
Physical performance, mean (SD)		
Gait speed[c] (m/s)	1.11 (0.20)	1.11 (0.20)
Grip strength (kg)	36.40 (12.15)	30.99 (10.56)[d]
STS[e] time (s)	9.24 (3.21)	10.52 (4.10)[d]
STS mean power (W)	360.55 (163.47)	312.96 (141.26)[d]
Self-perceived age and health (n=280)		
Subjective age (y), mean (SD)	42.28 (10.80)	49.57 (9.59)
Differences between true age and subjective age, mean (SD)[f]	9.36 (9.42)	11.35 (8.40)
Self-perceived general physical health, median (IQR)[g]	2 (2-3)	2 (2-3)
Self-perceived health compared with age peers, median (IQR)	2 (1-3)	2 (1-2)

[a]Chronological age (n=309) and subjective age (n=280) are reported for different analytic samples.

[b]Recurrent falls were defined as the self-reported occurrence of 2 or more falls in the 12 months preceding each assessment.

[c]Gait speed was calculated from the 50-foot walk test as distance divided by completion time.

[d]*P*<.001 for changes in physical performance between waves, analyzed using linear mixed-effects models adjusted for baseline age and the time interval between assessments.

[e]STS: sit-to-stand.

[f]The age-difference values represent the mean within-person differences between chronological and subjective age calculated within the n=280 subsample and therefore do not equal the differences between the means reported above.

[g]Self-perceived general and comparative health were measured on a 1 to 5 scale, with higher scores indicating worse perceived health.

Association of Physical Performance Measures at Wave 2 With Recurrent Falls at Wave 3

Among the physical performance measures examined, only gait speed was significantly associated with recurrent falls. Slower gait speed was associated with higher odds of 2 or more falls at follow-up in the total sample (OR 1.43, 95% CI 1.05-1.95; *P*=.02), corresponding to a 43% increase in the odds of recurrent falls per 1 SD decrease in gait speed. In the age-stratified adjusted models, the association was stronger among participants who were 65 years or older at follow-up (OR 1.72, 95% CI: 1.11-2.74; *P*=.02), whereas no significant association was observed among participants younger than 65 years at follow-up (OR 1.15, 95% CI 0.73-1.86; *P*=.55). Figure 2 illustrates the predicted probability of recurrent falls at follow-up according to baseline gait speed in the total sample and age-stratified analyses. In contrast to gait speed, no statistically significant associations were observed between sit-to-stand time, calculated sit-to-stand power, and grip strength and recurrent falls in the total sample or in the age-stratified analyses. Results from the unadjusted and adjusted models are presented in Table 2. In a sensitivity analysis restricted to the 254 participants who were not recurrent fallers at baseline, slower midlife gait speed remained significantly associated with recurrent falls in the total sample (adjusted OR 1.48, 95% CI 1.05-2.13; *P*=.03), and this association was stronger among participants 65 years or older (adjusted OR 1.89, 95% CI 1.16-3.17; *P*=.02). Grip strength also showed a stronger association with a narrower 95% CI, most pronounced among participants 65 years or older, where lower grip strength was significantly associated with becoming a new recurrent faller (adjusted OR 2.13, 95% CI 1.02-4.67; *P*=.049). This finding was not seen in the unadjusted model or the primary, nonrestricted analysis. However, this result did not meet the Bonferroni-corrected threshold (*P*<.016). Results of the sensitivity analysis are summarized in Table S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 2. Probability of recurrent falls (≥2) by baseline gait speed, adjusted for baseline covariates measured at Midlife in the United States (MIDUS) wave 2, including age, sex, BMI, height, baseline fall count, aggregate count of self-reported medical diagnoses and chronic conditions, and the time interval (years) between MIDUS waves.

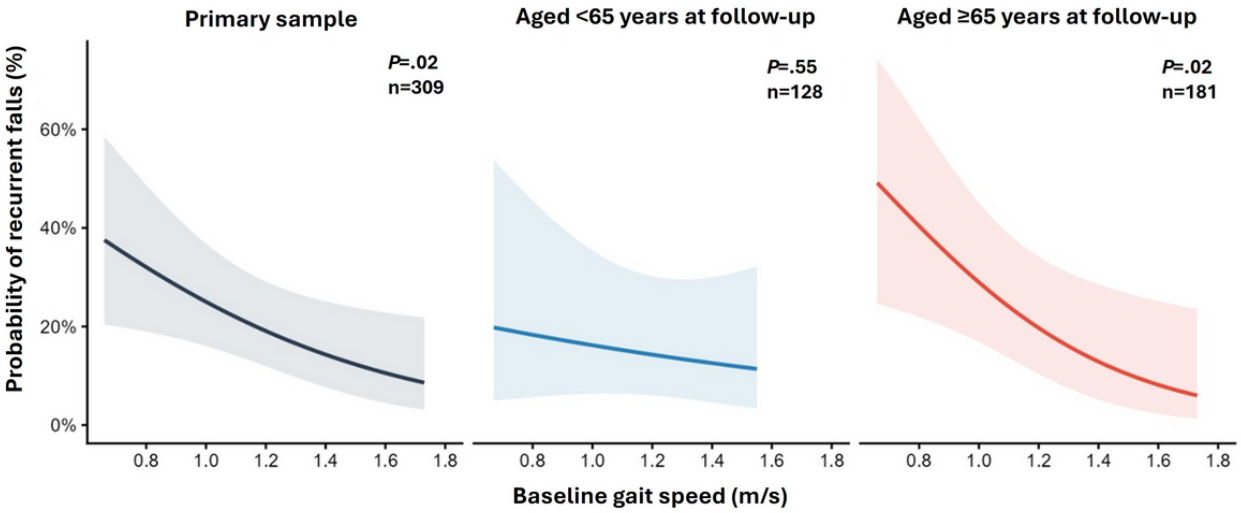


Table 2. Logistic regression models examining the association between baseline physical performance and recurrent falls at follow-up.

Independent variable[a] and total sample/age category at follow-up	Unadjusted OR[b] (95% CI)	*P* value	Adjusted OR (95% CI)[c]	*P* value
Gait speed (m/s)[d]				
Total sample	1.35 (1.02-1.81)	.04	1.43 (1.05-1.95)	.02
<65 y	1.20 (0.79-1.86)	.39	1.15 (0.73-1.86)	.55
≥65 y	1.48 (1.01-2.21)	.047	1.72 (1.11-2.74)	.02
Sit-to-stand time (s)				
Total sample	1.06 (0.80-1.39)	.67	1.10 (0.82-1.47)	.50
<65 y	1.15 (0.75-1.71)	.50	1.05 (0.66-1.64)	.80
≥65 y	0.99 (0.67-1.42)	.96	1.06 (0.69-1.60)	.76
Power during sit-to-stand (W)[d]				
Total sample	1.03 (0.79-1.38)	.82	1.14 (0.78-1.69)	.52
<65 y	0.89 (0.59-1.37)	.61	0.87 (0.50-1.51)	.62
≥65 y	1.16 (0.79-1.76)	.45	1.39 (0.78-2.60)	.28
Grip strength (kg)[d]				
Total sample	1.12 (0.85-1.49)	.43	1.34 (0.87-2.09)	.19
<65 y	1.05 (0.69-1.61)	.83	1.04 (0.54-1.95)	.90
≥65 y	1.19 (0.82-1.77)	.36	1.63 (0.86-3.19)	.14

[a]Age category represents the participant's chronological age at the time of follow-up (MIDUS wave 3).

[b]OR: odds ratio.

[c]Adjusted models controlled for baseline covariates measured at MIDUS wave 2, including age, sex, BMI, height, baseline fall count, aggregate count of self-reported medical diagnoses and chronic conditions, and the time interval (years) between MIDUS waves.

[d]For all physical performance measures except sit-to-stand time, variables were coded so that higher odds ratio indicate poorer performance.

Longitudinal Changes in Gait Speed and Muscle Performance

Gait speed did not show a statistically significant change between MIDUS waves, remaining stable at 1.11 (SD 0.20) m/s at both baseline and follow-up. In contrast, all muscle performance indicators declined over time. Grip strength showed the largest standardized decline, decreasing from 36.40 (SD 12.15) kg at baseline to 30.99 (SD 10.56) kg at follow-up, corresponding to an approximately 15% reduction (β=−0.47, 95% CI −0.53 to −0.40; *P*<.001). Sit-to-stand time increased from 9.24 (SD 3.21) seconds to 10.52 (SD 4.10) seconds, indicating slower task completion and worsening performance (β=0.34, 95% CI 0.21-0.48; *P*<.001). Sit-to-stand power

XSL•FO
RenderX

decreased from 360.55 (SD 163.47) W to 312.96 (SD 141.26) W, indicating reduced lower-extremity power (β=−0.31, 95% CI −0.41 to −0.21; *P*<.001). Together, these findings indicate that muscle performance measures were more sensitive to age-related change across the follow-up period than gait speed.

Association of Physical Performance Measures at Wave 2 With Subjective Age and Self-Rated Health at Wave 3

This analysis included 280 participants. In unadjusted models, slower gait speed was associated with poorer self-rated health on both follow-up health measures. Each SD decrease in gait speed was associated with 32% higher odds of poorer self-perceived physical health in the total sample (OR 1.32, 95% CI 1.05-1.64; *P*=.02) and 39% higher odds among participants younger than 65 years at follow-up (OR 1.39, 95% CI 1.06-1.82; *P*=.01). A similar pattern was observed for health compared with age peers in the total sample (OR 1.49, 95% CI 1.19-1.89; *P*<.001) and among participants younger than 65 years (OR 1.56, 95% CI 1.19-2.08; *P*=.001), whereas no significant associations were observed among participants aged 65 years or older. However, after adjustment for baseline covariates, these associations were no longer statistically significant.

Gait speed was not associated with subjective age in unadjusted analyses (β=−0.03, 95% CI −1.34 to 0.93; *P*=.59). However, in fully adjusted models, faster gait speed was independently associated with a younger subjective age at follow-up, corresponding to an approximately 1-year younger subjective age for every 0.20 m/s increase in gait speed (β=−0.10, 95% CI −0.20 to −0.006; *P*=.04). This association did not reach statistical significance in either age-stratified subgroup.

No significant associations were observed between the other physical performance measures and either self-rated health or subjective age in the total sample or age-stratified analyses.

Discussion

Principal Findings

This longitudinal analysis of the MIDUS cohort examined whether midlife gait speed and muscle performance were associated with recurrent falls later in life. Slower midlife gait speed was associated with a 43% increase in the odds of recurrent falls at follow-up, whereas none of the muscle performance measures showed a significant association with future falls.

The association between gait speed and recurrent falls emerged despite the cohort maintaining high functional status, as indicated by a stable mean walking speed between waves (1.11, SD 0.20 m/s) that remained, on average, above commonly referenced geriatric risk thresholds (0.8-1.0 m/s) [4,26,27], and despite adjustment for fall history at baseline. Other longitudinal cohorts have similarly reported relative stability of gait speed across midlife and early older adulthood, consistent with our findings [36,37]. Although this average gait speed suggests preserved mobility at the group level, it may obscure meaningful between-individual differences in walking performance. Thus, the observed association with future falls likely reflects differences already present during midlife rather than progressive age-related decline, supporting the potential role of slower midlife gait speed as an early marker of future fall vulnerability.

Prior research also indicates that gait speed reflects broader biological and functional processes already present in midlife. For example, Rasmussen et al [19] reported that slower gait speed at age 45 was associated with poorer physical and neurocognitive function. Similarly, findings from the Baltimore Longitudinal Study of Aging demonstrated that multiple motor impairments are linked to slow gait speed in mid-to-late life [28].

Evidence from longitudinal studies in midlife, however, remains limited and somewhat inconsistent. In a cohort of community-dwelling adults aged 50 to 65 years, Segaux et al [14] reported that gait speed was not associated with subsequent adverse events, including nonaccidental falls. Several methodological differences may explain this discrepancy. First, participants in the study by Segaux et al [14] had exceptionally high functional status, with a median gait speed of 1.43 m/s and very few participants meeting criteria for slowness (<1 m/s). Second, the follow-up period in that study was considerably shorter than in the present analysis, which followed participants for more than a decade. Finally, differences in gait assessment may also contribute, as the MIDUS 50-foot walk includes a turning phase that places greater demands on dynamic balance than the straight-path walking tests [38,39] used in the Segaux et al [14] study. Together, these differences may explain the contrasting findings and highlight the need for further longitudinal research focusing on midlife populations to clarify the role of gait speed in predicting later fall risk. It would also be informative for future studies to examine midlife populations stratified into early and late midlife groups.

Consistent with previous studies reporting differences between younger and older middle-aged adults [40,41], our age-stratified analyses showed that the association between midlife gait speed and recurrent falls was primarily evident among participants who were 65 years or older at follow-up, whereas no statistically significant association was observed among those younger than 65 years. Nevertheless, these findings should be interpreted cautiously and require confirmation in larger longitudinal cohorts.

In contrast to gait speed, muscle performance indices, including sit-to-stand time, lower-extremity power, and grip strength, are considered particularly sensitive to age-related changes during midlife [24,30,37,42,43], and declined significantly over follow-up. However, none of these measures were associated with recurrent falls in the present cohort. Falls in community-dwelling adults commonly occur during walking [44], a complex task that requires continuous dynamic balance supported by sensorimotor integration and cognitive processing [19,45]; thus, gait speed likely reflects the integrity of multiple physiological systems involved in maintaining stability. In comparison, isolated aspects of muscular capacity assessed under relatively controlled conditions may be less sensitive to the complex demands underlying walking stability, particularly

XSL•FO
RenderX

during midlife when substantial muscle deconditioning is often absent [46].

Unlike our findings, the China Health and Retirement Longitudinal Study (CHARLS) reported that weak grip strength was associated with increased fall risk over a 4-year follow-up period that included middle-aged adults [22]. However, the CHARLS cohort included participants with conditions such as stroke, rheumatoid arthritis, and kidney disease that independently increase fall risk [22], whereas the present study excluded participants with major neurological and mobility-impairing conditions. In addition, mean midlife grip strength in our cohort was 45 kg for men and 28 kg for women, whereas CHARLS identified weaker grip strength (<30 kg for men and <20 kg for women) as a fall risk factor [22]. These differences in participant characteristics and grip strength distribution likely contributed to the discrepant findings. Therefore, our findings suggest that slower midlife gait speed may reflect early vulnerability to future fall risk, even when conventional measures of muscle performance remain within normal ranges. These findings underscore the need for further longitudinal studies to identify indicators of fall risk before overt functional decline.

A secondary exploratory aim of the present investigation was to examine whether midlife physical performance was associated with subjective age and self-rated health later in life. Faster midlife gait speed was associated with younger subjective age at follow-up after accounting for baseline participant characteristics; however, the estimated magnitude was small, corresponding to approximately a 1-year younger subjective age for a 0.20 m/s increase in gait speed. Associations between gait speed and self-perceived health were observed only before baseline characteristics were accounted for. Although these secondary findings should be interpreted as complementary to the primary fall-related outcomes, they are in line with previous research showing that gait speed correlates with self-reported physical limitations, lower self-rated health, and poorer scores on health-related quality of life instruments such as the 36-Item Short Form Survey [19,47]. Together, these findings suggest that walking performance may relate not only to future fall risk but also to subjective perceptions of health and aging.

Notably, participants in both MIDUS waves generally reported feeling younger than their chronological age, with a mean subjective age approximately 10 years lower, a pattern often interpreted as reflecting psychological resilience [48]. This pattern may also reflect the relatively healthy composition of the analytic sample. Associations between objective gait performance and perceived aging have rarely been examined in such healthy midlife populations, making this a potentially novel observation that warrants further investigation. To obtain a more comprehensive understanding of subjective and objective functioning during midlife, future studies should examine the relationship between objective physical performance and midlife-specific self-perception tools, such as the 16-item Mobility in Middle-Age Questionnaire [49].

Limitations

Several limitations should be considered when interpreting our findings. First, the primary gait speed association should be interpreted in the context of the multiplicity approach used in this study. Although gait speed and muscle performance were defined in advance as separate domains, if a single Bonferroni correction had been applied across all 4 physical performance measures (P <.0125), the association between gait speed and recurrent falls would not have met that threshold. Second, the observational design and reliance on self-reported falls and health conditions may introduce recall bias. Third, although exclusion criteria were applied to restrict the analytic sample to generally independent participants and strengthen internal validity, they also limit generalizability to the broader middle-aged population and reduce the sample size and number of recurrent fall events, particularly in analyses stratified by age category. Because MIDUS is an established closed cohort and the sample size was fixed by the number of eligible participants with complete data at both waves, no a priori power calculation was applicable. Although participants with major neurological conditions were excluded, residual confounding from unreported or undiagnosed health conditions cannot be ruled out and may have influenced performance on the physical tests. In addition, sit-to-stand muscle power was estimated using a validated equation rather than measured directly. Although this approach is established [30], the assumption of a standard chair height may have reduced individual-level precision and potentially obscured subtle associations. Finally, baseline subjective measures were collected during the MIDUS telephone interview, which preceded the biomarker assessment by approximately 2 to 3 years. This temporal gap may have introduced some measurement misalignment between subjective and objective variables, potentially affecting the precision of model adjustment. Despite these limitations, the study benefits from a longitudinal design with over a decade of follow-up and the use of objective physical performance measures within a well-characterized population-based cohort, providing a valuable opportunity to examine early functional markers of fall risk beginning in midlife.

Conclusions

Using data from the MIDUS cohort, this study provides evidence that vulnerability to recurrent falls may be detectable as early as midlife. Slower self-selected walking speed was associated with higher odds of recurrent falls at follow-up, despite stable average gait speed over time. In contrast, muscle performance indices declined with age but were not significantly associated with recurrent falls in the primary analysis. Further longitudinal studies in larger midlife cohorts are warranted to confirm these findings.

Acknowledgments

The authors used ChatGPT (GPT 5.6; OpenAI) to assist with English grammar and phrasing, and with code development for figure generation. All AI-assisted content and code were reviewed and verified by the authors.

XSL•FO
RenderX

Funding

This research did not receive funding.

Data Availability

The data used in this study are publicly available from the Midlife in the United States (MIDUS) study via the Inter-university Consortium for Political and Social Research (ICPSR) repository [50].

Authors' Contributions

Conceptualization: RH, YN, DK, SS
Data curation: RH, YN
Formal analysis: YN
Investigation: RH, YN
Methodology: RH, YN, DK, SS
Project administration: SS
Supervision: SS
Validation: RH, YN
Visualization: DK
Writing—original draft: RH
Writing—review and editing: YN, SS
All authors reviewed and approved the final version of this paper.

Conflicts of Interest

The authors declare no conflict of interest.

Multimedia Appendix 1
STROBE checklist and sensitivity analysis results.
[PDF File (Adobe PDF File), 257 KB - aging_v9i1e103917_app1.pdf]

References

1. Chatterji S, Byles J, Cutler D, Seeman T, Verdes E. Health, functioning, and disability in older adults—present status and future implications. Lancet 2015 Feb 07;385(9967):563-575 [FREE Full text] [doi: 10.1016/s0140-6736(14)61462-8]
2. Brown CJ, Flood KL. Mobility limitation in the older patient: a clinical review. JAMA 2013 Sep 18;310(11):1168-1177. [doi: 10.1001/jama.2013.276566] [Medline: 24045741]
3. Montero-Odasso M, van der Velde N, Martin FC, Petrovic M, Tan MP, Ryg J, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 02;51(9):afac205 [FREE Full text] [doi: 10.1093/ageing/afac205] [Medline: 36178003]
4. Colón-Emeric CS, McDermott CL, Lee DS, Berry SD. Risk assessment and prevention of falls in older community-dwelling adults: a review. JAMA 2024 Apr 23;331(16):1397-1406 [FREE Full text] [doi: 10.1001/jama.2024.1416] [Medline: 38536167]
5. Garnett MF, Weeks JD, Spencer MR. Unintentional fall deaths among adults aged 65 and over: United States, 2020. NCHS Data Brief 2022 Nov(449):1-8 [FREE Full text] [Medline: 36454076]
6. Wu W, Zhang B, Li J, Zhao Y, Wang L, Zhao J, et al. Global and national burden of falls among individuals aged 55 years and older (1990-2021): analysis of the Global Burden of Disease Study 2021. BMC Public Health 2025 Oct 31;25(1):3681 [FREE Full text] [doi: 10.1186/s12889-025-25016-w] [Medline: 41174600]
7. Stolt L, Kolish D, Cardoso M, Tanaka C, Vasconcelos E, Pereira E, et al. Accidental falls in middle-aged women [Article in English, Portuguese]. Rev Saude Publica 2020 Dec 16;54:141 [FREE Full text] [doi: 10.11606/s1518-8787.2020054002579] [Medline: 33331487]
8. Peeters G, van Schoor NM, Cooper R, Tooth L, Kenny RA. Should prevention of falls start earlier? Co-ordinated analyses of harmonised data on falls in middle-aged adults across four population-based cohort studies. PLoS One 2018 Aug 07;13(8):e0201989 [FREE Full text] [doi: 10.1371/journal.pone.0201989] [Medline: 30086149]
9. Burns SD, Ailshire JA, Crimmins EM. Functional limitation among middle age and older adults: exploring cross-national gender disparities. Arch Gerontol Geriatr 2024 Aug;123:105410. [doi: 10.1016/j.archger.2024.105410] [Medline: 38503129]
10. Brennan TH, Lewis LK, Gordon SJ, Prichard I. Effectiveness of interventions to prevent or reverse pre-frailty and frailty in middle-aged community dwelling adults: a systematic review. Prev Med 2024 Aug;185:108008 [FREE Full text] [doi: 10.1016/j.ypmed.2024.108008] [Medline: 38797264]
11. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort profile: the Health and Retirement Study (HRS). Int J Epidemiol 2014 Apr;43(2):576-585 [FREE Full text] [doi: 10.1093/ije/dyu067] [Medline: 24671021]

XSL•FO
RenderX

12. Ryff C, Almeida D, Ayanian JZ, Carr DS, Cleary PD, Coe C, et al. Midlife in the United States (MIDUS 2), 2004-2006 (ICPSR 4652). Inter-university Consortium for Political and Social Research. 2021. URL: https://www.icpsr.umich.edu/web/NACDA/studies/4652/versions/V8 [accessed 2026-09-29]
13. Ryff C, Almeida D, Ayanian J, Binkley N, Carr DS, Coe C, et al. Midlife in the United States (MIDUS 3), 2013-2014 (ICPSR 36346). Inter-university Consortium for Political and Social Research. 2019. URL: https://www.icpsr.umich.edu/web/NACDA/studies/36346/versions/V3 [accessed 2026-03-02]
14. Segaux L, Broussier A, Oubaya N, Leissing-Desprez C, Laurent M, Naga H, et al. Several frailty parameters highly prevalent in middle age (50-65) are independent predictors of adverse events. Sci Rep 2021 Apr 22;11(1):8774 [FREE Full text] [doi: 10.1038/s41598-021-88410-5] [Medline: 33888851]
15. Finegood ED, Briley DA, Turiano NA, Freedman A, South SC, Krueger RF, et al. Association of wealth with longevity in US adults at midlife. JAMA Health Forum 2021 Jul 23;2(7):e211652 [FREE Full text] [doi: 10.1001/jamahealthforum.2021.1652] [Medline: 35977209]
16. Byun M, Kim J, Kim JE. Physical and psychological factors contributing to incidental falls in older adults who perceive themselves as unhealthy: a cross-sectional study. Int J Environ Res Public Health 2021 Apr 02;18(7):3738 [FREE Full text] [doi: 10.3390/ijerph18073738] [Medline: 33918455]
17. Fundenberger H, Stephan Y, Terracciano A, Dupré C, Bongue B, Hupin D, et al. Subjective age and falls in older age: evidence from two longitudinal cohorts. J Gerontol B Psychol Sci Soc Sci 2022 Oct 06;77(10):1814-1819 [FREE Full text] [doi: 10.1093/geronb/gbac094] [Medline: 35861191]
18. Bortone I, Sardone R, Lampignano L, Castellana F, Zupo R, Lozupone M, et al. How gait influences frailty models and health-related outcomes in clinical-based and population-based studies: a systematic review. J Cachexia Sarcopenia Muscle 2021 Apr;12(2):274-297 [FREE Full text] [doi: 10.1002/jcsm.12667] [Medline: 33590975]
19. Rasmussen LJ, Caspi A, Ambler A, Broadbent JM, Cohen HJ, d'Arbeloff T, et al. Association of neurocognitive and physical function with gait speed in midlife. JAMA Netw Open 2019 Oct 02;2(10):e1913123 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.13123] [Medline: 31603488]
20. Buckinx F, Aubertin-Leheudre M. Relevance to assess and preserve muscle strength in aging field. Prog Neuropsychopharmacol Biol Psychiatry 2019 Aug 30;94:109663 [FREE Full text] [doi: 10.1016/j.pnpbp.2019.109663] [Medline: 31176745]
21. Steiber N. Strong or weak handgrip? Normative reference values for the German population across the life course stratified by sex, age, and body height. PLoS One 2016 Oct 04;11(10):e0163917 [FREE Full text] [doi: 10.1371/journal.pone.0163917] [Medline: 27701433]
22. Liu H, Hou Y, Li H, Lin J. Influencing factors of weak grip strength and fall: a study based on the China Health and Retirement Longitudinal Study (CHARLS). BMC Public Health 2022 Dec 13;22(1):2337 [FREE Full text] [doi: 10.1186/s12889-022-14753-x] [Medline: 36514090]
23. Middleton A, Fritz SL, Lusardi M. Walking speed: the functional vital sign. J Aging Phys Act 2015 Apr;23(2):314-322 [FREE Full text] [doi: 10.1123/japa.2013-0236] [Medline: 24812254]
24. Coelho-Júnior HJ, Álvarez-Bustos A, Landi F, da Silva Aguiar S, Rodriguez-Mañas L, Marzetti E. Why are we not exploring the potential of lower limb muscle power to identify people with sarcopenia? Ageing Res Rev 2025 Feb;104:102662 [FREE Full text] [doi: 10.1016/j.arr.2025.102662] [Medline: 39818236]
25. Pham T, McNeil JJ, Barker AL, Orchard SG, Newman AB, Robb C, et al. Longitudinal association between handgrip strength, gait speed and risk of serious falls in a community-dwelling older population. PLoS One 2023 May 08;18(5):e0285530 [FREE Full text] [doi: 10.1371/journal.pone.0285530] [Medline: 37155689]
26. Montero-Odasso M, Schapira M, Soriano ER, Varela M, Kaplan R, Camera L, et al. Gait velocity as a single predictor of adverse events in healthy seniors aged 75 years and older. J Gerontol A Biol Sci Med Sci 2005 Oct;60(10):1304-1309 [FREE Full text] [doi: 10.1093/gerona/60.10.1304] [Medline: 16282564]
27. Abellan van Kan G, Rolland Y, Andrieu S, Bauer J, Beauchet O, Bonnefoy M, et al. Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and Aging (IANA) task force. J Nutr Health Aging 2009 Dec;13(10):881-889 [FREE Full text] [doi: 10.1007/s12603-009-0246-z] [Medline: 19924348]
28. Cai Y, Tian Q, Gross AL, Wang H, E JY, Agrawal Y, et al. Motor and physical function impairments as contributors to slow gait speed and mobility difficulty in middle-aged and older adults. J Gerontol A Biol Sci Med Sci 2022 Aug 12;77(8):1620-1628 [FREE Full text] [doi: 10.1093/gerona/glac001] [Medline: 34984437]
29. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349 [FREE Full text] [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
30. Alcazar J, Losa-Reyna J, Rodriguez-Lopez C, Alfaro-Acha A, Rodriguez-Mañas L, Ara I, et al. The sit-to-stand muscle power test: an easy, inexpensive and portable procedure to assess muscle power in older people. Exp Gerontol 2018 Oct 02;112:38-43 [FREE Full text] [doi: 10.1016/j.exger.2018.08.006] [Medline: 30179662]
31. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge; 1988.

XSL•FO
RenderX

32. Nieminen P. Application of standardized regression coefficient in meta-analysis. BioMedInformatics 2022 Aug 31;2(3):434-458 [FREE Full text] [doi: 10.3390/biomedinformatics2030028]
33. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2023. URL: https://www.R-project.org [accessed 2026-09-29]
34. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. J Open Source Softw 2019;4(43):1686 [FREE Full text] [doi: 10.21105/joss.01686]
35. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015;67(1):1-48 [FREE Full text] [doi: 10.18637/jss.v067.i01]
36. Chan LL, Lord SR, Brodie MA. Daily-life walking speed, quality and quantity derived from a wrist motion sensor: large-scale normative data for middle-aged and older adults. Sensors (Basel) 2024 Aug 10;24(16):5159 [FREE Full text] [doi: 10.3390/s24165159] [Medline: 39204870]
37. Mayhew AJ, So HY, Ma J, Beauchamp MK, Griffith LE, Kuspinar A, et al. Normative values for grip strength, gait speed, timed up and go, single leg balance, and chair rise derived from the Canadian Longitudinal Study on Ageing. Age Ageing 2023 Apr 01;52(4):afad054 [FREE Full text] [doi: 10.1093/ageing/afad054] [Medline: 37078755]
38. Gulley E, Ayers E, Verghese J. A comparison of turn and straight walking phases as predictors of incident falls. Gait Posture 2020 Jun;79:239-243 [FREE Full text] [doi: 10.1016/j.gaitpost.2020.05.002] [Medline: 32450510]
39. Adusumilli G, Lancia S, Levasseur VA, Amblee V, Orchard M, Wagner JM, et al. Turning is an important marker of balance confidence and walking limitation in persons with multiple sclerosis. PLoS One 2018;13(6):e0198178 [FREE Full text] [doi: 10.1371/journal.pone.0198178] [Medline: 29879144]
40. Baranes G, Hayek R, Gutman I, Frenkel-Toledo S, Springer S. Use of the Brief-BESTest partially instrumented with accelerometry to detect balance deterioration in middle-age. Aging Clin Exp Res 2024 Nov 09;36(1):214 [FREE Full text] [doi: 10.1007/s40520-024-02868-7] [Medline: 39520609]
41. Hayek R, Brown RT, Gutman I, Baranes G, Springer S. Smartphone-based analysis for early detection of aging impact on gait and stair negotiation: a cross-sectional study. Sensors (Basel) 2025 Apr 05;25(7):2310 [FREE Full text] [doi: 10.3390/s25072310] [Medline: 40218822]
42. Alcazar J, Aagaard P, Haddock B, Kamper RS, Hansen SK, Prescott E, et al. Assessment of functional sit-to-stand muscle power: cross-sectional trajectories across the lifespan. Exp Gerontol 2021 Sep;152:111448 [FREE Full text] [doi: 10.1016/j.exger.2021.111448] [Medline: 34118352]
43. Grgic J, Wazny V, Maier AB, Schoenfeld BJ, Pedisic Z. Reference values for handgrip strength in Europe: analysis of individual participant data from 27 countries. Geroscience 2026 Jun;48(3):4349-4371 [FREE Full text] [doi: 10.1007/s11357-025-01919-9] [Medline: 41034601]
44. Cai Y, Leveille SG, Andreeva O, Shi L, Chen P, You T. Characterizing fall circumstances in community-dwelling older adults: a mixed methods approach. J Gerontol A Biol Sci Med Sci 2023 Aug 27;78(9):1683-1691 [FREE Full text] [doi: 10.1093/gerona/glad130] [Medline: 37210687]
45. Britting S, Krumpoch S, Lindemann U, Müller A, Rohleder N, Ellmers TJ, et al. The interaction of gait characteristics and concerns about falling in community-dwelling older adults. Sci Rep 2025 Dec 12;15(1):43815 [FREE Full text] [doi: 10.1038/s41598-025-32000-2] [Medline: 41387511]
46. Stel V, Smit JH, Pluijm SM, Lips P. Balance and mobility performance as treatable risk factors for recurrent falling in older persons. J Clin Epidemiol 2003 Jul;56(7):659-668 [FREE Full text] [doi: 10.1016/s0895-4356(03)00082-9] [Medline: 12921935]
47. De la Cámara M, Higueras-Fresnillo S, Sadarangani KP, Esteban-Cornejo I, Martinez-Gomez D, Veiga Ó. Clinical and ambulatory gait speed in older adults: associations with several physical, mental, and cognitive health outcomes. Phys Ther 2020 Apr 17;100(4):718-727 [FREE Full text] [doi: 10.1093/ptj/pzz186] [Medline: 31899490]
48. Pinquart M, Wahl HW. Subjective age from childhood to advanced old age: a meta-analysis. Psychol Aging 2021 May;36(3):394-406 [FREE Full text] [doi: 10.1037/pag0000600] [Medline: 33829847]
49. Hayek R, Karvonen-Gutierrez CA, Bean JF, Guralnik JM, Covinsky K, Hoffman JR, et al. Development and validation of the Mobility in Middle-Age Questionnaire. J Gerontol A Biol Sci Med Sci 2025 Nov 06;80(12):glaf205 [FREE Full text] [doi: 10.1093/gerona/glaf205] [Medline: 41017642]
50. Inter-university Consortium for Political and Social Research (ICPSR) repository. URL: https://www.icpsr.umich.edu/sites/icpsr/home [accessed 2026-09-29]

Abbreviations

CHARLS: China Health and Retirement Longitudinal Study
MIDUS: Midlife in the United States
OR: odds ratio
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by KJ Moon; submitted 07.Jun.2026; peer-reviewed by M Ershadmanesh, J Lee; comments to author 21.Jul.2026; revised version received 06.Sep.2026; accepted 07.Sep.2026; published 06.Oct.2026.

Please cite as:
Hayek R, Nudelman Y, Keren D, Springer S
Associations of Midlife Gait Speed and Muscle Performance With Falls in Later Life: Secondary Analysis of the MIDUS Cohort
JMIR Aging 2026;9:e103917
URL: https://aging.jmir.org/2026/1/e103917
doi:10.2196/103917
PMID:

©Roee Hayek, Yaniv Nudelman, Dan Keren, Shmuel Springer. Originally published in JMIR Aging (https://aging.jmir.org), 06.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

An Ultra-Brief Informant Questionnaire for Case Finding of Cognitive Impairment Across Diverse Literacy: Diagnostic Accuracy Study

Tau Ming Liew[1,2,3,4], FRCPsych, PhD; King Fan Yip[5], MBBS, MRCP; Kaavya Narasimhalu[6], MD, PhD; Simon Kang Seng Ting[6], MBBS, FRCP; Weishan Li[6], MD, MRCP; Sze Yan Tay[7], MPsych; Way Inn Koay[7], MPsych

[1]Department of Psychiatry, Singapore General Hospital, Outram Road, Singapore, Singapore

[2]SingHealth Duke-NUS Medicine Academic Clinical Programme, Duke-NUS Medical School, Singapore, Singapore

[3]Health Services and Systems Research, Duke-NUS Medical School, Singapore, Singapore

[4]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

[5]Department of Geriatric Medicine, Singapore General Hospital, Singapore, Singapore

[6]Department of Neurology, National Neuroscience Institute, Singapore General Hospital, Singapore, Singapore

[7]Department of Psychology, Singapore General Hospital, Singapore, Singapore

Corresponding Author:
Tau Ming Liew, FRCPsych, PhD
Department of Psychiatry, Singapore General Hospital, Outram Road, Singapore, Singapore

Abstract

Background: Undiagnosed cognitive impairment poses a global challenge, prompting recent interest in ultra-brief screening questionnaires (comprising <2 to 3 items) to efficiently identify individuals needing further evaluation. However, evidence on ultra-brief questionnaires remains limited, particularly regarding their validity across diverse literacy levels.

Objective: This study aimed to develop an ultra-brief questionnaire that performs well in detecting mild cognitive impairment or dementia (MCI/dementia) across diverse literacy levels and to compare its performance with an established questionnaire (the 8-item Informant Interview to Differentiate Aging and Dementia [AD8]).

Methods: This diagnostic study involved 1856 participants aged ≥65 years (median education 10 y, range 0 - 23 y), prospectively recruited from community settings in Singapore. Participants and informants completed 15 cognition-related questions. MCI/dementia was diagnosed via a comprehensive assessment and consensus conference. The sample was randomly split 70/30—the training sample (70%) was used to derive an ultra-brief questionnaire from the 15 cognition-related questions (using an exhaustive search approach), and the test sample (30%) evaluated the area under the receiver operating characteristic curve (AUC).

Results: The new questionnaire comprised 2 informant questions (ie, *assistance with medications* and *worry about cognition*), plus age and years of education. It demonstrated excellent performance in detecting MCI/dementia (AUC 85%, 95% CI 80% - 90%), significantly better (P=.003) than a nested baseline model (comprising age and years of education; AUC 78%, 95% CI 73% - 83%). In contrast, the AD8 had an AUC of 76% (95% CI 70% - 83%), not significantly different (P>.99) from the baseline model. The questionnaire's performance was consistent across education subgroups and varying prevalence scenarios. Two optimal cutoffs were used—the lower cutoff provided 80% sensitivity and 96% negative predictive value, and the upper cutoff provided 99% specificity and 81% positive predictive value. A web-based calculator is available for public use.

Conclusions: This ultra-brief questionnaire enables rapid screening for cognitive impairment (in <1 min) by family members or as part of community geriatric assessments. Its excellent performance across literacy levels supports its utility for case finding in diverse populations, including underserved communities and lower- and middle-income countries.

(JMIR Aging 2026;9:e72963) doi:10.2196/72963

KEYWORDS

machine learning; Informant Questionnaire for cognitive impairment–two items plus demographics; IQ2+; informant questionnaire; subjective questionnaire; subjective cognitive decline; neurocognitive disorders

XSL•FO
RenderX

Introduction

Undiagnosed cognitive impairment is a global challenge [1], with 60% to 90% of affected individuals never receiving a formal diagnosis [2,3]. Those who remain undiagnosed miss out on timely clinical care [4], including management of reversible causes, prescription of cognitive enhancers, behavioral management, caregiver support, and advanced care planning [4-9]. All these can impact well-being [10,11] and increase the risk of premature nursing home placement [12-14]. Furthermore, undiagnosed individuals often do not receive adequate support to manage and coordinate care for their chronic diseases [15,16], resulting in suboptimal disease management, inappropriate health care utilization, and higher health care costs [17,18]. As an example, a modeling study estimated that timely management of cognitive impairment could potentially yield annual cost savings of US $13 to $41 billion in the United States [15]. Recently, the importance of early diagnosis has been heightened by emerging evidence supporting early interventions for cognitive impairment [19,20], such as risk factor modification [21,22] and antiamyloid monoclonal antibodies [23,24].

To address the challenge of undiagnosed cognitive impairment, various international bodies (eg, the Alzheimer's Disease International [25] and the International Association of Gerontology and Geriatrics) [26] have advocated for a systematic approach to case finding among high-risk individuals in the community [26]. In particular, a 2-stage strategy [27-29] has been proposed in recent literature to address resource constraints for community case finding. In the first stage, subjective reports (eg, Functional Activities Questionnaire and AD8 [the 8-item Informant Interview to Differentiate Aging and Dementia]) [30,31] are used to identify individuals with potential cognitive impairment. In the second stage, these individuals undergo brief cognitive tests (eg, Mini-Cog and short variants of Montreal Cognitive Assessment) [4,32-34] to confirm the presence of cognitive impairment. This 2-stage strategy is efficient and scalable—the first stage relies solely on subjective reports and does not draw on scarce health care resources for administration of cognitive tests, whereas the second stage reserves brief cognitive tests for a smaller subset of individuals [27]. Moreover, combining subjective reports and brief cognitive tests has been shown to improve the detection of subtle cognitive changes [27], making this approach optimal for case finding of early cognitive impairment.

In 2019, the 2-stage strategy was adopted by the World Health Organization (WHO) within the Integrated Care for Older People assessment tool to identify cognitive impairment, alongside assessments of other key components of intrinsic capacity (ie, mobility, nutrition, hearing, vision, and mood) [29]. For the first stage (subjective report), the WHO adopted an ultra-brief questionnaire based on a single question: "Do you have problems with memory or orientation (such as not knowing where one is or what day it is)?" [29] The decision to embed an ultra-brief cognitive questionnaire within Integrated Care for Older People is understandable, as it balances the need to assess a wide range of intrinsic capacity domains against the scarcity of community resources for comprehensive assessments. However, the validity of such ultra-brief questionnaires (ie, those comprising fewer than 2 - 3 items) remains unclear in the literature [28], especially when used across diverse levels of literacy [26,35]. This concern can be critical, as questionnaires with fewer items tend to have increased measurement variability and may be more susceptible to confounding factors such as educational attainment [36].

In this study, we sought to strengthen the evidence base supporting the use of ultra-brief questionnaires across diverse levels of literacy, potentially enhancing their utility in diverse populations across lower-, middle-, and higher-income countries. Specifically, we aimed to: (1) derive an ultra-brief questionnaire with high performance for detecting cognitive impairment (ie, the presence of mild cognitive impairment or dementia [MCI/dementia]) across diverse levels of literacy, using a contemporary, computationally intensive approach to identify the questions most discriminative of MCI/dementia; and (2) compare the performance of the new ultra-brief questionnaire to the well-established AD8 across participants with lower and higher educational attainment. We selected the AD8 as the benchmark because it is a widely used and extensively validated informant-based questionnaire, as demonstrated in recent systematic reviews [37,38]. Its use has been recommended by various international bodies, including the International Association of Gerontology and Geriatrics [26], the Gerontological Society of America [39], the US Alzheimer's Association [40,41], and the National Institute on Aging workgroup [42].

Of note, this study was conducted in Singapore, a city-state in Southeast Asia that provides a unique testbed of literacy diversity for developing the ultra-brief questionnaire. The current generation of older Singaporeans witnessed the country's transformation from a traditional, lower-income, Asian society to a more westernized, higher-income country [43]. Consequently, this cohort encompasses a wide range of educational backgrounds, from minimal formal education to tertiary education. By validating the ultra-brief questionnaire in such a heterogeneous population, we sought to demonstrate its potential for broader implementation in other literacy-diverse settings beyond Singapore, including populations across East and South Asia, and potentially, in some lower- and middle-income countries (LMICs).

Methods

Study Procedures

This study involved community-dwelling older persons recruited between March 2022 and September 2024, as part of a nationally funded project in Singapore aimed at developing artificial intelligence tools to detect early cognitive impairment in the community (Project PENSIEVE) [44]. Community-dwelling individuals were invited to participate if they met the following criteria: (1) higher risk of cognitive impairment (ie, aged ≥65 y [26]) and had at least one of the 3 chronic diseases (ie, diabetes mellitus, hypertension, or hyperlipidemia); this criterion was included a priori to focus on individuals with at least some risk of cognitive impairment, in line with current literature suggesting the limited benefit of screening among low-risk individuals) [26,45]; (2) ability to follow simple instructions in

XSL•FO
RenderX

English or Mandarin (due to limitations in the available assessment language); and (3) presence of an informant (eg, family member or friend) who knew the participants well. Individuals were excluded if they had severe visual impairment that would affect their ability to complete neuropsychological assessments (to ensure generalizability, participants were included as long as they could see pictures on a piece of paper held in front of them). No participants were excluded for reasons related to missing data, as we implemented strict data collection procedures (eg, mandatory data field) and routine data audits throughout the study.

Sources of recruitment included 14 community roadshows conducted by the study team, clients of our community partners, home visits by community volunteers who partnered with us, media publicity (radio, online articles, and posters), and word-of-mouth referrals from participants who had completed research assessments. To ensure the recruited samples were representative of the community, the study's publicity materials emphasized the key message of "detect dementia early" (along with direct referrals to memory clinics in the event of significant findings), rather than the conventional invitation to participate in research (which may inadvertently attract a distinct group of individuals). Examples of these publicity materials (eg, study banner, poster, and brochure) are provided in Method S1 in Multimedia Appendix 1.

All participants received comprehensive assessments, which included semistructured interviews with both participants and their informants, detailed neuropsychological testing, and observational notes of participants' behavior during assessments. Full descriptions of the comprehensive assessments are available in Method S2 in Multimedia Appendix 1, with further details on each assessment tool provided in Method S3 in Multimedia Appendix 1. Diagnoses of MCI and dementia were determined via consensus conference by 3 dementia specialists. Dementia was diagnosed using the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, criteria [46], which require the presence of cognitive concerns (reported by the individual or a knowledgeable informant), impairment in instrumental activities of daily living (eg, managing money or medications), and objective cognitive deficits. MCI was diagnosed using the modified Petersen criteria [47], which require the presence of cognitive concerns (reported by the individual or a knowledgeable informant), absence of impairment in instrumental activities of daily living, and the presence of objective cognitive deficits. Normal cognition was diagnosed when participants did not meet criteria for dementia or MCI.

Measures

AD8 [30] is an 8-item, informant-based questionnaire that assesses changes in a participant's cognition and function over the past few years. For each item, the informant rates whether there has been a change in the participant: 1=yes (a change) and 0=no (no change) or don't know. Responses to the 8 items are summed to provide a total score, with higher scores reflecting greater cognitive problems. AD8 has been shown to be useful in detecting varying severities of cognitive impairment [37,38]. Informant AD8 has also been previously validated in Singapore [48].

To derive an ultra-brief questionnaire, the study team focused on candidate questions that assess the two key criteria for diagnosing MCI/dementia [46,47]: (1) the presence of cognitive concerns as reported by the individual or a knowledgeable informant; and (2) impairment in instrumental activities of daily living (iADL). The presence of cognitive concerns was evaluated using validated questions related to subjective cognitive decline (SCD), by asking participants or informants: "Do you feel like your (or your family member's) memory is becoming worse?" and "Are you *worried* that your (or your family member's) memory is becoming worse?" These 2 sets of questions have been validated in previous studies [49,50] and have been shown to be useful for capturing early symptoms of cognitive decline [51-55].

Impairment in iADL was assessed using the locally validated modified Lawton scale [56], with informants asked about difficulties in various domains (public commuting, grocery shopping, managing money, using the telephone, taking medications, preparing meals, doing housework, and doing laundry). The original responses on the modified Lawton scale included 4 options (ie, unable to do at all, needs some help, needs no help, and never needed to do this), which were collapsed in this study into two options: (1) yes and (2) no/never needed to do this.

In total, 15 candidate questions were considered for the ultra-brief questionnaire: 12 cognition-related items and 3 basic demographic variables (eg, age, sex, and years of education), given their potential correlation with cognition. The exact wording and response options for all 15 items are provided in Table 1.

XSL•FO
RenderX

Table . A preselected list of 15 question items that are potentially related to cognition.

Question category [a]	Question item [a]	Response options
SCD[b] question—informant	Do you feel like your family member's memory is becoming worse?	1=yes; 0=no/not sure
Worry about cognition—informant	Are you *worried* that your family member's memory is becoming worse?	1=yes; 0=no/not sure
SCD question—participant	Do you feel like your memory is becoming worse?	1=yes; 0=no/not sure
Worry about cognition—participant	Are you *worried* that your memory is becoming worse?	1=yes; 0=no/not sure
iADL[c]—commute	Does your family member need help to take public transport or drive a car?	1=yes; 0=no/never needed to do this
iADL—grocery	Does your family member need help to do grocery shopping?	1=yes; 0=no/never needed to do this
iADL—money	Does your family member need help to manage money?	1=yes; 0=no/never needed to do this
iADL—telephone	Does your family member need help to use the telephone?	1=yes; 0=no/never needed to do this
iADL—medications	Does your family member need help to take medications?	1=yes; 0=no/never needed to do this
iADL—meals	Does your family member need help to prepare meals?	1=yes; 0=no/never needed to do this
iADL—housework	Does your family member need help to do housework?	1=yes; 0=no/never needed to do this
iADL—laundry	Does your family member need help to do laundry?	1=yes; 0=no/never needed to do this
Age	What is your family member's age?	Continuous variable
Sex	What is your family member's sex?	1=male; 0=female
Years of education	What is your family member's years of education?	Continuous variable. Count the years of full-time education, starting from elementary/primary school.

[a]Items in the list were selected to assess the 2 key criteria in the diagnosis of mild cognitive impairment and dementia: (1) the presence of cognitive concerns as reported by the individual or a knowledgeable informant and (2) impairment in iADL [46,47]. The presence of cognitive concerns was evaluated using validated questions related to SCD, given prior literature on the usefulness of SCD to reflect early symptoms of cognitive decline [51-55]. Impairment in iADL was assessed using the locally validated modified Lawton scale [56], with informants asked about difficulties in various domains of iADL. Three basic demographic variables (ie, age, sex, and years of education) were also included, given their potential correlation with cognition.

[b]SCD: subjective cognitive decline.

[c]iADL: instrumental activities of daily living.

Statistical Analyses

The study sample was randomly split into a 70% *training sample* and a 30% *test sample*. The *training sample* was used to develop an ultra-brief questionnaire that best distinguished MCI/dementia from normal cognition, whereas the *test sample* was used to evaluate the actual performance of this questionnaire.

In the *training sample*, a best-subset approach [57] with 5-fold cross-validation was used to select the optimal combination of items from the 15 candidate questions. The best-subset approach is an efficient, computationally intensive method for variable selection [4,20,58,59] in which logistic regression is used to exhaustively evaluate all possible combinations of the candidate questions, identifying models with the lowest prediction errors. Five-fold cross-validation was then used to select the most parsimonious model within 1 SE of the best-performing model. After identifying the optimal model from the best-subset approach, we further refined the model by considering potential inclusions of quadratic terms for continuous variables (eg, age and years of education); quadratic terms with $P<.05$ were incorporated into the final model to improve fit. The final selected model was then applied to the test sample to generate predicted probabilities of cognitive impairment (ie, MCI/dementia), with the model variables constituting the new ultra-brief questionnaire.

Predicted probabilities were computed from logistic regression using the following equation:

Predicted probability=eLogit1+eLogit

XSL•FO
RenderX

where

$$\text{logit} = \beta 0 + \beta 1 \cdot (\text{Variable } 1) + \beta 2 \cdot (\text{Variable } 2) + \beta 3 \cdot (\text{Variable } 3) + \ldots$$

with each β representing the regression coefficient for its respective variable in the model.

In the *test sample*, the predicted probabilities were used to compute the area under the receiver operating characteristic curve (AUC), thereby assessing the actual performance of the ultra-brief questionnaire in discriminating MCI/dementia from normal cognition. In general, an AUC of 0.7 to 0.8 is considered acceptable discrimination, 0.8 to 0.9 is considered excellent, and more than 0.9 is considered outstanding [60]. Comparisons of AUC were conducted using the nonparametric method proposed by DeLong et al [33,59,61,62], with analyses stratified by education subgroups (ie, ≤10 and >10 y of education based on median split). A 2-cutoff approach [63-67] was adopted for the ultra-brief questionnaire. The first cutoff was chosen to yield high sensitivity and negative predictive value (>80% each) and was used to rule out MCI/dementia (ie, when probability scores fell below the first threshold). The second cutoff was selected for high specificity and positive predictive value (>80% each), identifying those very likely to have MCI/dementia. This 2-cutoff approach is recommended in recent literature [67], as it enhances the performance of cognitive assessment tools [63-66], reduces the effects of prevalence on tool performance [64], and allows prioritization of scarce health care resources for individuals who truly require further cognitive assessments [63].

As a secondary analysis, the performance of the ultra-brief questionnaire was also evaluated for distinguishing dementia from nondementia. Additionally, 2 sensitivity analyses were conducted in the test sample to evaluate the robustness of results when the prevalence of MCI/dementia was readjusted to reflect the average prevalence in most communities:

- Prevalence of MCI/dementia was artificially readjusted to 20% based on prior meta-analytic findings that community prevalence is ~15% for MCI [68-70] and ~5% for dementia [71-73]. Readjustment of prevalence was done by randomly selecting a subset of participants with MCI and normal cognition—for each participant with dementia, 3 participants with MCI and 16 participants with normal cognition were randomly selected (ie, so that the final dataset corresponded to 5% prevalence for dementia and 15% prevalence for MCI).
- Prevalence of MCI/dementia was artificially readjusted to 35% based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI [69,70] and ~10% for dementia [72,73]. Readjustment of prevalence was done by randomly selecting only a subset of participants with MCI and normal cognition—for each participant with dementia, 2.5 participants with MCI and 6.5 participants with normal cognition were randomly selected (ie, so that the final dataset corresponded to 10% prevalence for dementia and 25% prevalence for MCI).

The best-subset approach was performed with the "bestglm" [57] package in R (version 4.4.0; R Foundation for Statistical Computing). All remaining analyses were conducted in Stata (version 18; StataCorp LLC). No a priori sample size calculation was performed; the final sample size was determined pragmatically based on participants recruited between March 2022 and September 2024. Post hoc power analyses confirmed that the test sample provided robust power (90%) to distinguish MCI/dementia from normal cognition (α=.05, two-sided test). Power was also reasonably sufficient for participants with ≤10 years of education (72%) but was limited in those with >10 years of education (19%) due to fewer positive cases. The power calculations were conducted using PASS software (version 15.0.5; NCSS, LLC) and the Hanley and McNeil formula [74], with further details provided in MethodS4 in Multimedia Appendix 1.

Ethical Considerations

The study received ethical approval from the SingHealth Centralized IRB of Singapore (reference number: 2021/2590). Informed consent was obtained from all participants. Before obtaining informed consent, the mental capacity of participants was briefly assessed in accordance with the Mental Capacity Act of Singapore [75]. In the event of concerns about mental capacity, informed consent by proxy was then obtained from the legally authorized next-of-kin. Participants who completed the research assessments received Singapore Dollar $80 as compensation for their time, inconvenience and transportation costs.

Results

Overview

A total of 1856 participants were included, of whom 255 (13.7%) had MCI/dementia. Participant characteristics are presented in Table 2, with a median age of 72 years and a median education of 10 years (range 0-23 y). Corresponding informant characteristics are provided in Table S1 in Multimedia Appendix 1, with informants primarily comprising spouses (897/1856, 48.3%) and children (506/1856, 27.3%). The sample was randomly split into a *training sample* (1299/1856, 70%) and a *test sample* (557/1856, 30%). Table S2 in Multimedia Appendix 1 shows that the training and test samples had comparable demographic characteristics (*P*>.05 across covariates).

XSL•FO
RenderX

Table . Characteristics of the study participants.

Variable	Overall sample (n=1856)	Normal cognition (n=1601)	MCI[a] (n=207)	Dementia (n=48)	*P* value [b]
Age, median (IQR) [range]	72 (68 to 76) [65 to 101]	71 (68 to 75) [65 to 93]	74 (70 to 79) [65 to 91]	80 (76 to 82) [66 to 101]	<.001
Years of education, median (IQR) [range]	10 (9 to 13) [0 to 23]	10 (10 to 13) [0 to 23]	10 (6 to 12) [0 to 21.5]	10 (2 to 10) [0 to 17]	<.001
Sex (male), n (%)	688 (37.1)	572 (35.7)	101 (48.8)	15 (31.2)	<.001
Ethnicity, n (%)					.39
Chinese	1735 (93.5)	1504 (93.9)	188 (90.8)	43 (89.6)	
Malay/Indian	96 (5.2)	77 (4.8)	15 (7.2)	4 (8.3)	
Eurasian/others	25 (1.3)	20 (1.2)	4 (1.9)	1 (2.1)	
MoCA[c] total score, median (IQR)	26 (24 to 28)	27 (25 to 28)	21 (17 to 24)	14 (9 to 19)	<.001
NTB[d] Global Z-scores, median (IQR)	−0.2 (−0.6 to 0.1)	−0.1 (−0.4 to 0.2)	−1.0 (−1.3 to 0.7)	−1.6 (−2.1 to 1.2)	<.001
Global CDR[e], n (%)					<.001
0	1570 (84.6)	1557 (97.3)	13 (6.3)	0 (0.0)	
0.5	255 (13.7)	44 (2.7)	194 (93.7)	17 (35.4)	
1	22 (1.2)	0 (0.0)	0 (0.0)	22 (45.8)	
2	8 (0.4)	0 (0.0)	0 (0.0)	8 (16.7)	
3	1 (0.1)	0 (0.0)	0 (0.0)	1 (2.1)	

[a]MCI: mild cognitive impairment.

[b]Test of difference across diagnoses: chi-square test for categorical variables and Kruskal-Wallis test for continuous variables.

[c]MoCA: Montreal Cognitive Assessment.

[d]NTB: neuropsychological battery.

[e]CDR: clinical dementia rating.

Development of the Ultra-Brief Questionnaire

Table 3 presents the top models identified through the exhaustive search method in the *training sample* (n=1299). *iADL–medications* and *worry about cognition–informant* were among the most useful items in detecting MCI/dementia, whereas *iADL–meals* and *iADL–grocery* were among the least useful. Following 5-fold cross-validation, the model with 4 items was identified as the most parsimonious among the top models (Figure S1 in Multimedia Appendix 1). This model included 2 informant questions (*iADL–medications* and *worry about cognition–informant*), plus 2 demographic variables (*age* and *years of education*). These items were then selected to constitute the new ultra-brief questionnaire, henceforth denoted as IQ2+ (Informant Questionnaire for cognitive impairment–2 items plus demographics). Further model refinement considered the inclusion of quadratic terms for *age* and *years of education*, resulting in the addition of a quadratic term for years of education (*P*=.004) in the final model. Thus, the final IQ2+ model comprised five predictors: iADL–medications, worry about cognition–informant, age, years of education, and the quadratic term for years of education. Responses from the IQ2+ questionnaire can be converted to predicted probabilities using an interactive web-based calculator we have created [76]. A screenshot of the calculator is shown in Figure 1.

XSL•FO
RenderX

Table . The top models that best discriminate mild cognitive impairment and dementia from normal cognition (as identified by the best-subset approach) in the training sample (n=1299).

Question items, rearranged by their usefulness in detecting MCI[a]/dementia[b]	Number of items in the top models														
	1	2	3	4[c]	5	6	7	8	9	10	11	12	13	14	15
iADL—medications[d]	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Worry about cognition—informant	N/A[e]	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Age	N/A	N/A	●	●	●	●	●	●	●	●	●	●	●	●	●
Years of education	N/A	N/A	N/A	●	●	●	●	●	●	●	●	●	●	●	●
Sex	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●	●	●	●	●
iADL—money	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●	●	●	●
SCD[f] question—informant	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●	●	●
iADL—commute	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●	●
iADL—telephone	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●
Worry about cognition—participant	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●
iADL—laundry	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●
SCD question—participant	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●
iADL—housework	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●
iADL—grocery	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●
iADL—meals	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●

XSL•FO
RenderX

[a]MCI/dementia: mild cognitive impairment or dementia.

[b]A full description of each item is available in Table 1. Briefly, items in the list were selected to assess the 2 key criteria in the diagnosis of mild cognitive impairment and dementia: (1) the presence of cognitive concerns as reported by the individual or a knowledgeable informant and (2) impairment in iADL [46,47]. The presence of cognitive concerns was evaluated using validated questions related to SCD, given prior literature on the usefulness of SCD to reflect early symptoms of cognitive decline [51-55]. Impairment in iADL was assessed using the locally validated modified Lawton scale [56], with informants asked about difficulties in various domains of iADL. Three basic demographic variables (ie, age, sex, and years of education) were also included, given their potential correlation with cognition.

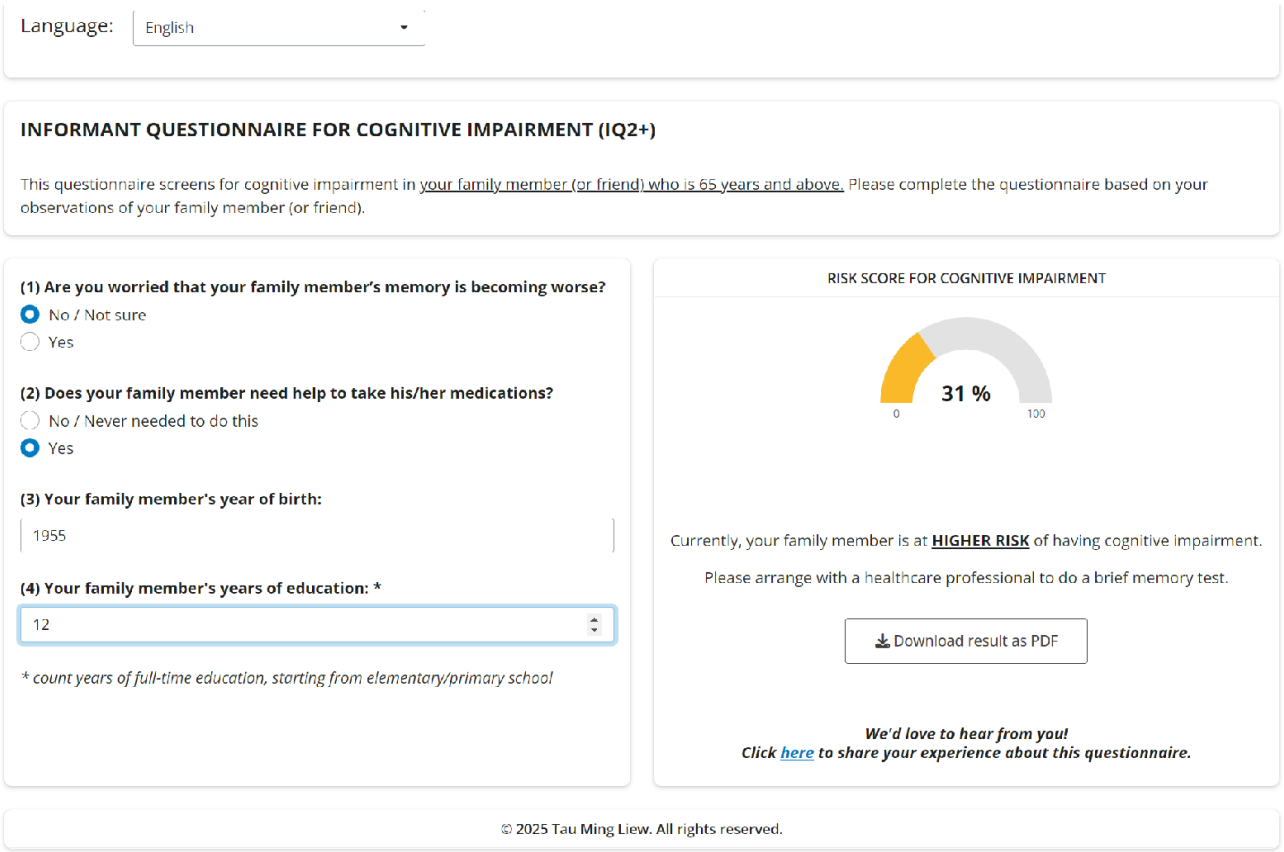
[c]The 4-item model was identified as the most parsimonious among the top models. Further details on model selection are available in Figure S1 in Multimedia Appendix 1.

[d]iADL: instrumental activities of daily living.

[e]N/A: not applicable.

[f]SCD: subjective cognitive decline

Figure 1. A sample screenshot of the interactive web-based calculator. The web-based calculator can be accessed at [76].



Performance of the Ultra-brief Questionnaire

Table 4 presents the AUC results for IQ2+ and AD8 in the *test sample* (n=557). Predicted probabilities from the new IQ2+ demonstrated excellent performance in distinguishing MCI/dementia from normal cognition (AUC 85.3%, 95% CI 80.4% - 90.2%), which was significantly better (P=.003) than a nested baseline model (comprising age, years of education, and the quadratic term of years of education; AUC 78.0%, 95% CI 72.6% - 83.4%). In contrast, AD8 had an AUC of 76.1% (95% CI 69.6% - 82.6%), which was not significantly different (P>.99) from that of the baseline model. For the detection of dementia, both IQ2+ and AD8 had AUCs >90%, which were significantly higher (P<.05) than the baseline model (83.2%). AUC results remained largely similar across education subgroups and in the 2 sensitivity analyses where the prevalence of MCI/dementia was increased to reflect average prevalence in most communities (ie, 20% [68-73] and 35% [69,70,72,73], respectively).

XSL•FO
RenderX

Table . Performance of IQ2+ for detecting cognitive impairment in the test sample (n=557) and a comparison with the performance of AD8.

Assessment tool	All education subgroups		≤10 y of education[a]		>10 y of education[a]	
	AUC[b], % (95% CI)	*P* value[c]	AUC, % (95% CI)	*P* value[c]	AUC, % (95% CI)	*P* value[c]
Detection of MCI[d]/dementia						
Baseline model (age and education)[e]	78.0 (72.6 - 83.4)	Ref[f]	76.7 (69.9 - 83.5)	Ref	74.8 (65.2 - 84.5)	Ref
IQ2+[g]	85.3 (80.4 - 90.2)	.003	84.5 (78.6 - 90.4)	.015	83.3 (72.5 - 94.0)	.095
AD8[h]	76.1 (69.6 - 82.6)	1.000	75.5 (68.0 - 83.0)	1.000	76.8 (62.7 - 91.0)	1.000
Detection of dementia						
Baseline model (age and education)[e]	83.2 (72.8 - 93.6)	Ref	77.8 (64.5 - 91.2)	Ref	94.1 (90.2 - 97.9)	Ref
IQ2+	96.7 (92.8 - 100)	.035	95.1 (89.6 - 100)	.034	99.8 (99.2 - 100)	.003
AD8	99.4 (98.8 - 99.9)	.005	99.6 (99.1 - 100)	.003	99.1 (97.8 - 100)	.059
Sensitivity analysis 1 (prevalence of MCI/dementia=20%)[i]						
Detection of MCI/dementia						
Baseline model (age and education)[e]	74.8 (68.3 - 81.3)	Ref	73.8 (65.7 - 81.9)	Ref	70.1 (58.4 - 81.7)	Ref
IQ2+	85.2 (79.6 - 90.8)	<.001	84.3 (77.4 - 91.2)	.005	83.6 (72.1 - 95.0)	.023
AD8	77.1 (70.1 - 84.2)	1.000	75.9 (67.6 - 84.2)	1.000	79.1 (64.7 - 93.5)	.525
Detection of dementia						
Baseline model (age and education)[e]	81.5 (70.7 - 92.3)	Ref	76.4 (62.9 - 90.0)	Ref	92.5 (87.7 - 97.4)	Ref
IQ2+	96.5 (92.1 - 100)	.023	94.8 (88.5 - 100)	.026	100 (100 - 100)	.005
AD8	99.1 (98.3 - 100)	.003	99.5 (98.9 - 100)	.002	98.6 (96.5 - 100)	.069
Sensitivity analysis 2 (prevalence of MCI/dementia=35%)[j]						
Detection of MCI/dementia						
Baseline model (age and education)[e]	77.0 (69.3 - 84.8)	Ref	75.6 (65.8 - 85.5)	Ref	73.2 (58.0 - 88.3)	Ref
IQ2+	85.1 (78.5 - 91.6)	.031	83.8 (75.7 - 91.9)	.135	82.4 (67.6 - 97.1)	.280
AD8	73.4 (65.0 - 81.9)	1.000	73.5 (63.5 - 83.4)	1.000	73.6 (56.3 - 91.0)	1.000
Detection of dementia						
Baseline model (age and education)[e]	80.1 (68.5 - 91.7)	Ref	73.6 (58.7 - 88.5)	Ref	95.8 (90.6 - 100)	Ref
IQ2+	94.4 (89.3 - 99.5)	.044	91.8 (84.4 - 99.2)	.046	99.2 (96.8 - 100)	.307
AD8	98.8 (97.5 - 100)	.004	99.2 (98.1 - 100)	.002	98.3 (95.2 - 100)	.893

[a]Education subgroups were stratified based on median split. This subgroup analysis has reduced statistical power (see Method S4 in Multimedia Appendix 1 for details) and is exploratory in nature.

[b]AUC: area under the receiver operating characteristic curve.

[c]*P* values were based on comparisons of AUC using the nonparametric method proposed by DeLong et al [61]. *P*<.05 indicates significant difference in AUC between the baseline model and the respective assessment tools. *P* values were Bonferroni-adjusted to minimize the risk of type 1 error in the

XSL•FO
RenderX

context of multiple testing.

[d]MCI/dementia: mild cognitive impairment or dementia.

[e]This baseline model was provided mainly for comparison purposes, by omitting IQ2+'s 2 core questions (ie, *assistance with medications* and *worry about cognition*) to examine the incremental utility of the 2 core questions beyond those provided by the demographic information of age and education. The baseline model was generated in the training sample using a logistic model with the dependent variable of MCI/dementia and with the independent variables of age, years of education, and the quadratic term of years of education. This baseline model was then applied to the test sample to generate predicted probabilities of MCI/dementia, with the predicted probabilities used for AUC comparisons.

[f]Ref: reference.

[g]IQ2+: the Informant Questionnaire for cognitive impairment–2 items (plus demographics).

[h]AD8: the 8-item Informant Interview to Differentiate Aging and Dementia.

[i]Prevalence of MCI/dementia was readjusted to 20% in the test sample based on prior meta-analytic findings that community prevalence was ~15% for MCI and ~5% for dementia. In the test sample, a subset of participants with MCI and dementia was randomly selected to readjust the prevalence in the dataset (see Methods section for further details). The resulting dataset comprised 256 participants with normal cognition (80%), 48 participants with MCI (15%), and 16 participants with dementia (5%).

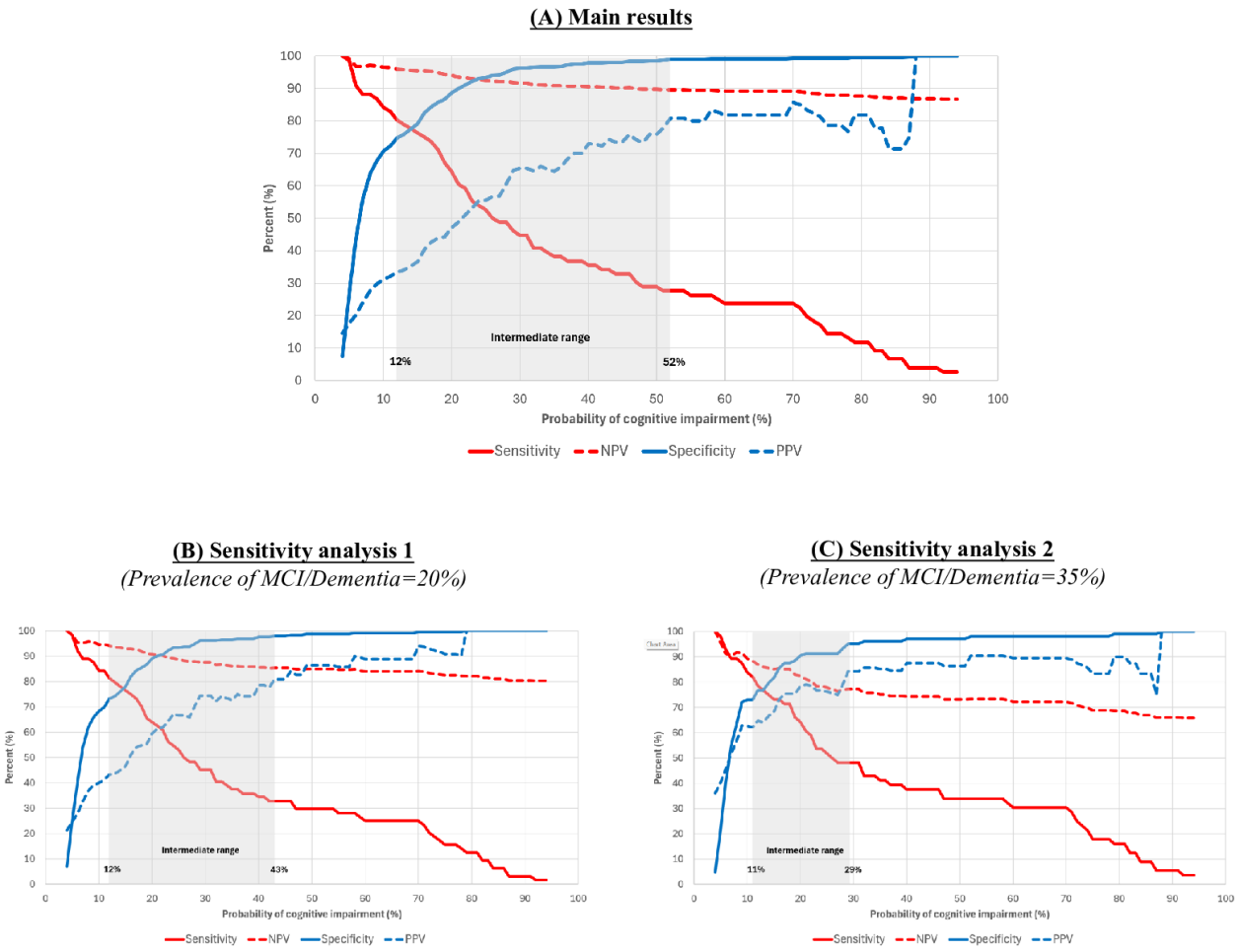
[j]Prevalence of MCI/dementia was readjusted to 35% in the test sample based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI and ~10% for dementia. In the test sample, a subset of participants with MCI and dementia was randomly selected to readjust the prevalence in the dataset (see Methods section for further details). The resulting dataset comprised 104 participants with normal cognition (65%), 40 participants with MCI (25%), and 16 participants with dementia (10%).

Optimal Cutoffs of the Ultra-Brief Questionnaire

Test statistics for IQ2+ are plotted in Figure 2A. Adopting a 2-cutoff approach, the lower cutoff (probability ≥12%) had 80.3% sensitivity and 96.0% negative predictive value and was used to rule out MCI/dementia (for individuals with probability scores below the cutoff), whereas the upper cutoff (probability ≥52%) had 99.0% specificity and 80.8% positive predictive value and identified those likely to have MCI/dementia (ie, to rule in MCI/dementia). These 2 cutoffs provide an intermediate range between them (grayed area in Figure 2A), identifying those who warrant further assessment. The optimal cutoffs varied slightly with changing prevalence of MCI/dementia, as shown in Figure 2B and 2C. Detailed results on test statistics for IQ2+ are also available in Tables S3–S5 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 2. Plot of sensitivity, specificity, NPV, and PPV based on probabilities of cognitive impairment generated from IQ2+ in the test sample (n=557). This plot is intended to demonstrate the 2-cutoff approach. The lower cutoff identifies sensitivity and NPV (red lines), which are >80% each and are used to rule out MCI/dementia (when probability scores fall below this threshold). The upper cutoff identifies specificity and PPV (blue lines), which are >80% each, and are used to rule in MCI/dementia (when probability scores exceed this threshold). The grayed area (demarcated by the lower and upper cutoffs) represents the intermediate range, identifying those who may warrant further assessment. Plot (A) was based on the main results from all the test samples (n=577). Plot (B) was based on results from sensitivity analysis 1, whereby the prevalence of MCI/dementia was readjusted to 20% in the test sample based on prior meta-analytic findings that community prevalence was ~15% for MCI and ~5% for dementia (in the test sample, a subset of participants with MCI and dementia were randomly selected in the test sample to readjust the prevalence in the dataset; see Methods section for further details). Plot (C) was based on results from sensitivity analysis 2, whereby the prevalence of MCI/dementia was readjusted to 35% in the test sample based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI and ~10% for dementia (in the test sample, a subset of participants with MCI and dementia were randomly selected in the test sample to readjust the prevalence in the dataset; see Methods section for further details). IQ2+: the Informant Questionnaire for Cognitive Impairment–2 Items (plus demographics); MCI/dementia: mild cognitive impairment or dementia; NPV: negative predictive value; PPV: positive predictive value.



Result Interpretation for the Ultra-Brief Questionnaire

The 2 cutoffs of IQ2+ effectively identified three risk categories for cognitive impairment: (1) less likely to have cognitive impairment, (2) higher risk of cognitive impairment, and (3) likely to have cognitive impairment. These 3 categories, along with their cross-tabulation with the final diagnoses, are presented in Table 5. In the first category (ie, less likely to have cognitive impairment), 88% to 96% of individuals had normal cognition. In the second category (ie, higher risk of cognitive impairment), 25% to 45% of individuals were diagnosed with MCI. In the third category (ie, likely to have cognitive impairment), 81% to 84% of the individuals had MCI/dementia, with a large proportion having dementia (47%-54%). Distinctions between these 3 risk categories of IQ2+ are also visible in the box plots in Figure 3A–3C. The first category (white region with probability scores below the lower cutoff) identified those with normal cognition, the third category (dark gray region with probability scores above the upper cutoff) identified almost all individuals with dementia, and the second category (light gray region between the lower and upper cutoffs) mostly captured those with MCI. In contrast, AD8 showed poorer discrimination between normal cognition and MCI, with discernible floor effects in the normal cognition and MCI groups as seen in the box plots in Figure 3D–3F. This is also reflected in the test statistics of AD8 (as presented in Tables S6–S8 in Multimedia Appendix 1), which demonstrated low sensitivity across its cutoff scores (with maximum sensitivity of 69.6% - 73.4%).

XSL•FO
RenderX

Table . Cross-tabulation between the output from IQ2+ and the final diagnosis in the test sample (n=557).

Result from IQ2+[ab]	Final diagnosis		
	Normal cognition	MCI[c]	Dementia
Less likely to have CI, n (%)[d]	360 (96.0)	14 (3.7)	1 (0.3)
Higher risk of CI, n (%)[d]	116 (74.4)	39 (25.0)	1 (0.6)
Likely to have CI, n (%) [d]	5 (19.2)	7 (26.9)	14 (53.8)
Sensitivity analysis 1 (prevalence of MCI/dementia=20%)[e]			
Less likely to have CI, n (%)[f]	188 (94.0)	11 (5.5)	1 (0.5)
Higher risk of CI, n (%)[f]	63 (67.0)	30 (31.9)	1 (1.1)
Likely to have CI, n (%)[f]	5 (19.2)	7 (26.9)	14 (53.8)
Sensitivity analysis 2 (Prevalence of MCI/dementia=35%)[g]			
Less likely to have CI, n (%)[h]	76 (88.0)	10 (12.0)	0 (0.0)
Higher risk of CI, n (%)[h]	23 (55.0)	18 (43.0)	1 (2.0)
Likely to have CI, n (%)[h]	5 (16.0)	12 (38.0)	15 (47.0)

[a]IQ2+, the Informant Questionnaire for cognitive impairment–2 items (plus demographics).

[b]Two-cutoff approach was adopted for IQ2+. The lower cutoff has high sensitivity and negative predictive value (>85% respectively) and is used to rule out MCI/dementia (for individuals with probability scores below the cutoff). The upper cutoff has high specificity and positive predictive value (>85% respectively) and identifies those who are likely to have MCI/dementia. These 2 cutoffs provide an intermediate range between them, identifying those who may be at higher risk and require further monitoring or assessment.

[c]MCI: mild cognitive impairment.

[d]Probability cutoff for the main results (ie, prevalence of MCI/dementia=14%):<12% (less likely to have CI), 12% to 51% (higher risk of CI), and ≥52% (likely to have CI).

[e]Prevalence of MCI/dementia was readjusted to 20% in the test sample based on prior meta-analytic findings that community prevalence was ~15% for MCI and ~5 % for dementia. In the test sample, a subset of participants with MCI and dementia was randomly selected to readjust the prevalence in the dataset (see Methods section for further details).

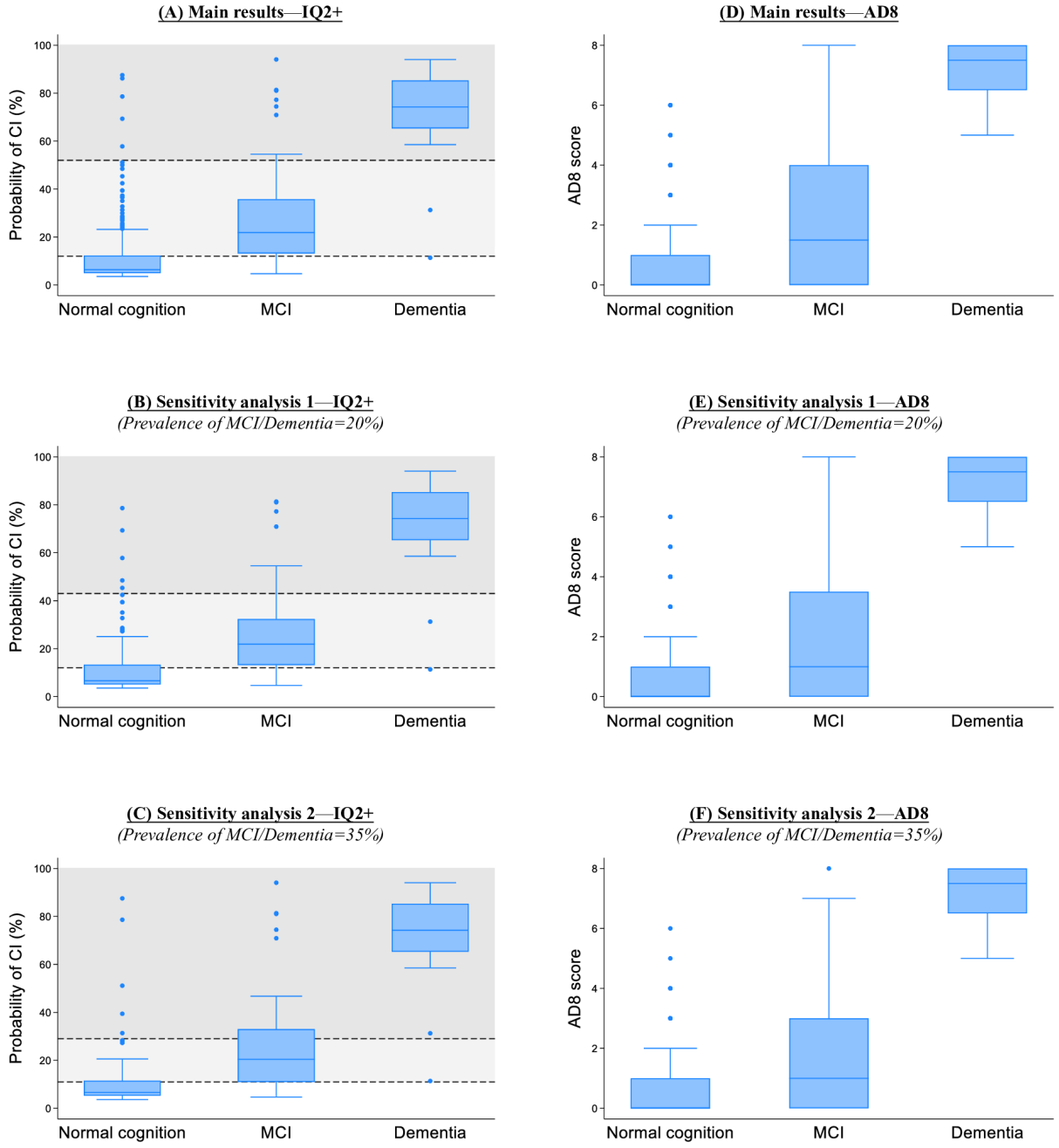
[f]Probability cutoff for the first sensitivity analysis (ie, prevalence of MCI/dementia=20%):<12% (less likely to have CI), 12% to 42% (higher risk of CI), and ≥43% (likely to have CI).

[g]Prevalence of MCI/dementia was readjusted to 35% in the test sample based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI and ~10% for dementia. In the test sample, a subset of participants with MCI and dementia was randomly selected to readjust the prevalence in the dataset (see Methods section for further details).

[h]Probability cutoff for the second sensitivity analysis (ie, prevalence of MCI/dementia=35%):<11% (less likely to have CI), 11% to 28% (higher risk of CI), and ≥29% (likely to have CI).

XSL•FO
RenderX

Figure 3. Box plots showing the distribution of IQ2+ and AD8 in the test sample (n=577). Plots (A) and (B) were based on the main results from all the test samples (n=577). Plots (C) and (D) were based on results from sensitivity analysis 1, whereby the prevalence of MCI/dementia was readjusted to 20% in the test sample based on prior meta-analytic findings that community prevalence was ~15% for MCI and ~5% for dementia (in the test sample, a subset of participants with MCI and dementia were randomly selected in the test sample to readjust the prevalence in the dataset; see Methods section for further details). Plots (E) and (F) were based on results from sensitivity analysis 2, whereby the prevalence of MCI/dementia was readjusted to 35% in the test sample based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI and ~10% for dementia (in the test sample, a subset of participants with MCI and dementia were randomly selected in the test sample to readjust the prevalence in the dataset; see Methods section for further details). In plots (B), (D), and (F), the 2 horizontal dashed lines represent the 2 optimal cutoffs for IQ2+. The lower cutoff has high sensitivity and negative predictive value (>80% each) and is used to rule out MCI/dementia when probability scores fall below this threshold (as shown by the white region). The upper cutoff has high specificity and positive predictive value (>80% each) and identifies individuals likely to have MCI/dementia (when probability scores exceed this threshold, as shown by the dark gray region). The light gray region (demarcated by the lower and upper cutoffs) represents the intermediate range, identifying those who may warrant further assessment. AD8: the 8-item Informant Interview to Differentiate Aging and Dementia; IQ2+: the Informant Questionnaire for Cognitive Impairment–2 Items (plus demographics); CI: cognitive impairment; MCI/dementia: mild cognitive impairment or dementia.



XSL•FO
RenderX

Discussion

Principal Findings

In this study, we developed and validated an ultra-brief informant questionnaire (IQ2+) for the detection of cognitive impairment in literacy-diverse communities. Using a rigorous best-subset approach with cross-validation, we identified a parsimonious 4-item model that was most useful in detecting MCI/dementia, comprising 2 informant-based questions (*assistance with medications* and *worry about cognition*), along with *age* and *years of education*. In the independent test sample, IQ2+ achieved an AUC of 85.3%, outperforming the widely used AD8 (AUC=76.1%) and a baseline demographic model (AUC=78.0%). In contrast, AD8 was not significantly better than the baseline demographic model. The robust performance of IQ2+ was consistent across education subgroups and under varying prevalence scenarios, highlighting its potential utility for case finding in diverse community settings. Adopting 2 optimal cutoffs [63-67], IQ2+ demonstrated high sensitivity, specificity, negative predictive value, and positive predictive value.

Interpretation of Findings

Despite being widely recommended [26,39-42], the AD8 performed poorly in detecting MCI/dementia and was not substantially better than a basic demographic model comprising age and education. This finding aligns with recent literature, which demonstrates AD8's lower performance for detecting early cognitive impairment (AUC 61% - 69%) compared to its performance in clear-cut dementia cases (AUC 89% - 93%) [48,77]. AD8 was originally developed for the purpose of detecting dementia [30], and its items tend to focus on identifying conspicuous changes in iADL (ie, hallmarks of the onset of dementia) [46]. As a result, AD8 may possibly not be as sensitive or consistent in detecting subtle changes associated with early cognitive impairment [30,78]. In contrast, the 2 informant questions in IQ2+ may plausibly be more sensitive to early cognitive changes. Worry about cognition—especially when reported by a knowledgeable informant—has been shown to be an early symptom of cognitive decline [51]. This item requires informants to compare current abilities with previous premorbid abilities, making it more likely to detect subtle cognitive decline over time. It also prompts informants to apply a "threshold" to decide whether they are worried about the cognitive decline, which serves as a filter to identify individuals with meaningful changes. Similarly, the ability to manage one's medications is a complex task that possibly involves multiple cognitive domains [16], such as language (for understanding medication instructions), executive function (for scheduling medication intake and problem-solving when medications are not taken on schedule), and memory (for tracking medication intake). A subtle decline in any of these domains may potentially manifest as increasing difficulty in managing medications independently, serving as an early symptom of cognitive decline [79].

Implications

At the population level, IQ2+ can serve as a quick and efficient risk stratification tool, facilitating appropriate triage for further assessment. As summarized in Table 6, low-risk individuals (<12% probability of MCI/dementia) may be reassured and advised to repeat the test in 3 to 5 years or if circumstances change (eg, appearance of new symptoms). Intermediate-risk individuals (~25% - 45% probability of MCI/dementia) can be directed to further assessments using brief cognitive tests [4,32-34] to provide more conclusive evidence of cognitive impairment [27]. High-risk individuals (>80% probability of MCI/dementia) may benefit from direct referral to memory clinics for further clinical management. This approach, as summarized in Table 6, aligns with the 2-stage strategy for active case finding of cognitive impairment (ie, subjective reports, followed by brief cognitive tests) [27,28]. It offers a scalable approach to case finding in large populations by conserving resources for cognitive testing and reserving them for individuals who truly need further verification [27]. To ensure that IQ2+ remains useful in diverse populations, its optimal cutoffs can also be adjusted depending on the prevalence of MCI/dementia in different populations (as shown in Figure 1), thus providing a more tailored solution for identifying individuals at varying risk levels.

Table . Potential clinical implications based on the result from IQ2+.

Result from IQ2+[a]	Risk communication	Potential implications
Less likely to have CI[b]	<12% chance to have CI	• Repeat IQ2+ in 3 to 5 years or when circumstances change (eg, appearance of new symptoms)
Higher risk of CI	~25% to 45% chance to have CI	• Arrange with health care professionals to do brief cognitive tests
Likely to have CI	>80% chance to have CI	• Referral to memory clinic for further assessment and management

[a]IQ2+: the Informant Questionnaire for cognitive impairment–2 items (plus demographics).

[b]CI: cognitive impairment.

Essentially, IQ2+ may have 2 plausible use cases in the community. First, it allows family members who have concerns about a loved one's cognition to complete the questionnaire online (accessible at [76]), providing an immediate result with risk score and a brief interpretation of its implications. The results can also be downloaded as a PDF file for further

XSL•FO
RenderX

discussion with health care providers. This approach mirrors a well-established practice in the field of diabetes mellitus, where the public is encouraged to complete an online risk test [80] as an initial screening to determine the need for further diagnostic tests [81]. By enabling self-appraisal, this use case of IQ2+ leverages family members for efficient case finding and empowers them to proactively support the brain health of older persons.

Second, IQ2+ can also be used by health care workers who routinely conduct comprehensive geriatric assessments in the community. Since 2019, the WHO has advocated for routine assessment of intrinsic capacity among community-dwelling older persons, which requires comprehensive evaluations of cognition, mobility, nutrition, hearing, vision, and mood [29]. Given the extensive range of components to cover, it can be challenging to include routine cognitive testing in all geriatric assessments. IQ2+, which can be completed in <1 minute, is well-suited to be embedded within initial geriatric assessments to prioritize individuals who require further cognitive testing. Given its excellent performance across education subgroups (Table 4), IQ2+ offers a practical tool that may potentially be broadly implemented in diverse populations, including in underserved populations in LMICs, which currently have the largest number of undiagnosed cognitive impairment [82]. This is also in line with the 2024 Lancet Commission's call to address the unmet need for cognitive screening tools that are also suited for individuals with lower literacy in LMICs [35].

Limitations

Several limitations are notable. First, IQ2+ was developed for older individuals aged ≥65 years. It is unclear whether IQ2+ would be useful for individuals aged <65 years. Second, IQ2+ would benefit from further validations in other cultures and languages, as cultural differences may affect how informants respond to its items and could potentially modify its performance in different settings. Third, IQ2+ requires the use of a web-based calculator to determine the probability of cognitive impairment. This approach leverages technology to automate test scoring, result visualization, and interpretation. However, it may limit accessibility compared to pen-and-paper tests that use raw scores as a cutoff (eg, AD8), particularly in settings with limited access to technology. Fourth, one of the core questions of IQ2+ involves difficulty managing medications. This item was identified as the most useful item in our exhaustive search (Table 3) and has some face validity (as it detects subtle changes in a complex task of managing medications [16,79], and the need to take regular medications may also be a proxy for higher risk due to the presence of chronic diseases). However, this question may be less useful among healthier older individuals who do not need to take medications regularly. Fifth, although IQ2+ can be a useful case finding tool, it is not intended to replace comprehensive clinical and neuropsychological assessments, which provide more definitive diagnoses as well as granular information on specific cognitive deficits [20,83,84].

Conclusions

Using rigorous methodology, this study developed an ultra-brief questionnaire that enables untrained laypersons to screen for cognitive impairment in <1 minute. Despite its brevity, the questionnaire demonstrated excellent performance in detecting MCI/dementia, outperforming the well-established AD8. The questionnaire can be completed by members of the public who have concerns about a family member's cognition (accessible at [76]) or embedded within community geriatric assessments to prioritize cognitive testing. Its excellent performance was consistent across education subgroups and varying prevalence scenarios, supporting its utility for case finding in diverse populations, including underserved communities and LMICs.

Acknowledgments

The authors would like to thank the following persons for their assistance in recruiting study participants: Yanling Tan, Alcantara Leicester Shawn, Kai Xin Choo, Megan Cheng Mun Choy, Xin Tong Tan, Lydia Jia En Cheong, Jane Mee Chin Liew, Xiao Hui Ng, and Spencer Peng Ming Yuen. The authors would also like to thank the following community organizations for their support in participant recruitment: Thye Hua Kwan Moral Charities (AMK 645, Bukit Merah View, Beo Crescent), Kreta Ayer Senior Activity Centres, NTUC Health Active Ageing Centre (Lengkok Bahru), Agency for Integrated Care (Caregiving & Community Mental Health Division), Silver Generation Office, People's Association, Precious Active Ageing Centre, Montfort Care, Yong-en Care Centre, Presbyterian Community Services, Lions Befrienders Service Association Singapore.

Funding

This research is funded by the Singapore Prime Minister Office's Smart Nation and Digital Government Office (grant I_20092346). Separately, TML is supported by the Singapore Ministry of Health's National Medical Research Council (grant HCSAINV23jul-0001, NMRC/CG2/005e/2022-SGH, and MOH-SEEDFD22apr-0001). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

The data used in this study contain sensitive information about the study participants, and they did not provide consent for public data sharing. The current approval by the SingHealth Centralized IRB of Singapore (reference number: 2021/2590) does not include data sharing. A minimal dataset could be shared by request from a qualified academic investigator for the sole purpose of replicating the present study, provided the data transfer is in agreement with Singapore's legislation on the general data protection regulation and a data sharing agreement has been signed. Contact person: corresponding author.

XSL•FO
RenderX

Authors' Contributions

Conceptualization – TML
Data curation – TML
Formal analysis – TML
Funding acquisition - TML
Investigation – consensus diagnosis – TML, KFY, KN, SKST, WL, SYT, WIK
Methodology – TML
Project administration – TML
Resources – training of research coordinators – TML, SYT, WIK
Software – TML
Supervision – TML
Validation – TML
Visualization – TML
Writing – original draft – TML
Writing – review & editing – TML, KFY, KN, SKST, WL, SYT, WIK

Conflicts of Interest

TML has provided consultation to Lundbeck and Eisai. KN has provided consultation to Takeda. The remaining authors declare no conflicts of interest.

Multimedia Appendix 1
Additional information on methodology and results.
[PDF File, 1207 KB - aging_v9i1e72963_app1.pdf]

References

1. Dementia: A Public Health Priority: World Health Organization; 2012. URL: https://www.who.int/publications/i/item/dementia-a-public-health-priority [accessed 2025-12-10]
2. Lang L, Clifford A, Wei L, et al. Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis. BMJ Open 2017 Feb 3;7(2):e011146. [doi: 10.1136/bmjopen-2016-011146] [Medline: 28159845]
3. Liu Y, Jun H, Becker A, Wallick C, Mattke S. Detection rates of mild cognitive impairment in primary care for the United States Medicare population. J Prev Alzheimers Dis 2024;11(1):7-12. [doi: 10.14283/jpad.2023.131] [Medline: 38230712]
4. Liew TM. A 4-item case-finding tool to detect dementia in older persons. J Am Med Dir Assoc 2019 Dec;20(12):1529-1534. [doi: 10.1016/j.jamda.2019.06.015] [Medline: 31399363]
5. Burns A, Iliffe S. Dementia. BMJ 2009 Feb 5;338:b75. [doi: 10.1136/bmj.b75] [Medline: 19196746]
6. Ying J, Yap P, Gandhi M, Liew TM. Iterating a framework for the prevention of caregiver depression in dementia: a multi-method approach. Int Psychogeriatr 2018 Aug;30(8):1119-1130. [doi: 10.1017/S1041610217002629] [Medline: 29223191]
7. Liew TM. Neuropsychiatric symptoms in early stage of Alzheimer's and non-Alzheimer's dementia, and the risk of progression to severe dementia. Age Ageing 2021 Sep 11;50(5):1709-1718. [doi: 10.1093/ageing/afab044] [Medline: 33770167]
8. Liew TM, Lee CS. Reappraising the efficacy and acceptability of multicomponent interventions for caregiver depression in dementia: the utility of network meta-analysis. Gerontologist 2019 Jul 16;59(4):e380-e392. [doi: 10.1093/geront/gny061] [Medline: 29860310]
9. Ang LC, Malhotra R, Roy Chowdhury A, Liew TM. Pre-and post-COVID-19 trends related to dementia caregiving on Twitter. Sci Rep 2025 Feb 12;15(1):5173. [doi: 10.1038/s41598-024-82405-8] [Medline: 39939632]
10. Thyrian JR, Hertel J, Wucherer D, et al. Effectiveness and safety of dementia care management in primary care: a randomized clinical trial. JAMA Psychiatry 2017 Oct 1;74(10):996-1004. [doi: 10.1001/jamapsychiatry.2017.2124] [Medline: 28746708]
11. Vickrey BG, Mittman BS, Connor KI, et al. The effect of a disease management intervention on quality and outcomes of dementia care: a randomized, controlled trial. Ann Intern Med 2006 Nov 21;145(10):713-726. [doi: 10.7326/0003-4819-145-10-200611210-00004] [Medline: 17116916]
12. Cepoiu-Martin M, Tam-Tham H, Patten S, Maxwell CJ, Hogan DB. Predictors of long-term care placement in persons with dementia: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2016 Nov;31(11):1151-1171. [doi: 10.1002/gps.4449] [Medline: 27045271]
13. Jennings LA, Laffan AM, Schlissel AC, et al. Health care utilization and cost outcomes of a comprehensive dementia care program for Medicare beneficiaries. JAMA Intern Med 2019 Feb 1;179(2):161-166. [doi: 10.1001/jamainternmed.2018.5579] [Medline: 30575846]

XSL•FO
RenderX

14. Spijker A, Vernooij-Dassen M, Vasse E, et al. Effectiveness of nonpharmacological interventions in delaying the institutionalization of patients with dementia: a meta-analysis. J Am Geriatr Soc 2008 Jun;56(6):1116-1128. [doi: 10.1111/j.1532-5415.2008.01705.x] [Medline: 18410323]
15. Bott NT, Sheckter CC, Yang D, et al. Systems delivery innovation for Alzheimer disease. Am J Geriatr Psychiatry 2019 Feb;27(2):149-161. [doi: 10.1016/j.jagp.2018.09.015] [Medline: 30477913]
16. Elliott RA, Goeman D, Beanland C, Koch S. Ability of older people with dementia or cognitive impairment to manage medicine regimens: a narrative review. Curr Clin Pharmacol 2015;10(3):213-221. [doi: 10.2174/1574884710666150812141525] [Medline: 26265487]
17. Persson S, Saha S, Gerdtham UG, Toresson H, Trépel D, Jarl J. Healthcare costs of dementia diseases before, during and after diagnosis: longitudinal analysis of 17 years of Swedish register data. Alzheimers Dement 2022 Dec;18(12):2560-2569. [doi: 10.1002/alz.12619] [Medline: 35189039]
18. Chay J, Koh WP, Tan KB, Finkelstein EA. Healthcare burden of cognitive impairment: evidence from a Singapore Chinese health study. Ann Acad Med Singap 2024 Apr 29;53(4):233-240. [doi: 10.47102/annals-acadmedsg.2023253] [Medline: 38920180]
19. Liew TM. Distinct trajectories of subjective cognitive decline before diagnosis of neurocognitive disorders: longitudinal modelling over 18 years. J Prev Alzheimers Dis 2025 May;12(5):100123. [doi: 10.1016/j.tjpad.2025.100123] [Medline: 40057463]
20. Liew TM. Developing a brief neuropsychological battery for early diagnosis of cognitive impairment. J Am Med Dir Assoc 2019 Aug;20(8):1054-1e11. [doi: 10.1016/j.jamda.2019.02.028] [Medline: 30992186]
21. Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. Lancet 2015 Jun 6;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5] [Medline: 25771249]
22. Baker LD, Espeland MA, Whitmer RA, et al. Structured vs self-guided multidomain lifestyle interventions for global cognitive function: the US POINTER randomized clinical trial. JAMA 2025 Aug 26;334(8):681-691. [doi: 10.1001/jama.2025.12923] [Medline: 40720610]
23. van Dyck CH, Swanson CJ, Aisen P, et al. Lecanemab in early Alzheimer's disease. N Engl J Med 2023 Jan 5;388(1):9-21. [doi: 10.1056/NEJMoa2212948] [Medline: 36449413]
24. Sims JR, Zimmer JA, Evans CD, et al. Donanemab in early symptomatic Alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. JAMA 2023 Aug 8;330(6):512-527. [doi: 10.1001/jama.2023.13239] [Medline: 37459141]
25. Prince M, Bryce R, Ferri C. World alzheimer report 2011: the benefits of early diagnosis and intervention. : Alzheimer's Disease International; 2011 URL: https://www.alzint.org/u/WorldAlzheimerReport2011.pdf [accessed 2025-12-10]
26. Morley JE, Morris JC, Berg-Weger M, et al. Brain health: the importance of recognizing cognitive impairment: an IAGG consensus conference. J Am Med Dir Assoc 2015 Sep 1;16(9):731-739. [doi: 10.1016/j.jamda.2015.06.017] [Medline: 26315321]
27. Liew TM. Active case finding of dementia in ambulatory care settings: a comparison of three strategies. Eur J Neurol 2020 Oct;27(10):1867-1878. [doi: 10.1111/ene.14353] [Medline: 32441837]
28. Hendry K, Hill E, Quinn TJ, Evans J, Stott DJ. Single screening questions for cognitive impairment in older people: a systematic review. Age Ageing 2015 Mar;44(2):322-326. [doi: 10.1093/ageing/afu167] [Medline: 25385272]
29. World Health Organization. Integrated Care for Older People (ICOPE): Guidance for Person-Centred Assessment and Pathways in Primary Care 2019. URL: https://www.who.int/publications/b/71300
30. Galvin JE, Roe CM, Powlishta KK, et al. The AD8: a brief informant interview to detect dementia. Neurology (ECronicon) 2005 Aug 23;65(4):559-564. [doi: 10.1212/01.wnl.0000172958.95282.2a] [Medline: 16116116]
31. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol 1982 May;37(3):323-329. [doi: 10.1093/geronj/37.3.323] [Medline: 7069156]
32. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The mini‐cog: a cognitive "vital signs" measure for dementia screening in multi‐lingual elderly. Int J Geriat Psychiatry 2000 Nov;15(11):1021-1027. [doi: 10.1002/1099-1166(200011)15:11<1021::AID-GPS234>3.0.CO;2-6] [Medline: 11113982]
33. Liew TM. The optimal short version of montreal cognitive assessment in diagnosing mild cognitive impairment and dementia. J Am Med Dir Assoc 2019 Aug;20(8):1055-10e1. [doi: 10.1016/j.jamda.2019.02.004] [Medline: 30910550]
34. McDicken JA, Elliott E, Blayney G, et al. Accuracy of the short‐form montreal cognitive assessment: systematic review and validation. Int J Geriat Psychiatry 2019 Oct;34(10):1515-1525. [doi: 10.1002/gps.5162]
35. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet Standing Commission. The Lancet 2024 Aug;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0]
36. Fayers PM, Machin D. Quality of Life: The Assessment, Analysis and Interpretation of Patient-Reported Outcomes, 2nd edition: John Wiley & Sons; 2007. [doi: 10.1002/9780470024522]
37. Tanwani R, Danquah MO, Butris N, et al. Diagnostic accuracy of ascertain dementia 8-item questionnaire by participant and informant-a systematic review and meta-analysis. PLoS ONE 2023;18(9):e0291291. [doi: 10.1371/journal.pone.0291291] [Medline: 37699028]

XSL•FO
RenderX

38. Taylor-Rowan M, Nafisi S, Owen R, et al. Informant-based screening tools for dementia: an overview of systematic reviews. Psychol Med 2023 Jan;53(2):580-589. [doi: 10.1017/S0033291721002002] [Medline: 34030753]
39. A 4-step process to detecting cognitive impairment and earlier diagnosis of dementia: approaches and tools for primary care providers. : The Gerontological Society of America; 2017 URL: https://issuu.com/gsastrategicalliances/docs/gsa_kaer_toolkit_4-step_process_to_detecting_cogni?fr=xKAE9_zU1NQ [accessed 2025-12-10]
40. Cordell CB, Borson S, Boustani M, et al. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. Alzheimers Dement 2013 Mar;9(2):141-150. [doi: 10.1016/j.jalz.2012.09.011] [Medline: 23265826]
41. Atri A, Dickerson BC, Clevenger C, et al. The Alzheimer's association clinical practice guideline for the diagnostic evaluation, testing, counseling, and disclosure of suspected Alzheimer's disease and related disorders (DETeCD-ADRD): validated clinical assessment instruments. Alzheimers Dement 2025 Jan;21(1):e14335. [doi: 10.1002/alz.14335] [Medline: 39713939]
42. Maslow K, Fortinsky RH. Nonphysician care providers can help to increase detection of cognitive impairment and encourage diagnostic evaluation for dementia in community and residential care settings. Gerontologist 2018 Jan 18;58(suppl_1):S20-S31. [doi: 10.1093/geront/gnx171] [Medline: 29361070]
43. The World Bank in Singapore. World Bank Group. 2024. URL: https://www.worldbank.org/en/country/singapore/overview [accessed 2025-12-02]
44. Liew TM, Foo JYH, Yang H, et al. PENSIEVE-AI a brief cognitive test to detect cognitive impairment across diverse literacy. Nat Commun 2025 Mar 23;16(1):2847. [doi: 10.1038/s41467-025-58201-x] [Medline: 40122854]
45. Patnode CD, Perdue LA, Rossom RC, et al. Screening for cognitive impairment in older adults: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2020 Feb 25;323(8):764-785. [doi: 10.1001/jama.2019.22258] [Medline: 32096857]
46. Diagnostic and Statistical Manual of Mental Disorders: DSM-5: American Psychiatric Association; 2013. [doi: 10.1176/appi.books.9780890425596]
47. Petersen RC, Morris JC. Mild cognitive impairment as a clinical entity and treatment target. Arch Neurol 2005 Jul;62(7):1160-1163. [doi: 10.1001/archneur.62.7.1160] [Medline: 16009779]
48. Kan CN, Zhang L, Cheng CY, et al. The informant AD8 can discriminate patients with dementia from healthy control participants in an Asian older cohort. J Am Med Dir Assoc 2019 Jun;20(6):775-779. [doi: 10.1016/j.jamda.2018.11.023] [Medline: 30661859]
49. Jessen F, Wiese B, Bachmann C, et al. Prediction of dementia by subjective memory impairment: effects of severity and temporal association with cognitive impairment. Arch Gen Psychiatry 2010 Apr;67(4):414-422. [doi: 10.1001/archgenpsychiatry.2010.30] [Medline: 20368517]
50. van Harten AC, Mielke MM, Swenson-Dravis DM, et al. Subjective cognitive decline and risk of MCI: the Mayo Clinic Study of Aging. Neurology (ECronicon) 2018 Jul 24;91(4):e300-e312. [doi: 10.1212/WNL.0000000000005863] [Medline: 29959257]
51. Jessen F, Amariglio RE, Buckley RF, et al. The characterisation of subjective cognitive decline. Lancet Neurol 2020 Mar;19(3):271-278. [doi: 10.1016/S1474-4422(19)30368-0] [Medline: 31958406]
52. Liew TM. Depression, subjective cognitive decline, and the risk of neurocognitive disorders. Alzheimers Res Ther 2019 Aug 9;11(1):70. [doi: 10.1186/s13195-019-0527-7] [Medline: 31399132]
53. Liew TM. Subjective cognitive decline, anxiety symptoms, and the risk of mild cognitive impairment and dementia. Alzheimers Res Ther 2020 Sep 11;12(1):107. [doi: 10.1186/s13195-020-00673-8] [Medline: 32917264]
54. Liew TM. Trajectories of subjective cognitive decline, and the risk of mild cognitive impairment and dementia. Alzheimers Res Ther 2020 Oct 27;12(1):135. [doi: 10.1186/s13195-020-00699-y] [Medline: 33109275]
55. Liew TM. Subjective cognitive decline, APOE e4 allele, and the risk of neurocognitive disorders: age- and sex-stratified cohort study. Aust N Z J Psychiatry 2022 Dec;56(12):1664-1675. [doi: 10.1177/00048674221079217] [Medline: 35229693]
56. Ng TP, Niti M, Chiam PC, Kua EH. Physical and cognitive domains of the instrumental activities of daily living: validation in a multiethnic population of Asian older adults. J Gerontol A Biol Sci Med Sci 2006 Jul;61(7):726-735. [doi: 10.1093/gerona/61.7.726] [Medline: 16870636]
57. Bestglm: best subset GLM. 2010. URL: https://cran.r-project.org/package=bestglm [accessed 2024-07-17]
58. Liew TM, Yap P. A brief, 6-item scale for caregiver grief in dementia caregiving. Gerontologist 2020;60:e1-e10. [doi: 10.1093/geront/gny161]
59. Liew TM, Yap P. A 3-item screening scale for caregiver burden in dementia caregiving: scale development and score mapping to the 22-item Zarit burden interview. J Am Med Dir Assoc 2019 May;20(5):629-633. [doi: 10.1016/j.jamda.2018.11.005] [Medline: 30591383]
60. Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression, 3rd edition: John Wiley & Sons; 2013. [doi: 10.1002/9781118548387]
61. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988 Sep;44(3):837-845. [Medline: 3203132]

XSL•FO
RenderX

62. Yu J, Yap P, Liew TM. The optimal short version of the Zarit Burden Interview for dementia caregivers: diagnostic utility and externally validated cutoffs. Aging Ment Health 2019 Jun;23(6):706-710. [doi: 10.1080/13607863.2018.1450841] [Medline: 29553806]
63. Dautzenberg GMFG, Lijmer JG, Beekman ATF. The Montreal Cognitive Assessment (MoCA) with a double threshold: improving the MoCA for triaging patients in need of a neuropsychological assessment. Int Psychogeriatr 2022 Jun;34(6):571-583. [doi: 10.1017/S1041610221000612] [Medline: 34463237]
64. Landsheer JA. Impact of the prevalence of cognitive impairment on the accuracy of the Montreal Cognitive Assessment: the advantage of using two MoCA thresholds to identify error-prone test scores. Alzheimer Dis Assoc Disord 2020;34(3):248-253. [doi: 10.1097/WAD.0000000000000365] [Medline: 31934880]
65. Thomann AE, Berres M, Goettel N, Steiner LA, Monsch AU. Enhanced diagnostic accuracy for neurocognitive disorders: a revised cut-off approach for the Montreal Cognitive Assessment. Alz Res Therapy 2020 Dec;12(1):39. [doi: 10.1186/s13195-020-00603-8]
66. Swartz RH, Cayley ML, Lanctôt KL, et al. Validating a pragmatic approach to cognitive screening in stroke prevention clinics using the Montreal Cognitive Assessment. Stroke 2016 Mar;47(3):807-813. [doi: 10.1161/STROKEAHA.115.011036] [Medline: 26903584]
67. Jack CR, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. Alzheimer's Dementia 2024 Aug;20(8):5143-5169. [doi: 10.1002/alz.13859]
68. Bai W, Chen P, Cai H, et al. Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: a meta-analysis and systematic review of epidemiology studies. Age Ageing 2022 Aug 2;51(8):afac173. [doi: 10.1093/ageing/afac173] [Medline: 35977150]
69. Hu C, Yu D, Sun X, Zhang M, Wang L, Qin H. The prevalence and progression of mild cognitive impairment among clinic and community populations: a systematic review and meta-analysis. Int Psychogeriatr 2017 Oct;29(10):1595-1608. [doi: 10.1017/S1041610217000473] [Medline: 28884657]
70. Song WX, Wu WW, Zhao YY, et al. Evidence from a meta-analysis and systematic review reveals the global prevalence of mild cognitive impairment. Front Aging Neurosci 2023;15:1227112. [doi: 10.3389/fnagi.2023.1227112] [Medline: 37965493]
71. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. Alzheimer's Dementia 2013 Jan;9(1):63-75. [doi: 10.1016/j.jalz.2012.11.007] [Medline: 23305823]
72. Fiest KM, Jetté N, Roberts JI, et al. The prevalence and incidence of dementia: a systematic review and meta-analysis. Can J Neurol Sci 2016 Apr;43:S3-S50. [doi: 10.1017/cjn.2016.18] [Medline: 27307127]
73. Cao Q, Tan CC, Xu W, et al. The prevalence of dementia: a systematic review and meta-analysis. J Alzheimers Dis 2020;73(3):1157-1166. [doi: 10.3233/JAD-191092] [Medline: 31884487]
74. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983 Sep;148(3):839-843. [doi: 10.1148/radiology.148.3.6878708] [Medline: 6878708]
75. Mental Capacity Act 2008. Singapore Statutes Online. URL: https://sso.agc.gov.sg/Act/MCA2008 [accessed 2024-07-17]
76. Informant questionnaire for cognitive impairment. URL: https://brainhealth-iq2plus.hf.space/ [accessed 2025-12-10]
77. Cai Y, Qiu P, Wan Y, et al. Establishing cut-off scores for the self-rating AD8 based on education level. Geriatr Nurs (Lond) 2021 Sep;42(5):1093-1098. [doi: 10.1016/j.gerinurse.2021.06.027]
78. Chen HH, Sun FJ, Yeh TL, et al. The diagnostic accuracy of the Ascertain Dementia 8 questionnaire for detecting cognitive impairment in primary care in the community, clinics and hospitals: a systematic review and meta-analysis. Fam Pract 2018 May 23;35(3):239-246. [doi: 10.1093/fampra/cmx098] [Medline: 29045636]
79. Barthold D, Marcum ZA, Chen S, et al. Difficulty with taking medications is associated with future diagnosis of Alzheimer's disease and related dementias. J Gen Intern Med 2021 Apr;36(4):863-868. [doi: 10.1007/s11606-020-06279-y] [Medline: 33037589]
80. Take the type 2 risk test. American Diabetes Association. 2024. URL: https://diabetes.org/diabetes-risk-test [accessed 2024-07-14]
81. American Diabetes Association Professional Practice Committee. 2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes-2024. Diabetes Care 2024 Jan 1;47(Suppl 1):S20-S42. [doi: 10.2337/dc24-S002] [Medline: 38078589]
82. Kalaria R, Maestre G, Mahinrad S, et al. The 2022 symposium on dementia and brain aging in low- and middle-income countries: Highlights on research, diagnosis, care, and impact. Alzheimers Dement 2024 Jun;20(6):4290-4314. [doi: 10.1002/alz.13836] [Medline: 38696263]
83. Jacova C, Kertesz A, Blair M, Fisk JD, Feldman HH. Neuropsychological testing and assessment for dementia. Alzheimer's Dementia 2007 Oct;3(4):299-317. [doi: 10.1016/j.jalz.2007.07.011] [Medline: 19595951]
84. Ang LC, Yap P, Tay SY, Koay WI, Liew TM. Examining the validity and utility of Montreal Cognitive Assessment domain scores for early neurocognitive disorders. J Am Med Dir Assoc 2023 Mar;24(3):314-320. [doi: 10.1016/j.jamda.2022.12.028] [Medline: 36758620]

XSL•FO
RenderX

Abbreviations

AD8: 8-item Informant Interview to Differentiate Aging and Dementia
AUC: area under the receiver operating characteristic curve
iADL: instrumental activities of daily living
IQ2+: Informant Questionnaire for cognitive impairment–2 items plus demographics
LMIC: lower- and middle-income country
MCI/dementia: mild cognitive impairment or dementia
SCD: subjective cognitive decline
WHO: World Health Organization

Edited by G Seçkin; submitted 21.Feb.2025; peer-reviewed by D Urbanski, YJ Leow; revised version received 14.Aug.2025; accepted 14.Nov.2025; published 06.Jan.2026.

Please cite as:
Liew TM, Yip KF, Narasimhalu K, Ting SKS, Li W, Tay SY, Koay WI
An Ultra-Brief Informant Questionnaire for Case Finding of Cognitive Impairment Across Diverse Literacy: Diagnostic Accuracy Study
JMIR Aging 2026;9:e72963
URL: https://aging.jmir.org/2026/1/e72963
doi:10.2196/72963

© Tau Ming Liew, King Fan Yip, Kaavya Narasimhalu, Simon Kang Seng Ting, Weishan Li, Sze Yan Tay, Way Inn Koay. Originally published in JMIR Aging (https://aging.jmir.org), 6.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Data-Driven Phenotyping Reveals Nonuniform Association Between Age and Mortality After Aortic Surgery: Retrospective Cohort Study of UK Biobank Data

Maria Elisabeth Leinweber, MD; Fadi Taher, MD; Miriam Kliewer, MD; Afshin Assadian, MD; Amun Georg Hofmann, MD

Department of Vascular and Endovascular Surgery, Klinik Ottakring, Montleartstraße 37, Vienna, Austria

Corresponding Author:
Amun Georg Hofmann, MD
Department of Vascular and Endovascular Surgery, Klinik Ottakring, Montleartstraße 37, Vienna, Austria

Abstract

Background: Life expectancy and age are frequently considered factors to assess perioperative and postoperative mortality risks in patients affected by aortic pathologies, which can affect the decision whether to suggest invasive treatment.

Objective: This study aims to investigate the association between age and all-cause mortality after invasive aortic treatment.

Methods: Unsupervised clustering (k-means) using data from the UK Biobank was conducted for patients with aortic pathologies (*International Classification of Diseases, Tenth Revision* [ICD-10] group I71) receiving endovascular or open surgical treatment. Clustering variables encompassed demographic and clinical parameters. Survival analyses (postoperative survival time in days to all-cause death) between clusters and cluster-derived age groups were conducted.

Results: The study included 1801 individuals undergoing surgical or endovascular repair for aortic aneurysms. Unsupervised cluster analysis identified distinct groups primarily based on age, both in models using 2 or 3 clusters. Clusters with older patients at surgery exhibited lower postoperative survival, with perioperative mortality disproportionately affecting these groups. While age was significantly associated with postoperative mortality overall (hazard ratio 1.07, 95% CI 1.05 - 1.08), this association diminished in older clusters after excluding perioperative deaths, a trend confirmed in analyses adjusted for relevant confounders.

Conclusions: Unsupervised cluster analysis revealed age as the primary factor distinguishing patient groups undergoing invasive treatment for aortic pathologies. However, age at surgery appears to have different consequences in certain age brackets, indicating a complex nonuniform relationship.

(JMIR Aging 2026;9:e75611) doi:10.2196/75611

KEYWORDS

aneurysm; aortic aneurysm; cluster analysis; survival analysis; outcomes

Introduction

Aortic aneurysms (AAs) encompass a group of pathologies affecting the aorta in any of its segments from the ascending aorta to the infrarenal abdominal aorta, which is the most prevalent type [1]. Apart from true degenerative aneurysms, dissections are commonly included in this group of conditions or summarized as aortic syndromes due to a similar enlargement of diameter and treatment options [2]. The aorta is now considered an organ rather than a simple conduit to other tissues due to its complex structure and function [3]. Considering the high rate of adverse outcomes during acute ruptures, preemptive treatment is favored when possible in elective settings and is associated with significantly lower rates of perioperative morbidity and mortality [4]. Over the past 2 decades, shifts in global cardiovascular risk factors, the establishment of national screening programs, and the widespread adoption of endovascular aneurysm repair have significantly altered the epidemiology and management of AA [5-7]. Most high-income countries have experienced a marked decline in both incidence and mortality rates for elective and ruptured AA. However, demographic changes and expanded treatment options for older and frail patients have introduced new challenges [5,8].

Estimated life expectancy is frequently considered to be a relevant factor in clinical decision-making as the risks and benefits of treatment options may be assessed differently for patients with varying life expectancies. It is broadly suggested that younger and healthier patients can reasonably be exposed to the higher perioperative risks of open surgical repair to balance out limitations of endovascular strategies including concerns with durability and imaging surveillance [4,9]. The tradeoff between short-term survival advantages of endovascular repair and potential long-term disadvantages is not yet fully elucidated and, in general, is more nuanced and influenced by

XSL•FO
RenderX

patient characteristics, expectations, anatomy, and comorbidities [10].

Analyses of large population-based datasets and administrative registries have identified several risk factors associated with adverse perioperative outcomes and impaired long-term survival. However, predicting mortality and serious adverse events is complex, and traditional association studies alone may not be sufficient [11-14].

Unsupervised clustering algorithms can be used to identify patterns within complex clinical data, enabling a different approach to risk stratification by grouping patients with similar phenotypes. Unlike traditional regression models, which isolate independent effects and may produce high hazard ratios for rare but extreme risks, cluster analysis can identify clinically coherent patient archetypes where multiple low and moderate risk factors can accumulate to define high-risk subgroups. This is particularly valuable in surgical cohorts, where multimorbidity, rather than selective patient factors, frequently determines outcomes.

In this work, we investigated whether an unsupervised cluster analysis based on demographic and clinical patient characteristics and naïve to postoperative outcomes can be leveraged to identify distinct postoperative survival patterns for specific patient groups.

Methods

Study Design and Dataset

In this retrospective observational analysis, an unsupervised cluster analysis was conducted using data from the UK Biobank [15]. The UK Biobank is a large population cohort study with 502,488 participants aged 40 to 69 years, with a recruitment phase from 2006 to 2010, spanning over 22 assessment centers in the United Kingdom. During recruitment, participants consented to synchronizing their UK Biobank patient file with their electronic medical records from the National Health Service (NHS) [16]. The World Health Organization (WHO)'s *International Classification of Diseases, Tenth Revision* (ICD-10) *codes* were used to code clinical diagnoses of hospital inpatient data. The *Office of Population, Censuses and Surveys: Classification of Interventions and Procedures* (OPCS-4) codes were used to code operations and procedures. Data from the NHS Information Center for participants from England and Wales and from the NHS Central Register Scotland for participants from Scotland were used to obtain date and cause of death [17,18]. The investigation was conducted in all patients with the ICD-10 group I71 (AA and dissection). Included OPSC-4 codes to classify procedures were: L18*, emergency replacement of aneurysmal segment of aorta; L19*, other replacement of aneurysmal segment of aorta; L25.4, operations on aneurysm of aorta, not elsewhere classified (NEC); L26.6, transluminal aortic stent graft with fenestration, NEC; L26.7, transluminal aortic branched stent graft, NEC; L27*, transluminal insertion of stent graft for aneurysmal segment of aorta; and L28*, transluminal operations on aneurysmal segment of aorta.

Ethical Considerations

This study accessed UK Biobank data. Ethical approval for the UK Biobank was obtained from the North West Multi-Centre Research Ethics Committee (11/NW/03820). All enrolled participants gave written informed consent prior to enrollment in accordance with the tenets of the Declaration of Helsinki.

Cluster Analysis

Cluster analysis was performed using the k-means clustering algorithm. Optimization of k was performed using the NbClust function (*NbClust* package [19]) in R (version 4.4.1; R Foundation for Statistical Computing), which uses 30 different indices. Clusters were formed based on 14 features, encompassing demographics and comorbidities. The list of variables included: sex, age at surgery, coronary heart disease, hypertension, heart failure, cardiac dysrhythmia, chronic kidney disease, BMI, smoking, peripheral aneurysms, history of cancer, aortic pathology type, aortic rupture, and type of surgery. Comorbidities and treatment were binary coded, sex was one-hot encoded, and aortic pathology was ordinal structured (dissections, thoracoabdominal, thoracic, and abdominal). Variables were chosen based on their effects on outcomes in aortic surgery [4] and life expectancy in general. Optimization suggested different k-values using the *NbClust* package with the same level of recommendation (5 metrics suggested 2 or 3 clusters each). Clusters were generated with the centers set to 2 and nstart set to 25. Optimization of k and cluster formation was carried out with a set seed. Since clustering requires no missing data, 18 (1%) out of the 1819 patients with aortic pathologies undergoing invasive treatment had to be excluded from the dataset. This was considered reasonable with a low risk of introducing significant selection bias.

Sensitivity Analyses

Multiple steps were included in the analytical workflow to ensure the robustness of our findings. First, 2 different cluster analyses were conducted fitting 2 and 3 clusters to investigate whether the same set of variables would separate clusters from each other. Second, age cutoffs used in the cluster analyses were shifted to 65/75 years to investigate the consistency of the cluster-derived results. Third, survival analyses were done using different subsets of the study population, stratifying by clusters and after removing deaths within the perioperative period (30 d). Fourth, varying sets of covariates were introduced to adjust survival models and establish a set of parameters for the final model. Potential confounders included demographics such as age at surgery and sex, comorbidities used during clustering, smoking history, aortic pathology, as well as type and year of surgery.

Statistical Analysis

Descriptive statistics were used to summarize baseline characteristics, with continuous variables reported as mean (SD) or median (IQR), depending on their distribution, and categorical variables presented as frequencies and percentages. Survival analyses (postoperative survival in days to all-cause death) were conducted using the Kaplan-Meier curves to estimate survival, stratified by clusters or indices derived from the clustering analysis. Log-rank tests were used to compare survival curves

XSL•FO
RenderX

between groups. Cox regression models were used to fit survival models with varying covariates as mentioned above and Wald Test *P* value and 95% CI were assessed for each covariate. Statistical significance was defined as *P* value <.05. Cubic splines (4 degrees of freedom) were fitted based on the survival models to calculate and visualize age-related hazard ratios. All analyses were conducted using R. Multimedia Appendix 1 illustrates the analytic workflow and included packages [20-25]. The code outline can be accessed at the public GitHub repository [26] or in Multimedia Appendix 1. The actual analyses were conducted in the UK Biobank Research Analysis Platform.

Results

Study Population

The study included 1801 individuals, predominantly male (n=1500, 82.3%) with a median age at surgery of 69 (IQR 65 - 74) years. Coronary heart disease and arterial hypertension were highly prevalent, affecting 56.6% (n=1019) and 81.5% (n=1468) of participants, respectively. Heart failure and cardiac dysrhythmia were observed in 36.7% (n=661) and 53.1% (n=956) of cases. Peripheral aneurysms occurred in 12.3% (n=222), and 82% (n=1477) reported a history of smoking. AAs were primarily abdominal (n=1204, 66.9%), followed by thoracic (n=394, 21.9%) and thoracoabdominal (n=9, 0.5%), with dissections reported in 10.8% (n=194). Ruptures were noted in 9.9% (n=179) of cases (Table 1). In summary, 29.5% of patients did not survive the observational period.

Table . Demographic and clinical characteristics of the study population (N=1801).

Characteristics	Total (N=1801)
Sex, n (%)	
Female	301 (16.7)
Male	1500 (82.3)
Age at surgery, mean (IQR)	69 (65-74)
Coronary heart disease, n (%)	1019 (56.6)
Arterial hypertension, n (%)	1468 (81.5)
Heart failure, n (%)	661 (36.7)
Cardiac dysrhythmia, n (%)	956 (53.1)
Chronic kidney disease, n (%)	726 (40.3)
BMI, median (IQR)	28.0 (25.6 - 30.9)
Peripheral aneurysms, n (%)	222 (12.3)
Smoking (ever), n (%)	1477 (82)
History of cancer, n (%)	229 (12.7)
Aortic aneurysm, n (%)	
Abdominal	1204 (66.9)
Thoracic	394 (21.9)
Thoracoabdominal	9 (0.5)
Dissection	194 (10.8)
Rupture	179 (9.9)

Unsupervised Cluster Analysis

Cluster analysis was performed fitting 2 and 3 cluster solutions (see “Methods” section). When fitting 2 clusters, age appeared to be the primary factor separating groups. Cluster 1/2 had a younger median age at surgery (62, IQR 58-65 y) compared to cluster 2/2 (73, IQR 70-76 y). Both clusters had similar rates of arterial hypertension (approximately 81%) and smoking (approximately 82%). Differences included a higher prevalence of cardiac dysrhythmia and chronic kidney disease in cluster 2/2 (55% and 42.8%, respectively) compared to cluster 1/2 (49.3% and 35.5%, respectively), corresponding to the higher age. Abdominal aneurysms predominated in both clusters, but cluster 1/2 showed a higher frequency of dissections (14.7% vs 8.7%) and strictly thoracic aneurysms (26.8% vs 19.3%), while ruptures were comparable (approximately 10%; Table 2)

XSL•FO
RenderX

Table . Fitting 2 clusters to the study population. Cluster 1/2 consists of younger patients (at the time of surgery) with more thoracic pathologies.

Variable	Cluster 1/2 (n=612)	Cluster 2/2 (n=1189)
Female (%), proportion (95% CI)	14.9 (12.0 - 18.4)	17.7 (15.6 - 19.9)
Age at surgery, median (IQR)	62 (58-65)	73 (70-76)
Coronary heart disease (%), proportion (95% CI)	58.8 (54.8 - 62.7)	55.4 (52.5 - 58.2)
Arterial hypertension (%), proportion (95% CI)	81.2 (77.8 - 84.2)	81.7 (79.4 - 83.8)
Heart failure (%), proportion (95% CI)	37.7 (33.9 - 41.7)	36.2 (33.5 - 39.0)
Cardiac dysrhythmia (%), proportion (95% CI)	49.3 (45.3 - 53.4)	55 (52.1 - 57.8)
Chronic kidney disease (%), proportion (95% CI)	35.5 (31.7 - 39.4)	42.8 (40.0 - 45.6)
BMI, median (IQR)	28.1 (25.5 - 31.3)	2 (25.6 - 30.6)
Peripheral aneurysms (%), proportion (95% CI)	12.7 (10.1 - 15.9)	12.1 (10.4 - 14.1)
Smoking (ever) (%), proportion (95% CI)	80.4 (76.9 - 83.5)	83.7 (81.5 - 85.7)
Connective tissue disease (%), proportion (95% CI)	3.1 (1.9 - 5.0)	3.4 (2.5 - 4.6)
Aortic aneurysm		
Abdominal, %	57.8	71.5
Thoracic, %	26.8	19.3
Thoracoabdominal, %	0.7	0.4
Dissection, %	14.7	8.7
Rupture (%), proportion (95% CI)	9.2 (7.0 - 11.9)	10.3 (7.0 - 11.9)

When grouped into 3 clusters, age differences were more pronounced: cluster 1/3 (n=700; median age at surgery 75, IQR 73-78 y), cluster 2/3 (n=811; 68, IQR 65-69 y), and cluster 3/3 (n=290; 57, IQR 53-60 y). Cluster 3/3 had the highest rates of heart failure (43.8%) and thoracic aneurysms (34.1%) and the lowest smoking prevalence (73.5%). Cluster 1/3 had the highest aortic rupture rate (12%), with abdominal aneurysms being most common (73.9%). Cluster 2/3 displayed intermediate characteristics across many variables, including rupture rate (8.8%) and the distribution of aortic pathologies (Table 3). Due to the predominant separation between clusters by age, this was further evaluated. It was identified that the first cluster analysis split patients at age 66/67 years, while the second cluster analysis used 61/62 years and 71/72 years as thresholds. All patients below or above the respective thresholds were exclusively grouped in the corresponding clusters.

XSL•FO
RenderX

Table . Fitting 3 clusters in the study population. Cluster 1/3 includes the oldest patients at the time of surgery. Cluster 3/3 contains the youngest patients who have a higher rate of cardiac comorbidities and more thoracic pathologies.

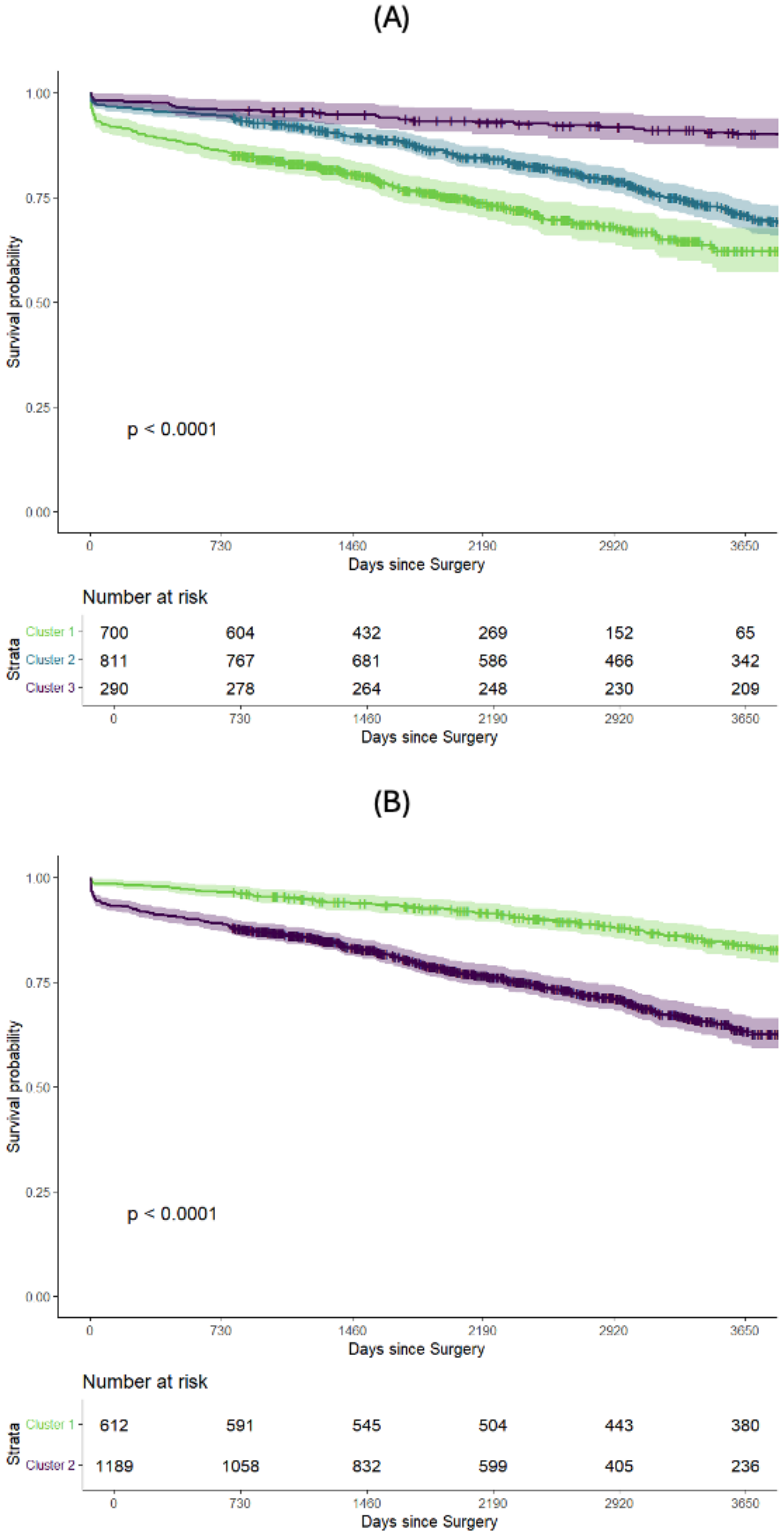
Variable	Cluster 1/3 (n=700)	Cluster 2/3 (n=811)	Cluster 3/3 (n=290)
Female (%), proportion (95% CI)	18 (15.3 - 21.0)	15.4 (13.0 - 18.2)	17.2 (13.3 - 22.0)
Age at surgery, median (IQR)	75 (73-78)	68 (65-69)	57 (53-60)
Coronary heart disease (%), proportion (95% CI)	52.1 (48.4 - 55.9)	59.6 (56.1 - 62.9)	59 (53.2 - 64.5)
Arterial hypertension (%), proportion (95% CI)	82.7 (79.8 - 85.3)	81 (78.2 - 83.6)	80 (75.0 - 84.3)
Heart failure (%), proportion (95% CI)	36.1 (32.6 - 39.8)	34.6 (31.4 - 38.0)	43.8 (38.1 - 49.6)
Cardiac dysrhythmia (%), proportion (95% CI)	56.4 (52.7 - 60.1)	49.2 (45.7 - 52.7)	55.9 (50.1 - 61.5)
Chronic kidney disease (%), proportion (95% CI)	45.7 (42.0 - 49.5)	37.9 (34.6 - 41.3)	34.1 (28.9 - 39.8)
BMI, median (IQR)	28.0 (25.7 - 30.5)	28.0 (25.4 - 31.1)	28.3 (25.3 - 31.3)
Peripheral aneurysms (%), proportion (95% CI)	13.6 (11.1 - 16.5)	10.9 (8.9 - 13.3)	13.4 (9.8 - 18.1)
Smoking (ever) (%), proportion (95% CI)	83.4 (80.5 - 86.0)	85.1 (82.4 - 87.4)	73.5 (68.1 - 78.3)
Connective tissue disorder (%), proportion (95% CI)	2.9 (1.8 - 4.5)	7.4 (5.8 - 9.4)	5.2 (3.1 - 8.5)
Aortic aneurysm			
Abdominal, %	73.9	67.7	46.9
Thoracic, %	17.7	21.1	34.1
Thoracoabdominal, %	0.4	0.5	0.7
Dissection, %	7.7	10.7	18.3
Rupture (%), proportion (95% CI)	12.0 (9.7 - 14.7)	8.8 (6.9 - 11.1)	8.3 (5.6 - 12.1)

Cluster Survival Analyses

Survival analyses were conducted for each cluster analysis. In both cases (2 or 3 clusters), age at death was higher in patients undergoing treatment at a later age, reflective of the selection bias of age at treatment. However, postoperative survival was significantly lower in these clusters (clusters 1/3 and 2/2; Figures 1A and 1B). It appeared that the main phase of separation in both cases lay within the perioperative period, where older clusters were associated with a higher all-cause mortality. Especially when comparing clusters 2/3 and 3/3, differences in long-term mortality appeared to stabilize between clusters.

XSL•FO
RenderX

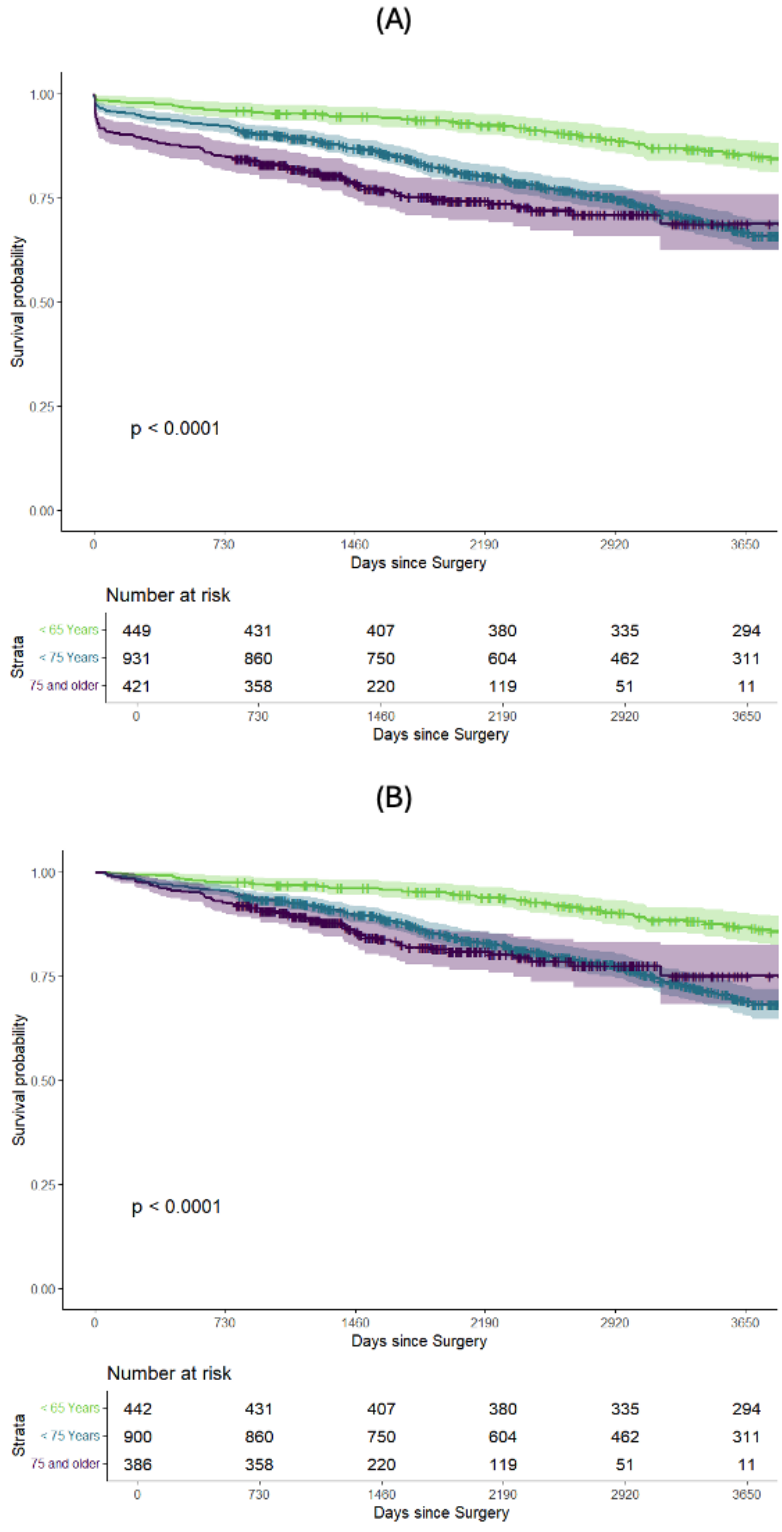
Figure 1. Kaplan-Meier curves illustrating postoperative survival for each cluster. Clusters with older patients at surgery had lower postoperative survival rates. (A) 3 clusters and (B) 2 clusters. The depicted *P* value refers to a log-rank test.



To explore whether these observations would remain constant after shifting age brackets, the study population was divided into 3 subgroups at 65 and 75 years. Postoperative survival was repeatedly lower in older age groups that were characterized by a higher perioperative mortality (Figure 2A). However, after removing deaths within the first 30 postoperative days, the survival estimates of all groups not only converged but, especially regarding the 2 older age groups, there was no evidence for a difference in all-cause mortality rates (Figure 2B).

XSL•FO
RenderX

Figure 2. Kaplan-Meier curve illustrating postoperative survival stratified by age. Patients with higher age at surgery have lower postoperative survival rates. (A) Full observational period and (B) after removing the first 30 postoperative days. The depicted *P* values refer to a log-rank test.



Age and Mortality

To further explore the relationship between age and postoperative survival, univariate survival analyses were conducted in different subsets of the study population (Table 4). In the full study population, age was significantly associated with increased all-cause mortality after invasive aneurysm treatment. When stratifying by clusters, this association faded for the cluster with the oldest patients (cluster 1/3), indicating that in very old patients, additional years of age do not significantly increase mortality risks. However, since the perioperative period was associated with disproportionate mortality in older people, the analyses were repeated removing patients who died during the first 30 postoperative days. Now, there was no evidence for an association between all-cause

XSL•FO
RenderX

mortality and age in both cluster analyses for the oldest clusters (cluster 2/2 and cluster 1/3).

Table . The association between age at surgery and mortality in varying patient groups in a single-predictor survival analysis (N=1801).

Group	Participants (n)	Coefficient	95% CI	*P* value[a]
Study population	1801	1.07	1.05 - 1.08	<.001
Cluster 1/2	612	1.06	1.03 - 1.09	<.001
Cluster 2/2	1189	1.04	1.01 - 1.07	.008
Cluster 1/3	700	1.02	0.96 - 1.07	.55
Cluster 2/3	811	1.07	1.02 - 1.12	.008
Cluster 3/3	290	1.08	1.03 - 1.13	.002
Study population (without first 30 d)	1728	1.06	1.05 - 1.08	<.001
Cluster 1/2 (without first 30 d)	604	1.07	1.03 - 1.10	<.001
Cluster 2/2 (without first 30 d)	1124	1.02	0.99 - 1.06	.19
Cluster 1/3 (without first 30 d)	653	0.97	0.90 - 1.03	.32
Cluster 2/3 (without first 30 d)	790	1.06	1.01 - 1.12	.02
Cluster 3/3 (without first 30 d)	285	1.08	1.03 - 1.13	.003

[a]The *P* value is derived from the Wald test.

Subsequently, the analyses were repeated adjusting for the year of surgery since overall mortality rates have decreased over time and this was therefore considered a relevant confounder (Table 5). There was again strong evidence for an association between age and post-operative mortality, except for cluster 1/3 after deaths were removed in the perioperative period. To fit a final model, varying sets of confounders were introduced, and nonsignificant parameters were removed. The final analyses adjusted for year of surgery, chronic kidney disease, heart failure, and type of aortic pathology (Table 6). However, there was strong evidence for an association between postoperative mortality and age in most subgroups. However, after excluding perioperative deaths, age was no longer statistically significantly associated with all-cause mortality in either cluster 1/3 or cluster 3/3 (the youngest cohort).

XSL•FO
RenderX

Table . The association between age at surgery and mortality in varying patient groups in a survival analysis adjusted for year of surgery (N=1801).

Group	Participants (n)	Coefficient	95% CI	*P* value[a]
Study population	1801	1.08	1.06 - 1.11	<.001
Cluster 1/2	612	1.06	1.03 - 1.09	<.001
Cluster 2/2	1189	1.10	1.06 - 1.15	<.001
Cluster 1/3	700	1.08	1.00 - 1.15	.04
Cluster 2/3	811	1.11	1.05 - 1.17	<.001
Cluster 3/3	290	1.07	1.02 - 1.12	.004
Study population (without first 30 d)	1728	1.08	1.07 - 1.11	<.001
Cluster 1/2 (without first 30 d)	604	1.07	1.03 - 1.10	<.001
Cluster 2/2 (without first 30 d)	1124	1.10	1.06 - 1.15	<.001
Cluster 1/3 (without first 30 d)	653	1.06	0.97 - 1.15	.18
Cluster 2/3 (without first 30 d)	790	1.10	1.04 - 1.17	<.001
Cluster 3/3 (without first 30 d)	285	1.07	1.02 - 1.13	.004

[a]The *P* value is derived from the Wald test.

Table . The association between age at surgery and mortality in varying patient groups in a survival analysis adjusted for year of surgery, chronic kidney disease, heart failure, and type of aortic pathology (N=1801).

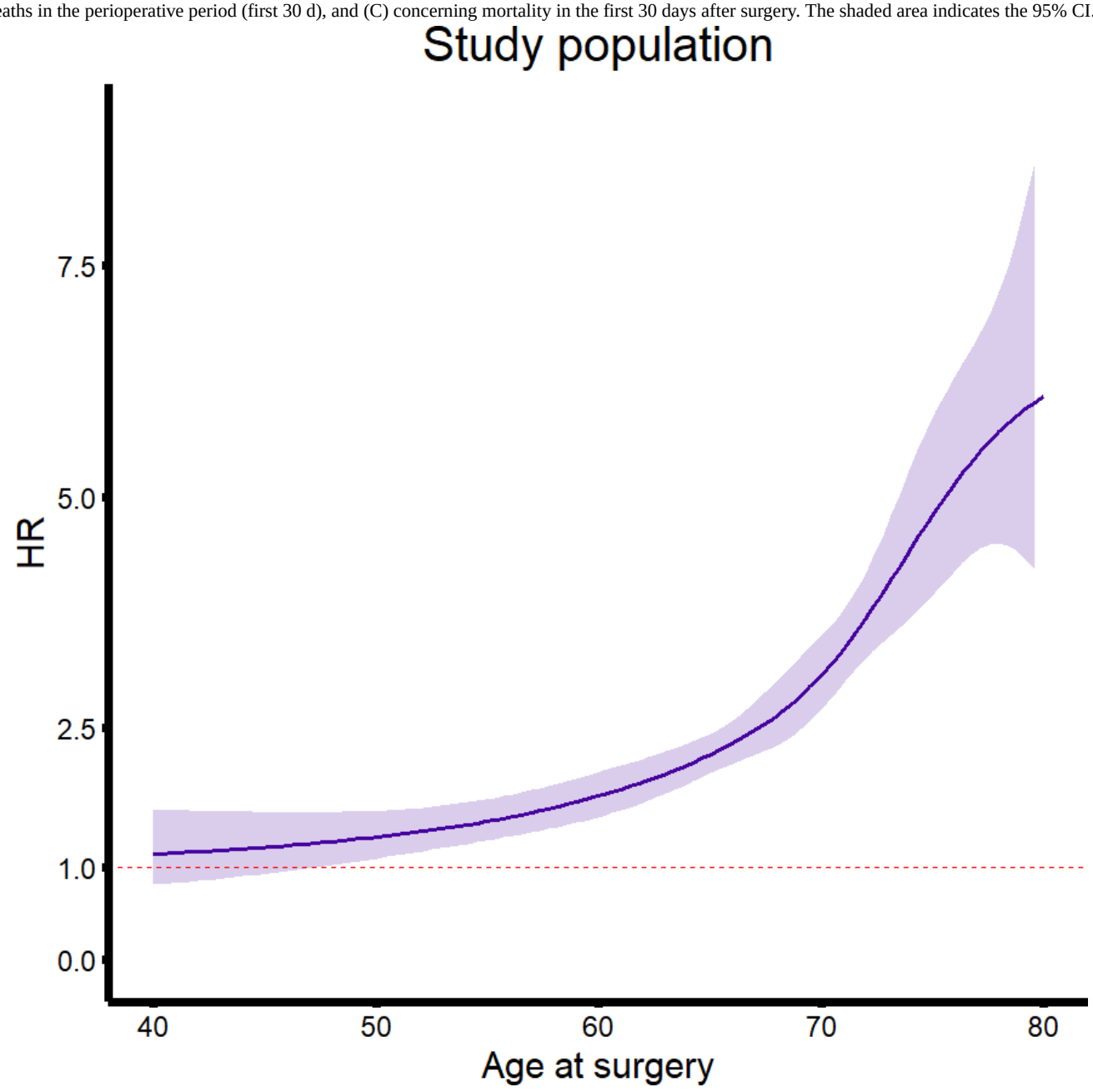
Group	Participants (n)	Coefficient	95% CI	*P* value[a]
Study population	1801	1.09	1.05 - 1.14	<.001
Cluster 1/2	612	1.04	1.01 - 1.08	.02
Cluster 2/2	1189	1.09	1.05 - 1.14	<.001
Cluster 1/3	700	1.08	1.00 - 1.15	.04
Cluster 2/3	811	1.10	1.04 - 1.16	<.001
Cluster 3/3	290	1.05	1.00 - 1.11	.04
Study population (without first 30 d)	1728	1.07	1.05 - 1.10	<.001
Cluster 1/2 (without first 30 d)	604	1.05	1.01 - 1.08	<.001
Cluster 2/2 (without first 30 d)	1124	1.09	1.05 - 1.14	<.001
Cluster 1/3 (without first 30 d)	653	1.07	0.98 - 1.16	.13
Cluster 2/3 (without first 30 d)	790	1.09	1.03 - 1.15	.002
Cluster 3/3 (without first 30 d)	285	1.05	1.00 - 1.11	.06

[a]The *P* value is derived from the Wald test.

Figure 3A-C illustrate the age-specific hazard ratios for all-cause mortality and the total study population, after removing perioperative deaths, and exclusively for death during the perioperative period.

XSL•FO
RenderX

Figure 3. Age-specific hazard ratios (HRs) for (A) all-cause mortality after invasive aortic treatment for the whole study population, (B) after removing deaths in the perioperative period (first 30 d), and (C) concerning mortality in the first 30 days after surgery. The shaded area indicates the 95% CI.



Discussion

The study analyzed 1801 individuals undergoing open surgical or endovascular repair for AAs. The majority were male (82.3%) with a median age of 69 years. Cardiovascular comorbidities and smoking were highly prevalent. AAs were primarily abdominal (66.9%). Overall, 29.5% of patients did not survive the observational period. Unsupervised cluster analysis revealed distinct groupings exclusively based on age. In a 2-cluster model, the clusters were separated at a median age of 62 and 73 years. A 3-cluster model further stratified the population, with median ages of 57, 68, and 75 years. Survival analyses highlighted lower postoperative survival in older clusters relative to treatment, with perioperative mortality disproportionately impacting these clusters. There was strong evidence that age was significantly associated with all-cause mortality overall but showed diminished significance in older clusters after excluding perioperative deaths. Adjusting for confounders confirmed this observation, with age losing significance in the youngest and oldest clusters in later analyses.

That clusters were primarily fitted based on age is not necessarily intuitive from a physician's perspective. We would have expected the presence of ruptures and the specific type of pathology to exert a larger influence on cluster formation, as these factors all have a relevant influence on clinical practice. Interestingly, the strictly data scientific approach of the analysis attributed heterogeneity between the included patients predominantly to age at the time of surgery. We hypothesize that a combination of collinearity, for example, a higher rate of dissections on younger patients, and homogeneity, with similarly distributed comorbidities, could explain this finding. Furthermore, aortic surgery, which is by definition a high-risk

XSL•FO
RenderX

procedure, may homogenize the investigated patient sample (at least for the electively treated patients) as patients must prove a certain level of health to qualify for such an operation. Age is strongly associated with postoperative mortality following invasive treatment of aortic pathologies, but there is solid evidence that this relationship varies across age groups. In the older clusters, the diminished association between age and all-cause mortality after excluding perioperative deaths suggests that perioperative risk plays a critical role in the age-related mortality disparity. Specifically, older patients demonstrated disproportionately higher all-cause mortality within the first 30 days after surgery, implying that the perioperative period is a key driver of mortality in this subgroup. This complies with previous evidence that reported a perioperative survival disadvantage after invasive aneurysm treatment in older patients [27,28].

This differential impact across age groups highlights the complexity of the relationship between age and mortality risk, which does not appear to be linear. While overall mortality increases with age, the cluster analysis revealed thresholds or transition zones in risk rather than a simple uniform progression. The higher perioperative mortality in older patients might reflect their greater vulnerability to surgical stress, emphasizing the importance of targeted perioperative management in this group. Additionally, the lack of a strong age-mortality association in the older clusters after the perioperative period suggests that factors beyond age, such as the type of aneurysm or co-existing conditions, may have a greater influence on long-term survival in these populations. Overall, the data support a nuanced interpretation of age as a risk factor, driven largely by perioperative mortality. Investigating nonlinear relationships between age and outcomes such as mortality in biomedical research has been of growing interest in the previous years [29-32]. This work contributes to the cumulative body of evidence that might gradually induce a paradigm shift in patient risk stratification.

From a statistical perspective, the findings reflect a nonlinear association in which the hazard function for age is dominated by early postoperative events in older patients, leading to attenuation of the age-mortality relationship once perioperative deaths are removed. This pattern is consistent with threshold-like behavior rather than a single global linear term, suggesting that conventional models assuming proportional and constant age effects may obscure important variation across the age distribution.

The clinical implications of this work emphasize the need for tailored perioperative management strategies that account for age-related differences in mortality risk. The necessity to identify solutions for an aging population is not confined to the field of healthcare, but is a subject of considerable interest within it. It is estimated that by 2050, 129.8 million individuals in Europe will be aged 65 years or older [33], with significant implications for the demand on national health care systems and the surgical workforce [34-36]. Therefore, the question of age as a limitation for operative procedures such as AA repair has attracted increasing attention in vascular surgery [27] but is discussed in other fields such as orthopedics with knee arthroplasty [37] as well. A substantial body of research has indicated an elevated risk of perioperative complications for older patients across various fields, including common abdominal [38-40] and vascular procedures [41,42]. In addition to adverse outcomes, surgical treatment in older patients has been demonstrated to be more resource-intensive, characterized by prolonged hospital stays [35,43]. The unsupervised cluster analysis identified distinct age thresholds at surgery, providing a framework for stratifying patients into risk groups. This approach moves beyond arbitrarily defined age-based subgroups, such as octogenarians or nonagenarians, which have traditionally been used as proxies for elderly populations in surgical research and have been previously questioned regarding their use [44].

However, the sensitivity analysis showed that alternative age cutoffs at 65 and 75 years produced similar Kaplan-Meier survival curves, suggesting that the transitions between age-based clusters are fluid rather than strictly defined. Consequently, these thresholds should not be interpreted as rigid cutoffs but rather as approximate transition zones that may guide clinical decision-making. For example, patients in the older age groups are at increased risk for perioperative mortality, highlighting the need for enhanced perioperative care, such as optimization of cardiovascular and respiratory function, and close monitoring during the immediate postoperative period. Similarly, younger patients may benefit from strategies aimed at mitigating long-term risk factors, such as smoking cessation and management of hypertension.

The identification of fluid but meaningful thresholds by the cluster analysis provides a valuable tool for clinicians, balancing the need for structured risk stratification with the recognition of patient-specific variability. From a group-level perspective, comorbidities were rather homogenously distributed among clusters. This may be due to case selection, as the analysis included only patients deemed suitable for surgery. As a result, eligibility for surgical treatment could have a homogenizing effect on this patient sample. Nevertheless, on an individual scale, this is certainly not the case and also requires attention when evaluating invasive treatment for a given patient. Age at surgery can give some guidance when evaluating perioperative and postoperative risks, and they vary for distinct age groups, but other factors still impose meaningful effects on survival as well.

The reliance on observational data introduces the potential for selection bias, particularly evident in the older clusters, where longer median survival age after surgery may simply reflect the fact that these patients underwent surgery at a later age. This delayed treatment may be indicative of slower disease progression, possibly due to reduced exposure to risk factors such as smoking or poorly controlled hypertension, but it complicates direct comparisons between age groups. Additionally, the UK Biobank lacks detailed information on aneurysm morphology, including diameters and growth rates, which are critical for understanding disease severity and progression. Such missing data may limit the granularity of risk stratification and affect the interpretation of clinical outcomes. Data from the UK Biobank are also subject to a volunteer selection bias. Furthermore, the use of administrative and diagnostic coding introduces the possibility of errors or misclassification in identifying aneurysm type, surgical

XSL•FO
RenderX

interventions, and comorbid conditions. Future studies incorporating more comprehensive clinical data and prospective designs could address these limitations and further refine the understanding of age-related risks in aortic aneurysm surgery. This will also address the need for further external validation of these findings.

This study highlights age as one of the key factors in postoperative all-cause mortality following AA surgery, with older patients experiencing higher perioperative mortality. Unsupervised cluster analysis revealed age thresholds that guide risk stratification reflecting transitions in age-related mortality risk.

Funding

Maria Elisabeth Leinweber received funding from the Austrian Society of Vascular Surgery (Wissenschaftsförderung 2024).

Data Availability

Data used in this manuscript are exclusively available through the UK Biobank.

Authors' Contributions

Conceptualization: MEL, AGH
Formal analysis: AGH
Funding acquisition: MEL
Investigation: MEL
Methodology: AGH
Resources: MK
Supervision: AA
Validation: FT, MK
Writing – original draft: MEL, AGH
Writing – review & editing: FT, MK, AA

Conflicts of Interest

None declared.

Multimedia Appendix 1
R code analytic workflow.
[TXT File, 9 KB - aging_v9i1e75611_app1.txt]

References

1. Cho MJ, Lee MR, Park JG. Aortic aneurysms: current pathogenesis and therapeutic targets. Exp Mol Med 2023 Dec;55(12):2519-2530. [doi: 10.1038/s12276-023-01130-w] [Medline: 38036736]
2. Bossone E, Eagle KA. Epidemiology and management of aortic disease: aortic aneurysms and acute aortic syndromes. Nat Rev Cardiol 2021 May;18(5):331-348. [doi: 10.1038/s41569-020-00472-6] [Medline: 33353985]
3. Authors/Task Force Members, Czerny M, Grabenwöger M, et al. EACTS/STS guidelines for diagnosing and treating acute and chronic syndromes of the aortic organ. Ann Thorac Surg 2024 Jul;118(1):5-115. [doi: 10.1016/j.athoracsur.2024.01.021] [Medline: 38416090]
4. Wanhainen A, Van Herzeele I, Bastos Goncalves F, et al. Editor's choice—European Society for Vascular Surgery (ESVS) 2024 clinical practice guidelines on the management of abdominal aorto-iliac artery aneurysms. Eur J Vasc Endovasc Surg 2024 Feb;67(2):192-331. [doi: 10.1016/j.ejvs.2023.11.002] [Medline: 38307694]
5. Sidloff D, Stather P, Dattani N, et al. Aneurysm global epidemiology study: public health measures can further reduce abdominal aortic aneurysm mortality. Circulation 2014 Feb 18;129(7):747-753. [doi: 10.1161/CIRCULATIONAHA.113.005457] [Medline: 24249717]
6. Dias-Neto M, Castro-Ferreira R, Mani K, Freitas A, Leite-Moreira A, Sampaio SM. Nationwide analysis of ruptured abdominal aortic aneurysm in Portugal (2000–2015). Eur J Vasc Endovasc Surg 2020 Jul;60(1):27-35. [doi: 10.1016/j.ejvs.2020.02.024] [Medline: 32307303]
7. Gilmore BF, Scali ST, D'Oria M, et al. Temporal trends and outcomes of abdominal aortic aneurysm care in the United States. Circ Cardiovasc Qual Outcomes 2024 Jun;17(6):e010374. [doi: 10.1161/CIRCOUTCOMES.123.010374] [Medline: 38775052]
8. Sampson UKA, Norman PE, Fowkes FGR, et al. Estimation of global and regional incidence and prevalence of abdominal aortic aneurysms 1990 to 2010. Glob Heart 2014 Mar;9(1):159-170. [doi: 10.1016/j.gheart.2013.12.009] [Medline: 25432125]
9. Isselbacher EM, Preventza O, Hamilton Black J 3rd, et al. 2022 ACC/AHA guideline for the diagnosis and management of aortic disease: a report of the American Heart Association/American College of Cardiology Joint Committee on Clinical

XSL•FO
RenderX

Practice Guidelines. Circulation 2022 Dec 13;146(24):e334-e482. [doi: 10.1161/CIR.0000000000001106] [Medline: 36322642]
10. AlOthman O, Bobat S. Comparison of the short and long-term outcomes of endovascular repair and open surgical repair in the treatment of unruptured abdominal aortic aneurysms: meta-analysis and systematic review. Cureus 2020 Aug 12;12(8):e9683. [doi: 10.7759/cureus.9683] [Medline: 32923276]
11. Schermerhorn ML, Buck DB, O'Malley AJ, et al. Long-term outcomes of abdominal aortic aneurysm in the Medicare population. N Engl J Med 2015 Jul 23;373(4):328-338. [doi: 10.1056/NEJMoa1405778] [Medline: 26200979]
12. Tan X, Jung G, Herrmann E, et al. Sex difference in early mortality after abdominal aortic aneurysm repair. J Vasc Surg 2023 Jun;77(6):1658-1668. [doi: 10.1016/j.jvs.2023.02.001] [Medline: 36773666]
13. Mani K, Björck M, Lundkvist J, Wanhainen A. Improved long-term survival after abdominal aortic aneurysm repair. Circulation 2009 Jul 21;120(3):201-211. [doi: 10.1161/CIRCULATIONAHA.108.832774] [Medline: 19581497]
14. Karthikesalingam A, Grima MJ, Holt PJ, et al. Comparative analysis of the outcomes of elective abdominal aortic aneurysm repair in England and Sweden. Br J Surg 2018 Apr;105(5):520-528. [doi: 10.1002/bjs.10749] [Medline: 29468657]
15. UK Biobank. 2025. URL: https://www.ukbiobank.ac.uk [accessed 2026-02-13]
16. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015 Mar;12(3):e1001779. [doi: 10.1371/journal.pmed.1001779] [Medline: 25826379]
17. Littlejohns TJ, Sudlow C, Allen NE, Collins R. UK Biobank: opportunities for cardiovascular research. Eur Heart J 2019 Apr 7;40(14):1158-1166. [doi: 10.1093/eurheartj/ehx254] [Medline: 28531320]
18. Welsh P, Welsh CE, Jhund PS, et al. Derivation and validation of a 10-year risk score for symptomatic abdominal aortic aneurysm: cohort study of nearly 500 000 individuals. Circulation 2021 Aug 24;144(8):604-614. [doi: 10.1161/CIRCULATIONAHA.120.053022] [Medline: 34167317]
19. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: an R package for determining the relevant number of clusters in a data set. J Stat Softw 2014;61(6):1-36. [doi: 10.18637/jss.v061.i06]
20. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. J Open Source Softw 2019 Nov 21;4(43):1686. [doi: 10.21105/joss.01686]
21. Kassambara A, Kosinski M, Biecek P. Survminer: drawing survival curves using "ggplot2". URL: https://cran.r-project.org/web/packages/survminer/index.html [accessed 2026-02-13]
22. Kassambara A, Mundt F. Factoextra: extract and visualize the results of multivariate data analyses. 2020. URL: https://CRAN.R-project.org/package=factoextra [accessed 2026-02-13]
23. Therneau TM. Survival: survival analysis. 2024. URL: https://CRAN.R-project.org/package=survival [accessed 2026-02-13]
24. Robinson D, Hayes A, Couch S. Broom: convert statistical objects into tidy tibbles. broom. 2025. URL: https://broom.tidymodels.org/reference/broom.html [accessed 2026-02-13]
25. Bates DM, Venables WN. Regression spline functions and classes. 2025. URL: https://r-universe.dev/manuals/splines.html [accessed 2026-02-13]
26. AHOpit/non-uniform-association-between-age-and-mortality-after-aortic-surgery-UK-biobank. GitHub. URL: https://github.com/AHOpit/Non-Uniform-Association-between-Age-and-Mortality-After-Aortic-Surgery-UK-Biobank/ [accessed 2026-02-21]
27. Sonesson B, Dias N, Resch T. Is there an age limit for abdominal aortic aneurysm repair? J Cardiovasc Surg (Torino) 2018 Apr;59(2):190-194. [doi: 10.23736/S0021-9509.18.10344-2] [Medline: 29327568]
28. Han Y, Zhang S, Zhang J, Ji C, Eckstein HH. Outcomes of endovascular abdominal aortic aneurysm repair in octogenarians: meta-analysis and systematic review. Eur J Vasc Endovasc Surg 2017 Oct;54(4):454-463. [doi: 10.1016/j.ejvs.2017.06.027] [Medline: 28822680]
29. Andersen LW, Bivens MJ, Giberson T, et al. The relationship between age and outcome in out-of-hospital cardiac arrest patients. Resuscitation 2015 Sep;94:49-54. [doi: 10.1016/j.resuscitation.2015.05.015] [Medline: 26044753]
30. Nguyen VK, Colacino J, Chung MK, Goallec AL, Jolliet O, Patel CJ. Characterising the relationships between physiological indicators and all-cause mortality (NHANES): a population-based cohort study. Lancet Healthy Longev 2021 Oct;2(10):e651-e662. [doi: 10.1016/s2666-7568(21)00212-9] [Medline: 34825242]
31. Monaghan C, de Andrade Moral R, Power JM. Modelling the non-linear associations between age and health: implications for care. Age Ageing 2024 Sep;53(Suppl 4):afae178.084. [doi: 10.1093/ageing/afae178.084]
32. Qian J, Yuan Y, Shang Z, et al. Association between emergency department to intensive care units time and in-hospital mortality: an analysis of the MIMIC-IV database. BMJ Open 2025 Apr 5;15(4):e090011. [doi: 10.1136/bmjopen-2024-090011] [Medline: 40187785]
33. Population structure and ageing. Eurostat. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_structure_and_ageing [accessed 2025-02-02]
34. Etzioni DA, Liu JH, Maggard MA, Ko CY. The aging population and its impact on the surgery workforce. Ann Surg 2003 Aug;238(2):170-177. [doi: 10.1097/01.SLA.0000081085.98792.3d] [Medline: 12894008]
35. McVeigh TP, Al-Azawi D, O'Donoghue GT, Kerin MJ. Assessing the impact of an ageing population on complication rates and in-patient length of stay. Int J Surg 2013;11(9):872-875. [doi: 10.1016/j.ijsu.2013.07.016] [Medline: 23917211]

XSL•FO
RenderX

36. Ellison EC, Pawlik TM, Way DP, Satiani B, Williams TE. The impact of the aging population and incidence of cancer on future projections of general surgical workforce needs. Surgery 2018 Mar;163(3):553-559. [doi: 10.1016/j.surg.2017.09.035] [Medline: 29179915]
37. Basa CD, Kaya Bicer E, Aydogdu S, Sur H. How long do octogenarians benefit from knee arthroplasty? Cureus 2021 May 13;13(5):e14997. [doi: 10.7759/cureus.14997] [Medline: 34131540]
38. Massarweh NN, Legner VJ, Symons RG, McCormick WC, Flum DR. Impact of advancing age on abdominal surgical outcomes. Arch Surg 2009 Dec;144(12):1108-1114. [doi: 10.1001/archsurg.2009.204] [Medline: 20026827]
39. Merga ZC, Lee JS, Gong CS. Outcomes of gastrectomy for gastric cancer in patients aged >80 years: a systematic literature review and meta-analysis. J Gastric Cancer 2023 Jul;23(3):428-450. [doi: 10.5230/jgc.2023.23.e23] [Medline: 37553130]
40. Chen B, Yu W, Ma Y, et al. Evaluation of the safety and efficacy of perform enterectomy in colorectal cancer patients aged 80 or older. A meta-analysis and a systematic review. Int J Colorectal Dis 2023 Jul 3;38(1):185. [doi: 10.1007/s00384-023-04461-2] [Medline: 37395836]
41. Wübbeke LF, Naves C, Daemen J, Jacobs MJ, Mees BME. Editor's choice—Mortality and major amputation after revascularisation in octogenarians versus non-octogenarians with chronic limb threatening ischaemia: a systematic review and meta-analysis. Eur J Vasc Endovasc Surg 2020 Aug;60(2):231-241. [doi: 10.1016/j.ejvs.2020.04.027] [Medline: 32709468]
42. Nana P, Spanos K, Brotis A, Fabre D, Mastracci T, Haulon S. Systematic review on early and follow-up mortality rate in octogenarians treated with a fenestrated and/or branched endovascular aortic repair. J Endovasc Ther 2025 Jun;32(3):567-577. [doi: 10.1177/15266028231182798] [Medline: 37341255]
43. Gelsomino S, Lorusso R, Livi U, et al. Cost and cost-effectiveness of cardiac surgery in elderly patients. J Thorac Cardiovasc Surg 2011 Nov;142(5):1062-1073. [doi: 10.1016/j.jtcvs.2011.02.013] [Medline: 21492881]
44. Hofmann AG, Leinweber ME, Assadian A, Falkensammer J, Taher F. The effect of age on peri-operative outcomes after FEVAR. J Clin Med 2023 Jun 5;12(11):3858. [doi: 10.3390/jcm12113858] [Medline: 37298053]

Abbreviations

AA: aortic aneurysm
ICD-10: *International Classification of Diseases, Tenth Revision*
NEC: not elsewhere classified
NHS: National Health Service
OPCS-4: *Office of Population, Censuses and Surveys: Classification of Interventions and Procedures*
WHO: World Health Organization

Edited by Q Fan; submitted 07.Apr.2025; peer-reviewed by H Wu, RA Moral, YT Cheung; revised version received 05.Dec.2025; accepted 17.Dec.2025; published 11.Mar.2026.

Please cite as:
Leinweber ME, Taher F, Kliewer M, Assadian A, Hofmann AG
Data-Driven Phenotyping Reveals Nonuniform Association Between Age and Mortality After Aortic Surgery: Retrospective Cohort Study of UK Biobank Data
JMIR Aging 2026;9:e75611
URL: https://aging.jmir.org/2026/1/e75611
doi:10.2196/75611

© Maria Elisabeth Leinweber, Fadi Taher, Miriam Kliewer, Afshin Assadian, Amun Georg Hofmann. Originally published in JMIR Aging (https://aging.jmir.org), 11.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations of Solid Fuel Use and Circadian Rhythm Syndrome With Physical Function and Muscle Strength in Middle-Aged and Older Adults: Nationwide Cohort Study in China

Haochen Wang[1,2*], MM; Pengsen Mou[3,4*], MM; Donghao Liu[1*], PhD; Ze Zhao[1], MM; Wenzheng Chen[1], MM; Yuxin Yao[3,4], MM; Haoyan Shi[1], MM; Jiaxin Guan[3,4], MM; Jing Dong[3,4*], PhD; Yingliang Wei[1], PhD

[1]Department of Orthopaedics, Shengjing Hospital, China Medical University, Shenyang, Liaoning, China

[2]Department of Orthopedics, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

[3]Key Laboratory of Environmental Stress and Chronic Disease Control and Prevention, Ministry of Education, China Medical University, Shenyang, Liaoning, China

[4]Department of Occupational and Environmental Health, School of Public Health, China Medical University, Shenyang, Liaoning, China

[*]these authors contributed equally

Corresponding Author:
Yingliang Wei, PhD
Department of Orthopaedics
Shengjing Hospital
China Medical University
36, San Hao Street
Shenyang, Liaoning, 110004
China
Phone: 86 18900911268
Email: ylwei@cmu.edu.cn

Abstract

Background: The relationships among circadian rhythm syndrome, physical function, and muscle strength remain unclear.

Objective: This study aimed to demonstrate the separate and combined deleterious effects of solid fuel use and circadian rhythm syndrome on physical function and muscle strength.

Methods: We used data from the China Health and Retirement Longitudinal Study cohort. The study population consisted of participants who underwent comprehensive assessments of metabolism, circadian rhythm, indoor air pollution, physical function, and muscle strength at the initial evaluation. Muscle strength was assessed using repeated grip strength measurements, and physical function was assessed using a composite score of muscle strength, physical performance, and balance. Circadian rhythm syndrome was derived from the 5 diagnostic components of metabolic syndrome combined with sleep duration and depression. Logistic regression and linear mixed models were used to assess the relationships among solid fuel use, circadian rhythm syndrome, physical function, and muscle strength. Furthermore, we analyzed the mediating role of circadian rhythm syndrome and its combined effect with solid fuel use on physical function and muscle strength.

Results: A total of 7934 participants were included in the study, most of whom used solid fuels. Solid fuel use was positively associated with circadian rhythm syndrome (odds ratio [OR] 1.078, 95% CI 1.031-1.125; P<.05). Circadian rhythm syndrome was found to be a significant risk factor for impairment of physical function (β=−0.475; P<.05) and muscle strength (β=−0.304; P<.05). Participants who used solid fuels and had circadian rhythm syndrome needed to pay more attention to changes in physical function (β=−0.698; P<.05) and muscle strength (β=−0.332; P<.05). A significant interaction was observed between solid fuel use and circadian rhythm syndrome on physical function ($P_{interaction}$=.03) and muscle strength ($P_{interaction}$=.02). Circadian rhythm syndrome partially mediated the association between solid fuel use and physical function, accounting for 2.51% of the total effect.

Conclusions: Circadian rhythm syndrome exacerbates the adverse effects of solid fuel use on physical function and muscle strength. Fuel cleanliness and regular work and rest habits are crucial for the health of middle-aged and older adults.

(JMIR Aging 2026;9:e78352) doi:10.2196/78352

XSL•FO
RenderX

KEYWORDS

age-related health decline; circadian rhythm syndrome; household air pollution; muscle strength; physical function

Introduction

Physical function assesses the ability to perform basic behavioral activities in terms of muscle strength, upper and lower extremity function, and balance [1]. Maintaining adequate physical function preserves the ability to live independently, optimizes quality of life, improves social participation, and reduces the risk of disability [2]. Low muscle strength and reduced physical function are associated with frailty, an increased risk of falls and fractures, and consequently, high morbidity and mortality [3-5]. Given that there are currently no effective pharmacological treatments to delay age-related decline in physical function [6], it is important to explore the risk factors associated with declining physical function and identify high-risk populations that are more likely to benefit from interventions to mitigate the decline in physical function.

The circadian rhythm is a biological rhythm with a cycle of approximately 24 hours that adapts to environmental changes and maintains stable physiological functions [7]. The suprachiasmatic nucleus of the hypothalamus is the central pacemaker of circadian rhythms and maintains synchronization with the external environment through neurohumoral signals in coordination with molecular oscillators in peripheral tissues [8]. Circadian rhythm is highly tissue-specific; each tissue of the human body, including the musculoskeletal system, has its own unique biological clock, which maintains physical function to meet daily demands by regulating the transcription and translation of clock genes [9]. The robustness of circadian rhythms decreases with age and the disruption of circadian rhythms; this is known as circadian rhythm syndrome and is a mechanism that potentially leads to an increased risk of various diseases [10,11]. Although maintaining a normal circadian rhythm is critical for skeletal muscle metabolic homeostasis, muscle fiber number and volume, and joint health [12], the impact of circadian rhythm syndrome on physical function and muscle strength lacks high-quality epidemiological studies for confirmation. As such, it is necessary to elucidate the relationship between circadian rhythm syndrome, physical function, and muscle strength and to identify special populations susceptible to circadian rhythm syndrome.

Household air pollution from solid fuel use is a common environmental safety problem worldwide [13]. According to statistics, the use of solid fuels is more common in developing countries. In China, approximately 450 million people rely heavily on solid fuels [14]. Numerous studies have shown that solid fuel use significantly increases the risk of physical and mental illnesses, such as cardiovascular disease and depression [14,15]. Additionally, our previous research also found adverse effects of solid fuel use and a combined adverse effect of solid fuel use and airborne particulate matter on physical function [1]. However, the association between solid fuel use and circadian rhythm syndrome remains unclear. Furthermore, solid fuel use and circadian rhythm syndrome are both risk factors for aging [16], and their combined effects on physical function and muscle strength require further elucidation.

To address these gaps, we performed a nationwide population-based cohort study to assess the independent and combined associations of solid fuel use and circadian rhythm syndrome with physical function and muscle strength in middle-aged and older adults. This study further aimed to explore the potential role of circadian rhythm syndrome in the relationship between solid fuel use and functional outcomes. The findings are expected to provide evidence relevant to public health strategies targeting environmental and lifestyle risk factors in aging populations.

Methods

Study Design and Data Source

This study used data from the China Health and Retirement Longitudinal Study (CHARLS), a large nationally representative cohort study of middle-aged and older adults in China [17]. CHARLS randomly selected 150 counties from 28 provinces in China, with a total of 17,708 participants as baseline and followed up every 2 years using face-to-face computer-assisted personal interviews (wave 2 in 2013, wave 3 in 2015, wave 4 in 2018, and wave 5 in 2020).

In the CHARLS survey, comprehensive assessments of circadian rhythm syndrome, physical function, and muscle strength were available only in the 2011, 2013, and 2015 waves. Therefore, data from 17,708 participants in 2011 (wave 1) were used as the baseline, with longitudinal analyses conducted using data from 2015 (wave 3). Exclusion criteria were: (1) being aged younger than 45 years (n=321); (2) no information on circadian rhythm syndrome (n=6067), physical function (n=2712), or household fuel use (n=674); and (3) no follow-up data (n=2916), comprising participants with missing information on circadian rhythm syndrome (n=929), physical function (n=1681), and household fuel use (n=306). A total of 7934 participants were finally included in the cross-sectional study, and 5018 participants in the longitudinal analysis (Multimedia Appendix 1). As shown in Multimedia Appendix 2, all participants in this study were distributed across 126 cities in 28 provinces in China, with the majority located in southern, central, and western China.

Household Solid Fuel Use Measurement

Household solid fuel use was assessed using a structured questionnaire. Participants were considered to use solid fuels when they answered “coal,” “crop residue/wood burning,” or “other” for heating and/or cooking, and clean fuels were considered to be used when they answered “natural gas,” “marsh gas,” “liquefied petroleum gas,” “solar,” or “electricity” for heating and/or cooking. Of note, due to the lack of data on the type of heating fuel used during follow-up, we conducted a longitudinal analysis of cooking fuels only and used participants using clean fuels as a reference to assess the impact of solid fuel use on circadian rhythm syndrome, physical function, and muscle strength. For the cross-sectional analysis, since data on both cooking and heating fuels were available, we adopted a

XSL•FO
RenderX

3-category classification to comprehensively reflect household fuel use patterns: clean fuel (reference group), partial solid fuel use (solid fuel used for at least 1 of cooking or heating), and solid fuel use.

Circadian Rhythm Syndrome Assessment

Based on previous studies, we used circadian rhythm syndrome as assessed by a combination of the 5 variables of the International Diabetes Federation assessing metabolic syndrome (elevated blood pressure, blood glucose, triglycerides, waist circumference, and lowered high-density lipoprotein cholesterol), combined with sleep duration and depression [11,18]. Detailed diagnostic criteria are shown in Multimedia Appendix 3 [19,20]. Participants were diagnosed with circadian rhythm syndrome if they met 4 of the 7 components [11,18].

Physical Function Assessment

Grip strength is a validated proxy for overall skeletal muscle strength and is widely used for sarcopenia screening in middle-aged and older adults [21-23]; therefore, it was used to assess muscle strength in this study. Each participant held the TM WL-1000 dynamometer (Nantong Yuejian Physical Measurement Equipment Co Ltd) in both hands successively at full strength for 2 separate measurements. The maximum of the 4 recordings was used and normalized by body weight [24]. If the participant was unable to perform the test with 1 hand due to special reasons, the maximum value of the other hand was recorded. In addition, the physical performance of each participant was assessed using gait speed and the chair stand test [23]. The above 3 tests were assigned a score of 1 to 4 to the participant based on the results from the first quartile to the fourth quartile. If the participant was unable to complete the test, the test was assigned a score of 0. Physical function was assessed using the above 3 tests plus the balance test [1]. For the balance test, participants were asked to stand and hold for 10 seconds in 3 increasingly difficult positions with their feet side-by-side, in semi-tandem, and in full-tandem. Scores were assigned based on the results of the tests (0: side-by-side <10 seconds; 1: side-by-side ≥10 seconds but semi-tandem <10 seconds; 2: semi-tandem ≥10 seconds and full-tandem <3 seconds; 3: semi-tandem ≥10 seconds and full-tandem between 3 and 10 seconds; 4: semi-tandem ≥10 seconds and full-tandem ≥10 seconds). Based on these assessments, the total physical function score was calculated as the sum of 4 domains: muscle strength, gait speed, chair-stand performance, and balance. Muscle strength, gait speed, and chair-stand test were each scored from 0 to 4 according to quartile-based performance, while balance was scored from 0 to 4 based on predefined criteria. Accordingly, the composite physical function score ranged from 0 to 16, with higher scores indicating better overall physical function [25-27].

Covariates

The following basic characteristics were included in the main analysis [1,11,18]: age, gender, education level, marital status, urbanization level, smoking status, drinking status, annual income, indoor temperature, housing type, BMI, and a comprehensive set of chronic diseases. Detailed information is provided in Multimedia Appendix 4.

Statistical Analysis

Statistical Summary of Participant Characteristics

The baseline characteristics of the participants were summarized based on the types of cooking and heating fuels used. Continuous variables are presented as mean (SD), while categorical variables are expressed as frequencies (percentage). Differences between groups were assessed using 2-tailed *t* tests for continuous variables and chi-square tests for categorical variables.

Cross-Sectional and Longitudinal Analysis of Separate Effects

We examined the relationships among solid fuel use, circadian rhythm syndrome, and physical function scores in both cross-sectional and longitudinal studies. Detailed information is provided in Multimedia Appendix 4. The missing rate for most individual characteristics was <5%; thus, missing values were imputed using a predictive mean-matching method to avoid excessive sample loss [18,28]. Upon comparing the summary statistics of the imputed dataset with those of the original dataset, no significant differences in the basic characteristics were observed (Multimedia Appendix 5). All statistical analyses were performed using R (version 4.3.1; R Foundation for Statistical Computing). Two-sided tests with a *P* value <.05 were considered significant.

Combined Effects and Interaction Analysis

We created a new variable by combining fuel type and circadian rhythm syndrome, resulting in 4 categories (2 × 2): fuel type (clean or solid fuels) and circadian rhythm syndrome (yes or no), to explore the collective effects of the 2 exposures on physical function scores. Moreover, we added multiplicative interaction terms to the model to analyze the interaction effect of solid fuel use and circadian rhythm syndrome on physical function scores. A *P* value <.05 for the interaction indicated the presence of a multiplicative interaction.

Stratified Analysis

Concurrently, we stratified by fuel type and circadian rhythm syndrome to explore the impact of other exposures on physical function scores, as well as the scores of the 4 tests.

Mediation Analysis

Finally, we conducted a mediation analysis using the R package *mediation* to assess whether circadian rhythm syndrome mediates the association between solid fuel use and physical function scores. Detailed information and sensitivity analyses are provided in Multimedia Appendix 4.

Ethical Considerations

Ethical approval was obtained from the Ethics Review Committee of Peking University (approval number IRB00001052-11015), and all participants consented to participate in this study.

XSL•FO
RenderX

Results

Statistical Summary of Participant Characteristics

Table 1 shows the baseline characteristics of the 7934 participants enrolled in the study. The findings revealed that solid fuel was the predominant household fuel, with 4820 (60.8%) and 4827 (60.8%) participants using it for cooking and heating, respectively. Participants who used solid fuels were generally older (aged 59.6, SD 8.9 years), female (2522/4820, 52.3%), less educated (4563/4820, 94.7%), living in rural areas (3827/4820, 79.4%), married (4078/4820, 84.6%), never smokers (3240/4820, 67.2%), never alcohol drinkers (3231/4820, 67%), with lower annual household incomes (3372/4820, 70%), and living in low-rise buildings (3639/4820, 75.5%) with unsuitable indoor temperatures. In terms of health status, participants who used solid fuels were more likely to have various chronic conditions, had a higher prevalence of circadian rhythm syndrome (1923/4820, 39.8%), and had lower physical function (12, SD 2.4), muscle strength (2.5, SD 1.1), gait speed (3.3, SD 1.1), chair stand test (2.4, SD 1.1), and balance (3.9, SD 0.5) scores. Characteristics of the participants in the longitudinal analysis are presented in Multimedia Appendix 6. Multimedia Appendix 7 illustrates the national distribution of solid fuel use, prevalence of circadian rhythm syndrome, physical function, and muscle strength. Northern China exhibited a higher prevalence of circadian rhythm syndrome and solid fuel dependency than Southern China, whereas physical function and grip strength scores were lower in the North.

XSL•FO
RenderX

Table 1. Baseline characteristics of the study participants according to household fuel types. Continuous and categorical variables were compared using analyses of independent sample Student 2-tailed *t* test and chi-square tests, respectively.

Variable	Cooking fuels		*P* value	Heating fuels		*P* value
	Clean (n=3114)	Solid (n=4820)		Clean (n=3107)	Solid (n=4827)	
Demographic characteristics						
Age (years), mean (SD)	57.3 (8.8)	59.6 (8.9)	<.001	58.5 (9.1)	58.9 (8.8)	.04
Gender, n (%)			.99			.71
Male	1484 (47.7)	2298 (47.7)		1473 (47.4)	2309 (47.8)	
Female	1630 (52.3)	2522 (52.3)		1634 (52.6)	2518 (52.2)	
Education level, n (%)			<.001			<.001
Middle school or below	2723 (87.4)	4563 (94.7)		2783 (89.6)	4503 (93.3)	
High or vocational school	349 (11.2)	240 (5)		282 (9.1)	307 (6.4)	
Above high school	42 (1.3)	17 (0.4)		42 (1.4)	17 (0.4)	
Residence, n (%)			<.001			<.001
Urban	1529 (49.1)	993 (20.6)		1376 (44.3)	1146 (23.7)	
Rural	1585 (50.9)	3827 (79.4)		1731 (55.7)	3681 (76.3)	
Marital status, n (%)			.14			.02
Married and living	2607 (83.7)	4078 (84.6)		2583 (83.1)	4102 (85)	
Married but separated	175 (5.6)	223 (4.6)		181 (5.8)	217 (4.5)	
Single or other	332 (10.7)	519 (10.8)		343 (11)	508 (10.5)	
Smoking status, n (%)			.005			.001
Never	2187 (70.2)	3240 (67.2)		2191 (70.5)	3236 (67)	
Current	927 (29.8)	1580 (32.8)		916 (29.5)	1591 (33)	
Alcohol drinking status, n (%)			.38			.63
Never	2058 (66.1)	3231 (67)		2081 (67)	3208 (66.5)	
Current	1056 (33.9)	1589 (33)		1026 (33)	1619 (33.5)	
Household characteristics, n (%)						
Housing type			<.001			<.001
One-story	1728 (55.5)	3639 (75.5)		1546 (49.8)	3821 (79.2)	
Multistory	1386 (44.5)	1181 (24.5)		1561 (50.2)	1006 (20.8)	
Indoor temperature			<.001			<.001
Bearable	2716 (87.2)	3943 (81.8)		2702 (87)	3957 (82)	
Hot	320 (10.3)	684 (14.2)		344 (11.1)	660 (13.7)	
Cold	78 (2.5)	193 (4)		61 (2)	210 (4.4)	
Annual income, RMB[a]			<.001			<.001
≤20,000	1523 (48.9)	3372 (70)		1633 (52.6)	3262 (67.6)	
>20,000	1591 (51.1)	1448 (30)		1474 (47.4)	1565 (32.4)	
Health status						
BMI, mean (SD), kg/m^2	24 (8.2)	23.4 (13.4)	.03	23.8 (11.7)	23.5 (11.6)	.24
Cancer, n (%)	28 (0.9)	38 (0.8)	.60	26 (0.8)	40 (0.8)	.97
Chronic lung disease, n (%)	263 (8.4)	541 (11.2)	<.001	283 (9.1)	521 (10.8)	.02
Heart problem, n (%)	336 (10.8)	543 (11.3)	.51	273 (8.8)	606 (12.6)	<.001
Stroke, n (%)	61 (2)	543 (2.3)	.25	56 (1.8)	118 (2.4)	.06
Psychiatric problems, n (%)	29 (0.9)	63 (1.3)	.13	28 (0.9)	64 (1.3)	.08

XSL•FO
RenderX

Variable	Cooking fuels		*P* value	Heating fuels		*P* value
	Clean (n=3114)	Solid (n=4820)		Clean (n=3107)	Solid (n=4827)	
Arthritis, n (%)	1021 (32.8)	1766 (36.6)	<.001	1025 (33)	1762 (36.5)	.001
Liver disease, n (%)	98 (3.1)	167 (3.5)	.44	104 (3.3)	161 (3.3)	>.99
Kidney disease, n (%)	161 (5.2)	309 (6.4)	.02	177 (5.7)	293 (6.1)	.49
Digestive disease, n (%)	670 (21.5)	1186 (24.6)	.001	731 (23.5)	1125 (23.3)	.82
Asthma, n (%)	115 (3.7)	260 (5.4)	<.001	120 (3.9)	255 (5.3)	.004
Memory disorder, n (%)	30 (1)	69 (1.4)	.07	32 (1)	67 (1.4)	.16
Circadian rhythm syndrome, n (%)	1181 (37.9)	1894 (39.3)	.22	1152 (37.1)	1923 (39.8)	.01
Circadian rhythm syndrome component, n (%)						
Reduced HDL-C[b]	1278 (41)	1839 (38.2)	.01	1200 (38.6)	1917 (39.7)	.33
Elevated waist circumference	1922 (61.7)	2561 (53.1)	<.001	1815 (58.4)	2668 (55.3)	.006
Hyperglycemia	1697 (54.5)	2837 (58.9)	<.001	1735 (55.8)	2799 (58)	.06
Raised blood pressure	1854 (59.5)	2830 (58.7)	.47	1809 (58.2)	2875 (59.6)	.24
Elevated triglyceride	940 (30.2)	1291 (26.8)	.001	870 (28)	1361 (28.2)	.85
Short sleep duration	861 (27.6)	1529 (31.7)	<.001	920 (29.6)	1470 (30.5)	.42
Depressive symptoms	921 (29.6)	2084 (43.2)	<.001	964 (31)	2041 (42.3)	<.001
Physical function, mean (SD)						
Physical function score	12.5 (2.4)	12 (2.5)	<.001	12.5 (2.4)	12 (2.4)	<.001
Muscle strength score	2.5 (1.1)	2.5 (1.1)	.24	2.5 (1.1)	2.5 (1.1)	.021
Gait speed score	3.5 (1)	3.2 (1.1)	<.001	3.4 (1)	3.3 (1.1)	.009
Chair stand test score	2.7 (1.1)	2.4 (1.1)	<.001	2.7 (1.1)	2.4 (1.1)	<.001
Balance score	3.9 (0.4)	3.9 (0.5)	<.001	3.9 (0.5)	3.9 (0.5)	.69

[a]RMB 1=US $0.15 as of June 10, 2026.

[b]HDL-C: high-density lipoprotein cholesterol.

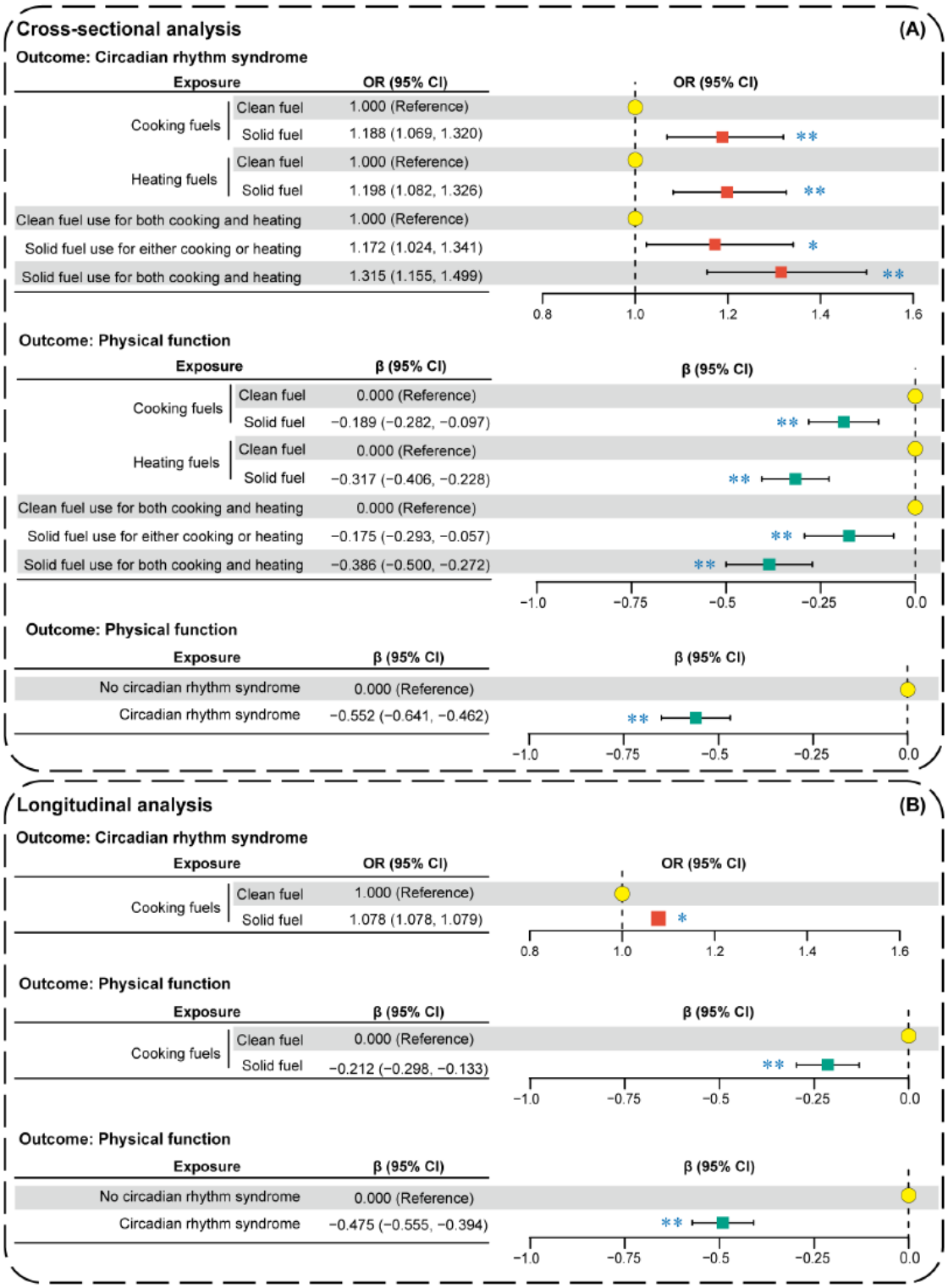
Cross-Sectional and Longitudinal Analysis of Separate Effects

Figure 1A displays the separate effects of solid fuel use, circadian rhythm syndrome, and physical function at baseline. Regarding solid fuel use and circadian rhythm syndrome, participants who used solid fuels for cooking and those who used solid fuels for heating exhibited an elevated risk of circadian rhythm syndrome by 18.8% (odds ratio [OR] 1.188, 95% CI 1.069 to 1.32) and 19.8% (OR 1.198, 95% CI 1.082 to 1.326), respectively. Compared with participants using clean fuels for cooking and heating, those partially or entirely reliant on solid fuels experienced an increased risk of circadian rhythm syndrome by 17.2% (OR 1.172, 95% CI 1.024 to 1.341) and 31.5% (OR 1.315, 95% CI 1.155 to 1.499), respectively. Additionally, we examined the influence of solid fuel use on physical function and observed a negative correlation between the use of solid fuels for cooking (β=–0.189, 95% CI –0.282 to –0.097) and heating (β=–0.317, 95% CI –0.406 to –0.228) and physical function compared with clean fuels. Partial and complete reliance on solid fuels led to reductions in physical function scores by 0.175 (95% CI –0.293 to –0.057) and 0.386 (95% CI –0.5 to –0.272) points, respectively. Further analysis revealed significant decreases in muscle strength (β=–0.061, 95% CI –0.117 to –0.005) and chair stand test scores (β=–0.273, 95% CI –0.335 to –0.211) among participants who used solid fuels for cooking and heating (Multimedia Appendix 8). After adjusting for all covariates, we found a negative correlation between circadian rhythm syndrome and physical function (β=–0.552, 95% CI –0.641 to –0.462). Patients with circadian rhythm syndrome experienced decreased muscle strength, balance, and chair stand test scores (β=–0.378, 95% CI –0.421 to –0.335; β=–0.032, 95% CI –0.053 to –0.01; and β=–0.123, 95% CI –0.172 to –0.073, respectively; Multimedia Appendix 8).

Consistent with the results of the cross-sectional analysis, as shown in Figure 1B, our longitudinal analyses revealed a positive correlation between solid fuel use and the risk of circadian rhythm syndrome (OR 1.078, 95% CI 1.031 to 1.125). Additionally, participants who used solid fuel (β=–0.212, 95% CI –0.298 to –0.133) and had circadian rhythm syndrome (β=–0.475, 95% CI –0.555 to –0.394) had notably lower physical function scores. Multimedia Appendix 9 further demonstrates that these participants also had notably lower chair stand test and balance scores. Although no correlation was detected in the analyses examining the impact of solid fuel use on muscle strength, lower muscle strength scores were observed in participants with circadian rhythm syndrome (β=–0.304, 95% CI –0.342 to –0.265).

XSL•FO
RenderX

Figure 1. Associations among solid fuel use, circadian rhythm syndrome, and physical function in (A) cross-sectional and (B) longitudinal analyses. Estimates were derived from multivariable regression models (binary logistic regression for circadian rhythm syndrome, linear models for cross-sectional analyses, and mixed-effects linear models for longitudinal analyses). Results shown are from model 4, adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type, indoor temperature, and chronic diseases. A single asterisk (*) indicates *P*<.05 and double asterisks (**) indicate *P*<.01. OR: odds ratio.

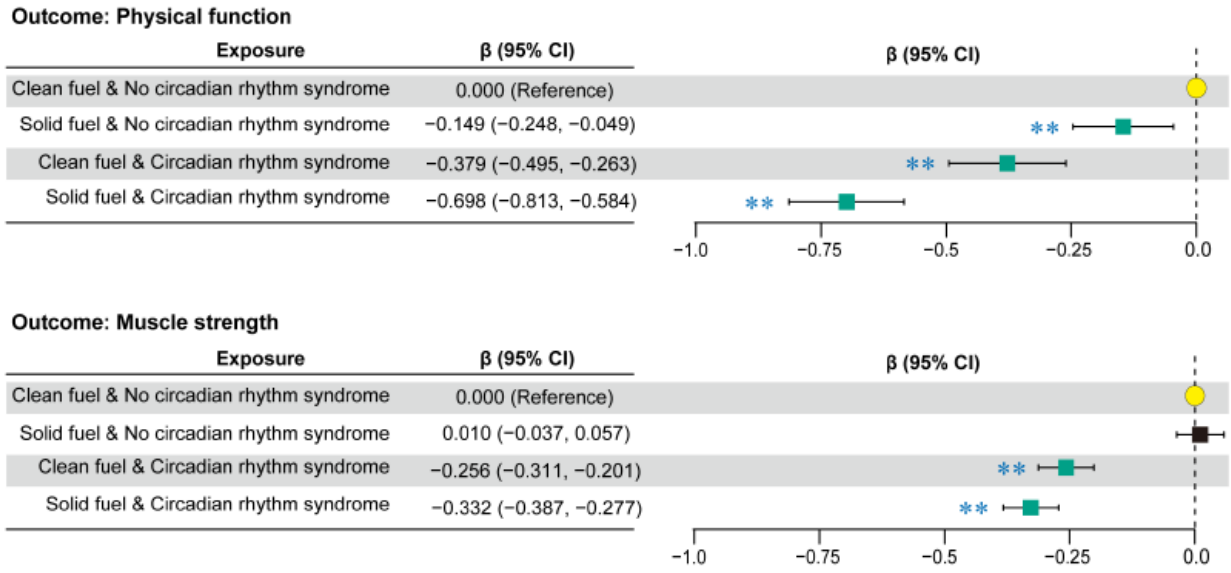


Combined Effects and Interaction Analysis

Figure 2 illustrates the combined effects of solid fuel use and circadian rhythm syndrome on physical function and muscle strength. Specifically, participants who used solid fuels and had circadian rhythm syndrome concurrently (solid fuels + circadian rhythm syndrome group) had the most pronounced reduction in physical function (β=–0.698, 95% CI –0.813 to –0.584) and muscle strength (β=–0.332, 95% CI –0.387 to –0.277) compared with those who did not use solid fuels or had only one of the risk factors.

XSL•FO
RenderX

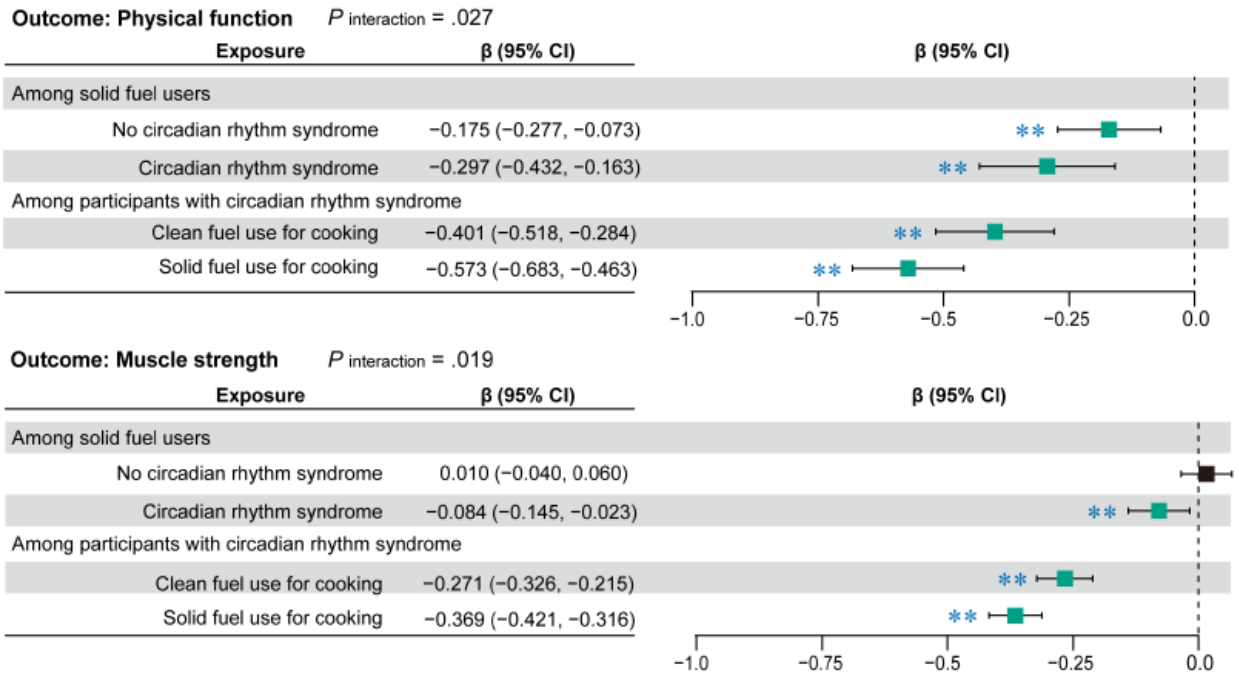
Figure 2. Combined effects of solid fuel use and circadian rhythm syndrome on physical function and muscle strength. Estimates were derived from multivariable linear regression models including a 2 × 2 combined exposure variable (fuel type × circadian rhythm syndrome) and a multiplicative interaction term to assess interaction effects. Adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type, indoor temperature, and chronic diseases. A single asterisk (*) indicates P<.05 and double asterisks (**) indicate P<.01.



Stratified Analyses

As we detected a multiplicative interaction between solid fuel use and circadian rhythm syndrome on physical function ($P_{interaction}$=.03) and muscle strength ($P_{interaction}$=.02), we conducted further stratified analyses (Figure 3). Our findings indicated that individuals with circadian rhythm syndrome were more vulnerable to the effects of solid fuel use on physical function (β=–0.297, 95% CI –0.432 to –0.163) and muscle strength (β=–0.084, 95% CI –0.145 to –0.023). Similarly, the negative correlation between circadian rhythm syndrome and physical function (β=–0.573, 95% CI –0.683 to –0.463) and muscle strength (β=–0.369, 95% CI –0.421 to –0.316) was more pronounced in participants who used solid fuels.

Figure 3. Stratified analyses were conducted separately for circadian rhythm syndrome and household fuel types. Estimates were derived from multivariable linear regression models stratified by fuel type and circadian rhythm syndrome. Adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type, indoor temperature, and chronic diseases. A single asterisk (*) indicates P<.05 and double asterisks (**) indicate P<.01.

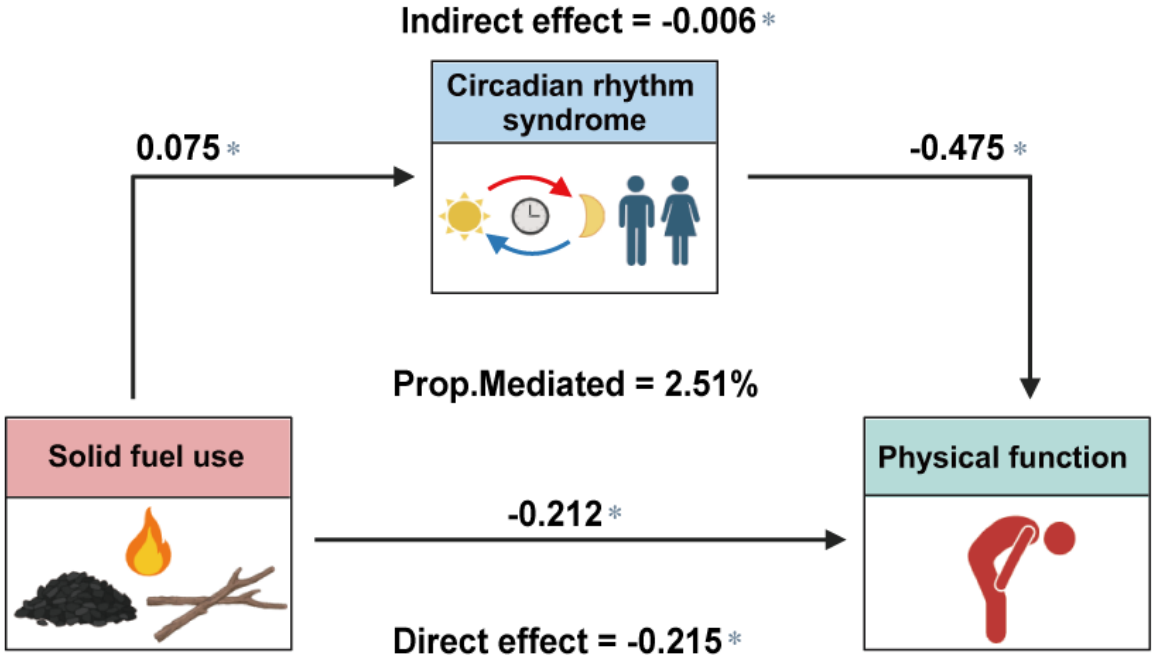


XSL•FO
RenderX

Mediation Analysis

The results of the mediation analysis are depicted in Figure 4. Circadian rhythm syndrome served as a significant mediator of the positive association between solid fuel use and physical function, with a mediation effect of 2.51%.

Figure 4. Mediation analysis of the association among solid fuel use, circadian rhythm syndrome, and physical function. Mediation analysis was conducted using the R package mediation to examine whether circadian rhythm syndrome mediated the association between solid fuel use and physical function scores. Regression models were fitted for both the mediator and the outcome, adjusting for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type, indoor temperature, and chronic diseases. Direct, indirect, and total effects were estimated using 1000 bootstrap simulations to derive 95% CIs. A single asterisk (*) indicates $P<.05$.



Sensitivity Analyses

Multiple sensitivity analyses were performed to verify the reliability and validity of the findings. When examining the association between solid fuel use and the number of circadian rhythm syndrome components, we found that participants who used solid fuels for both cooking and heating exhibited a 5.5% increase in the number of syndrome components compared with those using clean fuels (Multimedia Appendix 10). We also examined the effects of different types of solid fuels on physical function and circadian rhythm syndrome and found that coal use was more strongly correlated with both physical function (β=–0.276, 95% CI –0.411 to –0.141) and circadian rhythm syndrome (OR 1.221, 95% CI 1.02 to 1.463), whereas crop residue/wood burning was only significantly associated with physical function (β=–0.198, 95% CI –0.286 to –0.11) (Multimedia Appendices 11 and 12). The results of the subgroup analyses, shown in Multimedia Appendix 13, indicate that older adults are more susceptible to the impact of circadian rhythm syndrome on physical function scores, highlighting the importance of maintaining regular work and rest habits in this population. We excluded participants who reported “other” as their fuel type and reexamined the main associations among solid fuel use, circadian rhythm syndrome, and physical function in both cross-sectional and longitudinal analyses. The results were consistent with the primary findings (Multimedia Appendices 14-16), supporting the robustness of the observed associations.

Discussion

Principal Findings and Comparison With Prior Works

This study found that solid fuel use was positively associated with circadian rhythm syndrome, and that both solid fuel use and circadian rhythm syndrome were independently and jointly associated with poorer physical function and lower muscle strength. A significant multiplicative interaction was observed, indicating that the associations between solid fuel use and these functional outcomes were stronger among participants with circadian rhythm syndrome. In addition, circadian rhythm syndrome partially mediated the association between solid fuel use and physical function, suggesting that circadian rhythm disruption may represent one potential pathway underlying the observed associations.

To the best of our knowledge, this is the first study to investigate the relationship between household air pollution from solid fuel use and circadian rhythm syndrome, making it difficult to compare our results with those of previous studies. Although previous studies have not reported a direct link between solid fuel use and circadian rhythm syndrome, associations between solid fuel use and various components of the circadian rhythm syndrome have been demonstrated. Results from a cross-sectional study involving women suggested that household

XSL•FO
RenderX

air pollution from burning biomass fuel is associated with an increased risk of metabolic syndrome [29]. Studies from different regions have consistently confirmed the association between household air pollution and components of metabolic syndrome, such as blood pressure, blood glucose, and lipid profile, highlighting the potential risk of metabolic disorders that can be reduced by mitigating household air pollution [30-32]. The use of solid fuels also affects the mental health of older adults, with the results of a cohort study showing that using solid fuels for heating and cooking increases the risk of depression by 45% [33]. Similarly, the sleep-related health of middle-aged and older adults is affected by the use of solid fuels, especially in rural areas, where the vast majority of households rely on the combustion of biomass fuels for nighttime heating in winter. Studies from different regions have shown a significant positive association among solid fuel use, poor sleep quality, and reduced sleep duration [34,35]. Metabolic homeostasis, social stress, and sleep significantly affect the stability of the circadian clock, and exploring the effects of solid fuel use on one of them alone is one-sided. Therefore, this study examined the comprehensive impact of household air pollution on circadian rhythm disorders and revealed the importance of household fuel cleanliness on circadian rhythm homeostasis in middle-aged and older adults (OR 1.315, 95% CI 1.155 to 1.499 at baseline; OR 1.078, 95% CI 1.031 to 1.125 at follow-up). We also further analyzed the impact of household air pollution on circadian rhythm syndrome components; participants who used solid fuels for heating and cooking showed greater accumulation of circadian rhythm syndrome components than participants who used clean fuels. Cooking and heating are 2 common and often coexisting household activities; therefore, examining the independent and combined use of solid fuels in these settings is of particular relevance. We found that, compared with the use of clean fuels for both purposes, using solid fuels for either cooking or heating, as well as for both, was associated with a higher likelihood of circadian rhythm syndrome (either cooking or heating: OR 1.172, 95% CI 1.024 to 1.341; both: OR 1.315, 95% CI: 1.155 to 1.499) and lower physical function scores (either cooking or heating: β=–0.175, 95% CI –0.293 to –0.057; both: β=–0.386, 95% CI –0.5 to –0.272) in a graded manner. These findings suggest that different patterns of household fuel exposure may exert additive adverse effects on circadian regulation and physical function.

Disturbances in the circadian rhythm pose a major threat to the health of the musculoskeletal system in middle-aged and older adults. Real-life problems faced by older adults, such as a reduced metabolic rate, sleep disturbances, and depression caused by loneliness, affect the stability of circadian rhythms and impair muscle, joint, and bone health [9,36]. Pathophysiological changes in the musculoskeletal system significantly affect mobility and muscle strength and accelerate aging, which in turn leads to the disruption of circadian rhythms [3,37]. Although the vicious cycle of circadian rhythm syndrome and reduced physical function and muscle strength is detrimental to the health of middle-aged and older adults, no high-quality epidemiological studies have confirmed this association. Using a large nationwide cohort, we observed a significant inverse association among circadian rhythm syndrome, physical function (β=–0.386, 95% CI –0.5 to –0.272 at baseline; β=–0.475, 95% CI –0.555 to –0.394 at follow-up), and muscle strength (β=–0.061, 95% CI –0.117 to –0.005 at baseline; β=–0.304, 95% CI –0.342 to –0.265 at follow-up). We also evaluated the impact of circadian rhythm syndrome on physical performance (assessed by gait speed and chair stand tests) and showed that disruption of the circadian rhythm impairs mobility in middle-aged and older adults.

The decline in muscle strength and mobility due to aging causes most middle-aged and older adults to stay indoors for long periods. Therefore, it is important to explore the comprehensive impacts of indoor environmental factors and lifestyle. We found that solid fuel use and circadian rhythm syndrome had a synergistic effect on the impairment of physical function (β=–0.698, 95% CI –0.813 to –0.584) and muscle strength (β=–0.332, 95% CI –0.387 to –0.277). There was a multiplicative interaction between solid fuel use and circadian rhythm syndrome ($P_{interaction}$<.05), with circadian rhythm syndrome exacerbating the effects of solid fuel use on physical function (β=–0.297, 95% CI –0.432 to –0.163) and muscle strength (β=–0.084, 95% CI –0.145 to –0.023). Since middle-aged and older adults, especially in rural areas, mostly use solid fuels for nighttime heating, sleep disorders caused by circadian rhythm disorders are more likely to be exposed to household air pollution caused by solid fuels, which in turn leads to a more significant negative correlation with physical function and muscle strength. Furthermore, we elucidated the mediating role of circadian rhythm syndrome, whereby solid fuel use impaired physical function in part by increasing the risk of circadian rhythm syndrome (mediation effect: 2.51%).

Mechanistically, disruption of central and peripheral circadian regulatory networks may underlie the associations observed among solid fuel use and impaired physical function and muscle strength. The combustion of solid fuels generates particulate matter and toxic gases, which have been shown to enter systemic circulation and may affect the central nervous system, potentially disrupting the coordination between the central pacemaker in the suprachiasmatic nucleus and peripheral clocks through neurohumoral pathways [38-40]. In addition, household air pollution has been associated with increased systemic inflammation and oxidative stress. Air pollutants can stimulate the production of inflammatory cytokines and reactive oxygen species while impairing antioxidant defense systems [41,42]. These processes may contribute to circadian dysregulation, as both inflammation and oxidative stress have been implicated in altering clock gene expression and circadian homeostasis. Furthermore, chronic inflammation and oxidative stress are known to adversely affect joint integrity and muscle function, thereby contributing to declines in physical function [12]. Metabolic pathways may also play a role in linking circadian disruption to functional impairment. Experimental studies have shown that disturbances in skeletal muscle circadian rhythms can lead to metabolic dysregulation, including increased reliance on anaerobic glycolysis, reduced mitochondrial function, and increased apoptosis [9,43]. Such changes may accelerate muscle degradation and contribute to sarcopenia. Taken together, although the precise mechanisms remain to be fully elucidated, existing evidence suggests that household air pollution from solid fuel use may influence circadian regulation through

XSL•FO
RenderX

interconnected pathways involving neurohumoral disruption, systemic inflammation, oxidative stress, and metabolic imbalance, ultimately contributing to impaired physical function and reduced muscle strength. Consequently, although the precise mechanisms remain unclear, the relationship between household air pollution due to solid fuel use, circadian rhythm syndrome, physical function, and muscle strength warrants further investigation.

This study has several strengths: (1) novel findings, we identified an association between solid fuel use and circadian rhythm syndrome, with both solid fuel use and circadian rhythm syndrome separately and synergistically associated with declines in physical function and muscular strength; (2) representative and reliable results, we used a large national cohort of 126 cities across 28 provinces to conduct meticulous sensitivity analyses; (3) rigorous analytic methods, to minimize the recall and social desirability biases inherent in self-report measures, we used repeated-measures tests focusing on physical function, muscular strength, and physical performance in middle-aged and older adults; (4) guiding insights, mediation analyses revealed that solid fuel use affected physical function and muscular strength by increasing the risk of circadian rhythm syndrome, emphasizing the importance of household fuel cleaning and regular work and rest habits for the health of middle-aged and older adults.

Limitations

However, some limitations of this study should be noted. The CHARLS is a representative cohort study of middle-aged and older individuals. We only evaluated the impact of circadian rhythm syndrome on the health of middle-aged and older adults. The differences between the different age groups warrant further exploration. In addition, information on household solid fuel use was self-reported by participants, which may be subject to recall bias or reporting bias. This study relied solely on information on household fuel use for heating and cooking to assess indoor air pollution and may have overlooked other factors, such as ventilation efficiency and weather conditions. The degree of dependence on solid fuels varies seasonally. For example, in winter, the dependence on solid fuels for heating is higher. It is necessary to pay attention to the differences in the use of solid fuels in different seasons with respect to circadian rhythm syndrome, physical function, and muscle strength. Moreover, given the observational nature of this study, the identified associations among solid fuel use, circadian rhythm syndrome, and physical function or muscle strength should be interpreted with caution. Although we adjusted for a wide range of potential confounders, residual confounding cannot be completely ruled out, and causal relationships cannot be definitively established.

In conclusion, this nationwide prospective cohort study revealed a positive association between solid fuel use and circadian rhythm syndrome and found negative effects of solid fuel use and circadian rhythm syndrome, separately and jointly, on physical function and muscle strength. Disturbances in the circadian rhythm exacerbate the decline in physical function and muscle strength caused by solid fuel use. Circadian rhythm syndrome is an important mediator in the relationship between solid fuel use and physical function. This study highlights the combined impact of indoor environmental factors and lifestyle on middle-aged and older adults and recommends the use of clean fuels and the development of regular work and rest habits to improve physical health.

Acknowledgments

The data for this study come from the CHARLS dataset, and we sincerely thank the CHARLS research team and each of the participants.

Funding

This study was supported by the Joint Fund Project General Support Program Project of Liaoning Provincial Department of Science and Technology (grant number 2023-MSLH-386 to YW) and the 345 Talent Project of Shengjing Hospital of China Medical University (grant number M0271 to YW).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

HW contributed to conceptualization and writing–original draft. PM contributed to formal analysis and writing–original draft. DL, ZZ, and JG contributed to writing–review and editing. HS, WC, and YY contributed to visualization. JD contributed to conceptualization, supervision, formal analysis, and writing–review and editing. YW contributed to conceptualization, formal analysis, data curation, writing–review and editing, supervision, project administration, and funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Selection process of participants.
[PNG File , 822 KB - aging_v9i1e78352_app1.png]

Multimedia Appendix 2
Spatial distribution of participants.
[PNG File , 330 KB - aging_v9i1e78352_app2.png]

Multimedia Appendix 3
Assessment criteria and definition of circadian rhythm syndrome.
[PDF File (Adobe PDF File), 213 KB - aging_v9i1e78352_app3.pdf]

Multimedia Appendix 4
Supplementary methods for the study.
[PDF File (Adobe PDF File), 77 KB - aging_v9i1e78352_app4.pdf]

Multimedia Appendix 5
Comparison of basic information between the original dataset and the dataset used in the study. Continuous and categorical variables were compared using analyses of independent sample Student *t* test and chi-square tests, respectively.
[PDF File (Adobe PDF File), 100 KB - aging_v9i1e78352_app5.pdf]

Multimedia Appendix 6
Summary characteristics of 5018 participants in longitudinal analysis.
[PDF File (Adobe PDF File), 92 KB - aging_v9i1e78352_app6.pdf]

Multimedia Appendix 7
Spatial distribution of solid fuel use in households, scores of physical function and grip strength, and circadian rhythm syndrome participants Notes: Spatial distributions were estimated using ordinary kriging interpolation to construct a continuous prediction surface from observed provincial data. Spatial patterns were visualized using ordinary kriging interpolation based on observed CHARLS data. Tibet, Ningxia, Hainan, and Taiwan had no original participants in the analytic dataset; therefore, the values shown for these regions are interpolated estimates rather than observed data, whereas all other provinces reflect distributions based on observed data. The interpolated estimates were derived from spatial autocorrelation patterns of neighboring provinces. Color gradients represent predicted levels of the outcome variable.
[PNG File , 2785 KB - aging_v9i1e78352_app7.png]

Multimedia Appendix 8
Effects of solid fuel use on four test scores of physical functions after adjusting for all covariates. Notes: Adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type and indoor temperature, and chronic diseases.
[PDF File (Adobe PDF File), 64 KB - aging_v9i1e78352_app8.pdf]

Multimedia Appendix 9
Effects of solid fuel use and circadian rhythm syndrome on four test scores of physical functions after adjusting for all covariates. Notes: Adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type and indoor temperature, and chronic diseases.
[PDF File (Adobe PDF File), 58 KB - aging_v9i1e78352_app9.pdf]

Multimedia Appendix 10
Association between household solid fuel use and number of circadian rhythm syndrome components: negative binomial regression models. Notes: Adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type and indoor temperature, and chronic diseases. IRR: incidence rate ratio.
[PDF File (Adobe PDF File), 37 KB - aging_v9i1e78352_app10.pdf]

Multimedia Appendix 11
Effects of different solid fuel types on physical function and circadian rhythm syndrome. Notes: Adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type and indoor temperature, and chronic diseases.

XSL•FO
RenderX

[PDF File (Adobe PDF File), 57 KB - aging_v9i1e78352_app11.pdf]

Multimedia Appendix 12
Effects of solid fuel use on physical function in participants for whom data were available in 2011, 2013, and 2015. Notes: Adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type and indoor temperature, and chronic diseases.
[PDF File (Adobe PDF File), 51 KB - aging_v9i1e78352_app12.pdf]

Multimedia Appendix 13
Subgroup analysis of effects of solid fuel use and circadian rhythm syndrome on physical function. Notes: All models were adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type and indoor temperature, and chronic diseases, except for the stratifying variables.
[PNG File , 180 KB - aging_v9i1e78352_app13.png]

Multimedia Appendix 14
Sensitivity analysis excluding participants reporting “Other” fuel type: cross-sectional associations among solid fuel use, circadian rhythm syndrome, and physical function. Notes: Estimates were derived from multivariable regression models (binary logistic regression for circadian rhythm syndrome and linear models for cross-sectional analyses). Results shown are from model 4, adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type and indoor temperature, and chronic diseases.
[PDF File (Adobe PDF File), 39 KB - aging_v9i1e78352_app14.pdf]

Multimedia Appendix 15
Sensitivity analysis excluding participants reporting “Other” fuel type: Longitudinal associations among solid fuel use, circadian rhythm syndrome, and physical function. Notes: Estimates were derived from mixed-effects linear models for longitudinal analyses. Results shown are from model 4, adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type and indoor temperature, and chronic diseases.
[PDF File (Adobe PDF File), 37 KB - aging_v9i1e78352_app15.pdf]

Multimedia Appendix 16
Sensitivity analysis excluding participants reporting “Other” fuel type: combined associations among solid fuel use, circadian rhythm syndrome, and physical function. Notes: Estimates were derived from multivariable linear regression models including a 2 × 2 combined exposure variable (fuel type × circadian rhythm syndrome) and a multiplicative interaction term to assess interaction effects. Adjusted for age, gender, BMI, urbanization, education level, marital status, smoking status, drinking status, housing type and indoor temperature, and chronic diseases. Circs: circadian rhythm syndrome.
[PDF File (Adobe PDF File), 36 KB - aging_v9i1e78352_app16.pdf]

References

1. Wang H, Mou P, Yao Y, Su J, Guan J, Zhao Z, et al. Effects of different sizes of ambient particulate matter and household fuel use on physical function: national cohort study in China. Ecotoxicol Environ Saf 2024 May;276:116308 [FREE Full text] [doi: 10.1016/j.ecoenv.2024.116308] [Medline: 38593496]
2. Inouye SK, Peduzzi PN, Robison JT, Hughes JS, Horwitz RI, Concato J. Importance of functional measures in predicting mortality among older hospitalized patients. JAMA 1998;279(15):1187-1193. [doi: 10.1001/jama.279.15.1187] [Medline: 9555758]
3. Szulc P, Lewis JR, Chapurlat R. Decline in muscle strength and physical function after fracture in men - the prospective STRAMBO study. J Bone Miner Res 2024;39(3):252-259. [doi: 10.1093/jbmr/zjae014] [Medline: 38477746]
4. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
5. Koga HK, Grodstein F, Williams DR, Manson JE, Tindle HA, Shadyab AH. Longitudinal associations between optimism and objective measures of physical functioning in women. JAMA Psychiatry 2024;81(5):489-497. [doi: 10.1001/jamapsychiatry.2023.5068] [Medline: 38506825]
6. Landré B, Fayosse A, Ben Hassen C, Machado-Fragua MD, Dumurgier J, Kivimaki M. Terminal decline in objective and self-reported measures of motor function before death: 10 year follow-up of whitehall II cohort study. BMJ 2021;374:n1743 [FREE Full text] [doi: 10.1136/bmj.n1743] [Medline: 34348957]
7. Patke A, Young MW, Axelrod S. Molecular mechanisms and physiological importance of circadian rhythms. Nat Rev Mol Cell Biol 2020;21(2):67-84. [doi: 10.1038/s41580-019-0179-2] [Medline: 31768006]

XSL•FO
RenderX

8. Meyer N, Harvey AG, Lockley SW, Dijk D. Circadian rhythms and disorders of the timing of sleep. Lancet 2022;400(10357):1061-1078. [doi: 10.1016/S0140-6736(22)00877-7] [Medline: 36115370]
9. Martin RA, Viggars MR, Esser KA. Metabolism and exercise: the skeletal muscle clock takes centre stage. Nat Rev Endocrinol 2023;19(5):272-284. [doi: 10.1038/s41574-023-00805-8] [Medline: 36726017]
10. Bae S, Fang MZ, Rustgi V, Zarbl H, Androulakis IP. At the interface of lifestyle, behavior, and circadian rhythms: metabolic implications. Front Nutr 2019;6:132 [FREE Full text] [doi: 10.3389/fnut.2019.00132] [Medline: 31555652]
11. Shi Z, Tuomilehto J, Kronfeld-Schor N, Alberti GK, Stern N, El-Osta A, et al. The circadian syndrome predicts cardiovascular disease better than metabolic syndrome in Chinese adults. J Intern Med 2021 Jun;289(6):851-860 [FREE Full text] [doi: 10.1111/joim.13204] [Medline: 33340184]
12. Zhang H, Liang J, Chen N. Do not neglect the role of circadian rhythm in muscle atrophy. Ageing Res Rev 2020;63:101155. [doi: 10.1016/j.arr.2020.101155] [Medline: 32882420]
13. Yang L, Wu H, Zhao M, Magnussen CG, Xi B. Prevalence and trends in tobacco use, secondhand smoke exposure at home and household solid fuel use among women in 57 low- and middle-income countries, 2000-2018. Environ Int 2022;161:107142 [FREE Full text] [doi: 10.1016/j.envint.2022.107142] [Medline: 35180671]
14. Yu K, Qiu G, Chan KH, Lam KBH, Kurmi OP, Bennett DA, et al. Association of solid fuel use with risk of cardiovascular and all-cause mortality in rural China. JAMA 2018 Apr 03;319(13):1351-1361 [FREE Full text] [doi: 10.1001/jama.2018.2151] [Medline: 29614179]
15. Local Burden of Disease Household Air Pollution Collaborators. Mapping development and health effects of cooking with solid fuels in low-income and middle-income countries, 2000-18: A geospatial modelling study. Lancet Glob Health 2022;10(10):e1395-e1411 [FREE Full text] [doi: 10.1016/S2214-109X(22)00332-1] [Medline: 36113526]
16. Zhang B, Huang L, Zhu X, Ran L, Zhao H, Zhu Z. Impact of household solid fuel use on sarcopenia in China: a nationwide analysis. Sci Total Environ 2023;877:162814. [doi: 10.1016/j.scitotenv.2023.162814] [Medline: 36933714]
17. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
18. Hu X, Nie Z, Ou Y, Lin L, Qian Z, Vaughn MG. Long-term exposure to ambient air pollution, circadian syndrome and cardiovascular disease: a nationwide study in China. Sci Total Environ 2023;868:161696. [doi: 10.1016/j.scitotenv.2023.161696] [Medline: 36682545]
19. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National Sleep Foundation's sleep time duration recommendations: methodology and results summary. Sleep Health 2015 Mar;1(1):40-43. [doi: 10.1016/j.sleh.2014.12.010] [Medline: 29073412]
20. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994;10(2):77-84. [Medline: 8037935]
21. Huang Y, Bodnar D, Chen C, Sanchez-Andrade G, Sanderson M, Hurles ME. Rare genetic variants impact muscle strength. Nat Commun 2023;14(1):3449 [FREE Full text] [doi: 10.1038/s41467-023-39247-1] [Medline: 37301943]
22. Madden KM, Feldman B, Arishenkoff S, Meneilly G. A rapid point-of-care ultrasound marker for muscle mass and muscle strength in older adults. Age Ageing 2021;50(2):505-510 [FREE Full text] [doi: 10.1093/ageing/afaa163] [Medline: 32909032]
23. Alajlouni D, Tran T, Bliuc D, Blank RD, Cawthon PM, Orwoll ES. Muscle strength and physical performance improve fracture risk prediction beyond Garvan and FRAX: the Osteoporotic Fractures in Men (MrOS) study. J Bone Miner Res 2022;37(3):411-419 [FREE Full text] [doi: 10.1002/jbmr.4483] [Medline: 34842309]
24. Cooper R, Kuh D, Cooper CR, Gale C, Lawlor DA, Matthews F. Objective measures of physical capability and subsequent health: a systematic review. Age Ageing 2011;40(1):14-23 [FREE Full text] [doi: 10.1093/ageing/afq117] [Medline: 20843964]
25. Xie Y, Lou Y, Huang S, Jiang Q, Wang X, Wang L. Association between changes in physical functions and risk of stroke: a prospective cohort study. Age Ageing 2025;54(4):afaf087 [FREE Full text] [doi: 10.1093/ageing/afaf087] [Medline: 40202754]
26. Mayhew A, So H, Ma J, Beauchamp M, Griffith L, Kuspinar A, et al. Normative values for grip strength, gait speed, timed up and go, single leg balance, and chair rise derived from the Canadian longitudinal study on ageing. Age Ageing 2023 Apr 01;52(4):afad054 [FREE Full text] [doi: 10.1093/ageing/afad054] [Medline: 37078755]
27. Cooper R, Kuh D, Hardy R, Mortality Review Group, FALCon and HALCyon Study Teams. Objectively measured physical capability levels and mortality: systematic review and meta-analysis. BMJ 2010 Sep 09;341:c4467 [FREE Full text] [doi: 10.1136/bmj.c4467] [Medline: 20829298]
28. Yang J, Sakhvidi MJZ, de Hoogh K, Vienneau D, Siemiatyck J, Zins M, et al. Long-term exposure to black carbon and mortality: a 28-year follow-up of the GAZEL cohort. Environ Int 2021 Dec;157:106805 [FREE Full text] [doi: 10.1016/j.envint.2021.106805] [Medline: 34375941]
29. Rajkumar S, Young BN, Clark ML, Benka-Coker ML, Bachand AM, Brook RD, et al. Household air pollution from biomass-burning cookstoves and metabolic syndrome, blood lipid concentrations, and waist circumference in Honduran

XSL•FO
RenderX

women: a cross-sectional study. Environ Res 2019 Mar;170:46-55 [FREE Full text] [doi: 10.1016/j.envres.2018.12.010] [Medline: 30557691]
30. Baumgartner J, Schauer JJ, Ezzati M, Lu L, Cheng C, Patz JA, et al. Indoor air pollution and blood pressure in adult women living in rural China. Environ Health Perspect 2011;119(10):1390-1395 [FREE Full text] [doi: 10.1289/ehp.1003371] [Medline: 21724522]
31. Kang N, Song X, Zhang C, Li R, Yuchi Y, Liao W. Association of household air pollution with glucose homeostasis markers in Chinese rural women: effect modification of socioeconomic status. Ecotoxicol Environ Saf 2022;248:114283 [FREE Full text] [doi: 10.1016/j.ecoenv.2022.114283] [Medline: 36371884]
32. Xia X, Niu X, Chan K, Xu H, Shen Z, Cao J, et al. Effects of indoor air purification intervention on blood pressure, blood-oxygen saturation, and heart rate variability: a double-blinded cross-over randomized controlled trial of healthy young adults. Sci Total Environ 2023 May 20;874:162516. [doi: 10.1016/j.scitotenv.2023.162516] [Medline: 36868269]
33. Zhang H, Xia Y, Cao L, Chang Q, Zhao Y. Associations between long term exposures to outdoor air pollution and indoor solid fuel use and depression in China. J Environ Manage 2022;302(Pt A):113982. [doi: 10.1016/j.jenvman.2021.113982] [Medline: 34700082]
34. Chen C, Liu GG, Sun Y, Gu D, Zhang H, Yang H. Association between household fuel use and sleep quality in the oldest-old: evidence from a propensity-score matched case-control study in Hainan, China. Environ Res 2020;191:110229. [doi: 10.1016/j.envres.2020.110229] [Medline: 32950513]
35. Yu H, Luo J, Chen K, Pollitt KJG, Liew Z. Solid fuels use for cooking and sleep health in adults aged 45 years and older in China. Sci Rep 2021 Jun 25;11(1):13304. [doi: 10.1038/s41598-021-92452-0] [Medline: 34172775]
36. Swanson CM, Kohrt WM, Buxton OM, Everson CA, Wright KP, Orwoll ES, et al. The importance of the circadian system and sleep for bone health. Metabolism 2018 Jul;84:28-43 [FREE Full text] [doi: 10.1016/j.metabol.2017.12.002] [Medline: 29229227]
37. Morrison M, Halson SL, Weakley J, Hawley JA. Sleep, circadian biology and skeletal muscle interactions: implications for metabolic health. Sleep Med Rev 2022;66:101700 [FREE Full text] [doi: 10.1016/j.smrv.2022.101700] [Medline: 36272396]
38. Koronowski KB, Kinouchi K, Welz P, Smith JG, Zinna VM, Shi J. Defining the independence of the liver circadian clock. Cell 2019;177(6):1448-1462.e14 [FREE Full text] [doi: 10.1016/j.cell.2019.04.025] [Medline: 31150621]
39. You R, Ho YS, Chang RCC. The pathogenic effects of particulate matter on neurodegeneration: a review. J Biomed Sci 2022;29(1):15 [FREE Full text] [doi: 10.1186/s12929-022-00799-x] [Medline: 35189880]
40. Shi W, Zhang T, Li Y, Huang Y, Luo L. Association between household air pollution from solid fuel use and risk of chronic diseases and their multimorbidity among Chinese adults. Environ Int 2022;170:107635 [FREE Full text] [doi: 10.1016/j.envint.2022.107635] [Medline: 36413929]
41. Daiber A, Frenis K, Kuntic M, Li H, Wolf E, Kilgallen AB. Redox regulatory changes of circadian rhythm by the environmental risk factors traffic noise and air pollution. Antioxid Redox Signal 2022;37(10-12):679-703 [FREE Full text] [doi: 10.1089/ars.2021.0272] [Medline: 35088601]
42. Hanaoka BY, Ithurburn MP, Rigsbee CA, Bridges SL, Moellering DR, Gower B, et al. Chronic inflammation in rheumatoid arthritis and mediators of skeletal muscle pathology and physical impairment: a review. Arthritis Care Res (Hoboken) 2019 Feb;71(2):173-177 [FREE Full text] [doi: 10.1002/acr.23775] [Medline: 30295435]
43. Peek CB, Levine DC, Cedernaes J, Taguchi A, Kobayashi Y, Tsai SJ, et al. Circadian clock interaction with HIF1α mediates oxygenic metabolism and anaerobic glycolysis in skeletal muscle. Cell Metab 2017 Jan 10;25(1):86-92 [FREE Full text] [doi: 10.1016/j.cmet.2016.09.010] [Medline: 27773696]

Abbreviations

CHARLS: China Health and Retirement Longitudinal Study
OR: odds ratio

Edited by M Gray; submitted 01.Jun.2025; peer-reviewed by Q Zhao, K Hu; comments to author 09.Nov.2025; revised version received 25.Apr.2026; accepted 28.May.2026; published 29.Jun.2026.

Please cite as:
Wang H, Mou P, Liu D, Zhao Z, Chen W, Yao Y, Shi H, Guan J, Dong J, Wei Y
Associations of Solid Fuel Use and Circadian Rhythm Syndrome With Physical Function and Muscle Strength in Middle-Aged and Older Adults: Nationwide Cohort Study in China
JMIR Aging 2026;9:e78352
URL: https://aging.jmir.org/2026/1/e78352
doi:10.2196/78352
PMID:

XSL•FO
RenderX

©Haochen Wang, Pengsen Mou, Donghao Liu, Ze Zhao, Wenzheng Chen, Yuxin Yao, Haoyan Shi, Jiaxin Guan, Jing Dong, Yingliang Wei. Originally published in JMIR Aging (https://aging.jmir.org), 29.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations Between Indoor Musty Odors and Sleep Quality and Duration Among Older Adults: Longitudinal Study

Jinwei Bian[1,2], MPA; Pui Hing Chau[2], PhD; Kirsten Koehler[3], PhD; Youngmin Cho[1], PhD; Sofia Liu[1], MSN; Junxin Li[1], PhD

[1]School of Nursing, Johns Hopkins University, Baltimore, MD, United States

[2]School of Nursing, University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Department of Environmental and Health Engineering, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Junxin Li, PhD
School of Nursing
Johns Hopkins University
525 N. Wolfe Street
Baltimore, MD, 21205
United States
Phone: 1 410 502 2608
Email: junxin.li@jhu.edu

Abstract

Background: Indoor environmental air quality may influence sleep health among older adults. However, evidence on the association between indoor musty odors, a proxy for dampness- and mold-related exposures, and sleep outcomes remains limited. Susceptibility to environmental exposures may also vary by chronic lung disease status.

Objective: This study aimed to examine the associations of indoor musty odors with sleep quality and sleep duration among older adults and to explore these associations across subgroups defined by chronic lung disease status.

Methods: Data were drawn from 3 waves (2011-2018) of the Chinese Longitudinal Healthy Longevity Survey. Community-dwelling older adults aged ≥65 years without cognitive impairment or dementia were included. Self-reported sleep quality was categorized as good, fair, or poor, and sleep duration was categorized as short (<7 hours), normal (7-8 hours), or long (>8 hours). Generalized estimating equation models with a logit link and an exchangeable working correlation structure were applied. Stabilized inverse probability of attrition weights was used to account for attrition. Models were adjusted for sociodemographic, health-related, lifestyle, and residential environmental factors. Analyses were conducted overall and stratified by chronic lung disease status.

Results: A total of 5712 participants were included, contributing 10,449 observations across 3 survey waves. At baseline, participants had a mean age of 82.4 (SD 10.0) years; 39.8% (n=2271) were aged ≥85 years, 52.7% (n=3010) were male, and 12.8% (n=732) had chronic lung disease. Overall, 17.6% (n=1008) reported indoor musty odors; 24.6% (n=1404) and 11.4% (n=650) reported fair and poor sleep quality, respectively; and 32% (n=1829) and 27.9% (n=1591) reported short and long sleep durations, respectively. Exposure to indoor musty odors was associated with higher odds of fair sleep quality (odds ratio [OR] 1.19, 95% CI 1.04-1.36; *P*=.01) and poor sleep quality (OR 1.63, 95% CI 1.37-1.95; *P*<.001), as well as short sleep duration (OR 1.36, 95% CI 1.19-1.55; *P*<.001). These associations were statistically significant among individuals without chronic lung disease (*P*=.02, *P*<.001, and *P*<.001, respectively) but not among those with chronic lung disease (*P*=.20, *P*=.12, and *P*=.10, respectively).

Conclusions: Indoor musty odors were associated with adverse sleep outcomes among older adults, particularly among individuals without chronic lung disease. These findings highlight the importance of indoor air quality for sleep health and support interventions targeting dampness and mold in residential settings.

(JMIR Aging 2026;9:e107641) doi:10.2196/107641

KEYWORDS

indoor musty odors; sleep quality; sleep duration; indoor air quality; chronic lung disease

XSL•FO
RenderX

Introduction

Sleep plays a critical role in health and well-being among older adults [1]. However, sleep disturbances are highly prevalent in this population and have been linked with a range of adverse health outcomes, including increased risk of diabetes, falls, functional disability, and dementia [2]. A meta-analysis of 52 studies reported that the pooled prevalence of poor sleep quality among older adults reached 50% (95% CI 45%-55%) [3]. Sleep duration is another important dimension of sleep health. The National Sleep Foundation recommends 7 to 8 hours of sleep for older adults [4]. One umbrella review indicated that, among short sleepers, each 1-hour reduction in sleep duration was associated with a 3% to 11% increased risk of all-cause mortality, coronary heart disease, osteoporosis, stroke, and diabetes [5]. In contrast, among long sleepers, each 1-hour increase in sleep duration was associated with a 7% to 17% higher risk of stroke mortality, coronary heart disease, stroke, and diabetes [5].

Sleep is influenced by indoor environmental conditions, such as air quality, noise, temperature, and lighting. The US Environmental Protection Agency estimates that Americans spend about 90% of their time indoors [6], highlighting the importance of indoor environments for health, especially for older adults who tend to spend a substantial proportion of their time indoors. Indoor environmental conditions related to dampness and mold have received increasing attention in recent years. Humid conditions favor the growth of microbial species, such as fungi and bacteria, which release spores, cell fragments, and microbial volatile organic compounds into the air [7]. These emissions are associated with musty odors and may contribute to sensory irritation, discomfort, or physiological stress responses, thereby potentially disrupting sleep. Prior studies have shown that children exposed to visible mold or dampness at home have a higher risk of sleep problems, including difficulty falling asleep, difficulty maintaining sleep, shorter sleep duration, and sleep-disordered breathing [8,9]. In addition, a study in Canada reported that visible mold in the house was significantly associated with a higher risk of sleep deprivation in the general population [10]. Despite the potential impact on health, evidence regarding the association between dampness-related exposures and sleep outcomes remains limited among older adults. Furthermore, even when research is conducted, visible mold may not fully capture indoor microbial exposure, particularly when mold growth occurs in hidden or less visible areas.

Indoor musty odors, often indicative of dampness and mold growth, are commonly used as a proxy for underlying biological exposures in residential environments [7]. A prior study conducted in China reported that approximately 13.6% of respondents experienced indoor musty odors in their homes during the past year [11]. To date, only a few studies have examined the relationship between indoor musty odors and sleep health. One study reported that adults aged ≥20 years exposed to indoor musty odors were more likely to report sleep complaints [12]. Another cross-sectional study primarily examined the association between residential proximity to major roadways and sleep quality and found that the association was stronger among participants reporting indoor musty odors [13]. Despite these findings, evidence on the direct association between indoor musty odors and sleep remains limited, particularly among older adults.

In addition, individual susceptibility to environmental exposures may vary by chronic lung disease status, reflecting differences in respiratory sensitivity, underlying symptom burden, and behavioral responses [14]. These differences may influence both exposure levels and physiological responses and may also mask the observable impact of environmental exposures, thereby potentially attenuating the associations between indoor environmental exposures and sleep outcomes in certain subgroups. Therefore, examining these associations across subgroups defined by chronic lung disease status may provide additional insights into population-specific patterns.

Therefore, this study aimed to examine the associations between indoor musty odors and sleep outcomes, including sleep quality and sleep duration, among older adults and to further explore these associations across subgroups defined by chronic lung disease status.

Methods

Data and Sample

The data used in this study were derived from the Chinese Longitudinal Healthy Longevity Survey (CLHLS), a national survey of older adults aged ≥65 years [15]. A multistage cluster random sampling method was adopted to recruit participants from 23 of 34 provinces in China, providing representative data to investigate determinants of longevity. The baseline survey was conducted in 1998, and the follow-up surveys were conducted in 2000, 2002, 2005, 2008 to 2009, 2011 to 2012, 2014, and 2017 to 2018 [15].

Because information on indoor musty odors was first collected in the 2011 wave of the CLHLS, participants were included from this wave and followed up in the 2014 and 2017 to 2018 waves. Participants were included in the analytic cohort if they were aged ≥65 years; were community dwelling; and had nonmissing data on indoor musty odors, sleep quality, and sleep duration at baseline. Individuals with self-reported dementia or cognitive impairment at baseline, identified using age- and education-adjusted cutoff scores on the Mini-Mental State Examination, were excluded to minimize potential reporting bias [16]. Reports of sleeping less than 2 hours or more than 16 hours each night were considered implausible and were coded as missing [17]. During follow-up, person-wave observations were excluded if participants were living in institutional care; had self-reported dementia or cognitive impairment; or had missing data on indoor musty odors, sleep quality, or sleep duration at that wave. Details of the sample selection process are provided in Figure S1 in Multimedia Appendix 1.

Measurements

Indoor Musty Odors

Indoor musty odors were measured by asking the participants, “Does your home often have a smell of musty odors?” The responses to this question were categorized as either “no” or

XSL•FO
RenderX

"yes." Given the large-scale nature of this nationwide study, self-reported data were used as a practical and commonly accepted method for assessing environmental exposures. This approach has been validated in previous studies for evaluating indoor musty odors [11,12].

Sleep Quality and Sleep Duration

Sleep outcomes included sleep quality and sleep duration. Sleep quality was assessed using a single item, "How is your sleep quality now?" with responses recorded on a 5-point Likert scale. In this study, "very good" and "good" were classified as good, and "poor" and "very poor" were categorized as poor. The neutral response was considered fair [13]. Sleep duration was assessed using the question "How many hours do you usually sleep each day?" Information was provided by the participants or their close family members during face-to-face interviews. Sleep duration was initially treated as a continuous variable and categorized as short (<7 hours/day), normal (7-8 hours/day), or long (>8 hours/day), with 7 to 8 hours/day used as the reference range based on the National Sleep Foundation recommendation for older adults [4] and previous studies [18].

Covariates

On the basis of previous studies on sleep health [19,20], the following covariates were examined: sociodemographic, health-related, lifestyle, and residential environmental factors. Sociodemographic factors included age (65-84 or ≥85 years), gender (male or female), living area (urban or rural), marital status (married or others), literacy status (illiterate or literate), household income (at or above the median income or below the median income), and living alone (yes or no). Health-related factors included BMI, multimorbidity (yes or no), physical frailty (yes or no), and chronic lung disease (yes or no). BMI was calculated by dividing weight by the square of height (kg/m^2) and categorized according to the revised Asia-Pacific criteria proposed by the World Health Organization [21] into underweight (BMI<18.5 kg/m^2), normal weight (18.5kg/m^2≤BMI<23.0 kg/m^2), and overweight or obese (BMI≥23.0 kg/m^2). Multimorbidity was defined as the presence of 2 or more self-reported diseases, including hypertension, diabetes, heart disease, stroke, tuberculosis, cataract, glaucoma, cancer, gastric ulcer, Parkinson disease, arthritis, epilepsy, cholecystitis, blood disease, chronic nephritis, and hepatitis. Physical frailty was assessed using the frailty index (FI), with FI≥0.25 indicating physical frailty [22]. Chronic lung disease was assessed in the survey using a single self-reported question about the presence of any of the following conditions: bronchitis, emphysema, asthma, or pneumonia [23,24]. Lifestyle factors included current smoking (yes or no), current drinking (yes or no), and regular exercise (yes or no). Residential environmental factors included whether a respondent had their own bedroom (yes or no), their housing type (detached house or other), their cooking fuel type (solid fuel or other), and their exposure to secondhand smoke (yes or no). Cooking fuel type was categorized as solid if coal or coke, firewood or straw, or charcoal was used; otherwise, it was classified as other. Exposure to secondhand smoke was defined as having any co-residing household member who smoked more than 0 cigarettes per day inside the home.

Statistical Analysis

Descriptive statistics were used to summarize participants' baseline characteristics. Differences across sleep quality and sleep duration groups and by the presence of indoor musty odors were assessed using one-way ANOVA for continuous variables and χ^2 tests for categorical variables.

Given that the study population comprised older adults, a substantial proportion of participants experienced attrition during the study period due to death or loss to follow-up, which may be informative and introduce selection bias. To mitigate potential bias arising from informative attrition, stabilized inverse probability of attrition weights (IPAW) were constructed based on the probability of remaining under observation, conditional on indoor musty odors, sleep outcomes, and other covariates included in the analytic models [25].

Generalized estimating equations (GEEs) with a logit link and an exchangeable working correlation structure were used to account for within-subject correlation across repeated assessments [26,27]. As exposure and outcome were assessed concurrently at each visit, the estimated associations should be interpreted as repeated cross-sectional rather than strictly longitudinal effects. Although sleep quality and sleep duration were specified as 3-category outcomes, multinomial GEE models were not adopted because their combination with IPAW is not readily supported in standard R (R Foundation for Statistical Computing) implementations. Therefore, for sleep quality, 2 separate logistic IPAW-weighted GEE models were fitted to compare (1) fair vs good sleep quality and (2) poor vs good sleep quality. For sleep duration, another 2 logistic IPAW-weighted GEE models were fitted to compare (1) short vs normal sleep duration and (2) long vs normal sleep duration.

Both unadjusted and fully adjusted models were estimated. The fully adjusted models included the covariates described above and survey wave (modeled as dummy variables). Analyses were conducted both overall and stratified by chronic lung disease status, and interaction terms were included to assess effect modification by chronic lung disease status. Adjusted odds ratios (ORs) and 95% CIs were reported. Following previous CLHLS studies [28,29], sampling weights were not applied in the regression models because age, gender, and urban or rural residence, which were used to construct the CLHLS sampling weights, were included as covariates in the adjusted models. Applying sampling weights under these conditions may inflate SEs and reduce precision [30].

Given that alternative thresholds for short sleep have been used among older Chinese adults [31], we conducted a sensitivity analysis defining short sleep as <6 hours/day, with 6 to 8 hours/day as the reference category and >8 hours/day as long sleep.

Missing values among covariates were handled using multiple imputation by chained equations. Multicollinearity was assessed using the variance inflation factor (VIF), with VIF>4 indicating potential multicollinearity [32]. All statistical analyses were performed using RStudio (version 2025.09.0; Posit PBC), and a 2-sided P<.05 was considered statistically significant.

XSL•FO
RenderX

Ethical Considerations

The CLHLS was approved by the Biomedical Ethics Committee of Peking University (IRB00001052-13074). Written informed consent was obtained from all participants or their legal representatives. The present study was a secondary analysis of deidentified data, and participant privacy and confidentiality were maintained throughout the study.

Results

Sample Characteristics

A total of 5712 participants were included at baseline, contributing 10,449 observations across 3 survey waves. At baseline, the mean age of participants was 82.4 (SD 10.0) years; 39.8% (n=2271) were aged ≥85 years, 52.7% (n=3010) were male, 19% (n=1087) lived alone, 12.8% (n=732) had chronic lung disease, and 17.6% (n=1008) reported indoor musty odors. Regarding sleep outcomes, 64% (n=3658), 24.6% (n=1404), and 11.4% (n=650) of participants reported good, fair, and poor sleep quality, respectively. The mean sleep duration was 7.56 (SD 2.13) hours, with 40.1% (n=2292), 32% (n=1829), and 27.9% (n=1591) reporting normal, short, and long sleep duration, respectively. Table 1 presents the baseline characteristics of the study participants according to sleep quality and sleep duration.

Participants exposed to indoor musty odors were more likely to report poor sleep quality and short sleep duration than those not exposed (P<.001). Most covariates differed across sleep quality and sleep duration categories, except for living area, having one's own bedroom, housing type, and cooking fuel type. Baseline characteristics according to the presence of indoor musty odors are presented in Table S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Table 1. Distribution of sleep quality and sleep duration according to baseline characteristics (N=5712)[a].

Variables	Sleep quality				Sleep duration			
	Good	Fair	Poor	*P* value	Normal	Short	Long	*P* value
Participants, n (%)	3658 (64)	1404 (24.6)	650 (11.4)	—[b]	2292 (40.1)	1829 (32)	1591 (27.9)	—[b]
Indoor musty odors, n (%)				<.001				<.001
Yes (n=1008)	564 (56)	284 (28.2)	160 (15.8)		347 (34.4)	406 (40.3)	255 (25.3)	
No (n=4704)	3094 (65.8)	1120 (23.8)	490 (10.4)		1945 (41.3)	1423 (30.3)	1336 (28.4)	
Age (y), mean (SD)	82.3 (10.1)	82.8 (10.1)	82.3 (9.3)	.02	80.7 (9.7)	82.4 (9.7)	84.9 (10.2)	.03
65-84 (n=3441)	2209 (64.2)	831 (24.1)	401 (11.7)		1532 (44.5)	1124 (32.7)	785 (22.8)	
≥85 (n=2271)	1449 (63.8)	573 (25.2)	249 (11)		760 (33.5)	705 (31)	806 (35.5)	
Gender, n (%)				<.001				<.001
Male (n=3010)	2057 (68.3)	685 (22.8)	268 (8.9)		1291 (42.9)	845 (28.1)	874 (29)	
Female (n=2702)	1601 (59.3)	719 (26.6)	382 (14.1)		1001 (37.1)	984 (36.4)	717 (26.5)	
Living area, n (%)				.28				.20
Urban (n=2803)	1808 (64.5)	665 (23.7)	330 (11.8)		1157 (41.3)	886 (31.6)	760 (27.1)	
Rural (n=2909)	1850 (63.6)	739 (25.4)	320 (11)		1135 (39)	943 (32.4)	831 (28.6)	
Marital status, n (%)				.01				<.001
Married (n=2767)	1826 (66)	649 (23.5)	292 (10.5)		1248 (45.1)	834 (30.1)	685 (24.8)	
Others (n=2945)	1832 (62.2)	755 (25.6)	358 (12.2)		1044 (35.4)	995 (33.8)	906 (30.8)	
Literacy status, n (%)				<.001				<.001
Illiterate (n=2805)	1717 (61.2)	745 (26.6)	343 (12.2)		1035 (36.9)	953 (34)	817 (29.1)	
Literate (n=2907)	1941 (66.8)	659 (22.6)	307 (10.6)		1257 (43.3)	876 (30.1)	774 (26.6)	
Living alone, n (%)				.09				<.001
Yes (n=1087)	668 (61.5)	278 (25.5)	141 (13)		420 (38.6)	407 (37.5)	260 (23.9)	
No (n=4625)	2990 (64.6)	1126 (24.4)	509 (11)		1872 (40.5)	1422 (30.7)	1331 (28.8)	
Household income[c], n (%)				.001				<.001
Below the median income (n=2893)	1785 (61.7)	746 (25.8)	362 (12.5)		1115 (38.5)	1001 (34.6)	777 (26.9)	
At or above the median income (n=2819)	1873 (66.4)	658 (23.4)	288 (10.2)		1177 (41.8)	828 (29.3)	814 (28.9)	
BMI, n (%)				<.001				<.001
Normal (n=2657)	1709 (64.3)	653 (24.6)	295 (11.1)		1048 (39.4)	852 (32.1)	757 (28.5)	
Underweight (n=1198)	678 (56.6)	347 (29)	173 (14.4)		411 (34.3)	433 (36.2)	354 (29.5)	
Overweight or obese (n=1857)	1271 (68.4)	404 (21.8)	182 (9.8)		833 (44.9)	544 (29.3)	480 (25.8)	
Multimorbidity, n (%)				<.001				<.001
Yes (n=1378)	755 (54.8)	393 (28.5)	230 (16.7)		529 (38.4)	510 (37)	339 (24.6)	
No (n=4334)	2903 (67)	1011 (23.3)	420 (9.7)		1763 (40.7)	1319 (30.4)	1252 (28.9)	
Physical frailty, n (%)				<.001				.003
Yes (n=299)	135 (45.2)	88 (29.4)	76 (25.4)		99 (33.1)	122 (40.8)	78 (26.1)	
No (n=5413)	3523 (65.1)	1316 (24.3)	574 (10.6)		2193 (40.5)	1707 (31.5)	1513 (28)	
Chronic lung disease, n (%)				<.001				<.001
Yes (n=732)	412 (56.3)	195 (26.6)	125 (17.1)		249 (34)	285 (38.9)	198 (27.1)	
No (n=4980)	3246 (65.2)	1209 (24.3)	525 (10.5)		2043 (41)	1544 (31)	1393 (28)	

XSL•FO
RenderX

Variables	Sleep quality				Sleep duration			
	Good	Fair	Poor	*P* value	Normal	Short	Long	*P* value
Current smoking[d], n (%)				<.001				.005
Yes (n=1216)	845 (69.5)	258 (21.2)	113 (9.3)		524 (43.1)	344 (28.3)	348 (28.6)	
No (n=4496)	2813 (62.6)	1146 (25.5)	537 (11.9)		1768 (39.3)	1485 (33)	1243 (27.7)	
Current drinking[e], n (%)				<.001				.02
Yes (n=1106)	813 (73.5)	187 (16.9)	106 (9.6)		461 (41.7)	314 (28.4)	331 (29.9)	
No (n=4606)	2845 (61.8)	1217 (26.4)	544 (11.8)		1831 (39.8)	1515 (32.9)	1260 (27.3)	
Regular exercise[f], n (%)				<.001				.04
Yes (n=2315)	1561 (67.4)	503 (21.7)	251 (10.9)		971 (41.9)	703 (30.4)	641 (27.7)	
No (n=3397)	2097 (61.7)	901 (26.5)	399 (11.8)		1321 (38.9)	1126 (33.1)	950 (28)	
Own bedroom, n (%)				.12				.91
Yes (n=5410)	3480 (64.3)	1323 (24.5)	607 (11.2)		2174 (40.2)	1732 (32)	1504 (27.8)	
No (n=302)	178 (59)	81 (26.8)	43 (14.2)		118 (39.1)	97 (32.1)	87 (28.8)	
Housing type, n (%)				.33				.13
Detached house (n=3743)	2416 (64.5)	917 (24.5)	410 (11)		1484 (39.6)	1184 (31.6)	1075 (28.7)	
Others (n=1969)	1242 (63.1)	487 (24.7)	240 (12.2)		808 (41)	645 (32.8)	516 (26.2)	
Cooking fuel type, n (%)				.10				.88
Solid fuel (n=2767)	1742 (63)	715 (25.8)	310 (11.2)		1101 (39.8)	891 (32.2)	775 (28)	
Others (n=2945)	1916 (65.1)	689 (23.4)	340 (11.5)		1191 (40.4)	938 (31.9)	816 (27.7)	
Exposure to secondhand smoke, n (%)				.21				<.001
Yes (n=1634)	1034 (63.3)	395 (24.2)	205 (12.5)		615 (37.6)	477 (29.2)	542 (33.2)	
No (n=4078)	2624 (64.4)	1009 (24.7)	445 (10.9)		1677 (41.1)	1352 (33.2)	1049 (25.7)	

[a]For categorical variables, percentages were calculated using the total number of participants in each category as the denominator (row percentages).

[b]*P* value columns indicate that *P* values are reported for each variable overall and are not repeated for individual categories.

[c]For household income, participants with household incomes below the median for their respective living areas (east, central, and west) were classified as "below the median income," whereas those with incomes at or above the median were classified as "at or above the median income."

[d]Current smoking was assessed using the question "Do you smoke at the present?" and was classified as "yes" or "no."

[e]Current drinking was assessed using the question "Do you drink alcohol at the present?" and was classified as "yes" or "no."

[f]Regular exercise was assessed using the question "Do you do exercises regularly at present?" and was classified as "yes" or "no."

Associations Between Indoor Musty Odors and Sleep Outcomes

Table 2 and Figure 1 present the associations between indoor musty odors and sleep outcomes among older adults. In the fully adjusted model (10,449 person-wave observations), exposure to indoor musty odors was significantly associated with higher odds of fair sleep quality (OR 1.19, 95% CI 1.04-1.36; *P*=.01) and poor sleep quality (OR 1.63, 95% CI 1.37-1.95; *P*<.001), compared with good sleep quality. Exposure was also associated with a higher likelihood of short sleep duration (OR 1.36, 95% CI 1.19-1.55; *P*<.001), whereas no significant association was observed for long sleep duration (OR 1.04, 95% CI 0.89-1.21; *P*=.66), compared with optimal sleep duration. Full model results are presented in Table S2 in Multimedia Appendix 1.

XSL•FO
RenderX

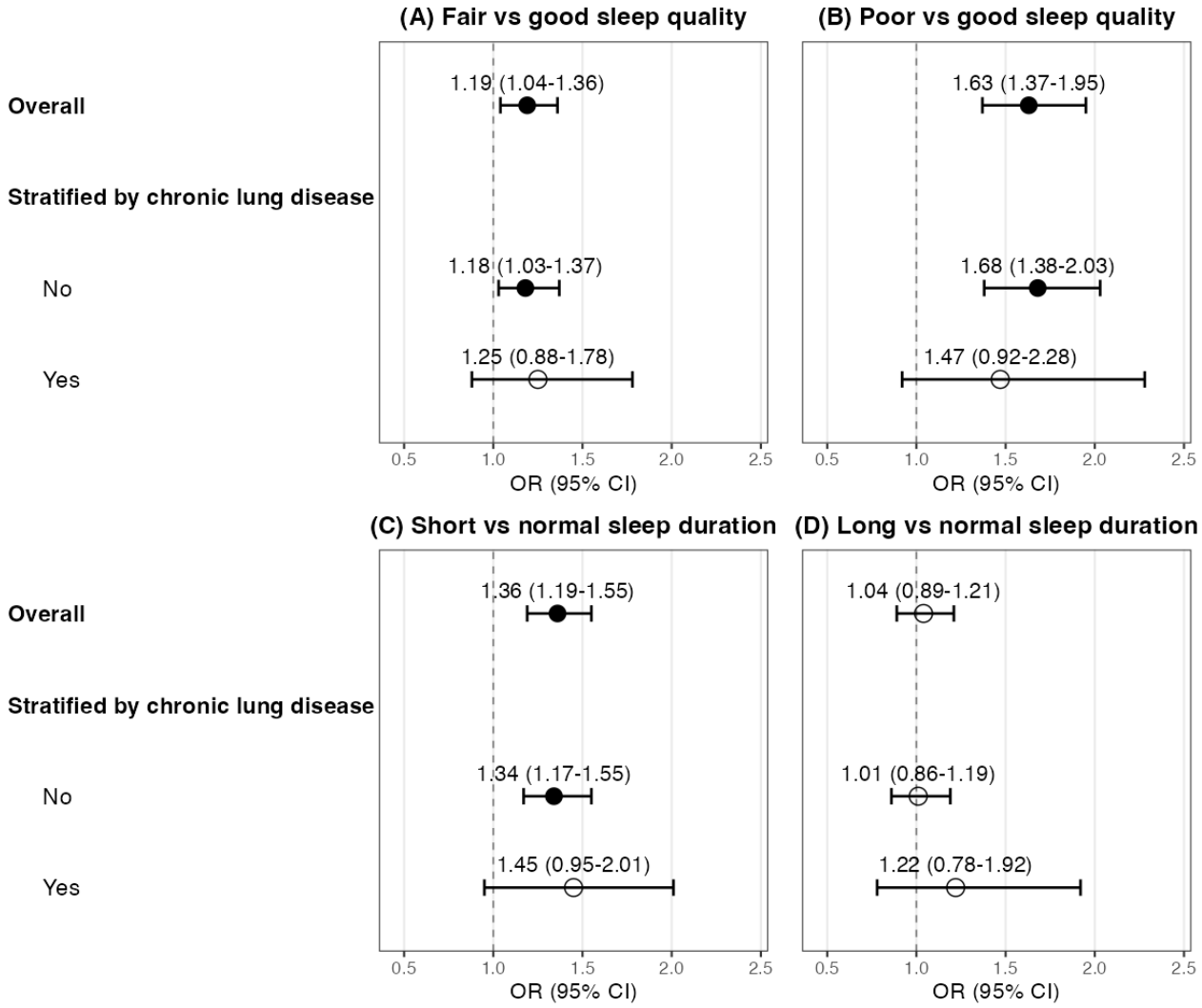
Table 2. Associations between indoor musty odors and sleep outcomes[a].

Model	Sleep quality				Sleep duration			
	Fair vs good		Poor vs good		Short vs normal		Long vs normal	
	OR[b] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Overall (10,449 person-wave observations)								
Unadjusted	1.25 (1.10-1.41)	.001	1.73 (1.47-2.03)	<.001	1.41 (1.24-1.59)	<.001	1.06 (0.92-1.22)	.43
Fully adjusted	1.19 (1.04-1.36)	.01	1.63 (1.37-1.95)	<.001	1.36 (1.19-1.55)	<.001	1.04 (0.89-1.21)	.66
Stratified by chronic lung disease								
Without chronic lung disease (9099 person-wave observations)								
Unadjusted	1.23 (1.07-1.40)	.003	1.73 (1.44-2.07)	<.001	1.37 (1.19-1.57)	<.001	1.04 (0.89-1.21)	.65
Fully adjusted	1.18 (1.03-1.37)	.02	1.68 (1.38-2.03)	<.001	1.34 (1.17-1.55)	<.001	1.01 (0.86-1.19)	.96
With chronic lung disease (1350 person-wave observations)								
Unadjusted	1.34 (0.90-1.87)	.08	1.61 (1.10-2.36)	.01	1.55 (1.02-2.08)	.04	1.20 (0.80-1.81)	.38
Fully adjusted	1.25 (0.88-1.78)	.20	1.47 (0.92-2.28)	.12	1.45 (0.95-2.01)	.10	1.22 (0.78-1.92)	.44

[a]Fully adjusted models were adjusted for sociodemographic factors, health-related factors, lifestyle factors, and residential environment factors.
[b]OR: odds ratio.

XSL•FO
RenderX

Figure 1. Associations between indoor musty odors and sleep outcomes among older adults, overall and stratified by chronic lung disease status. (A) Fair vs good sleep quality. (B) Poor vs good sleep quality. (C) Short vs normal sleep duration. (D) Long vs normal sleep duration. OR: odds ratio.



Stratified Analyses

In stratified analyses by chronic lung disease status, the associations remained statistically significant among participants without chronic lung disease (9099 person-wave observations). In the fully adjusted models, exposure to indoor musty odors was associated with higher odds of fair sleep quality (OR 1.18, 95% CI 1.03-1.37; *P*=.02), poor sleep quality (OR 1.68, 95% CI 1.38-2.03; *P*<.001), and short sleep duration (OR 1.34, 95% CI 1.17-1.55; *P*<.001), while no significant association was observed for long sleep duration (*P*=.96). In contrast, among participants with chronic lung disease (1350 person-wave observations), no statistically significant associations were observed for fair sleep quality (*P*=.20), poor sleep quality (*P*=.12), short sleep duration (*P*=.10), or long sleep duration (*P*=.44) in the fully adjusted models. No statistically significant interaction was observed by chronic lung disease status (*P* for interaction=.92, .56, .61, and .53 for fair sleep quality, poor sleep quality, short sleep duration, and long sleep duration, respectively).

Sensitivity Analysis

In the sensitivity analysis using <6 hours/day to define short sleep duration, the findings were generally consistent with those of the primary analysis. Indoor musty odors remained significantly associated with short sleep duration in the overall sample (OR 1.37, 95% CI 1.17-1.59; *P*<.001) and among participants without chronic lung disease (OR 1.36, 95% CI 1.15-1.61; *P*<.001) but not among those with chronic lung disease (*P*=.11). No significant associations with long sleep duration were observed (*P*=.52, .38, and .70, respectively) (Table S3 in Multimedia Appendix 1).

Discussion

Principal Results

On the basis of nationwide data from the CLHLS (2011-2018), this study found that exposure to indoor musty odors was associated with higher odds of poor sleep quality and short sleep duration among older adults. Similar patterns were observed across subgroups, although the associations appeared more evident among participants without chronic lung disease.

XSL•FO
RenderX

However, no statistically significant interaction by chronic lung disease status was identified.

In this study, 17.6% (1008/5712) of participants reported indoor musty odors at baseline. Similar prevalence has been reported in other studies, including 18.6% in Canada in 2010 [33] and 13.6% in China in 2018 [11]. In addition, studies using related indicators, such as the presence of mold, have reported prevalence estimates of 15.6% in Japan in 2001 [34] and 9.5% in the United States in 2017 to 2018 [35]. Although the reported prevalence varies across studies, potentially reflecting differences in study period and sample characteristics, the overall evidence suggests that mold-related conditions are relatively common in residential environments.

Our findings suggest that exposure to indoor musty odors was associated with higher odds of fair or poor sleep quality and short sleep duration among older adults, extending existing evidence from prior studies [12,13]. This may be explained by indoor musty odors, often indicative of dampness and mold growth, reflecting the presence of microbial volatile organic compounds and allergens that can induce respiratory irritation, inflammation, or discomfort, particularly affecting the upper airway [36]. Such irritation and inflammation may disrupt normal breathing during sleep and impair sleep continuity, thereby contributing to poorer sleep quality [37]. In addition, airway discomfort may lead to difficulty initiating sleep and early morning awakening, which may explain the higher likelihood of short sleep duration observed in this study. Unpleasant odors themselves may also act as sensory stressors, potentially activating the central nervous system and contributing to discomfort or difficulty initiating and maintaining sleep [38]. Furthermore, chronic exposure to poor indoor environmental conditions may be associated with psychological stress or reduced perceived environmental comfort, both of which are known to adversely affect sleep [39,40]. Taken together, these findings suggest that indoor musty odors may disrupt sleep through respiratory, sensory, and psychological pathways.

Interestingly, the associations between indoor musty odors and sleep outcomes were evident among individuals without chronic lung disease but not among those with chronic lung disease. Several explanations may account for this pattern. First, individuals with chronic lung disease are more likely to experience poorer sleep, as reflected in our sample by higher proportions of fair sleep quality (195/732, 26.6% vs 1209/4980, 24.3%), poor sleep quality (125/732, 17.1% vs 525/4980, 10.5%), and short sleep duration (285/732, 38.9% vs 1544/4980, 31%), compared with those without chronic lung disease at baseline. These underlying conditions, together with persistent respiratory symptoms, may attenuate or mask the additional impact of musty odors, thereby reducing the detectability of the marginal effect of indoor musty odors. Second, individuals with chronic lung conditions may adopt behavioral adaptations, such as improving indoor ventilation or avoiding damp environments to alleviate respiratory symptoms, which may reduce their actual exposure to mold-related pollutants and attenuate observable associations. Third, differential perception of odors and differential reporting of sleep quality and duration may also contribute to the observed differences across groups. Importantly, no statistically significant interaction was observed by chronic lung disease status, indicating no evidence of effect modification despite apparent differences in subgroup-specific estimates. These subgroup differences should therefore be interpreted with caution and may partly reflect limited statistical power, particularly in the subgroup with chronic lung disease. Future studies with larger sample sizes are needed to confirm these findings.

In contrast, no significant association was observed between indoor musty odors and long sleep duration. Previous research has linked dampness- and mold-related exposures primarily to sleep disturbances, including difficulties initiating or maintaining sleep [8], suggesting that these exposures may be more likely to disrupt sleep and contribute to shortened rather than prolonged sleep. Notably, the lack of association with long sleep duration persisted in the sensitivity analysis using an alternative sleep duration threshold, suggesting that this finding was unlikely to be explained solely by the choice of cutoff. However, evidence specifically examining indoor musty odors in relation to long sleep duration remains limited, and further studies are needed to clarify this association.

Implications

Our findings highlight that indoor environmental conditions may represent a practical and modifiable target for promoting sleep health among older adults. Air quality research has traditionally focused on health impacts among individuals with chronic respiratory diseases [41,42]. Extending this perspective, our findings further suggest that associations between indoor musty odors and adverse sleep outcomes may be more evident among individuals without chronic lung disease, although no statistically significant interaction was observed. These results underscore the need for broader public health attention to indoor musty odors, a dampness- and mold-related indicator of indoor air quality, beyond populations with preexisting respiratory conditions. Interventions such as improving household ventilation, reducing indoor dampness, optimizing building conditions, and minimizing exposure to mold-related pollutants may help mitigate the adverse effects of indoor musty odors and improve sleep outcomes [14].

Strengths and Limitations

To the best of our knowledge, this is the first study to examine the associations between indoor musty odors and both sleep quality and sleep duration among older adults, providing novel evidence in this area. In addition, the use of a nationally representative cohort with data collected across 3 waves over a 7-year period enhances the generalizability of the findings. We also conducted stratified analyses by chronic lung disease status, providing additional insights into subgroup-specific patterns. Nevertheless, this study has several limitations. First, sleep quality and sleep duration were assessed using self-reported measures, which may be subject to reporting bias or misclassification; however, we excluded extreme values of sleep duration and participants with cognitive impairment or dementia to minimize potential measurement error. Second, indoor musty odors were also self-reported, and the lack of objective measurements of specific indoor environmental pollutants may limit the precision of exposure assessment. In

XSL•FO
RenderX

addition, indoor musty odors may reflect broader housing conditions and socioeconomic disadvantage, which could not be fully disentangled from mold-related exposures. Future studies incorporating objective measurements are warranted to validate these findings and better understand the underlying mechanisms. Third, information on sensory-related conditions (eg, olfactory impairment) was not available, which may influence individuals' perception of indoor odors and potentially affect the observed associations. Additionally, residual confounding by psychosocial factors, such as depressive symptoms, anxiety, and perceived stress, cannot be excluded because these factors were not consistently measured across all study waves. Fourth, the relatively small sample size in certain subgroups, particularly in participants with chronic lung disease, may have limited statistical power and reduced the precision of the estimates. Finally, causal relationships cannot be inferred due to the observational nature of the study. However, compared with cross-sectional analyses, the use of GEE models incorporating repeated assessments across 3 waves allowed us to account for within-subject correlation and may provide more robust estimates of associations.

Conclusions

This study provides evidence that exposure to indoor musty odors was associated with higher odds of fair or poor sleep quality and short sleep duration among older adults. These associations appeared more evident among individuals without chronic lung disease, although no statistically significant interaction was observed. Our findings suggest that indoor environmental conditions may play a role in sleep health and could be considered as part of strategies to promote healthy sleep and support healthy aging.

Acknowledgments

The authors sincerely thank the staff of the provincial and county Centers for Disease Control and Prevention for conducting the fieldwork and the participants of the Chinese Longitudinal Healthy Longevity Survey (CLHLS) study.

Data Availability

The data that support the findings of this study are openly available through the Peking University Open Research Data Platform [43].

Funding

The study was funded by the Johns Hopkins University Catalyst Award and the Johns Hopkins School of Nursing Dorothy Evans Lyne (DEL) Fund. Sofia Liu was supported by a National Institute of Nursing Research training grant from the National Institutes of Health (T32NR020315). The funders had no role in the study design, data collection, analysis, interpretation of results, or manuscript preparation.

Generative AI Use

The authors declare that generative AI was not used in the preparation of this manuscript.

Authors' Contributions

Conceptualization: JB, JL
Data curation: JB, JL
Formal analysis: JB
Investigation: JB, JL
Methodology: JB, PHC, KK, YC, SL, JL
Project administration: JL
Supervision: JL
Validation: SL
Visualization: JB
Writing—original draft: JB
Writing—review & editing: PHC, KK, YC, SL, JL
All authors reviewed and approved the final version of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study sample selection (Figure S1), baseline characteristics according to the presence of indoor musty odors (Table S1), associations between indoor musty odors and sleep outcomes (Table S2), and sensitivity analysis of sleep duration (Table S3).
[DOCX File , 132 KB - aging_v9i1e107641_app1.docx]

XSL•FO
RenderX

References

1. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 01;37(1):9-17 [FREE Full text] [doi: 10.5665/sleep.3298] [Medline: 24470692]
2. Gao C, Guo J, Gong TT, Lv JL, Li XY, Liu FH, et al. Sleep duration/quality with health outcomes: an umbrella review of meta-analyses of prospective studies. Front Med (Lausanne) 2021 Jan 20;8:813943 [FREE Full text] [doi: 10.3389/fmed.2021.813943] [Medline: 35127769]
3. Kavousi P, Mali E, Seifhashemi N, Souri M, Pakravan L, Khalili F. Worldwide prevalence of poor sleep quality in older adults: a systematic review and meta-analysis. Iran J Psychiatry 2025 Apr;20(2):265-280. [doi: 10.18502/ijps.v20i2.18207] [Medline: 40521285]
4. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National Sleep Foundation's updated sleep duration recommendations: final report. Sleep Health 2015 Dec;1(4):233-243. [doi: 10.1016/j.sleh.2015.10.004] [Medline: 29073398]
5. Li J, Cao D, Huang Y, Chen Z, Wang R, Dong Q, et al. Sleep duration and health outcomes: an umbrella review. Sleep Breath 2022 Sep;26(3):1479-1501. [doi: 10.1007/s11325-021-02458-1] [Medline: 34435311]
6. Improving your indoor environment. United States Environmental Protection Agency. 2026. URL: https://www.epa.gov/indoor-air-quality-iaq/improving-your-indoor-environment [accessed 2026-11-11]
7. WHO guidelines for indoor air quality: dampness and mould. World Health Organization Regional Office for Europe. 2009 Jan 01. URL: https://iris.who.int/server/api/core/bitstreams/9261a374-7d3b-4825-a61e-0c09dea2af87/content [accessed 2026-11-11]
8. Tiesler CM, Thiering E, Tischer C, Lehmann I, Schaaf B, von Berg A, et al. Exposure to visible mould or dampness at home and sleep problems in children: results from the LISAplus study. Environ Res 2015 Feb;137:357-363 [FREE Full text] [doi: 10.1016/j.envres.2014.11.023] [Medline: 25601739]
9. Gokdemir Y, Civelek E, Cakir B, Demir A, Kocabas CN, Ikizoglu NB, et al. Prevalence of sleep-disordered breathing and associated risk factors in primary school children in urban and rural environments. Sleep Breath 2021 Jun;25(2):915-922. [doi: 10.1007/s11325-020-02206-x] [Medline: 33030645]
10. Karunanayake CP, Fenton M, Skomro R, Ramsden VR, Kirychuk S, Rennie DC, et al. Sleep deprivation in two Saskatchewan First Nation communities: a public health consideration. Sleep Med X 2021 Jun 02;3:100037 [FREE Full text] [doi: 10.1016/j.sleepx.2021.100037] [Medline: 34169273]
11. Liu X, Sun X, Wang X, Xu J, Zang S. Association between indoor musty odors and cognitive impairment among older adults. Sci Rep 2025 Aug 30;15(1):31943 [FREE Full text] [doi: 10.1038/s41598-025-12000-y] [Medline: 40883342]
12. Shiue I. Indoor mildew odour in old housing was associated with adult allergic symptoms, asthma, chronic bronchitis, vision, sleep and self-rated health: USA NHANES, 2005-2006. Environ Sci Pollut Res Int 2015 Sep;22(18):14234-14240. [doi: 10.1007/s11356-015-4671-8] [Medline: 25971810]
13. Zhang C, Zhao Y, Huang J, Hu M, Guo Y, Peng R, et al. Association between residential proximity to major roadways and sleep health in older Chinese people: the moderating role of living environments. BMC Public Health 2025 Oct 03;25(1):3337 [FREE Full text] [doi: 10.1186/s12889-025-24684-y] [Medline: 41044536]
14. Raju S, Siddharthan T, McCormack MC. Indoor air pollution and respiratory health. Clin Chest Med 2020 Dec;41(4):825-843 [FREE Full text] [doi: 10.1016/j.ccm.2020.08.014] [Medline: 33153698]
15. Center for Healthy Aging and Development Studies. The Chinese Longitudinal Healthy Longevity Survey (CLHLS)—longitudinal data (1998-2018). Peking University Open Research Data. 2020. URL: https://opendata.pku.edu.cn/dataset.xhtml?persistentId=doi:10.18170/DVN/WBO7LK [accessed 2026-11-11]
16. Zhang S, Qiu Q, Qian S, Lin X, Yan F, Sun L, et al. Determining appropriate screening tools and cutoffs for cognitive impairment in the Chinese elderly. Front Psychiatry 2021 Dec 02;12:773281 [FREE Full text] [doi: 10.3389/fpsyt.2021.773281] [Medline: 34925100]
17. Kripke DF, Garfinkel L, Wingard DL, Klauber MR, Marler MR. Mortality associated with sleep duration and insomnia. Arch Gen Psychiatry 2002 Feb 01;59(2):131-136 [FREE Full text] [doi: 10.1001/archpsyc.59.2.131]
18. Ren Y, Miao M, Yuan W, Sun J. Sleep duration and all-cause mortality in the elderly in China: a population-based cohort study. BMC Geriatr 2020 Dec 30;20(1):541 [FREE Full text] [doi: 10.1186/s12877-020-01962-5] [Medline: 33380318]
19. Miner B, Kryger MH. Sleep in the aging population. Sleep Med Clin 2016 Dec 20;12(1):31-38 [FREE Full text] [doi: 10.1016/j.jsmc.2016.10.008] [Medline: 28159095]
20. Grandner MA. Sleep, health, and society. Sleep Med Clin 2017 Mar;12(1):1-22 [FREE Full text] [doi: 10.1016/j.jsmc.2016.10.012] [Medline: 28159089]
21. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004 Jan 10;363(9403):157-163. [doi: 10.1016/S0140-6736(03)15268-3] [Medline: 14726171]
22. Zhang J, Wang Q, Hao W, Zhu D. Long-term food variety and dietary patterns are associated with frailty among Chinese older adults: a cohort study based on CLHLS from 2014 to 2018. Nutrients 2022 Oct 13;14(20):4279 [FREE Full text] [doi: 10.3390/nu14204279] [Medline: 36296962]

XSL•FO
RenderX

23. Zhao SQ, Zhao LP, Xu XP, You H. Individual-level health care costs attributable to noncommunicable diseases: a longitudinal study based on the older adults in China. Inquiry 2023;60:469580231214469 [FREE Full text] [doi: 10.1177/00469580231214469] [Medline: 38044620]
24. Li S, Tang X, Ge Y, Zhou D, Liu H, Sun T, et al. Association between frailty transitions and chronic lung disease incidence: findings from four national cohort studies. BMC Public Health 2025 Oct 03;25(1):3313 [FREE Full text] [doi: 10.1186/s12889-025-24361-0] [Medline: 41044483]
25. Weuve J, Tchetgen Tchetgen EJ, Glymour MM, Beck TL, Aggarwal NT, Wilson RS, et al. Accounting for bias due to selective attrition: the example of smoking and cognitive decline. Epidemiology 2012 Jan;23(1):119-128 [FREE Full text] [doi: 10.1097/EDE.0b013e318230e861] [Medline: 21989136]
26. Robins JM, Rotnitzky A. Semiparametric efficiency in multivariate regression models with missing data. J Am Stat Assoc 1995;90(429):122-129 [FREE Full text] [doi: 10.1080/01621459.1995.10476494]
27. Westgate PM, West BT. Tools for selecting working correlation structures when using weighted GEE to model longitudinal survey data. J Surv Stat Methodol 2021 Feb;9(1):141-158 [FREE Full text] [doi: 10.1093/jssam/smz048] [Medline: 37020583]
28. Xu H, Dupre ME, Gu D, Wu B. The impact of residential status on cognitive decline among older adults in China: results from a longitudinal study. BMC Geriatr 2017 May 15;17(1):107 [FREE Full text] [doi: 10.1186/s12877-017-0501-9] [Medline: 28506252]
29. Bai C, Xie Y, Gu D. Sleep and cognition at older ages and the moderating role of fruit and/or vegetable intake: the empirical evidence from China. Dietetics 2026 Jan 06;5(1):4. [doi: 10.3390/dietetics5010004]
30. Winship C, Radbill L. Sampling weights and regression analysis. Sociol Methods Res 1994;23(2):230-257 [FREE Full text] [doi: 10.1177/0049124194023002004]
31. Liu H, Byles JE, Xu X, Zhang M, Wu X, Hall JJ. Association between nighttime sleep and successful aging among older Chinese people. Sleep Med 2016 Jun;22:18-24. [doi: 10.1016/j.sleep.2016.04.016] [Medline: 27544831]
32. Mason G. Coping with collinearity. Can J Program Eval 1987 Mar 01;2(1):87-93 [FREE Full text] [doi: 10.3138/cjpe.02.010]
33. Janzen B, Karunanayake C, Rennie D, Pickett W, Lawson J, Kirychuk S, et al. Gender differences in the association of individual and contextual exposures with lung function in a rural Canadian population. Lung 2017 Feb;195(1):43-52. [doi: 10.1007/s00408-016-9950-8] [Medline: 27738826]
34. Saijo Y, Kishi R, Sata F, Katakura Y, Urashima Y, Hatakeyama A, et al. Symptoms in relation to chemicals and dampness in newly built dwellings. Int Arch Occup Environ Health 2004 Oct;77(7):461-470. [doi: 10.1007/s00420-004-0535-0] [Medline: 15558298]
35. Xiao S, Ngo AL, Mendola P, Bates MN, Barcellos AL, Ferrara A, et al. Household mold, pesticide use, and childhood asthma: a nationwide study in the U.S. Int J Hyg Environ Health 2021 Apr;233:113694 [FREE Full text] [doi: 10.1016/j.ijheh.2021.113694] [Medline: 33556714]
36. Coulburn L, Miller W. Prevalence, risk factors, and impacts related to mould-affected housing: an Australian integrative review. Int J Environ Res Public Health 2022 Feb 07;19(3):1854 [FREE Full text] [doi: 10.3390/ijerph19031854] [Medline: 35162876]
37. Cao B, Chen Y, McIntyre RS. Comprehensive review of the current literature on impact of ambient air pollution and sleep quality. Sleep Med 2021 Mar;79:211-219. [doi: 10.1016/j.sleep.2020.04.009] [Medline: 32912798]
38. Oberdörster G, Sharp Z, Atudorei V, Elder A, Gelein R, Kreyling W, et al. Translocation of inhaled ultrafine particles to the brain. Inhal Toxicol 2008;16(6-7):437-445 [FREE Full text] [doi: 10.1080/08958370490439597]
39. Han KS, Kim L, Shim I. Stress and sleep disorder. Exp Neurobiol 2012 Dec;21(4):141-150 [FREE Full text] [doi: 10.5607/en.2012.21.4.141] [Medline: 23319874]
40. Evans GW. The built environment and mental health. J Urban Health 2003 Dec;80(4):536-555 [FREE Full text] [doi: 10.1093/jurban/jtg063] [Medline: 14709704]
41. Raju S, Woo H, Koehler K, Fawzy A, Liu C, Putcha N, et al. Indoor air pollution and impaired cardiac autonomic function in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2023 Mar 15;207(6):721-730 [FREE Full text] [doi: 10.1164/rccm.202203-0523OC] [Medline: 36288428]
42. Zhang H, Ye R, Yang H, Liu Y, Zhao L, Zhao Y, et al. Long-term noise exposure and cause-specific mortality in chronic respiratory diseases, considering the modifying effect of air pollution. Ecotoxicol Environ Saf 2024 Sep 01;282:116740 [FREE Full text] [doi: 10.1016/j.ecoenv.2024.116740] [Medline: 39024955]
43. Peking University Open Research Data. URL: https://opendata.pku.edu.cn/ [accessed 2026-11-12]

Abbreviations

CLHLS: Chinese Longitudinal Healthy Longevity Survey
FI: frailty index
GEE: generalized estimating equation
IPAW: inverse probability of attrition weights
OR: odds ratio

XSL•FO
RenderX

Edited by A Mavragani; submitted 21.Jul.2026; peer-reviewed by OA Jaramillo-Morales, Z Chen; comments to author 22.Sep.2026; revised version received 09.Oct.2026; accepted 08.Nov.2026; published 23.Nov.2026.

Please cite as:
Bian J, Chau PH, Koehler K, Cho Y, Liu S, Li J
Associations Between Indoor Musty Odors and Sleep Quality and Duration Among Older Adults: Longitudinal Study
JMIR Aging 2026;9:e107641
URL: https://aging.jmir.org/2026/1/e107641
doi:10.2196/107641
PMID:

©Jinwei Bian, Pui Hing Chau, Kirsten Koehler, Youngmin Cho, Sofia Liu, Junxin Li. Originally published in JMIR Aging (https://aging.jmir.org), 23.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Latent Class Analysis for Identifying Patterns of Digital Device Use and Adaptation Difficulty and Their Associations With Service Needs Among Older Adults: Secondary Data Analysis of a Cross-Sectional Survey Study

SangA Lee[1], PhD; Minhee Yang[2], PhD; Jungwon Cho[3], MSN; Jiyoung Kim[3], BSN; Eunhee Cho[4], PhD

[1]Research Institute for Hospice and Palliative Care, College of Nursing, The Catholic University of Korea, Seoul, Republic of Korea

[2]Department of Nursing, Catholic Kwandong University, 24 Beomil-ro 579beon-gil, Gangneung, Republic of Korea

[3]College of Nursing and Brain Korea 21 FOUR Project, Yonsei University, Seoul, Republic of Korea

[4]Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
Minhee Yang, PhD
Department of Nursing, Catholic Kwandong University, 24 Beomil-ro 579beon-gil, Gangneung, Republic of Korea

Abstract

Background: In this digital era, digital devices play an increasingly vital role in health management beyond daily convenience. However, age-related physical changes and low digital literacy hinder technology use among older adults, widening the digital divide, with potential implications for health outcomes.

Objective: This study aimed to identify subgroups of older adults based on their digital device use and adaptation difficulty, examine the characteristics of each subgroup, and assess their preferred service needs for digital adaptation.

Methods: This is a cross-sectional secondary data analysis study. Data were drawn from 7470 older adults participating in the 2023 National Survey of Older Koreans. Digital device use and adaptation difficulty were assessed using 13 binary items and one 5-point Likert-scale item. Latent class analysis was conducted to identify subgroups of older adults based on these patterns. Service needs for digital adaptation were categorized as digital accessibility support, digital literacy enhancement, and age-friendly digital solutions. Additionally, bias-adjusted 3-step multinomial logistic regression approaches were used to examine subgroup characteristics and service preferences.

Results: The study sample included older adults aged 65 to 98 (mean 73.78, SD 6.65) years. Among the preferred service needs for digital adaptation, digital accessibility support was selected most frequently (n=2731, 36.6%), followed by digital literacy enhancement (n=2508, 33.6%) and age-friendly digital solutions (n=2231, 29.9%). Four distinct subgroups emerged from the analysis: digitally active (active; n=1112, 14.9%), moderately active (moderate; n=2492, 33.4%), text messages only (text-only; n=2082, 27.9%), and digitally inactive (inactive; n=1784, 23.9%). Compared to the active group, the inactive and text-only groups exhibited lower cognitive functioning (odds ratio [OR] 0.676, 95% CI 0.637-0.717 and OR 0.761, 95% CI 0.720-0.805, respectively). The inactive group required greater assistance in daily living (OR 1.165, 95% CI 1.024-1.326). Additionally, older age, rural residence, lower income, and lower educational attainment were significantly associated with patterns characterized by more limited digital device use. Regarding service needs, compared to the active group, groups reflecting greater digital limitations expressed a significantly stronger preference for age-friendly digital solutions over digital literacy enhancement, with ORs ranging from 1.294 (95% CI 1.051-1.593) to 1.978 (95% CI 1.510-2.591).

Conclusions: This study identified 4 distinct subgroups of digital device use and adaptation difficulty among older adults and examined their sociodemographic and health-related characteristics. Notably, individuals in classes reflecting greater digital limitations demonstrated a stronger preference for age-friendly digital services than for conventional digital literacy training, indicating that usability and design barriers might be more salient than skill-based limitations for many older adults. By delineating subgroup-specific needs, our findings can provide actionable guidance for developing targeted interventions. Tailored strategies are essential to support older adults in navigating an increasingly digital environment, thereby reducing digital divides and promoting digital inclusion with potential implications for healthy aging.

(JMIR Aging 2026;9:e91770) doi:10.2196/91770

KEYWORDS

digital adaptation; digital divide; digital inclusion; latent class analysis; older adults

XSL•FO
RenderX

Introduction

Technology-driven devices have been incorporated into many aspects of daily activities such as information exchange, communication with others, leisure, and, most importantly, health care [1-3]. With the widespread use of smartphones, wearable devices, and health-related apps, both health care professionals and patients have benefited from real-time health monitoring and personalized care, thereby supporting more effective preventive health behaviors [4]. The COVID-19 pandemic further accelerated the adoption of digital health technologies, as the demand for remote health care services surged [5]. Various digital functions used in everyday life, such as digital communication, information searching, video viewing, and application use, are also increasingly applicable in accessing and navigating digital health care services. Digital communication supports remote consultations, while information searching and video viewing provide access to health information and education. Apps are also used to enable engagement with various health care services [1,2]. As digital health care increasingly relies on functions that are similarly used in everyday digital activities, limited experience with these functions or difficulty adapting to digital technologies may pose additional challenges for older adults in accessing digital health care services.

Despite these technological advancements, older adults remain disproportionately affected by the digital divide, limiting their access to devices, internet usage, and overall engagement with technology compared to that observed with the younger generations [6-9]. Their lower digital literacy levels are often attributed to unfamiliarity and limited experience with digital devices [2,10,11]. In addition, age-related sensory impairments, such as diminished vision and hearing loss, present significant physical barriers to effective interactions with digital technologies [2,11-13]. Cognitive aging—manifesting as difficulties with memory, such as recalling passwords or acquiring new information—further compounds these challenges [11,12]. Previous studies have reported differences in digital technology use among older adults according to sociodemographic characteristics, including age, education, and income, as well as health-related factors such as cognitive and functional statuses [12,14,15]. As health care services increasingly rely on digital technologies, these differences in digital device use and adaptation difficulty may translate into unequal access to digital health resources and services, potentially widening existing health inequities [16-20]. In contrast, older adults often encounter barriers from the outset of the health care journey. For instance, since the onset of the COVID-19 pandemic, many hospitals have adopted app-based reservation systems, which have created substantial difficulties for older adults in the scheduling of medical appointments [10]. As digital health care solutions continue to evolve and become established as the standard of service delivery, older adults risk further marginalization in accessing essential health care services, thereby exacerbating existing health inequities.

This issue is particularly salient in Korea, one of the most digitally connected countries, with a smartphone penetration rate of 95% and a household internet penetration rate of 83% as of 2024 [21,22]. Within this highly digitalized context, many public services and welfare benefits have transitioned to exclusively digital formats, offering additional advantages to only online users [13]. However, a substantial proportion of older adults in Korea (50.0% - 67.2%) continue to experience difficulties with using digital devices, limiting their ability to benefit from these digital-based services [23,24]. Digital inclusion encompasses multiple dimensions, including access to digital technologies and individuals' ability to use them effectively [2,7]; therefore, it is particularly important in highly digitalized societies such as Korea. This gap between digital infrastructure and older adults' access to and use of digital services highlights the urgent need for more targeted support to ensure equitable access for older adults.

A growing body of research has examined the digital divide among older adults, with particular attention to digital device usage, literacy levels, and influencing factors [12,14,15,18,25]. Some qualitative studies have examined the challenges older adults face in adapting to digital technologies [10,13,26]. Notably, digital skill training and support have frequently been emphasized as strategies to promote digital inclusion among older adults by improving their ability to access and engage with digital technologies [14,18]. Nonetheless, improvements in digital knowledge or skills through training do not necessarily translate into increased digital device use or fewer difficulties in adapting to digital technologies [2,11]. For older adults, challenges in digital technology use may persist even after training, as improved knowledge does not always lead to confident or autonomous use [27]. Digital competency is a broader construct encompassing operational device skills, communication, safety, problem-solving, and attitudes toward technology [28]. Previous studies have identified distinct profiles of older adults primarily based on digital literacy, which focuses on knowledge and skills related to the use of digital technologies [29]. However, digital literacy reflects the knowledge and capabilities of older adults, rather than their engagement with digital devices or related difficulties. Little is known about the patterns of digital device use among older adults and the difficulties they face in adapting to digital technologies. Understanding this heterogeneity may provide a more nuanced view of digital device use and adaptation difficulty among older adults and inform more tailored digital inclusion strategies.

Therefore, this study aimed to explore heterogeneity in digital device use and adaptation difficulty among older adults and to examine the characteristics and service needs associated with the identified patterns. Specifically, this study addressed the following research questions: (1) What distinct patterns of digital device use and adaptation difficulty exist among older adults? (2) What sociodemographic and health-related factors are associated with the identified patterns? (3) How do service needs for digital adaptation differ across these patterns?

Methods

Study Design and Data

This was a cross-sectional secondary data analysis study using data sourced from the 2023 National Survey of Older Koreans (NSOK). The NSOK, conducted every 3 years since 2008, is

XSL•FO
RenderX

administered by the Ministry of Health and Welfare of South Korea and the Korean Institute for Health and Social Affairs under the Welfare of Senior Citizens Act. The survey serves as a foundational resource for policy development aimed at improving the quality of life for older Koreans by examining their diverse characteristics, living conditions, and needs. For the 2023 NSOK, 176 trained survey interviewers administered the survey via tablet-assisted personal interviews conducted face-to-face between September and November 2023.

Study Sample

The 2023 NSOK surveyed older adults aged 65 years and above across the nation. To ensure a representative sample, a proportional 2-stage stratified sampling method was used for the survey, using the 2021 Population and Housing Census of Korea as the primary sampling frame. Among the 10,178 older adults who participated in the survey, individuals with cognitive impairment, those who had a proxy respondent, and those with missing responses to key questions were excluded from the analysis. Consequently, data from a final sample of 7470 older adults were used for this study.

Measures

Patterns of Digital Device Use and Adaptation Difficulty

Patterns of digital device use and adaptation difficulty were assessed using 2 aspects of the 2023 NSOK: digital device activities and difficulty with digital adaptation. Digital device use was measured using 13 items, assessing whether participants used specific functions when they are using their digital devices: text messages, video calls, information searches, taking photos/videos, music, games, watching videos, social networking service (SNS), e-commerce, financial transactions, app search/installation, and machine-based ordering/check-in at service locations. Participants responded to each item with a "yes" or "no." Difficulty in digital adaptation was assessed using a single item on a 5-point Likert scale, ranging from "not difficult at all" (1) to "very difficult" (5).

Service Needs for Digital Adaptation

A single question was used to assess service needs for digital adaptation. Participants were asked to select the most necessary service from six options: provision of smart devices, support for smart device usage fees, expansion of public Wi-Fi, diversification of digital literacy education programs, development of senior-friendly smart devices and applications, and expansion of digital infrastructure to bridge the digital divide (eg, hands-on experience spaces). Previous studies have highlighted the importance of access to digital technologies, education and support for developing digital skills, and the usability of these technologies among older adults [2,6,11]. Based on these considerations, the six service needs were grouped into 3 broader categories according to their primary implementation purpose. Provision of smart devices, support for smart device usage fees, and expansion of public Wi-Fi were grouped as "digital accessibility support," which aims to improve access to digital technology by addressing financial and infrastructural barriers. Diversification of digital literacy education programs and expansion of education-based digital infrastructure were categorized as "digital literacy enhancement," which focuses on improving digital skills among older adults through experiential learning opportunities. Development of senior-friendly smart devices and apps was classified as "age-friendly digital solutions," which aims to make digital technology more tailored to older adults by reducing usability challenges.

Covariates

The covariates included participants' sociodemographic and health factors, identified from the literature [30]. Sociodemographic factors included age, sex, residential area, household type, annual household income, and education level [6,9,25]. Health factors comprised chronic diseases, cognitive function, and functional status, which have been directly linked to digital device use and adaptation difficulty [8,12,15], as well as smoking status, drinking, exercise, and self-rated health [14]. Chronic diseases were assessed as the number of physician-diagnosed chronic diseases. Cognitive function and functional status were evaluated using the Korean-Mini Mental State Examination, 2nd Edition (K-MMSE-2) and the Korean Instrumental Activities of Daily Living (K-IADL), respectively [31,32]. Questions on smoking, drinking, and exercise were answered with yes or no, while self-rated health was categorized as healthy, moderate, or unhealthy.

Statistical Analysis

To identify patterns of digital device use and adaptation difficulty among older adults, a latent class analysis (LCA) was conducted. LCA estimates the probability of individual responses to questions about digital device activities and difficulties in adapting to the digital society. The LCA model included 14 indicators: 13 binary items assessing digital device activities (yes/no) and 1 ordinal item assessing adaptation difficulty, measured on the original 5-point Likert scale ("not difficult at all" to "very difficult"). All the 14 indicators, including the 5-category difficulty item, were modeled according to the standard categorical LCA procedures, in which indicator variables—including those with more than 2 ordered response categories—are treated as categorical variables with class-specific response probabilities estimated for each category [33]. Multiple latent class models with an increasing number of classes were compared to optimize model fit and determine the most appropriate number of classes. Several statistical indicators, including the Akaike information criterion (AIC), Bayesian information criterion (BIC), sample size–adjusted BIC (aBIC), entropy, Lo-Mendell-Rubin (LMR) adjusted likelihood ratio test, bootstrapped likelihood ratio test (BLRT), and smallest estimated class proportions, were used to evaluate the model fit [33-35]. The preferred model was selected based on the lower AIC, BIC, or aBIC values, as lower values indicate better model fit [36]. Entropy was used to assess classification accuracy, with values above 0.80 generally considered acceptable [35,37]. The LMR and BLRT tests were conducted to compare models with k and k-1 classes, with a statistically significant difference ($P<.05$) indicating that a model with k classes provided a better fit than a model with k-1 classes [38]. The smallest estimated class proportion was set at 5% of the total sample size to ensure meaningful class sizes [39]. In addition to the statistical criteria, the interpretability and

XSL•FO
RenderX

parsimony of the identified classes were considered in selecting the final model [39-42]. Once the optimal latent class model was identified, latent class membership was incorporated into subsequent analyses using bias-adjusted 3-step approaches, which account for classification uncertainty rather than treating class assignment as an error-free observed variable [43]. Specifically, the 3-step approach with covariates (R3STEP) [44] was used to relate covariates to latent class membership, and the Bolck-Croon-Hagenaars (BCH) method [43,45] was used to derive classification-error-adjusted weights for examining differences in a distal outcome across latent classes.

Using the R3STEP method, multinomial logistic regression was performed to compare the sociodemographic and health characteristics of the identified latent classes. The odds ratio (OR) and corresponding 95% CIs were estimated to assess the likelihood of belonging to a specific class compared to the reference group. Finally, differences in service needs across the identified latent classes were examined using a BCH-weighted multinomial logistic regression model. Statistical significance was determined at a *P* value of <.05, and all multinomial logistic regression analyses were conducted after adjusting for covariates. The LCA was conducted in Mplus (version 8; Muthén & Muthén). R3STEP-based multinomial logistic regression and BCH-weighted multinomial logistic regression were also performed in Mplus. Descriptive statistical analyses were conducted in SPSS (version 27.0; IBM Corp).

Because the analyses in this study were conducted on a restricted subsample defined by specific inclusion criteria, survey weights calibrated for the whole NSOK population were not applied. Although omitting survey weights in analyses of nationally representative samples can lead to biased estimates and distorted statistical inference [46], using the original weights for a subsample without recalibration may itself introduce bias and increase estimation error [47]. Given that the analytic sample was no longer representative of the broader population, using the original weights could have led to misleading parameter estimates and incorrect variance estimates. Therefore, unweighted analyses were deemed more appropriate; this analytic approach was adopted in consultation with an experienced biostatistician to ensure sound analytic practice.

Ethical Considerations

Permission to use the data was granted by the Korea Institute for Health and Social Affairs, and this study was exempted from ethics approval by the institutional review board of the university affiliated with the research team (number 4-2025-0085).

Results

Sample Characteristics

Table 1 presents the sociodemographic and health-related characteristics of 7470 older adults aged 65 to 98 (mean 73.78, SD 6.65) years. More than half were female (n=4462, 59.7%) and lived with a spouse (n=4183, 56.0%). The majority resided in urban areas (n=5357, 71.7%). Regarding health factors, participants had an average of 2.11 (SD 1.57) chronic diseases (range 0‐13), scored an average of 26.25 (SD 3.16) on the K-MMSE-2 (range 10‐30), and reported a mean K-IADL score of 10.53 (SD 1.91; range 10‐33). Only 8.4% (n=625) of the participants were current smokers, 35.0% (n=2611) consumed alcohol, and 43.7% (n=3266) reported not engaging in regular exercise. Additionally, 45.0% (n=3363) of the participants rated their overall health as healthy.

XSL•FO
RenderX

Table . Sociodemographic and health-related characteristics (N=7470).

Variables	Values
Sociodemographic factors	
Age (y), mean (SD; range)	73.78 (6.65; 65 - 98)
Sex, n (%)	
Male	3008 (40.3)
Female	4462 (59.7)
Residential area, n (%)	
Urban	5357 (71.7)
Rural	2113 (28.3)
Household type, n (%)	
Living alone	2450 (32.8)
Living with spouse	4183 (56.0)
Multigenerational	837 (11.2)
Annual household income, n (%)	
1st quintile (low)	1541 (20.6)
2nd quintile	1553 (20.8)
3rd quintile	1524 (20.4)
4th quintile	1480 (19.8)
5th quintile (high)	1372 (18.4)
Education level, n (%)	
Elementary or less	3137 (42.0)
Middle-high	3823 (51.2)
College or above	510 (6.8)
Health factors	
Number of chronic diseases, mean (SD; range)	2.11 (1.57; 0 - 13)
K-MMSE-2[a], mean (SD; range)	26.25 (3.16; 10-30)
K-IADL[b], mean (SD; range)	10.53 (1.91; 10-33)
Smoking, n (%)	
Yes	625 (8.4)
No	6845 (91.6)
Drinking, n (%)	
Yes	2611 (35.0)
No	4859 (65.0)
Exercise, n (%)	
Yes	4204 (56.3)
No	3266 (43.7)
Self-rated health status, n (%)	
Healthy	3363 (45.0)
Moderate	2553 (34.2)
Unhealthy	1554 (20.8)

[a]K-MMSE-2: Korean-Mini-Mental State Examination, 2nd Edition.

[b]K-IADL: Korean Instrumental Activities of Daily Living.

XSL•FO
RenderX

Table 2 presents digital device use, adaptation difficulty, and perceived service needs for digital adaptation among older adults. Regarding overall participation in 13 digital device activities, approximately 80% of the participants could send and receive text messages, whereas more than 80% could not play games, use SNS, engage in e-commerce or financial transactions, search for or install apps, or use machines for ordering/check-in. Most participants reported either “difficult” (n=3116, 41.7%) or “very difficult” (n=1868, 25.0%) in their digital adaptation. In contrast, only a small proportion of participants found digital adaptation as “not difficult” (n=556, 7.4%) or “not difficult at all” (n=57, 0.8%). Regarding the preferred service needs for digital adaptation, the most frequently selected support area was digital accessibility support (n=2731, 36.6%), followed closely by digital literacy enhancement (n=2508, 33.6%) and age-friendly digital solutions (n=2231, 29.9%).

XSL•FO
RenderX

Table . Digital device use, adaptation difficulty, and service needs for digital adaptation (N=7470).

Variables	Participants, n (%)
Digital device use	
Receiving text messages	
Yes	6306 (84.4)
No	1164 (15.6)
Sending text messages	
Yes	5548 (74.3)
No	1922 (25.7)
Video calls	
Yes	3294 (44.1)
No	4176 (55.9)
Information search and browsing	
Yes	3759 (50.3)
No	3711 (49.7)
Taking photos/videos	
Yes	3925 (52.5)
No	3545 (47.5)
Listening to music	
Yes	2009 (26.9)
No	5461 (73.1)
Playing games	
Yes	1058 (14.2)
No	6412 (85.8)
Watching videos	
Yes	3308 (44.3)
No	4162 (55.7)
SNS[a]	
Yes	662 (8.9)
No	6808 (91.1)
E-commerce	
Yes	894 (12.0)
No	6576 (88.0)
Financial transactions	
Yes	1479 (19.8)
No	5991 (80.2)
Searching for and installing apps	
Yes	973 (13.0)
No	6497 (87.0)
Machine-based ordering and check-in at service locations	
Yes	1211 (16.2)
No	6259 (83.8)
Adaptation difficulty	
Not difficult at all	57 (0.8)

XSL•FO
RenderX

Variables	Participants, n (%)
Not difficult	556 (7.4)
Moderate	1873 (25.1)
Difficult	3116 (41.7)
Very difficult	1868 (25.0)
Service needs for digital adaptation	
Digital accessibility support	2731 (36.6)
Digital literacy enhancement	2508 (33.6)
Age-friendly digital solutions	2231 (29.9)

[a]SNS: social networking service.

Patterns of Digital Device Use and Adaptation Difficulty

Table 3 shows the model fit indices for the LCA, comparing models with 2 to 5 classes, to identify patterns of digital device use and adaptation difficulty among older adults. The optimal number of latent classes was determined to be 4 based on multiple criteria, including model fit statistics, interpretability, parsimony, and the elbow pattern observed in Figure 1. Although the AIC, BIC, and aBIC values consistently decreased as additional classes were added, the rate of improvement diminished beyond the 4-class model. Entropy values exceeded 0.8 for the 2-, 3-, and 4-class models, indicating acceptable classification accuracy, whereas the 5-class model fell below this threshold. Both the 3- and 4-class models demonstrated adequate statistical fit; nevertheless, the 4-class model offered superior interpretability by distinguishing a text-only group from the moderately active and inactive groups, thereby allowing for a more nuanced characterization of patterns of digital device use and adaptation difficulty. Additionally, the elbow plot exhibited an inflection at the 4-class point, with minimal information loss when moving from 4 to 5 classes, further supporting its selection. Furthermore, its smallest class accounted for 14.9% (n=1112) of the sample, indicating that all classes were reasonably sized. Accordingly, the 4-class model was retained for subsequent analysis.

Table . Model fit indices for latent class analysis.

Class	AIC[a]	BIC[b]	aBIC[c]	Entropy	LMR[d] *P* value	BLRT[e] *P* value	Smallest class size (%)
2	91,769.431	92,011.584	91,900.361	.91	<.001	<.001	49.3
3	85,215.300	85,581.988	85,413.565	0.887	<.001	<.001	19.0
4	82,669.662	83,160.886	82,935.263	0.874	<.001	<.001	14.9
5	82,250.775	82,866.535	82,583.712	0.799	.009	<.001	12.6

[a]AIC: Akaike information criterion.

[b]BIC: Bayesian information criterion.

[c]aBIC: sample size–adjusted BIC.

[d]LMR: Lo-Mendell-Rubin.

[e]BLRT: bootstrapped likelihood ratio test.

XSL•FO
RenderX

Figure 1. Elbow plot. aBIC: sample size–adjusted Bayesian information criterion; AIC: Akaike information criterion; BIC: Bayesian information criterion.

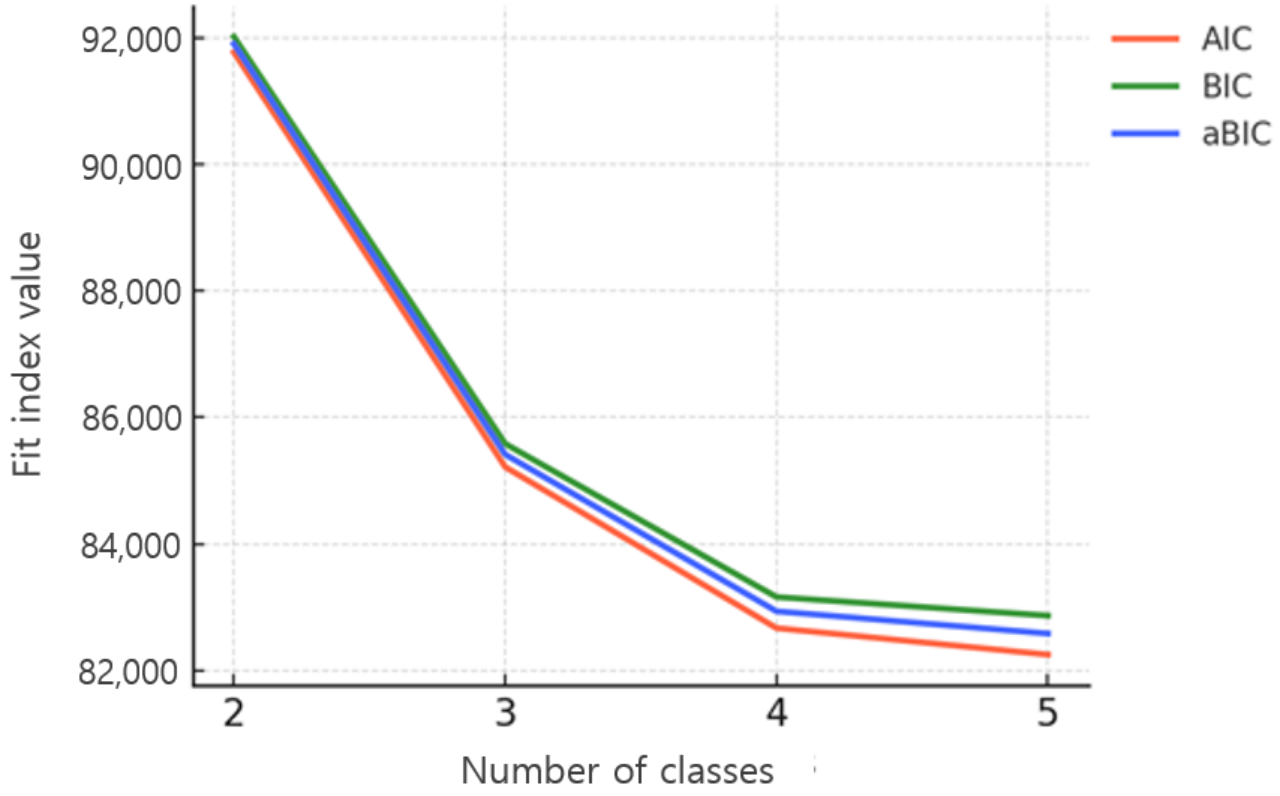


Figure 2 displays the latent class of digital device use and adaptation difficulty, along with estimated probabilities for each activity, and Multimedia Appendix 1 provides detailed characteristics by latent class. Participants were classified into to 4 classes based on their patterns of digital device use and adaptation difficulty. Class 1 (n=1112, 14.9%), labeled "digitally active with lower difficulty (active)," represented individuals with the broadest digital device use and the least difficulty, although it was the smallest group in size. Most participants in this class could perform nearly all digital activities. The least performed digital activities were playing games and using SNS, which were used by approximately half of the group. Regarding perceived difficulty, 45.0% (n=500) of the participants reported digital adaptation as moderate, and 25.9% (n=288) as not difficult. Class 2 (n=2492, 33.4%), the largest group, was labeled "moderately active (moderate)" and exhibited mid-range probabilities across all activities. The least performed tasks in this group were e-commerce and app search/installation. Most participants could engage in approximately half of the digital activities. However, approximately 60% (n=1469) reported difficulties with digital adaptation. Class 3 (n=2082, 27.9%), labeled "text messages only (text-only)," was characterized by a high probability of only sending and receiving text messages, with over 90% being able to perform these 2 tasks. A smaller subset (approximately 20%) could perform 4 activities—video calls, information searches, taking photos/videos, and watching videos. The remaining digital device activities were rarely performed. Only 4.2% (n=92) of the participants in this class reported no difficulty in digital adaptation, while 74.2% (n=1546) reported it as difficult. Class 4 (n=1784, 23.9%), labeled "digitally inactive with difficulty (inactive)," was characterized by the most limited digital device use and the greatest difficulty in digital adaptation. Nearly all digital activities were reported as unperformable, except for sending text messages, which 35.6% (n=635) of the participants could perform. A striking 92.2% (n=1645) of this group reported digital adaptation as difficult or very difficult. The estimated proportions of participants reporting digital adaptation as difficult or very difficult were 29.2% (n=324), 59% (n=1469), 74.2% (n=1546), and 92.2% (n=1645) across classes 1 (active) through 4 (inactive), respectively.

XSL•FO
RenderX

Figure 2. Four latent classes of digital device use and adaptation difficulty.

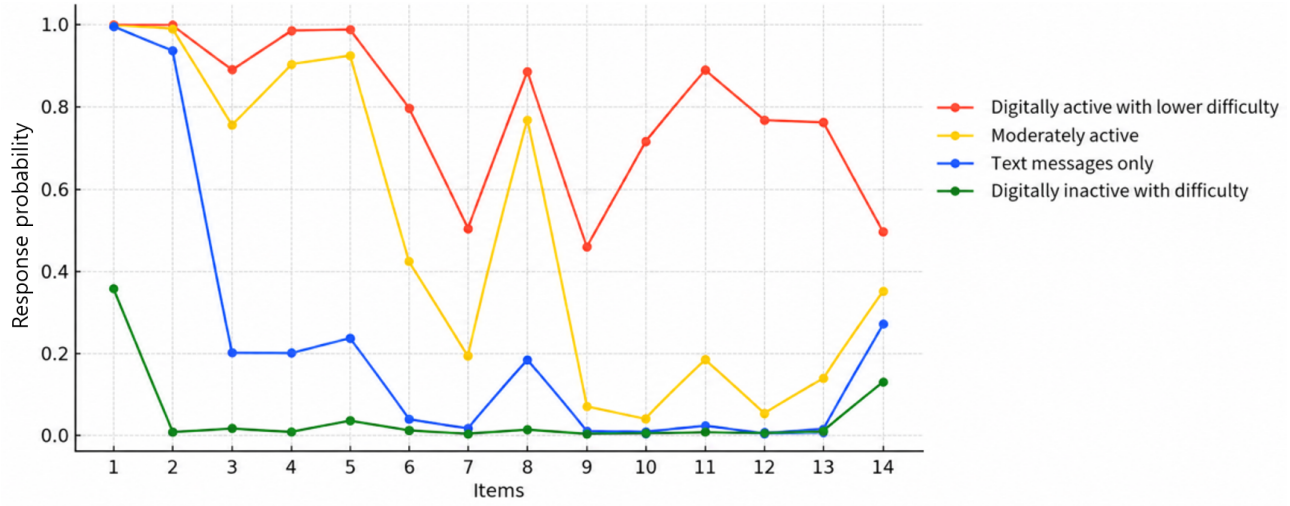


Table 4 presents the results of an R3STEP-based multinomial logistic regression analysis of sociodemographic and health-related factors associated with the identified latent classes, using the "active" class as the reference group. Older age was significantly associated with greater odds of belonging to classes characterized by more limited digital device use and greater adaptation difficulty across all comparisons, with ORs ranging from 1.136 (95% CI 1.110-1.162) to 1.335 (95% CI 1.301-1.370). Conversely, younger age was associated with a greater likelihood of membership in the active class. Living in urban areas was associated with a lower likelihood of being in the "inactive" class (OR 0.675, 95% CI 0.525 - 0.867). Regarding income, participants in lower-income quintiles, especially the first to third quintiles, were significantly more likely to be in the "inactive," "text-only," or "moderate" classes compared with that observed with those in the highest quintile. Similarly, individuals with lower educational attainment—particularly those with only elementary education—had substantially higher odds of belonging to classes characterized by more limited digital device use and greater adaptation difficulty (OR 22.956 for "inactive" vs "active").

XSL•FO
RenderX

Table . R3STEP (3-step approach with covariates)–based multinomial logistic regression of latent class membership by sociodemographic and health characteristics of older adults.

Variables	Inactive vs active (ref), OR[a] (95% CI)	Text-only vs active (ref), OR (95% CI)	Moderate vs active (ref), OR (95% CI)
Sociodemographic factors			
Age (y)	1.335 (1.301-1.370)	1.222 (1.192-1.253)	1.136 (1.110-1.162)
Sex			
Male	0.851 (0.656-1.104)	0.907 (0.727-1.132)	0.883 (0.725-1.076)
Female	__[b]	—	—
Residential area			
Urban	0.675 (0.525-0.867)	1.182 (0.938-1.491)	0.983 (0.794-1.218)
Rural	—	—	—
Household type			
Living alone	0.774 (0.516-1.161)	1.009 (0.714-1.427)	0.986 (0.722-1.347)
Living with spouse	0.813 (0.568-1.164)	1.017 (0.755-1.369)	1.181 (0.914-1.527)
Multigenerational	—	—	—
Annual household income			
1st quintile (low)	2.996 (1.968-4.563)	2.787 (1.980-3.922)	1.458 (1.072-1.981)
2nd quintile	4.391 (2.841-6.786)	4.051 (2.833-5.793)	2.586 (1.840-3.634)
3rd quintile	2.019 (1.368-2.980)	2.201 (1.627-2.979)	1.609 (1.238-2.091)
4th quintile	1.063 (0.725-1.560)	1.448 (1.100-1.907)	1.333 (1.076-1.650)
5th quintile (high)	—	—	—
Education level			
Elementary or less	22.956 (12.773-41.258)	15.370 (9.773-24.174)	5.074 (3.366-7.650)
Middle–high	2.549 (1.523-4.265)	2.894 (2.006-4.174)	1.962 (1.524-2.525)
College or above	—	—	—
Health factors			
Number of chronic diseases	1.081 (0.995-1.175)	1.067 (0.988-1.151)	1.086 (1.013-1.165)
K-MMSE-2[c]	0.676 (0.637-0.717)	0.761 (0.720-0.805)	0.837 (0.796-0.879)
K-IADL[d]	1.165 (1.024-1.326)	1.138 (1.000-1.294)	1.014 (0.894-1.150)
Smoking			
Yes	1.687 (1.118-2.546)	1.367 (0.982-1.903)	1.236 (0.924-1.654)
No	—	—	—.
Drinking			
Yes	0.469 (0.362-0.608)	0.548 (0.444-0.677)	0.822 (0.684-0.987)
No	—	—	—
Exercise			
Yes	0.386 (0.307-0.485)	0.413 (0.339-0.503)	0.700 (0.587-0.835)
No	—	—	—
Self-rated health status			
Healthy	0.182 (0.120-0.276)	0.325 (0.219-0.481)	0.466 (0.319-0.682)
Moderate	0.439 (0.288-0.668)	0.747 (0.496, 1.125)	0.680 (0.454-1.020)
Unhealthy	—	—	—

[a]OR: odds ratio.

XSL•FO
RenderX

[b]Not applicable.

[c]K-MMSE-2: Korean-Mini Mental State Examination, 2nd Edition.

[d]K-IADL: Korean Instrumental Activities of Daily Living.

Regarding health factors, a higher number of chronic diseases was associated with increased odds of belonging to the "moderate" class (OR 1.086, 95% CI 1.013 - 1.165). Older adults with lower cognitive function compared with the others, as measured by the K-MMSE-2, were significantly more likely to be digitally less active, with ORs of 0.676 (95% CI 0.637-0.717), 0.761 (95% CI 0.720-0.805), and 0.837 (95% CI 0.796-0.879) for the "inactive," "text-only," and "moderate" classes, respectively. Greater dependence in instrumental activities of daily living, as indicated by higher K-IADL scores, also increased the likelihood of being classified into the "inactive" class. Additionally, current smokers were more likely to belong to the "inactive" class (OR 1.687, 95% CI 1.118 - 2.546). Individuals who consumed alcohol or engaged in regular exercise had lower odds of belonging to classes characterized by more limited digital device use and greater adaptation difficulty than those who did not. Compared with participants who rated their health as unhealthy, those who perceived their health as healthy were significantly less likely to belong to the inactive (OR 0.182, 95% CI 0.120-0.276), text-only (OR 0.325, 95% CI 0.219-0.481), or "moderate" (OR 0.466, 95% CI 0.319-0.682) classes.

Service Needs for Digital Adaptation

Multimedia Appendix 1 presents tables that provide the descriptive analysis of sociodemographic characteristics, digital device use and adaptation difficulty, and preferred service needs for digital adaptation across the 4 latent classes. Among participants in the "inactive" class, the most frequently selected service need was age-friendly digital solutions (n=727, 40.8%), followed by digital accessibility support (n=544, 30.5%) and digital literacy enhancement (n=513, 28.7%). In the "text-only" class, a shift in preference was observed, with digital accessibility support emerging as the most selected option (n=787, 37.8%), slightly ahead of digital literacy enhancement (n=657, 31.6%) and age-friendly solutions (n=638, 30.6%). The "moderate" class expressed the highest preference for digital accessibility support (n=987, 39.6%), followed by digital literacy enhancement (n=870, 34.9%). Finally, participants in the "active" class demonstrated a clear preference for digital literacy enhancement (n=468, 42.1%), with digital accessibility support (n=413, 37.1%) also rated as high. Only 20.8% (n=231) of this class selected age-friendly solutions.

Table 5 presents the results of the BCH-weighted multinomial logistic regression examining the associations between the identified latent classes and preferred service needs for digital adaptation, with "digital literacy enhancement" used as the reference category for service needs. After adjusting for sociodemographic and health-related covariates, participants in the "inactive" class compared with the "active" class were significantly more likely to prefer age-friendly digital solutions over digital literacy enhancement (OR 1.978, 95% CI 1.510 - 2.591) but showed no significant difference in preference for digital accessibility support (OR 1.012, 95% CI 0.787 - 1.302). Similarly, members of the "text-only" class were also more likely to prefer age-friendly solutions (OR 1.536, 95% CI 1.218 - 1.937), with no significant difference observed for accessibility support (OR 1.158, 95% CI 0.938 - 1.429). The "moderate" class demonstrated significantly higher odds of preferring both digital accessibility support (OR 1.236, 95% CI 1.029 - 1.485) and age-friendly digital solutions (OR 1.294, 95% CI 1.051 - 1.593) compared to digital literacy enhancement.

Table . Bolck-Croon-Hagenaars (BCH)–weighted multinomial logistic regression of service needs for digital adaptation across identified latent classes.[a]

Class	Digital accessibility support vs digital literacy enhancement (ref), OR[b] (95% CI)	Age-friendly digital solutions vs digital literacy enhancement (ref), OR (95% CI)
Inactive	1.012 (0.787, 1.302)	1.978 (1.510-2.591)
Text-only	1.158 (0.938-1.429)	1.536 (1.218-1.937)
Moderate	1.236 (1.029-1.485)	1.294 (1.051-1.593)
Active	—[c]	—

[a]Analyses were adjusted for the following covariates: age, sex, residential area, household type, annual household income, education level, number of chronic diseases, Korean-Mini Mental State Examination-2, Korean Instrumental Activities of Daily Living, smoking, drinking, exercise, and self-rated health status.

[b]OR: odds ratio.

[c]Not applicable.

Discussion

Principal Findings

To the best of our knowledge, this study is the first to assess preferred service needs for digital adaptation across distinct patterns of digital device use and adaptation difficulty among older adults. Using LCA and multinomial logistic regression, we identified 4 distinct classes, examined the characteristics associated with each class, and evaluated their corresponding service preferences. The LCA revealed 4 distinct classes: the "active" class, characterized by the broadest digital device use and minimal difficulty in digital adaptation; the "moderate" class, with mid-range probabilities across most digital activities; the "text-only" class, defined by high engagement in only sending and receiving text messages; and the "inactive" class,

XSL•FO
RenderX

which exhibited the lowest digital activities and the greatest difficulty in digital adaptation. The results from the multinomial logistic regression indicated that older age, rural residence, lower household income, and lower educational attainment were consistently associated with classes characterized by more limited digital device use and greater adaptation difficulty. Regarding health factors, individuals with reduced cognitive functioning, greater dependence in daily living activities, and poorer self-rated health were significantly more likely to belong to the "inactive" or "text-only" classes. Finally, analysis of preferred service needs by latent class revealed that individuals in the "inactive" and "text-only" classes expressed a significantly greater preference for age-friendly digital solutions over digital literacy enhancement, while those with more active digital device users prioritized literacy and accessibility support.

Among the sociodemographic characteristics, age and education have long been recognized as key determinants of the digital divide among older adults. Prior research has consistently shown that older age is associated with lower levels of digital literacy, competency, and proficiency [48-50]. Similarly, higher educational attainment has been linked to more frequent use of e-services and digital technologies [14,48,50]. Our findings align with these established patterns. In our study, participants with older age and lower levels of education were significantly associated with membership in classes with lower levels of digital device use and greater adaptation difficulty. Notably, participants with only elementary-level education had the highest likelihood of belonging to the "inactive group" across all comparisons. These findings underscore the importance of targeted digital inclusion strategies that address educational disparities among older populations. Furthermore, our analysis revealed that older adults with higher household income levels and those living in urban areas were more likely to be members of the "active" class. This suggests that socioeconomic resources, including financial capacity and geographic location, might play an important role in shaping digital device use and adaptation among older adults [51-53]. Enhancing digital inclusion among older adults, therefore, requires addressing both individual-level factors and the broader structural and financial barriers that systematically create a digital disadvantage in increasingly digital societies. Specifically, improving the affordability of digital devices and internet service fees, alongside reducing disparities in rural digital infrastructure, might help to address structural barriers to digital access [7]. Expanding access to digital literacy education through community-based educators might further support skill development among vulnerable older adults [51]. Finally, addressing the adaptation and usage barriers among older adults when they are engaging with increasingly digitized public and health services might help to reduce their exclusion from these services [11,54].

Health factors were significantly associated with patterns of digital device use and adaptation difficulty among older adults. Previous research has shown that cognitive impairments mostly lead to functional limitations over time, further complicating digital engagement [55,56]. In particular, cognitive decline has been identified as a major barrier to digital engagement, as it can impair older adults' ability to learn new technologies, understand technical terminology, and comprehend or remember digital procedures [2,57]. On the other hand, emerging evidence has suggested that smartphone and tablet use may be associated with cognitive benefits among older adults, including those with and without cognitive impairment [58], suggesting a potential association between engagement and cognitive outcomes. Given that individuals with severe cognitive impairment or those requiring proxy respondents were excluded, the associations observed in this study likely reflect a relatively higher-functioning subset of older adults. Nonetheless, cognitive function may be an important consideration when tailoring digital support for this population. In addition, previous research has suggested that, among adults with chronic disease, older adults are generally less likely than younger adults to use telemedicine, often due to difficulties navigating telehealth platforms or adapting to small screens and text [11,59,60]. Consistent with these findings, our study revealed that individuals with more chronic diseases, reduced cognitive function, and higher levels of dependence in instrumental activities of daily living were significantly more likely to belong to classes reflecting greater digital limitations [14,61,62]. This disproportionate use of digital devices may be associated with a compounded burden on older adults—particularly those with cognitive and physical impairments—who often have the greatest health care needs. This pattern may suggest potential limitations in their access to digitally delivered health services.

Furthermore, our study found that self-rated health and health behaviors were closely associated with patterns of digital device use and adaptation difficulty. Participants who perceived themselves as unhealthy were significantly more likely to be classified as digitally inactive. Self-rated health, one of the most commonly used indicators in health and social research, serves as a comprehensive measure of overall health status—capturing aspects of health that may extend beyond clinical diagnoses [63]. Consistent with other health-related factors, poor self-rated health may be linked to a dual burden of health-related vulnerability and technological exclusion, which may reflect greater disparities in access to digital health resources, services, and information. Moreover, health behaviors, including smoking, alcohol consumption, and physical exercise, were also significantly associated with latent class membership. While these findings suggest that certain health behaviors may align with greater digital technology use and adaptation, the associations remain underexplored and, in some cases, inconsistent in the existing literature [14]. Further research is needed to clarify these relationships and determine whether the observed associations reflect health behaviors themselves or broader differences in health and functional status.

Analysis of service needs revealed that older adults in the "inactive" and "text-only" classes expressed a significantly greater preference for age-friendly digital solutions over digital literacy enhancement. Notably, members of the "inactive" class were nearly twice as likely as those in the "active" class to prioritize age-friendly features, suggesting that their barriers extend beyond knowledge gaps or lack of training. This is consistent with recent studies that have identified challenges such as difficulties navigating digital technologies, poor website usability, lack of adaptive features, and poorly designed user

XSL•FO
RenderX

interfaces as key obstacles to technology use among older adults [2,55]. Conversely, simple and customizable technologies aligned with older adults' abilities and preferences are considered facilitators of digital adoption—supporting our findings that older adults require more than training alone.

Although our survey did not assess preferences for specific design features, participants in the inactive and text-only classes were significantly more likely to prefer age-friendly digital solutions to digital literacy enhancement. These findings suggest that improving the usability of digital services may be as important as providing additional digital literacy education for older adults in these groups. In response to these limitations, the concept of an "age-friendly lens" has gained attention in recent literature. This approach moves beyond standardized assumptions of the "average" user and embraces the full spectrum of aging and diversity [64]. While traditional interventions have emphasized raising awareness, enhancing self-efficacy, and providing digital literacy training [65-67], age-friendly technology incorporates both the biological realities of aging and the broader values, aspirations, and lived experiences of older adults. It emphasizes inclusive, human-centered design rather than merely modifying existing technologies. Emerging technologies, including care robots, demonstrate this paradigm shift; older adults should be engaged from the outset through participatory or co-design processes rather than being treated as secondary end users [64,68]. Potential examples include simplified navigation, voice-controlled interfaces, and extended screen timeouts, which may reduce usability barriers for older adults with physical or cognitive limitations [64]. However, the specific design features required might differ according to the digital task. For example, booking a medical appointment, using a telehealth application, completing a financial transaction, and interacting with a self-service kiosk involve different interaction demands; therefore, they might require different age-friendly design strategies [64]. Importantly, these examples are illustrative rather than directly evaluated in the present study, and our findings do not indicate the design features that should be prioritized for specific digital tasks.

Age-friendly usability also spans domains such as providing access to a variety of services, simple and intuitive use, and security that ensures privacy and authenticity [69]. Collectively, these design principles highlight the practical implications of our findings and may inform the design of public services and health care platforms that better support digital inclusion and healthy aging. Although traditional alternatives, such as in-person services and telephone support, are available, they are increasingly becoming limited as public services and health care shift toward digital platforms. Consequently, the efficiency of digitalization might not be experienced by older adults who cannot use digital systems, further excluding them from the benefits of digitization. Improving digital inclusion is therefore essential to enable older adults to benefit from convenient access to public services, health information, and health care. These features should be carefully considered and implemented in public services to promote digital inclusion among the most digitally inactive older adults, with similar implications for health care platforms supporting healthy aging. Additionally, the "moderate" class exhibited a balanced demand for both digital accessibility and age-friendly support, indicating that their needs span both usability and skill development. These differentiated needs highlight the shortcomings of "one-size-fits-all" digital inclusion strategies and emphasize the importance of tailoring interventions to different patterns of digital technology use and adaptation. Future research should adopt longitudinal designs to examine changes in class membership over time, identify age-friendly design features that best support different digital tasks, and evaluate the effectiveness of tailored digital interventions for each latent class.

Recommendations for Implementation

Taken together, these findings have meaningful implications for advancing digital inclusion among older adults and guiding the development of responsive, subgroup-specific strategies to reduce digital divides. Local governments, health care systems, and welfare organizations could apply these latent classes to segment their programs more effectively, offering tiered digital support rather than relying on one-size-fits-all digital literacy workshops. This approach underscores the need for age-sensitive service design across both public and private sectors, ensuring that interventions are aligned with the specific needs and capabilities of diverse older adult groups. Prior studies provided concrete design examples to operationalize age-sensitive design principles. Specifically, for visual accessibility, larger high-contrast fonts and simplified navigation menus with minimal hierarchical steps can help reduce cognitive and visual load for those with sensory decline [11,70]. Moreover, regarding interaction and input, touch-friendly screens with larger clickable areas, voice-command functions, and audio feedback can enable older adults with reduced dexterity or vision to operate devices more independently [11,70,71]. To support ongoing engagement, tailoring interface complexity to users' abilities and offering clear privacy and security indicators can build confidence and promote sustained use [11]. These design considerations illustrate how digital services can be better aligned with older adults' functional needs. Building on these principles, subgroup-specific intervention strategies can be further refined. For the "inactive" group, individualized instruction or small group training sessions, supplemented with real-time technical assistance, may help address multifaceted barriers and support sustained digital engagement. For the "text-only" group, a stepwise introduction to more diverse digital tasks, such as video communication, mobile apps, and online services, can gradually expand their digital participation beyond messaging. Additionally, policy initiatives should focus on reducing structural barriers by increasing access to digital devices and strengthening infrastructure, particularly in socioeconomically disadvantaged or rural areas. Aligning intervention content, delivery approaches, and broader policy supports with the distinct needs of each subgroup will enable digital inclusion efforts to improve both digital literacy and practical device-use skills among older adults.

Limitations and Strengths

This study has some limitations. First, because this study used secondary data, the available measures were limited to 13 binary

XSL•FO
RenderX

items assessing digital device use and one 5-point Likert item assessing digital adaptation difficulty. As these items were from the secondary data rather than a validated digital competency scale, the identified patterns should be interpreted as an exploratory classification of digital device use and adaptation difficulty. Future studies using validated, multidimensional measures are needed to confirm and extend these patterns. Second, the 6 service-need options were grouped into 3 categories based on dimensions identified in the previous literature rather than on a validated classification framework. In particular, "expansion of digital infrastructure to bridge the digital divide (eg, hands-on experience spaces)" could plausibly be classified as either literacy- or access-related; this categorization involves interpretive judgment, and alternative classifications could yield different patterns in the analysis of service-preference findings. Third, because the measure relied on self-report, responses may be subject to social desirability bias, with some participants potentially overstating or understating their actual digital device use and adaptation difficulty in line with perceived expectations in the digital era. Fourth, individuals with severe cognitive impairment or those requiring proxy responses were excluded due to the self-reported nature of the measures. Consequently, our findings may understate the digital exclusion and needs of the most vulnerable older adults, and associations involving cognitive function should be interpreted within this narrower range. Finally, the cross-sectional design limits the ability to infer causal relationships. A longitudinal approach would be necessary to capture intra-individual change in digital device use and adaptation difficulty, particularly given that older adults' health conditions are likely to change over time. Such an approach would allow for the analysis of how patterns of digital device use and adaptation evolve with aging. Despite these limitations, this study has important strengths. Its primary contribution lies in its focus on older adults, who are at the highest risk of digital exclusion. By identifying heterogeneity within the older population and examining the corresponding service needs, the study provides meaningful insights for developing tailored digital inclusion strategies. These findings are particularly relevant for supporting the most vulnerable subgroup, that is, the digitally inactive group, by informing interventions that address specific barriers to digital inclusion.

Conclusions

This study identified 4 distinct latent classes of the patterns of digital device use and adaptation difficulty among older adults and examined the sociodemographic and health-related factors associated with digital inclusion. Notably, digitally inactive older adults were more likely to be older, less educated, and had lower incomes and reduced cognitive and functional abilities compared with their digitally active counterparts. Additionally, those in classes reflecting greater digital limitations expressed a stronger preference for age-friendly digital services over conventional digital literacy training, suggesting that usability and design-related barriers may be more relevant than skill-based limitations for many participants in this group. These findings underscore the need for differentiated and targeted digital inclusion strategies that go beyond basic skills training to address cognitive, educational, and usability challenges. Tailoring digital support to diverse patterns of digital use and adaptation difficulty may help policymakers and service providers to develop more responsive digital inclusion strategies and facilitate older adults' access to essential digital services, including health care services. Future studies should build on these findings to deepen the understanding of digital behavior trajectories among aging populations and inform evidence-based policies and practices that promote digital inclusion and support healthy aging.

Acknowledgments

Generative AI (ChatGPT-5.6 Sol [OpenAI] and Claude 5.5 [Anthropic]) tools were used for language editing and improving the clarity of the manuscript.

Funding

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education [Grant no. RS-2020-NR049581] and the Catholic Medical Center Research Foundation made in the program year of 2026. The funding source had no involvement in the study design, data collection, data analysis, interpretation of data, or manuscript writing.

Data Availability

The data used in this study are publicly available.

Authors' Contributions

Conceptualization: SL
Data curation: SL
Formal analysis: SL, MY
Funding acquisition: EC
Methodology: SL
Supervision: EC
Writing – original draft: SL, MY, JC, JK, EC

XSL•FO
RenderX

Writing – review & editing: SL, MY, JC, JK, EC
All the authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sociodemographic characteristics, digital device use, adaptation difficulty, and service needs by the identified latent classes.
[DOCX File, 29 KB - aging_v9i1e91770_app1.docx]

References

1. Abernethy A, Adams L, Barrett M, et al. The promise of digital health: then, now, and the future. NAM Perspect 2022;2022. [doi: 10.31478/202206e] [Medline: 36177208]
2. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
3. Global strategy on digital health 2020-2025. : World Health Organization; 2021 URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2025-05-26]
4. Kim BY, Lee J. Smart devices for older adults managing chronic disease: a scoping review. JMIR Mhealth Uhealth 2017;5(5):e69. [doi: 10.2196/mhealth.7141] [Medline: 28536089]
5. Ndayishimiye C, Lopes H, Middleton J. A systematic scoping review of digital health technologies during COVID-19: a new normal in primary health care delivery. Health Technol (Berl) 2023;13(2):273-284. [doi: 10.1007/s12553-023-00725-7] [Medline: 36628261]
6. Charness N, Boot WR. A grand challenge for psychology: reducing the age-related digital divide. Curr Dir Psychol Sci 2022;31(2):187-193. [doi: 10.1177/09637214211068144] [Medline: 35754678]
7. Lythreatis S, Singh SK, El-Kassar AN. The digital divide: a review and future research agenda. Technol Forecast Soc Change 2022 Feb;175:121359. [doi: 10.1016/j.techfore.2021.121359]
8. Reiners F, Sturm J, Bouw LJW, Wouters EJM. Sociodemographic factors influencing the use of eHealth in people with chronic diseases. Int J Environ Res Public Health 2019;16(4):645. [doi: 10.3390/ijerph16040645] [Medline: 30795623]
9. Walker DM, Hefner JL, Fareed N, Huerta TR, McAlearney AS. Exploring the digital divide: age and race disparities in use of an inpatient portal. Telemed J E Health 2020 May;26(5):603-613. [doi: 10.1089/tmj.2019.0065] [Medline: 31313977]
10. Hvalič-Touzery S, Laznik J, Petrovčič A. "I'm still struggling with it, and it scares me": a qualitative analysis of older adults' experiences with digital health portals during and after COVID-19. Digit Health 2024;10. [doi: 10.1177/20552076241282247] [Medline: 39372805]
11. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
12. Gell NM, Rosenberg DE, Demiris G, LaCroix AZ, Patel KV. Patterns of technology use among older adults with and without disabilities. Gerontologist 2015 Jun;55(3):412-421. [doi: 10.1093/geront/gnt166] [Medline: 24379019]
13. Pirhonen J, Lolich L, Tuominen K, Jolanki O, Timonen V. "These devices have not been made for older people's needs"—older adults' perceptions of digital technologies in Finland and Ireland. Technol Soc 2020 Aug;62:101287. [doi: 10.1016/j.techsoc.2020.101287]
14. Shin J, Kang H, Choi S, Chu SH, Choi J. Identifying profiles of digital literacy among community-dwelling Korean older adults: latent profile analysis. J Med Internet Res 2025;27:e57122. [doi: 10.2196/57122] [Medline: 39969960]
15. Wu YH, Lewis M, Rigaud AS. Cognitive function and digital device use in older adults attending a memory clinic. Gerontol Geriatr Med 2019;5. [doi: 10.1177/2333721419844886] [Medline: 31080848]
16. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37824584]
17. Ji H, Dong J, Pan W, Yu Y. Associations between digital literacy, health literacy, and digital health behaviors among rural residents: evidence from Zhejiang, China. Int J Equity Health 2024;23(1). [doi: 10.1186/s12939-024-02150-2] [Medline: 38594723]
18. Lee H, Choi JY, Kim SW, et al. Digital health technology use among older adults: exploring the impact of frailty on utilization, purpose, and satisfaction in Korea. J Korean Med Sci 2024;39(1):e7. [doi: 10.3346/jkms.2024.39.e7] [Medline: 38193326]
19. Liu S, Lu Y, Wang D, et al. Impact of digital health literacy on health-related quality of life in Chinese community-dwelling older adults: the mediating effect of health-promoting lifestyle. Front Public Health 2023;11. [doi: 10.3389/fpubh.2023.1200722] [Medline: 37415711]
20. Stellefson M, Paige SR, Alber JM, et al. Association between health literacy, electronic health literacy, disease-specific knowledge, and health-related quality of life among adults with chronic obstructive pulmonary disease: cross-sectional study. J Med Internet Res 2019;21(6):e12165. [doi: 10.2196/12165] [Medline: 31172962]

XSL•FO
RenderX

21. 2023 survey on broadcasting media usage behavior report [Article in Korean]. Korea Media and Communications Commission (KMCC). 2023. URL: https://www.kmcc.go.kr/user.do?boardId=1027&boardSeq=59157&mode=view&page=A02060100 [accessed 2026-11-18]
22. Household internet penetration and computer ownership rates [Article in Korean]. Indicator Nuri (e-National Indicators). 2024. URL: https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=1345 [accessed 2026-11-18]
23. Kang EN, Kim HS, Jung CW, et al. National survey of older adults [Report in Korean]. : Ministry of Health and Welfare, Republic of Korea (MOHW); 2023 URL: https://www.mohw.go.kr/board.es?mid=a10411010100&bid=0019&tag=&act=view&list_no=1483359 [accessed 2026-11-13]
24. Lee H. Analysis of the impact of digital literacy on life satisfaction (2019-2022) for older adults in South Korea: a national community-based panel study. Sci Rep 2024;14(1). [doi: 10.1038/s41598-024-71397-0] [Medline: 39223233]
25. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
26. Tsai HYS, Shillair R, Cotten SR. Social support and "playing around": an examination of how older adults acquire digital literacy with tablet computers. J Appl Gerontol 2017 Jan;36(1):29-55. [doi: 10.1177/0733464815609440] [Medline: 26491029]
27. Finkelstein R, Wu Y, Brennan-Ing M. Older adults' experiences with using information and communication technology and tech support services in New York City: findings and recommendations for post-pandemic digital pedagogy for older adults. Front Psychol 2023;14:1129512. [doi: 10.3389/fpsyg.2023.1129512] [Medline: 37138998]
28. Ilomäki L, Paavola S, Lakkala M, Kantosalo A. Digital competence—an emergent boundary concept for policy and educational research. Educ Inf Technol 2016 May;21(3):655-679. [doi: 10.1007/s10639-014-9346-4]
29. Pangrazio L, Godhe AL, Ledesma AGL. What is digital literacy? A comparative review of publications across three language contexts. E-Learn Digit Media 2020;17(6):442-459. [doi: 10.1177/2042753020946291]
30. Shi Z, Du X, Li J, Hou R, Sun J, Marohabutr T. Factors influencing digital health literacy among older adults: a scoping review. Front Public Health 2024;12. [doi: 10.3389/fpubh.2024.1447747] [Medline: 39555039]
31. Song M, Lee SH, Yu KH, Kang Y. Development and validation of the full version of story memory in the Korean-Mini Mental State Examination, 2nd Edition: Expanded Version (K-MMSE-2: EV). Dement Neurocogn Disord 2019;18(3):96-104. [doi: 10.12779/dnd.2019.18.3.96] [Medline: 31681444]
32. Won CW, Rho YG, Sunwoo D, Lee YS. The validity and reliability of Korean Instrumental Activities of Daily Living (K-IADL) scale [Article in Korean]. J Korean Geriatr Soc 2002;6(4):273-280 [FREE Full text]
33. Weller BE, Bowen NK, Faubert SJ. Latent class analysis: a guide to best practice. J Black Psychol 2020;46(4):287-311. [doi: 10.1177/0095798420930932]
34. Berlin KS, Williams NA, Parra GR. An introduction to latent variable mixture modeling (part 1): overview and cross-sectional latent class and latent profile analyses. J Pediatr Psychol 2014 Mar;39(2):174-187. [doi: 10.1093/jpepsy/jst084] [Medline: 24277769]
35. Vermunt JK. The Vuong-Lo-Mendell-Rubin test for latent class and latent profile analysis: a note on the different implementations in Mplus and LatentGOLD. Methodology 2024;20(1):72-83. [doi: 10.5964/meth.12467]
36. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equ Modeling 2007;14(4):535-569. [doi: 10.1080/10705510701575396]
37. Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. J Classif 1996 Sep;13:195-212. [doi: 10.1007/BF01246098]
38. Lo Y, Mendell NR, Rubin DB. Testing the number of components in a normal mixture. Biometrika 2001 Oct 1;88(3):767-778. [doi: 10.1093/biomet/88.3.767]
39. Nylund-Gibson K, Choi AY. Ten frequently asked questions about latent class analysis. Transl Issues Psychol Sci 2018;4(4):440-461. [doi: 10.1037/tps0000176]
40. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. Prev Sci 2013 Apr;14(2):157-168. [doi: 10.1007/s11121-011-0201-1] [Medline: 21318625]
41. Muthén B. Statistical and substantive checking in growth mixture modeling: comment on Bauer and Curran (2003). Psychol Methods 2003;8(3):369-377. [doi: 10.1037/1082-989X.8.3.369] [Medline: 14596497]
42. Sinha P, Calfee CS, Delucchi KL. Practitioner's guide to latent class analysis: methodological considerations and common pitfalls. Crit Care Med 2021;49(1):e63-e79. [doi: 10.1097/CCM.0000000000004710] [Medline: 33165028]
43. Bolck A, Croon M, Hagenaars J. Estimating latent structure models with categorical variables: one-step versus three-step estimators. Polit Anal 2004;12(1):3-27. [doi: 10.1093/pan/mph001]
44. Vermunt JK. Latent class modeling with covariates: two improved three-step approaches. Polit Anal 2010;18(4):450-469. [doi: 10.1093/pan/mpq025]
45. Asparouhov T, Muthén B. Auxiliary variables in mixture modeling: three-step approaches using Mplus. Struct Equ Modeling 2014;21(3):329-341. [doi: 10.1080/10705511.2014.915181]

XSL•FO
RenderX

46. Bell BA, Onwuegbuzie AJ, Ferron JM, Jiao QG, Hibbard ST, Kromrey JD. Use of design effects and sample weights in complex health survey data: a review of published articles using data from 3 commonly used adolescent health surveys. Am J Public Health 2012 Jul;102(7):1399-1405. [doi: 10.2105/AJPH.2011.300398] [Medline: 22676502]
47. Pfeffermann D. The use of sampling weights for survey data analysis. Stat Methods Med Res 1996 Sep;5(3):239-261. [doi: 10.1177/096228029600500303] [Medline: 8931195]
48. Elena-Bucea A, Cruz-Jesus F, Oliveira T, Coelho PS. Assessing the role of age, education, gender and income on the digital divide: evidence for the European Union. Inf Syst Front 2021;23:1007-1021. [doi: 10.1007/s10796-020-10012-9]
49. Hwang M, Kim G, Lee S, Park YH. Digital health literacy and associated factors among older adults living alone in South Korea: a cross-sectional study. Res Community Public Health Nurs 2024;35(4):389-400. [doi: 10.12799/rcphn.2024.00766]
50. Heponiemi T, Kaihlanen AM, Kouvonen A, Leemann L, Taipale S, Gluschkoff K. The role of age and digital competence on the use of online health and social care services: a cross-sectional population-based survey. Digit Health 2022;8:20552076221074485. [doi: 10.1177/20552076221074485] [Medline: 35111333]
51. Gruben M, Sheil A, Das S. "It's like not being able to read and write": narrowing the digital divide for older adults and leveraging the role of digital educators in ireland. In: Proceedings of the Nineteenth International Conference on Tangible, Embedded, and Embodied Interaction: Association for Computing Machinery (ACM); 2025:1-15. [doi: 10.1145/3689050.3704945]
52. Spante M, Hashemi SS, Lundin M, Algers A. Digital competence and digital literacy in higher education research: systematic review of concept use. Cogent Educ 2018;5(1):1519143. [doi: 10.1080/2331186X.2018.1519143]
53. Perifanou M, Economides AA. The digital competence actions framework. In: ICERI2019 Proceedings: 12th International Conference of Education, Research and Innovation: International Academy of Technology, Education and Development (IATED); 2019:11109-11116. [doi: 10.21125/iceri.2019.2743]
54. Frishammar J, Essén A, Bergström F, Ekman T. Digital health platforms for the elderly? Key adoption and usage barriers and ways to address them. Technol Forecast Soc Change 2023 Apr;189:122319. [doi: 10.1016/j.techfore.2023.122319]
55. Lau KM, Parikh M, Harvey DJ, Huang CJ, Farias ST. Early cognitively-based functional limitations predict loss of independence in instrumental activities of daily living in older adults. J Int Neuropsychol Soc 2015 Oct;21(9):688-698. [doi: 10.1017/S1355617715000818] [Medline: 26391766]
56. Gong B, Shang S, Wu C. Association between cognitive declines and disability in activities of daily living in older adults with COPD: evidence from the China health and retirement longitudinal study. BMJ Open 2020 Oct;10(10):e040098. [doi: 10.1136/bmjopen-2020-040098] [Medline: 33115903]
57. Gedvilaitė-Kordušienė M, Rapolienė G. Digital engagement of older adults in Lithuania: obstacles and support opportunities. Soc Teor Emp Polit Prakt 2023;27(27):38-57. [doi: 10.15388/STEPP.2023.27.3]
58. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
59. Horrell LN, Hayes S, Herbert LB, et al. Telemedicine use and health-related concerns of patients with chronic conditions during COVID-19: survey of members of online health communities. J Med Internet Res 2021;23(2):e23795. [doi: 10.2196/23795] [Medline: 33539307]
60. Lam K, Lu AD, Shi Y, Covinsky KE. Assessing telemedicine unreadiness among older adults in the United States during the COVID-19 pandemic. JAMA Intern Med 2020 Oct 1;180(10):1389-1391. [doi: 10.1001/jamainternmed.2020.2671] [Medline: 32744593]
61. Heponiemi T, Kainiemi E, Virtanen L, et al. Predicting internet use and digital competence among older adults using performance tests of visual, physical, and cognitive functioning: longitudinal population-based study. J Med Internet Res 2023;25:e42287. [doi: 10.2196/42287] [Medline: 37145836]
62. Lu X, Yao Y, Jin Y. Digital exclusion and functional dependence in older people: findings from five longitudinal cohort studies. eClinicalMedicine 2022;54. [doi: 10.1016/j.eclinm.2022.101708] [Medline: 36353265]
63. Kananen L, Enroth L, Raitanen J, et al. Self-rated health in individuals with and without disease is associated with multiple biomarkers representing multiple biological domains. Sci Rep 2021;11(1). [doi: 10.1038/s41598-021-85668-7] [Medline: 33731775]
64. McGinley C, Myerson J, Briscoe G, Carroll S. Towards an age-friendly design lens. J Popul Ageing 2022;15:541-556. [doi: 10.1007/s12062-022-09367-5]
65. Martínez-Alcalá CI, Rosales-Lagarde A, Alonso-Lavernia MDLÁ, et al. Digital inclusion in older adults: a comparison between face-to-face and blended digital literacy workshops. Front ICT 2018;5:21. [doi: 10.3389/fict.2018.00021]
66. Kim M, Oh J, Kim B. Experience of digital music services and digital self-efficacy among older adults: enjoyment and anxiety as mediators. Technol Soc 2021 Nov;67:101773. [doi: 10.1016/j.techsoc.2021.101773]
67. Pywell J, Vijaykumar S, Dodd A, Coventry L. Barriers to older adults' uptake of mobile-based mental health interventions. Digit Health 2020;6. [doi: 10.1177/2055207620905422] [Medline: 32110429]
68. Nedeljko M, Bogataj D, Perović BT, Kaučič BM. Smart and age-friendly communities: a review of research, policy and literature. IFAC-Pap 2023;56(2):9546-9551. [doi: 10.1016/j.ifacol.2023.10.255]
69. Chua CSW, Lim WM, Teh PL. Designing age-friendly mobile apps: insights from a mobility app study. Act Adapt Aging 2024;48(3):409-429. [doi: 10.1080/01924788.2023.2213028]

XSL•FO
RenderX

70. Choi S, Chlebek CJ. Exploring mHealth design opportunities for blind and visually impaired older users. Mhealth 2024;10. [doi: 10.21037/mhealth-23-65] [Medline: 38689610]
71. Amouzadeh E, Dianat I, Faradmal J, Babamiri M. Optimizing mobile app design for older adults: systematic review of age-friendly design. Aging Clin Exp Res 2025;37(1). [doi: 10.1007/s40520-025-03157-7] [Medline: 40804492]

Abbreviations

aBIC: sample size–adjusted Bayesian information criterion
AIC: Akaike information criterion
BCH: Bolck-Croon-Hagenaars
BIC: Bayesian information criterion
BLRT: bootstrapped likelihood ratio test
K-IADL: Korean Instrumental Activities of Daily Living
K-MMSE-2: Korean-Mini Mental State Examination, 2nd Edition
LCA: latent class analysis
LMR: Lo-Mendell-Rubin
NSOK: National Survey of Older Koreans
OR: odds ratio
R3STEP: 3-step approach with covariates
SNS: social networking service

Edited by PH Chau; submitted 19.Jan.2026; peer-reviewed by P Worthy, TD Grace; revised version received 04.Oct.2026; accepted 22.Oct.2026; published 24.Nov.2026.

Please cite as:
Lee S, Yang M, Cho J, Kim J, Cho E
Latent Class Analysis for Identifying Patterns of Digital Device Use and Adaptation Difficulty and Their Associations With Service Needs Among Older Adults: Secondary Data Analysis of a Cross-Sectional Survey Study
JMIR Aging 2026;9:e91770
URL: https://aging.jmir.org/2026/1/e91770
doi:10.2196/91770

© SangA Lee, Minhee Yang, Jungwon Cho, Jiyoung Kim, Eunhee Cho. Originally published in JMIR Aging (https://aging.jmir.org), 24.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Intergenerational Digital Learning and Relational Outcomes in Junior-Senior Dyads: Exploratory Dyadic Intervention Study

May O Lwin, PhD; Valerie Yu, PhD; Gelean Veronica Madrid Evangelista, MA; Peter J Schulz, PhD

Asian Centre for Health Behavioural Insights & Interventions (HABITS), Wee Kim Wee School of Communication & Information, Nanyang Technological University, 31 Nanyang Link, Singapore, Singapore

Corresponding Author:
May O Lwin, PhD
Asian Centre for Health Behavioural Insights & Interventions (HABITS), Wee Kim Wee School of Communication & Information, Nanyang Technological University, 31 Nanyang Link, Singapore, Singapore

Abstract

This study sought to examine whether and how an intergenerational learning mentorship program could be linked to individual and relational benefits. Outcomes studied through the actor-partner interdependence model (APIM) included social connectedness across generations, socioemotional gains, digital confidence, and well-being among junior partners and senior partners.

(JMIR Aging 2026;9:e92774) doi:10.2196/92774

KEYWORDS

digital literacy; digital inclusion; intergenerational interaction; intergenerational communication; socioemotional well-being; dyadic analysis

Introduction

Age-related digital divides increasingly affect older adults' access to information and services (eg, health resources), with implications for their autonomy and psychological well-being [1]. Previous research has shown that digital literacy interventions have been associated with psychosocial adaptation for older adults, including not only functional competency but also confidence development, empowerment, and the provision of new avenues for social support [2-4]. Intergenerational digital learning, where younger individuals support older adults in acquiring everyday digital skills, may therefore support not only skill acquisition but also relational connectedness through a socially embedded learning process.

Intergenerational learning occurs through relational exchanges, where outcomes are shaped by individuals' own experiences as well as their partner's perceptions and behaviors. Emotional engagement and reciprocal interaction between junior partners and senior partners may therefore also influence connectedness and relational experiences across dyads in addition to targeted learning outcomes. Hence, we used the actor-partner interdependence model (APIM) [5] as a useful framework for examining both individual (actor) and cross-partner influences within junior-senior dyads. This research contributes to dyadic research on intergenerational communication by examining the association between participation in a structured digital learning program and psychosocial and relational outcomes within junior-senior dyads in the Asian context, where societies typically hold reverence for older adults.

Methods

We conducted an exploratory intervention study examining the individual and dyadic effects of an intergenerational digital learning program in which junior partners mentored senior partners in developing digital literacy skills. We recruited 42 participants (21 junior partners and 21 senior partners), who were either socially close (ie, family and friends) or socially distant (ie, strangers). In total, 16 socially close and 5 socially distant pairs were recruited through convenience sampling. Given the exploratory nature and limited sample size of this study, dyads were analyzed collectively.

Junior partners were aged 12 to 24 years, resided in Singapore, and were familiar with the digital applications used in the digital literacy intervention, while senior partners were aged 60 to 85 years, also resided in Singapore, and owned their own smartphones. The intervention consisted of two sessions covering communication and social media (eg, making audio and video calls, posting content) and privacy management (eg, managing security configurations, blocking suspicious calls), and a pre- and postsurvey administered before and after the intervention. Through the APIM [5], we assessed how individual outcomes could influence junior partners' and senior partners' personal well-being, as well as their dyadic partner's relational experiences.

Paired-samples *t* tests were conducted individually for junior partners and senior partners using SPSS software (version 29.0.20; IBM Corp) to examine pre-post changes. APIMs were estimated using the *lavaan* package in R (version 0.6 - 20; R Foundation for Statistical Computing) to evaluate cross-partner

XSL•FO
RenderX

effects. The following items were measured pre- and postintervention: digital literacy (ω_y=.83; ω_s=.84); media use (ω_y=.87; ω_s=.56); eHealth literacy [6] (ω_y=.78; ω_s=.90); community involvement (ω_y=.89; ω_s=.81); intergroup empathy (single-item measure); intergenerational bonding (α_y=.55; α_s=.79; ω was not computed due to the 2-item structure); intergenerational relationship (ω_y=.80; ω_s=.88); social attraction as 2 distinct factors, a factor of positive social attraction (ω_y=.82; ω_s=.77) and a factor of negative social attraction (ω_y=.87; ω_s=.84); and knowledge exchange (ω_y=.76, ω_s=.91).

Ethical Considerations

Informed consent was obtained from all participants, and the study was approved by Nanyang Technological University's ethics board (IRB-2022 - 072).

Results

Paired-sample analyses indicated an increase across several digital and socioemotional outcomes. Senior partners demonstrated gains in digital literacy, eHealth literacy, and positive social attraction, while junior partners showed favorable changes in media use, relationship quality, and knowledge exchange (Table 1).

At the dyadic level, partner effects were generally limited, with the clearest effects observed for intergenerational bonding, where junior partners' bonding perceptions were positively related to senior partners' relational experiences (Table 2).

Table . Pre-post comparisons for senior partners and junior partners (paired *t* tests, including marginal results) in 21 dyads (42 participants).

Construct	Group	Mpre[a]	Mpost[b]	*t* test (20)	*P* value	Cohen *d*	Interpretation
Digital literacy	Senior	3.08	3.42	−2.74	.01	0.60	↑[c] moderate
eHEALS[d] literacy	Senior	3.42	3.71	−2.65	.02	0.58	↑ moderate
Social attraction (positive)	Senior	3.67	3.90	−2.23	.04	0.49	↑ small–moderate
Intergenerational empathy	Senior	3.29	3.48	−1.75	.09	0.35	↑ trend[e]
Media use	Junior	3.05	3.27	−5.59	<.001	1.22	↑ large
Relationship quality	Junior	3.80	3.85	−2.09	.049	0.46	↑ small–moderate
Knowledge exchange	Junior	3.96	4.00	−2.91	.009	0.64	↑ moderate
Digital literacy	Junior	4.26	4.23	1.91	.07	0.42	↓[f] trend

[a]Mpre: mean score preintervention.

[b]Mpost: mean score postintervention.

[c]Indicates improvement.

[d]eHEALS: eHealth Literacy Scale.

[e]*P*<.10.

[f]Indicates reduction.

XSL•FO
RenderX

Table . Summary of actor-partner effects across actor-partner independence models.

Predictor	Actor (junior)	Partner (senior→junior)	Actor (senior)	Partner (junior→senior)	Interpretation
Digital literacy					Weak, noncoupled trends
Unstandardized coefficient	0.009	−0.263	−0.276	0.190	
P value	.95	.23	.3	.25	
Bonding					Junior partners' bonding positively relates to senior partners' closeness
Unstandardized coefficient	0.115	−0.363	−0.729	0.216	
P value	.71	.13	.001	.02	
Community involvement					Marginal junior self-effect
Unstandardized coefficient	0.255	−0.283	0.063	0.02	
P value	.1	.17	.78	.9	

Discussion

Principal Findings

Participants in both age groups appeared to receive positive benefits from the program in complementary ways. Senior partners primarily demonstrated gains in digital competencies, while junior partners reported an increase in relational outcomes and knowledge exchange. These findings are consistent with previous research showing that digital skills can support senior well-being [7] and that sustained intergenerational interaction promotes bonding [8].

At the dyadic level, the APIM revealed limited but meaningful partner effects, primarily from junior partners to senior partners. Beyond functional skill acquisition (eg, digital literacy), intergenerational learning may also function as a psychosocial process involving emotional reciprocity and social connectedness between generations. The observed partner effects suggest that the quality of relational engagements could contribute to the learning experiences within junior-senior dyads. Similarly, a study found that the enthusiasm and commitment of a highly skilled mentor can influence the outcomes of intergenerational learning programs [9,10]. In aging societies where access to communication, health resources, and everyday services is increasingly digital, relational approaches may be particularly useful in helping senior partners to bridge this digital divide.

Considering the nature of our study and associated limitations related to the sample size, intervention duration, and measures with weaker reliability coefficients, these results should be considered exploratory. Future studies can (1) recruit larger pairs to improve statistical power, (2) consider a longitudinal structure to observe digital literacy, well-being, and relational satisfaction over time, (3) include follow-up assessments to determine relational outcomes postintervention, (4) examine how cultural contexts and communication styles may influence intergenerational learning, and (5) understand how preexisting relationship closeness impacts learning outcomes and partner effects.

Conclusions

Overall, our findings suggest that intergenerational digital learning is promising in supporting digital inclusion, psychological well-being, and social connectedness across generations. By adopting a dyadic perspective, this study suggests that reciprocal intergenerational interactions could be a positive contributing factor for encouraging digital adaptation among older adults.

Generative AI Use

The authors used ChatGPT (OpenAI) for proofreading, text editing, and language polishing.

Funding

This study is made possible through the Singapore Ministry of Education Academic Research Fund Tier 1 grant (RT04-21) from March 2022 to February 2026. The funder was not involved in any conceptualization of the study design, the analysis of data, or the writing of the paper.

XSL•FO
RenderX

Data Availability

Data are available upon request by contacting the corresponding author, subject to the university's data management guidelines.

Authors' Contributions

MOL provided overall supervision, leadership, conceptualization, mentoring, guidance, writing, and presentation of the published work. PJS provided conceptualization, mentorship on data analysis using relevant analytical methods, writing, and presentation of the published work. VY provided conceptualization and writing of the original draft, data analysis, processing of data, data cleaning, and responding to comments and feedback. GVME provided conceptualization, writing of the original draft, responding to comments, feedback, and implementing revisions during the subsequent stages leading to paper acceptance and publication.

Conflicts of Interest

None declared.

References

1. Seifert A, Hofer M, Rössel J. Older adults' perceived sense of social exclusion from the digital world. Educ Gerontol 2018 Dec 2;44(12):775-785. [doi: 10.1080/03601277.2019.1574415]
2. Carrasco-Dajer CM, Vera-Calzaretta AR, Ubillos-Landa S, Oyanedel JC, Díaz-Gorriti V. Impact of a culturally adapted digital literacy intervention on older people and its relationship with health literacy, quality of life, and well-being. Front Psychol 2024;15:1305569. [doi: 10.3389/fpsyg.2024.1305569] [Medline: 38686085]
3. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018 May 8;58(3):467-477. [doi: 10.1093/geront/gnw249] [Medline: 28201730]
4. Niu HJ, Li MH, Hsieh FY, Yu CC, Lin CT. From digital anxiety to empowerment in older adults: cross-sectional survey study on psychosocial drivers of digital literacy. JMIR Aging 2026 Jan 21;9:e75245. [doi: 10.2196/75245] [Medline: 41564350]
5. Kenny DA, Kashy DA, Cook WL. Dyadic Data Analysis, 2nd edition: Guilford Publications; 2020.
6. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
7. Liesa-Orús M, Latorre-Cosculluela C, Sierra-Sánchez V, Vázquez-Toledo S. Links between ease of use, perceived usefulness and attitudes towards technology in older people in university: a structural equation modelling approach. Educ Inf Technol 2023 Mar;28(3):2419-2436. [doi: 10.1007/s10639-022-11292-1]
8. So KM, Shek DTL. Elder lifelong learning, intergenerational solidarity and positive youth development: the case of Hong Kong. Int J Adolesc Med Health 2011;23(2):85-92. [doi: 10.1515/ijamh.2011.016] [Medline: 21870669]
9. Jarrott SE, Stremmel AJ, Naar JJ. Practice that transforms intergenerational programs: a model of theory- and evidence informed principles. J Intergener Relatsh 2019 Oct 2;17(4):488-504. [doi: 10.1080/15350770.2019.1579154]
10. Webster M, Norwood K, Waterworth J, Leavey G. Effectiveness of intergenerational exchange programs between adolescents and older adults: a systematic review. J Intergenerational Relatsh 2024 Oct;22(4):603-644. [doi: 10.1080/15350770.2023.2267532]

Abbreviations

APIM: actor-partner independence model

Edited by Y Jiang; submitted 03.Feb.2026; peer-reviewed by C Tziraki, HJ Niu; comments to author 20.May.2026; revised version received 19.Aug.2026; accepted 26.Aug.2026; published 19.Nov.2026.

Please cite as:
Lwin MO, Yu V, Evangelista GVM, Schulz PJ
Intergenerational Digital Learning and Relational Outcomes in Junior-Senior Dyads: Exploratory Dyadic Intervention Study
JMIR Aging 2026;9:e92774
URL: https://aging.jmir.org/2026/1/e92774
doi:10.2196/92774

© May O Lwin, Valerie Yu, Gelean Veronica Madrid Evangelista, Peter J Schulz. Originally published in JMIR Aging (https://aging.jmir.org), 19.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction

XSL•FO
RenderX

in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Gaps Between Self-Reported and Documented Use of a Health Care Service Portal Among Older Adults: Cross-Sectional Survey and Retrospective Electronic Data Study

Efrat Neter[1], PhD; Diane Levin-Zamir[2,3], MPH, PhD; Ester Eliyahu[2], MPH; Peter W S Chang[4], MD, MPH, ScD; Tuyen V Duong[5], PhD; Orna Baron-Epel[2], PhD

[1]Behavioral Sciences Department, Ruppin Academic Center, 3 Bait, Emeq Hefer, Israel

[2]School of Public Health, Faculty of Social Welfare and Health Sciences, University of Haifa, Haifa, Israel

[3]Department of Health Education and Promotion, Clalit Health Services, Tel-Aviv, Israel

[4]Chang Bing Show Chwan Memorial Hospital, Lukang, Taiwan

[5]School of Nutrition and Health Sciences, Taipei Medical University, Taipei, Taiwan

Corresponding Author:
Efrat Neter, PhD
Behavioral Sciences Department, Ruppin Academic Center, 3 Bait, Emeq Hefer, Israel

Abstract

Background: Health care systems increasingly offer services to citizens through online portals, which have been found to be associated with patient satisfaction and better management of health concerns. Few studies have compared self-reported use with objectively documented portal use, especially among older adults.

Objective: This study aimed to examine differences between documented and self-reported health portal use, and their associations with background variables and digital health literacy (DHL) among older adults.

Methods: A cross-sectional telephone survey was conducted among 600 participants aged 60 years and older, who were sampled from a registry of members of a national health maintenance organization (HMO). The participants reported their use of the HMO's portal, their DHL, and demographic and background variables. Participants' use of the portal was retrospectively retrieved from the HMO's administrative database.

Results: A large gap between self-reported use and documented use of the health portal was found: for example, while only 8.1% (46/566) reported not using the portal, the HMO's records indicated that 62% (372/600) did not make use of the portal. However, a positive correlation was found between self-reported use and documented use (r=0.22; P<.001). Significant positive associations were found between administrative and physician-related online services, both reported use (r=0.41; P<.001) and documented use (r=0.56; P<.001). A multivariate analysis, where documented use was the dependent variable, explained 12.5% of the variance. Experience with computers (β=0.13; P=.02) and ethnicity (β=−0.22; P<.001) were significantly associated with documented use. DHL, though significantly associated with documented use at the bivariate level (r=0.09; P=.04), was not significantly associated with documented use in this analysis (β=0.02; P=.97).

Conclusions: The findings indicate that users tend to overestimate their health portal use. Interpretations vary from measurement issues, such as social desirability bias regarding digital functioning or the perception that assisted use also constitutes usage, to signaling their identity as engaged digital health consumers. Future studies on digital health use should also include measures of electronically documented use and not only self-reported use.

(JMIR Aging 2026;9:e93057) doi:10.2196/93057

KEYWORDS

digital health literacy; ethnicity; portal use; online portal; older adults

Introduction

Digital transformation in health care, and its potential for improving patient outcomes, operational efficiency, and care quality, is fundamentally contingent on users' active engagement and their digital competencies. Consistent evidence demonstrates that the success of digital health technologies critically depends on users' willingness to participate in and effectively use these systems [1,2]. Individuals need sufficient digital health literacy (DHL) to access, understand, appraise, and apply health information from electronic sources to realize the benefits of digital health solutions [3]. The evidence indicates that without adequate user engagement and digital skill development, digital health technologies often produce limited clinical impact [4].

XSL•FO
RenderX

A patient health portal is a digital health application, defined as a secure, web-based platform through which patients can access and manage their personal health services, including, for example, communication with health care providers (telehealth), medical records, laboratory results, appointment scheduling, prescription requests, and referrals. The Danish national health portal, sundhed.dk, exemplifies the use of these platforms for understanding laboratory results and treatment plans, with the authors reporting on actual patient use [5]. Systematic reviews from recent years on the use of health portals and their outcomes [6-8] surveyed the different metrics, reported on positive associations with health processes (monitoring, interactions with health care providers, and health behaviors), and noted that patient portals are increasingly integral to health systems, offering functionalities that support patient engagement. Still, the quality of evidence on the effects of service use (eg, health care workers' workload, emergency department use, and hospitalization) and clinical outcomes was rated as low to moderate.

When discussing patient portal use, it is important to distinguish between self-reported use and documented use, and the current study focuses on the gap between the two. Reported use refers to data gathered directly from patients through surveys, interviews, or questionnaires, where they self-report their portal activities, perceptions, attitudes, or intentions [9]. Numerous studies have focused on reported use of portals in conjunction with self-reported eHealth literacy [10,11]. Documented use, on the other hand, refers to objective data tracked directly from patient portal systems, such as login frequency, session duration, specific functionalities accessed (eg, viewing test results and sending messages), digital health outcomes (health behaviors and clinical outcomes), and health care use patterns (interactions with health care providers) [12].

Few studies have reported on both documented and reported use. For example, Woods et al [10] examined patient factors associated with documented use among patients (n=270) who registered with a portal and self-identified as users and reported the percentage of documented users among those reporting use. It should be noted that studies in a related field of DHL, which assessed both reported and actual literacy levels, found a low to medium association between the two [11-13]. A comparison of self-reported use to documented use may serve as a form of validation of self-reported use. The question is, to what degree does self-reported use of portal services represent actual or documented use, as measured by computerized measures of use?

A related research effort that focused on the factors associated with adoption and engagement by patients with their portals categorized these factors into demographic (age, gender, and race/ethnicity), socioeconomic (education, income, insurance, and geography), health-related (chronic conditions/health status and patient activation), and technology-related characteristics (health literacy, digital access/skills, and DHL). There is considerable heterogeneity in the findings, and a systematic review [7], which analyzed 32 studies on this subject, identified only women's gender and age (older, though depending on the population studied) as consistently associated with portal use. Another umbrella review [6] did not find strong evidence of an association between patients' characteristics and portal use, and assessed the quality of the evidence (regarding private insurance, higher illness burden, race [White], age, and income) as moderate.

Studies that examined both self-reported and documented use of health portals are scarce and are often based on small samples, presenting a research gap. The objective of the present study was to address this gap by examining both types of use in a medium-sized sample, assessing the association between the two, and examining patient characteristics associated with documented and self-reported use. We focused on the population of older adults (aged >60 years), who represent a growing demographic that often faces challenges in managing complex and multiple chronic conditions, leading to a high demand for health care services. Due to their increased health care needs and frequent interactions with the health care system, they are also considered among the populations that may benefit significantly from patient portals for improved access to care, self-management, communication, and health information [14]. In addition to demographic and socioeconomic characteristics, we also included the often-studied skill of DHL, which has been found to be associated with engagement and use [15].

We hypothesized the following:

- H1: Reported use of the health portal will be higher than documented use, and the two use measures will be positively associated with each other to a weak-to-medium degree.
- H2: Patient characteristics will be associated with both reported and documented use. Based on the cited reviews, we hypothesize that, in the restricted age group of those older than 60 years, age will be negatively associated with use measures, while women will have higher documented and reported use than men.
- H3: DHL will be positively associated with both reported and documented use.

In addition, we explored the following research question (RQ): Are there associations between portal-documented use and socioeconomic status, self-rated health, and prior use of the internet and computers?

Methods

Sample

The eligibility criteria for the study included being a member of "Clalit" Health Services (CHS) in Israel, being aged 60 years or older, and being fluent in Hebrew, Arabic, or Russian.

CHS is Israel's largest health maintenance organization (HMO), serving more than half of the country's population. It is a nonprofit organization, sanctioned by Israel's National Health Insurance Law of 1994, operating a comprehensive network of community and hospital services throughout the country, and, over the past several decades, has maintained electronic health records that support clinical care, quality improvement, and both clinical and population-based research [16,17]. CHS has a digital portal that includes a large, diverse range of services and can be accessed from any type of digital device (mobile

XSL•FO
RenderX

phone, personal computer, laptop, or tablet). Access to the portal—for both patients and health care providers—is availed by using the patient's personal identification number.

A random list of potential participants (names and phone contact details) was generated by CHS, assigned a serial number, anonymized, and securely transferred to a survey company. The target number of 600 (actually, N=601) was reached after phone contact was initiated with 5688 individuals. The overall response rate was 20.4%, with a sampling error of 2.04%, discounting phone numbers with no response after at least 3 attempts. The phone interviews were conducted in Hebrew, Arabic, or Russian, according to the preference of the interviewee. There was no requirement for prior experience using the portal, as all CHS members are entitled to access the digital resources.

Design and Procedure

The telephone survey's design was cross-sectional, while the data on documented portal use were retrospective for the year 2024. The outcome (dependent) variables were the documented and reported use of the health portal. The independent variables were demographics, DHL, and use of digital resources and devices. The survey was conducted from June to September 2024. Following the telephone survey, information on the documented use of the CHS portal was extracted from the national CHS information system on "no-queue" services, matched to each individual who responded to the survey, and anonymized for analysis.

Measures

Documented Portal Use

CHS collects information on the use of its website/mobile application through which digital health services are provided. Service use is recorded upon completion of each service transaction. Portal use pertains to information regarding the types of services. We grouped services (with equal weights) into medical/health, administrative, and overall services. The medical/health services include medical authorizations/certificates (eg, sick leave) and clinical summaries; results of tests/procedures and referrals; and requests for prescriptions (eg, refills of routinely taken medications). The administrative services include consultations with clinic administrators/service staff, authorizations for funding of medical procedures implemented outside the CHS system, and payment refunds.

Reported Portal Use

Respondents were asked about the level of use of the above-mentioned services on the CHS portal during the past 6 months (Multimedia Appendix 1). The list of services was read out to the respondents, who estimated the number of times they accessed the portal. The 6-month time frame was chosen as this is a time frame widely used in health care use, found to produce more accurate reporting (than 12 mo) and lower the chances of recall bias [18]. Responses were recorded on a 5-point scale (not at all, less than once in this time period, 1‐2 times, 4‐6 times, and more than once in a given month). We included the response option "less than once in this time period" to convey the possibility that respondents have used the activity sometime in the past, but not in the time period asked about, as distinct from "not at all." Three indices of reported portal use were constructed by averaging across service items: reported medical/health-related use (Cronbach internal reliability was α=0.71), reported administrative use (Cronbach internal reliability was α=0.59), and overall reported use.

DHL

DHL was measured using the HLS_{19}-DIGI instrument [19] of the World Health Organization Action Network on Measuring Population and Organizational Health Literacy, which comprises 8 items assessing the difficulty or ease of the following skills: accessing, understanding, appraising, and applying online health information. The internal reliability (Cronbach α) was 0.89 in the current study.

As a correlate of DHL, an additional module (HLS_{19}-DIGI-DD), comprising 6 items, assessed the frequency of using digital health resources. Participants reported the frequency (ranging from 1="Not relevant or less than once per week" to 5="More than once per day") of using the following resources: websites, social media, digital devices, mobile health apps, digital communication with health services, and "other digital resources."

Demographic Characteristics

Participants provided information on gender, age, country of birth, ethnicity/religion (Jewish, Muslim, Christian, Druze, and other; grouped such that Muslims, Christians, and Druze were classified as Arab, and Jews and others were classified as Jewish), religiosity, education (grouped into less than high school, completed high school, and completed more than high school), perceived social status (the MacArthur Scale [20]), difficulty covering the cost of medications, employment status, marital status, household size, and self-rated health [21].

Ethical Considerations

This study was approved by the institutional ethics review boards of both the University of Haifa (#298/25) and CHS (#0101‐23-COM2). All respondents expressed explicit consent before their participation, and their consent was orally recorded; they were notified that any publication of the study would contain no identifying information. All data were anonymized for data analysis purposes. No compensation was offered or given to participants.

Statistical Analysis

Participants' characteristics and the study's primary measures (reported and documented portal use) were calculated for descriptive analysis. Categorical variables were analyzed using frequency and percentage distributions, while continuous variables were examined using means and SDs.

Indices of reported and documented use were computed by averaging the medical/health and administrative items. Bivariate associations between reported use and documented use of the health portal were computed, as well as the associations of these with explanatory (independent) variables (demographics and DHL). Bivariate associations were also computed between all

XSL•FO
RenderX

subindices (medical/health, administration, and overall) of reported and documented use.

For comparison between reported and documented use of services, both indices were aligned to qualitative ranges: none, low, and high. In the case of documented use, the transformation was none to none, ≥1 and ≤2 to low, and >2 to high. In reported use, the transformation was none to none, "less than once in this time period" and "1 to 2 times" to low, and "4 to 6 times" and "more than once in a given month" to high.

Lastly, we conducted linear hierarchical regressions to assess the association between documented and reported use and explanatory variables. Only variables that were significant in the bivariate analysis with either documented or reported use were entered. As years of computer use and using computers at work were highly correlated (r=0.50; P<.001), only the former was entered. The variables entered in the first step included the following: age, ethnicity, education, household size, self-rated health, and experience with computers. DHL was entered in the second step. We also reversed the order of the steps. Ethnicity was coded so that Jewish was 0 and Arab was 1. R^2 and coefficients were reported.

Results

Participant Characteristics

Of 600 study participants, 314 (52.3%) were women and 286 (47.7%) were men, with a mean age of 70.45 (SD 7.54) years, ranging in age from 60 to 80 years. Over half (n=351, 58.5%) were Jewish and were married or cohabiting (n=378, 63%). Over half (n=365, 61%) completed more than a high school education, ranked themselves in the upper half of the social ladder (6 and above on a 10-point McArthur scale: n=449, 74.8%), and a quarter of the sample (n=150, 25%) found paying for medications difficult or very difficult. Nearly a third (n=179, 29.8%) rated their health as very good and good, 41.5% (n=249) as average, and nearly one-third (n=172, 28.6%) as bad or not so good. Most people lived either with another person in their household (n=258, 44%), or lived in larger households (n=198, 33%) and the rest, about a quarter (n=144, 24%), lived alone. Lastly, 65% (n=114) of the sample reported regularly using/used computers in their work, with an average of 18.35 (SD 15.53) years of experience using computers, and 58.3% (n=349) endorsed being very good at surfing the internet (see Table 1 for more details).

XSL•FO
RenderX

Table . Participant characteristics (N=600) and the association with documented and reported portal use.

Characteristics	Values	Association with documented portal use[a]	Association with reported portal use[a]
Gender, n (%)		t_{598}=0.09[b]; *P*=.93	t_{568}=0.61; *P*=.54
Women	314 (52.3)		
Men	286 (47.7)		
Age (y), mean (SD)	70.45 (7.54)	*r*=−0.08; *P*=.03[c]	*r*=0.05; *P*=.21
Ethnicity, n (%)		$F_{4,595}$=7.13[d], *P*<.001[c]	$F_{4,561}$=6.45[d]; *P*<.001[c]
Jewish	351 (58.5)		
Muslim	164 (27.3)		
Christian	33 (5.7)		
Druze	13 (2.2)		
Other	38 (6.3)		
Marital status, n (%)		$F_{2,597}$=1.32[d]; *P*=.27	$F_{2,563}$=0.21[d]; *P*=.81
Single	29 (4.8)		
Married/cohabiting	378 (63)		
Divorced/widowed	193 (32.2)		
Education, n (%)		0.14, *P*<.001[c]	0.12, *P*=.003[c]
None, primary, or secondary	265 (44.2)		
Postsecondary	102 (17.0)		
BA degree	137 (22.8)		
MA degree and higher	96 (16.0)		
Paying for medications, n (%)		0.03; *P*=.54	–0.03; *P*=.51
Very easy	196 (32.7)		
Pretty easy	254 (42.3)		
Pretty difficult	88 (14.7)		
Very difficult	62 (10.3)		
Self-rated health		0.004; *P*=.92	−0.11; *P*=.01
Very bad	47 (7.8)		
Not so good	125 (20.8)		
Average	249 (41.5)		
Good	107 (17.8)		
Very good	72 (12)		
Household size		*r*=−0.08[d], *P*=.04[c]	*r*=−0.03[d], *P*=.57
1	144 (24)		
2	258 (44)		
3 and more	198 (33)		
Using/used computers at work, n (%)		$F_{1,172}$=5.85[d]; *P*=.02[c]	*F*=2.99[d]; *P*=.09
Yes	114 (65)		
No	60 (34.5)		
Missing	426 (71.0)		

XSL•FO
RenderX

Characteristics	Values	Association with documented portal use[a]	Association with reported portal use[a]
Years of using computers, mean (SD)	18.35 (15.53)	r=0.14; P<.001[c]	r=0.22; P<.001[c]

[a]Statistics reported are Pearson *r*, unless otherwise stated.

[b]Statistics is *t* test.

[c]Significant associations.

[d]Statistics is *F*; in the case of ethnicity and marital status, the 2 large groups were compared.

Documented and Reported Portal Use

Documented and reported use are presented for each service (Multimedia Appendix 2) and for the averaged indices: medical/health, administrative, and overall.

The overall reported and documented indices were aligned for comparative purposes, so that both reported and documented use included 3 categories: none, low, and high. The frequency scale of the averaged documented use of the types of services on the portal (measured as the number of times, continuous) was transformed into categories: none, low (≥1 and ≤2), and high (>2). Categories on the scale of reported use were transformed into none, low (including "less than once in this time period" and "1 - 2 times"), and high (the 2 upper response options). The documented and reported use of the portal is presented in Table 2.

Table . Reported and documented use of the HMO portal[a].

Percentage of portal use by services	Reported[a], n (%)	Documented, n (%)
None	46 (8.1)	372 (62.0)
Low	123 (21.7)	223 (37.2)
High	397 (70.1)	5 (0.8)

[a]There was missing data on the reported use for 34 respondents.

The data indicate that reported use is higher than documented use. The data from the HMO portal indicate that the majority of the participants (n=372, 62.0%) did not use the portal at all in the past year, whereas only 8.1% (n=46) reported not using it in the past 6 months. An opposite pattern was observed for use defined as low: documented use of the portal under this category was about a third (n=223, 37.2%), but only a fifth of participants reported using the portal at this use level (n=123, 21.7%). Most participants reported using the portal at high levels (n=397, 70.1%), whereas documented high use was reported for only a few participants (n=5, 0.8%). The data on specific types of services used (Multimedia Appendix 2) also consistently showed higher reported use than documented use.

The association between documented and reported use of the portal was positive and significant (r=0.22; P<.001), such that higher documented use was positively associated with higher reported use (significant associations are given in Table 3). This association was also found at the service-use level, both for medical/health-related services between documented and reported use (r=0.16; P<.001) and for administrative services between documented and reported use (r=0.19; P<.001), such that greater engagement in one activity was associated with greater use of the other activity. Table 3 also indicates a significant positive association between the 2 types of services in the portal, medical/health-related and administration-related, in both the reported (r=0.41; P<.001) and documented use (r=0.56; P<.001).

XSL•FO
RenderX

Table . Associations (Pearson *r* and 2-tailed *P* value) between use indices (reported and documented) and digital health literacy (DHL).

Associations	1	2	3	4	5	6	7	8
1. Reported medical/health services	—[a]							
2. Reported administrative services	0.41; *P*<.001	—						
3. Reported services index	0.88; *P*<.001	0.80; *P*<.001	—					
4. Documented medical/health services	0.16; *P*<.001	0.27; *P*<.001	0.24; *P*<.001	—				
5. Documented administrative services	0.10; *P*=.02	0.19; *P*<.001	0.16; *P*<.001	0.56; *P*<.001	—			
6. Documented services index	0.14; *P*<.001	0.25; *P*<.001	0.22; *P*<.001	0.83; *P*<.001	0.91; *P*<.001	—		
7. Digital health literacy	0.03; *P*=.54	0.20; *P*<.001	0.11; *P*=.01	0.17; *P*<.001	0.09; *P*=.04	0.17; *P*<.001	—	
8. Using digital resources	0.11; *P*=.02	–0.01[b]	0.06[b]	0.02[b]	0.05[b]	0.04[b]	0.04[b]	—

[a]Not applicable.

[b]All *P* values >.05.

Bivariate Associations With Portal Use

The association between documented portal use and background variables is displayed in Table 1 in the third column (with significant associations/differences mentioned in Table 1). Pearson coefficients were computed for continuous variables, and a 2-tailed *t* test or ANOVA was used for categorical variables. Age (r=−0.08; P=.03), higher education (r=0.14; P<.001), and ethnicity (being Jewish, $F_{4,595}$=7.13; P<.001) were significantly associated with documented portal use, as well as household size (r=−0.08; P=.04) and using or having used computers at work ($F_{1,172}$=5.85; P=.02). Documented portal use was not significantly associated with gender, marital status, being able to cover the expense of medications, and self-rated health (all P values >.05).

A somewhat similar pattern emerged between reported portal use and background variables (last column in Table 1): higher education (r=0.12; P=.003), ethnicity (being Jewish, $F_{4,561}$=6.45; P<.001), and years of experience with computers (r=0.22; P<.001) were significantly associated with reported portal use; unlike documented use, self-rated health was associated with reported use, with participants reporting worse health reporting more portal use (r=−0.11; P=.01).

We also examined all portal use indices in relation to DHL (both the cognitive capacities and the use of digital devices; Table 3). Documented portal use indices (administrative and medical-related indices, and the overall index) were significantly associated with DHL, but not with the use of various digital devices (all P values >.05), except for the association between digital device use and medical-related services in the portal (r=0.11; P=.02).

Multivariate Associations Explaining Portal Use (Reported and Documented)

We regressed background and DHL variables on reported and documented portal use across services. We included only variables that were significantly associated with combined services—either reported or documented—at the bivariate level. Only one variable, “using or having used computers at work,” was excluded from the analysis due to high shared variance with the related question of “experience with computers” (r=0.50; P<.001). Table 4 presents the results of the 2 regressions.

XSL•FO
RenderX

Table . Hierarchical linear regression on documented and reported portal use.

Characteristics	*B*	SE(*B*)	β	*P* value
Outcome: documented portal use[a]				
Age	0	0	−0.09	.04
Ethnicity	−0.17	0.04	−0.22	<.001
Education	0.01	0.01	0.05	.37
Household size	0.01	0.01	0.05	.36
Self-rated health	0	0.02	0.01	.86
Experience with computers	0	0	0.13	.02
DHL[b]	0.01	0.01	0.02	.97
Outcome: reported portal use[c]				
Age	0	0	0.04	.43
Ethnicity	−0.1	0.1	−0.06	.32
Education	0	0.03	0.01	.83
Household size	0	0.02	−0.01	.79
Self-rated health	-0.12	0.04	-.15	<.001
Experience with computers	0.01	0	.23	<.001
DHL	0	0.002	.04	.46

[a]R^2 for the documented portal use outcome is 12.5%.

[b]DHL: digital health literacy.

[c]R^2 for the reported portal use outcome is 7.1%.

As shown in Table 3, the documented use of the portal was explained better (adjusted R^2=12.5%) than reported use (adjusted R^2=7.1%), with documented use significantly explained by the ethnicity of patients (with higher use among patients who identified as Jewish than among those who identified as Arab) and reported experience with computers (the more experience, the more use), while reported use was significantly explained by experience with computers and self-rated health, with education, ethnicity, and DHL not significantly associated with reported use. Notably, DHL, while significantly associated at the bivariate level, did not reach significance at the multivariate level in both analyses, possibly due to shared variance with other explanatory variables (eg, r=0.46; P<.001, r=0.49; P<.001, and r=0.31; P<.001, for education, experience with computers, and self-rated health, respectively).

Discussion

Principal Results

This study focused on portal use—documented and reported—among a relatively vulnerable, growing in numbers, and highly relevant population segment—older adults. The design of the study—relying on both sources of information—allowed the researchers to identify and quantify a significant gap between patients' perceptions of their activity and their documented behavior. The study found a significant discrepancy between patients' self-reported use of services on their health portal and the documented use of these services by the HMO, supporting H1 that reported use will be higher than documented use. While only 8.1% (n=46) of participants reported not using the portal at all, the retrospective analysis of the CHS database indicated that 62% (n=372) did not use the services on the portal at all. This substantial gap underscores a potential overestimation of use of services when relying solely on self-reported data, suggesting that both indicators should be used in studies recording and explaining portal use.

Despite this disparity, a positive correlation was found between self-reported use and documented use of the portal (r=0.22; P<.001). This positive association suggests that although not compatible, each measure represents a different aspect of use of services. When examining factors associated with portal use in multivariate analyses, the hypotheses on age and gender (H2: negative association with age and more use among women) were not supported (though the age component of H2 was supported in a bivariate analysis). Key predictors of documented portal use included ethnicity (higher use among participants who identified as Jewish than among those who identified as Arab) and reported experience with computers (more experience correlated with greater use, addressing the RQ). In contrast, reported use was significantly explained by self-rated health (increased use among participants reporting worse health) and by reported experience with computers. Additionally, the model for documented use demonstrated better explanatory power (adjusted R^2=12.5%) compared to reported use (adjusted R^2=7.1%). DHL was associated with most use indices in a

XSL•FO
RenderX

bivariate analysis, but was not significantly associated with either type of use in multivariate analyses, thus lending mixed support for H3. Lastly, both medical/health-related and administrative services within the portal were positively correlated in both reported (r=0.41; P<.001) and documented use (r=0.56; P<.001).

Comparison With Prior Work and Interpretation

The substantial gap observed between self-reported and documented portal use in this study is a major finding, echoing concerns highlighted in previous literature regarding the accuracy of self-reported use of digital health services. Studies have consistently shown that self-reported data can often reflect an overestimation of portal use. For instance, Beal et al [8] indicated that measurements of patient portal use in research often vary widely, with inconsistent definitions and a reliance on simplified metrics, such as "at least one login," which may not reflect meaningful engagement or use of services. Tavares and Oliveira [22] also noted that awareness and registration do not necessarily translate to frequent use, finding that only 30% of users regularly checked their electronic health records portal. The current study's finding of 62% (n=372) documented nonuse compared with 8.1% (n=46) reported nonuse attests to the limitations of self-report and the need to augment self-reports with objective, tracked data to better understand patient engagement. The disparity between perceived and recorded measures is analogous to the low concordance between self-reported and externally measured DHL/eHealth literacy [23] as well as between self-reported and documented health care use for periods exceeding a month [23]. Another relevant comparison is the assessment of stress, which overwhelmingly encompasses both self-reported stress and physiological measures. These two are also only moderately correlated and are considered to capture different, distinct aspects of the stress response, each considered valid [24]. Other examples are smoking or physical activity, assessed by both self-reports and objective measures (accelerometers and the presence of cotinine [a metabolite of nicotine], respectively).

As there was no incentive for high reported use, interpretations can vary, ranging from measurement explanations, cognitive-motivational explanations, and broader identity signaling or symbolic engagement explanations. Measurement explanations address the possibilities that participants perceived surrogate/assisted use (clinic staff or family members) as their own use; that respondents may have experienced recall bias or estimation bias, so that they genuinely believe they use the portal more frequently than they actually do; or that an acquiescence bias occurred, whereby some respondents tended to answer "yes" to an interviewer. Conversely, database records may be incomplete because of the definition of a completed transaction or errors in linking activity to individual participants. Therefore, electronic records should not necessarily be regarded as an error-free gold standard, and thus the gap could be interpreted as measurement disagreement, as neither self-report nor database documentation necessarily captures precisely the same construct.

A more social-psychological interpretation pertains to constructs such as self-efficacy and self-perception. Self-efficacy theory could suggest that individuals with high digital health self-efficacy may report portal use partly based on confidence, perceived competence, or familiarity, rather than documented frequency [25]. Relatedly, Bem's Self-Perception Theory proposes that people infer attitudes and identities from observations of themselves, constructing narratives about who they are. Specifically, respondents may infer that portal use is consistent with a self-image of caring about their health, being engaged with their health, and using modern technology. Under this interpretation, over-reporting is identity-consistent self-description [22]. The technology acceptance model could also explain the findings through the intention-behavior gap: respondents may have reported intended use, expected use, or favorable attitudes, instead of documented behavior. Valuing and intending to use the portal may have become using the portal [26]; respondents understood something other than what was intended.

The reported-documented use gap may be broadly interpreted through identity signaling theory or digital participation/symbolic technology use research. Identity signaling theory proposes that behaviors communicate desired identities to others and even to oneself [27]. Digital health technologies increasingly function as symbols of responsibility, proactivity, health consciousness, and technological competence. Thus, respondents may have reported portal use in a manner consistent with their identity as an engaged, responsible, and digitally connected person, even when objective use remains low. While identity signaling has not been studied in patient portals specifically, it was used to explain other digital health behaviors, such as health-app adoption and wearable use [28].

The perspective of digital participation views portal use as part of a broader digital inclusion and citizenship [29]. Under this perspective, the central question becomes: "Does an individual feel they participate in digital health care?" rather than "How many times have they logged in?" Hence, portal use may function as a marker of digital inclusion, and respondents who see themselves as digitally connected may report higher use because they either identify as digitally engaged, feel connected to modern health care, or find that portal membership itself becomes meaningful. Thus, reporting use reflects perceived participation in digital health care ecosystems rather than observed activity.

In conclusion, on the interpretations of the reported-documented use gap, rather than viewing self-reported portal use solely as a proxy for behavior, "use" may be more appropriately understood as a multidimensional construct that reflects both documented engagement and participants' self-perceptions, identity signaling, and symbolic affiliation with digital health care, thereby explaining why subjective reports and objective usage metrics do not necessarily converge.

The identified demographic disparities in documented portal use align with, and also nuance, existing research. The finding that older age was associated with less documented portal use in the bivariate analysis is largely consistent with numerous studies reporting that older adults generally trail the general population in adopting digital health technologies. However, some literature presents mixed findings on age, with Carini et

XSL•FO
RenderX

al [7] noting both increasing use with age and a lack of homogeneity in results, while Woods et al [10] found no significant differences in portal logins by age among already enrolled users. The current study, focusing on an older adult population (mean age 70.45 y), records the challenge faced by this demographic group in portal use.

Regarding ethnicity, our finding of higher documented use among Jewish participants compared with Arab participants supports the broader understanding that patient portal use is often lower among racial and ethnic minorities. Previous research suggests that White patients are generally more likely to use patient portals and that patients' race and ethnic backgrounds are associated with perceptions of need for, and access to, portal services [2,6,30]. However, it is also important to note that some national studies [10,31], after adjusting for other factors, have not identified significant associations between patients' race and ethnic backgrounds and patient portal access or use. The significant association between ethnicity and *documented* portal use observed in our multivariate model (as opposed to no significant association of ethnicity with *reported* use) may be due to different degrees of acquiescence bias in reporting among the groups [32,33] or to differences in actual use associated with reduced internet infrastructure and resources available in rural areas, inhabited more by Arabs.

The significant association of reported experience with computers for both self-reported and documented portal use is well-supported in prior studies, either reviews or empirical work. Limited digital literacy, computer skills, and internet access are frequently cited as significant barriers to patient portal adoption [7,10,13,14]. This consistency across studies on prior experience with computers underscores the fundamental importance of previously developed digital skills for effective portal engagement.

Our findings on DHL—association with reported and documented use only at the bivariate level and not in a multivariate analysis—align with the mixed results in the literature. For example, van der Vaart et al [13], who examined documented portal use, found it associated with eHealth literacy, as well as other works [34,35] which identified eHealth literacy as positively associated with *reported* portal use. However, Xie concluded in her review [36] that the evidence is inconclusive, as other studies point to no association [37,38]. Still, this result could stem from the shared variance with other related variables in the multivariate analysis.

The difference in explanatory power between models for documented vs reported use (R^2=12.3% vs R^2=5.3%) could reflect a measurement issue or a substantive difference. Measurement-wise, documented portal use is an objective measure with relatively low noise, whereas self-reported use is subjective and likely contains greater measurement error. These differences in measurement properties alone can influence R^2 values, independent of any true discrepancy. Alongside, this result may indicate that objective measures capture a different angle of user behavior and its drivers. It should be noted that studies using the technology acceptance model and its extensions [22,34] explained higher variance than that found in our study, though they addressed reported usage and not documented usage. Still, most variation in portal engagement remains unexplained.

Strengths and Limitations

This study demonstrates methodological strengths. A primary strength is its robust dual-data approach, which directly addresses a paucity in the literature by comparing both self-reported portal use and objective, documented use data retrieved directly from the HMO database. Furthermore, the study strategically focused on a vulnerable, growing, and highly relevant population segment—older adults—and used a medium-sized sample of 600 individuals aged 60 years and older. This demographic is critical to study because older adults often need to manage complex chronic conditions, have high health care needs, and stand to benefit significantly from portal functionalities, yet often face unique challenges with technology adoption. Finally, the study moves beyond basic demographics by incorporating specific digital variables, such as DHL and reported experience with computers, into its analysis.

Several limitations should be considered when interpreting these findings. First, the nonidentical temporal frames and scales used for documented (frequency during the year of 2024) and reported use (addressed past 6 months, ranging from "not at all" to "more than once a month") should be considered. The latter time frame (6 months) was intended to reduce recall bias, yet it may have made direct comparisons difficult and introduced measurement bias. However, we did not compare exact quantities but rather qualitative ranges, and when comparing reported and documented use we focused on the category of no use, which was identical in both measures. Relatedly, though the time frame was constructed to be shorter to facilitate the accuracy of reported use (6 months rather than a year), some degree of recall error or forgetfulness may be attributed to age-related memory challenges among people aged 60 years and older. Second, while the study directly addressed the gap between self-reported and documented use, it relied on self-reported data for some variables, such as computer experience and DHL. Additionally, some participants might have had clinic staff or family members use the portal on their behalf (surrogate/assisted use), which could influence their self-perception of "use" without actual personal operation. Notably, this interpretation of surrogate use would emphasize the gap between reported and documented use even more. Third, the cross-sectional design of the survey limits the ability to establish causality or observe changes in portal use patterns and their explanatory variables over time. Longitudinal studies are essential for understanding the evolution from initial awareness to sustained adoption of patient portals. Fourth, the study was conducted within a single HMO in Israel, which—though serving more than half the country's population—may limit the generalizability of findings to other health care systems or populations with different cultural, socioeconomic, or technological contexts. Fifth, the response rate of 20.4% may be considered relatively low. Though this response rate is not uncommon in a population sample with no paid incentive [39], it may have introduced selection bias, which in turn may have biased estimates. Furthermore, the study analyzed the population aged 60 years and older as one homogeneous sample with regard to age. Further studies could

XSL•FO
RenderX

contribute more in-depth analyses of stratified age groups among the older population.

Conclusions

This study reveals a substantial gap between patients' perceptions of their health portal use and their documented use, with self-reported use significantly higher than documented use. This highlights a critical challenge for health care systems and reinforces the necessity of also using objective metrics to comprehensively assess patient engagement. The identification of background variables associated with documented portal use underscores the persistent digital divide, which affects vulnerable populations, including older adults and certain ethnic minority groups.

Bridging this divide and enhancing patient engagement calls for integrating interventions in health care that are culturally sensitive, tailored to specific demographic needs and digital experience levels, including DHL. By addressing these factors, health care systems can move toward more equitable and effective patient portal use, ultimately leveraging these tools to improve health outcomes for all.

Acknowledgments

We would like to acknowledge the contribution of Talish Razi from "Clalit" Community Medical Branch for her support regarding the data used in the study.

The first author (EN) used Microsoft Copilot (Microsoft Corp) during manuscript preparation to improve grammar, clarity, and readability. All AI-generated suggestions were reviewed, edited, and verified by the author, who takes full responsibility for the final content of the manuscript.

Funding

The study was funded by a grant from the Israeli Ministry of Science and Technology for a joint bi-national study with Taiwan.

Data Availability

Due to Clalit Health Services' data privacy regulations and per the institutional Helsinki and data utilization committee approvals for this study, the data used for this study cannot be shared.

Authors' Contributions

EN, DL-Z, PWSC, TVD, and OB-E conceived and planned the study. EE coordinated data collection, supported by DL-Z. EN conducted the data analysis and wrote the first draft. All authors commented on the drafts.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Use of portal (reported and documented).
[DOCX File, 19 KB - aging_v9i1e93057_app1.docx]

Multimedia Appendix 2
Frequency of portal-documented services (N=600).
[DOCX File, 15 KB - aging_v9i1e93057_app2.docx]

References

1. Mitzner TL, Savla J, Boot WR, et al. Technology adoption by older adults: findings from the PRISM trial. Gerontologist 2019 Jan 9;59(1):34-44. [doi: 10.1093/geront/gny113] [Medline: 30265294]
2. van Dijk J, van Deursen A. Impact: why digital skills are the key to the information society. In: Digital Skills: Palgrave Macmillan; 2014:43-62. [doi: 10.1057/9781137437037_3]
3. Western MJ, Smit ES, Gültzow T, et al. Bridging the digital health divide: a narrative review of the causes, implications, and solutions for digital health inequalities. Health Psychol Behav Med 2025;13(1):2493139. [doi: 10.1080/21642850.2025.2493139] [Medline: 40276490]
4. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR mHealth uHealth 2018 Jan 17;6(1):e23. [doi: 10.2196/mhealth.8873] [Medline: 29343463]
5. Nielsen NB, Sekkal CK, Yoganathan S. Online investigations on optimizing the Danish health portal sundhed.dk. Stud Health Technol Inform 2021 Nov 8;286:89-93. [doi: 10.3233/SHTI210644] [Medline: 34755696]

XSL•FO
RenderX

6. Antonio MG, Petrovskaya O, Lau F. The state of evidence in patient portals: umbrella review. J Med Internet Res 2020 Nov 11;22(11):e23851. [doi: 10.2196/23851] [Medline: 33174851]
7. Carini E, Villani L, Pezzullo AM, et al. The impact of digital patient portals on health outcomes, system efficiency, and patient attitudes: updated systematic literature review. J Med Internet Res 2021 Sep 8;23(9):e26189. [doi: 10.2196/26189] [Medline: 34494966]
8. Beal LL, Kolman JM, Jones SL, Khleif A, Menser T. Quantifying patient portal use: systematic review of utilization metrics. J Med Internet Res 2021 Feb 25;23(2):e23493. [doi: 10.2196/23493] [Medline: 33629962]
9. Shahrabani S, Mizrachi Y. Factors affecting compliance with use of online healthcare services among adults in Israel. Isr J Health Policy Res 2016;5(1):15. [doi: 10.1186/s13584-016-0073-8] [Medline: 27307983]
10. Woods SS, Forsberg CW, Schwartz EC, et al. The association of patient factors, digital access, and online behavior on sustained patient portal use: a prospective cohort of enrolled users. J Med Internet Res 2017 Oct 17;19(10):e345. [doi: 10.2196/jmir.7895] [Medline: 29042345]
11. Kim S, Park K, Jo HS. Gap between perceived eHealth literacy and ability to use online cancer-related information. J Korean Med Sci 2020;35(24):e187. [doi: 10.3346/jkms.2020.35.e187] [Medline: 32567257]
12. Neter E, Brainin E. Perceived and performed eHealth literacy: survey and simulated performance test. JMIR Hum Factors 2017 Jan 17;4(1):e2. [doi: 10.2196/humanfactors.6523] [Medline: 28096068]
13. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011 Nov 9;13(4):e86. [doi: 10.2196/jmir.1840] [Medline: 22071338]
14. Sakaguchi-Tang DK, Bosold AL, Choi YK, Turner AM. Patient portal use and experience among older adults: systematic review. JMIR Med Inform 2017 Oct 16;5(4):e38. [doi: 10.2196/medinform.8092] [Medline: 29038093]
15. Cheng C, Gearon E, Hawkins M, et al. Digital health literacy as a predictor of awareness, engagement, and use of a national web-based personal health record: population-based survey study. J Med Internet Res 2022 Sep 16;24(9):e35772. [doi: 10.2196/35772] [Medline: 36112404]
16. Balicer RD, Shadmi E, Lieberman N, et al. Reducing health disparities: strategy planning and implementation in Israel's largest health care organization. Health Serv Res 2011 Aug;46(4):1281-1299. [doi: 10.1111/j.1475-6773.2011.01247.x] [Medline: 21352224]
17. Balicer RD, Afek A. Digital health nation: Israel's global big data innovation hub. Lancet 2017 Jun 24;389(10088):2451-2453. [doi: 10.1016/S0140-6736(17)30876-0] [Medline: 28495105]
18. Marshall-Aiyelawo K, Bannick R, Beamer S, et al. Effect of change in the CG CAHPS survey instrument recall period on patient experience scores on healthcare utilization. Patient Exp J 2019;6(3):114-123. [doi: 10.35680/2372-0247.1359]
19. Levin-Zamir D, Van de Broucke S, Bíró É, et al. Measuring digital health literacy and its associations with determinants and health outcomes in 13 countries. Front Public Health 2025 Mar 20;13:1472706. [doi: 10.3389/fpubh.2025.1472706]
20. Adler NE, Epel ES, Castellazzo G, Ickovics JR. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy White women. Health Psychol 2000;19(6):586-592. [doi: 10.1037//0278-6133.19.6.586] [Medline: 11129362]
21. Benyamini Y. Why does self-rated health predict mortality? An update on current knowledge and a research agenda for psychologists. Psychol Health 2011 Nov;26(11):1407-1413. [doi: 10.1080/08870446.2011.621703] [Medline: 22111660]
22. Tavares J, Oliveira T. Electronic health record patient portal adoption by health care consumers: an acceptance model and survey. J Med Internet Res 2016 Mar 2;18(3):e49. [doi: 10.2196/jmir.5069] [Medline: 26935646]
23. Short ME, Goetzel RZ, Pei X, et al. How accurate are self-reports? Analysis of self-reported health care utilization and absence when compared with administrative data. J Occup Environ Med 2009 Jul;51(7):786-796. [doi: 10.1097/JOM.0b013e3181a86671] [Medline: 19528832]
24. Vaessen T, Rintala A, Otsabryk N, et al. The association between self-reported stress and cardiovascular measures in daily life: a systematic review. PLoS One 2021;16(11):e0259557. [doi: 10.1371/journal.pone.0259557] [Medline: 34797835]
25. Park J, Liang M, Alpert JM, Brown RF, Zhong X. The causal relationship between portal usage and self-efficacious health information-seeking behaviors: secondary analysis of the Health Information National Trends Survey data. J Med Internet Res 2021 Jan 27;23(1):e17782. [doi: 10.2196/17782] [Medline: 33502334]
26. Philpot LM, Ramar P, Roellinger DL, Njeru JW, Ebbert JO. Individual-level digital determinants of health and technology acceptance of patient portals: cross-sectional assessment. JMIR Form Res 2024 Jun 10;8:e56493. [doi: 10.2196/56493] [Medline: 38695754]
27. Guraya SS, Ennab F, Guraya SY. When the self "logs in"—a critical narrative review of digital identity in health professions education. Front Med (Lausanne) 2025;12:1715752. [doi: 10.3389/fmed.2025.1715752] [Medline: 41503353]
28. Burford K, Golaszewski NM, Bartholomew J. "I shy away from them because they are very identifiable": a qualitative study exploring user and non-user's perceptions of wearable activity trackers. Digit Health 2021;7:20552076211054922. [doi: 10.1177/20552076211054922] [Medline: 34820134]
29. Sieck CJ, Sheon A, Ancker JS, Castek J, Callahan B, Siefer A. Digital inclusion as a social determinant of health. NPJ Digit Med 2021 Mar 17;4:52. [doi: 10.1038/s41746-021-00413-8] [Medline: 33731887]

XSL•FO
RenderX

30. Clarke MA, Lyden ER, Ma J, et al. Sociodemographic differences and factors affecting patient portal utilization. J Racial Ethn Health Disparities 2021 Aug;8(4):879-891. [doi: 10.1007/s40615-020-00846-z] [Medline: 32839896]
31. El-Toukhy S, Méndez A, Collins S, Pérez-Stable EJ. Barriers to patient portal access and use: evidence from the Health Information National Trends Survey. J Am Board Fam Med 2020;33(6):953-968. [doi: 10.3122/jabfm.2020.06.190402] [Medline: 33219074]
32. Baron-Epel O, Friedman N, Lernau O. Validity of self-reported mammography in a multicultural population in Israel. Prev Med 2008 Jun;46(6):489-491. [doi: 10.1016/j.ypmed.2008.03.003] [Medline: 18407343]
33. Baron-Epel O, Kaplan G, Weinstein R, Green MS. Extreme and acquiescence bias in a bi-ethnic population. Eur J Public Health 2010 Oct;20(5):543-548. [doi: 10.1093/eurpub/ckq052] [Medline: 20439322]
34. Hoogenbosch B, Postma J, de Man-van Ginkel JM, Tiemessen NA, van Delden JJ, van Os-Medendorp H. Use and the users of a patient portal: cross-sectional study. J Med Internet Res 2018 Sep 17;20(9):e262. [doi: 10.2196/jmir.9418] [Medline: 30224334]
35. Price-Haywood EG, Harden-Barrios J, Cooper LA. Comparative effectiveness of audit-feedback versus additional physician communication training to improve cancer screening for patients with limited health literacy. J Gen Intern Med 2014 Aug;29(8):1113-1121. [doi: 10.1007/s11606-014-2782-4] [Medline: 24590734]
36. Xie L, Zhang S, Xin M, Zhu M, Lu W, Mo PKH. Electronic health literacy and health-related outcomes among older adults: a systematic review. Prev Med 2022 Apr;157:106997. [doi: 10.1016/j.ypmed.2022.106997] [Medline: 35189203]
37. Conroy SS, Harcum S, Keldsen L, Bever CT. Novel use of existing technology: a preliminary study of patient portal use for telerehabilitation. J Telemed Telecare 2022 Jun;28(5):380-388. [doi: 10.1177/1357633X20950172] [Medline: 32869689]
38. Khan D, Fjerbæk A, Andreasen JJ, Thorup CB, Dinesen B. Cardiac surgery patients' e-health literacy and their use of a digital portal. Health Educ J 2018 Jun;77(4):482-494. [doi: 10.1177/0017896918756435]
39. Vallance JK, Eurich DT, Gardiner PA, Taylor LM, Stevens G, Johnson ST. Utility of telephone survey methods in population-based health studies of older adults: an example from the Alberta Older Adult Health Behavior (ALERT) study. BMC Public Health 2014 May 22;14:486. [doi: 10.1186/1471-2458-14-486] [Medline: 24884997]

Abbreviations

CHS: Clalit Health Services
DHL: digital health literacy
HMO: health maintenance organization
RQ: research question

Edited by AF Abdulai; submitted 07.Feb.2026; peer-reviewed by E Mpofu, J Park; revised version received 10.Aug.2026; accepted 22.Sep.2026; published 13.Oct.2026.

Please cite as:
Neter E, Levin-Zamir D, Eliyahu E, Chang PWS, Duong TV, Baron-Epel O
Gaps Between Self-Reported and Documented Use of a Health Care Service Portal Among Older Adults: Cross-Sectional Survey and Retrospective Electronic Data Study
JMIR Aging 2026;9:e93057
URL: https://aging.jmir.org/2026/1/e93057
doi:10.2196/93057

© Efrat Neter, Diane Levin-Zamir, Ester Eliyahu, Peter W S Chang, Tuyen V Duong, Orna Baron-Epel. Originally published in JMIR Aging (https://aging.jmir.org), 13.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Fear-Specific Technophobia as a Mediator Between Daily Technical Support and Instrumental Versus Hedonic Technology Use in Later Life: Cross-Sectional Survey Study

Yu Yang[1,2], MSc; Shuo Dai, BA; Siyu Yin, BA; Wanyu Xi[3], PhD; Qi Li[4], PhD

[1]Hohai University, Nanjing, Jiangsu, China

[2]Jiangsu Open University, Nanjing, Jiangsu, China

[3]School of Aging Services and Management, Nanjing University of Chinese Medicine, No.138 Xianlin Road, Nanjing, Jiangsu, China

[4]School of Business Administration (MBA), Key Research Institute of Humanities and Social Sciences at Universities Modern Business Research Center, Zhejiang Gongshang University, Hangzhou, Zhejiang, China

Corresponding Author:
Wanyu Xi, PhD
School of Aging Services and Management, Nanjing University of Chinese Medicine, No.138 Xianlin Road, Nanjing, Jiangsu, China

Abstract

Background: Population aging and rapid digitalization have intensified concerns about older adults' digital participation. Although access disparities have narrowed, older adults continue to differ from younger groups in how they use digital technologies, with their use skewed toward instrumental rather than hedonic purposes. Technophobia is a salient psychological barrier; yet, its distinct fear dimensions and their links to different forms of technology use remain insufficiently differentiated. Daily technical support (DTS) may provide an important sociotechnical context for understanding these relationships.

Objective: This study aimed to test how DTS promotes older adults' instrumental and hedonic technology use by reducing multidimensional technophobia (fear of inconvenience, fear of personal failure, and human-machine ambiguity) and whether these mechanisms differ by gender.

Methods: A cross-sectional survey was conducted among Chinese adults aged ≥55 years recruited via online platforms and offline communities in Nanjing (June-October 2024). A total of 212 valid responses were included (mean age 66.96, SD 6.96, range 55 - 86 years; 142/212, 68.4% women). Measures assessed day-to-day technical support, 3 dimensions of technophobia, instrumental versus hedonic technology use, and demographic variables. We used partial least squares structural equation modeling (PLS-SEM) with 5000 bootstrap resamples to test direct, indirect, and moderated mediation effects, reporting 95% bias-corrected and accelerated (BCa) CI.

Results: DTS significantly reduced fear of inconvenience (β=−0.455, 95% BCa CI −0.632 to −0.229; *P*<.001) but did not significantly affect fear of personal failure (*P*=.12) or human-machine ambiguity (*P*=.26). Fear of inconvenience and human-machine ambiguity negatively predicted instrumental use (IU; β=−0.215, 95% BCa CI −0.410 to −0.009; *P*=.05 and β=−0.211, 95% BCa CI −0.389 to −0.001; *P*=.04). DTS had a positive direct effect on IU (β=0.104, 95% BCa CI 0.009 to 0.203; *P*=.04) but not on hedonic use (HU; β=0.075; *P*=.11). The indirect effect of DTS on IU through fear of inconvenience was significant (β=0.097, 95% BCa CI 0.020-0.204). Gender moderated the effect of DTS on fear of inconvenience (β=0.287, 95% BCa CI 0.052-0.518; *P*=.02), and such conditional indirect effect was significant for men (β=0.098, 95% BCa CI 0.001-0.219) but not for women (β=0.036, 95% BCa CI −0.002 to 0.116).

Conclusions: DTS was associated with greater instrumental technology use among older adults through lower fear of inconvenience, with a stronger indirect association among men, whereas no comparable pathway emerged for HU. This suggests that gender may shape how older adults benefit psychologically from technical support rather than reflect general differences in support needs. Interventions should move beyond one-size-fits-all approaches by providing targeted, task-oriented support that addresses older adults' specific technological fears and practical barriers.

(JMIR Aging 2026;9:e93063) doi:10.2196/93063

KEYWORDS

technophobia; daily technical support; technology adoption; older adults; digital inclusion

XSL•FO
RenderX

Introduction

Background

Population aging is accelerating worldwide, with the proportion of people aged 60 years and older projected to nearly double from 12% in 2015 to 22% by 2050 [1]. Meanwhile, with the rapid development of information and communication technologies (ICTs), digital technologies have become increasingly embedded in older adults' everyday lives, mediating routine activities such as banking, shopping, health management, and social interaction and making digital competence increasingly comparable to foundational literacy [2,3]. The first-level digital divide, which refers to inequalities in access to digital technologies [4], has narrowed substantially with the widespread diffusion of internet-enabled devices and improved connectivity [5-7]. In China, for example, the number of internet users aged 60 years and older increased from 119 million in December 2021 to approximately 173 million in December 2025, an increase of approximately 45.4% over 4 years [8,9]. Nevertheless, older adults remain disproportionately disadvantaged in digital access, with an internet penetration rate of 53.7%, compared with 80.1% for the population as a whole [8]. In addition, despite the narrowing first-level digital divide, substantial inequalities persist at the second and third levels, reflected in disparities in how older adults use digital technologies and in the digital skills and competencies needed for effective and meaningful engagement [10,11]. Indeed, older adults' online engagement remains markedly skewed toward instrumental uses (IUs; eg, online services and communication) rather than hedonic uses (HUs; eg, entertainment, gaming, and social media) [11-13]. For instance, in the United Kingdom, 58% of older internet users (ages 65 - 74 years) shop online and 65% bank online, whereas only about 32% play online games [14].

Barriers to ICT adoption among older adults are multifaceted, encompassing health-related limitations, inadequate digital skills, resource constraints such as cost and connectivity, insufficient support networks, and psychological factors [15]. While many of these barriers primarily limit older adults' opportunities or capacity to engage with technology, psychological factors may exert a more immediate influence on adoption decisions by shaping how older adults perceive their technological capabilities, evaluate the benefits and demands of technology, and respond to difficulties encountered during use. Accordingly, psychological barriers—particularly technophobia—are salient for older adults. Technophobia is commonly defined as a fear or anxiety about technology, and it is associated with delayed or hesitant adoption, especially among individuals with limited digital literacy or inaccurate beliefs about technology [11,16]. Although scholars have called for conceptualizing technophobia as multidimensional, much of the literature continues to treat it as a unidimensional construct, limiting fine-grained insight into how distinct fears map onto different forms of technology use [16-19]. Moreover, while prior research has examined demographic correlates of technophobia and shown that structured interventions (eg, training and intergenerational mentoring) can alleviate it [20,21], evidence remains scarce on whether and how spontaneous day-to-day technical support—more typical of naturalistic settings—shapes technophobia within a broader sociotechnical context.

To address these gaps, this study examines how daily technical support (DTS) influences older adults' instrumental versus hedonic technology use through multiple fear-related dimensions of technophobia. We further test whether these indirect effects vary by gender, given documented gender differences in support-seeking motivations and technology-related stereotypes [22]. By clarifying which dimensions of technophobia are most responsive to everyday support—and for whom—this research advances digital divide scholarship and informs more targeted, age-friendly support and training strategies.

Literature Review and Hypotheses Development

The Multidimensionality of Technophobia

Originated from the concept of computer anxiety, technophobia was initially conceptualized as a unidimensional construct capturing negative psychological responses to technology in general [23,24]. This early view closely resembles technology anxiety or digital anxiety, which refers primarily to restlessness or apprehensive tension when one has to use or learn a new technology [25]. Subsequent research has refined this perspective, arguing that technophobia is not merely a simple emotional reaction (ie, anxiety); rather, technology anxiety constitutes only 1 facet of technophobia [17,26]. Accordingly, technophobia is better understood as an umbrella, higher-order phenomenon comprising multiple dimensions that reflect the cognitive appraisals and beliefs underlying the sources of fear during technology adoption.

Some studies have approached technophobia from an affective perspective, conceptualizing it primarily in terms of anxiety, discomfort, or other negative emotional responses to technology [27-29], with measures including techno-anxiety, techno-paranoia, and privacy concerns [26]. A different line of research has focused on the specific sources from which such fear arises. Sinkovics et al (2002) [19] proposed 3 dimensions of technophobia from the perspective of the underlying sources of fear; it distinguishes self-related, internal sources of fear from externally situated concerns embedded in the usage context. Fear of personal failure refers to the anticipated threat to one's competence and self-worth when engaging with technology—specifically, the concern that one may be unable to use it correctly, may make errors, or may be perceived as incapable. Fear of inconvenience reflects a more instrumental concern about whether the effort required to use a technology is justified by its anticipated benefits. Human-machine ambiguity captures discomfort rooted in one's preferences and beliefs about everyday action and interpersonal exchange, particularly when human contact is substituted by machine-mediated interaction. Subsequent research has further reinforced the view of technophobia as a multidimensional construct. For example, Khasawneh (2018) [17] identified 5 dimensions of technophobia, including techno-anxiety, techno-paranoia, techno-fear, techno-cybernetic revolt, and communication device avoidance. Although more differentiated, broadly capture fear sources similar to those represented in Sinkovics et al [19] 3D framework.

XSL•FO
RenderX

Prior empirical work has consistently shown that older adults tend to report higher levels of technophobia than younger cohorts [16,18]. Technophobia has also been associated with avoidance and reduced technology engagement, such that older adults may use digital technologies less frequently, limit themselves to a narrower range of applications, and show less willingness to adopt unfamiliar technologies [30]. However, much of the existing literature has treated technophobia as a global or unidimensional construct, providing limited insight into which specific sources of fear are associated with particular technology-use behaviors. By differentiating among distinct fear sources, Sinkovics et al [19] framework offers a more fine-grained basis for identifying the psychological barriers underlying older adults' technology use and for considering forms of support that are better matched to those barriers. Accordingly, the current study adopts this 3D, fear-specific framework to examine how different forms of technophobia are associated with older adults' technology use.

The Instrumental Versus HU of Technology

Scholars of information systems commonly distinguish 2 types of IT according to their primary purposes and functional orientations, utilitarian (instrumental) systems that support productivity and task accomplishment, and hedonic systems that foreground enjoyment, entertainment, and self-expression [31,32]. This typology draws on the hedonic consumption tradition that differentiates utilitarian from hedonic products and experiences [33]. It underscores that information systems are embedded not only in work or education settings for job and study-related activities, but also in domestic and leisure contexts where technology use itself can provide enjoyment and experiential value. Correspondingly, users' purposes for daily technology use tend to cluster into 2 broad motivation-driven categories: instrumental purposes (eg, banking, information search, and health management) and hedonic purposes (eg, streaming, gaming, and social media). Although some digital technologies can serve both instrumental and hedonic purposes, this dichotomy remains analytically useful because the psychological beliefs underlying technology acceptance may differ depending on the primary purpose for which a technology is used.

The Technology Acceptance Model (TAM) provides the primary theoretical foundation for understanding why individuals adopt and use technology. TAM proposes that technology acceptance is largely determined by 2 key beliefs: perceived usefulness and perceived ease of use [32,34]. These 2 beliefs reflect different motivational considerations underlying technology use. Perceived usefulness is closely associated with extrinsic motivation because users adopt a technology in anticipation of instrumental outcomes, such as improved performance, greater efficiency, or task accomplishment [35]. Perceived ease of use, in contrast, reflects the effort required to interact with a technology and is more closely related to the experiential process of technology use; by reducing effort and facilitating a more fluent interaction, ease of use can support intrinsically motivated experiences such as enjoyment and playfulness [36]. Importantly, the relative importance of these motivational mechanisms varies across technology-use contexts. In instrumental contexts, technology is primarily used as a means to accomplish an external goal. Consequently, extrinsic motivation is particularly salient, and perceived usefulness becomes a central determinant of adoption because users evaluate whether the technology helps them perform tasks more effectively or efficiently. In hedonic contexts, by contrast, technology use is more strongly driven by intrinsic rewards derived from the use experience itself, such as enjoyment, pleasure, playfulness, and immersion. Accordingly, perceived enjoyment becomes particularly important, while ease of use facilitates hedonic motivation by reducing interactional effort and allowing users to experience these intrinsic rewards more readily [32,35]. Thus, from a TAM perspective, the distinction between instrumental and hedonic technology use reflects not merely differences in technological functions, but differences in the motivational beliefs that become most salient across use contexts.

The Impact of Technophobia on Different Use of Technology Among Older Adults

Building on TAM, which suggests that different technology-use contexts may be driven by different motivational beliefs, we further draw on the multidimensional view of technophobia proposed by Sinkovics et al [19] to examine how distinct sources of technological fear may map onto extrinsic versus intrinsic motivational orientations toward technology use [37,38]. In particular, fear of inconvenience aligns with extrinsic motivation, such that perceived usefulness from the TAM represents rational skepticism about the effort-benefit balance of using technology [35]. Older adults with a heightened fear of inconvenience perceive technology as overly burdensome, thereby reducing perceived value and discouraging adoption. Conversely, fear of personal failure relates closely to perceived ease of use and self-efficacy, which echoes the intrinsic motivation in TAM, where low self-confidence leads to anxiety about technological mistakes or inadequacies [34,39]. Older adults with a heightened fear of personal failure avoid technologies that require adaptation or learning, thus reinforcing digital exclusion [6,18]. In Sinkovics et al [19], the human-machine ambiguity dimension captures deep-seated preferences and beliefs about everyday action, particularly individuals' comfort with human interaction when the "human element" is replaced by a machine. Recent research suggests that such substitution can undermine perceived autonomy, thereby threatening intrinsic motivation [40-42].

We build on the foregoing review to argue that the fear of inconvenience, which is a rational, extrinsic, effort-benefit appraisal, should most strongly depress iIU, where perceived usefulness is decisive. By contrast, fear of personal failure, as well as human-machine ambiguity, which erodes competence beliefs and undermines intrinsic engagement, should strongly curb HU, where perceived enjoyment and comfort are pivotal. This distinction motivates our domain-specific expectation that different technophobia dimensions map onto different adoption purposes rather than exerting uniform effects across all forms of technology use. Therefore, we propose the hypotheses mentioned below.

XSL•FO
RenderX

Hypothesis 1

Specific dimensions of technophobia will be negatively associated with older adults' technology use, with different fear dimensions expected to be more relevant to different types of technology use. Specifically, (1) fear of inconvenience will be negatively associated with instrumental technology use, (2) fear of personal failure will be negatively associated with hedonic technology use, and (3) human-machine ambiguity will be negatively associated with hedonic technology use.

Technical Support and Its Positive Effects

In gerontological research, social support is widely recognized as a key social determinant of healthy aging, and it generally refers to the resources and assistance provided through social relationships, which help individuals feel cared for, valued, and part of a mutually obliging network [43]. It typically distinguishes between structural and functional dimensions of support. The structural aspect refers to the existence and quantity of social relationships, such as living arrangements, network size, and frequency of social interaction [44]. The functional aspect refers to the specific supportive functions that relationships serve, and 2 primary functional forms are studied. Emotional support involves empathy, caring, love, and someone to confide in during stress. Instrumental support, which is usually tangible, refers to practical help or services, such as assistance with daily tasks or financial aid [45]. A robust body of evidence suggests that a strong social support network of any functions is associated with better outcomes in late life, including improved health and longevity [46], better mental health (eg, reduced stress, loneliness, and depression) [47,48], and greater well-being (eg, life satisfaction and overall quality of life) [49,50].

Technical support, which often provides immediate resolution and encourages continued technology use, is primarily recognized as instrumental support. Such support can be structured through formal workshops, training sessions, or digital learning programs led by instructors; it can also be informal and embedded in daily life through ad-hoc help from peers, family, or volunteers [11,51]. Previous research has consistently demonstrated that technical support is a powerful facilitator for new technology adoption among older adults because it enables the immediate resolution of technical issues, thus underscoring the general importance of assistance in overcoming initial adoption hurdles [11,52]. Therefore, we propose hypothesis 2.

Hypothesis 2

DTS will promote the instrumental and HUs of technologies among older adults.

On the other hand, it is also found that daily family support that focuses on digital tasks helps to reduce older users' overall digital anxiety [25]. A set of intervention studies demonstrate that technical support from internet training programs effectively reduced older adults' technology anxiety while improving confidence [53], and similar improvements have been observed in collaborative peer-learning contexts [54]. With this basis, we argue that ongoing DTS can positively influence all 3 dimensions of technophobia. First, practical assistance alleviates the fear of inconvenience by simplifying the learning process and demonstrating that technology can make life easier [55]. Second, frequent support helps counter the fear of personal failure—knowing that mistakes can be fixed with help boosts self-efficacy through repeated successful experiences. Older adults particularly prefer assistance from trusted family or friends, as this reduces embarrassment and enhances comfort [56]. Third, having a human intermediary eases human-technology ambiguity, transforming a solitary, intimidating process into an interactive, supportive experience [57]. Therefore, we further proposed hypothesis 3 mentioned below.

Hypothesis 3

DTS will mitigate each dimension of technophobia among older adults, including (1) fear of inconvenience, (2) fear of personal failure, and (3) fear of human-machine ambiguity.

The Mediation Effect of Technophobia

While the direct positive impact of technical support on both perception and behavioral reaction toward technologies is well documented, research also suggested that this effect is often indirect and operates through users' perceptions of the technology. As articulated in the TAM, external factors (such as training or support available in one's environment) do not directly cause someone to adopt a technology; instead, they influence key perceptual drivers of adoption, such as the technology's perceived ease of use and perceived usefulness [39,58]. For instance, workplace studies have found that when support is available, employees report high ease-of-use and usefulness beliefs about new IT systems, which then predict great adoption and usage [39]. Therefore, we further proposed the mediation effect of technophobia (hypothesis 4).

Hypothesis 4

Each dimension of technophobia will mediate the effect of DTS on both HU and IU of technologies. Specifically, (1) fear of inconvenience, (2) fear of personal failure, and (3) human-technology ambiguity will mediate the effect of DTS on both IU and HU of technologies among older adults.

The Moderation Effect of Gender

Social role theory posits that societally prescribed gender roles shape help-seeking tendencies and coping behaviors [59,60]. Across cultural contexts, masculinity norms emphasize competence, making help-seeking or admitting difficulty potentially threatening to a "competent man" identity; consequently, men are often more reluctant to seek assistance and report higher self-efficacy and lower anxiety than women [61]. Men's achievement- and task-oriented roles further heighten their focus on performance payoffs and instrumental attributes [59]. In contrast, women are frequently stereotyped as less competent in masculine-typed domains such as technology and with comparatively lower self-efficacy may place greater value on environmental and interpersonal support to reduce anxiety and maintain positive self-perceptions [62]. Supporting evidence further showed that men's technology adoption is more strongly driven by extrinsic motivation (perceived usefulness), whereas women are more responsive to perceived ease of use and internalized social norms [63,64].

XSL•FO
RenderX

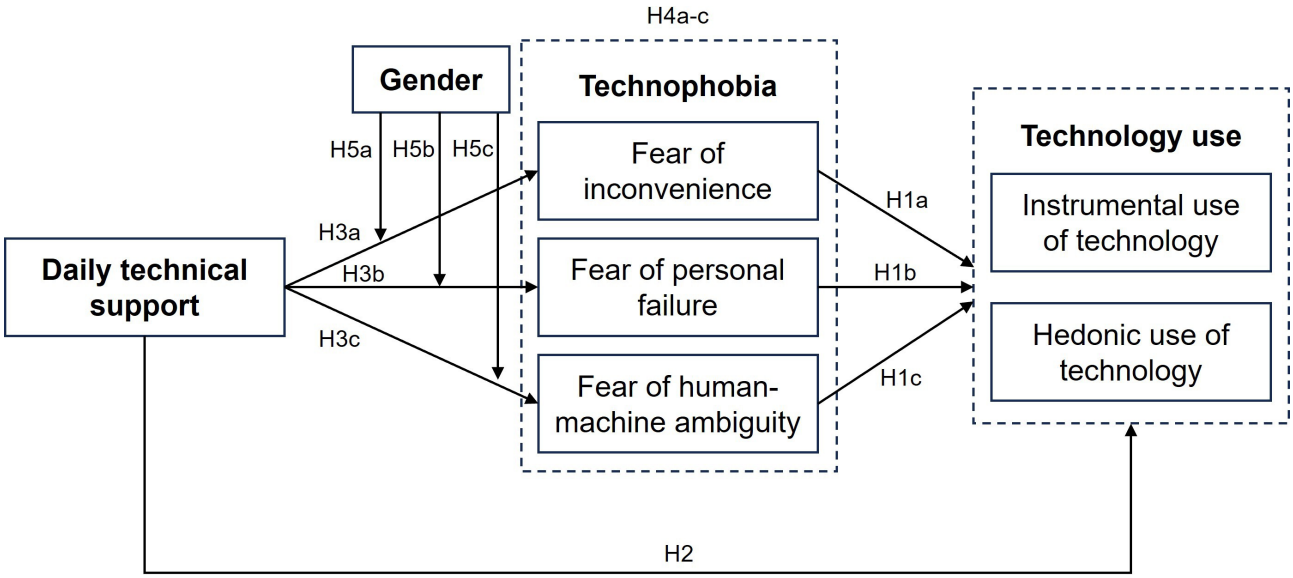
On this basis, older adults, who were often socialized under more traditional gender norms, may differ in how they respond to DTS [22]. Because older men may view seeking help from others as a threat to their competence and treat support mainly as an instrumental fix, DTS may be more likely to reduce extrinsic, cost-related fears such as fear of inconvenience, but less likely to alleviate intrinsic, self-efficacy related fears such as fear of personal failure [65]. By contrast, older women may be more inclined to rely on social support when coping with difficult tasks; such support can reduce anxiety and bolster confidence as learners [66]. Thus, DTS may more effectively address women's fear of personal failure. Finally, because fear of human-machine ambiguity reflects not only the interaction experience but also relational beliefs in everyday activities, and because women tend to be more attuned to internal feelings and social norms, technical support may more strongly reduce this fear among older women than older men. In sum, we propose hypothesis 5 as mentioned below.

Hypothesis 5

Gender moderates the effect of DTS on technophobia. Specifically, (1) technical support will likely reduce the fear of inconvenience among older men (vs women). By contrast, (2) technical support will likely reduce the fear of personal failure, and (3) the fear of human-machine ambiguity among older women (vs men).

Figure 1 shows the overall conceptual model of this research.

Figure 1. Conceptual model.



Methods

Sample and Procedures

We recruited Chinese adults aged 55 years and older through online platforms and offline communities in Nanjing, China. Inclusion criteria were (1) age of ≥55 years, (2) sufficient cognitive ability to complete the questionnaire independently or with assistance, and (3) the ability to understand the questionnaire content. The target sample size was determined based on recruitment feasibility during the predefined data collection period. To further evaluate whether the achieved sample size was adequate for the proposed moderated mediation analyses, we conducted a sensitivity analysis using G*Power 3.1 (Franz Faul). Because the hypothesized involved first-stage moderation, the analysis focused on the incremental variance explained by the interaction term in a linear multiple regression model (fixed model, R^2 increase). Each first-stage regression included 3 predictors and 9 covariates, resulting in a total of 12 predictors. With the final sample of 212 participants, α=.05, and power=.80, the study was capable of detecting an incremental interaction effect of approximately f^2=.037 or larger. Thus, the achieved sample provided sufficient sensitivity to detect relatively small moderation effects in the proposed model.

A total of 230 questionnaires were collected between June and October 2024, thus yielding 212 valid responses for analysis after data screening. No missing data were present in the final dataset, as all items were set as mandatory in the online questionnaire, while trained volunteers assisted participants with questionnaire completion during offline data collection.

Ethical Considerations

This study was approved by the Research Ethics Committee of the School of Aging Services and Management, Nanjing University of Chinese Medicine (number 20240523) and was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent before participation. No personally identifiable information was collected, thereby protecting participant privacy and confidentiality. Participants received either RMB 7 (approximately US $1) or a small gift of equivalent value as compensation for their participation.

Variables and Measures

All measures were adapted from validated scales and translated via forward-backward translation to ensure linguistic and cultural validity for Chinese older adults. Full item wording, outer loadings, reliability and validity statistics for all constructs are reported in Table 1. DTS was assessed using 3 items with

XSL•FO
RenderX

5-point ordered category and 6-point Likert-type response formats; scores from the 6-point scale were linearly rescaled to a 5-point metric for comparability, with higher scores indicating greater DTS. Instrumental and hedonic technology use were each measured using 2 items based on the frequency of use over the past week, with responses recorded on 5-point scales and higher scores indicating more frequent use. Technophobia was measured using Sinkovics et al [19] 13-item scale with a 6-point Likert-type agreement scale. Reverse-coded items were recoded prior to analysis, with higher scores indicating greater technophobia. All constructs were modeled as latent variables in the subsequent analyses. Covariates included age, marital status, residence, education, economic status, functional status, subjective health, number of children, and social activity participation, with their coding and categories reported in Table 2. Gender was coded as 0 for men and 1 for women.

XSL•FO
RenderX

Table . Assessment of construct reliability and validity.

Construct and items	L[a]	Cronbach α	ρ_a[b]	ρ_c[b]	AVE[c]
DTS[d]	—[e]	.689	0.778	0.832	0.634
DTS1: with how many peers around your age do you have regular contact about how to use digital technologies?	0.916				
DTS2: with how many younger adults (aged 50 years and younger) do you have regular contact about how to use digital technologies?	0.892				
DTS3: to what extent can you get help from these people when having difficulties in using digital technology?	0.819				
TP-TI[f]	—	.759	0.761	0.837	0.508
TP-TI1: digital devices make me feel irritated.	0.784				
TP-TI2: the ubiquity of digital devices in our daily life makes me dissatisfied.	0.771				
TP-TI3: I feel comfortable when using digital devices.	0.761				
TP-TI4: digital devices make daily activities easier.	0.709				
TP-TI5: I enjoy the convenience provided by digital devices.	0.785				
TP-PF[g]	—	.929	0.933	0.946	0.779
TP-PF1: I feel some anxiety when approaching digital devices.	0.784				
TP-PF2: I believe most people know better than I do how to use digital devices.	0.671				
TP-PF3: I feel frustrated when using digital devices.	0.661				
TP-PF4: just thinking about digital devices makes me nervous.	0.706				
TP-PF5: digital devices are frightening.	0.685				
TP-HMA[h]	—	.785	0.829	0.872	0.695

XSL•FO
RenderX

Construct and items	L[a]	Cronbach α	ρ_a[b]	ρ_c[b]	AVE[c]
TP-HMA1: I prefer to have people help me with daily activities rather than using digital devices.	0.878				
TP-HMA2: I feel more confident interacting with people than using digital services.	0.791				
TP-HMA3: our daily activities should not be solved by machines.	0.829				
IU[i]	—	.836	0.890	0.900	0.751
IU1: I used digital technologies for daily tasks (eg, information seeking and online payment).	0.916				
IU2: I used digital technologies for communication (eg, private messages, phone calls, and online communities).	0.901				
HU[j]	—	.668	0.706	0.856	0.748
HU1: I used digital technologies for individual leisure activities (eg, online games and short videos).	0.828				
HU2: I used digital technologies for group leisure activities (eg, online multiplayer games).	0.900				

[a]L: outer loadings.

[b]ρa and ρc: composite reliability.

[c]AVE: average variance extracted.

[d]DTS: daily technical support.

[e]Not available.

[f]TP-TI: technophobia-technology inconvenience. Items 3, 4, and 5 were reverse-coded.

[g]TP-PF: technophobia-personal failure.

[h]TP-HMA: technophobia-human-machine ambiguity.

[i]IU: instrumental use.

[j]HU: hedonic use.

XSL•FO
RenderX

Table . Description of participants (N=212).

Characteristics	Values, n (%)
Age (years)	
55 - 59	24 (11.3)
60 - 64	64 (30.2)
65 - 69	48 (22.6)
70 - 74	42 (19.8)
75+	34 (16.0)
Gender	
Men	67 (31.6)
Women	145 (68.4)
Marital status	
Married and/or living together with a partner	197 (92.9)
Divorced or separated	15 (7.1)
Residence	
City	203 (95.8)
Countryside or village	9 (4.2)
Economic status	
Very poor	0 (0.0)
Poor	21 (9.9)
Fair	127 (59.9)
Good	58 (27.4)
Very good	6 (2.8)
Number of children	
0	15 (7.1)
1	115 (54.2)
2	62 (29.2)
≥3	20 (9.4)
Physical function	
Greatly limits daily activities	5 (2.4)
Sometimes limits daily activities	19 (9.0)
Minimally limits daily activities	93 (43.9)
No health condition or health condition causes no limitation	95 (44.8)
Subjective health	
Excellent	20 (9.4)
Very good	42 (19.8)
Good	66 (31.1)
Fair	77 (36.3)
Poor	6 (2.8)
Education level	
Less than 12 years of education	126 (59.4)
12 years of education	31 (14.6)
Professional or vocational education	11 (5.2)
Higher education (university or college)	44 (20.8)

XSL•FO
RenderX

Characteristics	Values, n (%)
Social activity participation	
Yes	146 (68.9)
No	66 (31.1)
Device ownership[a]	
Smartphone	187 (88.2)
Smart TV	48 (22.6)
Smart watch or wearable device	26 (12.3)
Personal digital assistant	20 (9.4)
Laptop	4 (1.9)
Tablet	15 (7.1)
Other devices	16 (7.5)
Number of same-age technical supporters	
0	27 (12.7)
1 - 2	27 (12.7)
3 - 5	62 (29.2)
6 - 10	41 (19.3)
>10	55 (25.9)
Number of younger-adult technical supporters	
0	31 (14.6)
1 - 2	37 (17.5)
3 - 5	40 (18.9)
6 - 10	48 (22.6)
>10	56 (26.4)

[a]Percentages do not sum to 100% because participants could select more than one device type.

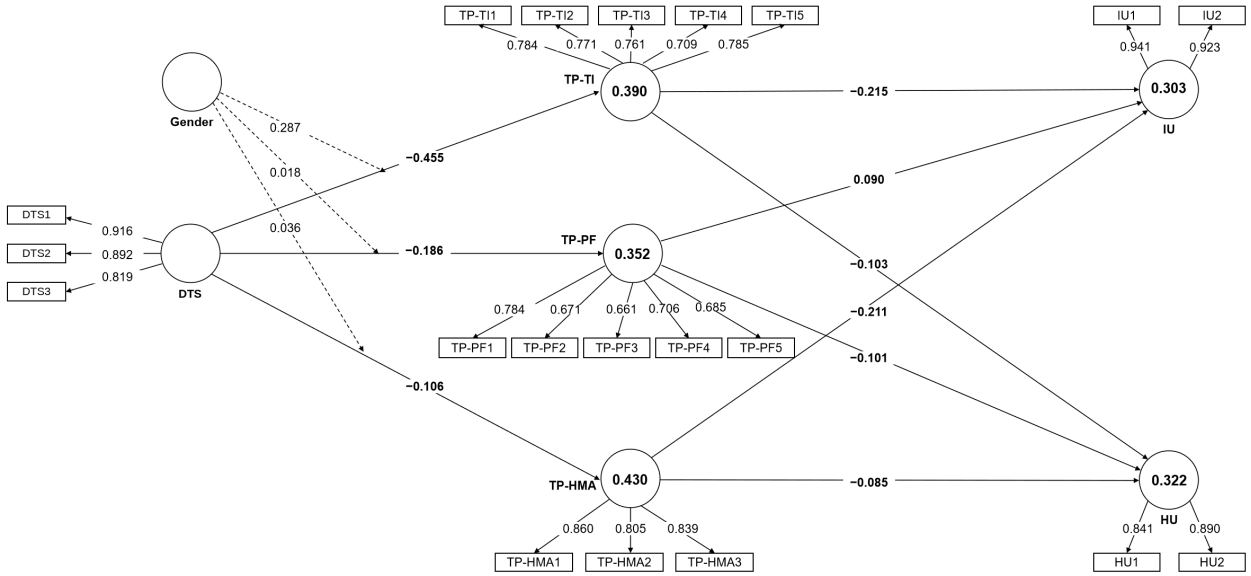
Common Method Bias

Procedural and statistical controls were implemented to address potential common method bias that might arise from self-reported data. Procedurally, a pilot test with older adults improved item clarity and reduced bias. The questionnaire items were phrased in neutral language to avoid leading responses. Statistically, Harman single-factor test was performed via exploratory factor analysis with all measurement items. The results indicated that no single factor accounted for more than 50% of the total variance. Interconstruct correlation coefficients ranged from 0.154 to 0.814 with all values being below the 0.90 threshold [67]. Variance inflation factors (VIFs) for all constructs were between 1.054 and 3.261, which were below the critical value of 3.3 [68], thus eliminating severe multicollinearity.

Figure 2 illustrates the final partial least squares structural equation modeling (PLS-SEM) model with standardized path coefficients.

XSL•FO
RenderX

Figure 2. PLS-SEM results. Standardized path coefficients are shown on structural paths, and values inside endogenous constructs represent explained variance (R^2). Outer values represent indicator loadings. Dashed lines indicate moderation effects of gender. HU: hedonic use; IU: instrumental use; DTS: daily technical support; TP-HMA: technophobia-human-machine ambiguity; TP-PF: technophobia-personal failure; TP-TI: technophobia-technology inconvenience.



Data Analysis

PLS-SEM was conducted in SmartPLS version 4.1.1.4 (SmartPLS GmbH) to estimate the measurement and structural models examining how DTS and technophobia shape older adults' technology engagement. PLS-SEM is well suited to this study because it accommodates complex models with multiple latent constructs, is robust to nonnormal survey data, and performs well with relatively small samples. Therefore, model parameters were first obtained using the partial least squares (PLS) algorithm. Statistical inference for all hypothesized effects—including direct paths, the moderation and indirect (mediated) effects—was then performed via a nonparametric bootstrapping procedure with 5000 resamples. For the moderated mediation hypothesis, conditional indirect effects were computed at different levels of the moderator, and their significance was evaluated using 95% bias-corrected and accelerated (BCa) bootstrap CIs. Effects were considered significant when the corresponding 95% CI did not include zero. This approach is consistent with recent applications of moderated-mediation designs examining sex or gender-differentiated pathways in health-related outcomes [69,70].

Results

Participants' Demographic Profile

Participants' demographic characteristics are reported in Table 2. The sample (N=212) had a mean age of 66.96 (SD 6.96; range 55 - 86) years and was predominantly women (142/212, 68.4%). Most respondents lived in urban areas (203/212, 95.8%) and were married or cohabiting (192/212, 92.9%). In addition, 68.9% (146/212) reported having volunteer service experience, indicating active social participation among the sample. Regarding family structure, most reported having 1 or 2 children. Overall health was generally moderate, with over 40% (95/212) reporting no or minimal limitations in daily activities and most rating their health as fair or good. Educational attainment was relatively low, with nearly 60% (126/212) reporting fewer than 12 years of schooling and about one-fifth having completed higher education. The digital devices used by participants varied, with smartphones being the most commonly owned device (187/212, 88.2%), followed by smart televisions (48/212, 22.6%) and smartwatches or wearable devices (26/212, 12.3%). Regarding DTS, participants reported support from both same-age peers and younger adults, with 25.9% (55/212) reporting more than 10 same-age technical supporters and 26.4% (56/212) reporting more than 10 younger-adult technical supporters; only 12.7% (27/212) and 14.6% (31/212), respectively, reported having no such supporters.

Measurement Model Analysis

Table 1 reports factor loadings, reliability, and convergent validity. Most outer loadings exceeded the recommended threshold of 0.708, ranging from 0.661 to 0.916, indicating adequate indicator reliability [71]. Convergent validity was supported, with Cronbach α ranging from 0.668 to 0.929, composite reliability (CR) from 0.832 to 0.946, and average variance extracted (AVE) from 0.508 to 0.779; all CR values exceeded 0.70 and AVE values exceeded 0.50 [72]. Discriminant validity was established using the Fornell–Larcker criterion and Heterotrait-Monotrait ratio (HTMT). As shown in Table 3, the square roots of AVE exceeded interconstruct correlations, and all HTMT ratios were below 0.90 [67].

XSL•FO
RenderX

Table . Assessment of discriminant validity[a].

Fornell-Larcker criterion	DTS[b]	IU[c]	HU[d]	TP-TI[e]	TP-PF[f]	TPHA[g]
DTS	0.797[a]	0.495	0.495	−0.351	−0.260	−0.259
IU	0.495	0.867[a]	0.814	−0.351	−0.351	−0.423
HU	0.459	0.814	0.865[a]	−0.364	−0.373	−0.391
TP-TI	−0.351	−0.351	−0.364	0.712[a]	0.673	0.459
TP-PF	−0.260	−0.351	−0.373	0.673	0.852[a]	0.750
TP-HMA	−0.259	−0.423	−0.391	0.450	0.750	0.833[a]
HTMT[h]						
DTS	—[i]	0.631	0.669	0.504	0.306	0.317
IU	0.631	—	0.795	0.405	−0.390	−0.492
HU	0.669	0.795	—	0.486	0.461	0.520
TP-TI	0.504	0.405	0.486	—	0.673	0.732
TP-PF	0.306	0.390	0.461	0.673	—	0.865
TP-HMA	0.317	0.492	0.520	0.732	0.865	—

[a]Diagonal values represent the square root of the average variance extracted (AVE) for each construct; off-diagonal values represent interconstruct correlations. HTMT values are reported in the lower panel.

[b]DTS: daily technical support.

[c]IU: instrumental use.

[d]HU: hedonic use.

[e]TP-TI: technophobia-technology inconvenience.

[f]TP-PF: technophobia-personal failure.

[g]TP-HMA: technophobia-human-machine ambiguity.

[h]HTMT: Heterotrait-Monotrait ratio.

[i]Not available.

Structural Model Analysis

Structural Model Evaluation

Before hypothesis testing, the structural model was assessed for collinearity, explanatory power, and predictive relevance (Table 4). Inner VIFs ranged from 1.009 to 3.404, indicating no serious multicollinearity concerns. The model explained 16.5% of variance in technophobia-technology inconvenience (TP-TI), 11.6% in technophobia-personal failure (TP-PF), and 10.6% in technophobia-human-machine ambiguity (TP-HMA), indicating modest levels of explained variance, and 27% of the variance in IU and 29.5% in HU, reflecting moderate explanatory power [71]. Effect size (f^2) results showed that technical support had its largest impact on TP-TI (f^2=0.126, small-to-moderate), with small effects on TP-PF (f^2=0.034) and TP-HMA (f^2=0.036), while effect sizes for instrumental and HU were negligible (all f^2<0.03). All Q^2 values were above zero, indicating adequate predictive relevance.

XSL•FO
RenderX

Table . Structural model evaluation[a].

Endogenous construct	R^2	f^2	Inner VIF[b] range	Q^2
TP-TI[c]	0.165	DTS[d] (0.126)	1.009 - 2.956	0.092
TP-PF[e]	0.116	DTS (0.034)	1.009 - 2.956	0.073
TP-HMA[f]	0.106	DTS (0.036)	1.009 - 2.956	0.067
IU[g]	0.270	TI[h] (0.029)	1.082 - 3.404	0.141
HU[i]	0.295	TI (0.007)	1.082 - 3.404	0.162

[a] VIF <5 (with most values <3.3) indicates no multicollinearity; f^2=0.02, 0.15, and 0.35 denote small, medium, and large effect sizes; Q^2>0 indicates predictive relevance.

[b]VIF: variance inflation factors.

[c]TP-TI: technophobia-technology inconvenience.

[d]DTS: daily technical support.

[e]TP-PF: technophobia-personal failure.

[f]TP-HMA: technophobia-human-machine ambiguity.

[g]IU: instrumental use.

[h]TI: technology inconvenience.

[i]HU: hedonic use.

Path Analysis

To further test the hypothesis, path analysis was conducted. As shown in Table 5, the effects of each dimension of technophobia on technology use (H1) showed that TP-TI significantly reduced IU (β=−0.215, 95% BCa CI −0.410 to - 0.009; *P*=.045), while TP-HMA also negatively predicted IU (β=−0.211, 95% BCa CI −0.389 to 0.001; *P*=.04). By contrast, TP-PF did not significantly predict either IU or HU, as the corresponding 95% BCa CI included zero. Human-machine ambiguity was also significantly negatively associated with IU (β=−0.211, 95% BCa CI −0.389 to −0.001; *P*=.04). Next, DTS was found to exert a significantly positive direct effect on IU (β=0.104, 95% BCa CI 0.009 to 0.203*; P*=.04), but its direct path to HU was not significant, with the 95% BCa CIs including zero, which partially supports H2. In addition, results showed that DTS negatively predicted the dimension of perceived TP-TI (β=−0.455, 95% BCa CI −0.632 to −0.229; *P*<.001), while the effects of DTS on TP-PF and on TP-HMA were not significant, as both CIs crossed zero. Hence, only H3a is supported.

XSL•FO
RenderX

Table . Hypothesis testing[a].

Hypotheses and paths	β[b] (95% BCa[c] CI)	*t* test[d]	*P* values	Supported
H1a				
TP-TI[e] → IU[f]	−0.215 (−0.410 to −0.009)	2.01	.05	Yes
H1b				
TP-PF[g] → HU[h]	−0.101 (−0.351 to 0.132)	0.82	.41	No
H1c				
TP-HMA[i] → HU	−0.085 (−0.273 to 0.145)	0.80	.43	No
TP-HMA → IU	−0.211 (−0.389 to −0.001)	2.10	.04	Yes
H2				
DTS[j] → IU	0.104 (0.009 to 0.203)	2.06	.04	Yes
DTS → HU	0.075 (−0.018 to 0.166)	1.58	.11	No
H3a				
DTS → TP-TI	−0.455 (−0.632 to −0.229)	5.47	.00	Yes
H3b				
DTS → TP-PF	−0.186 (−0.400 to 0.073)	1.57	.12	No
H3c				
DTS → TP-HMA	−0.106 (−0.286 to 0.075)	1.14	.26	No
H5a				
DTS × Gender→TP-TI	0.287 (0.052 to 0.518)	2.40	.02	Yes
H5b				
DTS × Gender→TP-PF	0.018 (−0.274 to 0.268)	0.13	.90	No
H5c				
DTS × Gender→TP-HMA	0.036 (−0.198 to 0.255)	0.32	.75	No

[a]CI values not including zero indicate statistical significance.

[b]β: standardized path coefficient.

[c]BCa: Bias-corrected and accelerated CI.

[d]*t* values were derived from PLS-SEM bootstrapping (5000 resamples); df is not applicable.

[e]TP-TI: technophobia-technology inconvenience.

[f]IU: instrumental use.

[g]TP-PF: technophobia-personal failure.

[h]HU: hedonic use.

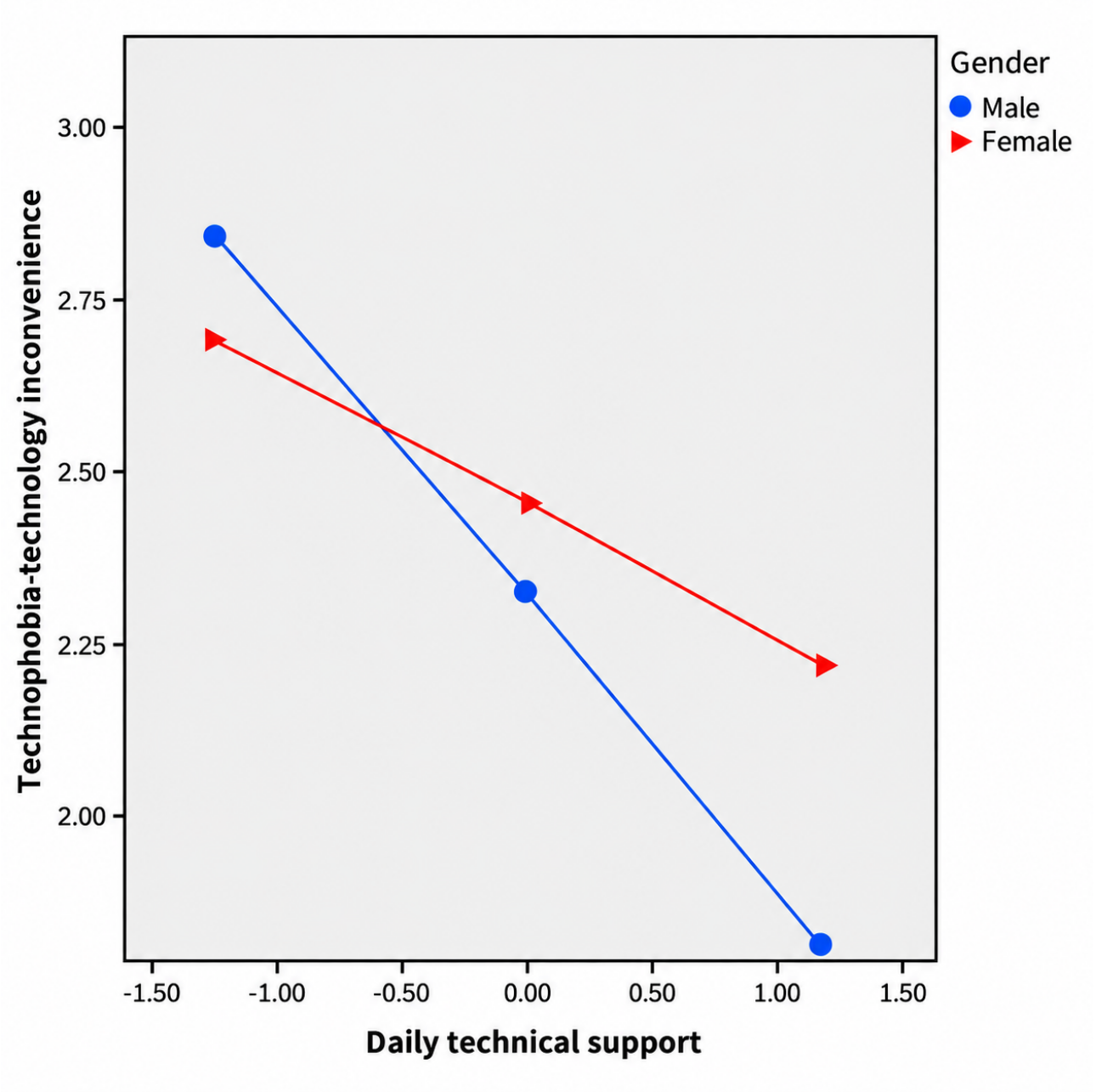
[i]TP-HMA: technophobia-human-machine ambiguity.

[j]DTS: daily technical support.

Lastly, the interaction effect between DTS and gender was significant only for TP-TI (β=0.287, 95% BCa CI 0.052-0.518; *P*=.02), but not for TP-PF or TP-HMA, as the corresponding CIs included zero. We further conducted the simple slope analysis. Figure 3 presents that fear of inconvenience declined with increasing DTS for both genders. However, the decline was substantially steeper among males, indicating that daily support more effectively reduced perceived inconvenience for men (β=−0.455; *P*<.001) than for women (β=−0.168; *P*=.07). H5a is therefore supported.

XSL•FO
RenderX

Figure 3. Simple slopes analysis.



Additional moderation analyses were conducted to examine whether gender moderated the overall associations between DTS and the 2 forms of technology use. Neither the interaction predicting IU (β=0.116, SE .115, *t*=1.014, 95% BCa CI −0.107 to 0.334; *P*=.31) nor that predicting HU (β=0.165, SE .114, *t*=1.444, 95% BCa CI −0.056 to 0.391; *P*=.15) was significant.

Moderated Mediation Effect Analysis

Furthermore, bootstrap analysis (5000 resamples, bias-corrected) was conducted to examine the moderated mediating effects of technophobia in the interaction effect between DTS and gender on 2 types of technology use (H4, H5). Gender was coded as 0=men (reference category) and 1=women (comparison category). The results indicated that only TP-TI significantly mediated the effect of DTS on IU (β=0.097, 95% BCa CI 0.020-0.204), thus suggesting that part of the positive influence of DTS on goal-oriented use operates through reducing perceptions of inconvenience. By contrast, the indirect effects through TP-PF (β=–0.017, 95% BCa CI –0.059 to 0.150) and TP-HMA (β=0.009, 95% BCa CI –0.046 to 0.146) were not supported, as the corresponding CIs included zero (Table 6). These findings indicate that among the 3 dimensions of technophobia, only TP-TI acts as a significant mediator that links DTS and instrumental technology use.

XSL•FO
RenderX

Table . Moderated mediation analysis[a].

Hypothesis	Group	β[b] (95% BCa CI[c])	Supported
Overall indirect effects			
H4a			
DTS[d] → TP-TI[e] → IU[f]	—[g]	0.097 (0.020 to 0.204)	Yes
DTS → TP-TI → HU[h]	—	0.047 (−0.046 to 0.146)	No
H4b			
DTS → TP-PF[i] → IU	—	−0.017 (−0.059 to 0.150)	No
DTS → TP-PF → HU	—	0.019 (−0.018 to 0.125)	No
H4c			
DTS → TP-HMA[j] → IU	—	0.009 (−0.046 to 0.146)	No
DTS → TP-HMA → HU	—	0.022 (−0.012 to 0.091)	No
Conditional direct effects			
DTS → TP-TI	Male (0)	−0.455 (−0.657 to - 0.251)	—
DTS → TP-TI	Female (1)	−0.168 (−0.354 to 0.001)	—
DTS → TP-PF	Male (0)	−0.186 (−0.400 to 0.073)	—
DTS → TP-PF	Female (1)	−0.168 (−0.335 to - 0.001)	—
DTS → TP-HMA	Male (0)	−0.106 (−0.286 to 0.075)	—
DTS → TP-HMA	Female (1)	−0.070 (−0.238 to 0.098)	—
Conditional indirect effects			
DTS → TP-TI → IU	Male (0)	0.098 (0.001 to 0.219)	Yes
DTS → TP-TI → IU	Female (1)	0.036 (−0.004 to 0.111)	No
DTS → TP-TI → HU	Male (0)	0.047 (−0.047 to 0.146)	No
DTS → TP-TI → HU	Female (1)	0.017 (−0.017 to 0.073)	No
DTS → TP-PF → IU	Male (0)	−0.017 (−0.074 to 0.057)	No
DTS → TP-PF → IU	Female (1)	−0.015 (−0.070 to 0.037)	No
DTS → TP-HMA → IU	Male (0)	0.022 (−0.014 to 0.089)	No
DTS → TP-HMA → IU	Female (1)	0.015 (−0.019 to 0.074)	No

[a]CI values not including zero indicate statistical significance. Conditional indirect effects were estimated using bias-corrected bootstrapping (5000 resamples). Gender was dummy-coded as 0=men and 1=women.

[b]β: Standardized path coefficient.

[c] BCa: Bias-corrected and accelerated CI.

[d]DTS: daily technical support.

[e]TP-TI: technophobia-technology inconvenience.

[f]IU: instrumental use.

[g]Not available.

[h]HU: hedonic use.

[i]TP-PF: technophobia-personal failure.

[j]TP-HMA: technophobia-human-machine ambiguity.

Conditional indirect effect analysis showed that the indirect effect of DTS on IU through TP-TI was significant for men (β=0.098, 95% BCa CI 0.001-0.219) but not significant for women (β=0.036, 95% BCa CI −0.002 to 0.116). All other indirect effects were nonsignificant for both genders, as the corresponding 95% BCa CI included zero.

Discussion

Principal Findings

Using a cross-sectional survey of community-dwelling older adults and PLS-SEM for hypothesis testing, the current research found that DTS lowers the technology-inconvenience dimension of technophobia, and this reduction subsequently promotes

XSL•FO
RenderX

instrumental—though not hedonic—technology use, especially among older men. By contrast, DTS was not significantly related to the personal-failure or human-machine ambiguity dimensions, but higher human-machine ambiguity was indeed linked to lower IU. These results refine theory by specifying a dimension-specific and gendered mechanism, and offer practical guidance to bridge the digital divide.

Distinct Fear Sources Linked to Instrumental Technology Use

The first finding of this research is that not all fears under the umbrella of technophobia equally predict older adults' actual technology use. Older individuals' instrumental (vs hedonic) use of technology was significantly influenced by their fear of inconvenience and by human-machine ambiguity but not by their fear of personal failure. Our finding is consistent with other research that has shown that extrinsic motivations, especially technology usefulness, often outweigh intrinsic motivations in determining the usage of involuntary technologies [73,74]. On the other hand, the finding from the dimension of human-machine ambiguity aligns with Smrke et al [27] findings, that older adults who are uneasy about technology's role tend to stay away from new health technologies.

Fear of personal failure did not significantly predict either instrumental or hedonic technology use among older adults in our study. One possible explanation is that our survey captured a general, self-reported belief about technological competence rather than a contextually primed perception [75]. Research on age stereotypes suggests that the activation of negative self-stereotypes is often context dependent and more likely to occur when salient age-related cues are present [76]. As a result, older adults may not explicitly endorse negative self-evaluations in a generic survey setting, which could attenuate observed associations with actual technology use.

The Role of DTS Among Older Adults

The key finding is that the effect of DTS on technophobia varies by dimension and gender. Daily support significantly reduced fear of inconvenience but did not significantly change fear of personal failure or human-machine ambiguity. These latter 2 dimensions are more closely related to internal psychological concerns than to the instrumental difficulties captured by fear of inconvenience. Previous research has indicated that fear of personal failure is often rooted in internalized age stereotypes about older adults' technological incompetence [6] and typically requires the development of self-efficacy through structured learning [77]. The nonsignificant association observed in the current study differs from findings reported in some intervention studies [53,54]. One possible explanation is that older adults may respond differently to different forms of technical support [53,54]. Informal, immediate help may function primarily as instrumental problem-solving, whereas structured programs may be more capable of changing underlying confidence and motivation [74,78]. Similarly, human-machine ambiguity reflects broader attitudinal or ideological concerns about technology and its potential to replace or reduce human interaction [42]. Because TS typically targets immediate operational issues rather than these deeper beliefs, its lack of a significant association with human-machine ambiguity is understandable.

More importantly, this research uncovered that the effect of DTS on reducing the fear of inconvenience is more salient among older men than among older women. Our research resonates with existing findings, that men are more likely to address the instrumental aspect of social support than women, thus the extrinsic motivation-related fear is affected [59]. However, our research did not yield significant gender differences in the effect of technical support on fear of personal failure [63,64]. This outcome may also support the point proposed above that older adults generally seek technical support primarily to solve problems (ie, extrinsic motivation) rather than to obtain internal psychological relief (ie, intrinsic motivation). Notably, additional analyses provided no evidence that gender moderated the overall associations between DTS and either instrumental or hedonic technology use. This pattern suggests that gender may function more specifically as a boundary condition of older adults' psychological responses to technical support rather than as a general moderator of its behavioral association with technology use.

On the other hand, DTS has a significantly positive effect on older adults' IU of technology but no significant effect on HU. This finding is consistent with prior observations that many older adults view technology, such as mobile apps, primarily as a tool for social engagement rather than a source of entertainment [79]. Therefore, DTS, by increasing technology accessibility and decreasing frustration, demonstrates the concrete utility of technology in everyday tasks [12]. Meanwhile, many older people have well-established leisure routines outside of digital technology; even with daily assistance available, they may not be motivated to explore new entertainment technologies [79].

Moderated Mediating Effect of Fear of Inconvenience

The gender-specific finding demonstrated that DTS boosts instrumental technology use by reducing fear of inconvenience to a greater extent among older men than among older women. This result is in line with the established TAM; that is, technical support serves as an external environmental factor that affects one's perception toward technology per se, which is further a key driver of technology adoption [34]. It bridges 3 dimensions of technophobia and 2 classic focal predictors of technology use (ie, technology usefulness and ease of use), as well as incorporates gender as a moderator. As such, it generalizes TAM into older adults and enables a nuanced understanding of how the social environment can influence older men and women's technology adoption behavior differently.

Implication and Limitation

Our findings suggest a shift in how interventions to promote technology use among older adults are designed. Older adults are often stereotyped as technophobic and lacking confidence in their ability to use technology, and interventions have therefore tended to emphasize intrinsic empowerment (skills training and confidence building). Our findings indicate that more instrumental forms of support may also be important, particularly for technologies that have become difficult to avoid

XSL•FO
RenderX

in everyday life. By reducing the inconvenience, effort, and practical burden associated with technology use, DTS may help older adults recognize the functional value of technology and facilitate IU. Accordingly, interventions could place greater emphasis on readily accessible, task-oriented assistance, such as just-in-time support, simplified workflows, default configurations, and age-friendly interface design with larger fonts and icons, simplified navigation, and fewer steps required to complete common tasks [80].

The stronger indirect association observed among men further suggests that older adults may not benefit from the same form of technical support to the same extent. Rather than treating gender itself as the basis for intervention, this finding highlights the importance of identifying the specific barriers that users experience and tailoring support accordingly. For individuals whose technology use is particularly constrained by perceived inconvenience, immediate and task-focused assistance may be especially effective, whereas those whose difficulties stem from other sources of technophobia may require different forms of support, such as structured learning, confidence-building, or greater reassurance about human-technology interaction. Gender may therefore serve as 1 contextual indicator of heterogeneity in how technical support is experienced, but intervention decisions should ultimately be guided by users' specific needs and sources of technological fear [22]. Such a barrier-specific approach may provide a more precise and sustainable way to promote technology engagement among older adults.

Despite these findings, our research is not without limitations. First, the survey sample was small and skewed toward urban older adults, thus limiting the generalizability of our findings to diverse socioeconomic and rural populations. As the sample was drawn from one country, the magnitude and relative salience of these associations may also vary across cultural and social settings. Second, the cross-sectional design precludes strong causal inference about the proposed relationship between variables. Third, measures relied on self-reports, which are susceptible to recall and social desirability biases. For instance, older men may underreport technophobia because of gendered norms about technological competence. Future research should use larger and more diverse samples, including cross-cultural comparisons, and longitudinal designs. Complementing surveys with objective behavioral indicators, such as smartphone or app usage logs and passively sensed interaction traces, could also provide fine-grained evidence on how daily support translates into reduced technophobia and sustained technology use. Future studies could further examine how emerging forms of online and real-time technical support can translate these findings into practice. As older adults increasingly engage with digital technologies as everyday users, platforms could potentially use user profiles and interaction data to provide more timely and tailored support, including gender-sensitive prompts or assistance. Finally, we failed to find a significant predictor of older adults' HU of technology; future research can focus on this specific technology domain.

Conclusion

This study examined how DTS relates to older adults' instrumental and hedonic technology use, with fear-specific technophobia as a mediating mechanism and gender as a moderator. TA was associated with lower fear of inconvenience, which in turn was associated with greater instrumental technology use, and this indirect effect was stronger among men than women. No comparable pathway was found for HU. Theoretically, these findings advance a more differentiated understanding of technophobia by highlighting the importance of examining specific sources of technological fear rather than treating technophobia as a unitary barrier. The absence of a comparable pathway for HU further suggests that different forms of technology use may be shaped by distinct psychological mechanisms and therefore warrant more use-specific investigation. Practically, the findings underscore the value of moving beyond one-size-fits-all approaches to digital inclusion by matching support strategies to the particular barriers older adults experience. Efforts to reduce technophobia should therefore complement generic skills training with targeted assistance that addresses specific sources of fear, practical barriers that can be alleviated through daily support, and differences in how older adults psychologically respond to such support. Such a more targeted approach may help improve the effectiveness of interventions designed to promote sustained and meaningful technology engagement in later life.

Acknowledgments

The authors would like to thank all participants for their time and contributions to this study and the trained volunteers who assisted with offline questionnaire administration.

ChatGPT (OpenAI) was used during the revision stage following receipt of the reviewers' comments, specifically for language editing and improving the clarity of the manuscript. All AI-assisted outputs were critically reviewed and revised by the authors, who take full responsibility for the accuracy, originality, and integrity of the final manuscript.

Funding

This work was supported by the Humanities and Social Sciences Research Project of the Ministry of Education in China (number 24YJC870013) and the Jiangsu Provincial Social Science Fund Project (number 26GLC028), the Philosophy and Social Sciences Research Project of Colleges and Universities in Jiangsu Province (number 2024SJYB0241), Key Project of Aging Civilization Think Tank (number 26LLWM06), and National Natural Science Foundation of China (number 72202210).

XSL•FO
RenderX

Data Availability

Data are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: YY, WX
Data curation: SY
Formal analysis: YY, SD
Investigation: SD, SY
Methodology: YY
Supervision: WX
Writing – original draft: YY
Writing – review & editing: SY, WX, QL

Conflicts of Interest

None declared.

References

1. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2026-09-07]
2. Quan-Haase A, Martin K, Schreurs K. Interviews with digital seniors: ICT use in the context of everyday life. Inf Commun Soc 2016 May 3;19(5):691-707. [doi: 10.1080/1369118X.2016.1140217]
3. Roffarello AM, De Russis L, Lottridge D, Cecchinato ME. Understanding digital wellbeing within complex technological contexts. Int J Hum Comput Stud 2023 Jul;175:103034. [doi: 10.1016/j.ijhcs.2023.103034]
4. Park C, Cho S. Exploring why and how long older adults use the internet. Act Adapt Aging 2026 Jan 2;50(1):117-132. [doi: 10.1080/01924788.2025.2492985]
5. Reisdorf BC, Petrovčič A, Grošelj D. Going online on behalf of someone else: characteristics of internet users who act as proxy users. New Media Soc 2021 Aug;23(8):2409-2429. [doi: 10.1177/1461444820928051]
6. Xi W, Zhang X, Ayalon L. When less intergenerational closeness helps: the influence of intergenerational physical proximity and technology attributes on technophobia among older adults. Comput Human Behav 2022 Jun;131:107234. [doi: 10.1016/j.chb.2022.107234]
7. Wang K, Chen XS, Gu D, Smith BD, Dong Y, Peet JZ. Examining first- and second-level digital divide at the intersection of race/ethnicity, gender, and socioeconomic status: an analysis of the national health and aging trends study. Gerontologist 2024 Sep 1;64(9):gnae079. [doi: 10.1093/geront/gnae079] [Medline: 38898816]
8. China Internet Network Information Center. The 57th statistical report on china's internet development. : Beijing, China: China Internet Network Information Center; 2026 URL: https://www.cnnic.cn/n4/2026/0304/c88-11549.html [accessed 2026-09-21]
9. China internet network information center. The 49th statistical report on China's internet development. : Beijing, China: China Internet Network Information Center; 2022 URL: https://www.cnnic.cn/n4/2022/0401/c88-1131.html [accessed 2026-09-21]
10. van Deursen AJ, Helsper EJ. A nuanced understanding of internet use and non-use among the elderly. Eur J Commun 2015 Apr;30(2):171-187. [doi: 10.1177/0267323115578059]
11. Hunsaker A, Hargittai E. A review of internet use among older adults. New Media Soc 2018 Oct;20(10):3937-3954. [doi: 10.1177/1461444818787348]
12. Wilson G, Gates JR, Vijaykumar S, Morgan DJ. Understanding older adults' use of social technology and the factors influencing use. Ageing Soc 2023 Jan;43(1):222-245. [doi: 10.1017/S0144686X21000490]
13. Individuals – frequency of internet use. Eurostat. URL: https://ec.europa.eu/eurostat/databrowser/view/isoc_ci_ifp_fu/default/table?lang=en [accessed 2026-09-21]
14. Facts and figures about digital inclusion and older people. : London, UK: Age UK; 2024 URL: https://www.ageuk.org.uk/siteassets/documents/reports-and-publications/reports-and-briefings/active-communities/internet-use-statistics-june-2024.pdf [accessed 2026-09-21]
15. Khan S, Webster S, Puxty JAH, Robertson M. Key challenges and barriers to digital literacy for older adults: scoping review. JMIR Aging 2026 Mar 16;9:e80647. [doi: 10.2196/80647] [Medline: 41838931]
16. Osiceanu ME. Psychological implications of modern technologies: "Technofobia" versus "Technophilia". Procedia Soc Behav Sci 2015 May;180:1137-1144. [doi: 10.1016/j.sbspro.2015.02.229]
17. Khasawneh OY. Technophobia: examining its hidden factors and defining it. Technol Soc 2018 Aug;54:93-100. [doi: 10.1016/j.techsoc.2018.03.008]
18. Martínez-Córcoles M, Teichmann M, Murdvee M. Assessing technophobia and technophilia: development and validation of a questionnaire. Technol Soc 2017 Nov;51:183-188. [doi: 10.1016/j.techsoc.2017.09.007]

XSL•FO
RenderX

19. Sinkovics RR, Stöttinger B, Schlegelmilch BB, Ram S. Reluctance to use technology‐related products: development of a technophobia scale. Thunderbird Intl Bus Rev 2002 Jul;44(4):477-494. [doi: 10.1002/tie.10033]
20. Shedletsky LJ. Internet training for older adult learners: an intergenerational mentoring approach. LLI Review 2006(1):34-43 [FREE Full text]
21. Lee OEK, Kim DH. Bridging the digital divide for older adults via intergenerational mentor-up. Res Soc Work Pract 2019 Oct;29(7):786-795. [doi: 10.1177/1049731518810798]
22. Xi W, Zhang X, Ayalon L. The framing effect of intergenerational comparison of technologies on technophobia among older adults. J Gerontol B Psychol Sci Soc Sci 2022 Jul 5;77(7):1179-1185. [doi: 10.1093/geronb/gbab199] [Medline: 34687290]
23. Rosen LD, Sears DC, Weil MM. Computerphobia. Behav Res Methods Instrum Comput 1987 Mar;19(2):167-179. [doi: 10.3758/BF03203781]
24. Maricutoiu LP. A meta-analysis on the antecedents and consequences of computer anxiety. Procedia Soc Behav Sci 2014 Apr;127:311-315. [doi: 10.1016/j.sbspro.2014.03.262]
25. Niu HJ, Li MH, Hsieh FY, Yu CC, Lin CT. From digital anxiety to empowerment in older adults: cross-sectional survey study on psychosocial drivers of digital literacy. JMIR Aging 2026 Jan 21;9:e75245. [doi: 10.2196/75245] [Medline: 41564350]
26. Sun E, Ye X. Older and fearing new technologies? The relationship between older adults' technophobia and subjective age. Aging Ment Health 2024 Apr;28(4):569-576. [doi: 10.1080/13607863.2023.2241017] [Medline: 37553781]
27. Smrke U, Špes T, Mlakar I, Musil B, Plohl N. Technophobia mediates the associations between age, education level, and readiness to adopt new (Health) technology among aging adults. J Appl Gerontol 2025 Mar;44(3):497-507. [doi: 10.1177/07334648241274260] [Medline: 39177432]
28. Lee OE, Kim DH, Lee H, Beum KA. Information communication technology use to improve eHealth literacy, technophobia, and social connection among community dwelling older adults. Educ Gerontol 2022 Oct 3;48(10):445-457. [doi: 10.1080/03601277.2022.2045727]
29. Koul S, Eydgahi A. The impact of social influence, technophobia, and perceived Ssafety on autonomous vehicle technology adoption. Period Polytech Transp Eng 2020;48(2):133-142. [doi: 10.3311/PPtr.11332]
30. Nimrod G. Technophobia among older internet users. Educ Gerontol 2018 Mar 4;44(2-3):148-162. [doi: 10.1080/03601277.2018.1428145]
31. Massey AP, Khatri V, Montoya‐Weiss MM. Usability of online services: the role of technology readiness and context. Decis Sci 2007 May;38(2):277-308. [doi: 10.1111/j.1540-5915.2007.00159.x]
32. van der Heijden H. User acceptance of hedonic information systems1. MIS Q 2004 Dec 1;28(4):695-704. [doi: 10.2307/25148660]
33. Hirschman EC, Holbrook MB. Hedonic consumption: emerging concepts, methods and propositions. J Mark 1982 Jul;46(3):92-101. [doi: 10.1177/002224298204600314]
34. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
35. Wu J, Lu X, California State University, East Bay. Effects of extrinsic and intrinsic motivators on using utilitarian, hedonic, and dual-purposed information systems: a meta-analysis. J Assoc Inf Syst 2013;14(3):153-191. [doi: 10.17705/1jais.00325]
36. Wang Q, Zhan H, Liang Y, Li P, Ai L. Perceived ease of use, perceived enjoyment, and continuance intention as predictors of medical students' well-being in hybrid learning: an extended TAM approach. Front Psychol 2025;16:1704678. [doi: 10.3389/fpsyg.2025.1704678]
37. Fagan MH, Neill S, Wooldridge BR. Exploring the intention to use computers: an empirical investigation of the role of intrinsic motivation, extrinsic motivation, and perceived ease of use. J Comput Inf Syst 2008 Mar;48(3):31-37. [doi: 10.1080/08874417.2008.11646019]
38. Malhotra Y, Galletta DF, Kirsch LJ. How.endogenous motivations influence user intentions: beyond the dichotomy of extrinsic and intrinsic user motivations. J Manage Inf Syst 2008 Jul;25(1):267-300. [doi: 10.2753/MIS0742-1222250110]
39. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view1. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
40. Deci EL, Ryan RM. The "What" and "Why" of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000 Oct;11(4):227-268. [doi: 10.1207/S15327965PLI1104_01]
41. Li C, Yanagisawa H. Intrinsic motivation in virtual assistant interaction for fostering spontaneous interactions. PLoS ONE 2021;16(4):e0250326. [doi: 10.1371/journal.pone.0250326]
42. Lee S(. The role of intrinsic motivations on customers' service robot use experience. J Qual Assur Hosp Tour 2025 Nov 2;26(6):1194-1217. [doi: 10.1080/1528008X.2023.2289381]
43. Vila J. Social support and longevity: meta-analysis-based evidence and psychobiological mechanisms. Front Psychol 2021;12:717164. [doi: 10.3389/fpsyg.2021.717164] [Medline: 34589025]
44. Dai S, Shi Y, Zhang Y, et al. Influencing factors of two-way social support for the old adults in China: a cross-sectional study. Geriatr Nurs 2023;54:192-198. [doi: 10.1016/j.gerinurse.2023.09.007] [Medline: 37801874]

XSL•FO
RenderX

45. Yin S, Yang Q, Xiong J, Li T, Zhu X. Social support and the incidence of cognitive impairment among older adults in China: findings from the Chinese longitudinal healthy longevity survey study. Front Psychiatry 2020;11:254. [doi: 10.3389/fpsyt.2020.00254] [Medline: 32317993]
46. Lindsay Smith G, Banting L, Eime R, O'Sullivan G, van Uffelen JGZ. The association between social support and physical activity in older adults: a systematic review. Int J Behav Nutr Phys Act 2017 Apr 27;14(1):56. [doi: 10.1186/s12966-017-0509-8] [Medline: 28449673]
47. Sulandari S, Coats RO, Miller A, Hodkinson A, Johnson J. A systematic review and meta-analysis of the association between physical capability, social support, loneliness, depression, anxiety, and life satisfaction in older adults. Gerontologist 2024 Nov 1;64(11):gnae128. [doi: 10.1093/geront/gnae128] [Medline: 39233622]
48. Samtani S, Mahalingam G, Lam BCP, et al. Emotional and instrumental social support and older adults' depressive symptoms: collaborative individual participant data meta-analysis of 11 population-based studies of aging. Am J Epidemiol 2025 Oct 7;194(10):3041-3049. [doi: 10.1093/aje/kwaf137] [Medline: 40643376]
49. Yeung GTY, Fung HH. Social support and life satisfaction among Hong Kong Chinese older adults: family first? Eur J Ageing 2007 Dec;4(4):219-227. [doi: 10.1007/s10433-007-0065-1] [Medline: 28794791]
50. Cheng ST. The effect of negative aging self-stereotypes on satisfaction with social support. J Gerontol B Psychol Sci Soc Sci 2020 Apr 16;75(5):981-990. [doi: 10.1093/geronb/gby113]
51. Portz JD, Bayliss EA, Bull S, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. J Med Internet Res 2019 Apr 8;21(4):e11604. [doi: 10.2196/11604] [Medline: 30958272]
52. Mitzner TL, Savla J, Boot WR, et al. Technology adoption by older adults: findings from the PRISM trial. Gerontologist 2019 Jan 9;59(1):34-44. [doi: 10.1093/geront/gny113] [Medline: 30265294]
53. Xie B, Bugg JM. Public library computer training for older adults to access high-quality Internet health information. Libr Inf Sci Res 2009 Sep 1;31(3):155. [doi: 10.1016/j.lisr.2009.03.004] [Medline: 20161649]
54. Xie B. Older adults, e-health literacy, and collaborative learning: an experimental study. J Am Soc Inf Sci 2011 May;62(5):933-946. [doi: 10.1002/asi.21507]
55. Dai S, Xi W, Xue X. Older, but not all fear: latent profile analysis on technophobia and technology adoption among older adults. 2025 Presented at: AMCIS 2025 Proceedings; Aug 14-16, 2025 URL: https://aisel.aisnet.org/amcis2025/intelfuture/intelfuture/41 [accessed 2026-10-14]
56. Hunsaker A, Nguyen MH, Fuchs J, Djukaric T, Hugentobler L, Hargittai E. "He Explained It to Me and I Also Did It Myself": how older adults get support with their technology uses. Socius 2019 Jan;5:5. [doi: 10.1177/2378023119887866]
57. Jin H, Qu Y. Association between intergenerational support, technology perception and trust, and intention to seek medical care on the internet among Chinese older adults: cross-sectional questionnaire study. J Med Internet Res 2025 Jan 6;27:e65065. [doi: 10.2196/65065] [Medline: 39761564]
58. Venkatesh V, Davis FD. A model of the antecedents of perceived ease of use: development and test. Decis Sci 1996 Sep;27(3):451-481. [doi: 10.1111/j.1540-5915.1996.tb01822.x]
59. Venkatesh V, Morris MG. Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior1. MIS Q 2000 Mar 1;24(1):115-139. [doi: 10.2307/3250981]
60. Eagly AH, Wood W, Diekman AB. Social role theory of sex differences and similarities: a current appraisal. In: The Developmental Social Psychology of Gender 2000:123-174 URL: https://www.routledge.com/The-Developmental-Social-Psychology-of-Gender/Eckes-Trautner/p/book/9780805831900 [accessed 2026-09-21]
61. Franz RL, Findlater L, Barbosa Neves B, Wobbrock JO. Gender and help seeking by older adults when learning new technologies. 2019 Oct 24 Presented at: Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility; Oct 25-28, 2026; Pittsburgh, PA p. 136-142. [doi: 10.1145/3308561.3353807]
62. Ayalon L, Tesch-Römer C, editors. Contemporary Perspectives on Ageism: Springer; 2018. [doi: 10.1007/978-3-319-73820-8]
63. Venkatesh V, Morris MG, Sykes TA, Ackerman PL. Individual reactions to new technologies in the workplace: the role of gender as a psychological construct. J Applied Social Pyschol 2004 Mar;34(3):445-467. [doi: 10.1111/j.1559-1816.2004.tb02556.x]
64. Venkatesh V V, Morris M, Ackerman P. A longitudinal field investigation of gender differences in individual technology adoption decision-making processes. Organ Behav Hum Decis Process 2000 Sep;83(1):33-60. [doi: 10.1006/obhd.2000.2896] [Medline: 10973782]
65. Prentice DA, Carranza E. What women and men should be, shouldn't be, are allowed to be, and don't have to be: the contents of prescriptive gender stereotypes. Psychol Women Q 2002 Dec;26(4):269-281. [doi: 10.1111/1471-6402.t01-1-00066]
66. Tamres LK, Janicki D, Helgeson VS. Sex differences in coping behavior: a meta-analytic review and an examination of relative coping. Pers Soc Psychol Rev 2002 Feb;6(1):2-30. [doi: 10.1207/S15327957PSPR0601_1]
67. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J of the Acad Mark Sci 2015 Jan;43(1):115-135. [doi: 10.1007/s11747-014-0403-8]
68. Kock N. Common method bias in PLS-SEM: a full collinearity assessment approach. Int J e-Collab 2015;11(4):1-10. [doi: 10.4018/ijec.2015100101]

XSL•FO
RenderX

69. Luis SE. Body mass index, chronic health conditions, and depression among US adults with disabilities: a modeling study of the 2020–2023 BRFSS. Am J Health Educ 2026 Mar 4;57(2):93-104. [doi: 10.1080/19325037.2025.2559652]
70. Luis SE. Bullying victimization and suicide among US high school adolescents: the mediating role of depressive symptoms and the moderating effects of physical activity and sex. J Sch Violence 2026;25:1-14. [doi: 10.1080/15388220.2025.2596991]
71. Hair JJ, Hult GTM, Ringle CM, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM): Thousand Oaks, CA: Sage Publications; 2014. URL: https://digitalcommons.kennesaw.edu/facbooks2014/39/ [accessed 2026-09-21]
72. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981 Feb;18(1):39-50. [doi: 10.1177/002224378101800104]
73. Vilnai-Yavetz I, Levina O. Motivating social sharing of e-business content: Intrinsic motivation, extrinsic motivation, or crowding-out effect? Comput Human Behav 2018 Feb;79:181-191. [doi: 10.1016/j.chb.2017.10.034]
74. Zafrani O, Nimrod G, Edan Y. Between fear and trust: older adults' evaluation of socially assistive robots. Int J Hum Comput Stud 2023 Mar;171:102981. [doi: 10.1016/j.ijhcs.2022.102981]
75. Wheeler SC, Petty RE. The effects of stereotype activation on behavior: a review of possible mechanisms. Psychol Bull 2001 Nov;127(6):797-826. [doi: 10.1037/0033-2909.127.6.797] [Medline: 11726072]
76. Abrams D, Eller A, Bryant J. An age apart: the effects of intergenerational contact and stereotype threat on performance and intergroup bias. Psychol Aging 2006 Dec;21(4):691-702. [doi: 10.1037/0882-7974.21.4.691] [Medline: 17201490]
77. Vines J, Pritchard GW, Wright PC, Olivier PL, Brittain K. An age-old problem: examining the discourses of ageing in HCI and strategies for future research. ACM Trans Comput Hum Interact 2015;22(1):2. [doi: 10.1145/2696867]
78. Selwyn N. Distrusting Educational Technology: Critical Questions for Changing Times: Abingdon, UK: Routledge; 2014. [doi: 10.4324/9781315886350]
79. Liu JYW, Sorwar G, Rahman MS, Hoque MR. The role of trust and habit in the adoption of mHealth by older adults in Hong Kong: a healthcare technology service acceptance (HTSA) model. BMC Geriatr 2023 Feb 4;23(1):73. [doi: 10.1186/s12877-023-03779-4] [Medline: 36737712]
80. Amouzadeh E, Dianat I, Faradmal J, Babamiri M. Optimizing mobile app design for older adults: systematic review of age-friendly design. Aging Clin Exp Res 2025 Aug 14;37(1):248. [doi: 10.1007/s40520-025-03157-7] [Medline: 40804492]

Abbreviations

AVE: average variance extracted
BCa: bias-corrected and accelerated
CR: composite reliability
DTS: daily technical support
HTMT: heterotrait-monotrait ratio
HU: hedonic use
ICT: information and communication technology
IU: instrumental use
PLS: partial least squares
PLS-SEM: partial least squares structural equation modeling
TAM: Technology Acceptance Model
TP-HMA: technophobia-human-machine ambiguity
TP-PF: technophobia-personal failure
TP-TI: technophobia-technology inconvenience
VIF: variance inflation factor

Edited by F Puga; submitted 07.Feb.2026; peer-reviewed by M Batra, S Luis, T Wells, W Luan; revised version received 17.Aug.2026; accepted 31.Aug.2026; published 23.Oct.2026.

Please cite as:
Yang Y, Dai S, Yin S, Xi W, Li Q
Fear-Specific Technophobia as a Mediator Between Daily Technical Support and Instrumental Versus Hedonic Technology Use in Later Life: Cross-Sectional Survey Study
JMIR Aging 2026;9:e93063
URL: https://aging.jmir.org/2026/1/e93063
doi:10.2196/93063

XSL•FO
RenderX

© Yu Yang, Shuo Dai, Siyu Yin, Wanyu Xi, Qi Li. Originally published in JMIR Aging (https://aging.jmir.org), 23.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient Satisfaction With Telehealth Visits in Rural Compared With Urban Communities: Single-Center Study

Corrin Hepburn[1], MD; Karolina Krawczyk[2], MD; Cara Joyce[2], PhD; Jonah Rubin[2], MD

[1]Division of Gastroenterology, Rush University Medical Center, 1620 W Harrison St, Chicago, IL, United States

[2]Division of Gastroenterology, Loyola University Medical Center, Maywood, IL, United States

Corresponding Author:
Corrin Hepburn, MD
Division of Gastroenterology, Rush University Medical Center, 1620 W Harrison St, Chicago, IL, United States

Abstract

Background: Studies performed in urban communities with access to technology suggest high patient satisfaction with telehealth. While virtual visits can increase the reach of clinical practice in rural communities, technological barriers may reduce patient satisfaction.

Objective: This study aimed to compare satisfaction with telehealth visits between patients living in rural and urban communities.

Methods: A telephone survey was developed and administered to hepatology patients seen at outpatient clinics from March 2020 through March 2021. Patient characteristics and survey responses were compared by urban and rural location as defined by the census tract based on zip code using univariable and multivariable logistic regression.

Results: Of 400 patients, 164 (41%) completed the survey. Compared with urban patients, rural patients had twice the transportation time to clinic (mean 59, SD 35 vs mean 30, SD 15 min) and were more likely to cancel due to transportation issues (21/48, 46% vs 15/116, 13%). Rural patients reported less proficiency with technology and more technical difficulties, including an inability to log on to the portal or access the camera or microphone (35/48, 75% vs 29/116, 25%) and less comfort with their devices (26/48, 54% vs 10/116, 9%). Overall, urban patients were more likely to prefer telehealth (adjusted odds ratio 5.20, 95% CI 2.15 - 13.7) and were more satisfied with telehealth vs in-person visits than rural patients (72/116, 62.1% vs 10/48, 20.8%).

Conclusions: Rural patients reported more technical challenges with telehealth and more transportation issues than urban patients but favored in-person hepatology visits. Urban patients were more satisfied with telehealth visits compared with in-person visits. Research is needed to improve telehealth delivery and satisfaction for rural patients.

(JMIR Aging 2026;9:e85018) doi:10.2196/85018

KEYWORDS

technology literacy; hepatology; patient experience; broadband; technological barriers

Introduction

Telehealth provides an alternative to in-person appointments, and the number of these visits soared during the COVID-19 pandemic [1]. Potential benefits of telehealth include eliminating the need for transportation and reducing the amount of time associated with office visits, which can increase patient convenience and reduce health care costs. Previous research showed high patient satisfaction with telehealth visits, but most of these studies were performed in urban communities [1-3]. Despite the benefits, patients may experience difficulties with these visits related to internet connectivity and technological demands [1,4]. These difficulties are noted to be increased in patients with low technology literacy. Between 55% and 85% of households have broadband internet access, with lower rates in rural communities [5,6]. Our hepatology practice covers a wide geography with clinics throughout the state of Illinois, including patients in both rural and urban communities. Given the disparity in access to broadband in rural areas, our objective was to examine patient satisfaction with telehealth between rural and urban communities.

Methods

Ethical Considerations

This study was reviewed and approved by the institutional review board at Loyola University Medical Center (IRB#LU 214739; approval date October 21, 2021). All participants were adults and provided verbal informed consent prior to participation in the telephone survey. Participants were informed of the voluntary nature of the study, that refusal to participate would not affect their medical care, and that they could withdraw at any time. No protected health information was collected beyond data necessary for study objectives. Survey responses were deidentified and securely stored to maintain confidentiality, in accordance with institutional and ethical guidelines.

XSL•FO
RenderX

Survey Development

A survey tool was developed to assess patient satisfaction with telehealth visits in addition to technology literacy and broadband internet access (Multimedia Appendix 1). Survey questions were identified from a literature search focusing on 4 main themes: demographic information, geographic information, telehealth literacy, and telehealth satisfaction. Telehealth literacy questions were developed to test similar concepts asked in prior validated questionnaires including the eHealth Literacy Scale [7] and the Mobile Device Proficiency Questionnaire [8]. Telehealth satisfaction questions were developed to test similar concepts asked in the Telehealth Usability Questionnaire and asked on a 5-point Likert scale [9].

Patients were approached for participation in this study if they had both a telehealth visit and an in-person appointment between March 2020 to March 2021, when our telehealth program was most active. We excluded patients who were liver transplant recipients and those with hepatocellular carcinoma, as their appointment criteria may require more frequent in-person visits with a multidisciplinary care team and therefore may bias the result. Patients were contacted by telephone by an investigator. Informed consent was obtained over the phone via a script. The survey was then read to these patients and responses were recorded in Research Electronic Data Capture (REDCap) software. For patients who were non–English speaking, a certified interpreter was used over the phone. Information on rural-urban commuting area code was calculated via the most recent US Department of Agriculture data by zip code for the 2010 census [10]. Demographic information was collected from the electronic medical record including age, gender, race, diagnosis, and zip code. Our primary outcome was patient satisfaction with the visit, measured on a 5-point Likert scale. Secondary outcomes included internet access, difficulties with telehealth use, device used, and preference for type of visit in the future.

Statistical Analysis

Descriptive variables were compared using Student *t* tests (2-tailed) for continuous variables and chi-square or Fisher exact tests for nominal variables by location (urban vs rural as defined by census tract) and visit preference (in person or telehealth). Univariable and multivariable analyses with logistic regression models were used to determine odds ratios (ORs) for patient characteristics associated with telehealth preference. Certain questions in the survey used the Likert scale to assess characteristics associated with satisfaction with telehealth which were analyzed with proportional odds logistic regression models. Analyses were performed using the R programming language (version 4.4.1; R Foundation for Statistical Computing).

Results

Of 400 patients who were contacted, 164 (41%) completed the survey. The mean age was 64 (SD 10) years, more than half were male participants (n=88, 53.7%), and three-quarters were non-Hispanic White (122/163, 74.8%). The most common causes of liver disease were alcohol (51/162, 31.5%), metabolic dysfunction–associated steatotic liver disease (49/162, 30.2%), and hepatitis C (36/162, 22.2%; Table 1).

Table . Patient demographics.

Characteristics	Overall (n=164)	Rural (n=48)	Urban (n=116)	*P* value
Age (y), mean (SD)[a]	64 (10)	65 (11)	63 (10)	.35
Sex, n (%)[c]				.86
Male	88 (53.7)	25 (52.1)	63 (54.3)	
Female	76 (46.3)	23 (47.9)	53 (45.7)	
Race, n (%)[a]				<.001
Black	14 (8.6)	0 (0)	14 (12.1)	
Hispanic	25 (15.3)	1 (2.1)	24 (20.7)	
Other	2 (1.2)	0 (0)	2 (1.7)	
White	122 (74.8)	46 (97.9)	76 (65.5)	
Diagnosis, n (%)[b]				.18
Alcohol-related liver disease	51 (31.5)	19 (40.4)	32 (27.8)	
Hepatitis C	36 (22.2)	12 (25.5)	24 (20.9)	
Metabolic dysfunction–associated steatotic liver disease	49 (30.2)	12 (25.5)	37 (32.2)	
Other	26 (16)	4 (8.5)	22 (19.1)	

[a]Data were available for 163 participants (overall),

[b]Data were available for 162 participants (overall),

[c]Data were available for 164 participants (overall)

XSL•FO
RenderX

Using univariate analysis, rural patients (n=48) compared with urban patients (n=116) had double the transportation time to the clinic (mean 59, SD 35 vs mean 30, SD 15 min; *P*<.001), were less likely to reliably attend appointments (36/48, 75% vs 110/116, 96%; *P*<.001), and were more likely to cancel visits due to transportation issues (21/48, 46% vs 15/116, 13%; *P*<.001; Table 2). Rural patients also reported more technical difficulties (35/48, 75% vs 29/116, 25%; *P*<.001), less comfort with their devices (26/48, 54% vs 10/116, 9%; *P*<.001), and less frequent computer use (14/48, 29% vs 78/116, 67%; *P*<.001). Despite these differences, similar proportions of rural and urban patients were willing to have telehealth visits in the future (38/48, 79% vs 100/116, 86%; *P*=.26).

XSL•FO
RenderX

Table . Patient responses to survey.

Characteristics	Overall (n=164)	Rural (n=48)	Urban (n=116)	*P* value
Time to clinic (min), mean (SD)[a]	39 (27)	59 (35)	30 (15)	<.001
Mode of transportation, n (%)				.56
Drive	161 (98.2)	48 (100)	113 (97.4)	
Public transportation	3 (1.8)	0 (0)	3 (2.6)	
Transportation, n (%)				.64
Drive yourself	117 (71.3)	33 (68.8)	84 (72.4)	
Depend on others	47 (28.7)	15 (31.3)	32 (27.6)	
Able to make appointments reliably, n (%)[a]	146 (89.6)	36 (75)	110 (95.7)	<.001
Had to cancel due to transportation issues, n (%)[b]	36 (22.2)	21 (45.7)	15 (12.9)	<.001
Internet connection, n (%)				<.001
Dial-up	1 (0.6)	0 (0)	1 (0.9)	
Broadband	122 (74.4)	26 (54.2)	96 (82.8)	
Satellite	7 (4.3)	4 (8.3)	3 (2.6)	
Other	1 (0.6)	0 (0)	1 (0.9)	
I do not know	33 (20.1)	18 (37.5)	15 (12.9)	
Owns a mobile phone, n (%)	162 (98.8)	48 (100)	114 (98.3)	>.99
Device with camera, n (%)[b]	154 (95.1)	43 (89.6)	111 (97.4)	.05
Device used for telehealth, n (%)				<.001
Computer	50 (30.5)	25 (52.1)	25 (21.6)	
Smartphone	70 (42.7)	14 (29.2)	56 (48.3)	
Tablet	44 (26.8)	9 (18.8)	35 (30.2)	
Technical difficulties, n (%)	65 (39.6)	36 (75)	29 (25)	<.001
How often uses a computer, n (%)				<.001
Daily	92 (56.1)	14 (29.2)	78 (67.2)	
Weekly	54 (32.9)	23 (47.9)	31 (26.7)	
Monthly	18 (11)	11 (22.9)	7 (6)	
Comfort with device, n (%)				<.001
Very uncomfortable	3 (1.8)	2 (4.2)	1 (0.9)	
Uncomfortable	33 (20.1)	24 (50)	9 (7.8)	
Neutral	25 (15.2)	12 (25)	13 (11.2)	
Comfortable	47 (28.7)	6 (12.5)	41 (35.3)	
Very comfortable	56 (34.1)	4 (8.3)	52 (44.8)	
Aware of myLoyola portal, n (%)[c]	148 (90.2)	40 (83.3)	108 (93.1)	.08
Communication with liver physician, n (%)[c]				<.001
Phone	87 (54)	37 (78.7)	50 (43.9)	
Email	49 (30.4)	0 (0)	49 (43)	
In person	25 (15.5)	10 (21.3)	15 (13.2)	
Communication with primary care provider, n (%)[a]				<.001

XSL•FO
RenderX

Characteristics	Overall (n=164)	Rural (n=48)	Urban (n=116)	*P* value
Phone	84 (51.5)	37 (77.1)	47 (40.9)	
Email	53 (32.5)	0 (0)	53 (46.1)	
In person	26 (16)	11 (22.9)	15 (13)	
Preferences for reminders, n (%)				<.001
Email	55 (33.5)	2 (4.2)	53 (45.7)	
Telephone	109 (66.5)	46 (95.8)	63 (54.3)	
Interested in new resources for telemedicine, n (%)[a]	131 (80.4)	37 (77.1)	94 (81.7)	.50
Easy to download app, n (%)	28 (17.1)	8 (16.7)	20 (17.2)	.93
How-to guide for accessing visits, n (%)	86 (52.4)	34 (70.8)	52 (44.8)	.002
App for symptoms to aid with medications, n (%)	64 (39)	8 (16.7)	56 (48.3)	<.001
Email reminders for myLoyola, n (%)	8 (4.9)	2 (4.2)	6 (5.2)	>.99
Important to have regular visits with liver physician, n (%)[a]	161 (98.8)	47 (100)	114 (98.3)	>.99
Understand more during virtual or in-person visits, n (%)				<.001
In person	59 (36)	28 (58.3)	31 (26.7)	
Virtual	3 (1.8)	0 (0)	3 (2.6)	
Equally between them	102 (62.2)	20 (41.7)	82 (70.7)	
Satisfaction with visit, n (%)				.97
Very unsatisfied	0 (0)	0 (0)	0 (0)	
Unsatisfied	2 (1.2)	0 (0)	2 (1.7)	
Neutral	3 (1.8)	1 (2.1)	2 (1.7)	
Satisfied	41 (25)	13 (27.1)	28 (24.1)	
Very satisfied	118 (72)	34 (70.8)	84 (72.4)	
Preference for visit type, n (%)				<.001
In person	88 (53.7)	39 (81.3)	49 (42.2)	
Virtual	76 (46.3)	9 (18.8)	67 (57.8)	
Receive same care during telehealth visit, n (%)[a]	126 (77.3)	31 (64.6)	95 (82.6)	.01
More satisfied with telehealth or in-person visits, n (%)				<.001
Telehealth	82 (50)	10 (20.8)	72 (62.1)	
In person	82 (50)	38 (79.2)	44 (37.9)	
Willing to have more telehealth visits in future, n (%)	138 (84.1)	38 (79.2)	100 (86.2)	.26

[a]Data were available for 163 participants (overall).

[b]Data were available for 162 participants (overall).

[c]Data were available for 161 participants (overall).

Multivariate analysis showed that 54% (88/164) of all participants reported a preference for in-person visits and 46% (76/164) for virtual visits. Furthermore, 81% (39/48) of rural patients preferred in-person visits compared with only 42% (49/116) of urban patients (*P*<.001). More urban patients were satisfied with telehealth visits compared with rural patients (72/116, 62% vs 10/48, 20%; *P*<.001). Both rural and urban patients had predominantly Medicare insurance (20/32, 63% vs 67/129, 51%); however, the proportion of Medicaid patients was greater in the urban population (26/129, 20% vs 4/32, 13%).

XSL•FO
RenderX

Patients with Medicaid insurance preferred virtual visits to in-person visits (19/30, 63% vs 11/30, 37%, *P*=.03). Even with shorter travel times and fewer barriers to in-person visits, urban patients were more likely to prefer virtual visits compared with rural patients (OR 5.93, 95% CI 2.73 - 14.1) and were more satisfied with telehealth visits compared with in-person visits (OR 6.22, 95%, CI 2.91 - 14.3; Table 3). Other factors associated with patient satisfaction with telehealth visits included age (OR 0.96, 95% CI 0.93 - 0.99) and comfort with their device (OR 6.08, 95% CI 3.75 - 10.9). However, after adjusting for age, sex, location, and comfort with the device in a multivariable logistic regression, comfort with the device was the only significant predictor of patient satisfaction with telehealth visits compared with in-person visits (adjusted OR 5.80, 95% CI 3.30 - 11.3; Table 3).

Table . Association of patient characteristics with telehealth satisfaction and preference.

	Prefers virtual visits, OR[a] (95% CI)		More satisfied with telehealth than in person, OR (95% CI)		Satisfaction with virtual visit, OR (95% CI)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Age (y)	0.96 (0.92 - 0.99)	0.99 (0.95 - 1.04)	0.96 (0.93 - 0.99)	1.01 (0.96 - 1.05)	1.00 (0.97 - 1.04)	1.02 (0.98 - 1.06)
Sex						
Female	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)
Male	2.53 (1.35 - 4.82)	1.77 (0.77 - 4.09)	2.72 (1.46 - 5.18)	2.04 (0.87 - 4.88)	3.38 (1.68 - 7.08)	2.84 (1.38 - 6.05)
Comfort with device	5.36 (3.37 - 9.33)	4.78 (2.81 - 8.88)	6.08 (3.75 - 10.9)	5.80 (3.30 - 11.3)	1.39 (1.05 - 1.86)	1.52 (1.03 - 2.25)
Location						
Rural	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)	1 (reference)
Urban	5.93 (2.73 - 14.1)	1.18 (0.36 - 3.79)	6.22 (2.91 - 14.3)	1.09 (0.32 - 3.51)	1.06 (0.49 - 2.18)	0.60 (0.23 - 1.54)

[a]OR: odds ratio.

Discussion

Rural patients in our study reported more difficulty attending in-person visits due to double the transportation time and were more likely to cancel these appointments due to transportation issues. Despite these challenges, rural patients reported higher satisfaction with in-person visits compared with urban patients. In contrast, rural patients experienced more technical difficulties during telehealth visits, reported less comfort with their devices, and used their smartphones less frequently. Notably, comfort with device use emerged as the strongest predictor of telehealth satisfaction, suggesting that lower technological literacy may underlie the lower preference for telehealth observed among rural patients.

Telehealth use has expanded rapidly, particularly during the post–COVID-19 pandemic era, due to its advantages in convenience, access, and timeliness [3,11,12]. Prior studies have demonstrated high satisfaction among both patients and health care providers, supporting telehealth as a viable alternative to in-person care [11,13,14]. Within the field of gastroenterology and hepatology, telehealth has been successfully integrated into inflammatory bowel disease clinics, with associated improvements in patient satisfaction and disease-specific quality of life [15]. In hepatology specifically, telehealth interventions have also been used to help support sleep, stress management, and nutrition [16-18]. Within academic gastroenterology and hepatology practices, factors associated with higher satisfaction have included younger age, race, and type of visit [14]. However, broader socioeconomic factors are increasingly recognized as important contributors [19,20]. Prior reviews have demonstrated that income and education level are key determinants of telehealth accessibility and are closely linked to lower electronic health literacy [3,21,22]. Consistent with this literature, our study found that preference for virtual visits was associated with younger age, urban location, smartphone or tablet use, and greater comfort with technology. We also observed significantly lower telehealth satisfaction among rural patients, which appeared to be primarily driven by reduced comfort with devices and lower technology literacy. Although age alone was not independently associated with visit-type preference in multivariable analysis (OR 0.97; *P*=.20), patients with Medicaid insurance, used as a proxy for income, were more likely to prefer virtual visits (*P*=.03), highlighting the complex interaction between socioeconomic factors and telehealth preferences.

Telehealth has been shown to increase access to care through methods of improving availability and affordability [23]. However, persistent challenges remain, particularly related to infrastructure limitations such as broadband access and gaps in technology literacy. Our results show that rural patients are disproportionately affected by both barriers. While the Infrastructure Act of 2021 promised US $65 billion toward improving broadband access, addressing connectivity alone may be insufficient. Targeted efforts to improve technology literacy are likely necessary to fully realize the benefits of telehealth in rural populations.

To better understand potential solutions, our survey asked patients to identify resources that might help them with their visits in the future. Proposed interventions included a step-by-step “how-to” guide for accessing visits, a loaned tablet with a preloaded telehealth app, and a hospital-based app for tracking weight or medication management. A total of 80%

XSL•FO
RenderX

(131/164) of patients expressed interest in additional resources, with the most endorsed intervention being a “how-to” guide focusing on accessing visits. These findings suggest that relatively low-cost educational interventions may meaningfully improve telehealth satisfaction. Future studies should focus on implementing such interventions, particularly in rural communities, to assess whether targeted technology training and assistance can help improve health care access and patient satisfaction.

The Veterans Affairs (VA) health care system was an early adopter of telehealth and provides a useful model. Through initiatives led by the VA Office of Connected Care, established in 2015, the VA has addressed common telehealth barriers by providing loaned devices and offering structured support to improve technology literacy [24]. These investments in telehealth have allowed the VA to reach a broader patient population [25]. Whether other health care systems and private insurers will adopt similar programs to address these barriers remains an important area for study.

Several limitations should be considered when interpreting our findings. This was a single-center study conducted during the COVID-19 pandemic, and some patients may have favored telehealth visits due to concerns about infection risk with in-person visits. Although our survey was not validated externally, it was based on prior studies and reviewed by our hepatology colleagues and statistician to ensure appropriate content. Visit types were not stratified by routine follow-ups and acute visits, which may have played a role in satisfaction. In addition, patient-physician relationships may have confounded satisfaction ratings, as patients’ comfort with a physician may have biased their opinion on visit type. Furthermore, given the complexity of our patient population, those who answered the survey may have had higher technology literacy at baseline, potentially limiting generalizability to the broader hepatology patient population. Finally, our study did not address health care provider burden or perceived quality of care [26]. Health care provider concerns regarding limitations of physical examinations during telehealth visits were not evaluated, nor were objective clinical outcomes assessed. Future studies should incorporate both patient- and health care provider–reported outcomes to better characterize the impact of telehealth on quality of care.

In conclusion, this study demonstrates lower satisfaction with telehealth among rural patients compared with urban patients. Our findings suggest that efforts to improve broadband access, provide appropriate devices, and enhance telehealth literacy may increase satisfaction and reduce barriers to care, including travel time and associated costs, for patients in rural areas.

Funding

The authors declared no financial support was received for this work.

Data Availability

The data that support the findings of this study are not publicly available due to privacy concerns and compliance with HIPAA (Health Insurance Portability and Accountability Act) regulations. Data are available upon reasonable request from the corresponding author.

Authors' Contributions

CH and JR conceived of the study design and methodology and developed the study protocol. CH and KK were responsible for data collection and sample preparation. JR oversaw data quality and ensured data integrity. CJ conducted the statistical analysis, interpreted the results, and generated the tables. CH and JR contributed to the interpretation of the results. CH drafted the manuscript. JR critically revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental materials including survey questionnaire and other data tables.
[DOCX File, 31 KB - aging_v9i1e85018_app1.docx]

References

1. Ramaswamy A, Yu M, Drangsholt S, et al. Patient satisfaction with telemedicine during the COVID-19 pandemic: retrospective cohort study. J Med Internet Res 2020 Sep 9;22(9):e20786. [doi: 10.2196/20786] [Medline: 32810841]
2. Tenforde AS, Borgstrom H, Polich G, et al. Outpatient physical, occupational, and speech therapy synchronous telemedicine: a survey study of patient satisfaction with virtual visits during the COVID-19 pandemic. Am J Phys Med Rehabil 2020 Nov;99(11):977-981. [doi: 10.1097/PHM.0000000000001571] [Medline: 32804713]
3. Chen K, Lodaria K, Jackson HB. Patient satisfaction with telehealth versus in-person visits during COVID-19 at a large, public healthcare system. J Eval Clin Pract 2022 Dec;28(6):986-990. [doi: 10.1111/jep.13770] [Medline: 36148479]

XSL•FO
RenderX

4. Triana AJ, Gusdorf RE, Shah KP, Horst SN. Technology literacy as a barrier to telehealth during COVID-19. Telemed J E Health 2020 Sep;26(9):1118-1119. [doi: 10.1089/tmj.2020.0155] [Medline: 32429770]
5. Computer and internet use in the United States: 2018. : United States Census Bureau; 2021 URL: www.census.gov/newsroom/press-releases/2021/computer-internet-use.html [accessed 2026-05-29]
6. Increase the proportion of adults with broadband internet — HC/HIT-05. Office of Disease Prevention and Health Promotion, Office of the Assistant Secretary for Health, Office of the Secretary, U.S. Department of Health and Human Services. URL: https://odphp.health.gov/healthypeople/objectives-and-data/browse-objectives/neighborhood-and-built-environment/increase-proportion-adults-broadband-internet-hchit-05 [accessed 2024-06-04]
7. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
8. Roque NA, Boot WR. A new tool for assessing mobile device proficiency in older adults: the Mobile Device Proficiency Questionnaire. J Appl Gerontol 2018 Feb;37(2):131-156. [doi: 10.1177/0733464816642582] [Medline: 27255686]
9. Parmanto B, Lewis AN Jr, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). Int J Telerehabil 2016;8(1):3-10. [doi: 10.5195/ijt.2016.6196] [Medline: 27563386]
10. Rural-Urban Commuting Area Codes. U.S. Department of Agriculture. URL: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes [accessed 2024-06-04]
11. Bensted K, Kim C, Freiman J, Hall M, Zekry A. Gastroenterology hospital outpatients report high rates of satisfaction with a telehealth model of care. J Gastroenterol Hepatol 2022 Jan;37(1):63-68. [doi: 10.1111/jgh.15663] [Medline: 34402105]
12. Capusan KY, Fenster T. Patient satisfaction with telehealth during the COVID-19 pandemic in a pediatric pulmonary clinic. J Pediatr Health Care 2021;35(6):587-591. [doi: 10.1016/j.pedhc.2021.07.014] [Medline: 34417077]
13. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 3;7(8):e016242. [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
14. Serper M, Nunes F, Ahmad N, Roberts D, Metz DC, Mehta SJ. Positive early patient and clinician experience with telemedicine in an academic gastroenterology practice during the COVID-19 pandemic. Gastroenterology 2020 Oct;159(4):1589-1591. [doi: 10.1053/j.gastro.2020.06.034] [Medline: 32565015]
15. Cross RK, Langenberg P, Regueiro M, et al. A randomized controlled trial of TELEmedicine for patients with inflammatory bowel disease (TELE-IBD). Am J Gastroenterol 2019 Mar;114(3):472-482. [doi: 10.1038/s41395-018-0272-8] [Medline: 30410041]
16. Mauro E, Marciano S, Torres MC, Roca JD, Novillo AL, Gadano A. Telemedicine improves access to hepatology consultation with high patient satisfaction. J Clin Exp Hepatol 2020;10(6):555-562. [doi: 10.1016/j.jceh.2020.04.017] [Medline: 33311892]
17. Evon DM, Liu C, Rollason C, et al. Group-based telehealth stress management-lifestyle interventions are feasible for patients with chronic liver disease. Hepatol Commun 2025;9(12):e0849. [doi: 10.1097/HC9.0000000000000849] [Medline: 41264908]
18. Chang WC, Wu MS, Chen YH, Yang SC. Effects of integrative telehealth-based nutrition care with and without oral nutritional supplements in patients with liver and colorectal cancer: a randomized controlled trial. Nutrition 2025 Jul;135:112768. [doi: 10.1016/j.nut.2025.112768] [Medline: 40233464]
19. Siegel CA. Transforming gastroenterology care with telemedicine. Gastroenterology 2017 Apr;152(5):958-963. [doi: 10.1053/j.gastro.2017.01.048] [Medline: 28192101]
20. Arora S, Thornton K. Novel models of hepatitis C virus care delivery: telemedicine, Project ECHO, and integrative care. Clin Liver Dis (Hoboken) 2020 Jul;16(1):5-7. [doi: 10.1002/cld.912] [Medline: 32714515]
21. Xu RH, Zhou LM, Wong EL, Wang D. The association between patients' eHealth literacy and satisfaction with shared decision-making and well-being: multicenter cross-sectional study. J Med Internet Res 2021 Sep 24;23(9):e26721. [doi: 10.2196/26721] [Medline: 34559062]
22. Campanozzi LL, Gibelli F, Bailo P, Nittari G, Sirignano A, Ricci G. The role of digital literacy in achieving health equity in the third millennium society: a literature review. Front Public Health 2023;11:1109323. [doi: 10.3389/fpubh.2023.1109323] [Medline: 36891330]
23. Anawade PA, Sharma D, Gahane S. A comprehensive review on exploring the impact of telemedicine on healthcare accessibility. Cureus 2024 Mar;16(3):e55996. [doi: 10.7759/cureus.55996] [Medline: 38618307]
24. Heyworth L, Shah N, Galpin K. 20 years of telehealth in the Veterans Health Administration: taking stock of our past and charting our future. J Gen Intern Med 2024 Feb;39(Suppl 1):5-8. [doi: 10.1007/s11606-024-08617-w] [Medline: 38378981]
25. Myers US, Birks A, Grubaugh AL, Axon RN. Flattening the curve by getting ahead of it: how the VA Healthcare System is leveraging telehealth to provide continued access to care for rural veterans. J Rural Health 2021 Jan;37(1):194-196. [doi: 10.1111/jrh.12449] [Medline: 32282955]
26. Bloomberg L, Hong P, Hepburn C, et al. Changes in provider responsibilities and associated outcomes for cirrhotic patients with telehealth: a single-center, retrospective study. J Telemed Telecare 2026 Feb;32(2):112-117. [doi: 10.1177/1357633X251323185] [Medline: 40080886]

XSL•FO
RenderX

Abbreviations

OR: odds ratio
REDCap: Research Electronic Data Capture
VA: Veterans Affairs

Edited by J Wang; submitted 29.Sep.2025; peer-reviewed by D Poddighe, RD Davidson, VJ Powell; revised version received 02.Mar.2026; accepted 11.May.2026; published 24.Jun.2026.

Please cite as:
Hepburn C, Krawczyk K, Joyce C, Rubin J
Patient Satisfaction With Telehealth Visits in Rural Compared With Urban Communities: Single-Center Study
JMIR Aging 2026;9:e85018
URL: https://aging.jmir.org/2026/1/e85018
doi:10.2196/85018

© Corrin Hepburn, Karolina Krawczyk, Cara Joyce, Jonah Rubin. Originally published in JMIR Aging (https://aging.jmir.org), 24.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Self-Efficacy, Satisfaction With the Daily Life Changes Stemming From Digital Transformation, and the Moderating Effect of Social Capital in Middle-Aged and Older Adults: Cross-Sectional Survey Study

Jeongeun Choi[1], RN, BSN; Hyeonmi Cho[2], RN, PhD; Hyangkyu Lee[3], RN, PhD

[1]Mo-Im Kim Nursing Research Institute, College of Nursing and Brain Korea 21 FOUR Project, Yonsei University, Seoul, Republic of Korea

[2]College of Nursing, Research Institute of AI and Nursing Science, Gachon University, Incheon, Gyeonggi-do, Republic of Korea

[3]Mo-Im Kim Nursing Research Institute, College of Nursing, Institute for Innovation in Digital Healthcare, Yonsei University, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Corresponding Author:
Hyangkyu Lee, RN, PhD
Mo-Im Kim Nursing Research Institute, College of Nursing, Institute for Innovation in Digital Healthcare, Yonsei University, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: The swift pace of digital transformation has heightened individuals' dependence on digital technologies. This makes it imperative to explore how individual factors such as digital self-efficacy and social capital affect satisfaction with the daily life changes stemming from digital transformation. This investigation is particularly essential among middle-aged and older adults because they encounter numerous challenges and opportunities from digital transformation.

Objective: This study examined the moderating effect of social capital on the relationship between digital self-efficacy and satisfaction with the daily life changes stemming from digital transformation among middle-aged and older adults.

Methods: This study used data from the 2022 Digital Divide Survey conducted by the National Information Society Agency in South Korea. The sample included 4155 individuals aged 40 years or older, of whom 71.8% (n=2985) were classified as middle-aged (40 - 64 y) and 28.2% (n=1170) as older adults (65 - 96 y). Data were analyzed using multiple linear regression models with IBM SPSS (version 27.0) and the PROCESS macro.

Results: Compared with the middle-aged group, the older adult group showed lower levels of digital self-efficacy, social capital, and satisfaction with the daily life changes stemming from digital transformation. The interaction term between digital self-efficacy and social capital was significant in the middle-aged group (B=−0.11, *P*=.002) and the older adult group (B=−0.24, *P*<.001). There was a positive association between digital self-efficacy and satisfaction in both groups, and this association was stronger among those with low social capital. Furthermore, this effect was more pronounced in the older adult group (B=0.71, *P*<.001 for low and B=0.47, *P*<.001 for high social capital) than in the middle-aged group (B=0.55, *P*<.001 for low and B=0.46, *P*<.001 for high social capital). Education level and digital competence were positively associated with satisfaction in both age groups; higher household income was positively associated with satisfaction only in the middle-aged group.

Conclusions: Individuals with low social capital showed a stronger association between digital self-efficacy and satisfaction with the daily life changes stemming from digital transformation. Enhancing digital self-efficacy among older adults, especially those with low levels of social capital, may improve their satisfaction with the daily life changes stemming from digital transformation.

(JMIR Aging 2026;9:e79845) doi:10.2196/79845

KEYWORDS

digital transformation; self-efficacy; social capital; accessibility; digital integration; digital divide

Introduction

Background

The COVID-19 pandemic expedited digital transformation, increasing the adoption of digital technology-based public and private services and accelerating innovation in information and communication technologies (ICTs) [1,2]. As of 2023, approximately 5.16 billion people (constituting 64.4% of the global population) were using the internet, and approximately 5.44 billion (or 68% of the global population) were using mobile phones [3]. South Korea is at the forefront of this global trend.

XSL•FO
RenderX

According to the Organisation for Economic Co-operation and Development (OECD) Digital Government Index published in 2025, South Korea ranked first overall, with a composite score of 0.94, substantially higher than the OECD average of 0.61 [4]. This reflects an environment in which digital platforms are no longer optional but function as essential infrastructure for daily life. This pervasive digitalization has fostered the development of a well-connected society in South Korea. However, this rapid digitalization coincides with an unprecedented pace of population aging.

The global proportion of individuals aged 65 years and older is projected to rise to 1 in 6 by 2050 [5]. Notably, South Korea is experiencing the world's most rapid pace of population aging and is expected to become one of the world's most aged societies by 2050, with older adults comprising approximately 40% of the total population [5]. In this context, where rapid digital transformation and population aging are occurring simultaneously, the digital divide extends beyond access to resources to include disparities in digital skills and perceptions of digital technologies [6].

The digital divide refers to the disparity between individuals with access to modern ICTs and those without such access [7]. It contributes to digital exclusion [2], particularly in societies where technology and the internet govern various aspects of daily activities [8,9]. This technological shift brings about changes in daily life that individuals perceive as either beneficial or burdensome. When individuals are digitally excluded, barriers to accessing information and services, as well as maintaining social participation, can undermine everyday autonomy, thereby reducing satisfaction with the daily life changes stemming from digital transformation [9].

While internet adoption has accelerated across all ages in recent years, middle-aged and older adults continue to account for a large proportion of those experiencing the digital divide [10,11]. This population is often less familiar with digital technologies, which makes them more vulnerable to the digital divide and increases their risk of digital exclusion [2,8-12]. Given that digital exclusion can detrimentally affect the quality of life and social participation [13,14], it is essential to address digital divide concerns among middle-aged and older adults. However, the digital divide is not a uniform phenomenon, as its underlying mechanisms and consequences often vary across life stages. From a life-course perspective, middle-aged adults often face strong demands to use digital technologies for work, family responsibilities, and social roles, whereas older adults may encounter more complex barriers due to retirement-related transitions, changes in social ties, and health or functional limitations [15,16].

Previous research suggests that digital technology adoption processes differ between middle-aged and older adults, particularly in how key adaptive resources such as digital self-efficacy and social capital are developed and used. Middle-aged adults, who have higher levels of participation in the workforce, tend to strengthen their digital self-efficacy through task-related experiences, training, and repeated ICT use in both occupational and everyday contexts [17]. In contrast, older adults are more likely to encounter digital technologies later in life and tend to rely more on relational support, such as encouragement, demonstration, and guidance from family members or social networks during the adaptation process [18]. The use of social capital in digital environments also differs across age groups. Middle-aged adults often use broader and more diverse networks, including professional and extended social ties, to access information and resources. In contrast, older adults tend to depend on close ties, such as family members and familiar contacts, for instrumental assistance and emotional support [19,20]. Taken together, structural differences in the formation of digital self-efficacy and the usage of social capital between middle-aged and older adults support distinguishing these groups to capture heterogeneity in digital adaptation processes across life stages.

The impact of digital transformation on daily life extends beyond technology use and reflects individuals' experiences of changes in the digital environment. Recent studies have shown that digital transformation is associated with a range of subjective life outcomes, including life satisfaction [21], quality of life [22], subjective well-being [23], and overall well-being [24]. These indicators capture general life outcomes and assess the effects of digital transformation only indirectly, without reflecting individuals' evaluations of specific changes in their daily lives. In this study, the dependent variable was measured as "satisfaction with the daily life changes stemming from digital transformation." This variable reflects how individuals evaluate the impact of accelerated digital changes across key domains of everyday life, including leisure opportunities, access to information and knowledge, social connection and exchange, and work and study efficiency.

This conceptualization is consistent with the third-level digital divide perspective, which emphasizes that digital inequalities are not limited to differences in internet access (the first-level digital divide) and digital skills (the second-level digital divide) [25], and are ultimately reflected in differences in the tangible benefits individuals derive from using technologies [26]. Previous studies further demonstrate that digital transformation is closely associated with key domains of everyday life. Digital transformation facilitates access to information and knowledge by enhancing opportunities for learning and information seeking [27], promotes social connection and exchange by reducing social isolation and supporting social well-being [14,28], and contributes to improved learning experiences and work-related productivity in educational and occupational contexts [29]. Furthermore, digital transformation has extended into leisure, reshaping recreational activities and influencing individuals' daily experiences [30,31]. Collectively, the dependent variable captures a theoretically grounded measure of perceived satisfaction with digital transformation–related changes in everyday life as a key outcome of digital engagement.

Digital self-efficacy refers to an individual's belief in their capability to use digital systems effectively [32]. It promotes the long-term use of digital devices, thereby reducing the digital divide and boosting the quality of life [33-35]. According to the self-efficacy theory, self-efficacy is one's belief in their ability to plan and successfully perform the activities needed to achieve a specific goal [36]. Self-efficacy acts as a resource in task attainment, as individuals regulate their effort and

XSL•FO
RenderX

perseverance based on their self-efficacy [37]. Insufficient digital self-efficacy can hamper the ability of middle-aged adults and older adults to effectively use technologies to address later-life needs, such as financial management, health care, and retirement planning [8,38]. Addressing the barriers to technology use and enhancing digital self-efficacy may improve their engagement and proficiency with digital devices, which can, in turn, aid them in achieving successful aging [33,39].

Social capital is defined as networks and relationships among individuals that are underpinned by the norms of reciprocity and trust. It encompasses two key functions: (1) bonding, which creates strong emotional connections; and (2) bridging, which produces weaker but broader connections [40]. As individuals transition from middle age to older adulthood, events such as retirement, the loss of spouses and close connections, limited educational opportunities, and mobility limitations diminish social capital. Specifically, there is a gradual reduction in one's social networks and the capacity to use social resources [41-43]. Social capital is a significant contributor to the use of digital devices and the maintenance of online engagement [38]. Additionally, it can function as a contextual resource that moderates the relationship between digital self-efficacy and satisfaction with the daily life changes stemming from digital transformation [44]. When social capital is high, individuals can access support that facilitates digital engagement and reduces reliance on their own digital self-efficacy, whereas when social capital is low, limited support increases dependence on digital self-efficacy for digital engagement [45]. Accordingly, the effect of digital self-efficacy on satisfaction with the daily life changes stemming from digital transformation may vary depending on the level of social capital.

According to the conservation of resources (COR) theory [46], individuals experiencing resource loss tend to protect and preserve their remaining resources. In technology-dependent societies, individuals with low social capital have limited access to external support, increasing their vulnerability to resource loss. In this context, digital self-efficacy may function as a key internal resource that compensates for the lack of external support. Specifically, higher digital self-efficacy may enable individuals to more effectively navigate digital environments despite constrained support, thereby facilitating a more adaptive appraisal of digital transformation as manageable rather than threatening [47]. Such appraisal may reduce maladaptive responses and enhance satisfaction with the daily life changes stemming from digital transformation. Taken together, these considerations suggest that digital self-efficacy may serve as a compensatory resource among middle-aged and older adults with limited social capital, with important implications for supporting adaptation to evolving digital environments and reducing the digital divide.

Purpose

This study aimed to evaluate the relationship between digital self-efficacy, social capital, and satisfaction with the daily life changes stemming from digital transformation. Specifically, it focused on how social capital moderates the relationship between digital self-efficacy and satisfaction with the daily life changes stemming from digital transformation among both middle-aged and older adults.

Methods

Data and Study Population

We used cross-sectional data from the 2022 Digital Divide Survey, an annual nationwide survey conducted by the Ministry of Science and ICT and the National Information Society Agency in South Korea. In this study, analyses were based on the general population subsample aged 40 years or older (n=4155). Adults aged 40 - 64 years were classified as the middle-aged group (n=2985, 71.8%), and those aged 65 - 96 years as the older adult group (n=1170, 28.2%), based on the official life-stage age classification used in Korea's national administrative statistics [48]. In 2022, the survey was conducted from September to December using standardized face-to-face interviews and used a stratified probability-proportional-to-size sampling design to ensure national representativeness. The survey assessed the effectiveness of policies aimed at reducing the digital divide and included a total of 15,000 respondents from the general population, along with additional samples from vulnerable groups such as older adults, individuals with disabilities, individuals in low-income households, immigrant women in international marriages, North Korean defectors, and individuals engaged in agriculture or fishing. The official methodology report of this survey did not provide an overall response rate; however, the analytic dataset contained no missing values for study variables.

Variables

Dependent Variable

Satisfaction with the daily life changes stemming from digital transformation was determined using 4 items rated on a 5-point Likert scale from 1 ("strongly disagree") to 5 ("strongly agree"). The items were as follows: (1) "The digital transformation has increased opportunities for leisure activities and made them more enjoyable (for instance, through online content and mobile applications)"; (2) "The digital transformation has enabled quicker access to new information or knowledge (for instance, through remote learning and online content platforms)"; (3) "The digital transformation has increased opportunities to engage with new people and exchange opinions"; and (4) "Efficiency in work or academic pursuits has improved due to practices like remote work and mobile learning." The total score was calculated as the mean of 4 items, with higher scores indicating greater satisfaction levels [49]. In this study, the scale exhibited excellent internal consistency, with a Cronbach α of 0.90.

Construct validity was examined using confirmatory factor analysis (CFA) in the total sample and separately in middle-aged and older adults, given that this 4-item measure was newly added to the 2022 survey. Model fit was evaluated using the comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root-mean-square residual (SRMR). The chi-square statistic was excluded from model fit evaluation because it is overly sensitive to large sample sizes [50]. Acceptable model fit was indicated by CFI and TLI≥0.90 and SRMR<0.08 [51],

XSL•FO
RenderX

with more stringent cutoffs of CFI and TLI≥0.95 suggested [52]. In addition to CFI, TLI, and SRMR, root-mean-square error of approximation (RMSEA)<0.08 is commonly used as a criterion for acceptable model fit; however, RMSEA tends to be upwardly biased in models with small dfs [53]. Since our measurement model for the 4-item scale has a small df (df=2), model fit was primarily evaluated using CFI, TLI, and SRMR, which are more robust in such cases [53]. Construct reliability was assessed using composite reliability (CR; ≥0.70) [54], and convergent validity was assessed using average variance extracted (AVE; ≥0.50) [55].

Independent Variable

Digital self-efficacy was assessed using 4 items rated on a 4-point scale ranging from 1 (“strongly disagree”) to 4 (“strongly agree”). These 4 items have been widely used in past studies [56-58]. The items are as follows: (1) “I am confident that I can learn about digital devices”; (2) “I am confident in using digital devices”; (3) “I can quickly learn how to use new digital devices”; and (4) “I want to use digital devices more frequently.” The digital self-efficacy score was calculated based on the average score on the 4 items, with higher scores reflecting higher digital self-efficacy. In this study, the scale demonstrated excellent internal consistency, with a Cronbach α of 0.88.

Moderating Variable

Social capital was measured with an abbreviated version of Williams’s social-capital scale [59] developed by the agency that conducted the 2022 Digital Divide Survey. It comprises 10 items—5 items each for bonding and bridging social capital. The items for bonding social capital were (1) “I know someone who can help me solve my problems”; (2) “I know someone who I can ask for advice when I have to make important decisions”; (3) “I know someone who I can talk to about my private matters”; (4) “I know someone who I can trust with important work”; and (5) “I know someone who can help me fight injustice.” The items for bridging social capital were (1) “I feel connected to a bigger world as I interact with people online and offline.”; (2) “I feel that everyone in the world is connected as I interact with people online and offline.”; (3) “I am willing to spend time on community activities”; (4) “Online and offline interactions allow me to communicate with new people”; and (5) “Online and offline interactions allow me to meet new people.” The items are rated on a 4-point scale that ranges from 1 (“strongly disagree”) to 4 (“strongly agree”). We used a combined social capital score (bonding plus bridging) to capture overall accessible social resources, consistent with previous research [44,58,60]. This operationalization aligned with our focus on moderation by overall social support and information pathways. The average of the scores on the 10 items indicates the total social capital score. Higher scores indicate greater social capital. This scale has been extensively adopted in previous studies [44,58,60]. In this study, the scale had a Cronbach α of 0.84, indicating great internal consistency.

Covariates

The general characteristics of the study population and digital competence were included as covariates. The general characteristics were age, sex, education level, living arrangements, the presence of disability, monthly household income, self-rated health, and digital competence. These covariates were selected based on previous studies [32,61,62]. The exchange rate in 2022 was KRW 1=US $0.000774, based on the average annual exchange rate reported by the World Bank [63]. Self-rated health was captured with a single question on how satisfied participants felt about their physical and mental health, with responses ranging from 1 (“not satisfied at all”) to 4 (“very satisfied”) [64]. Digital competence was measured with 14 items assessing personal computer and mobile phone use skills, each rated from 1 to 4, where higher scores indicated greater digital competence [65]. In this study, the scale exhibited a Cronbach α of 0.96, suggesting excellent internal consistency.

Ethical Considerations

This was a secondary analysis study that used publicly available data without any personally identifiable information. Therefore, this study was exempt from review by the Institutional Review Board (IRB 4-2023-1516).

Statistical Analysis

We compared the general characteristics of the study population and the study variables between middle-aged and older adults using chi-square tests and 2-tailed independent *t* tests. To examine the associations between digital self-efficacy, social capital, and satisfaction with the daily life changes stemming from digital transformation, we used multiple linear regression analyses. Three separate regression models were developed for the total sample, for middle-aged adults, and for older adults. All regression models were adjusted for covariates. The continuous variables were mean-centered before the regression models were constructed, and no variables were standardized. The moderation effect of social capital was assessed using an interaction term (digital self-efficacy×social capital) in the models. The PROCESS macro in IBM SPSS Statistics was used to determine significant interaction effects. Unstandardized simple slopes were calculated at 1 SD above (+1 SD) and below (−1 SD) the mean score for social capital.

Considering the potential heterogeneity across the dependent variable’s subdimensions, we performed an additional analysis using a multiple indicators multiple causes (MIMIC) model, a specific form of structural equation modeling (SEM) [66]. In this model, the dependent variable was specified as a latent factor indicated by 4 reflective items, while observed covariates were included as causal predictors [67].

Statistical analyses were performed using SPSS Statistics for Windows (version 27.0). The CFA and SEM models were estimated in R (version 4.5.1; R Core Team) using the *lavaan* package, using robust maximum likelihood estimation [68]. We evaluated normality by inspecting absolute skewness values for all continuous variables (with values ≥3 indicating substantial deviation from normality) and visually reviewing histograms and quantile-quantile (Q-Q) plots. All absolute skewness values were <1.5, and no marked deviations from normality were observed in the histograms or Q-Q plots. No missing values were present for any variables included in the analytic sample; therefore, no exclusions or imputations were required. No

XSL•FO
RenderX

multicollinearity issues were observed in any of the models because the variance inflation factor remained under 10.

Results

General Characteristics of the Study Population

The average age of middle-aged adults was 52.42 (SD 6.80) years, while that of older adults was 71.33 (SD 5.05) years (Table 1). The proportion of individuals with high school or higher educational attainment was higher in the middle-aged group (2814/2985, 94.3%) than that in the older adult group (510/1170, 43.6%). The proportion of individuals living with others was lower in the older adult group (927/1170, 79.2%) than that in the middle-aged group (2786/2985, 93.3%). Furthermore, older adults had lower household incomes, poorer self-rated health, and weaker digital competence than middle-aged adults.

Table . General characteristics of the study population (N=4155).

Variable	Total (n=4155)	Middle-aged adults (n=2985, 71.8%)	Older adults (n=1170, 28.2%)	*t* test (*df*) or chi-square test (*df*)	*P* value
Age (in y), mean (SD)	57.74 (10.62)	52.42 (6.80)	71.33 (5.05)	−97.93[a] (4,153)	<.001
Sex, n (%)					
Male	2054 (49.4)	1502 (50.3)	552 (47.2)	3.31[b] (1)	.07
Female	2101 (50.6)	1483 (49.7)	618 (52.8)		
Education level, n (%)					
Middle school or lower	831 (20)	171 (5.7)	660 (56.4)	1349.40[b] (1)	<.001
High school or above	3324 (80)	2814 (94.3)	510 (43.6)		
Living arrangements, n (%)					
Living alone	442 (10.6)	199 (6.7)	243 (20.8)	175.85[b] (1)	<.001
Living with others	3713 (89.4)	2786 (93.3)	927 (79.2)		
Presence of disability, n (%)					
No	4055 (97.6)	2911 (97.5)	1144 (97.8)	0.24[b] (1)	.63
Yes	100 (2.4)	74 (2.5)	26 (2.2)		
Monthly household income (KRW), n (%)[c]					
<4,000,000	1912 (46)	949 (31.8)	963 (82.3)	863.44[b] (1)	<.001
≥4,000,000	2243 (54)	2036 (68.2)	207 (17.7)		
Self-rated health, mean (SD)	2.76 (0.66)	2.88 (0.61)	2.47 (0.69)	17.53[a] (4,153)	<.001
Digital competence, mean (SD)	4.80 (1.65)	5.36 (1.41)	3.37 (1.33)	41.39[a] (4,153)	<.001
Digital self-efficacy, mean (SD)	2.46 (0.72)	2.67 (0.62)	1.93 (0.68)	32.13[a] (4,153)	<.001
Social capital, mean (SD)	2.82 (0.45)	2.89 (0.41)	2.65 (0.50)	14.76[a] (4,153)	<.001
Satisfaction with the daily life changes stemming from digital transformation, mean (SD)	3.01 (0.87)	3.23 (0.75)	2.44 (0.88)	27.17[a] (4,153)	<.001

[a]tested using independent *t* test.

[b]tested using the chi-square test.

[c]A currency exchange rate of KRW 4,000,000=US $3096.81 (KRW 1=US $0.000774; 2022 annual average, World Bank) was applied [63].

XSL•FO
RenderX

Differences in Digital Self-Efficacy, Social Capital, and Satisfaction With the Daily Life Changes Stemming From Digital Transformation Between Middle-Aged and Older Adults

The older adult group (mean 1.93, SD 0.68) exhibited lower levels of digital self-efficacy than the middle-aged group (mean 2.67, SD 0.62; $t_{4,153}$=32.13; P<.001; Table 1). The former (mean 2.65, SD 0.50) demonstrated lower levels of social capital than the latter (mean 2.89, SD 0.41; $t_{4,153}$=14.76; P<.001). The former (mean 2.44, SD 0.88) also reported lower levels of satisfaction with the daily life changes stemming from digital transformation than the latter (mean 3.23, SD 0.75; $t_{4,153}$=27.17; P<.001).

CFA Results for the 4-Item Measure of the Dependent Variable

CFA supported a unidimensional structure of the 4-item measure of satisfaction with the daily life changes stemming from digital transformation in the total sample and both age groups (Table 2). Standardized factor loadings were high (total: 0.82 - 0.86; middle-aged: 0.75 - 0.82; and older adults: 0.84 - 0.88), and fit indices indicated acceptable to excellent fit overall (CFI and TLI≥0.95; SRMR<0.08). Construct reliability and convergent validity were well-supported across groups (CR≥0.7; AVE≥0.5).

XSL•FO
RenderX

Table . Confirmatory factor analysis results for the dependent variable (N=4155).

Sample and item	Factor loading		Goodness-of-fit indicators				
	β[a]	*P* value	CFI[b]	TLI[c]	SRMR[d]	CR[e]	AVE[f]
Total (n=4155)			0.99	0.99	0.01	0.9	0.7
Item 1: Leisure opportunities	0.83	<.001					
Item 2: Access to information and knowledge	0.82	<.001					
Item 3: Social connection and exchange	0.85	<.001					
Item 4: Work and study efficiency	0.86	<.001					
Middle-aged adults (n=2985)			0.99	0.98	0.01	0.87	0.62
Item 1: Leisure opportunities	0.77	<.001					
Item 2: Access to information and knowledge	0.75	<.001					
Item 3: Social connection and exchange	0.80	<.001					
Item 4: Work and study efficiency	0.82	<.001					
Older adults (n=1170)			0.99	0.98	0.01	0.92	0.73
Item 1: Leisure opportunities	0.84	<.001					
Item 2: Access to information and knowledge	0.84	<.001					
Item 3: Social connection and exchange	0.88	<.001					
Item 4: Work and study efficiency	0.87	<.001					

[a]β: standardized regression coefficient.

[b]CFI: comparative fit index.

[c]TLI: Tucker-Lewis index.

[d]SRMR: standardized root-mean-square residual.

[e]CR: composite reliability.

[f]AVE: average variance extracted.

Associations Between Digital Self-Efficacy, Social Capital, and Satisfaction With the Daily Life Changes Stemming From Digital Transformation

Higher levels of digital self-efficacy were associated with higher levels of satisfaction with the daily life changes stemming from digital transformation in the middle-aged group (B=0.50, *P*<.001) and in the older adult group (B=0.59, *P*<.001; Table 3). Similarly, higher social capital was significantly associated with higher satisfaction with the daily life changes stemming from digital transformation among both middle-aged (B=0.28, *P*<.001) and older adults (B=0.12, *P*=.01).

XSL•FO
RenderX

Table . Associations between digital self-efficacy, social capital, and satisfaction with the daily life changes stemming from digital transformation (N=4155).

Variable[a]	Model 1 (total; N=4155)			Model 2 (middle-aged adults; n=2985, 71.8%)			Model 3 (older adults; n=1170, 28.2%)		
	B[b] (SE)	*t* test (*df*)	*P* value	B (SE)	*t* test (*df*)	*P* value	B (SE)	*t* test (*df*)	*P* value
Independent variable									
Digital self-efficacy	0.53 (0.02)	27.87 (4,143)	<.001	0.50 (0.02)	22.59 (2,973)	<.001	0.59 (0.04)	16.15 (1,158)	<.001
Moderation variable									
Social capital	0.22 (0.03)	8.44 (4,143)	<.001	0.28 (0.03)	9.08 (2,973)	<.001	0.12 (0.05)	2.60 (1,158)	.01
Interaction term									
Digital self-efficacy×social capital[c]	−0.11 (0.03)	−4.01 (4,143)	<.001	−0.11 (0.04)	−3.13 (2,973)	.002	−0.24 (0.06)	−4.05 (1,158)	<.001
Control variables[d]									
Age	−0.003 (0.001)	−2.22 (4,143)	.03	−0.002 (0.002)	−0.91 (2,973)	.36	−0.01 (0.01)	−1.65 (1,158)	.10
Sex (ref[e]: male)	0.02 (0.02)	1.04 (4,143)	.30	0.02 (0.02)	1.00 (2,973)	.32	−0.01 (0.04)	−0.32 (1,158)	.75
Education level (ref: middle school or lower)	0.14 (0.03)	4.32 (4,143)	<.001	0.12 (0.05)	2.43 (2,973)	.02	0.09 (0.05)	1.99 (1,158)	.047
Living arrangements (ref: living alone)	0.002 (0.03)	0.06 (4,143)	.96	−0.003 (0.05)	−0.07 (2,973)	.95	−0.04 (0.05)	−0.82 (1,158)	.41
Presence of disability (ref: without disability)	−0.01 (0.06)	−0.21 (4,143)	.83	−0.09 (0.07)	−1.23 (2,973)	.22	0.16 (0.14)	1.16 (1,158)	.25
Monthly household income (ref: <KRW 4,000,000)[f]	0.09 (0.02)	3.66 (4,143)	<.001	0.11 (0.03)	4.12 (2,973)	<.001	0.02 (0.06)	0.37 (1,158)	.71
Self-rated health	−0.03 (0.02)	−1.58 (4,143)	.12	−0.03 (0.02)	−1.56 (2,973)	.12	−0.03 (0.03)	−0.90 (1,158)	.37
Digital competence	0.08 (0.01)	9.40 (4,143)	<.001	0.07 (0.01)	6.51 (2,973)	<.001	0.13 (0.02)	7.02 (1,158)	<.001

[a]All continuous predictors were mean-centered prior to creating the interaction term.

[b]B: unstandardized regression coefficients.

[c]Conditional effects (simple slopes) of digital self-efficacy were probed at low (−1 SD) and high (+1 SD) levels of social capital and are presented in Figure 1.

[d]All models were adjusted for the same set of covariates (age, sex, education level, living arrangements, disability, household income, self-rated health, and digital competence).

[e]ref: reference category.

[f]A currency exchange rate of KRW 4,000,000=US $3096.81 (KRW 1=US $0.000774; 2022 annual average, World Bank) was applied [63].

XSL•FO
RenderX

Moderating Role of Social Capital

The interaction effect of digital self-efficacy and social capital was significant among both middle-aged (B=−0.11, *P*=.002) and older adults (B=−0.24, *P*<.001; Table 3). We estimated the unstandardized simple slopes of these relationships (Figure 1). High and low levels of social capital were determined using SD, and values 1 SD above and 1 SD below the mean indicated high and low levels of social capital, respectively. In both the middle-aged and older adult groups, digital self-efficacy displayed a stronger positive relationship with satisfaction with the daily life changes stemming from digital transformation (satisfaction) when social capital was low. Additionally, this relationship was stronger in the older adult group (B=0.71, *P*<.001 for low and B=0.47, *P*<.001 for high social capital) than in the middle-aged group (B=0.55, *P*<.001 for low and B=0.46, *P*<.001 for high social capital). Among covariates, education level and digital competence were positively associated with satisfaction in all groups; household income was positively associated with satisfaction only in the total sample and middle-aged adults, and age was negatively associated with satisfaction only in the total sample.

These results were largely consistent in additional SEM analyses using a MIMIC model (Multimedia Appendices 1-3), particularly with respect to the effects of digital self-efficacy and the digital self-efficacy×social capital interaction term across groups. The direct effect of social capital was significant in the total and middle-aged samples, but was attenuated and not statistically significant among older adults in SEM.

To further examine age heterogeneity, we conducted age-stratified analyses (40 - 50 y: n=1189, 51 - 64 y: n=1796, 65 - 74 y: n=872, and ≥75 y: n=298; Multimedia Appendix 4). In the 65 - 74 years group, digital self-efficacy was positively associated with satisfaction with the daily life changes stemming from digital transformation (B=0.60, *P*<.001), and the interaction effect was negative and statistically significant (B=−0.27, *P*<.001). In the ≥75 years group, the interaction term did not reach statistical significance (B=−0.21, *P*=.18), whereas the direct effect of digital self-efficacy remained significant (B=0.54, *P*<.001).

XSL•FO
RenderX

Figure 1. Graphical representation of the moderation effect of social capital in the relationship between digital self-efficacy and satisfaction with the daily life changes stemming from digital transformation.

XSL•FO
RenderX

Discussion

Principal Findings

This study examined the moderating effect of social capital on the relationship between digital self-efficacy and satisfaction with the daily life changes stemming from digital transformation among middle-aged and older adults. In both groups, digital self-efficacy was positively linked to satisfaction, and the effect was stronger among individuals with low social capital compared with those with high social capital. Notably, this positive relationship was more pronounced among older adults than among middle-aged adults. These results offer valuable insights into the challenges of digital transformation in aging populations. Importantly, satisfaction with the daily life changes stemming from digital transformation remains underexamined compared with broader outcomes such as quality of life. Focusing on this transformation-specific outcome helps capture individuals' subjective adaptation to rapidly changing digital environments, which may not be fully reflected in broader well-being indicators. This study extends previous research by demonstrating that social capital moderates the association between digital self-efficacy and satisfaction with the daily life changes stemming from digital transformation, with clear differences across age groups.

Our findings revealed that for individuals with low social capital, digital self-efficacy was more strongly associated with satisfaction with the daily life changes stemming from digital transformation. This was particularly evident among older adults, who may face greater challenges in adapting to digital advancements. These findings suggest that older adults with limited social connections may rely more heavily on their own digital self-efficacy to maintain social engagement, access services, and maintain satisfaction with the daily life changes stemming from digital transformation. Enhancing digital self-efficacy may support more positive perceptions of digital transformation, particularly among socially vulnerable groups, consistent with the COR theory [46]. Similarly, past studies have shown that the positive effect of frequent internet use on life satisfaction is stronger among older adults living alone compared with those living with others [61]. Likewise, the autonomy benefits of active internet use have been reported to be greater among older adults with poorer functional health than among those with better functional health [69].

The moderation effect of social capital reached statistical significance among adults aged 65 - 74 years but not among those aged ≥75 years. However, among the old-old (≥75 y; n=298; B=−0.21; *P*=.18), the coefficient was comparable in direction and magnitude to that observed in the young-old (65 - 74 y; n=872; B=−0.27; *P*<.001). This suggests that the nonsignificant finding in the old-old group may reflect limited statistical power rather than the absence of a meaningful moderating pattern. Although the directional pattern was broadly consistent across older adult subgroups, these age-stratified findings may also indicate heterogeneity within the older adult population. The young-old (65 - 74 y) often experience major life transitions such as retirement, accompanied by shifts in social roles and available resources; thus, digital self-efficacy may function more prominently as a compensatory resource [70]. In contrast, among adults aged ≥75 years, satisfaction with the daily life changes stemming from digital transformation may be shaped not only by digital self-efficacy and social capital but also by structural and functional constraints, such as physical limitations, cognitive burden, poorer health, and accessibility barriers [20,71].

The observed age-related differences may also be interpreted in light of a potential digital cohort effect, whereby digital adaptation is shaped more by the timing and context of initial exposure to digital technologies than by chronological age alone [25,72]. Earlier and sustained exposure to digital technologies may strengthen digital self-efficacy and its influence on satisfaction, whereas later or limited exposure may constrain digital skills and self-efficacy, making it more difficult to effectively use social capital in digital contexts. This interpretation is consistent with previous research emphasizing the role of experience and cumulative engagement with digital technologies in shaping the digital divide among older adults [20]. Although the moderation effect was not statistically significant in the old-old group, digital self-efficacy remained significantly associated with satisfaction. This suggests that digital self-efficacy still plays an important role in this group, consistent with previous findings showing that it continues to influence behavioral intention to use the internet in old-old age [73].

Our results highlight the importance of self-efficacy, as also suggested in social cognitive theory [36]. Self-efficacy is particularly relevant in later life, as this stage is characterized by resource losses [74]. Strategies such as persuasion, observational learning, and mastery experiences can effectively enhance self-efficacy, enabling individuals to balance their actions and acquire new skills [36]. Notably, previous research has consistently identified digital self-efficacy as a key factor in digital proficiency, learning outcomes, and life satisfaction [56,74,75]. Therefore, improving digital self-efficacy may foster acceptance of digital services and confidence in navigating them, supporting greater satisfaction with the daily life changes stemming from digital transformation.

Building upon theoretical foundations, our findings suggest that policy and program designs can benefit from differentiated strategies for middle-aged and older adults. For middle-aged adults, who are often required to engage with digital technologies in occupational and administrative contexts, interventions should focus on enhancing the efficient and context-specific use of digital services. Program impact may be strengthened by long-term outreach that enhances exposure and awareness, alongside a shift toward skill-building-based empowerment and low-barrier service pathways.

For older adults, particularly those with low social capital, interventions could prioritize strengthening digital self-efficacy through tailored support, including dedicated one-on-one mentorship to provide guided mastery experiences. Such support can emphasize confidence-building, achievable goal setting, and the development of organizational and self-management skills. In addition, heterogeneity within older adults should be considered. The young-old may be more likely than the old-old

XSL•FO
RenderX

to have previous experience with digital technologies before retirement, and such experience has been shown to be a strong predictor of internet use in later life [20]. Previous experience may facilitate continued learning and make the young-old more responsive to structured, skill-building support such as one-on-one mentorship, enabling the translation of existing familiarity into effective skills and rapid mastery. Therefore, interventions for the young-old may benefit from focusing on such skill-oriented, experience-based support. In contrast, interventions for the old-old should prioritize reducing barriers to access and use, such as simplifying interfaces and providing assistive support. They may also benefit from approaches that enhance digital self-efficacy through personalized, family-integrated support, with attention to accessibility, usability, and functional limitations, including visual and hearing impairments [20,71]. In South Korea, government initiatives to promote digital literacy have included kiosk usage training at senior welfare centers, along with efforts to transform over 2000 senior centers into smart facilities by 2024 [76]. Nonetheless, older adults continue to report low digital self-efficacy and limited awareness of online government services [77], indicating that further expansion and active dissemination are needed to enhance program reach and impact. These efforts may further prioritize subgroups facing compounded barriers, such as individuals with lower education or income, those living alone, people with disabilities, and rural residents. Targeted and accessible support for these groups may further enhance equity in digital adaptation.

Satisfaction with the daily life changes stemming from digital transformation may facilitate individuals' openness to engaging with digital environments and, in turn, may help them realize the potential benefits of digital technologies. The active usage of digital technology provides middle-aged and older adults with access to resources and services such as information, entertainment, shopping, and banking, thereby improving their independent living, health, and life satisfaction [78-80]. The use of digital services can also predict health outcomes [78]. Based on data from 5 international longitudinal cohort studies, Li [11] found that middle-aged and older adults who use the internet reported lower levels of frailty, fewer health deficits, and better overall health. Additionally, the networks formed online can foster social connections among individuals experiencing a gradual reduction in their social roles and relationships, thereby contributing to decreased loneliness and improved well-being [19,38,81,82]. This suggests a reciprocal relationship where active digital engagement not only requires social resources but also provides opportunities to maintain or enhance them, particularly for those with limited social networks [83]. Within this framework, improving digital self-efficacy does more than facilitate technical adaptation; it may empower socially isolated individuals to manage and expand their social capital through digital means.

By examining individuals' satisfaction with the daily life changes stemming from digital transformation, this study highlights the evolving challenges of the digital era. These challenges are particularly relevant in the post–COVID-19 period, as digital competencies have become essential for life satisfaction. South Korea provides a unique context for this study, as it is the fastest-aging society among the OECD countries [84] and is undergoing rapid digital transformation [6]. Our findings not only contribute to the literature on aging and digitalization but also provide actionable insights for addressing similar challenges worldwide.

Strengths and Limitations

This study used nationally representative survey data from 2985 middle-aged and 1170 older adults in South Korea. The survey was conducted by reputable organizations and government agencies, including the Ministry of Science and ICT in collaboration with the National Information Society Agency. Trained interviewers visited participants' residences to administer the questionnaires in a standardized manner. In addition, the dataset contained no missing values for the variables used in this study, eliminating the need for casewise deletion or imputation. Therefore, the large sample size and rigorous data collection methods enhance the reliability and generalizability of our findings, offering valuable insights for the older adult population.

Nevertheless, this study has several limitations. First, this study focused on the general population and did not include specific groups such as people with disabilities or those living alone, who may experience distinct challenges in digital integration. Future research should explicitly include these populations and use accessibility-focused measures, tailored digital interventions, or subgroup-specific analyses to better understand digital integration across diverse groups. Second, the instrument used to measure social capital assessed both online and offline social capital, and it was not possible to differentiate between the two in the results. Therefore, the findings reflect overall social capital rather than mode-specific resources, and future research should disentangle offline and online dimensions to examine whether their moderating roles differ. Third, the cross-sectional design precludes the establishment of a clear causal relationship between digital self-efficacy and satisfaction with the daily life changes stemming from digital transformation. It remains unclear whether higher levels of digital self-efficacy lead to greater satisfaction or vice versa. Future studies should use a longitudinal design to explore these causal relationships. Fourth, key variables were measured using self-reported questionnaires, which may be subject to recall or social desirability bias. Fifth, although we adjusted for digital competence, some conceptual overlap between digital competence and digital self-efficacy may remain. Future studies should further disentangle these constructs using more refined measures or alternative model specifications. Sixth, because this study was conducted in South Korea, which is characterized by rapid digital transformation and distinct sociocultural features, generalizability to other countries or cultural settings may be limited. Additionally, most participants had completed high school or above (approximately 80%), which may further limit generalizability to individuals with lower educational attainment. Seventh, although our study focused on digital self-efficacy in relation to digital transformation–related satisfaction, we should also consider the possibility of problematic internet use and addiction. Previous studies have also shown that spending excessive time on online activities increases the risk of depressive symptoms [85], sleep disturbance [86], and musculoskeletal pain [87] in this

XSL•FO
RenderX

population. Accordingly, future research should assess not only the benefits of digital engagement but also its negative consequences, such as digital burden, fatigue, and cognitive load, to provide a more balanced understanding of its impact on well-being.

Conclusion

The findings of this study demonstrate the important role of digital self-efficacy in improving satisfaction with the daily life changes stemming from digital transformation, particularly among older adults with low social capital. This study not only contributes to the literature on aging and digitalization but also offers practical insights for reducing the digital divide. The results emphasize the importance of fostering digital self-efficacy through strategies such as mentorship programs, observational learning, and tailored educational interventions. These strategies can empower vulnerable populations to adapt to digital transformation.

Acknowledgments

We thank the Ministry of Science and ICT and the National Information Society Agency (NIA) for sharing the data with the researchers.

We confirm that no generative AI tools were used in the preparation of this manuscript.

Funding

This study was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (No. 2020R1A6A1A03041989) and the Ministry of Science and ICT (MSIT) (RS-2025-00520935) to HL. JC received a scholarship from Brain Korea 21 FOUR Project fundedby NRF of Korea, Yonsei University College of Nursing.

Data Availability

This data is publicly available, and the authors do not have the authority to share it.

Authors' Contributions

JC, HC, and HL contributed to the conceptualization, study design, data interpretation, manuscript drafting, and manuscript revision. JC performed the data analysis. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Confirmatory factor analysis results for the latent variable used in the multiple indicators multiple causes (MIMIC) model.
[DOCX File, 17 KB - aging_v9i1e79845_app1.docx]

Multimedia Appendix 2
Goodness-of-fit indices for the multiple indicators multiple causes (MIMIC) model.
[DOCX File, 13 KB - aging_v9i1e79845_app2.docx]

Multimedia Appendix 3
Multiple indicators multiple causes (MIMIC) model for satisfaction with daily life changes stemming from digital transformation in the total sample, middle-aged adults, and older adults.
[DOCX File, 18 KB - aging_v9i1e79845_app3.docx]

Multimedia Appendix 4
Age-specific associations between digital self-efficacy, social capital, and satisfaction with the daily life changes stemming from digital transformation (N=4155).
[DOCX File, 17 KB - aging_v9i1e79845_app4.docx]

Checklist 1
STROBE checklist.
[PDF File, 95 KB - aging_v9i1e79845_app5.pdf]

References

XSL•FO
RenderX

1. Lee SM, Lee D. Opportunities and challenges for contactless healthcare services in the post-COVID-19 era. Technol Forecast Soc Change 2021 Jun;167:120712. [doi: 10.1016/j.techfore.2021.120712] [Medline: 33654330]
2. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. Inquiry 2022;59:469580221096272. [doi: 10.1177/00469580221096272] [Medline: 35471138]
3. Digital 2023: Global overview report. DataReportal. 2023. URL: https://datareportal.com/reports/digital-2023-global-overview-report [accessed 2026-02-13]
4. Government at a Glance 2025. Organisation for Economic Co-operation and Development. URL: https://www.oecd.org/en/publications/government-at-a-glance-2025_0efd0bcd-en/full-report/digital-government-index_1edec44e.html [accessed 2026-02-13]
5. World social report 2023: leaving no one behind in an ageing world. : United Nations | Department of Economic and Social Affairs Publications; 2023 URL: https://desapublications.un.org/publications/world-social-report-2023-leaving-no-one-behind-ageing-world [accessed 2026-02-13]
6. Nam Y, Lee ST. Behind the growth of FinTech in South Korea: digital divide in the use of digital financial services. Telematics and Informatics 2023 Jun;81:101995. [doi: 10.1016/j.tele.2023.101995]
7. van Dijk J. Digital divide research, achievements and shortcomings. Poetics 2006 Aug;34(4-5):221-235. [doi: 10.1016/j.poetic.2006.05.004]
8. Reneland-Forsman L. 'Borrowed access' – the struggle of older persons for digital participation. International Journal of Lifelong Education 2018 May 4;37(3):333-344. [doi: 10.1080/02601370.2018.1473516]
9. Seifert A. The digital exclusion of older adults during the COVID-19 pandemic. J Gerontol Soc Work 2020;63(6-7):674-676. [doi: 10.1080/01634372.2020.1764687] [Medline: 32401181]
10. Li P, Zhang C, Gao S, et al. Association between daily internet use and incidence of chronic diseases among older adults: prospective cohort study. J Med Internet Res 2023 Jul 17;25:e46298. [doi: 10.2196/46298] [Medline: 37459155]
11. Li L. Internet use and frailty in middle-aged and older adults: findings from developed and developing countries. Global Health 2024;20(1):53. [doi: 10.1186/s12992-024-01056-6]
12. Zapletal A, Wells T, Russell E, Skinner MW. On the triple exclusion of older adults during COVID-19: technology, digital literacy and social isolation. Soc Sci Humanit Open 2023;8(1):100511. [doi: 10.1016/j.ssaho.2023.100511] [Medline: 37021073]
13. Alhassan MD, Adam IO. The effects of digital inclusion and ICT access on the quality of life: a global perspective. Technol Soc 2021 Feb;64:101511. [doi: 10.1016/j.techsoc.2020.101511]
14. Seifert A, Cotten SR, Xie B. A double burden of exclusion? Digital and social exclusion of older adults in times of COVID-19. J Gerontol B Psychol Sci Soc Sci 2021 Feb 17;76(3):e99-e103. [doi: 10.1093/geronb/gbaa098] [Medline: 32672332]
15. Barreda Gutiérrez M, Cantarero-Prieto D, Pascual Sáez M. Age, technology, and the digital divide: are they directly related to mental health problems? Healthcare (Basel) 2024 Dec 5;12(23):2454. [doi: 10.3390/healthcare12232454] [Medline: 39685075]
16. Komp-Leukkunen K. A life-course perspective on older workers in workplaces undergoing transformative digitalization. Gerontologist 2023 Oct 17;63(9):1413-1418. [doi: 10.1093/geront/gnac181] [Medline: 36508331]
17. Li J, Wang Y, Han J. A study on the relationship between digital leadership, digital self-efficacy, and work engagement: the moderating role of technostress. Front Psychol 2026;17:1751207. [doi: 10.3389/fpsyg.2026.1751207]
18. Tian X, Mi S. A study of the deconstruction and construction of self-efficacy in internet use among older people. BMC Geriatr 2025 May 20;25(1):355. [doi: 10.1186/s12877-025-05892-y] [Medline: 40394477]
19. Barbosa Neves B, Fonseca JRS, Amaro F, Pasqualotti A. Social capital and internet use in an age-comparative perspective with a focus on later life. PLOS ONE 2018;13(2):e0192119. [doi: 10.1371/journal.pone.0192119] [Medline: 29481556]
20. Friemel TN. The digital divide has grown old: determinants of a digital divide among seniors. New Media & Society 2016 Feb;18(2):313-331. [doi: 10.1177/1461444814538648]
21. Fang X, Mo Z. The impact of educational digitalization on life satisfaction: an empirical analysis based on PISA 2022. Front Psychol 2026;17:1740535. [doi: 10.3389/fpsyg.2026.1740535] [Medline: 41693990]
22. Barykin SE, Sergeev SM, Kapustina IV, et al. Environmental sustainability and digital transformation of socio-economic: quality of life perspective. J Env Assmt Pol Mgmt 2023 Mar;25(1):2350001. [doi: 10.1142/S1464333223500011]
23. Li H, Li J, Yang Z. The impact of digital economy development on subjective well-being: micro-level evidence from China. Front Public Health 2025;13:1660857. [doi: 10.3389/fpubh.2025.1660857]
24. Hameed K, Naha R, Hameed F. Digital transformation for sustainable health and well-being: a review and future research directions. Discov Sustain 2024;5(1):104. [doi: 10.1007/s43621-024-00273-8]
25. Hargittai E. Second-level digital divide: differences in people's online skills. First Monday 2002;7(4). [doi: 10.5210/fm.v7i4.942]
26. Van Deursen AJ, Helsper EJ. Collateral benefits of internet use: explaining the diverse outcomes of engaging with the internet. New Media Soc 2018 Jul;20(7):2333-2351. [doi: 10.1177/1461444817715282] [Medline: 30581362]
27. Hunsaker A, Hargittai E. A review of internet use among older adults. New Media & Society 2018 Oct;20(10):3937-3954. [doi: 10.1177/1461444818787348]

XSL•FO
RenderX

28. Sen K, Prybutok G, Prybutok V. The use of digital technology for social wellbeing reduces social isolation in older adults: a systematic review. SSM Popul Health 2022 Mar;17:101020. [doi: 10.1016/j.ssmph.2021.101020] [Medline: 35024424]
29. Dhawan S. Online learning: a panacea in the time of COVID-19 crisis. Journal of Educational Technology Systems 2020 Sep;49(1):5-22. [doi: 10.1177/0047239520934018]
30. Gale T, Devine MA. Advancing understanding of digital well-being at the intersection of technology and leisure. J Leis Res 2023 Oct 20;54(5):667-683. [doi: 10.1080/00222216.2023.2247295]
31. Tutar Ö, Turhan FH. Digital leisure: transformation of leisure activities. Shanlax International Journal of Education 2023;11(S1-Oct):16-28. [doi: 10.34293/education.v11iS1-Oct.6365]
32. Ulfert-Blank AS, Schmidt I. Assessing digital self-efficacy: review and scale development. Comput Educ 2022 Dec;191:104626. [doi: 10.1016/j.compedu.2022.104626]
33. Gatti FM, Brivio E, Galimberti C. "The future is ours too": a training process to enable the learning perception and increase self-efficacy in the use of tablets in the elderly. Educ Gerontol 2017 Apr 3;43(4):209-224. [doi: 10.1080/03601277.2017.1279952]
34. Kim M, Oh J, Kim B. Experience of digital music services and digital self-efficacy among older adults: enjoyment and anxiety as mediators. Technol Soc 2021 Nov;67:101773. [doi: 10.1016/j.techsoc.2021.101773] [Medline: 34538985]
35. Spears J, Zheng R. Older adults' self-efficacy in computer use and the factors that impact their self-efficacy: a path analysis. Educ Gerontol 2020 Dec 1;46(12):757-767. [doi: 10.1080/03601277.2020.1815976]
36. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037/0033-295X.84.2.191]
37. Bandura A. Self-Efficacy: The Exercise of Control: New York, NY: W.H. Freeman; 1997.
38. Wilson G, Gates JR, Vijaykumar S, Morgan DJ. Understanding older adults' use of social technology and the factors influencing use. Ageing Soc 2023 Jan;43(1):222-245. [doi: 10.1017/S0144686X21000490]
39. Han S, Nam SI. Creating supportive environments and enhancing personal perception to bridge the digital divide among older adults. Educ Gerontol 2021 Aug 3;47(8):339-352. [doi: 10.1080/03601277.2021.1988448]
40. Putnam RD. Bowling Alone: The Collapse and Revival of American Community: New York, NY: Simon & Schuster; 2000.
41. Cudjoe TKM, Roth DL, Szanton SL, Wolff JL, Boyd CM, Thorpe RJ. The epidemiology of social isolation: national health and aging trends study. J Gerontol B Psychol Sci Soc Sci 2020 Jan 1;75(1):107-113. [doi: 10.1093/geronb/gby037] [Medline: 29590462]
42. Sauter J, Widmer E, Kliegel M. Changes in family composition and their effects on social capital in old age: evidence from a longitudinal study conducted in Switzerland. Ageing Soc 2023 Mar;43(3):724-742. [doi: 10.1017/S0144686X21000921]
43. Wen Z, Peng S, Yang L, et al. Factors associated with social isolation in older adults: a systematic review and meta-analysis. J Am Med Dir Assoc 2023 Mar;24(3):322-330. [doi: 10.1016/j.jamda.2022.11.008] [Medline: 36549651]
44. Jung H, Lee H, Nam EW. Mediating effect of social capital on the association between digital literacy and life satisfaction among older adults in South Korea: cross-sectional study. JMIR Form Res 2025 Feb 6;9:e68163. [doi: 10.2196/68163] [Medline: 39916363]
45. Yoo N, Jang SH. Enhancing or compensating? Role of on- and offline social capital and technological self-efficacy on subjective well-being among immigrants and natives. Cyberpsychol Behav Soc Netw 2024 Nov;27(11):846-855. [doi: 10.1089/cyber.2024.0152] [Medline: 39381858]
46. Hobfoll SE, Halbesleben J, Neveu JP, Westman M. Conservation of resources in the organizational context: the reality of resources and their consequences. Annu Rev Organ Psychol Organ Behav 2018 Jan 21;5(1):103-128. [doi: 10.1146/annurev-orgpsych-032117-104640]
47. Herdian, Qolbi ZN, Na'imah T, Meilani ER, Nuryana Z, Wijaya TT. Digital literacy, self-efficacy, and digital well-being: direct and indirect effects among final-year university students. Curr Psychol 2026 Apr;45(7):729. [doi: 10.1007/s12144-026-09313-0]
48. 2024 administrative statistics results by life stage. Statistics Korea. URL: https://www.kostat.go.kr/board.es?mid=a10301010000&bid=11895&act=view&list_no=442561 [accessed 2026-02-08]
49. Jun D. The effect of digital competences and digital supporters on life satisfaction: focusing on the moderating effect of aging. KSPA 2023 Nov 30;34(3):107-127. [doi: 10.53865/KSPA.2023.11.34.3.107]
50. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. Methods Psychol Res Online 2003;8(2):23-74 [FREE Full text]
51. Bagozzi RP, Yi Y. On the evaluation of structural equation models. JAMS 1988 Mar;16(1):74-94. [doi: 10.1007/BF02723327]
52. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
53. Kenny DA, Kaniskan B, McCoach DB. The performance of RMSEA in models with small degrees of freedom. Sociol Methods Res 2015 Aug;44(3):486-507. [doi: 10.1177/0049124114543236]
54. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis, 8th edition: Cengage; 2019.
55. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 1981 Feb;18(1):39-50. [doi: 10.1177/002224378101800104]

XSL•FO
RenderX

56. Jun M, Kim M, Han J. Examining the digital technology usage of older adults as determinants of life satisfaction in the digital divide. J Appl Gerontol 2025 Aug;44(8):1349-1357. [doi: 10.1177/07334648241302674] [Medline: 39626312]
57. Park JI. A study on factors affecting digital literacy among older adults. Local Gov Stud 2023;27:277-300. [doi: 10.20484/klog.27.2.12]
58. Kong H, Liu H. The relationship between ICT use and perceived life satisfaction among older people in Korea: the mediating effect of social capital. Sustainability 2023;15(12):9353. [doi: 10.3390/su15129353]
59. Williams D. On and off the 'Net: scales for social capital in an online era. J Comp Mediated Comm 2006 Jan;11(2):593-628. [doi: 10.1111/j.1083-6101.2006.00029.x]
60. Kim BS, Ko JH. A study on the impact of digital supporter on digital social capital in the elderly: focusing on the mediating effect of digital usage. DPSS 2020 Oct 31;13(2):27-57. [doi: 10.22417/DPSS.2020.10.13.2.27]
61. Cho H, Choi M, Lee H. Mobile internet use and life satisfaction among older adults: the moderating effect of living alone. J Appl Gerontol 2024 Jul;43(7):841-849. [doi: 10.1177/07334648231216383] [Medline: 37982396]
62. Oh J. Digital transformation and satisfaction with daily life change across different age groups after COVID-19: the role of digital competence and changes in internet service usage. JKGS 2024 Aug 31;44(4):493-507. [doi: 10.31888/JKGS.2024.44.4.493]
63. Official exchange rate (local currency units per US $, period average). World Bank. URL: https://data.worldbank.org/indicator/PA.NUS.FCRF?locations=KR [accessed 2026-03-30]
64. Lee J, Jang SN. Have changes in Internet use during the COVID-19 pandemic affected older adults' self-rated health? A cross-sectional study of young-old and old-old populations in Korea. Geriatr Nurs 2022;48:145-149. [doi: 10.1016/j.gerinurse.2022.09.012] [Medline: 36219933]
65. Jung HK, You YY, Kim JR. The effect of digital (PC, mobile) utilization competence and digital information utilization on digital technology awareness changes: focusing on seniors. APJCRI 2023 Apr 30;9(4):165-174. [doi: 10.47116/apjcri.2023.04.14]
66. Chang C, Gardiner J, Houang R, Yu YL. Comparing multiple statistical software for multiple-indicator, multiple-cause modeling: an application of gender disparity in adult cognitive functioning using MIDUS II dataset. BMC Med Res Methodol 2020 Nov 12;20(1):275. [doi: 10.1186/s12874-020-01150-4] [Medline: 33183226]
67. Cao L, Chongsuvivatwong V, McNeil EB. The sociodemographic digital divide in mobile health app use among clients at outpatient departments in inner Mongolia, China: cross-sectional survey study. JMIR Hum Factors 2022 May 19;9(2):e36962. [doi: 10.2196/36962] [Medline: 35587367]
68. Rosseel Y, Jorgensen TD, Rockwood N, Oberski D, Byrnes J, Vanbrabant L, et al. Package 'lavaan': latent variable analysis. CRAN. 2025. URL: https://cran.r-project.org/web/packages/lavaan/lavaan.pdf [accessed 2026-02-13]
69. Oswald V, Wagner M. Internet usage among the oldest-old: does functional health moderate the relationship between internet usage and autonomy? Eur J Ageing 2023 Feb 21;20(1):3. [doi: 10.1007/s10433-023-00748-z] [Medline: 36802304]
70. Lim-Soh JW, Lee Y. Social participation through the retirement transition: differences by gender and employment status. Res Aging 2023 Jan;45(1):47-59. [doi: 10.1177/01640275221104716] [Medline: 35694800]
71. Kainiemi E, Saukkonen P, Virtanen L, et al. Perceived benefits of digital health and social services among older adults: a population-based cross-sectional survey. Digit Health 2023;9:20552076231173559. [doi: 10.1177/20552076231173559] [Medline: 37312955]
72. DiMaggio P, Hargittai E, Celeste C, Shafer S. Digital inequality: from unequal access to differentiated use. In: Neckerman KM, editor. Social Inequality: Russell Sage Foundation; 2004:355-400.
73. Jokisch MR, Scheling L, Doh M, Wahl HW. Contrasting internet adoption in early and advanced old age: does internet self-efficacy matter? J Gerontol B Psychol Sci Soc Sci 2022 Feb 3;77(2):312-320. [doi: 10.1093/geronb/gbab096] [Medline: 34100932]
74. Jokisch MR, Schmidt LI, Doh M, Marquard M, Wahl HW. The role of internet self-efficacy, innovativeness and technology avoidance in breadth of internet use: comparing older technology experts and non-experts. Comput Human Behav 2020 Oct;111:106408. [doi: 10.1016/j.chb.2020.106408]
75. Lozoya SVM, Guirado MAZ, Zapata Gonzalez A, Lopez ABM. Use of technologies and self-efficacy in older adults. IEEE R Iberoamericana Tecnologias Aprendizaje 2022;17(2):125-130. [doi: 10.1109/RITA.2022.3166870]
76. Promoting healthy and happy aging for the era of 10 million seniors. Ministry of Health and Welfare Press Release. 2024. URL: https://www.mohw.go.kr/board.es?mid=a10503000000&bid=0027&list_no=1480732&act=view [accessed 2026-02-13]
77. Kim Y, Lee J. Digitally vulnerable populations' use of e-government services: inclusivity and access. Asia Pacific Journal of Public Administration 2024 Oct;46(4):422-446. [doi: 10.1080/23276665.2024.2321569]
78. Benda NC, Veinot TC, Sieck CJ, Ancker JS. Broadband internet access is a social determinant of health!. Am J Public Health 2020 Aug;110(8):1123-1125. [doi: 10.2105/AJPH.2020.305784] [Medline: 32639914]
79. Chiarelli TM, Batistoni SST. An analysis of socioemotional selectivity theory in the context of older adults' use of Facebook. Educ Gerontol 2021 Jan 2;47(1):13-24. [doi: 10.1080/03601277.2020.1849956]
80. Schehl B, Leukel J, Sugumaran V. Understanding differentiated internet use in older adults: a study of informational, social, and instrumental online activities. Comput Human Behav 2019 Aug;97:222-230. [doi: 10.1016/j.chb.2019.03.031]

XSL•FO
RenderX

81. Domahidi E. The associations between online media use and users' perceived social resources: a meta-analysis. J Comput Mediat Commun 2018 Jul 1;23(4):181-200. [doi: 10.1093/jcmc/zmy007]
82. Nguyen MH, Hunsaker A, Hargittai E. Older adults' online social engagement and social capital: the moderating role of internet skills. Information, Communication & Society 2022 May 19;25(7):942-958. [doi: 10.1080/1369118X.2020.1804980]
83. Grotto G, Buja A. Association between digital technology use and social capital among older adults: a systematic review. Digit Health 2025;11:20552076251336973. [doi: 10.1177/20552076251336973] [Medline: 40656846]
84. Old-age dependency ratio (indicator). Organisation for economic co-operation and development. 2023. URL: https://data.oecd.org/pop/old-age-dependency-ratio.htm [accessed 2026-02-13]
85. Ding L, Li Z, Jiang H, Zhang X, Xiong Z, Zhu X. Mobile phone problem use and depressive symptoms: the mediating role of social support and attitude to aging among Chinese older adults. BMC Psychiatry 2024 Feb 16;24(1):135. [doi: 10.1186/s12888-024-05565-x] [Medline: 38365625]
86. Bertocchi FM, De Oliveira AC, Lucchetti G, Lucchetti ALG. Smartphone use, digital addiction and physical and mental health in community-dwelling older adults: a population-based survey. J Med Syst 2022 Jun 18;46(8):53. [doi: 10.1007/s10916-022-01839-7] [Medline: 35716194]
87. Najafipour H, Sadeghi M, Kordestani Z, et al. The prevalence of musculoskeletal pain syndrome in 15-80 years old population of Kerman: the role of smart phone use and its related factors. J Kerman Univ Med Sci 2021;28(4):358-366. [doi: 10.22062/jkmu.2021.91716]

Abbreviations

AVE: average variance extracted
CFA: confirmatory factor analysis
CFI: comparative fit index
COR: conservation of resources
CR: composite reliability
ICT: information and communication technology
MIMIC: multiple indicators multiple causes
OECD: Organisation for Economic Co-operation and Development
Q-Q: quantile-quantile
RMSEA: root mean-square-error of approximation
SEM: structural equation modeling
SRMR: standardized root-mean-square residual
TLI: Tucker-Lewis index

Edited by J Wang; submitted 30.Jun.2025; peer-reviewed by A Neumann, H Shin, R Shaleha, R Xu; revised version received 20.May.2026; accepted 12.Jun.2026; published 31.Jul.2026.

Please cite as:
Choi J, Cho H, Lee H
Digital Self-Efficacy, Satisfaction With the Daily Life Changes Stemming From Digital Transformation, and the Moderating Effect of Social Capital in Middle-Aged and Older Adults: Cross-Sectional Survey Study
JMIR Aging 2026;9:e79845
URL: https://aging.jmir.org/2026/1/e79845
doi:10.2196/79845

© Jeongeun Choi, Hyeonmi Cho, Hyangkyu Lee. Originally published in JMIR Aging (https://aging.jmir.org), 31.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Inclusion Pathways for Older Chinese Adults in the Context of Active Aging: Secondary Analysis of 2023 China Longitudinal Aging Social Survey Data

Bo Zhao[1,2], PhD; Xiaoyan Wang[1], BMs; Ren Chen[1], PhD; Eun Woo Nam[1,2], Prof Dr

[1]Department of Health Services Management, School of Health Services Management, Anhui Medical University, 81 Meishan Road, Shushan District, Hefei, Anhui, China

[2]Yonsei Global Health Center, Yonsei University, 1 Yonseidae-gil, Wonju-si, Gangwo-do, Republic of Korea

Corresponding Author:
Bo Zhao, PhD
Department of Health Services Management, School of Health Services Management, Anhui Medical University, 81 Meishan Road, Shushan District, Hefei, Anhui, China

Abstract

Background: Rapid population aging and the intensifying digitalization of everyday life are unfolding simultaneously in China. While prior studies have largely examined pairwise associations among digital inclusion, social engagement, mental health, and overall health status, few have evaluated an integrated, theoretically grounded pathway linking these domains in later life.

Objective: This study aims to quantify the direct and indirect pathways through which digital inclusion influences older adults' overall health status, social engagement, and mental health, specified as sequential mediators.

Methods: We analyzed the newly released, nationally representative data from the 2023 wave of the China Longitudinal Aging Social Survey, comprising 9918 adults aged 60 years or older. Overall health status was assessed using 3 self-rated health (SRH) indicators: current SRH, SRH relative to age peers, and SRH relative to last year. Digital inclusion was measured through digital access, device proficiency, and digital ability. Social engagement captured social support, frequency of participation in community or voluntary activities, and nononline activities. Mental health included depressive symptoms, social adaptation, and life satisfaction. Analyses included descriptive statistics, multivariable hierarchical linear regressions, and structural equation modeling to estimate direct and mediated effects (2-sided; α=.05).

Results: Older age, chronic disease, and functional limitations were associated with poorer overall health status, whereas higher education and current employment were associated with better health status. Digital inclusion was positively associated with social engagement (β=.50), which in turn was positively associated with mental health. Mental health showed the strongest association with SRH (β=.74). The direct path from social engagement to overall health status was nonsignificant (*P*=.34), indicating that participation influences health primarily through psychological pathways. In regression analyses, digital inclusion modestly improved model fit for health status outcomes, while adding mental health produced a greater increase.

Conclusions: Digital inclusion promoted active aging indirectly, by expanding social engagement and enhancing mental health, thereby improving overall health status. Policy efforts should prioritize narrowing the digital divide by improving digital skills and capability, rather than access alone. Meaningful opportunities for social engagement should also be expanded to strengthen community-based mental health support. In addition, strategies should be tailored to the differing needs of urban and rural settings.

(JMIR Aging 2026;9:e83078) doi:10.2196/83078

KEYWORDS

digital inclusion; social engagement; mental health; overall health status; healthy aging; active aging; older adults; China

Introduction

Background

Population aging has become one of the most consequential demographic shifts worldwide. In China, this transition is particularly rapid: by the end of 2024, adults aged 60 years or older accounted for 22% of the national population [1], accompanied by rising multimorbidity, functional decline, and widening social and health disparities. As chronic disease burden and long-term care needs increase, understanding the mechanisms that support health and well-being in later life has become a central scientific and policy priority [2].

To guide global responses, the Active Ageing Framework of the World Health Organization (WHO) emphasized that healthy aging is not only about disease prevention but also about creating an environment that enables older adults to remain

XSL•FO
RenderX

healthy and socially engaged [3]. This perspective aligns with core gerontological theories: the activity theory highlights the importance of meaningful participation for psychological and social well-being [4], while the socioemotional selectivity theory underscores the increasing value of emotionally meaningful social ties in later life [5]. China has integrated these principles into national strategies, illustrated by the 2025 *Guiding Opinions on Supporting Elderly People's Social Participation and Promoting the Realization of a Proactive Age*, which call for the need to mobilize older adults' capacities, enhance opportunities for social engagement and overall health, and reduce psychosocial risks [6].

Meanwhile, the rapid digitalization of everyday life has profoundly reshaped how older adults access services, maintain social ties, and engage with community resources. By 2024, China had 157 million internet users aged 60 years or older, yet their online participation remains disproportionately low [7,8]. Digital inclusion today extends beyond access to encompass skills, motivation, trust, and the ability to meaningfully participate in digital environments. The digital divide theory conceptualizes these disparities as layered barriers—spanning access, ability, and confidence—that structure opportunities for social interaction and integration [9,10]. For many older adults, insufficient digital competence, low usability of platforms, and the absence of age-friendly design restrict meaningful digital participation and may exacerbate social and health inequalities.

An integrated theoretical perspective suggests that digital inclusion can influence well-being through psychosocial pathways [11]. The removal of access and skill barriers expands older adults' capacity to obtain information and mobilize social support, forming the structural basis for the participation mechanisms described by the activity theory. Technology acceptance frameworks explain how perceived usefulness, ease of use, and digital self-efficacy determine whether digital access translates into sustained digital engagement [12]. The biopsychosocial model further posits that psychological well-being—shaped in part by social engagement—is a central determinant of both perceived and actual health [13]. Together, these perspectives imply a sequential logic in which digital inclusion facilitates social engagement, which enhances mental health, ultimately contributing to better overall health [14].

However, most existing studies examine these components in isolation—for example, exploring associations between digital use and social connectedness, or between social engagement and mental health—without testing how these elements interact as part of a continuous mechanism [11]. The absence of an integrated analytic framework limits both theoretical coherence and the development of targeted interventions to promote active aging within increasingly digital environments. Empirical validation of the full pathway linking digital inclusion, social engagement, mental health, and their overall health is therefore essential for understanding how digitalization intersects with aging and for informing policy efforts that support healthy, meaningful, and socially integrated later life [15].

Literature Review

Existing research provides substantial but fragmented evidence on the individual links within the broader connecting digital inclusion to health in later life. Studies commonly assess digital inclusion across access, skills, use frequency, and trust and report significant positive effects on social interaction and meaningful forms of participation [16-18]. Digital engagement extends beyond leisure use to include health information seeking, online consultations, communication with family, and participation in community activities [19]. These digitally enabled interactions have been shown to reduce social isolation, strengthen support networks, and enhance older adults' ability to mobilize social resources. Subgroup analyses further indicate heterogeneous effects, with women, urban residents, and younger segments of the older adult population benefiting more from digital inclusion [20]. While literature consistently affirms the positive association between digital engagement and social participation, most studies examine direct associations and rarely evaluate how digital inclusion may activate deeper psychosocial processes that ultimately shape well-being.

Research on the linkage between social engagement and mental health provides equally strong evidence yet remains similarly limited. Social engagement is widely recognized as a cornerstone to active aging. Numerous empirical studies demonstrate that greater frequency and diversity of participation are associated with lower depressive symptoms, reduced loneliness, improved social adaptation, and higher life satisfaction [21-25]. Engagement enhances psychological resources by offering meaningful roles, reinforcing identity and purpose, and strengthening emotional connections—all of which align with core propositions of the biopsychosocial model. Mental health, in turn, as the central determinant of overall health and self-rated health (SRH), exerts substantial influence through mechanisms such as stress buffering, resilience, and positive cognitive appraisal [21-27]. Although these findings underscore mental health as a pivotal conduit between social participation and health outcomes, prior studies rarely embed mental health within a continuous pathway that begins with digital inclusion and culminates in overall health [17,28]. Consequently, the literature lacks an integrated explanation of how digitalization interacts with social processes and psychological well-being to shape health trajectories in later life.

To address these gaps, this study draws on the 2023 wave of the China Longitudinal Aging Social Survey (CLASS), the most recently released and nationally representative dataset on Chinese older adults. Importantly, the data were collected in the immediate post–COVID-19 period, during which digital technologies became deeply embedded in daily life and the digital divide among older adults intensified. This context offered a unique opportunity to examine the mechanisms through which digital inclusion influences social engagement, mental health, and overall health. By constructing and testing an integrated pathway model, this study offers timely empirical evidence to inform more precise and actionable strategies for promoting digital inclusion and active aging in China's rapidly aging society.

XSL•FO
RenderX

Methods

Study Design and Participants

This study used a cross-sectional design and analyzed data from the CLASS 2023. The CLASS project is a national, large-scale social survey conducted by the National Survey Research Center at Renmin University of China. The survey covered 28 provinces, 134 districts, and 462 villages across China, excluding Hong Kong, Taiwan, Macao, Hainan, Xinjiang, and Tibet. The survey collected detailed information on respondents' health status, family structure, social background, and economic conditions.

Ethical Considerations

The study of the CLASS database was approved by the ethics committee at the Capital University of Physical Education and Sport (ChiCTR-IOR-ChiCTR2200063177). All participants provided informed consent before, and for individuals with limited literacy skills, consent was obtained from their legally authorized representatives. Participants were informed of the study purpose and their right to withdraw at any time. All data were deidentified, stored securely, and accessible only to authorized research team members to ensure privacy and confidentiality. No financial compensation was provided, and the study involved minimal risk.

Survey Procedures

Data collection for the CLASS project was carried out using the China Social Survey Network. The survey used a stratified, multistage probability sampling method. County-level regions, including counties, county-level cities, and districts, were selected as primary sampling units. Villages and neighborhood committees served as secondary sampling units. With each secondary sampling unit, sample households were randomly selected, and 1 eligible older adult was interviewed per household. Data were collected through structured interviews and standardized questionnaires. To ensure data quality, on-site supervision and call-backs were employed as part of the quality control measures implemented by CLASS.

Variables

The dependent variable was older adults' overall health status, proxied by SRH using 3 items: current SRH (respondents rate their present health status, scores of 0 - 5), SRH relative to age peers (scores of 0 - 5), and SRH relative to last year (scores of 0 - 3). Item scores were summed (range 3 - 13), with higher scores indicating better overall health.

Independent variables included measures of digital inclusion and active aging indicators. Digital inclusion comprised 3 dimensions as follows:

- Digital access: This dimension reported whether the respondent could connect to the internet via available infrastructure and devices, measured by three items—network availability at the dwelling; current internet use; and ability to access the internet via a mobile phone, tablet, or computer. Summed scores reflected overall access, with higher scores indicating greater access.
- Device proficiency: This dimension reported self-rated proficiency in operating devices such as mobile phones and tablets, rated from 1 ("very unskilled") to 5 ("very skilled"). Higher summed scores indicated stronger digital skills.
- Digital ability: This dimension reported the performance of 11 common online activities (eg, voice or video chat, SMS text messaging, online shopping, and reading news), each scored from 0 to 1. Total scores (range 0 - 11) reflected functional digital competence.

Active aging comprises health, engagement, and security; because security is primarily a macro-level construct [29], this study focused on individual-level with health and engagement. Social engagement included 3 components as follows:

- Social support and network: This component was a six-item scale measuring (1) the number of family or relatives contacted or met at least monthly, (2) the number of family or relatives with whom one can confidently discuss private matters, (3) the number of family or relatives available to help when needed, and (4-6) the same three items for friends. Each item was scored from 0 to 5; total scores ranged from 0 to 30, with higher scores indicating stronger support and denser network.
- Activity participation frequency: This component measured involvement in 7 community or voluntary activities (community patrols, caring for older adults or children, environmental protection, dispute mediation, companion chatting, professional volunteering, and caring for the next generation). Each item was scored from 0 to 4 (0="never" to 4="almost daily"); total scores ranged from 0 to 28.
- Nononline activities: This component measured participation in 6 offline activities (religious activities; senior college or training courses; watching television, listening to the radio, reading newspapers or books, or watching an opera; singing or playing instruments; mahjong, chess, or cards; and square dancing). Each item was scored from 0 to 4; total scores ranged from 0 to 24, with higher scores indicating greater offline engagement.

Mental health included 3 measures as follows:

- Depressive symptoms: This component was measured using the 9-item short form of the Center for Epidemiologic Studies Depression Scale [30]. Items were scored from 1 to 3 (1="no," 2="sometimes," and 3="often"). Except for items 1, 4, and 9, all items were reverse-scored, producing a total score that ranged from 9 to 27, with higher scores indicating fewer depressive symptoms.
- Social adaptation: This component was measured using 8 items (eg, willingness to participate in village or neighborhood committee work, desire to contribute to society again, current enjoyment of learning, feeling useful to society, difficulty adapting to rapid social change, difficulty accepting more viewpoints, difficulty accepting new social policies, and perception that social change is increasingly unfavorable to older adults). Items were scored from 1 to 5 (1="strongly disagree" to 5="strongly agree"); total scores range from 8 to 40, with higher scores indicating better adaptation.

XSL•FO
RenderX

- Life satisfaction: This was a single global item measuring overall life satisfaction (1="very dissatisfied" to 5="very satisfied").

Following prior studies [18,31], this study included a set of covariates to control for demographic and health-related factors, including age, sex, educational level, marital status, number of coresidents, household registration (*hukou*), chronic disease (yes or no), and current employment status (yes or no). Functional status measures included as follows:

1. Activities of daily living (ADL): 11 items scored from 1 to 3 (1="independent," 2="needs some help," and 3="unable"); total scores ranged from 11 to 33.
2. Instrumental activities of daily living: 9 items scored from 1 to 3 (same as ADL); total scores range from 9 to 27, where lower scores indicated better instrumental functioning.

Statistical Analysis

Power

Data were curated in SPSS version 26.0 (IBM Corp). We conducted descriptive analyses, chi-square tests, correlation analyses, and multivariable hierarchical regression. Structural equation modeling (SEM) was specified in AMOS version 28.0 (IBM Corp); model fit was evaluated and modified as appropriate. It also compared pathways by which digital inclusion, social engagement, and mental health influence older adults' overall health status. Statistical significance was set at a 2-sided α of .05.

Data Exclusion

The 2023 CLASS dataset contains 11,670 valid cases. Based on the requirements of the research design, variables unrelated to the study aims were excluded, and missing data were processed accordingly. A final analytic sample of 9918 respondents was obtained.

Results

Descriptive Analysis

Table 1 showed significant differences in SRH across all demographic groups (P<.001). Younger respondents and those with higher education, smaller household sizes, nonrural *hukou*, absence of chronic diseases, and current employment all reported higher SRH. Men also showed slightly better SRH than women. Overall, health status was consistently better among individuals with greater socioeconomic resources and fewer health limitations (Figure 1).

XSL•FO
RenderX

Table . Baseline sociodemographic characteristics (n=9918).

Variable	Value, n (%)	SRH[a], mean (SD)	2-tailed *t*/*F* test (*df*)	*P* value
Sex			4.938[b] (9916)	<.001
Male	5121 (51.6)	8.81 (1.507)		
Female	4797 (48.4)	8.66 (1.519)		
Age (y)			15.395[b] (9916)	<.001
60-74	7491 (75.5)	8.87 (1.503)		
≥75	2427 (24.5)	8.33 (1.476)		
Marital status			–8.056[b] (9916)	<.001
Married or with spouse	8209 (82.8)	8.80 (1.490)		
Other	1709 (17.2)	8.46 (1.596)		
Education			88.915[c] (2, 9915)	<.001
Illiterate	1709 (17.2)	8.46 (1.596)		
Junior middle school or below	6791 (68.5)	8.72 (1.495)		
High school or above	1418 (14.3)	9.17 (1.406)		
Coresidents			13.404[b] (9916)	<.001
1–2 persons	7262 (73.2)	8.87 (1.460)		
≥3 persons	2656 (26.8)	8.39 (1.603)		
Household registration (*hukou*)			–3.613[b] (9916)	<.001
Rural	5183 (52.3)	8.69 (1.557)		
Nonrural	4735 (47.7)	8.80 (1.464)		
Chronic disease			28.674[b] (9916)	<.001
No	2047 (20.6)	9.56 (1.415)		
Yes	7871 (79.4)	8.53 (1.465)		
Current employment			−9.143[b] (9916)	<.001
Not working	7069 (71.3)	8.65 (1.529)		
Working	2849 (28.7)	8.95 (1.455)		

[a]SRH: Self-rated health.

[b]*t* test.

[c]*F* test.

XSL•FO
RenderX

Figure 1. Data exclusion flowchart.

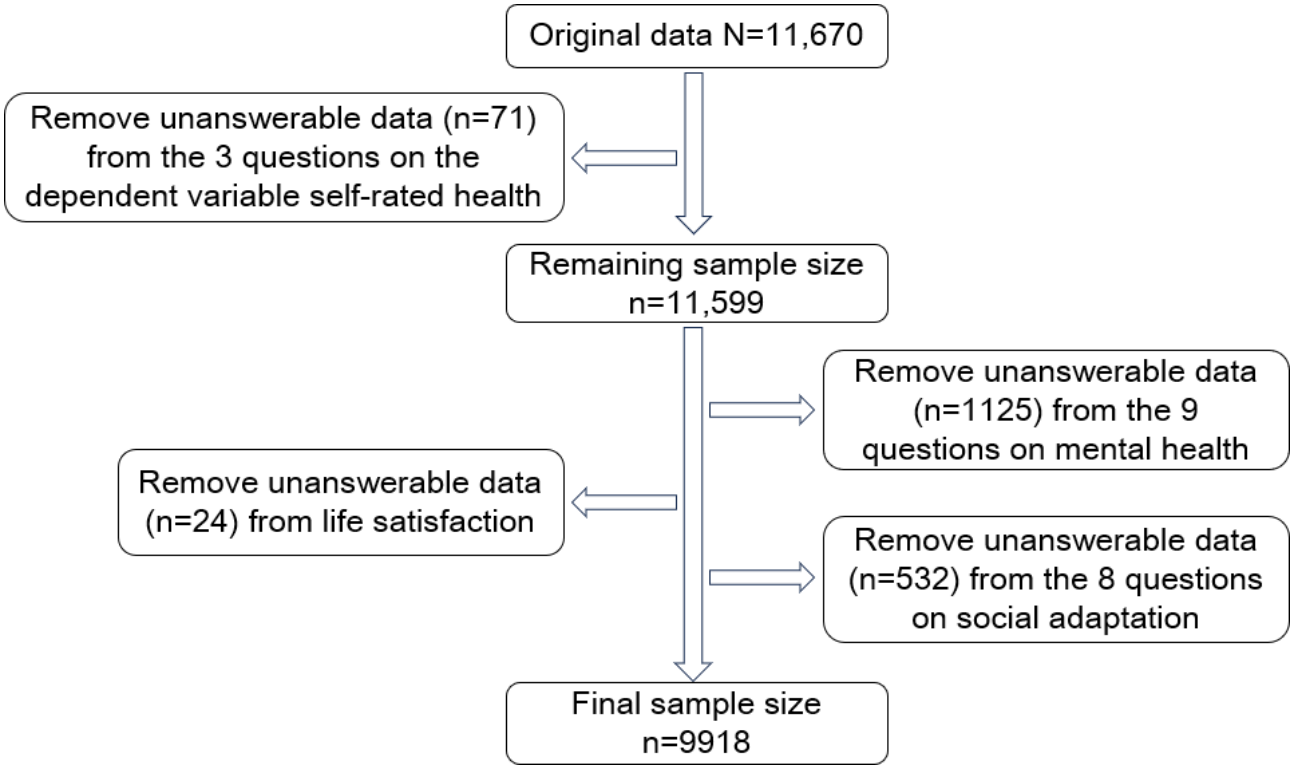


Table 2 presented descriptive statistics for the main study variables. Mean ADL and instrumental activities of daily living scores were 11.38 and 9.58, respectively. SRH measures (3.59, 3.23, and 1.92) showed moderate levels of perceived current and comparative health. Digital inclusion indicators (access, proficiency, and ability) displayed relatively low mean scores with wide variability. Social engagement levels were modest overall, while social support was relatively higher. Mental health indicators showed moderate depressive symptoms, average social adaptation, and relatively high life satisfaction.

XSL•FO
RenderX

Table . Descriptive statistics of key variables.

Domain	Score, minimum	Score, maximum	Score, mean (SD)
Functional status			
ADL[a]	11	33	11.38 (1.515)
IADL[b]	9	25	9.58 (1.802)
Overall health status (SRH[c])			
Current SRH (1-5)	1	5	3.59 (0.78)
SRH vs age peers (1-5)	1	5	3.23 (0.666)
SRH vs last year (1-3)	1	3	1.92 (0.391)
Digital inclusion			
Digital access (0-5)	0	5	1.77 (1.395)
Device proficiency (0-15)	0	15	3.84 (3.983)
Digital ability (0-11)	0	11	2.22 (2.627)
Social engagement			
Activity participation frequency (0-28)	0	25	1.86 (3.862)
Nononline activities (0-24)	0	23	5.68 (3.555)
Social support (0-30)	0	30	14.39 (4.955)
Mental health			
Depressive symptoms (CES-D[d] short, 9-27)[e]	9	27	15.48 (3.386)
Social adaptation (8-40)	8	40	24.53 (4.064)
Life satisfaction (1-5)	1	5	3.84 (0.733)

[a]ADL: activities of daily living.

[b]IADL: instrumental activities of daily living.

[c]SRH: self-rated health.

[d]CES-D: Center for Epidemiologic Studies Depression Scale.

[e]Higher scores indicate fewer depressive symptoms (lower depressive tendency).

Results of Regression Analysis

Table 3 presented the hierarchical regression results for overall health. Model 1, which included demographic and baseline health variables, explained 16.4% of the variance in overall health (ΔR^2=0.164). Older age, being unmarried, larger household size, poorer functional status, and having chronic conditions were all negatively associated with health, whereas higher education and being employed showed positive associations. Model 2 added digital inclusion variables and modestly increased explanatory power (ΔR^2=0.173). Digital proficiency and digital ability—rather than mere access—were positively associated with better perceived health. Model 3 incorporated social engagement indicators but did not meaningfully improve model fit (ΔR^2 unchanged). Model 4 added mental health variables and substantially increased explanatory power to 29.4% (ΔR^2=0.294). Lower depressive symptoms, higher life satisfaction, and better social adaptation emerged as strong predictors of overall health.

XSL•FO
RenderX

Table . Hierarchical regression for the overall health status (self-rated health).

Variables	Model 1		Model 2		Model 3		Model 4	
	Value	*P* value	Value	*P* value	Value	*P* value	Value	*P* value
Constant	13.083	<.001	12.324	<.001	12.327	<.001	8.178	<.001
Age	−.017	<.001	−.007	.01	−.007	.01	−.010	<.001
Education	.326	<.001	.209	<.001	.209	<.001	.104	.01
Marital status	−.198	.001	−.155	.01	−.155	.01	−0.085	.13
Coresidents	−.209	<.001	−.195	<.001	−.191	<.001	−.159	<.001
ADL[a]	−.033	.03	−.041	.006	−.040	.008	−.028	.04
IADL[b]	−.162	<.001	−.157	<.001	−.157	<.001	−.139	<.001
Chronic disease	−.835	<.001	−.851	<.001	−.850	<.001	−.585	<.001
Current Employment	.228	<.001	.267	<.001	.274	<.001	.277	<.001
Digital inclusion								
Digital access	—[c]	—	0.014	.54	0.019	.43	−0.009	.70
Device proficiency	—	—	.018	.03	0.017	.06	0.012	.15
Digital ability	—	—	.036	<.001	.032	.001	0.002	.79
Social engagement								
Activity frequency	—	—	—	—	0	.92	0.003	.35
Nononline activities	—	—	—	—	0.007	.13	−0.003	.51
Social support	—	—	—	—	−0.005	.09	−.009	.001
Mental health								
Depressive symptoms	—	—	—	—	—	—	.053	<.001
Life satisfaction	—	—	—	—	—	—	.636	<.001
Social adaptation	—	—	—	—	—	—	.014	<.001
F test (*df*)	243.780 (8, 9909)	<.001	190.043 (11, 9906)	<.001	149.679 (14, 9903)	<.001	244.467 (17, 9900)	<.001
R^2	0.164	—	0.174	—	0.175	—	0.296	—
$\triangle R^2$	0.164	—	0.173	—	0.173	—	0.294	—

[a]ADL: activities of daily living.

[c]Data not available.

To verify the robustness of the baseline regression results, a sensitivity analysis was also conducted. As some areas may introduce bias to the estimated digital inclusion effects, respondents from 3 regions with the highest levels of internet development during the survey period (Guangdong, Shanghai, and Beijing) were excluded. Using the remaining sample, we re-estimated the pathway linking digital inclusion, social engagement, mental health, and overall health. The results of the revised model were highly consistent with models in Table 3, indicating strong robustness of the estimated relationships (Table 4).

XSL•FO
RenderX

Table . Results of the robustness analysis.

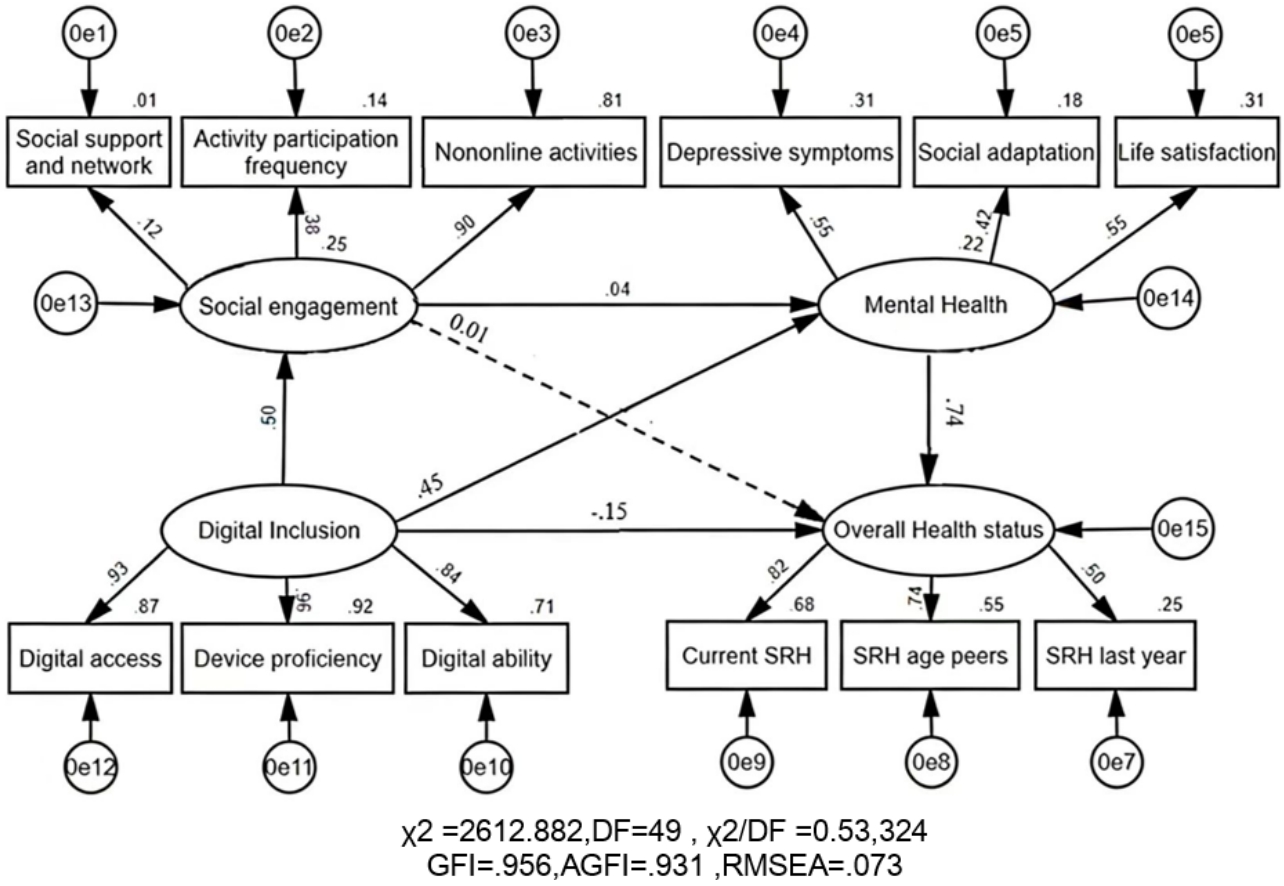
Variables	Model 5		Model 6		Model 7		Model 8	
	Value	*P* value	Value	*P* value	Value	*P* value	Value	*P* value
Digital inclusion								
Digital access	__[a]	—	−0.011	.68	−0.009	.73	−0.04	.10
Device proficiency	—	—	0.016	.10	0.013	.21	0.011	.23
Digital ability	—	—	.064	<.001	.061	<.001	.032	.002
Social engagement								
Activity frequency	—	—	—	—	−0.002	.66	−0.001	.81
Nononline activities	—	—	—	—	.014	.005	0.003	.52
Social support	—	—	—	—	−0.001	.76	−.006	.04
Mental health								
Depressive symptoms	—	—	—	—	—	—	.055	<.001
Life satisfaction	—	—	—	—	—	—	.617	<.001
Social adaptation	—	—	—	—	—	—	.019	<.001
Sample sizes	8533	—	8533	—	8533	—	8533	—
Control variables	Yes	—	Yes	—	Yes	—	Yes	—
F test (*df*)	209.194 (8, 8524)	<.001	167.196 (11, 8521)	<.001	132.007 (14, 8518)	<.001	211.313 (17, 8515)	<.001
R^2	0.164	—	0.178	—	0.178	—	0.297	—
$\triangle R^2$	0.163	—	0.176	—	0.177	—	0.295	—

SEM Analysis Outcomes

Figure 2 showed that the structural equation model identified multiple significant pathways linking digital inclusion, social engagement, mental health, and overall health status. Digital inclusion strongly predicted social engagement (β=.45) and influenced overall health indirectly through its effects on engagement and, subsequently, mental health. Although the direct path from social engagement to overall health was negligible (β=.01), its effect on mental health was significant (β=.04). Mental health had the strongest predictive power in the model (β=.74). Surprisingly, digital inclusion had a small negative direct effect on overall health (β=–0.15), although the indirect pathways through engagement and mental health were positive. Model fit indices indicated acceptable fit (goodness-of-fit index=0.956, adjusted goodness-of-fit index=0.931, and root mean square error of approximation=0.073).

XSL•FO
RenderX

Figure 2. Structural equation model linking digital inclusion, social engagement, mental health, and overall health status. DF: degree of freedom; GFI: goodness-of-fit index; AGFI: adjusted goodness-of-fit index; RMSEA: root mean square error of approximation; SRH: self-rated health.



Discussion

Principal Results

This study elucidated how demographic characteristics, digital inclusion, and social engagement and mental health jointly shape older adults' perceptions of overall health. First, results showed that advancing age, chronic disease burden, and functional limitations were associated with poorer SRH, consistent with longitudinal evidence from multiple countries. Specifically, SRH declines with age—likely reflecting physiological deterioration, cumulative adverse life events, and shifting social roles—and age is a robust predictor of SRH among older adults [32]. Chronic diseases also weakened perceived overall health, reinforcing findings that multimorbidity served as a risk factor for poor SRH [33]. Internationally, analyses using the Health and Retirement Study (HRS) and the English Longitudinal Study of Ageing (ELSA) showed that a higher number of conditions was tightly linked to lower SRH, even after psychological factors were considered [34]. Functional limitations further depressed health perceptions from a capability standpoint. By contrast, higher educational attainment and current employment operated as stable protective factors, suggesting that education22 and work [22] not only improved health literacy and economic resources but also expanded social support networks that buffer the health costs of aging. These associations aligned with cumulative advantage theory and underscore how late-life health inequalities reflect long-term disparities in socioeconomic and social resource accumulation.

Second, the findings clarified the mechanism of digital inclusion as one of "limited direct effects but substantial indirect effects." Despite high factor loadings for digital access, device proficiency, and digital ability—and a clear positive effect of digital inclusion on social engagement—digital inclusion did not directly translate into better overall health. This mirrors evidence that where uptake of digital health services (eg, teleconsultation and health management apps) remains modest, digital skills training without integrated health education or health intervention yields limited direct health gains. Survey of Health, Ageing and Retirement in Europe (SHARE) data indicated that internet use had only a weak direct association with SRH. Its benefits are realized mainly through enhanced social connectedness and life satisfaction [35]. US studies similarly suggested that digital technology improved health indirectly via enriched social interaction, greater information access, and increased cognitive stimulation. Moreover, internet use benefited older adults' mental health, with social capital partially mediating the effect [36]. In China, a persistent digital divide, especially in rural and low-education groups, further dampened any direct health effect [37].

What was notable in this study was the negative direct path from digital inclusion to overall health. This may reflect information overload, increased exposure to health misinformation [16], health perfectionism, driven social comparison, and health anxiety, all of which may exacerbate anxiety and offset positive spillovers transmitted through social engagement and mental health [25]. These findings emphasize that digitalization is not inherently health-promoting, while

XSL•FO
RenderX

containing both enabling and burdensome properties. Policy should therefore prioritize not only access but also content quality, digital safety, and age-friendly platform design, embedding digital literacy within broader behavioral and psychosocial intervention strategies to unlock the genuine health-promoting potential of digital inclusion.

Third, the study verified that social engagement mediated between digital inclusion and perceived health, while mental health was the core bridging mechanism. This underscores the pivotal role of mental health in shaping health perceptions. The mechanism suggested is that positive experiences derived from social support, activity participation, and social interaction are re-processed psychologically—via reduced depressive symptoms, greater social adaptation, and higher life satisfaction—before they translate into better SRH. Thus, social engagement serves as an upstream facilitator of psychological well-being, whereas mental health functions as the final common pathway through which digital inclusion affects overall health.

This mechanism accorded with the WHO Active Ageing chain of "social connectedness → mental health → overall health" [38] and was supported by diverse empirical evidence. For example, the longitudinal Chinese Longitudinal Healthy Longevity Survey showed that the lack of participation predicted declines in mental health, whereas active participation was linked to better psychological well-being [10]; leisure-oriented participation could improve mental health broadly and significantly [39]. Additionally, the Australian Longitudinal Study on Women's Health found that persistent or rising loneliness was closely tied to deterioration in mental health–related quality of life, highlighting the long-term protective value of social ties [40]. Korean aging panel studies showed that multimorbidity and disability depressed SRH, whereas social engagement partially mitigated these effects through psychological pathways [41]. In rural China, research on "empty-nest" older adults indicated that internet use directly raised life satisfaction and indirectly improved mental health by boosting participation, thereby further elevating health perceptions [42]. Moreover, a recent cross-national systematic review of digital interventions (randomized controlled trials) showed that online socialization and digital interaction significantly reduced loneliness and depressive symptoms, indirectly enhancing subjective health and life satisfaction [25]. These findings indicated that the dual mediation of social engagement and mental health is the key mechanism through which digital inclusion delivers health benefits.

Although the SEM results revealed significant links among digital inclusion, social engagement, mental health, and SRH, these pathways should not be interpreted as causal. The cross-sectional design prevents verification of temporal ordering, and reverse or bidirectional influences remain possible—for example, older adults with better health or stronger social ties may be more inclined to adopt digital technologies. Unobserved factors, such as cognitive status or personality traits, may also contribute to the associations observed. Thus, the mediation pathways identified here reflect theoretical plausibility rather than confirmed causal processes. Longitudinal analyses using future waves of CLASS will be necessary to assess whether these mechanisms hold over time.

Finally, the study revealed cross-national differences. Compared with high-income Western settings, the health benefits of digital inclusion among Chinese older adults relied more on indirect pathways (via social engagement and mental health) than on direct effects. Variations in social structure, welfare systems, and digital health ecosystem maturity likely contribute to these differences. In Europe and North America, digital tools were more deeply embedded in daily life and were closely integrated with community health services, social welfare, and family care, forming a "digital + in-person" support network that yielded more immediate perceived health gains [27]. In China, limited penetration of digital health services, uneven distribution of community digital resources, and the predominance of offline group activities among older adults meant that participation—although beneficial for mental health—had weaker direct links to concrete health behaviors (eg, screening, exercise, and self-management) [43]. Gaps also remained in the availability and quality of mental health services. Although mental health sat at the center of the causal chain, under-identification of emotional disorders in community settings and limited access to age-appropriate professional interventions likely attenuated its potential health benefits. By contrast, many high-income countries have embedded mental health supports within primary care and community systems, amplifying the health gains from both participation and digital technology.

Limitations

This study had several limitations. First, SRH was a subjective measure without objective clinical or biomarker data. As it was susceptible to cognitive frames and momentary affect, potentially biasing associations and attenuating the precision of mediation estimates among digital inclusion, social engagement, and mental health. Second, the negative direct effect of digital inclusion on SRH may partly reflect perceptual or cognitive mechanisms—such as upward social comparison, exposure to idealized health information online, or heightened health anxiety—rather than a true adverse effect of digital engagement. This analysis was unable to fully disentangle these psychological influences from structural digital factors. Third, this study relied on cross-sectional data; thus, the directionality of associations cannot be conclusively established. The mediation effects suggested by SEM may partially reflect reverse causality or unmeasured confounding. Future research using multiwave CLASS data will allow for longitudinal validation of these mechanisms and more rigorous evaluations of causal ordering.

Implications

On the basis of these findings, several implications emerge for policy, practice, and the design of future interventions. First, digital inclusion initiatives must extend beyond access provision to emphasize skills, usability, and sustained engagement, particularly for rural, low-education, and less digitally confident older adults. Integrating digital literacy with health education and psychosocial support may help ensure that digital engagement translates into meaningful health benefits rather than information overload or anxiety. Second, expanding opportunities for diverse and age-appropriate social

XSL•FO
RenderX

participation—both online and offline—is essential, as participation serves as an upstream determinant of psychological well-being and a conduit through which digital inclusion exerts its influence on health. Community organizations, social service agencies, and primary care institutions should co-develop engagement platforms that are socially meaningful, culturally appropriate, and accessible to older adults with varying functional capacities.

Conclusions

This study shows that digital inclusion shapes older adults' health perceptions primarily through indirect pathways, operating first by enhancing social engagement and subsequently by improving mental health. While demographic and health-related vulnerabilities such as aging, chronic diseases, and functional limitations remain dominant predictors of poorer SRH, digital competence and meaningful participation represent actionable levers that can mitigate health disadvantages in later life. Mental health emerged as the pivotal link in the causal chain, underscoring the need to position psychological well-being at the center of healthy aging strategies. These findings highlight the importance of moving beyond infrastructure-focused digital policies to approaches that build skills, support sustained social participation, and integrate mental health services within both community and digital ecosystems. As population aging and digital transformation continue to accelerate, coordinated interventions that align technological, social, and psychological resources will be essential for promoting active, equitable, and healthy aging.

Funding

This work was supported by the Ph.D. research initiation project of Anhui Medical University (0501054201) and the Youth Social Science Research Program of the Education Department of Anhui Province (2025AHGXSK40223).

Data Availability

The data for this study were taken from the China Longitudinal Aging Social Survey (CLASS) database. Data are available in the public domain.

Authors' Contributions

BZ and XW contributed to the design, data analysis, and writing of the paper. RC and EWM revised the manuscript. BZ contributed to funding acquisition, quality control, and data processing and revised the manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. China's population shrinks again. Caixin Global. 2025. URL: https://www.caixinglobal.com/2025-01-17/chinas-population-shrinks-again-102280359.html [accessed 2026-01-24]
2. Upasen R, Saengpanya W, Awae W, Prasitvej P, Stithyudhakarn S. Understanding how older adults sustain mental well-being when facing challenging in later life. Acta Psychol (Amst) 2025 Sep;259:105293. [doi: 10.1016/j.actpsy.2025.105293] [Medline: 40680646]
3. Decade of healthy ageing: baseline report. : World Health Organization; 2021 URL: https://www.who.int/publications/i/item/9789240017900 [accessed 2026-01-24]
4. Zhou J, Yang Y, Li S, et al. Patterns of social participation among older adults and their association with self-rated health: the mediating role of activities of daily living. Aging Clin Exp Res 2025 Jul 19;37(1):221. [doi: 10.1007/s40520-025-03131-3] [Medline: 40682656]
5. Rook KS, Charles ST. Close social ties and health in later life: strengths and vulnerabilities. Am Psychol 2017 Sep;72(6):567-577. [doi: 10.1037/amp0000104] [Medline: 28880103]
6. Joint issuance by the Ministry of Civil Affairs and 18 other departments of the "guiding opinions on supporting older adults' social participation to promote active ageing" [Title in Chinese]. The Central People's Government of the People's Republic of China. 2025. URL: https://www.gov.cn/lianbo/bumen/202505/content_7025054.htm [accessed 2026-01-24]
7. The 55th statistical report on China's internet development. : China Internet Network Information Center; 2025 URL: https://www.cnnic.com.cn/IDR/ReportDownloads/202505/P020250514564119130448.pdf [accessed 2026-01-24]
8. Zhou Y, Bai Y, Wang J. The impact of internet use on health among older adults in China: a nationally representative study. BMC Public Health 2024 Apr 17;24(1):1065. [doi: 10.1186/s12889-024-18269-4] [Medline: 38632588]
9. Xue Z, Liu C. The effect of internet use on the health of middle-aged and older adults: evidences from China general social survey. BMC Public Health 2024 Oct 24;24(1):2949. [doi: 10.1186/s12889-024-20313-2] [Medline: 39449133]
10. Du M, Dai W, Liu J, Tao J. Less social participation is associated with a higher risk of depressive symptoms among Chinese older adults: a community-based longitudinal prospective cohort study. Front Public Health 2022;10:781771. [doi: 10.3389/fpubh.2022.781771] [Medline: 35223728]

XSL•FO
RenderX

11. Hu J, Tan C. The impact of internet use on the subjective well-being of older adults: the mediating role of mental health. PLoS ONE 2025;20(10):e0335985. [doi: 10.1371/journal.pone.0335985] [Medline: 41171891]
12. Zhang S, Dong S, Yu Z, Bi S, Wen W, Li J. Internet usage, social participation, and depression symptoms among middle-aged and older adult Chinese individuals before and during the COVID-19 pandemic: evidence from the China Health and Retirement Longitudinal Observational Study. J Med Internet Res 2025 Sep 22;27:e67039. [doi: 10.2196/67039] [Medline: 41024412]
13. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
14. Li PQ, Sun GY, Zhao XX, Hu ZX, Gan KP. Internet use and self-rated health among Chinese older adults: the role of social engagement and social support. Front Public Health 2025;12:1504940. [doi: 10.3389/fpubh.2024.1504940] [Medline: 39975695]
15. Lyu C, Siu K, Xu I, Osman I, Zhong J. Social isolation changes and long-term outcomes among older adults. JAMA Netw Open 2024 Jul 1;7(7):e2424519. [doi: 10.1001/jamanetworkopen.2024.24519] [Medline: 39046736]
16. Zhong F, Gu C. The impact of health information echo chambers on older adults avoidance behavior: the mediating role of information fatigue and the moderating role of trait mindfulness. Front Psychol 2024;15:1412515. [doi: 10.3389/fpsyg.2024.1412515] [Medline: 39228876]
17. Hamplová D, Klusáček J, Mráček T. Assessment of self-rated health: the relative importance of physiological, mental, and socioeconomic factors. PLoS ONE 2022;17(4):e0267115. [doi: 10.1371/journal.pone.0267115] [Medline: 35436307]
18. Yang H, Deng Q, Geng Q, et al. Association of self-rated health with chronic disease, mental health symptom and social relationship in older people. Sci Rep 2021 Jul 19;11(1):14653. [doi: 10.1038/s41598-021-94318-x] [Medline: 34282246]
19. Yang R, Gao S, Jiang Y. Digital divide as a determinant of health in the U.S. older adults: prevalence, trends, and risk factors. BMC Geriatr 2024 Dec 21;24(1):1027. [doi: 10.1186/s12877-024-05612-y] [Medline: 39709341]
20. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
21. Roquet A, Martinelli P, Lampraki C, Jopp DS. Internet use as a moderator of the relationship between personal resources and stress in older adults: cross-sectional study. JMIR Aging 2024 Jul 19;7:e52555. [doi: 10.2196/52555] [Medline: 39028547]
22. Mu A, Liu Z. Assessing the impact of internet skills on depressive symptoms among Chinese middle-aged and older adults: cross-sectional instrumental variables analysis. JMIR Aging 2024 Mar 21;7:e50880. [doi: 10.2196/50880] [Medline: 38533782]
23. Chen XL, Li J, Sun SN, et al. Association between daily internet use and intrinsic capacity among middle-aged and older adults in China: large prospective cohort study. J Med Internet Res 2024 Nov 12;26:e54200. [doi: 10.2196/54200] [Medline: 39531641]
24. Sardareh M, Matlabi H, Shafiee-Kandjani AR, Bahreini R, Mohammaddokht S, Azami-Aghdash S. Interventions for improving health literacy among older people: a systematic review. BMC Geriatr 2024 Nov 5;24(1):911. [doi: 10.1186/s12877-024-05522-z] [Medline: 39501193]
25. Shi X, Zhang J, Wang H, Luximon Y. The effectiveness of digital interactive intervention on reducing older adults' depressive and anxiety symptoms: a systematic review and meta-analysis. Gerontology 2024;70(9):991-1012. [doi: 10.1159/000539404] [Medline: 38857587]
26. Qiu YF, Wu M, Liu JL, et al. Effectiveness of digital intelligence interventions on depression and anxiety in older adults: a systematic review and meta-analysis. Psychiatry Res 2024 Dec;342:116166. [doi: 10.1016/j.psychres.2024.116166] [Medline: 39243439]
27. Bevilacqua R, Stara V, Amabili G, et al. e-VITA study protocol: EU-Japan virtual coach for smart aging. Front Public Health 2024;12:1256734. [doi: 10.3389/fpubh.2024.1256734] [Medline: 38544729]
28. Cui S, Yu Y, Dong W, et al. Are there gender differences in the trajectories of self-rated health among Chinese older adults? An analysis of the Chinese Longitudinal Healthy Longevity Survey (CLHLS). BMC Geriatr 2021 Oct 18;21(1):563. [doi: 10.1186/s12877-021-02484-4] [Medline: 34663221]
29. Jin YA, Zhao MH. Internet use and active ageing among older adults in China: evidence from the 2016 Chinese Longitudinal Aging Social Survey [Article in Chinese]. Ren Kou Xue Kan 2019;41(6):44-55. [doi: 10.16405/j.cnki.1004-129X.2019.06.004]
30. Wang Y, Shen H, Hurwicz M. Measuring depressive symptoms with 8-item and 9-item CES-D: Does race/ethnicity matter? Innov Aging 2017 Jul 1;1(suppl_1):1149-1149. [doi: 10.1093/geroni/igx004.4196]
31. Kannan VD, Veazie PJ. US trends in social isolation, social engagement, and companionship—nationally and by age, sex, race/ethnicity, family income, and work hours, 2003-2020. SSM Popul Health 2023 Mar;21:101331. [doi: 10.1016/j.ssmph.2022.101331] [Medline: 36618547]
32. Xu T, Huang Z, Li B, et al. Association between home and community-based services utilization and self-rated health among Chinese older adults with chronic diseases: evidence from the 2018 China Health and Retirement Longitudinal Study. BMC Public Health 2024 Jan 8;24(1):117. [doi: 10.1186/s12889-023-17535-1] [Medline: 38191440]

XSL•FO
RenderX

33. Gao Y, Zeng J, Liao Z, Yang J. Social transition, socioeconomic status and self-rated health in China: evidence from a national cross-sectional survey (CGSS). Front Public Health 2024 Jun 6;12:1359609. [doi: 10.3389/fpubh.2024.1359609]
34. Jiao K, Sun Y. Selection and connection: the impact of internet use on the depression of Chinese older adults. Front Public Health 2025;13:1524276. [doi: 10.3389/fpubh.2025.1524276] [Medline: 40177088]
35. Fu J, Fu S, Wang X, Wang X. The association between Internet use and intrinsic capacity among older adults in China: the mediating role of social participation. Public Health Nurs 2024;41(6):1600-1611. [doi: 10.1111/phn.13405] [Medline: 39221697]
36. Du X, Liao J, Ye Q, Wu H. Multidimensional internet use, social participation, and depression among middle-aged and elderly Chinese individuals: nationwide cross-sectional study. J Med Internet Res 2023 Aug 30;25:e44514. [doi: 10.2196/44514] [Medline: 37647119]
37. Vivion M, Reid V, Dubé E, Coutant A, Benoit A, Tourigny A. How older adults manage misinformation and information overload—a qualitative study. BMC Public Health 2024 Mar 21;24(1):871. [doi: 10.1186/s12889-024-18335-x] [Medline: 38515081]
38. Jeon S, Charles ST. Internet-based social activities and cognitive functioning 2 years later among middle-aged and older adults: prospective cohort study. JMIR Aging 2024 Dec 10;7:e63907. [doi: 10.2196/63907] [Medline: 39656544]
39. Wang J, Xu J, Nie Y, et al. Effects of social participation and its diversity, frequency, and type on depression in middle-aged and older persons: evidence from china. Front Psychiatry 2022;13:825460. [doi: 10.3389/fpsyt.2022.825460] [Medline: 35546944]
40. HaGani N, Owen K, Clare PJ, Merom D, Smith BJ, Ding D. Long-term elevated levels of loneliness are linked to lower health-related quality of life in middle-aged Australian women. Commun Psychol 2025 May 26;3(1):85. [doi: 10.1038/s44271-025-00264-z] [Medline: 40419747]
41. Ryu D. The impact of physical functionality and activity level on the self-rated health status of older adults living alone: an analysis of the mediating effect of social engagement. Geriatr Nurs 2025;63:464-469. [doi: 10.1016/j.gerinurse.2025.03.035] [Medline: 40262405]
42. Liu Y, Wang Y, Yang Z, et al. Internet use and life satisfaction among empty nesters in rural areas of Yangzhou: based on propensity score matching and mediating effect model. BMC Geriatr 2025 Feb 25;25(1):128. [doi: 10.1186/s12877-025-05788-x] [Medline: 40000969]
43. Ye J, He L, Beestrum M. Implications for implementation and adoption of telehealth in developing countries: a systematic review of China's practices and experiences. NPJ Digit Med 2023 Sep 18;6(1):174. [doi: 10.1038/s41746-023-00908-6] [Medline: 37723237]

Abbreviations

ADL: basic activities of daily living
CLASS: China Longitudinal Aging Social Survey
ELSA: English Longitudinal Study of Ageing
HRS: Health and Retirement Study
SEM: structural equation modeling
SHARE: Survey of Health, Ageing and Retirement in Europe
SRH: self-rated health
WHO: World Health Organization

Edited by A Garcia; submitted 27.Aug.2025; peer-reviewed by SN Sun, Y Zhang, Z Wang; revised version received 17.Dec.2025; accepted 05.Jan.2026; published 17.Mar.2026.

Please cite as:
Zhao B, Wang X, Chen R, Nam EW
Digital Inclusion Pathways for Older Chinese Adults in the Context of Active Aging: Secondary Analysis of 2023 China Longitudinal Aging Social Survey Data
JMIR Aging 2026;9:e83078
URL: https://aging.jmir.org/2026/1/e83078
doi:10.2196/83078

© Bo Zhao, Xiaoyan Wang, Ren Chen, Eun Woo Nam. Originally published in JMIR Aging (https://aging.jmir.org), 17.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Broadband Access, Internet Use, and General Cognition in Middle-Aged and Older Adults: Longitudinal Study

Fangfang Cheng[1], MSc; Li Cao[2], PhD

[1]Student Affairs Office, Inner Mongolia Medical University, Hohhot, China

[2]Information Technology Department, Inner Mongolia Medical University, 5th floor, Teaching Building, Xinhua Road, Hui Min District, Hohhot, China

Corresponding Author:
Li Cao, PhD
Information Technology Department, Inner Mongolia Medical University, 5th floor, Teaching Building, Xinhua Road, Hui Min District, Hohhot, China

Abstract

Background: Internet use may support cognitive health in aging populations, but the digital divide in access and use might worsen inequalities.

Objective: This study aimed to examine whether internet use mediates the long-term relationship between broadband access and cognitive function among middle-aged and older adults in China.

Methods: Using 3-wave data (2015 - 2020) from the China Health and Retirement Longitudinal Study (N=6063), we applied parallel-process latent growth curve models to analyze trajectories of broadband access, internet use, and general cognition, adjusting for sociodemographic and time-varying covariates.

Results: Baseline broadband access predicted initial internet use (β=0.453; *P*<.001), and broadband access increases strongly forecasted rises in internet use (β=0.625; *P*<.001). Increased internet use was associated with improved baseline general cognition (β=0.470; *P*<.001) and more significant cognitive improvements over time (β=0.444; *P*=.006). Older age, lower educational level, and rural residence were correlated with lower initial digital engagement; however, participants from rural areas demonstrated a greater adoption rate. Health status and social activities initially positively influenced outcomes; however, their impact diminished by the final wave.

Conclusions: These findings demonstrate that internet use mediates the longitudinal relationship between broadband access and cognitive resilience. The observed cascading growth trajectories suggest that gradually bridging the digital divide may help mitigate cognitive decline.

(JMIR Aging 2026;9:e75947) doi:10.2196/75947

KEYWORDS

digital divide; cascade; cognitive function; trajectory; parallel-process latent growth curve modeling; PP-LGCM

Introduction

In the digital age, the internet is essential for accessing information, maintaining social connections, and managing daily tasks [1]. Approximately 2.6 billion people worldwide (32% of the population) remain offline as of 2024 [2]. The benefits of digital innovation are unevenly distributed, leading to a persistent "digital divide" [3]. The definition of "digital divide" has shifted from a simple binary notion of connectivity in the 1990s [4] to a 3-level framework of coverage and access, use, and real-world consequences [5,6], with significant effects on individual health and societal well-being [7,8].

Cognitive decline is one of the most substantial challenges faced by the older adult population in health and social care systems [9]. Cognitive decline, which can start as early as the age of 45 years [10], is a significant risk factor for disability, dementia, and death [11-13]. A notable increase in cognitive decline after the age of 65 years has been documented [14]. China, with the world's largest older adult population, faces an increasing challenge: almost 20% of adults aged >60 years have mild cognitive impairment, with 6% developing dementia each year [15-18].

Since 2015, China's Internet Plus initiative has rapidly improved digital infrastructure, especially broadband access [19,20]. Despite this progress, age remains a significant factor of inequality: older adults, particularly those living in rural areas, with less education, or lacking digital skills, often face exclusion from meaningful internet use [21-23]. A 2024 cohort study in Beijing revealed that nearly half of older adults experience an "internet use divide," which correlates with reduced processing speed and overall cognitive performance [24].

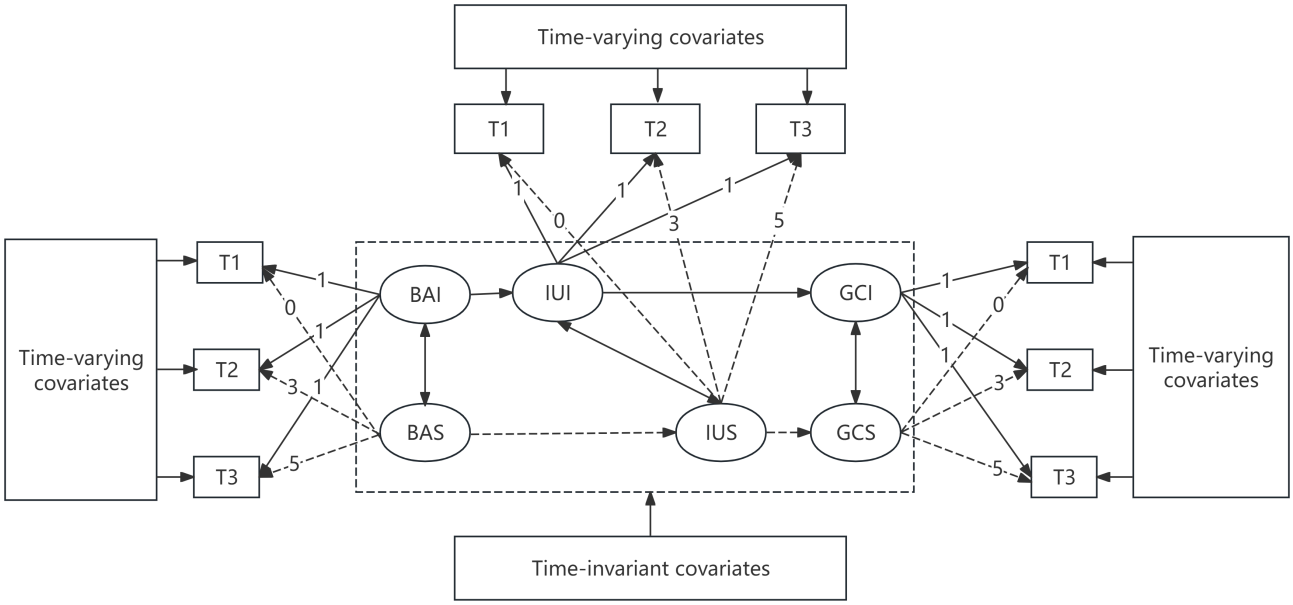
XSL•FO
RenderX

This challenge aligns with a vital global goal: the United Nations Decade of Healthy Ageing (2021 - 2030) [25]. The first progress report [26] highlights that, although many countries, including China, have developed laws and infrastructure for age-friendly environments, more than two-thirds still lack sufficient resources to fully implement their action plans. This problem is especially severe in low- and middle-income areas, where nearly 80% of the world's older adults are expected to live by 2050 [26,27]. The World Bank states that increasing access alone is insufficient to bridge the digital divide as technological progress may exclude the targeted populations [28]. Therefore, digital inclusion should be regarded as an ongoing process incorporated into age-inclusive communities and an essential component of healthy aging [29].

Empirical studies increasingly suggest that internet use—not merely access—is an active factor influencing cognitive outcomes in later stages of life. An ecological analysis found that higher broadband coverage correlates with lower prevalence of cognitive disorders, suggesting that digital infrastructure may help mitigate cognitive health disparities among older adults [30]. However, broadband access is commonly associated with an area's economic resources, enabling internet use and social interactions. For instance, one longitudinal study in Europe found that older adults who regularly used the internet exhibited slower cognitive decline over 2 years independent of baseline connectivity [31]. Similarly, research using nationally representative data from the United States demonstrated that changes in internet use had asymmetric effects on cognitive decline in older adults: initiating use slowed decline, whereas discontinuing use accelerated it [32]. Further supporting this distinction, another investigation using data from the China Family Panel Study found that the frequency and variety of internet activities positively contributed to cognitive functions [33]. These findings align with 2 complementary theoretical perspectives. First, a widely cited model of digital inequality posits that material access to technology constitutes only the initial layer of inclusion; its benefits emerge only when individuals acquire the skills, motivation, and opportunities to engage meaningfully with digital tools [5]. Second, the cognitive reserve hypothesis proposes that participation in cognitively stimulating activities builds neural resilience against age-related cognitive decline [34]. Thus, internet use—as a dynamic, interactive, and often socially embedded behavior—may serve as the behavioral bridge through which structural access translates into sustained cognitive benefit.

We hypothesize that internet use mediates the longitudinal association between household broadband access and cognitive function among middle-aged and older adults in China such that growth in access predicts increased use, which in turn predicts more favorable cognitive trajectories over time as shown in Figure 1. However, most existing research examines only isolated links. These include the relationships between broadband access and cognitive outcomes, and between technology use patterns and cognitive development. Most of this research relies on cross-sectional data. Few studies use longitudinal designs capable of modeling how disparities in digital infrastructure spread through behavior to affect cognitive aging. Using China's rapid digitization and population aging as a natural experiment, this study applied parallel-process latent growth curve modeling (PP-LGCM) with nationally representative panel data to estimate trajectories of broadband access, internet use, and general cognition simultaneously and test whether internet use changes mediate the relationship between digital infrastructure and cognitive outcomes among middle-aged and older Chinese adults.

Figure 1. The hypothesized model of the digital divide across 3 waves. In this model, intercept factors—broadband access intercept (BAI), internet use intercept (IUI), and general cognition intercept (GCI)—capture baseline levels of broadband access, internet use, and cognitive function, respectively. Slope factors—broadband access slope (BAS), internet use slope (IUS), and general cognition slope (GCS)—represent the rates of change in these outcomes over time. T1: 2015; T2: 2018; T3: 2020.



XSL•FO
RenderX

Methods

Ethical Considerations

This study conducted a secondary analysis of deidentified, publicly available China Health and Retirement Longitudinal Study (CHARLS) data, exempt from additional ethics review. The original CHARLS was approved by Peking University's Biomedical Ethics Review Committee (IRB00001052-11015), obtained written informed consent from all participants, and complied with the Declaration of Helsinki and China's Personal Information Protection Law. All procedures adhere to CHARLS protocols.

Participants

Participants were enrolled in 3 follow-ups of the CHARLS from 2015 to 2020. It is a nationally representative sample of the middle-aged and older adult population (aged ≥45 years) in China, a community-based longitudinal study conducted by the China Center for Economic Research at Peking University [35]. The CHARLS used a multistage, probability proportional to size sampling technique to select participants. The primary sampling units were administrative villages in rural areas and urban neighborhoods. The CHARLS data can be accessed through its official website [36]. It is designed to better understand the socioeconomic determinants and consequences of aging.

Data were drawn from 3 waves (2015 [T1], 2018 [T2], and 2020 [T3]) of the CHARLS. The baseline cohort included 23,459 participants. Of these 23,459 participants, we excluded individuals aged <45 years or >100 years and those with household consumption per capita equal to 0 at baseline (n=10,026, 42.7%), yielding 13,433 (57.3%) eligible respondents. From this group, we further excluded those with missing data on broadband access, internet use, or cognitive function (362/13,433, 2.7%) and those with a total cognitive score of 0 (143/13,433, 1.1%). An additional 1.1% (153/13,433) of the participants were excluded due to missing values for key time-invariant covariates (employment status, self-rated health, or household consumption per capita) at baseline. Of the remaining 12,775 participants with at least one valid cognitive measure, 6712 (52.5%) were lost to follow-up across the 3 waves, resulting in a final analytical sample of 6063 (47.5%) participants with sufficient longitudinal data for modeling.

Measurements

Dependent Variables

Cognitive function assessment tools included an adapted Chinese version of the Mini-Mental State Examination [37]. Four specific dimensions were assessed at each wave: orientation, episodic memory, calculation ability, and constructability. Orientation was measured by naming the date (day, month, season, and year) and day of the week (score range 0-5). For episodic memory, participants were first asked to immediately repeat as many as possible of 10 Chinese nouns that were read to them and then to recall the same list 5 minutes later (score range 0-10). Calculation ability was evaluated using the serial sevens subtraction task, which required participants to subtract 7 from 100 and continue subtracting 7 from each subsequent number 5 times (score range 0-5). Constructability was evaluated through a drawing task that assesses visuospatial skills and motor coordination (score range 0-1). General cognition was the sum of cognitive scores (range 0-21) [14] and was considered the digital outcome.

Independent and Mediating Variables

The primary independent variable was household broadband access. The digital divide was operationalized using 2 indicators aligned with the multilevel model by van Deursen and van Dijk [5] and Scheerder et al [38], measured using the question "Does your residence have a broadband internet connection?" (yes=1; no=0). The mediating variable was internet use, measured using the question "Have you accessed the Internet in the past month?" (yes=1; no=0).

Covariates

There were time-invariant sociodemographic factors (age, sex, educational level, rural residence, marital status, and household consumption per capita) and time-varying characteristics (self-rated health, social activity participation, and employment status) included, all measured at each wave.

Statistical Analysis

R (version 4.1.0; R Foundation for Statistical Computing) was used for the purposes of descriptive and correlational analyses. Descriptive statistics were used to summarize the characteristics of the participants, specifically frequencies with percentages for categorical variables and medians with IQRs for continuous variables. The *lavaan* package in R was used to perform structural equation modeling. PP-LGCM was implemented to investigate the cascading effects of digital divides on trajectories of cognitive function. For each domain, the baseline level (intercept) and the rate of change over time (slope) were modeled [39]. The cascade effect framework suggests that broadband access influences internet use, which subsequently affects the intercepts and slopes of cognitive function. Full information maximum likelihood was applied to address missing data by leveraging all available information from each participant. The following fit indexes were evaluated: the comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Conventional thresholds stipulate that CFI and TLI values above 0.90 are acceptable, with values exceeding 0.95 indicating a good fit. For RMSEA and SRMR, values below 0.05 indicate a good fit, whereas values between 0.05 and 0.08 are acceptable [40]. In the conditional latent growth curve modeling and PP-LGCM analyses, demographic variables were incorporated as covariates.

Results

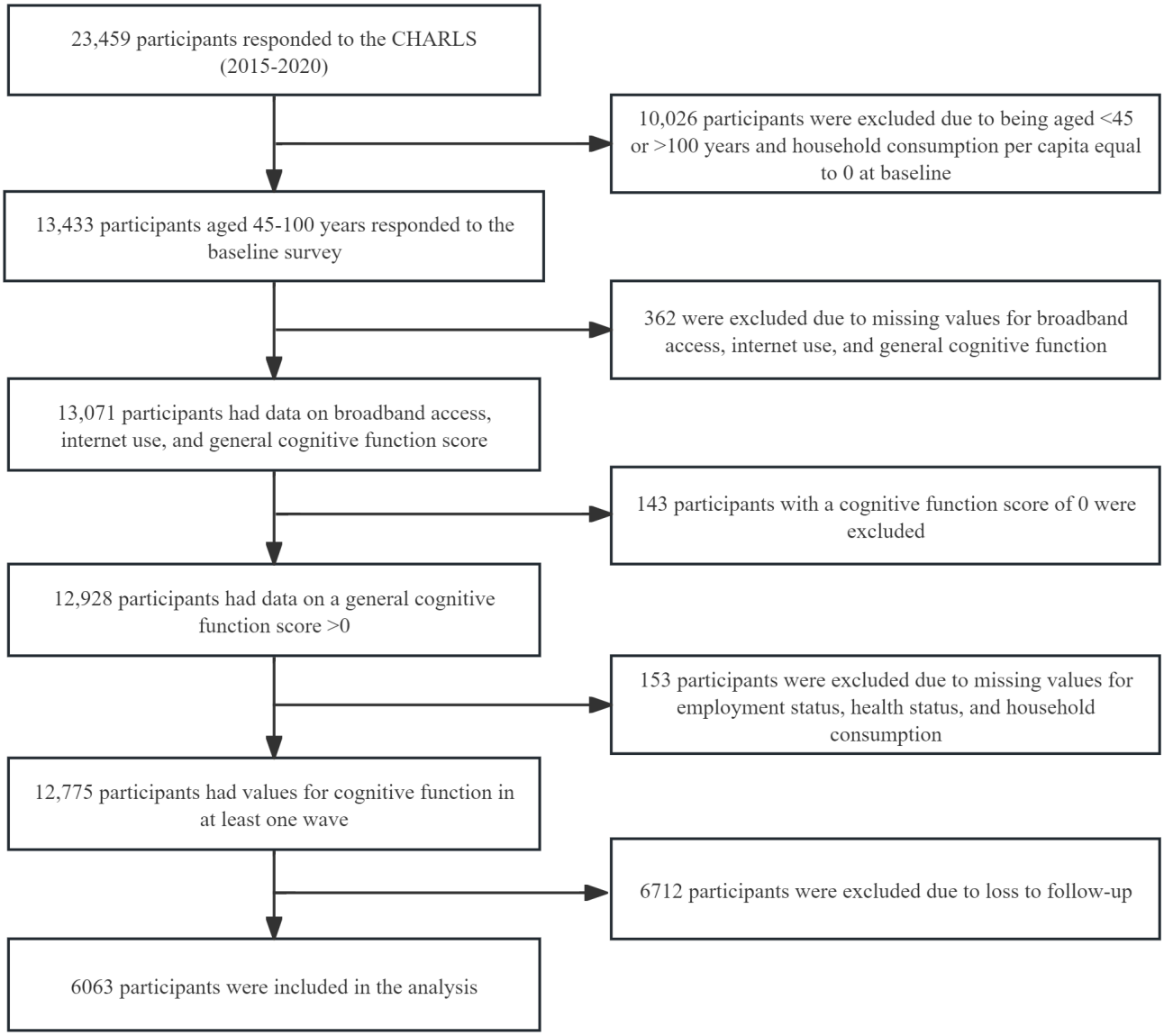
Descriptive Statistics

A total of 6063 Chinese middle-aged and older adults were included in this study (Figure 2). Table 1 shows significant sociodemographic disparities in digital access and use. Individuals with broadband access and internet use were significantly younger, more educated, and more likely to reside in urban areas; reported better health; engaged in social

XSL•FO
RenderX

activities; and had a higher household consumption (*P*<.001 in all cases). Men were more likely to be internet users (*P*<.001) but not among broadband owners (*P*=.79). Marital status and employment status showed minimal or no association with internet use (*P*=.10 and *P*=.29, respectively), although both differed slightly by broadband subscription (*P*<.001). Those who had broadband access and internet use also exhibited significantly higher cognitive function scores (median 15.5, IQR 14.5-16.5 vs median 12.5, IQR 10-14.5; *P*<.001), although the difference for broadband access was less pronounced (median 14.0, IQR 12-15.5 vs median 12.5, IQR 9.5-14.5; *P*<.001).

Figure 2. Flowchart of participant selection. CHARLS: China Health and Retirement Longitudinal Study.



XSL•FO
RenderX

Table . Baseline characteristics of participants by digital divide status. The rank sum test was used for age and general cognition, and the chi-square test was used for other variables. Household consumption per capita was divided into 4 quartiles based on the 25th percentile.

Variable	No broadband access (n=4247)	Broadband access (n=1816)	*P* value	No internet use (n=5545)	Internet use (n=518)	*P* value
Age (years), median (IQR)	59 (52-65)	53 (49-59)	<.001	58 (51-64)	51 (47-56)	<.001
Male sex, n (%)			.79			<.001
No	2022 (47.6)	872 (48.0)		2708 (48.8)	186 (35.9)	
Yes	2225 (52.4)	944 (52.0)		2837 (51.2)	332 (64.1)	
Educational level, n (%)			<.001			<.001
Primary school and lower	1738 (40.9)	317 (17.5)		2048 (36.9)	7 (1.4)	
Middle school	1115 (26.3)	354 (19.5)		1419 (25.6)	50 (9.7)	
High school	964 (22.7)	642 (35.4)		1419 (25.6)	187 (36.1)	
College and higher	430 (10.1)	503 (27.7)		659 (11.9)	274 (52.9)	
Rural residence, n (%)			<.001			<.001
No	1285 (30.3)	1087 (59.9)		2000 (36.1)	372 (71.8)	
Yes	2962 (69.7)	729 (40.1)		3545 (63.9)	146 (28.2)	
Marital status, n (%)			<.001			.10
Unmarried	353 (8.3)	89 (4.9)		414 (7.5)	28 (5.4)	
Married	3894 (91.7)	1727 (95.1)		5131 (92.5)	490 (94.6)	
Health status, n (%)			<.001			<.001
Excellent	197 (4.6)	44 (2.4)		232 (4.2)	9 (1.7)	
Very good	724 (17.0)	197 (10.8)		888 (16.0)	33 (6.4)	
Good	2331 (54.9)	980 (54.0)		3048 (55.0)	263 (50.8)	
Fair	468 (11.0)	309 (17.0)		663 (12.0)	114 (22.0)	
Poor	527 (12.4)	286 (15.7)		714 (12.9)	99 (19.1)	
Social activity, n (%)			<.001			<.001
No	2142 (50.4)	668 (36.8)		2710 (48.9)	100 (19.3)	
Yes	2105 (49.6)	1148 (63.2)		2835 (51.1)	418 (80.7)	
Employment status, n (%)			<.001			.29
Not employed	1024 (24.1)	515 (28.4)		1418 (25.6)	121 (23.4)	
Employed	3223 (75.9)	1301 (71.6)		4127 (74.4)	397 (76.6)	
General cognition score (0-21), median (IQR)	12.5 (9.5-14.5)	14 (12-15.5)	<.001	12.5 (10-14.5)	15.5 (14.5-16.5)	<.001
Household consumption per capita, n (%)			<.001			<.001
Quartile 1	1267 (29.8)	250 (13.8)		1484 (26.8)	33 (6.4)	
Quartile 2	1134 (26.7)	382 (21.0)		1438 (25.9)	78 (15.1)	
Quartile 3	1022 (24.1)	493 (27.1)		1377 (24.8)	138 (26.6)	
Quartile 4	824 (19.4)	691 (38.1)		1246 (22.5)	269 (51.9)	

Bivariate Correlations Regarding the Digital Divide

The correlation matrix presented in Table 2 delineates the relationships among broadband access, internet use, and general cognition at 3 distinct time points. The findings demonstrate moderate stability within each construct over time, with autocorrelation coefficients ranging from 0.30 to 0.62. Conversely, cross-construct correlations were consistently weaker, typically below 0.40, thereby providing evidence of adequate discriminant validity among the constructs. This pattern of associations indicates that the 3 constructs may be

XSL•FO
RenderX

validly treated as separate latent variables in longitudinal modeling, such as latent growth curve analyses.

Table . Bivariate correlations (Spearman *r* and 2-tailed *P* value)[a].

Variable	Broadband access			Internet use			General cognition		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Broadband access at T1									
r	1	0.45	0.34	0.35	0.31	0.33	0.21	0.22	0.23
P value	—[b]	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Broadband access at T2									
r	0.45	1	0.48	0.24	0.30	0.36	0.18	0.21	0.22
P value	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001	<.001
Broadband access at T3									
r	0.34	0.48	1	0.19	0.23	0.38	0.18	0.21	0.21
P value	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001
Internet use at T1									
r	0.35	0.24	0.19	1	0.44	0.30	0.19	0.19	0.20
P value	<.001	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001
Internet use at T2									
r	0.31	0.30	0.23	0.44	1	0.40	0.22	0.23	0.24
P value	<.001	<.001	<.001	<.001	—	<.001	<.001	<.001	<.001
Internet use at T3									
r	0.33	0.36	0.38	0.30	0.40	1	0.31	0.33	0.34
P value	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001	<.001
General cognition at T1									
r	0.21	0.18	0.18	0.19	0.22	0.31	1	0.58	0.59
P value	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001
General cognition at T2									
r	0.22	0.21	0.21	0.19	0.23	0.33	0.58	1	0.62
P value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001
General cognition at T3									
r	0.23	0.22	0.21	0.20	0.24	0.34	0.59	0.62	1
P value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—

[a]T1 to T3 refer to assessment time points 1 (2015), 2 (2018), and 3 (2020).

[b]Not applicable.

Conditional Latent Growth Curve Models for Each Element of the Digital Divide

The fit indexes presented in Table 3 strongly support the measurement models for each domain. Overall, these results indicate that the latent construct representations were adequate over time, confirming the validity of the developmental trajectories studied.

Table 4 shows that sociodemographic factors influenced initial levels and changes over time in broadband access, internet use, and cognitive function. Older age correlated with lower baseline levels (β=–0.327 to –0.144) and slower growth in internet use (β=–0.356) and general cognition (β=–0.151). Male individuals reported slightly less broadband access (β=–0.032) but higher baseline internet use and general cognition (β=0.107), with steeper declines over time (β=–0.152 and –0.160). Higher educational level boosted initial general cognition (β=0.518) but predicted slower growth in internet use (β=–0.332) and slightly faster cognitive gains (β=0.123). Rural residents started lower (eg, broadband; β=–0.220) but showed accelerated growth in broadband access (β=0.170) and internet use (β=0.202). Health status and social activity predicted higher levels of digital inclusion early on. For example, at T2, better health status was associated with greater broadband access (β=0.066), increased internet use (β=0.081), and higher general cognition scores (β=0.039). Social activity demonstrated a stronger link

XSL•FO
RenderX

specifically with internet use (β=0.122). By T3, these effects were weakened. Employment status reduced broadband access (β=–0.062 to –0.041 across waves) but had minimal impact on internet use and a small positive association with general cognition (β=0.034 at T3).

Table . Model fit indexes for conditional latent growth curve models.

Variable	Chi-square (*df*)	CFI[a]	TLI[b]	RMSEA[c]
Broadband access	93.0 (24)	0.973	0.997	0.020
Internet use	115.3 (24)	0.931	0.991	0.024
General cognition	107.6 (24)	0.991	0.983	0.025

[a]CFI: comparative fit index.

[b]TLI: Tucker-Lewis index.

[c]RMSEA: root mean square error of approximation.

XSL•FO
RenderX

Table . Conditional latent growth curve models (N=6063).

Covariate	Broadband access, β	Internet use, β	General cognition, β
Time-invariant covariates ~ latent intercept			
Log (age)[a]	–0.327[b]	–0.322[b]	–0.144[b]
Male sex	–0.032	0.107[b]	0.107[b]
Educational level	0.240[b]	0.457[b]	0.518[b]
Rural residence	–0.220[b]	–0.185[b]	–0.144[b]
Marital status	0.049[c]	–0.008	0.033[d]
Log (household consumption)	0.205[b]	0.213[b]	0.060[b]
Time-invariant covariates ~ latent slope			
Log (age)	–0.034	–0.356[b]	–0.151[c]
Male sex	–0.008	–0.152[c]	–0.160[b]
Educational level	–0.042	–0.332[b]	0.123[d]
Rural residence	0.170[b]	0.202[b]	0.090
Marital status	–0.018	0.027	–0.025
Log (household consumption)	–0.091[b]	–0.179[b]	–0.013
Time-varying covariates ~ observed outcomes			
Self-rated health status			
T1[e]	0.037	0.073[c]	0.046[b]
T2[f]	0.066[b]	0.081[b]	0.039[b]
T3[g]	0.016	0.006	0.028[c]
Social activity			
T1	0.034[d]	0.117[b]	0.030[b]
T2	0.074[b]	0.122[b]	0.038[b]
T3	0.020	0.109[b]	0.027[b]
Employment status			
T1	–0.062[b]	–0.003	0.007
T2	–0.073[b]	–0.014	0.013
T3	–0.041	–0.019	0.034[b]

[a]Log-transformed using the natural logarithm (base e).
[b]*P*<.001.
[c]*P*<.01.
[d]*P*<.05.
[e]T1: 2015.
[f]T2: 2018.
[g]T3: 2020.

PP-LGCM Model for the Digital Divide

The PP-LGCM model demonstrated an excellent fit (χ^2_{97}=335.1; *P*<.001; CFI=0.976; TLI=0.991; RMSEA=0.020; SRMR=0.026). This finding revealed a significant mediating pathway connecting broadband access to cognitive function through internet use, as illustrated in Table 5 and Figure 3. Consistent with our hypothesized cascade model, baseline broadband access strongly predicted initial internet use (β=0.453; *P*<.001), and increases in broadband access over time forecasted steeper growth in internet use (β=0.625; *P*<.001). In turn, higher baseline internet use was associated with better initial cognition (β=0.470; *P*<.001), and greater increases in use were associated with slower cognitive decline (β=0.444;

XSL•FO
RenderX

P=.006). Although older age, lower educational level, and rural residence were associated with lower initial digital engagement, rural participants exhibited significantly faster adoption rates (broadband slope: β=0.163; internet slope: β=0.102; *P*=.10), suggesting a partial narrowing of the use gap. Educational level boosted broadband access (β=0.238), internet use (β=0.337), and general cognition (β=0.299; *P*<.001 in all cases); and older age negatively affected broadband access (β=–0.324) and internet use (β=–0.167; *P*<.001 in both cases) and slowed internet use growth (β=–0.338; *P*<.001), with no significant impact on cognitive slope. Men reported higher baseline internet use (β=0.119; *P*<.001) and general cognition (β=0.071; *P*<.001), although they experienced steeper declines in internet use over time (β=–0.149; *P*=.02). Household consumption predicted higher baseline broadband access (β=0.203) and internet use (β=0.116; *P*<.001 in both cases) but was negatively related to their slopes. Social activity had a consistent association with internet use, with β coefficients of 0.117 at T1, 0.122 at T2, and 0.109 at T3. Furthermore, it positively predicted cognitive function (β=0.037 at T1, β=0.040 at T2, and β=0.035 at T3). Health status demonstrated a positive relationship with both internet use and cognition at T1 and T2; however, this association did not persist at T3. Employment displayed a negative correlation with broadband access across all waves (β=–0.062 at T1, β=–0.073 at T2, and β=–0.041 at T3), indicating lower adoption rates among employed individuals. Its influence on cognition remained negligible.

XSL•FO
RenderX

Table . Parallel-process latent growth curve model (N=6063).

Outcome and predictor	β (SE; 95% CI)	*P* value
Internet use intercept		<.001
Broadband access intercept	0.453 (0.022; 0.410 to 0.496)	
Internet use slope		<.001
Broadband access slope	0.625 (0.067; 0.494 to 0.756)	
General cognition intercept		<.001
Internet use intercept	0.470 (0.027; 0.417 to 0.523)	
General cognition slope		.006
Internet use slope	0.444 (0.161; 0.128 to 0.760)	.006
Broadband access intercept		
Log (age)[a]	–0.324 (0.020; –0.363 to –0.285)	<.001
Male sex	–0.032 (0.017; –0.065 to 0.001)	.07
Educational level	0.238 (0.018; 0.203 to 0.273)	<.001
Rural residence	–0.218 (0.018; –0.253 to 0.183)	<.001
Marital status	0.049 (0.018; 0.014 to 0.084)	.008
Log (household consumption)	0.203 (0.017; 0.170 to 0.236)	<.001
Broadband access slope		
Log (age)	–0.033 (0.030; –0.092 to 0.026)	.27
Male sex	–0.008 (0.028; –0.063 to 0.047)	.78
Educational level	–0.040 (0.028; –0.095 to 0.015)	.15
Rural residence	0.163 (0.028; 0.108 to 0.218)	<.001
Marital status	–0.018 (0.028; –0.073 to 0.037)	.51
Log (household consumption)	–0.087 (0.026; –0.138 to –0.036)	.001
Internet use intercept		
Log (age)	–0.167 (0.032; –0.229 to –0.105)	<.001
Male sex	0.119 (0.025; 0.070 to 0.168)	<.001
Educational level	0.337 (0.029; 0.280 to 0.394)	<.001
Rural residence	–0.081 (0.026; –0.132 to –0.030)	.001
Marital status	–0.030 (0.028; –0.085 to 0.025)	.28
Log (household consumption)	0.116 (0.025; 0.067 to 0.165)	<.001
Internet use slope		
Log (age)	–0.338 (0.073; –0.481 to –0.195)	<.001
Male sex	–0.149 (0.061; –0.269 to –0.029)	.02
Educational level	–0.309 (0.070; –0.446 to –0.172)	<.001
Rural residence	0.102 (0.062; –0.019 to 0.223)	.10
Marital status	0.038 (0.064; –0.087 to 0.163)	.55
Log (household consumption)	–0.126 (0.061; –0.246 to –0.006)	.04
General cognition intercept		
Log (age)	–0.021 (0.022; –0.064 to 0.022)	.33
Male sex	0.071 (0.018; 0.036 to 0.106)	<.001
Educational level	0.299 (0.023; 0.254 to 0.344)	<.001
Rural residence	–0.037 (0.019; –0.074 to 0.000)	.046
Marital status	0.043 (0.018; 0.008 to 0.078)	.01

XSL•FO
RenderX

Outcome and predictor	β (SE; 95% CI)	*P* value
Log (household consumption)	–0.046 (0.018; –0.081 to –0.011)	.01
General cognition slope		
Log (age)	0.006 (0.094; –0.178 to 0.190)	.95
Male sex	–0.091 (0.066; –0.220 to 0.038)	.17
Educational level	0.269 (0.087; 0.098 to 0.440)	.002
Rural residence	0.003 (0.061; –0.117 to 0.123)	.96
Marital status	–0.040 (0.057; –0.152 to 0.072)	.48
Log (household consumption)	0.062 (0.066; –0.067 to 0.191)	.35
Broadband access at T1[b]		
Health status	0.037 (0.019; 0.000 to 0.074)	.06
Social activity	0.034 (0.017; 0.001 to 0.067)	.049
Employment status	–0.062 (0.019; –0.099 to –0.025)	.001
Broadband access at T2[c]		
Health status	0.066 (0.019; 0.029 to 0.103)	<.001
Social activity	0.074 (0.016; 0.043 to 0.105)	<.001
Employment status	–0.073 (0.020; –0.112 to –0.034)	<.001
Broadband access at T3[d]		
Health status	0.016 (0.019; –0.021 to 0.053)	.40
Social activity	0.020 (0.019; –0.017 to 0.057)	.23
Employment status	–0.041 (0.021; –0.082 to 0.000)	.05
Internet use at T1		
Health status	0.073 (0.026; 0.022 to 0.124)	.006
Social activity	0.117 (0.024; 0.070 to 0.164)	<.001
Employment status	–0.003 (0.037; –0.076 to 0.070)	.93
Internet use at T2		
Health status	0.081 (0.023; 0.036 to 0.126)	<.001
Social activity	0.122 (0.019; 0.085 to 0.159)	<.001
Employment status	–0.014 (0.028; –0.069 to 0.041)	.56
Internet use at T3		
Health status	0.006 (0.019; –0.031 to 0.043)	.75
Social activity	0.109 (0.015; 0.080 to 0.138)	<.001
Employment status	–0.019 (0.022; –0.062 to 0.024)	.33
General cognition at T1		
Health status	0.052 (0.013; 0.027 to 0.077)	<.001
Social activity	0.037 (0.012; 0.013 to 0.061)	.002
Employment status	0.006 (0.015; –0.023 to 0.035)	.67
General cognition at T2		
Health status	0.043 (0.013; 0.018 to 0.068)	.001
Social activity	0.040 (0.012; 0.016 to 0.064)	<.001
Employment status	0.006 (0.017; –0.027 to 0.039)	.67
General cognition at T3		
Health status	0.023 (0.013; –0.002 to 0.048)	.07

XSL•FO
RenderX

Outcome and predictor	β (SE; 95% CI)	*P* value
Social activity	0.035 (0.012; 0.011 to 0.059)	.003
Employment status	0.028 (0.015; –0.001 to 0.057)	.046

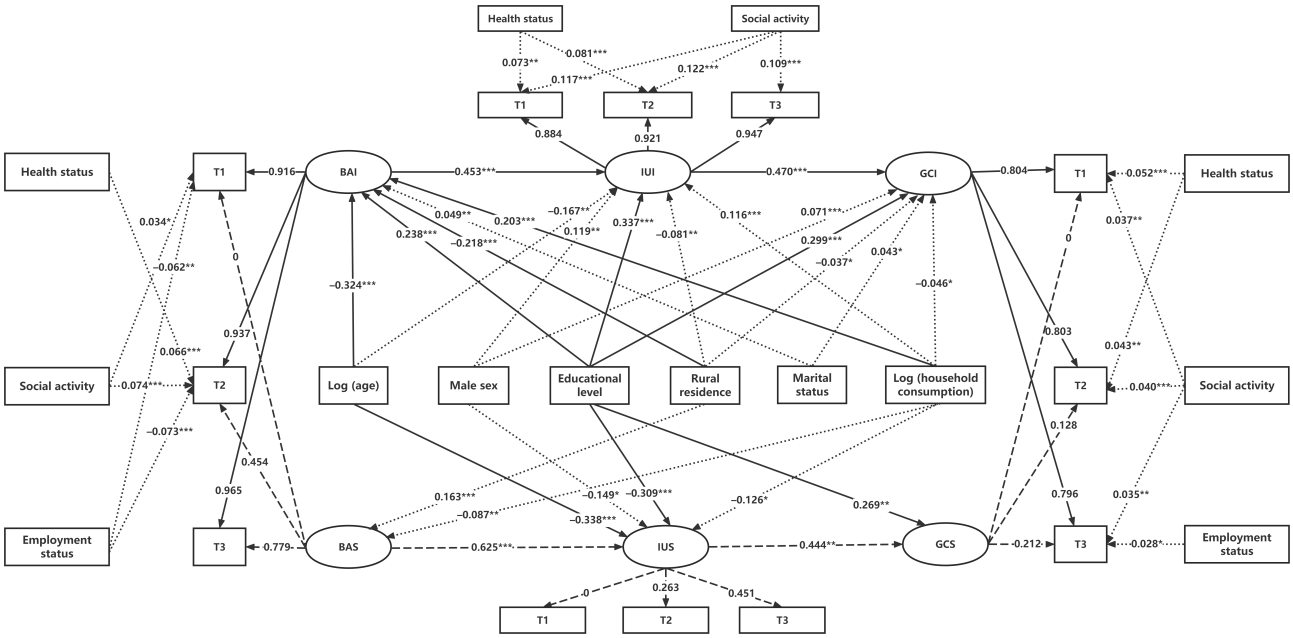
[a]Log-transformed using the natural logarithm (base e).

[b]T1: 2015.

[c]T2: 2018.

[d]T3: 2020.

Figure 3. Digital divide in cognitive function using parallel-process latent growth curve modeling. For latent variables, solid lines denote intercept relationships, while dashed lines represent slope relationships. Numerical values indicate loadings from latent variables to items. For regressions of sociodemographic variables, solid lines denote statistically significant paths with |β| ≥0.20; dotted lines denote statistically significant paths with |β| <0.20; nonsignificant paths are omitted. All the coefficients are standardized. ***P<.001; **P<.01; *P<.05; BAI: broadband access intercept; BAS: broadband access slope; GCI: general cognition intercept; GCS: general cognition slope; IUI: internet use intercept; IUS: internet use slope; T1: 2015; T2: 2018; T3: 2020.



Discussion

Principal Results

The findings of this study demonstrated that the cascading potential influence of household broadband access predicts internet use and subsequent codevelopment of this prediction, both of which are associated with a possible slower decline in cognitive trajectories among middle-aged and older Chinese adults. The PP-LGCM model confirmed the 3-level digital divide framework and extends the perspective from static views to a process-based approach, emphasizing how disparities in infrastructure could influence actual internet use and cognitive aging through codevelopmental pathways. Our findings substantially enhance the discussion of digital inclusion as a vital component of healthy aging, emphasizing the importance of not only physical access but also creating supportive environments that enable older adults to thrive.

This aligns with the first progress report of the United Nations Decade of Healthy Ageing, which underscores the vital role of digital engagement in sustaining intrinsic capacity in later life [26]. The report cautions that, despite investments in connectivity, the resources for engagement remain insufficient, especially among older adults in low- and middle-income countries. Van Dijk [41] similarly argues that access alone is insufficient; individuals require motivation, digital skills, and supportive opportunities to translate connectivity into benefit. Our results support this view: China's infrastructure expansion from 2015 to 2020 narrowed the rural-urban gap in broadband access, yet this did not translate into comparable gains in internet use or cognitive outcomes among rural respondents. Early advantages among urban and highly educated individuals likely reflect the Matthew effect [42,43], in which initial resources accumulate over time. However, what appears to be a cumulative advantage may stem from persistent conversion gaps [44,45]: the failure to translate access into meaningful use, especially in mid to late life, when cognitive resilience is more vulnerable [46,47].

Moreover, the association between internet use and cognitive resilience is plausibly mediated by neurocognitive mechanisms aligned with the cognitive reserve hypothesis [48,49]. Internet use is positively associated with cognitive stimulation and cognitive reserve [50,51]. Activities such as seeking health information, comparing services, or navigating digital interfaces engage executive functions, working memory, and attention [52,53]. Social relationship factors are correlated with cognitive

XSL•FO
RenderX

function and its decline, and engaging in social activities via calls or messaging may reduce social isolation, which is identified as a risk factor for cognitive decline [54,55]. The internet offers problem-solving opportunities through online tasks such as navigating websites, managing banking, or learning, which demand cognitive effort and decision-making, helping maintain mental agility [56] and thereby stimulating cognitive abilities [57]. Although our binary measure of internet use limits granularity, the findings are consistent with those of research linking online behaviors to delayed recall and processing speed [58], indicating that continuous engagement activates these neural pathways [59].

Structural inequalities shape both digital and cognitive trajectories [60]. Age is a predictor of disadvantage [23,43]: older adults tend to have less broadband access, lower internet use, and slower progress in these domains. Educational level functions as a lifelong enabler, increasing broadband access, internet use, and cognitive capacity [60]. Men in this study reported higher initial levels of internet use and general cognition. However, there is a significant decline in internet use over time, rather than in cognitive function, aligning with previous research [18,61], some of which is potentially attributable to cohort attrition in men. Household consumption predicted higher initial digital engagement but also correlated with slower growth, implying saturation among more advantaged individuals [22]. Self-rated health and social activity were associated with better outcomes in the initial wave but lost significance in later waves, potentially due to survivor bias or the stabilization of digital habits [32]. Employment status consistently showed a negative relationship with broadband access across waves, likely reflecting work-related constraints or measurement limitations.

While our PP-LGCM model supports a directional cascade from access to use to general cognition, we acknowledge the plausibility of reverse causation: individuals with higher baseline cognitive capacity may be more likely to adopt and sustain internet use [32]. Future studies using cross-lagged panel models or instrumental variables are recommended to further elucidate the causal direction.

Limitations

This study has several limitations. First, attrition (6712/12,775, 52.5%) may introduce bias if those lost to follow-up were disproportionately cognitively impaired or digitally excluded. Second, our binary measures of internet use could not distinguish between passive (eg, scrolling) and cognitively stimulating activities (eg, learning and problem-solving)—a key nuance for future research. Third, although the Mini-Mental State Examination cognitive composite is widely used in surveys among large aging populations, it has limited ceiling effects in high-functioning individuals, sensitivity to early or specific decline, and possible cultural or educational biases in literacy or numeracy items [62,63].

Conclusions

This study provides longitudinal evidence that internet use mediates the association between broadband access and general cognition among middle-aged and older Chinese adults. The growth trajectory in broadband access predicted increased internet use, which in turn predicted slower cognitive decline. These results argue for integrating digital literacy and socially engaging online opportunities into public health strategies for healthy aging.

Acknowledgments

The authors thank the China Health and Retirement Longitudinal Study research team and all respondents for their contribution.

Funding

This study was supported by grants from the PhD Research Startup Foundation of Inner Mongolia Medical University (YKD2024BSQD024).

Data Availability

The datasets generated or analyzed during this study are available on the China Health and Retirement Longitudinal Study website [36].

Authors' Contributions

Conceptualization: LC (lead)
Data curation: FC
Formal analysis: LC (lead), FC (supporting)
Funding acquisition: LC
Methodology: LC
Project administration: LC (lead), FC (supporting)
Resources: FC
Supervision: LC
Validation: LC
Visualization: FC (lead), LC (supporting)
Writing—original draft: FC (lead), LC (supporting)

XSL•FO
RenderX

Writing—review and editing: LC, FC (equal)

Conflicts of Interest

None declared.

References

1. Dutton WH, Graham M, editors. Society and the Internet: How Networks of Information and Communication Are Changing Our Lives: Oxford University Press; 2019.
2. ITU measuring digital development – facts and figures 2024. ITU. 2024. URL: https://www.itu.int/hub/publication/d-ind-ict_mdd-2024-4 [accessed 2026-03-09]
3. van Dijk J, Hacker K. The digital divide as a complex and dynamic phenomenon. Inf Soc 2003 Sep;19(4):315-326. [doi: 10.1080/01972240309487]
4. Falling through the net: defining the digital divide. National Telecommunications and Information Administration. 1999. URL: https://www.ntia.doc.gov/legacy/ntiahome/fttn99/contents.html [accessed 2026-03-09]
5. van Deursen AJ, van Dijk JA. The first-level digital divide shifts from inequalities in physical access to inequalities in material access. New Media & Society 2019 Feb;21(2):354-375. [doi: 10.1177/1461444818797082]
6. Wei KK, Teo HH, Chan HC, Tan BCY. Conceptualizing and testing a social cognitive model of the digital divide. Information Systems Research 2011 Mar;22(1):170-187. [doi: 10.1287/isre.1090.0273]
7. Hong YA, Zhou Z, Fang Y, Shi L. The digital divide and health disparities in China: evidence from a national survey and policy implications. J Med Internet Res 2017 Sep 11;19(9):e317. [doi: 10.2196/jmir.7786] [Medline: 28893724]
8. Yang R, Gao S, Jiang Y. Digital divide as a determinant of health in the U.S. older adults: prevalence, trends, and risk factors. BMC Geriatr 2024 Dec 21;24(1):1027. [doi: 10.1186/s12877-024-05612-y] [Medline: 39709341]
9. Pais R, Ruano L, P Carvalho O, Barros H. Global cognitive impairment prevalence and incidence in community dwelling older adults-a systematic review. Geriatrics (Basel) 2020 Oct 27;5(4):84. [doi: 10.3390/geriatrics5040084] [Medline: 33121002]
10. Singh-Manoux A, Kivimaki M, Glymour MM, et al. Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. BMJ 2012 Jan 5;344:d7622. [doi: 10.1136/bmj.d7622] [Medline: 22223828]
11. Prince MJ, Wu F, Guo Y, et al. The burden of disease in older people and implications for health policy and practice. Lancet 2015 Feb 7;385(9967):549-562. [doi: 10.1016/S0140-6736(14)61347-7] [Medline: 25468153]
12. Bradburn S, Murgatroyd C, Ray N. Neuroinflammation in mild cognitive impairment and Alzheimer's disease: a meta-analysis. Ageing Res Rev 2019 Mar;50:1-8. [doi: 10.1016/j.arr.2019.01.002] [Medline: 30610927]
13. Huynh QL, Negishi K, Blizzard L, et al. Mild cognitive impairment predicts death and readmission within 30 days of discharge for heart failure. Int J Cardiol 2016 Oct 15;221:212-217. [doi: 10.1016/j.ijcard.2016.07.074] [Medline: 27404677]
14. Guo S, Zheng XY. New evidence of trends in cognitive function among middle-aged and older adults in China, 2011-2018: an age-period-cohort analysis. BMC Geriatr 2023 Aug 21;23(1):498. [doi: 10.1186/s12877-023-04166-9] [Medline: 37605117]
15. Chen X, Giles J, Yao Y, et al. The path to healthy ageing in China: a Peking University–Lancet Commission. The Lancet 2022 Dec;400(10367):1967-2006. [doi: 10.1016/S0140-6736(22)01546-X]
16. Ding D, Zhao Q, Guo Q, et al. Prevalence of mild cognitive impairment in an urban community in China: a cross‐sectional analysis of the Shanghai Aging Study. Alzheimers Dement 2015 Mar;11(3):300. [doi: 10.1016/j.jalz.2013.11.002]
17. Ding D, Zhao Q, Guo Q, et al. Progression and predictors of mild cognitive impairment in Chinese elderly: a prospective follow‐up in the Shanghai Aging Study. Alzheimers Dement (Amst) 2016 Jan;4(1):28-36. [doi: 10.1016/j.dadm.2016.03.004] [Medline: 27489876]
18. Zeng Y, Feng Q, Hesketh T, Christensen K, Vaupel JW. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. The Lancet 2017 Apr;389(10079):1619-1629. [doi: 10.1016/S0140-6736(17)30548-2]
19. Wang Z, Chen C, Guo B, Yu Z, Zhou X. Internet plus in China. IT Prof 2016;18(3):5-8. [doi: 10.1109/MITP.2016.47]
20. The 55th statistical report on internet development in china. China Internet Network Information Center. 2025. URL: https://www.cnnic.net.cn/n4/2025/0117/c88-11229.html [accessed 2025-03-09]
21. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. Inquiry 2022;59:469580221096272. [doi: 10.1177/00469580221096272] [Medline: 35471138]
22. Ackermann K, Awaworyi Churchill S, Smyth R. Broadband internet and cognitive functioning. Economic Record 2023 Dec;99(327):536-563. [doi: 10.1111/1475-4932.12757]
23. Cao L, Chongsuvivatwong V, McNeil EB. The sociodemographic digital divide in mobile health app use among clients at outpatient departments in inner Mongolia, China: cross-sectional survey study. JMIR Hum Factors 2022 May 19;9(2):e36962. [doi: 10.2196/36962] [Medline: 35587367]
24. Li Y, Liu C, Sun J, Zhang J, Li X, Zhang Z. The digital divide and cognitive disparities among older adults: community-based cohort study in China. J Med Internet Res 2024;26:e59684. [doi: 10.2196/59684]

XSL•FO
RenderX

25. UN decade of healthy ageing: plan of action 2021-2030. World Health Organization. URL: https://cdn.who.int/media/docs/default-source/decade-of-healthy-ageing/decade-proposal-final-apr2020-en.pdf?sfvrsn=b4b75ebc_28 [accessed 2026-03-09]
26. Progress report on the United Nations decade of healthy ageing, 2021-2023. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240079694#:~:text=The%20purpose%20of%20this%20report%20is%20to%3A%20assess%20progress%20in%20interventions%20to%20ensure%20healthy%20ageing [accessed 2026-03-09]
27. World report on ageing and health. World Health Organization. 2015. URL: https://iris.who.int/server/api/core/bitstreams/5174879e-b0dc-43fc-b3a8-b1db31c51d4c/content [accessed 2026-03-09]
28. World development report 2016: digital dividends. World Bank Group. 2016. URL: https://www.worldbank.org/en/publication/wdr2016 [accessed 2026-03-09]
29. World social report 2023: leaving no one behind in an ageing world. United Nations. 2023. URL: https://www.un.org/development/desa/dspd/world-social-report/2023-2.html [accessed 2026-03-09]
30. Rajczyk JI, Burke JF, Xu WY, Wing JJ. Broadband availability and Alzheimer's disease and related dementias prevalence in central Appalachia. J Soc Serv Res 2026 Jan 19:1-10. [doi: 10.1080/01488376.2025.2610961] [Medline: 41658010]
31. Kamin ST, Lang FR. Internet use and cognitive functioning in late adulthood: longitudinal findings from the Survey of Health, Ageing and Retirement in Europe (SHARE). J Gerontol B Psychol Sci Soc Sci 2020 Feb 14;75(3):534-539. [doi: 10.1093/geronb/gby123] [Medline: 30346591]
32. Kim YK, Han SH. Internet use and cognitive functioning in later life: focus on asymmetric effects and contextual factors. Gerontologist 2022 Mar 28;62(3):425-435. [doi: 10.1093/geront/gnab149] [Medline: 34614179]
33. Yu X, Mu A, Wu X, Zhou L. Impact of internet use on cognitive decline in middle-aged and older adults in China: longitudinal observational study. J Med Internet Res 2022 Jan 24;24(1):e25760. [doi: 10.2196/25760] [Medline: 35072642]
34. Stern Y. Cognitive reserve. Neuropsychologia 2009 Aug;47(10):2015-2028. [doi: 10.1016/j.neuropsychologia.2009.03.004] [Medline: 19467352]
35. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
36. Datasets of CHARLS. CHARLS. URL: https://charls.charlsdata.com/pages/data/111/zh-cn.html [accessed 2025-11-11]
37. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
38. Scheerder A, van Deursen A, van Dijk J. Determinants of internet skills, uses and outcomes. A systematic review of the second- and third-level digital divide. Telemat Inform 2017 Dec;34(8):1607-1624. [doi: 10.1016/j.tele.2017.07.007]
39. Principles and Practice of Structural Equation Modeling, 4th edition: The Guilford Press; 2016:534.
40. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
41. van Dijk J. The Digital Divide: Polity; 2020.
42. Dannefer D. Cumulative advantage/disadvantage and the life course: cross-fertilizing age and social science theory. J Gerontol B Psychol Sci Soc Sci 2003 Nov;58(6):S327-S337. [doi: 10.1093/geronb/58.6.s327] [Medline: 14614120]
43. Li S, Ouyang Y, Hu M. Ten-year trends of the digital divides and its effect on healthy aging among older adults in China from 2011 to 2020. NPJ Digit Med 2025 Nov 21;8(1):710. [doi: 10.1038/s41746-025-02076-1] [Medline: 41272171]
44. van Dijk JAGM. Digital divide research, achievements and shortcomings. Poetics (Amst) 2006 Aug;34(4-5):221-235. [doi: 10.1016/j.poetic.2006.05.004]
45. van Deursen A, Helsper EJ, Eynon R, van Dijk JAGM. The compoundness and sequentiality of digital inequality. Int J Commun 2017;11:452-473 [FREE Full text]
46. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. The Lancet 2020 Aug;396(10248):413-446. [doi: 10.1016/S0140-6736(20)30367-6]
47. Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. Lancet 2015 Jun 6;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5] [Medline: 25771249]
48. Liang C, Subramaniam P, Mohd Ridzwan Goh NS, Kok Wai T, Moustafa AA. Digital device use, risk of cognitive impairment, and cognition in healthy older adults: the role of cognitive reserve. Healthcare (Basel) 2023 Oct 25;11(21):2822. [doi: 10.3390/healthcare11212822] [Medline: 37957967]
49. Savarimuthu A, Ponniah RJ. Cognition and cognitive reserve. Integr Psychol Behav Sci 2024 Jun;58(2):483-501. [doi: 10.1007/s12124-024-09821-3] [Medline: 38279076]
50. Deng Y, Liu L, Yang Q, Yu T. The impact of internet use on cognitive function in older adults: based on the cognitive reserve hypothesis. J Health Psychol 2025 Oct 7:13591053251371788. [doi: 10.1177/13591053251371788] [Medline: 41055540]
51. Chai Y, Xian G, Wang M, Guo L, Luo S. Aging wisely: the impact of internet use on older adults' mental health. J Affect Disord 2024 Nov 1;364:139-145. [doi: 10.1016/j.jad.2024.08.076] [Medline: 39147146]
52. Small GW, Lee J, Kaufman A, et al. Brain health consequences of digital technology use. Dialogues Clin Neurosci 2020 Jun;22(2):179-187. [doi: 10.31887/DCNS.2020.22.2/gsmall] [Medline: 32699518]

XSL•FO
RenderX

53. Small GW, Moody TD, Siddarth P, Bookheimer SY. Your brain on Google: patterns of cerebral activation during internet searching. Am J Geriatr Psychiatry 2009 Feb;17(2):116-126. [doi: 10.1097/JGP.0b013e3181953a02] [Medline: 19155745]
54. Duffner L, Deckers K, van Boxtel MPJ, de Vugt M, Köhler S. Social relationship factors, cognitive activity engagement and cognitive decline across 25 years in the Maastricht Ageing Study. Alzheimers Dement 2023 Dec;19(S22):19. [doi: 10.1002/alz.079509]
55. Ding X, Yuan L, Zhou Y. Internet access and older adults' health: evidence from China. China Economic Review 2023 Dec;82:102047. [doi: 10.1016/j.chieco.2023.102047]
56. Zheng RZ, Hill RD, Gardner MK. Engaging Older Adults with Modern Technology: Internet Use and Information Access Needs: IGI Global Scientific Publishing; 2013. [doi: 10.4018/978-1-4666-1966-1]
57. Benge JF, Scullin MK. A meta-analysis of technology use and cognitive aging. Nat Hum Behav 2025 Jul;9(7):1405-1419. [doi: 10.1038/s41562-025-02159-9] [Medline: 40229575]
58. Xavier AJ, d'Orsi E, de Oliveira CM, et al. English longitudinal study of aging: can internet/e-mail use reduce cognitive decline? J Gerontol A Biol Sci Med Sci 2014 Sep;69(9):1117-1121. [doi: 10.1093/gerona/glu105] [Medline: 25116923]
59. Firth J, Torous J, Stubbs B, et al. The "online brain": how the internet may be changing our cognition. World Psychiatry 2019 Jun;18(2):119-129. [doi: 10.1002/wps.20617] [Medline: 31059635]
60. Chen B, Yang C, Ren S, Li P, Zhao J. Relationship between internet use and cognitive function among middle-aged and older Chinese adults: 5-year longitudinal study. J Med Internet Res 2024 Dec 2;26:e57301. [doi: 10.2196/57301] [Medline: 39539034]
61. Ihle A, Bavelier D, Maurer J, Oris M, Kliegel M. Internet use in old age predicts smaller cognitive decline only in men. Sci Rep 2020 Jun 2;10(1):8969. [doi: 10.1038/s41598-020-65846-9] [Medline: 32488153]
62. Scott J, Mayo AM. Instruments for detection and screening of cognitive impairment for older adults in primary care settings: a review. Geriatr Nurs (Lond) 2018 May;39(3):323-329. [doi: 10.1016/j.gerinurse.2017.11.001]
63. Nielsen TR. Cognitive assessment in culturally, linguistically, and educationally diverse older populations in Europe. Am J Alzheimers Dis Other Demen 2022;37:15333175221117006. [doi: 10.1177/15333175221117006] [Medline: 36325840]

Abbreviations

CFI: comparative fit index
CHARLS: China Health and Retirement Longitudinal Study
PP-LGCM: parallel-process latent growth curve modeling
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index

Edited by F Puga; submitted 14.Apr.2025; peer-reviewed by C Liu, M Li, P Setthasuravich; revised version received 17.Jan.2026; accepted 19.Feb.2026; published 24.Mar.2026.

Please cite as:
Cheng F, Cao L
Broadband Access, Internet Use, and General Cognition in Middle-Aged and Older Adults: Longitudinal Study
JMIR Aging 2026;9:e75947
URL: https://aging.jmir.org/2026/1/e75947
doi:10.2196/75947

© Fangfang Cheng, Li Cao. Originally published in JMIR Aging (https://aging.jmir.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Associated With the Use of Digital Health and Well-Being Resources in Non–Memory-Led Dementias: Quantitative Survey Study

Emilie V Brotherhood[1*], PhD; Céline El Baou[2*], MSc; Caroline Fearn[2], MSc; Oliver Hayes[2], MSc; Nikki Zimmermann[1]; Anastasia Tsipa[2], DClinPsy, PhD; Sebastian J Crutch[1], DClinPsy, PhD; Joshua Stott[2], DClinPsy, PhD

[1]Dementia Research Centre, Queen Square Institute of Neurology, University College London, 8-11 Queen Square, London, United Kingdom

[2]Division of Psychology and Language Sciences, University College London, London, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Emilie V Brotherhood, PhD
Dementia Research Centre, Queen Square Institute of Neurology, University College London, 8-11 Queen Square, London, United Kingdom

Abstract

Background: Digital platforms disseminating health information and providing support for the experience of non-memory-led dementias (NMLDs) are invaluable. However, the factors influencing engagement with these resources in people affected by NMLDs are poorly understood. We conducted the world's largest survey exploring the experience of digital access in NMLDs to learn directly from people with a diagnosis, their care partners, and NMLD health care professionals (HCPs).

Objective: This study aimed to determine factors associated with digital health and well-being resource use in people with NMLD and their care partners and investigate differences in digital health and well-being resource use according to NMLD subtype.

Methods: Individuals diagnosed with NMLD—for example, those with frontotemporal lobar degeneration, posterior cortical atrophy (PCA), and primary progressive aphasia (PPA)—their care partners, and NMLD HCPs (N=450) responded to the survey. A subset of care partners provided two responses (carer-related and proxy), generating 4 survey groups with responses (N=538). The survey included demographics and basic clinical information, the outcome measure of usage behavior, and factors including constructs from the Short Senior Technology Acceptance Model (Short STAM) and additional constructs informed by a combination of previous literature and consultation with NMLD experts-by-experience (anxiety and depression [Patient Health Questionnaire–4-item], instrumental activities of daily living, web-related privacy and security concerns, and digital health literacy). Separate multiple linear regressions were run for each survey group to elucidate which variables predicted higher use of digital health and well-being resources. The use of digital resources for health and well-being was also explored across 3 NMLD subtypes: frontotemporal dementia (FTD), PPA, and PCA.

Results: Attitudinal belief was consistently the strongest predictor of digital health and well-being resource use in people with NMLD as well as caregivers and proxy respondents. Control belief was significantly associated with higher digital health and well-being resource use in the NMLD and proxy groups; a trend was observed in the carer group. The adjusted R^2 range (63.0% - 70.8%) denotes the variance in digital health and well-being resource use accounted for in the 3 models. Lower digital health and well-being resource use was associated with FTD diagnosis and caregiver groups relative to PPA and PCA.

Conclusions: Collectively, these findings indicate several factors are critical to consider when designing digital health and well-being resources for people with NMLD and their caregivers, in particular targeting practical and emotional perceptions of digital resource use for health and well-being. This should be undertaken in combination with design considerations that address condition-specific cognitive profiles encountered by those living with NMLD diagnoses and those who care for them.

(JMIR Aging 2026;9:e85863) doi:10.2196/85863

KEYWORDS

mobile health; telemedicine; telehealth; telemonitoring; dementia; atypical Alzheimer disease; frontotemporal dementia; primary progressive aphasia, gerontechnology; non–memory-led dementia; young-onset dementia; posterior cortical atrophy

XSL•FO
RenderX

Introduction

Background

Dementia is an umbrella term encompassing several diseases that cause progressive cognitive and functional decline, affecting over 57 million people worldwide, with 10 million new cases diagnosed each year [1]. While pharmacological therapeutic interventions are available for some forms of dementia, access remains inconsistent, and many individuals living with a dementia diagnosis—as well as their caregivers—rely heavily on psychosocial support and community-based services. eHealth and digitally delivered services designed for people affected by dementia are both an appealing way to disseminate relevant information and provide virtual opportunities for dementia care in the context of stretched resources (thus reducing reliance on geographical proximity to in-person services). The COVID-19 pandemic catalyzed a rapid shift toward digital health and well-being tools, accelerating the transition from paper-based and in-person community support provision to virtual platforms [2]. This shift opened numerous possibilities for dementia care, including increasing accessibility to dementia-related support groups [3]. Emerging evidence suggests that individuals with dementia and their caregivers engage with technology in similar ways, indicating that both audiences benefit from digital health and well-being provision [4].

The most common form of dementia is Alzheimer disease, with the majority of individuals presenting with episodic memory impairment as an early symptom [5]. However, a substantial proportion of people with dementia present with initial symptoms other than memory complaints. For example, individuals with an atypical variant of Alzheimer disease, posterior cortical atrophy (PCA), exhibit early visual dysfunction [6]. Changes in behavior and personality are characteristic of behavioral variant frontotemporal dementia (bvFTD [7]); language production and comprehension are a hallmark of the primary progressive aphasias (PPA [8]); and sudden-onset mood disorders are often accompanied by hallucinations in dementia with Lewy bodies [9]. These atypical presentations, collectively referred to as non–memory-led dementias (NMLDs), pose unique challenges for individuals with a diagnosis and their caregivers. In addition, symptoms often emerge before the age of 65 years (considered a young-onset dementia) with consequent implications for finances, work, and family [10]. The relative rarity of these conditions contributes to limited postdiagnostic service provision and low primary care professional and public awareness [11], while their geographical dispersion further complicates access to relevant face-to-face services [12]. Many of these challenges can be mitigated through virtual platforms, which can enable expert-led interventions [13] and peer support across geographical regions [14]. Capacity for digital engagement is high in people with NMLD as well as caregivers, owing to the fact that many affected individuals are younger than 65 years; in 2019, 73.9% of individuals aged between 50 and 64 years in England used the internet every day [15].

While digital services are a promising platform to meet the specific needs and reach of people affected by NMLDs, the factors influencing engagement with digital health and well-being resources—specifically among people living with NMLDs and their caregivers—remain underexplored, representing a critical gap in the literature. While digital platforms (eg, websites and videoconferencing) offer promising avenues for support, their usability may be directly affected by specific clinical and neuropsychological profiles associated with NMLDs. For example, anosognosia (lack of insight), which is characteristic of bvFTD, may influence likely engagement with condition-specific health and well-being materials if the individuals do not perceive themselves as having a disease. NMLD caregivers, too, face distinct potential barriers to digital health and well-being engagement. These include limited time to source and access digital health and well-being information due to demanding caregiving roles, often compounded by a lack of recognition or support from professionals and the broader public; insufficient privacy to participate in virtual support groups; and the emotional strain of navigating complex and frequently misunderstood symptoms. Understanding the specific factors associated with NMLD digital health and well-being resource use will optimize NMLD-specific digital resource and intervention design, maximizing subsequent engagement and uptake from the audiences with NMLD they aim to support.

Previous models explaining variance in technology use and behavioral intention to use technology have sought to do so in a consumer context, such as the Unified Theory of Acceptance and Use of Technology (UTAUT [16]), the revised version of which explains up to 74% of behavioral intention to use technology and 52% of actual use [17]. Combining this with another predominant model for workplace technology use (Technology Acceptance Model [18]), Chen and Chan [19] developed a 38-item seniors-focused model to understand gerontechnological acceptance by older people from Hong Kong. Alongside original constructs pertaining to attitudinal beliefs and control aspects, Chen and Chan [19] included additional constructs of age-related health and independent capabilities, revised in a shortened version of the Short Senior Technology Acceptance Model (14-item; Short STAM), which demonstrated its accuracy in relation to assessing the propensity of individuals older than 55 years to engage with new technology, appropriate for dementia populations with a lower age of onset and their caregivers [20]. Low variance for the Short STAM (14-item) explaining technology usage behavior in subsequent studies [21] indicates other population-specific factors that may be important to consider when examining factors contributing to technology-use behavior in NLMDs. Qualitative work focusing on technology use in older adults with hypertension, for example, highlights other factors potentially influencing engagement, such as privacy concerns [22]. This factor may even be arguably more pertinent for studies exploring technology usage behavior in dementia populations, owing to acknowledged vulnerability to financial and technological scams and for people with NMLDs in the necessity to divulge sensitive category data by virtue of explaining symptoms (eg, a person living with posterior cortical atrophy disclosing their clinical phenotype to allow for reasonable adjustments to be made). Additional aspects influencing digital health and well-being resource use in people with NMLDs may include factors related to major neurocognitive changes (eg,

XSL•FO
RenderX

changes in instrumental activities of daily living), shared experiences between people with dementia and their caregivers (eg, anxiety and depression), and more recently identified constructs such as digital health literacy that were not included in the Short STAM (14-item) model. Informed by these current models of technology use, we conducted the first quantitative study to our knowledge exploring these and additional constructs in people living with NMLDs and their caregivers (see Figure 1). We included subconstructs and item numbers from the Short STAM, the selection for which was derived from a sequential item-reduction strategy and confirmatory factor analysis of prior models (original STAM–38 items [19]). Standardized parameter estimates for interconstruct correlations for the Short STAM (14-item) are reported in full in the study by Chen and Lou [20]. The right-hand block indicates additional constructs added to the digital access survey in this study, with variable selection guided by technology-use investigations in other clinical cohorts [22] and patient and public involvement (PPI) consultation with NMLD experts-by-experience. Aligned with our focus on flexible health care provision, we focused on technology use relating to digital platforms providing resources in the domains of health and well-being (eg, websites, Zoom [Zoom Video Communications]) and remained agnostic to mode of access (eg, whether accessed via smartphone, tablet, or computer).

XSL•FO
RenderX

Figure 1. Construct of a map detailing the domains within the digital access survey. A: attitude toward aging; ATT: attitude toward using technology; C: cognitive ability; H: self-reported health conditions; FC: facilitating conditions; PEOU: perceived ease of use; PU: perceived usefulness; S: social relationships; SE: gerontechnology self-efficacy.

XSL•FO
RenderX

Study Objectives

To address this gap, the following research questions were formulated:

1. What are the factors associated with digital health and well-being resource use in people with NMLD and their care partners?
2. Is there a difference in digital health and well-being resource use according to NMLD subtype?

Methods

Participant Characterization

Background demographic variables previously reported to influence usage (eg, age in years, gender, and ethnicity [17,23]) were collected alongside clinical characteristics such as NMLD type (eg, primary progressive aphasia, posterior cortical atrophy, and frontotemporal dementia), time since first symptoms, and formal dementia diagnosis.

The final sample comprised 450 participants with 538 survey responses, collected between March 2023 and March 2024. This exceeded our initial power calculations for analysis of participation issues affecting people with NMLD, which was computed at n=385 (assuming 10 predictors; power=90%; R^2=0.33; α=.05). Taking a conservative approach, the minimum sample size for the 10 candidate predictor variables would be 385 (assuming outcome prevalence=50%, margin of error in intercept estimation =5%, 5% acceptable difference in apparent, and adjusted R^2 of 0.2). Survey response N exceeds sample size owing to a group of carers providing 2 responses, one relating to their own experiences and proxy responses for the person with NMLD. Participants included 63 people with a diagnosis of NMLD, 269 carer reporting on their own digital health and well-being resource use, 105 proxy responses, and 101 health care professionals (HCPs).

Survey Design and Measure Selection

The study was based on a cross-sectional questionnaire, using a participant-report approach. The survey was carefully designed to consider ease of administration and reduce participant burden where possible, in consultation with HCPs experienced in working with people with NMLD as well as caregivers and proxy respondents (eg, consultant neurologists, clinical psychologists, academic researchers, and support workers), and the Rare Dementia Support PPI group, comprising people living with and alongside NMLDs.

The theoretical constructs proposed for the survey were based on a rapid scoping review of studies exploring information and communication technology adoption and usage among older adults. Three theoretical models were identified as relevant for this study: (1) the UTAUT [16], (2) the Technology Acceptance Model [18], and (3) the Senior Technology Acceptance Model (STAM [19,20]). A variety of proposed measures were initially circulated in full to HCPs with relevant expertise in NMLD, PPI leads from two NMLD research studies, and the Rare Dementia Support PPI group. Following PPI consultation and feedback relating to length and potential participant burden, scale selection was finalized with the objective of balancing research requirements with participant burden and accessibility for people living with NMLDs. The questionnaire was available in English. The finalized survey scale items are outlined in full in the Table S1 in Multimedia Appendix 1 and described further below.

Outcome Measure: Usage Behavior

Usage behavior (UB) was measured using participant-reported technology use (TU) and behavioral intention to use. TU was measured by establishing frequency of use within the 3 months prior to data collection (7-point Likert scale, ranging from “not at all” [0] to “multiple times a day” [6]). Behavioral intention reflected a measure of usage behavior validated in previous work via participant report of 3 questions relating to intention to use technology, responses to which were established using a 10-point Likert scale (“disagree strongly” to “agree strongly” [17]). Usage behavior was calculated as a composite score of TU scores (where 0 - 6 frequency scores were rescaled as dummy variables) and behavioral intention to use scores (total score range$_{BI}$: 3 - 30), resulting in a total UB score range of 3 - 36 reflecting past use frequency and intention to use in the future. Higher scores indicate greater usage behavior of digital resources for health and psychological well-being in the surveyed population. Summing user variables to compute a composite outcome measure and subsequently treating this as a continuous outcome aligns with previous approaches in the literature using the UTAUT2 [17] and STAMs to assess smartphone and technology use in older adults [17,19,23].

Factors and Constructs

Predictors comprised constructs validated in the Short STAM 14-item (attitudinal beliefs, n=3; control beliefs, n=4; gerontechnology anxiety, n_{items}=2; health, n_{items}=5), which has been shown to explain up to 81.5% variance of actual technology usage in older adults [20]. Additional measures of depression and anxiety (Patient Health Questionnaire–4-item [PHQ-4]; [24]), instrumental activities of daily living [25], web-related privacy and security concerns (n_{items}=6 [26]) and the Digital Health Care Literacy Scale (n_{items}=3 [27]) were added to this survey following consultation of other technology-use studies in clinical older adult cohorts [22] and PPI consultation with NMLD experts-by-experience to further characterize the sample and explore their potential influence on technology use in people with NMLD as well as caregivers and proxy respondents. Likert scales were used for all scale responses, preserving the parameters and scale numbers according to each scale’s administration instructions.

Adaptation and Procedure

PPI consultation resulted in several adaptations being incorporated into the final survey. This included using the standardized phrase “web-based resources (eg, Zoom and websites) for your health and psychological well-being” across the scales, replacing varying terms such as “mobile Internet,” “ICT,” and “technology,” to create an accessible term that captured the digital platforms of interest in this study and reduce the number of questions. In addition, online survey branching was introduced to ensure participants would view only the

XSL•FO
RenderX

questions relevant to their experience, thus reducing completion time and participant burden.

The finalized scales were uploaded to Qualtrics and replicated in a hard-copy version where preferred. Four survey types were developed to tailor the questions to different possible NMLD experiences:

1. NMLD: for people living with a diagnosis of NMLD;
2. Proxy NMLD: for family members or friends of individuals with NMLD who provided proxy responses;
3. Proxy HCP: for HCPs who provided a proxy response on behalf of a person with NMLD who had capacity to access online resources;
4. Care partners: for family members or friends of those with a diagnosis of NMLD.

All electronic surveys were tested for technical functionality before data collection took place. Surveys were administered to individuals with experience of the following non–memory-led and young-onset dementias: Frontotemporal lobar degeneration (behavioral variant frontotemporal dementia; primary progressive aphasias [semantic variant primary progressive aphasia; nonfluent variant primary progressive aphasia]), atypical Alzheimer disease (posterior cortical atrophy; logopenic variant primary progressive aphasia); dementia with Lewy bodies, inherited dementias (familial frontotemporal dementia; familial Alzheimer disease) and young-onset Alzheimer disease.

Variable Derivation and Statistical Analysis

Data cleaning and item scoring were undertaken in SAS (SAS Institute) and RStudio (Posit). For analysis purposes, clinical phenotypes were pooled to reflect broad similarities in the presenting symptoms' cognitive domains. For example, the primary progressive aphasia subtypes (semantic, logopenic, and nonfluent variants, where specified) were subsumed under the label "PPA" to characterize digital health and well-being resource use for individuals with predominant language impairments. Where non–memory-led conditions were either reported to be copresenting (eg, frontotemporal dementia with dementia with Lewy bodies), individuals presented with other syndromes such as young onset Alzheimer disease or vascular dementia, or group sizes precluded condition-specific analyses (eg, individuals living with progressive supranuclear palsy; n=2), these were labeled as "Other."

Demographic information stratified by survey group is outlined in Table 1. To facilitate comparison across scales, composite scores were normalized, creating *z* scores for outcome and predictor variable responses. To elucidate which factors may be associated with higher usage of digital health and well-being resources, a linear multiple regression was fitted using the UB composite variable as an outcome. Separate multiple regressions were undertaken for each survey type (NMLD, Proxy, HCP, and Carer). To understand whether FTD, PCA, and PPA subtypes were associated with different usage behavior, a robust linear model was undertaken using the UB composite variable as the outcome, with diagnostic category as the main predictor variable and age and gender as covariates. This model was run twice; once collapsing the NMLD and proxy groups and once modeling FTD, PPA, and PCA carers only. Standardized coefficients, 95% CIs, and related significance values (where $P<.05$) are reported for each predictor. The default approach within the lm function in R (R Foundation for Statistical Computing) handles missing data by performing listwise (complete case) deletion. Therefore, cases with any missing data on the predictors were excluded from the regression model. The number of observations for each group is outlined in Multimedia (Table S2 in Multimedia Appendix 1). Acknowledging the smaller sample size, particularly in the NMLD group, model residuals were assessed for normality using statistical tests (Kolmogorov-Smirnov) and visual inspection via Q-Q plots. Testing for multicollinearity by calculating variance inflation factors for variables within each model confirmed that all variance inflation factor values were satisfactory (<5).

XSL•FO
RenderX

Table . Demographic characteristics and normalized survey scores, split by survey type.

Participant group	NMLD[a] (N=63)	Proxy NMLD (N=105)	Proxy HCP[b] (N=101)	Care partners (N=269)
Basic demographic and clinical information				
Age (years), mean (SD)	63.49 (10.54)	67.72 (9.80)	48.79 (11.52)	61.85 (12.88)
Sex (male), n (%)	30 (47.62)	55 (52.38)	77 (76.24)	192 (71.91)
Ethnicity (White), n (%)	62 (100.00)[c]	99 (95.19)	88 (87.13)	251 (93.31)
Time since diagnosis (years), mean (SD)	3.39 (3.05)	3.51 (4.70)	—[d]	3.87 (3.75)
Time since symptoms noticed (years), mean (SD)	5.74 (3.75)	5.97 (3.61)	—	6.81 (4.41)
Dementia subtype, n (%)				
FTD[e]	12 (19.05)	33 (31.43)	17 (16.83)	88 (32.71)
PPA[f]	15 (23.81)	21 (20.00)	67 (66.34)	74 (27.51)
PCA[g]	28 (44.44)	38 (36.19)	2 (1.98)	72 (26.77)
DLB[h]	0 (0.00)	7 (6.67)	1 (0.99)	12 (4.46)
Other	8 (12.70)	6 (5.71)	14 (13.86)	23 (8.55)
Survey completion format, n (%)				
Online	51 (80.95)	99 (94.29)	101 (100.00)	204 (76.40)
Digital health resource use (ever use = yes)	51 (80.95)	55 (52.38)	—	217(80.67)
Scale scores mean raw scores (SD)				
Digital resource use (Composite)	22.81 (10.29)	8.91 (9.20)	13.88 (6.35)	22.16 (9.11)
Attitudinal beliefs	19.84 (8.98)	9.28 (7.22)	21.12 (5.35)	19.66 (7.25)
Control beliefs	27.78 (9.43)	16.68 (8.46)	21.74 (5.75)	32.72 (6.39)
Gerontechnology anxiety	10.28 (6.31)	13.01 (6.83)	12.30 (3.78)	6.26 (5.06)
Health	34.44 (8.88)	25.97 (9.45)	—[d]	35.95 (7.55)
PHQ-4[i]	4.25 (3.16)	5.57 (3.73)	—	3.93 (3.20)
Instrumental activities of daily living	14.88 (7.19)	5.86 (6.34)	—	23.15(2.90)
Online security and privacy concerns	19.28 (10.43)	17.56 (10.32)	20.15 (7.14)	23.04 (11.32)
Digital literacy	5.93 (4.77)	0.89 (2.14)	4.04 (2.52)	10.03 (2.82)

[a]NMLD: non-memory-led dementia.

[b]HCP: health care professional.

[c]Does not reflect total Nowing to missing data.

[d]Not applicable.

[e]FTD: frontotemporal dementia.

[f]PPA: primary progressive aphasia.

[g]PCA: posterior cortical atrophy.

[h]DLB: dementia with Lewy bodies.

[i]PHQ-4: Patient Health Questionnaire-4 item.

Ethical Considerations

Ethical approval was obtained from the University College London (UCL) Research Ethics Committee (8545/004: Rare Dementia Support Impact Study [28]). Potential participants were identified by one of two mechanisms: (1) membership of the UCL-led Rare Dementia Support [29], or (2) membership of the Dementia Research Center participant database. Members of both received email correspondence containing a link to the participant information sheet, online consent form, and survey.

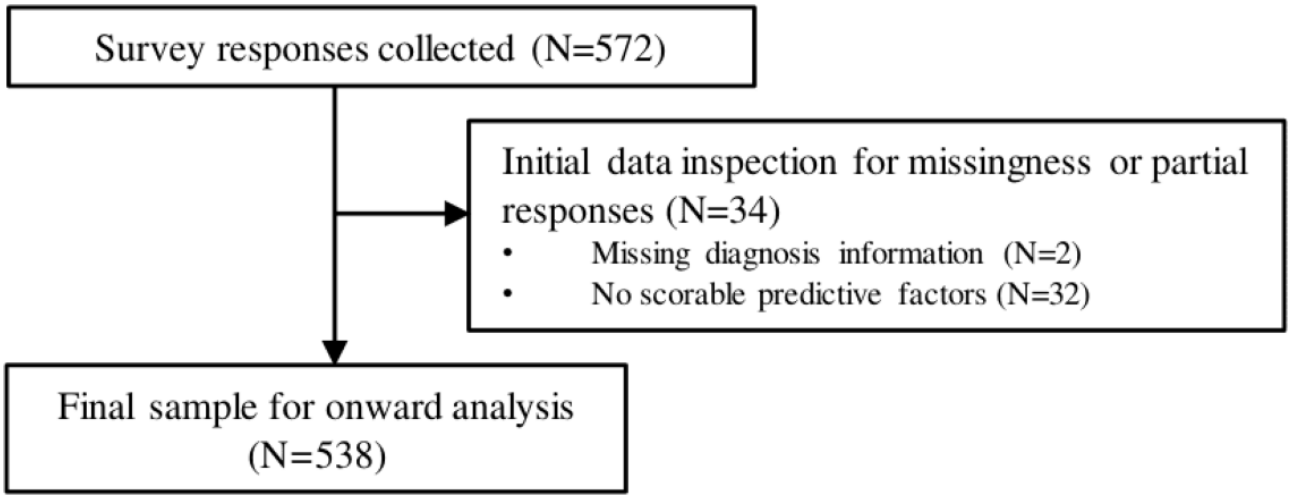
XSL•FO
RenderX

To ensure we included individuals who may not use digital health and well-being resources, the e-invitation was supplemented by brief research reminders in several in-person support group meetings, where attendees had the opportunity to read the participant information sheet and ask clarifying questions to the research team, provide written informed consent, and complete a hard copy version of the questionnaire (including a version adapted for posterior cortical atrophy in consultation with a specialist NMLD neuropsychologist).

Results

The final sample comprised 538 responses from 450 individuals (354 women, 92.9% White; mean age 60.75, SD 13.29 years; see Figure 2). All collected entries were used to calculate the overall survey completion rate (94.1%). A small proportion (0.74% - 3.39%) of data were missing in the remaining survey group datasets, the highest proportion observed for people living with an NMLD response data. Full characteristics of the sample are shown in Table 1, split by NMLD condition and participant group.

Figure 2. Flowchart depicting response exclusion and final sample selection.



Model outputs are available in Table S2 in Multimedia Appendix 1. A higher use of digital resources for health and well-being for NMLD responders was moderately associated with higher attitudinal beliefs (eg, thinking that digital resources might be helpful) (β=.43, 95% CI 0.24-0.63; *P*<.001), followed by control beliefs (eg, perceived ease of use; β=.30, 95% CI 0.05-0.54; *P*=.02). Adjusted R^2 values indicated the model accounted for 69.3% of the variance in use of health and well-being digital resources.

Proxy NMLD diagnoses were provided in 97.1% of caregiver responses by either a spouse or partner (67.6%) or first-degree relative (parent, child, or sibling; 29.5%). The proxy model output broadly recapitulated the predictors of digital health and well-being resources as reported by individuals living with a diagnosis (attitudinal beliefs β=.66, 95% CI 0.52-0.80; *P*<.001; control beliefs β=.20, 95% CI 0.05-0.35; *P*=.009). In contrast to the NMLD model, proxy PHQ-4 score was additionally associated with proxy digital health and well-being resource use (β=.15, 95% CI 0.03-0.26; *P*=.01), with the overall model accounting for 70.8% of the variance in proxy-reported digital health and well-being resource use.

Responses from HCPs working with people with NMLD demonstrated a range of disciplines, namely NMLD clinicians (neurologists, psychiatrists, and clinical neuropsychologists), nurses, occupational and speech and language therapists, academic researchers, dementia support workers, care home managers, and social workers. One observation in the HCPs' model, corresponding to an outlier, was identified as deviating from the theoretical normal distribution by approximately 1.0 in the upper tail of the Q-Q plot. No remarkable changes in results were observed when rerunning this model after having excluded this observation, at which point normality assumptions were upheld. The model's ability to account for the variance of digital health and well-being resource use in the proxy HCPs' responses was greatly reduced relative to the two previous groups (Adjusted R^2=0.284). Nevertheless, the proxy HCP model was consistent with the findings above in that control beliefs were moderately associated with proxy digital health and well-being resource use (β=.38, 95% CI 0.12-0.64; *P*=.005).

Carers responding to the survey about their own use of digital health and well-being resources were consistent with other models in reporting a strong association between digital health and well-being resource use and attitudinal beliefs (β=.70, 95% CI 0.61-0.79; *P*<.001). A small association was demonstrated between digital health and well-being resource use and control beliefs (β=.13, 95% CI −0.001 to 0.25; *P*=.05). Distinct from previous models, carers' perception of better health was associated with digital health and well-being resources (β=.12, 95% CI 0.02-0.23; *P*=.03). Adjusted R^2 values indicated the model accounted for 63.0% of the variance in use of digital health and well-being resources.

To provide further insight into clinical phenotype associations with digital health and well-being resource use, we conducted exploratory analyses profiling associations of diagnosis (FTD, PPA, and PCA phenotypes) with digital health and well-being resource use. Given the relative concordance between the NMLD and carer-proxy profiles, we considered it appropriate to pool data from these 2 groups and stratify this aggregated

XSL•FO
RenderX

group (NMLD+proxy) by clinical phenotype, alongside carer scores for each group (see Table 2). All observations were retained (n=147) for onward analysis. Controlling for covariates age and gender, a robust linear regression indicated that diagnosis category was associated with digital health and well-being resource use for self- and proxy-reported NMLD. Compared to participants with FTD (reference group), higher digital health and well-being resource use were associated with the PCA (β=.75, 95% CI 0.33-1.16) and PPA groups (β=.55, 95% CI 0.06-1.03). This was recapitulated in the carer groups; higher digital health and well-being resource use was associated with the PCA (β=.51, 95% CI 0.23-0.78) and PPA (β=.27, 95% CI 0.01-0.54) carer groups relative to FTD carers.

Table . Demographic, basic survey information, and outcome measure scores depicted for the subgroup of the initial sample, stratified by clinical phenotype of behavioral variant frontotemporal dementia, primary progressive aphasia, and posterior cortical atrophy.

Clinical phenotype	bvFTD[a]	PPA[b]	PCA[c]
Living with NMLD[d] diagnosis and proxy combined	n=45	n=36	n=66
Age (years), mean (SD)	63.9 (12.4)	71.3 (8.1)	65.4 (8.1)
Female:male (N)	15:30	16:20	42:24
Ethnicity (White), n (%)	43 (95.5)	33 (91.7)	65 (98.5)
Time since diagnosis (years), mean (SD)	3.5 (3.3)	3.1 (1.8)	2.9 (2.6)
Time since symptoms noticed (years), mean (SD)	6.6 (4.2)	5.5 (2.8)	5.3 (3.5)
Online, n (%)	44 (97.8)	31 (86.1)	56 (84.9)
Digital resource use (composite), mean raw score (SD)	9.80 (10.25)	14.44 (12.67)	16.91 (11.34)
Carers	n=88	n=74	n=72
Age (years), mean (SD)	61.5 (13.3)	63.6 (11.3)	59.8 (14.6)
Female:male (N)	75:12[e]	54:20	39:31[e]
Ethnicity (White) n (%)	83 (94.3)	67 (90.5)	68 (94.4)
Time since diagnosis (years), mean (SD)	3.6 (3.2)	4.0 (2.9)	3.5 (2.5)
Time since symptoms noticed (years), mean (SD)	7.3 (4.2)	7.2 (5.5)	5.8 (2.9)
Online, n (%)	67 (76.1)	55 (74.3)	56 (77.8)
Digital resource use (composite), mean raw score (SD)	19.91 (9.02)	22.14 (9.39)	24.10 (8.47)

[a]bvFTD: behavioral variant frontotemporal dementia.

[b]PPA: primary progressive aphasia.

[c]PCA: posterior cortical atrophy.

[d]NMLD: non-memory-led dementia.

[e]Does not reflect total N owing to missing data or respondents preferring to self-describe.

Discussion

Principal Results

This work was conducted with the aim of understanding the factors associated with digital health and well-being resource use in people with NMLD as well as caregivers and proxy respondents. The findings provide insight for HCPs who seek to maximize engagement with NMLD-specific health and well-being support and underscore considerations for digital NMLD resource design and dissemination. Our findings characterize the extent to which theoretically derived technology use factors associate with digital health and well-being resource use in people with NMLD as well as caregivers and proxy respondents. Here, we report a consistent association between digital health and well-being resource use and control and attitudinal beliefs, regardless of whether it is proxy- or self-reported for people living with NMLDs. Caregiver-proxy reports additionally identified that increased feelings of anxiety and depression in the person with NMLD are associated with their engagement with digital health and well-being resource use. For carers' own digital health and well-being resource use, perception of better health was an additional predictor of higher use. Carer-reported proxy responses estimated NMLD resource use was lower relative to the reports received directly from individuals with an NMLD diagnosis, and individuals with FTD and their carers demonstrated lower digital health and well-being resource use relative to their PPA and PCA counterparts.

XSL•FO
RenderX

Interpretations

Controls' beliefs were positively associated with digital health and well-being resource use across NMLD self-reported and proxy models. Examining the subordinate items within this construct, the likelihood of NMLD engagement with digital health and well-being offering is increased with a platform's perceived ease of use, the user's self-efficacy, and facilitating conditions being met (eg, accessibility factors and financial means). Promoting a user's self-efficacy maximizes the potential of digital platforms for health and well-being. Practically, this finding highlights the importance of training opportunities for people with NMLDs to optimize their experience with any bespoke digital health and well-being resource. This could be achieved by virtue of a tutorial or demonstration by another individual. Training may, for example, involve incorporating initial "onboarding" sessions, which aim to acclimatize the person affected by NMLD to a new digital health and well-being offer. These findings support previous research investigating videoconference-delivered support for people with dementia and their caregivers, whereby integrated technological support and training facilitated ongoing access by participants and users [30-33]. Digital provision of mental health support is available at the time of this writing in the United Kingdom for people with NMLD as well as caregivers if the need of the individuals is deemed to be low in complexity [34]. Evidence suggests that people with NMLD as well as caregivers benefit from accessing such services, therefore underscoring a critical need for NMLD-specific onboarding training as part of this and other digitally mediated mental health provisions [34,35]. In addition, the findings relating to control beliefs highlight the influence of basic accessibility factors in digital health and well-being engagement, which emphasizes the importance of designing for open access for people with NMLD as well as caregivers. Capitalizing on accessibility opportunities, such as working in partnership with community groups and institutions that offer tablet devices and computer access to the general public, will increase equity for individuals affected by NMLD who may otherwise be unable to access specialist digital resources. Taken together, these findings reflect recently reported themes from semistructured interviews exploring technology use in older adults with intellectual disabilities, resulting in recommendations around providing structured, repetitive, and gradual training opportunities for individuals with cognitive challenges, as well as promoting distribution of digital resources in rural areas [36].

Attitudinal beliefs were consistently the strongest predictor of digital health and well-being resources, arguably playing a crucial role in determining whether people living with dementia and their caregivers engage with these resources. This was demonstrated to a greater extent relative to the usability barriers outlined above. This construct reflects the individual user's perception of these resources being beneficial (eg, enhancing daily effectiveness) and whether they find their use appealing. For people with NMLD as well as caregivers and proxy respondents, the perceived functional value of digital health and well-being resources is critical, perhaps owing to this population seeking digital resources that specifically respond to their unique challenges, which often differ from amnestic dementia experiences. Practically, these findings point to a need for effective communication strategies at the point of offering these resources, which (1) provide a clear scaffold and framing of the digital resource in question and (2) emphasize the positive role that digital resources can play in supporting the navigation of NMLD symptoms, caregiving guidance, or health and well-being support for caregivers.

An additional insight from the NMLD-proxy group indicated that carer-perceived increased digital health and well-being resource use was associated with the person with NMLD's higher anxiety and depression. Acknowledging that causality may not be inferred from the study design and analysis approach, we venture a range of factors that may explain this finding. For example, depression and anxiety symptoms may be linked to digital health and well-being engagement by virtue of a different profile of help-seeking behavior, either in relation to the person with NMLD consulting the internet for potential mood-related symptom treatment or management through mental health and well-being platforms aimed to target these experiences; or, more broadly, to quell other non–dementia-specific concerns or worries that are exacerbated by virtue of increased anxiety and depression as a result of navigating the challenges dementia brings. Additionally, caregivers providing proxy responses may be aware of their own prompting behaviors, perhaps suggesting that a loved one displaying anxiety and depression may find online resources beneficial for managing their distress. Without further qualitative investigation; however, the mechanisms mediating this association remain speculative.

NMLD caregiver–specific findings indicated that higher self-reported health status predicted higher use of digital health and well-being resources. This differs from previous findings relating to dementia caregiver internet use, whereby caregivers of people primarily with Alzheimer disease or vascular dementia who appraised their own physical health as worse reported significantly more frequent internet use relative to those who appraised their physical health to be similar to a noncaregiver peer of the same age and gender [37]. This contrast may, however, reflect some specificity of the interplay of NMLD caregiver health and digital health resources relative to the same factor's influence for those caring for individuals with typical dementia variants. This supports the justification for stratifying the needs of these caregivers through bespoke post-diagnostic support and signposting as they form their own distinct profile of factors influencing digital health and well-being technology use.

The distinct cognitive hallmarks of FTD may explain the heterogeneity observed between NMLD phenotypes in relation to digital health and well-being resource use in the self- and proxy-reported group. The most common form of FTD, bvFTD, is characterized by changes in personality and behavior, coupled with early apathy and lack of insight into these symptoms [7]. The lower use of health and well-being digital resources relative to others presenting with different non–memory-led symptoms (as in PCA or PPA) may either reflect a lack of motivation to engage, a failure to even acknowledge or recognize there may be a health problem to consult about, or a combination. Addressing these 2 condition-specific barriers may be key in engaging people with FTD in health and well-being resources related to their condition. For FTD individuals in particular

XSL•FO
RenderX

(relative to PPA or PCA), it may be important to explore digital health and well-being resource use design that harnesses methods aiming to increase intrinsic motivation to engage, such as gamification approaches. Unequal diagnostic group sizes resulted in unequal precision as indicated by wider 95% CIs with lower bounds for PPA, consistent with a smaller group size and implying lower sensitivity to detect small effects. We, therefore, interpret this finding with caution. An interesting finding related to the recapitulation of these condition-specific web resource use profiles in the FTD carer groups in relation to the PCA and PPA groups. One potential explanation for this could be reflecting higher care demands specifically in bvFTD, with higher reports of caregiver stress and lower perceived control observed in this caregiver group relative to Alzheimer disease carers [38] and those caring for people with other NMLDs such as PPA [39]. This increase in stress and caregiver burden is likely to diminish FTD carers' dedicated time and resources to digital use in relation to their own health and well-being needs, in turn reflecting lower use relative to other NMLD caregiver peers.

Taken together, these findings highlight the utmost importance of condition-specific considerations when designing health and well-being digital materials for individuals living with and caring for individuals with NMLDs. However, as a recent systematic review of participatory methods in digital health interventions for dementia caregivers highlighted, subtyping is relatively rare, and many interventions are designed for the typical array of caregiver needs [40]. Approaches to adapt existing support platforms designed with more prevalent dementias in mind (eg, iSupport; World Health Organization) to cater to the specific needs of individuals with NMLDs included several co-design stages to enable iterative changes, including focus groups and validation surveys [41]. Similar approaches have been adopted for the co-development of forthcoming digital health interventions specifically designed by, with, and for specific dementia caregiver subtypes from the outset [13]. Co-design of digital resources and interventions with individuals living with or caring for NMLD is likely critical for ensuring ongoing engagement and interaction with digital provisions designed with people with NMLD as well as caregivers and proxy respondents in mind [42].

Limitations

Our health care professional model demonstrated far less predictive value of health and well-being digital resource use (only 28% of the variance). This may reflect the instruction given to HCPs to "imagine clients you have worked with," which does not give ample opportunity for HCPs to share the complexity or breadth of their experience when trying to engage NMLD audiences in digital health and well-being offers. This contrasts with the good explanatory power demonstrated by the NMLD, carer, and proxy models, perhaps reflecting the selective focus of these groups, who were instructed to respond in reference to only one individual (either themselves or one other in the proxy report). Given the HCP group comprised a multidisciplinary group, there is also the potential for differing frames of reference to have impacted model fit. A variability in the professional role within the HCP group, and therefore the clients or service users envisaged when HCP participants completed the survey, may have introduced greater heterogeneity, which in turn would weaken the predictive capacity of the HCP model in contrast to the other models reported here. We acknowledge that the quantitative format of the survey may have also restricted the potential for HCPs to share the depth of expertise that professionals working in this space have to offer. This reflection emphasizes the importance of mixed methods approaches incorporating qualitative inquiry with HCPs to harness the benefits of their range of experiences in this area and to elucidate any discipline-specific observations. This work has been undertaken, and findings are described in a separate publication (Fearn et al, forthcoming).

The definition of technology use in this study was intentionally narrow, concentrating specifically on digital health and well-being resources in order to capture information relating to engagement with NMLD digital offers that are currently available, such as virtual support groups and digital platforms (eg, Rare Dementia Support and Dyscover). This survey did not capture particular devices NMLD users preferred to access these resources, which may be mediated by NMLD audience (persons living with vs carers) or syndromic group and which may have notable influence on whether or not these digital resources are engaged with. FTD populations, for example, exhibit lower smartphone battery usage relative to clinically neurotypical controls [43], which may necessitate other means of presenting health and well-being information for these groups that seem not to engage with smartphone interaction. Together, these findings must also be considered alongside other contextual factors at play, including mediating factors that can influence telehealth success in these individuals, such as the engagement of the communication partner, which has been found to associate with functional outcomes in videoconference-delivered PPA speech and language therapies [44].

Although we endeavored to collect data using both hard-copy paper questionnaires and a complementary digital survey option to capture a broad range of digital access experiences, we acknowledge that the invitations for both formats were distributed electronically. As a result, this sample may overrepresent digitally literate older adults relative to the wider dementia and caregiver populations. Relatedly, internet use is associated most strongly with individuals with dementia specifically from middle-income backgrounds with high educational attainment, characteristics that are also prevalent in membership of specialist support services such as the one from which our sample was derived [45]. Socioeconomic factors similarly associate with internet access in caregiver populations; for instance, individuals with private health insurance are 8 times more likely to have internet access than those without [46]. Furthermore, our sample comprised a very high proportion of White participants, which did not provide the opportunity for exploring potential digital-access patterns and associated factors in more ethnically diverse people with NMLD as well as caregivers and proxy respondents. Our findings and their generalizability should therefore be considered in this context. As awareness for NMLDs grows, there may be opportunities for future research studies to use recruitment strategies that identify participants from more community-based or primary

XSL•FO
RenderX

care settings who, in turn, are likely to increase the representativeness relative to this sample. Finally, despite the large overall sample size, we acknowledge that in order to examine group-specific profiles, the number of available observations per group and number of predictors resulted in regressions that may be overfitting. While assumptions were upheld through normality testing with little evidence for multicollinearity, the findings should be interpreted with the understanding that these models are exploratory in nature.

Conclusions

Collectively, these findings provide insight into several factors that are critical to consider when designing digital health and well-being resources for people with NMLD and their caregivers. The insights from this work underscore the importance of marketing and resource design that resonate with NMLD audiences in terms of targeting practical and emotional perceptions of digital health and well-being offers. This should be undertaken in combination with design considerations that address technical and logistical barriers to access, as well as considering the specific intended audience in terms of specific cognitive and symptom profiles.

Acknowledgments

The authors would like to thank all our participants, without whom this work would not have been possible. We are also grateful to Mrs Valerie Mansfield as Co-Lead (PPI), Mrs Roberta McKee-Jackson, and the Rare Dementia Support PPI group for reviewing the surveys and providing feedback on drafts.

Funding

This study/project is funded by the National Institute for Health and Care Research (NIHR) under its Programme Grants for Applied Research Programme [NIHR203680]. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: EVB, CEB, CF, AT, SJC, JS.
Data curation: EVB, CEB
Formal analysis: EVB, CEB
Funding acquisition: JS, SJC
Investigation: EVB, CEB, CF, NZ, OH
Methodology: EVB, CEB, CF, PH, AT, JS, SJC
Project administration: EVB, CEB, CF, OH
Resources: CEB
Supervision: JS, SJC
Visualization: EVB
Writing – original draft: EVB (lead), CEB (supporting)
Writing – review & editing: EVB (lead), CEB (supporting), CF, OH, NZ, AT, SJC, JS

Conflicts of Interest

None declared.

Multimedia Appendix 1
Digital access survey items, full regression model reporting, and Cronbach alpha stratified by survey type.
[DOCX File, 44 KB - aging_v9i1e85863_app1.docx]

References

1. Dementia. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2026-02-15]
2. Patel AM, Schuldt R, Boudreau DM, Cobb BR, Win N, McGinley MP. Telemedicine use before and during the COVID-19 pandemic in people with Alzheimer's disease, multiple sclerosis, or Parkinson's disease: a cross-sectional study using US commercial claims data. Telemed Rep 2024;5(1):247-255. [doi: 10.1089/tmr.2024.0030] [Medline: 39184874]

XSL•FO
RenderX

3. Waddington C, Harding E, Brotherhood EV, et al. The development of videoconference-based support for people living with rare dementias and their carers: protocol for a 3-phase support group evaluation. JMIR Res Protoc 2022 Jul 20;11(7):e35376. [doi: 10.2196/35376] [Medline: 35857375]
4. Lee AR, McDermott O, Orrell M. Understanding barriers and facilitators to online and app activities for people living with dementia and their supporters. J Geriatr Psychiatry Neurol 2023 Sep;36(5):366-375. [doi: 10.1177/08919887221149139] [Medline: 36597870]
5. Dubois B. The emergence of a new conceptual framework for Alzheimer's disease. J Alzheimers Dis 2018;62(3):1059-1066. [doi: 10.3233/JAD-170536] [Medline: 29036825]
6. Crutch SJ, Schott JM, Rabinovici GD, et al. Consensus classification of posterior cortical atrophy. Alzheimers Dement 2017 Aug;13(8):870-884. [doi: 10.1016/j.jalz.2017.01.014] [Medline: 28259709]
7. Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain (Bacau) 2011 Sep;134(Pt 9):2456-2477. [doi: 10.1093/brain/awr179] [Medline: 21810890]
8. Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. Neurology (ECronicon) 2011 Mar 15;76(11):1006-1014. [doi: 10.1212/WNL.0b013e31821103e6] [Medline: 21325651]
9. McKeith IG, Boeve BF, Dickson DW, et al. Diagnosis and management of dementia with lewy bodies: fourth consensus report of the DLB Consortium. Neurology (ECronicon) 2017 Jul 4;89(1):88-100. [doi: 10.1212/WNL.0000000000004058] [Medline: 28592453]
10. Svanberg E, Spector A, Stott J. The impact of young onset dementia on the family: a literature review. Int Psychogeriatr 2011 Apr;23(3):356-371. [doi: 10.1017/S1041610210001353] [Medline: 20735892]
11. Tookey S, Greaves CV, Rohrer JD, Desai R, Stott J. Exploring experiences and needs of spousal carers of people with behavioural variant frontotemporal dementia (bvFTD) including those with familial FTD (fFTD): a qualitative study. BMC Geriatr 2022 Mar 7;22(1):185. [doi: 10.1186/s12877-022-02867-1] [Medline: 35255821]
12. Sullivan MP, Williams V, Grillo A, et al. Peer support for people living with rare or young onset dementia: an integrative review. Dementia (London) 2022 Nov;21(8):2700-2726. [doi: 10.1177/14713012221126368] [Medline: 36114712]
13. Suárez-González A, John A, Brotherhood E, et al. "Better Living with Non-memory-led Dementia": protocol for a feasibility randomised controlled trial of a web-based caregiver educational programme. Pilot Feasibility Stud 2023 Oct 11;9(1):172. [doi: 10.1186/s40814-023-01403-1] [Medline: 37821924]
14. Suárez-González A, Zimmermann N, Waddington C, et al. Non-memory led dementias: care in the time of covid-19. BMJ 2020 Jun 30;369:m2489. [doi: 10.1136/bmj.m2489] [Medline: 32606068]
15. Kung CSJ, Steptoe A. Changes in Internet use patterns among older adults in England from before to after the outbreak of the COVID-19 pandemic. Sci Rep 2023 Mar 9;13(1):3932. [doi: 10.1038/s41598-023-30882-8] [Medline: 36894600]
16. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view1. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
17. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology1. MIS Q 2012 Mar 1;36(1):157-178. [doi: 10.2307/41410412]
18. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
19. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
20. Chen K, Lou VWQ. Measuring senior technology acceptance: development of a brief, 14-item scale. Innov Aging 2020;4(3):igaa016. [doi: 10.1093/geroni/igaa016] [Medline: 32617418]
21. Ha J, Park HK. Factors affecting the acceptability of technology in health care among older Korean adults with multiple chronic conditions: a cross-sectional study adopting the senior technology acceptance model. Clin Interv Aging 2020;15:1873-1881. [doi: 10.2147/CIA.S268606] [Medline: 33061336]
22. Harris MT, Rogers WA. Developing a healthcare technology acceptance model (H-TAM) for older adults with hypertension. Ageing Soc 2023 Apr;43(4):814-834. [doi: 10.1017/s0144686x21001069] [Medline: 37007645]
23. DeLange Martinez P, Tancredi D, Pavel M, Garcia L, Young HM. Technology acceptance among low-income Asian American older adults: cross-sectional survey analysis. J Med Internet Res 2024 Nov 22;26:e52498. [doi: 10.2196/52498] [Medline: 39576987]
24. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ–4. Psychosomatics 2009 Nov;50(6):613-621. [doi: 10.1016/S0033-3182(09)70864-3]
25. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [Medline: 5349366]
26. Hong W, Thong JYL. Internet privacy concerns: an integrated conceptualization and four empirical studies1. MIS Q 2013 Mar 1;37(1):275-298. [doi: 10.25300/MISQ/2013/37.1.12]
27. Nelson LA, Pennings JS, Sommer EC, Popescu F, Barkin SL. A 3-item measure of digital health care literacy: development and validation study. JMIR Form Res 2022 Apr 29;6(4):e36043. [doi: 10.2196/36043] [Medline: 35486413]
28. Brotherhood EV, Stott J, Windle G, et al. Protocol for the rare dementia support impact study: RDS impact. Int J Geriatr Psychiatry 2020 Aug;35(8):833-841. [doi: 10.1002/gps.5253] [Medline: 31876030]

XSL•FO
RenderX

29. Advice community learning. Rare Dementia Support. URL: https://www.raredementiasupport.org/ [accessed 2026-02-15]
30. Austrom MG, Geros KN, Hemmerlein K, et al. Use of a multiparty web based videoconference support group for family caregivers: innovative practice. Dementia (London) 2015 Sep;14(5):682-690. [doi: 10.1177/1471301214544338] [Medline: 25062788]
31. Banbury A, Parkinson L, Gordon S, Wood D. Implementing a peer-support programme by group videoconferencing for isolated carers of people with dementia. J Telemed Telecare 2019 Oct;25(9):572-577. [doi: 10.1177/1357633X19873793] [Medline: 31631761]
32. Loseto-Gerritzen EV, McDermott O, Orrell M. Development of a best practice guidance on online peer support for people with young-onset dementia. Behav Sci (Basel) 2024 Aug 26;14(9):746. [doi: 10.3390/bs14090746] [Medline: 39335962]
33. Lundberg S. The results from a two-year case study of an information and communication technology support system for family caregivers. Disabil Rehabil Assist Technol 2014 Jul;9(4):353-358. [doi: 10.3109/17483107.2013.814170] [Medline: 23855387]
34. El Baou C, Saunders R, Buckman JEJ, et al. Effectiveness of psychological therapies for depression and anxiety in atypical dementia. Alzheimers Dement 2024 Dec;20(12):8844-8854. [doi: 10.1002/alz.14332] [Medline: 39439362]
35. Lattie EG, Stiles-Shields C, Graham AK. An overview of and recommendations for more accessible digital mental health services. Nat Rev Psychol 2022 Feb;1(2):87-100. [doi: 10.1038/s44159-021-00003-1] [Medline: 38515434]
36. Álvarez-Aguado I, Vega Córdova V, Muñoz La Rivera F, et al. Exploring technology use among older adults with intellectual disabilities: barriers, opportunities, and the role of advanced technologies. Disabil Rehabil Assist Technol 2025 Oct;20(7):2249-2260. [doi: 10.1080/17483107.2025.2498566] [Medline: 40293959]
37. Teles S, Paúl C, Costa-Santos C, Ferreira A. Use of dementia and caregiving-related internet resources by informal caregivers: a cross-sectional study. Front Med (Lausanne) 2022;9(978635):978635. [doi: 10.3389/fmed.2022.978635] [Medline: 36186787]
38. Wong C, Merrilees J, Ketelle R, Barton C, Wallhagen M, Miller B. The experience of caregiving: differences between behavioral variant of frontotemporal dementia and Alzheimer disease. Am J Geriatr Psychiatry 2012 Aug;20(8):724-728. [doi: 10.1097/JGP.0b013e318233154d] [Medline: 21941168]
39. Mioshi E, Foxe D, Leslie F, et al. The impact of dementia severity on caregiver burden in frontotemporal dementia and Alzheimer disease. Alzheimer Dis Assoc Disord 2013;27(1):68-73. [doi: 10.1097/WAD.0b013e318247a0bc] [Medline: 22314247]
40. Messina A, Annoni AM, Amati R, et al. Participatory methods in designing digital health interventions for informal caregivers of people with dementia. A systematic review. Internet Interv 2025 Mar;39:100799. [doi: 10.1016/j.invent.2024.100799] [Medline: 40161473]
41. Naunton Morgan B, Windle G, Lamers C, Brotherhood E, Crutch S, Rare Dementia Support (RDS) Impact Project Research Team. Adaptation of an eHealth intervention: iSupport for carers of people with rare dementias. Int J Environ Res Public Health 2023 Dec 28;21(1):47. [doi: 10.3390/ijerph21010047] [Medline: 38248512]
42. Talbot CV, Briggs P. The use of digital technologies by people with mild-to-moderate dementia during the COVID-19 pandemic: a positive technology perspective. Dementia (London) 2022 May;21(4):1363-1380. [doi: 10.1177/14713012221079477] [Medline: 35333111]
43. Paolillo EW, Casaletto KB, Clark AL, et al. Examining associations between smartphone use and clinical severity in frontotemporal dementia: proof-of-concept study. JMIR Aging 2024 Jun 26;7:e52831. [doi: 10.2196/52831] [Medline: 38922667]
44. Rogalski E, Roberts A, Salley E, et al. Communication partner engagement: a relevant factor for functional outcomes in speech-language therapy for aphasic dementia. J Gerontol B Psychol Sci Soc Sci 2022 Jun 1;77(6):1017-1025. [doi: 10.1093/geronb/gbab165] [Medline: 34528692]
45. Nakagomi A, Kondo K, Shiba K. Heterogeneity in the association between internet use and dementia among older adults: a machine-learning analysis. Arch Gerontol Geriatr 2025 Sep;136:105912. [doi: 10.1016/j.archger.2025.105912] [Medline: 40456200]
46. Cameron E, Mansfield E, Ampofo A, Coda A, Boyes A. Internet access and use among dementia carers and the people they support in Australia: cross-sectional survey. JMIR Form Res 2025 Aug 22;9:e68333. [doi: 10.2196/68333] [Medline: 40845330]

Abbreviations

bvFTD: behavioral variant frontotemporal dementia
HCP: health care professional
NMLD: non–memory-led dementia
PCA: posterior cortical atrophy
PHQ-4: Patient Health Questionnaire–4-item
PPA: primary progressive aphasia
PPI: patient and public involvement

XSL•FO
RenderX

Short STAM: Short Senior Technology Acceptance Model
TU: technology use
UB: usage behavior
UCL: University College London
UTAUT: Unified theory of acceptance and use of technology

Edited by Q Fan; submitted 15.Oct.2025; peer-reviewed by MM Sikder, X Ding; revised version received 22.Jan.2026; accepted 02.Feb.2026; published 30.Mar.2026.

Please cite as:
Brotherhood EV, El Baou C, Fearn C, Hayes O, Zimmermann N, Tsipa A, Crutch SJ, Stott J
Factors Associated With the Use of Digital Health and Well-Being Resources in Non–Memory-Led Dementias: Quantitative Survey Study
JMIR Aging 2026;9:e85863
URL: https://aging.jmir.org/2026/1/e85863
doi:10.2196/85863

© Emilie V Brotherhood, Céline El Baou, Caroline Fearn, Oliver Hayes, Nikki Zimmermann, Anastasia Tsipa, Sebastian J Crutch, Joshua Stott. Originally published in JMIR Aging (https://aging.jmir.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Attitudes Toward Digital Meal Assistance Services Among Older Adults in China: Cross-Sectional Survey

Zhe Chen[1], MSc; Tinghe Deng[2], PhD; Xiaoli Zhu[3]

[1]Department of Economics and Management, Century College, Beijing University of Posts and Telecommunications, Beijing, China

[2]Chinese Center for Health Education, Building 12, Anhua Xili First District, Chaoyang District, Beijing, China

[3]Overseas Chinese College, Capital University of Economics and Business, Beijing, China

Corresponding Author:
Tinghe Deng, PhD
Chinese Center for Health Education, Building 12, Anhua Xili First District, Chaoyang District, Beijing, China

Abstract

Background: Digital eldercare services are promoted to address gaps in traditional care provision, yet actual uptake among older adults remains limited. Prior research has largely focused on technological attributes, with less attention to older adults' context-specific behavioral attitudes toward digital services.

Objective: This study aimed to identify the multidimensional structure of older adults' behavioral attitudes toward digital meal assistance services and to examine their associations with usage intention and actual use.

Methods: A mixed-mode survey was conducted in 2020 among adults aged 60 years or older in a large metropolitan city in China. A total of 1019 valid questionnaires were collected (n=405 online and n=614 offline). Exploratory factor analysis was used to identify attitudinal dimensions, followed by binary logistic regression analyses conducted separately for online and offline samples, controlling for sociodemographic characteristics.

Results: Five attitudinal dimensions were identified, explaining 60.38% of the variance. Regression analyses showed that associations between attitudinal dimensions and service use differed by usage stage and sample type. Digital trust was the most consistent predictor, significantly associated with intention to use in both samples—particularly online (odds ratio [OR] 1.351, 95% CI 1.054 - 1.732; *P*=.02)—and with actual use in the offline sample (OR 1.068, 95% CI 1.011 - 1.128; *P*=.02), with a marginal association online (OR 1.347, 95% CI 0.986 - 1.840; *P*=.06).

Conclusions: Older adults' adoption of digital meal assistance services is shaped by multidimensional and context-dependent attitudes that extend beyond technological considerations. Policy efforts should therefore emphasize trust-building, affordability, and differentiated strategies tailored to diverse groups of older adults.

(JMIR Aging 2026;9:e84956) doi:10.2196/84956

KEYWORDS

digital meal assistance services; older adults; behavioral attitudes; usage intention; actual use; China

Introduction

Digital older care services are increasingly viewed as an important approach to addressing population aging and alleviating shortages in traditional eldercare provision. Prior studies indicate that digital technologies, such as the internet and big data, can integrate service resources through smart platforms. These platforms enable the delivery of diverse services tailored to older adults' needs, including home-based care, remote medical consultation, rehabilitation guidance, and health management [1,2]. In China, national policies have further promoted the integration of digital technologies and eldercare services. The Smart Health and Elderly Care Industry Development Action Plan (2021 - 2025) explicitly encourages the digitalization and age-friendly application of services such as meal delivery and household assistance, aiming to narrow the digital divide among older adults [3]. Despite these policy efforts and technological advances, the actual use of digital eldercare services among older adults remains limited [2,4].

Digital meal assistance services represent a key component of digital eldercare by linking daily nutritional support with the broader eldercare service delivery system through platform-based models. Meals on Wheels services were initially designed for older adults with difficulties in meal preparation and have gradually expanded to serve broader populations [5,6]. Existing research suggests that meal assistance services can improve nutritional intake and health outcomes, reduce family caregiving burden, and promote social interaction [7-9]. With digitalization, these services have evolved into multiple forms, including online meal delivery platforms and food delivery applications [10-12]. However, empirical evidence indicates that such services have not been widely adopted by older adults,

XSL•FO
RenderX

potentially undermining food security and health outcomes [2,7,10].

Previous studies have mainly relied on the Andersen behavioral model, the Technology Acceptance Model (TAM), and the theory of planned behavior (TPB) to explain digital service use, highlighting the roles of perceived usefulness, perceived ease of use, and behavioral attitudes in shaping usage intention [13-16]. Nevertheless, 2 gaps remain. First, many studies focus on the general population and give insufficient attention to older adults, who differ markedly in digital skills and risk perceptions [10,17]. Second, although behavioral attitudes are frequently emphasized, their underlying structure among older adults remains insufficiently specified, and the mechanisms through which attitudes translate into actual service use—the intention-behavior gap—are not fully understood [14,16,18].

To address these gaps, this study examines digital meal assistance services in urban China using survey data collected from older adults in 2020. We aim to (1) identify the structure of older adults' behavioral attitudes toward digital meal assistance services through exploratory factor analysis (EFA); (2) examine differences in attitudinal structures across population subgroups; and (3) assess the associations between different attitudinal dimensions, usage intention, and actual service use. By adopting a behavioral attitude perspective, this study seeks to clarify the mechanisms underlying digital eldercare service adoption among older adults and to provide empirical insights for service design and policy development.

Methods

Study Design and Setting

This study used data from a questionnaire survey examining older adults' intention to use smart eldercare services. The survey was conducted over a 2-month period in 2020 in 2 districts of a large metropolitan city in China. The survey collected information on respondents' sociodemographic characteristics and their awareness and use of digital meal assistance services (ie, online food ordering services). Given potential systematic differences between online and offline surveys in digital skills and service access, a mixed-mode design was adopted. Descriptive and regression analyses were conducted separately for online and offline samples. The final analytic sample comprised 1019 valid responses (n=405 online and n=614 offline). Additional details on the sampling framework and survey implementation are provided in Multimedia Appendix 1.

Offline Survey Procedures

The offline survey used a 2-stage stratified random sampling strategy. Streets within each district were first selected using simple random sampling. Communities within selected streets were then stratified by geographic accessibility to community eldercare service stations (approximately 15-min walkability, defined as about 1.0 - 1.2 km based on navigation distance) and randomly sampled within each stratum. Eligible older adults were recruited through intercept surveys, yielding 614 valid questionnaires. Further procedural details are provided in Multimedia Appendix 1.

Online Survey Procedures and Quality Control

The online survey was administered via the Wenjuanxing platform and yielded 405 valid responses. Survey links were primarily distributed by community staff, with additional responses collected through the platform's mutual-assistance service. No incentives were provided. Eligibility was ensured through an automated age-screening item (≥60 y), and duplicate submissions were prevented by restricting 1 response per IP address. Participation was voluntary, and informed consent was implied by questionnaire submission. Additional quality control details are provided in Multimedia Appendix 1, and the completed CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist is provided in Checklist 1.

Measures

Behavioral Attitudes Toward Digital Meal Services

According to the TPB, behavioral attitudes are derived from individuals' beliefs about behavioral outcomes and are influenced by perceived behavioral control [19]. When individuals perceive that they have sufficient ability and resources to perform a behavior, they are more likely to develop positive attitudes. The TAM further posits that perceived usefulness and perceived ease of use are core mechanisms underlying attitude formation, with external factors influencing attitudes indirectly through these perceptions [20-22]. Building on the TPB and the TAM, this study incorporates the digital divide experienced by older adults as well as the family-oriented eldercare context in China [23-25]. Behavioral attitudes are defined as older adults' overall evaluations of the value, acceptability, and safety of digital meal assistance services within a specific service context, as well as their perceived impacts on themselves and their families. From this perspective, attitudes toward digital meal assistance services encompass evaluations of usefulness and convenience. They also involve judgments related to family roles and caregiving responsibilities, trust-related risk assessments, and perceived constraints associated with resources and affordability.

Measurement items were developed based on a systematic review of prior studies on eldercare service use, digital technology adoption, and digital inequality. The initial item pool primarily covered perceived service value, service acceptance, safety concerns, and perceived impacts on individuals and families. To reduce social desirability and self-presentation bias, some items were framed in a projective manner (eg, asking how "older adults around you" perceive digital meal services). Reverse-coded items were also included to minimize acquiescence bias. All items were measured using a 5-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). Following semantic refinement and standardization prior to the formal survey, a total of 19 items were retained for analysis.

Based on the results of EFA, composite indices were constructed for each behavioral attitude dimension. Scores for each dimension were calculated as the mean of the corresponding items, with higher scores indicating more positive attitudes. Internal consistency for each attitudinal dimension was assessed using Cronbach α.

XSL•FO
RenderX

Outcome Variables

To assess the effects of behavioral attitudes, 2 outcome variables were examined: intention to use digital meal assistance services and actual service use. Usage intention was measured using the item "Would you (continue to) order meals through internet-based platforms?" on a 5-point Likert scale (1="definitely would not use," 5="definitely would use") and dichotomized as 1 for responses of 4 or 5 and 0 otherwise. Actual service use was measured by the frequency of online food ordering and coded as 1 for any reported use and 0 for "never."

Ethical Considerations

This study involved an anonymous questionnaire survey administered online (via Wenjuanxing) and offline face-to-face. No personally identifiable information (eg, names, contact details, identification numbers, or precise addresses) or sensitive personal data (eg, medical or financial information) was collected, and no intervention, clinical procedures, or biological sample collection was involved. Under Article 32 of the Measures for Ethical Review of Life Sciences and Medical Research Involving Human Subjects, research using anonymized information data that does not cause harm to participants and does not involve sensitive personal information or commercial interests may be exempt from ethics review. Accordingly, this anonymous, minimal-risk survey study met the criteria for exemption from institutional ethics review.

Participants were informed about the study purpose, procedures, voluntary participation, and data protection. For the online survey, consent was implied by submission of the completed questionnaire; for the offline survey, oral informed consent was obtained prior to participation. Participants could withdraw at any time without consequences. No compensation was provided for participation.

Data Privacy and Security

All survey data were collected anonymously. Online responses were exported and stored in password-protected files, and offline paper questionnaires (if applicable) were securely stored and then digitized. Only the research team had access to the data.

Data Analysis

EFA was conducted to identify the latent structure of older adults' behavioral attitudes toward digital meal assistance services. Prior to analysis, questionnaires with missing values on key variables or substantial nonresponse were excluded. The final analytic sample consisted of 1019 complete questionnaires; therefore, no imputation was applied.

EFA followed standard procedures of item screening, factor extraction, rotation, and interpretation. Factors were extracted using principal component analysis, with the number of factors determined by the Kaiser criterion (eigenvalues >1). Varimax orthogonal rotation was applied. Items with factor loadings below 0.50 were considered for removal; items with cross-loadings (loading differences <0.20 across factors) or communalities below 0.40 were also evaluated and excluded when appropriate.

Binary logistic regression models were used to examine associations between behavioral attitude dimensions and both usage intention and actual service use. Given differences between online and offline samples, all models were estimated separately for each subsample. The general model specification was

$$(1)\log(pi1-pi)=\beta0+\beta1Xi+\beta2Zi$$

where pi denotes the probability that individual i reports a higher usage intention or actual service use; Xi represents behavioral attitudes (or specific attitudinal dimensions); and Zi includes control variables (gender, age, education, and income). Results were reported as odds ratios (ORs) with 95% CIs. Given the sensitivity of the Hosmer-Lemeshow test in large samples, pseudo-R^2 and the Akaike information criterion were also reported as supplementary fit indicators.

Results

Participant Characteristics

Table 1 summarizes the sociodemographic characteristics of the sample. Overall, the age distribution was skewed toward younger older adults, with participants aged 60‐69 years accounting for 569 of 1019 (55.8%) of the sample, which may limit generalizability. Compared with the offline sample, the online sample was younger on average, whereas the offline sample included a higher proportion of the oldest-old (aged ≥80 y). In terms of socioeconomic characteristics, the online sample had higher educational attainment than the offline sample, with 272 of 405 (67.2%) reporting a high school education or above compared with 297 of 614 (48.4%) in the offline sample. The online sample also had a higher proportion of high-income respondents (≥8000 CNY, 1 CNY=US $0.14), whereas the offline sample included a relatively larger share of middle-income respondents (4000‐7999 CNY). Overall, the offline sample more closely resembled community-dwelling older adults, whereas the online sample was more likely to represent younger and more digitally skilled individuals. Accordingly, the findings were more applicable to younger and older adults and those with greater digital capacity, and caution was warranted when generalizing to the oldest-old or to older adults with limited digital skills; these limitations are discussed in the Limitations section.

XSL•FO
RenderX

Table . Sample characteristics of the study participants (n=1019).

Variables and categories	Total sample, n (%)[a]	Online survey, n (%)[a]	Offline survey, n (%)[a]
Gender			
Male	477 (46.815)	223 (55.062)	254 (41.368)
Female	542 (53.185)	182 (44.938)	360 (58.632)
Education level			
No formal education	97 (9.519)	3 (0.741)	94 (15.309)
Primary school	82 (8.048)	8 (1.975)	74 (12.052)
Junior high school	271 (26.595)	122 (30.123)	149 (24.266)
Senior high school or above	569 (55.838)	272 (67.160)	297 (48.373)
Age group (y)			
60 - 69	569 (55.838)	212 (52.346)	357 (58.143)
70 - 79	365 (35.819)	173 (42.716)	192 (31.270)
≥80	85 (8.342)	20 (4.938)	65 (10.587)
Self-rated health			
Very poor	106 (10.402)	38 (9.383)	68 (11.074)
Poor	40 (3.926)	24 (5.926)	16 (2.606)
Fair	139 (13.640)	91 (22.469)	48 (7.820)
Good	538 (52.796)	153 (37.778)	385 (62.703)
Very good	196 (19.235)	99 (24.444)	97 (15.798)
Marital status			
Married	850 (83.415)	361 (89.136)	489 (79.641)
Unmarried	63 (6.183)	23 (5.679)	40 (6.514)
Widowed	106 (10.402)	21 (5.185)	85 (13.844)
Monthly income (CNY[b])			
0 - 3999	213 (20.902)	92 (22.716)	121 (19.706)
4000 - 7999	485 (47.596)	124 (30.617)	361 (58.795)
≥8000	321 (31.502)	189 (46.667)	132 (21.498)

[a]Percentages may not sum to 100% due to rounding.

[b]CNY: Chinese yuan; 1 CNY=US $0.14.

Behavioral Attitude Structure

EFA was conducted to identify the latent structure of older adults' behavioral attitudes toward digital meal assistance services. Data suitability was confirmed by a significant Bartlett test of sphericity (χ^2=286.67, P<.01) and a Kaiser-Meyer-Olkin value of 0.894, indicating adequate interitem correlations. After excluding 3 items based on predefined criteria, 16 items were retained for analysis. Five factors with eigenvalues greater than 1 were extracted, with initial eigenvalues of 6.01, 2.43, 1.82, 1.39, and 1.03. Following Varimax orthogonal rotation, the 5 factors explained 14.60%, 14.13%, 11.64%, 11.09%, and 8.92% of the total variance, respectively, yielding a cumulative explained variance of 60.38%. Detailed results are presented in Table 2.

Table 3 presents the standardized factor loadings and communalities for each retained item. All items showed strong loadings on their corresponding factors, and each factor was represented by at least 3 items, indicating a clear and stable factor structure. Reliability analysis showed that the overall scale demonstrated good internal consistency, with a Cronbach α of 0.879. In addition, correlations between the mean scores of individual behavioral attitude dimensions and the overall scale mean ranged from 0.414 to 0.762, suggesting that the dimensions were moderately correlated while maintaining adequate discriminant validity (Multimedia Appendix 2).

XSL•FO
RenderX

Table . Rotated total variance explained.[a]

Factor	Eigenvalue (rotated)	Variance explained (%)	Cumulative variance explained (%)
F1	3.065	14.60	14.60
F2	2.968	14.13	28.73
F3	2.444	11.64	40.37
F4	2.330	11.09	51.46
F5	1.873	8.92	60.38

[a]Variance explained is based on principal component analysis with Varimax rotation.

XSL•FO
RenderX

Table . Exploratory factor analysis loading matrix.[a]

Item	F1	F2	F3	F4	F5	Communality
I am likely to use online meal ordering services when cooking on my own is difficult	0.773	—[b]	—	—	—	0.754
I am likely to use online meal ordering services when conditions for cooking at home are unfavorable or when I do not wish to cook	0.679	—	—	—	—	0.627
I am likely to use online meal ordering services when my children experience difficulty in providing me with meals	0.649	—	—	—	—	0.529
I am likely to use online meal ordering services when going out is inconvenient due to extreme weather	—	0.717	—	—	—	0.726
I am likely to use online meal ordering services for convenience and to save time and effort	—	0.718	—	—	—	0.664
I am likely to use online meal ordering services when the food I want is complicated to prepare	—	0.686	—	—	—	0.562
I avoid using online meal ordering services because of concerns about being perceived as lazy by others	—	—	0.769	—	—	0.779
I avoid using online meal ordering services because of concerns that it may reduce children's sense of responsibility	—	—	0.724	—	—	0.701
I avoid using online meal ordering services due to concerns about upsetting children	—	—	0.723	—	—	0.678

XSL•FO
RenderX

Item	F1	F2	F3	F4	F5	Communality
I avoid using online meal ordering services because I am not accustomed to ordering food online	—	—	—	0.608	—	0.612
I am more likely to use online meal ordering services when people of my age around me use such services	—	—	—	0.713	—	0.725
I avoid using online meal ordering services because I am concerned about making operational mistakes or being scammed	—	—	—	0.704	—	0.711
I avoid using online meal ordering services because of concerns about food quality or hygiene during preparation	—	—	—	0.691	—	0.606
I avoid using online meal ordering services if my children have to pay for the cost	—	—	—	—	0.652	0.662
I am more likely to use online meal ordering services when government subsidies are available	—	—	—	—	0.583	0.584
I am willing to use online meal ordering services only when my pension income is sufficient to cover the cost	—	—	—	—	0.671	0.681

[a]Only factor loadings with absolute values ≥0.50 are reported. Factor extraction was based on principal component analysis with Varimax rotation.

[b]Not reported (factor <0.50).

Based on the content characteristics of items with high factor loadings, 5 factors were labeled. Alleviating burden (F1) reflects value-based judgments regarding whether digital meal assistance services were worth using, particularly in terms of their potential to reduce the caregiving burden on adult children in daily life. When such services were perceived as helping to alleviate family caregiving pressure, older adults tended to hold more positive attitudes toward their use. Enhancing convenience (F2) captures perceived evaluations of the extent to which digital meal assistance services improve convenience in older adults' daily lives. This factor reflected the functional use of the service, consistent with perceived usefulness in the TAM. Service acceptance (F3) represented older adults' evaluations of the social meaning of digital meal assistance services. When service use was interpreted as indicating that adult children had not adequately fulfilled their caregiving responsibilities, attitudes toward the service tended to be negative. Digital trust (F4) reflected older adults' attitudinal judgments regarding the safety and reliability of digital meal assistance services. Funding sources (F5) captured older adults' perceived affordability of digital meal assistance services, consistent with perceived behavioral control in the TPB. Lower perceived affordability was associated with a lower likelihood of service use.

Notably, perceived ease of use, as defined in the TAM, did not emerge as an independent attitudinal dimension in the identified factor structure.

XSL•FO
RenderX

Differences in Behavioral Attitudes Across Groups

Table 4 summarizes mean scores across behavioral attitude dimensions by participant characteristics. Overall, older adults reported relatively higher evaluations on alleviating burden (F1) and enhancing convenience (F2). Middle- and higher-income groups generally scored higher on enhancing convenience (F2), service acceptance (F3), and digital trust (F4), whereas lower-income older adults reported higher scores on funding sources (F5) and alleviating burden (F1). Higher educational attainment was associated with more positive evaluations of alleviating burden (F1), enhancing convenience (F2), and service acceptance (F3), as well as lower concern regarding funding sources (F5). By contrast, differences in digital trust (F4) were modest across education groups. By survey mode, the online sample consistently reported higher scores across most attitude dimensions—particularly enhancing convenience (F2)—whereas differences in digital trust (F4) between online and offline samples were relatively small.

Table . Mean scores of behavioral attitude dimensions across different groups of older adults.[a]

Behavioral attitude dimension	Total sample	Middle- or high-income[b]	Low-income[b]	Higher education	Lower education[c]	Offline sample[d]	Online sample[d]
Alleviating burden (F1), mean (SD)	3.668 (1.412)	3.278 (1.357)	3.816 (1.602)	3.700 (1.288)	3.346 (1.492)	3.278 (1.525)	3.417 (1.064)
Enhancing convenience (F2), mean (SD)	3.588 (1.500)	3.763 (1.410)	3.237 (1.762)	3.604 (1.287)	3.251 (1.650)	3.054 (1.682)	3.611 (0.784)
Service acceptance (F3), mean (SD)	3.334 (1.890)	3.411 (1.999)	3.146 (1.35)	3.551 (2.029)	3.228 (1.681)	3.209 (2.00)	3.441 (0.784)
Digital trust (F4), mean (SD)	3.022 (1.226)	3.082 (1.023)	2.927 (1.79)	2.962 (1.228)	3.002 (1.219)	2.989 (1.32)	3.132 (1.063)
Funding sources (F5), mean (SD)	3.010 (1.947)	2.711 (1.953)	3.132 (1.812)	3.002 (2.140)	3.359 (1.703)	2.796 (1.961)	3.364 (1.807)

[a]Values represent mean scores on a 5-point Likert scale (1=strongly disagree, 5=strongly agree).

[b]Low-income refers to a monthly income of 0-3999 CNY; middle- or high-income refers to ≥4000 CNY; 1 CNY=US $0.14.

[c]Lower education refers to junior high school or below; higher education refers to high school or above.

[d]Offline and online samples were collected using different survey modes as described in the Methods section.

Associations Between Behavioral Attitudes, Usage Intention, and Actual Use

Tables 5 and 6 present logistic regression results examining associations between the 5 behavioral attitude dimensions and both intention to use and actual use of digital meal assistance services, estimated separately for the online and offline samples. All 4 models were statistically significant (P<.001), with pseudo-R^2 values ranging from 0.156 to 0.240 and Akaike information criterion values from 800.001 to 1002.470, indicating acceptable model fit. Hosmer-Lemeshow tests did not suggest a significant lack of fit (P value range .055-.062).

XSL•FO
RenderX

Table . Logistic regression results for the associations between behavioral attitude dimensions, intention to use, and actual use of digital meal assistance services.

Variables	Panel A, online sample (n=405)		Panel B, offline sample (n=614)	
	Intention to use	Actual use	Intention to use	Actual use
Alleviating burden (F1), OR[a] (95% CI; *P* value)	1.618[b] (1.235-2.118; <.001)	1.742[b] (1.286-2.359;<.001)	1.673[b] (1.420-1.969;<.001)	2.340 (0.710-7.727; .16)
Enhancing convenience (F2), OR (95% CI; *P* value)	1.107[c] (1.002-1.223; .046)	1.089 (0.900-1.318; .39)	1.129 (0.968-1.317; .13)	0.843 (0.652-1.088; .19)
Service acceptance (F3), OR (95% CI; *P* value)	1.128 (0.979-1.299; .10)	1.250[c] (1.000-1.557; .05)	1.123[c] (1.001-1.259; .049)	1.011 (0.991-1.031; .27)
Digital trust (F4), OR (95% CI; *P* value)	1.351[c] (1.054-1.732; .02)	1.347[d] (0.986-1.840; .06)	1.221[b] (1.083-1.376; <.001)	1.068 (1.011-1.128; .02)
Funding sources (F5), OR (95% CI; *P* value)	1.417[d] (0.957-2.100; .08)	0.800 (0.444-1.440; .46)	0.806[d] (0.642-1.020; .07)	0.600 (0.225-1.598; .31)
Prior use of online meal ordering, OR (95% CI; *P* value)	6.350[b] (3.668 - 10.973;<.001)	N/A[e]	4.773[b] (3.523-6.471;<.001)	N/A
Log likelihood ratio	202.94	67.1	354.25	285.85
Model significance (*P* value)	<.001[b]	<.001[b]	<.001[b]	<.001[b]
Pseudo R^2	0.212	0.156	0.24	0.228
AIC[f]	800.001	888.224	1002.47	979.838
Hosmer-Lemeshow goodness-of-fit (*P* value)[g]	.06	.06	.06	.06

[a]OR: odds ratio.

[b]*P*≤.001.

[c]*P*≤.05.

[d]*P*<.10.

[e]NA: not applicable.

[f]AIC: Akaike information criterion.

[g]The Hosmer-Lemeshow goodness-of-fit test is sensitive to large sample sizes; therefore, pseudo R^2 and Akaike information criterion are reported as complementary indicators of model fit.

Table . Control variables included in all models (definitions, distinctions, and coding).

Control variables[a]	What it means (to distinguish similarly named variables)	Coding/categories[b]
Gender	Binary indicator	Male=1; female=0
Education level	Highest attained education category (not y of schooling)	No formal education=1; primary school=2; junior high school=3; high school and above=4
Age group	Categorical age bands (not continuous age in y)	60-69=1; 70-79=2; ≥80=3
Self-rated health	5-point subjective health rating	Very poor=1 to very good=5
Marital status	Current marital status	Married=1; unmarried/widowed=0
Monthly household income (CNY)[c]	Household-level total monthly income (not personal income)	0-3999 CNY=1; 4000-7999 CNY=2; ≥8000 CNY=3

[a]Control variables were included in all models reported in Table 5.

[b]For binary variables, the reference category is coded as 0 (female; unmarried/widowed). For ordinal variables entered as numeric scores (education level, age group, self-rated health, and monthly household income), odds ratios represent the change in odds per 1-level increase in the coded category.

[c]Exchange rate for interpretation (study period Jul-Aug 2020): 1 CNY=US $0.14.

In the online sample, alleviating burden (F 1) was positively associated with both usage intention and actual use. A 1-unit increase in F1 was associated with a 61.8% increase in the odds of intending to use the service (OR 1.618, 95% CI 1.235 - 2.118; *P*<.001) and a 74.2% increase in the odds of actual use (OR 1.742, 95% CI 1.286 - 2.359; *P*<.001). In addition, digital trust

XSL•FO
RenderX

(F4) was positively associated with usage intention, with each 1-unit increase corresponding to a 35.1% increase in the odds of intention to use (OR 1.351, 95% CI 1.054 - 1.732; *P*=.02), and showed a marginal association with actual use (OR 1.347, 95% CI 0.986 - 1.840; *P*=.06). Service acceptance (F3) was not significantly associated with usage intention (*P*=.10) but demonstrated a marginal positive association with actual use (OR 1.250, 95% CI 1.000 - 1.557; *P*=.05).

In the offline sample, alleviating burden (F1), service acceptance (F3), and digital trust (F4) were all positively associated with usage intention. Notably, a 1-unit increase in digital trust (F4) was associated with a 22.1% increase in the odds of intention to use (OR 1.221, 95% CI 1.083 - 1.376; *P*<.001). In contrast, funding sources (F5) showed a marginal negative association with usage intention (OR 0.806, 95% CI 0.642 - 1.020; *P*=.07). With respect to actual use, digital trust (F4) was the only attitude dimension that remained significantly associated with service use in the offline sample, with a 1-unit increase corresponding to a 6.8% increase in the odds of actual use (OR 1.068, 95% CI 1.011 - 1.128; *P*=.02). Although the odds ratio for alleviating burden (F1) in predicting actual use appeared large in the offline sample, this association did not reach statistical significance, likely reflecting limited statistical power and greater heterogeneity in offline respondents.

Discussion

This study uses digital meal assistance services as a case to explain the phenomenon of "adequate supply but insufficient use" of digital eldercare services. Based on survey data from 1019 older adults in a large metropolitan city in China, we examined service adoption from a behavioral attitude perspective. The results show that older adults' attitudes toward digital meal assistance services have a multidimensional structure. Different attitudinal dimensions play distinct roles in shaping usage intention and actual service use. Evaluations of digital eldercare services are not determined by technology alone. They are also influenced by family caregiving contexts, perceived social norms, risk perceptions, and financial constraints. In addition, the effects of behavioral attitudes differ across the transition from intention to actual behavior. These findings provide empirical evidence to inform the promotion of digital services in an aging society.

Older adults' behavioral attitudes toward digital meal assistance services exhibited a clear multidimensional structure. Five dimensions were identified: Alleviating burden (F1), enhancing convenience (F2), service acceptance (F3), digital trust (F4), and funding sources (F5). These findings indicate that evaluations of digital eldercare services extend beyond functional attributes. Family caregiving relationships, perceived social norms, risk perceptions, and financial constraints also shape these evaluations. This pattern is consistent with prior studies showing that older adults' technology adoption is influenced by broader social and contextual factors [26-28]. A purely technology-rational framework that emphasizes usefulness or ease of use alone may therefore be insufficient to explain service choice. Notably, no independent factor of perceived ease of use emerged in this study. Ease of use appears to be embedded within broader evaluations of convenience, trust, and affordability. This suggests that, in specific eldercare service contexts, older adults' attitudinal structures toward technology may differ from those proposed in classical TAMs.

Attitudinal priorities toward digital meal assistance services varied across older adult groups. Most older adults focused on whether these services could help alleviate caregiving burdens, whereas younger-old adults placed greater emphasis on support for daily convenience. As shown in Table 4, the offline sample scored higher on alleviating burden (F1) than on other dimensions, indicating that many older adults tend to view digital meal assistance services as a supplement to the family caregiving system. This pattern is consistent with the Chinese eldercare context, in which limited family caregiving resources, particularly from spouses and adult children, have been an important foundation for the development of formal eldercare services [29,30]. Accordingly, the perceived ability of digital meal assistance services to reduce family caregiving pressure serves as a key reference in older adults' attitude formation. By contrast, the online sample scored higher on enhancing convenience (F2), consistent with previous studies [31-33]. This finding indicates that younger old adults place greater value on digital eldercare services that improve daily convenience and support autonomy. It also suggests that future cohorts of older adults may increasingly emphasize the role of digital services in enhancing quality of life and independent living.

Insufficient digital trust represents a key barrier to the use of digital meal assistance services, particularly among lower-income older adults. The findings show that although low-income older adults expressed relatively high levels of service acceptance (F3), their scores on digital trust (F4) were lower than those of other income groups. This attitudinal pattern is consistent with the regression results, which showed that digital trust was significantly associated with usage intention and selected aspects of actual service use. These findings align with previous research indicating that the use of eldercare services tends to be concentrated among middle- and high-income groups [34-36]. They also shed light on the mechanism through which the digital divide operates. Limited digital skills and resources not only constrain older adults' capacity to use digital services but may also weaken trust, thereby influencing usage decisions. When individuals face simultaneous deficits in digital skills and resources, psychological strain may increase, including reduced trust, which can further limit service use and exacerbate digital inequality [37,38]. Notably, regardless of income level, overall scores on digital trust remained relatively low among older adults, consistent with the distribution shown in Table 4. This pattern highlights trust as a fundamental condition that requires particular attention in the broader promotion of digital eldercare services.

Behavioral attitudes played an important but uneven role in the translation of usage intention into actual behavior. Although the TPB posits behavioral intention as a key predictor of behavior, prior studies have reported inconsistencies in how attitudes relate to intention and actual use [39,40]. In the online sample, alleviating burden (F1) was significantly associated with both usage intention and actual service use. Service

XSL•FO
RenderX

acceptance (F3) was associated only with actual use and showed no significant relationship with usage intention. This pattern suggests that some attitudinal dimensions may be less salient during intention formation but exert a more direct influence on usage decisions. By contrast, in the offline sample, multiple attitudinal dimensions, including alleviating burden (F1), service acceptance (F3), and digital trust (F4), were significantly associated with the usage intention, whereas only digital trust was further translated into actual service use. Funding sources (F5) primarily influenced usage intention and showed no significant association with actual behavior. Overall, behavioral attitudes toward digital eldercare services show clear context dependence across intention and behavior. This provides empirical evidence for the intention–behavior gap in older adults' use of digital services.

These findings have important implications for policy and practice. Greater attention should be given to the vulnerability of low-income older adults in digital environments. Public digital spaces and smart communities can improve access to digital technologies for groups affected by the digital divide [41,42]. Community-based courses and remote assistance programs can support tailored digital skills training [43]. Digital platforms should also reduce barriers through age-friendly interface design, such as larger fonts and simplified procedures [41,44]. In addition, efforts are needed to reduce technology-related anxiety among older adults [45-47]. Trust-enhancing mechanisms, including service quality certification, data security standards, and transparent pricing, can strengthen perceived safety and confidence. Finally, digital eldercare services should better match the needs of different groups. Services should emphasize burden reduction for most older adults and highlight independence and convenience for younger-old or higher-income groups. For low-income older adults, establishing basic trust remains a critical condition for service uptake.

This study has several limitations. First, the sample was skewed toward younger older adults, with relatively few participants aged 80 years or older. The findings may therefore be less representative of the oldest-old population. Second, differences between the online and offline samples may have introduced selection bias. Although key control variables were included, the results reflect associations rather than causal relationships. In addition, the sample was drawn from a single region, which may limit generalizability. Future studies should include larger and more diverse samples, increase representation of the oldest-old, and apply methods such as sample weighting or sensitivity analyses to better assess potential sampling bias.

Funding

The authors received no specific funding for this work.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to participant privacy considerations but are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sampling and fieldwork procedures.
[PDF File, 130 KB - aging_v9i1e84956_app1.pdf]

Multimedia Appendix 2
Correlations between behavioral attitude dimensions and the overall scale score.
[PDF File, 57 KB - aging_v9i1e84956_app2.pdf]

Checklist 1
CHERRIES checklist.
[PDF File, 115 KB - aging_v9i1e84956_app3.pdf]

References

1. FitzGerald C. Multi-level Meals on Wheels insights: a qualitative study. Work Older People 2024 Feb 8;28(1):41-48. [doi: 10.1108/WWOP-12-2022-0063]
2. Namasivayam B. AI for healthy meal preparation in smart cities. EAI Endorsed Trans Smart Cities 2022;6(4):e1. [doi: 10.4108/eetsc.v6i4.2267]
3. Action plan for the development of smart health and elderly care industry (2021-2025). : Ministry of Industry and Information Technology of the People's Republic of China, Ministry of Civil Affairs, National Health Commission; 2021 URL: https:/

XSL•FO
RenderX

/www.miit.gov.cn/cms_files/filemanager/1226211233/attach/202110/c9d0dfacc13d411cadd44bb6190db4f2.pdf [accessed 2026-03-03]
4. Hill R, Betts LR, Gardner SE. Older adults' experiences and perceptions of digital technology: (dis)empowerment, wellbeing, and inclusion. Comput Human Behav 2015 Jul;48:415-423. [doi: 10.1016/j.chb.2015.01.062]
5. Campbell AD, Godfryd A, Buys DR, Locher JL. Does participation in home-delivered meals programs improve outcomes for older adults? Results of a systematic review. J Nutr Gerontol Geriatr 2015;34(2):124-167. [doi: 10.1080/21551197.2015.1038463] [Medline: 26106985]
6. Zanetta LD, Hakim MP, Gastaldi GB, et al. The use of food delivery apps during the COVID-19 pandemic in Brazil: the role of solidarity, perceived risk, and regional aspects. Food Res Int 2021 Nov;149:110671. [doi: 10.1016/j.foodres.2021.110671] [Medline: 34600673]
7. Papadaki A, Ali B, Cameron A, et al. "It's not just about the dinner; it's about everything else that we do": a qualitative study exploring how Meals on Wheels meet the needs of self-isolating adults during COVID-19. Health Soc Care Community 2022 Sep;30(5):e2012-e2021. [doi: 10.1111/hsc.13634] [Medline: 34766667]
8. Zhu H, An R. Impact of home-delivered meal programs on diet and nutrition among older adults: a review. Nutr Health 2013 Apr;22(2):89-103. [doi: 10.1177/0260106014537146] [Medline: 24916974]
9. Tandon A, Kaur P, Bhatt Y, Mäntymäki M, Dhir A. Why do people purchase from food delivery apps? A consumer value perspective. J Retail Consum Serv 2021 Nov;63:102667. [doi: 10.1016/j.jretconser.2021.102667]
10. Tan SY, Lim SY, Yeo SF. Online food delivery services: cross-sectional study of consumers' attitude in Malaysia during and after the COVID-19 pandemic. F1000Res 2024;10:972. [doi: 10.12688/f1000research.73014.2]
11. Dickinson A, Wills W. Meals on Wheels services and the food security of older people. Health Soc Care Community 2022 Nov;30(6):e6699-e6707. [doi: 10.1111/hsc.14092] [Medline: 36300541]
12. Kaplan M, Hotle S, Heaslip K. How has COVID-19 impacted customer perceptions and demand for delivery services: an exploratory analysis. Transp Policy (Oxf) 2023 Apr;134:217-230. [doi: 10.1016/j.tranpol.2023.02.020] [Medline: 36855673]
13. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [doi: 10.2307/2137284] [Medline: 7738325]
14. Ajzen I. The theory of planned behavior: frequently asked questions. Hum Behav Emerg Technol 2020 Oct;2(4):314-324. [doi: 10.1002/hbe2.195]
15. Kaur P, Dhir A, Talwar S, Ghuman K. The value proposition of food delivery apps from the perspective of theory of consumption value. Transp Policy 2021 May 6;33(4):1129-1159. [doi: 10.1108/IJCHM-05-2020-0477]
16. Jeng MY, Pai FY, Yeh TM. Antecedents for older adults' intention to use smart health wearable devices-technology anxiety as a moderator. Behav Sci (Basel) 2022 Apr 18;12(4):114. [doi: 10.3390/bs12040114] [Medline: 35447686]
17. Jaana M, Tamim H, Paré G. eHealth and the digital divide among older Canadians: insights from a national cross-sectional study. J Med Internet Res 2025 Nov 25;27:e72274. [doi: 10.2196/72274] [Medline: 41289575]
18. Fu X, Sun N, Xu F, et al. Influencing factors of inequity in health services utilization among the elderly in China. Int J Equity Health 2018 Sep 15;17(1):144. [doi: 10.1186/s12939-018-0861-6] [Medline: 30219075]
19. Ajzen I, Fishbein M. Attitude-behavior relations: a theoretical analysis and review of empirical research. Psychol Bull 1977;84(5):888-918. [doi: 10.1037/0033-2909.84.5.888]
20. van Houwelingen CT, Ettema RG, Antonietti MG, Kort HS. Understanding older people's readiness for receiving telehealth: mixed-method study. J Med Internet Res 2018 Apr 6;20(4):e123. [doi: 10.2196/jmir.8407] [Medline: 29625950]
21. Naudé B, Rigaud AS, Pino M. Investigating the acceptability of an interactive television intervention promoting social links among older adults living at home and in care institutions: qualitative interview and questionnaire study. JMIR Hum Factors 2024 Dec 2;11(1):e58051. [doi: 10.2196/58051] [Medline: 39622505]
22. Guner H, Acarturk C. The use and acceptance of ICT by senior citizens: a comparison of Technology Acceptance Model (TAM) for elderly and young adults. Univ Access Inf Soc 2020 Jun;19(2):311-330. [doi: 10.1007/s10209-018-0642-4]
23. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. Inquiry 2022;59:469580221096272. [doi: 10.1177/00469580221096272] [Medline: 35471138]
24. Yap YY, Tan SH, Choon SW. Elderly's intention to use technologies: a systematic literature review. Heliyon 2022 Jan;8(1):e08765. [doi: 10.1016/j.heliyon.2022.e08765] [Medline: 35128090]
25. Hülür G, Macdonald B. Rethinking social relationships in old age: digitalization and the social lives of older adults. Am Psychol 2020;75(4):554-566. [doi: 10.1037/amp0000604] [Medline: 32378949]
26. Berkowsky RW, Sharit J, Czaja SJ. Factors predicting decisions about technology adoption among older adults. Innov Aging 2018 Jan;2(1):igy002. [doi: 10.1093/geroni/igy002] [Medline: 30480129]
27. Kavandi H, Jaana M. Factors that affect health information technology adoption by seniors: a systematic review. Health Soc Care Community 2020 Nov;28(6):1827-1842. [doi: 10.1111/hsc.13011] [Medline: 32378769]
28. Sun X, Yan W, Zhou H, et al. Internet use and need for digital health technology among the elderly: a cross-sectional survey in China. BMC Public Health 2020 Sep 11;20(1):1386. [doi: 10.1186/s12889-020-09448-0] [Medline: 32917171]
29. Wei Y, Zhang L. Analysis of the influencing factors on the preferences of the elderly for the combination of medical care and pension in long-term care facilities based on the Andersen model. Int J Environ Res Public Health 2020 Jul 28;17(15):5436. [doi: 10.3390/ijerph17155436] [Medline: 32731557]

XSL•FO
RenderX

30. Zhan HJ, Montgomery RJV. Gender and elder care in China: the influence of filial piety and structural constraints. Gender Soc 2003;17(2):209-229. [doi: 10.1177/0891243202250734]
31. Zhou Y, He T, Lin F. The digital divide is aging: an intergenerational investigation of social media engagement in China. Int J Environ Res Public Health 2022 Oct 10;19(19):12965. [doi: 10.3390/ijerph191912965] [Medline: 36232259]
32. Fu X, Hwang W, Hadi N, Brown MT, Silverstein M. Young adult grandchildren's intergenerational and digital solidarity with grandparents and depressive symptoms during the pandemic. Fam Relat 2025 May 1:229-244. [doi: 10.1111/fare.13195] [Medline: 40896472]
33. de Rivera L, Ozamiz-Etxebarria N, Dosil-Santamaría M, de Rivera-Monterrey L. Autogenic training improves the subjective perception of physical and psychological health and of interpersonal relational abilities: an electronic field survey during the COVID-19 crisis in Spain. Front Psychol 2021;12:616426. [doi: 10.3389/fpsyg.2021.616426] [Medline: 34393872]
34. Cao J, Kurata K, Lim Y, Sengoku S, Kodama K. Social acceptance of mobile health among young adults in Japan: an extension of the UTAUT model. Int J Environ Res Public Health 2022 Nov 17;19(22):15156. [doi: 10.3390/ijerph192215156] [Medline: 36429875]
35. Beck AM, Hansen KS. Meals served in Danish nursing homes and to Meals-on-Wheels clients may not offer nutritionally adequate choices. J Nutr Elder 2010 Jan;29(1):100-109. [doi: 10.1080/01639360903574742] [Medline: 20391045]
36. Wilson A, Dennison K. Meals on Wheels service: knowledge and perceptions of health professionals and older adults. Nutrition Dietetics 2011 Jun;68(2):155-160. [doi: 10.1111/j.1747-0080.2011.01522.x]
37. Matli W, Ngoepe M. Extending information poverty theory to better understand the digital access and inequalities among young people who are not in education, employment or training in South Africa. High Educ Skills Work-Based Learn 2022 May 9;12(3):419-436. [doi: 10.1108/HESWBL-05-2020-0107]
38. Farajpahlou AH, Koohi Rostami M, Beshlideh K, Pourkhalil N. Information poverty predictors among Ahvazi citizens; investigating some cases. J Librariansh Inf Sci 2022 Dec;54(4):678-691. [doi: 10.1177/09610006211036729]
39. Gómez-Ramirez I, Valencia-Arias A, Duque L. Approach to m-learning acceptance among university students: an integrated model of TPB and TAM. Int Rev Res Open Distrib Learn 2019;20(3):181-199. [doi: 10.19173/irrodl.v20i4.4061]
40. Holden RJ, Karsh BT. The Technology Acceptance Model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
41. Lian JW, Yen DC. Online shopping drivers and barriers for older adults: age and gender differences. Comput Human Behav 2014 Aug;37:133-143. [doi: 10.1016/j.chb.2014.04.028]
42. Saare MA, Hussain A, Yue WS. Conceptualizing mobile health application use intention and adoption among Iraqi older adults: from the perspective of expanded Technology Acceptance Model. Int J Interact Mob Technol 2019;13(10):28. [doi: 10.3991/ijim.v13i10.11285]
43. da Silva Sobrinho AC, da Silva Rodrigues G, de Oliveira GL, de Oliveira Gomes GA, Bueno Júnior CR. Everyday digital support to promote health and literacy among older adults: 14-week randomized digital pilot trial by engagement level. JMIR Form Res 2025 Dec 10;9:e77319. [doi: 10.2196/77319] [Medline: 41370788]
44. Bezirgani A, Lachapelle U. Online grocery shopping for the elderly in Quebec, Canada: the role of mobility impediments and past online shopping experience. Travel Behav Soc 2021 Oct;25:133-143. [doi: 10.1016/j.tbs.2021.07.001]
45. Wale A, Everitt J, Ayres T, et al. Effectiveness of interventions for addressing digital exclusion in older adults in the social care domain: rapid review. JMIR Aging 2025 Dec 30;8:e70377. [doi: 10.2196/70377] [Medline: 41468583]
46. Ha S, Stoel L. Consumer e-shopping acceptance: antecedents in a Technology Acceptance Model. J Bus Res 2009 May;62(5):565-571. [doi: 10.1016/j.jbusres.2008.06.016]
47. Hong C, Choi H, Choi EK, Joung HW. Factors affecting customer intention to use online food delivery services before and during the COVID-19 pandemic. J Hosp Tour Manag 2021 Sep;48:509-518. [doi: 10.1016/j.jhtm.2021.08.012]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
OR: odds ratio
TAM: Technology Acceptance Model

Edited by S Cacciatore; submitted 29.Sep.2025; peer-reviewed by MM Sikder, P Adedigba; revised version received 12.Jan.2026; accepted 09.Feb.2026; published 30.Mar.2026.

Please cite as:
Chen Z, Deng T, Zhu X
Attitudes Toward Digital Meal Assistance Services Among Older Adults in China: Cross-Sectional Survey
JMIR Aging 2026;9:e84956
URL: https://aging.jmir.org/2026/1/e84956
doi:10.2196/84956

XSL•FO
RenderX

© Zhe Chen, Tinghe Deng, Xiaoli Zhu. Originally published in JMIR Aging (https://aging.jmir.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of the Windmill Model for Mapping Older Adults' Intrinsic Capacity Using Digital Twin Technology: Descriptive Qualitative Study

Yirou Niu[1], MSc; Zehui Xuan[1], BSc; Yanling Wang[1], MSc; Xiaoya Zhang[1], BSc; Yibo Guo[1], BSc; Xin Li[1], BSc; Shuai Jin[1], PhD; Jie Zhao[2], MSc; Qian Xiao[1*], PhD; Hong Chang[2*], MSc

[1]School of Nursing, Capital Medical University, No.10 Xi-tou-tiao,You-an-men Wai Street, Fengtai District, Beijing, China

[2]Department of Neurology, Xuanwu Hospital, Capital Medical University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Qian Xiao, PhD
School of Nursing, Capital Medical University, No.10 Xi-tou-tiao,You-an-men Wai Street, Fengtai District, Beijing, China

Abstract

Background: Intrinsic capacity (IC) refers to the sum of the physical and mental capacities of an individual. Conventional IC assessment requires substantial temporal and human resources. Digital twin (DT) technology emerges as a promising solution for efficiently mapping ICs.

Objective: This study aims to explore older adults' perspectives on the DT technology and their perceptions of how it could effectively represent their ICs.

Methods: A qualitative study was used. Face-to-face semistructured interviews with 23 older adults were conducted. The interviews were transcribed verbatim and analyzed via content analysis approach.

Results: The analysis identified five themes and 16 subthemes: (1) "opt for or not my digital twin," revealing the older adults' decisions regarding whether to use DT technology for mapping ICs; (2) "my ideal digital avatar," describing the older adults' preferences for personalized digital avatar appearances; (3) "my digital twin maps my intrinsic capacity," highlighting how multimodal reminders and synchronized avatar changes enhanced their comprehension of ICs; (4) "the benefits my digital twin can deliver," emphasizing the potential of the DT system to provide feedback services to older adults; (5) "some expectations for my digital twin," outlining their expectations for DT technology. Based on the above insights, a conceptual model, "windmill" model, was further developed to better understand how to build DTs of older adults and map their ICs.

Conclusions: DT technology was a promising tool for mapping ICs of older adults. Furthermore, the "windmill" model provided a framework to build tailored DTs. The findings of this study could provide references to develop DT model to support IC management.

(JMIR Aging 2026;9:e81075) doi:10.2196/81075

KEYWORDS

intrinsic capacity; IC; aging; digital twin; DT; avatar; qualitative research

Introduction

As the global population continues to age, promoting healthy aging has become an increasingly urgent public health goal. The meta-analysis indicated that only 22% of the global older adults were expected to achieve healthy aging status by 2023 [1]. To address this issue, the World Health Organization (WHO) developed a health care pathway program—integrated care for older people (ICOPE) [2]. The primary focus of ICOPE was to optimize the trajectory of intrinsic capacity (IC) [3]. IC refers to the sum of the physical and mental capacities of an individual, comprising 5 domains, namely cognitive, locomotive, sensory, vitality, and psychological capacities [4]. The prevention of IC decline plays a vital role in promoting healthy aging. The rate of decline in IC of older adults has been found to reach as significant a proportion as 76.1% [5]. Furthermore, it was hypothesized that changes in IC preceded the process of disability development [3]. Therefore, it is essential for screening and assessing IC and implementing interventions to reverse IC decline, with the goal of fostering healthy aging.

Both the first and second versions of the ICOPE framework emphasized the importance of IC assessment [6,7]. IC assessment was typically conducted by health care providers. Specifically, assessment for IC was recommended to be carried out once every 4 to 6 months [8]. However, this approach was time-consuming and labor-intensive. Therefore, the exploration

XSL•FO
RenderX

of new assessment tools was warranted. Technological advancements suggest that digital devices might assist in screening and assessing IC [3]. For instance, ICOPE Monitor (IHU Health Age), a digital tool, has been applied to ICOPE implementation [9,10]. Additionally, conversational robots have been used to assist in screening the older adults for IC [8]. Despite these efforts, current digital devices remained limited in their ability to automatically transmit IC-related data, presenting challenges for older adults who faced difficulties in operating digital devices. In contrast, digital twin (DT) technology, with its potential for real-time data integration and predictive analytics [11], might offer a more efficient and comprehensive approach to the screening and assessment of IC.

DT refers to the integration of data from physical models, sensors, and operational history into a multidisciplinary, multiphysical, multiscale, and multiprobabilistic simulation process, which maps the entire lifecycle of the corresponding physical entities in virtual space [12]. The DT facilitated bidirectional data flow between physical and virtual entities to continuously optimize physical counterparts and leveraged historical data to acquire knowledge, predict outcomes, simulate behaviors, and provide valuable insights into the future performance of physical systems [13]. DT has been applied across various fields, including architecture, industry, and aerospace [12]. In recent years, DT has also been applied to the field of health care, demonstrating significant potential for the realization of precision medicine [14]. One representative example is human DT, which can dynamically reflect an individual's molecular profile, physiological status, emotional and psychological states, and lifestyle evolution [15]. In addition, the study showed DT models could simulate and replicate the wound healing process to facilitate wound management [16]. Within cardiovascular disease, DT technology seeks to predict disease risk and progression in patients and to help clinicians forecast how pathology will evolve against a particular environmental background [17]. Furthermore, the artificial intelligence–enabled DT technology was a lifestyle medicine intervention that results in significant improvements in hypertension [18]. These applications highlighted the versatility and potential of DT in health care.

The 5D DT model can be used to guide the implementation of DT technology. It was proposed in 2018, incorporating physical entities, the model of physical entities in the virtual space, data, services, and connections [19]. This 5D DT model is a general framework that has been applied across various fields [20]. However, existing studies have mainly focused on technical applications and system construction, with limited exploration of how this model can be tailored to represent the IC of older adults. To bridge this gap and adapt the 5D DT model to a health-related framework, this study proposed a preliminary customization based on the IC context. Accordingly, the preliminary design of each component in the customized 5D DT model was structured as follows. First, the physical entity referred to the older adults and their IC levels. Second, the data dimension covered all data related to the IC of older adults, including historical, real-time, and predictive data, involving the collection, storage, management, and analysis processes. Third, services represented the various functions provided within the DT, illustrating how real-world services could be supported and enhanced through digital means. Fourth, connections described the interaction and communication among different components. Finally, the model of physical entities in the virtual space specifically refers to the digital avatar of an older person, designed to accurately reflect his or her level of IC in this study.

This study represents a step toward developing a customized 5D DT model, with a specific focus on exploring the perspectives of older adults. When designing the customized 5D DT model, it is crucial to ensure that the model not only accurately reflects the assessment results of their ICs but also aligns with the self-perception of older adults. Moreover, the DT model must be designed in a way that is both comprehensible and acceptable to older adults, addressing their specific needs and preferences. To translate these theoretical considerations into a practical model, we conducted semistructured face-to-face interviews to explore older adults' perspectives on the DT concept and their views on how their ICs could be effectively represented within such a model. The findings from these interviews provide the foundational evidence that informs the construction of the customized 5D DT model tailored to map the ICs of older adults.

Methods

Study Design

This study used descriptive qualitative research design, which followed the principles of naturalistic inquiry [21]. Face-to-face semistructured interviews with the older adults were conducted between November 2024 and December 2024.

Study Setting and Participants

Participants were recruited from urban parks and leisure clubs for older adults in Beijing, China, where they engaged in regular recreational and social activities. These sites were chosen because they are popular gathering places for community-dwelling older adults, making them ideal for reaching individuals with diverse levels of ICs. However, we acknowledged that such sampling might introduce a selection bias toward relatively active older adults who are more socially involved. Purposive sampling was used in this study, which enabled researchers to select experienced participants based on the aims of the study [22]. The maximum variation sampling technique was also applied simultaneously to recruit heterogeneous samples of older adults with different levels of ICs, education, sex, and age [23]. The participants had different levels of ICs which were measured by traditional assessment methods (Table S1 in Multimedia Appendix 1), and the educational attainment of participants included primary, secondary, and higher education, to ensure diversity in cognitive and experiential perspectives, offering a wealth of diverse perspectives that could enrich the research topic.

Potential participants were selected according to the following criteria: (1) aged 60 years or older, (2) completed the IC assessment (Table S1 in Multimedia Appendix 1), (3) able to communicate effectively in daily life, and (4) informed consent and voluntary participation in this study.

XSL•FO
RenderX

Data Collection

Older adults were invited to participate in both the assessments of ICs and qualitative interviews. The IC assessment was conducted using a tablet computer, with specific measurement tools and scoring rules detailed in Table S1 in Multimedia Appendix 1. Interviews were scheduled with individuals who agreed to participate after completing the IC assessment. All participants were assured that their participation was entirely voluntary, and written informed consent was obtained prior to data collection.

Before the interview, the interviewer fully introduced the objective of this study to the older adults. The researcher prepared textual and graphic materials to explain the study to the older adults, helping them understand their ICs and DT technology. In addition, a short video introducing DT technology and the concept of a digital avatar was shown to enhance participants' comprehension. Everyday examples and scenario-based demonstrations were used to ensure comprehension, and participants were asked to restate their understanding before the discussion began. The older adults were encouraged to actively share their understanding of the applications of the avatar and DT technology, as well as to voice their perspectives and related knowledge during the discussion. Afterward, the older adults were invited to ask any questions they had about the avatar or DT technology, and the researchers provided clear and comprehensive responses to their inquiries. The participants were also told they could withdraw at any point during the interview. The interviews were conducted in a quiet and relaxed cafe environment that allowed for open and comfortable conversation. The interviews were conducted through Chinese conversations. The participants were encouraged to express themselves based on the interview guide. Audio recordings were made with the consent of the participants.

All of the interviews were conducted by the same researcher (YN) to ensure consistency. The researcher was trained by qualitative research experts and practiced interviews to refine their interviewing skills. The interviews were conducted using a semistructured interview guide (Table S2 in Multimedia Appendix 2), which was initially designed based on self-construal theory, providing a deeper understanding of individuals' self-perception [24]. Then, experts in qualitative research, IC, and DT technology were organized to discuss the interview guide, and the interview guide was revised based on feedback. One participant, who met the inclusion criteria, was selected for a preinterview, and the interview outline was modified based on the results of this preinterview. The data obtained from the preinterview were not included in the analysis. The principle of data saturation was a recognized criterion for determining sample size in qualitative research designs [22]. In this study, data collection and analysis processes were carried out simultaneously, and data collection was stopped when no new data was generated related to the topic of study. Specifically, after interviewing 21 participants, no additional themes related to the study topic were identified, and 2 further interviews were conducted to confirm saturation. Therefore, data collection was stopped once thematic redundancy was observed.

Data Analysis

The interview recordings were transcribed and organized verbatim within 24 hours of the end of each interview. NVivo 12.0 software (QSR International Pty Ltd) was used for coding and analysis, and data analysis was conducted in Chinese. This study adopted the content analysis approach [25]. The interviewer reviewed the text multiple times to ensure comprehensive understanding of its content. Subsequently, open coding was conducted to identify key ideas and concepts within the material. To ensure comprehensive capture of information, open coding was conducted in a line-by-line, sentence-by-sentence format. Then, similar and related codes were classified into categories and subcategories. Furthermore, themes and subthemes were generated based on categories and subcategories. This process was iteratively repeated until saturation was reached, with no new themes or subthemes emerging, and was independently conducted, confirmed, and discussed by 2 researchers. Finally, the themes and subthemes were translated into English using both direct translation and back-translation methods.

Rigor

Rigor focuses on 4 key dimensions: credibility, transferability, dependability, and confirmability [26]. The credibility was strengthened by the researcher's establishment of trusting relationships with older adults, which facilitated authentic self-expression during the interview process. The results of the study were also fed back to the participants, enabling them to confirm and add to the results. Additionally, the researcher used reflexive journaling to track the research process and identify potential biases, thereby enhancing credibility. To enhance transferability, a purposive sampling method was adopted to ensure that participants had rich perspectives. Moreover, rich interview data enhanced transferability by documenting participants' context, enabling evaluation of results' applicability to other settings. The dependability was addressed through regular discussions among the researchers. These discussions allowed for the resolution of disagreements and confusion during the analysis process. Finally, the confirmability was established by providing a description and rationale for the decisions made based on the results.

Ethical Considerations

This study was approved by the ethics committee of the Capital Medical University (batch: Z2023SY138). Informed consent was obtained from all participants, and the study adhered to all relevant privacy and confidentiality protocols. All patient-related information has been fully anonymized and de-identified. No compensation was provided to the participants. All procedures were conducted in accordance with the guidelines of the Institutional Review Board (IRB) and Research Ethics Board (REB).

Results

Participant Characteristics

A total of 23 older adults were interviewed face-to-face, with participant selection considering the age, sex, degree of education, and the level of IC. Interviews lasted between 27.75

XSL•FO
RenderX

and 49.58 minutes (mean 37.57, SD 6.77 min). The demographics were outlined in Table 1. Of the 23 older adults, 12 were women and 11 were men, with an average age of 68.30 (SD 5.45; range: 61-84) years. 17 had a high school education or higher, including 6 with a bachelor's degree or higher. The average score for IC was 4.04 (SD 0.98; range: 2-5).

XSL•FO
RenderX

Table . Participant information.

Number	Sex	Age (years)	Education level	Chronic diseases	Intrinsic capacity score	Cognition	Psychological capacity	Sensory	Vitality	Locomotive capacity
N1	Male	73	College graduate or higher	0	4	29/normal	1/normal	1/hearing failed	14/normal	12/normal
N2	Female	68	Secondary education level	2	4	28/normal	2/normal	2/normal	14/normal	8/decline
N3	Female	68	Secondary education level	1	5	26/normal	0/normal	2/normal	13/normal	11/normal
N4	Male	67	College graduate or higher	4	5	29/normal	3/normal	2/normal	12/normal	11/normal
N5	Male	64	Secondary education level	1	3	24/decline	4/normal	2/normal	13/normal	7/decline
N6	Female	75	College graduate or higher	1	5	28/normal	2/normal	2/normal	13/normal	11/normal
N7	Female	65	Secondary education level	0	5	29/normal	0/normal	2/normal	13/normal	10/normal
N8	Female	69	Secondary education level	1	5	29/normal	0/normal	2/normal	14/normal	10/normal
N9	Male	65	Secondary education level	3	4	28/normal	9/decline	2/normal	14/normal	12/normal
N10	Male	68	College graduate or higher	0	4	22/decline	2/normal	2/normal	14/normal	11/normal
N11	Female	67	Secondary education level	1	4	29/normal	3/normal	1/vision failed	13/normal	11/normal
N12	Male	73	Secondary education level	1	5	26/normal	2/normal	2/normal	14/normal	12/normal
N13	Female	84	Elementary school	0	2	18/decline	3/normal	0/hearing and vision failed	14/normal	9/decline
N14	Male	61	College graduate or above	1	4	28/normal	3/normal	2/normal	13/normal	4/decline
N15	Male	66	Secondary education level	1	5	28/normal	2/normal	2/normal	14/normal	12/normal
N16	Male	68	Elementary school	0	5	30/normal	0/normal	2/normal	14/normal	12/normal
N17	Male	75	Secondary education level	1	2	23/decline	3/normal	0/hearing and vision failed	13/normal	1/decline
N18	Female	67	Secondary education level	0	4	28/normal	8/decline	2/normal	14/normal	12/normal

XSL•FO
RenderX

Number	Sex	Age (years)	Education level	Chronic diseases	Intrinsic capacity score	Cognition	Psychological capacity	Sensory	Vitality	Locomotive capacity
N19	Male	75	Secondary education level	2	2	24/decline	6/decline	0/hearing and vision failed	12/normal	11/normal
N20	Female	61	Secondary education level	3	4	30/normal	4/normal	1/vision failed	13/normal	11/normal
N21	Female	61	Secondary education level	1	4	27/normal	4/normal	1/hearing failed	12/normal	11/normal
N22	Female	66	Secondary education level	2	4	28/normal	2/normal	1/vision failed	14/normal	11/normal
N23	Female	65	College graduate or higher	4	4	28/normal	4/normal	0/hearing and vision failed	14/normal	11/normal

Summary of Themes

Overview

The study identified five key themes related to build DTs of older adults to map their ICs: (1) “opt for or not my digital twin,” explored the willingness of older adults to adopt DTs as a means to reflect their ICs; (2) “my ideal digital avatar,” in which older adults depicted the characteristics of their ideal digital avatars; (3) “my digital twin maps my intrinsic capacity,” focused on how DTs could effectively represent and monitor the ICs of older adults; (4) “the benefits my digital twin can deliver,” in which older adults expressed their beliefs regarding the potential benefits that DTs could offer to them; and (5) “some expectations for my digital twin,” expressed the expectations that older adults had for the future development of DTs. From these insights, we further developed a conceptual model to better understand how to build DTs of older adults to map their ICs.

Theme 1: “Opt for or Not My Digital Twin”

DT technology, being an emerging concept, was unfamiliar to the majority of older adults. When older adults understood the concept of DT, they were presented with the option to use this technology for monitoring and reflecting on their levels of IC. This decision-making process yielded 4 response outcomes:

Subtheme 1: “Health Mirror”—Optional

Some older adults recognized and believed that DTs were optional as health mirrors. They believed that DTs were just like mirrors that could reflect their IC levels and health status. The presentations of DTs were real-time, with changes synchronizing with the IC levels of older adults:

> *If one day digital twins could become a reality and come into my life, that would be quite nice, I reckon.* [N3]

> *Digital twin is my soul, the soul that represents my physical body and health status. The digital twin is perfectly in sync with me, like two hearts beating as one.* [N9]

Subtheme 2: “Multidimensional Advantages”—Worthwhile

Some older adults believed that DTs were well worth choosing because of their multidimensional advantages in terms of convenience, intuitive visibility, scientific rigor, and timely alerts. Moreover, from the perspective of older adults, DTs might be better than physicians. The results generated by DTs were based on real data, which provided a rational basis for accuracy and reliability. However, it was possible for the physicians to selectively disclose certain information to older adults:

> *Digital twins can bring a lot of convenience to my life. It can remind you before you are in danger. It’s very useful, isn’t it? My son is not around me, and my health is poor. I am afraid that one day something bad will happen. No one reminds me. No one is around me. It will be much more convenient to have the digital twins* [N12]

> *Doctors, they have their own emotions and thoughts, and mixing them in with the process (clinical care trajectory) is very bad. Digital twins, they have no emotions. They just do what they are supposed to do, and you know what? They are based on data. Well, they are true.* [N2]

Subtheme 3: “Unknown and Unbelievable”—Nonessential

There remained a group of older adults who believed that constructing a DT to reflect their ICs and health status was of little consequence. Many demonstrated limited familiarity with DT and had no practical experience in its use, leading to skepticism or ambivalence toward its adoption. Additionally, a segment of the older adults perceived DT as being overly futuristic and unattainable, particularly within the near-term timeframe. For technologies perceived as infeasible, the question of choice was deemed irrelevant:

> *I didn’t know anything about digital twins before, and even now, I’ve only heard a little bit about it. I’ve never actually used it, so I can’t really say much*

XSL•FO
RenderX

about it. And as for using it? Well, I wouldn't have a clue how to. [N3]

Well, how should I say it? For someone like me, at this age, even talking about this stuff—digital twin technology—seems like wishful thinking. I don't waste my time on it because this kind of technology is not going to be ready anytime soon. This thing won't be around for at least a decade or two. It has nothing to do with me, so whether I choose it or not doesn't really matter. [N10]

Subtheme 4: "Proliferation Concerns"—Rejected

Not all older adults actively accepted new technologies, with a subset expressing opposition to the adoption of DTs. They argued that advancements in technology might reduce the need for human labor, leading to job losses or unemployment among younger generations. Without stable employment and income, young people might delay or forgo marriage and childbearing, ultimately contributing to a decline in fertility rates. These adverse outcomes caused by the development of technology could affect human survival and reproduction:

I surely don't wish to see such a thing come to pass. If this technology was applied, a few people could take care of all the older adults. What about other young people? How should they survive? Let me tell you this, the more advanced technology becomes, the more people find themselves out of work. But people must have a work and live on. Then they will choose to get married and have children. Why are there so many DINK families coming out? Because technology is too advanced." [N15]

Theme 2: "My Ideal Digital Avatar"

The appearance of the digital avatar served as the end point for the presentation to older adults and directly affected the effectiveness and adherence of the application. According to the concept of DT technology, the digital avatars were replicas of physical entities that synchronized with the changes in real older adults. Unexpectedly, they voiced some perspectives that were outside the box.

Subtheme 5: "Mirror Mapping" and "Visual Refinement"

To better understand their IC and health conditions, some older adults desired the digital avatar to be an exact replica of themselves, reflecting a mirror-image relationship. To avoid negative emotions, such as sadness or self-abasement, that might arise from confronting the realities of aging and poor health, others preferred the digital avatar to embody an idealized, glorified version of themselves. A feasible solution might be to show the real appearance to physicians and a beautified version to the older adults:

I don't want the real one. When you look at yourself against this appearance at home, it's necessary to have an appearance that's more pleasant to look at. You could design two versions—one that shows the real appearance to the doctors, and another that shows a more comforting and pleasant appearance to us. [N1]

Subtheme 6: Self-Preference Projection

The ideal digital avatar for older adults varied significantly based on personal preferences. Some chose representations that mirrored themselves, including their current appearance, an enhanced version, or a rejuvenated younger self. Moreover, some older adults said the digital avatar was portrayed as a young person of the opposite sex, which they found more visually appealing and engaging. Others indicated that the digital avatar represented specific organs, enabling them to better understand health conditions. Interestingly, idols, cute pets, and favorite cartoon characters were also regarded as their ideal digital avatar:

Everyone likes handsome men and beautiful women, don't they? If you're a man, of course you'd want a beautiful woman, especially as you get older, right? And for us old ladies, it's such a delight to see something pleasing to the eye - it just makes us happy! [N8]

Well, if that digital avatar looked like Daoming Chen (an actor), I'd be using it all the time, I tell you! I'd follow every suggestion he makes. [N20]

Theme 3: "My Digital Twin Maps My Intrinsic Capacity"

DT technology can map the ICs of older adults, and the shape and changes of the digital avatar indicate the level and changes in their ICs. The expectations among older adults regarding how DT should reflect the diversity and dynamics of their ICs vary widely.

Subtheme 7: Multimodal Presentation

The manifestation of IC variations in DTs of older adults was multimodal, encompassing auditory, visual, animated, and textual modalities. Some older adults indicated that auditory alerts from the DTs were considered an adequate response to IC variations, dimensional changes, or risk factor detection, and visual representations through images or animations were generally well-received among the older adults. Textual presentations, though beneficial for maintaining the privacy of IC results, were not user-friendly for older adults with vision loss:

My motor function is down. The digital twin of mine is just like the avatar in virtual game—when its energy runs low, it slowly starts to shrink away. It's like a little nudge, reminding me to get moving and recharge my batteries. [N14]

All I need is for it to show me in plain writing—what's the test about, what's the perfect score, how many points I got, what's considered normal range, and where exactly I stand. Just lay it out simple and clear for me. [N18]

Subtheme 8: Matching Behavior Mapping

Older adults tend to monitor external changes that align with the 5 dimensions of IC, using these changes as indicators to

XSL•FO
RenderX

identify potential issues (Table 2). For example, assistive devices indicated declining motor functions, lateral ear orientation suggested hearing loss, and squinting eyes represented visual impairment. There were still some older adults who said they did not need to have extrinsic changes when their IC levels changed:

> *When I can't move like I used to, my digital twin shows it too—starts walking slower, takes smaller steps, kind of shuffling along like my feet are stuck to the floor.* [N23]

> *Even if my IC decreased, I'd still want my digital twin to look happy and cheerful. You see, when the people around me are smiling and in good spirits, it makes me feel better too. But if they look sick and miserable, it just brings me down.* [N11]

Table . Digital avatar's external changes linked to intrinsic capacity abnormalities.

Dimension and abnormal outcomes	External changes
Cognition	
Cognitive decline	• Dull eyes • Forgetting behavior • Speech delay • Look unsure • Dementia-like behavior • Fixed posture • Unresponsive
Psychological capacity	
Mild depressive symptoms	• Less communication • Moodiness
Severe depressive symptoms	• Expression of fear • Expression of irritability • Expression of anger • Expression of giving up • Abnormal behavior • Social avoidance
Sensory	
Vision failed	• Wearing glasses • Squinting
Hearing failed	• Wearing hearing aids • Lean in to listen • Unresponsive
Vitality	
Malnutrition or risk of malnutrition	• Slim • Pale complexion • Listless • Increased wrinkles • Activity reduction
Locomotive capacity	
Decline	• With the aid of assistive devices • Stride reduction • Flustered gait

Theme 4: "The Benefits My Digital Twin Can Deliver"

From the perspective of older adults, DT offers the following potential benefits.

Subtheme 9: Health Advisor

Some older adults regarded DTs as personalized health advisors, capable of enhancing their IC through proactive interventions. Moreover, DTs could perform comprehensive health monitoring, provide early disease detection, support medical treatment, and facilitate health management:

> *Every morning when I open my phone and see my digital twin, it's like getting a little preview of how my day might go. It tells me what I should watch out for, if something's not quite right with my health, giving me a heads-up before any real trouble starts.*

XSL•FO
RenderX

It's like having a little guardian angel in my pocket. [N21]

Subtheme 10: Life Assistant

Beyond the medical applications, DT was also recognized as a comprehensive life assistant, offering support across various aspects of daily living, such as essential reminders, practical life guidance, and tailored solutions to address the unique challenges faced by seniors living independently:

In the future, if I am left alone, I will get such a digital twin. It'd be a real help. Like when I feel like cooking, it could show me recipes. And when I want to do my exercises, it could lead me through them. [N7]

Subtheme 11: Spiritual Mentor

Some older adults not only found joy in engaging with their DTs but also saw them as spiritual mentors in favor of their psychological well-being:

Sometimes we old folks aren't too keen on socializing with others. Maybe we don't quite get along with real people, but chatting with our own digital twin gives us something to do. What's more, the mental state of my digital twin is exceptionally good, and I consider it to be my role model. [N1]

Subtheme 12: Role Complement

DT was poised to become an important supplement to the lives of older adults, such as their children, caregivers, or health advisors. For instance, for a group of older adults experiencing geographical separation from their children or those belonging to childless households, DT offered comprehensive solutions by providing both practical assisted living services and essential emotional companionship, thereby effectively supplementing traditional filial support roles:

My children can't really take care of me, bless their hearts - they've got their own families to look after, and they live so far away. But I've got this digital twin of mine, it's been a real lifesaver. Whether it's helping me in life or giving me health guidance, it's always there for me. Honestly, it's just like having one of my own children by my side. [N9]

Theme 5: "Some Expectations for My Digital Twin"

Older adults expressed some expectations for their DTs, notably the following subthemes.

Subtheme 13: Intelligence and Knowledge Integration

The DT technology used to reflect the IC of older adults was expected to be intelligent and knowledge-integrated, enabling automated data transfer, rich functionality, and remote operation. The intelligence and knowledge integration of the DT technology were not only applied in health care, but also extended to various facets of older adults' daily life, encompassing daily assistance, entertainment, and so on:

All my information and data just get sent up there automatically—I don't have to bother with typing anything in myself. [N10]

My digital avatar can even dance and sing for me. When I'm in the mood for some Beijing opera, it can belt out those classic tunes too. It's got all sorts of little entertainment tricks up its sleeve. [N13]

Subtheme 14: Positive Feedback

Timely positive feedback could help motivate older adults to persist in using DT technology to improve their ICs and health status:

Every day, my digital avatar gives me exercise guidance. After the workout, it shows me all my vital signs right then and there, making me feel like I'm getting better. Getting this instant feedback really keeps me motivated to stay active and healthy. [N16]

Subtheme 15: Multisubjective Bond

DTs for older adults were anticipated to function as interconnected systems with multiple entities, serving as bonds that facilitated integration across domestic, community, and hospital networks:

If the doctor notices my digital avatar 's head turning red—that means my blood pressure's up—they can just send me a message right away, telling me what to do about it. [N6]

Let me give you an example—my digital avatar here, it's not just saw by me, but my spouse too. My son can check in on me anytime, through this system. [N14]

Subtheme 16: Privacy and Security

The assessment and management of older adults' IC using DT technology necessitated the huge sensitive personal data. This process should be private and secure:

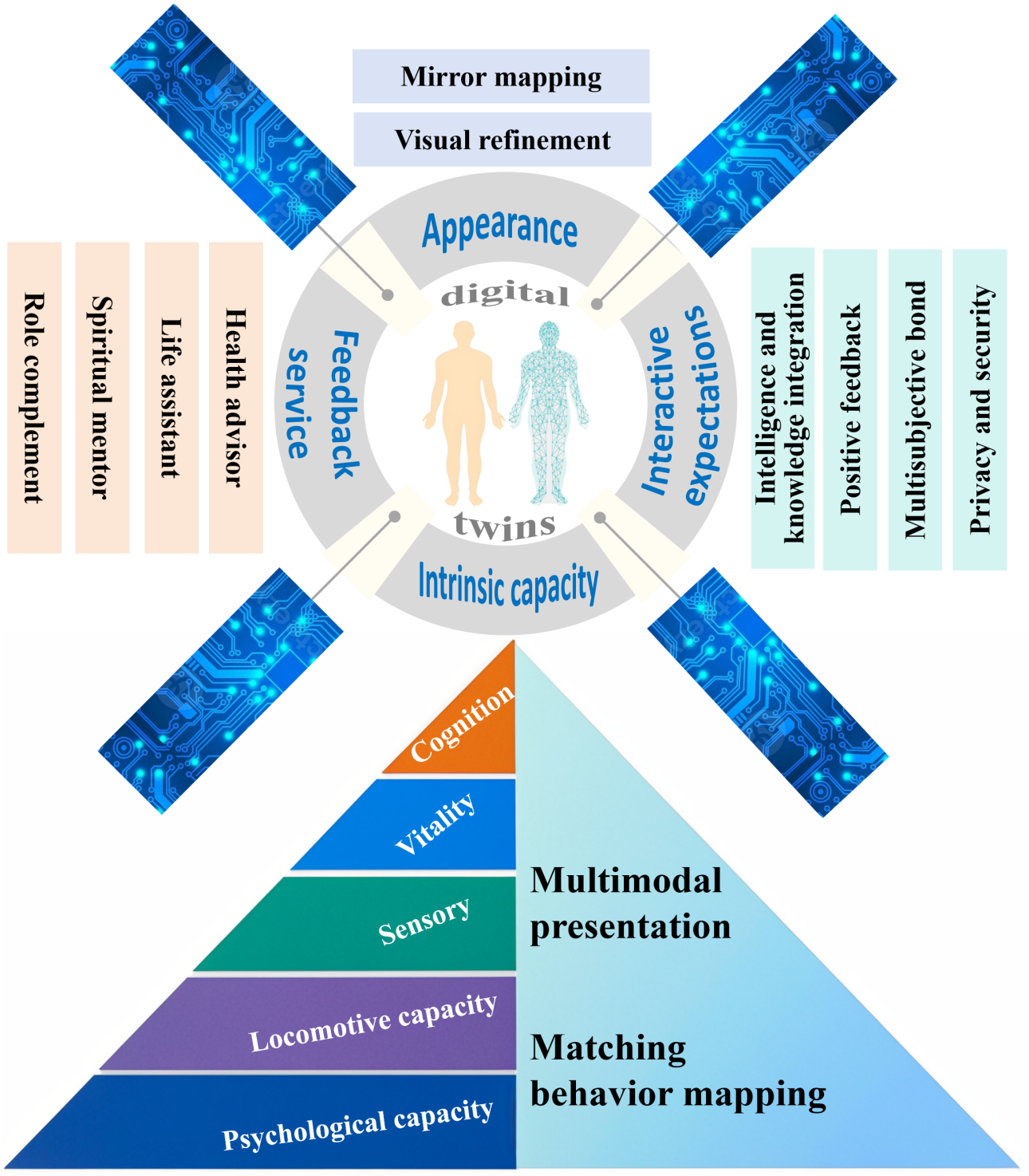
The digital twin should have proper security measures. If they go leaking all our private information, we seniors would lose trust in them. [N12]

Conceptual Model

Based on the insights, we further developed a conceptual model, the "windmill" model, to better understand how to build DTs of older adults to map their ICs (Figure 1). Centered on the evaluation of IC, the "windmill" model integrated DT technology and comprised 4 key components. First, the appearance of the DT, which reflects 2 main preferences: "mirror-like" and "idealized" forms. Second, the mechanism of mapping IC, which enables dynamic visualization of older adults' IC levels and changes through multimodal presentation and behavioral mapping. Third, feedback service, which includes 4 functional roles: health advisor, life assistant, spiritual mentor, and role supplement. Fourth, interactive expectations, which emphasize intelligence and knowledge integration, the provision of positive feedback, effective coordination among families, communities, and health care institutions, and ensuring the security of personal privacy information. Grounded in the perspective of older adults, this "windmill" model provided support for the application of DT technology in IC assessment.

XSL•FO
RenderX

Figure 1. Digital twins focusing on intrinsic capacity of older adults—windmill model.



Discussion

Principal Findings

The integration of DT technology into IC assessment held the potential to transform how older adults' IC was monitored and supported. Incorporating the perspectives of older adults, this study offered an initial exploration of how DT technology might be leveraged to support IC assessment. As for avatar appearance, the participants have demonstrated diverse perspectives regarding their envisioned characteristics of an ideal DT. They also described how the DT should represent and map their IC levels, highlighting the importance of multimodal reminders for enhancing IC comprehension. In addition, participants expected the DT to offer feedback services that could help maintain or improve IC. Based on these findings, a preliminary framework—the "windmill" model—was developed, incorporating 4 core components: appearance, intrinsic capacity, feedback services, and interactive expectations. The "windmill" model offered potential pathways for guiding the design of DT-supported IC assessment tools for older adults.

The appearance of digital avatars played a pivotal role in user engagement and acceptance, which was conducive to helping

XSL•FO
RenderX

older adults understand and maintain their ICs. The findings of this study emphasized the importance of personalized digital avatars, consistent with existing literature that indicated superior self-identity recognition with personalized digital representations versus generic avatars [27]. The digital avatars might be one-to-one reductions of physical entities. Individuals might demonstrate an inherent desire to protect avatars that mirrored themselves [28,29], which was beneficial for them to maintain their internal abilities. Digital avatars representing older adults were also expected to take other forms, such as younger versions of themselves, specific organs, or even their idols. It has been shown that whether or not digital avatars resemble the individuals themselves, both can have an impact [30]. Effective interaction, rather than mere personification, might be the key to the success of digital avatars [31]. Therefore, the design of digital avatars should be tailored to the unique characteristics and preferences of older adults to enable effective interactions, rather than strictly replicating their physical appearance.

DT technology realized the dynamic mapping of the ICs of older adults. DT technology could effectively reflect ICs in the following ways: (1) current IC levels and trends, presented through multimodal methods such as voice, text, images, and animations, which cater to diverse user preferences and cognitive abilities; and (2) synchronous changes between digital avatars and the older adults they represent, which were more easily understood. By bridging the gap between complex health data and actionable information, DT technology had the potential to empower older adults to better understand and manage their ICs, which could compensate for the health inequalities caused by the digital divide [32].

The digital avatar had a dual identity—"from avatars to agents" [27]. In this study, digital avatars functioned not only as visual representations to display the ICs of older adults but also as interactive agents capable of providing feedback services. These feedback services focused on health support, life support, and psychological and social support. In terms of health support, digital avatars could provide health status monitoring, early warning, and health education [14]. Additionally, when interacting with digital avatars representing themselves, individuals could adopt an interventionist perspective, leading to a more objective self-assessment [33]. Digital avatars were also expected to provide assistance in daily life for older adults, thereby increasing user engagement and stickiness. Regarding emotional and social support, older adults could identify with their digital avatars, experience meaningful emotions, and form authentic attachments with them. Digital avatars created a beneficial psychological distance that facilitated more objective self-reflection while maintaining emotional engagement [34]. Compared to interactions with real people, individuals perceived digital avatars as more private and secure, making them more willing to express themselves [34,35]. Avatar therapy has been used for the treatment of psychosocial diseases [35-37]. In summary, digital avatars were expected to serve as multifunctional tools.

Older adults expressed some expectations for DT technology in mapping their ICs. First, the process should be intelligent and equipped with sufficient knowledge. In addition to the foundational technologies such as sensors, Internet of Things, machine learning, and deep learning, the intelligence of DT can be further advanced through the integration of multisensor fusion, edge-cloud computing synergy, virtual reality or augmented reality interactive technologies, and generative adversarial networks for data augmentation [38,39]. In the future, DT could be further enhanced by integrating knowledge graphs [40] and expert knowledge [11] to ensure sufficient knowledge support. Second, older adults need positive feedback from DT. Positive feedback has been proven to improve self-efficacy and intrinsic motivation [41]. Third, the DT technology, as a multisubjective bond, connected the older adults, families, communities, and hospitals. From the perspective of older adults, remote health care services become more accessible. For families, it provides a clearer and more comprehensive understanding of older adults' well-being. Meanwhile, communities and hospitals can achieve batch management of older adults. Finally, the privacy and security of DT technology can be guaranteed by blockchain technology, federated learning, and other methods [11].

Overall, the "windmill" model provided insights from older adults on the construction of the virtual entity, the mapping of IC, the mechanisms of interaction between the virtual entity and his or her twin, and the types of services that a DT system should deliver. It also encompasses the core components of a DT system—the physical entity, its virtual twin, and the mapping relationship linking the two [42]. Compared with the 5D DT model, the "windmill" model offered a distinct theoretical contribution by integrating older adults' perspectives into the conceptualization of DT-based IC assessment. This user-centered orientation was absent from the 5D DT model, which primarily emphasized system architecture, data flows, and computational mechanisms but offered limited consideration of older adults' acceptance, perceptions, and needs. However, the "windmill" model presented a notable limitation in that it did not address the data and connectivity dimensions included in the 5D DT model. To enhance its comprehensiveness, future work should incorporate expert consultations to refine these technical dimensions and integrate them with the older adults-centered components.

The IC assessment based on DT technology guided by the "windmill" model showed considerable promise. The IC assessment is grounded in multidimensional data collected from older adults, such as physiological parameters, handgrip strength, and daily activity [43-46]. By integrating general population-level knowledge with real-time individual data from wearable devices, DT technology enabled personalized and continuously refined IC assessments for older adults [47,48]. The entire process of data collection and IC assessment results presentation through DT technology reflects the ongoing transition in modern medicine towards a data-driven paradigm [49]. Beyond assessment, DT technology informed by the "windmill" model also showed strong potential for IC-oriented interventions. The study demonstrated that personalized digital coaching presents a highly promising approach for interventions targeting IC [50]. Virtual entity could also be seen as a form of coaching. The study had shown that a virtual entity equipped with expert-level medical knowledge could serve as a personal physician capable of answering patients' questions around the

XSL•FO
RenderX

clock [51]. In addition, an older adult's DT could serve as a platform for computer-simulated therapeutic experiments, enabling the design of precisely tailored and personalized treatment plans [51].

However, the application of DT to the IC assessment in older adults still faces several major challenges, such as high implementation costs [52], limited usability for individuals with low digital literacy [53], and ethical concerns regarding data privacy and consent [54]. To overcome these limitations, future DT systems are expected to incorporate generative artificial intelligence–based interfaces to simplify user interaction and adopt low-cost wearable sensors and privacy-preserving data protocols to mitigate financial and ethical constraints [51,55,56]. These technological advances also provide a foundation for practical implementation in real-world settings. Take the Chinese community setting as an example, older adults could wear basic wearable sensors, and the collected data would automatically update their individual DTs. Based on these real-time data streams, the DT model could provide targeted guidance for daily IC management, while community health professionals could remotely assess and manage older adults' IC through the DT system.

There were several limitations to this study. First, this qualitative study sample involved participants from parks and leisure clubs for older adults but did not include older adults from homes, nursing homes, or hospitals. This limited the generalizability of our findings to the broader population of older adults. Second, all interviewees were from urban areas and had some level of familiarity with DT technology. This might have influenced their perceptions and suggests that the findings might not fully represent older adults living in rural areas or those with limited exposure to DT technology [57]. In the future, a more diverse sample should be considered, including participants from different regions and with varying levels of technological experience, to improve representativeness and capture a wider range of perspectives.

Conclusions

To integrate DT technology for mapping the ICs of older adults and enable comprehensive management of ICs, a preliminary conceptual model, termed the "windmill" model, was proposed. According to the "windmill" model, the ideal digital avatars of older adults had diverse personalized characteristics. To help older adults understand their ICs, multimodal reminders and synchronous changes of digital avatars were essential. Furthermore, the DT system was believed to provide various feedback services to older adults in terms of health, daily life, psychology, and society. Older adults also had some expectations for DT technology, such as intelligence, knowledge integration, and positive feedback. These findings provided guidance for the development of DT systems tailored to effectively support IC mapping in older adults, ultimately contributing to healthy aging.

Acknowledgments

The authors would like to express their deepest appreciation to all the older adults who participated in this study for their valuable involvement.

Funding

This study was supported by a grant from the National Key R&D Program of China (2023YFC3605200 and 2023YFC3605201).

Data Availability

The data are not publicly available due to privacy concerns and confidentiality agreements. For inquiries, please contact the corresponding author. This study was not preregistered.

Authors' Contributions

Conceptualization: YN, ZX, QX, HC
Data curation: YN, ZX
Formal analysis: YN, YW, QX, HC
Funding acquisition: QX, HC
Investigation: YN, XZ, YG, XL
Methodology: YN, YW, JZ, QX, HC
Project administration: YN
Supervision: YW, QX, HC
Visualization: YN
Writing – original draft: YN
Writing – review and editing: QX, HC
HC is co-corresponding author (email: changhong19791111@126.com, phone number:86 13701034090).

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
The assessment tools of intrinsic capacity.
[DOCX File, 13 KB - aging_v9i1e81075_app1.docx]

Multimedia Appendix 2
Semi-structured interview guide summary.
[DOCX File, 17 KB - aging_v9i1e81075_app2.docx]

References

1. Liu Y, Zhang X, Fu X, Li W, Wang Y, Lu W. Successful aging rates of global older adults: a systematic review and meta-analysis. J Am Med Dir Assoc 2025 Jun;26(6):105334. [doi: 10.1016/j.jamda.2024.105334] [Medline: 39521019]
2. Integrated care for older people: guidelines on community-level interventions to manage declines in intrinsic capacity. : World Health Organization; 2017 URL: https://iris.who.int/server/api/core/bitstreams/4639a0b9-f77e-48db-9617-987aabb143f2/content [accessed 2026-03-06]
3. Chhetri JK, Harwood RH, Ma L, Michel JP, Chan P. Intrinsic capacity and healthy ageing. Age Ageing 2022 Nov 2;51(11):afac239. [doi: 10.1093/ageing/afac239] [Medline: 36346737]
4. Zhou Y, Ma L. Intrinsic capacity in older adults: recent advances. Aging Dis 2022 Apr;13(2):353-359. [doi: 10.14336/AD.2021.0818] [Medline: 35371613]
5. Liu Y, Du Q, Jiang Y. Detection rate of decreased intrinsic capacity of older adults: a systematic review and meta-analysis. Aging Clin Exp Res 2023 Oct;35(10):2009-2017. [doi: 10.1007/s40520-023-02515-7] [Medline: 37543528]
6. Yan Wang N, Liu X, Kong X, et al. Implementation and impact of the World Health Organization integrated care for older people (ICOPE) program in China: a randomised controlled trial. Age Ageing 2024 Jan 2;53(1):afad249. [doi: 10.1093/ageing/afad249] [Medline: 38251736]
7. Integrated care for older people (ICOPE) guidance for person-centred assessment and pathways in primary care. : World Health Organization; 2024 URL: https://iris.who.int/server/api/core/bitstreams/7516b015-e205-43d1-883f-01f2c31af261/content [accessed 2026-03-06]
8. Tavassoli N, Piau A, Berbon C, et al. Framework implementation of the INSPIRE ICOPE-CARE program in collaboration with the World Health Organization (WHO) in the Occitania Region. J Frailty Aging 2021;10(2):103-109. [doi: 10.14283/jfa.2020.26] [Medline: 33575698]
9. Berbon C, Takeda C, Balardy L, et al. Implementing the WHO ICOPE program in clinical practice: three years of lessons from monitoring 27 082 participants using the ICOPE monitor digital tool. J Gerontol A Biol Sci Med Sci 2025 Feb 10;80(3):glae278. [doi: 10.1093/gerona/glae278] [Medline: 39549299]
10. Sanchez-Rodriguez D, Piccard S, Dardenne N, Giet D, Annweiler C, Gillain S. Implementation of the integrated care of older people (ICOPE) app and ICOPE monitor in primary care: a study protocol. J Frailty Aging 2021 Jul;10(3):290-296. [doi: 10.14283/jfa.2021.22] [Medline: 34105715]
11. Wu H, Ji P, Ma H, Xing L. A comprehensive review of digital twin from the perspective of total process: data, models, networks and applications. Sensors (Basel) 2023;23(19):8306. [doi: 10.3390/s23198306] [Medline: 37837136]
12. Mulder ST, Omidvari AH, Rueten-Budde AJ, et al. Dynamic digital twin: diagnosis, treatment, prediction, and prevention of disease during the life course. J Med Internet Res 2022 Sep 14;24(9):e35675. [doi: 10.2196/35675] [Medline: 36103220]
13. Dihan MS, Akash AI, Tasneem Z, et al. Digital twin: data exploration, architecture, implementation and future. Heliyon 2024 Mar 15;10(5):e26503. [doi: 10.1016/j.heliyon.2024.e26503] [Medline: 38444502]
14. Shen MD, Chen SB, Ding XD. The effectiveness of digital twins in promoting precision health across the entire population: a systematic review. NPJ Digit Med 2024 Jun 3;7(1):145. [doi: 10.1038/s41746-024-01146-0] [Medline: 38831093]
15. Chen J, Yi C, Okegbile SD, Cai J, Shen X. Networking architecture and key supporting technologies for human digital twin in personalized healthcare: a comprehensive survey. IEEE Commun Surv Tutorials 2023;26(1):706-746. [doi: 10.1109/COMST.2023.3308717]
16. Sarp S, Kuzlu M, Zhao Y, Gueler O. Digital twin in healthcare: a study for chronic wound management. IEEE J Biomed Health Inform 2023 Nov;27(11):5634-5643. [doi: 10.1109/JBHI.2023.3299028] [Medline: 37549083]
17. Coorey G, Figtree GA, Fletcher DF, Redfern J. The health digital twin: advancing precision cardiovascular medicine. Nat Rev Cardiol 2021 Dec;18(12):803-804. [doi: 10.1038/s41569-021-00630-4] [Medline: 34642446]
18. Shamanna P, Joshi S, Dharmalingam M, et al. Digital twin in managing hypertension among people with type 2 diabetes: 1-year randomized controlled trial. JACC Adv 2024 Sep;3(9):101172. [doi: 10.1016/j.jacadv.2024.101172] [Medline: 39372467]
19. Inamdar A, van Driel WD, Zhang G. Digital twin technology—a review and its application model for prognostics and health management of microelectronics. Electronics (Basel) 2024;13(16):3255. [doi: 10.3390/electronics13163255]
20. Tao F, Liu W, Zhang M. Five-dimension digital twin model and its ten applications. Comput Integr Manuf Syst 2019;25(1):1-18. [doi: 10.13196/j.cims.2019.01.001]

XSL•FO
RenderX

21. Colorafi KJ, Evans B. Qualitative descriptive methods in health science research. HERD 2016 Jul;9(4):16-25. [doi: 10.1177/1937586715614171] [Medline: 26791375]
22. Bradshaw C, Atkinson S, Doody O. Employing a qualitative description approach in health care research. Glob Qual Nurs Res 2017;4:2333393617742282. [doi: 10.1177/2333393617742282] [Medline: 29204457]
23. Holloway I, Galvin K. Qualitative Research in Nursing and Healthcare, 5th edition: John Wiley & Sons; 2023. URL: https://books.google.co.in/books/about/Qualitative_Research_in_Nursing_and_Heal.html?id=8ZHPEAAAQBAJ&redir_esc=y [accessed 2026-03-28]
24. Voyer BG, Franks B. Toward a better understanding of self-construal theory: an agency view of the processes of self-construal. Rev Gen Psychol 2014 Jun;18(2):101-114. [doi: 10.1037/gpr0000003]
25. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
26. Guba EG, Lincoln YS. Fourth Generation Evaluation: Sage Publications; 1989. URL: https://uk.sagepub.com/en-gb/eur/fourth-generation-evaluation/book2748 [accessed 2026-03-06]
27. Fiedler ML, Wolf E, Döllinger N, et al. From avatars to agents: Self-related cues through embodiment and personalization affect body perception in virtual reality. IEEE Trans Visual Comput Graphics 2024 Nov;30(11):7386-7396. [doi: 10.1109/TVCG.2024.3456211]
28. Baysden E, Mendoza N, Callender C, Deng Z, Thompson D. Teen reactions to a self-representational avatar: a qualitative exploration. J Sport Health Sci 2022 Mar;11(2):157-163. [doi: 10.1016/j.jshs.2021.07.004] [Medline: 34293497]
29. Xie P, Huang C, Wang X, Sang H, Zhou A. The effect of avatar identity on spontaneous perspective-taking in patients with schizophrenia. Schizophr 2025;11(1). [doi: 10.1038/s41537-024-00551-4]
30. Wang X, Liu Q, Li X, Zheng X, Hommel B, Ma K. Sadness is reduced by virtual hugging. Appl Psychol Health Well Being 2025 Feb;17(1):e70007. [doi: 10.1111/aphw.70007] [Medline: 39934083]
31. Carter OBJ, Loft S, Visser TAW. Meaningful communication but not superficial anthropomorphism facilitates human-automation trust calibration: the human-automation trust expectation model (HATEM). Hum Factors 2024 Nov;66(11):2485-2502. [doi: 10.1177/00187208231218156] [Medline: 38041565]
32. Wu M, Xue Y, Ma C. The association between the digital divide and health inequalities among older adults in China: nationally representative cross-sectional survey. J Med Internet Res 2025 Jan 15;27:e62645. [doi: 10.2196/62645] [Medline: 39813666]
33. Neyret S, Bellido Rivas AI, Navarro X, Slater M. Which body would you like to have? The impact of embodied perspective on body perception and body evaluation in immersive virtual reality. Front Robot AI 2020;7:31. [doi: 10.3389/frobt.2020.00031] [Medline: 33501199]
34. Jang B, Yuh C, Lee H, et al. Exploring user experience and the therapeutic relationship of short-term avatar-based psychotherapy: qualitative pilot study. JMIR Hum Factors 2025 Feb 5;12:e66158. [doi: 10.2196/66158] [Medline: 39924261]
35. Bryan EG, Chen H, Vilaro M, et al. Developing a supportive virtual human to deliver clinical trial education for older women and other populations historically excluded from research. Patient Educ Couns 2025 Jan;130:108485. [doi: 10.1016/j.pec.2024.108485] [Medline: 39476460]
36. Garety PA, Edwards CJ, Jafari H, et al. Digital AVATAR therapy for distressing voices in psychosis: the phase 2/3 AVATAR2 trial. Nat Med 2024 Dec;30(12):3658-3668. [doi: 10.1038/s41591-024-03252-8] [Medline: 39468363]
37. Karsan S, Diaz-Fong JP, Ly R, Hellemann G, Feusner JD. Quantifying body size estimation accuracy and body dissatisfaction in body dysmorphic disorder using a digital avatar. Compr Psychiatry 2024 Nov;135:152529. [doi: 10.1016/j.comppsych.2024.152529] [Medline: 39241374]
38. Lv Z. Digital twins in industry 5.0. Research (Lambertville) 2023;6:0071. [doi: 10.34133/research.0071] [Medline: 36930777]
39. Wu K, Chen J, Chen L, et al. QoE-aware joint visual and haptic signal transmission with adaptive data compression for immersive interactions in human digital twin. IEEE Trans Netw Serv Manage 2025;22(3):2780-2794. [doi: 10.1109/TNSM.2025.3558358]
40. Stavropoulou G, Tsitseklis K, Mavraidi L, et al. Digital twin meets knowledge graph for intelligent manufacturing processes. Sensors (Basel) 2024 Apr 19;24(8):2618. [doi: 10.3390/s24082618] [Medline: 38676238]
41. Plunkett A. Embracing excellence in healthcare: the role of positive feedback. Arch Dis Child Educ Pract Ed 2022 Oct;107(5):351-354. [doi: 10.1136/archdischild-2020-320882] [Medline: 34426538]
42. Okegbile SD, Cai J, Zheng H, Chen J, Yi C. Differentially private federated multi-task learning framework for enhancing human-to-virtual connectivity in human digital twin. IEEE J Select Areas Commun 2023;41(11):3533-3547. [doi: 10.1109/JSAC.2023.3310106]
43. Guo W, Meng L, Han J, et al. Intrinsic capacity and its association with predictors among Chinese empty nest older adults in communities: a latent class analysis. BMC Geriatr 2024 Dec 19;24(1):1008. [doi: 10.1186/s12877-024-05583-0] [Medline: 39702076]
44. Liu Q, Li X, Hu M, Zhao Y, Wu S, Feng H. Factors influencing the self-management ability among older adults experiencing intrinsic capacity decline: a cross-sectional study. Front Aging Neurosci 2024;16:1456167. [doi: 10.3389/fnagi.2024.1456167] [Medline: 39660337]

XSL•FO
RenderX

45. Chen XL, Li J, Sun SN, et al. Association between daily internet use and intrinsic capacity among middle-aged and older adults in China: large prospective cohort study. J Med Internet Res 2024 Nov 12;26:e54200. [doi: 10.2196/54200] [Medline: 39531641]
46. Xuekelati S, Maimaitiwusiman Z, Xiang H, Wumaer A, Bai X, Wang H. Handgrip strength: a simple and effective tool to recognize decreased intrinsic capacity in Chinese older adults. Exp Gerontol 2024 Oct 15;196:112567. [doi: 10.1016/j.exger.2024.112567] [Medline: 39236871]
47. Yang Y, Shi Y, Yi C, et al. Dynamic human digital twin deployment at the edge for task execution: a two-timescale accuracy-aware online optimization. IEEE Trans on Mobile Comput 2024;23(12):12262-12279. [doi: 10.1109/TMC.2024.3406607]
48. Pellegrino G, Gervasi M, Angelelli M, Corallo A. A conceptual framework for digital twin in healthcare: evidence from a systematic meta-review. Inf Syst Front 2025 Feb;27:7-32. [doi: 10.1007/s10796-024-10536-4]
49. Shomorony I, Cirulli ET, Huang L, et al. An unsupervised learning approach to identify novel signatures of health and disease from multimodal data. Genome Med 2020 Jan 10;12(1):7. [doi: 10.1186/s13073-019-0705-z] [Medline: 31924279]
50. Stara V, Soraci L, Takano E, et al. Intrinsic capacity and active and healthy aging domains supported by personalized digital coaching: survey study among geriatricians in Europe and Japan on eHealth opportunities for older adults. J Med Internet Res 2023 Oct 12;25:e41035. [doi: 10.2196/41035] [Medline: 37824183]
51. Chen J, Shi Y, Yi C, Du H, Kang J, Niyato D. Generative-AI–driven human digital twin in IoT healthcare: a comprehensive survey. IEEE Internet Things J 2024;11(21):34749-34773. [doi: 10.1109/JIOT.2024.3421918]
52. Jaana M, Lévesque Ryan M, Tamim H, Riachy E, Paré G. Perspectives of older adults on assistive technology: qualitative study. JMIR Hum Factors 2025 Nov 12;12:e74214. [doi: 10.2196/74214] [Medline: 41222704]
53. Pani J, Fiorini L, Rovini E, et al. Identifying and evaluating factors influencing technology adoption: a multi-pilot study within the Pharaon project. Int J Med Inform 2026 Feb;206:106179. [doi: 10.1016/j.ijmedinf.2025.106179] [Medline: 41237513]
54. Shen S, Qi W, Liu X, et al. From virtual to reality: innovative practices of digital twins in tumor therapy. J Transl Med 2025 Mar 19;23(1):348. [doi: 10.1186/s12967-025-06371-z] [Medline: 40108714]
55. Chen J, Li Y, Yi C, Gong S. Generative AI-aided qoe maximization for RIS-assisted digital twin interaction. Presented at: IEEE INFOCOM 2025 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS); May 19-22, 2025. [doi: 10.1109/INFOCOMWKSHPS65812.2025.11152999]
56. Chen J, Yi C, Du H, et al. A revolution of personalized healthcare: enabling human digital twin with mobile AIGC. IEEE Netw 2024;38(6):234-242. [doi: 10.1109/MNET.2024.3366560]
57. Nguyen NS, Ngo QD, Phan TMC, Tao VA, Bui TNL. Understanding mHealth adoption among older adults: a novel NCA-fsQCA analysis. J Health Organ Manag 2025 Nov 10;39(8):1565-1579. [doi: 10.1108/JHOM-01-2025-0053] [Medline: 41201198]

Abbreviations

DT: digital twin
IC: intrinsic capacity
ICOPE: integrated care for older people
WHO: World Health Organization

Edited by Y Jiang; submitted 22.Jul.2025; peer-reviewed by E Maranesi, J Chen; revised version received 03.Dec.2025; accepted 01.Feb.2026; published 13.Apr.2026.

Please cite as:
Niu Y, Xuan Z, Wang Y, Zhang X, Guo Y, Li X, Jin S, Zhao J, Xiao Q, Chang H
Development of the Windmill Model for Mapping Older Adults' Intrinsic Capacity Using Digital Twin Technology: Descriptive Qualitative Study
JMIR Aging 2026;9:e81075
URL: https://aging.jmir.org/2026/1/e81075
doi:10.2196/81075

© Yirou Niu, Zehui Xuan, Yanling Wang, Xiaoya Zhang, Yibo Guo, Xin Li, Shuai Jin, Jie Zhao, Qian Xiao, Hong Chang. Originally published in JMIR Aging (https://aging.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited.

XSL•FO
RenderX

The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Segmentation of Older People's Needs and Readiness for Smart Homes by Residentially Based Lifestyles in Spain: Survey Study

Jiyeon Yu*, PhD; Angélica de Antonio*, PhD; Elena Villalba-Mora*, PhD

Centre for Biomedical Technology, Universidad Politécnica de Madrid, Parque Científico y Tecnológico de la UPM, Campus de Montegancedo, Crta. M40, Km. 38, Pozuelo de Alarcón, Madrid, Spain

*all authors contributed equally

Corresponding Author:
Elena Villalba-Mora, PhD
Centre for Biomedical Technology, Universidad Politécnica de Madrid, Parque Científico y Tecnológico de la UPM, Campus de Montegancedo, Crta. M40, Km. 38, Pozuelo de Alarcón, Madrid, Spain

Abstract

Background: Globally, the older population is increasing rapidly, becoming one of the most significant demographic trends of the 21st century. This growth poses important social, health, and technological challenges for societies that must adapt their environments and services to promote independent and healthy aging. In Spain, the population aged 65 years and older reached 18% of the total population in 2020, and projections indicate that this proportion will continue to rise in the coming decades. Within this context, smart homes have emerged as one of the most promising avenues to support aging in place and improve the quality of life. Smart homes encompass a wide variety of functions, including environmental control, safety monitoring, communication, and other assistive technologies, that may help older people stay healthy, safe, and independent in their own homes. However, older people are not a homogeneous group. Their lifestyles, health conditions, and technological experiences differ substantially, which means that, as with any assistive technology, smart home functions must match the real and perceived needs of the target users to ensure acceptance, adoption, and long-term use.

Objective: In this study, as a step forward toward the adaptability of smart home technology, we present a method to analyze the practical needs of smart home functions for older people. Specifically, we aim to understand the Spanish older population's readiness and needs for smart homes and to provide insights that can guide the design of more adaptive and user-centered solutions.

Methods: We conducted an online survey focusing on residentially based lifestyles, health conditions, and preferences for smart home functions, targeting older adults living in Spain. The survey collected information about participants' demographic profiles, daily activities, health self-assessment, and attitudes toward technology. A total of 102 valid responses were analyzed. We then classified the older adults according to their residentially based lifestyles using clustering techniques and analyzed the preferences and needs for smart home functions in each identified group.

Results: Four clusters emerged based on the information provided by the participants: (1) high quality of life and independent life, (2) poor quality of life, (3) social-centered life, and (4) creative and personal-centered hobbies at home. On the basis of this classification, we explored each group's specific needs for smart homes and estimated their readiness to embrace different aspects of technology. As a result, the top-priority smart home functions for each group were identified and compared.

Conclusions: This research contributes to understanding the practical user needs of smart homes as assistive technologies for older people. It provides a methodological approach to anticipate and prioritize functions according to user characteristics, supporting the development of personalized, adaptive, and more acceptable smart home solutions for aging populations.

(JMIR Aging 2026;9:e75110) doi:10.2196/75110

KEYWORDS

older people; smart homes; assistive technology; user needs; user segmentation; residentially based lifestyles

Introduction

Population aging is a global challenge [1]. It is of utmost importance for older people to stay healthy and independent in their homes, where they feel comfortable. To support independent living, technology that can quickly detect hazards and respond appropriately in emergency situations becomes crucial for older people living in their own homes [2].

Smart home technology has made significant advances in the fields of remote management and older adult assistance [3]. Smart home services improve comfort, safety, and security at home with Internet-of-Things technology and allow older adults to receive medical treatment or check their health at home [4]. It is becoming a very important technology to support healthy and independent lives for older people [5]. Ultimately, this

XSL•FO
RenderX

technology can help improve the quality of life in the older population [6].

However, most older adults perceive smart technologies to be difficult to use, and there may be many barriers to adopting the technology in real life. In addition, older people may not even be aware of how smart home functions can meet their needs. For these reasons, it is expected that commercialization and widespread use of smart homes for older adults will take still some time [6]. In order for this technology to pervade their lives, older users should not be classified into a single group but should be subdivided according to the service purpose, so their specific needs can be identified and satisfied accordingly [7].

To address this challenge, we segmented older people according to their lifestyles within residential facilities, where smart homes are used. Our segmentation was based on residentially based lifestyles (RBL). Some studies claim that lifestyles related to a specific domain can provide meaningful results for segmentation of the population under study [8]. For example, Thøgersen [9] considered food-related lifestyles for analysis related to food consumption, and Sanquist et al [10] proposed the use of energy-related lifestyles as a domain to identify patterns of energy consumption.

In this study, we used a questionnaire on RBL and smart home needs that we have previously developed and validated in a Korean population, in which we confirmed that this is a tool that can lead to meaningful results [11]. Building upon that prior work, the present research applies the instrument to a sample of older adults in Spain, with the aim of proposing a segmentation method and analyzing the smart home requirements of each subgroup. This approach enables the assessment of both readiness and specific needs, thereby generating insights that can inform strategies to enhance the adoption and usability of smart home technologies among older adults.

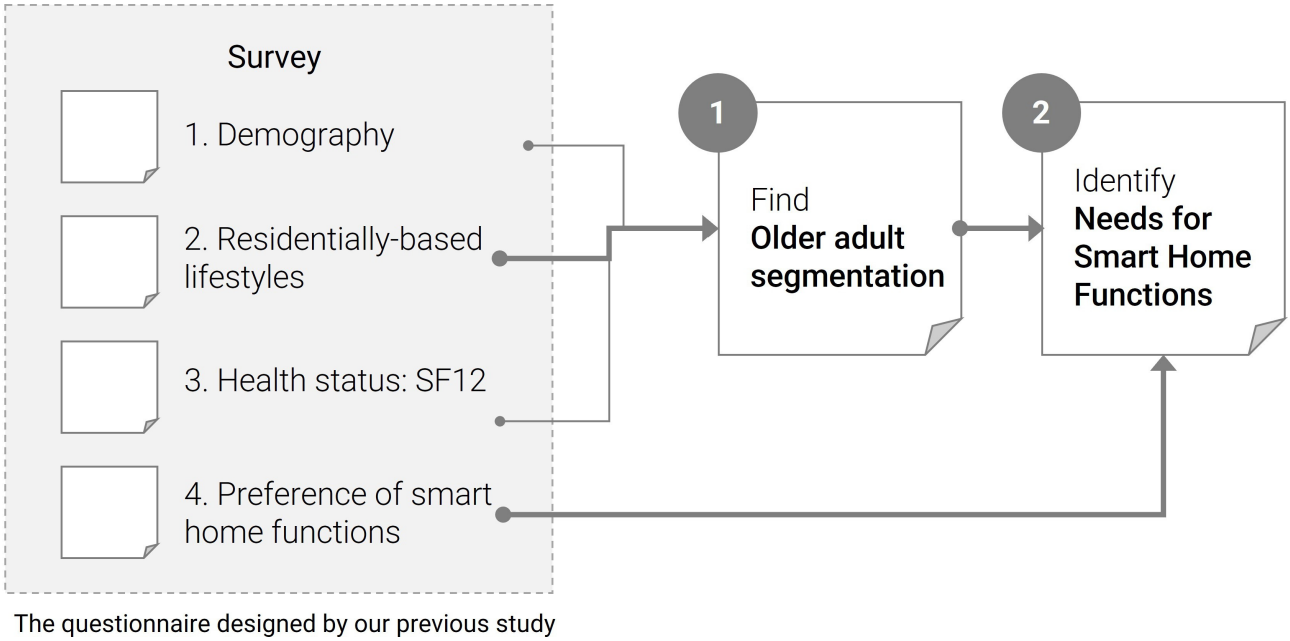
The remainder of the paper is organized as follows. The *Methods* section presents the study process, the participants, the questionnaire used, and the data analysis. Then, the *Results* section presents the principal component analysis (PCA) of the RBL, the segmentation, and the needs for the smart home functions. The *Discussion* section states first the main findings, comparing them with related works, and then it discusses the limitations. Finally, the *Conclusions* section closes the paper with the main contribution and future research.

Methods

Process

Figure 1 presents the study process. The resulting outcomes are (1) RBL-based segmentation of older adults in Spain and (2) analysis of the needs of smart home functions for each segment. This study proves that segmentation according to RBL and analysis of smart home needs through the questionnaire and method presented are useful to analyze the perception and detailed needs of smart homes, which are essential as assistive technologies for older adults.

Figure 1. Study process. SF12: 12-item Short-Form Health Survey.



Participants

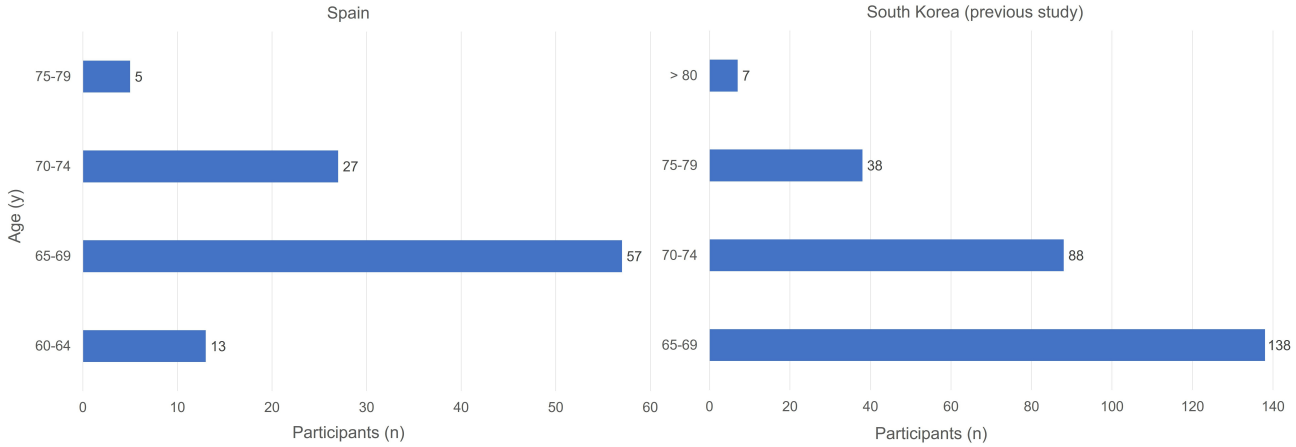
This study targeted older adults in Spain, a representative country facing an aging society. In the previous Korean study, the survey was conducted with older adults aged 65 years or older. Considering that the internet usage of older adults in South Korea decreases at higher ages than in Spain [12,13], we set the lower age limit at 60 years, 5 years below that of South Korea.

Therefore, the survey participants were designated as senior citizens aged 60 years and over, living in Spain (Figure 2). The criteria for a valid response were (1) age of 60 years or older, (2) no missing answers to questions, and (3) responses received within the specified survey period. We collected a total of 102 valid answers over 5 consecutive weeks. Regarding the

XSL•FO
RenderX

proportion of women and men, our sample included 41 men and 61 women (40.2% vs 59.8%), similar to that of Spain (45% vs 55%) [14]. The average age of the respondents was 68 (SD 3.8) years.

Figure 2. Age comparison of participants in 2 countries.



Instruments

We used a questionnaire that was developed and used in our previous Korean study [11], which consisted of a total of four sections: (1) demographics, (2) RBL, (3) health status, and (4) smart home function preferences. The demographics section consisted of eight questions: gender, age, marital status, residence type, caregiver, time spent out of the home, education, and monthly income. Questions for RBL were constructed by analyzing existing relevant studies [15-23] and explored the following: (1) quality of home life and needs for daily life, (2) daily home activities, and (3) social events at home.

In addition, the health status was checked through the Short-Form Health Survey of 12 items, so that the physical and mental health of each group could be compared after segmentation by RBL questions. With proven reliability and validity, the 12-item Short-Form Health Survey was developed to measure the health status of respondents in a short time. This survey calculates a physical component summary (PCS) and a mental component summary (MCS), which can be assessed based on the US average of 50 points. A high score means good status of physical or mental health [24].

Finally, regarding smart home function preferences, functions considered were based on a systematic review of smart home research [25]. On the basis of the results of this study, we could extract 26 functions in 8 categories. This section of the questionnaire demands the participant to select from 1 (not necessary) to 5 (most necessary) how much each of the 26 smart home functions is needed.

All questionnaires were carefully revised by experts in gerontology research and smart home technology.

Procedure

Data collection was conducted as an online survey, a non-face-to-face survey tool. This survey used convenience sampling, which is one of the nonprobability sampling methods. The questionnaire was distributed among older adults in different areas of the Madrid region via day care centers. The questionnaire was translated into Spanish, and it was distributed via Google Forms, together with a guide on the purpose of the study and the information that the survey responses would not be used for any other purpose except for research. Responses were collected over 5 weeks.

Ethical Considerations

We did not gather any personal data that could be linked in any way to the identity of the respondents; therefore, according to the rules of our ethical committee at Universidad Politécnica de Madrid [26], there was no need to request the prior approval of the committee. The survey does not require any personally identifiable data (eg, name, phone, email, and personal address) and does not require going through a log-in process in which personal information may be collected. In addition, respondents could voluntarily choose whether to participate in the survey or not; therefore, starting the survey constitutes a valid consent to participate. The collected responses were stored and analyzed anonymously. Moreover, the collected answers were not used for any purpose other than research analysis, and all initial data from the survey were deleted after analysis. No interventions, compensation, or treatments were included in this research.

Data Analysis

The items on RBL were dimensionally reduced to main factors through PCA. By applying PCA before clustering, the original 17 survey items were reduced to five interpretable dimensions: (1) independent living and quality of life, (2) social events at home, (3) regular life routine, (4) creative and healthy, and (5) time spent alone. This dimensional reduction served 2 main purposes. First, it minimized noise by removing items with low communalities and condensed correlated items into coherent factors, thereby improving the stability of the clustering solution. Second, it enhanced interpretability, as clusters could be explained in terms of meaningful lifestyle dimensions rather than raw questionnaire items. In high-dimensional spaces (many variables), distance metrics become less discriminating, and clustering algorithms such as K-means can perform poorly or yield unstable clusters. Compared to direct clustering on the original 17 variables, clustering on the 5 extracted components provided clearer group separation and yielded more distinct, interpretable lifestyle-based clusters. In contrast, clusters formed directly on all 17 items would be harder to interpret because

XSL•FO
RenderX

each cluster's distinguishing features would involve possibly many items with varying contributions. Moreover, parsimony in model building (ie, fewer variables) tends to generalize better when replicating or applying to new data. It is important to highlight that PCA does sacrifice some information present in the original variables, especially if lower variance components (or items with unique variance) are dropped; some cluster-defining features might lie in those discarded dimensions. In addition, PCA assumes linear relationships among items and may not capture complex nonlinear interactions.

For reliability analysis of the resulting factors, internal consistency was verified with Cronbach α value. K-means clustering based on these RBL factors could segment Spanish older respondents with similar patterns of RBL and characteristics. After segmentation by RBL, the demographics and health status of each group were compared. Then, through the questions on smart home function preferences, the needs for smart home functions for each group were measured.

To verify significant differences among groups, chi-square test validation was applied to the demographic data, and ANOVA analysis was used on the health status scores (PCS and MCS) and smart home needs scale.

Results

PCA of RBL

The mean Kaiser-Meyer-Olkin value for the RBL items was 0.604. After removing factors with commonalities of low scores, PCA was performed. Bartlett's test of sphericity was significant (χ^2_{136}=455.893; P<.001), indicating that the correlation matrix was not an identity matrix and was therefore suitable for factor analysis. PCA was selected to reduce the dimension to the main factors of each assessment item and analyzed by Varimax rotation. Table 1 presents the PCA results of the RBL items. Five principal factors were extracted by allowing factor loadings greater than 0.4: (1) independent living and quality of life, (2) social event at home, (3) regular life routine, (4) creative and healthy, and (5) time spent alone.

XSL•FO
RenderX

Table . Principal component analysis on residentially based lifestyles (RBL questions [11]).

RBL question	Independent living and quality of life	Social event at home	Regular life routine	Creative and healthy	Time spent alone
Independent living and quality of life (factor loadings)					
My chronic disease doesn't affect the quality of my daily life.	0.828	—[a]	—	—	—
My medication or treatments doesn't affect the quality of my daily life.	0.827	—	—	—	—
Generally, I feel satisfied with my daily life.	0.718	—	—	—	—
I don't need help to clean the house, wash the dishes or do the laundry.	0.695	—	—	—	—
I am satisfied with how I share my time.	0.599	—	—	—	—
I eat at least 3 times a day.	0.415	—	—	—	—
Social event at home (factor loadings)					
Invite friends	—	0.788	—	—	—
Visits of another type	—	0.784	—	—	—
Family visit	—	0.636	—	—	—
Regular life routine (factor loadings)					
My diet is varied and balanced.	—	—	0.732	—	—
I am not sleepy during the day.	—	—	0.707	—	—
Take care of my house	—	—	0.489	—	—
Creative and healthy (factor loadings)					
Cook or bake	—	—	—	0.732	—
Creative and artistic activities	—	—	—	0.646	—
Do light exercise	—	—	—	0.539	—
Time spent alone (factor loadings)					
Use a computer or smartphone or tablet at home	—	—	—	—	0.775
Read books, newspapers or magazines	—	—	—	—	0.626
Sum of squares loadings	3.051	2.034	1.674	1.625	1.593
Proportion variance	17.949	11.965	9.847	9.56	9.372
Cumulative variance	17.949	29.914	39.761	49.32	58.692
Cronbach α	0.8	0.651	0.489	0.424	0.449

[a]Not applicable.

XSL•FO
RenderX

The reliability of each factor was confirmed by the Cronbach α value. Although Cronbach α values of 3 factors were measured somewhat lower, those were accepted considering that the questionnaire was newly developed in the previous study and the survey was conducted with a limited number of people. Each factor has at least an eigenvalue of 1, and the total variance explained by these 5 factors is 58.69%. According to methodological studies in the social sciences, cumulative explained variance values of approximately 50% to 60% are considered adequate for survey-based instruments.

Segmentation Based on RBL

On the basis of 5 RBL factors, a hierarchical cluster analysis via the Ward method was conducted, and it was judged that classification into 4 clusters was the most stable and convincing. After that, final clustering was performed through K-mean cluster analysis. Table 2 presents the clustering results segmented into the 4 groups. ANOVA analysis was conducted to explore differences in RBL factors by clusters. The analysis shows that all factors are significantly different: (1) independent living and quality of life (50.80; P<.001), (2) social event at home (6.83; P<.01), (3) regular life routine (21.29; P<.001), (4) creative and healthy (5.98; P<.01), and (5) time spent alone (25.43; P<.001). Figure 3 shows the differences among the factors for each cluster.

The lifestyle patterns and characteristics of each group were defined according to the 5 factor values in the RBL.

- Cluster 1: “High quality and independent life” is a group that shows the highest quality of life and independent living ability.
- Cluster 2: “Poor quality of life” has the lowest quality of life compared to other groups and the least amount of time spent on hobbies at home alone. It also displays the most regular life routine.
- Cluster 3: “Social-centered life” has a lifestyle with the most frequent visits by friends and family to their homes and the least regular life routine.
- Cluster 4: “Creative and personal-centered hobbies at home” is a group of people who have many creative hobbies and activities that they can do alone.

Table . Cluster analysis for residentially based lifestyles (RBL) of older people.

	Cluster 1 (n=47), factor scores (mean)	Cluster 2 (n=5), factor scores (mean)	Cluster 3 (n=25), factor scores (mean)	Cluster 4 (n=25), factor scores (mean)	*F* test (*df*)
Independent living and quality of life	0.45	−3.16	0.07	−0.28	50.80[a] (3, 98)
Social event at home	−0.36	−0.19	0.67	0.04	6.83[a] (3, 98)
Regular life routine	0.36	1.05	−1.05	0.16	21.29[a] (3, 98)
Creative and healthy	−0.08	0.07	−0.48	0.62	5.98[b] (3, 98)
Time spent alone	−0.47	−1.08	0.04	1.06	25.43[a] (3, 98)

[a]P<.001.

[b].001≤P<.01.

XSL•FO
RenderX

Figure 3. Clusters for older people by the residentially based lifestyles.

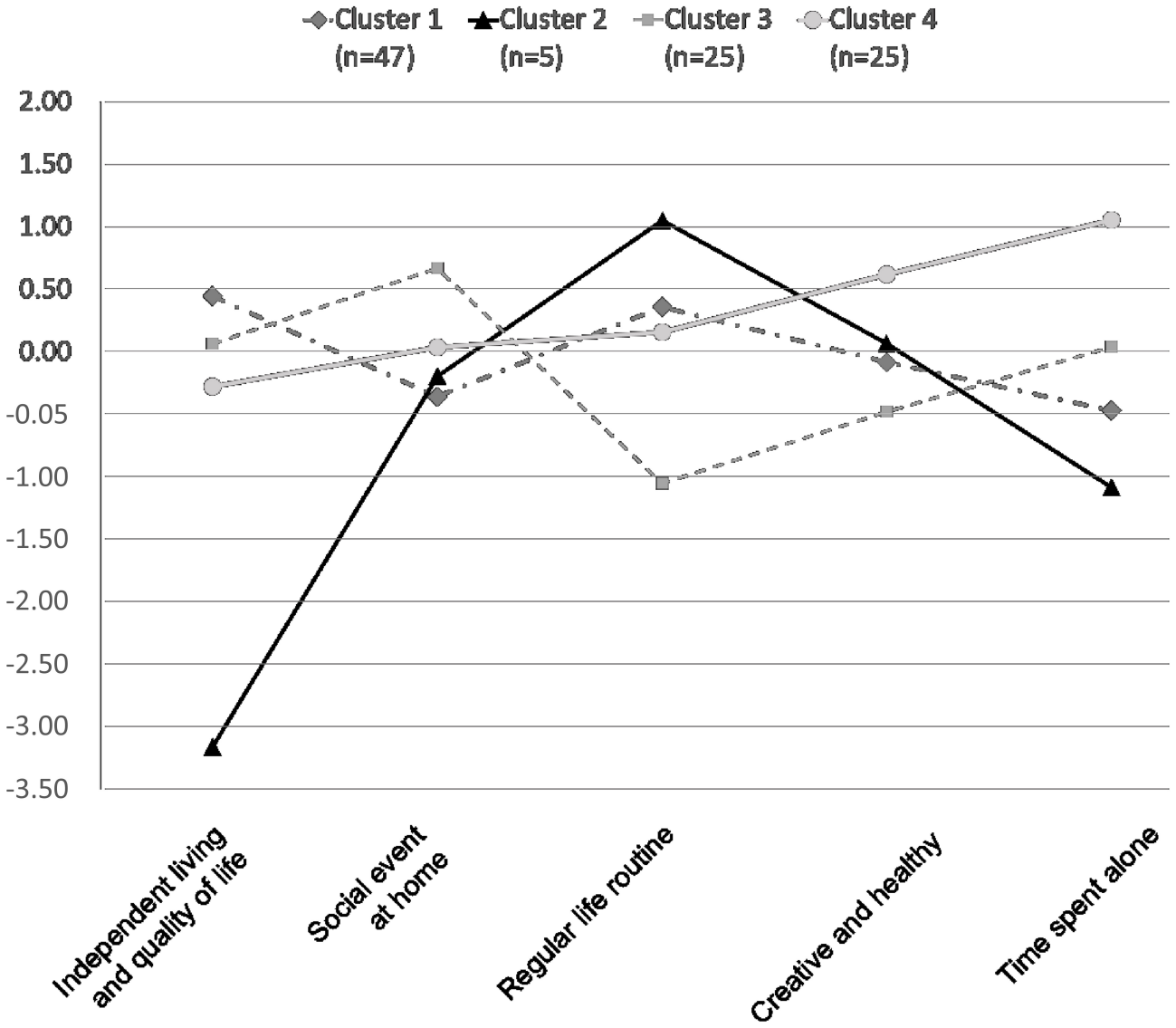


Table 3 presents the demographic characteristics of each group. There is a significant difference in gender of each group. Clusters 1 and 4 had more females with a male-to-female ratio of approximately 1:2, and cluster 2 is made up of 100% females. Cluster 3 was the only group in which the proportion of men was higher.

The MCS score of cluster 1 “high quality and independent life” was the highest, and the PCS score of cluster 4 “creative and personal-centered hobbies at home” is the highest. On the other hand, cluster 2 “poor quality of life” had the lowest both PCS and MCS ().

A one-way analysis of variance was conducted to examine differences in health status across the four residentially based lifestyle clusters. The results indicated statistically significant differences in physical component scores among the clusters (P=.01). The mean physical component scores were 50.33 for Cluster 1 (high quality of life and independent), 25.73 for Cluster 2 (poor quality of life), 47.74 for Cluster 3 (social-centered life), and 50.45 for Cluster 4 (creative and personal-centered hobbies at home). Specifically, Cluster 2 demonstrated substantially lower physical component scores compared with the other clusters, whereas Clusters 1, 3, and 4 showed relatively higher and comparable scores.

In contrast, differences in mental component scores did not reach statistical significance (P=.15). The mean mental component scores were 52.88 (SD 8.42) for Cluster 1, 43.48 (SD 10.12) for Cluster 2, 52.87 (SD 7.95) for Cluster 3, and 51.66 (SD 9.03) for Cluster 4, suggesting that mental health status was relatively similar across clusters.

XSL•FO
RenderX

Table . Demographics of each cluster.

	Cluster 1 (high quality of life and independent life; n=47), n (%)[a]	Cluster 2 (poor quality of life; n=5), n (%)[b]	Cluster 3 (social-centered life; n=25), n (%)[c]	Cluster 4 (creative and personal-centered hobbies at home; n=25), n (%)[d]	Chi-square (*df*)
Gender					12.9 (3)[e]
Male	15 (31.9)	0 (0.0)	17 (68.0)	9 (36.0)	
Female	32 (68.1)	5 (100.0)	8 (32.0)	16 (64.0)	
Age (y)					6.6 (9)
60-64	6 (12.8)	0 (0)	4 (16.0)	3 (12.0)	
65-69	26 (55.3)	4 (80.0)	13 (52.0)	14 (56.0)	
70-74	14 (29.8)	0 (0.0)	6 (24.0)	7 (28.0)	
75-79	1 (2.1)	1 (20.0)	2 (8.0)	1 (4.0)	
Marital status					16.3 (9)
Married	32 (68.1)	2 (40.0)	18 (72.0)	13 (52.0)	
Single	3 (6.4)	2 (40.0)	5 (20.0)	2 (8.0)	
Widowed	3 (6.4)	1 (20.0)	2 (8.0)	4 (16.0)	
Separated	9 (19.1)	0 (0)	0 (0.0)	6 (24.0)	
Resident type					11.9 (9)
Alone	12 (25.5)	2 (40.0)	6 (24.0)	11 (44.0)	
With spouse	30 (63.8)	2 (40.0)	14 (56.0)	10 (40.0)	
With family	4 (8.5)	0 (0)	4 (16.0)	4 (16.0)	
Others	1 (2.1)	1 (20.0)	1 (4.0)	0 (0)	
Caregiver					12.2 (12)
Spouse	33 (70.2)	2 (40.0)	18 (72.0)	13 (52.0)	
Son or daughter	3 (6.4)	1 (20.0)	2 (8.0)	5 (20.0)	
Other family	1 (2.1)	0 (0)	3 (12.0)	1 (4.0)	
Friends or neighbor	5 (10.6)	1 (20.0)	1 (4.0)	2 (8.0)	
Others	5 (10.6)	1 (20.0)	1 (4.0)	4 (16.0)	
Spent time out of home (h)					10.5 (9)
<1	2 (4.3)	2 (40.0)	2 (8.0)	4 (16.0)	
1-3	23 (48.9)	3 (60.0)	12 (48.0)	10 (40.0)	
3-6	18 (38.3)	0 (0.0)	8 (32.0)	8 (32.0)	
>6	4 (8.5)	0 (0.0)	3 (12.0)	3 (12.0)	
Education					8.3 (6)
Basic	1 (2.1)	1 (20.0)	2 (8.0)	0 (0)	
Intermediate	8 (17.0)	2 (40.0)	4 (16.0)	4 (16.0)	
University	38 (80.9)	2 (40.0)	19 (76.0)	21 (84.0)	
Income (€)[f]					5.5 (9)
<1500	5 (10.6)	1 (20.0)	2 (8.0)	2 (8.0)	
1500-2000	7 (14.9)	1 (20.0)	3 (12.0)	4 (16.0)	
2001-2999	15 (31.9)	3 (60.0)	7 (28.0)	10 (40.0)	
>3000	20 (42.6)	0 (0)	13 (52.0)	9 (36.0)	

[a]Residentially based lifestyles: They have independent life, high quality of life, and low social activity at home.

XSL•FO
RenderX

[b]Residentially based lifestyles: They have low quality of life and less personal free time.

[c]Residentially based lifestyles: They have many social events and low time spent alone.

[d]Residentially based lifestyles: They spend a lot of time at home and do various activities alone.

[e].001≤*P*<.01.

[f]€1=US $1.18.

Needs for Smart Home Functions

In this section, we compared the needs for smart home functions of each cluster. The respondents could choose from 1 (not necessary) to 5 (most necessary) for 26 detailed smart home functions in 8 categories. Higher numbers should be interpreted to mean "more necessary functions" for their lives (Table 4). As a result of ANOVA, there was a significant difference in the preference—F(3, 98)—of the *sensors you can wear as clothes* function. The smart home needs of cluster 2 "poor quality of life" are the highest, with an average score of 2.06 for a total of 26 functions, while the needs of cluster 1 "high quality of life and independent life" are the lowest, with an average score of 1.75.

We compared the needs of smart home functions above the median, based on the Spanish responses (value=2) (highlighted in italic in Table 4 and ordered by decreasing priority in Table 5). Cluster 2 "poor quality of life" needed more smart home functions compared to other groups. On the other hand, "high quality and independent life" and "social-centered life" had relatively few needs for the smart home functions.

There was also difference in the most essential functions for each group. Cluster 1 "high quality and independent life" required the *gas leak detection or smoke detection* function the most, cluster 2 "poor quality of life" required the *telehealthcare* function, cluster 3 "social-centered life" required the *remote-controlled electricity* function, and cluster 4 "creative and personal-centered hobbies at home" required the *automatic heating* function the most.

However, it is important to highlight that there are three functions—*automatic heating*, *telehealthcare*, and *call the family in case of emergency*—which are essential in all 4 groups. This indicates that the Spanish older population is ready to adopt such services. On the other hand, we observed the functions that were not considered important in any of the groups: *automatic watering in the garden*, *automatically checked postbox*, *monitoring in your absence*, and *robot like a friend*. Table 5 lists which smart home functions are needed according to the perceived priority in each group.

XSL•FO
RenderX

Table . The needs for smart home functions (function list [11]).

Smart home functions	Cluster 1 (n=47, 46.1%; high quality of life and independent life), mean (SD)	Cluster 2 (n=5, 4.9%; poor quality of life), mean (SD)	Cluster 3 (n=25, 24.5%; social-centered life), mean (SD)	Cluster 4 (n=25, 24.5%; creative and personal-centered hobbies at home), mean (SD)	*F* test (*df*)
Smart washing machine	1.85 (0.83)	*2.2* [a] *(1.10)*	1.84 (0.85)	1.8 (0.76)	0.045 (3, 98)
Smart refrigerator	1.45 (0.65)	*2.2 (0.84)*	1.6 (0.71)	1.76 (0.78)	1.374 (3, 98)
Smart cooker and smart coffee maker	1.43 (0.62)	*2.4 (1.14)*	1.72 (0.79)	1.76 (0.83)	2.468 (3, 98)
Schedule reminder	1.49 (0.66)	*2.4 (1.14)*	1.44 (0.65)	1.76 (0.83)	0.673 (3, 98)
Reminder to take medicine	*1.55 (0.71)*	*2.6 (1.34)*	1.32 (0.56)	1.52 (0.71)	0.409 (3, 98)
Open windows or doors detection	1.62 (0.74)	*2 (1.00)*	1.48 (0.65)	1.64 (0.76)	0.035 (3, 98)
Automatic watering in the garden	1.77 (0.84)	1 (0)	1.52 (0.71)	1.96 (0.93)	0.128 (3, 98)
Automatically checked postbox	1.72 (0.80)	1 (0)	1.68 (0.75)	1.88 (0.88)	0.366 (3, 98)
Automatic heating	*2.28 (1.08)*	*2.8 (1.30)*	*2.24 (1.05)*	*2.52 (1.20)*	0.359 (3, 98)
Gas leak detection or smoke detection	*2.38 (1.12)*	*1.8 (0.84)*	*2.12 (1.01)*	*2.48 (1.16)*	0 (3, 98)
Remote-controlled electricity	*2.06 (0.95)*	*1.4 (0.55)*	*2.28 (1.08)*	*2.4 (1.12)*	1.93 (3, 98)
Remote-controlled lights	1.77 (0.84)	1.8 (0.84)	1.8 (0.82)	*2.08 (0.95)*	1.269 (3, 98)
Light mode settings	1.77 (0.84)	1.8 (0.84)	1.88 (0.88)	*2.24 (1.05)*	3.23 (3, 98)
TV auto-play function, notifications from my channel	1.66 (0.77)	2.4 (1.14)	1.72 (0.79)	1.88 (0.88)	0.779 (3, 98)
Personalized learning TV content	1.79 (0.86)	1.6 (0.55)	1.44 (0.65)	*2 (1.00)*	0.072 (3, 98)
Automatic cinema mode setting	1.72 (0.80)	2 (1.00)	1.8 (0.82)	1.76 (0.78)	0.053 (3, 98)
Fall detection	*2.23 (1.08)*	*2.8 (1.30)*	1.92 (0.91)	2.00 (1)	0.985 (3, 98)
Monitoring in your absence	1.74 (0.81)	1.8 (0.84)	1.56 (0.71)	1.96 (0.94)	0.216 (3, 98)
Measure sleep health	1.77 (0.84)	1.8 (0.84)	*2 (1)*	*2.04 (1.02)*	1.77 (3, 98)
Sensors you can wear as clothes	1.53 (0.70)	*2.2 (1.1)*	1.88 (0.88)	1.96 (0.94)	3.965[b] (3, 98)
Mental health detection	1.47 (0.66)	*2.4 (1.14)*	1.84 (0.85)	1.84 (0.85)	3.035 (3, 98)
Telehealthcare	*2.06 (1)*	*3 (1.41)*	*2.16 (1.03)*	*2.4 (1.12)*	0.002 (3, 98)
Home management assistant robot	1.6 (0.75)	*2.4 (1.14)*	*2 (1)*	1.76 (0.83)	1.115 (3, 98)
Equipment to help you up and down stairs	1.34 (0.57)	*2 (1)*	1.2 (0.41)	1.68 (0.79)	1.313 (3, 98)
Robot like a friend	1.28 (0.54)	1.2 (0.45)	1.28 (0.54)	1.32 (0.56)	0.073 (3, 98)
Call the family in case of emergency	*2.13 (1.02)*	*2.6 (1.34)*	*2.12 (1.01)*	*2.04 (1.02)*	0.098 (3, 98)
Average of needs	1.75	*2.06*	1.76	1.94	—[c]

[a]Italics indicate values higher than the median of 2.0.

XSL•FO
RenderX

[b]Statistical significance at the $P<.05$ level.

[c]Not applicable.

Table . Smart home function priority for each cluster.

Priority	Cluster 1 (n=47, 46.1%; high quality of life and independent life)	Cluster 2 (n=5, 4.9%; poor quality of life)	Cluster 3 (n=25, 24.5%; social-centered life)	Cluster 4 (n=25, 24.5%; creative and personal-centered hobbies at home)
1	Gas leak detection or smoke detection	Telehealthcare	Remote-controlled electricity	Automatic heating
2	Automatic heating	Automatic heating and fall detection	Automatic heating	Gas leak detection or smoke detection
3	Fall detection	Reminder to take medicine and call the family in case of emergency	Telehealthcare	Remote-controlled electricity and telehealthcare
4	Call the family in case of emergency	Smart cooker and smart coffee maker, schedule reminder, TV auto-play function, notifications from my channel, mental health detection, and home management assistant robot	Gas leak detection or smoke detection and call the family in case of emergency	Light mode settings
5	Remote-controlled electricity	Smart washing machine	Measure sleep health	Remote-controlled lights

Discussion

Principal Findings

This study identifies the smart home needs of older adults in Spain through a segmentation according to RBL. Through this approach, we can assess the readiness of older people to use and adopt concrete smart home functions. Concretely, we found 4 clusters of the Spanish older adults, identified the smart home functions that each group considers important, and analyzed how the smart home needs of older users relate to RBL.

To derive meaningful lifestyle clusters from the original 17 survey items, we first applied PCA to reduce dimensionality before performing K-means clustering on the extracted component scores. Specifically, after removing items with low communalities, PCA with Varimax rotation yielded 5 interpretable components (ie, independent living and quality of life, social events at home, regular life routine, creative and healthy, and time spent alone), each having an eigenvalue of 1 or greater. These 5 components explained approximately 58.7% of the total variance in the data.

Using these components for clustering rather than the full set of original variables offers several methodological advantages:

1. Reduction of redundancy and noise: Some of the 17 items are correlated and share underlying variance; PCA consolidates that shared variance and filters out items or variance that contribute less to the common structure. This reduces noise (ie, irrelevant or weakly contributing variation) that might obscure true cluster structure.
2. Mitigation of the curse of dimensionality: In high-dimensional space (many variables), distance metrics become less discriminating, and clustering algorithms such as K-means can perform poorly or yield unstable clusters. By reducing dimensions to those that capture the majority of variance and represent interpretable dimensions, we improve the cluster algorithm's ability to separate groups meaningfully.
3. Improved interpretability: Clusters defined on principal component scores are easier to characterize: each component is labeled (eg, "regular routine," "social at home," etc), so the resulting clusters can be described in terms of these meaningful dimensions. In contrast, clusters formed directly on all 17 items would be harder to interpret because each cluster's distinguishing features would involve possibly many items with varying contributions.
4. Stability and parsimony: Working with fewer, more stable dimensions tends to improve the reliability of clustering; clusters are less likely to be driven by measurement error or idiosyncratic items. Also, parsimony in model building (ie, fewer variables) tends to generalize better when replicating or applying to new data.
5. Facilitation of validity checks: Components allow checking internal consistency (Cronbach α) for each extracted dimension, which helps ensure that the reduced dimensions are coherent constructs. Additionally, when the components are reliably measured, using them in clustering gives more confidence in subsequently interpreting differences among clusters.

In our case, after PCA, we used K-means clustering on the 5 component scores and identified 4 clusters with distinct lifestyle profiles. Compared to what clustering on the full 17 variables would likely have yielded, the component-based clustering produced more coherent, stable, and interpretable clusters. For example, rather than clusters being defined by numerous item-specific idiosyncrasies, each cluster's profile can be described in terms of high or low scores on the 5 factors.

XSL•FO
RenderX

It should be noted that PCA does sacrifice some information present in the original variables, especially if lower variance components (or items with unique variance) are dropped; some cluster-defining features might lie in those discarded dimensions. Also, PCA assumes linear relationships among items and may not capture complex nonlinear interactions. Nonetheless, given our sample size and the moderate Kaiser-Meyer-Olkin (~0.60), the approach of PCA followed by K-means was judged appropriate.

Visutsak and Daoudi [27] attempted to identify the perceptions, accessibility factors, and needs of the older adults for smart homes through a survey. They asked about 7 smart home features in total and received responses from 20 stakeholders (doctors, nurses, the older adults, and their adult children). In a study on perception of smart home technologies [4], a survey was conducted with the general population, including working people in hospitals and nursing homes, to identify awareness of 6 smart home technologies for the older adults and concerns for using. This identified perceptions of usefulness for key technologies and their concerns from 18 respondents. Unlike previous studies, our study focuses on the older adults who will be the actual primary users. In addition, our questionnaire asked about the needs and perceptions of 26 detailed functions in a total of 8 categories of smart homes, covering a wider range of functions than previous studies. In addition, we compared and analyzed the particular needs of each group by segmentation based on the RBL.

In this study, rather than targeting older adults with a specific status (cognitive level and health status), the participants were segmented according to RBL. In addition, after segmentation, the average physical and mental health status of each group could be analyzed in connection with smart home needs. Although smart home technology could have great benefits for older people who need physical assistance or have mild cognitive impairment, we believe that smart home technology will help maintain independent living both for the healthy older adults and those with poor health.

Four segmented groups were identified in Spain: (1) high quality of life and independent life, (2) poor quality of life, (3) social-centered life, and (4) creative and personal-centered hobbies at home. We figured out the priority of smart home functions for each group and found that the priority of each group and the number of required functions were different. Although the 4 clusters provided meaningful and interpretable lifestyle profiles, 1 cluster (cluster 2, n=5) was substantially smaller than the others. At first glance, such a small subgroup may raise concerns about representativeness and the robustness of comparisons across clusters. However, several factors justify retaining this cluster in the analysis. First, the emergence of a small cluster was not an artifact of the analytic procedure but rather a reflection of genuine heterogeneity within the data. Even after multiple reclusterings, cluster 2 consistently appeared, suggesting that it represents a stable and reproducible pattern rather than a random byproduct of the algorithm. Second, the value of clustering lies not only in identifying large, dominant groups but also in revealing smaller, distinctive subgroups. In many areas of social science and consumer research, rare or niche groups play a crucial role in theory development and practice, despite their limited size [28,29]. Third, prior work in clustering methodology emphasizes that the interpretability and stability of clusters are more critical than their absolute size [30]. From this perspective, the consistent detection of cluster 2 across replications strengthens its validity as a meaningful subgroup. Finally, while the small number of respondents in this cluster warrants caution in generalization, its unique profile provides important insight into lifestyle heterogeneity. Future research with larger and more diverse samples can further verify the existence and characteristics of this subgroup, but its inclusion in this study enhances rather than detracts from the comprehensiveness of the clustering solution.

Cluster 2 “poor quality of life,” which is the group with the weakest mental and physical health, required more smart home functions than the other clusters. Their top priority is *Telehealth*, a system that allows them to receive medical care without going to a hospital. Besides that, they valued almost all categories of features as important; especially, they considered the *automation of daily routines* category and the *assist mobility* category as more important than the other groups. Also, as they spend more time at home than other groups, it was confirmed that they were interested in setting up TV programs and movie modes to enjoy at home.

Although only a limited number of functions showed statistically significant differences across the 4 clusters, this does not necessarily diminish the value of the segmentation. Clustering methods aim primarily to uncover underlying heterogeneity in patterns rather than to maximize between-group statistical significance on individual variables [28,31]. Even when differences in specific functions are not statistically significant, the cluster solution may still provide meaningful insights.

First, the overall profiles of the clusters reflect distinct combinations of lifestyle and technology needs that are not captured when examining single variables in isolation. Prior work has emphasized that segmentation should be evaluated on interpretability and theoretical coherence, rather than significance tests alone [29,32]. Second, the relatively small sample size in some clusters may have limited statistical power to detect differences. This limitation is particularly relevant for cluster 2, but the stability of this subgroup across repeated clustering iterations supports its substantive validity.

Thus, while statistical tests provide one form of validation, the meaningful differentiation in the patterns of technology needs across clusters—combined with theoretical plausibility and methodological robustness—justifies the retention and interpretation of the 4-cluster solution. Comparing these results with the segmentation and smart home needs analyzed previously in the Korean study [11], we noticed that the median value of smart home needs in South Korea was 3 (IQR 1-4), which is higher than in Spain (2, IQR 1-2). We believe that the difference in the popularization of smart homes in the 2 countries [33,34] and the internet use among the older adults can affect overall smart home perception and needs. It is possible to establish some parallels between clusters in Spain and in Korea. For instance, cluster 1 “high quality of life and independent life” in Spain and Korea’s “high quality of life and healthy” have similar RBL and had the lowest demands for

XSL•FO
RenderX

smart home functions compared to other groups. However, we can highlight that although Spain and Korea have similarities, the segmentation extracted different clusters and needs. This difference arises because instead of imposing a top-down classification, we obtained the clusters in a bottom-up fashion, maintaining in this way the particularities of each country. We believe this is a very useful tool to understand the idiosyncrasy of each culture and to better design smart home technologies that satisfy older adults' needs.

Limitations

A first limitation can be pointed out regarding the number and heterogeneity of participants: Due to the characteristics of the non–face-to-face survey, we could not target all the older adults in Spain but only those older adults who could respond to an internet survey. Also, the number of responses collected is related to unfamiliarity with online surveys for the older adults in this country. However, the use of information and communication technologies is increasing year by year in Spain. In particular, the daily internet usage rate among those aged 65 to 74 years increased to 50.7% in 2020, compared to 43.3% in 2019 [35,35]. We believe it is a practical method to quickly gather data and reach conclusions for the older population, and we are confident that it will become a more effective method in Spain.

The channel for the distribution of the survey may also have an impact on the educational level of respondents. We collected 78.43% of responses from older people with university education, which is significantly higher than the percentage of Spanish adults who have reached tertiary education (approximately 38%). In addition, because participants were able to complete this online survey alone, we acknowledge there could be a bias in our sample to older adults with good cognitive status, familiarity with technology, and possibly a higher willingness to adopt smart home technologies. This can pose a threat to the representativeness of the sample, which is important, as smart home technology could especially benefit older adults with cognitive decline. In future research, we aim to collect feedback from older adults with lower educational levels and more varied cognitive status, to increase the representativeness of the results. Further, given that the survey was anonymous, and we were not controlling who the person answering the questionnaire was, there was a risk of collecting responses from participants who do not meet the target profile. However, the channels that we used for the distribution of the questionnaire increase our confidence in effectively being able to reach the target population.

Another issue to be highlighted is the low number of respondents who were classified into cluster 2 "poor quality of life" (n=5). As a result, the demographic characteristics and responses of each individual member have a considerable influence in the definition of the cluster. We were concerned about the small dimension of cluster 2, so we tested grouping in various ways, but these 5 adults could not be tied to other groups. Although the cluster with 5 people was a small group, we think it is a meaningful group because of their distinct characteristics and RBL. In addition, small but stable clusters have been shown in prior literature to capture unique patterns or extreme profiles that larger clusters cannot reveal [29]. Retaining this cluster therefore enhances our ability to reflect the full heterogeneity of the data, even if statistical generalizations must be made with caution. Following established recommendations [36], we decided to keep the 4-cluster solution, while explicitly noting the interpretive limitations associated with the small sample size of cluster 2.

We believe the reduced number of participants in this group may be due to limited internet access of older people with this profile. We observe that the proportion of members aged older than 75 years in this cluster is the highest (20%), whereas the representation of this age segment in the other 3 clusters is much lower (with a maximum of 8%). Considering that the quality of life tends to decrease with age, this cluster might be biased to be defined by the low quality of life of its members. We would like to explore this issue by collecting more responses on the older age segment.

A more fundamental limitation of this study (but also one of its strengths) is that the results are inherently dependent on the target population. Spanish RBLs may differ from those in other countries and cultures. Being a data-driven research process, the factors that emerge through PCA are different, and the clustering process also leads to groups that are not necessarily comparable to the ones in another population. Direct generalization of the conclusions to another population should therefore be avoided. However, we consider this as a major strength of our method. The independent results of the segmentation by countries (Spain and South Korea) are evidence that smart home technology must be tailored to each specific context. This method may help to investigate the context and guide the tailoring process in technology development and deployment. This is also an interesting tool to guide policymakers in deciding concrete actuations to foster the integration of smart home technologies with existing services.

Conclusions

A smart home is an assistive technology that can improve the quality of life in old age and lead to an independent life. In order for this rapidly growing technology to fulfill its role in the lives of older people, it is necessary to first identify and analyze the actual needs of potential users. Older people are generally going through stages that may be physically and mentally debilitating, but their lifestyles at home cannot be defined as homogeneous. If the specific needs of older people are met by analyzing and segmenting life patterns at home, this assistive technology can fully perform its role. Our main contribution is a method that can be applied to any older adult group to analyze their smart home needs. In this study, we applied previously defined methods and questionnaires in a different geographical context. As confirmed through the analysis results, there was a difference in the priority of smart home functions and the number of functions required according to segmentation by RBL. It was also confirmed that RBL-based smart home needs analysis is meaningful even for studies conducted in different countries with different cultures. This second study demonstrates that the proposed questionnaire and research method are generally useful and not limited to specific countries or contexts. The results of this study and the previous study in South Korea showed that

XSL•FO
RenderX

this method can be used to investigate the perception of smart home technology in older adults and identify the actual needs of the groups that emerge.

Future research may explore a more sophisticated yet comprehensive categorization of smart home features, exploring the possibility of hierarchically grouping the 26 individual smart home functions into function categories or themes, to improve the interpretation of the preferences of different user groups.

Funding

The authors declared no financial support was received for this work.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Conflicts of Interest

None declared.

References

1. Global health and aging. : World Health Organization; 2011 URL: https://www.nia.nih.gov/sites/default/files/2017-06/global_health_aging.pdf [accessed 2026-02-04]
2. Das R, Tuna G, Tuna A. Design and implementation of a smart home for the elderly and disabled. Int J Comput Netw Appl 2015;2(6):242-246 [FREE Full text]
3. Chan M, Estève D, Escriba C, Campo E. A review of smart homes- present state and future challenges. Comput Methods Programs Biomed 2008 Jul;91(1):55-81. [doi: 10.1016/j.cmpb.2008.02.001] [Medline: 18367286]
4. Chernbumroong S, Atkins AS, Yu H. Perception of smart home technologies to assist elderly people. 2010 Presented at: Proceedings of the 4th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2010); Aug 25-27, 2010; Paro, Bhutan p. 90-97 URL: https://files01.core.ac.uk/download/pdf/18270182.pdf [accessed 2026-02-04]
5. Deen MJ. Information and communications technologies for elderly ubiquitous healthcare in a smart home. Pers Ubiquit Comput 2015 Jul;19(3-4):573-599. [doi: 10.1007/s00779-015-0856-x]
6. Pal D, Triyason T, Funikul S. Smart homes and quality of life for the elderly: a systematic review. Presented at: 2017 IEEE International Symposium on Multimedia (ISM); Dec 11-13, 2017; Taichung, Taiwan p. 413-419. [doi: 10.1109/ISM.2017.83]
7. Dolnicar S, Grün B. Challenging “Factor–Cluster Segmentation”. J Travel Res 2008 Aug;47(1):63-71. [doi: 10.1177/0047287508318910]
8. van Raaij WF, Verhallen TMM. Domain‐specific market segmentation. Eur J Mark 1994 Oct 1;28(10):49-66. [doi: 10.1108/03090569410075786]
9. Thøgersen J. Sustainable food consumption in the nexus between national context and private lifestyle: a multi-level study. Food Qual Prefer 2017 Jan;55:16-25. [doi: 10.1016/j.foodqual.2016.08.006]
10. Sanquist TF, Orr H, Shui B, Bittner AC. Lifestyle factors in U.S. residential electricity consumption. Energy Policy 2012 Mar;42:354-364. [doi: 10.1016/j.enpol.2011.11.092]
11. Yu J, de Antonio A, Villalba-Mora E. Older adult segmentation according to residentially-based lifestyles and analysis of their needs for smart home functions. Int J Environ Res Public Health 2020 Nov 16;17(22):8492. [doi: 10.3390/ijerph17228492] [Medline: 33207793]
12. Población que ha usado internet de manera frecuente en los últimos tres meses por nivel de estudios [Article in Spanish]. Instituto Nacional de Estadística (INE). URL: https://www.ine.es/jaxi/Datos.htm?path=/t00/mujeres_hombres/tablas_1/&file=c05002.px#!tabs-grafico [accessed 2026-02-04]
13. 2020 internet usage survey in south Korea. : Ministry of science and ICT; 2020 URL: https://www.msit.go.kr/publicinfo/view.do?sCode=user&mId=63&mPid=62&publictSeqNo=406&publictListSeqNo=8&formMode=R&referKey=406%2C8 [accessed 2026-02-27]
14. Población residente según sexo, grupo de edad, estado civil y nacionalidad [Article in Spanish]. Instituto Nacional de Estadística (INE). URL: https://www.ine.es/jaxi/Datos.htm?path=/t20/p274/serie/prov/p04/l0/&file=01001.px [accessed 2026-02-04]
15. Bone PF. Identifying mature segments. J Consum Mark 1991 Apr;8(4):19-32. [doi: 10.1108/07363769110035126]
16. Moschis GP, Lee E, Mathur A. Targeting the mature market: opportunities and challenges. J Consum Mark 1997 Aug 1;14(4):282-293. [doi: 10.1108/07363769710188536]
17. Sudbury L, Simcock P. A multivariate segmentation model of senior consumers. J Consum Mark 2009 Jun 26;26(4):251-262. [doi: 10.1108/07363760910965855]
18. Du C. Study on lifestyle evaluation systems for the health of Chinese elderly. Int J Environ Res Public Health 2019 Jan 21;16(2):2. [doi: 10.3390/ijerph16020284] [Medline: 30669558]

XSL•FO
RenderX

19. Fallah Mehrabadi E, Pakgohar M, Asadi S, Haghani H. Lifestyle of elderly people with osteoporosis and its related factors. Iran J Ageing 2017;12(2):132-145. [doi: 10.21859/sija-1202132]
20. Saint Martin M, Sforza E, Barthélémy JC, et al. Long-lasting active lifestyle and successful cognitive aging in a healthy elderly population: The PROOF cohort. Rev Neurol (Paris) 2017 Dec;173(10):637-644. [doi: 10.1016/j.neurol.2017.05.009] [Medline: 29100612]
21. Shahboulaghi FM, Karbalaee Z, Zarghi A, Moghaddam AG, Mohammadi M. Lifestyle of the elderly and its predictive factors. J Res Med Dent Sci 2018;6(1):219-225. [doi: 10.5455/jrmds.20186136]
22. Yuan M, Chen J, Han Y, et al. Associations between modifiable lifestyle factors and multidimensional cognitive health among community-dwelling old adults: stratified by educational level. Int Psychogeriatr 2018 Oct;30(10):1465-1476. [doi: 10.1017/S1041610217003076] [Medline: 29444740]
23. Zhang X, Xia R, Wang S, et al. Relative contributions of different lifestyle factors to health-related quality of life in the elderly. Int J Environ Res Public Health 2018 Feb 3;15(2):256. [doi: 10.3390/ijerph15020256] [Medline: 29401658]
24. Ware JE Jr, Keller SD, Kosinski M. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales, 3rd edition: Health Institute, New England Medical Center; 1998. URL: https://books.google.co.in/books/about/SF_12.html?id=OM0pcAAACAAJ&redir_esc=y [accessed 2026-02-04]
25. Marikyan D, Papagiannidis S, Alamanos E. A systematic review of the smart home literature: a user perspective. Technol Forecast Soc Change 2019 Jan;138:139-154. [doi: 10.1016/j.techfore.2018.08.015]
26. Regulations of the Ethics Committee for R&D+i Activities at the Universidad Politécnica De Madrid. : Universidad Politécnica de Madrid; 2017 URL: https://www.upm.es/gsfs/SFS36248 [accessed 2026-02-27]
27. Visutsak P, Daoudi M. The smart home for the elderly: perceptions, technologies and psychological accessibilities: the requirements analysis for the elderly in thailand. Presented at: 2017 XXVI International Conference on Information, Communication and Automation Technologies (ICAT); Oct 26-28, 2017; Sarajevo, Bosnia and Herzegovina p. 1-6. [doi: 10.1109/ICAT.2017.8171625]
28. Dolnicar S. A review of data-driven market segmentation in tourism. J Travel Tour Mark 2002 Jul 12;12(1):1-22. [doi: 10.1300/J073v12n01_01]
29. Ketchen DJ, Shook CL. The application of cluster analysis in strategic management research: an analysis and critique. Strat Mgmt J 1996;17(6):441-458. [doi: 10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G]
30. Jain AK. Data clustering: 50 years beyond K-means. Pattern Recogn Lett 2010 Jun;31(8):651-666. [doi: 10.1016/j.patrec.2009.09.011]
31. Hennig C. Cluster-wise assessment of cluster stability. Comput Stat Data Anal 2007 Sep;52(1):258-271. [doi: 10.1016/j.csda.2006.11.025]
32. Simon K. Digital 2021: Spain. DataReportal. 2021. URL: https://datareportal.com/reports/digital-2021-spain [accessed 2026-02-04]
33. Equipamiento y uso de TIC en los hogares [Report in Spanish]. : Instituto Nacional de Estadística (INE); 2019 URL: https://www.ine.es/prensa/tich_2019.pdf [accessed 2026-02-04]
34. Equipamiento y uso de TIC en los hogares [Report in Spanish]. : Instituto Nacional de Estadística (INE); 2020 URL: https://www.ine.es/prensa/tich_2020.pdf [accessed 2026-02-04]
35. Everitt BS, Landau S, Leese M, Stahl D. Cluster Analysis, 5th edition: John Wiley & Sons, Ltd; 2011. [doi: 10.1002/9780470977811]

Abbreviations

MCS: mental component summary
PCA: principal component analysis
PCS: physical component summary
RBL: residentially based lifestyles

Edited by M Mardini; submitted 28.Mar.2025; peer-reviewed by J Yang, LF Belo; revised version received 30.Oct.2025; accepted 08.Dec.2025; published 05.May.2026.

Please cite as:
Yu J, de Antonio A, Villalba-Mora E
Segmentation of Older People's Needs and Readiness for Smart Homes by Residentially Based Lifestyles in Spain: Survey Study
JMIR Aging 2026;9:e75110
URL: https://aging.jmir.org/2026/1/e75110
doi:10.2196/75110

XSL•FO
RenderX

© Jiyeon Yu, Angélica de Antonio, Elena Villalba-Mora. Originally published in JMIR Aging (https://aging.jmir.org), 5.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Web-Based Application to Support Caregivers in the Use of Learning Optimization Methods: Participatory Action Research Study

Chantal Viscogliosi[1], PhD; Hubert Kenfack Ngankam[2], PhD; Romane Duquette-Laplante[3], BA; Sarah Rahimaly[3], MA; Hélène Pigot[4], PhD; Bernard Chassé[5], PhD; Sara Ait Bouziaren[3], BSc; Yves Couturier[6], PhD; Jessica Déry[7], MSc; Nicolas Biard[3], BSc; Jordan Duchesne[1], MSc; Dominique Giroux[8], PhD; Sabrina Jacob[1], MSc; Kim Routhier-Chevrier[9], DEC; Jordan Mino-Roy[1], MSc; Mélanie Caron[1], MSc; Véronique Provencher[1], PhD

[1]École de réadaptation, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, 3001, 12e avenue Nord, Sherbrooke, QC, Canada

[2]Département d'informatique, Faculté des sciences, Université de Sherbrooke, Sherbrooke, QC, Canada

[3]Département de psychologie, Faculté de lettres et sciences humaines, Université de Sherbrooke, Sherbrooke, QC, Canada

[4]Département d'informatique, Faculté des sciences, Université de Sherbrooke, Sherbrooke, QC, Canada

[5]Société Alzheimer des Maskoutains - Vallée des Patriotes, Saint-Hyacinthe, QC, Canada

[6]École de travail social, Faculté des lettres et sciences humaines, Université de Sherbrooke, Sherbrooke, QC, Canada

[7]CIUSSS Mauricie-Centre-du-Québec, Trois-Rivières, QC, Canada

[8]École des sciences de la réadaptation, Université Laval, Québec, QC, Canada

[9]CEGEP de Sherbrooke, Sherbrooke, QC, Canada

Corresponding Author:
Chantal Viscogliosi, PhD
École de réadaptation, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, 3001, 12e avenue Nord, Sherbrooke, QC, Canada

Abstract

Background: Major neurocognitive disorders (MNCDs) frequently lead to difficulties in performing activities. Several studies have shown that people living with an MNCD benefit from the use of learning optimization methods from cognitive rehabilitation, such as error-free learning, motor encoding, spaced retrieval, and fading, which promote the safe pursuit of their meaningful activities. However, while the principles of learning optimization methods are relatively straightforward, the personalized application of these methods to the specific situations encountered can be more difficult.

Objective: The aim of this study was to describe the codevelopment process, including the validation of a web application called "Aide-Mémoire-Interactif (AMI)," a tool to help in personalizing learning optimization methods.

Methods: To design the web application, participatory action research based on a codevelopment and validation process was carried out. The various stages leading to the codevelopment of the AMI web application were grouped into 3 phases: (1) assessment of caregivers' needs (identification of situations frequently encountered by people living with an MNCD); (2) production of the first version of the AMI web application; and (3) validation of the AMI application. Individual interviews and workshops were conducted with 20 participants (caregivers, health and social service professionals, and community organization workers) to obtain a diversity of viewpoints and application contexts of use.

Results: The AMI web application was developed through an iterative, participatory process involving caregivers and professionals. Participants identified daily situations requiring the use of learning optimization methods and provided feedback that guided successive refinements to the content, answer choices, navigation, and personalization features. The logical architecture was built according to the rationale underlying learning optimization methods. The final version enables users to generate tailored strategies for specific situations, access educational capsules, and save personalized strategies. Participants reported improved clarity, usability, and relevance.

Conclusions: This participatory action research enabled the codevelopment and validation of the AMI web application to support the operationalization of learning optimization methods for people living with an MNCD by personalizing them to the situations encountered daily as well as to the underlying cognitive difficulties. By involving caregivers, professionals, and users, this process contributed to the development of a web application that meets user needs and their appreciation.

(JMIR Aging 2026;9:e76543) doi:10.2196/76543

XSL•FO
RenderX

KEYWORDS

web application; major neurocognitive disorder; family caregivers; intervention; learning-optimizing methods; qualitative research

Introduction

Major neurocognitive disorders (MNCDs) frequently lead to difficulties in performing activities [1]. These difficulties can be attributed particularly to the deterioration of certain cognitive functions, which hinders the learning of new tasks or the maintenance of known tasks [2]. These difficulties increase as the MNCD progresses, ranging from difficulties in performing complex activities such as managing finances and planning meals to activities of daily living such as hygiene and eating. However, procedural memory and implicit retrieval processes are relatively well preserved compared to other cognitive functions, enabling the continuation of automatic, well-integrated sequences of actions in everyday life. Learning methods that build on these preserved capacities and compensate for the cognitive difficulties arising from MNCDs have been shown to be effective in maintaining the ability to perform activities [2,3].

Several studies have shown that people living with an MNCD benefit from the use of learning optimization methods from cognitive rehabilitation, such as error-free learning, motor encoding, spaced retrieval, and fading, which promote the safe pursuit of their meaningful activities [3-8]. These learning optimization methods provide successful experiences by ensuring that the right challenge is offered [3], thus helping to foster their engagement in the pursuit of meaningful activities and increasing the sense of competence of the caregiver of a person living with an MNCD [7]. These learning optimization methods rely on the cognitive capabilities generally preserved in people living with MNCDs, namely procedural memory and implicit retrieval processes [9]. These methods focus on the automation of correct action sequences during learning (errorless learning and motor encoding [10,11]), the gradually spaced repetition of these sequences (spaced retrieval [12,13]), and the progressive reduction of motor, visual, and verbal cues provided during their performance (vanishing cues [14]). To integrate these learning optimization methods, the involvement of a caregiver or an intervener is necessary to ensure consistency in the sequence of activities and to enable learning [15]. Since the integration of these methods, supported by the caregiver, supports the pursuit of meaningful activities, it fosters a positive relationship within the dyad, which may explain why the introduction of these methods does not increase caregiver burden [16].

However, while the principles of learning optimization methods are relatively straightforward, the personalized application of these methods to specific situations encountered can be more difficult [12,17]. In addition to considering the person's specific cognitive difficulties (eg, difficulty sequencing tasks, forgetting to eat, or missing appointments), caregivers must also take into account the person's habits and daily routine [16]. It is therefore essential to offer support or tools to caregivers to help them apply learning optimization methods to the specific situations experienced daily by the person living with an MNCD [17]. Support and tools for caregivers in the personalized application of learning optimization methods could optimize the pursuit of meaningful activities by people living with an MNCD by promoting the use of their preserved abilities.

Many support resources and tools are available for caregivers of people living with MNCDs. According to a recent comprehensive meta-review of systematic reviews and meta-analyses, many face-to-face, nonpharmacological interventions are beneficial for caregivers [18]. This category of intervention includes psychoeducation, psychotherapy, occupational therapy, mindfulness interventions, and multicomponent interventions (eg, combining psychoeducation and psychotherapy components). Positive effects on depression, quality of life, sense of mastery, and communication skills have been attributed to various nonpharmacological interventions [18]. Although face-to-face interventions are effective, caregivers do not always use them and often find them difficult to access. This can be due to challenges such as being unable to leave the person alone, difficulty traveling, believing they do not need help, reluctance to seek services, limited knowledge of available resources, lack of time, or concerns about sharing personal information about their loved one [19].

To reduce some of these barriers, studies recommend that technological tools could make certain types of support more accessible to caregivers [20]. Among these, online interventions are effective and cost-effective and could favor accessibility [21,22]. A systematic review of online support interventions (eg, websites) for caregivers, including 12 studies [23], showed that these interventions can have positive effects on caregivers' well-being. Among the interventions included in this systematic review, those that combine information, personalized strategies, and contact with other caregivers could increase caregivers' knowledge and skills for caring for a loved one living with an MNCD, their confidence and self-efficacy, and reduce their stress, burden, and depressive symptoms [24-29]. Online interventions can easily incorporate a variety of formats (eg, videos, graphics, text), as well as support (eg, specific recommendations) [30]. The use of health information technologies can be relevant to family caregivers if adapted [31]. These technologies make it possible to reach isolated and stigmatized groups, reduce health care costs, and access information and advice at any time [32]. Among online support interventions, personalized ones (based on responses to a questionnaire or feedback following the intervention) show greater positive effects than nonpersonalized intervention [23,33,34].

In addition to online interventions, several technology applications (eg, sensors, wristwatches, cameras, voice recorders, calendars, automatic night lamps, gas cooker devices, picture button telephones, alarms, and electronic pillboxes) can support caregivers and people living with MNCDs [35]. In their scoping review, Huisman et al (2022) found that these technology applications help to reduce the burden and social isolation of caregivers and consequently promote better sleep. Technology applications for caregivers are mainly designed to

XSL•FO
RenderX

provide information and resources, offer general practical tips for frequently encountered problems (eg, managing behavioral problems or medication), or help coordinate care among multiple caregivers (eg, sharing calendars or to-do lists) [36].

Although online interventions and technology applications have shown benefits and have many advantages when compared with face-to-face interventions, caregivers may find it difficult to personalize the content and translate the information provided into concrete strategies that fit specific everyday situations [37]. For example, the transposition of methods based on analysis of the situation encountered can represent a challenge [17]. Thus, to our knowledge, no tool that enables caregivers to learn how to use learning optimization methods to support the realization of meaningful activities by building on the person's preserved abilities.

A previous study from our team highlighted the difficulty caregivers had in appropriating learning optimization methods, as well as their need for support tools to use these methods [38]. Based on this previous study, a knowledge mobilization study was carried out to codevelop a variety of support tools, including a web application, which is the scope of the present article. These support tools were designed to help caregivers learn, understand, and use learning optimization methods (eg, error-free learning). More precisely, the project aimed to support caregivers in adapting these methods to specific situations encountered daily and underlying difficulties to help people living with an MNCD maintain autonomy in meaningful activities by building on preserved abilities. This article describes the codevelopment and iterative steps involved in the design and validation process of a web application called "Aide-Mémoire-Interactif (AMI)," a support tool developed as part of this knowledge mobilization study.

Methods

Participatory Action Research Approach

To design the AMI web application, participatory action research based on a codevelopment and validation process was carried out. In the health care field, the codevelopment of such a web application requires a partnership among health care professionals, the people who will use the application, and the team of researchers and designers [39]. Caregivers, health and social service professionals, and workers from community organizations providing support to caregivers were involved in identifying situations frequently encountered by people living with an MNCD [16] and throughout the process of codeveloping and validating the AMI web application. The stages of this codevelopment were overseen by at least 1 coinvestigator with expertise in the stage in question and the principal investigator. For example, during the computer coding stage to operationalize the application, 2 project coinvestigators with expertise in this field supervised the computer programming work, and the principal investigator took part in all meetings with these coinvestigators and the computer scientist to ensure that user needs were met. The various stages leading to the codevelopment of the AMI web application can be grouped into 3 phases: (1) assessment of caregivers' needs (identification of situations frequently encountered by people living with an MNCD) [16]; (2) production of appropriation tools, including the first version of the AMI web application [40]; and (3) validation of the tools, leading to the final version of the AMI application. Caregivers, health and social service professionals, and workers from community organizations experimented with the first version of the application during workshops with an occupational therapist and have made suggestions that contributed to revisions as the codevelopment progressed.

Participants, Recruitment, and Data Collection

For the first stage of the project (ie, the identification of frequently encountered situations), 20 participants (caregivers, health and social service professionals, and community organization workers) were targeted to obtain a diversity of viewpoints and application contexts and data saturation [41]. Participants were recruited by one of the coinvestigators, a practitioner from the partner community organization. An interview guide for individual semistructured interviews (preplanned duration of approximately 60 min) was developed to identify meaningful activities and lived situations in which the cognitive difficulties of people living with an MNCD cause difficulties and would thus benefit from learning optimization methods. The results of this stage are reported in a previously published article [16].

For the 2 subsequent stages (ie, codevelopment and validation of the support tools), a target number of 8 to 10 participants was set for the learning optimization methods appropriation workshops. To ensure diversity of viewpoints in terms of application contexts while keeping the workshop group small enough to promote interactions, 3 participant groups were targeted: family caregivers, stakeholders from community organizations working with family caregivers, and occupational therapists in health and social services who work with people living with an MNCD in a variety of daily situations. A purposive sampling strategy was used to recruit women and men from different age groups and whether they were cohabiting or not with a person living with an MNCD. For family caregivers, inclusion criteria were (1) providing day-to-day support to a person living with an MNCD and (2) not having an MNCD as perceived by the referring caregiver. For occupational therapists and caregivers from community organizations, to meet the inclusion criteria, they had to work with people living with MNCDs and their caregivers. Occupational therapists were recruited by the head of the home care department based on their knowledge of home care issues. The coresearcher from the partner organization (Société Alzheimer des Maskoutains–Vallée des Patriotes) recruited community workers and caregivers by contacting caregivers who attended the organization's activities and who, he felt, could contribute to the development of the appropriation tools by experimenting with and critiquing them.

Production of Appropriation Tools

Several tools for appropriating learning methods were designed during the project, including workshops for caregivers [34], the AMI web application, companion leaflets, and video capsules showing examples of putting learning optimization methods into action, to support their operationalization in specific situations and contexts encountered. This study describes more

XSL•FO
RenderX

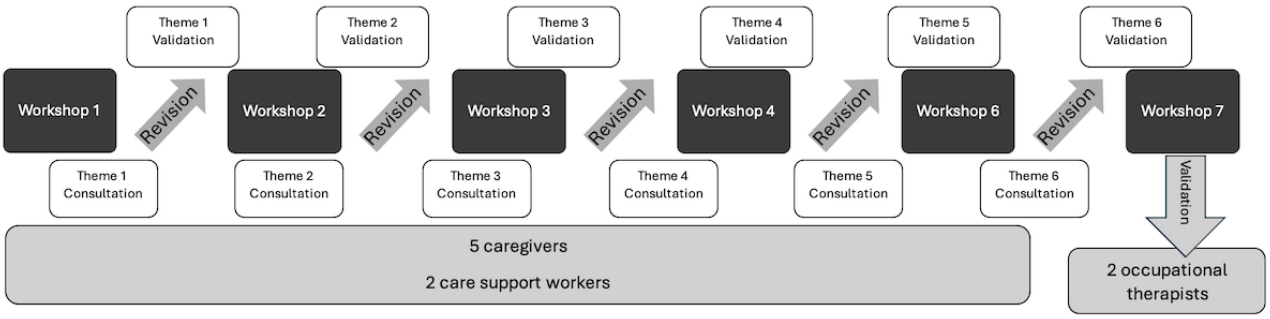
specifically the codevelopment of the AMI application using a software development process based on the Agile Scrum methodology [42]. The Agile Scrum method is a project management methodology that follows an iterative process with users to adjust to changing objectives and needs. It is divided into short cycles, called "sprints," which enable features to be developed progressively while adjusting priorities based on feedback from stakeholders (caregivers, occupational therapists, stakeholders from partner organizations, coresearchers, and research assistants). This methodology requires teams with the necessary complementary expertise and an operating structure based on defined roles to ensure communication and ongoing project progress. The design is therefore based on principles of user friendliness and accessibility to provide a practical and effective solution for specific contexts [43]. To promote accessibility, the application is accessed via a web browser, making it available on different devices and operating systems. Moreover, platform independence reduces development effort and costs involved in supporting multiple user installations. Finally, updates and the addition of new functionalities are possible, enabling the application to evolve in line with emerging needs.

The Agile Scrum methodology chosen was built around 4 elements targeted by the team. First, an ordered list of features for the AMI application was created based on the developed logic algorithm and usage requirements. Second, planning of the functionalities to be programmed was divided into subtasks per 2-week period. Third, daily meetings between the programmer and 2 coinvestigators (an occupational therapist and a computer scientist) were held to identify problems and find solutions, which were implemented by the programmer. Finally, demonstrations of developed functionalities and reviews of the latest iteration were carried out by the programmer and the 2 coinvestigators during the first workshop with caregivers and occupational therapists, enabling participants to experiment with the AMI application, criticize it, and validate the changes made as they experimented during the series of workshops.

Tool Validation

A first version of the AMI web application was designed by the research team, followed by several feedback loops aimed at adjusting the application according to the comments received (Figure 1). Before the experimentation by caregivers, the application was tested by members of the research team and representatives of partner organizations. This allowed us to clarify some terms and add some answer choices. During the trial period, caregivers participating in the workshops provided feedback during demonstrations at each of the workshops and by trying out the application for themselves at home with their loved one afterwards [36]. After each workshop on a specific theme (eg, memory or executive function difficulties), feedback on the use of the AMI web application was collected, and modifications were made on an ongoing basis. For example, caregivers were asked to provide specific situations they wanted to be included in the app. For each situation, they answered the questions asked by the AMI web application and reviewed the suggested strategy. Participants were invited to provide feedback on various aspects of the app, including the variety of situations or suggestions for additional or missing ones, answer choices that could be missing for specific situations, cognitive difficulties encountered for a specific situation, and the actions included in the learning sequence. Modifications were then validated with participants during the following workshops, leading to the final version of the application after the workshop series. This approach, which included several iterations, made it possible to quickly adjust the development according to needs. Each caregiver had the opportunity to provide suggestions for the application based on specific situations they had encountered. The approach of valuing diversity and the richness of complementary or divergent points of view was presented right from the first workshop. It was specified that everyone's experience was essential to the codevelopment of a relevant web application. Opposing comments were discussed as a group to identify a solution that would meet the needs of all participants. The principal researcher, who cofacilitated the workshops, took note of the participants' comments and then discussed their implementation with the programmer and the information technology coresearcher.

Figure 1. Codevelopment and validation process of the Aide-Mémoire-Interactif application.



Ethical Considerations

This study was approved by the CIUSSS de l'Estrie-CHUS Research Ethics Committee (MP-31-2019-3211). Written informed consent was obtained from all participants. To ensure privacy, any information that could lead to identification was removed. Information gathered was kept confidential. No financial compensation was given to the participants.

XSL•FO
RenderX

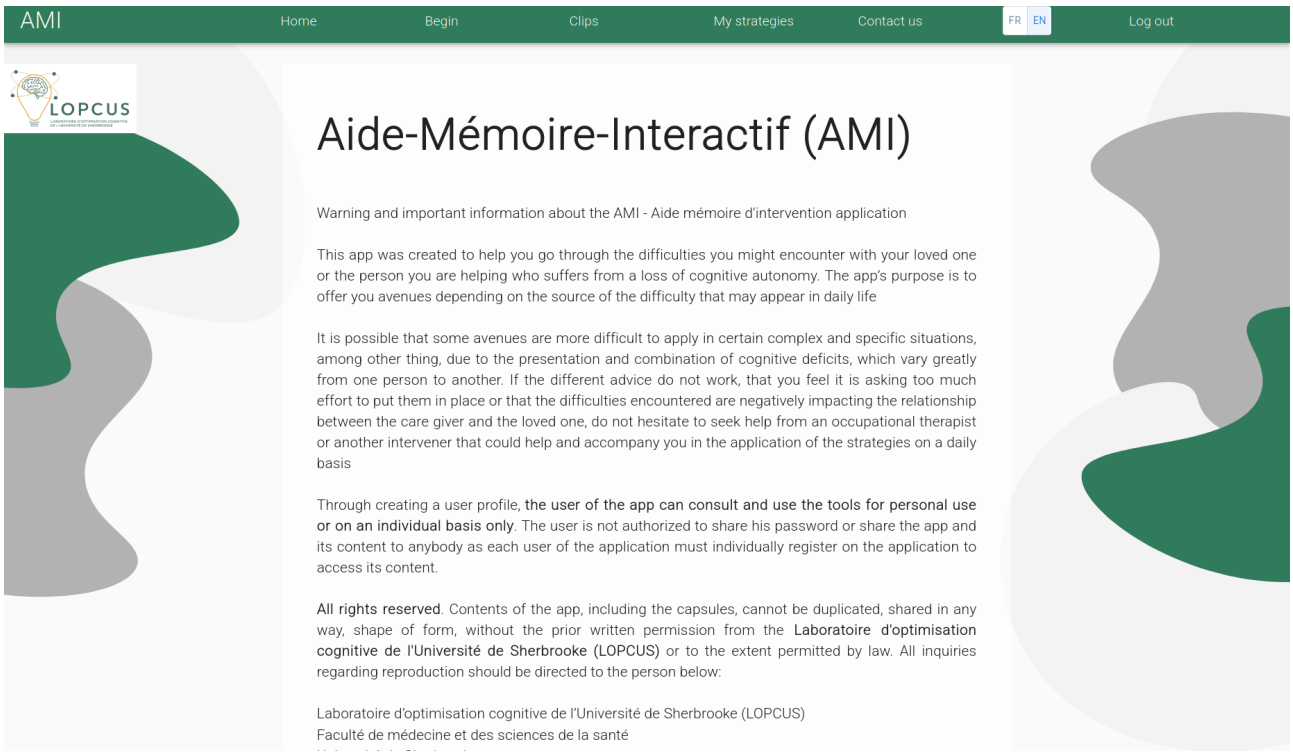
Results

Process and Participants

This section describes the different steps that led to the development of the AMI web application and focuses on 2 complementary aspects: the participatory process through which the application was developed and the adjustments made to the tool based on users' feedback. These elements illustrate how the final version of the application, which is described in detail, emerged through iterative refinement.

For the first part, caregivers (n=11), health and social service professionals (n=7; social workers and occupational therapists), and community organization workers (n=7) participated in identifying situations for which the AMI web application should support the operationalization of learning optimization methods. For the subsequent codevelopment and validation of the AMI web application, caregivers (n=5), caregiver support workers from community organizations (n=2), and occupational therapists working with people living with MNCDs (n=2) were recruited. The application was designed to be easy to navigate for users who are mainly caregivers or health care professionals working with people living with an MNCD [44]. The application home page (Figure 2) provides summary information on navigation, its various functionalities, and contact details for training in its optimal use. Indeed, the AMI web application, designed to support the operationalization of learning optimization methods, requires a basic understanding of the difficulties present in neurocognitive disorders and of these methods.

Figure 2. Aide-Mémoire-Interactif web application home page.



Codevelopment of the AMI Web Application: Logical Architecture

Technological choices were made to ensure that the web application would be easy to adapt to the realities of partner environments. Thus, the logical architecture was built according to the rationale underlying learning optimization methods and is structured around the following questions: considering the specific situation in which the difficulty occurs, (1) What is the difficulty that can explain it? (2) What is the desired action? (3) After which action should the desired action be done? (4) What action is performed instead of the desired action? The application was codeveloped based on real-life situations requiring the use of learning optimization methods [39]. Question 1 asks the caregiver to identify, based on their perception, the primary cognitive difficulty that may explain why learning optimization methods are needed for this specific situation. The cognitive difficulty guides how the 4 learning optimization methods are combined to rely on preserved abilities. Question 2 requires the caregiver to analyze the sequence of actions, identify the usual error, and specify which action should be learned through motor encoding to replace this error. Question 3 also draws on the analysis of the action sequence and, in line with errorless learning principles, identifies the action that immediately precedes the error as well as the action that needs to be integrated using learning optimization methods. Finally, question 4 is essential to prevent the error before it occurs, supporting the errorless learning method. The strategy suggested by AMI includes a progression that integrates fading and spaced retrieval methods.

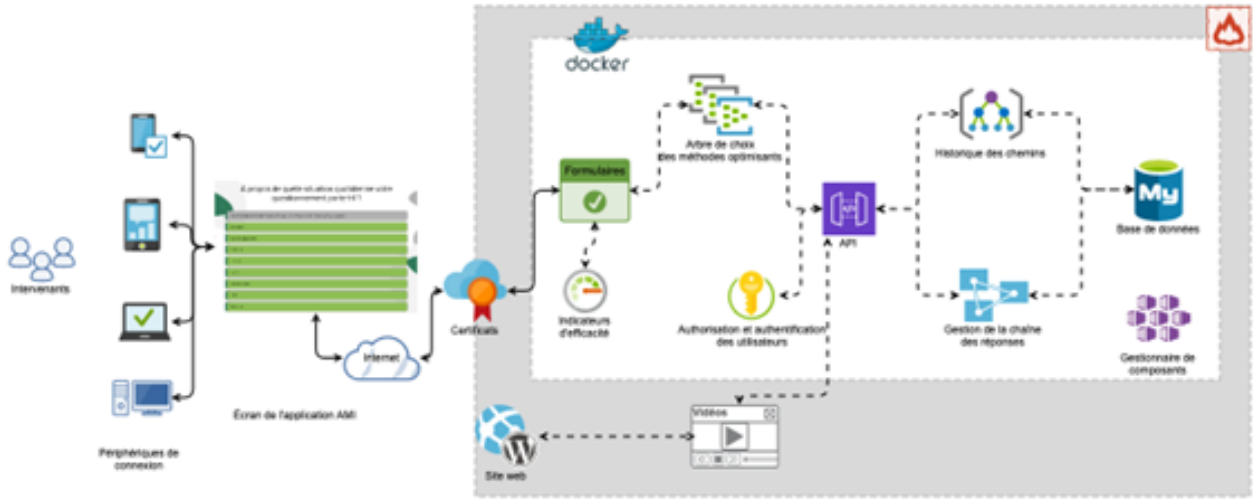
Being user-centered, the AMI web application is built to offer each user dynamic interaction with the generated content so that the strategies generated are personalized to the situation encountered and the specific context. To optimize accessibility, the application is accessed using a web browser via a computer, tablet, or smartphone. As illustrated in Figure 3, AMI is

XSL•FO
RenderX

organized around modular, adaptable software components comprising the following:

- a server secured by an SSL (Secure Sockets Layer) certificate, enabling access to data when the user interacts with the application, regardless of the device used;
- an API (application programming interface) that offers flexibility to integrate new components based on needs expressed by users during validation and scaling in different types of environments and with people with different levels of impairment;
- a Docker-type application virtualization environment, enabling the application to be used on any device;
- a MySQL database to store data for later access to customized strategies;
- a centralized Keycloak authentication and authorization component to facilitate user management [39] and ensure access to personalized strategies;
- an HTML interface for visualizing situation and context information entered by users;
- a ReactJS framework for rapid updating of suggested strategies based on user input;
- a high-performance, open-source FastAPI engine, enabling development, particularly with the Python language, to facilitate integration of new components and easily evolve the application in line with user needs.

Figure 3. Logical architecture diagram of the Aide-Mémoire-Interactif web application.



The application is composed of scripts written in several programming languages. This plurality of programming languages allows greater modularity based on the requirements of each tool for programming learning methods. A 4-GB RAM server with 4 cores covers all application needs.

Validation of the Tools: Perceptions and Feedback From Participants

As the workshops progressed, participants provided feedback that directly informed several modifications to the tool. Their suggestions mainly concerned the following: (1) formulation of everyday situations to make them easier to understand and choose from; (2) formulation of the difficulties underlying the problems in everyday situations to make them easier to understand and choose from; (3) addition of frequently encountered answer choices for desired actions and actions performed in place of desired actions; and (4) visual presentation for navigating questions, choices, and proposed strategies.

Following systematic verification of the vocabulary used in each situation category by participants, clarifications were made to facilitate understanding. Regarding the difficulties underlying the problems in everyday situations, certain categories of cognitive difficulties initially established were subdivided into 2 distinct categories to facilitate understanding and more direct application to the difficulties encountered. Regarding desired actions and actions performed in place of the desired action, additions were made to the answer choices to cover a wider range of contexts for operationalization in everyday situations. Additions were made to the list of answer choices, and if none of the options corresponded to the person's situation, an "other" button led to an answer box for entering the context-specific element subsequently used by the AMI web application for personalized operationalization of learning optimization methods.

Some participants reported that certain dichotomous "yes or no" questions did not fit their situation, so they did not know which option to select. As a result, the "not applicable" option was added. Regarding the visual presentation of the AMI web application, feedback from participants helped to make navigation more intuitive (eg, replacing the "questionnaire" button with "start"). In response to participants' request for the opportunity to save the strategies proposed for a specific situation, a "Favorites" button was created to save the individual history of strategies based on the individual answers provided to the questions. The AMI web application underwent refinements to support its ongoing improvement. Enhancements included the ability to save personalized strategies using the initials of both the caregiver and the care recipient. A printable version of the strategy, with editable sections to tailor intervention sequences, was also implemented. In parallel, visual and technical adjustments were made to optimize system performance, streamline navigation, and improve clarity for

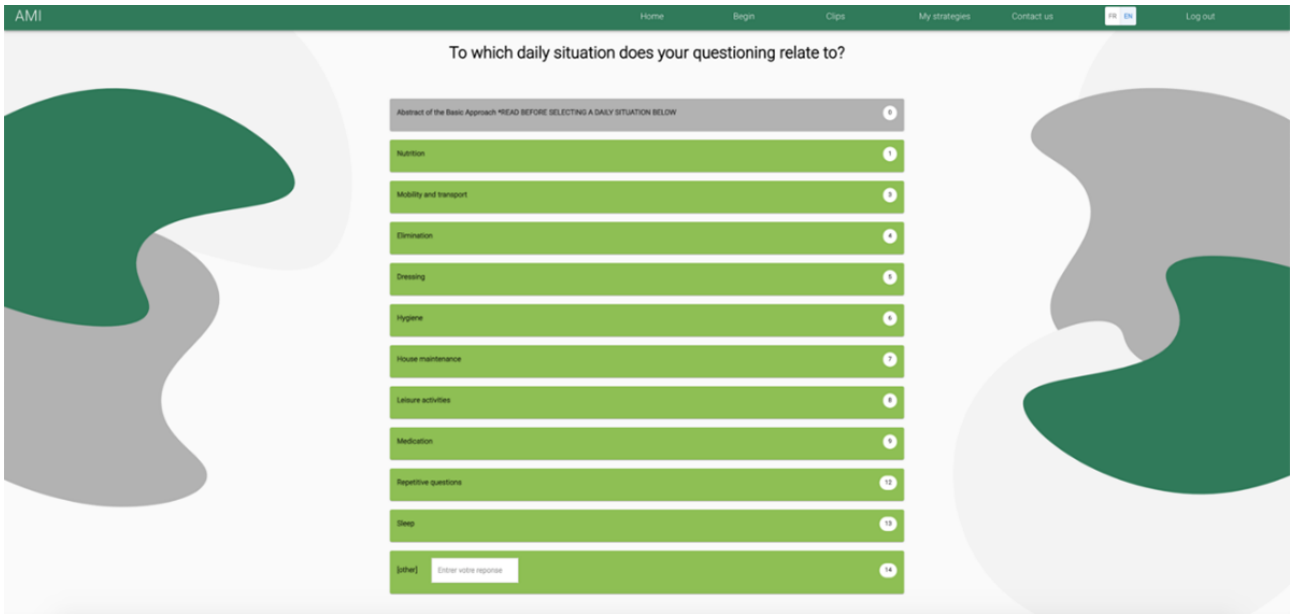
XSL•FO
RenderX

users. Participants confirmed that the latest version, shaped through iterative feedback, better met their needs.

Final Version of the Application: Description of Functionalities

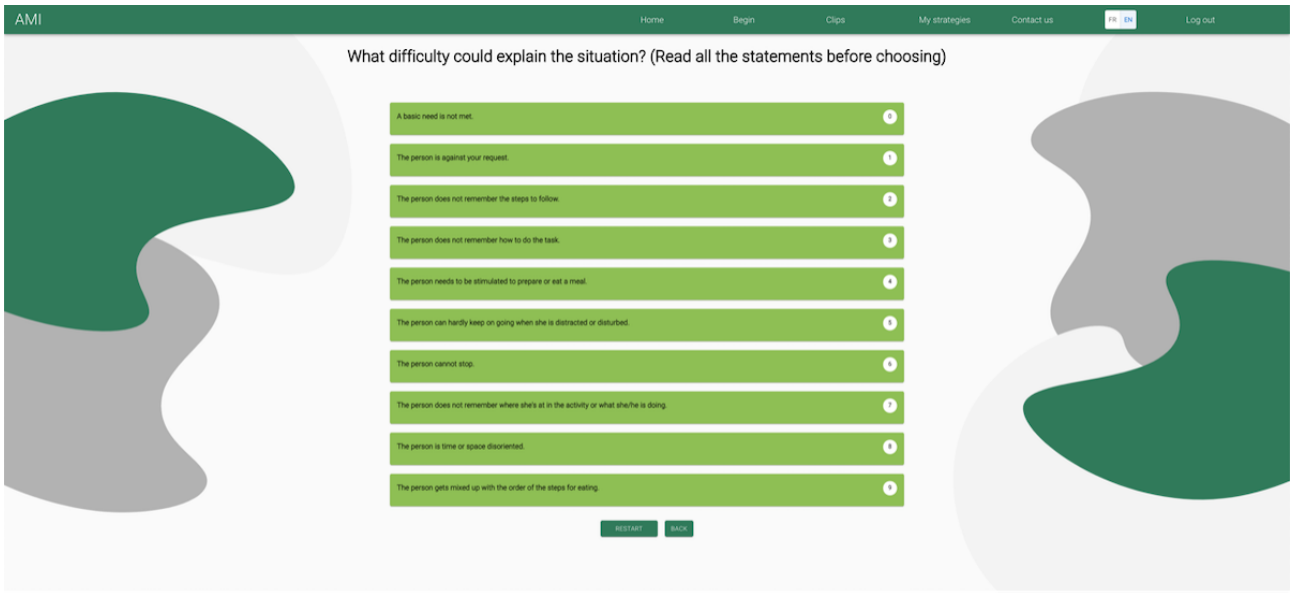
The home page of the final version of the AMI web application lets the user choose among the following: (1) starting navigation to operationalize learning optimization methods in a specific situation, (2) viewing video capsules describing these methods, or (3) consulting strategies saved from previous navigations. By selecting the Start option, users can select a daily situation from a range of possible options (Figure 4).

Figure 4. List of daily situations.



Once a situation is selected, users identify the difficulty that could explain it. For example, for the Nutrition situation, 10 difficulties are presented (Figure 5). For each situation, questions and response options vary depending on the difficulty encountered.

Figure 5. List of difficulties that could explain the situation.



Depending on the answers given about the situation (activity, cognitive difficulties) and the context in which it occurs, the user obtains suggestions for possible solutions to the situation (Figure 6).

XSL•FO
RenderX

Figure 6. Example of a possible solution related to the situation.



Users can obtain a reminder of the theoretical foundations by clicking on Capsules. Twelve capsules are available, covering the following topics: (1) basic needs; (2) approaches (communication, validation, diversion, refusal); (3) methods to promote learning (error-free learning, motor encoding, spaced retrieval, blurring); (4) reminders of how memory, praxis, and executive functions work; (5) preserved abilities; and (6) complementarity with compensatory measures.

Finally, the *My Strategies* option in the drop-down menu includes 2 tabs: Lately, which allows the user to view the most recent sequences he or she has performed, and Favorites, which saves the operationalization of preferred strategies.

Discussion

Main Results

This participatory action research enabled us to codevelop and validate the AMI web application to support the operationalization of learning optimization methods for people living with an MNCD, by personalizing them to situations encountered daily as well as to the underlying cognitive difficulties. It enabled the reformulation of everyday situations and their underlying difficulties, the addition of frequently encountered answer choices for desired actions and actions performed in place of desired actions, and finally a visual presentation facilitating navigation. By involving caregivers [45], professionals, and users, this research contributed to the development of a web application that aligns with their needs and preferences.

Comparison With Previous Studies

To date, a number of applications have been developed specifically for caregivers, with a variety of functions to meet the diverse needs of this population [20,31,32]. For its part, the AMI application is, to our knowledge, the first that facilitates the use of learning optimization methods by caregivers and that proposes solutions adapted to the specific context of everyday situations by promoting the use of the preserved abilities of people living with an MNCD. Indeed, when caregivers find themselves in a situation where they do not know how to deal with the impact of MNCDs on the performance of daily activities, the AMI web application provides them, after a series of questions on the specific context, with a strategy adapted to the problematic activity, the underlying cognitive difficulty, and the context in which it is performed [16]. Considering that several negative effects on physical and mental health can be caused by the role of caregiver, such as chronic stress and depression [46], feelings of burden [47], and poor-quality sleep [48], the possibility of having a web application that supports their role can help foster a positive aspect of the caregiver role.

The difficulties and activities performed by people living with MNCDs are highly variable, as are the coping skills of their caregivers, justifying the relevance and superior effectiveness of personalized interventions [23]. Warmoth et al [49] developed a personalized cognitive rehabilitation intervention and found that people living with MNCDs and their caregivers appreciated the personalized approach to intervention and that the flexibility allowed by this approach played an important role in their appropriation of learning optimization methods. Personalization is essential when using learning optimization methods, given the difficulty caregivers have in transposing these methods to

XSL•FO
RenderX

another context [17]. The inclusion of users in studies aimed at developing services offered to them is essential [50] so that they can contribute to improving the products that meet their needs [51]. In the present study, the involvement of caregivers at all stages of the research enabled the application to be developed in line with the situations they encounter daily. In this way, the AMI application fills a gap for caregivers, enabling them to support the pursuit of day-to-day activities of people living with MNCDs.

Although the application is adapted to the situations encountered, studies report that lack of time and low digital literacy can limit the use of web applications by caregivers [52,53]. Since it can be difficult for some caregivers to find information [54], easy navigation and recording of information can be supportive. Clarity of language is also essential [52]. This participatory action research, which led to the codevelopment of the AMI application and involved seniors at every stage, enabled us to take these potential issues into account and make the necessary adjustments based on feedback from caregivers, mainly by adding the My Strategies feature to enable easy retrieval of personalized strategies.

Finally, studies tend to show that peer support and psychosocial support are also important elements in the interventions offered to caregivers [55]. Although the AMI web application was created to meet another equally important need of caregivers, the workshops held to help them appropriate the web application enabled several caregivers in similar situations to interact. This enabled them to share their experiences and possibly feel less alone, a feeling often experienced by people in this role [54]. In fact, 1 caregiver had recommended the addition of a function that would enable caregivers to obtain psychosocial support or communicate with other caregivers. Due to the extent of the research and programming required, which went beyond the primary objective of codevelopment, this recommendation could not be integrated into the current version of the application, despite its great relevance. These additional functionalities would be relevant given that the mobile web applications currently available for caregivers rarely address their personal needs, such as their health, financial security, or the problems they may encounter at work [56].

It would also fill a gap in web applications that currently do not offer support for caregivers' emotional well-being, notably for stress and anxiety management, as well as depression [23-25,29].

Strengths and Limitations

One of the main strengths of this study is the codevelopment process to ensure that the application meets needs and is user-friendly [37]. All types of potential users contributed to the codevelopment, including health care professionals and caregivers who worked in partnership with the team of researchers and designers (occupational therapy, social work, leisure, and information technology). Caregivers and health care professionals were involved throughout the process, from the definition of needs to the final validation of the application. Codevelopment ensures that the needs, preferences, and functionalities required by users are considered, and that the result is finally tailored to them [56]. Participants were caregivers and health care professionals mainly involved in home care. Although some have experience of intervention in a residential setting, the application answer choices are mainly based on situations occurring at home, in the community. Caregivers and interveners of people living with MNCDs in residential settings whose situations are not included in the application's answer choices can, however, choose the *Other* option for the various questions and enter the contextual elements specific to their situation to obtain a personalized strategy proposal. These situations have not yet been validated. Although the number of caregivers involved was small, the multiple iterations allowed for the inclusion of a wide variety of situations. Future studies may adapt the application to other practice contexts (eg, hospital settings) or living environments (eg, residential resources offering long-term care) for people living with MNCDs. Although recruitment by reasoned choice targeted a diversity of experience, as caregivers and care providers were linked to a community organization and to health care network stakeholders working in collaboration with this organization, cultural representativeness and socioeconomic status were not present. More, since the caregivers were recruited targeting people who would be critical for a web application, since they had good digital literacy, they are not representative of all caregivers who may have more difficulty using technological tools. During implementation, introductory tools will have to consider the low digital literacy of some caregivers to guide their use of the web application. Although caregivers shared positive comments regarding the usefulness, accessibility, and effects of the AMI web application on autonomy and quality of the caregiving relationships, these effects were not assessed. Future studies should therefore assess the effects and usefulness of the AMI web application for caregivers and health professionals in supporting the performance of activities using the preserved abilities of people living with MNCDs. Future studies could evaluate the application's implementation and its effects on caregivers, health professionals, and people living with an MNCD.

Recommendations and Avenues of Research

The aim of this study was to describe the codevelopment process of the AMI web application, which supports the operationalization of learning optimization methods using the preserved abilities of people living with MNCDs. Suggestions received from caregivers and health care professionals have enabled us, in an iterative process, to improve the application so that it meets the specific needs of using preserved abilities in daily activities, thus compensating for difficulties encountered at home. Considering the physical and mental impacts of the caregiver role mentioned above, it would also be relevant to measure whether daily use of the AMI application has an impact on caregivers, particularly on their sense of burden. In the same way, the impact on the person living with an MNCD would also be interesting to assess. Second, it would also be relevant to look at the prolonged use of the AMI application to assess the sustainability of its use daily and to evaluate its long-term effectiveness [57], particularly in view of the progression of difficulties in daily activities attributable to the evolution of MNCD. Finally, as mentioned, additions to the web application would be relevant to make in future studies, such as offering peer support and psychosocial support via the application [53,58]

XSL•FO
RenderX

and adding features that would support the personal needs of caregivers [55].

Conclusions

This study details the steps that led to the development of the AMI web application, which enables caregivers to appropriate learning optimization methods for the specific situations encountered daily. The AMI application stands out for its ability to be customized to everyone's context and for its codevelopment process, which involved caregivers, health care professionals, community workers, and researchers with expertise in the field. The AMI application was developed to provide a practical, high-performance solution for managing interventions. Its user-friendly interface, flexible web architecture, and advanced functionalities make it a valuable tool for professionals working in this specific field. What's more, the technology used to develop the AMI web application has enabled optimal scalability and performance, guaranteeing a fluid user experience even with many simultaneous users. The chosen methodology favors the development of a web application that meets the needs of all those involved and thus optimizes its day-to-day use.

Acknowledgments

The research team thanks caregivers and support workers who generously contributed to the codevelopment of the Aide-Mémoire-Interactif (AMI) application by sharing examples of their experience to optimize the application's response to their needs. The team also thanks the Ordre des ergothérapeutes du Québec for its involvement and pertinent comments. A special thank you to the DOMUS team for its information technology contribution, through the contribution of several people with specific expertise at different stages of codevelopment.

Authors' Contributions

Viscogliosi, C: wrote the proposal, conceptualize the prototype, collaborate with informatic team for the development, writing of the article, coordinate: data collection, analysis; Kenfack Ngankam, H: participated to the supervision of informatics students, participated to article writing; Duquette-Laplante, R: participated in data analysis and the writing of the article; Rahimaly, S: participated in data analysis and the writing of the article; Pigot, H: participated in the conception of the proposal, participated to the supervision of informatics students; Chassé, B: participated in the conception of the proposal, recruitment, data collection, conceptualization of the prototype; Ait Bouziaren, S: participated in the informatic programming of the web application; Couturier, Y: participated in the conception of the proposal, participated to the conceptualization of the web application; Déry, J: participated in the conception of the proposal and conceptualization of the web application; Biard, N: participated in informatic programming of the web application; Duchesne, J: participated in data analysis; Giroux, D: participated in the conception of the proposal and the conceptualization of the web application; Jacob, S: participated in the conceptualization of web application and data analysis; Mino-Roy, J: participated in the conceptualization of web application and literature review; Caron, M: participated in data collection; Provencher, V: participated in the conception of the proposal and to the conceptualization of web application

Conflicts of Interest

None declared.

References

1. Allison TA, Gubner JM, Oh A, et al. Meaningful activities and sources of meaning for community-dwelling people living with dementia. J Am Med Dir Assoc 2022 Jul;23(7):1191-1196. [doi: 10.1016/j.jamda.2021.08.009] [Medline: 34481793]
2. Van der Linden M, Juillerat AC. La revalidation neuropsychologique dans la maladie d'Alzheimer à un stade précoce: principes, méthodes et perspectives [Article in French]. Rev Neurol (Paris) 2004 Apr;160(4):64-70. [doi: 10.1016/S0035-3787(04)70945-5]
3. Crowe J, Gabriel L. Errorless learning and spaced retrieval training for clients with Alzheimer's dementia. Phys Occup Ther Geriatr 2013 Sep;31(3):254-267. [doi: 10.3109/02703181.2013.796037]
4. Kudlicka A, Martyr A, Bahar-Fuchs A, Sabates J, Woods B, Clare L. Cognitive rehabilitation for people with mild to moderate dementia. Cochrane Database Syst Rev 2023 Jun 29;6(6):CD013388. [doi: 10.1002/14651858.CD013388.pub2] [Medline: 37389428]
5. Hildebrandt H. Treatment of severely impaired, amnestic patients. In: Cognitive Rehabilitation of Memory: Elsevier, Vol. 2019:161-177. [doi: 10.1016/B978-0-12-816981-0.00004-7]
6. Small JA, Cochrane D. Spaced retrieval and episodic memory training in Alzheimer's disease. Clin Interv Aging 2020;15:519-536. [doi: 10.2147/CIA.S242113] [Medline: 32368019]
7. Choi J, Twamley EW. Cognitive rehabilitation therapies for Alzheimer's disease: a review of methods to improve treatment engagement and self-efficacy. Neuropsychol Rev 2013 Mar;23(1):48-62. [doi: 10.1007/s11065-013-9227-4] [Medline: 23400790]
8. Clare L, Kudlicka A, Oyebode JR, et al. Goal-oriented cognitive rehabilitation for early-stage Alzheimer's and related dementias: the GREAT RCT. Health Technol Assess 2019 Mar;23(10):1-242. [doi: 10.3310/hta23100] [Medline: 30879470]

XSL•FO
RenderX

9. De Wit L, Marsiske M, O'Shea D, et al. Procedural learning in individuals with amnestic mild cognitive impairment and Alzheimer's dementia: a systematic review and meta-analysis. Neuropsychol Rev 2021 Mar;31(1):103-114. [doi: 10.1007/s11065-020-09449-1] [Medline: 32897482]
10. Clare L, Wilson BA, Carter G, Breen K, Gosses A, Hodges JR. Intervening with everyday memory problems in dementia of Alzheimer type: an errorless learning approach. J Clin Exp Neuropsychol 2000 Feb;22(1):132-146. [doi: 10.1076/1380-3395(200002)22:1;1-8;FT132] [Medline: 10649552]
11. de Werd MME, Boelen D, Rikkert M, Kessels RPC. Errorless learning of everyday tasks in people with dementia. Clin Interv Aging 2013;8:1177-1190. [doi: 10.2147/CIA.S46809] [Medline: 24049443]
12. Bier N, Provencher V, Gagnon L, Van der Linden M, Adam S, Desrosiers J. New learning in dementia: transfer and spontaneous use of learning in everyday life functioning. Two case studies. Neuropsychol Rehabil 2008 Apr;18(2):204-235. [doi: 10.1080/09602010701406581] [Medline: 18350414]
13. Lipinska B, Bäckman L. Encoding-retrieval interactions in mild Alzheimer's disease: the role of access to categorical information. Brain Cogn 1997 Jul;34(2):274-286. [doi: 10.1006/brcg.1997.0916] [Medline: 9220090]
14. Bier N, Desrosiers J, Gagnon L. Prise en charge cognitive de la mémoire dans le vieillissement normal, les troubles cognitifs légers et la démence de type Alzheimer [Article in French]. Can J Occup Ther 2006 Feb;73(1):26-35. [doi: 10.2182/cjot.05.0004]
15. Clare L, Linden DEJ, Woods RT, et al. Goal-oriented cognitive rehabilitation for people with early-stage Alzheimer disease: a single-blind randomized controlled trial of clinical efficacy. Am J Geriatr Psychiatry 2010 Oct;18(10):928-939. [doi: 10.1097/JGP.0b013e3181d5792a] [Medline: 20808145]
16. Viscogliosi C, Desrosiers J, Belleville S. Optimizing participation of older adults with cognitive deficits post-stroke: types of help and caregiver burden. Can J Aging 2019 Jun;38(2):222-235. [doi: 10.1017/S0714980818000521] [Medline: 30773163]
17. Amieva H, Robert PH, Grandoulier AS, et al. Group and individual cognitive therapies in Alzheimer's disease: the ETNA3 randomized trial. Int Psychogeriatr 2016 May;28(5):707-717. [doi: 10.1017/S1041610215001830] [Medline: 26572551]
18. Cheng ST, Zhang F. A comprehensive meta-review of systematic reviews and meta-analyses on nonpharmacological interventions for informal dementia caregivers. BMC Geriatr 2020 Apr 15;20(1):137. [doi: 10.1186/s12877-020-01547-2] [Medline: 32293325]
19. Brodaty H, Thomson C, Thompson C, Fine M. Why caregivers of people with dementia and memory loss don't use services. Int J Geriatr Psychiatry 2005 Jun;20(6):537-546. [doi: 10.1002/gps.1322] [Medline: 15920707]
20. Hu C, Kung S, Rummans TA, Clark MM, Lapid MI. Reducing caregiver stress with internet-based interventions: a systematic review of open-label and randomized controlled trials. J Am Med Inform Assoc 2015 Apr;22(e1):e194-e209. [doi: 10.1136/amiajnl-2014-002817] [Medline: 25125686]
21. O'Connor MF, Arizmendi BJ, Kaszniak AW. Virtually supportive: a feasibility pilot study of an online support group for dementia caregivers in a 3D virtual environment. J Aging Stud 2014 Aug;30:87-93. [doi: 10.1016/j.jaging.2014.03.001] [Medline: 24984911]
22. Blom MM, Bosmans JE, Cuijpers P, Zarit SH, Pot AM. Effectiveness and cost-effectiveness of an internet intervention for family caregivers of people with dementia: design of a randomized controlled trial. BMC Psychiatry 2013 Jan 10;13(1):17. [doi: 10.1186/1471-244X-13-17] [Medline: 23305463]
23. Leng M, Zhao Y, Xiao H, Li C, Wang Z. Internet-based supportive interventions for family caregivers of people with dementia: systematic review and meta-analysis. J Med Internet Res 2020 Sep 9;22(9):e19468. [doi: 10.2196/19468] [Medline: 32902388]
24. Boots LMM, de Vugt ME, van Knippenberg RJM, Kempen G, Verhey FRJ. A systematic review of Internet-based supportive interventions for caregivers of patients with dementia. Int J Geriatr Psychiatry 2014 Apr;29(4):331-344. [doi: 10.1002/gps.4016] [Medline: 23963684]
25. Beauchamp N, Irvine AB, Seeley J, Johnson B. Worksite-based internet multimedia program for family caregivers of persons with dementia. Gerontologist 2005 Dec;45(6):793-801. [doi: 10.1093/geront/45.6.793] [Medline: 16326661]
26. Fowler CN, Haney T, Lemaster M. Helping dementia caregivers through technology. Home Healthc Now 2016 Apr;34(4):203-209. [doi: 10.1097/NHH.0000000000000372] [Medline: 27023296]
27. Lewis ML, Hobday JV, Hepburn KW. Internet-based program for dementia caregivers. Am J Alzheimers Dis Other Demen 2010 Dec;25(8):674-679. [doi: 10.1177/1533317510385812] [Medline: 21131674]
28. Marziali E, Garcia LJ. Dementia caregivers' responses to 2 Internet-based intervention programs. Am J Alzheimers Dis Other Demen 2011 Feb;26(1):36-43. [doi: 10.1177/1533317510387586] [Medline: 21282276]
29. Gitlin LN, Belle SH, Burgio LD, et al. Effect of multicomponent interventions on caregiver burden and depression: the REACH multisite initiative at 6-month follow-up. Psychol Aging 2003 Sep;18(3):361-374. [doi: 10.1037/0882-7974.18.3.361] [Medline: 14518800]
30. Godwin KM, Mills WL, Anderson JA, Kunik ME. Technology-driven interventions for caregivers of persons with dementia: a systematic review. Am J Alzheimers Dis Other Demen 2013 May;28(3):216-222. [doi: 10.1177/1533317513481091] [Medline: 23528881]

XSL•FO
RenderX

31. Fischer SH, David D, Crotty BH, Dierks M, Safran C. Acceptance and use of health information technology by community-dwelling elders. Int J Med Inform 2014 Sep;83(9):624-635. [doi: 10.1016/j.ijmedinf.2014.06.005] [Medline: 24996581]
32. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006 Jun 23;8(2):e10. [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
33. Núñez-Naveira L, Alonso-Búa B, de Labra C, et al. UnderstAID, an ICT platform to help informal caregivers of people with dementia: a pilot randomized controlled study. Biomed Res Int 2016;2016:5726465. [doi: 10.1155/2016/5726465] [Medline: 28116300]
34. Williams KN, Perkhounkova Y, Shaw CA, Hein M, Vidoni ED, Coleman CK. Supporting family caregivers with technology for dementia home care: a randomized controlled trial. Innov Aging 2019 Jul;3(3):igz037. [doi: 10.1093/geroni/igz037] [Medline: 31660443]
35. Huisman C, Huisman E, Kort H. Technological applications contributing to relieve care burden or to sleep of caregivers and people with dementia: a scoping review from the perspective of social isolation. Front Public Health 2022;10:797176. [doi: 10.3389/fpubh.2022.797176] [Medline: 35425752]
36. Brown AF, Ma GX, Miranda J, et al. Structural interventions to reduce and eliminate health disparities. Am J Public Health 2019 Jan;109(S1):S72-S78. [doi: 10.2105/AJPH.2018.304844] [Medline: 30699019]
37. Fox BI, Felkey BG. Identifying and addressing technology challenges among older adults. Hosp Pharm 2014 Sep;49(8):780-781. [doi: 10.1310/hpj4908-780] [Medline: 25477604]
38. Viscogliosi C, Dame N, Duquette-Laplante R, et al. Accompany caregivers to optimize learning with people living with a major neurocognitive disorder: a participatory action research. Educ Gerontol 2024 Nov;50(11):1029-1044. [doi: 10.1080/03601277.2024.2371529]
39. Ward ME, De Brún A, Beirne D, et al. Using co-design to develop a collective leadership intervention for healthcare teams to improve safety culture. Int J Environ Res Public Health 2018 Jun 5;15(6):1182. [doi: 10.3390/ijerph15061182] [Medline: 29874883]
40. Laboratoire d'optimisation cognitive de l'Université de sherbrooke [Website in French]. URL: https://lopcus.recherche.usherbrooke.ca/ [accessed 2026-01-19]
41. Paillé P, Mucchielli A. L'analyse Qualitative En Sciences Humaines et Sociales [Book in French], 5th edition: Armand Colin; 2021. URL: https://www.dunod.com/sciences-humaines-et-sociales/analyse-qualitative-en-sciences-humaines-et-sociales-1 [accessed 2026-01-19]
42. Schwaber K. SCRUM development process. In: Sutherland J, Casanave C, Miller J, Patel P, Hollowell G, editors. Business Object Design and Implementation: Springer; 1997:117-134. [doi: 10.1007/978-1-4471-0947-1_11]
43. Christie MA, Bhandar A, Nakandala S, et al. Managing authentication and authorization in distributed science gateway middleware. Future Gener Comput Syst 2020 Oct;111:780-785. [doi: 10.1016/j.future.2019.07.018]
44. AMI. URL: https://ami.lopcus.ca [accessed 2026-01-27]
45. Corrado AM, Benjamin-Thomas TE, McGrath C, Hand C, Laliberte Rudman D. Participatory action research with older adults: a critical interpretive synthesis. Gerontologist 2020 Jul 15;60(5):e413-e427. [doi: 10.1093/geront/gnz080] [Medline: 31264680]
46. Davies J, Read J. A systematic review into the incidence, severity and duration of antidepressant withdrawal effects: are guidelines evidence-based? Addict Behav 2019 Oct;97:111-121. [doi: 10.1016/j.addbeh.2018.08.027] [Medline: 30292574]
47. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. JAMA 2014 Mar 12;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]
48. Brewster GS, Molinari V, McCrae C, Beckstead J, D'Aoust R, Rowe M. Cognitive function and sleep in caregivers of persons living with dementia. West J Nurs Res 2022 Mar;44(3):260-268. [doi: 10.1177/0193945921104116 3] [Medline: 34467789]
49. Warmoth K, Morgan-Trimmer S, Kudlicka A, et al. Reflections on a personalized cognitive rehabilitation intervention: experiences of people living with dementia and their carers participating in the GREAT trial. Neuropsychol Rehabil 2022 Mar;32(2):268-286. [doi: 10.1080/09602011.2020.1820876] [Medline: 32972311]
50. Ploeg J, McAiney C, Duggleby W, et al. A web-based intervention to help caregivers of older adults with dementia and multiple chronic conditions: qualitative study. JMIR Aging 2018 Apr 23;1(1):e2. [doi: 10.2196/aging.8475] [Medline: 31518230]
51. Hassan AYI. Challenges and recommendations for the deployment of information and communication technology solutions for informal caregivers: scoping review. JMIR Aging 2020 Jul 29;3(2):e20310. [doi: 10.2196/20310] [Medline: 32723720]
52. Duggleby W, Ploeg J, McAiney C, et al. A comparison of users and nonusers of a web-based intervention for carers of older persons with Alzheimer disease and related dementias: mixed methods secondary analysis. J Med Internet Res 2019 Oct 17;21(10):e14254. [doi: 10.2196/14254] [Medline: 31625947]
53. Teles S, Paúl C, Lima P, Chilro R, Ferreira A. User feedback and usability testing of an online training and support program for dementia carers. Internet Interv 2021 Sep;25:100412. [doi: 10.1016/j.invent.2021.100412] [Medline: 34401371]

XSL•FO
RenderX

54. Hajek A, Kretzler B, König HH. Informal caregiving, loneliness and social isolation: a systematic review. Int J Environ Res Public Health 2021 Nov 18;18(22):12101. [doi: 10.3390/ijerph182212101] [Medline: 34831857]
55. Bidenko K, Bohnet-Joschko S. Supporting family care: a scoping app review. BMC Med Inform Decis Mak 2022 Jun 21;22(1):162. [doi: 10.1186/s12911-022-01906-6] [Medline: 35729573]
56. Scariot CA, Heemann A, Padovani S. Understanding the collaborative-participatory design. Work 2012;41 Suppl 1:2701-2705. [doi: 10.3233/WOR-2012-0656-2701] [Medline: 22317129]
57. Sala-González M, Pérez-Jover V, Guilabert M, Mira JJ. Mobile apps for helping informal caregivers: a systematic review. Int J Environ Res Public Health 2021 Feb 10;18(4):1702. [doi: 10.3390/ijerph18041702] [Medline: 33578819]
58. Hopwood J, Walker N, McDonagh L, et al. Internet-based interventions aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018 Jun 12;20(6):e216. [doi: 10.2196/jmir.9548] [Medline: 29895512]

Abbreviations

AMI: Aide-Mémoire-Interactif
MNCD: major neurocognitive disorder

Edited by AF Abdulai; submitted 27.May.2025; peer-reviewed by M Cozad, S Law; revised version received 14.Dec.2025; accepted 07.Jan.2026; published 26.Feb.2026.

Please cite as:
Viscogliosi C, Ngankam HK, Duquette-Laplante R, Rahimaly S, Pigot H, Chassé B, Ait Bouziaren S, Couturier Y, Déry J, Biard N, Duchesne J, Giroux D, Jacob S, Routhier-Chevrier K, Mino-Roy J, Caron M, Provencher V
Web-Based Application to Support Caregivers in the Use of Learning Optimization Methods: Participatory Action Research Study
JMIR Aging 2026;9:e76543
URL: https://aging.jmir.org/2026/1/e76543
doi:10.2196/76543

© Chantal Viscogliosi, Hubert Kenfack Ngankam, Romane Duquette-Laplante, Sarah Rahimaly, Hélène Pigot, Bernard Chassé, Sara Ait Bouziaren, Yves Couturier, Jessica Déry, Nicolas Biard, Jordan Duchesne, Dominique Giroux, Sabrina Jacob, Jordan Mino-Roy, Kimberley Routhier-Chevrier, Mélanie Caron, Véronique Provencher. Originally published in JMIR Aging (https://aging.jmir.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Level of eHealth Literacy and Its Associations With Health Behaviors and Outcomes in Chinese Older Adults: Cross-Sectional Analysis of Baseline Data From a Large-Scale Community Project

Siu Long Chau[1*], PhD; Wanjia He[1*], PhD; Tzu Tsun Luk[2], PhD; Sophia Siu Chee Chan[1*], PhD

[1]School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Metro South Tower 1, 39 Wong Chuk Hang Road, Wong Chuk Hang, Hong Kong, China (Hong Kong)

[2]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[*]these authors contributed equally

Corresponding Author:
Sophia Siu Chee Chan, PhD
School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Metro South Tower 1, 39 Wong Chuk Hang Road, Wong Chuk Hang, Hong Kong, China (Hong Kong)

Abstract

Background: eHealth literacy is important for older adults to be able to seek and evaluate online health information. However, there is a scarcity of large-scale data on their eHealth literacy levels, particularly among the oldest older individuals (aged >75 years) in unique, high-income Asian regions such as Hong Kong. A comprehensive understanding of how eHealth literacy is associated with specific health behaviors, mental well-being, and physical health outcomes in this population is lacking.

Objective: This study aims to assess the level of eHealth literacy and its associations with health behaviors and health-related outcomes among older adults in Hong Kong.

Methods: We conducted a cross-sectional analysis of baseline data from the Generations Connect Project. This is an ongoing large-scale community-based project, where we trained university students to conduct home visits and assess the health status of older adults (N=6704) in Hong Kong. eHealth literacy was measured using the eHealth Literacy Scale (eHEALS; score: 8 - 40). Health behavior measurements included physical activity levels (metabolic equivalent of task minutes per week) and smoking, drinking, and eating habits. Mental well-being was measured using the World Health Organization-Five Well-Being Index (percentage score: 0 - 100) and UCLA 3-item Loneliness Scale (score: 3 - 9). Physical health was assessed on the basis of self-reported medical diagnosis of noncommunicable diseases (yes/no), including hypertension, diabetes, cardiovascular disease, and stroke. Adjusted unstandardized coefficients (b) and odds ratios (ORs) were calculated to determine the associations between variables.

Results: Among the 6704 participants (mean age 77.8, SD 7.0 years), the mean eHEALS score was 18.2 (SD 10.2), and 44.1% (2897/6566) of the participants had inadequate eHealth literacy (score: 8 - 15.99). Increasing age (adjusted b −0.32, 95% CI −0.35 to −0.28; *P*<.001), support from the Comprehensive Social Security Assistance Scheme (adjusted b −1.49, 95% CI −2.04 to −0.95; *P*<.001), and living in public housing (adjusted b −1.60, 95% CI −2.69 to −0.50; *P*=.004) were associated with a lower eHEALS score. Participants with moderate eHealth literacy (score: 24 - 31.99) were less likely to be current smokers (adjusted OR 0.60, 95% CI 0.38-0.95; *P*=.04), more physically active (adjusted b 39.83, 95% CI 2.04-77.62; *P*=.04), more likely to be community health center members (adjusted OR 1.52, 95% CI 1.30-1.77; *P*<.001) and to have healthy diets (adjusted b 0.034, 95% CI 0.006-0.063; *P*=.04), and less likely to have a medical diagnosis of diabetes (adjusted OR 0.73, 95% CI 0.62-0.85; *P*<.001). Moreover, they had a higher score on the World Health Organization-Five Well-being Index (adjusted b 2.89, 95% CI 1.42-4.36; *P*<.001) and a lower score on the UCLA 3-item Loneliness Scale (adjusted b −0.26, 95% CI −0.37 to −0.15; *P*<.001).

Conclusions: The level of eHealth literacy was low among older adults in Hong Kong. eHealth literacy was associated with positive health behaviors and health-related outcomes. Interventions are warranted to boost their eHealth literacy in the future.

(JMIR Aging 2026;9:e74110) doi:10.2196/74110

KEYWORDS

eHealth literacy; older adults; health behaviors; health-related outcomes; healthy aging

XSL•FO
RenderX

Introduction

Older adults (aged >60 years) accounted for 5% of the global population in the last few decades, and this estimate is projected to rise to 16% by 2050 [1]. In China, the population aged ≥80 years reached 36 million in 2020 and will increase to 115 million in 2050 [2]. This global demographic shift toward an older population is accompanied by an increased burden of noncommunicable diseases (NCDs) [3] and mental illnesses [4], which places immense strain on public health care systems. More than 20% of older adults in the world have 1 or more mental illnesses [4], and the prevalence increases with population size. In 2017, 8.6% of older adults in Hong Kong (N=4368) had depression and anxiety [5]. Mental illnesses among older people remained underrecognized and undertreated, which accounted for 10.6% of the total loss of disability-adjusted life years in older adults [6]. In response, and in line with the World Health Organization's call for primary health care reform [7], there is an urgent need for innovative strategies to support healthy aging. The rapid development of digital technologies, accelerated by the COVID-19 pandemic, offers a key opportunity. Digital tools have the potential to support older adults' intrinsic capacity, promote active and healthy aging, and alleviate the economic burden on health care sectors, making them an increasingly recognized priority in geriatric care [8].

The emergence of digital health offers a new avenue for improving individuals' quality of life. Rapid advances in information and communications technology contribute to the increasing innovation and upgrading of health service modes [9]. Digital health enables low-cost, timely health care services for older adults and supports their physical and mental health by providing access to reliable health information, self-monitoring tools for chronic conditions, and social connection platforms [9]. Although accessing online medical services and information is convenient, the quality of online information varies [10]. Older adults need eHealth literacy to navigate and evaluate the accuracy of medical information, especially in the context of primary health care reform in Hong Kong, where the government is committed to enhancing digital primary health care services to shift the emphasis from tertiary to primary care.

eHealth literacy is defined as the ability of an individual to find and evaluate health information from online platforms and apply health-related knowledge to address health problems [11]. Although smartphone ownership has increased among older adults [12], they often possess lower technological skills than the younger generations and are consequently less likely to use these devices to search for health information [12]. eHealth literacy impacts how older adults search for health information online, which might influence health-related decisions [12]. With the rapid and continuous development of eHealth resources, it is crucial to assess older adults' eHealth literacy and examine its association with health-related outcomes.

Previous studies suggested that higher eHealth literacy was associated with some healthy aging components (eg, less cognitive impairment and functional limitation) in older adults [13]. However, database searches in PubMed and CINAHL (using keywords such as "eHealth literacy," "older adults," "aging population," "health behavior," "smoking," and "physical activity") suggest that studies on the association of eHealth literacy with specific health behaviors (eg, smoking, drinking, and physical exercise) and health status are scarce. Only 1 recent meta-analysis showed a positive association between eHealth literacy and general health-related behaviors (overall estimate of the correlation: 0.31, 95% CI: 0.25 - 0.34), with a similar effect size observed in older adult populations [14]. Another systematic review showed that higher eHealth literacy was associated with better health knowledge and attitude in older adults, but the associations with physical and psychosocial outcomes were still inconsistent [15]. The relationships between eHealth literacy and specific health-related outcomes (eg, mental and physical well-being) need to be assessed to further explore their underlying mechanisms. In addition, studies on digital health often recruit participants from more accessible, community-dwelling populations [16], potentially underrepresenting the most vulnerable, such as homebound or socioeconomically disadvantaged older adults, who may face the greatest digital divide.

While existing research provides a foundational understanding of eHealth literacy, several critical gaps remain, limiting the development of effective, evidence-based interventions for older adults. First, eHealth literacy levels are known to vary significantly across different sociocultural and demographic contexts, with studies in the United States [16], South Korea [17], and Norway [18] reporting higher average scores than those reported in mainland China [19]. However, there is a scarcity of large-scale, representative data from Hong Kong, a unique high-income region characterized by extreme longevity, high digital penetration, and a government-led push for digital primary care reform. Second, much of the existing literature focuses on younger older populations (eg, those aged 60-65 years); however, the challenges and capabilities of the oldest older individuals (ie, those aged ≥75 years) may differ substantially. Third, while a study has confirmed a general positive association between eHealth literacy and composite health-related behaviors [14], there is a lack of research examining the associations between eHealth literacy and a comprehensive suite of specific, modifiable behaviors (eg, smoking, physical activity, and diet) and health service use. Finally, outcomes are often studied in silos. A holistic understanding requires a simultaneous investigation of the links between eHealth literacy and a broad spectrum of outcomes, including health behaviors, mental well-being, social connectedness, and physical health status.

To address these gaps, this study aimed to (1) assess the level of eHealth literacy in a large, community-based sample of underprivileged older adults in Hong Kong, with a particular focus on the oldest older population and (2) comprehensively examine the associations of eHealth literacy with specific health behaviors, mental health outcomes (well-being and loneliness), and physical health outcomes (prevalence of multiple NCDs).

XSL•FO
RenderX

Methods

Study Design

This study used a cross-sectional design, using baseline data from the Generations Connect Project. The parent project is an ongoing, large-scale, community-based, quasi-experimental pre-post study that trained over 1000 nursing students to provide home visits and perform a baseline health assessment and to deliver a health intervention (face-to-face and digital) to older adults in Hong Kong. The present study focused on the analysis of the baseline data collected before any intervention took place.

Older adults were recruited through collaborations with 20 local nongovernmental organizations covering 18 districts across Hong Kong. Between November 2022 and December 2024, we recruited participants aged ≥65 years who could read and communicate in Chinese and were cognitively and physically capable of understanding and answering survey questions. Older adults who were bedbound and had a history of mental illnesses were excluded. The exclusion of older adults with self-reported mental illness was to minimize potential confounding effects on the mental health outcomes measured in this study. In contrast, older adults with physical limitations (eg, visual or hearing impairments) were not excluded, as these conditions are prevalent in the target population, and their inclusion is crucial for the generalizability of our findings regarding the digital divide among the oldest older population. We recruited University of Hong Kong students to be student ambassadors through open recruitment via internal university emails and on-campus promotional materials (eg, posters and flyers) as well as selected curriculum (medical, nursing, dental, and common core courses) integration. Before making a home visit, all student ambassadors received a 3-hour training to ensure the intervention's fidelity and data quality during the data collection process. Details of the training session and intervention are reported elsewhere [20].

Ethical Considerations

Ethical approval was sought from the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (IRB/REC reference: UW 22 - 693). The institutional review board also allowed secondary analysis without additional consent. The reporting of this study strictly followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (Checklist 1). The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Informed consent was obtained using comprehensive information sheets and verbal explanations, and signed consent forms were obtained from participants before any procedures began. The consent form included information about the study's purpose and procedures and the participant's right to withdraw at any time without penalty.

Participants enrolled in the study received gift bags worth HK $50 (US $6.44) as an incentive during the initial student visits.

All data collected during the study were deidentified to protect the participants' privacy, with all identifying information removed. Access to the data was restricted to the research team, and all data were stored securely.

Measurements

We collected and analyzed the baseline data on the sociodemographic characteristics (eg, sex, age, highest educational attainment, marital status, and family monthly income), health behaviors, mental health, physical health, and eHealth literacy of participants.

Health behaviors included physical activity levels in metabolic equivalent of task minutes (MET min) per week (measured using the International Physical Activity Questionnaire Short Form) [21], number of days of vegetable and fruit consumption in the past week, smoking status (never smoker, ex-smoker, and current smoker), intention to quit within 30 days (yes/no) for current smokers, and drinking frequency in the past year (never, monthly or less, 2-4 times per month, 2-3 times per week, and ≥4 times per week).

Mental health and perceived level of loneliness were assessed using the World Health Organization-Five (WHO-5) Well-Being Index [22] and UCLA 3-item Loneliness Scale, respectively [23]. The raw score on the WHO-5 ranges from 0 to 25, and the percentage score ranges from 0 to 100 (raw score multiplied by 4). A percentage score of 0 represents the worst possible mental health status, and 100 represents the best possible mental health status [22]. The WHO-5 has been validated for the Chinese population [24]. In our sample, the scale demonstrated good internal consistency, with a Cronbach α of 0.89. The UCLA 3-item Loneliness Scale measures 3 dimensions of loneliness (relational and social connectedness and self-perceived isolation) with a score ranging from 3 to 9, where a higher score indicates a higher level of loneliness [23]. The UCLA 3-item Loneliness Scale was validated for older Chinese adults [25]. The scale showed high reliability in the present study, with a Cronbach α of 0.88. Physical health was assessed by asking whether the participants had been medically diagnosed with each NCD (yes/no), including hypertension, diabetes, cardiovascular disease, and stroke. A multimorbidity variable was constructed (defined as having been medically diagnosed with ≥2 of the following 4 conditions: hypertension, diabetes, cardiovascular disease, and stroke).

eHealth literacy was measured using the eHealth Literacy Scale (eHEALS) [11]. The scale measures participants' perceived skills, knowledge, and comfort toward eHealth [11]. The scale consists of 8 items, which are rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The total score ranges from 8 to 40, with higher scores indicating higher self-perceived eHealth literacy [11]. On the basis of the score, the respondents can be categorized into 4 groups: inadequate eHealth literacy (score: 8 - 15.99), low eHealth literacy (score: 16 - 23.99), moderate eHealth literacy (score: 24 - 31.99), and high eHealth literacy (score: 32 - 40) [26]. The scale has high internal consistency, with a Cronbach α of 0.96 [26], and has been validated for Chinese older adults, with a Cronbach α of 0.88 [26].

XSL•FO
RenderX

Statistical Analysis

Inverse probability weighting based on the sex distribution of Hong Kong older adults in 2023 (from the census) was conducted to make the sample more representative of Hong Kong's older adult population. Descriptive statistics were used to describe the participants' sociodemographic characteristics, eHealth literacy, health behaviors, mental health, and physical health at baseline.

The associations of sociodemographic characteristics (sex, age, highest educational attainment, family monthly income, Comprehensive Social Security Assistance [CSSA] Scheme status, type of housing, and living with family members) with the eHEALS score were analyzed using multivariable linear regression (adjusted unstandardized coefficient [b]) with mutual adjustment for baseline sociodemographic characteristics.

Multiple logistic regression (adjusted odds ratio [AOR]) was used to analyze the associations of the level of eHealth literacy with smoking and drinking status (yes/no), District Health Center member status (yes/no), multimorbidity of NCDs (yes/no), and medical diagnosis of NCDs (yes/no). Multivariable linear regression (adjusted unstandardized coefficient [b]) was used to examine the associations of the level of eHealth literacy with total physical activity (MET min/week), the WHO-5 Well-Being Index score, and the UCLA 3-item Loneliness Scale score. The associations of the level of eHealth literacy with the number of days of vegetable and fruit consumption in the past week were analyzed using Poisson regression. The models were adjusted for baseline sociodemographic characteristics. All analyses were conducted in Stata (version 15.1; StataCorp LLC). A 2-sided *P* value <.05 was considered statistically significant.

Results

Between November 2022 and December 2024, a total of 7087 potential participants were screened and found eligible, and 6704 agreed to participate and were included in the baseline analysis. The mean time to conduct baseline assessments and deliver interventions was 102 (SD 7.3) minutes.

Table 1 shows that the mean age of the participants was 77.8 (SD 7.0) years, and 26.9% (1805/6704) of the participants were men. Most participants had primary education or below (4003/6676, 60.0%), were married (3173/6626, 47.9%), and had a family monthly income of HK $25,000 (US $3205.9) or less (5127/6653, 77.1%). Moreover, 25.4% (1691/6650) of participants were supported by the CSSA Scheme. Most participants were retired (6460/6694, 96.5%), living in public housing (4829/6696, 72.1%), and living with family members (3479/6696, 52.0%). In total, 71.2% (4774/6704) of participants were not District Health Center members. The mean eHEALS score of participants was 18.2 (SD 10.2), and most participants (2897/6566, 44.1%) had inadequate eHealth literacy (score: 8 - 15.99).

XSL•FO
RenderX

Table . Sociodemographic characteristics of participants (N=6704)[a].

Characteristic	Participants, n (%)
Sex	
Male	1805 (26.9)
Female	4894 (73.1)
Age[b] (years)	77.8 (7.0)
Highest educational attainment	
Primary education or below	4003 (60)
Secondary education	2288 (34.2)
Tertiary education	385 (5.8)
Marital status	
Single	397 (6)
Married	3173 (47.9)
Divorced	454 (6.8)
Widowed	2602 (39.3)
Family monthly income	
None	204 (3.1)
HK $25,000 (US $ 3205.9) or less	5127 (77.1)
More than HK $25,000 (US $ 3205.9)	1322 (19.9)
Support from the Comprehensive Social Security Assistance Scheme	
No	4959 (74.6)
Yes	1691 (25.4)
Occupational status	
Full-time (employed)	35 (0.5)
Part-time (employed)	90 (1.3)
Unemployed	13 (0.2)
Retired	6460 (96.5)
Caregiver/housewife	96 (1.4)
Type of housing	
Public	4829 (72.1)
Private	1546 (23.1)
Other	321 (4.8)
Living with family members	
No	3046 (45.5)
Yes	3479 (52)
Other	171 (2.5)
District Health Center member status	
No	4774 (71.2)
Yes	1930 (28.8)
eHealth literacy (eHEALS score[b]: 8 - 40)[c]	18.2 (10.2)
Inadequate (score: 8 - 15.99)	2897 (44.1)
Low (score:16 - 23.99)	1322 (20.1)
Moderate (score: 24 - 31.99)	1182 (18)

XSL•FO
RenderX

Characteristic	Participants, n (%)
High (score: 32 - 40)	1165 (17.7)

[a]The proportions were weighted by sex distribution of older adults in Hong Kong 2023. The observations (n) were unweighted.

[b]Reported as mean (SD).

[c]eHealth literacy is measured on a 5-point eHealth Literacy Scale (eHEALS), with ratings ranging from 1 (strongly disagree) to 5 (strongly agree). The overall score ranges from 8 to 40, with a higher score indicating more perceived skills in finding, evaluating, and using electronic information to make health decisions.

Table 2 shows that the participants' mean total physical activity was 628.7 (SD 542.9) MET min per week. The mean number of days on which vegetables and fruits were consumed in the past week was 6.6 (SD 1.3) and 5.9 (SD 2.0), respectively. In total, 3% (198/6686) of participants were current smokers, and 91% (163/179) had no intention to quit. Moreover, 8.5% (570/6704) of participants were current drinkers. The mean percentage score on the WHO-5 Well-being Index was 68.1 (SD 21.3), and the mean score on the UCLA 3-item Loneliness Scale was 4.0 (SD 1.6). Notably, 45.2% (3020/6674) of participants had multimorbidity of NCDs (medically diagnosed with ≥2 NCDs). Of the 6704 participants, 60.5% (n=4059) were medically diagnosed with hypertension, 27.1% (n=1816) had diabetes, 13.3% (n=894) had cardiovascular disease, and 4.6% (n=309) had a medical diagnosis of stroke.

XSL•FO
RenderX

Table . Health behaviors, mental health, and physical health of participants (N=6704)[a].

Parameter	Outcome
Physical activity, IPAQ-SF continuous score (MET[b] min/week), mean (SD)[c]	
Walking	463.5 (443.8)
Moderate	139.9 (222.6)
Vigorous	25.2 (108.6)
Total	628.7 (542.9)
Number of days of vegetable consumption in the past week, mean (SD)	6.6 (1.3)
Number of days of fruit consumption in the past week, mean (SD)	5.9 (2.0)
Smoking status, n (%)	
Never smoker	5812 (86.9)
Ex-smoker	676 (10.1)
Current smoker	198 (3)
Intention to quit for current smokers, n (%)	
No	163 (91)
Yes	16 (9)
Drinking frequency in the past year, n (%)	
Never	6104 (91.5)
Monthly or less	289 (4.3)
2 - 4 times per month	129 (1.9)
2 - 3 times per week	46 (0.7)
>4 times per week	106 (1.6)
WHO-5[d] Well-Being Index percentage score (0 - 100)[e], mean (SD)	68.1 (21.3)
UCLA 3-item Loneliness Scale score: 3 - 9)[f], mean (SD)	4.0 (1.6)
Multimorbidity[g] of NCDs[h], n (%)	
No	3654 (54.7)
Yes	3020 (45.2)
Medical diagnosis of hypertension, n (%)	
No	2645 (39.5)
Yes	4059 (60.5)
Medical diagnosis of diabetes, n (%)	
No	4888 (72.9)
Yes	1816 (27.1)
Medical diagnosis of cardiovascular disease, n (%)	
No	5810 (86.7)
Yes	894 (13.3)
Medical diagnosis of stroke, n (%)	
No	6395 (95.4)
Yes	309 (4.6)

[a]The proportions were weighted by sex distribution of older adults in Hong Kong 2023. The observations (n) were unweighted.

[b]MET: metabolic equivalent of task.

[c]The International Physical Activity Questionnaire Short Form (IPAQ-SF) continuous scores are expressed in MET min/week: MET level × minutes of activity × events per week. The total physical activity is computed as the sum of walking, moderate, and vigorous scores (MET min/week).

XSL•FO
RenderX

[d]WHO-5: World Health Organization-Five Well-Being Index.

[e]The raw score, ranging from 0 to 25, is multiplied by 4 to obtain a percentage score. A percentage score of 0 represents the worst possible quality of life, whereas a score of 100 represents the best possible quality of life.

[f]The 3-point response scale for each item ranges from "hardly ever or never" (1 point) to "often" (3 points), and the total score is the sum of all items, which ranges from 3 to 9, with higher scores indicating a higher level of perceived loneliness.

[g]Multimorbidity was defined as having been medically diagnosed with ≥2 of the following 4 conditions: hypertension, diabetes, cardiovascular disease, and stroke.

[h]NCD: noncommunicable disease.

Table 3 shows that increasing age (adjusted b −0.32, 95% CI −0.35 to −0.28; *P*<.001), support from the CSSA Scheme (adjusted b −1.49, 95% CI −2.04 to −0.95; *P*<.001), and living in public housing (adjusted b −1.60, 95% CI −2.69 to −0.50; *P*=.004) were associated with a lower eHEALS score, after mutual adjustment for baseline sociodemographic characteristics. Secondary education (adjusted b 4.31, 95% CI 3.81-4.82; *P*<.001) and tertiary education (adjusted b 9.04, 95% CI 8.04-10.06; *P*<.001) were associated with higher eHEALS scores.

Table . Associations of sociodemographic characteristics with eHealth literacy of participants (N=6704).

Characteristic	Association with eHealth Literacy Scale score (8 - 40)			
	Crude b (95% CI)	*P* value	Adjusted b (95% CI)[a]	*P* value
Sex (reference: male)				
Female	0.05 (−0.50 to 0.61)	.85	0.17 (−0.38 to 0.71)	.89
Age	−0.40 (−0.43 to −0.36)	<.001	−0.32 (−0.35 to −0.28)	<.001
Highest educational attainment (reference: primary education or below)				
Secondary education	5.37 (4.86 to 5.87)	<.001	4.31 (3.81 to 4.82)	<.001
Tertiary education	9.05 (8.02 to 10.08)	<.001	9.04 (8.04 to 10.06)	<.001
Family monthly income (reference: none)				
HK $25,000 (US $3205.9) or less	0.50 (−0.12 to 1.12)	.12	0.24 (−0.34 to 0.82)	.16
More than HK $25,000 (US $3205.9)	2.27 (0.81 to 3.73)	.002	0.91 (−0.46 to 2.29)	.06
Support from the Comprehensive Social Security Assistance Scheme (reference: no)				
Yes	−2.39 (−2.96 to −1.82)	<.001	−1.49 (−2.04 to −0.95)	<.001
Type of housing (reference: other)				
Public	−2.77 (−3.93 to −1.62)	<.001	−1.60 (−2.69 to −0.50)	.004
Private	0.24 (−0.98 to 1.47)	.70	0.13 (−0.73 to 1.41)	.69
Living with family members (reference: other)				
No	0.06 (−1.72 to 1.84)	.94	−0.79 (−2.46 to 0.87)	.87
Yes	0.99 (−0.78 to 2.76)	.95	−0.43 (−2.10 to 1.23)	.89

[a]The model was mutually adjusted for all sociodemographic characteristics listed in the table, including sex, age, highest educational attainment, family monthly income, Comprehensive Social Security Assistance status, type of housing, and living with family members.

Descriptive analysis showed that 93% (6227/6704) of participants were smartphone users. In total, 57.5%(3549/6173) used a smartphone for 1 hour or more daily, 91.5% (6135/6704) had instant messaging apps installed on the smartphone, and 60.2% (3992/6634) had not searched for health-related information from online sources using the smartphone (Multimedia Appendix 1). Table 4 shows that participants with moderate eHealth literacy (score: 24 - 31.99) were less likely to be current smokers (AOR 0.60, 95% CI 0.38-0.95; *P*=.04) and more likely to be District Health Center members (AOR 1.52, 95% CI 1.30-1.77; *P*<.001), and moderate eHealth literacy was associated with higher total physical activity levels (adjusted

XSL•FO
RenderX

b 39.83, 95% CI 2.04-77.62; *P*=.04) and a higher number of days of fruit consumption in the past week (adjusted b 0.034, 95% CI 0.006-0.063; *P*=.04). Participants with high eHealth literacy (score 32 - 40) were less likely to be current smokers (AOR 0.57, 95% CI 0.36-0.92; *P*=.03) and more likely to be District Health Center members (AOR 1.76, 95% CI 1.51-2.05; *P*<.001), and high eHealth literacy was associated with a higher number of days of fruit consumption in the past week (adjusted b 0.046, 95% CI 0.017-0.075; *P*=.003).

Table . Associations of eHealth literacy with health behaviors of participants (N=6704).

Health behavior	eHealth literacy (eHEALS[a] score: 8 - 40) (reference: inadequate eHealth literacy [score: 8-15.99], n=2897)					
	Low (score:16 - 23.99) n=1322	*P* value	Moderate (score: 24 - 31.99) n=1182	*P* value	High (score: 32 - 40) n=1165	*P* value
Current smoker						
Crude OR (95% CI)	0.94 (0.65 to 1.35)	.73	0.69 (0.45 to 1.06)	.09	0.66 (0.42 to 1.02)	.06
Adjusted OR (95% CI)[b]	0.90 (0.61 to 1.34)	.67	0.60 (0.38 to 0.95)	.04	0.57 (0.36 to 0.92)	.03
Current drinker						
Crude OR (95% CI)	1.23 (0.98 to 1.55)	.07	1.33 (1.06 to 1.68)	.02	1.01 (0.79 to 1.30)	.91
AOR (95% CI)[b]	1.16 (0.92 to 1.48)	.10	1.19 (0.93 to 1.53)	.06	0.85 (0.65 to 1.11)	.93
Physical activity (total MET min/week)						
Crude b (95% CI)	5.85 (−29.77 to 41.47)	.75	74.47 (37.54 to 111.39)	<.001	80.12 (42.96 to 117.28)	<.001
Adjusted b (95% CI)[b]	−10.07 (−45.96 to 25.82)	.80	39.83 (2.04 to 77.62)	.04	36.51 (−2.16 to 75.18)	.06
District Health Center member status						
Crude OR (95% CI)	1.25 (1.07 to 1.45)	.003	1.87 (1.61 to 2.16)	<.001	2.27 (1.97 to 2.63)	<.001
AOR (95% CI)[c]	1.13 (0.97 to 1.32)	.06	1.52 (1.30 to 1.77)	<.001	1.76 (1.51 to 2.05)	<.001
Number of days of vegetable consumption in the past week[d]						
Crude b (95% CI)	0.01 (−0.02 to 0.03)	.68	0.01 (−0.01 to 0.04)	.72	0.02 (−0.01 to 0.05)	.44
Adjusted b (95% CI)[b]	0.01 (−0.02 to 0.03)	.70	0.02 (−0.01 to 0.04)	.83	0.02 (−0.01 to 0.05)	.45
Number of days of fruit consumption in the past week[d]						
Crude b (95% CI)	−0.003 (−0.030 to 0.024)	.56	0.034 (0.006 to 0.062)	.04	0.047 (0.019 to 0.074)	.003
Adjusted b (95% CI)[b]	−0.006 (−0.033 to 0.022)	.57	0.034 (0.006 to 0.063)	.04	0.046 (0.017 to 0.075)	.003

[a]eHEALS: eHealth Literacy Scale.

[b]Adjusted for sex, age, highest educational attainment, family monthly income, living with family members, and being a District Health Center member.

[c]Adjusted for sex, age, highest educational attainment, family monthly income, and living with family members.

[d]Poisson regression was used to calculate the unstandardized coefficient.

XSL•FO
RenderX

Table 5 shows that participants with moderate eHealth literacy (score 24 - 31.99) were less likely to have a medical diagnosis of diabetes (AOR 0.73, 95% CI 0.62-0.85; *P*<.001) and stroke (AOR 0.58, 95% CI 0.40-0.85; *P*=.004), and moderate eHealth literacy was associated with a higher score on the WHO-5 Well-being Index (adjusted b 2.89, 95% CI 1.42-4.36; *P*<.001) and a lower score on the UCLA 3-item Loneliness Scale (adjusted b −0.26, 95% CI −0.37 to −0.15; *P*<.001). Participants with high eHealth literacy (score: 32 - 40) were less likely to have a medical diagnosis of hypertension (AOR 0.77, 95% CI 0.67-0.90; *P*=.004) and diabetes (AOR 0.69, 95% CI 0.58-0.81; *P*<.001), and high eHealth literacy was associated with a higher score on the WHO-5 Well-being Index (adjusted b 5.37, 95% CI 3.86-6.87; *P*<.001) and lower score on the UCLA 3-item Loneliness Scale (adjusted b −0.34, 95% CI −0.45 to −0.23; *P*<.001).

XSL•FO
RenderX

Table . Associations of eHealth literacy with the physical health and mental health of participants (N=6704).

Parameter	eHealth literacy (eHEALS[a] score: 8 - 40) (reference: inadequate eHealth literacy [score: 8-15.99], n=2897)					
	Low (score:16 - 23.99) n=1322	*P* value	Moderate (score: 24 - 31.99) n=1182	*P* value	High (score: 32 - 40) n=1165	*P* value
Medical diagnosis of hypertension						
Crude OR[b] (95% CI)	1.01 (0.88 to 1.15)	.92	0.77 (0.67 to 0.89)	<.001	0.65 (0.56 to 0.74)	<.001
Adjusted OR (95% CI)[c]	1.06 (0.93 to 1.22)	.95	0.89 (0.77 to 1.02)	.06	0.77 (0.67 to 0.90)	.004
Medical diagnosis of diabetes						
Crude OR (95% CI)	0.84 (0.73 to 0.97)	.02	0.73 (0.63 to 0.85)	<.001	0.69 (0.59 to 0.80)	<.001
Adjusted OR (95% CI)[c]	0.83 (0.71 to 0.96)	.03	0.73 (0.62 to 0.85)	<.001	0.69 (0.58, 0.81)	<.001
Medical diagnosis of cardiovascular disease						
Crude OR (95% CI)	0.99 (0.83 to 1.20)	.98	0.81 (0.66 to 1.00)	.06	0.93 (0.76 to 1.14)	.49
Adjusted OR (95% CI)[c]	1.10 (0.91 to 1.34)	.98	0.97 (0.79 to 1.21)	.07	1.12 (0.91 to 1.40)	.52
Medical diagnosis of stroke						
Crude OR (95% CI)	0.76 (0.55 to 1.04)	.08	0.58 (0.40 to 0.83)	.003	0.77 (0.55 to 1.06)	.11
Adjusted OR (95% CI)[c]	0.74 (0.54 to 1.03)	.08	0.58 (0.40 to 0.85)	.004	0.80 (0.56 to 1.12)	.13
Multimorbidity[f] of NCDs[g]						
Crude b (95% CI)	0.83 (0.57 to 1.09)	.12	0.67 (0.44 to 1.07)	.10	0.65 (0.50 to 1.12)	.14
Adjusted b (95% CI)[c]	0.86 (0.62 to 1.12)	.14	0.69 (0.48 to 1.10)	.10	0.66 (0.52 to 1.14)	.14
WHO-5 Well-Being Index score[d]						
Crude b (95% CI)	0.96 (−0.42 to 2.34)	.17	2.75 (1.32 to 4.19)	<.001	5.65 (4.21 to 7.09)	<.001
Adjusted b (95% CI)[c]	0.87 (−0.52 to 2.27)	.15	2.89 (1.42 to 4.36)	<.001	5.37 (3.86 to 6.87)	<.001
UCLA 3-item Loneliness Scale score[e]						
Crude b (95% CI)	−0.19 (−0.29 to −0.08)	.003	−0.27 (−0.38 to −0.15)	<.001	−0.40 (−0.52 to −0.29)	<.001
Adjusted b (95% CI)[c]	−0.17 (−0.28 to −0.07)	.003	−0.26 (−0.37 to −0.15)	<.001	−0.34 (−0.45 to −0.23)	<.001

[a]eHEALS: eHealth Literacy Scale.

XSL•FO
RenderX

[b]OR: odds ratio.

[c]Adjusted for sex, age, highest educational attainment, family monthly income, living with family members, and District Health Center member.

[d]WHO-5: World Health Organization-Five Well-Being Index. The score ranges from 0 to 25, with 0 representing the worst possible and 25 representing the best possible quality of life. The score is multiplied by 4. A percentage score of 0 represents the worst possible quality of life, whereas a score of 100 represents the best possible quality of life.

[e]The 3-point response scale for each item ranges from "hardly ever or never" (1 point) to "often" (3 points), and the total score is the sum of all items, which ranges from 3 to 9, with higher scores indicating a higher level of perceived loneliness.

[f]Multimorbidity was defined as having been medically diagnosed with ≥2 of the following 4 conditions: hypertension, diabetes, cardiovascular disease, and stroke.

[g]NCD: noncommunicable disease.

Discussion

Principal Findings

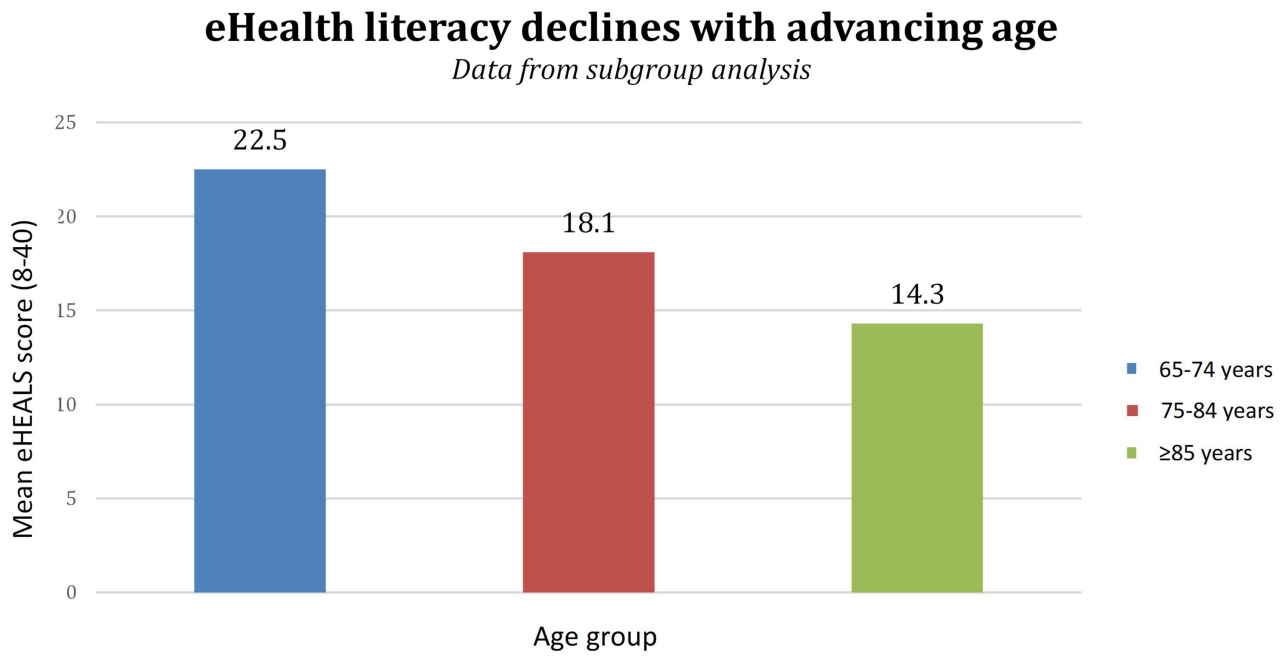
We analyzed the baseline data from an ongoing, large-scale, quasi-experimental, pre-post study to investigate the level of eHealth literacy and its associations with health behaviors, primary care service use, mental health, and physical health among Hong Kong Chinese older adults (N=6704). The results suggested that their eHealth literacy level was generally low. In addition, our findings showed that older adults with higher eHealth literacy were associated with positive health behaviors and better health-related outcomes at baseline. A key strength of this study is its examination of specific, modifiable health behaviors (eg, smoking, physical activity, and diet) and health outcomes (eg, mental health and NCD risks). This granular approach provides more actionable insights for public health interventions. For example, identifying a link between eHealth literacy and smoking status allows for more targeted health promotion campaigns. This specificity moves the field beyond broad associations toward evidence that can directly inform the content of programs that leverage digital tools for health improvement in aging populations.

The mean eHealth literacy score of the study participants was 18.2 (SD 10.2), which was lower than the scores reported in the United States (30.9) [16], South Korea (30.5) [17], Norway (25.7) [18], and China (21.4) [19]. The notably higher age (mean 77.8, SD 7.0 years) of our sample is likely a key contributor to this low average score, a conclusion supported by our finding that increasing age was the most significant demographic predictor of lower eHealth literacy. This suggests that interventions to enhance older adults' eHealth literacy are much needed. Our study indicated that 44.1% of older adults (mean age 77.8, SD 7.0 years) lacked eHealth literacy despite having a smartphone with internet access. This can be attributed to older adults' attitudes toward the internet, including self-efficacy in managing digital technology and a preference for in-person interaction with health practitioners [27]. eHealth literacy may also be lower due to age-related problems and cognitive decline [27]. These findings are relevant to delivering eHealth resources to older adults [27]. Overcoming these sociodemographic and psychological barriers may increase eHealth literacy and help older adults improve their general health. In addition, despite 93% (6227/6697) of participants owning a smartphone, 60.2% (3992/6634) reported never having used it to search for health information (Multimedia Appendix 1). This highlights a critical gap between device ownership and meaningful health-related use, underscoring that access to technology alone is insufficient. This "know-do" gap underscores the urgent need for interventions that build skills and confidence in searching for health information with technological devices among older adults.

Furthermore, we performed subgroup analyses of eHealth literacy by age group (65 - 74 years, 75 - 84 years, and ≥85 years) in our study sample (Figure 1). The results showed a stepwise decline in eHealth literacy across age strata. Older adults with increasing age had the highest burden of chronic disease and the greatest need for health care services, yet they are the most likely to be excluded from an increasingly digitized health care system. Our study's findings can serve as a data point for policymakers and health care providers, highlighting the urgent need for tailored, age-appropriate support systems to prevent the digital divide from becoming a health equity crisis.

XSL•FO
RenderX

Figure 1. Mean eHealth literacy Scale (eHEALS) score by age group in the study sample (N=6704).



Our findings showed that older adults with higher education were associated with higher eHealth literacy, and older adults with indicators of lower socioeconomic status (receiving support from the CSSA Scheme and living in public housing) were associated with lower eHealth literacy. Lower-income older adults may have poorer access to digital technology, such as limited computer and internet access, leading to skill disparity [28]. The relationship between socioeconomic status and eHealth literacy needs to be investigated in depth to prevent the potential loss of the silver economy in a digitalized society. It is important to consider the social disparities and the digital divide factors when promoting eHealth services to ensure that those needing these services are not left out [28]. Primary health care is the cornerstone of the health care system, and we should optimize the primary health care services to enhance older adults' eHealth literacy. In addition, training should be provided to carers and the younger generation (eg, university students) to guide underprivileged older adults in using electronic information technology and meeting their health needs in the digital age.

We found that older adults with moderate and high eHealth literacy were less likely to be smokers and more likely to engage in positive health behaviors. Higher eHealth literacy may enhance an individual's self-efficacy in managing their health. With an enhanced ability to find, evaluate, and apply online health information, older adults may feel more confident in making informed decisions [29], leading to better health behaviors such as improved diet or increased physical activity and more effective self-management of chronic conditions. In addition, greater access to eHealth resources increases the awareness of health maintenance and encourages older adults to adopt healthy lifestyles [30]. Older adults skilled in using digital health services will be able to communicate with health care professionals in advance and manage their health more efficiently, enabling them to receive more comprehensive NCD prevention services.

Our findings revealed a strong association between higher eHealth literacy and better mental health outcomes, specifically higher levels of well-being and lower levels of loneliness. This finding aligns with a growing body of literature demonstrating the positive role of digital engagement in the psychosocial health of older adults [31]. For instance, a study in Hong Kong during the COVID-19 pandemic found that the use of instant messaging apps was a key factor in mitigating loneliness [32]. Our results expand on this by suggesting that eHealth literacy, as a core component of digital competency, is a crucial enabler of such beneficial social connections. The mechanism is likely 2-fold. First, digitally literate older adults can more easily use technology to maintain contact with family and friends, thereby strengthening their social support networks. Second, those with low eHealth literacy may struggle to differentiate credible information from misinformation online [33], leading to increased anxiety and negative emotional states. The implication for practice is significant: interventions aimed at improving mental health in older adults should consider incorporating digital literacy training not merely as a technical skill but as a fundamental tool for fostering social connectedness and resilience in a digital world.

Implications and Future Directions

Based on the findings identified in this cross-sectional study, several avenues for future research are warranted to deepen our understanding and inform effective interventions. First, there is a critical need for longitudinal studies to establish temporal relationships between eHealth literacy and health outcomes. For instance, a recent 3-wave longitudinal study demonstrated that higher baseline eHealth literacy predicted a better health-promoting lifestyle over time among Chinese older adults [34]. Future research should build on this finding by following up older adults to determine whether improving eHealth literacy leads to sustained health behavior change and better long-term

XSL•FO
RenderX

health outcomes and to explore the mediating pathways (eg, self-efficacy) in these relationships.

Second, to move beyond association and test for causality, future work should include intervention studies, such as randomized controlled trials. These studies could evaluate the effectiveness of tailored eHealth literacy training programs on specific outcomes, including health behavior change, chronic disease self-management, and mental well-being in older adults. Such interventions should be multifaceted and co-designed with older adults, moving beyond basic technology access to build practical skills and confidence. Strategies could include (1) tailored, hands-on training workshops; (2) intergenerational peer-tutoring models; (3) the development of age-friendly user interfaces for health apps; and (4) the integration of eHealth literacy support within primary care settings and community centers.

Third, the measurement of eHealth literacy itself requires advancement. This study, similar to many others, relied on a self-report scale. Future research would benefit from incorporating objective, performance-based assessments. This would provide a more accurate measure of actual competency and help bridge the gap between perceived and demonstrated skills. Additionally, research is needed to validate eHealth literacy assessment tools and their cut-off scores for the older and oldest older populations.

Finally, to enhance generalizability, future research should include older adults with a history of mental illness. While they were excluded from our analysis to avoid confounding the mental well-being outcomes, research focus on this vulnerable group is essential for developing tailored interventions that address their specific barriers to digital engagement and health management [35], thus further narrowing the digital divide.

Limitations

This study has several limitations. First, the findings were based on an analysis of the baseline data, and causal relationships between variables could not be confirmed. Given the initial observational data, intervention studies (eg, randomized controlled trials) should be conducted to determine causal relationships between eHealth literacy, health behaviors, and health-related outcomes. Follow-up data and qualitative research may provide more insights into these mechanisms in the future. Second, our study excluded older adults with a self-reported history of mental illness (eg, depression and anxiety), and this exclusion may limit the generalizability of our findings. Mental health status may be associated with an individual's motivation and ability to engage with digital technology [31]. Further research is needed to investigate the relationships between eHealth literacy and health outcomes, specifically within this vulnerable subgroup of older adults. Third, individual eHealth literacy and health status are closely related to the local economic development level [36], and the study findings may not be generalizable to regions with different socioeconomic contexts, cultural backgrounds, or levels of digital infrastructure. In addition, the study was conducted in Hong Kong, a high-income region with advanced digital infrastructure. The observed levels of eHealth literacy and their associations with health outcomes may not be directly generalizable to older adults in low-income or middle-income settings, where access to technology and digital skills may be substantially different. Fourth, this study used the eHEALS to assess eHealth literacy. While widely used and validated, this scale has significant limitations. Developed in 2006 [11], eHEALS may not fully capture the complex skills required to navigate today's ecosystem of mobile apps and interactive health technologies [37]. Furthermore, newer instruments such as the Digital Health Literacy Instrument (DHLI) have been developed and show good utility in Chinese older adults [38]. Recent systematic reviews also highlight the need for more robust, performance-based tools, as eHEALS assesses self-perceived skills rather than actual competence, which may be subject to self-report bias [37,39]. Finally, the generalizability of our findings to the entire older adult population in Hong Kong may be limited. While our sample was large and weighted by sex distribution to match the Hong Kong census, the mean age of our participants was 77.8 (SD 7.0) years, which is higher than that of the general older adult population in Hong Kong [40].

Conclusions

eHealth literacy is an essential skill in a rapidly digitalizing world. The findings showed that the eHealth literacy of older adults in Hong Kong was low and needs improvement, especially in the context of global primary health care reform. We also observed that higher levels of eHealth literacy were associated with health-promoting behaviors, primary care service use, and better physical and mental health outcomes. With the application of the results from this study, tailored interventions should be implemented to improve eHealth literacy and narrow the digital divide among older adults.

Acknowledgments

We would like to thank the participants who enrolled in this study.

Funding

This research was supported by the Sino Group, Ng Teng Fong Charitable Foundation, China.

Data Availability

Data from this study will be made available upon reasonable request. For data access, please contact the corresponding author, Sophia Siu Chee CHAN (University of Hong Kong, Hong Kong).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of smartphone usage of participants (N=6704).
[DOCX File, 22 KB - aging_v9i1e74110_app1.docx]

Checklist 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist.
[PDF File, 91 KB - aging_v9i1e74110_app2.pdf]

References

1. Kassebaum NJ. Global, regional, and national burden of diseases and injuries for adults 70 years and older: systematic analysis for the Global Burden of Disease 2019 Study. BMJ 2022;376:e068208. [doi: 10.1136/bmj-2021-068208]
2. Lv J, Zhang ZF. Progress and challenges in NCD prevention and control in China. BMJ 2024 Oct 18;387:q2098. [doi: 10.1136/bmj.q2098] [Medline: 39424318]
3. Bai J, Cui J, Shi F, Yu C. Global epidemiological patterns in the burden of main non-communicable diseases, 1990-2019: relationships with socio-demographic index. Int J Public Health 2023;68:1605502. [doi: 10.3389/ijph.2023.1605502] [Medline: 36726528]
4. Petrova NN, Khvostikova DA. Prevalence, structure, and risk factors for mental disorders in older adults [Article in Russian]. Adv Gerontol 2021;34(1):152-159. [doi: 10.1134/S2079057021040093] [Medline: 33993676]
5. Fung AWT, Chan WC, Wong CSM, et al. Prevalence of anxiety disorders in community dwelling older adults in Hong Kong. Int Psychogeriatr 2017 Feb;29(2):259-267. [doi: 10.1017/S1041610216001617] [Medline: 27766997]
6. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/S2215-0366(21)00395-3]
7. Global momentum on primary health care: time to unite. World Health Organization. 2025 Jun 4. URL: https://www.who.int/news/item/04-06-2025-global-momentum-on-primary-health-care-time-to-unite [accessed 2025-10-07]
8. Stara V, Soraci L, Takano E, et al. Intrinsic capacity and active and healthy aging domains supported by personalized digital coaching: survey study among geriatricians in Europe and Japan on eHealth opportunities for older adults. J Med Internet Res 2023 Oct 12;25:e41035. [doi: 10.2196/41035] [Medline: 37824183]
9. Junaid SB, Imam AA, Balogun AO, et al. Recent advancements in emerging technologies for healthcare management systems: a survey. Healthcare (Basel) 2022 Oct 3;10(10):1940. [doi: 10.3390/healthcare10101940] [Medline: 36292387]
10. Fahy E, Hardikar R, Fox A, Mackay S. Quality of patient health information on the Internet: reviewing a complex and evolving landscape. Australas Med J 2014;7(1):24-28. [doi: 10.4066/AMJ.2014.1900] [Medline: 24567763]
11. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
12. Ye W. The impact of internet health information usage habits on older adults' e-health literacy. Digit Health 2024;10:20552076241253473. [doi: 10.1177/20552076241253473] [Medline: 38726215]
13. Tsai YIP, Beh J, Ganderton C, Pranata A. Digital interventions for healthy ageing and cognitive health in older adults: a systematic review of mixed method studies and meta-analysis. BMC Geriatr 2024 Mar 4;24(1):217. [doi: 10.1186/s12877-023-04617-3] [Medline: 38438870]
14. Kim K, Shin S, Kim S, Lee E. The relation between eHealth literacy and health-related behaviors: systematic review and meta-analysis. J Med Internet Res 2023 Jan 30;25:e40778. [doi: 10.2196/40778] [Medline: 36716080]
15. Xie L, Zhang S, Xin M, Zhu M, Lu W, Mo PKH. Electronic health literacy and health-related outcomes among older adults: a systematic review. Prev Med 2022 Apr;157:106997. [doi: 10.1016/j.ypmed.2022.106997] [Medline: 35189203]
16. Sanchez R, Kay H, Srikanth P, Sandow L, Zhang M. Health behaviors and eHealth literacy among older adults, HINTS 2019. Innov Aging 2020 Dec 16;4(Supplement_1):226-226. [doi: 10.1093/geroni/igaa057.730]
17. Yang E, Chang SJ, Ryu H, Kim HJ, Jang SJ. Comparing factors associated with eHealth literacy between young and older adults. J Gerontol Nurs 2020 Aug 1;46(8):46-56. [doi: 10.3928/00989134-20200707-02] [Medline: 32936926]
18. Brørs G, Wentzel-Larsen T, Dalen H, et al. Psychometric properties of the Norwegian version of the Electronic Health Literacy Scale (eHEALS) among patients after percutaneous coronary intervention: cross-sectional validation study. J Med Internet Res 2020 Jul 28;22(7):e17312. [doi: 10.2196/17312] [Medline: 32720900]
19. Jiang X, Wang L, Leng Y, et al. The level of electronic health literacy among older adults: a systematic review and meta-analysis. Arch Public Health 2024 Nov 7;82(1):204. [doi: 10.1186/s13690-024-01428-9] [Medline: 39511667]
20. He AWJ, Yuan R, Luk TT, Wang KMP, Chan SSC. Boosting digital health engagement among older adults in Hong Kong: pilot pre-post study of the Generations Connect Project. JMIR Form Res 2025 May 8;9:e69611. [doi: 10.2196/69611] [Medline: 40340793]

XSL•FO
RenderX

21. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8:115. [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
22. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176. [doi: 10.1159/000376585] [Medline: 25831962]
23. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
24. Du J, Jiang Y, Lloyd C, et al. Validation of Chinese version of the 5-item WHO well-being index in type 2 diabetes mellitus patients. BMC Psychiatry 2023 Nov 29;23(1):890. [doi: 10.1186/s12888-023-05381-9] [Medline: 38031007]
25. Liu T, Lu S, Leung DKY, et al. Adapting the UCLA 3-item loneliness scale for community-based depressive symptoms screening interview among older Chinese: a cross-sectional study. BMJ Open 2020 Dec 10;10(12):e041921. [doi: 10.1136/bmjopen-2020-041921] [Medline: 33303463]
26. Ghazi SN, Berner J, Anderberg P, Sanmartin Berglund J. The prevalence of eHealth literacy and its relationship with perceived health status and psychological distress during Covid-19: a cross-sectional study of older adults in Blekinge, Sweden. BMC Geriatr 2023 Jan 4;23(1):5. [doi: 10.1186/s12877-022-03723-y] [Medline: 36597040]
27. Jung SO, Son YH, Choi E. E-health literacy in older adults: an evolutionary concept analysis. BMC Med Inform Decis Mak 2022 Jan 31;22(1):28. [doi: 10.1186/s12911-022-01761-5] [Medline: 35101005]
28. Fung KK, Hung SSL, Lai DWL, Shum MHY, Fung HW, He L. Access to information and communication technology, digital skills, and perceived well-being among older adults in Hong Kong. Int J Environ Res Public Health 2023 Jun 23;20(13):6208. [doi: 10.3390/ijerph20136208] [Medline: 37444058]
29. Wang Y, Song Y, Zhu Y, Ji H, Wang A. Association of eHealth literacy with health promotion behaviors of community-dwelling older people: the chain mediating role of self-efficacy and self-care ability. Int J Environ Res Public Health 2022;19(10):6092. [doi: 10.3390/ijerph19106092]
30. Zhao YC, Zhao M, Song S. Online health information seeking among patients with chronic conditions: integrating the health belief model and social support theory. J Med Internet Res 2022 Nov 2;24(11):e42447. [doi: 10.2196/42447] [Medline: 36322124]
31. Ao SH, Zhu Y, Wang J, Zhao X. ehealth use and psychological health improvement among older adults: the sequential mediating roles of social support and self-esteem. Digit Health 2025;11:20552076251346659. [doi: 10.1177/20552076251346659] [Medline: 40487887]
32. Yang C, Lai DWL, Sun Y, Ma CY, Chau AKC. Mobile application use and loneliness among older adults in the digital age: insights from a survey in Hong Kong during the COVID-19 pandemic. Int J Environ Res Public Health 2022;19(13):7656. [doi: 10.3390/ijerph19137656]
33. Paige SR, Krieger JL, Stellefson ML. The influence of eHealth literacy on perceived trust in online health communication channels and sources. J Health Commun 2017 Jan;22(1):53-65. [doi: 10.1080/10810730.2016.1250846] [Medline: 28001489]
34. Xie L, Mo PKH. Prospective relationship between electronic health literacy and health-promoting lifestyle among Chinese older adults: a three-wave longitudinal study. Soc Sci Med 2025 Jan;364:117166. [doi: 10.1016/j.socscimed.2024.117166] [Medline: 39586136]
35. Tan G, Chao J, Jin S, et al. The association between digital health literacy and health inequalities among Chinese older adults: a multicenter cross-sectional study. Digit Health 2025;11:20552076251349901. [doi: 10.1177/20552076251349901] [Medline: 40547434]
36. Milanti A, Chan DNS, Parut AA, So WKW. Determinants and outcomes of eHealth literacy in healthy adults: a systematic review. PLoS One 2023;18(10):e0291229. [doi: 10.1371/journal.pone.0291229] [Medline: 37792773]
37. Lee J, Lee EH, Chae D. eHealth literacy instruments: systematic review of measurement properties. J Med Internet Res 2021;23(11):e30644. [doi: 10.2196/30644]
38. Xie L, Mo PKH. Comparison of eHealth Literacy Scale (eHEALS) and Digital Health Literacy Instrument (DHLI) in assessing electronic health literacy in Chinese older adults: a mixed-methods approach. Int J Environ Res Public Health 2023 Feb 13;20(4):3293. [doi: 10.3390/ijerph20043293] [Medline: 36833987]
39. Huang YQ, Liu L, Goodarzi Z, Watt JA. Diagnostic accuracy of eHealth literacy measurement tools in older adults: a systematic review. BMC Geriatr 2023 Mar 29;23(1):181. [doi: 10.1186/s12877-023-03899-x] [Medline: 36978033]
40. Population estimates. Census and Statistics Department.: The Government of the Hong Kong Special Administrative Region; 2020. URL: https://www.censtatd.gov.hk/en/scode150.html [accessed 2025-10-20]

Abbreviations

AOR: adjusted odds ratio
CSSA: Comprehensive Social Security Assistance
DHLI: Digital Health Literacy Instrument
eHEALS: eHealth Literacy Scale
MET: metabolic equivalent of task

XSL•FO
RenderX

NCD: noncommunicable disease
OR: odds ratio
WHO-5: World Health Organization-Five Well-Being Index

Edited by H Köttl; submitted 18.Mar.2025; peer-reviewed by E Maranesi, J Chen, L Xie; accepted 29.Dec.2025; published 30.Jan.2026.

Please cite as:
Chau SL, He W, Luk TT, Chan SSC
Level of eHealth Literacy and Its Associations With Health Behaviors and Outcomes in Chinese Older Adults: Cross-Sectional Analysis of Baseline Data From a Large-Scale Community Project
JMIR Aging 2026;9:e74110
URL: https://aging.jmir.org/2026/1/e74110
doi:10.2196/74110

© Siu Long Chau, Wanjia He, Tzu Tsun Luk, Sophia Siu Chee Chan. Originally published in JMIR Aging (https://aging.jmir.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

From Digital Anxiety to Empowerment in Older Adults: Cross-Sectional Survey Study on Psychosocial Drivers of Digital Literacy

Han-Jen Niu[1*], PhD; Ming-Hsuan Li[1*], MS; Feng-Yu Hsieh[2], PhD; Chun-Chieh Yu[3], PhD; Chun-Ting Lin[1], MS

[1]Management Science, College of Business and Management, Tamkang University, No. 151 Ying-Chung Rd., New Taipei City, Taiwan

[2]Physical Education Instruction and Activities Section, Office of Physical Education, Tamkang University, New Taipei City, Taiwan

[3]AgriSense Agri-Food Education Association, New Taipei City, Taiwan

[*]these authors contributed equally

Corresponding Author:
Han-Jen Niu, PhD
Management Science, College of Business and Management, Tamkang University, No. 151 Ying-Chung Rd., New Taipei City, Taiwan

Abstract

Background: Amid the convergence of global population aging and accelerating digital transformation, older adults' digital adaptability has emerged as a critical indicator of their quality of life, autonomy, and capacity for successful aging. However, digital disparities, technology-related anxiety, and insufficient support systems continue to hinder older individuals from fully participating in digital society. Particularly in modern family structures—where children often live apart from aging parents—the diminishing role of family support further underscores the importance of broader social influences.

Objective: This study aims to examine how environmental factors (family support and social influence) and psychological factors (digital anxiety and sense of achievement) are associated with older adults' intention to use Assistive Digital Tools and Services (ADTS), and how these relationships contribute to the development of digital literacy. Drawing upon an integrative framework that combines constructs from the Technology Acceptance Model, the Unified Theory of Acceptance and Use of Technology, and social cognitive theory, the study also investigates the mediating and moderating mechanisms underlying these effects, offering strategic insights to support older adults in moving from social isolation to digital empowerment.

Methods: A structured questionnaire survey was conducted using a convenience sampling method among adults aged 55 years and older in Shenyang, Liaoning Province, China, yielding 480 valid responses. Structural equation modeling, bootstrapping, and moderation analysis were used to test the proposed integrative framework.

Results: For both family support and social influence, their associations with digital literacy were fully mediated by ADTS. Higher family support was associated with lower digital anxiety, which in turn correlated with greater intention, while stronger social influence was directly associated with higher intention. Digital anxiety showed a strong negative association with intention; however, this relationship was significantly weaker among those reporting a higher sense of achievement. These findings highlight usage intention as a central pathway through which environmental and psychological conditions are related to digital competence.

Conclusions: Digital literacy in later life is more than a technical skill set—it represents a vital form of psychological and social capital that empowers autonomy, well-being, and social integration. Strengthening older adults' intention to engage with digital tools through emotional reinforcement, achievement-oriented experiences, and supportive social environments is key to narrowing the digital divide. Beyond its personal benefits, fostering digital competence contributes to successful aging, which in turn brings profound advantages for families, strengthens community cohesion, and supports national goals in public health, economic participation, and social sustainability. Intergenerational learning initiatives, community-based engagement programs, and leveraging social influence to offset weakened family support can create a more inclusive, resilient, and age-friendly digital ecosystem—one that benefits not only older individuals but society at large.

(JMIR Aging 2026;9:e75245) doi:10.2196/75245

KEYWORDS

digital literacy; successful aging; family support; social influence; technology acceptance; older adults; psychological moderation; digital anxiety

XSL•FO
RenderX

Introduction

Global Aging and the Challenge of Digital Inclusion

The intersection of rapid global population aging and accelerated digital transformation has rendered digital inclusion among older adults a pressing concern for researchers and policymakers alike. According to the United Nations [1], by 2050, the global population aged 65 years and older is projected to surpass 1.6 billion, accounting for nearly 16% of the total population. This demographic shift not only places strain on health care systems and labor markets but also redefines the societal roles and needs of aging populations in an increasingly digital world.

Within this context, digital literacy—encompassing not only operational skills but also digital confidence and information navigation—has been widely recognized as a key enabler for bridging this divide and promoting the effective use of Assistive Digital Tools and Services (ADTS) [2].

Prior studies have linked digital literacy to improved psychological well-being, life satisfaction, and learning motivation among older adults [3]. Yet, how digital literacy develops under the influence of environmental and psychological factors—especially in the context of aging—remains underexplored.

Digital exclusion exacerbates social isolation and loneliness among older adults, increasing the risk of depression and cognitive decline—while digital literacy offers a potential buffer against these outcomes [4]. In digitally mediated societies, older adults may experience compounded vulnerability—excluded not only socially but technologically.

Crucially, digital literacy holds promise as a remedy to both social and digital isolation. Older adults with strong digital competencies are more likely to engage in video communication, access health services remotely, and maintain active social networks [5,6]. In this sense, digital literacy is not merely a technical skill—it is a bridge to successful aging.

COVID-19: A Double-Edged Catalyst for Digital Transformation

The COVID-19 pandemic served as a global stress test for digital readiness, rapidly shifting key aspects of daily life—health care, communication, commerce—into digital spaces. While digital services such as telemedicine and online grocery delivery became lifelines for many, they also exposed and widened the digital gap among older populations [7].

Older adults demonstrated significantly lower intention to adopt digital services during and after the pandemic, largely due to digital anxiety, low confidence, and lack of digital trust [8]. Digital anxiety encompasses fears about making mistakes, information overload, and concerns about fraud and data breaches—factors particularly salient for older first-time users [9].

While digital literacy training can alleviate anxiety and improve adoption, sustained engagement often depends on ongoing emotional and environmental support—particularly from family and community contexts [10]. However, the long-term sustainability of digital engagement often depends on external support structures such as family guidance or community-based learning. Without sustained environmental and emotional support, even trained users may regress into avoidance behaviors. This underscores the importance of examining how family and social influences shape not only technology adoption but also digital confidence and persistence among older adults.

Research Gaps and Objectives

While widely adopted models such as the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) have contributed significantly to understanding digital behavior, they were primarily developed for younger, working-age users [11]. As such, these models often overlook emotional vulnerabilities and contextual factors that are particularly salient among older adults—such as digital anxiety, family support, and social influence.

At the same time, the Digital Literacy Framework highlights competencies essential for navigating the digital world but does not explicitly address how these skills develop within psychosocial environments. Moreover, existing studies tend to examine these psychological and environmental variables in isolation, without a cohesive explanatory structure [5,11].

To address this gap, we propose an integrative framework that draws on TAM, UTAUT, and social cognitive theory (SCT) [12]. In this model:

- Family support and social influence reflect environmental enablers of technology adoption (UTAUT);
- Digital anxiety and sense of achievement represent emotional and cognitive mechanisms (SCT);
- Usage intention (from TAM) serves as a behavioral mediator;
- Digital literacy is positioned as a dynamic outcome representing both skill and empowerment.

Furthermore, we incorporate a moderated mediation structure, positing that sense of achievement may buffer the negative impact of digital anxiety on usage intention, thereby shaping digital literacy development [12].

This integrative approach also aligns with Successful Aging Theory, framing digital literacy as a form of “aging capital” that enhances autonomy, social participation, and psychological well-being.

Based on this framework, our study addresses the following research questions:

1. To what extent are family support and social influence associated with older adults’ intention to adopt ADTS?
2. Does digital anxiety negatively affect usage intention, and is this effect moderated by sense of achievement?
3. Does usage intention mediate the relationship between environmental/psychological factors and digital literacy?

By empirically validating this multilevel framework, we aim to offer new insights into how older adults move from digital exclusion to empowerment in an aging digital society.

XSL•FO
RenderX

In applying UTAUT as a guiding framework, we intentionally retained social influence and conceptualized family support as an environmental enabler, while excluding performance expectancy and effort expectancy. We do so for parsimony and contextual salience: (1) their predictive roles have been robustly established in prior work, so retesting them would add replication rather than novel insight; (2) prestudy interviews suggested many older adults lacked sufficient hands-on ADTS experience to meaningfully assess usefulness or ease of use; and (3) our theoretical focus is the age-specific psychosocial mechanisms (family support, social influence, anxiety, and achievement) that drive intention in later life. This selective adaptation contextualizes UTAUT rather than diluting it.

We also model digital literacy as an outcome—an accumulated form of "aging capital"—to trace how environmental and psychological conditions translate into competence via usage intention. While prior studies have treated literacy as an antecedent that can reduce anxiety and strengthen intention, our outcome-focused specification clarifies the developmental pathway we test here; it does not preclude reciprocal dynamics, which we note as an avenue for longitudinal research.

Literature Review

Family Support

Family support plays a critical role in facilitating older adults' digital engagement. Defined broadly, family support encompasses emotional encouragement, technical guidance, and intergenerational interaction that collectively enhance older adults' adaptation to digital life [13]. Caplan [14] viewed the family as a key provider of values and behavioral norms, offering essential mediation when individuals encounter challenges. Casper et al [15] further divided family support into instrumental, financial, and emotional dimensions, all of which contribute meaningfully to older adults' ability to learn and apply new technologies.

In the digital context, family members frequently act as facilitators in older adults' learning process, offering real-time help and reassurance [16]. Older adults are often influenced by their children's or grandchildren's attitudes toward technology, making family encouragement a significant determinant of their intention to engage with ADTS [17].

Empirical research supports this relationship. Xiong and Zuo [13] showed that emotional support and technical instruction from family members enhance digital literacy by reducing fear and uncertainty. Similarly, Meng et al [18] and Sosa Díaz [19] found that positive intergenerational communication increases digital confidence and learning motivation. Sosa Díaz [19] highlighted that assistance from younger family members strengthens adaptability, while Roman et al [20] emphasized how strong family bonds can boost learning confidence.

Thus, this study highlights family support as a foundational environmental factor influencing both psychological readiness and behavioral engagement in digital contexts. The focus on intergenerational interaction aims to deepen our understanding of how familial dynamics contribute to digital inclusion among older adults.

Social Influence

Social influence refers to the degree to which individuals' behaviors, attitudes, or decisions are shaped by those in their social environment—such as family, friends, or peers [11]. Rooted in the UTAUT, social influence is recognized as a key driver of technology adoption, particularly through normative pressure and perceived expectations from significant others [11,21].

From a broader psychological perspective, social influence encompasses beliefs, emotions, and behavioral patterns formed through interactions within social networks [21]. Morosan et al [22] highlighted that individuals' perceptions of others' expectations can significantly affect their own technology acceptance decisions. Karahanna et al [21] expanded this view by discussing how proximity, contact frequency, and interpersonal dynamics influence behavior, emphasizing that social structures can either facilitate or inhibit digital engagement.

Among older adults, social influence is particularly potent. Compared with younger individuals, older adults are more likely to rely on interpersonal cues and social norms when evaluating new technologies [11]. The impact of social influence is also more pronounced in collectivist cultures—such as in many East Asian societies—where conformity to group norms and maintaining social harmony are especially valued.

The role of social influence has become even more salient in the postpandemic digital era. During COVID-19, social distancing measures drove many older adults to adopt digital tools for health, communication, and daily tasks, often under the encouragement or guidance of their social circles [23]. Positive social reinforcement can lead to greater confidence and higher willingness to engage with digital services, whereas skepticism or lack of support may result in avoidance or anxiety [24-26].

Furthermore, recent studies show that social influence not only affects direct behavioral intention but also moderates psychological factors such as digital anxiety and self-efficacy [27]. These findings underscore the need to understand social influence as both an external motivator and a psychological buffer or amplifier in older adults' digital adaptation.

By focusing on the multifaceted nature of social influence—including social norms, peer encouragement, and perceived expectations—this study aims to clarify how social context shapes older adults' engagement with ADTS and the broader development of digital literacy.

Digital Anxiety: Concept and Development

Digital anxiety refers to the emotional discomfort, fear, or stress that individuals—particularly older adults—experience when interacting with digital technologies. Unlike simple unfamiliarity, digital anxiety reflects a deeper psychological resistance often rooted in low confidence, fear of failure, and perceptions of complexity or risk [26,28]. Wang and Zhang [29] defined it as a form of state anxiety, varying with task demands and context. Among older adults, this anxiety is

XSL•FO
RenderX

especially pronounced due to the widening digital divide and their limited prior exposure to emerging technologies.

Rook [30] highlighted that older individuals frequently report elevated anxiety when using ADTSs such as telemedicine platforms, smartphones, or mobile payments. This is compounded when they lack the digital self-efficacy or training needed to operate these tools confidently. Straub [26] found that prior experience and perceived self-efficacy can significantly reduce anxiety and increase willingness to adopt digital tools. Similarly, the TAM posits that perceived ease of use and usefulness directly shape user attitudes and behaviors.

From the perspective of Rogers' Diffusion of Innovations Theory, perceived complexity and compatibility with past experiences are key barriers for older adults. Even simple digital interfaces can feel cognitively taxing, potentially triggering avoidance behaviors rather than active engagement [31,32].

One of the most consistently supported buffers against digital anxiety is family support. Beyond emotional reassurance, family members often provide practical technical guidance and encouragement, helping older adults feel less overwhelmed and more motivated to engage with digital tools. Studies have shown that such support fosters emotional security, strengthens digital self-confidence, and decreases feelings of uncertainty in digital environments [33]. For instance, when family members offer direct assistance—such as walking through digital tasks step-by-step—older adults are more likely to persist in their learning process and overcome initial apprehension.

However, overly reliant support can have mixed effects. While moderate, empowering support improves outcomes, excessive dependence may inadvertently signal incompetence or fuel learned helplessness, thereby reinforcing anxiety [7,34]. Nevertheless, the preponderance of evidence suggests a net protective role of family involvement in alleviating digital stress, leading to the following hypothesis:

H1: Family support negatively influences digital anxiety.

In addition to family, broader social influence plays a crucial role in shaping digital anxiety. When older adults perceive positive expectations or encouragement from peers, neighbors, or community members, it can reduce fear and promote digital exploration. On the contrary, negative social feedback—such as expressions of doubt, impatience, or age-related stereotypes—can heighten self-doubt and anxiety. Recent studies have demonstrated that social environments that are judgmental or unsupportive intensify older adults' digital apprehension, especially when they fear being seen as "incompetent" or "too old to learn" [35,36].

Conversely, positive social modeling and group-based learning environments—such as community digital workshops—can boost older adults' digital confidence and reduce anxiety levels. Venkatesh et al [11] and Cambre and Cook [27] have also pointed out that perceived social expectations (a key dimension of UTAUT) can influence both emotional responses and behavioral intentions related to technology.

H2: Social influence positively influences digital anxiety.

In summary, digital anxiety among older adults is shaped not only by individual cognitive appraisals but also by the presence—or absence—of emotional and environmental support. Understanding how family support and social influence interact with psychological states like anxiety is essential for building more inclusive and effective digital interventions.

Usage Intention of Assistive Digital Tools and Services

The intention to use ADTS among older adults is a central construct in models of technology adoption. Usage intention refers to an individual's motivational readiness and willingness to adopt or engage with a specific technology [37]. In the context of aging populations, intention plays a critical intermediary role, linking psychological and environmental factors with actual digital behavior and literacy development.

Family Support and Usage Intention

Family support has been widely documented as a catalyst for technology acceptance among older adults. Beyond alleviating anxiety, supportive family environments contribute to a more proactive stance toward digital learning. Emotional encouragement, technical guidance, and shared experiences with family members strengthen older adults' belief in their own capability, thereby enhancing their readiness to use ADTS.

Studies suggest that family members who model digital behavior or provide hands-on help not only increase access but also shape attitudes of usefulness and ease of use—2 key predictors in the TAM [38]. For instance, Zhang [17] found that the opinions of close family members significantly influenced older adults' willingness to try new technologies. Likewise, Selwyn [16] and Park [39] noted that active family involvement enhances digital confidence and curiosity, which translates into a stronger intention to use tools like mobile banking or telemedicine services.

While over-dependence on family for digital engagement may hinder independent learning, moderate and empowering support appears to foster technology acceptance. Therefore, based on the strong empirical and theoretical link between family support and technology adoption, the following hypothesis is proposed:

H3: Family support positively influences older adults' intention to use ADTS.

Social Influence and Usage Intention

Social influence extends beyond the family unit to include peers, community members, and broader societal norms that shape older adults' willingness to engage with digital technologies. In UTAUT, social influence is considered a direct antecedent of usage intention, reflecting how individuals perceive the expectations and behaviors of others as relevant to their own decision-making [11].

Research demonstrates that when older adults observe peers successfully using ADTS, they are more likely to emulate those behaviors, especially in collectivist cultures where conformity and social approval are emphasized [40,41]. Additionally, informational and normative forms of social influence—such as digital learning groups, community workshops, or

XSL•FO
RenderX

word-of-mouth recommendations—play a crucial role in shaping perceived accessibility and relevance of digital services [42,43].

Moreover, positive social reinforcement enhances self-efficacy, reduces uncertainty, and increases perceived behavioral control—factors that strongly predict technology use. Conversely, negative perceptions within one's social network may lead to skepticism or hesitation. Given this evidence, we propose:

H4: Social influence positively influences older adults' intention to use ADTS.

Digital Anxiety and Usage Intention

Digital anxiety, as established in prior sections, serves as a psychological inhibitor that can undermine willingness to engage with ADTS. Fear of making mistakes, concerns over privacy, and low digital confidence create a barrier that suppresses both motivation and perceived ability.

Drawing from TAM and SCT, when anxiety increases, perceived ease of use diminishes, leading to lower adoption intention [3,9]. In real-world contexts, this means older adults who feel overwhelmed or unsupported are more likely to avoid digital tools, even when such tools offer substantial benefits.

Studies in health care technology and fintech show consistent negative correlations between digital anxiety and usage intention [44,45]. Lee et al [9] emphasized that enhancing digital self-efficacy can reduce anxiety and indirectly boost usage intention, but when anxiety remains unaddressed, intention significantly declines. Given this robust evidence, we hypothesize:

H5: Digital anxiety negatively influences older adults' intention to use ADTS.

Digital Literacy

Digital literacy is a multidimensional competency that encompasses not only the ability to use digital tools, but also to evaluate, create, and communicate digital content effectively [46,47]. It includes technical, cognitive, and socioemotional skills, all of which are crucial for navigating digital environments [48,49]. In aging populations, digital literacy has emerged as a key determinant of autonomy, participation, and well-being [50].

Research has shown that the intention to use digital tools often precedes and fosters the development of digital literacy. When older adults are motivated to use digital tools—such as smartphones, e-payment systems, or telehealth platforms—they are more likely to develop the necessary skills and confidence through experiential learning [51,52]. This iterative process suggests that usage intention is not merely an outcome but also a catalyst in digital literacy acquisition.

H6: Usage intention of ADTS positively influences digital literacy.

Beyond direct effects, usage intention also serves as a mediating mechanism through which environmental support influences digital literacy. For instance, when family members provide technical guidance and emotional encouragement, they boost older adults' confidence and motivation. This, in turn, leads to increased digital usage, which fosters skill development and digital empowerment [38,53].

H7: Usage intention of ADTS mediates the relationship between family support and digital literacy.

Similarly, social influence plays a significant role in digital skill development. Older adults who are embedded in digitally engaged social networks are more likely to receive informal training, tips, and positive reinforcement, which enhance both their usage intention and their digital competence [35,41,42]. This cascading effect suggests that social influence indirectly contributes to digital literacy via behavioral intention.

H8: Usage intention of ADTS mediates the relationship between social influence and digital literacy.

Sense of Achievement

Sense of achievement refers to the intrinsic satisfaction and psychological reward individuals experience upon successfully accomplishing a task, overcoming a challenge, or attaining a self-defined goal [54,55]. This construct is closely tied to self-efficacy, motivation, and emotional resilience—factors that are especially relevant when individuals are navigating unfamiliar or cognitively demanding environments.

For older adults adapting to digital life, a sense of achievement plays a pivotal role. Unlike younger users who often grow up immersed in digital environments, older adults frequently approach technology as late adopters. When they successfully use digital tools—such as conducting a video call, managing a mobile payment, or accessing web-based health services—this sense of accomplishment becomes a powerful motivator for further learning and sustained engagement [56]. It also fosters autonomy and contributes to their psychological well-being, aligning with key principles of successful aging.

However, the digital environment presents numerous barriers that can threaten or diminish this feeling of accomplishment. These include perceived complexity, lack of intuitive design, cybersecurity concerns, and insufficient support systems [57]. When older adults encounter repeated failure or confusion, their confidence may erode, reinforcing negative beliefs about their ability to master technology. In this context, a diminished sense of achievement can magnify digital anxiety and lead to avoidance behaviors.

Recent studies suggest that a strong sense of achievement can buffer the negative effects of digital anxiety. Older adults who perceive themselves as capable learners are more likely to persist through technological challenges and less likely to internalize failure as a reflection of their overall competence [9]. In contrast, those who lack this sense of accomplishment are more prone to technostress and self-doubt—factors that reduce usage intention [58].

Given these dynamics, this study positions sense of achievement as a moderator in the relationship between digital anxiety and the intention to use ADTS. Specifically, we propose that achievement-oriented individuals will experience a weaker negative effect of digital anxiety on usage intention, as their

XSL•FO
RenderX

internal sense of progress and mastery helps to offset fear and hesitation.

H9: Sense of achievement moderates the relationship between digital anxiety and usage intention of ADTS, such that the negative effect is weaker at higher levels of achievement.

Methods

Participants

This study focuses on older adults in Shenyang, Liaoning Province, China, considering the region's demographic aging trends and digital adaptation status. As one of the major cities in Northeast China, Shenyang has a rapidly aging population, with individuals aged 60 years and older accounting for over 20% of its residents, meeting the United Nations' definition of a "deeply aging society"[96]. With the advancement of digital technologies, the internet behaviors and technological adaptation of older adults have become critical research topics. However, despite Shenyang's well-developed urban infrastructure, older adults in the region still face significant challenges in using ADTS (smartphones and e-services, etc).

Previous research has highlighted the existence of a digital divide among China's aging population, which is particularly pronounced in the northeastern region. Chu (2010) further pointed out that older individuals' digital learning and technology adaptation are influenced by educational background, social support, and technology anxiety, all of which contribute to lower adoption rates of e-services. These factors may limit older adults' ability to access web-based resources, subsequently affecting their digital quality of life [13].

The socioeconomic structure of Shenyang also plays a crucial role in shaping older adults' digital behaviors. As a traditional industrial city, Shenyang has a significant population of retirees from state-owned enterprises, whose social support networks and digital behavior patterns differ from those in coastal cities or rural areas [5]. Research by Chen et al [5] has further demonstrated a correlation between older adults' socioeconomic status, online activity types, and susceptibility to digital fraud, emphasizing the importance of digital literacy in ensuring their safety on the web. Given these factors, this study selects Shenyang as its research site to gain deeper insights into older adults' behavioral patterns and psychological influences in digital environments.

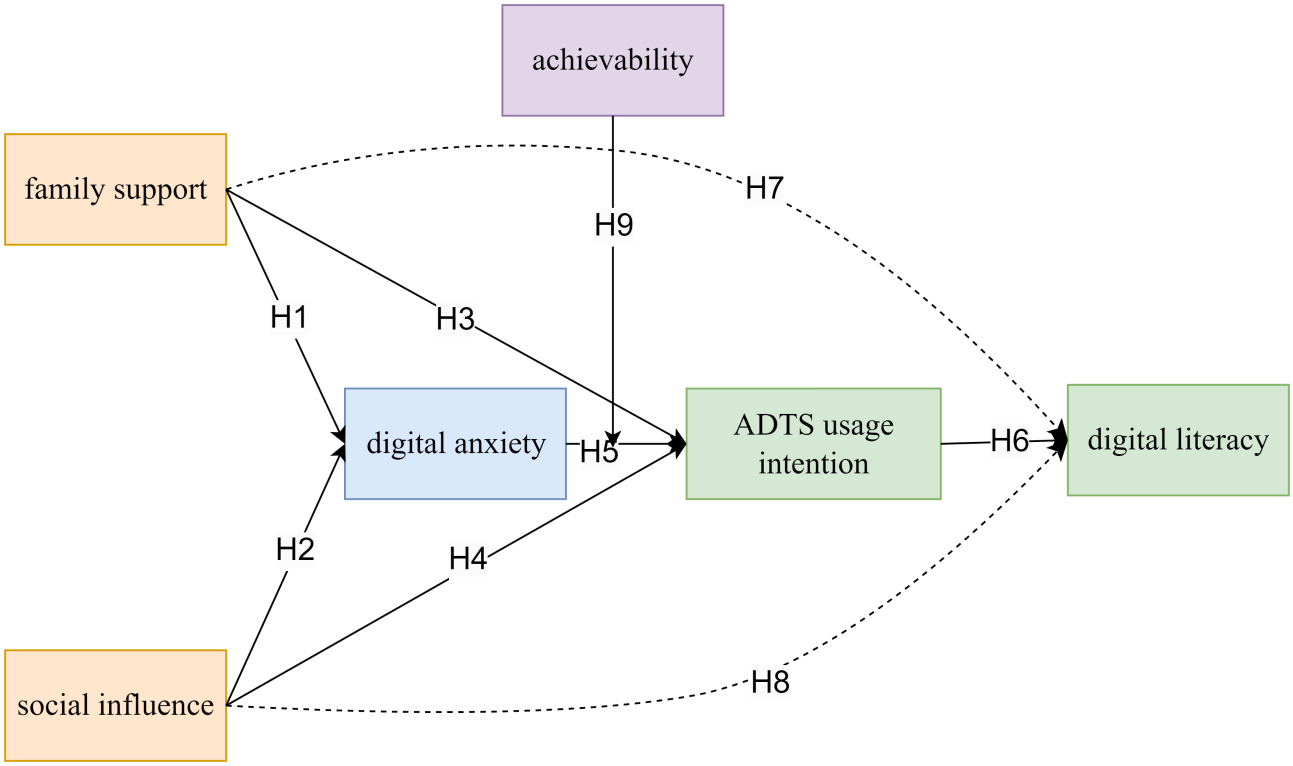
Additionally, this study includes individuals aged 55 years and older rather than limiting the sample to traditional older adult groups. There are 2 key reasons for this decision. First, individuals in this age group are rapidly adopting internet technologies, making them a highly relevant demographic for this study. Second, early retirement is common in China, with many individuals experiencing retirement before the official retirement age. By incorporating participants aged 55 years and older, this study effectively captures the early stages of aging and the transition into digital adaptation [5].

This study used a convenience sampling method, with surveys distributed across community centers, senior activity venues, and elderly care institutions in Shenyang. Individuals aged 55 years and older were invited to voluntarily complete the questionnaire. A total of 541 responses were collected. After excluding 61 incomplete questionnaires, 480 valid responses were retained, resulting in an effective valid completion rate among respondents of 88.7%. All participants were residents of Shenyang, Liaoning Province, China.

Demographic data were collected on participants' gender, age, education level, marital status, family structure, place of residence, and smartphone usage behavior. To ensure accurate comprehension and response quality, the research team provided detailed instructions and on-site assistance during the survey process. The conceptual framework guiding this study is presented in Figure 1.

XSL•FO
RenderX

Figure 1. Research Framework. ADTS: Assistive Digital Tools and Services.



Measures

This study primarily investigates the psychological states and digital behaviors of older adults. A 6-point Likert scale was used, ranging from "strongly disagree" to "strongly agree," to measure key variables, including social influence, family support, digital anxiety, digital literacy, adoption intention of ADTS, and sense of achievement. The measurement instruments were derived from well-established scales in the existing literature:

Family support was adapted from Wang and Wu [28], originally developed in the context of older adults' digital adaptation in urban China. The scale assesses 3 core dimensions of support provided by family members: emotional encouragement, technical assistance, and informational guidance. While Wang and Wu's [28] study focused on digital literacy outcomes, the psychometric properties of the family support scale were independently validated through confirmatory factor analysis (CFA) and internal consistency measures (Cronbach α=0.82). This scale was chosen for its cultural contextual relevance and conceptual alignment with our study focus. In our sample, Cronbach α was 0.802.

Social influence was adapted from Venkatesh et al [11] in the UTAUT. The scale evaluates the extent to which individuals perceive that important others—such as family members, peers, or community figures—believe they should use digital technologies. Each item was rated on a 6-point Likert scale. This construct has been validated in numerous technology acceptance studies across diverse populations, including older adults. In this study, the scale demonstrated good internal consistency (Cronbach α=0.798).

Digital anxiety was also adapted from Venkatesh et al [11]. The items reflect the UTAUT construct of anxiety, which captures individuals' apprehension, fear, or discomfort when interacting with digital technology. Although originally tested in younger user populations, the scale has been applied and adapted in aging research. In this study, internal consistency was acceptable (Cronbach α=0.762), supporting its reliability among older adult users.

Intention to use ADTS was adapted from the UTAUT model (Venkatesh et al [11]), assessing respondents' willingness to engage with digital services such as e-health platforms, digital payments, and digital communication tools. Each item was rated on a 6-point Likert scale. The original scale has been validated in both general and older adult populations. In this study, Cronbach α was .789, indicating acceptable reliability.

Digital literacy was evaluated using the eHealth Literacy Scale developed by Norman and Skinner [59], consisting of 8 items that measure perceived ability to find, understand, evaluate, and apply digital health information. This scale has been widely adopted in aging, health, and digital divide studies and has demonstrated strong psychometric properties across diverse populations. In our study, the scale yielded a Cronbach α of 0.871, indicating high internal consistency.

Sense of achievement was measured using a 3-item subscale from the achievement motivation framework developed by Janke and Lüftenegger [50]. The items reflect individuals' feelings of satisfaction and competence when overcoming digital challenges. The original instrument was validated in academic and motivational contexts and has since been adapted in studies of digital learning among older adults. The internal consistency for this scale in our sample was strong (Cronbach α=0.851).

XSL•FO
RenderX

To ensure the statistical adequacy of structural equation modeling (SEM), this study adhered to the sample size recommendations proposed by Bentler and Chou [60], which suggest that the sample size should be at least 5 times the number of estimated parameters. Additionally, data screening procedures, including missing data analysis and normality tests, were conducted to ensure the suitability of the dataset for subsequent analysis. Ultimately, a total of 480 valid responses were collected, meeting the theoretical requirements for robust statistical analysis.

Reliability (Cronbach α), CFA, and SEM were conducted using SPSS (version 26.0; IBM Corp) and AMOS (version 24.0; IBM Corp). Bootstrapping (n=5000) and PROCESS Macro (Model 1) were used to test mediation and moderation effects, ensuring robust statistical validity.

Ethical Considerations

This study utilized fully anonymized survey data that contained no personally identifiable information. As such, ethical approval and informed consent were not required under institutional and national research ethics regulations. All research procedures adhered to internationally recognized ethical standards, ensuring transparency, data integrity, and the responsible handling of anonymized information.

Results

Sample Profile

Regarding gender distribution, 69% (n=331) of respondents were male, and 31% (n=149) were female. In terms of age, the majority were aged between 60 and 65 years (263/480, 54.8%), followed by those aged 65 to 70 years (167/480, 34.8%), while only 1.0% (n=5) were aged 75 years and older. Residential distribution indicated that 60.4% (n=290) of participants lived in urban areas, whereas 39.6% (n=190) resided in rural regions. In terms of educational attainment, 57.1% (n=274) of respondents had completed middle school or below, 23.5% (n=111) held a university degree, and 1.5% (n=7) had attained a postgraduate degree or higher. Regarding marital status, 87.5% (n=420) of respondents were married, 4.8% (n=23) were unmarried, 2.9% (n=14) were divorced, and 4.8% (n=23) were widowed. As for family structure, 74.8% (n=359) lived with their spouse, 12.5% (n=56) lived with their children, and 12.7% (n=61) lived alone.

Regarding the usage of ADTS, 85.0% (n=408) of respondents primarily used smartphones for voice calls, followed by text messaging (381/480, 79.4%) and video calls (357/480, 74.4%). In contrast, mobile shopping (261/480, 54.3%) and digital payments (271/480, 56.5%) were less frequently used. Notably, only 53.4% (n=256) of participants used their smartphones for digital information retrieval, suggesting that a considerable portion of older adults remains unfamiliar with internet-based search functions.

Reliability Analysis

To assess the internal consistency of the measurement scales, this study calculated Cronbach α for each variable. The results, presented in Table 1, indicate that all Cronbach α values exceeded the recommended threshold of 0.7, demonstrating good reliability across the constructs.

Table . The results of reliability analysis (N=480). All constructs were measured on a 6-point Likert scale (1=strongly disagree, 6=strongly agree).

Variable	Values, mean (SD)	Values, Cronbach α
Social influence	2.850 (1.023)	0.798
Family support	4.960 (0.803)	0.802
Digital anxiety	4.980 (0.784)	0.762
ADTS[a] usage intention	5.030 (0.844)	0.789
Digital literacy	4.510 (0.919)	0.871
Achievability	5.380 (0.817)	0.851

[a]ADTS: Assistive Digital Tools and Services.

Among the variables, Digital Literacy exhibited the highest internal consistency, with a Cronbach α of 0.871, suggesting strong reliability. Meanwhile, Digital Anxiety had a Cronbach α of 0.762, which remains within an acceptable range, indicating sufficient internal consistency for further analysis.

Confirmatory Factor Analysis

To ensure convergent validity and discriminant validity of the measurement model, this study conducted CFA. Following the recommendations of Hair et al [61], composite reliability should exceed 0.7, and average variance extracted should be above 0.5 to confirm the internal consistency and validity of the measurement instruments.

As shown in Table 2, all constructs in this study met these criteria, with composite reliability values exceeding 0.7 and average variance extracted values above 0.5, indicating strong convergent validity. Additionally, all standardized factor loadings were statistically significant (>0.60), demonstrating that the observed variables effectively captured their respective latent constructs.

XSL•FO
RenderX

Table . Results of convergent validity analysis (N=480).

Concept and item	Standard factor loading	Composite reliability	Average variance extracted
Social influence (SI)		0.802	0.506
SI1	0.692		
SI3	0.818		
SI4	0.631		
F1	0.670		
Family support (F)		0.803	0.504
F3	0.705		
F4	0.748		
F5	0.716		
D2	0.586		
Digital literacy (D)		0.872	0.535
D4	0.690		
D5	0.696		
D6	0.808		
D7	0.798		
D8	0.785		
ADTS1[a]	0.765		
ADTS usage intention (ADTS)		0.802	0.578
ADTS2	0.865		
ADTS3	0.633		
DA1	0.704		
Digital anxiety (DA)		0.803	0.506
DA2	0.732		
DA3	0.729		
DA4	0.678		
A1	0.804		
Achievability (A)		0.880	0.711
A2	0.950		
A3	0.765		

[a]ADTS: Assistive Digital Tools and Services.

Overall, the results confirm that the measurement tools used in this study exhibit high reliability and validity, ensuring their suitability for further SEM analysis.

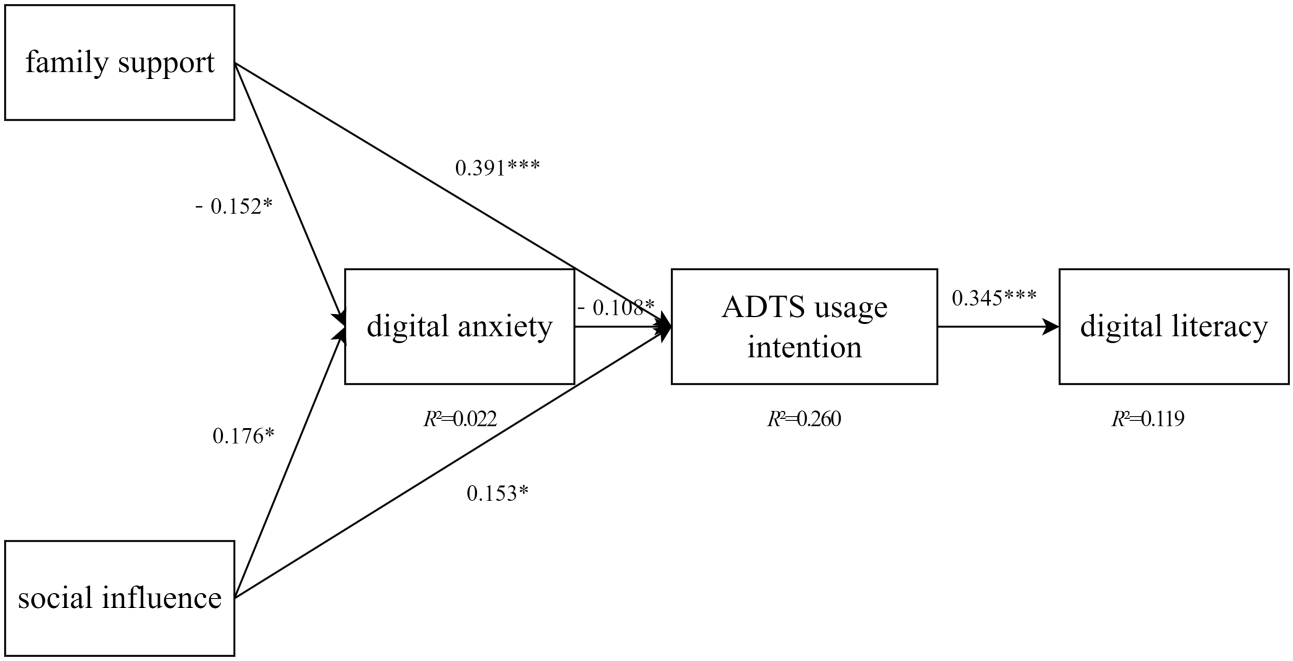
Structural Equation Modeling (SEM) Analysis

Model Fit Assessment

This study used SEM to validate the proposed research framework, as illustrated in Figure 2. Prior to conducting SEM analysis, we assessed the model fit indices to ensure the suitability of the model.

XSL•FO
RenderX

Figure 2. The model of structural equation modeling. ADTS: Assistive Digital Tools and Services. *P<.001, **P<.01, ***P<.05.



By examining the Modification Index, we identified and removed items with large residual discrepancies, including 2 items from the Family Support scale and items 1 and 3 from the Digital Literacy scale. After these modifications, the revised model exhibited improved goodness-of-fit indices, confirming its appropriateness for further SEM analysis (Table 3).

XSL•FO
RenderX

Table . Goodness of model fit indices (N=480).

Goodness of Fit Index	Scholars	Goodness fit range	Modified model (Delete SI2, D1, D3)
CMIN/DF[a]	Joreskog and Sorbom [52]	<3.00	2.499
GFI[b]	Doll, Xia, and Torkzadeh [53]	>0.8	0.925
AGFI[c]	MacCallum and Hong [62]	>0.8	0.902
SRMR[d]	Hu and Bentler [63]	<0.08	0.057
RMSEA[e]	MacCallum, Browne, and Sugawara [54]	<0.08	0.053
NFI[f]	Hair et al [61]	>0.8	0.902
NNFI[g]	Bentler and Bonett [55]	>0.8	0.884
CFI[h]	Hair et al [61]	>0.8	0.941
PNFI[i]	Mulaik et al [64]	>0.5	0.751
PGFI[j]	Mulaik et al [64]	>0.5	0.705
CN[k]	Mulaik et al [64]	>200	259

[a]CMIN/DF: chi-square minimum/degrees of freedom.
[b]GFI: Goodness-of-Fit Index.
[c]AGFI: Adjusted Goodness-of-Fit Index.
[d]SRMR: Standardized Root Mean Square R.
[e]RMSEA: Root Mean Square Error of Approximation.
[f]NFI: Normed Fit Index.
[g]NNFI: Non-Normed Fit Index.
[h]CFI: Comparative Fit Index.
[i]PNFI: Parsimony Normed Fit Index.
[j]PGFI: Parsimony Goodness-of-Fit Index.
[k]CN: Hoelter's Critical N.

Structural Equation Modeling (SEM) Analysis

The SEM analysis in this study was conducted in 2 stages. First, we examined the overall model fit indices to ensure the model's suitability. Second, we tested the path coefficients to validate the research hypotheses.

By performing path analysis on the latent variables, we assessed the relationships among key constructs. The results are presented in Table 4 and Figure 2.

Table . The results of structural equation modeling path analysis.

Hypothesis	Independent variable	Dependent variable	β	SE	Composite reliability	*P* value
H1	F[a]	A[b]	−.152	0.080	–2.087	<.037
H2	SI[c]	A	.176	0.081	2.362	<.018
H3	F	ADTS[d]	.391	0.077	5.356	<.001
H4	SI	ADTS	.153	0.071	2.235	<.025
H5	A	ADTS	−.108	0.049	–2.112	<.035
H6	ADTS	D[e]	.345	0.041	6.031	<.001

[a]F: Family support.
[b]A: Achievability.
[c]SI: Social influence.
[d]ADTS: Assistive Digital Tools and Services.
[e]D: Digital literacy.

XSL•FO
RenderX

Mediating Effect

This study used the PROCESS macro in SPSS (version 26.0; IBM Corp) with the bootstrap method to examine mediating effects. The analysis tested whether usage intention of ADTS mediated the relationships between (1) family support and digital literacy, and (2) social influence and digital literacy.

For the family support → digital literacy pathway (Table 5), the indirect effect was 0.071, with a 95% CI of (0.095-0.262), indicating statistical significance as the interval did not include zero. In contrast, the direct effect was 0.397, with a 95% CI of (–0.130 to 0.082), which was nonsignificant. This pattern reflects full mediation, meaning that family support influences digital literacy entirely through its impact on usage intention.

Table . Mediating effect analysis (family support and digital literacy).

F→ADTS→D[a,b,c]	Effect	Bootstrapping (BC[d] 95% CI)
Indirect effect	0.071	0.095 to 0.262
Direct effect	0.397	–0.130 to 0.082
Total effect	0.468	0.003 to 0.052

[a]F: family support.

[b]ADTS: Assistive Digital Tools and Services.

[c]D: digital literacy.

[d]BC: bias-corrected percentile method.

For the social influence → digital literacy pathway (Table 6), the indirect effect was also significant (95% CI excluding zero), whereas the direct effect was −0.418 with a 95% CI of (−0.141 to 0.186), which was nonsignificant. This likewise indicates full mediation, suggesting that social influence affects digital literacy only indirectly via usage intention.

Table . Mediating effect analysis (social influence and digital literacy).

SI→ADTS→D[a,b,c]	Effect	Bootstrapping (BC[d] 95% CI)
Indirect effect	0.128	0.069 to 0.194
Direct effect	–0.418	–0.141 to 0.186
Total effect	0.086	0.014 to 0.186

[a]SI: social influence.

[b]ADTS: Assistive Digital Tools and Services.

[c]D: digital literacy.

[d]BC: bias-corrected percentile method.

Overall, these findings demonstrate that both family support and social influence enhance older adults' digital literacy not by directly improving their skills, but by strengthening their intention to use ADTS—thereby facilitating subsequent skill acquisition and competence development.

Moderation Effect

The moderating effect of sense of achievement on the relationship between digital anxiety and ADTS usage intention was examined using PROCESS Model 1 with 5000 bootstrap samples. The overall model was significant (R^2=0.389, $F_{1, 476}$=101.222, P=.003), and the interaction term (Digital Anxiety×Sense of Achievement) was also significant, indicating the presence of a moderation effect.

Simple-slopes analysis revealed that the negative association between digital anxiety and ADTS usage intention was weaker at higher levels of sense of achievement and stronger at lower levels. This suggests that a greater sense of achievement can buffer the detrimental impact of digital anxiety on the intention to use ADTS. Figures 3 and 4 depict the slope differences across varying levels of sense of achievement, providing visual confirmation of the moderation effect.

In line with Self-Determination Theory, these results indicate that perceived competence and mastery help sustain motivation when engaging in tasks that may provoke anxiety. Accordingly, H9 is supported: sense of achievement attenuates the negative effect of digital anxiety on ADTS usage intention.

XSL•FO
RenderX

Figure 3. Moderation path coefficients. ADTS: Assistive Digital Tools and Services. *P<.05.

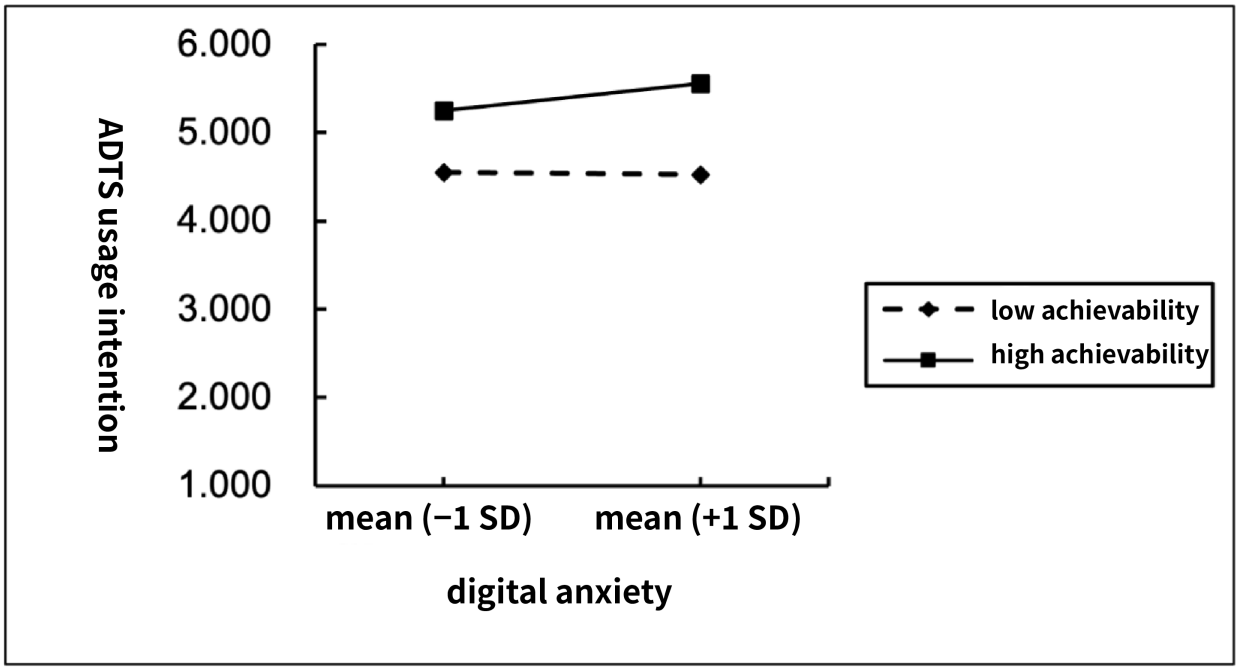
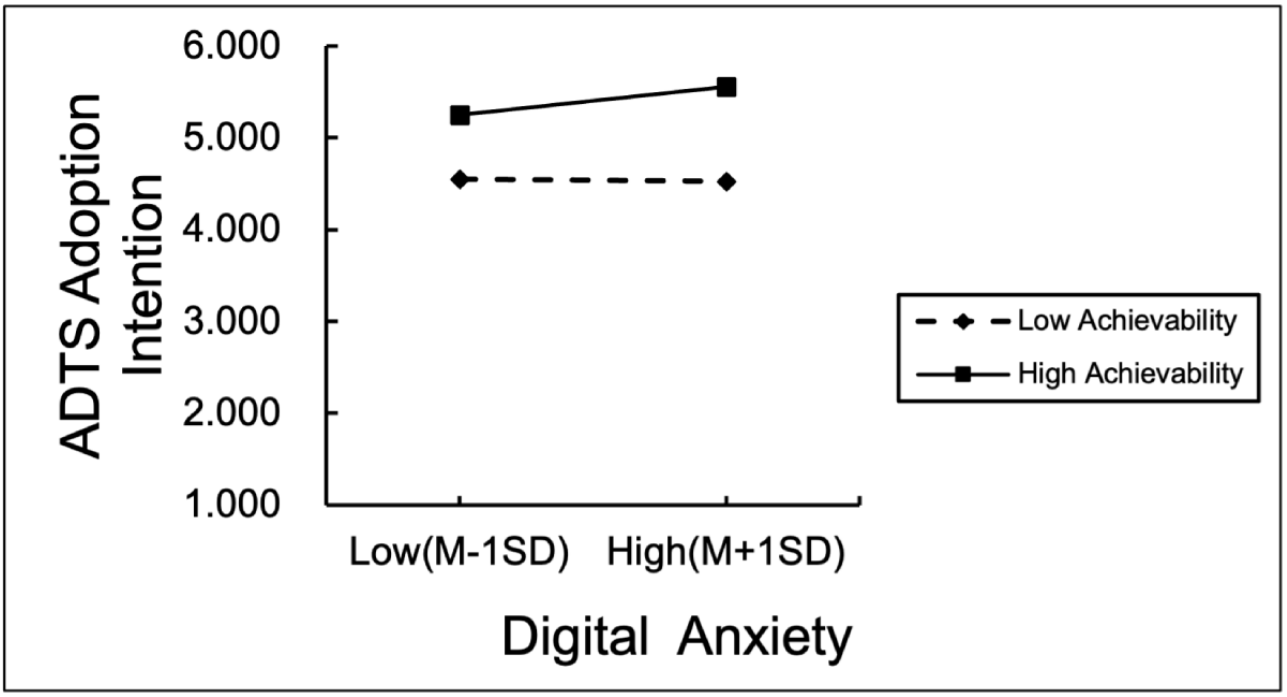


Figure 4. Moderation slope analysis. ADTS: Assistive Digital Tools and Services.



Discussion

Summary of Key Findings

In an era marked by rapid digital transformation and global population aging, older adults' ability to adapt to digital technologies has become a critical determinant of quality of life, autonomy, and the trajectory of successful aging [2,3]. Grounded in the UTAUT and the Digital Literacy Framework, this cross-sectional study tested an integrated model that incorporated environmental, psychological, and behavioral factors. Specifically, it examined how family support, social influence, digital anxiety, and sense of achievement are associated with older adults' intention to use ADTS, and how these factors, in turn, relate to the development of digital literacy. The findings provide new insights into how digital engagement can be transformed into real-life competence and social connection.

XSL•FO
RenderX

Family Support and Social Influence Operate Through Intention

Both factors were positively associated with ADTS usage intention (β=.413 and β=.160, respectively, *P*<.05). Mediation analysis showed that their effects on digital literacy were fully mediated by usage intention, with nonsignificant direct paths after accounting for the mediator.

Family Support Was Associated With Lower Digital Anxiety

Family support was negatively associated with digital anxiety (β=−0.168, *P*<.05), which in turn was negatively associated with usage intention (β=−.103, *P*<.05). This suggests that family support may help older adults overcome apprehension toward technology, indirectly promoting adoption.

Sense of Achievement Attenuates the Association of the Negative Effect of Digital Anxiety

The interaction term (Digital Anxiety × Sense of Achievement) was significant, indicating that a higher sense of achievement attenuates the negative relationship between digital anxiety and usage intention.

Usage Intention Is a Central Behavioral Pathway

Across environmental and psychological predictors, usage intention emerged as the pivotal link to higher digital literacy scores (β=.345, *P*<.001).

Theoretical Implications

This study advances the literature by showing how psychosocial factors operate within an integrative, later-life adaptation of UTAUT. Specifically, it contributes three distinctive insights—each grounded in the SEM results:

1. Extending UTAUT with age-specific psychosocial mechanisms: Beyond utilitarian appraisals, digital anxiety exerts a negative effect on intention (β=−0.103, *P*<.05), while sense of achievement attenuates this link (significant interaction), indicating that affective constraints and motivational buffers are integral to older adults' adoption processes.
2. Unpacking the pathway from environmental enablers to competence: The effects of family support (β=.413) and social influence (β=.160) on digital literacy are fully mediated by usage intention (indirect effects with 95% CIs excluding zero), positioning intention as a necessary conduit from context to competence.
3. Reframing successful aging for the digital era: Digital literacy functions as aging capital—a capability accumulated through intentional engagement and shaped by psychosocial conditions—thereby warranting its incorporation into the theoretical core of successful aging in technology-mediated societies.

Practical / Managerial Implications

The findings translate into the following action priorities for policymakers, community organizations, and technology designers:

1. Leverage family and peer networks as complementary pillars: Given the stronger association of family support with intention (β=.413) relative to social influence (β=.160), intergenerational pairings can serve as high-impact entry points. To avoid over-dependence, pair these with peer-led workshops, senior tech clubs, and neighborhood "digital ambassadors" for continuity.
2. Design for progressive achievement to counter anxiety: Because achievement buffers the anxiety-intention link, use milestone-based modules, micro-goals (eg, a first mobile payment), and instant positive feedback to convert apprehension into mastery and sustained engagement.
3. Prioritize sustained, habitual use over one-off adoption: With literacy gains operating through intention, build reinforcement sessions, periodic "digital check-ins," and adaptive nudges (reminders, personalized suggestions, calibrated difficulty) to routinize use and consolidate skills. These recommendations move beyond generic calls for support, advancing age-sensitive, culturally grounded strategies that directly target the psychosocial dynamics revealed by our model.

Future Research and Practice Directions

Limitations and Future Directions

The cross-sectional design precludes causal inference; results reflect associations within one metropolitan sample, limiting generalizability. Demographic covariates (eg, age, education, marital status) were collected but excluded from the structural model to preserve parsimony; future work can incorporate them to refine estimates.

Modeling Choices

We intentionally excluded performance expectancy and effort expectancy to foreground later-life psychosocial mechanisms and avoid redundancy with well-established findings; future studies should reincorporate these UTAUT constructs to test robustness and generalizability. We also modeled digital literacy as an outcome to trace competence formation; longitudinal or cross-lagged designs should examine possible bidirectional effects whereby literacy reduces anxiety, enhances self-efficacy, and strengthens intention.

Next Steps

We recommend longitudinal/experimental designs, cross-cultural comparisons, and the inclusion of digital trust/privacy and emerging AI/IoT tools to evaluate scalable, age-friendly pathways to digital inclusion.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Department of Economic and Social Affairs. World Population Prospects 2024: Summary of Results: New York: United Nations; 2024.
2. Radovanović D, Holst C, Belur SB, et al. Digital literacy key performance indicators for sustainable development. SI 2020;8(2):151-167. [doi: 10.17645/si.v8i2.2587]
3. Lee H. Analysis of the impact of digital literacy on life satisfaction (2019–2022) for older adults in South Korea: a national community-based panel study. Sci Rep 2024;14(1):20399. [doi: 10.1038/s41598-024-71397-0]
4. Hawkley LC, Cacioppo JT. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. Ann Behav Med 2010 Oct;40(2):218-227. [doi: 10.1007/s12160-010-9210-8] [Medline: 20652462]
5. Chen L, Jia J, Xiao M, Wu C, Zhang L. A study on the influence of digital literacy on elderly user's intention to identify social media false information. EL 2024 Sep 23;42(5):701-721. [doi: 10.1108/EL-10-2023-0257]
6. Wang D, Duan Y, Jin Y. Navigating online perils: socioeconomic status, online activity lifestyles, and online fraud targeting and victimization of old adults in China. Comput Human Behav 2025 Jan;162:108458. [doi: 10.1016/j.chb.2024.108458]
7. Majeed A, Zhang X. On the adoption of modern technologies to fight the COVID-19 pandemic: a technical synthesis of latest developments. COVID 2023;3(1):90-123. [doi: 10.3390/covid3010006]
8. Alassaf P, Szalay ZG. The impact of 'Compulsory' shifting to use e-services during COVID-19 pandemic restrictions period on e-services users' future attitude and intention "Case Study of Central European Countries/Visegrád Group (V4)". Sustainability 2022;14(16):9935. [doi: 10.3390/su14169935]
9. Lee K, Kim S, Kim SH, et al. Digital health interventions for adult patients with cancer evaluated in randomized controlled trials: scoping review. J Med Internet Res 2023 Jan 6;25:e38333. [doi: 10.2196/38333] [Medline: 36607712]
10. Miller LMS, Callegari RA, Abah T, Fann H. Digital literacy training for low-income older adults through undergraduate community-engaged learning: single-group pretest-posttest study. JMIR Aging 2024 May 14;7:e51675. [doi: 10.2196/51675] [Medline: 38599620]
11. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
12. Crowther MR, Parker MW, Achenbaum WA, Larimore WL, Koenig HG. Rowe and Kahn's model of successful aging revisited: positive spirituality--the forgotten factor. Gerontologist 2002 Oct;42(5):613-620. [doi: 10.1093/geront/42.5.613] [Medline: 12351796]
13. Xiong J, Zuo M. How does family support work when older adults obtain information from mobile internet? ITP 2019 Dec 2;32(6):1496-1516. [doi: 10.1108/ITP-02-2018-0060]
14. Caplan N. Social research and national policy: what gets used, by whom, for what purposes, and with what effects. Int Soc Sci J 1976;28(1):187-194 [FREE Full text]
15. Chopdar PK, Korfiatis N, Sivakumar VJ, Lytras MD. Mobile shopping apps adoption and perceived risks: a cross-country perspective utilizing the unified theory of acceptance and use of technology. Comput Human Behav 2018 Sep;86:109-128. [doi: 10.1016/j.chb.2018.04.017]
16. Selwyn N. Reconsidering political and popular understandings of the digital divide. New Media Soc 2004 Jun;6(3):341-362. [doi: 10.1177/1461444804042519]
17. Zhang M. Older people's attitudes towards emerging technologies: a systematic literature review. Public Underst Sci 2023 Nov;32(8):948-968. [doi: 10.1177/09636625231171677]
18. Meng Q, Yan Z, Abbas J, Shankar A, Subramanian M. Human–computer interaction and digital literacy promote educational learning in pre-school children: mediating role of psychological resilience for kids' mental well-being and school readiness. International Journal of Human–Computer Interaction 2025 Jan 2;41(1):16-30. [doi: 10.1080/10447318.2023.2248432]
19. Sosa Díaz MJ. Emergency remote education, family support and the digital divide in the context of the COVID-19 lockdown. Int J Environ Res Public Health 2021 Jul 28;18(15):7956. [doi: 10.3390/ijerph18157956] [Medline: 34360248]
20. Roman NV, Balogun TV, Butler-Kruger L, et al. Strengthening family bonds: a systematic review of factors and interventions that enhance family cohesion. Soc Sci (Basel) 2025;14(6):371. [doi: 10.3390/socsci14060371]
21. Karahanna E, Straub DW, Chervany NL. Information technology adoption across time: a cross-sectional comparison of pre-adoption and post-adoption beliefs. MIS Q 1999 Jun;23(2):183. [doi: 10.2307/249751]
22. Morosan C, DeFranco A. It's about time: revisiting UTAUT2 to examine consumers' intentions to use NFC mobile payments in hotels. Int J Hosp Manag 2016 Feb;53:17-29. [doi: 10.1016/j.ijhm.2015.11.003]
23. Peral-Peral B, Villarejo-Ramos Á, Arenas-Gaitán J. Self-efficacy and anxiety as determinants of older adults' use of internet banking services. Univ Access Inf Soc 2020 Nov;19(4):825-840. [doi: 10.1007/s10209-019-00691-w]
24. Arenas-Gaitán J, Villarejo Ramos AF, Peral-Peral B. A posteriori segmentation of elderly internet users: applying PLS-POS. MIP 2020 May 13;38(3):340-353. [doi: 10.1108/MIP-01-2019-0057]
25. Igbaria M, Iivari J. The effects of self-efficacy on computer usage. Omega (Westport) 1995 Dec;23(6):587-605. [doi: 10.1016/0305-0483(95)00035-6]
26. Straub ET. Understanding technology adoption: theory and future directions for informal learning. Rev Educ Res 2009 Jun;79(2):625-649. [doi: 10.3102/0034654308325896]
27. Cambre MA, Cook DL. Computer anxiety: definition, measurement, and correlates. Journal of Educational Computing Research 1985 Feb;1(1):37-54. [doi: 10.2190/FK5L-092H-T6YB-PYBA]

XSL•FO
RenderX

28. Wang CH, Wu CL. Bridging the digital divide: the smart TV as a platform for digital literacy among the elderly. Behav Inf Technol 2022 Sep 10;41(12):2546-2559. [doi: 10.1080/0144929X.2021.1934732]
29. Wang W, Zhang Y, Zhao J. Technological or social? influencing factors and mechanisms of the psychological digital divide in rural Chinese elderly. Technol Soc 2023 Aug;74:102307. [doi: 10.1016/j.techsoc.2023.102307]
30. Rook KS. Parallels in the study of social support and social strain. J Soc Clin Psychol 1990 Mar;9(1):118-132. [doi: 10.1521/jscp.1990.9.1.118]
31. Carlo S, Nanetti S. Understanding ICTS in older life: a scope review of 'The Gerontologist' and 'Research on Aging'. 2023 Jul Presented at: Human Aspects of IT for the Aged Population: 9th International Conference, ITAP 2023, Held as Part of the 25th HCI International Conference, HCII 2023; Jul 23-28, 2023; Copenhagen, Denmark p. 426-442. [doi: 10.1007/978-3-031-34917-1_29]
32. Valta M, Hildebrandt Y, Maier C. Fostering the digital mindset to mitigate technostress: an empirical study of empowering individuals for using digital technologies. INTR 2024 Nov 25;34(6):2341-2369. [doi: 10.1108/INTR-09-2022-0766]
33. Mariano J, Marques S, Ramos MR, et al. Too old for technology? Stereotype threat and technology use by older adults. Behav Inf Technol 2022 May 19;41(7):1503-1514. [doi: 10.1080/0144929X.2021.1882577]
34. Helsper EJ. The social relativity of digital exclusion: applying relative deprivation theory to digital inequalities. Commun Theor 2017 Aug;27(3):223-242. [doi: 10.1111/comt.12110]
35. Pejić Bach M, Ivančić L, Bosilj Vukšić V, Stjepić AM, Milanović Glavan L. Internet usage among senior citizens: self-efficacy and social influence are more important than social support. JTAER 2023;18(3):1463-1483. [doi: 10.3390/jtaer18030074]
36. Jang HW, Moon C, Jung HS, Cho M, Bonn MA. Normative and informational social influence affecting digital technology acceptance of senior restaurant diners: a technology learning perspective. Int J Hosp Manag 2024 Jan;116:103626. [doi: 10.1016/j.ijhm.2023.103626]
37. Balki E, Hayes N, Holland C. Loneliness and older adults: psychological resilience and technology use during the COVID-19 pandemic-a cross sectional study. Front Aging 2023;4:1184386. [doi: 10.3389/fragi.2023.1184386] [Medline: 37434741]
38. Eshet Y. Digital literacy: a new terminology framework and its application to the design of meaningful technology-based learning environments. Assoc Adv Comput Educ 2002:493-498 [FREE Full text]
39. Park H, Kim HS, Park HW. A scientometric study of digital literacy, ICT literacy, information literacy, and media literacy. Journal of Data and Information Science 2021 Apr 4;6(2):116-138. [doi: 10.2478/jdis-2021-0001]
40. Fu J. ICT in education: a critical literature review and its implications. Int J Educ Dev Using Inf Commun Technol 2013;9(1):112-125 [FREE Full text]
41. Rogers SE. Bridging the 21st century digital divide. TechTrends 2016 May;60(3):197-199. [doi: 10.1007/s11528-016-0057-0]
42. Tang X, Sun Y, Zhang B, et al. I never imagined grandma could do so well with technology. Proc ACM Hum-Comput Interact 2022 Nov 7;6(CSCW2):1-29. [doi: 10.1145/3555579]
43. Zhu R, Yu X, Krever R. The double burden: the digital exclusion and identity crisis of elderly patients in rural China. MaC ;12. [doi: 10.17645/mac.8106]
44. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: a systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124. [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
45. Magen-Nagar N, Cohen L. Learning strategies as a mediator for motivation and a sense of achievement among students who study in MOOCs. Educ Inf Technol 2017 May;22(3):1271-1290. [doi: 10.1007/s10639-016-9492-y]
46. Martin A. DigEuLit–a European framework for digital literacy: a progress report. J eLiteracy 2005;2(2):130-136 [FREE Full text]
47. Demirdis B. Integrating digital literacy to enhance emotional and social skills in education. In: In Innovative Educational Frameworks for Future Skills and Competencies: IGI Global Scientific Publishing; 2025:1-38. [doi: 10.4018/979-8-3693-7555-6.ch001]
48. Bianchi C. Exploring how internet services can enhance elderly well-being. JSM 2021 Sep 1;35(5):585-603. [doi: 10.1108/JSM-05-2020-0177]
49. Safarov N. Personal experiences of digital public services access and use: older migrants' digital choices. Technol Soc 2021 Aug;66:101627. [doi: 10.1016/j.techsoc.2021.101627]
50. Janke S, Bardach L, Oczlon S, Lüftenegger M. Enhancing feasibility when measuring teachers' motivation: a brief scale for teachers' achievement goal orientations. Teaching and Teacher Education 2019 Jul;83:1-11. [doi: 10.1016/j.tate.2019.04.003]
51. National Bureau of Statistics of China. China Statistical Yearbook 2023: Aging and Demographic Trends: China Statistics Press; 2023.
52. Jöreskog KG, Sörbom D. LISREL 8.80. : Scientific Software International Inc; 2006 URL: https://www.ssicentral.com/lisrel [accessed 2025-12-11]
53. Doll WJ, Xia W, Torkzadeh G. A confirmatory factor analysis of the end-user computing satisfaction instrument. MIS Q 1994 Dec;18(4):453. [doi: 10.2307/249524]

XSL•FO
RenderX

54. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1996;1(2):130-149. [doi: 10.1037//1082-989X.1.2.130]
55. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull 1980;88(3):588-606. [doi: 10.1037//0033-2909.88.3.588]
56. Deci EL, Ryan RM. The general causality orientations scale: self-determination in personality. J Res Pers 1985 Jun;19(2):109-134. [doi: 10.1016/0092-6566(85)90023-6]
57. Badr J, Motulsky A, Denis JL. Digital health technologies and inequalities: a scoping review of potential impacts and policy recommendations. Health Policy 2024 Aug;146:105122. [doi: 10.1016/j.healthpol.2024.105122] [Medline: 38986333]
58. Farmania A, Elsyah RD, Fortunisa A. The phenomenon of technostress during the COVID-19 pandemic due to work from home in Indonesia. Sustainability ;14(14):8669. [doi: 10.3390/su14148669]
59. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006;8(4):e27. [doi: 10.2196/jmir.8.4.e27]
60. Bentler PM, Chou CP. Practical issues in structural modeling. Sociol Methods Res 1987 Aug;16(1):78-117. [doi: 10.1177/0049124187016001004]
61. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis, 7th edition: Pearson; 2009.
62. MacCallum RC, Hong S. Power analysis in covariance structure modeling using GFI and AGFI. Multivariate Behav Res 1997 Apr;32(2):193-210. [doi: 10.1207/s15327906mbr3202_5]
63. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
64. Mulaik SA, James LR, Van Alstine J, Bennett N, Lind S, Stilwell CD. Evaluation of goodness-of-fit indices for structural equation models. Psychol Bull 1989;105(3):430-445. [doi: 10.1037/0033-2909.105.3.430]

Abbreviations

ADTS: Assistive Digital Tools and Services
CFA: confirmatory factor analysis
SCT: social cognitive theory
SEM: structural equation modeling
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by R Bjarnadottir; submitted 31.Mar.2025; peer-reviewed by GHY Wong, S Shapira; revised version received 03.Sep.2025; accepted 20.Oct.2025; published 21.Jan.2026.

Please cite as:
Niu HJ, Li MH, Hsieh FY, Yu CC, Lin CT
From Digital Anxiety to Empowerment in Older Adults: Cross-Sectional Survey Study on Psychosocial Drivers of Digital Literacy
JMIR Aging 2026;9:e75245
URL: https://aging.jmir.org/2026/1/e75245
doi:10.2196/75245

© Han-Jen Niu, Ming-Hsuan Li, Feng-Yu Hsieh, Chun-Chieh Yu, Chun-Ting Lin. Originally published in JMIR Aging (https://aging.jmir.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smartphone Addiction, Use Preferences, and Depression Among Older Adults in the Digital Context: Machine Learning Analysis of Survey Data

Sheng Chen[1], PhD; Yue Song[1], BA; Chien-chung Huang[2], PhD

[1]School of Public Administration, Guangdong University of Foreign Studies, 2 Baiyun Avenue, Baiyun District, Guangzhou, China

[2]School of Social Work, Rutgers University, New Brunswick, NJ, United States

Corresponding Author:
Yue Song, BA
School of Public Administration, Guangdong University of Foreign Studies, 2 Baiyun Avenue, Baiyun District, Guangzhou, China

Abstract

Background: The increasing use of smartphones among older adults offers new opportunities for social connection but may also pose risks associated with adverse mental health outcomes, including depression.

Objective: This study examined the relationship between smartphone use and depression among older adults in Guangzhou, China, to identify key predictors and complex configurations associated with depression.

Methods: Using a hybrid analytic approach applied to survey data from 2585 older adults in Guangzhou, machine learning methods first identified the strongest predictors of depression. Subsequent fuzzy-set qualitative comparative analysis delineated distinct configurations associated with depression.

Results: The analysis identified high levels of smartphone addiction and a low preference for interactive use as central configurations associated with depression. These patterns appeared to operate through 2 distinct mechanisms: structural exclusion related to resource limitations and agentic maladaptation associated with relational deficiencies. This underlying vulnerability, in combination with specific sociodemographic factors, delineated distinct depression-related typologies. High-risk profiles included unmarried men with limited educational attainment and socially withdrawn individuals with greater socioeconomic resources.

Conclusions: Problematic smartphone use, particularly when it displaces social interaction, is significantly associated with depression among older adults. Interventions should therefore promote healthier interactive digital behaviors and provide targeted support for high-risk groups.

(JMIR Aging 2026;9:e84703) doi:10.2196/84703

KEYWORDS

depression; fuzzy-set qualitative comparative analysis; machine learning; smartphone addiction; smartphone use preference

Introduction

The rapid development of digital technologies, coupled with an aging population, has resulted in increasing smartphone use among older adults. The proportion of individuals aged 60 years and older in China increased from 14.9% in 2013 to 21.1% in 2023, reflecting the country's transition toward a moderately aging society [1]. According to the 56th Statistical Report on China's Internet Development, released in July 2025, the internet penetration rate among older adults in China reached 52%, with 161 million internet users aged 60 years and older [2]. These trends reflect the growing digital engagement of older adults in contemporary society.

The expanding prevalence of smartphone use has generated increasing scholarly interest in its implications for the mental health of older adults. Studies suggest that smartphones fulfill daily needs for communication, entertainment, and information [3]. However, problematic smartphone use has been associated with reduced social connectedness, intergenerational tension, and sleep disturbances [4], all of which may contribute to psychological distress. Thus, smartphone use presents both potential benefits and risks for older adults: moderate and purposeful engagement may support well-being, whereas excessive or maladaptive use may exacerbate mental health risks.

Despite a growing body of research, many studies on late-life depression tend to examine physiological, psychological, or social predictors in isolation, often underemphasizing the complex interaction among these dimensions and their combined effects. The limitation of single-factor studies constrains

XSL•FO
RenderX

understanding of how digital behaviors, social contexts, and individual vulnerabilities interact to produce divergent outcomes. The influence of smartphone use intensity and preference does not occur in isolation but is embedded within broader life circumstances. To address this theoretical and methodological gap, this study developed a configurational model to identify the most salient predictors and examine how these factors combined to form patterns associated with depressive outcomes. The findings provide empirically grounded insights to inform strategies aimed at promoting mental well-being and digital literacy among older adults.

Depression in Later Life

Late-life depression is a multifactorial condition influenced by physiological, psychological, and social factors. Physiological contributors include chronic illness and functional decline [5,6]. Psychological factors such as negative thinking and loneliness are also associated with increased risk [7,8]. Social support systems and community engagement play important roles. Social participation theory suggests that engagement in activities such as volunteering and recreational activities protects mental health by reducing loneliness and building supportive networks [9,10].

Smartphone Use and Depression Among Older Adults

In this study, smartphone addiction is defined as a behavioral addiction characterized by compulsive and excessive use that disrupts daily functioning and is associated with adverse psychological outcomes [11,12]. Smartphone addiction has been positively associated with depression, particularly with prolonged use [13]. Excessive reliance may create a cycle in which loneliness drives escapist use, further deepening social isolation [14].

Patterns of smartphone use are associated with mental health outcomes. Purposeful and socially oriented use—categorized as interactive functions (eg, social networking and online task engagement)—may enhance social connectedness [15]. In contrast, entertainment-driven use—categorized as self-entertainment functions (eg, playing games, scrolling social media, watching videos)—is linked to higher levels of depression and anxiety, often mediated by escapism [16,17]. Although much of the evidence derives from studies of younger populations [18-20], older adults may be particularly vulnerable due to cognitive decline, which may increase the likelihood of passive use reinforcing isolation and depressive symptoms.

Conceptual Framework

The conceptual framework of this study integrates socioemotional selectivity theory (SST) and the digital inequality theory (DIT). Socioemotional selectivity theory posits that older adults prioritize emotionally meaningful relationships, thus smartphone use may be beneficial only when it supports these goals rather than displaces them [21]. Digital inequality theory emphasizes that structural disadvantages limit digital literacy and access, potentially transforming technology into a source of exclusion rather than connection [22]. Together, these perspectives inform the hypotheses in this study regarding the conditions under which depression may arise among older adults. Accordingly, this study tested the hypothesis that depression among digitally active older adults is associated with maladaptive patterns of smartphone use. It further examined whether socioeconomic barriers limit beneficial engagement, contributing to exclusion and depressive symptoms, consistent with DIT, and whether individuals with greater socioeconomic resources substitute smartphone use for meaningful social participation, potentially resulting in emotional dissatisfaction despite material security, consistent with SST.

This study tested 3 hypotheses. First, high smartphone addiction was hypothesized to represent a necessary condition for high levels of depression. Second, high smartphone addiction in combination with low educational attainment and low preference for interactive use was hypothesized to constitute a sufficient configuration for depression. Third, high smartphone addiction combined with low social participation was hypothesized to form a sufficient configuration for depression, particularly among individuals with greater socioeconomic resources.

Methods

Data

Data were collected using a combined convenience and quota sampling strategy through a study-specific survey titled Smartphone Use and Mental Health Among Older Adults, administered between October and November 2024. To ensure data quality, structured, on-site assisted data collection was implemented across 87 social workstations in 5 districts of Guangzhou. Each workstation recruited 30 community-dwelling older adults. Social workers organized in-person sessions at community centers, guided participants through the questionnaires, clarified items, assisted with smartphone use, and reviewed responses for completeness. This assisted approach helped address technical and cognitive barriers, resulting in a high valid response rate and a consistent sample size.

Measures

Depression was measured using the 15-item Geriatric Depression Scale, adapted by Tang [23] for Chinese older adults. Items were answered in binary form (yes=1, no=0), yielding total scores ranging from 0 to 15. Following Tang's guidelines, scores were categorized as 0 to 4 (no depression), 5 to 9 (mild depression), and 10 points or more (clinical depression). Internal consistency was good (Cronbach α=0.790). Smartphone addiction was assessed using the 17-item Mobile Phone Addiction Index [24]. Items were rated on a 5-point Likert scale, with higher scores reflecting greater severity. In this study, the scale scores were summed and standardized. Reliability in this sample was excellent (Cronbach α=0.968).

Smartphone use preferences were assessed with 2 multiple-choice items: 1 captured the general smartphone functions used and another identified the 3 smartphone functions most frequently used. Exploratory factor analysis of 12 functions yielded three factors: (1) self-entertainment function use (eg, streaming and music; explained variance=3.627), (2) interactive function use (eg, WeChat, short videos; explained variance=1.264), and (3) basic telecommunication (eg, calls, texts; explained variance=1.122). Short video platforms, such as TikTok and Xiaohongshu, play an increasingly prominent role in fostering social connectivity and interaction. These

XSL•FO
RenderX

platforms enable users to share videos with family members and engage in interactive livestreams within local community contexts [25,26]. Self-entertainment use and interactive use were the 2 functions retained for analysis. The use of self-entertainment or interactive functions was operationalized as the occurrence of engagement with self-entertainment or interactive features, with higher counts indicating stronger preferences.

Social participation was measured using the 22-item scale developed by Ren et al [27]. Responses were recorded on a 5-point Likert scale, with higher scores indicating greater participation. In this study, scores were summed and standardized. Internal consistency was high (Cronbach α=0.946). Covariates included gender, age, education, employment, self-rated health, income, marital status, and average daily smartphone use. Multicollinearity diagnostics indicated no concerns (all variance inflation factors <2).

Methodological Approach

We used a multimethod strategy that combined machine learning and fuzzy-set qualitative comparative analysis (fsQCA). The integration of machine learning and fsQCA followed a predictive-explanatory sequence. Machine learning was used to handle high-dimensional feature selection and objectively identify factors most relevant for prediction, whereas fsQCA was used to uncover how these factors combine to produce depressive outcomes. This sequential design leveraged the strengths of both methods: machine learning identified the most influential predictors of depression, and fsQCA revealed complex configurations underlying depressive outcomes. Together, these provided a more nuanced understanding of complex relationships.

First, 4 machine learning algorithms—logistic regression, support vector machine, random forest (RF), and extreme gradient boosting—were applied using the web-based Python Jupyter Notebook development environment to identify salient predictors of depression. To prevent overfitting, 10-fold cross-validation was used, and performance metrics were reported as the mean across folds to provide stable and generalizable estimates. To enhance interpretability, the Shapley additive explanations algorithm was used, quantifying each variable's contribution to model predictions.

Following predictive modeling, fsQCA was conducted to test hypotheses on the necessary and sufficient conditions for depression. Fuzzy-set qualitative comparative analysis is a set-theoretic method that evaluates causal complexity by identifying configurations of factors, rather than the net effects of individual variables [28]. Based on fuzzy-set theory, it assesses necessary conditions, sufficient pathways, and conjunctural causation, allowing for equifinality (ie, multiple combinations of conditions leading to the same outcome). This approach is well-suited for exploring how structural vulnerabilities, smartphone use patterns, and social participation jointly shape mental health in later life.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and approved by the Research Review Committee of the School of Public Administration, Guangdong University of Foreign Studies (2024003). Informed consent was obtained prior to survey administration. Participants were informed that participation was voluntary and anonymous, that they could withdraw at any time, and that measures were in place to ensure data confidentiality, such as deidentification, secure server storage, and restricted access to data.

Results

Of 2610 questionnaires distributed, 2585 were completed and deemed valid, yielding a 99% response rate. Demographic characteristics are presented in Table 1.

XSL•FO
RenderX

Table . Statistics of demographic characteristics (N=2585).

Characteristics	Participants, n (%)
Depression	
1 - 4 points, no depression	1429 (55.3)
5 - 9 points, mild depression	983 (38.0)
10 points and above, clinical depression	173 (6.7)
Use interactive functions	
No	387 (15.0)
Yes	2198 (85.0)
Use self-entertainment functions	
No	2056 (79.5)
Yes	529 (20.5)
Gender	
Female	1719 (66.5)
Male	866 (33.5)
Age (years)	
60-69	1490 (57.6)
70-79	729 (28.2)
≥80	366 (14.2)
Education level	
Primary school or below	1056 (40.9)
Junior or senior high school	1205 (46.6)
College and above	324 (12.5)
Employed	
No	2337 (90.4)
Yes	248 (9.6)
Health status	
Unhealthy	1935 (74.9)
Healthy	650 (25.1)
Monthly income[a]	
Less than 2000 RMB	797 (30.8)
2000 - 5000 RMB	1224 (47.4)
More than 5000 RMB	564 (21.8)
Marital status	
Single	971 (37.6)
Married	1614 (62.4)
Smartphone use duration	
1 hour or less	520 (20.1)
2 - 6 hours	1711 (66.2)
7 hours or more	354 (13.7)

[a]A currency exchange rate of 6.9 RMB=US $1 is applicable.

XSL•FO
RenderX

Predictive Analysis Using Machine Learning Models

A comparative analysis was conducted using logistic regression, support vector machine, RF, and extreme gradient boosting models to predict depression based on smartphone use variables, social participation, and sociodemographic covariates. The dataset was partitioned into a training set comprising 80% of the data and a test set comprising the remaining 20%. During feature engineering, the outcome variable was binarized and the predictor variables were standardized. This process transformed all features to a common scale; this ensured their influence was balanced and prevented certain features from having an excessive impact on the model due to differences in magnitude. To reduce the risk of model overfitting, bootstrapping was applied within the machine learning framework to generate training sets for each tree, resulting in 1740 additional resampled observations, which were used exclusively for model training; all subsequent analyses were conducted using the original dataset.

To address observed imbalances in the outcome variable (depression categories) within the training data, the synthetic minority over-sampling technique was applied to the training set only and not to the test set used for final area under the curve computation. After applying the synthetic minority over-sampling technique, the distribution of depression levels was fully balanced, providing a more equitable foundation for model training. Grid search with 10-fold cross-validation was used for hyperparameter tuning and model optimization. Among the 4 machine learning analyses conducted, the RF model demonstrated superior performance on the test set as shown in Table 2; therefore, RF was selected for feature importance analysis.

Table . Performance of machine learning models on the test set.

Model	Accuracy	Precision	Recall	F_1-score	AUC[a]
RF[b]	0.8483	0.8258	0.8258	0.8258	0.9261
XGBoost[c]	0.8186	0.7895	0.7955	0.7925	0.8742
SVM[d]	0.7585	0.7140	0.7424	0.7280	0.8103
LR[e]	0.6735	0.6130	0.6780	0.6439	0.7293

[a]AUC: area under the curve.

[b]RF: random forest.

[c]XGBoost: extreme gradient boosting.

[d]SVM: support vector machine.

[e]LR: logistic regression.

Ranking the Importance of Features Across Different Variables

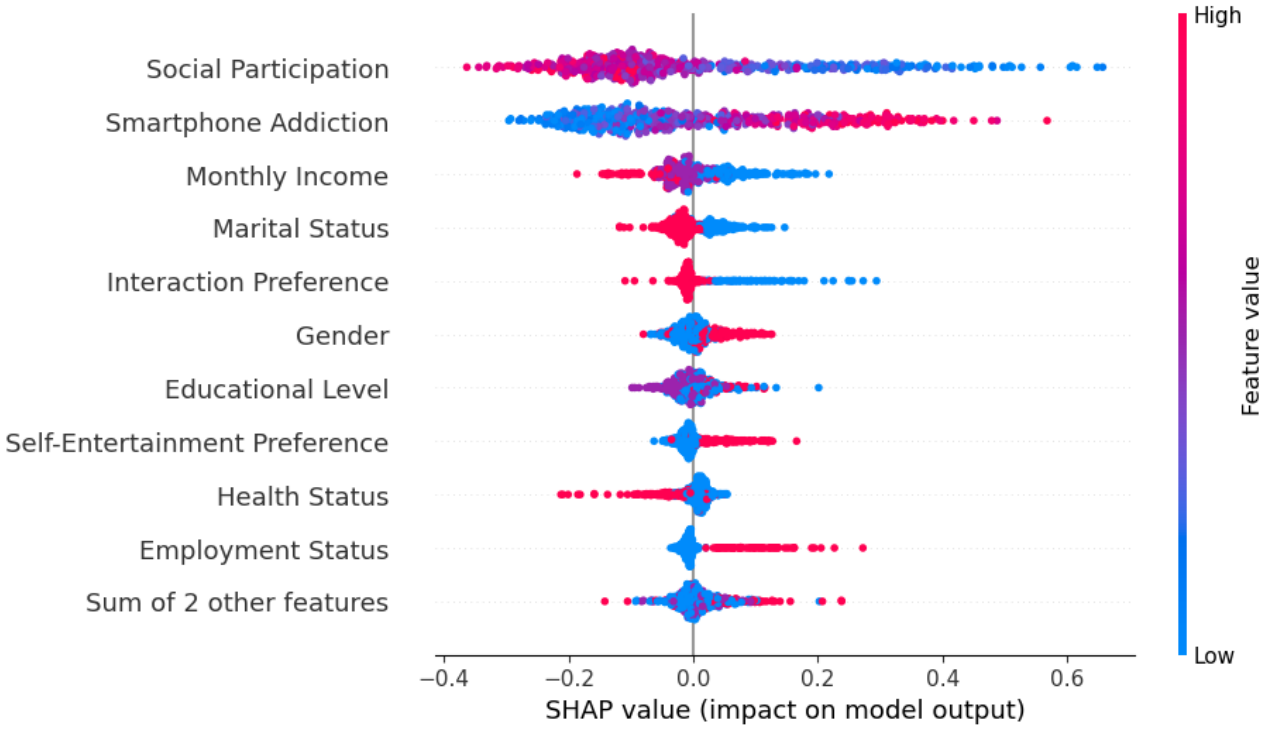
As shown in Table 3, the most influential predictors of depression, in descending order of importance, were social participation, smartphone addiction, monthly income, marital status, interaction preference (ie, the use of interactive functions), education level, gender, health status, self-entertainment preference (ie, the use of self-entertainment functions), and age, with social participation and smartphone addiction being the most salient factors. Figure 1 presents the Shapley additive explanations algorithm values, indicating that social participation had the strongest negative association with depression, whereas higher levels of participation were associated with a significantly lower probability of depression. Smartphone addiction was positively associated with depression. Higher income and being married were associated with a lower risk of depression, and use of interactive functions preference was also associated with reduced risk. Educational attainment showed a relatively modest association with a risk of depression. Female sex, poorer health status, a stronger preference for self-entertainment functions, and older age were each associated with a slight increase in depression risk. Because fsQCA typically examines 4 to 8 antecedent conditions, the top 7 predictors—social participation, smartphone addiction, monthly income, marital status, use of interactive functions, education, and gender—were selected as key conditions for configurational analysis of late-life depression.

XSL•FO
RenderX

Table . Top 10 features ranked by Shapley additive explanations (SHAP) algorithm values.

Importance rank	Feature	SHAP value
1	Social participation	0.1613
2	Smartphone addiction	0.1441
3	Monthly income	0.0421
4	Marital status	0.0272
5	Preference for interactive functions	0.0230
6	Educational level	0.0211
7	Gender	0.0201
8	Health status	0.0199
9	Preference for self-entertainment functions	0.0194
10	Age	0.0192

Figure 1. Predictors of depression. SHAP: Shapley additive explanations.



FsQCA Analysis

This study used fsQCA 4.0 software to convert antecedent and outcome variables into fuzzy-set membership scores using the direct calibration method. Following the literature in FsQCA analysis, the 95th, 50th, and 5th percentiles were set as the 3 anchor points for full membership, the crossover point, and full nonmembership, respectively [29,30]. To address the issue that cases with a fuzzy membership score of exactly 0.5 cannot be included in the analysis, the method proposed by Fiss [31] was followed, replacing the 0.5 value with 0.501. Table 4 presents the results of the variable calibration.

XSL•FO
RenderX

Table . Calibration analysis.

Variables	Anchor points[a]		
	Full nonmembership	Crossover point	Full membership
Condition variables			
Social participation	–1.722	0.662	1.790
Smartphone addiction	–1.224	0.018	1.624
Monthly income	0	1	2
Marital status	0	1	1
Preference for interactive functions	0	1	1
Education level	0	1	2
Gender	0	0	1
Outcome variable			
Depression	0	0	2

[a]Anchor points were calibrated using the 5th percentile (full nonmembership), 50th percentile (crossover), and 95th percentile (full membership).

To test the first hypothesis, which posited that high smartphone addiction is a necessary condition for high depression, a necessity analysis was conducted. As shown in Table 5, the consistency score for smartphone addiction was 0.991, exceeding the commonly used threshold of 0.9. This suggests that nearly all cases of high depression also exhibited high smartphone addiction. Therefore, the first hypothesis is strongly supported. The consistency scores for the remaining variables were all below 0.9, indicating that none met the threshold for depression in individual cases. The combined effects of multiple variables warrant further examination using a configurational approach.

Table . Fuzzy-set qualitative comparative analysis necessity analysis.

Condition variables	Consistency	Coverage
Social participation		
Yes	0.490	0.887
No	0.824	0.841
Smartphone addiction		
Yes	0.991	0.974
No	0.508	0.985
Monthly income		
Yes	0.675	0.757
No	0.616	0.744
Marital status		
Married	0.030	0.789
Unmarried	0.634	0.665
Use of interactive functions		
Yes	0.013	0.918
No	0.994	0.658
Education		
Any formal education	0.560	0.780
No formal education	0.680	0.707
Gender		
Male	0.745	0.746
Female	0.534	0.639

XSL•FO
RenderX

The analysis identified 3 distinct configurations (Table 6), which collectively achieved a high overall fsQCA configurational (sufficiency) solution of 0.956. The first configuration, termed the digitally disconnected and vulnerable profile, is characterized by high smartphone addiction, low interaction preference, lower educational attainment, and male gender. The raw coverage was 0.499, indicating that approximately 49.9% of cases of depression were represented by this configuration. This finding supports the hypothesized risk pattern in contexts where structural constraints may amplify behavioral vulnerabilities. The second and third configurations highlight more complex, paradox-driven mechanisms. The socially withdrawn profile with greater socioeconomic resources was characterized by high smartphone addiction, low social participation, and no use of interaction functions, despite high educational attainment and high income. This configuration demonstrated a raw coverage of 0.560. This supports the hypothesis that greater socioeconomic resources may not necessarily protect against depression when digital use is maladaptive, suggesting a potential paradox among socioeconomically advantaged older adults. The third configuration, the emotionally isolated and digitally dependent profile, was specifically characterized by unmarried men with high smartphone addiction and low interaction preference, even when overall social participation appeared adequate. The raw coverage was 0.462. Collectively, these 3 configurations illustrated how smartphone addiction, in combination with specific socioeconomic and marital contexts, was associated with late-life depression.

The analysis revealed an additional sufficient configuration that was not prespecified in the hypotheses. This emergent configuration, termed the emotionally isolated and digitally dependent profile, demonstrated a consistency of 0.910 and unique coverage of 0.011. It was characterized by unmarried older men who, despite relatively high social participation, exhibited high smartphone addiction and no use of interaction functions. The identification of this configuration is an important finding. Whereas the initial second and third hypotheses focused on vulnerabilities associated with limited socioeconomic resources or low social participation, this configuration suggested a distinct pattern associated with relational and gender-related vulnerabilities. The theoretical implications of this unanticipated finding are discussed further in the Discussion section. The robustness of the solution was assessed by (1) raising the consistency threshold from 0.80 to 0.85, (2) tightening the proportional reduction in inconsistency cutoff, and (3) modifying the case frequency threshold. The core configurations for depression remained substantively unchanged, confirming the stability of the solution.

Table . Configurations associated with depression.[a,b]

Variables and metrics	Depression		
	Configuration 1	Configuration 2	Configuration 3
Condition variables			
Social participation		⊗ [c]	⊗ [d]
Marital status	⊗		⊗ [e]
Interaction preference		⊗	⊗
Education	⊗	⊗	⊗
Gender	⊗		⊗
Monthly income	⊗ [f]	⊗	
Smartphone addiction	⊗	⊗	⊗
Solution metrics			
Consistency	0.928	0.978	0.910
Raw coverage	0.499	0.560	0.462
Unique coverage	0.160	0.043	0.011

[a]Overall solution consistency was 0.956.

[b]Overall solution coverage was 0.243.

[c]Indicates the absence of core conditions.

[d]Indicates the presence of core conditions.

[e]Indicates the absence of peripheral conditions.

[f]Indicates the presence of peripheral conditions.

XSL•FO
RenderX

Discussion

Configurational Patterns of Late-Life Depression in the Digital Context

This study confirmed smartphone addiction as a foundational vulnerability and necessary condition for late-life depression. However, this individual-level risk factor appeared to operate through specific configurational pathways rather than in isolation. Thus, this study synthesized the findings into a dual-mechanism integrated framework of depression caused by digital technology to clarify heterogeneity in etiological patterns across different groups.

The first mechanism, structural exclusion and deficit, reflects a resource-deficit pattern of depression primarily observed among older men with low educational attainment. This pattern may be shaped by digital inequality and gender-related social norms [32]. Limited educational attainment may constrain digital literacy, contributing to a greater reliance on smartphone use to engage in passive entertainment rather than socially connective functions. In contrast to women, who often maintain broader kin-keeping networks, older men may rely more heavily on spouse-centered networks. Following widowhood or social isolation, no use of interaction functions and lower digital literacy may hinder the formation of online connections, while fewer kin-keeping practices may reduce social buffers against loneliness. These intersecting factors are associated with an elevated risk of depression in this subgroup [33,34]. In this context, smartphones may not function as bridges to social integration and instead may reinforce existing structural marginalization.

The second mechanism, agentic maladaptation and relational void, reflects a maladaptation-based pattern of depression observed among individuals with greater socioeconomic resources and unmarried men. For socioeconomically advantaged individuals, material resources such as income and education may not provide effective psychological protection in the absence of socially meaningful connections. Consistent with SST, passive digital consumption without interpersonal interaction cannot substitute for emotionally meaningful engagement, potentially leading to material abundance coexisting with emotional deprivation [35]. Among unmarried men, the absence of a primary attachment figure may create a unique emotional void. Even when overall social participation seems adequate, reliance on noninteractive smartphone use may represent an attempt to use technology for unmet emotional needs. In this configuration, depression was associated not with limited capability or access but with patterns of digital engagement that may reflect efforts to cope with offline relational deficits through compulsive digital escapism [36].

Theoretical and Practical Implications

This study contributes to DIT in two ways. First, it refines DIT by identifying gendered patterns among men with lower educational attainment. Second, it extends SST by specifying contextual conditions under which unmet social connection needs are associated with depression, even among groups with greater socioeconomic resources, highlighting the importance of meaningful engagement for well-being.

From a practical perspective, these findings suggest that interventions should be targeted and configuration-specific rather than uniform. For digitally disconnected older men at elevated risk, intervention strategies may extend beyond basic technical support to strengthen digital family connections. Community-based workshops (eg, grandparent-grandchild connection programs) can teach functional skills such as video calling and photo sharing, transforming passive screen time into active intergenerational engagement. Peer-led interest groups may also promote social engagement.

For socioeconomically advantaged but socially isolated older adults, greater resources may not prevent depression associated with a diminished sense of purpose. Meaning-oriented engagement, such as recruiting older adults as online mentors, may shift their role from passive consumer to active contributor. Among unmarried older men, the combination of unmarried status and high screen time may warrant proactive screening. Programs such as digital workshops for older men could create low-pressure online spaces that foster weak-tie networks, reduce loneliness, and decrease emotional dependency on smartphones. Finally, given that smartphone addiction was identified as a necessary condition across depressive profiles, technology developers should consider incorporating age-sensitive design principles and responsible algorithm practices to support digital well-being.

These findings should be considered alongside several limitations. First, the cross-sectional design and nonprobability sampling from Guangzhou limit causal inference and generalizability. Longitudinal studies are needed to clarify causality and reduce potential measurement overlap between smartphone addiction and depression. Second, the measure of social participation did not distinguish between online and offline engagement, which may obscure their distinct associations with mental health. Future research should use instruments that differentiate these modalities. Third, smartphone addiction in older adults may require more nuanced assessment. Behaviors resembling addiction may reflect limited digital literacy rather than true dependency. Mixed-methods studies could help distinguish between compulsive use and unskilled engagement. Fourth, the high internal consistency (Cronbach α=0.968) of the Mobile Phone Addiction Index in this sample suggests it may be perceived as a unidimensional construct by older adults; further validation of its factor structure in this population is warranted. Finally, this study did not account for medication use, a potential confounder. Future research should incorporate medication-related variables to better isolate the association between digital engagement and depressive symptoms.

Conclusion

This study suggests that late-life depression and the increasing use of smartphones is not a uniform condition but reflects complex interactions. Smartphone addiction was consistently associated with an increased risk of depression and appeared to operate through 2 distinct mechanisms: structural exclusion and maladaptive engagement. The former was characterized by

XSL•FO
RenderX

resource deficits among vulnerable men, whereas the latter was associated with relational isolation among individuals with greater socioeconomic resources or limited social integration. These results support the use of a configurational approach rather than single-factor explanations to promote healthy aging in an increasingly digital society.

Acknowledgments

During the preparation of this manuscript, the authors used Google Gemini for the sole purpose of language editing and grammatical refinement. After using this tool, the authors reviewed and edited the content as needed, and take full responsibility for the accuracy and integrity of the final publication.

Funding

This work was supported by the 2023 Disciplinary Co-construction Project of Guangdong Philosophy and Social Sciences Planning, "Research on the Collaborative Governance of Psychological Poverty among the Recently Poverty-Alleviated Population in Rural Guangdong" (GD23XSH09).

Conflicts of Interest

None declared.

References

1. 2023 national aging development bulletin. Ministry of Civil Affairs. 2023. URL: https://www.mca.gov.cn/n156/n2679/c1662004999980001751/attr/360830.pdf [accessed 2024-11-08]
2. The 56th statistical report on China's internet development [Report in Chinese]. : China Internet Network Information Center; 2025 Jul 21 URL: https://www.cnnic.net.cn/n4/2025/0721/c88-11328.html [accessed 2025-07-21]
3. Belic I, Winskel H, Allen K, Longstaff MG. Communication preferences and factors predicting smartphone addiction among four generations of Australians: gender and generational differences. J Technol Behav Sci 2025;10(2):373-381. [doi: 10.1007/s41347-024-00437-3]
4. Kim HR, Han MA. Associations between problematic smartphone use, unhealthy behaviors, and mental health status in Korean adolescents: based on data from the 13th Korea Youth Risk Behavior Survey (2017). Psychiatry Investig 2020 Dec;17(12):1216-1225. [doi: 10.30773/pi.2020.0007] [Medline: 34724602]
5. Luis SE. Body mass index, chronic health conditions, and depression among US adults with disabilities: a modeling study of the 2020–2023 BRFSS. Am J Health Educ 2026 Mar 4;57(2):93-104. [doi: 10.1080/19325037.2025.2559652]
6. Karaş H, Küçükparlak İ, Özbek MG, Yılmaz T. Addictive smartphone use in the elderly: relationship with depression, anxiety and sleep quality. Psychogeriatrics 2023 Jan;23(1):116-125. [doi: 10.1111/psyg.12910] [Medline: 36366976]
7. Wister A, Li L, Levasseur M, Kadowaki L, Pickering J. The effects of loneliness on depressive symptoms among older adults during COVID-19: longitudinal analyses of the Canadian Longitudinal Study on Aging. J Aging Health 2023 Jun;35(5-6):439-452. [doi: 10.1177/08982643221129686] [Medline: 36383045]
8. Oliveira LD, Abrantes GD, Ribeiro GDS, Cunha NM, Pontes MDLDF, Vasconcelos SC. Loneliness in senescence and its relationship with depressive symptoms: an integrative review. Rev Bras Geriatr Gerontol 2019;22(6):e190241. [doi: 10.1590/1981-22562019022.190241]
9. Ho M, Pullenayegum E, Fuller-Thomson E. Is social participation associated with successful aging among older Canadians? Findings from the Canadian Longitudinal Study on Aging (CLSA). Int J Environ Res Public Health 2023 Jun 6;20(12):6058. [doi: 10.3390/ijerph20126058] [Medline: 37372645]
10. Demir OB, Bilgin A, Yilmaz FT. Smartphone addiction among elderly individuals: its relationship with physical activity, activities of daily living, and balance levels. BMC Public Health 2025 Mar 11;25(1):965. [doi: 10.1186/s12889-025-22145-0] [Medline: 40069692]
11. Nawaz S. Rethinking classifications and metrics for problematic smartphone use and dependence: addressing the call for reassessment. Comput Human Behav Rep 2023 Dec;12:100327. [doi: 10.1016/j.chbr.2023.100327]
12. Park CS. Examination of smartphone dependence: functionally and existentially dependent behavior on the smartphone. Comput Human Behav 2019 Apr;93:123-128. [doi: 10.1016/j.chb.2018.12.022]
13. Ge J, Liu Y, Cao W, Zhou S. The relationship between anxiety and depression with smartphone addiction among college students: the mediating effect of executive dysfunction. Front Psychol 2022;13:1033304. [doi: 10.3389/fpsyg.2022.1033304] [Medline: 36710811]
14. Marinucci M, Pancani L, Aureli N, Riva P. Online social connections as surrogates of face-to-face interactions: a longitudinal study under Covid-19 isolation. Comput Human Behav 2022 Mar;128:107102. [doi: 10.1016/j.chb.2021.107102]
15. Caliandro A, Garavaglia E, Sturiale V, Di Leva A. Older people and smartphone practices in everyday life: an inquire on digital sociality of Italian older users. Commun Rev 2021 Jan 2;24(1):47-78. [doi: 10.1080/10714421.2021.1904771]

XSL•FO
RenderX

16. Sun JJ, Chang YJ. Associations of problematic binge-watching with depression, social interaction anxiety, and loneliness. Int J Environ Res Public Health 2021 Jan 28;18(3):1168. [doi: 10.3390/ijerph18031168] [Medline: 33525732]
17. Starosta J, Izydorczyk B, Wontorczyk A. Anxiety-depressive syndrome and binge-watching among young adults. Front Psychol 2021;12:689944. [doi: 10.3389/fpsyg.2021.689944] [Medline: 34335407]
18. Miller D, Abed Rabho L, Awondo P, et al. The Global Smartphone: Beyond a Youth Technology: UCL Press; 2021.
19. Luis SE. Bullying victimization and suicide among us high school adolescents: the mediating role of depressive symptoms and the moderating effects of physical activity and sex. J Sch Violence 2025:1-14. [doi: 10.1080/15388220.2025.2596991]
20. Luis SE, Veremu G, Namacala EL, Gama-Krummel A. A narrative theoretical review of research on factors associated with media use and mental health among adolescents and emerging adults. Mental Health Digital Technol 2025 Nov 27:1-23. [doi: 10.1108/MHDT-03-2025-0020]
21. Carstensen LL. Socioemotional selectivity theory: the role of perceived endings in human motivation. Gerontologist 2021 Nov 15;61(8):1188-1196. [doi: 10.1093/geront/gnab116] [Medline: 34718558]
22. Hu H, Xu W. Socioeconomic differences in digital inequality among Chinese older adults: results from a nationally representative sample. PLoS One 2024;19(4):e0300433. [doi: 10.1371/journal.pone.0300433] [Medline: 38564613]
23. Tang D. Use of the Geriatric Depression Scale-15 (GDS-15) in Chinese elderly. Chin J Clin Psychol 2013;21(3):402-405. [doi: 10.16128/j.cnki.1005-3611.2013.03.036]
24. Leung L. Linking psychological attributes to addiction and improper use of the mobile phone among adolescents in Hong Kong. J Child Media 2008 Jul;2(2):93-113. [doi: 10.1080/17482790802078565]
25. Li Y, Liu Q, Wang Y. The research on the usage behavior of TikTok short video platform in the elderly group. EHSS 2022;5:189-197. [doi: 10.54097/ehss.v5i.2901]
26. Sun T, Jin T, Zhang Q, Liu Y. Shared social account: understanding intergenerational collaboration and interaction on Xiaohongshu. Innov Aging 2024 Dec 31;8(Supplement_1):844-844. [doi: 10.1093/geroni/igae098.2733]
27. Ren HT, Zhang Z, Zhao YL, Du P. Development and validation of the Social Participation Scale for the elderly in community settings. Med Soc (Berkeley) 2024 Mar;37(3):92-99. [doi: 10.13723/j.yxysh.2024.03.015]
28. Pappas IO, Woodside AG. Fuzzy-set qualitative comparative analysis (fsQCA): guidelines for research practice in information systems and marketing. Int J Inf Manage 2021 Jun;58:102310. [doi: 10.1016/j.ijinfomgt.2021.102310]
29. Zhang S, Wang X, Wang H. Exploring configurations of social determinants for enhancing older adult health in China: an fuzzy-set qualitative comparative analysis based on 31 provinces in China. Front Public Health 2023 Oct 3;11:1255877. [doi: 10.3389/fpubh.2023.1255877]
30. Wang Y, Lu L, Zhang R, Ma Y, Zhao S, Liang C. The willingness to continue using wearable devices among the elderly: SEM and FsQCA analysis. BMC Med Inform Decis Mak 2023 Oct 16;23(1):218. [doi: 10.1186/s12911-023-02336-8] [Medline: 37845659]
31. Fiss PC. Building better causal theories: a fuzzy set approach to typologies in organization research. AMJ 2011 Apr;54(2):393-420. [doi: 10.5465/amj.2011.60263120]
32. Xu Y, Zeng K, Dong L, Zheng X, Si Y. Understanding older adults' smartphone addiction in the digital age: empirical evidence from China. Front Public Health 2023;11:1136494. [doi: 10.3389/fpubh.2023.1136494] [Medline: 37483945]
33. Ratcliffe J, Galdas P, Kanaan M. Older men and loneliness: a cross-sectional study of sex differences in the English Longitudinal Study of Ageing. BMC Public Health 2024 Feb 2;24(1):354. [doi: 10.1186/s12889-024-17892-5] [Medline: 38308255]
34. Tang D, Mair CA, Hu Q. Widowhood, social networks, and mental health among Chinese older adults: the moderating effects of gender. Front Psychol 2023;14:1142036. [doi: 10.3389/fpsyg.2023.1142036] [Medline: 37077844]
35. Liu Z, Li Z. Relationships between digital engagement and the mental health of older adults: evidence from China. PLoS One 2024;19(8):e0308071. [doi: 10.1371/journal.pone.0308071] [Medline: 39106268]
36. Dey M, Studer J, Schaub MP, et al. Problematic smartphone use in young Swiss men: its association with problematic substance use and risk factors derived from the pathway model. J Behav Addict 2019 Jun 1;8(2):326-334. [doi: 10.1556/2006.8.2019.17] [Medline: 31079472]

Abbreviations

DIT: digital inequality theory
fsQCA: fuzzy-set qualitative comparative analysis
RF: random forest
SST: socioemotional selectivity theory

XSL•FO
RenderX

Edited by M Mardini; submitted 23.Sep.2025; peer-reviewed by S Luis, Z Ye; revised version received 22.Dec.2025; accepted 10.Feb.2026; published 12.Mar.2026.

Please cite as:
Chen S, Song Y, Huang CC
Smartphone Addiction, Use Preferences, and Depression Among Older Adults in the Digital Context: Machine Learning Analysis of Survey Data
JMIR Aging 2026;9:e84703
URL: https://aging.jmir.org/2026/1/e84703
doi:10.2196/84703

© Sheng Chen, Yue Song, Chien-chung Huang. Originally published in JMIR Aging (https://aging.jmir.org), 12.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Self-Identified Age Cohorts and Personal Experience Sharing in Pseudonymous Online Spaces: Natural Language Processing Analysis of Reddit

Ern Chern Khor[1], MS; Moon Choi[1,2], PhD

[1]Graduate School of Science and Technology Policy, Korea Advanced Institute of Science and Technology, Daejeon, Daejeon, Republic of Korea
[2]Graduate School of Data Science, Korea Advanced Institute of Science and Technology, Daejeon, Daejeon, Republic of Korea

Corresponding Author:
Moon Choi, PhD
Graduate School of Science and Technology Policy
Korea Advanced Institute of Science and Technology
N4 Rm 1233
KAIST, 291 Daehak-ro, Yuseong-gu
Daejeon, Daejeon, 34141
Republic of Korea
Phone: 82 0423504850
Email: moonchoi365@gmail.com

Abstract

Background: While older adults' social media use has been widely studied for its instrumental benefits, such as accessing health information or maintaining family ties, research has largely focused on identity-based platforms that mirror offline social networks, leaving pseudonymous, interest-driven environments such as Reddit underexplored. Although older adults actively participate in these spaces to share personal narratives and engage beyond their existing social roles, the literature has yet to center their own voices, with most existing work focusing on caregivers or younger users discussing older adults rather than older adults speaking for themselves.

Objective: Integrating a life course perspective with the concept of subjective age cohort, this study examines how Reddit users who explicitly self-identify as being in their 50s, 60s, or 70s engage in self-expression and personal experience sharing within pseudonymous online spaces.

Methods: Using data collected via the Reddit application programming interface, this study analyzed posts and comments from 848 self-identified older Reddit users, including 488 future older adults in their 50s (57.5%) and 360 current older adults in their 60s and 70s (42.5%) identified through age-based flairs in the subreddit *r/AskOldPeople*. A fine-tuned large language model (BLOOM-560m; F_1-score=0.96) classified 4,055,275 sentences into three personal experience domains: (1) health and wellness, (2) personal relationships and identity, and (3) professional and financial life. Chi-square analysis compared domain distributions across age groups. BERTopic topic modeling identified thematic patterns within each domain.

Results: Of 569,107 personal experience sentences identified, personal relationships and identity comprised the largest share (n=268,212, 47.1%), followed by professional and financial (n=186,768, 32.8%) and health and wellness (n=114,127, 20.1%). Chi-square analysis revealed significant between-group differences (χ^2_2=34.7; P<.001): current older adults shared proportionally more about health and wellness, whereas future older adults shared more about relationships and identity. Topic modeling further revealed qualitatively distinct emphases within each domain. Future older adults' posts frequently discussed menopause, depression, and friendships, while current older adults more frequently addressed chronic illness, aging, and financial security. Both groups used pseudonymity to disclose sensitive or stigmatized topics.

Conclusions: This study demonstrates that older adults' personal experience sharing on pseudonymous platforms is shaped by the simultaneous interplay of life course stage, age cohort background, and platform affordances, rather than by any single factor. Pseudonymous environments such as Reddit enable disclosures of health vulnerabilities, contested social identities, and emotional experiences that are structurally discouraged on identity-based networks, and this has direct implications for how digital inclusion is conceptualized and practiced. Beyond operational literacy, supporting older adults' expressive inclusion in interest-based, pseudonymous online communities represents a meaningful yet underaddressed dimension of digital participation policy.

XSL•FO
RenderX

(JMIR Aging 2026;9:e90103) doi:10.2196/90103

KEYWORDS

older adults; age cohort; life course; social media; Reddit; pseudonymity; self-identification; self-disclosure; natural language processing; machine learning; social support; digital inclusion

Introduction

Background and Research Gap

Older adults' use of social media is often examined through the lens of instrumental benefits, such as accessing health-related information or maintaining contact with family and friends [1-3]. Yet, this perspective largely centers on identity-based platforms that replicate offline social networks. A growing body of research distinguishes between platforms that sustain identity-based networks, such as Facebook or Instagram, and those that are pseudonymous and interest-based, such as Reddit (Reddit, Inc), where users connect with strangers through shared topics rather than preexisting ties [4,5]. Unlike social networking sites grounded in real-name connections, these interest-driven environments encourage interaction that is less constrained by existing social roles and expectations.

Despite growing recognition of older adults' digital participation, a critical gap remains in the literature: a distinct lack of research that prioritizes older adults' own voices and personal sharing within pseudonymous online environments. While some studies have examined aging-related discussions on these platforms, most have focused on caregivers or younger users talking about older adults [6,7]. Given that older adults actively participate in various online environments to learn from and interact with younger generations and to share their own narratives and reflections [8,9], such pseudonymous spaces allowing users to interact free from existing offline social roles may therefore provide meaningful yet overlooked contexts for older adults to communicate and sustain social contact in digital spaces. This study addresses this gap by shifting the focus to older adults' self-expression on Reddit, examining how they navigate and share personal experiences when untethered from identity-based networks.

Theoretical Framework

From a social relations perspective, the convoy model provides a rationale for why such spaces can matter in later life [10,11]. As individuals age, their social networks, or convoys, tend to shrink due to retirement, mobility limitations, or loss of peers. Supplementary, low-barrier online exchanges can partially compensate for this contraction by maintaining a sense of embeddedness and continuity. While the convoy model explains why older adults may be drawn to online pseudonymous platforms, a life course perspective is needed to understand how they engage with them. In this perspective, age is understood not only as biological age but also as a series of transitional life phases [12]. Rather than viewing age as fixed, the life course approach conceptualizes it as a dynamic process in which each stage is shaped by distinct social and developmental characteristics [13]. These age-related characteristics influence media use alongside changing social circumstances [14]. Unlike rigid biological age, the life course perspective attempts to capture the social dynamics of individuals in relation to their broader societal context [15]. An individual's current life stage, such as transitioning out of the workforce, brings specific social conditions that can influence the types of experiences they wish to discuss with others.

Self-Identification on Pseudonymous Platforms

In pseudonymous spaces like Reddit, the intersection of life course and age cohort becomes highly visible through self-identification. Because Reddit relies on anonymity, a user's age group is invisible unless they explicitly disclose it. This act of voluntary disclosure can be understood through the lens of subjective age identification, which examines the age group to which a person self-identifies, regardless of their strict chronological age [13]. Often referred to as "identity age," this self-categorization grounds age at a subjectively experienced level rather than an objective one [13]. In this sense, one's age cohort self-identified on Reddit signals the social realities of their life course stage.

Within this integrated framework, pseudonymous, topic-based platforms such as Reddit offer an important setting for examining how these identities are expressed in practice. Their affordances, namely, default anonymity, topic segmentation, and interest-based exchange, enable older adults to share personal experiences outside the boundaries of identity-based networks, creating new conditions under which age-group–related patterns of sharing can be reconsidered [16]. On Reddit, users who explicitly self-identify as "60-something" or "70-something" publicly articulate a sense of belonging to an age cohort within a pseudonymous environment. This study focuses on people who consciously identify with these groups and make that identity visible online. Such acts of self-identification involve 2 complementary processes: *individuation*, through which members recognize common traits and shared experiences within the group, and *differentiation*, through which they mark their distinction from other groups [9,17].

Research Objective

The primary objective of this study is to examine how older adults engage in self-expression and personal experience sharing in pseudonymous online spaces. Rather than treating older adults as a monolithic age group, this study moves beyond demographic categories by integrating a life course perspective with the concept of subjective age cohort. Specifically, it investigates how Reddit users who explicitly self-identify as being in their 50s, 60s, or 70s share personal experiences online. Focusing on users who make their age cohort membership publicly visible, this study interprets such disclosures as performative expressions of subjective age cohort identity. This framing illuminates how people navigating the later stages of midlife and early older adulthood draw on pseudonymous

XSL•FO
RenderX

platforms to adapt their modes of self-expression to the particular demands and transitions of their current life stage.

Methods

Platform Context

This study adopted a mixed methods approach combining large-scale computational text analysis, statistical comparisons, and topic modeling. The analysis focused on data from Reddit users who explicitly self-identify as being in their 50s, 60s, or 70s. Reddit is a large-scale, forum-based platform where users post and comment within topic-specific communities called *subreddits*. Each subreddit is user-created and moderated, with its own rules, culture, and norms. Users can join and participate in multiple subreddits, allowing diverse discussions across interests and experiences. As of 2024, Reddit ranked as the ninth-most-used online platform in the United States [18]. Unlike identity-based social networks, Reddit operates under pseudonymity, where users participate under consistent usernames that are not tied to real-world identities. This allows them to maintain a recognizable online presence while preserving privacy [19]. Subreddits also allow the use of user flairs, which are short labels that appear beside usernames within a given subreddit. These flairs can indicate characteristics such as age group (eg, “60-something”), experience level (eg, “newbie”), or location (eg, “NYC”).

Data Collection

This study focused on *r/AskOldPeople*, a subreddit with a strong presence of older adults that facilitates intergenerational and peer discussions about aging and everyday life. As of June 2025, *r/AskOldPeople* had more than 800,000 members and ranked within the top 1% of all subreddits by size. To identify users’ age groups in this pseudonymous environment, the study used the subreddit’s flair system. In *r/AskOldPeople*, users may select flairs such as “50-something” or “70-something” to indicate their age. Data were collected via the Reddit Developer Platform application programming interface (API) in January 2024, retrieving users who posted or commented on the 1000 most recent posts, reflecting the limit imposed by Reddit’s API. Although this constrained the historical range of data, it ensured the inclusion of recently active users with up-to-date age flairs and minimized overrepresentation of highly active users. A total of 848 users were identified for analysis: 360 (42.5%) current older adults (“60-something” and “70-something” flairs) and 488 (57.5%) future older adults (“50-something”). Of the 360 current older adults, 304 (84.4%) were in their 60s, and 56 (15.6%) were in their 70s. These 2 age groups were combined into a single category to ensure a comparable sample size with the future older adults’ group.

Although *r/AskOldPeople* was used to identify older Reddit users, the analysis was not limited to that subreddit. Reddit’s structure allows public posts and comments across all subreddits to be retrieved for a given user. Using the Reddit API in February 2024, up to 1000 of the most recent posts and 1000 of the most recent comments were collected for each identified user. To ensure comparability across varying content lengths, all text was tokenized into sentences using the Punkt sentence tokenizer, which detects sentence boundaries in natural language text. This process yielded a dataset of 4,055,275 sentences: 1,671,400 (41.2%) from current older adults and 2,383,875 (58.8%) from future older adults.

Personal Sharing Classification

Given the dataset’s scale, a machine learning model was fine-tuned to classify each sentence into one of four categories: (1) health and wellness, (2) personal relationships and identity, (3) professional and financial, or (4) not related to personal experience sharing. After excluding the sentences unrelated to personal experience sharing, only the first 3 categories were analyzed. The classification schema was adapted from Geetha et al [20] and refined to include only personally reflective sentences. When sentences referenced multiple domains, the dominant theme was selected. For example, “My wife was diagnosed with stage IV cancer” was classified as *health and wellness*, while “I kept talking to my therapist when I started dating again after I divorced” was classified as *personal relationships and identity*. Sentences not describing personal experiences (eg, “The International Lyme and Associated Diseases Society recommends 4-6 weeks of antibiotics” or “Thank you so much”) were excluded.

The model used was BLOOM-560m [21], an open-source large language model available via Hugging Face. A balanced training set of 2000 manually labeled sentences (500 per category) was constructed through stratified random sampling and split into training (80%), validation (10%), and test (10%) subsets. The fine-tuned model achieved an F_1-score of 0.96 on the test set and was subsequently applied to label the full dataset. The fine-tuned model is publicly accessible at [22].

Ethical Considerations

This study used publicly available data collected via the official Reddit API. Data collection complied with the platform’s Terms of Service, and the study qualified for an Institutional Review Board exemption due to the lack of direct human subject interaction. Any direct quotes used for illustration were carefully paraphrased to prevent reverse searches.

Analytic Strategy

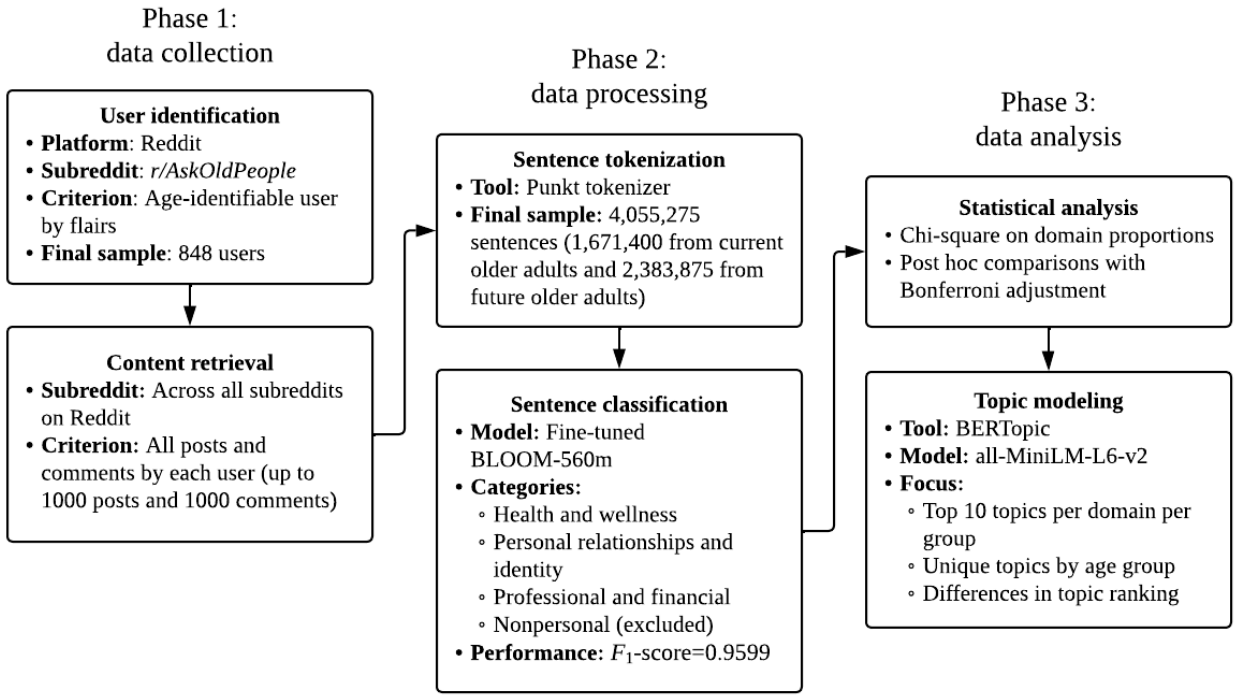
To examine age cohort differences in personal sharing, the proportions for each domain were calculated for both groups. A chi-square test was performed to assess whether the distribution of domains differed between age cohorts, with Bonferroni-adjusted post hoc comparisons conducted when significant.

To further explore thematic variation within each domain, topic modeling was applied using BERTopic, which uses transformer-based embeddings to cluster semantically related sentences into interpretable themes [23]. The SentenceTransformer model *all-MiniLM-L6-v2* was used to generate embeddings. Topic modeling was conducted separately for each domain and group. For each group, the top 10 topics were identified by frequency, represented by characteristic keywords automatically extracted by the model. The analysis focused on identifying topics unique to one group and differences in topic prominence between current and future older adults. This approach enabled a detailed comparison of

XSL•FO
RenderX

how distinct groups of older adults share personal experiences across life domains in pseudonymous online contexts. A visual summary of the data processing and analysis workflow is provided in Figure 1.

Figure 1. Data collection and analysis process.



Results

Overview of Personal Experience Sentences

After excluding nonpersonal content, a total of 569,107 sentences were identified as related to personal experiences. Among these, the "personal relationships and identity" domain accounted for the largest proportion (n=268,212, 47.1%), followed by the "professional and financial" (n=186,768, 32.8%) and "health and wellness" (n=114,127, 20.1%) domains.

Chi-Square Analysis of Domain Distribution

A chi-square test revealed a significant difference in the distribution of domains between future and current older adults (χ^2_2=34.7; P<.001), with Bonferroni-adjusted post hoc comparisons indicating that the professional and financial domain did not differ significantly between groups. Current older adults in their 60s and 70s shared proportionally more about health and wellness (46,379/227,458, 20.4%; P<.001), whereas future older adults in their 50s shared more about personal relationships and identity (161,925/341,649, 47.4%; P<.001). Table 1 summarizes the domain distribution and standardized residuals.

Table 1. Distribution of personal experience domains by age group.

Age group	Health and wellness		Personal relationships and identity		Professional and financial	
	Values, n (%)	R [a]	Values, n (%)	R	Values, n (%)	R
Future older adults (50s; n=341,649)	67,748 (19.8)	–5.17[b]	161,925 (47.4)	4.94[b]	111,976 (32.8)	–0.84
Current older adults (60s and 70s; n=227,458)	46,379 (20.4)	5.17[b]	106,287 (46.7)	–4.94[b]	74,792 (32.9)	0.84
Total (n=569,107)	114,127 (20.1)	—[c]	268,212 (47.1)	—	186,768 (32.8)	—

[a]R represents standardized residuals from the chi-square test.

[b]Residuals with Bonferroni-adjusted significance levels P<.001.

[c]Standardized residuals are not applicable to marginal totals.

Thematic Patterns by Domain

Health and Wellness

Table 2 presents the 10 most frequent topics by domain and age group. In the health and wellness domain, leading themes for future older adults included *smoking and nicotine use* (615/41,134, 1.5%), *cancer experiences* (460/41,134, 1.1%), *hormonal changes* (335/41,134, 0.8%), mental health discussions, such as *depression* (326/41,134, 0.8%), and *health care access* (280/41,134, 0.7%). Sentences commonly described

XSL•FO
RenderX

experiences such as treatment side effects (“The estrogen patch stopped my hot flashes”) or emotional responses to illness (“I am much more lonely now and prone to depression”).

For current older adults, the most frequent themes were *smoking* (409/28,376, 1.4%), *keto diet* (365/28,376, 1.3%), *positive health reflections* (296/28,376, 1.0%), *vaccination* (254/28,376, 0.9%), and *aging awareness* (206/28,376, 0.7%). Additional frequent topics included *dental issues* (195/28,376, 0.7%) and *Parkinson’s disease* (190/28,376, 0.7%). Example sentences referenced daily management of age-related conditions (“As I get older, I feel tired and discouraged”) and lifestyle adjustments (“I transitioned from the Standard American Diet to keto gradually”). While some topics overlapped (eg, smoking, vaccination, and cancer), their relative prominence varied between groups.

Personal Relationships and Identity

Both groups frequently discussed hobbies and everyday interests, but the emphasis of the discussion differed. For future older adults, top-ranked topics were *biking* (604/83,150, 0.7%), *racial identity* (527/83,150, 0.6%), *LGBTQ+ identity* (445/83,150, 0.5%), *photography* (419/83,150, 0.5%), and *books and reading* (383/83,150, 0.5%). Additional themes included *hairstyles* (353/83,150, 0.4%), *friendship* (302/83,150, 0.4%), and *movies* (283/83,150, 0.3%). Sentences often included midlife reflections on friendships (“I have five friends in my inner circle, and it’s all I need”).

Current older adults similarly mentioned *biking* (597/62,124, 1.0%) and *reading* (364/62,124, 0.6%), but also featured *gun ownership* (232/62,124, 0.4%) and *cooking* (230/62,124, 0.4%) among top topics. Examples included sharing recipes or memories of learning to cook for family (“I’m putting together a cookbook for my sons”). Across both groups, discussions of racial and LGBTQ+ identities were frequent, indicating engagement with broader conversations about social identity on the platform.

Professional and Financial Life

Work, income, and retirement were common themes for both groups, though with different foci. Leading topics for future older adults included *ticketing and consumer transactions* (345/63,052, 0.5%), *legal issues* (302/63,052, 0.5%), *work attire* (263/63,052, 0.4%), *retirement planning* (246/63,052, 0.4%), and *tipping practices* (238/63,052, 0.4%). Additional topics involved *job interviews* (205/63,052, 0.3%) and *COVID-related work disruptions* (204/63,052, 0.3%). Example sentences included practical reflections on everyday financial activity and employment: “I already have my tickets for next year,” “I’m a big tipper, so I usually go closer to 30 percent,” and “I have not gotten one single interview this year.”

For current older adults, frequent topics were *legal matters* (326/42,626, 0.8%), *retirement* (312/42,626, 0.7%), *401(k) and pension savings* (241/42,626, 0.6%), *social security* (193/42,626, 0.5%), and *charitable giving* (181/42,626, 0.4%). Other topics included *gender at work* (157/42,626, 0.4%) and *air travel* (148/42,626, 0.3%). Example sentences included: “I am 67, retired in August, and most basic expenses are covered by Social Security and a pension,” “I donate to charities every year,” and “I used to fly weekly for work for many years.” Overall, both groups contributed extensive discourse around financial security and employment experiences, though the specific subthemes differed in prominence.

Across domains, both current and future older users produced diverse personal narratives on Reddit. The largest quantitative differences appeared in the health and wellness and personal relationships and identity domains, whereas professional and financial content showed similar proportions. Topic modeling further revealed that overlapping themes, such as health conditions, hobbies, and work, were framed through distinct emphases in frequency and rank, reflecting the breadth of personal sharing among older Reddit users.

XSL•FO
RenderX

Table 2. Top topics by personal experience domain and age group with representative keywords.

Category and topic rank	Future older adults (50s)		Current older adults (60s and 70s)	
	Topic keywords (topic ID)	Value, n (%)[a]	Topic keywords (topic ID)	Value, n (%)
Health and wellness				
Top 1	“Smoking,” “nicotine,” “cigarettes” (T0)	615 (1.5)	“Smoking,” “nicotine,” “cigarettes” (T0)	409 (1.4)
Top 2	“Cancer,” “breast,” “stage” (T2)	460 (1.1)	“Keto,” “carb,” “recipes” (T5)	365 (1.3)
Top 3	“Estrogen,” “hormone,” “progesterone” (T10)	335 (0.8)	“Healthy,” “thanking,” “happiness” (T1)	296 (1.0)
Top 4	“Healthy,” “thanking,” “happiness” (T1)	327 (0.8)	“Vaccine,” “vaccines,” “vaccinated” (T3)	254 (0.9)
Top 5	“Depression,” “depressed,” “depressive” (T6)	326 (0.8)	“Aging,” “age,” “old” (T12)	206 (0.7)
Top 6	“Health care,” “insurance,” “universal” (T4)	280 (0.7)	“Health care,” “insurance,” “universal” (T4)	200 (0.7)
Top 7	“Nurses,” “nurse,” “nursing” (T7)	276 (0.7)	“Dentists,” “dental,” “teeth” (T9)	195 (0.7)
Top 8	“Vaccine,” “vaccines,” “vaccinated” (T3)	271 (0.7)	“Parkinson disease,” “cardinal,” “disease” (T55)	190 (0.7)
Top 9	“Diagnosed,” “officially,” “rediagnosed” (T8)	242 (0.6)	“Fasting,” “intermittent,” “fasts” (T29)	183 (0.6)
Top 10	“Keto,” “carb,” “recipes” (T5)	238 (0.6)	“Cancer,” “breast,” “stage” (T2)	176 (0.6)
Personal relationships and identity				
Top 1	“Bike,” “bikes,” “ride” (T0)	604 (0.7)	“Bike,” “bikes,” “ride” (T0)	597 (1.0)
Top 2	“White,” “racist,” “black” (T1)	527 (0.6)	“Book,” “read,” “library” (T3)	364 (0.6)
Top 3	“Gay,” “straight,” “LGBTQ” (T2)	445 (0.5)	“White,” “racist,” “black” (T1)	363 (0.6)
Top 4	“Photos,” “photo,” “pictures” (T4)	419 (0.5)	“Gay,” “straight,” “LGBTQ” (T2)	309 (0.5)
Top 5	“Book,” “read,” “library” (T3)	383 (0.5)	“Photos,” “photo,” “pictures” (T4)	286 (0.5)
Top 6	“Hair,” “curly,” “bangs” (T5)	353 (0.4)	“Hair,” “curly,” “bangs” (T5)	280 (0.5)
Top 7	“Blonde,” “gray,” “hair” (T7)	322 (0.4)	“Gun,” “rifle,” “shooting” (T10)	232 (0.4)
Top 8	“Friendships,” “acquaintances,” “friends” (T9)	302 (0.4)	“Cook,” “recipes,” “cookbook” (T8)	230 (0.4)
Top 9	“Wedding,” “ceremony,” “bride” (T6)	295 (0.4)	“Blonde,” “gray,” “hair” (T7)	223 (0.4)
Top 10	“Movie,” “theater,” “film” (T13)	283 (0.3)	“Wedding,” “ceremony,” “bride” (T6)	206 (0.3)
Professional and financial				
Top 1	“Tickets,” “ticketing,” “queue” (T3)	345 (0.5)	“Lawyer,” “attorney,” “legal” (T0)	326 (0.8)
Top 2	“Lawyer,” “attorney,” “legal” (T0)	302 (0.5)	“Retire,” “retirement,” “enjoying” (T1)	312 (0.7)
Top 3	“Wear,” “dress,” “suit” (T4)	263 (0.4)	“401k,” “ira,” “roth” (T2)	241 (0.6)
Top 4	“Retire,” “retirement,” “enjoying” (T1)	246 (0.4)	“Security,” “social,” “collect” (T26)	193 (0.5)
Top 5	“Tip,” “tipping,” “delivery” (T6)	238 (0.4)	“Donate,” “charity,” “charities” (T5)	181 (0.4)
Top 6	“401k,” “ira,” “roth” (T2)	229 (0.4)	“Wear,” “dress,” “suit” (T4)	164 (0.4)
Top 7	“Name,” “named,” “nickname” (T7)	229 (0.4)	“Women,” “gender,” “men” (T14)	157 (0.4)
Top 8	“Paid,” “pay,” “skeptic” (T9)	206 (0.3)	“Covid,” “precovid,” “lockdown” (T8)	148 (0.3)
Top 9	“Interview,” “interviews,” “interviewed” (T12)	205 (0.3)	“Fly,” “airline,” “pilot” (T10)	148 (0.3)
Top 10	“Covid,” “precovid,” “lockdown” (T8)	204 (0.3)	“Paid,” “pay,” “skeptic” (T9)	141 (0.3)

[a]Percentages represent the distribution among sentences successfully assigned to a theme, excluding unassigned outlier sentences (topic -1).

XSL•FO
RenderX

Discussion

Principal Findings

This study investigated how self-identified older Reddit users share personal experiences across 3 domains (ie, health and wellness, personal relationships and identity, and professional and financial life) and found that age-group differences emerged both at the domain level and within specific topics. The central analytic contribution is not merely that the 2 groups discuss different things but that the nature of those differences points to distinct underlying forces: life course position most powerfully structures sharing in health and professional domains, while age cohort socialization more visibly shapes expression in the relationships and identity domain. Meanwhile, pseudonymous platform affordances function as an enabling condition across all 3 domains, facilitating disclosures that identity-based networks structurally discourage. Taken together, these patterns suggest that older adults' online self-expression cannot be adequately explained by any single framework; rather, it reflects the simultaneous, domain-specific interplay among aging, cohort, and platform.

Health and Professional Domains: Life Course Position as the Primary Driver

The sharpest between-group contrasts in health and professional content are best understood as reflections of where each cohort stands in the life course. In health and wellness, the topics that distinguish the 2 groups correspond closely to the physiological and institutional realities of their respective life stages: hormonal change, depression, and nursing interactions for those in their 50s; Parkinson disease, dental decline, and aging awareness for those in their 60s and 70s. This pattern is consistent with life course theory's premise that age-related transitions bring distinct social and bodily concerns that shape what individuals perceive as meaningful to communicate [24-28]. Crucially, neither group's health discourse is simply about illness; both articulate experiences of bodily negotiation, that is, managing a changing body and its social consequences. What differs is the temporal orientation: future older adults write from a position of midlife adjustment and medical navigation, while current older adults write from one of chronic management and embodied constraint. This temporal dimension suggests that health-related sharing on Reddit may serve less as information exchange and more as a form of life-stage sense-making.

A parallel logic applies to professional and financial content. Future older adults' emphasis on job interviews, workplace roles, and consumer transactions reflects continued embeddedness in the labor market, whereas current older adults' focus on Social Security, pensions, and charitable giving reflects the financial and moral landscape of postretirement life. Notably, the shift from employment-centered to giving-centered discourse among current older adults is not merely a change in topic. It signals a reorientation of economic identity from productivity to stewardship. This aligns with life course research on the transition from midcareer to later adulthood, in which individuals increasingly frame financial decisions through a lens of legacy and responsibility rather than accumulation [29-32]. The presence of this discourse on a pseudonymous platform is itself significant: it implies that Reddit serves not only as a space for practical information-seeking but also as a venue for working through identity transitions that are rarely voiced in face-to-face settings.

Relationships and Identity: Cohort Culture as the Primary Driver

If life course position explains health and professional differences, it is less adequate for understanding the divergences in personal relationships and identity. Their expressive content in this domain differs in ways that point more directly to age cohort socialization, which includes the enduring cultural habits, values, and media landscapes formed during each cohort's formative years [33]. Future older adults' more frequent discussions of friendships and film suggest that this cohort's sense of self is significantly organized around shared cultural memory and relational intimacy. This cohort, consisting of Generation X, came of age during an era of expanded media consumption and shifting norms around personal expressiveness, and these formative experiences appear to persist as frameworks for identity and connection in midlife [34]. Current older adults' greater emphasis on cooking and gun ownership is analytically distinct: rather than nostalgia or relational affirmation, these topics function as articulations of responsibility, continuity, and moral selfhood. Research on boomer generational culture documents a persistent orientation toward family stewardship and civic identity that was instilled during early socialization and continues to anchor self-narratives in later life [35-37]. On Reddit, these values are not merely expressed; they are performed for an anonymous audience, suggesting that even in pseudonymous spaces, generational identity remains a meaningful resource for self-presentation.

The cross-group similarities in this domain are equally important analytically. Both groups engaged substantively with race and LGBTQ+ identity, topics that carry heightened social risk for older cohorts in offline contexts. This convergence challenges assumptions that older adults' online engagement is primarily conservative or identity-confirming, and points instead to Reddit's capacity to elicit discursive engagement with contested social identities across age lines. This finding is better explained by platform affordance than by either life course or generational factors, which sets up the argument in the following section.

Pseudonymity as an Enabling Condition for Authentic Self-Expression

Across all 3 domains, the patterns of sharing observed in this study are not fully intelligible without accounting for Reddit's pseudonymous architecture. Pseudonymity lowers inhibitions and restructures the environment of self-disclosure, enabling users to share content that would carry social risk in identity-bound contexts [38,39]. This mechanism is most visible in health-related sharing in this study. Both age cohort groups disclosed experiences involving vulnerability, such as mental health struggles, bodily decline, and stigmatized conditions, topics that research consistently shows to be suppressed in face-to-face interactions and on real-name platforms due to stigma, self-restraint norms, and the fear of burdening others [40,41]. The willingness to share such content on Reddit is not simply a function of digital confidence; it reflects the structural

XSL•FO
RenderX

condition that pseudonymity enables disclosure without consequences for one's offline social standing. This aligns with the conceptualization of networked privacy by De Choudhury and De [42], in which users actively calibrate visibility and intimacy through selective, context-dependent disclosure, a practice that pseudonymous platforms uniquely support.

In the relationships and identity domains, pseudonymity's enabling role takes a different form. It expands the boundaries of moral and social expression, not only emotional disclosure. The engagement of both groups with race and LGBTQ+ identity, sensitive topics that generate social friction in many older adults' offline communities, suggests that Reddit functions as a low-risk venue for articulating positions and experiences that would otherwise be carefully managed in identity-tied settings. This expressive expansion extends to loneliness, marital life, nostalgia, and everyday frustration—contents that are affectively authentic but socially inadmissible on platforms where one's full social identity is visible [43,44].

Taken together, these patterns position Reddit not as a neutral medium but as a structurally distinctive space that actively reconfigures what older adults are willing to say, and to whom. The act of self-identifying with an age cohort through a flair compounds this effect as users are not merely anonymous individuals but members of a loosely assembled, voluntarily constituted community of peers—a social context that provides the dual affordance of relatability and anonymity. This combination may be especially consequential for older adults, whose offline social networks, as convoy theory predicts, are contracting precisely at the life stage when the need for resonant peer exchange may be intensifying [10,11].

Design and Policy Implications

These findings carry implications that go beyond platform research to inform *how digital environments are designed* and *how digital inclusion is conceptualized for older populations*. The results suggest that conventional digital inclusion frameworks, which tend to center operational literacy and family connectivity via identity-based platforms, leave a critical dimension unaddressed: what might be called *expressive inclusion*—the capacity to communicate authentically within a social frame that is not constrained by one's offline identity and relationships. The finding that pseudonymous spaces enable disclosures about mental health, physical vulnerability, and contested social identities suggests that access to such digital spaces constitutes a distinct and meaningful dimension of digital participation that identity-centric interventions cannot replace. This has practical implications for both platform design and community-based programming. Designers of eHealth platforms and digital aging services should consider integrating pseudonymous or semianonymous communication options alongside identity-tied profiles, particularly for forums addressing sensitive topics such as mental health, midlife transitions, or chronic illness. Evidence from this study suggests that such affordances may substantially increase candor and peer engagement among older users. For digital literacy practitioners, the implication is similarly concrete: curricula could incorporate "pseudonymous navigation," teaching older adults how to identify, evaluate, and safely participate in interest-based online communities alongside the operational and security competencies that currently dominate training programs. Equipping older adults to find resonant peer communities outside their immediate offline networks would directly address the social contraction that convoy theory associates with later life.

Limitations

Several limitations should be acknowledged when interpreting the findings. First, the study operationalizes age groups through user-selected age-based flairs. It is important to acknowledge that these labels reflect a voluntary self-categorization within a digital space, rather than objectively verified demographic membership. This choice may also bias the sample toward Reddit users who are both more digitally confident and more comfortable identifying publicly with their age. These individuals likely represent a distinctive subset of older adults. Findings should therefore be understood as characterizing active, self-identifying older Reddit users, who are a distinctive and probably more digitally agentic subset rather than older adults in general.

A second and more fundamental limitation is the cross-sectional design, which makes it impossible to disentangle life course effects from cohort effects. The differences observed between the 50s and the 60s and 70s groups could reflect either the distinct life-stage position of each group, the different generational cultures that shaped them, or their interaction. This study treats the two explanatory frameworks as complementary rather than competing but cannot empirically adjudicate their relative contributions. Doing so would require a longitudinal or cohort-sequential design capable of tracking how the same cohort's expressive patterns shift as it moves through subsequent life stages, and this is a methodologically demanding yet theoretically essential direction for future research.

Acknowledgments

A preliminary version of this work was presented at the Gerontological Society of America 2024 Annual Scientific Meeting, Seattle, Washington, in November 2024.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT; RS-2024-00348379). ECK gratefully acknowledges the financial support from the Hyundai Motor Chung Mong-Koo Foundation Global Scholarship.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Jung EH, Walden J, Johnson AC, Sundar SS. Social networking in the aging context: why older adults use or avoid Facebook. Telemat Inform 2017 Nov;34(7):1071-1080. [doi: 10.1016/j.tele.2017.04.015]
2. Myhre JW, Mehl MR, Glisky EL. Cognitive benefits of online social networking for healthy older adults. J Gerontol B Psychol Sci Soc Sci 2017 Sep 01;72(5):752-760. [doi: 10.1093/geronb/gbw025] [Medline: 26984523]
3. Teano AL, Scott A, Gipson C, Albert M, Pettigrew C. Social media programs for outreach and recruitment supporting aging and alzheimer disease and related dementias research: longitudinal descriptive study. JMIR Aging 2024 Jul 09;7:e51520 [FREE Full text] [doi: 10.2196/51520] [Medline: 38981112]
4. Chen M, Li W. Social capital development on interest-based networks: examining its antecedents, process, and consequences. Humanit Soc Sci Commun 2024 Jan 31;11:190. [doi: 10.1057/s41599-024-02609-1]
5. Vasist PN, Krishnan S, Agnihotri P. The unfolding of geopolitical tensions on social networks: a social network analysis of Twitter and Reddit conversations. Internet Res 2026;36(2):789-821. [doi: 10.1108/INTR-02-2024-0155]
6. Huang NM, Timothy B, Ho SS. Stressors disclosed on Reddit by caregivers of older adults and social support received: content analysis. JMIR Aging 2025 Sep 05;8:e71452 [FREE Full text] [doi: 10.2196/71452] [Medline: 40911350]
7. Pickett AC, Valdez D, Sinclair KL, Kochell WJ, Fowler B, Werner NE. Social media discourse related to caregiving for older adults living with Alzheimer disease and related dementias: computational and qualitative study. JMIR Aging 2024 Jun 19;7:e59294 [FREE Full text] [doi: 10.2196/59294] [Medline: 38896462]
8. Harley D, Fitzpatrick G. Creating a conversational context through video blogging: a case study of Geriatric1927. Comput Hum Behav 2009 May;25(3):679-689. [doi: 10.1016/j.chb.2008.08.011]
9. Siibak A, Vittadini N. Editorial: Introducing four empirical examples of the “generationing” process. Cyberpsychology 2012 Sep 01;6(2). [doi: 10.5817/cp2012-2-1]
10. Antonucci TC, Ajrouch KJ, Manalel JA. Social relations and technology: continuity, context, and change. Innov Aging 2017 Nov;1(3):igx029 [FREE Full text] [doi: 10.1093/geroni/igx029] [Medline: 29795794]
11. Wang S, Carmeline A, Kolko B, Munson SA. Understanding the role of technology in older adults' changing social support networks. Proc ACM Hum Comput Interact 2024 Apr 26;8(CSCW1):1-28. [doi: 10.1145/3641027]
12. Hutchison ED. A life course perspective. In: Dimensions of Human Behavior: The Changing Life Course. Thousand Oaks, CA: SAGE Publication; 2010.
13. Settersten RA, Mayer KU. The measurement of age, age structuring, and the life course. Annu Rev Sociol 1997;23:233-261. [doi: 10.1146/annurev.soc.23.1.233]
14. Bolin G, Skogerbø E. Age, generation and the media. North Lights 2013;11(1):3-14. [doi: 10.1386/nl.11.1.3_2]
15. Burnett J. Generations: The Time Machine in Theory and Practice. Milton Park, UK: Routledge; 2016.
16. Massanari AL. Participatory Culture, Community, and Play: Learning from Reddit. Lausanne, Switzerland: Peter Lang; 2015.
17. Corsten M. The time of generations. Time Soc 1999 Sep;8(2-3):249-272. [doi: 10.1177/0961463X99008002003]
18. Social media fact sheet. Pew Research Center. 2025. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2026-04-22]
19. Panek ET. Understanding Reddit. Milton Park, UK: Routledge; 2021.
20. Geetha R, Karthika S, Kumaraguru P. Tweet-scan-post: a system for analysis of sensitive private data disclosure in online social media. Knowl Inf Syst 2021 Jul 31;63:2365-2404. [doi: 10.1007/s10115-021-01592-2]
21. BigScience Workshop. BLOOM: a 176B-parameter open-access multilingual language model. arXiv Preprint posted online on November 9, 2022. [doi: 10.48550/arXiv.2211.05100]
22. Model card: BLOOM-560m for personal sharing classification. Hugging Face. URL: https://huggingface.co/ernchern/personal_info_classification [accessed 2026-04-22]
23. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online on March 11, 2022. [doi: 10.48550/arXiv.2203.05794]
24. Bamford A, Henderson EJ. Parkinson's disease in older people. Medicine 2021 Jan;49(1):56-61. [doi: 10.1016/j.mpmed.2020.10.008]
25. Chan AK, Tamrakar M, Jiang CM, Lo EC, Leung KC, Chu CH. Common medical and dental problems of older adults: a narrative review. Geriatrics (Basel) 2021 Aug 06;6(3):76 [FREE Full text] [doi: 10.3390/geriatrics6030076] [Medline: 34449647]
26. Domenichiello AF, Ramsden CE. The silent epidemic of chronic pain in older adults. Prog Neuropsychopharmacol Biol Psychiatry 2019 Jul 13;93:284-290 [FREE Full text] [doi: 10.1016/j.pnpbp.2019.04.006] [Medline: 31004724]
27. Soares CN. Menopause and mood: the role of estrogen in midlife depression and beyond. Psychiatr Clin North Am 2023 Sep;46(3):463-473. [doi: 10.1016/j.psc.2023.04.004] [Medline: 37500244]

XSL•FO
RenderX

28. Wood K, McCarthy S, Pitt H, Randle M, Arnot G, Thomas SL. "Hiding symptoms and balancing work, family and relationships": Australian women discuss menopause and the midlife collision. Soc Sci Med 2025 Dec;387:118681 [FREE Full text] [doi: 10.1016/j.socscimed.2025.118681] [Medline: 41110246]
29. Kuitto K, Kuivalainen S, Möhring K. Inequalities in pensions and retirement from a life course perspective: an introduction. Soc Policy Adm 2021 Apr 07;55(3):403-408. [doi: 10.1111/spol.12663]
30. Mitchell OS, Clark R, Lusardi A. Income trajectories in later life: longitudinal evidence from the Health and Retirement Study. J Econ Ageing 2022 Jun;22:100371 [FREE Full text] [doi: 10.1016/j.jeoa.2022.100371] [Medline: 36156898]
31. Okulicz-Kozaryn A, Morawski L. A similar effect of volunteering and pensions on subjective wellbeing of elderly. Econ Sociol 2021 Jun;14(2):11-39. [doi: 10.14254/2071-789X.2021/14-2/1]
32. Thomas T. Remaining active in the labour market: Trends and characteristics of the over 50s. Ir J Manag 2019;38(1):16-25. [doi: 10.2478/ijm-2018-0007]
33. Bolin G. Media Generations: Experience, Identity and Mediatised Social Change. Milton Park, UK: Routledge; 2016.
34. Brabazon T. From Revolution to Revelation: Generation X, Popular Memory and Cultural Studies. Milton Park, UK: Routledge; Sep 29, 2017.
35. Simmons D, Chapman GE. The significance of home cooking within families. Br Food J 2012 Aug;114(8):1184-1195. [doi: 10.1108/00070701211252110]
36. Yamane D. The sociology of U.S. gun culture. Sociol Compass 2017 Jun 16;11(7):e12497. [doi: 10.1111/soc4.12497]
37. Edmunds J, Turner · BS. Generations, Culture and Society. London, UK: Open University Press; 2002.
38. Xu H, Zhang P. The influence of anonymity and social ties on personal experience sharing: A comprehensive mixed-methods study. Proc ACM Hum Comput Interact 2025 Jan 10;9(1):1-22. [doi: 10.1145/3701210]
39. Clark-Gordon CV, Bowman ND, Goodboy AK, Wright A. Anonymity and online self-disclosure: A meta-analysis. Commun Rep 2019 May 07;32(2):98-111. [doi: 10.1080/08934215.2019.1607516]
40. Moore A, Grime J, Campbell P, Richardson J. Troubling stoicism: Sociocultural influences and applications to health and illness behaviour. Health (London) 2013 Mar;17(2):159-173 [FREE Full text] [doi: 10.1177/1363459312451179] [Medline: 22763794]
41. Rydberg Sterner T, Dahlin-Ivanoff S, Gudmundsson P, Wiktorsson S, Hed S, Falk H, et al. 'I wanted to talk about it, but I couldn't', an H70 focus group study about experiencing depression in early late life. BMC Geriatr 2020 Dec 07;20(1):528 [FREE Full text] [doi: 10.1186/s12877-020-01908-x] [Medline: 33287708]
42. De Choudhury M, De S. Mental health discourse on Reddit: self-disclosure, social support, and anonymity. Proc Int AAAI Conf Weblogs Soc Media 2014 May 16;8(1):71-80. [doi: 10.1609/icwsm.v8i1.14526]
43. Suler J. The online disinhibition effect. Cyberpsychol Behav 2004 Jun;7(3):321-326. [doi: 10.1089/1094931041291295] [Medline: 15257832]
44. Bargh JA, McKenna KY, Fitzsimons GM. Can you see the real me? Activation and expression of the “true self” on the internet. J Soc Issues 2002 Dec 17;58(1):33-48. [doi: 10.1111/1540-4560.00247]

Abbreviations

API: application programming interface

Edited by AF Abdulai; submitted 21.Dec.2025; peer-reviewed by A Motel-Klingebiel, P Rehor; comments to author 12.Mar.2026; revised version received 10.Apr.2026; accepted 15.Apr.2026; published 12.May.2026.

Please cite as:
Khor EC, Choi M
Self-Identified Age Cohorts and Personal Experience Sharing in Pseudonymous Online Spaces: Natural Language Processing Analysis of Reddit
JMIR Aging 2026;9:e90103
URL: https://aging.jmir.org/2026/1/e90103
doi:10.2196/90103
PMID:

©Ern Chern Khor, Moon Choi. Originally published in JMIR Aging (https://aging.jmir.org), 12.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Developing Customized Personas to Capture Intrinsic Capacity Profiles and Digital Monitoring Intentions in Older Adults: Mixed Methods Study

Zehui Xuan[1], BSc; Yirou Niu[1], MM; Ruifu Kang[1], BSc; Qian Xiao[1], PhD; Shuai Jin[1], PhD; Jie Zhao[2], MM; Yanling Wang[1*], MM; Hong Chang[2*], MM

[1]School of Nursing, Capital Medical University, Beijing, China

[2]Department of Neurology, Xuanwu Hospital, Capital Medical University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Yanling Wang, MM
School of Nursing
Capital Medical University
No.10 Xi-tou-tiao, You-an-men Wai, Fengtai District
Beijing, 100069
China
Phone: 86 010 83911644
Email: wangyanling@ccmu.edu.cn

Abstract

Background: Integrated Care for Older People (ICOPE), focused on monitoring and optimizing the intrinsic capacity (IC) of older adults, is a new model of geriatric care that is currently being accelerated globally. Digital health technologies are recommended for longitudinal IC monitoring to provide precise and timely interventions. However, little is known about the psychological intentions of engaging in digital monitoring of IC according to the profile heterogeneity of IC among older adults.

Objective: This study aims to map a set of customized personas to capture the profiles of IC and match psychological intentions that support personalized digital IC monitoring.

Methods: An explanatory sequential mixed methods study was conducted at 16 sites in Beijing, China. Older adults aged ≥60 years (n=481) were selected to complete the quantitative survey. Latent profile analysis, descriptive statistics, and logistic regression analyses were performed to cluster subgroups using Mplus (Muthén & Muthén) and SPSS (IBM Corp). A subsample of participants from each profile (n=25) was purposively sampled for semistructured interviews. An inductive-deductive content analysis was used to identify similar attributes and to affirm the personas gradually. A joint statistical and thematic visualization method was used to integrate the customized personas.

Results: Three profiles of IC patterns emerged: “multisubdomain decline–IC imbalance group,” “multisubdomain moderate–sensory deficit group,” and “multisubdomain robust–whole balance group.” The distribution of latent profiles was influenced by age, education, monthly per capita household income, self-rated health, and number of chronic diseases, while positively impacting older adults’ functional ability. The following customized personas were captured regarding established themes: “affects my mood—anxious evader,” characterized by avoidance and anxiety, low digital interest, and perceived social isolation; “capitalize on what comes—accommodative adopter,” pragmatically oriented toward disease detection, with moderate digital openness but limited self-efficacy; and “more autonomy—active improver,” who exhibited proactive engagement, high digital literacy, and motivation rooted in self-management and social participation.

Conclusions: This study is the first to integrate latent profile analysis with customized qualitative personas to link the heterogeneity of IC with the psychological intentions underlying digital monitoring. The resulting personas model provides an actionable framework for tailoring digital IC monitoring strategies in community-based integrated care. The findings emphasize the need to align monitoring approaches with older adults’ IC characteristics, psychological readiness, digital literacy, and social support to enhance engagement in digital IC monitoring.

(JMIR Aging 2026;9:e82867) doi:10.2196/82867

XSL•FO
RenderX

KEYWORDS

digital monitoring; integrated care for older people; intrinsic capacity; latent profile analysis; mixed methods study; older adults; personas

Introduction

The Aging Challenge and the Centrality of Intrinsic Capacity

In the face of increasing aging trends, health care systems are confronted with a diverse and intricate array of geriatric care demands [1]. The transition from a model of health care that is centered on responding to diseases to one that focuses on maintaining functions is an urgent strategy for improving and coordinating health care services [2]. Intrinsic capacity (IC), commonly consisting domains of cognition, psychological capacity, sensory capacity (including vision and hearing), vitality, and locomotor capacity, is the physiological reserve of an individual's functional ability and plays an important role in quality of later life [3,4]. It is defined as "the composite of all the physical and mental capacities that an individual can draw on" [1]. To operationalize this concept, the World Health Organization (WHO) has released 2 editions of the "Integrated Care for Older People (ICOPE): Guidance for Person-centered Assessment and Pathways in Primary Care" handbook to guide the assessment and monitoring of older adults' IC in community settings [1,5].

IC, as a health metric reflecting the degree of healthy aging in a country or region, is of special reference value [6]. A systematic review stated that the global prevalence of IC decline among community-dwelling older adults was 67.8% [7]. In China, Ma et al [8] reported that 69.1% of participants showed a decline in 1 or more domains of IC. A longitudinal study of retired Chinese older adults revealed that 32.1% experienced a decline in IC over a 4-year follow-up period [9]. According to previous studies, IC is a robust predictor of adverse health outcomes, including frailty, falls, health care use, incident disability, hospitalization, and mortality [10,11]. In general, optimizing IC in older adults can enhance functional autonomy and predict adverse events. Consequently, dynamic monitoring of cognitive impairment is needed to identify risk factors that can be ameliorated and to develop individualized intervention strategies that delay or prevent the deterioration of IC.

Digital Health for IC Monitoring: Potential and the Challenge of Equity

Digital health technology is an efficient and manageable approach for expanding the study and monitoring of IC in older adults [12]. An ICOPE program in France used digital tools to establish a timely feedback mechanism and support network for early warning IC monitoring in community settings, enabling interaction between older adults and health care professionals and providing an interface for subsequent steps in ICOPE [13-15]. A qualitative study indicated that implementing eHealth interventions within the IC framework may promote the social participation of older adults and communication with physicians, enhancing subjective well-being and enriching the monitoring activities of medical staff [16]. Currently underway in China, the "Comprehensive Evaluation of Intrinsic Capacity in China Study" (HEALTHY) aims to develop a consolidated digital technology–based platform for assessing IC and tailoring personalized interventions to promote proactive health behaviors among Chinese older adults [12]. However, older adults' motivations for participating in digital health programs are complex and influenced by multiple factors [17]. Several studies have shown that older adults who engage in IC-focused digital health programs tend to be younger and healthier [13,14], which appears to conflict with the health equity goals advocated by ICOPE. The psychological intentions of older adults have a profound effect on their interaction with technology and their ability to use digital tools [18]. Therefore, establishing evidence-based links between IC heterogeneity and digital engagement patterns is an important process for scaling up the individual-centered ICOPE model and promoting equity in digital health for older adults.

A Persona-Based Design for Customized IC Monitoring Strategies

The concepts of people-oriented integrated care and user-centered digital design achieve similar purposes and effects, both of which aim to meet the specific needs and preferences of the target group and achieve a deep and intuitive understanding [19]. Personas are a valuable approach for capturing older adults' preferences, abilities, and attitudes toward self-management of health, particularly in the use of digital technology [20]. The novel paradigm of IC as a health criterion undoubtedly impacts the traditional disease-centered perceptions of older adults. Previous studies exploring the heterogeneity of IC subgroups have neglected to provide insight into older adults' psychological intentions regarding IC monitoring and management [21,22]. Strategies for adopting digital health technologies may be hindered by a lack of demand or skills if only the external characteristics of older adults are targeted and the underlying behavioral determinants of accepting IC monitoring are overlooked [23]. Customized personas, with the advantage of incorporating biopsychosocial characteristics that encompass needs, experiences, and expectations, are an important reference for clustering older adults for digital IC monitoring [24].

Study Purpose

Few studies have combined quantitative analysis with additional data to construct accurate customized personas, which are usually generated through gradual affinity grouping after labeling qualitative data [25,26]. To the best of the authors' knowledge, this study is the first to adopt a mixed methods approach aimed at exploring the psychological intentions of older adults to receive integrated care focused on digital IC monitoring, thereby supporting personalized monitoring activities for community health care providers.

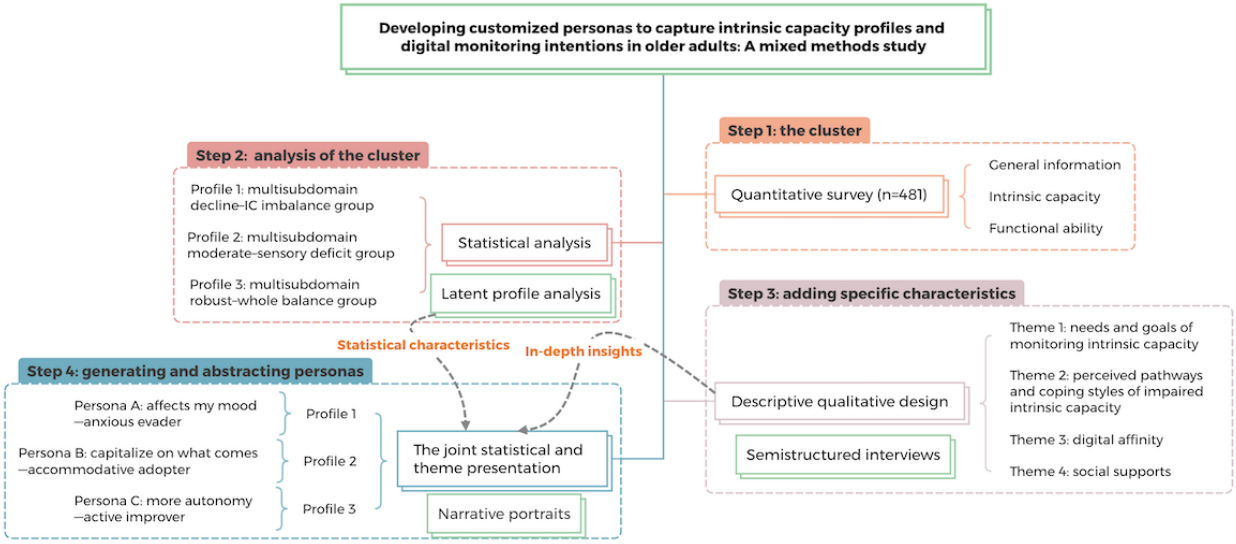
XSL•FO
RenderX

Methods

Study Procedures

This study used an explanatory sequential mixed methods design, which was implemented in 2 consecutive phases [27]. First, a quantitative phase was conducted using structured questionnaires to collect objective data and identify key patterns or clinically significant issues. Subsequently, a qualitative phase followed, in which in-depth interviews were carried out with a purposively selected subgroup of participants to explore and interpret the underlying reasons and contextual meanings behind the quantitative findings [28]. The following hypotheses were proposed in this study: (1) by using the 5 domains of IC as observational indicators, older adults will be clustered into distinct groups, reflecting differences in IC levels and subdomain characteristics; and (2) older adults in differently characterized IC subgroups will express different motivations, goals, and concerns regarding digital monitoring of IC. A total of 3 PhD candidates in nursing jointly developed the research protocol under the guidance of 2 nursing professors, 1 associate professor, 1 lecturer, and 1 senior nurse practitioner, all of whom possesses extensive research or clinical practice experience in geriatric or community nursing. Figure 1 illustrates the 4 steps of the study [29], including: (1) the cluster: questionnaire survey on IC of older adults; (2) analysis of the cluster: to gain an initial understanding of differential characteristics and classification of subgroups based on latent profile analysis (LPA); (3) qualitative interviews to explore the attitudes, goals, needs, and potential concerns of older adults for digital monitoring under the different profiles of IC; and (4) integration and corroboration of quantitative and qualitative findings to portray a set of personas of older adults' IC profiles and matched digital monitoring intentions, laying the groundwork for the future development of personalized monitoring strategies. Reporting adhered to the Good Reporting of a Mixed Methods Study (GRAMMS) checklist (Multimedia Appendix 1) [30].

Figure 1. Study steps and integration process.



Quantitative Phase

Setting, Participants, and Recruitment

The study took place at a total of 16 sites, including 14 community committees, 1 nursing home, and 1 community health service center in 6 districts of Beijing, China, from March 2024 to October 2024. Convenience sampling was applied during the quantitative survey phase, as it aligns with the community-level screening and voluntary nature of IC assessment under the ICOPE framework. The authors coordinated and communicated with community workers and heads of health facilities before each questionnaire survey and recruited a consecutive sample of older adults by posting group notices and conducting onsite outreach after consent was obtained. Older adults who met the following inclusion criteria were eligible to participate in the survey: (1) aged 60 years or older; (2) with normal reading, writing, and comprehension skills; and (3) voluntarily participated in this survey after providing informed consent. The exclusion criteria were as follows: (1) participants with severe functional impairments that prevented them from safely completing an exercise test, (3) participants with medically diagnosed severe cognitive impairment, (3) participants with medically diagnosed serious psychiatric disorders, and (4) participants currently experiencing an acute or worsening illness.

The sample size was estimated based on 2 primary considerations. First, the initial target sample was calculated using the Kendall model, which recommended at least 5 to 10 times the number of variables [31]. With 17 variables in this study, and considering the possibility of 20% invalid questionnaires, the required sample size was 102 to 204 cases. Second, to ensure the stability and reliability of the subsequent LPA, a sample size exceeding the commonly recommended minimum of 300 is advised [32]. A final sample of 481 older adults therefore met both the adjusted recruitment target and the methodological requirement for robust LPA.

Variables and Instruments

Based on the literature review [33,34], a structured questionnaire was designed by the research team to collect general information

XSL•FO
RenderX

as predictor variables, including sex, age, education level, marital status, family structure, type of labor force before retirement, per capita monthly family income, self-rated health using a Likert scale, and the number of chronic diseases.

The functional ability of participants, which results from the interplay between an individual's IC, supportive environment, and personal characteristics [1], was assessed using 3 established scales as key outcome measures: the Barthel Index [35], Fried Frailty Phenotype [36], and HALFT scales [37] were used to assess activities of daily living, frailty, and social frailty, respectively.

The 5 key domains of IC were assessed following the instruments recommended in the new version of the WHO ICOPE guidelines [1]. The Mini-Mental State Examination [38], Geriatric Depression Scale-15 [39], WHO vision screening chart [1], whisper voice test [1], Mini Nutritional Assessment Short-Form [40], and Short Physical Performance Battery test [41] were used to assess cognition, psychological capacity, sensory capacity (vision and hearing), vitality, and locomotor capacity, respectively. The scoring rules of these instruments are summarized in Multimedia Appendix 2.

Data Collection

Before the formal questionnaire, the investigative team, comprising postgraduate nursing students, underwent a uniform offline training on the purpose of the study, assessment procedures, questionnaire content, assessment site setup, and language and behavioral norms. For participants who had difficulty completing the questionnaire, such as those experiencing presbyopia or slow reading speeds, the investigators helped them complete it after explaining the meaning of each item. The survey was conducted using either a paper version of the questionnaire or an application installed on tablet computers, with the content being the same in both formats. The paper version of the questionnaire was entered into the application after validity was established so that the data could be accessed for statistical analysis. At the end of the questionnaire survey, the authors invited older adults who were interested in this study to participate in follow-up qualitative interviews and retained their contact details.

Data Analysis

The IC of older adults is influenced by many factors and exhibits differential characteristics [42]. Therefore, we attempted to find a people-oriented clustering approach to provide an objective population classification basis for customized personas [32]. LPA is a statistical method that identifies characteristics between individuals based on the estimation of their response patterns on observed indicators, providing unique insights into the heterogeneity of the target population [43].

First, Mplus (version 8.3, Muthén & Muthén) was used to analyze the latent profiles of the IC of older adults. Starting with 1 category, categories were added progressively until the maximum likelihood method was used to determine the best solution. The main evaluation indices included Akaike information criterion (AIC), Bayesian information criterion (BIC), adjusted BIC (aBIC), and entropy. The lower the AIC, BIC, and aBIC are, the better the model fit [44]. Entropy ranges from 0 to 1; the closer the value is to 1, the more accurate the model classification is, and it is generally believed that values of 0.80 or greater indicate that the classification accuracy of the model reaches 90% [45]. The Lo-Mendell-Rubin likelihood ratio test (LMRT) and the bootstrapped likelihood ratio test (BLRT) were used to compare the fit of 2 adjacent models, where low and significant *P* values suggest a *k*-class model is superior to *k-1*−class model [46]. To visualize the latent profiles, the estimated class-specific means from the Mplus LPA model output were exported and plotted using line charts in Microsoft Excel version 16.1.

Second, latent profiles were named and assigned values to classify groups. Descriptive statistics were analyzed using SPSS (version 26.0, IBM Corp) to understand the data distribution characteristics of the profiles. Variables that conformed to a normal distribution were described as mean (SD), and comparisons between groups were made using 1-way ANOVA. Count data were described as frequency and percentage. Differences in sociodemographic and health-related factors between groups were assessed by chi-square tests. Finally, multivariate logistic regression analyses were performed with latent profiles of IC as the dependent variable and variables that were statistically significant in the univariate analyses as the independent variables. To avoid modeling errors, dummy variables were used for categorical data. The impact of the predictor variables and their statistical significance were assessed through regression coefficients (β), *P* values, odds ratios (ORs), and 95% CIs. A *P* value <.05 indicated a statistically significant difference.

Qualitative Phase

Setting and Participants

A naturalistic, descriptive qualitative design is well suited to mixed methods research [47], enabling older adults to provide a rich and direct description of the goals, needs, and concerns of digital monitoring regarding their IC level. Stratified purposive sampling based on the categorical results of the LPA ensured the representation of participants originating from different IC profiles. Given the sequential continuity of the mixed methods study, the inclusion and exclusion criteria established for the quantitative phase were likewise applied to this phase. The first author contacted the older adults who had accepted the invitation to be interviewed during the quantitative phase, while explaining the purpose, content, and methodology of the further study and setting the time and place of the interviews.

Data Collection

Semistructured interviews were conducted between October and December 2024 to collect data. In accordance with the objectives of the study and the literature review [48], the research team developed the following interview outline: (1) "What is your idea of IC?" (2) "How would you evaluate your IC? Can you share some stories from your daily life where you have felt IC come into play or decline?" (3) "What do you think about using digital health tools for IC monitoring?" (4) "What do you hope the use of digital health tools for IC monitoring will do for you?" (5) "Do you have any concerns and doubts

XSL•FO
RenderX

about using digital health tools for IC monitoring?" and (6) "Do you have any suggestions or expectations regarding the digital monitoring of IC?"

The first author reviewed the results of the quantitative phase of the IC assessment with the participants before the commencement of the interviews. Participants were interviewed in relaxed and comfortable parks, cafes, and clinic rooms. A total of 2 additional individual interviews were conducted via voice calls on the Tencent Conference platform. Each interview, led by the first author and assisted by another author for observation and field notes, lasted 20 to 68 minutes. Audio interviews were transcribed into text within 24 hours using Lark intelligent transcription. The first 2 authors (ZX and YN) and 2 qualitative experts (QX and YW) proofread the transcript and verified its contextual integrity. Data collection and analysis were synchronized in this study. Recruitment of participants was stopped when no further significant new information emerged, implying data saturation [49].

Data Analysis

The qualitative phase of data analysis was conducted with the main aim of obtaining the needs, intentions, and challenges of older adults regarding digital monitoring originating from the different IC profiles. Inductive combined with deductive content analysis was used to hand code the data to compare and integrate the similarities and differences between the different profiles horizontally [50]. Initially, the research team repeatedly reviewed interview transcripts to gain a holistic understanding of the data. Second, after interviewing 2 participants from each latent profile, the authors conducted open coding, assigning descriptive codes to meaningful text segments to develop a preliminary coding framework for guiding the remaining transcripts. Next, data from the same latent profiles were clustered and extracted contemporaneously by the first 2 authors (ZX and YN) following the attributes in the coding framework, which enabled the creation of customized personas. The framework was continuously refined and clarified throughout this process as new insights emerged or existing categories required redefinition. The final version of the framework constituted the thematic results. To capture the group-level commonalities within each persona, the primary author noted interview logs after each interview. These logs documented the author's immediate impressions and holistic understanding of the participants, which later played a pivotal role in identifying the shared attributes that define each persona. Finally, the research team collaborated to define and describe the customized personas, outlining the participants' differing attitudes, goals, needs, and available support that ultimately formed the subthemes for the 3 customized personas. Discrepancies between the first 2 authors (ZX and YN) during data analysis were addressed by an expert in qualitative nursing research to achieve harmonization of the results.

Data Integration Procedure

As mentioned above, this study achieved methodological integration in that a quantitative LPA identified subgroups of IC that guided the qualitative phase of cluster sampling. Furthermore, the joint statistical and thematic presentation was used as a visualization tool to integrate and cross corroborate the main quantitative and qualitative findings at the level of results interpretation [51], providing a reference for future digital personalized interventions. The integrated customized personas were described jointly by the first 2 authors and iteratively discussed and revised in regular meetings of the research team, and finally confirmed by all authors.

Ethical Considerations

This study was undertaken in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Capital Medical University (approval number Z2023SY138). All participants enrolled in this study voluntarily and signed informed consent before the formal start of the assessment after being informed that they could withdraw at any time. Each older adult who participated in the quantitative survey received a gift worth approximately RMB 20 (US $2.81; a currency exchange rate of RMB 1≈US $0.1403 is applicable), while those who participated in the interviews received an additional gift worth RMB 40 (US $5.58). All data, audio recordings, and written records are stored on a password-protected laptop and backed up in the cloud; only the research team has access to them. In addition, all data were used for research purposes only and could not be disclosed without the participants' consent. To ensure privacy protection, the participants' personal information was anonymized and labeled "P1-P25."

Results

Quantitative Phase

Participant Characteristics

A total of 481 older adults, aged 60-97 years, with a mean age of 70.7 (SD 6.8) years, participated in the quantitative phase. The results in Table 1 show that most participants were female (320/481, 66.5%), aged 60-69 years (234/481, 48.6%), currently married (356/481, 74%), and living in a married couple's household (224/481, 46.6%). Most participants had a secondary education (318/481, 66.1%), a per capita monthly household income of RMB 3000-RMB 5999 (260/481, 54.1%), and engaged in light labor (241/481, 50.1%) before retirement. In addition, the majority of older adults rated their health as "very good or good" (258/481, 53.6%) and had 1 to 3 chronic diseases (323/481, 67.2%).

XSL•FO
RenderX

Table 1. Sociodemographic characteristics and univariate analysis of latent profiles (N=481).

Sociodemographic characteristics	Total, n (%)	IC[a] impairment patterns, n (%)			Chi-square (*df*)	*P* value
		Profile 1 (n=92)[b]	Profile 2 (n=140)[c]	Profile 3 (n=249)[d]		
Demographic variables						
Sex					3.0 (2)	.22
Male	161 (33.5)	29 (31.5)	55 (39.3)	77 (30.9)		
Female	320 (66.5)	63 (64.5)	85 (60.7)	172 (69.1)		
Age (years)					47.2 (4)	<.001
60-69	234 (48.6)	35 (38)	53 (37.9)	146 (58.6)		
70-79	190 (39.5)	35 (38)	59 (42.1)	96 (38.6)		
≥80	57 (11.9)	22 (23.9)	28 (20)	7 (2.8)		
Marital status					8.2 (2)	.02
Currently married	356 (74)	62 (67.4)	96 (68.6)	198 (79.5)		
Divorced, widowed, or unmarried	125 (26)	30 (32.6)	44 (31.4)	51 (20.5)		
Family structures					6.2 (4)	.19
Lives alone	79 (16.4)	20 (21.7)	27 (19.3)	32 (12.9)		
A married couple’s household	224 (46.6)	44 (47.8)	60 (42.9)	120 (48.2)		
Multigeneration household	178 (37)	28 (30.4)	53 (37.8)	97(39)		
Socioeconomic variables						
Education level					65.4 (6)	<.001
Primary and below	85 (17.7)	33 (35.9)	38 (27.1)	14 (5.6)		
Junior high school	163 (33.9)	30 (32.6)	49 (35)	84 (33.7)		
Senior high school	155 (32.2)	23 (25)	29 (20.7)	103 (41.4)		
College and above	78 (16.2)	6 (6.5)	24 (17.1)	48 (19.3)		
Type of preretirement labor force					20.6 (4)	<.001
Mental labor	149 (31)	22 (23.9)	45 (32.1)	82 (32.9)		
Light labor	241 (50.1)	39 (42.4)	66 (47.1)	136 (54.6)		
Heavy labor	91 (18.9)	31 (33.7)	29 (20.7)	31 (12.4)		
Monthly per capita household income (RMB)[e]					52.6 (6)	<.001
<1000	56 (11.6)	24 (26.1)	22 (15.7)	10 (4)		
1000-2999	85 (17.7)	21 (22.8)	29 (20.7)	35 (14.1)		
3000-5999	260 (54.1)	41 (44.6)	59 (42.1)	160 (64.3)		
≥6000	80 (16.6)	6 (6.5)	30 (21.4)	44 (17.7)		
Health-related variables						
Self-rated health					75.2 (8)	<.001
Very good	61 (12.7)	4 (4.3)	11 (7.9)	46 (18.5)		
Good	197 (41)	18 (19.6)	53 (37.9)	126 (50.6)		
General	169 (35.1)	45 (48.9)	58 (41.4)	66 (26.5)		
Poor	40 (8.3)	17 (18.5)	14 (10)	9 (3.6)		
Very poor	14 (2.9)	8 (8.7)	4 (2.9)	2 (0.8)		
Number of chronic diseases					25.8 (4)	<.001
0	99 (20.6)	10 (10.9)	22 (15.7)	67 (26.9)		
1-3	323 (67.2)	60 (65.2)	100 (71.4)	163 (65.5)		

XSL•FO
RenderX

Sociodemographic characteristics	Total, n (%)	IC[a] impairment patterns, n (%)			Chi-square (*df*)	*P* value
		Profile 1 (n=92)[b]	Profile 2 (n=140)[c]	Profile 3 (n=249)[d]		
≥4	59 (12.3)	22 (23.9)	18 (12.9)	19 (7.6)		

[a]IC: intrinsic capacity.

[b]Percentage of the total=19.1 (92/481).

[c]Percentage of the total=29.1 (140/481).

[d]Percentage of the total=51.8 (249/481).

[e]A currency exchange rate of RMB 1≈US $0.1403 is applicable.

LPA Results

The scores in the 5 domains of IC were used as exogenous variables to model the latent profiles. The *z* score was used to eliminate the effects of different magnitudes on data analysis during data preprocessing. The fitting statistics for modeling 1-5 latent profiles sequentially are shown in Table 2. After comprehensive consideration of the latent profile metrics and clinical significance, we identified the 3-profile model as the best option for the following reasons: (1) the values of AIC, BIC, and aBIC decreased gradually as the number of profile subgroups increased, but beginning with the 4-profile model, the LMRT was greater than .05 and the minimum classification probability was less than 10%, signifying that this category may be spurious; (2) the LMRT and BLRT results were statistically significant (P<.05) for the 3-profile model, indicating that it was better than the 2-profile solution; and (3) the entropy value of the 3-profile model exceeded 0.8, indicating good model accuracy. Therefore, the 3-profile model was selected as the most appropriate model. To verify the reliability of the LPA, we calculated the average posterior probabilities for the 3 profiles. The results are shown in Multimedia Appendix 3, where the posterior probabilities were greater than 0.8, indicating that the model results for the 3 profiles were well separated [52].

Table 2. Latent profile analysis model fit of intrinsic capacity impairment patterns.

Model	Log(L)[a]	AIC[b]	BIC[c]	aBIC[d]	Entropy	LMRT[e]	BLRT[f]	Classification probability
1	–3409.625	6839.251	6881.010	6849.271	—[g]	—	—	1.00
2	–3085.278	6202.556	6269.370	6218.588	1.000	.2054	<.001	0.291/0.709
3	–2988.495	6020.991	6112.860	6043.034	0.906	.0174	<.001	0.191/0.291/0.518
4	–2946.724	5949.448	6066.372	5977.503	0.918	.1141	<.001	0.089/0.102/0.291/0.518
5	–2928.674	5925.348	6067.327	5959.414	0.920	.1226	<.001	0.046/0.080/0.080/0.291/0.503

[a]Log(L): Log-likelihood.

[b]AIC: Akaike information criterion.

[c]BIC: Bayesian information criterion.

[d]aBIC: adjusted BIC.

[e]LMRT: Lo-Mendell-Rubin likelihood ratio test.

[f]BLRT: bootstrapped likelihood ratio test.

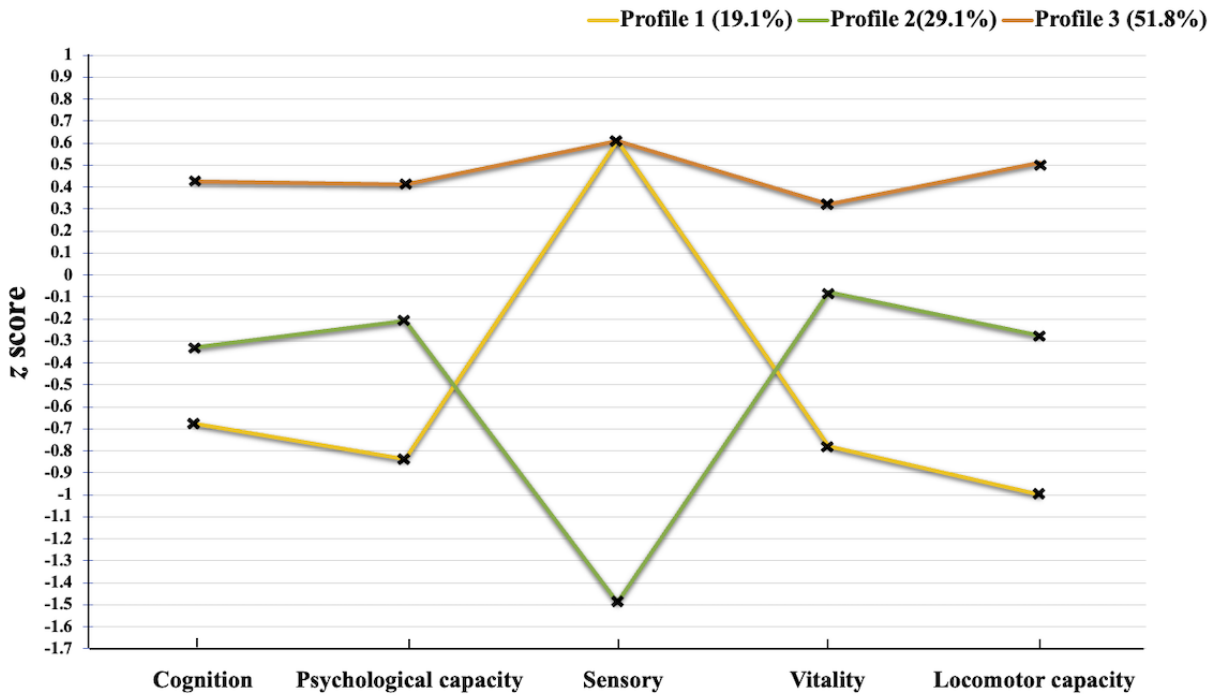
[g]Not applicable.

Characteristics and Naming of Latent Profiles

Figure 2 illustrates a graphical presentation of the *z* score for the 5 domains of IC in the 3 impairment profiles. Each profile was named according to its characteristics. Profile 1, constituting 19.1% (92/481) of the sample, was designated the "multisubdomain decline–IC imbalance group." The most obvious characteristic of this group was the overall decline in IC, while only the sensory domain remained at a high level, resulting in an imbalance in IC. It should be noted that higher scores on the psychological domain indicated reduced psychological capacity. Profile 2, the "multisubdomain moderate–sensory deficit group," included 29.1% (140/481) of the participants. This group showed a moderate IC level that was higher than those in profile 1 and was characterized by significant sensory deficits, reflecting a pattern of decline with single deficits. Profile 3 was the most prevalent in the sample, accounting for 51.8% (249/481) of cases. It was designated the "multisubdomain robust–whole balance group" owing to its exceptional performance in IC, with no significant loss of ability.

XSL•FO
RenderX

Figure 2. Characteristic distribution of 3 latent profiles. Profile 1: multisubdomain decline–IC imbalance group; profile 2: multisubdomain moderate–sensory deficit group; and profile 3: multisubdomain robust–whole balance group.



Comparison of Predictive Variables in Each Latent Profile

As shown in Table 1, the findings derived from the univariate analyses provide insights into the differences among the 3 latent profiles by using sociodemographic and health-related factors as predictor variables. Among these variables, age, marital status, education level, type of labor force prior to retirement, per capita monthly household income, self-rated health, and number of chronic diseases differed significantly (P<.05) among the 3 IC impairment patterns. No statistically significant differences were observed in sex and family structure.

Only variables that were statistically different between the IC subgroups in the univariate analyses were included in the multivariate logistic regression analyses (Table 3). Using profile 3, the "multisubdomain robust–whole balanced group," as the reference group, older adults aged 80 years and older are more frequently found in either profile 1 (OR 9.190, 95% CI 2.986-28.284; P<.001) or profile 2 (OR 7.018, 95% CI 2.559-19.247, P<.001). As detailed in Table 3, as education levels rose, older adults were less likely to be categorized in profile 1 or profile 2. Those with a per capita monthly household income of ≥RMB 6000, indicating a higher quality of life and social status, were less likely to be attributed to profile 1 (OR 0.163, 95% CI 0.040-0.670; P<.05). Those with better self-rated health were more likely to belong to profile 3. For example, the probability of older adults rating their self-health as "very good" joining profile 1 (OR 0.037, 95% CI 0.005-0.265; P<.001) and profile 2 (OR 0.115, 95% CI 0.017-0.773; P<.05) was reduced. Older adults with no chronic diseases were less likely to be enrolled in profile 1 (OR 0.315, 95% CI 0.105-0.943; P<.05). The results, when viewed through profile 2 as the reference group, indicated a lower possibility of older adults with a per capita monthly household income ≥RMB 6000 (OR 0.219, 95% CI 0.061-0.782; P<.05) and self-rated health of "good" (OR 0.236, 95% CI 0.060-0.935; P<.05) belonging to profile 1.

XSL•FO
RenderX

Table 3. Multivariate logistic regression analysis of predictor variables of 3 latent profiles (N=481).

Predictor variables	Profile 1[a] vs profile 3[b]		Profile 2[c] vs profile 3		Profile 1 vs profile 2	
	β (SE)	OR[d] (95% CI)	β (SE)	OR (95% CI)	β (SE)	OR (95% CI)
Demographic variables						
Age (years)						
60-69 (ref[e])	N/A[f]	N/A	N/A	N/A	N/A	N/A
70-79	–0.079 (0.350)	0.924 (0.465-1.833)	0.119 (0.272)	1.126 (0.661-1.919)	–1.198 (0.353)	0.820 (0.410-1.639)
≥80	2.218 (0.574)[g]	9.190 (2.986-28.284)	1.949 (0.515)[g]	7.018 (2.559-19.247)	0.270 (0.445)	1.309 (0.548-3.130)
Marital status						
Currently married (ref)	N/A	N/A	N/A	N/A	N/A	N/A
Divorced, widowed, or unmarried	–0.076 (0.352)	0.927 (0.465-1.847)	–0.162 (0.289)	0.850 (0.482-1.498)	0.086 (0.341)	1.090 (0.558-2.128)
Socioeconomic variables						
Education level						
Primary and below (ref)	N/A	N/A	N/A	N/A	N/A	N/A
Junior high school	–1.004 (0.492)[h]	0.366 (0.140-0.961)	–1.041 (0.429)[h]	0.353 (0.152-0.819)	0.036 (0.422)	1.037 (0.454-2.371)
Senior high school	–1.484 (0.549)[i]	0.227 (0.077-0.665)	–1.899 (0.481)[g]	0.150 (0.058-0.384)	0.415 (0.510)	1.514 (0.557-4.117)
College and above	–1.755 (0.713)[h]	0.173 (0.043-0.699)	–1.241 (0.536)[h]	0.289 (0.101-0.827)	–0.514 (0.679)	0.598 (0.158-2.263)
Type of preretirement labor force						
Mental labor (ref)	N/A	N/A	N/A	N/A	N/A	N/A
Light labor	–0.441 (0.392)	0.644 (0.299-1.387)	–0.295 (0.305)	0.744 (0.409-1.353)	–0.145 (0.406)	0.865 (0.391-1.915)
Heavy labor	0.218 (0.491)	1.244 (0.475-3.260)	–0.238 (0.420)	0.788 (0.346-1.797)	0.465 (0.483)	1.578 (0.612-4.071)
Monthly per capita household income (RMB) [j]						
<1000 (ref)	N/A	N/A	N/A	N/A	N/A	N/A
1000-2999	–0.558 (0.567)	0.572 (0.189-1.737)	–0.151 (0.529)	0.860 (0.305-2.424)	–0.407 (0.463)	0.665 (0.268-1.650)
3000-5999	–1.056 (0.538)	0.348 (0.121-0.999)	–0.780 (0.503)	0.458 (0.171-1.230)	–0.276 (0.448)	0.759 (0.315-1.825)
≥6000	–1.813 (0.721)[h]	0.163 (0.040-0.670)	–0.295 (0.570)	0.744 (0.243-2.275)	–1.518 (0.649)[h]	0.219 (0.061-0.782)
Health-related variables						
Self-rated health						
Very good	–3.290 (1.000)[i]	0.037 (0.005-0.265)	–2.166 (0.974)[h]	0.115 (0.017-0.773)	–1.124 (0.882)	0.325 (0.058-1.833)
Good	–3.408 (0.886)[i]	0.033 (0.006-0.188)	–1.964 (0.923)[h]	0.140 (0.023-0.858)	–1.444 (0.703)[h]	0.236 (0.060-0.935)
General	–2.065 (0.871)[h]	0.127 (0.023-0.699)	–1.304 (0.924)	0.271 (0.044-1.661)	–0.761 (0.682)	0.467 (0.123-1.778)
Poor	–0.665 (0.952)	0.514 (0.080-3.323)	–.0.514 (1.004)	0.598 (0.084-4.278)	–0.151 (0.751)	0.860 (0.197-3.748)

XSL•FO
RenderX

Predictor variables	Profile 1[a] vs profile 3[b]		Profile 2[c] vs profile 3		Profile 1 vs profile 2	
	β (SE)	OR[d] (95% CI)	β (SE)	OR (95% CI)	β (SE)	OR (95% CI)
Very poor (ref)	N/A	N/A	N/A	N/A	N/A	N/A
Number of chronic diseases						
0	–1.155 (0.559)[h]	0.315 (0.105-0.943)	–0.413 (0.477)	0.661 (0.260-1.684)	–0.741 (0.539)	0.477 (0.166-1.370)
1-3	–0.357 (0.430)	0.700 (0.301-1.626)	0.195 (0.411)	1.215 (0.543-2.718)	–0.552 (0.391)	0.576 (0.268-1.239)
≥4 (ref)	N/A	N/A	N/A	N/A	N/A	N/A

[a]Profile 1: multisubdomain decline–IC imbalance group.

[b]Profile 3: multisubdomain robust–whole balance group.

[c]Profile 2: multisubdomain moderate–sensory deficit group.

[d]OR: odds ratio.

[e]ref: reference category

[f]N/A: not applicable.

[g]*P*<.001.

[h]*P*<.05.

[i]*P*<.01.

[j]A currency exchange rate of RMB 1≈US $0.1403 is applicable.

Outcome Variables of Latent Profiles

As demonstrated in Table 4, there are significant differences in activities of daily living, frailty, and social frailty among the identified latent profiles of IC impairment. Subsequent post hoc comparisons revealed a gradual decline in functioning for latent profiles 3, 2, and 1, using the 3 functional abilities as the outcome variables.

Table 4. Effect of the 3 latent profiles on outcome variables (N=481).

Outcome variables (score range)	IC[a] impairment patterns, mean (SD)			Post hoc comparison	*F* test (*df*)	*P* value
	Profile 1[b] (n=92)	Profile 2[c] (n=140)	Profile 3[d] (n=249)			
BI[e] (0-100)	88.70 (1.801)	93.32 (1.239)	99.20 (0.168)	3>2>1	33.866 (2, 478)	<.001
FFP[f] (0-5)	1.82 (0.105)	1.35 (0.094)	0.49 (0.043)	3>2>1	90.299 (2, 478)	<.001
HALFT scale (0-5)	1.71 (0.126)	1.19 (0.103)	0.51 (0.049)	3>2>1	52.867 (2, 478)	<.001

[a]IC: intrinsic capacity.

[b]Profile 1: multisubdomain decline–IC imbalance group.

[c]Profile 2: multisubdomain moderate–sensory deficit group.

[d]Profile 3: multisubdomain robust–whole balance group.

[e]BI: Barthel index.

[f]FFP: Fried frailty phenotype.

Qualitative Phase

Interview Sample Characteristics

A total of 25 older adults volunteered to participate in the interviews, including 5 in profile 1, 10 in profile 2, and 10 in profile 3. Most participants were female (17/25, 68%). All participants had a mean age of 70.1 (SD 7.03, range 60-85) years, and had experienced between 0 and 4 IC domain declines. Detailed information is provided in Multimedia Appendix 4.

Developing Customized Personas

A total of 4 themes emerged in this study. In the hierarchical data analysis, 3 customized persons corresponding to different IC profiles were depicted, and subthemes under the themes that matched the persons were illustrated in Figures 3-5. Representative quotations were translated into English by the authors and confirmed by back-translation. In this section, the findings are reported as an integration of customized personas as follows.

XSL•FO
RenderX

Figure 3. Persona A: affects my mood—anxious evader.

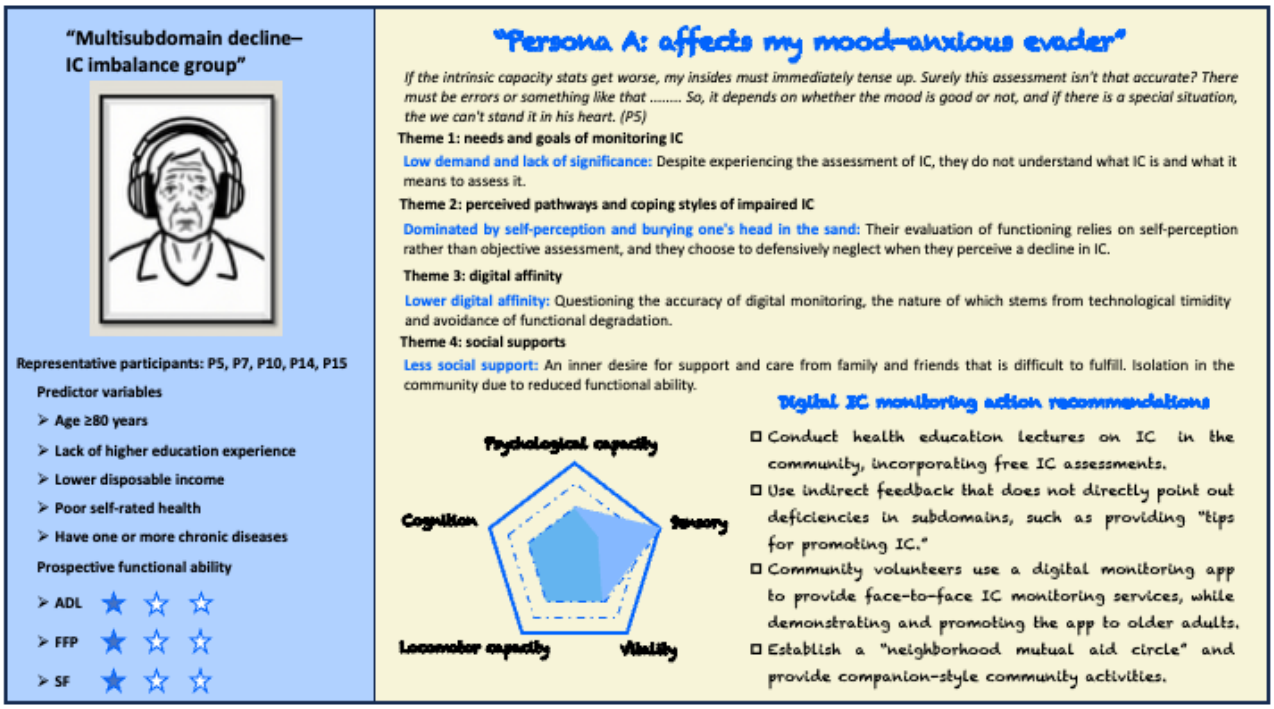
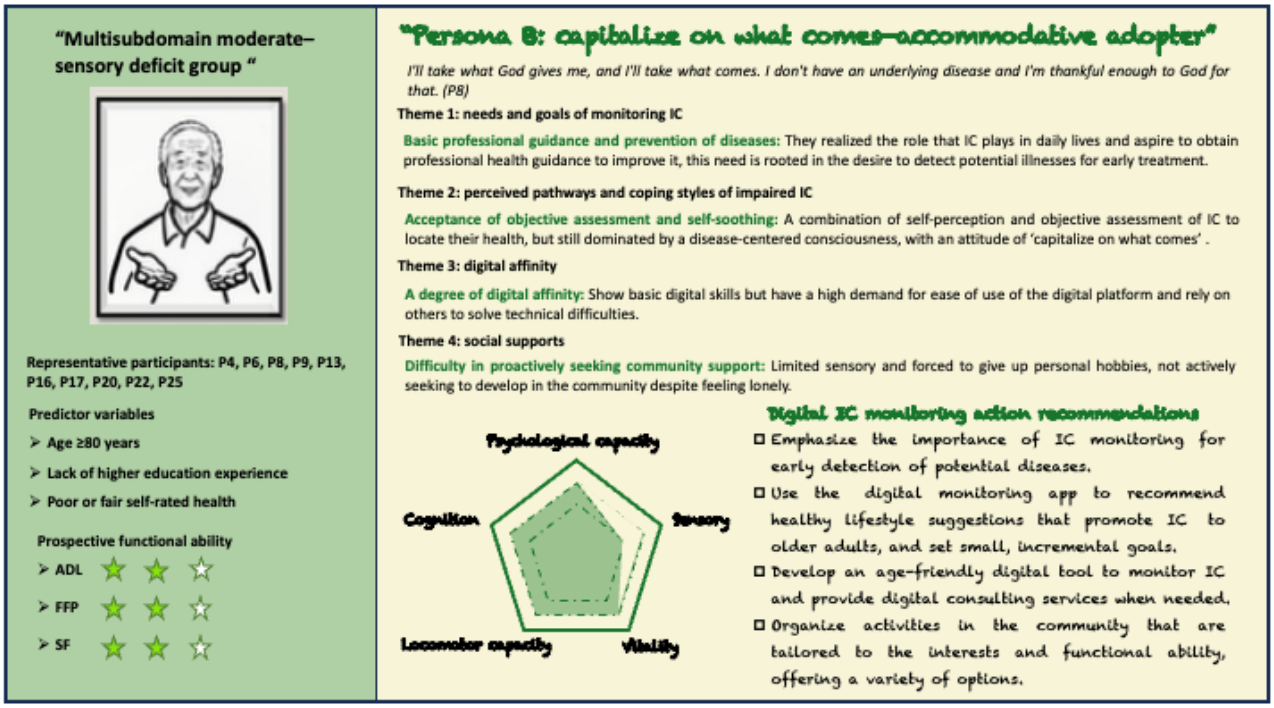
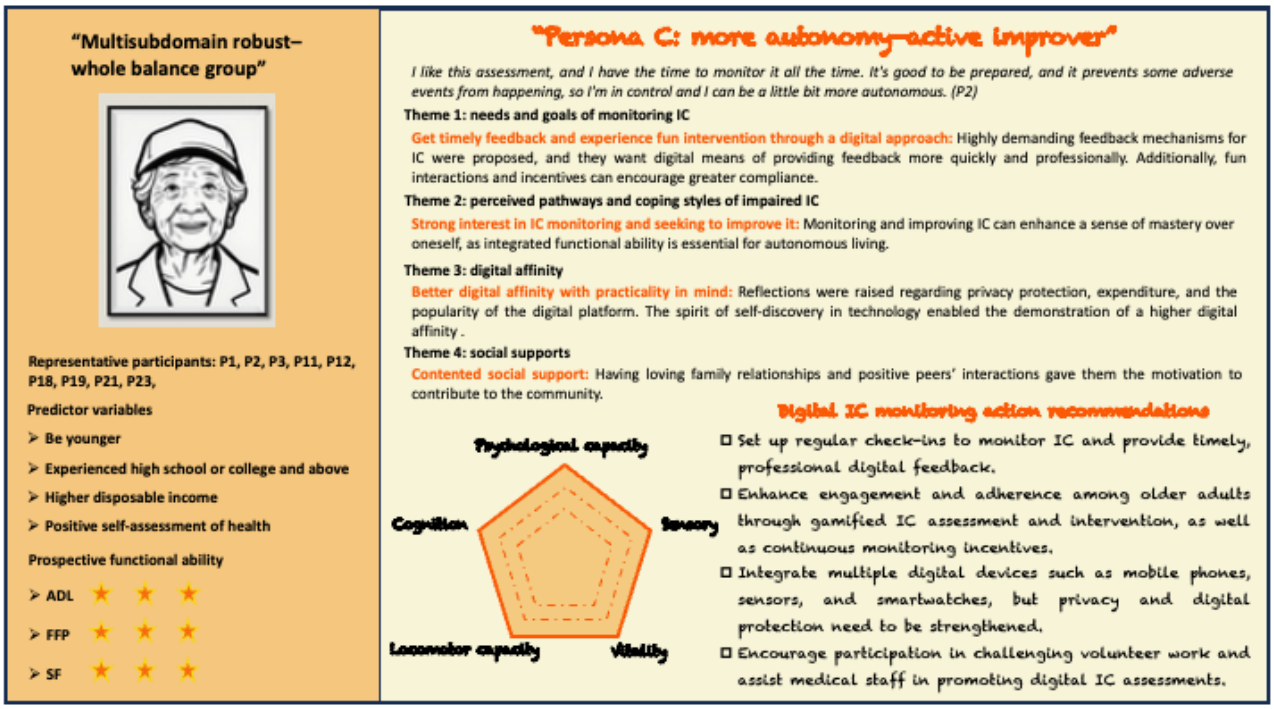


Figure 4. Persona B: capitalize on what comes—accommodative adopter.



XSL•FO
RenderX

Figure 5. Persona C: more autonomy—active improver.



Persona A: Affects My Mood—Anxious Evader

Narrative persona: affects my mood—anxious evader are female and belong to profile 1. They are older than those in other profiles, have lower education and disposable income, rate their health negatively, and are more likely to be chronically ill. Insights from the interviews suggest that they are psychologically fragile and stay away from anything that may bring about emotional stress. Moreover, the lack of social support for persona A, as evidenced by inadequate family care, a desire for companionship, and a neglected position in social groups, may further diminish their physical and psychological capacity in the face of potential stress. In addition, they were skeptical about using digital technology to assess IC and reluctant to ask for help based on past experiences. As a result, persona A showed less interest in digital monitoring of IC, which may have been a psychological coping mechanism that prompted the adoption of an avoidance strategy.

> *It certainly can't improve, this time this way, next time do better than this time, I don't think it can be done, people can only get older and worse.* [P15]
>
> *I know all about it, the results of this assessment are pretty much the same as how I feel about myself, the assessment is not very useful, and I'm doing enough to build muscle at home.* [P10]
>
> *I can type, but it might be difficult for other older adults. If I ask my child to help, she always tells me to take a break and not mess around. So, um..., but I don't want people to think I'm stupid if you ask for help, I'd be embarrassed and stressed.* [P14]
>
> *I used to be the lead dancer of the dance team, I thought I could still teach people how to dance when I got better, but I didn't realize that now I'm the most insignificant person in the dance team, and I can't dance .....I wish someone would tell me how to face society.* [P7]

Persona B: Capitalize on What Comes—Accommodative Adopter

Narrative persona: older adults in persona B, derived from profile 2, have a slight decline in IC, characterized by an impairment in sensory function. Compared to persona C, older adults in persona B were older, less educated, and more likely to rate their health as neutral or slightly negative. Participants expressed a desire to receive assistance from IC in the identification of potential illnesses and the acquisition of professional guidance, which was a passive strategy for treating illnesses. The acceptable attitude they displayed was derived from the Chinese proverb "capitalize on what comes," which implied that since illness is unavoidable, it should be accepted calmly. Seemingly, this was not consistent with the Healthy Aging Framework, which posited the maintenance of IC to promote functional ability and quality of life. Moreover, they were willing to take advantage of digital health technology to assess IC but were marginally deficient in digital literacy, resulting in greater demands for functionality and ease of use. Regarding social support, they reported feeling cared for by family and friends. However, the activities and hobbies that they were compelled to abandon due to declining functionality also engendered helplessness and isolation, echoing the findings that they may be in the prefrailty and presocial frailty stages.

> *I'd like to know, what I should eat, and what I should watch out for to make me cognitively aware if I've lost my ability in a certain area.* [P8]
>
> *The disease is inevitable, no one will not be sick, old age is sicker, this is also in the natural, normal. Of course, when things go wrong, maybe it's just a minor problem, who doesn't do that in life? I can still sing,*

XSL•FO
RenderX

my head is not confused, and I don't have any trouble recognizing people, so I'm very satisfied with myself. [P6]

In any case, it should be simplified. Technology is something that makes things simpler, it is 1 + 1 = 2, but why can't some older adults accept it? That's why I think it should be simplified; too many functions don't match the energy of the older adults. [P13]

I have to give up a lot of activities and interests because of declining physical abilities.……The last six months I've been feeling really lonely and unenergetic. When I get up early, I go out and sit for two hours, and then the afternoon is particularly torturous. [P16]

Persona C: More Autonomy—Active Improver

Narrative persona: participants in persona C are derived from profile 3 and represent older adults with high levels of IC. Compared to personas A and B, they are likely to be younger, more educated, have better disposable income, and a more positive assessment of their health. They showed a high level of interest in monitoring IC and were keen to maintain communication with their community health care team through age-friendly digital platforms for regular reminders, feedback on professional advice, and facilitation of functional improvement. Older adults in the technology-rich exploratory persona C also needed fun interventions and motivational approaches to help improve monitoring adherence, confidence, and sense of achievement. They suggested that free and community outreach is essential to increase access to IC monitoring. In addition, the interviews revealed that they were more likely to live in well-functioning family relationships and maintained good relationships with neighbors and peers through developing shared hobbies and were more autonomous and socially active.

After all, it needs to be assessed regularly, so could you design some mini-games? Design mini-games that exercise cognitive and finger dexterity and are similar to completing tasks. Older adults enjoy playing games just as much as children do. [P12]

I think vitality, cognition, psychology and so on make up a state of equilibrium, and good IC is the most appropriate balance. For example, sometimes we may encounter some things, and our mental ability may be reduced, but we can regulate and restore it through some methods. [P23]

Some of the digital stuff reflected ..... Actually, to the point where we use it quite simple. I have no difficulty at all, may have something to do with my previous work exposure to digital technology earlier. [P3]

I often participate in social activities and volunteer, and I want to do what I can. I have been instructing residents on rubbish sorting and participating in community work, which I am willing to do as long as I have enough time. [P11]

Discussion

Principal Findings

This study used an explanatory sequential mixed methods design to elucidate the data-driven link between heterogeneous IC profiles and psychological intentions toward digital IC monitoring in older adults. Specifically, an explanatory sequential mixed methods design was used, in which quantitative LPA provided the sampling framework and identified group-centered persona predictors. This informed subsequent qualitative interviews, which uncovered the underlying psychosocial narratives and specific digital monitoring intentions unique to each persona. “Persona A: affects my mood—anxious evader,” “persona B: capitalize on what comes—accommodative adopter,” and “persona C: more autonomy—active improver” represent differentiated IC characteristics among older adults and their digital monitoring intentions. These findings provide person-centered typology that can inform the development of tailored digital monitoring strategies within community-based integrated care for older adults.

While digital health tools offer efficient ways to monitor IC, it is vital that implementation strategies remain cautious and respect older adults’ willingness for digital engagement [29]. This caution is particularly pertinent to the “persona A: affects my mood—anxious evader” in this study, who was revealed to have the lowest level of IC and a negative willingness to engage in digital monitoring. They exhibited a higher prevalence of individuals aged 80 years or older, with low levels of education, low income, and poor self-rated health compared to the other 2 personas, echoing results reported in other studies showing a clear correlation between sociodemographic determinants and significant decline in IC [34,53]. A scoping review suggested that the 3 key factors contributing to digital exclusion among older adults are sociodemographic, physiological, and psychological factors, resulting in “unable” and “unwilling” attitudes [17], which further validate the trait of technological timidity in persona A.

The Health Belief Model can also provide an indirect explanation for this phenomenon, whereby older adults recognize that a decline in IC poses risks to their functional capacity and quality of life, yet the perceived benefits of digital IC monitoring are overshadowed by perceived barriers, including technological complexity, technophobia, and threats to self-esteem. As far as psychological states were concerned, the interview findings revealed that older adults within persona A exhibit a strong sense of self-protection, which is manifested in their avoidance of confronting their actual IC level. They expressed anxiety about the implication that an IC decline reflects age-related functional ability lapse, a condition that signifies aging. This phenomenon may be partially explained by Guo et al [54], who claimed that self-neglecting older adults are more exposed to delayed health actions and mental health damage and have a multisubdomain impairment of IC. Furthermore, social participation and family relationships are potential ameliorative factors in the decline of IC [53]. Persona A yearns for familial care and community engagement, though concerns about self-esteem and public image hinder proactive

XSL•FO
RenderX

action, resulting in unmet social needs such as reminders and companionship [55].

Therefore, intervention strategies for Persona A should still be initiated with traditional IC monitoring to build trust via stable, low-pressure collaborative relationships. Community care providers should conduct comprehensive assessments of older adults' daily living conditions, home environments, and social networks to design tailored supportive measures, which may involve gradual exposure to digital IC assessment tools and the provision of accessible social activities.

Persona B are characterized as "capitalize on what comes—accommodative adopter," who have made peace with the decline in IC and demonstrate a more pragmatic, disease-centered demand for digital monitoring. They believed that the IC could provide early warning of hidden diseases and guide them to seek timely medical treatment by obtaining professional advice from the community health care team. ICOPE, focused on monitoring IC, is a key practice in the shift from disease-centered to function-centered approaches to healthy aging [56]. It is an innovative model that not only facilitates the transformation of community health care models but also urges older adults to overcome the cognitive limitations of a disease-focused mindset.

However, older adults in persona B are subconsciously constrained by self-agism, whereby they perceive the decline in IC as inevitable, which may reduce their perceived benefits of digital IC monitoring and lead to passive technological application [57]. A further characteristic is a notable sensory deficit that may result in diminished self-efficacy during digital interactions. As one grows older, the sensory domain is where the prevalence of impaired IC is most evident [21]. A previous study confirmed that augmented digital access could help address unmet medical needs and increase use of preventive care [58]. As participants desired, digital platforms integrated with wearable devices for real-time monitoring, capturing metrics such as step speed and gait, may mitigate interaction challenges arising from visual impairment. Besides, a flat structure and clear interface design of the digital platform can reduce technological dependence and enhance adherence to long-term IC monitoring.

Therefore, while monitoring IC through digital platforms, it is also essential to recommend and track lifestyle factors such as nutrition, physical activity, and social engagement for persona B. These recommendations should be tailored rather than generic health advice to alleviate their concerns about disease risks and functional decline. By linking behaviors that support IC to the motivation of disease detection, digital monitoring of IC becomes more proactive.

Persona C is an active collaborator in the community integrated care team's use of digital IC monitoring; hence, it is designated as "more autonomy—active improver." They were characterized by relatively younger age, higher educational attainment, disposable income, and more positive self-assessments of health, in line with previous studies that found that older adults with these characteristics had better IC and lower risk of decline [59,60]. Older adults in this persona were particularly proactive in mastering their IC to seek autonomy in later life. They believed that maintaining self-management abilities is essential for older adults' health and independence, especially in the face of IC decline, and that it also moderates the relationship between quality of life and physical function [33]. Moreover, persona C showed higher digital literacy, as evidenced by their constructive comments on the needs for digital platforms, such as timely and professional feedback and interesting incentives and interventions to stimulate adherence and engagement in health management. As confirmed by prior practical experience studies, timely feedback mechanisms and support networks enhanced interactions between older adults and community integrated care teams, providing an efficient interface for subsequent steps in ICOPE [14,15]. Additionally, another characteristic of this persona was that they fully perceived and used social support and further contributed to society to the best of their ability. Similar findings indicated that when older adults had a greater sense of control, they were able to actively use perceived social support to achieve goals they deemed important [61]. On the other hand, these older adults unintentionally accessed more resources or strategies while participating in social activities, creating a positive feedback loop for the improvement of IC [62].

Consequently, digital monitoring of persona C's IC should incorporate motivational elements to encourage long-term engagement, thereby reinforcing a sense of competence and autonomy. To sustain engagement, the monitoring platform could incorporate gamification features such as progress tracking, challenge missions, and reward mechanisms for achieving IC-related goals, which could directly enhance a sense of accomplishment and confidence in self-management, aligning with their proactive approach to healthy aging.

Strengths and Limitations

The primary contribution of this study is the application of a data-driven persona framework to elucidate older adults' psychological intentions regarding digital IC monitoring. These personas synthesize quantitative predictors of IC profiles with qualitative insights into attitudes, goals, ability reserves, and support resources. Methodologically, the initial LPA established a sampling and analytical framework ensuring that interviews focused on revealing rich contextual narratives behind each statistically defined profile. This approach advances the discourse from describing subgroup characteristics to explaining psychological intentions, constructing personas that are both statistically grounded and vividly descriptive. Additionally, customized personas will inform targeted recruitment strategies and the development of IC monitoring systems oriented toward real abilities and goals, ensuring monitoring pathways align with the readiness and needs of diverse groups.

However, this study has several limitations. First, the participants in this study were recruited from urban communities in Beijing, China, and were selected through convenience sampling. Therefore, the findings may not be generalizable to older adults in rural or remote settings, where factors such as health care accessibility, digital literacy, socioeconomic status, and community support structures may present distinct profiles that could influence both IC trajectories and engagement with digital monitoring. Second, sex differences in the value of health

XSL•FO
RenderX

and social engagement resulted in a majority of participants being female, which has an unknown effect on depicting persona characteristics. Third, the interview sample based on the results of the quantitative phase has differences in the number of participants, which can be explained by the poorer health and lower interest of older adults in profile 1. The final limitation is that the 3 customized personas provide only a momentary sketch of a specific population. To keep the personas up to date, future studies will address this by collecting additional data through follow-up visits.

Conclusion

This study integrates quantitative feature analysis with qualitative narrative exploration to investigate the heterogeneous IC characteristics of older adults and their psychological intentions toward digital IC monitoring. Through a comprehensive review, 3 distinct IC profiles were identified, and corresponding customized personas were developed. These personas not only reflect the multidimensional nature of IC but also capture underlying attitudes, personal goals, digital affinity, and social support. The findings provide practical, persona-based tools for community health care providers to tailor communication, support, and technology adoption strategies, enhancing the acceptance and effectiveness of digital IC monitoring among older adults.

Acknowledgments

YW and HC made equal contributions to the manuscript and are considered co-corresponding authors. The time, sincerity, and wisdom of the participants in this study are gratefully acknowledged. Sincere gratitude is extended to the community workers and community health care workers who assisted in recruitment for this study. All authors declare that generative AI was used only for language polishing of the manuscript and was not used in any part of the study.

Funding

This work was supported by the National Key Research and Development Program of China (grant numbers 2023YFC3605200 and 2023YFC3605201). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due privacy or ethical restrictions, but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Good Reporting of a Mixed Methods Study (GRAMMS) checklist.
[DOCX File , 18 KB - aging_v9i1e82867_app1.docx]

Multimedia Appendix 2
Instruments used to assess intrinsic capacity and functional ability.
[DOCX File , 26 KB - aging_v9i1e82867_app2.docx]

Multimedia Appendix 3
Mean attribution probabilities for each latent profile.
[DOCX File , 17 KB - aging_v9i1e82867_app3.docx]

Multimedia Appendix 4
Sociodemographic characteristics of interview participants (n=25).
[DOCX File , 19 KB - aging_v9i1e82867_app4.docx]

References

1. Integrated Care for Older People (ICOPE): Guidance for Person-Centred Assessment and Pathways in Primary Care. 2nd Edition. Geneva: World Health Organization; 2024.
2. Goh SHE, Zhang D, Tan KH, Koh SLS. Nurses' perception of their role in leading nurse-led interventions in intrinsic capacity assessment to improve nursing care of older adults. J Adv Nurs 2025;81(8):5066-5079. [doi: 10.1111/jan.16729] [Medline: 39749867]

XSL•FO
RenderX

3. Ahmadi S, Afshar PF, Malakouti K, Azadbakht M. The relationship between intrinsic capacity and functional ability in older adults. BMC Geriatr 2025;25(1):57 [FREE Full text] [doi: 10.1186/s12877-025-05709-y] [Medline: 39863832]
4. Beard JR, Jotheeswaran AT, Cesari M, Araujo de Carvalho I. The structure and predictive value of intrinsic capacity in a longitudinal study of ageing. BMJ Open 2019;9(11):e026119 [FREE Full text] [doi: 10.1136/bmjopen-2018-026119] [Medline: 31678933]
5. Integrated Care for Older People (ICOPE): Guidance for Person-Centred Assessment and Pathways in Primary Care. Geneva: World Health Organization; 2019.
6. Liu M, Zhang M, Zhou J, Song N, Zhang L. Research on the healthy life expectancy of older adult individuals in China based on intrinsic capacity health standards and social stratification analysis. Front Public Health 2023;11:1303467 [FREE Full text] [doi: 10.3389/fpubh.2023.1303467] [Medline: 38356656]
7. Cao X, Yi X, Chen H, Tian Y, Li S, Zhou J. Prevalence of intrinsic capacity decline among community-dwelling older adults: a systematic review and meta-analysis. Aging Clin Exp Res 2024;36(1):157. [doi: 10.1007/s40520-024-02816-5] [Medline: 39088112]
8. Ma L, Chhetri JK, Zhang Y, Liu P, Chen Y, Li Y. Integrated care for older people screening tool for measuring intrinsic capacity: preliminary findings from ICOPE pilot in China. Front Med (Lausanne) 2020;7:576079 [FREE Full text] [doi: 10.3389/fmed.2020.576079] [Medline: 33330532]
9. Zhou Y, Wang G, Li J, Liu P, Pan Y, Li Y, et al. Trajectory of intrinsic capacity among community-dwelling older adults in China: the China Health and Retirement Longitudinal Study. Arch Gerontol Geriatr 2024;124:105452. [doi: 10.1016/j.archger.2024.105452] [Medline: 38728820]
10. Sánchez-Sánchez JL, Lu W, Gallardo-Gómez D, Del Pozo Cruz B, de Souto Barreto P, Lucia A. Association of intrinsic capacity with functional decline and mortality in older adults: a systematic review and meta-analysis of longitudinal studies. Lancet Healthy Longev 2024;5(7):e480-e492 [FREE Full text] [doi: 10.1016/S2666-7568(24)00092-8] [Medline: 38945130]
11. Hwang AC, Chen LY, Tseng SH, Huang CY, Yen KH, Chen LK. Intrinsic capacity transitions predict overall and cause-specific mortality, incident disability, and healthcare utilization. J Nutr Health Aging 2024;28(10):100359 [FREE Full text] [doi: 10.1016/j.jnha.2024.100359] [Medline: 39277970]
12. Ma L. Physical activity, sedentary behaviour, and intrinsic capacity at older ages: get active!. Lancet Healthy Longev 2025;6(5):100687 [FREE Full text] [doi: 10.1016/j.lanhl.2025.100687] [Medline: 40414231]
13. Berbon C, Rolland Y, Takeda C, Lafont C, Tavassoli N, De Kerimel J. WHO ICOPE programme adherence of 8672 older age people over 2-years of follow-up. J Adv Nurs 2025;81(9):5925-5936. [doi: 10.1111/jan.16740] [Medline: 40195235]
14. Berbon C, Takeda C, Balardy L, Lafont C, Tavassoli N, Carrie I. Implementing the WHO ICOPE program in clinical practice: three years of lessons from monitoring 27,082 participants using the ICOPE monitor digital tool. J Gerontol A Biol Sci Med Sci 2025;80(3):glae278. [doi: 10.1093/gerona/glae278] [Medline: 39549299]
15. Tavassoli N, de Souto Barreto P, Berbon C, Mathieu C, de Kerimel J, Lafont C. Implementation of the WHO integrated care for older people (ICOPE) programme in clinical practice: a prospective study. Lancet Healthy Longev 2022;3(6):e394-e404 [FREE Full text] [doi: 10.1016/S2666-7568(22)00097-6] [Medline: 36098317]
16. Stara V, Soraci L, Takano E, Kondo I, Möller J, Maranesi E. Intrinsic capacity and active and healthy aging domains supported by personalized digital coaching: survey study among geriatricians in Europe and Japan on eHealth opportunities for older adults. J Med Internet Res 2023;25:e41035 [FREE Full text] [doi: 10.2196/41035] [Medline: 37824183]
17. Ge H, Li J, Hu H, Feng T, Wu X. Digital exclusion in older adults: a scoping review. Int J Nurs Stud 2025;168:105082. [doi: 10.1016/j.ijnurstu.2025.105082] [Medline: 40279791]
18. Özgül E, Döner NH, Çürük GN. Investigating older adults' technology attitudes: psychometric evaluation and cross-cultural adaptation of the TechPH scale. Int J Med Inform 2025;203:106001. [doi: 10.1016/j.ijmedinf.2025.106001] [Medline: 40479776]
19. Li X, Zhang N, Yang J, Geng Z, Zhou J, Zhang J. Weight management personas of breast cancer patients undergoing chemotherapy in China: a multi-method study. BMC Med Inform Decis Mak 2024;24(1):108 [FREE Full text] [doi: 10.1186/s12911-024-02515-1] [Medline: 38664653]
20. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013;82(11):e251-E268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
21. Yu J, Si H, Jin Y, Qiao X, Ji L, Bian Y. Patterns of intrinsic capacity among community-dwelling older adults: identification by latent class analysis and association with one-year adverse outcomes. Geriatr Nurs 2022;45:223-229. [doi: 10.1016/j.gerinurse.2022.04.021] [Medline: 35576783]
22. Ma L, Zheng E, Fang Y, Chen H, Cai S, Luo F. Intrinsic capacity loss rates and protective factors among individuals aged 80 years and older in Chinese nursing homes: a latent class analysis. Geriatr Nurs 2024;60:42-51 [FREE Full text] [doi: 10.1016/j.gerinurse.2024.08.019] [Medline: 39217841]
23. Groos SS, Linn AJ, Kuiper JI, van Schoor NM, van der Velde N, van Weert JC. Combining user-centered design and behavioral theory to enhance health technologies: a personas-based approach for a primary-care based multifactorial falls risk assessment tool. Int J Med Inform 2024;186:105420 [FREE Full text] [doi: 10.1016/j.ijmedinf.2024.105420] [Medline: 38518678]

XSL•FO
RenderX

24. Holden RJ, Kulanthaivel A, Purkayastha S, Goggins KM, Kripalani S. Know thy eHealth user: development of biopsychosocial personas from a study of older adults with heart failure. Int J Med Inform 2017;108:158-167 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.006] [Medline: 29132622]
25. Yang Z, Xu L, Gao Y, Zhang C, Wang A. Tailored personas for self-management in home-based cardiac rehabilitation for patients with coronary heart disease: a qualitative study. Int J Nurs Stud 2025;163:105000. [doi: 10.1016/j.ijnurstu.2025.105000] [Medline: 39854909]
26. Jolliff A, Hill J, Zuraw M, Elliott C, Werner NE. Representing the needs of rural caregivers of people living with Alzheimer's disease and related dementias through user personas. Innov Aging 2024;8(12):igae096 [FREE Full text] [doi: 10.1093/geroni/igae096] [Medline: 39687843]
27. Ivankova NV, Creswell JW, Stick SL. Using mixed-methods sequential explanatory design: from theory to practice. Field Methods 2006;18(1):3-20. [doi: 10.1177/1525822X05282260]
28. Robinson P. Designing and conducting mixed methods research. Aust N Z J Public Health 2007;31(4):388. [doi: 10.1111/j.1753-6405.2007.00096.x]
29. Schäfer K, Rasche P, Bröhl C, Theis S, Barton L, Brandl C. Survey-based personas for a target-group-specific consideration of elderly end users of information and communication systems in the German health-care sector. Int J Med Inform 2019;132:103924 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.07.003] [Medline: 31569006]
30. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
31. Chow SC, Shao J, Wang H, Lokhnygina Y. Sample Size Calculations in Clinical Research. 3rd Edition. Boca Raton: CRC Press; 2017.
32. Nylund-Gibson K, Garber AC, Carter DB, Chan M, Arch DAN, Simon O. Ten frequently asked questions about latent transition analysis. Psychol Methods 2023;28(2):284-300. [doi: 10.1037/met0000486] [Medline: 35834194]
33. Liu Q, Li X, Hu M, Zhao Y, Wu S, Feng H. Factors influencing the self-management ability among older adults experiencing intrinsic capacity decline: a cross-sectional study. Front Aging Neurosci 2024;16:1456167 [FREE Full text] [doi: 10.3389/fnagi.2024.1456167] [Medline: 39660337]
34. Yu S, Wang J, Xia Y, Tang Q. The status quo and influencing factors of intrinsic capacity among community-dwelling older adults from the perspective of ecological systems theory: a cross-sectional study. BMC Geriatr 2024;24(1):934 [FREE Full text] [doi: 10.1186/s12877-024-05499-9] [Medline: 39533175]
35. Mahoney FI, Barthel DW. Functional evaluation: The Barthel index. Md State Med J 1965;14:61-65. [doi: 10.1037/t02366-000]
36. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
37. Ma L, Sun F, Tang Z. Social frailty is associated with physical functioning, cognition, and depression, and predicts mortality. J Nutr Health Aging 2018;22(8):989-995 [FREE Full text] [doi: 10.1007/s12603-018-1054-0] [Medline: 30272104]
38. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
39. Rubenstein LZ, Harker JO, Salvà A, Guigoz Y, Vellas B. Screening for undernutrition in geriatric practice: developing the short-form mini-nutritional assessment (MNA-SF). J Gerontol A Biol Sci Med Sci 2001;56(6):M366-M372. [doi: 10.1093/gerona/56.6.m366] [Medline: 11382797]
40. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
41. Yesavage JA, Sheikh JI. Geriatric depression scale (GDS). Clin Gerontol 2008;5(1-2):165-173. [doi: 10.1300/J018v05n01_09]
42. Huang ZT, Lai ET, Luo Y, Woo J. Social determinants of intrinsic capacity: a systematic review of observational studies. Ageing Res Rev 2024;95:102239. [doi: 10.1016/j.arr.2024.102239] [Medline: 38382677]
43. Leung CLK, Li K, Wei VWI, Tang A, Wong SYS, Lee SS. Profiling vaccine believers and skeptics in nurses: a latent profile analysis. Int J Nurs Stud 2022;126:104142 [FREE Full text] [doi: 10.1016/j.ijnurstu.2021.104142] [Medline: 34923316]
44. Tofighi D, Enders CK. Identifying the correct number of classes in growth mixture models. In: Hancock GR, Samuelsen KM, editors. Advances in Latent Variable Mixture Models. Greenwich: Information Age Publishing; 2007:317-341.
45. Tein J, Coxe S, Cham H. Statistical power to detect the correct number of classes in latent profile analysis. Struct Equ Modeling 2013;20(4):640-657 [FREE Full text] [doi: 10.1080/10705511.2013.824781] [Medline: 24489457]
46. Nylund KL, Asparouhov T, Muthén BO. Erratum: deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study (Struct Equ Modeling. 2007;14(4):535-569). Struct Equ Modeling 2007;14(4):535-569. [doi: 10.1080/10705510701575396]
47. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]

XSL•FO
RenderX

48. Möller J, Stara V, Amabili G, Barbarossa F, Riccardi GR, Martella C. Toward innovation in healthcare: an analysis of the digital behavior of older people in Europe and Japan for the introduction of a technological coaching system. Healthcare (Basel) 2024;12(2):143 [FREE Full text] [doi: 10.3390/healthcare12020143] [Medline: 38255032]
49. Rahimi S, Khatooni M. Saturation in qualitative research: an evolutionary concept analysis. Int J Nurs Stud Adv 2024;6:100174 [FREE Full text] [doi: 10.1016/j.ijnsa.2024.100174] [Medline: 38746797]
50. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
51. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015;13(6):554-561 [FREE Full text] [doi: 10.1370/afm.1865] [Medline: 26553895]
52. Weller BE, Bowen NK, Faubert SJ. Latent class analysis: a guide to best practice. J Black Psychol 2020;46(4):287-311. [doi: 10.1177/0095798420930932]
53. Zhao Y, Chen Y, Xiao LD, Liu Q, Nan J, Li X. Intrinsic capacity trajectories, predictors and associations with care dependence in community-dwelling older adults: a social determinant of health perspective. Geriatr Nurs 2024;56:46-54. [doi: 10.1016/j.gerinurse.2023.12.022] [Medline: 38237340]
54. Guo W, Meng L, Han J, Yang B, Sun J, Guo Y. Intrinsic capacity and its association with predictors among Chinese empty nest older adults in communities: a latent class analysis. BMC Geriatr 2024;24(1):1008 [FREE Full text] [doi: 10.1186/s12877-024-05583-0] [Medline: 39702076]
55. Ciuffreda I, Amabili G, Casaccia S, Benadduci M, Margaritini A, Maranesi E. Design and development of a technological platform based on a sensorized social robot for supporting older adults and caregivers: GUARDIAN ecosystem. Int J of Soc Robotics 2023;17(5):803-822. [doi: 10.1007/s12369-023-01038-5]
56. Hoogendijk EO, Dent E, Koivunen K. Intrinsic capacity: an under-researched concept in geriatrics. Age Ageing 2023;52(10):afad183. [doi: 10.1093/ageing/afad183] [Medline: 37782890]
57. Ishikawa M. Internalization of negative societal views on old age into self-perceptions of aging: exploring factors associated with self-directed ageism. Front Sociol 2023;8:1291325 [FREE Full text] [doi: 10.3389/fsoc.2023.1291325] [Medline: 38045535]
58. Zhang Z, Yang L, Cao H. Effect of inter-generational living arrangement and digital exclusion on unmet healthcare needs among older adults: findings from two national cohort studies. J Clin Nurs 2026;35(5):2159-2180. [doi: 10.1111/jocn.17638] [Medline: 39789828]
59. Beard JR, Si Y, Liu Z, Chenoweth L, Hanewald K. Intrinsic capacity: validation of a new WHO concept for healthy aging in a longitudinal Chinese study. J Gerontol A Biol Sci Med Sci 2022;77(1):94-100. [doi: 10.1093/gerona/glab226] [Medline: 34343305]
60. Wei X, Chen Y, Qin J, Yang Y, Yang T, Yan F. Factors associated with the intrinsic capacity in older adults: a scoping review. J Clin Nurs 2024;33(5):1739-1750. [doi: 10.1111/jocn.17017] [Medline: 38345142]
61. Yu J, Jin Y, Si H, Bian Y, Liu Q, Qiao X. How does social support interact with intrinsic capacity to affect the trajectory of functional ability among older adults? Findings of a population-based longitudinal study. Maturitas 2023;171:33-39. [doi: 10.1016/j.maturitas.2023.03.005] [Medline: 37001477]
62. Liu M, Chang Y, Zhao S, Guo W, Ji X, Liu Y. The effect of the interaction between intrinsic capacity and social support on the trajectories of activities of daily living in older adults. Geriatr Nurs 2024;60:231-240 [FREE Full text] [doi: 10.1016/j.gerinurse.2024.09.007] [Medline: 39298925]

Abbreviations

aBIC: adjusted BIC
AIC: Akaike information criterion
BIC: Bayesian information criterion
BLRT: bootstrapped likelihood ratio test
GRAMMS: Good Reporting of a Mixed Methods Study
HEALTHY: Comprehensive Evaluation of Intrinsic Capacity in China Study
IC: intrinsic capacity
ICOPE: Integrated Care for Older People
LMRT: Lo-Mendell-Rubin likelihood ratio test
LPA: latent profile analysis
OR: odds ratio
WHO: World Health Organization

XSL•FO
RenderX

Edited by J Wang; submitted 22.Aug.2025; peer-reviewed by G Amabili, X Yuan; comments to author 13.Jan.2026; revised version received 18.Mar.2026; accepted 11.May.2026; published 27.May.2026.

Please cite as:
Xuan Z, Niu Y, Kang R, Xiao Q, Jin S, Zhao J, Wang Y, Chang H
Developing Customized Personas to Capture Intrinsic Capacity Profiles and Digital Monitoring Intentions in Older Adults: Mixed Methods Study
JMIR Aging 2026;9:e82867
URL: https://aging.jmir.org/2026/1/e82867
doi:10.2196/82867
PMID:

©Zehui Xuan, Yirou Niu, Ruifu Kang, Qian Xiao, Shuai Jin, Jie Zhao, Yanling Wang, Hong Chang. Originally published in JMIR Aging (https://aging.jmir.org), 27.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Aging Like Milk or Fine Wine: Content Analysis of Reactions to One's Future Self on TikTok

Reuben Ng[1], PhD; Nicole Indran[1], BSocSci (Hons)

Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore, Singapore

Corresponding Author:
Reuben Ng, PhD
Lee Kuan Yew School of Public Policy
National University of Singapore
469C Bukit Timah Road
Singapore, 259772
Singapore
Phone: 65 6601 5229
Email: spprng@nus.edu.sg

Abstract

Background: The "aged" filter has gone viral on TikTok. This filter leverages AI to superimpose signs of aging, such as wrinkles, crow's feet, and sagging skin, onto a user's face. Notably, several dermatologists have attested to the accuracy of this filter in simulating the effects of aging. Reactions among the TikTok community have been mixed, with some utterly dismayed but others wholeheartedly embracing the sight of their aged visages.

Objective: This study takes a deep dive into the ways in which TikTok users react to their aged selves through a content analysis of videos featuring the aged filter. Understanding how they respond to the filter can shed light on the types of meanings commonly attributed to old age as a social construct. The following research questions form the premise of our content analysis of videos on TikTok: (1) How do individuals react to the way their possible future selves look? (2) What types of attitudes toward old age are being conveyed in these videos? (3) What do these responses say about prevailing cultural beliefs, values, and attitudes toward old age?

Methods: To build a comprehensive dataset, we compiled all publicly available videos (N=2682) using the following hashtags as search queries: "#AgedFilter," "#OldFilter," and "#AgeFilter." Collectively, these hashtags amassed >2,068,100,000 views at the time of writing this manuscript. After applying our exclusion criteria, the dataset comprised 681 videos, which received a total of 34,382,236 likes, 1,204,741 favorites, and 211,006 comments. Both deductive and inductive approaches informed our content analysis.

Results: A total of 7 themes emerged from our content analysis. Half of the videos (n=342, 50%) fell under theme 1, "anxiety about aging." The second-largest theme (n=95, 14%) focused on the "pursuit of youthfulness" (theme 2). Theme 3 was about "aging with confidence" (n=84, 12%) and theme 4 was about the "beauty of aging" (n=54, 8%). Theme 5 examined the "privilege of aging" (n=48, 7%) and theme 6 examined "generational connections" (n=43, 6%). Theme 7 covered the "inevitability of aging" (n=15, 2%).

Conclusions: This study marks one of the first attempts to analyze videos using the aged filter on TikTok. Our findings suggest that ageist beliefs may be prevalent among some users. Going forward, there is value in adopting a definition of beauty that transcends age. By doing so, aging can be reframed as a natural part of life to be embraced rather than as a grim reality to be dreaded.

(JMIR Aging 2026;9:e70794) doi:10.2196/70794

KEYWORDS

aging; TikTok; social media; artificial intelligence; AI; ageism; stigma; cultural beliefs

XSL•FO
RenderX

Introduction

TikTok boasts a presence in more than 160 countries and lays claim to approximately a billion monthly active users [1]. One defining characteristic that sets TikTok apart from other social media services is its extensive array of creative tools—filters, sounds, and stickers, among others—which users can experiment with when creating videos. In July 2023, the "aged" filter went viral on TikTok [2]. This filter essentially leverages AI to superimpose signs of aging, such as wrinkles, crow's feet, and sagging skin, onto a user's face. Notably, several dermatologists have attested to the accuracy of this filter in simulating the effects of aging [2]. Reactions among the TikTok community have been mixed, with some utterly dismayed and others wholeheartedly embracing the sight of their aged visages [2]. This study takes a deep dive into the ways TikTok users react to their aged selves through a content analysis of videos using the aged filter.

From a conceptual viewpoint, this study marks one of the first attempts to analyze videos using the aged filter on TikTok. Other gerontological studies that have dealt with social media focus on other topics such as stereotypes about older adults [3-7]. Understanding how individuals respond to the aged filter can shed light on their attitudes toward the physical changes that attend the aging process. From a practical viewpoint, this study can illuminate the types of meanings people attribute to old age as a social construct. It is well documented that TikTok is populated primarily by younger individuals. On the basis of data from July 2023, more than 37% of TikTok's global user base is composed of individuals aged 18 to 24 years. Those aged 25 to 34 years accounted for approximately 32% of the TikTok populace [8]. Insights from this study can support broader efforts to improve societal perceptions of old age.

In accordance with the theory of social constructionism, all human knowledge is produced, communicated, and sustained in social contexts [9]. To conceive of age as a social construct is to acknowledge that although aging itself is a biological fact, the meanings assigned to it are by no means cast in stone [10]. Rather, these meanings are colored by the values, attitudes, and narratives that prevail in society at a particular time [10].

In virtually all contemporary societies, standards of attractiveness are predicated on having a youthful appearance [11-13]. Specifically, youth has been constructed as the epitome of health, vitality, and beauty. In contrast, aging is typically associated with a loss of appearance-related assets, especially in terms of social and sexual desirability [11,14]. Perhaps the phenomenon that best exemplifies the hegemonic ideal of youthfulness is the enduring profitability of the antiaging industry, which purveys a plenitude of products, regimens, and services often peddled as elixirs of youth [15,16]. The global antiaging market was valued at approximately US $58.5 billion in 2020 and is forecast to experience a compound annual growth rate of 7% from 2021 to 2026 [17].

Critical feminists have censured the antiaging industry for the way it has problematized a natural transition, cast old age as the enemy, and ultimately engineered a climate of fear surrounding the aging process [18]. Many of them have argued that the construct of antiaging is a capitalistic framework designed to cash in on feelings of insecurity [19]. These feelings of insecurity are implanted through marketing that casts wrinkles and fine lines as issues to be "treated" or "prevented," and aging as a process to be "delayed," "repaired," or "inhibited" [20]. A striking example of how the antiaging industry instills a fear of aging is the way Elizabeth Arden advertised its PREVAGE moisturizer. Specifically, a fencing mask was used to symbolize the protective nature of the brand's antiaging skincare technology. One of the lines in the advertisement, "[p]repare yourself for the beauty battlefield," further underscores the message that aging ought to be fought and conquered [21].

Another camp emphasizes the need to pay attention to the role of agency when assessing one's reasons for conducting antiaging work. Some postulate that this qualifies as disidentification [22,23], an adaptive strategy that involves a reconceptualization of the self in a way that minimizes one's aging identity as a basis of self-evaluation. More precisely, in a society where beauty is frequently entwined with youthfulness, partaking in antiaging work can serve as a way to elevate one's self-esteem and confidence [24]. Research states that cosmetic treatments can enhance psychological well-being and ameliorate appearance-based distress [25]. Engaging in antiaging work may also function as a way to sidestep the marginalization that ensues when one is visibly older [12,26].

It is widely accepted that a sense of anxiety accompanies the notion of aging [27-29]. Lasher and Faulkender [30] define this anxiety as the anticipation of losses that aging brings about. They consider that anxiety about aging encompasses 4 distinct dimensions: physical, psychological, social, and transpersonal. Particularly apposite to this study is the physical dimension, which includes factors such as perceptions of one's health status, concerns about physical self-efficacy, as well as the physical changes that come with growing older. Anxiety about aging may manifest in 3 types of fears: fear of aging, fear of being old, and fear of older people [30]. Anxiety about aging is said to be intertwined with anxiety about death [31,32].

The perceived loss of physical attractiveness can pose a threat to one's self-esteem and psychological well-being, especially in a world where good looks constitute a form of social currency [33]. Additionally, as per the theory of stereotype embodiment [34], the negativity ascribed to societal meanings of old age can be hazardous to one's health. When assimilated as part of one's culture, age stereotypes may eventually become self-definitions that shape health outcomes in later life. Since the 1990s, research done in regions including America, Europe, and Asia has found that older adults who hold negative age stereotypes have a lower sense of efficacy, a higher risk of depression, and poorer cardiovascular health [34-36]. Conversely, those who embrace positive age stereotypes generally enjoy better functional health, enhanced well-being, and increased longevity [34,37,38]. There is also evidence that people who endorse negative age stereotypes in their younger years confront a greater likelihood of experiencing a cardiovascular event in later life than those who hold positive age stereotypes [34].

The following research questions form the premise of our content analysis of videos on TikTok: (1) How do individuals

XSL•FO
RenderX

react to the way their possible future selves look? (2) What types of attitudes toward old age are being conveyed in these videos? (3) What do these responses say about prevailing cultural beliefs, values, and attitudes toward old age?

Methods

Dataset

Consistent with previous research [39], a new TikTok (ByteDance Ltd) account was created to compile the videos. This was done to reduce bias since videos on the app are arranged based on a complex algorithm, which takes into account the popularity of the video (measured by views, likes, comments, and shares), the popularity of the creator (measured by followers and engagement), any previous content that was engaged with, and the geographic location of the device where the app was accessed [39]. No content was previously engaged with to ensure a regular user's experience when using the app [39].

The aged filter on TikTok is a type of visual effect designed to simulate an aged appearance. Several versions of this filter exist on the platform, but they all serve the same purpose. We compiled all publicly available videos (N=2682) using the following hashtags as search queries: "#AgedFilter" (n=901), "#OldFilter" (n=877), and "#AgeFilter" (n=904). Hashtags enhance the searchability of content by organizing it into categories. Collectively, these hashtags amassed >2,068,100,000 views at the time of writing this manuscript. Videos were compiled from August 21, 2023, to September 21, 2023. We removed (1) videos where the filter served purposes other than the self-examination of one's future appearance, such as role-playing or performing skits (n=745); (2) videos not in English (n=682); (3) videos where the filter was applied to participants other than the user, such as celebrities or animals (n=246); (4) videos with a different filter or without a filter (n=218); (5) videos where the user had no discernible reaction (n=90); and (6) videos that appeared more than once in the dataset, in which case only one was retained (n=20). After applying the exclusion criteria, our dataset consisted of 681 videos, which received a total of 34,382,236 likes, 1,204,741 favorites, and 211,006 comments. Table 1 shows the data collection process.

Table 1. Themes identified in TikTok content featuring the aged filter based on a content analysis of 681 videos.

Themes	Videos, n (%)
Anxiety about aging	342 (50)
Pursuit of youthfulness	95 (14)
Aging with confidence	84 (12)
Beauty of aging	54 (8)
Privilege of aging	48 (7)
Generational connections	43 (6)
Inevitability of aging	15 (2)

Content Coding of TikTok Videos

Following past research [40], the coding rubric was designed using both deductive and inductive modes of reasoning [41]. Analyses led by a directed or deductive approach begin with the identification of an initial set of codes based on past scholarship [41]. Meanwhile, in inductive content analyses, codes are derived directly from the data [41]. We used both deductive and inductive approaches to ensure certain pertinent assumptions informed the analysis, even as we were mindful that new categories would emerge inductively.

To develop a preliminary codebook, we identified categories based on scholarship concerning perceptions of physical appearance in later life [12,14,33]. All the videos were viewed twice by the authors to ensure familiarity with the data. The objective of the first reading was to ascertain the validity of the initial set of categories, as well as to modify the codebook until all variables were refined and defined clearly. During this first reading, a new category was added whenever a video featured a particular attribute that could not be suitably coded into any of the existing categories and appeared frequently enough to warrant its own category. The objective of the second reading was to ensure we had a framework sufficiently representative of the different types of videos to finalize the coding rubric.

To maintain analytic rigor, the 2 coders had regular discussions during which any discrepancies were reviewed and adjudicated. Areas of major overlap were identified and sectioned into broader themes. The percentage agreement between the 2 raters was 96.5% with a weighted Cohen κ of 0.93 (P<.001), showing high interrater reliability. Seven themes emerged from this iterative procedure. It must be highlighted that a video may be classified under more than one theme. As mentioned in previous research, categories in a content analysis need not be mutually exclusive, although they should be internally homogeneous (ie, coherent within themes) and externally heterogeneous (ie, distinct from each other) as far as possible [42,43].

Ethical Considerations

This study analyzed publicly available TikTok data and did not involve interaction with human participants. According to institutional guidelines, ethical review board approval was not required for this study. Informed consent was not obtained as all data were publicly accessible and anonymized prior to analysis.

XSL•FO
RenderX

Results

Summary of Insights From Content Analysis of TikTok Videos

Of the 681 videos, half (n=342, 50%) fell under theme 1, "anxiety about aging." The "pursuit of youthfulness" was the focal point of theme 2 (n=95, 14%). Theme 3 was about "aging with confidence" (n=84, 12%), while theme 4 was about the "beauty of aging" (n=54, 8%). The "privilege of aging" was discussed in 7% (n=48) of the videos. Theme 6, "generational connections," made up 6% (n=43) of the dataset. Theme 7 covered the "inevitability of aging" (n=15, 2%). Table 1 summarizes the themes.

Theme 1: Anxiety About Aging

In this theme, users reacted to their aged selves with varying levels of displeasure. "Shocked," "traumatized," "disturbed," "depressed," "devastated," "terrified," and "enraged" are but some of the words TikTokers used upon seeing themselves through the filter. Many were convinced the filter did them "dirty." One said the filter "lower[ed] [her] self-esteem," and another asserted that the creator of the filter essentially "amplified her body dysmorphia." An audio clip used by multiple users featured a woman uttering, "I need to buy a gun."

Some evoked death-related imagery to describe their future appearances. For instance, one TikToker likened herself to someone who had "just been dug up from the grave." Another bewailed, "It looks like I am [expletive] decaying, bro. I don't even look human, bro, I look like a [expletive] corpse, bro." One insisted he looked like he was "two seconds away from [his] final breath."

Comparisons to movie characters were commonplace. One referred to herself as a "wrinkly Voldemort," and another as the "after version of Rose from Titanic." One poster claimed she looked like "something from The Conjuring."

References to witches emerged in a couple of uploads. A case in point is a video where a user asked, "Why do I look like the witch in the [expletive] Bog that steals your firstborn child? Why do I look like I put a curse on the princess in the castle?" Another opined that she bore a resemblance to a "witch living in the woods."

There were posters who took issue with how fellow users seemed to age "well" or "gracefully." In particular, one individual lamented, "Some people age like fine wine, some people age like milk. I am of the milk variety."

Many were vexed by how old they appeared. "I look like I fought in World War II," exclaimed one user. Another griped, "I look like I am 110 years old and was left out in the sun for an amount of time." One user groaned, "I look like someone left me in the pool for ages."

A few TikTokers took a more dramatic approach, declaring their intent to withdraw from the aging process altogether. Statements that encapsulated this sentiment include "Old isn't for me, guys. Peace out," "So, youth is not in my future," "I need to die young," and "I will not be partaking in aging."

Some were discomfited by how unapproachable or unfriendly they appeared. For instance, one exclaimed, "I look like I eat children." Another complained that she looked like "she would bite the nursing home staff."

Theme 2: Pursuit of Youthfulness

Posts under theme 2 featured users attempting to stave off the aging process through enhanced skincare regimens, cosmetic procedures, or antiaging supplements. Such videos typically depicted individuals frantically applying sunscreen or moisturizer, occasionally using copious amounts for dramatic effect. Similar videos showcased users incorporating facial rollers into their skincare routines.

Many users conveyed a desire to undergo cosmetic procedures such as Botox (AbbVie Inc) or face-lifts. "Give me the botox right now. I want whatever Kris Jenner has.... Before you start commenting 'aging is beautiful', shut the [expletive] up, okay, this is not. This is hideous," commanded one user. Another video contained a text sticker stating, "Running to put on liquid botox and sunscreen [because] I refuse to believe this is what I will look like when I'm old."

One TikToker reached for a tub of collagen peptides and other supplements, vowing to "prove the filter wrong." Another shared that she would be "switching from 0.05% to 0.1% tretinoin" in order not to "age like spoiled milk."

Theme 3: Aging With Confidence

Videos in this category featured users feeling content or confident with their older selves. Some enthused over how "sophisticated" and "powerful" they looked. One effused unapologetically, "I'm gonna be such a hot granny." Another was certain of her ability to continue "luring men" in her advanced age. One proudly declared he was "aging like fine wine" and dubbed himself a "good-looking man." There was even a playful warning from one poster to "hide the abuelas," presumably because of his attractiveness. Several TikTokers called themselves "cute." For example, one commented, "I hope I age this good. I look like a cute old lady."

Theme 4: Beauty of Aging

This theme revolved around the positive aspects of aging, focusing on the personal growth and enrichment it brings.

Several users embraced aging as a path toward increased wisdom, inner strength, and personal fulfillment. "I can't wait to grow old...to have experienced so much life, hold powerful wisdom, and witness myself blossom and change over time. I get so emotional and filled with love seeing the older version of me," one remarked. Another shared, "In my family, wrinkles show laughter and grey hair means wisdom. Age is beauty. I'm so excited to meet her. She looks like she gives the best hugs and runs the local crafts store." One believed that her older self appeared fulfilled, affirming, "She looks like she's lived." Another individual volunteered that being able to reach old age means having "lived and survived."

A group of posters dwelled on the warmth radiated by their future selves. Some relished the prospect of becoming grandparents. One user mentioned that she looked like "she

XSL•FO
RenderX

gives the best hugs, is loved by all the grandkids, and always slips you money every time she sees you." Another focused on the pleasant demeanor of her aged self, saying, "I like her. She has kind eyes. She's lived a full life. She's the type of woman who will comfort you when you're sad. She deeply understands people."

Theme 5: Privilege of Aging

In these posts, users reflected on aging as a "gift," "blessing," or "privilege." A fair share of individuals spoke about how their struggles with mental illness made them grateful for the opportunity to envision life in the future. For example, one confessed, "I've been so mentally ill my whole life. I've genuinely never pictured living this long before." Similarly, a self-proclaimed cancer survivor announced that it would be the "greatest gift" to live till old age. A paramedic said the filter made him think about "all the people [he] was not able to save," believing these people would have desired to live a long life. He concluded, "Aging is a gift. It's a gift to be able to lead a fulfilled, long life."

Some TikTokers were disheartened by the overwhelming negativity surrounding the filter. One maintained that aging is a "privilege," not a "punishment." Another criticized the TikTok community for being "wildly ageist," stressing that anyone would be "lucky to live this long."

Theme 6: Generational Connections

Videos under theme 6 had users reveling in the fact that their aged selves were the spitting image of their parents or grandparents. A representative video contained a caption that read, "What's so wonderful about this is I immediately see my grandma, dad, mum, aunties, and uncles in my features. We are combinations of so many people."

A handful of posters had a deep sense of admiration for their parents or grandparents, which in turn fueled their excitement about growing older. For instance, one gushed, "I look like my grandma who I think is the most beautiful woman." Another anticipated "aging gracefully" like her grandmother. This sentiment was especially pronounced for those who had lost their parents or grandparents.

Theme 7: Inevitability of Aging

In this theme, users either normalized aging as a natural part of life or resigned themselves to the inevitability of its implications. One TikToker spoke reassuringly, "You're gonna age and that's okay." Another put forth, "We all have to get old one day."

One individual was upfront about becoming less attractive. "I wanna embrace the fact that we're all gonna look like this someday and nobody's gonna die looking sexy," she proclaimed.

Discussion

Principal Findings

In this study, we examined the reactions of individuals to the appearance of their possible future selves by performing a content analysis of videos using the aged filter on TikTok. Our analysis unveiled 7 themes: "anxiety about aging" (theme 1), "pursuit of youthfulness" (theme 2), "aging with confidence" (theme 3), "beauty of aging" (theme 4), "privilege of aging" (theme 5), "generational connections" (theme 6), and "inevitability of aging" (theme 7).

Collectively, themes 1 and 2 accounted for more than two-thirds of the videos, which underscores the widespread negativity surrounding the concept of aging among users. In particular, old age is perceived as entailing a loss of one's physical attractiveness [14,30]. An outsized number of users couched their future selves in words that were highly pejorative. Many also conveyed their intent to pay more attention to their skincare routines or to undergo cosmetic procedures to prolong their youthful countenance and quell the anxiety of growing older. These themes reflect deeply ingrained beliefs that equate youth with beauty and old age with undesirability. Such beliefs can exact a toll on one's self-esteem and overall psychological well-being [33]. Furthermore, through the lens of the theory of stereotype embodiment, these negative beliefs about aging can jeopardize one's overall health outcomes [34,37,38].

Theme 3 portrayed users with a positive self-image and sense of self-assurance regarding growing older, attributes that will likely stand them in good stead as they eventually advance into later life. Posts under this theme offer a powerful and refreshing counternarrative that reimagines old age as a period characterized by continued appeal. However, this sense of contentment or empowerment could very well be tethered to the perception of one's physical appearance. One person commented that he aged like "fine wine," insinuating that less appealing ways to age do exist. It is possible that their reactions would be less favorable had the filter presented a version of themselves with more pronounced signs of aging.

Posts under theme 4 reflect the belief that old age is a phase marked by the accrual of experience and wisdom, as well as the development of a mellower personality. Implicit in these posts is a recognition that aging is a multifaceted journey that goes beyond mere physical transformation. In particular, growing older involves certain emotional, social, and psychological changes [44]. Similarly, theme 5 illustrates that for some members of the TikTok population, the experience of aging is not so much a matter of one's changing visage as it is an opportunity to enjoy a long and fulfilling life. Both themes 4 and 5 substantiate findings showing that as people age, their priorities may evolve such that appearance no longer figures into their self-evaluations as prominently [45-47].

Videos under theme 6 depicted users getting emotional over the striking resemblance between their future selves and their parents or grandparents. A growing body of evidence documents that many individuals tend to perceive their grandparents more positively than the general older populace [48,49]. Moreover, being close to one's grandparents has been linked to more positive views of older adults [50,51]. The excitement expressed by certain individuals about being able to look like their older family members may be anchored in strong emotional ties.

Theme 7 demonstrated users having come to terms with the reality of the aging process. Cognizant of the fact that they would eventually deviate from the hegemonic ideal of youthfulness, some TikTokers adopted cognitive strategies that involved adjusting their expectations about their eventual selves

XSL•FO
RenderX

[47]. For instance, a user immediately reconciled herself to the fact that she would no longer be considered attractive by societal standards, thereby decoupling her self-worth from youth-based ideals and accepting her older self.

Implications

Our findings suggest multiple avenues for combating ageism. First, there is a need to step up efforts to reframe aging. Over the years, various social movements, such as the "body positivity" movement, have sprouted up to counter prevailing beauty standards deemed harmful. The principles of inclusivity underlying these movements could be extended to the way society perceives old age. Recently, a number of beauty companies have eschewed the term "antiaging." Instead, these companies now opt for seemingly anodyne alternatives, such as "regeneration," "renewal," and "vitality," which, at the end of the day, however, continue to promulgate the idea that aging skin is undesirable [52]. True inclusivity could involve accepting natural signs of aging. Even as some may rely on antiaging solutions as a way to cope with the aging process [12,24,26], the reality is that all individuals will eventually reach a point where physical signs of aging can no longer be camouflaged by cosmetic interventions [45].

Second, given its preponderant role in stigmatizing old age, the media bears significant responsibility for improving portrayals of older adults [12]. Proactive steps could be taken to increase their representation. For example, media outlets and advertising agencies could actively include older individuals in beauty-related content. This may help transmit the message that beauty is not the prerogative of youth and that aging is but a natural and inevitable part of life. It could also foster positive age stereotypes and lead to more favorable health outcomes [34].

Third, it is paramount that society unlearns the notion that aging brings with it a loss of physical attractiveness. Instead, aging should be recognized as a time of increased experience [53] and emotional intelligence [54]. More importantly, it should be seen as a privilege since not all individuals will live to see old age [55]. Promoting the idea of self-love and self-acceptance may be worthwhile as well [56].

Fourth, tapping into the deep emotional bonds that people share with their parents or grandparents may inspire a more positive outlook on aging [48,49]. When individuals discern aspects of their family members in themselves, they may navigate their own aging journey more favorably. Policymakers could involve grandparents or aging parents in interventions to combat ageism [57,58], harnessing them as proof that aging can indeed be a time of continued growth.

Finally, our findings make a case for implementing initiatives to reframe aging from an early age so that individuals are exposed to positive narratives on aging long before they reach later life. To this end, efforts to reframe aging could be integrated into schools and workplaces.

Limitations and Future Directions

This study is subject to several limitations. First, our findings are unique to a segment of TikTok users whose videos happened to be included in this study and are therefore not generalizable to the wider population. Second, as we did not have access to the motivations that drove users to create the videos, our interpretations were molded by our own thoughts and assumptions. It is worth noting that TikTok is a platform that thrives on entertaining content. Hence, some may have reacted in specific ways to bolster engagement. Third, we lacked information on the demographic characteristics of users, such as their gender. We also refrained from using automated recognition systems given their potential to perpetuate gender biases [59]. Thus, we could not comment on how demographic factors influence individuals' attitudes toward their older selves. It is well established that men and women are subject to different sets of expectations as far as aging is concerned [11,60]. This double standard of aging places women in a position where they contend with greater losses as growing older is often perceived to entail an erosion of their most valued social asset: their physical allure. Men, in contrast, are sometimes regarded as more attractive as they age [61]. Fourth, a substantial number of videos had to be excluded as they were not in English. Future studies could examine this topic in other cultural contexts [62]. Finally, although we recorded the overall number of videos for each theme, we did not track the number of videos associated with the various subthemes. Future research could adopt a quantitative approach [63] by noting down the number of videos for different subthemes. This would lay the groundwork for a regression analysis to explore potential relationships in the data [64, 65].

Other avenues for further gerontological inquiry include an investigation of the psychological effects of using the aged filter. In addition, there is room for an analysis of how reactions to one's aged self differ across age groups. Studying the comments section of videos that use the aged filter presents another promising direction for subsequent research.

Conclusions

This study has shed light on the prevalence of ageist beliefs among some users of TikTok. Going forward, there is value in adopting a definition of beauty that transcends age. By doing so, aging can be reframed as a natural part of life to be embraced rather than as a grim reality to be dreaded.

Acknowledgments

The authors thank G Koh for research assistance.

Funding

The authors gratefully acknowledge the Commonwealth Fund's Harkness Fellowship in Health Care Policy and Practice. The funders had no role in the study design, data collection, analysis, writing, or the decision to publish.

XSL•FO
RenderX

Data Availability

Data are publicly available on TikTok [66].

Authors' Contributions

RN designed the study, developed the methodology, analyzed the data, wrote the paper, provided supervision, and acquired the funding. NI co-designed the study, coanalyzed the data, and cowrote the paper.

Conflicts of Interest

None declared.

References

1. Kumar N. Quora statistics 2026 (active users and demographics data). Demand Sage. URL: https://www.demandsage.com/quora-statistics/ [accessed 2023-03-29]
2. Butera I. A new TikTok filter predicts Gen Z's future wrinkles, and insecurities, in a viral trend. USA Today. 2023. URL: https://www.usatoday.com/story/life/health-wellness/2023/07/17/tiktok-old-age-filter/70409997007/ [accessed 2023-08-29]
3. Jimenez-Sotomayor MR, Gomez-Moreno C, Soto-Perez-de-Celis E. Coronavirus, ageism, and Twitter: an evaluation of tweets about older adults and COVID-19. J Am Geriatr Soc 2020 Aug;68(8):1661-1665 [FREE Full text] [doi: 10.1111/jgs.16508] [Medline: 32338787]
4. Levy BR, Chung PH, Bedford T, Navrazhina K. Facebook as a site for negative age stereotypes. Gerontologist 2014 Apr;54(2):172-176. [doi: 10.1093/geront/gns194] [Medline: 23392643]
5. Ng R, Indran N, Liu L. Ageism on Twitter during the COVID-19 pandemic. J Soc Issues 2022 Jul 20;78(4):842-859 [FREE Full text] [doi: 10.1111/josi.12535] [Medline: 35942488]
6. Ng R, Indran N. Videos about older adults on TikTok. PLoS One 2023 Aug 2;18(8):e0285987 [FREE Full text] [doi: 10.1371/journal.pone.0285987] [Medline: 37531317]
7. Ng R, Indran N. Hostility toward Baby Boomers on TikTok. Gerontologist 2022 Sep 07;62(8):1196-1206 [FREE Full text] [doi: 10.1093/geront/gnac020] [Medline: 35106587]
8. Distribution of TikTok users worldwide as of October 2025, by age and gender. Statista. URL: https://www.statista.com/statistics/1299771/tiktok-global-user-age-distribution/ [accessed 2026-07-29]
9. Berger PL, Luckmann T. The Social Construction of Reality: A Treatise in the Sociology of Knowledge. New York, NY: Knopf Doubleday Publishing Group; 1966.
10. Gergen MM, Gergen KJ. Positive aging: new images for a new age. Ageing Int 2001 Dec;27:3-23. [doi: 10.1007/s12126-001-1013-6]
11. Åberg E, Kukkonen I, Sarpila O. From double to triple standards of ageing. Perceptions of physical appearance at the intersections of age, gender and class. J Aging Stud 2020 Dec;55:100876 [FREE Full text] [doi: 10.1016/j.jaging.2020.100876] [Medline: 33272447]
12. Clarke LH, Griffin M. Visible and invisible ageing: beauty work as a response to ageism. Ageing Soc 2008 Nov 03;28(5):653-674. [doi: 10.1017/s0144686x07007003]
13. Clarke LH. Facing Age: Women Growing Older in Anti-Aging Culture. London, UK: Bloomsbury Publishing; 2010.
14. Hurd Clarke L, Griffin M. The body natural and the body unnatural: beauty work and aging. J Aging Stud 2007 Aug;21(3):187-201. [doi: 10.1016/j.jaging.2006.11.001]
15. Calasanti T. Ageism, gravity, and gender: experiences of aging bodies. Generations 2005 Sep;29(3):8-12 [FREE Full text]
16. Fishman JR, Binstock RH, Lambrix MA. Anti-aging science: the emergence, maintenance, and enhancement of a discipline. J Aging Stud 2008 Dec 01;22(4):295-303 [FREE Full text] [doi: 10.1016/j.jaging.2008.05.010] [Medline: 23264719]
17. Size of the anti-aging products market worldwide from 2022 to 2032. Statista. URL: https://www.statista.com/statistics/509679/value-of-the-global-anti-aging-market/ [accessed 2022-07-15]
18. Flatt MA, Settersten RAJ, Ponsaran R, Fishman JR. Are "anti-aging medicine" and "successful aging" two sides of the same coin? Views of anti-aging practitioners. J Gerontol B Psychol Sci Soc Sci 2013 Nov;68(6):944-955 [FREE Full text] [doi: 10.1093/geronb/gbt086] [Medline: 24022620]
19. Williams CM. Behind the mask: an analysis of women's perceptions and rationale toward the purchase and use of anti-aging products. University of Northern Iowa. 2013. URL: https://scholarworks.uni.edu/etd/35/ [accessed 2026-05-15]
20. Ng R, Indran N. Innovations for an aging society through the lens of patent data. Gerontologist 2024 Feb 01;64(2):gnad015 [FREE Full text] [doi: 10.1093/geront/gnad015] [Medline: 37497634]
21. Ellison KL. Age transcended: a semiotic and rhetorical analysis of the discourse of agelessness in North American anti-aging skin care advertisements. J Aging Stud 2014 Apr;29:20-31 [FREE Full text] [doi: 10.1016/j.jaging.2013.12.003] [Medline: 24655670]
22. Richeson JA, Shelton JN. A social psychological perspective on the stigmatization of older adults. In: When I'm 64. Washington, DC: National Academies Press (US); 2006.

XSL•FO
RenderX

23. Steele CM. A threat in the air: how stereotypes shape intellectual identity and performance. Am Psychol 1997;52(6):613-629. [doi: 10.1037/0003-066X.52.6.613]
24. Sadick NS. The impact of cosmetic interventions on quality of life. Dermatol Online J 2008 Aug 15;14(8):2 [FREE Full text] [Medline: 19061562]
25. McKeown DJ. Impact of minimally invasive aesthetic procedures on the psychological and social dimensions of health. Plast Reconstr Surg Glob Open 2021 Apr 28;9(4):e3578 [FREE Full text] [doi: 10.1097/GOX.0000000000003578] [Medline: 33936919]
26. Negrin L. Cosmetic surgery and the eclipse of identity. Body Soc 2002 Dec;8(4):21-42. [doi: 10.1177/1357034X02008004002]
27. Lynch SM. Measurement and prediction of aging anxiety. Res Aging 2000 Sep;22(5):533-558. [doi: 10.1177/0164027500225004]
28. Ramírez L, Palacios‐Espinosa X. Stereotypes about old age, social support, aging anxiety and evaluations of one's own health. J Soc Issues 2016 Mar 09;72(1):47-68. [doi: 10.1111/josi.12155]
29. Allan LJ, Johnson JA. Undergraduate attitudes toward the elderly: the role of knowledge, contact and aging anxiety. Educ Gerontol 2008 Dec 12;35(1):1-14. [doi: 10.1080/03601270802299780]
30. Lasher KP, Faulkender PJ. Measurement of aging anxiety: development of the Anxiety about Aging Scale. Int J Aging Hum Dev 1993 Dec;37(4):247-259. [doi: 10.2190/1u69-9au2-v6lh-9y1l]
31. Benton JP, Christopher AN, Walter MI. Death anxiety as a function of aging anxiety. Death Stud 2007 Apr;31(4):337-350. [doi: 10.1080/07481180601187100] [Medline: 17378111]
32. Greenberg J, Pyszczynski T, Solomon S. The causes and consequences of a need for self-esteem: a terror management theory. In: Baumeister RF, editor. Public Self and Private Self. New York, NY: Springer; 1986.
33. Clarke LC. Beauty in later life: older women's perceptions of physical attractiveness. Can J Aging 2010 Mar 31;21(3):429-442. [doi: 10.1017/s0714980800001744]
34. Levy B. Stereotype embodiment: a psychosocial approach to aging. Curr Dir Psychol Sci 2009 Dec 01;18(6):332-336 [FREE Full text] [doi: 10.1111/j.1467-8721.2009.01662.x] [Medline: 20802838]
35. Wurm S, Tesch-Römer C, Tomasik MJ. Longitudinal findings on aging-related cognitions, control beliefs, and health in later life. J Gerontol B Psychol Sci Soc Sci 2007 May;62(3):P156-P164. [doi: 10.1093/geronb/62.3.p156] [Medline: 17507583]
36. Chang ES, Kannoth S, Levy S, Wang SY, Lee JE, Levy BR. Global reach of ageism on older persons' health: a systematic review. PLoS One 2020 Jan 15;15(1):e0220857 [FREE Full text] [doi: 10.1371/journal.pone.0220857] [Medline: 31940338]
37. Levy BR, Hausdorff JM, Hencke R, Wei JY. Reducing cardiovascular stress with positive self-stereotypes of aging. J Gerontol B Psychol Sci Soc Sci 2000 Jul;55(4):P205-P213. [doi: 10.1093/geronb/55.4.p205] [Medline: 11584877]
38. Levy BR, Slade MD, Kunkel SR, Kasl SV. Longevity increased by positive self-perceptions of aging. J Pers Soc Psychol 2002 Aug;83(2):261-270. [doi: 10.1037//0022-3514.83.2.261] [Medline: 12150226]
39. Herrick SS, Hallward L, Duncan LR. "This is just how I cope": an inductive thematic analysis of eating disorder recovery content created and shared on TikTok using #EDrecovery. Int J Eat Disord 2021 Apr;54(4):516-526 [FREE Full text] [doi: 10.1002/eat.23463] [Medline: 33382136]
40. Ng R, Indran N. Not too old for TikTok: how older adults are reframing aging. Gerontologist 2022 Sep 07;62(8):1207-1216. [doi: 10.1093/geront/gnac055] [Medline: 35512401]
41. Armat MR, Assarroudi A, Rad M, Sharifi H, Heydari A. Inductive and deductive: ambiguous labels in qualitative content analysis. Qual Rep 2018 Jan 26;23(1):219-221. [doi: 10.46743/2160-3715/2018.2872]
42. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8-14. [doi: 10.1016/j.npls.2016.01.001]
43. Tesch R. Qualitative Research: Analysis Types and Software. London, UK: Routledge; 1990.
44. Crosnoe R, Elder GHJ. Successful adaptation in the later years: a life course approach to aging. Soc Psychol Q 2002 Dec;65(4):309-328. [doi: 10.2307/3090105]
45. Gosselink CA, Cox DL, McClure SJ, De Jong ML. Ravishing or ravaged: women's relationships with women in the context of aging and Western beauty culture. Int J Aging Hum Dev 2008 Apr 07;66(4):307-327. [doi: 10.2190/ag.66.4.c]
46. Hurd LC. Older women's body image and embodied experience: an exploration. J Women Aging 2008 Oct 22;12(3-4):77-97. [doi: 10.1300/j074v12n03_06]
47. Tiggemann M. Body image across the adult life span: stability and change. Body Image 2004 Jan;1(1):29-41. [doi: 10.1016/s1740-1445(03)00002-0]
48. Liao T, Zhuoga C, Chen X. Contact with grandparents and young people's explicit and implicit attitudes toward older adults. BMC Psychol 2023 Sep 26;11(1):289 [FREE Full text] [doi: 10.1186/s40359-023-01344-7] [Medline: 37752575]
49. Ng R, Indran N. Reframing aging: foregrounding familial and occupational roles of older adults is linked to decreased ageism over two centuries. J Aging Soc Policy 2024 Sep 02;36(5):749-766 [FREE Full text] [doi: 10.1080/08959420.2023.2238538] [Medline: 37534573]
50. Augustin F, Freshman B. The effects of service-learning on college students' attitudes toward older adults. Gerontol Geriatr Educ 2016;37(2):123-144 [FREE Full text] [doi: 10.1080/02701960.2015.1079705] [Medline: 26679428]

XSL•FO
RenderX

51. Pecchioni LL, Croghan JM. Young adults' stereotypes of older adults with their grandparents as the targets. J Commun 2006 Jan 10;52(4):715-730. [doi: 10.1111/j.1460-2466.2002.tb02570.x]
52. Hess A. The ever-changing business of ‘anti-aging’. The New York Times Magazine. 2017 Sep 12. URL: https://www.nytimes.com/2017/09/12/magazine/the-ever-changing-business-of-anti-aging.html [accessed 2022-07-19]
53. Verhaeghen P, Hertzog C. The Oxford Handbook of Emotion, Social Cognition, and Problem Solving in Adulthood. Oxford, UK: Oxford University Press; 2014.
54. Chen Y, Peng Y, Fang P. Emotional intelligence mediates the relationship between age and subjective well-being. Int J Aging Hum Dev 2016 Jul;83(2):91-107 [FREE Full text] [doi: 10.1177/0091415016648705] [Medline: 27199490]
55. Davidovic M, Erceg P, Trailov D, Djurica S, Milosevic D, Stevic R. The privilege to be old. Gerontology 2003;49(5):335-339. [doi: 10.1159/000071716] [Medline: 12920355]
56. Jenull B, Mayer C, Awenius L, Bostjancic G, Kollitsch K, Knobel P. Accepting the aging body: protective factors of body acceptance in persons of advanced age. Gerontol Geriatr Med 2023 Jan 13;9:23337214221150067 [FREE Full text] [doi: 10.1177/23337214221150067] [Medline: 36655235]
57. Ng R, Indran N. Reframing aging during COVID-19: familial role-based framing of older adults linked to decreased ageism. J Am Geriatr Soc 2022 Jan;70(1):60-66. [doi: 10.1111/jgs.17532] [Medline: 34674224]
58. Ng R, Indran N. Role-based framing of older adults linked to decreased ageism over 210 years: evidence from a 600-million-word historical corpus. Gerontologist 2022 Apr 20;62(4):589-597 [FREE Full text] [doi: 10.1093/geront/gnab108] [Medline: 34323967]
59. Fosch-Villaronga E, Poulsen A, Søraa RA, Custers BH. A little bird told me your gender: gender inferences in social media. Inf Process Manag 2021 May;58(3):102541. [doi: 10.1016/j.ipm.2021.102541]
60. Krekula C, Nikander P, Wili?ska M. Multiple marginalizations based on age: gendered ageism and beyond. In: Ayalon L, Tesch-Römer C, editors. Contemporary Perspectives on Ageism. Cham, Switzerland: Springer; May 23, 2018.
61. Halliwell E, Dittmar H. A qualitative investigation of women's and men's body image concerns and their attitudes toward aging. Sex Roles 2003 Dec;49:675-684. [doi: 10.1023/B:SERS.0000003137.71080.97]
62. Ng R, Lim-Soh JW. Ageism linked to culture, not demographics: evidence from an 8-billion-word corpus across 20 countries. J Gerontol B Psychol Sci Soc Sci 2021 Oct 30;76(9):1791-1798 [FREE Full text] [doi: 10.1093/geronb/gbaa181] [Medline: 33099600]
63. Ng R. Societal age stereotypes in the U.S. and U.K. from a media database of 1.1 billion words. Int J Environ Res Public Health 2021 Aug 21;18(16):8822 [FREE Full text] [doi: 10.3390/ijerph18168822] [Medline: 34444578]
64. Ng R, Indran N, Yang W. Portrayals of older adults in over 3000 films around the world. J Am Geriatr Soc 2023 Sep;71(9):2726-2735. [doi: 10.1111/jgs.18400] [Medline: 37314013]
65. Giest S, Ng R. Big Data Applications in Governance and Policy. PaG 2018 Nov 21;6(4):1-4. [doi: 10.17645/pag.v6i4.1810]
66. TikTok. URL: https://www.tiktok.com/ [accessed 2026-07-29]

Edited by J Wang; submitted 02.Jan.2025; peer-reviewed by A Zheluk, S Just; comments to author 20.Mar.2025; revised version received 04.Jul.2025; accepted 26.Sep.2025; published 18.Aug.2026.

Please cite as:
Ng R, Indran N
Aging Like Milk or Fine Wine: Content Analysis of Reactions to One’s Future Self on TikTok
JMIR Aging 2026;9:e70794
URL: https://aging.jmir.org/2026/1/e70794
doi:10.2196/70794
PMID:

©Reuben Ng, Nicole Indran. Originally published in JMIR Aging (https://aging.jmir.org), 18.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Factors Influencing Older Adults' Sustained Engagement in Digital Psychological Interventions in Resource-Limited Settings: Qualitative Study Guided by the Social Ecological Model

Jingying Wang[1,2], MM; Yating Luo[2], PhD; Sipei Jia[2], PhD; Wenling Yang[3], BS; Dan Liu[1], MM; Jiarong Li[1,2], MM; Gang Gan[1], MM; Yi Zhou[1], MM; Sha Wang[1], PhD; Jianfei Xie[1,4], PhD

[1]Nursing Department, The Third Xiangya Hospital, Central South University, Changsha, China

[2]Xiangya Nursing School, Central South University, Changsha, China

[3]Mental Health Department, Huayuan People's Hospital, Xiangxi Autonomous Prefecture, China

[4]Research Centre for Clinical Nursing Safety Management, Central South University, Changsha, China

Corresponding Author:
Jianfei Xie, PhD
Nursing Department, The Third Xiangya Hospital, Central South University
No.138, Tongzipo Road, Yuelu District
Changsha, 410013
China
Phone: 86 1 387 496 2861
Email: xiejianfei@csu.edu.cn

Abstract

Background: Digital behavioral activation therapy (DBAT) is effective in reducing late-life depression (LLD). However, older adults in resource-limited settings often show low adherence and retention. This pattern is largely driven by the digital divide. Digital inclusion offers a pathway to address this gap by recognizing older adults as active participants rather than passive recipients. Despite this potential, limited research has examined low engagement from the perspective of older adults' lived experiences. This gap restricts understanding of the complex factors that shape engagement in digital psychological interventions.

Objective: This study aimed to systematically explore, through the lens of digital inclusion and guided by the social ecological model, the multilevel factors influencing older adults' sustained engagement in digital psychological interventions within resource-limited settings, specifically focusing on the lived experiences of retained users who completed the full program.

Methods: A descriptive phenomenological design was used. Semistructured interviews were conducted with older adults in Western Hunan, China. Data were transcribed using machine-assisted methods and verified manually. Data were analyzed using the Colaizzi phenomenological method combined with thematic and framework analysis guided by the social ecological model. Two researchers independently coded the data and refined themes through discussion.

Results: Nineteen older adults with LLD from resource-limited settings were included. Engagement in DBAT was shaped by factors across multiple levels. At the individual level, factors included health needs, learning motivation, physiological and cognitive limitations, and low digital literacy. At the interpersonal level, factors included intergenerational support, emotional burden, established social networks, and digital exclusion. At the society level, available devices and networks sometimes hindered engagement due to poor age-friendly design and unstable connectivity. Personal aspirations for digital participation often conflicted with social resistance. At the policy level, financial burden, standardized policy approaches, and unmet individual needs limited engagement. These factors interacted to influence engagement behavior.

Conclusions: This study identifies multilevel factors shaping older adults' sustained engagement in digital psychological interventions within resource-limited settings, reflecting the experiences of retained users. The findings advocate for integrated, context-sensitive strategies grounded in a digital inclusion framework. Crucially, nurses play a key role in promoting technology access, improving digital literacy, supporting social participation, and advocating for equitable services. These efforts can enhance accessibility and adherence, ultimately fostering more equitable and sustainable digital mental health care for vulnerable aging populations.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400091193; https://tinyurl.com/y5e6wt7p

XSL•FO
RenderX

(JMIR Aging 2026;9:e84350) doi:10.2196/84350

KEYWORDS

older adults; digital inclusion; social ecological model; resource-limited settings; digital divide; late-life depression; behavioral activation; mobile phone

Introduction

Background

The rapid growth of the aging population has made mental health in later life a major global public health concern [1]. The World Health Organization (WHO) projects that by 2050 the number of people aged 60 years and older will reach 2.1 billion, with a large proportion living in resource-limited settings such as low-income countries, rural areas, and remote regions [2]. Depression is one of the most common mental health conditions in this population, with an estimated prevalence of 31.8% [3]. The burden is even greater in resource-limited settings [4]. Late-life depression (LLD) is strongly associated with poor health outcomes, cognitive decline, reduced quality of life [5], and increased risk of suicide [6]. It also places a substantial burden on families, communities, and health systems. However, these settings often face severe shortages of mental health professionals, unequal distribution of health resources, and limited access to psychological care. These constraints make timely and effective management of depression difficult [7].

The rapid development of digital health technologies offers new opportunities to address these service gaps [1,8]. Digital psychological interventions, including digital behavioral activation therapy (DBAT), have shown effectiveness comparable to face-to-face treatments in reducing depressive symptoms. They also offer advantages such as lower cost, greater flexibility, and improved accessibility [9-11]. Behavioral activation (BA) was first introduced by Ferster in 1973 [12]. It is an evidence-based approach that promotes engagement in meaningful activities to increase positive reinforcement and improve mood [13,14]. This study selected BA for digital delivery because its structure is simple and suitable for self-guided use [10]. It requires limited clinical resources and has demonstrated effectiveness for LLD [11]. Compared with other digital therapies, its action-oriented design fits well with the daily routines and capacities of older adults in underserved settings. Recent advances in digital delivery have further expanded its reach beyond geographic and service barriers, making it a promising option for this population. Despite these advantages, the full potential of digital mental health interventions has not yet been realized. A key barrier is the digital divide.

The digital divide includes more than unequal access to technology. It also involves limited digital literacy, low motivation to use technology, inadequate infrastructure, financial constraints, and lack of intergenerational support [8,15,16]. Older adults in resource-limited settings are particularly affected [17]. This divide contributes to low engagement in digital psychological interventions, even though engagement is essential for both effectiveness and efficient use of resources [18]. Most existing studies have focused on feasibility and clinical outcomes. They often rely on basic indicators such as adherence and retention rates. For example, a study in Hong Kong reported retention rates of 54.8% at 1 month and 49.9% at 3 months after the intervention [19]. A randomized controlled trial in Brazil found an attrition rate of 19.4% at 5 months, which eliminated differences between treatment groups [10]. These findings suggest that although digital psychological interventions are effective in theory, their real-world impact is limited by low engagement. This leads to underuse of already scarce health resources [18,20].

While the digital divide literature traditionally emphasizes a deficit model focused on lacking access and technical skills, this study uses the digital inclusion framework to shift the analytic focus toward agency, relational dynamics, and equitable participation [21-23]. Rather than viewing nonuse simply as a technical failure to be corrected, the digital inclusion lens allows us to examine how older adults actively negotiate meaning, manage emotional burdens, and navigate structural constraints to sustain their engagement. This framework moves beyond merely cataloging material barriers, providing a critical lens to understand the complex psychosocial processes required to achieve meaningful digital citizenship in resource-limited settings. However, the lived experiences of navigating these processes remain underexplored from the perspective of older adults themselves. In resource-limited settings, there is a lack of theory-driven qualitative research that explains the psychological, experiential, and contextual factors shaping engagement in digital interventions. This gap limits the development of targeted and sustainable strategies.

To address this gap, this study used the social ecological model as a guiding framework. The social ecological model is particularly suitable for this context, as it posits that health behaviors are shaped by the interplay of individual, interpersonal, societal, and policy factors, which aligns directly with the multilevel nature of the digital divide and the concept of digital inclusion [24,25]. In the context of digital interventions, it provides a structured way to examine digital inclusion as a complex and multidimensional process. It considers individual capacity, interpersonal support, societal resources, and policy environments. In resource-limited settings, these levels are closely interconnected and jointly influence engagement in digital care.

Research Aim

This study used a descriptive phenomenological approach to explore the lived experiences of older adults participating in DBAT. The aim was to identify and understand the multilevel factors that influence sustained engagement among retained users, guided by the social ecological model and the concept of digital inclusion. By focusing on the lived experiences of older adults, this study moves beyond a descriptive inventory of structural barriers to advance the conceptual understanding of digital inclusion. Specifically, it refines the social ecological

XSL•FO
RenderX

model by conceptualizing digital engagement as a dynamic process of navigating cross-level paradoxical tensions, such as the conflict between macropolicy pushes and microsocial resistance, and the simultaneous empowering and burdening effects of intergenerational support. Highlighting these psychosocial interactions offers a deeper, theory-informed basis for improving digital mental health interventions. The goal is to support interventions that are more inclusive, context-sensitive, and person-centered rather than to evaluate a specific implementation strategy. For nursing practice, these insights can guide improvements in intervention design, enhance accessibility, and support sustained engagement. More broadly, this study contributes both theoretical and practical evidence to inform digital health strategies in the context of population aging. It supports efforts to promote health equity and aligns with the goal of achieving accessible digital health services for all. Ultimately, it aims to improve mental health and well-being among older adults in resource-limited settings.

Methods

Overview

This study adopted a qualitative design grounded in a descriptive phenomenological approach [26]. The aim was to explore factors that influence older adults' engagement with DBAT in resource-limited settings from their own perspectives. This approach focuses on understanding lived experience while minimizing researcher assumptions. Semistructured interviews were conducted to capture participants' perceptions, emotional responses, and cognitive experiences during their engagement with digital mental health programs. This approach allowed the identification of barriers and facilitators related to technology access, interface use, and understanding of intervention content. Phenomenology was well suited to examine complex psychosocial processes, including perceived difficulty with technology, management of anxiety related to digital use, and gradual adaptation with support from family and social networks. The study followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) to ensure transparency and rigor [27]. An initial interview guide was developed based on a literature review and refined through pilot interviews. All research staff completed standardized training before data collection. Inclusion and exclusion criteria were clearly defined. Before each interview, participants were informed about the study purpose and procedures, and consent for audio recording was obtained. All interviews were transcribed verbatim and coded for analysis. The analysis combined thematic analysis with framework analysis. Thematic analysis identified key patterns and meanings in participants' experiences. Framework analysis, guided by the social ecological model, organized findings across 4 levels, including individual, interpersonal, society, and policy. This approach preserved participants' lived experiences while linking them to a structured theoretical framework.

Intervention Program

The intervention was a DBAT program delivered through a digital platform and supported by research nurses trained in geriatric psychological care. The platform was designed to meet the needs of older adults in resource-limited settings, with features such as large text, voice prompts, multimedia content, and low data requirements. The program included mood and activity monitoring, identification of pleasant activities, individualized goal setting, gradual engagement in meaningful activities, and strategies to reduce avoidance and negative behaviors [13,14]. Each session lasted 45 to 60 minutes and was conducted once per week for 8 weeks. Participants accessed the program through a mobile-friendly website and completed weekly modules independently. These modules included educational videos, monitoring tools, and interactive exercises. During weekly sessions, research nurses guided participants through the content, explained instructions, demonstrated navigation, and reviewed progress. Outside these sessions, participants used the platform independently. Technical support was available through phone or text message when needed. The intervention therefore combined self-guided digital learning with structured human support. All sessions were delivered by trained research nurses with experience in geriatric psychological care.

Setting

This study was conducted at Xiangxi Autonomous Prefecture People's Hospital in Hunan Province, China, which stands as the sole Grade III Class A hospital in Western Hunan. The region is economically underdeveloped and has a high proportion of older adults, accounting for 19.88% (472,419/2,376,353) of the population. Mental health services in this area are limited and fall below the provincial average, making it a typical resource-limited setting [28].

Participants and Recruitment

Participants were recruited from the intervention group of a larger clinical trial evaluating the DBAT program. All individuals who met the eligibility criteria for this qualitative study were invited to participate. Purposive sampling was used to achieve variation in age, gender, education, and income. Participants were approached directly by research nurses involved in the intervention. All invited participants agreed to take part. Sample size was determined by data saturation. Data collection and analysis were conducted concurrently. After 16 interviews, no new themes emerged. Three additional interviews confirmed saturation, resulting in a final sample of 19 participants. All participants had completed the full 8-week intervention before the interviews. Their accounts therefore reflect the full engagement process from initial participation to completion.

Participants were eligible if they were aged 60 years or older, had lived in the region for at least two years, had mild to moderate depressive symptoms defined by a score of 5 to 11 on the Geriatric Depression Scale with 15 items (GDS-15), could communicate in basic Mandarin, and were able to use a smartphone independently. Some participants used a family member's device to access the program. Participants were excluded if they had severe physical illness that limited participation, major cognitive or psychiatric disorders, or severe visual or hearing impairments that interfered with digital use.

XSL•FO
RenderX

Data Collection

Data were collected through one-on-one semistructured interviews that explored factors influencing engagement in DBAT across multiple levels. The interview guide was informed by prior literature, field experience, and the social ecological model. It covered individual, interpersonal, society, and policy domains [24,25]. Pilot interviews were conducted to refine the guide (see Textbox 1 for interview outline).

Textbox 1. Interview outline.

Individual level

- Could you share what motivated you to begin engaging in the digital behavioral activation therapy (DBAT) program?
- During your engagement in DBAT, what aspects of the process have been most challenging for you? How do you usually manage or adapt to these challenges?

Interpersonal level

- Are your family members, friends, or partner aware of your engagement in DBAT? In what ways, if any, have they provided support?
- When you need assistance from people around you, how do you usually seek help? How does this influence your willingness to continue engaging in DBAT?

Society level

- How common is it for older adults in your community to engage in similar online psychological health services?
- How would you describe the overall attitude in your community toward older adults using online psychological health services?

Policy level

- Are you aware of any policies or community initiatives that have supported or promoted similar online psychological health services?
- If this type of program were to continue in the future, what additional forms of support would make it easier for you to engage?

All interviews were conducted by the first author, a female graduate student trained in qualitative research and geriatric psychology. Interviews took place in a private office to ensure comfort and confidentiality. Written informed consent was obtained before each interview. Each interview lasted between 30 and 60 minutes. A research assistant recorded nonverbal observations. A local health care worker assisted with translation when regional dialects were used. After each interview, the research team recorded reflective notes. Audio recordings were transcribed within 24 hours and checked by 2 researchers for accuracy. When needed, participants were contacted to clarify unclear statements. Field notes and reflective journals were also transcribed. Data collection continued until saturation was reached [29].

Analysis

Data were analyzed using the Colaizzi phenomenological method combined with thematic and framework analysis [30,31]. The social ecological model guided the overall analytical process. To explicitly distinguish between inductive findings and theoretical interpretation, a 3-stage analytical process was used. First, initial themes reflecting participants' raw lived experiences emerged inductively through open coding. Second, these empirically derived themes were systematically organized within the predefined individual, interpersonal, society, and policy levels of the social ecological model framework. Third, rather than treating these categories as isolated silos, the research team conducted a broader theoretical interpretation to identify how these themes dynamically interacted across ecological levels. Constant comparison was used to refine this cross-level thematic structure. Data management was supported by NVivo 12, and final themes were established through team consensus.

Rigor

Trustworthiness was ensured using the criteria of credibility, transferability, dependability, and confirmability [32]. Purposeful sampling and data saturation supported the adequacy of the sample. Detailed transcripts, field notes, and reflective journals preserved contextual depth. Two researchers independently coded the data, and differences were resolved through discussion. The research process was documented in detail to allow transparency and potential replication.

Ethical Considerations

Ethical approval was obtained from the Institutional Review Board of the Third Xiangya Hospital, Central South University (Approval No.K24901; October 18, 2024). Written informed consent was obtained from all participants. They were informed about the study purpose, procedures, and potential risks. The study followed the principles of the Declaration of Helsinki and prioritized participants' psychological well-being. A brief mental health check was conducted before each interview. If distress occurred, interviews were paused, and supportive communication was provided. Participants who needed additional support were referred to local counseling services or community resources. At the end of the study, all participants received information on available mental health resources to support their continued well-being.

Results

Participant Characteristics

A total of 19 older adults from resource-limited settings participated in the interviews. The sample included 10 men and 9 women. Most participants were younger than 80 years.

XSL•FO
RenderX

Educational attainment was generally low, with most participants having a primary school education or less. The majority were married, while a small number were unmarried, divorced, or widowed. Monthly income (MI) levels were low across the sample. Most participants reported having at least one physical illness, and about three-quarters were covered by basic health insurance (HI). Detailed demographic information is presented in Table 1.

Table 1. Participant characteristics.

ID	Gender	Age (years)	Education	MS[a]	MI[b]	PD[c]	HI[d]
01	Female	64	Junior high school	Married	1000-3000	Yes	Yes
02	Female	72	Junior high school	Married	＞5000	Yes	Yes
03	Female	78	Senior high school	Married	＞5000	Yes	Yes
04	Male	61	Above senior high school	Married	＞5000	Yes	Yes
05	Female	82	Primary school or below	Divorced	1000-3000	Yes	No
06	Male	65	Primary school or below	Married	＜1000	Yes	Yes
07	Male	74	Primary school or below	Married	3000-5000	Yes	Yes
08	Male	60	Primary school or below	Single	＜1000	Yes	No
09	Male	79	Above senior high school	Married	＞5000	Yes	Yes
10	Female	60	Primary school or below	Married	1000-3000	No	Yes
11	Female	73	Primary school or below	Married	＜1000	Yes	Yes
12	Male	75	Primary school or below	Married	＜1000	Yes	No
13	Male	78	Primary school or below	Widowed	1000-3000	Yes	Yes
14	Male	64	Primary school or below	Single	1000-3000	Yes	Yes
15	Male	67	Junior high school	Married	1000-3000	Yes	No
16	Male	66	Primary school or below	Married	3000-5000	No	Yes
17	Female	79	Primary school or below	Married	1000-3000	Yes	Yes
18	Female	83	Primary school or below	Married	1000-3000	Yes	Yes
19	Female	72	Junior high school	Married	1000-3000	No	No

[a]MS: marital status.

[b]MI: monthly income (CNY). A currency exchange rate of CNY 1 = US $0.15 was applicable.

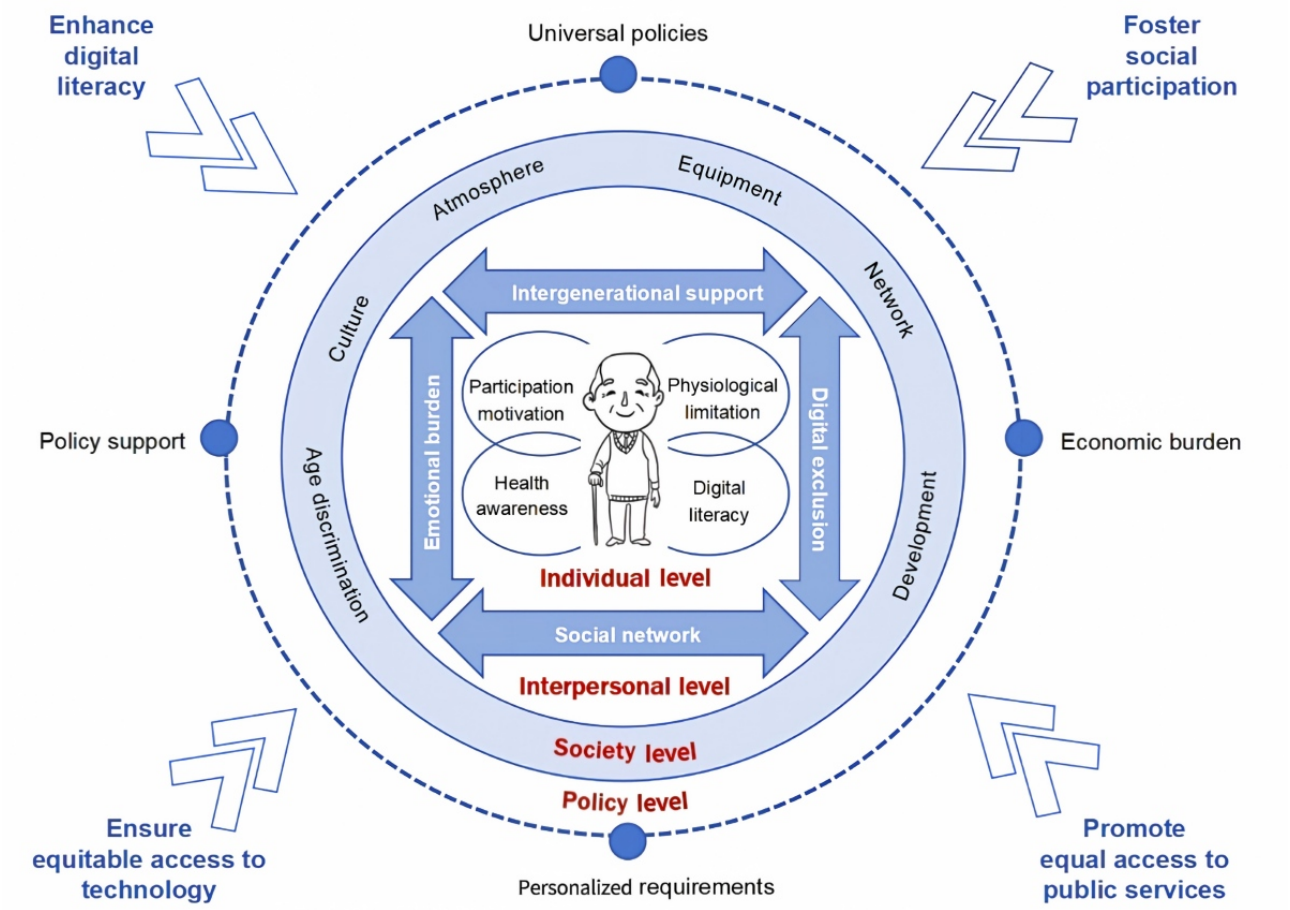
[c]PD: physical disease.

[d]HI: health insurance.

Guided by the social ecological model, the inductively derived themes shaping sustained engagement in DBAT were mapped onto the individual, interpersonal, society, and policy levels. While presented sequentially below for clarity, participants' narratives consistently revealed that these factors did not operate in isolation but dynamically interacted across levels to shape behavioral outcomes (Figure 1).

XSL•FO
RenderX

Figure 1. Factors at 4 levels influencing older adults' engagement.



Individual-Level Factors

Health Needs Versus Physiological and Cognitive Constraints

Fourteen participants reported a strong desire to improve their mental health through DBAT. However, this motivation was often constrained by sensory decline, memory impairment, and low confidence in using digital tools. The mismatch between strong health needs and limited functional capacity emerged as a key barrier to sustained engagement. Participants showed intrinsic motivation driven by the need for social connection and avoidance of marginalization.

> *Seeing other older people actively chatting in the course video's comment section made me anxious. If only I could learn to type and leave a message, it would ease my mind a lot.* [ID 01, aged 64 years]

Three participants noted that course videos and activities reduced loneliness and increased enjoyment: "Watching course videos in my spare time makes life less monotonous and gives me something to look forward to" (ID 19, aged 72 years). Another stated, "Doing exercises with the videos every day helps me stay healthy, and I feel much more positive" (ID 03, aged 78 years). Despite this motivation, physiological constraints were prominent. Nine participants reported difficulties.

> *My eyesight is failing, the small text on the screen looks like a foreign language, and I often press the wrong buttons. At this age, my physical functions have declined, and I can barely handle a smartphone.* [ID 16, aged 66 years]

Cognitive barriers further limited engagement: "I always feel like my brain is too slow for this. These new things are too complicated, and just looking at them makes me anxious" (ID 02, aged 72 years). These perceptions reduced confidence and persistence, reinforcing digital exclusion.

Learning Motivation Versus Deficits in Digital Literacy

Seven participants expressed a strong willingness to learn smartphone use, motivated by mental health needs, social connectedness, and independent information access. However, prevalent digital literacy deficits—such as difficulty recognizing interface elements, poor operational execution, and an inability to troubleshoot errors—created a persistent gap between their intentions and functional abilities.

> *I heard that watching videos can teach psychological health skills, and I can chat with others in the comments. I'm curious about it. If I could learn this, I'd finally have something interesting to do, instead of feeling every day is hard to get through.* [ID 14, aged 64 years]

Another interviewee noted, "I know it's difficult, but I'm willing to learn slowly so I can handle my depressive moods by myself and not worry my family" (ID 17, aged 79 years). Despite this, the lack of digital literacy significantly hindered progress. Many participants felt overwhelmed when first using digital devices.

XSL•FO
RenderX

> *The course website has so many functions that confuse me. I can't tell what they all mean. Even answering calls from the nurse or exiting the website requires someone to teach me several times. I feel disconnected from this era.* [ID 03, aged 78 years]

Operational errors were common. Another participant described, "My hands are clumsy. I often press the wrong buttons. I try to send a message but accidentally exit the course, and I don't know how to go back. I just get anxious" (ID 06, aged 65 years). These experiences reflect deficits in functional understanding and error recovery, limiting effective engagement.

Interpersonal-Level Factors

Intergenerational Digital Support and Emotional Burden

Six participants identified intergenerational support as key to initial engagement. Assistance from children or grandchildren facilitated access to digital interventions. However, a similar number reported feelings of burden. Limited time or patience from younger family members often led to negative emotions and reduced engagement. Thirteen participants emphasized the emotional benefits of such support.

> *The happiest moment was when my granddaughter taught me how to use the course website. Seeing her lovely smile on the screen made me feel truly warm inside.* [ID 04, aged 61 years]

> *My grandson kept praising me when I learned how to play the videos. It made me so happy. The process brought us closer, and I felt as if he was right there with me.* [ID 10, aged 60 years]

At the same time, some participants felt pressure due to slow learning. "My child taught me several times, but I still couldn't remember. I felt terribly guilty, as if I were bothering them. When I saw their frown, I just wanted to give up" (ID 01, aged 64 years). This emotional tension highlights that intergenerational digital support is not an unconditional resource. Its effectiveness largely depends on the investment and patience of younger family members.

> *They are busy with work, yet they have to spend time teaching me. I feel bad about it. I want to learn quickly, but these functions are really too complicated.* [ID 02, aged 72 years]

This dynamic shows that intergenerational support can both enable and hinder engagement. When support is limited, older adults may withdraw due to self-doubt. Paradoxically, this form of digital caregiving could exacerbate feelings of anxiety and inadequacy, further widening the psychological gap in their digital engagement.

Entrenched Social Networks and Latent Digital Exclusion

More than half of the participants relied on face-to-face interaction and showed resistance to digital communication. This reduced motivation to engage in online interventions. Economic constraints and weak infrastructure further limited access. Six participants preferred in-person interaction.

> *I'm used to talking face-to-face. Sitting together makes me feel secure. Online chatting makes me worry no one will reply, and then I feel neglected.* [ID 05, aged 82 years]

This attachment to traditional forms of socialization was reinforced by the similar habits of peers. One participant said, "Few people around me use this, so I don't want to either. I feel we don't share common topics in that area" (ID 14, aged 64 years). This makes it difficult to integrate digital interaction into their daily lives, thus reducing the intrinsic motivation for online intervention. Material constraints also contributed to exclusion. Many older adults lacked the financial means to own smart devices, while remote geographical locations often resulted in weak network coverage and unstable connections, further impeding use even among those with devices.

> *Smartphones are too expensive for my family. My grandson bought me one a few years ago, but the signal in our village is poor, and the videos kept cutting off. I tried using it for a while, but I felt both disappointed and frustrated, realizing how far behind we are compared to city people.* [ID 17, aged 79 years]

Society-Level Factors

The Paradox of Device and Network Availability

Participants reported mixed experiences with devices and network access. When systems were simple and adapted to their needs, engagement improved. When complex or unstable, they created frustration.

> *These videos do help me maintain my mental health, but the icons on the website are so small, and the steps are too complicated. I tap around for a long time and still can't find the assignment section. Sometimes I just lose patience and confidence to keep trying.* [ID 08, aged 60 years]

Conversely, design improvements increased usability.

> *At first, I found it troublesome. Later, when they changed to larger fonts and added arrows and voice prompts, I felt it was much friendlier for older people. These improvements are really necessary. If all programs were like this, we'd be more willing to learn.* [ID 17, aged 79 years]

Yet, network instability remained a major barrier.

> *The internet in our village is unstable and keeps cutting out. Videos keep buffering, and sentences are incomplete. Sometimes I click on the next page, and it just keeps loading, loading… in the end, nothing shows up. I sigh and quit; it feels like a waste of time.* [ID 12, aged 75 years]

These findings show that technology can either enable or hinder engagement depending on design and infrastructure.

Aspirations for Progress Versus Social Resistance

Participants described a tension between the desire to adopt digital health and a resistant social environment. Public messages promoted digital health as part of modern life, which

XSL•FO
RenderX

encouraged some participants to try new technologies. However, local gossip under the old locust tree and the recurring household refrain of "you cannot do this" often acted as an invisible web, dampening emerging intentions to engage in online psychological care.

> *The TV keeps saying how advanced digital health care is, that everyone in the new era sees doctors online. I'm really afraid of being left behind, unable to adapt, or even laughed at by others. So I wanted to try this course to keep up, but I also fear I won't learn it. I want to, but I don't dare—it's very conflicting inside.* [ID 18, aged 83 years]

Even when individuals mustered the courage to use their devices in public, peers gathered in community spaces often discouraged them:

> *They say looking at the phone strains your eyes and wastes time you could spend farming. When I mention mental health, they just shake their heads and say it's useless. They trust the village doctor more than any online program.* [ID 18, aged 83 years]

Family-level discouragement was also reported.

> *What hurts the most is my own son often discourages me. He says the course videos are too complicated, that I might break the phone. If I press the wrong button, I could leak personal information or even get scammed. I feel so wronged. I just want to learn about mental health online, but every time I bring it up, I'm shut down.* [ID 11, aged 73 years]

This social tension suspended many older adults in a state of wanting to try but not daring to, leaving online psychological interventions in a fragile limbo between aspiration and hesitation.

Policy-Level Factors

Policy Orientation Versus Economic Burden

Twelve participants were aware of national digital health policies and recognized their intent. However, most reported financial barriers related to device purchase and internet costs. Limited subsidies and complex procedures reduced policy impact. Seven participants endorsed the overall direction of national efforts, perceiving them as a sign of progress.

> *The television says the government encourages rural older adults to use smartphones for health consultations and social connections. I have thought about trying it—at my age, fewer trips to the hospital would be nice. But in the end, these policies rarely reach us.* [ID 18, aged 83 years]

Such ambivalence of hope mingled with skepticism was common. Another participant noted, "The village loudspeaker promotes digital health every day, and it makes me want to try. But we see no real action. It still feels empty" (ID 04, aged 61 years). Direct economic barriers were identified as the primary obstacle to policy implementation. Five participants cited the prohibitive cost of smart devices.

> *A smartphone easily costs over a thousand yuan, more than I spend in a year. I can't afford it, and even if I bought one, I wouldn't dare use it for fear of breaking it.* [ID 15, aged 67 years]

Internet costs were also burdensome. An interviewee stated, "Twenty yuan a month for data may not sound like much, but for someone like me who relies on a few acres of farmland to eat, it's unbearable" (ID 19, aged 72 years). Subsidies, where available, were inconsistently distributed and procedurally cumbersome. One participant lamented, "I heard the government is giving subsidies for digital health, but I haven't seen a penny" (ID 08, aged 60 years). Such administrative bottlenecks and local implementation gaps undermined the intended mitigating effects of these policies, reinforcing inequality in access to digital resources and limiting the feasibility of leveraging digital tools for psychological well-being.

Universal Policies Versus Personalized Needs

Four participants reported that current policies rely on standardized approaches that do not meet individual needs. These include uniform device distribution and centralized training. Such strategies often fail to consider variations in physical ability, digital skills, and geographic access. Two participants explicitly criticized the lack of age-friendly design considerations. One remarked about the distributed devices:

> *The policy just says they'll give us smartphones, but they don't think about our shaky hands and blurry eyes. The icons are tiny like ants, and one wrong press takes me back to the start. It makes me want to throw the phone into the ditch.* [ID 10, aged 60 years]

Thus, without adaptive design, even hardware provision fails to meaningfully enhance digital engagement. Access barriers also persisted.

> *They hold classes at the training center every week, but my legs aren't good. It takes two buses and half a day to get there. The round-trip fare costs half a month of my blood pressure medicine. So I'd rather do nothing.* [ID 04, aged 61 years]

In contrast, participants commonly expressed a strong desire for precise and personalized policy measures. The ideal policy, in their view, would not only provide age-appropriate devices but also include ongoing, context-sensitive guidance and sustained financial support.

> *If someone could come to my home, set me up with a senior-friendly phone, teach me step by step, and give me enough data every month, I'd definitely watch the course videos every day and join the online chats. I wouldn't feel lonely anymore.* [ID 17, aged 79 years]

This reflects the need for policies that are tailored, accessible, and responsive to individual circumstances.

Cross-Level Interactions

Participants' narratives consistently revealed that factors across the 4 ecological levels did not operate independently but dynamically interacted to shape sustained engagement. A barrier at 1 level frequently compounded difficulties at another. For instance, as previously noted in the participants' experiences,

XSL•FO
RenderX

macrolevel policy initiatives providing standardized devices often clashed with individual physiological limitations (ID 10, aged 60 years). This policy-individual friction subsequently exacerbated emotional frustration and increased the burden on interpersonal assistance. Furthermore, robust individual learning motivation was frequently neutralized by intersecting societal and interpersonal barriers. This was evident when an older adult's intrinsic desire to engage with digital health was directly suppressed by family discouragement, which was in turn fueled by broader societal concerns about digital scams (ID 11, aged 73 years). These empirical observations illustrate that older adults did not experience these barriers in isolation. Rather, their sustained engagement was determined by continuously navigating intersecting vulnerabilities, where the failure to address a structural barrier could easily neutralize supportive resources at the microlevel.

Discussion

Principal Findings

To our knowledge, this study is among the first to apply the social ecological model to systematically examine factors influencing older adults' engagement in DBAT within resource-limited settings. While the identified determinants—such as limited digital literacy and infrastructural barriers—align with well-established digital divide literature, this study advances existing conceptual frameworks by revealing the cross-level paradoxical tensions that characterize digital inclusion. Rather than viewing digital exclusion as a static lack of resources, our findings conceptually refine digital engagement in resource-limited settings as a highly relational and emotionally fraught process. For instance, intergenerational support acts as a double-edged sword, providing essential technical access while simultaneously generating emotional burden and guilt that can deter sustained use. Similarly, older adults are often caught in a limbo between macrolevel aspirations for digital progress and rigid microlevel social resistance. These findings are consistent with previous large-scale research showing that single-level interventions are insufficient to sustain engagement [33]. They support the core principles of digital inclusion, which emphasize equitable access, improved digital literacy, increased social participation, and fair access to services. Effective strategies require coordinated action across levels. Interventions should combine individual capacity building, family and peer support, community-based delivery, and responsive policy design. Such integrated efforts can improve intervention outcomes, support efficient use of resources, and promote equitable and sustainable digital health development in line with global health priorities [34].

Translating Barriers Into Digital Inclusion Strategies

The analytic value of applying a digital inclusion framework, rather than a conventional digital divide lens, lies in shifting the focus from deficit correction to relational empowerment. A standard barrier analysis might conclude that simply providing devices and basic training resolves digital exclusion. In contrast, evaluating our findings through the digital inclusion framework reveals that overcoming material deficits is insufficient without addressing the psychosocial and relational dimensions of engagement. For instance, while providing a smartphone addresses a structural deficit at the policy level, a digital inclusion perspective highlights that true participation requires mitigating the emotional guilt associated with intergenerational training and countering entrenched community stigma. Consequently, translating these findings into inclusion strategies demands interventions that move beyond mere hardware provision to cultivate digital autonomy, emotional safety, and meaningful social embeddedness across all ecological levels.

In resource-limited settings, older adults often show low adherence to digital interventions due to physical decline, cognitive limitations, and limited health literacy [8,20]. Interventions should therefore extend beyond technical adaptation. Health education can increase awareness of mental health and the value of digital care. Structured digital literacy training with large text, voice guidance, and gradual progression can reduce operational difficulty [35]. Leite et al [36] reported that such improvements foster technological inclusion and subjective well-being among older adults. Education on privacy and data security is also important to build trust. In addition, helping older adults develop realistic expectations and critical awareness can reduce frustration and dropout. In nursing practice, BA strategies should be combined with continuous support, follow-up, and remote guidance. This approach allows for effective and accessible care. Evidence from a Chinese cohort demonstrated that bridging the digital divide not only improved intervention outcomes directly but also indirectly enhanced cognitive function and general health [37]. Training nurses in both psychological care and digital skills can support the implementation of these strategies.

At the interpersonal level, intergenerational digital support plays a dual role by both facilitating and constraining engagement. Family members often provide essential technical guidance and encouragement, which can improve participation in digital interventions. However, excessive reliance may lead to emotional burden, reduced autonomy, and avoidance of help-seeking among older adults [17,20]. This reflects the structural influence of family resources and education highlighted in prior studies [38,39]. Recent evidence indicates that internet altruistic behaviors, such as offering help and sharing information online, can enhance affective well-being [40]. These findings suggest that extending support beyond the family to broader digital networks may reduce emotional strain and improve engagement. Nurses should provide caregiver education that promotes supportive yet autonomy-preserving interactions. Encouragement and confidence building are essential, while assistance should remain limited to maintain self-efficacy [41]. As Qiu et al [42] noted, intergenerational digital gaps temporarily narrowed during the pandemic but lacked long-term sustainability measures. Thus, nurse-led peer support and digital navigator programs can expand nonfamilial support, foster mutual assistance, and strengthen digital social networks [35]. These strategies may reduce dependence on family members and promote more sustainable engagement.

Despite gradual improvements in digital infrastructure, older adults in resource-limited settings continue to face significant barriers. Limited access to devices, unstable internet connections, and inadequate maintenance reduce the usability

XSL•FO
RenderX

of digital psychological interventions [43]. In their review of older adults' digital engagement, Kebede et al [44] highlighted key intervention strategies, including training, equipment support, and social participation. In nursing practice, community-based models offer a practical solution. Digital health access points can be established in local centers where nurses and trained volunteers provide guidance and support. Regular feedback and program adjustment can further improve effectiveness. At a broader level, digital connectivity should also be recognized as an essential resource. Investment in infrastructure and data security is necessary to ensure safe and equitable access [15,20]. Aly et al [45] emphasized the importance of privacy education tailored to regional and age-specific needs, alongside customized digital literacy training. Age-friendly applications, simplified interfaces, and combined online and offline support can improve sustained engagement [17,41]. Collaboration between government, community organizations, and local stakeholders is essential. Peer learning and incentive mechanisms can further encourage participation and create a positive cycle of engagement.

Current policies often focus on broad coverage but do not fully address individual needs. Financial barriers, limited subsidies, and fragmented services remain key challenges. Future policies should adopt a more targeted approach. Resources should be directed toward underserved populations to reduce inequality. Suggested strategies include flexible subsidy systems, device support programs, and stronger legal protection for data security [46,47]. Wu et al [48] also emphasize the necessity of policy transformation. Integrating digital skills training into public health and eldercare services can improve long-term engagement. Simplified procedures and user-friendly tools can also support individuals with low literacy [15,49]. Nurses can play a central role in implementing these policies. They can support coordination, provide follow-up care, and contribute to ongoing evaluation. Their involvement is essential for translating policy into practice and ensuring effective outcomes [41].

Interaction Between Multilevel Factors

A key conceptual contribution of this study is demonstrating that engagement in digital interventions is fundamentally shaped by dynamic and often contradictory interactions across social ecological model levels. This challenges the static notion that resolving a barrier at 1 level automatically facilitates inclusion. Instead, we observed compensatory and canceling cross-level effects. For example, robust interpersonal support can temporarily compensate for low individual digital literacy, but entrenched society-level resistance such as community stigma can entirely cancel out both individual motivation and policy-level infrastructural improvements. Based on these theoretical insights, effective strategies must move beyond addressing isolated factors to resolving these cross-level frictions [50]. Therefore, we propose a multistakeholder, multilevel intervention framework. At the individual level, improving knowledge, skills, and confidence is essential. At the interpersonal level, balanced support from family and peers can promote sustained use. At the society level, accessible services and supportive environments can reduce barriers. At the policy level, fair resource allocation and integrated services can ensure long-term impact.

These levels are closely connected. Policy support can improve infrastructure and reduce financial barriers, but without local adaptation, resources may remain underused. Interpersonal support can compensate for low skills but may also create dependence if not balanced. Positive social environments that combine technical support and peer interaction can reduce pressure on families and promote independent engagement [51,52]. Feedback across levels can further improve system design. Increased engagement at the individual level can also strengthen support from families and communities, creating a reinforcing cycle.

To ensure long-term impact, evaluation should be conducted in stages. Early phases should focus on engagement, followed by retention and sustained use, and finally long-term outcomes such as mental health and well-being. Continuous feedback and adaptation are essential to maintain effectiveness.

Limitations

This study has several limitations. The sample was drawn from a single region in China, which may limit generalizability. Cultural, social, and policy contexts differ across settings and may influence engagement. Multi-site and cross-cultural comparisons are therefore needed in future work. As a qualitative study, this research did not quantify the relative importance of each factor. Mixed-methods approaches could provide further insight. The study focused only on older adults and did not include perspectives from policymakers, health care providers, or family members. A more comprehensive view would benefit from including multiple stakeholders in future investigations. Participants were recruited from a clinical trial and had completed the intervention. This may limit transferability to individuals who do not engage or who drop out early. Their experiences may emphasize successful engagement rather than barriers to initial participation. Including individuals who discontinue interventions is an important direction for future inquiry. In addition, individuals with severe cognitive, sensory, or psychiatric conditions were excluded. As a result, the findings may underestimate the extent of digital exclusion. Exploring the perspectives of these underrepresented groups remains a necessary next step.

Conclusions

Guided by the social ecological model, this study systematically identified multilevel factors that influence older adults' sustained engagement in DBAT in resource-limited settings, specifically capturing the adaptive experiences of users who completed the intervention. Beyond merely categorizing known barriers, including cognitive limitations and infrastructural deficits, this study advances the conceptualization of digital inclusion by revealing the cross-level paradoxical tensions that older adults must navigate. The findings refine the social ecological model by illustrating how emotional burdens of interpersonal support and societal resistance deeply complicate digital engagement. Ultimately, the digital inclusion framework adds critical analytic value by demonstrating that sustained engagement is not merely a technical fix of the digital divide but a profound socioemotional transition. Achieving true inclusion requires moving beyond hardware distribution to cultivate digital autonomy and supportive relational environments. Nurses are

XSL•FO
RenderX

well positioned to support this process. They can coordinate care, provide guidance, and contribute to the development of integrated strategies across levels. Such efforts can improve engagement and treatment outcomes while promoting equitable and sustainable digital mental health services. This study provides practical and theoretical insights to support the development of digital health strategies for aging populations, particularly in resource-limited settings.

Data Availability

The data that supported the findings of this study are available on request from the first author. The data are not publicly available due to privacy restrictions.

Funding

This work was supported by the Wisdom Accumulation and Talent Cultivation Project of the Third Xiangya Hospital of Central South University (No.BJ202205) and the Yale-China Association Chia Community Health Services Program (No.2024YC01). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of this paper.

Generative AI Use

No generative AI tools were used at any stage in the preparation of this manuscript.

Authors' Contributions

Conceptualization: JW, JX
Data curation: YL, GG
Formal analysis: YL, SJ, JL, GG, SW
Funding acquisition: JX
Investigation: WY, DL
Methodology: YL, SJ, DL, YZ, SW
Project administration: JX
Resources: JW, WY, DL, JL, GG
Supervision: JW, SJ, SW, JX
Validation: YL, JL, YZ
Visualization: JW, WY, GG, YZ
Writing – original draft: JW
Writing – review & editing: YL, JX

Conflicts of Interest

None declared.

References

1. Yang Y, Li W, Zhang Q, Zhang L, Cheung T, Xiang YT. Mental health services for older adults in China during the COVID-19 outbreak. Lancet Psychiatry 2020;7(4):e19 [FREE Full text] [doi: 10.1016/S2215-0366(20)30079-1] [Medline: 32085843]
2. Ageing and health. World Health Organization. 2025 Oct 1. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-10-14]
3. Wu Y, Lei P, Ye R, Sunil TS, Zhou H. Prevalence and risk factors of depression in middle-aged and older adults in urban and rural areas in China: a cross-sectional study. Lancet 2019;394:S53. [doi: 10.1016/s0140-6736(19)32389-x]
4. Liu S, Yang L, Zhang C, Xiang Y, Liu Z, Hu S, et al. Online mental health services in China during the COVID-19 outbreak. Lancet Psychiatry 2020;7(4):e17-e18 [FREE Full text] [doi: 10.1016/S2215-0366(20)30077-8] [Medline: 32085841]
5. Fancourt D, Aughterson H, Finn S, Walker E, Steptoe A. How leisure activities affect health: a narrative review and multi-level theoretical framework of mechanisms of action. Lancet Psychiatry 2021;8(4):329-339 [FREE Full text] [doi: 10.1016/S2215-0366(20)30384-9] [Medline: 33581775]
6. Matison AP, Mather KA, Flood VM, Reppermund S. Associations between nutrition and the incidence of depression in middle-aged and older adults: a systematic review and meta-analysis of prospective observational population-based studies. Ageing Res Rev 2021;70:101403. [doi: 10.1016/j.arr.2021.101403] [Medline: 34246793]
7. Arnold C. Innovations from rural communities are improving health care. Nature 2024;634(8035):S30-S33. [doi: 10.1038/d41586-024-03411-4] [Medline: 39443783]
8. Riadi I, Kervin L, Dhillon S, Teo K, Churchill R, Card KG, et al. Digital interventions for depression and anxiety in older adults: a systematic review of randomised controlled trials. Lancet Healthy Longev 2022;3(8):e558-e571 [FREE Full text] [doi: 10.1016/S2666-7568(22)00121-0] [Medline: 36102765]

XSL•FO
RenderX

9. Choi NG, Marti CN, Wilson NL, Chen GJ, Sirrianni L, Hegel MT, et al. Effect of telehealth treatment by lay counselors vs by clinicians on depressive symptoms among older adults who are homebound: a randomized clinical trial. JAMA Netw Open 2020;3(8):e2015648 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.15648] [Medline: 32865577]
10. Nakamura CA, Seward N, Peters TJ, Nadaleto Didone TV, Moretti FA, Oliveira da Costa M, et al. A self-help mobile messaging intervention to improve subthreshold depressive symptoms among older adults in a socioeconomically deprived region of Brazil (PRODIGITAL): a pragmatic, two-arm randomised controlled trial. Lancet Reg Health Am 2024;39:100897 [FREE Full text] [doi: 10.1016/j.lana.2024.100897] [Medline: 39435262]
11. Scazufca M, Nakamura CA, Seward N, Didone TVN, Moretti FA, Oliveira da Costa M, et al. Self-help mobile messaging intervention for depression among older adults in resource-limited settings: a randomized controlled trial. Nat Med 2024;30(4):1127-1133 [FREE Full text] [doi: 10.1038/s41591-024-02864-4] [Medline: 38486075]
12. Ferster CB. A functional analysis of depression. Am Psychol 1973;28(10):857-870. [doi: 10.1037/h0035605] [Medline: 4753644]
13. Arjadi R, Nauta MH, Scholte WF, Hollon SD, Chowdhary N, Suryani AO, et al. Internet-based behavioural activation with lay counsellor support versus online minimal psychoeducation without support for treatment of depression: a randomised controlled trial in Indonesia. Lancet Psychiatry 2018;5(9):707-716. [doi: 10.1016/S2215-0366(18)30223-2] [Medline: 30006262]
14. Zabihi S, Lemmel FK, Orgeta V. Behavioural activation for depression in informal caregivers: a systematic review and meta-analysis of randomised controlled clinical trials. J Affect Disord 2020;274:1173-1183. [doi: 10.1016/j.jad.2020.03.124] [Medline: 32663948]
15. Yao Y, Zhang H, Liu X, Liu X, Chu T, Zeng Y. Bridging the digital divide between old and young people in China: challenges and opportunities. Lancet Healthy Longev 2021;2(3):e125-e126 [FREE Full text] [doi: 10.1016/S2666-7568(21)00032-5] [Medline: 36098110]
16. Min D, Yun J, Parslow C, Jajodia A, Han H. Online-based recruitment methods for community-dwelling older adults: scoping review and lessons learned from the PLAN trial. J Med Internet Res 2025;27:e55082 [FREE Full text] [doi: 10.2196/55082] [Medline: 39998873]
17. Guo L, Li Y, Cheng K, Zhao Y, Yin W, Liu Y. Impact of internet usage on depression among older adults: comprehensive study. J Med Internet Res 2025;27:e65399 [FREE Full text] [doi: 10.2196/65399] [Medline: 39890098]
18. Andrews G, Basu A, Cuijpers P, Craske M, McEvoy P, English C, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
19. Kwok JYY, Jiang D, Yeung DY, Choi NG, Ho RTH, Warner LM, et al. Layperson-delivered telephone-based behavioral activation among low-income older adults during the COVID-19 pandemic: the HEAL-HOA randomized clinical trial. JAMA Netw Open 2024;7(6):e2416767 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.16767] [Medline: 38888923]
20. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
21. Daniel K. e-Inclusion: new challenges and policy recommendations. European Commission. 2009 Apr 1. URL: https://fing.org/e-inclusion-nouveaux-enjeux-nouvelles-politiques/ [accessed 2026-10-14]
22. Li X, Chen C, Li X, Li Y, Xiao S, Han J, et al. Examining the impact of digital inclusion on depression among older adults in China: mediating role of noncognitive abilities. J Med Internet Res 2025;27:e71441 [FREE Full text] [doi: 10.2196/71441] [Medline: 40577713]
23. Liu H, Zhou Q, Liang S. Digital inclusion in public services for vulnerable groups: a systematic review for research themes and goal-action framework from the lens of public service ecosystem theory. Gov Inform Q 2025 Jun;42(2):102019. [doi: 10.1016/j.giq.2025.102019]
24. Wang YC, Miao NF, Wang FT, Chang SR. Factors associated with providers' culturally competent care for sexual and gender minority older adults in Taiwan: a qualitative study applying the social ecological model. Int J Nurs Stud 2022;135:104331. [doi: 10.1016/j.ijnurstu.2022.104331] [Medline: 35981449]
25. Ottie-Boakye D, Ayivor S, Appiah M, Ayivor JS. Predictors of food insecurity among older persons in selected cash-grant communities in Ghana: application of social-ecological model. BMC Public Health 2025;25(1):2028 [FREE Full text] [doi: 10.1186/s12889-025-23183-4] [Medline: 40457302]
26. Jackson C, Vaughan DR, Brown L. Discovering lived experiences through descriptive phenomenology. IJCHM 2018;30(11):3309-3325. [doi: 10.1108/ijchm-10-2017-0707]
27. Hu H, Li Y, Lu L, Li Z, Zhang J, Chang L, et al. Experiences and role adaptation of nursing PhDs in hospital settings: a qualitative study. J Adv Nurs 2026;82(5):4291-4309. [doi: 10.1111/jan.70115] [Medline: 40767124]
28. Luo Y, Wang J, Gan G, Liu M, Lu Z, Jia S, et al. Effectiveness of digital modified behavioral activation treatment program for rural older adults with depressive symptoms in Western Hunan: study protocol for a multi-center randomized controlled trial. BMC Psychol 2024;12(1):783 [FREE Full text] [doi: 10.1186/s40359-024-02215-5] [Medline: 39722075]

XSL•FO
RenderX

29. Lin FF, Peet J, Murray L, Ramanan M, Jacobs K, Brailsford J, et al. Who gets the bed: factors influencing the intensive care exit block: a qualitative study. Int J Nurs Stud 2025;161:104949 [FREE Full text] [doi: 10.1016/j.ijnurstu.2024.104949] [Medline: 39536612]
30. Colaizzi PF. Psychological research as the phenomenologist views it. In: Valle RS, King M, editors. Existential Phenomenological Alternatives for Psychology. New York: Oxford University Press; 1978:48-71.
31. Zhang S, Wu M, Sun R, Cui C, Zhang Z, Liao J, et al. Exploring the discontinuous usage behavior of digital cognitive training among older adults with mild cognitive impairment and their family members: qualitative study using the extended model of IT continuance. J Med Internet Res 2025;27:e66393 [FREE Full text] [doi: 10.2196/66393] [Medline: 40132189]
32. Lincoln YS, Guba EG. But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. New Dir Program Eval 1986;1986(30):73-84. [doi: 10.1002/ev.1427]
33. Yu W, Ji Y, Li Z, Wang K, Jiang X, Chang C. Study on the "digital divide" in the continuous utilization of internet medical services for older adults: combination with PLS-SEM and fsQCA analysis approach. Int J Equity Health 2025;24(1):71 [FREE Full text] [doi: 10.1186/s12939-025-02418-1] [Medline: 40075379]
34. Global strategy on digital health 2020-2025. World Health Organization. 2021 Aug 18. URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2026-10-14]
35. Shen K, Sun H. Aging in the digital age: exploring the relationship between internet use and age identity among Chinese older adults. Telematics Inform 2025;98:102260. [doi: 10.1016/j.tele.2025.102260]
36. Leite H, Hodgkinson IR, Volochtchuk AVL, Nascimento TC. 'It's not the boogeyman': how voice assistant technology is bridging the digital divide for older people. Technovation 2024;136:103080. [doi: 10.1016/j.technovation.2024.103080]
37. Li Y, Liu C, Sun J, Zhang J, Li X, Zhang Z. The digital divide and cognitive disparities among older adults: community-based cohort study in China. J Med Internet Res 2024;26:e59684 [FREE Full text] [doi: 10.2196/59684] [Medline: 39602795]
38. Gan H, Lin C, Zhou Y, Zhuo Z. Adult children's education and elderly parents' energy poverty: evidence from China. Energy Policy 2025;202:114604. [doi: 10.1016/j.enpol.2025.114604]
39. Chang J, Mai Y, Zhang D, Yang X, Li A, Yan W, et al. Media use behavior mediates the association between family health and intention to use mobile health devices among older adults: cross-sectional study. J Med Internet Res 2024;26:e50012 [FREE Full text] [doi: 10.2196/50012] [Medline: 38373031]
40. Liu S, Zheng J, Yang F, Li W, Zhou Y, Lu S, et al. Internet altruistic behaviors and affective well-being among family caregivers: a natural language processing approach. J Gerontol B Psychol Sci Soc Sci 2026;81(1):gbaf230. [doi: 10.1093/geronb/gbaf230] [Medline: 41224706]
41. Fiorini L, Pani J, Rovini E, Delgado MB, Pérez-Martos S, Russo S, et al. Investigating the impact of sociodemographic factors on training efficacy and its correlation with technology usability in older adults: lessons learned in Italian and Murcian pilots. Int J Med Inform 2025;202:105973 [FREE Full text] [doi: 10.1016/j.ijmedinf.2025.105973] [Medline: 40409170]
42. Qiu Y, Huang H, Gai J, De Leo G. The effects of the COVID-19 pandemic on age-based disparities in digital health technology use: secondary analysis of the 2017-2022 Health Information National Trends Survey. J Med Internet Res 2024;26:e65541 [FREE Full text] [doi: 10.2196/65541] [Medline: 39631070]
43. Wang W, Zhang Y, Zhao J. Technological or social? Influencing factors and mechanisms of the psychological digital divide in rural Chinese elderly. Technol Soc 2023;74:102307 [FREE Full text] [doi: 10.1016/j.techsoc.2023.102307]
44. Kebede AS, Ozolins L, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022;24(12):e40192 [FREE Full text] [doi: 10.2196/40192] [Medline: 36477006]
45. Aly H, Liu Y, Khan S, Anaraky RG, Byrne K, Knijnenburg B. Digital privacy education: customized interventions for U.S. older and younger adults in rural and urban settings. Technol Soc 2025;81:102805. [doi: 10.1016/j.techsoc.2024.102805]
46. Kannan L, Bhatt T. Authors' reply: addressing the digital divide among the older population presents a substantial challenge. J Med Internet Res 2025;27:e72565 [FREE Full text] [doi: 10.2196/72565] [Medline: 40153783]
47. Xie S. Addressing the digital divide among the older population presents a substantial challenge. J Med Internet Res 2025;27:e69482 [FREE Full text] [doi: 10.2196/69482] [Medline: 40153778]
48. Wu M, Xue Y, Ma C. The association between the digital divide and health inequalities among older adults in China: nationally representative cross-sectional survey. J Med Internet Res 2025;27:e62645 [FREE Full text] [doi: 10.2196/62645] [Medline: 39813666]
49. Agrawal L, DaSouza RO, Mulgund P, Chaudhary P. Frequency of electronic personal health record use in US older adults: cross-sectional study of a national survey. JMIR Aging 2025;8:e71460 [FREE Full text] [doi: 10.2196/71460] [Medline: 40720810]
50. Smith C. An analysis of digital inclusion projects: three crucial factors and four key components. JITE: Res 2015;14:179-188. [doi: 10.28945/2262]
51. Jung J, You J, Kim D. Effective but sustainable? A case of a digital literacy program for older adults. Educ Inform Technol 2025;30(10):13309-13330. [doi: 10.1007/s10639-025-13364-4]
52. Knotnerus HR, Ngo HTN, Maarsingh OR, van Vugt VA. Understanding older adults' experiences with a digital health platform in general practice: qualitative interview study. JMIR Aging 2024;7:e59168 [FREE Full text] [doi: 10.2196/59168] [Medline: 39212599]

XSL•FO
RenderX

Abbreviations

BA: behavioral activation
COREQ: Consolidated Criteria for Reporting Qualitative Research
DBAT: digital behavioral activation therapy
GDS-15: Geriatric Depression Scale with 15 items
HI: health insurance
LLD: late-life depression
MI: monthly income
WHO: World Health Organization

Edited by J Wang; submitted 18.Sep.2025; peer-reviewed by E Granholm, F Yang, D Yu; comments to author 24.Mar.2026; revised version received 16.Sep.2026; accepted 09.Oct.2026; published 19.Nov.2026.

Please cite as:
Wang J, Luo Y, Jia S, Yang W, Liu D, Li J, Gan G, Zhou Y, Wang S, Xie J
Factors Influencing Older Adults' Sustained Engagement in Digital Psychological Interventions in Resource-Limited Settings: Qualitative Study Guided by the Social Ecological Model
JMIR Aging 2026;9:e84350
URL: https://aging.jmir.org/2026/1/e84350
doi:10.2196/84350
PMID:

©Jingying Wang, Yating Luo, Sipei Jia, Wenling Yang, Dan Liu, Jiarong Li, Gang Gan, Yi Zhou, Sha Wang, Jianfei Xie. Originally published in JMIR Aging (https://aging.jmir.org), 19.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Real-Time Associations Between Subjective Cognitive Concerns and Daily Activity Participation in Persons Aging With Stroke: Smartphone-Based Ecological Momentary Assessment Study

Jiali He[1,2], PhD; Quoc Bui[3], MD, MS; Mandy W M Fong[1,4], PhD; David C Mohr[5,6], PhD; Christopher L Metts[7], MD; Alex W K Wong[1,2,8], PhD, DPhil

[1]Department of Physical Medicine and Rehabilitation, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

[2]Center for Rehabilitation Outcomes Research, Shirley Ryan AbilityLab, 355 E Erie Street, Chicago, IL, United States

[3]Graduate Medical Education Consortium, South Texas Health System, Edinburg, TX, United States

[4]Michigan Avenue Neuropsychologists, Chicago, IL, United States

[5]Center for Behavioral Intervention Technologies, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

[6]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

[7]Department of Pathology and Laboratory Medicine, Medical University of South Carolina, Charleston, SC, United States

[8]Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

Corresponding Author:
Alex W K Wong, PhD, DPhil
Department of Physical Medicine and Rehabilitation, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

Abstract

Background: Stroke is a leading cause of global disability in the aging population, with cognitive impairments playing a significant role. Prior research has shown that subjective cognitive concerns (SCCs) can predict later dementia and serve as an essential indicator for poststroke functional rehabilitation. The use of smartphone-based ecological momentary assessment (EMA) in real-world environments may help us understand how SCCs relate to daily functioning in individuals aging with stroke, thereby guiding cognitive rehabilitation and prevention efforts.

Objective: Our study aimed to use EMA to examine the real-time associations between SCCs and daily activity participation in persons aging with stroke.

Methods: This longitudinal observational study used smartphone-based EMA for real-time assessment of individual cognitive concerns and participation in various daily activities. EMA survey items, including SCCs (concentration and learning) and participation in daily activities (location, company, current activity, and self-appraisals of performance, help needed, satisfaction, and engagement), were collected 5 times per day for 2 weeks. Multilevel models were used to analyze the data.

Results: A total of 202 individuals with mild-to-moderate chronic stroke participated in the study (n=90, 44.6% female; n=89, 44.1% Black; n=182, 90.1% ischemic stroke; mean age 59.7, SD 11.7 years). SCCs were concurrently lower when participants engaged in activities of daily living (ADL; B=−0.04, 95% CI −0.07 to −0.01; *P*=.02), instrumental ADL (B=−0.05, 95% CI −0.07 to −0.02; *P*<.001), cognitively stimulating activities (B=−0.05, 95% CI −0.08 to −0.02; *P*<.001), and social activities (B=−0.05, 95% CI −0.08 to −0.02; *P*=.002); when participants were located in a friend's home (B=−0.10, 95% CI −0.17 to −0.02; *P*=.001); and when they spent time with family members (B=−0.07, 95% CI −0.10 to −0.04; *P*<.001), friends (B=−0.05, 95% CI −0.10 to −0.01; *P*=.01), and spouse or partners (B=−0.04, 95% CI −0.07 to −0.01; *P*=.02). Conversely, SCCs were higher when participants were in the hospital (B=0.39, 95% CI 0.25 - 0.53; *P*<.001). Additionally, greater SCCs were concurrently associated with worse ratings of performance (B=−0.05, 95% CI −0.06 to −0.05; *P*<.001), satisfaction (B=−0.05, 95% CI −0.06 to −0.05; *P*<.001), and activity engagement (B=−0.05, 95% CI −0.06 to −0.04; *P*<.001).

Conclusions: EMA provides an effective means of understanding the links between poststroke cognition and participation in daily activities. Our findings suggest that ADL, instrumental ADL, cognitively demanding activities, and socially engaging activities may lessen cognitive concerns among stroke survivors, implying that clinicians should schedule these activities to help reduce poststroke cognitive issues. Conversely, interventions that enhance cognition may increase participation in these challenging activities. Tracking cognition, everyday activity involvement, and their interactions in real-world settings could ultimately help develop rehabilitation and prevention strategies for individuals at risk of dementia due to stroke.

(JMIR Aging 2026;9:e80222) doi:10.2196/80222

XSL•FO
RenderX

KEYWORDS

stroke; aging; cognition; activities of daily living; participation; ecological momentary assessment; smartphone; technology

Introduction

Stroke is the second most common cause of death and the third leading cause of disability worldwide [1,2], especially in the older population, in which stroke incidence is highest [3]. Alarmingly, stroke incidence in younger age groups (eg, <55 years) has been gradually increasing over the past few decades [4]. Although stroke mortality rates have fallen with advancements in medicine, the burden of disability and dependence on activities of daily living (ADL) after stroke remains substantial. Existing evidence has shown that the incidence of cognitive impairment in the first few weeks following a stroke exceeds 70% [5]. Although many cognitive issues in acute stroke survivors improve within weeks or months after onset [6], longitudinal studies have documented a notable cognitive decline among stroke survivors over long-term follow-up [7], suggesting the importance of investigating the long-term effect of stroke (ie, chronic stroke) and its impact.

Identifying cognitive impairments in research and clinical practice has often relied on objective assessments (eg, neuropsychological testing). In contrast, subjective cognitive concerns (SCCs), defined as whether people self-report cognitive difficulties and, if so, whether and how these problems disturb or annoy them in everyday life, are often neglected [8]. SCCs are commonly reported in people after stroke, with prevalence rates ranging from 28.6% to 90.2%, and these problems often increase over time [9]. The literature suggests that SCCs serve as early indicators of Alzheimer disease and other forms of dementia [10], which are another important source of disability in old age [11,12]. Given that stroke (acute brain tissue damage) itself is a known risk factor for subsequent dementia [13], persons after stroke, in particular, necessitate ongoing monitoring for early detection of cognitive decline. Moreover, cognitive problems prevent stroke survivors from resuming their prestroke roles or achieving independent living [14]. Therefore, monitoring subjective cognitive functioning in everyday life not only helps predict later dementia but also serves as an important indicator of stroke recovery [15].

Research has shown that time spent in passive leisure activities (eg, watching television) is negatively related to cognitive functioning in healthy older adults and increases their risk of developing cognitive impairment in subsequent years [16]. Meanwhile, participation in cognitively, physically, or socially stimulating activities is positively associated with cognitive functioning in healthy older adults and those with medical conditions [17,18]. However, these studies have primarily been conducted in laboratory or clinical settings, often assuming that cognitive characteristics are constant and stable. In fact, cognitive functioning can be influenced by daily life contexts. Furthermore, these studies exploring the link between cognitive difficulties and participation in daily activities have mostly relied on retrospective self-report questionnaires, typically collecting data at one or a few points using cross-sectional or panel designs. These methods may miss subtle changes that happen within and between days and are prone to recall bias. Therefore, it is essential to have a tool that can capture the dynamic interplay of cognition in day-to-day contexts within the study population, such as an ecological momentary assessment (EMA) often implemented via a smartphone. EMA is an intensive longitudinal data collection approach that captures an individual's daily experiences in real time to minimize recall bias in natural environments, thereby enhancing external validity [19]. Multiple studies have supported the use of EMA in stroke. For example, prior poststroke studies have used EMA to understand real-time relationships between depressed mood and daily activity participation, between somatic and mental symptoms and social interactions [20-22], and to classify poststroke latent cognitive memberships [23]. Although there are studies on these topics, no study has used EMA to identify real-time associations between SCCs and daily activity patterns in stroke survivors. Hence, this study aimed to use EMA to investigate SCCs related to daily activity participation in a sample with mild-to-moderate chronic stroke. We used EMA to evaluate SCCs and the time spent on, as well as participants' perceptions of, daily activity participation. Our first research question investigated the momentary relationships between SCCs and various factors, including activity types, locations, and interactions with other people. We hypothesized that participants would report higher SCCs when staying at home, being alone, and participating in passive leisure activities. Our second research question examined whether SCCs were momentarily associated with perceived qualities of activities. We hypothesized that SCCs would be higher during activities that participants rated as more difficult, felt less satisfied with, were less engaged in, and required more help.

Methods

Participants

Participants were recruited from a local hospital database with active enrollment from October 2018 to January 2021. The inclusion criteria were (1) mild-to-moderate stroke defined by a National Institutes of Health Stroke Scale score of 13 or less at the time of stroke onset, either hemorrhagic or ischemic stroke; (2) at least 3 months after stroke before enrollment; (3) no or mild prestroke disability; and (4) English fluency. People with premorbid neurological or psychiatric disorders, severe communication difficulties, severe apraxia, unilateral visual inattention, or limited visual acuity were excluded from this study.

Procedures

Participants completed the study protocol, which included an initial laboratory visit, 14-day EMA monitoring, and a post-EMA laboratory visit.

At the initial laboratory visit, all participants provided written informed consent, completed screening to ensure eligibility, and participated in a 20-minute tutorial on using an EMA survey mobile app [24] on either their iPhone or an iPod touch provided by our research team. Participants completed at least 1 practice

XSL•FO
RenderX

session to ensure that there were no issues using the app. They were given an EMA manual with hotlines to call should questions arise. As part of the assessment protocol, participants also completed in-laboratory measures, such as the Quality of Life in Neurological Disorders [25] and the Patient-Reported Outcome Measurement Information System [26]. Detailed information on the in-laboratory measures was published elsewhere [20].

The EMA monitoring started within 1 week of the initial visit. Participants completed 5 surveys per day for 14 consecutive days. Surveys were delivered randomly at approximately 2.5-hour intervals per day from 8 AM to 10 PM. Participants had 1 hour to complete each survey, with up to 3 reminders sent before it expired. During the first 2 days and periodically thereafter, the research team contacted participants by phone to address any technical issues or questions they might have while completing the EMA surveys at home.

After EMA monitoring, participants revisited the laboratory to uninstall the EMA app or return the device (if borrowed) and received an honorarium.

Ethical Considerations

The ethics committees of Washington University (201704024) and Northwestern University (STU00215308) approved this longitudinal observational study. All participants provided written informed consent at the initial laboratory visit. Participants received an honorarium of up to US $125 for their participation in the research.

Measures

Participants completed a validated EMA survey tailored to the unique needs of stroke survivors (refer to Bui et al [20] for detailed EMA survey questions). Each EMA survey included 23 questions, and participants took about 3 to 5 minutes to complete it. The survey was primarily composed of checkbox- or slider-format questions. At each prompt, participants answered multiple items about their daily activity participation and current levels of SCCs and other poststroke symptoms.

To assess participation in daily activities, participants first reported their current location (eg, home, workplace, and other) and indicated who they spent time with (eg, alone, family members, or others). The subsequent items were customized based on the participant's reported location, prompting queries about their current activities (eg, working, gardening, etc). On the basis of prior classification models [27], we categorized activities into 7 domains: physical, cognitive, social, ADL, instrumental ADL (IADL), vocational, and passive leisure activities, and further grouped them into total productive and nonproductive activities. After reporting their current activities, participants rated the activity quality in terms of the help needed, performance, satisfaction, and engagement using a 7-point scale, ranging from "no help," "not well," "not satisfied," or "not engaged" to "a lot of help," "very well," "very satisfied," or "very engaged."

To assess SCCs, participants reported their current levels using the following questions: (1) "I am having trouble concentrating" and (2) "I am having difficulty learning new tasks or instructions," with a 5-point scale ranging from "not at all" to "very much." We generated a composite score of SCCs by summing the 2 items, such that high scores indicated more cognitive concerns. These 2 items were moderately correlated (r=0.69; P<.001) [23]. Our prior research demonstrated a strong association between the EMA-measured SCCs and the Quality of Life in Neurological Disorders cognitive function (r=−0.64; P<.001) [20].

Statistical Analysis

We used descriptive statistics to examine the demographics of the study sample. Multilevel models (MLMs) examined the real-time associations between SCCs and participation in daily activities. MLMs treated data on 2 levels; EMA observations were modeled as level 1 (within-person level) nested within individuals modeled as level 2 (between-person level). All statistical tests were 2-sided, with a significance level of P<.05.

We conducted models focusing on concurrent associations, including activity type, location, the person participants were interacting with, and quality of activity as predictors of same-survey SCCs. EMA variables were centered to generate person-centered deviation scores. The centered value reflected momentary changes relative to each individual's weekly average. Centering allows the examination of within-subject and between-subject variances separately [28]. We used PROC MIXED in SAS (version 9.4; SAS Institute) to model both between-subject and within-person variances and account for autocorrelation between adjacent observations. PROC MIXED is robust in handling missing data. Thus, all data were analyzed, and random missingness was assumed [1].

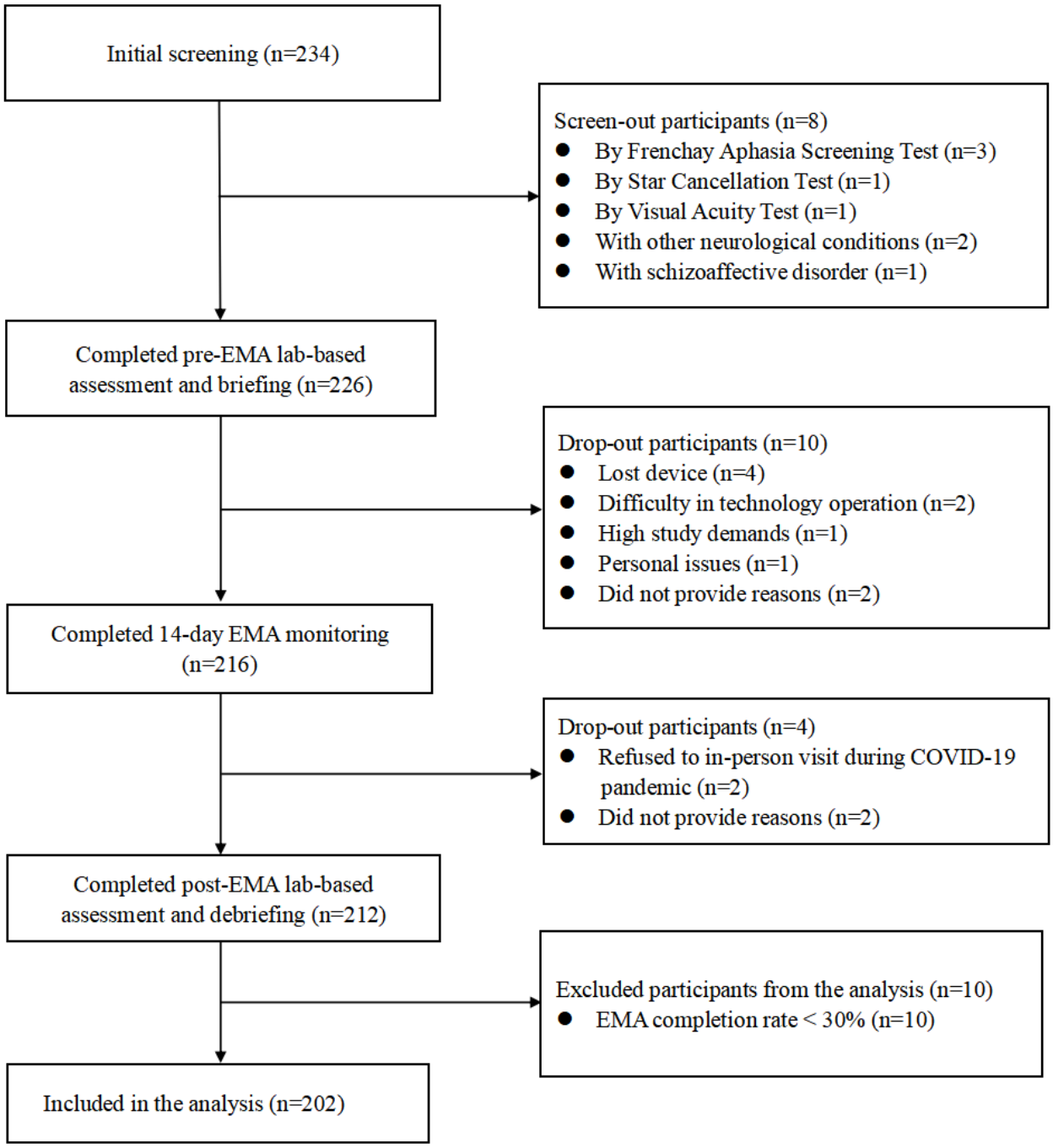
Results

Participant Characteristics

Figure 1 presents the study flow diagram from enrollment to analysis. Of the 212 participants who completed the full study protocol, 10 (4.7%) were excluded because they had an EMA completion rate below 30%. The final sample consisted of 202 (95.3%) stroke survivors (n=90, 44.6% women; mean age 59.7, SD 11.7 years). We observed no significant differences in most demographic and clinical characteristics between the included and excluded participants, except that the included participants had higher global cognitive function (P=.02) and more experience with mobile device use (P=.02). Table 1 presents the details of participant characteristics. We included a total of 14,140 EMA observations for subsequent analyses.

XSL•FO
RenderX

Figure 1. The flow chart of the ecological momentary assessment (EMA) study.



XSL•FO
RenderX

Table . Clinicodemographic information of participants.

Variables	Overall (n=212)	Excluded participants (n=10)	Included participants (n=202)	*P* value[a]
Time since stroke (months), mean (SD)	48.31 (34.18)	54.63 (45.96)	48.00 (33.61)	.84
Age (years), mean (SD)	59.85 (11.69)	62.10 (11.45)	59.74 (11.72)	.46
Gender, n (%)				.52
Female	93 (43.9)	3 (30)	90 (44.6)	
Male	119 (56.1)	7 (70)	112 (55.4)	
Race, n (%)				.09
Asian	1 (0.5)	0 (0)	1 (0.5)	
Black	97 (45.8)	8 (80)	89 (44.1)	
White	114 (53.8)	2 (20)	112 (55.4)	
Marital status, n (%)				.08
Married or cohabitating	107 (50.5)	2 (20)	105 (52)	
Separated, divorced, or widowed	66 (31.1)	6 (60)	60 (29.7)	
Single	39 (18.4)	2 (20)	37 (18.3)	
Residential status, n (%)				.06
Alone	51 (24.1)	5 (50)	46 (22.8)	
With others	161 (75.9)	5 (50)	156 (77.2)	
Stroke diagnosis, n (%)				.60
Hemorrhagic	20 (9.4)	0 (0)	20 (9.9)	
Ischemic	192 (90.6)	10 (100)	182 (90.1)	
Stroke side[b], n (%)				.28
Bilateral	9 (4.2)	0 (0)	9 (4.5)	
Left	84 (39.6)	3 (30)	81 (40.1)	
Right	83 (39.2)	3 (30)	80 (39.6)	
Unknown	35 (16.5)	4 (40)	31 (15.3)	
Premorbid disability (mRS[c]), n (%)				.19
0	177 (83.5)	8 (80)	169 (83.7)	
1	20 (9.4)	0 (0)	20 (9.9)	
2	14 (6.6)	2 (20)	12 (5.9)	
Number of previous strokes, mean (SD)	1.65 (1.76)	1.78 (1.09)	1.65 (1.78)	.23
Stroke severity (NIHSS[d]), mean (SD)	3.23 (3.55)	3.50 (3.17)	3.21 (3.57)	.55
Comorbidity, mean (SD)	4.56 (2.87)	5.11 (2.26)	4.53 (2.90)	.31
Education (years), mean (SD)	14.13 (2.63)	12.80 (2.97)	14.20 (2.60)	.07
Employment status, n (%)				.20
Full time	57 (26.9)	0 (0)	57 (28.2)	
Not employed	125 (59)	9 (90)	116 (57.4)	
Part time	27 (12.7)	1 (10)	26 (12.9)	
Volunteer	3 (1.4)	0 (0)	3 (1.5)	

XSL•FO
RenderX

Variables	Overall (n=212)	Excluded participants (n=10)	Included participants (n=202)	*P* value[a]
Global cognition function (MoCA[e]), mean (SD)	25.86 (3.26)	22.10 (5.32)	26.04 (3.02)	.02
Mobile device use experience, n (%)				.02
No	6 (2.9)	2 (22)	4 (2)	
Yes	203 (97.1)	7 (78)	196 (98)	

[a]Fisher exact test; Wilcoxon rank-sum test.

[b]n=1 participant missing information.

[c]mRS: modified Rankin Scale.

[d]NIHSS: National Institutes of Health Stroke Scale.

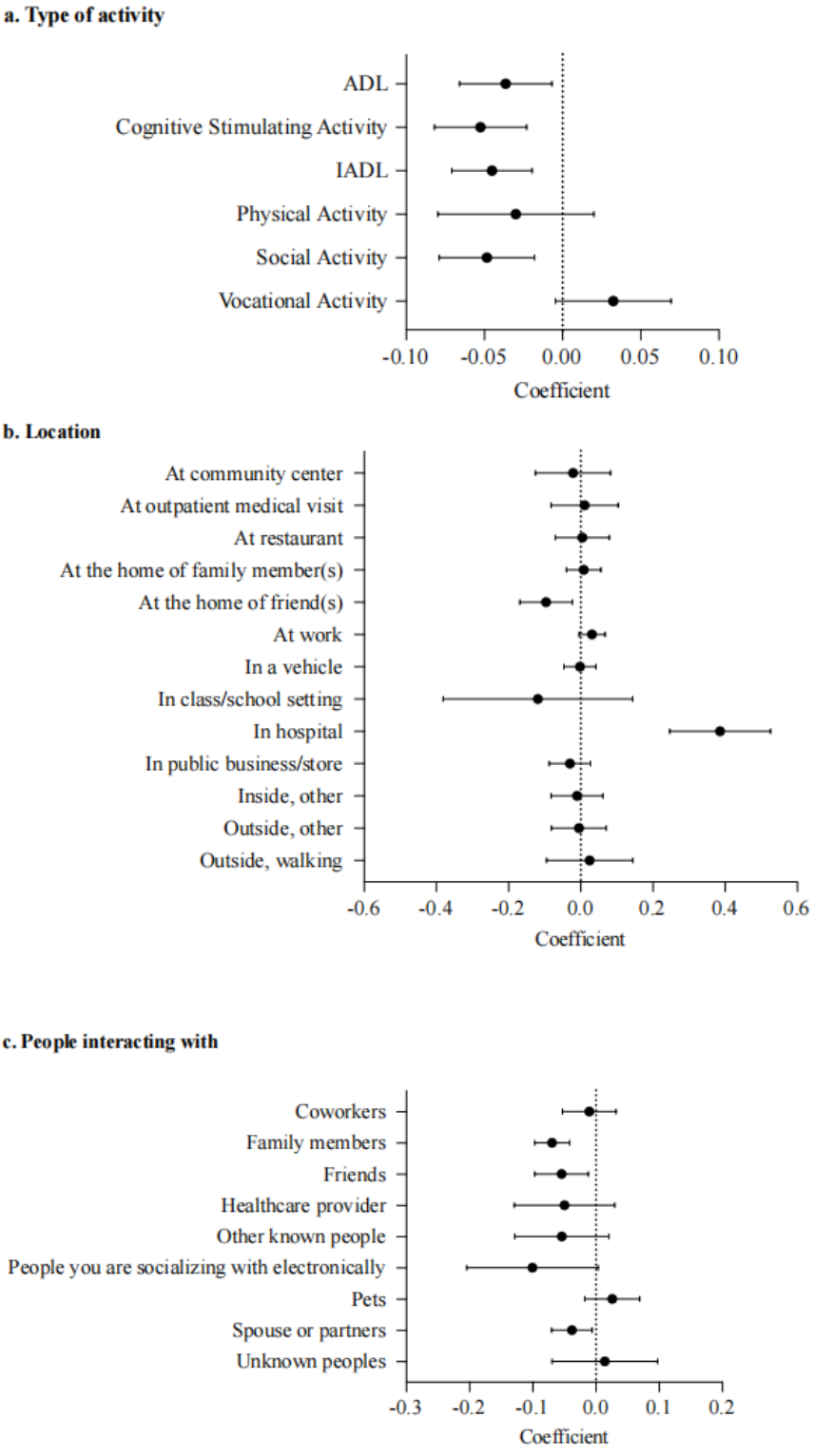
[e]MoCA: Montreal Cognitive Assessment.

Concurrent Relationships Between SCCs and Daily Activity Participation

Descriptive statistics on the daily activity participation, measured by EMA, were reported in our previous publication [20]. By testing the activity type from a total of 11,580 EMA observations, MLM results showed that lower SCCs were significantly associated with more time spent performing ADL (B=−0.04, 95% CI −0.07 to −0.01; *P*=.02), IADL (B=−0.05, 95% CI −0.07 to −0.02; *P*<.001), cognitively stimulating activity (B=−0.05, 95% CI −0.08 to −0.02; *P*<.001), and social activity (B=−0.05, 95% CI −0.08 to −0.02; *P*=.002) relative to passive leisure activity as the reference. We also found no significant relationships between SCCs and physical or vocational activity (Figure 2A and Table 2).

XSL•FO
RenderX

Figure 2. Concurrent relationships of subjective cognitive concerns with daily activity participation are as follows: (A) type of activity (reference: passive leisure activity), (B) location (reference: being at home), and (C) person interacting with (reference: being alone). ADL: activities of daily living; IADL: instrumental activities of daily living.



XSL•FO
RenderX

Table . Results of multilevel models examining concurrent relationships between subjective cognitive concerns and daily activity participation.

Variables	EMA[a] surveys, n (%)	B (SE; 95% CI)	*P* value
Type of activity (n=11,580)			
Intercept	—[b]	0.38 (0.04; 0.30 to 0.46)	<.001
Activities of daily living	1419 (12.3)	−0.04 (0.02; −0.07 to −0.01)	.02
Cognitive stimulating activity	1467 (12.7)	−0.05 (0.02; −0.08 to −0.02)	<.001
Instrumental activities of daily living	2505 (21.6)	−0.05 (0.01; −0.07 to −0.02)	<.001
Physical activity	415 (3.6)	−0.03 (0.03; −0.08 to 0.02)	.24
Social activity	1523 (13.2)	−0.05 (0.02; −0.08 to −0.02)	.002
Vocational activity	993 (8.6)	0.03 (0.02; −0.00 to 0.07)	.09
Passive leisure activity	3258 (28.1)	Referent	—
Location (n=11,642)			
Intercept	—	0.35 (0.04; 0.27 to 0.43)	<.001
At community center	93 (0.8)	−0.02 (0.05; −0.13 to 0.08)	.69
At outpatient medical visit	103 (0.9)	0.01 (0.05; −0.08 to 0.10)	.82
At restaurant	186 (1.6)	0.00 (0.04; −0.07 to 0.08)	.91
At the home of family members	552 (4.7)	0.01 (0.02; −0.04 to 0.06)	.74
At the home of friends	181 (1.6)	−0.10 (0.04; −0.17 to −0.02)	.01
At work	1002 (8.6)	0.03 (0.02; −0.01 to 0.07)	.09
In a vehicle	439 (3.8)	0.00 (0.02; −0.05 to 0.04)	.92
In class or school setting	18 (0.2)	−0.12 (0.13; −0.38 to 0.14)	.37
In hospital	57 (0.5)	0.39 (0.07; 0.25 to 0.53)	<.001
In public business or store	293 (2.5)	−0.03 (0.03; −0.09 to 0.03)	.30
Inside, other	187 (1.6)	−0.01 (0.04; −0.08 to 0.06)	.78
Outside, other	177 (1.5)	−0.01 (0.04; −0.08 to 0.07)	.89
Outside, walking	56 (0.5)	0.02 (0.06; −0.10 to 0.14)	.69
At my home	8298 (71)	Referent	—
Person interacting with (n=8860)			
Intercept	—	0.38 (0.04; 0.30 to 0.46)	<.001
Coworkers	581 (5.1)	−0.01 (0.02; −0.05 to 0.03)	.63
Family members	1940 (17)	−0.07 (0.01; −0.10 to −0.04)	<.001
Friends	441 (3.9)	−0.05 (0.02; −0.10 to −0.01)	.01
Health care provider	118 (1)	−0.05 (0.04; −0.13 to 0.03)	.22
Other known people	128 (1.1)	−0.05 (0.04; −0.13 to 0.02)	.16
People you are socializing with electronically	64 (0.6)	−0.10 (0.05; −0.20 to 0.00)	.06
Pets	861 (7.6)	0.03 (0.02; −0.02 to 0.07)	.24
Spouse or partners	1518 (13.3)	−0.04 (0.02; −0.07 to −0.01)	.02
Unknown people	91 (0.9)	0.01 (0.04; −0.07 to 0.10)	.74
Alone	3118 (27.4)	Referent	—

[a]EMA: ecological momentary assessment.

[b]Not applicable.

XSL•FO
RenderX

By testing the location from a total of 11,642 EMA observations, MLM results revealed that participants reported lower SCCs at the home of friends (B=−0.10, 95% CI −0.17 to −0.02; *P*=.001) but higher SCCs in the hospital, including inpatient and outpatient visits (B=0.39, 95% CI 0.25 - 0.53; *P*<.001) compared with being at home. However, the relationships between SCCs and other locations (eg, community centers and workplaces) were not robust enough to achieve significance compared with being at home (Figure 2B and Table 2).

By testing the person participants were interacting with from a total of 8860 EMA observations, MLM results indicated that participants had lower SCCs when they spent time with their family members (B=−0.07, 95% CI −0.10 to −0.04; *P*<.001), friends (B=−0.05, 95% CI −0.10 to −0.01; *P*=.01), and spouse or partners (B=−0.04, 95% CI −0.07 to −0.01; *P*=.02) when compared with being alone. We found no significant relationships between SCCs and interactions with coworkers, health care providers, and others (Figure 2C and Table 2).

Concurrent Relationships Between SCCs and Activity Qualities

By testing the activity quality from a total of 11,580 EMA observations, MLM results revealed significant associations between greater perceived performance (B=−0.05, 95% CI −0.06 to −0.05; *P*<.001), satisfaction (B=−0.05, 95% CI −0.06 to −0.05; *P*<.001), and engagement with lower SCCs (B=−0.05, 95% CI −0.06 to −0.04; *P*<.001); however, perceived help did not show a statistically significant relationship with SCCs (B=0.00, 95% CI −0.01 to 0.00; *P*=.31; Table 3).

Table . Results of multilevel models examining concurrent relationships between subjective cognitive concerns and activity quality.

Variables	B (SE; 95% CI)	*P* value
Help from someone while doing activities		
Intercept	0.36 (0.05; 0.25 to 0.46)	<.001
Help score (between-person)	−0.01 (0.03; −0.08 to 0.05)	.63
Help score (within-person)	0.00 (0.00; −0.01 to 0.00)	.31
Performance of activities		
Intercept	1.86 (0.17; 1.52 to 2.20)	<.001
Performance score (between-person)	−0.30 (0.03; −0.36 to −0.23)	<.001
Performance score (within-person)	−0.05 (0.00; −0.06 to −0.05)	<.001
Satisfaction with doing activities		
Intercept	1.83 (0.17: 1.51 to 2.16)	<.001
Satisfaction score (between-person)	−0.29 (0.03: −0.36 to −0.23)	<.001
Satisfaction score (within-person)	−0.05 (0.00: −0.06 to −0.05)	<.001
Engagement in activities		
Intercept	1.57 (0.16: 1.25 to 1.88)	<.001
Engagement score (between-person)	−0.25 (0.03: −0.31 to −0.19)	<.001
Engagement score (within-person)	−0.05 (0.00: −0.06 to −0.04)	<.001

Discussion

Principal Findings

This study provides the first in-the-moment examination of the relationships between SCCs and daily activity participation reported on an EMA survey in stroke survivors in naturalistic environments. With respect to our first hypothesis, the results indicated that after stroke, SCCs were not static but were experienced dynamically in daily life. Specifically, SCCs were significantly lower when performing ADL or IADL, engaging in cognitively stimulating activity, and attending social activities than when engaging in passive leisure activities among stroke survivors. Our findings align with a previous real-time assessment study by Moore et al [29], which found greater cognition and multitasking abilities were associated with less time spent in passive leisure activities among older adults living with HIV. Unexpectedly, despite existing evidence consistently indicating that cognitive impairments are linked to high levels of sedentary behavior and low levels of physical activity [30], such a relationship was not observed in this EMA study. Indeed, poststroke cognitive impairment may influence participation in physical activity [31], but the relationship is complex and may not always be direct. Other factors, such as physical limitations, psychological issues, and social support, may mediate how cognitive impairment affects physical activity participation in stroke survivors [32,33]. Further research is necessary to identify these factors as moderators or mediators and to unravel the relationship between after stroke SCCs and physical activity participation in real time.

The results revealed that SCCs were significantly lower among stroke survivors when they were at a friend's home, while higher SCCs were observed in the hospital compared with being at home. This difference may be attributed to psychological factors. As we know, people likely feel more comfortable and relaxed in the familiar surroundings of a friend's home. In contrast, an unanticipated visit to the hospital may impose additional

XSL•FO
RenderX

psychological burdens, such as stress and anxiety, leading to increased concerns of cognitive difficulties [34].

This argument can be extended to another significant finding: stroke survivors exhibited lower SCCs when spending time with family members, friends, or partners compared with when they were alone. Individuals may be more susceptible to psychological challenges when alone than when surrounded by their families and friends [35]. Other factors, such as a lack of social support and limited mobility, may further contribute to heightened SCCs in stroke survivors when they are alone [36].

It is important to recognize that the relationship between SCCs and participation in daily activities may be bidirectional. Cognitive difficulties experienced by stroke survivors can be triggered by certain activities, but these difficulties can also affect their willingness to engage in and choice of activities. For example, stroke survivors who reported higher levels of SCCs were more likely to avoid cognitively demanding activities or social interactions beyond family members and friends. Because the time-lagged relationships between SCCs and daily activity participation, such as responses from a previous EMA survey predicting responses on the current survey, have not been examined, the causal direction remains unclear. Future research could use time-lagged models to investigate these cause-and-effect relationships.

Regarding the second hypothesis, our results showed that SCCs were significantly lower when stroke survivors rated their performance, satisfaction, and engagement in daily activities as higher. It is consistent with previous research [37] indicating that cognitive impairment after stroke is closely linked to poorer performance in daily activities, suggesting that how stroke survivors rate their everyday functional limitations is closely related to their ratings of SCCs. Contrary to our hypothesis, seeking help from others was not significantly associated with SCCs. A prior study on help-seeking behaviors found that many community-dwelling stroke survivors with cognitive difficulties often hesitate to seek assistance from health care professionals. This reluctance can stem from fears of receiving a dementia diagnosis, denial of symptoms as part of “natural” aging, and limited access to care in their community [38]. These results indicate a need for screening and risk assessment of stroke survivors to prevent more severe cognitive impairments or the onset of dementia within the health care system. Future interventions should focus on ensuring that stroke survivors and their caregivers have access to support and information, as well as facilitating referrals back to specialized services if cognitive symptoms arise or persist.

Despite the contributions of this study, several limitations should be acknowledged. First, regarding our EMA survey, in-the-moment SCCs were assessed using only 2 single items. These 2 items may not provide a comprehensive measurement of cognitive functioning compared with a multi-item survey. Future studies should incorporate additional items that assess diverse cognitive domains, such as processing speed, working memory, executive functioning, and receptive and expressive language, to determine whether all types of cognitive concerns are equally significant for participation in daily activities. Furthermore, the 2 items in our study were moderately correlated, suggesting that these 2 cognitive domains may be differently associated with the daily activity predictors included in the study. Because we did not provide an option for participants to respond “not applicable” in this study, we are unable to determine whether the absence of subjective reports of cognitive difficulty reflects a lack of situational demands related to attention or learning, rather than preserved disability. This further underscores the importance of treating the association between perceived difficulties in each cognitive domain and daily activities in considerable detail to achieve a more accurate understanding of the dynamics of cognition and activity participation.

Second, two-thirds of our participants were community-dwelling individuals with neurologically mild stroke, which limits the generalizability of our results to a broader stroke population. Notably, we found that the cognitive function of participants who dropped out was significantly worse than that of participants who were included in the study, suggesting that the EMA method may not be particularly well-suited for stroke populations with more severe cognitive impairment. Similar concerns also apply to populations that lack experience with mobile devices, which significantly contributed to the dropout rate in this EMA study. Third, there may be other confounding variables affecting SCCs in stroke survivors that were not accounted for in this study, such as negative affect [21]. Future research should investigate the specific risk factors associated with developing SCCs after stroke. Additionally, while the use of EMA is a strength of this study, we only monitored participants for 14 days. It remains unclear whether these findings would hold true over a longer monitoring period.

Conclusions

The findings of this study, which used dynamic data collection methods in natural settings and involved a large sample size, support the idea that SCCs fluctuate among stroke survivors in their daily lives. Furthermore, our study highlights the importance of ongoing monitoring of cognitive concerns and daily activity functioning. This monitoring enables us to identify cognitive and functional challenges faced by individuals aging with stroke. Monitoring these factors and their dynamics in real-world contexts may ultimately help identify rehabilitation and prevention strategies for individuals at risk of dementia due to stroke.

Acknowledgments

We thank all personnel at the Shirley Ryan AbilityLab, Northwestern University, and Washington University for their assistance in data collection. The contents do not necessarily represent the funding agencies' policies. We certify that all financial support for this research is identified in the manuscript. The authors affirm that all intellectual content and writing of this manuscript were performed by the authors. Grammarly (AI-assisted software) was used exclusively for grammar and stylistic review. All

XSL•FO
RenderX

recommendations generated by the tool were critically assessed, and only those deemed appropriate were incorporated by the authors.

Funding

The Eunice Kennedy Shriver National Institute of Child Health and Human Development/National Center for Medical Rehabilitation Research (NICHD/NCMRR; K01HD095388) and the NICHD and National Institute of Neurological Disorders and Stroke (NINDS)–funded Center for Smart Use of Technologies to Assess Real-world Outcomes (C-STAR; P2CHD101899).

Data Availability

The data that support the findings of this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JH, MWMF, DCM, AWKW
Formal analysis: QB, MWMF, AWKW
Funding acquisition: AWKW
Writing—original draft: JH
Writing—review and editing: QB, MWMF, DCM, CLM, AWKW
All authors read and approved the final draft of the manuscript.

Conflicts of Interest

DCM reported personal fees from Apple and other support from Actualize Therapy and Otsuka Pharmaceuticals. DCM also has an ownership interest in Adaptive Health and has a pending patent to US Patent 15/654,245, 2018. CLM has an ownership interest in Infinite Arms. No other disclosures were reported.

References

1. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820. [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
2. GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Neurol 2017 Nov;16(11):877-897. [doi: 10.1016/S1474-4422(17)30299-5] [Medline: 28931491]
3. Katan M, Luft A. Global burden of stroke. Semin Neurol 2018 Apr;38(2):208-211. [doi: 10.1055/s-0038-1649503] [Medline: 29791947]
4. Scott CA, Li L, Rothwell PM. Diverging temporal trends in stroke incidence in younger vs older people: a systematic review and meta-analysis. JAMA Neurol 2022 Oct 1;79(10):1036-1048. [doi: 10.1001/jamaneurol.2022.1520] [Medline: 35943738]
5. Nys GM, van Zandvoort MJ, de Kort PL, Jansen BP, de Haan EH, Kappelle LJ. Cognitive disorders in acute stroke: prevalence and clinical determinants. Cerebrovasc Dis 2007;23(5-6):408-416. [doi: 10.1159/000101464] [Medline: 17406110]
6. Gottesman RF, Hillis AE. Predictors and assessment of cognitive dysfunction resulting from ischaemic stroke. Lancet Neurol 2010 Sep;9(9):895-905. [doi: 10.1016/S1474-4422(10)70164-2] [Medline: 20723846]
7. Lo JW, Crawford JD, Desmond DW, et al. Long-term cognitive decline after stroke: an individual participant data meta-analysis. Stroke 2022 Apr;53(4):1318-1327. [doi: 10.1161/STROKEAHA.121.035796] [Medline: 34775838]
8. Hun Sung J, Kim S, Shin JH. Subjective memory complaints and their relationship with the objective cognitive performance of stroke patients. Dement Geriatr Cogn Disord 2022;51(6):475-484. [doi: 10.1159/000527685] [Medline: 36657424]
9. van Rijsbergen MW, Mark RE, de Kort PL, Sitskoorn MM. Subjective cognitive complaints after stroke: a systematic review. J Stroke Cerebrovasc Dis 2014 Mar;23(3):408-420. [doi: 10.1016/j.jstrokecerebrovasdis.2013.05.003] [Medline: 23800498]
10. Mendonça MD, Alves L, Bugalho P. From subjective cognitive complaints to dementia: who is at risk?: a systematic review. Am J Alzheimers Dis Other Demen 2016 Mar;31(2):105-114. [doi: 10.1177/1533317515592331] [Medline: 26142292]
11. Sousa RM, Ferri CP, Acosta D, et al. Contribution of chronic diseases to disability in elderly people in countries with low and middle incomes: a 10/66 Dementia Research Group population-based survey. Lancet 2009 Nov 28;374(9704):1821-1830. [doi: 10.1016/S0140-6736(09)61829-8] [Medline: 19944863]
12. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. Alzheimers Dement 2013 Jan;9(1):63-75.e2. [doi: 10.1016/j.jalz.2012.11.007] [Medline: 23305823]
13. Pendlebury ST, Rothwell PM. Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: a systematic review and meta-analysis. Lancet Neurol 2009 Nov;8(11):1006-1018. [doi: 10.1016/S1474-4422(09)70236-4] [Medline: 19782001]

XSL•FO
RenderX

14. Fride Y, Adamit T, Maeir A, et al. What are the correlates of cognition and participation to return to work after first ever mild stroke? Top Stroke Rehabil 2015 Oct;22(5):317-325. [doi: 10.1179/1074935714Z.0000000013] [Medline: 26461878]
15. McDonald MW, Black SE, Copland DA, et al. Cognition in stroke rehabilitation and recovery research: consensus-based core recommendations from the second Stroke Recovery and Rehabilitation Roundtable. Int J Stroke 2019 Oct;14(8):774-782. [doi: 10.1177/1747493019873600] [Medline: 31514685]
16. Fancourt D, Steptoe A. Television viewing and cognitive decline in older age: findings from the English Longitudinal Study of Ageing. Sci Rep 2019 Feb 28;9(1):2851. [doi: 10.1038/s41598-019-39354-4] [Medline: 30820029]
17. Verghese J, Lipton RB, Katz MJ, et al. Leisure activities and the risk of dementia in the elderly. N Engl J Med 2003 Jun 19;348(25):2508-2516. [doi: 10.1056/NEJMoa022252] [Medline: 12815136]
18. Fazeli PL, Woods SP, Heaton RK, et al. An active lifestyle is associated with better neurocognitive functioning in adults living with HIV infection. J Neurovirol 2014 Jun;20(3):233-242. [doi: 10.1007/s13365-014-0240-z] [Medline: 24554483]
19. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
20. Bui Q, Kaufman KJ, Pham V, et al. Ecological momentary assessment of real-world functional behaviors in individuals with stroke: a longitudinal observational study. Arch Phys Med Rehabil 2022 Jul;103(7):1327-1337. [doi: 10.1016/j.apmr.2022.02.011] [Medline: 35278467]
21. Bui Q, Kaufman KJ, Munsell EG, et al. Smartphone assessment uncovers real-time relationships between depressed mood and daily functional behaviors after stroke. J Telemed Telecare 2024 Jun;30(5):871-884. [doi: 10.1177/1357633X221100061] [Medline: 35549589]
22. Neff AJ, Lee Y, Metts CL, Wong AW. Ecological momentary assessment of social interactions: associations with depression, anxiety, pain, and fatigue in individuals with mild stroke. Arch Phys Med Rehabil 2021 Mar;102(3):395-405. [doi: 10.1016/j.apmr.2020.08.007] [Medline: 32918909]
23. Munsell EG, Bui Q, Kaufman KJ, et al. Intraindividual variability in post-stroke cognition and its relationship with activities of daily living and social functioning: an ecological momentary assessment approach. Top Stroke Rehabil 2024 Sep;31(6):564-575. [doi: 10.1080/10749357.2024.2307203] [Medline: 38278142]
24. Nys GM, van Zandvoort MJ, van der Worp HB, et al. Early cognitive impairment predicts long-term depressive symptoms and quality of life after stroke. J Neurol Sci 2006 Sep 25;247(2):149-156. [doi: 10.1016/j.jns.2006.04.005] [Medline: 16716359]
25. Cella D, Lai JS, Nowinski CJ, et al. Neuro-QOL: brief measures of health-related quality of life for clinical research in neurology. Neurology 2012 Jun 5;78(23):1860-1867. [doi: 10.1212/WNL.0b013e318258f744] [Medline: 22573626]
26. Amtmann D, Cook KF, Johnson KL, Cella D. The PROMIS initiative: involvement of rehabilitation stakeholders in development and examples of applications in rehabilitation research. Arch Phys Med Rehabil 2011 Oct;92(10 Suppl):S12-S19. [doi: 10.1016/j.apmr.2011.04.025] [Medline: 21958918]
27. Granholm E, Holden JL, Mikhael T, et al. What do people with schizophrenia do all day? Ecological momentary assessment of real-world functioning in schizophrenia. Schizophr Bull 2020 Feb 26;46(2):242-251. [doi: 10.1093/schbul/sbz070] [Medline: 31504955]
28. Enders CK, Tofighi D. Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. Psychol Methods 2007 Jun;12(2):121-138. [doi: 10.1037/1082-989X.12.2.121] [Medline: 17563168]
29. Moore RC, Kaufmann CN, Rooney AS, et al. Feasibility and acceptability of ecological momentary assessment of daily functioning among older adults with HIV. Am J Geriatr Psychiatry 2017 Aug;25(8):829-840. [doi: 10.1016/j.jagp.2016.11.019] [Medline: 28065496]
30. Barnes DE, Blackwell T, Stone KL, et al. Cognition in older women: the importance of daytime movement. J Am Geriatr Soc 2008 Sep;56(9):1658-1664. [doi: 10.1111/j.1532-5415.2008.01841.x] [Medline: 18662201]
31. Viktorisson A, Andersson EM, Lundström E, Sunnerhagen KS. Levels of physical activity before and after stroke in relation to early cognitive function. Sci Rep 2021 Apr 27;11(1):9078. [doi: 10.1038/s41598-021-88606-9] [Medline: 33907260]
32. Stolwyk RJ, Mihaljcic T, Wong DK, Chapman JE, Rogers JM. Poststroke cognitive impairment negatively impacts activity and participation outcomes: a systematic review and meta-analysis. Stroke 2021 Jan;52(2):748-760. [doi: 10.1161/STROKEAHA.120.032215] [Medline: 33493048]
33. Gothe NP, Bourbeau K. Associations between physical activity intensities and physical function in stroke survivors. Am J Phys Med Rehabil 2020 Aug;99(8):733-738. [doi: 10.1097/PHM.0000000000001410] [Medline: 32167953]
34. Smit D, Koerts J, Bangma DF, Fuermaier AB, Tucha L, Tucha O. Look who is complaining: psychological factors predicting subjective cognitive complaints in a large community sample of older adults. Appl Neuropsychol Adult 2024;31(3):203-217. [doi: 10.1080/23279095.2021.2007387] [Medline: 34882062]
35. Taylor HO, Taylor RJ, Nguyen AW, Chatters L. Social isolation, depression, and psychological distress among older adults. J Aging Health 2018 Feb;30(2):229-246. [doi: 10.1177/0898264316673511] [Medline: 28553785]
36. McKevitt C, Fudge N, Redfern J, et al. Self-reported long-term needs after stroke. Stroke 2011 May;42(5):1398-1403. [doi: 10.1161/STROKEAHA.110.598839] [Medline: 21441153]
37. Fitri FI, Fithrie A, Rambe AS. The impact of cognitive impairment after stroke on quality of life and daily life activities. Int J Res Sci Manag 2020;7(8):6-11. [doi: 10.29121/ijrsm.v7.i8.2020.2]

XSL•FO
RenderX

38. Tang EY, Price C, Stephan BC, Robinson L, Exley C. Post-stroke memory deficits and barriers to seeking help: views of patients and carers. Fam Pract 2019 Jul 31;36(4):506-510. [doi: 10.1093/fampra/cmy109] [Medline: 30452613]

Abbreviations

ADL: activities of daily living
EMA: ecological momentary assessment
IADL: instrumental activities of daily living
MLM: multilevel model
SCC: subjective cognitive concern

Edited by A Quialheiro; submitted 07.Jul.2025; peer-reviewed by L Zając-Lamparska, Z Kunicki; revised version received 30.May.2026; accepted 04.Jun.2026; published 14.Jul.2026.

Please cite as:
He J, Bui Q, Fong MWM, Mohr DC, Metts CL, Wong AWK
Real-Time Associations Between Subjective Cognitive Concerns and Daily Activity Participation in Persons Aging With Stroke: Smartphone-Based Ecological Momentary Assessment Study
JMIR Aging 2026;9:e80222
URL: https://aging.jmir.org/2026/1/e80222
doi:10.2196/80222

© Jiali He, Quoc Bui, Mandy W M Fong, David C Mohr, Christopher L Metts, Alex W K Wong. Originally published in JMIR Aging (https://aging.jmir.org), 14.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Prediction of 2-Year Incident Clinical Knee Osteoarthritis From Extensor Mechanism Mechanical Properties in Asymptomatic Community-Dwelling Older Adults: Prospective Cohort Study

Ui-jae Hwang[1], PhD; Zongpan Li[1], PhD; Peng Xia[1], PhD; Tianxiang Fan[1], PhD; Arnold Wong[1], PhD; Oh-yun Kwon[2], PhD; Siu-ngor Fu[1], PhD

[1]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, QT 508, The Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, Kowloon, Hong Kong, China (Hong Kong)

[2]Department of Physical Therapy, College of Software and Digital Healthcare Convergence, Yonsei University - Mirae Campus, Wonju, Republic of Korea

Corresponding Author:
Siu-ngor Fu, PhD
Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, QT 508, The Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, Kowloon, Hong Kong, China (Hong Kong)

Abstract

Background: Knee osteoarthritis (KOA) lacks disease-modifying treatment, making early identification of modifiable risk factors in asymptomatic older adults a priority. Patellar tendon stiffness (PTS) and knee extensor strength (KESt) may independently influence KOA development, yet their prognostic value in individuals free of knee symptoms is unestablished.

Objective: This study aimed to determine whether baseline extensor mechanism mechanical properties predict 2-year incident clinical KOA in community-dwelling older adults who were asymptomatic at enrollment, and to quantify the relative contribution of each parameter.

Methods: This 2-year prospective cohort study enrolled 116 community-dwelling older adults (220 knees) free of knee pain at baseline. Incident clinical KOA was diagnosed according to the American College of Rheumatology criteria. Prognostic performance was evaluated across 7 survival algorithms using 5-fold nested cross-validation with the Harrell C-index. Predictor importance was quantified using Shapley Additive Explanations (SHAP) values from the best-performing model, and absolute KOA risk was estimated across tertiles of the linear predictor score.

Results: Over 2 years, 55 (25%) knees developed incident clinical KOA. PTS and KESt were consistently selected across all cross-validation folds, whereas quadriceps muscle passive stiffness was not selected in any fold. The gradient boosted survival model achieved the highest discrimination (C-index=0.720, 95% CI 0.686-0.754), significantly exceeding chance (P<.05). SHAP analysis identified KESt as the dominant predictor (mean |SHAP| 0.314, 95% CI 0.251-0.388), with PTS as the second most important predictor (0.166, 95% CI 0.128-0.200). Knees in the highest-risk tertile showed a 24-month KOA incidence of 38.4% (28/73), which was more than twice that in the lowest-risk tertile (13/74, 17.6%).

Conclusions: PTS and KESt independently and consistently predicted 2-year incident clinical KOA in asymptomatic community-dwelling older adults. Extensor mechanism mechanical profiling may identify individuals at elevated KOA risk prior to symptom onset, although these exploratory findings require confirmation in larger cohorts before KESt and PTS can be considered for screening or preventive intervention.

(JMIR Aging 2026;9:e84494) doi:10.2196/84494

KEYWORDS

arthritis; elderly; machine learning; musculoskeletal; rehabilitation

Introduction

Knee osteoarthritis (KOA) is among the most prevalent musculoskeletal conditions affecting older adults worldwide, with an estimated 595 million individuals affected globally in 2020 and projections indicating a 74.9% increase by 2050 [1]. This condition is associated with substantial functional limitation and currently lacks disease-modifying pharmacological treatment [2], making early identification of modifiable prognostic markers in asymptomatic community-dwelling older adults a research priority. Among the candidate domains, biomechanical factors governing tibiofemoral load distribution are among the most modifiable determinants of KOA onset, and the extensor mechanism represents a primary target given its central role in dynamic knee joint stabilization during weight-bearing activities [3,4].

XSL•FO
RenderX

The extensor mechanism of the knee, comprising the quadriceps muscle group and the patellar tendon, functions to extend the tibiofemoral joint through coordinated muscular force transmission to the tibial tuberosity [5]. Although historically conceptualized as a simple pulley system, biomechanical evidence indicates that it operates as a mechanically complex unit in which the quadriceps musculature and patellar tendon function interdependently [3,5,6]. Three mechanical properties within this integrated system have each been independently supported by prospective evidence as contributors to KOA-related pathology: knee extensor strength (KESt), quadriceps muscle passive stiffness (QMPS), and patellar tendon stiffness (PTS).

KESt is a modifiable risk factor for KOA development. A systematic review and meta-analysis of 46,819 participants demonstrated that quadriceps weakness is a robust prospective risk factor for KOA [7], and a meta-analysis of 15 longitudinal studies found that lower KESt was associated with symptomatic progression in established KOA [8]. Reduced KESt has been associated with increased peak tibiofemoral contact forces and impaired joint stabilization during dynamic loading, which may contribute to articular cartilage degeneration [9-11]. Increased QMPS is a recognized consequence of aging arising from changes in viscoelastic tissue properties and intramuscular connective tissue deposition [12,13], and prospective evidence demonstrates that baseline QMPS is independently associated with incident clinical KOA over 12 to 24 months in community-dwelling older adults [14,15]. Although PTS generally declines with age due to structural degeneration [16], substantial interindividual variation persists even among asymptomatic older adults [17]. This interindividual variation in PTS among asymptomatic older adults may reflect subclinical differences in tendon mechanical integrity prior to symptom onset, and reduced patellar tendon elasticity has been observed in individuals with established KOA compared to asymptomatic controls [18].

Although these 3 parameters have each been examined as individual predictors in separate prospective studies, no study has simultaneously evaluated their combined predictive value, and existing approaches have relied predominantly on conventional regression methods that assume prespecified linear relationships and may not adequately characterize the nonlinear interactions likely to be present among mechanically interdependent predictors [19,20]. Unlike conventional Cox regression, which requires prespecified functional forms for each predictor, ensemble-based machine learning survival algorithms adapt to the structure of the data without assumptions about linearity or additivity and are capable of detecting complex interactions among simultaneously entered predictors. When applied within nested cross-validation frameworks with permutation-based significance testing, these methods provide discrimination estimates that are less susceptible to optimism bias than conventional regression concordance statistics [19], and Shapley Additive Explanations (SHAP) values further allow the relative contribution of each predictor to be quantified within the context of the full predictor set [21].

The purpose of this study was to determine whether baseline KESt, QMPS, and PTS, assessed in community-dwelling older adults who were asymptomatic at enrollment, predict the 2-year incidence of clinical KOA. Secondary aims were to quantify the relative prognostic contribution of each mechanical parameter through SHAP decomposition and to estimate absolute clinical KOA risk across predicted risk tertiles. We hypothesized that at least 1 of the fitted survival models would achieve a C-index of ≥0.70, indicating acceptable discrimination for the 2-year incidence of clinical KOA.

Methods

Study Design and Participants

This study used data from a 2-year prospective cohort study conducted between November 2018 and October 2019. Community-dwelling older adults aged 60 to 80 years who were able to walk 10 meters independently were recruited from local communities. Individuals were excluded if they reported hip, knee, or ankle pain at baseline, had a history of knee injury requiring medical treatment, demonstrated cognitive impairment based on the Montreal Cognitive Assessment, or had a diagnosed neurological condition. Of 225 participants (450 knees) assessed for eligibility, 142 (63%) participants (271/450, 60% knees) met the inclusion criteria. Following the exclusion of knees with missing follow-up data for clinical KOA, 49% (220/450) of knees from 52% (116/225) of participants were included in the present analysis.

A formal a priori sample size calculation was not performed, as this study used data from an existing prospective cohort. The study sample comprised 55 incident clinical KOA events, corresponding to an events-per-variable (EPV) ratio of 9.2 for the primary model. Although this was marginally below the commonly cited threshold of 10, simulation evidence demonstrates that EPV values of 5 to 10 produce acceptable parameter estimation when predictors are continuous and prespecified on theoretical grounds rather than identified through data-driven exploration [22], both of which apply to the present predictor set. Because knees were clustered within participants, we additionally quantified the effective information size. The 55 knee-level events occurred in 31% (36/116) of participants (17/36, 47% unilateral and 19/36, 53% bilateral). The within-participant correlation of the clinical KOA outcome between the 2 knees of a participant was high (intraclass correlation coefficient [ICC] 0.91), corresponding to a design effect of 1.82 and a clustering-adjusted effective EPV of 5.0, which we regarded as the lower bound for interpreting model stability.

Ethical Considerations

The study was approved by the Human Subjects Ethics Sub-committee of the Hong Kong Polytechnic University (HSEARS20180110001), and all participants provided written informed consent prior to participation in accordance with the Declaration of Helsinki.

Baseline Predictor Measurements

The primary predictor set comprised 3 extensor mechanism mechanical properties, each selected on the basis of independent prospective evidence linking it to KOA-related pathology: PTS, QMPS, and KESt [7,14,15,18]. Age, sex, and BMI were

XSL•FO
RenderX

included as covariates. Full measurement protocols, equipment specifications, and reliability data are provided in Multimedia Appendix 1.

PTS (kPa) was measured using shear-wave ultrasound elastography (Aixplorer version 4.2; Hologic Inc), with participants positioned supine and the knee flexed to 30° [23]. After 5 minutes of rest to unload tendon tension, the patellar tendon was located longitudinally using B-mode ultrasonography. Shear elastic modulus was calculated using the equation $E=3\rho c^2$, where ρ is tissue density (1000 kg/m³), and c is shear wave velocity. Three measurements were obtained and averaged. Intrarater reliability was excellent (ICC 0.98) [23].

QMPS (kPa) was defined as the unweighted mean shear modulus of the rectus femoris, vastus lateralis, and vastus medialis, each measured by shear-wave ultrasound elastography at standardized anatomical sites with the knee at 60° flexion [15]. This composite formulation was adopted to reduce the predictor count relative to the number of outcome events, thereby improving the EPV ratio. Intercorrelations among the 3 component muscles were moderate (r=0.34-0.47), supporting their aggregation into a single composite index. Intrarater reliability was excellent across all sites (ICC 0.89-0.94) [15].

KESt (Nm/kg) was assessed as isokinetic peak torque at 60°/s using an isokinetic dynamometer (Cybex Co), with participants seated and the dynamometer axis aligned with the lateral femoral epicondyle [10]. Peak torque values from the middle 3 of 5 maximal concentric contractions were averaged and normalized to body mass. All 3 mechanical predictors were measured separately on the left and right limbs, and each knee's predictor values were linked to the incident clinical KOA outcome of the corresponding knee, so that every knee contributed a self-contained set of limb-specific predictors and an outcome.

Outcome: Incident Clinical KOA

All participants were invited for in-person clinical examination between 12 and 24 months after baseline assessment, with the specific timing of each visit recorded for time-to-event analysis. Clinical examinations were performed by a single experienced physiotherapist blinded to all baseline measurements. Clinical KOA was diagnosed according to the American College of Rheumatology (ACR) clinical classification criteria [15,24,25]. Given that all participants were aged >60 years, a knee was classified as having clinical KOA if it demonstrated knee pain with an intensity of Visual Analog Scale (VAS) ≥3/10 during the examination and any 2 of the following 5 features: morning stiffness lasting <30 minutes, crepitus on active joint motion, bony tenderness on palpation, bony enlargement on inspection or palpation, or absence of palpable warmth of the synovium [15]. Knees that did not meet these criteria at the time of examination were recorded as event-free, with follow-up time defined as the time from baseline to the clinical examination visit. Pain intensity and each physical examination finding were assessed and recorded separately for the left and right knee, and the classification criteria were applied to each knee independently, so that a participant could be classified as having incident clinical KOA in 1 knee, both knees, or neither.

Machine Learning Approaches and Statistical Analysis

All analyses were performed using R (version 4.5.2; R Foundation for Statistical Computing). Missing predictor data (maximum 2.73% per variable) were handled using multiple imputation by chained equations (m=20 datasets, predictive mean matching), with estimates pooled using Rubin rules [26]. The primary predictor set comprised PTS, QMPS, and KESt, adjusted for age, sex, and BMI, producing an EPV ratio of 9.2, considered acceptable for prespecified predictors with strong theoretical priors [22].

Seven survival prediction algorithms were compared: cluster-robust Cox proportional hazards regression (Cox_CR) SEs as the statistical reference, accounting for within-person correlation between bilateral knees [27]; Least Absolute Shrinkage and Selection Operator (LASSO)–penalized Cox regression (LASSO_Cox) [28]; ElasticNet-penalized Cox regression (ElasticNet_Cox) [29]; random survival forest (RSF); gradient boosting survival (GBS) and componentwise LASSO GBS (LGBS) [30]; and a regression-type survival support vector machine (SurvSVM) with a radial basis function kernel [31]. All hyperparameters were optimized within each outer training fold via dedicated inner cross-validation, ensuring that no test-fold observations influenced parameter selection (Multimedia Appendix 2). Prior to hyperparameter tuning, a LASSO-penalized Cox model with 5-fold inner cross-validation (lambda.1se criterion) was applied within each outer training fold to select a parsimonious feature subset, ensuring that variable selection was entirely independent of test-fold observations.

All algorithms were evaluated using 5-fold nested cross-validation, with folds stratified at the participant level by event status to prevent data leakage from bilateral knee observations (Multimedia Appendix 3). Predictor subsets were identified within each outer fold via inner LASSO_Cox, and the selected predictors were then passed to all algorithms. Discrimination was quantified using Harrell C-index (mean across 5 folds; 95% CIs from fold-level SEs), and a participant-level permutation test (999 permutations) was used to confirm that the best-performing model exceeded chance.

Calibration and predictor importance were assessed in a participant-level held-out 20% test set. Calibration at 24 months was evaluated by comparing Kaplan-Meier–observed with model-predicted clinical KOA probabilities across 3 risk tertiles, supplemented by the inverse probability of censoring-weighted Brier score [32]. Predictor contributions were quantified using SHAP values with a Monte Carlo approximation, and directionality was confirmed from hazard ratios (HRs) estimated in a cluster-robust Cox model on the same training partition. The study sample was additionally stratified into low, intermediate, and high clinical KOA risk tertiles by linear predictor score, with 24-month Kaplan-Meier incidence estimated for each tertile. Sex-stratified C-indices were reported with bootstrap 95% CIs, and a sex-by-predictor interaction term was tested in a cluster-robust Cox model. Within-participant correlation between bilateral knees was handled consistently across all components of the analysis. Both knees of a participant were always kept together within the same partition: the nested

XSL•FO
RenderX

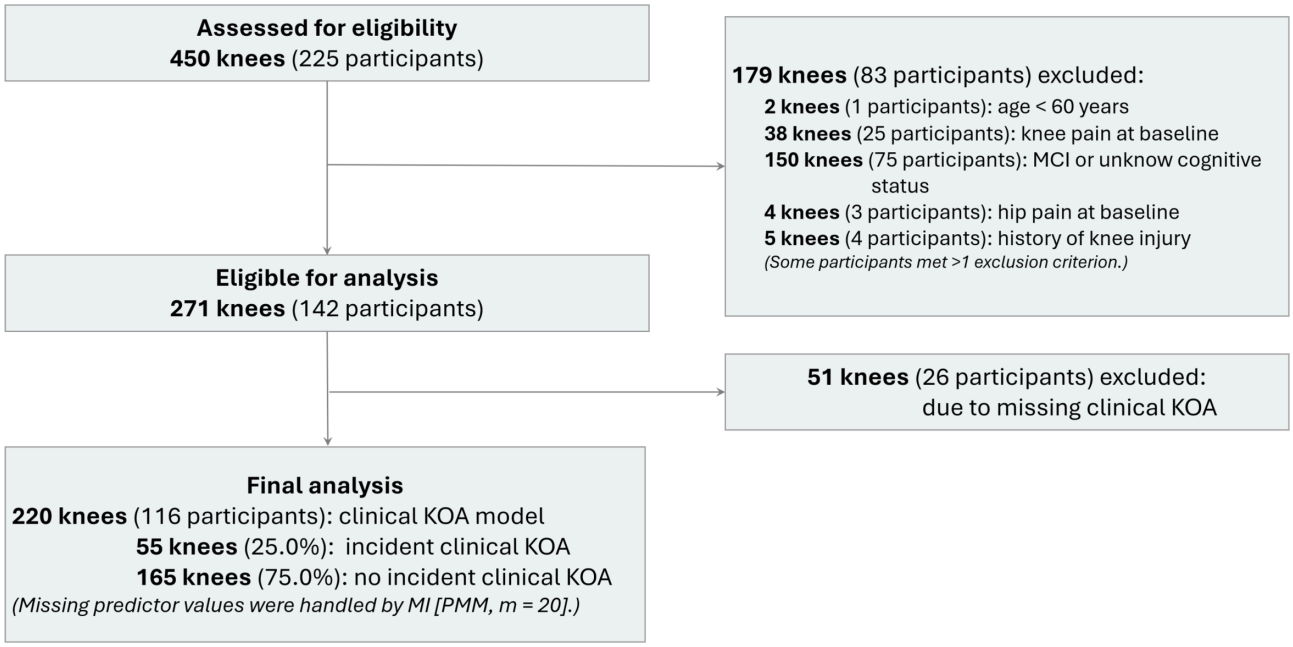
cross-validation folds and the final 80-20 held-out split were assigned at the participant level, so that no participant contributed knees to both the training and evaluation sets. Consequently, all machine learning algorithms, the C-index, calibration, and the inverse probability of censoring-weighted Brier score were trained and evaluated on participant-disjoint partitions, and SHAP values were computed only on the held-out test set. In the permutation test, outcome labels were permuted at the participant level so that both knees of a participant received the same permuted status, preserving the clustered data structure under the null. The Cox reference model additionally used cluster-robust SEs. As a sensitivity analysis, we repeated the modeling with each participant contributing a single observation, collapsing each participant to their more affected knee (the knee with incident clinical KOA, or, when both knees were concordant, the knee with the higher PTS). We then compared the discrimination of this participant-level model with that of the primary knee-level model on a common participant-level held-out test set and tested the difference in C-index using a bootstrap resampling procedure at the participant level.

Results

Participant Characteristics

The analytic sample comprised 220 knees from 116 participants, of whom 55 (25%) knees developed incident clinical KOA over the 2-year follow-up (Figure 1). Missing predictor data were minimal (maximum 2.73% per variable; Multimedia Appendix 4) and handled using multiple imputation. Groups did not differ significantly in age, sex, or BMI. At baseline, knees that subsequently developed clinical KOA showed significantly higher PTS (51.1 vs 40.0 kPa; *P*=.046) and lower KESt (1.14 vs 1.29 Nm/kg; *P*=.01) than knees that did not develop clinical KOA. QMPS did not differ between groups (*P*=.54; Table 1). Intercomponent correlations among the 3 quadriceps stiffness measures were moderate (Spearman *r*=0.344-0.472), supporting their aggregation into a single composite variable (Multimedia Appendix 5). Because clinical KOA is diagnosed at the level of the individual knee and the 2 knees of a participant can differ in both their mechanical properties and their outcome, each knee was treated as a separate observational unit, and incident clinical KOA was ascertained and counted per knee. The 55 incident knee-level events occurred in 36 participants, 47% (17/36) developed clinical KOA in 1 knee (unilateral) and 53% (19/36) in both knees (bilateral); within-participant correlation between paired knees was accounted for in all analyses.

Figure 1. Participant flow diagram. A total of 450 knees (225 participants) were screened for eligibility. Following application of the exclusion criteria, 60% (271/450) of knees (142/225, 63% participants) were eligible for analysis. After exclusion of knees with missing follow-up data for the clinical KOA outcome (51/450, 11% knees from 26/225, 12% participants), 49% (220/450) of knees (116/225, 52% participants) were included in the final analytic sample. Missing predictor data (maximum 2.73% per variable) were handled using multiple imputation by chained equations. KOA: knee osteoarthritis; MCI: mild cognitive impairment; MI: multiple imputation; PMM: predictive mean matching.



XSL•FO
RenderX

Table . Baseline characteristics of the analytic sample stratified by 2-year incident clinical knee osteoarthritis (KOA; N=220 knees from 116 participants)[a].

Characteristics	No KOA (n=165)	KOA (n=55)	*P* value[b]	Effect size (95% CI)[c]
Age (years), mean (SD)	65.7 (3.6)	66.0 (3.6)	.67	0.07 (−0.24 to 0.37)
Sex: female, n (%)	102 (62)	35 (64)	.94	0.01 (1.08; 0.57 to 2.04)
BMI (kg/m^2), mean (SD)	23.5 (3.2)	23.9 (3.1)	.46	0.11 (−0.19 to 0.42)
Quadriceps composite stiffness (kPa), mean (SD)	7.21 (1.77)	7.40 (2.04)	.54	0.10 (−0.20 to 0.41)
Patellar tendon stiffness (kPa), mean (SD)	40.0 (22.3)	51.1 (38.2)	.046[d]	0.41 (0.10 to 0.71)
Knee extension strength (Nm/kg), mean (SD)	1.29 (0.32)	1.14 (0.36)	.01[d]	−0.43 (−0.74 to −0.12)

[a]Both knees from each participant were included as independent observational units; 116 participants contributed 220 knees. Demographic variables reflect participant-level data, whereas mechanical variables reflect knee-level data.

[b]Between-group comparisons were performed using an independent *t* test for continuous variables and a chi-square test for categorical variables.

[c]Effect size is Cohen *d* with 95% CI for continuous variables; positive values indicate a higher mean in the KOA group. For sex, Cramer V (odds ratio; 95% CI) is reported. Standardized effect sizes with CIs are reported in preference to post hoc statistical power.

[d]*P*<.05.

Feature Selection

LASSO-penalized inner Cox regression consistently selected PTS and KESt in all 5 outer cross-validation folds (5/5 each). QMPS was not selected in any fold (0/5). In 1 fold, age and sex were additionally selected alongside the 2 mechanical predictors (Multimedia Appendix 6).

Model Discrimination

All penalized and ensemble algorithms exceeded the null model (C=0.500; Figure 2). GBS achieved the highest discrimination (C=0.720, 95% CI 0.686-0.754), followed closely by LGBS (C=0.718, 95% CI 0.683-0.752) and Cox_CR (C=0.715, 95% CI 0.679-0.750; Table 2). RSF showed comparable mean discrimination (C=0.665) but substantially greater fold-level variability (SD 0.185) than all other algorithms (Multimedia Appendix 7). SurvSVM performed below chance level (C=0.415). Both GBS and Cox_CR significantly exceeded chance-level discrimination by permutation test (*P*<.01; Multimedia Appendix 8).

XSL•FO
RenderX

Figure 2. Machine learning model performance and predictor importance for clinical knee osteoarthritis (KOA) prediction. (A) Forest plot of Harrell C-index values for 7 algorithms (5-fold nested cross-validation; n=220 knees, 55 events). Diamond points and bars represent the mean C-index and 95% CI across folds; the dashed line represents the null model (C=0.50). Gradient boosting survival (GBS) achieved the best discrimination (C=0.720, 95% CI 0.686-0.754; permutation *P*<.01). (B) Mean absolute Shapley Additive Explanations (SHAP) values for the GBS model in the held-out test set; error bars represent the bootstrap 95% CI. Knee extensor strength (KESt) was the dominant predictor (0.314, 95% CI 0.251-0.388), followed by patellar tendon stiffness (PTS), ranked second (0.166, 95% CI 0.128-0.200). (C) SHAP beeswarm plot; each point represents 1 knee, with color indicating feature value (red=high and blue=low). (D) Partial dependence plots for KESt (left) and PTS (right); shaded bands represent the 95% CI. Cox_CR: cluster-robust Cox proportional hazards regression; LGBS: LASSO gradient boosting survival; RSF: random survival forest; SurvSVM: survival support vector machine.

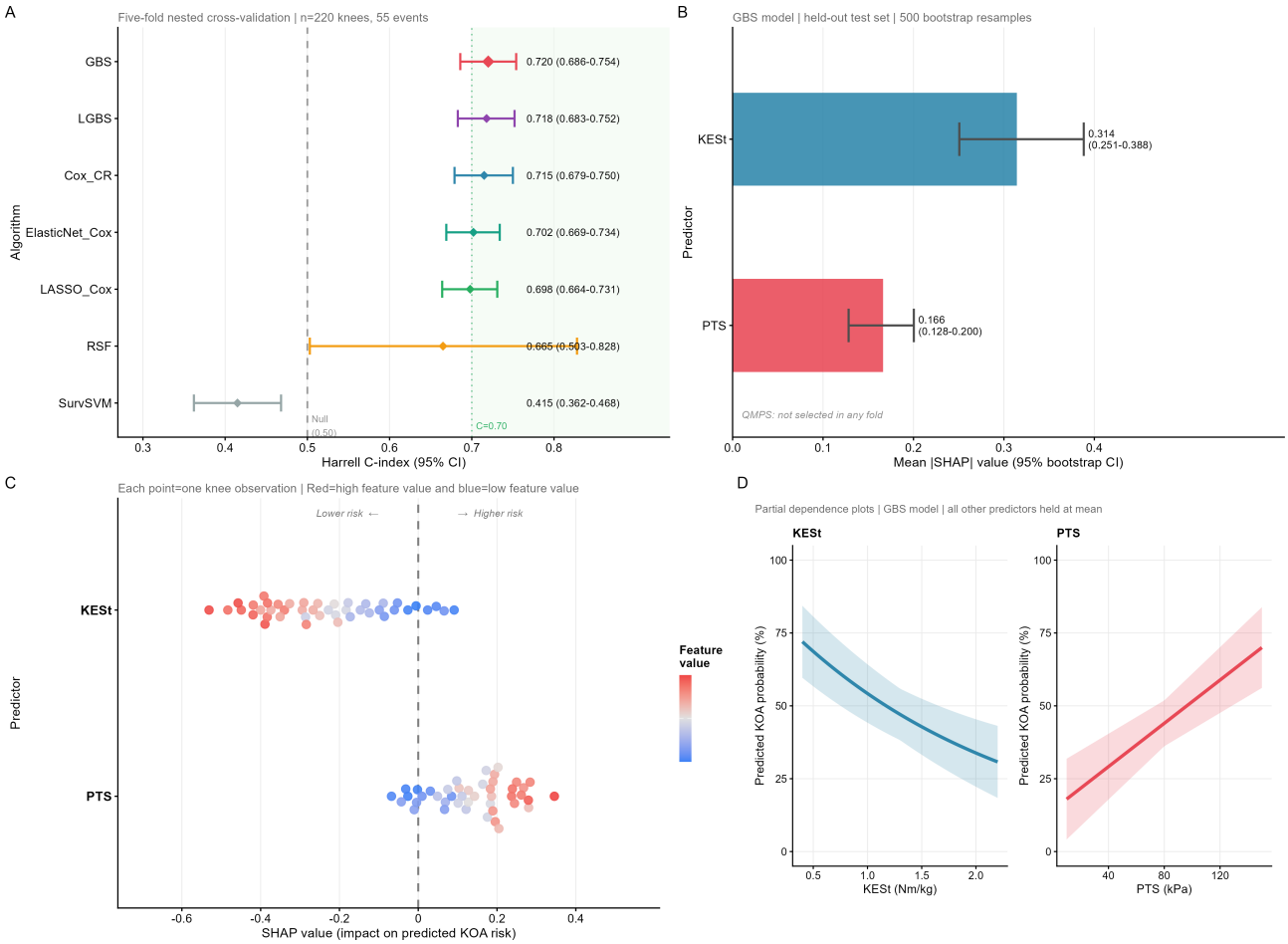


Table . Discrimination performance (Harrell C-index) of 7 survival prediction algorithms evaluated using 5-fold nested cross-validation (n=220 knees, 55 events)[a].

Algorithms	C-index, mean (SD)	95% CI
Cluster-robust Cox proportional hazards regression	0.715 (0.040)	0.679-0.750
LASSO[b]-penalized Cox regression	0.698 (0.038)	0.664-0.731
ElasticNet-penalized Cox regression	0.702 (0.037)	0.669-0.734
Random survival forest	0.665 (0.185)	0.503-0.828
Gradient boosting survival[c]	*0.720 (0.039)*	*0.686-0.754*
Componentwise LASSO gradient boosting survival	0.718 (0.040)	0.683-0.752
Survival support vector machine	0.415 (0.060)	0.362-0.468

[a]C-indices are means across 5 outer folds; 95% CI derived from fold-level SEs. Folds were stratified at the participant level by event status.

[b]LASSO: Least Absolute Shrinkage and Selection Operator.

[c]Italicized text represents the best-performing algorithm.

XSL•FO
RenderX

Predictor Importance and Directionality

SHAP analysis of the GBS model in the held-out test set identified KESt as the dominant predictor (mean |SHAP|=0.314, 95% CI 0.251-0.388), with higher strength consistently associated with lower clinical KOA risk across individual observations (Figure 2). PTS was the second most important contributor (mean |SHAP|=0.166, 95% CI 0.128-0.200), with higher stiffness associated with higher clinical KOA risk. Partial dependence plots confirmed monotonic dose-response relationships for both predictors in the expected directions (Figure 2). These directional findings were further confirmed by cluster-robust Cox HRs estimated on the same training partition: HR of 0.221 (95% CI 0.081-0.601; *P*=.02) for KESt and HR of 1.012 (95% CI 1.002-1.021; *P*=.049) for PTS (Table 3).

Table . Feature selection frequency, Shapley Additive Explanations (SHAP) predictor importance, and directionality for the gradient boosting survival model[a].

Predictors	Selection frequency	SHAP rank	Mean |SHAP| (95% CI)[b]	Hazard ratio (95% CI)[c]	Direction
Knee extension strength (Nm/kg)	5/5	1	0.314 (0.251-0.388)	0.221 (0.081-0.601)[d]	Higher to lower risk
Patellar tendon stiffness (kPa)	5/5	2	0.166 (0.128-0.200)	1.012 (1.002-1.021)[d]	Higher to higher risk
Quadriceps composite stiffness (kPa)	0/5	—[e]	—	—	Not selected

[a]Feature selection was performed via Least Absolute Shrinkage and Selection Operator–penalized Cox regression (lambda.1se) independently within each outer training fold.

[b]Mean absolute SHAP values and 95% CI were derived from 500 bootstrap resamples (200 Monte Carlo simulations).

[c]Hazard ratios and 95% CI were estimated from a cluster-robust Cox model on the same training partition.

[d]*P*<.05.

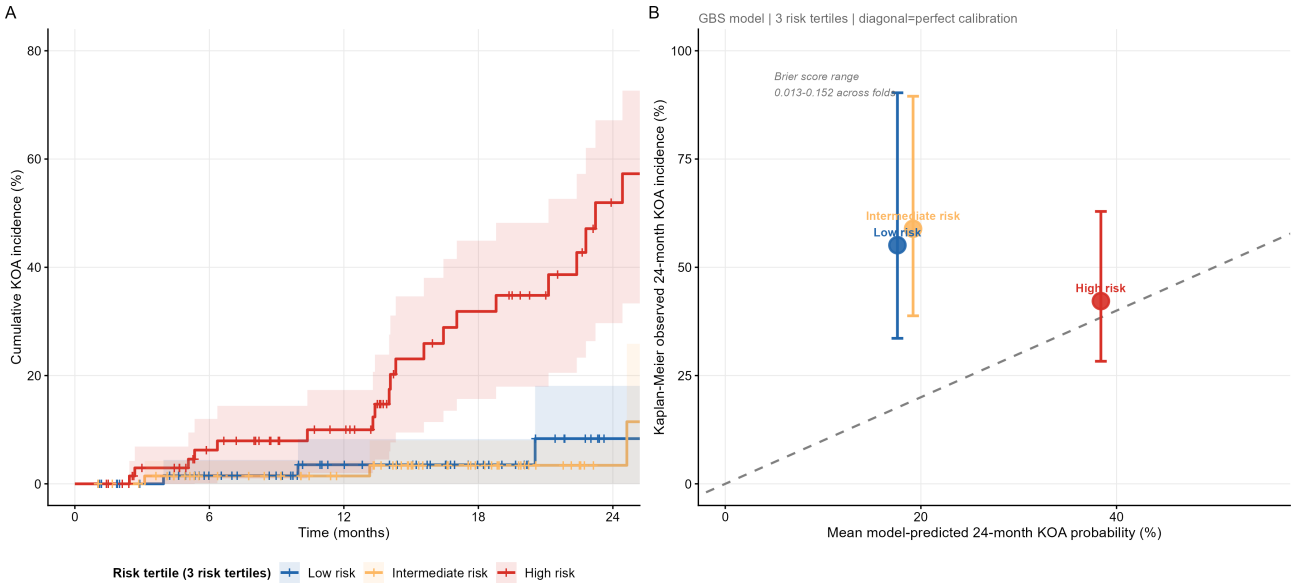
[e]Not applicable (because the predictor was not selected in any fold).

Absolute Risk Stratification and Calibration

The inverse probability of censoring-weighted Brier score at 24 months ranged from 0.013 to 0.152 across folds (Multimedia Appendix 9). Although fold-level variability in the Brier score was observed, mean model-predicted probabilities were broadly consistent with Kaplan-Meier–observed incidence across risk tertiles, indicating acceptable calibration at the group level. Stratification into low, intermediate, and high clinical KOA risk tertiles by GBS linear predictor score generated 24-month clinical KOA incidence rates of 17.6% (13/74), 19.2% (14/73), and 38.4% (28/73), respectively, with the high-risk group showing more than twice the event rate of the low-risk group (Figure 3, Multimedia Appendix 10). Model-predicted probabilities were broadly consistent with Kaplan-Meier–observed incidence across tertiles (Figure 3, Multimedia Appendix 11).

XSL•FO
RenderX

Figure 3. Absolute risk stratification and model calibration. (A) Kaplan-Meier knee osteoarthritis (KOA)–free survival curves stratified by gradient boosting survival (GBS) linear predictor risk tertile (n=220 knees, 116 participants). Low-risk tertile (n=74 knees, 47 participants, 13 events), intermediate-risk tertile (n=73 knees, 57 participants, 14 events), and high-risk tertile (n=73 knees, 48 participants, 28 events). The 24-month KOA incidence was 17.6% (13/74), 19.2% (14/73), and 38.4% (28/73) in the low-, intermediate-, and high-risk tertiles, respectively. Shaded areas represent the 95% CI. Risk tables below the panel indicate the number of knees remaining at risk at each time point. (B) Calibration plot at 24 months comparing mean model-predicted KOA probability with Kaplan-Meier–observed incidence across the 3 risk tertiles. Error bars represent 95% CI. The diagonal dashed line represents perfect calibration. The inverse probability of censoring-weighted Brier score at 24 months ranged from 0.013 to 0.152 across folds.



Sensitivity Analysis

The GBS model demonstrated consistent discrimination in male (C=0.727, 95% CI 0.564-0.835) and female (C=0.748, 95% CI 0.663-0.826) subgroups. A formal sex-by-predictor interaction term in a cluster-robust Cox model was nonsignificant (*P*=.49), indicating no meaningful effect modification by sex (Multimedia Appendix 12, Multimedia Appendix 13). In the participant-level sensitivity analysis, in which each participant contributed a single knee, discrimination on the common held-out test set was similar to that of the primary knee-level model (C=0.685 vs 0.731). The difference in C-index was not statistically significant (difference=0.046, 95% CI −0.096 to 0.162; *P*=.50). The direction of the predictor associations was preserved, with higher PTS associated with higher risk (HR 1.013; *P*=.01) and higher KESt associated with lower risk (HR 0.244; *P*=.02). These results indicate that the main findings are not an artifact of treating the 2 knees of a participant as independent observations.

Discussion

Principal Findings

In this 2-year prospective cohort study of community-dwelling older adults who were asymptomatic at enrollment, KESt and PTS were consistently identified as prognostic predictors of incident clinical KOA across all cross-validation folds, with the prespecified discrimination threshold of C-index ≥0.70 being met. To our knowledge, prospective time-to-event evidence linking higher PTS to incident clinical KOA has not previously been reported, and the present findings extend the existing literature by simultaneously quantifying the relative prognostic contributions of KESt and PTS within the same predictor set.

KESt emerged as the dominant predictor across all model interpretability metrics, with higher strength associated with lower KOA risk. This direction is consistent with prospective evidence that quadriceps weakness predicts KOA onset and progression [7,8,11] and extends those findings by quantifying the contribution of KESt within a simultaneously entered extensor mechanism predictor set. The cluster-robust Cox HR on the training partition (HR 0.221, 95% CI 0.081-0.601) was consistent in direction with previous studies [33,34]. Partial dependence analysis confirmed a monotonic dose-response relationship between KESt and predicted KOA risk, indicating a graded rather than threshold effect, which supports the inference that KESt may be a modifiable factor worth investigating as a potential target for preventive intervention in asymptomatic older adults at elevated KOA risk.

Greater baseline PTS was independently associated with a higher 2-year incidence of clinical KOA. A previous study examining PTS in participants with KOA was limited to cross-sectional designs, which cannot establish whether altered tendon mechanical properties precede or follow structural joint changes [18]. The present findings address this gap by demonstrating that PTS measured in asymptomatic older adults prior to any KOA diagnosis carries independent prognostic information over a 2-year follow-up period. From a mechanical perspective, a stiffer patellar tendon has been proposed to absorb less of the peak force produced during dynamic loading, resulting in higher tibial contact forces during locomotion [16]. Sustained elevation of tibiofemoral contact stress is a recognized contributor to articular cartilage degeneration, and patellar tendon enthesis abnormalities assessed on magnetic resonance imaging (MRI) have been shown to be associated with both knee pain and structural abnormalities in community-dwelling older adults [35], indicating that mechanical alterations at the tendon-bone

XSL•FO
RenderX

interface may contribute to the pathological process leading to clinical KOA. Structural patellar tendon abnormalities are prevalent even in asymptomatic individuals, estimated at 22% to 36% across imaging modalities and populations [36], and the present findings suggest that mechanical stiffness carries prognostic information that is not fully explained by structural abnormality alone. It should be noted that shear-wave elastography provides a measure of tissue shear modulus as an index of passive mechanical stiffness, which is distinct from direct assessment of tendon viscoelasticity, thickness, or tensile properties. Future studies incorporating multiple indices of patellar tendon mechanical behavior may more comprehensively characterize the relationship between tendon properties and KOA risk. Whether age-related changes in tendon elasticity, cumulative mechanical loading history, or physical activity levels contribute to the elevation in PTS that preceded KOA onset in the present cohort warrants investigation in future studies with longer follow-up and serial tendon assessments. In contrast, QMPS was not selected in any cross-validation fold, suggesting that passive muscle stiffness at rest may lack independent prognostic value once active force-generating capacity and tendon mechanical properties are simultaneously accounted for. This pattern indicates that the prognostic signal within the extensor mechanism is carried primarily by its active and structural components rather than by passive muscle stiffness in isolation.

The application of survival analysis with predictor importance decomposition to this prediction problem produced 3 findings of clinical relevance. First, the consistency of predictor selection across all 5 cross-validation folds indicates that PTS and KESt carry a stable prognostic signal independent of data partition. Second, SHAP-based decomposition quantified the relative contribution of each predictor without imposing prespecified functional forms, identifying KESt as contributing approximately twice the predictive weight of PTS. Partial dependence analysis confirmed that both relationships were monotonic, indicating that the associations are adequately characterized by graded, directional effects rather than nonlinear interactions. The primary value of the analytical approach in the present study lies in its capacity to simultaneously rank predictor contributions and verify functional form assumptions without requiring a priori specification. Finally, risk stratification by linear predictor score identified a high-risk tertile with a 24-month KOA incidence more than twice that of the low-risk tertile. This gradient suggests that older adults with concurrently low KESt and high PTS constitute an identifiable subgroup for whom targeted preventive strategies, including quadriceps strengthening and tendon loading programs, may warrant priority consideration, pending confirmation in prospective interventional designs.

Several limitations of this study should be acknowledged. First, clinical KOA was diagnosed according to ACR clinical classification criteria without radiographic confirmation, which precludes direct comparison with studies using Kellgren-Lawrence grading as the outcome and may have introduced misclassification at the diagnostic boundary. Second, the study was conducted in a single community-based cohort, and external validation in an independent cohort is required before the risk stratification thresholds reported here can be considered for clinical application. Third, the 12-to-24–month follow-up window may be insufficient to capture the full spectrum of KOA incidence in an asymptomatic older population, given the typically prolonged natural history of the disease. Fourth, PTS was measured at 30° knee flexion with the tendon in a partially unloaded state, which may not fully reflect the functional stiffness experienced during weight-bearing activities. Finally, the nominal EPV ratio of 9.2 was calculated from the 55 knee-level events, marginally below the conventional threshold of 10, which may have introduced instability in the penalized regression estimates, particularly in folds with fewer events. Moreover, because the 55 knee-level events were clustered within 36 participants, the effective number of independent outcome events was smaller than the nominal count (clustering-adjusted effective EPV 5.0). Although a participant-level sensitivity analysis produced discrimination that did not differ significantly from the knee-level model, this clustering reduces the effective information available for model estimation, and the findings should therefore be regarded as exploratory pending validation in larger independent cohorts. Replication in larger prospective cohorts with longer follow-up and radiographically confirmed outcomes is required to determine the generalizability of the extensor mechanism mechanical profiling approach.

Conclusions

In community-dwelling older adults free of knee symptoms at baseline, KESt and PTS were independently and consistently associated with 2-year incident clinical KOA, whereas QMPS was not retained as a prognostic predictor once KESt and PTS were simultaneously accounted for. Risk stratification based on extensor mechanism mechanical properties identified a subgroup with a 24-month KOA incidence more than twice that of the lowest-risk group, suggesting that baseline mechanical profiling of the extensor mechanism may provide clinically meaningful prognostic information prior to symptom onset. These findings support KESt and PTS as candidate targets that warrant further investigation. Given the modest sample, the clustered knee-level structure, and the reduced effective event count, these findings should be interpreted as exploratory and preliminary rather than as support for immediate use of these parameters in screening or preventive intervention. Confirmation in larger independent cohorts with radiographically confirmed outcomes is required before any clinical application.

Acknowledgments

The authors thank all participants in the study for their active participation and cooperation. The authors also gratefully acknowledge the donation of the Aixplorer ultrasound by Dr Lui Che Woo and family to Prof Fu, which was instrumental in conducting this

XSL•FO
RenderX

research. During the preparation of this manuscript, the authors used Claude (Anthropic) to assist with language editing and to improve the clarity of the text. The tool was not used to generate study data, design the study, or perform the statistical analyses. All content was reviewed, verified, and edited by the authors, who take full responsibility for the integrity and accuracy of the work.

Funding

This research was supported by the Research Centre for Innovative Technologies for Chronic Musculoskeletal Pain (grant 1-CE0K) at the Hong Kong Polytechnic University.

Data Availability

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: UJH, SNF
Data curation: UJH, ZL, PX, TF, ZL, AYW
Formal analysis: ZL, AYW
Methodology: UJH, LZ, OYK, AYW
Project administration: SNF
Software: UJH, ZL, PX, TF
Supervision: SNF
Validation: UJH, ZL, PX, TF
Visualization: UJH, SNF
Writing—original draft: UJH, SNF
Writing—review and editing: UJH, LZ, OYK, AYW

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed measurement protocols for shear-wave elastography, quadriceps stiffness, patellar tendon stiffness, and knee extensor strength.
[DOCX File, 16 KB - aging_v9i1e84494_app1.docx]

Multimedia Appendix 2
Hyperparameter tuning grids and selected values for all algorithms.
[DOCX File, 14 KB - aging_v9i1e84494_app2.docx]

Multimedia Appendix 3
Schematic of the 5-fold nested cross-validation design.
[DOCX File, 99 KB - aging_v9i1e84494_app3.docx]

Multimedia Appendix 4
Missing data summary for predictor variables (N=220 knees).
[DOCX File, 14 KB - aging_v9i1e84494_app4.docx]

Multimedia Appendix 5
Spearman correlations among the 3 quadriceps muscle stiffness components.
[DOCX File, 14 KB - aging_v9i1e84494_app5.docx]

Multimedia Appendix 6
Feature selection history across the 5 outer cross-validation folds.
[DOCX File, 15 KB - aging_v9i1e84494_app6.docx]

Multimedia Appendix 7
Fold-level C-index variability across the 5 outer cross-validation folds for all 7 algorithms.
[DOCX File, 116 KB - aging_v9i1e84494_app7.docx]

XSL•FO
RenderX

Multimedia Appendix 8
Permutation test results for model discrimination.
[DOCX File, 14 KB - aging_v9i1e84494_app8.docx]

Multimedia Appendix 9
Fold-level C-indices and Brier scores across the 5 outer cross-validation folds.
[DOCX File, 15 KB - aging_v9i1e84494_app9.docx]

Multimedia Appendix 10
Twenty-four–month knee osteoarthritis incidence by gradient boosting survival linear predictor risk tertile.
[DOCX File, 15 KB - aging_v9i1e84494_app10.docx]

Multimedia Appendix 11
Overall Kaplan-Meier knee osteoarthritis–free survival for the full analytic sample.
[DOCX File, 89 KB - aging_v9i1e84494_app11.docx]

Multimedia Appendix 12
Sex-stratified C-indices for the gradient boosting survival model.
[DOCX File, 14 KB - aging_v9i1e84494_app12.docx]

Multimedia Appendix 13
Sex-stratified Kaplan-Meier knee osteoarthritis–free survival and C-index comparison.
[DOCX File, 194 KB - aging_v9i1e84494_app13.docx]

References

1. GBD 2021 Osteoarthritis Collaborators. Global, regional, and national burden of osteoarthritis, 1990-2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Rheumatol 2023;5(9):e508-e522. [doi: 10.1016/S2665-9913(23)00163-7] [Medline: 37675071]
2. Hunter DJ, March L, Chew M. Osteoarthritis in 2020 and beyond: a Lancet Commission. Lancet 2020 Nov;396(10264):1711-1712. [doi: 10.1016/S0140-6736(20)32230-3] [Medline: 33159851]
3. Davis JE, Harkey MS, Ward RJ, et al. Accelerated knee osteoarthritis is associated with pre-radiographic degeneration of the extensor mechanism and cruciate ligaments: data from the Osteoarthritis Initiative. BMC Musculoskelet Disord 2019 Jun 29;20(1):308. [doi: 10.1186/s12891-019-2685-y] [Medline: 31253142]
4. Gray HA, Guan S, Thomeer LT, Pandy MG. Moment arm of the knee-extensor mechanism measured in vivo across a range of daily activities. J Biomech 2021 Jun 23;123:110484. [doi: 10.1016/j.jbiomech.2021.110484] [Medline: 34062347]
5. Dan M, Parr W, Broe D, Cross M, Walsh WR. Biomechanics of the knee extensor mechanism and its relationship to patella tendinopathy: a review. J Orthop Res 2018 Dec;36(12):3105-3112. [doi: 10.1002/jor.24120] [Medline: 30074265]
6. Halonen KS, Mononen ME, Jurvelin JS, et al. Importance of patella, quadriceps forces, and depthwise cartilage structure on knee joint motion and cartilage response during gait. J Biomech Eng 2016 Jul 1;138(7). [doi: 10.1115/1.4033516] [Medline: 27138135]
7. Øiestad BE, Juhl CB, Culvenor AG, Berg B, Thorlund JB. Knee extensor muscle weakness is a risk factor for the development of knee osteoarthritis: an updated systematic review and meta-analysis including 46 819 men and women. Br J Sports Med 2022 Mar;56(6):349-355. [doi: 10.1136/bjsports-2021-104861] [Medline: 34916210]
8. Culvenor AG, Ruhdorfer A, Juhl C, Eckstein F, Øiestad BE. Knee extensor strength and risk of structural, symptomatic, and functional decline in knee osteoarthritis: a systematic review and meta-analysis. Arthritis Care Res (Hoboken) 2017 May;69(5):649-658. [doi: 10.1002/acr.23005] [Medline: 27563843]
9. Claiborne TL, Armstrong CW, Gandhi V, Pincivero DM. Relationship between hip and knee strength and knee valgus during a single leg squat. J Appl Biomech 2006 Feb;22(1):41-50. [doi: 10.1123/jab.22.1.41] [Medline: 16760566]
10. Segal NA, Torner JC, Felson D, et al. Effect of thigh strength on incident radiographic and symptomatic knee osteoarthritis in a longitudinal cohort. Arthritis Rheum 2009 Sep 15;61(9):1210-1217. [doi: 10.1002/art.24541] [Medline: 19714608]
11. Øiestad BE, Juhl CB, Eitzen I, Thorlund JB. Knee extensor muscle weakness is a risk factor for development of knee osteoarthritis. A systematic review and meta-analysis. Osteoarthritis Cartilage 2015 Feb;23(2):171-177. [doi: 10.1016/j.joca.2014.10.008] [Medline: 25450853]
12. Alfuraih AM, Tan AL, O'Connor P, Emery P, Wakefield RJ. The effect of ageing on shear wave elastography muscle stiffness in adults. Aging Clin Exp Res 2019 Dec;31(12):1755-1763. [doi: 10.1007/s40520-019-01139-0] [Medline: 30762201]

XSL•FO
RenderX

13. Xu J, Fu SN, Hug F. Age-related increase in muscle stiffness is muscle length dependent and associated with muscle force in senior females. BMC Musculoskelet Disord 2021 Sep 27;22(1):829. [doi: 10.1186/s12891-021-04519-8] [Medline: 34579696]
14. Li Z, Leung KL, Huang C, et al. Associations amongst dynamic knee stiffness during gait, quadriceps stiffness, and the incidence of knee osteoarthritis over 24 months: a cohort study with a mediation analysis. BMC Musculoskelet Disord 2024;25(1):511. [doi: 10.1186/s12891-024-07618-4] [Medline: 38961407]
15. Li Z, Leung KL, Huang C, Huang X, Chung R, Fu SN. Passive stiffness of the quadriceps predicts the incidence of clinical knee osteoarthritis in twelve months. Eur J Phys Rehabil Med 2023 Feb;59(1):65-74. [doi: 10.23736/S1973-9087.22.07634-1] [Medline: 36598343]
16. Carroll CC, Dickinson JM, Haus JM, et al. Influence of aging on the in vivo properties of human patellar tendon. J Appl Physiol (1985) 2008 Dec;105(6):1907-1915. [doi: 10.1152/japplphysiol.00059.2008] [Medline: 18927271]
17. Hsiao MY, Chen YC, Lin CY, Chen WS, Wang TG. Reduced patellar tendon elasticity with aging: in vivo assessment by shear wave elastography. Ultrasound Med Biol 2015 Nov;41(11):2899-2905. [doi: 10.1016/j.ultrasmedbio.2015.07.008] [Medline: 26304500]
18. Dickson DM, Smith SL, Hendry GJ. Strain sonoelastography in asymptomatic individuals and individuals with knee osteoarthritis: an evaluation of quadriceps and patellar tendon. Rheumatol Int 2022 Dec;42(12):2241-2251. [doi: 10.1007/s00296-022-05184-3] [Medline: 35974116]
19. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol 2016 Jan;69:245-247. [doi: 10.1016/j.jclinepi.2015.04.005] [Medline: 25981519]
20. Makridakis S, Spiliotis E, Assimakopoulos V. Statistical and machine learning forecasting methods: concerns and ways forward. PLoS One 2018;13(3):e0194889. [doi: 10.1371/journal.pone.0194889] [Medline: 29584784]
21. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems: Curran Associates Inc; 2017:4768-4777. [doi: 10.5555/3295222.3295230]
22. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 2007 Mar 15;165(6):710-718. [doi: 10.1093/aje/kwk052] [Medline: 17182981]
23. Zhang ZJ, Fu SN. Shear elastic modulus on patellar tendon captured from Supersonic Shear Imaging: correlation with tangent traction modulus computed from material testing system and test-retest reliability. PLoS One 2013;8(6):e68216. [doi: 10.1371/journal.pone.0068216] [Medline: 23826378]
24. Wang Q, Runhaar J, Kloppenburg M, et al. Evaluation of the diagnostic performance of American College of Rheumatology, EULAR, and National Institute for Health and Clinical Excellence Criteria against clinically relevant knee osteoarthritis: data from the CHECK Cohort. Arthritis Care Res (Hoboken) 2024 Apr;76(4):511-516. [doi: 10.1002/acr.25270] [Medline: 37933434]
25. Altman R, Asch E, Bloch D, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum 1986 Aug;29(8):1039-1049. [doi: 10.1002/art.1780290816] [Medline: 3741515]
26. Van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45(3):1-67. [doi: 10.18637/jss.v045.i03]
27. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. J Am Stat Assoc 1989;84(408):1074-1078. [doi: 10.1080/01621459.1989.10478874]
28. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med 1997 Feb 28;16(4):385-395. [doi: 10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3] [Medline: 9044528]
29. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Series B Stat Methodol 2005 Apr;67(2):301-320. [doi: 10.1111/j.1467-9868.2005.00503.x]
30. Taylor JM. Random survival forests. J Thorac Oncol 2011 Dec;6(12):1974-1975. [doi: 10.1097/JTO.0b013e318233d835] [Medline: 22088987]
31. Van Belle V, Pelckmans K, Van Huffel S, Suykens JA. Improved performance on high-dimensional survival data by application of survival-SVM. Bioinformatics 2011 Jan 1;27(1):87-94. [doi: 10.1093/bioinformatics/btq617] [Medline: 21062763]
32. Gerds TA, Schumacher M. Consistent estimation of the expected Brier score in general survival models with right-censored event times. Biom J 2006 Dec;48(6):1029-1040. [doi: 10.1002/bimj.200610301] [Medline: 17240660]
33. Patterson BE, Girdwood MA, West TJ, et al. Muscle strength and osteoarthritis of the knee: a systematic review and meta-analysis of longitudinal studies. Skeletal Radiol 2023 Nov;52(11):2085-2097. [doi: 10.1007/s00256-022-04266-4] [Medline: 36562820]
34. Kemnitz J, Wirth W, Eckstein F, Ruhdorfer A, Culvenor AG. Longitudinal change in thigh muscle strength prior to and concurrent with symptomatic and radiographic knee osteoarthritis progression: data from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2017 Oct;25(10):1633-1640. [doi: 10.1016/j.joca.2017.07.003] [Medline: 28698106]

XSL•FO
RenderX

35. Mattap SM, Aitken D, Wills K, et al. Patellar tendon enthesis abnormalities and their association with knee pain and structural abnormalities in older adults. Osteoarthritis Cartilage 2019 Mar;27(3):449-458. [doi: 10.1016/j.joca.2018.11.009] [Medline: 30529466]
36. Docking SI, Rio E, Girdwood MA, Hannington MC, Cook JL, Culvenor AG. Physical activity and investigation with magnetic resonance imaging partly explain variability in the prevalence of patellar tendon abnormalities: a systematic review with meta-analysis of imaging studies in asymptomatic individuals. J Orthop Sports Phys Ther 2021 May;51(5):216-231. [doi: 10.2519/jospt.2021.10054] [Medline: 33779217]

Abbreviations

ACR: American College of Rheumatology
Cox_CR: cluster-robust Cox proportional hazards regression
ElasticNet_Cox: ElasticNet-penalized Cox regression
EPV: events-per-variable
GBS: gradient boosting survival
HR: hazard ratio
ICC: intraclass correlation coefficient
KESt: knee extensor strength
KOA: knee osteoarthritis
LASSO: Least Absolute Shrinkage and Selection Operator
LASSO_Cox: LASSO-penalized Cox regression
LGBS: LASSO gradient boosting survival
MRI: magnetic resonance imaging
PTS: patellar tendon stiffness
QMPS: quadriceps muscle passive stiffness
RSF: random survival forest
SHAP: Shapley Additive Explanations
SurvSVM: survival support vector machine
VAS: Visual Analog Scale

Edited by M Gray; submitted 20.Sep.2025; peer-reviewed by W Wang, X Cen, Y Hashimoto; revised version received 28.Jul.2026; accepted 27.Aug.2026; published 24.Sep.2026.

Please cite as:
Hwang UJ, Li Z, Xia P, Fan T, Wong A, Kwon OY, Fu SN
Prediction of 2-Year Incident Clinical Knee Osteoarthritis From Extensor Mechanism Mechanical Properties in Asymptomatic Community-Dwelling Older Adults: Prospective Cohort Study
JMIR Aging 2026;9:e84494
URL: https://aging.jmir.org/2026/1/e84494
doi:10.2196/84494

© Ui-jae Hwang, Zongpan Li, Peng Xia, Tianxiang Fan, Arnold Wang, Oh-yun Kwon, Siu-ngor Fu. Originally published in JMIR Aging (https://aging.jmir.org), 24.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Subsequent Hypertension's Mediation of the Association Between Sleep Duration Trajectories and New-Onset Cardiovascular Disease: Population-Based Cohort Study

Hongwei Liu[1*], MD, PhD; Minghui Wu[1*], MD; Minheng Zhang[2], MD; Jian Pan[3], MD; Haiyan Wu[4], MD; Haixia Fan[4], MD, PhD

[1]Department of Neurology, Taiyuan City Central Hospital, taiyuan, China

[2]Department of Gerontology, The First People's Hospital of Jinzhong, Jinzhong, China

[3]Department of Anesthesiology, Shenzhen People's Hospital Longhua Branch, ShenZhen, China

[4]Department of Sleep Center, First Hospital of Shanxi Medical University, 85, Jiefang South Road, Taiyuan, China

[*]these authors contributed equally

Corresponding Author:
Haixia Fan, MD, PhD
Department of Sleep Center, First Hospital of Shanxi Medical University, 85, Jiefang South Road, Taiyuan, China

Abstract

Background: Prior research has rarely explored the link between sleep duration trajectories and cardiovascular disease (CVD) either in cross-sectional or longitudinal data.

Objective: We aimed to analyze the relationship between sleep duration trajectories and the incidence of CVD.

Methods: Data from 5603 participants in the China Health and Retirement Longitudinal Study (CHARLS) were used in the analysis. Sleep duration was self-reported at multiple time points, and group-based trajectory modeling identified distinct patterns of total and nocturnal sleep duration over time. To investigate the link among sleep duration trajectories and CVD, Cox proportional hazards models, restricted cubic splines, and mediation analyses were used, with sensitivity analyses ensuring the results' robustness.

Results: Four sleep duration trajectories were identified, including steady high, inverted U-shaped, steady low, and U-shaped. In fully adjusted models, an inverted U-shaped sleep trajectory was associated with a 34% higher risk of hypertension compared with the steady high group (total sleep: hazard ratio [HR]=1.34, 95% CI 1.07 - 1.69; *P*=.01). Notably, the inverted U-shaped sleep trajectory showed a 110% increased stroke risk during nocturnal periods compared to the steady high group (nocturnal sleep: HR=2.10, 95% CI 1.08 - 4.05; *P*=.03). The findings remained consistent across a range of sensitivity analyses. Hypertension was identified as a partial mediator of the association between sleep instability and stroke risk.

Conclusions: The findings indicate that an inverted U-shaped sleep pattern is associated with increased cardiovascular risk, partially mediated by hypertension, highlighting sleep stability as a potential preventive target.

(JMIR Aging 2026;9:e78914) doi:10.2196/78914

KEYWORDS

sleep duration trajectory; hypertension; cardiovascular disease; stroke; heart disease; longitudinal study; mediation analysis

Introduction

Due to longer life expectancy and declining fertility rates, the global population aged 60 years and older is expected to rise from 12% to 22% between 2015 and 2050 [1]. China, one of the fastest-aging countries, will see the percentage of people aged >60 years grow from 12.4% in 2010 to 28% by 2040, reaching around 397 million [2]. Cardiovascular diseases (CVD) [3], including hypertension [4], stroke [5], and heart diseases [6], are particularly significant, as they represent leading causes of morbidity and mortality in older adults [7]. As the aging population increases, the burden of CVD will intensify, placing additional pressure on health care systems globally. In light of this, strategies to improve cardiovascular health and promote healthy aging are essential to reduce the impact of these diseases and enhance the quality of life for older adult individuals.

Insufficient or excessive sleep has been linked to adverse outcomes [8], including mood disorders [9], cognitive frailty [10], cancers [11], and even mortality [12]. Recent studies have shown that sleep abnormalities, including both deprivation and excessive sleep, negatively impact physiological functions and notably increase the risk of CVDs. Both short and long sleep durations are strongly linked to an elevated risk of coronary heart disease [13] and stroke [14]. Short sleep duration is

XSL•FO
RenderX

strongly associated with a higher incidence of hypertension, while the link between long sleep duration or difficulty falling asleep and hypertension remains inconclusive [15]. Although these studies explored the link between sleep duration and CVD, their conclusions were mixed. Most research relies on cross-sectional surveys and neglects the continuous development of sleep duration, offering an incomplete view of how it changes throughout life. This gap is particularly pertinent in middle-aged and older adults, where sleep patterns often change due to physiological, psychological, and social factors. Tracking these changes over time, rather than using a single measurement, is crucial for understanding their role in CVD development.

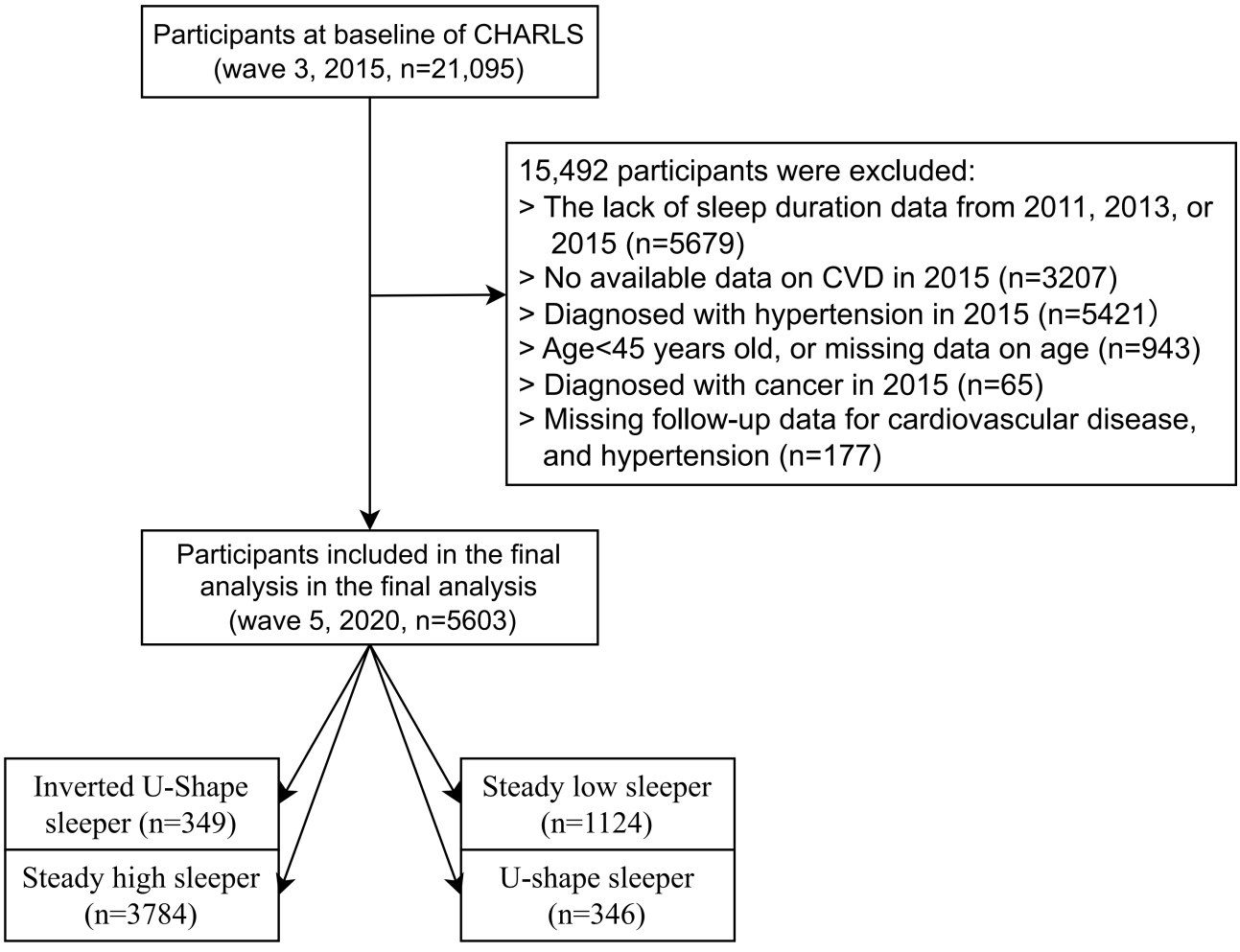
In light of these concerns, this study seeks to investigate the association between sleep duration trajectories and CVD outcomes in Chinese middle-aged and older adults (the China Health and Retirement Longitudinal Study [CHARLS]). By using a longitudinal dataset, we aim to identify distinct sleep trajectories over time and examine their potential role in the onset of CVD. This research not only contributes to the growing body of literature on sleep and health outcomes but also has important implications for public health strategies aimed at improving the quality of life and reducing the burden of CVD among the aging population.

Methods

Study Design and Participants

This nationally representative longitudinal survey, CHARLS, is designed to explore the health and economic status of Chinese people aged 45 years and older [16]. The CHARLS study's design, methods, and participant information have been extensively described in prior publications [16]. The 2011 baseline survey (Wave 1) involved 17,708 participants from over 10,000 households, with follow-up surveys conducted in 2013 (Wave 2), 2015 (Wave 3), 2018 (Wave 4), and 2020 (Wave 5). The research adhered to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) standards for reporting observational epidemiological studies. Data from Waves 1, 2, and 3 (2011, 2013, and 2015) were used in this study to track sleep duration trajectories. The 2015 Wave 3 of the CHARLS initially assessed 21,095 participants. Participants were excluded sequentially for the following reasons: missing sleep duration data in 2011, 2013, or 2015 (n=5679); lack of CVD data in 2015 (n=3207); preexisting hypertension in 2015 (n=5421); age <45 years or missing age information (n=943); diagnosis of cancer in 2015 (n=65); and missing follow-up information on CVD or hypertension (n=177). With these exclusions, 5603 participants were included in the final longitudinal analysis, which was followed up through Wave 5 (2020) (Figure 1).

Figure 1. Flowchart of participants' selection. CVD: cardiovascular disease; CHARLS: China Health and Retirement Longitudinal Study.



XSL•FO
RenderX

Measurements of Sleep Duration

Sleep duration trajectories refer to patterns of change in an individual's total sleep duration, which includes both nighttime sleep and daytime naps, over a specific period [17]. By using longitudinal trajectories, we were able to capture long-term sleep behavior patterns, including the consistency and changes in sleep duration over multiple years, providing a more detailed characterization of sleep patterns than any single metric. Detailed data on nighttime sleep and daytime naps were obtained from the 2011, 2013, and 2015 CHARLS waves, and the total sleep duration over 24 hours was calculated for each participant at each wave. The selection of these 3 waves was intentional, as they cover all sleep assessments available before the 2015 landmark, ensuring complete observation of exposure information before follow-up starts. Moreover, latent class and longitudinal trajectory models require at least 3 repeated measurements to reliably distinguish stable, monotonic, and nonlinear patterns, meaning that using only the 2011 and 2013 waves would not provide sufficient temporal variation for meaningful trajectory identification. Surveys conducted in 2018 or 2020, which are part of the post-2015 waves, were excluded due to methodological considerations. Data on sleep duration collected after 2015 may have been influenced by intervening hypertension or cardiovascular events, leading to misclassification of exposure after the outcome and violating the required temporal order for valid conclusions. By extending the trajectory window to 2018, substantial survivor bias would be introduced, as only those who remained event-free and continued follow-up until 2018 could contribute full sleep data, resulting in a healthier, nonrepresentative sample. Focusing trajectory construction on the pre–follow-up period (2011 - 2015) offers the temporal resolution required to capture significant sleep dynamics while upholding a valid exposure-outcome framework for subsequent cardiovascular evaluations. Applying latent class mixed models (LCMM), we identified 4 distinct sleep duration trajectory groups, including inverted U-shaped sleeper, steady low sleeper, steady high sleeper, and U-shaped sleeper. Each group represents a unique sleep pattern and trend. Inverted U-shaped sleepers show an initial increase in sleep duration, followed by a peak and a gradual decrease, potentially reflecting changes in sleep patterns during different physiological or life stages in middle-aged and older adults. Steady low sleepers maintain consistently low sleep durations with minimal fluctuation, which may indicate chronic sleep deprivation, potentially linked to cognitive decline, mental health issues, and other health concerns. Steady high sleepers consistently maintain higher sleep durations, which could reflect healthy sleep habits or specific health needs, such as restorative sleep, although prolonged high sleep duration may also be associated with chronic diseases or health risks, including metabolic disorders. Finally, U-shaped sleepers exhibit a decrease in sleep duration followed by an increase, forming a U-shaped trajectory, which may reflect dynamic changes in sleep needs across different age groups or health conditions, highlighting physiological, psychological, and environmental adaptations during the aging process.

Outcome Identification

Data on cardiovascular outcomes came from self-reported physician diagnoses collected in Waves 4 (2018) and 5 (2020). In accordance with established epidemiological practice, CVD was defined strictly as the first occurrence of stroke or heart disease, and hypertension was treated as a separate end point rather than included within the CVD composite [18-20]. Participants were classified as having incident CVD if they answered "yes" to either of the following physician-diagnosis questions, "Have you ever been diagnosed by a doctor with a heart attack, angina, coronary heart disease, heart failure, or other cardiovascular problems?" or "Have you received a diagnosis of a stroke from a doctor?" [18-20]. The event date was defined as the first survey wave (2018 or 2020) at which the condition was reported. If the same condition was reported in later waves, these were treated as duplicate events, and only the earliest report was used for analysis. When concurrent events (stroke and heart disease reported in the same wave) occurred, the participant was counted as having experienced a CVD event, with the shared date assigned as the event time. Each individual could contribute a maximum of one event to the hypertension end point and one event to the cardiovascular end points. In defining CVD as a combination of stroke and heart disease, only the first diagnosis of either was acknowledged as the incident CVD event, with subsequent reports considered duplicates.

In Waves 4 (2018) and 5 (2020), data regarding hypertension were obtained through a standardized self-reported physician-diagnosis questionnaire. Instead of incorporating it into the CVD composite, hypertension was addressed as an independent end point. Participants were classified as having incident hypertension if they answered "yes" to the question, "Have you been told by a doctor that you have high blood pressure?" The initial survey wave (2018 or 2020) that recorded a physician's diagnosis of hypertension was used to define the event date. Subsequent reports were regarded as duplicates, and only the earliest one was included in the analysis. Each participant was limited to contributing one hypertension event, irrespective of their classification for stroke, heart disease, or composite CVD.

Covariates

Covariates were selected based on established literature and clinical expertise. These covariates included age, gender, marital status, education level, smoking status, alcohol consumption, hypertension, dyslipidemia, heart disease, obesity, systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate, waist circumference, waist-to-height ratio, high-density lipoprotein cholesterol, total cholesterol, hemoglobin A_{1c} (HbA_{1c}), triglycerides, low-density lipoprotein cholesterol, creatinine, blood urea nitrogen (BUN), fasting plasma glucose (FPG), BMI, and estimated glomerular filtration rate (eGFR). Hypertension was defined as a self-reported physician diagnosis, use of antihypertensive medications, and/or an average SBP of ≥140 mm Hg and/or DBP of ≥90 mm Hg [21]. Diabetes mellitus (DM) was defined as FPG of ≥126 mg/dL or HbA_{1c} level of ≥6.5%, and/or a self-reported physician diagnosis, and/or the use of glucose-lowering medications. Pre-DM was identified

XSL•FO
RenderX

by FPG levels of 100 - 125 mg/dL or HbA_{1c} level between 5.7% and 6.4%. Individuals without DM or pre-DM were classified as having normal glucose regulation [22]. Dyslipidemia was diagnosed based on a self-reported physician diagnosis, the use of lipid-lowering medications, and/or triglycerides >150 mg/dL, high-density lipoprotein cholesterol level of <40 mg/dL, low-density lipoprotein cholesterol level of >160 mg/dL, or total cholesterol level of >240 mg/dL [23]. Chronic kidney disease was defined as a self-reported physician diagnosis or an eGFR of <60 mL/min/1.73 m^2, in accordance with previous CHARLS studies [24]. Obesity was assessed by calculating BMI using the formula: weight (kg) / height^2 (m^2) [25]. Consistent with our earlier research [26,27], socioeconomic status (SES) was assessed through 4 indicators, including yearly household income, educational achievement, employment status, and medical insurance coverage. Family annual income was determined by the total household income from the prior year and divided into 2 groups, above and below the median, with the sample median set at US $2877.7 or CNY 18,130.00. Education was sorted into 3 tiers, including below high school, high school or equivalent (General Educational Development), and college or more advanced. Educational attainment was classified into 3 categories, including junior high school or below, senior high school, and college or above. Employment status was sorted into employed (covering those with paid employment, self-employment, or retired) and unemployed. Insurance coverage was divided into 2 categories, insured, which includes public or private health insurance, and uninsured.

Statistical Analyses

For continuous variables that did not conform to a normal distribution, medians and IQRs were used, and categorical variables were displayed as frequencies and percentages. The Kruskal-Wallis rank-sum test was used to compare nonnormally distributed continuous variables across the 4 groups, and the chi-square test was used for categorical variables. With the assumption that the missing data followed a random pattern, we used *multiple imputation by chained equations* (MICE) in R (R Foundation for Statistical Computing) to estimate the missing covariates. Incorporating all analytic covariates and auxiliary variables, the imputation model aimed to maximize the use of information. To achieve algorithmic stability and convergence, 5 complete datasets were generated, each undergoing 50 iterations. Predictive mean matching ("pmm") was used for continuous variables, logistic regression ("logreg") for binary variables, multinomial models for nominal categorical variables, and proportional-odds models for ordinal variables. The evaluation of convergence involved inspecting trace plots for the mean and SD of the imputed values. Each imputed dataset was analyzed with Cox proportional hazards models, and the parameter estimates were merged using Rubin rules to generate pooled coefficients and standard errors. Using the *LCMM* package in R, we applied LCMMs to study longitudinal variations in total and nighttime sleep duration. Models with different polynomial specifications (linear, quadratic, and cubic) were compared to evaluate goodness of fit, and the optimal model was determined according to the principle of Bayesian Information Criterion minimization, ensuring both statistical adequacy and practical interpretability (Table S1 and S2 in Multimedia Appendix 1). To maintain enough sample representation in each trajectory and prevent overfitting, each latent class needed to account for at least 5% of the total sample. According to these guidelines, a 4-class model provided the optimal fit, resulting in 4 distinct sleep trajectory patterns, including inverted U-shaped sleepers, steady low sleepers, steady high sleepers, and U-shaped sleepers. Using Schoenfeld residuals, the proportional hazards assumption in the Cox models was evaluated, and no significant violations were identified (all P>.05). Variance inflation factors were used to evaluate potential multicollinearity among covariates, and since all values were below 5, it indicated no serious collinearity. These diagnostic checks validated that the Cox proportional hazards models complied with crucial assumptions and were apt for the analyses. Potential confounders were pinpointed using prior epidemiological evidence and clinical relevance, encompassing demographic features, lifestyle practices, and metabolic indicators. Using Cox proportional hazards models, we investigated the relationship between sleep patterns and hypertension, as well as subsequent cardiovascular events, including stroke and heart disease. To elucidate the progressive impact of confounding, 3 hierarchical Cox proportional hazards models were fitted. Model 1 served as the crude model; Model 2 adjusted for sociodemographic and lifestyle variables (age, sex, education level, marital status, smoking, and alcohol use); and Model 3 extended the adjustment to biological variables, namely dyslipidemia, BUN, eGFR, DM, and obesity status, thereby capturing potential physiological confounding. Follow-up time from Wave 4 (2018) to Wave 5 (2020) or event occurrence served as the analytic time scale. Participants who remained free of cardiovascular events were administratively censored at their last observed follow-up. Restricted cubic spline models were applied to examine the potential nonlinear associations between sleep duration and cardiovascular outcomes. Based on established recommendations, 4 knots were set at the 5th, 35th, 65th, and 95th percentiles of the sleep-duration distribution to fit restricted cubic spline models. To enhance clinical interpretability, contrasts were determined at sleep durations of 6, 7, 9, and 10 hours. These values were selected based on both clinical relevance and the empirical distribution observed in the trajectory patterns. These values were selected a priori because they correspond to commonly used sleep duration categories in epidemiological research, with approximately 6 hours representing short sleep, 7 - 8 hours reflecting recommended sleep duration, and 9 - 10 hours representing extended sleep. Using these established thresholds allows the spline-estimated associations to be interpreted against clinically meaningful reference points without imposing additional assumptions on the functional form. An exploratory mediation analysis was conducted to see if hypertension could potentially serve as an intermediate pathway between total or nocturnal sleep trajectories and cardiovascular outcomes. Using nonparametric bootstrapping, direct and indirect effects were estimated, but the temporal structure of the data hinders causal interpretation, indicating that the findings are exploratory. For this analysis, the *mediation* package was used with 1000 iterations of simulation. To mitigate potential reverse causality and enhance the robustness of our findings, participants who experienced cardiovascular events within the first 2 years of

XSL•FO
RenderX

follow-up were excluded from the analysis. In addition, to account for uncertainty in sleep trajectory assignment, sensitivity analyses used misclassification-adjusted Cox regression models with posterior-probability weighting. Further sensitivity analyses included additional adjustments for socioeconomic status to evaluate the possible influence of leftover socioeconomic confounding. R software (version 4.4.1) facilitated all statistical analyses, with a 2-sided *P* value less than .05 considered statistically significant.

Ethical Considerations

The Biomedical Ethics Review Committee of Peking University (IRB00001052-11015) approved the study, and all participants provided written informed consent. All data used in this study were obtained from publicly available, deidentified datasets. These datasets were generated by original investigators who obtained informed consent from participants and conducted their studies in accordance with the Declaration of Helsinki and applicable institutional review board approvals. No new human participants were enrolled for this research, and no additional ethical approval was required for this secondary analysis. All data accessed are publicly available and fully anonymized prior to deposition. No personally identifiable information (such as names, addresses, contact details, or unique personal identifiers) was used or disclosed in any part of the analysis. We strictly adhered to the terms of use and data sharing policies of the original repositories. All analyses were conducted on aggregated or deidentified data, ensuring that individual privacy and confidentiality were fully maintained throughout the study. As this research involved the analysis of preexisting, deidentified data, no participants were directly involved in study procedures and no compensation was provided or required as part of this work.

Results

Baseline Characteristics of Participants and Sleep Duration Trajectories

Table 1 summarized the baseline demographic and clinical characteristics of 5603 participants, grouped by total sleep duration trajectories, including inverted U-shaped (n=349), steady low (n=1124), steady high (n=3784), and U-shaped (n=346). Figures 2A and 2B further illustrated the longitudinal trends in total and nocturnal sleep duration from 2011 to 2015, confirming 4 distinct trajectories. The overall median age was 60 (IQR 53.00 - 66.00) years, with group-specific medians ranging from 62 (IQR 52.00 - 68.00) years in the inverted U-shaped group to 61 (IQR 53.00 - 68.00) years in the U-shaped group. The cohort included 2765 (47.74%) males and 2928 (52.26%) females, with the proportion of females highest in the U-shaped group (65.90%).

XSL•FO
RenderX

Table . Baseline characteristics of participants categorized by total sleep duration trajectories.

Characteristic	Overall (N=5603)	Inverted U-shape sleeper (n=349)	Steady high sleeper (n=3784)	Steady low sleeper (n=1124)	U-shaped sleeper (n=346)	*P* value[a]
Sleep (total), median (IQR)	7.00 (6.00-8.33)	4.50 (3.00-6.00)	8.00 (7.00-9.00)	5.00 (4.00-6.00)	7.00 (5.50-9.00)	<.001
Sleep (night), median (IQR)	6.50 (5.00-8.00)	4.00 (3.00-6.00)	7.00 (6.00-8.00)	5.00 (4.00-6.00)	7.00 (5.00-8.00)	<.001
Age (years), median (IQR)	60.00 (53.00-66.00)	62.00 (56.00-68.00)	59.00 (53.00-65.00)	61.00 (54.00-67.00)	61.00 (53.00-68.00)	<.001
Sex, n (%)						<.001
Male	2675.00 (47.74)	136.00 (38.97)	1963.00 (51.88)	458.00 (40.75)	118.00 (34.10)	
Female	2928.00 (52.26)	213.00 (61.03)	1,821.00 (48.12)	666.00 (59.25)	228.00 (65.90)	
Marital status, n (%)						<.001
Married and living with spouse	4977.00 (88.83)	304.00 (87.11)	3,418.00 (90.33)	958.00 (85.23)	297.00 (85.84)	
Other	626.00 (11.17)	45.00 (12.89)	366.00 (9.67)	166.00 (14.77)	49.00 (14.16)	
Education level, n (%)						.008
Junior high school and below	886.00 (15.81)	61.00 (17.48)	561.00 (14.83)	195.00 (17.35)	69.00 (19.94)	
Senior high school	326.00 (5.82)	13.00 (3.72)	240.00 (6.34)	62.00 (5.52)	11.00 (3.18)	
College and above	4,391.00 (78.37)	275.00 (78.80)	2,983.00 (78.83)	867.00 (77.14)	266.00 (76.88)	
Smoking status, n (%)						<.001
Never smoked	3,298.00 (58.86)	233.00 (66.76)	2120.00 (56.03)	709.00 (63.08)	236.00 (68.21)	
Smoking at the time of this writing	1659.00 (29.61)	78.00 (22.35)	1209.00 (31.95)	291.00 (25.89)	81.00 (23.41)	
Ever smoked	646.00 (11.53)	38.00 (10.89)	455.00 (12.02)	124.00 (11.03)	29.00 (8.38)	
Drinking status, n (%)						.005
Never drink	3774.00 (67.36)	245.00 (70.20)	2506.00 (66.23)	777.00 (69.13)	246.00 (71.10)	
Drinking at the time of this writing	1493.00 (26.65)	73.00 (20.92)	1063.00 (28.09)	277.00 (24.64)	80.00 (23.12)	
Ever drunk	336.00 (6.00)	31.00 (8.88)	215.00 (5.68)	70.00 (6.23)	20.00 (5.78)	
DM[b], n (%)						.05
No	3558.00 (63.50)	227.00 (65.04)	2378.00 (62.84)	747.00 (66.46)	206.00 (59.54)	
Yes	2045.00 (36.50)	122.00 (34.96)	1406.00 (37.16)	377.00 (33.54)	140.00 (40.46)	
Dyslipidemia, n (%)						.17
No	4044.00 (72.18)	260.00 (74.50)	2703.00 (71.43)	836.00 (74.38)	245.00 (70.81)	
Yes	1559.00 (27.82)	89.00 (25.50)	1081.00 (28.57)	288.00 (25.62)	101.00 (29.19)	
Obesity, n (%)						<.001
Normal weight	3119.00 (55.67)	229.00 (65.62)	2030.00 (53.65)	660.00 (58.72)	200.00 (57.80)	
Overweight	1833.00 (32.71)	94.00 (26.93)	1293.00 (34.17)	343.00 (30.52)	103.00 (29.77)	
Obesity	651.00 (11.62)	26.00 (7.45)	461.00 (12.18)	121.00 (10.77)	43.00 (12.43)	

XSL•FO
RenderX

Characteristic	Overall (N=5603)	Inverted U-shape sleeper (n=349)	Steady high sleeper (n=3784)	Steady low sleeper (n=1124)	U-shaped sleeper (n=346)	*P* value[a]
Creatinine (mg/dl), median (IQR)	0.85 (0.70-1.79)	0.84 (0.68-1.23)	0.86 (0.70-1.93)	0.84 (0.69-1.32)	0.85 (0.67-1.79)	<.001
BUN[c] (mg/dl), median (IQR)	15.69 (12.89-19.89)	15.41 (12.61-19.89)	15.69 (13.17-20.17)	15.69 (12.89-19.33)	15.41 (12.61-20.17)	.04
eGFR[d] (ml/min/1.73m2), median (IQR)	86.80 (38.43-98.00)	86.51 (57.05-96.28)	87.27 (33.43-98.43)	86.48 (55.74-97.24)	82.86 (33.89-97.63)	.26
Hypertension incident, n (%)						.11
No	5165.00 (92.18)	313.00 (89.68)	3508.00 (92.71)	1025.00 (91.19)	319.00 (92.20)	
Yes	438.00 (7.82)	36.00 (10.32)	276.00 (7.29)	99.00 (8.81)	27.00 (7.80)	
Stroke incident, n (%)						.02
No	5503.00 (98.22)	340.00 (97.42)	3731.00 (98.60)	1096.00 (97.51)	336.00 (97.11)	
Yes	100.00 (1.78)	9.00 (2.58)	53.00 (1.40)	28.00 (2.49)	10.00 (2.89)	
Heart incidence, n (%)						.11
No	5165.00 (92.18)	313.00 (89.68)	3508.00 (92.71)	1025.00 (91.19)	319.00 (92.20)	
Yes	438.00 (7.82)	36.00 (10.32)	276.00 (7.29)	99.00 (8.81)	27.00 (7.80)	
CVD[e] incident, n (%)						.006
No	5081.00 (90.68)	304.00 (87.11)	3465.00 (91.57)	1002.00 (89.15)	310.00 (89.60)	
Yes	522.00 (9.32)	45.00 (12.89)	319.00 (8.43)	122.00 (10.85)	36.00 (10.40)	

[a]Kruskal-Wallis rank sum test; Pearson χ^2 test; Data were presented as mean (SD), median (IQR), or as n (%).

[b]DM: diabetes mellitus.

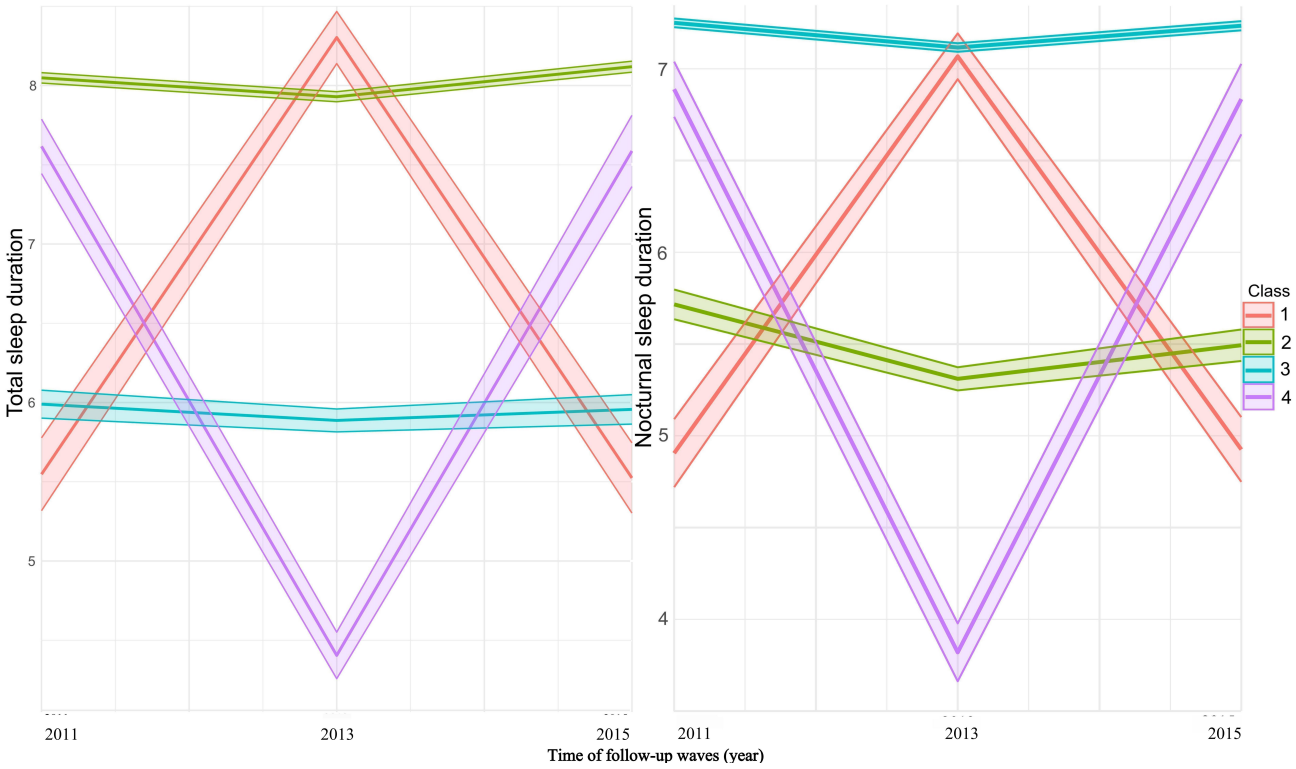
[c]BUN: blood urea nitrogen.

[d]eGFR: estimated glomerular filtration rate.

[e]CVD: cardiovascular disease.

XSL•FO
RenderX

Figure 2. Total sleep duration trajectories from 2011 to 2015 in Chinese middle-aged and older adults. (**A**) Total sleep duration trajectories. (**B**) Nocturnal sleep duration trajectories.



Both total sleep duration and nocturnal sleep duration differed significantly across sleep trajectory groups (*P*<.001). Participants in the inverted U-shaped sleep group exhibited the shortest median total sleep duration (4.5, IQR 3.0 - 6.0 h) and nocturnal sleep duration (4.0, IQR 3.0 - 6.0 h), whereas those in the steady high sleep group showed the longest median total sleep duration (8.0, IQR 7.0 - 9.0 h) and nocturnal sleep duration (7.0, IQR 6.0 - 8.0 h). Participants differed significantly across trajectory groups regarding age, gender, marital status, education level, smoking status, drinking status, creatinine, BUN, and obesity (all *P*<.05; Table 1). Although no significant differences were observed in estimated glomerular filtration rate (*P*=.26), DM (*P*=.05), dyslipidemia (*P*=.17), or hypertension (*P*=.11; Table 1).

Association of Sleep Duration Trajectories With New-Onset Hypertension

In the course of the follow-up, 1123 new hypertension cases were identified, making up 20% (Table 2). Differences were observed in the onset of hypertension among the sleep trajectory groups. The distribution of total sleep trajectories (Figure S1A in Multimedia Appendix 2) was 19% for steady high sleepers, 22% for steady low sleepers, 24% for inverted U-shaped sleepers, and 20% for U-shaped sleepers. Nocturnal sleep trajectories (Figure S1B in Multimedia Appendix 2) exhibited a similar pattern, with incidences of 24%, 21%, 19%, and 22% across the respective categories. For total sleep duration trajectories, compared with the steady high group (reference), participants in the inverted U-shaped group had a significantly higher risk of developing hypertension (hazard ratio [HR]=1.34, 95% CI 1.07 - 1.69; *P*=.01) in the fully adjusted model (Model 3; Table 2). The steady low and U-shaped groups showed modest but nonsignificant increases in hypertension risk HR=1.13, 95% CI 0.97 - 1.31; *P*=.11; and HR=0.98, 95% CI 0.76 - 1.26; *P*=.87, respectively; Table 2). Figure 3A illustrates the association between total sleep duration and the risk of incident hypertension. The overall association was statistically significant (for overall *P*=.04), and the test for nonlinearity also reached significance (for nonlinear *P*=.04).

XSL•FO
RenderX

Table . Association of the sleep trajectories and new-onset hypertension.

Characteristic	N	Event, n (%)	Model 1[a]		Model 2[b]		Model 3[c]	
			HR[d] (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Total sleep trajectories	5603	1123 (20.04)						
Steady high sleeper	3784	726 (19)	Reference[e]		Reference		Reference	
Steady low sleeper	1124	242 (22)	1.14 (0.98-1.31)	.09	1.11 (0.96-1.29)	.17	1.13 (0.97-1.31)	.11
Inverted U-shaped sleeper	349	85 (24)	1.31 (1.05-1.64)	.02	1.26 (1.01-1.59)	.04	1.34 (1.07-1.69)	.01
U-shaped sleeper	346	70 (20)	1.05 (0.82-1.34)	.69	1.03 (0.80-1.31)	.84	0.98 (0.76-1.26)	.87
Nocturnal sleep trajectories	5603	1123 (20.04)						
Steady high sleeper	3849	744 (19)	Reference		Reference		Reference	
Steady low sleeper	1073	224 (21)	1.09 (0.94-1.26)	.26	1.07 (0.92-1.24)	.41	1.06 (0.91-1.24)	.43
Inverted U-shaped sleeper	386	91 (24)	1.24 (1.00-1.54)	.05	1.19 (0.96-1.48)	.12	1.26 (1.01-1.58)	.04
U-shaped sleeper	295	64 (22)	1.13 (0.87-1.45)	.37	1.10 (0.85-1.42)	.48	1.09 (0.84-1.41)	.50

[a]Model 1: crude model.

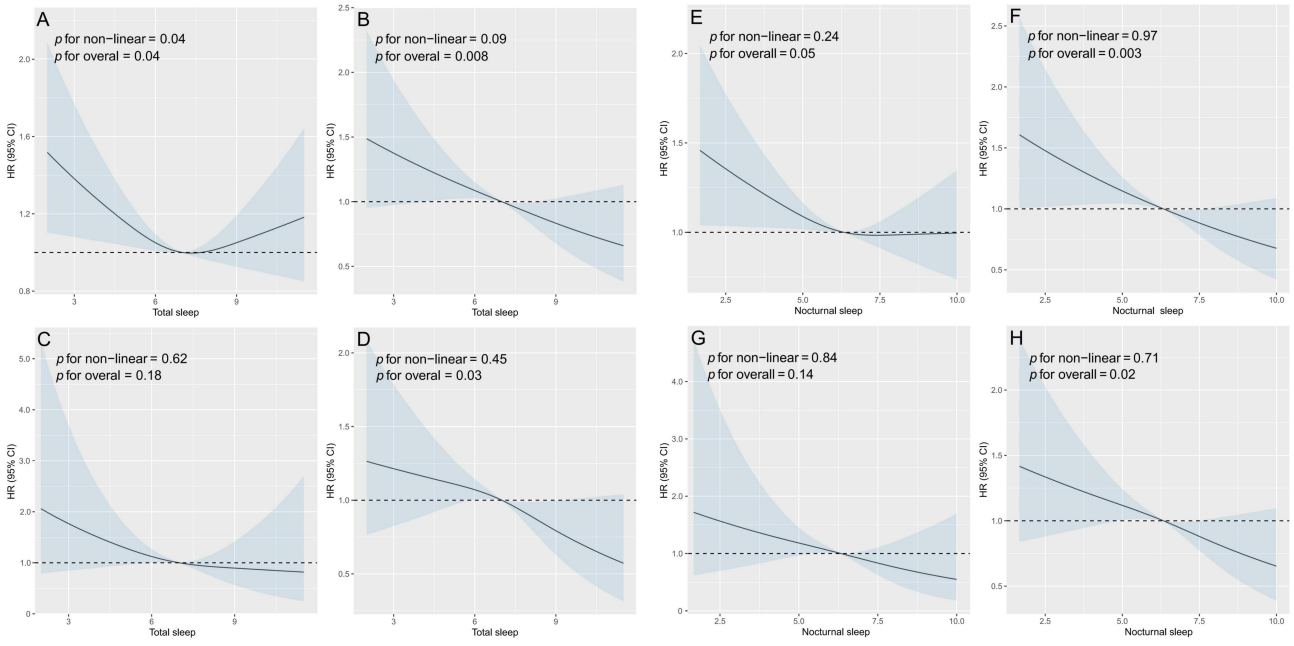
[b]Model 2: adjusted for age, gender, education level, marital status, drinking status, and smoking status.

[c]Model 3: adjusted for age, gender, education level, marital status, drinking status, smoking status, dyslipidemia (current use of lipid-lowering medications, and/or triglycerides>150 mg/dL, high-density lipoprotein cholesterol [HDL-C]<40 mg/dL, low-density lipoprotein cholesterol [LDL-C]>160 mg/dL, or total cholesterol [TC]>240 mg/dL), blood urea nitrogen, estimated glomerular filtration rate, diabetes mellitus, and obesity status.

[d]HR: hazard ratio.

[e]Participants with steady high sleep duration, which is used as the reference for comparison when reporting hazard ratios (HR) for the other sleep trajectory groups.

Figure 3. Curvilinear association between total sleep/nocturnal sleep and risk of cardiovascular outcomes. (A, E) New onset of hypertension; (B, F) new onset of cardiovascular diseases (CVD); (C, G) new onset of stroke; and (D, H) new onset of heart diseases. HR: hazard ratio; CVD: cardiovascular disease.



XSL•FO
RenderX

For nocturnal sleep duration trajectories, similar patterns were observed. The inverted U-shaped group had a significantly elevated risk of new-onset hypertension compared to the steady-high group (HR=1.26, 95% CI 1.01 - 1.58; *P*=.04 in Model 3). No significant associations were found for the steady low (HR=1.13, 95% CI 0.97 - 1.30; *P*=.11) or U-shaped group (HR=0.97, 95% CI 0.74 - 1.26; *P*=.81). Figure 3E depicts the association between nocturnal sleep duration and hypertension risk. The overall association was statistically significant (*P* for overall=.05), whereas the test for nonlinearity was not significant (*P* for nonlinear=.24).

Association of Sleep Duration Trajectories With New-Onset CVD

Over the follow-up duration, 522 participants, which is 10.17%, developed new-onset CVD. Different sleep trajectory groups also experienced varying rates of new CVD onset. The sleep trajectory analysis (Figure S1C in Multimedia Appendix 2) showed that 8.4% of participants were classified as steady high sleepers, 11.0% as steady low sleepers, 13.0% as inverted U-shaped sleepers, and 10.0% as U-shaped sleepers. For nocturnal sleep trajectories (Figure S1D in Multimedia Appendix 2), the corresponding rates were 12%, 11%, 8.5%, and 11%, respectively. Compared to steady high sleepers (reference), those categorized as inverted U-shaped showed an increased chance of experiencing new-onset CVD. Compared to steady high sleepers, the inverted U-shaped group showed a nearly statistically significant elevated risk in the fully adjusted model, with an HR of 1.46 (95% CI 1.07 - 2.00; *P*=.02) in Model 3 (Table 3). Higher cardiovascular risk was observed in the U-shaped and steady low groups, but the associations were not statistically significant (HR=1.08 95% CI 0.76 - 1.54; *P*=.66; HR=1.22, 95% CI 0.99 - 1.50; *P*=.07; Table 3). In Figure 3B, the correlation between total sleep duration and the occurrence of new CVD is shown. The overall association was statistically significant (*P* for overall=.008), while the test for nonlinearity did not indicate evidence of a nonlinear relationship (*P* for nonlinearity=.09).

Table . Association of the sleep trajectories and new-onset cardiovascular disease.

Characteristic	N	Event, n (%)	Model 1[a]		Model 2[b]		Model 3[c]	
			HR[d] (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Total sleep trajectories	5603	522 (9.32)						
Steady high sleeper	3784	319 (8.4)	Reference[e]		Reference		Reference	
Steady low sleeper	1124	122 (11)	1.31 (1.06-1.61)	.01	1.21 (0.98-1.49)	.08	1.22 (0.99-1.50)	.07
Inverted U-shaped sleeper	349	45 (13)	1.58 (1.15-2.15)	.004	1.40 (1.02-1.92)	.04	1.46 (1.07-2.00)	.02
U-shaped sleeper	346	36 (10)	1.25 (0.88-1.76)	.21	1.11 (0.79-1.57)	.55	1.08 (0.76-1.54)	.65
Nocturnal sleep trajectories	5603	522 (9.32)						
Steady high sleeper	3849	329 (8.5)	Reference		Reference		Reference	
Steady low sleeper	1073	117 (11)	1.30 (1.05-1.60)	.02	1.20 (0.97-1.48)	.10	1.21 (0.98-1.50)	.08
Inverted U-shaped sleeper	386	45 (12)	1.40 (1.02-1.91)	.04	1.24 (0.90-1.70)	.18	1.30 (0.95-1.79)	.10
U-shaped sleeper	295	31 (11)	1.24 (0.85-1.79)	.26	1.11 (0.76-1.60)	.59	1.11 (0.77-1.61)	.57

[a]Model 1: crude model.

[b]Model 2: adjusted for age, gender, education level, marital status, drinking status, and smoking status.

[c]Model 3: adjusted for age, gender, education level, marital status, drinking status, smoking status, dyslipidemia (current use of lipid-lowering medications, and/or triglycerides>150 mg/dL, high-density lipoprotein cholesterol [HDL-C]<40 mg/dL, low-density lipoprotein cholesterol [LDL-C]>160 mg/dL, or total cholesterol [TC]>240 mg/dL), blood urea nitrogen, estimated glomerular filtration rate, diabetes mellitus, and obesity status.

[d]HR: hazard ratio.

[e]Participants with steady high sleep duration, which is used as the reference for comparison when reporting hazard ratios (HR) for the other sleep trajectory groups.

When evaluating nocturnal sleep duration trajectories, the overall association patterns remained consistent. Compared with steady high sleepers (reference), the steady low sleep trajectory was associated with an elevated risk of incident CVD in the fully adjusted model, although the association did not reach statistical significance (HR=1.21, 95% CI 0.98 - 1.50;

XSL•FO
RenderX

P=.08; Table 3). Similarly, individuals with an inverted U-shaped sleep trajectory showed a higher risk of new-onset CVD, but this association was marginal and not statistically significant (HR=1.30, 95% CI 0.95 - 1.79; *P*=.10). No significant association was observed for the U-shaped sleep trajectory (HR=1.11, 95% CI 0.77 - 1.61; *P*=.57; Table 3). Figure 3F depicts the association between nocturnal sleep duration and heart disease risk. The nonlinearity test did not indicate evidence of a nonlinear relationship (*P* for nonlinear=.96), whereas the overall association was highly significant (*P* for overall=.003), mirroring the pattern observed for total sleep duration.

Association of Sleep Duration Trajectories With New-Onset Stroke

In the course of the study, 100 individuals, representing 1.77% of the total, experienced a new-onset stroke. The incidence of new-onset stroke differed notably across sleep trajectory groups. For total sleep trajectories, the incidence was 1.4% among steady high sleepers, 2.5% among steady low sleepers, 2.6% among inverted U-shaped sleepers, and 2.9% among U-shaped sleepers (Figure S1E in Multimedia Appendix 2). The pattern observed in nocturnal sleep trajectories (Figure S1F in Multimedia Appendix 2) was similar, with incidences recorded at 2.8%, 2.7%, 1.4%, and 2.7%. Individuals with inverted U-shaped, steady-low, or U-shaped sleep patterns showed an increased risk of new-onset stroke compared to those with consistently high sleep durations. For total sleep duration trajectories, compared with steady high sleepers (reference), individuals in the steady low sleep trajectory exhibited a significantly higher risk of incident stroke in the fully adjusted model (HR=1.71, 95% CI: 1.07 - 2.72; *P*=.02; Table 4). In contrast, participants with an inverted U-shaped sleep trajectory showed an elevated but statistically nonsignificant risk of stroke (HR=1.85, 95% CI 0.91 - 3.79; *P*=.09). Similarly, no statistically significant association was observed for the U-shaped sleep trajectory (HR=1.77, 95% CI 0.87 - 3.60; *P*=.12; Table 4). Figure 3C illustrates the association between total sleep duration and new-onset stroke, with neither the overall association (*P* for overall=.18) nor the nonlinearity test (*P* for nonlinear=.62) reaching statistical significance.

XSL•FO
RenderX

Table . Association of the sleep trajectories and new-onset stroke.

Characteristic	Total (N)	Event, n (%)	Model 1[a]		Model 2[b]		Model 3[c]	
			HR[d] (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Total sleep trajectories	5603	100 (1.78)						
Steady high sleeper	3784	53 (1.4)	Reference[e]		Reference		Reference	
Steady low sleeper	1124	28 (2.5)	1.79 (1.13, 2.83)	.01	1.67 (1.05-2.66)	.03	1.71 (1.07-2.72)	.02
Inverted U-shaped sleeper	349	9 (2.6)	1.85 (0.91-3.75)	.09	1.68 (0.82-3.42)	.16	1.85 (0.91-3.79)	.09
U-shaped sleeper	346	10 (2.9)	2.09 (1.06-4.10)	.03	1.93 (0.97-3.82)	.06	1.77 (0.87-3.60)	.12
Nocturnal sleep trajectories	5603	100 (1.78)						
Steady high sleeper	3849	52 (1.4)	Reference		Reference		Reference	
Steady low sleeper	1073	29 (2.7)	2.02 (1.28-3.18)	.002	1.87 (1.18-2.96)	.008	1.83 (1.15-2.91)	.01
Inverted U-shaped sleeper	386	11 (2.8)	2.12 (1.11-4.06)	.02	1.93 (1.00-3.73)	.05	2.10 (1.08-4.05)	.03
U-shaped sleeper	295	8 (2.7)	2.02 (0.96-4.26)	.06	1.89 (0.89-4.00)	.10	1.96 (0.92-4.18)	.08

[a]Model 1: crude model.

[b]Model 2: adjusted for age, gender, education level, marital status, drinking status, and smoking status.

[c]Model 3: adjusted for age, gender, education level, marital status, drinking status, smoking status, dyslipidemia (current use of lipid-lowering medications, and/or triglycerides>150 mg/dL, high-density lipoprotein cholesterol [HDL-C]<40 mg/dL, low-density lipoprotein cholesterol [LDL-C]>160 mg/dL, or total cholesterol [TC]>240 mg/dL), blood urea nitrogen, estimated glomerular filtration rate, diabetes mellitus, and obesity status.

[d]HR: hazard ratio.

[e]Participants with steady high sleep duration, which is used as the reference for comparison when reporting hazard ratios (HR) for the other sleep trajectory groups.

For nocturnal sleep duration trajectories, the associations were more pronounced. Compared with steady high sleepers (reference), individuals in the steady low-sleep trajectory exhibited a significantly increased risk of new-onset stroke in the fully adjusted model (HR=1.83, 95% CI 1.15 - 2.91; *P*=.01; Table 4). In contrast, participants with an inverted U-shaped sleep trajectory also showed a significantly elevated risk of incident stroke (HR=2.10, 95% CI 1.08 - 4.05; *P*=.03). By comparison, although the U-shaped sleep trajectory was associated with a higher hazard ratio, the association did not reach statistical significance (HR=1.96, 95% CI 0.92 - 4.18; *P*=.08). Figure 3G illustrates the association between nocturnal sleep duration and new-onset stroke, with neither the overall association (*P* for overall=0.14) nor the nonlinearity test (*P* for nonlinear=0.84) reaching statistical significance.

Association of Sleep Duration Trajectories With New-Onset Heart Disease

Over the follow-up duration, 438 participants (7.8%) were found to have newly developed heart disease. The incidence of new-onset heart disease differed modestly across sleep trajectory groups. For total sleep trajectories (Figure S1G in Multimedia Appendix 2), the lowest incidence was observed among steady high sleepers (7.3%), followed by steady low sleepers (8.8%) and U-shaped sleepers (7.8%), whereas inverted U-shaped sleepers exhibited the highest incidence (10.0%). A similar

XSL•FO
RenderX

pattern was observed for nocturnal sleep duration trajectories (Figure S1H in Multimedia Appendix 2). The incidence of new-onset heart disease was 7.5% among steady high sleepers, increased to 8.5% among steady low sleepers, and peaked among inverted U-shaped sleepers (8.8%) before declining in the U-shaped group (7.8%). For total sleep duration trajectories, compared with steady high sleepers (reference), no statistically significant associations were observed across trajectory groups in the fully adjusted model. Specifically, the steady low sleep trajectory was associated with a modestly elevated but nonsignificant risk of heart disease (HR=1.14, 95% CI 0.90 - 1.43; *P*=.27; Table 5). Similarly, participants with an inverted U-shaped sleep trajectory showed an increased HR that did not reach statistical significance (HR=1.34, 95% CI: 0.94 - 1.90; *P*=.11). The U-shaped sleep trajectory was not associated with heart disease risk (HR=0.93, 95% CI 0.63 - 1.39; *P*=.73). Figure 3D illustrates the association between total sleep duration and heart disease risk. The overall association reached statistical significance (*P* for overall=.03); however, there was no evidence supporting a nonlinear relationship (*P* for nonlinear=.45).

Table . Association of the sleep trajectories and new-onset heart disease.

Characteristic	Total (N)	Event, n (%)	Model 1[a]		Model 2[b]		Model 3[c]	
			HR[d] (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Total sleep trajectories	5603	438 (7.82)						
Steady high sleeper	3784	276 (7.3)	Reference[e]		Reference		Reference	
Steady low sleeper	1124	99 (8.8)	1.22 (0.97-1.53)	.09	1.11 (0.88-1.40)	.36	1.14 (0.90-1.43)	.29
Inverted U-shaped sleeper	349	36 (10)	1.44 (1.02-2.04)	.04	1.27 (0.89-1.80)	.18	1.34 (0.94-1.90)	.11
U-shaped sleeper	346	27 (7.8)	1.07 (0.72-1.59)	.75	0.95 (0.64-1.41)	.79	0.93 (0.63-1.39)	.73
Nocturnal sleep trajectories	5603	438 (7.82)						
Steady high sleeper	3849	290 (7.5)	Reference		Reference		Reference	
Steady low sleeper	1073	91 (8.5)	1.13 (0.89-1.43)	.30	1.04 (0.82-1.32)	.72	1.06 (0.84-1.35)	.62
Inverted U-shaped sleeper	386	34 (8.8)	1.19 (0.83-1.69)	.35	1.04 (0.73-1.49)	.83	1.09 (0.76-1.56)	.63
U-shaped sleeper	295	23 (7.8)	1.03 (0.67-1.57)	.90	0.91 (0.60-1.40)	.68	0.91 (0.60-1.40)	.68

[a]Model 1: crude model.

[b]Model 2: adjusted for age, gender, education level, marital status, drinking status, and smoking status.

[c]Model 3: adjusted for age, gender, education level, marital status, drinking status, smoking status, dyslipidemia (current use of lipid-lowering medications, and/or triglycerides>150 mg/dL, high-density lipoprotein cholesterol [HDL-C]<40 mg/dL, low-density lipoprotein cholesterol [LDL-C]>160 mg/dL, or total cholesterol [TC]>240 mg/dL), blood urea nitrogen, estimated glomerular filtration rate, diabetes mellitus, and obesity status.

[d]HR: hazard ratio.

[e]Participants with steady high sleep duration, which is used as the reference for comparison when reporting hazard ratios (HR) for the other sleep trajectory groups.

XSL•FO
RenderX

Sensitivity Analysis

In the sensitivity analysis excluding participants who developed hypertension within the first 2 years of follow-up (Table S3 in Multimedia Appendix 1), compared with steady high sleepers (reference), individuals with an inverted U-shaped sleep trajectory exhibited a significantly increased risk in the fully adjusted total sleep trajectory model (HR=1.34, 95% CI 1.06 - 1.71; *P*=.02). In contrast, compared with the steady high sleepers (reference), the association remained statistically significant in the fully adjusted nocturnal sleep trajectory model (HR=1.28, 95% CI 1.02 - 1.62; *P*=.03; Table S3 in Multimedia Appendix 1). In the sensitivity analysis excluding early CVD within the first 2 years of follow-up (Table S4 in Multimedia Appendix 1), the association between the inverted U-shaped total sleep trajectory and incident CVD was weakened and did not reach statistical significance after full adjustment (HR=1.38, 95% CI 0.93 - 2.04; *P*=.11). Excluding strokes that occurred within the first 2 years of follow-up (Table S5 in Multimedia Appendix 1), the association for total sleep trajectories was attenuated. Specifically, although individuals with an inverted U-shaped total sleep trajectory showed an elevated HR for incident stroke, the association did not reach statistical significance in the fully adjusted model (HR=2.23, 95% CI 0.90 - 5.51; *P*=.08). In contrast, for nocturnal sleep trajectories, the inverted U-shaped sleep pattern remained robustly and independently associated with an increased risk of new-onset stroke after full adjustment (HR=2.91, 95% CI 1.28 - 6.61; *P*=.01).

In misclassification-adjusted analyses using posterior-probability-weighted Cox regression (Table S6A and S6B in Multimedia Appendix 1), the overall patterns of association were largely consistent with those observed in the primary analyses. In particular, the inverted U-shaped sleep trajectory remained associated with an increased risk of adverse CVD including stroke (Total sleep trajectories-CVD- HR=1.47, 95% CI 1.07 - 2.02; *P*=.02; Nocturnal sleep trajectories-stroke-HR=1.29, 95% CI: 1.06 - 2.71, *P*=.03; Table S6B in Multimedia Appendix 1) and hypertension (Total sleep trajectories: HR=1.35, 95% CI 1.07 - 1.69; *P*=.01; Table S6A in Multimedia Appendix 1), although the strength and statistical significance of the associations varied across specific outcomes.

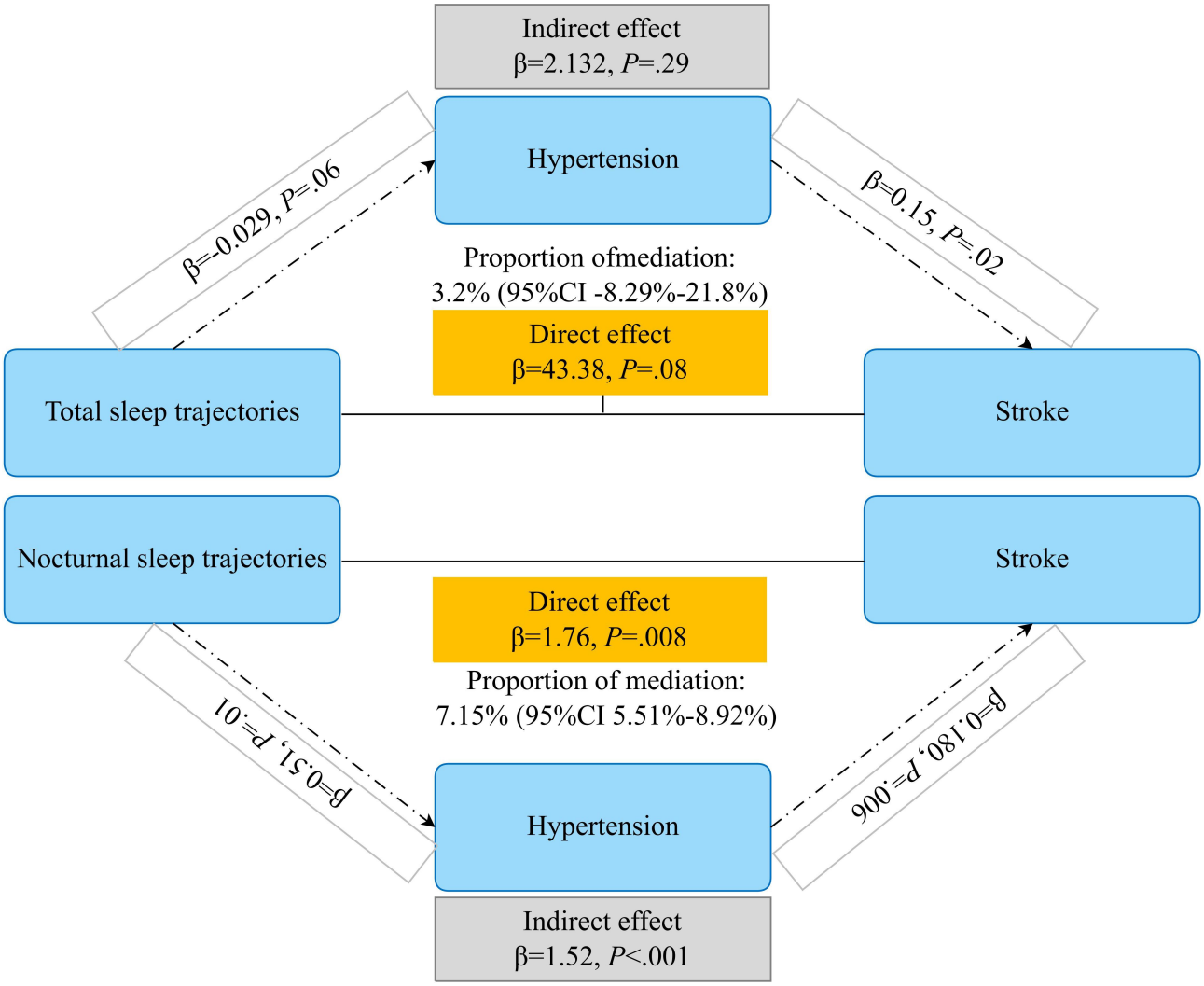
In sensitivity analyses with additional adjustment for SES <4 (Table S7A and S7B in Multimedia Appendix 1), the overall patterns of association between sleep trajectories and cardiometabolic outcomes were largely consistent with the primary analyses. Specifically, the inverted U-shaped pattern of total sleep maintained a significant association with a higher risk of new CVD (HR=1.58, 95% CI: 1.06 - 2.34; *P*=.02), but the link with hypertension was weakened after SES adjustment (HR=1.45, 95% CI 1.09 - 1.93; *P*=.01; Table S7A in Multimedia Appendix 1). In the same socioeconomic stratum, among nocturnal sleep trajectories, a significant association was observed between the inverted U-shaped nocturnal sleep pattern and incident stroke among participants with lower SES of <4 (HR=3.07, 95% CI 1.24 - 7.58; *P*=.01; Table S7B in Multimedia Appendix 1).

Mediator Analysis

Mediation analyses were conducted to assess whether hypertension mediated the associations between sleep duration trajectories and incident CVD, stroke, and heart disease. For total sleep duration trajectories, the indirect effect mediated through hypertension was not statistically significant (β=2.13; *P*=.30), accounting for 3.2% of the total effect (95% CI −8.29% to 21.8%). The direct effect of total sleep trajectories on stroke risk was also not statistically significant (β=43.38; *P*=.08; Figure 4). In contrast, for nocturnal sleep duration trajectories, a significant indirect effect through hypertension was observed (β=1.52; *P*<.001), with hypertension mediating 7.15% of the total association between nocturnal sleep trajectories and stroke risk (95% CI 5.51% - 8.92%; Figure 4). Notably, the direct effect of nocturnal sleep trajectories on stroke remained statistically significant after accounting for hypertension (β=1.76; *P*=.008; Figure 4). Hypertension partially mediated the association between sleep duration trajectories and stroke. For total sleep duration trajectories, neither the indirect effect through hypertension nor the direct effect on CVD reached statistical significance (indirect effect: β=0.050; *P*=.37; direct effect: β=0.029; *P*=.09), despite a modest estimated mediation proportion (8.7%; Figure S2 in Multimedia Appendix 2). Similar null findings were observed for nocturnal sleep duration trajectories, with nonsignificant indirect and direct effects on CVD. Consistent results were observed for heart disease, where neither total nor nocturnal sleep duration trajectories showed significant indirect effects mediated by hypertension, and all corresponding direct effects were also nonsignificant (Figure 3 in Multimedia Appendix 2).

XSL•FO
RenderX

Figure 4. Mediation analysis of the association between the total sleep/nocturnal sleep duration trajectories and stroke.



Discussion

Principal Findings

In this population-based study that is both longitudinal and nationally representative, researchers identified 4 distinct trajectories of total and nocturnal sleep duration. It was found that an inverted U-shaped sleep pattern was related to a greater risk of CVD, irrespective of typical risk factors. Multiple sensitivity analyses showed that these associations remained largely stable, supporting the strength of the results. We also explored the potential role of hypertension as an intermediate factor and observed preliminary evidence suggesting that hypertension may partly account for the relationship between unstable sleep trajectories and incident CVD.

This analysis highlights a clear inverted U-shaped association between sleep amount and the chance of developing hypertension. The risk of hypertension is considerably higher in individuals with an inverted U-shaped sleep duration trajectory than in those who consistently sleep for extended periods. In Model 3, the inverted U-shaped total sleep trajectory was associated with an increased risk of hypertension (HR=1.34, 95% CI 1.07 - 1.69; *P*=.01) [28]. In comparison, groups with a consistently low amount of sleep and those with a U-shaped sleep pattern did not have a significant association with the risk of hypertension. The important link identified in the inverted U-shape group is in line with past research that has related irregular or disrupted sleep patterns to an increased likelihood of hypertension. The nonlinear relationship between total sleep time and hypertension risk indicates that variability in sleep duration, rather than consistently short or long sleep, might be an important factor in the onset of hypertension. This association may be partly explained by the physiological stress induced by unstable sleep patterns, which can contribute to dysregulation of the autonomic nervous system, heightened inflammatory responses, and disruptions of circadian rhythms, all of which have been implicated in the development of hypertension [29].

With regard to the development of new-onset CVD, our findings indicate individuals with inverted U-shaped sleep duration trajectories exhibited a significantly elevated CVD risk. In the fully adjusted Model 3, the inverted U-shaped trajectory group showed an HR of 1.46 (95 %CI 1.07 - 2.00; *P*=.02) for total sleep duration, indicating a trend toward increased risk. Conversely, the U-shaped and steady low trajectory showed no significant association, likely due to limited statistical power from a smaller sample size. These results are consistent with meta-analytic data, which consistently show an irregular or insufficient relationship between sleep duration and CVD risk,

XSL•FO
RenderX

with optimal cardiovascular outcomes occurring between 7 and 8 hours of sleep per night [30]. The underlying physiological mechanisms for these associations likely involve multiple factors. Short sleep may increase stress hormone levels and inflammation, leading to endothelial dysfunction and atherosclerosis [29], while long sleep durations may be indicative of comorbid conditions such as depression or metabolic syndrome, both of which are known to be significant risk factors for CVD [30]. The study further suggests that individuals with erratic sleep patterns over the long term have a substantially higher risk of CVD, even when their sleep duration falls within the recommended range, emphasizing the role of sleep variability in cardiovascular risk.

In terms of stroke risk, there was a significant connection between the duration of sleep at night and the onset of new strokes in those with an inverted U-shaped sleep pattern, with a HR of 2.10 (95% CI 1.08 - 4.05; *P*=.03). This observation is consistent with findings from large cohort studies that have identified short sleep as a key risk factor for stroke, likely due to its impact on raising blood pressure and altering circadian rhythms [28,30]. The steady low sleep trajectory was associated with a significantly increased risk of stroke compared with the steady high sleep trajectory, both for total sleep duration (HR=1.71, 95% CI 1.07 - 2.72; *P*=.02) and nocturnal sleep duration (HR=1.83, 95% CI 1.15 - 2.91; *P*=.01). The evidence points to the possibility that chronic sleep disruption, either from insufficient sleep or prolonged sleeping, may heighten the likelihood of stroke. However, after excluding strokes occurring within the first 2 years of follow-up, the steady low sleep trajectory remained significantly associated with stroke risk, supporting the robustness of the observed association in longer-term follow-up.

No significant relationships were found between new-onset heart disease and any sleep duration trajectories, whether for overall sleep or night sleep duration. Unlike the associations seen with hypertension, CVD, and stroke, this suggests that the link between sleep duration and heart disease could differ from other cardiovascular conditions. Underlying health issues or genetic factors could potentially have a greater influence on the development of heart disease. In addition, the insignificant correlation may result from the relatively low prevalence of heart disease in our study population, highlighting the importance of conducting future research with more extensive sample sizes and prolonged follow-up durations.

Furthermore, our exploratory mediation analysis indicated that hypertension was responsible for only a slight portion of the association between nocturnal sleep patterns and stroke, comprising a modest fraction of the overall effect. The pattern suggests a potential relationship between sleep habits and stroke risk, but the temporal structure of the data does not allow for a definitive causal interpretation. No notable mediation was detected for CVD or heart disease, which suggests that the association between sleep and cardiovascular health might operate through mechanisms other than hypertension. Overall, these findings are consistent with the possibility that sleep health influences cardiovascular risk partly through its relationship with blood pressure regulation, although confirmation will require longitudinal data with more precise temporal ordering [31].

These findings are consistent with previous research linking irregular or insufficient sleep to negative cardiovascular outcomes, although the lack of significant associations for heart disease in Table 5 suggests that the effects may be specific to certain outcomes. The inverted U-shaped sleep duration trajectory consistently showed the strongest associations across outcomes, particularly for stroke, with a HR of 2.10 for nocturnal sleep. These findings are consistent with existing evidence suggesting that chronic sleep deprivation may contribute to endothelial dysfunction, oxidative stress, and inflammation, which are key mechanisms underlying cardiovascular and cerebrovascular pathogenesis. The borderline associations between the inverted U-shaped sleep trajectory and CVD and stroke may indicate potential involvement of circadian disruption or heightened sympathetic activity, whereas the nonsignificant findings for heart disease could be attributable to reduced statistical power or heterogeneity in disease pathophysiology. The absence of significant associations for the U-shaped sleep trajectory across outcomes may suggest that compensatory increases in sleep duration following periods of reduction could partially offset adverse effects. However, this hypothesis warrants further investigation in larger cohorts. The observed significant associations, especially for steady low sleepers, persisted after extensive adjustment for demographic, lifestyle, and clinical confounders, supporting the robustness of the results.

Identifying harmful trends in sleep duration has important implications for public health policies and medical practices. From the viewpoint of population health, those who regularly experience short, long, or erratic sleep durations form a high-risk group. Observing their sleep patterns over time could aid in early preventive interventions. Longitudinal data reveal that unhealthy sleep trends are associated with a heightened risk of cardiovascular conditions and death [32]. For example, a recent cohort study revealed that inconsistent sleep duration over many years significantly elevated the risk of cardiovascular problems [33]. Thus, ongoing monitoring of sleep duration and its time-related trends might contribute to risk assessment and efforts to reduce cardiovascular burden on a population scale. In a clinical context, recognizing distinct sleep-duration trajectories supports the inclusion of sleep pattern assessments in regular cardiovascular risk evaluations. Incorporating sleep-trajectory data alongside conventional risk factors (hypertension, dyslipidemia, and obesity) enables targeted interventions, including behavioral counseling on sleep hygiene, cognitive behavioral therapy for insomnia, and wearable-device monitoring, with the aim of stabilizing trajectories and mitigating subsequent cardiometabolic risk. From a mechanistic standpoint, ongoing deviations from normal sleep patterns could amplify pathophysiological processes such as sympathetic overactivity, chronic inflammation, and endothelial dysfunction, which are involved in the progression of CVD [34-36]. Further studies should integrate objective sleep-monitoring tools, such as actigraphy or polysomnography, with precision medicine approaches to enhance risk prediction and investigate if

XSL•FO
RenderX

modifying detrimental sleep patterns affects cardiovascular health.

The strengths of the current study lie in the large, nationally representative sample used for the analysis, which enhances the generalizability of the findings. Additionally, the study results remained robust across several sensitivity analyses, which further supports the reliability of the findings. The use of LCMM allowed for the clustering of individuals with similar underlying trajectories, minimizing potential misclassification and providing a more accurate representation of sleep duration patterns over time. However, several limitations must be acknowledged. First, as an observational study, the inability to establish causality is a fundamental limitation. Despite controlling for numerous potential confounders, residual confounding due to unmeasured variables cannot be ruled out, and the nonrandomized nature of the data further complicates causal inference. Although we excluded participants with prevalent CVD prior to the follow-up period and conducted a value analysis to address this concern, the observational design still precludes definitive causal conclusions. Second, the research was based on self-reported data for baseline measurements, particularly in terms of sleep duration and health outcomes, which might introduce recall bias or misclassification. The relevance of this bias is heightened among older adults, whose cognitive or memory challenges may compromise the accuracy of their self-reported information. Misreporting could result in nondifferential measurement errors, weakening the associations between sleep duration trajectories and cardiovascular outcomes, and possibly impacting the precision of trajectory classification. Nonetheless, self-reported sleep duration has been widely confirmed and applied in large-scale epidemiological studies like CHARLS, and it still serves as a practical and reliable measure for research on populations. To substantiate these findings and delve deeper into the connection between sleep patterns and cardiovascular health, future research should incorporate objective measures such as actigraphy or polysomnography. Third, although residual confounding cannot be fully ruled out, CHARLS lacks detailed dietary data, and the physical activity information has significant missing elements that could introduce bias if incorporated into the models. When interpreting the current findings, these data constraints should be taken into account. Fourth, due to the study population consisting of older adults who survived to the baseline, survivor bias might have occurred, leading to an underrepresentation of those with severe comorbidities or early death, which may have lessened the impact of the findings. A thorough examination of bias suggests that selection bias, information bias, and residual confounding together may have slightly influenced the effect estimates, underscoring the importance of careful interpretation. Fifth, the absence of cause-specific mortality records in the CHARLS dataset makes it impossible to identify whether deaths were due to cardiovascular or noncardiovascular causes. Consequently, the application of competing-risk models was not possible, possibly affecting the precision of cardiovascular event risk estimation. Finally, the sample comprised primarily middle-aged and older Chinese adults, limiting the external validity of the findings. The results may not be fully generalizable to younger populations or individuals from other countries or ethnic groups. Future research should seek to replicate these findings in more diverse populations and consider additional factors, such as sleep quality, sleep disorders, and other contextual variables, which may further refine our understanding of the relationship between sleep duration and cardiovascular health.

Conclusions

Our findings indicate that the longitudinal stability of sleep duration is an important marker of cardiovascular health. Sleep trajectories with continuous insufficiency or marked fluctuations were related to higher chances of negative cardiovascular events. In exploratory mediation analyses, hypertension accounted for a small proportion of the association between nocturnal sleep trajectories and stroke, suggesting a possible pathway through which sleep disruption may influence cerebrovascular risk. These results highlight the potential relevance of maintaining consistent sleep patterns as part of strategies to support cardiovascular health, while underscoring the need for future longitudinal research with clearer temporal ordering to clarify underlying mechanisms.

Acknowledgments

The authors would like to thank the editors and the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper. We confirm that no generative artificial intelligence (AI) tools were used at any stage in the preparation of this manuscript. The authors take full responsibility for the accuracy, originality, and integrity of all content, including all references and citations.

Funding

Support for this research was given by the Shanxi Health Committee Foundation Project (2024057) and Fundamental Research Program of Shanxi Province (202403021222432).

Data Availability

Original data supporting the conclusions of this paper will be available in the article or Supplementary Material from the authors. For further inquiries, reach out to the corresponding authors.

Authors' Contributions

Conceptualization: HL, MW

XSL•FO
RenderX

Methodology: HL, MZ
Data Curation: HL, HF
Investigation: HL, JP, HW
Writing—Original Draft Preparation: HL, HF, HW
Writing—Review & Editing: HL, MW, MZ, JP, HW, HF
Supervision: HL, MW
Project Administration: MZ
Funding Acquisition: HF
The manuscript reflects equal contributions from all authors, who have also read and agreed to its published form. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables on latent class mixed model fit indices, sensitivity analyses, misclassification-adjusted Cox regression analyses, and socioeconomic status–adjusted associations between sleep trajectories and incident hypertension, cardiovascular disease (CVD), stroke, and heart disease.
[DOCX File, 55 KB - aging_v9i1e78914_app1.docx]

Multimedia Appendix 2
Supplementary figures showing new-onset rates of hypertension, cardiovascular disease (CVD), stroke, and heart disease according to sleep duration trajectories, and mediation analyses of the associations between sleep trajectories and CVD/heart disease.
[DOCX File, 2042 KB - aging_v9i1e78914_app2.docx]

References

1. World Health Organization. UN Decade of Healthy Ageing: Plan of Action 2021–2030: World Health Organization; 2020:1-31 URL: https://www.who.int/publications/m/item/decade-of-healthy-ageing-plan-of-action [accessed 2026-05-15] [doi: 10.1177/10547738211065790]
2. World Health Organization. China country assessment report on ageing and health. World Health Organization. 2015 URL: https://iris.who.int/handle/10665/194271 [accessed 2026-05-15]
3. North BJ, Sinclair DA. The intersection between aging and cardiovascular disease. Circ Res 2012 Apr 13;110(8):1097-1108. [doi: 10.1161/CIRCRESAHA.111.246876] [Medline: 22499900]
4. Buford TW. Hypertension and aging. Ageing Res Rev 2016 Mar;26:96-111. [doi: 10.1016/j.arr.2016.01.007] [Medline: 26835847]
5. Yousufuddin M, Young N. Aging and ischemic stroke. Aging (Albany NY) 2019 May 1;11(9):2542-2544. [doi: 10.18632/aging.101931] [Medline: 31043575]
6. Lakatta EG, Levy D. Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: part II: the aging heart in health: links to heart disease. Circulation 2003 Jan 21;107(2):346-354. [doi: 10.1161/01.cir.0000048893.62841.f7] [Medline: 12538439]
7. Ford ES. Trends in mortality from all causes and cardiovascular disease among hypertensive and nonhypertensive adults in the United States. Circulation 2011 Apr 26;123(16):1737-1744. [doi: 10.1161/CIRCULATIONAHA.110.005645] [Medline: 21518989]
8. Chattu VK, Manzar MD, Kumary S, Burman D, Spence DW, Pandi-Perumal SR. The global problem of insufficient sleep and its serious public health implications. In: Kumary S, editor. Healthcare (Basel) 2018 Dec 20;7(1):1. [doi: 10.3390/healthcare7010001] [Medline: 30577441]
9. Dong L, Xie Y, Zou X. Association between sleep duration and depression in US adults: a cross-sectional study. J Affect Disord 2022 Jan 1;296:183-188. [doi: 10.1016/j.jad.2021.09.075] [Medline: 34607059]
10. Cai R, Chao J, Gao C, Gao L, Hu K, Li P. Association between sleep duration and cognitive frailty in older Chinese adults: prospective cohort study. JMIR Aging 2025 Apr 23;8:e65183. [doi: 10.2196/65183] [Medline: 40267503]
11. Erren TC, Morfeld P, Foster RG, Reiter RJ, Groß JV, Westermann IK. Sleep and cancer: synthesis of experimental data and meta-analyses of cancer incidence among some 1,500,000 study individuals in 13 countries. Chronobiol Int 2016;33(4):325-350. [doi: 10.3109/07420528.2016.1149486] [Medline: 27003385]
12. Ungvari Z, Fekete M, Varga P, et al. Imbalanced sleep increases mortality risk by 14-34%: a meta-analysis. Geroscience 2025 Jun;47(3):4545-4566. [doi: 10.1007/s11357-025-01592-y] [Medline: 40072785]
13. Wang D, Li W, Cui X, et al. Sleep duration and risk of coronary heart disease: a systematic review and meta-analysis of prospective cohort studies. Int J Cardiol 2016 Sep 15;219:231-239. [doi: 10.1016/j.ijcard.2016.06.027] [Medline: 27336192]

XSL•FO
RenderX

14. Li W, Wang D, Cao S, et al. Sleep duration and risk of stroke events and stroke mortality: a systematic review and meta-analysis of prospective cohort studies. Int J Cardiol 2016 Nov 15;223:870-876. [doi: 10.1016/j.ijcard.2016.08.302] [Medline: 27584562]
15. Meng L, Zheng Y, Hui R. The relationship of sleep duration and insomnia to risk of hypertension incidence: a meta-analysis of prospective cohort studies. Hypertens Res 2013 Nov;36(11):985-995. [doi: 10.1038/hr.2013.70] [Medline: 24005775]
16. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
17. Fang J, Wen Z, Ouyang J, Wang H. Modeling the change trajectory of sleep duration and its associated factors: based on an 11-year longitudinal survey in China. BMC Public Health 2021 Oct 30;21(1):1963. [doi: 10.1186/s12889-021-12017-8] [Medline: 34717596]
18. Yang M, Liu J, Shen Q, et al. Body roundness index trajectories and the incidence of cardiovascular disease: evidence from the China health and retirement longitudinal study. J Am Heart Assoc 2024 Oct;13(19):e034768. [doi: 10.1161/JAHA.124.034768] [Medline: 39319466]
19. Zhang Z, Zhao L, Lu Y, Xiao Y, Zhou X. Insulin resistance assessed by estimated glucose disposal rate and risk of incident cardiovascular diseases among individuals without diabetes: findings from a nationwide, population based, prospective cohort study. Cardiovasc Diabetol 2024 Jun 6;23(1):194. [doi: 10.1186/s12933-024-02256-5] [Medline: 38844981]
20. Liu H, Xu F, Zhang M, Chen X, Fan H, Hou M. Association between body roundness index and new-onset stroke risk in middle-aged and older adults with varying glucose metabolism status: a longitudinal study using CHARLS data. Diabetes Res Clin Pract 2025 Aug;226:112367. [doi: 10.1016/j.diabres.2025.112367] [Medline: 40645343]
21. Williams B, Mancia G, Spiering W, et al. 2018 ESC/ESH guidelines for the management of arterial hypertension. Eur Heart J 2018 Sep 1;39(33):3021-3104. [doi: 10.1093/eurheartj/ehy339] [Medline: 30165516]
22. Yu J, Yi Q, Chen G, et al. The visceral adiposity index and risk of type 2 diabetes mellitus in China: a national cohort analysis. Diabetes Metab Res Rev 2022 Mar;38(3):e3507. [doi: 10.1002/dmrr.3507] [Medline: 34679251]
23. Zhou Z, Ong KL, Whelton SP, et al. Impact of blood lipids on 10-year cardiovascular risk in individuals without dyslipidemia and with low risk factor burden. Mayo Clin Proc 2022 Oct;97(10):1883-1893. [doi: 10.1016/j.mayocp.2022.03.025] [Medline: 35760597]
24. Zheng X, Han L, Shen S. Hypertension, remnant cholesterol and cardiovascular disease: evidence from the China health and retirement longitudinal study. J Hypertens 2022 Nov 1;40(11):2292-2298. [doi: 10.1097/HJH.0000000000003259] [Medline: 35950988]
25. Pan XF, Wang L, Pan A. Epidemiology and determinants of obesity in China. Lancet Diabetes Endocrinol 2021 Jun;9(6):373-392. [doi: 10.1016/S2213-8587(21)00045-0] [Medline: 34022156]
26. Li C, Ma Y, Yang C, Hua R, Xie W, Zhang L. Association of cystatin C kidney function measures with long-term deficit-accumulation frailty trajectories and physical function decline. JAMA Netw Open 2022 Sep 1;5(9):e2234208. [doi: 10.1001/jamanetworkopen.2022.34208] [Medline: 36178684]
27. Li C, He D, Liu Y, Yang C, Zhang L. Anti-hypertensive medication adherence, socioeconomic status, and cognitive aging in the Chinese community-dwelling middle-aged and older adults ≥ 45 years: a population-based longitudinal study. BMC Med 2025 Feb 25;23(1):121. [doi: 10.1186/s12916-025-03949-8] [Medline: 40001139]
28. Han B, Chen WZ, Li YC, Chen J, Zeng ZQ. Sleep and hypertension. Sleep Breath 2020 Mar;24(1):351-356. [doi: 10.1007/s11325-019-01907-2] [Medline: 31402441]
29. Gangwisch JE, Malaspina D, Boden-Albala B, Heymsfield SB. Inadequate sleep as a risk factor for obesity: analyses of the NHANES I. Sleep 2005 Oct;28(10):1289-1296. [doi: 10.1093/sleep/28.10.1289] [Medline: 16295214]
30. Zhao Y, Li J, Li Y, et al. Sleep duration and risk of cardio-cerebrovascular disease: A meta-analysis. Front Cardiovasc Med 2022;9:907990. [doi: 10.3389/fcvm.2022.907990] [Medline: 36237900]
31. Li J, Li Y, Li Y, et al. Relationship of sleep duration with incident cardiovascular outcomes: a cohort study. BMC Public Health 2023;23(1):15042. [doi: 10.1186/s12889-023-15042-x] [Medline: 36653782]
32. Zhang J, Chen L, Zhang S, et al. Associations of sleep patterns with dynamic trajectory of cardiovascular multimorbidity and mortality: a multistate analysis of a large cohort. J Am Heart Assoc 2023 Oct 3;12(19):e029463. [doi: 10.1161/JAHA.123.029463] [Medline: 37776189]
33. Thurston RC, Chang Y, Kline CE, et al. Trajectories of sleep over midlife and incident cardiovascular disease events in the study of women's health across the nation. Circulation 2024 Feb 13;149(7):545-555. [doi: 10.1161/CIRCULATIONAHA.123.066491] [Medline: 38284249]
34. Greenlund IM, Carter JR. Sympathetic neural responses to sleep disorders and insufficiencies. Am J Physiol Heart Circ Physiol 2022 Mar 1;322(3):H337-H349. [doi: 10.1152/ajpheart.00590.2021] [Medline: 34995163]
35. Irwin MR. Sleep and inflammation: partners in sickness and in health. Nat Rev Immunol 2019 Nov;19(11):702-715. [doi: 10.1038/s41577-019-0190-z] [Medline: 31289370]
36. Calvin AD, Covassin N, Kremers WK, et al. Experimental sleep restriction causes endothelial dysfunction in healthy humans. J Am Heart Assoc 2014 Nov 25;3(6):e001143. [doi: 10.1161/JAHA.114.001143] [Medline: 25424573]

XSL•FO
RenderX

Abbreviations

BUN: blood urea nitrogen
CHARLS: China Health and Retirement Longitudinal Study
CVD: cardiovascular disease
DBP: diastolic blood pressure
DM: diabetes mellitus
eGFR: estimated glomerular filtration rate
FPG: fasting plasma glucose
HbA_{1c}: hemoglobin A_{1c}
HR: hazard ratio
LCMM: latent class mixed models
MICE: multiple imputation by chained equation
SBP: systolic blood pressure
SES: socioeconomic status
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by J Wang; submitted 12.Jun.2025; peer-reviewed by DG Mo, JW Kim, Y Tang, Z cai; revised version received 04.Jan.2026; accepted 02.Apr.2026; published 28.May.2026.

Please cite as:
Liu H, Wu M, Zhang M, Pan J, Wu H, Fan H
Subsequent Hypertension's Mediation of the Association Between Sleep Duration Trajectories and New-Onset Cardiovascular Disease: Population-Based Cohort Study
JMIR Aging 2026;9:e78914
URL: https://aging.jmir.org/2026/1/e78914
doi:10.2196/78914

© Hongwei Liu, Minghui Wu, Minheng Zhang, Jian Pan, Haiyan Wu, Haixia Fan. Originally published in JMIR Aging (https://aging.jmir.org), 28.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Deep Learning–Based Estimated Pulmonary Biological Age From Chest Computed Tomography Images in Healthy Adults: Model Development and Validation Study

Liping Zuo[1*], MD; Na Zhu[2*], MM; Bowen Wang[1], MD; Donglai Li[3], MD; Jinlei Fan[1], MD; Zhaolei Fan[4], MSc; Yongsheng Shang[4], MSc; Yongxiang Wang[4], MSc; Lei Xu[5], MM; Peng Zhou[6], MM; Wangshu Cai[6], MM; Dexin Yu[1], MD

[1]Department of Radiology, Qilu Hospital of Shandong University, 107 Wenhua Xilu, Lixia District, Jinan 250012, Shandong, China

[2]Department of Clinical Laboratory, Qilu Hospital of Shandong University (Qingdao), Qingdao 266035, Shandong, China

[3]Department of Emergency Medicine, Qilu Hospital of Shandong University, Shandong Provincial Clinical Research Center for Emergency and Critical Care Medicine, Jinan 250012, Shandong, China

[4]Msun Health Technology Group Co., Ltd, Jinan 250014, Shandong, China

[5]Medical Imaging Department, Shengli Oilfield Central Hospital, Dongying 257100, Shandong, China

[6]Department of Radiology, Central Hospital Affiliated to Shandong First Medical University, Jinan 250013, Shandong, China

[*]these authors contributed equally

Corresponding Author:
Dexin Yu, MD
Department of Radiology, Qilu Hospital of Shandong University, 107 Wenhua Xilu, Lixia District, Jinan 250012, Shandong, China

Abstract

Background: Estimated pulmonary biological age (ePBA) has emerged as a more reliable indicator for disease progression and mortality than chronological age, with chest computed tomography (CT) as a promising tool for calculating ePBA. However, the lack of models trained and validated with large-scale healthy adults hinders the generalizability of the CT-based ePBA.

Objective: This study aims to develop an aging biomarker (ePBA) from multicenter chest CTs of healthy adults using deep learning and investigate the association between age gap (ePBA - chronological age) and pulmonary function as well as all-cause mortality in patients with chronic obstructive pulmonary disease (COPD).

Methods: We used 11,187 chest CT scans from healthy adults at 3 health management centers and used multiple deep learning models. Of these, 7726 scans from institution A were used for model development. The remaining CT scans from institutions B (n=1506) and C (n=1955) served as external test datasets. To examine whether ePBA provided information beyond chronological age in patients with the disease, we investigated the association of age gap with lung function and all-cause mortality among 138 patients with COPD hospitalized at the same time period in institution A.

Results: The deep learning models demonstrated acceptable applicability for this task and exhibited a strong correlation between ePBA and chronological age. Age gap was significantly associated with forced expiratory volume in 1 second expressed as percentage of predicted values reduction (r_s=−0.18; P=.03) and an increased risk of all-cause mortality (hazard ratio: 1.16, 95% CI 1.08-1.25) in patients with COPD.

Conclusions: This study developed and validated a biomarker of aging—ePBA—with deep learning models based on chest CT. Age gap could serve as a novel clinical biomarker in patients with COPD.

(JMIR Aging 2026;9:e78243) doi:10.2196/78243

KEYWORDS

estimated pulmonary biological age; chest computed tomography; chest CT; deep learning; age gap; chronic obstructive pulmonary disease

Introduction

With the acceleration in global population aging, people who have chronic respiratory diseases have increased in number, and chronic respiratory disease has become the third leading cause of death, causing substantial economic and health burdens to society. Longitudinal studies have emphasized the critical need to precisely assess pulmonary senescence [1,2], which is essential for extending health span and reducing age-related respiratory morbidity. Thus, there is an urgent necessity to

XSL•FO
RenderX

develop a biomarker of aging—estimated pulmonary biological age (ePBA)—which is applicable to large-scale populations.

Models have been developed to estimate ePBA based on genes, proteins, and cells, but their expensive and complex assays prevent clinical translation [3-6]. Imaging techniques, such as noninvasive and highly accessible diagnostic tools, are widely used across health care settings, from primary care to tertiary medical centers. Despite the common use of chest radiographs for initial respiratory assessments, their limited spatial resolution contrasts with the detailed structural and physiological insights provided by computed tomography (CT) scans. Chest CT offers a detailed assessment of changes in lung parenchyma, small airways, and pulmonary vasculature associated with declining lung function and aging [7,8]. Therefore, an ePBA assessment model based on chest CT is both highly feasible and generalizable, holding great promise for managing chronic respiratory diseases [6].

Deep learning can automatically extract features from massive training data, which is helpful for exploring the signs of aging in medical images. Although Azarfar et al [9] have explored a deep learning model for estimating biological age using chest CT, it was trained on the National Lung Screening Trial dataset, with participants having a smoking history of at least 30 pack-years and comorbidities and thus could not represent a healthy elderly population. In a previous study [10], ePBA was calculated using chest CT reconstructed bone from 2500 healthy participants, but due to the single-center nature and limited sample size in the validation dataset, the stability and generalization of the model needed to be improved. Thus, the lack of models trained and validated with large-scale healthy adults hinders the generalizability of the CT-based ePBA.

This study aimed to develop and validate ePBA using deep learning models based on chest CT scans from healthy individuals. The age gap—defined as the difference between ePBA and chronological age—was used to represent the extent of lung aging in an individual relative to a standard population. Furthermore, we examined the correlation between age gap and pulmonary function in patients with chronic obstructive pulmonary disease (COPD) and evaluated its ability to assess all-cause mortality risk in these patients, which may have potential applications for the management of patients with chronic respiratory diseases.

Methods

Participants in Model Development and Test

Participants underwent chest CT scans between January 2022 and November 2023 as periodic health check-ups in 3 health management centers (Qilu Hospital of Shandong University, Jinan [institution A]; Central Hospital Affiliated to Shandong First Medical University, Jinan [institution B]; and Shengli Oilfield Central Hospital, Dongying [institution C]), who were enrolled to develop and validate a deep learning model for calculating ePBA.

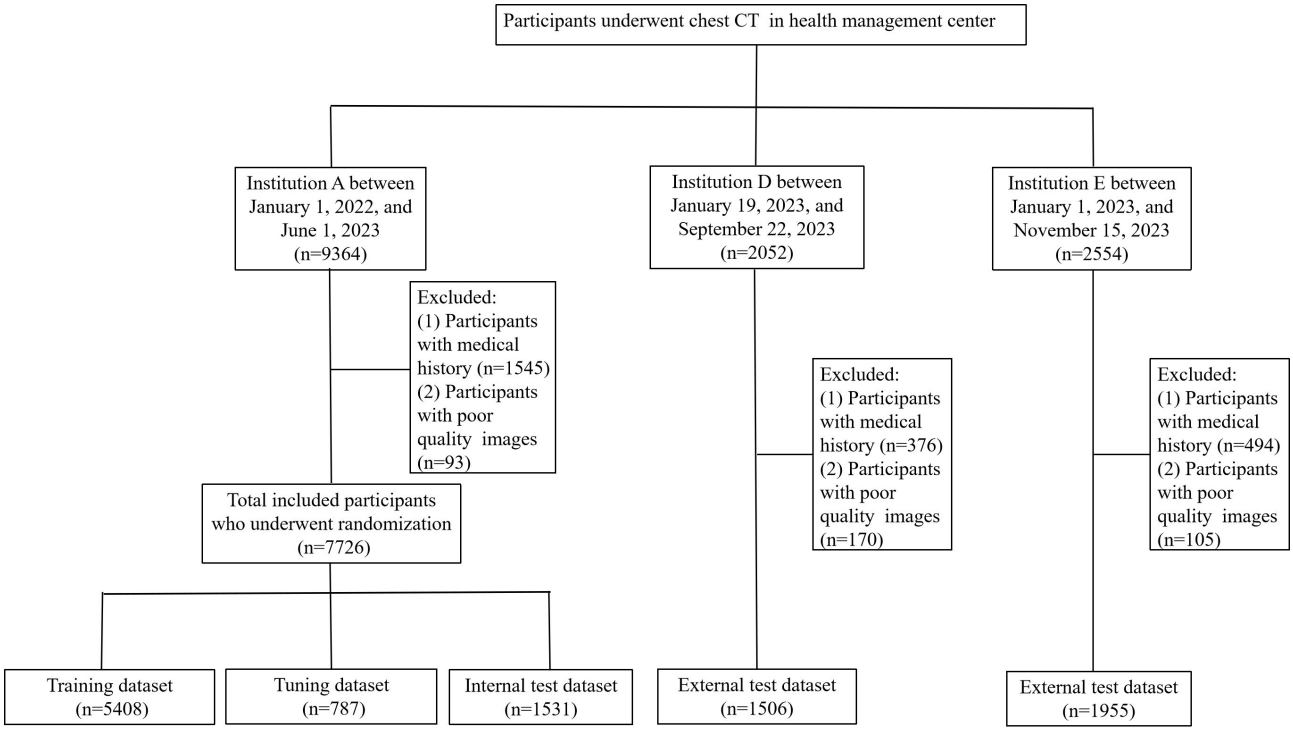
If participants underwent multiple chest CT scans during the data collection period, only the initial scan was selected. Disease registration was completed for all participants on the day of the CT examination at the 3 hospitals. The disease registration was established in accordance with the International Classification of Diseases and Related Health Problems, Tenth Revision [11,12]. Participants with disease registrations or compromised poor image quality were excluded (Figure 1).

Ethical Consideration

This retrospective and multicenter study was approved by the ethics committee of Qilu Hospital of Shandong University. No invasive operations were used or human tissue was obtained in this study, so the requirement for informed consent was waived (KYLL-202310‐027). The study was conducted in accordance with the principles of the Declaration of Helsinki. All chest CT images and clinical data were deidentified and anonymized prior to analysis to protect patient privacy. The confidentiality of participants' personal information was strictly maintained throughout the study. As this was a retrospective study, participants were not compensated.

XSL•FO
RenderX

Figure 1. The flowchart of the included and excluded individuals in model development and test. CT: computed tomography.



CT Acquisition

Noncontrast CT images of healthy participants were acquired following complete inspiration. Images of the entire lung, with a slice thickness of 1 - 5 mm, were reconstructed using standard kernels. The CT scanners used were Siemens Force, SOMATOM Definition AS (both from Siemens Healthineers), Philips Incisive CT, Philips Ingenuity CT, Philips IQon Spectral CT (all from Philips Healthcare), United Imaging uCT710 (United Imaging Healthcare), GE Revolution CT, and GE Discovery CT750 HD (both from GE Healthcare). The scanning parameters are shown in Table S1 in Multimedia Appendix 1.

Ground Truth Labeling and Data Partition

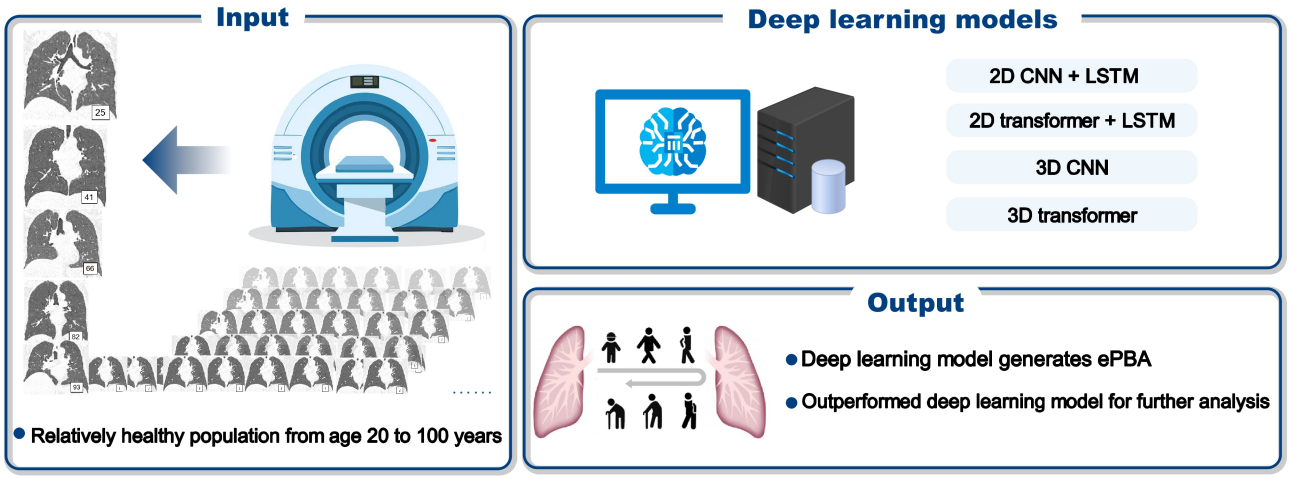
Chest CT images of healthy participants were labeled using chronological age. All labeled CTs from institution A were randomly assigned in a 7:1:2 ratio into training, tuning, and internal test datasets, and age distribution consistency was preserved. The tuning dataset was dedicated to hyperparameter optimization, whereas the internal test dataset evaluated preliminary model performance. Additionally, the labeled CTs from institutions B and C were used for external test datasets.

Model Development and Test

A total of 10 deep learning models (visual geometry group [11 layers] [VGG11] [13]+long short-term memory [LSTM]; residual network (18 layers) [14]+LSTM; ConvNeXt [15]+LSTM; vision transformer [16]+LSTM; shifted windows transformer [17]+LSTM; 3D VGG11 [18]; 3D residual network (18 layers) [19]; 3D ConvNeXt [15]; 3D vision transformer [20]; 3D shifted windows transformer [21]), including 4 major categories (2D convolutional neural network [CNN]/transformer+LSTM, 3D CNN or transformer), were trained to calculate ePBA from healthy individuals (Figure 2). Further details on model training and parameter settings are available in supplementary methods in Multimedia Appendix 1. The model most suitable for this task was selected for subsequent analysis.

XSL•FO
RenderX

Figure 2. Development of deep learning estimated pulmonary biological age (ePBA) models. Deep learning–based ePBA models were developed using chest computed tomography (CT) scans from relatively healthy individuals aged 20 to 100 years with chronological age labels. CNN: convolutional neural network; LSTM: long short-term memory.



The training and tuning datasets were used to train and tune this model. Data augmentation included random horizontal flips, vertical flips, rotations, and offsets to increase the diversity of the data. The loss function used mean square error, and the optimizer used Adam. The initial learning rate was set to 10^{-4}. We used a cosine annealing schedule over 500 epochs and used early stopping to prevent overfitting. Through training and tuning, we obtained an optimal model. The performance of the optimal model was evaluated on internal and external test datasets. To empirically validate the computational efficiency and performance of the 2D CNN+LSTM framework relative to 3D CNNs, we extended our ablation studies to a suite of 3D convolutional architectures. The input processing for these 3D models aligns with that of the 2D pipeline, ensuring fair comparisons across architectures.

The development process was performed using the PyTorch framework (version 1.8.1), and the hardware was an NVIDIA TITAN V GPU with 12 GB of memory.

To investigate the decision-making process of our best-performing model, we used score-weighted class activation mapping (score-CAM) [22] to visualize the discriminative regions it focused on during validation. This method represents a novel gradient-free visual interpretation technique aimed at explaining the classification decisions made by deep learning models. By highlighting important image regions in the visualizations, score-CAM reveals where the model allocates attention during the classification process. Notably, it serves as a bridge between perturbation-based and CAM-based approaches while providing intuitive and accessible insights into how the activation map weights are derived.

Application of ePBA in Patients with COPD

To evaluate the clinical applicability of the ePBA, we retrospectively enrolled patients hospitalized with COPD at institution A between January 1, 2022 and June 1, 2023. Data on demographics, pulmonary function, and medical history were collected from medical records. Dyspnea severity was assessed using the modified Medical Research Council (mMRC) scale. We also evaluated the ADO (age, dyspnea, and airflow obstruction) index [23], an abbreviated version of the BODE index (including body-mass index, airflow obstruction, dyspnea, and exercise capacity) and a simpler COPD prognostic risk index based on age, dyspnea, and airflow obstruction. Chest CT images from patients with COPD were input into the ePBA model, and the age gap (the difference between ePBA and chronological age) was subsequently calculated. Age gap above 0 indicated a higher level of aging than their healthy peers of the same chronological age, while age gap below 0 indicated a lower level of aging. The vital status was collected by reviewing medical records or through telephone calls to the patient or family. After at least 6 months of follow-up, a total of 138 patients with COPD were included in the study. The detailed inclusion and exclusion criteria are described in Figure S1 in Multimedia Appendix 1. An investigation was conducted to determine whether the age gap was related to the results of pulmonary function measurements and their risk of death.

Statistical Analysis

Categorical variables were presented as numbers and percentages. Continuous variables were summarized as mean (SD) or median and IQR, based on their distribution. Group comparisons were performed using the 2-tailed *t* test or Mann-Whitney *U* test for continuous variables and the chi-square or Fisher exact test for categorical variables, as appropriate.

The Pearson correlation coefficient (*r*), coefficient of determination (R^2), mean absolute error (MAE), mean square error, and root mean square error between the ePBA and chronological age were calculated to evaluate the regression performance of the deep learning model. The 99% CIs for these metrics were estimated using the nonparametric bootstrap method with 1000 replicates. This process involved generating 1000 independent bootstrap samples through random resampling with replacement from the original dataset, each of the same size as the original. The CIs were then derived using the percentile method, based on the 0.5th and 99.5th percentiles of the bootstrap sampling distribution.

In patients with COPD, the correlation between age gap and pulmonary function measurements was calculated. To examine the relationship between age gap and the mortality risk in

XSL•FO
RenderX

patients with COPD, Cox proportional hazards regression models were conducted with age gap used as a continuous variable and a binary variable, respectively (biologically older and younger represent age gap >0 and 0 [24], respectively). Adjustments were made for chronological age, sex, mMRC scale, smoking status, and medical history. Variance inflation factor (VIF) testing was performed to assess multicollinearity among the included variables. The proportional hazard assumption for the Cox model was checked using Schoenfeld residuals. For survival analysis, to assess the discriminatory ability of the age gap and ADO index, we performed Harrell concordance index and area under the curve, with 95% CIs yielded by 1000 bootstrap resamples. To visualize the discriminative performance of age gap, we also plotted the Kaplan-Meier curves of the biologically older and younger groups.

Statistical analyses were performed using the R software (version 3.3.3; R Foundation for Statistical Computing) and SPSS 26.0 for Windows (IBM Corp.). Statistical significance was set at $P<.05$.

Results

Demographic Characteristics of Healthy Individuals

A total of 13,970 participants with health check-ups (9364 in institution A, 2052 in institution B, and 2554 in institution C) were eligible for inclusion; 11,187 (7726 in institution A, 1506 in institution B, and 1955 in institution C) were ultimately enrolled after exclusion (Figure 1). Participants from institution A were randomly assigned to the training (n=5408, 2588 male participants), tuning (n=787, 385 male participants), and internal test datasets (n=1531, 710 male participants) in a ratio of 7:1:2. Additionally, participants from institutions B (n=1506, 814 male participants) and C (n=1955, 1310 male participants) were assigned to the external test datasets. The median ages (quartiles) of male and female participants were similar in the training, tuning, internal validation test, and the external test datasets (Table 1).

Table . Demographic characteristics of individuals.

Characteristics	Training dataset	Tuning dataset	Internal test dataset	External test dataset	
	Institution A	Institution A	Institution A	Institution B	Institution C
Total individuals	5408	787	1531	1506	1955
Male, n (%)	2588 (47.9)	385 (48.9)	710 (46.4)	814 (54.1)	1310 (67)
Female, n (%)	2820 (52.1)	402 (51.1)	821 (53.6)	692 (45.9)	645 (33)
Age (y), (min–max)					
Male	53 (20-98)	53 (22-100)	54 (23-94)	54 (20-97)	47 (21-85)
Female	53 (20-97)	53 (22-98)	54 (22-96)	54 (21-94)	47 (23-83)

Performance of Deep Learning Models for Age Estimation

All deep learning models achieved acceptable performance in this task (Table 2, Table S2-S11 in Multimedia Appendix 1). In total healthy individuals, the ePBAs derived from the VGG11+LSTM deep learning model were significantly associated with chronological ages across all datasets (training, tuning, internal test, and external test; all r >0.95, Table S2 in Multimedia Appendix 1). The external test dataset showed a strong correlation between ePBA and chronological age (r=0.97, 99% CI 0.96-0.97 in institution B, Figure 3A; r=0.98, 99% CI 0.96-0.99 in institution C, Figure 3B) and good model performance (institution B: R^2=0.93, 99% CI 0.92-0.94, MAE=4.59, 99% CI 4.34-4.83; institution C: R^2=0.93, 99% CI 0.84-0.94, MAE=3.64, 99% CI 3.47-3.81). Notably, the VGG11-LSTM model also exhibited a favorable balance between complexity and inference efficiency, containing 134.02 million parameters and achieving an average inference time of 18 ms per sample, rendering it suitable for potential clinical deployment. Thus, this model most suitable for this task was selected for subsequent analysis in patients with COPD.

XSL•FO
RenderX

Table . Regression performance parameters in the external test dataset, parameter counts, and inference time of deep learning models[a].

Model and backbone	*r*		R^2		MAE[b]		MSE[c]		RMSE[d]		Parameters (million)	Inference (ms)
	Institution B	Institution C	Institution B	Institution C	Institution B	Institution C	Institution B	Institution C	Institution B	Institution C		
2D CNN+LSTM[fg]												
VGG11[e]	0.97 (0.96-0.97)	0.98 (0.96-0.99)	0.93 (0.92-0.94)	0.93 (0.84-0.94)	4.59 (4.34-4.83)	3.64 (3.47-3.81)	34.46 (34.12-34.80)	21.64 (21.37-21.91)	5.87 (5.53-6.21)	4.65 (4.38-4.92)	134.02	18
Resnet18[h]	0.97 (0.96-0.97)	0.88 (0.87-0.89)	0.93 (0.92-0.94)	0.78 (0.75-0.80)	5.00 (4.76-5.24)	3.80 (3.63-3.97)	38.01 (37.66-38.36)	22.77 (22.52-23.03)	6.17 (5.82-6.51)	4.77 (4.52-5.03)	14.59	6
ConvNeXt	0.95 (0.95-0.96)	0.83 (0.81-0.85)	0.91 (0.90-0.92)	0.69 (0.66-0.72)	6.78 (6.46-7.10)	4.86 (4.65-5.08)	68.96 (68.57-69.36)	37.48 (37.15-37.80)	8.30 (7.91-8.70)	6.12 (5.80-6.45)	31.37	24
2D transformer+LSTM												
ViT[i]	0.95 (0.95-0.96)	0.85 (0.84-0.87)	0.90 (0.89-0.91)	0.66 (0.62-0.69)	5.02 (4.80-5.27)	4.32 (4.12-4.52)	39.53 (36.30-43.38)	29.59 (26.75-32.26)	6.29 (6.03-6.21)	5.44 (5.17-5.68)	25.02	19
Swin transformer	0.96 (0.96-0.97)	0.88 (0.86-0.89)	0.89 (0.88-0.91)	0.73 (0.69-0.76)	5.07 (4.83-5.36)	3.78 (3.61-3.95)	41.05 (37.02-45.95)	23.37 (21.50-25.56)	6.41 (6.08-6.78)	4.83 (4.64-5.06)	31.07	32
3D CNN												
VGG11	0.95 (0.95-0.96)	0.85 (0.83-0.86)	0.91 (0.89-0.92)	0.58 (0.51-0.63)	4.58 (4.32-4.85)	4.77 (4.56-4.99)	35.34 (31.65, 39.98)	36.37 (33.12-40.38)	5.95 (5.63-6.32)	6.03 (5.75-6.35)	178.66	20
Resnet18	0.95 (0.94-0.96)	0.84 (0.83-0.86)	0.90 (0.89-0.91)	0.61 (0.55-0.66)	4.90 (4.65-5.17)	4.56 (4.35-4.77)	38.54 (34.95-42.74)	33.81 (30.74-37.25)	6.21 (5.91-6.54)	5.82 (5.54-6.10)	33.21	64
ConvNeXt	0.89 (0.88-0.91)	0.71 (0.67-0.74)	0.80 (0.77-0.82)	0.10 (0.04-0.22)	6.80 (6.46-7.19)	6.78 (6.47-7.12)	78.01 (68.96-87.46)	77.56 (70.11-85.93)	8.83 (8.30-9.35)	8.81 (8.37-9.27)	29.78	197
3D transformer												
ViT	0.86 (0.85-0.88)	0.64 (0.60-0.67)	0.74 (0.71-0.77)	0.01 (0.14-0.12)	8.04 (7.62-8.44)	7.33 (7.02-7.67)	99.82 (90.17-110.19)	86.75 (79.62-94.39)	9.99 (9.50-10.50)	9.31 (8.92-9.72)	53.85	9
Swin transformer	0.92 (0.91-0.93)	0.78 (0.75-0.80)	0.85 (0.83-0.87)	0.34 (0.23-0.42)	5.87 (5.55, 6.20)	5.90 (5.63-6.17)	57.80 (51.21-64.28)	57.21 (52.05-63.42)	7.60 (7.16-8.02)	7.56 (7.21-7.96)	9.72	15

[a]All the data in parentheses are 99% CI.

[b]MAE: mean absolute error.

[c]MSE: mean square error.

[d]RMSE: root mean square error.

[e]VGG11: visual geometry group (11 layers).

[f]CNN: convolutional neural network.

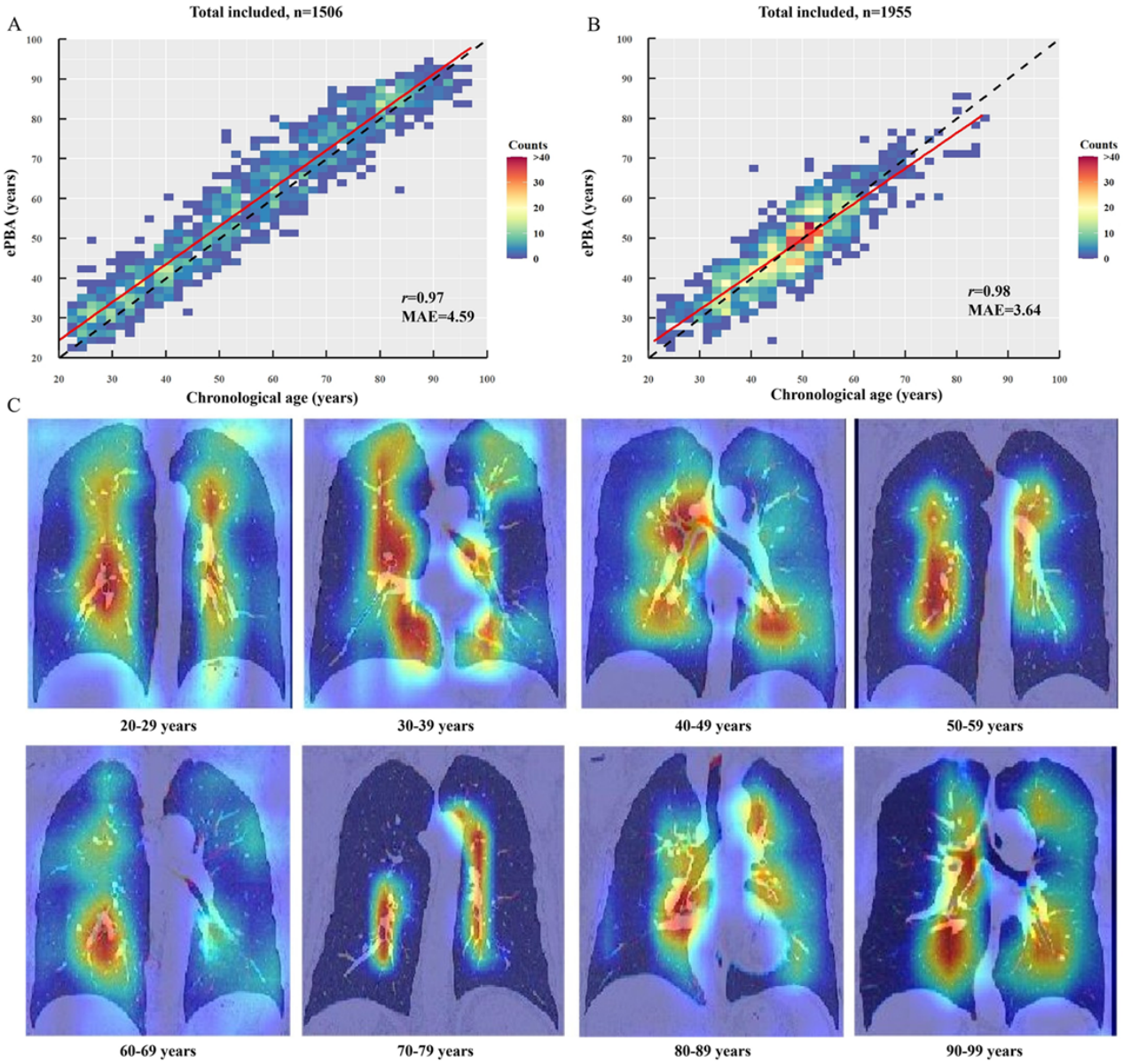
[g]LSTM: long short-term memory.

[h]ResNet18: residual network (18 layers).

[i]ViT: vision transformer.

XSL•FO
RenderX

Figure 3. Scatterplot between estimated pulmonary biological age (ePBA) derived from the VGG11-LSTM deep learning model and chronological age in (A) institution B and (B) institution C. (C) Score-weighted class activation mapping (score-CAM) original image class activation maps for healthy individuals. LSTM: long short-term memory; MAE: mean absolute error; VGG11: visual geometry group (11 layers).



To identify critical regions for optimizing feature localization during model training, we performed a visualization analysis using score-CAM. Figure 3C shows the score-CAM original image class activation maps of ePBA derived from VGG11-LSTM. The highlighted important image regions were located around the tracheal and bronchial vascular bundles after the visualization in healthy individuals at different age group distribution.

Application of ePBA in Patients With COPD

The ePBA of the 138 patients with COPD was significantly greater than their chronological age (ePBA: median 78.65, IQR 74.11-84.00 y vs chronological age: median 74.00, IQR 68.00-81.00 y; $P<.001$), resulting in a median age gap of median 5.43 (IQR 0.49-7.79) years, which was also significantly greater than 0 ($P<.001$). These findings indicate that these patients are in a state of accelerated aging. Table 3 provides the demographics, pulmonary function measurements, and medical history of survivors and deceased patients with COPD. As shown in Figure 4A, the age gap of deceased patients with COPD (median 6.96, IQR 3.00-11.00 y) was significantly greater than that of survivors (median 5.05, IQR –1.04 to 6.89 y; $P=.004$), and the mMRC scale score (median 4.00, IQR 3.25-4.00) was also significantly higher than that of survivors (median 3.00, IQR 2.00-4.00; $P<.001$; Table 3). There were no significant differences in other aspects.

XSL•FO
RenderX

Table . Demographics, pulmonary function, ADO (age, dyspnea, and airflow obstruction) index, and medical history in survivors and deceased patients with chronic obstructive pulmonary disease (COPD).

Characteristics	All (n=138)	Survivors (n=88)	Deceased (n=50)	Statistics	*P* value
Chronological age (y), median (Q1, Q3)	74.00 (68.00 to 81.00)	74.00 (69.00 to 81.00)	73.50 (66.00 to 82.00)	–0.55[a]	.58
ePBA[b] (y), median (Q1, Q3)	78.65 (74.11 to 84.00)	77.66 (73.09 to 83.11)	79.97 (76.23 to 85.00)	–1.79[a]	.07
Age gap (y), median (Q1, Q3)	5.43 (0.49 to 7.79)	5.05 (–1.04 to 6.89)	6.96 (3.00 to 11.00)	–2.85[a]	.004
Male participants, n (%)	115 (83.3)	74 (84.1)	41 (82.0)	0.10 (*df*=1)[c]	.75
BMI (kg/m^2), median (Q1, Q3)	22.47 (20.25 to 24.74)	22.13 (20.53 to 24.38)	22.96 (19.42 to 25.06)	–0.58[a]	.56
Smoking status, n (%)				3.80 (*df*=2)[c]	.15
Never smoker	37 (26.8)	25 (28.4)	12 (24.0)		
Current smoker	41 (29.7)	30 (34.1)	11 (22.0)		
Ex-smoker	60 (43.5)	33 (37.5)	27 (54.0)		
mMRC[d], median (Q1, Q3)	4.00 (3.00 to 4.00)	3.00 (2.00 to 4.00)	4.00 (3.25 to 4.00)	3.77[a]	<.001
FEV1%[e], median (Q1, Q3)	0.42 (0.33 to 0.55)	0.44 (0.34 to 0.56)	0.38 (0.31 to 0.54)	–1.52[a]	.13
FVC%[f], median (Q1, Q3)	0.67 (0.51 to 0.79)	0.70 (0.50 to 0.80)	0.65 (0.51 to 0.74)	–0.91[a]	.36
FEV1 or FVC[h, g], median (Q1, Q3)	0.50 (0.43 to 0.59)	0.51 (0.46 to 0.60)	0.49 (0.42 to 0.58)	–1.48[a]	.14
ADO index, median (Q1, Q3)	7.00 (5.00 to 8.00)	6.50 (5.00 to 8.00)	7.00 (6.00 to 8.00)	–2.11[a]	.03
Medical history, n (%)					
Hypertension	68 (49.3)	42 (47.7)	26 (52.0)	0.23 (*df*=1)[c]	.63
Diabetes	22 (15.9)	14 (15.9)	8 (16.0)	0.00 (*df*=1)[c]	.99
Cerebrovascular disease	25 (18.1)	18 (20.5)	7 (14.0)	0.90 (*df*=1)[c]	.34
Follow-up period (mo), median (Q1, Q3)	9.97 (4.83 to 16.93)	13.40 (8.68 to 18.42)	4.50 (0.76 to 11.26)	–5.78[a]	<.001

[a]z-score.

[b]ePBA: estimated pulmonary biological age.

[c]Chi-square.

[d]mMRC: modified Medical Research Council scale.

[e]FEV1%: forced expiratory volume in 1 second expressed as percentage of predicted values.

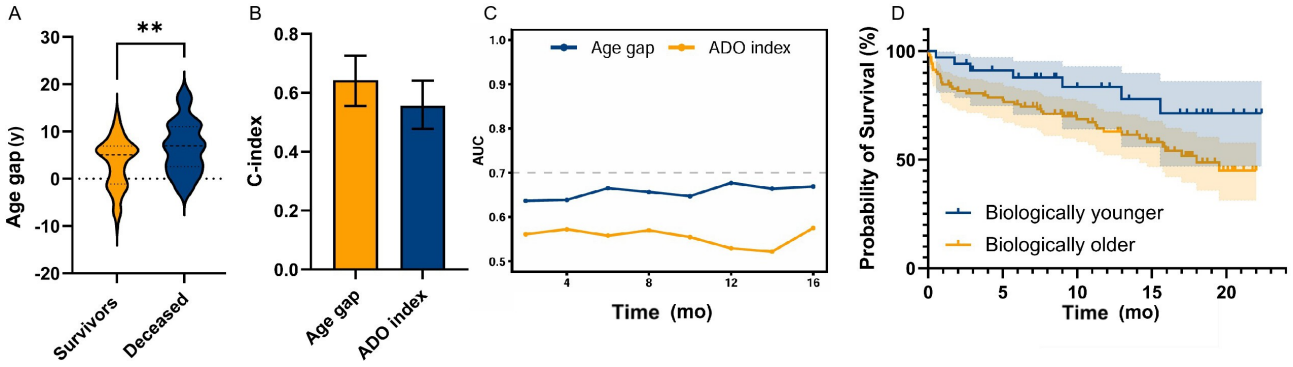
[f]FVC%: forced vital capacity expressed as percentage of predicted values.

[g]FEV1: forced expiratory volume in 1 second.

[h]FVC: forced vital capacity.

XSL•FO
RenderX

Figure 4. (A) The distribution of age gap between survivors and deceased patients with chronic obstructive pulmonary disease (COPD); (B) comparison of the performance of age gap and ADO (age, dyspnea, and airflow obstruction) index in predicting the survival outcome of patients with COPD; (C) time-dependent area under the curve (AUC) curves for age gap and ADO index in patients with COPD; and (D) Kaplan-Meier curve of overall survival for biologically younger and older groups in patients with COPD. **P<.01.



Spearman correlation analysis showed that age gap was negatively correlated with forced expiratory volume in 1 second expressed as percentage of predicted values (FEV1%; r_s=−0.18; P=.03), while no significant correlation was observed between age gap and forced vital capacity expressed as percentage of predicted values (r_s=−0.11; P=.21) or FEV1 or forced vital capacity (r_s=−0.14; P=.09).

Before the multivariate Cox regression analysis, VIF testing was performed to assess multicollinearity among the included variables (Table S12 in Multimedia Appendix 1). As VIF values above 10 suggest significant multicollinearity, all variables demonstrated VIF values below 2, indicating the absence of severe multicollinearity. These results confirm the reliability and validity of the subsequent multivariate regression models for evaluating the independent association between age gap and mortality in patients with COPD. In the Cox regression analysis model, age gap was significantly associated with an increased risk of mortality in patients with COPD (hazard ratio [HR]: 1.12, 95% CI 1.06-1.19). Further adjustments for chronological age, sex, mMRC scale, smoking status, and medical history showed similar HR (HR: 1.16, 95% CI 1.08-1.25; Table 4). Compared with biologically younger patients, biologically older patients had a significantly higher risk of death (HR: 2.24, 95% CI 1.01-4.98), and this remained significant even after adjusting for covariates (HR: 2.78, 95% CI 1.17-6.58; Table 4). The Schoenfeld residuals test results (Table S13 in Multimedia Appendix 1) indicated that both the global test and individual covariates in the Cox proportional hazards models of this study satisfied the proportional hazards assumption (P>.05). The global Schoenfeld residuals test P values for model 1, model 2, and model 3 were .39, .75, and .42, respectively. The concordance index of age gap was 0.643 (95% CI 0.555-0.726), which was higher than the ADO index 0.556 (95% CI 0.478-0.641) in differentiating the overall survival rate of patients with COPD (Figure 4B). Age gap maintained a consistently high area under the curve value throughout the entire prediction interval (Figure 4C). The Kaplan-Meier overall survival curve demonstrated an excellent distinction in the risk of death between the younger and older biological age groups, and the survival time of the older biological age group was significantly shorter than that of the younger biological age group (P=.047; Figure 4D).

Table . Association between age gap and mortality risk in patients with chronic obstructive pulmonary disease (COPD)[a].

	Model 1[b]	P value	Model 2[c]	P value	Model 13[d]	P value
Age gap (y)	1.12 (1.06-1.19)	<.001	1.16 (1.08-1.25)	<.001	1.16 (1.08-1.25)	<.001
Age gap category						
Biologically younger	Reference		Reference		Reference	
Biologically older	2.24 (1.01-4.98)	.048	2.72 (1.16-6.35)	.02	2.78 (1.17-6.58)	.02

[a]Biologically older and younger represent age gap >0 and ≤0, respectively. All data in parentheses are 95% CI.

[b]Unadjusted.

[c]Adjusting for chronological age, sex, BMI, and the modified Medical Research Council scale.

[d]Adjusting for covariates in model 2 plus smoking status, and medical history (hypertension, diabetes, and cerebrovascular disease).

Discussion

In this study, we developed and validated multiple deep learning models to calculate ePBA using chest CT scans from healthy individuals. Second, when applied to a COPD cohort, the derived “age gap” (ePBA—chronological age) emerged as a significant clinical biomarker, showing a negative correlation with FEV1% and proving to be a stronger predictor of all-cause mortality than the established ADO index.

In previous studies on calculating ePBA, gene-based, protein-based, and cellular marker–based ePBA calculation

XSL•FO
RenderX

models have been reported [24-26]. However, due to the high cost and complex workflow, these models were not suitable for large-scale cohort applications and were more suitable for exploring the biological mechanisms of lung aging. Liang et al [4] developed an estimation equation for lung age based on pulmonary function indicators (FEV1, FEF50%, FEF75%, and height) by the spline method (male: R^2=0.661, female: R^2=0.690). This equation might have a higher estimation error, as the coefficient of determination of this equation was lower than that of the CT-based deep learning model in this study. Furthermore, they failed to accomplish widespread clinical application due to the great difficulty in conducting pulmonary function tests in primary medical units. Among imaging tests, X-ray is often used as the first-line recommended test for chest disease. Mitsuyama et al [6] developed a biomarker of aging from chest radiography and examined the correlation between estimated age and chronological age (r=0.95). However, due to the low-density resolution of chest X-ray, chest CT is preferred for the full characterization of the aging lung in identifying signs of lung aging. The chest imaging features related to aging show vascular curvature and calcification, mediastinal lipomatosis, diaphragmatic bulge and protrusion, and musculoskeletal features, such as chest osteophyte and costal cartilage calcification [27]. Azarfar et al [9] developed an age estimation model using chest CT and deep learning and demonstrated its association with the risk of lung cancer. However, this model was trained using the National Lung Screening Trial dataset, which contains a large number of heavy smokers, and the participants have comorbidities, such as chronic lung diseases, making it difficult for this model to be generalized to patients with chronic respiratory diseases. In the study of Lu et al [10], ePBA was calculated using chest CT reconstructed bone from 2500 healthy participants, but due to the single-center nature and limited sample size in the validation dataset (n=200), the stability and generalization of the model needed to be improved. In this study, chest CT images of 11,187 healthy adults from 3 institutions were used to train the model to ensure that the model training dataset could represent a healthy population, and it was externally validated on 3461 healthy adults from 2 institutions that were included in the external test datasets, ensuring its stability and generalizability.

Importantly, we verified the application of ePBA in patients with COPD. It was observed that an increase in the ePBA-derived age gap was significantly associated with a reduction in FEV_1 % in patients with COPD. A similar negative association with lung function was also observed for phenotypic age acceleration, which was derived from phenotypic age based on chronological age and 9 clinical biomarkers as described by Wang et al [24]. After adjusting for chronological age and other covariates, the age gap remained independently associated with an elevated risk of all-cause mortality in patients with COPD. Notably, the predictive performance of age gap for mortality surpassed that of the established ADO index. Given that ePBA can be derived solely from routinely acquired chest CT images without requiring additional clinical parameters, it shows promise for improving the management of patients with chronic respiratory diseases and the implementation of targeted interventions aimed at addressing aging-related issues.

There are several limitations in this study. First, we did not compare this model with some established biological age markers (eg, epigenetic clock, DNAmGrimAge) due to the complexity and higher costs. Despite the continued refinement of epigenetic clocks over the last decade, standardized measures of organ age remain to be developed [28,29]. Second, participants were enrolled in China from a single ethnic background. The inclusion of participants with diverse ethnic, demographic, and socioeconomic backgrounds limits the generalizability of our current findings. Third, we did not construct a hybrid deep learning model based on clinical phenotype and chest CT. This was because clinical data, such as pulmonary function tests, are typically not collected during routine health check-ups, limiting their availability for model development. Finally, the COPD validation cohort has certain limitations, including its relatively small size, short follow-up, and predominance of severe cases. Although suitable for assessing mortality risk, this composition limits the direct evaluation of ePBA's ability to distinguish early pathological pulmonary aging from healthy aging. Therefore, future studies in community-based cohorts with milder disease and longer follow-up are warranted to verify the model's broader applicability.

In summary, this study constructed deep learning models to calculate the biomarker of aging—ePBA—based on chest CT images of multicenter healthy adults. The results showed that the age gap (ePBA—chronological age), representing the extent of lung aging in an individual relative to a standard population, may have potential applications for the management of patients with COPD.

Acknowledgments

We are grateful to Peng Xuewen, Zhao Weijia, Yanmin Yang, and Zhou Zhanyu for providing data from Qilu Hospital of Shandong University for this study. We are grateful to Gangming Ji for providing data from Shengli Oilfield Central Hospital for this study.

Data Availability

The study protocol and metadata are available upon request to the corresponding author. However, chest computed tomographies are not available, as these have been withheld by the hospitals that participated in this study in order to protect participant privacy. The code used to develop our model is available at [30].

XSL•FO
RenderX

Conflicts of Interest

ZF, YS, and YW declare relationships with the following company: Msun Health Technology Group Co Ltd. The other authors declare no relationships with any companies whose products or services may be related to the subject matter of the study.

Multimedia Appendix 1
Supplementary methods, results, and figures.
[DOCX File, 204 KB - aging_v9i1e78243_app1.docx]

References

1. Schuliga M, Pechkovsky DV, Read J, et al. Mitochondrial dysfunction contributes to the senescent phenotype of IPF lung fibroblasts. J Cell Mol Med 2018 Dec;22(12):5847-5861. [doi: 10.1111/jcmm.13855] [Medline: 30255990]
2. Schafer MJ, Zhang X, Kumar A, et al. The senescence-associated secretome as an indicator of age and medical risk. JCI Insight 2020 Jun 18;5(12):e133668. [doi: 10.1172/jci.insight.133668] [Medline: 32554926]
3. Aging Biomarker Consortium, Bao H, Cao J, et al. Biomarkers of aging. Sci China Life Sci 2023 May;66(5):893-1066. [doi: 10.1007/s11427-023-2305-0] [Medline: 37076725]
4. Liang X, Xie Y, Gao Y, et al. Estimation of lung age via a spline method and its application in chronic respiratory diseases. NPJ Prim Care Respir Med 2022 Sep 29;32(1):36. [doi: 10.1038/s41533-022-00293-9] [Medline: 36175436]
5. Tian YE, Cropley V, Maier AB, Lautenschlager NT, Breakspear M, Zalesky A. Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality. Nat Med 2023 May;29(5):1221-1231. [doi: 10.1038/s41591-023-02296-6] [Medline: 37024597]
6. Mitsuyama Y, Matsumoto T, Tatekawa H, et al. Chest radiography as a biomarker of ageing: artificial intelligence-based, multi-institutional model development and validation in Japan. Lancet Healthy Longev 2023 Sep;4(9):e478-e486. [doi: 10.1016/S2666-7568(23)00133-2] [Medline: 37597530]
7. Nam JG, Kang HR, Lee SM, et al. Deep learning prediction of survival in patients with chronic obstructive pulmonary disease using chest radiographs. Radiology 2022 Oct;305(1):199-208. [doi: 10.1148/radiol.212071] [Medline: 35670713]
8. Olesen ASO, Miger K, Ørting SN, et al. AI-based algorithm to detect heart and lung disease from acute chest computed tomography scans: protocol for an algorithm development and validation study. JMIR Res Protoc 2025 Sep 19;14:e77030. [doi: 10.2196/77030] [Medline: 40973115]
9. Azarfar G, Ko SB, Adams SJ, Babyn PS. Deep learning-based age estimation from chest CT scans. Int J Comput Assist Radiol Surg 2024 Jan;19(1):119-127. [doi: 10.1007/s11548-023-02989-w] [Medline: 37418109]
10. Lu T, Diao YR, Tang XE, et al. Deep learning enables automatic adult age estimation based on CT reconstruction images of the costal cartilage. Eur Radiol 2023 Nov;33(11):7519-7529. [doi: 10.1007/s00330-023-09761-3] [Medline: 37231070]
11. Weiner MG. POINT: is ICD-10 diagnosis coding important in the era of big data? Yes. Chest 2018 May;153(5):1093-1095. [doi: 10.1016/j.chest.2018.01.025] [Medline: 29410203]
12. ICD-10: there's a code for that. Lancet 2015 Oct 10;386(10002):1420. [doi: 10.1016/S0140-6736(15)00453-5] [Medline: 26466024]
13. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2015 Presented at: 3rd International Conference on Learning Representations, 2015; May 7-9, 2015 URL: https://www.scirp.org/reference/referencespapers?referenceid=3764205 [accessed 2026-02-17]
14. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 27-30, 2016 URL: https://ieeexplore.ieee.org/document/7780459 [accessed 2026-02-17]
15. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A ConvNet for the 2020s. Presented at: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 18-24, 2022 URL: https://ieeexplore.ieee.org/document/9879745 [accessed 2026-02-17]
16. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Houlsby N. An image is worth 16x16 words: transformers for image recognition at scale. arXiv. Preprint posted online on Oct 22, 2020. [doi: 10.48550/arXiv.2010.11929]
17. Liu Z, Lin Y, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows. Presented at: 2021 IEEE/CVF International Conference on Computer Vision (ICCV); Oct 10-17, 2021 URL: https://ieeexplore.ieee.org/document/9710580 [accessed 2026-02-17]
18. Tran D, Bourdev L, Fergus R, Torresani L, Paluri M. Learning spatiotemporal features with 3D convolutional networks. Presented at: 2015 IEEE International Conference on Computer Vision (ICCV); Dec 7-13, 2015 URL: https://ieeexplore.ieee.org/document/7410867 [accessed 2026-02-17]
19. Hara K, Kataoka H, Satoh Y. Can spatiotemporal 3D CNNs retrace the history of 2D CNNs and ImageNet? 2018 Presented at: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 18-23, 2018 URL: https://ieeexplore.ieee.org/document/8578783 [accessed 2026-02-17]

XSL•FO
RenderX

20. Perera S, Navard P, Yilmaz A. SegFormer3D: an efficient transformer for 3D medical image segmentation. 2024 Presented at: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); Jun 17-18, 2024 URL: https://www.computer.org/csdl/proceedings-article/cvprw/2024/654700e981/20A1jcB3z20 [accessed 2026-02-17]
21. Cai Y, Long Y, Han Z, et al. Swin Unet3D: a three-dimensional medical image segmentation network combining vision transformer and convolution. BMC Med Inform Decis Mak 2023 Feb 14;23(1):33. [doi: 10.1186/s12911-023-02129-z] [Medline: 36788560]
22. Li J, Zhang D, Meng B, Li Y, Luo L. FIMF score‐CAM: fast score‐CAM based on local multi‐feature integration for visual interpretation of CNNs. IET Image Process 2023 Feb;17(3):761-772. [doi: 10.1049/ipr2.12670]
23. Puhan MA, Garcia-Aymerich J, Frey M, et al. Expansion of the prognostic assessment of patients with chronic obstructive pulmonary disease: the updated BODE index and the ADO index. Lancet 2009 Aug 29;374(9691):704-711. [doi: 10.1016/S0140-6736(09)61301-5] [Medline: 19716962]
24. Wang T, Duan W, Jia X, et al. Associations of combined phenotypic ageing and genetic risk with incidence of chronic respiratory diseases in the UK Biobank: a prospective cohort study. Eur Respir J 2024 Feb;63(2):2301720. [doi: 10.1183/13993003.01720-2023] [Medline: 38061785]
25. Sanders JL, Putman RK, Dupuis J, et al. The association of aging biomarkers, interstitial lung abnormalities, and mortality. Am J Respir Crit Care Med 2021 May 1;203(9):1149-1157. [doi: 10.1164/rccm.202007-2993OC] [Medline: 33080140]
26. Hernandez Cordero AI, Yang CX, Milne S, et al. Epigenetic blood biomarkers of ageing and mortality in COPD. Eur Respir J 2021 Dec;58(6):2101890. [doi: 10.1183/13993003.01890-2021] [Medline: 34561282]
27. Ensor RE, Fleg JL, Kim YC, de Leon EF, Goldman SM. Longitudinal chest X-ray changes in normal men. J Gerontol 1983 May;38(3):307-314. [doi: 10.1093/geronj/38.3.307] [Medline: 6841924]
28. Elliott ML, Caspi A, Houts RM, et al. Disparities in the pace of biological aging among midlife adults of the same chronological age have implications for future frailty risk and policy. Nat Aging 2021 Mar;1(3):295-308. [doi: 10.1038/s43587-021-00044-4] [Medline: 33796868]
29. McGreevy KM, Radak Z, Torma F, et al. DNAmFitAge: biological age indicator incorporating physical fitness. Aging (Albany NY) 2023 May 31;15(10):3904-3938. [doi: 10.18632/aging.204538] [Medline: 36812475]
30. Lipingzuo/chestAge. GitHub. URL: https://github.com/lipingzuo/chestAge [accessed 2026-02-21]

Abbreviations

ADO: age, dyspnea, and airflow obstruction
CNN: convolutional neural network
COPD: chronic obstructive pulmonary disease
ePBA: estimated pulmonary biological age
FEV1: forced expiratory volume in 1 second
FEV1%: forced expiratory volume in 1 second expressed as percentage of predicted values
HR: hazard ratio
LSTM: long short-term memory
MAE: mean absolute error
mMRC: the modified Medical Research Council
Score-CAM: score-weighted class activation mapping
VGG11: visual geometry group (11 layers)
VIF: variance inflation factor

Edited by Y Jiang; submitted 29.May.2025; peer-reviewed by CC Xue, W Li; accepted 01.Feb.2026; published 12.Mar.2026.

Please cite as:
Zuo L, Zhu N, Wang B, Li D, Fan J, Fan Z, Shang Y, Wang Y, Xu L, Zhou P, Cai W, Yu D
Deep Learning–Based Estimated Pulmonary Biological Age From Chest Computed Tomography Images in Healthy Adults: Model Development and Validation Study
JMIR Aging 2026;9:e78243
URL: https://aging.jmir.org/2026/1/e78243
doi:10.2196/78243

© Liping Zuo, Na Zhu, Bowen Wang, Donglai Li, Jinlei Fan, Zhaolei Fan, Yongsheng Shang, Yongxiang Wang, Lei Xu, Peng Zhou, Wangshu Cai, Dexin Yu. Originally published in JMIR Aging (https://aging.jmir.org), 12.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR

XSL•FO
RenderX

Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Literacy and Heart Failure Self-Care in Older Patients and Their Caregivers: Dyadic Analysis Using the Actor-Partner Interdependence Model

Hokon Kim[1], PhD; Misook L Chung[2], PhD; Yeonsoo Jang[1,3], PhD; JiYeon Choi[1,3,4], PhD; Byung Su Yoo[5], MD, PhD; Sang Hui Chu[1,3,4], PhD

[1]College of Nursing, Yonsei University, Seoul, Republic of Korea

[2]School of Nursing, Vanderbilt University, Nashville, TN, United States

[3]Mo-Im Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

[4]Yonsei University Institute for Innovation in Digital Healthcare, Yonsei University, Seoul, Republic of Korea

[5]Division of Cardiology, Department of Internal Medicine, Wonju Severance Christian Hospital, Wonju, Republic of Korea

Corresponding Author:
Sang Hui Chu, PhD
College of Nursing
Yonsei University
50-1 Yonsei-ro
Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 02 2228 3257
Fax: 82 02 2228 3257
Email: shchu@yuhs.ac

Abstract

Background: The population of patients with heart failure (HF) is rapidly aging, and the prevalence of HF continues to rise among older adults. Effective HF self-care is essential for improving survival and reducing hospital readmissions, and the role of family caregivers in supporting and reinforcing these behaviors has become increasingly important. With the growing integration of digital health technologies into HF management, technology-assisted self-care is becoming more common. However, many older adults experience difficulties in adopting and effectively using digital tools, which may limit the potential benefits of digital health interventions. As both patients' and caregivers' levels of digital literacy may jointly shape HF self-care behaviors, a dyadic analytic approach is warranted to clarify their interdependent effects.

Objective: This study aimed to compare digital literacy between older patients with HF and their caregivers and to examine how digital literacy influences HF self-care behaviors within patient-caregiver dyads.

Methods: This cross-sectional study included 102 patients with HF–caregiver dyads recruited from outpatient clinics in South Korea. Digital literacy was measured using the Everyday Digital Literacy Questionnaire. HF self-care, encompassing 3 key dimensions—self-care maintenance, symptom perception, and self-care management—was assessed using the Self-Care of Heart Failure Index (version 7.2) for patients and the Caregiver Contribution to Self-Care of Heart Failure Index (version 2.0) for caregivers. Dyadic associations were analyzed using the Actor-Partner Interdependence Model with maximum likelihood estimation.

Results: Patients (age: mean 79.4, SD 9.1 years) had significantly lower digital literacy than caregivers (age: mean 59.0, SD 13.1 years). Caregivers primarily consisted of adult children (63/102, 61.8%), followed by spouses (33/102, 32.4%). The Actor-Partner Interdependence Model results revealed significant actor effects of digital literacy on symptom perception for both patients (β=0.26; *P*=.008) and caregivers (β=0.32; *P*=.002). For self-care management, a significant actor effect was observed only for patients (β=0.24; *P*=.02). No significant actor effects were found for self-care maintenance, and no partner effects reached statistical significance across any dimension.

Conclusions: Digital literacy significantly influenced individuals' own HF self-care behaviors, particularly symptom perception, but cross-partner effects were not observed within dyads. These findings suggest that digital health interventions should assess

XSL•FO
RenderX

and address patients' and caregivers' digital skills as distinct targets rather than assuming spillover effects within dyads. To optimize HF outcomes in aging populations, culturally sensitive and dyad-focused strategies that consider individual digital literacy are essential.

(JMIR Aging 2026;9:e85976) doi:10.2196/85976

KEYWORDS

digital literacy; heart failure; self-care; caregiver; symptom perception

Introduction

Heart failure (HF) is a prevalent and progressive cardiovascular syndrome among older adults, affecting approximately 18.6% of adults aged 80 years and older in Korea [1]. It is driven by age-related structural and functional changes in the cardiovascular system as well as cumulative cardiac diseases across the life course [2]. Older patients with HF often present with multiple comorbidities, frailty, and functional limitations, making disease management increasingly complex [3]. As populations age globally, the burden of HF continues to rise with high readmission and mortality rates among older adults, resulting in substantial health care use and economic burden [1,2,4,5].

Effective self-care is central to slowing HF progression, preventing irreversible cardiac damage, and improving long-term outcomes [5-7]. Substantial evidence consistently demonstrates that appropriate HF self-care reduces hospital readmissions and mortality and enhances the quality of life among patients with HF [8-10]. HF self-care is theoretically grounded in situation-specific theories of HF management and conceptualized as a multidimensional process encompassing self-care maintenance, symptom perception, and self-care management. These dimensions enable patients to maintain physiological stability, detect changes in symptoms, and respond appropriately to exacerbations [6,7,11]. Particularly, symptom perception—the ability to recognize, interpret, and evaluate bodily changes—has gained increasing attention because of its strong association with lower mortality and fewer readmissions [4,12,13].

Despite its importance, effective self-care is particularly challenging for older adults with HF [3]. Cognitive decline, physical limitations, multimorbidity, and reduced self-efficacy often hinder adherence to recommended behaviors and delay timely symptom perception and management [8,12,14,15]. In the context of aging, successful self-care increasingly depends on external support, as age-related functional decline limits patients' ability to independently manage complex treatment regimens [3,14,16]. This interdependence highlights the essential role of caregivers in facilitating HF self-care among older adults [3,17]. Caregivers assist with medication management, symptom monitoring, and decision-making, compensating for cognitive or functional limitations [3,16]. Consequently, HF self-care in later life should be conceptualized as a dyadic process embedded within patient-caregiver relationships rather than as isolated individual efforts and behaviors [17-19].

As HF management increasingly incorporates digital health technologies—including mobile apps, telemonitoring systems, and wearable devices—these tools have become integral components of contemporary care, enabling remote monitoring and promoting patient-centered self-care beyond traditional clinical settings [20,21]. However, the effectiveness of digital health technologies depends largely on users' acceptance and digital literacy [22,23], particularly among older adults who often experience difficulties with technology adoption and use [24,25].

Digital literacy, defined as the ability to access, understand, evaluate, and apply digital information, is a key determinant of successful engagement with digital health interventions [22]. Older adults generally exhibit lower levels of digital literacy than younger populations, which can limit effective technology use, undermine self-care adherence, and exacerbate digital health disparities in HF care [23-25]. Because caregivers frequently assist older patients in navigating digital environments [26,27], understanding digital literacy within patient-caregiver dyads is essential for developing inclusive and sustainable HF self-care strategies [3,17].

Given the frequent and ongoing interactions between patients and caregivers in older populations, analyses focused solely on individuals are insufficient to capture the complexity of HF self-care processes [3,17,18]. A dyadic analytic framework, such as the Actor-Partner Interdependence Model (APIM), allows the simultaneous evaluation of how individual characteristics influence both one's own outcomes (actor effects) and those of one's partner (partner effects), capturing bidirectional dynamics that are often overlooked in individual-level analyses [19,28].

Importantly, patient-caregiver relationships vary substantially across cultural contexts [18,29]. While previous research in Western settings has predominantly focused on spousal caregivers [30-32], East Asian societies, influenced by Confucian values, emphasize family-centered caregiving, with adult children frequently assuming the primary caregiving role [18,29]. Despite these cultural distinctions, empirical research examining dyadic processes and the role of digital literacy in HF self-care within Asian contexts remains remarkably scarce [19]. This gap is especially salient in East Asian contexts, where family-centered caregiving and age-related digital literacy challenges frequently coexist.

Therefore, this study aimed to (1) compare digital literacy levels between older patients with HF and their caregivers and (2) examine how digital literacy influences HF self-care behaviors within patient-caregiver dyads using a dyadic analytic framework. By identifying actor and partner effects, this study seeks to inform the design of culturally sensitive, dyad-focused digital health interventions that enhance collaborative HF

XSL•FO
RenderX

self-care among older adults with HF and their caregivers in Korea and similar East Asian contexts.

Methods

Participants and Study Design

This cross-sectional study recruited 102 dyads of patients with HF and their caregivers from the outpatient clinic of a tertiary hospital in South Korea between March and September 2023. Eligible patients were adults aged 19 years or older with a confirmed diagnosis of HF. Caregivers were adults aged 19 years or older who provided unpaid care to the participating patients. Patients were excluded if they had a history of heart transplantation or left ventricular assist device implantation or had cognitive impairment that could interfere with study participation. Caregivers who were paid or formally employed by the patient or the patient's family were also excluded.

Ethical Considerations

This study was approved by the institutional review board of Yonsei University Health System (approval 4-2022-1622). All participants provided written informed consent before participation. The survey was conducted anonymously, and no personally identifiable information was collected. As compensation for participation, both patients and their caregivers received a small gift equivalent to approximately US $7.

Measurements

Digital Literacy

Digital literacy was assessed using the Everyday Digital Literacy Questionnaire (EDLQ), a 22-item instrument developed and validated for older adults in South Korea [33]. The EDLQ measures 3 domains: information and communication (9 items), content creation and management (4 items), and safety and security (9 items). Items are rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree), with higher total scores indicating greater overall digital literacy. The EDLQ has demonstrated excellent internal consistency, with a Cronbach α of 0.98 at the time of development [33] and 0.96 in this study.

HF Self-Care and Caregiver Contribution to HF Self-Care

Patients' HF self-care was assessed using the Self-Care of Heart Failure Index (version 7.2), a 29-item instrument comprising 3 dimensions: self-care maintenance (11 items), symptom perception (10 items), and self-management (8 items). Items are rated on a 5-point Likert scale. Following the scoring guidelines provided by the instrument developers, raw scores of each dimension were transformed into standardized scores ranging from 0 to 100, with higher standardized scores indicating better self-care. The Self-Care of Heart Failure Index (version 7.2) was originally developed by Riegel et al [6] and has been culturally adapted and validated for use in Korea [34]. Reported internal consistency ranges from Cronbach α of 0.70 in the original version [6] to 0.71 to 0.96 in the Korean version [34]. In this study, Cronbach α values for patients' HF self-care ranged from 0.76 to 0.84 across dimensions. Additionally, the caregiver contribution to HF self-care was measured using the Caregiver Contribution to Self-Care of Heart Failure Index (version 2.0), a parallel instrument designed to assess caregivers' support for patients' HF self-care across the same 3 dimensions. Scores were standardized on a 0 to 100 scale. The Caregiver Contribution to Self-Care of Heart Failure Index (version 2.0) has demonstrated strong validity in previous studies (Cronbach α=0.90) [35]. In this study, Cronbach α values for caregiver contribution were 0.72 for maintenance, 0.85 for symptom perception, and 0.76 for self-care management.

Covariates

Dyadic mutuality was assessed using the 15-item Mutuality Scale from the Family Caregiving Inventory, which evaluates love, shared values, reciprocity, and enjoyable activities on a 5-point scale. Higher scores indicate greater perceived mutuality [36]. The scale has demonstrated strong internal consistency in previous studies (Cronbach α=0.91-0.96) [31,37] and in this study (Cronbach α=0.95). Although mutuality was assessed for both dyad members, patient and caregiver scores were highly correlated (r>0.70; P<.001). To maintain model parsimony and avoid multicollinearity within the APIMs [38], only caregiver-perceived mutuality was included as a covariate, reflecting the caregiver's central role in shaping HF self-care support and caregiving behaviors [37,39]. Caregiver relationship type (spouse vs adult children) was included to account for differences in caregiving roles and responsibilities [29,30]. Patient comorbidity burden was assessed using the age-adjusted Charlson Comorbidity Index (ACCI), which incorporates age as a weighted component in addition to comorbid conditions to better capture overall disease complexity in older adults [40]. Additionally, patients' perceived economic status and residential areas (urban vs rural) were included to capture potential socioeconomic and contextual disparities in digital literacy, health resource access, and HF self-care behaviors in older adults [22,24,25].

Data Analysis

Data analyses were conducted using SPSS Statistics (version 26.0; IBM Corp) and SPSS AMOS (version 30.0; IBM Corp). Descriptive statistics were used to summarize participant characteristics and study variables.

Preliminary Analyses and Assumptions

Before model estimation, distributional assumptions for all primary continuous variables were examined using skewness and kurtosis statistics, histogram inspection, and normal *Q-Q* plots. These diagnostic procedures were conducted to assess the suitability of maximum likelihood (ML) estimation for APIM-based structural equation models. Although several variables demonstrated mild departures from normality—common for bounded Likert-type measures in behavioral and health research—all values remained within acceptable ranges for ML estimation in structural equation modeling (|skewness|<2.0; |kurtosis|<7.0) [41-43]. No evidence of severe nonnormality or influential outliers was identified. Given these findings and the moderate sample size (102 dyads), no variable transformations were applied, and ML estimation was retained for all analyses.

XSL•FO
RenderX

Bivariate associations among key variables were examined using Pearson correlation coefficients for continuous variables. Spearman correlations were applied to assess the association involving the ordinal variables (ie, perceived economic status). A complete correlation matrix with exact sample size, correlation coefficients, and *P* values is presented in Multimedia Appendix 1.

APIM Model Specification

Dyadic relationships were analyzed using the APIM, treating dyad members as a priori distinguishable by role (patients with HF vs caregivers) [38]. Each dyad consisted of 1 patient and 1 identified primary caregiver, recruited and coded separately by role. Digital literacy was measured using the same instrument for both members, whereas HF self-care outcomes were assessed using parallel role-specific instruments (patient HF self-care vs caregiver contribution to HF self-care).

Separate APIMs were estimated for each HF self-care dimension (eg, maintenance, symptom perception, and management) using ML estimation in AMOS. In each model, four structural paths were freely estimated: (1) two actor effects (patient digital literacy predicting patient HF self-care and caregiver digital literacy predicting caregiver contribution to HF self-care) and (2) two partner effects (patient digital literacy predicting caregiver contribution to HF self-care and caregiver digital literacy predicting patient HF self-care).

Patient and caregiver digital literacy scores, specified as the primary predictors, were allowed to covary to account for interdependence within dyads. Residuals of patient HF self-care and caregiver contribution to HF self-care were also correlated to capture dyadic interdependence not explained by the predictors. Five covariates—caregiver-perceived mutuality, caregiver relationship type (spouse vs adult child), patient comorbidity burden (ACCI), patient's perceived economic status, and residential area—were specified as exogenous variables, and covariances among all exogenous variables were freely estimated. No equality constraints were imposed across patient and caregiver paths. APIMs were estimated using a structural equation modeling framework to account for dyadic interdependence. Consistent with previous APIM studies [44], global model fit indices were not reported because the model was saturated. Model evaluation focused on the magnitude, direction, and statistical significance of estimated actor and partner effects, covariance, and explained variance (R^2) for patient and caregiver outcomes. The analyses represent concurrent (cross-sectional) APIMs. No mediators were specified, and no indirect effects were estimated.

Results

Characteristics of Patients With HF and Their Caregivers

A total of 102 patient-caregiver dyads were included in the analysis. Patients with HF had a mean age of 79.4 (SD 9.1) years, and 55 (53.9%) patients were female. Caregivers were younger (age: mean 59.0, SD 13.1 years) and predominantly female (n=67, 65.7%). Most caregivers were adult children, including sons, daughters, daughters-in-law, and sons-in-law (n=63, 61.7%), followed by spouses (n=33, 32.4%). Nearly half (n=49, 48%) of the patients lived with their caregivers. Patients generally had lower educational attainment compared with caregivers, and half (n=51, 50%) of the patients resided in rural areas.

Mean mutuality score indicated moderate levels of perceived mutuality for both patients (mean 2.8, SD 0.7) and caregivers (mean 2.7, SD 0.9). Patient and caregiver mutuality were strongly correlated (r=0.70; P<.001; Multimedia Appendix 1), reflecting the interdependent nature of patient-caregiver dyads.

Regarding clinical characteristics, 55 (53.9%) patients had HF with preserved ejection fraction. The mean left ventricular ejection fraction was 52.25% (SD 12.2%). Most patients presented with mild HF symptoms, classified as New York Heart Association class I to II, accounting for 89 (87.3%) samples. Patients exhibited a substantial comorbidity burden, with a mean ACCI score of 6.4 (SD 2.2). Detailed demographic and clinical characteristics are summarized in Table 1.

XSL•FO
RenderX

Table 1. Characteristics of patients with heart failure and their caregivers[a].

Characteristics	Patients (n=102)	Caregivers (n=102)
Age (years), mean (SD)	79.4 (9.1)	59.0 (13.1)
Female, n (%)	55 (53.9)	67 (65.7)
Education level, n (%)		
Below elementary school	67 (65.7)	19 (18.6)
Middle-to-high school	30 (29.4)	48 (47.1)
Above college	5 (4.9)	35 (34.3)
Perceived economic status, n (%)		
Low	65 (63.7)	30 (29.4)
Moderate to high	37 (36.3)	72 (70.6)
Residential area (urban), n (%)	51 (50.0)	—[b]
Relationship type, n (%)		
Spouse	—	33 (32.4)
Adult child	—	63 (61.7)
Living arrangement (cohabitation), n (%)	49 (48.0)	—
New York Heart Association class, n (%)		
Class I-II (mild symptoms)	89 (87.3)	—
Class III (moderate symptoms)	13 (12.7)	—
Left ventricular ejection fraction (%), mean (SD)	52.3 (12.2)	—
≤40% (heart failure with reduced ejection fraction), n (%)	18 (17.6)	—
41%-49% (heart failure with mildly reduced ejection fraction), n (%)	29 (28.4)	—
≥50% (heart failure with preserved ejection fraction), n (%)	55 (53.9)	—
Comorbidity (age-adjusted Charlson Comorbidity Index), mean (SD)	6.4 (2.2)	—
Mutuality score, mean (SD)	2.8 (0.7)	2.7 (0.9)

[a]Patients and caregivers are members of the same dyad; therefore, values are presented for descriptive purposes only.

[b]Not applicable.

Digital Literacy and HF Self-Care Outcomes Within Dyads

Patients demonstrated significantly lower digital literacy than caregivers (mean 31.9, SD 20.6 vs mean 77.8, SD 30.8: $P<.001$). HF self-care outcomes were assessed across 3 dimensions: maintenance, symptom perception, and management. Symptom perception was the lowest-scoring dimension for both patients' HF self-care (mean 46.9, SD 16.6) and caregivers' contribution to HF self-care (mean 52.5, SD 19.7). Descriptive statistics of HF self-care outcomes across the 3 dimensions are presented in Table 2.

XSL•FO
RenderX

Table 2. Comparisons of digital literacy and heart failure (HF) self-care outcomes between patients and caregivers (N=102 dyads)[a].

Variables	Patients, mean (SD)	Caregivers, mean (SD)	*t* test (*df*)	*P* value
Digital literacy				
Total score	31.9 (20.6)	77.8 (30.8)	–12.60 (101)	<.001
Safety and security	13.5 (9.5)	33.3 (12.9)	–12.43 (101)	<.001
Information and communication	13.0 (8.1)	31.9 (12.6)	–12.98 (101)	<.001
Contents creation and management	5.4 (3.3)	12.6 (6.1)	–10.51 (101)	<.001
HF self-care				
Maintenance	62.6 (10.2)	59.0 (17.4)	2.05 (101)	.04
Symptom perception	46.6 (16.6)	52.5 (19.7)	–2.66 (101)	.009
Management	62.1 (16.1)	65.0 (18.1)	–1.39 (101)	.17

[a]Comparisons between patients and caregivers were conducted using paired *t* tests (2-tailed), as they were members of the same dyad.

Dyadic Associations Between Digital Literacy and HF Self-Care (APIM Results)

Across the 3 HF self-care dimensions, distinct patterns of association between digital literacy and self-care outcomes were observed. For HF self-care maintenance, digital literacy was not significantly associated with patient or caregiver outcomes, indicating no interpretable actor-partner structure in this dimension. In contrast, for symptom perception, significant actor effects of digital literacy were observed for both members of the dyad. Higher patient digital literacy was associated with better patient symptom perception (B=0.21; β=0.26, SE 0.08, 95% CI 0.06-0.37; *P*=.008), and higher caregiver digital literacy was associated with greater caregiver contribution to symptom perception (B=0.21; β=0.32, SE 0.07, 95% CI 0.08-0.34; *P*=.002), reflecting an actor-dominant pattern with no evidence of cross-partner effects. For self-care management, only the patient's digital literacy demonstrated a significant actor effect on patient self-care outcome (B=0.19; β=0.24, SE 0.08, 95% CI 0.03-0.35; *P*=.02), whereas no other actor or partner effects reached statistical significance for either member of the dyad. Detailed APIM results are presented in Table 3. The model accounted for approximately 22.6% of the variance in patients' symptom perception, 35.3% in caregivers' symptom perception, and 14.6% in patients' self-care management. A summary of standardized actor and partner effects across the 3 APIMs is presented in Figure 1.

XSL•FO
RenderX

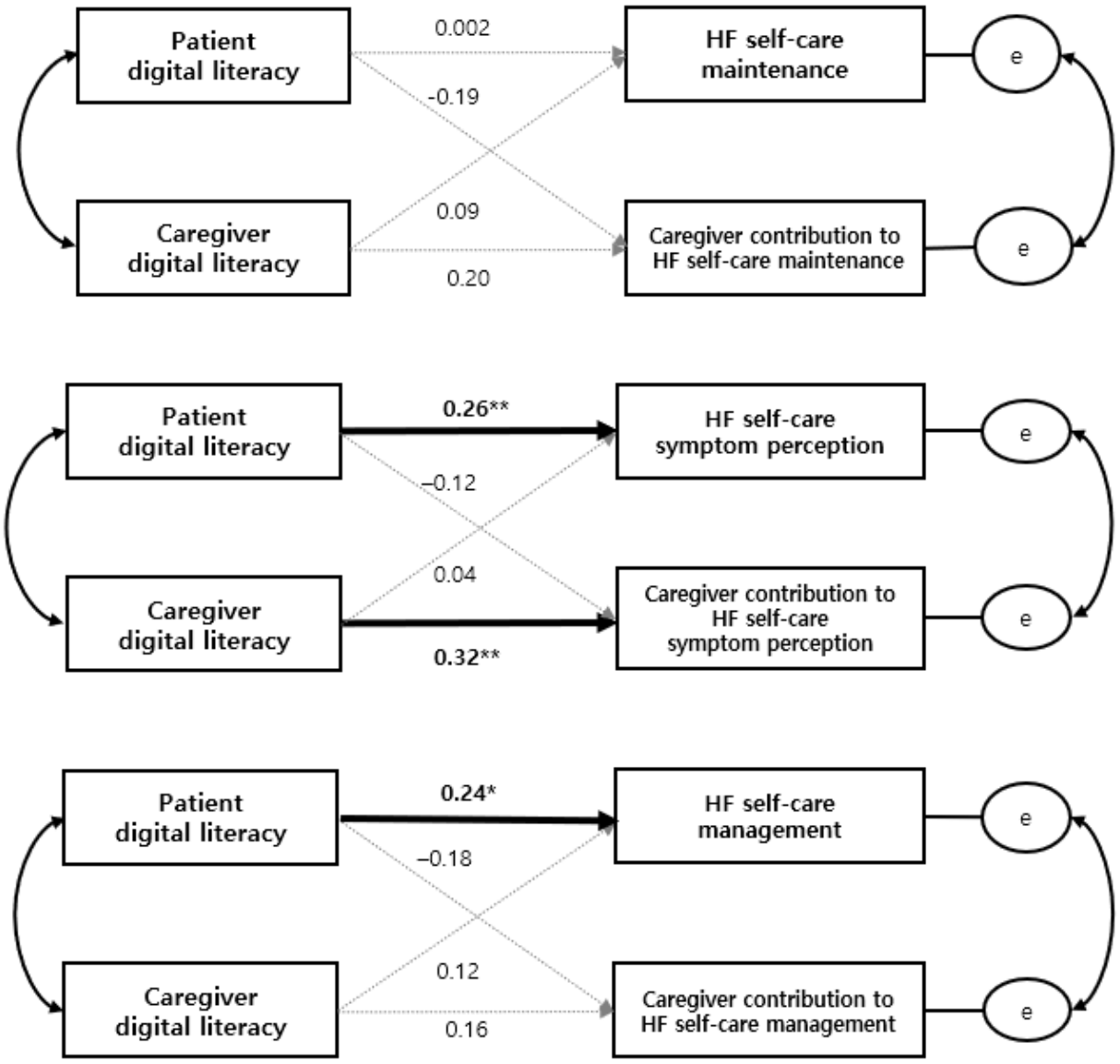
Table 3. Actor and partner effects of digital literacy on heart failure (HF) self-care based on the Actor-Partner Interdependence Model (N=102 dyads).

Outcome or effect, role	B	β (SE; 95% CI)	*P* value
HF self-care maintenance			
Actor effect			
Patient	0.001	0.002 (0.05; –0.10 to 0.11)	.98
Caregiver	0.11	0.20 (0.06; –0.01 to 0.24)	.08
Partner effect			
Patient→caregiver	–0.16	–0.19 (0.08; –0.32 to 0.004)	.06
Caregiver→patient	0.03	0.09 (0.04; –0.05 to 0.11)	.48
HF self-care symptom perception			
Actor effect			
Patient	*0.21* [a]	*0.26 (0.08; 0.06 to 0.37)*	*.008*
Caregiver	*0.21*	*0.32 (0.07; 0.08 to 0.34)*	*.002*
Partner effect			
Patient→caregiver	–0.11	–0.12 (0.09; –0.29 to 0.06)	.19
Caregiver→patient	0.02	0.04 (0.06; –0.10 to 0.14)	.73
HF self-care management			
Actor effect			
Patient	*0.19*	*0.24 (0.08; 0.03 to 0.35)*	*.02*
Caregiver	0.09	0.16 (0.07; –0.03 to 0.22)	.15
Partner effect			
Patient→caregiver	–0.15	–0.18 (0.08; –0.32 to 0.01)	.07
Caregiver→patient	0.06	0.12 (0.06; –0.06 to 0.18)	.34

[a]Italicized values indicate statistically significant effects (*P*<.05).

XSL•FO
RenderX

Figure 1. Actor-Partner Interdependence Model examining the associations between patient and caregiver digital literacy and heart failure (HF) self-care outcomes across 3 dimensions: HF self-care maintenance, HF self-care symptom perception, and HF self-care management. Solid arrows indicate statistically significant actor effects, whereas dashed arrows indicate nonsignificant paths. Standardized coefficients (β) are shown. *P<.05, **P<.01.



Discussion

Principal Findings

To the best of our knowledge, this is the first study to examine the influence of digital literacy on HF self-care using APIM within patient-caregiver dyads. Significant actor effects of digital literacy were found for both patients and caregivers on the symptom perception dimension of HF self-care, whereas no partner effects were observed. Patients with higher digital literacy demonstrated a better ability to detect and interpret symptoms and to engage in effective self-care management. Likewise, caregivers with higher digital literacy showed greater competence in identifying HF-related changes, although their digital literacy did not significantly affect patients' overall maintenance or management behaviors. These findings highlight the independent role of digital literacy in shaping HF self-care behaviors and underscore the need for strategies that enhance digital literacy in both patients and caregivers [45].

Previous research indicates that educational attainment and access to health information are closely related to both digital literacy and HF self-care, particularly symptom perception [12,15,46]. Lower educational levels have been linked to reduced self-care autonomy and delayed recognition of symptom changes, often through the pathway of limited digital literacy [22,25]. In this study, the generally low education level of the older patients, combined with the fact that half of the participants (51/102, 50%) resided in rural areas, likely reflects a broader digital divide [22,25]. These sociocontextual factors may help explain the overall low digital literacy observed in our sample and emphasize the persistent digital disparities faced by older adults with HF [22,23,25].

In this context, caregivers with higher digital literacy demonstrated greater contributions to symptom perception,

XSL•FO
RenderX

enabling faster identification of changes in patients' health status [3,11,12,47]. However, contrary to our initial hypothesis, caregivers' digital literacy did not directly affect maintenance or management dimensions of HF self-care. These findings may also reflect the current design orientation of many digital health tools for HF, which predominantly emphasize information provision and symptom monitoring rather than sustained lifestyle modification [20,21,48,49]. Accordingly, digital literacy may facilitate symptom perception without necessarily translating into long-term behavioral change in HF self-care maintenance and management [50].

This finding may also be explained by entrenched lifestyle patterns among older adults, such as long-standing dietary and physical activity habits that are less amenable to behavioral change [8,16]. Furthermore, as most caregivers in this study were patients' adult children, their capacity to significantly alter patients' long-standing lifestyle patterns—especially regarding diet and physical activity routines—is likely limited within the context of family-centered caregiving norms in East Asian societies [18,29].

These cultural caregiving norms are particularly evident in Korea. In many East Asian societies shaped by Confucian values that prioritize filial piety, adult children frequently take primary responsibility for caregiving [18,29]. However, rapid industrialization and changing social structures have reduced opportunities for intergenerational coresidence, resulting in caregiving that is frequently provided without daily cohabitation [18,51]. In such contexts, caregivers may play more of a technical or assisting role—helping with device use or symptom monitoring—rather than engaging in continuous motivational or educational support [26,27]. This relational pattern may reduce the potential for partner effects on self-care behaviors. This finding contrasts with previous research that has emphasized caregiver support—particularly within spousal dyads—as an important determinant of adherence [9,31,32].

Finally, we adjusted for mutuality—the quality of reciprocal understanding and rapport between patient and caregiver—to isolate the effects of digital literacy [31,36]. Because mutuality is known to influence communication, decision-making, and joint self-care [31,36,37,39], future studies should explore how digital literacy and mutuality interact over time or across caregiving stages to shape dyadic engagement in digital health–based management.

Limitations

Despite the valuable insights provided by this study, several limitations should be noted. First, caregiving in South Korea often involves multiple family members rather than a single designated caregiver, which may limit the transferability of the findings. Second, the study was conducted in a single center with mixed urban and rural characteristics, and the limited sample size may affect generalizability. Third, other relevant factors, including cognitive function and depressive symptoms, were not examined and should be considered in future dyadic studies of digital health–supported self-care. Finally, the reliance on self-reported data may have introduced bias; future studies should incorporate objective measures of digital literacy, self-care behaviors, and clinical outcomes (eg, hospitalization or mortality) to better assess health impact.

Conclusions

This dyadic analysis indicates that digital literacy plays a significant role in HF self-care, particularly in enhancing symptom perception for both patients and caregivers, while cross-partner influences were not evident. The findings suggest that digital literacy interventions should engage both members of the dyad, with strategies tailored to caregiving relationships, cultural norms, and individual readiness. To bridge the gap between symptom perception and behavior change, future digital health interventions should integrate more interactive, habit-forming features beyond simple monitoring. In addition, considering older adults' ingrained habits, mutuality, and sociocultural caregiving patterns will be essential in designing effective, equitable digital self-care programs. Future longitudinal and intervention studies are needed to validate these pathways and optimize collaborative self-care in diverse populations with HF.

Acknowledgments

The authors thank the outpatient clinic staff who assisted with participant recruitment and coordination of data collection, as well as the research assistants who supported data management. The authors also appreciate the collaborative efforts of all coauthors in completing this research.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Funding

This research was supported by the Brain Korea 21 FOUR Project through the National Research Foundation of Korea, Yonsei University College of Nursing, and a grant from the National Research Foundation of Korea funded by the Ministry of Education (grant 2340046871).

Authors' Contributions

HK conceptualized the study, developed the overall research design, and drafted the manuscript. SHC supervised the research, provided overall guidance, and managed the study process. MLC contributed to the theoretical framework and refinement of the

XSL•FO
RenderX

research concept and provided statistical consultation. JYC and YJ reviewed the study design and manuscript draft. BSY assisted with data collection and manuscript review. All authors contributed substantially to the work and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Bivariate correlations among key variables (N=102 dyads).
[DOCX File , 18 KB - aging_v9i1e85976_app1.docx]

References

1. Lee CJ, Lee H, Yoon M, Chun KH, Kong MG, Jung MH, et al. Heart failure statistics 2024 update: a report from the Korean Society of Heart Failure. Int J Heart Fail 2024 Apr;6(2):56-69 [FREE Full text] [doi: 10.36628/ijhf.2024.0010] [Medline: 38694933]
2. Bui AL, Horwich TB, Fonarow GC. Epidemiology and risk profile of heart failure. Nat Rev Cardiol 2011 Jan;8(1):30-41 [FREE Full text] [doi: 10.1038/nrcardio.2010.165] [Medline: 21060326]
3. Buck HG, Howland C, Stawnychy MA, Aldossary H, Cortés YI, DeBerg J, et al. Caregivers' contributions to heart failure self-care: an updated systematic review. J Cardiovasc Nurs 2024;39(3):266-278 [FREE Full text] [doi: 10.1097/JCN.0000000000001060] [Medline: 38306302]
4. Anderson KM, Murphy D, Groninger H, Kolm P, Wang H, Barton-Maxwel V. Perceived symptoms as the primary indicators for 30-day heart failure readmission. PLoS One 2022 May 5;17(5):e0267820 [FREE Full text] [doi: 10.1371/journal.pone.0267820] [Medline: 35511916]
5. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur J Heart Fail 2016 Aug;18(8):891-975 [FREE Full text] [doi: 10.1002/ejhf.592] [Medline: 27207191]
6. Riegel B, Dickson VV, Faulkner KM. The situation-specific theory of heart failure self-care: revised and updated. J Cardiovasc Nurs 2016;31(3):226-235. [doi: 10.1097/JCN.0000000000000244] [Medline: 25774844]
7. Riegel B, Dickson VV. A situation-specific theory of heart failure self-care. J Cardiovasc Nurs 2008;23(3):190-196. [doi: 10.1097/01.JCN.0000305091.35259.85] [Medline: 18437059]
8. Aghajanloo A, Negarandeh R, Janani L, Tanha K, Hoseini-Esfidarjani SS. Self-care status in patients with heart failure: systematic review and meta-analysis. Nurs Open 2021 Sep 23;8(5):2235-2248 [FREE Full text] [doi: 10.1002/nop2.805] [Medline: 33619877]
9. Deek H, Chang S, Newton PJ, Noureddine S, Inglis SC, Arab GA, et al. An evaluation of involving family caregivers in the self-care of heart failure patients on hospital readmission: randomised controlled trial (the FAMILY study). Int J Nurs Stud 2017 Oct;75:101-111. [doi: 10.1016/j.ijnurstu.2017.07.015] [Medline: 28772186]
10. Ryou C, Kang SM, Jang Y. Factors associated with self-care behaviours among Koreans with heart failure. Eur J Cardiovasc Nurs 2021 Mar 01;20(3):276-284. [doi: 10.1177/1474515120934060] [Medline: 33570596]
11. Riegel B, De Maria M, Barbaranelli C, Matarese M, Ausili D, Stromberg A, et al. Symptom recognition as a mediator in the self-care of chronic illness. Front Public Health 2022 May 17;10:883299 [FREE Full text] [doi: 10.3389/fpubh.2022.883299] [Medline: 35655456]
12. Santos GC, Liljeroos M, Dwyer AA, Jaques C, Girard J, Strömberg A, et al. Symptom perception in heart failure: a scoping review on definition, factors and instruments. Eur J Cardiovasc Nurs 2020 Feb;19(2):100-117 [FREE Full text] [doi: 10.1177/1474515119892797] [Medline: 31782668]
13. Lee S, Riegel B. State of the science in heart failure symptom perception research: an integrative review. J Cardiovasc Nurs 2018;33(3):204-210. [doi: 10.1097/JCN.0000000000000445] [Medline: 28858886]
14. Wu JR, Lin CY, Latimer A, Hammash M, Moser DK. Mediators of the association between cognitive function and self-care behaviors in patients hospitalized with an exacerbation of heart failure. J Cardiovasc Nurs 2024;39(3):237-244. [doi: 10.1097/JCN.0000000000001066] [Medline: 38099586]
15. Jiang Y, Zhang C, Hong J, Tam WW, Ramachandran HJ, Wang W. Relationships of person-related, problem-related, and environment-related factors to self-care behaviours and the mediating role of self-confidence among patients with heart failure: cross-sectional analysis of structural equation modelling. Int J Nurs Stud 2023 Nov;147:104590. [doi: 10.1016/j.ijnurstu.2023.104590] [Medline: 37741260]
16. Jaarsma T, Hill L, Bayes-Genis A, La Rocca HB, Castiello T, Čelutkienė J, et al. Self-care of heart failure patients: practical management recommendations from the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2021 Jan;23(1):157-174 [FREE Full text] [doi: 10.1002/ejhf.2008] [Medline: 32945600]

XSL•FO
RenderX

17. Vellone E, Riegel B, Alvaro R. A situation-specific theory of caregiver contributions to heart failure self-care. J Cardiovasc Nurs 2019;34(2):166-173. [doi: 10.1097/JCN.0000000000000549] [Medline: 30363017]
18. Son YJ. Dyadic research of patients and their family caregivers in the context of chronic illness: current status and challenges. Korean J Adult Nurs 2021 Jan;33(5):415. [doi: 10.7475/kjan.2021.33.5.415]
19. Uchmanowicz I, Faulkner KM, Vellone E, Siennicka A, Szczepanowski R, Olchowska-Kotala A. Heart failure care: testing dyadic dynamics using the actor-partner interdependence model (APIM)-a scoping review. Int J Environ Res Public Health 2022 Feb 09;19(4):1919 [FREE Full text] [doi: 10.3390/ijerph19041919] [Medline: 35206131]
20. Singhal A, Cowie MR. Digital health: implications for heart failure management. Card Fail Rev 2021 May 11;7:e08 [FREE Full text] [doi: 10.15420/cfr.2020.28] [Medline: 34035953]
21. Mortara A, Vaira L, Palmieri V, Iacoviello M, Battistoni I, Iacovoni A, et al. Would you prescribe mobile health apps for heart failure self-care? An integrated review of commercially available mobile technology for heart failure patients. Card Fail Rev 2020 Mar;6:e13 [FREE Full text] [doi: 10.15420/cfr.2019.11] [Medline: 32537246]
22. Oh SS, Kim KA, Kim M, Oh J, Chu SH, Choi J. Measurement of digital literacy among older adults: systematic review. J Med Internet Res 2021 Feb 03;23(2):e26145 [FREE Full text] [doi: 10.2196/26145] [Medline: 33533727]
23. Rodríguez Parrado IY, Achury Saldaña DM. Digital health literacy in patients with heart failure in times of pandemic. Comput Inform Nurs 2022 Nov 01;40(11):754-762 [FREE Full text] [doi: 10.1097/CIN.0000000000000883] [Medline: 35234702]
24. Raja M, Bjerkan J, Kymre IG, Galvin KT, Uhrenfeldt L. Telehealth and digital developments in society that persons 75 years and older in European countries have been part of: a scoping review. BMC Health Serv Res 2021 Oct 26;21(1):1157 [FREE Full text] [doi: 10.1186/s12913-021-07154-0] [Medline: 34696789]
25. Arcury TA, Sandberg JC, Melius KP, Quandt SA, Leng X, Latulipe C, et al. Older adult internet use and eHealth literacy. J Appl Gerontol 2020 Feb;39(2):141-150 [FREE Full text] [doi: 10.1177/0733464818807468] [Medline: 30353776]
26. Baik D, Reeder B, Coats H, Baker C, Jankowski C. Perceptions and attitudes toward a proposed digital health physical activity program among older family caregivers of persons with heart failure: a qualitative study. Inform Health Soc Care 2023 Jul 03;48(3):239-251 [FREE Full text] [doi: 10.1080/17538157.2023.2227704] [Medline: 37417465]
27. Wali S, Keshavjee K, Nguyen L, Mbuagbaw L, Demers C. Using an electronic app to promote home-based self-care in older patients with heart failure: qualitative study on patient and informal caregiver challenges. JMIR Cardio 2020 Nov 09;4(1):e15885 [FREE Full text] [doi: 10.2196/15885] [Medline: 33164901]
28. Kenny DA, Cook W. Partner effects in relationship research: conceptual issues, analytic difficulties, and illustrations. Pers Relatsh 1999 Dec;6(4):433-448. [doi: 10.1111/j.1475-6811.1999.tb00202.x]
29. Wang Q, Xiao X, Zhang J, Jiang D, Wilson A, Qian B, et al. The experiences of East Asian dementia caregivers in filial culture: a systematic review and meta-analysis. Front Psychiatry 2023 Apr 21;14:1173755 [FREE Full text] [doi: 10.3389/fpsyt.2023.1173755] [Medline: 37151975]
30. Steinberg RS, Nayak A, Burke MA, Aldridge M, Raja Laskar S, Bhatt K, et al. Association of race and gender with primary caregiver relationships and eligibility for advanced heart failure therapies. Clin Transplant 2022 Jan;36(1):e14502 [FREE Full text] [doi: 10.1111/ctr.14502] [Medline: 34634150]
31. Hooker SA, Schmiege SJ, Trivedi RB, Amoyal NR, Bekelman DB. Mutuality and heart failure self-care in patients and their informal caregivers. Eur J Cardiovasc Nurs 2018 Feb;17(2):102-113 [FREE Full text] [doi: 10.1177/1474515117730184] [Medline: 28868917]
32. Trivedi RB, Piette J, Fihn SD, Edelman D. Examining the interrelatedness of patient and spousal stress in heart failure: conceptual model and pilot data. J Cardiovasc Nurs 2012;27(1):24-32. [doi: 10.1097/JCN.0b013e3182129ce7] [Medline: 21743348]
33. Choi J, Choi S, Song K, Baek J, Kim H, Choi M, et al. Everyday digital literacy questionnaire for older adults: instrument development and validation study. J Med Internet Res 2023 Dec 14;25:e51616 [FREE Full text] [doi: 10.2196/51616] [Medline: 38095999]
34. Kim J, Kim KH, Lim YH, Heo S, Moon K, Oh MS, et al. Validity and reliability of the Korean version of the revised Self-Care of Heart Failure Index V7.2. Clin Nurs Res 2022 Sep;31(7):1296-1307. [doi: 10.1177/10547738221106590] [Medline: 35726492]
35. Vellone E, Barbaranelli C, Pucciarelli G, Zeffiro V, Alvaro R, Riegel B. Validity and reliability of the Caregiver Contribution to Self-Care of Heart Failure Index version 2. J Cardiovasc Nurs 2020;35(3):280-290. [doi: 10.1097/JCN.0000000000000655] [Medline: 32084084]
36. Archbold PG, Stewart BJ, Greenlick MR, Harvath T. Mutuality and preparedness as predictors of caregiver role strain. Res Nurs Health 1990 Dec;13(6):375-384. [doi: 10.1002/nur.4770130605] [Medline: 2270302]
37. Vellone E, Chung ML, Alvaro R, Paturzo M, Dellafiore F. The influence of mutuality on self-care in heart failure patients and caregivers: a dyadic analysis. J Fam Nurs 2018 Nov 19;24(4):563-584. [doi: 10.1177/1074840718809484] [Medline: 30453797]
38. Cook WL, Kenny DA. The actor-partner interdependence model: a model of bidirectional effects in developmental studies. Int J Behav Dev 2005;29(2):101-109. [doi: 10.1080/01650250444000405]

XSL•FO
RenderX

39. Dellafiore F, Chung ML, Alvaro R, Zeffiro V, Ercole V, Pucciarelli G. Influence of mutuality on quality of life in heart failure patient with inadequate self-care and caregiver dyads: an actor-partner interdependence model analysis. Eur J Cardiovasc Nurs 2022 Jun 02;21(4):366-373. [doi: 10.1093/eurjcn/zvab089] [Medline: 34661635]
40. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index. J Clin Epidemiol 1994 Nov;47(11):1245-1251. [doi: 10.1016/0895-4356(94)90129-5] [Medline: 7722560]
41. Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychol Methods 1996;1(1):16-29. [doi: 10.1037/1082-989x.1.1.16]
42. Kline RB. Principles and Practice of Structural Equation Modeling 4th Edition. New York, NY: The Guilford Press; 2016.
43. Li CH. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. Behav Res Methods 2016 Sep;48(3):936-949. [doi: 10.3758/s13428-015-0619-7] [Medline: 26174714]
44. Wolff JC, Thomas SA, Hood E, Bettis AH, Rizzo CJ, Liu RT. Application of the actor-partner interdependence model in parent-adolescent emotion regulation and depression. J Affect Disord 2020 Dec 01;277:733-741 [FREE Full text] [doi: 10.1016/j.jad.2020.08.075] [Medline: 32919294]
45. Buck HG, Stromberg A, Chung ML, Donovan KA, Harkness K, Howard AM, et al. A systematic review of heart failure dyadic self-care interventions focusing on intervention components, contexts, and outcomes. Int J Nurs Stud 2018 Jan;77:232-242 [FREE Full text] [doi: 10.1016/j.ijnurstu.2017.10.007] [Medline: 29128777]
46. Wu JR, Lin CY, Hammash M, Moser DK. Heart failure knowledge, symptom perception, and symptom management in patients with heart failure. J Cardiovasc Nurs 2022 Dec 20 [FREE Full text] [doi: 10.1097/JCN.0000000000000961] [Medline: 36542682]
47. Bidwell JT, Hostinar CE, Higgins MK, Abshire MA, Cothran F, Butts B, et al. Caregiver subjective and physiological markers of stress and patient heart failure severity in family care dyads. Psychoneuroendocrinology 2021 Nov;133:105399 [FREE Full text] [doi: 10.1016/j.psyneuen.2021.105399] [Medline: 34482256]
48. Zisis G, Carrington MJ, Oldenburg B, Whitmore K, Lay M, Huynh Q, et al. An m-Health intervention to improve education, self-management, and outcomes in patients admitted for acute decompensated heart failure: barriers to effective implementation. Eur Heart J Digit Health 2021 Nov 13;2(4):649-657 [FREE Full text] [doi: 10.1093/ehjdh/ztab085] [Medline: 36713108]
49. Portz JD, Vehovec A, Dolansky MA, Levin JB, Bull S, Boxer R. The development and acceptability of a mobile application for tracking symptoms of heart failure among older adults. Telemed J E Health 2018 Feb;24(2):161-165 [FREE Full text] [doi: 10.1089/tmj.2017.0036] [Medline: 28696832]
50. Wright PJ, Raynor PA, Bowers D, Combs EM, Corbett CF, Hardy H, et al. Leveraging digital technology for social connectedness among adults with chronic conditions: a systematic review. Digit Health 2023 Oct 3;9:20552076231204746 [FREE Full text] [doi: 10.1177/20552076231204746] [Medline: 37799504]
51. Lee EE, Lee SY. Caregiving experiences of Korean family caregivers of cancer patients: an integrative literature review. Psychooncology 2020 Oct;29(10):1486-1503. [doi: 10.1002/pon.5498] [Medline: 32720400]

Abbreviations

ACCI: age-adjusted Charlson Comorbidity Index
APIM: Actor-Partner Interdependence Model
EDLQ: Everyday Digital Literacy Questionnaire
HF: heart failure
ML: maximum likelihood

Edited by J Wang; submitted 16.Oct.2025; peer-reviewed by JR de-Moya-Romero; comments to author 20.Nov.2025; accepted 16.Feb.2026; published 12.Mar.2026.

Please cite as:
Kim H, Chung ML, Jang Y, Choi J, Yoo BS, Chu SH
Digital Literacy and Heart Failure Self-Care in Older Patients and Their Caregivers: Dyadic Analysis Using the Actor-Partner Interdependence Model
JMIR Aging 2026;9:e85976
URL: https://aging.jmir.org/2026/1/e85976
doi:10.2196/85976
PMID:41736640

©Hokon Kim, Misook L Chung, Yeonsoo Jang, JiYeon Choi, Byung Su Yoo, Sang Hui Chu. Originally published in JMIR Aging (https://aging.jmir.org), 12.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons

XSL•FO
RenderX

Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Dynamic Relationship Between Sleep Patterns and Behavioral and Psychological Symptoms of Dementia: Longitudinal Observational Study

Eunhee Cho[1], PhD; Sinwoo Hwang[1,2], MSN; Minhee Yang[1], PhD; Eunkyo Kim[1,3], MSN; Jungwon Cho[1,3], MSN; Chang Park[4], PhD

[1]Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

[2]Army Cadet Military School, Goesan, Republic of Korea

[3]College of Nursing and Brain Korea 21 FOUR Project, Yonsei University, Seoul, Republic of Korea

[4]College of Nursing, University of Illinois Chicago, Chicago, IL, United States

Corresponding Author:
Sinwoo Hwang, MSN
Mo-Im Kim Nursing Research Institute
College of Nursing
Yonsei University
50-1 Yonsei-ro
Seoul, 03722
Republic of Korea
Phone: 82 01037217432
Email: sinwoo3034@gmail.com

Abstract

Background: A higher prevalence of behavioral and psychological symptoms of dementia is associated with a greater caregiver burden and increased mortality in people with dementia. Considering the possibility of a reciprocal relationship between sleep disturbances and these symptoms, time series analyses are necessary to explore the associated temporal dynamics.

Objective: This study aimed to examine dynamic interdependencies between sleep disturbances and behavioral and psychological symptoms of dementia in older adults.

Methods: Daily interactions between sleep patterns and behavioral and psychological symptoms of dementia were analyzed over a 14-day period using a panel vector autoregressive model. Data were collected from June 2018 to June 2020 in community and institutional settings. A total of 154 older adults with dementia wore wrist actigraphy devices continuously for 2 weeks for sleep data, and caregivers recorded behavioral and psychological symptoms of dementia in a daily symptom diary.

Results: Using a panel vector autoregressive model, we analyzed data from 154 older adults living with dementia and their caregivers. The results showed unidirectional Granger causality running from the number of awakenings on the previous day to irritability (P=.03) and appetite or eating disorders (P=.04) on the following day. Conversely, some of the previous day's behavioral and psychological symptoms of dementia temporally preceded subsequent changes in sleep patterns. Specifically, delusions had a Granger-causality effect on total sleep time (P<.001), wake after sleep onset (P=.01), and the number of awakenings (P=.006), while irritability had a Granger causality effect on the number of awakenings (P=.007). Notably, bidirectional Granger causality was observed between irritability and the number of awakenings.

Conclusions: This study demonstrates that the relationship between the behavioral and psychological symptoms of dementia and sleep patterns is dynamic and forms a vicious cycle. Consequently, early intervention to alleviate symptoms is imperative, and strategies to enhance sleep quality and address sleep disturbances should be prioritized.

(JMIR Aging 2026;9:e80422) doi:10.2196/80422

KEYWORDS

behavioral and psychological symptoms; dementia; granger causality; panel vector autoregressive model; sleep

XSL•FO
RenderX

Introduction

Behavioral and psychological symptoms of dementia encompass a range of noncognitive disturbances [1]. The Neuropsychiatric Inventory, a scale widely used for assessing the behavioral and psychological symptoms of dementia, identifies 12 subsymptoms, including delusion, hallucination, anxiety, irritability, depression, sleep disturbances, apathy, disinhibition, changes in appetite, and agitation [1,2]. These behavioral and psychological symptoms can manifest at any stage of dementia, with their prevalence, frequency, and severity often increasing as the condition progresses [3]. Recent evidence indicates that approximately 90% of people living with dementia exhibit at least 1 behavioral and psychological symptom [4]. A higher prevalence of such symptoms is associated with a higher care burden in caregivers and increased mortality among people living with dementia [5-7]. In light of these observations, securing effective management of the behavioral and psychological symptoms of dementia is crucial to alleviating caregiver burden and improving the quality of life for those living with dementia. Among the factors contributing to behavioral and psychological symptoms of dementia, unmet physical and psychological needs have been identified as key drivers [8,9], with sleep disturbances playing a particularly important role. Individuals living with dementia commonly experience sleep disturbances [10,11], and accumulating evidence indicates that poor sleep quality is associated with behavioral and psychological symptoms of dementia, such as apathy and depression [12-14]. The existing evidence suggests that sleep patterns may correlate with the occurrence of behavioral and psychological symptoms of dementia [15-17].

Several studies have explored the relationship between sleep disturbances and the behavioral and psychological symptoms of dementia. Some studies suggest that sleep disturbances contribute to the development of such symptoms [18-20], while others claim that such symptoms may affect sleep disturbances the following day [12]. Cho et al [20] identified that sleep disturbances are a risk factor for behavioral and psychological symptoms of dementia in patients living with the condition. Nonetheless, these results have definite drawbacks because they assume a unidirectional relationship. Given the complexity of these interactions, there is a need to examine the temporal dynamics and directionality of the relationship between sleep patterns and behavioral and psychological symptoms of dementia.

Accordingly, this study used a panel vector autoregressive approach to examine day-to-day temporal dynamics between actigraphy-based sleep patterns and caregiver-recorded behavioral and psychological subsymptoms of dementia, using data collected over a 14-day period. Given the potential variability in the associations between behavioral and psychological subsymptoms of dementia and sleep patterns, we initially used a linear mixed model to account for between-participant variability. However, this approach is limited in its ability to capture dynamic changes over time and directional temporal associations. To address these limitations, we subsequently used a panel vector autoregressive model, which treats all variables as endogenous and enables the examination of reciprocal temporal interdependencies and autoregressive structures over time. This approach is well suited for exploratory analyses of complex temporal relationships without imposing predefined causal directions [21]. Accordingly, this approach allowed for an exploratory examination of day-to-day temporal dynamics between objectively measured sleep disturbances and behavioral and psychological symptoms of dementia, including 12 specific subsymptoms. By applying this framework, this study aims to provide a more nuanced understanding of the temporal interdependence between behavioral and psychological symptoms of dementia and sleep disturbances at the daily level.

Methods

Design

This study used a prospective observational design and applied a panel vector autoregressive model to explore the dynamic interdependencies between behavioral and psychological symptoms of dementia and sleep patterns.

Participants

Older adults with dementia were recruited from neurology outpatient departments of tertiary hospitals, daycare centers, and nursing homes in South Korea. The inclusion criteria for the participants were as follows: age 65 years or older, a diagnosis of dementia, a Korean Mini-Mental State Examination score less than 24 points, and the presence of behavioral and psychological symptoms of dementia at least once a week at baseline. Primary caregivers were also included if they provided the majority of care for the recruited older adults and were able to read and write in Korean. The data were collected from June 2018 to June 2020.

Procedure

Eligibility screening was conducted by trained researchers through medical record reviews and participant interviews. Two researchers visited the participants' homes or nursing homes on 2 occasions to collect data. Demographic and health data were collected during the first visit. Moreover, older adults living with dementia were instructed to wear an actigraphy device on their wrists continuously for 2 weeks, and caregivers were asked to record the behavioral and psychological symptoms of dementia in a daily diary for 2 weeks. The researchers conducted telephone follow-ups with participants and their caregivers at least 3 times during the 2-week period to monitor participation status and document any notable observations or issues. At the second visit, which was conducted 2 weeks later, the behavioral and psychological symptoms of dementia were retrieved from the diaries and actigraphy devices for data extraction and analysis.

Measurements

Demographic and Health Data

Demographic data, including age, sex, marital status, and education level, along with information on activities of daily living, Mini-Mental State Examination scores, use of sedative medication, and type of dementia diagnosis, were collected from the caregivers and participants' medical records.

XSL•FO
RenderX

Behavioral and Psychological Symptoms of Dementia

A diary consisting of checklists was developed based on the Neuropsychiatric Inventory [22], a standardized instrument for measuring 12 neuropsychiatric symptoms in people living with dementia. Caregivers were asked to record behavioral and psychological symptoms of dementia in a diary daily for 2 weeks. These symptoms were delusions, hallucinations, agitation or aggression, depression, anxiety, euphoria or elation, apathy, disinhibition, irritability, aberrant motor behavior, sleep and nighttime behavior, and appetite or eating disorders.

Sleep Patterns

Sleep patterns were objectively measured using an actigraphy device, the wGT3X-BT activity monitor (ActiGraph) [23]. Participants wore the device on their wrists for 2 weeks, and sleep data were extracted by applying the Cole-Kripke algorithm in ActiLife (version 6.13.3; Ametris) software. The following variables were derived from the daily sleep pattern data, with midnight as the reference point: total sleep time (ie, the time measured as sleep in a day), wake after sleep onset (ie, the time measured as wakefulness during sleep), the number of awakenings, and mean awakening length (ie, the average length of each awakening in minutes).

Ethical Considerations

The data collection procedures for the original study received ethics approval from the institutional review boards of Severance Hospital (4-2018-0348, 4-2019-0314, and 4-2020-0454) and Ilsan Hospital (2018-10-002-001 and 2019-08-012-001). In the original study, informed consent procedures included explicit permission for future secondary analyses of the collected data without the need for additional consent. Participants who completed the study received approximately US $38 as compensation for their participation. All data used in this study were deidentified before access by the research team. Study-related documents and electronic files were stored on a secure, password-protected computer with access limited to members of the authorized research team.

Statistical Analyses

All statistical analyses were performed using Stata (version 16.1; Stata Corp). Given the substantial individual variability in the behavioral and psychological symptoms of dementia and sleep patterns, we first applied a linear mixed model to account for between-participant differences and examine individual-level heterogeneity. We calculated the intraclass correlation coefficient (ICC) to quantify the proportion of the total variance attributable to individual differences. This coefficient indicates the extent to which the variability in a construct is explained by between-participant differences, rather than within-individual fluctuations. A higher ICC suggests that individual characteristics contribute to the observed variability in the behavioral and psychological symptoms of dementia and sleep patterns. However, as the linear mixed model has limitations in capturing bidirectional relationships, fully reflecting dynamic interactions over time, and testing causality, we further used the panel vector autoregressive model to analyze the dynamic interplay between behavioral and psychological symptoms of dementia and sleep patterns.

For the panel vector autoregressive estimation, the optimal lag length was determined as lag 1 based on the minimum modified Bayesian information criterion, modified Akaike information criterion, and modified Hannan-Quinn information criterion statistics [24] (Table S1 in Multimedia Appendix 1). To assess causality within this framework, we conducted a Granger causality test, where the null hypothesis assumed no causal relationship, and a P value <.05 indicated rejection of the null hypothesis (ie, there was a Granger causal effect). This test also allowed us to determine whether the relationship was unidirectional or bidirectional. Statistical significance was set at P<.05.

Results

Participants' General Characteristics, Behavioral and Psychological Symptoms of Dementia, and Sleep Patterns

We initially recruited 175 older adults with dementia, but 18 (10.3%) were excluded owing to refusal to wear the actigraphy device (n=10, 55.6%), hospitalization (n=6, 33.3%), death (n=1, 5.6%), and absence of a caregiver during data collection (n=1, 5.6%). We also excluded 3 (1.7%) participants with insufficient data from the actigraphy device, leaving a final sample of 154 (88%) older adults living with dementia and their caregivers.

Table S2 in Multimedia Appendix 1 shows the baseline characteristics of the participants. The mean age was 81.34 (SD 5.97) years, 93 (60.4%) were women, and most had an elementary school education or less (n=70, 49.4%). The mean activities of daily living score was 10.68 (SD 3.87), indicating difficulty in performing activities of daily living. The mean Mini-Mental State Examination score was 16.86 (SD 5.84), indicating moderate cognitive impairment. Regarding the type of dementia diagnosis, 76 (49.4%) participants were diagnosed with Alzheimer disease, 60 (39%) with Lewy body dementia, 24 (15.6%) with vascular dementia, and 34 (22.1%) with other types of dementia.

Table 1 shows the participants' behavioral and psychological symptoms of dementia and sleep patterns, with a mean of 12.85 (SD 2.40) days of data recorded per participant. Of the 2145 days on which symptoms were recorded, apathy was the most common symptom (n=293, 13.7%), followed by sleep and nighttime behavior (n=259, 12.1%), anxiety (n=238, 11.1%), irritability (n=231, 10.8%), depression (n=230, 10.7%), agitation or aggression (n=194, 9%), and delusion (n=192, 8.6%). The mean total sleep time was 458.21 (SD 253.05) minutes, with a mean of 29.37 (SD 27.47) minutes spent awake during sleep. The total number of awakenings was 12.40 (SD 10.25) per night, with each awakening lasting approximately 2.16 (SD 1.09) minutes.

XSL•FO
RenderX

Table 1. Participants' behavioral and psychological symptoms of dementia and sleep patterns (N=2145 days).

Variables	Values
Number of days of behavioral and psychological symptoms of dementia and actigraphy recordings per person, mean (SD)	12.85 (2.40)
Behavioral and psychological symptoms of dementia, n (%)	
Delusion	192 (8.6)
Hallucination	129 (5.9)
Agitation or aggression	194 (9.0)
Depression	230 (10.7)
Anxiety	238 (11.1)
Euphoria or elation	108 (4.9)
Apathy	293 (13.7)
Disinhibition	41 (1.9)
Irritability	231 (10.8)
Aberrant motor behavior	73 (3.4)
Sleep and nighttime behavior	259 (12.1)
Appetite or eating disorders	157 (7.3)
Sleep factors, mean (SD)	
Total sleep time (minutes)	458.21 (253.05)
Wake after sleep onset (minutes)	29.37 (27.47)
Number of awakenings	12.40 (10.25)
Mean awake length (minutes)	2.16 (1.09)

Linear Mixed Model Estimation Results for Total Behavioral and Psychological Symptoms of Dementia and Sleep Patterns

On the basis of the linear mixed model analysis (Table S3 in Multimedia Appendix 1), there was substantial between-participant variability in the behavioral and psychological symptoms of dementia; however, the fixed effects of sleep patterns on these symptoms were not statistically significant in this model. The ICC values suggested that approximately 49% to 50% of the total variance in the behavioral and psychological symptoms of dementia can be attributed to between-participant differences rather than within-participant fluctuations over time. The random slopes were statistically significant, indicating that the effect of the previous day's symptoms on the following day's symptoms varied between individuals.

Similarly, the total behavioral and psychological symptoms of dementia did not have a statistically significant fixed effect on sleep patterns (Table S4 in Multimedia Appendix 1). However, the ICCs ranged from 0.097 to 0.391, indicating that individual-level differences contributed moderately to the overall variance in sleep patterns, particularly for wake after sleep onset (0.339) and the number of awakenings (0.391), where the proportion of variance due to individual differences was relatively high.

Given the statistically significant random slopes, the relationship between the behavioral and psychological symptoms of dementia and sleep may be strongly influenced by individual differences. This required further analysis using a panel vector autoregressive model to capture the within-individual temporal dynamics.

Panel Vector Autoregressive Model Estimation Results for Total Behavioral and Psychological Symptoms of Dementia and Sleep Patterns

Table 2 shows the results of the panel vector autoregressive model estimation, with behavioral and psychological symptoms of dementia as the response variable. The previous day's wake after sleep onset (P=.01), the number of awakenings (P=.008), and mean awakening length (P=.008) were temporally associated with higher levels of the following day's total behavioral and psychological symptoms of dementia.

XSL•FO
RenderX

Table 2. Panel vector autoregressive model estimation results for total behavioral and psychological symptoms of dementia by sleep pattern.

Impulse variables	Response variables: total behavioral and psychological symptoms of dementia	
	β (SE)	*P* value[a]
Total sleep time	0.00089 (0.00056)	.11
Wake after sleep onset	0.00495 (0.00200)	.01
Number of awakenings	0.01580 (0.00597)	.008
Mean awakening length	0.15242 (0.05767)	.008

[a]*P* value represent the results of the Granger causality test within the panel vector autoregressive model.

Table 3 shows the results of the panel vector autoregressive model estimation using sleep patterns as the response variable. The previous day's behavioral and psychological symptoms of dementia were temporally associated with higher total sleep time (*P*<.001), greater wake after sleep onset (*P*=.003), and a higher number of awakenings (*P*=.001) on the following day.

Table 3. Panel vector autoregressive model estimation results for sleep pattern by total behavioral and psychological symptoms of dementia.

Impulse variables: total behavioral and psychological symptoms of dementia	Response variables	
	β (SE)	*P* value[a]
Total sleep time	90.49515 (25.21297)	<.001
Wake after sleep onset	4.34126 (1.47666)	.003
Number of awakenings	1.83004 (0.56304)	.001
Mean awakening length	0.00914 (0.07713)	.91

[a]*P* value represent the results of the Granger causality test within the panel vector autoregressive model.

Granger causality tests (Tables 2 and 3) were conducted to test for causality between the variables that had a statistically significant relationship. The tests revealed unidirectional Granger causality, consistent with the panel vector autoregressive model estimates.

Linear Mixed Model Estimation Results for Behavioral and Psychological Subsymptoms of Dementia and Sleep Patterns

Tables S5 and S6 in Multimedia Appendix 1 show the relationship between the 12 behavioral and psychological subsymptoms of dementia and sleep patterns as analyzed by the linear mixed model. The findings resemble those of the linear mixed model analysis of total symptoms and sleep patterns. Few of the fixed effects of sleep patterns on the behavioral and psychological subsymptoms of dementia were statistically significant, but the ICC values ranged from 0.57 to 0.83, indicating that much of the variation in the subsymptoms could be attributed to between-individual differences (Table S5 in Multimedia Appendix 1).

Similarly, the 12 behavioral and psychological subsymptoms of dementia did not exhibit a statistically significant fixed effect on sleep patterns (Table S6 in Multimedia Appendix 1). However, the ICC values, ranging from 0.099 to 0.391, suggest that individual differences account for a moderate proportion of the overall variance in sleep patterns. Notably, wake after sleep onset (ICC 0.337-0.340) and the number of awakenings (ICC 0.389-0.391) showed relatively high ICC values, indicating that a substantial portion of the variance in these sleep parameters is attributable to between-participant variability.

The high ICC values and statistically significant random slopes indicate substantial individual variability in the relationships between behavioral and psychological subsymptoms of dementia and sleep patterns. Consequently, it was necessary to use a panel vector autoregressive model to analyze these within-individual dynamic, reciprocal relationships over time.

Panel Vector Autoregressive Model Estimation Results for Behavioral and Psychological Subsymptoms of Dementia and Sleep Patterns

Tables 4-6 show the results of the panel vector autoregressive model estimation for behavioral and psychological subsymptoms of dementia as response variables. The previous day's sleep patterns were temporally associated with higher levels of some subsymptoms: irritability was temporally associated with higher total sleep time (*P*=.02), a greater number of awakenings (*P*=.03), and longer mean awakening length (*P*=.04) on the previous day, and appetite or eating disorders were temporally associated with a higher number of awakenings on the previous day (*P*=.04).

XSL•FO
RenderX

Table 4. Panel vector autoregressive model estimation results for the behavioral and psychological subsymptoms (delusion, hallucination, agitation or aggression, and depression) of dementia by sleep pattern.

Response and impulse variables	β (SE)	*P* value[a]
Delusion		
TST[b]	0.00026 (0.00016)	.10
WASO[c]	0.00090 (0.00058)	.12
NoA[d]	0.00249 (0.00167)	.13
MAL[e]	0.01977 (0.01654)	.23
Hallucination		
TST	0.00004 (0.00013)	.78
WASO	0.00048 (0.00058)	.40
NoA	0.00197 (0.00178)	.27
MAL	0.01163 (0.01269)	.36
Agitation or aggression		
TST	0.00003 (0.00016)	.87
WASO	0.00018 (0.00055)	.75
NoA	0.00061 (0.00177)	.73
MAL	–0.00100 (0.01699)	.95
Depression		
TST	–0.00016 (0.00018)	.37
WASO	–0.00033 (0.00070)	.64
NoA	–0.00113 (0.00208)	.59
MAL	0.00441 (0.01972)	.82

[a]*P* value represent the results of the Granger causality test within the panel vector autoregressive model.

[b]TST: total sleep time.

[c]WASO: wake after sleep onset.

[d]NoA: number of awakenings.

[e]MAL: mean awakening length.

XSL•FO
RenderX

Table 5. Panel vector autoregressive model estimation results for the behavioral and psychological subsymptoms (anxiety, euphoria or elation, apathy, and disinhibition) of dementia by sleep pattern.

Response and impulse variables	β (SE)	*P* value[a]
Anxiety		
TST[b]	–0.00018 (0.00018)	.33
WASO[c]	0.00015 (0.00057)	.79
NoA[d]	0.00022 (0.00162)	.89
MAL[e]	0.01034 (0.01991)	.60
Euphoria or elation		
TST	0.00014 (0.00013)	.28
WASO	0.00059 (0.00047)	.20
NoA	0.00207 (0.00138)	.13
MAL	0.02514 (0.01582)	.11
Apathy		
TST	0.00020 (0.00018)	.25
WASO	0.00058 (0.00074)	.43
NoA	0.00288 (0.00221)	.19
MAL	0.00412 (0.01818)	.82
Disinhibition		
TST	–0.00013 (0.00009)	.15
WASO	–0.00038 (0.00040)	.34
NoA	0.87058 (2.85687)	.76
MAL	–0.00547 (0.00925)	.55

[a]*P* value represent the results of the Granger causality test within the panel vector autoregressive model.

[b]TST: total sleep time.

[c]WASO: wake after sleep onset.

[d]NoA: number of awakenings.

[e]MAL: mean awakening length.

XSL•FO
RenderX

Table 6. Panel vector autoregressive model estimation results for the behavioral and psychological subsymptoms of dementia (irritability, aberrant motor behavior, sleep and nighttime behavior, and appetite or eating disorders) by sleep pattern.

Response and impulse variables	β (SE)	*P* value[a]
Irritability		
TST[b]	0.00048 (0.00021)	.02
WASO[c]	0.00126 (0.00065)	.05
NoA[d]	0.00428 (0.00199)	.03
MAL[e]	0.03738 (0.01821)	.04
Aberrant motor behavior		
TST	0.00004 (0.00009)	.61
WASO	0.00001 (0.00038)	.97
NoA	0.00112 (0.0011232)	.94
MAL	0.00049 (0.01183)	.97
Sleep and nighttime behavior		
TST	0.00008 (0.00018)	.66
WASO	0.00130 (0.00078)	.09
NoA	0.00347 (0.00222)	.12
MAL	0.03416 (0.01947)	.08
Appetite or eating disorders		
TST	0.00020 (0.00013)	.13
WASO	0.00091 (0.00048)	.06
NoA	0.00292 (0.00141)	.04
MAL	0.02892 (0.01586)	.07

[a]*P* value represent the results of the Granger causality test within the panel vector autoregressive model.

[b]TST: total sleep time.

[c]WASO: wake after sleep onset.

[d]NoA: number of awakenings.

[e]MAL: mean awakening length.

Table 7 shows the results of the panel vector autoregressive model estimation for sleep patterns as the response variable. Some of the previous day's behavioral and psychological subsymptoms of dementia were temporally associated with changes in sleep patterns on the following day: the previous day's delusion (*P*<.001), hallucination (*P*=.04), agitation or aggression (*P*=.02), depression (*P*=.002), anxiety (*P*=.02), and apathy (*P*=.04) were temporally associated with higher total sleep time; the previous day's delusion (*P*=.01), agitation or aggression (*P*=.02), and apathy (*P*=.02) were temporally associated with greater wake after sleep onset; and the previous day's delusions (*P*=.006), apathy (*P*=.02), and irritability (*P*=.007) were temporally associated with a higher number of awakenings on the following day.

XSL•FO
RenderX

Table 7. Panel vector autoregressive model estimation results for sleep pattern by the behavioral and psychological subsymptoms of dementia.

Response and impulse variables	β (SE)	*P* value[a]
Total sleep time		
Delusion	207.12760 (59.37694)	<.001
Hallucination	146.51250 (69.72166)	.04
Agitation or aggression	148.64660 (61.71230)	.02
Depression	186.97850 (59.76790)	.002
Anxiety	166.45380 (71.14920)	.02
Euphoria or elation	59.388490 (59.88787)	.32
Apathy	128.96060 (61.87693)	.04
Disinhibition	–44.88155 (149.0512)	.76
Irritability	82.21647 (53.16712)	.12
Aberrant motor behavior	141.56140 (82.73073)	.09
Sleep and nighttime behavior	113.56550 (59.04195)	.05
Appetite or eating disorders	11.843420 (60.79683)	.85
Wake after sleep onset		
Delusion	8.55337 (3.44196)	.01
Hallucination	7.06856 (4.91151)	.15
Agitation or aggression	9.17265 (4.03159)	.02
Depression	1.671489 (3.45359)	.63
Anxiety	–1.32085 (4.10093)	.75
Euphoria or elation	3.72705 (4.00864)	.35
Apathy	10.78888 (4.53788)	.02
Disinhibition	5.65713 (8.89465)	.80
Irritability	5.657127 (3.06763)	.07
Aberrant motor behavior	7.76621 (6.07391)	.20
Sleep and nighttime behavior	6.09917 (3.51403)	.08
Appetite or eating disorders	5.15131 (3.93965)	.19
Number of awakenings		
Delusion	3.79514 (1.39367)	.006
Hallucination	2.83285 (1.48000)	.14
Agitation or aggression	2.75541 (1.43263)	.05
Depression	0.90918 (1.30923)	.49
Anxiety	–0.60624 (1.52338)	.69
Euphoria or elation	1.73797 (1.53151)	.26
Apathy	4.07411 (1.67891)	.02
Disinhibition	0.87058 (2.85687)	.76
Irritability	3.31969 (1.23878)	.007
Aberrant motor behavior	3.79111 (2.51553)	.13
Sleep and nighttime behavior	2.28761 (1.32894)	.09
Appetite or eating disorders	2.23741 (1.51776)	.14
Mean awakening length		
Delusion	0.20611 (0.18999)	.28
Hallucination	–0.08000 (0.22148)	.72

XSL•FO
RenderX

Response and impulse variables	β (SE)	*P* value[a]
Agitation or aggression	0.12385 (0.21356)	.56
Depression	0.09569 (0.15077)	.53
Anxiety	0.04655 (0.24929)	.85
Euphoria or elation	–0.02885 (0.19814)	.88
Apathy	0.13391 (0.18566)	.47
Disinhibition	–0.25250 (0.45123)	.58
Irritability	–0.12203 (0.16155)	.45
Aberrant motor behavior	0.03143 (0.29223)	.91
Sleep and nighttime behavior	–0.04485 (0.20128)	.82
Appetite or eating disorders	–0.21659 (0.20477)	.29

[a]*P* value represent the results of the Granger causality test within the panel vector autoregressive model.

Granger causality tests were subsequently conducted for variable pairs that demonstrated statistically significant temporal associations in the panel vector autoregressive models (Tables 5-7). The results showed unidirectional Granger causality, consistent with the panel vector autoregressive model estimation. Bidirectional Granger causality was observed only between irritability and the number of awakenings.

Discussion

Principal Findings

Using panel data and a panel vector autoregressive model, this study longitudinally analyzed the dynamic relationships between behavioral and psychological symptoms of dementia and sleep patterns among older adults with dementia. Regarding the main findings, first, the number of sleep awakenings on the previous day was temporally associated with higher levels of irritability and appetite or eating disorders on the following day. Second, some of the previous day's behavioral and psychological subsymptoms of dementia were temporally associated with subsequent changes in sleep patterns (eg, an increase in total sleep time, the number of awakenings, and mean awakening length).

A linear mixed model analysis was conducted to contextualize these findings, revealing substantial between-participant variability in both behavioral and psychological symptoms of dementia and sleep patterns. Although the fixed effects were not statistically significant, the random slopes were statistically significant, and there were moderate to high ICCs, indicating that individual differences contributed meaningfully to variability in both variables; that is, the relationships between behavioral and psychological symptoms of dementia and sleep patterns were not uniform across individuals and may involve complex, time-dependent interactions. Accordingly, a panel vector autoregressive model was implemented, revealing statistically significant temporal associations in which sleep patterns preceded changes in irritability and appetite or eating disorders. This indicates that interventions to improve sleep quality should be implemented in the early stages of sleep disturbance to alleviate the behavioral and psychological symptoms of dementia that may stem from poor sleep quality. Importantly, our evidence implies that effectively preventing the occurrence of irritability and appetite or eating disorders in older adults with dementia requires careful and consistent monitoring and management of increases in total sleep time, the number of awakenings, and mean awakening length. Previous studies have reported that 22.6% of people with dementia have excessive daytime sleepiness and that sleep disturbances are associated with the onset of behavioral and psychological symptoms of dementia [18,19]. Furthermore, to avoid excessive total sleep time, people with dementia should engage in sufficient physical and social activity and limit daytime napping [25,26].

In particular, the number of awakenings during sleep and their duration demonstrated significant temporal associations with subsequent behavioral and psychological symptoms. These findings support other studies that have found a relationship between excessive awakenings and the occurrence of behavioral and psychological symptoms of dementia [27]. Education on sleep hygiene practices, such as maintaining consistent sleep-wake routines and optimizing the sleep environment [25], and combined nonpharmacological interventions, including physical or social activities and caregiver support interventions [28], are reportedly effective in reducing awakenings during sleep and alleviating the behavioral and psychological symptoms of dementia.

Conversely, behavioral and psychological symptoms of dementia onset were temporally associated with subsequent sleep patterns. In particular, irritability had bidirectional Granger causality with the number of awakenings, implying that poor sleep patterns may contribute to the occurrence of symptoms the next day and that the occurrence of irritability may increase the number of awakenings, their duration, and excessive total sleep time. Previous studies provide mechanistic plausibility for this reciprocal relationship, showing that increased nocturnal awakenings and sleep fragmentation can impair emotion regulation through the hyperactivation of the amygdala [29]. Notably, this bidirectional effect was specific to irritability, likely because irritability is more directly coupled with the "arousal system" than internalized symptoms such as depression. Unlike cognitive-heavy symptoms, nocturnal awakenings trigger immediate amygdala hyperreactivity and a failure in top-down

XSL•FO
RenderX

prefrontal control [30], manifesting as acute reactive irritability. This heightened physiological arousal further disrupts sleep maintenance, reinforcing a self-perpetuating reciprocal loop [31]. Clinically, these findings suggest that the relationship between behavioral and psychological symptoms of dementia and sleep can be characterized as a vicious circle, where nocturnal awakenings serve as an early warning signal of worsening irritability.

Consistent with this interpretation, previous studies in patients with Alzheimer disease have shown that on nights when patients awakened their caregivers, they exhibited more behavioral and psychological symptoms of dementia the following morning. Additionally, patients who had difficulty falling asleep displayed more behavioral and psychological symptoms of dementia at night [27]. Similarly, Kuzmik et al [12] demonstrated that higher behavioral and psychological symptoms of dementia were associated with sleep disturbances, especially prolonged sleep latency. Thus, the relationship between the behavioral and psychological symptoms of dementia and sleep patterns seems to be complex and mutually reinforcing, necessitating the integration of sleep management into dementia symptom care programs.

The behavioral and psychological symptoms of dementia represent a primary burden for caregivers of older adults with dementia [7], and sleep disturbances affect the quality of life of both the primary caregiver and the older adult with dementia, contributing to cognitive decline and difficulties with community-based care [32]. Therefore, interventions to decrease these symptoms should be initiated early in the course of the symptoms, and interventions to improve sleep quality should be promoted simultaneously. Pharmacological interventions to address the behavioral and psychological symptoms of dementia are less accessible owing to the prescription process and may have side effects; therefore, nonpharmacologic interventions are recommended, such as reminiscence and music therapy, as they can be easily implemented by caregivers [33,34]. In particular, recent reports suggest that nonpharmacological interventions are more effective when customized to reflect the characteristics and preferences of older adults with dementia, compared with standardized interventions that are applied uniformly across all [35]. This highlights the need to develop and test nonpharmacological interventions aimed at improving sleep quality and alleviating behavioral and psychological symptoms of dementia in older adults with dementia.

This study has important implications for the field. To the best of our knowledge, this is the first study to explore the relationship between the behavioral and psychological symptoms of dementia and sleep patterns through panel vector autoregressive analysis. Sleep patterns were also observed using actigraphy technology, not subjective surveys.

Limitations

This study has some limitations. First, behavioral and psychological symptoms of dementia were assessed using caregiver-recorded daily diaries, which may be vulnerable to recall bias owing to retrospective reporting. Future studies may consider ecological momentary assessment approaches using mobile apps to enable real-time symptom recording and reduce recall bias. Second, data on medications used by older adults to manage behavioral and psychological symptoms of dementia and sleep, as well as symptom-driven medication changes over time, were not available. Although the panel vector autoregressive model addresses time-invariant individual heterogeneity by modeling within-person temporal dynamics, it cannot fully adjust for unmeasured time-varying confounders, such as changes in medication use or comorbidity severity over time. Consequently, such unobserved temporal changes may have influenced both behavioral and psychological symptoms of dementia and sleep patterns and should be considered when interpreting the findings.

Third, the relatively short 14-day observation period limited the evaluation of longer lag structures and seasonal effects on sleep patterns. Future studies using longer-term longitudinal data are needed to examine extended lag dynamics and potential seasonal influences. Furthermore, the lack of observed bidirectionality in behavioral and psychological subsymptoms of dementia, other than irritability, may be due to their longer-term trajectories, which might necessitate more extended observation and larger cohorts to detect subtle daily fluctuations. Some statistically significant associations were characterized by small effect sizes, likely reflecting the chronic and progressive nature of dementia, in which abrupt day-to-day changes in behavioral and psychological symptoms of dementia and sleep are uncommon. In this context, the short observation period may have captured only subtle temporal fluctuations, and future studies with longer follow-up periods are needed to examine whether such small changes accumulate over time and become clinically meaningful.

Finally, this study relied on convenience sampling and included a relatively small sample, which may limit the generalizability of the findings. Moreover, although the overall participation rate (154/175, 88%) was relatively high for studies involving older adults with dementia, participant attrition occurred during the study period owing to reasons such as refusal to continue using the actigraphy device or hospitalization. Given that these reasons for dropout may have been more common among participants with worsening symptoms, the final analytic sample may underrepresent individuals with more severe conditions. This potential selection bias, combined with the modest sample size and wide 95% CIs for several estimates, may have reduced statistical power and increased the risk of both type I and type II errors.

Conclusions

By integrating a linear mixed model and a panel vector autoregressive model, this study provides comprehensive evidence that accounts for individual variability and examines the bidirectional relationships between behavioral and psychological symptoms of dementia and sleep disturbances. This approach enhances the robustness of our Granger causal inferences. Panel vector autoregressive analysis demonstrated that the variables had an intricate and bidirectional relationship. This suggests that addressing sleep disturbances may mitigate certain behavioral and psychological symptoms of dementia, and vice versa.

XSL•FO
RenderX

The findings imply the importance of holistic approaches to dementia care management. One potential pathway to mitigate symptoms involves detecting behavioral and psychological symptoms of dementia onset and implementing timely interventions to improve both sleep and these symptoms of dementia. Given that these findings highlight the impact of sleep patterns on behavioral and psychological symptoms of dementia, there is a critical need for interventions aimed at improving sleep quality in older adults with dementia. However, the availability of effective nonpharmacologic interventions to enhance sleep quality in older adults with dementia remains limited, highlighting the need for future research to develop and validate such interventions.

Funding

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (grant RS-2020-NR049581).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables contain the lag order of the panel vector autoregression models, participants' baseline characteristics, and linear mixed model results examining the relationships between sleep patterns and behavioral and psychological symptoms of dementia.
[DOCX File , 54 KB - aging_v9i1e80422_app1.docx]

References

1. Cloak N, Schoo C, Al Khalili Y. Behavioral and psychological symptoms in dementia. In: StatPearls. Treasure Island, FL: StatPearls Publishing; 2024.
2. IPA Complete Guides to Behavioral and Psychological Symptoms of Dementia (BPSD). Northfield, IL: International Psychogeriatric Association; 2012.
3. Kwon CY, Lee B. Prevalence of behavioral and psychological symptoms of dementia in community-dwelling dementia patients: a systematic review. Front Psychiatry 2021 Oct 21;12:741059 [FREE Full text] [doi: 10.3389/fpsyt.2021.741059] [Medline: 34744832]
4. Laganà V, Bruno F, Altomari N, Bruni G, Smirne N, Curcio S, et al. Neuropsychiatric or behavioral and psychological symptoms of dementia (BPSD): focus on prevalence and natural history in Alzheimer's disease and frontotemporal dementia. Front Neurol 2022 Jun 24;13:832199 [FREE Full text] [doi: 10.3389/fneur.2022.832199] [Medline: 35812082]
5. Bränsvik V, Granvik E, Minthon L, Nordström P, Nägga K. Mortality in patients with behavioural and psychological symptoms of dementia: a registry-based study. Aging Ment Health 2021 Jun 18;25(6):1101-1109 [FREE Full text] [doi: 10.1080/13607863.2020.1727848] [Medline: 32067466]
6. Feast A, Moniz-Cook E, Stoner C, Charlesworth G, Orrell M. A systematic review of the relationship between behavioral and psychological symptoms (BPSD) and caregiver well-being. Int Psychogeriatr 2016 Nov;28(11):1761-1774 [FREE Full text] [doi: 10.1017/S1041610216000922] [Medline: 27345942]
7. Mukherjee A, Biswas A, Roy A, Biswas S, Gangopadhyay G, Das S. Behavioural and psychological symptoms of dementia: correlates and impact on caregiver distress. Dement Geriatr Cogn Dis Extra 2017 Nov 1;7(3):354-365 [FREE Full text] [doi: 10.1159/000481568] [Medline: 29282408]
8. Cohen-Mansfield J. Behavioral and psychological symptoms of dementia. In: Lichtenberg PA, Mast BT, Carpenter BD, Loebach Wetherell J, editors. APA Handbook of Clinical Geropsychology. Washington, DC: American Psychological Association; 2015:271-317.
9. Kovach CR, Noonan PE, Schlidt AM, Wells T. A model of consequences of need-driven, dementia-compromised behavior. J Nurs Scholarsh 2005 Jun;37(2):134-40; discussion 140. [doi: 10.1111/j.1547-5069.2005.00025_1.x] [Medline: 15960057]
10. Bombois S, Derambure P, Pasquier F, Monaca C. Sleep disorders in aging and dementia. J Nutr Health Aging 2010 Mar;14(3):212-217 [FREE Full text] [doi: 10.1007/s12603-010-0052-7] [Medline: 20191256]
11. Okuda S, Tetsuka J, Takahashi K, Toda Y, Kubo T, Tokita S. Association between sleep disturbance in Alzheimer's disease patients and burden on and health status of their caregivers. J Neurol 2019 Jun 9;266(6):1490-1500 [FREE Full text] [doi: 10.1007/s00415-019-09286-0] [Medline: 30968171]
12. Kuzmik A, Boltz M, BeLue R, Galvin JE, Arendacs R, Resnick B. Factors associated with sleep quality in hospitalized persons with dementia. Alzheimer Dis Assoc Disord 2022;36(3):253-258. [doi: 10.1097/WAD.0000000000000502] [Medline: 36001764]

XSL•FO
RenderX

13. Shin HY, Han HJ, Shin DJ, Park HM, Lee YB, Park KH. Sleep problems associated with behavioral and psychological symptoms as well as cognitive functions in Alzheimer's disease. J Clin Neurol 2014 Jul;10(3):203-209 [FREE Full text] [doi: 10.3988/jcn.2014.10.3.203] [Medline: 25045372]
14. Zhou G, Liu S, Yu X, Zhao X, Ma L, Shan P. High prevalence of sleep disorders and behavioral and psychological symptoms of dementia in late-onset Alzheimer disease: a study in Eastern China. Medicine (Baltimore) 2019 Dec;98(50):e18405 [FREE Full text] [doi: 10.1097/MD.0000000000018405] [Medline: 31852160]
15. García-Alberca JM, Lara JP, Cruz B, Garrido V, Gris E, Barbancho M. Sleep disturbances in Alzheimer's disease are associated with neuropsychiatric symptoms and antidementia treatment. J Nerv Ment Dis 2013 Mar;201(3):251-257. [doi: 10.1097/NMD.0b013e3182848d04] [Medline: 23407210]
16. Balouch S, Dijk DA, Rusted J, Skene SS, Tabet N, Dijk DJ. Night-to-night variation in sleep associates with day-to-day variation in vigilance, cognition, memory, and behavioral problems in Alzheimer's disease. Alzheimers Dement (Amst) 2022 May 16;14(1):e12303 [FREE Full text] [doi: 10.1002/dad2.12303] [Medline: 35603140]
17. Kabeshita Y, Adachi H, Matsushita M, Kanemoto H, Sato S, Suzuki Y, et al. Sleep disturbances are key symptoms of very early stage Alzheimer disease with behavioral and psychological symptoms: a Japan multi-center cross-sectional study (J-BIRD). Int J Geriatr Psychiatry 2017 Feb 21;32(2):222-230. [doi: 10.1002/gps.4470] [Medline: 27001907]
18. Petrovsky DV, McPhillips MV, Li J, Brody A, Caffeé L, Hodgson NA. Sleep disruption and quality of life in persons with dementia: a state-of-the-art review. Geriatr Nurs 2018 Nov;39(6):640-645 [FREE Full text] [doi: 10.1016/j.gerinurse.2018.04.014] [Medline: 29803606]
19. Rongve A, Boeve BF, Aarsland D. Frequency and correlates of caregiver-reported sleep disturbances in a sample of persons with early dementia. J Am Geriatr Soc 2010 Mar;58(3):480-486. [doi: 10.1111/j.1532-5415.2010.02733.x] [Medline: 20398116]
20. Cho E, Kim S, Hwang S, Kwon E, Heo SJ, Lee JH, et al. Factors associated with behavioral and psychological symptoms of dementia: prospective observational study using actigraphy. J Med Internet Res 2021 Oct 29;23(10):e29001 [FREE Full text] [doi: 10.2196/29001] [Medline: 34714244]
21. Canova F, Ciccarelli M. Panel vector autoregressive models: a survey. European Central Bank. 2013. URL: https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1507.pdf [accessed 2026-03-04]
22. Cummings JL. The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. Neurology 1997 May;48(5 Suppl 6):S10-S16. [doi: 10.1212/wnl.48.5_suppl_6.10s] [Medline: 9153155]
23. Camargos EF, Louzada FM, Nóbrega OT. Wrist actigraphy for measuring sleep in intervention studies with Alzheimer's disease patients: application, usefulness, and challenges. Sleep Med Rev 2013 Dec;17(6):475-488. [doi: 10.1016/j.smrv.2013.01.006] [Medline: 23669093]
24. Andrews DW, Lu B. Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models. J Econom 2001 Mar;101(1):123-164. [doi: 10.1016/s0304-4076(00)00077-4]
25. Ooms S, Ju YE. Treatment of sleep disorders in dementia. Curr Treat Options Neurol 2016 Sep;18(9):40 [FREE Full text] [doi: 10.1007/s11940-016-0424-3] [Medline: 27476067]
26. Woods DL, Kim H, Yefimova M. To nap or not to nap: excessive daytime napping is associated with elevated evening cortisol in nursing home residents with dementia. Biol Res Nurs 2013 Apr 13;15(2):185-190. [doi: 10.1177/1099800411420861] [Medline: 21998447]
27. Connelly C, Kim K, Liu Y, Zarit SH. Temporal patterns of behavioral and psychological symptoms of dementia and caregiver distress: associated daily and individual factors. J Appl Gerontol 2024 Jun;43(6):786-796. [doi: 10.1177/07334648231216382] [Medline: 37987697]
28. Wilfling D, Calo S, Dichter MN, Meyer G, Möhler R, Köpke S. Non-pharmacological interventions for sleep disturbances in people with dementia. Cochrane Database Syst Rev 2023 Jan 03;1(1):CD011881 [FREE Full text] [doi: 10.1002/14651858.CD011881.pub2] [Medline: 36594432]
29. Walker MP. The role of sleep in cognition and emotion. Ann N Y Acad Sci 2009 Mar 25;1156(1):168-197. [doi: 10.1111/j.1749-6632.2009.04416.x] [Medline: 19338508]
30. Yoo SS, Gujar N, Hu P, Jolesz FA, Walker MP. The human emotional brain without sleep--a prefrontal amygdala disconnect. Curr Biol 2007 Oct 23;17(20):R877-R878 [FREE Full text] [doi: 10.1016/j.cub.2007.08.007] [Medline: 17956744]
31. Vandekerckhove M, Wang YL. Emotion, emotion regulation and sleep: an intimate relationship. AIMS Neurosci 2018;5(1):1-17 [FREE Full text] [doi: 10.3934/Neuroscience.2018.1.1] [Medline: 32341948]
32. Lee DR, Thomas AJ. Sleep in dementia and caregiving--assessment and treatment implications: a review. Int Psychogeriatr 2011 Mar;23(2):190-201 [FREE Full text] [doi: 10.1017/S1041610210001894] [Medline: 20946702]
33. Cho E, Kim MJ, Yang M, Jang J, Cho J, Lee JY. Symptom-specific non-pharmacological interventions for behavioural and psychological symptoms of dementia: protocol of an umbrella review of systematic reviews of randomised controlled trials. BMJ Open 2023 Feb 15;13(2):e070317 [FREE Full text] [doi: 10.1136/bmjopen-2022-070317] [Medline: 36792326]
34. Overshott R, Burns A. Non-pharmacological treatment of severe dementia: an overview. In: Burns A, Winblad B, editors. Severe Dementia. Hoboken, NJ: John Wiley & Sons; 2006.

XSL•FO
RenderX

35. Cho E, Yang M, Jang J, Cho J, Kang B, Jang Y, et al. A mobile app to predict and manage behavioral and psychological symptoms of dementia: development, usability, and users' acceptability. Digit Health 2024 Apr 30;10:20552076241249277 [FREE Full text] [doi: 10.1177/20552076241249277] [Medline: 38698833]

Abbreviations

ICC: intraclass correlation coefficient

Edited by R Lin; submitted 10.Jul.2025; peer-reviewed by Z Ye, M Yildiz; comments to author 26.Nov.2025; revised version received 26.Jan.2026; accepted 26.Feb.2026; published 23.Mar.2026.

Please cite as:
Cho E, Hwang S, Yang M, Kim E, Cho J, Park C
Dynamic Relationship Between Sleep Patterns and Behavioral and Psychological Symptoms of Dementia: Longitudinal Observational Study
JMIR Aging 2026;9:e80422
URL: https://aging.jmir.org/2026/1/e80422
doi:10.2196/80422
PMID:

©Eunhee Cho, Sinwoo Hwang, Minhee Yang, Eunkyo Kim, Jungwon Cho, Chang Park. Originally published in JMIR Aging (https://aging.jmir.org), 23.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Experiences of Older Adults and Caregivers With Home Telemonitoring for Heart Failure in Canada: Qualitative Study

Guy Paré[1*], PhD; Marie-Pierre Moreault[1*], MS; Philippe Voyer[2*], PhD; Alexandre Castonguay[3*], PhD; Marie-Soleil Hardy[2*], PhD; Mickael Ringeval[4*], PhD

[1]Department of Information Technologies, HEC Montréal, Montreal, QC, Canada

[2]Faculty of Nursing, Université Laval, Québec, QC, Canada

[3]Faculty of Nursing, Université de Montréal, Montreal, QC, Canada

[4]Department of Computer Science, Bentley University, Waltham, MA, United States

[*]all authors contributed equally

Corresponding Author:
Guy Paré, PhD
Department of Information Technologies
HEC Montréal
3000 Cote-Ste-Catherine Road
Montreal, QC, H3T2A7
Canada
Phone: 1 5143406812
Email: guy.pare@hec.ca

Abstract

Background: Home telemonitoring programs are increasingly used to support older adults living with chronic conditions such as heart failure (HF). While these interventions show promise for improving health outcomes and reducing care burden, their effectiveness depends largely on how patients and caregivers integrate digital technologies into everyday life and care relationships. However, relatively few studies have examined these experiences using conceptual frameworks that capture both functional and relational dimensions of care.

Objective: This study aimed to explore the experiences of older adults and their informal caregivers participating in a home telemonitoring program for HF. Drawing on the Person-Based Approach and the Person-Centered Practice frameworks, we examined how participants engaged with both the technofunctional and relational aspects of the intervention.

Methods: We conducted a qualitative study involving 34 patients, 28 informal caregivers, and 20 nurses across 3 primary care organizations in Quebec, Canada. The 6-month intervention included 4 connected devices used by patients (smartwatch, Bluetooth-enabled scale, voice-activated tablet, and a smart pill dispenser [xPill; Domedic]) and a mobile app for caregivers, complemented by remote nursing follow-up. Nurses reviewed patient data through a clinical dashboard at least once daily during weekday daytime shifts. Data were collected through semistructured interviews and field notes and analyzed using directed content analysis.

Results: Participants' experiences revealed both enabling and constraining factors across 2 key dimensions. Technofunctional engagement was shaped by digital literacy, emotional responses to the technology, alignment with daily routines, and access to technical or caregiver support. Relational aspects of care were influenced by perceived professional presence, opportunities for communication and shared decision-making, and the degree of emotional reassurance provided by remote monitoring. While many participants reported increased confidence and a sense of being supported, others experienced frustration, fatigue, or disengagement when the system disrupted routines or when feedback from clinicians was perceived as limited.

Conclusions: Engagement with home telemonitoring technologies among older adults depends not only on usability but also on the relational context in which these technologies are embedded. Combining technofunctional and relational perspectives provides a more comprehensive understanding of how telemonitoring interventions are experienced and highlights the importance of personalized support, reliable technology, and sustained clinical engagement to promote meaningful adoption.

(JMIR Aging 2026;9:e79797) doi:10.2196/79797

XSL•FO
RenderX

KEYWORDS

digital health; heart failure; home telemonitoring; informal caregivers; older adults; patient experience; person-centered care; qualitative research

Introduction

Overview

Heart failure (HF) remains one of the leading causes of hospitalization and mortality among older adults globally [1,2]. As health care systems grapple with the dual pressures of aging populations and rising rates of chronic illness, the demand for scalable, patient-centered models of care has become increasingly urgent. Among the most promising innovations in this space are home telemonitoring programs, which aim to support early detection of clinical deterioration, foster self-management, and reduce reliance on in-person visits or rehospitalization [3,4].

Recent evidence supports the clinical value of these interventions. A systematic review and meta-analysis of mobile health technologies for home-based cardiac rehabilitation demonstrated significant improvements in patient outcomes and engagement, suggesting the potential of these tools to enhance chronic disease management at scale [5]. However, findings across studies remain mixed. The effectiveness of home telemonitoring programs often hinges on contextual factors such as system design, user characteristics, and the broader care environment [6]. This variability underscores the importance of understanding how these technologies are experienced and navigated by those who use them in their daily lives.

Older adults, in particular, represent a user group with unique needs and vulnerabilities when it comes to digital health technologies [7,8]. Differences in digital literacy, cognitive capacity, and physical function can pose challenges to sustained engagement. Emotional responses, such as anxiety, fatigue, or frustration, may also shape how older adults relate to technology over time. Moreover, their experience is often mediated by informal caregivers, whose involvement can critically influence the adoption, maintenance, and perceived usefulness of telemonitoring interventions. Yet, despite growing interest in user-centered evaluation, relatively few studies have provided in-depth, real-world accounts of how both patients and caregivers experience home telemonitoring programs.

This study seeks to address this gap by examining the lived experiences of older adults and their informal caregivers who participated in a home telemonitoring intervention [9] for HF across 3 primary care institutions in Quebec, Canada. In this study, we use the term home telemonitoring to refer to digital systems that collect and transmit patients' physiological and self-reported health data from the home environment to clinicians for remote monitoring and follow-up.

The intervention involved 4 interconnected digital devices used by patients: a connected smartwatch, a Bluetooth-enabled scale, a voice-activated virtual assistant tablet, and a smart pill dispenser (xPill; Domedic). In addition, informal caregivers had access to the Proximity mobile app (Virtuose Technologies), which allowed them to receive alerts, monitor patient activity, and communicate with both patients and nurses. These digital tools were complemented by remote nursing follow-up, during which nurses reviewed patient data through a clinical dashboard, responded to alerts, and communicated with patients and caregivers when necessary.

Our objective was to understand how participants perceived and interacted with these technologies, how they navigated their technical and relational demands, and how the system was integrated into the routines and realities of everyday life. To achieve this goal, we draw on 2 complementary conceptual frameworks, namely, the Person-Based Approach (PBA) developed by Yardley et al [10] and the Person-Centered Practice (PCP) framework by McCance and McCormack [11]. The PBA provides a structured lens for understanding users' cognitive, emotional, and behavioral engagement with digital health tools, emphasizing the importance of contextual relevance, emotional acceptability, and iterative user-centered adaptation. The PCP framework complements this by highlighting the relational and experiential dimensions of care, such as perceived professional presence, shared decision-making, and therapeutic alliance. By analyzing user experience through these dual lenses, this study offers a more nuanced understanding of the opportunities and limitations associated with home telemonitoring for older adults with HF. In doing so, it provides actionable insights for clinicians, designers, and managers aiming to improve the effectiveness, equity, and person-centeredness of these interventions. The following section introduces each framework in more detail and outlines its relevance to this study.

Conceptual Background

The PBA Framework

To explore the techno-functional experience, we adopt the PBA to digital health intervention development [10]. The framework emphasizes the centrality of users' lived experiences and psychosocial contexts in shaping how they interact with and respond to digital health tools. It is especially well suited to evaluating complex interventions such as multidevice telemonitoring programs, where sustained engagement is contingent not only on technical usability but also on the personal significance and practical fit of the technology in everyday life. Key elements of the PBA relevant to this study include the following: contextual tailoring, emotional and cognitive response to technology, perceived relevance and burden of technology use, and autonomy, control, and iterative engagement.

The PBA stresses the importance of designing or implementing interventions in ways that reflect the diversity of users' real-life circumstances. This includes recognizing variation in cognitive functioning, literacy levels, physical capabilities, caregiving arrangements, and daily schedules. A telemonitoring system that is too rigid or uniform in its expectations may be experienced as alienating or unmanageable by users whose life

XSL•FO
RenderX

contexts do not match the assumptions embedded in the system's design.

The PBA underscores that user experience includes both affective and cognitive reactions. Feelings of reassurance, anxiety, confidence, or fatigue can influence whether and how users persist with an intervention. In the context of older adults, such responses are closely tied to familiarity with digital devices, past experiences with health care, and the presence or absence of technical or social support.

Perceived relevance and burden of technology use refers to how meaningful and manageable users perceive the intervention to be. For an older adult managing a chronic condition like HF, a telemonitoring device must clearly align with their health goals and offer value in their day-to-day life. At the same time, if the tool is perceived as cognitively demanding, repetitive, or disruptive to routines, it may generate a burden that outweighs perceived benefits, leading to frustration, disengagement, or early withdrawal.

Central to the PBA is the idea that users should feel a sense of agency when interacting with the intervention. Technologies that allow for flexibility and customization can reinforce users' self-efficacy and support their identity as active participants in care. Conversely, systems perceived as overly prescriptive or difficult to modify may reassert feelings of passivity or dependence, particularly among older individuals already accustomed to clinician-led care.

In short, by applying the PBA, we aim to capture the nuanced ways in which older adults and their informal caregivers relate to the telemonitoring devices, how they negotiate its integration into their daily lives, and how usability, meaning, and burden are interpreted differently depending on their capabilities, routines, and support networks.

The PCP Framework

Originally developed to support quality improvement and professional reflection in nursing and allied health contexts, the PCP framework offers a comprehensive model of care that integrates professional ethos, care context, clinical process, and patient outcomes. It is especially relevant for understanding how care is experienced by patients in nontraditional or hybrid contexts, such as digitally mediated care environments, where therapeutic presence, trust, and relational continuity must be maintained at a distance. While the full framework includes 4 domains, namely, prerequisites, care environment, person-centered processes, and person-centered outcomes, this study focuses on the last 2, as these most directly pertain to the interpersonal and experiential dimensions of care in a home telemonitoring model.

Person-centered processes describe how health care professionals enact person-centered values in clinical practice. They represent the core relational and therapeutic mechanisms through which patients experience empathy, partnership, and a sense of agency in their care. One such process is sympathetic presence, which refers to the professional's ability to be emotionally attuned and fully present with the patient. This presence is conveyed through attentive listening, sensitive responses, and compassionate behavior. In remote care settings, where nonverbal cues are absent, sympathetic presence must often be expressed through tone of voice, verbal phrasing, and consistent follow-up. Another key process is authentic engagement, which entails a genuine effort to understand the patient as a person, beyond their clinical profile or transmitted data. This means adapting interactions to align with each patient's preferences, values, and life context, even when mediated by digital systems.

Shared decision-making is also central to PCP. It involves actively involving patients in decisions about their care by presenting options, seeking informed consent, and ensuring mutual understanding. In the context of home telemonitoring, this process often intersects with choices around the timing and content of reminders, the frequency of follow-ups, and the interpretation of patient-reported data. Finally, person-centered care involves working with patient beliefs and values, which requires acknowledging and respecting diverse worldviews, cultural backgrounds, and personal care goals, even when these may differ from standardized clinical routines or digital system defaults.

In a home telemonitoring context, these processes are especially vulnerable to being eroded or routinized, making their presence (or absence) highly meaningful to patients and caregivers.

Person-centered outcomes represent the experiential consequences of person-centered processes and serve as key indicators for evaluating the quality of care from the patient's perspective. One such outcome is the degree of involvement in care, which reflects whether patients feel that their voice is heard and that they have meaningful input into how care is delivered. In the context of home telemonitoring, this may involve being consulted about preferred methods and times for communication or having a say in how their personal health data are used. Another important outcome is the sense of well-being generated through care interactions. This encompasses emotional comfort, reduced anxiety, and a strengthened sense of self-efficacy. Expressions of reassurance, feelings of being understood, and the perception of being genuinely "cared for" at a distance are strong indicators of success in this domain. Finally, the development of a therapeutic alliance, a relationship characterized by trust, continuity, and mutual respect, remains fundamental even in digitally mediated care. Building this alliance requires proactive engagement from clinicians, timely responses to patient inputs, and communication practices that maintain the human element of care despite the technological interface.

By foregrounding these processes and outcomes, McCance and McCormack's [11] framework helps us evaluate not just whether care is delivered, but how it is lived and interpreted by patients and informal caregivers. It provides a valuable counterbalance to the task-oriented lens of technology implementation by highlighting the relational texture of remote care.

Methods

Study Design and Setting

We adopted a qualitative study design guided by user-centered evaluation principles to explore the experiences of patients,

XSL•FO
RenderX

informal caregivers, and nurses who participated in the home telemonitoring program. This approach emphasizes understanding how end users experience and engage with technology in the context of their everyday lives and care relationships [10,12]. It prioritizes the elicitation of first-hand accounts from those directly interacting with the intervention—namely, patients and caregivers—while also incorporating nurses' perspectives to triangulate findings, enrich interpretation, and provide insight into how clinical support shaped user experiences.

Participants were recruited from 3 health care organizations in Quebec: Centre intégré de santé et de services sociaux (CISSS) Chaudière-Appalaches, CISSS Gaspésie, and CISSS Outaouais. These sites were selected by the project sponsor, the Quebec Ministry of Health and Social Services, to reflect the diversity and typical characteristics of primary care organizations across the province. Importantly, from the outset of the project, a formal governance structure was established to ensure coordination and collaboration among the 3 participating sites. Clinical managers responsible for the project in each organization worked closely together throughout its implementation, deployment, and follow-up. As such, the study was not designed to support cross-site comparisons. There was no a priori reason to expect meaningful differences in participant experiences between locations, and our analysis confirmed this assumption. Indeed, patients, informal caregivers, and nurses across sites reported broadly similar perceptions in relation to the 2 central dimensions of analysis, technofunctional use and relational care. Accordingly, findings are presented in an aggregated form.

Intervention

The home telemonitoring intervention consisted of 4 interconnected digital devices installed in the residences of older adults diagnosed with HF. All participating patients were supplied with the same digital devices. The goal was to support daily self-monitoring of symptoms, vital signs, and medication adherence, while enabling remote follow-up by nurses. Each device was preconfigured to collect specific clinical data at defined frequencies, as follows:

- Patients interacted daily with a voice-activated tablet (Samsung [13]) that administered a structured symptom questionnaire (anamnesis) focused on HF warning signs and related physiological indicators. Items included: breathing difficulty, cough, dizziness, sleep disturbances, fatigue, retrosternal pain, digestion issues, urinary and bowel elimination, and disease-related anxiety. The system also issued prompts and reminders for weight measurement and medication intake.
- A smartwatch (Biobeat [14]) was used to continuously collect biometric data, including heart rate, blood pressure, and oxygen saturation, provided the patient wore it day and night. Data were passively captured and transmitted to the nurse monitoring platform.
- Patients were asked to weigh themselves using a Bluetooth-enabled weight scale (A&D Medical [15]) daily upon waking. No specific instructions were provided regarding attire during weighing (eg, nightwear vs daytime clothing), as the objective was to integrate the measurement into patients' usual daily routines. The digital scale automatically transmitted weight data, which served as a proxy for fluid retention and clinical decompensation risk.
- A smart pill dispenser (xPill [16]) was used to issue audio-visual cues to prompt medication intake and record the date and time of each compartment opening to track adherence.

Each morning, patients received an audible alert from the voice-activated tablet signaling that it was time to complete a symptom questionnaire. They also had the option to initiate the questionnaire up to 1 hour before the scheduled alert. The questionnaire could be completed either by selecting responses directly on the tablet or by answering verbally. When responding orally, users were able to elaborate on their answers. In addition to completing the questionnaire, patients were expected to weigh themselves daily using the connected scale. The smart xPill automatically logged medication intake and allowed users to view as-needed prescriptions and detailed information about their medication regimen. The purpose of the xPill was to support medication adherence among patients and to enable nurses to monitor compliance remotely. Communication between the user and the nurse could be conducted through written messages or video calls initiated via the tablet, allowing for responsive and flexible clinical follow-up.

Together, these devices offered a comprehensive view of the patient's evolving health status. Data were integrated into a secure clinical dashboard monitored by nurses, who could initiate follow-up interventions when anomalies or gaps in reporting were detected. Participating nurses were required to review their assigned patients' data at least once daily during weekday daytime shifts. Although data were collected continuously (24 hours a day, 7 days a week), active monitoring occurred only during these daytime hours (8 AM to 4 PM), Monday through Friday. Outside of these hours, including evenings and weekends, patients experiencing health concerns were instructed to contact Quebec's provincial health hotline (Info-Santé 811) or seek care at the nearest emergency department. Data collected overnight during the week were reviewed by nurses the following morning, while data from the weekend were examined on Monday morning.

The home telemonitoring program was designed as a 6-month intervention for participating patients. Patients who remained enrolled for the full duration used the monitoring devices continuously during this period. However, participation length varied for some individuals because several patients withdrew from the program after deployment of the devices. As reported in the Results section, the duration of participation among those who withdrew ranged from less than 1 week to approximately 7 weeks.

Informal caregivers were also integrated into the intervention through the Proximity mobile app. The app allowed them to view automated alerts designed to support the patient's adherence to telemonitoring routines (eg, "your relative has not weighed themselves for 4 days," or "the scale's battery needs charging") and to stay informed about the patient's clinical status. Additionally, the app offered features for communicating

XSL•FO
RenderX

via text message or secure videoconferencing with the patient and the assigned nurse.

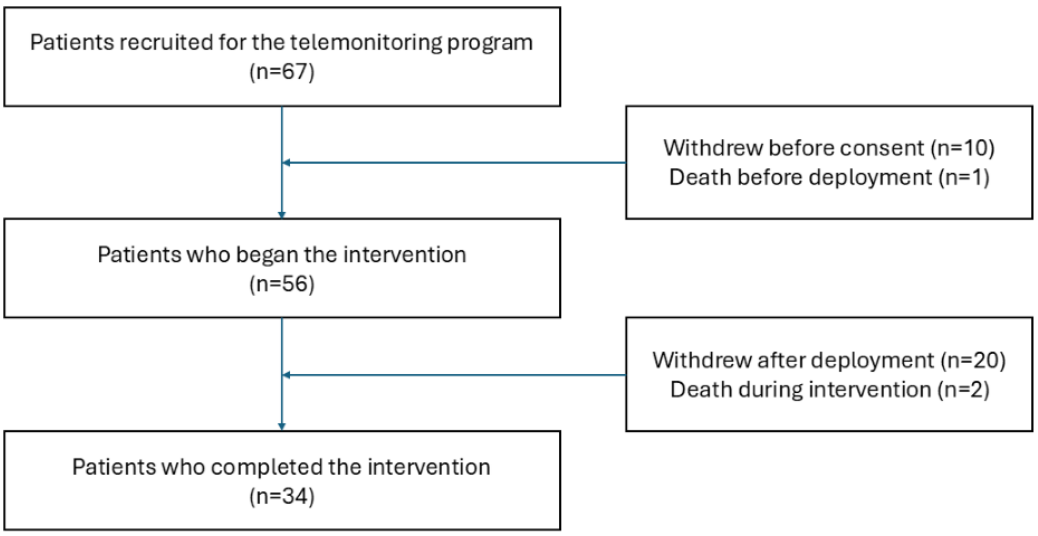
Participant Recruitment

A total of 20 nurses from the participating health care organizations were recruited by their supervisors. To participate in the project, they were required to complete an online training module on the clinical cardiac assessment of older adults [17]. This training consisted of 10 instructional videos, totaling approximately 4 hours, and concluded with a knowledge assessment.

Each nurse was responsible for identifying eligible patients from their caseload and initiating recruitment. Eligibility criteria for patients included the following: being aged 65 years or older, being francophone, having a diagnosis of HF, and currently receiving home care services related to that condition. These 4 criteria were mandatory. An additional, nonmandatory criterion was having an informal caregiver willing to participate in the study, aged 18 years or older, and possessing a smartphone or tablet with internet access.

A total of 67 patients were initially recruited into the study. Before the telemonitoring devices could be installed, however, 10 participants withdrew, while 1 patient died. For those who consented, the nurse proceeded to install the devices in the patient's home and provided basic instructions on their use. Technical support throughout the study was provided by Virtuose Technologies [18], the project's primary technology partner, alongside IT staff from each participating organization and the nurses involved in the intervention. During the intervention, 2 additional patients died, and 20 others chose to discontinue their participation. These withdrawals were observed across all participating sites. Among those who withdrew, the duration of participation varied widely, ranging from less than 1 week to more than 7 weeks. The reasons underlying early withdrawal are examined in the following section. Figure 1 summarizes patient recruitment, participation, and retention throughout the intervention.

Figure 1. Participant flowchart in the home telemonitoring program.



Recruited patients were asked to provide the name and contact information of their informal caregiver, when applicable. This information was shared with both the research team and the primary technology provider. Caregivers were then contacted by a member of the research team, who explained the study procedures and obtained their informed consent. Among the 34 patients who completed the intervention, 28 were supported by an informal caregiver who also took part in the study. Once consent was obtained, caregivers received an email from Virtuose Technologies with instructions on how to download the Proximity mobile app on their smartphone or tablet.

Lastly, as part of the onboarding process, the patient's community pharmacy was identified, and the pharmacy owner was invited to collaborate with the project to allow the prescription management system to interface with the smart xPill dispenser. Connection costs were fully covered by the company providing the smart xPill to enable seamless data integration.

Data Collection

Semistructured interviews were conducted at the end of the study with all 34 patients who completed the intervention, 16 of the 20 patients who withdrew early, 21 informal caregivers, and 20 nurses. Among informal caregivers, 13 participated in individual telephone interviews, while 8 took part in joint interviews conducted together with the patient they supported. Among the nursing staff, 18 were interviewed face-to-face, and 2 participated via telephone.

Of the 34 patients who completed the program, 32 were interviewed in person in their homes, while the remaining 2 participated via telephone. All 16 patients who withdrew early were interviewed by telephone. These interviews were analyzed alongside those of participants who completed the intervention. Although generally shorter, they provided valuable insights into barriers to sustained engagement, including emotional fatigue, technical difficulties, and misalignment with daily routines. Consequently, they were particularly informative in identifying

XSL•FO
RenderX

factors associated with disengagement and early withdrawal from the home telemonitoring program.

All interviews were conducted by a member of the research team with training in qualitative research and extensive prior experience studying digital health interventions. Conducting all interviews with the same interviewer helped ensure consistency in the interview process and familiarity with the study context. The interviewer had no prior clinical relationship with the participants.

Interview guides were developed by the research team to explore participants' experiences with the intervention while remaining closely aligned with the study's conceptual framing. Separate guides were prepared for patients, nurses, and informal caregivers to reflect their distinct roles in the program. The guides were informed by the 2 conceptual frameworks used in this study. While these frameworks helped structure the domains explored during interviews, the guides were designed to remain flexible and open-ended to allow participants to describe their experiences in their own terms and to surface issues not anticipated by the frameworks. The interview guides are provided in Multimedia Appendix 1.

No repeat interviews were conducted. However, the combination of interviews across multiple participant groups and iterative analytic discussions allowed the research team to refine and validate emerging interpretations.

The average duration of interviews varied by participant type: 45 to 60 minutes for patients who completed the program, 20 to 30 minutes for nurses, and 5 to 10 minutes for patients who withdrew early and informal caregivers. All interviews conducted in person were audio-recorded with participants' consent, except for 2 patients who declined recording. Telephone interviews were not recorded, but detailed notes were taken immediately after each interview to capture key points, striking statements, and unexpected reactions. Audio recordings were listened to in their entirety by the interviewer and the lead author to ensure familiarity with the full content of each interview.

Rather than producing full verbatim transcripts for all interviews, segments of the recordings directly relevant to the study's research objective were transcribed verbatim. These excerpts included both confirmatory and disconfirming observations related to participants' experiences with the home telemonitoring intervention. Field notes and analytic discussions among all team members were used to ensure that themes emerging outside the conceptual frameworks were not overlooked during the coding process.

Data Analysis

We conducted a directed content analysis, following the approach outlined by Hsieh and Shannon [19]. This method enabled the systematic examination of qualitative data using existing theoretical frameworks to guide the initial coding process while remaining open to insights emerging from participants' accounts. Interview notes and transcripts were first reviewed in their entirety to foster immersion in the data and develop a comprehensive understanding of participants' experiences.

A preliminary coding scheme was developed based on the study's main objective and sensitizing concepts drawn from the 2 guiding frameworks: the PBA [10] and the PCP framework [11]. The 2 frameworks were applied in a complementary manner rather than as independent analytic lenses. During coding and interpretation, particular attention was given to how participants' technofunctional experiences with the devices interacted with their relational experiences of care. For example, technical frustrations or usability challenges sometimes influenced perceptions of clinical support, whereas timely and empathetic follow-up from nurses often mitigated negative reactions to the technology. This integrative perspective helped illuminate how digital tools and relational care processes jointly shaped participants' overall experience of telemonitoring.

The interviewer also led the coding process and was therefore highly familiar with the interview material and study context. The coding process was highly iterative. Interview notes and transcripts were reviewed to identify segments relevant to the study's objective, which were then coded using sensitizing concepts derived from the 2 frameworks. Throughout the analytic process, coding decisions and emerging interpretations were discussed during regular team meetings. These discussions allowed the research team to examine the coder's analytic reasoning, question assumptions, and refine the organization of themes. Differences in interpretation were addressed through discussion until a shared understanding of the coding structure was reached.

Although the analysis was initially guided by the 2 conceptual frameworks, attention was paid to identifying experiential elements that did not fit neatly within these categories. The coding process, therefore, remained flexible and iterative, allowing the coding scheme to evolve as new insights emerged from the data.

Rigor and Reflexivity

Two strategies were used to enhance the trustworthiness of the analysis: triangulation of perspectives and collaborative interpretation. Specifically, we triangulated insights across the 3 participant groups interviewed (patients, informal caregivers, and nurses) to capture multiple perspectives on the home telemonitoring experience. Reflexivity was supported through regular analytic discussions among team members with diverse disciplinary backgrounds, including health informatics, nursing, and qualitative research.

We also drew on member-checking opportunities during follow-up meetings with clinical managers from each participating organization, during which preliminary interpretations were discussed and validated in light of their knowledge of the program's implementation and patient population.

Analytic transparency was supported by providing a detailed description of the coding procedures, explaining how the conceptual frameworks guided the analysis, and illustrating themes with representative excerpts from participants across the 3 groups interviewed. Data coding and theme development were conducted primarily by 1 member of the research team, with iterative peer debriefing sessions held with other team

XSL•FO
RenderX

members throughout the analytic process. These discussions helped challenge preliminary interpretations, ensure that coding decisions remained grounded in participants' accounts, and refine the organization of themes. Particular attention was given to identifying experiences that diverged from dominant patterns in the data, such as frustration, fatigue, or disengagement with the technology.

Although the research team was not involved in the initial design of the telemonitoring intervention, 2 authors participated in monthly meetings of the project's executive committee, and 1 author regularly attended weekly operational meetings in each participating organization. While their role in these settings was primarily observational and advisory, we acknowledge that this presence, and occasional input, may have influenced implementation dynamics or participant behavior. To mitigate this potential bias, we relied on triangulation of data sources, consensus-based interpretation, and a sustained effort to ground our analysis in participants' experiences rather than in project objectives or institutional priorities.

Ethical Considerations

This study received ethics approval from the ethics board of the CIUSSS Chaudière-Appalaches (#MP-23-2023-1037) on July 11, 2023. Ethical approval for the 2 other participating sites (CISSS Outaouais and CISSS Gaspésie) was granted through institutional agreements recognizing the lead site's ethics review. As noted above, all participants received information sheets and provided written or verbal informed consent prior to participation. Identifying information was removed from notes and transcripts, and data were stored on secure, password-protected servers. Participation in this study was voluntary and uncompensated.

As part of the informed consent process, patients were asked to identify the informal caregiver they wished to involve in the intervention. Only the caregiver explicitly authorized by the patient was granted access to the patient's health-related data via the Proximity mobile app. Patients were informed about the nature of the data shared, and participation remained entirely voluntary. For participants with limited digital literacy, additional support was provided to ensure they understood the implications of data sharing and retained autonomy in decision-making.

Results

Profile Participants

As shown in Table 1, 44% (15/34) of the patients who completed the intervention were female, with a mean age of 83 (SD 7.7) years. A total of 38% (13/34) resided in private seniors' residences, and more than half lived with a spouse or one of their children. Most participants (28/34, 82%) received active support from an informal caregiver as part of the project. In most cases, the caregiver was one of the user's children, followed by spouses and members of the extended family. In terms of perceived health status, 47% (16/34) of users described their health as fair, while the others considered it to be good or very good.

XSL•FO
RenderX

Table 1. Profile of the patients who completed the intervention (n=34).

Characteristic	Value
Female, n (%)	15 (44)
Age (years), mean (SD)	83 (7.7)
Reside in private seniors' residences, n (%)	13 (38)
Live with a spouse or relative, n (%)	21 (62)
Have an informal caregiver, n (%)	28 (82)
Perceive health status as, n (%)	
Fair	16 (47)
Good or very good	18 (53)
Perceive their level of digital literacy as, n (%)	
Low or very low	18 (53)
Average	9 (26)
High or very high	7 (21)
Use of digital devices, n (%)	
Smartwatch	34 (100)
Virtual assistant tablet	34 (100)
Bluetooth-enabled scale	31 (91)
Smart xPill dispenser	21 (62)
Perceive their autonomy level in the use of digital devices as, n (%)	
Full autonomy	27 (79)
Used with help of caregiver	4 (12)
Mainly used by caregiver	3 (9)

Regarding digital literacy, more than half of the participants reported having low or very low comfort with digital technologies. In this context, approximately 20% (7/34) were assisted by their informal caregiver in completing the symptom questionnaire (anamnesis) and logging their daily medication intake using the smart xPill dispenser. All patients used the connected smartwatch, and the vast majority also used the connected scale. However, about one-third of participants were unable to use the smart xPill, as their community pharmacy was not yet connected to the system.

Among the individuals who initially agreed to participate but ultimately did not sign the consent form, several expressed concerns about their ability to manage the technological devices during formal follow-up. Some described their health as too fragile to take on the additional demands of the project and expressed fears that participation might cause further stress or anxiety. Others felt overwhelmed after reviewing the information sheet or watching the onboarding videos, which led to doubts about their ability to manage the devices. In some cases, these concerns were discussed with informal caregivers, resulting in a mutual decision to decline participation. Overall, these early withdrawals appear to have been driven by emotional apprehension, perceived cognitive or physical limitations, and a subjective sense of unpreparedness to engage with digital health technologies.

Another group consisted of patients who withdrew after the telemonitoring devices had been installed in their homes. This group included a higher proportion of women (12/20, 60%) and had a slightly higher mean age of 84 (SD 6.9) years. As shown in Table 2, 60% (12/20) of these participants had an informal caregiver involved in the project. On average, they remained enrolled for 4.6 weeks before deciding to discontinue their participation. Compared with the 34 patients who completed the intervention, the postdeployment withdrawal group tended to be older and had lower rates of caregiver involvement.

XSL•FO
RenderX

Table 2. Demographic characteristics of patient subgroups by participation status.

Characteristic	Surviving patients who withdrew early after deployment (n=20)	Patients who completed the 6-month intervention (n=34)
Female, n (%)	12 (60)	15 (44)
Age (years), mean (SD)	84 (6.9)	83 (7.7)
With an informal caregiver, n (%)	12 (60)	28 (82)
Duration of participation (weeks), mean (SD)	4.6 (1.1)	24 (0.2)

Interviews with the 20 patients who withdrew after deployment revealed a distinct set of challenges arising from their direct experience with the system. Many reported emotional fatigue from repetitive tasks, cognitive overload due to the use of multiple devices, and frustration with technical issues such as unreliable connectivity or unintuitive interfaces. Some also found it difficult to adjust their daily routines to accommodate the system's requirements, such as scheduled prompts or weighing times, especially in the absence of regular caregiver support. Taken together, these accounts suggest that advanced age, lower digital confidence, and limited support networks were important contextual factors that contributed to early disengagement from the intervention.

As shown in Table 3, the participating nurses were predominantly women, with a mean age of 40 (SD 6.1) years. A total of 65% (13/20) held clinical nurse positions, and, on average, they had been working full-time in home care for approximately 8 years.

Table 3. Profile of the participating nurses (n=20).

Characteristic	Value
Female, n (%)	19 (95)
Age (years), mean (SD)	40 (6.1)
Job position, n (%)	
Registered nurse	7 (35)
Clinical nurse	13 (65)
Average experience in home care (years), mean (SD)	8 (2.1)
Full-time job, n (%)	20 (100)

Technofunctional Experience With Home Telemonitoring

Overview

We apply here the 4 core elements of the PBA to interpret the technofunctional experiences of patients and caregivers as they interacted with the multidevice home telemonitoring intervention. As mentioned earlier, this framework allows us to examine how participants engaged with the technologies cognitively, emotionally, and behaviorally within their daily lives. Following this thematic analysis, we present device-specific accounts to illustrate how these broader experiential patterns manifested across the 4 digital tools used in the intervention. These accounts help contextualize the variability in usability, perceived burden, and emotional response linked to each device.

Contextual Tailoring

Participants' ability to engage with the home telemonitoring program depended strongly on how well the devices aligned with their daily routines, physical capabilities, and living environments. Many participants described entering unfamiliar territory when first encountering the digital tools. Initial apprehension was common, with patients voicing fears of breaking the equipment or failing to use it correctly. As 1 patient shared, "At first, I was nervous. I didn't want to mess anything up, but the more I used it, the more confident I got" (Patient #7). These early fears diminished with hands-on practice, clear instructions, and encouragement from caregivers, demonstrating the importance of tailoring onboarding to individual learning styles and contexts.

However, integration into daily life was more difficult for participants whose routines were less predictable or socially constrained. For example, patients living in senior residences often reported difficulty aligning measurement times with scheduled meals or medication routines:

> *Sometimes I would miss the time to fill out the questionnaire or take my medication because I was still in the cafeteria. It depended on the day.* [Patient #6]

These observations highlight how routines must be considered in designing and delivering home telemonitoring interventions.

Connectivity issues compounded the challenge of learning. Several participants, particularly those living in rural settings, reported problems with Bluetooth synchronization or weak cellular signals, which led to repeated prompts or lost data. As 1 patient explained: "When it glitched, I would lose patience. I eventually figured out how to restart the system myself, but at first, I was overwhelmed" (Patient #23). These experiences

XSL•FO
RenderX

illustrate the iterative process of digital learning, where contextual barriers often dictate the pace and success of adaptation.

Emotional and Cognitive Responses to Technology

Participants expressed a wide range of emotional responses to the home telemonitoring program, ranging from reassurance and increased confidence to frustration, fatigue, and anxiety. Many felt reassured knowing their health data were monitored in real time by a clinical team, "Even if she [nurse] wasn't here, I felt like she was looking after me every day," said one patient (Patient #11). This sense of remote presence contributed to perceived security, especially for those living alone.

Others, however, experienced anxiety, irritation, or emotional fatigue. This was especially true when dealing with repetitive tasks, frequent prompts, or nighttime alerts. As 1 patient remarked:

> *It felt useless to me, because it's always the same questions… and the same answers. So, I eventually stopped responding to the robot.* [Patient #1]

A nurse confirmed these frustrations:

> *Some of my patients even mentioned being woken up repeatedly by the tablet's sound alerts, which was experienced as a source of frustration.* [Nurse #17]

When feedback loops were missing, the technology was often perceived as intrusive or indifferent, diminishing its credibility and value.

Perceived Relevance and Burden of Technology Use

Participants evaluated the usefulness of the telemonitoring system in relation to the effort required to use it, and their experiences varied substantially across the different devices. The smartwatch, which required minimal user input, was generally well received. "I liked that I didn't have to do much. But I never really knew what it was tracking or if it was working" (Patient #10). In contrast, the connected scale, though simple to use, lost value when readings were not acknowledged: "I weighed myself every morning, but it didn't seem to matter. No one ever mentioned the numbers" (Patient #5). For its part, the voice-activated tablet received mixed reviews. While some appreciated the reminders, others struggled with the interface. "It would talk to me, but I couldn't always hear it or figure out how to answer. It made me nervous" (Patient #21).

These examples suggest that bundling multiple devices under a single intervention may obscure the highly variable user experiences across technologies. Emotional and functional reactions depended not only on individual digital skills but also on how well each device fit into the participant's everyday life.

Autonomy, Control, and Iterative Engagement

Over time, some participants developed greater confidence and autonomy in using the devices, whereas others experienced persistent difficulties that led to disengagement. For some participants, learning to troubleshoot devices fostered a sense of control and accomplishment. "After a while, I knew how to fix it myself. That made me feel more in control" (Patient #17). Others, however, experienced persistent problems that eroded their confidence. "I didn't know what was wrong with the xPill, so I stopped using it. It was too stressful" (Patient #12). This divergence in experience illustrates the need for systems to support users not only during onboarding but throughout their engagement with technological devices.

Among participants who withdrew early, emotional and cognitive overload were frequent themes. "I just knew I wasn't going to be able to manage it. It's too much stress for someone like me" (Patient #20). In some cases, caregiver frustration contributed to early dropout: "My husband got frustrated with the machine after the second week. It was just easier to go back to the usual way" (Caregiver #4). These narratives point to the importance of iterative and adaptive support strategies to retain engagement over time.

Relational and Experience Aspects of Remote Care

Perceived Professional Presence at a Distance

Many participants reported that the home telemonitoring program created a sense of ongoing clinical presence, even when direct communication with nurses was limited. This theme aligns with the person-centered processes domain of the PCP framework, particularly the dimension of sympathetic presence. Many participants described feeling reassured by the knowledge that their health status was being monitored, even in the absence of direct, face-to-face contact. One patient expressed this sense of invisible support: "Even if she [nurse] wasn't here, I felt like she was looking after me every day" (Patient #11).

Nurses also made intentional efforts to convey a sense of presence despite the absence of face-to-face interactions. Several described how, without visual cues, they learned to attune more closely to patients' vocal tone and inflection. As 1 nurse noted, "You have to 'read' the patient's tone of voice… I could tell when someone wasn't doing well just by how they said hello" (Nurse #2). To maintain an empathetic connection, some nurses also relied on open-ended questions and lengthier conversations. One nurse explained, "When I didn't see their body language, I compensated by asking more open-ended questions" (Nurse #4).

These accounts illustrate that while digital devices introduce new barriers to physical proximity, person-centered care practices can still be upheld through deliberate and thoughtful communication strategies.

Shared Decision-Making and Tailored Interactions

Participants differed in the extent to which they wished to be involved in decisions related to the home telemonitoring program and their care. Shared decision-making represents a critical process within McCance and McCormack's [11] framework. Participants' experiences varied. Some felt genuinely involved when nurses contacted them to discuss their health status or device readings. One patient recounted: "She would call and ask how I felt before deciding anything. It felt personal" (Patient #5). Similarly, another participant explained how their input was acknowledged: "She asked me if it was okay to continue using the smart xPill since it was giving me trouble. I appreciated that" (Patient #9).

XSL•FO
RenderX

Others, however, preferred a more passive role and deferred entirely to clinical expertise. As 1 participant explained: "I just followed what they said. I trusted them" (Patient #13). A second participant echoed this sentiment: "I'm not the kind to ask questions. If the nurse says it's good, I go with that" (Patient #18).

These differing accounts underscore the importance of communication strategies that are responsive to individual preferences for autonomy and involvement. While some users welcomed collaborative decision-making, others derived comfort from trusting clinical judgment, highlighting the need for flexibility in how shared decision-making is operationalized in remote care settings.

Emotional Reassurance and Well-Being

For many patients and caregivers, the intervention generated a sense of reassurance by signaling that clinical professionals were overseeing the patient's condition. This theme maps onto the person-centered outcomes domain of McCance and McCormack's [11] framework, specifically the importance of emotional well-being and reassurance. Several participants expressed comfort in knowing they were being watched over, even in the absence of direct communication. One patient put it simply: "It was reassuring, knowing someone was checking in, even if we didn't talk every day" (Patient #1). Another noted: "It was nice to know someone was checking my data every day. Even if we didn't talk, I felt watched over in a good way" (Patient #8). In short, this quiet sense of presence, mediated through the smart devices, was a recurring source of comfort for many.

Informal caregivers also shared this emotional benefit. Knowing that clinical professionals were monitoring their loved ones provided peace of mind and alleviated the perceived burden of constant vigilance. As 1 caregiver explained: "I knew that if something went wrong, the nurse would see the data and call. That took a weight off my shoulders" (Caregiver #4).

Limitations of Remote Engagement

Despite these benefits, participants also recognized important limits to digitally mediated care, particularly when compared with in-person interactions. This aligns with the contextual sensitivity emphasized in McCance and McCormack's [11] framework. Some users acknowledged that the technological interface, while functional, could not replace the warmth and nuance of in-person interaction. One patient captured this sentiment: "The technology was good, but it didn't replace a nurse sitting next to you, talking face to face" (Patient #12). Nurses echoed this concern. Some reported emotional strain stemming from the lack of tangible presence and the uncertainty of whether patients felt genuinely supported. One nurse reflected: "Sometimes I wondered if they [patients] felt alone with the devices. It's not easy to know if your presence is felt" (Nurse #3).

Yet, many nurses believed that their efforts to stay in contact, personalize their communication, and respond promptly still conveyed a sense of relational continuity. One nurse recounted positive feedback from patients: "Several told me, 'It's like you're still here, even if we don't see you.' That's what we wanted" (Nurse #1).

These perspectives reinforce the idea that, while digital health technologies alter the medium of interaction, person-centered care can still be practiced and felt if thoughtfully adapted.

Discussion

Principal Results

Using the complementary lenses of the PBA and the PCP frameworks, the study findings highlight three central insights: (1) engagement with home telemonitoring technologies depends strongly on users' routines, digital confidence, and perceived burden; (2) relational support from nurses plays a critical role in shaping how patients interpret and tolerate technological challenges; and (3) informal caregivers often act as key mediators of sustained engagement with the system.

From a technofunctional perspective, findings highlight the importance of designing home telemonitoring programs that are emotionally acceptable, contextually adaptable, and responsive to user needs and constraints. Participants' engagement varied widely based on their digital literacy, cognitive capacity, comfort with wearable or interactive devices, and the fit of the system with their daily routines. While some developed confidence over time, especially when supported by informal caregivers, others experienced persistent difficulties, particularly when their routines were irregular or when onboarding was not sufficiently individualized.

Participants often reacted emotionally to the devices, reporting reassurance, frustration, or anxiety depending on their ability to interact successfully with a particular technology. Repetitive symptom questionnaires and poorly timed alerts contributed to emotional fatigue, especially when feedback was absent or unclear. These reactions were compounded when technical issues such as Bluetooth disconnections or connectivity failures disrupted user routines or sleep. In some cases, persistent problems led to disengagement or early withdrawal.

Importantly, the intervention did not impact all devices equally. The smartwatch and scale were generally well accepted due to their ease of use and alignment with daily habits. In contrast, the tablet and smart pill dispenser posed greater usability and cognitive challenges, particularly for participants with lower digital confidence or limited caregiver support. These device-specific variations underscore the need to avoid one-size-fits-all solutions and instead tailor devices to user profiles and contextual factors.

From a relational standpoint, participants often found the experience of being monitored reassuring, provided that follow-up from nurses was timely and personalized. Short, routine interactions (eg, phone calls or messages) reinforced the sense of being cared for, even at a distance. However, when follow-up was perceived as inconsistent or impersonal, patients reported feeling overlooked or abandoned. Nurses themselves reported challenges in maintaining relational presence without visual cues and in addressing technical issues beyond their clinical role.

XSL•FO
RenderX

Informal caregivers emerged as critical, if underrecognized, mediators of technology use and continuity. Their availability, digital literacy, and relationship to the patient shaped the sustainability of engagement. Caregivers often provided troubleshooting support and emotional reassurance, especially in cases in which users felt overwhelmed by the technology.

Lastly, while the PBA and PCP frameworks helped structure our analysis along technofunctional and relational dimensions, participants' experiences often spanned both domains simultaneously. In practice, engagement with the telemonitoring system emerged from the interaction between how users experienced the devices and how they experienced the care relationships surrounding those devices.

As shown in Table 4, participants' experiences with the telemonitoring intervention cannot be fully understood through a purely technological or purely relational lens. Instead, many of the themes identified in the analysis reflect the interaction between technofunctional engagement with the devices and the relational dynamics of care. For instance, the sense of reassurance reported by many patients did not stem solely from the monitoring devices themselves, but from the perception that nurses were actively reviewing the transmitted data and would intervene if necessary. Similarly, frustrations associated with repetitive tasks or technical difficulties were often amplified when feedback from the clinical team was perceived as limited or delayed. Conversely, timely communication and empathetic follow-up frequently mitigated negative reactions to the technology and encouraged sustained engagement. Taken together, these findings suggest that the effectiveness and acceptability of home telemonitoring programs depend not only on the usability of digital tools but also on the quality of relational support surrounding their use.

Table 4. Integration of technofunctional and relational dimensions of participants' experiences with home telemonitoring.

Integrated theme	PBA[a] framework (technofunctional lens)	PCP[b] framework (relational lens)	Combined insight
Reassurance through monitoring	Users interpreted device monitoring as health surveillance	Nurses' follow-up created a sense of presence	Reassurance emerged from the interaction between technology and relational care, not from the device alone
Burden of repeated tasks	Repetitive questionnaires and alerts created fatigue	Lack of feedback reduced perceived care responsiveness	Techno-functional burden was amplified when relational feedback was missing
Adaptation and confidence	Users developed autonomy through learning the devices	Supportive nurses reinforced confidence	Sustained engagement depended on both usability and relational encouragement
Disengagement and withdrawal	Technical frustration and cognitive overload	Limited caregiver support and relational reassurance	Withdrawal often resulted from combined technical and relational gaps

[a]PBA: Person-Based Approach.

[b]PCP: Person-Centered Practice.

The following section situates these findings within the broader literature on home telemonitoring and digitally mediated care.

Comparison With Prior Work

Our findings reinforce and extend prior research on home telemonitoring in aging populations, particularly along 2 key dimensions: technofunctional engagement and relational and experiential aspects of care. On the functional side, earlier studies have emphasized the importance of usability, ease of integration, and access to technical support in facilitating telehealth adoption [19,20]. This study builds on this foundation by identifying additional factors that influence sustained engagement, including emotional burden, alignment with daily routines, and users' capacity to adapt and troubleshoot technical issues—insights that closely reflect the principles of the PBA [10]. In particular, this framework highlights the importance of contextual tailoring, emotional acceptability, and perceived autonomy, all of which were evident in our participants' experiences.

Our findings also confirm prior evidence that alert fatigue, lack of personalized feedback, and the perceived irrelevance of prompts can erode motivation over time [20,21]. More broadly, the literature on older adults and digital health underscores the importance of designing interventions that are emotionally acceptable and minimally burdensome, rather than merely functionally efficient [8].

Although Feroz and Ahmad [22] did not explicitly apply the PBA, their systematic review of usability factors in secure mobile health apps aligns with many of its core tenets, including the role of emotional responses to technology, the challenges posed by unfamiliar digital environments, and the need for robust support structures. Our findings complement theirs by illustrating how these dynamics unfold in a complex, real-world home telemonitoring context involving multiple devices and interdependent roles among patients, caregivers, and nurses. In doing so, this study extends person-centered usability principles beyond app-based systems to the broader ecosystem of digitally mediated chronic care.

Turning to the relational and experiential dimensions of care, our findings align with prior research showing that digital systems can sustain a sense of connection between patients and health care professionals when supported by consistent follow-up and empathetic communication [23,24]. At the same time, few studies have examined these dynamics simultaneously from the perspectives of patients, informal caregivers, and nurses. By triangulating insights across these groups, this study highlights that the perceived quality of care relationships depends not only on the presence of user-friendly technological

XSL•FO
RenderX

devices but also on the continuity and sensitivity of communication, factors essential to maintaining therapeutic engagement at a distance.

Our results also resonate with those of Murphy et al [25], who used McCance and McCormack's [11] framework to examine older adults' experiences with integrated care. Both studies emphasize the central role of sympathetic presence, authentic engagement, and shared decision-making in shaping how care is experienced. While Murphy et al [25] focused on in-person multidisciplinary care, our findings suggest that these person-centered processes can also be preserved and meaningfully experienced within digitally mediated care contexts.

Finally, the variation observed in participants' preferences for involvement in decision-making can be interpreted through Savoli et al's [26] typology of patient self-management styles. In their study of patient portal use, Savoli et al [26] identified 3 profiles—autonomous, engaged, and reliant—based on how patients attribute responsibility for self-management and perceive the role of technology. Participants in this study who welcomed dialogue and shared decision-making resembled the engaged profile, whereas those who preferred to defer decisions to nurses reflected the reliant profile. These parallels highlight the importance of tailoring communication and support strategies to individual preferences for autonomy and relational engagement in home telemonitoring programs.

Implication for Practice

Building on these findings, several practical implications emerge for the design and implementation of home telemonitoring programs. Drawing on the principles of the PBA and PCP framework, we emphasize the importance of designing interventions that are emotionally acceptable, contextually tailored, and responsive to the lived experiences of patients and their caregivers.

First, onboarding and ongoing support must be individualized. Older adults vary widely in their levels of digital literacy, cognitive capacity, and learning preferences. Training should therefore be incremental and reinforced through scheduled follow-ups, with clear, accessible materials to help users navigate common technical challenges. Providing personalized guidance during the adaptation phase can reduce anxiety and build early confidence in system use.

Second, home telemonitoring tools must accommodate diverse daily routines. Our results highlight the friction created when device prompts conflict with users' eating habits, medication schedules, or lifestyle preferences. Data entry and alert schedules should be easily configurable by patients or caregivers to minimize disruption and support sustainable engagement. Third, relational continuity remains essential, even in digitally mediated care. Participants valued consistent, personalized interactions with nurses, and expressed distress when follow-up felt irregular or impersonal. Assigning a dedicated referent nurse or implementing minimal standards for proactive communication may help reinforce the therapeutic relationship and enhance users' sense of being seen and supported.

Fourth, regular feedback to patients is crucial for motivation. The absence of acknowledgment for submitted data diminished the perceived value of system use. Even brief, automated messages (eg, "Your data looks stable today") can reinforce engagement, provide emotional reassurance, and maintain participation over time. Fifth, interfaces should allow for user-controlled modes of interaction. While some patients appreciated detailed access to their health data, others preferred to limit their exposure to potentially anxiety-inducing information. Home telemonitoring interventions should offer flexible configurations to respect these differences and promote a sense of control.

Sixth, caregivers require tailored onboarding and communication strategies. Their ability to support patients often depended on their understanding of how the technology worked and their preferred communication channel (eg, SMS, phone, or app). Dedicated support and training for caregivers, including technical troubleshooting and role clarification, can strengthen the triad of care among patients, caregivers, and clinicians. Seventh, technical reliability is foundational, especially in rural or low-connectivity environments. Participants reported repeated frustrations linked to connectivity failures, alerts triggered by errors, and synchronization issues. Hybrid network strategies (eg, Wi-Fi and cellular backup), intuitive error messaging, and simple troubleshooting instructions are critical to sustaining trust in the system.

Eighth, integration with existing clinical workflows and community resources must be prioritized. The effectiveness of tools like the smart xPill dispenser depends on accurate and updated medication profiles, which in turn require coordination with community pharmacists. Home telemonitoring programs should not operate in isolation, but rather as part of a broader, interprofessional care ecosystem.

Finally, while many participants reported positive experiences, it is important to acknowledge that frustration, emotional fatigue, and technical difficulties were common and, for some, led to early withdrawal. These negative experiences were often exacerbated by repetitive prompts, lack of real-time feedback, or device malfunction, especially in the absence of caregiver support. For several patients, the cognitive and emotional effort required to engage with the devices outweighed the perceived benefits. These findings underscore a critical tension in distant care. Even well-designed interventions can generate unintended burdens that erode engagement over time. As such, scaling similar home telemonitoring programs without addressing these experiential barriers may exacerbate disparities in access, increase dropout rates, and strain frontline support systems. Future implementations should therefore prioritize sustained user support, adaptive personalization, and continuous feedback mechanisms to mitigate these risks.

Study Limitations

Three main limitations should be considered when interpreting the findings of this study. First, the transferability of our results is constrained by the study's geographic scope and sample size. Data were collected from 3 health care organizations located within a single Canadian province, and only 34 patients and their caregivers completed the full duration of the pilot program.

XSL•FO
RenderX

Although the participating sites were selected to represent rural, semiurban, and urban contexts, the relatively small number of participants and the contextual specificity of the intervention limit the extent to which the findings can be generalized to broader populations or settings.

At the same time, the goal of this qualitative study is not statistical generalization but analytical transferability. By providing detailed descriptions of the intervention, participant characteristics, and care context, we aim to enable readers to assess the extent to which the patterns identified in this study—such as the interplay between technological usability, caregiver support, and relational continuity with clinicians—may resonate with similar home telemonitoring initiatives in other health care environments.

The transferability of our findings to contexts outside Canada should also be interpreted with caution. The home telemonitoring program was implemented within the publicly funded Canadian health care system, where home care services, nursing roles, and digital health initiatives may be organized differently than in other countries. For instance, the involvement of home care nurses in reviewing patient data and conducting follow-up reflects organizational features that may not be present in all health systems. Nevertheless, several of the dynamics identified in this study, including the influence of digital literacy, caregiver involvement, technological burden, and the relational aspects of remote care, have been reported in home telemonitoring initiatives in other settings, suggesting that these findings may still offer useful insights for similar programs implemented in different health care contexts.

Second, the study may be subject to selection bias. Indeed, many patients agreed to participate based on a strong preexisting relationship with their home care nurse, suggesting that trust in the health care professional played a central role in motivating enrollment. Others were motivated by altruistic values such as a desire to contribute to research or "give back" to the health system. These motivations, while valuable, may have skewed the participant sample toward individuals more likely to engage with the intervention or tolerate its demands, potentially underrepresenting the perspectives of those with lower motivation, weaker support networks, or more limited digital literacy.

Third, although we used a rigorous approach to directed content analysis [19], the inherently interpretive nature of qualitative research carries the risk of subjectivity in coding and theme development. Even though coding was performed rigorously, the analytical process was inevitably shaped by our prior assumptions, research objectives, and theoretical framing. The risk of interpretive bias is especially salient given that our analysis was guided by 2 preexisting conceptual frameworks, which, while helpful in structuring the analysis, may have constrained the emergence of unexpected insights. We sought to mitigate this risk by maintaining an iterative coding process that remained open to experiential elements not fully captured by the frameworks. Divergent accounts, including frustration, disengagement, and early withdrawal from the program, were examined alongside more positive experiences to ensure that the analysis reflected the diversity of participants' perspectives.

Conclusions

Home telemonitoring programs that involve multiple digital devices can empower older adults and their caregivers, but only when they are flexibly integrated into daily routines, emotionally acceptable, and consistently supported both technically and relationally. The PBA provided a valuable framework for interpreting how users engaged with the intervention across functional, emotional, and contextual dimensions. It emphasized the importance of personalization, perceived relevance, and iterative adaptation in shaping technology use over time. Complementing this, the PCP framework helped illuminate how participants experienced presence, shared decision-making, and relational continuity, even in the absence of face-to-face contact. This framework foregrounded the relational and emotional texture of care, drawing attention to the processes through which users felt acknowledged, supported, and empowered.

By combining these 2 lenses, this study captures both the tangible functionality and the intangible relational dynamics of digitally mediated care. Together, they underscore the need for patient- and caregiver-centered design principles, sustained interpersonal engagement, and robust interprofessional coordination. Evaluating home telemonitoring interventions through both technofunctional and relational perspectives offers a fuller understanding of how such programs are lived and how they can be more effectively implemented at scale.

Acknowledgments

The authors would like to express our sincere gratitude to all the patients, informal caregivers, and nurses who generously participated in this study and shared their experiences. We also thank the Quebec Ministry of Health and Social Services for its financial support, which made this research possible. Finally, we extend our thanks to Anne-Marie Veillette for her valuable support during the conceptualization of the study.

In preparing this manuscript, the authors used ChatGPT 5.3 (OpenAI) to assist with proofreading, improving the structure of the text, and drafting the response to reviewers. All artificial intelligence (AI)–generated content was critically reviewed, edited, and approved by the authors to ensure its accuracy and compliance with ethical standards. This use of generative AI complies with JMIR Publications' guidelines on AI use in scholarly writing [27]. All authors had full access to the data and approved the initial and revised versions of this manuscript.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Funding

This work was conducted as part of a pilot initiative sponsored and funded by the Quebec Ministry of Health and Social Services. Budget planning and financial oversight were managed by the health autorities. The budget covered device acquisition, software licensing, training, technical support, and coordination activities. Detailed financial breakdowns were not shared with the research team.

Authors' Contributions

GP contributed to conceptualization, methodology, investigation, formal analysis, project administration, and writing of the original draft, as well as review and editing of the manuscript. MPM contributed to methodology, investigation, and manuscript review and editing. PV contributed to conceptualization, funding acquisition, resources, project administration, supervision, and manuscript review and editing. AC contributed to conceptualization, methodology, and manuscript review and editing. MSH contributed to conceptualization, supervision, and manuscript review and editing. MR contributed to the investigation and manuscript review and editing.

Conflicts of Interest

AC is an Associate Editor of the *Journal of Medical Internet Research* at the time of publication.

Multimedia Appendix 1
Interview guides.
[DOCX File , 18 KB - aging_v9i1e79797_app1.docx]

References

1. Goyal P, Maurer MS, Roh J. Aging in heart failure: embracing biology over chronology: JACC family series. JACC Heart Fail 2024 May;12(5):795-809 [FREE Full text] [doi: 10.1016/j.jchf.2024.02.021] [Medline: 38597865]
2. Liu G, Nguyen NQH, Wong KE, Agarwal SK, Boerwinkle E, Chang PP. Metabolomic association and risk prediction with heart failure in older adults. Circ Heart Fail 2024;17(3):e010896 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.123.010896] [Medline: 38426319]
3. Jaana M, Sherrard H, Paré G. A prospective evaluation of telemonitoring use by seniors with chronic heart failure: adoption, self-care, and empowerment. Health Informatics J 2019;25(4):1800-1814 [FREE Full text] [doi: 10.1177/1460458218799458] [Medline: 30247080]
4. Kitsiou S, Paré G, Jaana M. Effects of home telemonitoring interventions on patients with chronic heart failure: an overview of systematic reviews. J Med Internet Res 2015;17(3):e63 [FREE Full text] [doi: 10.2196/jmir.4174] [Medline: 25768664]
5. Li L, Ringeval M, Wagner G, Paré G, Ozemek C, Kitsiou S. Effectiveness of home-based cardiac rehabilitation interventions delivered via mHealth technologies: a systematic review and meta-analysis. Lancet Digit Health 2025;7(4):e238-e254 [FREE Full text] [doi: 10.1016/j.landig.2025.01.011] [Medline: 40023729]
6. Paré G, Moqadem K, Pineau G, St-Hilaire C. Clinical effects of home telemonitoring in the context of diabetes, asthma, heart failure and hypertension: a systematic review. J Med Internet Res 2010;12(2):e21 [FREE Full text] [doi: 10.2196/jmir.1357] [Medline: 20554500]
7. Jaana M, Paré G. Comparison of mobile health technology use for self-tracking between older adults and the general adult population in Canada: cross-sectional survey. JMIR Mhealth Uhealth 2020;8(11):e24718 [FREE Full text] [doi: 10.2196/24718] [Medline: 33104517]
8. Yan M, Sun W, Tan C, Wu Y, Liu Y. Analysis of factors influencing the willingness of Chinese older adults to use mHealth devices: nationwide cross-sectional survey study. J Med Internet Res 2025;27:e66804 [FREE Full text] [doi: 10.2196/66804] [Medline: 40053781]
9. Castonguay A, Paré G, Moreault MP, Hardy MS, Ringeval M, Voyer P. Implementation of a multidevice telemonitoring program for home-based nursing care in Quebec: qualitative report. JMIR Med Inform 2026;14:e83615 [FREE Full text] [doi: 10.2196/83615]
10. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
11. McCance T, McCormack B. The person-centred nursing framework: a mid-range theory for nursing practice. J Res Nurs 2025;30(1):47-60. [doi: 10.1177/17449871241281428] [Medline: 40093819]
12. Bate P, Robert G. Experience-based design: from redesigning the system around the patient to co-designing services with the patient. Qual Saf Health Care 2006;15(5):307-310 [FREE Full text] [doi: 10.1136/qshc.2005.016527] [Medline: 17074863]

XSL•FO
RenderX

13. Voice activated tablet. Samsung. URL: https://www.samsung.com/ca/ [accessed 2026-05-25]
14. Smartwatch. Biobeat. URL: https://www.bio-beat.com/ [accessed 2026-05-25]
15. Bluetooth-enabled weight scale. A&D Medical. URL: https://medical.andonline.com/ [accessed 2026-05-25]
16. xPill pro. Domedic. URL: https://domedic.ca/en/xpill-pro/ [accessed 2026-05-25]
17. Voyer P. L'examen clinique de l'aîné. Montréal: Pearson ERPI; 2017.
18. Virtuose Technologies. URL: https://virtuosetechnologies.com/en/ [accessed 2026-05-25]
19. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
20. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
21. Pols J. Care at a Distance: On the Closeness of Technology. Amsterdam: Amsterdam University Press; 2012.
22. Feroz I, Ahmad N. Systematic review of usability factors, models, and frameworks with blockchain integration for secure mobile health (mHealth) applications. Blockchain Healthc Today 2024;7(3):1-20. [doi: 10.30953/bhty.v7.357] [Medline: 39995486]
23. Green CA, Polen MR, Janoff SL, Castleton DK, Wisdom JP, Vuckovic N. Understanding how clinician-patient relationships and relational continuity of care affect recovery from serious mental illness: STARS study results. Psychiatr Rehabil J 2008;32(1):9-22 [FREE Full text] [doi: 10.2975/32.1.2008.9.22] [Medline: 18614445]
24. Morris ME. Enhancing relationships through technology: directions in parenting, caregiving, romantic partnerships, and clinical practice. Dialogues Clin Neurosci 2022;22(2):151-160. [doi: 10.31887/dcns.2020.22.2/mmorris]
25. Murphy S, Doyle M, Gallagher N. "They talk to you like you're a person": validating McCormack and McCance's person-centred practice framework through a qualitative study of older people's experiences of person-centred integrated care. Int J Integr Care 2023 [FREE Full text] [doi: 10.5334/ijic.icic23233]
26. Savoli A, Barki H, Pare G. Examining how chronically ill patients? Reactions to and effective use of information technology can influence how well they self-manage their illness. MIS Quarterly 2020;44:351-389. [doi: 10.25300/MISQ/2020/15103]
27. Leung TI, de Azevedo Cardoso T, Mavragani A, Eysenbach G. Best practices for using AI tools as an author, peer reviewer, or editor. J Med Internet Res 2023;25:e51584 [FREE Full text] [doi: 10.2196/51584] [Medline: 37651164]

Abbreviations

CISSS: Centre intégré de santé et de services sociaux
HF: heart failure
PBA: Person-Based Approach
PCP: Person-Centered Practice

Edited by R Yang; submitted 28.Jun.2025; peer-reviewed by J Edwards, A Roundtree; comments to author 20.Feb.2026; revised version received 12.Apr.2026; accepted 15.May.2026; published 16.Jun.2026.

Please cite as:
Paré G, Moreault MP, Voyer P, Castonguay A, Hardy MS, Ringeval M
Experiences of Older Adults and Caregivers With Home Telemonitoring for Heart Failure in Canada: Qualitative Study
JMIR Aging 2026;9:e79797
URL: https://aging.jmir.org/2026/1/e79797
doi:10.2196/79797
PMID:

©Guy Paré, Marie-Pierre Moreault, Philippe Voyer, Alexandre Castonguay, Marie-Soleil Hardy, Mickael Ringeval. Originally published in JMIR Aging (https://aging.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Early-Life Experiences as Moderators of the Relationship Between Extreme Heat and Alcohol Consumption among Older Adults: Quantitative Study

Hee Yun Lee[1], PhD; Su Hyun Shin[2], PhD; Yeon Jin Choi[3], PhD; Eun Young Choi[4], PhD; Hyunjung Ji[5], PhD

[1]School of Social Work, University of Georgia, 279 Williams Street, Athens, GA, United States

[2]Department of Family and Consumer Studies, University of Utah, Salt Lake City, UT, United States

[3]College of Social Work, University of Kentucky, Lexington, KY, United States

[4]School of Gerontology, University of Southern California, Los Angeles, CA, United States

[5]Department of Political Science, The University of Alabama, Tuscaloosa, AL, United States

Corresponding Author:
Hee Yun Lee, PhD
School of Social Work, University of Georgia, 279 Williams Street, Athens, GA, United States

Abstract

Background: Climate change has intensified the frequency and duration of extreme heat events worldwide, posing growing public health risks, particularly for older adults who are physiologically more susceptible to heat-related illnesses. Concurrently, alcohol consumption among older adults in the United States has risen significantly over the past 2 decades, increasing vulnerability to dehydration, cardiovascular strain, and cognitive impairment during heat exposure. Emerging evidence suggests that environmental stressors such as extreme heat may exacerbate maladaptive coping behaviors, including alcohol use; however, few studies have examined this association in aging populations. Moreover, little is known about how early-life experiences such as childhood adversity or positive parental relationships shape behavioral responses to environmental stressors later in life.

Objective: This study examines the relationship between extreme heat and alcohol consumption among older Americans, emphasizing the moderating role of early-life experiences within a life course framework.

Methods: Using data from individuals aged >50 years in the Health and Retirement Study (1996 - 2018), we analyzed the association between extreme heat exposure (>95 °F) and alcohol consumption, while examining whether early-life experiences such as parental substance abuse, law enforcement encounters, and relationships with fathers moderate this relationship.

Results: Extreme heat exposure was significantly associated with increased alcohol consumption (0.21% per additional extreme heat day, $P<.001$). A positive father-child relationship buffered this effect, while adverse early-life experiences, including law enforcement encounters (0.08%, $P<.001$) and parental substance abuse (0.05%, $P<.001$), exacerbated it.

Conclusions: Given the link between extreme heat and alcohol use in older adults, further longitudinal research and targeted interventions are needed to mitigate associated health risks. Strengthening positive childhood experiences may offer long-term protective effects by fostering resilience and healthier coping mechanisms that persist into later life.

(JMIR Aging 2026;9:e76904) doi:10.2196/76904

KEYWORDS

alcohol use; adverse childhood experiences; ACEs; parental substance abuse; life course perspective framework; psychosocial vulnerability

Introduction

Background

The increasing prevalence of alcohol consumption, particularly high-risk and excessive drinking, among older adults has emerged as a growing public health concern in the United States [1-4]. Between 2001 and 2013, there was a marked rise in alcohol use patterns among US adults aged ≥65 years, characterized by a 22% increase in past-year alcohol consumption, a 65% increase in high-risk drinking, and a staggering 107% increase in alcohol use disorder diagnoses [2]. More recent data reveal that over half of adults in this age group reported alcohol consumption within the past year, with approximately 11% engaging in binge drinking within the past month [5]. These trends are particularly alarming, given the heightened susceptibility of older adults to alcohol-related health complications, owing to age-related physiological changes, coexisting chronic conditions, and potential interactions with medications—factors that collectively elevate risks for unintentional injuries, toxic reactions, and exacerbation of preexisting health conditions [6,7]. As the aging population

XSL•FO
RenderX

continues to grow, so too will the public health implications of alcohol consumption among older adults, underscoring the urgent need for risk mitigation and intervention strategies.

Extreme Heat Exposure and Alcohol Consumption

While previous research has predominantly centered on individual-level determinants of alcohol use among older adults, such as demographic characteristics, socioeconomic status, and health conditions [3-8], a growing body of literature highlights the influence of environmental factors, particularly climate-related stressors such as extreme heat. The hypothesis that climate change–induced extreme heat may exacerbate substance use disorders is gaining empirical support [9]. Scoping reviews suggest that climate change contributes to the worsening of mental health conditions, including heightened levels of anxiety, depression, and psychological distress [10], which may in turn lead to increased substance use as a maladaptive coping mechanism. Notably, empirical evidence from Shenzhen, China, the first study to investigate the relationship between extreme heat and acute excessive drinking, identified a positive correlation between heat exposure and the likelihood of engaging in heavy drinking episodes [11]. Similarly, a nationwide survey in Australia reported that 11% of respondents consumed more alcohol on extremely hot days [11]. Despite these important contributions, prior studies have not specifically targeted older adults nor used nationally representative samples, leaving critical gaps in understanding how heat exposure may influence alcohol consumption patterns in aging populations.

Several physiological and psychological mechanisms may explain the link between high temperatures and increased alcohol use. On a physiological level, alcohol can alleviate thermal discomfort by promoting vasodilation, enhancing perspiration, and facilitating heat dissipation [12,13]. Cold carbonated alcoholic beverages, such as beer, may further provide immediate relief and perceived hydration [14]. Psychologically, alcohol's mild anxiolytic and analgesic properties may offer temporary relief from heat-induced discomfort, irritability, or anxiety [15,16]. However, relying on alcohol as a coping strategy during heatwaves may pose serious health risks, exacerbating dehydration, impairing thermoregulation, and amplifying the risk of heat-related illnesses [11].

Early-Life Experiences as Moderators

Importantly, individual susceptibility to environmental stressors such as extreme heat is not uniform. Psychosocial resources and early-life experiences play a pivotal role in shaping vulnerability and resilience. The existing literature underscores that climate-related health risks are not solely determined by exposure intensity but are also mediated by underlying vulnerability factors (ie, predisposition to adverse outcomes) and resilience capacities [17]. While studies have highlighted personal socioeconomic status, education, income, and social isolation as key determinants of resilience [18], the influence of early-life experiences, both positive childhood experiences (PCEs) and adverse childhood experiences (ACEs), remains underexplored in this context.

PCEs such as parental care, emotional support, and stable family relationships foster resilience and serve as protective factors for mental health in later life [19,20]. For instance, receiving unconditional love and experiencing nurturing parenting contribute to enhanced coping skills, reducing reliance on harmful behaviors such as substance abuse [21]. Conversely, ACEs such as exposure to parental substance misuse, neglect, abuse, or encounters with law enforcement are strongly associated with heightened psychosocial vulnerability and maladaptive coping strategies, including alcohol abuse, in adulthood [22-25]. There is extensive evidence linking ACEs to increased risk of mental health challenges such as depression, psychiatric disorders, and substance misuse later in life [26-29]. Early adverse experiences may thus compound the impact of environmental stressors, amplifying health risks through entrenched behavioral and psychological pathways.

This Study

Grounded in a life course perspective, this study aims to investigate the relationship between extreme heat exposure and alcohol consumption among older adults in the United States, with a focus on how early-life experiences shape risk disparities. Leveraging over 2 decades of nationally representative data integrated with historical temperature records, we provide a comprehensive analysis of this association, addressing a critical gap in the literature. While previous research has predominantly emphasized adulthood socioeconomic and psychosocial factors, our study extends this framework by elucidating how both PCEs and ACEs moderate older adults' behavioral responses to extreme heat. This inquiry is particularly timely given the projected rise in the frequency and severity of heatwaves associated with climate change [30].

The intersection of extreme heat and alcohol consumption is a public health issue of growing urgency, as alcohol use during heatwaves may exacerbate vulnerability to acute health crises, including dehydration, hyperthermia, and heat stroke [31,32]. Older adults, already facing age-related declines in thermoregulatory capacity, are at heightened risk. By identifying how early-life experiences shape susceptibility to maladaptive coping behaviors under environmental stress, our findings offer actionable insights for designing targeted interventions and health policies aimed at mitigating alcohol-related harms in the context of climate change. To our knowledge, this is the first study to use a nationally representative dataset to explore the intersection of extreme heat exposure, substance use, and early-life experiences among older adults in the United States.

Methods

Data and Sample

As individual-level data, this research used data collected from a nationally representative sample of the Health and Retirement Study (HRS), sampling over 20,000 older adults aged ≥50 years, along with their spouses, living in the United States. Since 1992, the HRS has conducted biennial surveys. This comprehensive survey covers diverse information, including demographics, physical health, health-related behaviors, family dynamics, disabilities, housing, wealth, work history, health care

XSL•FO
RenderX

accessibility, and more [33]. This study used the 1996 to 2018 waves of the HRS.

The National Environmental Public Health Tracking Network, a Center for Disease Control and Prevention division, offers yearly county-level data concerning extreme weather incidents. This data source provides historical and projected metrics linked to temperature, heat index, heat-related ailments, and previous and predicted precipitation. To explore the relationship between heat events and the drinking behaviors of older adults, we merged the county-level data on extreme weather incidents with the individual-level data of the HRS using the participant's county of residence and the years in which the survey was conducted.

The study included participants aged ≥50 years with geographical information about their county of residence across time. The detailed sample characteristics are presented in Table 1. The average age of the sample was 68 (SD 11) years. More than half of the sample were women (57%), non-Hispanic White (64%), and coupled (61%). Approximately two-thirds of the respondents completed a high school degree or lower (76%), were unemployed or retired (73%), and owned a primary residence (74%). A total of 65% of the older adults reported good or better health, and <1% reported activity of daily living and instrumental activity of daily living with difficulty. The mean number of household members and living children was 2 (SD 1) and 3 (SD 2), respectively.

XSL•FO
RenderX

Table . Sample characteristics.[a]

Variables	Participants
Outcome variable, mean (SD)	
Drinks per week	0.38 (2.97)
Explanatory variables, mean (SD)	
Extreme heat days ≥95	16.63 (30.43)
Extreme heat days ≥100	7.18 (19.15)
Extreme heat days ≥105	2.44 (9.03)
Moderators: early-life experiences	
Relational quality with father, mean (SD)	4.21 (1.18)
Police encounter (%)	5.28
Parental drink or drug use (%)	16.08
Covariates	
Age (years), mean (SD)	68.19 (11.12)
Gender (%)	
Men	42.70
Women	57.30
Race or ethnicity (%)	
Black	19.58
Hispanic	12.69
Other	3.78
White	63.95
Education (%)	
Less than high school	25.03
High school	50.66
Some college	5.27
Bachelor of Arts	12.03
Graduate	7.01
Marital status (%)	
Coupled	60.68
Separated, divorced, or widowed	34.97
Never married	4.36
Self-reported health (%)	
Poor	10.91
Fair	23.73
Good	31.77
Very good	24.84
Excellent	8.75
ADLs[b], mean (SD)	0.49 (1.13)
IADLs[c], mean (SD)	0.45 (1.11)
Health insurance (%)	
Yes	93.22
No	6.78

XSL•FO
RenderX

Variables	Participants
Employment status (%)	
Employed	27.29
Unemployed	14.74
Retired	57.97
Homeownership (%)	
Yes	73.76
No	26.24
Household income (US $), mean (SD)	62,372 (228,780)
Financial assets (US $), mean (SD)	159,621 (813,402)
People in a household, mean (SD)	2.27 (1.29)
Living children, mean (SD)	3.20 (2.15)
Extremely rainy days, mean (SD)	7.78 (4.87)

[a]Unweighted estimates. 1996 to 2018 waves of the Health and Retirement Study (N=35,261; observations [respondent wave], observations=148,696, N=148,696).

[b]ADL: activities of daily living.

[c]IADL: instruments of activities of daily living.

Variables

The study's dependent variable was older adults' alcohol consumption behaviors measured using the natural log-transformed number of drinks consumed per week. To create the number of drinks taken per week, the study combined two questions asking how many days per week a respondent has had any alcohol to drink and how many drinks she had on the days of drinking in the previous 3 months. Responses to these questions were multiplied. The average number of alcoholic beverages consumed in a week was estimated as <1 (Table 1). The ratio of observations who had at least 1 drink per week was 5%. Among drinkers, the mean number of drinks consumed per week was estimated as 7.85 (SD 11.06).

The primary explanatory variable in the study was instances of extreme heat events taking place in the HRS respondents' county. This factor was proxied by examining the annual frequency of extreme heat days, by counting the days when the highest daily temperatures reached 95 °F, 100 °F, and 105 °F between May and September. On average, there were 17 days per year with peak temperatures of 95 °F, 7 days per year with peak temperatures of 100 °F, and 2 days per year with peak temperatures of 105 °F (Table 1).

The study considered early-life experiences as a moderator of the relationship between extreme heat days and alcohol consumption. The early-life events included paternal relational quality, police encounters, and parental substance use problems before age 18 years, all of which were included in the Leave Behind section of the HRS. Since 2006, 50% of the main survey participants have randomly been assigned to complete the Leave Behind section, and the other half was required to do so in 2008, resulting in respondents' participation in the section in 4-year intervals. The paternal relational quality before age 18 years was assessed using the question, "I had a good relationship with my father before age 18," rated on a 5-point Likert scale [34]. The mean score was 4.21 (SD 1.18), with higher values indicating better relationship quality. As the measure asked about early relational quality, the measure was also averaged across waves to have a time-invariant individual factor. The score was further standardized to have a mean of 0 (SD 1). The other early-life events associated with later-life alcohol consumption were included: "Before you were 18 years old, were you ever in trouble with the police?" "Before you were 18 years old, did either of your parent drink or use drugs so often that it caused problems in the family?" [35]. The percentage of the observations that encountered law enforcement and had parents with drinking and drug use problems was 5.28% and 16.06%, respectively. The two variables were also coded as time-invariant factors, using the mode of the responses.

The study included a comprehensive set of time-variant individual and household characteristics as covariates. These include age centered on 50 years, its squared term, marital status, self-reported health, number of activities of daily living and instrumental activities of daily living with difficulty performing, health insurance ownership, employment status, home ownership, the natural logarithm of household income and financial assets, and number of household members and living children. The study further controlled for individual and wave fixed effects. The study finally controlled for the yearly number of days experiencing intense precipitation, quantified by the annual heavy rainfall exceeding the absolute threshold of 0.01 inches. The study found that, on average, there were 8 such extreme precipitation days per year during the study period.

Empirical Models

To estimate the effect of extreme heat events on alcohol consumption behaviors, this study used individual fixed effects models. By controlling for individual fixed effects, the models were designed to capture within-individual changes in behaviors in response to changes in extreme heat events. The following is the specified model:

XSL•FO
RenderX

(1)ACiw=α0+α1EHct+α2Xiw+ii+ww+ iw

AC_{iw} denotes a respondent *i*'s alcohol consumption behaviors measured using the natural logarithm of the number of drinks consumed per week in wave *w*. Coefficients from this model can be interpreted as percentage changes in weekly alcohol consumption associated with a 1-unit increase in the number of extreme heat days in a given year. For example, a coefficient of 0.01 indicates that each additional extreme heat day is associated with a 1% increase in the number of drinks consumed per week. EH_{ct} is the measure for extreme heat events proxied for using the annual counts of days with daily maximum temperatures of 95 °F, 100 °F, and 105 °F in county *c* in year *t*. X_{iw} indicates a vector of time-variant individual and household characteristics measured in wave *w*. i_i and w_w are individual and wave fixed effects. i_i reflects all the time-invariant individual *i*'s factors such as gender, race or ethnicity, educational attainment, etc. $\in_{iw}$ is the idiosyncratic error for individual *i* in wave *w*.

The study further explores whether early-life experiences, known to be correlated with alcohol consumption and alcoholism, moderated the association between extreme heat events and alcohol consumption behaviors. The following is the individual fixed effects model after interacting the count of extreme heat days with early-life experience factors:

(2)ACiw=β0+β1EHct+β2EHctLEi+β3Xiw+ii+ww+ iw

LE_i denotes the early–life course events of respondent *i*. The early-life experiences included relational quality with a father, police encounters, and parental substance use problems before the age of 18 years. All the life course events were time invariant because the events occurred early in life. In this case, the standalone effects of life course events are absorbed in the individual effects i_i, and the combined effect therefore would be measured as $\beta_1+\beta_2$.

Ethical Considerations

This study was approved by the institutional review board (IRB) at the University of Utah (IRB Protocol #00194769). All respondents in the HRS provided informed consent prior to participation. The analyses used restricted HRS data containing county-of-residence information, which were accessed under a Restricted Data Use Agreement and analyzed in accordance with applicable data security and confidentiality requirements. No direct participant compensation was provided by the authors because this study relied on the secondary analysis of existing survey data.

Results

Extreme Heat Events on Alcohol Consumption Behaviors

Table 2 presents the estimates from individual fixed effects models where the dependent variable was the logarithm of the number of drinks per week. Regardless of the thresholds used to define extreme heat temperatures, the number of extreme heat days was positively correlated with the number of alcoholic beverages consumed per week. A 1-unit increase in the extreme heat day counts defined as the maximum temperatures of 95 °F, 100 °F, and 105 °F related to increases in the number of drinks per week by 0.21%, 0.02%, and 0.04%, respectively. Interpreting the coefficients in more meaningful units, a seasonal increase of 30 extreme heat days (rather than a single day) is associated with a 6.3% increase in weekly alcohol consumption when extreme heat is defined as temperatures ≥95 °F, a 0.6% increase when defined as ≥100 °F, and a 1.2% increase when defined as ≥105 °F.

Table . The effect of extreme heat on alcohol consumption.[a]

	Dependent variable (log of number of alcoholic drinks/week)		
	Coefficient B (SE)[b]	Coefficient B (SE)[c]	Coefficient B (SE)[d]
Extreme heat days			
≥95	0.0021[e] (0.0001)	—[f]	—
≥100	—	0.0002[g] (0.0001)	—
≥105	—	—	0.0004[h] (0.0001)

[a]Unweighted individual fixed effects model estimators. 1998 to 2018 waves of the Health and Retirement Study (N=35,261; observations, N=148,696). The covariates include all the factors listed in the *Methods* section.

[b]R^2=0.0531.

[c]R^2=0.0530.

[d]R^2=0.0531.

[e]*P*<.001.

[f]Not applicable.

[g]*P*<.01.

[h]*P*<.05.

As a sensitivity test, we restricted the sample to respondents who participated in the survey between May and September of each year and reestimated the individual fixed effects models. This procedure addresses the potential concern that the timing of drinking behavior may not align with the period during which extreme heat events are measured. As noted in the *Methods*

XSL•FO
RenderX

section, extreme heat is defined using temperature observations from May till September, whereas alcohol consumption is measured throughout the year based on the respondent's survey date. If the effect of extreme heat is short-lived and primarily influences contemporaneous or near-term drinking behavior—rather than exerting long-lasting effects over the entire year—then restricting the sample to the survey period overlapping with the heat exposure window is appropriate.

The results of this analysis are presented in Table S1 in Multimedia Appendix 1. The estimates are consistent with those from the main specification, although the significance level is reduced to $P<.05$ for models in which extreme heat days are defined as ≥95 °F and ≥100 °F. When extreme heat days are defined as ≥105 °F, the coefficient becomes statistically insignificant. Importantly, the direction and magnitude of the estimated effects remain similar, suggesting that the primary findings are not driven by the inclusion of out-of-season observations. For this reason, we retain the full analytic sample in the main analysis to preserve variation in within-person drinking behavior and maintain statistical power, while presenting the seasonal restriction as a robustness check.

Early-Life Experiences as a Moderator

From this section onward, the study used a threshold of 95 °F to define the extreme heat days. Table 3 presents the estimates from individual fixed effects models after interacting extreme heat days with early-life experiences (equation 2). All the early-life experiences appeared to moderate the association between extreme heat events and the number of drinks consumed per week. Specifically, early-life relational quality with a father could offset the adverse effect of extreme heat days on alcohol consumption (panel A in Table 3). A 30-day increase in extreme heat days is associated with approximately a 0.6% increase in weekly alcohol consumption among older adults with average levels of paternal relational quality (0.0002 × 30 days = 0.006 or 0.6%). However, for older adults whose paternal relational quality is 1 SD above the mean, this increase is effectively offset. In other words, stronger early-life relationships with fathers buffer the effect of heat exposure on later-life alcohol use. Furthermore, older adults without law enforcement encounters did not change their alcohol consumption in response to changes in extreme heat days (panel B in Table 3). In contrast, among those who experienced adverse police encounters before the age of 18 years, the same seasonal increase in extreme heat is associated with a 2.4% increase in weekly consumption (0.0008×30 days=0.024, or 2.4%). Finally, older adults who did not grow up with parents who had drinking or drug problems show no meaningful change in alcohol consumption in response to additional extreme heat. In contrast, those who did experience parental alcohol or drug problems exhibit a 1.5% increase in weekly alcohol consumption for each 30 additional extreme heat days (0.0005×30 days=0.015, or 1.5%; panel C in Table 3).

Table . Life course events for alcohol consumption as a moderator.[a]

	Dependent variable (log of number of alcoholic drinks/week), coefficient B (SE)
Panel A: relational quality with a father before the age of 18 years (N=16,018; observations, N=91,701)[b]	
Extreme heat days	0.0002[c] (0.0001)
Heat × relational quality with a father	−0.0002[c] (0.0001)
Panel B: police encounter before the age of 18 years (N=13,949; observations, N=47,364)[d]	
Extreme heat days	0.0000 (0.0001)
Heat × police encounter	0.0008[e] (0.0002)
Panel C: parental drinking or drug problems before the age of 18 years (N=16,081; observations, N=64,291)[f]	
Extreme heat days	−0.0000 (0.0001)
Heat × parental drinking or drug problem	0.0005[e] (0.0001)

[a]Unweighted individual fixed effects model estimators. 1998 to 2018 waves of the Health and Retirement Study. The covariates include all the factors listed in the *Methods* section.

[b]R^2=0.0570.

[c]$P<.01$.

[d]R^2=0.0731.

[e]$P<.001$.

[f]R^2=0.0565.

Discussion

Principal Findings

This study offers novel, population-level insights into the intersection of environmental stressors and behavioral health, specifically examining how extreme heat exposure is associated with alcohol consumption among older adults, with particular attention to the moderating role of early-life experiences. To our knowledge, this is the first analysis within the United States to explore this relationship through a life course framework,

XSL•FO
RenderX

highlighting the enduring impact of early ACEs or PCEs on later-life health behaviors.

While some studies have reported potential health benefits of light to moderate alcohol use, such as enhanced cardiovascular function and delayed cognitive decline [36,37], the deleterious consequences of excessive or binge drinking are well established. These include elevated risks of falls, cardiovascular disease, cognitive impairment, and mortality [38], risks that may be exacerbated under extreme heat conditions. Our findings demonstrate a significant association between heat exposure and increased alcohol consumption, aligning with earlier evidence from non-US contexts [39]. However, unlike prior studies, which lacked representative samples and focused on younger populations, our analysis centers on older adults, thereby filling a critical gap in understanding how environmental stress may drive maladaptive coping behaviors in aging populations.

More specifically, our findings indicate that a seasonal increase of 30 extreme heat days is associated with a 1.2% (≥105 °F) to 6.3% (≥95 °F) increase in weekly alcohol consumption. Studies have shown that alcohol consumption during heat exposure can physiologically lower core body temperature by enhancing sweating and vasodilation, which are key mechanisms of thermoregulation [13]. In addition, cold alcoholic beverages may provide immediate psychological relief and a sense of hydration, making them particularly appealing in hot weather [13]. This perceived cooling effect may encourage greater alcohol intake during high temperatures. However, given alcohol's potential to exacerbate dehydration, impair thermoregulation, and increase vulnerability to heat-related illness [11], it is not a healthy coping strategy during extreme heat. Evidence also shows that higher temperatures are associated with increased hospital visits for alcohol-related disorders [40]. This is especially salient for older adults, as even small behavioral shifts may have clinically meaningful consequences for those already vulnerable to impaired thermoregulation and hydration with aging.

Given the significant health risks associated with excessive alcohol consumption, our findings underscore the need to prioritize populations most vulnerable to both extreme heat and substance use. Targeted interventions for older adults exposed to prolonged heat should consider the compounding effect of environmental stress on health behaviors. Possible strategies include providing subsidies for home retrofitting (eg, air conditioning installation), disaster relief funds, and establishing accessible cooling centers such as community centers, public pools, and shelters. Such measures may not only alleviate heat-induced physiological stress but also reduce the likelihood of maladaptive coping mechanisms such as increased alcohol consumption.

Beyond immediate environmental interventions, this study underscores the enduring influence of early-life experiences on later-life behavioral responses to stress. Specifically, we found evidence for the protective role of PCEs, nurturing, supportive, and emotionally secure environments during early life that contribute to resilience and healthier aging trajectories. In particular, a strong paternal relationship during childhood mitigated the association between extreme heat exposure and alcohol consumption in older adulthood, suggesting that early supportive relationships may foster adaptive coping mechanisms that persist across the life course. Prior research has demonstrated that nurturing paternal relationships buffer against alcohol misuse [21] and enhance stress resilience [20]. Our findings extend this literature by suggesting that the psychological advantages conferred by early positive parenting not only shape interpersonal and emotional health but also influence behavioral responses to environmental stressors later in life. This underscores the value of policies and programs promoting positive parenting practices and mental health education aimed at fostering resilience from childhood through adulthood.

However, the nonsignificant findings for maternal relationships should not be interpreted as an absence of influence (Table S2 in Multimedia Appendix 1). This may reflect limited measurement sensitivity, recall variability, or contextual differences in how maternal warmth and caregiving were reported. Future research should use multidimensional assessments of PCEs that capture diverse caregiving roles and family contexts. Such approaches would better illuminate how early relational environments, across both paternal and maternal domains, can buffer environmental stressors and promote adaptive aging trajectories.

Conversely, the presence of ACEs, including exposure to abuse, neglect, or household dysfunction, was associated with higher alcohol consumption during extreme heat events. This pattern suggests that early-life adversity, particularly exposure to parental substance use, may contribute to enduring susceptibilities in stress regulation and coping behaviors that persist into older adulthood. This finding aligns with previous research highlighting the heightened vulnerability of individuals exposed to ACEs [41]. The influence of parental substance misuse is twofold: it both serves as a behavioral model for maladaptive coping strategies and introduces potential genetic predispositions for substance use disorders [42,43]. These factors collectively increase susceptibility to problematic drinking behaviors in response to stressors such as extreme heat. Addressing these risks necessitates a multifaceted approach—incorporating trauma-informed care, family support programs, and community-based interventions aimed at breaking the intergenerational transmission of substance abuse and fostering healthier coping mechanisms.

Additionally, our findings reveal that negative childhood interactions with law enforcement were associated with increased alcohol consumption in adulthood during heatwaves. Such early encounters often contribute to heightened stress reactivity and may normalize maladaptive coping behaviors within disadvantaged environments [29,44]. This underscores the critical need for trauma-informed policing practices and early intervention programs that mitigate the long-term psychological impacts of ACEs. Integrating trauma-awareness training for law enforcement and expanding mental health services for affected communities could serve as pivotal steps in reducing long-term health disparities.

XSL•FO
RenderX

Limitations

While this study advances the understanding of the interplay between extreme heat exposure, alcohol consumption, and early-life experiences, certain limitations warrant consideration. The use of HRS data, although robust, restricts our ability to ascertain the precise mechanisms underlying the observed associations. Future research should explore the mediating psychological or physiological pathways in more depth. Additionally, our analysis of PCEs was limited to paternal relationships, as maternal relationships did not yield statistically significant findings in relation to heat exposure and alcohol consumption (Table S2 in Multimedia Appendix 1), highlighting an avenue for further inquiry. An important caveat is that retrospective questions on early-life parental relationships were only administered to respondents who reported having a mother or father present during childhood. Individuals who lacked such parental figures did not receive these items, resulting in nonrandom missingness for that measure. Finally, the absence of detailed environmental variables (eg, home cooling resources and neighborhood infrastructure) represents a limitation, as these factors likely influence the degree of heat-related stress experienced by older adults and shape how they buffer heat-related stress. Future work could incorporate proxy information from ancillary HRS modules such as electricity spending from the Consumption and Activities Mail Survey or community facility measures from the Neighborhood and Network Data to examine potential moderating effects. Although these proxies do not directly capture older adults' adaptation to extreme heat and linking these subpopulation modules substantially reduces sample size and statistical power, incorporating such environmental factors in future studies will enrich our understanding of the multi-level determinants of health behavior under climate-related stressors.

Conclusions

This study identifies a clear association between extreme heat events and increased alcohol consumption among older adults, with early-life experiences serving as significant moderators. Positive paternal relationships appear to buffer against the adverse behavioral impacts of heat stress, while adverse experiences such as parental substance use and negative interactions with law enforcement exacerbate vulnerability. These findings emphasize the long-lasting influence of childhood environments on later-life resilience and underscore the need for integrated, trauma-informed policies and interventions. By addressing both environmental stressors and the psychosocial foundations established early in life, policymakers and practitioners can more effectively mitigate the compounding health risks facing older populations in an era of increasing climate volatility.

Acknowledgments

This research was conducted using restricted data files from the Health and Retirement Study obtained under a data use agreement with the University of Michigan. The opinions and conclusions expressed herein are those of the authors and do not represent the views of the Health and Retirement Study, the University of Michigan, the National Institute on Aging, the US Census Bureau, or any other funding agency. All results were reviewed to avoid disclosure of confidential information.

Funding

The authors declared no financial support was received for this work.

Data Availability

The data used in this study are publicly available through the Health and Retirement Study, sponsored by the National Institute on Aging (grant U01AG009740) and conducted by the University of Michigan. Access to restricted Health and Retirement Study geographic data were provided through a special data agreement with the University of Michigan. These data include sensitive geographic identifiers and were accessed under strict confidentiality protocols.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary sensitivity and moderation analyses examining the association between extreme heat exposure and alcohol consumption among older adults.
[DOCX File, 15 KB - aging_v9i1e76904_app1.docx]

References

1. Breslow RA, Castle IJP, Chen CM, Graubard BI. Trends in alcohol consumption among older Americans: national health interview surveys, 1997 to 2014. Alcohol Clin Exp Res 2017 May;41(5):976-986. [doi: 10.1111/acer.13365] [Medline: 28340502]
2. Grant BF, Chou SP, Saha TD, et al. Prevalence of 12-month alcohol use, high-risk drinking, and DSM-IV alcohol use disorder in the United States, 2001-2002 to 2012-2013: results from the National Epidemiologic Survey on alcohol and

XSL•FO
RenderX

related conditions. JAMA Psychiatry 2017 Sep 1;74(9):911-923. [doi: 10.1001/jamapsychiatry.2017.2161] [Medline: 28793133]
3. Han BH, Moore AA, Sherman S, Keyes KM, Palamar JJ. Demographic trends of binge alcohol use and alcohol use disorders among older adults in the United States, 2005-2014. Drug Alcohol Depend 2017 Jan 1;170:198-207. [doi: 10.1016/j.drugalcdep.2016.11.003] [Medline: 27979428]
4. Han BH, Moore AA, Ferris R, Palamar JJ. Binge drinking among older adults in the United States, 2015 to 2017. J Am Geriatr Soc 2019 Oct;67(10):2139-2144. [doi: 10.1111/jgs.16071] [Medline: 31364159]
5. 2021 National Survey on Drug Use and Health: Section 5 PE Tables. Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality (CBHSQ). 2022. URL: https://www.samhsa.gov/data/sites/default/files/reports/rpt39441/NSDUHDetailedTabs2021/NSDUHDetailedTabs2021/NSDUHDetTabsSect5pe2021.htm [accessed 2024-05-18]
6. Breslow RA, Dong C, White A. Prevalence of alcohol-interactive prescription medication use among current drinkers: United States, 1999 to 2010. Alcohol Clin Exp Res 2015 Feb;39(2):371-379. [doi: 10.1111/acer.12633] [Medline: 25597432]
7. Moore AA, Whiteman EJ, Ward KT. Risks of combined alcohol/medication use in older adults. Am J Geriatr Pharmacother 2007 Mar;5(1):64-74. [doi: 10.1016/j.amjopharm.2007.03.006] [Medline: 17608249]
8. Rao R, Schofield P, Ashworth M. Alcohol use, socioeconomic deprivation and ethnicity in older people. BMJ Open 2015 Aug 24;5(8):e007525. [doi: 10.1136/bmjopen-2014-007525] [Medline: 26303334]
9. Vergunst F, Berry HL, Minor K, Chadi N. Climate change and substance-use behaviors: a risk-pathways framework. Perspect Psychol Sci 2023 Jul;18(4):936-954. [doi: 10.1177/17456916221132739] [Medline: 36441663]
10. Middleton J, Cunsolo A, Jones-Bitton A, Wright CJ, Harper SL. Indigenous mental health in a changing climate: a systematic scoping review of the global literature. Environ Res Lett 2020 May 1;15(5):053001. [doi: 10.1088/1748-9326/ab68a9]
11. Zander KK, Mathew S, Carter S. Behavioural (mal)adaptation to extreme heat in Australia: implications for health and wellbeing. Urban Climate 2024 Jan;53:101772. [doi: 10.1016/j.uclim.2023.101772]
12. Wasielewski JA, Holloway FA. Alcohol's interactions with circadian rhythms. a focus on body temperature. Alcohol Res Health 2001;25(2):94-100 [FREE Full text] [Medline: 11584555]
13. Yoda T, Crawshaw LI, Nakamura M, et al. Effects of alcohol on thermoregulation during mild heat exposure in humans. Alcohol 2005 Jul;36(3):195-200. [doi: 10.1016/j.alcohol.2005.09.002] [Medline: 16377461]
14. Peyrot des Gachons C, Avrillier J, Gleason M, et al. Oral cooling and carbonation increase the perception of drinking and thirst quenching in thirsty adults. PLoS ONE 2016;11(9):e0162261. [doi: 10.1371/journal.pone.0162261] [Medline: 27685093]
15. Baum-Baicker C. The psychological benefits of moderate alcohol consumption: a review of the literature. Drug Alcohol Depend 1985 Aug;15(4):305-322. [doi: 10.1016/0376-8716(85)90008-0] [Medline: 4053968]
16. Rony MK, Alamgir HM. High temperatures on mental health: recognizing the association and the need for proactive strategies-a perspective. Health Sci Rep 2023 Dec;6(12):e1729. [doi: 10.1002/hsr2.1729] [Medline: 38059052]
17. Few R. Health and climatic hazards: framing social research on vulnerability, response and adaptation. Glob Environ Change 2007 May;17(2):281-295. [doi: 10.1016/j.gloenvcha.2006.11.001]
18. Li A, Toll M, Bentley R. Mapping social vulnerability indicators to understand the health impacts of climate change: a scoping review. Lancet Planet Health 2023 Nov;7(11):e925-e937. [doi: 10.1016/S2542-5196(23)00216-4] [Medline: 37940212]
19. Bethell C, Jones J, Gombojav N, Linkenbach J, Sege R. Positive childhood experiences and adult mental and relational health in a statewide sample: associations across adverse childhood experiences levels. JAMA Pediatr 2019 Nov 1;173(11):e193007. [doi: 10.1001/jamapediatrics.2019.3007] [Medline: 31498386]
20. Morris AS, Hays-Grudo J. Protective and compensatory childhood experiences and their impact on adult mental health. World Psychiatry 2023 Feb;22(1):150-151. [doi: 10.1002/wps.21042] [Medline: 36640391]
21. Jacob T, Johnson S. Parenting influences on the development of alcohol abuse and dependence. Alcohol Health Res World 1997;21(3):204-209 [FREE Full text] [Medline: 15706770]
22. Fish JN, Russell BS, Watson RJ, Russell ST. Parent-child relationships and sexual minority youth: implications for adult alcohol abuse. J Youth Adolesc 2020 Oct;49(10):2034-2046. [doi: 10.1007/s10964-020-01299-7] [Medline: 32772330]
23. Widom CS, White HR, Czaja SJ, Marmorstein NR. Long-term effects of child abuse and neglect on alcohol use and excessive drinking in middle adulthood. J Stud Alcohol Drugs 2007 May;68(3):317-326. [doi: 10.15288/jsad.2007.68.317] [Medline: 17446970]
24. Lee RD, Chen J. Adverse childhood experiences, mental health, and excessive alcohol use: examination of race/ethnicity and sex differences. Child Abuse Negl 2017 Jul;69:40-48. [doi: 10.1016/j.chiabu.2017.04.004] [Medline: 28448813]
25. Schalkwijk F, Van Someren EJ, Nicolai NJ, Uijttewaal JL, Wassing R. From childhood trauma to hyperarousal in adults: the mediating effect of maladaptive shame coping and insomnia. Front Hum Neurosci 2023;17:990581. [doi: 10.3389/fnhum.2023.990581] [Medline: 36875235]
26. Gondek D, Patalay P, Lacey RE. Adverse childhood experiences and multiple mental health outcomes through adulthood: a prospective birth cohort study. SSM Mental Health 2021 Dec;1:100013. [doi: 10.1016/j.ssmmh.2021.100013]

XSL•FO
RenderX

27. Shin SH, Wright C, Johnston S. Early life experiences moderate the relationship between genetic risk of autism and current and lifetime mental health. Front Psychiatry 2021;12:772841. [doi: 10.3389/fpsyt.2021.772841] [Medline: 34916975]
28. Anda RF, Whitfield CL, Felitti VJ, et al. Adverse childhood experiences, alcoholic parents, and later risk of alcoholism and depression. Psychiatr Serv 2002 Aug;53(8):1001-1009. [doi: 10.1176/appi.ps.53.8.1001] [Medline: 12161676]
29. Dube SR, Anda RF, Felitti VJ, Edwards VJ, Croft JB. Adverse childhood experiences and personal alcohol abuse as an adult. Addict Behav 2002;27(5):713-725. [doi: 10.1016/s0306-4603(01)00204-0] [Medline: 12201379]
30. Ebi KL, Capon A, Berry P, et al. Hot weather and heat extremes: health risks. Lancet 2021 Aug 21;398(10301):698-708. [doi: 10.1016/S0140-6736(21)01208-3] [Medline: 34419205]
31. Luber G, McGeehin M. Climate change and extreme heat events. Am J Prev Med 2008 Nov;35(5):429-435. [doi: 10.1016/j.amepre.2008.08.021] [Medline: 18929969]
32. Page LA, Hajat S, Kovats RS, Howard LM. Temperature-related deaths in people with psychosis, dementia and substance misuse. Br J Psychiatry 2012 Jun;200(6):485-490. [doi: 10.1192/bjp.bp.111.100404] [Medline: 22661680]
33. Fisher GG, Ryan LH. Overview of the health and retirement study and introduction to the special issue. Work Aging Retire 2018 Jan;4(1):1-9. [doi: 10.1093/workar/wax032] [Medline: 29423243]
34. Rossi AS, editor. Caring and Doing for Others: Social Responsibility in the Domains of Family, Work, and Community: University of Chicago Press; 2001. URL: https://archive.org/details/caringdoingforot0000unse [accessed 2025-05-25]
35. Yang MS, Hedeker D. A life-span approach to examining older vulnerable population's subjective well-being: the role of adversity and trauma. Aging Ment Health 2020 Dec;24(12):2043-2052. [doi: 10.1080/13607863.2019.1652245] [Medline: 31402675]
36. O'Keefe JH, Bhatti SK, Bajwa A, DiNicolantonio JJ, Lavie CJ. Alcohol and cardiovascular health: the dose makes the poison…or the remedy. Mayo Clin Proc 2014 Mar;89(3):382-393. [doi: 10.1016/j.mayocp.2013.11.005] [Medline: 24582196]
37. Stampfer MJ, Kang JH, Chen J, Cherry R, Grodstein F. Effects of moderate alcohol consumption on cognitive function in women. N Engl J Med 2005 Jan 20;352(3):245-253. [doi: 10.1056/NEJMoa041152] [Medline: 15659724]
38. Liu AB, Lin YX, Li GY, et al. Associations of frailty and cognitive impairment with all-cause and cardiovascular mortality in older adults: a prospective cohort study from NHANES 2011-2014. BMC Geriatr 2025 Feb 22;25(1):124. [doi: 10.1186/s12877-025-05752-9] [Medline: 39987017]
39. Liu X, Wen Y, Zhang K, et al. Examining the association between apparent temperature and incidence of acute excessive drinking in Shenzhen, China. Sci Total Environ 2020 Nov;741:140302. [doi: 10.1016/j.scitotenv.2020.140302]
40. Parks RM, Rowland ST, Do V, et al. The association between temperature and alcohol- and substance-related disorder hospital visits in New York State. Commun Med (Lond) 2023 Sep 26;3(1):118. [doi: 10.1038/s43856-023-00346-1] [Medline: 37752306]
41. Kobayashi LC. Adverse childhood experiences and later-life cognitive aging: persistent methodological challenges limit the evidence base. J Gerontol A Biol Sci Med Sci 2024 Aug 1;79(8):glae138. [doi: 10.1093/gerona/glae138] [Medline: 38920298]
42. Goldman D, Oroszi G, Ducci F. The genetics of addictions: uncovering the genes. Nat Rev Genet 2005 Jul;6(7):521-532. [doi: 10.1038/nrg1635] [Medline: 15995696]
43. Santiago CD, Etter EM, Wadsworth ME, Raviv T. Predictors of responses to stress among families coping with poverty-related stress. Anxiety Stress Coping 2012 May;25(3):239-258. [doi: 10.1080/10615806.2011.583347] [Medline: 21614698]
44. Kerig PK, Becker SP. Trauma and girls' delinquency. In: Leve LD, Chamberlain P, editors. Girls' Antisocial Behavior and Relationships: A Developmental Perspective: Springer; 2012:143-160. [doi: 10.1007/978-1-4614-0415-6_8]

Abbreviations

ACE: adverse childhood experience
HRS: Health and Retirement Study
IRB: institutional review board
PCE: positive childhood experience

Edited by G Seçkin; submitted 03.May.2025; peer-reviewed by D Ajayi, X Ding; revised version received 25.Dec.2025; accepted 13.Feb.2026; published 10.Jun.2026.

Please cite as:
Lee HY, Shin SH, Choi YJ, Choi EY, Ji H
Early-Life Experiences as Moderators of the Relationship Between Extreme Heat and Alcohol Consumption among Older Adults: Quantitative Study
JMIR Aging 2026;9:e76904
URL: https://aging.jmir.org/2026/1/e76904
doi:10.2196/76904

XSL•FO
RenderX

© Hee Yun Lee, Su Hyun Shin, Yeon Jin Choi, Eun Young Choi, Hyunjung Ji. Originally published in JMIR Aging (https://aging.jmir.org), 10.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Global Trends and Research Hot Spots in Medication Regimen Simplification: Bibliometric Analysis

Chuwen Dou[1,2], MSN; Bei Wu[2], PhD; Leyi Wu[1], MSN; An Gu[1], MSN; An Huang[1], MSN; Chen Zhang[3], BA; Xichenhui Qiu[4], PhD; Jianlin Lou[1], PhD; Lina Wang[1], PhD

[1]Huzhou Key Laboratory of Precise Prevention and Control of Major Chronic Diseases, School of Medicine, Huzhou Normal University, No. 168 Wuxing Road, Wuxing District, Huzhou, Zhejiang, China

[2]Faculty of Arts and Science, New York University Shanghai, Shanghai, China

[3]Department of General Medicine, Community Health Service Center of Renhuangshan, Huzhou, Huzhou, China

[4]Health Science Center, Shenzhen University, Shenzhen, China

Corresponding Author:
Lina Wang, PhD
Huzhou Key Laboratory of Precise Prevention and Control of Major Chronic Diseases, School of Medicine, Huzhou Normal University, No. 168 Wuxing Road, Wuxing District, Huzhou, Zhejiang, China

Abstract

Background: Medication regimen simplification has gained increasing attention as a strategy to reduce treatment burden and improve medication use. However, the overall development, knowledge structure, and emerging themes of this field have not been systematically mapped, hindering efforts to identify clear research priorities and support strategies that facilitate the translation of simplified approaches into optimized medication management.

Objective: This study aimed to examine global research trends, collaboration patterns, and research hot spots in medication regimen simplification.

Methods: A bibliometric analysis was conducted to characterize the research landscape of medication regimen simplification. Publications from the Web of Science Core Collection (January 1, 2005, to December 31, 2025) were analyzed using Microsoft Excel 365, VOSviewer, CiteSpace, and Bibliometrix in R for publication trends, countries and regions, institutions, authors, cocited references, keyword co-occurrence, clustering, and thematic evolution.

Results: A total of 468 publications were included in this bibliometric analysis. Annual publication output showed an upward trend, although fluctuations occurred across years. The United States ranked first in publication output and total citations, and Monash University was the most productive institution. Collaboration was concentrated in a limited number of country, institutional, and author clusters, with the strongest international links between North America and Europe. Cocitation and keyword analyses showed that the field evolved from a focus on HIV-related regimen optimization and medication adherence to broader themes in chronic disease management, including type 2 diabetes, heart failure, long-term care, and medication regimen complexity.

Conclusions: Research on medication regimen simplification has grown over time, but collaboration remains relatively concentrated. Current hot spots center on its use in specific disease settings and on medication complexity and adherence, whereas recent trends highlight real-world implementation. Future research should further examine the long-term impact and sustainability of simplification strategies.

(JMIR Aging 2026;9:e82274) doi:10.2196/82274

KEYWORDS

medication regimen simplification; bibliometric analysis; polypharmacy; medication adherence; medication regimen complexity

Introduction

With advances in medical care, more patients with chronic conditions such as hypertension, diabetes, and cardiovascular disease require multiple medications for long-term pharmacotherapy. However, the benefits of treatment are frequently limited by poor medication adherence, and complex drug regimens are an important contributing factor [1,2]. More complex medication regimens are associated with poorer adherence, which may compromise disease control and increase the risk of adverse outcomes [3]. Older adults, particularly those with cognitive impairment or limited health literacy, often experience difficulties with medication management, which may contribute to unintentional nonadherence, medication errors, and reduced quality of life [4-6].

Complex medication regimens in older adults, particularly those involving multiple drugs and frequent dosing, can make adherence more difficult [2]. Polypharmacy may also increase

XSL•FO
RenderX

the risk of drug-drug interactions and health care costs [7]. As a result, medication regimen simplification has become a key strategy to optimize pharmacotherapy, improve adherence, and reduce treatment burden [8,9]. Its importance is further underscored by the World Health Organization's advocacy for the rational use of medicines, ensuring that patients receive medications appropriate to their clinical needs and in correct doses for effective outcomes [10]. Similarly, a Centers for Disease Control and Prevention Grand Rounds report highlighted improved medication adherence as a key component of chronic disease management and optimal prescribing practices [11]. In clinical practice, regimen simplification may be considered as part of broader medication review and optimization processes [12].

Medication regimen simplification refers to reducing the complexity of medication use through strategies such as consolidating dosing times, reducing dosing frequency, standardizing administration routes, substituting long-acting for short-acting formulations, and using fixed-dose or other combination products when possible without altering the therapeutic goals, efficacy, or safety of treatment [8]. Importantly, medication regimen complexity is not determined solely by the number of medicines prescribed but also by dosage forms, dosing frequency, administration times, routes of administration, and additional instructions required for proper use [13-15]. Accordingly, medication regimen simplification aims to reduce the practical and cognitive burden of medication use for patients or caregivers [8]. Its application spans a wide range of clinical contexts, including polypharmacy in older adults; chronic disease management; HIV care; and other settings characterized by high regimen complexity, such as long-term care, home-based care, and multi-morbidity [16,17]. This concept is distinct from deprescribing, which primarily involves discontinuing inappropriate or unnecessary medications, whereas medication regimen simplification seeks to make an otherwise appropriate regimen easier to follow without changing its therapeutic intent [16,18]. It should also be distinguished from broader concepts such as medication optimization, medication review, rational prescribing, and adherence-enhancing interventions [19-21]. In this sense, simplification may be regarded as one component of medication optimization and one possible outcome of medication review, but it is not synonymous with either. In this study, reductions in medication regimen complexity index scores or simplification of medication administration processes were considered relevant only when they reflected a deliberate effort to simplify the medication regimen actually used by patients or caregivers [22,23].

Despite growing interest in this area, related research remains dispersed across populations, disease settings, and intervention types, and the overall intellectual structure and development of the field have not been systematically characterized. Bibliometric analysis offers a systematic and visual approach to map the structure, trends, and knowledge gaps of a research field [24,25]. This study used VOSviewer (Centre for Science and Technology Studies, Leiden University) and CiteSpace to analyze the global literature on medication regimen simplification and identify key contributors, major research themes, and emerging trends with the aim of informing future research and clinical practice.

Methods

Search Strategy and Eligibility Criteria

Publications were retrieved from the Web of Science Core Collection (WoSCC). The search was conducted on March 6, 2026, and covered the period from January 1, 2005, to December 31, 2025. All records were obtained from the Science Citation Index Expanded within WoSCC. The search was performed in the title, author keywords, and abstract fields. Only articles and reviews published in English were included. Because "medication regimen simplification" is neither a standardized indexing term nor a MeSH (Medical Subject Headings) term, the search strategy relied on a combination of free-text terms informed by related expressions reported in previous studies and by the conceptual scope of this study [8,9,26]. These terms included simplification-related expressions for regimens, medications, and therapies, as well as expressions related to reduced medication regimen complexity. The full search strategy is shown in Table 1.

XSL•FO
RenderX

Table . Summary of data source and study selection.

Item	Description
Data source	WoSCC[a]
Search date	March 6, 2026
Time span	2005-2025
Language	English
Document type	Articles and reviews
Search field	Title, author keywords, and abstract
Search strategy	TI/AK/AB = (“regimen simplification” OR “treatment simplification” OR “therapy simplification” OR “medication regimen simplification” OR “simplified regimen” OR “simplified medication regimen” OR (simplif* NEAR/3 regimen*) OR (simplif* NEAR/3 medication*) OR (simplif* NEAR/3 therap*) OR (“medication regimen complexity” NEAR/3 reduc*) OR (reduce* NEAR/3 “medication regimen complexity”))
Initial records retrieved, n	1466
Screening criteria	Records were screened by title and abstract. Publications were excluded if they did not describe medication regimen simplification as a primary focus or if any reported reduction in medication regimen complexity was not presented as an intended simplification strategy.
Final records included, n	468

[a]WoSCC: Web of Science Core Collection.

In this study, medication regimen simplification was operationally defined as strategies intended to reduce the complexity of a medication regimen while preserving its therapeutic intent. On the basis of this definition, this bibliometric analysis included publications that explicitly described such strategies as medication regimen simplification or clearly characterized them as a reduction in medication regimen complexity. Interventions such as reduced dosing frequency, consolidated administration times, long-acting formulations, and fixed-dose combinations were included only when they were presented as intended to simplify the regimen or reduce its complexity. Deprescribing was not considered equivalent to medication regimen simplification. Similarly, lower medication regimen complexity scores or simpler medication administration schedules were considered relevant only when they reflected an explicit strategy to simplify the regimen followed by patients or caregivers.

Deduplication and Screening Reliability

Two researchers independently screened the titles and abstracts of all retrieved records. Full texts of potentially eligible studies were then reviewed independently by the same 2 researchers according to the predefined inclusion and exclusion criteria. Any disagreements were resolved through discussion with a third researcher. Manual checking was performed to identify any remaining duplicate records or overlapping reports of the same study. In addition, the reference lists of included studies were screened manually to identify potentially relevant articles. Interrater agreement between the 2 researchers was assessed using the Cohen κ, which indicated a high level of agreement (κ=0.83).

Data Extraction and Collection

Bibliometric data were extracted from the included publications, including publication year, country or region, institution, author, journal, keywords, cited references, and research area information, as well as annual publication counts. Before analysis, limited manual corrections were performed to improve the consistency of country, institution, and author information. Institution names were merged when spelling variants, abbreviations, or formatting differences referred to the same organization. Country names were standardized when the WoSCC classification was evidently inconsistent. Author names were standardized only for obvious variants referring to the same individual based on publication context and affiliation information. In addition, the names of the top authors and institutions reported in the main tables were manually cross-checked against the original WoSCC records and affiliation information.

Data Processing

CiteSpace (version 6.4.R1 Advanced) [27] and VOSviewer (version 1.6.20) [28] were used for bibliometric visualization and network analysis. CiteSpace was used to analyze the dual-map overlay of journals, reference cocitation, keyword clustering, and burst detection. VOSviewer was used to visualize collaboration networks among countries and regions, institutions, and authors, as well as keyword co-occurrence. In addition, Bibliometrix in R (version 4.4.1; K-Synth Srl) [29] was used for supplementary visualization of countries and regions and keyword data, and Microsoft Excel 365 was used to present annual publication trends.

For network construction in VOSviewer, the following thresholds were applied: keywords with at least 3 occurrences, countries and regions with at least 5 publications, institutions

XSL•FO
RenderX

with at least 3 publications, and authors with at least 2 publications were included. These thresholds were chosen to keep the networks interpretable and reduce noise from low-frequency items [30]. In CiteSpace, the time span was set from 2005 to 2025 with 1 year per slice, and Pathfinder pruning was applied to simplify the network structure and highlight the most meaningful links. Betweenness centrality was calculated by CiteSpace using its default settings and was used to identify important nodes in the networks.

Ethical Considerations

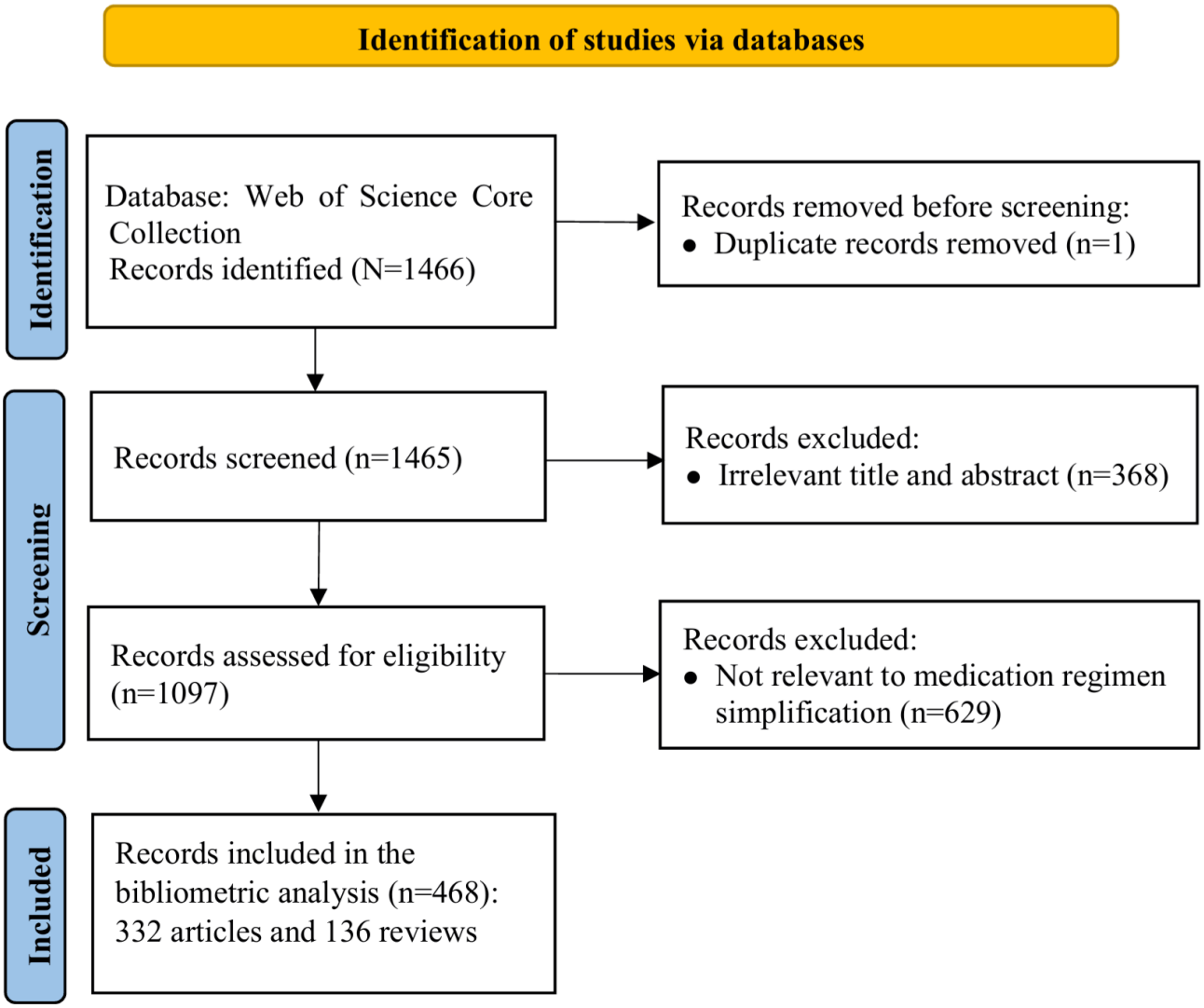
This study was based on publicly available literature and did not involve human or animal participants; thus, no ethics approval was required.

Results

Study Selection

A total of 1466 records were identified. After title and abstract screening, publications without an explicit focus on medication regimen simplification or reduction in medication regimen complexity were excluded. Ultimately, of the 1466 records, 468 (31.9%) publications were included in the final analysis. The study selection process is shown in Figure 1.

Figure 1. Flowchart of the study selection process.

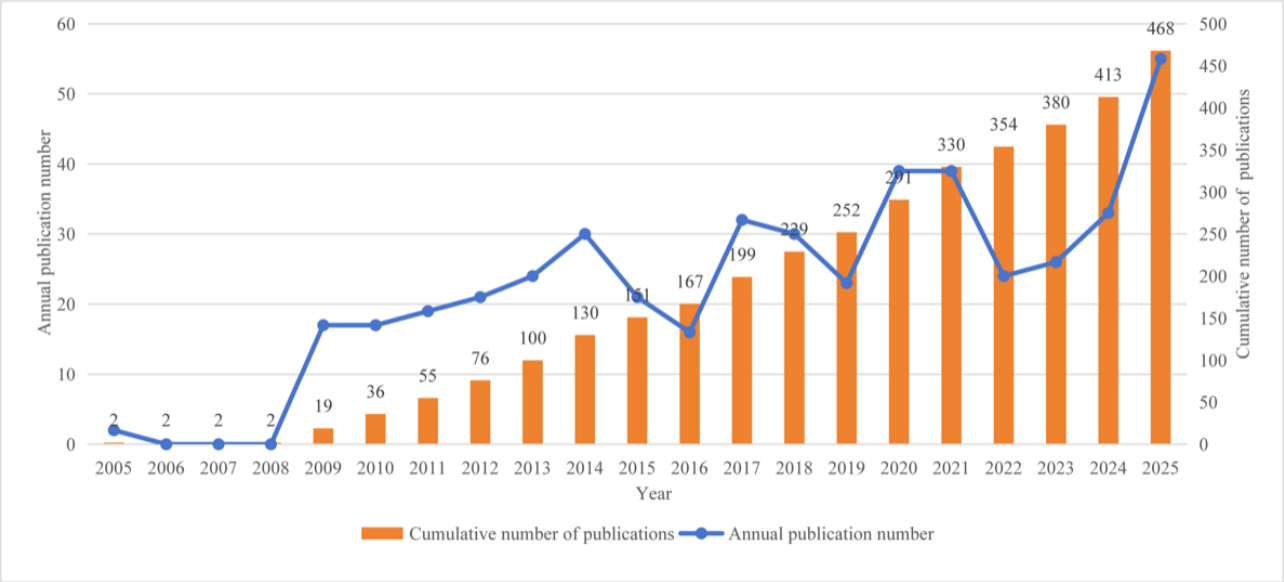


Annual Publication Statistics

Figure 2 illustrates the annual and cumulative number of publications on medication regimen simplification from 2005 to 2025. Annual publication output was low from 2005 to 2008 and then increased from 2009 onward, with fluctuations across years. The cumulative number of publications increased throughout the study period, from 2 in 2005 to 468 in 2025.

XSL•FO
RenderX

Figure 2. Annual volume of publications.



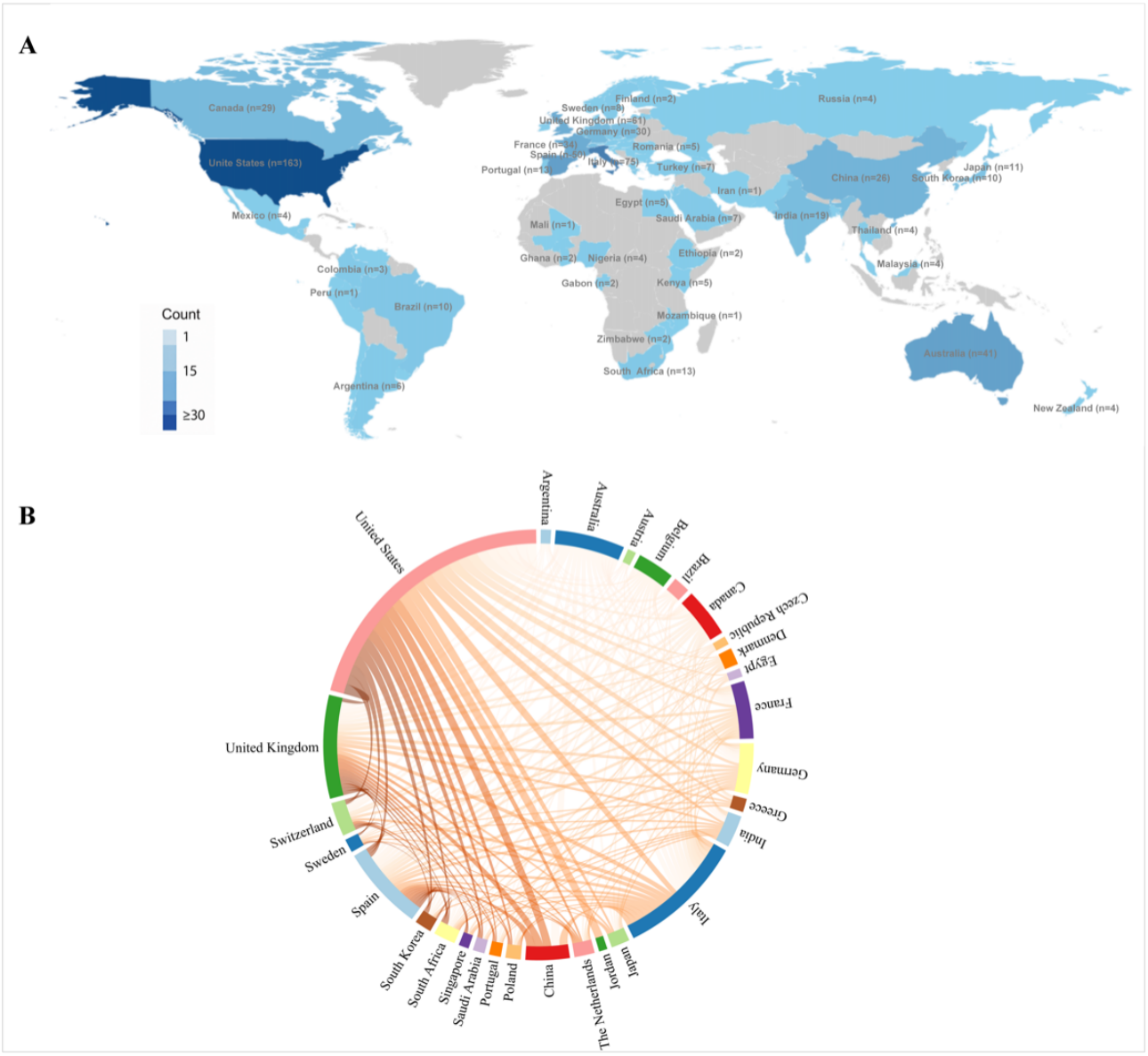
Global Contribution of Countries and Regions, Institutions, and Authors to the Field

Countries and Regions

Figure 3A shows the distribution of publications on medication regimen simplification across countries and regions from 2005 to 2025. The 10 most productive countries and regions and their citation counts are listed in Multimedia Appendix 1. The United States contributed the largest number of publications (163/468, 34.8%), followed by Italy (75/468, 16.0%) and the United Kingdom (61/468, 13.0%). The United States also had the highest citation count (n=6720), followed by Italy (n=3800). Figure 3B shows the country and region collaboration network, in which the United States, the United Kingdom, Italy, Germany, China, and Australia had larger nodes and multiple links with other countries and regions. Cross-national links were most evident between North America and Europe, with additional links involving East Asia and Oceania.

XSL•FO
RenderX

Figure 3. (A) Global distribution of publications (2005-2025) and (B) networks of country cooperation. UK: United Kingdom.



Institutions

Table 2 presents the 10 institutions with the highest publication counts and total citations. A total of 2234 institutions were identified. Monash University had the highest number of publications (16/468, 3.4%), followed by the University of Sydney (13/468, 2.8%). Among the top 10 most productive institutions, 5 (50%) were from Australia, 2 (20%) were from the United States, and 2 (20%) were from Spain. The University of Sydney had the highest total citation count (n=1764), followed by Harvard University (n=1740) and the University of Barcelona (n=550). Figure 4A shows the institutional collaboration network, in which multiple clusters are visible. Figure 4B presents the overlay visualization based on average normalized citations. Harvard University, Duke University, and Vivli are located in relatively warmer areas of the network, indicating higher average normalized citations.

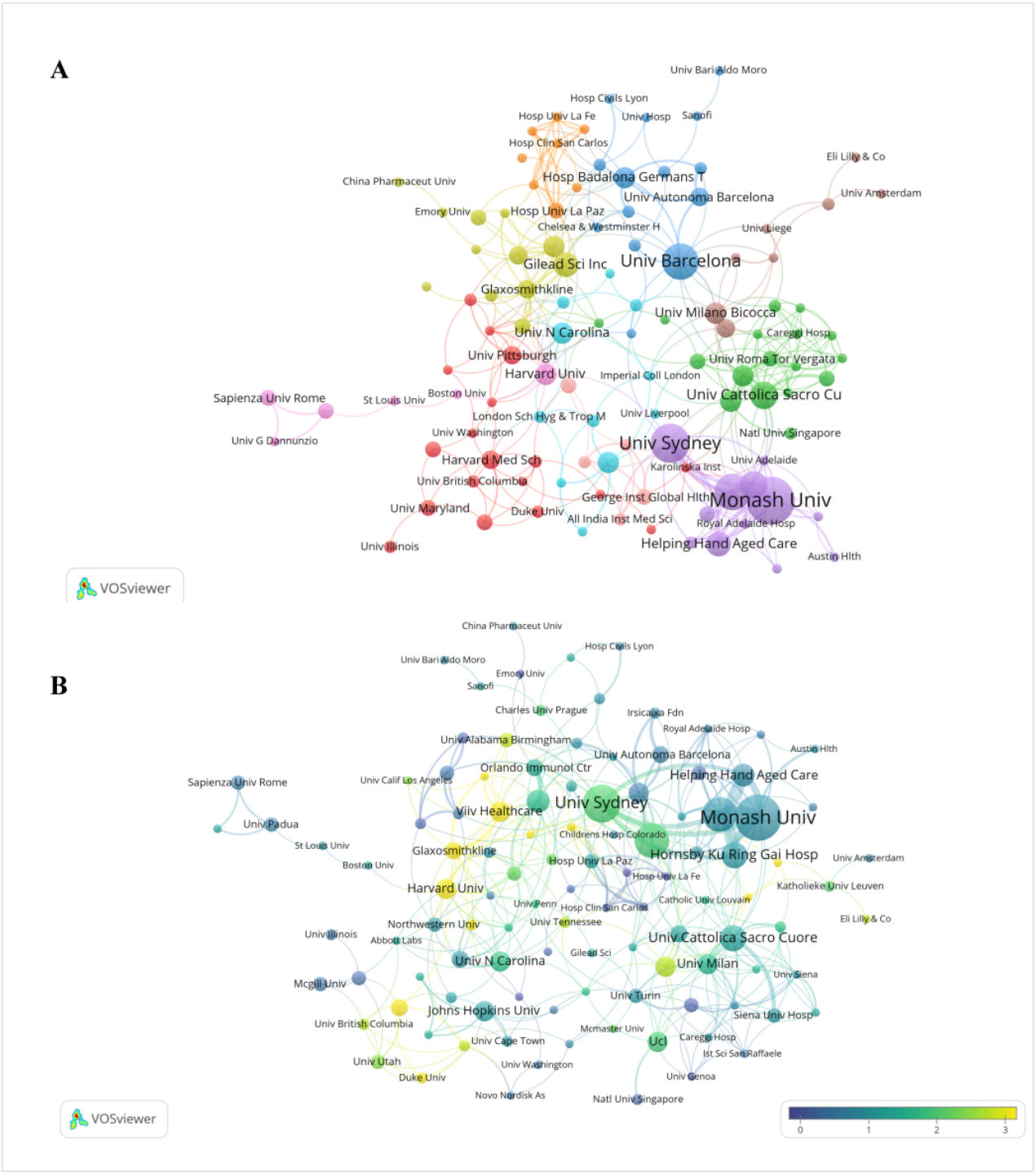
XSL•FO
RenderX

Table . Top 10 institutions by publication count and total citations.

Rank	Institution	Country	Studies (n=468), n (%)	Total citations, n
1	Monash University	Australia	16 (3.4)	333
2	University of Sydney	Australia	13 (2.8)	1764
3	University of Barcelona	Spain	12 (2.6)	550
4	University of South Australia	Australia	12 (2.6)	220
5	Hornsby Ku-ring-gai Hospital	Australia	9 (1.9)	213
6	Catholic University of the Sacred Heart	Italy	9 (1.9)	329
7	Gilead Sciences	United States	8 (1.7)	440
8	Helping Hand	Australia	8 (1.7)	149
9	Harvard University	United States	7 (1.5)	1740
10	Hospital Germans Trias i Pujol	Spain	7 (1.5)	125

XSL•FO
RenderX

Figure 4. (A) Institutional cooperation network and (B) overlay visualization of the institution network based on average normalized citations.

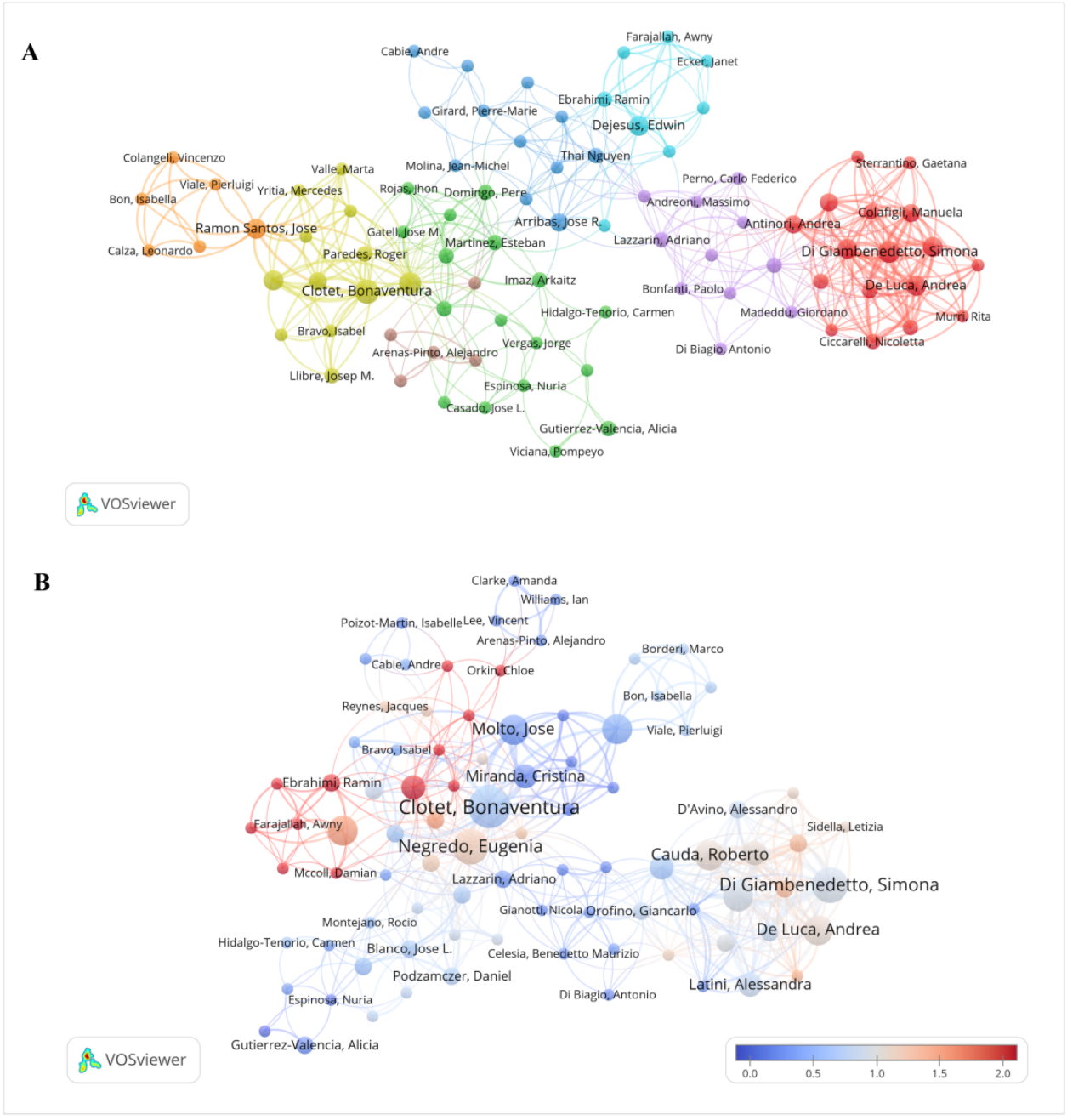


Authors

A total of 3109 authors contributed to the 468 publications on medication regimen simplification. Figure 5A shows the author collaboration network, in which multiple clusters are visible. Clotet, Di Giambenedetto, and De Luca are located in connected parts of the network and are linked with multiple coauthors. Figure 5B shows the overlay visualization of the author network based on average normalized citations. Warmer colors indicate higher citation impact. Clotet, Ebrahim Ramin, Farajzadeh Awny, and Di Giambenedetto are represented by relatively warmer-colored nodes. Table 3 shows author-level metrics. Bell had the highest number of publications (12/468, 2.6%), followed by Sluggett (11/468, 2.4%), and both authors had an H-index of 9. Di Giambenedetto had the highest total citation count (n=260), followed by De Luca (n=254), Bell (n=239), and Sluggett (n=237).

XSL•FO
RenderX

Figure 5. (A) Network map of author cooperation (2005-2025) and (B) overlay visualization of the author network based on average normalized citations.



XSL•FO
RenderX

Table . Most prolific authors based on author-level metrics.

Author	Publications (n=468), n (%)	H-index[a]	TC[b], n	PY_start[c]
Bell	12 (2.6)	9	239	2018
Sluggett	11 (2.4)	9	237	2018
Chen	8 (1.7)	8	155	2018
Corlis	9 (1.9)	8	150	2018
De Luca	7 (1.5)	7	254	2013
Di Giambenedetto	8 (1.7)	7	260	2013
Van Emden	7 (1.5)	7	121	2018
Caporale	6 (1.3)	6	112	2018
Cauda	6 (1.3)	6	217	2013
Clotet	7 (1.5)	6	132	2009

[a]H-index: Hirsch index; h publications cited at least h times each.

[b]TC: total citations.

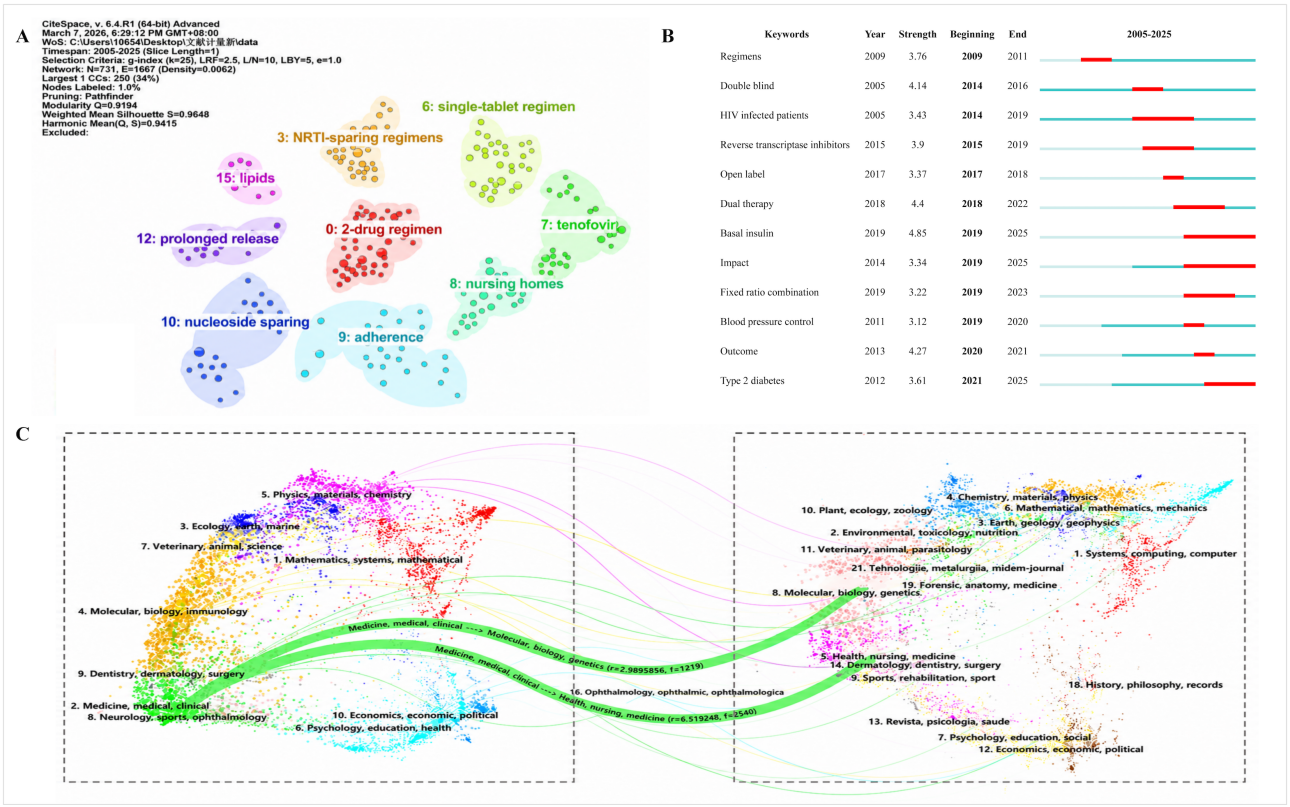
[c]PY_start: author's first publication year in the dataset.

Reference Cocitation and Keyword Burst Analysis

Figure 6A shows the reference cocitation clusters. Using CiteSpace with a 1-year time slice from 2005 to 2025, the network included 731 nodes and 1667 links, with a modularity *Q* of 0.919 and a silhouette score of 0.964. The main labeled clusters included 2-drug regimen (cluster 0), nucleoside or nucleotide reverse transcriptase inhibitors–sparing regimens (cluster 3), single-tablet regimen (cluster 6), tenofovir (cluster 7), nursing homes (cluster 8), and adherence (cluster 9). Figure 6B shows the 12 keywords with the strongest citation bursts. Burst keywords occurring earlier in the study period included "regimens," "double blind," "HIV infected patients," "reverse transcriptase inhibitors," and "open label," whereas burst keywords extending into later years included "dual therapy," "basal insulin," "fixed ratio combination," "blood pressure control," "impact," and "type 2 diabetes." The burst keywords "dual therapy" and "regimens" overlapped with the regimen-related cocitation clusters in Figure 6A, including 2-drug regimen (cluster 0), nucleoside or nucleotide reverse transcriptase inhibitors–sparing regimens (cluster 3), and single-tablet regimen (cluster 6).

XSL•FO
RenderX

Figure 6. (A) Reference cocitation clusters, (B) top 12 keywords with the strongest citation bursts (2005-2025), and (C) the dual-map overlay of journals publishing studies on medication regimen simplification from 2005 to 2025 (left: citing journals; right: cited journals). NRTI: nucleoside or nucleotide reverse transcriptase inhibitors.



Changing Trends in Research Disciplines

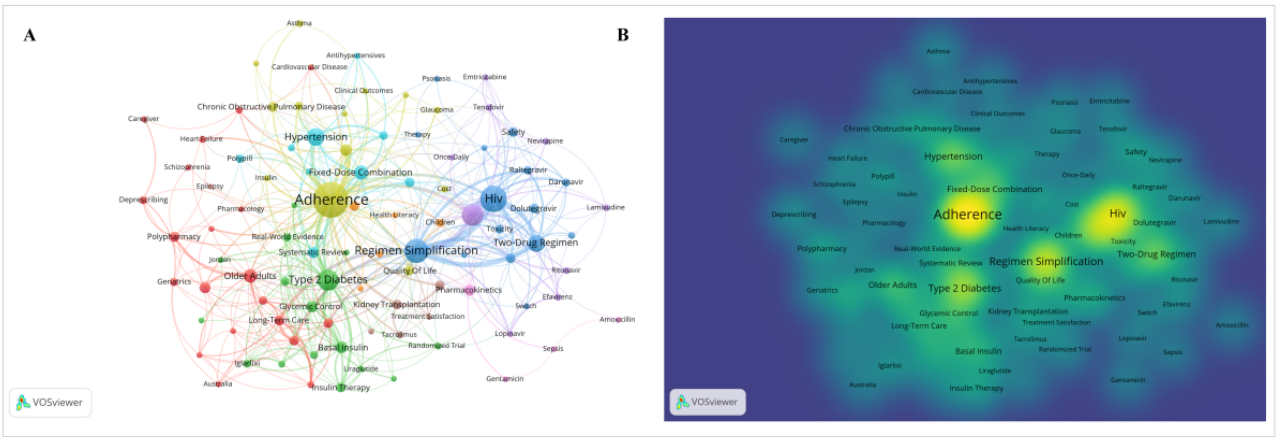
As shown in Figure 6C, the dual-map overlay reveals the knowledge flow between disciplines in the field of medication regimen simplification. The main citation paths originated from journals in the medicine, medical, and clinical area and extended to journals in the molecular, biology, and genetics area and the health, nursing, and medicine area. The most prominent paths are shown in green.

Keyword Analysis

Keyword Co-Occurrence

Figure 7A shows the keyword co-occurrence network generated using VOSviewer. Using all keywords as the unit of analysis, and after removal of irrelevant terms, 100 keywords with at least 3 occurrences were included. The most frequent keywords were “efficacy” (n=63), “therapy” (n=61), and “adherence” (n=61; Multimedia Appendix 2). In Figure 7A, “adherence,” “HIV,” “regimen simplification,” “type 2 diabetes,” and “fixed-dose combination” are located in connected parts of the network and have multiple links with other keywords. Figure 7B shows the density visualization of keyword co-occurrence. Areas of higher density are visible around “adherence,” “HIV,” and “regimen simplification,” with additional dense areas around “type 2 diabetes” and “fixed-dose combination.”

Figure 7. (A) Keyword co-occurrence network and (B) density visualization of keyword co-occurrence.

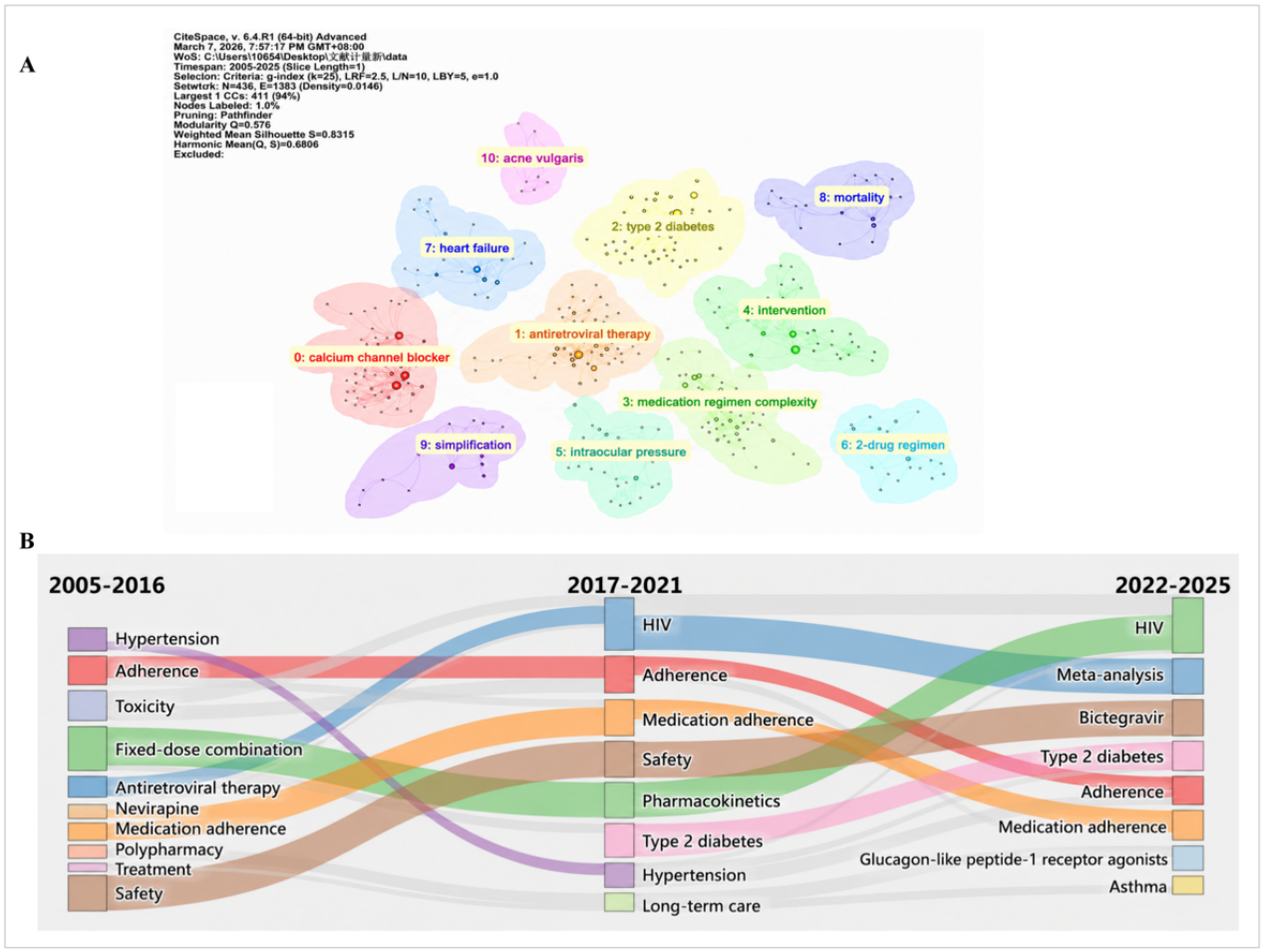


XSL•FO
RenderX

Keyword Cluster Analysis

Figure 8A shows the keyword clustering network generated by CiteSpace. Using the log-likelihood ratio algorithm, 11 clusters were identified. The network had a modularity *Q* of 0.576 and silhouette score of 0.831. The labeled clusters were calcium channel blocker (cluster 0), antiretroviral therapy (cluster 1), type 2 diabetes (cluster 2), medication regimen complexity (cluster 3), intervention (cluster 4), intraocular pressure (cluster 5), 2-drug regimen (cluster 6), heart failure (cluster 7), mortality (cluster 8), simplification (cluster 9), and acne vulgaris (cluster 10).

Figure 8. (A) Keyword clustering map and (B) thematic evolution of keywords (2005 - 2025).



Thematic Evolution

Figure 8B shows the thematic evolution of keywords across 3 periods: 2005 to 2016, 2017 to 2021, and 2022 to 2025. From 2005 to 2016, the main keywords included “hypertension,” “adherence,” “toxicity,” “fixed-dose combination,” “antiretroviral therapy,” “nevirapine,” “medication adherence,” “polypharmacy,” “treatment,” and “safety.” From 2017 to 2021, the main keywords included “HIV,” “adherence,” “medication adherence,” “safety,” “pharmacokinetics,” “type 2 diabetes,” “hypertension,” and “long-term care.” From 2022 to 2025, the main keywords included “HIV,” “meta-analysis,” “bictegravir,” “type 2 diabetes,” “adherence,” “medication adherence,” “glucagon-like peptide-1 receptor agonists,” and “asthma.” Links between periods are shown for several keywords, including “adherence,” “medication adherence,” “HIV,” “type 2 diabetes,” and “safety.”

Discussion

Principal Findings

This study provides a bibliometric overview of research on medication regimen simplification from 2005 to 2025. The findings show increasing publication output over time, with research activity concentrated in a limited number of countries, institutions, and author groups, whereas the intellectual structure and evolving keyword patterns indicate a field centered on regimen simplification, adherence, and disease-specific applications.

At the country level, research on medication regimen simplification was dominated by a limited number of contributors. The United States ranked first in both publication output and total citations, possibly reflecting not only the size of its research community and funding environment [31,32] but also broader factors that influence bibliometric visibility, including English-language publication and database coverage. At the institutional level, Monash University and the University

XSL•FO
RenderX

of Sydney were among the leading contributors, and 50% (5/10) of the most productive institutions were from Australia. Monash University's prominent contribution may reflect both its sustained research on medication use, medication safety, healthy aging, and aged care [33] and the broader Australian policy and practice focus on polypharmacy and medication management in older adults [34,35]. Notably, several productive institutions were hospitals, aged care providers, or industry-affiliated institutions, suggesting that medication regimen simplification is often closely linked to clinical practice and service delivery. This interpretation is supported by aged care simplification studies, including the Medication Regimen Simplification Guide for Residential Aged Care and related Australian care-based research [16,36], as well as by HIV simplification studies and treatment switch trials conducted by hospital-based and industry-affiliated teams [37,38]. Meanwhile, institution-level productivity may also be influenced by organizational structures, affiliation patterns, and database indexing practices and, therefore, should be interpreted with caution. At the collaboration level, collaboration patterns were similarly concentrated. International links were most evident between North America and Europe, with additional connections involving East Asia and Oceania, whereas institutional and author networks were organized into several clusters. This pattern suggests that the field has been shaped by a limited number of active research groups and broader cross-regional and cross-institutional collaboration may be needed to expand the evidence base across different health systems and care settings.

In this study, keyword burst analysis was used to identify terms receiving rapidly increasing attention over specific periods, whereas the interpretation of broader research hot spots was based on their consistency with cocitation structure, keyword clustering, and thematic evolution. A notable feature of medication regimen simplification is its extension across multiple disease areas and care settings. The complementary analyses suggest that medication regimen simplification has developed from an initially focused literature centered largely on antiretroviral therapy, fixed-dose combinations, and treatment-related toxicity into a broader area that also encompasses chronic disease management, long-term care, and medication regimen complexity. Within this broader development, HIV remained the most consistent research direction, with persistent signals from clusters related to "2-drug regimen," "NRTI-sparing regimens," and "single-tablet regimen," together with thematic evolution results including "antiretroviral therapy," "HIV," and "bictegravir." This continuity is likely related to the long-term nature of HIV treatment and the need to reduce regimen burden while maintaining adherence and efficacy [39,40]. Beyond HIV, the field has expanded into other long-term treatment settings, particularly diabetes and cardiovascular disease [41-43], with increasing attention to "type 2 diabetes," "heart failure," "blood pressure control," "basal insulin," and "glucagon-like peptide-1 receptor agonists." However, this expansion remains uneven across disease areas. HIV, diabetes, and selected cardiovascular topics were more prominently represented in the keyword and clustering analyses, whereas depression and Alzheimer disease were present in the keywords but did not form independent clusters or clearly defined thematic pathways. It may reflect variation in dominant research priorities across disease areas, with some fields historically placing greater emphasis on efficacy or survival-oriented outcomes rather than focusing on adherence or simplification [44,45]. It may also reflect the greater difficulty evaluating regimen simplification in mental health disorders and neurodegenerative diseases, where treatment and medication management are often individualized and shaped by cognitive impairment and caregiving involvement [46,47]. Further research should examine regimen simplification more explicitly in mental health, cognitive impairment, and multi-morbidity to clarify the indications, implementation, and long-term outcomes.

A central challenge in this field is no longer whether medication regimen simplification is feasible in principle but whether it can be implemented and sustained in routine care. Existing studies have focused mainly on populations with long-term medication use and high treatment burden, particularly older adults, people receiving community care, and those in long-term care settings [23,48,49]. In these populations, implementation depends not only on the regimen itself but also on cognitive and functional status, caregiver involvement, and the level of existing medication support [8,23,50]. Furthermore, simplification may also remain a clinical and pharmaceutical balancing process. Its real-world application requires consideration of effectiveness, safety, adherence, dosage form, dosing frequency, drug interactions, and pharmacokinetic constraints [8,13]. In this sense, regimen simplification is better understood as part of comprehensive medication management than as a simple act of discontinuing or switching treatment. Additionally, system-level factors may further shape implementation. In routine practice, medication simplification strategies may be limited by unclear workflows, insufficient multidisciplinary collaboration, and poor integration into everyday care processes [51,52]. These barriers may be more prominent in long-term care and community settings, where decision-making often depends on coordination among physicians, pharmacists, nurses, and caregivers [8]. These findings suggest that the future development of this field may depend not only on identifying effective simplification strategies but also on understanding how they can be embedded, coordinated, and sustained across real-world care settings.

The wider implementation of medication regimen simplification is likely to depend on policy and system conditions that support its integration into routine care. Reimbursement and reporting requirements may facilitate this process by providing institutional support for service delivery. For example, in the United States, Medicare Part D requires sponsors to establish medication therapy management programs and incorporate them into plan benefit structures and annual reporting requirements [53]. Quality measurement is another key element to routine implementation [54]. Although existing medication-related indicators can support accountability for medication use and safety [55,56], they seldom capture whether regimens have become simpler, whether simplification recommendations have been implemented, or whether treatment burden has been reduced. Therefore, more targeted indicators of regimen complexity, implementation, and treatment burden may be

XSL•FO
RenderX

needed. Furthermore, implementation is also likely to depend on clearer professional roles and better workflow integration. In real-world practice, simplification is rarely a single prescription event but, rather, an ongoing process of assessment, decision-making, implementation, and follow-up. Clearer role allocation across physicians, pharmacists, and nurses may improve implementation [8,23,52]. In this context, the formalization of pharmacist prescribing pathways in Australia may support a more structured role for pharmacists in simplification-related care [57]. Additionally, digital support should be considered within the same implementation framework. Its value is likely to depend less on stand-alone tools than on the integration of complexity assessment, simplification prompts, and follow-up monitoring into existing workflows. Previous studies suggest that electronic decision support is more likely to be scalable when combined with routine structured medication review in long-term care [58,59]. Therefore, a more feasible approach may be to embed simplification-related functions within established care processes, including electronic health records, postdischarge medication reconciliation, and medication review in long-term care and chronic disease follow-up. Beyond reviews that focus on prescribing decisions or regimen optimization in specific patient groups [60,61], this study highlights a broader implementation-oriented perspective by linking bibliometric patterns to policy conditions for routine care.

Limitations

Several limitations should be considered when interpreting the findings. First, this analysis was restricted to English-language publications indexed in the WoSCC. Therefore, relevant studies from other databases or in other languages may not have been captured, which could affect the completeness of the research landscape reported. Second, because medication regimen simplification is not a standardized indexing term, the search strategy relied on free-text terms and an operational definition, which may have resulted in the exclusion of conceptually relevant studies using different terminology. Third, bibliometric indicators describe patterns of publication, citation, and thematic structure, and recently published articles may have been underestimated because of citation lag. In addition, despite software-assisted standardization and manual checking, residual errors in author and institutional name disambiguation may have remained. Such errors can be reduced but are difficult to eliminate completely in bibliometric datasets, particularly for common surnames and institutions with multiple name variants, which could have affected the identification of prolific authors or institutions and the collaboration network analysis.

Conclusions

This bibliometric analysis indicates a clear increase in publications on medication regimen simplification over the past 2 decades, particularly in recent years. Quantitative mapping further indicated that the literature has concentrated mainly on HIV, diabetes, and cardiovascular-related management, whereas some areas such as mental health and cognitive disorders remain comparatively underrepresented. Overall, the field is expanding, but its thematic development is still uneven. Future studies should further clarify the definition and scope of medication regimen simplification, pay more attention to underexplored diseases and care settings, and strengthen the evidence on implementation and patient-relevant outcomes. Cross-national and cross-system research will also be important for developing quality indicators, implementation pathways, and policy frameworks to support the wider use of medication regimen simplification in real-world practice.

Funding

This work was supported by the National Natural Science Foundation of China (grants 72174061 and 71704053), China Scholarship Council (grant 202308330251), Key Research and Development Program of Zhejiang Province (grants 2026C02A1146 and 2025C02106), and General Project of Zhejiang Primary Health Association (grant 2024ZJWX-B014).

Data Availability

The datasets analyzed in this study are available from the corresponding author on justified request.

Authors' Contributions

Conceptualization: BW, L Wang
Data curation: L Wang
Funding acquisition: CZ, L Wang
Investigation: CD, XQ
Methodology: CD, AG, L Wang
Resources: CZ, JL
Software: CD, L Wu, AG, AH
Supervision: BW, XQ, JL, L Wang
Visualization: CD, L Wu
Writing—original draft: CD
Writing—review and editing: BW, AH, L Wang

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Top 10 countries in terms of number of publications and citations.
[DOCX File, 18 KB - aging_v9i1e82274_app1.docx]

Multimedia Appendix 2
Top 20 keywords in terms of frequency of occurrence and their corresponding centrality.
[DOCX File, 16 KB - aging_v9i1e82274_app2.docx]

References

1. Brown MT, Bussell JK. Medication adherence: WHO cares? Mayo Clin Proc 2011 Apr;86(4):304-314. [doi: 10.4065/mcp.2010.0575] [Medline: 21389250]
2. Claxton AJ, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. Clin Ther 2001 Aug;23(8):1296-1310. [doi: 10.1016/s0149-2918(01)80109-0] [Medline: 11558866]
3. Hughes CM. Medication non-adherence in the elderly: how big is the problem? Drugs Aging 2004;21(12):793-811. [doi: 10.2165/00002512-200421120-00004] [Medline: 15382959]
4. Elliott RA, Goeman D, Beanland C, Koch S. Ability of older people with dementia or cognitive impairment to manage medicine regimens: a narrative review. Curr Clin Pharmacol 2015;10(3):213-221. [doi: 10.2174/1574884710666150812141525] [Medline: 26265487]
5. Ghassab-Abdollahi N, Ghorbani Z, Kheirollahi N, Nadrian H, Hashemiparast M. Exploring the reasons for self-administration medication errors among illiterate and low-literate community-dwelling older adults with polypharmacy: a qualitative study. BMC Geriatr 2024 Dec 19;24(1):1010. [doi: 10.1186/s12877-024-05595-w] [Medline: 39702061]
6. Kim CY, Choi BY, Ryoo SW, Son SY, Min JY, Min KB. Health literacy and health-related quality of life in older adults with mild cognitive impairment. J Am Med Dir Assoc 2024 Nov;25(11):105253. [doi: 10.1016/j.jamda.2024.105253] [Medline: 39265633]
7. Maher RL, Hanlon J, Hajjar ER. Clinical consequences of polypharmacy in elderly. Expert Opin Drug Saf 2014 Jan;13(1):57-65. [doi: 10.1517/14740338.2013.827660] [Medline: 24073682]
8. Bell JS, McInerney B, Chen EY, Bergen PJ, Reynolds L, Sluggett JK. Strategies to simplify complex medication regimens. Aust J Gen Pract 2021;50(1-2):43-48. [doi: 10.31128/AJGP-04-20-5322] [Medline: 33543163]
9. Elnaem MH, Irwan NA, Abubakar U, Syed Sulaiman SA, Elrggal ME, Cheema E. Impact of medication regimen simplification on medication adherence and clinical outcomes in patients with long-term medical conditions. Patient Prefer Adherence 2020;14:2135-2145. [doi: 10.2147/PPA.S268499] [Medline: 33173282]
10. Promoting rational use of medicines. World Health Organization. URL: https://www.who.int/activities/promoting-rational-use-of-medicines [accessed 2026-03-16]
11. Neiman AB, Ruppar T, Ho M, et al. CDC grand rounds: improving medication adherence for chronic disease management - innovations and opportunities. MMWR Morb Mortal Wkly Rep 2017 Nov 17;66(45):1248-1251. [doi: 10.15585/mmwr.mm6645a2] [Medline: 29145353]
12. Cole JA, Gonçalves-Bradley DC, Alqahtani M, et al. Interventions to improve the appropriate use of polypharmacy for older people. Cochrane Database Syst Rev 2023 Oct 11;10(10):CD008165. [doi: 10.1002/14651858.CD008165.pub5] [Medline: 37818791]
13. George J, Phun YT, Bailey MJ, Kong DCM, Stewart K. Development and validation of the medication regimen complexity index. Ann Pharmacother 2004 Sep;38(9):1369-1376. [doi: 10.1345/aph.1D479] [Medline: 15266038]
14. Bell DS. Combine and conquer: advantages and disadvantages of fixed-dose combination therapy. Diabetes Obes Metab 2013 Apr;15(4):291-300. [doi: 10.1111/dom.12015] [Medline: 23013323]
15. Oh CK, Bang JB, Kim SJ, et al. Improvement of medication adherence with simplified once-daily immunosuppressive regimen in stable kidney transplant recipients: a prospective cohort study. Asian J Surg 2020 Jun;43(6):660-667. [doi: 10.1016/j.asjsur.2019.07.011] [Medline: 31353239]
16. Chen EY, Sluggett JK, Ilomäki J, et al. Development and validation of the Medication Regimen Simplification Guide for Residential Aged CarE (MRS GRACE). Clin Interv Aging 2018;13:975-986. [doi: 10.2147/CIA.S158417] [Medline: 29844664]
17. Nachega JB, Rosenkranz B, Pham PA. Twice-daily versus once-daily antiretroviral therapy and coformulation strategies in HIV-infected adults: benefits, risks, or burden? Patient Prefer Adherence 2011;5:645-651. [doi: 10.2147/PPA.S27558] [Medline: 22259241]
18. Reeve E, Gnjidic D, Long J, Hilmer S. A systematic review of the emerging definition of “deprescribing” with network analysis: implications for future research and clinical practice. Br J Clin Pharmacol 2015 Dec;80(6):1254-1268. [doi: 10.1111/bcp.12732] [Medline: 27006985]
19. Carollo M, Boccardi V, Crisafulli S, et al. Medication review and deprescribing in different healthcare settings: a position statement from an Italian scientific consortium. Aging Clin Exp Res 2024 Mar 8;36(1):63. [doi: 10.1007/s40520-023-02679-2] [Medline: 38459218]

XSL•FO
RenderX

20. Kuntz JL, Safford MM, Singh JA, et al. Patient-centered interventions to improve medication management and adherence: a qualitative review of research findings. Patient Educ Couns 2014 Dec;97(3):310-326. [doi: 10.1016/j.pec.2014.08.021] [Medline: 25264309]
21. Maxwell SR. Rational prescribing: the principles of drug selection. Clin Med (Lond) 2016 Oct;16(5):459-464. [doi: 10.7861/clinmedicine.16-5-459] [Medline: 27697811]
22. Elliott RA, O'Callaghan C, Paul E, George J. Impact of an intervention to reduce medication regimen complexity for older hospital inpatients. Int J Clin Pharm 2013 Apr;35(2):217-224. [doi: 10.1007/s11096-012-9730-3] [Medline: 23212732]
23. Sluggett JK, Ooi CE, Gibson S, et al. Simplifying medication regimens for people receiving community-based home care services: outcomes of a non-randomized pilot and feasibility study. Clin Interv Aging 2020;15:797-809. [doi: 10.2147/CIA.S248377] [Medline: 32581521]
24. Huang HC, Wang CH, Chen PC, Lee YD. Bibliometric analysis of medication errors and adverse drug events studies. J Patient Saf 2019 Jun;15(2):128-134. [doi: 10.1097/PTS.0000000000000205] [Medline: 26237238]
25. Kumar R, Goel P. Exploring the Domain of Interpretive Structural Modelling (ISM) for sustainable future panorama: a bibliometric and content analysis. Arch Computat Methods Eng 2022 Aug;29(5):2781-2810. [doi: 10.1007/s11831-021-09675-7]
26. Elliott RA. Reducing medication regimen complexity for older patients prior to discharge from hospital: feasibility and barriers. J Clin Pharm Ther 2012 Dec;37(6):637-642. [doi: 10.1111/j.1365-2710.2012.01356.x] [Medline: 22607618]
27. Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. J Am Soc Inf Sci 2006 Feb;57(3):359-377. [doi: 10.1002/asi.20317]
28. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010 Aug;84(2):523-538. [doi: 10.1007/s11192-009-0146-3] [Medline: 20585380]
29. Aria M, Cuccurullo C. bibliometrix: an R-tool for comprehensive science mapping analysis. J Informetr 2017 Nov;11(4):959-975. [doi: 10.1016/j.joi.2017.08.007]
30. Muis A, Sintang S. Bibliometric analysis on local wisdom research in Southeast Asia. Discov glob soc 2026;4(1):30. [doi: 10.1007/s44282-026-00366-x]
31. Budget. National Institutes of Health. 2025. URL: https://www.nih.gov/about-nih/organization/budget [accessed 2026-03-16]
32. US R&D totaled $892 billion in 2022; estimate for 2023 indicates further increase to $940 billion. National Center for Science and Engineering Statistics. 2025. URL: https://ncses.nsf.gov/pubs/nsf25327 [accessed 2026-03-16]
33. Centre for medicine use and safety. Monash University. 2025. URL: https://www.monash.edu/medicine/ehcs/marc/groups/cmus [accessed 2026-03-16]
34. Guiding principles for medication management in residential aged care facilities. Australian Government Department of Health and Aged Care. 2022. URL: https://www.health.gov.au/resources/publications/guiding-principles-for-medication-management-in-residential-aged-care-facilities?language=en [accessed 2026-03-16]
35. Polypharmacy, 75 years and over. Australian Commission on Safety and Quality in Health Care. 2021. URL: https://www.safetyandquality.gov.au/our-work/healthcare-variation/fourth-atlas-2021/medicines-use-older-people/61-polypharmacy-75-years-and-over [accessed 2026-03-16]
36. Dugré N, Bell JS, Hopkins RE, et al. Impact of medication regimen simplification on medication incidents in residential aged care: SIMPLER randomized controlled trial. J Clin Med 2021 Mar 6;10(5):1104. [doi: 10.3390/jcm10051104] [Medline: 33800845]
37. Negredo E, Bonjoch A, Clotet B. Benefits and concerns of simplification strategies in HIV-infected patients. J Antimicrob Chemother 2006 Aug;58(2):235-242. [doi: 10.1093/jac/dkl191] [Medline: 16807250]
38. Molina JM, Ward D, Brar I, et al. Switching to fixed-dose bictegravir, emtricitabine, and tenofovir alafenamide from dolutegravir plus abacavir and lamivudine in virologically suppressed adults with HIV-1: 48 week results of a randomised, double-blind, multicentre, active-controlled, phase 3, non-inferiority trial. Lancet HIV 2018 Jul;5(7):e357-e365. [doi: 10.1016/S2352-3018(18)30092-4] [Medline: 29925489]
39. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in adults and adolescents with HIV. Clinicalinfo, National Institutes of Health. 2025. URL: https://clinicalinfo.hiv.gov/sites/default/files/guidelines/documents/adult-adolescent-arv/guidelines-adult-adolescent-arv.pdf [accessed 2026-03-16]
40. Nachega JB, Parienti JJ, Uthman OA, et al. Lower pill burden and once-daily antiretroviral treatment regimens for HIV infection: a meta-analysis of randomized controlled trials. Clin Infect Dis 2014 May;58(9):1297-1307. [doi: 10.1093/cid/ciu046] [Medline: 24457345]
41. Lee JE, Lee J, Shin R, Oh O, Lee KS. Treatment burden in multimorbidity: an integrative review. BMC Prim Care 2024 Sep 28;25(1):352. [doi: 10.1186/s12875-024-02586-z] [Medline: 39342121]
42. Pantuzza LL, Ceccato MDG, Silveira MR, Junqueira LMR, Reis AMM. Association between medication regimen complexity and pharmacotherapy adherence: a systematic review. Eur J Clin Pharmacol 2017 Nov;73(11):1475-1489. [doi: 10.1007/s00228-017-2315-2] [Medline: 28779460]
43. Ab Rahman NA, Lim MT, Thevendran S, Ahmad Hamdi NA, Sivasampu S. Medication regimen complexity and medication burden among patients with type 2 diabetes mellitus: a retrospective analysis. Front Pharmacol 2022;13:808190. [doi: 10.3389/fphar.2022.808190] [Medline: 35387353]

XSL•FO
RenderX

44. Lam RW, McIntosh D, Wang J, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: section 1, disease burden and principles of care. Can J Psychiatry 2016 Sep;61(9):510-523. [doi: 10.1177/0706743716659416] [Medline: 27486151]
45. Geldmacher DS. Treatment guidelines for Alzheimer's disease: redefining perceptions in primary care. Prim Care Companion J Clin Psychiatry 2007;9(2):113-121. [doi: 10.4088/pcc.v09n0205] [Medline: 17607333]
46. Barkhof E, Meijer CJ, de Sonneville LMJ, Linszen DH, de Haan L. Interventions to improve adherence to antipsychotic medication in patients with schizophrenia--a review of the past decade. Eur Psychiatry 2012 Jan;27(1):9-18. [doi: 10.1016/j.eurpsy.2011.02.005] [Medline: 21561742]
47. Lim RH, Sharmeen T. Medicines management issues in dementia and coping strategies used by people living with dementia and family carers: a systematic review. Int J Geriatr Psychiatry 2018 Dec;33(12):1562-1581. [doi: 10.1002/gps.4985] [Medline: 30270451]
48. Aggarwal P, Woolford SJ, Patel HP. Multi-morbidity and polypharmacy in older people: challenges and opportunities for clinical practice. Geriatrics (Basel) 2020 Oct 28;5(4):85. [doi: 10.3390/geriatrics5040085] [Medline: 33126470]
49. Chen EY, Bell JS, Ilomaki J, et al. Medication regimen complexity in 8 Australian residential aged care facilities: impact of age, length of stay, comorbidity, frailty, and dependence in activities of daily living. Clin Interv Aging 2019;14:1783-1795. [doi: 10.2147/CIA.S216705] [Medline: 31695348]
50. Pereira F, Bieri M, Martins MM, Del Río Carral M, Verloo H. Safe medication management for polymedicated home-dwelling older adults after hospital discharge: a qualitative study of older adults, informal caregivers and healthcare professionals' perspectives. Nurs Rep 2022 May 31;12(2):403-423. [doi: 10.3390/nursrep12020039] [Medline: 35736616]
51. Sluggett JK, Hughes GA, Ooi CE, et al. Process evaluation of the Simplification of Medications Prescribed to Long-Term Care Residents (SIMPLER) cluster randomized controlled trial: a mixed methods study. Int J Environ Res Public Health 2021 May 27;18(11):5778. [doi: 10.3390/ijerph18115778] [Medline: 34072223]
52. Anrys P, Strauven G, Roussel S, et al. Process evaluation of a complex intervention to optimize quality of prescribing in nursing homes (COME-ON study). Implement Sci 2019 Dec 11;14(1):104. [doi: 10.1186/s13012-019-0945-8] [Medline: 31829252]
53. Medication therapy management. Centers for Medicare & Medicaid Services. 2026. URL: https://www.cms.gov/medicare/coverage/prescription-drug-coverage-contracting/medication-therapy-management [accessed 2026-03-16]
54. Medicare 2026 part c & d star ratings technical notes. Centers for Medicare & Medicaid Services. 2025. URL: https://www.cms.gov/files/document/2026-star-ratings-technical-notes.pdf [accessed 2025-09-25]
55. Gutteridge DS, Calder AH, Stasinopoulos J, et al. Quality indicators for safe and effective use of medications in long-term care settings: a systematic review. Br J Clin Pharmacol 2025 Nov;91(11):3054-3069. [doi: 10.1002/bcp.70242] [Medline: 40825531]
56. Expansion of the residential aged care quality indicators: evidence review summary report. Australian Government Department of Health and Aged Care. 2021. URL: https://www.health.gov.au/resources/publications/expansion-of-the-residential-aged-care-quality-indicators-evidence-review-summary-report?language=en [accessed 2026-03-16]
57. New accreditation standards for pharmacist prescriber education programs. Pharmacy Board of Australia. 2023. URL: https://www.pharmacyboard.gov.au/News/2023-12-21-New-accreditation-standards-for-pharmacist-prescriber-education-programs.aspx [accessed 2026-03-16]
58. McDonald EG, Estey JL, Davenport C, et al. Electronic decision support for deprescribing in older adults living in long-term care: a stepped-wedge cluster randomized trial. JAMA Netw Open 2025 May 1;8(5):e2512931. [doi: 10.1001/jamanetworkopen.2025.12931] [Medline: 40445620]
59. Perri GA, Bortolussi-Courval É, Brinton CD, et al. MedSafer to support deprescribing for residents of long-term care: a mixed-methods study. Can Geriatr J 2022 Jun;25(2):175-182. [doi: 10.5770/cgj.25.545] [Medline: 35747414]
60. Growdon ME, Smith AK. The maelstrom of medications-optimization of prescribing during the course of dementia. JAMA Intern Med 2023 Oct 1;183(10):1108-1110. [doi: 10.1001/jamainternmed.2023.3584] [Medline: 37603360]
61. Keller MS, Qureshi N, Mays AM, Sarkisian CA, Pevnick JM. Cumulative update of a systematic overview evaluating interventions addressing polypharmacy. JAMA Netw Open 2024 Jan 2;7(1):e2350963. [doi: 10.1001/jamanetworkopen.2023.50963] [Medline: 38198136]

Abbreviations

MeSH: Medical Subject Headings
WoSCC: Web of Science Core Collection

XSL•FO
RenderX

Edited by M Antoniou; submitted 04.Sep.2025; peer-reviewed by F Xiao, X Luo; revised version received 30.Mar.2026; accepted 12.May.2026; published 29.May.2026.

Please cite as:
Dou C, Wu B, Wu L, Gu A, Huang A, Zhang C, Qiu X, Lou J, Wang L
Global Trends and Research Hot Spots in Medication Regimen Simplification: Bibliometric Analysis
JMIR Aging 2026;9:e82274
URL: https://aging.jmir.org/2026/1/e82274
doi:10.2196/82274

© Chuwen Dou, Bei Wu, Leyi Wu, An Gu, An Huang, Chen Zhang, Xichenhui Qiu, Jianlin Lou, Lina Wang. Originally published in JMIR Aging (https://aging.jmir.org), 29.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Determinants of Medical Help–Seeking Behavior Following Case Finding of Early Cognitive Impairment: Semistructured Interview Study of Patients and Caregivers

Shu Rong Lim[1], MSc; Xiaohui Xin[1], MA; Gerald Choon Huat Koh[2], MMed, PhD; Julian Thumboo[1,3,4], MMed; Tau Ming Liew[2,4,5,6], FRCPsych, PhD

[1]Health Services Research Unit, Singapore General Hospital, Outram Road, Singapore, Singapore

[2]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

[3]Department of Rheumatology and Immunology, Singapore General Hospital, Singapore, Singapore

[4]SingHealth Duke-NUS Medicine Academic Clinical Programme, Duke-NUS Medical School, Singapore, Singapore

[5]Department of Psychiatry, Singapore General Hospital, Outram Road, Singapore, Singapore

[6]Health Services Research and Population Health, Duke-NUS Medical School, Singapore, Singapore

Corresponding Author:
Tau Ming Liew, FRCPsych, PhD
Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

Abstract

Background: Timely medical follow-up after a diagnosis of cognitive impairment, such as mild cognitive impairment (MCI) or dementia, is imperative for initiating appropriate medical treatment and accessing comprehensive care management and psychosocial support. However, many community-dwelling older adults who receive a positive case-finding result default on their medical follow-up appointments. This persistent challenge undermines early detection and active case-finding efforts and increases the risk of early institutionalization. Understanding the determinants is important for developing effective interventions in community-based case-finding.

Objective: This qualitative study aimed to explore the barriers and facilitators influencing medical help–seeking behavior among community-dwelling individuals in Singapore diagnosed with MCI or dementia following a positive case-finding result. The COM-B (capability, opportunity, motivation-behavior) framework of the Behavior Change Wheel (BCW) was used to systematically identify these determinants.

Methods: Using maximum variation sampling based on age, gender, ethnicity, living arrangement, relationship with caregiver, prior memory concerns, and whether participants had sought medical follow-up after their cognitive impairment diagnosis, we conducted 21 unique household semistructured interviews with 26 individuals (comprising 5 participant-caregiver dyads of MCI, 6 participants diagnosed with MCI, 1 caregiver of a participant with MCI, and 9 caregivers of participants diagnosed with dementia). These participants were a subset of individuals and caregivers recruited from a community-based study validating an artificial intelligence (AI)–based dementia case-finding tool. Interviews were audio-recorded, transcribed verbatim, and analyzed using the framework analysis method. Two authors (SRL and XX) performed inductive coding before mapping to COM-B, with subsequent discussion and review by the remaining authors. The COM-B components informed the selection of the relevant intervention functions from the BCW.

Results: Barriers to medical help–seeking behavior after cognitive impairment diagnosis included physical impairments, low health literacy to navigate the health care processes, inaccurate disease knowledge, complicated health care processes, inadequate physical infrastructure to navigate health care organizations, oversimplified cultural representation of dementia, perceived inaccuracy of the case-finding tool, and lay beliefs about seeking medical care when sick. Facilitators included adopting strategies to track medical appointments, a case-finding result letter as a cue for action, caregivers with profamily workplace policies, and valuing proactive medical disease management. Notably, social capital, (in)ability to recognize symptoms, and strong affective states triggered by the case-finding results were both barriers and enablers.

Conclusions: These findings highlight a complex interplay of individual, social, and structural determinants influencing medical help–seeking behavior following a cognitive impairment diagnosis. Rather than a linear trajectory, medical help–seeking is shaped by a dynamic interplay across the COM-B domains. Drawing on these findings, we developed a set of comprehensive and actionable recommendations grounded in the BCW intervention functions to support timely medical follow-up after a cognitive impairment diagnosis. To maximize their impact, these recommendations should be collaboratively refined and evaluated with stakeholders using the acceptability, practicability, effectiveness, affordability, side effects, and equity (APEASE) criteria.

XSL•FO
RenderX

(JMIR Aging 2026;9:e79386) doi:10.2196/79386

KEYWORDS

behaviour change wheel; COM-B; dementia; framework analysis; medical help-seeking; mild cognitive impairment; qualitative research methods

Introduction

Background

The global prevalence of dementia is projected to increase substantially, from 57.4 million cases in 2019 to 152.8 million in 2050, primarily due to population growth and an aging population, with a greater proportion of cases occurring among women than men [1]. International bodies such as Alzheimer's Disease International [2] and the International Association of Gerontology and Geriatrics [3] have advocated for active case-finding of dementia among high-risk older adults. This involves extending beyond tertiary hospitals to proactively identify high-risk individuals through routine annual cognitive evaluations in community-dwelling older adults [3]. However, the potential benefits of these cognitive evaluations in the community are diminished when individuals do not pursue timely medical follow-up after receiving a positive case-finding result for cognitive impairment (ie, mild cognitive impairment [MCI] or dementia). Timely medical follow-up is essential, not only for initiating appropriate medical treatments (eg, cognitive enhancers), but also for accessing comprehensive care management (eg, behavioral, financial, legal, and end-of-life issues), and social support (eg, support for caregivers) [4]. Delayed follow-up can have far-reaching consequences, including compromised well-being [5,6] and increased risk of early institutionalization in long-term care facilities [7-9].

Despite these risks, little is known about the frequency of actual medical help–seeking following a diagnosis of cognitive impairment. A large national survey conducted in China between 2015 and 2018 revealed that 71.4% (1974/2766) of individuals with dementia and 97.2% (6926/7125) of individuals with MCI did not seek medical treatment [10]. In Singapore, unpublished data from a pilot case-finding study [11] by the corresponding author (TML) as part of a larger implementation project to train and implement PENSIEVE-AI, an artificial intelligence (AI) tool for early cognitive impairment detection in the community, further illustrate this challenge. Among community-dwelling older adults diagnosed with cognitive impairment and referred directly to the Singapore General Hospital (SGH) memory clinics, 20.8% (10/48) defaulted on their appointments. Of these individuals, 8 were diagnosed with MCI and 2 with dementia. These missed follow-ups occurred despite deliberate efforts to facilitate medical follow-up, including direct referrals to SGH memory clinics for individuals living near SGH, text message appointment reminders, and proactive coordination by a clinic case manager before follow-up appointments. Although these data were derived from a single, relatively well-resourced acute tertiary hospital in Singapore, they likely underestimate the extent of the default follow-up rates in the broader population. Together, these findings highlight a persistent and multifaceted challenge of timely medical follow-up after receiving a diagnosis of cognitive impairment. Addressing this gap requires a better understanding of the determinants of medical help–seeking behavior in this context.

Existing Evidence on Medical Help–Seeking Determinants

A systematic review [12] identified 5 key illness perception factors associated with delayed help-seeking among individuals with dementia and their caregivers, including cultural beliefs of caregiving duties, acceptance of diagnosis, lack of knowledge about dementia, complex health care systems (including both the positive and negative experiences from prior health care professionals and services), and perceived threat to personal independence. However, the review focused on help-seeking intentions rather than actual help-seeking behaviors. This distinction is important due to the well-established intention-action or intention-behavior gap in health behaviors [13-15]. For clarity, throughout this paper, help-seeking without the qualifier medical refers to both formal (ie, health care professionals) and informal (ie, family and friends) sources of help-seeking support. In contrast, medical help–seeking behavior refers specifically to the actual actions taken to seek and obtain care from formal health care providers, such as neurologists, psychiatrists, or general practitioners.

Building on this, a recent qualitative study [16] extended medical help–seeking intentions in individuals with dementia and identified 8 key determinants of medical help–seeking behavior among individuals with MCI. These determinants can be conceptually clustered into 3 broad categories, including psychological processes, comprising disease knowledge, perceived disease threat, symptom attribution, negative valence emotions, and reliance on cognitive strategies; socioeconomic and structural conditions, that is, financial stability and accessible health care services; and social support. However, participants in this study were recruited from a memory clinic and were already on medical follow-up. As such, the determinants influencing individuals with MCI who did not seek medical help remain underexplored, leaving an important unaddressed gap in the context of medical help–seeking following positive case-finding in the community.

Study Aim

To address these methodological and practical knowledge gaps, this qualitative study aims to explore the barriers and facilitators influencing medical help–seeking behavior among individuals diagnosed with MCI or dementia following positive community-based case-finding results in Singapore. Specifically, what are the barriers and facilitators to medical help–seeking behavior among community-dwelling individuals diagnosed with cognitive impairment following a positive case-finding result? This study forms part of a larger implementation effort involving the deployment of PENSIEVE-AI but is methodologically independent from the PENSIEVE-AI tool development study [11]. In other words, although the

XSL•FO
RenderX

participants in this qualitative study were recruited from the PENSIEVE-AI pilot study [11], this study focuses specifically on exploring the determinants of medical help–seeking after MCI or dementia diagnosis. Given the complex, multifactorial nature of medical help–seeking behavior, a theory-informed approach was adopted to guide data collection and analysis. This approach enables the systematic identification of key determinants grounded in an established theoretical framework to inform the future development of targeted interventions and implementation strategies to improve medical help–seeking follow-up [17].

Theoretical Framework

This study applies the capability, opportunity, motivation-behavior (COM-B) framework [18,19], a widely used systematic framework for understanding behavior, designing targeted interventions, and achieving health behavior change across diverse settings, including dementia screening [20]. Within this framework, capability refers to the physical and psychological ability to perform the behavior; opportunity pertains to the physical and social environments; and motivation encompasses both automatic processes, such as emotions and impulses, and reflective processes, such as evaluation and plans [18,19]. Its utility lies not only in identifying behavioral determinants but also in linking the COM-B diagnosis with the 9 intervention functions that encircle the COM-B framework of the behavior change wheel (BCW): education, persuasion, enablement, incentivization, coercion, training, restrictions, environmental restructuring, and modeling [18,19,21]. This allows for the systematic development of tailored strategies specific to the behavioral determinants [18,19].

Methods

Study Setting

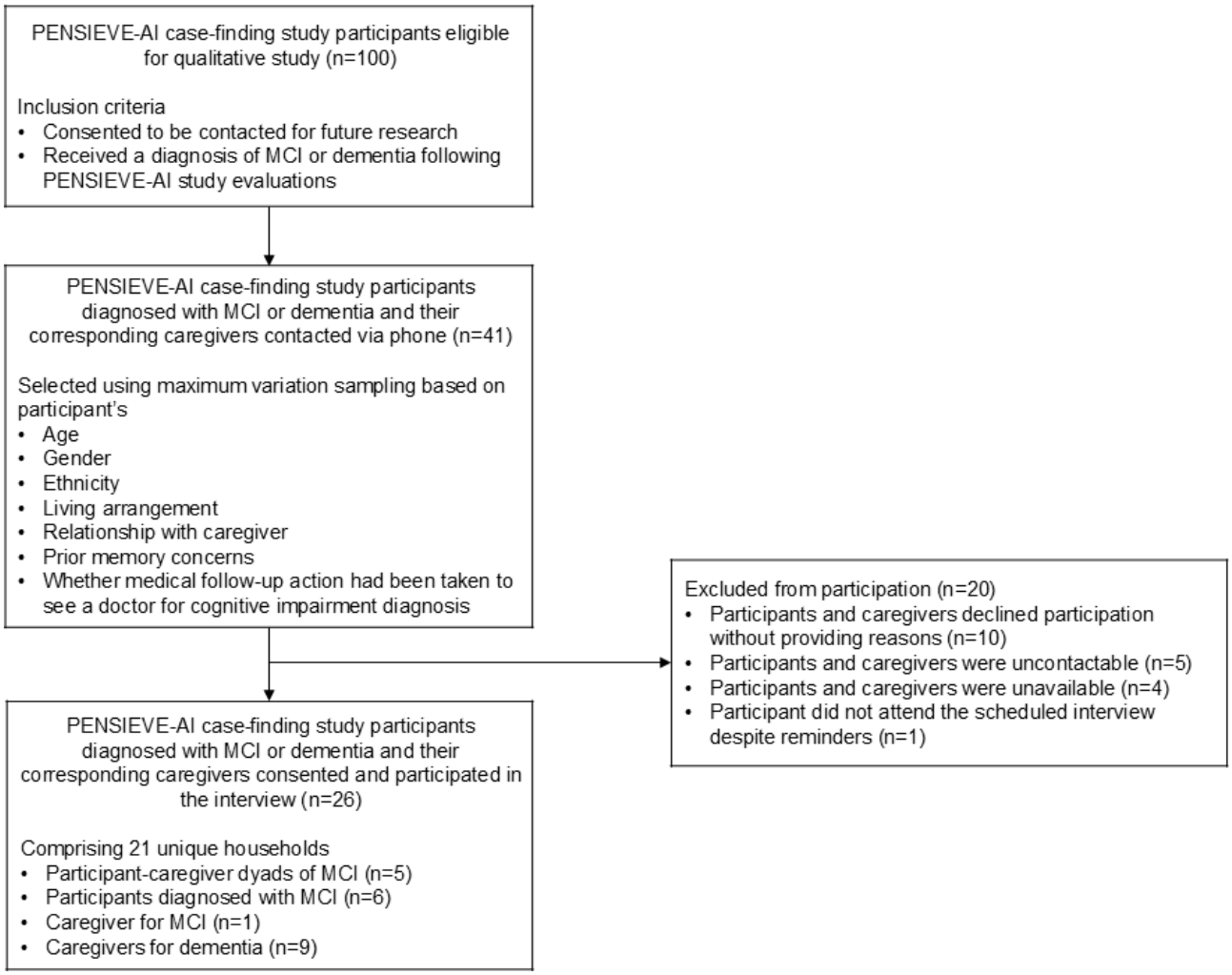
This study arose from a recent pilot case-finding study in Singapore that screened approximately 1800 community-dwelling older adults for early cognitive impairment using the PENSIEVE-AI tool [11]. Multimedia Appendix 1 outlines the steps of the PENSIEVE-AI case-finding pilot study. Community-dwelling older adults were recruited and underwent comprehensive cognitive assessments. Assessments of MCI, dementia, or normal cognition were made by 3 dementia specialists using an internationally recognized gold-standard consensus approach based on all the information from the comprehensive assessment. Community-dwelling older adults living in the SGH catchment area with positive case-finding results for MCI or dementia received direct medical appointments for follow-up at SGH memory clinics, whereas those living outside the SGH catchment area were strongly advised to seek medical attention from primary care doctors for referral to memory clinics within their home vicinity.

Study Design and Participant Selection

Forty-one eligible PENSIEVE-AI participants diagnosed with MCI or dementia and their corresponding caregivers from the pilot case-finding study [11] were contacted for the semistructured interviews over the phone. Participants who declined to participate did not preclude their caregivers from participating in the study. Interviewees were selected using maximum variation sampling to capture diverse experiences of medical help–seeking following positive case-finding. Sampling variation was based on participants' age, gender, ethnicity, living arrangement, relationship with caregiver, prior memory concerns, and whether participants had sought medical follow-up after their PENSIEVE cognitive impairment diagnosis. These characteristics were selected because they are expected to influence the cognitive impairment appraisal, access to formal and informal support for medical help–seeking, and caregiver involvement in decision-making regarding medical help–seeking. Figure 1 details the recruitment and selection process.

XSL•FO
RenderX

Figure 1. Qualitative study recruitment and selection flowchart for participants and caregivers. MCI: mild cognitive impairment.



For participants diagnosed with dementia, only their caregivers were interviewed due to practical considerations such as impaired decision-making capacity, reliance on interviewer's skill for consistent self-reports, and risk of distress during lengthy interviews. For participants diagnosed with MCI, interviews were conducted with participants alone, with participants and caregivers together, or with caregivers only, depending on participant preferences. This strategy aimed to improve interview participation rates. Interviews were conducted in either English or Mandarin Chinese. All interviews were audio recorded, anonymized, transcribed verbatim, and verified by another author (SRL or XX). Mandarin Chinese interviews were transcribed verbatim in the source language, coded in English, and only the extracts were translated into English by SRL and verified by XX. All authors are bilingual and proficient in English and Mandarin. All interviews were conducted by SRL, with XX accompanying one interview when a male participant was home alone. Interviews continued until data saturation was reached. An interview guide based on COM-B [18] was developed and used for the semistructured interviews (Multimedia Appendix 2). The COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist was used as a guide for reporting our study [22]. Informed consent was taken at the start of the interview.

Data Analysis

The data were analyzed using framework analysis [23] involving the following steps: (1) familiarizing with the interviews and transcripts; (2) independent inductive coding of the transcripts (ie, open coding) using Microsoft Word's comments function and the coding macro [24] to export the codes and the associated text to Microsoft Excel; (3) grouping codes into defined facilitator and barrier categories; (4) indexing the transcripts according to the existing codes and categories; (5) charting the categories into the COM-B matrix, including quotation references; and (6) performing data interpretation.

The identified COM-B components were used to guide the selection of relevant intervention functions from the BCW. The coding, indexing, charting, and selection were performed by SRL and XX, who have respective backgrounds in psychology and sociology. The categories were developed together through discussion with TML, who has expertise in geriatric psychiatry and public health, and subsequently reviewed by the remaining authors with expertise in geriatric medicine, medicine, health services research, and public health. Disagreements in the categorization of barriers and facilitators, mapping to COM-B components, and selection of intervention functions were resolved through discussion. We maintained an audit trail of coding comments, categorization and selection decisions, and engaged in reflexive discussions throughout the analysis.

XSL•FO
RenderX

Ethical Consideration

This study was approved by the SingHealth Centralized Institutional Review Board (reference number 2021/2590). Informed consent procedures for participants with cognitive impairment included a simplified capacity assessment in accordance with the Mental Capacity Act. Participants were required to demonstrate their ability to understand, retain, and use the study information to make an informed decision, and to communicate their decision to participate by paraphrasing key information, including the voluntary nature of participation and the right to withdraw at any time. The interviewer remained attentive to verbal and nonverbal cues of distress and paused interviews when necessary to ensure participants' and caregivers' well-being and to accommodate caregiving responsibilities. Participants and caregivers were reminded that they could terminate the interview at any time, and breaks were offered when needed. Data were kept confidential, with access restricted to the study team. Interview data were anonymized during transcription, with identifiable information removed and pseudonyms assigned where appropriate. Participants received Singapore $40 (US $1=1.26 Singapore Dollars) as a token of appreciation.

Results

Overview

Between May 2023 and January 2024, 26 individuals were interviewed across 21 unique households, comprising 12 households with a family member diagnosed with MCI and 9 households with a family member diagnosed with dementia. The mean interview duration was 53 minutes, ranging from 25 to 71 minutes. Of the 21 household interviews, 17 were conducted in person, and 4 by telephone. The median age group for participants with MCI was 70 - 79 years, while the median age group for the interviewed caregivers was 60 - 69 years. Most interviewees were Chinese, living with family members in a 4 - 5 room public housing apartment. Notably, 23.8% of households did not take further medical follow-up action after the positive findings from the pilot case-finding program, with 25.0% for MCI households and 22.2% for dementia households. Detailed sociodemographic characteristics of the interviewees are presented in Table 1. Five interviews were conducted in Mandarin Chinese.

XSL•FO
RenderX

Table . Sociodemographic characteristics of participants diagnosed with mild cognitive impairment (MCI) and caregivers of participants with MCI or dementia (n=26, across 21 unique household interviews: 5 participant-caregiver dyads of MCI, 6 participants diagnosed with MCI, 1 caregiver of a participant with MCI, and 9 caregivers of participants diagnosed with dementia).

Characteristic[a]	Participants with MCI[b] n=11	Caregivers of participants with MCI n=6	Caregivers of participants with dementia n=9
Age group (years)			
30 - 39	—[c]	1	—
50 - 59	—	1	4
60 - 69	4	3	1
70 - 79	4	1	3
80 - 89	3	—	1
Sex			
Female	6	4	8
Male	5	2	1
Ethnicity			
Chinese	9	4	8
Malay	—	1	—
Indian	1	1	1
Eurasian	1	—	—
Marital status			
Widowed/divorced/separated	3	1	—
Married	5	4	5
Single	3	1	4
Type of housing			
1 - 2 room public housing apartment	3	2	—
3-room public housing apartment	2	—	1
4 - 5 room public housing apartment	5	4	4
Private apartment/condominium	1	—	3
Landed property	—	—	1
Work status			
Retired/unemployed	9	2	6
Working	2	4	3
Living arrangement			
Lives alone	3	1	—
Lives with family	8	5	9
Attended medical follow-up for cognitive impairment			
Yes	8	1[d]	7[e]
No	3	—	2[e]
Type of caregiver's relationship with the participant			
Spouse	—	2	2
Child	—	3	5

XSL•FO
RenderX

Characteristic[a]	Participants with MCI[b] n=11	Caregivers of participants with MCI n=6	Caregivers of participants with dementia n=9
Sibling	—	—	2
Neighbor	—	1	—
Interview setting			
Home interview	4	1	2
Meeting room interview on the hospital campus	6	3	5
Phone interview	1	2	2

[a]Interviewees comprised 5 participant-caregiver dyads of MCI, 6 participants diagnosed with MCI, 1 caregiver of a participant with MCI, and 9 caregivers of participants with dementia.

[b]MCI: mild cognitive impairment.

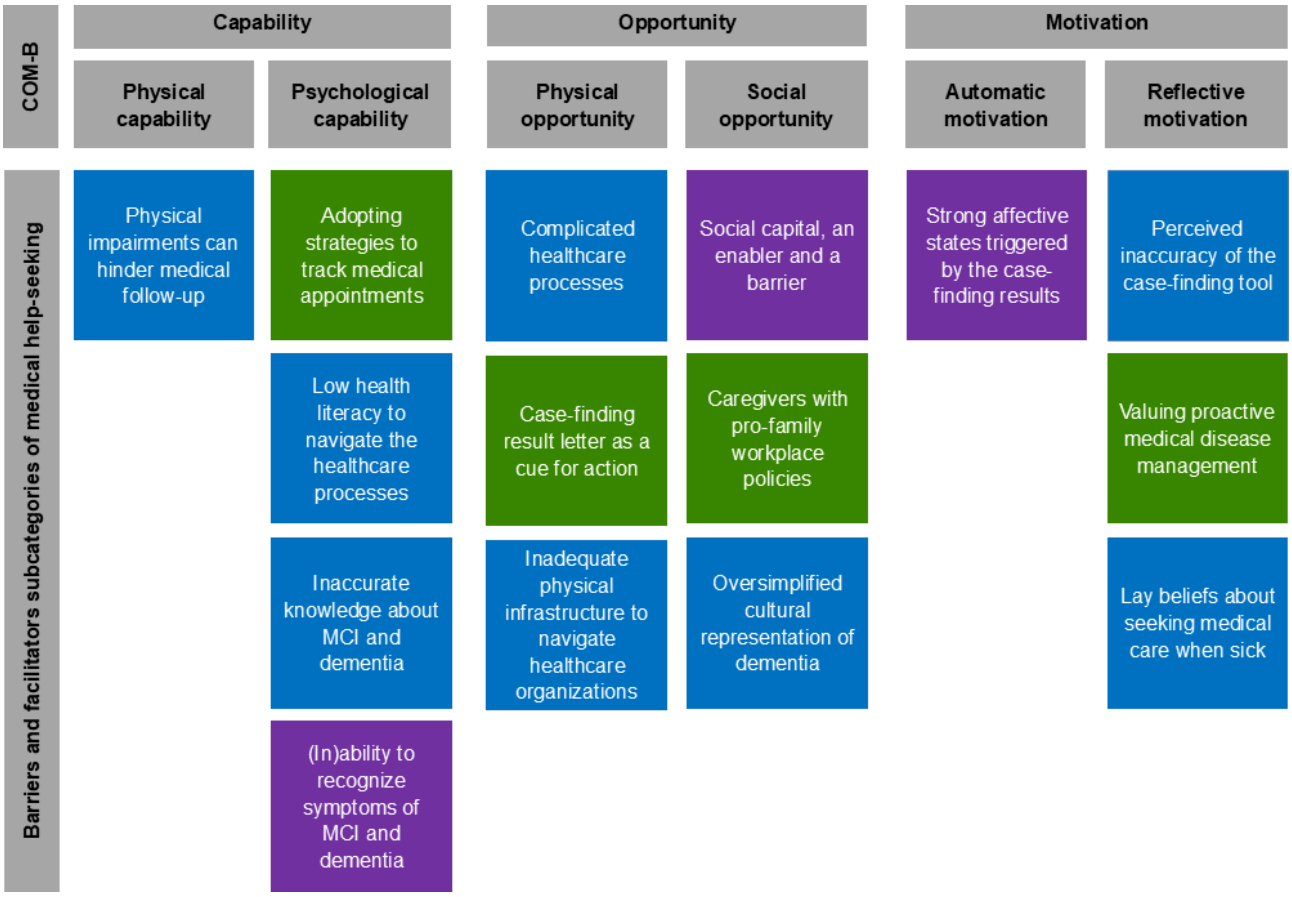
[c]Not applicable.

[d]Denotes a participant who was not interviewed due to a language barrier.

[e]Denotes participants with dementia who were not interviewed.

Based on the in-depth interviews, we identified 15 determinants encompassing both facilitators and barriers of medical help–seeking behavior. These determinants were mapped onto the COM-B domains. Detailed descriptions of each determinant, along with illustrative verbatim quotes, are detailed in the sections below. Figure 2 provides an overview of these determinants as mapped onto the COM-B domains and subdomains.

Figure 2. Determinants of medical help–seeking behavior following mild cognitive impairment (MCI) or dementia diagnosis, mapped onto the capability, opportunity, motivation-behavior (COM-B) domains and subdomains. Blue-colored boxes indicate barriers; green-colored boxes indicate facilitators; and purple-colored boxes indicate determinants that are both barriers and facilitators. COM-B: capability, opportunity, motivation-behavior; MCI: mild cognitive impairment.



XSL•FO
RenderX

Capability

Physical Impairments Can Hinder Medical Follow-Up

Two types of physical impairment were identified as barriers to timely medical follow-up after cognitive impairment case-finding results. First, mobility limitations necessitated reliance on caregivers for transportation, making the decision to attend medical follow-up appointments not solely the participant's, but also contingent on the availability of social support (Extract 1).

Extract 1

> *… she can't walk very well … so most of the time my husband … will drive her. So I have to also cater to my [husband's availability too]- sometimes my husband cannot make it.* [Caregiver 8 of participant with dementia]

Second, speech and hand impairments, particularly among stroke survivors, can impede participants' ability to independently manage appointment scheduling by phone or text messages (Extract 2).

Extract 2

> *Speaking to a person sometime[s] may have a problem … (In dialect) one duck one chicken … you communicate they can't get you. Sometimes over the phone … either I cannot … Because I have (sic) stroke … now I still can talk la but the slur is still there … [When typing on the phone,] it goes to the next alpha numeral … [as] I can't type with this [disabled] hand …* [Participant 5 with MCI]

Adopting Strategies to Track Medical Appointments

A notable facilitator of psychological capability was the adoption of practical strategies to track medical appointments. To mitigate forgetfulness and minimize missed medical appointments, participants and caregivers used resourceful strategies, such as diligently recording appointment dates on physical or digital calendars (Extracts 3 and 4) and prominently displaying appointment letters in a conspicuous part of their living spaces.

Extract 3

> *… when [I have an] appointment … I must note it immediately [on the physical calendar]. If I delay, I will tend to forget …* [Participant 1 with MCI]

Extract 4

> *… when I have the [doctor's] appointment … I surely will keep it up … I will put it down [in my phone calendar] …* [Participant 2 with MCI]

Low Health Literacy to Navigate the Health Care Processes

Low health literacy, defined as the limitation in the degree to which individuals have the ability to find, understand, and use information and services to make informed health-related decisions and actions for themselves or others [25], posed a barrier to timely follow-up after case-finding of cognitive impairment. Participants and caregivers, who lacked the familiarity with, and necessary skills to navigate, the health care processes (eg, rescheduling medical appointments to follow up on cognitive impairment case-finding results) had to rely on family members living elsewhere for assistance, thereby causing further delays or defaults on their medical appointments (Extract 5).

Extract 5

> *No, I don't know [we can bring forward medical appointments]. My appointments also I don't know. My sister [who does not live with me] will do [it] for me … when she comes [by to visit us] in the evening.* [Caregiver 9 of participant with dementia]

Inaccurate Knowledge About MCI and Dementia

Three types of inaccurate knowledge about cognitive impairment were identified: (1) misperceiving symptoms as normal aging, (2) misinterpreting cognitive function improvement, and (3) misunderstanding disease progression. First, participants and caregivers' understanding of MCI and dementia was primarily based on traditional lay beliefs rather than scientific knowledge, rooted in the common misperception that the symptoms of MCI and early dementia were a normal part of aging rather than a medical condition (Extract 6). As a result, these symptoms were not considered sufficient to warrant medical attention.

Extract 6

> *Caregiver: He just said he's 85 [years old], he's just forgetful …*
>
> *Participant: I feel that I'm just being forgetful only [due to my age]* [Caregiver 7 and participant 7 with MCI]

Second, participants and their caregivers described perceived subjective improvement in participants' cognition following treatment for other health conditions, which contributed to delays or defaults in seeking medical follow-up on case-finding results (Extract 7).

Extract 7

> *It is nothing unusual for him [to have MCI] as he's over 70 years already … you will definitely have it (MCI) when you are old … it is just a matter of degree … there is no need to see a doctor for it … because he's receiving regular vitamin [B12] injection at the polyclinic, I feel that his recent memory is better, more stable … there's [an] improvement [in his memory].* [Caregiver 10 of participant with MCI]

The third inaccuracy in the lay understanding of cognitive impairment arises from a life course perspective of the disease trajectory. While participants and their caregivers recognized the value of medical care to delay disease progression, they did not anticipate the cognitive impairment progressing to the disabling state of severe dementia within the participants' remaining lifetime. They anticipated other health issues to take precedence, diminishing the perceived benefit in seeking medical treatment for cognitive impairment (Extract 8).

Extract 8

XSL•FO
RenderX

> *… so after all, she's 83. You don't know, maybe before she reached that [advanced] stage [of dementia], she has (sic) some [other] diseases.* [Caregiver 16 of participant with dementia]

(In)ability to Recognize Symptoms of MCI and Dementia

The ability to recognize the signs and symptoms of MCI and early dementia as abnormal was a key facilitator in legitimizing the need to prompt participants to seek further medical attention (Extract 9).

Extract 9

> […. I left the stove on and I went out … I may lock the door, but I will leave my key … at the gate … there were a number of times I did that … Then the other thing … was that I am not certain how to pronounce this word … I don't think it is normal. That was the trigger to seek medical attention for MCI … that was my motivation.] [Participant 1 with MCI]

Caregivers who recognized these signs used the positive case-finding result to confirm their suspicions and used the diagnosis to encourage resistant family members who disagreed about the nature of their symptoms to pursue medical follow-up. However, denial remained a barrier despite these efforts (Extract 10).

Extract 10

> *… [the PENSIEVE-AI tool result] was expected … it's just to confirm the diagnosis and then at the same time hopefully can make her wake up and acknowledge that she has this issue. But she is still in denial. Up to today she still says … it's normal aging … old folks around her age are like that so she doesn't think it's an issue.*- [Caregiver 16 of participant with dementia]

In other cases, both the participants and the caregivers failed to recognize the symptoms of cognitive impairment and dismissed them, leading to inaction, which constitutes a major barrier (Extract 11).

Extract 11

> *… there's no signs like he's getting worse off or he's getting bad … he also thinks that it's nothing wrong … it's basically more like probably because he's not occupied …* [Caregiver 6 of participant with MCI]

Opportunities

Complicated Health Care Processes

Three barriers within the health care processes impeded timely medical follow-up on cognitive impairment case-finding results. The first was variability in the referral process. Participants living near SGH received direct appointments arranged by a hospital coordinator to follow up on the case-finding results with a specialist at the SGH memory clinic (Extract 12).

Extract 12

> *This was an appointment [date] given by [the hospital]. [If the hospital did not provide an appointment date], we will not go [to see the doctor]. We will wait for the doctor, [the doctor] can message me if they have something to tell me.* [Caregiver 3 of participant with MCI]

In contrast, those living further away were advised to obtain a referral from a primary care doctor to see a memory clinic specialist by themselves. This process introduced at least 2 potential risks for appointment delays and defaults: oversight by the participants, their caregivers, or both due to other commitments (Extract 13), and potential dismissal of the referral letter by primary care doctors (Extract 14).

Extract 13

> *… I [am] also forgetful … even though I went to the polyclinic with her [recently for urine infection], I forgot to ask the doctor for a referral [to see a specialist for dementia].* [Caregiver 8 of participant with dementia]

Extract 14

> *But because the doctor is the one that told him there is no need to come … so he never see[s] that doctor again …* [Caregiver 6 of participant with MCI]

The second barrier was the perceived high cost of laboratory investigations (eg, blood tests, brain imaging; Extract 15) that deterred medical follow-up.

Extract 15

> *The issue … [for] her is [the brain imaging scan is] going to cost me a lot of money.* [Caregiver 3 of participant with MCI]

Some participants and their caregivers were reluctant to make out-of-pocket cash payments for further investigations, with some preferring to use funds from their MediSave accounts, Singapore's national medical savings scheme. The MediSave annual Singapore $300 (US $236) withdrawal limit for outpatient diagnostic and treatment imaging per person influenced decisions to medical follow-up, depending on the availability of sufficient funds in their MediSave accounts for some of these individuals (Extract 16).

Extract 16

> *As long as not using hard cash it's okay, I don't think she minds [going for a scan] if [the hospital is] able to deduct the medical expenses from the C[entral]P[rovident]F[und] MediSave … so if we can postpone her scan till the next year so that she can pay via her MediSave as currently for this year her MediSave is almost all used up.* [Caregiver 3 of participant with MCI]

The financial considerations indirectly reflected how socioeconomic constraints shaped participants' and caregivers' decisions and behaviors about medical follow-up. Finally, commonly known problems such as long waiting times for a specialist appointment, extended clinic waiting periods, and difficulty in rescheduling a specialist appointment were acutely felt by participants and caregivers (Extracts 17 and 18). Improving these processes would make follow-up appointments

XSL•FO
RenderX

for cognitive impairment less onerous and more manageable for participants and caregivers.

Extract 17

> *… The problem is I couldn't get for her [an appointment] … It is something [of a] nuisance right we have to call them and check every day … It's not so easy, that day I call[ed] so many times they pick[ed] up and then they say (sic) no, you have to call every day. Really, it's a very disturbing nuisance what. How can I call the hospital every day? You all put [her] on the waiting list and then it's better for you all that … you check your computer system and then you tell us [when there is an available slot].* [Caregiver 12 of participant with MCI]

Extract 18

> *… the appointment is as good as don't have appointment … Because we every time have to wait [for] very long … your appointment is at 2 o'clock but you're seeing the doctor at 4 plus 5. I think it's ridiculous cause sometimes I have to work, so I will [have] no choice try to bring my laptop with me and work in the hospital … All hospitals [are] the same, no matter where you go, you have to wait. Not just [the] hospital[s], [the] polyclinic[s] also have to wait [for a] very long [time].* [Caregiver 18 of participant with MCI]

Case-Finding Result Letter as a Cue for Action

In this study, all PENSIEVE-AI participants received a letter titled "Feedback on Research Participation," which conveyed their positive case-finding result (ie, MCI or dementia), informed them of upcoming follow-up appointments or provided instructions to arrange for a specialist appointment through a primary care doctor, and included lifestyle advice to mitigate further memory deterioration. This letter was intended to function as a physical cue, prompting participants and caregivers to take necessary follow-up actions based on the positive case-finding results. However, one caregiver reported not receiving the letter. Although this was an isolated case, it demonstrated the importance such letters can serve in prompting action or, conversely, how the absence of the result letter might contribute to inaction following a positive case-finding result. In another instance, a participant and caregiver mistakenly perceived the result letter solely as a referral letter intended for the polyclinic doctors, viewing their role as mere couriers responsible for delivering the letter to the polyclinic doctor without engaging with its content (Extract 19).

Extract 19

> *… they said must go to the poly[clinic], give this letter to the poly[clinic doctor] … I didn't actually read [this letter] properly … because I thought this is for [the] doctor, [so] I don't want to read, I thought it might be confidential, for [the] doctor only. Now you [are] talking [to me] then I opened up and read … [previously] I asked John to bring to the doctor la (sic). So the doctor already opened up the envelope so now I'm reading [it] to you …* [Caregiver 7 of participant with MCI (The name (John) has been replaced with a pseudonym.)]

This example illustrates that the letter's effectiveness as a cue to prompt medical help–seeking behavior depends on participants' and caregivers' engagement with its content. When unread, the result letter could lose its cue effect.

Inadequate Physical Infrastructure to Navigate Health Care Organizations

Caregivers reported that participants often avoided using escalators due to mobility concerns (Extract 20), necessitating planning routes to medical appointments that use elevators instead.

Extract 20

> *… initially, she was still okay with the escalator, stepping on the escalator that is going up. She was just afraid to step on the escalator that's coming down. But these days, she dares not step on anything that moves.* [Caregiver 21 of participant with dementia]

Elevators became essential in the commute, yet issues such as limited availability of elevators, lengthy wait times, and overcrowded elevators further hindered accessibility for participants and caregivers. For example, accessing elevators during peak travel hours when taking the train to medical appointments. These barriers, common to many older adults with mobility limitations, extend beyond cognitive impairment. Additional barriers include the lack of easily accessible ramps for wheelchairs and pushchairs, insufficient space for mobility devices on public transport, and poorly lit parking lots with unclear and difficult-to-find signage within health care facilities (Extract 21).

Extract 21

> *the roads … are not built … for easy access or handicapped people, don't talk about even dementia people, just handicapped (sic). You see a lot of these places they are steps instead of ramps and even the ramps are not built in such a way [that is] properly designed … For the purpose of wheelchair-bound people or people having difficulty in walking, even my wife have (sic) difficulty in getting down, up and down the escalator, she gets very scared … when you are a bit older, your reaction is different. Your mental reaction … and this moving escalator is very very difficult for older people … and to be very frank as a driver I find Singapore the road … signage is very bad … The [carpark] exit, their signage is so bad … the directional signs are so bad … the underground carpark, you don't know east, west, north, south … want to go to [the] medical centre, [there is] no sign.* [Caregiver 14 of participant with dementia]

Social Capital, Both an Enabler and a Barrier

Social capital, defined as the resources derived from a person's social relationships to enable a desired social action [26], functioned as both an enabler and a barrier to medical follow-up

XSL•FO
RenderX

after a cognitive impairment case-finding result. As an enabler, social capital provided direct instrumental support, such as booking or rescheduling follow-up appointments on participants' behalf (Extract 22), reminding participants of medical appointments (Extract 23), driving and accompanying them to their appointments, and providing financial aid for medical bills (Extract 24).

Extract 22

> *… my son is arranging [for his mum] to see the doctor …* [Caregiver 14 of participant with dementia]

Extract 23

> *Caregiver: I grumbled and nagged [at him to see the primary care doctor for his MCI referral letter to the specialist].*
>
> [Participant: I cannot stand it, the nagging. That's why I went eventually.] [Caregiver 7 and participant 7 with MCI]

Extract 24

> *(When asked about difficulty paying for further investigation of the case-finding of cognitive impairment result) my friend always says anytime if you need financial assistance let me know.* [Participant 5 with MCI]

Social capital also played an indirect role in facilitating timely medical follow-up among socially active participants. Social activity increased the likelihood that others with accurate knowledge about cognitive impairment would notice signs of cognitive decline and encourage medical follow-up (Extract 25).

Extract 25

> *… her friend called me … during the [Japan] trip [and informed] … that she didn't bath[e], and then personal hygiene is (sic) an issue, and then erm forgetfulness or something but then she said it's not normal … I think your mum got to get her [to go for a] check-up because she's behaving differently from before.* [Caregiver 16 of participant with dementia]

This external feedback can reinforce the positive case-finding results; together, they may be more persuasive in motivating participants and their caregivers to acknowledge cognitive decline and pursue medical follow-up. Additionally, active social engagement enables participants to vicariously experience dementia in others via witnessing the lived experiences of dementia by someone they know. This exposure may heighten their sense of personal susceptibility to the disease and the importance of early medical intervention to delay the disease progression (Extracts 26 and 27).

Extract 26

> *… my own sister in her 90s had advanced dementia. But hers was I think from … the age of 80 … my younger brother who was discovered after me, suddenly have (sic) the advanced dementia … and completely now couldn't talk … he can't recognize anybody, advanced dementia, yeah he can't control his bowel everything. Yeah so his was very fast within five years of discovering … So I don't want to follow the same route as them. In the sense that I'm aware, I can still take control of my life.* [Participant 1 with MCI]

Extract 27

> *I went to [stroke] rehab [and got to know a] man … [with] dementia … then I came to know [he was] a newspaper editor… wah very sad and pitiful when … [at] this age you have this problem … he's educated … his mind suddenly slow[ed] down … the impact [on] normal person I don't mind. But he's an editor you know, that means basically he's educated … so [you] never know, it may happen to me … the possibility [of me having dementia] is always there …* [Participant 5 of MCI]

However, social capital can be a double-edged sword. Despite its facilitative role, it may also act as a barrier. When individuals with inaccurate knowledge about cognitive impairment normalize participants' symptoms of cognitive decline as typical signs of aging (Extract 28), and when members of the participants' immediate social circle (eg, primary caregivers, noncaregiving family members, friends, neighbors) disagree on the severity of the case-finding results and the importance of seeking medical attention, such discord can lead to conflict and hinder timely medical follow-up (Extract 29).

Extract 28

> *I can't remember things and my close friend … said oh no la (sic) old age la (sic) people will get old you know, will forget, yeah, so nothing la (sic) …* [Participant 1 with MCI]

Extract 29

> *… a lot of [my] neighbours, they keep telling me your father has no memory loss, dementia. See, he still can recognize us. He can still remember us, but I guess it's because they didn't (sic) live with him. They didn't talk to him … if you look at him you will not think that he has dementia … So it's actually very different [for] a caretaker and a non-caretaker, their views will be very different … it's only when you live with them then you know that how difficult it is … [my brother who does not live with us,] it's just like the same as the outsiders. They will say that my dad is so well but it's only the outside, but inside?* [Caregiver 17 of participant with dementia]

Caregivers With Profamily Workplace Policies

Many participants required family members to accompany them to follow-up appointments, both for instrumental support with transportation to and from the hospital and to help communicate symptoms accurately to the doctors. As many of the caregivers were still working, having a profamily employer was found to be an important facilitator that enabled them to fulfill this caregiving commitment and responsibility (Extract 30).

Extract 30

XSL•FO
RenderX

Fortunately, my current company is quite understanding, so I just take half-day leave … [Caregiver 17 of participant with dementia]

Oversimplified Cultural Representation of Dementia

Dementia was often portrayed in mass media as involving severe symptoms such as an inability to return home or failure to recognize family members. While such portrayals increased public awareness about the disease, they inadvertently promoted the false notion that dementia was absent unless these severe symptoms were present. This misunderstanding posed as a barrier for the participants and caregivers in recognizing the seriousness of the symptoms of MCI and early-stage dementia, as daily functioning was only minimally disrupted. As a result, participants' cognitive decline was not readily associated with dementia risk due to the mismatch with media representations. Rather than prompting timely medical help–seeking, the prevailing oversimplified cultural representation of dementia may have unintentionally justified the delays in medical follow-up (Extract 31).

Extract 31

… she say no la no la, I only old only la (sic) … everybody [who] get[s] old [will] get this kind of forgetfulness. See, I will never get this, not like this (referring to media portrayal of advanced dementia). So she denies. [Caregiver 16 of participant with dementia]

Motivation

Strong Affective States Triggered by the Case-Finding Results

Participants and caregivers expressed feelings of fear, fright, hopelessness, and sadness upon learning about the case-finding results. For some, these affective states acted as catalysts for seeking medical follow-up, prompting proactive actions when sufficiently intense (Extracts 32 and 33).

Extract 32

… since COVID (sic) when I see my [cognitive] decline and my forgetfulness, that really frightens me … you (sic) seem so hopeless like there's no cure no nothing … I see my own decline during this COVID time and then seeing my sister's … sharp … [cognitive] decline … So that makes me worried uh scared you know … then sometimes the trouble is there is no one to point it out to you. [Participant 1 with MCI]

Extract 33

Sad because I'm afraid that she can't remember me (caregiver was tearing) … I hope she won't reach that stage … [Caregiver 8 of participant with dementia]

However, for others, these affects were more ephemeral and did not spur action toward seeking medical follow-up (Extract 34).

Extract 34

she was a bit moody for about one or two days then after that she's back to her old self in denial again … I think [the moodiness was due to] when she do (sic) the [case-finding of cognitive impairment] test then she realized a lot of things that she cannot remember, a lot of things she cannot recall … [Caregiver 16 of participant with dementia]

Participants and caregivers used both emotions and moods to describe their feelings, often interchangeably, despite these being distinct constructs. Emotions, such as fear, are more concentrated, have specific sources, contain cognitive meaning, and are often short-lived, whereas moods, such as feeling bad, are relatively enduring, more diffuse, lower in intensity, lack specific sources, and have little cognitive meaning [27,28]. To capture participants' and caregivers' descriptions, we used the term affect as the generic psychological term that encompasses both emotions and moods [28].

Perceived Inaccuracy of the Case-Finding Tool

One rationale for not seeking medical attention following the positive case-finding results was rooted in the perception that the case-finding tool was biased, thus rendering the test result or diagnosis invalid. Although diagnoses in the pilot case-finding program were based on gold-standard practices (ie, consensus diagnosis by 3 specialists, including semistructured interviews, neuropsychological testing, and behavioral observations), the pilot program was embedded within a larger AI study with concurrent administration of a separate AI tool. This led to confusion among some participants and caregivers, who mistakenly assumed that the diagnosis was generated by an unvalidated AI tool still under development and testing, rather than the gold standard method for diagnosing cognitive impairment. As a result, trust in the legitimacy of the diagnosis was undermined and the need for medical follow-up was dismissed. Notwithstanding this confusion, it provided valuable insights into how perceptions of the case-finding process influence medical help–seeking behavior. Skepticism was particularly pronounced among participants with hand movement impairment due to stroke (Extract 35) and those with low technological literacy (Extract 36).

Extract 35

I can tell you I can easily fail [the case-finding of cognitive impairment test] … [as my mobility is] left [with] only the left hand … I think … I'm not able to pass … Because of my handicap [due to stroke] … [Participant 5 with MCI]

Extract 36

She wasn't able to answer the questions because of the [case-finding tool] technology, she doesn't know how to use [the tool] … we don't know the [diagnostic] accuracy of the [case-finding] tool. But the brain scan is very accurate … which we haven't done. [Caregiver 3 of participant with MCI]

In an extreme case, the case-finding process per se was perceived to be the direct cause of the MCI or dementia result (Extract 37).

XSL•FO
RenderX

Extract 37

> *I guess that she doesn't want to know her condition … she ever commented … the more you test, the more you see all this sickness, the better not to know that kind … Don't test, don't know, don't have.* [Caregiver 16 of participant with dementia]

Valuing Proactive Medical Disease Management

Several reasons were identified that facilitated timely medical follow-up after a positive case-finding result. These included the adage that prevention is better than cure, obtaining professional medical advice on disease management, and close monitoring of disease progression by a medical professional (Extract 38).

Extract 38

> *I don't see it as unnecessary [to see the doctor even though there is no treatment] … the specialists … at least [can] check … whether her condition deteriorates or improves or whatever … I don't see it as pointless because we're not doctors, we're not specialized, you (sic) don't even know how to evaluate the situation and her situation. So I think it would be good if the doctor in their professionalism they can actually advise her what is the condition, or at least we are prepared [for] what is going to happen. Rather than you just judge by yourself, you don't know whether it is right or wrong [and] make the wrong judgment.* [Caregiver 18 of participant with MCI]

Participants and caregivers also sought professional medical advice to support practical future planning and informed decision-making ahead of further disease deterioration. This included making a Lasting Power of Attorney to organize affairs and downsizing to a smaller apartment with dementia-friendly facilities to better accommodate the needs of the family member with dementia (Extract 39).

Extract 39

> *It is still [a] doubt or maybe [to hear from the doctor] mention by at what stage or what? … Simple things like … how? Should we downsize [our house or] relocate? And then [if we do] relocate, what kind of facilities should we look for to deal with the dementia patient? We have [a] general idea, but then we need something even more concrete [like concrete advice from the doctors] … and then which route to take or the best route to take, [what are the] possible route[s] to take. And then from there we will say, okay the route we could change later on but then at least how do we get started regarding the new route …* [Caregiver 14 of participant with dementia]

Lay Beliefs About Seeking Medical Care When Sick

The term "lay beliefs" is used here to refer to participants' and caregivers' understandings of when medical attention is needed. These beliefs may differ from clinical recommendations that promote timely and appropriate medical help–seeking for cognitive concerns. Rather than being uninformed or irrational, such beliefs are shaped by the personal, social, cultural, and economic contexts. Within this interpretive frame, several lay beliefs about medical help–seeking when sick were identified. Among participants who defaulted on their medical follow-up, there was a common belief that one should only visit the doctor when severely sick. Here, severe sickness is defined as a medical condition that markedly interferes with daily living and functional abilities (Extract 40).

Extract 40

> *For her current condition, I don't think there is a need for her to go for a brain scan … Because the doctor told me there is mild, moderate, and severe for this condition. To me, her mild condition isn't affecting her life much. But if I were to notice her condition worsens, then yes, even if have to pay cash for the [brain] scan we would go … How to differentiate mild, moderate, and severe … if for example … she went out via M[ass] R[apid] T[ransit] (train network) and can return home on her own, I think it's okay. But if she is unable to return home on her own, then I think it is problematic … and we must take action [to see the doctor].* [Caregiver 3 of participant with MCI]

Because MCI and early dementia often progress insidiously and may not noticeably disrupt daily living and functional abilities in the initial stages, some participants and caregivers perceived seeking medical attention for such conditions as unnecessary or even financially wasteful, particularly in the context of Singapore's high cost of living and the considerable out-of-pocket expenses associated with consulting a primary care doctor (Extract 41).

Extract 41

> *… I just wanted to bring her to the polyclinic for referral to [the] hospital but she is in denial … She says she doesn't [want treatment] … she's quite independent, quite stubborn, so I cannot force her to do anything … she said it's a waste of money [to see doctor] … in general. So she's kind of averse to seeing a doctor because nowadays seeing the G[eneral] P[ractitioner] is not cheap …* [Caregiver 16 of participant with dementia]

This belief was particularly salient among participants managing comorbid conditions (eg, cancer, cardiovascular diseases, and respiratory diseases), which were perceived as more severe, urgent, or life-threatening than cognitive impairment (Extracts 42 and 43).

Extract 42

> *… there is nothing worse than cancer … When people get sick, very sick or what, it's okay … What I mean is that when people die, people got cancer or what, take it easy lah. No need to worry [about the problems associated with MCI] …* [Participant 4 with MCI]

Extract 43

> *It's more of his asthma and his episodes of the asthma … it's more of his concern [than MCI] …* [Caregiver 6 of patient with MCI]

XSL•FO
RenderX

Among participants with comorbidities, disease prioritization was shaped by subjective interpretations of illness severity and financial considerations. Although these lay beliefs may diverge from clinical guidelines and recommendations advocating for early detection and intervention for cognitive impairment, they reflect a coherent lay understanding of disease severity, disease burden, and the health priorities, shaped by the lived experiences and concerns of both participants and caregivers.

Discussion

Principal Findings

This qualitative study identified 15 barriers and facilitators to medical help–seeking behavior following a diagnosis of MCI or dementia, categorized using the COM-B model. The discussion first contextualizes the findings within the existing literature, followed by integrated evidence- and practice-based recommendations.

Barriers

Our analyses on inaccurate knowledge about MCI and dementia, inability to recognize the symptoms of MCI and dementia, complicated health care processes, and low health literacy to navigate the health care processes aligned with commonly reported themes in the literature, including normalization of problems, lack of knowledge about dementia, and the complexity and inaccessibility of health care services [12,16,29,30]. Lay beliefs about seeking medical care when sick mirrored the finding of the review by Hill et al [30] on the perceived lack of benefit in medical help–seeking. Both reflect mismatches between clinical recommendations and the lay understandings that can undermine the perceived value of medical treatment, highlighting the need to increase public awareness of the benefits of timely medical intervention and care.

Strong affective states triggered by positive case-finding results were consistent with the findings of Jiao et al [16], who noted that individuals diagnosed with MCI perceived the disease as inherently threatening, characterized by negative valence emotions, such as anxiety and fear regarding disease progression. While Devoy and Simpson [29] identified fear as a barrier to medical help–seeking, our findings suggest a more nuanced, dual role, in which fear, conceptualized as a strong affective state triggered by the case-finding results, functions as both a facilitator and a barrier (eg, Extract 32). This duality reflects the multidimensional nature of fear, encompassing affective, cognitive, and psychobiological dimensions [31], with each influencing medical help–seeking behavior in different ways [32]. In the context of cancer screening, for instance, the cognitive and affective dimensions of fear facilitated cancer screening participation, while the psychobiological dimension hindered it [32].

A unique barrier identified in this study was the perceived inaccuracy of the case-finding tool. As participants were recruited from the larger case-finding pilot program using an AI tool, concerns about the tool's accuracy arose despite participants receiving diagnoses through the gold-standard consensus method. Future implementation of the PENSIEVE-AI tool [11] will need to explicitly address accuracy concerns and clarify the role of the AI tool in the cognitive impairment diagnosis process, particularly given that the consensus approach by memory specialists is not typical in routine clinical practice.

Medical help–seeking was further complicated by anosognosia in individuals with dementia and symptom fluctuation among individuals with MCI. Anosognosia, the lack of insight into one's own cognitive impairment, is a characteristic of dementia [33] that may contribute to denial and resistance to medical follow-up. This is evident in Extracts 10 and 41, where a caregiver described how her mother was in denial and misattributed cognitive symptoms to normal aging, features potentially indicative of the presenting symptoms of anosognosia. Similarly, among individuals with MCI, a clinically and neuropathologically heterogeneous group [34], the progression to dementia varies; while approximately 46% of individuals with MCI aged 65 years and older progress to dementia within 3 years [35], rising to 65% for those aged 75 years and older [36], others reported symptom improvement [37,38]. This symptom fluctuation makes early signs of cognitive impairment harder to recognize, contributing to uncertainty and delays in seeking medical help, as reflected in Extracts 7, 11, and 29.

Facilitators

Despite the identified barriers, several facilitators supported medical help–seeking behaviors. These included the ability to recognize cognitive impairment symptoms and the adoption of practical strategies to navigate symptom-related challenges, such as tracking medical appointments. Strong social capital, particularly reliance on caregivers and friends to reschedule medical appointments, provide transportation, and accompany individuals to medical appointments, facilitated medical follow-up behavior. Additionally, caregivers with profamily workplace policies and the case-finding result letter served as a physical cue to prompt medical follow-up action. For some individuals diagnosed with cognitive impairment, the strong affective states triggered by the case-finding results also prompted medical follow-up, particularly among those who valued proactive disease management through medical professional monitoring of disease progression. These facilitators align with the literature highlighting the importance of positive social support, disease knowledge, and the value of early detection in promoting medical help–seeking [16,29,30].

Our findings contrast with the paradoxical finding of Jiao et al [16] that individuals with MCI were less likely to seek medical help when using cognitive strategies. Our study found that individuals with MCI actively adopted strategies to track medical appointments, reflecting problem-solving efforts by the diagnosed individuals and their families to manage cognitive impairment symptoms. Notably, those with strong social capital and support networks were able to overcome some of the current barriers identified in this study, including physical impairments, low literacy to navigate health care processes, inadequate physical infrastructure, complicated health care processes, and lay beliefs about seeking medical care when sick. Facilitators unique to this study included caregivers with profamily workplace policies and the case-finding results letter as a cue

XSL•FO
RenderX

for action. The former strengthened social capital, while the latter is a distinctive finding in this qualitative study on determinants of medical help–seeking behavior following a cognitive impairment diagnosis via a community-based AI case-finding tool.

Integration With the Alzheimer Disease System Framework

Our findings extend the Alzheimer disease (AD) system framework of Ball et al [39] by specifying the barriers and facilitators influencing medical help–seeking behavior following a cognitive impairment diagnosis. These determinants can be situated within the AD system framework, between "health care system patient flow" and "access to support, care, and delivery," specifically following the "timely and accurate diagnosis" node [39]. Notably, current results demonstrate that the pathway from cognitive impairment case-finding or diagnosis to medical help–seeking is nonlinear. Instead, it is a process punctuated by interrelated determinants at multiple junctures, through a series of interactions between individual capacities and motivations, physical and social dynamics and opportunities, and broader structural conditions.

Although COM-B provides a systematic structure for identifying these determinants, several exhibit cross-domain interdependence. For instance, while physical impairments hindering medical follow-up are categorized under the physical capability subdomain, this barrier is closely linked to the social opportunity subdomain; many participants often relied on caregivers for transportation and company, a form of social capital that is both an enabler and a barrier. Such overlaps highlight the complexity and dynamic nature of medical help–seeking behavior following cognitive impairment case-finding results. Recognizing these interplays is essential for a nuanced understanding of medical help–seeking behavior and for designing interventions that address barriers within their broader systemic and contextual realities rather than in isolation.

Actionable Recommendations

In view of this complexity, we developed actionable recommendations by mapping BCW intervention functions onto our analytic findings. In addition to barriers, facilitators across the COM-B domains were integrated to bolster support for timely medical follow-up among individuals with cognitive impairment. Each determinant was mapped to one or more BCW intervention functions, as detailed in the matrix in Table 2. Training, coercion, and restrictions of the behavior change intervention functions were excluded; training addresses knowledge and skill gaps that are already covered by the education and enablement intervention functions, while coercion and restrictions are not suitable in this context. Further details of the recommendations are provided in Textbox 1.

XSL•FO
RenderX

Table . Matrix of the key determinants mapped onto the capability, opportunity, motivation-behavior (COM-B) domains and the suggested behavior change intervention functions.

Key determinants mapped onto COM-B[a] domains	Behavior change intervention functions[b]					
	Education	Persuasion	Role modeling	Enablement	Incentivization	Environmental restructuring
Capabilities						
Physical impairments can hinder medical follow-up			✓	✓	✓	✓
Adopting strategies to track medical appointments	✓		✓	✓		
Low health literacy to navigate the health care processes	✓		✓	✓		
Inaccurate knowledge about MCI[c] and dementia	✓	✓		✓		
(In)ability to recognize symptoms of MCI and dementia	✓	✓		✓		
Opportunities						
Complicated health care processes				✓	✓	
Case-finding result letter as a cue for action		✓		✓		
Inadequate physical infrastructure to navigate health care organizations				✓		✓
Social capital, an enabler and a barrier	✓	✓	✓	✓	✓	
Caregivers with profamily workplace policies					✓	
Oversimplified cultural representation of dementia	✓		✓			
Motivation						
Strong affective states triggered by the case-finding results	✓	✓	✓	✓		
Perceived inaccuracy of the case-finding tool		✓		✓		
Valuing proactive medical disease management	✓	✓	✓		✓	

XSL•FO
RenderX

Key determinants mapped onto COM-B[a] domains	Behavior change intervention functions[b]					
	Education	Persuasion	Role modeling	Enablement	Incentivization	Environmental restructuring
Lay beliefs about seeking medical care when sick	✓	✓	✓	✓		

[a]COM-B: capability, opportunity, motivation-behavior.

[b]Training, coercion, and restrictions of the behavior change intervention functions are excluded.

[c]MCI: mild cognitive impairment.

XSL•FO
RenderX

Textbox 1. Actionable recommendations based on the behavior change intervention functions.

Education through promoting awareness and understanding of early cognitive impairment among the public, consumers, and health care professionals is pivotal. Educational efforts can be delivered through public outreach initiatives (eg, television programs, hospital websites, public roadshows, newsletters, informational brochures included with the case-finding result letters) with a focus on the following:

1. **Recognizing symptoms**: educating individuals about the symptoms of mild cognitive impairment (MCI) and the various stages of dementia to enable early recognition of cues and to seek timely medical help.
2. **Future planning**: highlighting the importance and value of early planning for care and decision-making for the present and the future to motivate individuals to address cognitive health.
3. **Reframing MCI and dementia as a cognitive disability**: reframing cognitive impairment, as cognitive disability as this is not different from physical disabilities. Such reframing can plausibly destigmatize cognitive impairment and rally a whole-of-society approach to improve medical help–seeking behavior. Individuals may be more likely to seek help when cognitive impairment is not perceived as a liability (ie, akin to a physical disability that requires prostheses to maintain normalcy), while immediate family and the rest of society will be more willing to step up to provide support to individuals with cognitive impairment, akin to providing a prosthetic environment for the cognitive disability.
4. **Facilitating conversations**: strategies for broaching the topic of medical help–seeking with loved ones who are resistant to medical help–seeking for their cognitive impairment (eg, expressing concern, validating feelings, addressing misconceptions, and collaboratively exploring alternatives).

Persuasion through the following strategies:

1. **Providing result letters after case-finding efforts**: the letters should be designed using psychological and behavioral science theories and principles to encourage medical follow-up. The letters should:

- Clearly explain the benefits of further follow-ups and diagnostic tests, such as brain imaging, to rule out cognitive impairments that may be treatable.
- Highlight the various subtypes of cognitive impairment, including those that may be treatable or reversible.
- Include personalized risk of dementia to increase the urgency of the desired action.
- Provide a clear, concise, and easy-to-understand description of the follow-up procedures, including the estimated cost after government subsidies, and the expected duration of each procedure.
- Highlight the benefits of early diagnosis and management of cognitive impairment, which extend beyond direct immediate health benefits, such as enabling individuals with cognitive impairment to secure a better quality of life with the necessary help and support and allowing them to plan for the future and prioritize what matters most to them even in the face of cognitive impairment.

1. **Featuring dementia specialists and community leaders on television programs and radio talk shows.**

Role modeling by showcasing interviews and testimonials from individuals who sought medical help for early cognitive impairment and are living physically, mentally, and socially active lifestyles can serve as powerful role models. These interviews should ideally incorporate the following key topics:

1. **Rationale for proactive disease management**: their reasons for taking a proactive approach to managing their condition.
2. **Support systems**: the support they received from health care professionals, caregivers, friends, and the community.
3. **Strategies to cope and live with the disease**: how they manage the strong negative emotions triggered by the case-finding results and the strategies they use to keep track of their medical appointments.

Interactive and social media platforms can be used to disseminate these videos, fostering a more proactive approach to medical care.

Enablement through:

1. **Providing information on frequently asked questions about the case-finding tool**: addressing confusion, doubts, and skepticism of the validity of the case-finding results by offering clear and accessible information on the frequently asked questions.
2. **Simplifying the referral process via direct referrals to memory clinics**: streamlining the referral process to increase follow-up rates and minimize delays by implementing direct referrals to specialized memory clinics.
3. **Formulating practice guidelines**: the practice guidelines should include simplified key points on when primary care physicians should be concerned about cognitive impairment (as opposed to normal aging) and simplified criteria on when referrals to memory clinics are considered appropriate.
4. **Deploying trained community workers/volunteers** (eg, from Active Ageing Centres and Silver Generation Office) for engagement with individuals with positive case-finding results. These community workers/volunteers can serve the dual purpose of monitoring well-being and providing educational and persuasive interventions and support, including:

- Reinforcing the importance of seeking medical help for early cognitive impairment to both individuals and caregivers.
- Supporting individuals in managing the strong negative affective states triggered by the case-finding results.
- Guiding individuals and caregivers on the strategies for tracking medical appointments.

XSL•FO
RenderX

- Coaching caregivers on managing resistant behaviors.
- Assisting with rescheduling of medical follow-up appointments if necessary.
- Accompanying individuals to the memory clinics, particularly those without immediate caregivers.
- Escalating more complex behaviors or challenges to the community health care workers (eg, community nurses).
- Linking individuals to available community programs that address the risk factors for cognitive decline (eg, programs that target physical activities, diet, and cognitive and social engagement).

Incentivization is particularly important given that financial concerns and socioeconomic constraints can deter medical help–seeking behavior. Strategies should therefore prioritize affordability and equity, including policies aimed at reducing out-of-pocket costs and addressing indirect financial burdens to improve access for individuals and families with constrained financial resources. Incentivization can be implemented in the form of:

1. **Financial incentives for medical follow-up**: providing financial incentives to individuals with positive case-finding results to encourage attending memory clinics (eg, indirect financial rewards).
2. **Increasing the withdrawal limit for national medical savings**: increasing the annual withdrawal limit of health insurance or MediSave for diagnostic testing thereby reducing out-of-pocket payment for diagnostic imaging and subsequent treatments.
3. **Subsidy for initial consultation fees**: waiving the consultation fee for the initial visit to memory clinics or providing a higher subsidy for specialist follow-ups.
4. **Providing subsidized transportation**: offering subsidized transportation to memory clinics, particularly for individuals who have physical difficulties. This should also include the accompanying caregiver and community volunteer.
5. **Profamily workplace policies** as a type of nonfinancial incentive by providing work-from-home options for caregivers, additional leave for caregiving responsibilities, increased sick leave for individuals with MCI and dementia, and emergency time off (eg, when an individual with dementia gets lost and fails to return home).

Environmental restructuring to improve accessibility, safety, and user-friendliness:

1. **Wheelchair and pushchair accessibility**: improving wheelchair and pushchair accessibility (eg, using elevators instead of escalators, and ramps instead of steps) in hospitals, public spaces, and public transport.
2. **Clear and prominent signage**: installing clear and prominent signage in public spaces, public transport, and memory clinics to aid wayfinding.
3. **Wheelchair-friendly toilets**: increasing the availability of wheelchair-friendly toilets in public spaces, near memory clinics, and at train stations.

The education behavior change intervention function addresses barriers related to psychological capability, social opportunity, and both automatic and reflective motivation. Consistent with Ball et al [39], educational content should be tailored to different audiences (ie, the general public, consumers, and health care professionals). Points 1-3 outlined in Textbox 1 apply across all audiences, whereas Point 4 targets the general public and health care professionals. Building on evidence showing improved public dementia knowledge and attitude [40], education efforts, public campaigns, and media should evolve beyond depicting individuals with severe dementia to include individuals living with MCI and early stages of dementia. In doing so, they can illustrate how individuals adapt and continue to live positively and meaningfully despite cognitive decline. This shift will enable the public to gain a more accurate understanding of early cognitive impairment, emphasizing that dementia is more than just memory loss; it also involves language difficulties, problem-solving challenges, and changes in mood and behavior. By providing a broader perspective, it can increase psychological capability and reflective motivation, thereby encouraging timely medical help–seeking behavior even in the early stages of the disease.

In tandem with education, persuasion strategies can activate both automatic and reflective motivation processes. For individuals reluctant to seek medical help after receiving positive case-finding results, carefully designed result letters may serve as a behavioral nudge, drawing on psychological and behavioral science principles, such as nudges [41-43] and loss-framed messaging that emphasizes the negative consequences of delaying medical follow-up [44]. Point 2, engaging the public through expert figures and community leaders to offer insights and perspectives on early cognitive impairment, is a powerful persuasion strategy that draws on the principle of authority to engage the public and encourage medical help–seeking behavior by deferring to the experts (eg, memory specialists) [45,46], and positioning these experts as catalysts for behavior change.

The enablement intervention function is necessary to reduce structural and procedural barriers. Our study supports simplifying the referral pathways; participants receiving direct referrals to SGH memory clinics achieved a 100% attendance rate (ie, for those living in the SGH vicinity, n=10), compared to 54.6% for those with indirect referrals to memory clinics within their home vicinity (ie, participants living outside of the SGH vicinity, n=6). The differences point to the value of simplifying the referral process to reduce cognitive load and barriers in navigating the health care system. At the organizational and policy level, establishing clear practice guidelines for referrals from primary care to memory clinics would reduce practice variations across the health care system and improve the opportunity to timely access medical care.

Finally, environmental restructuring is essential to address physical opportunity barriers, particularly for individuals living with mobility challenges. Limitations in Singapore's public transport can hinder attendance at follow-up medical

XSL•FO
RenderX

appointments, particularly during peak commuting hours when public transport is crowded and less accessible. For instance, the Singapore public buses are only able to accommodate one or 2 wheelchair users at a time [47], requiring additional wheelchair users to wait for the next available bus. Environmental restructuring and structural modifications, such as increasing wheelchair capacity on buses and improving clinic accessibility, would not only ease travel burdens but also signal a societal commitment to inclusion by proactively supporting the diverse needs of people living with cognitive and physical impairments.

Limitations and Strengths

This study has several limitations. First, the sample was drawn from participants in the PENSIEVE-AI study [11], raising the possibility of self-selection bias whereby participants may have been more inclined to follow up on their cognitive impairment. To mitigate this, we purposively included individuals who had not sought medical follow-up and expressed no intention to do so unless their symptoms became more severe (eg, difficulty returning home using a familiar route). Second, the sample lacked ethnic diversity, with limited representation from Malay and Indian minority groups relative to the Chinese majority group. Third, the long intervals between memory clinic follow-ups, typically 6-12 months, introduce uncertainty regarding continued long-term medical follow-up, particularly for individuals with MCI whose symptoms are typically less severe and may be perceived as less disruptive than those associated with dementia.

A further limitation concerns the exclusion of individuals with dementia from direct interviews. Although based on ethical and practical considerations regarding potential distress and communication challenges, this warrants critical reflection on the representation gap inherent in relying on proxy accounts. This reliance placed emphasis on caregivers' experiences and perspectives in identifying barriers and facilitators of the medical help–seeking journey following dementia diagnosis, inadvertently glossing over the affective experiences and personal autonomy of the individuals with dementia themselves. In particular, their personal decision-making in medical help–seeking, sense-making of the dementia diagnosis, and challenges in navigating health care services, as well as the interactions with caregivers and health care professionals may be underrepresented, despite being highlighted in the broader dementia literature [48-50].

Prior research emphasizes the importance of including individuals with dementia to address epistemic injustice and support person-centered approaches in research [51-56]. While caregiver accounts were sufficiently informative for understanding the barriers and facilitators to medical help–seeking in this study, we acknowledge the importance and value of including the perspectives of individuals with dementia. Future studies should adopt more inclusive and adapted approaches, such as process consent, to ethically and effectively engage individuals with dementia. Finally, evaluating the feasibility and effectiveness of our recommendations was beyond the scope of this qualitative study. Future research should apply the acceptability, practicability, effectiveness, affordability, side-effects, and equity (APEASE) criteria [19] in collaboration with stakeholders, including policymakers and community representatives (eg, patient advocates, representatives from caregiver organizations, community health care workers, or volunteers), to refine and evaluate these recommendations.

Despite these limitations, this study makes several important contributions. First, while prior research has primarily focused on medical help–seeking intentions [12,29], this study explored determinants of actual medical help–seeking behavior following a cognitive impairment diagnosis. Second, by including individuals with MCI and the caregivers of individuals with MCI or dementia who had not sought medical follow-up, we addressed a key sampling gap in earlier qualitative work [16]. This provided a more nuanced understanding of the determinants of medical help–seeking behavior, particularly among those not yet engaged in follow-up care. In doing so, this study extends the AD system framework [39] by highlighting the nonlinear processes involved in seeking medical treatment.

Conclusions

This study identified key barriers and facilitators to medical help–seeking behavior following a cognitive impairment diagnosis, demonstrating the dynamic and nonlinear processes in accessing medical treatment. Building on these findings, we developed actionable recommendations grounded in the COM-B framework and the BCW intervention functions. We advocate for a whole-of-society approach, with implications extending to policymakers, primary care providers, tertiary hospitals, and the transport and media authorities. To maximize the impact of these recommendations, ongoing collaboration with stakeholders and careful evaluation using the APEASE criteria are needed to address practical considerations and promote equity in facilitating timely medical follow-up after positive case-finding results or diagnosis.

Acknowledgments

We sincerely thank the participants and caregivers for generously sharing their experiences and insights. Their willingness to contribute and the valuable input they provided were instrumental in deepening our understanding of the research topic. We are also grateful to A/Prof Philip Yap for his insightful feedback on an earlier draft of this paper, which helped refine our work. Generative artificial intelligence tools were not used at any stage in the preparation of this manuscript.

XSL•FO
RenderX

Funding

This work was part of a larger study that received funding from the Smart Nation and Digital Government Office, Prime Minister's Office, Singapore (grant I_20092346), awarded to the corresponding author, A/Prof Tau Ming Liew. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The authors confirm that the data supporting the results of this study are available within the article and/or its supplementary materials.

Conflicts of Interest

TML has provided consultation to Lundbeck and Eisai. TML is a co-inventor on a patent for a digital cognitive test (PENSIEVE-AI; patent application number PCT/SG2026/050015).

Multimedia Appendix 1
Flowchart of participants in PENSIEVE-AI tool case-finding pilot study.
[DOCX File, 289 KB - aging_v9i1e79386_app1.docx]

Multimedia Appendix 2
Interview guide using the capability, opportunity, motivation-behavior (COM-B) framework.
[DOCX File, 20 KB - aging_v9i1e79386_app2.docx]

References

1. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. The Lancet 2022;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8]
2. Prince M, Bryce R, Ferri C. World alzheimer report 2011: the benefits of early diagnosis and intervention. Alzheimer's Disease International. 2011. URL: https://www.alzint.org/u/WorldAlzheimerReport2011.pdf [accessed 2024-02-03]
3. Morley JE, Morris JC, Berg-Weger M, et al. Brain health: the importance of recognizing cognitive impairment: an IAGG consensus conference. J Am Med Dir Assoc 2015 Sep 1;16(9):731-739. [doi: 10.1016/j.jamda.2015.06.017] [Medline: 26315321]
4. Burns A, Iliffe S. Dementia. BMJ 2009 Feb 5;338(feb05 1):b75. [doi: 10.1136/bmj.b75] [Medline: 19196746]
5. Thyrian JR, Hertel J, Wucherer D, et al. Effectiveness and safety of dementia care management in primary care: a randomized clinical trial. JAMA Psychiatry 2017 Oct 1;74(10):996-1004. [doi: 10.1001/jamapsychiatry.2017.2124] [Medline: 28746708]
6. Vickrey BG, Mittman BS, Connor KI, et al. The effect of a disease management intervention on quality and outcomes of dementia care: a randomized, controlled trial. Ann Intern Med 2006 Nov 21;145(10):713-726. [doi: 10.7326/0003-4819-145-10-200611210-00004] [Medline: 17116916]
7. Cepoiu-Martin M, Tam-Tham H, Patten S, Maxwell CJ, Hogan DB. Predictors of long-term care placement in persons with dementia: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2016 Nov;31(11):1151-1171. [doi: 10.1002/gps.4449] [Medline: 27045271]
8. Jennings LA, Laffan AM, Schlissel AC, et al. Health care utilization and cost outcomes of a comprehensive dementia care program for medicare beneficiaries. JAMA Intern Med 2019 Feb 1;179(2):161-166. [doi: 10.1001/jamainternmed.2018.5579] [Medline: 30575846]
9. Spijker A, Vernooij-Dassen M, Vasse E, et al. Effectiveness of nonpharmacological interventions in delaying the institutionalization of patients with dementia: a meta-analysis. J Am Geriatr Soc 2008 Jun;56(6):1116-1128 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.01705.x] [Medline: 18410323]
10. Jia L, Du Y, Chu L, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. Lancet Public Health 2020 Dec;5(12):e661-e671. [doi: 10.1016/S2468-2667(20)30185-7] [Medline: 33271079]
11. Liew TM, Foo JYH, Yang H, et al. PENSIEVE-AI a brief cognitive test to detect cognitive impairment across diverse literacy. Nat Commun 2025 Mar 23;16(1):2847. [doi: 10.1038/s41467-025-58201-x] [Medline: 40122854]
12. Gregg JE, Simpson J, Nilforooshan R, Perez-Algorta G. What is the relationship between people with dementia and their caregiver's illness perceptions post-diagnosis and the impact on help-seeking behaviour? A systematic review. Dementia (London) 2021 Oct;20(7):2597-2617. [doi: 10.1177/1471301221997291] [Medline: 33787368]
13. Conner M, Norman P. Understanding the intention-behavior gap: the role of intention strength. Front Psychol 2022;13:923464. [doi: 10.3389/fpsyg.2022.923464] [Medline: 35992469]
14. Faries MD. Why we don't "Just Do It". Am J Lifestyle Med 2016 Sep;10(5):322-329. [doi: 10.1177/1559827616638017]
15. Sheeran P, Webb TL. The nitention–behavior gap. Soc Personal Psych 2016 Sep;10(9):503-518. [doi: 10.1111/spc3.12265]

XSL•FO
RenderX

16. Jiao YC, Chang J, Liu C, Zhou SY, Ji Y, Meng Y. Factors influencing the help-seeking behavior in patients with mild cognitive impairment: a qualitative study. BMC Health Serv Res 2023 Dec 2;23(1):1345. [doi: 10.1186/s12913-023-10281-5] [Medline: 38042819]
17. Glanz K, Rimer BK, Viswanath K. Health Behavior: Theory, Research, and Practice: John Wiley & Sons; 2015.
18. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation Sci 2011 Dec;6(1). [doi: 10.1186/1748-5908-6-42]
19. West R, Michie S, Atkins L, et al. Achieving behaviour change: a guide for local government and partners. Public Health England. 2019. URL: https://assets.publishing.service.gov.uk/media/5e7b4e85d3bf7f133c923435/PHEBI_Achieving_Behaviour_Change_Local_Government.pdf [accessed 2024-02-03]
20. Gong N, Yang D, Zou J, et al. Exploring barriers to dementia screening and management services by general practitioners in China: a qualitative study using the COM-B model. BMC Geriatr 2023 Jan 31;23(1):55. [doi: 10.1186/s12877-023-03756-x] [Medline: 36717818]
21. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions: Silverback; 2014.
22. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
23. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13(1):117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
24. Babbage D, Terry G. Thematic analysis coding management macro. : Open Science Framework; 2025. [doi: 10.17605/OSF.IO/ZA7B6]
25. What is health literacy? Centers for Disease Control and Prevention. URL: https://www.cdc.gov/health-literacy/php/about/?CDC_AAref_Val=https://www.cdc.gov/healthliteracy/learn/index.html [accessed 2024-02-03]
26. Kawachi I, Kennedy BP, Glass R. Social capital and self-rated health: a contextual analysis. Am J Public Health 1999 Aug;89(8):1187-1193. [doi: 10.2105/ajph.89.8.1187] [Medline: 10432904]
27. Forgas JP. Affect in social judgments and decisions: a multiprocess model. In: Zanna MP, editor. Advances in Experimental Social Psychology 1992:227-275. [doi: 10.1016/S0065-2601(08)60285-3]
28. Forgas JP. Mood and judgment: the affect infusion model (AIM). Psychol Bull 1995 Jan;117(1):39-66. [doi: 10.1037/0033-2909.117.1.39] [Medline: 7870863]
29. Devoy S, Simpson EEA. Help-seeking intentions for early dementia diagnosis in a sample of Irish adults. Aging Ment Health 2017 Aug;21(8):870-878. [doi: 10.1080/13607863.2016.1179262] [Medline: 27149181]
30. Hill NL, Bratlee-Whitaker E, Sillner A, Brautigam L, Mogle J. Help-seeking for cognitive problems in older adults without dementia: a systematic review. Int J Nurs Stud Adv 2021 Nov;3:100050. [doi: 10.1016/j.ijnsa.2021.100050] [Medline: 38746728]
31. Öhman A. Fear and anxiety. In: Lewis M, Haviland-Jones JM, Barrett LF, editors. Handbook of Emotions: Guilford Press; 2008:709-729.
32. Vrinten C, Waller J, von Wagner C, Wardle J. Cancer fear: facilitator and deterrent to participation in colorectal cancer screening. Cancer Epidemiol Biomarkers Prev 2015 Feb;24(2):400-405. [doi: 10.1158/1055-9965.EPI-14-0967] [Medline: 25634890]
33. Newens AJ, Forster DP, Kay DW. Referral patterns and diagnosis in presenile Alzheimer's disease: implications for general practice. Br J Gen Pract 1994 Sep;44(386):405-407. [Medline: 8790653]
34. Cooper C, Sommerlad A, Lyketsos CG, Livingston G. Modifiable predictors of dementia in mild cognitive impairment: a systematic review and meta-analysis. Am J Psychiatry 2015 Apr;172(4):323-334. [doi: 10.1176/appi.ajp.2014.14070878] [Medline: 25698435]
35. Tschanz JT, Welsh-Bohmer KA, Lyketsos CG, et al. Conversion to dementia from mild cognitive disorder: the Cache County Study. Neurology (ECronicon) 2006 Jul 25;67(2):229-234. [doi: 10.1212/01.wnl.0000224748.48011.84] [Medline: 16864813]
36. Busse A, Angermeyer MC, Riedel-Heller SG. Progression of mild cognitive impairment to dementia: a challenge to current thinking. Br J Psychiatry 2006 Nov;189(5):399-404. [doi: 10.1192/bjp.bp.105.014779] [Medline: 17077428]
37. Mitchell AJ, Shiri-Feshki M. Temporal trends in the long term risk of progression of mild cognitive impairment: a pooled analysis. J Neurol Neurosurg Psychiatry 2008 Dec;79(12):1386-1391. [doi: 10.1136/jnnp.2007.142679] [Medline: 19010949]
38. Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia--meta-analysis of 41 robust inception cohort studies. Acta Psychiatr Scand 2009 Apr;119(4):252-265. [doi: 10.1111/j.1600-0447.2008.01326.x] [Medline: 19236314]
39. Ball DE, Mattke S, Frank L, et al. A framework for addressing Alzheimer's disease: without a frame, the work has no aim. Alzheimers Dement 2023 Apr;19(4):1568-1578 [FREE Full text] [doi: 10.1002/alz.12869] [Medline: 36478657]
40. Hansra GK, Lim H, Cheong CY, Yap P. Knowledge and attitudes towards dementia among the general public in Singapore: a comparative analysis. Dement Geriatr Cogn Disord 2023;52(4):214-221. [doi: 10.1159/000530271] [Medline: 37080177]

XSL•FO
RenderX

41. Levers of behavior change: a guide to the science and applications. Rare's Center for Behavior & the Environment. 2024. URL: https://behavior.rare.org/wp-content/uploads/2024/08/Levers-of-Behavior-Change-Guide-8-2024.pdf [accessed 2024-02-03]
42. Samson A. The behavioral economics guide 2024. URL: https://www.behavioraleconomics.com/be-guide/the-behavioral-economics-guide-2024/ [accessed 2024-02-03]
43. Sunstein CR. Nudging: a very short guide. J Consum Policy 2014 Dec;37(4):583-588. [doi: 10.1007/s10603-014-9273-1]
44. Tannenbaum MB, Hepler J, Zimmerman RS, et al. Appealing to fear: a meta-analysis of fear appeal effectiveness and theories. Psychol Bull 2015 Nov;141(6):1178-1204. [doi: 10.1037/a0039729] [Medline: 26501228]
45. Cialdini RB. Harnessing the science of persuasion. Harv Bus Rev 2001(9):72-79 [FREE Full text]
46. Cialdini RB. The Science of Persuasion. Sci Am 2001 Feb;284(2):76-81 [FREE Full text] [doi: 10.1038/scientificamerican0201-76]
47. Special assistance. Singapore Bus Services Transit. URL: https://www.sbstransit.com.sg/special-assistance [accessed 2024-02-03]
48. Górska S, Forsyth K, Maciver D. Living with dementia: a meta-synthesis of qualitative research on the lived experience. Gerontologist 2018 May 8;58(3):e180-e196. [doi: 10.1093/geront/gnw195] [Medline: 28069886]
49. Mammen JR, Goldman JG, Tyo M, Xiao Y. Examining the lived experience of dementia with Lewy bodies through qualitative research: a systematic review. Alzheimers Dement 2025 May;21(5):e70217. [doi: 10.1002/alz.70217] [Medline: 40371732]
50. Prorok JC, Horgan S, Seitz DP. Health care experiences of people with dementia and their caregivers: a meta-ethnographic analysis of qualitative studies. CMAJ 2013 Oct 1;185(14):E669-E680. [doi: 10.1503/cmaj.121795] [Medline: 24003093]
51. Beuscher L, Grando VT. Challenges in conducting qualitative research with individuals with dementia. Res Gerontol Nurs 2009 Jan;2(1):6-11. [doi: 10.3928/19404921-20090101-04] [Medline: 20077988]
52. Diaz-Gil A, Brooke J, Kozlowska O, Jackson D, Appleton J, Pendlebury S. A human rights-based framework for qualitative dementia research. Nurs Ethics 2023;30(7-8):1138-1155. [doi: 10.1177/09697330231161687] [Medline: 37247413]
53. Halonen U, Aaltonen M, Aerschot LV, Pirhonen J. Participation of persons living with dementia in research: a means to address epistemic injustice. Dementia (London) 2025 Jul;24(5):850-865. [doi: 10.1177/14713012241299015] [Medline: 39510100]
54. McKeown J, Clarke A, Ingleton C, Repper J. Actively involving people with dementia in qualitative research. J Clin Nurs 2010 Jul;19(13-14):1935-1943. [doi: 10.1111/j.1365-2702.2009.03136.x] [Medline: 20920020]
55. Murphy K, Jordan F, Hunter A, Cooney A, Casey D. Articulating the strategies for maximising the inclusion of people with dementia in qualitative research studies. Dementia (London) 2015 Nov;14(6):800-824. [doi: 10.1177/1471301213512489] [Medline: 24403314]
56. Novek S, Wilkinson H. Safe and inclusive research practices for qualitative research involving people with dementia: a review of key issues and strategies. Dementia (London) 2019 Apr;18(3):1042-1059. [doi: 10.1177/1471301217701274] [Medline: 28350179]

Abbreviations

AD: Alzheimer disease
AI: artificial intelligence
APEASE: acceptability, practicability, effectiveness, affordability, side-effects, equity
BCW: behavioral change wheel
COM-B: capability, opportunity, motivation-behavior
COREQ: Consolidated Criteria for Reporting Qualitative Research
MCI: mild cognitive impairment
SGH: Singapore General Hospital

Edited by M Gray; submitted 25.Jun.2025; peer-reviewed by AV Bosse; revised version received 26.Feb.2026; accepted 11.Mar.2026; published 19.May.2026.

Please cite as:
Lim SR, Xin X, Koh GCH, Thumboo J, Liew TM
Determinants of Medical Help–Seeking Behavior Following Case Finding of Early Cognitive Impairment: Semistructured Interview Study of Patients and Caregivers
JMIR Aging 2026;9:e79386
URL: https://aging.jmir.org/2026/1/e79386
doi:10.2196/79386

XSL•FO
RenderX

© Shu Rong Lim, Xiaohui Xin, Gerald Choon Huat Koh, Julian Thumboo, Tau Ming Liew. Originally published in JMIR Aging (https://aging.jmir.org), 19.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Roadmap for Implementing Virtual Geriatric Mental Health Services for Rural Veterans: Qualitative Longitudinal Evaluation

Christine E Gould[1,2,3], PhD; Marika Blair Humber[1,2], PhD; Chalise Carlson[1,2], MA; Ranak B Trivedi[3,4], PhD; Marisa-Francesca Lindsey[1,2], BS; Althea Lloyd[5], PhD; Lynsay Paiko[1], PhD; Amanda D Peeples[6], PhD

[1]Geriatric Research Education and Clinical Center, VA Palo Alto Health Care System, 3801 Miranda Ave, Palo Alto, CA, United States

[2]Veterans Rural Health Resource Center - Salt Lake City (VRHRC-SLC), VA Salt Lake Health Care System, Salt Lake City, UT, United States

[3]Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, United States

[4]Center for Innovation to Implementation, VA Palo Alto Health Care System, Menlo Park, CA, United States

[5]Clinical Resource Hub, Veterans Integrated Service Network (VISN) 7, Atlanta, GA, United States

[6]Center for Healthcare Engagement, Research, and Promotion (CHERP), Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, United States

Corresponding Author:
Christine E Gould, PhD
Geriatric Research Education and Clinical Center, VA Palo Alto Health Care System, 3801 Miranda Ave, Palo Alto, CA, United States

Abstract

Background: Access to geriatric mental health (GMH) care is limited in rural areas. To meet this need, the Veterans Health Administration provides specialty tele-GMH care for aging rural veterans via regional telehealth hubs.

Objective: This study aims to create a roadmap describing key phases and determinants underlying the implementation and sustainment of tele-GMH services as part of a qualitative longitudinal evaluation of tele-GMH teams.

Methods: Semistructured interviews were conducted with clinicians from all 8 tele-GMH teams (n=25) at 3 time points across a 3-year period (October 2021-September 2024). Interview (n=46) data were summarized into key domains using a templated rapid qualitative approach, guided by the Consolidated Framework for Implementation Research (CFIR) 2.0. Further thematic analysis and team discussion elucidated the findings.

Results: We identified key activities and determinants of success in three phases: (1) preimplementation (engaging leaders, securing funding/hiring, and defining services); (2) implementation scale-up and expansion (advertising, addressing challenges, seeking feedback, refining, and growth); and (3) sustainment (maintenance). Activities within each phase were cyclical and iterative (ie, nonlinear). Barriers to implementation included unfamiliarity with local aging resources; facilitators included tailoring strategies and engaging referring clinicians.

Conclusions: Similar processes emerged across regions in the development and sustainment of tele-GMH services, allowing for the creation of a unified roadmap. Limitations including sampling bias are discussed. Further work could apply and evaluate the utility of the roadmap to guide creation of tele-GMH services in new regions to enhance access to specialty care for aging rural veterans.

(JMIR Aging 2026;9:e95286) doi:10.2196/95286

KEYWORDS

CFIR; rural health; older adults; telehealth; virtual care

Introduction

Over half (54%) of US veterans residing in rural areas are aged 65 years or older and face challenges in accessing and coordinating health care services [1]. These problems are compounded for aging rural veterans who experience complex mental health, neurocognitive, and medical needs that warrant specialized geriatric mental health (GMH) services. Access to GMH specialists is scarce, particularly in rural areas [2,3], but virtual care (ie, video/audio telehealth services and e-consultation) overcomes geographical barriers to accessing these services [4-6]. Virtual care has been demonstrated to be acceptable, feasible, and efficacious for older patients, caregivers, and referring clinicians [7-9]. However, barriers such as access to broadband internet and digital literacy attenuate the positive impact of virtual care for some rural veterans [10].

The Veterans Health Administration (VHA), the largest integrated health care system in the United States, is a leader in the provision of telehealth services [11] and in addressing barriers to care through national programs, such as a Digital

XSL•FO
RenderX

Divide Consult through which clinicians can request that veterans receive loaned video-enabled tablets with internet access [12]. The VHA also created regional telehealth centers to deliver virtual care to multiple health care systems in a VHA region. These regional telehealth centers, expanded broadly under the nomenclature of clinical resource hubs (CRHs), address gaps in VHA primary and mental health care access due to clinical staffing shortages caused by resignations, extended leave periods, difficulty hiring needed clinicians, or increasing enrollment of veterans [13,14]. CRH services, which initially focused on temporary clinical support based on requests from local VHA medical centers or clinics, have expanded to encompass specialty care such as cardiology [15], sleep medicine [16], and specialty mental health models (eg, tele-neuropsychology) that provide consultative care [17]. Other VHA programs that provide virtual care to older veterans occur as part of regular care from VHA medical centers, outpatient clinics, and geriatrics centers of excellence to rural clinics [18,19]. However, CRH-based services differ from other services in VHA in that these regional hubs have clinicians focused only on delivering virtual care and embedded support staff specifically trained to facilitate telehealth services [13].

One CRH specialty model is tele-GMH. Tele-GMH refers to the provision of pharmacological and nonpharmacological management of neurocognitive and psychiatric conditions in older adults with multiple medical conditions by one or more GMH clinicians via virtual modalities. One such tele-GMH consultation model led by a geriatric psychiatrist in a multistate rural region was the pioneering innovation that led to this evaluation [20,21]. This model of care mitigated perceived barriers impeding access to specialty geriatric psychiatry care, including geographic barriers and limited workforce [2,3], and was valued by caregivers and referring clinicians [22]. Regional needs and leadership preferences contributed to the initial development of diverse approaches to tele-GMH care, including single clinician-led services and interdisciplinary teams and models of care [23].

Best practices [24,25] and drivers of telehealth use, such as patient and clinician factors [26,27], have been identified; however, systematic examination of factors supporting implementation and growth of tele-GMH services is needed. We conducted a naturalistic prospective program evaluation (ie, systematic collection of data on the program to assess impact, acceptability, and implementation determinants) using qualitative methods guided by the Consolidated Framework for Implementation Research (CFIR) 2.0 [28]. CFIR allows for consideration of the complex setting within which care is delivered, spanning local clinics, virtual telemedicine hubs, and the larger VHA system. The objective of the evaluation was to identify barriers and facilitators of implementation and to synthesize findings across multiple sites to create an implementation roadmap to facilitate the future spread of tele-GMH services to additional VHA regions.

Methods

Study Design

This program evaluation used a qualitative longitudinal design, collecting data across different stages of development and implementation of the tele-GMH teams over a 3-year period. Interviews were conducted between February 2022 and September 2024 via convenience sampling. The initial study design was guided by CFIR 1.0 [29]; however, the study design was updated, and all analysis was conducted based on CFIR 2.0 [28]. Findings reported are based on CFIR 2.0.

Setting/Participants

We used a convenience sampling approach, focused on trying to reach all the clinicians from the 8 CRHs that provided GMH services or specialty geriatrics services between October 2021 and September 2024 to participate in one or more interviews. We invited 26 of the 27 clinicians; we were unable to invite one clinician because they left their VHA position before the invitation was sent. All clinicians served rural veterans, as identified using the rural-urban commuting area codes used by the VHA Office of Rural Health [1]. We refer to the services as tele-GMH, as the majority of clinicians and services are focused on GMH.

Measures

Led by the project director (CEG), the analysis team (MBH, ADP, and RBT) developed semistructured interview guides to be administered to clinicians at 3 time points corresponding to each service's startup (time point 1), development (time point 2), and sustainment (time point 3) stages. Interview guides are available from the first author (CEG) by request; the overarching purpose of these guides was related to the multiyear evaluation with varying data collected at each time point.

During the time point 1 interviews, clinicians were asked about factors related to the development of their service, reasons why they joined the CRH, services provided, common reasons for referral and referral sources, regularly used measures, team composition, and team collaboration. Overall, time point 1 focused on CFIR domains related to the innovation and characteristics of the inner setting.

Time point 2 interviews asked clinicians about planning and implementation of services, facilitators and barriers of implementation, service impacts on rural veterans and their caregivers, service impacts on referring clinicians and others in the health care system, outcomes to track, and reflections on the sustainment of the service. This interview guide included questions that mapped onto constructs in CFIR 1.0. Time point 2 interviews gathered information about changes to the innovation and the inner setting and the initial implementation process.

Time point 3 interviews focused on relationships with referral sources, service impacts on referring clinicians and rural veterans, barriers and facilitators to service delivery, and geriatric medicine needs [30] for rural veterans. This interview guide asked questions about any changes to innovation, the inner setting, and the implementation process.

XSL•FO
RenderX

We aimed to interview all clinicians at time point 1, then used more purposive sampling at time points 2 and 3 to ensure that perspectives were gathered from multiple disciplines and from every CRH team. This purposive approach for time points 2 and 3 allowed for gathering perspectives from every region with services while minimizing time burden on the clinicians delivering services.

Ethical Considerations

The project was reviewed by the Stanford University IRB, the IRB of record for VA Palo Alto Health Care System, and determined that this work did not meet the definition of human subjects research (review numbers: IRB-65150, IRB-70700, and IRB-73787) and program evaluation was conducted to inform VA operations. Union concurrence was obtained prior to conducting interviews. All participants provided verbal consent prior to their virtual interviews.

Procedures

In accordance with VHA policy, interview guides were sent to the unions for review and concurrence. Participants were emailed an invitation to engage in interviews approximately annually. One reminder email was sent to participants who did not respond to the first invitation. Interviews were conducted and recorded via videoconferencing technology. All interviewers were trained in qualitative interviewing by the qualitative methodologist (ADP).

Analysis

The evaluation team consisted of a geropsychologist/project director (CEG), project coordinators (CC and MFL), a qualitative analyst (MBH), a qualitative methodologist (ADP), geropsychologist (AL), a health psychologist (RBT), and a geropsychology postdoctoral fellow (LP). Interviews were conducted by 5 authors (CEG, MBH, CC, MFL, and LP). All evaluation team members participated in the analysis (see Multimedia Appendix 1 for reflexivity statements from the evaluation team). Biases were managed through discussion, weekly qualitative analysis meetings, and multiple authors reviewing narrative summaries.

Interviews were auto-transcribed using Microsoft Teams software. Team members proofread transcriptions against the audio recordings to ensure that transcripts were verbatim. Led by the project director (CEG), the team developed templates to summarize transcripts into domains (key topics of interest) that corresponded to the interview guides. Using rapid qualitative template-based analysis [31], one individual summarized each transcript, and a second individual reviewed the transcript summary; then both met to review the summary and resolve any discrepancies by consensus. Data saturation was determined through review of the transcript summaries and the written domain narratives that synthesized data across each summary. Review occurred in pairs; the full team discussed findings during twice-monthly meetings. Domain narratives with similar content were combined. This analysis focused on 8 domain summaries, including joining CRH; name of service; planning process; advertising; determinants; relationships with referral sources; services/challenges; and outcomes, sustainability, and service impacts. Two authors (CEG and MBH) used a matrix template based on CFIR 2.0 [28,32] to code the domain summaries to identify the key components of CFIR underlying our findings. These 2 authors also used the 8 domain narratives to generate the implementation roadmap. The roadmap was reviewed and refined during qualitative team meetings. The finalized implementation roadmap was reviewed with the project advisory team, which included 2 CRH geriatric psychiatrists, a CRH director, and a national leader in GMH.

The evaluation team developed materials summarizing key aspects related to the development and implementation of tele-GMH services in the CRHs, outlining steps sites may consider in the planning and implementation process. These materials were reviewed with CRH clinicians during monthly consultation meetings and were distributed to CRH clinicians via email. Minor changes (focused primarily on CRH-specific processes) were made based on feedback. This process enabled the evaluation team to engage in member checking [33] through sharing our understanding of implementation processes with the interviewees themselves.

Results

Overview

Forty-six interviews were conducted, representing perspectives from 25 clinicians across 8 CRHs. Most clinicians were geriatric psychiatrists (10/25, 40%) or geropsychologists (6/25, 24%). Other clinicians included neuropsychologists, clinical pharmacists, and a geriatrician (see Table 1). The CRH services primarily encompassed tele-GMH services, with later expansions including geriatric medicine. All models (n=23) included at least one GMH clinician and delivered synchronous telehealth services to home or clinic (see Table S1 in Multimedia Appendix 1). Teams provided services for veterans at spoke sites, including community-based outpatient clinics, smaller VHA medical centers, and video to home, including for patients receiving home-based primary care (HBPC) services. Two teams also provided telehealth services to community living centers (CLCs), VHA's nursing home equivalent.

XSL•FO
RenderX

Table . Interviewee characteristics (n=25).

Characteristics	Interview phases		
	Time point 1 (n=23)	Time point 2 (n=13)	Time point 3 (n=11)
Discipline			
Geriatric psychiatrist	9 (39.1)	5 (38.5)	6 (54.5)
Geropsychologist	6 (26.1)	4 (30.8)	3 (27.3)
Neuropsychologist	4 (17.4)	1 (7.7)	0 (0)
Other clinician type[a]	4 (17.4)	3 (23.1)	2 (18.2)
Model type			
Consultative (2 regions)	8 (34.8)	3 (23.1)	4 (36.4)
Hybrid (4 regions)	10 (43.5)	7 (53.8)	5 (45.5)
Continuity (1 region)	2 (8.7)	2 (15.4)	2 (18.2)
Other model type[b] (2 regions)	3 (13.0)	1 (7.7)	0 (0)
Interview duration (minutes), mean (SD)	39.35 (0.42)	38.06 (0.42)	47.35 (0.36)
Invitation response rate[c,d], n/N (%)	23/24 (95.8)	12 (92.3)	11/12 (91.7)

[a]Other category was created to help keep data anonymized for professions with lower numbers (ie, clinical pharmacist practitioners, social workers, geriatricians).

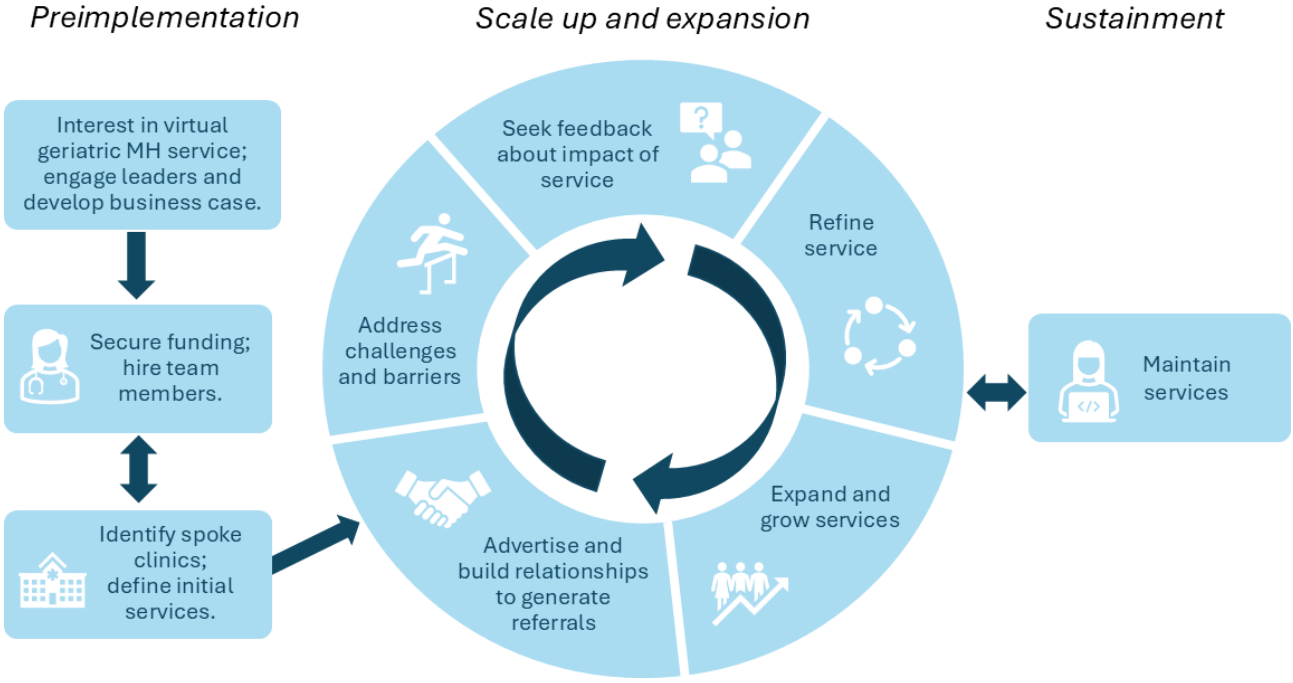
[b]Other model types include gap coverage and tele-neuropsychology services.

[c]One interviewee completed Part 1 and Part 2 interviews at the same time point as they were transitioning to a different clinical resource hub (CRH) position; the interview duration is included in mean for Part 1 interviews.

[d]Another interviewee who was the founding clinician of tele-GMH delivered services prior to 2021 and only completed a Part 3 interview.

We identified the key steps taken to implement tele-GMH services, organized into three main phases: (1) preimplementation, (2) implementation scale-up and expansion, and (3) sustainment. As seen in Figure 1, the steps were cyclical with iterative changes over time. The following describes the steps and relevant determinants of implementation in each phase, with particular attention to the implementation of care for rural veterans.

Figure 1. Overview of roadmap for implementation of tele-geriatric mental health services. MH: mental health



XSL•FO
RenderX

Preimplementation

Defining and Adapting the Innovation

The first tele-GMH service within the CRHs focused on a consultative model encompassing assessment, treatment recommendations, and limited follow-up appointments [20]. Variations on this approach emerged to include continuity-based services and a hybridized model of consultative and continuity services (termed "hybrid model") [23]. During time point 1 interviews, we asked tele-GMH clinicians whether they had followed the originating consultative model in developing their services (guidebook available internally to VHA). Clinicians reported considering the model but making adaptations based on their own clinical experience in VHA settings to develop their service.

> *…ultimately, we decided to do more of a shared intake and offer different modalities and then also provide some follow-up. We are staffed differently which allowed for the uniqueness of being able to provide follow-up.* [ID 23, clinical pharmacist practitioner]

Others described developing their approach based on specific needs for services in their region, such as neuropsychology, continuity care for older veterans who needed care following the retirement of a local psychiatrist, or CLCs that needed psychiatry coverage. Thus, tele-GMH models can adapt to fit regional needs at start-up and to meet emergent needs over time (see Refining Service).

Building the Tele-GMH Team

Building the tele-GMH team required buy-in from and ongoing relationships with regional leadership in mental health and geriatrics to secure funding for clinical positions. In some regions, CRH leadership initiated building the tele-GMH teams, whereas in other regions, a self-identified champion (often a geriatrics or GMH clinician) advocated for establishing virtual services. Tele-GMH teams often attracted innovators with GMH expertise who were motivated to design new services. The first hires (ie, founding members of the teams) described how being involved in the design of the service gave them a sense of ownership and satisfaction in expanding access to GMH services for rural populations.

The CFIR characteristics of capability and motivation emerged as important constructs among clinicians hired into tele-GMH positions. Many described their prior VHA experience, including in settings such as HBPC, CLCs, and community-based outpatient clinics. Others who were hired from outside the VHA described their experience working in the community (eg, private practice and hospital) with older adults. Motivations included clinicians being drawn to serve the geriatric community and having the opportunity to use their training and work exclusively with older patients while addressing the national GMH clinician shortage. Many also acknowledged seeking out these positions to find a better work/life balance. One geropsychologist (ID 11) described their motivation for joining their CRH team:

> *I've always enjoyed doing things with older adults that expand services for them, so I thought that this was a nice extension of that. Also, it was a completely virtual position, so I was hopeful that would give me more life balance and that kind of thing as well… I had been in a CLC setting prior and I was ready for a change after the whole COVID year and a half where nursing homes were really just hit so hard by that. So, all of those factors combined plus it just seemed like a good, exciting opportunity to do something new with geriatrics.* [ID 11, a geropsychologist]

Others described the challenges of private practice, dealing with frequent cuts to Medicare/Medicaid reimbursement, and noting that it was "just getting harder and harder" with an influx of "corporate medicine" (ID 13, geriatric psychiatrist). Attracting clinicians into tele-GMH positions most often occurred through word of mouth from professional colleagues.

Specifying the Available Services and Identifying Initial Local Sites for Service Delivery

Prospective tele-GMH service offerings were decided at the outset by CRH leaders or the tele-GMH founding members through conversations with regional stakeholders about perceived needs and specific gaps in care pathways. Regional stakeholders included leaders in geriatrics and extended care, mental health, and primary care and their teams. Subject matter experts at sites were consulted, when possible, to avoid overlap with existing GMH services in a region. In some cases, needs assessments were conducted with the general mental health clinicians in the CRH, as they were the primary mental health clinicians for some rural clinics that experience substantial turnover of clinicians and thus would be the main referral sources for some teams. Table 2 summarizes the often-delivered services by the tele-GMH teams.

XSL•FO
RenderX

Table . Tele-geriatric mental health services. Other clinical services such as geriatrics evaluations for dementia, frailty, or multicomplexity may be delivered by geriatricians or advanced practice providers (APPs) along with other interdisciplinary team members.

Clinical functions/services	Clinicians who may deliver services[a]
Behavioral/environmental interventions to address disruptive or distress behaviors	Geropsychologist, geriatric psychiatrist, or geriatrician
Capacity evaluations[b]	Geriatrician, geriatric psychiatrist, neuropsychologist, or geropsychologist
Caregiver education related to psychiatric/neurocognitive disorders	Nurse, social worker, geropsychologist, geriatric psychiatrist, geriatrician, or advanced practice provider (APP) with geriatrics training
Cognitive assessment/neuropsychological assessment	Nurse, APP with geriatrics training, geriatrician, or geriatric psychiatrist could conduct brief cognitive assessments. Neuropsychologist or geropsychologist could conduct brief cognitive assessments or multidomain cognitive assessment batteries.
Diagnostic questions / second opinions related to neuropsychiatric presentations	Geriatric psychiatrist, geriatrician, APP with geriatrics training, neuropsychologist, or geropsychologist
Psychiatric medication recommendations and/or management	Geriatric psychiatrist, clinical pharmacist, geriatrician, or APP with geriatrics training
Psychotherapy specifically tailored for aging veterans	Geropsychologist, geriatric psychiatrist (may do brief counseling), or social worker
Psychotherapy for caregivers related to caregiver strain	Geropsychologist, geriatric psychiatrist (may do brief counseling), or social worker

[a]Clinicians from other disciplines, such as occupational therapists or speech and language pathologists, may also deliver these services.

[b]Not all teams provide this clinical service. It may depend on the specific question and state-specific rules and regulations.

A key facilitator to preimplementation was having a strong planning process, including establishing a sequence of steps and identifying key players at each local site. Telehealth coordinators played important roles in the startup as well to set up the necessary agreements for services. Delegating responsibility for each step and setting a timeline for the tasks streamlined implementation of services. Open lines of communication with other CRH clinicians who previously implemented services at a spoke site helped guide implementation. Barriers to implementation emerged when the specific processes, such as creating telehealth service agreements or identifying key points of contact, were not directly followed, which then led to delays in moving forward with service delivery.

Implementation Scale-Up and Expansion

The scale-up and expansion of the tele-GMH services was an iterative process, as represented by the cyclical phase of Figure 1. See Table 3 for examples of the CFIR constructs related to the implementation process domain.

XSL•FO
RenderX

Table . Consolidated Framework for Implementation Research (CFIR) 2.0 implementation process domain analysis. One construct (assessing context) was not used when coding the data.

Implementation process domain construct	Summary of constructs within interview data	Illustrative quote
Teaming	Collaborating intentionally on tasks with referring clinicians (eg, attending staff meetings, discussing cases) helped develop rapport and established strong working relationships.	"[I attend] all of their [CLC][a] team meetings once a week, and discuss cases with them, active ones or referrals that are coming my way and that's done for an hour every week" (ID 20, geriatric psychiatrist)
Assessing needs–innovation deliverers (ie, geriatric mental health [GMH] clinicians)	Assessing needs of innovation deliverers (GMH clinicians) guided expansion of services over time. Sometimes GMH clinicians required input from local clinicians to obtain information about local services.	GMH clinicians describe the barrier of being unfamiliar with local resources: "…we're trying to refer them for adult daycare or, even a Senior Center or something like that. And we don't know anything about what is there. I think that's an ongoing challenge when not local and don't know the local resources." (ID 20, geriatric psychiatrist)
Assessing needs–innovation recipients (ie, veterans, caregivers, and referring clinicians)	Assessing needs of innovation recipients, such as patients with cognitive and/or sensory impairments, informed implementation and service delivery including time needed to troubleshoot technology and help patients connect to telehealth appointments.	"Many of these [patients] would just get canceled or no-showed… because we are aware, we take extra effort, I have none of my consults just get canceled because I can't reach [them]. We have extra steps in place where we have someone on our team to go through the chart, just making sure we're [calling] the right person and, maybe 75% of the time, turns out we're not." (ID 8, geriatrician)
Planning	Planning the steps in implementation was evident throughout the preimplementation and early implementation stages. This included planning the desired service to deliver, initiating services at a spoke site, or advertising services.	"There's been some times where some sites seemed really eager for us to get started. And so, we'll do things out of order, and then it just kind of like backfires because then we realized, oh, we don't have all the safety contacts in place, or we didn't finalize CVT availability is going to be …" (ID 4, geriatric psychiatrist)
Tailoring strategies	Tailoring strategies occurred regarding altering the planned menu of services offered to sites, adjusting time allotted to see patients, collaborating with other clinicians on shared patients, or adjusting other aspects of implementation.	"I am starting to do more briefer therapy now… while folks are inpatient. So that has been a huge change from the start of it." (ID 7, geropsychologist)
Engaging–innovation deliverers (ie, CRH[b] clinicians)	The opportunity to build an innovative service from the ground up attracted some CRH clinicians to their roles.	What drew some GMH clinicians to their role was: "the ability to kind of build something from the ground up, kind of like building a[n] airplane while flying it in a sense and being a part of something innovative." (ID 14, social worker)
Engaging–innovation recipients (ie, veterans, caregivers, and referring clinicians)	Engaging referring clinicians and continuing to remind them of the service was a critical aspect of implementation. Advertising directly to veterans was found to be challenging due to the lack of a clear VA procedure on how to do this.	Continuing to remind referring clinicians of the service's existence and purpose is key for awareness. "If the consult service rests on its laurels and doesn't put their names out there all the time, we are private practice within the VA and they [referring clinicians] aren't going to remember us or use us if they're infrequent users of us and small outpatient CBOCs …so, we need to continue to let people know that we exist." (ID 18, geriatric psychiatrist)

XSL•FO
RenderX

Implementation process domain construct	Summary of constructs within interview data	Illustrative quote
Doing	CRH clinicians adjust clinical processes and procedures throughout implementation to optimize efficiency and workflow.	Adding a phone line for patients improved coverage and responsiveness: "We've added is a phone line where patients can call in and leave a message and then our MSAs[c] will give us the message and we can follow up… And then our nurse also created a dementia education packet that included a lot of the consistent resources that we were pulling for patients and information and educational tool." (ID 23, clinical pharmacist practitioner)
Reflecting and evaluating–implementation	GMH clinicians reflected that note templates need to be standardized across sites for better workflow. GMH clinicians also described difficulties getting homebound or rural patients to attend in-person appointments like labs or EKGs.	Challenges came from trying to schedule lab work or an electrocardiogram for homebound or rural patients who are unable to come into the facility: "Lab work is difficult. They can go to lab and lab won't come to them. And EKG [electrocardiogram], workups are difficult in [the] beginning if patient cannot go and nobody can come to them." (ID 10, geriatric psychiatrist)
Reflecting and evaluating innovation	Telehealth services have been effective for older, rural patients and are valued by patients and caregivers alike. Referrals often come from clinicians who have referred to the service previously.	"[Caregivers] love [the support group] because hearing from other caregivers about how they're surviving and that it's normal to feel overwhelmed and all, that really helps." (ID 15, geropsychologist)
Adapting	Adapting includes making modifications in response to changes that have been observed during implementation. Some sites decreased advertising efforts as the referral flow became steady and robust. Others added clinicians to the CRH GMH team or provided education.	"We did reach out to just a variety of primary care staff and it was very exhausting. It was like a press tour. It was a lot of extra work outside of work hours to do so much extra effort on top of patient care. But, I think it's valuable because we've been on cruise control for a while, like we haven't advertised for a long time, and we've had plenty of volume." (ID 8, geriatrician)

[a]CLC: community living center.

[b]CRH: clinical resource hub.

[c]MSA: medical support assistant.

Advertising and Building Relationships

Generating referrals was crucial to the scale-up process. Referrals were initially increased through advertising and marketing to VA clinicians, and ongoing referrals were maintained through care collaboration with local teams. Outreach strategies included attending local and regional leadership meetings, attending local staff meetings, connecting individually with potential referring clinicians, and initiating regular contact with local sites (see Table S2 in Multimedia Appendix 1). Several tele-GMH clinicians used marketing materials, such as flyers, with service descriptions and other necessary information for sites to place consultation requests.

Despite the many facilitative strategies used to increase referrals, inconsistency in the naming of the tele-GMH service emerged as a potential barrier. Specifically, two key contributors were identified: (1) the slightly different naming conventions for consult orders across different instances of the VHA electronic health record (EHR) in each health care system, and (2) the uncertainty expressed by some interviewees when asked directly about the name of the service. This could have implications for advertising and maintaining a referral network.

Addressing Challenges and Barriers

The challenges that arose during implementation were grouped into 4 areas, including virtual care delivery, patient-related factors, availability of local services, and logistical challenges. Virtual care delivery challenges included needing to troubleshoot technology (eg, setting up assistive devices to be compatible with the technology, connectivity, and use of video platforms) before and during visits. Tele-GMH clinicians used different strategies to help facilitate access to and use of technology, including loaning VHA internet-enabled tablets (ie, Digital Divide Consult) and connecting patients with the VHA's helpdesk to troubleshoot connecting to video visits. Telehealth clinical technicians or nurse case managers helped by conducting practice calls with patients and/or their caregivers as well.

Patient-related challenges emerged for specific populations, such as patients with cognitive impairment who do not have a caregiver present to assist with troubleshooting technology or to corroborate the patient's medical history, or when working with patients with hearing impairment who are unable to reliably hear the clinicians through telehealth. Other complications arose when the point of contact listed in the electronic health record was not involved in the patient's care. To overcome some of these challenges, tele-GMH clinicians collaborated with other

XSL•FO
RenderX

VHA teams to obtain accurate information or coordinate visits. The CFIR construct of teaming, or intentional collaboration on tasks, included having local HBPC clinicians assist patients in logging on to a tablet or phone for the tele-GMH appointment.

Additional challenges arose regarding the availability of local resources. Some rural veterans had limited access to local primary care or specialty care. This made it challenging for those veterans to obtain timely and consistent medical care, which also impacts their mental health care. One geropsychologist (ID 15) described difficulties in trying to get in touch with local clinics.

> *It's scary to me, the risks that are inherent in that [scarce resources in rural areas] and feeling so separate from it and limited in what I can do, you know, because I refer people to the patient advocate. I provide those contact numbers. I really encourage them to speak up for themselves and then they'll tell me like, 'I've called [local VHA clinic] 12 times and no one's called me back.' And you know, you just feel like your hands are tied and that's really hard for my morale and their morale.* [ID 15, a geropsychologist]

Potential solutions included working with local social workers or with the CRH nurse case managers to identify local resources.

The last challenge centered on logistics, clinical processes, and procedures. Tele-GMH clinicians described making small adjustments after they began providing services. These iterative changes, captured by the CFIR construct of doing, led to changes in patient-clinician communication (eg, adding a phone line where patients could leave messages for a medical support assistant who triages requests), workload balance among team members (eg, adjusting intake appointments), improving within-team communication (eg, developing a format for internal huddles and patient discussions), and communication with sites (eg, developing internal communication channels with each local site).

Seeking Feedback About Impact of Service

Tele-GMH clinicians described ways in which they receive feedback about the impact of their services, captured by the CFIR construct of reflecting and evaluating. Positive feedback about the service encompassed effects on veterans (acceptability of telehealth, convenience of telehealth, and positive impact on quality of life); caregivers (involvement in care and their appreciation of services); referring clinicians (assistance with patients with complex needs); and access to specialty care (see Table 4).

XSL•FO
RenderX

Table . Tele-geriatric mental health service impacts.

Theme	Definition	Illustrative quote
Increased acceptability of telehealth for older veterans	Older veterans had positive responses to telehealth technology and preferred telehealth. Referring clinicians expressed initial skepticism about telehealth being effective; however, this was overcome as the virtual team proved to be helpful.	"There's still this culture or stigma around the use of technology and older adults…I think we're seeing that, yes, there's barriers and challenges for sure. It's a unique population with unique needs, but we absolutely can make this work." (ID 6, geropsychologist)
Telehealth is convenient for older veterans	The telehealth service enables patients with age-related health conditions to avoid driving far distances by offering video-to-home appointments (VA Video Connect [VVC]) and video appointments to local clinics (Clinical Video Telehealth [CVT]).	"…saving them having to drive to places that are further away, whether they're doing VVC visits from their home, which obviously is the easiest way to do it… it's been a lot easier for them not to have to drive as far as they might otherwise." (ID 20, geriatric psychiatrist)
Caregivers are often involved in and value the services	Patients are the focus of the service; and many have stressed and overwhelmed caregivers. Caregivers benefit from direct services such as caregiver groups or brief therapy, but also from education from clinicians during the veterans' appointments.	"[Caregivers] don't feel like they can be honest about how scary [caregiving] is and how taxing and overwhelming it is because they don't want to burden other people. And they feel a sense of shame sometimes because they feel like they're supposed to just be up for it all and never complain." (ID 15, geropsychologist)
Positive effects on quality of life for veterans, particularly those with cognitive impairment	The service has resulted in patients getting high quality care managing dementia, behaviors and neuropsychiatric symptoms. This care extends to quality of life for the veterans.	"…with dementia there's no cure, but we can give them the best quality of life possible and the best time with their loved ones and caregivers, the best time with their family member with dementia that we can and, and that's the ultimate goal for us." (ID 23, clinical pharmacist practitioner)
Referring clinicians appreciate the services for patients with complex needs	This service benefits referring clinicians by reinforcing their decisions and providing additional guidance when needed. Education to direct care staff also helps them manage veterans with disruptive or challenging behaviors and conflicts with staff and other patients.	"…anticipating questions that they might not realize they might have. And I think also just trying to alleviate concerns or kind of anxieties about, you know, in particular with dementia-related behaviors to manage, just reinforcing that I'm there to provide help when needed. So, not having to worry alone." (ID 4, geriatric psychiatrist)
Services increase access to needed expertise	These services increased access to geriatrics care for highly rural patients in under-resourced areas who may otherwise "fall through the cracks" in health care systems.	"[Local clinicians] may want to refer that Veteran on to a higher level of care or to community care, and that usually delays care a lot versus having that geriatric mental health specialist or geropsychologist available. You can see that patient right away. So again, it's just really makes care a lot more accessible and efficient." (ID 1, geropsychologist)

Constructive feedback was elicited when tele-GMH clinicians connected with referring clinicians to discuss the recommendations given. During those conversations, tele-GMH clinicians learned more about situations in which the referring clinicians did not follow the recommendations. One geriatric psychiatrist (ID 20) described learning that the consultative care approach conflicted with referring clinicians' preferences.

> *I've gotten some feedback from the providers or at least in the consultation that they were really hoping for somebody to kind of take over the prescribing and be the provider for the patients going forward. I just had a consult right before this call up in [STATE] where it's a primary care doctor who is not at all comfortable prescribing the meds and was really looking for me to kind of take over that… my role hasn't changed, but maybe, we should rethink that and see what's best for it. That's probably the biggest thing.* [ID 20, a geriatric psychiatrist]

The tele-GMH team reflected on the impact of their service model on patients. For example, one clinician using a consultative care approach noted that this model allows patients to be seen within 1-2 days compared with 1-2 months. Clinicians using a hybrid model focused on stabilization or more continuity services described the challenges with longer wait times (4 - 8 wk). This latter observation led to tailoring strategies of increasing the number of intake slots and broader adaptations such as advocating to hire more clinicians to the team.

Refining Service

Refining the services largely consists of the CFIR constructs of tailoring and adapting the service delivery in response to feedback, observations, and emergent needs. One such change was to tailor the specific services offered according to the local needs. For example, a geropsychologist who initially intended

XSL•FO
RenderX

to provide behavior plans and assessments expanded their services to include brief psychotherapy (4‐5 sessions). Tele-GMH clinicians described setting aside time to intentionally collaborate on care with local clinicians through teaming strategies such as attending weekly behavioral huddle meetings in CLCs. Other adjustments focused on how the team divided up cases from the referring sites. For example, one service consisting of multiple geriatric psychiatrists learned that assigning each psychiatrist a CLC to cover was less confusing and prevented mixing of treatment styles and medication choices. Teams also described that refinement encompassed learning what expertise each discipline may bring to the team, which helped their team work more effectively together. One site found that their processes became more efficient when the newly hired nurse case manager began conducting prescreening questions prior to the patient's initial appointment.

Growth and Expansion

Growth and expansion of services encompassed adding spoke sites to be served by tele-GMH teams, offering new services (eg, caregiver groups), and hiring more clinicians. One team developed a virtual interprofessional cognitive assessment clinic in response to an emergent need within a health care system following their loss of a dementia assessment clinic after the lead clinician's retirement. For teams experiencing increasing referrals or additional sites, they met these demands by hiring additional clinicians, such as psychologists, geriatric psychiatrists, clinical pharmacists, and nurse case managers. However, moving toward interprofessional care led to logistical challenges using different disciplines (eg, how neuropsychology fits into the service, how to efficiently use a new nurse manager). A geriatric psychiatrist (ID 16) described challenges in equitably assigning workload and corresponding with referring clinicians.

> *We had to do some work because you've got three different providers covering five states, multiple different sites. And if we're all seeing one or two people here or there per week, it gets challenging for the primary providers there to know 'Who just saw my patient?' and 'Who do I reach out to?'* [ID 16, a geriatric psychiatrist]

Notably, not all teams focused on expansion. Some described stability in their workload, which led to the team decreasing advertising due to a steady stream of consultations. Others reported identifying systems-based challenges, namely serving homebound or rural patients. One geriatric psychiatrist (ID 10) described the need for a social worker to help with identifying local aging services.

> *I'm a single man army here. I don't have any social worker… I know there's [a] psychologist, but sometimes getting information about resources locally [is hard]. Again, it's not one clinic, [it's] multiple clinic[s], I go to so many, so I don't know what exactly is there.* [ID 10, a geriatric psychiatrist]

A geropsychologist (ID 15) noted the absence of geriatric medicine services as a challenge:

> *…if you're going to integrate geriatric mental health, and neuropsychology, you need the primary care in place. It's kind of crazy to be siloed as geriatric mental health when there's no geriatric physical health services in the [facility].* [ID 15, a geropsychologist]

Maintenance and Sustainment of Service

The maintenance and sustainment phase encompassed stability in the interprofessional teams and focused on collaboration with local clinicians. Other key aspects of the maintenance phase were reflections on the impact of the service on patients, caregivers, referring clinicians, and an overall increased access to geriatrics specialty care, as displayed in Table 4, and found within the CFIR construct of reflecting and evaluating.

Key facilitators to maintenance of services were positive interprofessional relationships, including mentoring new team members. One geropsychologist (ID 7) said of the founding tele-GMH team member:

> *I think that she has been a gift to me with regard to just her assistance in taking me under her wing and bringing me along with her in an already established little program that she had.* [ID 7, a geropsychologist]

Similarly, a clinical pharmacy practitioner (ID 12) felt their team nurse was "fantastic… she's been instrumental for some success" because the nurse holds multiple roles, like helping psychology with overflow work, triaging patients, and contacting veterans and facilitating connecting them with tech help. Hiring of support staff, such as a dedicated medical support assistant to assist with scheduling, was deemed to be very supportive, enabling the teams' ongoing success.

Discussion

Principal Findings

Using qualitative methods with multiple interviews conducted with tele-GMH clinicians across a 3-year period, we identified the key activities and determinants of success across three phases: (1) preimplementation, (2) implementation scale-up and expansion, and (3) sustainment of tele-GMH services for older veterans.

The preimplementation phase encompassed the crucial steps of engaging leaders, securing funding/hiring, and defining services. A key determinant of success during this phase was the use of adaptations to the tele-GMH model to meet local needs and priorities. Also important were the characteristics of the individuals involved in developing and delivering the tele-GMH services. The CFIR domains of capability (ie, expertise in GMH) and motivation to care for rural and underserved older veterans were factors that led clinicians to these CRH positions. Expanding virtual positions may be one way to address the persistent geriatrics workforce shortage and enable specialist clinicians to serve the population they trained to serve. Moreover, remote work has been found to be associated with low burnout, particularly for female health care workers [34]. Thus, expanding virtual positions may help individual clinicians while also filling the geographic need for geriatrics care in rural regions [35].

XSL•FO
RenderX

The implementation scale-up and expansion phase was cyclical and iterative, focused on building a referral network and expanding or maintaining services offered to older veterans. Key determinants in this phase included building and maintaining relationships with local clinicians, teaming (or intentional collaboration), doing (making iterative changes to processes), tailoring services for sites, and adapting services in response to feedback. This fits with findings on the importance of partnership alignment, communication, and feedback as determinants of implementing comprehensive geriatrics assessment [36]. A key determinant to both preimplementation and implementation was having the tele-GMH team embedded within CRH, which allowed the clinicians to take advantage of CRH operational knowledge to initiate services. This likely allowed the tele-GMH clinicians to focus on building relationships and designing their program while CRH staff guided them in the logistics of telehealth service delivery. This echoes previous work that identified the critical role of administrative staff in supporting virtual services [37].

Tele-GMH clinicians described periods of maintenance followed by refinement and growth in the third phase. Barriers to implementation, including patient-level factors, virtual care obstacles, and unfamiliarity with local aging resources, were addressed using refinement to processes and engaging nurse case managers, social workers, and telehealth clinical technicians. As found in our previous work with surveys of referring clinicians [22], several barriers impede the uptake of services, including lack of clarity about consultation process logistics (ie, who can place consultations and from which settings), perceptions that older veterans may not benefit from telemedicine, and not following tele-GMH clinician recommendations. Facilitators of the maintenance phase include clear messaging, teaming, relationship building within teams, and care coordination. The maintenance phase warrants future study to better understand the mechanisms underlying high-functioning remote health care teams and how these teams communicate with local VHA and community clinicians. This communication and care coordination is critical as many rural veterans use both VHA and community care [38]. Furthermore, there is a need to understand patient perspectives as they pertain to working with remote teams. A mixed methods study found unclear and overlapping health care clinician roles and limited care coordination to be barriers to the effective virtual primary care for patients with chronic diseases [39].

The growing shortage of both mental health and geriatric clinicians demands innovative and scalable solutions and programs to address the needs of older adults. This work dovetails with calls for integrated geriatric clinical service lines rather than disease-focused service lines [35]. However, due to workforce shortages [2,3] and the limited availability of geriatrics clinicians in rural areas, virtual care models must be considered. Tele-GMH has the potential to be such a program due to the adaptability of the model, sustainability of service delivery with embedded administrative support, and attraction of specialist clinicians to serve older veterans. Strengths of our study include using CFIR to identify determinants of success and barriers to implementation that can be applied in similar program development in other health care systems. Although the tele-GMH service may be adapted to care for veterans with complex needs, broader initiatives to transform health care, such as “Age Friendly Health Systems,” must be pursued to improve care for all older adults [40].

Limitations

Several limitations to this qualitative evaluation warrant consideration. First, the findings are limited to perspectives on VHA and thus may not directly apply to other health care systems. Second, we do not include quantitative outcomes or perspectives from patients or caregivers, which limits the interpretation of the acceptability of the findings and impact on service recipients. Third, our findings rely on tele-GMH clinician self-report, which may not fully capture all VHA perspectives in this report (ie, referring clinician, VHA leadership), although those perspectives have been reported elsewhere [21-23]. Fourth, we did not interview every tele-GMH clinician at each time point, potentially biasing our findings. Fifth, our findings may be biased by sampling of clinicians who chose to work in these roles in the CRH. Sixth, the rapid analytic approach fit with our team-based synthesis; however, it does not allow for the line-by-line analytic depth of traditional qualitative coding approaches.

Conclusions

Our findings provide evidence of implementation factors related to virtual GMH care, which was noted as a main weakness in the broader literature on geriatrics care models [41]. Similar processes emerged across regions in the development and sustainment of tele-GMH services, which allowed for the creation of a unified roadmap. The roadmap provides a pathway that other regions may follow to create tele-GMH services to enhance access to specialty care for aging rural veterans. Our findings underscore the role of VHA as a learning health system [42] where experts and innovators can pilot and refine new models of care.

Acknowledgments

Portions of these data were presented as posters at the American Association of Geriatric Psychiatry annual meeting in March 2023 in New Orleans, LA, United States, the Academy Health Dissemination & Implementation Conference in December 2023 in Washington, DC, and the American Association of Geriatric Psychiatry annual meeting in April 2026 in Arlington, VA, United States. The authors wish to acknowledge the assistance of Ana J. Alfaro, PsyD and Rachel Van Putten, PhD, in assisting with some interviews and summarizing of interview transcripts. We appreciate the guidance of Michele Karel, PhD, ABPP, throughout the course of the project. Views expressed in this article are those of the authors and not necessarily those of the Department of Veterans Affairs or the Federal Government.

XSL•FO
RenderX

Generative artificial intelligence (AI) tools were not used in the preparation of this manuscript.

Funding

Department of Veterans Affairs, Veterans Health Administration, Office of Rural Health (NOMAD PROJFY-009747 [FY2021-2023]; NOMAD PROJFY-010206 [FY2024]).

Data Availability

The data used in this project are part of a program evaluation initiative within the Veterans Health Administration and are not publicly accessible. Those interested in the project's methodology or general findings can contact the corresponding author for further information.

Authors' Contributions

Conceptualization: CEG, ADP
Data curation: MBH, CC, MFL
Formal analysis: CEG, MBH, CC, MFL, RBT, AL, LP, ADP
Funding acquisition: CEG
Methodology: CEG, RBT, ADP
Project administration: CEG
Supervision: CEG
Visualization: CEG
Writing – original draft: CEG, MBH
Writing – review & editing: CEG, MBH, CC, MFL, RBT, AL, LP, ADP

Conflicts of Interest

All authors are employed by the Department of Veterans Affairs. No other competing interests to report.

Multimedia Appendix 1
Supplemental materials contain researcher reflexivity statement and supplemental tables 1 and 2.
[DOCX File, 25 KB - aging_v9i1e95286_app1.docx]

References

1. VHA office of rural health. US Department of Veterans Affairs. 2016 Jan 19. URL: https://www.ruralhealth.va.gov/aboutus/rural-veterans.asp [accessed 2026-06-22]
2. Conroy ML, Wilkins KM, van Dyck LI, Yarns BC. Geriatric psychiatry across the spectrum. Psychiatr Clin North Am 2022 Dec;45(4):765-777. [doi: 10.1016/j.psc.2022.07.008]
3. Juul D, Colenda CC, Lyness JM, Dunn LB, Hargrave R, Faulkner LR. Subspecialty training and certification in geriatric psychiatry: a 25-year overview. Am J Geriatr Psychiatry 2017 May;25(5):445-453. [doi: 10.1016/j.jagp.2016.12.018] [Medline: 28214074]
4. Gentry MT, Lapid MI, Rummans TA. Geriatric telepsychiatry: systematic review and policy considerations. Am J Geriatr Psychiatry 2019 Feb;27(2):109-127. [doi: 10.1016/j.jagp.2018.10.009] [Medline: 30416025]
5. Sivakumar PT, Mukku SSR, Kar N, et al. Geriatric telepsychiatry: promoting access to geriatric mental health care beyond the physical barriers. Indian J Psychol Med 2020 Oct;42(5 Suppl):41S-46S. [doi: 10.1177/0253717620958380] [Medline: 33354062]
6. Volle D, Rosen B, Doumlele K, Newkirk T, Shakhau A, Moak G. Asynchronous consultations in geriatric psychiatry: experiences in a rural academic health system. Am J Geriatr Psychiatry Open Sci Educ Pract 2024 Jun;2:43-50. [doi: 10.1016/j.osep.2024.07.002]
7. de Oliveira PBF, Dornelles TM, Gosmann NP, Camozzato A. Efficacy of telemedicine interventions for depression and anxiety in older people: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2023 May;38(5):e5920. [doi: 10.1002/gps.5920] [Medline: 37204341]
8. Dryden EM, Kennedy MA, Conti J, et al. Perceived benefits of geriatric specialty telemedicine among rural patients and caregivers. Health Serv Res 2023 Feb;58 Suppl 1(Suppl 1):26-35. [doi: 10.1111/1475-6773.14055] [Medline: 36054487]
9. Nkodo JA, Gana W, Debacq C, et al. The role of telemedicine in the management of the behavioral and psychological symptoms of dementia: a systematic review. Am J Geriatr Psychiatry 2022 Oct;30(10):1135-1150. [doi: 10.1016/j.jagp.2022.01.013] [Medline: 35241355]
10. Kaboli P, Blaine A, Mares J, Fortney J, Ono S, O'Shea AMJ. Health care access from the rural perspective: a narrative review. J Rural Health 2026 Jan;42(1):e70119. [doi: 10.1111/jrh.70119] [Medline: 41685945]

XSL•FO
RenderX

11. Heyworth L, Shah N, Galpin K. 20 years of telehealth in the veterans health administration: taking stock of our past and charting our future. J Gen Intern Med 2024 Feb;39(Suppl 1):5-8. [doi: 10.1007/s11606-024-08617-w] [Medline: 38378981]
12. Ferguson JM, Van Campen J, Slightam C, Greene L, Heyworth L, Zulman DM. Evaluation of the veterans health administration's digital divide consult for tablet distribution and telehealth adoption: cohort study. J Med Internet Res 2024 Sep 9;26:e59089. [doi: 10.2196/59089] [Medline: 39250183]
13. Burnett K, Stockdale SE, Yoon J, et al. The clinical resource hub initiative: first-year implementation of the veterans health administration regional telehealth contingency staffing program. J Ambul Care Manage 2023;46(3):228-239. [doi: 10.1097/JAC.0000000000000468] [Medline: 37079357]
14. Jaske E, Wheat CL, Rubenstein LV, et al. Understanding how contingency staffing programs can support mental health services in the veterans health administration. Telemed J E Health 2024 Jun;30(7):1857-1865. [doi: 10.1089/tmj.2023.0573] [Medline: 38563753]
15. Tisdale RL, Purmal C, Kalwani N, et al. Opportunities to address specialty care deserts and the digital divide through the veterans health administration's telehealth hub-and-spoke cardiology clinic: retrospective cohort study. J Med Internet Res 2024 Nov 28;26:e53932. [doi: 10.2196/53932] [Medline: 39607997]
16. Belkora JK, Reichert J, Williams K, et al. Implementing telesleep at veterans healthcare administration: an organizational case study of adaptation and sustainment. Front Sleep 2024 Sep 10;3:1444689. [doi: 10.3389/frsle.2024.1444689] [Medline: 39507772]
17. Messler AC, Hargrave DD, Gould CE, Carlson C, Sordahl JA. Staff perspectives on the VISN 20 tele-neuropsychology program. Fed Pract 2025 Nov;42(11):1-6. [doi: 10.12788/fp.0652] [Medline: 41884070]
18. Pimentel CB, Dryden EM, Nearing KA, et al. The role of Department of Veterans Affairs community-based outpatient clinics in enhancing rural access to geriatrics telemedicine specialty care. J Am Geriatr Soc 2024 Feb;72(2):520-528. [doi: 10.1111/jgs.18703] [Medline: 38032320]
19. Connolly SL, Stolzmann KL, Heyworth L, et al. Patient and provider predictors of telemental health use prior to and during the COVID-19 pandemic within the Department of Veterans Affairs. Am Psychol 2022;77(2):249-261. [doi: 10.1037/amp0000895] [Medline: 34941310]
20. Gould CE, Carlson C, Alfaro AJ, et al. Supporting veterans, caregivers, and providers in rural regions with tele-geriatric psychiatry consultation: a mixed methods pilot study. Am J Geriatr Psychiatry 2023 Apr;31(4):279-290. [doi: 10.1016/j.jagp.2023.01.005] [Medline: 36754647]
21. Humber MB, Carlson C, Peeples AD, et al. "She Made Me Feel Like a Valuable Part of the Care": a qualitative study of the impact of tele-geriatric psychiatry consultation on rural veterans' clinicians and caregivers. J technol behav sci 2025. [doi: 10.1007/s41347-025-00554-7]
22. Carlson C, Davis CH, Humber MB, et al. Connecting aging rural veterans with tele-geriatrics and geriatric mental health services: referring clinicians' surveyed perspectives. Telemed Rep 2026 Mar;7:26924366261431529. [doi: 10.1177/26924366261431529]
23. Gould CE, Paiko L, Carlson C, et al. Implementation of tele-geriatricmental healthcare for rural veterans: factors influencing care models. Front Health Serv 2024;4:1221899. [doi: 10.3389/frhs.2024.1221899] [Medline: 39678158]
24. Holliday AM, Hashmi AZ, Okoli-Umeweni AO, et al. American geriatrics society position statement: telehealth policy for older adults. J Am Geriatr Soc 2025 Dec;73(12):3646-3654. [doi: 10.1111/jgs.70004] [Medline: 40650623]
25. Wardlow L, Leff B, Biese K, et al. Development of telehealth principles and guidelines for older adults: a modified Delphi approach. J Am Geriatr Soc 2023 Feb;71(2):371-382. [doi: 10.1111/jgs.18123] [Medline: 36534900]
26. Connolly SL, Adusumelli Y, Azario RP, Ferris SD, Hwang AR, Miller CJ. A qualitative evidence synthesis of patient and provider attitudes toward audio-only telemental health care. Telemed J E Health 2025 Jan;31(1):3-17. [doi: 10.1089/tmj.2024.0355] [Medline: 39234786]
27. Connolly SL, Ferris SD, Azario RP, Miller CJ. Patient and provider attitudes toward video and phone telemental health care during the COVID-19 pandemic: a systematic review. Clin Psychol (New York) 2024;31(4):488-503. [doi: 10.1037/cps0000226] [Medline: 40129719]
28. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
29. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4(1):50. [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
30. Davis CH, Peeples AD, Carlson C, et al. Geriatrics needs among rural older veterans receiving virtual mental health services. J Rural Health 2026 Jan;42(1):e70124. [doi: 10.1111/jrh.70124] [Medline: 41641930]
31. Kowalski CP, Nevedal AL, Finley EP, et al. Planning for and Assessing Rigor in Rapid Qualitative Analysis (PARRQA): a consensus-based framework for designing, conducting, and reporting. Implement Sci 2024 Oct 11;19(1):71. [doi: 10.1186/s13012-024-01397-1] [Medline: 39394597]

XSL•FO
RenderX

32. Nevedal AL, Reardon CM, Opra Widerquist MA, et al. Rapid versus traditional qualitative analysis using the Consolidated Framework for Implementation Research (CFIR). Implement Sci 2021 Jul 2;16(1):67. [doi: 10.1186/s13012-021-01111-5] [Medline: 34215286]
33. Saldaña J. The Coding Manual for Qualitative Researchers, 3rd edition: SAGE Publications Ltd; 2016.
34. Brault ME, Laudermith A, Kroll-Desrosiers A. Telemedicine during COVID-19 response: a welcome shift for younger female healthcare workers. J Gen Intern Med 2023 Feb;38(3):627-632. [doi: 10.1007/s11606-022-07785-x] [Medline: 36127541]
35. Colenda CC, Applegate WB. Gluing together a fragmented healthcare system for geriatrics will be hard. It's time for united action. J Am Geriatr Soc 2024 Apr;72(4):993-1003. [doi: 10.1111/jgs.18814] [Medline: 38494999]
36. Sum G, Nicholas SO, Nai ZL, Ding YY, Tan WS. Health outcomes and implementation barriers and facilitators of comprehensive geriatric assessment in community settings: a systematic integrative review [PROSPERO registration no.: CRD42021229953]. BMC Geriatr 2022 Apr 29;22(1):379. [doi: 10.1186/s12877-022-03024-4] [Medline: 35488198]
37. Cottrell M, Burns CL, Jones A, et al. Sustaining allied health telehealth services beyond the rapid response to COVID-19: Learning from patient and staff experiences at a large quaternary hospital. J Telemed Telecare 2021 Dec;27(10):615-624. [doi: 10.1177/1357633X211041517] [Medline: 34726993]
38. Garvin LA, Pugatch M, Gurewich D, Pendergast JN, Miller CJ. Interorganizational care coordination of rural veterans by veterans affairs and community care programs: a systematic review. Med Care 2021 Jun 1;59(Suppl 3):S259-S269. [doi: 10.1097/MLR.0000000000001542] [Medline: 33976075]
39. McGraw M, Couturier Y, Gaboury I, Poirier MD, Poitras ME. The process of interprofessional collaboration in primary healthcare in a telehealth context: perspectives of patients living with chronic disease. J Interprof Educ Pract 2026 Jun;43:100813. [doi: 10.1016/j.xjep.2026.100813]
40. Wozneak KA, Jindal SK, Munro S, et al. Lessons from the Department of Veterans Affairs: a continuum of age-friendly care for older adults. J Am Geriatr Soc 2025 Feb;73(2):358-366. [doi: 10.1111/jgs.19228] [Medline: 39422433]
41. Dadich A, Lan A, Shanmugarajan S, Childs S, Alford J, Chróinín DN. Models of care for older people: a scoping review. J Am Geriatr Soc 2025 May;73(5):1588-1597. [doi: 10.1111/jgs.19371] [Medline: 39831543]
42. Kilbourne AM, Evans E, Atkins D. Learning health systems: driving real-world impact in mental health and substance use disorder research. FASEB Bioadv 2021 Aug;3(8):626-638. [doi: 10.1096/fba.2020-00124] [Medline: 34377958]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
CLC: community living center
CRH: clinical resource hub
GMH: geriatric mental health
HBPC: home-based primary care
VHA: Veterans Health Administration

Edited by H LaMonica; submitted 13.Mar.2026; peer-reviewed by BL Felker, TA Wani; revised version received 13.May.2026; accepted 13.May.2026; published 16.Jun.2026.

Please cite as:
Gould CE, Humber MB, Carlson C, Trivedi RB, Lindsey MF, Lloyd A, Paiko L, Peeples AD
A Roadmap for Implementing Virtual Geriatric Mental Health Services for Rural Veterans: Qualitative Longitudinal Evaluation
JMIR Aging 2026;9:e95286
URL: https://aging.jmir.org/2026/1/e95286
doi:10.2196/95286

© Christine E Gould, Marika Blair Humber, Chalise Carlson, Ranak B Trivedi, Marisa-Francesca Lindsey, Althea Lloyd, Lynsay Paiko, Amanda D Peeples. Originally published in JMIR Aging (https://aging.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Changes in Patterns of Health Care Use Among Older Adults in Singapore Across a Decade: Repeated Nationwide Cross-Sectional Survey

Edimansyah Abdin[1*], PhD; AshaRani P V[1*], PhD; Bernard Tan[1*], BSc; Ke Ning[1*], PhD; Peizhi Wang[1*], MPH; Saleha Shafie[1*], BSocSci; Anitha Jeyagurunathan[1*], MPhil; Janhavi Ajit Vaingankar[1*], PhD; Li Ling Ng[2*], MBBS, MMed; Kelvin Bryan Tan[3*], PhD; Paul McCrone[4*], PhD; Siow Ann Chong[1*], MBBS, MD; Mythily Subramaniam[1,5,6*], PhD

[1]Research Division, Institute of Mental Health, 10 Buangkok View, Singapore, Singapore

[2]Psychological Medicine, Changi General Hospital, Singapore, Singapore

[3]Health Economist's Office, Ministry of Health, Singapore, Singapore

[4]Institute for Lifecourse Development, Faculty of Education, Health & Human Sciences, University of Greenwich, London, United Kingdom

[5]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

[6]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[*]all authors contributed equally

Corresponding Author:
Edimansyah Abdin, PhD
Research Division, Institute of Mental Health, 10 Buangkok View, Singapore, Singapore

Abstract

Background: There is limited evidence on the patterns of health care use among older adults in Singapore over the past decade.

Objective: This study examines whether the proportion of specific services used within sociodemographic subgroups, the factors associated with health care use, and the network structure of health care use have changed over a decade.

Methods: We used data from 2 repeated nationwide cross-sectional surveys of Singapore residents aged 60 years and older, conducted in 2013 (n=2565) and 2023 (n=2010). Health care service use was obtained from an adapted version of the Client Service Receipt Inventory and analyzed using multivariable logistic regression and network analysis.

Results: The most significant changes between 2013 and 2023 were in the proportions of older adults who had consulted private hospital/clinic doctors (decreased by 4.6 percentage points; n=655, 26.5% vs n=425, 21.9%; *P*=.02), consulted public and private dentists (increased by 6.4 percentage points; n=215, 11.6% vs n=269, 18%; *P*<.001), and had health insurance (increased by 26.9 percentage points; n=1312, 59.9% vs n=1697, 86.8%; *P*<.001). Those diagnosed with high blood pressure and diabetes were more likely to consult primary care doctors, while those diagnosed with dementia and heart problems were more likely to consult public hospital doctors and health workers, visit emergency departments, and require hospitalization services in both 2013 and 2023. We also found that those diagnosed with dementia were more likely to visit intermediate and long-term care services in both 2013 and 2023. Network analysis revealed that there were strong connections between respite care, nursing homes, and hospital admissions in 2023.

Conclusions: This study provides important data on population-level changes in health care usage among older adults across the decade from 2013 to 2023, which will help policymakers ensure better health care delivery for older adults.

(JMIR Aging 2026;9:e80157) doi:10.2196/80157

KEYWORDS

health care use; dementia; aging; network analysis; medical service; health care delivery

Introduction

Background

The number of older adults aged 65 years and older is projected to more than double from 771 million in 2022 to 1.6 billion by 2050 [1], suggesting that the world's population is aging. Aging has been associated with cognitive problems, such as mild cognitive impairment (MCI) and dementia [2,3]. A previous systematic review and meta-analysis found that the global prevalence of MCI was 15.56% [3], while the Global Burden of Disease study suggests that the number of people living with dementia will increase from 57.4 million in 2019 to 152.8 million by 2050 [4]. It has also been reported that aging is associated with an increase in multimorbidity, which requires complex medical management [5]. The complexity of managing

XSL•FO
RenderX

multiple health issues and the need for frequent medical attention and management have led to significant increases in health care use, including ambulatory services [6].

Health care use patterns among aging populations are often interpreted through the Andersen Behavioral Model of Health Services Use, which emphasizes the interplay of predisposing characteristics (such as age, gender, ethnicity, and education level), enabling factors (including health insurance coverage, income, and geographical access to care), and need factors (eg, the presence of chronic conditions, such as dementia, hypertension, and diabetes) [7,8]. Although this framework has guided international research for decades, population-level evidence on how these factors evolve over time in multiethnic Asian contexts remains limited. Singapore provides a unique opportunity to examine how predisposing, enabling, and need factors shape health care use patterns across a decade of rapid demographic change and substantive policy reform.

Singapore is a city-state in Southeast Asia inhabited by a multiethnic population that is predominantly Chinese (74%), followed by Malay (13.5%), Indian (9%), and other ethnic groups (3.4%). Singapore is among the fastest-aging countries in the world. The country is experiencing rapid population aging, with the proportion of citizens aged 65 years and above projected to increase from 12.4% in 2014 to 24.1% by 2030 [9,10]. To meet the needs of an aging population, the government has formed a Ministerial Committee on Aging (MCA) to coordinate government policies and programs relating to population aging in Singapore. The MCA has launched 2 Action Plans for Successful Aging in 2015 and 2023 as part of the national blueprint to enable Singaporeans to age well. These include increasing access to aged care services, improving health care affordability through the Pioneer Generation (PG) and Merdeka Generation (MG) packages, providing more senior-friendly homes, and enhancing public infrastructure [9,11].

Singapore's health care system is a mixed public-private model [12]. Public institutions such as primary care polyclinics and hospitals provide most of the subsidized care through schemes like MediShield Life and CareShield Life, while private clinics and hospitals operate on a fee-for-service basis. To strengthen preventive care, the Ministry of Health launched HealthierSG, a national strategy that encourages enrollment with family doctors and closer integration of public and private providers [13]. The intent is to ensure regular screening, referrals to lifestyle interventions, and management of chronic conditions.

PG and MG packages were introduced by the government to support older adults with their medical bills and provide additional benefits, including annual MediSave top-ups, MediShield Life and CareShield Life premium subsidies, and outpatient care subsidies. CareShield Life, introduced in 2020, is the enhanced successor to ElderShield. It offers higher monthly payouts and lifetime coverage for individuals with severe disabilities; however, it applies only to those who turned 30 years old in 2020 onward or later. ElderShield is a long-term care insurance scheme for Singapore citizens and permanent residents who turned 40 years old before 2020, providing monthly cash payouts to support long-term care needs. Older cohorts may remain on ElderShield unless they opt into CareShield Life. MediShield Life and CareShield Life are national health and long-term care insurance schemes, respectively.

Given the demographic changes and numerous government initiatives implemented in recent years, it is important to examine how health care use patterns among older adults have evolved across this decade of demographic and policy change and what sociodemographic and clinical factors explain persistent disparities in service use.

Objectives

The objective of the study is to examine changes in health care use patterns among older adults between 2013 and 2023. Specifically, we aim to determine whether the proportion of specific services used within sociodemographic subgroups has changed; the factors associated with health care use have shifted; and the network structure of health care use has evolved over a decade.

Methods

Study Design and Setting

The Wellbeing of the Singapore Elderly (WiSE) 2023 study was conducted following the same methodology as the WiSE 2013 study [14,15]. WiSE 2023 and WiSE 2013 are population-based, cross-sectional epidemiological studies that included Singapore residents aged 60 years and older living in Singapore. These surveys used a sampling frame based on a national population registry database of all citizens and permanent residents in Singapore, which is updated regularly. A probability sample was selected using a disproportionate stratified sampling design with 12 strata defined by ethnicity (Chinese, Malay, Indian, or others) and age group (60 - 74 y, 75 - 84 y, 85 y and older). Residents aged 85 years and older, Malays, and Indians were oversampled to ensure sufficient sample size for subgroup analysis.

Ethical Considerations

The study procedures were approved by the relevant ethics committees (National Healthcare Group [NHG] Domain Specific Review Board [DSRB]; Ref: 2020/01400), and all participants or their legally acceptable representatives (where applicable) gave written informed consent before the initiation of the study. All participant data were deidentified prior to analysis, with no personally identifiable information retained. Data were stored on secure, password-protected servers accessible only to authorized members of the research team. Participants received a cash reimbursement of SGD $50 (SGD $1=US $0.78 as of October 12, 2026) upon completion of the interview.

Outcomes

Information on health care service use was obtained from an adapted version of the Client Service Receipt Inventory (CSRI) [16]. The CSRI is a standardized instrument used to collect detailed information on the types and frequency of health and social care services used by respondents over the past 3 months. Data on health care use were self-reported by participants. In this study, we focused on direct medical care services, which

XSL•FO
RenderX

refer to care provided by health professionals in the public or private sector, in outpatient or inpatient settings. This included health care services from public primary care doctors (polyclinic doctors), public hospital doctors (ie, doctors employed by the government), other public hospital health care workers (eg, physiotherapists, psychologists, nurses, medical social workers), private hospital/clinic doctors (eg, general practitioners [GPs]), other private health care workers (eg, physiotherapists, psychologists, nurses), public and private dentists, traditional healers, hospitalization in public or private hospitals, health insurance, and admissions to the emergency department.

Covariates

Data on health conditions were captured by asking participants whether a doctor had ever diagnosed them with any of the following conditions: dementia, hypertension (high blood pressure [HBP]), heart problems, stroke, diabetes, or transient ischemic attacks (TIAs). Multimorbidity (label as Morbidity in Figure 1) was defined as the presence of three or more chronic conditions. Data on sex, age group (60-74, 75-84, ≥85), ethnicity (Chinese, Malay, Indian, and others), marital status (never married, married, widowed, divorced/separated), educational level (none, some but did not complete primary, primary, secondary, and tertiary), employment status (employed, unemployed, and economically inactive, ie, homemakers and retirees), income (received income or no income), and smoking status(ever or never smoked) were also collected. Consistent with the Andersen Behavioral Model of Health Services Use, the sociodemographic variables (sex, age, ethnicity, educational level, and employment status) were treated as predisposing factors, and doctor diagnosed chronic conditions (dementia, hypertension, heart problems, stroke, diabetes, and TIAs) as need factors, whereas enabling factors (national health insurance and long-term care insurance coverage and related subsidies) were examined at the population and policy levels; this framework guided the selection of correlates entered into the multivariable models and the interpretation of the observed changes.

Statistical Analysis

Statistical analyses were carried out using the Stata (version 13.0; StataCorp LLC) and R (version 4.3.1; R Foundation for Statistical Computing). To account for the stratified, disproportionate sampling design and ensure that the survey estimates were representative of the older adult population in Singapore, weighted survey data were used to estimate health care service use and to describe the sociodemographic profile of the older adult population. A χ^2 test was used to examine changes in the proportion of health care use between the 2 surveys and within each sociodemographic subgroup. A series of multivariable logistic regression analyses were used to examine sociodemographic and clinical correlates of each health care service use within each survey year. Standard errors for logistic regression analyses were estimated using a Taylor series linearization method, while χ^2 tests were Rao-Scott corrected to obtain design-adjusted test statistics and *P* values, accounting for the complex survey design. A network analysis was used to explore the network structure of health care service use and the centrality indices by fitting Ising models to the data using the *IsingFit* R-package [17] and subsequently visualized and assessed the accuracy of edge weights by nonparametric bootstrapping using the *qgraph* and *bootnet* R-packages [18] version 4.3.1. All differences were evaluated using 2-sided tests, with statistical significance set at $P<.05$.

Results

Descriptive Statistics

A total of 4575 older adults were included in the final sample, with response rates of 65.6% and 62.7% in 2013 and 2023, respectively. Multimedia Appendix 1 shows the sociodemographic distribution of the older adult population in 2013 and 2023. The most frequently reported health care use categories among older adults in 2023 were primary care doctors (n=932, 44.5%), followed by public hospital doctors (n=678, 31.3%) and private hospital/clinic doctors (n=425, 21.9%; Table 1).

XSL•FO
RenderX

Table . Percentage of health care use among older adults in 2013 and 2023.

Service	2013, n (%)	2023, n (%)	% Difference	*P* value
Primary care doctors	1082 (42.2)	932 (44.5)	2.4	.28
Public hospital doctors	749 (28.6)	678 (31.3)	2.7	.17
Other public hospital health care workers	121 (4.0)	157 (5.9)	1.9	.04
Any public hospital doctors or health care workers	777 (29.4)	698 (32.2)	2.8	.16
Private hospital/clinic doctors	655 (26.5)	425 (21.9)	–4.6	.02
Other private health care workers	29 (1.6)	31 (1.6)	–0.1	.90
Any private hospital/clinic doctors or health care workers	673 (27.6)	438 (22.3)	–5.4	.005
Public and private dentists	215 (11.6)	269 (18.0)	6.4	<.001
Traditional healers	144 (10.4)	105 (8.4)	–2.0	.14
Day care center	19 (0.7)	32 (1.2)	0.5	.22
Respite care and nursing home	24 (0.7)	33 (0.7)	0.0	.94
Intermediate or long-term care	43 (1.4)	65 (1.9)	0.5	.26
Medications	2109 (80.9)	1620 (78.1)	–2.8	.13
Hospitalizations	156 (4.5)	152 (4.4)	–0.1	.95
A&E[a]	135 (4.3)	129 (5.1)	0.7	.40
Health insurance	443 (25.7)	528 (36.8)	11.2	<.001
MediShield/MediShield Life (from November 1, 2015, onward)	1127 (56.6)	1626 (89.7)	33.1	<.001
ElderShield	442 (24.1)	783 (51.6)	27.5	<.001
Any health insurance	1312 (59.9)	1697 (86.8)	26.9	<.001

[a]A&E: accident and emergency.

Changes in the Proportion of Health Care Use Among Older Adults Between 2013 and 2023

The percentage of older adults who consulted other public hospital health care workers increased significantly from 4.0% in 2013 to 5.9% in 2023 (increased by 1.9 percentage points; *P*=.04). The most significant changes between 2013 and 2023 were in the proportions of older adults who had consulted private hospital/clinic doctors (decreased by 4.6 percentage points; *P*=.02), who had consulted public or private dentists (increased by 6.4 percentage points; *P*<.001), and who had health insurance (increased by 26.9 percentage points; *P*<.001), including MediShield and ElderShield.

Table 2 shows the changes in the percentage of health care use within each sociodemographic subgroup between 2013 and 2023. The percentage of Indians who had consulted primary care doctors increased significantly from 46.8% in 2013 to 54.2% in 2023 (an increase of 7.4 percentage points; *P*=.02), while the percentage of Malays who had consulted public hospital doctors and health workers increased significantly from 22.2% in 2013 to 28.3% in 2023 (an increase of 6.1 percentage points; *P*=.02). The largest reduction in private hospital/clinic consultations was observed among those who had below primary education, those with secondary education, those who were employed, and those aged 60 to 74 years. The largest increase in consultations with government and private dentists was among homemakers, those with no formal education, those with

XSL•FO
RenderX

below primary education, and the widowed. For intermediate and long-term care, the largest increase was observed among those with no formal education, those with below primary education, those aged 75‐84 years old, and Indians.

XSL•FO
RenderX

Table . Changes in the percentage of health care use within each sociodemographic subgroup between 2013 and 2023.

Characteristics	Primary care doctors			Public hospital doctors and health workers			Private hospital/clinic doctors and health workers			Public and private dentists			Hospital admissions			A&E[a]			Traditional healers			Intermediate and long-term care			Health insurance		
	T1[b]	T2[c]	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]
Age group (y)																											
60-74	408	441	3.3	273	30	2.7	287	215	–72[e]	134	193	59[e]	3.1	3.4	0.3	3.7	3.9	0.2	116	8.6	–3	0.8	0.5	–03	678	889	211[e]
75-84	477	48	0.3	351	385	3.4	251	248	–03	7.2	149	77[e]	7.9	6.2	–17	5	8.1	3.1	6.7	9.2	2.5	1.5	5.5	4[e]	407	805	398[e]
≥85	418	391	–27	375	395	2	222	247	2.5	3.1	119	88[e]	105	113	0.8	11	107	–03	6.5	3.3	–32	9.2	8.6	–06	19	805	615[e]
Sex																											
Male	425	456	3.1	283	33	4.7	252	202	–5	13	175	4.5	5.7	5.7	0	4	6.2	2.2	7.7	6.1	–16	1.6	1.6	0	665	895	23[e]
Female	419	436	1.7	302	315	1.3	296	241	–55[e]	106	185	79[e]	3.5	3.3	–02	4.6	4	–06	125	104	–21	1.2	2.2	1	546	844	298[e]
Ethnicity																											
Chinese	425	442	1.7	296	321	2.5	278	227	–51[e]	124	20	76[e]	4.2	3.9	–03	4.3	5	0.7	121	9.5	–26	1.5	2	0.5	608	865	257[e]
Malay	368	409	4.1	222	283	61[e]	261	20	–61[e]	5.1	7.1	2	4.4	5.9	1.5	3.6	4.5	0.9	2	3.2	1.2	0.7	1.5	0.8	487	854	367[e]
Indian	468	542	74[e]	347	365	1.8	256	248	–08	9.6	149	53[e]	6.5	9.1	2.6	5.7	8.5	2.8	2.5	5.1	26[e]	0.7	1.8	11[e]	613	894	281[e]
Others	392	464	7.2	441	483	4.2	381	3.4	–34[e]	194	2.6	–168[e]	8.7	001	–869[e]	7.5	0	–75[e]	—[f]	—	—	—	—	—	722	993	271[e]
Marital status																											
Never married	309	423	114	227	281	5.4	236	237	0.1	174	147	–27	3	2.2	–08	2	3.9	1.9	7.2	122	5	0.1	1.8	1.7	64	873	233[e]
Married	437	446	0.9	297	325	2.8	299	232	–67[e]	133	192	59[e]	4.8	5.1	0.3	4.1	5	0.9	108	7.9	–29	1	1.5	0.5	649	884	235[e]
Widowed	429	449	2	309	337	2.8	249	209	–4	6.2	161	99[e]	4.4	3.9	–05	5.9	6.7	0.8	102	8.1	–21	3.2	4.1	0.9	427	806	379[e]
Divorced/separated	371	472	101	305	311	0.6	197	145	–52	6.9	163	9.4	3.3	2.3	–1	5.1	3	–21	116	9.2	–24	0.3	0.2	–01	656	89	234[e]
Education																											
None	442	404	–38	301	318	1.7	247	307	6	4	155	115[e]	5.2	7.8	2.6	7.8	5.6	–22	9.8	131	3.3	2.2	6.1	39[e]	388	712	324[e]
Some, but did not complete primary	446	432	–14	248	294	4.6	29	204	–86[e]	7.2	161	89[e]	2.9	6.1	3.2	3.8	6.3	2.5	9.7	5.5	–42	1.4	4.4	3[e]	476	81	334[e]
Primary	416	475	5.9	275	319	4.4	248	228	–2	6.3	9.8	3.5	4.7	3.7	–1	3.3	4	0.7	106	8.1	–25	1.6	0.6	–1	566	805	239[e]

XSL•FO
RenderX

Characteristics	Primary care doctors			Public hospital doctors and health workers			Private hospital/clinic doctors and health workers			Public and private dentists			Hospital admissions			A&E[a]			Traditional healers			Intermediate and long-term care			Health insurance		
	T1[b]	T2[c]	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]	T1	T2	%[d]
Secondary	412	475	6.3	292	312	2	29	202	–88[e]	195	203	0.8	5.4	3.1	–23	4.1	4.4	0.3	139	9.8	–41	1	1.1	0.1	788	944	156[e]
Tertiary	367	391	2.4	416	38	–36	315	237	–78	263	30	3.7	4.2	4.2	0	3.2	6.1	2.9	5.8	8.2	2.4	0.2	0.8	0.6	84	978	138[e]
Employment																											
Employed	371	416	4.5	219	254	3.5	307	224	–83[e]	159	192	3.3	2.2	2.5	0.3	2.8	3.9	1.1	123	8.8	–35	—	—	—	749	91	161[e]
Unemployed	648	538	–11	331	366	3.5	246	9.2	–64	155	156	0.1	0	156	156	0	139	139	7.9	122	4.3	—	—	—	697	863	166
Homemaker	402	468	5.6	298	286	–12	321	27	–51	7.7	202	125[e]	3.8	5.2	1.4	5.2	2.6	–26	113	9.2	–21	1.2	2.8	1.6	471	788	317[e]
Retired	469	471	0.2	353	41	5.7	219	207	–12	107	161	5.4[e]	7.1	5.7	–14	5.2	7.2	2	8.1	7.6	–05	2.7	3.7	1	551	861	31[e]
Income																											
Received no income	438	34	–98	381	284	–97	197	124	–73	6.7	143	7.6	8.6	7.3	–13	6.2	9	2.8	5.5	5	–05	6.9	108	3.9	474	74	266[e]
Received income	42	453	3.3	284	325	4.1	286	231	–55[e]	122	184	6.2[e]	3.9	4.2	0.3	4.1	4.8	0.7	11	8.7	–23	0.7	1.3	0.6	614	881	267[e]

[a]A&E: accident and emergency.

[b]T1: 2013.

[c]T2: 2023.

[d]% Difference

[e]$P<.05$.

[f]Not available.

Sociodemographic and Clinical Correlates of Health Care Service Use in 2013 and 2023

Figure 1 presents the significant sociodemographic and clinical correlates of each type of health care service use in 2013 and 2023, whereas Multimedia Appendix 2 shows the complete set of sociodemographic and clinical factors for each health care service use in both years. We found that certain sociodemographic and clinical factors remained significantly associated with health care service use over time. For example, after adjusting for all covariates in the multivariable logistic regression analysis, those of Malay ethnicity (vs Chinese) were less likely to consult primary care doctors, while those who had HBP and diabetes were more likely to consult primary care doctors in both 2013 and 2023. We found that those diagnosed with dementia or heart problems and those retired (vs employed) were more likely to consult public hospital doctors and health workers in both 2013 and 2023. We also found that those diagnosed with dementia were more likely to use intermediate and long-term care in both 2013 and 2023. Furthermore, those of Malay and Indian ethnicity (vs Chinese) and with lower education (none, some, primary vs tertiary) were less likely to consult public and private dentists in both 2013 and 2023.

XSL•FO
RenderX

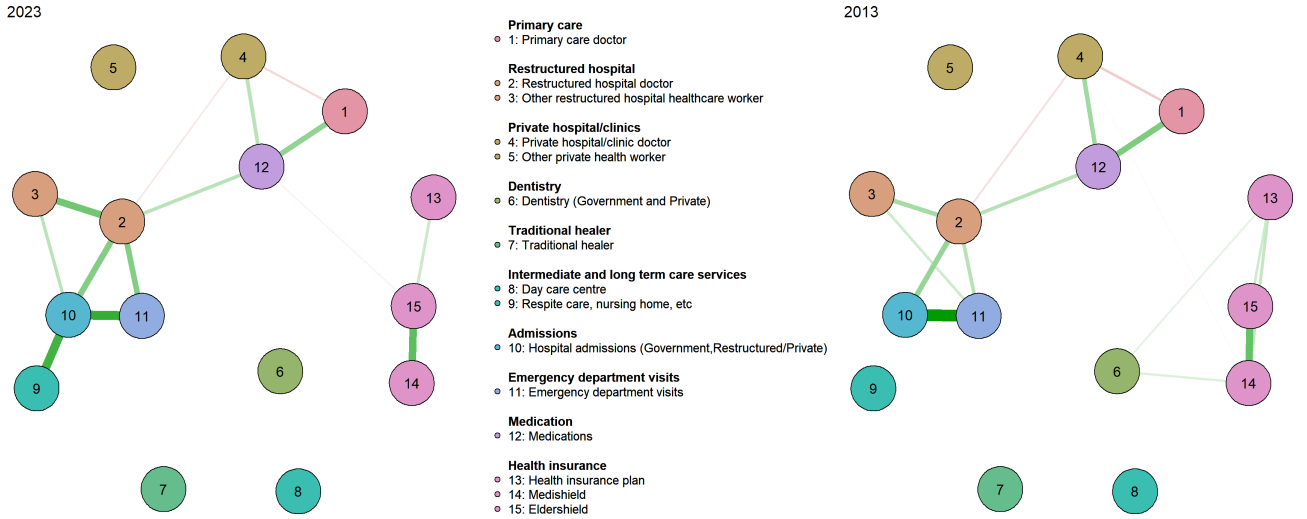
Figure 1. Significant sociodemographic and clinical correlates of each health care service use in 2013 and 2023. A&E: accident and emergency; HBP: high blood pressure; TIA: transient ischemic attack.

XSL•FO
RenderX

The Network Structure of Health Care Service Use in 2013 and 2023

Figure 2 shows the pattern of the network structure of health care service use among older adults in 2013 and 2023, while Multimedia Appendix 3 shows the 95% bootstrapped CIs for all edge weights or connections among health care services. Each node in the network depicts a health care service use, and each edge represents a bidirectional partial relationship between services, after controlling for all other associations in the network. In 2023, there was a strong positive connection between the use of emergency department visits and hospital admissions, followed by strong connections between hospital admissions and respite care and nursing homes and between MediShield and ElderShield insurance. In 2013, a strong positive connection was observed between emergency department visits and hospital admissions, followed by strong positive connections between MediShield and ElderShield and between primary care doctor and medication use.

Figure 2. The pattern of the network structure of health care service use among older adults in 2013 and 2023.



Discussion

Principal Findings

In this study, we compared 2 nationwide cross-sectional surveys of older adults conducted 10 years apart and identified three main findings. First, the types of health care older adults used changed over the decade. Fewer people went to private hospitals or clinics, while more people used dental services and were covered by national health insurance and long-term care insurance. Second, the factors associated with health care use were broadly consistent across the 2 surveys; health problems such as dementia, HBP, and diabetes were the strongest correlates of service use. Meanwhile, certain groups such as Malays, Indians, and people with lower education continued to face gaps in dental and preventive care. Third, the network structure of health care use changed over the decade, showing stronger and more stable connections among services in 2023. Strong networks appeared between respite care, nursing homes, and hospital admissions, showing that these services were being used together more often in 2023.

Our results align closely with the Andersen Behavioral Model of Health Services Use [7], where need factors such as dementia, hypertension, and diabetes were the strongest and most consistent predictors of service use across both survey waves, while predisposing factors such as ethnicity and education explained persistent disparities in dental and preventive care use. The substantial expansion of enabling factors over the decade, particularly with MediShield Life and CareShield Life coverage, corresponds directly with the observed population-level shifts in insurance uptake and the reduction in private provider use.

The findings of this study have relevance beyond Singapore. Singapore's experience offers valuable lessons, especially for other rapidly aging middle- and high-income settings, including South Korea, Taiwan, Japan, and parts of the Middle East, where demographic aging is accelerating but repeated, population-level data on health care usage patterns remain sparse. Four insights may have broad generalizability. First, mandatory enrollment in basic health insurance such as MediShield Life, combined with expanded voluntary uptake of long-term care insurance among older cohorts, can rapidly and substantially expand population-level coverage within a short policy time frame [19,20]. Second, expansion of health care subsidies and financial incentives can substantially shift use patterns across public and private sectors, highlighting how strongly cost drives health care decisions [19,20]. Third, network analysis provides a valuable and underused tool for detecting structural changes in care integration pathways at the population level. Finally, the persistent ethnic disparities in dental service use observed in this study are relevant to multiethnic societies globally, where minority populations continue to face access barriers that policy interventions have not fully overcome [21].

Changes in the Proportion of Health Care Use

This study examined population-level changes in health care use among the older adult population in Singapore between 2013 and 2023. Over this 10-year period, primary care doctors and public hospital doctors remained the most commonly used health care providers among older adults in Singapore, underscoring their central role in the health care delivery system

XSL•FO
RenderX

in Singapore. We found that the proportion of older adults who visited private hospitals or clinic doctors declined significantly in 2023 (a decrease of 4.6 percentage points; *P*=.02), suggesting a shift in health care–seeking behavior. This trend may reflect the impact of public health initiatives, including expanded subsidies and improved accessibility and capacity of primary care services and community hospitals nationwide. The increasing reliance on public health care resources among older adults aligns with national strategies such as HealthierSG [19] and the expansion of Community Health Assist Scheme (CHAS) subsidies. Such a shift has the potential to enhance equity in health care access, reduce out-of-pocket expenditures, and optimize resource use. Policymakers should continue to monitor these trends to ensure that public health care services remain adequately resourced and responsive to the evolving needs of Singapore's aging population.

Notably, we observed a significant increase in the use of dental services, likely driven by government subsidies for private clinics through the introduction of CHAS dental benefits in 2012. Additionally, we observed a substantial increase in health insurance coverage among older adults following the implementation of national health insurance schemes, including MediShield Life [22]. MediShield Life was introduced on November 1, 2015, and covers all Singapore Citizens and Permanent Residents. The marked increase in long-term care insurance coverage in 2023 (27.5 percentage points) may reflect both continued ElderShield coverate among younger cohorts, and voluntary uptake of CarShiled Life along generous subsidies and incentives may have further increased long-term care insurance coverage.

Emergency department visits and hospitalizations were among the least frequently used direct medical services among older adults in Singapore. The proportion of hospitalizations (4.5% in 2013 and 4.4% in 2023) in our study was slightly higher than that in South Korea (2%) and lower than in Mexico and the United States (>5%) [23]. The lower rates of hospitalizations and emergency department visits among the older adult population may reflect the combined effects of system-level and population-level factors, including improvements in the quality of primary care as well as broader health determinants, population health improvements, expanded insurance coverage, and changes in care integration [19,24]. Previous studies have suggested that stronger primary care systems are generally associated with better population health outcomes, including lower rates of hospitalization and mortality [25,26]. These findings highlight the importance of continued investment in primary care to support healthy aging and reduce the burden on acute care services.

Subgroup-Level Changes and Factors Associated With Health Care Use

In the multivariate analysis, we found that older adults with high blood pressure and diabetes were more likely to consult primary care doctors in both 2013 and 2023, suggesting a significant role of primary care in treating patients with these conditions. In Singapore, primary care institutions have adopted the Chronic Care Model (CCM) for care delivery [27]. Previous studies have reported that the CCM is an effective framework for improving the quality of diabetes and hypertension care in primary care settings [28]. Moreover, diabetes is one of the chronic conditions projected to increase health care expenditure and lead to greater productivity losses in Singapore [29] due to an uptrend in diabetes-related complications [30]. Hence, improving service delivery and increasing awareness through ongoing national campaigns such as the "War on Diabetes" [31] to promote active lifestyles, early screening and intervention, and better disease control should continue to combat diabetes and help the older adult population live healthier.

Our findings showed that those diagnosed with dementia and heart problems were more likely to consult public hospital doctors and health care workers in both 2013 and 2023. We also found that those diagnosed with dementia were more likely to visit intermediate and long-term care services in both 2013 and 2023. Dementia is a complex condition that requires a multidisciplinary team with specialized expertise and resources, including geriatric specialists from public hospitals. It has been suggested that patients with dementia are frequently admitted as emergency cases [32,33]. Similarly, heart problems are complex conditions that often require cardiologists and specialized services, such as cardiac rehabilitation programs that are available in public hospitals.

Malay older adults were consistently less likely to consult primary care doctors and public or private dentists in both 2013 and 2023. They were also less likely to use public hospital doctors and health workers and health insurance in 2013. This pattern suggests lower overall health care use among Malay older adults compared to their Chinese counterparts, potentially reflecting disparities in health literacy, life expectancy, and cultural differences in health beliefs and health-seeking behavior. For example, the National Population Health Survey 2024 [34] reported that the Malay ethnic group consistently exhibited the lowest rates of screening uptake for diabetes, hypertension, and high blood cholesterol compared to their Chinese and Indian counterparts [34]. Guo et al [35] similarly projected that the lowest lifetime healthcare expenditure among Malay older adults, despite a higher chronic disease burden, which attributed to shorter life expectancy. In contrast, Indian older adults were more likely than their Chinese counterparts to be admitted in 2023. This finding is consistent with Guo et al [35] projection that lifetime healthcare expenditure would be the highest among Indian older adults in Singapore, which may be attributed to the relatively higher prevalence of chronic medical conditions such as diabetes and hypertension in this group [36,37], as well as lifestyle factors.

Older adults with lower education showed a larger reduction in private hospital/clinic consultations and a greater increase in government and private dental consultations, as well as intermediate and long-term care. These patterns suggest a shift away from private providers toward subsidized public services, likely influenced by CHAS dental benefits and expanded subsidies under HealthierSG. The increased use of long-term care reflects the vulnerability of less educated groups, underscoring the importance of schemes like CHAS dental benefits and HealthierSG in reducing barriers to accessing care. Importantly, these initiatives are supported by a multipronged communication effort. For example, eligible seniors often

XSL•FO
RenderX

receive direct mail and SMS invitations to enroll in HealthierSG and to understand their CHAS, PG, and MG card benefits. Information is also disseminated through official government websites (eg, Ministry of Health [MOH] and HealthierSG), as well as the HealthHub and Healthy 365 mobile apps, which provide digital health plans and activity tracking. At the same time, traditional methods remain crucial for reaching less tech-savvy seniors, complemented by dedicated helplines such as the MOH hotline and the CHAS hotline, which offer personalized assistance for enrollment and subsidy queries.

Network Analyses

Our network analysis shows that there were strong connections among emergency department visits, hospital admissions, and seeing public hospital doctors in both 2013 and 2023. Interestingly, we only observed strong connections between respite care and nursing homes and hospital admissions in 2023. This finding is consistent with a previous study, which indicated that higher hospital admissions and transfers to hospital for acute care were common among nursing home residents [38]. This suggests a significant shift in service use, which warrants further investigation. Bootstrapped CI findings showed that there were more significant and stable network connections between health care services in 2023 than in 2013. This finding could be related to the broadened health care coverage and expansion of care integration between multidisciplinary teams from primary care, public hospitals, and intermediate and long-term care institutions for the older adult population in Singapore. However, further studies using network analysis from a qualitative perspective might be needed to better understand this relationship [39].

Limitations and Strengths

The present study has some limitations that must be considered while interpreting the findings. First, data on direct health care use were self-report, which may be influenced by recall and memory biases. Second, the edge weights in the network analysis, which represent the strength of co-utilization between health care services, may be influenced by small sample sizes. Health care services with low usage (eg, day care, respite care, or intermediate and long-term care) were reported by relatively few respondents, which could affect the stability and interpretability of their associated edges. While all edges reflect observed co-occurrence patterns, those derived from smaller subsamples should be interpreted with caution. To address this, we applied bootstrapping to enhance the robustness of our estimates. Notwithstanding the limitations, the present study has notable strengths: the data comprise 2 large, nationally representative, multiethnic samples, and a comprehensive assessment of service use was conducted using standardized instruments for comparability with other population-based studies.

Conclusions

Our study findings address important gaps in the literature on health care use in aging populations. To our knowledge, this is one of the few studies to provide repeated cross-sectional, population-level estimates of direct health care service use a decade apart in a multiethnic Asian older adult population and the first to apply network analysis methods to understand the links between different types of health care services used by older adults in Singapore. These findings contribute to 3 bodies of evidence: health care system responsiveness to population aging, equity research on persistent and emergent disparities in service access, and methodological work on network analysis approaches to understanding care integration. Future research should examine whether observed trends are sustained and whether care integration initiatives translate into improved health outcomes for older adults.

Funding

The study was funded by a grant from the Ministry of Health, Singapore. The funder had no role in the study design, in the collection, analysis, or interpretation of the data, in the writing of the report, or in the decision to submit the article for publication.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to institutional data protection policies and participant confidentiality requirements. Researchers who wish to access the anonymized data for research purposes may contact the corresponding author (EA) or the senior author (MS). Access may be granted upon reasonable request and is subject to approval by the relevant institutional review boards.

Authors' Contributions

Conceptualization: EA, LLN, KBT, PM, SAC, MS
Data curation: PVA, BT, PW, SS, AJ, JAV
Formal analysis: EA
Funding acquisition: MS
Investigation: PVA, BT, KN, PW, SS, AJ, JAV
Methodology: EA
Project administration: PVA, MS
Supervision: PM, SAC, MS
Writing – original draft: EA
Writing – review & editing: EA, PVA, BT, KN, PW, SS, AJ, JAV, LLN, KBT, PM, SAC, MS

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sociodemographic distribution of the older adults in 2013 and 2023.
[DOCX File, 19 KB - aging_v9i1e80157_app1.docx]

Multimedia Appendix 2
Sociodemographic and clinical correlates of health care service use in 2013 and 2023.
[DOCX File, 34 KB - aging_v9i1e80157_app2.docx]

Multimedia Appendix 3
95% bootstrapped CI results for all edge weights or connections among health care service use.
[PNG File, 146 KB - aging_v9i1e80157_app3.PNG]

References

1. United Nations Department of Economic and Social Affairs, Population Division. World population prospects 2022: summary of results. : United Nations; 2022 URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf [accessed 2026-10-01]
2. Chen P, Cai H, Bai W, et al. Global prevalence of mild cognitive impairment among older adults living in nursing homes: a meta-analysis and systematic review of epidemiological surveys. Transl Psychiatry 2023 Mar 11;13(1):88. [doi: 10.1038/s41398-023-02361-1] [Medline: 36906613]
3. Bai W, Chen P, Cai H, et al. Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: a meta-analysis and systematic review of epidemiology studies. Age Ageing 2022 Aug 2;51(8):afac173. [doi: 10.1093/ageing/afac173] [Medline: 35977150]
4. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
5. Kadambi S, Abdallah M, Loh KP. Multimorbidity, function, and cognition in aging. Clin Geriatr Med 2020 Nov;36(4):569-584. [doi: 10.1016/j.cger.2020.06.002] [Medline: 33010895]
6. Bierman AS, Wang J, O'Malley PG, Moss DK. Transforming care for people with multiple chronic conditions: Agency for Healthcare Research and Quality's research agenda. Health Serv Res 2021 Oct;56 Suppl 1(Suppl 1):973-979. [doi: 10.1111/1475-6773.13863] [Medline: 34378192]
7. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [doi: 10.2307/2137284] [Medline: 7738325]
8. Yin Z, Stratton L, Song Q, et al. Perceptions and utilization of online peer support among informal dementia caregivers: survey study. JMIR Aging 2024 May 31;7:e55169. [doi: 10.2196/55169] [Medline: 38825836]
9. 2023 action plan for successful ageing. : Ministry of Health, Singapore; 2023 URL: https://isomer-user-content.by.gov.sg/3/b1fd5713-8ff9-46d5-9911-0f233f2a8b31/refreshed-action-plan-for-successful-ageing-2023.pdf [accessed 2026-10-01]
10. Longevity. National Population and Talent Division. URL: https://www.population.gov.sg/our-population/population-trends/longevity/ [accessed 2026-10-01]
11. Schleich AM, editor. Perspectives of Two Island Nations: Singapore – New Zealand: World Scientific; 2024. [doi: 10.1142/13709]
12. Lim MK. Transforming Singapore health care: public-private partnership. Ann Acad Med Singap 2005 Aug;34(7):461-467. [Medline: 16123824]
13. Foo CD, Woon TH, Chia HX, et al. Healthier SG: a gateway for evolving public-private-population partnerships in population health. Lancet Reg Health West Pac 2025 Jul;60:101606. [doi: 10.1016/j.lanwpc.2025.101606] [Medline: 40605968]
14. Subramaniam M, Chong SA, Vaingankar JA, et al. Prevalence of dementia in people aged 60 years and above: results from the WiSE study. J Alzheimers Dis 2015;45(4):1127-1138. [doi: 10.3233/JAD-142769] [Medline: 25672767]
15. Subramaniam M, Abdin E, Asharani PV, et al. Prevalence of dementia in Singapore: changes across a decade. Alzheimers Dement 2025 Feb;21(2):e14485. [doi: 10.1002/alz.14485] [Medline: 39868432]
16. Beecham J, Knapp M. Costing psychiatric interventions. In: Thornicroft G, Brewin CR, Wing J, editors. Measuring Mental Health Needs: Gaskell; 1992:163-183.
17. van Borkulo CD, Borsboom D, Epskamp S, et al. A new method for constructing networks from binary data. Sci Rep 2014 Aug 1;4:5918. [doi: 10.1038/srep05918] [Medline: 25082149]
18. Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph: network visualizations of relationships in psychometric data. J Stat Soft 2012;48(4):1-18. [doi: 10.18637/jss.v048.i04]

XSL•FO
RenderX

19. Foo CD, Chia HX, Teo KW, et al. Healthier SG: Singapore's multi-year strategy to transform primary healthcare. Lancet Reg Health West Pac 2023 Aug;37:100861. [doi: 10.1016/j.lanwpc.2023.100861] [Medline: 37529631]
20. White paper on healthier SG. : Ministry of Health, Singapore; 2022 URL: https://file.go.gov.sg/healthiersg-whitepaper-pdf.pdf [accessed 2026-06-11]
21. Chen AM. Practical barriers to health equity: challenges, opportunities, and solutions. Adv Public Health 2025 Jul;2025(1):2257356. [doi: 10.1155/adph/2257356]
22. Earn LC. International profiles of health care systems, 2020. : The Commonwealth Fund; 2020 URL: https://www.commonwealthfund.org/sites/default/files/2020-12/International_Profiles_of_Health_Care_Systems_Dec2020.pdf [accessed 2026-10-01]
23. Macinko J, Cristina Drumond Andrade F, Bof de Andrade F, Lima-Costa MF. Universal health coverage: are older adults being left behind? Evidence from aging cohorts in twenty-three countries. Health Aff (Millwood) 2020 Nov;39(11):1951-1960. [doi: 10.1377/hlthaff.2019.01570] [Medline: 33136504]
24. Expanding healthcare capacity and transforming the healthcare workforce. Ministry of Health, Singapore. 2025. URL: https://www.moh.gov.sg/newsroom/expanding-healthcare-capacity-and-transforming-the-healthcare-workforce/ [accessed 2025-12-09]
25. Shi L. The impact of primary care: a focused review. Scientifica (Cairo) 2012 Dec 31;2012:432892. [doi: 10.6064/2012/432892] [Medline: 24278694]
26. Niti M, Ng TP. Avoidable hospitalisation rates in Singapore, 1991-1998: assessing trends and inequities of quality in primary care. J Epidemiol Community Health 2003 Jan;57(1):17-22. [doi: 10.1136/jech.57.1.17] [Medline: 12490643]
27. Zheng LF, Koh YLE, Sankari U, Tan NC. Asthma care based on Chronic Care Model in an aging Asian community. NPJ Prim Care Respir Med 2019 May 3;29(1):16. [doi: 10.1038/s41533-019-0130-1] [Medline: 31053715]
28. Goh LH, Siah CJR, Tam WWS, Tai ES, Young DYL. Effectiveness of the chronic care model for adults with type 2 diabetes in primary care: a systematic review and meta-analysis. Syst Rev 2022 Dec 15;11(1):273. [doi: 10.1186/s13643-022-02117-w] [Medline: 36522687]
29. Png ME, Yoong J, Phan TP, Wee HL. Current and future economic burden of diabetes among working-age adults in Asia: conservative estimates for Singapore from 2010-2050. BMC Public Health 2016 Feb 16;16(1):153. [doi: 10.1186/s12889-016-2827-1] [Medline: 26880337]
30. Tan JK, Salim NNM, Lim GH, Chia SY, Thumboo J, Bee YM. Trends in diabetes-related complications in Singapore, 2013-2020: a registry-based study. PLoS One 2022;17(10):e0275920. [doi: 10.1371/journal.pone.0275920] [Medline: 36219616]
31. Speech by minister for health, Mr Gan Kim Yong, at the MOH Committee of Supply debate 2016. Ministry of Health, Singapore. 2016. URL: https://www.moh.gov.sg/newsroom/speech-by-minister-for-health-mr-gan-kim-yong-at-the-moh-committee-of-supply-debate-2016/ [accessed 2026-10-01]
32. Natalwala A, Potluri R, Uppal H, Heun R. Reasons for hospital admissions in dementia patients in Birmingham, UK, during 2002-2007. Dement Geriatr Cogn Disord 2008;26(6):499-505. [doi: 10.1159/000171044] [Medline: 19005254]
33. Gerlach LB, Martindale J, Bynum JPW, Davis MA. Characteristics of emergency department visits among older adults with dementia. JAMA Neurol 2023 Sep 1;80(9):1002-1004. [doi: 10.1001/jamaneurol.2023.2244] [Medline: 37486693]
34. National population health survey 2024 (household interview and health examination). : Disease Policy and Strategy Division and Health Analytics Division, Ministry of Health; Policy, Research & Surveillance Group, Health Promotion Board, Singapore; 2025 URL: https://isomer-user-content.by.gov.sg/3/38583351-2fdc-4a01-8bf7-88d24889af78/NPHS%202024%20Survey%20Report.pdf [accessed 2026-10-01]
35. Guo X, Tysinger B, Wee HL, et al. Disease burden, lifetime healthcare cost and long-term intervention impact projections among older adults in Singapore. Nat Aging 2025 Jul;5(7):1358-1369. [doi: 10.1038/s43587-025-00915-0] [Medline: 40664834]
36. Seow LSE, Subramaniam M, Abdin E, Vaingankar JA, Chong SA. Hypertension and its associated risks among Singapore elderly residential population. Journal of Clinical Gerontology and Geriatrics 2015 Dec;6(4):125-132. [doi: 10.1016/j.jcgg.2015.05.002]
37. Yeo KK, Tai BC, Heng D, et al. Ethnicity modifies the association between diabetes mellitus and ischaemic heart disease in Chinese, Malays and Asian Indians living in Singapore. Diabetologia 2006 Dec;49(12):2866-2873. [doi: 10.1007/s00125-006-0469-z] [Medline: 17021918]
38. Goh HS, Tan S, Khaw AK, Ow CC, Devi MK. Hospital admissions among nursing home residents in Singapore: a retrospective study. Singapore Nurs J 2018;45(3):10-16 [FREE Full text]
39. Vos J, Gerling K, Linehan C, Siriwardena AN, Windle K. Understanding care navigation by older adults with multimorbidity: mixed-methods study using social network and framework analyses. JMIR Aging 2018 Nov 14;1(2):e11054. [doi: 10.2196/11054] [Medline: 31518236]

Abbreviations

CCM: Chronic Care Model

XSL•FO
RenderX

CHAS: Community Health Assist Scheme
CSRI: Client Service Receipt Inventory
DSRB: Domain Specific Review Board
GP: general practitioner
HBP: high blood pressure
MCA: Ministerial Committee on Aging
MCI: mild cognitive impairment
MG: Merdeka Generation
MOH: Ministry of Health
NHG: National Health Care Group
PG: Pioneer Generation
TIA: transient ischemic attack
WiSE: Wellbeing of the Singapore Elderly

Edited by C Lee, Y Jiang; submitted 05.Jul.2025; peer-reviewed by H Rafman, VH Gaidhane; revised version received 22.Sep.2026; accepted 22.Sep.2026; published 05.Nov.2026.

Please cite as:
Abdin E, P V A, Tan B, Ning K, Wang P, Shafie S, Jeyagurunathan A, Vaingankar JA, Ng LL, Tan KB, McCrone P, Chong SA, Subramaniam M
Changes in Patterns of Health Care Use Among Older Adults in Singapore Across a Decade: Repeated Nationwide Cross-Sectional Survey
JMIR Aging 2026;9:e80157
URL: https://aging.jmir.org/2026/1/e80157
doi:10.2196/80157

© Edimansyah Abdin, AshaRani P V, Bernard Tan, Ke Ning, Peizhi Wang, Saleha Shafie, Anitha Jeyagurunathan, Janhavi Ajit Vaingankar, Li Ling Ng, Kelvin Bryan Tan, Paul McCrone, Siow Ann Chong, Mythily Subramaniam. Originally published in JMIR Aging (https://aging.jmir.org), 5.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Factors Associated With Patient Portal and Synchronous Telehealth Use Across Age Groups in the Postpandemic Era: Retrospective Analysis of the Health Information National Trends Survey

Jaeyoung Park, PhD; Su-I Hou, DrPH, CPH, MCHES, RN, FACHE, FGSA

School of Global Health Management and Informatics, University of Central Florida, 528 W Livingston St, Orlando, FL, United States

Corresponding Author:
Su-I Hou, DrPH, CPH, MCHES, RN, FACHE, FGSA
School of Global Health Management and Informatics, University of Central Florida, 528 W Livingston St, Orlando, FL, United States

Abstract

Background: Since the COVID-19 pandemic, telehealth has become a core component of modern health care, encompassing both synchronous services (real-time video or phone) and patient portals, which offer a wide range of online features to streamline care delivery, including asynchronous communication and access to medical records. Older adults often face greater barriers to these technologies, potentially widening health care access gaps.

Objective: This study examines factors influencing the adoption of patient portals and synchronous telehealth in the postpandemic era. It compares the 2 modalities and identifies distinct usage patterns across age groups to inform targeted strategies.

Methods: We analyzed data from the Health Information National Trends Survey (HINTS) Cycles 6 and 7. Outcomes over the past 12 months included (1) frequency of patient portal use (continuous), (2) use of synchronous telehealth (binary), and (3) no engagement with any telehealth services (binary). Key variables included demographics, socioeconomic status, technology familiarity, and health care behaviors. Age groups were categorized as young (18–49 years), middle-aged (50–64 years), and older adults (≥65 years). Interaction terms between age groups and other variables were included to uncover age-specific patterns. To mitigate multicollinearity arising from the interaction terms, we used penalized linear and logistic regression models using the least absolute shrinkage and selection operator (LASSO) penalty without survey weights, supplemented by bootstrapping to evaluate the stability. Using the selected variables, we then fitted survey-weighted generalized linear models to account for the complex design of the HINTS survey and more accurately represent the US population.

Results: Among 12,865 respondents, 36.1% (n=4638) were 65 years or older. Overall, 65.2% used patient portals and 38.4% used synchronous telehealth. Age alone was not a significant factor, but older adults exhibited distinct patterns. Older adults with higher care frequency were less likely to use patient portals (coefficient −0.07, 95% CI −0.10 to −0.04; P<.001). Those who completed the survey online in this age group were more likely to use patient portals (coefficient 0.32, 95% CI 0.18-0.47; P<.001) than to use a paper form. Older Black or African American patients were less likely to use patient portals (coefficient −0.30, 95% CI −0.45 to −0.16; P<.001). These age-specific patterns were identified only in patient portal use, not in synchronous telehealth use.

Conclusions: Telehealth usage among older adults differs in subtle ways from other age groups, and the factors driving adoption vary between patient portals and synchronous care. These findings underscore the need for age- and context-specific strategies to enhance the uptake of both types of services.

(JMIR Aging 2026;9:e83730) doi:10.2196/83730

KEYWORDS

telehealth; eHealth; mHealth; virtual healthcare; health services for the aged

Introduction

Telehealth has emerged as a cost-effective and integral component of modern health care, offering both patient portals and synchronous care (eg, real-time video or phone consultations) to improve access to care [1]. Synchronous telehealth facilitates live communication with health care providers without the need for long-distance travel, while patient portals streamline care delivery and reduce unnecessary office visits through various features such as asynchronous communications, access to medical records, medication refills, and appointment scheduling.

XSL•FO
RenderX

As the demand for remote health care continues to grow, telehealth has expanded beyond primary care into numerous specialties, with many providers now offering both in-person and virtual care options [2]. Despite its many advantages, telehealth has also introduced new challenges, particularly in exacerbating existing disparities in health care access [3]. For instance, individuals living in areas with limited internet infrastructure have struggled to benefit from telehealth services [4,5]. Moreover, the digital divide, driven by gaps in digital literacy, has hindered telehealth adoption, especially among older adults [5,6]. Additional barriers such as limited health literacy [7] and English proficiency [8] further complicate access and use.

Although research on telehealth is expanding, critical gaps remain. The barriers to telehealth use may be intricately linked to the modality itself, with patient portal and synchronous services presenting distinct challenges and opportunities. This suggests that tailored strategies may be necessary to enhance use across different modalities. Furthermore, patterns of telehealth use may vary significantly by age group, indicating the need for age-specific approaches to promote equitable access.

Using survey data collected in the postpandemic period, this study investigates the complex relationships between 2 types of telehealth modalities (ie, patient portals and synchronous telehealth) and various patient factors, including demographic factors, socioeconomic status, medical conditions, care behaviors, and technological proficiency. For the development of effective telehealth interventions, we compare the key drivers of use across modalities. Additionally, we explore how these factors differ among age groups, particularly among older adults, testing the hypothesis that telehealth usage patterns are uniquely shaped by age-related characteristics.

Methods

Data

In this study, we used the Health Information National Trends Survey (HINTS) Cycles 6 and 7, where the surveys were conducted in 2022 and 2024, respectively [9]. These surveys include several categories of questions, such as, “Looking for Health Information, Internet and Technology Use” (in Cycle 7; referred to as “Using the Internet to Find Information” in Cycle 6), “Your Health Care,” “Telehealth,” and “You and Your Household.” Assuming no period effects, as all the survey periods were conducted postpandemic, we harmonized the 2 datasets without accounting for the survey cycles. We used only overlapping variables with necessary name and type adjustments (see Table S1 in Multimedia Appendix 1) and combined the full sample weight and 50 replicate weights following HINTS guidelines [10]. Details on data preparation are available in the GitHub repository [11].

Outcomes

To identify the use of patient portals and synchronous telehealth, we used the following survey questions: “How many times did you access your online medical record or patient portal in the last 12 months?” (AccessOnlineRecord2 in Cycle 6 and AccessOnlineRecord3 in Cycle 7) for patient portal usage, and “In the past 12 months, did you receive care from a doctor or health professional using telehealth?” (ReceiveTelehealthCare) for synchronous telehealth usage. The patient portal question captures usage frequency, with 5 response options, including 0, 1 - 2, 3 - 5, 6 - 9, and 10 or more times. Patients who selected “I do not have an online medical record or patient portal” were classified as having no use (ie, 0 times). For simplicity in implementing parametric models, this outcome was treated as continuous in the model [12], with the 5 response categories coded from 0 to 4. On the other hand, the question about synchronous telehealth focuses on the method of experience, with possible answers being no, yes by video, yes by phone call, and yes by both. For analysis, all “yes” responses were combined and treated as a binary outcome. Additionally, we created a dichotomous variable “Neither” to indicate patients who had not used any form of telehealth, meaning neither patient portals nor synchronous telehealth. We excluded those who did not answer at least one of the telehealth questions (ie, AccessOnlineRecord2/3 or ReceiveTelehealthCare) to ensure consistent analytic samples across all outcomes.

Variables

Based on existing studies [5,6,13-15], variables included demographic and socioeconomic status information (ie, age, race, ethnicity, birth sex, total number of households, and residence), tech-savviness (ie, internet use and social media use), surveys, and care behaviors (ie, frequency of care services, ability to take good care of health). The 5 age categories in the data (ie, 18 - 34 years, 35 - 49 years, 50 - 64 years, 65 - 74 years, and ≥75 years) were combined into the following 3 groups, including young (18-49 years), middle (50-64 years), and old (≥65 years). Missing values for the independent variables were imputed using multivariate imputation by chained equations (MICE) [16] without survey weights. This imputation process included all independent variables and both dependent variables (patient portal and synchronous telehealth use) and produced converged imputed values after 5 iterations. Table S1 in Multimedia Appendix 1 presents the missing rates.

Models

To examine how the 3 outcomes (ie, patient portals, synchronous, and Neither) were associated with various variables across age groups, our models incorporated interaction terms between age groups and the other variables. The 3 models used the same analytic cohort and the same set of independent variables. For the patient portal and synchronous telehealth models, use of the other telehealth modality was also added to the set of independent variables (ie, synchronous telehealth uses in the patient portal model, and vice versa). To facilitate straightforward interpretation, ordinal independent variables, in addition to the patient portal frequency outcome, were treated as continuous in the models [12].

We used a two-step approach: (1) variable selection using unweighted penalized regression models and (2) statistical inference with survey-weighted generalized linear models. As the first step, we selected variables to prevent multicollinearity and overfitting from interaction terms. This screening process was conducted solely for model specification and did not involve

XSL•FO
RenderX

the use of survey weights. We built penalized linear or logistic regression models with the least absolute shrinkage and selection operator (LASSO) penalty for each outcome [17]. We applied LASSO linear regression for the patient portal outcome and LASSO logistic regression for the synchronous telehealth and Neither outcomes. The LASSO hyperparameter, penalty magnitude (λ), was determined using 5-fold cross-validation and selected based on the penalty value that produced the smallest average error across the 5 validation sets. To assess stability, we bootstrapped 500 samples, fitted penalized linear models with LASSO, and calculated the selection frequency, 95% CIs, and *P* values for each variable based on *t* tests, defined as the mean divided by the SD [18]. Variables with a *P* value of less than .05 were selected.

In the second stage, we fitted survey-weighted generalized linear models using the selected variables to account for the complex design of the HINTS survey and ensure accurate representation of the US population [19]. We used the full sample weight to produce population-level estimates and the replicate weights to calculate the standard error and variance of those estimates. For model evaluation, adjusted R^2 (for the patient portal model) and pseudo-R^2 (for the synchronous and Neither models) were used to assess goodness of fit [20,21]. In human subject research, these values between 0.10 and 0.30 are generally considered acceptable [21,22]. We reported coefficients for linear regression, odds ratio (OR) for logistic regression, and 95% CI and *P* value for both. For inference, a positive coefficient indicates being more likely to use patient portals, while an OR over 1 indicates being more likely to use synchronous telehealth in the synchronous model or never use any form of telehealth in the Neither model. The variable significance was determined when the variable's *P* value was smaller than the significance level (.05).

All analyses were conducted using R programming (version 4.4.3; R Core Team) with the following packages: *forestplot* [23], *glmnet* [17,24], *mice* [16], and *survey* [19].

Sensitivity Analyses

We conducted sensitivity analyses under 3 different scenarios to assess potential bias arising from data cleaning. First, to assess potential bias introduced by imputation, we created 5 different imputed datasets using the same imputation method (ie, MICE) with different random seeds. Models were then built on each imputed dataset, and significant variables were compared across them. Second, since survey form type may reflect technology familiarity, we excluded this variable from the entire analysis (from imputing missing values through constructing models) and then investigated changes in the identifications of significant factors among other technology familiarity variables (ie, social media interaction and internet use). Finally, since we treated ordinal patient portal use as continuous, we applied an additional modification to confirm its reasonableness, categorizing portal use into 2 groups (<3 and ≥3 times), treating it as binary, and rerunning the entire analysis for this outcome. This new outcome type required a different regression model than the original, making the results not directly comparable in size or magnitude. Instead, we focused on whether the same variables were significant and whether their effects pointed in the same direction (positive or negative).

Ethical Considerations

This study used publicly available, deidentified HINTS data to conduct a secondary analysis. The dataset contains no personal identifiers. Because the dataset is publicly available and deidentified, this research does not constitute human subjects research as defined by federal regulations (45 CFR 46.102) and is therefore exempt from institutional ethics review. Accordingly, institutional review board review at the authors' institution was not required.

Results

Data

Excluding missing responses to either patient portal or synchronous telehealth usage, we identified 12,865 patients in this study. They were comprised of 4714 (36.6%) young-age (18-49 years), 3513 (27.3%) middle-age (50-64 years), and 4638 (36.1%) old-age (≥65 years) patients. Non-Hispanic White patients accounted for more than half (7204/12,865, 56%), followed by Hispanic (2519/12,865, 19.6%), Non-Hispanic Black or African American (1979/12,865, 15.4%), and Non-Hispanic Asian (680/12,865, 5.3%). More female patients (7739/12,865, 60.2%) participated in this survey than male (5126/12,865, 39.8%). Additionally, 37.7% of patients (4854/12,865) resided in large metropolitan areas and 22.0% (2828/12,865) in large fringe metropolitan areas, while 13.4 % lived in nonmetropolitan areas (1730/12,865). Table 1 provides an overview of the unweighted distributions by telehealth modalities, based on the imputed dataset.

XSL•FO
RenderX

Table . Variables in this study and their unweighted distributions in total and by the outcomes (N=12,865).

Variables	Total, n (%)	Patient portals					Synchronous	
		0 times, n (%)	1-2 times, n (%)	3-5 times, n (%)	6-9 times, n (%)	≥10 times, n (%)	No use, n (%)	Use, n (%)
Telehealth use								
Demographic and socioeconomic status, n (%)	12,865 (100)	4482 (34.8)	2590 (20.1)	2605 (20.2)	1382 (10.7)	1806 (14.0)	7923 (61.6)	4942 (38.4)
Age (years), n (%)								
18-49	4714 (36.6)	1339 (28.4)	1168 (24.8)	1002 (21.3)	516 (10.9)	689 (14.6)	2721 (57.7)	1993 (42.3)
50-64	3513 (27.3)	1136 (32.3)	682 (19.4)	746 (21.2)	410 (11.7)	539 (15.3)	2138 (60.9)	1375 (39.1)
≥65	4638 (36.1)	2007 (43.3)	740 (16.0)	857 (18.5)	456 (9.8)	578 (12.5)	3064 (66.1)	1574 (33.9)
Sex, n (%)								
Female	7739 (60.2)	2433 (31.4)	1567 (20.2)	1604 (20.7)	908 (11.7)	1227 (15.9)	4524 (58.5)	3215 (41.5)
Male	5126 (39.8)	2049 (40.0)	1023 (20.0)	1001 (19.5)	474 (9.2)	579 (11.3)	3399 (66.3)	1727 (33.7)
Race and ethnicity, n (%)								
Non-Hispanic White	7204 (560)	2163 (30.0)	1405 (19.5)	1562 (21.7)	872 (12.1)	1202 (16.7)	4517 (62.7)	2687 (37.3)
Non-Hispanic Black or African American	1979 (15.4)	799 (40.4)	387 (19.6)	377 (19.1)	189 (9.6)	227 (11.5)	1216 (61.4)	763 (38.6)
Hispanic	2519 (19.6)	1151 (45.7)	531 (21.1)	409 (16.2)	197 (7.8)	231 (9.2)	1482 (58.8)	1037 (41.2)
Non-Hispanic Asian	680 (5.3)	197 (29.0)	167 (24.6)	165 (24.3)	70 (10.3)	81 (11.9)	426 (62.6)	254 (37.4)
Non-Hispanic other	483 (3.8)	172 (35.6)	100 (20.7)	92 (19.0)	54 (11.2)	65 (13.5)	282 (58.4)	201 (41.6)
Total household, n (%)								
1	3893 (30.3)	1718 (44.1)	651 (16.7)	686 (17.6)	354 (9.1)	484 (12.4)	2504 (64.3)	1389 (35.7)
>2	8972 (69.7)	2764 (30.8)	1939 (21.6)	1919 (21.4)	1028 (11.5)	1322 (14.7)	5419 (60.4)	3553 (39.6)
Education, n (%)								
College graduate or more	6168 (47.9)	1361 (22.1)	1367 (22.2)	1507 (24.4)	814 (13.2)	1119 (18.1)	3537 (57.3)	2631 (42.7)
Less than high school	797 (6.2)	545 (68.4)	120 (15.1)	72 (9.0)	34 (4.3)	26 (3.3)	548 (68.8)	249 (31.2)
High school graduate	2202 (17.1)	1199 (54.5)	367 (16.7)	299 (13.6)	165 (7.5)	172 (7.8)	1531 (69.5)	671 (30.5)
Some college	3698 (28.7)	1377 (37.2)	736 (19.9)	727 (19.7)	369 (10.0)	489 (13.2)	2307 (62.4)	1391 (37.6)
Neighborhood, n (%)								
Large metro	4854 (37.7)	1597 (32.9)	1009 (20.8)	991 (20.4)	535 (11.0)	722 (14.9)	2773 (57.1)	2081 (42.9)
Large fringe metro	2828 (22.0)	778 (27.5)	616 (21.8)	689 (24.4)	322 (11.4)	423 (15.0)	1682 (59.5)	1146 (40.5)

XSL•FO
RenderX

Variables	Total, n (%)	Patient portals					Synchronous	
		0 times, n (%)	1-2 times, n (%)	3-5 times, n (%)	6-9 times, n (%)	≥10 times, n (%)	No use, n (%)	Use, n (%)
Medium small metro	3453 (26.8)	1354 (39.2)	683 (19.8)	608 (17.6)	364 (10.5)	444 (12.9)	2252 (65.2)	1201 (34.8)
Non-metro	1730 (13.4)	753 (43.5)	282 (16.3)	317 (18.3)	161 (9.3)	217 (12.5)	1216 (70.3)	514 (29.7)
Medical condition								
Depression, n (%)								
No	9430 (73.3)	3492 (37.0)	2006 (21.3)	1909 (20.2)	924 (9.8)	1099 (11.7)	6351 (67.3)	3079 (32.7)
Yes	3435 (26.7)	990 (28.8)	584 (17.0)	696 (20.3)	458 (13.3)	707 (20.6)	1572 (45.8)	1863 (54.2)
Diabetes, n (%)								
No	10,066 (78.2)	3405 (33.8)	2148 (21.3)	2057 (20.4)	1070 (10.6)	1386 (13.8)	6362 (63.2)	3704 (36.8)
Yes	2799 (21.8)	1077 (38.5)	442 (15.8)	548 (19.6)	312 (11.1)	420 (15.0)	1561 (55.8)	1238 (44.2)
High blood pressure, n (%)								
No	7266 (56.5)	2388 (32.9)	1655 (22.8)	1484 (20.4)	764 (10.5)	975 (13.4)	4537 (62.4)	2729 (37.6)
Yes	5599 (43.5)	2094 (37.4)	935 (16.7)	1121 (20.0)	618 (11.0)	831 (14.8)	3386 (60.5)	2213 (39.5)
Heart condition, n (%)								
No	11,552 (89.8)	3946 (34.2)	2418 (20.9)	2365 (20.5)	1247 (10.8)	1576 (13.6)	7165 (62.0)	4387 (38.0)
Yes	1313 (10.2)	536 (40.8)	172 (13.1)	240 (18.3)	135 (10.3)	230 (17.5)	758 (57.7)	555 (42.3)
Lung disease, n (%)								
No	11,147 (86.6)	3942 (35.4)	2298 (20.6)	2270 (20.4)	1185 (10.6)	1452 (13.0)	7070 (63.4)	4077 (36.6)
Yes	1718 (13.4)	540 (31.4)	292 (17.0)	335 (19.5)	197 (11.5)	354 (20.6)	853 (49.7)	865 (50.3)
Tech savviness								
Social media interact, n (%)								
Never	9895 (76.9)	3861 (39.0)	1939 (19.6)	1915 (19.4)	966 (9.8)	1214 (12.3)	6508 (65.8)	3387 (34.2)
Less than once a month	1700 (13.2)	367 (21.6)	371 (21.8)	400 (23.5)	243 (14.3)	319 (18.8)	897 (52.8)	803 (47.2)
A few times a month	710 (5.5)	138 (19.4)	160 (22.5)	169 (23.8)	104 (14.6)	139 (19.6)	305 (43.0)	405 (57.0)
At least once a week or more	560 (4.4)	116 (20.7)	120 (21.4)	121 (21.6)	69 (12.3)	134 (23.9)	213 (38.0)	347 (62.0)
Use internet, n (%)								
No	1240 (9.6)	1014 (81.8)	99 (8.0)	71 (5.7)	27 (2.2)	29 (2.3)	883 (71.2)	357 (28.8)
Yes	11,625 (90.4)	3468 (29.8)	2491 (21.4)	2534 (21.8)	1355 (11.7)	1777 (15.3)	7040 (60.6)	4585 (39.4)
Survey								
Form type, n (%)								
Paper	4686 (36.4)	2528 (53.9)	706 (15.1)	731 (15.6)	343 (7.3)	378 (8.1)	3181 (67.9)	1505 (32.1)

XSL•FO
RenderX

Variables	Total, n (%)	Patient portals					Synchronous	
		0 times, n (%)	1-2 times, n (%)	3-5 times, n (%)	6-9 times, n (%)	≥10 times, n (%)	No use, n (%)	Use, n (%)
Web	8179 (63.6)	1954 (23.9)	1884 (23.0)	1874 (22.9)	1039 (12.7)	1428 (17.5)	4742 (58.0)	3437 (42.0)
Care behavior								
Self-care ability, n (%)								
Not confident at all or a little	749 (5.8)	334 (44.6)	129 (17.2)	111 (14.8)	75 (10.0)	100 (13.4)	411 (54.9)	338 (45.1)
Somewhat	3149 (24.5)	1191 (37.8)	614 (19.5)	601 (19.1)	319 (10.1)	424 (13.5)	1900 (60.3)	1249 (39.7)
Completely	3307 (25.7)	1126 (34.0)	709 (21.4)	643 (19.4)	365 (11.0)	464 (14.0)	2123 (64.2)	1184 (35.8)
Very	5660 (44.0)	1831 (32.3)	1138 (20.1)	1250 (22.1)	623 (11.0)	818 (14.5)	3489 (61.6)	2171 (38.4)
Care frequency, n (%)								
None	1459 (11.3)	1018 (69.8)	249 (17.1)	106 (7.3)	35 (2.4)	51 (3.5)	1219 (83.6)	240 (16.4)
1 time	1687 (13.1)	710 (42.1)	569 (33.7)	267 (15.8)	80 (4.7)	61 (3.6)	1221 (72.4)	466 (27.6)
2 times	2329 (18.1)	808 (34.7)	630 (27.1)	553 (23.7)	193 (8.3)	145 (6.2)	1544 (66.3)	785 (33.7)
3 times	2015 (15.7)	583 (28.9)	430 (21.3)	560 (27.8)	252 (12.5)	190 (9.4)	1184 (58.8)	831 (41.2)
4 times	1854 (14.4)	535 (28.9)	336 (18.1)	455 (24.5)	284 (15.3)	244 (13.2)	1078 (58.1)	776 (41.9)
5-9 times	2236 (17.4)	510 (22.8)	266 (11.9)	471 (21.1)	396 (17.7)	593 (26.5)	1145 (51.2)	1091 (48.8)
10 or more times	1285 (10.0)	318 (24.7)	110 (8.6)	193 (15.0)	142 (11.1)	522 (40.6)	532 (41.4)	753 (58.6)

Within this cohort, 8383 (65.2%) and 4942 (38.4%) used patient portals and synchronous telehealth at least once in the past 12 months, respectively. Patients who used patient portals used synchronous telehealth more, and vice versa. For example, of the 4942 patients who used synchronous telehealth, 3877 (78.5%) also used patient portals at least once. Moreover, the likelihood of using synchronous telehealth increased progressively with higher levels of patient portal use. Compared to those who never used patient portals, individuals who used it at least 10 times were significantly more likely to use synchronous telehealth, with an OR of 4.9 (See Table 2).

Table . Unweighted joint distributions of patient portal and synchronous telehealth use.

Variables	Total, n (%)	Patient portals				
		0 time	1-2 times	3-5 times	6-9 times	≥10 times
Telehealth use, n (%)	12865 (100)	4482 (34.8)	2590 (20.1)	605 (20.2)	1382 (10.7)	1806 (14.0)
Synchronous, n (%)						
No	7923 (61.6)	3417 (43.1)	1665 (21.0)	1453 (18.3)	674 (8.5)	714 (9.0)
Yes	4942 (38.4)	1065 (21.5)	925 (18.7)	1152 (23.3)	708 (14.3)	1092 (22.1)
Odds ratio (95% CI)	__[a]	reference	1.8 (1.6-2.0)	2.5 (2.3-2.8)	3.4 (3.0-3.8)	4.9 (4.4-5.5)

[a]Not applicable.

By age group, younger participants were more likely to use both patient portals and synchronous telehealth services (see age rows in Table 1). Among those aged 18 - 49 years (n=4714), 3375 (71.6%) participants used patient portals at least once, compared to 2377 (67.7%) in the 50 - 64 years age group (n=3513) and 2631 (56.7%) among those aged 65 years and older (n=4638). Similarly, synchronous telehealth was used by 1993 (42.3%) participants in the youngest group, 1375 (39.1%) in the middle-aged group, and 1574 (33.9%) in the oldest group. Furthermore, the mode of synchronous telehealth varied by age. Among those who used synchronous telehealth, use of phone calls only (without video) was significantly more common in the older group (672/1574, 42.7%) than in the younger group

XSL•FO
RenderX

(547/1993, 27.4%) and the middle-aged group (423/1375, 30.8%; P<.001; χ^2; see Table 3).

Table . Synchronous telehealth modes by age groups among those who used synchronous telehealth.

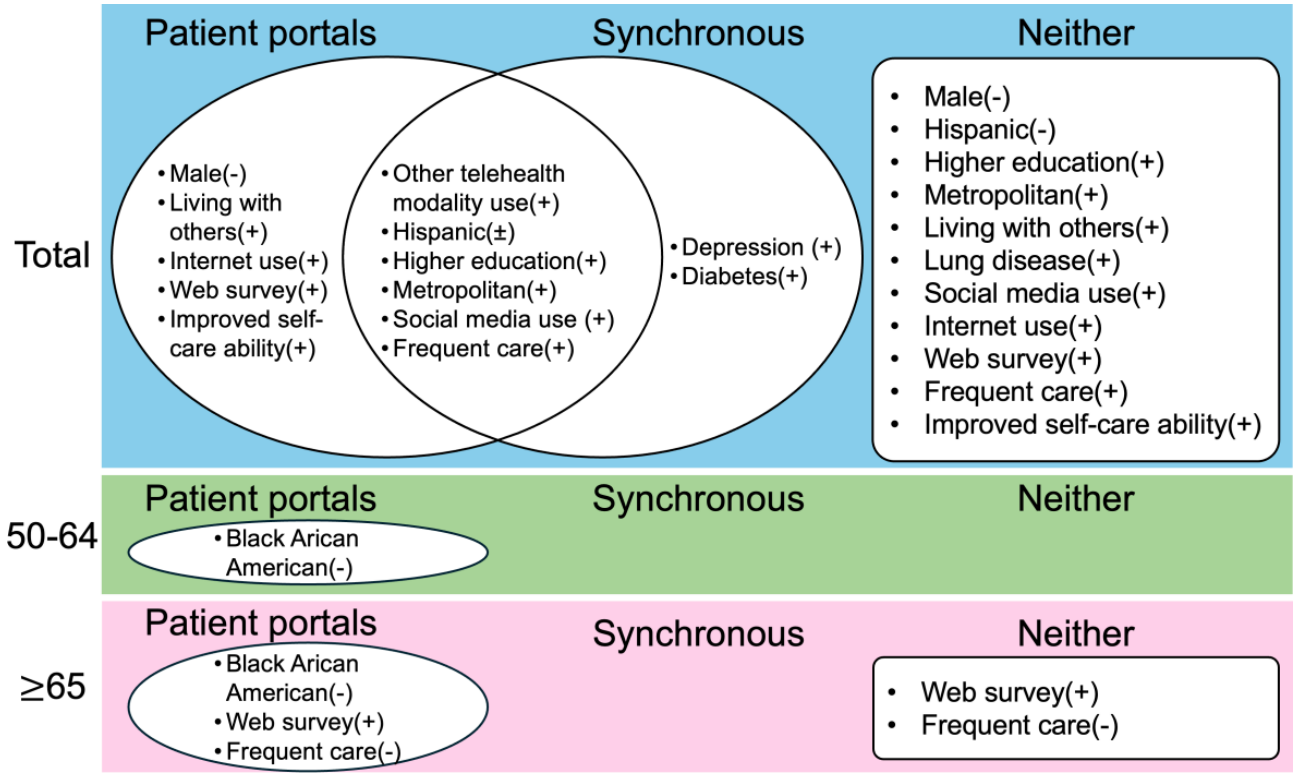
Variables	Total	Synchronous telehealth mode, n (%)		
		By video, n (%)	By phone call (voice only with no video), n (%)	Some by video and some by phone call, n (%)
Age (years), n (%)	4942 (100)	2248 (45.5)	1642 (33.2)	1052 (21.3)
18 - 49	1993 (42.3)	1000 (50.2)	547 (27.4)	446 (22.4)
50 - 64	1375 (39.1)	671 (48.8)	423 (30.8)	281 (20.4)
≥65	1574 (33.9)	577 (36.7)	672 (42.7)	325 (20.6)

Two-Step Approach for Variable Screening and Model Inference

Unweighted penalized regression models with bootstrapping were applied to select variables: 21 for the patient portal model, 13 for the synchronous model, and 20 for the Neither model, as shown in Tables S2-S4 in Multimedia Appendix 1. Many interaction terms were not selected because the age-specific subgroups had smaller sample sizes. Even though variables within the subgroups were consistently selected more than 480 out of 500 times, their P value was sometimes larger than .05 due to the sample sizes (eg, survey type for the middle-aged group in the patient portal model). Estimates, confidence intervals, P values, and selection frequencies for all variables are reported in Table S2-S4 in Multimedia Appendix 1.

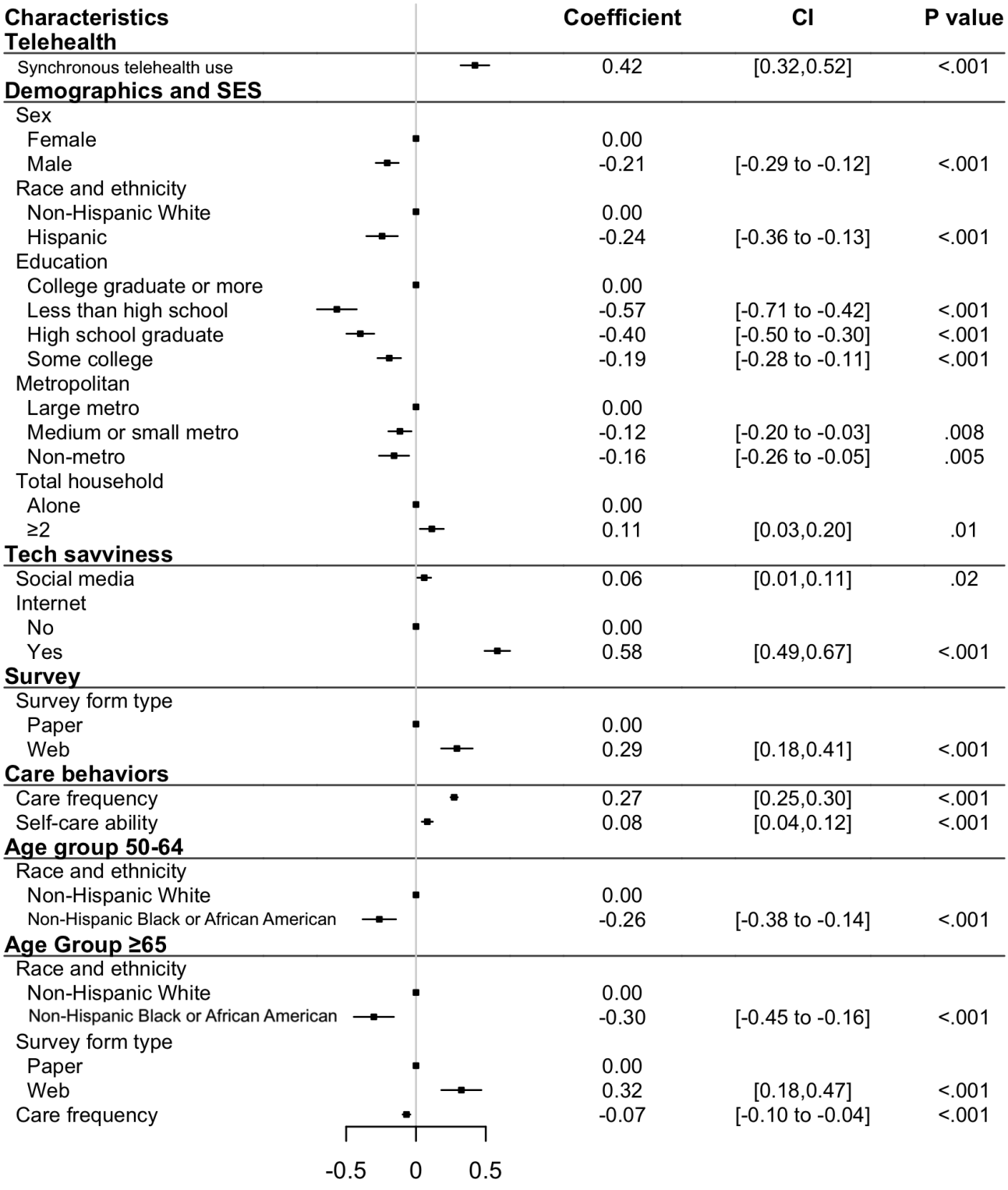
Using the selected variables, we constructed survey-weighted generalized linear models for the 3 outcomes. The adjusted R^2 of the patient portal model was 0.31, while the pseudo-R^2 values of the synchronous telehealth model and the Neither model were 0.15 and 0.22, respectively. Based on the acceptable range of 0.10-0.30 [21,22], all models exhibited a satisfactory goodness of fit. Figure 1 presents the unique and shared significant variables identified in the models, and Figures 2-4 detail the direction and strength of associations for each model.

Figure 1. Shared and unique significant variables associated with patient portal and synchronous telehealth use, both overall and across different age groups. The figure also highlights significant variables identified in the Neither model. A plus (+) symbol indicates greater telehealth use, a minus (−) indicates less use, and a plus or minus (±) symbol denotes mixed directions of association.



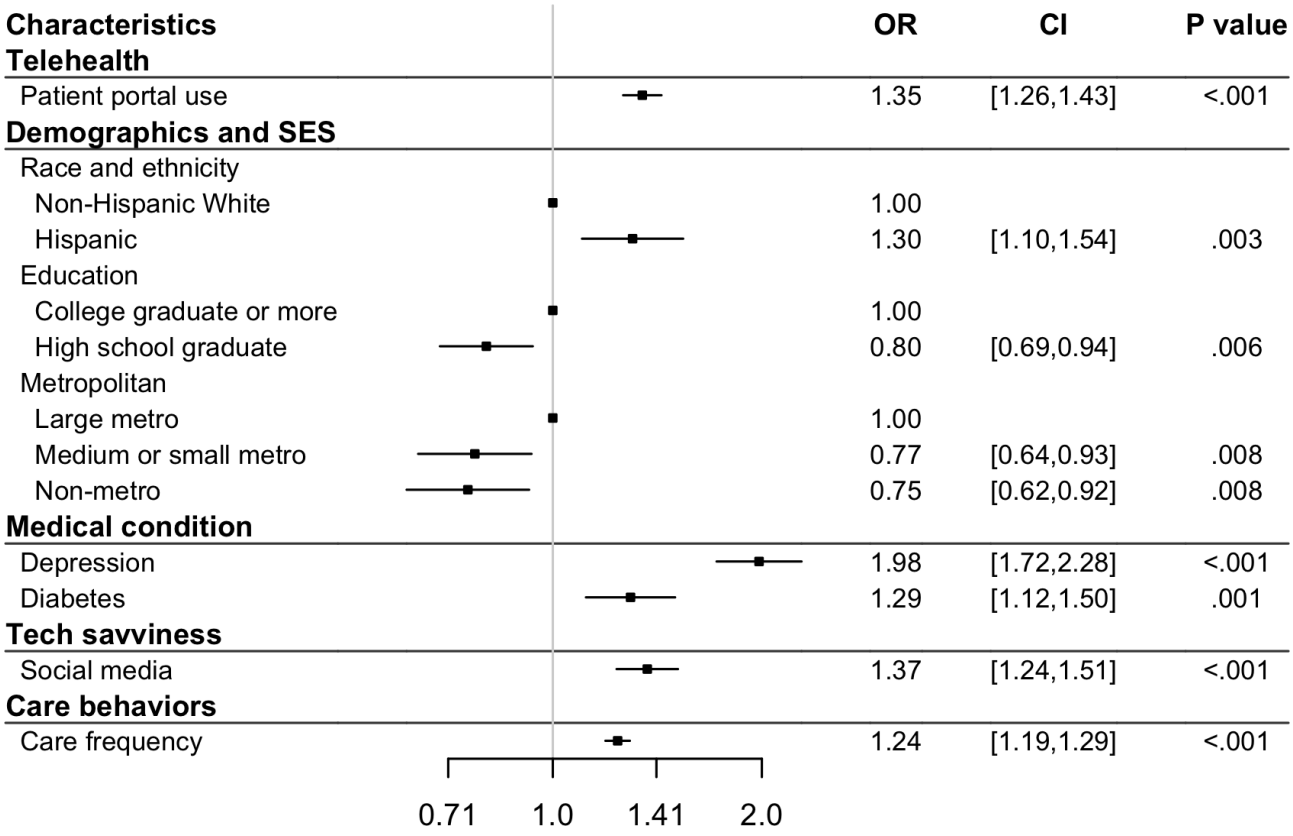
XSL•FO
RenderX

Figure 2. Significant variables identified from the patient portal model. OR: odds ratio; SES: socioeconomic status.



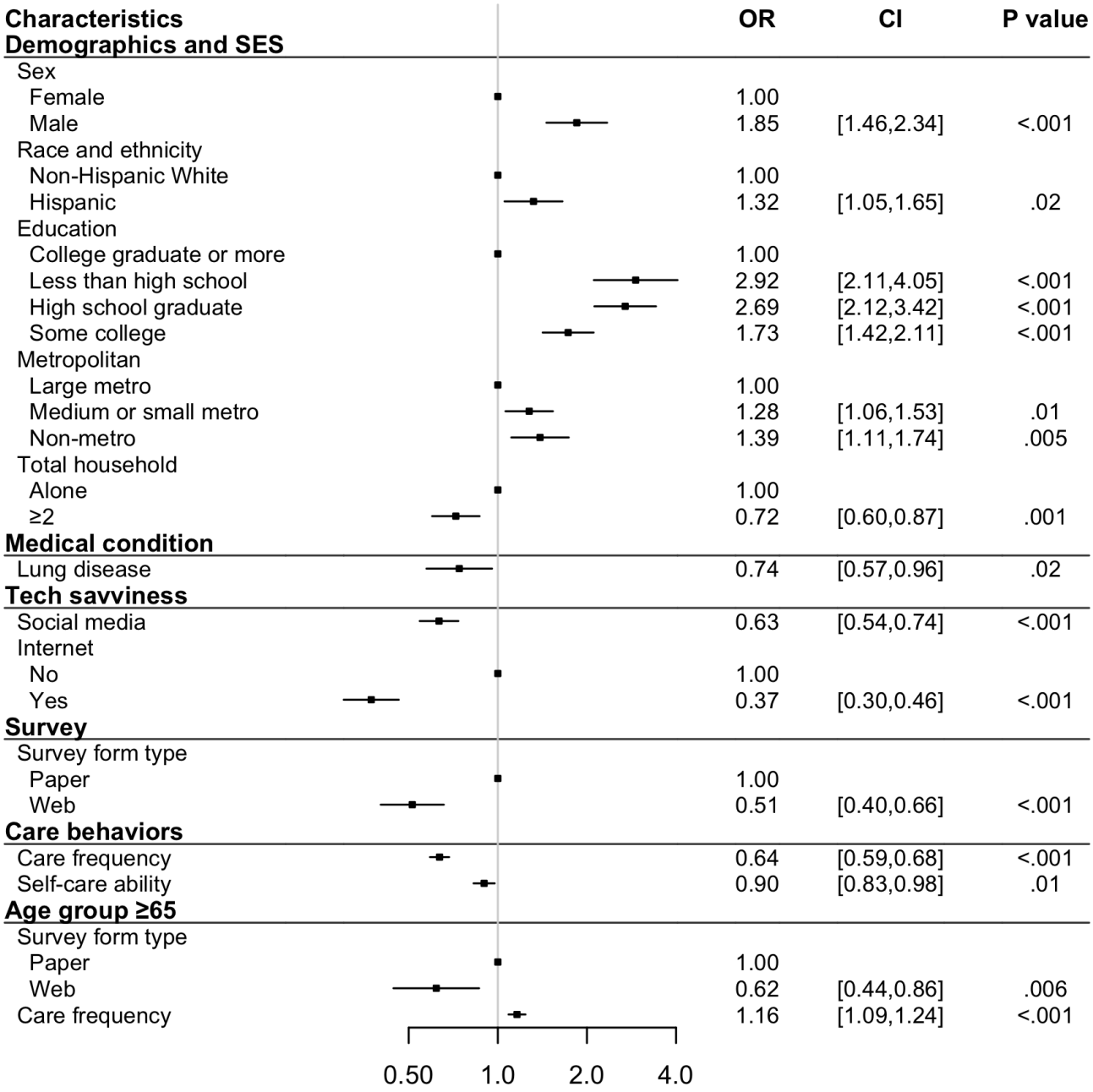
XSL•FO
RenderX

Figure 3. Significant variables identified from the synchronous telehealth model. OR: odds ratio; SES: socioeconomic status.



XSL•FO
RenderX

Figure 4. Significant variables identified from the neither model. OR: odds ratio; SES: socioeconomic status.



Use of one type of telehealth service was positively associated with use of the other modality (coefficient 0.42, 95% CI 0.32-0.53; *P*<.001 for patient portals; OR 1.35, 95% CI 1.26-1.43; *P*<.001 for synchronous). In terms of demographics and socioeconomic status, male patients were less likely to use patient portals than female (coefficient −0.21, 95% CI −0.29 to −0.12; *P*<.001), there was no significant difference in synchronous use between male and female patients (*P*=.10; see Table S6 in Multimedia Appendix 1). Additionally, Hispanic patients showed a contrasting pattern: they were more likely to use more synchronous telehealth (OR 1.30, 95% CI 1.10-1.54; *P*=.003) but less likely to use patient portals (coefficient −0.24, 95% CI −0.36 to −0.13; *P*<.001). Furthermore, middle-aged and older non-Hispanic Black or African American patients (aged 50 years and older) were less likely to use patient portals (coefficient −0.26, 95% CI −0.38 to −0.14; *P*<.001 for 50 - 64 years; coefficient −0.30, 95% CI −0.45 to −0.16; *P*<.001 for ≥65 years).

Compared to patients in large metropolitan areas, those in medium or small metropolitan and nonmetropolitan areas were less likely to use either telehealth modality (patient portals: coefficient −0.12, 95% CI −0.20 to −0.03; *P*=.008 for medium or small metropolitan, coefficient −0.16, 95% CI −0.26 to −0.05; *P*=.005 for nonmetropolitan; synchronous: OR 0.77, 95% CI 0.64-0.93; *P*=.008 for medium or small metropolitan, OR 0.75, 95% CI 0.62-0.92; *P*=.008 for nonmetropolitan). Finally, among the demographic and socioeconomic variables, individuals living with other household members were more likely to use only patient portals (coefficient 0.11, 95% CI 0.03-0.20; *P*=.01). Those with higher educational attainment were more likely to use both services (patient portals: coefficient range [−0.57 to −0.19], 95% CI −0.71 to −0.11; *P*<.001 for all the other categories [less than high school, high school graduate, some

XSL•FO
RenderX

college]; synchronous: OR 0.80, 95% CI 0.69-0.94; *P*=.006 for high school graduate; see Figures 2 and 3).

Medical conditions played a role mainly for synchronous telehealth. Patients with depression were more likely to use synchronous telehealth services (OR 1.98, 95% CI 1.72-2.28; *P*<.001). Diabetes was associated with increased use of synchronous telehealth (OR 1.29, 95% CI 1.12-1.50; *P*=.001), and patients with lung disease were less likely to avoid telehealth altogether (OR 0.74, 95% CI 0.57-0.96; *P*=.02).

In the tech-savviness category, social media use was positively associated with telehealth use across both modalities (patient portals: coefficient 0.06, 95% CI 0.01-0.11; *P*=.02; synchronous: OR 1.37, 95% CI 1.24-1.51; *P*<.001). Moreover, those who responded using the internet were more likely to use patient portals (coefficient 0.58, 95% CI 0.49-0.67; *P*<.001). For the survey type, participants who completed web-based surveys were more likely to use patient portals than those who completed paper-based surveys (coefficient 0.29, 95% CI 0.18-0.41; *P*<.001), with this trend being even stronger among older adults (coefficient 0.32, 95% CI 0.18-0.47; *P*<.001).

Regarding care behaviors, patients who saw health care professionals more frequently were more likely to use telehealth (coefficient 0.27, 95% CI 0.25-0.30; *P*<.001 for patient portals; OR 1.24, 95% CI 1.19-1.29; *P*<.001 for synchronous). However, among older patients, this association was less pronounced for patient portals (coefficient −0.07, 95% CI −0.10 to −004; *P*<.001). Additionally, patients who reported being confident in managing their own health were more likely to use patient portals (coefficient 0.08, 95% CI 0.04-0.12; *P*<.001).

Sensitivity Analyses

Table S5-S7 in Multimedia Appendix 1 exhibit the comparisons of the significant variables across models using the imputed dataset in the main text and 5 different imputations. Most significant variables identified in the main text models remained significant. However, due to the small number of samples in the age subgroups, the interaction terms sometimes became insignificant, non-Hispanic Black or African American in the patient portal models was identified as significant 3 out of 6 instances for both the 50 - 64 years and ≥65 years age groups (see Table S6 in Multimedia Appendix 1).

As illustrated in Table S8-S10 in Multimedia Appendix 1, after excluding the form type variable, the effect of internet use became more pronounced in both the patient portal telehealth model (see Table S8 in Multimedia Appendix 1) and the Neither model (see Table S10 in Multimedia Appendix 1), where an OR closer to “0” indicates a much higher likelihood of using one or both telehealth options. For synchronous telehealth, the form type variable was not identified as significant in the original model, resulting in minimal change (see Table S9 in Multimedia Appendix 1).

Finally, the model treating patient portal use as a continuous variable largely shared significant predictors with the binary model, as shown in Table S11 in Multimedia Appendix 1. Each model, however, had a few unique significant variables: for the continuous model, nonmetro, social media use, and non-Hispanic Black or African American individuals aged 50 - 64 years; for the binary model, high blood pressure and nonmetro status among those aged 65 years or older.

Discussion

Principal Findings

Using the harmonized HINTS Cycles 6 and 7 data, including national representative patient populations, this study offers a unique examination of both patient portal and synchronous telehealth use. While both services are delivered remotely and therefore share foundational similarities, their distinct functionalities lead to divergent patterns of use. By analyzing a range of factors, including demographics, socioeconomic status, and technological proficiency, we identified several common predictors associated with the use of both modalities, as well as unique factors specific to each. Furthermore, we discovered the variations across age groups, particularly among those aged 50 - 64 years and those 65 years and older. These findings contribute to a deeper understanding of age-related differences in telehealth use and could offer insights for enhancing patient experience and promoting broader telehealth adoption.

Methodologically, our 2-step approach effectively addresses potential multicollinearity during variable selection and ensures nationally representative estimates. The bootstrapped variable selection process helps build a parsimonious set of variables while accommodating multiple interaction terms. With this reduced set of variables, the survey-weighted generalized linear models enable robust population-level statistical inference, providing both nationally representative estimates and measures of their variability.

Unique Patterns of Telehealth Use Among Older Adults

While previous research has suggested that older adults are generally less likely to use telehealth [5,6], our models did not find age itself to be a significant predictor. Instead, our findings reveal that older adults exhibit unique usage patterns that differ from the general population in several ways; some are entirely new (eg, race-related differences), some are more pronounced (eg, survey form type), and others are weaker (eg, care frequency). These nuances underscore the importance of targeted investigations into telehealth use among older adults.

As an example of an emerging pattern, our data show non-Hispanic Black or African American patients overall used patient portals at rates similar to non-Hispanic White, which contradicts findings from prior research [25,26]. However, within this race group, individuals aged 50 years and older were less likely to use patient portals in some models trained on different imputed datasets. On average, responses shifted 0.26 points toward the “no-use” end of the scale for those aged 50 - 64 years and 0.30 points for those aged 65 years and older. This suggests that the digital divide in patient portal services may not be solely age-based but could intersect with race, underscoring the possible need for multidimensional strategies to address disparities in the use.

XSL•FO
RenderX

In terms of stronger patterns in the old patient group (≥65 years), old patients who completed the web-based survey were more likely to use patient portals than the general population, with their reported usage increasing by an average of 0.32 points toward the "10 or more than 10 times" end of the scale. In the sensitivity analysis, completing a web-based survey likely reflects greater familiarity with technology. Taken together, these findings suggest that comfort with technology can significantly improve engagement with patient portal services. As such, targeted technology education initiatives for older adults could play a vital role in promoting patient portal adoption and improving health care accessibility for this population, as suggested in prior research [27].

Regarding weaker patterns observed in the old population (≥65 years), patients who frequently visited health care providers were more likely to use patient portals in the general population, but this association was weaker among older adults. For each one-point increase in visit frequency toward the "10 or more times" end of the scale, responses for patient portal use were, on average, 0.07 points lower. Patients who receive care more often are generally exposed to more educational opportunities about patient portals and have greater chances to use patient portals. However, older adults may prefer addressing all concerns during in-person visits rather than through patient portals.

Key Differences in Drivers of Patient Portals and Synchronous Telehealth

The use of patient portals and synchronous telehealth shares some common drivers, such as residing in metropolitan areas and having higher levels of education. Particularly, patients who used one type of telehealth were more likely to use the other, consistent with previous studies suggesting that telehealth experience can enhance patient satisfaction and encourage further use [28,29]. However, important differences exist between the two. Patient portal use is largely shaped by patient characteristics, whereas synchronous telehealth depends primarily on the availability of real-time care options. Patients with diabetes and those with depression were more likely to use synchronous telehealth, with odds that were 98% higher and 29% higher, respectively, compared to those without these conditions. Additionally, those with lung disease were more likely to use either type or both, as their odds of not using any form of telehealth decreased by 26%. This finding likely reflects the nature of care required. For example, psychiatric care and diabetes care can involve synchronous consultations and telemonitoring [30,31]. Chronic obstructive pulmonary disease patients may involve a wide range of services, such as telerehabilitation, health education, telemonitoring, early detection of exacerbations, psychosocial support, and smoking cessation, delivered through both modalities [32,33]. These findings highlight the potential for synchronous telehealth patients who require long-term care and frequent visits.

On the other hand, the patient portal model included more variables related to tech-savviness (ie, internet use and completion of web-based surveys) in addition to social media use, which was commonly identified as significant for both modalities. Patients who used the internet scored, on average, 0.58 points higher toward the "10 or more than 10 times" end of the scale in patient portal usage, and those who completed the HINTS online scored 0.29 points higher. This may be due to the multiple functions available in patient portals, such as sending images and secure messages, which require a certain level of technological familiarity. In contrast, synchronous telehealth can be conducted not only via video but also through phone calls, and voice-only visits demand far fewer technical skills. Additionally, a one-unit increase in patients' self-care ability toward the "very confident" end of the scale was associated with greater patient portal use, corresponding to an average increase of 0.08 points, but showed no association with synchronous telehealth use. This may be because patient portal features, such as viewing lab results or refilling prescriptions, are more directly aligned with self-management behaviors [34,35].

In terms of demographic and socioeconomic factors, prior research shows that male patients are generally less likely to use telehealth [36,37]. However, this study confirms that this pattern applies only to patient portal use, with an average decrease of 0.21 points toward the "no use" end of the scale, but not to synchronous telehealth. Additionally, patients living with household members were also more likely to use patient portals only. This pattern may be linked to the shared access features of patient portals [38]. Family members or care partners can log into the patient's account to help manage their care, which may encourage patients to engage with the portal and participate more actively in their health care. Finally, age-specific effects were identified only for patient portal use. These findings suggest that promoting patient portal adoption should be particularly tailored to individual patient characteristics.

Interestingly, Hispanic patients showed contrasting usage patterns; they were more likely to use synchronous telehealth but less likely to use patient portals. Their odds of using synchronous telehealth increased by 30%, while their responses for patient portals were, on average, 0.24 points lower toward the "no use" end of the scale. This aligns with one side of the existing literature, which remains inconclusive [39,40]. One possible explanation is that while health care providers may offer services in Spanish, patient portal platforms often lack adequate translation features, creating barriers for patients with limited English proficiency [41]. To address this, multilingual support and culturally informed design are essential for effective telehealth delivery.

Limitations

This study is based on observational, retrospective data from surveys not specifically designed for this research. As such, we were unable to explore in-depth questions, such as patient satisfaction with telehealth across both modalities and the underlying reasons for choosing patient portals and synchronous care. Furthermore, we applied survey weights to the final inference models to improve sample representativeness, but they could also be used for imputing missing values [42]. Additionally, despite the complex survey design approach, the generalizability of our findings remains limited. Survey weights adjust nonresponse based on observed characteristics, but

XSL•FO
RenderX

unobserved factors may still influence participation. Residual confounding, such as broadband connectivity and state-level telehealth policies, both known to affect remote care use, was not directly measured. For example, counties with greater broadband availability have been shown to exhibit significantly higher telehealth use [43,44], and policy variations across states continue to impact access in rural areas [45]. Therefore, future research using multilevel designs or incorporating additional factors (eg, infrastructure and policy indicators) is warranted.

Another potential limitation is endogeneity between patient portal use and health care use frequency. Individuals who have more in-person health care service encounters may receive more provider recommendations to use patient portals, leading to higher portal adoption and engagement. Such unobserved factors may simultaneously drive greater portal use [46,47]. Future analyses could apply lagged models to test temporal ordering (eg, whether prior care use predicts later portal engagement) or use instrumental variable approaches, such as broadband access or provider-level promotion, to better account for endogeneity and unobserved confounding.

While harmonizing Cycles 6 and 7, we assumed no survey period effects. However, as shown in Table S12 in Multimedia Appendix 1, the use of patient portals and synchronous telehealth varied across cycles. In Cycle 7, participants reported more frequent use of patient portals and less frequent use of synchronous telehealth compared to Cycle 6 (all P values <.001 and chi-square tests). Future research should account for these period differences.

As shown in Table 2 and suggested in existing research [5,48], we recognize that disaggregating synchronous telehealth into video and phone visits could provide valuable insights. However, splitting synchronous encounters in our dataset resulted in small cell sizes within key age strata, leading to unstable estimates and reduced statistical power. Therefore, we retained a combined synchronous category in the primary analyses while noting potential modality differences as an important direction for future research with larger samples.

Finally, in the sensitivity analysis assessing potential bias from missing data imputation, race was sometimes identified as significant and at other times not. A possible explanation is the relatively higher missing rate for races (8.6%). Obtaining more complete observations would strengthen the robustness of our findings.

Conclusions

This study identifies both shared and distinct factors significantly associated with patient portal and synchronous telehealth use, as well as age-specific use patterns. These insights contribute to a deeper understanding of how to effectively promote both modalities, particularly among aging populations. Notably, the increased use of patient portals among older adults who are more technologically adept suggests that ongoing digital education initiatives may enhance telehealth adoption in this demographic. Furthermore, targeted outreach to specific demographic groups among old patients, such as non-Hispanic Black or African American, may be necessary to address disparities in telehealth use. Collectively, these findings offer valuable guidance for optimizing telehealth implementation strategies and advancing equitable telehealth delivery.

Funding

There is no funding support for this research.

Data Availability

The dataset is available on the HINTS website [9].

Authors' Contributions

Conceptualization: JP, SH
Data curation: JP
Formal analysis: JP
Investigation: JP
Methodology: JP, SH
Visualization: JP
Writing – original draft: JP, SH
Writing – review & editing: JP, SH

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental tables (Table S1 through S12) providing detailed results and additional analyses.
[DOCX File, 98 KB - aging_v9i1e83730_app1.docx]

References

XSL•FO
RenderX

1. Hilty DM, Torous J, Parish MB, et al. A literature review comparing clinicians' approaches and skills to in-person, synchronous, and asynchronous care: moving toward competencies to ensure quality care. Telemed J E Health 2021 Apr;27(4):356-373. [doi: 10.1089/tmj.2020.0054] [Medline: 32412882]
2. Connolly SL, Miller CJ, Gifford AL, Charness ME. Perceptions and use of telehealth among mental health, primary, and specialty care clinicians during the COVID-19 pandemic. JAMA Netw Open 2022 Jun 1;5(6):e2216401. [doi: 10.1001/jamanetworkopen.2022.16401] [Medline: 35671053]
3. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. JMIR Med Inform 2020 Aug 12;8(8):e20359. [doi: 10.2196/20359] [Medline: 32784177]
4. Kobeissi MM, Hickey JV. An infrastructure to provide safer, higher-quality, and more equitable telehealth. Jt Comm J Qual Patient Saf 2023 Apr;49(4):213-222. [doi: 10.1016/j.jcjq.2023.01.006] [Medline: 36775714]
5. Kim J, Cai ZR, Chen ML, Onyeka S, Ko JM, Linos E. Telehealth utilization and associations in the United States during the third year of the COVID-19 pandemic: population-based survey study in 2022. JMIR Public Health Surveill 2024 Apr 26;10(1):e51279. [doi: 10.2196/51279] [Medline: 38669075]
6. Mojtahedi Z, Sun I, Shen JJ. Telehealth utilization and good care among informal caregivers: health information national trends survey, 2022. Healthcare (Basel) 2023 Dec 18;11(24):3193. [doi: 10.3390/healthcare11243193] [Medline: 38132083]
7. Brown SH, Griffith ML, Kripalani S, Horst SN. Association of health literacy and area deprivation with initiation and completion of telehealth visits in adult medicine clinics across a large health care system. JAMA Netw Open 2022 Jul 1;5(7):e2223571. [doi: 10.1001/jamanetworkopen.2022.23571] [Medline: 35877125]
8. Rodriguez JA, Saadi A, Schwamm LH, Bates DW, Samal L. Disparities in telehealth use among California patients with limited English proficiency. Health Aff (Millwood) 2021 Mar;40(3):487-495. [doi: 10.1377/hlthaff.2020.00823] [Medline: 33646862]
9. Health information national trends survey. HINTS. URL: https://hints.cancer.gov/ [accessed 2023-10-24]
10. How to merge HINTS data part 3 - creating a merged dataset and verifying your data. Youtube. 2018. URL: https://www.youtube.com/watch?v=a5VuMoJHZh0 [accessed 2025-12-08]
11. Park J. Jaeyoung-jae-park/ptportal_syncth_hints67. GitHub. URL: https://github.com/jaeyoung-jae-park/PtPortal_SyncTH_HINTS67 [accessed 2026-01-10]
12. Huh I, Gim J. Exploration of Likert scale in terms of continuous variable with parametric statistical methods. BMC Med Res Methodol 2025 Sep 29;25(1):218. [doi: 10.1186/s12874-025-02668-1] [Medline: 41023822]
13. Spaulding EM, Fang M, Commodore-Mensah Y, Himmelfarb CR, Martin SS, Coresh J. Prevalence and disparities in telehealth use among US adults following the COVID-19 pandemic: national cross-sectional survey. J Med Internet Res 2024 May 10;26(1):e52124. [doi: 10.2196/52124] [Medline: 38728070]
14. Bhuyan SS, Kalra S, Mahmood A, et al. Motivation and use of telehealth among people with depression in the United States. J Prim Care Community Health 2024;15:21501319241266515. [doi: 10.1177/21501319241266515] [Medline: 39279333]
15. Raj M, Iott B. Characterizing telehealth use in the US: analysis of the 2022 Health Information National Trends Survey. Am J Manag Care 2024 Jan;30(1):50-56. [doi: 10.37765/ajmc.2024.89486] [Medline: 38271582]
16. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw 2011;45(3):1-67. [doi: 10.18637/jss.v045.i03]
17. Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc Series B Stat Methodol 1996 Jan 1;58(1):267-288. [doi: 10.1111/j.2517-6161.1996.tb02080.x]
18. DiCiccio TJ, Efron B. Bootstrap confidence intervals. Statist Sci 1996;11(3):189-228. [doi: 10.1214/ss/1032280214]
19. Lumley T. Analysis of complex survey samples. J Stat Soft 2004;9(8):1-19. [doi: 10.18637/jss.v009.i08]
20. Mbachu HI, Nduka EC, Nja ME. Designing a pseudo R-squared goodness-of-fit measure in generalized linear models. JMR 2012;4(2):148. [doi: 10.5539/jmr.v4n2p148]
21. Ozili PK. The acceptable r-square in empirical modelling for social science research. In: Social Research Methodology and Publishing Results: A Guide to Non-Native English Speakers 2023:134-143. [doi: 10.4018/978-1-6684-6859-3.ch009]
22. Gupta A, Stead TS, Ganti L. Determining a meaningful R-squared value in clinical medicine. Acad Med Surg 2024;27. [doi: 10.62186/001c.125154]
23. Gordon M, Lumley T. Forestplot: advanced forest plot using "grid. CRAN (Comprehensive R Archive Network). 2024. URL: https://CRAN.R-project.org/package=forestplot [accessed 2026-02-16]
24. Tay JK, Narasimhan B, Hastie T. Elastic net regularization paths for all generalized linear models. J Stat Softw 2023;106:1-31. [doi: 10.18637/jss.v106.i01] [Medline: 37138589]
25. Iasiello JA, Rajan A, Zervos E, Parikh AA, Snyder RA. Racial differences in patient-reported access to telehealth: an important and unmeasured social determinant of health. JCO Oncol Pract 2023 Dec;19(12):1215-1223. [doi: 10.1200/OP.23.00006] [Medline: 37844269]
26. Marcondes FO, Normand SLT, Le Cook B, et al. Racial and ethnic differences in telemedicine use. JAMA Health Forum 2024 Mar 1;5(3):e240131. [doi: 10.1001/jamahealthforum.2024.0131] [Medline: 38517424]

XSL•FO
RenderX

27. Mace RA, Mattos MK, Vranceanu AM. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022 Dec 30;12(12):1102-1105. [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
28. Sorondo B, Allen A, Fathima S, Bayleran J, Sabbagh I. Patient portal as a tool for enhancing patient experience and improving quality of care in primary care practices. EGEMS (Wash DC) 2016;4(1):1262. [doi: 10.13063/2327-9214.1262] [Medline: 28203611]
29. Ezeamii VC, Okobi OE, Wambai-Sani H, et al. Revolutionizing healthcare: how telemedicine is improving patient outcomes and expanding access to care. Cureus 2024 Jul;16(7):e63881. [doi: 10.7759/cureus.63881] [Medline: 39099901]
30. Scott AM, Clark J, Greenwood H, et al. Telehealth v. face-to-face provision of care to patients with depression: a systematic review and meta-analysis. Psychol Med 2022 Oct;52(14):2852-2860. [doi: 10.1017/S0033291722002331] [Medline: 35959559]
31. Dhediya R, Chadha M, Bhattacharya AD, Godbole S, Godbole S. Role of telemedicine in diabetes management. J Diabetes Sci Technol 2023 May;17(3):775-781. [doi: 10.1177/19322968221081133] [Medline: 35227105]
32. Barbosa MT, Sousa CS, Morais-Almeida M. Telemedicine in the management of chronic obstructive respiratory diseases: an overview. In: Linwood SL, editor. Digital Health Exon Publications 2022. URL: http://www.ncbi.nlm.nih.gov/books/NBK580636 [accessed 2025-08-29] [doi: 10.36255/exon-publications-digital-health-respiratory-diseases]
33. Jezewski E, Miller A, Eusebio M, Potter J. Targeted telehealth education increases interest in using telehealth among a diverse group of low-income older adults. Int J Environ Res Public Health 2022 Oct 16;19(20):13349. [doi: 10.3390/ijerph192013349] [Medline: 36293929]
34. Park J, Guo S, Liang M, Zhong X. Investigation of the causal relationship between patient portal utilization and patient's self-care self-efficacy and satisfaction in care among patients with cancer. BMC Med Inform Decis Mak 2025 Jan 8;25(1):12. [doi: 10.1186/s12911-024-02837-0] [Medline: 39780146]
35. Kujala S, Hörhammer I, Kaipio J, Heponiemi T. Health professionals' expectations of a national patient portal for self-management. Int J Med Inform 2018 Sep;117:82-87. [doi: 10.1016/j.ijmedinf.2018.06.005]
36. Fischer SH, Ray KN, Mehrotra A, Bloom EL, Uscher-Pines L. Prevalence and characteristics of telehealth utilization in the United States. JAMA Netw Open 2020 Oct 1;3(10):e2022302. [doi: 10.1001/jamanetworkopen.2020.22302] [Medline: 33104208]
37. Williams C, Shang D. Telehealth usage among low-income racial and ethnic minority populations during the COVID-19 pandemic: retrospective observational study. J Med Internet Res 2023 May 12;25:e43604. [doi: 10.2196/43604] [Medline: 37171848]
38. Gleason KT, Peereboom D, Wec A, Wolff JL. Patient portals to support care partner engagement in adolescent and adult populations: a scoping review. JAMA Netw Open 2022 Dec 1;5(12):e2248696. [doi: 10.1001/jamanetworkopen.2022.48696] [Medline: 36576738]
39. Cockrell H, Wayne D, Wandell G, et al. Understanding Hispanic patient satisfaction with telehealth during COVID-19. J Pediatr Surg 2023 Sep;58(9):1783-1788. [doi: 10.1016/j.jpedsurg.2022.12.006] [Medline: 36635160]
40. White-Williams C, Liu X, Shang D, Santiago J. Use of telehealth among racial and ethnic minority groups in the United States before and during the COVID-19 pandemic. Public Health Rep 2023;138(1):149-156. [doi: 10.1177/00333549221123575] [Medline: 36113138]
41. Solomou T, Mappouras S, Kyriacou E, et al. Bridging language barriers in healthcare: a patient-centric mobile app for multilingual health record access and sharing. Front Digit Health 2025;7:1542485. [doi: 10.3389/fdgth.2025.1542485] [Medline: 40041125]
42. Kalpourtzi N, Carpenter JR, Touloumi G. Handling missing values in surveys with complex study design: a simulation study. J Surv Stat Methodol 2024 Jan 25;12(1):105-129. [doi: 10.1093/jssam/smac039]
43. Quinton JK, Ong MK, Vangala S, et al. The association of broadband internet access and telemedicine utilization in rural Western Tennessee: an observational study. BMC Health Serv Res 2021 Aug 3;21(1):765. [doi: 10.1186/s12913-021-06746-0] [Medline: 34344377]
44. Wilcock AD, Rose S, Busch AB, et al. Association between broadband internet availability and telemedicine use. JAMA Intern Med 2019 Nov 1;179(11):1580-1582. [doi: 10.1001/jamainternmed.2019.2234] [Medline: 31355849]
45. Examining rural telehealth during the public health emergency. : Centers for Medicare & Medicaid Services URL: https://www.cms.gov/files/document/examining-rural-telehealth-jan-2023.pdf [accessed 2026-02-23]
46. Hong YR, Xie Z, Nguyen OT, Turner K, Walker AF. Telehealth service use and quality of care among US adults with diabetes: a cross-sectional study of the 2022 health information national trends survey. BMJ Open 2024 Nov 3;14(10):e086418. [doi: 10.1136/bmjopen-2024-086418] [Medline: 39489536]
47. Silver RA, Subramaniam C, Stylianou A. Association between patient portal use and office visits. J Med Syst 2021 Jun 7;45(7):74. [doi: 10.1007/s10916-021-01749-0] [Medline: 34097149]
48. Bhatla A, Ding J, Mhaimeed O, et al. Patterns of telehealth visits after the COVID-19 pandemic among individuals with or at risk for cardiovascular disease in the United States. J Am Heart Assoc 2024 Sep 3;13(17):e036475. [doi: 10.1161/JAHA.124.036475] [Medline: 39206726]

XSL•FO
RenderX

Abbreviations

HINTS: Health Information National Trends Survey
LASSO: Least Absolute Shrinkage and Selection Operator
MICE: multivariate imputation by chained equation
OR: odds ratio

Edited by PH Chau; submitted 13.Sep.2025; peer-reviewed by S Mohanadas, X Ding; revised version received 21.Jan.2026; accepted 21.Jan.2026; published 03.Mar.2026.

Please cite as:
Park J, Hou SI
Identifying Factors Associated With Patient Portal and Synchronous Telehealth Use Across Age Groups in the Postpandemic Era: Retrospective Analysis of the Health Information National Trends Survey
JMIR Aging 2026;9:e83730
URL: https://aging.jmir.org/2026/1/e83730
doi:10.2196/83730

© Jaeyoung Park, Su-I Hou. Originally published in JMIR Aging (https://aging.jmir.org), 3.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital-Human Public Community Care Integration for Chronic Pain in Low-Income Older Adults in a 6-Week Living Lab Setting: Quasi-Experimental Feasibility Study

Sunmi Song[1,2,3,4], MA, PhD; Seo-Yeon Hwang[5], MPH; Hae-Young Kim[4,5,6], DDS, PhD; Junesun Kim[1,2,3,4], PT, PhD

[1]College of Health Science, Department of Physical Therapy, Korea University, Seoul, Republic of Korea

[2]Rehabilitation Science Program, Department of Health Sciences, Graduate School, Korea University, Seoul, Republic of Korea

[3]Department of Health and Environmental Science, Undergraduate School, College of Health Science, Korea University, Seoul, Republic of Korea

[4]BK21FOUR Program: Learning Health Systems, College of Health Science, Korea University, Seoul, Republic of Korea

[5]Department of Public Health Sciences, Graduate School, Korea University, Seoul, Republic of Korea

[6]Department of Health Policy and Management, College of Health Science, Korea University, Seoul, Republic of Korea

Corresponding Author:
Junesun Kim, PT, PhD
College of Health Science, Department of Physical Therapy
Korea University
665 Hana Health Science B BLD, 145 Anam-ro, Sungbuk-gu
Seoul, 02841
Republic of Korea
Phone: 82 232905689
Email: junokim@korea.ac.kr

Abstract

Background: Digital health technologies offer promising solutions for managing chronic pain and depression in older adults, yet low-income populations with limited digital literacy face substantial barriers to access. Community-based approaches that leverage existing care infrastructure may bridge this digital divide, but evidence remains limited on effective integration strategies for digitally excluded populations.

Objective: We evaluated a multimodal digital health monitoring platform for this underserved population through 6-week living lab trials with 86 low-income Korean older adults (43 intervention and 43 age- and gender-matched controls) and 25 community caregivers.

Methods: Participants were recruited offline from recipients of government-supported community care services. The platform integrated chatbot surveys, smartwatch monitoring, and motion sensors, generating personalized traffic-light alerts when health indicators deviated from individual baselines and triggering real-time caregiver notifications for persistent declines. Mobile apps for older adult users and caregivers, along with a centralized monitoring system for community center managers, were provided to the intervention group, while controls received standard community care. Pre- and posttest surveys assessed changes in pain-related functional limitations, depressive symptoms, sleep quality, and system usability through face-to-face assessments. Caregivers documented platform-triggered interventions.

Results: Following attrition due to illness (n=3) and participation burden (n=4), we analyzed 1318 days of continuous monitoring data from 35 intervention participants (mean 37.7, SD 8.8 days per participant), with 77 participants (35 intervention and 42 control) completing pre-post assessments. Distinct patterns emerged in 24-hour heart rate profiles: participants experiencing higher-than-usual pain demonstrated elevated heart rates during early morning (5-8 AM) and late evening (10-11 PM) hours compared to their low-pain days. Multilevel modeling revealed significant within-person associations between digital biomarkers and symptoms. Heart rate variability (P=.02) and moderate physical activity (P=.05) were associated with same-day pain. Heart rate variability (P=.03) and long wake episodes during sleep (P=.03) predicted next-day pain. Shorter sleep duration (P=.03) and lower sleep efficiency (P=.05) were associated with same-day depressive symptoms. A significant group-by-time interaction was observed for pain-related functional limitations (P=.03): the intervention group maintained baseline levels while the comparison group reported increased functional limitations. Depressive symptoms, sleep quality, and platform usability did not show significant changes. Community caregivers successfully conducted 37 health decline–triggered interventions during regular service hours (25 hours per week), demonstrating integration of digital monitoring with existing care workflows.

XSL•FO
RenderX

Conclusions: Integrating digital health monitoring within existing community care infrastructure may be a feasible approach to supporting vulnerable older adults. The intervention was associated with maintained pain-related functional limitations, while depressive symptoms and sleep quality showed no significant changes, possibly due to the 6-week duration. This feasibility study provides preliminary support for digital-human integration models to address health equity, though randomized controlled trials are needed to establish causal relationships and isolate active intervention components.

Trial Registration: ClinicalTrials.gov NCT06270121; https://clinicaltrials.gov/study/NCT06270121

(JMIR Aging 2026;9:e85611) doi:10.2196/85611

KEYWORDS

aging in place; community caregivers; depression; digital divide; digital health; health equity; pain management; quasi-experimental study; wearable sensors

Introduction

The global demographic shift toward population aging has created an urgent need for innovative health care solutions [1,2]. Digital health technologies offer unprecedented opportunities for continuous monitoring and personalized care [3,4], particularly valuable for older adults managing multiple chronic conditions, including pain and depression [5,6]. However, the populations who could benefit most from these innovations—low-income older adults with limited digital literacy [7,8]—face the greatest barriers to accessing them [2,9]. This digital health divide is especially pronounced among socially isolated older adults, who experience higher rates of depression and chronic pain while lacking the support networks essential for technology adoption and health management [10,11].

Digital phenotyping—the continuous collection of behavioral and physiological data through digital devices—has shown promise in detecting and managing pain and mental health conditions in various populations [12-17]. Studies have demonstrated associations between wearable sensor data and depression symptoms, with heart rate variability (HRV), sleep patterns, and physical activity (PA) serving as reliable indicators of mental health status. Similarly, smartphone-based assessments have successfully captured daily fluctuations in pain and functional limitations among adults with chronic conditions. Recent evidence suggests that continuous monitoring can detect symptom changes before clinical deterioration, enabling timely interventions [3,18-20]. However, most digital phenotyping research has focused on younger, digitally literate populations, leaving a critical gap in understanding how these technologies might benefit older adults who face both high disease burden and significant technology barriers [5,21].

Multiple factors contribute to the digital exclusion of older adults. Age-related sensory and cognitive changes create difficulties in learning new technologies [22], while low income limits access to devices and internet connectivity [7,23,24]. Social isolation compounds these challenges, as older adults lack the informal support networks that facilitate technology learning [2]. Previous digital health interventions for older adults have shown mixed results, often failing due to high dropout rates and usability issues [25,26]. Even well-designed platforms struggle when they require direct interaction from users with limited digital skills and cognitive impairments [27,28]. These failures highlight a fundamental design flaw: most digital health solutions assume a level of digital literacy that many vulnerable older adults do not possess, perpetuating rather than addressing health inequities.

Emerging evidence suggests that community-based approaches may overcome these barriers. Studies have shown improved outcomes when digital interventions are delivered through trusted intermediaries such as nurses or trained caregivers [29-31]. Living lab methodologies, which embed technology development within users' natural environments, have demonstrated success in creating solutions that are feasible and sustainable in real-world settings [29,32]. In many countries, existing community care infrastructure already provides regular contact between formal caregivers and vulnerable older adults. These established relationships and routine visits represent an untapped opportunity for digital health delivery. By positioning caregivers as technology intermediaries rather than expecting direct use by older adults, digital health platforms could leverage existing human infrastructure to bridge the digital divide.

This study addresses this critical gap by evaluating the feasibility and preliminary effectiveness of integrating multimodal digital health monitoring within existing community care infrastructure for low-income, socially isolated older adults in South Korea. Using a 6-week living lab quasi-experimental design with an intervention and matched comparison group, we aim to answer the following three research questions: (1) Can digital biomarkers derived from continuous monitoring predict daily variations in pain and depression among older adults? (2) Does digital monitoring integrated with existing care services improve functional health outcomes compared to usual care? (3) Is it feasible to integrate digital monitoring within existing community care infrastructure without disrupting established workflows? We hypothesized that daily fluctuations in HRV, sleep quality, and PA would be associated with same-day and next-day pain and depressive symptoms (hypothesis 1 [H1]); participants receiving the integrated intervention would demonstrate improved or maintained functional health outcomes compared to those receiving usual care alone (hypothesis 2 [H2]); and community caregivers would successfully integrate platform-generated health alerts into their existing care workflows, enabling timely interventions for detected health declines without requiring additional work hours or service costs (hypothesis 3 [H3]. This quasi-experimental design would allow feasibility evaluation of an integrated digital-human care model rather than isolated digital platform effects.

XSL•FO
RenderX

Based on previous formative and pilot study [33], our approach integrates continuous digital phenotyping through smartwatches and daily chatbot surveys with existing community care services, where caregivers serve as technology intermediaries (Figure 1).

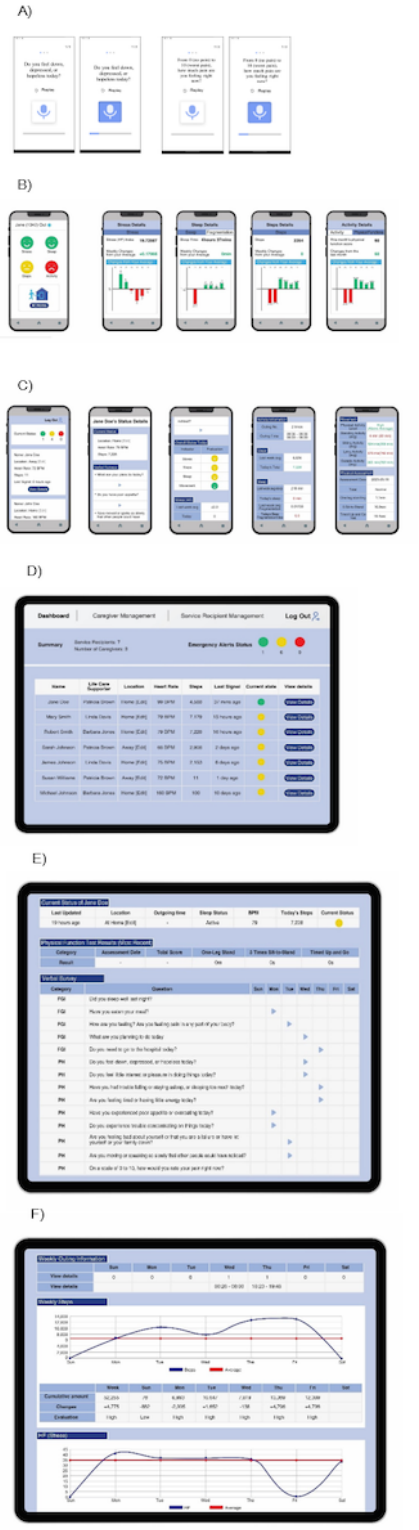
The platform's traffic-light feedback system simplifies complex health data into actionable information for both users and caregivers (Figure 2).

Figure 1. Community-mediated digital health platform architecture for vulnerable older adults. The 3-tiered service delivery system enables comprehensive health monitoring and care coordination. At the individual level, older adults in their homes interact with smartphone apps, smartwatches, and motion sensors for continuous health monitoring. When personalized algorithms detect health deterioration signals, assigned community caregivers provide targeted support through phone calls or home visits. For severe health declines or emergencies requiring coordinated response, the system escalates to the community center for comprehensive intervention. This hierarchical architecture ensures appropriate care intensity while maintaining older adults' independence in their home environment.



XSL•FO
RenderX

Figure 2. Platform interfaces demonstrating the traffic-light feedback system. (A) User application showing morning chatbot interaction for pain and depressive symptom reporting. (B) Personalized health status display using intuitive traffic-light indicators (green: improvement; yellow: stable; and red: decline from baseline) and detailed weekly pattern views for 4 health domains (stress, sleep, steps, and activity). (C) Caregiver dashboard displaying multiple recipients' health status. (D-F) Administrative interface for center managers showing aggregate health trends, alert history, and intervention tracking. This design simplifies complex health data into actionable information for major stakeholders. A higher resolution version of this image is available in Multimedia Appendix 1.



XSL•FO
RenderX

Methods

Overview

This study represents the culmination of a multiphase research program aimed at developing and evaluating a community-integrated digital health monitoring platform for socially vulnerable older adults. The development process began with formative research involving surveys (n=99 older adults) and focused group interviews with older adults (n=16), community caregivers (n=12), and social workers (n=3) to identify primary health concerns, understand community care workflows, and determine essential platform features. This informed the platform design, which was subsequently refined through an informal prepilot test with study staff (n=6) and a 6-week single-arm living lab pilot study (n=31) to establish feasibility and optimize functionality, which was previously reported [33].

Building on these foundational phases, this study used a 6-week quasi-experimental design with matched comparison groups to evaluate the feasibility and preliminary effectiveness of integrating digital monitoring within existing community care services. The evaluation encompassed not only the digital monitoring components but also the integration with in-person community care services and the institutional emergency alert system that enables coordinated responses from caregivers and center managers when health decline is detected. The research was conducted at a community senior welfare center in Seoul, South Korea, from October to December 2023, following formative and pilot phases that spanned the COVID-19 pandemic period (2021-2022). This iterative development approach—progressing from needs assessment through pilot testing to quasi-experimental evaluation—ensured the platform was both technically robust and contextually appropriate for the target population of low-income older adults with limited digital literacy. This iterative process was guided by a participatory, user-centered approach that engaged older adult participants, community caregivers, and welfare center administrators during and after the pandemic period. Their input informed not only platform design but also key study design decisions, including the adoption of a matched comparison rather than a randomized design, which stakeholders collectively identified as essential for ethical engagement that respects the right of choice and dignity of this vulnerable population.

Older adult participants received the Korean Government–supported Community Care Service that was provided free of charge: twice-weekly phone calls and weekly home visits from assigned caregivers, with additional follow-up services provided for acute illness episodes. Caregiver work hours are regulated at 25 hours per week under the service protocol, and poststudy caregiver surveys confirmed adherence to this standardized schedule in the intervention group without overtime work. The platform-triggered interventions in the intervention group aimed to facilitate earlier detection of health decline events; caregivers were briefly trained to use the platform to identify emerging health issues through digital alerts before these became apparent during scheduled contacts or escalated to acute episodes.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Boards of Korea University (approval number KUIRB-2021-0324) and Sahmyook University (approval number SYU 2022-06-003-001). Written informed consent was obtained from all individual participants included in the study prior to enrollment. Safety monitoring was conducted through community caregivers who served as technology intermediaries. Caregivers received daily health status updates via the traffic-light dashboard and responded to health deterioration alerts with phone calls or in-person visits, enabling continuous safety monitoring and timely response to detected health concerns. Privacy was protected through secure data transmission and restricted access to participant health information, with only assigned caregivers, a community center manager, and authorized research staff permitted to view individual health data.

Recruitment

Participants were identified through the Community Care service database at a municipal senior welfare facility in Seoul. Community caregivers facilitated recruitment by identifying eligible older adults from their existing caseloads and arranging initial contact with the research team.

Inclusion criteria comprised (1) age of 65 years or older, (2) living alone with low income qualifying for government-supported care services, (3) currently receiving biweekly phone calls and weekly home visits through the Community Care service, (4) possession of an Android smartphone compatible with the study application, and (5) capacity to provide informed consent. Individuals were excluded if they had cognitive impairment preventing study participation, severe sensory impairments interfering with technology use, cohabitation with others (due to motion sensor limitations), or medical conditions requiring intensive care.

The intervention group (n=43) consisted of older adults whose caregivers agreed to use the digital platform, while the comparison group (n=43) included age- (±3 years), gender-, and residential area–matched older adults whose caregivers maintained standard care practices. A total of 25 community caregivers participated as technology intermediaries for the intervention group.

Digital Health Monitoring Platform

The platform architecture integrated 3 synchronized applications: a user interface providing personalized health feedback, a caregiver dashboard for monitoring multiple recipients, and an administrative system for center managers (Figure 1). Daily health indicators were visualized using an intuitive traffic-light system (green: better than baseline, yellow: similar to baseline, red: worse than baseline), calculated from each participant's first-week average (Figure 2). Detailed technical specifications and user interface designs of the platform have been previously described in our pilot study [33]. The results from the motion sensing camera, which was integrated with experimental sessions outside the living lab, are reported separately from this study [34].

XSL•FO
RenderX

Technical components included a chatbot-enabled smartphone app for daily surveys, continuous physiological monitoring via Fitbit Sense smartwatches, and Azure Kinect motion sensors for home-based activity detection. While motion-sensing cameras were deployed as part of the platform infrastructure, their biomechanical data were analyzed in a concurrent validation study and are presented separately [34].

Physiological measures were derived from the Fitbit Sense smartwatch, a consumer-grade device. While Fitbit devices have demonstrated acceptable validity for PA, resting heart rate, and irregular heart rhythms compared to research or clinical-grade sensors [35,36], heart rate measurements during active periods and sleep staging algorithms show greater variability against polysomnography gold standards, particularly in populations with sleep disorders or arrhythmias [37,38]. Our older adult participants' high prevalence of chronic conditions may have introduced measurement noise. Thus, the use of a consumer-grade sensor in this study may have attenuated observed associations or introduced measurement error.

Intervention Protocol

Following informed consent and baseline assessment, research staff and caregivers conducted home visits to install applications and provide initial training. Participants engaged with the chatbot each morning, responding to 2 randomly selected Patient Health Questionnaire-9 (PHQ-9) items for depressive symptom screening and rating pain levels (0-10 visual analog scale [VAS]). Continuous smartwatch monitoring captured HRV, sleep patterns, and PA throughout the study period.

Community caregivers received daily updates on participants' health status through the traffic-light system displayed on their dashboard. When participants exhibited multiple red indicators (≥2 domains) for 2 or more consecutive days, caregivers were instructed to conduct follow-up interventions via phone calls or in-person visits based on severity assessment. All caregiver-initiated interventions triggered by health decline signals were systematically documented and analyzed (Figure S1 in Multimedia Appendix 2 for examples of health decline signals). Table S3 in Multimedia Appendix 2 shows the records of the community caregiver's follow-up reports after detecting health decline signals. This quasi-experimental study with both an intervention and a comparison group has been registered on ClinicalTrials.gov (NCT06270121).

Daily Digital Measures of Pain and Depression

Intervention participants completed daily assessments via the chatbot interface that was developed in our pilot study [33]. For depressive symptom screening, 2 items were randomly selected from the PHQ-9 questionnaire and presented in a yes/no format each day. Participants who responded "yes" to one or both items were coded as experiencing depressive symptoms for that day. Pain intensity was assessed using a VAS (0-10, where 0=no pain, 10=worst imaginable pain). Response rates and patterns were analyzed as engagement indicators.

Digital Biomarker Measurements

Continuous physiological monitoring via Fitbit Sense smartwatches captured 3 domains of digital biomarkers throughout the 6-week intervention period, as previously described [33]. Sleep metrics included total sleep time, sleep fragmentation index (awakening frequency divided by total sleep time), number of long wake episodes (over 5 minutes of awakening), and sleep efficiency (total sleep time divided by time in bed), calculated using validated algorithms that applied 7 correction criteria to improve actigraphy-based sleep detection accuracy [39,40].

HRV was computed every 5 minutes using the open-source "Aura-healthcare" Python package [41], generating both time-domain indicators (SD of N-N intervals and root-mean-square of successive differences) and frequency-domain measures (high-frequency/low-frequency ratio) from R-R interval data [41,42].

PA levels were quantified through daily step counts and minutes of light, moderate, and intense activity, leveraging Fitbit's validated 3-axis accelerometer algorithms [35,43,44]. These multidomain digital biomarkers provided comprehensive 24-hour physiological profiles for examining within-person associations with daily pain and depressive symptoms.

Pre- and Postsurvey Measures

Primary and secondary outcomes were assessed through structured home visits at baseline, week 3 (midintervention), and week 6 (postintervention) by trained research assistants. Caregivers also participated in brief surveys at pretest and posttest.

Pain-Related Functional Limitations

The Korean version of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) was administered to assess pain-related functional impairment [45,46]. WOMAC comprises 24 items evaluating 3 dimensions: pain (5 items), stiffness (2 items), and physical function (17 items). Each item is scored on a 5-point Likert scale (0-4), with higher total scores indicating greater functional limitations.

Geriatric Depression Scale Short-Form

The 15-item Geriatric Depression Scale Short-Form (GDS-SF) was used to screen for depressive symptoms [47,48]. Each item requires a yes/no response, with scores ranging from 0 to 15. Clinical cutoffs indicate normal (0-4), mild depression (5-9), and severe depression (≥10). The GDS-SF has shown high sensitivity (92%) and specificity (89%) for detecting depression in Korean older adults.

Pittsburgh Sleep Quality Index

The Pittsburgh Sleep Quality Index (PSQI) assessed subjective sleep quality over the past month [49,50]. This 19-item questionnaire evaluates 7 components: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medications, and daytime dysfunction. Global PSQI scores range from 0 to 21, with scores >5 indicating poor sleep quality. The Korean PSQI has demonstrated good internal consistency (Cronbach α=0.84) and test-retest reliability (r=0.85).

XSL•FO
RenderX

Technology Usability

System usability was measured using a simplified System Usability Scale (SUS) adapted for cognitively impaired and older adults [51]. The scale contains 10 items rated on a 5-point scale (strongly disagree to strongly agree), yielding scores from 0 to 100. Higher scores indicate better perceived usability, with scores >68 considered above average. Preintervention SUS assessed general technology use, while postintervention SUS specifically evaluated the digital health platform. To assist participants unfamiliar with digital technology, examples (smartphones and hospital kiosks) were provided during baseline assessment.

In addition to the SUS, participants were asked about the frequency of platform use and the number of functions used during the midterm assessment. At postintervention, participants provided qualitative feedback for platform improvement, which research assistants documented verbatim. Community caregivers similarly completed the SUS and reported their usage frequency and range of functions used at the postintervention assessment.

Covariates

Covariates were determined a priori based on extant literature summarizing demographic and health risk factors for daily depressive symptoms of older adults [52,53]. In addition to the baseline levels of depression, demographic and health factors linked with depression (age, sex, and chronic health conditions) were used as covariates.

Caregiver Assessments

Caregivers participated in brief surveys at pretest about their demographic and job-related information, and at posttest about their usage of the caregiver app. Poststudy surveys assessed the SUS survey, caregiver platform engagement, including working hours compliance, usage frequencies, function use, and qualitative feedback on workflow integration and user experience.

Statistical Analysis

After analyzing missing data and assessing the distributions of key variables, descriptive analyses were performed. Group comparisons between older adults with depression and without depression were conducted using *t* tests for continuous variables (eg, age, BMI, PSQI total score, daily depressive symptoms, wearable sensor assessments of HRV, sleep, and PA indicators), chi-square tests for categorical variables (eg, sex, education, income, chronic disease, and sleep disorder categories). Subsequently, bivariate correlations were explored between baseline depression and person mean variables (daily pain and depressive symptoms, and wearable sensor assessments) over 6 weeks.

Within-person associations between digital biomarkers and daily symptoms were examined using multilevel models (SAS/STAT version 9.4; SAS Institute Inc). Level 1 examined daily variations in symptoms and sensor metrics; level 2 incorporated participant characteristics. It allows for the use of participants as their own controls [54].

For pain as outcome, the model included daily depressive symptoms (binary coded: 0=no symptoms, 1=responded “yes” to at least 1 of 2 daily PHQ-9 items), sleep, HRV, and PA indices were examined as predictors of same-day and next-day pain levels. For depressive symptoms as outcome, the model included daily pain levels (VAS 0-10), sleep, HRV, and PA indices as predictors of same-day and next-day depressive symptoms using logistic multilevel models, given the binary nature of the outcome.

All level 1 predictors were person mean–centered, while level 2 covariates (age, sex, and chronic conditions) were grand mean-centered. This approach allows interpretation of estimates as within-person deviations from individual averages across the 6-week period. Bonferroni corrections were applied for multiple testing (α=.005).

The pretest and posttest results were assessed using a mixed model with time (pretest and posttest) at level 1, personal characteristics (covariates of age, gender, and chronic disease conditions) at level 2, and intervention group at level 3. All plots of results were drawn using the R software (version 4.3.1; R Foundation for Statistical Computing) ggplot2 package [55].

Power Analysis

Simulation studies demonstrate that 35 participants with >30 days of continuous monitoring yields adequate statistical power for detecting within-person effects in multilevel models [56]. For pre-post comparisons, 72 participants (32 intervention and 40 comparison) completed end point assessments, providing 80% power to detect medium effect sizes (d=0.50) in group × time interactions. Although there is no prior study to estimate effect sizes from older adult populations, this is consistent with meta-analytic evidence from human-guided digital interventions for depression in general adult populations (Hedges g=0.52-0.63) [57], which represents the closest available benchmark for digital-human integration models. Given the modest sample size, findings should be interpreted as preliminary rather than definitive evidence of intervention effectiveness.

Data Exclusion and Missing Data

Of 86 enrolled participants, 77 provided analyzable data (35 intervention and 42 comparison), with 9 exclusions due to withdrawal (n=4 participation burden), illness (n=2), relocation (n=1), or incomplete baseline data (n=2). The intervention group generated 1318 observation days (mean 37.7, SD 8.8 days per participant) for multilevel analyses.

Missing data analyses showed missing in daily surveys (455/1318, 34.5%), HRV (356/1318, 27%), sleep monitoring (625/1218, 47.4%), and PA (126/1318, 9.6%). Missing data primarily resulted from forgetting to wear devices or charging issues, rather than technical failures or dropout. Full information maximum likelihood estimation addressed missing data, a method shown to be robust against bias with partial missingness [58].

Protocol Deviations

The trial registration (ClinicalTrials.gov NCT06270121) originally specified the Short-Form 12 health survey as the primary quality of life outcome. During implementation, we modified the primary outcome to the WOMAC to focus specifically on pain-related functional limitations, as baseline

XSL•FO
RenderX

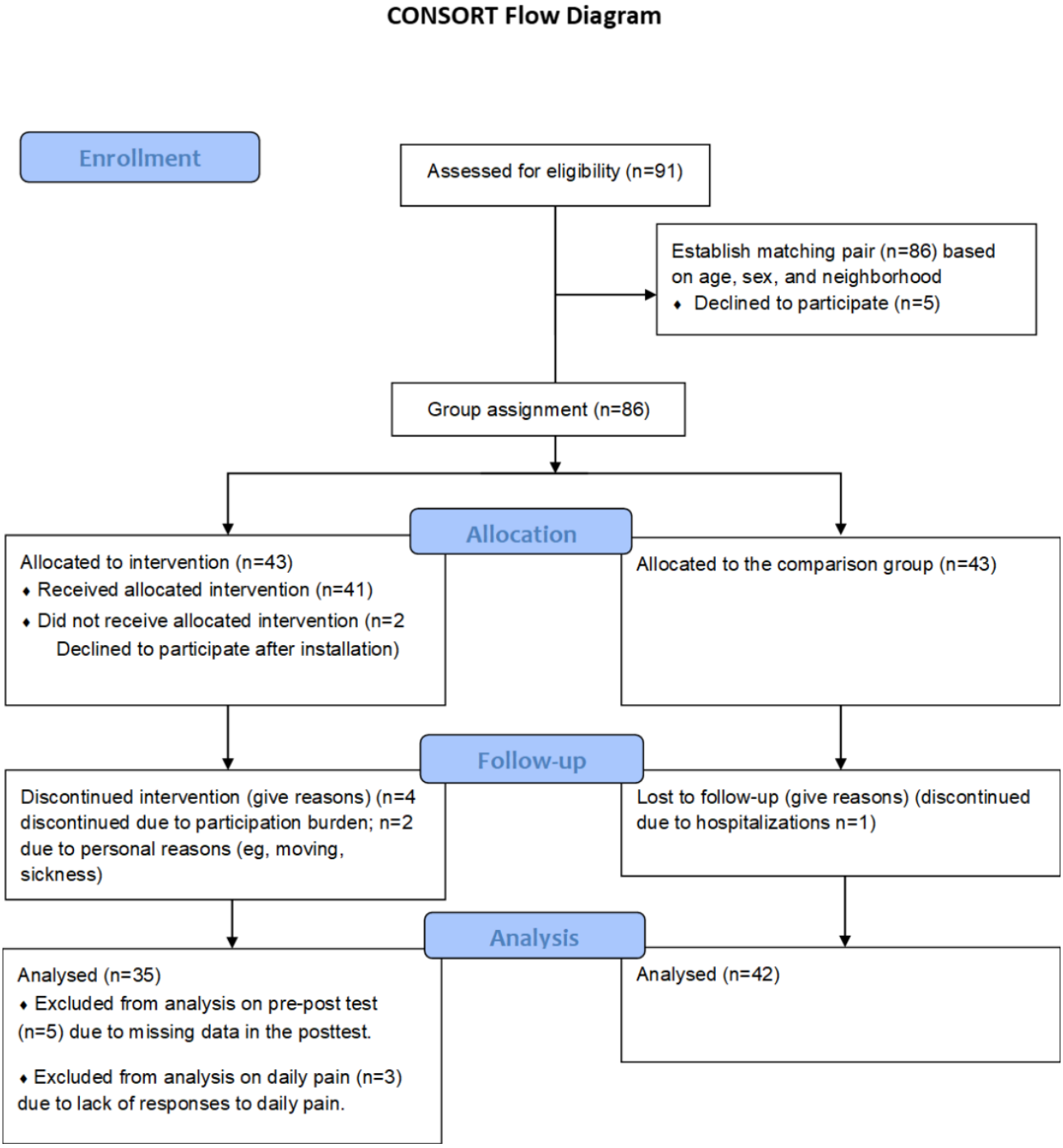
assessments revealed that 90% (70/73) of participants experienced chronic pain, making functional outcomes more clinically relevant and responsive to our 6-week intervention than generic quality of life measures. Additionally, the PSQI was added as a secondary outcome after initial data showed that 69.2% (54/78) of participants exceeded the clinical cutoff for sleep disturbances (PSQI >5), highlighting the importance of sleep assessment in understanding the pain-mood-sleep triad in this population.

Results

Participant Characteristics

Between April and November 2023, we enrolled 78 low-income older adults living alone in Seoul, South Korea (35 intervention and 43 comparison) and 25 community caregivers. The digital health monitoring platform design and service delivery model are illustrated in Figures 1 and 2. Participant flow through the study is shown in Figure 3.

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of participant recruitment and retention. Study enrollment occurred in October-November 2023 in Seoul, South Korea. Of 91 eligible older adults approached, 86 enrolled (43 matched pairs). The intervention group received digital health monitoring integrated with community care services; the comparison group received usual care. Attrition was primarily due to illness (n=3) or participation burden (n=4). Analysis includes 35 intervention and 42 comparison participants for pre-post outcomes, with 35 intervention participants contributing 1318 days of digital phenotyping data.



XSL•FO
RenderX

Baseline characteristics demonstrated comparable demographics and high disease burden across groups (Table 1). Mean age was 78 (SD 4.1) years, 85% (66/78) were women, and participants had limited education (mean 8.2, SD 4.7 years). Monthly income averaged KRW 900,000 (approximately US $670), confirming socioeconomic vulnerability. Participants reported an average of 4 chronic conditions, predominantly hypertension (46/78, 59%), diabetes mellitus (45/78, 58%), cataracts (43/78, 55%), and arthritis (25/78, 49%). Notably, 90% (70/78) experienced chronic pain lasting over 3 months, primarily affecting the back (50/78, 64%) and knees (39/78, 50%). Baseline assessments revealed mild depressive symptoms (mean GDS-SF score 4.3, SD 3.5), poor sleep quality (mean PSQI score 8.0, SD 3.8, exceeding the clinical cutoff of 5), and moderate pain-related functional limitations (mean WOMAC score 29.1, SD 19.0). Groups showed no significant baseline differences (all $P>.05$). Characteristics of caregivers and their platform usage were also reported in Table 2.

Table 1. Baseline characteristics of older adult participants demonstrate comparable demographics and high disease burden across groups.

Characteristics	Total (N=78)	Group assignments		*P* value
		Intervention group (n=35)	Comparison group (n=43)	
Demographic characteristics				
Age (years), mean (SD)	78.03 (4.10)	78.34 (3.15)	77.77 (4.76)	.53
Sex: women, n (%)	66 (84.62)	28 (80)	38 (88.37)	.31
Years of education, mean (SD)	8.21 (4.70)	7.17 (4.39)	9.05 (4.83)	.08
Monthly income (KRW[a]), n (%)				
Less than 500,000	17 (21.79)	8 (22.86)	9 (20.93)	.19
500,001-1,000,000	40 (51.28)	20 (57.14)	20 (46.51)	—[b]
1,000,001-1,500,000	9 (11.54)	5 (14.29)	4 (9.30)	—
1,500,001 or more	12 (15.38)	2 (5.71)	10 (23.26)	—
Currently employed	14 (18.18)	8 (22.86)	6 (14.29)	.33
Health characteristics, mean (SD)				
BMI	24.73 (3.53)	24.05 (3.06)	25.27 (3.81)	.14
Number of chronic diseases	4.24 (1.83)	4.31 (2.05)	4.18 (1.65)	.76
Baseline depressive symptoms (GDS-SF[c], 0~15)	4.28 (3.53)	4.42 (3.66)	4.16 (3.46)	.74
Baseline sleep quality (PSQI[d], 0~21)	8.02 (3.82)	8.06 (3.63)	8.00 (4.01)	.95
Baseline pain status				
Pain duration, n (%)				
Chronic pain >3 months	70 (89.74)	29 (82.86)	41 (95.35)	.07
Acute pain	4 (5.13)	4 (11.43)	0 (0)	—
No pain	4 (5.13)	2 (5.71)	2 (4.65)	—
Pain locations, n (%)				
Back	50 (64.10)	24 (68.57)	26 (60.47)	.46
Knee	39 (50)	13 (37.14)	26 (60.46)	.04
Neck and shoulder	21 (26.92)	7 (20)	14 (32.56)	.21
Calf	17 (21.79)	11 (31.43)	6 (14)	.06
Others	34 (43.59)	16 (45.71)	18 (41.86)	.73
Baseline pain severity (WOMAC[e], 0-100)	29.06 (18.96)	29.57 (19.88)	28.65 (18.41)	.83

[a]1 KRW ≈ US $0.00074.

[b]Not available.

[c]GDS-SF: Geriatric Depression Scale Short-Form.

[d]PSQI: Pittsburgh Sleep Quality Index.

[e]WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

XSL•FO
RenderX

Table 2. Characteristics and the platform usage of community caregivers.

Characteristics	Total (n=25)
Demographics characteristics	
Age (years), mean (SD)	61.24 (10.1)
Sex (women), n (%)	25 (100)
Education, n (%)	
Elementary school or less	2 (8)
Middle school	2 (8)
High school	6 (24)
College or above	13 (52)
Choose not to report	2 (8)
Marital status, n (%)	
Married	17 (68)
Widowed	4 (16)
Divorced	3 (12)
Never married	1 (4)
Job characteristics	
Work experiences (months), mean (SD)	63.56 (45.20)
Senior care provider (yes), n (%)	5 (20)
Number of care-providing older adults (persons), mean (SD)	10.96 (2.09)
Weekly hours of work during intervention, mean (SD)	25.4 (1.38)
Caregiver app usage[a]	
System Usability Scale (0-100) at posttest, mean (SD)	70.80 (16.67)
Frequency of use, n (%)	
More than 2-3 times per day	5 (21.7)
2-3 times per day	11 (47.8)
Once a day	6 (24)
Once every 2-3 days	1 (4.3)
Number of functions used, mean (SD)	6 (3.92)
Functions used, n (%)	
Listening to the voice recording of older adults	19 (82.6)
Steps of older adults	14 (60.9)
Stress levels of older adults	13 (56.5)
Activity at the home of older adults	12 (48)
Sleep of older adults	12 (52.2)
Emergency status	6 (27.3)
Etc (eg, weekly patterns or health status alerts)	14 (56.2)

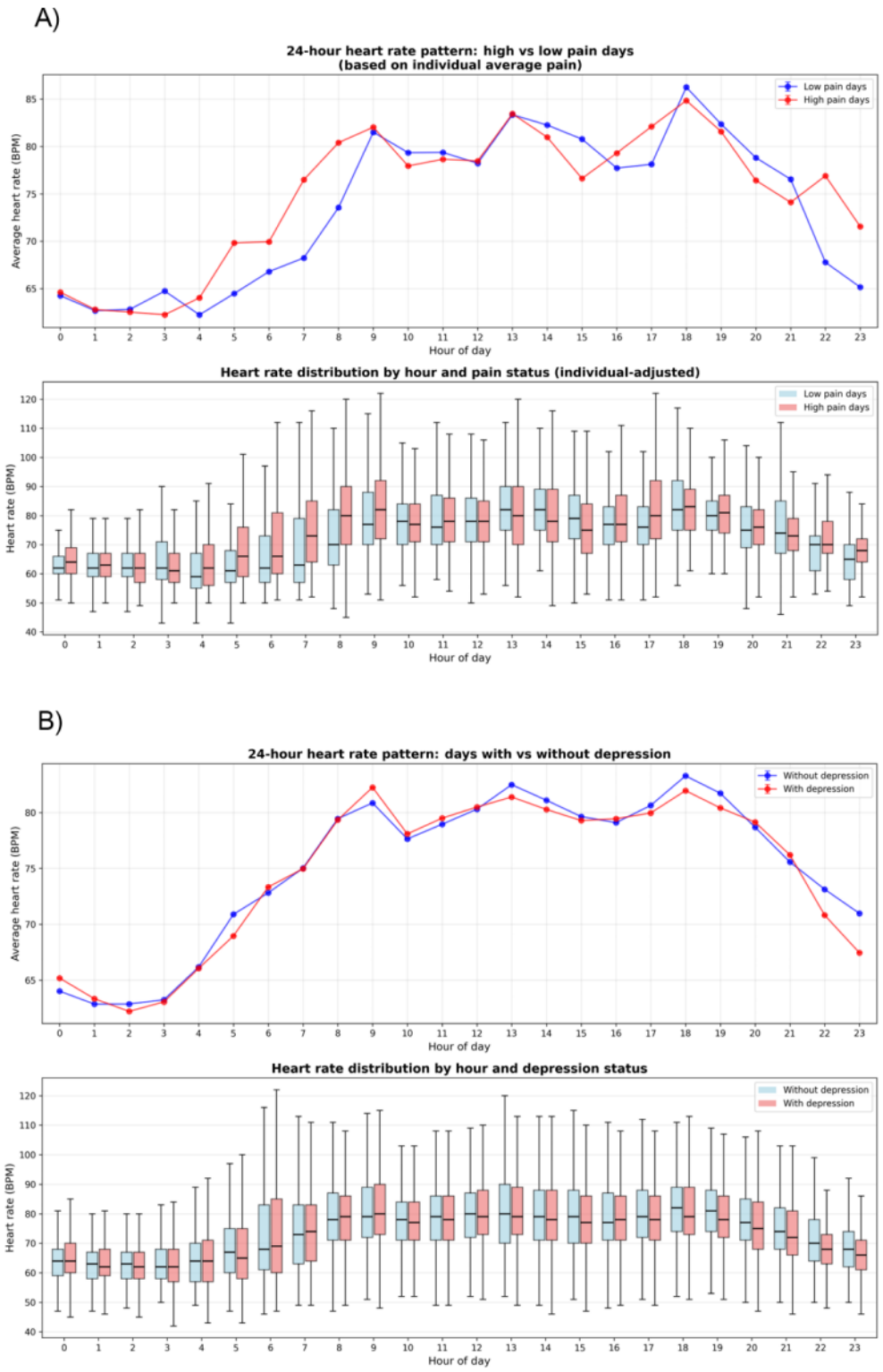
[a]2 caregivers were not able to participate in the posttest surveys due to schedule conflicts.

Descriptive Results

We analyzed 1318 days of continuous monitoring data from 35 intervention participants (mean 37.7, SD 8.8 days per participant). Distinct patterns associated with daily pain states were revealed by 24-hour heart rate profiles (Figure 4). Participants experiencing higher-than-usual pain demonstrated elevated heart rates during early morning (5-8 AM) and late evening (10-11 PM) hours compared to their low-pain days (Figure 4A). In contrast, days with and without depressive symptoms showed similar heart rate patterns throughout the day (Figure 4B).

XSL•FO
RenderX

Figure 4. Distinct patterns associated with pain in older adults emerging in 24-hour heart rate profiles. (A) Heart rate patterns on days with higher (red) versus lower (blue) pain relative to individual average demonstrate elevated rates during early morning (5-8 AM) and late evening (10-11 PM) hours, (B) heart rate patterns on days with (red) versus without (blue) self-reported depressive symptoms show minimal variation across the day. Data represent mean (standard error of the mean) from 35 participants over 1318 observation days. Top panels show hourly averages with error bars; bottom panels display distributions as box plots with median, IQR, and outliers. BPM: beats per minute.



Correlation analyses revealed complex within-person and between-person associations among digital biomarkers and symptoms. Within-person correlations showed that daily fluctuations in sleep time (r=–0.112; P=.02) and sleep fragmentation (r=–0.114; P=.01) significantly correlated with depressive symptoms (Table S1 in Multimedia Appendix 2). Between-person correlations indicated that individuals with higher average daily pain levels had higher baseline WOMAC scores (r=0.499; P=.003), greater average daily depressive symptoms (r=0.475; P=.006), and greater baseline depressive symptoms (r=0.593; P<.001; Table S2 in Multimedia Appendix 2). Higher daily depressive symptoms are also associated with

XSL•FO
RenderX

root-mean-square of successive differences (r=0.356; P=.05), moderate PA (r=0.462; P=.02), and baseline depressive symptoms (r=0.476; P=.006).

Digital Predictors of Daily Pain and Depressive Symptoms

Multilevel modeling identified specific digital biomarkers predicting daily pain and depressive symptoms (Table 3). For pain, HRV (mean of the normal-to-normal intervals) significantly predicted both same-day pain (β=–.007, SE=0.003; P=.02) and next-day pain (β=.007, SE=0.003; P=.03). Moderate PA was associated with reduced same-day pain (β=–.601, SE=0.298; P=.047), while increased long wake episodes during sleep predicted higher next-day pain (β=.343, SE=0.151; P=.03). For depressive symptoms, participants were more likely to report symptoms on days following shorter sleep duration (odds ratio [OR] 0.998, 95% CI 0.997-1.000; P=.03) and lower sleep efficiency (OR 0.988, 95% CI 0.975-1.000; P=.05). Higher pain increased the likelihood of same-day depressive symptoms (OR 1.222, 95% CI 1.000-1.494; P=.05).

Table 3. Digital biomarkers predict daily pain and depressive symptoms in older adults. Multilevel models examine within-person associations between sensor-derived measures and self-reported symptoms over 1318 observation days from 35 participants. All predictors were person mean–centered to re-present deviations from individual baselines. Models adjusted for age, sex, and chronic conditions.

Digital predictors	Same-day VAS[a] pain		Next-day VAS pain		Same-day depressive symptoms		Next-day depressive symptoms	
	Estimate (SE)	*P* value	Estimate (SE)	*P* value	OR[b] (95% CI)	*P* value	OR (95% CI)	*P* value
Daily depressive symptoms	0.683 (0.392)	.08	0.585 (0.408)	.16	—[c]	—	—	—
Daily pain	—	—	—	—	1.222 (1.000-1.494)	.05	1.132 (.901-1.423)	.28
Heart rate variability								
HF/LF[d]	–0.165 (1.584)	.92	1.258 (2.039)	.54	0.969 (0.170-5.514)	.97	1.060 (0.129-8.727)	.96
SDNN[e]	–0.005 (0.006)	.44	–0.007 (0.007)	.27	0.450 (0.141-1.443)	.18	1.002 (0.996-1.008)	.56
Mean NNI[f]	–0.007 (0.003)	.02	0.007 (0.003)	.03	0.998 (0.996-1.001)	.28	1.000 (0.997-1.003)	.85
RMSSD[g]	–1.668 (0.989)	.09	1.752 (1.285)	.18	0.450 (0.141-1.442)	.18	0.288 (0.079-1.057)	.06
Sleep								
Total sleep time	0.001 (0.001)	.34	0.000 (0.002)	.80	0.998 (0.997-1.000)	.03	1.001 (0.999-1.002)	.32
Sleep fragmentation index	0.324 (0.234)	.17	0.508 (0.385)	.19	1.142 (0.815-1.602)	.44	1.317 (0.891-1.948)	.17
Number of long sleep fragmentations	0.148 (0.117)	.21	0.343 (0.151)	.03	1.019 (0.902-1.153)	.76	1.031 (0.902-1.177)	.66
Sleep efficiency (%)	–0.005 (0.012)	.65	–0.011 (0.013)	.42	0.988 (0.975-1.000)	.05	0.992 (0.979-1.005)	.20
Physical activity								
Light PA[h]	0.002 (0.002)	.25	–0.003 (0.002)	.09	1.000 (0.998-1.002)	.82	1.002 (0.999-1.004)	.16
Moderate PA	–0.601 (0.298)	.047	0.264 (0.364)	.47	0.996 (0.730-1.358)	.98	1.315 (0.934-1.851)	.12
Intense PA	–0.256 (0.307)	.41	0.228 (0.354)	.52	0.852 (0.637-1.139)	.28	1.283 (0.921-1.789)	.14

[a]VAS: visual analog scale.

[b]OR: odds ratio.

[c]Not available.

[d]HF/LF: high frequency/low frequency ratio.

[e]SDNN: SD of the normal-to-normal intervals.

[f]Mean NNI: mean of the normal-to-normal intervals.

[g]RMSSD: root-mean-square of successive differences.

[h]PA: physical activity.

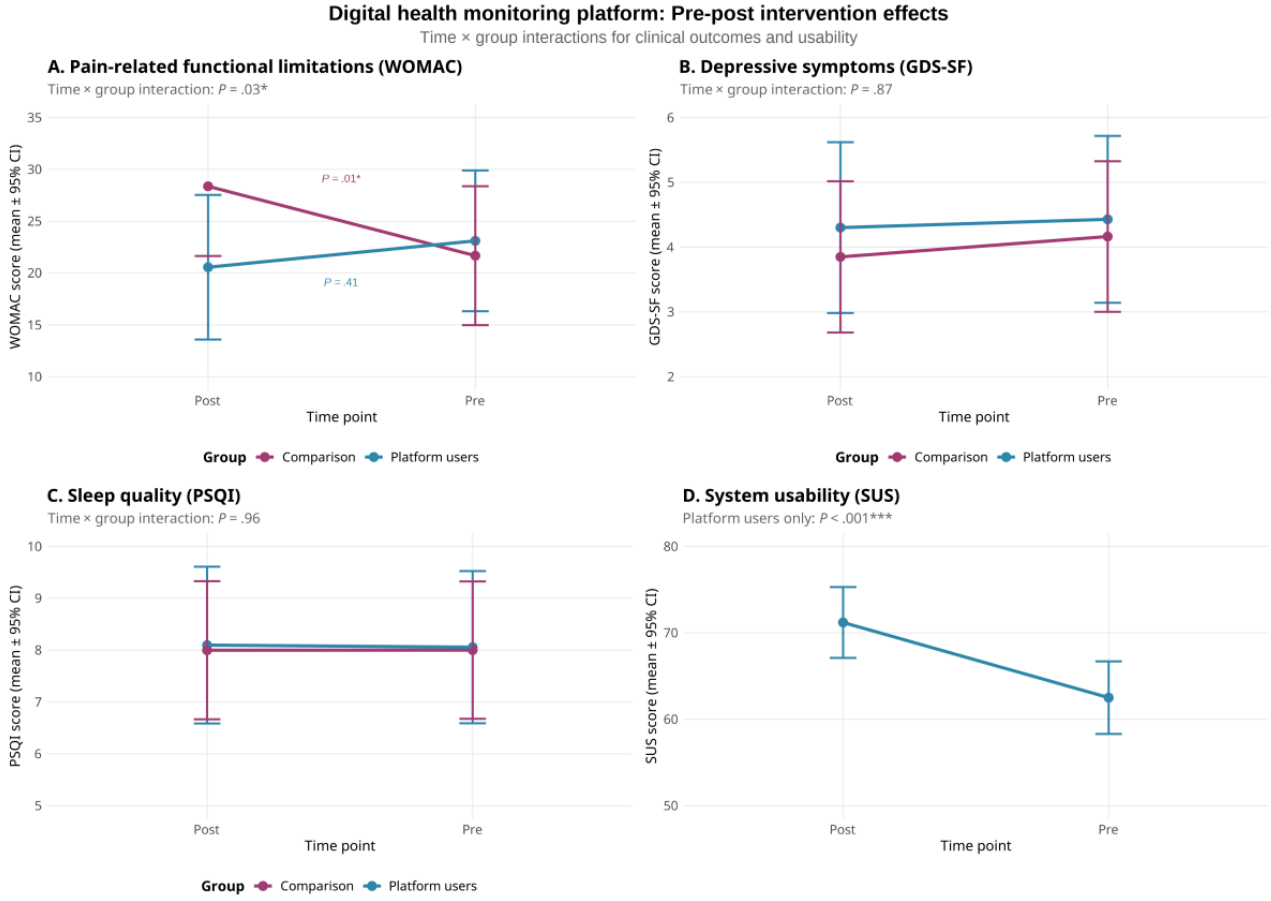
XSL•FO
RenderX

In unadjusted analyses, HRV (mean of the normal-to-normal intervals) was associated with both same-day and next-day pain (P<.05), and sleep duration was associated with next-day depressive symptoms (P<.05), providing initial support for H1. However, after Bonferroni correction for multiple testing (adjusted α=.005), none of these associations remained statistically significant. PA showed limited associations with mood outcomes across all analyses. Therefore, H1 was not supported.

Intervention Effects on Health Outcomes

The integrated intervention (digital monitoring plus enhanced caregiver support) yielded differential effects on health outcomes (Figure 5). Intervention participants maintained baseline levels of functional limitations (baseline: mean 23.1, SD 3.4 to post: mean 20.6, SD 3.5; change –2.6, 95% CI –7.8 to 2.6; P=.41), while the comparison group reported a significant increase in functional limitations (baseline: mean 21.7, SD 3.4 to post: mean 28.4, SD 3.4; change +6.7, 95% CI 2.3-11.1; P=.01). The between-group difference at postintervention was 7.8 (95% CI –16.3 to 0.6; P=.07) points, representing a potentially clinically meaningful difference that warrants confirmation in adequately powered trials (Figure 5A).

Figure 5. Changes in health outcomes over 6 weeks comparing integrated intervention with usual care. Pre-post intervention changes comparing platform users (blue) with the usual care comparison group (red). (A) Pain-related functional limitations (Western Ontario and McMaster Universities Osteoarthritis Index [WOMAC]) showed significant group × time interaction ($F_{1,72}$=5.26; ^{P}=.03), with platform users maintaining function (mean change –2.6, 95% CI –7.8 to 2.6) while the comparison group reported increased functional limitations (mean change +6.7, 95% CI 2.3 to 11.1). (B) Depressive symptoms (Geriatric Depression Scale Short-Form [GDS-SF]), (C) sleep quality (Pittsburgh Sleep Quality Index [PSQI]), and (D) system usability (System Usability Scale [SUS], platform users only) showed no significant changes. Data represent estimated marginal means (95% CI) adjusted for age, sex, and number of chronic conditions. Asterisk (*) indicates a significant interaction effect. Lower scores indicate better outcomes for WOMAC, GDS-SF, and PSQI; higher scores indicate better usability for SUS.



No significant group differences emerged for depressive symptoms (Figure 5B) or sleep quality (Figure 5C). Both groups maintained stable scores on these measures. System usability among intervention participants showed no significant changes from pre (mean 46.8, SD 20.5) to post (mean 44.9, SD 18.3) intervention (Figure 5D).

These findings partially support H2. Intervention participants demonstrated maintained functional health while the comparison group reported increased functional limitations, supporting our hypothesis for pain-related outcomes. However, the hypothesis was not supported for depressive symptoms or sleep quality, which showed no significant group differences.

Platform Engagement and Care Coordination

Despite limited baseline digital literacy, platform engagement remained robust throughout the intervention. Daily chatbot survey completion averaged 65% (863/1318), with 73%

XSL•FO
RenderX

(962/1318) compliance for heart rate monitoring, 53% (693/1318) for sleep tracking, and 90% (1192/1318) for activity monitoring. Missing data primarily resulted from device charging requirements or forgetting to wear devices rather than technical difficulties.

Table 2 presents caregiver's app usage at posttest. While maintaining the regular working hours (mean 25, SD 1.3 hours per week), community caregivers in the intervention group demonstrated strong platform engagement, with acceptable usability scores (mean SUS score 70.8, SD 16.7) and frequent daily use (16/23, 69.5% used ≥2-3 times daily). Caregivers actively used multiple functions (mean 6.0, SD 3.9), most commonly accessing voice recordings (19/23, 82.6%), step monitoring (14/23, 60.9%), and stress level tracking (13/23, 56.5%), indicating comprehensive adoption of the platform's digital monitoring features.

The platform appeared to enhance existing care services through 37 health decline–triggered interventions (Table S3 in Multimedia Appendix 1 for case reports and Figure S1 in Multimedia Appendix 1 for platform alert screenshots). When the platform detected concerning patterns (multiple red indicators for ≥2 consecutive days), caregivers conducted targeted interventions addressing: severe postprocedure pain (n=3), dizziness requiring medical attention (n=5), medication nonadherence (n=2), infections requiring hospitalization (n=1), and physical inactivity (n=26). One critical intervention identified an infected abscess requiring immediate surgical drainage, potentially preventing sepsis. All 25 community caregivers completed the study and actively responded to health alerts, demonstrating successful integration of digital monitoring with in-person care delivery.

Regarding safety, no serious adverse events were directly attributable to the platform. While 4 intervention participants withdrew citing burden and 3 participants discontinued due to unrelated illness or relocation, the 37 health decline–triggered interventions represent early detection rather than platform-induced harms.

H3 was supported. All 25 community caregivers completed the study and successfully integrated platform alerts into their workflows, conducting 37 health decline–triggered interventions. Poststudy surveys confirmed that caregivers did not exceed their contracted 25-hour work week, demonstrating that platform integration was achieved without additional work hours or service costs beyond the standardized care protocol.

Discussion

Principal Findings

This feasibility study provides preliminary evidence that integrating digital health monitoring with existing community care infrastructure may help bridge the digital divide for low-income older adults. The primary observation—that participants receiving the integrated intervention maintained functional health while the comparison group reported increased functional limitations—is consistent with potential benefits of the integrated digital-human care model, though the nonrandomized, quasi-experimental design precludes causal inference regarding which intervention components contributed to this differential pattern. The 37 health decline–triggered interventions by digital monitoring illustrate how technology may complement existing, rather than replace, human care relationships. However, it is important to note that the intervention group received caregiver interventions triggered by digital monitoring, while controls received standard care only. This fundamental confounding prevents isolation of digital platform effects from enhanced human contact. Our findings support the feasibility and preliminary effectiveness of digital-human integration but cannot establish independent clinical efficacy of the digital platform itself.

Our digital phenotyping results provide novel insights into pain and mood fluctuations among community-dwelling older adults. The distinct 24-hour heart rate patterns associated with pain—elevated rates during early morning and late evening—may reflect circadian variations in pain perception, previously observed only in controlled laboratory settings [59,60] but now demonstrated in older adults' daily living conditions. The association between HRV and both same-day and next-day pain extends previous cross-sectional findings of lower HRV in chronic pain patients [61,62], providing the first evidence that within-person HRV fluctuations may predict daily pain variations in older adults, suggesting autonomic dysregulation as an early indicator of pain exacerbation. Similarly, the relationship between sleep quality metrics and depressive symptoms aligns with established bidirectional links between sleep and mood in geriatric populations [63,64].

Regarding our hypotheses, results provided support for H2 and H3, but H1 was not supported after statistical correction. The digital phenotyping hypothesis (H1) was not supported: while HRV and sleep metrics showed initial associations with pain and depressive symptoms in unadjusted analyses, these relationships did not survive Bonferroni correction for multiple testing (adjusted α=.005). The effectiveness hypothesis (H2) received partial preliminary support: a differential pattern in functional health was observed, favoring the intervention group, but depression and sleep outcomes showed no group differences, possibly due to the short intervention duration and the confounding effects of enhanced caregiver support. The feasibility hypothesis (H3) was fully supported: caregivers successfully integrated digital monitoring into existing workflows without additional time burden, as confirmed by poststudy surveys verifying adherence to the standardized 25-hour weekly schedule. This finding is relevant for potential scalability, raising the hypothesis that integrated interventions may enhance care quality through better timing of interventions rather than requiring additional caregiver burden. This hypothesis requires confirmation through randomized designs.

Comparison With Prior Work

Previous digital health interventions for older adults have reported mixed results, with high attrition rates often limiting effectiveness [65,66]. A recent systematic review found that internet-delivered interventions for geriatric depression achieved modest effects but struggled with engagement, particularly among those with lower digital literacy [67-69]. Our approach differs fundamentally by not requiring direct technology use by

XSL•FO
RenderX

older adults themselves. This intermediary model extends traditional community health worker programs—proven effective in resource-limited settings [70-73]—by integrating digital monitoring tools with trained lay caregivers who facilitate early detection and timely intervention for mental and physical health concerns among vulnerable populations.

Our findings extend the digital phenotyping literature, which has predominantly focused on younger populations with sufficient digital literacy [4,16,74-76]. While previous studies have shown associations between PA and mood in controlled settings [77-80], our real-world data revealed more complex and weaker patterns that did not reach statistical significance. The observed trend toward a protective effect of moderate—but not intense—PA on pain aligns with recommendations for gradual, individualized exercise in older adults [81,82], though further research is needed to establish whether digital monitoring could help identify optimal activity ranges for personalized pain management [5].

Implications for Digital Health Equity

This study offers a preliminary model that, if validated through randomized trials, may be scalable for addressing digital health inequities. Rather than attempting to overcome multiple barriers simultaneously—teaching digital skills, providing devices, ensuring connectivity—our approach leverages 1 existing asset: regular contact with community caregivers. This transforms the digital literacy requirement from end users to intermediaries who typically possess better digital skills and can translate complex health data into actionable care for service-receiving older adults.

The current implementation of the developed platform in South Korea, where public community care services are well-established, suggests potential adaptability to other contexts with similar infrastructure. Many countries have comparable programs—community health worker programs operating globally [83,84]—that offer ready infrastructure for deploying digital health technologies to support aging populations.

Clinical and Policy Implications

Our findings suggest that digital health equity may require more than providing devices or training programs. Successful implementation depends on integrating technology within existing care relationships and workflows [23,70]. For clinicians, this means recognizing that digital health tools may be most effective when delivered through care teams rather than directly to patients. For policymakers, investing in community care infrastructure may paradoxically be essential for realizing the benefits of digital health investments.

The cost-effectiveness and scalability of this approach warrant future investigation. Implementation costs include (1) 1-time expenses for smartwatches (approximately US $200 per participant), smartphone app development, and server infrastructure; (2) recurring costs for device replacement, data storage, and technical support; and (3) caregiver time and wages for platform monitoring and responsive interventions. Notably, this study confirmed that these costs may be partially offset by integration with existing community care workflows—caregivers already conduct scheduled contacts, and platform monitoring required approximately 10-15 additional minutes daily per 10 assigned recipients for reviewing dashboard alerts. The critical economic question is whether early detection prevents costly downstream events. In our 6-week trial, the 37 triggered interventions included 1 case of infected abscess identified before sepsis development, suggesting substantial potential for preventing emergency department visits and hospitalizations that would far exceed platform costs. Formal cost-effectiveness analysis comparing platform-enhanced care versus usual care, incorporating quality-adjusted life years and health care use outcomes, is essential before recommending widespread implementation [85].

The below-threshold system usability scores (mean SUS score 46, SD 21.5 versus the cutoff score of 68 for acceptable usability) warrant consideration. These low scores likely reflect our target population's limited baseline digital literacy rather than platform-specific usability failures, as participants reported difficulty identifying any digital technology for reference during baseline assessment and no available internet connections at home. Nevertheless, suboptimal usability may have attenuated intervention effects by limiting engagement depth or creating frustration that offset potential benefits. In contrast, caregivers demonstrated acceptable usability scores and high engagement, with 69.5% (16/23) using the platform 2-3 times daily. Future iterations should prioritize simplified interfaces, voice-first interaction modalities, and extended onboarding support for older adults. Additionally, the low SUS scores suggest that the intermediary model—where caregivers translate platform outputs rather than expecting direct user proficiency—may be essential rather than merely convenient for this population.

Limitations

Several limitations warrant consideration. First, the quasi-experimental, nonrandomized design with unequal intervention exposure introduces fundamental confounding that limits causal inference. The intervention group received platform-triggered caregiver interventions, while the control group received standard scheduled care. This design makes it impossible to isolate the effect of the digital platform from the effect of possibly increased human contact, limiting conclusions to the feasibility and preliminary effectiveness of the integrated care model rather than independent platform efficacy. Randomization was not feasible in part because the multiphase research program spanned the COVID-19 pandemic and postpandemic periods (2021-2023), during which protecting vulnerable participants' safety and informed decision-making were prioritized. Voluntary participation through prior consultation with trusted caregivers was a prerequisite for enrollment, consistent with the participatory development process. This necessitated a matched comparison design based on caregiver and participant willingness rather than random assignment. Although the current living lab approach provides ecological validity for understanding real-world implementation within existing care systems, it was not feasible to match care services for the structured control group or require the provision of detailed reports of caregiving service types and the amount of time for each participant in both the intervention and control groups.

XSL•FO
RenderX

The nonrandomized design may have introduced additional potential selection bias, though groups were well-matched on measured characteristics. Particularly, intervention group caregivers volunteered to adopt the platform, potentially selecting for caregivers with greater technological comfort, higher motivation, or stronger relationships with their care recipients—factors that could independently affect health outcomes. The open-label nature was also unavoidable given the intervention's visibility to both participants and caregivers. Notably, both groups were drawn from older adults with established caregiver relationships who consented to face-to-face research assessments during the postpandemic period—a decision that itself reflects a degree of trust and social engagement uncommon in this population. The comparison group thus represents a relatively engaged subgroup rather than a passive usual-care population, which may partially mitigate concerns about selection bias favoring the intervention group.

Additionally, our multicomponent intervention does not allow experimental isolation of active mechanisms, and findings should not be interpreted as identifying which specific components drove the observed differential patterns. The platform integrated (1) continuous physiological monitoring via smartwatches, (2) daily chatbot surveys for symptom reporting, (3) personalized traffic-light feedback to users, (4) real-time dashboard alerts to caregivers, and (5) protocol-driven caregiver responses to health decline signals. Any or all components may have contributed to observed effects, individually or synergistically. For example, improvements could result from increased self-monitoring awareness, the supportive effect of daily chatbot interaction, enhanced caregiver responsiveness, or the earlier timing of follow-up enabled by digital alerts, or from complex interactions among these elements. Future research should use factorial designs or sequential multiple assignment randomized trials to systematically isolate component contributions and identify optimal intervention configurations for different patient subgroups.

While the multicomponent design precludes experimental isolation of active mechanisms, several observations generate testable hypotheses for future studies designed to isolate component effects. The 37 caregiver-initiated interventions were triggered by digital alerts rather than scheduled contacts, suggesting that the platform's primary mechanistic contribution may lie in the timing of care delivery—enabling earlier detection and response to health deterioration—rather than in the volume of caregiver contact. Caregivers maintained their standard 25-hour weekly schedule, further suggesting that differential outcomes may reflect improved allocation of existing care resources rather than increased service intensity. Additionally, the contrast between low older adult usability scores (mean SUS score 46, SD 21.5) and acceptable caregiver usability scores (mean SUS score 70.8, SD 16.7) suggests that intervention effects likely operated primarily through the caregiver-mediated pathway rather than through direct older adult engagement with the platform. These observations point toward caregiver-mediated early detection as a candidate mechanism, but this hypothesis requires formal testing through factorial designs that systematically vary platform components.

In addition, the 6-week intervention period represents a significant constraint. While adequate for demonstrating feasibility and detecting changes in functional limitations, this duration is insufficient for assessing depression outcomes, which meta-analytic evidence suggests require longer intervention periods [86]. Furthermore, observed effects may partly reflect novelty or Hawthorne effects that could diminish over time as participants habituate to monitoring. Long-term studies are needed to determine whether functional benefits persist, whether initially null effects on depression emerge with extended exposure, and whether engagement levels remain sustainable beyond the initial phase.

Also, missing data rates for some sensors (particularly sleep monitoring, 693/1318, 52.6%) may have limited our ability to detect associations, and nonrandom missingness cannot be excluded. Our sample of urban-dwelling Korean older adults limits generalizability to rural populations or other cultural contexts. The marginal significance of several findings after multiple testing corrections suggests the need for larger studies to confirm these associations.

Finally, our daily depression assessment—2 randomly selected binary PHQ-9 items—has limited sensitivity to mood fluctuations compared to full validated scales or ecological momentary assessment approaches with continuous response scales. This measurement constraint may have reduced our ability to detect subtle associations between digital biomarkers and mood states, potentially contributing to the null findings for depression outcomes. Future studies should use more sensitive daily mood measures, such as VASs or brief validated instruments designed for repeated administration.

Conclusions

This feasibility study provides preliminary evidence suggesting that integrating digital monitoring within existing community infrastructure may advance digital health support to low-income older adult populations. By positioning formal caregivers as technology intermediaries, vulnerable older adults were able to participate in continuous multimodal health monitoring while addressing digital literacy barriers. This approach demonstrated feasibility and raised the question of whether digital health equity may be better addressed by reimagining how, not just what, technologies are delivered within the existing community care infrastructure. Future research should use randomized controlled designs with appropriate control conditions to establish causal relationships and isolate the specific contributions of digital monitoring, enhanced feedback, and caregiver responsiveness to the observed differential patterns. Such studies will be essential for determining optimal intervention and implementation strategies across diverse settings and evaluating the long-term sustainability of this integrated care model.

XSL•FO
RenderX

Acknowledgments

We are very grateful to older adults and care service providers for their participation and the support of the staff at the Seong-buk City Senior Welfare Center.

Funding

This study was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Republic of Korea (grant HI21C0572), the Basic Research Lab Program of the National Research Foundation of Korea funded by the Ministry of Science and ICT (RS-2025-02303227), and the Basic Science Research Program through the National Research Foundation of Korea, funded by the Ministry of Science and ICT (grant RS-2024-00348012).

Conflicts of Interest

SS and JK developed the health monitoring platform software in collaboration with M2S+ Co and hold intellectual property rights for the platform. The Ministry of Health and Welfare of the Republic of Korea, which provided funding for platform development, retains rights to use the developed platform as a public service for older adults. HYK and SYH declare no competing interests.

Multimedia Appendix 1
Higher resolution version of Figure 2.
[PNG File , 415 KB - aging_v9i1e85611_app1.png]

Multimedia Appendix 2
Within- and between-person correlations of daily health measures; follow-up records after detected deterioration; and examples of caregiver app health decline reports.
[DOCX File , 382 KB - aging_v9i1e85611_app2.docx]

References

1. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8] [Medline: 37464029]
2. Xie B, Charness N, Fingerman K, Kaye J, Kim MT, Khurshid A. When going digital becomes a necessity: Ensuring older adults' needs for information, services, and social inclusion during COVID-19. J Aging Soc Policy 2020;32(4-5):460-470 [FREE Full text] [doi: 10.1080/08959420.2020.1771237] [Medline: 32507061]
3. Batra S, Baker RA, Wang T, Forma F, DiBiasi F, Peters-Strickland T. Digital health technology for use in patients with serious mental illness: a systematic review of the literature. Med Devices (Auckl) 2017;10:237-251 [FREE Full text] [doi: 10.2147/MDER.S144158] [Medline: 29042823]
4. Fortuna KL, Torous J, Depp CA, Jimenez DE, Areán PA, Walker R, et al. A future research agenda for digital geriatric mental healthcare. Am J Geriatr Psychiatry 2019;27(11):1277-1285 [FREE Full text] [doi: 10.1016/j.jagp.2019.05.013] [Medline: 31196619]
5. Collins JT, Walsh DA, Gladman JRF, Patrascu M, Husebo BS, Adam E, et al. The difficulties of managing pain in people living with frailty: the potential for digital phenotyping. Drugs Aging 2024;41(3):199-208 [FREE Full text] [doi: 10.1007/s40266-024-01101-4] [Medline: 38401025]
6. Aging and health. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2023-12-12]
7. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
8. Md Fadzil NH, Shahar S, Singh DKA, Rajikan R, Vanoh D, Mohamad Ali N, et al. Digital technology usage among older adults with cognitive frailty: a survey during COVID-19 pandemic. Digit Health 2023;9:20552076231207594 [FREE Full text] [doi: 10.1177/20552076231207594] [Medline: 37868158]
9. Seifert A, Cotten SR, Xie B. A double burden of exclusion? Digital and social exclusion of older adults in times of COVID-19. J Gerontol B Psychol Sci Soc Sci 2021;76(3):e99-e103 [FREE Full text] [doi: 10.1093/geronb/gbaa098] [Medline: 32672332]
10. Kim GE, Jo MW, Shin YW. Increased prevalence of depression in South Korea from 2002 to 2013. Sci Rep 2020 Oct 12;10(1):16979. [doi: 10.1038/s41598-020-74119-4] [Medline: 33046758]
11. Choi JW, Kim TH, Shin J, Han E. Poverty and suicide risk in older adults: a retrospective longitudinal cohort study. Int J Geriatr Psychiatry 2019;34(11):1565-1571. [doi: 10.1002/gps.5166] [Medline: 31276241]
12. Sano A, Taylor S, McHill AW, Phillips AJ, Barger LK, Klerman E, et al. Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. J Med Internet Res 2018;20(6):e210 [FREE Full text] [doi: 10.2196/jmir.9410] [Medline: 29884610]

XSL•FO
RenderX

13. Masud MT, Mamun MA, Thapa K, Lee DH, Griffiths MD, Yang SH. Unobtrusive monitoring of behavior and movement patterns to detect clinical depression severity level via smartphone. J Biomed Inform 2020;103:103371 [FREE Full text] [doi: 10.1016/j.jbi.2019.103371] [Medline: 31935462]
14. Rozgonjuk D, Levine JC, Hall BJ, Elhai JD. The association between problematic smartphone use, depression and anxiety symptom severity, and objectively measured smartphone use over one week. Comput Hum Behav 2018;87:10-17. [doi: 10.1016/j.chb.2018.05.019]
15. BinDhim NF, Shaman AM, Trevena L, Basyouni MH, Pont LG, Alhawassi TM. Depression screening via a smartphone app: cross-country user characteristics and feasibility. J Am Med Inform Assoc 2015;22(1):29-34 [FREE Full text] [doi: 10.1136/amiajnl-2014-002840] [Medline: 25326599]
16. Onnela JP, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology 2016 Jun;41(7):1691-1696. [doi: 10.1038/npp.2016.7] [Medline: 26818126]
17. Difrancesco S, Lamers F, Riese H, Merikangas KR, Beekman ATF, van Hemert AM, et al. Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: a 2-week ambulatory assessment study. Depress Anxiety 2019;36(10):975-986 [FREE Full text] [doi: 10.1002/da.22949] [Medline: 31348850]
18. Torous J, Staples P, Onnela JP. Realizing the potential of mobile mental health: new methods for new data in psychiatry. Curr Psychiatry Rep 2015;17(8):602 [FREE Full text] [doi: 10.1007/s11920-015-0602-0] [Medline: 26073363]
19. Richards D, Timulak L, Doherty G, Sharry J, Colla A, Joyce C, et al. Internet-delivered treatment: its potential as a low-intensity community intervention for adults with symptoms of depression: protocol for a randomized controlled trial. BMC Psychiatry 2014;14:147 [FREE Full text] [doi: 10.1186/1471-244X-14-147] [Medline: 24886179]
20. Fang ML, Walker M, Wong KLY, Sixsmith J, Remund L, Sixsmith A. Future of digital health and community care: exploring intended positive impacts and unintended negative consequences of COVID-19. Healthc Manage Forum 2022;35(5):279-285 [FREE Full text] [doi: 10.1177/08404704221107362] [Medline: 35775162]
21. Willems SH, Rao J, Bhambere S, Patel D, Biggins Y, Guite JW. Digital solutions to alleviate the burden on health systems during a public health care crisis: COVID-19 as an opportunity. JMIR Mhealth Uhealth 2021;9(6):e250 [FREE Full text] [doi: 10.2196/28264]
22. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
23. Peng R, Guo Y, Zhang C, Li X, Huang J, Chen X, et al. Internet-delivered psychological interventions for older adults with depression: a scoping review. Geriatr Nurs 2024;55:97-104. [doi: 10.1016/j.gerinurse.2023.10.026] [Medline: 37976561]
24. Hunsaker A, Hargittai E. A review of internet use among older adults. New Media Soc 2018;20(10):3937-3954. [doi: 10.1177/1461444818787348]
25. Kunonga TP, Spiers GF, Beyer FR, Hanratty B, Boulton E, Hall A, et al. Effects of digital technologies on older people's access to health and social care: umbrella review. J Med Internet Res 2021;23(11):e25887 [FREE Full text] [doi: 10.2196/25887] [Medline: 34821564]
26. Dubbeldam R, Stemplewski R, Pavlova I, Cyma-Wejchenig M, Lee S, Esser P, et al. Technology-assisted physical activity interventions for older people in their home-based environment: scoping review. JMIR Aging 2025;8:e65746 [FREE Full text] [doi: 10.2196/65746] [Medline: 40952779]
27. Mitzner TL, Savla J, Boot WR, Sharit J, Charness N, Czaja SJ, et al. Technology adoption by older adults: findings from the PRISM trial. Gerontologist 2019;59(1):34-44 [FREE Full text] [doi: 10.1093/geront/gny113] [Medline: 30265294]
28. Sawatzky R, Schick-Makaroff K, Ratner PA, Kwon J, Whitehurst DGT, Öhlén J, et al. Did a digital quality of life (QOL) assessment and practice support system in home health care improve the QOL of older adults living with life-limiting conditions and of their family caregivers? A mixed-methods pragmatic randomized controlled trial. PLoS One 2025;20(5):e0320306 [FREE Full text] [doi: 10.1371/journal.pone.0320306] [Medline: 40327663]
29. Korman M, Weiss PL, Kizony R. Living labs: overview of ecological approaches for health promotion and rehabilitation. Disabil Rehabil 2016;38(7):613-619. [doi: 10.3109/09638288.2015.1059494] [Medline: 26138020]
30. Verloo H, Lorette A, Rosselet Amoussou J, Gillès de Pélichy E, Matos Queirós A, von Gunten A, et al. Using living labs to explore needs and solutions for older adults with dementia: scoping review. JMIR Aging 2021;4(3):e29031 [FREE Full text] [doi: 10.2196/29031] [Medline: 34420916]
31. Callari T, Moody L, Saunders J, Ward G, Holliday N, Woodley J. Exploring participation needs and motivational requirements when engaging older adults in an emerging living lab. Technol Innov Manag Rev 2019;9(3):38-49. [doi: 10.22215/timreview/1223]
32. What are living labs? European Network of Living Labs. 2023. URL: https://enoll.org/about-us/ [accessed 2023-12-12]
33. Song S, Seo Y, Hwang SY, Kim HY, Kim J. Digital phenotyping of geriatric depression using a community-based digital mental health monitoring platform for socially vulnerable older adults and their community caregivers: 6-week living lab single-arm pilot study. JMIR Mhealth Uhealth 2024;12:e55842 [FREE Full text] [doi: 10.2196/55842] [Medline: 38885033]
34. Jo S, Song S, Kim J, Song C. Agreement between Azure Kinect and marker-based motion analysis during functional movements: a feasibility study. Sensors (Basel) 2022;22(24):9819 [FREE Full text] [doi: 10.3390/s22249819] [Medline: 36560187]

XSL•FO
RenderX

35. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth 2018;6(8):e10527 [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]
36. Lubitz SA, Faranesh AZ, Selvaggi C, Atlas SJ, McManus DD, Singer DE, et al. Detection of atrial fibrillation in a large population using wearable devices: the Fitbit heart study. Circulation 2022;146(19):1415-1424 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.122.060291] [Medline: 36148649]
37. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Digit Med 2020;3:18. [doi: 10.1038/s41746-020-0226-6] [Medline: 32047863]
38. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of wristband Fitbit models in assessing sleep: systematic review and meta-analysis. J Med Internet Res 2019;21(11):e16273 [FREE Full text] [doi: 10.2196/16273] [Medline: 31778122]
39. Fekedulegn D, Andrew ME, Shi M, Violanti JM, Knox S, Innes KE. Actigraphy-based assessment of sleep parameters. Ann Work Expo Health 2020;64(4):350-367 [FREE Full text] [doi: 10.1093/annweh/wxaa007] [Medline: 32053169]
40. Webster JB, Kripke DF, Messin S, Mullaney DJ, Wyborney G. An activity-based sleep monitor system for ambulatory use. Sleep 1982;5(4):389-399. [doi: 10.1093/sleep/5.4.389] [Medline: 7163726]
41. Aura Healthcare. AURA healthcare HRV-analysis program. GitHub. 2024. URL: https://github.com/Aura-healthcare/hrv-analysis [accessed 2026-03-19]
42. Pappano AJ, Wier WG. Cardiovascular Physiology. Philadelphia, PA: Elsevier Health Sciences; 2018.
43. Natarajan A, Pantelopoulos A, Emir-Farinas H, Natarajan P. Heart rate variability with photoplethysmography in 8 million individuals: a cross-sectional study. Lancet Digit Health 2020;2(12):e650-e657 [FREE Full text] [doi: 10.1016/S2589-7500(20)30246-6] [Medline: 33328029]
44. Redenius N, Kim Y, Byun W. Concurrent validity of the Fitbit for assessing sedentary behavior and moderate-to-vigorous physical activity. BMC Med Res Methodol 2019;19(1):29 [FREE Full text] [doi: 10.1186/s12874-019-0668-1] [Medline: 30732582]
45. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15(12):1833-1840. [Medline: 3068365]
46. Ko T, Kim S, Lee JS. Reliability and validity of the Korean Western Ontario and McMaster universities (WOMAC) osteoarthritis index in patients with osteoarthritis of the knee. J Korean Acad Rehabil Med 2009;19(2):251-260. [doi: 10.1007/978-3-211-79280-3_1165]
47. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a Geriatric Depression Screening Scale: a preliminary report. J Psychiatr Res 1982;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4] [Medline: 7183759]
48. Jung IK, Kwak DI, Shin DK, Lee MS, Lee HS, Kim JY. A reliability and validity study of geriatric depression scale. J Korean Neuropsychiatr Assoc 1997;36(1):103-112.
49. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
50. Sohn SI, Kim DH, Lee MY, Cho YW. The reliability and validity of the Korean version of the Pittsburgh Sleep Quality Index. Sleep Breath 2012;16(3):803-812. [doi: 10.1007/s11325-011-0579-9] [Medline: 21901299]
51. Holden RJ. A simplified system usability scale (SUS) for cognitively impaired and older adults. 2020 Presented at: Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care; 2020, May 18-21; Toronto, Ontario, Canada p. 180-182. [doi: 10.1177/2327857920091021]
52. Zenebe Y, Akele B, W/Selassie M, Necho M. Prevalence and determinants of depression among old age: a systematic review and meta-analysis. Ann Gen Psychiatry 2021 Dec 18;20(1):55 [FREE Full text] [doi: 10.1186/s12991-021-00375-x] [Medline: 34922595]
53. Kim H, Lee S, Lee S, Hong S, Kang H, Kim N. Depression prediction by using ecological momentary assessment, Actiwatch data, and machine learning: observational study on older adults living alone. JMIR Mhealth Uhealth 2019;7(10):e14149 [FREE Full text] [doi: 10.2196/14149] [Medline: 31621642]
54. Raudenbush SW, Bryk AS. Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks, CA: Sage Publications; 2002.
55. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.
56. Hecht M, Zitzmann S. Sample size recommendations for continuous-time models: compensating shorter time series with larger numbers of persons and vice versa. Struct Equ Modeling 2020;28(2):229-236. [doi: 10.1080/10705511.2020.1779069]
57. Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, et al. Digital interventions for the treatment of depression: a meta-analytic review. Psychol Bull 2021;147(8):749-786. [doi: 10.1037/bul0000334] [Medline: 34898233]
58. Martens MJ, Logan BR. A unified approach to sample size and power determination for testing parameters in generalized linear and time-to-event regression models. Stat Med 2021;40(5):1121-1132 [FREE Full text] [doi: 10.1002/sim.8823] [Medline: 33210321]

XSL•FO
RenderX

59. Daguet I, Raverot V, Bouhassira D, Gronfier C. Circadian rhythmicity of pain sensitivity in humans. Brain 2022;145(9):3225-3235. [doi: 10.1093/brain/awac147] [Medline: 36103407]
60. Mun CJ, Burgess HJ, Sears DD, Parthasarathy S, James D, Altamirano U, et al. Circadian rhythm and pain: a review of current research and future implications. Curr Sleep Med Rep 2022;8(4):114-123. [doi: 10.1007/s40675-022-00228-3]
61. Tracy LM, Ioannou L, Baker KS, Gibson SJ, Georgiou-Karistianis N, Giummarra MJ. Meta-analytic evidence for decreased heart rate variability in chronic pain implicating parasympathetic nervous system dysregulation. Pain 2016;157(1):7-29. [doi: 10.1097/j.pain.0000000000000360] [Medline: 26431423]
62. Rampazo É, Rehder-Santos P, Catai AM, Liebano RE. Heart rate variability in adults with chronic musculoskeletal pain: a systematic review. Pain Pract 2024;24(1):211-230. [doi: 10.1111/papr.13294] [Medline: 37661339]
63. Becker NB, Jesus SN, João KADR, Viseu JN, Martins RIS. Depression and sleep quality in older adults: a meta-analysis. Psychol Health Med 2017;22(8):889-895. [doi: 10.1080/13548506.2016.1274042] [Medline: 28013552]
64. Nielson SA, Kay DB, Dzierzewski JM. Sleep and depression in older adults: a narrative review. Curr Psychiatry Rep 2023;25(11):643-658. [doi: 10.1007/s11920-023-01455-3] [Medline: 37740851]
65. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
66. Gulliford M, Alageel S. Digital health intervention at older ages. Lancet Digit Health 2019;1(8):e382-e383 [FREE Full text] [doi: 10.1016/S2589-7500(19)30194-3] [Medline: 33323216]
67. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
68. Riadi I, Kervin L, Dhillon S, Teo K, Churchill R, Card KG, et al. Digital interventions for depression and anxiety in older adults: a systematic review of randomised controlled trials. Lancet Healthy Longev 2022;3(8):e558-e571 [FREE Full text] [doi: 10.1016/S2666-7568(22)00121-0] [Medline: 36102765]
69. Peng R, Li X, Guo Y, Ning H, Huang J, Jiang D, et al. Barriers and facilitators to acceptance and implementation of eMental-health intervention among older adults: a qualitative systematic review. Digit Health 2024;10:20552076241234628 [FREE Full text] [doi: 10.1177/20552076241234628] [Medline: 38444518]
70. Naslund JA, Shidhaye R, Patel V. Digital technology for building capacity of nonspecialist health workers for task sharing and scaling up mental health care globally. Harv Rev Psychiatry 2019;27(3):181-192 [FREE Full text] [doi: 10.1097/HRP.0000000000000217] [Medline: 30958400]
71. Murayama H, Taguchi A, Spencer MS, Yamaguchi T. Efficacy of a community health worker-based intervention in improving dietary habits among community-dwelling older people: a controlled, crossover trial in Japan. Health Educ Behav 2020;47(1):47-56. [doi: 10.1177/1090198119891975] [Medline: 31933395]
72. Kwok JYY, Jiang D, Yeung DYL, Choi NG, Ho RTH, Warner LM, et al. Layperson-delivered telephone-based behavioral activation among low-income older adults during the COVID-19 pandemic: the HEAL-HOA randomized clinical trial. JAMA Netw Open 2024;7(6):e2416767 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.16767] [Medline: 38888923]
73. Scazufca M, Nakamura CA, Seward N, Moreno-Agostino D, van de Ven P, Hollingworth W, et al. A task-shared, collaborative care psychosocial intervention for improving depressive symptomatology among older adults in a socioeconomically deprived area of Brazil (PROACTIVE): a pragmatic, two-arm, parallel-group, cluster-randomised controlled trial. Lancet Healthy Longev 2022;3(10):e690-e702 [FREE Full text] [doi: 10.1016/S2666-7568(22)00194-5] [Medline: 36202129]
74. Gianfredi V, Blandi L, Cacitti S, Minelli M, Signorelli C, Amerio A, et al. Depression and objectively measured physical activity: a systematic review and meta-analysis. Int J Environ Res Public Health 2020;17(10):3738 [FREE Full text] [doi: 10.3390/ijerph17103738] [Medline: 32466242]
75. Kamath J, Leon Barriera R, Jain N, Keisari E, Wang B. Digital phenotyping in depression diagnostics: integrating psychiatric and engineering perspectives. World J Psychiatry 2022;12(3):393-409 [FREE Full text] [doi: 10.5498/wjp.v12.i3.393] [Medline: 35433319]
76. De Angel V, Lewis S, White K, Oetzmann C, Leightley D, Oprea E, et al. Digital health tools for the passive monitoring of depression: a systematic review of methods. NPJ Digit Med 2022;5(1):3 [FREE Full text] [doi: 10.1038/s41746-021-00548-8] [Medline: 35017634]
77. Mercier HW, Hamner JW, Torous J, Onnela JP, Taylor JA. Digital phenotyping to quantify psychosocial well-being trajectories after spinal cord injury. Am J Phys Med Rehabil 2020;99(12):1138-1144 [FREE Full text] [doi: 10.1097/PHM.0000000000001506] [Medline: 32576743]
78. Jacobson NC, Weingarden H, Wilhelm S. Digital biomarkers of mood disorders and symptom change. NPJ Digit Med 2019;2:3 [FREE Full text] [doi: 10.1038/s41746-019-0078-0] [Medline: 31304353]
79. Price GD, Heinz MV, Song SH, Nemesure MD, Jacobson NC. Using digital phenotyping to capture depression symptom variability: detecting naturalistic variability in depression symptoms across one year using passively collected wearable movement and sleep data. Transl Psychiatry 2023;13(1):381 [FREE Full text] [doi: 10.1038/s41398-023-02669-y] [Medline: 38071317]

XSL•FO
RenderX

80. Hackett K, Xu S, McKniff M, Paglia L, Barnett I, Giovannetti T. Mobility-based smartphone digital phenotypes for unobtrusively capturing everyday cognition, mood, and community life-space in older adults: feasibility, acceptability, and preliminary validity study. JMIR Hum Factors 2024;11:e59974 [FREE Full text] [doi: 10.2196/59974] [Medline: 39576984]
81. Izquierdo M, Duque G, Morley JE. Physical activity guidelines for older people: knowledge gaps and future directions. Lancet Healthy Longev 2021;2(6):e380-e383 [FREE Full text] [doi: 10.1016/S2666-7568(21)00079-9] [Medline: 36098146]
82. Izquierdo M, Merchant RA, Morley JE, Anker SD, Aprahamian I, Arai H, et al. International exercise recommendations in older adults (ICFSR): expert consensus guidelines. J Nutr Health Aging 2021;25(7):824-853 [FREE Full text] [doi: 10.1007/s12603-021-1665-8] [Medline: 34409961]
83. LMBDE M, Santos RCD, Albuquerque PC. Community health workers: what do international studies tell us? Cien Saude Colet 2023;28(2):501-520 [FREE Full text] [doi: 10.1371/journal.pone.0280410] [Medline: 36800366]
84. Perry HB, Zulliger R, Rogers MM. Community health workers in low-, middle-, and high-income countries: an overview of their history, recent evolution, and current effectiveness. Annu Rev Public Health 2014;35:399-421. [doi: 10.1146/annurev-publhealth-032013-182354] [Medline: 24387091]
85. Gentili A, Failla G, Melnyk A, Puleo V, Tanna GLD, Ricciardi W, et al. The cost-effectiveness of digital health interventions: a systematic review of the literature. Front Public Health 2022;10:787135 [FREE Full text] [doi: 10.3389/fpubh.2022.787135] [Medline: 36033812]
86. Tong AC, Ho FS, Chu OH, Mak WW. Time-dependent changes in depressive symptoms among control participants in digital-based psychological intervention studies: meta-analysis of randomized controlled trials. J Med Internet Res 2023;25:e39029 [FREE Full text] [doi: 10.2196/39029] [Medline: 37043276]

Abbreviations

GDS-SF: Geriatric Depression Scale Short-Form
H1: hypothesis 1
H2: hypothesis 2
H3: hypothesis 3
HRV: heart rate variability
OR: odds ratio
PA: physical activity
PHQ-9: Patient Health Questionnaire-9
PSQI: Pittsburgh Sleep Quality Index
SUS: System Usability Scale
VAS: visual analog scale
WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index

Edited by J Wang; submitted 10.Oct.2025; peer-reviewed by K Cho; comments to author 14.Nov.2025; revised version received 21.Feb.2026; accepted 17.Mar.2026; published 20.Apr.2026.

Please cite as:
Song S, Hwang SY, Kim HY, Kim J
Digital-Human Public Community Care Integration for Chronic Pain in Low-Income Older Adults in a 6-Week Living Lab Setting: Quasi-Experimental Feasibility Study
JMIR Aging 2026;9:e85611
URL: https://aging.jmir.org/2026/1/e85611
doi:10.2196/85611

©Sunmi Song, Seo-Yeon Hwang, Hae-Young Kim, Junesun Kim. Originally published in JMIR Aging (https://aging.jmir.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Locomotive Syndrome Digital Therapeutics Provided via a Smartphone App: Proof-of-Concept Single-Group Trial Study

Tatsuru Sonobe[1], MD, PhD; Itaru Ogawa[1], MD; Takahiro Seki[1], MD; Kosuke Watanabe[1], MD; Yota Kaneko[1], MD, PhD; Takeru Yokota[1], MD, PhD; Taro Mawatari[2], MD, PhD; Satoru Harada[2], MD, PhD; Yasumichi Kadowaki[3]; Youhei Takenaka[3]; Yoshihiro Matsumoto[1], MD, PhD

[1]Department of Orthopaedic Surgery, School of Medicine, Fukushima Medical University, 1 Hikariga-oka, Fukushima, Japan

[2]Department of Orthopaedic Surgery, KKR Hamanomachi Hospital, Fukuoka, Japan

[3]Future Corporation, Tokyo, Japan

Corresponding Author:
Tatsuru Sonobe, MD, PhD
Department of Orthopaedic Surgery, School of Medicine, Fukushima Medical University, 1 Hikariga-oka, Fukushima, Japan

Abstract

Background: Individuals with locomotive syndrome (LS) have muscle weakness and reduced motor function due to musculoskeletal disorders that cause reduced mobility and physical function. In Japan, musculoskeletal disorders are the most common reason for requiring home support or nursing care, highlighting the need for preventing and ameliorating LS. Middle-aged and older adults sometimes encounter difficulty making a habit of exercise therapy (the mainstay of LS treatment).

Objective: We investigated whether digital therapeutics (DTx) can prevent or ameliorate LS in middle-aged and older adults.

Methods: We conducted a prospective, longitudinal, nonrandomized, single-group study of Japanese aged 40 years and older who were eligible for LS checkups (N=47). Each participant underwent an 8-week locomotion training intervention based on DTx supervised by medical staff. We objectively assessed the participants' subjective and objective motor abilities and motor awareness with the Timed Up and Go (TUG) test, 25-question Geriatric Locomotive Function Scale (GLFS-25), and Behavioral Regulation in Exercise Questionnaire-3 at baseline (before the DTx), an interim point (4 wk after the DTx initiation), and a final evaluation (8 wk post-DTx initiation). We compared the scores of the 3 tests at the 3 time points as dependent variables in a 3-factor ANOVA with Bonferroni correction (significance defined as 0.05/3=0.0167).

Results: No increase in amotivation to exercise or refusal to exercise was observed. Significant improvements at 8 weeks versus the baseline were observed in the TUG scores (baseline: 9.0, 95% CI 8.4 - 9.6; 8 wk: 7.5, 95% CI 7.1 - 8.0; *P*=.001) and GLFS-25 results (baseline: 18.7, 95% CI 14.5 - 22.8; 8 wk: 11.7, 95% CI 8.8 - 14.7; *P*=.004). The Behavioral Regulation in Exercise Questionnaire-3 and its subscale data did not differ significantly at any assessment time point.

Conclusions: These results indicate that an 8-week locomotive training intervention using DTx significantly improved middle-aged and older adults' TUG and GLFS-25 scores and will help prevent and ameliorate LS and establish better exercise habits among them.

Trial Registration: University Hospital Medical Information Network Clinical Trial Registry UMIN000053922; https://center6.umin.ac.jp/cgi-open-bin/ctr/ctr_view.cgi?recptno=R000061550

(JMIR Aging 2026;9:e86174) doi:10.2196/86174

KEYWORDS

locomotive syndrome; LS; digital therapeutic; DTx; locomotion training; LT; Timed Up and Go test; TUG test; 25-question Geriatric Locomotive Function Scale; 25-GLFS; Behavioral Regulation in Exercise Questionnaire-3; BREQ-3

Introduction

The Japanese Orthopaedic Association (JOA) proposed the concept of locomotive syndrome (LS) in 2007 [1]. The term "LS" refers to a condition in which a musculoskeletal disorder results in muscle weakness and motor weakness, leading to reduced mobility and physical function [1]. The development of LS causes a requirement for long-term care [2]. According to the 2022 Comprehensive Survey of Living Conditions in Japan, the leading cause requiring nursing care or support was musculoskeletal disorders at 26.3%, followed by falls and fractures at 13.9% and joint disorders at 10.2% [3]. The JOA recommends the early detection of a decline in motor function based on locomotive syndrome checks (LSCs), in part to prevent the requirement of assistance or care [4]. LSCs were scored as follows: 1 point (cannot put on a sock for 1 foot), 2 points (stumbles in their home), 3 points (needs a handrail to climb

XSL•FO
RenderX

stairs), 4 points (difficulty using a vacuum cleaner at home), 5 points (difficulty carrying parcels weighing >2 kg), 6 points (difficulty walking continuously for >15 min), and 7 points (cannot cross a crosswalk with a green light) [4]. The JOA also advocates locomotion training (LT), which involves standing on 1 leg and squatting, to prevent LS and lessen existing LS [5]. Currently, the risk of developing LS is assessed based on three items: (1) a stand-up test, (2) a 2-step test, and (3) the 25-question Geriatric Locomotive Function Scale (GLFS-25) [5-7].

Exercises such as LT are effective for preventing and mitigating LS [5], but it is difficult for many middle-aged and older adults to start or continue an exercise routine. Previous studies have reported that the average adherence rate for older adults continuing a prescribed home exercise program for 6 weeks was 65% [8]. As one solution, we advocate an approach to the prevention and mitigation of LS in middle-aged and older adults who meet the criteria for LS or qualify for LSCs using IT devices.

Apart from the uses of IT in the medical field, such as electronic medical records, telemedicine, and surgical robots, the uses of mobile devices for medical procedures and support are called mobile health [9]. Digital therapeutics (DTx), defined by the Digital Therapeutics Alliance as “evidence-based therapeutic interventions driven by high-quality software programs to prevent, manage, or treat a medical disorder or disease” [10], is an area of mobile health that is receiving particular attention [11]. A variety of DTx apps have been approved by health insurance programs in some countries, including those for diabetes, hypertension self-management, and smoking cessation support [12-14]. Regarding the effects of DTx on the musculoskeletal system, the effectiveness of DTx in musculoskeletal diseases has been reported for patients with pain associated with musculoskeletal diseases and osteoarthritis [15,16]. DTx has been shown to be effective in reducing musculoskeletal pain, improving function in patients with osteoarthritis, and increasing self-management abilities. However, there have been no reports on the use of DTx intervention for LS. If DTx intervention can establish exercise habits and lead to the prevention and improvement of LS, it may lead to a reduction in the number of people in Japan who require nursing care or support due to musculoskeletal diseases. We conducted this study to determine whether this use of DTx via a smartphone app can lead to behavioral change in middle-aged and older people and prevent or ameliorate their LS.

Methods

Participants

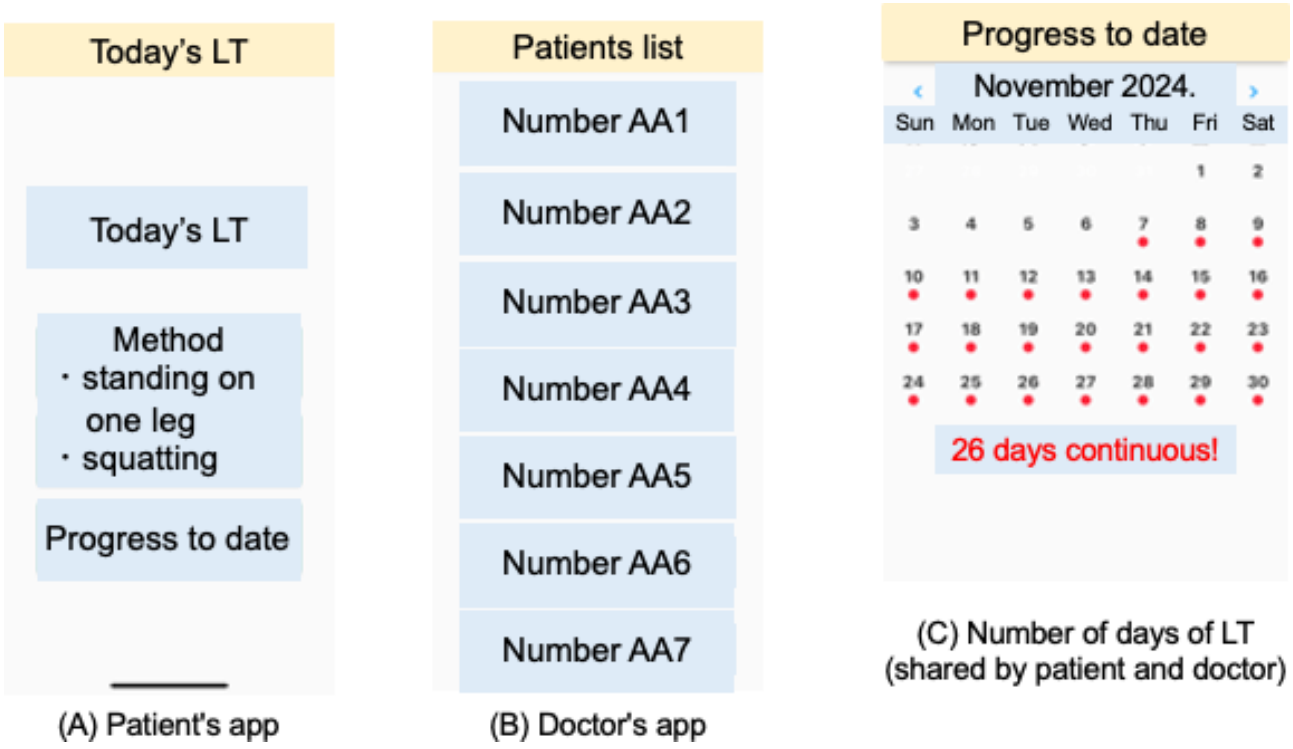
The details of the research protocol and participant recruitment have been published [17]. The study population was 47 Japanese participants aged more than 40 years who were eligible for the LSCs. The study’s inclusion criteria were as follows: individuals aged 40 years and older who generally begin to experience a decline in lower-extremity muscle strength [18] and met any of the LSC criteria. Additional inclusion criteria were: a score of 3 points or more on the Mini-Cog. The Mini-Cog is a screening test for cognitive function assessment combining 3-word immediate recall, delayed recall, and clock drawing. A score of less than 3 points indicates suspected dementia, with reported sensitivity of 76% to 99% and specificity of 83% to 93% [19,20]. The ability to use smartphone apps and being able to provide informed consent for study participation were additional requirements. Exclusion criteria were difficulty using smartphone apps due to cognitive decline or visual or hearing impairment, withdrawal during the study period, and withdrawal during the study period due to hospitalization or accidents. The participants’ demographic information was recorded, including age, gender, musculoskeletal disorders considered potential causes of LS, and underlying conditions corresponding to LSCs.

DTx Intervention

We have created a smartphone app for middle-aged and older people who have no exercise habits and whose physical functions have declined or are suspected to be declining. The smartphone app used by the participants is a new prototype app created for this research field. We provided a smartphone equipped with a communication environment with the app already downloaded to each participant. This smartphone app has functions that support the implementation of 2 types of LT with the use of video and audio. The LT app is simple to use and is set up so that by touching the screen, a voice announcement is played, and the user can perform 2 LT programs: the 1-leg stand program and the squat program. After the completion of the day’s LT, the participant touched the screen to check the LT record on both the participant’s smartphone and his or her physician’s smartphone (Figure 1). The smartphone app automatically records the number of days of LT.

XSL•FO
RenderX

Figure 1. Screenshots of the app. (A) The app for the patients or participants has a simple design and enabled them to perform the locomotion training (LT). (B) The physician's app can check each patient's LT status based on the ID allocated to that patient. (C) The number of days of LT is reflected on the same screen for both the patient and the physician.



Participants received direct instruction on how to use the app and perform LT from their attending physician during an outpatient visit at the medical institution involved in this study. The prior instruction took an average of 10 minutes. Furthermore, after the instruction, participants were able to use the DTx without any problems. Although there were several inquiries regarding its use, participants contacted the research staff directly by phone, and all inquiries were resolved over the phone. Each participant was scheduled for the daily LT intervention for 8 weeks. A single LT session consisted of a program of standing on 1 leg for 60 seconds (or <60 s if that was the participant's best effort) plus 5 squats, performed 3 times per day, as recommended by the JOA for the prevention and improvement of LS [5]. For participants with weak legs and/or backs who would find it difficult to do the regular LT, the physician responsible for explaining the DTx advised that they stand on 1 leg while supporting their body with their hands or fingers on a desk or table, or that they sit on a chair and stand up by putting their hands on a desk or table instead of doing squats, or that they reduce the number of times they do each LT from 3 times to 1 time [21].

The advantage of this system is that healthcare providers can check the progress of their patients' LT remotely, enabling them to prepare feedback in advance and freeing up time. It also provides patients with the reassurance that they are exercising under the support of medical staff.

Study Protocol

Briefly, the participants' objective mobility was assessed using the Timed Up and Go (TUG) test score; their subjective mobility was evaluated with the GLFS-25, and the changes in exercise-related motivation were investigated based on the Behavioral Regulation in Exercise Questionnaire-3 (BREQ-3). The GLFS25 is a questionnaire that assesses physical pain and difficulties in daily life using 25 questions. Each question is rated on a scale of 0 to 4, resulting in a total score of 0 to 100 [5-7]. The self-determination theory (SDT) [22] is widely applied as a theoretical foundation for evaluating the effectiveness of behavioral change in participants. The BREQ-3 is designed based on SDT and assesses motivation for exercise with 18 questions in 6 areas: intrinsic motivation, integrative adjustment, identification adjustment, incorporative adjustment, external adjustment, and demotivation [23]. Previous studies have adopted the BREQ-3 as an indicator for evaluating behavioral change [24]. In this study, the impact of DTx on motivation for exercise and exercise spontaneity was also evaluated using the BREQ-3. The participants underwent the TUG test, GLFS-25, and BREQ-3 measurements at baseline (before the DTx intervention was started), and at 4 weeks (intermediate evaluation) and 8 weeks (final evaluation) after the start of the DTx intervention. In addition, the implementation status and compliance rate of the LT during the survey period were analyzed. This study primarily aimed to explore the efficacy, feasibility, and behavioral impact of a DTx intervention for LT. Accordingly, a single-group pre-post comparative design was adopted to efficiently evaluate within-subject changes in objective mobility, subjective mobility, and exercise-related motivation. Outcome measures and the intervention duration were set at a feasible timeframe capable of detecting clinical effects while minimizing participant burden.

Statistical Analyses

Continuous data are summarized as the mean and SD, and dichotomous or categorical data are presented as proportions. As the main study outcomes, the TUG test and GLFS-25 results

XSL•FO
RenderX

were measured at 0 weeks before the DTx intervention (baseline) and at 4 and 8 weeks after the daily DTx intervention was started. TUG and GLFS-25 data at each assessment point were compared using ANOVA. Furthermore, the Pearson correlation coefficients were calculated between the participants' baseline TUG scores and the difference in their TUG scores (Δ) after 8 weeks of DTx in each case. Similarly, the Pearson correlation coefficients were calculated for the difference between baseline and post-8 weeks DTx GLFS-25 scores. As a secondary outcome measure in this study, participants' BREQ-3 scores obtained at 3 time points were also compared using the Wilcoxon signed-rank test. The *P* values after Bonferroni correction that were <.05/3=0.0167 were considered significant. All analyses were conducted using JMP PRO 16 (SAS Institute).

The reported minimal clinically important difference (MCID) in the TUG test among patients with degenerative disc disease is 3.4 (SD 5.0) seconds [25]. Considering that finding, we set the MCID at 3.0 seconds for the present investigation. The required sample size was calculated using G*power (version 3.1.9.7; Heinrich-Heine-Universität Düsseldorf).[26]. The corresponding 2-tailed *t* test was selected, and the effect size was calculated as 0.6 based on a mean difference of 3.0 and a SD of difference of 5.0. Assuming an α-value of .05 and .80 as the power (1−β), the required estimated sample size was 24. When taking into account both the possibility that the SD is smaller than those in the previous studies and the cases of dropout, we concluded that the study had the required sample size.

Ethical Considerations

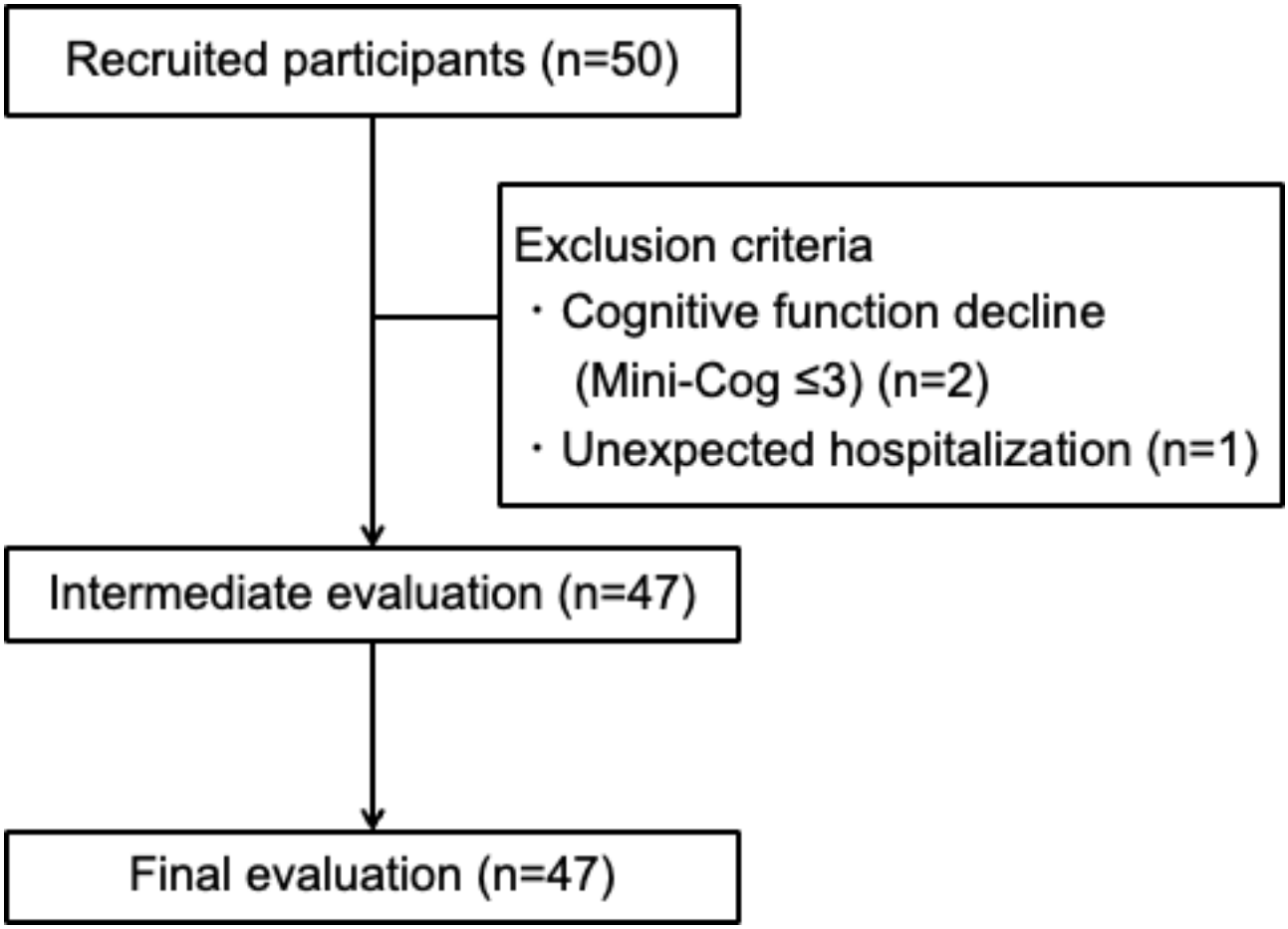
The Fukushima Medical University Ethics Review Committee provided ethical approval for this study (REC 2023 - 196). All study methods followed the relevant guidelines and regulations, and written informed consent was obtained from participants. This study was preregistered as a trial registration (University Hospital Medical Information Network Clinical Trial Registry UMIN000053922). Data was deidentified to protect participant information. Participants were not paid compensation and cooperated voluntarily.

Results

Recruitment Results Flowchart

Figure 2 illustrates the flow of participants over the course of the study. Two participants for whom cognitive function decline was observed and 1 participant who was unexpectedly hospitalized were excluded from the study. There were no dropouts after the start of the study, and all participants completed the final evaluation.

Figure 2. Flowchart of participants. Two individuals with cognitive impairment and one who was unexpectedly hospitalized were excluded from the study. After the start of the study, there were no dropouts, and 47 participants were followed up until the intermediate and final evaluations.



XSL•FO
RenderX

Participants' Characteristics

The characteristics of the 47 participants are summarized in Table 1. Most of the participants (55.3%, n=26) were aged under 65 years, and 87.2% of the participants were women. All of the participants had preserved cognitive function. The most common cause of LS in this population was knee osteoarthritis (51.1%).

Table . The participants' characteristics (N=47).

Characteristics	Value
Days of LT[a], mean (SD, 95% CI)	52.1 (5.9, 50.4 - 53.8)
Age (y), mean (SD, 95% CI)	63.5 (10.1, 60.5 - 66.5)
<65, n (%)	26 (55.3)
65 - 74, n (%)	13 (27.7)
≥75, n (%)	8 (17)
Sex, n (%)	
Male	6 (12.8)
Female	41 (87.2)
Mini-Cog score (points), n (%)	
4	11 (23.4)
5	36 (76.6)
Disease causing LS[b], n (%)	
Knee osteoarthritis	24 (51.1)
Hip osteoarthritis	5 (10.6)
Lumbar spondylosis	14 (29.7)
Hip fracture	2 (4.3)
Ankle osteoarthritis	2 (4.3)

[a]LT: locomotion training.

[b]LS: locomotive syndrome.

Primary Outcome Measures

TUG Test

The changes in the participants' TUG scores from baseline to the end of the 8-week intervention are depicted in Table 2. Notably, with the DTx, all but 1 of the 47 participants achieved an improvement in their TUG score. The TUG score at baseline was 9.0 (95% CI 8.4 - 9.6), reduced significantly to 8.0 (95% CI 7.6 - 8.5) at 4 weeks (*P*=.01). Furthermore, compared to baseline, the TUG score significantly decreased to 7.5 (95% CI 7.1 - 8.0) at week 8 (*P*=.001).

Table . The participants' Timed Up and Go (TUG) test data at baseline, 4 weeks, and 8 weeks[a].

	TUG, mean (SD, 95% CI)	Adjusted *P* value[b]
Baseline	9.0 (2.1, 8.4 - 9.6)	—[c]
4 wk	8.0 (1.5, 7.6 - 8.5)	.01[d]
8 wk	7.5 (1.5, 7.1 - 8.0)	.001[d]

[a]The data are the mean, n (%), and ANOVA (95% CI) results.

[b]Adjusted *P* values represent comparisons with baseline.

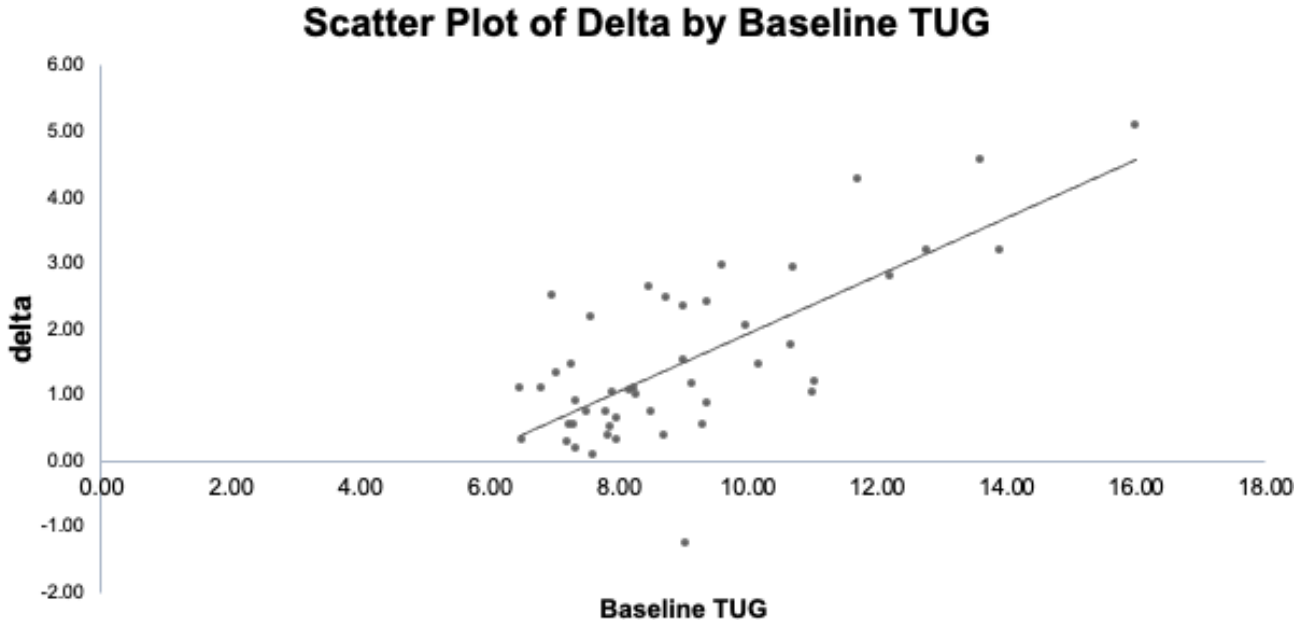
[c]Not applicable.

[d]*P*<.05.

We next examined the correlation. The analyses revealed that the baseline TUG score was significantly positively correlated with the improvement in TUG score obtained after 8 weeks of DTx (Pearson correlation coefficient, ie, *r*=0.728; *P*<.001; Figure 3).

XSL•FO
RenderX

Figure 3. The correlation between the baseline Timed Up and Go (TUG) score and the difference in the TUG score after 8 weeks of digital therapeutics (DTx) in each participant's case. The TUG score at baseline was positively correlated with the improvement in the TUG score obtained after 8 weeks of DTx (r=0.728; P<.001).



25-Question Geriatric Locomotive Function Scale

The results on this scale revealed that the GLFS-25 score improved in 43 of the 47 patients, with mean scores of 18.7 (95% CI 14.5 - 22.8) points at baseline, 12.9 (95% CI 10.4 - 15.4) points at 4 weeks, and 11.7 (95% CI 8.8 - 14.7) points at 8 weeks. The GLFS-25 at 8 weeks was significantly lower than the baseline values (P=.004; Table 3).

Table . The participants' 25-question Geriatric Locomotive Function Scale (GLFS-25) scores at baseline, 4 weeks, and 8 weeks of digital therapeutics (DTx)[a].

	GLFS-25, mean (SD, 95% CI)	Adjusted P value[b]
Baseline	18.7 (14.2, 14.5 - 22.8)	—[c]
4 wk	12.9 (8.4, 10.4 - 15.4)	.02[d]
8 wk	11.7 (10.1, 8.8 - 14.7)	.004[d]

[a]The data are the mean, n (%), and ANOVA (95% CI) results.

[b]Adjusted P values represent comparisons with baseline.

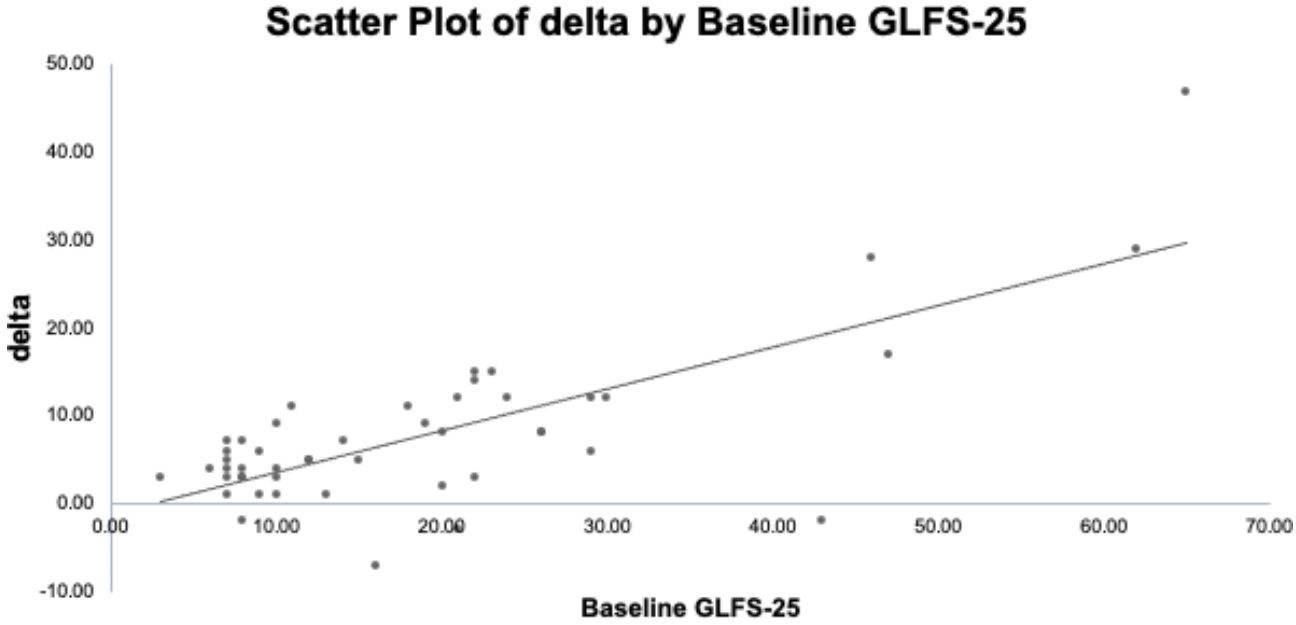
[c]Not applicable.

[d]P<.05.

When the correlation between the Δ in the GLFS-25 score from baseline to 8 weeks and the initial value was examined, a higher correlation was observed, with the participants who had higher initial GLFS-25 scores showing greater improvement at 8 weeks (r=0.749; P<.001; Figure 4).

XSL•FO
RenderX

Figure 4. The correlation between the baseline 25-question Geriatric Locomotive Function Scale (GLFS-25) score and the difference in the GLFS-25 score after 8 weeks of digital therapeutics (DTx) in each participant's case. A high correlation was observed, indicating that participants with higher initial GLFS-25 scores showed greater improvement after the 8-week intervention (r=0.749; P<.001).



Secondary Outcome Measures

BREQ-3

Our comparison of the BREQ-3 data between the baseline and end of the 8-week DTx intervention revealed no significant difference in these data (Table 4).

XSL•FO
RenderX

Table . The participants' Behavioral Regulation in Exercise Questionnaire-3 (BREQ-3) data at baseline, 4 weeks, and 8 weeks (mean and ANOVA).

	Value, mean (SD, 95% CI)	Adjusted *P* value[a]
Intrinsic motivation		
Baseline	11.2 (3.0, 10.4 - 11.2)	—[b]
4 wk	11.5 (2.8, 10.7 - 12.3)	.69
8 wk	11.8 (3.0, 10.9 - 12.7)	.53
Integrated regulation		
Baseline	10.5 (3.2, 9.6 - 11.5)	—
4 wk	10.5 (3.2, 9.5 - 11.4)	.95
8 wk	10.6 (3.4, 9.6 - 11.6)	.95
Identified regulation		
Baseline	13.3 (1.5, 12.9 - 13.8)	—
4 wk	13.3 (2.0, 12.7 - 13.9)	.99
8 wk	13.5 (2.3, 12.9 - 14.2)	.64
Introjected regulation		
Baseline	8.8 (3.3, 7.8 - 9.8)	—
4 wk	8.7 (3.3, 7.7 - 9.6)	.90
8 wk	8.4 (3.5, 7.3 - 9.8)	.52
External regulation		
Baseline	7.1 (3.4, 6.1 - 8.1)	—
4 wk	7.4 (2.6, 6.6 - 8.1)	.68
8 wk	7.1 (3.0, 6.2 - 7.9)	.82
Amotivation		
Baseline	4.8 (2.0, 4.2 - 5.4)	—
4 wk	4.8 (2.0, 4.2 - 5.4)	.88
8 wk	4.9 (2.3, 4.2 - 5.6)	.85

[a]Adjusted *P* values represent comparisons with baseline.

[b]Not applicable.

Adverse Events in This Study

No adverse events such as falls occurred.

Adherence to the Study

The average LT completion rate was 93% (52.1/56 days, 95% CI 50.4 - 53.9), demonstrating high adherence to the smartphone-based DTx. Of the 47 participants, 46 continued using the app until the end of the study period, for a retention rate of 97%. Only 1 participant abandoned the app before the end of the study period.

Discussion

Principal Results

We conducted this study to determine whether the use of DTx via a smartphone app can induce behavioral change and contribute to the prevention or improvement of LS among middle-aged and older adults. To our knowledge, this is the first study assessing the efficacy of DTx for both modifying exercise behavior and managing LS. It has been reported that continued LT for a 6-month period shortened (ie, improved) participants' TUG times [23] and that practicing 1-leg standing prevents falls [27]. It was also observed that the performance of squats enhanced the muscle strength and functional capacity of older women [28]. Another study revealed that among older adults with stage 1 LS, those who underwent LT and aerobic exercise for 8 weeks showed a significant increase in grip strength from before to after the intervention compared to those who underwent aerobic exercise alone for 8 weeks [29]. Based on this evidence, we selected LT to target lower-limb strength and balance, both of which are essential for ambulation and daily functioning.

Few studies have attempted home-based LT interventions. Aoki et al [30] reported that a 3-month program with 3×/week phone calls significantly improved their study participants' physical function and health-related quality of life; they also described an 85.4% adherence rate for the program. Similarly, Ito et al [31] demonstrated the achievement of improved physical function through regular phone-based guidance in a preventive care setting. Although such prior studies demonstrated the

XSL•FO
RenderX

efficacy of home-based LT, they were highly dependent on human resources. In contrast, our findings indicate that an app-based intervention can markedly reduce this burden, providing a scalable approach to LS management.

Our participants' motor function was assessed by the TUG test, which showed a significant 1.5-second mean improvement post intervention. Although this change did not reach our study's predefined MCID of 3 seconds, the mean postintervention TUG time was 7.5 seconds, which is below the average TUG time between the ages of 40 and 59 years [32]. Furthermore, the minimum detectable change for patients with knee osteoarthritis aged 45 to 70 years has been reported to be 1.10 to 1.14 [33]. The results of this study suggest that the observed changes are clinical rather than measurement errors. Considering that the MCID of 3 seconds demonstrated in previous studies was calculated from TUG scores before and after invasive procedures such as lumbar discectomy or lumbar decompression surgery [25], the results obtained in this study through a noninvasive intervention may indicate clinically meaningful improvements related to real-world activities.

In Japan, LS corresponds to the musculoskeletal ambulation disability symptom complex (MADS), which is covered by national health insurance. A TUG time ≥11 seconds is one diagnostic criterion for MADS. Notably, all participants (7/47, 14.8%) with baseline TUG times (≥11 s) improved to below this threshold after the DTx intervention. Moreover, those with longer baseline TUG times showed greater improvement. These findings suggest that our smartphone app may have potential to improve mobility for high-risk individuals meeting the MADS criteria and eligible for outpatient rehabilitation. In this study, the sample size was designed based on the MCID of TUG, and with only 7 cases exceeding TUG 11 seconds, stratified analysis proved challenging. To optimize this intervention, future research incorporating stratified analysis with a larger sample size is required.

The results of the present DTx were further supported by significant improvements in another primary outcome, the GLFS-25. After the 8-week intervention, GLFS-25 scores were markedly reduced compared with baseline. By the end of the program, 29 of 47 (62%) participants improved by at least 1 LS stage, and 16 (34%) no longer met the diagnostic criteria for LS. Among the 10 participants initially classified as stage 3, 7 improved to stage 2, suggesting potential efficacy even in advanced cases. These findings highlight the therapeutic potential of this DTx for LS management.

The effect of DTx on individuals' exercise behavior has been assessed based on SDT [22,34]. Although our analyses of BREQ-3 results revealed no significant changes across motivational indices, some meaningful trends emerged. The number of participants reporting increased "external regulation" (a BREQ-3 subitem) remained unchanged, whereas the subitem "intrinsic motivation" and the relative autonomy index [35] tended to improve at 8 weeks compared to the baseline. These trends suggest that extended intervention may enhance participants' autonomous motivation, warranting further investigation.

In this study, all participants received an exercise program based on LT without individualization. The exercise program demonstrated a favorable continuation rate, and no increase in amotivation or external regulation was observed on the BREQ-3; consequently, the exercise intensity is considered to have been appropriate.

Limitations

This study has several limitations. It was a proof-of-concept investigation without a control group and involved a short observation period. Participants were likely to have high baseline health interest, introducing potential selection bias. Future studies should therefore use a randomized controlled design with a larger sample size, longer follow-up, and a control group receiving only verbal exercise instruction. In addition, the study intervention was a standardized LT protocol, which may not have been sufficiently tailored to the individual physical conditions and needs of each participant. The primary focus of this app's functionality was on the remote management of participants' exercise habits. However, had it been able to assess each participant's individual exercise capacity and implement more personalized exercise programs, the improvement in exercise capacity could have been greater. Further consideration is required regarding the optimization of the app's functionality in this regard. Moreover, the intervention did not produce a clear behavioral change in exercise habits. Without sustained behavioral modification, the long-term outcomes of app-based interventions may diminish. Integrating behavioral science–based strategies will be important to maintain user engagement and therapeutic efficacy. Finally, there may have been a measurement error in the evaluation of our participants' TUG times. A related measurement error of 0.22 sec has been reported [36]. The measurement error in a TUG assessment is greater in frail older adults, and the longer it takes to perform the TUG test, the greater the variability is [37].

Conclusions

The results of our analyses demonstrate that an 8-week LT intervention using DTx for middle-aged and older adults significantly improved their TUG and GLFS-25 scores. Medical staff supervised the participants' use of the smartphone DTx, and no increase in amotivation to exercise or refusal to exercise was observed. Our findings will support the prevention and amelioration of LS in middle-aged and older adults, and the establishment of better exercise habits in this population.

Acknowledgments

The authors thank all of the study participants, and we thank the staff of the Bange-Kosei General Hospital, Hanawa-Kosei General Hospital, and Minami Tohoku Fukushima Hospital for their help with participant recruitment and data acquisition.

XSL•FO
RenderX

Funding

This study was supported by funding from the Japan Orthopaedic Association (Subsidized Science Project Research 2022-1).

Data Availability

The datasets generated and analyzed in this study are available from the corresponding author upon reasonable request.

Authors' Contributions

T Sonobe carried out the study and drafted the manuscript. TM, SH, and YM participated in the creation of the study design and helped to draft the manuscript. T Sonobe, IO, TY, T Seki, KW, and Y Kaneko recruited the participants. Y Kadowaki and YT contributed to the creation of the smartphone app. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Nakamura K. Locomotive syndrome: disability-free life expectancy and locomotive organ health in a “super-aged” society. J Orthop Sci 2009 Jan;14(1):1-2. [doi: 10.1007/s00776-008-1302-y] [Medline: 19214680]
2. Kitaura Y, Nishimura A, Senga Y, Sudo A. Locomotive syndrome affects the acquisition of long-term care insurance system certification. J Orthop Sci 2024 Jan;29(1):321-326. [doi: 10.1016/j.jos.2022.11.021] [Medline: 36577591]
3. Summary report of Comprehensive Survey of Living Conditions 2022. : Ministry of Health, Labour and Welfare; 2022 URL: https://www.mhlw.go.jp/english/database/db-hss/dl/report_gaikyo_2022.pdf? [accessed 2026-02-23]
4. Nakamura K, Ogata T. Locomotive syndrome: definition and management. Clin Rev Bone Miner Metab 2016;14(2):56-67. [doi: 10.1007/s12018-016-9208-2] [Medline: 27375370]
5. Yoshimura N, Muraki S, Oka H, et al. Association between new indices in the locomotive syndrome risk test and decline in mobility: third survey of the ROAD study. J Orthop Sci 2015 Sep;20(5):896-905. [doi: 10.1007/s00776-015-0741-5] [Medline: 26104219]
6. Ogata T, Muranaga S, Ishibashi H, et al. Development of a screening program to assess motor function in the adult population: a cross-sectional observational study. J Orthop Sci 2015 Sep;20(5):888-895. [doi: 10.1007/s00776-015-0737-1] [Medline: 26008771]
7. Taniguchi M, Ikezoe T, Tsuboyama T, et al. Prevalence and physical characteristics of locomotive syndrome stages as classified by the new criteria 2020 in older Japanese people: results from the Nagahama Study. BMC Geriatr 2021 Sep 9;21(1):489. [doi: 10.1186/s12877-021-02440-2] [Medline: 34503459]
8. Teng B, Gomersall SR, Hatton AL, Khan A, Brauer SG. Predictors of real-world adherence to prescribed home exercise in older patients with a risk of falling: a prospective observational study. Aging Med (Milton) 2023 Dec;6(4):361-369. [doi: 10.1002/agm2.12270] [Medline: 38239715]
9. Free C, Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of M-health technologies for improving health and health services: a systematic review protocol. BMC Res Notes 2010 Oct 6;3:250. [doi: 10.1186/1756-0500-3-250] [Medline: 20925916]
10. Digital Therapeutics Alliance. URL: https://dtxalliance.org [accessed 2025-09-25]
11. Patel NA, Butte AJ. Characteristics and challenges of the clinical pipeline of digital therapeutics. NPJ Digit Med 2020 Dec 11;3(1):159. [doi: 10.1038/s41746-020-00370-8] [Medline: 33311567]
12. Kario K, Nomura A, Harada N, et al. Efficacy of a digital therapeutics system in the management of essential hypertension: the HERB-DH1 pivotal trial. Eur Heart J 2021 Oct 21;42(40):4111-4122. [doi: 10.1093/eurheartj/ehab559] [Medline: 34455443]
13. Nomura A, Tateno H, Masaki K, et al. A novel smoking cessation smartphone app integrated with a mobile carbon monoxide checker for smoking cessation treatment: protocol for a randomized controlled trial. JMIR Res Protoc 2019 Feb 11;8(2):e12252. [doi: 10.2196/12252] [Medline: 30741641]
14. Ramakrishnan P, Yan K, Balijepalli C, Druyts E. Changing face of healthcare: digital therapeutics in the management of diabetes. Curr Med Res Opin 2021 Dec;37(12):2089-2091. [doi: 10.1080/03007995.2021.1976737] [Medline: 34511002]
15. Long J, You J, Yang Y. Effect of digital exercise therapy on the pain and physical function of patients with osteoarthritis: systematic review and meta-analysis. J Med Internet Res 2025 Apr 25;27:e66037. [doi: 10.2196/66037] [Medline: 40279572]
16. Valentijn PP, Tymchenko L, Jacobson T, et al. Digital health interventions for musculoskeletal pain conditions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022 Sep 6;24(9):e37869. [doi: 10.2196/37869] [Medline: 36066943]
17. Sonobe T, Matsumoto Y. Locomotive syndrome digital therapeutics provided via a smartphone app: protocol for a single-group trial. JMIR Res Protoc 2025 Apr 17;14:e70163. [doi: 10.2196/70163] [Medline: 40246298]
18. Keller K, Engelhardt M. Strength and muscle mass loss with aging process. Age and strength loss. Muscles Ligaments Tendons J 2013 Oct;3(4):346-350. [Medline: 24596700]

XSL•FO
RenderX

19. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The Mini-Cog: a cognitive "vital signs" measure for dementia screening in multi-lingual elderly. Int J Geriatr Psychiatry 2000 Nov;15(11):1021-1027. [doi: 10.1002/1099-1166(200011)15:11<1021::aid-gps234>3.0.co;2-6] [Medline: 11113982]
20. Borson S, Scanlan JM, Chen P, Ganguli M. The Mini-Cog as a screen for dementia: validation in a population-based sample. J Am Geriatr Soc 2003 Oct;51(10):1451-1454. [doi: 10.1046/j.1532-5415.2003.51465.x] [Medline: 14511167]
21. Locomotive syndrome. : The Locomotive Challenge! Council, Japanese Orthopaedic Association; 2015 Oct 1 URL: https://locomo-joa.jp/assets/files/index_english.pdf [accessed 2026-02-23]
22. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
23. Cid L, Monteiro D, Teixeira D, et al. The Behavioral Regulation in Exercise Questionnaire (BREQ-3) Portuguese-version: evidence of reliability, validity and invariance across gender. Front Psychol 2018;9:1940. [doi: 10.3389/fpsyg.2018.01940] [Medline: 30364299]
24. de Oliveira Vanini J, Karloh M, Coelho Bosco R, de Souza MG, Karsten M, Matte DL. The Brazilian Portuguese version of the Behavioral Regulation in Exercise Questionnaire 3 (BREQ-3) is reliable and valid for assessing motivational regulations and self-determination in exercise among adults aged 50 years or older: a methodological study. Int J Environ Res Public Health 2025 Jan 9;22(1):82. [doi: 10.3390/ijerph22010082] [Medline: 39857535]
25. Gautschi OP, Stienen MN, Corniola MV, et al. Assessment of the minimum clinically important difference in the Timed Up and Go test after surgery for lumbar degenerative disc disease. Neurosurgery 2017 Mar 1;80(3):380-385. [doi: 10.1227/NEU.0000000000001320] [Medline: 27352275]
26. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
27. Sakamoto K, Nakamura T, Hagino H, et al. Effects of unipedal standing balance exercise on the prevention of falls and hip fracture among clinically defined high-risk elderly individuals: a randomized controlled trial. J Orthop Sci 2006 Oct;11(5):467-472. [doi: 10.1007/s00776-006-1057-2] [Medline: 17013734]
28. Young CM, Weeks BK, Beck BR. Simple, novel physical activity maintains proximal femur bone mineral density, and improves muscle strength and balance in sedentary, postmenopausal Caucasian women. Osteoporos Int 2007 Oct;18(10):1379-1387. [doi: 10.1007/s00198-007-0400-6] [Medline: 17572834]
29. Nayasista AH, Tinduh D, Alit Pawana IP, Mei Wulan SM, Utomo DN, Soenarnatalina M. Effect of combined locomotor training and aerobic exercise on increasing handgrip strength in elderly with locomotive syndrome: a randomised controlled trial. Ann Med Surg (Lond) 2022 Jun;78:103800. [doi: 10.1016/j.amsu.2022.103800] [Medline: 35734655]
30. Aoki K, Sakuma M, Ogisho N, Nakamura K, Chosa E, Endo N. The effects of self-directed home exercise with serial telephone contacts on physical functions and quality of life in elderly people at high risk of locomotor dysfunction. Acta Med Okayama 2015;69(4):245-253. [doi: 10.18926/AMO/53561] [Medline: 26289916]
31. Ito S, Hashimoto M, Aduma S, Yasumura S. Effectiveness of locomotion training in a home visit preventive care project: one-group pre-intervention versus post-intervention design study. J Orthop Sci 2015 Nov;20(6):1078-1084. [doi: 10.1007/s00776-015-0760-2] [Medline: 26290221]
32. Kear BM, Guck TP, McGaha AL. Timed Up and Go (TUG) test: normative reference values for ages 20 to 59 years and relationships with physical and mental health risk factors. J Prim Care Community Health 2017 Jan;8(1):9-13. [doi: 10.1177/2150131916659282] [Medline: 27450179]
33. Alghadir A, Anwer S, Brismée JM. The reliability and minimal detectable change of Timed Up and Go test in individuals with grade 1-3 knee osteoarthritis. BMC Musculoskelet Disord 2015 Jul 30;16:174. [doi: 10.1186/s12891-015-0637-8] [Medline: 26223312]
34. Marcos-Pardo PJ, Martínez-Rodríguez A, Gil-Arias A. Impact of a motivational resistance-training programme on adherence and body composition in the elderly. Sci Rep 2018 Jan 22;8(1):1370. [doi: 10.1038/s41598-018-19764-6] [Medline: 29358716]
35. Martinez Kercher VM, Burton D, Pickering MA, Kercher K. Profiling physical activity motivation based on reasons for exercise: a cluster analysis approach. Psychol Rep 2024 Feb;127(1):124-141. [doi: 10.1177/00332941221119413] [Medline: 35968560]
36. Gautschi OP, Smoll NR, Corniola MV, et al. Validity and reliability of a measurement of objective functional impairment in lumbar degenerative disc disease: the Timed Up and Go (TUG) test. Neurosurgery 2016 Aug;79(2):270-278. [doi: 10.1227/NEU.0000000000001195] [Medline: 26702840]
37. Nordin E, Rosendahl E, Lundin-Olsson L. Timed "Up & Go" test: reliability in older people dependent in activities of daily living--focus on cognitive state. Phys Ther 2006 May;86(5):646-655. [Medline: 16649889]

Abbreviations

BREQ-3: Behavioral Regulation in Exercise Questionnaire-3
DTx: digital therapeutics
GLFS-25: 25-question Geriatric Locomotive Function Scale

XSL•FO
RenderX

JOA: Japanese Orthopaedic Association
LS: locomotive syndrome
LSC: locomotive syndrome check
LT: locomotive training
MADS: musculoskeletal ambulation disability symptom complex
MCID: minimal clinically important difference
SDT: self-determination theory
TUG: Timed Up and Go

Edited by S Cacciatore; submitted 20.Oct.2025; peer-reviewed by A Nomura, K tabira; revised version received 06.Jan.2026; accepted 09.Feb.2026; published 13.Mar.2026.

Please cite as:
Sonobe T, Ogawa I, Seki T, Watanabe K, Kaneko Y, Yokota T, Mawatari T, Harada S, Kadowaki Y, Takenaka Y, Matsumoto Y
Locomotive Syndrome Digital Therapeutics Provided via a Smartphone App: Proof-of-Concept Single-Group Trial Study
JMIR Aging 2026;9:e86174
URL: https://aging.jmir.org/2026/1/e86174
doi:10.2196/86174

© Tatsuru Sonobe, Itaru Ogawa, Takahiro Seki, Kosuke Watanabe, Yota Kaneko, Takeru Yokota, Taro Mawatari, Satoru Harada, Yasumichi Kadowaki, Youhei Takenaka, Yoshihiro Matsumoto. Originally published in JMIR Aging (https://aging.jmir.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Smartphone-Based Field Assessment of Trunk Stability and Its Relationship With Whole-Body Balance in Older Adults: Cross-Sectional Study

Javier De Los Ríos-Calonge[1*], MSc; Casto Juan-Recio[1*], PhD; Amaya Prat-Luri[1*], PhD; Aarón Miralles-Iborra[1*], PhD; Jose LL Elvira[1*], PhD; David Barbado[1,2*], PhD; Francisco Jose Vera-Garcia[1,2*], PhD

[1]Sports Research Centre, Department of Sport Sciences, Miguel Hernández University of Elche, Avda. de la Universidad s/n., Elche, Spain

[2]Alicante Institute for Health and Biomedical Research (ISABIAL), Alicante, Spain

[*]all authors contributed equally

Corresponding Author:
David Barbado, PhD
Sports Research Centre, Department of Sport Sciences, Miguel Hernández University of Elche, Avda. de la Universidad s/n., Elche, Spain

Abstract

Background: Maintaining balance is essential for older adults to preserve independence and reduce fall risk. However, empirical evidence linking trunk stability with balance and gait is scarce, partly due to the lack of accessible field tests. Smartphone-based accelerometry field tests offer a promising approach to assess trunk stability outside the laboratory.

Objective: The study aimed to determine (1) the association between trunk stability and whole-body balance and gait and (2) the extent to which trunk stability and whole-body balance field tests align with laboratory-based assessments in older adults.

Methods: A total of 53 physically active older adults performed 2 trunk stability tests (ie, the lumbopelvic stability [field] and the unstable sitting posturographic [laboratory] tests), 2 whole-body balance tests (ie, the tandem stance posturographic [laboratory] and the whole-body static balance [field] tests), and 2 gait-related tests (ie, the straight-line gait and the modified Timed Up and Go [field] tests) 2 times. A Spearman correlation analysis was performed after assessing the reliability of the variables.

Results: Significant positive low-to-high correlations were found between the lumbopelvic stability and the whole-body static balance tests ($0.380 \leq r \leq 0.700$; $P<.01$). A low positive correlation ($r=0.436$; $P<.01$) was also observed between the unstable sitting posturographic test and the tandem stance posturographic test. No correlations were found between trunk stability and any gait-related test. No moderate-to-high relationships ($0.289 \leq r \leq 0.487$; $P<.05$) were found between laboratory and field tests (both for trunk stability and whole-body balance).

Conclusions: Better static lumbopelvic control during exercises like the back bridge and the bird-dog is linked to improved whole-body balance in active older adults, suggesting that these positions be included in future interventions to enhance trunk stability. Accessible smartphone-based field tests for lumbopelvic stability offer a promising approach for the standardized assessment of trunk stability. However, direct extrapolation from field tools to laboratory findings is unfeasible due to a lack of correlation. These findings highlight the potential of smartphone-based field tests to complement clinical and research assessments of balance in older adults.

(JMIR Aging 2026;9:e83546) doi:10.2196/83546

KEYWORDS

mHealth; smartphone accelerometry; field tests; association; core stability; postural balance; screening; aging

Introduction

Maintaining balance and gait in older adults is crucial for functional independence and reducing fall risk [1]. Biomechanical research suggests that proper trunk neuromuscular control could play a key role in preserving motor function in this population, as it has been linked with the ability to maintain an upright posture [2], transfer loads [3], and coordinate movements during daily activities [4]. Previous studies that measured trunk stability (ie, understood as the capability to maintain or resume a relative position or trajectory of the trunk following internal or external forces) [5] in both laboratory [6] and field settings [7] support this idea, as they found reduced voluntary trunk control among older people in seated positions compared to young individuals [6], and they were able to discriminate between fallers and nonfallers [7]. More recent research has used smartphone-based accelerometry at the thoracic level to examine associations between trunk

XSL•FO
RenderX

accelerations during upright balance and functional tasks, such as the sit-to-stand task, in older adults [8-10]. While these approaches provide valuable insights into higher global postural control during standing tasks, trunk behavior is evaluated within movements in which control is distributed across multiple joints and segments. Consequently, there is a lack of empirical evidence directly linking trunk-focused stability outcomes (ie, derived from tasks designed to impose relatively high demands on lumbopelvic control) with whole-body balance and gait-related parameters in older adults. In this sense, while trunk-focused exercise programs have proven to be useful to improve balance and gait in this population [11], it remains unclear to what extent these benefits stem from enhanced trunk stability, as this capability is not typically assessed. Moreover, although correlational studies have shown positive relationships of trunk muscle strength [12-14], endurance [13], muscle thickness [15,16], and spinopelvic range of motion [12,17,18] with whole-body balance (ie, static and dynamic) [12,18], gait parameters (ie, step length, gait speed) [14,17,18], and functional capacity (ie, the time taken to complete the sit-to-stand and the Timed Up and Go tests) [12,14], correlational evidence specifically addressing trunk stability remains scarce. To the authors' knowledge, only 2 studies have linked trunk-focused stability outcomes from a seated dynamic postural control test on a force platform (in which lower-limb and arm strategies are largely constrained, making task performance predominantly dependent on trunk control) with observational-based assessments of whole-body balance and gait in older adults [19,20].

A key reason for the limited research in this field is the absence of field tests that objectively quantify trunk stability. Inertial sensors have emerged as cost-effective field-setting tools to reliably detect postural control impairments in older adults [21]. While laboratory-based force platform assessments remain the standard reference for biomechanical analysis, their limited accessibility has motivated the development of field-based alternatives. Based on this approach, our research group has developed a novel accelerometry-based field method [22] to assess trunk-focused stability outcomes during common floor-based trunk exercises (eg, bridges and bird-dogs). While this method has shown to be promising in young adults [23], its applicability and correlation with whole-body balance and gait in older adults remain unexplored. Since these exercises are used in exercise programs that aim to preserve motor function in older people [24,25], introducing an accelerometry-based assessment of trunk stability would help to establish which exercises are most closely associated with whole-body balance and gait and thus should be prioritized within training programs.

Given this scarcity of direct empirical evidence, the primary aim of this study was to explore the associations between biomechanically based trunk stability, whole-body balance, and gait outcomes in physically active older adults. This was managed by using both well-established laboratory protocols and novel inertial-sensor-based field tests: (1) trunk stability tests: the unstable sitting posturographic test (laboratory-based) [5] and the lumbopelvic stability tests (field-based) [22]; (2) whole-body balance tests: the tandem stance dynamic posturographic test (laboratory-based) [26] and the whole-body static balance test (field-based) [27]; and (3) gait-related tests: the straight-line gait [27] and the modified Timed Up and Go (field-based) [28] tests. A secondary exploratory aim was to analyze the relationships between these laboratory and field assessments to inform practitioners about their alignment. Finally, this study aimed to determine the between-session test-retest reliability (both relative and absolute) of all outcomes, to support interpretation of the correlational findings, and to guide the selection of robust and reproducible performance metrics. Based on the principle of specificity and on previous studies [23,26], it was first hypothesized that trunk stability would show moderate associations with biomechanical whole-body balance outcomes only in those tests that share task requirements. Specifically, we hypothesized that lumbopelvic stability tests would correlate with whole-body static balance tests, given the shared demand for maintaining steady positions (eg, standing, bridging, and bird-dog). Thereafter, we hypothesized that the unstable sitting posturographic test would correlate with the tandem stance posturographic test due to their common requirement to anticipate and respond to predicted disturbances. Regarding gait-related outcomes, we hypothesized that there would be no correlation with trunk stability tests, except between the unstable sitting posturographic test and the modified Timed Up and Go test, since voluntary trunk control may play a key role in standing, sitting, and turning.

Methods

Participants

In order to check the proposed hypotheses, a correlation coefficient (r) of 0.5 was expected for the sample size estimation, considering previous studies that examined the relationship between trunk stability tests and whole-body dynamic balance tests [23,26]. The software G*Power 3.1 (v3.1, University of Düsseldorf) was used with the following parameters: r=0.5; α=.006; 1–β=.8. Calculations suggested a sample size of approximately 46 participants. An additional 20% of the target sample was recruited to anticipate the loss of subjects due to the inability to implement some of the tests presented. A total of 53 volunteer participants aged 65 years or older (mean age 70.3, SD 3.8 y; 21 male participants and 32 female participants) were recruited for the study. The main sociodemographic and anthropometric characteristics and the physical activity levels of the participants are shown in Table 1. Moreover, Table S1 in Multimedia Appendix 1 shows the clinical data of the participants. Participants were asked to present a medical certificate signed by their primary care physician to guarantee that they had no contraindications to perform trunk stability, whole-body balance, and gait-related tests.

XSL•FO
RenderX

Table . Main sociodemographic and anthropometric characteristics and physical activity levels of the participants.

	Male participants (n=21)	Female participants (n=32)	Total (N=53)
Age and anthropometric data, mean (SD)			
Age (y)	70.3 (4.1)	70.3 (3.7)	70.3 (3.8)
Body mass (kg)	77.1 (9.0)	64.6 (7.9)[a]	69.6 (10.3)
Height (m)	1.70 (0.07)	1.56 (0.06)[a]	1.62 (0.09)
Trunk height (m)	0.57 (0.03)	0.50 (0.04)[a]	0.53 (0.05)
Body moment of inertia ($kg \cdot m^2$)	67.6 (10.3)	47.7 (7.0)[a]	55.6 (12.9)
Trunk moment of inertia ($kg \cdot m^2$)	9.8 (1.7)	6.5 (1.5)[a]	7.8 (2.3)
Sociodemographic data, n (%)			
Education			
Primary studies	3 (14)	12 (38)	15 (28)
Secondary studies	12 (57)	12 (38)	24 (45)
Higher studies	3 (14)	2 (6)	5 (9)
Vocational training	0 (0)	5 (16)	5 (9)
None	3 (14)	1 (3)	4 (8)
Living alone	0 (0)	8 (25)	8 (15)
Financial difficulties	1 (5)	3 (9)	4 (8)
Physical activity data, n (%)			
Physically active for the past 20 y			
>150 mins/wk	10 (48)	15 (47)	25 (47)
≤150 mins/wk	11 (52)	17 (53)	28 (53)
Performing strength training	4 (19)	15 (47)	19 (36)
Currently physically active			
>150 mins/wk	14 (67)	20 (63)	34 (64)
≤150 mins/wk	7 (33)	12 (38)	19 (36)
Performing strength training	9 (43)	21 (66)	30 (57)

[a]Pairwise comparisons between sexes: significant differences (P<.05).

Ethical Considerations

The University Office for Research Ethics (DCD.FBM.01.23) approved this study. The study was conducted in accordance with the Declaration of Helsinki, informing the participants of the characteristics and risks of the study. Each participant provided written informed consent. Data were handled to ensure privacy and confidentiality and were accessible only to the research team. No financial compensation was provided. Written informed consent was obtained for the publication of the participant's images.

Experimental Procedure

Each participant's mass (mean 69.6, SD 10.3 kg); height (mean 1.62, SD 0.10 m); and height of the center of mass position of the head, arms, and trunk segment (estimated at 62.6% [29] of the trunk height, measured as the distance between the greater trochanter and the glenohumeral joint: 0.53, SD 0.05 m) were registered before testing. The participant's preferred leg was established as the leg that the participant would use to kick a ball. Participants carried out 2 testing sessions with a day's rest in-between, and these sessions were repeated 1 week later. The first assessment session consisted of a trunk stability assessment through lumbopelvic stability tests (using a randomized block design to avoid fatigue influence) during the back bridge, the preferred side bridge (determined by the side of the preferred leg), and the front bridge and the bird-dog positions [22]. The second assessment session consisted of trunk stability, dynamic balance, static balance, and gait-related tests (carrying out first those tests in which performance is most likely to be affected by fatigue) performed in the following order: (A) the unstable sitting posturographic test, (B) the tandem dynamic balance stance posturographic test, (C) whole-body static balance tests, (D) straight-line gait tests, and (E) the modified Timed Up and Go test. A 5 - to 7-minute warm-up was performed at the beginning of each testing session to enhance awareness of spine and pelvis alignment and to facilitate the understanding and practice of neutral spine posture. The following exercises were included in this warm-up: (1) isolated scapular retractions and

XSL•FO
RenderX

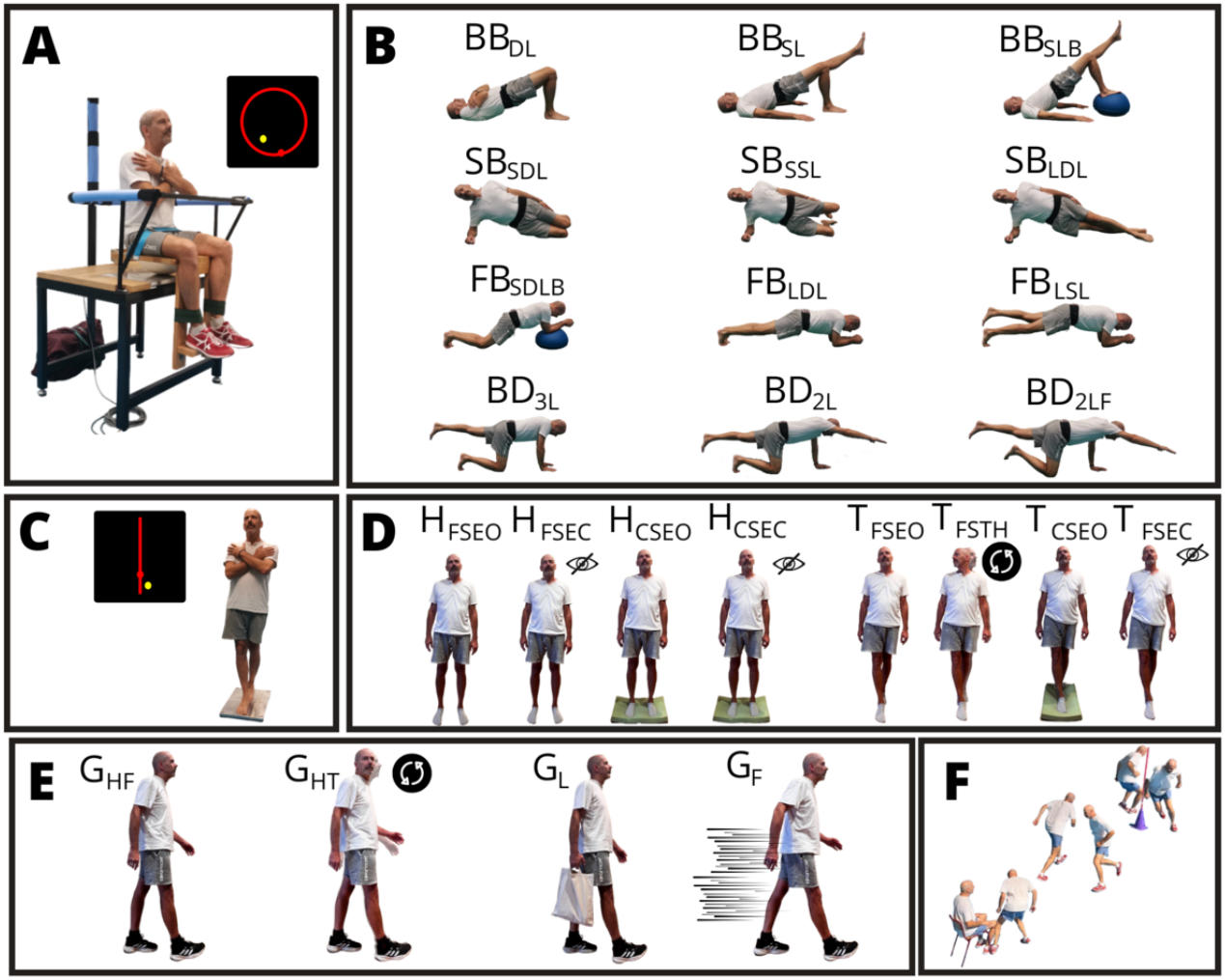
protractions; (2) isolated pelvic retroversions and anteversions; and (3) the cat-camel exercise.

Testing Protocol Description and Data Processing

Trunk Stability Tests

Lumbopelvic linear accelerations were recorded to assess the postural control challenge imposed on the participants as an index of exercise intensity during the lumbopelvic stability tests (field-based), which involved performing three 15-second back bridge, preferred side bridge, front bridge, and bird-dog positions (Figure 1B) using a novel "ad hoc" smartphone app (CoreMaker) created by our research group. A smartphone was placed with Velcro above the participant's sacrum on an elastic pelvic belt to reduce smartphone movement caused by muscle contractions similar to the protocol previously described by Heredia-Elvar et al [22]. The accelerations were recorded at 100 samples per second from a 3-axis accelerometer embedded in the smartphone (iPhone SE, 2nd generation, model MHGQ3QL/A; Apple Inc.), and it was configured using iOS version 15. Lumbopelvic stability was assessed through the mean acceleration ($m \cdot s^{-2}$) calculated as the average of the acceleration magnitude data series obtained from the 3-axis (vertical, anterior-posterior, and medial-lateral) of the smartphone accelerometer using the CoreMaker app mentioned earlier.

Figure 1. Trunk stability, whole-body balance, and gait-related tests: (A) The unstable sitting posturographic test. (B) The lumbopelvic stability tests performing 3 back bridge (back bridge with double-leg support [BB_{DL}], back bridge with single-leg support [BB_{SL}], and back bridge with single-leg support on a hemisphere ball [BB_{SLB}]), 3 preferred side bridge (side bridge with short double-leg support [SB_{SDL}], side bridge with short single-leg support [SB_{SSL}], and side bridge with long double-leg support [SB_{LDL}]), 3 front bridge (front bridge with short double-leg support and arms on hemisphere ball [FB_{SDLB}], front bridge with long double-leg support [FB_{LDL}], and front bridge with long single-leg support [FB_{LSL}]), and bird-dog (3-limb bird-dog position [BD_{3L}], 2-limb bird-dog position [BD_{2L}], and 2-limb bird-dog position with the forefoot of the supporting leg elevated [BD_{2LF}]) positions. (C) The tandem stance posturographic test. (D) The whole-body static balance tests performed with feet hip-width apart (firm surface and eyes open [H_{FSEO}], firm surface and eyes closed [H_{FSEC}], compliant surface and eyes open [H_{CSEO}], and compliant surface and eyes closed [H_{CSEC}]) and tandem (firm surface and eyes open [T_{FSEO}], firm surface and turning the head [T_{FSTH}], compliant surface and eyes open [T_{CSEO}], and firm surface and eyes closed [T_{FSEC}]) positions. (E) Straight-line gait tests (head-forward gait [G_{HF}]; head turning gait [G_{HT}]; 1-sided 10% body weight load gait [G_L]; fast gait without running [G_F]). (F) The modified Timed Up and Go test performed as fast as possible, being allowed to run. Written informed consent was obtained for the publication of the participant's images.



XSL•FO
RenderX

The unstable sitting posturographic test (laboratory-based) [5] was used to assess the participants' ability to control their trunks during a circular tracking task while sitting on an unstable seat placed on a force platform (9287CA, Kistler; Figure 1A). Real-time feedback of the participants' center of pressure displacement was displayed on a screen. Additionally, a target point (which moved in a circular trajectory) was presented to assess the participants' ability to adjust their center of pressure position to this point. A 4° inclination angle of the head, arms, and trunk center of mass (estimated at 62.6% [29] of the distance between the greater trochanter and the glenohumeral joint) was calculated to establish the movement amplitude of the target point. The participants' arms were crossed over their chest, and they performed five 70-second trials with a 60-second rest between trials. Center of pressure displacement was collected and filtered by a low-pass-Butterworth filter with a cutoff frequency of 5 Hz. The initial 10 seconds were discarded due to the nonstationarity of the signal. The mean radial error was used to quantify the trunk performance, calculated as the average of the vector distance magnitude (mm) of the center of pressure from the target point using an "ad hoc" software developed by our research group through LabView 9.0 environment (v9.0, National Instruments) [23].

Whole-Body Dynamic Balance Test

The tandem stance posturographic test (laboratory-based; Figure 1C) was used to measure the whole-body dynamic balance through an anterior-posterior tracking task (with the same characteristics as those used in the unstable sitting posturographic test but with a linear tracking task) while standing on a force platform (9286AA, Kistler) [26]. A 2° inclination angle of the whole-body center of mass (estimated at 55% [29] of each participant's height) was calculated to establish the movement amplitude of the target point. The participants performed the tandem test with their preferred lower limb placed ahead of the other limb. Their arms were crossed over their chest during both tests, and they performed each task 3 times with a 1-min rest between tasks, with a trial duration of 70 seconds. Similar to the unstable sitting posturographic test, center of pressure displacement was filtered (low-pass Butterworth filter; cutoff frequency: 5 Hz), and the initial 10 seconds were discarded [5]. Then, the mean radial error (mm) was calculated to quantify balance performance.

Whole-Body Static Balance Tests

The whole-body static balance tests (field-based) evaluated participants' postural sway through the manipulation of sensory inputs required for balance [27]. Participants wore an elastic sacroiliac belt fastened with Velcro around the lower back (around L5/S1) [27], which housed a smartphone (iPhone SE, 2nd generation, model MHGQ3QL/A). Data from the smartphone inertial sensors (accelerometer) were recorded during the tasks by the G&B smartphone app [27] at 100 samples per second. Participants were asked to stand as still as possible with their feet hip-width apart (H) and in tandem (T; Figure 1D) positions with their arms by their sides. The hip-width stance posturographic-field tests under steady-state conditions included standing on a firm surface and eyes open, on a firm surface and eyes closed, on a compliant surface and eyes open, and on a compliant surface and eyes closed. The tandem stance posturographic-field tests under steady-state conditions consisted of standing on a firm surface and eyes open, on a firm surface while turning the head, on a compliant surface and eyes open, and on a firm surface and eyes closed. The total of 8 static balance tasks (Figure 1D) were performed twice. Each trial lasted 30 seconds. The 8 tasks were performed in the order shown in Figure 1 with a rest period of 30 seconds between tasks. After a 2-min resting period, the tasks were performed again in the same order. A medium density foam mat (52 kg/m^3, 50×28×5 cm) was used as the compliant surface. The original postural stability outcome obtained from the G&B was inverse log transformed to compute the mean absolute acceleration in each balance task.

Straight-Line Gait Tests

Four gait tests in a straight line were proposed to assess walking balance (gait periodicity index) and walking speed (Figure 1E). These included head-forward gait, head turning gait, one-sided 10% body weight load gait, and fast gait without running. The gait tasks were designed to be performed at the user's preferred walking speed, except for the fast gait without running. Each task consisted of four 6-second walks. This task design allows the gait assessment to be carried out over approximately 10 m. The 4 tasks were performed in the order shown in Figure 1 with a rest period of 30 seconds between tasks, and a second set was performed in the same order after a 2-minute rest. As in the whole-body static balance tests, participants wore an elastic sacroiliac belt fastened with Velcro around the lower back (around L5/S1) [27] to house a smartphone (iPhone SE, 2nd generation, model MHGQ3QL/A). Data from the smartphone inertial sensors (accelerometer and gyroscope) were recorded by the G&B app [27] at 100 samples per second during the tasks. Specifically, a wavelet-based step-event detection algorithm and a double-pendulum gait model were used to analyze gait data by obtaining walking balance (periodicity index, %) and walking speed (m/s) outcomes through the G&B app. All straight-line gait and whole-body static balance assessments administered using the G&B smartphone app were adapted from the protocol described by Rashid et al [27], introducing additional task conditions to increase relevant task-specific postural and locomotor demands for older adults. These conditions were designed to progressively perturb balance control pathways, thereby increasing demands on segmental postural control, including the trunk.

Functional Mobility Test

The modified Timed Up and Go test consisted of standing up from a chair, moving forward 3 m, turning around a cone, returning to the chair, and sitting back down [28] (Figure 1F). Participants performed the task as fast as possible, including running if they were able, to better discriminate between participants and to avoid ceiling effects, particularly in active populations. The time (s) was recorded using a digital chronometer (HS-30W-N1V, CASIO).

If a participant failed an attempt (eg, in the lumbopelvic stability tests, a failed attempt was defined as being able to assume the required initial position but not being able to maintain proper body alignment throughout the entire trial, assuming the initial

XSL•FO
RenderX

position but discontinuing the task before completion, or not being able to assume the required initial position), the attempt was classified as invalid and repeated. Only the failed attempt was repeated, not the entire test sequence. If the participant failed again or reported being unable to safely perform the task, the trial was recorded as "not feasible," and no further attempts were requested for that specific condition. This procedure was applied consistently across all tests.

For the main analyses, we used a peak-performance criterion based on session 2 for each test condition. This approach was selected to emphasize participants' optimal task execution while limiting the impact of temporary, nonsystematic fluctuations in performance (eg, momentary attentional variation or short-term fatigue), which may occur during physically demanding tasks in older adults. Accordingly, for the lumbopelvic stability tests and for the hip-width and tandem stance posturographic-field tests, the best of the 2 recorded trials (ie, the lowest mean acceleration) was selected for further analyses. For the unstable sitting and the tandem stance dynamic posturographic laboratory tests, the average of the 2 best attempts (ie, lowest mean radial error) was used, based on 5 and 3 trials, respectively. In the straight-line gait tests, the best of the 2 recorded trials (ie, highest speed and percentage values) was used. Finally, for the modified Timed Up and Go test, participants performed 3 trials, and the best trial (ie, lowest completion time) was retained for statistical analyses.

We acknowledge that peak-performance selection can lead to an overestimation of typical functional capacity in older adults. Consequently, to evaluate the robustness and representativeness of the findings, we complemented these primary analyses with sensitivity analyses based on mean performance averaged across sessions 1 and 2. For these sensitivity analyses, performance was computed as the mean of all recorded trials within each session and then averaged across sessions. The only exception was the unstable sitting posturographic test, for which mean performance was calculated as the average of the three best attempts out of 5 trials.

Statistical Analyses

Statistical analyses were performed to address the study hypotheses and evaluate the reliability of the measurements. Data (presented as mean and SD) from those participants who carried out all the testing sessions were used for the statistical analyses. The normality of the data distribution was explored through the Shapiro-Wilk test. Since the data did not follow a normal distribution, nonparametric statistics were applied. The Wilcoxon test was performed for each test score to explore the existence of statistically significant mean differences between sessions, and the Spearman's correlation coefficient was used to analyze the relationship between the variables.

In order to avoid possible bias due to the low consistency of certain variables, typical error (TE) and intraclass correlation coefficient ($ICC_{3,1}$) were calculated (confidence limits set at 95%) to evaluate the absolute and relative test-retest reliability, respectively, and to determine which variables could be used in subsequent correlational analyses. The TE was calculated as the SD of the difference between the 2 testing sessions divided by $\sqrt{2}$. Furthermore, TE values were also expressed as percentages to facilitate data extrapolation, interpretation, and comparison with the pertinent literature. Percentage TE was calculated as the TE×100 and divided by the average of test-retest means. ICC values were interpreted according to the following criteria: poor (<0.50), moderate (0.50 - 0.75), good (0.75 - 0.90), and excellent (>0.90) [30]. Both absolute and relative reliability indexes were calculated through the spreadsheet proposed by Hopkins [31].

If some variables obtained a poor level of relative reliability [30] (ie, ICC <0.50), the Tukey outlier classification test was applied to the change between sessions. Outliers are based on the IQR (ie, [25th percentile]–1.5×IQR and [75th percentile]+1.5×IQR). The data obtained in the second assessment session for the variables that obtained a moderate-to-excellent level of relative reliability [30] (ie, ICC >0.50) were used to perform a correlation analysis (r) between them. Correlational analyses were performed with JASP 0.16.2 software (Eric-Jan Wagenmakers, Department of the Psychological Methods, University of Amsterdam, Nieuwe Achtergracht 129B). The correlation coefficient was interpreted as low (0.30 - 0.49), moderate (0.50 - 0.69), high (0.70 - 0.89), and very high (≥0.90) [32]. In order to minimize the probability of obtaining significant correlations by chance, the prespecified significance level (P<.05) was divided by the number of comparisons analyzed (Bonferroni correction).

To examine whether the observed correlations were influenced by the choice of data reduction strategy (session 2 peak-performance vs mean performance averaged across sessions 1 and 2), we conducted complementary sensitivity analyses as described in the previous section. First, we repeated the reliability analyses using mean performance across attempts within each session. Second, we repeated the correlational analyses using the average of session 1 and session 2 mean performance as a more representative estimate of typical performance and to reduce potential measurement noise. The pattern of results across these analyses was then compared with the original session 2 peak-performance analyses.

Given that the lumbopelvic stability tests represent the primary outcome of the study and that they are also the tests most affected by feasibility constraints at higher difficulty levels, an additional focused sensitivity analysis was performed to determine whether test-retest reliability metrics and the key correlations observed in the main analyses based on the session 2 peak-performance criterion were robust enough when analyses were restricted to a more homogeneous group of participants with higher physical fitness who successfully completed all 3 test levels.

Results

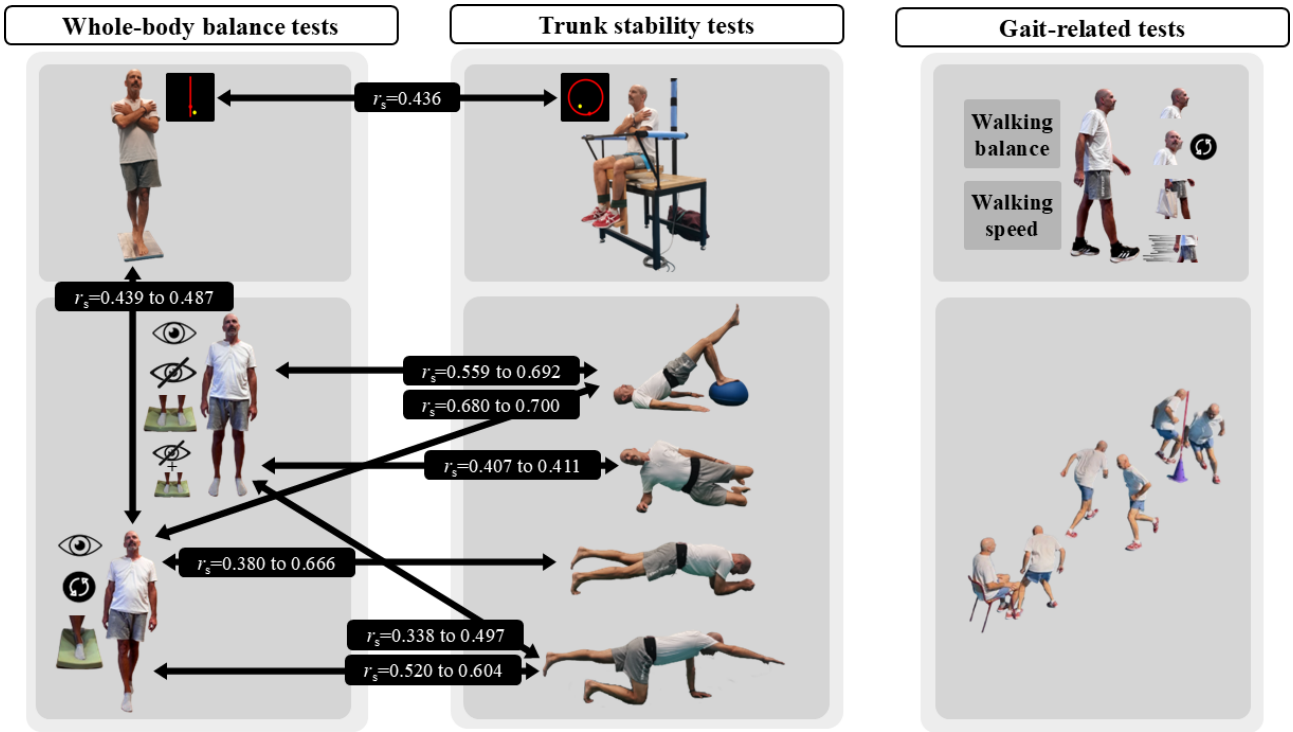
The reliability analyses demonstrated varying levels of test-retest reliability across the assessed variables. Trunk stability tests generally showed a moderate-to-excellent reliability (0.68≤ICC≤0.88; 12.44%≤TE≤31.54%; Table S2 in Multimedia Appendix 1) with the exception of the easiest variations of the back bridge (with double-leg support) and the preferred side bridge (with short double-leg support) and the easiest and most difficult variation of the bird-dog (3-limb position and 2-limb

XSL•FO
RenderX

position with the forefoot of the supporting leg elevated, respectively), which showed a poor level of reliability (0.19≤ICC≤0.39, 29.24%≤TE≤38.55%; Table S2 in Multimedia Appendix 1). A moderate-to-excellent level of reliability was shown regarding the reliability analyses for the whole-body balance and gait-related tests (0.51≤ICC≤0.85; 2.55%≤TE≤29.27%; Table S3 in Multimedia Appendix 1) except for the tandem on a firm surface with eyes closed static balance test, the periodicity index of the head-forward gait test, and the head turning gait test (0.37≤ICC≤0.43; 3.06%≤TE≤29.98%; Table S3 in Multimedia Appendix 1). These reliability values, particularly those for the tandem on firm surface and eyes closed static balance tests and the gait test periodicity index, were obtained after the application of the Tukey outlier classification test to address initial low reliability. The Wilcoxon test showed significant between-sessions differences for the mean radial error during the unstable sitting posturographic test and the tandem stance posturographic test, the accelerations of some lumbopelvic stability tests (side bridge with long double-leg support and 2-limb bird-dog positions), the whole-body static balance test with feet hip-width apart on firm surface and eyes open, and the walking speed during all the straight-line gait tests. These showed that there was a learning or repetition effect in these tests and therefore a significantly higher performance in these tests in session 2. For in-depth results, Table S2 in Multimedia Appendix 1 and Table S3 in Multimedia Appendix 1 show the descriptive statistics and the absolute and relative between-session reliability for trunk stability and whole-body balance, gait, and functional mobility tests, respectively.

Given the established reliability and observed learning effects, subsequent correlational analyses were performed using data from the second assessment session. These analyses explored relationships within and between different capabilities, as detailed below. Correlational analyses between the trunk stability tests (Figure 2) generally revealed no statistically significant relationships (P<.05) in the moderate-to-high range between the unstable sitting posturographic test and the lumbopelvic stability tests (Table S4 in Multimedia Appendix 1).

Figure 2. Main correlations between trunk stability tests, whole-body balance tests, and gait-related tests in physically active older adults. The bridge positions displayed represent those that showed the greatest number of correlations with the methodology used to evaluate whole-body static balance in field settings. The 2-limb bird-dog position was included because it was the only one with reliable values. The *tandem stance posturographic test* during a dynamic tracking task showed low but statistically significant correlations with the *tandem stance posturographic-field tests* under steady-state conditions, but no significant correlations with the *hip-width stance posturographic-field tests*. Overall, correlations between the *unstable sitting posturographic test* and the *lumbopelvic stability tests* were not statistically significant. The *unstable sitting* and *tandem stance posturographic tests* during a dynamic tracking task were lowly correlated. *Lumbopelvic stability tests* demonstrated low-to-high correlations with several whole-body balance tests under steady-state conditions, particularly for the back bridge and bird-dog positions. No significant associations were found between trunk stability tests and gait-related tests. Written informed consent was obtained for the publication of the participant's images.



As seen in Figure 2, broadly, no statistically significant relationships (P<.05) in the moderate-to-high range were found between the tandem stance posturographic test and the whole-body static balance tests. Focusing on some significant comparisons, whole-body static balance tests in the tandem position had a low significant relationship with the tandem stance posturographic test (0.439≤r≤0.487; P<.05; Table S5 in Multimedia Appendix 1).

Concerning the correlation analysis of the trunk stability tests with the whole-body balance performance tests, two of the four comparisons showed a significant association (Figure 2). A low significant association (r=0.436; P<.013) was found when

XSL•FO
RenderX

comparing the unstable sitting posturographic test with the tandem stance posturographic test. A range of significant associations, from low to high ($0.380 \leq r \leq 0.700$; $P<.013$; Table S6 in Multimedia Appendix 1), was observed when lumbopelvic stability tests were compared with whole-body static balance tests (Figure 2). Specifically, moderate correlations ($0.559 \leq r \leq 0.643$; $P<.013$) were found during the back bridge position with the 4 variations of the static balance tests in the hip-width position, some approaching high levels ($0.673 \leq r \leq 0.692$; $P<.013$). In addition, the most difficult back bridge (with single-leg support on a hemisphere ball) showed a high correlation with the moderate (firm surface and turning the head) and high (compliant surface and eyes open) difficulty tandem positions ($0.680 \leq r \leq 0.700$; $P<.013$). A low correlation was found with the variations in the hip-width position ($0.338 \leq r \leq 0.497$; $P<.013$) during the bird-dog position, and there was a moderate correlation with the variations in tandem position ($0.520 \leq r \leq 0.604$; $P<.013$). A low correlation was found with the low (firm surface and eyes open) and moderate (firm surface and turning the head) variations in tandem position ($0.380 \leq r \leq 0.427$; $P<.013$) during the low (with short double-leg support and arms on hemisphere ball) and moderate (with long double-leg support) variations of front bridge position. The most difficult front bridge (with long single-leg support) showed a high correlation with the most difficult tandem position (compliant surface and eyes open; $r=0.666$; $P<.013$).

Finally, no association was found between trunk stability and any gait-related test (Figure 2). The whole-body static balance tests in the hip-width position showed a consistently low association with walking speed in two of the four straight-line gait tests ($0.377 \leq r \leq 0.407$; $P<.006$) and a low-moderate inverse association with the modified Timed Up and Go test ($-0.501 \leq r \leq -0.444$; $P<.006$; Table S7 in Multimedia Appendix 1). This inverse relationship suggests that those participants with higher acceleration values (indicating worse balance control) tended to be faster in these gait-related tests. To try to understand this association better, it seems to be necessary to know the association of height (as a representation of the human inverted pendulum model) with whole-body static balance tests in the hip-width position ($0.366 \leq r \leq 0.489$; $P<.008$), with the modified Timed Up and Go test ($r=-0.527$; $P<.001$) and with the walking speed in the 1-sided 10% body weight load test and in the fast gait without running straight-line gait test ($0.451 \leq r \leq 0.503$; $P<.001$).

Sensitivity analyses of test-retest reliability based on mean performance are provided in Tables S8-S9 in Multimedia Appendix 1. Overall, reliability estimates were largely comparable to those obtained using the peak-performance approach, supporting the robustness of the measurement properties across different data reduction strategies. Sensitivity analyses of the correlational findings using mean performance and the average across both sessions revealed a consistent overall pattern of results, with the original analysis based on peak performance from session 2. Most associations between trunk stability, whole-body balance, and gait-related outcomes remained low and task-specific. However, a small number of correlations increased from nonsignificant or low to low-moderate and reached statistical significance (Figure 2 vs Figure S1 in Multimedia Appendix 1). Specifically, low but significant associations emerged between the unstable sitting posturographic test and the front bridge with long single-leg support position as well as with the 2-limb bird-dog position of the lumbopelvic stability tests. In addition, the association between the unstable sitting and the tandem stance dynamic posturographic laboratory tests increased from low to moderate, and small but significant associations were observed between these 2 latter tests and the modified Timed Up and Go test (Tables S10-S13 and Figure S1 in Multimedia Appendix 1). These changes are interpreted as the expected consequence of using more stable estimates of performance, which reduce random measurement error and day-to-day variability. Consequently, the primary findings and qualitative interpretation of the results remain unchanged, supporting predominantly low, selective, and task-dependent relationships between trunk stability, balance, and gait-related performance.

The additional focused sensitivity analysis conducted exclusively in participants who were able to successfully complete all 3 test levels yielded results consistent with the main analyses. For each lumbopelvic stability test position, analyses were restricted to the subset of participants who successfully completed all 3 difficulty levels within that test, and test-retest reliability metrics were recomputed accordingly. Although some estimates were attenuated due to a reduced between-subject variability, the overall directional pattern and structure of the findings were preserved. Variations demonstrating good-to-excellent reliability in the full sample retained comparable ICC values in the restricted subsamples, whereas variations with poorer reliability did not systematically improve (Tables S14 in Multimedia Appendix 1). Likewise, associations between trunk stability and whole-body balance were not strengthened (Tables S15 in Multimedia Appendix 1), indicating that the relationships observed in the main analyses were not driven by differences in physical capacity or task feasibility and supporting the robustness of both the reliability estimates and the observed associations.

Discussion

Principal Findings

The primary findings of this study highlight a significant association between superior lumbopelvic control during specific bridge and bird-dog exercises and enhanced whole-body static balance. Conversely, no such associations were observed between trunk stability and whole-body dynamic balance or gait-related outcomes. Importantly, consistent with existing literature [5,23], the lack of correlation between trunk stability and whole-body balance in laboratory and field tests underscores a critical point for sport science professionals and practitioners: the direct extrapolation of data obtained with the tools available in field settings to laboratory findings remains unfeasible.

Influence of Trunk Stability on Whole-Body Static Balance: Task Specificity

Our main results confirmed the first hypothesis, so maintaining a static position in floor-based lumbopelvic stability exercises was associated with the capability to maintain a static upright

XSL•FO
RenderX

posture [33]. Thus, incorporating these trunk stability exercises, which require trunk, shoulder, and leg strength to reach the posture, short-term muscle endurance, and muscle coactivation for postural control (via feedforward and/or feedback mechanisms) [34], into a workout routine may enhance whole-body static balance in older adults. However, not all stability exercises showed the same degree of association. Back bridge and bird-dog variations showed more and stronger correlations, emphasizing the role of the trunk-hip extensor (ie, some of the antigravity muscles of the posterior kinetic chain) to maintain static posture. These associations were more evident when balance was increasingly challenged by narrowing the support base through changes in foot position (tandem vs hip-width position), suggesting a greater involvement of the lumbopelvic complex. The asymmetrical tandem stance tends to induce body rotation and lateral adjustments, which may align with the demands of the single-leg back bridge and conventional bird-dog, both involving asymmetry, rotation, and multiplanar control. Conversely, frontal and lateral bridge variations showed fewer and lower correlations. The lack of relevant correlations for the side bridge variations was particularly unexpected as tandem stance tasks seem to have high frontal plane demands and previous studies have linked mediolateral balance to trunk control [34,35]. These results may be due to the fact that control during lateral bridges is heavily mediated by the physical condition of the shoulder complex [36], which could be more affected in older adults by age-related issues such as sarcopenia, tendinopathies, and arthritis [37].

While a significant positive correlation was observed between the well-established trunk stability and whole-body dynamic balance laboratory tests, our second hypothesis was partially rejected, as the strength of these correlations was unexpectedly low (r=0.436). This suggests that, although a statistical link exists, trunk stability as measured by the unstable sitting posturographic test may only have a limited practical impact on whole-body dynamic balance in physically active older adults, particularly compared to populations with greater balance deficits in which stronger associations have been reported [26]. In this sense, tandem stance performance showed stronger associations with trunk stability in people with multiple sclerosis [26] (who present neuromuscular deficits) when balance was greatly challenged by placing the weaker leg behind (r=0.540; P<.05), compared to when the stronger leg was behind (r=0.433; P<.05), suggesting a greater trunk involvement when postural demands increase. In our sample of physically active older adults, the demands in the unstable sitting posturographic test may not be as determinant for whole-body dynamic balance compared with more disabled populations. In this sense, once a minimum level of trunk stability is reached in a seated position, it no longer contributes to improving whole-body dynamic balance in tandem stance unless a greater balance challenge is imposed on this population [23]. This interpretation should be taken with caution, as higher associations were found (r>0.629; P<.05) in other samples of older adults of similar characteristics [38], suggesting that the role of trunk stability may have a different impact on whole-body dynamic balance in other scenarios.

Finally, as expected and given the lack of significant associations, maintaining static positions during bridging and bird-dog exercises seems to play a limited role in whole-body dynamic balance during voluntary weight-shifting tasks and also voluntary control during the unstable sitting posturographic test seems less relevant for whole-body static balance. The complex scenario presented in this section aligns with a previous study [39] that confirmed the involvement of task-specific cortical and subcortical pathways after analyzing the excitability of corticospinal inputs to trunk muscles during both voluntary and involuntary tasks. The study [39] also highlighted that this involvement could not be easily predicted based solely on whether the task was classified as voluntary or more “automatic” in nature. This highlights the need for personalized trunk stability assessments that take into account individual characteristics and the underlying neurophysiological mechanisms related to task demands, which may functionally impact whole-body balance in older adults.

Trunk Stability and Gait-Related Outcomes: Absence of Association

As expected, no association was observed between trunk stability and gait-related outcomes. However, the third hypothesis was partially rejected, as the anticipated correlation between the unstable sitting posturographic test and the modified Timed Up and Go test was not observed. As a previous study on physically active older adults showed [38], trunk stability does not seem to play a significant role in gait performance or even in a gait-related task, such as the Timed Up and Go test, which requires standing, sitting, and turning actions in this population. A previous study in minimally and moderately impaired people with multiple sclerosis [26] found an association with the Timed Up and Go test but not with a straight-line gait test, which is justified by the greater relevance of trunk control needed for its performance. Therefore, in populations with neuromuscular deficits, voluntary trunk control may have a more relevant role during the sit-to-stand, turn, and stand-to-sit phases of the Timed Up and Go test. Again, this interpretation should be taken with caution, as another study involving individuals with greater disability, such as those with hemiparesis following a stroke [40], found no association between an unstable sitting posturographic test and the Timed Up and Go test.

The seemingly counterintuitive inverse association between whole-body static balance (higher acceleration indicating worse control) and walking speed (straight-line gait and Timed Up and Go) might be influenced by anthropometric factors, such as participant height. Taller individuals, while potentially achieving higher walking speeds due to longer strides, often face greater challenges in maintaining static postural control due to a higher center of mass and increased pendulum-like dynamics [41]. This suggests a complex interplay wherein static balance demands differ from dynamic mobility requirements, particularly during phases like sit-to-stand, turning, and stand-to-sit in the Timed Up and Go test [42].

XSL•FO
RenderX

Discrepancy Between Field and Laboratory Assessments

The lack of correlation between tests grouped under the same capability (ie, trunk stability and whole-body balance) reinforces the notion of the multidimensional nature of postural control in older adults, as previously documented in studies involving other populations [5,23]. Therefore, sport science professionals and practitioners should be aware that the lack of correlations between the novel inertial-sensor–based field tools (CoreMaker and G&B apps) and the well-established force-platform-based assessments makes it unfeasible to extrapolate research results found through laboratory tests to the field. Despite the lack of relationship between laboratory and field tools, the technologically based trunk stability and balance assessment through field tools such as CoreMaker and G&B apps will enhance experimental studies, allowing postural control during training exercises to be assessed in larger, more representative samples, leading to more meaningful clinical findings [21]. In addition, measuring trunk stability during the exercises used for training may allow us to determine to what extent the potential benefits caused by trunk-focused exercise programs (bridges and bird-dogs [24,25]) on balance and gait may be due to an improvement in trunk stability.

Study Limitations

As with any research, it is important to highlight some limitations. A much larger sample would be desirable in the most difficult variations of bridges (Table S2 in Multimedia Appendix 1) to minimize the probability of a type II error due to undersampling. Second, our results are based on a sample of physically active older adults and of both sexes, so future studies would need to examine the relationships of these tests in more fragile populations and also potential sex differences. Although our study focused on well-established force platform and inertial sensor protocols to assess trunk stability, balance, and gait, it is important to admit that other assessment tools exist and that using different tests might have led to different results. Future studies should analyze and develop field tests to evaluate the reactive control of the trunk due to its apparent potential relationship with the mechanisms required during a fall situation in older people. Moreover, trunk-focused stability outcomes should also be analyzed during specific tests designed to quantify whole-body balance, gait, and functional mobility, particularly under conditions imposing high demands on trunk control, to better understand how performance in these tasks relates to direct measurements of trunk involvement [42,43]. Although for the accelerometry-based field tests postural control was summarized using the average of the acceleration magnitude data series obtained from the 3-axis as a proxy for performance, future studies with more focused designs should incorporate axis-specific and advanced signal analyses to further elucidate control mechanisms during the studied tasks. Although the main conclusions were robust across multiple analytical strategies and the qualitative interpretation of the results remained similar, as with any complex analytical framework, the choice of data reduction strategy may have had a modest influence on the magnitude of some associations. Importantly, the convergence of findings across these contrasting analytical strategies strengthens confidence in the robustness, task-specificity, and interpretability of the observed relationships, indicating that the primary analysis represents a conservative estimate rather than a biased one.

Conclusions

This study provides novel evidence linking lumbopelvic control in floor-based exercises (mainly back bridge and bird-dog) to whole-body static balance in physically active older adults, supporting trunk stability as a crucial factor for maintaining steady standing positions in physically active older adults. Crucially, this research highlights a notable disconnection between novel inertial sensor–based field assessments (CoreMaker and G&B apps) and traditional force platform–based laboratory protocols for trunk stability and balance. While direct extrapolation between these methods is not advisable, the accessibility of field tools like CoreMaker offers a promising avenue [21] for standardized lumbopelvic stability assessment and the monitoring of training load in broader populations. Future interventions targeting static balance in older adults should consider incorporating specific bridging and bird-dog positions [24,25] to enhance trunk control.

Acknowledgments

The authors would like to thank all the study participants for their support. No generative AI tools were used in the preparation of this manuscript.

Funding

The project I+D+I PID2022-140323OB-100 was funded by MICIU/AEI/10.13039/501100011033 and by European Regional Development Fund, European Union.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

JDLR-C contributed to the literature review, data collection, data analysis and interpretation, and drafting of the manuscript. CJ-R was involved in data collection, analysis, and interpretation, as well as reviewing and editing the manuscript draft. AP-L was involved in data collection, analysis, and interpretation, as well as reviewing and editing the manuscript draft. AM-I was involved

XSL•FO
RenderX

in data collection, analysis, and interpretation, as well as reviewing and editing the manuscript draft. JLLE contributed to software development, including program design, code implementation, and testing. DB supported the research concept and study design, contributed to programming and software development, and participated in reviewing and editing the manuscript draft. FJV-G contributed to the research concept and study design and participated in reviewing and editing the manuscript draft. All authors read and approved the final version of the manuscript and agreed with the order of submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Clinical data, reliability, associations, and sensitivity analyses.
[DOCX File, 426 KB - aging_v9i1e83546_app1.docx]

Checklist 1
STROBE checklist.
[PDF File, 198 KB - aging_v9i1e83546_app2.pdf]

References

1. Seidler RD, Bernard JA, Burutolu TB, et al. Motor control and aging: links to age-related brain structural, functional, and biochemical effects. Neurosci Biobehav Rev 2010 Apr;34(5):721-733. [doi: 10.1016/j.neubiorev.2009.10.005] [Medline: 19850077]
2. Haddas R, Satin A, Lieberman I. What is actually happening inside the “cone of economy”: compensatory mechanisms during a dynamic balance test. Eur Spine J 2020 Sep;29(9):2319-2328. [doi: 10.1007/s00586-020-06411-w] [Medline: 32306304]
3. Quirk DA, Hubley-Kozey CL. Age-related changes in trunk neuromuscular activation patterns during a controlled functional transfer task include amplitude and temporal synergies. Hum Mov Sci 2014 Dec;38:262-280. [doi: 10.1016/j.humov.2014.08.013] [Medline: 25457424]
4. McGill A, Houston S, Lee RYW. Effects of a ballet intervention on trunk coordination and range of motion during gait in people with Parkinson’s. In: Gao Y, editor. Cogent Med 2019 Jan 1;6(1):1583085. [doi: 10.1080/2331205X.2019.1583085]
5. Vera-Garcia FJ, López-Plaza D, Juan-Recio C, Barbado D. Tests to measure core stability in laboratory and field settings: reliability and correlation analyses. J Appl Biomech 2019 Jun 1;35(3):223-231. [doi: 10.1123/jab.2018-0407] [Medline: 30860424]
6. Saito K, Matsunaga T, Iwami T, Shimada Y. Evaluation of trunk stability in the sitting position using a new device. Biomed Res 2014;35(2):127-131. [doi: 10.2220/biomedres.35.127] [Medline: 24759180]
7. Akizuki K, Echizenya Y, Kaneno T, Ohashi Y. Usefulness of an unstable board balance test to accurately identify community-dwelling elderly individuals with a history of falls. J Rehabil Med 2019 Jan 1;51(1):71-76. [doi: 10.2340/16501977-2504] [Medline: 30406266]
8. Monteiro PHM, Valenciano PJ, Mendes PHS, Teixeira LA. Association of 30-s sit-to-stand power test outcome with body balance in physically active older adults. J Aging Phys Act 2025 Aug 1;33(4):370-378. [doi: 10.1123/japa.2023-0373] [Medline: 39753121]
9. Valenciano PJ, Castan VE, Monteiro PHM, Teixeira LA. Symmetric unipedal balance in quiet stance and dynamic tasks in older individuals. Brain Res 2024 May 1;1830:148850. [doi: 10.1016/j.brainres.2024.148850] [Medline: 38460718]
10. Valenciano PJ, Monteiro PHM, Lazzaro IM, et al. Validation of the Equidyn protocol for evaluation of dynamic balance in older adults through a smartphone application. Gait Posture 2024 Jun;111:59-64. [doi: 10.1016/j.gaitpost.2024.04.004] [Medline: 38643633]
11. Granacher U, Gollhofer A, Hortobágyi T, Kressig RW, Muehlbauer T. The importance of trunk muscle strength for balance, functional performance, and fall prevention in seniors: a systematic review. Sports Med 2013 Jul;43(7):627-641. [doi: 10.1007/s40279-013-0041-1] [Medline: 23568373]
12. Granacher U, Lacroix A, Roettger K, Gollhofer A, Muehlbauer T. Relationships between trunk muscle strength, spinal mobility, and balance performance in older adults. J Aging Phys Act 2014 Oct;22(4):490-498. [doi: 10.1123/japa.2013-0108] [Medline: 24152575]
13. Suri P, Kiely DK, Leveille SG, Frontera WR, Bean JF. Trunk muscle attributes are associated with balance and mobility in older adults: a pilot study. PM R 2009 Oct;1(10):916-924. [doi: 10.1016/j.pmrj.2009.09.009] [Medline: 19854420]
14. Stotz A, Mason J, Groll A, Zech A. Which trunk muscle parameter is the best predictor for physical function in older adults? Heliyon 2023 Oct;9(10):e20123. [doi: 10.1016/j.heliyon.2023.e20123] [Medline: 37822629]
15. Acar E, Çankaya T, Öner S. The relationship between trunk muscle thickness and static postural balance in older adults. J Aging Phys Act 2020 Apr 24;28(2):269-275. [doi: 10.1123/japa.2019-0127] [Medline: 31722293]

XSL•FO
RenderX

16. Shahtahmassebi B, Hebert JJ, Hecimovich MD, Fairchild TJ. Associations between trunk muscle morphology, strength and function in older adults. Sci Rep 2017 Sep 7;7(1):10907. [doi: 10.1038/s41598-017-11116-0] [Medline: 28883555]
17. Zhai M, Huang Y, Zhou S, et al. Effects of age-related changes in trunk and lower limb range of motion on gait. BMC Musculoskelet Disord 2023 Mar 28;24(1):234. [doi: 10.1186/s12891-023-06301-4] [Medline: 36978129]
18. Haddas R, Hu X, Lieberman IH. The correlation of spinopelvic parameters with biomechanical parameters measured by gait and balance analyses in patients with adult degenerative scoliosis. Clin Spine Surg 2020 Feb;33(1):E33-E39. [doi: 10.1097/BSD.0000000000000939] [Medline: 31996610]
19. Takahashi Y, Saito K, Matsunaga T, et al. Relationship between dynamic trunk balance and the Balance Evaluation Systems Test in elderly women. Prog Rehabil Med 2020;5:20200004. [doi: 10.2490/prm.20200004] [Medline: 32789272]
20. Lee KYT, Hui-Chan CWY, Tsang WWN. Reliability and validity of the sequential weight-shifting test: a new functional approach to the assessment of the sitting balance of older adults. J Phys Ther Sci 2016 Dec;28(12):3444-3450. [doi: 10.1589/jpts.28.3444] [Medline: 28174470]
21. Bohlke K, Redfern MS, Rosso AL, Sejdic E. Accelerometry applications and methods to assess standing balance in older adults and mobility-limited patient populations: a narrative review. Aging Clin Exp Res 2023 Oct;35(10):1991-2007. [doi: 10.1007/s40520-023-02503-x] [Medline: 37526887]
22. Heredia-Elvar JR, Juan-Recio C, Prat-Luri A, Barbado D, de los Ríos-Calonge J, Vera-Garcia FJ. Exercise intensity progressions and criteria to prescribe core stability exercises in young physically active men: a smartphone accelerometer-based study. J Strength Cond Res 2024;38(2):266-273. [doi: 10.1519/JSC.0000000000004630]
23. De Los Ríos‐Calonge J, Barbado D, Prat‐Luri A, et al. Are trunk stability and endurance determinant factors for whole‐body dynamic balance in physically active young males? A multidimensional analysis. Scandinavian J Med Sci Sports 2024 Mar;34(3). [doi: 10.1111/sms.14588]
24. Granacher U, Lacroix A, Muehlbauer T, Roettger K, Gollhofer A. Effects of core instability strength training on trunk muscle strength, spinal mobility, dynamic balance and functional mobility in older adults. Gerontology 2013;59(2):105-113. [doi: 10.1159/000343152] [Medline: 23108436]
25. Lima M, Silva B, Rocha-Rodrigues S, Bezerra P. The impact of an 8-week pilates-based physical training program on functional mobility: data from a septuagenarian group. Biomed Hum Kinet 2021 Jan 1;13(1):11-19. [doi: 10.2478/bhk-2021-0002]
26. Barbado D, Gomez-Illan R, Moreno-Navarro P, Valero-Conesa G, Reina R, Vera-Garcia FJ. Postural control quantification in minimally and moderately impaired persons with multiple sclerosis: the reliability of a posturographic test and its relationships with functional ability. J Sport Health Sci 2020 Dec;9(6):677-684. [doi: 10.1016/j.jshs.2018.06.008] [Medline: 33308819]
27. Rashid U, Barbado D, Olsen S, et al. Validity and reliability of a smartphone app for gait and balance assessment. Sensors (Basel) 2021 Dec 25;22(1):124. [doi: 10.3390/s22010124] [Medline: 35009667]
28. Christopher A, Kraft E, Olenick H, Kiesling R, Doty A. The reliability and validity of the Timed Up and Go as a clinical tool in individuals with and without disabilities across a lifespan: a systematic review. Disabil Rehabil 2021 Jun;43(13):1799-1813. [doi: 10.1080/09638288.2019.1682066] [Medline: 31656104]
29. Winter DA. Biomechanics and Motor Control of Human Movement: John Wiley & Sons; 2009.
30. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
31. Hopkins WG. Spreadsheets for Analysis of Validity and Reliability. Sportscience 2015;19:36-42 [FREE Full text]
32. Hinkle DE, Wiersma W, Jurs SG. Applied Statistics for the Behavioral Sciences: Houghton Mifflin; 2003.
33. Kiss R, Schedler S, Muehlbauer T. Associations between types of balance performance in healthy individuals across the lifespan: a systematic review and meta-analysis. Front Physiol 2018;9(1366):1366. [doi: 10.3389/fphys.2018.01366] [Medline: 30323769]
34. Shumway-Cook A, Woollacott MH. Motor Control: Translating Research Into Clinical Practice: Lippincott Williams & Wilkins; 2007.
35. Haruyama K, Kasai K, Makino R, Hoshi F, Nishihara K. Quantification of trunk segmental coordination and head stability in laterally unstable sitting identifies aging and cerebellar ataxia. Clin Biomech 2019 Mar;63:127-133. [doi: 10.1016/j.clinbiomech.2019.01.013]
36. Juan-Recio C, Prat-Luri A, Galindo A, Manresa-Rocamora A, Barbado D, Vera-Garcia FJ. Is the side bridge test valid and reliable for assessing trunk lateral flexor endurance in recreational female athletes? Biology (Basel) 2022 Jul 12;11(7):1043. [doi: 10.3390/biology11071043] [Medline: 36101422]
37. Minetto MA, Giannini A, McConnell R, Busso C, Torre G, Massazza G. Common musculoskeletal disorders in the elderly: the Star triad. J Clin Med 2020 Apr 23;9(4):1216. [doi: 10.3390/jcm9041216] [Medline: 32340331]
38. Hernández-Sánchez S, De Los Ríos-Calonge J, Juan-Recio C, Prat-Luri A, Vera-Garcia FJ, Barbado D. Understanding the effects of aging on whole-body dynamic balance, trunk stability, functional mobility, and hip strength. Sci Rep 2026 Jan 20;16(1):5828. [doi: 10.1038/s41598-026-35724-x] [Medline: 41559158]
39. Chiou SY, Gottardi SEA, Hodges PW, Strutton PH. Corticospinal excitability of trunk muscles during different postural tasks. PLoS ONE 2016;11(1):e0147650. [doi: 10.1371/journal.pone.0147650] [Medline: 26807583]

XSL•FO
RenderX

40. Bruyneel AV, Mesure S, Reinmann A, et al. Validity and reliability of center of pressure measures to quantify trunk control ability in individuals after stroke in subacute phase during unstable sitting test. Heliyon 2022 Oct;8(10):e10891. [doi: 10.1016/j.heliyon.2022.e10891] [Medline: 36237978]
41. Alonso AC, Mochizuki L, Silva Luna NM, Ayama S, Canonica AC, Greve JMDA. Relation between the sensory and anthropometric variables in the quiet standing postural control: is the inverted pendulum important for the static balance control? Biomed Res Int 2015;2015:985312. [doi: 10.1155/2015/985312] [Medline: 26539550]
42. Rezende LS, Monteiro PH, Oliveira JA, et al. Do Timed Up and Go and five times sit to stand test outcomes correlate with trunk stability? A pilot-study. Braz J Motor Behav 2023;17(4):150-157. [doi: 10.20338/bjmb.v17i3.358]
43. Naito D, Honda K, Sekiguchi Y, Izumi SI, Ebihara S. Characteristics of trunk acceleration and angular velocity in turning movement in post-stroke patients with high risk of falling. Sensors (Basel) 2025 Apr 24;25(9):2689. [doi: 10.3390/s25092689] [Medline: 40363128]

Abbreviations

ICC: intraclass correlation coefficient
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TE: typical error

Edited by M Gray; submitted 04.Sep.2025; peer-reviewed by K Fujita, PHM Monteiro; revised version received 03.Feb.2026; accepted 11.Mar.2026; published 06.May.2026.

Please cite as:
De Los Ríos-Calonge J, Juan-Recio C, Prat-Luri A, Miralles-Iborra A, Elvira JLL, Barbado D, Vera-Garcia FJ
Smartphone-Based Field Assessment of Trunk Stability and Its Relationship With Whole-Body Balance in Older Adults: Cross-Sectional Study
JMIR Aging 2026;9:e83546
URL: https://aging.jmir.org/2026/1/e83546
doi:10.2196/83546

© Javier De Los Ríos-Calonge, Casto Juan-Recio, Amaya Prat-Luri, Aarón Miralles-Iborra, Jose LL Elvira, David Barbado, Francisco Jose Vera-Garcia. Originally published in JMIR Aging (https://aging.jmir.org), 6.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Feasibility and Acceptability of a Remote Sleep–Dependent Memory Assessment in Older Adults With Cognitive Concerns: Pilot Cross-Sectional Study

Clara Cher Yi Tan[1,2], MSc; Aaron Lam[1,2,3], PhD; David Ireland[4], PhD; Dana Bradford[4], PhD; Nathan Cross[1,2], PhD; Melvern Ted Kurniawan[1,2], BPsych(Hons); Simone Simonetti[1,2], PhD; Jurgen Fripp[4], PhD; Sharon L Naismith[1,2], PhD

[1]Healthy Brain Ageing Program, The Brain and Mind Centre, The University of Sydney, 100 Mallett Street, Camperdown, NSW, Australia

[2]School of Psychology, The University of Sydney, Camperdown, NSW, Australia

[3]Woolcock Institute of Medical Research, Macquarie University, Macquarie Park, NSW, Australia

[4]Australian eHealth Research Centre, Commonwealth Scientific and Industrial Research Organisation, Herston, Australia

Corresponding Author:
Clara Cher Yi Tan, MSc
Healthy Brain Ageing Program, The Brain and Mind Centre, The University of Sydney, 100 Mallett Street, Camperdown, NSW, Australia

Abstract

Background: Sleep-dependent memory consolidation (SDMC), the process by which sleep supports the transfer of memories into long-term storage, declines with age but remains underexplored in older adults with subjective cognitive decline and mild cognitive impairment. Traditional SDMC assessments are typically conducted in lab settings, with limited evidence for feasibility to do these assessments at home for this clinical population.

Objective: To address this, we co-designed the *Sleep Memories* app, which assesses overnight memory for a previously validated 32-item word-pairs task. This pilot study first explored the feasibility and acceptability of the app, including willingness to participate. Second, we explored various demographic, clinical, and subjective sleep factors associated with task completion and SDMC performance in a memory clinic sample.

Methods: Within an 8-month period, we invited 141 older adults aged 50 years and above (mean 71.27, SD 7.51 y) from the Healthy Brain Ageing clinic, a specialist brain health and memory clinic in Sydney, to pilot the *Sleep Memories* app. Of these, 76 (mean 70.19, SD 7.75) agreed to participate. All participants underwent a full neuropsychological test battery, medical assessment, and mood assessment.

Results: Sixty-eight participants completed at least 1 test of the word-pairs task. The word-pairs task completion rate for all trials was over 50%. There was 57% (39/68) completion of both evening and morning delayed recall tests. Lower willingness to participate was associated with lower global cognitive scores, poorer sleep behavior, and clinical factors. Higher task completion was associated with greater education and greater anxiety levels. User feedback indicated that the app was well-accepted and liked, although some participants reported minor technical difficulties.

Conclusions: These findings support the feasibility of mobile app–based SDMC assessment in older adults at risk of cognitive decline and underscore the importance of considering individual characteristics (eg, subjective sleep factors, clinical characteristics, and education) when designing digital SDMC tools.

(JMIR Aging 2026;9:e87926) doi:10.2196/87926

KEYWORDS

sleep-dependent memory consolidation; digital cognitive assessment; aging; cognition; sleep

Introduction

Sleep is a fundamental biological process that is essential for both physical and cognitive health [1,2]. In particular, sleep plays an important role in sleep-dependent memory consolidation (SDMC). This process involves the reactivation and replay of temporary memories during sleep, which stabilizes them into long-term storage and facilitates their transfer from the hippocampus to the neocortex [3,4]. These features are thought to support memory consolidation; however, due to methodological challenges, SDMC remains underexamined in older adults, particularly those with subjective or mild cognitive impairment (MCI). Only a few studies have explicitly investigated SDMC in these populations, albeit in small sample sizes [5-8]. These scalability barriers have limited the field's

XSL•FO
RenderX

understanding of how factors such as modifiable dementia risk factors and global cognitive impairment influence SDMC.

Several factors may account for this gap in the literature, which currently constrains research on SDMC in populations at risk of dementia. For example, traditional SDMC assessment tasks were designed for younger or cognitively intact adults. Therefore, they are not suited for older adults with cognitive impairments, jeopardizing feasibility and producing floor effects. Furthermore, SDMC tasks have typically required assessment in the sleep laboratory before and after sleep [9-11]. The laboratory environment may not reflect naturalistic sleep conditions [12], limiting both the ecological validity and scalability of SDMC assessment for research and clinical use. While this setting offers a controlled environment, the requirement to stay overnight in the laboratory deems data collection resource-intensive, costly, and a limiting factor for larger-scale studies.

Digitalization of SDMC assessments offers an opportunity to address these limitations by enabling participants to complete memory tasks remotely and in their habitual sleeping environment. However, the feasibility of transiting from the laboratory to home assessments for SDMC tasks and willingness to participate in the research need to be tested. Both feasibility and willingness to participate may be influenced by demographics, clinical, and sleep-related factors as well as barriers such as technical issues, psychological concerns, and workload demands [13]. These considerations are especially important in memory clinic settings and specialist services that assess and support individuals with cognitive concerns, where digital tools are increasingly being considered to support assessment and intervention [14,15].

Recognizing the need for feasible SDMC tasks for older people with cognitive concerns, we initially developed a 32-item word-pairs task [5] in the sleep laboratory and demonstrated its use in older adults, including those with MCI. We then co-designed a mobile app named *Sleep Memories* [12,16] that delivers the 32-item task remotely, enabling at-home unsupervised assessment of SDMC. This research tool allows participants to complete the SDMC task at their convenience in their natural sleep environment, increases accessibility for those unable to attend the laboratory, and reduces research costs by eliminating the need for in-lab supervision. In this pilot study, the primary aim was to evaluate *willingness to participate* and the *feasibility* of implementing the *Sleep Memories* app in a memory clinic setting. Secondarily, we explored various demographic, clinical, and subjective sleep-related factors that may influence both willingness to participate and feasibility of the app. Furthermore, a subset of participants was asked for *user experiences and barriers* related to using the app for assessing the *acceptability of the app* in a memory clinic setting. Next, we explored the associations of demographics, clinical, subjective sleep-related factors, modifiable dementia risk factors, and global cognitive functioning with SDMC performance. Finally, we further explored differences in characteristics between low and high SDMC performers and compared the characteristics of participants who provided user feedback with the main sample. Feasibility was defined, in accordance with prior research, as the completion rate above 50% [17] and was hypothesized to be achievable in our study. We also hypothesized that willingness to participate would depend on demographics, clinical, and subjective sleep-related factors; specifically, it would be negatively associated with higher levels of cognitive impairment and lower global cognitive functioning, as well as poorer sleep behaviors.

Methods

Participants Recruitment

Between April and December 2024, we approached all consecutive participants attending the Healthy Brain Ageing clinic at the Brain and Mind Centre, University of Sydney, for participation. This clinic is a specialist memory and cognition clinic for older adults (aged 50 y or older) with new-onset cognitive or mood concerns. As previously described [18], the exclusion criteria for the clinic were: insufficient English to complete neuropsychological assessment, Mini-Mental State Examination (MMSE) score <20 [19], current substance dependence, intellectual disability, presence of nonaffective psychiatric disorder (eg, attention-deficit/hyperactivity disorder and schizophrenia), presence of neurological disease (eg, Parkinson and epilepsy), history of head injury with loss of consciousness >30 minutes, and history of stroke. Additional exclusion criteria for this study were the absence of a mobile phone and a diagnosis of dementia. Four participants were excluded for not having a mobile phone.

Procedure

Overview

All participants underwent comprehensive and standardized neuropsychological, medical, and mood assessments conducted by a neuropsychologist, a geriatrician or neurologist, and a research psychologist, respectively. The *Sleep Memories* app was introduced to participants during their mood assessment. Those who expressed interest and provided informed consent were assisted by a researcher in installing the app on their mobile phones. Participants were also given written instructions on its use and subsequently completed memory tasks at-home in the evening and the following morning. Additionally, participants received text reminders to complete the task in the evening. Given that the sleep benefits for sleep-dependent consolidation of verbal memory are well established [3,4], and as this study was designed to assess feasibility rather than to test sleep-dependent memory effects, we did not include a control condition (eg, a wake retention interval).

Neuropsychological Assessment

A standardized battery of tests was administered by a neuropsychologist, covering a range of domains including premorbid IQ, attention, processing speed, working memory, verbal learning and memory, language, visuospatial and executive functions, and general cognitive functioning. The neuropsychological scores were adjusted for each participant's age and level of education using normative data. For this study, we reported only the premorbid IQ, as assessed by the Wechsler Test of Adult Reading [20], global cognition, as assessed by the MMSE [19], and verbal learning (trials 1 - 5) and memory

XSL•FO
RenderX

(trial 7), as assessed by the Rey Auditory Verbal Learning Test [21].

Participants who were referred to the clinic by their general practitioner for the assessment of new-onset cognitive or mood concerns were classified as either subjective cognitive decline (SCD) or MCI based on standardized criteria through a consensus meeting involving a medical specialist and 2 neuropsychologists. Participants were classified as having MCI when they scored at least 1.50 SDs below normative values for their age and education on one or more neuropsychological tests, while overall functional independence remained intact. Participants meeting criteria for MCI were further subtyped based on the domain of cognitive impairment [22]. Those with deficits primarily in the memory domain were classified as having amnestic MCI, while those with impairments in nonmemory domains (eg, language, executive function, and visuospatial skills) were classified as having nonamnestic MCI. Participants were classified as having SCD if they self-reported a persistent decline in cognitive capacity compared to a previously normal status, in the absence of objective cognitive impairment on standardized neuropsychological testing. The decline was not caused by an acute event, and was not explained by MCI, prodromal Alzheimer disease, or other psychiatric, neurological, medical, or substance-related causes [23,24].

Medical Assessment

A medical assessment was conducted by a geriatrician or a neurologist who recorded the medical history and rated the medical burden using the Cumulative Illness Rating Scale-Geriatric [25]. Each system is scored from 0 to 4, and the total score of Cumulative Illness Rating Scale-Geriatric ranges from 0 to 52, with higher scores indicating greater illness severity and comorbidity.

Mood Assessment

A research psychologist assessed lifetime and current mood and anxiety symptoms using the Mini International Neuropsychiatric Interview [26]. They also used the Cogsleep Semi-structured Interview Revised Version [27] to probe for various sleep concerns. For this study, we reported the data on bedtime, wake-up time, and frequency (days) of trouble waking up in the morning from the semistructured interview, as individual sleep patterns and morning alertness might influence willingness to participate and task completion [28]. In addition, self-reported sleep quality was assessed using the Pittsburgh Sleep Quality Index [29], with higher global scores indicating poorer sleep quality. Insomnia symptoms were assessed using the Insomnia Severity Index [30], with higher scores indicating greater severity of insomnia symptoms. For depression and anxiety symptoms, the Geriatric Depression Scale-15 [31] and Geriatric Anxiety Inventory [32] were administered, respectively, with higher scores indicating greater levels of depression and anxiety.

Sleep Memories App

Development

As previously described [12,16], the *Sleep Memories* mobile app was designed to deliver the 32-item word-pairs task via a chatbot named “Aurora.” Aurora used a low-cost internal chatbot engine that ran on the user’s device and required no Internet access. The engine was developed at Commonwealth Scientific and Industrial Research Organisation (CSIRO) and has been used extensively in other health applications [33,34] requiring ethical and safe operation. The app was co-designed by the authors and Healthy Brain Ageing researchers who provided the word-pairs task list and conducted focus groups and CSIRO software engineers who created and tested the chatbot architecture. They gathered inputs on the app’s functionality and accessibility from CSIRO staff (who tested it on their devices) [16] and older adults (4/11 were male, mean age 68.50 [SD 5.10] years; 8 had SCD and 3 had MCI) from the Healthy Brain Ageing clinic (focus group discussion) [12].

Evening Procedure

Figure 1 shows the time course and details of the app setup and completion. Upon starting the app, the chatbot collected basic demographic information, including age, gender, sleep time, and wake-up time from participants. Participants were then provided with a trial demonstration run of 3 word pairs to become familiar with the task and the app’s functionality. Participants were verbally instructed to complete the following tasks at home in a quiet, distraction-free environment and to refrain from using external aids.

XSL•FO
RenderX

Figure 1. *Sleep Memories* app task flow. The figure was modified from the past study [5].

XSL•FO
RenderX

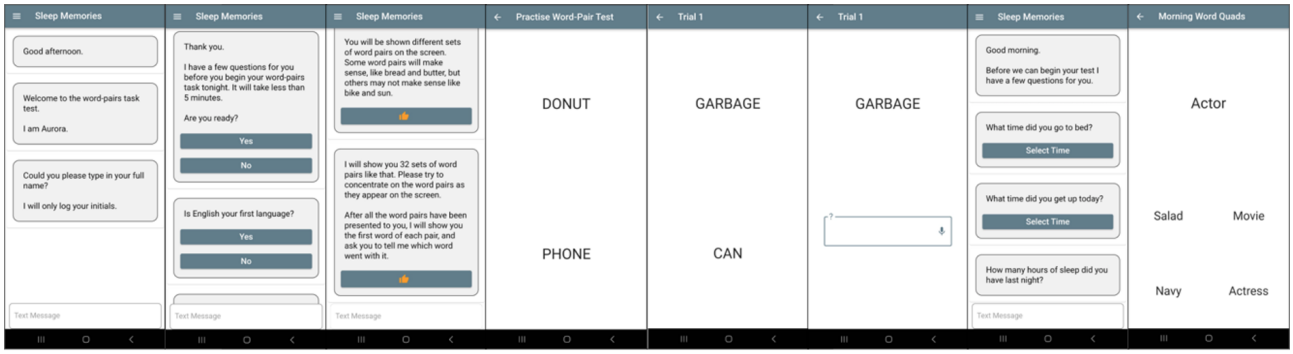
After the demonstration, the chatbot informed participants to begin the word-pairs task within the 3-hour period preceding their reported sleep time. At the scheduled time, the participants received a notification to begin. They initially completed 4 learning trials. Each learning trial entailed a random presentation of 16 semantically related pairs (eg, syrup-maple) and 16 semantically unrelated word pairs (eg, groom-minimum). Each pair was presented for 10 seconds. At the end of each trial (ie, after all the word pairs had been presented), the participants were presented with the first word of each pair (eg, syrup) and were asked to provide the matching word (eg, maple) of each pair either by text or an audio response. The participants had 15 seconds to respond for each word pair. This was repeated for all 4 trials.

Following a 30-minute break, the participants were asked again to recall the word pairs during the evening recall test. The first word of each pair was presented for 15 seconds, and participants were asked to recall the matched word that had been paired with it. The entire evening session, excluding the break, lasted approximately 45 minutes. Participants were then able to sleep as usual.

Morning Procedure

Approximately within an hour after sleep offset, based on the wake-up time set in the app, the chatbot prompted participants to answer a few questions about their sleep and to confirm their sleep time and wake time. This was followed by two tests: (1) a morning delayed recall test whereby the first word of each pair was shown, and participants had to type or verbalize the matching word pair; and (2) a morning multiple-choice recognition test, whereby the first word of each pair was shown, along with 1 correct answer and 3 incorrect answers, and participants were asked to choose the correct matching word pair. To encourage task completion, the *Sleep Memories* app sent 4 notifications reminding participants to complete the tasks (see Figure 2 for screenshots of the *Sleep Memories* app). *Sleep Memories* logged all data on a Google Firebase service, with the server located within Australia.

Figure 2. Screenshots of the *Sleep Memories* app. (A) The Aurora chatbot introducing themselves to the participant. (B) The app probing for demographic questions, sleep time, and wake time before learning trials. (C) The app providing instructions for the word-pairs task. (D) Trial run of the word-pairs task. (E) Learning component of the word-pairs task. (F) Recall component of the word-pairs task. (G) The app requesting details of the participants' sleep before the morning tasks. (H) Multiple-choice testing component of the word pairs. The figure was modified from the past study [5].



A B C D E F G H

Participants reported their estimated sleep-onset and wake-up times when setting up the app in the evening of testing, and these were confirmed again upon waking. In addition, the app automatically recorded the timing of the encoding and recall sessions, allowing us to derive the intervals between them.

At the data analysis stage, the SDMC score (% retention) was calculated by dividing the morning delayed recall test score by the evening recall test score and then multiplying by 100 (the evening recall test measured memory retention 30 min after learning the word pairs; the morning delayed recall test assessed memory retention after a night of sleep, approximately within an hour after waking).

Follow-Up

Within 2 weeks of task completion, a subset of participants was contacted to provide feedback on their user experiences (they were selected based on availability and successful contact). Researchers conducted phone interviews using a semistructured questionnaire as a guide (see Multimedia Appendix 1). During these interviews, the researchers summarized participants' verbal responses, rather than recording or transcribing their verbatim as we did not have ethical approval for voice recording. Thus, verbatim transcription was not available. Importantly, the researchers involved in the questionnaire development and the qualitative analysis were independent of the *Sleep Memories* app development.

Feasibility and Acceptability of the *Sleep Memories* App

The feasibility of the *Sleep Memories* app was operationally defined as a completion rate of over 50% across all tasks, including evening learning trials, evening recall test, morning delayed recall test, morning multiple-choice test, and both evening recall and morning delayed recall tests [17]. This threshold was set based on prior studies of app-based cognitive assessments in older populations [35-37], reflecting the challenges of unsupervised digital cognitive assessment in older adults. It aligns with the Pragmatic-Explanatory Continuum Indicator Summary framework, which emphasizes pragmatic feasibility over idealized adherence [38]. While a scoping review of mobile-based cognitive assessments in healthy older adults reported completion rates around 60% [39], we set a slightly lower threshold of 50% as our sample included older adults with cognitive concerns who may face greater barriers to engagement with digital cognitive assessment [40].

The acceptability of the app was assessed qualitatively using a semistructured questionnaire (see Multimedia Appendix 1)

XSL•FO
RenderX

developed through a bottom-up approach based on findings from prior focus group research [12]. While the Theoretical Framework of Acceptability [41] was not used as a blueprint during the initial design phase, the questionnaire sections align with the theory. Specifically, it captured: Intervention Coherence (sections 1 and 2 on app setup, layout, and clarity of instructions), Self-efficacy (sections 3 and 4 on the SDMC task and feasibility), Affective Attitude (section 5 on subjective feelings toward the app), and Burden (any reported task demands and obstacles). Furthermore, the questionnaire was reviewed by an experienced neuropsychologist and a researcher specializing in qualitative methodologies.

Modifiable Dementia Risk Factors

The Lifestyle for Brain Health (LIBRA) index is a composite measure of 12 modifiable dementia risk factors, developed through a systematic review and expert consensus using the Delphi method [42]. The LIBRA 2 index is an updated version that incorporates additional factors, including hearing impairment, social contact, and sleep disturbances [42]. In this study, we used the validated LIBRA2 index to calculate a composite risk score by summing the available factor scores for each participant, with a theoretical range of 0 to 100. Positive weights were assigned when risk factors were present or protective factors were absent [42] (Table 1), with higher scores indicating greater dementia risk. In our sample, data were available for 13 of the 15 LIBRA2 factors, including high physical activity, sleep disturbances, high alcohol consumption, hearing impairment, kidney disease, coronary heart disease, diabetes, high cholesterol, smoking, hypertension, obesity, depression, and high cognitive activity. Data for healthy diet and low social contact were not available. Following previously published methodology [43], when less than 4 factors were missing, scores were calculated using the available factors. Four participants had 4 or more factors missing; therefore, these participants were excluded from LIBRA 2 analyses.

Table . Weights of the modifiable dementia risk factors included in Lifestyle for Brain Health (LIBRA) 2 Index[a].

Factors	Type of factors	Measurements	Weight assigned
High physical activity	Protective	At least 150 minutes of activities and at least 5 sessions of activity over 1 week	6
Sleep disturbances	Risk	PSQI[b] global score ≥5	3.3
High alcohol consumption	Risk	AUDIT-C[c] total score (≥3 for women and ≥4 for men)	3.1
Hearing impairment	Risk	Self-report (yes)	7.6
Kidney disease	Risk	Self-report (yes)	5.7
Coronary heart disease	Risk	Self-report (yes)	8.3
Diabetes	Risk	Self-report (yes)	6.8
High cholesterol	Risk	Self-report (yes)	8.2
Smoking	Risk	Self-report (current smoking)	7.9
Hypertension	Risk	Self-report (yes)	3.5
Obesity	Risk	BMI ≥30	7
Depression	Risk	GDS-15[d] total score ≥6	13
High cognitive activity	Protective	Highest tertile of CSA[e] total score	9.4
Healthy diet	Protective	—[f]	—
Low social contact	Risk	—	—

[a]Positive weights were allocated whenever the risk factors were present, or the protective factors were absent.

[b]PSQI: Pittsburgh Sleep Quality Index.

[c]AUDIT-C: Alcohol Use Disorders Identification Test.

[d]GDS-15: Geriatric Depression Scale-15.

[e]CSA: Cognitively Stimulating Activities Questionnaire.

[f]Not available.

Participant Group Definitions

We categorized participants into several groups. First, those who agreed versus those who declined to use the app. Second, those who completed all recall trials (full completers) versus those who did not complete 1 or both of the following: evening recall test and morning delayed recall test (partial completers). For additional exploratory data analysis, the median score for SDMC % retention was 92.40% (IQR 84.31-100.00). From this, we categorized participants as either low (n=18, <92.40) or high (n=18, ≥92.40) SDMC performers.

XSL•FO
RenderX

Ethical Considerations

This study was approved by the University of Sydney Human Ethics Committee (2019/HE000271) and conducted in accordance with the World Medical Association Declaration of Helsinki. All participants provided informed consent before participating and were given the option to withdraw from the study at any time. The collected data were deidentified to ensure participant privacy and confidentiality. No compensation was provided to participants.

Data Analysis

We conducted both quantitative data analysis and summarized user feedback on the *Sleep Memories* app. Quantitative data analysis was conducted using IBM SPSS Statistics (version 29) [44]. All continuous variables were inspected for normality using visual inspection (histograms), while the assumption of homogeneity of variance for these data was checked using the Levene test for equality of variances. For group comparisons, independent samples *t* tests were used for continuous variables that met assumptions of normality and homogeneity of variance. When the assumption of equal variances was violated, the Welch *t* test was applied. For nonnormally distributed variables, the nonparametric Mann-Whitney *U* test was used. Categorical variables were compared using chi-square tests. Prior to examining associations between SDMC performance and other variables (demographics, clinical factors, sleep behaviors, modifiable dementia risk factors, and global cognitive functioning), linear relationships were assessed visually using scatterplots. Following these assessments, the relationships between variables were analyzed using Pearson *r* for normally distributed variables and Spearman ρ for nonnormally distributed variables. We then compared the demographics, clinical factors, and subjective sleep-related factors between high and low SDMC performers as well as between participants who provided user feedback and those in the main sample.

Next, we summarized user feedback on the app. This analysis was based on open-ended questions at the end of each questionnaire section (see Multimedia Appendix 1), allowing participants to describe their experiences, usability, and challenges. The Likert-scale items were used only to guide discussion and provide context; no formal statistical analysis was conducted due to the small sample size (n=22), and participants' responses were optional. We reported only the descriptive summary of selected items. The data analysis was conducted in an unbiased manner, with independent coding and theme identification and development carried out by 2 researchers (CCYT and SS). Following this, a meeting was held to compare and discuss their interpretations. Any discrepancies in coding or pattern identification and development were resolved through discussion, with a third researcher, AL, facilitating consensus when needed. To ensure impartiality, CCYT and SS were blinded to participants' age, clinical diagnosis, and other personal information. Researchers identified patterns first; only then did they examine participants' age, clinical diagnosis, and task completion to assess specific difficulties completing the tasks. The summarization of user feedback was conducted based on a four-step procedure: (1) becoming familiar with the data, (2) generating initial codes, (3) clearly defining and naming them, and finally, (4) producing writing that presents a coherent narrative supported by data extracts [45].

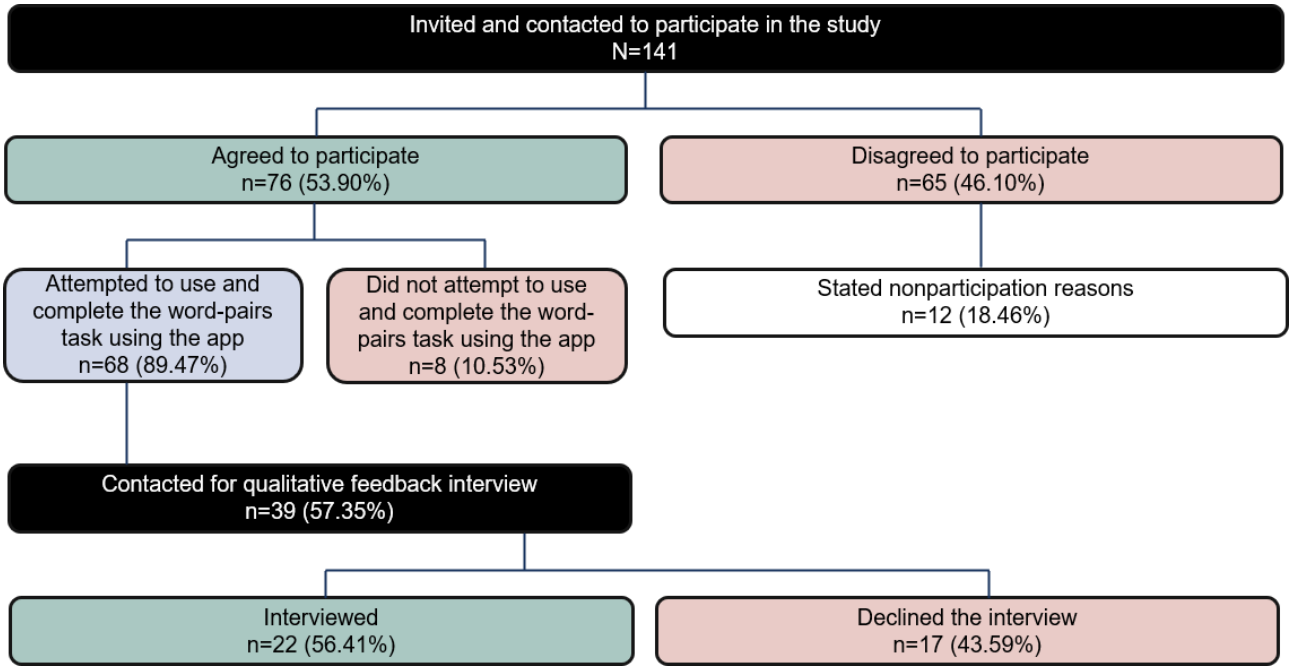
Results

Participants and Willingness to Participate

A total of 141 consecutive clinic attendees during the data collection period were invited to participate. Of these, 54% (n=76) agreed to participate. Of those who agreed, 89% (n=68) attempted the word-pairs task using the app. A subset (n=22) was contacted for postuse feedback (see Figure 3, which details the recruitment process, including numbers who agreed or declined, task completion, and participation in interviews). Additionally, 11 participants provided explicit reasons for nonparticipation during their clinic visit, including task too lengthy (n=1), lack of time (n=2), feeling overwhelmed by the task (n=2), difficulty downloading the app (n=2), struggling with technology (n=1), inability of the phone to scan the QR code (n=1), perceived lack of memory benefits (n=1), and forgetting to bring their phone (n=1).

XSL•FO
RenderX

Figure 3. Recruitment and completion summary.



Feasibility of the *Sleep Memories* App

Of the 68 participants who attempted to use the *Sleep Memories* app to complete the word-pairs task, 68% (n=46) completed the evening recall test, 69% (n=47) completed the morning delayed recall test, and 57% (n=39) completed both tests. Out of the 46 participants who completed the evening recall test, 15% (n=7) completed only the evening recall test and not the morning delayed recall test. Similarly, of the 47 participants who completed the morning delayed recall test, 17% (n=8) completed only the morning delayed recall test and not the evening recall test.

The feasibility of the *Sleep Memories* app was assessed by examining the completion rate of the tasks: evening learning trials 1 to 4 (58/68, 85%; 57/68, 84%; 60/68, 88%; 60/68, 88%; respectively), evening recall test (46/48, 68%), morning delayed recall test (47/68, 69%), morning multiple-choice test (45/68, 66%), and both evening and morning recall tests (39/68, 57%)—all tasks had completion rates above 50% (Table 2). As described in the *Data Analysis* section, after adjusting for outliers, we computed the mean SDMC score of 92.76 (SD 13.85). In addition, the chi-square test showed no significant differences in completion rates between SCD and MCI groups (Table 3).

Table . User uptake of the *Sleep Memories* app (N=68).

Type of task	Values, n (%)[a]	95% CI
Evening learning trial 1	58 (85.29)	74.15 - 92.35
Evening learning trial 2	57 (83.82)	72.47 - 91.27
Evening learning trial 3	60 (88.24)	77.59 - 94.42
Evening learning trial 4	60 (88.24)	77.59 - 94.42
Evening recall test	46 (67.65)	55.09 - 78.20
Morning delayed recall test	47 (69.12)	56.60 - 79.46
Morning multiple-choice test	45 (66.18)	53.59 - 76.93
Completed both evening recall and morning delayed recall tests	39 (57.35)	44.80 - 69.07

[a]n (%) represents number of users completed and completion rate.

XSL•FO
RenderX

Table . User uptake of the *Sleep Memories* app in participants with subjective cognitive decline (SCD) versus mild cognitive impairment (MCI).

Type of task	χ^2 (*df*)	*P* value	ϕ
Evening learning trial 1	0.61 (1)	.44	−0.09
Evening learning trial 2	0.13 (1)	.72	−0.04
Evening learning trial 3	0.05 (1)	.82	−0.03
Evening learning trial 4	0.05 (1)	.82	−0.03
Evening recall test	2.40 (1)	.12	0.19
Morning delayed recall test	0.12 (1)	.73	−0.04
Morning multiple-choice test	0.06 (1)	.81	−0.03
Completed both evening recall and morning delayed recall tests	0.001 (1)	.98	0.004

Among those who completed both evening recall and morning delayed recall tests, on average, the interval between evening encoding and sleep onset was 98.49 (SD 56.55) minutes. In the morning, on average, the interval between wake-up time and morning delayed recall was 33.61 (SD 37.85) minutes.

Factors Influencing Willingness to Participate and Feasibility of the *Sleep Memories* App

Characteristics of Individuals Who Agreed and Disagreed to Use the Sleep Memories App

Participants who agreed to use the *Sleep Memories* app were more likely to have SCD than MCI, with higher MMSE scores, and tended to experience less difficulty waking up in the morning compared to those who declined participation (Table 4). There were no significant differences in other demographics, clinical factors, and sleep behaviors between those who agreed and those who disagreed to participate.

XSL•FO
RenderX

Table . Characteristics of participants who agreed to participate (n=76) and disagreed to participate (n=65).

Variable domains and variable	Agreed (n=76)	Disagreed (n=65)	Test statistic (*df*)	*P* value	Effect size
Demographics					
Age (y), mean (SD)	70.19 (7.75)	72.53 (7.08)	−1.86 (139)[a]	.07	0.31[b]
Education[c] (y), mean (SD)	14.72 (2.77)	14.67 (3.46)	0.08 (131)[a]	.94	0.02[b]
Sex: females, n (%)	55 (72.37)	43 (66.15)	0.64 (1)[d]	.42	−0.07[e]
Vocational: working, n (%)	30 (39.47)	20 (30.77)	0.88 (1)[d]	.35	0.08[e]
Clinical factors					
Cognitive functioning[c]: MMSE[f], mean (SD)[g]	28.95 (1.56)	27.90 (2.73)	1725.50[h]	.03[i]	0.03[j]
Depression[g]: GDS-15[k], mean (SD)	2.95 (2.88)	3.34 (3.74)	2438.00[h]	.89[i]	0.00[j]
Anxiety[g]: GAI[l], mean (SD)	4.01 (3.98)	4.83 (5.67)	2416.00[h]	.82[i]	0.00[j]
Clinical diagnosis: MCI[m], n (%)	34 (44.74)	46 (70.77)	13.10 (1)[d]	<.001	−0.31[e]
Sleep-related factors, mean (SD)					
Sleep quality: PSQI[n]	6.47 (3.33)	6.88 (3.62)	−0.69 (139)[a]	.49	0.12[b]
Cogsleep[o]: bedtime	22:27 (1.02)	22:16 (1.34)	0.80 (139)[a]	.43	0.09[b]
Cogsleep[o]: wake time	6:27 (1.09)	6:51 (1.38)	−1.63 (112.75)[a]	.11	0.20[b]
Cogsleep[o]: trouble with waking up in the morning, number of days[c]	3.67 (2.40)	5.04 (2.25)	−2.04 (46)[a]	.047	0.59[b]
Insomnia[g]: ISI[gp]	7.17 (5.36)	7.14 (6.18)	2353.00[h]	.63	0.002[j]

[a]*t* test.

[b]Hedges *g*.

[c]Missing data for some participants.

[d]χ^2 test.

[e]Phi.

[f]MMSE: Mini-Mental State Examination.

[g]Mann-Whitney *U* test was conducted for these nonnormally distributed variables. The assumptions of Mann-Whitney *U* test, including dependent variables at a continuous level, an independent variable consisting of 2 independent groups, and similarly shaped distributions between groups, were assumed. The test compared medians between the groups since the distributions between groups had the same shape.

[h]Mann-Whitney *U* test.

[i]*P* value for Mann-Whitney *U* test, the asymptotic significance was reported for ≥20 per group.

[j]Eta-squared.

[k]GDS-15: Geriatric Depression Scale-15.

[l]GAI: Geriatric Anxiety Inventory.

[m]MCI: mild cognitive impairment.

[n]PSQI: Pittsburgh Sleep Quality Index.

[o]Cogsleep: Cogsleep Semi-structured Interview Revised Version.

[p]ISI: Insomnia Severity Index.

XSL•FO
RenderX

Characteristics of Full Completers and Partial Completers

Full completers (completed both the evening recall and morning recall tests) had significantly more years of education and higher anxiety levels compared to partial completers (Table 5). There were no significant differences in other demographics, clinical factors, and sleep behaviors between these 2 groups.

XSL•FO
RenderX

Table . Characteristics of full completers (n=39) and partial completers (n=29).

Variable domains and variables	Full completer (n=39)	Partial completer (n=29)	Test statistic (*df*)	*P* value	Effect size
Demographics					
Age (y), mean (SD)	69.87 (8.02)	71.61 (7.41)	−0.91 (66)[a]	.37	0.22[b]
Education (y), mean (SD)[c]	15.19 (2.65)	13.76 (2.85)	2.06 (62)[a]	.04	0.52[b]
Sex: females, n (%)	26 (86.67)	23 (79.31)	1.32 (1)[d]	.25	0.14[e]
Vocational: working, n (%)	17 (43.59)	10 (34.48)	0.42 (1)[d]	.52	0.08[e]
Clinical factors					
Cognitive functioning[c]: MMSE[f], mean (SD)[g]	29.03 (1.48)	28.79 (1.66)	466.00[h]	.28[i]	0.02[j]
Depression[g]: GDS-15[k], mean (SD)	2.46 (2.52)	3.10 (2.60)	456.00[h]	.17[i]	0.03[j]
Anxiety[g]: GAI[l], mean (SD)	4.64 (4.08)	2.69 (3.61)	387.50[h]	.03[i]	0.07[j]
Clinical diagnosis: MCI[m], n (%)	16 (41.03)	12 (41.38)	0.001 (1)[d]	.98	−0.004[e]
Sleep-related factors, mean (SD)					
Sleep quality: PSQI[n]	6.21 (3.24)	6.62 (3.44)	−0.51 (66)[a]	.61	0.12[b]
Cogsleep[o]: bedtime	22:25 (0:59)	22:21 (1:09)	0.26 (66)[a]	.79	0.05[b]
Cogsleep[o]: wake time	6:20 (1:12)	6:36 (1:07)	−0.91 (66)[a]	.37	0.15[b]
Cogsleep[o]: trouble with waking up in the morning, number of days[c]	2.55 (2.24)	3.75 (2.19)	−1.14 (16)[a]	.27	0.54[b]
Insomnia[g]: ISI[p]	7.31 (4.85)	6.21 (5.46)	470.00[h]	.24	0.02[j]

[a]*t* test.

[b]Hedges *g*.

[c]Missing data for some participants.

[d]χ^2 test.

[e]ϕ value.

[f]MMSE: Mini-Mental State Examination.

[g]Mann-Whitney *U* test was conducted for these nonnormally distributed variables. The assumptions of Mann-Whitney *U* test, including dependent variables at a continuous level, an independent variable consisting of 2 independent groups, and similarly shaped distributions between groups, were assumed. The test compared medians between the groups since the distributions between groups had the same shape.

[h]Mann-Whitney *U* test.

[i]*P* value for Mann-Whitney *U* test, the asymptotic significance was reported for ≥20 per group.

[j]Eta-squared.

[k]GDS-15: Geriatric Depression Scale-15.

[l]GAI: Geriatric Anxiety Inventory.

[m]MCI: mild cognitive impairment.

[n]PSQI: Pittsburgh Sleep Quality Index.

[o]Cogsleep: Cogsleep Semi-structured Interview Revised Version.

[p]ISI: Insomnia Severity Index.

XSL•FO
RenderX

User Feedback and Acceptability of the *Sleep Memories* App in a Memory Clinic Setting

First, we present a brief descriptive summary of selected Likert-scale items from the Sleep Memories App Feedback Questionnaire (see Multimedia Appendix 1). Overall, 82% (18/22) of participants agreed that the app instructions were easy to understand and that the task was straightforward. A majority (20/22, 91%) indicated a preference for completing the task using the app rather than in a sleep laboratory. Additionally, 86% (19/22) reported an overall positive experience with the app. Notably, none of the participants reported needing help from their family to complete the SDMC task, while only 14% (3/22) reported needing family assistance to navigate the app.

Next, qualitative feedback from participants described age-related physical and cognitive changes that influenced their experience with the app. More specifically, 2 participants commented that the app appeared more user-friendly for individuals who are younger or more computer-literate, rather than for older users. One participant noted that being from a generation familiar with computers, but not mobile phones, made completing the task on an unfamiliar device challenging. Another suggested that improvements could be made to better accommodate individuals with vision or dexterity impairments.

Participants generally provided positive feedback, describing the app as easy to use, user-friendly, and well-designed, with some finding it fun, interesting, and convenient to complete tasks at home. One participant even shared their positive experiences with others, and another described it as one of the best apps they had ever used. Less positive feedback included feelings of nervousness or stress during testing, concerns about performance, and perceptions that the app was less suitable or engaging for older users. Given that the app is designed for cognitive testing, some level of anxiety is expected.

In addition, participants reported several technical challenges related to the functionality of the *Sleep Memories* app and the clarity of the accompanying instructions. Participants also raised concerns regarding the app's accessibility and suitability for older adults and individuals with additional needs. The most frequently mentioned issue was tasks being delivered too quickly, making it difficult to keep up. Nevertheless, they appreciated the voice input feature, making task completion easier (see Table 6 for more details).

XSL•FO
RenderX

Table . Summary of user feedback.

Patterns and comments	Participants, n
Subjective experiences	
Positive	
Easy to use, user-friendly, and straightforward	3
Well set up	1
Good layout	1
Convenient as it allowed tasks to be completed at home	1
Simple tap-based multiple-choice test	1
Interesting process	1
Shared positive experience with others	1
Best apps they had ever used	1
Less positive	
Nervous or stressed while completing the tests	3
Worried about their performance	2
Not suitable for older users	2
The app was “dry”	1
Technical challenges	
Instruction sheet provided on how to use the app was unclear	1
Needed more explanation for the morning tasks	1
Misunderstood the evening task notification as requiring immediate completion	2
Unclear wording of the second morning task (multiple-choice)	2
Unawareness that questions would automatically skip after a time limit	3
Lack of instruction on what to do if they did not know an answer	1
Confusion over morning task timing based on wake-up time	1
Delays in morning task delivery	2
Suggested the need to visually distinguish cues from response options in the multiple-choice task	2
App’s accessibility	
Tasks being delivered too quickly, making it difficult to keep up	5
Challenges using the microphone	2
Difficulty with touch functions on mobile devices due to aging	1
Needed assistance from family members to navigate the app and difficulty using it independently	1
Experienced illness during testing	2
Questioned whether individuals with dementia might become frustrated using the app	1
Voice input feature was helpful and made task completion easier	1

Predictors of SDMC Performance

There were no differences in SDMC across gender (t_{34}=−0.65, P=.52, Hedges g=0.23) and clinical diagnostic groups (t_{34}=0.55, P=.59, Hedges g=0.18). We also examined the associations between SDMC performance and several variables: demographics, clinical factors, sleep behaviors, modifiable dementia risk factors, and global cognitive functioning. We first examined these relationships using all available SDMC data points. As shown in Table 7, there was a significant and positive correlation between premorbid IQ (as measured by Wechsler Test of Adult Reading) and SDMC. Nevertheless, there were no statistically significant correlations between SDMC and other demographics, clinical factors, sleep behaviors, and modifiable dementia risk factors. Since this was an exploratory study with a small sample size, we did not correct for multiple comparisons.

XSL•FO
RenderX

Table . Predictors of sleep-dependent memory consolidation (SDMC).

Variable domains and variables	Effect size	*P* value
Demographic		
Age	0.14[a]	.41
Years of education	0.10[a]	.57
Clinical factors		
Depression: GDS-15[b]	0.001[a]	.10
Anxiety: GAI[c]	0.25[a]	.13
Chronic disease: CIRS-G[d]	−0.006[a]	.97
Sleep behaviors		
Sleep quality: PSQI[e]	−0.11[a]	.51
Insomnia: ISI[f]	−0.05[a]	.74
Modifiable dementia risk factors		
LIBRA[g] 2 index	0.01[a]	.94
Global cognitive functioning		
WTAR[h] (premorbid IQ)	0.38[a]	.02
MMSE[i]	0.11[j]	.51
RAVLT[k] (7/5% retention)	0.05[a]	.79

[a]Pearson *r*.
[b]GDS-15: Geriatric Depression Scale-15.
[c]GAI: Geriatric Anxiety Inventory.
[d]CIRS-G: Cumulative Illness Rating Scale-Geriatric.
[e]PSQI: Pittsburgh Sleep Quality Index.
[f]ISI: Insomnia Severity Index.
[g]LIBRA: Lifestyle for Brain Health.
[h]WTAR: Wechsler Test of Adult Reading.
[i]MMSE: Mini-Mental State Examination.
[j]Spearman ρ.
[k]RAVLT: Rey Auditory Verbal Learning Test.

Next, we conducted a sensitivity analysis to assess whether potential outliers on the evening recall task might influence the results. We excluded 3 evening recall data points that were SD 1.50 below the mean. These data points were excluded for memory consolidation-specific analyses (SDMC% retention) but retained for all other analyses. The SD 1.50 cutoff was selected as a pragmatic adaptation of common outlier detection rule [46]. Results with this exclusion are presented in Table 8, and findings were consistent both with and without the exclusion.

XSL•FO
RenderX

Table . Predictors of sleep-dependent memory consolidation (SDMC) excluding evening recall scores 1.50 SD below the mean.

Variable domains and variables	Effect size	*P* value
Demographic		
Age	0.01[a]	.95
Years of education	0.17[a]	.34
Clinical factors		
Depression: GDS-15[b]	−0.03[a]	.88
Anxiety: GAI[c]	0.23[a]	.19
Chronic disease: CIRS-G[d]	−0.06[a]	.73
Sleep behaviors		
Sleep quality: PSQI[e]	−0.12[a]	.48
Insomnia: ISI[f]	0.01[a]	.95
Modifiable dementia risk factors		
LIBRA[g] 2 index	−0.10[a]	.57
Global cognitive functioning		
WTAR[h] (premorbid IQ)	0.36[a]	.04
MMSE[i]	0.05[j]	.74
RAVLT[k] (7/5% retention)	0.03[a]	.88

[a]Pearson *r*.
[b]GDS-15: Geriatric Depression Scale-15.
[c]GAI: Geriatric Anxiety Inventory.
[d]CIRS-G: Cumulative Illness Rating Scale-Geriatric.
[e]PSQI: Pittsburgh Sleep Quality Index.
[f]ISI: Insomnia Severity Index.
[g]LIBRA: Lifestyle for Brain Health.
[h]WTAR: Wechsler Test of Adult Reading.
[i]MMSE: Mini-Mental State Examination.
[j]Spearman ρ.
[k]RAVLT: Rey Auditory Verbal Learning Test.

Exploratory Analyses

In terms of comparing the characteristics between low and high SDMC performers (Table 9), low performers had significantly less education and lower anxiety levels compared to high performers. There were no other differences between these 2 groups.

XSL•FO
RenderX

Table . Characteristics of low (n=18) and high (n=18) performers of sleep-dependent memory consolidation (SDMC) task.

Variable domains and variables	Low performer (n=18)	High performer (n=18)	Test statistic (*df*)	*P* value	Effect size
Demographics					
Age (y), mean (SD)	70.83 (6.77)	68.95 (8.96)	0.71 (34)[a]	.48	0.24[b]
Education (y), mean (SD)[c]	14.11 (2.35)	16.25 (2.59)	−2.52 (32)[a]	.02	0.87[b]
Sex: females, n (%)	10 (55.56)	14 (77.78)	2.00 (1)[d]	.16	0.24[e]
Vocational: working, n (%)	9 (50)	8 (44.44)	0.11 (1)[d]	.74	0.06[e]
Clinical factors					
Cognitive functioning: MMSE[f], mean (SD)[g]	28.94 (1.51)	29.17 (1.51)	147.50[h]	.65[i]	0.006[j]
Depression[g]: GDS-15[k], mean (SD)	2.50 (2.94)	2.56 (2.28)	146.50[h]	.63[i]	0.01[j]
Anxiety[g]: GAI[l], mean (SD)	3.39 (3.99)	6.11 (3.94)	88.50[h]	.02[i]	0.15[j]
Clinical diagnosis: MCI[m], n (%)	8 (44.44)	7 (38.89)	0.11 (1)[d]	.74	0.06[e]
Sleep-related factors, mean (SD)					
Sleep quality: PSQI[n]	6.00 (3.01)	6.61 (3.62)	−0.55 (34)[a]	.59	0.19[b]
Cogsleep[o]: bedtime	22:38 (1:04)	22:09 (0:53)	1.42 (34)[a]	.17	0.35[b]
Cogsleep[o]: wake time	6:21 (1:01)	6:12 (1:20)	0.40 (34)[a]	.69	0.08[b]
Cogsleep[o]: trouble with waking up in morning, number of days[c]	1.00 (0.00)	3.25 (2.64)	−2.09 (5)[a]	.09	1.21[b]
Insomnia[g]: ISI[p]	7.00 (4.41)	8.11 (5.05)	144.00[h]	.58[i]	0.009[j]

[a]*t* test.

[b]Cohen *d*.

[c]Missing data for some participants.

[d]χ^2 test.

[e]Phi.

[f]MMSE: Mini-Mental State Examination.

[g]Mann-Whitney *U* test was conducted for these nonnormally distributed variables. The assumptions of the Mann-Whitney *U* test, including dependent variables at a continuous level, an independent variable consisting of 2 independent groups, and similarly shaped distributions between groups, were assumed. The test compared medians between the groups since the distributions between groups had the same shape.

[h]Mann-Whitney *U* test.

[i]*P* values for the Mann-Whitney *U* test, the exact significance was reported for the small sample size.

[j]Eta-squared.

[k]GDS-15: Geriatric Depression Scale-15.

[l]GAI: Geriatric Anxiety Inventory.

[m]MCI: mild cognitive impairment.

[n]PSQI: Pittsburgh Sleep Quality Index.

[o]Cogsleep: Cogsleep Semi-structured Interview Revised Version.

[p]ISI: Insomnia Severity Index.

XSL•FO
RenderX

When comparing the characteristics of participants who provided user feedback (subsample) with those of the main sample, the subsample included a higher proportion of individuals with SCD than the main sample. Nevertheless, there was no difference between the groups in the proportion of full and partial completers. There were also no other differences between these 2 groups (Table 10).

We also reported task completion rates among participants who reported difficulty in using the app (Table 11).

XSL•FO
RenderX

Table . Characteristics of user feedback provider (n=22) compared to the main sample (n=119).

Variable domains and variables	User feedback provider (n=22)	Main sample (n=119)	Test statistic (*df*)	*P* value	Effect size
Demographics					
Age (y), mean (SD)	72.18 (6.45)	71.10 (7.71)	0.61 (139)[a]	.54	−0.14[b]
Education (y), mean (SD)[c]	14.39 (2.99)	14.76 (3.12)	−0.51 (131)[a]	.61	0.12[b]
Sex: females, n (%)	19 (86.36)	79 (66.39)	3.50 (1)[d]	.06	−0.16[e]
Vocational: working, n (%)	7 (31.82)	43 (36.13)	0.25 (1)[d]	.62	−0.04[e]
Clinical factors					
Cognitive functioning[c]: MMSE[f], mean (SD)[g]	29.18 (1.22)	28.35 (2.33)	1002.50[h]	.16[i]	0.01[j]
Depression[g]: GDS-15[k], mean (SD)	2.64 (2.70)	3.22 (3.40)	1236.50[h]	.68[i]	0.00[j]
Anxiety[g]: GAI[l], mean (SD)	4.05 (3.87)	4.45 (5.00)	1241.00[h]	.70[i]	0.00[j]
Clinical diagnosis: MCI[m], n (%)	7 (31.82)	73 (61.34)	7.62 (1)[d]	.006	−0.24[e]
Sleep-related factors, mean (SD)					
Sleep quality: PSQI[n]	6.95 (2.34)	6.61 (3.64)	0.58 (42.44)[a]	.56	−0.10[b]
Cogsleep[o]: bedtime	22:33 (0.59)	22:20 (1.21)	0.72 (139)[a]	.47	−0.11[b]
Cogsleep[o]: wake time	6:25 (1.15)	6:41 (1.26)	−0.82 (139)[a]	.41	0.13[b]
Cogsleep[o]: trouble with waking up in the morning, number of days[c]	3.25 (2.87)	4.49 (2.36)	−0.99 (46)[a]	.33	0.52[b]
Insomnia[g]: ISI[p]	7.09 (4.21)	7.17 (5.98)	1212.00[h]	.58	0.002[j]
Word-pairs task completion, n (%)					
Task completion: completers	12 (57.14)[q]	27 (57.45)[r]	0.001 (1)[d]	.98	−0.003[e]

[a]*t* test.

[b]Hedges *g*.

[c]Missing data for some participants.

[d]χ^2 test.

[e]Phi.

[f]MMSE: Mini-Mental State Examination.

[g]Mann-Whitney *U* test was conducted for these nonnormally distributed variables. The assumptions of the Mann-Whitney *U* test, including dependent variables at a continuous level, an independent variable consisting of 2 independent groups, and similarly shaped distributions between groups, were assumed. The test compared medians between the groups since the distributions between groups had the same shape.

[h]Mann-Whitney *U* test.

[i]*P* values for the Mann-Whitney *U* test, the asymptotic significance was reported for ≥20 per group.

[j]Eta-squared.

[k]GDS-15: Geriatric Depression Scale-15.

[l]GAI: Geriatric Anxiety Inventory.

[m]MCI: mild cognitive impairment.

XSL•FO
RenderX

[n]PSQI: Pittsburgh Sleep Quality Index.

[o]Cogsleep: Cogsleep Semi-structured Interview. Revised Version.

[p]ISI: Insomnia Severity Index.

[q]n=21.

[r]n=47.

Table . Task completion rates among participants reporting difficulty.

ID	Age (y)	Clinical diagnosis	Completion of 4 evening learning trials	Completion of evening test	Completion of morning test	Completion of morning multiple choice task	Notes
1	73.46	MCI[a]	Yes (all 4)	No (no score)	No (no score)	No (no score)	Experienced app malfunction and believed their data were incomplete
2	72.05	SCD[b]	No (all 4)	No (no score)	No (no score)	No (no score)	Unable to complete the task
3	70.88	SCD	Yes (all 4)	Yes	Yes	Yes	The screen went blank due to inactivity (not touching the screen)
4	76.37	SCD	Yes (all 4)	Yes	Yes	Yes	Encountered several issues with the app and could not complete the final morning task
5	73.77	MCI	Yes (all 4)	No (no score)	No (no score)	No (no score)	Thought there were 3 trials, and as a result, lost all previous results
6	70.61	SCD	Yes (all 4)	Yes	Yes	Partial (score=1)	Unable to complete the morning task
7	64.5	SCD	Yes (all 4)	Yes	Yes	Yes	Encountered an issue during the first attempt and had to clear the app data and try again

[a]MCI: mild cognitive impairment.

[b]SCD: subjective cognitive decline.

Discussion

Principal Findings

This study aimed to evaluate the willingness to engage with and the feasibility of using the *Sleep Memories* app to assess SDMC in older adults with cognitive concerns and impairments within a brain health and memory clinic setting. It is important to note that this study focused on the feasibility of a digital delivery platform for the SDMC task rather than providing objective physiological sleep metrics, such as those obtained through actigraphy or polysomnography.

Overall, the app was deemed feasible (>50% completion) as determined a priori. Notably, our study involved greater logistical demands than prior app-based cognitive studies. Sleep-dependent memory protocols require tighter scheduling around sleep periods and may involve overnight laboratory stays in unfamiliar environments, which are not always feasible for older adults with cognitive impairment. Therefore, the 57% completion rate, while modest, indicates practical feasibility for real-world implementation. Willingness to participate was higher among those with SCD than MCI, those with higher global cognitive scores, and those with less difficulty waking up in the morning. Additionally, task completion was higher among those with more years of education and higher anxiety levels.

Interestingly, while similar completion rates in the SCD and MCI groups support the app's feasibility, the higher initial refusal rate among individuals with MCI and lower cognitive scores suggests a barrier to participation. This aligns with prior

XSL•FO
RenderX

work, which showed that, relative to those with more severe cognitive impairments, individuals with SCD are often motivated to participate in research studies due to concern about cognitive decline, despite normal objective performance [47]. Future research could adopt a dyadic approach by involving a care partner to assist with technical issues at home, or potentially offering on-call technical support, helping reduce barriers for those who might otherwise decline participation.

In this study, we recorded reasons for nonparticipation from a subset of participants, which suggested that nonparticipation was influenced by a combination of perceived task burden (eg, participant perceived the word-pairs task as lengthy), technical barriers (eg, phone unable to download the app), and motivational factors (eg, participant perceived the task as lacking memory benefit). These findings are in accordance with the Unified Theory of Acceptance and Use of Technology [48], where "effort expectancy" and "facilitating conditions" correspond to the perceived task burden and technical barriers observed in this study, and "performance expectancy" reflects participants' motivation to engage with the app based on perceived cognitive benefits. Moreover, consistent with past research linking higher global cognitive functioning to greater health literacy [49], our results revealed that individuals agreeing to participate had more superior global cognition compared to those with cognitive decrements. This may be suggesting that participants with higher global cognitive functioning might be more motivated to engage in research perceived as supporting sleep and cognitive health. Difficulty with sleep also appears to affect willingness to participate in sleep-dependent memory tasks. Waking difficulties often linked to disrupted sleep or fatigue. This may reduce motivation or capacity to engage in research activities in the morning, especially those requiring structured participation such as completing tasks via a mobile app with strict timings (ie, 3 h before sleep and upon waking). Analyses of other sleep-related apps reveal similar issues. For instance, "SleepFit," which aims to enhance sleep hygiene among paramedics [50], had a substantial dropout rate of 91.40%. The authors attributed this low engagement and retention to fatigue and disrupted sleep, particularly making it difficult for users to adhere to the app's scheduled tasks.

The *Sleep Memories* app demonstrated completion rates comparable to other remote cognitive assessments in older populations. For example, a study of older adults at risk of cognitive decline and dementia (mean age 57.3, SD 5.3 y), using the smartphone-based iVitality app (featuring tests adapted from conventional neuropsychological assessments) reported completion rates of approximately 60% [36]. Similarly, an unsupervised app−based study in healthy older adults, assessing memory recall and object and scene discrimination, reported completion rates ranging from 54% to 69% [35]. Notably, the completion rate observed in this pilot study exceeded that reported for the Linus Health Core Cognitive Evaluation app in older adults (mean age 74.0, SD 6.0 y), which achieved a substantially lower rate of around 10% [37]. In this study, we found a strong initial task completion rate, particularly during the learning phase. However, the task completion rate declined during the evening recall and morning recall tests, suggesting that the completion of delayed recall tasks administered either 30 minutes later (evening recall task) or the following morning (morning delayed recall and multiple-choice tasks) may pose greater challenges in remote, unsupervised contexts. Re-engagement with the app after a delay may be influenced by contextual factors such as interruptions, competing demands, or diminished motivation to return to the task [51,52]. Notably, completion rates for all learning trials, evening, and morning tasks did not differ significantly between participants with SCD and MCI. This finding is consistent with a past study that similarly showed no significant differences in online cognitive test completion based on cognitive status [53]. Additionally, based on user feedback provided, aging—especially physical and cognitive changes—affects task completion. For instance, for some, a decline in fine motor skills appeared to hinder typing accurately on a mobile screen, especially on smaller devices. Slowed reaction times and decreased hand-eye coordination further contributed to difficulties with tasks requiring precise tapping. Additionally, reduced visual acuity common in later life [54] might also impact the ability to read small on-screen instructions and navigate through mobile apps. These sensory and motor limitations are often compounded by a general age-related unfamiliarity with digital interfaces, resulting in reduced confidence in using mobile apps. Importantly, our results showed no differences in age between full completers and partial completers, suggesting that it is not chronological age itself but the effects of aging such as changes in motor, visual, and cognitive abilities that influence task completion. Other than physical and cognitive changes due to aging, education and psychological factors also play a role in affecting the completion of app-based SDMC tasks, as we found full completers tended to have more years of education and higher anxiety levels compared to partial completers. Education fosters cognitive discipline and familiarity with structured tasks [55,56], potentially making these individuals more likely to persist through challenges. Meanwhile, elevated anxiety (when managed effectively) can enhance task vigilance, increase motivation to avoid failure, and drive individuals to complete tasks with greater urgency [57]. A similar pattern was observed when we compared the characteristics of low and high performers of SDMC tasks—high performers had more years of education and higher anxiety levels than low performers. Together, these findings suggest that education equips individuals with strategies to manage tasks, and that anxiety can provide motivation for completion and performance in SDMC tasks. When examining user feedback from full completer and partial completer groups, full completers generally reported more positive subjective experiences, noting that the app was easy to use, enjoyable, and fun. Two participants from the full completer group mentioned feeling nervous during the task, possibly due to the evaluative or confrontational nature of cognitive testing and expressed concerns about their task performance. Partial completers reported similar experiences, indicating little difference between full completers and partial completers in terms of subjective experiences. In terms of technical challenges, some full completers reported that the instructional sheet was unclear and that the task progressed too quickly. It is important to note that substantially slowing down task delivery may risk introducing ceiling effects, making the task overly easy and unnecessarily

XSL•FO
RenderX

lengthy. In contrast, partial completers provided little feedback regarding the app or technical difficulties. This lack of feedback may reflect lower motivation, earlier disengagement before issues arose, or a simple disinclination to comment. One participant in the partial completer group (diagnosed with MCI) reported needing assistance from family members to navigate the app. Although this was an isolated case, and the task had originally been designed and validated for people with MCI within the sleep laboratory [5], this observation suggests that caregiver support may be necessary to facilitate engagement with the *Sleep Memories* app among some users with greater cognitive impairment.

Having considered factors influencing willingness to participate and the feasibility of the app based on task completion, it is also important to understand those affecting SDMC performance. Notably, cognitive reserve affects SDMC performance. We observed a positive correlation between premorbid IQ and SDMC performance, which aligns with past research showing that cognitive reserve is linked to better cognitive outcomes [58]. Our findings suggest that premorbid IQ influences task performance and retention, underscoring the importance of considering cognitive reserve in remote assessments.

Given that the *Sleep Memories* app was used in an unsupervised, remote capacity, we acknowledge that contextual and behavioral factors may contribute to variability in SDMC scores. However, we found that except for premorbid IQ, none of the demographic, clinical, sleep behavior, modifiable dementia risk factors, or cognitive variables were significantly associated with overnight memory consolidation. This suggests that while such factors may introduce variability, they do not systematically account for individual differences in SDMC performance.

Research Impact

This was the first study to demonstrate the feasibility of SDMC data collection delivered unsupervised through a chatbot and mobile app. While not a replacement for laboratory studies, the *Sleep Memories* app demonstrates potential as a complementary research tool that may support more accessible and cost-effective data collection.

The comparable task completion rates between participants with SCD and those with MCI indicate that the *Sleep Memories* app and the word-pairs task can be feasibly administered across these populations. Nevertheless, as with other digital cognitive tools, unsupervised SDMC testing presents challenges related to cognitive load and user engagement. In this study, we also highlighted key factors that may facilitate the adoption of the app, which could inform the development of personalized interventions aimed at optimizing healthy brain aging in the future.

Participant feedback emphasized the value of clear, concise instructions, and well-defined expectations regarding task timing and completion. To support usability among older adults, accessible design strategies are essential, such as larger icons, high-contrast text, intuitive audio cues, and customizable interface settings. Ultimately, refining and validating remote memory tools such as the *Sleep Memories* app across global cognitive and demographic profiles will be critical for advancing equitable, scalable cognitive health monitoring in clinical and community settings.

Limitations

There are several limitations to this study. First, we did not record participants' feedback. Future qualitative studies should consider audio-recording responses and conducting formal thematic analyses of verbatim data. Nonetheless, we identified several factors that influenced participation and completion rates. Second, we did not systematically assess the digital literacy of participants, which may have influenced their ability to use the app and interpret results. Third, as widely documented in the literature, the validity of remote cognitive testing must be carefully considered, given the potential for uncontrolled distractions and external assistance. Reassuringly, the observed scores in this study were comparable to those reported in similar lab-based SDMC studies involving comparable cohorts [5]. Nonetheless, incorporating additional validity safeguards and conducting larger-scale evaluations will be important to further mitigate these limitations. Fourth, the study did not directly validate against supervised lab-based SDMC assessments. However, complete equivalence should not be expected, given differences in delivery format and response conditions, and the observed scores were comparable to those reported in similar lab-based SDMC research in older and MCI samples [5]. Finally, sleep and wake times were self-reported, and no objective measures were collected to verify these estimates. Notwithstanding these limitations, the study has strengths worth mentioning. For instance, it demonstrates high ecological validity and that recruitment and SDMC task delivery via the *Sleep Memories* app are feasible for a memory clinic setting.

Conclusions

This study demonstrated the feasibility of using the *Sleep Memories* app to assess SDMC in older adults within a memory clinic context. These findings offer valuable guidance for the development of accessible digital tools that can effectively monitor cognitive health across varying levels of impairment.

The team is currently continuing to collect SDMC data using the app from a larger sample recruited through a brain health and memory clinic setting to support ongoing refinement and larger-scale validation studies. These efforts aim to further establish the feasibility, reliability, and clinical utility of the app, the characterization of neurobiological correlates, and sensitivity to interventions and to inform scalability for community-based implementation.

Acknowledgments

The authors would like to thank all the participants for their involvement in this study, the clinical and research team of the Healthy Brain Ageing Program, University of Sydney, and the research team from Commonwealth Scientific and Industrial

XSL•FO
RenderX

Research Organisation (CSIRO) for their support. The authors used Microsoft Copilot and ChatGPT to suggest language improvements within the manuscript.

Funding

Professor Sharon L Naismith is supported by the NHMRC Leadership II Fellowship (2008064), Miss Clara Cher Yi Tan is supported by the NHMRC Synergise Integrate and Enhance Sleep Research to Transform Brain Ageing (SIESTA) Synergy Program scholarship, and Dr Aaron Lam is supported by SIESTA Synergy Program as a postdoctoral researcher.

Data Availability

The dataset of this study is available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sleep Memories App Qualitative Feedback Questionnaire.
[DOCX File, 76 KB - aging_v9i1e87926_app1.docx]

References

1. Charest J, Grandner MA. Sleep and athletic performance: impacts on physical performance, mental performance, injury risk and recovery, and mental health. Sleep Med Clin 2020 Mar;15(1):41-57. [doi: 10.1016/j.jsmc.2019.11.005] [Medline: 32005349]
2. De Longis E, Kassis A, Rémond-Derbez N, et al. Cognitive benefits of sleep: a narrative review to explore the relevance of glucose regulation. Sleep Adv 2025;6(1):zpae095. [doi: 10.1093/sleepadvances/zpae095] [Medline: 39850251]
3. Marshall L, Cross N, Binder S, Dang-Vu TT. Brain rhythms during sleep and memory consolidation: neurobiological insights. Physiology (Bethesda) 2020 Jan 1;35(1):4-15. [doi: 10.1152/physiol.00004.2019] [Medline: 31799908]
4. Squire LR, Genzel L, Wixted JT, Morris RG. Memory consolidation. Cold Spring Harb Perspect Biol 2015 Aug 3;7(8):a021766. [doi: 10.1101/cshperspect.a021766] [Medline: 26238360]
5. Lam A, Haroutonian C, Grummitt L, et al. Sleep-dependent memory in older people with and without MCI: the relevance of sleep microarchitecture, OSA, hippocampal subfields, and episodic memory. Cereb Cortex 2021 May 10;31(6):2993-3005. [doi: 10.1093/cercor/bhaa406] [Medline: 33565576]
6. Leong RLF, Lo JC, Chee MWL. Sleep-dependent prospective memory consolidation is impaired with aging. Sleep 2021 Sep 13;44(9):zsab069. [doi: 10.1093/sleep/zsab069] [Medline: 33755184]
7. Pace-Schott EF, Spencer RMC. Sleep-dependent memory consolidation in healthy aging and mild cognitive impairment. Curr Top Behav Neurosci 2015;25:307-330. [doi: 10.1007/7854_2014_300] [Medline: 24652608]
8. Westerberg CE, Mander BA, Florczak SM, et al. Concurrent impairments in sleep and memory in amnestic mild cognitive impairment. J Int Neuropsychol Soc 2012 May;18(3):490-500. [doi: 10.1017/S135561771200001X] [Medline: 22300710]
9. Atherton KE, Nobre AC, Zeman AZ, Butler CR. Sleep-dependent memory consolidation and accelerated forgetting. Cortex 2014 May;54(100):92-105. [doi: 10.1016/j.cortex.2014.02.009] [Medline: 24657478]
10. Harding CD, Holloway BM, DeYoung PN, et al. Subjective daytime sleepiness, not sleep quality or hypoxia, predicts sleep-dependent memory consolidation in a cohort of older adults. J Clin Sleep Med 2025 Jul 1;21(7):1217-1226. [doi: 10.5664/jcsm.11648] [Medline: 40135690]
11. Wang SY, Baker KC, Culbreth JL, et al. “Sleep-dependent” memory consolidation? Brief periods of post-training rest and sleep provide an equivalent benefit for both declarative and procedural memory. Learn Mem 2021 Jun;28(6):195-203. [doi: 10.1101/lm.053330.120] [Medline: 34011516]
12. Lam A, Simonetti S, D’Rozario A, et al. Perceptions of the use of mobile apps to assess sleep-dependent memory in older adults with subjective and objective cognitive impairment: focus group approach. JMIR Aging 2025 Apr 28;8(1):e68147. [doi: 10.2196/68147] [Medline: 40294432]
13. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161. [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
14. LaMonica HM, English A, Hickie IB, et al. Examining internet and eHealth practices and preferences: survey study of Australian older adults with subjective memory complaints, mild cognitive impairment, or dementia. J Med Internet Res 2017 Oct 25;19(10):e358. [doi: 10.2196/jmir.7981] [Medline: 29070481]
15. Reid G, Vassilev P, Irving J, et al. The usability and reliability of a smartphone application for monitoring future dementia risk in ageing UK adults. Br J Psychiatry 2024 Jun;224(6):245-251. [doi: 10.1192/bjp.2024.18] [Medline: 38356396]
16. Ireland D, Lam A, Bradford D, Naismith S. A new role for chatbots: automation of a sleep-dependent memory task. Stud Health Technol Inform 2024 Jan 25;310:519-523. [doi: 10.3233/SHTI231019] [Medline: 38269863]

XSL•FO
RenderX

17. Miller MJ, Diaz A, Conti C, et al. The ADNI4 digital study: a novel approach to recruitment, screening, and assessment of participants for AD clinical research. Alzheimer's Dement 2024 Oct;20(10):7232-7247. [doi: 10.1002/alz.14234]
18. Duffy SL, Lagopoulos J, Hickie IB, et al. Glutathione relates to neuropsychological functioning in mild cognitive impairment. Alzheimers Dement 2014 Jan;10(1):67-75. [doi: 10.1016/j.jalz.2013.01.005] [Medline: 23688577]
19. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
20. Wechsler D. Wechsler Test of Adult Reading: WTAR: Psychological Corporation; 2001. [doi: 10.1037/t15180-000]
21. Lezak MD. Neuropsychological Assessment, 2nd edition: Oxford University Press; 1983. URL: https://www.scirp.org/reference/referencespapers?referenceid=1560922 [accessed 2026-05-29]
22. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment–beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J Intern Med 2004 Sep;256(3):240-246. [doi: 10.1111/j.1365-2796.2004.01380.x] [Medline: 15324367]
23. Jessen F, Amariglio RE, van Boxtel M, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alzheimers Dement 2014 Nov;10(6):844-852. [doi: 10.1016/j.jalz.2014.01.001] [Medline: 24798886]
24. Duffy SL, Lagopoulos J, Terpening Z, et al. Association of anterior cingulate glutathione with sleep apnea in older adults at-risk for dementia. Sleep 2016 Apr 1;39(4):899-906. [doi: 10.5665/sleep.5650] [Medline: 26856906]
25. Miller MD, Paradis CF, Houck PR, et al. Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale. Psychiatry Res 1992 Mar;41(3):237-248. [doi: 10.1016/0165-1781(92)90005-n] [Medline: 1594710]
26. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59(Suppl 20):22-33. [Medline: 9881538]
27. Naismith SL, Duffy S, Gordon C. CogSleep Semi-Structured Sleep-Wake Interview for older adults (version 2.0). : University of Sydney; 2022 URL: https://cogsleep.sydney.edu.au/wp-content/uploads/2022/04/CogSleep_Questionnaire_DTA_Version_March_2022.pdf [accessed 2026-05-29]
28. Ghadimi S, Ereso J, Kaula AJ, et al. Feasibility of cognitive testing and ecological momentary assessments using smartphones in middle aged and older adults with insomnia. BMC Digit Health 2025;3(1):21. [doi: 10.1186/s44247-025-00158-4] [Medline: 40606001]
29. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
30. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 2011 May 1;34(5):601-608. [doi: 10.1093/sleep/34.5.601] [Medline: 21532953]
31. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. In: Clinical Gerontology: Routledge; 1986:165-173 URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9781315826233-11/geriatric-depression-scale-gds-javaid-sheikh-jerome-yesavage [accessed 2026-05-29]
32. Pachana NA, Byrne GJ, Siddle H, Koloski N, Harley E, Arnold E. Development and validation of the Geriatric Anxiety Inventory. Int Psychogeriatr 2007 Feb;19(1):103-114. [doi: 10.1017/S1041610206003504] [Medline: 16805925]
33. Ireland D, Atay C, Liddle J, et al. Hello Harlie: enabling speech monitoring through chat-bot conversations. Stud Health Technol Inform 2016;227:55-60. [Medline: 27440289]
34. Ireland D, Bradford D, Szepe E, et al. Introducing Edna: a trainee chatbot designed to support communication about additional (secondary) genomic findings. Patient Educ Couns 2021 Apr;104(4):739-749. [doi: 10.1016/j.pec.2020.11.007] [Medline: 33234441]
35. Öhman F, Berron D, Papp KV, et al. Unsupervised mobile app-based cognitive testing in a population-based study of older adults born 1944. Front Digit Health 2022;4:933265. [doi: 10.3389/fdgth.2022.933265] [Medline: 36426215]
36. Jongstra S, Wijsman LW, Cachucho R, Hoevenaar-Blom MP, Mooijaart SP, Richard E. Cognitive testing in people at increased risk of dementia using a smartphone app: the iVitality proof-of-principle study. JMIR Mhealth Uhealth 2017 May 25;5(5):e68. [doi: 10.2196/mhealth.6939] [Medline: 28546139]
37. Fowler NR, Hammers DB, Perkins AJ, et al. Feasibility and acceptability of implementing a digital cognitive assessment for Alzheimer disease and related dementias in primary care. Ann Fam Med 2025 May 27;23(3):191-198. [doi: 10.1370/afm.240293] [Medline: 40300815]
38. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. J Clin Epidemiol 2009 May;62(5):464-475. [doi: 10.1016/j.jclinepi.2008.12.011] [Medline: 19348971]
39. Koo BM, Vizer LM. Mobile technology for cognitive assessment of older adults: a scoping review. Innov Aging 2019 Jan;3(1):igy038. [doi: 10.1093/geroni/igy038] [Medline: 30619948]
40. Zhou C, Ai Y, Wang S, et al. Barriers and facilitators to participation in electronic health interventions in older adults with cognitive impairment: an umbrella review. BMC Geriatr 2024 Dec 26;24(1):1037. [doi: 10.1186/s12877-024-05645-3] [Medline: 39725926]

XSL•FO
RenderX

41. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
42. Rosenau C, Köhler S, van Boxtel M, Tange H, Deckers K. Validation of the updated "LIfestyle for BRAin health"(LIBRA) index in the English Longitudinal Study of Ageing and Maastricht Aging Study. J Alzheimers Dis 2024;101(4):1237-1248. [doi: 10.3233/JAD-240666] [Medline: 39302374]
43. Pons A, LaMonica HM, Mowszowski L, Köhler S, Deckers K, Naismith SL. Utility of the LIBRA index in relation to cognitive functioning in a clinical health seeking sample. J Alzheimers Dis 2018;62(1):373-384. [doi: 10.3233/JAD-170731] [Medline: 29439337]
44. Downloading IBM SPSS statistics 29. IBM. 2024. URL: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-29 [accessed 2025-04-28]
45. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
46. Schwertman NC, Owens MA, Adnan R. A simple more general boxplot method for identifying outliers. Comput Stat Data Anal 2004 Aug;47(1):165-174. [doi: 10.1016/j.csda.2003.10.012]
47. Perrotin A, La Joie R, de La Sayette V, et al. Subjective cognitive decline in cognitively normal elders from the community or from a memory clinic: differential affective and imaging correlates. Alzheimers Dement 2017 May;13(5):550-560. [doi: 10.1016/j.jalz.2016.08.011] [Medline: 27693187]
48. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view1. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
49. Serper M, Patzer RE, Curtis LM, et al. Health literacy, cognitive ability, and functional health status among older adults. Health Serv Res 2014 Aug;49(4):1249-1267. [doi: 10.1111/1475-6773.12154] [Medline: 24476068]
50. Shriane AE, Vincent GE, Ferguson SA, Rebar A, Kolbe-Alexander T, Rigney G. Improving sleep health in paramedics through an app-based intervention: a randomised waitlist control pilot trial. BMC Public Health 2024 Sep 3;24(1):2395. [doi: 10.1186/s12889-024-19823-w] [Medline: 39227826]
51. Leiva L, Böhmer M, Gehring S, Krüger A. Back to the app: the costs of mobile application interruptions. Presented at: MobileHCI '12: Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services; Sep 21-24, 2012. [doi: 10.1145/2371574.2371617]
52. Shabir H, D'Costa M, Mohiaddin Z, et al. The barriers and facilitators to the use of lifestyle apps: a systematic review of qualitative studies. Eur J Investig Health Psychol Educ 2022 Jan 27;12(2):144-165. [doi: 10.3390/ejihpe12020012] [Medline: 35200235]
53. Ashford MT, Aaronson A, Kwang W, et al. Unsupervised online paired associates learning task from the Cambridge Neuropsychological Test Automated Battery (CANTAB®) in the Brain Health Registry. J Prev Alzheimers Dis 2024;11(2):514-524. [doi: 10.14283/jpad.2023.117] [Medline: 38374758]
54. Sjöstrand J, Laatikainen L, Hirvelä H, Popovic Z, Jonsson R. The decline in visual acuity in elderly people with healthy eyes or eyes with early age-related maculopathy in two Scandinavian population samples. Acta Ophthalmol 2011 Mar;89(2):116-123. [doi: 10.1111/j.1755-3768.2009.01653.x] [Medline: 19845558]
55. Reder LM, Liu XL, Keinath A, Popov V. Building knowledge requires bricks, not sand: the critical role of familiar constituents in learning. Psychon Bull Rev 2016 Feb;23(1):271-277. [doi: 10.3758/s13423-015-0889-1] [Medline: 26139355]
56. Shi Y, Qu S. The effect of cognitive ability on academic achievement: the mediating role of self-discipline and the moderating role of planning. Front Psychol 2022;13:1014655. [doi: 10.3389/fpsyg.2022.1014655] [Medline: 36275240]
57. Cheng BH, McCarthy JM. Understanding the dark and bright sides of anxiety: a theory of workplace anxiety. J Appl Psychol 2018 May;103(5):537-560. [doi: 10.1037/apl0000266] [Medline: 29355338]
58. Merema MR, Speelman CP, Kaczmarek EA, Foster JK. Age and premorbid intelligence suppress complaint-performance congruency in raw score measures of memory. Int Psychogeriatr 2012 Mar;24(3):397-405. [doi: 10.1017/S1041610211001918] [Medline: 22018288]

Abbreviations

CIRS-G: Cumulative Illness Rating Scale-Geriatric
CSIRO: Commonwealth Scientific and Industrial Research
LIBRA: Lifestyle for Brain Health
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
SCD: subjective cognitive decline
SDMC: sleep-dependent memory consolidation

XSL•FO
RenderX

Edited by Q Fan; submitted 10.Dec.2025; peer-reviewed by A Lynham, V Raj; revised version received 15.May.2026; accepted 20.May.2026; published 11.Jun.2026.

Please cite as:
Tan CCY, Lam A, Ireland D, Bradford D, Cross N, Kurniawan MT, Simonetti S, Fripp J, Naismith SL
Feasibility and Acceptability of a Remote Sleep–Dependent Memory Assessment in Older Adults With Cognitive Concerns: Pilot Cross-Sectional Study
JMIR Aging 2026;9:e87926
URL: https://aging.jmir.org/2026/1/e87926
doi:10.2196/87926

© Clara Cher Yi Tan, Aaron Lam, David Ireland, Dana Bradford, Nathan Cross, Melvern Ted Kurniawan, Simone Simonetti, Jurgen Fripp, Sharon L Naismith. Originally published in JMIR Aging (https://aging.jmir.org), 11.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perceived Algorithmic Care Intensity and Older Adults' Autonomy in Community Smart Care: Mixed Methods Study

Xianghui Huang, PhD

Logistics and E-commerce College, Zhejiang Wanli University, 8 Qianhu S Rd, Ningbo, Zhejiang, China

Corresponding Author:
Xianghui Huang, PhD
Logistics and E-commerce College, Zhejiang Wanli University, 8 Qianhu S Rd, Ningbo, Zhejiang, China

Abstract

Background: As smart older adult care shifts from basic information support to more continuous algorithm-driven care, concerns have grown about whether such systems support older adults' independence or weaken their sense of autonomy. Although previous research has focused mainly on technology acceptance, less is known about how algorithmic care shapes perceived autonomy and through which psychological pathways this occurs.

Objective: This study examined whether perceived algorithmic care intensity reduces older adults' perceived autonomy in community smart care settings, whether decisional substitution and perceived surveillance mediate this relationship, and whether digital literacy moderates these effects.

Methods: A 3-stage design was adopted. First, a qualitative prestudy with 15 older adults was conducted to refine the focal constructs and measures. Second, a vignette-based experiment with 233 valid participants tested the effects of algorithmic care intensity on decisional substitution, perceived surveillance, and perceived autonomy. Third, a community-based survey with 298 valid participants examined the moderated mediation model in real-world smart older adult care settings.

Results: In study 1, compared with the low-intensity condition, the high-intensity condition significantly increased decisional substitution and perceived surveillance, while reducing perceived autonomy. Indirect effects through decisional substitution (effect=−0.072, 95% CI −0.144 to −0.016) and perceived surveillance (effect=−0.084, 95% CI −0.161 to −0.019) were both significant. In study 2, perceived algorithmic care intensity positively predicted decisional substitution (β=.400, *P*<.001) and perceived surveillance (β=.440, *P*<.001), and negatively predicted perceived autonomy (β=−.356, *P*<.001). The indirect effects through decisional substitution (effect=−0.112, 95% CI −0.172 to −0.065) and perceived surveillance (effect=−0.149, 95% CI −0.215 to −0.093) were significant. Digital literacy significantly weakened the effect of algorithmic care intensity on decisional substitution (β=−.174, *P*=.002), but not on perceived surveillance (β=−.071, *P*=.18).

Conclusions: Algorithmic care may undermine older adults' perceived autonomy both directly and indirectly through decisional substitution and perceived surveillance. Digital literacy appears to be a selective rather than universal buffer. These findings extend smart older adult care research beyond technology acceptance and clarify the tension between empowerment and control in algorithmic care.

(JMIR Aging 2026;9:e94559) doi:10.2196/94559

KEYWORDS

community smart eldercare; perceived algorithmic care intensity; decisional substitution; perceived surveillance; perceived autonomy

Introduction

Smart Older Adult Care and the Emerging Autonomy Question

As population aging deepens and digital technologies become increasingly embedded in community care, smart older adult care is shifting from basic information support to more continuous, algorithm-driven care [1]. Wearable devices, home sensor systems, remote health platforms, and intelligent alert tools are now widely used for fall detection, chronic disease management, risk identification, and daily support. These technologies can ease pressure on care resources, improve response efficiency, and support aging in place [2]. At the same time, as data collection, condition recognition, and automated reminders become part of older adults' daily lives, smart older adult care also raises concerns about excessive monitoring, technological dependence, privacy exposure, and weakened care relationships [2-5]. The key issue is therefore no longer only whether such technologies are usable, but whether they change how older adults experience themselves as active subjects in care [6].

XSL•FO
RenderX

Existing research has examined smart older adult care and digital health technologies mainly through technology adoption, usage intention, functional performance, and implementation barriers. Previous studies show that perceived usefulness, ease of use, digital literacy, external support, organizational resources, deployment conditions, and service fit shape older adults' acceptance and use of digital technologies [7,8]. Other studies emphasize the benefits of smart monitoring and digital care tools for health management, risk warning, and continuity of care [9,10]. Although this literature has advanced understanding of smart older adult care, its main focus remains whether older adults are willing and able to use technology. Less attention has been paid to a more psychological question: when care increasingly relies on system recommendations, automated reminders, and continuous monitoring, do older adults still feel that they are directing their own lives, or do they begin to feel managed by the system [11]?

From Technology Adoption to Psychological Mechanisms

Research grounded in self-determination theory suggests that autonomy, competence, and relatedness are central to older adults' engagement with digital health technologies. Technology is more likely to support independent living when it preserves choice, navigation, and personalized control [12]. At the same time, studies of monitoring technologies show that older adults' concerns extend beyond informational privacy to dignity, control, and the feeling of being watched [13]. Research on intelligent care systems similarly suggests that technologies designed to improve efficiency and safety may also strengthen discipline, relational distance, and external control [14]. However, these concerns are often discussed separately. Existing research has not yet provided a clear model explaining how algorithmic care may undermine older adults' perceived autonomy and through which psychological pathways this occurs. Nor has it adequately examined whether these effects vary according to older adults' digital capabilities.

This study responds to these gaps by moving the discussion of community smart older adult care beyond technology adoption and toward perceived autonomy. It examines how perceived algorithmic care intensity influences older adults' perceived autonomy through 2 parallel psychological pathways, decisional substitution and perceived surveillance, and whether digital literacy moderates these effects. Decisional substitution refers to the feeling that care arrangements, health judgments, and everyday decisions are increasingly preset, guided, or replaced by the system, whereas perceived surveillance refers to the sense that one's behavior, physical condition, and daily routines are continuously observed and recorded by the technological system. The former reflects weakened control over action and decision making, whereas the latter reflects disruption of psychological boundaries.

This Study and Its Contributions

In this study, perceived algorithmic care intensity does not refer to a single device category or fixed technical function. Rather, it captures older adults' overall perception of how deeply algorithmic systems are involved in the surrounding care environment, especially through proactive system participation, automated evaluative input, and system-shaped follow-up arrangements. On this basis, the study develops a parallel mediation framework with first-stage moderation to examine a central tension in smart older adult care—whether algorithmic care supports independent living or instead weakens older adults' agency and control in the name of safety and efficiency.

This study makes 3 main contributions. First, it extends smart older adult care research beyond technology adoption by showing how algorithmic care shapes older adults' perceived autonomy. Second, it identifies decisional substitution and perceived surveillance as 2 distinct psychological pathways through which algorithmic care affects autonomy. Third, it shows that digital literacy functions as a selective boundary condition, buffering decisional substitution more than perceived surveillance, thereby offering a more refined basis for the design of human-centered community smart older adult care.

Literature Review

Algorithmic Care and Older Adults' Perceived Autonomy

Hypothesis 1: The higher the perceived intensity of algorithmic care, the lower older adults' perceived autonomy.

According to self-determination theory, autonomy does not mean acting without external support or completing all activities independently [15]. Rather, it refers to the subjective experience that one's actions are voluntary, self-endorsed, and guided by personal choice [16,17]. In community smart older adult care, perceived autonomy refers to older adults' perceived ability to remain the primary agents of everyday care-related arrangements, daily rhythms, and responses to health and living needs while using smart care services. It therefore concerns care-related everyday autonomy within technology-supported community living, rather than formal medical autonomy over clinical treatment or legal data autonomy over personal information governance.

As people age, physical decline, chronic disease management, and increasing dependence on care can make daily life more shaped by family members, community services, and institutional arrangements [18-20]. In this context, perceived autonomy is closely linked to dignity, self-worth, life continuity, and the ability to maintain an ordinary life on one's own terms [21]. Research further suggests that positive experiences with digital health systems depend not only on functional convenience, but also on whether technology supports autonomy, competence, and social connection [12]. When systems are easy to navigate, interactive, and customizable, older adults are more likely to view them as resources for self-management rather than sources of external control [22].

However, algorithmic care creates new challenges for perceived autonomy. Unlike traditional care, it increasingly intervenes in risk detection, health assessment, anomaly alerts, behavioral reminders, and daily management [23]. In older adult care and digital monitoring settings, such technologies may improve safety and efficiency while also intensifying concerns about control, dignity, constant observation, discipline, relational distance, and external control [14]. Thus, the more strongly older adults perceive algorithmic care to shape daily care, health management, and life arrangements through system recognition,

XSL•FO
RenderX

automated judgment, and system-initiated follow-up, the more likely they are to experience reduced control and choice in the care process.

Algorithmic Care and Decisional Substitution

Hypothesis 2: The higher the perceived intensity of algorithmic care, the stronger older adults' sense of decisional substitution.

In community smart older adult care, algorithmic care deserves attention not only because it can improve efficiency, but also because it may reshape older adults' role in decision-making [11]. Decisional substitution refers to the subjective feeling that matters once requiring one's own judgment, weighing, and choice are increasingly preset, guided, or handled by an external system [24,25]. Rather than treating technology as a simple source of information support, this concept captures a psychological experience in which one's room for participation, judgment, and control is reduced [26,27]. Research on shared decision-making shows that active participation in decisions is central to maintaining agency and self-direction; when decisions are shaped mainly by external authority, procedural rules, or technical logic, individuals are more likely to shift from active decision makers to passive recipients [16].

Algorithmic care increasingly performs functions such as risk detection, anomaly alerts, health assessment, behavioral reminders, and task prioritization [28]. As such systems move from supportive tools to active participants in judgment, older adults may feel that important decisions are no longer truly made by themselves but are prestructured or partly replaced by system logic [29]. Recent studies have therefore emphasized deliberation, customization, and human-centered loops in preserving users' decision-making role [30]. Without adequate room for participation, excessive reliance on automation and monitoring may create tension between supporting care and replacing the care recipient's own judgment [31].

From this perspective, higher perceived algorithmic care intensity implies not only deeper technological involvement, but also a shift in judgmental authority. The more older adults feel that care, health management, and daily arrangements depend on system recognition, automated triggers, and programmed recommendations, the more their space for self-judgment and initiative may be reduced. As a result, they are more likely to develop a stronger sense of decisional substitution.

Algorithmic Care and Perceived Surveillance

Hypothesis 3: The higher the perceived intensity of algorithmic care, the stronger older adults' perceived surveillance.

Beyond decisional substitution, perceived surveillance represents another key pathway through which algorithmic care may shape psychological outcomes. Perceived surveillance refers to the subjective feeling that one's behavior, physical condition, and daily routines are continuously observed, recorded, and analyzed by a technological system [3,32]. Unlike ordinary information collection, it captures a situated experience in which individuals feel that their lives are being constantly seen, interpreted, and even anticipated. Previous studies suggest that in health monitoring and assisted living settings, users' concerns extend beyond data leakage to dignity, control, autonomy, and the experience of intrusive observation [4,33-35].

This logic is especially salient in community smart older adult care. Traditional care relies mainly on face-to-face observation and periodic assessment, whereas algorithmic care is typically built on continuous sensing, anomaly detection, and real-time feedback [36]. Wearable devices and home sensor systems transform physiological states, movement patterns, daily routines, and unusual behaviors into data that can be tracked, recorded, and evaluated [37]. Although such systems may improve risk warning and service response, they may also blur the boundary between supportive care and constant observation [38].

This risk arises not simply because information is collected, but because technology turns everyday life into something continuously visible, recordable, and interpretable. As monitoring devices intensify this visibility, older adults may evaluate technology not only in terms of safety, but also in terms of whether it leaves them feeling constantly seen [39]. Perceived surveillance may therefore reshape how power is experienced within the care relationship [40], especially when smart technologies become deeply embedded in care institutions and daily routines [14]. Thus, the more older adults feel that their health status, activity patterns, and daily arrangements depend on continuous data capture and automated recognition, the more likely they are to feel watched rather than merely supported [41].

The Mediating Roles of Decisional Substitution and Perceived Surveillance in Perceived Autonomy

Hypothesis 4a: Decisional substitution mediates the relationship between perceived algorithmic care intensity and older adults' perceived autonomy.

Hypothesis 4b: Perceived surveillance mediates the relationship between perceived algorithmic care intensity and older adults' perceived autonomy.

Although algorithmic care may provide timely alerts, support continuity in health management, and reduce care uncertainty, its implications extend beyond convenience and safety. As argued in previous sections, stronger algorithmic care may heighten both decisional substitution and perceived surveillance. Self-determination theory suggests that autonomy depends not on complete independence from external support, but on whether individuals feel that their actions reflect their own will and that important choices are made through understanding, participation, and endorsement [16,17]. When an external system repeatedly narrows a person's space for judgment, choice, and interpretation, perceived autonomy is likely to decline even if the system is intended to protect, assist, or optimize [3,33].

Decisional substitution is likely to reduce perceived autonomy because it weakens the sense that older adults remain active participants in directing their own lives [42]. Research on shared decision-making shows that involvement in understanding information, weighing options, and selecting outcomes is essential for maintaining agency and self-direction [43]. When key matters are preset, prioritized, or automatically triggered by an external system, individuals are more likely to move from

XSL•FO
RenderX

active participants to passive executors [44,45]. In community smart older adult care, this shift may emerge as older adults become accustomed to relying on system reminders, assessments, and generated priorities [46]. Over time, when the system judges first and the individual responds afterward, the experience of self-directed action is likely to weaken [29]. Related research in virtual care and long-term care similarly suggests that technology-mediated care can narrow deliberative space, which is an early sign of declining perceived autonomy [47].

Perceived surveillance is likely to reduce perceived autonomy through a different mechanism: the disruption of psychological boundaries. When bodily states, movement patterns, and daily routines are continuously observed, recorded, and analyzed, the experience is no longer simply one of information collection, but of being constantly seen and interpreted [32,33]. This ongoing visibility reshapes private space, behavioral freedom, and self-control [4]. Previous research shows that resistance to monitoring technologies often centers on privacy, intrusiveness, stigma, and the feeling of being watched [48]. When technology conflicts with users' values or everyday practices, systems intended to support independent living may instead be experienced as intrusions into personal boundaries and daily rhythms [5]. In this sense, perceived surveillance is not only a privacy concern, but also a challenge to whether one can still live freely in one's own way.

The Moderating Role of Digital Literacy and Conditional Indirect Effects

Hypothesis 5a: Digital literacy moderates the relationship between algorithmic care intensity and decisional substitution, such that this relationship is stronger for older adults with lower digital literacy than for those with higher digital literacy.

Hypothesis 5b: Digital literacy moderates the relationship between algorithmic care intensity and perceived surveillance, such that this relationship is stronger for older adults with lower digital literacy than for those with higher digital literacy.

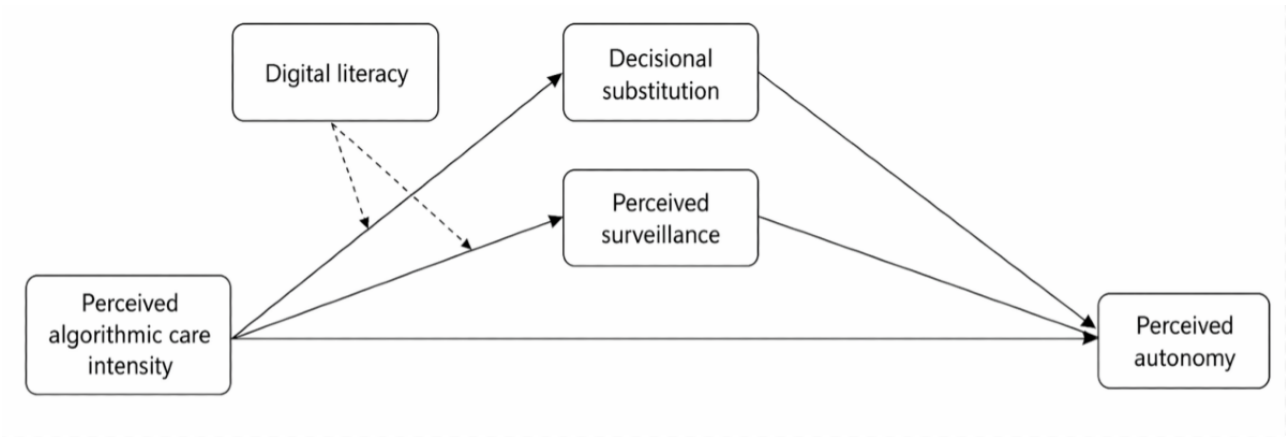
The preceding arguments suggest that algorithmic care intensity affects perceived autonomy through decisional substitution and perceived surveillance. However, these mechanisms may not operate uniformly across individuals. Older adults differ in how well they understand digital systems, how confidently they use them, and how they interpret system outputs. Digital literacy has been widely recognized as a key factor shaping older adults' experiences with and evaluations of technology [49]. Older adults with higher digital literacy are generally better able to understand how digital systems work, interpret system outputs, and use technology with confidence, whereas lower digital literacy is associated with greater difficulty interpreting digital information and engaging with technology-based services [50,51]. Because digital literacy reflects not only tool use but also basic understanding of system operation and data processing [52], it may serve as an important boundary condition in responses to algorithmic care.

Previous research suggests that individuals with higher digital literacy are more likely to view technological systems as instrumental support rather than as substitutes for their own control, and are therefore better able to maintain a sense of agency when using them [51,53]. In smart older adult care settings, stronger algorithmic care may generate reminders, warnings, and recommended courses of action. For older adults with higher digital literacy, such interventions are more likely to be interpreted as supportive suggestions because they are better able to understand system logic and evaluate system outputs [54]. For those with lower digital literacy, however, frequent prompts and automated judgments may be experienced as a transfer of decision-making authority, thereby strengthening decisional substitution.

Digital literacy may also shape interpretations of monitoring, because older adults with stronger technical understanding may be better able to distinguish data recording from social surveillance [55]. Therefore, higher digital literacy may weaken the association between algorithmic care intensity and perceived surveillance.

The research model is illustrated in Figure 1.

Figure 1. Research model.



XSL•FO
RenderX

Methods

Overview

This study adopted a multistage design combining a qualitative prestudy, a vignette-based experiment, and a field survey of older adults in real community settings. This design was used to improve contextual validity, establish causal direction, and test the proposed model in real-world smart older adult care contexts. Previous work shows that older adults' experiences of digital health and smart care technologies are shaped by everyday contexts, especially in relation to autonomy, privacy, surveillance, and control [12,56]. Reviews of monitoring-based assisted living technologies have also noted limitations related to ecological validity, end user orientation, and weak identification of underlying mechanisms [32,48]. Accordingly, we first used semistructured interviews to refine measurement items and vignette materials, then conducted an experiment to test causal effects, and finally validated the model using a community-based survey. In the main studies, participants were recruited in participating communities by the author and the author's student research team, who also assisted with on-site study administration.

Prestudy: Semistructured Interviews and Measure Development

Before the main studies, we conducted semistructured interviews to examine whether older adults described community smart older adult care in terms such as "things are no longer decided by me," "the system makes judgments first," or "I feel watched." The goal was not to test hypotheses, but to strengthen content validity and contextual fit by incorporating older adults' own language and everyday experiences [12,56].

Using purposive sampling, we recruited 15 community residents aged 60 years or older from an older adult care community in Hangzhou, China. All participants had used or been exposed to community smart older adult care services or related devices within the past year. To increase heterogeneity, we sought variation in age, education, living arrangement, chronic disease status, and prior experience with digital technologies. Interviews followed a semistructured guide covering system involvement, perceived loss of decision control, continuous recording or observation, and perceived autonomy, comfort, control, and respect. Each interview lasted about 30-50 minutes. With consent, interviews were audio recorded and transcribed.

The transcripts were analyzed using thematic analysis. Moreover, 2 members of the research team independently reviewed the transcripts, conducted open coding, compared coding results, and resolved differences through iterative discussion until agreement was reached on higher-order themes. The resulting themes were used to revise item wording and inform the vignette materials for study 1, thereby reducing the gap between researchers' conceptual assumptions and older adults' lived experiences [57].

Study 1: Vignette-Based Experiment

Study 1 tested whether perceived algorithmic care intensity affects decisional substitution, perceived surveillance, and perceived autonomy. A vignette-based experiment was used because cross-sectional survey data cannot rule out reverse causality, and older adults' responses to monitoring and digital care technologies are strongly shaped by specific use scenarios [13,56].

Study 1 used a 1-factor, 2-level between-subjects design. The independent variable was algorithmic care intensity, manipulated as a high intensity versus low intensity scenario. In the high intensity condition, the community smart older adult care system continuously collected information about older adults' activities, routines, and health status, automatically generated reminders, risk alerts, and care suggestions, and used these outputs to shape subsequent support and care arrangements. In the low intensity condition, the system mainly provided basic recording and information display, while subsequent support and care arrangements relied primarily on discussion between the older adult and community staff. In both conditions, community staff remained present as part of the care context. The 2 vignettes were matched in length, writing style, and background context.

We recruited 255 older adults aged 60 years or older from 2 large communities in Hangzhou, China, and randomly assigned them to one of the 2 conditions. Participants first received a brief explanation of the study and provided informed consent. They were then presented with 1 vignette at random and completed manipulation check items and the main measures. The manipulation check assessed whether participants perceived the system as highly reliant on recognition, automated reminders, and programmed judgment. The main measures assessed perceived algorithmic care intensity, decisional substitution, perceived surveillance, and perceived autonomy. Age, gender, education, self-rated health, and previous experience with digital technologies were also collected as control variables.

Study 2: Community-Based Field Survey

Study 2 examined whether the proposed model holds in real community smart older adult care settings and whether digital literacy moderates the proposed mechanisms. Because older adults' evaluations of smart care technologies depend on actual environments, interaction patterns, and the extent to which technology is embedded in daily life, vignettes alone cannot demonstrate whether the model applies to real use contexts [12].

Participants were recruited through on-site distribution with researcher-assisted completion in 3 large communities, day care centers, and community health service stations in Hangzhou and Ningbo, China. These 2 cities were selected because community smart older adult care services are relatively well developed there. A total of 330 older adults participated. Eligibility criteria were (1) aged 60 years or older; (2) having used or been exposed to at least 1 type of community smart older adult care service in the past year, such as a health monitoring system, remote consultation platform, smart emergency call device, or data management platform; and (3) having basic reading ability and being able to complete the questionnaire with assistance if needed. Before the survey, researchers provided a standardized explanation of "smart older adult care services" and instructed participants to answer based on their overall experience of algorithmic involvement across

XSL•FO
RenderX

the services they had used, rather than focusing on a single device or application.

Study 2 measured the same core constructs as study 1, that is, perceived algorithmic care intensity, decisional substitution, perceived surveillance, and perceived autonomy. Digital literacy was measured as the moderator. To reduce common method bias, the questionnaire visually separated different constructs into sections and included filler items to weaken consistency responding.

We estimated a moderated mediation model to test whether digital literacy attenuates the positive effect of algorithmic care intensity on decisional substitution and perceived surveillance. Multiple regression and bias-corrected bootstrap procedures were used to estimate conditional indirect effects, and moderation was tested through interaction terms. Age, gender, and education were included as control variables in the main regression analyses.

Statistical Analysis and Measures

This study measured five core constructs: (1) perceived algorithmic care intensity [13,56], (2) decisional substitution [12,13], (3) perceived surveillance [32,48], (4) perceived autonomy [22], and (5) digital literacy [58]. Because older adults' experiences with digital health and smart care technologies are highly contextual and subjective, the items were adapted from previous studies and refined through the prestudy interviews to fit community smart older adult care settings [12,56].

All formal measures used a 7-point Likert scale, ranging from 1=strongly disagree to 7=strongly agree. Study 1 and study 2 followed the same general measurement framework, with minor wording adjustments to fit the research context. In study 1, items referred to "the community smart older adult care system described above," whereas in study 2, they referred to "the community smart older adult care services you are currently using." This approach helped maintain construct consistency while improving contextual relevance and clarity.

Perceived algorithmic care intensity measured the extent to which older adults felt that the surrounding care process relied on system recognition, automated reminders, programmed recommendations, backend judgment, and technology-shaped follow-up arrangements. The construct captured older adults' higher-order perception of how deeply algorithmic systems were embedded in the care environment, rather than any single device type, alert frequency, or task criticality [13,56].

Decisional substitution measured the extent to which older adults felt that matters requiring their own judgment and choice were instead preset, guided, or handled by the system. This measure was grounded in self-determination theory and refined with reference to research on shared decision making and user participation [12,13].

Perceived surveillance measured the subjective feeling that one's behavior, physical condition, and daily routines were continuously observed, recorded, and analyzed by the technological system. This measure was informed by research on monitoring technologies, camera-based assisted living, and digital care systems for older adults [32,48].

Perceived autonomy captured whether older adults still felt that they remained the primary agents of everyday care-related arrangements, daily routines, and responses to health and living needs in community smart older adult care. The construct focused on agency, choice, and behavioral self-direction in ordinary technology-supported care, rather than formal medical decision autonomy or legal data autonomy [22].

Digital literacy measured older adults' ability to understand, access, and use digital technologies and digital health information [58]. A shortened digital literacy scale was adopted for this purpose.

In addition to the focal constructs, study 2 collected background information on age, gender, education level, marital status, self-rated health, number of chronic conditions, living arrangement, and frequency of smart older adult care service use. Age, gender, and education were included as controls in the reported regression analyses.

Ethical Considerations

This study received ethical approval from the institutional ethics committee of the authors' affiliated university (approval 2025-HUM-021) and was conducted in accordance with the Declaration of Helsinki. Data were collected between March and June 2025 from adults aged 60 years and older, and no minors were involved. All participants provided informed consent before participation. Participation was voluntary, and all responses were collected anonymously and used solely for academic research purposes. This study was not registered in a clinical trial registry because it was not designed as a prospective clinical intervention trial. Instead, it used a 3-stage design combining a qualitative prestudy, a vignette-based experiment, and a community-based field survey.

Results

Prestudy: Semistructured Interviews and Measure Development

The qualitative prestudy confirmed that the focal constructs were meaningful and recognizable in community smart older adult care settings. A total of 15 older adults who had used or been exposed to community smart older adult care services participated in semistructured interviews. Detailed participant characteristics are reported in Multimedia Appendix 1. The interviews examined how participants understood system involvement, loss of decision control, continuous recording or observation, and perceived autonomy, comfort, control, and respect. The purpose was not to test hypotheses, but to generate empirical input for item refinement and vignette development [12,56].

The thematic analysis yielded 6 representative themes, which are summarized in Multimedia Appendix 1. Overall, participants did not evaluate smart older adult care only in terms of convenience or technological advancement. Instead, they repeatedly focused on whether algorithm-enabled involvement changed their position in the care process. Moreover, 3 patterns

XSL•FO
RenderX

were especially salient. First, participants recognized safety and convenience benefits, especially for alerts, health monitoring, chronic disease management, and reducing burdensome medical visits. Second, they became more sensitive when the system moved from information display to active judgment, reminders, and arrangement shaping, which was associated with a perceived weakening of their own decision-making role. Third, continuous monitoring was often experienced not simply as data collection, but as ongoing visibility, boundary disruption, and reduced personal control. Across these accounts, a common concern emerged, that is, whether older adults could still remain the person in charge.

These findings informed both measure refinement and vignette design. Expressions such as "many things are no longer decided by me first" informed decisional substitution items, whereas expressions such as "it feels like there are always eyes in the house" informed perceived surveillance items. Items for perceived algorithmic care intensity emphasized proactive identification, automated reminders, programmed judgments, and system-shaped care arrangements. Items for perceived autonomy focused on whether participants still felt able to decide, retain freedom of choice, and live according to their own rhythm within smart care arrangements. The high-intensity vignette emphasized continuous monitoring, automated alerts, backend analysis, and system-shaped follow-up arrangements, whereas the low-intensity vignette emphasized basic recording and information display, with subsequent support relying more on discussion between older adults and community staff.

Study 1: Results of the Vignette-Based Experiment

Overview

After the prestudy confirmed that the focal constructs fit the smart older adult care context and after the measures and vignette materials were refined, we conducted study 1 to test the effects of perceived algorithmic care intensity on decisional substitution, perceived surveillance, and perceived autonomy. A total of 255 older adults were recruited. Responses with substantial missing data on the main study variables were excluded before analysis. As a result, 233 valid responses were retained for analysis, including 115 participants in the low algorithmic care intensity group and 118 in the high algorithmic care intensity group. The final analyses were conducted on complete cases, and no statistical imputation procedure was applied. The mean age of the sample was 69.70 (SD 6.24) years, and 127 out of 233 (54.5%) participants were female. Randomization checks showed no significant between-group differences in age, number of chronic conditions, gender, education, living arrangement, or digital literacy (all P values>.05), suggesting that random assignment was effective and that the 2 groups were comparable.

Manipulation Check

To examine whether the experimental manipulation was successful, we first compared the 2 groups on the manipulation check scale. The results showed that the high algorithmic care intensity group reported significantly higher perceived algorithmic care intensity than the low intensity group. Specifically, the mean score was 4.79 (SD 0.95) in the high intensity condition and 4.11 (SD 0.91) in the low intensity condition, and the difference was significant (2-tailed t_{231}=5.63; Cohen d=0.74; P<.001).

Reliability, Descriptive Statistics, and Correlation Analysis

Before testing the hypotheses, we examined the internal consistency of all measures. The results showed acceptable reliability for all scales: Cronbach α was 0.753 for the manipulation check scale, 0.695 for decisional substitution, 0.716 for perceived surveillance, and 0.782 for perceived autonomy. Although the α for decisional substitution was close to the conventional threshold, it remained within an acceptable range for a brief vignette-based measure, and the scale was retained because it captured the focal construct with adequate conceptual coverage in the study 1 context.

Tables 1 and 2 present the means, SDs, reliabilities, and correlations among the variables. Perceived algorithmic care intensity was positively correlated with decisional substitution (r=0.27, P<.001) and perceived surveillance (r=0.25, P<.001), and negatively correlated with perceived autonomy (r=−0.22, P<.001). In addition, decisional substitution was negatively correlated with perceived autonomy (r=−0.25, P<.001), and perceived surveillance was also negatively correlated with perceived autonomy (r=−0.29, P<.001). Notably, decisional substitution and perceived surveillance were positively related, but the correlation was not significant (r=0.05, P=.41). This suggests that although both reflect negative psychological responses to algorithmic care, they do not fully overlap in meaning. This pattern also supports treating them as 2 parallel mediators.

Table . Descriptive statistics and reliability for study 1 variables.

Variable	Mean (SD)	Cronbach α
Perceived algorithmic care intensity	4.45 (0.99)	0.753
Decisional substitution	4.28 (0.90)	0.695
Perceived surveillance	4.11 (0.93)	0.716
Perceived autonomy	3.86 (0.99)	0.782

XSL•FO
RenderX

Table . Correlation analysis for study 1 variables.

Variable	Perceived algorithmic care intensity	Decisional substitution	Perceived surveillance	Perceived autonomy
Perceived algorithmic care intensity				
r	1	0.27	0.25	−0.22
P value	—[a]	<.001	<.001	<.001
Decisional substitution				
r	0.27	1	0.05	−0.25
P value	<.001	—	.41	<.001
Perceived surveillance				
r	0.25	0.05	1	−0.29
P value	<.001	.41	—	<.001
Perceived autonomy				
r	−0.22	−0.25	−0.29	1
P value	<.001	<.001	<.001	—

[a]Not applicable.

Hypothesis Testing

To test hypotheses 1-3, we used experimental condition (low algorithmic care intensity=0, high algorithmic care intensity=1) as the independent variable and compared the 2 groups on decisional substitution, perceived surveillance, and perceived autonomy. As presented in Tables 3 and 4, compared with the low intensity condition, the high algorithmic care intensity condition significantly increased decisional substitution (mean 4.43, SD 0.94 vs mean 4.13, SD 0.84; t_{231}=2.64; Cohen *d*=0.35; *P*=.009) and perceived surveillance (mean 4.27, SD 0.91 vs mean 3.95, SD 0.95, t_{231}=2.58; Cohen *d*=0.34; *P*=.01), while significantly reducing perceived autonomy (mean 3.66, SD 1.03 vs mean 4.06, SD 0.91, t_{231}=−3.17; Cohen *d*=−0.42; *P*=.002). These results support hypotheses 1-3, indicating that stronger algorithmic care increased decisional substitution and perceived surveillance while reducing perceived autonomy.

Table . Manipulation check and hypothesis testing results.

Variable	Low algorithmic care intensity, mean (SD)	High algorithmic care intensity, mean (SD)	*t* test (*df*)	Cohen *d*	*P* value	Conclusion
Perceived algorithmic care intensity (manipulation check)	4.11 (0.91)	4.79 (0.95)	5.63 (231)	0.74	<.001	Manipulation successful
Decisional substitution	4.13 (0.84)	4.43 (0.94)	2.64 (231)	0.35	.009	H2 supported
Perceived surveillance	3.95 (0.95)	4.27 (0.91)	2.58 (231)	0.34	.01	H3 supported
Perceived autonomy	4.06 (0.91)	3.66 (1.03)	−3.17 (231)	−0.42	.002	H1 supported

Table . Bootstrap results for the parallel mediation effects in study 1.

Path	Effect (95% CI)	Conclusion
Algorithmic care intensity → decisional substitution → perceived autonomy	−0.072 (−0.144 to −0.016)	H4a supported
Algorithmic care intensity → perceived surveillance → perceived autonomy	−0.084 (−0.161 to −0.019)	H4b supported
Total indirect effect	−0.156 (−0.262 to −0.066)	Significant
Direct effect	−0.247 (−0.491 to −0.006)	Still significant, partial mediation
Total effect	−0.403 (−0.650 to −0.153)	Significant

XSL•FO
RenderX

We next tested the parallel mediating roles of decisional substitution and perceived surveillance using bootstrap analysis with 5000 resamples. Experimental condition was entered as the independent variable, perceived autonomy as the dependent variable, and decisional substitution and perceived surveillance as parallel mediators. As illustrated in Tables 3 and 4, the indirect effect through decisional substitution was significant (effect=−0.072, 95% CI −0.144 to −0.016), supporting hypothesis 4a. The indirect effect through perceived surveillance was also significant (effect=−0.084, 95% CI −0.161 to −0.019), supporting hypothesis 4b. The total indirect effect was −0.156 (95% CI −0.262 to −0.066). After both mediators were included, the direct effect of algorithmic care intensity on perceived autonomy remained significant (effect=−0.247, 95% CI −0.491 to −0.006), indicating partial mediation.

Study 2: Results of the Community-Based Field Survey

Overview

Study 2 was designed to further test the applicability and robustness of the proposed model in real community smart older adult care settings and to examine the moderating role of digital literacy in the first-stage paths. Given that participants had experience with different smart older care services and devices, the findings should be interpreted as reflecting the perceived psychological consequences of algorithmic care in a heterogeneous real-world service environment rather than the effects of any single technology category.

A total of 298 valid responses were retained for analysis. Responses with substantial missing data on the focal variables were excluded before analysis. The final analyses were conducted on complete cases, and no statistical imputation procedure was applied. The mean age of the participants was 71.8 (SD 6.202) years, and 158 out of 298 (53%) participants were female. In terms of education, 221 out of 298 (74%) participants had completed high school or below, whereas 77 out of 298 (26%) participants had completed junior college or above. Although most participants had completed high school or below, the proportion with junior college education or above and the urban locations of Hangzhou and Ningbo suggest that the sample reflects older adults in relatively developed community smart care settings. This context is substantively appropriate for examining the focal mechanisms of algorithmic care, although the strength and form of these relationships may vary across regions with different levels of digital development and care infrastructure.

Measurement Model

Before hypothesis testing, we examined the descriptive statistics and correlations among the focal variables. Perceived algorithmic care intensity was positively correlated with decisional substitution and perceived surveillance, and negatively correlated with perceived autonomy. Decisional substitution and perceived surveillance were also positively related to each other and both were negatively associated with perceived autonomy.

Because study 2 relied on self-reported questionnaire data, we also assessed the potential risk of common method bias. In addition to the procedural remedies applied during questionnaire design, we conducted Harman's single factor test and compared the proposed 5-factor model with a single factor model. The results indicated that common method bias was unlikely to materially distort the measurement structure.

We next assessed the reliability and validity of the measures. All Cronbach α values exceeded the conventional threshold of 0.70, indicating strong internal consistency. Confirmatory factor analysis showed that the 5-factor measurement model fit the data well, with χ^2_{179}=353.88, comparative fit index=0.947, Tucker-Lewis index=0.938, root-mean-square error of approximation=0.057, and standardized root-mean-square residual=0.049, supporting the structural distinctiveness of perceived algorithmic care intensity, decisional substitution, perceived surveillance, perceived autonomy, and digital literacy. In addition, all standardized factor loadings exceeded 0.80, and both composite reliability and average variance extracted were above recommended thresholds, supporting convergent validity. Detailed results are reported in Table 5.

To further assess discriminant validity, we used the heterotrait-monotrait ratio (HTMT). As illustrated in Table 6, all HTMT values were below 0.85, indicating adequate discriminant validity across constructs.

Table . Measurement model results for study 2.

Constructs	Number of items	Factor loading range	CR[a]	AVE[b]	Cronbach α
Perceived algorithmic care intensity	3	0.812 - 0.895	0.919	0.791	0.920
Decisional substitution	4	0.838 - 0.877	0.919	0.740	0.918
Perceived surveillance	4	0.840 - 0.868	0.916	0.732	0.915
Perceived autonomy	4	0.852 - 0.882	0.923	0.750	0.922
Digital literacy	6	0.841 - 0.901	0.954	0.776	0.954

[a]CR: composite reliability.

[b]AVE: average variance extracted.

XSL•FO
RenderX

Table . Heterotrait-monotrait ratio results for discriminant validity in study 2.

Variable	Perceived algorithmic care intensity	Decisional substitution	Perceived surveillance	Perceived autonomy	Digital literacy
Perceived algorithmic care intensity					
r	1	0.389	0.445	−0.356	0.065
P value	—[a]	<.001	<.001	<.001	.27
Decisional substitution					
r	0.389	1	0.285	−0.425	0.085
P value	<.001	—	<.001	<.001	.14
Perceived surveillance					
r	0.445	0.285	1	−0.463	0.037
P value	<.001	<.001	—	<.001	.52
Perceived autonomy					
r	−0.356	−0.425	−0.463	1	0.045
P value	<.001	<.001	<.001	—	.44
Digital literacy					
r	0.065	0.085	0.037	0.045	1
P value	.27	.14	.52	.44	—

[a]Not applicable.

Tests of Main and Mediating Effects

After confirming the measurement model, we tested the main effects and mediation effects. Consistent with the reported regression specification, age, gender, and education were included as controls in these models. As presented in Table 7, controlling for these variables, perceived algorithmic care intensity significantly and positively predicted decisional substitution (β=.400, SE 0.053; t_{293}=7.550; *P*<.001) and perceived surveillance (β=.440, SE 0.053; t_{293}=8.250; *P*<.001), and significantly and negatively predicted perceived autonomy (β=−.356, SE 0.053; t_{293}=−6.501; *P*<.001), supporting hypotheses 1-3. When decisional substitution and perceived surveillance were entered simultaneously, both negatively predicted perceived autonomy (β=−.294, SE 0.053; t_{291}=−5.533; *P*<.001; β=−.344, SE 0.055; t_{291}=−6.287; *P*<.001), whereas the direct effect of perceived algorithmic care intensity became nonsignificant (β=−.087, SE 0.055; t_{291}=−1.526; *P*=.13), indicating that the effect operated mainly through the 2 mediators. As presented in Table 8, bootstrap analysis further showed significant indirect effects through decisional substitution (effect=−0.112, 95% CI −0.172 to −0.065) and perceived surveillance (effect=−0.149, 95% CI −0.215 to −0.093). The total indirect effect was −0.261 (95% CI −0.346 to −0.189), whereas the direct effect was not significant (effect=−0.085, 95% CI −0.185 to 0.025), indicating full mediation.

Table . Regression results for main and mediating effects in study 2.

Dependent variable	Independent variable	β (SE)	*t* test (*df*)	*P* value	R^2
Decisional substitution	Perceived algorithmic care intensity	.400 (0.053)	7.550 (293)	<.001	0.193
Perceived surveillance	Perceived algorithmic care intensity	.440 (0.053)	8.250 (293)	<.001	0.217
Perceived autonomy (total effect)	Perceived algorithmic care intensity	−.356 (0.053)	−6.501 (293)	<.001	0.130
Perceived autonomy (with mediators)	Perceived algorithmic care intensity	−.087 (0.055)	−1.526 (291)	.13	0.319
Perceived autonomy (with mediators)	Decisional substitution	−.294 (0.053)	−5.533 (291)	<.001	—[a]
Perceived autonomy (with mediators)	Perceived surveillance	−.344 (0.055)	−6.287 (291)	<.001	—

[a]Not applicable.

XSL•FO
RenderX

Table . Bootstrap results for the parallel mediation effects in study 2.

Path	Effect (95% CI)	Conclusion
Perceived algorithmic care intensity → decisional substitution → perceived autonomy	−0.112 (−0.172 to −0.065)	H4a supported
Perceived algorithmic care intensity → perceived surveillance → perceived autonomy	−0.149 (−0.215 to −0.093)	H4b supported
Total indirect effect	−0.261 (−0.346 to −0.189)	Significant
Direct effect	−0.085 (−0.185 to 0.025)	Not significant
Total effect	−0.346 (−0.446 to −0.246)	Significant

Test of the Moderating Role of Digital Literacy

After obtaining support for the main model, we further examined whether digital literacy moderated the effects of algorithmic care intensity on the 2 first-stage paths. To reduce multicollinearity, both perceived algorithmic care intensity and digital literacy were mean-centered before constructing the interaction terms. We then estimated regression models with decisional substitution and perceived surveillance as dependent variables, including the main effects, the interaction term, and the same control variables used in the main effect models, namely age, gender, and education.

As presented in Table 9, the interaction between perceived algorithmic care intensity and digital literacy was significant for decisional substitution (β=−.174, SE 0.055; t_{291}=−3.178; *P*=.002), supporting hypothesis 5a. Simple slope analysis showed that the positive effect of algorithmic care intensity on decisional substitution was strongest at low digital literacy (simple slope=0.560), moderate at the mean level (simple slope=0.398), and weakest at high digital literacy (simple slope=0.235). By contrast, the interaction between perceived algorithmic care intensity and digital literacy was not significant for perceived surveillance (β=−.071, SE 0.053; t_{291}=−1.338; *P*=.18), so hypothesis 5b was not supported. This pattern suggests that digital literacy may buffer the sense that decision making is being replaced by the system, but not sensitivity to continuous monitoring and the accumulation of digital traces.

Table . Moderating effects of digital literacy in study 2.

Dependent variable and predictor	β (SE)	*t* test (*df*)	*P* value	R^2
Decisional substitution				0.193
Perceived algorithmic care intensity	.413 (0.051)	7.763 (291)	<.001	
Digital literacy	.075 (0.054)	1.380 (291)	.17	
Algorithmic care intensity×digital literacy	−.174 (0.055)	−3.178 (291)	.002	
Perceived surveillance				0.214
Perceived algorithmic care intensity	.448 (0.051)	8.534 (291)	<.001	
Digital literacy	−.018 (0.053)	−0.339 (291)	.73	
Algorithmic care intensity×digital literacy	−.071 (0.053)	−1.338 (291)	.18	

Discussion

Principal Results

This study examined how algorithmic care shapes older adults' perceived autonomy in community smart older adult care settings. Using a qualitative prestudy, a vignette experiment, and a community-based field survey, we tested whether perceived algorithmic care intensity reduces perceived autonomy through decisional substitution and perceived surveillance, and whether digital literacy moderates these effects.

The results show that algorithmic care is not merely a neutral arrangement for improving efficiency and safety. When care relies more heavily on system recognition, automated reminders, programmed recommendations, and system-shaped follow-up arrangements, older adults are more likely to feel that decisions are no longer made by them first and that their daily lives are continuously recorded and observed. These 2 experiences, in turn, reduce perceived autonomy. This pattern was supported in both the vignette experiment and the community-based field survey.

The findings also show that decisional substitution and perceived surveillance are 2 distinct pathways. Decisional substitution reflects a perceived weakening of older adults' role in judgment and choice, whereas perceived surveillance reflects the disruption of psychological boundaries caused by continuous

XSL•FO
RenderX

observation and data recording. Digital literacy weakened the effect of algorithmic care intensity on decisional substitution, but not on perceived surveillance. This suggests that digital competence may help older adults interpret system outputs as supportive inputs rather than replacements for their own judgment, but it does not necessarily remove discomfort caused by continuous monitoring and digital traces.

Limitations

This study has several limitations. First, perceived algorithmic care intensity was examined as older adults' overall perception of algorithmic involvement in the care environment. Different technologies may produce different psychological responses. For example, passive monitoring tools may more directly trigger feelings of being watched, whereas user-initiated advisory applications may allow more decisional participation. Future research could compare passive monitoring systems, user-initiated health applications, and hybrid care platforms.

Second, the study 2 sample was drawn from Hangzhou and Ningbo, where digital infrastructure and smart older adult care implementation are relatively developed. The findings may therefore reflect older adults' experiences in urban community smart care settings. Future research could test the model in counties, rural areas, and regions with different levels of digital development and care infrastructure.

Third, this study relied mainly on self-report data. Although self-reports are appropriate for capturing subjective experiences such as autonomy, decisional substitution, and perceived surveillance, they do not provide strong behavioral evidence. Future studies could include platform usage logs, alert response records, family intervention frequency, or actual choices made after system recommendations.

Fourth, although this study combined a qualitative prestudy, an experiment, and a field survey, study 2 was still cross-sectional. It therefore could not capture the longer-term adaptation process through which older adults respond to algorithmic care. Future research could use longitudinal or staged designs to examine whether older adults gradually adapt to algorithmic care, become more dependent on it, or develop stronger resistance over time.

Comparison With Previous Work

Much of the existing smart older adult care literature has focused on technology acceptance, perceived usefulness, usage intention, and implementation barriers [7,8,56]. This work has helped explain whether older adults are willing and able to use digital technologies, but it says less about what happens psychologically after such technologies become embedded in everyday care. Previous work has argued that older adults' digital health experiences are closely tied to autonomy, competence, and relatedness rather than usefulness alone [12]. The findings of this study extend this view by showing that algorithmic care may support daily living while simultaneously weakening older adults' sense of self-direction.

This study also extends previous research on monitoring technologies. Existing studies have shown that older adults' concerns about smart monitoring often go beyond data privacy and include dignity, control, intrusiveness, and the feeling of being watched [33,48]. Research on AI in care settings has similarly noted the tension between efficient care and humane care, especially when automation increases discipline and relational distance [13]. However, these concerns are often discussed together. By distinguishing decisional substitution from perceived surveillance, this study shows that the psychological consequences of algorithmic care are multidimensional. Decisional substitution concerns whether older adults remain part of the process of deciding, whereas perceived surveillance concerns whether they feel constantly seen, recorded, and made legible to the system.

The moderating findings also refine current assumptions about digital inclusion. Policy and practice often assume that improving digital literacy will reduce older adults' difficulties with smart care technologies. Our results only partly support this view. Digital literacy reduced the likelihood that older adults experienced system outputs as replacing their own judgment, but it did not significantly reduce perceived surveillance. This suggests that some autonomy-related risks are interpretive and may be partly buffered by technological understanding, whereas others are experiential and arise from the monitoring logic of the technology itself, especially its intrusion into private boundaries, everyday rhythms, and subjective comfort [5,32].

Conclusions

This study shows that algorithmic care can reduce older adults' perceived autonomy in community smart older adult care settings by increasing decisional substitution and perceived surveillance. It also shows that digital literacy is a selective rather than universal buffer: it weakens decisional substitution but does not significantly reduce perceived surveillance. These findings move smart older adult care research beyond technology acceptance by showing how algorithmic systems can reorganize older adults' judgment, participation, and psychological boundaries in everyday care.

For smart older adult care to support rather than replace older adults, system design and service governance should preserve choice, clarify monitoring boundaries, strengthen opt-out possibilities, and keep human discussion within the care process. System prompts should rely more on suggestive than commanding logic, and care workers or family members should avoid treating algorithmic outputs as final decisions.

Acknowledgments

The authors thank the reviewers and the manuscript editor for their valuable time, constructive comments, and support during the review process.

XSL•FO
RenderX

The authors declare the use of generative AI (GenAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: translation. The GenAI tool used was ChatGPT5.4 (OpenAI). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

Funding Support of Zhejiang Philosophy and Social Sciences Planning Project (24NDQN035YB).

Data Availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
participant characteristics, first-order themes identified from semi-structured interviews, and measurement items.
[DOCX File, 18 KB - aging_v9i1e94559_app1.docx]

References

1. Shi J, Zhang N, Wu K, Wang Z. Application status, challenges, and development prospects of smart technologies in home-based elder care. Electronics (Basel) 2025;14(12):2463. [doi: 10.3390/electronics14122463]
2. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
3. Ho A. Are we ready for artificial intelligence health monitoring in elder care? BMC Geriatr 2020 Sep 21;20(1):358. [doi: 10.1186/s12877-020-01764-9] [Medline: 32957946]
4. Percy Campbell J, Buchan J, Chu CH, Bianchi A, Hoey J, Khan SS. User perception of smart home surveillance among adults aged 50 years and older: scoping review. JMIR Mhealth Uhealth 2024;12:e48526. [doi: 10.2196/48526]
5. Kruithof K, Ruehle RC, Dörenberg V, van Etten-Jamaludin F, Frederiks B. Digital assistive technologies to support community-dwelling older adults: an interpretative synthesis of experienced benefits and harms. SSM - Qualitative Research in Health 2025 Dec;8:100584. [doi: 10.1016/j.ssmqr.2025.100584]
6. Hung J. Smart elderly care services in China: challenges, progress, and policy development. Sustainability 2023;15(1):178. [doi: 10.3390/su15010178]
7. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6(1):e44564. [doi: 10.2196/44564] [Medline: 37191976]
8. Martín-García AV, Redolat R, Pinazo-Hernandis S. Factors influencing intention to technological use in older adults. The TAM model aplication. Res Aging 2022;44(7-8):573-588. [doi: 10.1177/01640275211063797] [Medline: 34962846]
9. Groeneveld SWM, van Os-Medendorp H, van Gemert-Pijnen J, et al. Essential competencies of nurses working with AI-driven lifestyle monitoring in long-term care: a modified Delphi study. Nurse Educ Today 2025 Jun;149:106659. [doi: 10.1016/j.nedt.2025.106659] [Medline: 40056483]
10. Kumar V, Jain P. Transforming healthcare delivery: the role of digital technologies in disease prevention and health monitoring. In: Reshaping Health Promotion and Disease Prevention through Digital Innovation: IGI Global Scientific Publishing; 2026:255-298. [doi: 10.4018/979-8-3373-3531-5.ch009]
11. Khalil RA, Ahmad K, Ali H. Redefining elderly care with agentic AI: challenges and opportunities. IEEE Open J Comput Soc 2026;7:326-342. [doi: 10.1109/OJCS.2026.3650842]
12. Cotter LM, Shah D, Brown K, et al. Decoding the influence of eHealth on autonomy, competence, and relatedness in older adults: qualitative analysis of self-determination through the motivational technology model. JMIR Aging 2024 Oct 30;7:e56923. [doi: 10.2196/56923] [Medline: 39476377]
13. Joseph J. Designing for dignity: ethics of AI surveillance in older adult care. Front Digit Health 2025;7:1643238. [doi: 10.3389/fdgth.2025.1643238] [Medline: 40919326]
14. Neves BB, Omori M, Petersen A, Vered M, Carter A. Navigating artificial intelligence in care homes: competing stakeholder views of trust and logics of care. Soc Sci Med 2024 Oct;358:117187. [doi: 10.1016/j.socscimed.2024.117187]
15. Ryan RM, Deci EL. Self-determination theory. In: Maggino F, editor. Encyclopedia of Quality of Life and Well-Being Research: Springer; 2023. [doi: 10.1007/978-3-031-17299-1_2630]
16. Deci EL, Olafsen AH, Ryan RM. Self-determination theory in work organizations: the State of a Science. Annu Rev Organ Psychol Organ Behav 2017 Mar 21;4(1):19-43. [doi: 10.1146/annurev-orgpsych-032516-113108]

XSL•FO
RenderX

17. Ryan RM, Duineveld JJ, Di Domenico SI, Ryan WS, Steward BA, Bradshaw EL. We know this much is (meta-analytically) true: a meta-review of meta-analytic findings evaluating self-determination theory. Psychol Bull 2022;148(11-12):813-842. [doi: 10.1037/bul0000385]
18. Gnädinger M, Herzig L, Ceschi A, et al. Chronic conditions and multimorbidity in a primary care population: a study in the Swiss Sentinel Surveillance Network (Sentinella). Int J Public Health 2018 Dec;63(9):1017-1026. [doi: 10.1007/s00038-018-1114-6] [Medline: 29786762]
19. Dambha-Miller H, Simpson G, Hobson L, et al. Integrated primary care and social services for older adults with multimorbidity in England: a scoping review. BMC Geriatr 2021 Dec 3;21(1):674. [doi: 10.1186/s12877-021-02618-8] [Medline: 34861831]
20. Yi YM, Park YH, Cho B, et al. Development of a community-based integrated service model of health and social care for older adults living alone. Int J Environ Res Public Health 2021 Jan 19;18(2):825. [doi: 10.3390/ijerph18020825] [Medline: 33478027]
21. Clancy A, Simonsen N, Lind J, Liveng A, Johannessen A. The meaning of dignity for older adults: a meta-synthesis. Nurs Ethics 2021 Sep;28(6):878-894. [doi: 10.1177/0969733020928134] [Medline: 32613895]
22. Knotnerus HR, Ngo HTN, Maarsingh OR, van Vugt VA. Understanding older adults' experiences with a digital health platform in general practice: qualitative interview study. JMIR Aging 2024 Aug 30;7(1):e59168. [doi: 10.2196/59168] [Medline: 39212599]
23. Barrett M, Boyne J, Brandts J, et al. Artificial intelligence supported patient self-care in chronic heart failure: a paradigm shift from reactive to predictive, preventive and personalised care. EPMA J 2019 Dec;10(4):445-464. [doi: 10.1007/s13167-019-00188-9] [Medline: 31832118]
24. Ng CJ, Lee PY, Lee YK, et al. An overview of patient involvement in healthcare decision-making: a situational analysis of the Malaysian context. BMC Health Serv Res 2013 Oct 11;13(1):408. [doi: 10.1186/1472-6963-13-408] [Medline: 24119237]
25. Bouniols N, Leclère B, Moret L. Evaluating the quality of shared decision making during the patient-carer encounter: a systematic review of tools. BMC Res Notes 2016 Aug 2;9:382. [doi: 10.1186/s13104-016-2164-6] [Medline: 27485434]
26. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc 2012;19(1):121-127. [doi: 10.1136/amiajnl-2011-000089] [Medline: 21685142]
27. Torenholt R, Langstrup H. Between a logic of disruption and a logic of continuation: Negotiating the legitimacy of algorithms used in automated clinical decision-making. Health (London) 2023 Jan;27(1):41-59. [doi: 10.1177/1363459321996741] [Medline: 33685260]
28. Deshpande LK. AI for personalized healthcare: predicting risk and recommending the right care. 2025 Presented at: International Conference on Computer Science, Artificial Intelligence, Machine Learning p. 111-117. [doi: 10.56472/ICCSAIML25-113]
29. Moldt JA, Festl-Wietek T, Fuhl W, et al. Exploring the social dimensions of AI integration in healthcare: a qualitative study of stakeholder views on challenges and opportunities. BMJ Open 2025 Jun 27;15(6):e096208. [doi: 10.1136/bmjopen-2024-096208] [Medline: 40578890]
30. Lukkien DRM, Stolwijk NE, Ipakchian Askari S, et al. AI-assisted decision-making in long-term care: qualitative study on prerequisites for responsible innovation. JMIR Nurs 2024 Jul 25;7(1):e55962. [doi: 10.2196/55962] [Medline: 39052315]
31. Aldenhoven CM, Milton M, D'Ambrosio L, Lee C, Ashebir S, André E. An accessible multifunctional system to support safe and independent aging in place: iterative development and qualitative analysis. JMIR Aging 2025 Oct 16;8:e72579. [doi: 10.2196/72579] [Medline: 41100691]
32. Mujirishvili T, Maidhof C, Florez-Revuelta F, Ziefle M, Richart-Martinez M, Cabrero-García J. Acceptance and privacy perceptions toward video-based active and assisted living technologies: scoping review. J Med Internet Res 2023 May 1;25:e45297. [doi: 10.2196/45297] [Medline: 37126390]
33. Zhang W, Yin J, Chan KI, et al. Beyond digital privacy: uncovering deeper attitudes toward privacy in cameras among older adults. Int J Hum Comput Stud 2024 Dec;192:103345. [doi: 10.1016/j.ijhcs.2024.103345]
34. Mondal R. Ethical consideration in assisted living technology in healthcare. In: Ambient Assisted Living: CRC Press; 2025:309-327. [doi: 10.1201/9781003570134-17]
35. Grabler R, Koeszegi ST. Privacy beyond data: assessment and mitigation of privacy risks in robotic technology for elderly care. J Hum-Robot Interact 2025 Mar 31;14(1):1-23. [doi: 10.1145/3689216]
36. Shaik T, Tao X, Higgins N, et al. Remote patient monitoring using artificial intelligence: current state, applications, and challenges. WIREs Data Min & Knowl 2023 Mar;13(2):e1485. [doi: 10.1002/widm.1485]
37. van Kasteren Y, Bradford D, Zhang Q, Karunanithi M, Ding H. Understanding smart home sensor data for ageing in place through everyday household routines: a mixed method case study. JMIR Mhealth Uhealth 2017 Jun 13;5(6):e52. [doi: 10.2196/mhealth.5773] [Medline: 28611014]
38. Parkinson ME, Smith RM, Tanious K, et al. Experiences with home monitoring technology in older adults with traumatic brain injury: a qualitative study. BMC Geriatr 2024 Sep 30;24(1):796. [doi: 10.1186/s12877-024-05397-0] [Medline: 39350122]

XSL•FO
RenderX

39. Torres-Hernandez CM, Garduño-Aparicio M, Rodriguez-Resendiz J. Smart homes: a meta-study on sense of security and home automation. Technologies (Basel) 2025;13(8):320. [doi: 10.3390/technologies13080320]
40. Stabler L, Cunningham E, Mannay D, et al. "I probably wouldn't want to talk about anything too personal": a qualitative exploration of how issues of privacy, confidentiality and surveillance in the home impact on access and engagement with online services and spaces for care-experienced young people. Adopt Foster 2023 Oct;47(3):277-294. [doi: 10.1177/03085759231203019] [Medline: 37873026]
41. Fowe IE, Boot WR. Understanding older adults' attitudes toward mobile and wearable technologies to support health and cognition. Front Psychol 2022;13:1036092. [doi: 10.3389/fpsyg.2022.1036092] [Medline: 36571059]
42. Hawley-Hague H, Boulton E, Hall A, Pfeiffer K, Todd C. Older adults' perceptions of technologies aimed at falls prevention, detection or monitoring: a systematic review. Int J Med Inform 2014 Jun;83(6):416-426. [doi: 10.1016/j.ijmedinf.2014.03.002] [Medline: 24798946]
43. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367. [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
44. Mack DP, Greenhawt M, Bukstein DA, Golden DBK, Settipane RA, Davis RS. Decisions with patients, not for patients: shared decision-making in allergy and immunology. J Allergy Clin Immunol Pract 2024 Oct;12(10):2625-2633. [doi: 10.1016/j.jaip.2024.05.046] [Medline: 38851489]
45. Giap TTT, Park M, Jung J, Kim J, Doan TTT. Shared decision-making in long-term care: a systematic review of assessment instruments. Geriatric Nursing 2024 Nov;60:156-176. [doi: 10.1016/j.gerinurse.2024.08.020]
46. Abdelwanis M, Alarafati HK, Tammam MMS, Simsekler MCE. Exploring the risks of automation bias in healthcare artificial intelligence applications: a bowtie analysis. Journal of Safety Science and Resilience 2024 Dec;5(4):460-469. [doi: 10.1016/j.jnlssr.2024.06.001]
47. Expósito-Jiménez A, Alcaide-Leyva JM, Jiménez-Mérida MDR, Martínez-Angulo P. Health communication and shared decision-making between nurses and older adults in community setting: an integrative review. J Clin Nurs 2024 Aug;33(8):2922-2935. [doi: 10.1111/jocn.17152] [Medline: 38573001]
48. Tham NAQ, Brady AM, Ziefle M, Dinsmore J. Barriers and facilitators to older adults' acceptance of camera-based active and assisted living technologies: a scoping review. Innov Aging 2025;9(2):igae100. [doi: 10.1093/geroni/igae100] [Medline: 39968358]
49. Oh SS, Kim KA, Kim M, Oh J, Chu SH, Choi J. Measurement of digital literacy among older adults: systematic review. J Med Internet Res 2021 Feb 3;23(2):e26145. [doi: 10.2196/26145] [Medline: 33533727]
50. Miller LMS, Callegari RA, Abah T, Fann H. Digital literacy training for low-income older adults through undergraduate community-engaged learning: single-group pretest-posttest study. JMIR Aging 2024 May 14;7:e51675. [doi: 10.2196/51675] [Medline: 38599620]
51. van Deursen A, Helsper EJ, Eynon R. Development and validation of the internet skills scale (ISS). Information, Communication & Society 2016 Jun 2;19(6):804-823. [doi: 10.1080/1369118X.2015.1078834]
52. Reddy P, Sharma B, Chaudhary K. Digital literacy: a review of literature. Int J Technoethics 2020;11(2):65-94. [doi: 10.4018/IJT.20200701.oa1]
53. Hargittai E. Digital na(t)ives? Variation in internet skills and uses among members of the "Net Generation". Sociol Inq 2010 Feb;80(1):92-113. [doi: 10.1111/j.1475-682X.2009.00317.x]
54. Kyaw MY, Aung MN, Koyanagi Y, et al. Sociodigital determinants of eHealth literacy and related impact on health outcomes and eHealth use in Korean older adults: community-based cross-sectional survey. JMIR Aging 2024 Aug 13;7:e56061. [doi: 10.2196/56061] [Medline: 39140239]
55. Livingstone S, Helsper E. Gradations in digital inclusion: children, young people and the digital divide. New Media & Society 2007 Aug;9(4):671-696. [doi: 10.1177/1461444807080335]
56. Jaana M, Tamim H, Paré G. eHealth and the digital divide among older Canadians: insights from a national cross-sectional study. J Med Internet Res 2025 Nov 25;27:e72274. [doi: 10.2196/72274] [Medline: 41289575]
57. Leong QY, Lee VV, Ng WY, et al. Older adults' perspectives and experiences with digital health in Singapore: qualitative study. JMIR Hum Factors 2024 Nov 11;11(1):e58641. [doi: 10.2196/58641] [Medline: 39527787]
58. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]

Abbreviations

HTMT: heterotrait-monotrait ratio

XSL•FO
RenderX

Edited by AF Abdulai; submitted 03.Mar.2026; peer-reviewed by F Yang, L Dai; revised version received 25.Apr.2026; accepted 12.Jun.2026; published 17.Jul.2026.

Please cite as:
Huang X
Perceived Algorithmic Care Intensity and Older Adults' Autonomy in Community Smart Care: Mixed Methods Study
JMIR Aging 2026;9:e94559
URL: https://aging.jmir.org/2026/1/e94559
doi:10.2196/94559

© Xianghui Huang. Originally published in JMIR Aging (https://aging.jmir.org), 17.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Preparation for Future Care in Community-Dwelling Older Adults With Prefrailty and Frailty: Development and Pilot Test of a Digital Intervention

Qiqi Ni*, RN, MD; Jing Dong*, PhD; Weilin Jiang, MM; Jia Yi, BS; Ran Yan, BS; Yuhan Chen, BS; Li Wang, MM; Yiyu Zhuang, MSN

Nursing Department, Sir Run Run Shaw Hospital, No. 3, Qingchun East Road, Shangcheng District, Hangzhou, Zhejiang, China

*these authors contributed equally

Corresponding Author:
Yiyu Zhuang, MSN
Nursing Department, Sir Run Run Shaw Hospital, No. 3, Qingchun East Road, Shangcheng District, Hangzhou, Zhejiang, China

Abstract

Background: China's rapidly aging population and high burden of frailty make proactive preparation for future care increasingly urgent, yet few older adults engage in such preparation. Existing interventions often overlook the heterogeneity between prefrail and frail populations and lack precision-oriented digital strategies.

Objective: Building on a series of prior empirical studies, this study aimed to develop and pilot-test a digital intervention to support preparation for future care among community-dwelling older adults with prefrailty and frailty.

Methods: The "Yi Yang Plan," a WeChat mini-program, was developed through a rigorous multiphase process, including a scoping review, a convergent mixed methods study, and a structural equation model. These findings informed the design of a tailored, stage-specific intervention targeting the distinct needs of prefrail and frail older adults. An interdisciplinary team subsequently refined the intervention using an iterative design approach. The intervention modules comprised 4 processes: Awareness Enhancement, Care Resources, Care Decision-Making, and Care Planning. Pilot testing involved expert panel consultations and think-aloud tests. Expert characteristics, engagement, and authority coefficients were assessed, and feedback was synthesized using content analysis. Feasibility and acceptability among older adults and their family members were evaluated through task completion metrics, satisfaction surveys, think-aloud protocols, and semistructured interviews.

Results: Seven experts participated in the consultation, with an engagement coefficient of 100% and an authority coefficient of 0.88. Recommendations were synthesized into four key themes: (1) establishing mechanisms for care plan updating and review, (2) strengthening user-centered design, (3) implementing dynamic resource management and robust data security measures, and (4) enhancing integration with community care systems and policy frameworks. A total of 20 older adults and 20 family members were recruited. All participants successfully completed the assigned tasks, with a mean completion time of 20.7 (SD 5.2) minutes. Satisfaction ratings were generally favorable. The qualitative findings indicated that the intervention was perceived as useful and professionally designed, while also identifying several challenges, including variability in preparation for future care readiness and educational levels, limited interactivity, insufficient practical content, and reliance on support from adult children. Suggested improvements included enhanced personalization, additional supportive tools, improved usability, greater involvement of adult children, and stronger integration with offline services.

Conclusions: The "Yi Yang Plan" demonstrated preliminary scientific validity, feasibility, and acceptability as a multidisciplinary, collaboratively developed digital intervention to support preparation for future care among older adults with prefrailty and frailty. By translating prior empirical and theoretical findings into a differentiated, precision-oriented digital strategy, this study advances interventions beyond conventional one-size-fits-all approaches. Future iterations should focus on optimizing system architecture, user interface design, platform functionality, and implementation strategies. Further research should conduct higher-quality randomized controlled trials to evaluate the platform's effectiveness.

(JMIR Aging 2026;9:e92492) doi:10.2196/92492

KEYWORDS

preparation for future care; community-dwelling; digital intervention; frailty; pilot test

XSL•FO
RenderX

Introduction

In China,[1] older adults are defined as those aged 60 years and above [2]. By the end of 2024, the population of older adults in China reached approximately 310 million, accounting for 22% of the total population [3]. By 2035, older adults are expected to account for more than 30% of the population [4]. With advancing age, physiological reserve capacity progressively declines or becomes dysregulated, leading to increased vulnerability, reduced resilience to stressors, and impaired maintenance of internal homeostasis; consequently, older adults are more likely to develop frailty [5,6]. Frailty is a complex geriatric syndrome characterized by multifactorial etiologies and diverse influencing factors. Its hallmark features include reduced muscle strength, decreased endurance, and progressive physiological functional decline, which collectively increase the risks of dependency and mortality [7,8]. Evidence indicates that frailty may be associated with a 15% to 50% increase in all-cause mortality risk [9] and that for every 0.1-unit increase in the frailty index, the risk of adverse health outcomes rises by 68% [10]. Epidemiological studies suggest that the prevalence of frailty is approximately 15% among adults aged 65 years and older, increasing to 25% to 50% among those aged 80 years and above [11-13]. A systematic review further reported that the prevalence of frailty among community-dwelling older adults in China ranges from 5.9% to 17.4%, with a pooled prevalence of 10.0% [14]. The progression of frailty leads to a gradual decline in older adults' ability to live independently and may ultimately result in disability, necessitating reliance on external care and imposing substantial caregiving burdens on families and society [15]. Projections indicate that by 2050, approximately 91.4 million older adults in China will require long-term care services [16], reflecting a growing and sustained demand for formal care provision.

Under the influence of the cultural norm of filial piety, the traditional belief of "raising children for old-age support," and legal as well as moral regulations mandating adult children's caregiving responsibilities, older adults in China have traditionally relied predominantly on family-based care models [17,18]. However, as a result of the long-standing one-child policy, family structures have undergone profound changes, with nearly 70% of households nationwide becoming single-child families [19]. At present, the parents of the first generation of only children have entered old age, and over the next 3 decades, the number of older and oldest-old parents with a single child is expected to continue to increase. China is therefore approaching a peak period of "only-child-supported aging" [20]. Nevertheless, only children face substantial practical challenges in fulfilling eldercare responsibilities. Due to work-related pressures and obligations related to raising the next generation, they often struggle to provide adequate financial support, daily care, and emotional companionship for their aging parents [21]. In addition, with accelerating modernization and urbanization, increasing female labor force participation, and rising life expectancy among older adults, the traditional family caregiving function has been further weakened. Against this backdrop, older adults are increasingly compelled to shift from traditional family-based care models to transitional or more formalized, socially provided care arrangements [22]. This transition necessitates that older adults proactively plan for future care to mitigate potential caregiving risks. Taken together, in the face of the complex care needs associated with frailty, the traditional family care model is no longer sufficient to meet current demands. There is an urgent need for frail older adults to play a more active role by engaging in early planning, mobilizing available resources, and preparing for future care.

Preparation for future care (PFC) refers to the process of preparing for potential needs for assistance with activities of daily living resulting from functional decline due to chronic illness or health-related events in the future [23]. A substantial body of literature indicates that PFC promotes active aging [24]. Older adults who engage in more systematic planning for future care report a greater sense of control over their lives [25], higher life satisfaction [26], and lower levels of anxiety and depression [27]. Moreover, as a form of long-term risk management, PFC enables advance arrangements for future care needs, such as identifying potential caregivers, selecting care institutions, or planning for transitions in care modalities, which can substantially alleviate caregiver burden and generate benefits for caregivers as well [24]. In contrast, a lack of PFC is associated with higher risks of illness and increased financial and emotional strain [28]. Therefore, actively guiding older adults to engage in PFC is a critical component of community nursing practice in the context of population aging, with significant practical relevance and social value.

Over the past 2 decades, most interventions have focused on short-term care planning for patients at the end of life, with limited attention given to long-term care preparation among older adults [29,30]. Existing PFC interventions are predominantly delivered through face-to-face communication or group meetings and rely largely on experiential guidance, often lacking robust empirical evidence and theoretical frameworks. In addition, many studies are characterized by methodological limitations and poor scalability and feasibility for widespread implementation [31]. Consequently, developing scientifically grounded, feasible, and easily disseminated strategies to promote PFC has become a key challenge in this field. With the rapid advancement of information technology, digital nursing models have emerged as an important area of research. For example, Hoffman et al [32] developed a web-based platform that provides education, values clarification, resources, and behavioral planning for long-term care, which effectively facilitated PFC engagement among older adults and has been adopted by health care organizations. However, most existing digital PFC interventions have been developed in Western countries, particularly the United States [33]. These platforms often lack health care professional–led guidance and are not grounded in explicit theoretical frameworks, limiting their capacity to address the individualized needs of frail older adults [34].

In recent years, the Chinese government has issued a series of policy documents [35-37] emphasizing the development of nursing information networks and service platforms to deliver efficient and accessible digital nursing services to older adults in need. In light of both practical demands and policy priorities, there is an urgent need to develop a theory- and

XSL•FO
RenderX

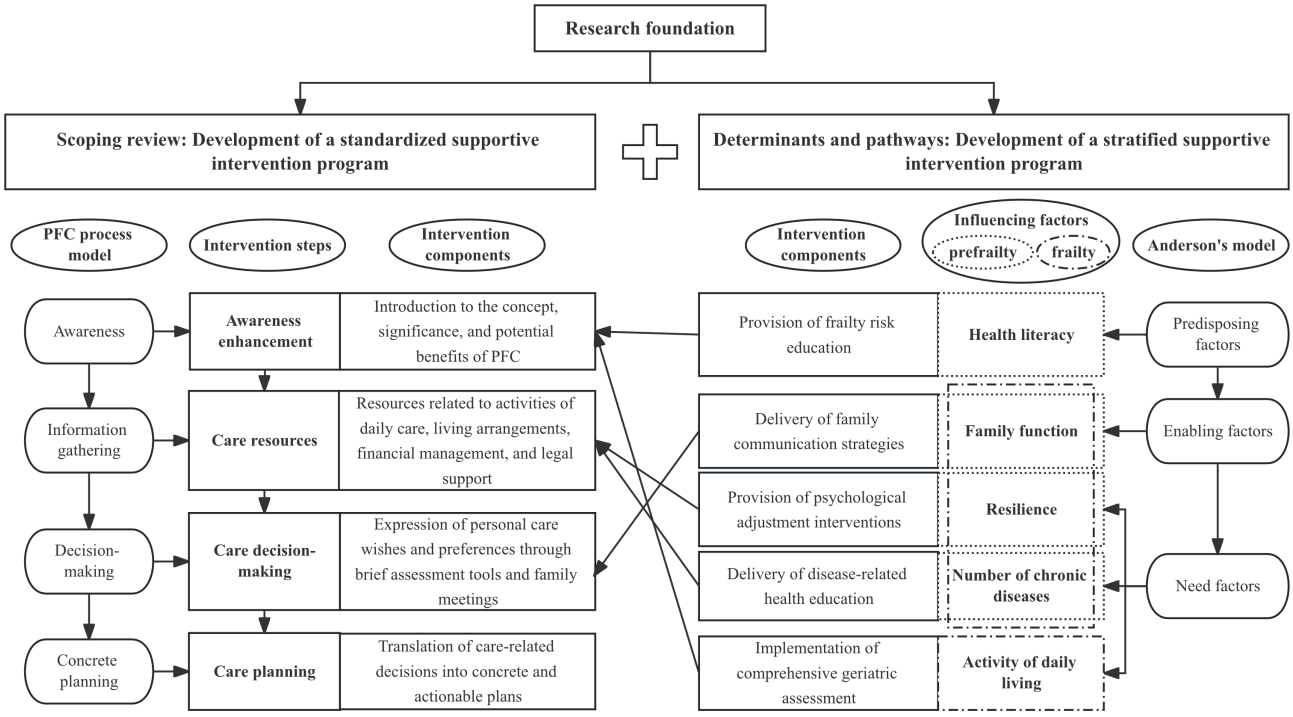
evidence-informed digital PFC intervention platform tailored to the Chinese context. Such a platform could enhance frail older adults' awareness of the importance of PFC and facilitate the implementation of practical long-term care plans, thereby optimizing the prospective allocation of aging-related resources. Therefore, the objectives of this study are twofold: (1) to develop a localized and stratified digital intervention for supporting PFC among community-dwelling older adults with prefrailty and frailty and (2) to assess the scientific validity, feasibility, and acceptability of this digital intervention using expert panel consultations and think-aloud tests. These efforts will provide a foundation for future refinements of the platform and a subsequent randomized controlled trial.

Methods

Digital Program Development: Research Foundation

To inform the development of a digital intervention for PFC among community-dwelling older adults with frailty, we conducted a series of sequential studies, including a scoping review [33], a convergent mixed methods study [38], and a structural equation model (SEM) based on triangulated findings from the mixed methods study. In brief, the scoping review identified 4 core components of PFC-supportive interventions (Awareness Enhancement, Care Resources, Care Decision-Making, and Care Planning), which are consistent with the PFC process model proposed by Sörensen et al [39]. The mixed methods study, guided by the Andersen Behavioral Model of Health Services Use, identified key determinants of PFC among prefrail and frail older adults. Furthermore, the SEM revealed that influencing pathways of PFC differ between prefrail and frail older adults, providing empirical support for a stratified intervention design (Figure 1).

Figure 1. Research foundation of digital intervention development. PFC: preparation for future care.

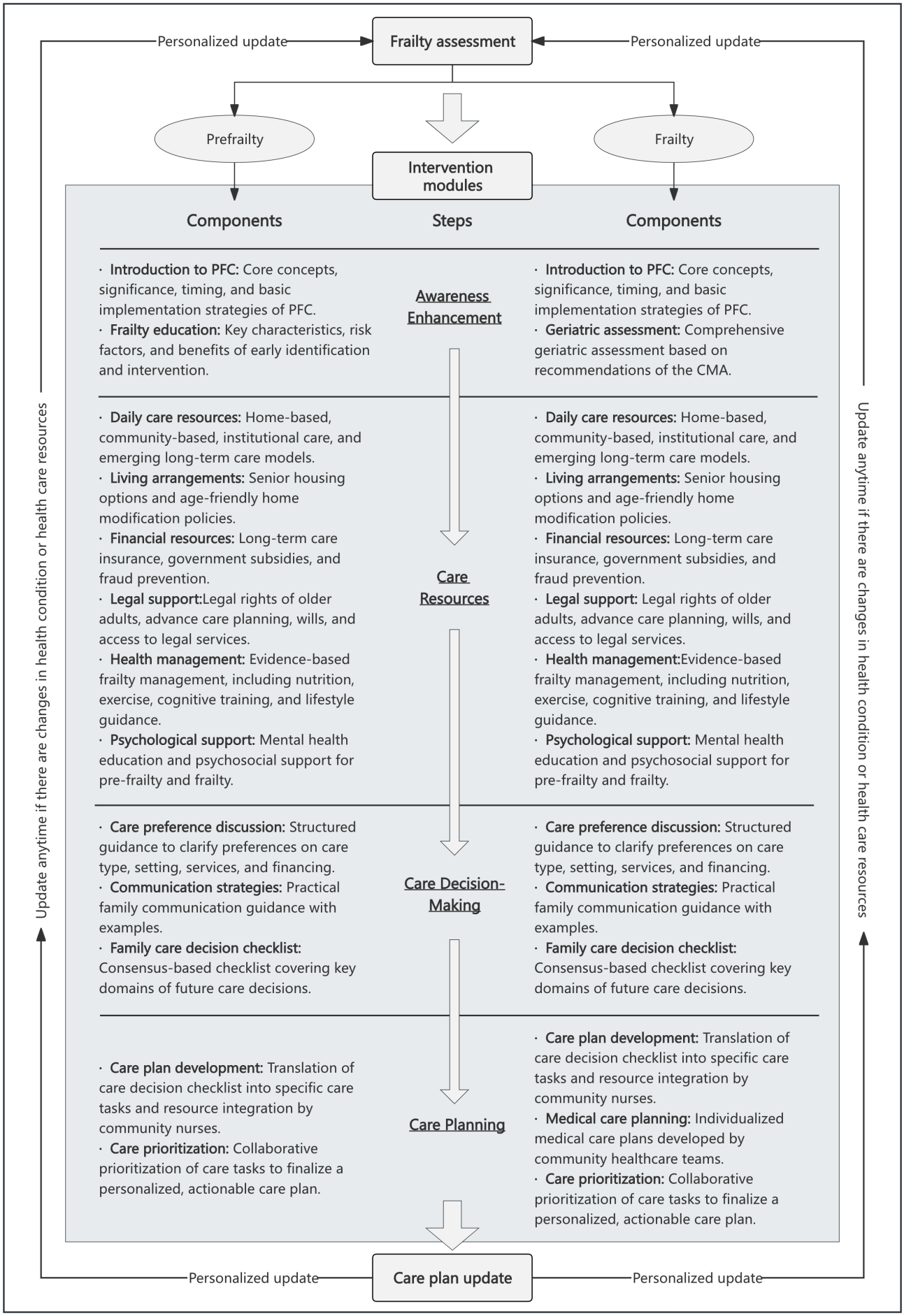


Based on these findings, a supportive intervention program for PFC among community-dwelling frail older adults was developed (Figure 2). The program emphasizes tailoring intervention strategies according to frailty status to enable targeted and differentiated support. Specifically, for prefrail older adults, interventions focus on frailty education, health management, psychological support, and family communication. For frail older adults, interventions focus on comprehensive geriatric assessment, health management, psychological support, and medical care planning.

XSL•FO
RenderX

Figure 2. Supportive intervention program for preparation for future care among community-dwelling prefrail and frail older adults. CMA: Chinese Medical Association; PFC: preparation for future care.

XSL•FO
RenderX



XSL•FO
RenderX

Digital Program Construction

Platform Selection and Conceptualization

Based on the intervention program, the digital platform was developed as a WeChat mini-program, given that WeChat is the most widely used mobile app in China and offers high accessibility for older adults. The mini-program was named "Yi Yang Plan." In Chinese, "Yi Yang" refers to maintaining physical and mental well-being through appropriate care, reflecting the program's goal of promoting healthy and meaningful aging.

Research Team Formation

An interdisciplinary team was established, comprising 2 geriatric care experts with experience in digital health app design, 1 interface designer specializing in age-friendly design, 1 software engineer, and 1 doctoral student in nursing. The team collaboratively conducted the design, development, and refinement of the mini-program.

Interface Design and Development Process

The platform interface was designed using an iterative approach. Initially, the interface designer created preliminary prototypes based on the intervention protocol. These prototypes were reviewed by the research team on a page-by-page basis, followed by iterative revisions focusing on content clarity, navigation logic, and user experience. After multiple rounds of revisions and optimization, the finalized design was implemented by the software engineer. The interface design was developed using the Lanhu platform (version 6.0.4; Beijing Jinwei Zhiguang Information Technology Co; Multimedia Appendix 1).

Development Environment

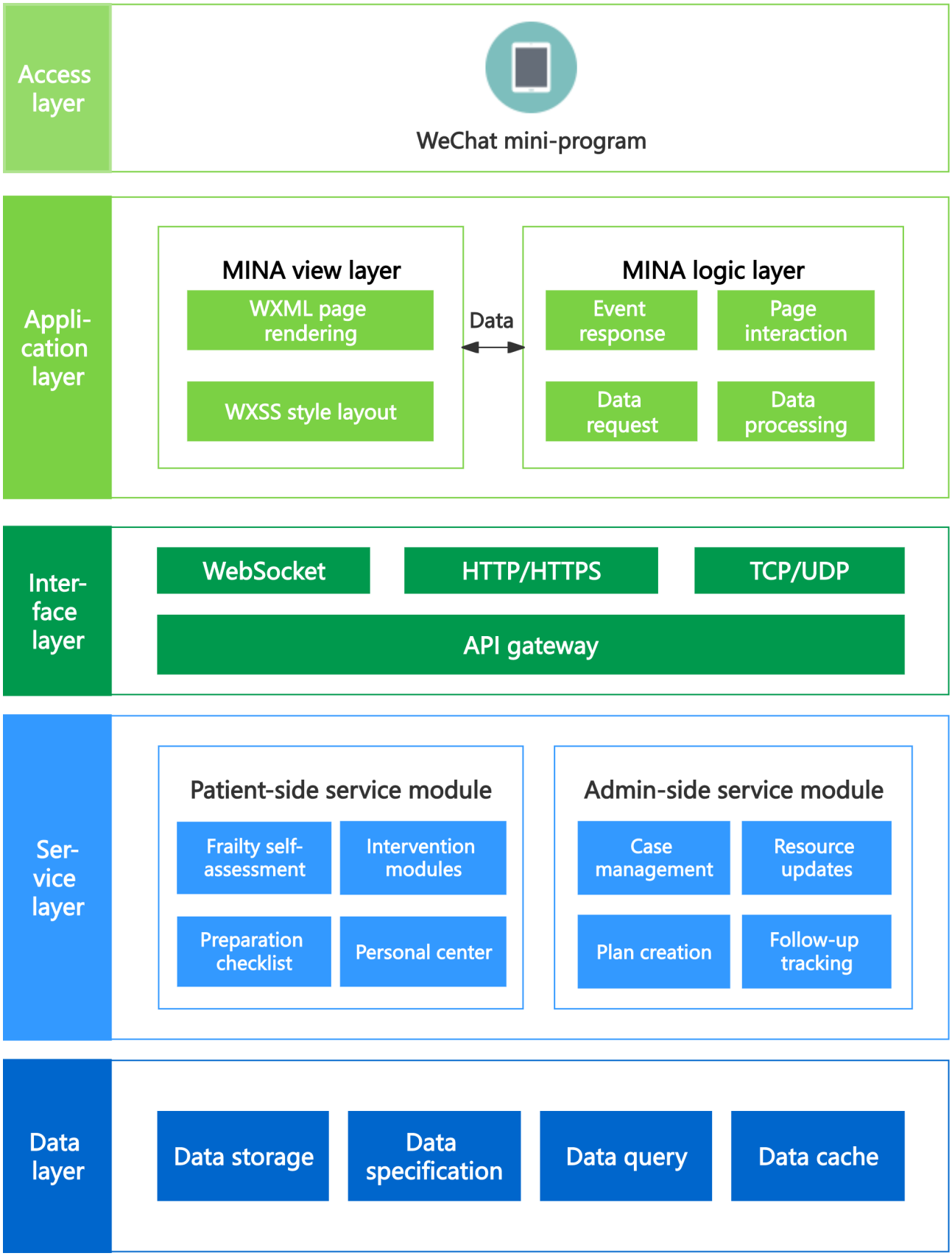
The mini-program was developed within the official WeChat Mini Program development environment. A developer account was registered, and a unique application identification was obtained. The WeChat Developer Tools were used to support coding, debugging, and deployment, with access to standardized frameworks, components, and technical documentation. Upon completion, the software engineer uploaded and managed the program code.

System Architecture

The system adopts a front-end or back-end separated architecture. The front-end was developed using the uni-app framework with uView components, while the back-end was implemented using C#-based server-side services. Data exchange between the 2 layers is achieved through network communication protocols. The platform architecture is organized into 5 hierarchical layers: Data Layer, Service Layer, Interface Layer, Application Layer, and Access Layer. This layered architecture enhances system scalability, maintainability, and security (Figure 3).

XSL•FO
RenderX

Figure 3. System architecture of the "Yi Yang Plan" mini-program. MINA: the WeChat Mini Program; TCP: Transmission Control Protocol; UDP: User Datagram Protocol; WXML: WeiXin Markup Language; WXSS: WeiXin Style Sheets.



Cognitive Load-Oriented Design

Complex content, such as financial and legal information, was simplified through multimodal presentation formats. Specifically, short videos were used to explain key concepts, text-to-speech functionality was incorporated for essential

XSL•FO
RenderX

content, and illustrative images were added to facilitate comprehension.

Digital Platform Functions

Overview

The "Yi Yang Plan" consists of 2 primary components: a community management terminal and a patient or family app terminal. The community management terminal supports case registration, list management, care plan development, resource updates, and follow-up tracking (Multimedia Appendix 2). The patient or family app terminal comprises 4 main modules: self-assessment, intervention modules, care preparation checklist, and personal center.

Self-Assessment

After account registration and binding (Multimedia Appendix 3A), users complete a frailty self-assessment (Multimedia Appendix 3B). Based on the results, users are classified as prefrail or frail, and a corresponding personalized intervention pathway is initiated (Multimedia Appendix 3C).

Intervention Modules

Guided by the PFC process model [39], the intervention is operationalized through four sequential and functionally integrated modules:

Awareness Enhancement

This module introduces the concept, importance, timing, and strategies of PFC through standardized short videos (Multimedia Appendix 4A). For prefrail users, targeted frailty education is provided, focusing on risk factors and early intervention (Multimedia Appendix 4B). For frail older adults, the module incorporates an electronic comprehensive geriatric assessment recommended by the Chinese Medical Association [40], including the Simplified Nutritional Appetite Questionnaire, SARC-F (Strength, Assistance in walking, Rise from a chair, Climb stairs, Falls), and Rapid Cognitive Screen (Multimedia Appendix 4C).

Care Resources

This module provides a structured repository of evidence-informed care resources (Multimedia Appendix 5A). Resources include daily care services, housing and home modification policies, financial planning, legal support, health management, and psychological support (Multimedia Appendix 5B-P). Information is delivered in multiple formats, including text, images, audio, and video.

Care Decision-Making

This module facilitates shared decision-making and family engagement. Structured discussion guides (Multimedia Appendix 6A) and communication strategies (Multimedia Appendix 6B) support family discussions. Users then complete a 12-domain family care decision checklist (Multimedia Appendix 7), reflecting consensus among family members (Multimedia Appendix 6C).

Care Planning

Based on the completed checklist, community nurses develop an initial care plan using the management terminal. For frail older adults, individualized medical care plans are additionally formulated. Users review, adjust, and prioritize care tasks within the mini-program, resulting in a finalized personalized care plan (Multimedia Appendix 8A-B). The system also supports ongoing plan updates (Multimedia Appendix 8C).

Care Preparation Checklist

Upon the completion of the intervention modules, the system automatically generates a comprehensive care preparation checklist integrating care plans and related resources. This checklist can be exported in PDF format for family communication or offline consultations (Multimedia Appendix 9A).

Personal Center

The Personal Center module allows users to manage their personal profiles and maintain relevant information. It also includes a "favorites" function, enabling users to save and access important resources for future reference (Multimedia Appendix 9B).

Expert Panel Consultations

Participants

A purposive sampling method was used to recruit experts from 6 fields: nursing management, community nursing, geriatric nursing, general medicine, geriatric medicine, and sociology. Experts were eligible if they met the following criteria: (1) demonstrated recognized academic expertise and authority in the treatment, care, or long-term care of community-dwelling frail older adults; (2) had experience in leading or participating in digital health-related projects; (3) held an associate senior professional title or higher; and (4) had at least 10 years of relevant professional experience.

Procedure

Prior to the consultation, all experts received the study protocol, a trial access link to the mini-program, and the meeting agenda. The consultation was conducted in a hospital conference room. At the beginning of the session, the project leader introduced the study background and objectives. With participants' informed consent, the entire session was audio-recorded. Experts were then guided through a 20-minute hands-on use of the "Yi Yang Plan" mini-program. The consultation focused on 3 key domains: scientific validity, operability, and feasibility of the intervention. During the discussion, the facilitator ensured that all questions were clearly understood and that ambiguities were resolved to accurately capture expert perspectives. Within 24 hours, audio recordings were transcribed verbatim and independently verified by 2 researchers. The transcripts were subsequently returned to the experts for member checking to confirm the accuracy and completeness of their statements.

Data Collection

Expert Characteristics

Data included gender, age, highest degree, professional title, field of expertise, and years of experience.

XSL•FO
RenderX

Expert Engagement Coefficient

Expert engagement was assessed using the response rate, reflecting the level of participation and attentiveness of experts in the consultation process. A higher response rate indicates greater engagement and contributes to the objectivity of the consultation results.

Expert Authority Coefficient

The authority coefficient was used to evaluate the credibility and expertise of the participating experts. It reflects their professional judgment and familiarity with the topic. An authority coefficient of 0.70 or above is generally considered indicative of reliable expert consultation. Detailed calculation methods are provided in Multimedia Appendix 10.

Data Analysis

Descriptive analysis was conducted to summarize expert demographics, engagement coefficients, and authority coefficients. Qualitative data from the consultation were analyzed using conventional content analysis. Transcripts were independently coded by 2 coders, followed by face-to-face discussions to modify and reconcile coding discrepancies. Data analysis was performed using MAXQDA 2024 (VERBI Software).

Think-Aloud Test

Participants

Using convenience sampling, we recruited 20 dyads consisting of community-dwelling older adults with prefrailty and frailty and their family members from a community health center in Hangzhou, China. According to Faulkner [41], 20 participants are sufficient to identify the majority of usability issues. The inclusion criteria for older adult participants were as follows: (1) aged 60 or above years; (2) self-reported as prefrail or frail based on the FRAIL (Fatigue, Resistance, Ambulation, Illness, Loss of weight) scale (1‐2 criteria indicating prefrail; ≥3 criteria indicating frail); (3) residence in the community for at least 6 months [42]; (4) ability to read, write, and understand Mandarin; and (5) ability to use a smartphone. Individuals with a clinical diagnosis history of severe psychiatric or neurological disease, such as schizophrenia or Alzheimer disease, were excluded. Each older adult was encouraged to involve 1 family member (aged ≥18 years), who was related to the patient by blood, marriage, adoption, or affinity as a significant other, and met the language and smartphone-use criteria.

Procedure

Think-aloud tests were conducted in a meeting room arranged to resemble a typical living room environment. Before each session, we prepared the necessary equipment: smartphones with the WeChat preinstalled for the participants, audio recorders, timers, and Resistance and pens for observers. We also enabled screen-recording software on the smartphones to capture on-screen interactions. Based on the core functions provided by the “Yi Yang Plan” mini-program and the typical operations required for preparing for future care, 5 core tasks were designed in consultation with a software engineering expert (Multimedia Appendix 11).

At the beginning of each session, researchers introduced the study’s purpose. Participants were informed that their verbalizations and screen operations would be recorded during the testing process, and their data were fully confidential. Participants were also informed about the think-aloud protocol and instructed to verbalize their thoughts continuously while using the mini-program. As they navigated through each feature of the mini-program, they were encouraged to speak aloud their current thoughts; any decisions they were considering; and any feelings, confusions, or difficulties they experienced. Importantly, researchers did not intervene or answer questions during the task; participants were advised beforehand that any questions or issues they raised would be addressed only after the completion of the session. If a participant became stuck or hesitant, researchers offered neutral encouragement to continue but provided no direct assistance. If certain tasks could not be completed by the participant, these were noted as uncompleted tasks in our observations.

Each test session was recorded in multiple ways. A digital audio recorder captured all spoken comments from participants, the smartphone’s screen-recording software logged the exact on-screen actions and navigation, and the timer recorded the duration of each completed task. Researchers also objectively observed and documented each participant’s spoken words and nonverbal behaviors without attempting to interpret them.

Measures

General Characteristics

For older adults, the data included age, gender, residence, education level, and prior experience with mobile health apps. For family members, the data included relationship to the older adults, age, education level, and prior experience with mobile health apps.

Platform Use Satisfaction

Platform use satisfaction was assessed using a self-developed questionnaire comprising 6 items: interface design, features and functions, ease of operation, content quality, system stability, and overall experience. Each item was rated on a 5-point Likert scale, with scores ranging from 1 (“very dissatisfied”) to 5 (“very satisfied”).

Platform User Experience Interviews

A semistructured interview guide was developed based on the technology acceptance model (TAM) [43], including three open-ended questions: (1) Do you perceive this platform as useful? (2) Do you find this platform easy to use? (3) What suggestions do you have for improving the platform?

Data Analysis

Quantitative data, including task completion time and satisfaction scores, were analyzed using descriptive statistics. Categorical variables were summarized as frequencies and percentages, and continuous variables were reported as means (SD). The qualitative data from think-aloud sessions and interviews were transcribed verbatim and analyzed within the framework of the TAM. Thematic analysis and coding were conducted with a focus on the core dimensions of perceived usefulness and perceived ease of use in order to elucidate

XSL•FO
RenderX

participants' acceptance motivations, perceived barriers, and suggestions related to their use of the "Yi Yang Plan" mini-program. Data analysis was performed using MAXQDA 2024 (VERBI Software). Data analysis was conducted in Chinese. Themes and related quotations were translated into English for report writing by the first author. Translations were reviewed by the research team, and discrepancies were resolved through discussion.

Ethical Considerations

This study followed the principles of the Declaration of Helsinki and was approved by the Ethics Research Committee of Sir Run Run Shaw Hospital Affiliated to Medical College of Zhejiang University (2024-2759-01). Before enrollment, the responsible researcher explained the possible benefits and risks of the study to each potential participant in a face-to-face conversation, and they could withdraw from the study at any time without penalty. Sufficient time was provided for individuals to consider their participation, and only those who voluntarily agreed and signed the informed consent form were included.

Results

Expert Panel Consultations

Expert Characteristics

A total of 7 experts were recruited, including 6 female experts and 1 male expert. The mean age was 52.9 SD (7.38) years, and the average year of professional experience was 30.4 (SD 10.62) years. Detailed demographic characteristics are presented in Table 1.

Table . Demographics characteristics of the experts (n=7).

No.	Age (y)	Gender	Highest degree	Professional title	Field of expertise	Years of experience
1	50	Female	Master	Chief physician	General medicine	27
2	57	Female	Master	Chief physician	Geriatrics	34
3	60	Female	Master	Professor	Geriatric nursing	41
4	45	Female	Master	Chief physician	General medicine	22
5	40	Male	PhD	Professor	Sociology	10
6	58	Female	Master	Associate chief nurse	Community nursing	39
7	60	Female	Master	Professor	Nursing management	40

Expert Engagement Coefficient

All experts actively participated in the consultation and provided constructive feedback on the digital intervention. The engagement coefficient was 100%, indicating full participation and a high level of involvement.

Expert Authority Coefficient

The calculated authority coefficient was 0.88, indicating a high level of expertise and credibility among the participating experts. Therefore, the consultation results can be considered reliable.

Expert Recommendations

Establish a Mechanism for Care Plan Updates and Reviews

Experts emphasized that future care planning is a dynamic and evolving process. Expert 3 emphasized that care plans should be adjusted according to changes in older adults' physical conditions, family structures, personal beliefs, and policy directions. Accordingly, a regular update mechanism is necessary, with experts recommending updates every 6 months. This frequency balances the need for timely adaptation without overburdening older adults or their families. Experts 6 and 7 agreed with this recommendation. Additionally, expert 1 suggested incorporating a feature that allows users to review past care plans, enabling older adults and their families to compare changes over time and understand the rationale behind adjustments. Therefore, the platform should implement a comprehensive mechanism for both updating and reviewing care plans to enhance their timeliness, continuity, and adaptability.

Strengthen User-Centered Design

Several experts highlighted the lack of user-centered design in the current platform. Expert 2 strongly recommended that the platform architecture should address the real needs of older adults, enabling users to quickly access personalized and valuable information upon opening the app. Expert 5 noted that the platform's text-heavy content is not suitable for older adults with limited digital literacy. Excessive text undermines usability, while a lack of personalized content reduces usefulness. This expert suggested using more videos, intuitive interactions, and case stories to reduce cognitive load. Expert 4, from a community practice perspective, emphasized that the platform must address practical problems faced by older adults, reinforcing the necessity of a needs-driven approach. Therefore, user experience design must go beyond functionality, aiming to create a digital environment that is friendly, inclusive, and capable of effectively conveying complex information to older adults.

Implement Dynamic Resource Management and Data Security Measures

The platform's long-term viability depends on content accuracy and technological advancement. Experts expressed concerns about static resources and limited technical support. Expert 2 highlighted that outdated resources compromise credibility and usefulness, prompting users to seek alternative information

XSL•FO
RenderX

sources. Expert 7 underscored the importance of data security, continuous technical support, and ongoing updates, noting that nursing researchers may lack the technical expertise required, necessitating interdisciplinary collaboration. This suggests that a platform lacking dynamic maintenance and professional technical support would have limited real-world effectiveness. Therefore, the platform should ensure dynamic content updates and implement strict data security measures.

Enhancing Integration With Community Care Systems and Policy Frameworks

The ultimate value of the platform lies in its ability to address real-world needs. Expert 6, a community practitioner, emphasized that the platform's value increases if it can provide actionable guidance and integrate with offline services. Expert 1 highlighted the importance of aligning the platform with national public health policies. This indicates that the platform should not function as an isolated information source; rather, it should be embedded within the existing care ecosystem, serving as a bridge between user needs and professional services. Active integration with public service systems, enabling data sharing and online appointment scheduling, would enhance accessibility and efficiency. Additionally, collaboration with community organizations and the selection of representative communities for pilot studies are recommended to evaluate effectiveness in practice.

Think-Aloud Test

Participant Characteristics

A total of 20 dyads were recruited, including 20 community-dwelling older adults and 20 family members. Among the older adults, 15 (75%) were classified as prefrail, and 13 (65%) were male participants. The mean age was 72.2 (SD 8.65) years, with 9 (45%) participants aged between 70 and 79 years. Regarding educational attainment, 8 (40%) participants had completed junior high school. Thirteen (65%) participants reported prior experience using mobile health apps. Among family members, 8 (40%) were daughters of the older adults. The mean age was 51.9 (SD 13.04) years, with 7 (35%) participants aged between 40 and 49 years. Eleven (55%) had a university-level education or higher, and 17 (85%) had prior experience using mobile health apps (Table 2).

XSL•FO
RenderX

Table . Characteristics of the participants (n=40).

Characteristics	Values
Older adults (n=20)	
Frailty status, n (%)	
Prefrailty	15 (75)
Frailty	5 (25)
Gender, n (%)	
Male	13 (65)
Female	7 (35)
Age (y),	
Mean (SD)	72.2 (8.65)
60 - 69, n (%)	8 (40)
70 - 79, n (%)	9 (45)
80 - 89, n (%)	2 (10)
≥90, n (%)	1 (5)
Education level, n (%)	
Primary school	3 (15)
Junior high school	8 (40)
High school	5 (25)
University and above	2 (10)
Have you used a mobile health care application before?, n (%)	
Yes	13 (65)
No	7 (35)
Family members (n=20)	
Kinship with patient, n (%)	
Spouse	5 (25)
Son	7 (35)
Daughter	8 (40)
Age (y)	
Mean (SD)	51.9 (13.04)
30 - 39, n (%)	4 (20)
40 - 49, n (%)	7 (35)
50 - 59, n (%)	2 (10)
60 - 69, n (%)	4 (20)
≥70, n (%)	3 (15)
Education level, n (%)	
Primary school	3 (15)
Junior high school	2 (10)
High school	4 (20)
University and above	11 (55)
Have you used a mobile health care application before?, n (%)	
Yes	17 (85)
No	3 (15)

XSL•FO
RenderX

Duration of Task Test and Satisfaction Ratings

All the set tasks were completed by the older adults and their family members. The mean total time required to complete all tasks was 20.7 (SD 5.24) minutes. Among the individual tasks, participants spent the least time on system familiarization (task 1; mean 0.8, SD 0.48 min) and the most time on exploring care resources (task 3; mean 6.9, SD 1.81 min; Table 3).

User satisfaction ratings were generally favorable across system attributes. The highest rating was observed for ease of operation, with a rate of 4.1 (SD 0.69), while content quality received the lowest rating, with a rate of 3.7 (SD 0.81), and overall experience was rated at a satisfactory level, with a rate of 3.7 (SD 0.73; Table 3).

Table . The results of the duration of task test and the satisfaction ratings.

Item	Values, mean (SD)
Duration of task test (min)	
Task 1	0.8 (0.48)
Task 2	2.9 (0.95)
Task 3	6.9 (1.81)
Task 4	6.0 (1.31)
Task 5	4.0 (1.55)
Overall	20.7 (5.24)
Satisfaction ratings	
Page design	3.8 (0.89)
Features and functions	3.8 (0.70)
Ease of operation	4.1 (0.69)
Content quality	3.7 (0.81)
System stability	3.9 (0.72)
Overall experience	3.7 (0.73)

Views on the "Yi Yang Plan" Mini-Program

Perceived Usefulness of the Mini-Program

Participants expressed mixed views about the acceptability of the "Yi Yang Plan" mini-program. Some users acknowledged the program's value. Participants noted that those with greater awareness of care planning found the content engaging and described the mini-program as professionally designed and useful. One older adult with a caregiving background strongly emphasized that such proactive planning tools are highly meaningful and necessary, highlighting that everyone will ultimately face care-related challenges. The participant stressed that the tool can help older adults recognize the importance of PFC and take preparatory actions. Some participants reported that the mini-program provided access to information and resources that had previously been difficult to obtain, which they found highly beneficial. In addition, several users also underscored that the mini-program was developed by a multidisciplinary professional team, which reinforced the credibility of its design and content:

> *As aging becomes increasingly severe ... the situation of elderly care is becoming more severe … this (mini-program) is a good tool to participate in this elderly care industry earlier.* [Male, 78 years old, high school graduate]
>
> *This mini-program provides valuable information … I think that's really good. For example, policies on age-friendly home modifications are very helpful for older adults. The way houses were renovated in the past might not suit them anymore.* [Son, 48 years old, university graduate]
>
> *I think this mini-program is very good … since it was developed by nursing professionals, and it is highly necessary.* [Female, 76 years old, high school graduate]

Nevertheless, some users adopted a cautious or reserved stance, perceiving the direct benefits of the program for older adults as limited. Certain participants demonstrated insufficient awareness of PFC and lacked the inclination to engage in proactive planning, leading them to view the mini-program as unnecessary. In contrast, others felt that they had already completed most of their preparations and therefore regarded the program as offering little additional support. This perspective underscores that older adults occupy different stages of PFC, which shapes their varying needs for the program. In addition, some pointed out that older adults with lower levels of education were unable to understand or process the information presented in the program, leaving them without the ability to use it effectively or appreciate its value:

> *A lot of older adults in rural areas can't accept the idea of actively preparing for future care … so it's pretty hard to promote and use this kind of mini-program.* [Daughter, 39 years old, high school graduate]
>
> *I've been planning for a long time. I already have information about a lot of nursing homes saved on*

XSL•FO
RenderX

my phone, and I've even called them to learn more. [Female, 62 years old, university graduate]

I think there is some usefulness … (but) for older adults with low education levels, its role may not be very significant. [Male, 63 y old, university graduate]

Perceived Usability of the Mini-Program

Many participants emphasized that the program demonstrated a high level of usability, aligning with older adults' operational habits and cognitive abilities, with logically structured, easy-to-operate, and comprehensive content. By providing a clear and accessible design, the mini-program was also perceived as enhancing older adults' confidence in using digital technologies:

It's just a tree structure (pointing at the screen), the operation is easy too, and it feels quite straightforward. Maybe younger people think it's not fancy enough, but for us, it's very clear. [Male, 76 years old, middle school graduate]

The pages are very clear, and the content is quite comprehensive. For older adults like me, even if we're not super skilled with smartphones, we can still use them (with confident nodding). [Female, 76 years old, high school graduate]

However, many older adults reported difficulties in operating and understanding complex information due to declining vision, limited education, and insufficient digital skills. For example, some noted operational barriers within the mini-program, such as unclear return or next-page buttons that required additional effort to navigate. Others highlighted barriers in page presentation, including excessively small font sizes and overly dense text, which made reading and information acquisition challenging. In addition, participants pointed out that disease-related knowledge was described as overly specialized, with numerous technical terms that hindered comprehension. Similarly, while resources such as government documents and updated laws and regulations were viewed as authoritative, they were considered too theoretical and lacking in practical applicability:

The return in this mini-program is a bit troublesome. [Male, 63 years old, university graduate]

The words are so small. Even with glasses, I cannot see them. [Female, 62 years old, university graduate]

You should explain things in simple terms, and be more concrete. Older people do not think as actively as younger ones. [Male, 78 years old, high school graduate]

In addition, older adults indicated that when using the "Yi Yang Plan" mini-program, they preferred the participation of their adult children. This was because they were not sufficiently skilled at operating health-related digital tools, whereas younger family members were more adept at using them, leading older adults to rely on their support. In addition, some participants pointed out that PFC is a complex intergenerational family issue that requires the joint involvement of adult children in decision-making, ensuring that older adults' preferences are considered while also securing the support of their children:

Many older adults with low education level cannot use the mini-program and understand it. It must be the children or younger siblings who are able to process this information. [Female, 62 years old, university graduate]

For people around my age, born in the 70 s, we can think about things like preparing for future care on our own. But for my dad's generation, it's much harder. Most of them rely on children (shake the head helplessly). [Son, 48 years old, university graduate]

Suggestions for Improving the Mini-Program

Providing Personalized Recommendations

As highlighted in the preparation for future care process model proposed by Sörensen et al [39], older adults may be at different stages of preparation. Therefore, individualized planning processes and components should be recommended based on their current level of preparation. Although the mini-program includes an assessment function, it does not yet generate personalized resources or stage-specific recommendations from the assessment results. Hence, strengthening the function of personalized recommendations is necessary:

I've already thought it through, so I can make decisions directly without going through all the earlier parts. ... These resources are more like a way to fill in the gaps. I just look at what I need, and I can skip whatever I don't. [Female, 76 years old, high school graduate]

Enhancing Supportive Tools or Assistance

Some participants suggested that many older adults with lower levels of education or limited smartphone skills showed low acceptance of using the mini-program independently. To address this, they recommended providing additional supportive measures, such as organizing offline community-based training sessions or distributing instructional videos and user manuals. These forms of assistance could help older adults better understand care preparation and use the mini-program more smoothly:

I think you need to organize in-person classes for those older adults with low education levels … They can't understand it just by looking (at the mini-program). You'd have to explain everything to them carefully … If they could attend in person and listen, the effect would be much better. [Female, 62 years old, university graduate]

Improving Interaction and Content

Several participants emphasized the importance of simplifying navigation and enhancing content readability. They suggested establishing clearer logic for return and exit functions, as well as providing prompt guidance at key steps to minimize confusion. Participants also recommended increasing font size and adopting plain, accessible language to accommodate older users. In addition, they proposed breaking down lengthy passages into bullet points or lists and presenting information in a simplified format similar to PowerPoint slides, which they felt would improve clarity and ease of understanding:

XSL•FO
RenderX

> *Reduce the amount of text, and present the main points first so that they can be quickly understood … a format similar to PowerPoint would be easier for older adults to follow. In slides, you can use visual elements, extract the key content, enlarge the font, and keep the information brief.* [Daughter, 44 years old, university graduate, with work experience in IT]

Enhancing the Involvement of Adult Children

Several participants emphasized that preparation for future care is a complex issue that requires the involvement of multiple generations. Older adults often find it difficult to process the complex information contained in the mini-program and to make related decisions. As a result, they preferred that their adult children participate in the discussions and provide assistance with using the mini-program.

> *It is better for older adults to use it together with their children, and the wishes of the older adults themselves are more important.* [Male, 48 years old, university graduate]

Integrating Practical Service Resources

Some participants expressed a desire for the program to be linked with offline services to improve the feasibility of planning. For example, they recommended direct connections to local nursing homes, home care providers, domestic service agencies, and medical resources. Such integration was seen as a way to help older adults move from planning to action, although participants recognized that this would require cross-sector collaboration and long-term development:

> *It would be better if the connection to the local community could be stronger ... Which communities have connected with certain services, such as domestic services, medical services ... If you can do a good job with this connection ... the experience would definitely be better.* [Son, 48 years old, university graduate]

Discussion

Principal Findings

This study developed and pilot-tested the "Yi Yang Plan" mini-program to support PFC among community-dwelling older adults with prefrailty and frailty. The findings from expert panel consultations demonstrated the scientific rigor of the program, while think-aloud tests indicated that the mini-program was feasible to deliver and generally acceptable to users. Nevertheless, expert recommendations, usability barriers, and user-derived suggestions highlighted several areas for further refinement.

Advantages of the Digital Intervention

The "Yi Yang Plan" mini-program was developed through a rigorous, theory- and evidence-based process integrating evidence synthesis, identification of influencing factors, and multidisciplinary collaboration. A prior scoping review identified 4 key elements of PFC [33], which are consistent with the established PFC process model [39], thereby providing a strong conceptual foundation. In addition, a convergent mixed methods study identified key determinants of PFC among frail older adults, and a structural equation model based on the Andersen Behavioral Model of Health Services Use informed the development of stratified intervention pathways. This approach enabled the design of a targeted and differentiated intervention tailored to varying frailty levels. Furthermore, iterative refinement by a multidisciplinary team enhanced both usability and content quality. Collectively, this structured development process offers a replicable framework for designing digital interventions to support PFC.

Expert consultation results indicated high credibility and reliability, reflected by strong engagement and authority coefficients. Experts with substantial clinical and theoretical expertise confirmed the scientific validity of the platform and provided actionable recommendations, supporting the robustness of the development process. The findings from think-aloud testing further demonstrated good feasibility and acceptability. Participants expressed generally positive perceptions and recognized the value of the mini-program, suggesting a willingness among older adults to engage with digital health tools. These findings align with previous studies indicating that older adults perceive supportive digital interventions as useful and acceptable [32,44]. One explanation is that the platform provides accessible information and resources, thereby enhancing the awareness and understanding of PFC [45]. In addition, the involvement of health care professionals in the development process may have strengthened users' trust, as older adults tend to rely on information from credible and familiar sources [46,47].

Barriers to Digital Intervention

Despite overall positive findings, several important challenges emerged. First, heterogeneity in PFC readiness influenced user needs. Participants noted that individuals at different stages of PFC require different types of support, which is consistent with prior research [48]. This highlights the importance of tailoring intervention pathways. For individuals with low awareness, interventions should focus on early-stage awareness enhancement, whereas those with higher readiness may benefit more from decision-making and planning support.

Second, disparities in education and digital literacy posed barriers to engagement. Older adults with lower educational attainment may have limited health literacy and digital skills, which can hinder technology adoption [49-51]. Therefore, assessing digital readiness prior to intervention delivery is essential [52]. Complementary support strategies, such as printed manuals, instructional videos, and offline guidance, are necessary to improve accessibility and inclusiveness [53].

Third, although usability received the highest satisfaction ratings and all participants completed the tasks, specific usability issues were identified, including unclear navigation and suboptimal font size. Age-related declines in sensory, motor, and cognitive functions may contribute to these challenges [47,54,55]. Accordingly, future designs should simplify navigation structures, provide clearer operational logic (eg, return and exit functions), and incorporate step-by-step guidance. Established age-friendly design principles, such as larger fonts, high contrast,

XSL•FO
RenderX

concise content, and clear labeling, should be consistently applied [56].

Fourth, content quality received relatively lower ratings. While participants acknowledged the comprehensiveness of the information, they reported difficulties in understanding and applying the content. This is likely because the resource module includes policy-level and professionally complex information. Simplifying the language and incorporating patient-friendly expressions may improve comprehension [57]. Additionally, multimedia formats (eg, videos and audio) can enhance engagement and accommodate diverse learning preferences [56]. Furthermore, emerging technologies such as AI could be leveraged to deliver personalized recommendations, improve information filtering, and enhance intervention precision and scalability [58]. Importantly, participants emphasized the need for more practical, actionable resources. Strengthening partnerships with community organizations and service providers is therefore critical to ensure that digital planning can be effectively translated into real-world care arrangements.

Finally, participants highlighted the essential role of adult children in facilitating technology use and care planning. This finding is consistent with previous research showing that older adults rely heavily on family support in PFC [1,44,54]. Integrating adult children as active participants in the intervention may enhance engagement, usability, and overall effectiveness.

Optimization Strategies

Based on these findings, several optimization strategies are warranted. At the system level, a comprehensive security and privacy protection framework should be implemented, including encryption, role-based access control, data deidentification, and audit logging, to ensure the safe handling of sensitive health information. At the interface level, improvements should focus on enhancing usability and emotional engagement. This includes adopting age-friendly layouts, increasing font size and contrast, simplifying navigation, and incorporating supportive visual and narrative elements to improve the user experience. At the functional level, the platform should further strengthen personalization by integrating periodic assessments, dynamic care plan updates, resource linkage, and historical tracking functions. At the implementation level, additional support tools should be provided to users with limited digital literacy. Moreover, the formal integration of adult children and stronger collaboration with community health care providers are essential to ensure continuity between digital planning and real-world service delivery.

Limitations

This study has several limitations that should be acknowledged. First, the small sample size limits the generalizability of the findings. As a pilot study, the primary aim was to assess feasibility and acceptability rather than effectiveness. Future research should conduct adequately powered randomized controlled trials to evaluate intervention efficacy. Second, although an iterative design approach was adopted, older adults and their family members were not involved as co-designers during the early development phase. Consequently, some usability issues were only identified during pilot testing. Future studies should incorporate more principles from human factors engineering and adopt participatory design methods to better align the intervention with the needs and capabilities of end users. Third, many participants had prior experience with digital health apps, which may limit the applicability of the findings to populations with lower digital literacy. Future research should intentionally recruit individuals with lower levels of digital skills to ensure that the tool is responsive to the needs of those who require the greatest support.

Conclusion

Based on a scoping review, analysis of influencing factors, stratified path analysis, as well as a multidisciplinary collaborative development process, this study developed a digital PFC platform for community-dwelling prefrail and frail older adults. It integrates standardized core procedures, stratified precision interventions, and digital platform delivery. The pilot test showed that the “Yi Yang Plan” mini-program demonstrated initial scientific validity, usability, and feasibility. Future versions should focus on improving system architecture, user interface design, platform features, and implementation strategies. The pilot results also provide a solid foundation for further thorough evaluation through randomized controlled trials.

Acknowledgments

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to generative AI (GenAI) tools under full human supervision: proofreading and editing. The GenAI tool used was ChatGPT-4.5. Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Declaration submitted by: Collective responsibility.

Funding

This work was supported by the National Health Commission of the People’s Republic of China (WKJ-ZJ-2304 to YZ).

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: QN
Data curation: JD, WJ
Formal analysis: RY
Methodology: QN, JD, WJ
Project administration: YZ
Resources: LW
Supervision: YZ
Visualization: RY, JY
Writing – original draft: QN
Writing – review and editing: YC, LW
QN and JD are co-first authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The interface design on the “Lanhu” platform.
[PNG File, 467 KB - aging_v9i1e92492_app1.png]

Multimedia Appendix 2
“Yi Yang Plan” community management terminal interface.
[PNG File, 128 KB - aging_v9i1e92492_app2.png]

Multimedia Appendix 3
User login and frailty assessment interface.
[PNG File, 246 KB - aging_v9i1e92492_app3.png]

Multimedia Appendix 4
Awareness enhancement interface.
[PNG File, 438 KB - aging_v9i1e92492_app4.png]

Multimedia Appendix 5
Care resources interface.
[PNG File, 494 KB - aging_v9i1e92492_app5.png]

Multimedia Appendix 6
Care decision-making interface.
[PNG File, 238 KB - aging_v9i1e92492_app6.png]

Multimedia Appendix 7
Family care decision checklist.
[DOCX File, 15 KB - aging_v9i1e92492_app7.docx]

Multimedia Appendix 8
Care plan and updates interface.
[PNG File, 423 KB - aging_v9i1e92492_app8.png]

Multimedia Appendix 9
Care preparation checklist and personal center interface.
[PNG File, 112 KB - aging_v9i1e92492_app9.png]

Multimedia Appendix 10
The methods of the expert authority coefficient.
[DOCX File, 15 KB - aging_v9i1e92492_app10.docx]

XSL•FO
RenderX

Multimedia Appendix 11
Task list for think-aloud tests.
[DOCX File, 16 KB - aging_v9i1e92492_app11.docx]

References

1. Jokisch MR, Schmidt LI, Doh M. Acceptance of digital health services among older adults: findings on perceived usefulness, self-efficacy, privacy concerns, ICT knowledge, and support seeking. Front Public Health 2022;10:1073756. [doi: 10.3389/fpubh.2022.1073756] [Medline: 36582385]
2. What is the aging of the population and its measurement standard? National Bureau of Statistics. 2025. URL: https://www.stats.gov.cn/zs/tjws/tjbz/202301/t20230101_1903949.html [accessed 2026-07-23]
3. Statistical communiqué on the national economic and social development of the People's Republic of China in 2024. National Bureau of Statistics. 2025 Feb. URL: https://www.stats.gov.cn/sj/zxfb/202502/t20250228_1958817.html [accessed 2025-04-16]
4. The 7th national census bulletin (no 5). Nation Bureau of Statistics. 2021. URL: http://www.stats.gov.cn/sj/tjgb/rkpcgb/qgrkpcgb/202302/t20230206_1902005.html [accessed 2023-03-01]
5. World report on ageing and health. : World Health Organization; 2015 URL: https://iris.who.int/server/api/core/bitstreams/5174879e-b0dc-43fc-b3a8-b1db31c51d4c/content [accessed 2026-07-08]
6. Zhao Y, Li CY, Xiao ZY, Li WH, Cai YT, Lyu QY. Community care needs and their influencing factors in older adults with multimorbidity in different frail states: a cross-sectional study. Geriatr Nurs 2025;65:103503. [doi: 10.1016/j.gerinurse.2025.103503] [Medline: 40645143]
7. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
8. Morley JE, Vellas B, van Kan GA, et al. Frailty consensus: a call to action. J Am Med Dir Assoc 2013 Jun;14(6):392-397. [doi: 10.1016/j.jamda.2013.03.022] [Medline: 23764209]
9. Shamliyan T, Talley KMC, Ramakrishnan R, Kane RL. Association of frailty with survival: a systematic literature review. Ageing Res Rev 2013 Mar;12(2):719-736. [doi: 10.1016/j.arr.2012.03.001] [Medline: 22426304]
10. Fan J, Yu C, Guo Y, et al. Frailty index and all-cause and cause-specific mortality in Chinese adults: a prospective cohort study. Lancet Public Health 2020 Dec;5(12):e650-e660. [doi: 10.1016/S2468-2667(20)30113-4] [Medline: 33271078]
11. Abete P, Basile C, Bulli G, et al. The Italian version of the "frailty index" based on deficits in health: a validation study. Aging Clin Exp Res 2017 Oct;29(5):913-926. [doi: 10.1007/s40520-017-0793-9] [Medline: 28688080]
12. Dent E, Morley JE, Cruz-Jentoft AJ, et al. Physical frailty: ICFSR international clinical practice guidelines for identification and management. J Nutr Health Aging 2019;23(9):771-787. [doi: 10.1007/s12603-019-1273-z] [Medline: 31641726]
13. Liguori I, Russo G, Bulli G, et al. Validation of "(fr)AGILE": a quick tool to identify multidimensional frailty in the elderly. BMC Geriatr 2020 Sep 29;20(1):375. [doi: 10.1186/s12877-020-01788-1] [Medline: 32993569]
14. He B, Ma Y, Wang C, et al. Prevalence and risk factors for frailty among community-dwelling older people in China: a systematic review and meta-analysis. J Nutr Health Aging 2019;23(5):442-450. [doi: 10.1007/s12603-019-1179-9] [Medline: 31021361]
15. Simpson KN, Seamon BA, Hand BN, et al. Effect of frailty on resource use and cost for Medicare patients. J Comp Eff Res 2018 Aug;7(8):817-825. [doi: 10.2217/cer-2018-0029] [Medline: 29808714]
16. Zhang L, Fang Y. Prediction study on the scale of disability in urban and rural elderly population in China and the cost of care from 2020 to 2050 [Article in Chinese]. Chin J Health Stat 2021;38(1):39-42 [FREE Full text] [doi: 10.3969/j.issn.1002-3674.2021.01.010]
17. Fu YY, Chui EWT. Determinants of patterns of need for home and community-based care services among community-dwelling older people in urban China: the role of living arrangement and filial piety. J Appl Gerontol 2020 Jul;39(7):712-721. [doi: 10.1177/0733464819871875] [Medline: 32517576]
18. Song Y, Yan ECW, Sörensen S. The effects of familism on intended care arrangements in the process of preparing for future care among one-child parents in urban China. Ageing Soc 2017 Aug;37(7):1416-1434. [doi: 10.1017/S0144686X16000349]
19. Li H, Wang R, Ren Y. Analysis of the number of only children and family structure since the implementation of the family planning policy [Article in Chinese]. Stat Decis 2018;34(13):99-104. [doi: 10.13546/j.cnki.tjyjc.2018.13.022]
20. Lu JH, Feng HD. Characteristics of parental care preferences among China's only-child generation and policy implications [Article in Chinese]. Youth Explor 2022(6):5-16. [doi: 10.13583/j.cnki.issn1004-3780.2022.06.001]
21. Li QJ. A study on the elderly care preparation and its impact on elderly care stress in only-child families [Dissertation]. : Zhejiang University; 2018 URL: https://d.wanfangdata.com.cn/thesis/Y3518881 [accessed 2023-10-28]
22. Lu J, Zhang L. Patterns of care needs and their influencing factors among older adults in China: an empirical study based on the China Longitudinal Aging Social Survey [Article in Chinese]. Popul J 2018(2):22-33. [doi: 10.16405/j.cnki.1004-129X.2018.02.002]
23. Friedemann ML, Newman FL, Seff LR, Dunlop BD. Planning for long-term care: concept, definition, and measurement. Gerontologist 2004 Aug;44(4):520-530. [doi: 10.1093/geront/44.4.520] [Medline: 15331809]

XSL•FO
RenderX

24. Sörensen S, Mak W, Pinquart M. Planning and decision making for care transitions. Annu Rev Gerontol Geriatr 2011;31(1):111-142. [doi: 10.1891/0198-8794.31.111] [Medline: 26207079]
25. Li Y, Sun Y, Li J, Jia J. Patterns and predictors of preparation for future care among older rural Chinese adults: a latent profile analysis. J Appl Gerontol 2022 Jan;41(1):207-216. [doi: 10.1177/0733464820964148] [Medline: 34886703]
26. Brinkman-Stoppelenburg A, Rietjens JAC, van der Heide A. The effects of advance care planning on end-of-life care: a systematic review. Palliat Med 2014 Sep;28(8):1000-1025. [doi: 10.1177/0269216314526272] [Medline: 24651708]
27. Sörensen S, Mak W, Chapman B, Duberstein PR, Lyness JM. The relationship of preparation for future care to depression and anxiety in older primary care patients at 2-year follow-up. Am J Geriatr Psychiatry 2012 Oct;20(10):887-894. [doi: 10.1097/JGP.0b013e31822ccd8c] [Medline: 21952122]
28. Chari AV, Engberg J, Ray KN, Mehrotra A. The opportunity costs of informal elder-care in the United States: new estimates from the American Time Use Survey. Health Serv Res 2015 Jun;50(3):871-882. [doi: 10.1111/1475-6773.12238] [Medline: 25294306]
29. Davies N, Maio L, Rait G, Iliffe S. Quality end-of-life care for dementia: what have family carers told us so far? A narrative synthesis. Palliat Med 2014 Jul;28(7):919-930. [doi: 10.1177/0269216314526766] [Medline: 24625567]
30. Nakanishi M, Endo K, Hirooka K, et al. Dementia behaviour management programme at home: impact of a palliative care approach on care managers and professional caregivers of home care services. Aging Ment Health 2018 Aug;22(8):1057-1062. [doi: 10.1080/13607863.2017.1332160] [Medline: 28553880]
31. Final report on the "own your future" consumer survey. : Office of the Assistant Secretary for Planning and Evaluation (ASPE), U.S. Department of Health and Human Services; 2006 URL: https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//42301/OYFsurvey.pdf [accessed 2026-07-08]
32. Hoffman AS, Bateman DR, Ganoe C, et al. Development and field testing of a long-term care decision aid website for older adults: engaging patients and caregivers in user-centered design. Gerontologist 2020 Jul 15;60(5):935-946. [doi: 10.1093/geront/gnz141] [Medline: 31773140]
33. Qiqi N, Li W, Weijing S, Yiyu Z. Supportive interventions for preparation for future care of community-dwelling older adults: a scoping review. J Adv Nurs 2025 Mar;81(3):1195-1207. [doi: 10.1111/jan.16394] [Medline: 39304322]
34. Ni Q, Wang L, Sui W, Zhuang Y. Research progress on preparation for future care among older adults [Article in Chinese]. J Nurs Sci 2024;39(16):105-110. [doi: 10.3870/j.issn.1001-4152.2024.16.105]
35. The CPC central committee and the state council issued the "Outline of the 'Healthy China 2030' plan" [Article in Chinese]. The State Council of the People's Republic of China. 2016. URL: https://www.gov.cn/zhengce/2016-10/25/content_5124174.htm [accessed 2023-10-10]
36. Notice from the General Office of the National Health Commission on further promoting the pilot work of "internet + nursing services" [Article in Chinese]. The State Council of the People's Republic of China. 2020. URL: https://www.gov.cn/zhengce/zhengceku/2020-12/16/content_5569982.htm [accessed 2023-10-10]
37. Notice from the national health commission on issuing the "national nursing development plan (2021-2025)" [Article in Chinese]. The State Council of the People's Republic of China. 2022. URL: https://www.gov.cn/zhengce/zhengceku/2022-05/09/content_5689354.htm?eqid=89a1eabf00048c8a000000066462e387 [accessed 2023-10-10]
38. Ni Q, Wang L, Sui W, Zhuang Y. Determinants of preparation for future care among community-dwelling older adults with chronic diseases: a mixed-method study. Geriatr Nurs 2025;66(Pt A):103704. [doi: 10.1016/j.gerinurse.2025.103704] [Medline: 41202436]
39. Sörensen S, Chapman BP, Duberstein PR, Pinquart M, Lyness JM. Assessing future care preparation in late life: two short measures. Psychol Assess 2017 Dec;29(12):1480-1495. [doi: 10.1037/pas0000446] [Medline: 28301191]
40. Editorial Board of Chinese Journal of Geriatrics, Chinese Geriatrics Society. Chinese expert consensus on prevention of frailty in the elderly (2022). Chin J Geriatr 2022;41(5):503-511. [doi: 10.3760/cma.j.issn.0254-9026.2022.05.001]
41. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]
42. How to distinguish between the resident population and the mobile population. National Bureau of Statistics of People's Republic of China. 2025. URL: https://www.stats.gov.cn/zs/tjws/tjbz/202301/t20230101_1903796.html [accessed 2026-07-23]
43. Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Int J Man Mach Stud 1993 Mar;38(3):475-487. [doi: 10.1006/imms.1993.1022]
44. Kozlov E, McDarby M, Duberstein P, Carpenter BD. The feasibility and acceptability of an intergenerational, web-based intervention to enhance later-life family care planning. Gerontologist 2021 Sep 13;61(7):1153-1163. [doi: 10.1093/geront/gnaa137] [Medline: 32946554]
45. Lindquist LA, Ramirez-Zohfeld V, Sunkara PD, et al. PlanYourLifeSpan.org - an intervention to help seniors make choices for their fourth quarter of life: results from the randomized clinical trial. Patient Educ Couns 2017 Nov;100(11):1996-2004. [doi: 10.1016/j.pec.2017.06.028] [Medline: 28689855]
46. Chaudhuri S, Le T, White C, Thompson H, Demiris G. Examining health information-seeking behaviors of older adults. Comput Inform Nurs 2013 Nov;31(11):547-553. [doi: 10.1097/01.NCN.0000432131.92020.42] [Medline: 23974574]
47. Cajita MI, Hodgson NA, Lam KW, Yoo S, Han HR. Facilitators of and barriers to mHealth adoption in older adults with heart failure. Comput Inform Nurs 2018 Aug;36(8):376-382. [doi: 10.1097/CIN.0000000000000442] [Medline: 29742549]

XSL•FO
RenderX

48. Lee J, Rafferty K, Peitz L, Bormann E. Qualitative evaluation of a future care planning program. J Ext 2020 Aug 1;58(4). [doi: 10.34068/joe.58.04.13]
49. Shao Y, Yang X, Chen Q, et al. Determinants of digital health literacy among older adult patients with chronic diseases: a qualitative study. Front Public Health 2025;13:1568043. [doi: 10.3389/fpubh.2025.1568043] [Medline: 40231178]
50. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
51. Litchfield I, Shukla D, Greenfield S. Impact of COVID-19 on the digital divide: a rapid review. BMJ Open 2021 Oct 12;11(10):e053440. [doi: 10.1136/bmjopen-2021-053440] [Medline: 34642200]
52. Gordon NP, Hornbrook MC. Older adults' readiness to engage with eHealth patient education and self-care resources: a cross-sectional survey. BMC Health Serv Res 2018 Mar 27;18(1):220. [doi: 10.1186/s12913-018-2986-0] [Medline: 29587721]
53. Lee JE, Kim D, Kahana E, Kahana B. Feasibility and acceptability of the community-based program: plan ahead. Aging Ment Health 2023 Apr;27(4):811-819. [doi: 10.1080/13607863.2022.2046692] [Medline: 35285774]
54. Navabi N, Ghaffari F, Jannat-Alipoor Z. Older adults' attitudes and barriers toward the use of mobile phones. Clin Interv Aging 2016;11:1371-1378. [doi: 10.2147/CIA.S112893] [Medline: 27757025]
55. Nymberg VM, Bolmsjö BB, Wolff M, Calling S, Gerward S, Sandberg M. 'Having to learn this so late in our lives…' Swedish elderly patients' beliefs, experiences, attitudes and expectations of e-health in primary health care. Scand J Prim Health Care 2019 Mar;37(1):41-52. [doi: 10.1080/02813432.2019.1570612] [Medline: 30732519]
56. Hou SI. Health literacy online: a guide to writing and designing easy-to-use health web sites. Health Promot Pract 2012 Sep;13(5):577-580. [doi: 10.1177/1524839912446480] [Medline: 22763891]
57. Clearer conversations, better care: new guide highlights best practices for communicating with older adults. Gerontological Society of America. 2025. URL: https://www.geron.org/News-Events/GSA-News/Press-Room/Press-Releases/clearer-conversations-better-care-new-guide-highlights-best-practices-for-communicating-with-older-adults [accessed 2025-09-28]
58. Arioz U, Allsop MJ, Goodman WD, et al. Artificial intelligence-based approaches for advance care planning: a scoping review. BMC Palliat Care 2025 Oct 23;24(1):268. [doi: 10.1186/s12904-025-01827-x] [Medline: 41131546]

Abbreviations

FRAIL: Fatigue, Resistance, Ambulation, Illness, Loss of weight
PFC: preparation for future care
SARC-F: Strength, Assistance in walking, Rise from a chair, Climb stairs, Falls
SEM: structural equation model
TAM: technology acceptance model

Edited by AF Abdulai; submitted 30.Jan.2026; peer-reviewed by A Farmer, K Zhang; revised version received 07.May.2026; accepted 05.Jun.2026; published 05.Aug.2026.

Please cite as:
Ni Q, Dong J, Jiang W, Yi J, Yan R, Chen Y, Wang L, Zhuang Y
Preparation for Future Care in Community-Dwelling Older Adults With Prefrailty and Frailty: Development and Pilot Test of a Digital Intervention
JMIR Aging 2026;9:e92492
URL: https://aging.jmir.org/2026/1/e92492
doi:10.2196/92492

© Qiqi Ni, Jing Dong, Weilin Jiang, Jia Yi, Ran Yan, Yuhan Chen, Li Wang, Yiyu Zhuang. Originally published in JMIR Aging (https://aging.jmir.org), 5.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Practice Effects on Repeated Self-Administered Digital Cognitive Assessments in Cognitively Intact Older Adults from the Framingham Heart Study: Longitudinal Observational Study

Kieffer Christianson[1,2], MA; Chenglin Lyu[2], MS; Cody Karjadi[2,3], MS; Sherral Devine[3], PhD; Ting Fang Alvin Ang[2,4], MPH, MD; Ileana De Anda-Duran[5], MPH, MD; Vijaya B Kolachalama[6,7], PhD; Honghuang Lin[8], PhD; Huitong Ding[2,3], PhD; Katherine A Gifford[2,3], PsyD; Rhoda Au[2,3,4,6,9,10], PhD, MBA; Phillip H Hwang[2,3,9], MPH, PhD; Preeti Sunderaraman[11], PhD

[1]Department of Psychology, University of Nevada, Reno, 1664 N. Virginia St, Reno, NV, United States
[10]Boston University Alzheimer's Disease Research Center, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States
[11]Department of Psychiatry and Human Behavior, Memory and Aging Program, Butler Hospital, Brown University, Providence, RI, United States
[2]Department of Anatomy and Neurobiology, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States
[3]Framingham Heart Study, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States
[4]Department of Neurology, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States
[5]Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, United States
[6]Department of Medicine, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States
[7]Department of Computer Science, Boston University, Boston, MA, United States
[8]Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States
[9]Department of Epidemiology, Boston University School of Public Health, Boston, MA, United States

Corresponding Author:
Kieffer Christianson, MA
Department of Psychology, University of Nevada, Reno, 1664 N. Virginia St, Reno, NV, United States

Abstract

Background: Remote, self-administered digital assessments provide scalable approaches for monitoring cognition in older adults. Practice effects (PEs), once viewed as measurement noise, are increasingly recognized as potential markers of cognitive decline. As disease-modifying therapies for neurodegenerative conditions emerge, longitudinal digital assessments may capture subtle cognitive changes. Although the majority of PE studies have focused on memory tasks, it is critical to determine whether PEs occur in other domains. Clarifying the potential presence of PEs across multiple time points on the Food and Drug Administration–cleared Defense Automated Neurocognitive Assessment (DANA) platform remains critical for determining whether PEs can serve as a measure of cognitive decline longitudinally.

Objective: This study aimed to examine whether PEs occur and whether demographics moderate them across quarterly self-administrations of digital cognitive tasks spanning different cognitive domains in cognitively intact older adults.

Methods: Cognitively intact older adults from the Framingham Heart Study self-administered DANA tasks at baseline and every 3 months for 1 year (4 assessments: T1-T4). Tasks included Simple Reaction Time (SRT), Go/No-Go (GNG), and Code Substitution (CS). These tasks varied in complexity, from simple motor responding (SRT) to inhibitory control (GNG) and associative learning with working memory demands (CS). The primary outcome was Cognitive Efficiency, a derived measure of speed and accuracy, which are also recorded as secondary measures. PEs were quantified using linear mixed-effects models, and the change in mean score across time points was expressed in pooled SD units as Cohen *d*.

Results: A total of 717 participants completed all 4 time points (mean age 64.1, SD 9.5 years; n=469, 65.4% female participants; n=481, 67.1% college-educated). PEs were observed on all tasks, with small effects on CS (*d*=0.43) and GNG (*d*=0.32) and a below-threshold effect on SRT (*d*=0.16). Demographics influenced baseline performance: older age predicted lower baseline scores on all tasks, females scored lower on SRT and GNG, and higher education predicted better performance on all tasks. Age was the only demographic moderator of PEs, with smaller improvements on CS among participants aged 65 years or older (B=−0.39; *P*<.001).

Conclusions: Quarterly self-administered digital cognitive assessments elicited measurable PEs across processing speed, inhibitory control, and visual scanning, with stronger effects on more complex tasks. Demographics influenced baseline performance, while age moderated PEs on the most demanding task. These findings provide baseline data against which PEs in clinical populations can be compared in future work, complementing prior evidence on other cognitive instruments.

XSL•FO
RenderX

(JMIR Aging 2026;9:e87208) doi:10.2196/87208

KEYWORDS

practice effects; brain smartphone; aging; digital assessment; cognitive assessment; neuropsychology

Introduction

Practice effects (PEs) refer to performance gains across repeated cognitive test administrations due to prior exposure to test materials and/or procedures [1]. Historically regarded as measurement noise, PEs are increasingly recognized as a meaningful index of learning capacity [2-4], and attenuated PEs have emerged as a candidate indicator of preclinical Alzheimer disease (AD) and prospective cognitive decline [5-8]. Failure to account for PEs can lead to misinterpretation of stable cognitive scores as preserved cognition rather than as reflecting deficits in learning capacity [9], and correcting for PEs alters rates of incident mild cognitive impairment detection [10].

PEs are mediated by test familiarity, whether with its stimuli or its procedures [2]. Stimulus-specific learning accumulates with each exposure [11], producing PEs that are largest on verbal memory tests and smallest on tests where new stimuli minimally alter task demands [12]. Alternate forms reduce but do not eliminate PEs on many tests [12,13]. The residual PEs reflect procedural familiarity with the test's instructions and response format, which generalizes across stimuli. More complex tasks plausibly offer more procedural elements to learn, although task complexity has not been formally operationalized in PE research.

Moderators of PEs include the interval between sessions and individual characteristics. Recent digital research has used high-frequency, measurement-burst, and daily ecological momentary assessment designs that produce PEs within days to weeks [14-16], whereas earlier studies of legacy paper-and-pencil tests have documented PEs across intervals ranging from months to years [17-19]. Among individual characteristics, cognitive status is the strongest moderator of PEs [5,12,20,21]. Age also moderates PEs among healthy adults, with correlations between mean age and estimated PEs ranging from −0.66 to −0.89 across reasoning, spatial, memory, and speed domains [22]. Evidence on sex and education effects has been mixed in both traditional tests [23-25] and digital studies, with some reporting no demographic moderation [26] and others identifying task-specific associations [27-29].

The Framingham Heart Study (FHS) began collecting digital, self-administered, and remote serial cognitive data in 2020 using the Defense Automated Neurocognitive Assessment (DANA) application. Subsequent studies have demonstrated DANA's feasibility and user acceptability [30,31], concurrent validity with traditional paper-and-pencil–based neuropsychological tests [32], and association with cardiovascular risk factors [33]. Although designed to minimize PE [34], Bellitti et al [27] reported modest response time (RT) PEs on 3 of 6 DANA tasks across 2 sessions spaced approximately 90 days apart in a 116-participant mixed-diagnostic sample. Whether these effects persist over additional administrations at a quarterly interval in a larger cognitively intact sample has not been established.

The primary aim of the current study was to characterize PEs across 4 quarterly assessments on 3 DANA tasks spanning a graded range of task demands (Simple Reaction Time [SRT], Go/No-Go [GNG], and Code Substitution [CS]) in a large sample of cognitively intact older adults. The secondary aim was to determine whether demographic characteristics moderate the observed PEs. The present study was not designed to test the prognostic value of PEs but rather to establish baseline PE patterns against which such questions can be evaluated in future longitudinal work. We anticipated that (1) PEs would be detectable across the 3 tasks but vary in magnitude, with more cognitively demanding tasks showing larger effects than motor-dominant tasks; and (2) older age would attenuate PE on the more demanding tasks [22], while sex and education effects would be modest given mixed prior evidence [23].

Methods

Participants

The study sample consisted of participants enrolled in the FHS, a longitudinal, population - based cohort study that began in 1948 in Framingham, Massachusetts. The FHS focuses on studying risk factors for cardiovascular disease and other major chronic diseases spanning 3 generations of participants and their descendants. Neuropsychological evaluations began in 1976 with the collection of a brief cognitive screen and progressed to a comprehensive cognitive evaluation in 1999. The FHS began enrolling participants from diverse racial and ethnic backgrounds into multiethnic cohorts (Omni cohorts) in 1990, while the generation 2 (Gen 2) and generation 3 (Gen 3) cohorts are mostly White. This study included participants from 5 subcohorts of the FHS: Gen 2, Gen 3, Omni 1, Omni 2, and New Offspring Spouse. The inclusion criteria for this study were owning a smartphone, proficiency in spoken English, and having Wi-Fi access. All participants recruited for the study were presumed to be cognitively unimpaired based on established FHS dementia surveillance protocols, which flag participants for cognitive concerns based on Mini-Mental State Examination performance (score ≤24 or decline ≥3 points between visits), self-reported or informant-reported cognitive symptoms, or clinician referral, followed by consensus dementia review by a panel with at least one neurologist and one neuropsychologist [35]. Sociodemographic data were collected in accordance with established FHS protocols [32]. Variables included self-reported age (years) at the first assessment, sex (female/male), and education (high school or less, some college, and college graduate).

A total of 1756 participants completed at least one cognitive assessment during the data collection period (July 14, 2022, to March 28, 2025) and had the opportunity to complete all 4 quarterly assessments. The primary analytic sample comprised the 717 participants who completed all 4 time points of CS; corresponding completer counts were 701 for GNG and 686 for SRT, reflecting the fixed task order in which early

XSL•FO
RenderX

discontinuation within a session resulted in retention of data for tasks completed earlier. The remaining 1039 participants completed 1 to 3 time points (partial completers) and served as the comparison sample for selection-bias analyses.

DANA Tasks

Participants were asked to complete the DANA tasks every 3 months for a year, resulting in 4 time points (baseline plus 3 additional time points). DANA tasks were delivered via the Linus Health mobile app. For each assessment interval, participants had a 42-day window to complete their assessments. Participants received automated daily notifications if tasks remained incomplete, which they had the option to mute. Research assistants provided up to 2 to 3 additional reminders per assessment window via phone, email, or text, based on participant preference.

The DANA app is a Food and Drug Administration–cleared computerized neurocognitive test battery accessible as a mobile application on iOS and Android devices. Previous work has demonstrated DANA's internal consistency and reliability [36]. Three subtests were selected for this study that collectively take approximately 5 minutes to complete. The first subtest, SRT, assesses RT by measuring latency between when a large target appears on the screen and when the participant taps the target across 40 trials. The second subtest, GNG, assesses sustained attention and impulsivity by requiring the participant to either tap or omit a response based on the color of stimuli that appear in varying locations on the screen across 30 trials. The third subtest, CS, assesses visual scanning, attention, and learning by asking the participant whether a sequence of symbol-digit pairs matches the 9 pairs shown in the key [34]. The same 9 symbol-digit pairs appeared at every administration; stimuli were not varied across the 4 time points. Task complexity is not a formally operationalized or quantified construct but rather reflects the polyfactorial demands imposed by a task, that is, the extent to which multiple cognitive domains are simultaneously engaged during task performance [27,37,38]. For reference, SRT requires only a simple motor response to a single stimulus and is therefore the least complex. GNG introduces additional inhibitory control and rule-based responding, while CS incorporates a learning component and subtle working memory demands, attributes recognized as markers of greater task complexity [37,38]. This progression is also reflected in average RT, with SRT producing the fastest responses, followed by GNG, and then CS [33].

The primary metric for DANA tasks is Cognitive Efficiency (CE), computed as percent correct divided by the mean RT for correct responses, multiplied by 60,000 to express the result in correct responses per minute units. This formulation derives from the throughput metric defined by Thorne [39], with DANA using correct-trial RTs only rather than the cumulative RTs across correct and incorrect trials in Thorne's original metric. Most neuropsychological tests prioritize either speed or accuracy, leaving the other dimension less rigorously quantified. CE integrates both into a single performance index. CE was specified as the primary outcome for all PE analyses. Where CE-based analyses revealed a significant pattern, RT (milliseconds) and accuracy (percentage of correct responses) were examined as secondary outcomes to clarify which dimension was driving the observed effect.

Analyses

Descriptive statistics are presented for the largest test sample (CS; n=717), which represents the union of all participants included in the analytic sample, with demographic characteristics summarized in Table 1. Descriptive statistics for demographic characteristics are presented as means and SDs for continuous variables and counts and percentages for categorical variables. Test descriptive statistics include means for RT, accuracy, and CE.

PEs were defined as improvements in task performance across repeated administrations and were operationalized in two ways. First, effect sizes (ie, Cohen *d*) were calculated to quantify the magnitude of observed change across time points in pooled SD units, providing a standardized comparison across tasks. Second, linear mixed-effects models (LMM) were used to test the significance of PEs and to examine potential demographic moderators. Thus, model coefficients (unstandardized B) represent the estimated average change in CE per time point, with CE specified as the primary outcome variable and time (4 assessments) as the primary predictor. For any observed demographic moderation, additional analyses were conducted examining RT and accuracy as secondary outcomes.

XSL•FO
RenderX

Table . Clinical characteristics (N=717).

Characteristic	Sample values
Age (years), mean (SD)	64.1 (9.5)
Education level, n (%)	
College graduate	481 (67.1)
Some college	172 (24.0)
High school or less	64 (8.9)
Sex, n (%)	
Female	469 (65.4)
Male	248 (34.6)
Race, n (%)	
White	669 (93.3)
Black	14 (2.0)
Asian	13 (1.8)
Asian Indian/Native Hawaiian/Pacific Islander	1 (0.1)
Other[a]	20 (2.8)
FHS[b] cohort, n (%)	
Gen 3	515 (71.8)
Gen 2	131 (18.3)
Omni 2	41 (5.7)
Omni 1	20 (2.8)
NOS[c]	10 (1.4)

[a]Other includes Hispanic, more than one race, and prefer not to answer.

[b]FHS: Framingham Heart Study.

[c]NOS: New Offspring Spouse.

The main effect of time, representing PE while adjusting for age, sex, and education, was first tested using LMMs (*lme4::lmer*; R, version 4.3.1; R Foundation for Statistical Computing). Subsequent demographic models examined the influence of demographic variables at baseline and whether they moderated PE through interaction terms (time × demographic variable). To further evaluate age effects, an additional analysis dichotomized participants at 65 years of age for each DANA task. Age was modeled continuously in primary analyses; the dichotomized analysis used a conventional threshold of 65 years for older adulthood, yielding approximately balanced subgroups (55% aged 65 years or younger, 45% aged 65 years or older). Models included random intercepts for participants. Sensitivity analyses fitting random intercepts and slopes for time produced consistent fixed-effect estimates for the time term on CS and GNG, and an attenuated but directionally consistent estimate for SRT. Results from the random-intercept models are reported in the main text, and the results from the random intercepts and slopes are provided in Multimedia Appendix 1 (Tables S1 and S4).

To evaluate whether the analytic-sample restriction to participants who completed all 4 time points affected estimated PE slopes, the no-interaction main-effect-of-time model was refitted within subgroups of participants who completed 2 (n=302) or 3 (n=349) consecutive time points, separately for CS, GNG, and SRT. Results are reported in Multimedia Appendix 1 (Tables S1 and S3).

To formally test whether PE magnitude differed across tasks, a stacked LMM combining data from all 3 tasks was fit with task and time as fixed-effect predictors and a task × time interaction. Outcomes were standardized within each task by that task's pooled SD across time points, such that interaction coefficients are interpretable in standardized effect size units comparable to the Cohen *d* values reported for individual tasks. Pairwise contrasts (SRT vs CS, SRT vs GNG, GNG vs CS) were derived from the interaction model (Tables S1 and S5 in Multimedia Appendix 1).

Baseline characteristics were compared between completers (n=717) and partial completers (n=1039) to evaluate selection bias. Demographic comparisons used unadjusted 2-group tests (Welch *t* test for continuous variables and chi-square tests for categorical variables). Baseline (T1) DANA cognitive scores (CE, RT, and accuracy) on each task were compared using linear models adjusted for age, sex, and education to isolate cognition-specific differences from those driven by demographic imbalance (Tables S1 and S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Ethical Considerations

The procedures and protocols of the FHS were approved by the institutional review board of the Boston University Medical Campus (H-40147). Written informed consent was obtained from all participants. The analyzed data did not include personal identifiers and were linked to a randomly generated identification number that cannot be used to identify participants personally. Study investigators never received or had access to identifying information. Since no identifying information was collected and no linkage file exists, there is no reasonable way to deduce participant identities from the data. The standard operating procedures used to protect the security of the data, along with other protections, are more than adequate to ensure the privacy of research participants and the confidentiality of the data. Participants were not financially compensated for participating in the study.

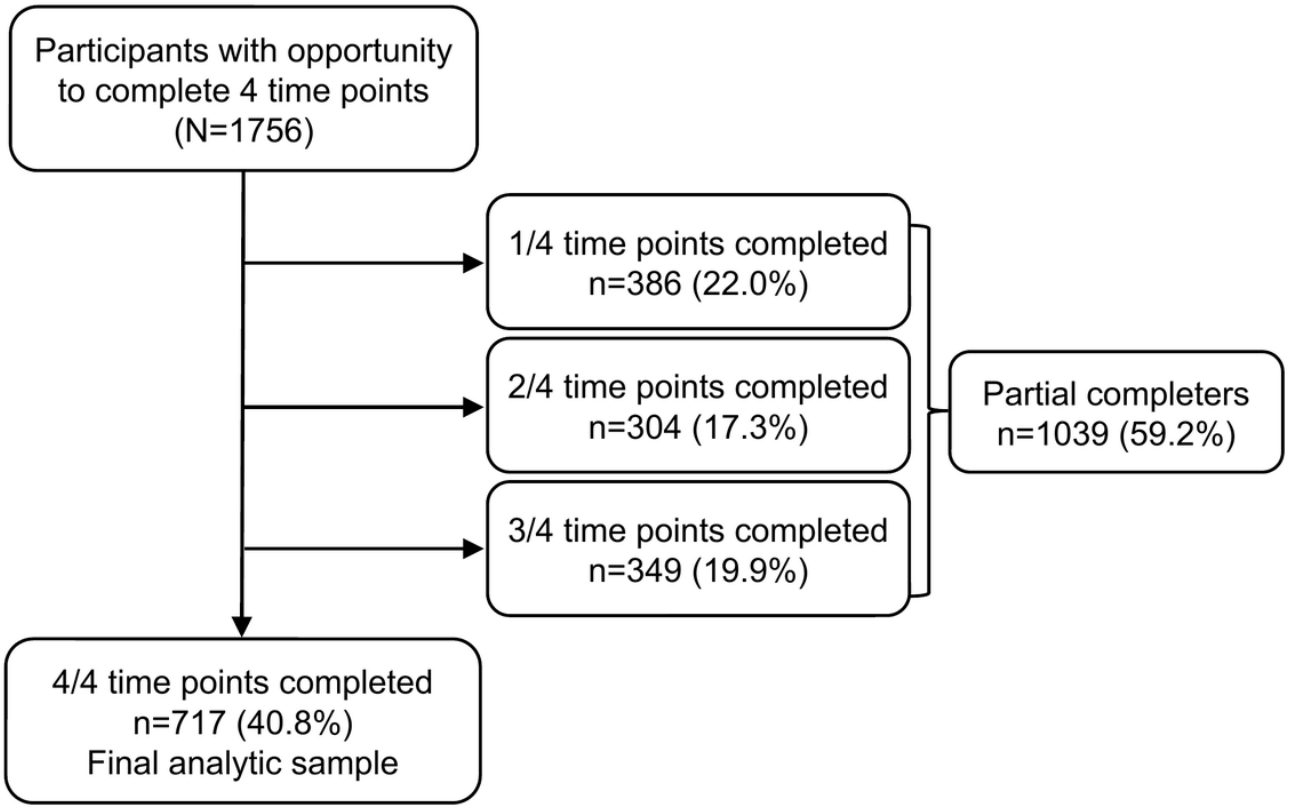
Results

Descriptive Characteristics

The sample comprising 717 participants (Figure 1; mean age 64.1, SD 9.5 years) was mostly White (n=669, 93.3%), female (n=469, 65.4%), and college graduates (n=481, 67.1%), primarily from the Gen 3 cohort (n=515, 71.8%, Table 1). Mean accuracy for T1 and T4 and the change in RT were as follows: for SRT, accuracy was 98.83% at T1 and 99.44% at T4, and RT decreased from 375.76 to 363.80 milliseconds; for GNG, accuracy was 96.81% at T1 and 97.36% at T4, and RT decreased from 883.46 to 843.95 milliseconds; and for CS, accuracy was 95.96% at T1 and 97.27% at T4, and RT decreased from 1941.46 to 1788.82 milliseconds (Table 2). Accuracy and RT showed an inverse relationship across tasks such that CS had the lowest accuracy and highest RT, SRT the highest accuracy and lowest RT, and GNG was intermediate on both. These descriptive trends corresponded to higher composite CE scores across assessments (Table 2).

Compared with partial completers (n=1039), the participants who completed all 4 time points were older (mean age 64.1, SD 9.5 years vs mean age 61.4, SD 10.1 years; $P<.001$), more often female (469/717, 65.4% vs 566/1039, 54.5%; $P<.001$), and more highly educated (481/717, 67.1% vs 101/1039, 9.7% college graduates; $P<.001$), with no difference in racial composition ($P=.39$; see Table S1 in Multimedia Appendix 1). After adjustment for age, sex, and education, groups did not differ in T1 CE, the primary outcome, on any task (see Tables S1 and S2 in Multimedia Appendix 1).

Figure 1. Participant selection based on adherence to the assessment schedule.



XSL•FO
RenderX

Table . DANA[a] subtest descriptives.

Task and visit	RT[b] (milliseconds), mean (SD)	Accuracy (%), mean (SD)	CE[c], mean (SD)
SRT[d]			
T1	375.76 (78.01)	98.83 (6.27)	164.84 (30.23)
T2	366.78 (69.80)	99.37 (2.49)	168.09 (27.67)
T3	365.48 (68.44)	99.35 (4.31)	168.71 (27.31)
T4	363.80 (66.70)	99.44 (1.95)	169.43 (26.86)
Average[e]	367.95 (70.74)	99.25 (3.76)	167.77 (28.02)
GNG[f]			
T1	883.46 (134.08)	96.81 (12.02)	93.52 (22.91)
T2	865.96 (130.40)	96.86 (11.68)	96.42 (23.23)
T3	851.95 (118.77)	97.59 (9.51)	99.50 (22.19)
T4	843.95 (127.36)	97.36 (9.17)	100.68 (22.61)
Average[e]	861.33 (127.65)	97.16 (10.59)	97.53 (22.74)
CS[g]			
T1	1941.46 (404.04)	95.96 (9.08)	31.22 (6.91)
T2	1857.82 (388.85)	96.50 (6.90)	32.77 (7.27)
T3	1814.56 (393.21)	97.00 (5.52)	33.82 (7.50)
T4	1788.82 (382.85)	97.27 (3.84)	34.34 (7.58)
Average[e]	1850.67 (392.24)	96.68 (6.34)	33.04 (7.32)

[a]DANA: Defense Automated Neurocognitive Assessment.

[b]RT: response time.

[c]CE: Cognitive Efficiency.

[d]SRT: Simple Reaction Time.

[e]Average: mean of the means across the 4 time points; the value in parentheses is the mean of the SDs across the 4 time points.

[f]GNG: Go/No-Go.

[g]CS: Code Substitution.

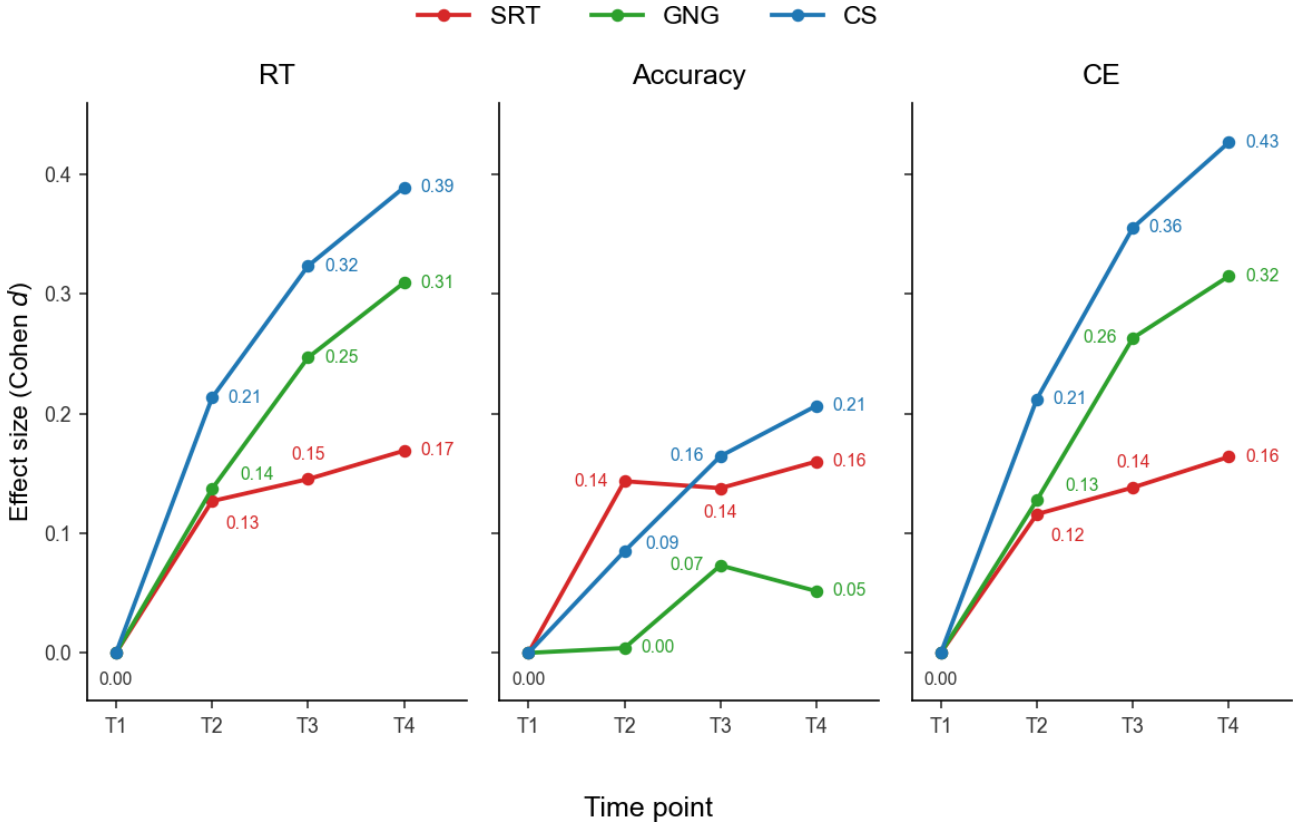
PEs

Without considering demographic influences, PEs on CE scores were observed across all subtests, reflected in significant main effects of time in the no-interaction models (Table 3). The magnitude of these PEs is shown in Figure 2, where improvements in CE from T1 to T4 were small for CS (d=0.43) and GNG (d=0.32) and below the small threshold for SRT (d=0.16). A stacked LMM with a task × time interaction confirmed that practice-effect magnitude differed across tasks, following the same ordering as the effect sizes, with all pairwise contrasts statistically significant (CS>GNG>SRT; see Tables S1 and S5 in Multimedia Appendix 1). Effect size trends in speed and accuracy indicated that changes for CS were driven by improvements in both RT and accuracy, while changes for GNG were driven almost exclusively by RT (Figure 2), reflecting baseline accuracy levels that were already at or near ceiling on GNG (mean 96.81%, SD 12.02%) and SRT (mean 98.83%, SD 6.27%). The slight decrease in cumulative effect size at T3 for SRT accuracy reflects a transient decline in performance between T2 and T3, though accuracy remained improved relative to T1 baseline. CS was the only subtest where change in accuracy reached the threshold for a small effect (d=0.21). For SRT, RT (d=0.17) and accuracy (d=0.16) contributed more evenly.

Sensitivity analyses, including those with partial completers (Tables S1 and S3 in Multimedia Appendix 1) and those fitting random intercepts and slopes for time (Tables S1 and S4 in Multimedia Appendix 1), showed no differences from the primary findings for CS or GNG; estimates for SRT were attenuated.

XSL•FO
RenderX

Figure 2. DANA subtest descriptives by effect size. Figures show the cumulative effect size. Cohen *d* thresholds: 0.20 (small effect), 0.50 (medium effect), and 0.80 (large effect). CE: Cognitive Efficiency; CS: Code Substitution; GNG: Go/No-Go; RT: response time; SRT: Simple Reaction Time.



XSL•FO
RenderX

Table . Practice effects models with demographic interactions.

Model and predictors	SRT[a]		GNG[b]		CS[c]	
	B[d] (SE)	*P* value	B[d] (SE)	*P* value	B[d] (SE)	*P* value
No-interaction models						
Intercept	213.21 (6.92)	<.001	155.86 (4.77)	<.001	56.07 (1.57)	<.001
Time	1.64 (0.28)	<.001	2.73 (0.23)	<.001	1.15 (0.05)	<.001
Age models						
Intercept	212.41 (7.49)	<.001	152.59 (5.30)	<.001	53.53 (1.65)	<.001
Time	2.18 (1.93)	.26	4.91 (1.56)	.002	2.83 (0.34)	<.001
Age	−0.80 (0.10)	<.001	−1.03 (0.07)	<.001	−0.40 (0.02)	<.001
Time × age	−0.01 (0.03)	.78	−0.03 (0.02)	.16	−0.03 (0.01)	<.001
Sex models						
Intercept	214.24 (6.95)	<.001	156.08 (4.80)	<.001	56.08 (1.57)	<.001
Time	0.97 (0.49)	.045	2.58 (0.39)	<.001	1.15 (0.08)	<.001
Sex (female)	−5.80 (2.02)	.004	−3.64 (1.44)	.01	0.69 (0.44)	.10
Time × sex	1.02 (0.60)	.09	0.22 (0.48)	.65	0.00 (0.10)	.97
Education models						
Intercept	212.67 (7.07)	<.001	156.63 (4.91)	<.001	56.39 (1.59)	<.001
Time	2.02 (1.00)	.04	2.22 (0.79)	.005	0.93 (0.17)	<.001
Some college	3.26 (3.90)	.40	6.24 (2.75)	.02	2.31 (0.82)	.005
College graduate	9.85 (3.60)	.006	10.33 (2.52)	<.001	2.77 (0.75)	<.001
Time × some college	0.72 (1.15)	.53	0.43 (0.92)	.64	0.21 (0.20)	.29
Time × college graduate	−0.82 (1.06)	.44	0.61 (0.84)	.47	0.25 (0.18)	.16

[a]SRT: Simple Reaction Time.

[b]GNG: Go/No-Go.

[c]CS: Code Substitution.

[d]B: unstandardized fixed-effect estimate.

Age, sex, and education models showed varied patterns of demographic-related differences in baseline performance and in their influence on PE. Older age, modeled continuously, predicted lower baseline CE scores on all tasks. The only significant age × time interaction was on CS, indicating smaller PE with increasing age. To examine categorical age effects, participants were dichotomized into younger (≤65 years; 55% of the sample) and older (>65 years) groups (Table 4). This approach confirmed what was observed in the continuous models: age-related moderation of PE occurred only for CS (B=−0.39, SE=0.10; *P*<.001). Additional analyses examining differences in speed and accuracy demonstrated that this moderation was driven by accuracy (B=0.46, 95% CI 0.09 - 0.83; *P*=.02) rather than RT (B=2.41, 95% CI −8.60 to 13.42; *P*=.67). Sex-related findings showed that at baseline, females scored lower than males on SRT and GNG, whereas no baseline sex-related differences were observed on CS. However, no significant sex × time interactions were observed, indicating that PE did not differ by sex. Education effects showed that, relative to participants with a high school degree or less, those with higher education scored better at baseline on all tasks. No significant education × time interactions were observed, indicating that PE did not vary by education.

XSL•FO
RenderX

Table . Difference in practice effects slopes for older participants (aged >65 years) compared with younger participants (aged ≤65 years).

Task	B[a] (SE)	*P* value
SRT[b]	−0.93 (0.58)	.11
GNG[c]	−0.56 (0.47)	.24
CS[d]	−0.39 (0.10)	<.001

[a]B: unstandardized fixed-effect estimate.

[b]SRT: Simple Reaction Time.

[c]GNG: Go/No-Go.

[d]CS: Code Substitution.

Discussion

Principal Findings

PEs were observed across all 3 DANA tasks. The magnitude of PE increased with task complexity, likely due to the complex tasks provide greater opportunity for improvement. Among demographic variables, only age moderated performance over time, with older adults (aged 65 years or older) showing smaller PE on the CS task. This age-related moderation effect on the CS was driven primarily by accuracy rather than speed, which contrasted with the general trend of speed driving improvements in composite CE scores.

Early studies of DANA reported little evidence of PE [34,36,40], likely due to small samples, short retest intervals, preliminary practice sessions, and reliance on younger, predominantly male cohorts with high baseline performance. More recent work by Bellitti et al [27] found modest RT improvements across 2 sessions in older adults, with CS showing the largest effects and emerging as the only task differentiating cognitively intact from impaired participants. This pattern may reflect the polyfactorial nature of CS, that is, its simultaneous recruitment of multiple cognitive domains [37]. Our results are consistent with this pattern, with CS showing the largest PE, followed by GNG and then SRT in the current study. Direct comparison across studies is limited because Bellitti et al [27] did not report raw RTs and used a mixed sample of cognitively intact and impaired participants. Our larger sample revealed PE on both GNG and SRT, whereas the previous study observed no change on SRT. Prior studies have also found that CS performance relates to a broader set of abilities than SRT or GNG, with the CS task being associated with executive function, attention, processing speed, and memory [32]. Another study also found that the CS task was more strongly associated with a global measure of cognition, the Mini-Mental State Examination, than were SRT and GNG [41]. Interestingly, in a previous study, when the DANA app was administered remotely, the CS task could not be reliably completed by patients diagnosed with AD [42]. One reason could be that this task was particularly challenging to complete remotely without help for those with deteriorating cognitive abilities. Together, these findings support our description of CS as more complex than the other DANA tasks and suggest that this task might be a more sensitive task for detecting subtle impairment over time. Future longitudinal studies tracking these cohorts will be required to validate these suppositions.

Our findings also suggest that task complexity contributes to the presence of PE, but the unique contributions of speed and accuracy remain less clear. Bellitti et al [27] analyzed only speed, citing uniformly high accuracy at baseline as the reason for not examining accuracy. Our results confirm that most improvements across time were speed-driven but also indicate that the age-related moderation of PE on CS was more strongly linked to accuracy.

Speed vs Accuracy

The ability of digital assessments to measure both speed and accuracy and to combine them into composite throughput measures [39] not only creates opportunities but also introduces new analytic challenges in interpreting higher-dimensional data. Cross-sectionally, variability in RT has been more closely linked to AD than accuracy [43], and tests with a majority speed-based component tend to produce the largest PEs [44]. However, Bergman et al [44] defined speed tasks using computerized iterations of legacy measures (ie, traditional neuropsychological tests) such as Trails A (approximately 20 - 30 seconds), distinguishing them from simple reaction tasks such as SRT (approximately 3 - 4 seconds), which, in the current study, is considered the most direct measure of speed. This distinction underscores the importance of carefully translating legacy constructs into digital formats. Relative to other measures in their respective eras, both Trails A and SRT are considered prototypical indicators of processing speed. Yet their susceptibility to PE differs, as Trails A is both more complex due to its demands on visual scanning and motor coordination and is substantially longer in duration. Bergman et al [44] also administered their tasks on laptops with external mice, whereas the present study relied on smartphones. Even within modern digital platforms, operating system and hardware differences (eg, iOS vs Android) can influence response capture and timing precision, raising additional concerns about generalizability across devices [45,46]. Thus, task duration, complexity, and device type likely increase susceptibility to PE, but disentangling their relative contributions is challenging. Whereas task duration is easy to quantify, task complexity is more abstract and difficult to measure directly. More broadly, this demonstrates that although tests are operationalized precisely, they often recruit multiple domains, and the relationship between tests and cognitive constructs is fluid [47].

PE and Defining Change

This study adds to the growing literature on PE, including studies using digital cognitive assessments [20], traditional tests

XSL•FO
RenderX

across short intervals [48], and those examining cognitive trajectories alongside concurrent biomarkers or lifestyle factors [49-51]. Much of the current high-frequency cognitive testing literature emphasizes memory tasks due to their sensitivity to neuropathological changes, particularly in AD [28,29]. In contrast, this study demonstrates that digital tasks measuring other predominantly nonmemory abilities, including processing speed, sustained attention, and visual scanning, also exhibit detectable PE.

Characterizing PEs in nonmemory tasks within a large sample of cognitively intact adults provides foundational data for research examining diminished PEs as a potential indicator of cognitive decline. Documenting that PEs occur in this population establishes a baseline for evaluating whether the absence or attenuation of PEs can distinguish impaired from intact individuals across neurodegenerative conditions and neurological injuries, where small but measurable deviations from normal trajectories may have clinical significance.

The utility of PE as a potential subtle indicator of cognitive impairment depends on careful consideration of assessment parameters, as PEs vary based on testing intervals, the specific cognitive domains assessed, the types of devices used, and how performance is quantified (speed, accuracy, or composite measures). Digital cognitive assessments vary considerably not only in the types of tests and cognitive abilities they measure but also in the time intervals between assessments, ranging from intensive longitudinal or measurement burst designs with multiple assessments in one day to weekly, monthly, quarterly, or annual follow-ups. Validation in cognitively impaired groups is a natural next step before PE from these brief digital tasks can be advanced as candidate clinical indicators.

The changes in performance observed in this study were significant but small in magnitude (SRT: d=0.16, GNG: d=0.32, CS: d=0.43). Given that effect sizes are expressed in SD units, these values reflect improvements of less than half an SD across 4 time points. Conventional neuropsychological interpretation relies on a ±1.645 SD threshold to define reliable change, a widely used but arbitrary cutoff intended to detect overt dysfunction [52]. Meanwhile, contemporary perspectives emphasize that neuropathological changes emerge before overt cognitive disruption, underscoring the potential relevance of smaller within-person shifts in cognition [7]. Progress toward this more nuanced perspective is reflected in recent efforts toward a “change-on-change” approach that interprets cognitive change alongside neuropathological change [49]. The current study clarified task-specific patterns of PE on DANA in a cognitively intact sample. Future research should aim to determine thresholds of clinical significance for PE, recognizing that meaningful change may be subtle and interpreted differently if occurring concurrently with neuropathological change.

Strengths and Limitations

This study has several strengths. It represents the largest investigation of PE using the DANA platform, with participants completing 4 assessments across a 1-year period. This study contributes to the understanding of quarterly assessments, which is a relatively understudied temporal frequency situated between intensive longitudinal burst paradigms and long-interval observational studies. Another strength is the use of composite efficiency metrics that integrate both speed and accuracy, allowing both variables to contribute to the overall score.

There are also important limitations. The sample was predominantly White, female, and highly educated, which constrains generalizability. Analyses were conducted on summary scores generated by the DANA platform rather than item-level or trial-level data, which may have limited sensitivity to finer-grained practice-related changes. Also, to ensure consistency in examining PE, we only included participants who completed all 4 time points. Although sensitivity analyses of 2- and 3-time-point subsets (Tables S1 and S3, Multimedia Appendix 1) revealed statistically significant PE on CS and GNG among noncompleters, mirroring the ordinal pattern observed in the primary sample, the smaller effect magnitudes in these subsets suggest that estimates derived from the current completer sample may represent an upper bound on population-level PE among cognitively intact older adults. However, adherence to longitudinal testing is itself an important behavioral variable [53], and it remains unclear how missed or off-schedule assessments might influence the occurrence of PE. A comparison of baseline cognition between completers and partial completers showed no difference in CE after adjustment for age, sex, and education (Tables S1 and S2 in Multimedia Appendix 1). We also did not measure contextual factors such as comfort with the testing application [31] or momentary states such as arousal and distraction, which recent evidence indicates can affect cognitive performance in remote settings [54]. Finally, this study focused exclusively on cognitively intact participants without integrating biomarkers, limiting the ability to evaluate how PEs interact with underlying neuropathological changes. Future work could examine whether baseline cognitive performance moderates the magnitude of PE on these tasks.

Conclusions

This study demonstrates that PEs are detectable across multiple digital cognitive tasks, with the largest effects observed on more complex tasks. Older adults showed smaller PEs on the CS task, and this trend was driven more by accuracy than speed. These results underscore the importance of considering task demands, as well as the relative contributions of speed and accuracy, when interpreting digital cognitive data. More broadly, the findings highlight the need to move beyond legacy thresholds for reliable change, which were designed to detect overt functional decline, and toward models that evaluate incremental within-person variability embedded within contextual factors as a meaningful signal in the aging process.

XSL•FO
RenderX

Acknowledgments

The authors acknowledge the FHS participants for their dedication, as this study would not have been possible without them. The authors also thank the researchers at FHS for their efforts over the years in examining participants. The authors declare the use of generative AI to assist with editing the manuscript text and formatting tables (Claude, Anthropic).

Funding

This research was supported by the National Institute on Aging of the National Institutes of Health under award numbers U01-AG068221, R01-AG062109, R01-AG072654, U19-AG068753, R00 AG062783, and R01-AG083735. Additional support was provided under contract HHSN269201500001C and by the Alzheimer Drug Discovery Foundation.

Data Availability

Data, methods, and materials used to conduct the research in this paper, along with additional information, can be shared upon reasonable request through a formal research application to the Framingham Heart Study. Requests to access these datasets should be directed to the Framingham Heart Study [55].

Authors' Contributions

KC conceptualized the project, drafted the original manuscript, proposed an analytical plan, and created visualizations. CL conducted the formal analyses and created visualizations. CK was responsible for data curation. HD provided supervision and guidance on the analytic approach. VBK, HL, and RA contributed to funding acquisition, methodology, conceptualization, and critical review of the manuscript content. SD, AFAA, IDA-D, and KAG also provided critical review of the manuscript content. PHH provided supervision and guidance on the analytic approach. PS contributed to project conceptualization, drafting of the original manuscript, data collection, study design, analytical approach, project administration, and supervision.

Conflicts of Interest

RA is a scientific advisor for Signant Health and Novo Nordisk, which are unrelated to the work presented.

Multimedia Appendix 1
Analyses of sample characteristics, baseline cognitive scores, time effects, model specifications, and cross-task practice effects.
[DOCX File, 20 KB - aging_v9i1e87208_app1.docx]

References

1. McCaffrey RJ, Ortega A, Orsillo SM, Nelles WB, Haase RF. Practice effects in repeated neuropsychological assessments. ABPP-CN 1992 Jan;6(1):32-42. [doi: 10.1080/13854049208404115]
2. Goldberg TE, Harvey PD, Wesnes KA, Snyder PJ, Schneider LS. Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. Alzheimers Dement (Amst) 2015 Mar;1(1):103-111. [doi: 10.1016/j.dadm.2014.11.003] [Medline: 27239497]
3. Jutten RJ, Grandoit E, Foldi NS, et al. Lower practice effects as a marker of cognitive performance and dementia risk: a literature review. Alzheimers Dement (Amst) 2020;12(1):e12055. [doi: 10.1002/dad2.12055] [Medline: 32671181]
4. Salthouse TA. Robust cognitive change. J Int Neuropsychol Soc 2012 Jul;18(4):749-756. [doi: 10.1017/S1355617712000380] [Medline: 22595019]
5. Hassenstab J, Ruvolo D, Jasielec M, Xiong C, Grant E, Morris JC. Absence of practice effects in preclinical Alzheimer's disease. Neuropsychology 2015 Nov;29(6):940-948. [doi: 10.1037/neu0000208] [Medline: 26011114]
6. Machulda MM, Pankratz VS, Christianson TJ, et al. Practice effects and longitudinal cognitive change in normal aging vs. incident mild cognitive impairment and dementia in the Mayo Clinic Study of Aging. Clin Neuropsychol 2013;27(8):1247-1264. [doi: 10.1080/13854046.2013.836567] [Medline: 24041121]
7. Jack CJ, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. Alzheimer's & Dementia 2024 Aug;20(8):5143-5169. [doi: 10.1002/alz.13859] [Medline: 38934362]
8. Papp KV, Samaroo A, Chou HC, et al. Unsupervised mobile cognitive testing for use in preclinical Alzheimer's disease. Alzheimers Dement (Amst) 2021;13(1):e12243. [doi: 10.1002/dad2.12243] [Medline: 34621977]
9. Sanderson-Cimino M, Chen R, Tu XM, Elman JA, Jak AJ, Kremen WS. Misinterpreting cognitive change over multiple timepoints: when practice effects meet age-related decline. Neuropsychology 2023 Jul;37(5):568-581. [doi: 10.1037/neu0000903] [Medline: 37079809]
10. Sanderson-Cimino M, Elman JA, Tu XM, et al. Cognitive practice effects delay diagnosis of MCI: implications for clinical trials. Alzheimers Dement (N Y) 2022;8(1):e12228. [doi: 10.1002/trc2.12228] [Medline: 35128027]
11. Bartels C, Wegrzyn M, Wiedl A, Ackermann V, Ehrenreich H. Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing. BMC Neurosci 2010 Sep 16;11:118. [doi: 10.1186/1471-2202-11-118] [Medline: 20846444]

XSL•FO
RenderX

12. Calamia M, Markon K, Tranel D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. Clin Neuropsychol 2012;26(4):543-570. [doi: 10.1080/13854046.2012.680913] [Medline: 22540222]
13. Aschenbrenner AJ, Hassenstab J, Wang G, et al. Avoid or embrace? Practice effects in Alzheimer's disease prevention trials. Front Aging Neurosci 2022;14:883131. [doi: 10.3389/fnagi.2022.883131] [Medline: 35783127]
14. Oravecz Z, Harrington KD, Hakun JG, et al. Accounting for retest effects in cognitive testing with the Bayesian double exponential model via intensive measurement burst designs. Front Aging Neurosci 2022;14:897343. [doi: 10.3389/fnagi.2022.897343] [Medline: 36225891]
15. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and validity of ambulatory cognitive assessments. Assessment 2018 Jan;25(1):14-30. [doi: 10.1177/1073191116643164] [Medline: 27084835]
16. Weizenbaum EL, Soberanes D, Hsieh S, et al. Capturing learning curves with the multiday Boston Remote Assessment of Neurocognitive Health (BRANCH): feasibility, reliability, and validity. Neuropsychology 2024 Feb;38(2):198-210. [doi: 10.1037/neu0000933] [Medline: 37971862]
17. Basso MR, Carona FD, Lowery N, Axelrod BN. Practice effects on the WAIS-III across 3- and 6-month intervals. Clin Neuropsychol 2002 Feb;16(1):57-63. [doi: 10.1076/clin.16.1.57.8329] [Medline: 11992227]
18. Estevis E, Basso MR, Combs D. Effects of practice on the Wechsler Adult Intelligence Scale-IV across 3- and 6-month intervals. Clin Neuropsychol 2012;26(2):239-254. [doi: 10.1080/13854046.2012.659219] [Medline: 22353021]
19. Salthouse TA, Schroeder DH, Ferrer E. Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. Dev Psychol 2004 Sep;40(5):813-822. [doi: 10.1037/0012-1649.40.5.813] [Medline: 15355168]
20. Duff K, Hammers DB, Koppelmans V, King JB, Hoffman JM. Short-term practice effects on cognitive tests across the late life cognitive spectrum and how they compare to biomarkers of Alzheimer's disease. J Alzheimers Dis 2024;99(1):321-332. [doi: 10.3233/JAD-231392] [Medline: 38669544]
21. Hammers DB, Suhrie KR, Porter SM, Dixon AM, Duff K. Validation of one-year reliable change in the RBANS for community-dwelling older adults with amnestic mild cognitive impairment. Clin Neuropsychol 2022 Aug;36(6):1304-1327. [doi: 10.1080/13854046.2020.1807058] [Medline: 32819188]
22. Salthouse TA. Influence of age on practice effects in longitudinal neurocognitive change. Neuropsychology 2010 Sep;24(5):563-572. [doi: 10.1037/a0019026] [Medline: 20804244]
23. Duff K, Callister C, Dennett K, Tometich D. Practice effects: a unique cognitive variable. Clin Neuropsychol 2012;26(7):1117-1127. [doi: 10.1080/13854046.2012.722685] [Medline: 23020261]
24. Eliassen IV, Kirsebom BE, Fladby T, et al. Regression-based cognitive change norms applied in biochemically defined predementia Alzheimer's disease. Neuropsychology 2023 Jan;37(1):32-43. [doi: 10.1037/neu0000853] [Medline: 36136790]
25. Gavett BE, Gurnani AS, Saurman JL, et al. Practice effects on story memory and list learning tests in the neuropsychological assessment of older adults. PLoS One 2016;11(10):e0164492. [doi: 10.1371/journal.pone.0164492] [Medline: 27711147]
26. Jutten RJ, Rentz DM, Fu JF, et al. Monthly at-home computerized cognitive testing to detect diminished practice effects in preclinical Alzheimer's disease. Front Aging Neurosci 2022;13:800126. [doi: 10.3389/fnagi.2021.800126] [Medline: 35095476]
27. Bellitti M, Lauber MV, Searls E, Lin H, Au R, Kolachalama VB. Practice effects on digital cognitive assessment tools: insights from the defense automated neurobehavioral assessment battery. Alzheimers Dement 2025 Sep;21(9):e70644. [doi: 10.1002/alz.70644] [Medline: 40911708]
28. Samaroo A, Amariglio RE, Burnham S, et al. Diminished Learning Over Repeated Exposures (LORE) in preclinical Alzheimer's disease. Alzheimers Dement (Amst) 2021;12(1):e12132. [doi: 10.1002/dad2.12132] [Medline: 33426266]
29. Young CB, Mormino EC, Poston KL, et al. Computerized cognitive practice effects in relation to amyloid and tau in preclinical Alzheimer's disease: results from a multi-site cohort. Alzheimers Dement (Amst) 2023;15(1):e12414. [doi: 10.1002/dad2.12414] [Medline: 36950699]
30. Christianson K, Prabhu M, Popp ZT, et al. Feasibility of frequent self‐administered mobile cognitive assessments in adults with and without cognitive impairment. Alzheimers Dement 2023 Jun;19(S4):e066006. [doi: 10.1002/alz.066006]
31. Sunderaraman P, De Anda-Duran I, Karjadi C, et al. Design and feasibility analysis of a smartphone-based digital cognitive assessment study in the Framingham Heart Study. J Am Heart Assoc 2024 Jan 16;13(2):e031348. [doi: 10.1161/JAHA.123.031348] [Medline: 38226510]
32. De Anda-Duran I, Sunderaraman P, Searls E, et al. Comparing cognitive tests and smartphone-based assessment in 2 US community-based cohorts. J Am Heart Assoc 2024 Jan 16;13(2):e032733. [doi: 10.1161/JAHA.123.032733] [Medline: 38226519]
33. Ding H, Kim M, Searls E, et al. Digital neuropsychological measures by Defense Automated Neurocognitive Assessment: reference values and clinical correlates. Front Neurol 2024;15:1340710. [doi: 10.3389/fneur.2024.1340710] [Medline: 38426173]
34. Lathan C, Spira JL, Bleiberg J, Vice J, Tsao JW. Defense Automated Neurobehavioral Assessment (DANA)-psychometric properties of a new field-deployable neurocognitive assessment tool. Mil Med 2013 Apr;178(4):365-371. [doi: 10.7205/MILMED-D-12-00438] [Medline: 23707818]

XSL•FO
RenderX

35. Yuan J, Maserejian N, Liu Y, et al. Severity distribution of Alzheimer's disease dementia and mild cognitive impairment in the Framingham Heart Study. J Alzheimers Dis 2021;79(2):807-817. [doi: 10.3233/JAD-200786] [Medline: 33361590]
36. Russo CR, Lathan CE. An evaluation of the consistency and reliability of the Defense Automated Neurocognitive Assessment Tool. Appl Psychol Meas 2015 Oct;39(7):566-572. [doi: 10.1177/0146621615577361] [Medline: 29881027]
37. Jaeger J. Digit Symbol Substitution Test: the case for sensitivity over specificity in neuropsychological testing. J Clin Psychopharmacol 2018 Oct;38(5):513-519. [doi: 10.1097/JCP.0000000000000941] [Medline: 30124583]
38. Aziz-Safaie T, Müller VI, Langner R, Eickhoff SB, Cieslik EC. The effect of task complexity on the neural network for response inhibition: an ALE meta-analysis. Neurosci Biobehav Rev 2024 Mar;158:105544. [doi: 10.1016/j.neubiorev.2024.105544] [Medline: 38220034]
39. Thorne DR. Throughput: a simple performance index with desirable characteristics. Behav Res Methods 2006 Nov;38(4):569-573. [doi: 10.3758/bf03193886] [Medline: 17393825]
40. Haran FJ, Dretsch MN, Bleiberg J. Performance on the Defense Automated Neurobehavioral Assessment across controlled environmental conditions. Appl Neuropsychol Adult 2016;23(6):411-417. [doi: 10.1080/23279095.2016.1166111] [Medline: 27182844]
41. Hollinger KR, Woods SR, Adams-Clark A, et al. Defense Automated Neurobehavioral Assessment accurately measures cognition in patients undergoing electroconvulsive therapy for major depressive disorder. J ECT 2018 Mar;34(1):14-20. [doi: 10.1097/YCT.0000000000000448] [Medline: 28991066]
42. Lathan C, Coffman I, Shewbridge R, Marissa L, Fonzetti P, Doraiswamy PM. A pilot to investigate the feasibility of mobile cognitive assessment of elderly patients and caregivers in the home. J Geriatr Palliat Care 2016;4(1). [doi: 10.13188/2373-1133.1000017]
43. Christ BU, Combrinck MI, Thomas KGF. Both reaction time and accuracy measures of intraindividual variability predict cognitive performance in Alzheimer's disease. Front Hum Neurosci 2018;12:124. [doi: 10.3389/fnhum.2018.00124] [Medline: 29686610]
44. Bergman I, Franke Föyen L, Gustavsson A, Van den Hurk W. Test-retest reliability, practice effects and estimates of change: a study on the Mindmore digital cognitive assessment tool. Scand J Psychol 2025 Feb;66(1):1-14. [doi: 10.1111/sjop.13054] [Medline: 39072723]
45. Germine L, Reinecke K, Chaytor NS. Digital neuropsychology: challenges and opportunities at the intersection of science and software. Clin Neuropsychol 2019 Feb;33(2):271-286. [doi: 10.1080/13854046.2018.1535662] [Medline: 30614374]
46. Germine L, Strong RW, Singh S, Sliwinski MJ. Toward dynamic phenotypes and the scalable measurement of human behavior. Neuropsychopharmacology 2021 Jan;46(1):209-216. [doi: 10.1038/s41386-020-0757-1] [Medline: 32629456]
47. Bilder RM, Reise SP. Neuropsychological tests of the future: how do we get there from here? Clin Neuropsychol 2019 Feb;33(2):220-245. [doi: 10.1080/13854046.2018.1521993] [Medline: 30422045]
48. Duff K, Dixon A, Embree L. A closer look at practice effects in mild cognitive impairment and Alzheimer's disease. Arch Clin Neuropsychol 2024 Jan 19;39(1):1-10. [doi: 10.1093/arclin/acad046] [Medline: 37323010]
49. Klinger HM, Healy BC, Hanseeuw BJ, et al. Latent change-on-change between amyloid accumulation and cognitive decline. Alzheimers Dement 2024 Dec;20(12):8728-8738. [doi: 10.1002/alz.14326] [Medline: 39470175]
50. De Vito AN, Kunicki ZJ, Joyce HE, Huey ED, Jones RN. Parallel changes in cognition, neuropsychiatric symptoms, and amyloid in cognitively unimpaired older adults and those with mild cognitive impairment. Alzheimers Dement 2025 Feb;21(2):e14568. [doi: 10.1002/alz.14568] [Medline: 39936256]
51. Kekäläinen T, Luchetti M, Terracciano A, et al. Physical activity and cognitive function: moment-to-moment and day-to-day associations. Int J Behav Nutr Phys Act 2023 Nov 22;20(1):137. [doi: 10.1186/s12966-023-01536-9] [Medline: 37993862]
52. Duff K. Evidence-based indicators of neuropsychological change in the individual patient: relevant concepts and methods. Arch Clin Neuropsychol 2012 May;27(3):248-261. [doi: 10.1093/arclin/acr120] [Medline: 22382384]
53. Christianson K, Prabhu M, Popp ZT, et al. Adherence timescale impacts completion rates of high-frequency mobile cognitive assessments among older adults. Explor Med 2025;6. [doi: 10.37349/emed.2025.1001356] [Medline: 41050829]
54. Brearly TW, Elbich DB, Dhinojwala M, Hakun JG. Contextualized cognition: clarifying associations between remote unsupervised performance and clinically relevant contextual factors. Clin Neuropsychol 2026 Jul;40(5):1443-1468. [doi: 10.1080/13854046.2025.2534720] [Medline: 40749706]
55. FHS – for researchers. Framingham Heart Study. 2026. URL: https://www.framinghamheartstudy.org/fhs-for-researchers/ [accessed 2026-10-07]

Abbreviations

AD: Alzheimer disease
CE: Cognitive Efficiency
CS: Code Substitution
DANA: Defense Automated Neurocognitive Assessment
FHS: Framingham Heart Study
GNG: Go/No-Go

XSL•FO
RenderX

LMM: linear mixed-effects model
PE: practice effect
RT: response time
SRT: Simple Reaction Time

Edited by J Wang; submitted 06.Nov.2025; peer-reviewed by K Papp, L Zając-Lamparska; revised version received 25.Sep.2026; accepted 26.Sep.2026; published 19.Oct.2026.

Please cite as:
Christianson K, Lyu C, Karjadi C, Devine S, Ang TFA, De Anda-Duran I, Kolachalama VB, Lin H, Ding H, Gifford KA, Au R, Hwang PH, Sunderaraman P
Practice Effects on Repeated Self-Administered Digital Cognitive Assessments in Cognitively Intact Older Adults from the Framingham Heart Study: Longitudinal Observational Study
JMIR Aging 2026;9:e87208
URL: https://aging.jmir.org/2026/1/e87208
doi:10.2196/87208

© Kieffer Christianson, Chenglin Lyu, Cody Karjadi, Sherral Devine, Ting Fang Alvin Ang, Ileana De Anda-Duran, Vijaya B Kolachalama, Honghuang Lin, Huitong Ding, Katherine A Gifford, Rhoda Au, Phillip H Hwang, Preeti Sunderaraman. Originally published in JMIR Aging (https://aging.jmir.org), 19.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Detection of Age-Related Cognitive Decline Over Short Time Frames Using Digital Multiday Learning Assessments: Longitudinal Study

Iris Blotenberg[1], PhD; Hairin Kim[2], PhD; Daniel Soberanes[2], MSc; Mark A Dubbelman[2], PhD; Lex R Minto[2], PhD; Onyinye J Udeogu[2], BSc; Vivienne Lubbat[2], BA; Rachel F Buckley[2,3], PhD; Gad A Marshall[2], MD; Keith A Johnson[2,4], MD; Dorene M Rentz[2], PsyD; Reisa A Sperling[2], MD; Kathryn V Papp[2], PhD; Rebecca E Amariglio[2], PhD

[1]German Center for Neurodegenerative Diseases (DZNE), Greifswald, Germany

[2]Department of Neurology, Mass General Brigham, Harvard Medical School, 60 Fenwood Road, Boston, MA, United States

[3]Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, VIC, Australia

[4]Department of Radiology, Mass General Brigham, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Rebecca E Amariglio, PhD
Department of Neurology, Mass General Brigham, Harvard Medical School, 60 Fenwood Road, Boston, MA, United States

Abstract

Background: High-resolution digital cognitive tools administered at frequent intervals enable the study of age-related cognitive change at a previously unattainable level of granularity. Computational modeling further enables the simultaneous examination of multiple aspects of cognitive performance.

Objective: The objective of this study was to examine whether repeated digital multiday learning assessments can detect age-related cognitive decline over short time frames in cognitively unimpaired older adults and to quantify age-related differences in process-based cognitive markers derived from computational modeling.

Methods: A total of 209 cognitively unimpaired older adults (mean age 73.6, SD 8.2 years; n=136, 65.1% female) from the Harvard Aging Brain Study and affiliated studies were administered the Boston Remote Assessment for Neurocognitive Health, which includes the face-name task, a paired associative learning task. Daily median response times for correct responses were used as the primary outcome. A new version of the face-name task was administered across up to 4 testing intervals over a mean of 2.8 (SD 0.4) years. Using computational modeling, we estimated longitudinal changes in peak performance, a practice-adjusted marker of an individual's best expected response time, as an indicator of age-related cognitive decline. We further examined age-related differences in peak performance, stability (within-person variability), learning within and across testing intervals, and between-interval forgetting, with embedded linear regressions for age, sex, years of education, and device type.

Results: Peak performance slowed by 0.04 seconds per year (95% credible interval 0.02-0.06; approximately 2% per year), indicating age-related cognitive decline. Higher age (+1-SD) was associated with slower peak performance (+0.19 seconds, 95% credible interval 0.12-0.26), steeper annual slowing in peak performance (+0.036 seconds per year), greater within-person variability (+0.043 seconds), reduced learning across testing intervals, and more forgetting.

Conclusions: Our findings indicate that a digital multiday learning paradigm combined with computational modeling of response times sensitively detects age-related cognitive decline over short timescales (<3 years), which are difficult to capture in conventional study designs relying on infrequent assessments. In addition, this approach identifies specific process-based cognitive markers associated with aging, including slower performance, reduced stability, and diminished learning across testing intervals.

(JMIR Aging 2026;9:e100917) doi:10.2196/100917

KEYWORDS

cognitive aging; older adults; digital cognitive assessment; remote assessment; response time; computational modeling; longitudinal study; associative memory; learning; mobile health; digital biomarkers; Alzheimer disease

Introduction

Even in the absence of Alzheimer disease or other neurodegenerative disorders, older adults experience gradual cognitive decline, particularly in fluid abilities such as processing speed and episodic memory—a phenomenon commonly referred to as normative cognitive aging [1,2]. These age-related changes manifest not only in reduced task accuracy but also in decreased efficiency and stability of underlying cognitive processes. Core markers of cognitive aging include

XSL•FO
RenderX

slower and more variable performance as well as reduced capacity to learn and retain new information [3-6]. Such process-level changes can be sensitively captured using response times, which provide fine-grained, temporally resolved measures of cognitive function and allow for the assessment of the efficiency and stability of underlying cognitive processes [7], extending information beyond accuracy-based measures.

A precise characterization of cognitive aging in cognitively unimpaired older adults is essential as it provides an important reference for distinguishing age-related change from that of pathological trajectories such as Alzheimer disease. In many cognitive aging cohorts, biomarker data are unavailable, making it difficult to determine whether observed longitudinal changes may be partly attributable to early preclinical Alzheimer disease pathology [8]. As a result, it remains unclear whether age-related cognitive decline can be reliably detected over short timescales in cognitively unimpaired older adults and to what extent distinct cognitive processes—such as speed and stability of performance, as well as learning and retention—are differentially affected by age.

Addressing this question—particularly detecting gradual cognitive change over relatively short timescales—is methodologically challenging when using traditional longitudinal study designs [9-12]. These studies typically rely on infrequent, single-administration cognitive assessments administered years apart and focus on global accuracy or mean response time scores [13]. As a consequence, short-term fluctuations due to sleep, mood, or environmental factors may obscure long-term cognitive change. At the same time, repeated testing introduces substantial practice effects, which may equal or exceed true age-related decline, especially over brief follow-up intervals of only a few years [3,9,14,15]. Consequently, observed changes in test scores reflect a composite of practice-related gains, day-to-day variability, and underlying cognitive decline, limiting interpretability and reducing sensitivity to individual differences.

Recent advances in digital cognitive assessment [16,17] provide a promising avenue for addressing these challenges. Remote cognitive testing conducted repeatedly over multiple days enables the capture of cognitive processes with high temporal resolution. In parallel, computational modeling approaches applied to intensive longitudinal response time data allow practice-related gains to be separated from cognitive change and enable the decomposition of performance into theoretically meaningful process-based markers, including speed and stability of performance as well as learning rate and forgetting [14,15]. A key advantage of this approach is that change in peak performance across testing intervals can be interpreted as a practice-adjusted estimate of within-person cognitive change over time. This approach leverages the fine-grained temporal information provided by response time data to characterize the efficiency and stability of underlying cognitive processes, complementing information obtained from accuracy-based measures [7].

In the present study, we leveraged a multiday associative face-name memory task from the Boston Remote Assessment for Neurocognitive Health (BRANCH) battery [18,19] administered daily over 7 consecutive days and repeated approximately every 6 months.

Our aims were twofold: (1) to evaluate whether this paradigm detects age-related cognitive decline over short time frames in cognitively unimpaired older adults and (2) to estimate the direction and magnitude of age associations with latent process-based cognitive markers.

Methods

Participants

Participants were recruited from 3 Mass General Brigham observational cohort studies: the Harvard Aging Brain Study (2P01AG036694-11; principal investigators: RAS and KAJ), the Instrumental Activities of Daily Living study (R01AG053184 and R01AG067021; principal investigator: GAM), and the Subjective Cognitive Decline study (R01AG058825-01A; principal investigator: REA). Participants were community dwelling and required to remain cognitively unimpaired throughout the study period. Cognitive status was assessed at annual in-clinic visits and was defined based on a Clinical Dementia Rating global score of 0, a Mini-Mental State Examination score greater than 25, and a logical memory delayed recall score above education-adjusted thresholds (>4 for 8-15 years of education; >8 for ≥16 years of education). Only participants who met these criteria at every available clinical assessment during follow-up were included. The neuropsychological assessment closest to the first BRANCH testing interval was conducted, on average, 29.7 (SD 121) days before BRANCH testing. The corresponding mean time differences were 98.8 (SD 135) days for the second interval, 202.0 (SD 149) days for the third interval, and 196.0 (SD 183) days for the fourth interval. Exclusion criteria included a primary psychiatric diagnosis, unstable medical conditions, severe head trauma, and a recent history of alcohol or substance abuse.

Materials and Procedure

Digital cognitive assessment was conducted using BRANCH, a previously validated web-based platform for remote cognitive testing that includes 2 associative memory tasks and 1 processing speed task [20]. Participants completed the tests independently and remotely on a daily basis over 7 consecutive days, enabling the characterization of learning dynamics.

The present study focused on the Face-Name Test, an associative memory task, which is illustrated in Figure 1. Participants were instructed to learn 20 face-name pairs. Following a 5-minute delay, they completed two tasks: (1) *first-letter name recall* (FNLT), in which they identified the first letter of the associated name for each face; and (2) *face-name matching* (FNMT), in which they selected the correct name from 3 multiple-choice options. The primary outcome of interest in the present study was response time for correct responses, which was recorded at the item level. First, median response times for correctly answered items were calculated separately for FNLT and FNMT. These task-specific median response times were then averaged to ensure equal weighting of both tasks. Across testing days, daily response time estimates were based on an average of 28.7

XSL•FO
RenderX

(SD 10.4) correctly answered trials across both subtasks out of 40 possible trials per testing day.

Figure 1. Illustration of the face-name matching task. Although written consent for publication was obtained, the facial image has been blurred in accordance with journal policy to protect participant privacy.



Participants completed up to 4 multiday BRANCH testing intervals. Each interval used the same face-name paradigm but a different set of face-name pairs, thereby serving as an alternate-form version of the task. Alternate forms were developed according to prespecified stimulus construction rules to maintain consistency in task structure and difficulty. In an independent pilot study, the first 3 forms were administered in counterbalanced order and showed broadly comparable accuracy and modest differences in response time, with no systematic increase in response time across forms. Further details are provided in Multimedia Appendix 1. Within each interval, the same stimulus set was administered repeatedly over 7 consecutive days to characterize multiday learning. Testing intervals were designed to include 7 consecutive testing days; participants contributed an average of 6.5 (SD 1.2) testing days per interval. The second testing interval followed an average of 1.2 (SD 0.4) years after the first testing interval, the third followed an average of 2.2 (SD 0.6) years after, and the fourth followed an average of 2.8 (SD 0.4) years after.

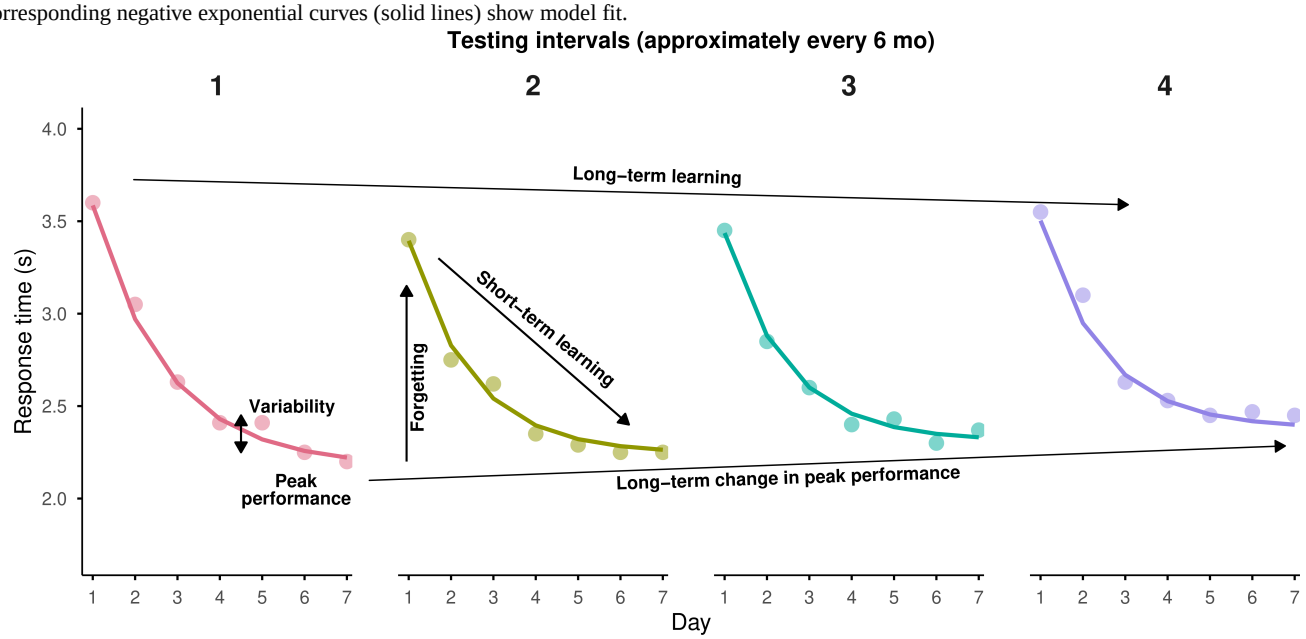
XSL•FO
RenderX

Statistical Analysis

We used a computational phenotyping approach to model daily median response times from correct trials across repeated 7-day testing intervals administered over multiple years. No additional response time trimming or exclusion procedures were applied before calculating the daily median. Specifically, we estimated a Bayesian double-negative exponential model (BDEM) proposed by Oravecz et al [14,15]. Full model specification and estimation details are provided in Multimedia Appendix 1.

Conceptually, the model assumes that observed performance reflects several processes occurring simultaneously: short-term learning across days within a testing interval, longitudinal practice-related gains across repeated testing intervals, forgetting of prior task-related practice gains between intervals, and longer-term change in peak performance. The 2 exponential components of the model capture learning on different timescales: one models rapid learning within each 7-day testing interval, whereas the other models slower practice-related gains that accumulate across repeated testing intervals and reflect increasing task familiarity. Within each testing interval, performance is modeled as approaching an individual's peak performance. By modeling learning separately, change in peak performance over time can be interpreted as a practice-adjusted indicator of within-person cognitive change. Figure 2 provides a conceptual illustration of the model and the process-based cognitive markers derived from it.

Figure 2. Illustration of the Bayesian double-negative exponential model with 4 testing intervals for 1 synthetic participant. Data are displayed as dots. Corresponding negative exponential curves (solid lines) show model fit.



Peak performance represents an individual's estimated best expected response time within a testing interval, corresponding to the asymptote approached during short-term learning. *Change in peak performance* across testing intervals reflects longer-term, practice-adjusted within-person cognitive change and is interpreted as an indicator of cognitive aging. *Short-term learning rate* quantifies how quickly individuals approach their peak performance across repeated exposure within a 7-day testing interval. In contrast, *long-term learning rate* captures the rate of task familiarization that accumulates across repeated testing intervals and unfolds over a longer timescale than learning within an interval. In addition, the model estimates within-person variability, indexing performance inconsistency around the predicted trajectory, and forgetting between testing intervals, reflecting the loss of prior task-related practice gains.

To examine age effects and further potential associations between cognitive markers and sociodemographic variables, age, sex, and years of education were included as predictors, along with a control variable for the device used (computer or smartphone). Age and years of education were *z* standardized to facilitate comparison of their associations across the process-based cognitive markers. In the present sample, 1 SD corresponded to 8.2 years of age and 2.5 years of education. These regression parameters were embedded in the BDEM in a 1-step analysis such that all reported results represented partial regression coefficients, summarizing the effect of each predictor conditional on all others.

The model used all available median response times. Missing testing days and testing intervals contributed no likelihood term. Accordingly, the primary analysis assumed that missingness was conditionally independent of unobserved response times after accounting for the observed covariates and the longitudinal performance patterns captured by the model.

All analyses were conducted in R (version 4.4.1; R Foundation for Statistical Computing) [21]. Model fitting was performed in Stan using the *rstan* interface [22]. Convergence was assessed using standard diagnostics, and model adequacy was evaluated using posterior predictive checks [23].

Sensitivity Analysis

Several additional analyses were conducted to assess the robustness of the findings and examine the time frame over which age-related cognitive change could be detected. First, to examine whether the age-related findings persisted in the absence of elevated amyloid burden, an important marker of

XSL•FO
RenderX

preclinical Alzheimer disease, analyses were repeated in a subset of participants with available amyloid positron emission tomography, restricting analyses to those classified as amyloid beta–negative (n=167). Details on positron emission tomography image acquisition and processing are provided in Multimedia Appendix 1. Second, the analyses were repeated using 3 and 2 testing intervals to evaluate the timescale at which age-related cognitive changes can be detected within the present paradigm. Third, the analyses were repeated separately for FNLT and FNMT to examine whether observed effects were driven by one specific subtask. Fourth, to examine whether the findings were sensitive to selective attrition, we repeated the analysis in participants who completed the first 3 of the 4 testing intervals using only data from these 3 intervals. This restriction balanced completeness of follow-up against retention of a sufficiently large sample. Fifth, to examine whether the number of correct responses used to calculate the daily median response times varied systematically with age, we fitted a beta-binomial mixed-effects model to account for overdispersion in the daily accuracy data. Daily accuracy was modeled using the number of correct responses out of 40 trials per testing day, with age, testing interval, study day, sex, education, and device type as fixed effects and a random intercept for participant. Age and education were *z* standardized.

Ethical Considerations

All participants provided written informed consent prior to study participation. The study protocol was approved by the Mass General Brigham Institutional Review Board (2019P000476). Data were handled in accordance with applicable data protection regulations and were stored and analyzed in deidentified form. Participants were compensated for each attempted assessment session.

Results

Sample Characteristics

Data from 209 participants who completed up to 4 testing intervals were analyzed (mean age at study entry 73.6, SD 8.2 years; n=136, 65.1% female). Sample characteristics are summarized in Table 1. Of these participants, 62.7% (n=131) completed the second testing interval, 51.7% (n=108) completed the third, and 34.9% (n=73) completed the fourth (Figure S1 in Multimedia Appendix 1). Baseline characteristics were broadly similar between participants who completed all 4 testing intervals and those who did not, although completion status differed by cohort source (Table S1 in Multimedia Appendix 1).

Table . Demographic characteristics (N=209).

Variables	Values
Age (y), mean (SD)	73.6 (8.2)
Sex, n (%)	
Female	136 (65.1)
Male	73 (34.9)
Race, n (%)	
Asian	5 (2.4)
Black	23 (11)
White	180 (86.1)
Other	1 (0.5)
Ethnicity, n (%)	
Hispanic	4 (1.9)
Non-Hispanic	205 (98.1)
Years of education, mean (SD)	16.5 (2.5)
MMSE[a], mean (SD)	28.9 (1.2)
Global CDR[b], n (%)	
0	209 (100)

[a]MMSE: Mini-Mental State Examination.

[b]CDR: Clinical Dementia Rating.

Description of Reaction Times

Correct response counts and accuracy by testing interval and study day are provided in Table S2 in Multimedia Appendix 1. To examine whether cognitive decline was observable in the raw response time data, we calculated the average response times for each participant and testing interval. Response times initially decreased substantially from the first to the second testing interval (mean 2.63, SD 0.65 seconds in the first; mean 2.47, SD 0.61 seconds in the second) before reaching a plateau in subsequent intervals (mean 2.46, SD 0.56 seconds in the third; mean 2.46, SD 0.50 seconds in the fourth). Over a mean observation period of 2.8 (SD 0.4) years, no clear age-related slowing was evident in the raw response time data.

XSL•FO
RenderX

BDEM

The observation-level R^2 was 0.87, indicating a good descriptive fit of the BDEM to the observed face-name response time data. Group-level summaries are shown in Table 2. Within the Bayesian framework, 95% credible intervals represent the range within which a parameter value lies with 95% probability, thereby reflecting the uncertainty surrounding the estimate [23]. For each cognitive marker, 2 population-level estimates are reported: the population mean across all participants and the population SD, which reflects the extent of variability in that marker.

Table . Group-level (population) estimates of Bayesian double-negative exponential model parameters.

Group-level cognitive marker	Posterior mean (SD; 95% credible interval)[a]
Peak performance (s)	2.10 (0.37; 2.05-2.15)
Change in peak performance (s/y)	0.04 (0.07; 0.02-0.06)
Forgetting between intervals (s)	1.05 (0.44; 0.99-1.10)
Short-term learning rate[b]	0.71 (0.31; 0.65-0.79)
Long-term learning rate[b]	0.64 (0.27; 0.58-0.70)
Within-person variability (s)	0.24 (0.10; 0.24-0.25)

[a]This represents the posterior mean of the group-level mean and the between-person SD.

[b]Learning rates are exponential model parameters reflecting short-term relearning within testing intervals and cumulative practice-related learning across testing intervals.

Model-estimated peak performance, derived from daily response times across the 7-day testing intervals, was 2.10 (SD 0.37) seconds on average. Peak performance slowed by 0.04 seconds per year (95% credible interval 0.02-0.06). Relative to mean peak performance, this corresponds to an average slowing of approximately 2% per year. At the beginning of a new testing interval, performance was substantially slower, as captured in the forgetting parameter: participants started, on average, 1.05 (SD 0.44) seconds slower compared to the end of the previous testing interval. Mean learning rates within (mean 0.71, SD 0.31) and across (mean 0.64, SD 0.27) testing intervals were similar. Mean within-person variability in performance was 0.24 (SD 0.10) seconds.

Results from the embedded linear regressions are shown in Table 3. Associations with age and educational attainment are reported per 1-SD increase as both variables were *z* standardized. Chronological age was associated with multiple cognitive markers. A 1-SD higher age was associated with 0.19 seconds slower peak performance. Higher age was also associated with more pronounced annual slowing in peak performance (mean 0.036 seconds per year). Older participants further showed greater within-person variability (mean 0.043 seconds), reduced benefit from prior test exposure as reflected in greater forgetting (mean 0.136 seconds), and a lower long-term learning rate (mean –0.15). Short-term learning rate was not clearly associated with age.

Additional associations were observed with sex and educational attainment. Women showed faster peak performance in the Face-Name Test than men (mean –0.177 seconds). A 1-SD increase in years of education (equivalent to 2.5 years in the present sample) was associated with faster peak performance (mean –0.052 seconds). Moreover, individuals with more years of education exhibited less pronounced slowing of peak performance over time (mean –0.017 seconds per year), indicating less cognitive aging, as well as a higher long-term learning rate (mean 0.06).

XSL•FO
RenderX

Table . Summary of results on associations between cognitive markers and person-level predictors.[a]

Cognitive marker and predictor	Results, mean (95% credible interval)
Peak performance (s)	
Age	0.19 (0.12 to 0.26)
Sex	–0.18 (–0.31 to –0.05)
Education	–0.05 (–0.11 to –0.01)
Device	0.04 (–0.07 to 0.16)
Change in peak performance (s/y)	
Age	0.04 (0.01 to 0.06)
Sex	0.00 (–0.04 to 0.05)
Education	–0.02 (–0.03 to –0.01)
Device	0.01 (–0.02 to 0.04)
Within-person variability (s)	
Age	0.04 (0.03 to 0.06)
Sex	–0.04 (–0.07 to 0.00)
Education	–0.01 (–0.02 to 0.01)
Device	0.00 (–0.04 to 0.03)
Forgetting between intervals (s)	
Age	0.14 (0.05 to 0.22)
Sex	0.10 (–0.08 to 0.29)
Education	–0.03 (–0.10 to 0.05)
Device	–0.02 (–0.18 to 0.14)
Short-term learning rate[b]	
Age	–0.08 (–0.16 to 0.00)
Sex	0.10 (–0.07 to 0.25)
Education	–0.03 (–0.10 to 0.04)
Device	0.02 (–0.13 to 0.18)
Long-term learning rate[b]	
Age	–0.15 (–0.21 to –0.09)
Sex	0.13 (–0.02 to 0.24)
Education	0.06 (0.01 to 0.11)
Device	–0.03 (–0.14 to 0.08)

[a]Age and years of education coefficients are reported per 1-SD increase. One SD corresponded to 8.2 years of age and 2.5 years of education. Reference categories were male for sex and computer (vs smartphone) for device type.

[b]Learning rates are exponential model parameters reflecting short-term relearning within testing intervals and cumulative practice-related learning across testing intervals.

Sensitivity Analyses

First, findings were replicated in the amyloid beta–negative subsample (n=167; Tables S3 and S4 in Multimedia Appendix 1), indicating that observed age-related effects are unlikely to be primarily driven by preclinical Alzheimer disease pathology. Second, analyses restricted to 2 and 3 testing intervals (Tables S5-S8 in Multimedia Appendix 1) yielded comparable group-level estimates and replicated the pattern of age and educational attainment effects observed in the full model, with aging-related slowing emerging from 3 testing intervals onward. Third, analyses conducted separately for the FNLT and FNMT (Tables S9-S12 in Multimedia Appendix 1) showed broadly similar patterns, including longitudinal slowing in peak performance and consistent associations between higher age and slower peak performance, greater within-person variability, and lower long-term learning. An age association with between-interval forgetting was observed for the FNLT but not for the FNMT. Fourth, the completer-only analysis restricted to participants who completed the first 3 testing intervals largely replicated the pattern observed in the primary analysis (Tables S13 and S14 in Multimedia Appendix 1). Peak performance

XSL•FO
RenderX

again slowed over time, and the direction and magnitude of the associations between higher age and slower peak performance, greater slowing in peak performance, greater within-person variability, more forgetting between testing intervals, and lower long-term learning were preserved. Fifth, in the beta-binomial mixed-effects model, higher age was associated with lower accuracy (odds ratio 0.72 per 1-SD increase in age, 95% CI 0.64 - 0.81; $P<.001$). Accordingly, older participants tended to contribute fewer correct trials to the calculation of the daily median response times. To examine whether the findings were robust to variation in the number of correct responses, we therefore repeated the analyses including only daily median response times based on at least 15 correct trials (Tables S15 and S16 in Multimedia Appendix 1). The resulting estimates and pattern of associations were highly similar to those observed in the primary analysis.

Discussion

Principal Results

In this study, we showed that a digital multiday learning paradigm combined with computational modeling of response times can sensitively detect age-related cognitive decline over relatively short timescales of less than 3 years in cognitively unimpaired older adults. Across repeated testing intervals, peak performance—a measure indexing cognitive performance after accounting for practice effects—declined by approximately 2% per year, with steeper declines observed at higher ages. In addition, higher age was associated with a broader process profile characterized by slower peak performance, greater within-person variability, lower long-term learning, and greater between-interval forgetting of prior task-related practice gains. These findings replicate and extend previous work addressing a long-standing methodological challenge in cognitive aging research: detecting gradual cognitive decline over brief intervals despite practice effects and day-to-day variability.

The modeling approach applied here was specifically developed to disentangle aging-related decline from other sources of within-person change but has thus far been used primarily in the context of processing speed or working memory tasks [9,14,15]. Our study extends this work in several important ways. First, we applied the model to an associative memory task, a cognitive domain particularly vulnerable to age-related decline. Second, the persistence of the age-related findings in an amyloid beta–negative subsample reduces the likelihood that they were primarily driven by preclinical Alzheimer disease pathology. All participants also remained cognitively unimpaired throughout the study period, as confirmed by repeated in-clinic neuropsychological assessments. Together, these findings strengthen the interpretation that the observed cognitive changes are related to aging rather than primarily attributable to preclinical Alzheimer disease, although contributions from other pathologies cannot be excluded. Finally, the availability of up to 4 repeated testing intervals over nearly 3 years allowed us to assess the minimum time frame over which age-related cognitive change becomes detectable. Systematic aging effects were evident with as few as 3 testing intervals, spanning a mean period of approximately 2.2 years.

A key advantage of the modeling approach is its ability to simultaneously estimate multiple latent, process-based markers of cognitive aging [15]. Chronological age was associated with a coherent pattern of changes across these markers. Older age was linked to slower peak performance, greater within-person variability, lower long-term learning, and greater forgetting of prior task-related practice gains. Notably, short-term learning within testing intervals showed no clear association with age. These findings suggest that cognitive aging in cognitively unimpaired older adults involves coordinated changes across multiple cognitive processes that would be overlooked in analyses based solely on global accuracy or mean response times.

Importantly, substantial interindividual differences in these process-based cognitive markers were not explained by chronological age alone. In particular, higher educational attainment was associated with faster peak performance, a higher long-term learning rate, and less pronounced slowing of peak performance over time, suggesting that educational attainment may be linked to greater resilience against age-related cognitive decline. These findings indicate that the applied paradigm and computational model may also be informative for studying factors that buffer age-related cognitive aging. Additionally, women showed faster peak performance than men, potentially reflecting task-specific advantages in associative face-name processing.

Comparison With Prior Work

The observed annual slowing in peak performance is consistent with prior work demonstrating gradual, small but systematic age-related cognitive slowing [3,7,24], although direct comparisons are limited due to differences in tasks, outcome measures, and study designs. In particular, our findings are consistent with the processing speed theory of cognitive aging, which posits generalized slowing as a primary mechanism of cognitive decline [2,25]. This slowing was more pronounced at older ages, aligning with trajectories of cognitive aging as described in the literature [26,27]. Greater within-person variability is consistent with prior evidence linking aging to reduced cognitive control and increased neural noise [4]. Similarly, lower long-term learning and greater forgetting of prior task-related practice gains are in line with theoretical accounts of age-related inefficiencies in encoding, consolidation, and retrieval of new information [28,29]. An interesting exception was short-term learning within testing intervals, which showed no clear association with age. This finding is notable in light of prior work suggesting that reduced multiday learning—often assessed via accuracy—is sensitive to early Alzheimer disease–related changes [19], raising the possibility that different process-based cognitive markers may show differential associations with cognitive aging vs early Alzheimer disease–related change.

The associations with educational attainment are consistent with theories of cognitive reserve and with prior evidence suggesting that life course factors shape trajectories of cognitive aging [8,30-32]. Against this background, the finding that higher educational attainment was linked not only to faster peak performance but also to less pronounced slowing over time and

XSL•FO
RenderX

better long-term learning suggests that process-based markers derived from intensive digital assessments may be useful for studying factors that buffer age-related cognitive decline. More broadly, such process-based cognitive markers are particularly relevant for prevention and intervention research as they may serve as more sensitive outcome measures in trials aimed at preserving or enhancing cognitive functioning in later life.

Limitations

Several limitations should be acknowledged. First, our sample consisted predominantly of highly educated, non-Hispanic White individuals, limiting generalizability to more diverse populations. Moreover, because BRANCH requires digital literacy and sustained adherence to repeated remote testing, its applicability to individuals with less experience using digital tools remains uncertain. Second, participants were not explicitly instructed to respond as quickly as possible; therefore, individual response strategies may have contributed to variability in response times despite item time limits and automatic progression. Third, alternate task forms were administered in a fixed sequence such that form and testing interval could not be disentangled. In the counterbalanced pilot study, the first 3 forms showed broadly comparable accuracy and modest differences in response time, with no systematic increase in response time across forms. Moreover, longitudinal slowing persisted when analyses were restricted to these first 3 forms. Nevertheless, residual alternate-form effects on response time cannot be fully separated from longitudinal change in the present longitudinal sample. Fourth, participant numbers decreased across testing intervals, raising the possibility of selective attrition. Although the main pattern of findings was preserved in the completer-only analysis, attrition associated with unobserved cognitive change or other unmeasured factors related to performance cannot be excluded. Fifth, because response times were analyzed for correct trials only and higher age was associated with lower accuracy, estimates for older participants tended to be based on fewer trials. However, older participants showed both slower response times and lower accuracy, which does not suggest a conventional speed-accuracy trade-off. Moreover, the findings were consistent across the 2 task-specific analyses and remained robust when the analysis was restricted to daily median response times based on at least 15 correct trials.

Conclusions

A digital multiday learning paradigm combined with computational modeling of response times enabled the detection of age-related cognitive decline over short time frames in cognitively unimpaired older adults. Beyond longitudinal slowing in peak performance, the approach revealed a broader pattern of aging-related differences, including slower and more variable performance, lower long-term learning, and greater forgetting of prior task-related gains. Higher educational attainment was associated with a more favorable profile across several of these markers. These findings demonstrate the potential of intensive digital assessment and computational modeling to characterize cognitive aging and support the further evaluation of process-based cognitive markers as sensitive outcomes in prevention and intervention research.

Acknowledgments

The authors would like to thank all study participants of the Harvard Aging Brain Study, Instrumental Activities of Daily Living study, and Subjective Cognitive Decline study, as well as everyone involved in the data collection.

Generative AI Use

ChatGPT (GPT-5.2; OpenAI) was used solely for language editing of author-written text. All AI-generated revisions were reviewed and approved by the authors.

Funding

The Harvard Aging Brain Study is supported by the National Institute on Aging of the National Institutes of Health (grant 2P01AG036694-11; principal investigators: RAS and KAJ). The Instrumental Activities of Daily Living study (grants R01AG053184 and R01AG067021; principal investigator: GAM) and the Subjective Cognitive Decline study (grant R01AG058825-01A1; principal investigator: REA) are also funded by the National Institute on Aging of the National Institutes of Health. The current study was supported by the Davis Alzheimer Prevention Program and the Vettel Alzheimer Innovation Fund. IB is funded by the German Research Foundation (project number: 557339238).

Data Availability

Demographic, cognitive, clinical, and neuroimaging data from the Harvard Aging Brain Study are publicly available on request through an electronic data request form available at the study website [33]. To use the Boston Remote Assessment for Neurocognitive Health in academic research, please contact the corresponding author.

Authors' Contributions

Conceptualization: IB, KVP, REA
Data curation: IB, HK, DS
Formal analysis: IB
Resources: GAM, KAJ, RAS, KVP, REA

XSL•FO
RenderX

Visualization: IB
Writing—original draft: IB, KVP, REA
Writing—review and editing: all authors

Conflicts of Interest

GAM reports research salary support for serving as site principal investigator for clinical trials funded by Eisai, Eli Lilly and Company, and Tiziana Life Sciences and consulting fees from Ono Pharma USA. KAJ reports consulting fees from Novartis, Merck, and Janssen. RAS reports consulting fees from AbbVie, AC Immune, Acumen, Alector, Apellis, Biohaven, Bristol Myers Squibb, Genentech, Ionis, Janssen, Oligomerix, Prothena Biosciences, Roche, and Vaxxinity and research funding from Eisai and Eli Lilly and Company. She also receives grant support from the GHR Foundation, the Alzheimer's Association, and the National Institute on Aging of the National Institutes of Health. KVP reports consulting fees from Biogen, Prothena Biosciences, and Cogstate. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Supplementary methods describing Face–Name Test stimulus development, Bayesian model specification and evaluation, and amyloid PET procedures, along with supplementary figures, descriptive statistics, and sensitivity analyses.
[DOCX File, 174 KB - aging_v9i1e100917_app1.docx]

References

1. Salthouse TA. What and when of cognitive aging. Curr Dir Psychol Sci 2004;13(4):140-144. [doi: 10.1111/j.0963-7214.2004.00293.x]
2. Tucker-Drob EM, Brandmaier AM, Lindenberger U. Coupled cognitive changes in adulthood: a meta-analysis. Psychol Bull 2019 Mar;145(3):273-301. [doi: 10.1037/bul0000179] [Medline: 30676035]
3. Salthouse TA. Selective review of cognitive aging. J Int Neuropsychol Soc 2010 Sep;16(5):754-760. [doi: 10.1017/S1355617710000706] [Medline: 20673381]
4. MacDonald SW, Nyberg L, Bäckman L. Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. Trends Neurosci 2006 Aug;29(8):474-480. [doi: 10.1016/j.tins.2006.06.011] [Medline: 16820224]
5. Nyberg L, Lövdén M, Riklund K, Lindenberger U, Bäckman L. Memory aging and brain maintenance. Trends Cogn Sci 2012 May;16(5):292-305. [doi: 10.1016/j.tics.2012.04.005] [Medline: 22542563]
6. Hasher L, Zacks RT. Working memory, comprehension, and aging: a review and a new view. Psychol Learn Motiv 1988;22:193-225. [doi: 10.1016/S0079-7421(08)60041-9]
7. Salthouse TA. Aging and measures of processing speed. Biol Psychol 2000 Oct;54(1-3):35-54. [doi: 10.1016/s0301-0511(00)00052-1] [Medline: 11035219]
8. Deary IJ, Corley J, Gow AJ, et al. Age-associated cognitive decline. Br Med Bull 2009;92:135-152. [doi: 10.1093/bmb/ldp033] [Medline: 19776035]
9. Sliwinski M, Hoffman L, Hofer S. Modeling retest and aging effects in a measurement burst design. In: Molenaar PC, Newell KM, editors. Individual Pathways of Change: Statistical Models for Analyzing Learning and Development: American Psychological Association; 2010. [doi: 10.1037/12140-003]
10. Salthouse TA. Influence of age on practice effects in longitudinal neurocognitive change. Neuropsychology 2010 Sep;24(5):563-572. [doi: 10.1037/a0019026] [Medline: 20804244]
11. Hoffman L, Hofer SM, Sliwinski MJ. On the confounds among retest gains and age-cohort differences in the estimation of within-person change in longitudinal studies: a simulation study. Psychol Aging 2011 Dec;26(4):778-791. [doi: 10.1037/a0023910] [Medline: 21639642]
12. Rabbitt P, Diggle P, Smith D, Holland F, McInnes L. Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. Neuropsychologia 2001;39(5):532-543. [doi: 10.1016/s0028-3932(00)00099-3] [Medline: 11254936]
13. Sliwinski MJ. Measurement ‐ burst designs for social health research. Soc Pers Psychol Compass 2008 Jan;2(1):245-261. [doi: 10.1111/j.1751-9004.2007.00043.x]
14. Oravecz Z, Harrington KD, Hakun JG, et al. Accounting for retest effects in cognitive testing with the Bayesian double exponential model via intensive measurement burst designs. Front Aging Neurosci 2022;14:897343. [doi: 10.3389/fnagi.2022.897343] [Medline: 36225891]
15. Oravecz Z, Vandekerckhove J, Hakun JG, et al. Computational phenotyping of cognitive decline with retest learning. J Gerontol B Psychol Sci Soc Sci 2025 Jun 10;80(7):gbaf030. [doi: 10.1093/geronb/gbaf030] [Medline: 39964977]
16. Öhman F, Hassenstab J, Berron D, Schöll M, Papp KV. Current advances in digital cognitive assessment for preclinical Alzheimer's disease. Alzheimers Dement (Amst) 2021;13(1):e12217. [doi: 10.1002/dad2.12217] [Medline: 34295959]
17. Polk SE, Öhman F, Hassenstab J, et al. A scoping review of remote and unsupervised digital cognitive assessments in preclinical Alzheimer's disease. NPJ Digit Med 2025 May 10;8(1):266. [doi: 10.1038/s41746-025-01583-5] [Medline: 40348870]

XSL•FO
RenderX

18. Papp KV, Samaroo A, Chou HC, et al. Unsupervised mobile cognitive testing for use in preclinical Alzheimer's disease. Alzheimers Dement (Amst) 2021;13(1):e12243. [doi: 10.1002/dad2.12243] [Medline: 34621977]
19. Papp KV, Jutten RJ, Soberanes D, et al. Early detection of amyloid-related changes in memory among cognitively unimpaired older adults with daily digital testing. Ann Neurol 2024 Mar;95(3):507-517. [doi: 10.1002/ana.26833] [Medline: 37991080]
20. Weizenbaum EL, Soberanes D, Hsieh S, et al. Capturing learning curves with the multiday Boston Remote Assessment of Neurocognitive Health (BRANCH): feasibility, reliability, and validity. Neuropsychology 2024 Feb;38(2):198-210. [doi: 10.1037/neu0000933] [Medline: 37971862]
21. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2022. URL: https://www.R-project.org/ [accessed 2026-10-09]
22. RStan: the R interface to Stan. The Stan Development Team. 2025. URL: https://mc-stan.org/rstan/articles/rstan.html [accessed 2026-10-09]
23. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis, 3rd edition: CRC Press; 2013.
24. Schaie KW. The course of adult intellectual development. Am Psychol 1994 Apr;49(4):304-313. [doi: 10.1037//0003-066x.49.4.304] [Medline: 8203802]
25. Salthouse TA. The processing-speed theory of adult age differences in cognition. Psychol Rev 1996 Jul;103(3):403-428. [doi: 10.1037/0033-295x.103.3.403] [Medline: 8759042]
26. Sliwinski MJ, Stawski RS, Hall CB, Katz M, Verghese J, Lipton R. Distinguishing preterminal and terminal cognitive decline. Eur Psychol 2006 Jan;11(3):172-181. [doi: 10.1027/1016-9040.11.3.172]
27. Wilson RS, Beckett LA, Bienias JL, Evans DA, Bennett DA. Terminal decline in cognitive function. Neurology 2003 Jun 10;60(11):1782-1787. [doi: 10.1212/01.wnl.0000068019.60901.c1] [Medline: 12796531]
28. Craik FI, Rose NS. Memory encoding and aging: a neurocognitive perspective. Neurosci Biobehav Rev 2012 Aug;36(7):1729-1739. [doi: 10.1016/j.neubiorev.2011.11.007] [Medline: 22155274]
29. Muehlroth BE, Rasch B, Werkle-Bergner M. Episodic memory consolidation during sleep in healthy aging. Sleep Med Rev 2020 Aug;52:101304. [doi: 10.1016/j.smrv.2020.101304] [Medline: 32278267]
30. Cabeza R, Albert M, Belleville S, et al. Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. Nat Rev Neurosci 2018 Nov;19(11):701-710. [doi: 10.1038/s41583-018-0068-2] [Medline: 30305711]
31. Stern Y. Cognitive reserve. Neuropsychologia 2009 Aug;47(10):2015-2028. [doi: 10.1016/j.neuropsychologia.2009.03.004] [Medline: 19467352]
32. Lindenberger U. Human cognitive aging: corriger la fortune? Science 2014 Oct 31;346(6209):572-578. [doi: 10.1126/science.1254403] [Medline: 25359964]
33. Harvard Aging Brain Study Public Dataset. Synapse. URL: https://www.synapse.org/Synapse:syn53910452/wiki/626438 [accessed 2026-10-09]

Abbreviations

BDEM: Bayesian double-negative exponential model
BRANCH: Boston Remote Assessment for Neurocognitive Health
FNLT: first-letter name recall
FNMT: face-name matching

Edited by H Köttl; submitted 10.May.2026; peer-reviewed by A Staffini, J Hakun; revised version received 30.Sep.2026; accepted 30.Sep.2026; published 27.Oct.2026.

Please cite as:
Blotenberg I, Kim H, Soberanes D, Dubbelman MA, Minto LR, Udeogu OJ, Lubbat V, Buckley RF, Marshall GA, Johnson KA, Rentz DM, Sperling RA, Papp KV, Amariglio RE
Detection of Age-Related Cognitive Decline Over Short Time Frames Using Digital Multiday Learning Assessments: Longitudinal Study
JMIR Aging 2026;9:e100917
URL: https://aging.jmir.org/2026/1/e100917
doi:10.2196/100917

© Iris Blotenberg, Hairin Kim, Daniel Soberanes, Mark A Dubbelman, Lex R Minto, Onyinye J Udeogu, Vivienne Lubbat, Rachel F Buckley, Gad A Marshall, Keith A Johnson, Dorene M Rentz, Reisa A Sperling, Kathryn V Papp, Rebecca E Amariglio. Originally published in JMIR Aging (https://aging.jmir.org), 27.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited.

XSL•FO
RenderX

The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Experiences of Ageism in mHealth App Usage Among Older Adults: Interview Study Among Older Adults Based on Extended Unified Theory of Acceptance and Use of Technology and Risks of Ageism Models

Jiayi Sun[1*], BSc; Yawen Liu[1*], BSc; Chengrui Zhang[2], BSc; Ying Xing[3], BSc; Wanqiong Zhou[2], BSc; Wei Luan[2], MD

[1]Shuguang Clinical Medical College, Shanghai University of TCM, Shanghai, China

[2]Nursing Department, Shuguang Hospital Affiliated to Shanghai University of TCM, No. 528, Zhangheng Road, Pudong New Area, Shanghai, China

[3]Shanghai Jiao Tong University School of Nursing, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Wei Luan, MD
Nursing Department, Shuguang Hospital Affiliated to Shanghai University of TCM, No. 528, Zhangheng Road, Pudong New Area, Shanghai, China

Abstract

Background: As the global aging population accelerates, mobile health (mHealth) apps have emerged as critical tools in the health management of older people. However, the promotion of mHealth apps has faced multiple obstacles, including insufficient technological adaptation to aging, digital resistance, and ageism. The impact of ageism on technology usage experiences among older adults is influenced by mechanisms such as stereotypes and biases. Notably, extant research has not adequately explored the subjective experiences of older adults in the context of mHealth app usage scenarios.

Objectives: The present study was predicated on the extended unified theory of acceptance and use of technology model and the risks of ageism model to systematically explore and understand older adults' ageism experiences in mHealth app usage. Our objectives were to provide a reference for optimizing age-friendly design and enhancing digital health management capabilities for older adults.

Methods: This qualitative study utilized an interpretive phenomenological design and was conducted between February and April 2025. Purposive sampling was employed to select older adults with experience using mHealth apps in a Shanghai community for semistructured interviews. This study used Colaizzi's phenomenological method to analyze and summarize older adults' experiences and perceptions of ageism and to extract themes.

Results: The study identified 3 core themes: (1) internalized age stereotypes, which manifest as technological uselessness and learning barriers; (2) anxiety and avoidance behaviors caused by stereotype threat; and (3) external unfair treatment (such as age-friendly design flaws and inadequate support systems), which inhibits usage. These experiences significantly impact older adults' intention to use mHealth apps.

Conclusions: Ageism profoundly affects the engagement of older adults with mHealth apps. It is advisable to execute systematic interventions to improve digital inclusion and health self-management capabilities, including strategies to challenge age stereotypes, optimize intergenerational support, refine age-friendly design, and establish strong social support networks.

(JMIR Aging 2026;9:e79457) doi:10.2196/79457

KEYWORDS

mHealth apps; ageism; technology use; older adults; healthy aging

Introduction

Background

The aging process of China's population is accelerating and has reached an advanced stage of development. Recent statistics indicate that by the conclusion of 2024, China's population aged 60 years and older will approximate 310 million, with those aged 65 years and older reaching around 220 million. These groups represent 22% and 15.6% of the total population, respectively [1]. As the aging process advances, the population of older internet users continues to expand. The 55th Statistical Report on China's Internet Development indicates that the population of internet users aged 60 years and above has

XSL•FO
RenderX

increased from 7.3 million in 2009 to 157.25 million in 2024 [2]. The internet has progressively emerged as a significant platform for older adults to obtain health information and services. Additionally, the incidence of chronic diseases in older adults has attained 75%, necessitating increasingly individualized and varied health care requirements [3]. This requires transcending conventional health care approaches to address the personalized health needs of the digital era [4]. The State Council of China has released a medium- to long-term plan (2017 - 2025) on the prevention and treatment of chronic diseases, proposing the use of information technology to promote health management [5]. Mobile health (mHealth) apps enable remote diagnostics and personalized therapies [6], particularly aiding older adults by enhancing health care access and fostering health autonomy [7]. Nevertheless, older adults exhibit low awareness and utilization rates of mobile health care [8], attributable to inadequate age-appropriate design, technological resistance, and insufficient social support [8-10]. To resolve this issue, it is necessary to improve nationwide access to digital health services. This necessitates the systematic elimination of obstacles to digital health adoption among older adults through the refinement of technical standards, optimization of service systems, and enhancement of social support.

Notably, ageism has been demonstrated to intensify the digital divide [11,12]. The World Health Organization [13] defines ageism as stereotyping, prejudice, or discrimination based on actual age [14]. This includes self-directed ageism, which refers to negative internalized beliefs about aging [15], as well as benevolent forms, such as overprotection, and hostile forms, including neglect and judgment [16,17]. Ageism impacts older adults via 3 mechanisms: internalized stereotypes, avoidance strategies, and direct discrimination experiences [18]. These mechanisms undermine the psychological well-being of older adults and exert various negative impacts on their physical and mental health [19]. Research demonstrates that ageism markedly diminishes life satisfaction among older adults, intensifies social isolation, hinders chronic disease management outcomes, and elevates the risk of depression and cognitive decline [20]. Moreover, ageism may instigate apprehension regarding operational mistakes and reduce self-efficacy by leading older adults to internalize adverse stereotypes such as “techno-phobia.” This markedly diminishes their willingness to engage with mHealth programs, thereby hindering their capacity to access health resources through mHealth apps [21]. Previous research has mostly concentrated on quantitative analyses of influencing factors [22-24]; however, it has insufficiently addressed the subjective experiences of older adults. Research on mHealth apps has identified issues such as inadequate aging [25,26], yet it seldom explores the subjective psychological effects resulting from ageism.

Objective

In order to investigate how older persons actually perceive ageism in the use of mHealth apps, this study takes an interpretive phenomenological method. It provides a reference and foundation for enhancing the age-friendly mHealth app design and creating an inclusive social support network, which will encourage older individuals’ digital integration and health management capabilities.

Theoretical Framework

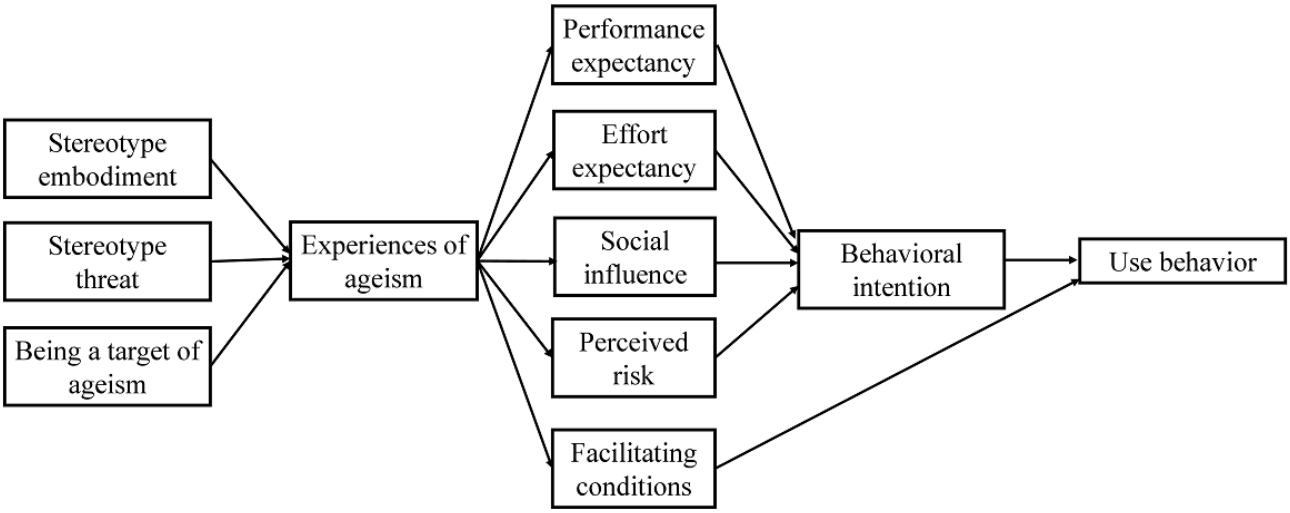
In 2003, Venkatesh et al [27] proposed the unified theory of acceptance and use of technology (UTAUT) model, which posits that user behavior in adopting technology is explained by four core dimensions: performance expectation (PE), effort expectation (EE), social influence (SI), and facilitation conditions (FC). The model demonstrated an explanatory power of 70%. It has been extensively used in research examining the adoption of new technologies among older adults [28]. This study examines privacy risks in mobile health care services by including a perceived risk (PR) dimension into the UTAUT model to create an enhanced framework. That improvement has been substantiated in the domain of medical technology [29,30].

The risks of ageism model (RAM) proposed by Swift et al explains the obstacles to positive aging through 3 pathways [18]: stereotype embodiment (negative labels imposed on older people by society), stereotype threat (self-doubt caused by older people internalizing negative labels), and being a target of ageism. The RAM model’s psychosocial mechanisms are not covered by the extended UTAUT model, although it can evaluate older adults’ desire to use mHealth apps due to its five dimensions: PEs, EEs, SI, enabling factors, and PR. The RAM model focuses on the manifestation of ageism and is not directly related to technology acceptance behavior.

The extended UTAUT model, primarily utilized in quantitative research, encompasses 5 core dimensions that systematically address essential factors influencing technology acceptance behavior, thereby offering a thorough theoretical framework for examining the digital health usage behavior of older adults. This study applies qualitative research to examine the influence of ageism on older adults’ perceptions and experiences within the UTAUT dimensions, aiming to reveal underlying mechanisms that quantitative research may not address. To establish a theoretical framework, as seen in Figure 1, this study combines the 2 models and maps the 3 routes of RAM to the 5 dimensions of UTAUT. This reveals how ageism affects the mechanism of interaction between the intention to use mHealth apps and behavior through these dimensions.

XSL•FO
RenderX

Figure 1. Extended unified theory of acceptance and use of technology and risks of ageism models.



Methods

Design

We conducted face-to-face semistructured interviews from February to April 2025. This study employed a qualitative research design based on interpretive phenomenology. This approach was chosen to deeply explore the subjective experiences and psychological sentiments of older adults concerning ageism in their usage of mHealth apps from their own viewpoints. This study used Colaizzi's method to understand individuals' real experiences and experiences [31]. We followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) reporting guideline to guarantee rigor and transparency.

Participant Selection

We employed purposive sampling to select older adults in a community in Shanghai based on gender, age, and mHealth app usage experience. The selection criteria for interviewees were as follows: (1) individuals aged 60 years or older, who have resided in the community for a minimum of 5 years (or at least 10 mo/y); (2) individuals with normal cognitive function, adequate physical strength to participate in the interview, and the ability to communicate in Mandarin; (3) regular use of mHealth apps for at least 3 months; and (4) voluntary consent obtained through the signing of an informed consent form. Exclusion criteria included the presence of mental illness or cognitive impairment, severe hearing or language impairment, and withdrawal from the study. We adopted the principle of maximum differentiation sampling, with sample size determined by information saturation (ie, no new themes appeared in the interview content).

Eligible interviewees were recruited through community hospital nurses. The study conducted face-to-face semistructured interviews from February to April 2025, with the interview outline designed based on the RAM and UTAUT theoretical frameworks. The research design emphasized ethical approval to ensure the protection of participants' privacy.

Ethical Considerations

This study was approved by the hospital ethics committee (RA-2021 - 465). All participants were informed that their participation was voluntary and anonymous and that no adverse consequences would result from the interview. All participants signed written informed consent forms after being informed of the study's purpose, procedures, risks, and benefits, and received guidance on mHealth apps usage as compensation. We refer to all interviewees with a number (N) and a letter to ensure their anonymity.

Setting

To facilitate comprehension among older adults regarding the study, information leaflets were handed out, and interviews were conducted in soundproof rooms to ensure privacy and minimize interference from unrelated individuals. The interviewer initially articulated the study's purpose and significance to the interviewee, ensuring that their personal information would remain confidential. Following that, after establishing trust through a 30-minute warm-up, the interviewer conducted a semistructured interview lasting 25 to 35 minutes, employing the "dual recording method" (audio recording along with written notes). The interviewer dynamically adjusted the questioning strategy and documented nonverbal cues, such as facial expressions and body language, using information saturation as the criterion for termination. The interviewer (first author) transcribed the recorded interviews verbatim within 24 hours and invited the interviewees to verify the transcripts. Data were encrypted and stored, accessible only by the project team within an ethical framework, in full compliance with the Declaration of Helsinki requirements.

Data Collection

After obtaining approval from the community hospital, we contacted the hospital staff to determine the visit time. Before inviting eligible residents to sign the informed consents, the researcher (first author) explained the study to them. After signing the informed consents, data were collected through semistructured face-to-face interviews and observations. The first author conducted individual semistructured face-to-face interviews from February to April 2025.

XSL•FO
RenderX

Following the identification of participants, researchers engaged with them to observe their daily utilization of mHealth apps, including the types of apps and their level of proficiency in usage. Using the extended UTAUT and RAM theoretical framework, we employed an interpretive phenomenological approach to develop semistructured interview outline. Following 2 rounds of Delphi expert consultation (n=5) and 3 discussions among the research team, along with practical feedback from 2 preliminary interviews (n=2), we developed a semistructured interview outline consisting of 6 theoretical dimensions: (1) EE: Which mHealth apps have you used? Can you give us a brief overview of your learning and usage experience? (2) PE: What is your attitude toward mHealth apps? Have they met your expectations? (3) SI: What is the attitude of your family, friends, or health care providers toward your usage of mHealth apps, and how does their attitude influence your usage? (4) FC and PR: What factors do you think influence your usage of mHealth apps? (5) In your opinion, what kind of mHealth app would be most suitable for you and your peers to use? (6) What kind of help would you like to receive when using mHealth apps? The research process strictly followed qualitative research standards. All interview questions were optimized for readability (Flesch-Kincaid index≤6.0) and cognitive adaptability testing to ensure that they were understandable to older interviewees.

Probing questions within the interview allowed participants to raise unexpected issues and provided flexibility to follow-up on these issues. By asking follow-up questions based on the answers to previous questions, interviewees were encouraged to freely share their experiences. The interview guide was used to ensure that all topics were covered.

Data Analysis

This study employed a systematic qualitative data analysis approach: initially, interview data was anonymized and assigned identification codes, while data collection and analysis occurred concurrently. Transcription was finalized within 24 hours post-interviews, with accuracy confirmed by a research team member and observational notes incorporated as supplementary data. The Colaizzi's method [31] was employed alongside NVivo 14.0 software to facilitate the analysis, resulting in a thematic map established through a 3-tier quality control process involving analysis by principal investigators, expert supervision, and review by the research team. Disputed content was addressed through consensus meetings to ensure the research's rigor and the interpretation's reliability.

Rigor

The content of the interview outline was determined through literature review and theoretical framework, and semistructured interviews were conducted with the target group in advance (not included in the study) to ensure that the interview outline was rigorous and easy to understand. The interviews were carried out by master's degree nursing students trained in qualitative research methodologies. To ensure the reliability of the collected data, we employed a combination of prolonged exposure, comprehensive analysis, diverse information sources, and various data collection methods, including interviews, field notes, and member checking by colleagues and participants. In the data analysis, efforts were made to incorporate the interviewees' emotions while minimizing the influence of the researchers' preconceived notions.

Results

Overview of Data and Analysis

The study achieved data saturation after conducting interviews with 12 participants, at which point no new codes emerged, leading to the termination of the interviews. A total of 12 older adults were included in the study. The total duration of the interviews was approximately 370 minutes, with the transcripts of the relevant themes totaling around 50,000 words. Table 1 presents the general information of the 12 interviewees. The analysis process is shown in Figure 2, which only showed the analysis process of Theme 3.

XSL•FO
RenderX

Table . General information of the interviewees (n=12).

Code	Gender	Age, y	Monthly household income[a]	Education	Living arrangement	Marital status	Health insurance	Self-reported health status
N1	Female	66	3000 - 5000 (US $420-700)	Middle school	Three-generation co-residence	Married	URBMI[b]	Excellent
N2	Female	67	3000 - 5000 (US $420-700)	Middle school	Living with spouse	Married	UEBMI	Excellent
N3	Female	77	5000 - 10,000 (US $700-1400)	High school	Living alone	Widowed	UEBMI[c]	Good
N4	Male	73	5000 - 10,000 (US $700-1400)	Middle school	Living with spouse	Married	URBMI	Fair
N5	Male	75	1000 - 3000 (US $140-420)	High school	Living with spouse	Married	URBMI	Good
N6	Female	69	5000 - 10,000 (US $700-1400)	High school	Living alone	Divorced	URBMI	Fair
N7	Female	65	5000 - 10,000 (US $700-1400)	High school	Living with spouse	Married	URBMI	Good
N8	Male	72	3000 - 5000 (US $420-700)	High school	Living with children	Married	NCMS[d]	Fair
N9	Female	74	3000 - 5000 (US $420-700)	High school	Living with spouse	Married	UEBMI	Good
N10	Male	63	1000 - 3000 (US $140-420)	Middle school	Three-generation co-residence	Married	NCMS	Fair
N11	Male	75	5000 - 10,000 (US $700-1400)	High school	Living with spouse	Married	URBMI	Fair
N12	Female	66	5000 - 10,000 (US $700-1400)	Middle school	Living with spouse	Married	UEBMI	Good

[a]A currency exchange rate of CNY 1 = US$ 0.14 (as of December 2025) is applicable for converting monthly household income.

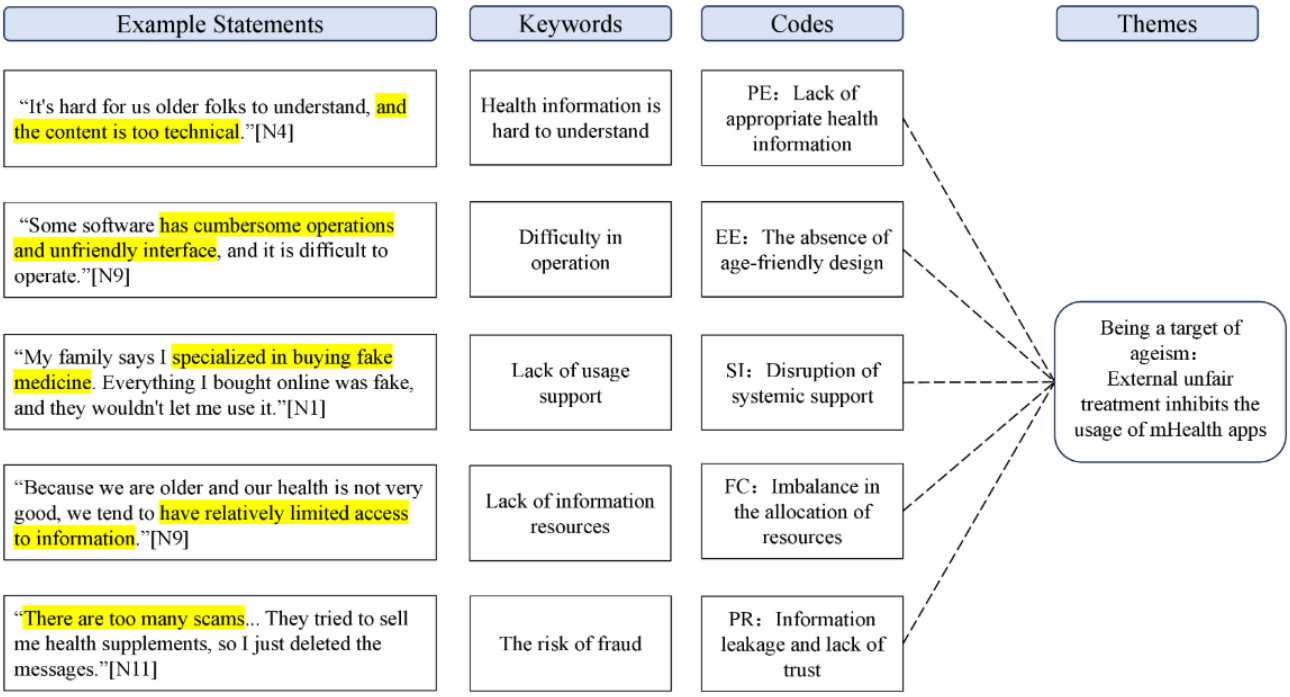
[b]URBMI: Urban Resident Basic Medical Insurance.

[c]UEBMI: Urban Employees Basic Medical Insurance.

[d]NCMS: New Cooperative Medical Scheme.

XSL•FO
RenderX

Figure 2. Process of thematic analysis (theme 3). EE: effort expectation; FC: facilitation condition; PE: performance expectation; PR: perceived risk; SI: social influence.



Theme 1: Stereotype Embodiment

PE: Age Labeling Weakens the Perceived Value of mHealth Apps

The internalization of age-related stereotypes caused participants to associate advanced age with a decrease in technological skills, resulting in a self-perception of "technological uselessness." The reduction in self-efficacy significantly impaired their acknowledgment of the potential health management advantages provided by mHealth apps. As N11 articulated:

> *I'm 75 years old. Why should I learn anything new? I'm getting old. These mHealth apps won't help much.*

It is worth noting that some interviewees actively use technology to compensate for age-related cognitive decline, reflecting compensatory strategies in individual behavior under the threat of stereotypes.

> *If there is something wrong with my body, I will look for answers.* [N9]
>
> *The doctor also talked about it (health-related knowledge), but we don't have very excellent recollections, so we can't remember anything. That's why we look for information online on our own.* [N10]

EE: Magnification Technology Usage Barriers Attributed to Age

Participants frequently ascribe operational challenges to the unavoidable cognitive and physical decline linked to aging, rather than to design deficiencies or insufficient guidance, thus exacerbating perceived obstacles to usage. N12 stated:

> *It's best if it's simple (in terms of operation and functions). We (elderly people) sometimes don't really understand (complex) programs.*

N11's attribution to educational background: "We have a relatively low level of education, so we don't really know how to use many of the functions, and it's quite difficult to learn," forming self-imposed barriers to technical learning. This cognitive pattern of attributing operational failures to age is essentially a concrete manifestation of implicit age stereotypes in the field of technology.

SI: Family Environment Reinforces Negative Expectations About Technology

External stereotypes are internalized as self-perceptions through significant individuals, particularly family members. Respondents reported that specific attitudes held by family members—such as the belief that "older adults are easily misled" (N10) and the idea that basic mobile phones are adequate for seniors (N8)—contributed to their negative self-perceptions about technological abilities. This indicates that the perpetuation of negative stereotypes within the social milieu substantially impedes the inclination to engage with technology.

> *My family told me not to believe what I read online because it's all scams aimed at old people, and they're worried that I'll get scammed.* [N10]
>
> *My kids said that basic phones for seniors are fine, so why do I need a smartphone?* [N8]

FC: Self-Imposed Constraints Hinder the Utilization of External Resources

Internalized stereotypes are evident in 2 specific behavioral patterns: intentional avoidance of acquiring new technologies and an overdependence on intergenerational support. The self-identification as "unlearnable" served as a significant obstacle, hindering the effective conversion of available external support resources into realized technical skills. As a result, a disconnect arose between the availability of facilitating

XSL•FO
RenderX

conditions and their actual behavioral application, thereby constraining the practical adoption of mHealth apps.

> *We're not like young people; we don't want to learn new functions; these functions are enough for us.* [N12]

> *I'm lazy. I won't learn how to use (mHealth apps) unless I have to. My kids will help us if I need it.* [N5]

Theme 2: Stereotype Threat

PE: Avoidance of Perceived Value of mHealth Apps

Internalized age stereotypes diminish older individuals' confidence in the health advantages of technology, thereby creating a cycle of self-fulfilling prophecy. N12 said,

> *I can't find the answers I want (in the software) … I like going to the hospital to visit a doctor more because I trust what they say more.*

This preference indicates a strategic avoidance of digital health information, motivated by a concern that unsuccessful attempts may reinforce the stereotype that older individuals are unable to effectively use technology. Meanwhile, N10's prudent disposition ("My mind works slowly, so I'm not very willing to try new things unless they're truly useful") illustrated the inhibition of perceived technological efficacy induced by stereotype threat, stemming from the allocation of working memory resources to anxiety.

EE: Psychological Attributions for Avoiding Technical Learning

Older adults attributed operational challenges to the inevitability of aging, creating a cognitive pattern of technical inertia. The assertions "I'm getting older and don't want to learn new things" in N7 and "I'm not familiar with it (mHealth apps) and am too lazy to learn; my son will assist me" in N4 suggest that age-related labels generate negative psychological expectations and significantly reduce the motivation to develop technological skills. This attribution corresponds with the behavioral pattern of stereotype threat characterized by low expectations and low investment, leading to additional impairment of executive processes, such as cognitive flexibility.

SI: Usage Avoidance Under Interpersonal Pressure

External stereotypes were internalized as self-restrictive ideas through intergenerational exchanges. N8 expressed that his children's belief that "the elderly are easily deceived" made him feel "afraid to look, lest my children get into trouble," while N10 stopped asking questions after encountering his children's "impatient" attitude, illustrating how adverse feedback within the familial context reinforced avoidance behavior. This social pressure may be viewed as a manifestation of "group identity threat," leading older adults to preserve interpersonal harmony by refraining from technological involvement.

PR: Safety Concerns Intensify the Crisis of Confidence

Technical risk issues combined with traditional age biases resulted in a dual barrier to trust. The primary concerns for older adults using mHealth apps were the security of personal information and the protection of property. N3's awareness of concealed charges ("they charge you after asking a few questions"), N2's apprehension and concern regarding potential fraud ("you will be scammed if you click on it"), and N12's doubt regarding the accuracy of information ("it's a mix of truth and falsehood, so it's better to go to the hospital") illustrated the spillover effect of stereotype threat; anxiety consistently undermined trust in technology, even when detached from specific contexts.

Theme 3: Being a Target of Ageism

PE: Lack of Appropriate Health Information

Systematically disregarding the cognitive traits of older adults resulted in ineffective health content. N3 indicated that "the software content (in the software) is too difficult to understand… too technical and hard to remember," while N4 complained that "the content is too technical." These perceptions indicate a failure in information design to accommodate age-related changes in cognitive processing, such as declines in working memory capacity, thereby reducing the perceived usefulness and accessibility of the information provided. N5 revealed deceptive advertising ("the recommended products claim to lower blood sugar, but there is no scientific basis for this"), which decreased trust in technology and highlighted the exploitation of older adults for profit.

EE: The Absence of Age-Friendly Design

There were considerable age-related deficiencies in interface interaction. N4 emphasized the lack of physiological adaptation, stating that "There is no version specifically designed for the elderly... the font should be enlarged." N8 emphasized the necessity for multimodal requirements, asserting that "the voice version should be directly audible." N9 criticized the "cumbersome operations and unfriendly interface," specifically highlighting the design issue of prioritizing "technology-centric" over "user-centric" in product development. The lack of age-appropriate design increased the learning burden for older adults and reduced the user experience.

SI: Disruption of Systemic Support

The intergenerational influence within families significantly impacted technology usage. N1 was designated as "specialized in buying fake medicine" and prohibited from using mHealth apps, while N4 encountered limitations due to his children's concerns regarding possible financial exploitation, demonstrating that protective measures resulted in technological deprivation.

The health care system offered minimal professional support. N10 noted, "The medical staff didn't mention it and told me to go to the hospital," whereas N1 stated, "Sometimes when I ask them, they get a little impatient."

FC: Imbalance in the Allocation of Resources

Public health education resources were insufficiently integrated into digital platforms. The lack of resources intensified the technological marginalization of older adults. Some older adults exhibited confidence in official information sources and a desire to gain knowledge.

XSL•FO
RenderX

There are no such platforms available on a regular basis, so I don't know what software is available or which software is suitable for us. Apart from what the nurses at the hospital recommend, I don't know anything else. [N8]

Because we are older and our health is not very good, we tend to have relatively limited access to information. [N9]

PR: Information Leakage and Lack of Trust

Older adults often mentioned challenges associated with breaches of personal information, excessive promotional communications, and misleading marketing practices from unregistered medical institutions.

I don't know how they got my personal information. Many "doctors" send me text messages. [N10]

They recommend products that claim to cure diabetes, but that's a scam… We also try to verify the authenticity of such information. [N8]

There are too many scams… They tried to sell me health supplements, so I just deleted the messages. [N11]

Discussion

Principal Findings

This study examined the ageism encountered by older adults in the community while using mHealth apps and the mechanisms through which it impacts them. The primary findings indicate that ageism obstructs the adoption and utilization of mHealth apps via 3 principal pathways of the RAM model: Stereotype internalization in older adults leads to the attribution of operational difficulties to aging, resulting in a self-perception of technological incompetence. This perception diminishes their acknowledgment of the benefits of mHealth and establishes a challenging learning threshold to surpass. Stereotype threat induces anxiety avoidance, characterized by fears of privacy breaches, worries regarding operational failures, and excessive protection or negative feedback from family members. These factors collectively result in the active avoidance of technology exploration and usage. External unfair treatment, evident in age-inappropriate design flaws in apps, including complex interfaces, limited information availability, and insufficient support systems, creates barriers to usage and diminishes the perceived value of these apps. Research indicates that ageism operates at various levels, including individual cognition, social interaction, and the technological environment, creating substantial barriers that impede older adults' integration into digital health.

Reconstructing Digital Health Cognition in Older Adults: A Dual-Pathway Intervention to Enhance mHealth App Usage Effectiveness

Older adults typically internalize the stereotype that "age dictates technological ability." This belief leads to attributing operational failures to age, avoiding the learning of new technologies, and concerns about inconveniencing others. Consequently, this significantly diminishes their willingness to use mHealth apps and undermines their sense of self-efficacy [32]. This cognitive pattern establishes obstacles in 3 dimensions—PE (diminished perceived health value), EE (reduced learning motivation), and FC (increased perceived learning costs)—illustrating the self-fulfilling prophecy of stereotype threat within the RAM model [33-35], thereby compromising well-being and impacting physical health [36]. Interventions should implement a dual approach to address this phenomenon. Positive narratives should be employed to highlight success stories among peers, such as in chronic disease management, underscoring that proficiency with technology is contingent upon practice rather than age, while minimizing the notion of "technological disadvantage." Meanwhile, an educational approach characterized by "low threshold+high feedback" should be adopted. This involves deconstructing essential functions, such as blood pressure monitoring and medication reminders, while offering voice navigation and immediate feedback to build successful experiences, improve self-efficacy, and encourage proactive health behaviors [37].

From Substitution to Empowerment: Developing a Novel Paradigm for Digital Health Support for Older Adults

Research indicates that benevolent ageism, characterized by overprotective family members and insufficient support from medical personnel, exacerbates older individuals' perceptions of incompetence [16]. The original intention may be to protect older adults; however, it implicitly stereotypes them as deficient in information, judgment, and technical skills, thereby diminishing their willingness to utilize mHealth apps [17,18]. This discrimination establishes a vicious cycle of "external rejection-self-rejection-behavioral withdrawal" through the "being a target of ageism" path in the RAM model and the "social influence" dimension in the UTAUT model. Studies indicate that the advice and trust that kids give notably affect older adults' willingness to adopt technology, while positive interactions between generations may improve perceptions of aging [38], Studies indicate that the advice and trust that kids give notably affect older adults' willingness to adopt technology, while positive interactions between generations may improve perceptions of aging [39]. Digital reverse mentoring should be promoted to enhance older adults' information literacy, while improving medical staff's negative perceptions of older adults [40]. The objective is to shift the support framework from "substitution" to "empowerment and accompaniment," motivating older adults to assume control over device operation, acknowledging their advancements promptly, eliminating their self-perception as "technologically disadvantaged," and enhancing their sense of worth as participants to digital health.

Eliminating Invisible Technological Barriers: A Dual-Path Approach to Aging-Friendly Design and Risk Prevention in mHealth Apps

Research indicates that only 40% of mHealth apps incorporate older adults in the design process [41], and challenges such as intricate interfaces, superfluous operations, and insufficient information availability highlight invisible ageism at the technology level. Analysis based on the UTAUT model: in the EE dimension, design deficiencies such as small font sizes and

XSL•FO
RenderX

poor-quality push notifications elevate learning costs [42]; in the PE dimension, an overabundance of technical terms and insufficient personalized guidance diminish perceived utility [43]; in the PR dimension, apprehensions regarding privacy breaches and misleading advertisements undermine usage intentions [44,45], all of which exacerbate older adults' self-perceptions of technological discrimination [41,46]. We advocate adopting a dual-faceted strategy: in terms of technology, implement age-appropriate features such as adjustable font sizes and voice navigation [47], create specialized and accessible health information [48], and engage older adults in the initial design stages to mitigate stereotypes [41]. In terms of risk prevention and control, establish data protection methods to eliminate hazardous information [49], while strengthening governmental regulation to create a trustworthy environment. By systematically optimizing design and management processes, obstacles to utilization by older adults can be efficiently reduced, consequently enhancing digital health inclusivity.

Multistakeholder Collaborative Empowerment: Constructing a Social Support System for Older Adults' Digital Health

Findings indicate that older adults frequently avoid engaging with mHealth apps due to a lack of informational resources and prevailing stereotypes (eg, "incapable of learning," "poor judgment") [50], which results in a diminished perception of resource support in the FC dimension. A multitiered support network is essential: medical institutions should incorporate mHealth guidance into health management services, and medical staff must acquire aging-friendly instructional skills to establish trust through recommendations from authoritative institutions. Research indicates that 55.5% of older internet users have encountered online risks [51], while also demonstrating significant vigilance [52]. Governments need to encourage platforms to enhance the digital environment, optimize recommendation algorithms, and implement antifraud training to bolster information discrimination capabilities. Through the combination of professional support from the medical system and social risk prevention and control, it is feasible to address older adults' demand for authoritative information while mitigating digital risks. This approach enhances the perceived benefits of mHealth apps and promotes healthy aging.

This study systematically elucidates the psychosocial mechanisms by which age discrimination affects older adults' adoption of mHealth apps, integrating and extending the UTAUT and RAM models. This study extends the application of the UTAUT model, confirming its efficacy and potential as a qualitative research framework for examining the subjective experiences of older adults. The integration of the RAM model with UTAUT elucidates the impact of ageism's psychosocial mechanisms on critical aspects of technology acceptance, providing fresh insights into the digital health barriers faced by older adults. These findings have practical implications for improving the delivery of digital health services in the communities and hospitals of Shanghai. Community health promotion activities must challenge stereotypes regarding older adults, enhancing their confidence through peer modeling and prompt feedback. Second, families and health care institutions ought to transition from "substitute operation" to "empowering accompaniment," thereby reducing psychological barriers through enhanced communication and initial guidance. Ultimately, the optimization of technology and services should be closely aligned with distinct user requirements, including interface complexity, information credibility, and the availability of continuous support.

Limitations

First, the sample size of this qualitative study is relatively small. Although small, the sample has reached theoretical saturation, meaning that adding more participants is unlikely to yield new insights. The current sample is drawn exclusively from a single community in Shanghai, which may introduce geographical and cultural biases that limit the generalizability of the findings. Second, despite using purposive sampling, selection bias may still be present, as participants who are more positive or more persistent in their usage of mHealth apps are more likely to participate. To mitigate this effect, we deliberately recruited participants who reported lower levels of mHealth apps usage. Additionally, the study was limited to older adults living in the community and did not include older adults in institutions such as nursing homes, which may have overlooked the experiences of ageism among more vulnerable groups. Therefore, future studies should include a broader population for comparative analysis to gain a more comprehensive understanding of the experiences of ageism in the use of mHealth apps among different groups.

In addition, this study used qualitative research methods, which are inherently subjective. To improve the scientific rigor of our research, we referenced qualitative research quality standards and focused on the following areas: researchers engaged in all interviews, recordings, and transcription processes to establish trust with participants, thereby enhancing credibility. Purposeful sampling was utilized to enhance the transferability of findings, accompanied by comprehensive descriptions of the sample characteristics. To ensure dependability, all interviews were audio-recorded, and comprehensive interview notes, transcripts, and research reflection journals were preserved. To ensure confirmability, 2 researchers performed independent data analyses, sought third-party consensus in cases of discrepancies, and provided organized data to research participants for verification. Although we used rigorous qualitative analysis to minimize bias, incorporating quantitative methods could provide more objective data, thereby improving the credibility and scientific rigor of the results. Future studies should adopt a mixed-method approach, combining qualitative and quantitative data to gain a more comprehensive understanding, provide objective indicators to enhance the reliability of the research results, and quantify the impact of ageism on the usage of mHealth apps. Finally, the duration of this study may limit our ability to track long-term changes in older adults' experiences of ageism. Future studies should include long-term follow-up to assess the temporal changes in ageism experiences and technology acceptance behavior, thereby providing more comprehensive insights and practical guidance for intervention design.

XSL•FO
RenderX

Conclusions

In this study, based on the extended UTAUT and RAM models, we identified 3 significant issues related to ageism in the usage of mHealth apps among older adults in the community: internalized stereotypes, benevolent ageism, and aging-friendly design flaws. We propose intervention strategies throughout 4 dimensions: psychological cognition, support systems, technological optimization, and social networks. This research is confined to a singular community sample located in Shanghai. Future studies should enhance the geographical and cultural diversity of the sample, implement longitudinal tracking and quantitative methods, and perform comprehensive research on the causal relationship between ageism and mHealth apps usage behavior to establish a foundation for creating an inclusive digital health environment.

Funding

This work was supported by the Open fund of key laboratory of ministry of education for long-term care of the older adults (LNYB-2023- - 12), the Shanghai Educational Science Research Project (C2023136), and the Shanghai Shenkang Hospital Development Center Project, the fifth batch of standardized management and promotion projects—Acute and Critical Illnesses Integrated Chinese and Medicine Professional Nursing Technique Training Base (SHDC22024209).

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JS (lead), WL (equal), YX (supporting)
Data curation: JS (lead), YL (equal)
Formal analysis: JS (lead), YL (equal), YX (supporting), CZ (supporting)
Funding acquisition: WL
Investigation: JS
Methodology: JS (lead), WL (equal), YX (supporting)
Project administration: WL (lead), WZ (supporting)
Resources: WL
Supervision: WL (lead), WZ (supporting)
Validation: JS (lead), WL (supporting)
Visualization: JS (lead), YL (equal)
Writing – original draft: JS (lead), YL (equal), YX (supporting), CZ (supporting)
Writing – review & editing: WL (lead), WZ (supporting)

Conflicts of Interest

None declared.

References

1. Population size decline narrows, population quality continuously improves. National Bureau of Statistics. 2025. URL: https://www.stats.gov.cn/sj/sjjd/202501/t20250117_1958337.html [accessed 2025-11-16]
2. The 55th statistical report on internet development in China [Article in Chinese]. China Internet Network Information Center (CNNIC). 2024. URL: https://www.cnnic.net.cn/NMediaFile/2025/0428/MAIN17458061595875K4FP1NEUO.pdf [accessed 2025-12-17]
3. Su B, Li D, Xie J, et al. Chronic disease in China: geographic and socioeconomic determinants among persons aged 60 and older. J Am Med Dir Assoc 2023 Feb;24(2):206-212. [doi: 10.1016/j.jamda.2022.10.002] [Medline: 36370750]
4. Spinean A, Mladin A, Carniciu S, Stănescu AMA, Serafinceanu C. Emerging methods for integrative management of chronic diseases: utilizing mHealth apps for lifestyle interventions. Nutrients 2025 Apr 29;17(9):1506. [doi: 10.3390/nu17091506] [Medline: 40362815]
5. Notice of the general office of the state council on issuing the medium- and long-term plan for the prevention and control of chronic diseases in China [Article in Chinese]. General Office of the State Council. 2017. URL: https://www.gov.cn/zhengce/content/2017-02/14/content_5167886.htm [accessed 2025-12-17]
6. Alkhaldi O, McMillan B, Maddah N, Ainsworth J. Interventions aimed at enhancing health care providers' behavior toward the prescription of mobile Health apps: systematic review. JMIR Mhealth Uhealth 2023 Feb 27;11:e43561. [doi: 10.2196/43561] [Medline: 36848202]
7. Sobrinho AS, Gomes GO, Bueno Júnior CR. Developing a multiprofessional mobile app to enhance health habits in older adults: user-centered approach. JMIR Form Res 2024 Apr 15;8:e54214. [doi: 10.2196/54214] [Medline: 38619865]
8. Askari M, Klaver NS, van Gestel TJ, van de Klundert J. Intention to use medical apps among older adults in the Netherlands: cross-sectional study. J Med Internet Res 2020 Sep 4;22(9):e18080. [doi: 10.2196/18080] [Medline: 32624465]

XSL•FO
RenderX

9. Ahmad NA, Mat Ludin AF, Shahar S, Mohd Noah SA, Mohd Tohit N. Willingness, perceived barriers and motivators in adopting mobile applications for health-related interventions among older adults: a scoping review. BMJ Open 2022 Mar 9;12(3):e054561. [doi: 10.1136/bmjopen-2021-054561] [Medline: 35264349]
10. Verloo H, Kampel T, Vidal N, Pereira F. Perceptions about technologies that help community-dwelling older adults remain at home: qualitative study. J Med Internet Res 2020 Jun 4;22(6):e17930. [doi: 10.2196/17930] [Medline: 32496197]
11. Mariano J, Marques S, Ramos MR, Gerardo F, de Vries H. Too old for computers? The longitudinal relationship between stereotype threat and computer use by older adults. Front Psychol 2020;11:568972. [doi: 10.3389/fpsyg.2020.568972] [Medline: 33123050]
12. McDonough CC, University of Massachusetts Lowell. The effect of ageism on the digital divide among older adults. GGM 2016 Jun 16;2(1):1-7 [FREE Full text] [doi: 10.24966/GGM-8662/100008]
13. Global report on ageism. World Health Organization. 2021. URL: https://www.who.int/teams/social-determinants-of-health/demographic-change-and-healthy-ageing/combatting-ageism/global-report-on-ageism [accessed 2025-12-17]
14. Officer A, Schneiders ML, Wu D, Nash P, Thiyagarajan JA, Beard JR. Valuing older people: time for a global campaign to combat ageism. Bull World Health Organ 2016 Oct 1;94(10):710-710A. [doi: 10.2471/BLT.16.184960] [Medline: 27843156]
15. Henry JD, Coundouris SP, Craik FIM, von Hippel C, Grainger SA. The cognitive tenacity of self-directed ageism. Trends Cogn Sci 2023 Aug;27(8):713-725. [doi: 10.1016/j.tics.2023.03.010] [Medline: 37147237]
16. Apriceno M, Lytle A, Monahan C, Macdonald J, Levy SR. Prioritizing health care and employment resources during COVID-19: roles of benevolent and hostile ageism. Gerontologist 2021 Jan 21;61(1):98-102. [doi: 10.1093/geront/gnaa165] [Medline: 33119089]
17. Cary LA, Chasteen AL, Remedios J. The ambivalent ageism scale: developing and validating a scale to measure benevolent and hostile ageism. Gerontologist 2016 Aug 12:gnw118. [doi: 10.1093/geront/gnw118]
18. Swift HJ, Abrams D, Lamont RA, Drury L. The risks of ageism model: how ageism and negative attitudes toward age can be a barrier to active aging. Soc Issues Policy Rev 2017 Jan;11(1):195-231 [FREE Full text] [doi: 10.1111/sipr.12031]
19. Jackson SE, Hackett RA, Steptoe A. Associations between age discrimination and health and wellbeing: cross-sectional and prospective analysis of the English longitudinal study of ageing. Lancet Public Health 2019 Apr;4(4):e200-e208. [doi: 10.1016/S2468-2667(19)30035-0] [Medline: 30954145]
20. Dix E, Van Dyck L, Adeyemo S, et al. Ageism in the mental health setting. Curr Psychiatry Rep 2024 Nov;26(11):583-590. [doi: 10.1007/s11920-024-01531-2] [Medline: 39278983]
21. Mannheim I, Köttl H. Ageism and (successful) digital engagement: a proposed theoretical model. Gerontologist 2024 Sep 1;64(9):gnae078. [doi: 10.1093/geront/gnae078] [Medline: 38874215]
22. Knuutila M, Lehti TE, Karppinen H, Kautiainen H, Strandberg TE, Pitkala KH. Associations of perceived poor societal treatment among the oldest-old. Arch Gerontol Geriatr 2021;93:104318. [doi: 10.1016/j.archger.2020.104318] [Medline: 33310658]
23. Tice-Brown D, Kelly P, Heyman JC, Phipps C, White-Ryan L, Davis HJ. Older adults' perceptions of ageism, discrimination, and racism. Soc Work Health Care 2024;63(6-7):415-432. [doi: 10.1080/00981389.2024.2365136] [Medline: 38899560]
24. Wang X, Wang Y, Zhang F, Ge D, Guo Z. Associations between social isolation, perceived ageism and subjective well-being among rural Chinese older adults: a cross-sectional study. Geriatr Nurs 2024;59:598-603. [doi: 10.1016/j.gerinurse.2024.08.014] [Medline: 39178626]
25. Kaihlanen AM, Virtanen L, Buchert U, et al. Towards digital health equity - a qualitative study of the challenges experienced by vulnerable groups in using digital health services in the COVID-19 era. BMC Health Serv Res 2022 Feb 12;22(1):188. [doi: 10.1186/s12913-022-07584-4] [Medline: 35151302]
26. Schroeder T, Seaman K, Nguyen AD, Gewald H, Georgiou A. Social determinants of mobile Health app adoption - a qualitative study of older adults' perceptions in Australia. Stud Health Technol Inform 2023 Jun 22;304:81-85. [doi: 10.3233/SHTI230376] [Medline: 37347575]
27. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
28. Maswadi K, Ghani NA, Hamid S. Factors influencing the elderly's behavioural intention to use smart home technologies in Saudi Arabia. PLoS ONE 2022;17(8):e0272525. [doi: 10.1371/journal.pone.0272525] [Medline: 36040877]
29. Wu W, Zhang B, Li S, Liu H. Exploring factors of the willingness to accept AI-assisted learning environments: an empirical investigation based on the UTAUT model and perceived risk theory. Front Psychol 2022;13:870777. [doi: 10.3389/fpsyg.2022.870777] [Medline: 35814061]
30. Cimperman M, Makovec Brenčič M, Trkman P. Analyzing older users' home telehealth services acceptance behavior-applying an extended UTAUT model. Int J Med Inform 2016 Jun;90:22-31. [doi: 10.1016/j.ijmedinf.2016.03.002] [Medline: 27103194]
31. Englander M. The phenomenological method in qualitative psychology and psychiatry. Int J Qual Stud Health Well-being 2016;11(1):30682. [doi: 10.3402/qhw.v11.30682] [Medline: 26968361]
32. Kim J, Jeon SW, Byun H, Yi E. Exploring E-Health literacy and technology-use anxiety among older adults in Korea. Healthcare (Basel) 2023 May 25;11(11):1556. [doi: 10.3390/healthcare11111556] [Medline: 37297696]

XSL•FO
RenderX

33. Okun S, Ayalon L. The paradox of subjective age: age(ing) in the self-presentation of older adults. Int Psychogeriatr 2023 Oct;35(10):566-575. [doi: 10.1017/S1041610222000667] [Medline: 35968843]
34. Ostermeier R, Rothermund K. Age differences in age stereotypes: the role of life domain and cultural context. GeroPsych (Bern) 2022 Dec;35(4):177-188. [doi: 10.1024/1662-9647/a000272]
35. Marquet M, Missotten P, Dardenne B, Adam S. Interactions between stereotype threat, subjective aging, and memory in older adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2019 Jan;26(1):121-143. [doi: 10.1080/13825585.2017.1413166] [Medline: 29221428]
36. Bergman YS, Palgi Y. Ageism, personal and others' perceptions of age awareness, and their interactive effect on subjective accelerated aging. J Appl Gerontol 2021 Dec;40(12):1876-1880. [doi: 10.1177/0733464820967209] [Medline: 33095076]
37. Wang Y, Song Y, Zhu Y, Ji H, Wang A. Association of eHealth literacy with health promotion behaviors of community-dwelling older people: the chain mediating role of self-efficacy and self-care ability. Int J Environ Res Public Health 2022 May 17;19(10):6092. [doi: 10.3390/ijerph19106092] [Medline: 35627627]
38. Almulhem JA. Factors, barriers, and recommendations related to mobile health acceptance among the elderly in Saudi Arabia: a qualitative study. Healthcare (Basel) 2023 Nov 23;11(23):3024. [doi: 10.3390/healthcare11233024] [Medline: 38063592]
39. Mikton C, de la Fuente-Núñez V, Officer A, Krug E. Ageism: a social determinant of health that has come of age. Lancet 2021 Apr 10;397(10282):1333-1334. [doi: 10.1016/S0140-6736(21)00524-9] [Medline: 33743845]
40. Crutzen C, Missotten P, Adam S, Schroyen S. Does caring lead to stigmatisation? The perception of older people among healthcare professionals and the general population: a cross-sectional study. Int J Older People Nurs 2022 Sep;17(5):e12457. [doi: 10.1111/opn.12457] [Medline: 35267232]
41. Mannheim I, Wouters EJM, Köttl H, van Boekel LC, Brankaert R, van Zaalen Y. Ageism in the discourse and practice of designing digital technology for older persons: a scoping review. Gerontologist 2023 Aug 24;63(7):1188-1200. [doi: 10.1093/geront/gnac144] [Medline: 36130318]
42. Herkert C, Graat-Verboom L, Gilsing-Fernhout J, Schols M, Kemps HMC. Home-based exercise program for patients with combined advanced chronic cardiac and pulmonary diseases: exploratory study. JMIR Form Res 2021 Nov 9;5(11):e28634. [doi: 10.2196/28634] [Medline: 34751655]
43. Kim S, Chow BC, Park S, Liu H. The usage of digital health technology among older adults in Hong Kong and the role of technology readiness and eHealth literacy: path analysis. J Med Internet Res 2023 Apr 12;25:e41915. [doi: 10.2196/41915] [Medline: 37043274]
44. Klaver NS, van de Klundert J, van den Broek R, Askari M. Relationship between perceived risks of using mHealth applications and the intention to use them among older adults in the Netherlands: cross-sectional study. JMIR Mhealth Uhealth 2021 Aug 30;9(8):e26845. [doi: 10.2196/26845] [Medline: 34459745]
45. Quan‐Haase A, Ho D. Online privacy concerns and privacy protection strategies among older adults in East York, Canada. J Assoc Inf Sci Technol 2020 Sep;71(9):1089-1102 [FREE Full text] [doi: 10.1002/asi.24364]
46. Shew A. Ableism, technoableism, and future AI. IEEE Technol Soc Mag 2020 Mar;39(1):40-85. [doi: 10.1109/MTS.2020.2967492]
47. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
48. Kraaijkamp JJM, van Dam van Isselt EF, Persoon A, Versluis A, Chavannes NH, Achterberg WP. eHealth in geriatric rehabilitation: systematic review of effectiveness, feasibility, and usability. J Med Internet Res 2021 Aug 19;23(8):e24015. [doi: 10.2196/24015] [Medline: 34420918]
49. Bonomi L, Huang Y, Ohno-Machado L. Privacy challenges and research opportunities for genomic data sharing. Nat Genet 2020 Jul;52(7):646-654. [doi: 10.1038/s41588-020-0651-0] [Medline: 32601475]
50. Choi YK, Thompson HJ, Demiris G. Internet-of-Things smart home technology to support aging-in-place: older adults' perceptions and attitudes. J Gerontol Nurs 2021 Apr;47(4):15-21. [doi: 10.3928/00989134-20210310-03] [Medline: 34038251]
51. Kemp S, Erades Pérez N. Consumer fraud against older adults in digital society: examining victimization and its impact. Int J Environ Res Public Health 2023 Apr 5;20(7):5404. [doi: 10.3390/ijerph20075404] [Medline: 37048017]
52. Nolte J, Hanoch Y, Wood S, Hengerer D. Susceptibility to COVID-19 scams: the roles of age, individual difference measures, and scam-related perceptions. Front Psychol 2021;12:789883. [doi: 10.3389/fpsyg.2021.789883] [Medline: 34975685]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
EE: effort expectation
FC: facilitation conditions
mHealth: mobile health
PE: performance expectation
PR: perceived risk

XSL•FO
RenderX

RAM: risks of ageism model
SI: social influence
UTAUT: unified theory of acceptance and use of technology

Edited by Y Luo; submitted 21.Jun.2025; peer-reviewed by A Hu, Y Sun; accepted 08.Dec.2025; published 07.Jan.2026.

Please cite as:
Sun J, Liu Y, Zhang C, Xing Y, Zhou W, Luan W
Experiences of Ageism in mHealth App Usage Among Older Adults: Interview Study Among Older Adults Based on Extended Unified Theory of Acceptance and Use of Technology and Risks of Ageism Models
JMIR Aging 2026;9:e79457
URL: https://aging.jmir.org/2026/1/e79457
doi:10.2196/79457

© Jiayi Sun, Yawen Liu, Chengrui Zhang, Ying Xing, Wanqiong Zhou, Wei Luan. Originally published in JMIR Aging (https://aging.jmir.org), 7.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Designing an mHealth App to Encourage Uptake of Muscle-Strengthening Exercise in Older Adults: Co-Design Focus Group Study

Ethan CJ Berry[1], MRes; Nicholas F Sculthorpe[1], PhD; Ashley Warner[1], PhD; Nilihan EM Sanal-Hayes[2], PhD; Lawrence D Hayes[3], PhD

[1]Division of Sport, Exercise and Health, University of the West of Scotland, Glasgow, Scotland, United Kingdom

[2]School of Health and Society, University of Salford, Manchester, England, United Kingdom

[3]Lancaster Medical School, Lancaster University, Lancaster, England, United Kingdom

Corresponding Author:
Lawrence D Hayes, PhD
Lancaster Medical School
Lancaster University
Health Innovation One, Faculty of Health and Medicine
Sir John Fisher Drive
Lancaster, England, LA1 4YW
United Kingdom
Phone: 1 44 1524 65
Email: l.hayes4@lancaster.ac.uk

Abstract

Background: Sarcopenia, the age-related decline in muscle mass and strength, poses a significant threat to functional independence in older adults. Despite strong evidence supporting resistance training as a preventive and therapeutic strategy, adherence to muscle-strengthening guidelines remains low. Mobile health (mHealth) technologies offer a promising avenue to bridge this gap; however, few apps are tailored to older adults or designed with their input.

Objective: This study aimed to identify key features that a muscle-strengthening exercise app should include to enhance engagement and uptake among older adults. Secondary aims were to explore perceived barriers and facilitators to app use and to inform the development of an evidence-based, co-designed mHealth intervention.

Methods: We used a qualitative co-design approach, involving 4 focus groups with 18 older adults (aged 60-83 years); each group comprised 3 to 6 older adults, stratified by experience with mHealth apps. Sessions were conducted online via Microsoft Teams and guided by a semistructured protocol informed by prior mHealth research and behavior change theory. Transcripts were analyzed using deductive thematic analysis, underpinned by the Technology Acceptance Model, focusing on perceived usefulness and perceived ease of use.

Results: A total of 4 overarching themes and 10 subthemes were identified. Theme 1, mHealth as a tool for supporting health and well-being, highlighted participants' recognition of digital tools in promoting activity and overcoming accessibility barriers. Theme 2, motivation and engagement through app features, revealed the importance of reminders, progress tracking, and feedback, although views on gamification were mixed. Theme 3, drawbacks of current mobile apps, captured concerns around complexity, poor usability, and lack of age-appropriate content, with skepticism regarding safety and evidence base. Theme 4, desired app elements and features, emphasized the need for customizable reminders, clear instructional videos, adaptable exercise options, and optional social features. Participants stressed the importance of simplicity, personalization, and relatable content to foster trust and sustained engagement.

Conclusions: Older adults are receptive to mHealth interventions for muscle-strengthening when design is user centered and grounded in their lived experiences. This study provides a framework for future app development, highlighting the need for intuitive interfaces, personalized features, and credible educational content. By aligning design with Technology Acceptance Model constructs and co-design principles, mHealth apps can better support healthy aging and sarcopenia prevention. These findings offer actionable guidance for developers and researchers aiming to enhance digital health equity and effectiveness in older populations.

XSL•FO
RenderX

Trial Registration: Open Science Framework 10.17605/OSF.IO/J64ER; https://osf.io/j64er/overview

(JMIR Aging 2026;9:e87332) doi:10.2196/87332

KEYWORDS

aged; co-design; mHealth; muscle-strengthening; older adults; telemedicine

Introduction

Background

Aging is a universal phenomenon, and in recent years, the global population has aged at an unprecedented rate [1]. Projections indicate that by 2050, 22% of the world's population will be aged 65 years or older [1]. Aging brings several health challenges, including increased risks of dementia and cognitive impairment, elevated blood pressure, heart disease, and diabetes mellitus [2,3]. These health issues significantly impact the ability of older adults to perform daily activities. For example, cardiovascular issues experienced in older age can lead to fatigue and shortness of breath, making everyday tasks difficult and in turn setting off a cascade of inactivity [4]. Diabetes can lead to peripheral neuropathy, vision loss, and fatigue in older adults, in turn impacting balance, fine motor tasks, driving, reading, and cooking [5]. Consequently, health and well-being have become priority areas, as reflected in the United Nations Sustainable Development Goals, particularly the goal of ensuring healthy lives and well-being for all ages [6].

One key age-related health issue is sarcopenia, which is defined as the progressive deterioration of muscle mass and strength [7,8]. This phenomenon occurs with age as type II muscle fibres and satellite cells decline, which are crucial for muscle repair [9]. Furthermore, aging causes mitochondrial dysfunction, meaning energy production in the muscle is reduced and atrophy occurs [10]. Factors such as neuromuscular decline and reduced physical activity (PA) also play a role in accelerating the onset of sarcopenia [11]. Maximum strength capacity is reached in the third decade of life, but by the time humans reach their 50s, a steady decline has already begun [12]. Reviews of skeletal muscle aging show that the trajectory and magnitude of muscle aging differ by sex, implying women may experience some aspects of functional decline earlier [13]. Indeed, it is suggested that sarcopenia has its roots in middle age [14,15]. After age 50 years, adults lose between 5% and 10% of muscle mass per decade; given muscle mass accounts for up to 60% of body mass, this loss can exert profound effects on functional independence [16]. Sarcopenia is heavily associated with several comorbidities such as diabetes, heart disease, and obesity, as well as leading to difficulties in performing activities of daily living [17].

Existing evidence points to exercise [18-22], specifically muscle-strengthening exercise, [23-25]. Kanda et al [26] demonstrated a 12-week bodyweight resistance training program undertaken twice weekly. Participants completed 2 sets of 6 exercises (targeting thighs, lower legs, buttocks, abdomen, chest, and back) for 8 repetitions (the extremities were elevated for 3 seconds and then lowered for 3 seconds), with the sets separated by 1-minute breaks. Findings demonstrated positive effects on lower limb muscle strength in a cohort of 97 participants aged 65 years or older. Muscle plasticity is well known to persist even into the ninth decade of life, as demonstrated by Wilkinson et al [27]. Despite this, guidelines for muscle-strengthening among older adults are rarely met. According to Scottish health survey data, 14% of adults aged 75 years or older complete this exercise regularly [28]. These figures are similarly reflected across the United Kingdom (17% of men and 14% of women) [29] and the United States (22% of men and 17% of women) [30].

A potential route for bridging this muscle-strengthening gap in older adults is by harnessing digital technology [31]. One particularly promising avenue is mobile health (mHealth), which incorporates mobile apps. Approximately 90% of older adults own a laptop or computer, and in the United Kingdom, approximately 70% of people aged 60 years or older own a smartphone (approximately 67% worldwide) [32]. This indicates that older adults are increasingly digitally literate [32]. mHealth can leverage features such as push notifications, daily reminders, support, and feedback [33-36]. The capability for mHealth to deliver interventions remotely is particularly beneficial for older adults, enhancing the acceptability, efficacy, and sustainability of exercise interventions within this age group [37]. The use of mHealth represents a potentially scalable, accessible delivery mechanism that can support, augment, and reinforce muscle-strengthening behaviors when appropriately designed [38]. Using mHealth as a delivery mechanism can overcome well-documented barriers to engagement in muscle-strengthening exercise among older adults, such as access to facilities, tailored guidance, and financial barriers [39]. When grounded in behavioral theory and co-designed with older adults, mHealth interventions may offer a pragmatic means of supporting self-monitoring, motivation, and adherence, which are all critical to sustained exercise engagement. Furthermore, modern cross-platform frameworks support multiple device types (eg, smartphones, tablets, personal computers, and other web-enabled platforms such as televisions). This enables access via devices that align with preferences and existing digital practices. This reduces the risk of digital exclusion, broadens reach, and helps overcome key accessibility barriers associated with device-specific mHealth apps. However, a recent scoping review concluded there was a lack of mHealth interventions focusing on muscle-strengthening for older adults, and more research is needed to refine these approaches [40]. Furthermore, there is currently no evidence-based framework to guide the design, development, and deployment of mobile technology in older adults relating to PA or exercise [41]. Without this valuable framework, mHealth interventions in older adults could encounter unintentional barriers that may impact efficacy [42]. Of more than 350,000 health-related apps on the market [43], fewer than 2000 are designed specifically for older adults [41]. Among those, few are evidence-based or involved older adults in their design and development, and given the potential of

XSL•FO
RenderX

mHealth interventions, it is prudent that further research develops a key framework to ensure future interventions are deployed effectively.

Aims and Objectives

To address these gaps, this study represents the initial phase of a broader co-design process aimed at developing an evidence-based mHealth intervention for older adults. Co-design refers to the active involvement of end users in the creation of interventions or products and is often an approach taken in mHealth development [44]. It frequently involves focus groups, interviews, and surveys [45]. The main aim of this study was to identify key features that a muscle-strengthening exercise app should contain to enhance engagement and uptake among older adults. Secondary aims were to understand limitations of existing health apps to inform the development of a more effective app for older populations that addresses these limitations and enhances engagement. This workstream was aligned with Phase 1 ("Predesign") of the co-design framework adapted from that described by Noorbergen et al [46].

Our specific research questions were:

- What do older adults perceive as positive aspects of current mobile apps?
- What do older adults perceive as barriers to current mobile apps?
- What features and considerations should a mobile app centered on exercise for older adults include?

In line with previous co-design approaches [47,48], the study conducted 4 co-design focus groups with older adults in groups of 6 to 8 participants, each lasting 40 to 80 minutes. The sessions were informed by a guided topic list that asked questions on general design preferences for an app interface as well as specific features of previous mHealth interventions and how these can be leveraged appropriately for older adults.

Methods

Overview

The study was preregistered on the Open Science Framework. Throughout the study, the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines were adhered to [49].

Co-Design Workflow

The study used a co-design methodology. Co-design refers to "the creativity of designers and people not trained in design, working collaboratively in the design and development process"[44]. It is a process in which researchers and stakeholders work together, focusing on incorporating end-user input into the design to ensure usability, relevance, and accessibility [44,50]. Older adult participants were involved in identifying key problems in current mHealth approaches, prioritizing app features, and refining concepts through structured group discussions and visual design prompts. Participants were encouraged to critique existing app features, propose alternatives, and discuss trade-offs of competing design options, with the aim of designing a suitable app for delivering muscle-strengthening exercise. Problem framing was achieved through older adults identifying barriers to mHealth engagement, which informed feature prioritization through consensus-building discussions. Visual design prompts were used during relevant discussions to allow participants to visualize specific features, garner discussion, and collectively refine concepts in real time. The workflow comprised 5 main stages.

Stage 1: Preparatory Work and Recruitment

Initial preparatory work involved identifying the evidence-based mHealth features and behavior change techniques from previous app approaches and existing literature. Older adult participants with varying levels of experience with mHealth and general apps were recruited to offer diverse perspectives.

Stage 2: Problem Framing

In the early portion of the co-design focus group discussions, participants actively shaped the problem space by discussing physical, motivational, technological, and contextual challenges relevant to aging that prevent engagement with both apps and muscle-strengthening exercise.

Stage 3: Feature Prioritization

Participants critiqued current approaches, proposed alternative solutions, and discussed the trade-offs between usability, personalization, and engagement.

Stage 4: Concept Refinement

Participants helped to refine proposed app concepts by suggesting changes to interface layout, notifications, app messaging content, notification timing, feedback mechanisms, and exercise delivery.

Stage 5: Information Gathering and Translation

Following the data collection stages indicated above, findings were deductively thematically analyzed and used to shape a concrete app prototype. Insights were mapped directly to design requirements, ensuring the prototype reflected the priorities highlighted by participants. Furthermore, recommendations were established to guide future mHealth exercise intervention development for older adults.

Study Setting and Procedures

Demographic data including, name, age, sex, occupation, education, smartphone use per day, and self-reported digital skill level were collected from participants (Table 1). A total of 2 email reminders were sent to participants prior to the focus group sessions. Four focus group sessions were completed: 2 sessions with older adults who had experience using mHealth apps and 2 focus groups with older adults without experience using mHealth apps. Co-design focus groups were chosen as the desired data collection method because they help leverage group dynamics and offer insights into shared perspectives, encouraging open discussion, which can lead to a rich qualitative dataset [51]. Each focus group session lasted approximately 40 to 80 minutes and began with an introduction to the study's purpose and the ground rules of the session. Time was then given to allow each participant to unmute themselves and give a short introduction to the group. Four focus groups were deemed sufficient because previous systematic reviews have

XSL•FO
RenderX

shown that 4 sessions typically reach saturation when addressing well-defined topics in a structured format [52].

Table 1. Participant information.

Participant	Sex	Age (years)	Occupation	Education (ISCED[a] [53])	Smartphone use per day	Self-reported digital skill level
1	Male	60	Retired	ISCED 6	30 minutes	Competent
2	Female	60	Retired	ISCED 6	30 minutes	Competent
3	Male	62	Local government officer	ISCED 7	3 hours	Good
4	Male	60	Supply chain manager	ISCED 6	30 minutes	Competent
5	Male	63	Retired	ISCED 8	2 hours 30 minutes	Good
6	Female	61	Retired	ISCED 3	2 hours	Competent
7	Female	63	Retired	ISCED 6	4 hours	Competent
8	Female	63	Retired	ISCED 6	4 hours	Competent
9	Female	64	Retired	ISCED 3	2 hours	Competent
10	Male	60	NHS[b] manager	ISCED 6	1 hour 30 minutes	Competent
11	Male	74	Taxi driver	ISCED 3	2 hours	Competent
12	Male	62	IT consultant	ISCED 7	1 hour	Very good
13	Female	83	Part time student	ISCED 7	0 hours	Poor
14	Female	82	Retired	ISCED 3	3 hours	Competent
15	Female	64	Yoga teacher	ISCED 63	3 hours	Competent
16	Male	64	Fitness instructor	ISCED 6	3 hours	Competent
17	Female	61	Retired	ISCED 6	3 hours	Good
18	Female	60	Nurse	ISCED 6	5 hours	Competent

[a]ISCED: International Standard Classification of Education.

[b]NHS: National Health Service.

There were 2 semistructured focus group protocols, 1 for the experienced groups and 1 for the inexperienced groups (Multimedia Appendices 1 and 2). The protocol itself included 7 different topics (6 for experienced participants and 7 for inexperienced participants), and each topic comprised 2 to 4 questions. The inexperienced protocol included an additional topic on privacy and data sharing to delve deeper into participants' apprehensions regarding using mHealth technology. The topics were developed based on previous mHealth intervention features [35] and consultation of the Behavior Change Technique Taxonomy version 1 (BCTTv1) [54] to gain an understanding of perceptions of common mobile app features and behavior change techniques to take forward into the development of a muscle-strengthening app. The sessions facilitated discussions using PowerPoint (Microsoft Corp) slides with images relating to app features and design, allowing participants to visualize how the app interface may look and how specific features may function within it. This resource is provided in Multimedia Appendix 3.

Research Team and Reflexivity

All focus group discussions were designed and carried out by the lead investigator of the study (ECJB). Prior consultation with the wider research team took place to ensure all members were satisfied with the focus group structure and content. Participants in each session included some individuals who were previously known to the investigator from previous research projects, and all were made aware of the broader purpose of the study. The investigator was a novice qualitative researcher with greater expertise in quantitative methods.

Participants and Recruitment

Participants were recruited using a combination of methods. Emails were sent to control group participants of previous studies conducted by the research team who had consented to be contacted for future studies. The Men's Sheds organization, the University of the Third Age (U3A), and the LEAP project were contacted to ascertain interest. These organizations fundamentally exist for older adults and work in proximity to the target population and, as such, were seen as appropriate contacts. Further recruitment was conducted through social media advertising.

To be included, participants had to meet the following criteria (1) be English speaking, (2) community dwelling (not residing in a skilled nursing facility), (3) aged 60 years or older, (4) have access to a mobile device capable of downloading apps, and (5) have access to a device capable of accessing the Microsoft Teams software and understand how to use the platform. Participants were asked prior to recruitment whether they were experienced or inexperienced with mHealth or exercise apps.

XSL•FO
RenderX

We aimed to recruit approximately 20 older adults, as this number allowed us to have 6 to 10 participants in each focus group, which facilitates good discussion in which all participants can contribute. The final number of participants included was 18; a total of 2 participants withdrew because the study timeline did not align with their personal schedules. Of the 18 participants included, 9 had experience with mHealth or exercise apps and 9 had no experience. Inexperienced participants were included to ensure a highly inclusive and informed design process, with the view to increasing scalability.

Data Collection

Each focus group session lasted approximately 40 to 80 minutes. Initially, the ground rules of the session were outlined. Each participant was asked to provide verbal consent for the session to be recorded. Each session was recorded, and raw transcripts were produced. The lead investigator checked transcripts for accuracy, grammar, and spelling. A clean verbatim approach to transcription was selected to balance accuracy with clarity by capturing the essential meaning of participants' words while omitting unnecessary filler words, false starts, and stutters. This approach has been previously justified within focus group literature as an ideal method to ensure richness of the discussion is captured while removing excerpts that may make the transcript difficult to read [55]. Finalized transcripts were distributed to participants for member checking. All participants agreed that the transcript was an accurate depiction.

Data Analysis

All data collection and analysis were conducted by the lead investigator (ECJB) with guidance from the wider research team. Once transcribed and validated, deductive thematic analysis was undertaken to generate themes and gain an understanding of the data [56]. As part of the analysis the Technology Acceptance Model (TAM) [57] was used as the theoretical framework to guide coding and development of themes. The TAM is a widely used framework that explains how users adopt and use technology. It suggests that 2 key factors influence technology acceptance: perceived usefulness, defined as the extent to which a person believes a piece of technology will benefit them; and perceived ease of use, defined as the degree to which a person believes that using the system will be free from effort [57].

Thematic analysis followed Braun and Clarke's [56,58] 6-phase framework. To operationalize the TAM within the thematic analysis approach, participants' language was first coded inductively at the semantic level to ensure the codes remained grounded in their experiences and meaning. A deductive, theory-driven coding approach was then used, guided by the TAM and the study's research questions, by mapping the codes onto the TAM constructs. Semantic coding focused on surface-level meanings to remain close to participants' language and experiences. Codes reflecting perceived benefits, value, or relevance to participants' health were grouped under the first main construct, perceived usefulness, while codes relating to effort, complexity, navigation, and usability were grouped under the second main construct, perceived ease of use. This allowed for the development of key themes grounded in the TAM. This 2-step process ensured participants' language informed the analysis prior to interpretation and meant the TAM functioned as an organizing framework. This approach was appropriate given the aim to identify app features, barriers, and facilitators relevant to older adults. The analysis revealed positive aspects of current apps, limitations affecting mHealth uptake, and key design considerations to enhance usability. These insights directly informed the app refinement phase, ensuring alignment with older adults' preferences and needs. Finally, the report was written to present a coherent narrative linking the themes to TAM and the study aims.

Ethical Considerations

Upon expression of interest, participant information documents were shared. Participants were given the opportunity to ask any additional questions for clarity before agreeing to participate. Electronic consent forms were completed prior to the focus groups. Participants were then given details of the focus group (date and time) and emailed a meeting link along with telephone instructions on how to access the Microsoft Teams session. During the session, participants were allowed technical support from a family member or friend.

The focus groups were completed on Microsoft Teams for participant convenience; this also allowed a broader geographical sample. Participants joined the Teams session using the meeting link that had been emailed to them previously and using their own smartphone, tablet, laptop, or computer device. Each participant was asked to provide verbal consent for the session to be recorded for subsequent thematic analysis. Participants were not provided with any financial or other form of compensation for their participation. All data were treated confidentially and anonymized using participant identification numbers. Focus groups were video recorded online (with all participants' permission) to facilitate subsequent thematic analysis. Recordings and associated data were stored on password-protected computers available only to the principal investigator (ECJB).

This study was carried out in accordance with the Declaration of Helsinki and approved by the University of the West of Scotland Health and Life Sciences Academic Integrity and Ethics Committee (approval number 22512).

Results

Overview

A total of 4 main themes were identified from the analysis, with several subthemes (Table 2). These themes encompass the previous lived experience of participants with mHealth technology, areas in which participants have struggled with engagement, and key considerations to take forward to enhance usability and intervention effectiveness. Full excerpts from focus group discussions are included in Multimedia Appendix 4.

XSL•FO
RenderX

Table 2. Four main themes and the corresponding subthemes derived from thematic analysis.

Theme	Subtheme
Theme 1: mHealth[a] as a tool for supporting health and well-being	• Overcoming barriers to accessibility • Potential for educational support via mHealth
Theme 2: motivation and engagement through app features	• Progress tracking as a source of motivation • Rewards and streaks as reinforcement
Theme 3: drawbacks of current mobile apps	• Complexity and poor usability • Skepticism over effectiveness, safety, and evidence base
Theme 4: desired app elements and features	• Customizable reminders • Clear video demonstrations of muscle-strengthening exercise • Customizable exercise options and feedback • Preferences regarding social features of mHealth

[a]mHealth: mobile health.

Theme 1: mHealth as a Tool for Supporting Health and Well-Being

This theme encapsulates participants perceptions of mHealth to support health and well-being in their daily lives. Across both experienced and inexperienced users, mHealth technologies were understood to be mechanisms that help promote health behavior through increasing awareness, providing feedback, and prompting action.

Participants who were experienced with mobile technology drew upon their experiences of mHealth promoting activity via push notifications and how advanced features support them during exercise. Experienced participants highlighted that feedback mechanisms and prompts help to break long bouts of sedentary behavior and regulate activity, as one participant noted:

> *A part of my app. It tells you if you've been sat for a long time and it just basically says maybe it's time to get up and move.* [Participant 6]

These features were perceived as highly important in the context of aging, where participants noticed a steady decline in their own activity and had a desire to maintain functional independence.

Instant feedback features such as heart rate and activity summaries were described as reassuring, motivating, and a useful mechanism for tracking exercise effort and intensity with greater confidence. The theme highlights a broader understanding, primarily among experienced participants, of their own deterioration during the aging process and that mHealth interventions may encourage them to participate in exercise, which would attenuate their decline.

This first theme led to the development of the subtheme, overcoming barriers to accessibility. The subtheme broadens understanding of the challenges older adult participants have faced in relation to muscle-strengthening exercise, namely physical and financial limitations. Participants identified digital resources as a means of overcoming these barriers. One participant noted:

> *I used to go to YouTube. I've discovered fabulous 50s. For people who are over 50 and then it's just as a range of different exercises, you know, usually they're just walking included in it, and it means you can do it in the house or whatever.* [Participant 8]

Furthermore, an experienced participant reflected that an app aid accessibility by stating:

> *Rather than expecting people to have, you know, weight sets or whatever get, you know, used items that people have readily accessible round their houses.* [Participant 1]

A further subtheme identified in relation to Theme 1 was the potential for educational support via mHealth. Participants consistently emphasized the potential for mHealth technologies to provide educational support related to exercise benefits, technique, and progression. Experienced participants drew upon the self-quantifying nature of digital exercise tools, acknowledging the empowerment of personal metrics, which encouraged continued engagement. Participants highlighted the value of clear guidance to support learning at their own pace:

> *I've not got the sort of strength that I probably want to have and just want to sort of educate myself into what the next steps could be just to sort of bring myself on.* [Participant 4]

This subtheme highlights the understanding that this technology has the potential to educate older adults on muscle-strengthening exercise.

Theme 2: Motivation and Engagement Through App Features

This theme captures how mHealth features were perceived to influence motivation, routine development, and sustained engagement with exercise behavior. Participants described how features such as prompts, feedback, and reinforcement mechanisms supported adherence by fostering accountability and a sense of progress.

Participants drew upon familiar examples from fitness apps that use nudge reminders to increase engagement. Similarly, experienced participants highlighted the benefit of wearable and app technologies that use haptic nudges to prompt users to

XSL•FO
RenderX

break up sedentary time and send motivational messages to urge users to meet step goals. These features were perceived as supportive, with one participant stating:

> *That's what I like about it. I just realised I got a reminder today that I haven't done my third run of the week, so that's good.* [Participant 3]

Personalized feedback on goal progression was also viewed as central to maintaining engagement.

A subtheme generated from Theme 2 was progress tracking as a source of motivation. This subtheme identified motivators and features that engage participants, such as the ability to track their progress. Participants highlighted that tracking their own progress when working toward an activity goal and recording improvements motivated and encouraged them to continue. Inexperienced participants highlighted that progress tracking could provide reassurance and validation.

A final subtheme developed as part of Theme 2 was rewards and streaks as reinforcement. This subtheme encapsulates experienced participants' thoughts, feelings, and encounters with reward systems (ie, gamification systems). Participants expressed mixed views, highlighting a clear trade-off between motivation and engagement. For some experienced participants, rewards such as badges, medals, and streaks were described as powerful motivators for sustained engagement:

> *I got fanatical about continuing my streak.* [Participant 12]

In contrast, those who had little experience with these features viewed the prospect of gamification as childish and condescending. These participants felt that simplistic rewards such as badges undermined autonomy and preferred personalized meaningful feedback linked to functional improvement. This highlights the challenge of designing engagement features that accommodate varying levels of digital familiarity and differing motivational preferences.

Theme 3: Drawbacks of Current Mobile Apps

The third theme reinforced the key issues that older adults face when engaging with mobile apps. Across both experienced and inexperienced users, drawbacks were not described as isolated issues but as indicators that mobile apps in their current form are inadequately designed for older adults' needs and priorities. Participants highlighted a general reluctance to input personal information, reflecting broader concerns around privacy and emotional burden of mHealth. Some participants described negative experiences with continuous monitoring devices, such as sleep tracking, which were perceived to exacerbate anxiety rather support well-being:

> *I've got a problem with sleep trackers from work perspective. They induce more anxiety than they help to be perfectly honest.* [Participant 10]

These sentiments highlight broader skepticism among participants about whether app engagement would meaningfully enhance health. Mainly inexperienced participants also felt that existing apps lacked relevance for older adults. Current apps were described as cluttered and filled with unnecessary features, creating confusion and discouraging sustained use. Several participants noted that while health and fitness apps appear abundant, those targeting muscle strength while also catering to age-related needs are scarce, further highlighting the accessibility gap between older adults and app design.

A subtheme that subsequently developed was complexity and poor usability. This subtheme explored specific difficulties participants encountered when engaging with mobile apps. Experienced participants described current fitness apps as being overcomplicated and overloaded with information, which they found hard to understand. Inexperienced participants felt that even when they have attempted to use currently available apps, they found them time-consuming and burdensome. Older adults indicated that simplicity was key and expressed a preference to engage only with activities and information that were personally relevant. Interestingly, even those who self-identified as digitally competent felt that current exercise apps were complex and difficult to use effectively without becoming overloaded with information.

A further subtheme that developed was skepticism regarding effectiveness, safety, and evidence base. Participants expressed widespread skepticism regarding the effectiveness and safety of current exercise apps, particularly for older adults with musculoskeletal and chronic conditions, such as sarcopenia. Despite recognizing the importance of strength training, sometimes informed by professional health care practice (eg, nursing), some participants questioned whether current apps were evidence-based or capable of safely supporting strength improvements. Concerns were also raised about the credibility of online exercise content. As one participant noted:

> *Now I look at a lot of videos and I think these people are promoting that and it's dangerous things they are doing, you know, when they're promoting it for older folk.* [Participant 17]

These concerns contributed to hesitation around adoption and highlighted the importance of clinical credibility and age-appropriate design in future mHealth interventions.

Theme 4: Desired App Elements and Features

The final theme identified was desired app elements and features. This theme was developed in conjunction with the positive and negative aspects of current mobile apps. It synthesizes participants' perspectives on how mHealth interventions should be designed to support sustained engagement and safety among older adults. Rather than describing isolated feature requests, participants articulated broader design priorities centered on simplicity, autonomy, and relevance to older adults' needs.

Participants consistently emphasized the importance of simple interface design. Features such as clear navigation, high-contrast colors, and adjustable font sizes were described as essential for reducing cognitive load. Participants described positive experiences with apps and websites that prioritized clarity over complexity, reinforcing the need to accommodate age-related sensory decline. One participant noted:

> *Some of them are better than others, you know, and most of them are kind of intuitively simplistic. So*

XSL•FO
RenderX

probably like those, not too busy and easy to follow. Is the kind of rule of thumb. [Participant 10]

A subtheme that developed in conjunction with Theme 4 was customizable reminders. Personalization emerged as a central mechanism for supporting sustained engagement. Participants described a strong preference for control over reminder frequency and timing, highlighting the importance of aligning prompts with individual routines. The timing of notifications was encapsulated within this subtheme, with many experienced and inexperienced participants stating that a reminder in the morning to engage in exercise would suit their routine. One participant noted that it was important for this reminder to come at the start of a new week. However, counterpoints were raised, with 1 participant stating that they often find their life is busy in the morning and, as such, a reminder later in the afternoon would be beneficial because they would have more time to engage in exercise at this time of day. This subtheme highlights autonomy as a key design principle, with flexibility enabling mHealth interventions to accommodate diverse lifestyles.

A further subtheme generated was clear video demonstrations of muscle-strengthening exercise. When discussing the specific features related to enhancing usability, both experienced and inexperienced participants emphasized a need for clear video demonstrations to improve safety and effectiveness. Some participants highlighted the importance of being able to take the learning process at their own pace and felt it would be useful to have the option of pausing, fast-forwarding, and rewinding a video demonstration to ensure they could adequately perform the exercise. Furthermore, participants who had experience using similar apps for gym-based exercise recalled that being able to click on certain exercises with which they were unfamiliar and access video demonstrations helped them build confidence with the exercise. Additionally, all participants emphasized the ability to resonate with the video demonstrations, suggesting the demonstrator in current available apps is often much younger. The consensus was that if the demonstrator within the videos were an older adult, this would add a sense of realism with which users could relate.

A further subtheme developed in relation to Theme 4 was customizable exercise options and feedback. Autonomy over exercise selection and progression was viewed as essential for maintaining engagement. Participants built on this idea by highlighting that users have a range of preferences for muscle strengthening–based exercise, and allowing them to select the activity may enhance their experience. Discussion developed further to generate the idea of customization to allow users of different abilities to be signposted to exercises appropriate for them. Participants described frustration with one-size-fits-all approaches commonly experienced by older adults in gym and digital environments. A staged structure, such as tiered programs guided by front-loaded questioning, was identified to personalize exercise without requiring users to label or compare themselves to others. Participants noted that inappropriate labeling may negatively impact confidence if their own progression stalled. However, participants appreciated that their app experience would be personal to them and that their individual needs would be catered to.

A final subtheme developed in relation to Theme 4 was preferences regarding social features of mHealth. Discussion surrounded the notion of including social features within the app to allow participants to complete exercise together with family or friends. A few experienced participants underlined that they felt the option to exercise together with peers would enhance their motivation, as they often find it hard to exercise alone. Others felt social features were unnecessary, as they become competitive and distract from the initial goal of encouraging exercise. The general feeling was that having the option to engage in social aspects would be beneficial but should be secondary to encouraging the uptake of strengthening exercise. Insights gathered particularly in this theme directly informed app prototype development, ensuring older adults' identified needs were translated into design features.

Discussion

Overview

The aim of the focus groups was to determine essential features that a muscle-strengthening exercise app should include to improve engagement among older adults. This study examined the limitations of current mHealth apps that hinder effectiveness and contribute to low adoption. During deductive thematic analysis, the research questions were considered alongside the TAM framework, which helped to produce 4 interrelated themes and 10 subthemes. Participants' insights highlighted their nuanced experience of both mHealth technology and muscle-strengthening exercise. The development of these main themes helps to build on broader practical implications and recommendations for the future of muscle-strengthening app development for older adults, contributing to healthy aging.

This manuscript represents a significant contribution to the field of digital health and aging through its originality, significance, and methodological rigor. It introduces a transformative approach by applying a co-design methodology to the development of a muscle-strengthening mHealth intervention for older adults, a population historically underserved in app design. We believe this is the first study to commence co-design of a muscle-strengthening mHealth intervention for older adults. The integration of the TAM into thematic analysis offers a novel theoretical lens for interpreting older adults' engagement with digital exercise tools, advancing conceptual understanding from gerontology to digital health and beyond.

This study's significance lies in its potential to reshape how mHealth interventions are designed and deployed for older populations. By directly involving older adults with varying levels of digital literacy and app experience, the research generates actionable insights that challenge assumptions about usability, motivation, and accessibility. The findings have immediate practical implications for developers, clinicians, and policymakers seeking to address sarcopenia and promote healthy aging through scalable, evidence-based digital solutions. Methodologically, the study demonstrates exceptional rigor. The sampling strategy ensures representation across digital experience levels, and saturation is robustly evidenced. The alignment of findings with TAM constructs provides a coherent

XSL•FO
RenderX

analytical framework, and the study is transparently reported in accordance with the COREQ guidelines.

Theme 1: mHealth as a Tool for Supporting Health and Well-Being

The first theme conveyed participants' views on how mHealth technology can support daily health and well-being. Although direct experience varied across focus groups, participants generally appreciated its potential to encourage healthier behaviors and improve access to PA, particularly muscle-strengthening exercise.

Those with experience described how mHealth guided healthier behavior and increased awareness of health metrics. These findings align with nudge theory, a concept in behavioral economics suggesting that positive reinforcement and indirect suggestions can influence behavior [59]. Nudge theory has been applied in studies using mHealth to promote PA in older adults [60-62]. For example, Yamada et al [61] showed that tablet-based nudges encouraged healthier behaviors in older men. Other research has similarly demonstrated that mHealth features can be leveraged to support health and well-being in older adults. Kwan et al [63] trialed a brisk walking intervention with 33 participants using the Samsung Health app and WhatsApp, delivering programs, reminders, and rewards. The intervention enhanced moderate-to-vigorous PA uptake by 86 minutes per week post intervention, supporting consensus in the literature that mHealth promotes health behavior.

Older adults often face accessibility barriers to exercise [64], a theme echoed by participants. Research has identified extrinsic barriers such as lack of transport, supervision, or facilities, and intrinsic barriers including perceptions of being too old, fatigue, or fear of injury [65]. Numerous studies have shown that digital delivery can help reduce such barriers [66-68]. Sciammana et al [69] delivered a 24-week, 1-minute bodyweight intervention via weekly emails to 24 older adults with a mean age of 71.6 (SD 8.9) years. The findings outlined that 75% of participants completing at least half of the prescribed sessions. Extending such approaches to mHealth, with an evidence-based approach, may allow for enhanced intervention delivery methods and further increase accessibility and adherence in muscle-strengthening contexts.

A further consideration from participants was the suggestion of incorporating household items, such as filled bottles or shopping bags, into mHealth exercise interventions to overcome limited gym access. While this could provide a training stimulus [70], safety concerns arise, as equipment use increases injury risk, especially given that most older adults lack resistance training experience [71,72]. Thus, equipment-free exercises may be more appropriate for mHealth resistance programs in older adult contexts.

Participants also valued the educational potential of mHealth. They emphasized the need for simple apps providing essential information to facilitate muscle-strengthening without unnecessary features. This reflects evidence on the importance of informational support in behavior change interventions for older adults [73,74]. Concerned about age-related muscle loss, participants welcomed resistance training and mHealth initiatives but highlighted a knowledge gap that left them ill-equipped to train independently. Prior research has also identified poor understanding of resistance training's purpose and safe practice among older adults [66]. Despite guidelines recommending twice-weekly strength training [75], messaging is often vague [76]. mHealth could bridge this gap by offering educational support, such as video demonstrations, to enable safe and effective training [40].

Overall, this theme underscores mHealth as a promising avenue for promoting muscle-strengthening behavior change. Participant insights link closely to the construct of perceived usefulness, which remains a strong predictor of mHealth engagement, particularly when technology supports daily life and independence [77].

Theme 2: Motivation and Engagement Through App Features

A crucial determinant of engagement with mHealth among older adults was the presence of motivational app features. Participants consistently discussed elements such as prompts, reminders, and progress feedback, explaining how these supported the integration of health behaviors into daily routines. These features were perceived to provide structure for engaging in PA and other health-related behaviors. Specifically, prompts and reminders acted as "nudges" that encouraged consistent PA engagement and supported long-term behavior change, aligning with broader evidence identifying these functions as key to sustained mHealth use [78,79].

Participants highlighted reminder functions, typically delivered via push notifications, as behavioral cues promoting consistent app use and engagement in health behaviors. These findings correspond with literature demonstrating the effectiveness of prompts for PA promotion, particularly among older adults who benefit from external structure [80,81]. Reminders and reinforcement messages can enhance exercise adherence when they become integrated into users' routines [82,83].

Progress tracking emerged as another major motivator. Both experienced and inexperienced users described viewing their progress over time as encouraging and essential for maintaining motivation. This reflects prior work emphasizing the value of feedback and progress monitoring in behavioral interventions [84]. For instance, Geraedts et al [85] combined wearable sensors and tablet-based strength exercises, enabling participants to view improvements, which enhanced motivation. Such findings relate to the concept of self-quantification, which refers to the practice of methodically collecting, measuring, and analyzing personal data using technology [86], wherein monitoring one's data fosters engagement and motivation [87].

While mHealth interventions targeting aerobic activity benefit from automated data collection via wearables [63,88,89], muscle-strengthening interventions face greater challenges. Accurately capturing resistance or repetitions often requires manual input, increasing user burden. Nonetheless, the motivational benefits of visualizing progress may justify this effort if it enhances adherence [90].

Views on rewards and streaks were mixed. Participants familiar with such features viewed them positively, while inexperienced

XSL•FO
RenderX

users often perceived them as "condescending." Although these negative attitudes may stem from unfamiliarity [91], the literature supports the effectiveness of gamified features, particularly in step-based interventions [92-94]. It may therefore be beneficial to personalize gamification for older adults. Although studies have shown the general benefits of gamification [95-97], these features must be context-sensitive and tailored to older users. Moreover, step-based gamification relies on automated tracking, whereas applying similar approaches to strength training poses technical and usability challenges. Streak-based systems may also risk shifting focus away from meaningful engagement by becoming the main focus of the user rather than the exercise intervention at hand [98].

These findings connect closely with the TAM construct of perceived usefulness, a key predictor of intention to use. The theme encapsulates features that translate app functions into tangible benefits by motivating users to exercise or track health behaviors. When older adults perceive clear value in how an app supports their activity or progress, they are more likely to adopt and sustain its use [99]. Overall, this study highlights the importance of motivational features while urging careful, context-aware integration of gamification in mHealth design for older adults.

Theme 3: Drawbacks of Current Mobile Apps

The third theme highlighted several disadvantages older adults face that hinder uptake and sustained engagement with mHealth technology. Data from this study identified drawbacks such as poor usability, irrelevant content, and doubts regarding safety and effectiveness—issues echoed in the broader literature as key barriers to mHealth adoption among older adults [100]. Participants, particularly those inexperienced with apps, felt current designs often neglect older adults' needs. A perceived lack of personal relevance reduced engagement, consistent with research on the poor alignment between mobile app design and older users [101]. Participants also noted limited accessible options for home-based muscle-strengthening, with most apps failing to account for reduced physical capacity [102]. Moreover, distrust in tracking technologies emerged as a barrier, consistent with existing findings on digital health skepticism in this demographic [103].

Complex app interfaces were another major source of frustration. Even digitally confident participants described current exercise apps as "too advanced" or "overloaded with information." This aligns with evidence identifying cluttered design and poor navigation as key deterrents for older users [101,104]. While design guidance often focuses on elements such as button size, textual and navigational clarity remain neglected [101]. The experiences in this study reinforce the need for simplified navigation, clear signposting, and concise exercise programming to reduce user burden and flatten the learning curve—particularly for those wary of technology [105].

Participants also expressed concerns about safety and effectiveness, particularly for unsupervised, home-based exercise. Many questioned whether current or future apps could ensure correct technique and injury prevention—concerns amplified among those with chronic conditions such as osteoporosis or osteoarthritis. Experienced users further noted the risks of non–evidence-based exercise advice, contrasting such apps with the reassurance of supervised exercise. These concerns, rooted in participants' experiences, align with literature highlighting trust as a critical factor in mHealth adoption [106]. The lack of a transparent evidence base underpinning many apps remains a key limitation [107] and may particularly discourage those with long-term conditions who fear injury [108].

Participants also questioned the credibility of some exercise apps, noting that a lack of visible evidence base or endorsement from recognized organizations undermines trust [42]. The perceived absence of personalization and adaptive programming reinforced the need for tailored content and clear signposting of safe options for older adults.

Negative prior experiences and limited understanding of strength training principles led some participants to doubt apps' ability to accurately track progress. Unlike aerobic activity, which can be easily quantified via automated measures such as step counts or heart rate, tracking strength improvements often requires manual input, increasing user burden [105]. Simplifying self-reporting—for example, logging completed sets rather than repetitions or using automated sensing technologies such as accelerometers—could reduce this burden [109,110]. However, the cost of smart resistance devices poses a scalability issue [111], limiting feasibility given that mHealth interventions are intended to reduce, not increase, costs [112].

Overall, these findings correspond to the TAM constructs of perceived usefulness and perceived ease of use. Apprehension regarding safety, evidence base, and design complexity reduced participants' belief that exercise apps can effectively support health and well-being. Such perceptions lower both perceived usefulness and ease of use—key predictors of technology adoption [113]. Consequently, addressing these design and trust-related drawbacks will be essential to enhance older adults' acceptance and engagement with future mHealth interventions.

Theme 4: Desired App Elements and Features

The final theme highlights actionable recommendations for future mHealth intervention development in older adults, broadening the understanding of muscle-strengthening apps for this population. These findings help direct researchers and app developers to enhance uptake and long-term engagement with muscle-strengthening exercise and broader mHealth interventions [114].

The focus group discussions reinforced that app usability is a central priority for older adults. Participants emphasized that many existing mHealth resources do not adequately address their needs, which negatively influences uptake and sustained engagement. Clear navigation, a simplified and uncluttered interface, and high-contrast color schemes were identified as key design features that would enhance usability. Participants drew on previous negative experiences with apps featuring unnecessary functions, visually dense layouts, and small text, aligning with literature underscoring the importance of inclusive digital design [115]. Font size was also highlighted as a critical consideration; although text can be adjusted within device settings, ensuring an appropriate default size may improve

XSL•FO
RenderX

perceived ease of use, as simpler interfaces can reduce cognitive load and support engagement among older adults [116].

Participants emphasized that personalization should be a core component of the app experience. Customizable reminders were viewed as essential, with participants noting that reminder frequency and timing should align with individual routines to prevent disengagement. This reflects previous findings that excessive notifications can contribute to technology-related stress and reduced adherence among older adults [117]. As also highlighted in Theme 2, the timing and tone of prompts were considered crucial: participants generally preferred a single reminder on the morning of an exercise day, delivered in supportive and nonjudgmental language. This aligns with literature demonstrating the importance of encouraging tones in digital prompts for older adults [118]. Evidence from Wolner-Strohmeyer et al [119], who conducted a 12-week PA intervention with adults aged 60 years or older, showed that twice-weekly reminders significantly increased activity levels, further supporting the value of well-calibrated reminders. Consistent with these findings, participants suggested that one appropriately timed reminder would be sufficient. Consequently, future app development should prioritize adaptable reminder systems that can be tailored to users' daily routines to enhance engagement.

Participants expressed a preference for greater autonomy within the muscle-strengthening program, including the ability to select exercises and create personalized routines. They believed that having control over their workouts would increase enjoyment, support longer-term engagement, and allow better alignment with individual physical abilities. This aligns with evidence from previous mHealth research demonstrating that autonomy can enhance adherence. For example, Mair et al [35] reported improved adherence (67%) when participants set their own step and activity goals, while Daniels et al [120] observed a 135% adherence rate in a 5-week intervention where users selected activities and customized workouts—although this figure reflects adherence relative to self-chosen goals. However, such customization is more readily implemented in aerobic-based interventions; strength-based programs present additional challenges regarding safety and effectiveness [121]. Limited knowledge of resistance training in older adults [122] further increases the risk of unbalanced or unsafe exercise choices. Additionally, providing too many options may increase cognitive load, adding to the usability burden identified in mobile app design guidelines [101].

A key message from these discussions was the need for varied entry points in strength programs based on individual capacity. Older adults are not a homogenous group [102], and age-related conditions amplify disparities in musculoskeletal health. Sarcopenia prevalence rises from 5%-13% in adults aged 60-70 years to around 50% in those aged 80 years and older [123], with substantial declines in muscular capacity; for example, leg press 1-repetition maximum dropping from 237 kg in adults aged 60-64 years to 172 kg in those aged 80-85 years in men [124]. These findings highlight the need for structured difficulty levels within mHealth apps, enabling participants to select a starting level aligned with their ability.

Participants stressed the need for instructional clarity through clear video demonstrations of muscle-strengthening exercises. These calls were rooted in safety and confidence in the exercise materials and were an extension of the previous discussion encapsulated in Theme 3. Participants valued video demonstrations due to their accessible features (pause, rewind, and replay), emphasizing that older adults feel it is important to learn muscle-strengthening exercises at their own pace. These specific needs have been catered to in previous interventions, which have demonstrated the benefits of using video demonstrations to assist the provision of muscle-strengthening exercise for older adults [125,126]. These findings also link to the broader literature, which highlights that older adults value personalized learning and the ability to revisit material as needed [127]. A crucial consideration that emerged from discussing the provision of video demonstrations was the demand from participants that videos should be relatable, ideally carried out by an older adult with a relatable physique. Participants felt that having an older adult demonstrate the exercise would help them resonate better and reinforce that muscle-strengthening exercise is not unattainable. These findings are comparable with broader literature, which outlines how age-related promotion of exercise can help foster identification in older adults and enhance engagement [128].

Participants had mixed views on social features of apps. While some felt group-based exercise or peer support would boost their motivation to engage, other participants felt that exercise was a personal experience and did not want social exposure. These thoughts are mirrored in the literature; a previous review by Tong and Laranjo [129] highlighted that while some participants benefit from social features in mHealth activity interventions, many are concerned about direct comparisons to peers, which may become off-putting. The consensus from participants was that social features should be included but optional rather than a default requirement, thereby catering to all needs. The findings from this study link to older adult literature, which highlights that social connectedness can help boost long-term adherence to exercise programs [130]; however, imposing this in an mHealth intervention risks alienating some participants.

The fourth theme is clearly rooted in both TAM key constructs of perceived usefulness and perceived ease of use. The emphasis on simple navigation and high-contrast visuals supports perceived ease of use by reducing the physical and cognitive requirements to engage with the app. Meanwhile, personalization of reminders, exercise selection, and instructional clarity enhances participants' perceived usefulness, ensuring that the app can deliver meaningful benefits.

Strengths and Limitations

One key strength of the study is that the findings offer insight into the applicability of mHealth interventions for muscle-strengthening among older adults. The findings provide a perspective on older adults' needs, wants, thoughts, and feelings regarding engagement with mobile technology, which in turn may enhance participation. By mapping the themes to the TAM core constructs of perceived usefulness and perceived ease of use, the study offers a clear understanding of the factors

XSL•FO
RenderX

influencing mHealth uptake in older adults and presents a framework embedded within a broader co-design process. To the best of the authors' knowledge, this is the first study to outline specific design and intervention recommendations exclusively for muscle-strengthening mHealth interventions for older adults. A further strength of this study was the representation of both experienced and inexperienced older adults in engaging with mHealth technology. This helps guide future interventions to ensure they cater to all needs and maximize intervention effectiveness. Moreover, the inclusion of both males and females ensured diverse perspectives and key design considerations.

The study is not without limitations, primarily recruitment bias. The sample was recruited mainly from the central Scotland area, and the majority were well-educated. Participants' affluent backgrounds likely reduced diversity, which may limit the applicability of the framework to the broader population. Furthermore, due to participants' affluence, it is likely they had a baseline digital literacy higher than that of older adults of lower socioeconomic status [131,132].

Participants in this study may have been more likely to complete exercise, as they may have been attracted to participate in this research. This is therefore not representative of the broader older adult population, in which fewer than 2% to 20% regularly complete muscle-strengthening exercise [29,133]. The study did not formally assess participants' baseline activity levels or adherence to muscle-strengthening guidelines; therefore, it was not possible to contextualize participants' perceptions or completion of current PA guidelines. The sample may have been more health-conscious or physically active than the general older adult population. This limitation should be considered when interpreting findings, as attitudes toward mHealth and exercise may differ among the more physically inactive individuals.

Moreover, the age profile of the participants largely reflects the "young-old" older adult population [134]. The mean age of the sample was 65 years, and participants were generally healthier and potentially more digitally capable than anticipated from older age groups. This may limit the transferability of these findings to older adults living with frailty, multimorbidity, or cognitive impairment.

Practical Recommendations

In line with the principles of the co-design process [44], this study points to several key design considerations for future mHealth interventions aimed at supporting muscle-strengthening exercise in older adults. These can be distilled into five guiding recommendations (1) simplicity and clarity in interface design; (2) personalization of exercise content, reward functions, and reminders; (3) safe and relatable instructional resources, such as video demonstrations with older adult demonstrators; (4) flexible exercise customization; and (5) optional social features.

Conclusion

Overall, this study aimed to identify key features that a muscle-strengthening exercise app should include to enhance engagement and uptake among older adults and to examine the limitations of existing health apps to inform more effective design. Findings highlight both positive and negative aspects of current apps, as well as essential considerations for future development. Across 4 themes, participants recognized the potential of mHealth to support health and well-being, identified motivational features, outlined usability challenges, and proposed desired app elements. Design preferences included large font, clear navigation, personalized exercise options, customizable reminders, instructional videos, and sensitivity to age-related needs. These insights offer a user-informed framework for app refinement. In conclusion, this study provides a foundation for impactful mHealth design targeting muscle-strengthening and sarcopenia prevention in older adults, with the potential to inform future interventions and improve digital health equity.Top of Form

Data Availability

The datasets generated or analyzed during this study are available in the Open Science Framework repository.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Experienced question list.
[DOCX File , 19 KB - aging_v9i1e87332_app1.docx]

Multimedia Appendix 2
Nonexperienced question list.
[DOCX File , 20 KB - aging_v9i1e87332_app2.docx]

Multimedia Appendix 3

XSL•FO
RenderX

Codesign focus group PowerPoint.
[PPTX File , 14123 KB - aging_v9i1e87332_app3.pptx]

Multimedia Appendix 4
Quote table.
[DOCX File , 34 KB - aging_v9i1e87332_app4.docx]

References

1. Lu W, Yuan J, Liu Z, Su Z, Shen Y, Li S. Worldwide trends in mortality for hypertensive heart disease from 1990 to 2019 with projection to 2034: data from the Global Burden of Disease 2019 study. Eur J Prev Cardiol 2024;31(1):23-37. [doi: 10.1093/eurjpc/zwad262] [Medline: 37665956]
2. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021;21(1):1556 [FREE Full text] [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
3. Hayes LD, Herbert P, Sculthorpe N, Grace F. High intensity interval training (HIIT) produces small improvements in fasting glucose, insulin, and insulin resistance in sedentary older men but not masters athletes. Exp Gerontol 2020;140:111074. [doi: 10.1016/j.exger.2020.111074] [Medline: 32860864]
4. Del Buono MG, Arena R, Borlaug BA, Carbone S, Canada JM, Kirkman DL. Exercise intolerance in patients with heart failure: JACC state-of-the-art review. J Am Coll Cardiol 2019;73(17):2209-2225 [FREE Full text] [doi: 10.1016/j.jacc.2019.01.072] [Medline: 31047010]
5. Tavares NMB, Silva JM, Silva MDMD, Silva LDT, Souza JND, Ithamar L. Balance, gait, functionality and fall occurrence in adults and older adults with type 2 diabetes mellitus and associated peripheral neuropathy. Clin Pract 2024;14(5):2044-2055 [FREE Full text] [doi: 10.3390/clinpract14050161] [Medline: 39451876]
6. Ageing. United Nations. 2024. URL: https://www.un.org/en/global-issues/ageing [accessed 2026-02-27]
7. Cho MR, Lee S, Song SK. A review of sarcopenia pathophysiology, diagnosis, treatment and future direction. J Korean Med Sci 2022;37(18):e146 [FREE Full text] [doi: 10.3346/jkms.2022.37.e146] [Medline: 35535373]
8. Hayes LD, Elliott BT, Yasar Z, Bampouras TM, Sculthorpe NF, Sanal-Hayes NEM. High intensity interval training (HIIT) as a potential countermeasure for phenotypic characteristics of sarcopenia: a scoping review. Front Physiol 2021;12:715044 [FREE Full text] [doi: 10.3389/fphys.2021.715044] [Medline: 34504439]
9. Huo F, Liu Q, Liu H. Contribution of muscle satellite cells to sarcopenia. Front Physiol 2022;13:892749 [FREE Full text] [doi: 10.3389/fphys.2022.892749] [Medline: 36035464]
10. Kubat GB, Bouhamida E, Ulger O, Turkel I, Pedriali G, Ramaccini D. Mitochondrial dysfunction and skeletal muscle atrophy: causes, mechanisms, and treatment strategies. Mitochondrion 2023;72:33-58. [doi: 10.1016/j.mito.2023.07.003] [Medline: 37451353]
11. Hämäläinen O, Tirkkonen A, Savikangas T, Alén M, Sipilä S, Hautala A. Low physical activity is a risk factor for sarcopenia: a cross-sectional analysis of two exercise trials on community-dwelling older adults. BMC Geriatr 2024;24(1):212 [FREE Full text] [doi: 10.1186/s12877-024-04764-1] [Medline: 38424514]
12. Narici M, McPhee J, Conte M, Franchi MV, Mitchell K, Tagliaferri S. Age-related alterations in muscle architecture are a signature of sarcopenia: the ultrasound sarcopenia index. J Cachexia Sarcopenia Muscle 2021;12(4):973-982 [FREE Full text] [doi: 10.1002/jcsm.12720] [Medline: 34060717]
13. Okabe T, Suzuki M, Goto H, Iso N, Cho K, Hirata K. Sex differences in age-related physical changes among community-dwelling adults. J Clin Med 2021;10(20):4800 [FREE Full text] [doi: 10.3390/jcm10204800] [Medline: 34682922]
14. Sayer AA, Syddall H, Martin H, Patel H, Baylis D, Cooper C. The developmental origins of sarcopenia. J Nutr Health Aging 2008;12(7):427-432 [FREE Full text] [doi: 10.1007/BF02982703] [Medline: 18615224]
15. Samoilova YG, Matveeva MV, Khoroshunova EA, Kudlay DA, Oleynik OA, Spirina LV. Markers for the prediction of probably sarcopenia in middle-aged individuals. J Pers Med 2022;12(11):1830 [FREE Full text] [doi: 10.3390/jpm12111830] [Medline: 36579539]
16. Rodrigues F, Domingos C, Monteiro D, Morouço P. A review on aging, sarcopenia, falls, and resistance training in community-dwelling older adults. Int J Environ Res Public Health 2022;19(2):874 [FREE Full text] [doi: 10.3390/ijerph19020874] [Medline: 35055695]
17. Papadopoulou SK. Sarcopenia: a contemporary health problem among older adult populations. Nutrients 2020;12(5):1293 [FREE Full text] [doi: 10.3390/nu12051293] [Medline: 32370051]
18. Herbert P, Hayes L, Sculthorpe N, Grace F. HIIT produces increases in muscle power and free testosterone in male masters athletes. Endocr Connect 2017;6(7):430-436 [FREE Full text] [doi: 10.1530/EC-17-0159] [Medline: 28794164]
19. Sculthorpe NF, Herbert P, Grace F. One session of high-intensity interval training (HIIT) every 5 days, improves muscle power but not static balance in lifelong sedentary ageing men: a randomized controlled trial. Medicine (Baltimore) 2017;96(6):e6040 [FREE Full text] [doi: 10.1097/MD.0000000000006040] [Medline: 28178145]

XSL•FO
RenderX

20. Sculthorpe N, Herbert P, Grace FM. Low-frequency high-intensity interval training is an effective method to improve muscle power in lifelong sedentary aging men: A randomized controlled trial. J Am Geriatr Soc 2015;63(11):2412-2413. [doi: 10.1111/jgs.13863] [Medline: 26603065]
21. Herbert P, Hayes LD, Sculthorpe N, Grace FM. High-intensity interval training (HIIT) increases insulin-like growth factor-I (IGF-I) in sedentary aging men but not masters' athletes: an observational study. Aging Male 2017;20(1):54-59 [FREE Full text] [doi: 10.1080/13685538.2016.1260108] [Medline: 28042739]
22. Hayes LD, Elliott BT. Short-term exercise training inconsistently influences basal testosterone in older men: a systematic review and meta-analysis. Front Physiol 2018;9:1878 [FREE Full text] [doi: 10.3389/fphys.2018.01878] [Medline: 30692929]
23. Vlietstra L, Hendrickx W, Waters DL. Exercise interventions in healthy older adults with sarcopenia: a systematic review and meta-analysis. Australas J Ageing 2018;37(3):169-183. [doi: 10.1111/ajag.12521] [Medline: 29638028]
24. Fiatarone MA, Marks EC, Ryan ND, Meredith CN, Lipsitz LA, Evans WJ. High-intensity strength training in nonagenarians. Effects on skeletal muscle. JAMA 1990;263(22):3029-3034. [Medline: 2342214]
25. Reeves ND, Maganaris CN, Narici MV. Plasticity of dynamic muscle performance with strength training in elderly humans. Muscle Nerve 2005;31(3):355-364. [doi: 10.1002/mus.20275] [Medline: 15654690]
26. Kanda K, Yoda T, Suzuki H, Okabe Y, Mori Y, Yamasaki K. Effects of low-intensity bodyweight training with slow movement on motor function in frail elderly patients: a prospective observational study. Environ Health Prev Med 2018;23(1):4 [FREE Full text] [doi: 10.1186/s12199-018-0693-4] [Medline: 29385984]
27. Wilkinson DJ, Piasecki M, Atherton PJ. The age-related loss of skeletal muscle mass and function: measurement and physiology of muscle fibre atrophy and muscle fibre loss in humans. Ageing Res Rev 2018;47:123-132 [FREE Full text] [doi: 10.1016/j.arr.2018.07.005] [Medline: 30048806]
28. Scottish Government. The Scottish Health Survey 2021: Volume 1—Main Report. 2021. URL: https://www.gov.scot/publications/scottish-health-survey-2021-volume-1-main-report/pages/11/ [accessed 2026-02-27]
29. NHS Digital. Health Survey for England 2021: Part 2. 2023. URL: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2021-part-2 [accessed 2026-02-27]
30. CDC. QuickStats: percentage of adults aged ≥18 years who met the federal guidelines for muscle-strengthening physical activity, by age group and sex—National Health Interview Survey, United States, 2020. MMWR Morb Mortal Wkly Rep 2022;71(18):642 [FREE Full text] [doi: 10.15585/mmwr.mm7118a6] [Medline: 35511726]
31. Langhammer B, Bergland A, Rydwik E. The importance of physical activity exercise among older people. Biomed Res Int 2018;2018:1-3 [FREE Full text] [doi: 10.1155/2018/7856823] [Medline: 30627571]
32. Changizi M, Kaveh MH. Effectiveness of the mHealth technology in improvement of healthy behaviors in an elderly population-a systematic review. Mhealth 2017;3:51 [FREE Full text] [doi: 10.21037/mhealth.2017.08.06] [Medline: 29430455]
33. Stockwell S, Schofield P, Fisher A, Firth J, Jackson SE, Stubbs B. Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults: a systematic review and meta-analysis. Exp Gerontol 2019;120:68-87. [doi: 10.1016/j.exger.2019.02.020] [Medline: 30836130]
34. Sculthorpe NF, McLaughlin M, Cerexhe L, Macdonald E, Dello Iacono A, Sanal-Hayes NEM. Tracking persistent symptoms in Scotland (TraPSS): a longitudinal prospective cohort study of COVID-19 recovery after mild acute infection. BMJ Open 2025;15(1):e086646 [FREE Full text] [doi: 10.1136/bmjopen-2024-086646] [Medline: 39819953]
35. Mair JL, Hayes LD, Campbell AK, Buchan DS, Easton C, Sculthorpe N. A personalized smartphone-delivered just-in-time adaptive intervention (JitaBug) to increase physical activity in older adults: mixed methods feasibility study. JMIR Form Res 2022;6(4):e34662 [FREE Full text] [doi: 10.2196/34662] [Medline: 35389348]
36. Mair JL, Hayes LD, Campbell AK, Sculthorpe N. Should we use activity tracker data from smartphones and wearables to understand population physical activity patterns? J Meas Phys Behav 2022;5(1):3-7. [doi: 10.1123/jmpb.2021-0012]
37. Cunningham C, O'Sullivan R. Why physical activity matters for older adults in a time of pandemic. Eur Rev Aging Phys Act 2020;17(1):16 [FREE Full text] [doi: 10.1186/s11556-020-00249-3] [Medline: 32983273]
38. Baumann H, Singh B, Staiano AE, Gough C, Ahmed M, Fiedler J. Effectiveness of mHealth interventions targeting physical activity, sedentary behaviour, sleep or nutrition on emotional, behavioural and eating disorders in adolescents: a systematic review and meta-analysis. Front Digit Health 2025;7:1593677 [FREE Full text] [doi: 10.3389/fdgth.2025.1593677] [Medline: 40761810]
39. Gluchowski A, Bilsborough H, Mcdermott J, Hawley-Hague H, Todd C. ‘A lot of people just go for walks, and don’t do anything else’: older adults in the UK are not aware of the strength component embedded in the Chief Medical Officers’ physical activity guidelines—a qualitative study. Int J Environ Res Public Health 2022;19(16):10002 [FREE Full text] [doi: 10.3390/ijerph191610002] [Medline: 36011634]
40. Berry ECJ, Sculthorpe NF, Warner A, Mather JD, Sanal-Hayes NEM, Hayes LD. A scoping review of the feasibility, usability, and efficacy of digital interventions in older adults concerning physical activity and/or exercise. Front Aging 2025;6:1516481. [doi: 10.3389/fragi.2025.1516481] [Medline: 40290578]
41. Portenhauser AA, Terhorst Y, Schultchen D, Sander LB, Denkinger MD, Stach M. Mobile apps for older adults: systematic search and evaluation within online stores. JMIR Aging 2021;4(1):e23313 [FREE Full text] [doi: 10.2196/23313] [Medline: 33605884]

XSL•FO
RenderX

42. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024;24(1):506 [FREE Full text] [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
43. Sadare O, Melvin T, Harvey H, Vollebregt E, Gilbert S. Can Apple and Google continue as health app gatekeepers as well as distributors and developers? NPJ Digit Med 2023;6(1):8 [FREE Full text] [doi: 10.1038/s41746-023-00754-6] [Medline: 36721023]
44. Noorbergen TJ, Adam MTP, The University of Newcastle, Roxburgh M, Teubner T, Technical University Berlin. Co-design in mHealth systems development: insights from a systematic literature review. AIS Transactions on Human-Computer Interaction 2021;13(2):175-205. [doi: 10.17705/1thci.00147]
45. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
46. Noorbergen TJ, Adam MTP, Teubner T, Collins CE. Using co-design in mobile health system development: a qualitative study with experts in co-design and mobile health system development. JMIR Mhealth Uhealth 2021;9(11):e27896 [FREE Full text] [doi: 10.2196/27896] [Medline: 34757323]
47. Son BK, Miura T, Yabu K, Sumikawa Y, Kim D, Lyu W. The co-design/co-development and evaluation of an online frailty check application for older adults: participatory action research with older adults. Int J Environ Res Public Health 2023;20(12):6101 [FREE Full text] [doi: 10.3390/ijerph20126101] [Medline: 37372688]
48. Thornton EJ, Hayes LD, Goodwin DS, Sculthorpe N, Prior Y, Sanal-Hayes NE. Managing energy, and shaping care: insights from adults with myalgic encephalomyelitis/chronic fatigue syndrome through co-production workshops. Am J Med 2025;138(6):1001-1009. [doi: 10.1016/j.amjmed.2025.02.008] [Medline: 39961545]
49. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
50. Sanz MF, Acha BV, García MF. Co-design for people-centred care digital solutions: a literature review. Int J Integr Care 2021;21(2):16 [FREE Full text] [doi: 10.5334/ijic.5573] [Medline: 33981193]
51. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017;8:1687 [FREE Full text] [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
52. Hennink MM, Kaiser BN, Weber MB. What influences saturation? Estimating sample sizes in focus group research. Qual Health Res 2019;29(10):1483-1496 [FREE Full text] [doi: 10.1177/1049732318821692] [Medline: 30628545]
53. EuroStat. International Standard Classification of Education (ISCED). Europe; 2014. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=International_Standard_Classification_of_Education_(ISCED) [accessed 2026-02-27]
54. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
55. Hill Z, Tawiah-Agyemang C, Kirkwood B, Kendall C. Are verbatim transcripts necessary in applied qualitative research: experiences from two community-based intervention trials in Ghana. Emerg Themes Epidemiol 2022;19(1):5 [FREE Full text] [doi: 10.1186/s12982-022-00115-w] [Medline: 35765012]
56. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019;11(4):589-597. [doi: 10.1080/2159676x.2019.1628806]
57. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319. [doi: 10.2307/249008]
58. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
59. Arno A, Thomas S. The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis. BMC Public Health 2016;16(1):676 [FREE Full text] [doi: 10.1186/s12889-016-3272-x] [Medline: 27475752]
60. Cammisuli D, Pietrabissa G, Castelnuovo G. Improving wellbeing of community-dwelling people with mild cognitive impairment: the SENIOR (SystEm of Nudge theory based ICT applications for OldeR citizens) project. Neural Regen Res 2021;16(5):963-966 [FREE Full text] [doi: 10.4103/1673-5374.297063] [Medline: 33229736]
61. Yamada Y, Uchida T, Sasaki S, Taguri M, Shiose T, Ikenoue T. Nudge-based interventions on health promotion activity among very old people: a pragmatic, 2-arm, participant-blinded randomized controlled trial. J Am Med Dir Assoc 2023;24(3):390-394.e5 [FREE Full text] [doi: 10.1016/j.jamda.2022.11.009] [Medline: 36529275]
62. Peng R, Cao Z, Hu S, Liu X, Guo Y, Li X. Frail older adults' needs and preferences for mobile health exercise interventions guided by nudge theory: a qualitative analysis. J Clin Nurs 2025;34(5):1830-1839. [doi: 10.1111/jocn.17419] [Medline: 39215431]
63. Kwan RY, Lee D, Lee PH, Tse M, Cheung DS, Thiamwong L. Effects of an mHealth brisk walking intervention on increasing physical activity in older people with cognitive frailty: pilot randomized controlled trial. JMIR Mhealth Uhealth 2020;8(7):e16596 [FREE Full text] [doi: 10.2196/16596] [Medline: 32735218]
64. Zhang J, Bloom I, Dennison EM, Ward KA, Robinson SM, Barker M. Understanding influences on physical activity participation by older adults: a qualitative study of community-dwelling older adults from the Hertfordshire Cohort Study, UK. PLoS One 2022;17(1):e0263050 [FREE Full text] [doi: 10.1371/journal.pone.0263050] [Medline: 35077522]

XSL•FO
RenderX

65. Rúa-Alonso M, Bovolini A, Costa-Brito AR, Vaz C, Marques E, Serra N. Exploring perceived barriers to physical activity among older adults living in low-population density regions: gender differences and associations with activity dimensions. Healthcare (Basel) 2023;11(22):2948. [doi: 10.3390/healthcare11222948] [Medline: 37998440]
66. Hurst C, Dismore L, Granic A, Tullo E, Noble JM, Hillman SJ. Attitudes and barriers to resistance exercise training for older adults living with multiple long-term conditions, frailty, and a recent deterioration in health: qualitative findings from the Lifestyle in Later Life–Older People's Medicine (LiLL-OPM) study. BMC Geriatr 2023;23(1):772 [FREE Full text] [doi: 10.1186/s12877-023-04461-5] [Medline: 38001414]
67. Kilgour AHM, Rutherford M, Higson J, Meredith SJ, McNiff J, Mitchell S. Barriers and motivators to undertaking physical activity in adults over 70-a systematic review of the quantitative literature. Age Ageing 2024;53(4):afae080 [FREE Full text] [doi: 10.1093/ageing/afae080] [Medline: 38651329]
68. Zhou F, Zhang H, Wang HY, Liu LF, Zhang XG. Barriers and facilitators to older adult participation in intergenerational physical activity program: a systematic review. Aging Clin Exp Res 2024;36(1):39 [FREE Full text] [doi: 10.1007/s40520-023-02652-z] [Medline: 38345651]
69. Sciamanna CN, Ladwig MA, Conroy DE, Schmitz KH, Silvis ML, Ballentine NH. Feasibility and impact of a 1-minute daily functional exercise regimen prescribed to older adults by their primary care physician. Prev Med Rep 2021;21:101307 [FREE Full text] [doi: 10.1016/j.pmedr.2020.101307] [Medline: 33520612]
70. Ghram A, Briki W, Mansoor H, Al-Mohannadi AS, Lavie CJ, Chamari K. Home-based exercise can be beneficial for counteracting sedentary behavior and physical inactivity during the COVID-19 pandemic in older adults. Postgrad Med 2021;133(5):469-480. [doi: 10.1080/00325481.2020.1860394] [Medline: 33275479]
71. Solis-Navarro L, Gismero A, Fernández-Jané C, Torres-Castro R, Solá-Madurell M, Bergé C. Effectiveness of home-based exercise delivered by digital health in older adults: a systematic review and meta-analysis. Age Ageing 2022;51(11):afac243 [FREE Full text] [doi: 10.1093/ageing/afac243] [Medline: 36346736]
72. Hayes EJ, Stevenson E, Sayer AA, Granic A, Hurst C. Recovery from resistance exercise in older adults: a systematic scoping review. Sports Med Open 2023;9(1):51 [FREE Full text] [doi: 10.1186/s40798-023-00597-1] [Medline: 37395837]
73. Chiu CJ, Kuo SE, Lin DC. Technology-embedded health education on nutrition for middle-aged and older adults living in the community. Glob Health Promot 2019;26(3):80-87. [doi: 10.1177/1757975917732351] [Medline: 29099337]
74. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022;24(2):e34790 [FREE Full text] [doi: 10.2196/34790] [Medline: 35171099]
75. Physical activity guidelines for adults aged 19 to 64. National Health Service, UK. 2024. URL: https://www.nhs.uk/live-well/exercise/physical-activity-guidelines-for-adults-aged-19-to-64/ [accessed 2026-02-27]
76. Mulhare B, O'Hagan C, Harrison M, Fagan S, Kehoe B. 92 Enhancing self-efficacy for and adherence to resistance training in older adults: a multi-stakeholder perspective of a supervised-to-unsupervised step-down model. Eur J Public Health 2024;34(Supplement_2):ckae114.061. [doi: 10.1093/eurpub/ckae114.061]
77. Adnan A, Irvine RE, Williams A, Harris M, Antonacci G. Improving acceptability of mHealth apps-the use of the technology acceptance model to assess the acceptability of mHealth apps: systematic review. J Med Internet Res 2025;27:e66432 [FREE Full text] [doi: 10.2196/66432] [Medline: 40334265]
78. Zhu Y, Long Y, Wang H, Lee KP, Zhang L, Wang SJ. Digital behavior change intervention designs for habit formation: systematic review. J Med Internet Res 2024;26:e54375 [FREE Full text] [doi: 10.2196/54375] [Medline: 38787601]
79. Kim SK, Hwang HR, Byun KS, Park SY, Moon SH, Lee H. Development of a digital intervention incorporating habit formation techniques for medication adherence. Inquiry 2025;62:469580251343785 [FREE Full text] [doi: 10.1177/00469580251343785] [Medline: 40542667]
80. Sohaib Aslam A, van Luenen S, Aslam S, van Bodegom D, Chavannes NH. A systematic review on the use of mHealth to increase physical activity in older people. Clinical eHealth 2020;3:31-39. [doi: 10.1016/j.ceh.2020.04.002]
81. Bowen PG, Affuso O, Opoku-Agyeman W, Mixon VR, Clay OJ. Texting older sisters to step to manage obesity in older black women: a feasibility study. Am J Prev Med 2022;63(1):S56-S66 [FREE Full text] [doi: 10.1016/j.amepre.2022.03.014] [Medline: 35725141]
82. Subedi N, Rawstorn JC, Gao L, Koorts H, Maddison R. Implementation of telerehabilitation interventions for the self-management of cardiovascular disease: systematic review. JMIR Mhealth Uhealth 2020;8(11):e17957 [FREE Full text] [doi: 10.2196/17957] [Medline: 33245286]
83. Martínez de la Cal J, Fernández-Sánchez M, Matarán-Peñarrocha GA, Hurley DA, Castro-Sánchez AM, Lara-Palomo IC. Physical therapists' opinion of E-Health treatment of chronic low back pain. Int J Environ Res Public Health 2021;18(4):1889 [FREE Full text] [doi: 10.3390/ijerph18041889] [Medline: 33669249]
84. Krukowski RA, Denton AH, König LM. Correction: impact of feedback generation and presentation on self-monitoring behaviors, dietary intake, physical activity, and weight: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2024;21(1):20 [FREE Full text] [doi: 10.1186/s12966-024-01569-8] [Medline: 38395892]
85. Geraedts HAE, Zijlstra W, Zhang W, Spoorenberg SLW, Báez M, Far IK. A home-based exercise program driven by tablet application and mobility monitoring for frail older adults: feasibility and practical implications. Prev Chronic Dis 2017;14:E12 [FREE Full text] [doi: 10.5888/pcd14.160227] [Medline: 28152361]

XSL•FO
RenderX

86. Feng S, Mäntymäki M, Dhir A, Salmela H. How self-tracking and the quantified self promote health and well-being: systematic review. J Med Internet Res 2021;23(9):e25171. [doi: 10.2196/25171]
87. Maltseva K, Lutz C. A quantum of self: a study of self-quantification and self-disclosure. Comput Human Behav 2018;81:102-114. [doi: 10.1016/j.chb.2017.12.006]
88. Ghanvatkar S, Kankanhalli A, Rajan V. User models for personalized physical activity interventions: scoping review. JMIR Mhealth Uhealth 2019;7(1):e11098 [FREE Full text] [doi: 10.2196/11098] [Medline: 30664474]
89. Hvalič-Touzery S, Šetinc M, Dolničar V. Benefits of a wearable activity tracker with safety features for older adults: an intervention study. Int J Environ Res Public Health 2022;19(23):15723 [FREE Full text] [doi: 10.3390/ijerph192315723] [Medline: 36497796]
90. Razaghizad A, McKee T, Malhamé I, Friedrich MG, Giannetti N, Coristine A. Mobile health fitness interventions: impact of features on routine use and data sharing acceptability. JACC Adv 2023;2(8):100613 [FREE Full text] [doi: 10.1016/j.jacadv.2023.100613] [Medline: 38938369]
91. Cajita MI, Hodgson NA, Lam KW, Yoo S, Han HR. Facilitators of and barriers to mHealth adoption in older adults with heart failure. Comput Inform Nurs 2018;36(8):376-382 [FREE Full text] [doi: 10.1097/CIN.0000000000000442] [Medline: 29742549]
92. Nurmi J, Knittle K, Ginchev T, Khattak F, Helf C, Zwickl P. Engaging users in the behavior change process with digitalized motivational interviewing and gamification: development and feasibility testing of the precious app. JMIR Mhealth Uhealth 2020;8(1):e12884 [FREE Full text] [doi: 10.2196/12884] [Medline: 32003750]
93. Wang M, Xu J, Zhou X, Li X, Zheng Y. Effectiveness of gamification interventions to improve physical activity and sedentary behavior in children and adolescents: systematic review and meta-analysis. JMIR Serious Games 2025;13:e68151-e68151 [FREE Full text] [doi: 10.2196/68151] [Medline: 40966596]
94. Mazéas A, Forestier C, Harel G, Duclos M, Chalabaev A. The impact of a gamified intervention on daily steps in real-life conditions: retrospective analysis of 4800 individuals. J Med Internet Res 2024;26:e47116 [FREE Full text] [doi: 10.2196/47116] [Medline: 39133533]
95. Xu L, Shi H, Shen M, Ni Y, Zhang X, Pang Y. The effects of mHealth-based gamification interventions on participation in physical activity: systematic review. JMIR Mhealth Uhealth 2022;10(2):e27794 [FREE Full text] [doi: 10.2196/27794] [Medline: 35113034]
96. Dawson J, Nee R, Ramirez C, Reyes S, Sanchez D, Sukhadia T. Gamification in mHealth apps for rehabilitation: protocol for a scoping review. JMIR Res Protoc 2025;14:e63600 [FREE Full text] [doi: 10.2196/63600] [Medline: 40294403]
97. Yu T, Parry M, Yu T, Xu L, Wu Y, Zeng T. Effectiveness of mobile health-based gamification interventions for improving physical activity in individuals with cardiovascular diseases: systematic review and meta-analysis of randomized controlled trials. JMIR Serious Games 2025;13:e64410 [FREE Full text] [doi: 10.2196/64410] [Medline: 39854099]
98. Carvalho R, Machado D, Brandao P. Gamification, mHealth and user adherence. : IEEE; 2020 Presented at: 2020 IEEE Symposium on Computers and Communications (ISCC); June 29-July 2, 2020; Rennes p. 1-6. [doi: 10.1109/iscc50000.2020.9219674]
99. Pan J, Dong H, Bryan-Kinns N. Perception and initial adoption of mobile health services of older adults in London: mixed methods investigation. JMIR Aging 2021;4(4):e30420 [FREE Full text] [doi: 10.2196/30420] [Medline: 34807836]
100. Liu BCP. The impact of intersectionality of multiple identities on the digital health divide, quality of life and loneliness amongst older adults in the UK. Br J Soc Work 2021;51(8):3077-3097. [doi: 10.1093/bjsw/bcaa149]
101. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023;11:e43186 [FREE Full text] [doi: 10.2196/43186] [Medline: 37733401]
102. Meredith SJ, Cox NJ, Ibrahim K, Higson J, McNiff J, Mitchell S. Factors that influence older adults' participation in physical activity: a systematic review of qualitative studies. Age Ageing 2023;52(8):afad145 [FREE Full text] [doi: 10.1093/ageing/afad145] [Medline: 37595070]
103. Frishammar J, Essén A, Bergström F, Ekman T. Digital health platforms for the elderly? Key adoption and usage barriers and ways to address them. Technol Forecast Soc Change 2023;189:122319. [doi: 10.1016/j.techfore.2023.122319]
104. Zhang S, Song J. An empirical investigation into the preferences of the elderly for user interface design in personal electronic health record systems. Front Digit Health 2023;5:1289904 [FREE Full text] [doi: 10.3389/fdgth.2023.1289904] [Medline: 38348367]
105. Nebeker C, Zlatar ZZ. Learning from older adults to promote independent physical activity using mobile health (mHealth). Front Public Health 2021;9:703910 [FREE Full text] [doi: 10.3389/fpubh.2021.703910] [Medline: 34476229]
106. Liu JYW, Sorwar G, Rahman MS, Hoque MR. The role of trust and habit in the adoption of mHealth by older adults in Hong Kong: a healthcare technology service acceptance (HTSA) model. BMC Geriatr 2023;23(1):73 [FREE Full text] [doi: 10.1186/s12877-023-03779-4] [Medline: 36737712]
107. McGarrigle L, Boulton E, Todd C. Map the apps: a rapid review of digital approaches to support the engagement of older adults in strength and balance exercises. BMC Geriatr 2020;20(1):483 [FREE Full text] [doi: 10.1186/s12877-020-01880-6] [Medline: 33208117]

XSL•FO
RenderX

108. Cavill NA, Foster CEM. Enablers and barriers to older people's participation in strength and balance activities: a review of reviews. J Frailty Sarcopenia Falls 2018;03(02):105-113. [doi: 10.22540/jfsf-03-105]
109. Mohieldin S, Batsis JA, Minor CM, Halter RJ, Petersen CL. BandPass: a Bluetooth-enabled remote monitoring device for sarcopenia. Canada: IEEE; 2021 Presented at: 2021 IEEE International Conference on Communications Workshops (ICC Workshops); June 14-23, 2021; Montreal p. 1-6 URL: https://europepmc.org/abstract/MED/34745771 [doi: 10.1109/iccworkshops50388.2021.9473520]
110. Tiggemann CL, Pietta-Dias C, Schoenell MCW, Noll M, Alberton CL, Pinto RS. Rating of perceived exertion as a method to determine training loads in strength training in elderly women: a randomized controlled study. Int J Environ Res Public Health 2021;18(15):7892 [FREE Full text] [doi: 10.3390/ijerph18157892] [Medline: 34360184]
111. Middleton A, Simpson KN, Bettger JP, Bowden MG. COVID-19 pandemic and beyond: considerations and costs of telehealth exercise programs for older adults with functional impairments living at home-lessons learned from a pilot case study. Phys Ther 2020;100(8):1278-1288 [FREE Full text] [doi: 10.1093/ptj/pzaa089] [Medline: 32372072]
112. Bommakanti KK, Smith LL, Liu L, Do D, Cuevas-Mota J, Collins K. Requiring smartphone ownership for mHealth interventions: who could be left out? BMC Public Health 2020;20(1):81 [FREE Full text] [doi: 10.1186/s12889-019-7892-9] [Medline: 31959145]
113. Lee AT, Ramasamy RK, Subbarao A. Understanding psychosocial barriers to healthcare technology adoption: a review of TAM technology acceptance model and unified theory of acceptance and use of technology and UTAUT frameworks. Healthcare (Basel) 2025;13(3):250 [FREE Full text] [doi: 10.3390/healthcare13030250] [Medline: 39942440]
114. Van Acker J, Maenhout L, Compernolle S. Older adults' user engagement with mobile health: a systematic review of qualitative and mixed-methods studies. Innov Aging 2023;7(2):igad007 [FREE Full text] [doi: 10.1093/geroni/igad007] [Medline: 37007638]
115. Van Calis JFF, Naaldenberg J, Van der Cruijsen AW, Koks-Leensen MC, Leusink GL, Bevelander KE. Inclusive digital platforms: designing for and with users with mild intellectual disabilities or low literacy skills. Comput Hum Behav Rep 2025;17:100617. [doi: 10.1016/j.chbr.2025.100617]
116. Kanakuze J. The effect of adaptive user interfaces on task completion time in mobile health applications for elderly users in Rwanda. Int J Comput Eng 2025;4(3):31-41. [doi: 10.47941/ijce.2842]
117. Cao Y, Li J, Qin X, Hu B. Examining the effect of overload on the MHealth application resistance behavior of elderly users: an SOR perspective. Int J Environ Res Public Health 2020;17(18):6658 [FREE Full text] [doi: 10.3390/ijerph17186658] [Medline: 32932679]
118. Janols R, Sandlund M, Lindgren H, Pettersson B. Older adults as designers of behavior change strategies to increase physical activity-report of a participatory design process. Front Public Health 2022;10:988470 [FREE Full text] [doi: 10.3389/fpubh.2022.988470] [Medline: 36620266]
119. Wolner-Strohmeyer G, Keilani M, Mähr B, Morawetz E, Zdravkovic A, Wagner B. Can reminders improve adherence to regular physical activity and exercise recommendations in people over 60 years old?: a randomized controlled study. Wien Klin Wochenschr 2021;133(11-12):620-624 [FREE Full text] [doi: 10.1007/s00508-020-01699-6] [Medline: 32591933]
120. Daniels K, Vonck S, Robijns J, Quadflieg K, Bergs J, Spooren A. Exploring the feasibility of a 5-week mHealth intervention to enhance physical activity and an active, healthy lifestyle in community-dwelling older adults: mixed methods study. JMIR Aging 2025;8:e63348 [FREE Full text] [doi: 10.2196/63348] [Medline: 39869906]
121. Bennie JA, Shakespear-Druery J, De Cocker K. Muscle-strengthening exercise epidemiology: a new frontier in chronic disease prevention. Sports Med Open 2020;6(1):40 [FREE Full text] [doi: 10.1186/s40798-020-00271-w] [Medline: 32844333]
122. Di Lorito C, Long A, Byrne A, Harwood RH, Gladman JR, Schneider S. Exercise interventions for older adults: a systematic review of meta-analyses. J Sport Health Sci 2021;10(1):29-47 [FREE Full text] [doi: 10.1016/j.jshs.2020.06.003] [Medline: 32525097]
123. Tu DY, Kao FM, Tsai ST, Tung TH. Sarcopenia among the elderly population: a systematic review and meta-analysis of randomized controlled trials. Healthcare (Basel) 2021;9(6):650 [FREE Full text] [doi: 10.3390/healthcare9060650] [Medline: 34072617]
124. Parrino RL, Strand KL, Hockman AC, Signorile JF. Leg press and chest press strength normative values by half-decades in older persons. Exp Gerontol 2021;150:111401. [doi: 10.1016/j.exger.2021.111401] [Medline: 33974938]
125. Daly RM, Gianoudis J, Hall T, Mundell NL, Maddison R. Feasibility, usability, and enjoyment of a home-based exercise program delivered via an exercise app for musculoskeletal health in community-dwelling older adults: short-term prospective pilot study. JMIR Mhealth Uhealth 2021;9(1):e21094 [FREE Full text] [doi: 10.2196/21094] [Medline: 33439147]
126. Zhang L, Ge Y, Zhao W, Shu X, Kang L, Wang Q, et al. A 4-week mobile app-based telerehabilitation program vs conventional in-person rehabilitation in older adults with sarcopenia: randomized controlled trial. J Med Internet Res 2025;27:e67846 [FREE Full text] [doi: 10.2196/67846] [Medline: 39854716]
127. Sharit J, Czaja SJ. Overcoming older adult barriers to learning through an understanding of perspectives on human information processing. J Appl Gerontol 2020;39(3):233-241 [FREE Full text] [doi: 10.1177/0733464818794574] [Medline: 30160573]

XSL•FO
RenderX

128. Chia F, Huang WY, Huang H, Wu CE. Promoting healthy behaviors in older adults to optimize health-promoting lifestyle: an intervention study. Int J Environ Res Public Health 2023;20(2):1628 [FREE Full text] [doi: 10.3390/ijerph20021628] [Medline: 36674395]
129. Tong HL, Laranjo L. The use of social features in mobile health interventions to promote physical activity: a systematic review. NPJ Digit Med 2018;1(1):43 [FREE Full text] [doi: 10.1038/s41746-018-0051-3] [Medline: 31304323]
130. Suragarn U, Hain D, Pfaff G. Approaches to enhance social connection in older adults: an integrative review of literature. Aging Health Res 2021;1(3):100029. [doi: 10.1016/j.ahr.2021.100029]
131. Yuen E, Winter N, Savira F, Huggins CE, Nguyen L, Cooper P. Digital health literacy and its association with sociodemographic characteristics, health resource use, and health outcomes: rapid review. Interact J Med Res 2024;13:e46888 [FREE Full text] [doi: 10.2196/46888] [Medline: 39059006]
132. Fang W, Nie C. Inequalities in digital literacy: exploring the disparity in tangible outcomes of internet use among college students in China. Front Commun 2025;10:1601240. [doi: 10.3389/fcomm.2025.1601240]
133. Strain T, Fitzsimons C, Kelly P, Mutrie N. BMC Public Health 2016;16(1):1108 [FREE Full text] [doi: 10.1186/s12889-016-3774-6] [Medline: 27769211]
134. Min SH, Topaz M, Lee C, Schnall R. Understanding changes in mental health symptoms from young-old to old-old adults by sex using multiple-group latent transition analysis. Geroscience 2023;45(3):1791-1801 [FREE Full text] [doi: 10.1007/s11357-023-00729-1] [Medline: 36626018]

Abbreviations

BCTTv1: Behavior Change Technique Taxonomy version 1
COREQ: Consolidated Criteria for Reporting Qualitative Research
mHealth: mobile health
PA: physical activity
TAM: Technology Acceptance Model
U3A: University of the Third Age

Edited by P Jansons; submitted 07.Nov.2025; peer-reviewed by L Kang, L Wilkstrom; comments to author 14.Jan.2026; revised version received 02.Feb.2026; accepted 24.Feb.2026; published 12.Mar.2026.

Please cite as:
Berry ECJ, Sculthorpe NF, Warner A, Sanal-Hayes NEM, Hayes LD
Designing an mHealth App to Encourage Uptake of Muscle-Strengthening Exercise in Older Adults: Co-Design Focus Group Study
JMIR Aging 2026;9:e87332
URL: https://aging.jmir.org/2026/1/e87332
doi:10.2196/87332
PMID:

©Ethan CJ Berry, Nicholas F Sculthorpe, Ashley Warner, Nilihan EM Sanal-Hayes, Lawrence D Hayes. Originally published in JMIR Aging (https://aging.jmir.org), 12.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Remote Asynchronous Smartphone 5 Times Sit-to-Stand Testing in Older Adults: Pragmatic Validity Trial

Ethan CJ Berry[1], BSc, MSc; Lawrence D Hayes[2], BSc, MSc, PhD; Ashley Warner[3], BSc, MSc, PhD; Nicholas F Sculthorpe[3], BSc, MSc, PhD

[1]School of Medicine, Medical Sciences & Nutrition, University of Aberdeen, Aberdeen, Scotland, United Kingdom

[2]Lancaster Medical School, Faculty of Health & Medicine, Lancaster University, Lancaster, England, United Kingdom

[3]Department of Sports and Sciences, Faculty of Computing, Sciences and Engineering, University of the West of Scotland, Glasgow, Scotland, United Kingdom

Corresponding Author:
Lawrence D Hayes, BSc, MSc, PhD
Lancaster Medical School
Faculty of Health & Medicine
Lancaster University
Sir John Fisher Drive
Lancaster, England, LA1 4YW
United Kingdom
Phone: 44 7743 061291
Email: l.hayes4@lancaster.ac.uk

Abstract

Background: Age-related declines in muscle function threaten independence and increase the risk of falls in community-dwelling older adults. Muscle-strengthening exercises delivered via mobile health (mHealth) apps could help address access barriers, but scalable programs require valid remote assessments of lower-limb function. The 5 times sit-to-stand (5TSTS) test is a well-established indicator of lower-limb strength and functional capacity, yet its validity in asynchronous, smartphone video formats remains unclear.

Objective: This study aimed to examine validity, intermethod relative agreement, intermethod measurement variability, and convergent validity of a single smartphone-recorded, asynchronously scored 5TSTS.

Methods: A total of 32 older adults aged >60 years attended a single laboratory visit for isokinetic dynamometry of dominant-limb knee extensors and standardized 5TSTS testing and then completed a home-based 5TSTS test after a washout period of at least 3 days, recording the test on their own smartphone. Videos were timed offline by the lead author (EB) using video analysis software. Validity was assessed using a Bland-Altman analysis comparing laboratory and remote 5TSTS tests, while intermethod relative agreement was assessed using the intraclass correlation coefficient (ICC). Intermethod measurement variability was quantified using SEM. Pearson correlations were used to examine associations between body weight–normalized peak torque and both 5TSTS measures to test convergent validity.

Results: A total of 29 participants provided complete laboratory and remote data. Bland-Altman analysis identified that remote 5TSTS times were, on average, 0.3 seconds faster than laboratory values, with 95% limits of agreement of −2.24 to 2.84 seconds, which is promising. Intermethod relative agreement and measurement variability metrics indicated good agreement between laboratory and remote 5TSTS (ICC=0.89; SEM=0.9 seconds). Both laboratory and remote 5TSTS showed moderate negative correlations with normalized knee extensor torque, indicating that greater strength was associated with faster performance in both conditions.

Conclusions: A single asynchronous smartphone video 5TSTS demonstrates promising pragmatic validity, relative agreement with laboratory assessment, and moderate convergent validity with isokinetic knee extensor strength. This could potentially support its use as a future pragmatic remote tool for monitoring progress within mHealth muscle-strengthening programs for community-dwelling older adults at the group level; however, research on test-retest reliability and responsiveness is required before it can be implemented.

Trial Registration: OSF Registries osf.io/xqt52; https://osf.io/xqt52

XSL•FO
RenderX

(JMIR Aging 2026;9:e108345) doi:10.2196/108345

KEYWORDS

mobile health; mHealth; muscle strengthening; older adults; 5 times sit-to-stand test; remote assessment; aged

Introduction

Background

The 5 times sit-to-stand (5TSTS) test is a performance-based and clinical test that is well standardized and devised by Csuka and McCarty [1]. The main purpose of using the test is to assess lower-limb strength, balance, and postural control in older adults, and it has high reliability [2]. The 5TSTS test is widely used due to its simplicity; minimal equipment requirements; and strong evidence supporting its validity, reliability, and safety [3-5]. The 5TSTS test measures the power of hip and knee extensors and provides information on balance and postural control relevant to fall prevention [3,6]. Its clinical utility lies in its ability to objectively track changes in strength and mobility [7]. Previous research has used this test in older adults in both community and laboratory or clinical settings due to its quick and easy implementation and low time or space requirements [8-11]. The test has been previously validated for predicting adverse health outcomes in community-dwelling older adults, with 1 study finding times greater than 10 seconds were associated with increased risk of developing disability [12]. van Cappellen-van Maldegem et al [13] highlighted the 5TSTS test's ability to act as a screening tool for frailty in clinical populations (intraclass correlation coefficient [ICC]=0.70; SEM=2.2; minimally important clinical difference (MID)=3.8). Trivedi et al [14] demonstrated the 5TSTS test as responsive to exercise-based interventions, highlighting its utility for tracking improvements. Moreover, previous work has highlighted a minimal detectable change at the 95% CI (MDC95) of 3.12 seconds in older adults, indicating the magnitude of change required to exceed measurement error [15].

Steffens et al [5] demonstrated that remote delivery of the 5TSTS test via Zoom was feasible, reliable, and safe in 37 patients with cancer preparing for surgery. In older adults, van Schooten et al [16] similarly reported good agreement between face-to-face and remote (Zoom-based) 5TSTS assessments, highlighting their suitability for large-scale remote muscle strength testing. Within mobile health (mHealth), the PowerFrail app developed by Alcazar et al [17] enables remote sit-to-stand assessment to estimate relative muscle power and strength and assess frailty. The mobile app estimates relative muscle power using a validated equation and the user's anthropometric data. The app can work using any sit-to-stand test, including both the 5TSTS and 30-second sit-to-stand tests. It can take the result either from direct entry or video recording, compare it with normative values to identify low power, and flag which features may need improvement. It also assesses frailty using the validated Frailty Trait Scale—Short Form, which produces a frailty score from 0 to 50 and provides tailored physical activity recommendations. The app saves data and generates a report, making it suitable for clinical, research, and public health settings. Other studies have used smartphones attached to the back or chest to capture accelerometry data during the 5TSTS test [18,19]. Moreover, Marques et al [20] used a smartphone app to quantify the validity and reliability of sit-to-stand performance among older adults living in care. Taken together, existing approaches rely either on synchronous videoconferencing with live supervision or on additional hardware or attachment requirements, and tools such as PowerFrail have been designed primarily for classification of low relative muscle power and frailty, rather than for monitoring strength adaptations within exercise interventions in community-dwelling older adults. Moreover, research using smartphone apps to quantify the validity and reliability of sit-to-stand performance has focused on older adults living in care, limiting generalizability to community settings.

There remains limited evidence on the validity and agreement between laboratory 5TSTS assessment and asynchronous smartphone video 5TSTS measurement in community-dwelling older adults. Establishing this is an essential first step toward simple, unsupervised, app-enabled monitoring of functional performance in remote smartphone-delivered muscle-strengthening programs. This will help move toward future automated 5TSTS assessment. We were recently involved in co-designing an mHealth tool to help older adults perform muscle-strengthening exercises at home [21]. The app includes an integrated remote strength and functional assessment using smartphone video to record the 5TSTS test, which is manually but asynchronously scored by a researcher. This provides a practical, at-home indicator of lower-limb strength and function. While the app is under development, it is prudent to assess the validity of smartphone camera–based remote, asynchronous 5TSTS assessment.

Aims and Objectives

The overarching aim of this study was to pragmatically assess the validity and relative agreement between remote 5TSTS performance measured via smartphone camera and laboratory-based assessment using Bland-Altman analysis. By "pragmatically," we mean a single 5TSTS recording to reduce user burden. The following objectives were addressed: evaluation of the validity of remote asynchronous 5TSTS assessment from a single trial and evaluation of the correlation between remote and laboratory-based assessments of 5TSTS and knee extensor strength assessed via isokinetic dynamometry.

Methods

Study Preregistration

The study was preregistered on OSF [22], which included full details on study design, approach, and data analysis plan prior to the commencement of recruitment and data collection. The raw data can be found in the associated project linked to the OSF registration. The study was reported in line with the Guidelines for Reporting Reliability and Agreement Studies [23]. This 15-item checklist covers essential reporting

XSL•FO
RenderX

requirements for the Introduction, Methods, Results, and Discussion sections.

Ethical Considerations

Ethics approval was obtained from the University of the West of Scotland's Health and Life Sciences Academic Integrity and Ethics Committee (23548). The study was conducted in accordance with the Declaration of Helsinki (2024 version) issued by the World Medical Association.

Study Design

This study was a cross-sectional study evaluating the validity, intermethod relative agreement, intermethod measurement validity, and convergent validity of laboratory and remote 5TSTS tests. The primary outcome was validity, quantified using Bland-Altman analysis. The secondary outcomes were intermethod relative agreement (by ICC) and intermethod measurement variability (SEM). The study's tertiary outcome was convergent validity, assessed through correlations between remote and laboratory-based 5TSTS tests and knee extensor strength measured using isokinetic dynamometry.

Participants and Recruitment Strategy

Participants were recruited through local organizations (eg, South and North Lanarkshire councils, the University of the Third Age, Men's Sheds, and Age Scotland), which circulated study information to eligible older adults, and via the social media of the lead author (EB) and posters with a QR code linking to a Microsoft Forms sign-up page. Interested individuals could leave contact details, and some participants also recruited peers by word of mouth through community groups and sports clubs.

Inclusion criteria were as follows: community-dwelling adults aged ≥60 years, with access to a mobile device, and no serious musculoskeletal injury preventing muscle-strengthening exercise. Potential participants received a detailed information sheet and, if willing to take part, provided informed consent, provided demographic data, and completed a Physical Activity Readiness Questionnaire; any contraindications or relevant injuries led to exclusion, and new health issues during the study required medical clearance.

All participants attended a single laboratory at the UWS Lanarkshire Campus Physiotherapy Laboratory (June-August 2025) for measurements and procedures. After a washout period of at least 3 days, they completed the remote part of data collection at home, using their smartphone to video-record themselves performing a functional capacity test. This conservative period was selected to mitigate the learning effects of the laboratory effort. As reported in the literature, older adults do not demonstrate a learning effect on the 5TSTS test with repeated attempts [24]. Participants received instructions in the laboratory to minimize their burden.

Procedures

Preparation and Warm-Up—Laboratory Visit

Prior to all the following procedures, participants were asked not to consume caffeine or alcohol and not to take part in vigorous physical exertion 24 hours prior to the testing for both the laboratory and remote arms of data collection. They complied according to self-report. Participants were advised to wear appropriate clothing for the testing (ie, loose-fitting clothing and shorts). Upon entry to the laboratory, participants were greeted by the lead author (EB), given an outline of what would be expected of them, and asked to read and complete all necessary paperwork if not already completed. Participants then provided demographic data, including age, height, body mass, BMI, blood pressure, sex, ethnicity, nationality, occupation, and education. Height, body mass, and blood pressure were measured in the laboratory as a means of profiling the participants.

A standardized warm-up was completed by each participant. The warm-up consisted of 5 minutes on a cycle ergometer (Wattbike Pro; Wattbike) with the fan resistance progressively increased over time to achieve 100 W and 60 r/min, followed by dynamic stretches of the major muscle groups performed on a matted area. A rest period of at least >5 minutes was given to participants between each of the laboratory assessments to mitigate adverse effects on performance in the tests. Isokinetic dynamometry was performed first, as it was anticipated to be the most physically demanding procedure for participants.

Isokinetic Dynamometry

Dominant-limb knee extensor strength was assessed using isokinetic dynamometry (HUMAC NORM; CSMi). The isokinetic dynamometer (IKD) was calibrated and set up in a standardized position [25-27], and participant positioning was carefully aligned to isolate the quadriceps during testing. Participants sat with the hip flexed at 90°, the knee aligned to the IKD axis, and the lower leg and trunk secured with straps to minimize compensatory movement. The test was performed at an angular velocity of 60°/s, with 3 submaximal familiarization repetitions (20%, 40%, and 60% effort), followed by a 1-minute rest, and then 1 set of 5 maximal knee extensions on the dominant leg with at least 30-second rest as needed between repetitions. Participants were instructed to extend the knee as hard and fast as possible through full range, with verbal encouragement provided. Body weight–normalized torque (N m/kg) was recorded, with the best of the 5 trials used for analysis, and testing was stopped if pain, instability, or significant fatigue occurred.

5TSTS Test

Laboratory

Participants were given a further 5-minute warm-up after their rest time upon completion of the dynamometry testing. The warm-up included 3 main phases. First, participants engaged in a general warm-up, which included marching on the spot, standing leg swings, and ankle rotations. Next, participants were advised to complete seated knee extensions and body weight squats. Finally, participants practiced the sit-to-stand test using slow, controlled movements.

A standard chair (46 cm high) was used for the test. It was placed on a flat, stable surface and against a wall to avoid sliding during the test. Participants were given the opportunity to complete further practice repetitions. Participants were then instructed to sit with their feet flat on the floor, back against the

XSL•FO
RenderX

chair, arms crossed over their chests, and knees at 90° flexion. On researcher command, the participants rose to full standing position and then sat back down on the chair. This movement was completed 5 times as quickly as possible without deviation from the technique. Two full trials were completed with a 1-minute rest period between trials, and the mean time of both trials was used to score the test; time was recorded using a stopwatch. If participants fell or lost balance during the test, the test was voided and restarted after a rest period [5,28].

Remote

Participants were asked to replicate the abovementioned protocol as closely as possible in their own homes, completing a single trial. Verbal instructions were reinforced to participants to ensure they were clear on how to safely and effectively complete the test. Instruction was given to participants to ensure they found a suitable, spacious, and hazard-free area in the home to complete the test and to use a chair of the same height as the one used in the laboratory, positioned against a wall to prevent it from sliding during the test. Although the chair height was standardized in the laboratory, this could not be controlled in participants' homes. Participants were therefore instructed to use a stable chair of similar height, where possible, aiming for 90° knee flexion, as was the case in the laboratory, reflecting the pragmatic conditions under which the app would be used in real-world settings. This approach prioritized feasibility and ecological validity while reducing participant burden, for example, by eliminating the need to measure chair height. Similar approaches have been used in previous protocols [5]. Participants were also encouraged to have someone present while performing the test in case of an injury. Participants were asked to start recording on their mobile devices prior to completing the test to ensure the whole test was recorded. Participants were asked to ensure that their phone cameras were set to 30 frames per second, as is typically the smartphone default, to send this file to the research team. Video files were scored asynchronously, remotely, and manually by the lead author (EB) using the DaVinci Resolve video analysis software (Blackmagic Design Pty Ltd) [29]. EB was not blinded to the laboratory-based 5TSTS scores because he conducted the in-person testing. The score by EB was taken as the ground truth for the remote 5TSTS value for comparison with laboratory testing, but to test EB's scoring, a second rater (LH) also scored all video files. LH was blinded to the scores by EB and to the laboratory scores. The inter-rater agreement for video files was 0.94 (95% CI 0.88-0.97), indicating excellent inter-rater relative agreement for the remote manual scoring of 5TSTS.

Statistical Analysis

Data were recorded in the laboratory using a Microsoft 365 Excel spreadsheet on a secure password-protected device. Data on knee extensor strength were automatically saved on the IKD. All data analysis and visualization were completed in RStudio (version 2026.01.0+392; Posit PBC) [30]. Descriptive statistics were produced, including mean (SD), to characterize the included sample. To test validity, the agreement between remote and laboratory-based 5TSTS was assessed using Bland-Altman analysis. This included the mean bias between both measures and the limits of agreement around the mean bias. The maximum clinically acceptable disagreement between laboratory and remote assessments was not prespecified because established thresholds for this specific study were scarce. Therefore, agreement between measures was assessed in conjunction with previously reported 5TSTS literature. Intermethod relative agreement was assessed via 2-way random-effects, absolute-agreement ICC(2,1). A value <0.50 was interpreted as poor, 0.50-0.75 as moderate, 0.75-0.90 as good, and ≥0.90 as excellent. Moreover, intermethod SEM was reported to express intermethod measurement variability. Two-way random-effects, absolute-agreement ICC(2,1) was used to assess relative inter-rater agreement after a second rater (LH) carried out a blind scoring of the remote 5TSTS video files. Pearson correlation (r) quantified the strength and direction of linear associations between body weight–normalized IKD torque and each 5TSTS measure (IKD vs remote 5TSTS and IKD vs laboratory 5TSTS) to express convergent validity. Where data met assumptions of linearity and normality (assessed using Shapiro-Wilk tests), Pearson r was used; otherwise, the Spearman rank correlation was applied. Correlations were interpreted as r<0.3 (weak), 0.3-0.5 (moderate), and >0.5 (strong), with statistical significance set at P<.05.

An a priori sample size calculation was performed in RStudio (version 2026.01.0+392) for the primary outcome of agreement between the remote and laboratory 5TSTS tests using Bland-Altman analysis. Sample size was determined based on the precision of the estimated 95% limits of agreement. Assuming an SD of paired differences of 3 seconds [5], a sample of 26 participants was required to estimate each limit of agreement, with an approximate 95% CI half-width of ≤2 seconds. This precision-based approach followed recommendations for the design of method comparison studies [31].

Results

Overview

A total of 40 participants were recruited to take part in the study. Of these, 8 (20%) initially agreed but declined to participate due to practicality and scheduling issues that prevented them from attending both the laboratory and remote arms of the trial. A total of 32 (80%) participants attended at least the laboratory-based elements of data collection. However, 1 (2.5%) participant had medical issues in the time between the laboratory visit and remote data collection, which were unrelated to any aspects of the data collection, and therefore could not participate in the remote aspect of the study. A further 2 (5%) participants dropped out after attending the laboratory due to practicality and scheduling issues that prevented them from being able to participate in the remote data collection. Thus, 29 participants were included in the final analysis. Participant characteristics and physical performance are reported in Tables 1 and 2, respectively.

Of the 2 participants who did not continue participation between the laboratory and remote arms of data collection, 1 had technical issues that prevented them from recording the 5TSTS. No further technical issues were reported by participants, and all participants were able to record and export their 5TSTS video

XSL•FO
RenderX

for analysis. Of the 29 participants included, 17 (59%) were Android (Google) device users, and 12 (41%) were iOS (Apple) device users.

Table 1. Demographic characteristics of included participants (N=29).

Characteristics	Participants
Sex, n (%)	
Female	17 (59)
Male	12 (41)
Age (y), mean (SD)	66.9 (6.0)
Body mass (kg), mean (SD)	81.6 (17.9)
Stature (cm), mean (SD)	165.1 (8.5)
BMI (kg/m^2), mean (SD)	29.9 (5.6)
Blood pressure, mean (SD)	
Systolic	147 (17)
Diastolic	81 (11)

Table 2. Physical performance measures and results of laboratory-based and remote 5TSTS assessments and normalized peak torque.

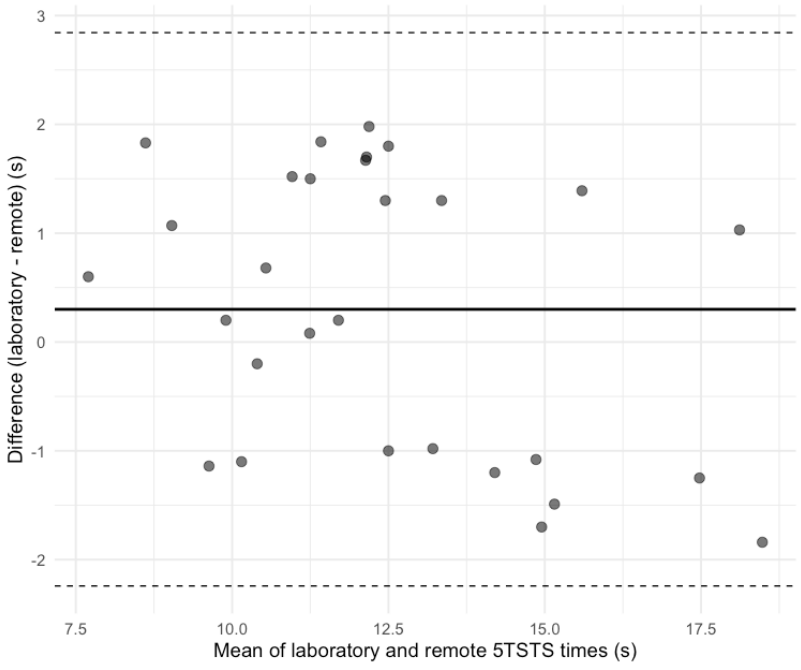
Variables	Values
Laboratory 5TSTS[a] (s), mean (SD)	12.6 (2.6)
Remote 5TSTS (s), mean (SD)	12.3 (3.0)
Normalized peak torque (N m/kg), mean (SD)	1.4 (0.4)

[a]5TSTS: 5 times sit-to-stand.

Validity

Bland-Altman analysis showed that the mean difference (bias) between laboratory-based and remote 5TSTS was 0.3 (95% CI −0.19 to 0.76) seconds, indicating that remote assessment times were, on average, slightly shorter than laboratory-based assessment times. The 95% limits of agreement ranged from −2.24 (95% CI −3.08 to −1.38) seconds to 2.84 (95% CI 1.99 to 3.67) seconds. Figure 1 highlights that there appears to be no obvious visual evidence of proportional bias at low or high values (ie, the spread around the mean bias is consistent across measured values).

Figure 1. Bland-Altman plot for laboratory-based vs remote 5 times sit-to-stand (5TSTS) assessment.

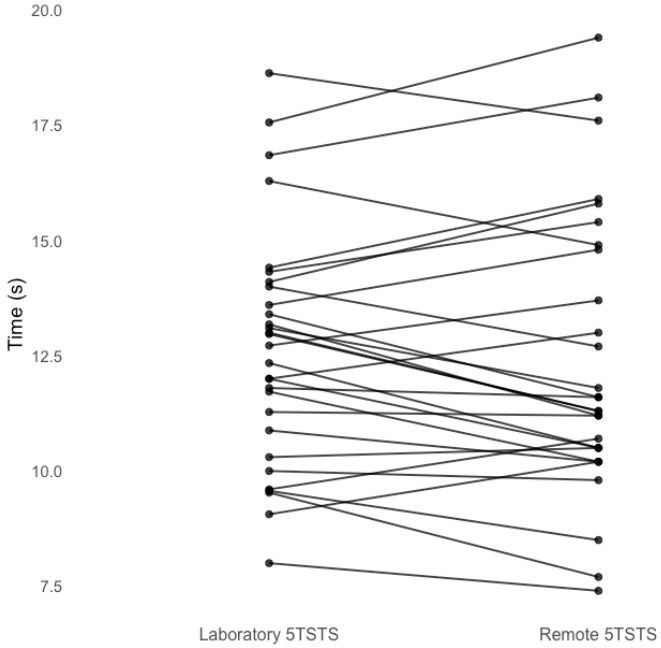


XSL•FO
RenderX

Intermethod Relative Agreement and Intermethod Measurement Variability

The ICC was 0.89 (95% CI 0.81-0.94), indicating good intermethod relative agreement between the remote 5TSTS measurement and the laboratory reference. The SEM was 0.9 seconds, corresponding to an intermethod measurement variability of less than 1 second. These metrics collectively indicate strong relative and absolute agreement between remote and laboratory-based 5TSTS measurements, while also reflecting measurement variability (Figure 2).

Figure 2. Individual-level comparison of laboratory-based and remote 5 times sit-to-stand (5TSTS) performance.



Convergent Validity Correlative Analysis

Pearson correlations revealed a moderate negative relationship between normalized peak torque and laboratory 5TSTS time and between normalized peak torque and remote 5TSTS time, indicating that higher knee extensor strength was associated with faster 5TSTS performance (Table 3). These correlation coefficients were similar in magnitude.

Table 3. Pearson correlation matrix between laboratory and remote 5TSTS[a] (seconds) and normalized peak torque (N m/kg).

	Laboratory 5TSTS	Remote 5TSTS	Normalized peak torque
Laboratory 5TSTS			
r	—[b]	—	—
P value	—	—	—
Remote 5TSTS			
r	0.907	—	—
P value	<.001	—	—
Normalized peak torque			
r	−0.519	−0.518	—
P value	0.004	0.004	—

[a]5TSTS: 5 times sit-to-stand.

[b]Not applicable.

Discussion

Main Findings

The objectives of this study were to evaluate the validity of remote asynchronous 5TSTS assessment from a single trial. An additional objective was to assess the convergent validity of remote and laboratory-based 5TSTS with knee extensor strength. The current study highlighted that both remote and laboratory-based 5TSTS assessments showed promising pragmatic agreement, indicated by a small mean bias and promising limits of agreement of <3 seconds. We interpret these findings to potentially support remote 5TSTS as a valid means for evaluating lower-limb muscle strength and functional mobility in community-dwelling older adults at the group level. However, further research on reliability and responsiveness is required. We found good intermethod relative agreement

XSL•FO
RenderX

between the remote smartphone assessment and the laboratory assessment, as indicated by ICC, and good intermethod measurement validity, as indicated by SEM. Both laboratory and remote 5TSTS showed moderate negative correlations with knee extensor torque.

Validity Analysis

The 5TSTS test is a well-established test of lower-limb muscle strength and functional mobility in community-dwelling older adults [7]. In practice, the findings from the current study could indicate that the mean difference from our Bland-Altman analysis is not clinically important [4], meaning the trivial mean bias between the laboratory and remote 5TSTS observed in this study could potentially support the use of remote smartphone 5TSTS for assessing lower-limb functional mobility in older adults at the group level when further test-retest evaluation has been conducted. The 95% limits of agreement indicated that remote 5TSTS times differed from laboratory-based assessment by −2.24 seconds (lower) and 2.84 seconds (upper). Previous research in older populations has highlighted similar results. For instance, van Schooten et al [16] assessed a variety of physical function tests in laboratory and remote setting (via Zoom) in 37 community-dwelling older adults, reporting lower and upper limits of agreement of more than 2 seconds for the 5TSTS test. Fyfe et al [32] included 19 community-dwelling older adults and assessed the reliability and feasibility of remote physical function testing. They conducted 3 separate tests across 3 videoconferencing calls and highlighted lower and upper limits of agreement of 2.59 and 4.41 seconds between tests 1 and 2, and 3.14 and 3.13 seconds between tests 2 and 3, respectively. These previous findings highlight that agreement between 5TSTS tests in older adults can vary due to several factors, including test familiarity and physical demands [32]. As such, the mean bias and limits of agreement comparing remote and laboratory testing reported within the present study are encouraging.

It must be noted that our sample included a substantial proportion of participants who exceeded the 12-second clinical cut-off and the 15-second threshold associated with an increased risk of falls [12]. Although a more homogeneous sample may have resulted in narrower limits of agreement, it would also have introduced spectrum bias and reduced the clinical relevance of the findings. In practice, the purpose of the test is to identify functional status without prior knowledge of an individual's risk profile; therefore, the heterogeneity of the sample likely widened the limits of agreement but also improved the ecological validity and representativeness of the study.

It should also be noted that, in this protocol, the laboratory-based 5TSTS time was calculated as the mean of 2 attempts, whereas only 1 attempt was scored for the remote 5TSTS assessment to minimize the technical burden on participants associated with recording multiple attempts. Therefore, the reduced measurement variability resulting from averaging 2 attempts in the laboratory may have contributed to the observed bias and limits of agreement.

The findings of the current study are consistent with those of previous studies that used videoconferencing software to remotely assess 5TSTS performance [16,32]. However, the current study helps advance the previous literature by introducing an asynchronous smartphone-based alternative to live videoconferencing assessment, therefore not requiring a researcher at that moment and reducing the resources required for remote assessment. As mentioned, previous apps have used remote 5TSTS to estimate relative muscle power, assess frailty, and store results, showing that smartphone-based 5TSTS testing is clinically feasible. Other studies have attached smartphones to the trunk to collect accelerometer data during 5TSTS, but these methods are technically demanding and less scalable.

Intermethod Relative Agreement and Intermethod Measurement Variability

The intermethod relative agreement data highlighted that the remote measurement of the 5TSTS test showed a moderate to strong agreement with the laboratory assessment. The ICC of 0.89 indicates good intermethod relative agreement [7]. This finding sits at the upper end of the threshold of good intermethod agreement and is close to excellent, according to widely used criteria [33]. The results of the current study indicate an SEM of approximately 0.9 seconds; however, this reflects intermethod measurement variability between laboratory-based and remote 5TSTS assessments and should not be interpreted as remote test-retest error, which requires further investigation. The intermethod relative agreement and measurement variability results in the current study suggest that remote 5TSTS could potentially be suitable for several applications. Measurement theory emphasizes that a specific measure does not need to be perfect to be useful but must be sufficiently stable for the decisions being made from it [34]. In this context, the good relative intermethod relative agreement findings potentially support the utility of remote 5TSTS for evaluating group-level responses to muscle-strengthening interventions in older adults [35]. However, it is important to note that research on reliability and responsiveness is required before this can be applied. This is relevant for mHealth apps and interventions, where repeated remote monitoring may be used to follow progress across a 12- to 16-week program or to compare average changes between groups [36]. As the present study did not examine repeated remote assessments, the measurement error and MDC95 are still to be established. Further research using repeated remote assessments is required to establish test-retest reliability and ability to detect meaningful individual changes.

Convergent Validity

The correlations between knee extensor strength and 5TSTS performance were moderate and negative for both the laboratory and remote 5TSTS assessments (laboratory: r=−0.519; remote: r=−0.518). These results indicate that participants with greater quadriceps torque tended to complete both 5TSTS tests more quickly [37]. This magnitude of association is consistent with the concept that the 5TSTS test is a functional proxy for lower-limb muscle strength in older adults rather than a "pure" strength measure. The performance of the 5TSTS test in older adults reflects force production but also other factors such as balance, coordination, movement strategy, and familiarity or confidence [38]. Previous work by Yee et al [10] in community-dwelling older adults confirmed that the performance of the 5TSTS test is a functional maneuver that

XSL•FO
RenderX

relies in part on the strength of the knee extensors and also on the stability of the trunk, sensation, and balance. Ruiz-Cárdenas et al [39] validated a smartphone sit-to-stand app as a diagnostic tool for sarcopenia in older adults. This work considered the velocity of hip extension as a key factor in the performance of the 5TSTS, indicating that while the ability of the knee extensor muscles to produce required force is essential, their ability to produce this force quickly (power) is also a key determinant [40]. In relation to the study's current findings, the moderate strength of correlations suggests that knee extensor strength makes an important but not exclusive contribution to 5TSTS performance in community-dwelling older adults. Encouragingly, the remote assessment *r* was analogous to the laboratory assessment *r*.

Strengths and Limitations

The current study exhibits several strengths, including preregistration, which enhances transparency. Additionally, the precision-based sample size calculation undertaken a priori ensures our conclusions are robust. The sample therefore provides adequate precision of the Bland-Altman limits of agreement, which is consistent with recommendations for method comparison studies [31]. In addition, we recruited a sample exceeding the minimum number, which further supports the adequacy of the study for its primary analysis.

An additional important strength is that participants performed the 5TSTS test at home, which may have reduced apprehension and encouraged a more natural testing environment where the participant could proceed at their own pace without external pressures [25]. The laboratory setting may be perceived as more formal and more intimidating [41]. Thus, the remote 5TSTS results may be considered more ecologically valid.

Every participant in the sample was considered an older adult (aged ≥60 years), with an age range from 60 to 80 years. While this could result in differing physical abilities across these ages, widening the limits of agreement, this enhances the validity and applicability of the present study, as opposed to examining agreement in younger adults. Our participants differed in exercise training experience, which could have led to differences in strength and power, balance and postural control, joint pain and stiffness, and confidence [28]. When conditions of the tests were changed (laboratory vs remote), the aforementioned factors may be exacerbated. For example, a trained 60-year-old individual may perform the 5TSTS test almost identically in both laboratory and remote settings; however, a frail 80-year-old individual with marginal balance might be more cautious in one setting, in this case the laboratory (unknown environment), but not at home, where they were comfortable [41]. We consider the rigorous assessment of agreement between laboratory-based and remote 5TSTS assessments a strength of the present study.

The study was not without limitations. First, as the data collection required both laboratory and remote 5TSTS assessments, there was an inherent variability between the laboratory environment and participants' homes. Participants received clear instructions on completing and recording the remote 5TSTS assessment to replicate the laboratory conditions as closely as possible. Because this was a pragmatic evaluation of real-world 5TSTS scores, chair height, chair characteristics, floor stability, number of trials, and timing procedures were not analogous to laboratory testing. However, this was necessary to collect data from participants in real-world settings. Therefore, the remote test should be interpreted as a pragmatic home-based measure, not a standardized laboratory equivalent. Most notably, the laboratory assessment was the mean of 2 trials, and the remote assessment was only 1 trial, which may contribute to between-method variability. That said, the approach reflected real-world conditions in which the app will be deployed. Requiring participants to use a chair of a fixed height at home or to source, measure, and precisely adjust a chair to a specified height would increase participant burden, reduce adherence, and be impractical for older adults.

Importantly, recruitment potential was limited by travel burden, as participants had to travel to the university campus to complete the laboratory-based data collection. This limited recruitment to those who lived near the campus or had the means to access it, meaning the sample may not be entirely representative. However, results include scores ranging between good and below thresholds, suggesting a range of abilities.

It must be noted that laboratory assessment always preceded remote assessment; order and familiarization effects cannot be excluded. The small tendency toward faster remote effects may therefore reflect learning, environmental comfort, or other environmental factors rather than the protocol alone, and future studies should consider counterbalancing the order of assessments.

It is important to note that when interpreting the findings of this study, the mean age of the sample was 66.9 (6.0) years. While age ranged from 60 to 80 years, participants were community dwelling and mainly healthy. Therefore, the findings may not apply to older adults aged >80 years or to those with frailty, chronic conditions, and limited digital experience. As such, this should be a key focus of future research.

Finally, the remote arm of the data collection required manual scoring for the 5TSTS tests. Although manageable for a sample of 29 participants, this approach may impose a substantial researcher burden in studies involving larger sample sizes.

Future Research

Given the current study's promising findings, there are avenues for future research to build on these findings. First, research is needed to explore the automation of remote 5TSTS assessment in a cost-effective manner. One way to do this is to investigate AI for video analysis in remote assessment. This would permit 5TSTS to be automatically scored and categorized upon uploading, reducing researcher burden and resulting in real-time feedback. Furthermore, larger scale mHealth muscle-strengthening interventions are needed to provide longitudinal data.

As mentioned, future research in this area should also seek to address the usability of remote 5TSTS assessment to ensure that this strength assessment type does not lead to digital exclusion. Future trials may benefit from investigating usability and acceptability of smartphone-based 5TSTS assessment in older adults with lower digital literacy. One avenue that has proven successful when it comes to digital tools in this

XSL•FO
RenderX

population is co-design, and well-designed trials should inform adaptations to digital platforms to ensure a smooth, user-designed platform [21]. In turn, the above future direction would help position remote, smartphone-based strength assessment as a core component of a broader digital strategy for the support and promotion of muscle-strengthening exercise in older adults.

Conclusions

The current study demonstrates that remote smartphone-based 5TSTS assessment provides promising pragmatic agreement with laboratory assessment under controlled conditions in community-dwelling older adults, supporting its potential future application for group-level evaluation once future reliability and responsiveness research has been conducted. Our data suggest that remote 5TSTS performance provides comparable estimates of lower-limb functional performance to laboratory assessment and preserves the observed relationship between 5TSTS performance and lower-limb strength. These findings could potentially support remote 5TSTS assessment as a practical and scalable alternative to laboratory-based assessment for monitoring lower-limb function in older adults at the group level in the future. However, as mentioned, future research is required to confirm test-retest reliability and responsiveness before this conclusion can be applied. The study has highlighted good intermethod relative agreement but has not yet established interchangeability for monitoring change at the individual level. As repeated remote assessments were not examined, further research is required to establish test-retest reliability, responsiveness to longitudinal change, and the ability of remote 5TSTS assessment to detect meaningful individual changes over time.

Data Availability

The dataset analyzed during this study is available in the OSF repository [22].

Funding

The authors declared no financial support was received for this work.

Generative AI Use

The authors declare that generative AI was not used in the preparation of this manuscript.

Authors' Contributions

Conceptualization: ECJB (lead), LDH (equal), AW (equal), NFS (equal)
Data curation: ECJB
Formal analysis: ECJB (lead), LDH (supporting), AW (supporting)
Funding acquisition: LDH
Investigation: ECJB
Methodology: ECJB, LDH, AW, NFS
Project administration: ECJB
Resources: ECJB
Supervision: LDH (equal), AW (equal), NFS (equal)
Validation: ECJB
Visualization: ECJB (lead), LDH (supporting), AW (supporting), NFS (supporting)
Writing – original draft: ECJB (lead), LDH (supporting), AW (supporting), NFS (supporting)
Writing – review & editing: ECJB (lead), LDH (supporting)

Conflicts of Interest

In the interest of complete transparency, the lead author (EB) was involved in the creation, design, and development of the Strength Buddy mobile app as part of their wider PhD project alongside the other authors of this manuscript. The app exists as a test prototype and currently has no commercial value. The submitted work was not carried out in the presence of any personal, professional, or financial relationships that could potentially be construed as conflicts of interest.

References

1. Csuka M, McCarty DJ. Simple method for measurement of lower extremity muscle strength. Am J Med 1985 Jan;78(1):77-81. [doi: 10.1016/0002-9343(85)90465-6] [Medline: 3966492]
2. Lord SR, Murray SM, Chapman K, Munro B, Tiedemann A. Sit-to-stand performance depends on sensation, speed, balance, and psychological status in addition to strength in older people. J Gerontol A Biol Sci Med Sci 2002 Aug;57(8):M539-M543. [doi: 10.1093/gerona/57.8.m539] [Medline: 12145369]
3. De Melo TA, Silva Guimarães F, Lapa E Silva JR. The five times sit-to-stand test: safety, validity and reliability with critical care survivors's at ICU discharge. Arch Physiother 2022 Dec 18;13(1):2 [FREE Full text] [doi: 10.1186/s40945-022-00156-z] [Medline: 36597159]

XSL•FO
RenderX

4. Albalwi AA, Alharbi AA. Optimal procedure and characteristics in using five times sit to stand test among older adults: a systematic review. Medicine (Baltimore) 2023 Jun 30;102(26):e34160 [FREE Full text] [doi: 10.1097/MD.0000000000034160] [Medline: 37390277]
5. Steffens D, Pocovi NC, Bartyn J, Delbaere K, Hancock MJ, Koh C, et al. Feasibility, reliability, and safety of remote five times sit to stand test in patients with gastrointestinal cancer. Cancers (Basel) 2023 Apr 24;15(9):2434 [FREE Full text] [doi: 10.3390/cancers15092434] [Medline: 37173899]
6. Porto JM, Peres-Ueno MJ, de Matos Brunelli Braghin R, Scudilio GM, de Abreu DC. Diagnostic accuracy of the five times stand-to-sit test for the screening of global muscle weakness in community-dwelling older women. Exp Gerontol 2023 Jan;171:112027 [FREE Full text] [doi: 10.1016/j.exger.2022.112027] [Medline: 36395978]
7. Muñoz-Bermejo L, Adsuar JC, Mendoza-Muñoz M, Barrios-Fernández S, Garcia-Gordillo MA, Pérez-Gómez J, et al. Test-retest reliability of five times sit to stand test (FTSST) in adults: a systematic review and meta-analysis. Biology (Basel) 2021 Jun 09;10(6):510 [FREE Full text] [doi: 10.3390/biology10060510] [Medline: 34207604]
8. Hayes LD, Elliott BT, Yasar Z, Bampouras TM, Sculthorpe NF, Sanal-Hayes NE, et al. High intensity interval training (HIIT) as a potential countermeasure for phenotypic characteristics of sarcopenia: a scoping review. Front Physiol 2021 Aug 24;12:715044 [FREE Full text] [doi: 10.3389/fphys.2021.715044] [Medline: 34504439]
9. Alcazar J, Losa-Reyna J, Rodriguez-Lopez C, Alfaro-Acha A, Rodriguez-Mañas L, Ara I, et al. The sit-to-stand muscle power test: an easy, inexpensive and portable procedure to assess muscle power in older people. Exp Gerontol 2018 Oct 02;112:38-43. [doi: 10.1016/j.exger.2018.08.006] [Medline: 30179662]
10. Yee XS, Ng YS, Allen JC, Latib A, Tay EL, Abu Bakar HM, et al. Performance on sit-to-stand tests in relation to measures of functional fitness and sarcopenia diagnosis in community-dwelling older adults. Eur Rev Aging Phys Act 2021 Jan 08;18(1):1 [FREE Full text] [doi: 10.1186/s11556-020-00255-5] [Medline: 33419399]
11. Søgaard D, Lund MT, Scheuer CM, Dehlbaek MS, Dideriksen SG, Abildskov CV, et al. High-intensity interval training improves insulin sensitivity in older individuals. Acta Physiol (Oxf) 2018 Apr;222(4):e13009. [doi: 10.1111/apha.13009] [Medline: 29197155]
12. Makizako H, Shimada H, Doi T, Tsutsumimoto K, Nakakubo S, Hotta R, et al. Predictive cutoff values of the five-times sit-to-stand test and the timed "up and go" test for disability incidence in older people dwelling in the community. Phys Ther 2017 Apr 01;97(4):417-424 [FREE Full text] [doi: 10.2522/ptj.20150665] [Medline: 28371914]
13. van Cappellen-van Maldegem SJ, Hoedjes M, Seidell JC, van de Poll-Franse LV, Buffart LM, Mols F, et al. Self-performed five times sit-to-stand test at home as (pre-)screening tool for frailty in cancer survivors: reliability and agreement assessment. J Clin Nurs 2023 Apr;32(7-8):1370-1380. [doi: 10.1111/jocn.16299] [Medline: 35332600]
14. Trivedi P, Patel S, Edwards G, Jenkins T, Man WD, Nolan CM. Five-repetition sit-to-stand test: responsiveness and minimal important difference in idiopathic pulmonary fibrosis. Ann Am Thorac Soc 2024 Apr;21(4):577-584 [FREE Full text] [doi: 10.1513/AnnalsATS.202306-561OC] [Medline: 37847730]
15. Puthoff ML, Saskowski D. Reliability and responsiveness of gait speed, five times sit to stand, and hand grip strength for patients in cardiac rehabilitation. Cardiopulm Phys Ther J 2013 Mar;24(1):31-37 [FREE Full text] [Medline: 23754937]
16. van Schooten KS, Steffens D, Engeler A, Letton M, Brown A, Perram A, et al. Remote assessment of physical function in older people: feasibility, safety and agreement with in-person administration. Age Ageing 2025 Aug 29;54(9):afaf266 [FREE Full text] [doi: 10.1093/ageing/afaf266] [Medline: 41014506]
17. Alcazar J, Lopez-Moranchel I, Losa-Reyna J, Baltasar-Fernandez I, Garcia-Garcia FJ, Alegre LM, et al. App Review: PowerFrail for the assessment of relative muscle power and frailty in older people. BJSM Blog. 2021. URL: https://blogs.bmj.com/bjsm/2021/02/04/app-review-powerfrail-for-the-assessment-of-relative-muscle-power-and-frailty-in-older-people/#_ENREF_2 [accessed 2026-11-13]
18. Cerrito A, Bichsel L, Radlinger L, Schmid S. Reliability and validity of a smartphone-based application for the quantification of the sit-to-stand movement in healthy seniors. Gait Posture 2015 Feb;41(2):409-413. [doi: 10.1016/j.gaitpost.2014.11.001] [Medline: 25467428]
19. Hayek R, Gutman I, Baranes G, Nudelman Y, Springer S. Smartphone-based sit-to-stand analysis for mobility assessment in middle age. Innov Aging 2024 Aug 28;8(10):igae079 [FREE Full text] [doi: 10.1093/geroni/igae079] [Medline: 39391811]
20. Marques DL, Neiva HP, Pires IM, Zdravevski E, Mihajlov M, Garcia NM, et al. An experimental study on the validity and reliability of a smartphone application to acquire temporal variables during the single sit-to-stand test with older adults. Sensors (Basel) 2021 Mar 15;21(6):2050 [FREE Full text] [doi: 10.3390/s21062050] [Medline: 33803927]
21. Berry EC, Sculthorpe NF, Warner A, Sanal-Hayes NE, Hayes LD. Designing an mHealth app to encourage uptake of muscle-strengthening exercise in older adults: co-design focus group study. JMIR Aging 2026 Mar 12;9:e87332 [FREE Full text] [doi: 10.2196/87332] [Medline: 41818741]
22. Berry E, Hayes LD, Sculthorpe N, Warner A. Supporting and tracking muscle strengthening exercise in older adults: a validity and feasibility trial. Open Science Framework. 2025. URL: https://osf.io/xqt52/overview [accessed 2026-11-13]
23. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 2011 Jan;64(1):96-106. [doi: 10.1016/j.jclinepi.2010.03.002] [Medline: 21130355]

XSL•FO
RenderX

24. Vilarinho R, Montes AM, Melo C. Age-related influence on reliability and learning effect in the assessment of lower limb strength using sit-to-stand tests: a cross-sectional study. Health Sci Rep 2024 May 20;7(5):e2064 [FREE Full text] [doi: 10.1002/hsr2.2064] [Medline: 38774592]
25. Hayes LD, Morse CI. The effects of progressive dehydration on strength and power: is there a dose response? Eur J Appl Physiol 2010 Mar;108(4):701-707. [doi: 10.1007/s00421-009-1288-y] [Medline: 19908058]
26. Bloch-Ibenfeldt M, Theil Gates A, Karlog K, Demnitz N, Kjaer M, Boraxbekk CJ. Heavy resistance training at retirement age induces 4-year lasting beneficial effects in muscle strength: a long-term follow-up of an RCT. BMJ Open Sport Exerc Med 2024 Jun 18;10(2):e001899 [FREE Full text] [doi: 10.1136/bmjsem-2024-001899] [Medline: 38911477]
27. Guadalupe-Grau A, Aznar-Laín S, Mañas A, Castellanos J, Alcázar J, Ara I, et al. Short- and long-term effects of concurrent strength and HIIT training in octogenarians with COPD. J Aging Phys Act 2017 Jan;25(1):105-115. [doi: 10.1123/japa.2015-0307] [Medline: 27402660]
28. Hayes LD, Sanal-Hayes NE, Mclaughlin M, Berry EC, Sculthorpe NF. People with long COVID and ME/CFS exhibit similarly impaired balance and physical capacity: a case-case-control study. Am J Med 2025 Jan;138(1):140-147 [FREE Full text] [doi: 10.1016/j.amjmed.2023.06.028] [Medline: 37490948]
29. DaVinci resolve 21. Blackmagic Design. URL: https://www.blackmagicdesign.com/uk/products/davinciresolve/ [accessed 2026-11-13]
30. RStudio IDE user guide. Posit Software. URL: https://posit.co/download/rstudio-desktop/ [accessed 2026-11-13]
31. Monti CB, Ambrogi F, Sardanelli F. Sample size calculation for data reliability and diagnostic performance: a go-to review. Eur Radiol Exp 2024 Jul 05;8(1):79 [FREE Full text] [doi: 10.1186/s41747-024-00474-w] [Medline: 38965128]
32. Fyfe JJ, Sousa F, Watson-Mackie K, Jansons P, Scott D, Daly RM. Remote physical function testing in older adults: a mixed methods study exploring test reliability, feasibility, and perceptions of participants and assessors. PLoS One 2025 Sep 19;20(9):e0332691 [FREE Full text] [doi: 10.1371/journal.pone.0332691] [Medline: 40971978]
33. Benhida M, Zeghari L, El Mouahid K, El Morchidy S, El Madhi Y, Enneya N, et al. Physical performance profiling in Moroccan women's football: a functional assessment from the Rabat-Salé-Kénitra regional league. BMC Sports Sci Med Rehabil 2025 Dec 03;18(1):13 [FREE Full text] [doi: 10.1186/s13102-025-01453-3] [Medline: 41331491]
34. Dimitrov D, Rumrill P, Fitzgerald S, Hennessey M. Reliability in rehabilitation measurement. Work 2001 Mar;16(2):159-164. [doi: 10.3233/wor-2001-00148]
35. Takimoto K, Takebayashi H, Kondo H, Ikeda K. Association between muscle quality assessed by the 5-repetition sit-to-stand test and falls in community-dwelling older adults in Japan: a cross-sectional study. Geriatrics (Basel) 2025 Jun 07;10(3):78 [FREE Full text] [doi: 10.3390/geriatrics10030078] [Medline: 40558617]
36. Jungreitmayr S, Kranzinger C, Venek V, Ring-Dimitriou S. Effects of an app-based physical exercise program on selected parameters of physical fitness of females in retirement: a randomized controlled trial. Front Physiol 2022 Mar 4;13:821773 [FREE Full text] [doi: 10.3389/fphys.2022.821773] [Medline: 35317213]
37. Steinbrink GM, Martinez J, Swartz AM, Strath SJ. Sit-to-stand power is a stronger predictor of gait speed than knee extension strength. J Funct Morphol Kinesiol 2024 Jun 13;9(2):103 [FREE Full text] [doi: 10.3390/jfmk9020103] [Medline: 38921639]
38. Niklasson J, Fagerström C, Backåberg S, Bergman P, Lindberg T. Exploring predictors of the five-time sit-to-stand test based on cross-sectional findings from the Swedish National Study on Aging and Care (SNAC). BMC Geriatr 2025 Feb 04;25(1):79 [FREE Full text] [doi: 10.1186/s12877-025-05737-8] [Medline: 39905293]
39. Ruiz-Cárdenas JD, Montemurro A, Martínez-García MD, Rodríguez-Juan JJ. Sit-to-stand video analysis-based app for diagnosing sarcopenia and its relationship with health-related risk factors and frailty in community-dwelling older adults: diagnostic accuracy study. J Med Internet Res 2023 Dec 08;25:e47873 [FREE Full text] [doi: 10.2196/47873] [Medline: 38064268]
40. Mato L, Chankavee N, Amatachaya S, Thaweewannakij T. Knee extensor muscle strength to measure the ability of five times sit to stand independently in patients with incomplete spinal cord injury. Malays J Med Sci 2022 Oct;29(5):74-82 [FREE Full text] [doi: 10.21315/mjms2022.29.5.8] [Medline: 36474537]
41. Nebeker C, Zlatar ZZ. Learning from older adults to promote independent physical activity using mobile health (mHealth). Front Public Health 2021 Aug 12;9:703910 [FREE Full text] [doi: 10.3389/fpubh.2021.703910] [Medline: 34476229]

Abbreviations

5TSTS: 5 times sit-to-stand
ICC: intraclass correlation coefficient
IKD: isokinetic dynamometer
MDC95: minimal detectable change at the 95% CI
mHealth: mobile health
MID: minimally important clinical difference
OSF: Open Science Framework

XSL•FO
RenderX

Edited by KJ Moon; submitted 30.Jul.2026; peer-reviewed by PR Venna, J Loureiro, B Unver; comments to author 14.Sep.2026; revised version received 10.Nov.2026; accepted 10.Nov.2026; published 23.Nov.2026.

Please cite as:
Berry ECJ, Hayes LD, Warner A, Sculthorpe NF
Remote Asynchronous Smartphone 5 Times Sit-to-Stand Testing in Older Adults: Pragmatic Validity Trial
JMIR Aging 2026;9:e108345
URL: https://aging.jmir.org/2026/1/e108345
doi:10.2196/108345
PMID:

©Ethan CJ Berry, Lawrence D Hayes, Ashley Warner, Nicholas F Sculthorpe. Originally published in JMIR Aging (https://aging.jmir.org), 23.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Egocentric Spatial Memory Deficit in Amnestic Mild Cognitive Impairment Revealed Through Virtual Reality: Cross-Sectional Study

Cosimo Tuena[1,2], PhD; Silvia Serino[3*], PhD; Chiara Stramba-Badiale[2*], PsyD; Diana Biondi[2], PsyD; Karine Marie Goulene[4], MD, PhD; Marco Stramba-Badiale[4], MD, PhD; Giuseppe Riva[2,5], PhD

[1]Department of Theoretical and Applied Sciences, eCampus University, Novedrate, Italy

[2]Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, 2 Alessandro Magnasco, Milan, Italy

[3]Department of Psychology, Università degli Studi di Milano-Bicocca, Milan, Lombardy, Italy

[4]Department of Medicine, Neurology and Rehabilitation, IRCCS Istituto Auxologico Italiano, Milan, Italy

[5]Humane Technology Lab, Università Cattolica del Sacro Cuore, Milan, Italy

[*]these authors contributed equally

Corresponding Author:
Cosimo Tuena, PhD
Department of Theoretical and Applied Sciences, eCampus University, Novedrate, Italy

Abstract

Background: Spatial navigation relies on egocentric and allocentric frames of reference, with the latter critically impaired in Alzheimer disease (AD) due to hippocampal involvement. Recent evidence suggests that egocentric representations may coexist within medial temporal lobe regions; however, their relative impairment in the early stages of AD remains unclear. Virtual reality offers a promising approach to bridge this gap by assessing spatial navigation abilities and providing sensitive digital biomarkers for AD.

Objective: This study investigates spatial memory performance in individuals with amnestic mild cognitive impairment (aMCI) using a virtual reality object-location memory task that manipulated available cues during recall. An environmental boundary probed hippocampal-dependent allocentric processing, while a discrete landmark probed striatal-dependent egocentric strategies.

Methods: Eighty participants (40 individuals with aMCI and 40 healthy controls [HC]) encoded object locations in a virtual arena and recalled them using either landmark or boundary cues.

Results: Regardless of the spatial frame of reference, patients with aMCI demonstrated significantly poorer overall spatial memory than the HC group. Surprisingly, virtual egocentric error emerged as a predictor of aMCI diagnosis. No differences were observed in frame-switching, object-location binding, or across recall trials. Analysis of virtual encoding paths revealed that patients with aMCI exhibited centrifugal navigation away from the discrete landmark and toward the boundary of the arena. During recall, patients with aMCI demonstrated less effective usage of boundary and landmark cues than HC, as revealed by the distribution of response coordinates.

Conclusions: These findings indicate that, in addition to well-documented allocentric deficits, egocentric spatial memory impairments provide a complementary and potentially sensitive marker within the AD continuum. This dual impairment of spatial cognition systems may enhance our understanding of navigational difficulties in aMCI and offers insight into the nature of spatial representations in aging.

(JMIR Aging 2026;9:e79224) doi:10.2196/79224

KEYWORDS

spatial navigation; dementia; mild cognitive impairment; spatial cognition; virtual reality; digital health

Introduction

Spatial navigation and memory represent cornerstone cognitive abilities that underpin our everyday functioning, playing a crucial role across the lifespan. These interconnected abilities enable self-orientation, object localization, wayfinding, and, in general, the maintenance of independence within one's surroundings. Central to these abilities is how spatial information is represented. The environment can be represented using 2 different spatial frames of reference [1,2]. The egocentric frame uses a body-dependent representation (eg, "the chair is to my right"), while the allocentric frame uses a body-independent representation (eg, "the chair is close to the north-facing wall"). Crucially, the flexible switch and use of both spatial frames of

XSL•FO
RenderX

reference sustain successful spatial navigation and memory [3-6].

Research has demonstrated that different types of cues in the environment tend to promote the use of different reference frames [7-9]. Specifically, extended environmental boundaries (such as walls) tend to promote allocentric processing, as they provide stable, viewpoint-independent reference frames. In contrast, discrete landmarks (such as a monument) often promote egocentric processing, as individuals tend to encode and recall target locations in relation to their own position relative to these salient features. Allocentric spatial memory relying on boundaries is mainly sustained by medial temporal regions, prevalently the right hippocampus and boundary cells in the subiculum and entorhinal cortex. In a virtual reality (VR) task, Lee et al [10] showed that encoding target locations close to the boundaries, compared to the center, elicited stronger theta oscillations than for target locations at the center of the environment, suggesting that the human subiculum is a key region associated with integrating boundary information into spatial memory, facilitating precise navigation and spatial recall. Conversely, egocentric spatial memory relying on landmarks is supported by the dorsal striatum and parietal cortex, with the medial prefrontal cortex serving as a critical hub for top-down modulation when both representations are similarly active [7,8,10-12].

As we age, the integrity of these spatial functions becomes increasingly significant, serving as a window into brain health. The literature indicates that spatial navigation and memory decline in normal aging, with older adults exhibiting increased use of egocentric navigation and diminished frame-switching abilities [13-15]. For instance, Schuck et al [16] manipulated geometric boundaries and local landmarks within a VR object-location (OL) memory task. Young adults demonstrated behavior patterns and medial temporal lobe activation consistent with hippocampal boundary-processing models. In contrast, older adults relied predominantly on landmark information, with their spatial learning correlating more strongly with caudate nucleus activity than hippocampal engagement, suggesting an age-related shift from a primarily hippocampal-dependent boundary-processing system to a caudate-mediated, landmark-oriented approach.

In pathological aging associated with Alzheimer disease (AD), this pattern progresses to a more severe impairment, with earlier and more marked deficits in allocentric navigation and eventual deterioration of compensatory egocentric strategies. Consistent with AD-specific neuropathological processes [17], significant allocentric deficits and frame-switching impairments have been consistently reported in AD dementia [18], even in the early stages [19]. Understanding when these deficits emerge is therefore critical for early detection. Mild cognitive impairment (MCI) has been proposed as a transitional stage from normal aging to dementia [20]. Specifically, individuals with objective impairment on memory tests, namely amnestic mild cognitive impairment (aMCI), have a heightened risk of developing AD compared to non-aMCI [21]. Consistent with this progression, individuals with aMCI not only report subjective navigational complaints but also manifest impairments in allocentric and egocentric spatial navigation, as well as in frame-switching abilities, especially from the allocentric to egocentric frame [22-25]. More recent research has begun to better reveal complex patterns of spatial impairment in this specific population. Using a VR boundary-based OL memory task, Castegnaro et al [26] found that individuals with aMCI recalled object locations less accurately than healthy controls (HC). They also exhibited increased object-binding errors, that is, difficulties in correctly associating each object with its original location, with objects frequently assigned to incorrect positions. Crucially, the distance between the recalled and encoded item locations achieved a good diagnostic classification performance (area under the curve [AUC]=0.89). However, aMCI with positive AD biomarkers performed similarly to individuals with negative AD biomarkers. The authors concluded that a possible egocentric strategy might have compensated for allocentric deficits in aMCI with AD. Nevertheless, regardless of the biomarker profile, participants with aMCI exhibited OL impairments compatible with entorhinal lesions.

Beyond allocentric deficits, there is some evidence for egocentric spatial impairments in aMCI. Studies using paradigms such as the human analog of the Morris water maze (MWM) [27] have found that individuals with aMCI plus multiple-domain impairments (aMCImd) and those with a hippocampal amnestic profile also manifested deficits in egocentric spatial memory [28,29]. In contrast, patients with focal impairment in memory (amnestic mild cognitive impairment, single-domain [aMCIsd]) were particularly impaired in allocentric spatial memory. Intriguingly, research has begun to characterize specific behavioral patterns during spatial navigation tasks. Individuals with aMCI showed responses less clustered around the target location during allocentric and egocentric conditions compared to HC and non-aMCI [27-29]. Behavioral markers during navigation tasks can reveal these underlying impairments.

Indeed, spatial behaviors tracked during a navigation task in response to specific environmental cues could express altered allocentric-hippocampal (eg, almost direct trajectory between starting and target location) and egocentric-nonhippocampal (eg, wall-hugging-centrophobic navigation, random exploration) strategies, which unveil impaired spatial navigation in AD [30]. Additionally, studies of AD risk carriers have identified distinctive navigational behaviors using VR tasks, including wall-hugging behavior when environmental boundaries were displayed and a preference for navigating closer to local landmarks to reach target locations when egocentric cues were presented [31-33].

In summary, while existing literature has demonstrated spatial navigation deficits in aMCI, significant questions remain regarding the precise nature of these impairments. Specifically, how patients with aMCI navigate when different types of spatial cues are available (environmental boundaries vs discrete landmarks) and the relative impairment in allocentric vs egocentric processing strategies remains poorly investigated.

To overcome these gaps, this study aims to exploit digital markers from the VR OL spatial memory task to evaluate boundary- and landmark-based recall strategies in aMCI compared to HC. Based on the existing literature, our study

XSL•FO
RenderX

addresses two main objectives: (1) to evaluate spatial memory performance in aMCI compared to HC. Despite some evidence suggesting possible landmark-based (egocentric) impairments, we expect a focal allocentric spatial recall (boundary-based) deficit in aMCI, which we hypothesize will predict aMCI diagnosis; (2) to investigate navigational behavioral patterns. We expect patients with aMCI to display distinct navigation strategies compared to HC, specifically wall-hugging behavior in boundary-based conditions and centripetal navigation toward discrete landmarks in landmark-based conditions.

Finally, we explore these predictions in relation to aMCI subtypes, specifically aMCIsd and aMCImd variants.

Methods

Participants

Eighty older adults were recruited for this study and divided into 40 individuals meeting the core clinical criteria for MCI [34] and 40 HC. The MCI phenotype was established [34] in the patient group (aMCIsd=10 and aMCImd=30). The presence of cognitive impairment was determined by a team of neuropsychologists (CT, CSB, and DB) and a physician (KMG), based on the patient's anamnesis, medical report, and neuropsychological diagnosis according to a screening battery. Objective cognitive impairment was determined by clinical judgment, including qualitative analysis of the participant's test performance [35], and the psychometric equivalent scores method for Italian neuropsychological tests [36]. To reduce diagnostic bias, neuropsychological tests for each participant were evaluated by a team of neuropsychologists.

Inclusion criteria for the cognitively HC group were (1) age ≥65 years, (2) absence of cognitive impairment, (3) independence in functional abilities, and (4) absence of cognitive deterioration as determined by the Italian Mini-Mental State Examination (MMSE) cut-off [37].

MCI was diagnosed according to the core clinical criteria: (1) concern about a change in cognition expressed by the participant, an informant, or a clinician; (2) cognitive impairment in one or more domains; (3) independence in functional abilities; and (4) absence of a dementia diagnosis. Additional inclusion criteria were (1) age ≥65 years and (2) absence of cognitive deterioration as determined by the MMSE cut-off [37]. aMCIsd was determined when only the memory domain (verbal and/or visuospatial) was impaired on formal neuropsychological testing, whereas aMCImd was defined when, in addition to memory (verbal and/or visuospatial), participants had an objective impairment in an additional cognitive domain.

The following exclusion criteria were applied to both HC and MCI: (1) presence of an acute stroke or transient ischemic attack in the last 6 months, (2) diagnosis of neurological (eg, multiple sclerosis and Parkinson disease) and psychiatric disorders (eg, schizophrenia) or untreated psychiatric conditions (eg, mood and anxiety disorders), (3) history of traumatic brain injury with loss of consciousness, (4) physical or functional impairments (eg, musculoskeletal disorders and limb paresis/paresthesia) that could impair the use of the VR procedure, (5) severe or uncorrected visual impairment, and (6) presence of self-reported recurrent vertigo.

To compute the sample size, we used an effect size of f=0.32 [19]. For a mixed design (2 within-participant levels and 2 between-participant levels) with a power of 0.8 and an α of .05, the total sample size was 79 (calculated using the *pwr* R package).

Participants were recruited at the Outpatient Clinic of the Department of Medicine, Neurology, and Rehabilitation, IRCCS Istituto Auxologico Italiano—Mosè Bianchi, Milan. Table 1 reports the demographic and clinical baseline characteristics of the participants included in the analyses. An extended summary table is reported in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Summary of participant characteristics (N=80).

Variable	aMCI[a] (n=40)	HC[b] (n=40)	*P* value
Age (years), mean (SD)	74 (6)	72 (6)	.30
Sex, n (%)			.10
Female	30 (75)	23 (58)	
Male	10 (25)	17 (42)	
Education, mean (SD)	12.3 (4.5)	13.9 (3.4)	.10
MMSE[c] (points), mean (SD)	27.48 (2.10)	28.52 (1.88)	.02
FAB[d] (points), mean (SD)	14.99 (2.17)	17.05 (1.12)	<.001
TMT-A[e] (s), mean (SD)	36 (23)	28 (14)	.30
CF[f] (points), mean (SD)	4.73 (1.59)	5.06 (1.00)	.50
DF[g] (points), mean (SD)	5.69 (1.14)	6.26 (0.95)	.02
CSS[h] (points), mean (SD)	12 (7)	19 (7)	<.001
PM-I (points), mean (SD)	4.17 (2.27)	5.82 (1.28)	<.001
PM-D[i] (points), mean (SD)	4.11 (2.22)	5.84 (1.19)	<.001
GDS[j] (points), mean (SD)	2.65 (2.62)	1.93 (2.12)	.20
ADL[k] (points), mean (SD)	5.95 (0.22)	5.92 (0.27)	.70
IADL[l] (points), mean (SD)	7.93 (0.47)	7.93 (0.27)	.30
VR[m] encoding time (minute), mean (SD)	8.87 (5.14)	7.01 (3.73)	.07
PC[n] expertise (points), mean (SD)	2.43 (1.19)	2.74 (1.27)	.27

[a]aMCI: amnestic mild cognitive impairment.
[b]HC: healthy control.
[c]MMSE: Mini-Mental State Examination.
[d]FAB: frontal assessment battery.
[e]TMT-A: trail-making test part A.
[f]CF: Corsi forward.
[g]DF: digit forward.
[h]CSS: Corsi supra-span learning.
[i]PM: prose memory immediate (I) and delayed (D).
[j]GDS: geriatric depression scale-short.
[k]ADL: activities of daily living.
[l]IADL: instrumental activities of daily living.
[m]VR: virtual reality.
[n]PC: personal computer.

Materials

Neuropsychological Measures

The battery we administered comprised memory, language, attention, executive function, and visuospatial domains. The full set of neuropsychological tests and descriptions is reported in Multimedia Appendix 1.

OL Spatial Memory Task

We used a VR landmark- and boundary-based spatial memory recall similar to that described by Tuena et al [38]. Participants were seated in a comfortable position at a distance of 50 cm from the computer screen. In this task, a 3dRudder and a VR-ready computer (15.6-inch MSI GF63 Thin 11UC Intel Core i7-11800H, NVIDIA GeForce RTX 3050 4 GB, 16 GB RAM, and 512 GB SSD) were used to interact with the virtual environment. The 3dRudder is a circular platform designed for use while seated in a comfortable and safe position. To operate it, the user places the feet on top of the device, which is mounted on a semispherical base. This design enables intuitive movement control by tilting the feet in the desired direction or by pushing the toes or heels to move forward or backward. The device is equipped with inertial and pressure sensors that detect the user's movements and convert them into virtual actions. The use of a

XSL•FO
RenderX

motion foot pad instead of a classic joypad is intended to recruit lower-limb motor commands instead of hand and finger motor efferent copies, which is closer to real-world navigation [39].

The structure of the spatial memory task was as follows. During the encoding phase, participants were instructed to collect 4 items at a time and encode their locations within a circular arena (with a diameter of 50 virtual meters). Critically, during encoding, both the arena boundary (wall) and a fixed intra-arena landmark (obelisk; coordinates [−10.5, 10.5]) were simultaneously visible, allowing participants to encode object locations using both cues.

Items were presented randomly, and each object was collected 4 times in a random order. The 4 objects were the same for all the participants and were presented at the same fixed coordinates (cat [−5, −14], chair [−4, 20], bike [18, 10], and carrot [5, 17]). To visualize the object, participants had to navigate to its exact location. Upon reaching it, the object disappeared, and they proceeded to find the next one. The starting position (coordinates [0, 0]) of this phase was always the center of the circular arena, facing an arbitrary north.

During the immediate recall phase, participants had to remember and navigate to the exact location where the item (cued at the bottom of the screen) had previously been collected and then press a button on the keyboard to respond. The next object was then presented. Either the wall or the obelisk was randomly removed, forcing participants to rely primarily on one type of spatial cue. In the allocentric (wall) condition, the obelisk was removed, but the visible wall remained, encouraging boundary-based spatial memory. In the egocentric (obelisk) condition, the wall became invisible, but the obelisk remained visible, encouraging landmark-based spatial memory. Fixed distal cues (ie, mountain range and clouds) remained visible in both recall conditions as orientation cues. Each object was tested twice under both egocentric and allocentric conditions, resulting in a total of 16 trials (8 trials per recall condition).

An invisible wall was implemented as a constraint during the obelisk condition to ensure comparability between the 2 recall cue conditions. Figure 1 provides an overview of the task.

Figure 1. Object-location spatial memory task structure.In the encoding phase (A), participants were asked to learn the locations of 4 consecutive objects in a circular virtual arena containing both walls (allocentric cues) and a discrete landmark (egocentric cue). Participants collected each object 4 times in random order. During the recall phase (B and C), participants were cued to navigate to previously learned object locations; however, either the obelisk (B) or the wall (C) was randomly removed, forcing reliance on egocentric or allocentric spatial recall cues, respectively. While distal orientation cues remained visible in both conditions, each object was tested twice under both cue conditions (8 trials per spatial recall cue and 16 total trials).



The recorded virtual variables were the encoding time to collect and encode the object locations; the Cartesian coordinates (x, y) of the recalled location; the Cartesian coordinates (x, y) of the paths made by participants during the encoding phase; and the error for each trial at recall, measured as the Euclidean distance (expressed in virtual Unity software meters) between the recalled position and the actual location. Details of the setup are shown in Figure S1 in Multimedia Appendix 1.

Procedure

First, participants read and signed the consent forms to participate in the study. The study was conducted across 2 sessions. In the first session, participants were administered a comprehensive neuropsychological battery (refer to "Neuropsychological Measures" subheading). In the second session, participants completed the OL spatial memory task. Before the actual task, participants were instructed by the experimenter on how to use the 3dRudder in a training virtual environment provided by the 3dRudder software. Then, the encoding phase instructions were read: "Now, you will find yourself in a circular virtual arena where your task is to collect objects and memorize their locations, as you will be asked to recall them later. You will see one object at a time, and each object will appear four times in the same position to help you remember its location more easily. To aid in memorization, you can use the obelisk, the wall, the mountain range, and the fixed clouds as reference points. You can navigate within the arena using the 3dRudder. To collect an object, move directly over it. Once collected, it will disappear, and the next object will be presented."

After this phase, the recall instructions were given: "You will be asked to place each object in the position where you originally collected it. However, either the obelisk or the wall will be randomly removed. Once you reach the location you believe is correct, press the spacebar on the keyboard to place the object and proceed to the next one. You will be asked to reposition each object multiple times, regardless of the accuracy of your previous answers." Finally, a brief debriefing on the study's aims was provided (the experimental sessions lasted approximately 1 hour and 15 minutes in total).

Data Preparation and Statistical Analyses

R (version 3.6.3; R Foundation for Statistical Computing) software was used for the analyses. The main independent variables were group (HC vs aMCI) and spatial recall cue

XSL•FO
RenderX

(egocentric vs allocentric). This resulted in a mixed between- and within-participants design.

Our study consisted of 2 sets of analyses corresponding to our 2 main research objectives: linear mixed-effects ANOVAs and generalized linear regression models (Objective 1), and binomial tests for density distributions (data points count) of Cartesian encoding and response coordinates (Objective 2).

The main dependent variable of interest for Objective 1 was the recall measure of Euclidean distance (error between encoded items' Cartesian coordinates and recalled items' Cartesian coordinates; hereafter referred to as error). In addition, to gain further insight into spatial memory performance, we assessed OL binding errors and frame-switching cost. A thorough description of the statistical methods for Objective 1 is reported in Multimedia Appendix 1. To reduce bias, we also controlled for computer expertise (refer to Table 1 and Multimedia Appendix 1 for details) in the groups and carried out analyses adjusted for potential confounding variables (eg, age, gender, and global cognition). Objective 2 details are reported in the "Data Preparation and Statistical Analyses for Navigational Behavioral Patterns" section. The significance level was set to .05 in all analyses.

Data Preparation and Statistical Analyses for Navigational Behavioral Patterns

First, we analyzed path length in the HC and aMCI groups at encoding. To quantify participants' movement within the arena, we computed the Euclidean distance between consecutive position samples (path Cartesian coordinates) for each participant. For each participant, we then computed the mean step distance (average distance traveled per time step) and the total distance (sum of all stepwise distances). Group differences in mean step distance between participants with HC and aMCI were assessed using an independent samples *t* test (with Welch correction for unequal variances). The same procedure was adopted for paths at recall. Encoding and recall path coordinates were sampled at 1 Hz (2/80 files could not be saved due to technical issues).

We also computed the Euclidean distance of each path point at encoding from the fixed landmark and the center of the arena and averaged this distance for each participant to carry out an independent samples *t* test (Welch correction for unequal variances) between the 2 groups.

Second, to analyze the spatial distribution of encoding and response patterns in the arena, we developed a systematic approach that allowed us to examine potential biases toward specific reference points in both egocentric and allocentric conditions, similar to that described by Lee et al [10]. Recall coordinates corresponded to each participant's response.

Regarding the allocentric condition data point densities (encoding paths and recall data points), we divided the circular arena (radius=25 units) into inner (radius≤12.5 units) and outer annular regions to examine biases toward the center or boundary. We calculated point densities in each region by dividing point counts by their respective areas (inner area=π x 12.5^2 and annular area=π x [25^2 – 12.5^2]). To determine whether participants showed significant deviations from the expected distribution based on relative areas (25% inner and 75% outer), we conducted binomial tests comparing observed proportions to the expected proportions. Additionally, we calculated density ratios (inner to outer) to quantify preferences toward central or peripheral regions.

For the egocentric condition densities (encoding paths and recall data points), we compared the quadrant containing the egocentric (obelisk) landmark (coordinates [−10.5, 10.5]) with the diametrically opposite quadrant. Specifically, we assessed whether participants placed objects more frequently in the upper-left quadrant (containing the obelisk) compared to the lower-right quadrant. We performed binomial tests to determine whether the proportion of points in the landmark quadrant significantly differed from 0.5 (the expected proportion under the null hypothesis of equal distribution between the 2 quadrants). We also calculated density ratios between these quadrants to quantify the strength of any spatial recall cue–based bias. In addition, we also calculated the ratio of density ratios between groups to quantify the relative strength of spatial biases across clinical populations.

Finally, before proceeding to compute data point densities for recall, we ensured that responses were unbiased to fixed object locations. We computed an intra-arena landmark attractor index, normalized by the distance of each object from the landmark. This index quantifies how much a response is shifted toward the intra-arena landmark ($[d_{target} - d_{responseLM}]/d_{target}$), where d_{target} is the Euclidean distance from the object's original location to the landmark and $d_{responseLM}$ is the distance from the participant's response to the landmark. Positive values indicate a bias toward the landmark, whereas negative values indicate a bias toward the object's true location. We also computed an arena center attractor index, normalized by the distance of each object from the arena center. This index quantifies how much a response is shifted toward the center of the arena ($[d_{target} - d_{responseC}]/d_{target}$), where d_{target} is the distance from the object's original location to the arena center, and $d_{responseC}$ is the distance from the response to the center. Negative values indicate a bias toward the center of the arena, whereas positive values indicate a bias toward the object's true location. These 2 attractor indices were analyzed separately with a linear mixed-effects model, with participants and objects as random effects and group (HC vs aMCI) as a fixed predictor.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Istituto Auxologico Italiano (approval no 2023_01_31_11). Written informed consent was obtained from all participants before their participation in the study. All data were de-identified prior to analysis. Participants were assigned a unique alphanumeric code, and any directly identifying information was removed and stored separately in compliance with ethical and data protection regulations.

XSL•FO
RenderX

Results

Overview

Table 1 reports the key characteristics of the 2 groups. The extended table with all the neuropsychological measures and multiple Pearson correlations in the whole sample between spatial error performance and those measures is reported in Tables S1 and S2 in Multimedia Appendix 1.

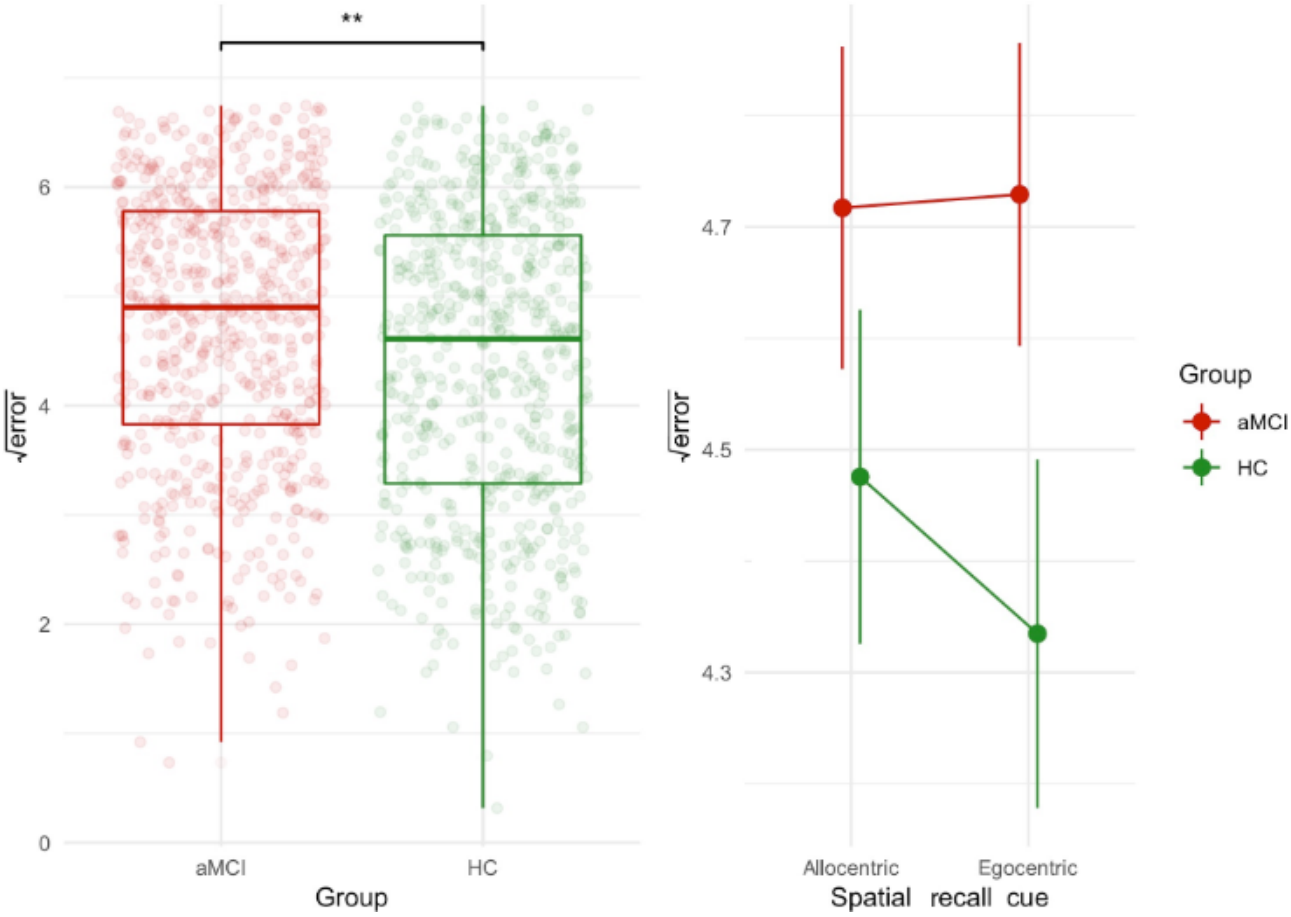
Objective 1

Differences Between aMCI and HC and Spatial Memory Recall Cues

To study group and spatial recall cue differences, we performed a linear mixed-effects ANOVA for the error metric.

We found a significant main effect of the group variable ($F_{1,77}$=8.37; *P*=.005; and *f*=0.33). Regardless of the spatial recall cue, patients with aMCI demonstrated a higher mean error of 23.9 (SD 11.3) virtual meters than HC, who showed a mean error of 21.3 (SD 11.7). We did not find a main effect of spatial recall cue (*P*=.37) or the interaction between spatial recall cue and group (*P*=.27). The significant effect of the group variable (*P*=.03) was also retained after adjusting for the covariates [26], namely age, years of education, gender (*P*=.03), retrieval time for each object, global cognition, age by time, and age by gender (*P*=.02). Figure 2 shows the results of these analyses with (A) a significant main effect of the group variable on spatial memory error, and (B) nonsignificant spatial memory performance by spatial recall cue in the groups (mean and 95% CIs), with a detailed summary of these analyses reported in Multimedia Appendix 1.

Figure 2. Group differences in the spatial memory task. aMCI: amnestic mild cognitive impairment; HC: healthy controls. **P<.01



Crucially, an ANOVA was conducted to explore any differences in recall time. We did not find a main effect of group (*P*=.07), spatial recall cue (*P*=.92), nor the interaction between group and spatial recall cue (*P*=.65). Results were unchanged when the model was adjusted for age, years of education, gender, and global cognition.

We also performed a polynomial linear mixed-effects model to investigate the performance across testing trials.

Concerning the error, the results of the group main effect above were replicated (*P*=.005). Crucially, we did not find any improvement or deterioration across testing trials (*P*=.40). The main effect of spatial recall cue (*P*=.37), the interaction between group and spatial recall cue (*P*=.27), the interaction between trials and group (*P*=.43), between trials and spatial recall cue (*P*=.98), and among trials, spatial recall cue, and group (*P*=.37) were not significant.

OL Memory Binding in aMCI and HC and Spatial Memory Recall Cues

To analyze the effects of group and spatial recall cues on binding errors, we performed a Poisson regression. Neither the main effect of group (*P*=.24) nor spatial recall cue type (*P*=.85) reached statistical significance. The interaction between group

XSL•FO
RenderX

and spatial recall cue type was also nonsignificant (P=.51). Results remained unchanged after adjusting for the covariates suggested by Castegnaro et al [26], namely age, education, gender, MMSE, and gender by age. Refer to Figure S2 for the binding errors in the groups. Figure S2 in Multimedia Appendix 1 shows the binding errors for the 4 objects.

Frame-Switching Cost in aMCI and HC

To analyze the effects of group (aMCI vs HC) and recall cue switching condition (ego-to-allo vs allo-to-ego vs no switch) on error, we performed a linear mixed-effects ANOVA.

We found a significant main effect of group ($F_{1,78}$=4.71; P=.03; and f=0.25). Regardless of the switching condition, HC had higher retrieval accuracy (mean 21.3, SD 6.89) compared to the aMCI group (mean 23.5, SD 6.89). The switching condition (P=.97) and the group-by-switching condition (P=.18) were not significant. We also performed a model adjusting for age, education, gender, retrieval time, MMSE, age by gender, and age by time; however, in this case, the main effect of group did not survive (P=.11). No other significant effects were found for the adjusted model. See Figure S3 in Multimedia Appendix 1 shows the unadjusted frame-switching cost results.

In addition, we performed a logistic regression analysis to explore the predictive power of the recall cue switching condition in classifying aMCI vs HC. We found that only the no-switch error (β=.12; P=.01) was positively associated with aMCI diagnosis (OR 1.13, 95% CI 1.03-1.25). The other switch predictors were not significant (ego-to-allo: P=.78 and allo-to-ego: P=.57). The association for no-switch error remained significant after adding relevant covariates (P=.04), such as age, education, gender, MMSE (β=−.4; P=.01), and age by gender. This suggests that frame switching is impaired in both conditions, as the no-switch condition was a good predictor of aMCI, whereas the other 2 conditions were not.

aMCI and HC Classification and Spatial Memory Recall Cues Performance

We investigated the predictive power of group classification of egocentric and allocentric performances using a logistic regression model.

Regarding the error, our model revealed that egocentric error was a significant predictor of group status (β=.98; P=.01), with higher error associated with aMCI diagnosis. In contrast, allocentric error did not significantly predict group membership (β=.33; P=.39). The intercept of the logistic regression was significant (β=−5.96; P=.01); when egocentric and allocentric errors were 0, there was a greater chance of being classified as HC.

The egocentric error effect also remained significant (β=1.11; P=.03) after controlling for relevant covariates, namely age, education, gender (β_{male}=−1.65; P=.02), retrieval time, and MMSE (β=−.43; P=.01) [26]. This pattern of results suggests a selective impairment in egocentric spatial navigation in patients with aMCI compared with HCs, while allocentric spatial representations appear relatively preserved. Table 2 reports the full model results.

Table . Full model logistic regression results. Reference class: aMCI[a].

Predictor	OR[b] (95% CI)	*P* value
Egocentric error (vm)[c]	3.02 (1.20-8.96)	.03
Allocentric error (vm)	1.25 (0.51-3.15)	.60
Age (years)	1.00 (0.91-1.09)	.90
Education (years)	0.93 (0.81-1.07)	.30
Sex		
Male	0.19 (0.04-0.71)	.02
MMSE[d] (points)	0.65 (0.46-0.88)	.008
Retrieval time (s)	0.99 (0.95-1.02)	.50

[a]aMCI: amnestic mild cognitive impairment.

[b]OR: odds ratio.

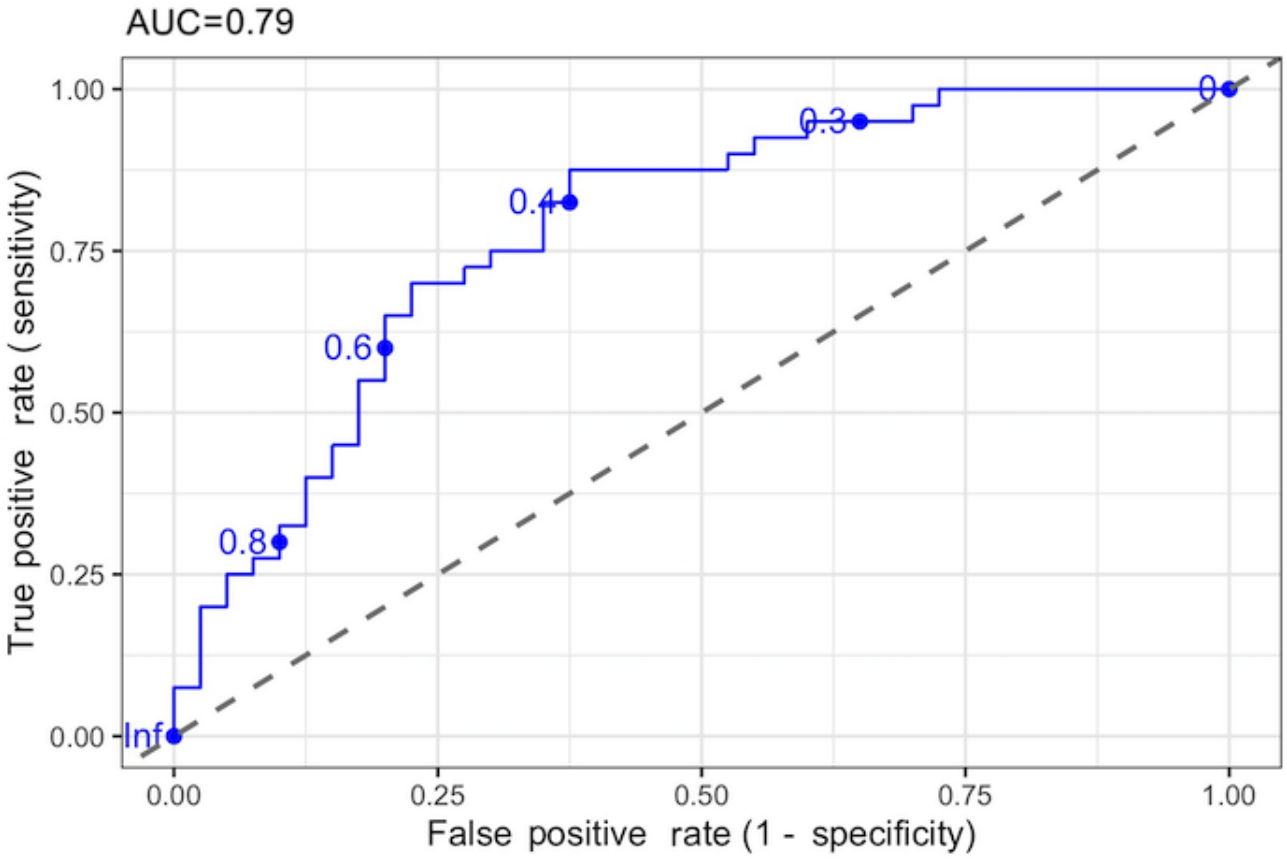
[c]vm: virtual meters.

[d]MMSE: Mini-Mental State Examination.

A receiver operating characteristic curve analysis was conducted to evaluate the diagnostic performance of the full model for the classification of aMCI. The AUC was 0.79, indicating moderate discriminatory ability (Figure 3). The optimal threshold was determined using the Youden index, which maximizes the sum of sensitivity and specificity. The analysis yielded an optimal probability cutoff value of 0.43. At this threshold, the test demonstrated a sensitivity of 0.88 and a specificity of 0.63, with a global accuracy of 0.75.

XSL•FO
RenderX

Figure 3. Receiver operating characteristic (ROC) curve of the full logistic regression model. AUC: area under the curve.



We also analyzed the cognitive measure of the model separately outside the logistic regression model. A receiver operating characteristic curve analysis was conducted to evaluate the diagnostic performance of the egocentric error measure. The AUC was 0.70, indicating moderate discriminatory ability. The optimal threshold was determined using the Youden index, which maximizes the sum of sensitivity and specificity. The analysis yielded an optimal cutoff value of 4.35. At this threshold, the test demonstrated a sensitivity of 0.80 and a specificity of 0.58. This means that egocentric error is particularly effective for ruling out aMCI, given the high sensitivity and, consequently, the low number of false negatives (ie, when egocentric error is below the cutoff, the participant is likely to be HC). Conversely, values above the cutoff may result in false-positive classifications because of the lower specificity.

Finally, MMSE showed an AUC of 0.68, indicating poor discriminatory ability. The analysis yielded an optimal cutoff value of 28.15. At this threshold, the test demonstrated a sensitivity of 0.68 and a specificity of 0.65. Consequently, the egocentric error is more accurate than MMSE for classifying healthy older adults.

Despite not being a significant predictor, we also computed a receiver operating characteristic curve analysis for the allocentric error. The AUC was 0.61, and the best cutoff was 4.87, which yielded a sensitivity of 0.5 and specificity of 0.75. This is specific to the egocentric error measure; in this case, the higher specificity is a better predictor of aMCI given the lower number of false positives, and a positive score above the cutoff may indicate aMCI.

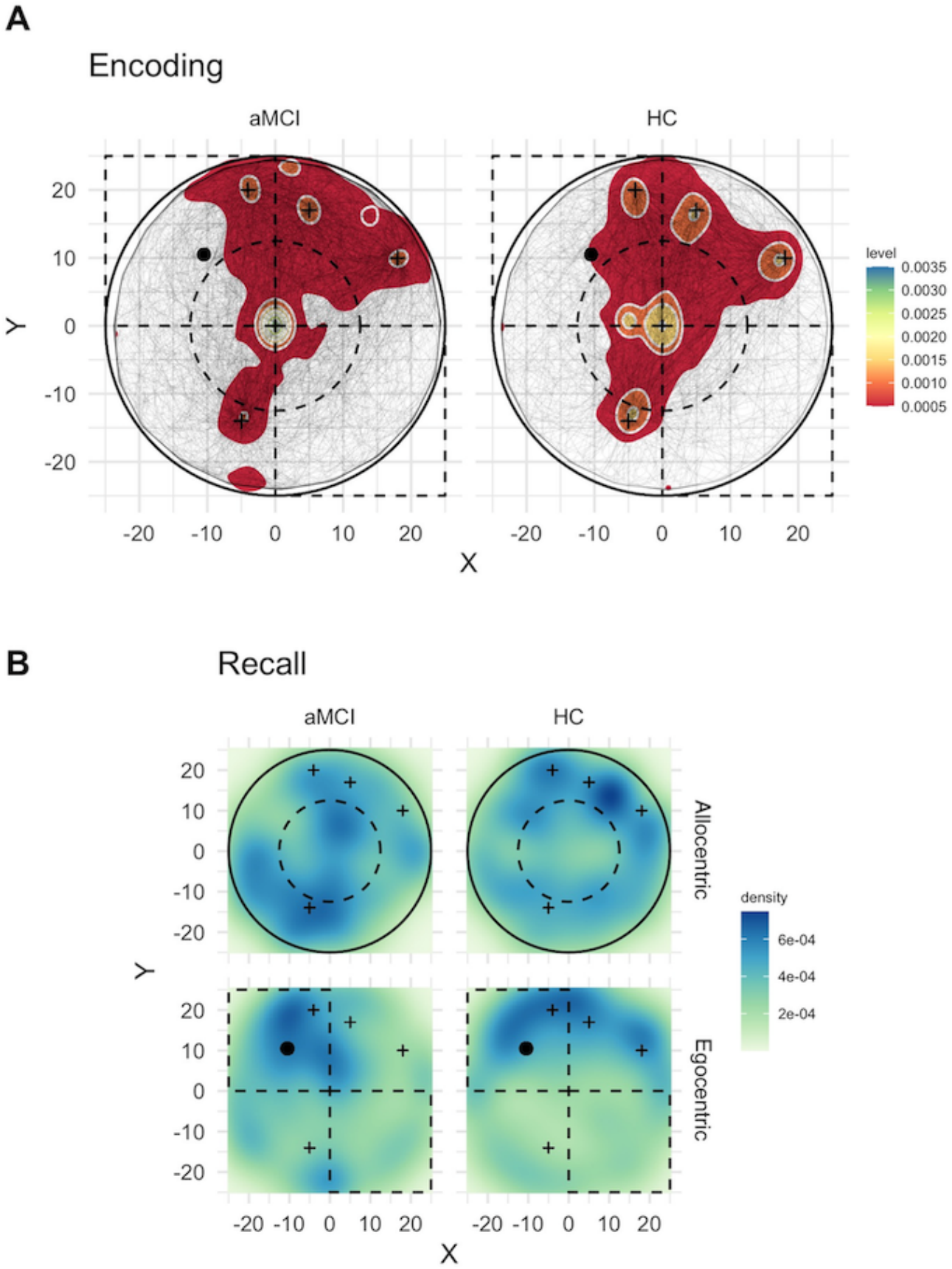
Objective 2

Encoding Paths in aMCI and HC

We computed the patterns of encoding Cartesian (path) coordinates to explore navigation behavior in this phase. This data type allows us to explore the paths within the virtual environment during the encoding phase of the object locations. Refer to Figure 4 for the regions used in these analyses and the path density distributions in the groups at encoding. In the figure, (A) shows heat maps representing the density of encoding path points, with dark blue indicating higher concentration. Black lines represent paths taken by participants during the encoding phase, and black cross markers show object positions. The black dot represents the obelisk landmark (–10.5, 10.5), with dashed lines delineating the analyzed upper-left and lower-right quadrants; the solid circle represents the arena boundary (radius=25 units), with a dashed inner circle (radius=12.5 units) separating central and annular regions. (B) Shows heat maps of response point densities at recall, with dark blue indicating higher concentration. Black cross markers show encoding object positions. In the egocentric condition, the black dot represents the obelisk landmark (–10.5, 10.5), and in the allocentric condition, the solid circle represents the arena boundary (radius=25 units), with a dashed inner circle (radius=12.5 units) separating central and annular regions. Both cues were presented during encoding. Allocentric refers to a boundary cue presented, but no local landmark, and egocentric refers to local landmark presented, but no boundary. Legends show density levels.

XSL•FO
RenderX

Figure 4. Density distributions within the virtual environment. aMCI: amnestic mild cognitive impairment; HC: healthy control.



First, we calculated whether the groups differed in terms of path length. A *t* test showed that groups did not differ in path length (*P*=.08). Along with encoding time (refer to Table 1), this result shows that time and path do not differ quantitatively in aMCI from normal cognitive aging.

XSL•FO
RenderX

Second, a Welch 2-sample *t* test was conducted to compare the mean distance from the intra-arena landmark between participants with aMCI and HC. There was a significant difference in mean distance between the 2 groups (t_{63}=2.35; *P*=.02). Participants with aMCI (mean 19.57, SD 1.34) maintained a greater average distance from the landmark during encoding navigation compared to HC participants (mean 18.65, SD 2.05). Similarly, a Welch 2-sample *t* test was conducted to compare the mean distance from the center of the arena between participants with aMCI and HC. There was a significant difference in mean distance between the 2 groups (t_{76}=2.35; *P*=.02). Participants with aMCI (mean 15.45, SD 1.96) maintained a greater average distance from the center compared to HC participants (mean 14.45, SD 1.77). These analyses showed a tendency to navigate closer to the boundary of the arena and away from the landmark, which is replicated in the data point density analyses.

HC demonstrated significantly higher point density in the inner circle compared to patients with aMCI (36.4% vs 29.16%, binomial test *P*<.001), suggesting a centripetal pattern during the encoding phase for HC. Conversely, patients with aMCI showed significantly higher concentration in the outer ring (70.83% vs 63.6%, binomial test *P*<.001), suggesting a centrifugal bias toward the boundary during the encoding phase.

Quadrant analysis revealed that HC had significantly higher point density in the top-left quadrant (region that contained the egocentric cue [obelisk]) compared to participants with aMCI (27.75% vs 24.37%; *P*<.001), suggesting a bias toward the egocentric landmark for HC. Conversely, participants with aMCI showed a significantly higher concentration in the bottom-right quadrant (14.98% vs 13.64%; *P*<.001), suggesting a centrifugal pattern away from the egocentric cue.

Overall, the results indicate a centrifugal pattern of encoding navigation for aMCI away from the egocentric cue and toward the boundary (away from the center of the arena) and a more clustered pattern for HC toward the center of the arena and the egocentric cue.

Point Pattern Analysis and Spatial Recall Cue in HC and aMCI

We analyzed the response patterns to investigate behavioral recall differences between the groups.

First, we calculated whether the groups differed in terms of path length. Results showed that aMCI had significantly (t_{74}=3.64; *P*<.001) longer paths (mean 2.83, SD 0.89) compared to HC (mean 2.16, SD 0.74).

Second, normalized attractor index analyses for egocentric and allocentric performance showed that the HC group had a lower bias toward the landmark (β=−.21; SE=0.07; *P*=.007) and toward the center of the arena (β=−.14; SE=0.06; *P*=.02) compared to the aMCI group. This ensured that the following analyses were relatively unbiased by fixed object locations. Indeed, this result was replicated in the data point densities analyses reported in the next paragraph: aMCI recalled locations in the allocentric condition toward the center of the arena and toward the landmark during the egocentric condition.

Regarding data point densities, we started by investigating the point pattern for allocentric performance. We analyzed the spatial distribution of response points in the 2 groups by examining the density patterns relative to the structure of the circular arena. For this purpose, we divided the circular arena into an inner region (within the inner circle of radius 12.5 units, far from the wall) and an outer annular region (between the inner and outer circles, 12.5< radius ≤25 units, closer to the wall). The inner area constituted 25% of the total area, while the outer area represented 75%.

Refer to Figure 4 for the regions used in these analyses, the groups' density point distributions, and recall cue conditions.

The HC group placed 17.81% of their recall points in the inner area (57/320 points) and 82.19% in the outer area (263/320 points). This distribution significantly differed from the expected proportions based on the inner and the outer portion of the arena (binomial test; *P*=.002).

In contrast, the aMCI group placed 23.75% (76/320) of their recall points in the inner area and 76.25% (244/320) in the outer area. This distribution did not significantly differ from the expected proportions for the inner and outer sections of the arena (binomial test; *P*=.65). To account for the different sizes of the inner and outer areas, we calculated point densities (points per square unit). In the HC group, the density was 0.11 points/unit^2 in the inner area and 0.18 points/unit^2 in the outer area, yielding an inner-to-outer density ratio of 0.65.

For the aMCI group, the density (points per unit area) was 0.15 points/unit^2 in the inner area and 0.17 points/unit^2 in the outer area, resulting in an inner-to-outer density ratio of 0.93. The difference in density ratios between groups was −0.28, indicating that patients with aMCI showed a more balanced distribution of points between inner and outer areas than HC participants, who demonstrated a stronger preference for the outer area. This suggests a potential alteration in allocentric spatial memory strategies in aMCI, characterized by reduced usage of the circular boundary as a spatial reference point compared with healthy older adults.

We then inspected the egocentric performance. We examined the spatial distribution of response points across opposing arena quadrants in patients with aMCI and HC. We focused on comparing the upper left quadrant, which contained the egocentric cue (obelisk at coordinates [−10.5, 10.5]), with the diametrically opposite lower right quadrant.

In the HC group, 32.19% (103/320) of recall points were positioned in the upper left quadrant, while 17.19% (55/320) were located in the lower right quadrant. This asymmetric distribution between the 2 quadrants was statistically significant (binomial test; *P*<.001).

Similarly, the aMCI group placed 32.50% (104/320) of recall points in the upper left quadrant and 20.00% (64/320) in the lower right quadrant. This distribution also showed a significant preference for the upper left quadrant (binomial test; *P*=.003).

To account for the spatial area, point densities were calculated for each quadrant. In the HC group, the point density was 0.21

XSL•FO
RenderX

points/unit^2 in the upper left quadrant and 0.11 points/unit^2 in the lower right quadrant, yielding an upper-left–to–lower-right density ratio of 1.87.

For the aMCI group, the point density was 0.21 points/unit^2 in the upper left quadrant and 0.13 points/unit² in the lower right quadrant, resulting in a density ratio of 1.62. The difference in density ratios between groups was 0.25, indicating that HC participants showed a slightly stronger preference for the upper left quadrant compared to patients with aMCI. The absolute differences in point densities between groups were minimal in both the upper left (−0.01 points/unit^2) and lower right (−0.01 points/unit^2) quadrants, with patients with aMCI showing slightly higher densities in both regions.

Explorative Analyses in aMCIsd and aMCImd

We used robust methods to compare encoding time and spatial error differences, and robust logistic regression between aMCI phenotypes, given unequal group sizes. A comprehensive description of aMCI phenotype differences is reported in Multimedia Appendix 1. To summarize, participants with aMCIsd spent significantly more time (P=.004) in the encoding phase compared to participants in the aMCImd group. Egocentric and allocentric errors could not predict aMCIsd or aMCImd diagnosis. Regarding spatial behavior during encoding, aMCIsd showed a statistically significant centrifugal (P<.001) pattern away from the local landmark, whereas aMCImd navigated closer to the obelisk; no differences were observed between the central and annular regions of the arena. Regarding spatial behavior during recall, no differences were reported again during the boundary-recall condition between the 2 groups in the central and annular regions. The aMCImd group showed a significantly higher response density closer to the discrete landmark compared to the control region (P=.008); this pattern was not observed in aMCIsd.

Discussion

Principal Findings

In this study, we sought to profile aMCI spatial memory performance using a VR OL task that manipulated available spatial cues during recall: environmental boundaries (engaging allocentric processing) vs discrete landmarks (engaging egocentric processing). Participants were asked to collect objects and encode their locations in a virtual arena with discrete landmarks and boundaries. At retrieval, they were asked to put objects in place where they were previously collected, either using the discrete landmark or the boundary. We expected patients with aMCI to show impaired spatial memory, particularly in boundary-based recall conditions, and to exhibit distinct navigational strategies, including wall-hugging behavior and preferential use of local landmarks.

Our predictions were only partially confirmed by the study results. Concerning the first objective, we found that regardless of the spatial recall cue, patients with aMCI exhibited lower spatial memory performance than HCs; surprisingly, the egocentric error, rather than the allocentric error, was a significant predictor of aMCI diagnosis. No differences were found between aMCIsd and aMCImd. Second, we did not find deterioration in testing trials or any OL memory binding deficits in aMCI. Finally, contrary to the previous literature, we found no impairment in switching abilities (from egocentric to allocentric and from allocentric to egocentric).

Concerning our second research objective, we found a wall-hugging–like behavior in aMCI and navigation away from the discrete landmark (especially in aMCIsd) during the encoding phase.

Interestingly, patients with aMCI showed no quantitative encoding deficits (ie, comparable encoding time and path length to HCs), but differed qualitatively in their encoding strategies. Regarding the recall phase, patients with aMCI gathered the responses at the center of the environment (in both phenotypes) in the allocentric condition and toward the landmark (especially in aMCImd) during the egocentric recall cue condition. Quantitatively, the 2 groups differed in path length but not in retrieval time.

Objective 1

While Castegnaro et al [26] showed that boundary-based recall has promising discriminative performance between aMCI and HC groups, intriguingly, we also found that higher egocentric error was associated with aMCI diagnosis. First, this is in line with a recent systematic review [22] that showed both egocentric and allocentric deficits in aMCI in real-world and VR tasks, extending the traditional focus on allocentric impairments [26-29,40], including egocentric impairments across the AD continuum. Our findings are consistent with the well-established association between AD and allocentric spatial memory deficits, yet they also highlight a complementary impairment emerging along the AD continuum, especially in its early stages. Although allocentric performance did not reach significance in the logistic model, it proved useful for ruling in the condition, whereas egocentric error was more informative for ruling out an aMCI diagnosis. This pattern aligns with a recent diagnostic meta-analysis conducted by our group [41], which showed that both egocentric and allocentric measures contribute valuable diagnostic information for AD, supporting a complementary, rather than competing, approach.

Our findings on egocentric impairment also align with recent results [42-45] that have provided compelling evidence for the coexistence of boundary egocentric vectors in medial temporal structures with allocentric representations. These body-based representations can integrate head direction signals to form a map of allocentric space, aiding navigation toward significant locations within an environment. Despite the right hemisphere, and particularly the hippocampal region, being the dedicated substrate for allocentric spatial computations [7,46], the human left hippocampus supports landmark-based spatial memory by sequencing spatiotemporal navigation (eg, the sequence of body-based movements at a given location and time) in an episodic-like format [11,46]. Aligned with this, patients with aMCI with hippocampal memory deficit, compared to those with frontal memory deficit, showed a greater deficit in egocentric spatial memory [28,29]. In addition, a critical diagnostic consideration is that current MCI neuropsychological classifications do not consider the modality of memory

XSL•FO
RenderX

impairment (eg, verbal vs visuospatial aMCI), potentially masking important cognitive differences that affect spatial navigation abilities. Given the heterogeneity of MCI, caution should be adopted when designing spatial navigation and memory tasks and interpreting their results. Finally, our results confirmed that egocentric spatial memory is preserved in healthy aging and is a good metric to distinguish between aMCI and HC groups; this aligns in particular with a previous study that showed preserved landmark-based compared to boundary-based spatial memory in aging [16]. It is also worth noting that, although gender showed a significant effect as a covariate in both the ANOVAs and the logistic regression, the main results remained significant. This suggests that our findings are not biased by participants' sex. Nevertheless, future research should further investigate potential gender differences in aMCI to better elucidate possible distinctions between egocentric and allocentric processes.

Our study revealed no performance improvement or deterioration in either the aMCI or HC groups. This contradicts previous research suggesting consolidation deficits in aMCI [27-29]. The absence of deterioration, paired with similar encoding times across groups, may indicate frontal involvement in retrieval strategies rather than hippocampal consolidation deficits. This aligns with evidence highlighting dysexecutive memory impairments in aging, aMCI, and AD, as well as the critical role of executive functions in supporting spatial navigation [47-49].

Unlike findings from a previous study [26] showing performance differences during allocentric recall, we found similar OL binding performance across cue types. Our stricter age criterion (≥65 years) likely explains why aMCI and HC performed comparably, as normal aging involves associative memory binding deficits linked to hippocampal and frontal dysfunction [49-51]. However, unimpaired OL binding was also found in AD genetic risk carriers [52].

Finally, while HC demonstrated lower switching costs than aMCI overall, we observed comparable frame-switching costs between groups, contradicting previous findings [24] of differential impairments in allocentric-to-egocentric vs egocentric-to-allocentric switching. This suggests that shifting between spatial representations represents a fundamental process affected by both normal aging and AD [5,15,18,53].

Objective 2

Concerning our second objective, we confirmed that the pattern of encoding paths and responses during retrieval could be a cognitive marker of aMCI [27-29]. However, we extended previous research by demonstrating specific biases in virtual navigation according to spatial recall cues and group.

Concerning the encoding paths, we found that individuals with aMCI, at the group level, tended to explore the arena closer to the boundaries, avoiding the center of the environment. In contrast, the inverse pattern was found in HC. A similar centripetal pattern was found for the discrete landmark in aMCI, while HC tended to navigate closer to the landmark. The former finding is compatible with a thigmotaxis-like (wall-hugging) navigation behavior observed in animals in the MWM [30]. This spatial behavior has been observed in several species while exploring a novel environment and is considered a safety-seeking strategy in the early phases [54]. However, thigmotaxis could also reflect pathological spatial behavior. Indeed, rats with lesions to the hippocampus exhibit perseverative thigmotaxis and an inability to learn more successful strategies [55]. Similarly, lesions to the dorsomedial striatum lead to increased thigmotaxis due to the inability to switch from this strategy to a more efficient one [56]. Importantly, thigmotaxis has been observed in humans as an early strategy in novel environments and is associated with higher levels of fear and avoidance bias toward fear-mobilizing situations or agoraphobia [57,58]. Extensive or pathological use of egocentric behavior, such as thigmotaxis, hampers the capacity to switch between spatial navigation strategies, which is crucial to forming a gestalt-like representation of the space, also known as the cognitive map [57,59]. In line with this, research on path integration (ie, the ability to track one's current position using current self-motion cues) has provided compelling evidence that environmental cues, particularly boundaries and, to a lesser extent, discrete landmarks, can correct navigation errors [8,60,61]. The boundary-hugging bias might reflect a compensatory egocentric mechanism, possibly also driven by spatial anxiety due to disorientation, that leads individuals with aMCI to walk closer to the boundaries virtually. Complementary to this, the local landmark is not perceived as a useful affordance [39] for navigation purposes in the aMCI group. This latter pattern is particularly true for aMCIsd, which also showed longer encoding times than aMCImd, suggesting a specific deficit during the learning process. The landmark-based deficit is compatible with findings on the human analog of the MWM in the egocentric recall test [28] and path integration movements toward the landmark in individuals with genetic risk for AD [33]. In contrast, HC had higher path density in the center of the arena, hinting at an allocentric-hippocampal strategy between object locations, akin to the "direct path" observed in MWM [30]. Collectively, the results show that in addition to an allocentric navigation deficit, aMCI, especially in aMCIsd, is accompanied by an egocentric navigation deficit at encoding.

Concerning the recalled Cartesian coordinates, we found that HC effectively leverages environmental boundaries for spatial memory recall; this boundary-anchored strategy may be partially compromised in older people with aMCI, who tended to remember object locations at the center of the arena. Indeed, the human hippocampus contains boundary cells that help to encode item locations with respect to distances from a boundary and support effective recall [10]. Conversely, patients with aMCI (especially aMCImd) and HC effectively used the egocentric reference point (obelisk) when performing spatial memory recall, with a strong preference for placing objects in the quadrant containing this landmark, especially for the HC. The results were similar in both aMCI phenotypes. This is partially in line with path integration based on boundary and landmark cues in individuals at risk for AD [33]. Bierbrauer et al [33] found that the distance of the landmark to the target condition is critical for recall performance during the landmark cue condition. At the same time, for young adults with genetic risk, this is true for the boundary cue. In addition, they showed

XSL•FO
RenderX

that AD risk carriers walked toward the landmark while on the path toward the target location.

However, this finding must be interpreted considering the encoding path patterns, error performance, and group discrimination. In the former case, we demonstrated that participants with aMCI during the encoding phase tended to navigate closer to the boundary. At the same time, at recall, they remembered object locations in the center of the arena. The opposite pattern was found for HC. This hints at an allocentric spatial memory impairment that, however, is not particularly discriminative between the 2 groups.

In the latter case, the bias toward the landmark at recall, in the absence of an environmental boundary, is indicative of defective egocentric recall performance, which arises as a first step in the encoding phase, possibly due to a primary bias toward the boundary in aMCI, leading participants to navigate away from the local landmark. This behavior was not observed in HC. Consequently, while the saliency of the landmark with respect to object locations is partially preserved during the testing phase, the recall accuracy for the actual location is impaired, allowing for better discrimination between groups. Finally, both MCI phenotypes have, in addition to an allocentric, a different egocentric deficit: aMCIsd does not use the egocentric cue as an affordance at encoding, while aMCImd does at recall.

Crucially, we found that qualitative measures of navigation during encoding revealed specific deficits that were not detected by quantitative measures such as total time spent in the environment (encoding time) or path length. This underscores the importance of analyzing behavioral strategies, which may capture the distinctive spatial navigation deficits observed in aMCI and potentially in AD. Similarly, although recall time did not differ between the groups, the combination of path length and qualitative patterns of spatial memory responses indicates that participants with aMCI exhibit both quantitative and qualitative impairments in spatial navigation. This aligns with the long, and sometimes overlooked, tradition of qualitative analysis in the neuropsychology of memory [35]. Quantitative scores alone provide limited information; they should always be complemented by an examination of how patients perform a given task. Previous research in spatial navigation has demonstrated that response patterns reveal underlying spatial memory behaviors, providing valuable insights into the deficits observed in aMCI [27-29].

Limitations

Our study has some limitations that should be acknowledged. First, we lack AD biomarkers that could help to define the underlying pathology better, and we collected participants in only 1 center. Second, we used a nonimmersive VR setup, which lacks crucial bodily (eg, vestibular) information, and this could have affected spatial memory results. Future studies could overcome these limitations by using an immersive apparatus equipped with an eye-tracking system to control for visual exploration, thereby discovering innovative digital biomarkers for aMCI diagnosis in larger cohorts.

Conclusion

To conclude, our findings suggest that suspected AD pathology affects not only allocentric spatial memory in aMCI but also impairs egocentric spatial representations. The latter constitutes a complementary critical process in navigation and could represent a useful digital biomarker in the progression of AD. This dual impairment of spatial cognition provides a more comprehensive understanding of the navigational difficulties in aMCI and may offer new avenues for detection and intervention with digital health technologies.

Acknowledgments

We thank all the participants who participated in the study, as well as Sarah Todisco and Alessandra Magrelli, who supported the study during data collection.

Funding

The publication fee has been supported by Ricerca Corrente from Italian Ministry of Health. The project was supported by grant RF-2021-12374099.

Data Availability

The datasets generated and analyzed during this study are available in the Zenodo repository [62].

Code is available upon request to the corresponding author.

Authors' Contributions

CT conceptualized the study, wrote the first draft of the manuscript, and conducted the statistical analyses. CT, SS, and CSB contributed to writing and editing. CSB and DB managed the study and data. CSB, DB, and KMG collected the data. GR and MSB supervised the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Neuropsychological measures and spatial memory tasks details and analyses.
[DOCX File, 498 KB - aging_v9i1e79224_app1.docx]

References

1. Burgess N. Spatial cognition and the brain. Ann N Y Acad Sci 2008 Mar;1124:77-97. [doi: 10.1196/annals.1440.002] [Medline: 18400925]
2. Poulter S, Hartley T, Lever C. The neurobiology of mammalian navigation. Curr Biol 2018 Sep 10;28(17):R1023-R1042. [doi: 10.1016/j.cub.2018.05.050] [Medline: 30205053]
3. Zhang H, Ekstrom A. Human neural systems underlying rigid and flexible forms of allocentric spatial representation. Hum Brain Mapp 2013 May;34(5):1070-1087. [doi: 10.1002/hbm.21494] [Medline: 22786703]
4. Ekstrom AD, Arnold A, Iaria G. A critical review of the allocentric spatial representation and its neural underpinnings: toward a network-based perspective. Front Hum Neurosci 2014;8:803. [doi: 10.3389/fnhum.2014.00803] [Medline: 25346679]
5. Vann SD, Aggleton JP, Maguire EA. What does the retrosplenial cortex do? Nat Rev Neurosci 2009 Nov;10(11):792-802. [doi: 10.1038/nrn2733] [Medline: 19812579]
6. Serino S, Riva G. How different spatial representations interact in virtual environments: the role of mental frame syncing. Cogn Process 2015 May;16(2):191-201. [doi: 10.1007/s10339-015-0646-4] [Medline: 25663561]
7. Doeller CF, King JA, Burgess N. Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. Proc Natl Acad Sci U S A 2008 Apr 15;105(15):5915-5920. [doi: 10.1073/pnas.0801489105] [Medline: 18408152]
8. Chersi F, Burgess N. The cognitive architecture of spatial navigation: hippocampal and striatal contributions. Neuron 2015 Oct 7;88(1):64-77. [doi: 10.1016/j.neuron.2015.09.021] [Medline: 26447573]
9. Guderian S, Dzieciol AM, Gadian DG, et al. Hippocampal volume reduction in humans predicts impaired allocentric spatial memory in virtual-reality navigation. J Neurosci 2015 Oct 21;35(42):14123-14131. [doi: 10.1523/JNEUROSCI.0801-15.2015] [Medline: 26490854]
10. Lee SA, Miller JF, Watrous AJ, et al. Electrophysiological signatures of spatial boundaries in the human subiculum. J Neurosci 2018 Mar 28;38(13):3265-3272. [doi: 10.1523/JNEUROSCI.3216-17.2018] [Medline: 29467145]
11. Iglói K, Doeller CF, Berthoz A, Rondi-Reig L, Burgess N. Lateralized human hippocampal activity predicts navigation based on sequence or place memory. Proc Natl Acad Sci U S A 2010 Aug 10;107(32):14466-14471. [doi: 10.1073/pnas.1004243107] [Medline: 20660746]
12. Shine JP, Valdés-Herrera JP, Tempelmann C, Wolbers T. Evidence for allocentric boundary and goal direction information in the human entorhinal cortex and subiculum. Nat Commun 2019 Sep 5;10(1):4004. [doi: 10.1038/s41467-019-11802-9] [Medline: 31488828]
13. Rodgers MK, Sindone JA III, Moffat SD. Effects of age on navigation strategy. Neurobiol Aging 2012 Jan;33(1):202. [doi: 10.1016/j.neurobiolaging.2010.07.021] [Medline: 20832911]
14. Lester AW, Moffat SD, Wiener JM, Barnes CA, Wolbers T. The aging navigational system. Neuron 2017 Aug 30;95(5):1019-1035. [doi: 10.1016/j.neuron.2017.06.037] [Medline: 28858613]
15. Colombo D, Serino S, Tuena C, et al. Egocentric and allocentric spatial reference frames in aging: a systematic review. Neurosci Biobehav Rev 2017 Sep;80:605-621. [doi: 10.1016/j.neubiorev.2017.07.012] [Medline: 28760627]
16. Schuck NW, Doeller CF, Polk TA, Lindenberger U, Li SC. Human aging alters the neural computation and representation of space. Neuroimage 2015 Aug 15;117:141-150. [doi: 10.1016/j.neuroimage.2015.05.031] [Medline: 26003855]
17. Jagust W. Imaging the evolution and pathophysiology of Alzheimer disease. Nat Rev Neurosci 2018 Nov;19(11):687-700. [doi: 10.1038/s41583-018-0067-3] [Medline: 30266970]
18. Serino S, Cipresso P, Morganti F, Riva G. The role of egocentric and allocentric abilities in Alzheimer's disease: a systematic review. Ageing Res Rev 2014 Jul;16:32-44. [doi: 10.1016/j.arr.2014.04.004] [Medline: 24943907]
19. Iachini T, Rapuano M, Ruotolo F, Iavarone A, Iuliano S, Ruggiero G. Early spatio-temporal and cognitive deficits in Alzheimer's disease. J Clin Med 2025 Jan 17;14(2):1-15. [doi: 10.3390/jcm14020579] [Medline: 39860585]
20. Petersen RC. Mild cognitive impairment. N Engl J Med 2011 Jun 9;364(23):2227-2234. [doi: 10.1056/NEJMcp0910237]
21. Dubois B, Albert ML. Amnestic MCI or prodromal Alzheimer's disease? Lancet Neurol 2004 Apr;3(4):246-248. [doi: 10.1016/S1474-4422(04)00710-0] [Medline: 15039037]
22. Tuena C, Mancuso V, Stramba-Badiale C, et al. Egocentric and allocentric spatial memory in mild cognitive impairment with real-world and virtual navigation tasks: a systematic review. J Alzheimers Dis 2021;79(1):95-116. [doi: 10.3233/JAD-201017] [Medline: 33216034]
23. Cerman J, Andel R, Laczo J, et al. Subjective spatial navigation complaints - a frequent symptom reported by patients with subjective cognitive decline, mild cognitive impairment and Alzheimer's disease. Curr Alzheimer Res 2018;15(3):219-228. [doi: 10.2174/1567205014666171120145349] [Medline: 29165083]
24. Ruggiero G, Iavarone A, Iachini T. Allocentric to egocentric spatial switching: impairment in aMCI and Alzheimer's disease patients? Curr Alzheimer Res 2018;15(3):229-236. [doi: 10.2174/1567205014666171030114821] [Medline: 29086696]

XSL•FO
RenderX

25. Tangen GG, Engedal K, Bergland A, Moger TA, Hansson O, Mengshoel AM. Spatial navigation measured by the Floor Maze Test in patients with subjective cognitive impairment, mild cognitive impairment, and mild Alzheimer's disease. Int Psychogeriatr 2015 Aug;27(8):1401-1409. [doi: 10.1017/S1041610215000022]
26. Castegnaro A, Howett D, Li A, et al. Assessing mild cognitive impairment using object-location memory in immersive virtual environments. Hippocampus 2022 Sep;32(9):660-678. [doi: 10.1002/hipo.23458] [Medline: 35916343]
27. Hort J, Laczó J, Vyhnálek M, Bojar M, Bures J, Vlcek K. Spatial navigation deficit in amnestic mild cognitive impairment. Proc Natl Acad Sci U S A 2007 Mar 6;104(10):4042-4047. [doi: 10.1073/pnas.0611314104] [Medline: 17360474]
28. Laczó J, Vlcek K, Vyhnálek M, et al. Spatial navigation testing discriminates two types of amnestic mild cognitive impairment. Behav Brain Res 2009 Sep 14;202(2):252-259. [doi: 10.1016/j.bbr.2009.03.041] [Medline: 19463709]
29. Laczó J, Andel R, Vyhnalek M, et al. Human analogue of the morris water maze for testing subjects at risk of Alzheimer's disease. Neurodegener Dis 2010;7(1-3):148-152. [doi: 10.1159/000289226] [Medline: 20197695]
30. Curdt N, Schmitt FW, Bouter C, et al. Search strategy analysis of Tg4-42 Alzheimer Mice in the Morris Water Maze reveals early spatial navigation deficits. Sci Rep 2022 Mar 31;12(1):5451. [doi: 10.1038/s41598-022-09270-1] [Medline: 35361814]
31. Coughlan G, Coutrot A, Khondoker M, Minihane AM, Spiers H, Hornberger M. Toward personalized cognitive diagnostics of at-genetic-risk Alzheimer's disease. Proc Natl Acad Sci U S A 2019 May 7;116(19):9285-9292. [doi: 10.1073/pnas.1901600116] [Medline: 31015296]
32. Kunz L, Schröder TN, Lee H, et al. Reduced grid-cell-like representations in adults at genetic risk for Alzheimer's disease. Science 2015 Oct 23;350(6259):430-433. [doi: 10.1126/science.aac8128] [Medline: 26494756]
33. Bierbrauer A, Kunz L, Gomes CA, et al. Unmasking selective path integration deficits in Alzheimer's disease risk carriers. Sci Adv 2020 Aug;6(35):eaba1394. [doi: 10.1126/sciadv.aba1394] [Medline: 32923622]
34. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J Intern Med 2004 Sep;256(3):240-246. [doi: 10.1111/j.1365-2796.2004.01380.x] [Medline: 15324367]
35. Lezak MD, Howieson, DB, Bigler ED, Tranel D. Neuropsychological Assessment: Oxford University Press; 2004. URL: https://global.oup.com/academic/product/neuropsychological-assessment-9780195395525?cc=it&lang=en& [accessed 2026-02-17]
36. Capitani E, Laiacona M. Outer and inner tolerance limits: their usefulness for the construction of norms and the standardization of neuropsychological tests. Clin Neuropsychol 2017 Oct 3;31(6-7):1219-1230. [doi: 10.1080/13854046.2017.1334830]
37. Magni E, Binetti G, Bianchetti A, Rozzini R, Trabucchi M. Mini - Mental State Examination: a normative study in Italian elderly population. Euro J of Neurology 1996 May;3(3):198-202. [doi: 10.1111/j.1468-1331.1996.tb00423.x]
38. Tuena C, Serino S, Goulene KM, Pedroli E, Stramba-Badiale M, Riva G. Bodily and visual-cognitive navigation aids to enhance spatial recall in mild cognitive impairment. J Alzheimers Dis 2024;99(3):899-910. [doi: 10.3233/JAD-240122] [Medline: 38701150]
39. Tuena C, Serino S, Pedroli E, Stramba-Badiale M, Riva G, Repetto C. Building embodied spaces for spatial memory neurorehabilitation with virtual reality in normal and pathological aging. Brain Sci 2021 Aug 14;11(8):1067. [doi: 10.3390/brainsci11081067] [Medline: 34439686]
40. Coughlan G, Laczó J, Hort J, Minihane AM, Hornberger M. Spatial navigation deficits - overlooked cognitive marker for preclinical Alzheimer disease? Nat Rev Neurol 2018 Aug;14(8):496-506. [doi: 10.1038/s41582-018-0031-x] [Medline: 29980763]
41. Serino S, Stramba-Badiale C, Magni G, Castegnaro A, Riva G, Tuena C. The complementary role of egocentric and allocentric spatial navigation tasks for the diagnosis of Alzheimer's disease: a diagnostic meta-analysis. Neurosci Biobehav Rev 2025 Nov;178:106379. [doi: 10.1016/j.neubiorev.2025.106379] [Medline: 40983243]
42. Wirth S. A place with a view: a first-person perspective in the hippocampal memory space. Hippocampus 2023 May;33(5):658-666. [doi: 10.1002/hipo.23537] [Medline: 37013360]
43. Wang C, Chen X, Lee H, et al. Egocentric coding of external items in the lateral entorhinal cortex. Science 2018 Nov 23;362(6417):945-949. [doi: 10.1126/science.aau4940] [Medline: 30467169]
44. Hinman JR, Chapman GW, Hasselmo ME. Neuronal representation of environmental boundaries in egocentric coordinates. Nat Commun 2019;10(1):1-8. [doi: 10.1038/s41467-019-10722-y]
45. Long X, Bush D, Deng B, Burgess N, Zhang SJ. Allocentric and egocentric spatial representations coexist in rodent medial entorhinal cortex. Nat Commun 2025 Jan 3;16(1):356. [doi: 10.1038/s41467-024-54699-9] [Medline: 39753542]
46. Burgess N, Maguire EA, O'Keefe J. The human hippocampus and spatial and episodic memory. Neuron 2002 Aug 15;35(4):625-641. [doi: 10.1016/s0896-6273(02)00830-9] [Medline: 12194864]
47. Baddeley A, Wilson BA. Prose recall and amnesia: implications for the structure of working memory. Neuropsychologia 2002;40(10):1737-1743. [doi: 10.1016/s0028-3932(01)00146-4] [Medline: 11992661]
48. Blanco Martín E, Ugarriza Serrano I, Elcoroaristizabal Martín X, et al. Dysexecutive syndrome in amnesic mild cognitive impairment: a multicenter study. BMC Neurol 2016 Jun 4;16(1):88. [doi: 10.1186/s12883-016-0607-2] [Medline: 27260328]
49. Craik FIM, Rose NS. Memory encoding and aging: a neurocognitive perspective. Neurosci Biobehav Rev 2012 Aug;36(7):1729-1739. [doi: 10.1016/j.neubiorev.2011.11.007] [Medline: 22155274]

XSL•FO
RenderX

50. Greene NR, Naveh-Benjamin M. Adult age-related changes in the specificity of episodic memory representations: a review and theoretical framework. Psychol Aging 2023 Mar;38(2):67-86. [doi: 10.1037/pag0000724] [Medline: 36729498]
51. Plancher G, Gyselinck V, Nicolas S, Piolino P. Age effect on components of episodic memory and feature binding: a virtual reality study. Neuropsychology 2010 May;24(3):379-390. [doi: 10.1037/a0018680] [Medline: 20438215]
52. Gellersen HM, Coughlan G, Hornberger M, Simons JS. Memory precision of object-location binding is unimpaired in APOE ε4-carriers with spatial navigation deficits. Brain Commun 2021;3(2):fcab087. [doi: 10.1093/braincomms/fcab087] [Medline: 33987536]
53. Byrne P, Becker S, Burgess N. Remembering the past and imagining the future: a neural model of spatial memory and imagery. Psychol Rev 2009;114(2):340-375. [doi: 10.1037/0033-295X.114.2.340]
54. Besson M, Martin JR. Centrophobism/thigmotaxis, a new role for the mushroom bodies in Drosophila. J Neurobiol 2005 Feb 15;62(3):386-396. [doi: 10.1002/neu.20111] [Medline: 15547935]
55. Day LB, Weisend M, Sutherland RJ, Schallert T. The hippocampus is not necessary for a place response but may be necessary for pliancy. Behav Neurosci 1999;113(5):914-924. [doi: 10.1037/0735-7044.113.5.914]
56. Lee AS, André JM, Pittenger C. Lesions of the dorsomedial striatum delay spatial learning and render cue-based navigation inflexible in a water maze task in mice. Front Behav Neurosci 2014;8(42):42. [doi: 10.3389/fnbeh.2014.00042] [Medline: 24592223]
57. Kallai J, Makany T, Csatho A, et al. Cognitive and affective aspects of thigmotaxis strategy in humans. Behav Neurosci 2007 Feb;121(1):21-30. [doi: 10.1037/0735-7044.121.1.21] [Medline: 17324048]
58. Walz N, Mühlberger A, Pauli P. A human open field test reveals thigmotaxis related to agoraphobic fear. Biol Psychiatry 2016 Sep 1;80(5):390-397. [doi: 10.1016/j.biopsych.2015.12.016] [Medline: 26876946]
59. O'Keefe J, Nadel L. The Hippocampus as a Cognitive Map: Clarendon Press; 1978. URL: https://www.cmor-faculty.rice.edu/~cox/neuro/HCMComplete.pdf [accessed 2026-02-17]
60. Hardcastle K, Ganguli S, Giocomo LM. Environmental boundaries as an error correction mechanism for grid cells. Neuron 2015 May 6;86(3):827-839. [doi: 10.1016/j.neuron.2015.03.039] [Medline: 25892299]
61. Segen V, Ying J, Morgan E, Brandon M, Wolbers T. Path integration in normal aging and Alzheimer's disease. Trends Cogn Sci 2022 Feb;26(2):142-158. [doi: 10.1016/j.tics.2021.11.001] [Medline: 34872838]
62. Zenodo. Dataset for Egocentric Spatial Memory Deficit in Amnestic Mild Cognitive Impairment Revealed Through Virtual Reality: Cross-Sectional Study. 2026. URL: https://zenodo.org/records/15119995

Abbreviations

AD: Alzheimer disease
aMCI: amnestic mild cognitive impairment
aMCImd: amnestic mild cognitive impairment, multiple-domain
aMCIsd: amnestic mild cognitive impairment, single-domain
AUC: area under the curve
HC: healthy control
MCI: mild cognitive impairment
MMSE: Mini-Mental Score Examination
MWM: Morris water maze
OL: object-location
VR: virtual reality

Edited by H LaMonica; submitted 23.Jun.2025; peer-reviewed by CY Cheong, H Schwarb; revised version received 05.Jan.2026; accepted 05.Jan.2026; published 25.Feb.2026.

Please cite as:
Tuena C, Serino S, Stramba-Badiale C, Biondi D, Goulene KM, Stramba-Badiale M, Riva G
Egocentric Spatial Memory Deficit in Amnestic Mild Cognitive Impairment Revealed Through Virtual Reality: Cross-Sectional Study
JMIR Aging 2026;9:e79224
URL: https://aging.jmir.org/2026/1/e79224
doi:10.2196/79224

© Cosimo Tuena, Silvia Serino, Chiara Stramba-Badiale, Diana Biondi, Karine Marie Goulene, Marco Stramba-Badiale, Giuseppe Riva. Originally published in JMIR Aging (https://aging.jmir.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited.

XSL•FO
RenderX

The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Long-Term Effects of Mobile-Based Metamemory Cognitive Training in Older Adults With Mild Cognitive Impairment: 15-Month Prospective Single-Arm Longitudinal Study

Jung-In Lim[1], MA; Yeeun Byeon[2], MS; Sunyoung Kang[3], MD; Hyeonjin Kim[4], PhD; Keun You Kim[3], MD, PhD; Lukas Stenzel[5], PhD; So Yeon Jeon[4,6], MD, PhD; Jun-Young Lee[4,6], MD, PhD

[1]Interdisciplinary Program in Cognitive Science, Seoul National University, Seoul, Republic of Korea

[2]Department of Medical Device Development, Seoul National University College of Medicine, Seoul, Republic of Korea

[3]Department of Psychiatry, Institute of Behavioral Science in Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea

[4]Department of Psychiatry, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul, Republic of Korea

[5]Cogthera GmbH, Munich, Germany

[6]Department of Psychiatry, Seoul National University College of Medicine, Seoul, Republic of Korea

Corresponding Author:
Jun-Young Lee, MD, PhD
Department of Psychiatry
Seoul Metropolitan Government-Seoul National University Boramae Medical Center
20, Boramae-ro 5-gil, Dongjak-gu
Seoul,
Republic of Korea
Phone: 82 2 870 2462
Email: benji@snu.ac.kr

Abstract

Background: Mild cognitive impairment (MCI) is an intermediate state between normal aging and dementia, characterized by subjective cognitive decline and objective memory impairment. Cognitive training has consistently shown short-term benefits for individuals with MCI, but evidence on the long-term effectiveness is extremely limited. Given the progressive nature of MCI and the need for sustainable strategies to delay cognitive decline, research on the long-term impact of cognitive training is necessary and timely. Mobile-based platforms offer a promising solution by enhancing accessibility and adherence, but their durability of effect over extended periods remains underexplored.

Objective: This study aimed to evaluate the long-term effects of a mobile-based cognitive training app on the cognitive function of older adults with MCI.

Methods: In total, 28 older adults with MCI used Cogthera, a mobile cognitive training app based on metamemory training. Participants completed 2 training sessions daily for 3 months, and 9 (32%) continued for an additional 12 months. Cognitive function and quality of life were assessed using the Alzheimer's Disease Assessment Scale-Cognitive Subscale 14 and EQ-5D-5L.

Results: Cognitive function improved over 15 months, as measured by Alzheimer's Disease Assessment Scale-Cognitive Subscale ($F_{2,35.56}$=7.08; P=.003). EQ-5D-5L scores increased at 3 months but did not show sustained change at 15 months ($F_{2,42.14}$=3.40; P=.04). Greater cognitive improvements were associated with younger age, higher functional status, and lower baseline cognitive function.

Conclusions: This study showed that long-term use of a mobile-based metamemory cognitive training app was associated with cognitive improvements over 15 months. Although limited by the small sample size and the absence of a control group, these findings suggest potential for mobile cognitive training as a sustainable intervention that warrants validation in larger trials.

(JMIR Aging 2026;9:e81648) doi:10.2196/81648

KEYWORDS

Alzheimer disease; cognitive training; digital technology; mild cognitive impairment; metamemory

XSL•FO
RenderX

Introduction

Background

Mild cognitive impairment (MCI) is an intermediate stage between normal aging and Alzheimer disease (AD). It has emerged as a key target for early intervention to delay AD. MCI is characterized by subjective cognitive decline along with objective memory impairment without functional impairment [1]. The prevalence of MCI is 6.7% among individuals aged between 60 and 64 years and 25.5% among those aged between 80 and 84 years [2]. In addition, approximately 10% of individuals with MCI progress to AD annually [3].

Nonpharmacological interventions, such as physical activity, social engagement, and cognitive training, are widely recommended to delay AD onset in MCI [4-7]. Cognitive training has demonstrated superior efficacy in improving cognitive function in short-term treatment [8-11]. Despite these short-term benefits, few studies have examined whether cognitive training produces sustained effects beyond 1 year because of methodological difficulties [12]. Traditional cognitive interventions also rely on in-person, paper-based sessions, which can limit accessibility and adherence [13]. Digital cognitive training based on a mobile app enables home-based, self-paced engagement and has demonstrated greater participation [14].

We have previously reported the short-term efficacy of metamemory training (MMT) [15,16]. Metamemory refers to an individual's awareness and understanding of their memory functions, including their contents and processes [17,18]. MMT is based on this concept and has been shown to be effective in teaching mnemonic strategies and improving cognitive function [19-22]. However, its long-term efficacy may need to be investigated, especially with a mobile cognitive training app.

Objectives

This study aimed to examine the long-term effects of a cognitive training app based on MMT in older adults with MCI. We hypothesized that prolonged use of the app would lead to sustained improvements in cognitive function.

Methods

Participants

Participants were recruited from 1 memory clinic. Inclusion criteria were as follows: (1) aged between 55 and 85 years; (2) diagnosed with MCI by trained psychiatrists or neurologists according to the criteria proposed by Petersen [23]; (3) ownership of a personal smartphone and no difficulties using mobile apps; (4) if taking cognitive enhancers (acetylcholinesterase inhibitors or memantine), a stable dose maintained for at least 12 weeks before randomization; (5) availability of a caregiver who spends more than 8 hours per week with the participant and agrees to assist with follow-up and clinical evaluations; (6) ability to independently make phone calls to a caregiver using a smartphone; (7) no difficulties with reading or writing in Korean; and (8) adequate vision and hearing to participate in the clinical trial. MCI diagnosis was based on the following criteria by Petersen [23]: (1) reported concerns regarding cognitive changes, (2) impairment in 1 or more cognitive domains, (3) preservation of independence in functional abilities, and (4) absence of dementia [24]. Participants underwent cognitive assessments conducted by clinical psychologists using the Clinical Dementia Rating (CDR) and the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Assessment Battery (CERAD-NP). Eligibility criteria required a global CDR score of 0.5 and performance at least 1.0 SD below the mean in 1 or more memory domains of the CERAD-NP.

Exclusion criteria included the presence of serious physical illness or psychiatric disorders that could interfere with study procedures; the use of medications known to affect cognitive function (except those taken consistently for at least 3 months); and any neurological or medical conditions associated with cognitive decline other than MCI, such as stroke, central nervous system infection, head trauma, alcohol dependence, or depression. Of the 40 participants screened, 10 (25%) were excluded, 1 (3%) dropped out, and 1 (3%) experienced a stroke. Data from 28 (70%) participants were analyzed.

Ethical Considerations

This study was reviewed and approved by the institutional review board of Seoul Metropolitan Government-Seoul National University Boramae Medical Center (20-2022-48). All participants provided informed consent before participation. All collected data were deidentified and stored securely, and no personally identifiable information was accessible to the research team. Participants received KRW 200,000 (US $136) as compensation for their participation.

Intervention

Participants underwent cognitive training using Cogthera, a smartphone-based app based on MMT developed by Youn et al [16]. The structure and content of the Cogthera intervention, including its implementation of metamemory-based strategies, have been previously described in detail. The program was designed to enhance key cognitive functions, including attention, imagery, and association, which support effective memory encoding and retrieval. Attention training helped participants focus and concentrate on target information to facilitate deeper encoding. Imagery training encouraged vivid visualization to reinforce memory consolidation and retrieval. Association training promoted the integration of new information into existing semantic memory networks. Throughout the training process, Cogthera provides personalized feedback and dynamically adjusts task difficulty, enabling users to observe and evaluate their cognitive processes independently.

A 15-month single-arm longitudinal study was conducted to investigate the long-term efficacy of mobile-based cognitive training. Participants completed a 3-month cognitive training program using Cogthera, with 9 (32%) of the 28 participants continuing an additional 12 months. The training program consisted of 2 daily sessions, 7 days per week, with each session lasting approximately 15 minutes. The first session comprised 3 core cognitive exercises targeting attention, imagery, and association. The second session included 4 additional exercises, selected from 9 available options, excluding those already used

XSL•FO
RenderX

in the first session. A personalized algorithm determined daily exercise composition to optimize engagement. All participants were able to use the Cogthera program independently without external assistance. The detailed content of each exercise can be found in Multimedia Appendix 1.

Measures

Alzheimer's Disease Assessment Scale-Cognitive Subscale 14

The Alzheimer's Disease Assessment Scale-Cognitive Subscale 14 (ADAS-Cog 14) was used to assess the severity of cognitive dysfunction [25]. This measure consists of fourteen tasks: (1) word recall, (2) commands, (3) constructional praxis, (4) delayed word recall, (5) naming, (6) ideational praxis, (7) orientation, (8) word recognition, (9) maze, (10) number cancelation, (11) remembering instructions, (12) comprehension, (13) word finding difficulty, and (14) spoken language ability. The ADAS-Cog 14 subdomains were categorized into 3 cognitive domains: memory, language, and praxis [26]. The total ADAS-Cog 14 score ranges from 0 to 90 points, with higher scores indicating greater cognitive impairment, as the score reflects the number of errors made across tasks.

CERAD-NP Assessment

The CERAD-NP was administered to assess cognitive function [27,28]. This battery included the following neuropsychological tests: (1) verbal fluency, (2) Boston Naming Test, (3) Mini-Mental State Examination, (4) word list learning, (5) constructional praxis, (6) word list recall, (7) word list recognition, (8) constructional praxis recall, (9) Trail Making Test, and (10) Stroop test. The CERAD-NP total score was calculated as the sum of raw scores, with higher scores indicating better cognitive function.

CDR Scale

The CDR scale was used to stage dementia severity [29]. This scale is informed by semistructured interviews conducted with both participants and informants, covering 6 domains: memory, orientation, judgment and problem-solving, community affairs, home and hobbies, and personal care. Each domain is scored from 0 to 3, with the scores used in calculating the global CDR score and the CDR-Sum of Boxes (CDR-SB). The global CDR score is an ordinal scale ranked from 0 to 3 as follows: 0=no cognitive impairment, 0.5=questionable or MCI, 1=mild dementia, 2=moderate dementia, 3=severe dementia. The CDR-SB is a continuous measure ranging from 0 to 18, calculated by summing the individual domain scores. Higher scores on both the global CDR and CDR-SB indicate greater cognitive and functional impairment.

EQ-5D-5L Scale

The EQ-5D-5L was administered to evaluate health-related quality of life [30]. This measure consists of five dimensions: (1) mobility, (2) self-care, (3) usual activities, (4) pain and discomfort, and (5) anxiety and depression. Each dimension is rated on a 5-level scale, with response options ranging from *no problem* to *extreme problem*. Responses were converted into a single index score using a national-specific value set [31]. This value set, derived from stated preference data collected from the general population, assigns weights to each health dimension level. The index score ranges from 0 to 1, with higher scores indicating better overall health status. In contrast, a lower score on each subdimension reflects a higher level of health in that specific domain.

Data Collection

Participants who provided informed consent were screened using global CDR and CERAD-NP. During the initial visit, participants underwent cognitive function assessments using ADAS-Cog 14 and completed self-reported questionnaires, such as EQ-5D-5L. After completing these assessments, they were provided with a smartphone preinstalled with Cogthera. Follow-up data on ADAS-Cog 14 and EQ-5D-5L were collected after 3 months of using Cogthera. In addition, further assessments were conducted on 9 (32%) of the 28 participants who continued using Cogthera for 15 months. Training adherence was assessed using compliance, which was defined as the proportion of completed sessions relative to the total number of assigned sessions. The assigned frequency was 2 sessions per day, and compliance was calculated separately for the initial 3-month period and the 12-month extension.

Statistical Analysis

Linear mixed model analyses were performed to assess changes in ADAS-Cog 14 and EQ-5D-5L scores over time while accounting for both intraindividual and interindividual variations in longitudinal data. The models included time (baseline, 3 mo, and 15 mo) as a fixed factor and participant as a random effect, allowing individual variability in intercepts. To evaluate the overall effect of time, type III ANOVA was conducted.

To identify explanatory variables predicting intervention effects, the least absolute shrinkage and selection operator (LASSO) regression was used. This approach was used to identify the most relevant predictors while minimizing overfitting [32]. The outcome variable was the change in ADAS-Cog 14 at follow-up, with LASSO models including age, sex, education, CDR-SB, baseline ADAS-Cog 14 scores, and baseline EQ-5D-5L scores as predictors. All variables were standardized by centering each variable around the mean and scaling by the SD. The regularization parameter, λ, was optimized using 9-fold cross-validation to balance model complexity and predictive performance. The optimal λ value was then applied to estimate the coefficients of the selected predictors.

To contextualize cognitive changes observed in this study, descriptive data from the placebo cohort of EXPEDITION studies were used as a historical reference [33,34]. EXPEDITION and EXPEDITION-2 were multicenter, double-blind, phase 3 trials of solanezumab, including 663 patients with MCI treated with a placebo. Cognitive function was assessed using ADAS-Cog 14 at baseline and at 6 follow-up points every 3 months over 18 months. Welch *t* test (2-tailed) was conducted to compare data from this study with the EXPEDITION placebo cohort at baseline, 3 months, and 15 months, accounting for unequal variances and sample sizes.

All statistical analyses were conducted using R software (version 4.4.1; R Foundation for Statistical Computing). Statistical significance was set at $P<.05$, with all tests being 2-tailed. For

XSL•FO
RenderX

the post hoc analysis of the linear mixed model, Bonferroni-adjusted *P* values were reported. Effect sizes for the fixed time effects were also reported, with partial eta-squared (η^2_p) computed from the corresponding *F* statistics. Missing data were addressed through complete case analysis without imputation. To ensure analytic independence, all data access, monitoring, and statistical analyses were conducted exclusively by authors not affiliated with the company. Company-affiliated authors had no access to raw data and did not participate in data cleaning, analytic decisions, or interpretation of results.

Results

Demographic Characteristics

A total of 28 participants were included in the analysis (Table 1). The mean age of participants was 72.8 (SD 6.7) years, and the mean education level was 11.4 (SD 5.0) years. The sample comprised a higher proportion of female individuals (20/28, 71%) and was predominantly composed of nondrinkers (27/28, 96%) and nonsmokers (26/28, 93%). The mean CDR-SB score was 2.0 (SD 1.1), ranging from 0.5 to 4.0.

Table 1. Baseline demographic data for the participants (N=28).

Demographic characteristic	Values
Age (y), mean (SD)	72.8 (6.7)
Sex, n (%)	
Female	20 (71)
Male	8 (29)
Education (y), mean (SD)	11.4 (5.0)
Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Assessment Battery (*z* score[a]), mean (SD)	
Word list learning	–0.5 (0.9)
Word list recall	–1.6 (1.0)
Word list recognition	–1.7 (1.4)
Constructional praxis recall	–1.3 (0.9)
Clinical Dementia Rating-Sum of Boxes (score), mean (SD)	2.0 (1.1)

[a]A standardized score indicating how many SDs a value is from the mean.

Training Adherence

During the initial 3-month training period, participants completed an average of 85.3% (SD 23.6%) of assigned sessions. Among 9 (32%) of the 28 participants who continued for an additional 12 months, mean compliance during the extension period was 51.7% (SD 25.1%).

Intervention Effects of Cogthera

Linear mixed model analysis revealed a significant decrease in total ADAS-Cog 14 scores over time, indicating improved cognitive function ($F_{2,35.56}$=7.08; *P*=.003; η^2_p=0.28; Table 2). Specifically, a significant decrease in total ADAS-Cog 14 scores was observed over 15 months (Bonferroni-adjusted *P*=.001), while the change over 3 months was not significant (Bonferroni-adjusted *P*=.27; Table 3).

Table 2. Effects of time on cognitive function and quality of life.

	Baseline (n=28), mean (SD)	3 mo (n=28), mean (SD)	15 mo (n=9), mean (SD)	*F* test (*df*)	*P* value
Alzheimer's Disease Assessment Scale-Cognitive Subscale[a]					
Total score	28.18 (8.21)	27.07 (7.73)	25.56 (8.28)	7.08 (2, 35.56)	.003
Memory[b]	22.25 (6.24)	21.86 (6.29)	20.44 (8.56)	2.53 (2, 35.83)	.09
Language[c]	0.86 (0.93)	0.64 (0.83)	0.56 (0.73)	1.64 (2, 35.83)	.21
Praxis[d]	2.32 (1.19)	1.96 (1.10)	1.33 (1.00)	7.11 (2, 39.33)	.002
EQ-5D-5L					
Index score	0.81 (0.08)	0.85 (0.04)	0.84 (0.07)	3.40 (2, 42.14)	.04

[a]Lower scores represent better performance.

[b]Sum of word recall, delayed word recall, orientation, and word recognition task scores.

[c]Sum of naming, remembering instructions, comprehension, word finding difficulty, and spoken language ability task scores.

[d]Sum of commands, constructional praxis, and ideational praxis task scores.

XSL•FO
RenderX

Table 3. Longitudinal changes in cognitive function and quality of life.

	Estimate (SE)	*t* test (*df*)	*P* value[a]
Alzheimer's Disease Assessment Scale-Cognitive Subscale			
Total score			
3 months vs baseline	–1.11 (0.72)	–1.53 (35.06)	.27
15 months vs baseline	–4.33 (1.16)	–3.74 (35.96)	.001
Memory			
3 months vs baseline	–0.39 (0.68)	–0.58 (35.17)	.99
15 months vs baseline	–2.42 (1.08)	–2.25 (36.38)	.06
Language			
3 months vs baseline	–0.21 (0.12)	–1.81 (34.65)	.16
15 months vs baseline	–0.09 (0.19)	–0.47 (36.83)	.99
Praxis			
3 months vs baseline	–0.36 (0.23)	–1.53 (35.78)	.27
15 months vs baseline	–1.35 (0.36)	–3.76 (42.63)	.001
EQ-5D-5L			
Index score			
3 months vs baseline	0.04 (0.02)	2.58 (35.29)	.03
15 months vs baseline	0.03 (0.02)	1.27 (49.19)	.42

[a]Bonferroni-adjusted *P* values.

In the ADAS-Cog 14 subdomain analysis, praxis scores, representing executive function, showed a significant decline over time ($F_{2,39.33}$=7.11; *P*=.002; η^2_p=0.27), particularly over 15 months (Bonferroni-adjusted *P*=.001). The reduction in memory scores was marginal ($F_{2,35.83}$=2.53; *P*=.09; η^2_p=0.12), with a trend-level decrease observed at 15 months (Bonferroni-adjusted *P*=.06). The overall effect of time on language scores was not significant.

A significant change in the EQ-5D-5L index score was observed over time ($F_{2,42.14}$=3.40; *P*=.04; η^2_p=0.14), with a statistically significant increase from baseline to 3 months (Bonferroni-adjusted *P*=.03). Among the EQ-5D-5L subdimensions, mobility showed a marginal overall effect of time ($F_{2,\ 39.05}$=2.59; *P*=.09; η^2_p=0.12), with a trend-level decrease observed at 3 months (Bonferroni-adjusted *P*=.06). Detailed results for the EQ-5D-5L subdimensions are provided in Tables S1 and S2 in Multimedia Appendix 2).

Factors Associated With Intervention Effects on Cognitive Function

According to the LASSO results, CDR-SB and baseline ADAS-Cog 14 scores predicted score changes over 3 months, while age, CDR-SB, and baseline ADAS-Cog 14 scores predicted score changes over 15 months. The multiple linear regression model, which included sex, CDR-SB, and baseline ADAS-Cog 14 scores, explained a significant proportion of the variance in score changes over 3 months (R^2=0.40; adjusted R^2=0.33; *P*=.01; Table 4). Specifically, lower CDR-SB scores (β=–0.51; *P*=.02), and higher baseline ADAS-Cog 14 scores significantly predicted greater improvements (β=0.61; *P*=.01).

XSL•FO
RenderX

Table 4. Factors predicting cognitive function improvements.

	Values, estimate (SE)	β	*t* test *(df)*	R^2 (adjusted R^2)	*P* value
Change over 3 months[a]				0.40 (0.33)	.01
Sex	2.60 (1.32)	0.69	1.97		.06
CDR-SB[b]	–1.78 (0.70)	–0.51	–2.53		.02
ADAS-Cog 14[c]	0.28 (0.09)	0.61	3.01		.01
Change over 15 months[d]				0.81 (0.70)	.03
Age	–0.44 (0.14)	–0.62	–3.23		.02
CDR-SB	–4.80 (1.39)	–1.10	–3.47		.02
ADAS-Cog 14	0.61 (0.14)	1.05	4.25		.01

[a]Score obtained by subtracting the Alzheimer's Disease Assessment Scale-Cognitive Subscale score at 3 months from the baseline Alzheimer's Disease Assessment Scale-Cognitive Subscale score.

[b]CDR-SB: Clinical Dementia Rating-Sum of Boxes.

[c]ADAS-Cog 14: Alzheimer's Disease Assessment Scale-Cognitive Subscale.

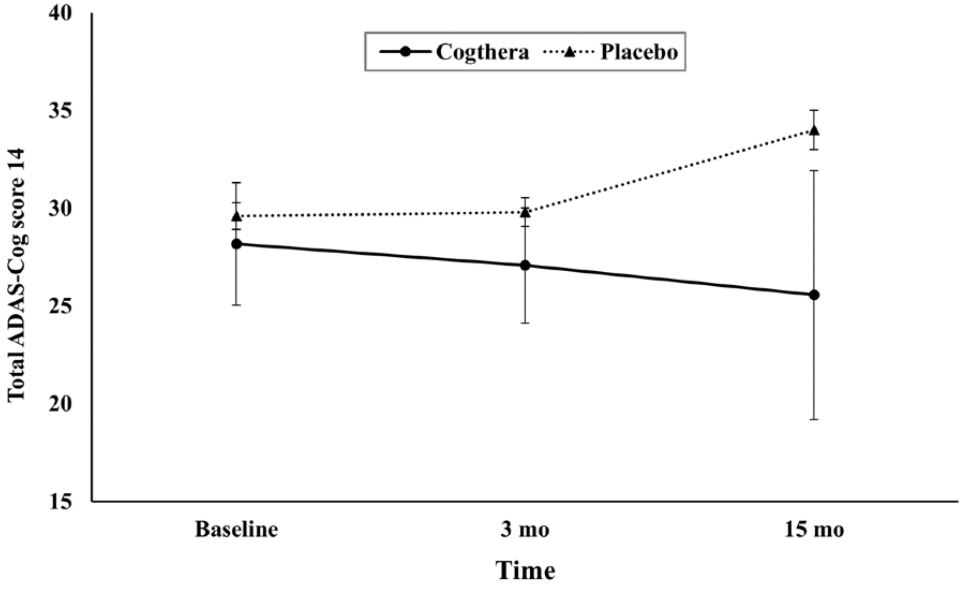
[d]Score obtained by subtracting the Alzheimer's Disease Assessment Scale-Cognitive Subscale score at 15 months from the baseline Alzheimer's Disease Assessment Scale-Cognitive Subscale score.

The regression model, which included age, CDR-SB, and baseline ADAS-Cog 14 scores, accounted for a significant proportion of the variance in score changes observed over 15 months (R^2=0.81; adjusted R^2=0.70; *P*=.03). Specifically, younger age (β=–0.62; *P*=.02), lower CDR-SB scores (β=–1.10; *P*=.02), and higher baseline ADAS-Cog 14 scores significantly predicted greater improvements (β=1.05; *P*=.01).

Comparison With Historical Placebo Data

The mean and SD data from a previously reported placebo cohort were used for comparison at each time point. No significant differences were observed between the Cogthera group and the placebo group at baseline ($t_{29.68}$=0.89; *P*=.38; Figure 1). At 3 months, ADAS-Cog 14 scores in the Cogthera group were marginally lower than those in the placebo group ($t_{30.63}$=1.81; *P*=.08). By 15 months, the difference became statistically significant, suggesting greater cognitive improvement in the Cogthera group ($t_{8.55}$=3.01; *P*=.02).

Figure 1. Comparison of cognitive function between the Cogthera group and the historical placebo group. Error bars represent 95% CIs of the mean Alzheimer's Disease Assessment Scale-Cognitive Subscale 14 (ADAS-Cog 14) scores at each time point.



XSL•FO
RenderX

Discussion

Principal Findings

This study examined the long-term effects of a mobile-based cognitive training app in older adults with MCI. ADAS-Cog 14 scores decreased over time, and EQ-5D-5L scores increased in the early phase of training. Age, CDR-SB, and baseline ADAS-Cog 14 scores were associated with the degree of cognitive change. Although comparisons with a historical placebo cohort are limited by differences in design and sample characteristics, the observed reduction in ADAS-Cog 14 scores over 15 months suggests a pattern of cognitive change associated with long-term training use.

The primary finding of this study is the reduction in ADAS-Cog 14 total, memory, and praxis scores associated with long-term use of Cogthera. While no meaningful change was observed at 3 months, a clearer reduction emerged at 15 months. This delayed pattern may reflect the cumulative effects of training, but it could also be influenced by practice effects, familiarity with test procedures, or natural fluctuations in cognitive performance. These possibilities underscore the need for cautious interpretation, particularly in the absence of a control group. Even so, the overall pattern is consistent with previous studies, suggesting that longer training durations are more likely to yield detectable cognitive improvements [35]. Unlike other ADAS-Cog 14 subdomains, language scores did not show meaningful change. The small numerical variation observed is likely to reflect measurement variability or stabilization rather than the true effect of the intervention. This aligns with previous research indicating that language tends to remain relatively preserved in the early stage of cognitive decline compared with other ADAS-Cog 14 subdomains [36,37].

Another finding is a rise in the EQ-5D-5L index score during the first 3 months of training. This early change is consistent with previous studies reporting short-term benefits in quality of life among individuals with MCI, but the effect did not persist at 15 months in this study [38-41]. This pattern suggests that changes in perceived well-being may be limited to the initial phase of training. Such early gains may also reflect nonspecific factors, including placebo effects or heightened engagement at the beginning of training.

Predictor analyses revealed that both short- and long-term improvements in cognitive function were associated with lower baseline cognitive function and higher functional status. Participants with higher baseline ADAS-Cog 14 scores, which indicate poorer initial cognitive performance, and those with lower CDR-SB scores, which indicate less functional impairment, tended to show larger reductions over time. Previous research has reported conflicting findings regarding the relationship between baseline cognitive performance and cognitive training effectiveness [42]. Two competing hypotheses have been proposed to explain this relationship: the compensation effect and the magnification effect. The compensation effect suggests that individuals with higher baseline cognitive performance benefit less from training due to limited room for improvement. In contrast, the magnification effect proposes that they benefit more by leveraging greater cognitive efficiency and plasticity. The findings of this study align with the compensation effect, as participants with lower baseline cognitive performance showed greater training-related gains. This pattern suggests that Cogthera may have provided compensatory cognitive stimulation. Such stimulation could have helped individuals with lower baseline cognitive performance make better use of their remaining cognitive resources. CDR-SB reflects the pathological progression of cognitive and functional decline and reliably differentiates MCI from very early AD while predicting future progression [43-45]. The association between lower CDR-SB scores and greater cognitive gains may suggest that individuals with preserved daily function but greater cognitive vulnerability were more responsive to training. These observations are broadly consistent with compensatory mechanisms, while other explanations remain possible.

Younger age was associated with modestly better cognitive outcomes over the 15-month period. One possible explanation is that younger participants may engage more consistently in long-term digital interventions. Previous research has shown that they tend to complete more training sessions and demonstrate higher adherence. Such patterns of adherence may help account for the cognitive changes observed in this group [46-48].

These findings suggest that MMT delivered through a mobile app may offer a scalable and accessible approach for individuals with MCI. Mobile apps provide a practical alternative to traditional cognitive training by allowing users to engage in cognitive exercises at their convenience. As memory decline has already begun in individuals with MCI, this approach may serve as a supportive digital complement to existing strategies. A potential strength of this study is the observation that cognitive changes became more evident with prolonged use, which may help inform future work on long-term digital training. Given the lack of known adverse effects, mobile-based cognitive training could represent a feasible and well-tolerated adjunctive approach for older adults with MCI. It may be particularly useful in cases where pharmacological treatments carry a higher risk of side effects.

Limitations

Despite the observational insights provided by this study, several limitations warrant careful consideration. First, the absence of a control group limits internal validity, as it is not possible to distinguish training-related changes from those that may occur naturally over time. Without contemporaneous controls, causal inferences cannot be drawn.

Second, although 28 participants initiated the study, only 9 (32%) continued to the 15-month assessment. Because long-term participation was voluntary, those who remained may differ in motivation or engagement. However, baseline comparisons between participants who continued to 15 months and those who participated only in the initial 3-month period showed no significant differences in demographic characteristics, cognitive measures, quality of life scores, or training adherence, as presented in Table S3 in Multimedia Appendix 2. In addition, baseline-to–3-month change scores did not differ between the 2 groups, indicating that early improvement was not greater

XSL•FO
RenderX

among those who remained in the study. These findings suggest that systematic baseline differences were limited, although unmeasured factors related to motivation or persistence may still have influenced long-term retention. The small number of long-term participants also restricts the generalizability of the 15-month findings.

Finally, the overall sample size was small, constraining statistical power and limiting the ability to detect subtle effects. Future studies with larger, randomized, and more diverse samples will be essential to determine whether long-term mobile-based cognitive training produces reliable and sustained effects in individuals with MCI.

Conclusions

This study suggests that a mobile-based MMT app may offer supportive benefits for cognitive function in older adults with MCI. Although methodological limitations constrain definitive interpretation, the observed patterns indicate that mobile-based cognitive training could be a feasible and accessible approach for individuals seeking strategies to maintain cognitive health. Continued evaluation in larger, controlled studies will help determine the extent and durability of these potential effects.

Funding

This work was supported by the Technology Innovation Program (or Industrial Strategic Technology Development Program—the bioindustry technology development; grant 20018143, development and clinical application of digital therapeutic devices user experience, untact clinical cloud platforms, and real-world evidence technologies based on the brain and body of the elderly) funded by the Ministry of Trade, Industry, and Energy (South Korea).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

YB is employed by Emocog Inc. SK was employed in the R&D department of Emocog Inc. during the conduct of the study and remained employed until August 31, 2025. JYL is the chief executive officer of Emocog Inc. LS is the chief operating officer of Cogthera GmbH, a subsidiary of Emocog Inc. Emocog Inc owns the app evaluated in this study and has a commercial interest in its future distribution. Cogthera GmbH is responsible for the commercialization and planned distribution of a modified version of this app in Germany. However, YB, SK, JYL, and LS were not involved in the data collection or analysis, which was conducted independently by other authors. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Description of the cognitive exercises included in Cogthera.
[XLSX File (Microsoft Excel File), 11 KB - aging_v9i1e81648_app1.xlsx]

Multimedia Appendix 2
Supplementary EQ-5D-5L subdimension results and baseline user comparisons.
[DOCX File , 29 KB - aging_v9i1e81648_app2.docx]

References

1. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 1999 Mar 01;56(3):303-308. [doi: 10.1001/archneur.56.3.303] [Medline: 10190820]
2. Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, et al. Practice guideline update summary: mild cognitive impairment [RETIRED]: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology 2018 Jan 16;90(3):126-135 [FREE Full text] [doi: 10.1212/WNL.0000000000004826] [Medline: 29282327]
3. Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia--meta-analysis of 41 robust inception cohort studies. Acta Psychiatr Scand 2009 Apr 03;119(4):252-265. [doi: 10.1111/j.1600-0447.2008.01326.x] [Medline: 19236314]
4. Baker LD, Frank LL, Foster-Schubert K, Green PS, Wilkinson CW, McTiernan A, et al. Effects of aerobic exercise on mild cognitive impairment: a controlled trial. Arch Neurol 2010 Jan 01;67(1):71-79 [FREE Full text] [doi: 10.1001/archneurol.2009.307] [Medline: 20065132]
5. Suzuki T, Shimada H, Makizako H, Doi T, Yoshida D, Tsutsumimoto K, et al. Effects of multicomponent exercise on cognitive function in older adults with amnestic mild cognitive impairment: a randomized controlled trial. BMC Neurol 2012 Oct 31;12:128 [FREE Full text] [doi: 10.1186/1471-2377-12-128] [Medline: 23113898]
6. Li H, Li J, Li N, Li B, Wang P, Zhou T. Cognitive intervention for persons with mild cognitive impairment: a meta-analysis. Ageing Res Rev 2011 Apr;10(2):285-296. [doi: 10.1016/j.arr.2010.11.003] [Medline: 21130185]

XSL•FO
RenderX

7. Hughes TF, Flatt JD, Fu B, Chang CC, Ganguli M. Engagement in social activities and progression from mild to severe cognitive impairment: the MYHAT study. Int Psychogeriatr 2013 Apr;25(4):587-595 [FREE Full text] [doi: 10.1017/S1041610212002086] [Medline: 23257280]
8. Gavelin HM, Dong C, Minkov R, Bahar-Fuchs A, Ellis KA, Lautenschlager NT, et al. Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. Ageing Res Rev 2021 Mar;66:101232. [doi: 10.1016/j.arr.2020.101232] [Medline: 33249177]
9. Intzandt B, Vrinceanu T, Huck J, Vincent T, Montero-Odasso M, Gauthier CJ, et al. Comparing the effect of cognitive vs. exercise training on brain MRI outcomes in healthy older adults: a systematic review. Neurosci Biobehav Rev 2021 Sep;128:511-533. [doi: 10.1016/j.neubiorev.2021.07.003] [Medline: 34245760]
10. Raimo S, Cropano M, Gaita M, Maggi G, Cavallo ND, Roldan-Tapia MD, et al. The efficacy of cognitive training on neuropsychological outcomes in mild cognitive impairment: a meta-analysis. Brain Sci 2023 Oct 25;13(11):1510 [FREE Full text] [doi: 10.3390/brainsci13111510] [Medline: 38002471]
11. Hampstead BM, Stringer AY, Iordan AD, Ploutz-Snyder R, Sathian K. Toward rational use of cognitive training in those with mild cognitive impairment. Alzheimers Dement 2023 Mar 06;19(3):933-945 [FREE Full text] [doi: 10.1002/alz.12718] [Medline: 35791724]
12. Belleville S, Cuesta M, Bier N, Brodeur C, Gauthier S, Gilbert B, et al. Five-year effects of cognitive training in individuals with mild cognitive impairment. Alzheimers Dement (Amst) 2024 Sep 06;16(3):e12626. [doi: 10.1002/dad2.12626] [Medline: 39246830]
13. Chan AT, Ip RT, Tran JY, Chan JY, Tsoi KK. Computerized cognitive training for memory functions in mild cognitive impairment or dementia: a systematic review and meta-analysis. NPJ Digit Med 2024 Jan 03;7(1):1 [FREE Full text] [doi: 10.1038/s41746-023-00987-5] [Medline: 38172429]
14. Rossetto F, Isernia S, Realdon O, Borgnis F, Blasi V, Pagliari C, et al. A digital health home intervention for people within the Alzheimer's disease continuum: results from the Ability-TelerehABILITation pilot randomized controlled trial. Ann Med 2023 Dec 17;55(1):1080-1091 [FREE Full text] [doi: 10.1080/07853890.2023.2185672] [Medline: 36929703]
15. Youn JH, Ryu SH, Lee JY, Park S, Cho SJ, Kwon H, et al. Brain structural changes after multi-strategic metamemory training in older adults with subjective memory complaints: a randomized controlled trial. Brain Behav 2019 May 27;9(5):e01278 [FREE Full text] [doi: 10.1002/brb3.1278] [Medline: 30916450]
16. Youn JH, Park S, Lee JY, Cho SJ, Kim J, Ryu SH. Cognitive improvement in older adults with mild cognitive impairment: evidence from a multi-strategic metamemory training. J Clin Med 2020 Jan 28;9(2):362 [FREE Full text] [doi: 10.3390/jcm9020362] [Medline: 32013035]
17. McDougall GJ, Kang J. Memory self-efficacy and memory performance in older males. Int J Mens Health 2003 May 1;2(2):131-147 [FREE Full text] [doi: 10.3149/jmh.0202.131] [Medline: 19043600]
18. Flavell JH. Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry. Am Psychol 1979 Oct;34(10):906-911. [doi: 10.1037/0003-066x.34.10.906]
19. Tarricone P. The Taxonomy of Metacognition. London, UK: Psychology Press; 2011.
20. Kim J, Shin E, Han K, Park S, Youn JH, Jin G, et al. Efficacy of smart speaker-based metamemory training in older adults: case-control cohort study. J Med Internet Res 2021 Feb 16;23(2):e20177 [FREE Full text] [doi: 10.2196/20177] [Medline: 33591276]
21. Sella E, Carbone E, Vincenzi M, Toffalini E, Borella E. Efficacy of memory training interventions targeting metacognition for older adults: a systematic review and meta-analysis. Aging Ment Health 2023 Apr 11;27(4):674-694. [doi: 10.1080/13607863.2022.2122931] [Medline: 36218025]
22. Youn JH, Lee JY, Kim S, Ryu SH. Multistrategic memory training with the metamemory concept in healthy older adults. Psychiatry Investig 2011 Dec;8(4):354-361 [FREE Full text] [doi: 10.4306/pi.2011.8.4.354] [Medline: 22216046]
23. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004 Sep 20;256(3):183-194 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01388.x] [Medline: 15324362]
24. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association Workgroups on Diagnostic Guidelines for Alzheimer’s Disease. Focus 2013 Jan 1;11(1):96-106. [doi: 10.1176/appi.focus.11.1.96]
25. Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11 Suppl 2:S13-S21. [Medline: 9236948]
26. Verma N, Beretvas SN, Pascual B, Masdeu JC, Markey MK, Alzheimer’s Disease Neuroimaging Initiative. New scoring methodology improves the sensitivity of the Alzheimer's Disease Assessment Scale-Cognitive subscale (ADAS-Cog) in clinical trials. Alzheimers Res Ther 2015 Nov 12;7(1):64 [FREE Full text] [doi: 10.1186/s13195-015-0151-0] [Medline: 26560146]
27. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 1989 Sep;39(9):1159-1165. [doi: 10.1212/wnl.39.9.1159] [Medline: 2771064]

XSL•FO
RenderX

28. Lee JH, Lee KU, Lee DY, Kim KW, Jhoo JH, Kim JH, et al. Development of the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment Packet (CERAD-K): clinical and neuropsychological assessment batteries. J Gerontol B Psychol Sci Soc Sci 2002 Jan 01;57(1):P47-P53. [doi: 10.1093/geronb/57.1.p47] [Medline: 11773223]
29. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993 Nov;43(11):2412-2414. [doi: 10.1212/wnl.43.11.2412-a] [Medline: 8232972]
30. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec 9;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
31. Kim SH, Ahn J, Ock M, Shin S, Park J, Luo N, et al. The EQ-5D-5L valuation study in Korea. Qual Life Res 2016 Jul 10;25(7):1845-1852. [doi: 10.1007/s11136-015-1205-2] [Medline: 26961008]
32. Hassan MM, Hassan MM, Yasmin F, Khan MA, Zaman S, Galibuzzaman, et al. A comparative assessment of machine learning algorithms with the Least Absolute Shrinkage and Selection Operator for breast cancer detection and prediction. Decis Anal J 2023 Jun;7:100245. [doi: 10.1016/j.dajour.2023.100245]
33. Doody RS, Thomas RG, Farlow M, Iwatsubo T, Vellas B, Joffe S, et al. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. N Engl J Med 2014 Jan 23;370(4):311-321. [doi: 10.1056/nejmoa1312889]
34. Liu-Seifert H, Siemers E, Price K, Han B, Selzler KJ, Henley D, et al. Cognitive impairment precedes and predicts functional impairment in mild Alzheimer's disease. J Alzheimers Dis 2015;47(1):205-214 [FREE Full text] [doi: 10.3233/JAD-142508] [Medline: 26402769]
35. Eggenberger P, Schumacher V, Angst M, Theill N, de Bruin ED. Does multicomponent physical exercise with simultaneous cognitive training boost cognitive performance in older adults? A 6-month randomized controlled trial with a 1-year follow-up. Clin Interv Aging 2015;10:1335-1349 [FREE Full text] [doi: 10.2147/CIA.S87732] [Medline: 26316729]
36. Raghavan N, Samtani MN, Farnum M, Yang E, Novak G, Grundman M, et al. The ADAS-Cog revisited: novel composite scales based on ADAS-Cog to improve efficiency in MCI and early AD trials. Alzheimers Dement 2013 Feb 05;9(1 Suppl):S21-S31 [FREE Full text] [doi: 10.1016/j.jalz.2012.05.2187] [Medline: 23127469]
37. Cogo-Moreira H, Krance SH, Wu CY, Lanctôt KL, Herrmann N, Black SE, et al. State, trait, and accumulated features of the Alzheimer's Disease Assessment Scale Cognitive Subscale (ADAS-Cog) in mild Alzheimer's disease. Alzheimers Dement (N Y) 2023 Mar 26;9(1):e12376 [FREE Full text] [doi: 10.1002/trc2.12376] [Medline: 36994227]
38. Maki Y, Yamaguchi T, Yamagami T, Murai T, Hachisuka K, Miyamae F, et al. The impact of subjective memory complaints on quality of life in community-dwelling older adults. Psychogeriatrics 2014 Sep 21;14(3):175-181. [doi: 10.1111/psyg.12056] [Medline: 25142381]
39. Giuli C, Papa R, Lattanzio F, Postacchini D. The effects of cognitive training for elderly: results from my mind project. Rejuvenation Res 2016 Dec;19(6):485-494 [FREE Full text] [doi: 10.1089/rej.2015.1791] [Medline: 26952713]
40. Han JW, Lee H, Hong JW, Kim K, Kim T, Byun HJ, et al. Multimodal cognitive enhancement therapy for patients with mild cognitive impairment and mild dementia: a multi- center, randomized, controlled, double-blind, crossover trial. J Alzheimers Dis 2017;55(2):787-796. [doi: 10.3233/JAD-160619] [Medline: 27802233]
41. Fu E, Compton M. Non-pharmacological interventions that improve quality of life in people with mild cognitive impairment. Biomed J Sci Tech Res 2021 Oct 28;39(4):31543-31547 [FREE Full text] [doi: 10.26717/BJSTR.2021.39.006338]
42. Shaw J, Hosseini SH. The effect of baseline performance and age on cognitive training improvements in older adults: a qualitative review. J Prev Alzheimers Dis 2021 Jan;8(1):100-109 [FREE Full text] [doi: 10.14283/jpad.2020.55] [Medline: 33336231]
43. Perneczky R, Hartmann J, Grimmer T, Drzezga A, Kurz A. Cerebral metabolic correlates of the clinical dementia rating scale in mild cognitive impairment. J Geriatr Psychiatry Neurol 2007 Jun 01;20(2):84-88. [doi: 10.1177/0891988706297093] [Medline: 17548777]
44. O'Bryant SE, Lacritz LH, Hall J, Waring SC, Chan W, Khodr ZG, et al. Validation of the new interpretive guidelines for the clinical dementia rating scale sum of boxes score in the national Alzheimer's coordinating center database. Arch Neurol 2010 Jun 01;67(6):746-749 [FREE Full text] [doi: 10.1001/archneurol.2010.115] [Medline: 20558394]
45. Tzeng RC, Yang YW, Hsu KC, Chang HT, Chiu PY. Sum of boxes of the clinical dementia rating scale highly predicts conversion or reversion in predementia stages. Front Aging Neurosci 2022 Sep 23;14:1021792 [FREE Full text] [doi: 10.3389/fnagi.2022.1021792] [Medline: 36212036]
46. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR Mhealth Uhealth 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]
47. Belleville S, Gilbert B, Fontaine F, Gagnon L, Ménard E, Gauthier S. Improvement of episodic memory in persons with mild cognitive impairment and healthy older adults: evidence from a cognitive intervention program. Dement Geriatr Cogn Disord 2006 Nov 3;22(5-6):486-499. [doi: 10.1159/000096316] [Medline: 17050952]
48. Turunen M, Hokkanen L, Bäckman L, Stigsdotter-Neely A, Hänninen T, Paajanen T, et al. Computer-based cognitive training for older adults: determinants of adherence. PLoS One 2019 Jul 10;14(7):e0219541 [FREE Full text] [doi: 10.1371/journal.pone.0219541] [Medline: 31291337]

XSL•FO
RenderX

Abbreviations

AD: Alzheimer disease
ADAS-Cog 14: Alzheimer's Disease Assessment Scale-Cognitive Subscale 14
CDR: Clinical Dementia Rating
CDR-SB: Clinical Dementia Rating-Sum of Boxes
CERAD-NP: Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Assessment Battery
LASSO: least absolute shrinkage and selection operator
MCI: mild cognitive impairment
MMT: metamemory training

Edited by M O'Connell; submitted 31.Jul.2025; peer-reviewed by AS Rigaud, P Kooncumchoo; comments to author 19.Oct.2025; accepted 05.Dec.2025; published 02.Jan.2026.

Please cite as:
Lim JI, Byeon Y, Kang S, Kim H, Kim KY, Stenzel L, Jeon SY, Lee JY
Long-Term Effects of Mobile-Based Metamemory Cognitive Training in Older Adults With Mild Cognitive Impairment: 15-Month Prospective Single-Arm Longitudinal Study
JMIR Aging 2026;9:e81648
URL: https://aging.jmir.org/2026/1/e81648
doi:10.2196/81648
PMID:

©Jung-In Lim, Yeeun Byeon, Sunyoung Kang, Hyeonjin Kim, Keun You Kim, Lukas Stenzel, So Yeon Jeon, Jun-Young Lee. Originally published in JMIR Aging (https://aging.jmir.org), 02.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Community Readiness Model Applied in Older Adults for the Promotion of Digital Cognitive Training: Mixed Methods Feasibility Study

Cíntia Monteiro Carvalho[1,2*], PhD; Bruno Costa Poltronieri[1,2,3*], PhD; Beatriz Pereira da Silva Lima[1], BSc; Rafaela Guilherme Ferreira[1,2,3], BSc; Maria Eduarda Alves Reis[1,2], BSc; Christine FitzGerald[4,5], PhD; Brian Lawlor[4], MD; Rogério Panizzutti[1,2,4], MD, PhD

[1]Instituto de Psiquiatria, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[2]Instituto de Ciências Biomédicas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[3]Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro, Rio de Janeiro, Brazil

[4]Global Brain Health Institute, Trinity College Dublin, Dublin, Ireland

[5]Ageing Research Centre, University of Limerick, Limerick, Ireland

[*]these authors contributed equally

Corresponding Author:
Rogério Panizzutti, MD, PhD
Instituto de Psiquiatria
Universidade Federal do Rio de Janeiro
Avenida Venceslau Brás 71
Rio de Janeiro, 22290140
Brazil
Phone: 55 21999090498
Email: rogerio.panizzutti@ipub.ufrj.br

Abstract

Background: Digital cognitive training, which involves structured, digital exercises designed to enhance cognitive functions, has shown potential benefits for older adults. While digital cognitive training has shown potential benefits for older adults, successfully incorporating it into their daily routines remains a challenge. Community readiness refers to the group's ability and capacity for a behavior change to be more effective and sustainable. In this study, readiness specifically refers to the community's preparedness to engage in and sustain digital cognitive training.

Objective: This study aimed to evaluate the feasibility of the Community Readiness Model (CRM) to identify facilitators and barriers in implementing training and applying supportive strategies to increase readiness.

Methods: This mixed methods feasibility study was conducted as part of a stepped-wedge randomized controlled trial examining the effects of digital cognitive training on functionality and cognition in older adults. Fifty-six participants were recruited from a nursing home and the community and were allocated to one of two training protocols: one group played digital leisure games for 10 hours followed by 10 hours of digital cognitive training, while the other group completed 20 hours of digital cognitive training. Readiness levels were assessed using CRM through semistructured individual interviews conducted before the intervention. Additionally, the CRM was administered to a sample of stakeholders to evaluate community-level readiness. Following the implementation of strategies aimed at enhancing readiness, as well as the delivery of the digital cognitive training, the interviews were readministered to evaluate changes in readiness levels.

Results: Before the training, participants demonstrated low levels of readiness, with a median CRM total score of 3.6 (IQR 2.2-4.6) in the 10-hour training group and 3.7 (IQR 2.7-4.3) in the 20-hour group. To address this, several strategies were implemented: dissemination of psychoeducational content via social media, distribution of cognitive exercise guides, stakeholder engagement to raise project awareness, and provision of necessary devices along with comprehensive support throughout the training. Following the implementation of these strategies and the training sessions, a significant increase in readiness was observed. Both groups achieved a median CRM total score of 6 (IQR 5-7). Participants identified key facilitators as interest in the training, noticeable cognitive improvements, team support, motivation to continue, the challenge presented by the training, and overall satisfaction. The primary barrier reported was difficulty using the technology.

XSL•FO
RenderX

Conclusions: The CRM demonstrates strong feasibility as a tool for identifying facilitators and barriers that can inform strategies to enhance readiness for digital cognitive training in older adults. The observed increase in readiness scores following the implementation of targeted strategies highlights the CRM's potential to guide the development of effective, supportive interventions. These findings emphasize the importance of addressing technological challenges while harnessing motivating factors to promote the successful adoption of digital cognitive training within this population.

(JMIR Aging 2026;9:e69434) doi:10.2196/69434

KEYWORDS

healthy aging; cognitive training; internet-based interventions; behavior change; community participation

Introduction

Aging is associated with a progressive decline in several brain functions including sensory perception, attention, memory, and motor control, which is termed "age-related cognitive decline." Age-related cognitive decline is associated with a poorer quality of life, less independence, and a higher cost for the health system [1]. Therefore, investing in preventive measures to mitigate cognitive decline is necessary. Based on robust research that has elucidated the fundamental principles of brain plasticity, it is now well established that dedicated behavioral training can bring about substantial improvement in cognitive function or recovery [2-6].

Digital cognitive training has emerged as a promising intervention to enhance cognition in older adults. By engaging in these training programs, individuals can either increase their neural capacity or improve the efficiency of existing brain resources [3]. This approach generally focuses on improving general cognitive processes such as processing speed, attention, and working memory, with the premise that these improvements will transfer to real-world cognitive gains [7]. Systematic reviews and meta-analyses of digital cognitive training for older adults have concluded that these training interventions can be effective for older adults and thus should be recommended [8-11].

While the benefits associated with cognitive training for older adults are well recognized, a challenge lies in effectively implementing digital cognitive training in their lives. Previous studies that provided cognitive training to older adults were conducted under the constraints of research protocols and did not consider the resources available to the participants nor the effective behavioral and societal changes needed to ensure engagement with users. Our previous experience with clinical trials and other studies using digital cognitive training has revealed that several training-related factors affect engagement, including personal support, location, attitudes toward new learning and technology, and communication between the users [12-16].

The Community Readiness Model (CRM) is a theory-based strategic model used to assess and build community capacity to tackle social issues. The CRM allows the community's influence and knowledge to assume a central role in the application of the model, allowing researchers to meet communities where they are and on their own terms [17]. The concept of CRM revolves around the idea that starting a new program or intervention is only appropriate when the community is ready. For community interventions or programs to be successful, a community must be consistent with its awareness of the problem and its readiness for change [16,18]. The concept of community adopted in this study focuses on a community of interest defined by shared interests regardless of individuals' location or social group [19,20]. By tailoring interventions to a community's specific needs and readiness, CRM not only identifies facilitators and barriers but also promotes skill development among community members. This collaborative approach is essential for fostering sustainable change and empowering individuals to tackle the challenges they face. Engaging the community in this way ensures that solutions are relevant and supported, ultimately leading to more successful outcomes [21].

Previous applications of the CRM were effective in addressing a variety of health issues, such as health service delivery [22,23], substance use [24], childhood obesity [25], and mental health awareness [21]. On the other hand, for the design and implementation of community-based interventions, the involvement of stakeholders (eg, health promotion professionals) is important, since they can help ensure that the intervention planning is suitable given the real conditions [26,27].

Given the complexity of promoting affective engagement in digital cognitive training—an essential factor for effectively integrating such interventions into the daily routines of older adults [28]—we chose to use the CRM as a framework to support this behavior change. In the context of our study, readiness refers to the community's preparedness to adopt and support training initiatives. The CRM offers valuable insights into both behavioral factors (eg, familiarity with technological tools) and personal factors (eg, an individual's interest in participating in a specific intervention). Unlike other models, the CRM has the distinct advantage of incorporating a community-based perspective, providing a comprehensive assessment of a community's readiness. These features are critical for tailoring strategies that align with the community's current level of readiness, thereby enhancing the likelihood of implementing an effective and sustainable intervention.

We hypothesized that the CRM can be a useful and viable tool to identify barriers and facilitators in the training, which can then be used to develop strategies to increase community readiness for digital cognitive training. To examine this hypothesis, the primary objective of this study was to provide evidence for the feasibility of using the CRM to evaluate the readiness of older adults undergoing digital cognitive training and identify facilitators and barriers associated with the training. The secondary objective was to implement strategies aimed at

XSL•FO
RenderX

enhancing older adults' readiness for training and to evaluate changes in their readiness levels after completing a digital cognitive training program. To achieve the proposed objectives, we applied the CRM to a sample of older adults before and after the implementation of the training. The model was also applied to a sample of stakeholders. We used the results of the preintervention quantitative and qualitative analysis to develop and implement strategies to increase the readiness level of older adults.

Methods

Design

This mixed methods feasibility study is part of an ongoing double-blind, stepped-wedge randomized controlled trial investigating the effects of different doses (10 or 20 hours) of digital cognitive training on cognitive function and functionality in older adults. We hypothesized that participants who received 20 hours of digital cognitive training would exhibit sustained improvements in both cognitive function and functionality. All participants underwent a comprehensive initial assessment, including clinical history, as well as assessments of cognition, functionality, and symptoms of anxiety and depression. They were randomized using a hierarchical stratification by gender, age, years of education, Montreal Cognitive Assessment test score, and the Technology Activities of Daily Living Questionnaire score. They were allocated to one of two training protocols: one group played digital games for leisure for 10 hours followed by 10 hours of digital cognitive training, while the other group engaged in 20 hours of digital cognitive training. The stepped-wedge randomized controlled trial involves random and sequential crossover of groups from control to intervention until all groups are exposed to the intervention [29]. This allowed all participants to perform digital cognitive training during the study. Participants and researchers were blinded throughout the study.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and was approved by the University Hospital Clementino Fraga Filho Ethics Committee (CAAE: 82713418.6.0000.5257). All participants (older adults and stakeholders) provided written informed consent prior to participation. Participant privacy was protected through data de-identification and secure data storage, with access restricted to the research team. No financial compensation or incentives were provided.

Participants and Training Conditions

Participants were recruited by advertising on social media and in a nursing home in Rio de Janeiro, or through referrals from health professionals, between April 2019 and November 2021. We invited older adults aged 60 years or above and included those who (1) provided written informed consent, (2) had normal or corrected vision and hearing, (3) had access to a computer and internet at home, and (4) were native Portuguese speakers and literate. Participants were excluded if they (1) had a diagnosis of dementia, (2) were dependent on activities of daily living, or (3) had major medical conditions that would prevent them from participating in the study.

The training platform used was BrainHQ from Posit Science, Inc, and the exercises were previously selected and aimed to improve executive functions. The exercises for cognitive training were as follows: *divided attention*, *eye for detail*, *target tracker*, *juggle factor*, and *mind-bender*. These exercises dynamically adjusted their speed and stimulus type based on participant performance. As participants improved, task difficulty systematically increased, and an algorithm maintained individual success rates around 85%. This threshold was incorporated into the design of the training program as a means of achieving an optimal balance between cognitive challenge and sustained learner engagement [4]. To enhance motivation, correct responses were rewarded with engaging visual and auditory feedback, alongside the accumulation of stars. After completing the intervention, participants received feedback provided by the BrainHQ platform on their progress in trained cognitive domains.

The leisure games condition served as a control for potential confounding factors such as computer exposure, researcher interaction, and nonspecific cognitive engagement—including attention, executive functions, and motivation—stimulated by graphics-based digital games. The leisure games were commercial online: *Bubble Poke*, *Smarty Bubbles*, *Dominoes Classic*, *Puzzle*, *Find 500 Differences*, *Find the Birds*, *Let's Clean Up*, and *Crosswords*.

In both groups, participants engaged in 15-minute sessions of 4 predetermined exercises, following the established protocol. They were encouraged to train at least twice a week. The intervention of both groups was accompanied by a video call by a research group member who explained each exercise or game in advance and supported the user by answering questions when necessary.

Stakeholders' Participation

Based on previous studies indicating that stakeholder engagement is associated with improved community readiness for facilitating change [30,31], we determined that stakeholder participation was essential for understanding community readiness and identifying potential barriers and facilitators to intervention implementation. Prior to launching the digital cognitive training program, we conducted interviews with 7 stakeholders, including 4 health professionals (a physical therapist, an occupational therapist, a physician, and a speech therapist) and 3 family members of study participants. All stakeholders had direct contact with the participants. Recruitment took place at a nursing home in Rio de Janeiro and through referrals from professionals in the field. Interviews were conducted prior to the intervention, and the CRM was applied at this stage to gather additional insights into participants' initial level of readiness.

Community Readiness Model

The CRM was applied during the initial assessment and after participants completed either 10 or 20 hours of training, through semistructured individual interviews conducted via video call. The semistructured interview had 11 questions that allowed

XSL•FO
RenderX

participants to talk about access, perceptions, expectations, and their experiences regarding the use of technological resources and digital cognitive training. The questions were developed following the criteria established in the manual proposed by Stanley [32], which assesses the community's readiness by addressing 5 key dimensions that are crucial to initiate and sustain healthy changes. These dimensions are *knowledge of the issue*, which refers to how much the community knows about the issue; *knowledge of efforts*, which refers to how much the community knows about the programs and activities of digital cognitive training; *community climate*, which reports the community's attitude toward addressing the issue; *leadership*, which indicates the leadership's attitude toward addressing the issue; and *resources*, which states how much the resources are being used or could be used to address the issue. The "leadership" dimension was not included in the interview, as the description of this dimension did not apply to the community context.

Each dimension is scored on a scale of 1 to 9, corresponding to a stage. In the initial stage of readiness, "no awareness" (1), the community lacks information about local initiatives, resulting in little awareness of the problem. As the community progresses to "denial/resistance" (2), there is resistance to acknowledging the issue due to a lack of accurate information. Community perception improves slightly in the "vague awareness" stage (3), but knowledge remains superficial. "Preplanning" (4) involves recognizing the concern more concretely, though knowledge and resources are limited. During "preparation" (5), awareness increases, and leadership supports efforts with directed resources. In "initiation" (6), most members have a basic understanding, and leadership plays a central role in planning additional initiatives. Community knowledge solidifies in the "stabilization" stage (7), and leadership ensures long-term viability with continuous resources. "Confirmation/expansion" (8) sees high community knowledge and leadership expanding initiatives, indicating lasting commitment. In the final stage, "high level of community ownership" (9), the community shows substantial knowledge, with active involvement and support from leadership, adjusting resources based on continuous evaluations. Each dimension is scored on a rating scale from 1 to 9, corresponding to these stages.

The CRM interviews were audio-recorded, transcribed, and then analyzed and scored by two independent researchers. The CRM scoring process followed the CRM handbook using anchored rating scales. After establishing a score by each researcher, a consensus was reached on the final score.

Development of Behavior Change Strategies

We implemented strategies to enhance readiness levels, drawing on each dimension of the CRM and incorporating behavior change techniques (BCTs). BCTs are commonly used to promote adherence to specific interventions [33]. The BCT taxonomy, a widely recognized framework, organizes these techniques into 16 main categories encompassing 93 distinct methods. For instance, it includes techniques such as "graded tasks," which involve starting with simple, easy-to-perform activities and progressively increasing their difficulty in a manageable way until the desired behavior is successfully achieved [34]. Strategies were informed by pretraining interviews and community feedback, with the aim of ensuring relevance and effectiveness. Implementation occurred concurrently across both groups, with no distinctions made between them.

Clinical Data and Assessments

We collected clinical and demographic variables from the participants. In addition, we assessed cognition through the application of the Montreal Cognitive Assessment [35] test; to measure functionality, we administered the Technology Activities of Daily Living Questionnaire [36]; to evaluate anxiety symptoms, we used the Geriatric Anxiety Inventory with 20 agree/disagree items [37]; and to measure the symptoms of depression, we used the 15-item version of the Geriatric Depression Scale [38].

Feasibility Outcomes

The primary results regarding the acceptability of the intervention were adapted from the study by Proctor et al [39].

- Recruitment rate: the sample size of recruited participants followed the manual's recommendation of conducting at least 6 interviews, with the possibility of more to obtain a comprehensive understanding of the community [32].
- Attendance: overall attendance rate of >60% [40].
- Attrition: retention rate of at least 75% of participants to the follow-up [40].
- Acceptability of intervention: it was assessed through the responses obtained from semistructured interviews using the CRM.

Data Analysis

The Shapiro-Wilk test was applied to determine data normality. For nonnormally distributed data, nonparametric statistics were used in the analyses. The independent sample *t* test was used to compare the two groups at baseline, the chi-square test was used to assess previous contact with a computer and gender, and the Wilcoxon test was used to verify differences of dimensions between pre- and postintervention. To identify the facilitators and barriers in the baseline and postintervention, we executed the descending hierarchical classification. This method analyzes the lexical context and generates a hierarchical scheme of word classes that are grouped into similarity classes by the chi-square test. High chi-square values characterize strong associations between word and class; we considered words with a $\chi^2_1>3.8$, corresponding to $P<.05$, for the class groupings. This analysis uses words in their reduced forms, and a minimum use of 75% of the total text segments (TS) of the original corpus is required [41]. The rate of textual segments below 75% generated by the software suggests a less homogeneous textual corpus, more dispersed concerning the analyzed content, and less representative [42]. From the TS assigned to each of the word classes revealed by the software, the data were analyzed and interpreted by the researchers to identify barriers and facilitators. The general corpus of textual analysis performed at baseline comprised 56 texts, separated into 652 TS, using 497 TS (76.2%). The postintervention analysis comprised 21 texts for the 10-hour training group, separated into 221 TS, using 172 TS (77.8%), and 20 texts for the 20-hour training group, separated into 245 TS, using 185 TS (75.5%). The data that

XSL•FO
RenderX

support the findings of this study are available on request from the corresponding author. We performed all quantitative statistical analyses using SPSS software (v. 26.0; IBM Inc) and the qualitative analysis using the IRAMUTEQ software (Université de Toulouse).

Results

Feasibility Outcomes

Recruitment Rate and Participant Characteristics

A total of 56 participants were recruited, adhering to the criteria outlined in the recruitment rate guidelines. The cohort was predominantly female, with 43 out of 56 participants (77%), and had an average age of 76.98 (SD 7.7) years and 12.77 (SD 5.03) years of education. Table 1 shows the clinical and demographic characteristics of the total sample and by group, and compares the groups concerning age, gender, education, weekly physical activity practice, number of falls in the last year, previous contact with a computer, cognition, and mood. There were no significant differences between the groups that performed 10 or 20 hours of training regarding any demographic or clinical variables.

Table 1. Participant characteristics.

	Total (n=56)	Group 10 hours training (n=31)	Group 20 hours training (n=25)	*P* value
Age (years), mean (SD)	76.98 (7.70)	77.52 (7.91)	76.32 (7.54)	.56
Female, n (%)	43 (77)	24 (77)	19 (76)	.90
Education (years), mean (SD)	12.77 (5.03)	12.55 (4.97)	13.04 (5.18)	.72
Physical activity (times per week), mean (SD)	1.77 (1.89)	2.00 (1.98)	1.48 (1.78)	.31
Number of falls (past year), mean (SD)	0.29 (0.78)	0.27 (0.94)	0.32 (0.55)	.80
Previous contact with computer, n (%)				.93
None	13 (23)	7 (23)	6 (24)	
Low	26 (46)	14 (45)	12 (48)	
Medium	6 (11)	3 (10)	3 (12)	
High	11 (20)	7 (23)	4 (16)	
MoCA[a] score, mean (SD)	20.88 (2.97)	21.10 (3.03)	20.60 (2.94)	.54
GDS[b] score, mean (SD)	3.79 (2.54)	3.48 (2.56)	4.16 (2.52)	.32
GAI[c] score, mean (SD)	6.46 (4.88)	5.78 (4.52)	7.44 (5.35)	.30

[a]MoCA: Montreal Cognitive Assessment.

[b]GDS: Geriatric Depression Scale.

[c]GAI: Geriatric Anxiety Inventory.

Attrition Rate

In the group that underwent 10 hours of training, 7 participants dropped out of the study during the first 10 hours of games: 4 due to lack of interest and 3 due to health problems, resulting in an attrition rate of 22%. However, none of the participants dropped out during the 10 hours of digital cognitive training. Additionally, we missed data for another 3 participants in this group. While in the group that completed 20 hours of training, 3 participants dropped out of the study. During the first 10 hours of training, 1 participant left due to lack of interest, resulting in an attrition rate of 4%. During the next 10 hours of cognitive training, 1 participant left due to lack of interest and 1 due to health problems, resulting in an attrition rate of 8%. Furthermore, we had missing data for 2 more participants in this group. Missing data resulted from participants' unavailability for the postintervention interview despite having completed the training.

Attendance

Attendance was 100% among participants in both groups.

Community Readiness Model Score

First, we analyzed the readiness level for both groups before starting the training. Before the training, the participants had a low level of readiness both for the CRM total and for the dimensions separately, being in the "denial/resistance," "vague awareness," or "initiation" stages (Table 2).

XSL•FO
RenderX

Table 2. Readiness level at baseline.

	Group 10 hours training (n=21)		Group 20 hours training (n=20)	
	Median (IQR)	Stage of readiness	Median (IQR)	Stage of readiness
CRM[a] total	3.6 (2.2-4.6)	Vague awareness	3.7 (2.7-4.3)	Vague awareness
Knowledge of issue	2.0 (3.0-4.0)	Denial/resistance	2.0 (3.0-4.0)	Denial/resistance
Knowledge of efforts	2.0 (1.0-3.0)	Denial/resistance	2.0 (2.0-3.0)	Denial/resistance
Community climate	3.5 (3.0-4.0)	Vague awareness	3.5 (3.0-4.0)	Vague awareness
Resources	6.0 (3.0-7.0)	Initiation	6.0 (4.0-7.0)	Initiation

[a]CRM: Community Readiness Model.

Readiness of Stakeholders

To gain a broader understanding of the community's readiness, we also assessed the readiness of stakeholders and found it to be consistently low in all areas. The stage and readiness level for each dimension were as follows: knowledge of efforts, "initiation" (5.71); knowledge of the subject, "initiation" (5.57); community climate, "preplanning" (4.14); leadership, "vague conscience" (3.28); and resources, "vague conscience" (3.42).

Facilitators and Barriers Related to Digital Cognitive Training at Baseline

To get closer to the real needs of older adults in the face of digital cognitive training, we went beyond the quantitative analysis and performed qualitative analysis of the interviews to identify the facilitators and barriers at baseline. The content generated by the IRAMUTEQ software was categorized into specific classes and classified as either a facilitator or a barrier. Examples of quotes for each class are described in a later section.

At baseline, the facilitators were the following. (1) "Curiosity" (TS=100/599, 16.7%): refers to the curiosity to know and understand how cognitive training works and the associated benefits. (2) "Search for cognitive stimulation" (TS=103/599, 17.2%): refers to the participants' search for resources to improve cognitive abilities. (3) "Getting results" (TS=130/599, 21.7%): refers to participants' interest in training gains because they perceive cognitive difficulties. (4) "Interest in the training" (TS=193/599, 32.2%): pertains to participants' interest in training to improve brain health. The barrier was "difficulty in handling technology" (TS=73/599, 12.2%), which refers to the difficulties older adults have in using technologies that are not part of their daily lives.

Strategies Adopted

Based on these findings in the baseline with participants and stakeholders, we adopted the following targeted strategies with BCTs to increase the level of readiness in each dimension.

1. Knowledge of issue: we shared psychoeducational information through social media about the potential benefits associated with the training (BCT: information about health consequences and problem-solving).
2. Knowledge of efforts: we provided a guide explaining how each cognitive exercise works (BCT: graded tasks, social support).
3. Community climate: we contacted stakeholders and disseminated information about this project (BCT: information about health consequences).
4. Resources: we provided the hardware (computer or tablet) according to the necessity of the participant and provided full support during training and a practical guide on the use of technology (BCT: material incentive and social support [practical]).

Postintervention and Implementation of the Strategies

Subsequently, we evaluated the readiness levels of both groups following the completion of the training. The results indicated that postintervention participants advanced to the preparation, initiation, and stabilization stages (Table 3).

Table 3. Level of readiness postintervention and the application of strategies.

	Group 10 hours training (n=21)		Group 20 hours training (n=20)	
	Median (IQR)	Stage of readiness	Median (IQR)	Stage of readiness
CRM[a] total	6.0 (5.0-7.0)	Initiation	6.0 (5.0-7.0)	Initiation
Knowledge of issue	5.0 (4.0-6.5)	Preparation	5.0 (4.5-5.5)	Preparation
Knowledge of efforts	5.0 (3.5-6.0)	Preparation	5.0 (4.0-6.0)	Preparation
Community climate	5.0 (4.5-6.5)	Preparation	5.75 (4.5-7.0)	Preparation
Resources	7.0 (7.0-8.0)	Stabilization	7.25 (7.0-8.0)	Stabilization

[a]CRM: Community Readiness Model.

Subsequently, we evaluated whether the readiness levels changed postintervention. The analysis revealed statistically

XSL•FO
RenderX

significant increases in all domains of readiness postintervention in both groups, with the exception of resources in the 10-hour training group (Table 4).

Table 4. Changes in readiness levels postintervention.

	Baseline, median (IQR)	Postintervention, median (IQR)	Wilcoxon	
			Z	*P* value
Group 10 hours training (n=21)				
CRM[a] total	3.6 (2.2-4.6)	6.0 (5.0-7.0)	–4.01	<.001
Knowledge of issue	2.0 (3.0-4.0)	5.0 (4.5-5.5)	–3.63	<.001
Knowledge of efforts	2.0 (1.0-3.0)	5.0 (4.0-6.0)	–3.78	<.001
Community climate	3.5 (3.0-4.0)	5.0 (4.5-6.5)	–4.02	<.001
Resources	6.0 (3.0-7.0)	7.0 (7.0-8.0)	–2.78	.51
Group 20 hours training (n=20)				
CRM total	3.7 (2.7-4.3)	6.0 (5.0-7.0)	–3.92	<.001
Knowledge of issue	2.0 (3.0-4.0)	5.0 (4.5-5.5)	–3.70	<.001
Knowledge of efforts	2.0 (2.0-3.0)	5.0 (4.0-6.0)	–3.72	<.001
Community climate	3.5 (3.0-4.0)	5.75 (4.5-7.0)	–3.92	<.001
Resources	6.0 (4.0-7.0)	7.25 (7.0-8.0)	–3.65	<.001

[a]CRM: Community Readiness Model.

Facilitators and Barriers Related to Digital Cognitive Training Postintervention

The facilitators post intervention for the 10-hour training group were as follows.

1. Desire to keep training (TS=34/188, 18.1%): refers to participants' desire to continue training or to start other cognitive stimulation activities after the completion of the training.
2. Team support (TS=46/188, 24.5%): refers to the importance of support and encouragement offered by the research team.
3. Cognitive improvement (TS=58/188, 30.8%): refers to participants' perception of cognitive improvement after training and sharing information with the community about the benefits of training. The main barrier was "difficulty in handling technology" (50/188, 26.6%). Although the older adults used a computer during training, there were complaints regarding difficulty with access and familiarity in handling technology.

The facilitators postintervention for the 20-hour training group were as follows.

1. Cognitive improvement (TS=26/176, 14.8%): addresses the participants' perception of cognitive improvement, better performance during the training, and sharing information to the community about the benefits of training.
2. Satisfaction (TS=62/176, 35.2%): refers to the positive experience reported by participants upon completing a new activity.
3. Training challenge (TS=26/176, 14.8%): refers to the participants' motivation to overcome the challenges of training exercises. The main barrier was "difficulty in handling technology" (TS=62/176, 35.2%): this class reports the difficulty in handling technology and the participants' need for assistance from a family member during the training.

The qualitative analyses also revealed that participants continued to report difficulties in dealing with technology, even after completing the training, in both the 10-hour and 20-hour training groups. Before the training, older adults reported that they were searching for cognitive stimulation, and postintervention, both groups reported perceived cognitive improvement.

Table 5 complements these findings by providing examples of quotes for each class of facilitators and barriers, offering further insights into the personal experiences of the participants. Additionally, examples of quotes for each dimension from both groups are available in the supplementary material.

XSL•FO
RenderX

Table 5. Examples of quotes for each class.

	Quotes
Baseline	
Facilitators	
Interest in the training (193/599, 32.2%)	*…To improve my memory. My attention, I think my attention is very bad, I can't read books.*
Curiosity (100/599, 16.7%)	*…I want to do the training because I found it interesting. I think that curiosity, we are curious to know what it is, what the result is.*
Getting results (130/599, 21.7%)	*…Because I want to work my brain to see if I don't get dementia like my mother did.*
Search for cognitive stimulation (103/599, 17.2%)	*…that's how it is, memorizing, it's the medication schedule thing, these things, I look for it, I even avoid writing it down so I don't get complacent.*
Barrier	
Difficulty in handling technology (73/599, 12.2%)	*…I have a computer, but I don't know how to deal with it very well.*
Postintervention 10 hours	
Facilitators	
Cognitive improvement (58/188, 30.8%)	*…I even noticed that there was some improvement in memory. It's... to remember things. I thought it improved a lot.*
Team support (46/188, 24.5%)	*…I want to thank the girls for their patience with us.*
Desire to keep training (34/188, 18.1%)	*…I started to think it was good, I thought it was worth it and that I want to continue if I get the chance.*
Postintervention 20 hours	
Facilitators	
Cognitive improvement (26/176, 14.8%)	*…I think it improved my mental agility because I came out of the doldrums.*
Satisfaction (62/176, 35.2%)	*…I think it was wonderful, it is fantastic to put the brains to work. I'm even recommending my mother and sister to participate.*
Training challenge (26/176, 14.8%)	*…I became so obsessed with this task that I kept doing exercises to do better, the reverse. I think I got on with that, doing my silly things, like putting some cream on the toothbrush, that was common thing, a while ago.*
Barrier	
Difficulty in handling technology (62/176, 35.2%)	*…I didn't understand a lot, I think I was doing it all wrong.*

Discussion

Our study explored the feasibility of using the CRM to identify barriers and facilitators in implementing digital cognitive training and applying strategies to increase readiness. Older adults were assessed for their initial readiness, which was low but showed significant improvement following targeted strategies and training. Key facilitators included curiosity, motivation to improve cognition, and team support. The main barrier was difficulty with technology, which remained even postintervention. Participants reported cognitive benefits and showed interest in continuing the training. We found that it was feasible to use CRM to identify facilitators and barriers in the implementation of digital cognitive training for older adults, aiming to guide support strategies to enhance readiness levels.

Although our study had a different objective, the observed increase in community readiness for digital cognitive training aligns with findings from institutional delivery services in Ethiopia. While our study addressed individual barriers and motivators, the Ethiopian study used a structured, village-based intervention with education, training, and feedback over 15 months. This highlights CRM's adaptability to different contexts and strategies [43]. Supporting this, a review study indicated that the CRM was effective in identifying stakeholders, informing them about the intervention, and enhancing community awareness of the issue [44]. Findings suggest that extended exposure to the intervention is associated with higher levels of community readiness, whereas previous research [44,45] has documented annual increases of 0.5 to 1 stage in readiness resulting from sustained intervention efforts.

We successfully recruited a target audience that exceeded the minimum required number of participants. The attrition rate was 7 out of 31 participants (22%) in the group that completed 10 hours of games, with no participants dropping out during the digital cognitive training. In the group that completed 20 hours of training, 1 out of 25 participants (4%) left during the first 10 hours. During the next 10 hours, the attrition rate was 2 out of 24 participants (8%). Attendance was high, with 100% of participants attending all 10 or 20 sessions. Attrition patterns differed between groups, with dropouts in the 10-hour group

XSL•FO
RenderX

occurring during the initial game-based activities, which may have been less stimulating than digital cognitive training. In contrast, the 20-hour group had fewer dropouts early on and remained more consistent throughout, possibly due to higher motivation or perceived benefits. External factors, such as health issues and missing data, also contributed to retention differences.

Although CRM has been used in various studies across different fields to promote behavioral changes for better health, such as preventing traumatic brain injury, reducing alcohol use, and supporting older adults, few studies have evaluated readiness before and after the intervention, as we did [23,25,26]. We found that participants had a low level of readiness before the training. However, after implementing the targeted strategies, we observed an increase in readiness, both in the overall CRM score and across its individual dimensions.

The improvement in readiness may be attributed to older adults' self-perception of benefiting from potential gains associated with training, leading to a greater awareness of the importance of brain training. A systematic review of BCT in computerized cognitive training for older adults described that feedback given by a person about behavioral outcomes can influence adherence and effectiveness of training [11]. In this study, participants received feedback from monitors who accompanied them throughout the sessions on their progress in trained cognitive domains as part of a behavior change strategy, which possibly increased their awareness of the potential benefits obtained from training. The choice to have more personalized feedback from the monitor was based on a systematic review [11] that indicated that when this feedback was provided by a computer, it tended to be less effective.

In terms of the "knowledge of the issue" dimension, participants progressed from the "vague awareness" stage at baseline to the "preparation" stage postintervention. Upon completion, they showed an improved understanding of cognitive exercise and the potential for recovering lost abilities. A qualitative study with older adults with mild cognitive impairment revealed that the computer program used in the study not only enhanced their memory and attention but also assisted in managing daily life challenges from the point of view of the participants [46]. Another study indicated that most older adult participants believed the program was beneficial for their health [47]. Motivators related to cognitive effectiveness, such as the perceived improvement in cognition and tangible evidence of that improvement, were also significant [48]. These findings suggest that as participants recognize the value of digital cognitive training and its impact on their lives, they are more motivated to engage in it.

Participants in this study also demonstrated a significant improvement in the "knowledge of the efforts" dimension. This change was observed as their initial state of "denial" or "resistance" transformed into a state of "preparation." Over time, participants expressed satisfaction with the potential benefits of cognitive training and recommended it to others, including family and community members. This is further supported by their quotes (Table 4) reflecting satisfaction following the completion of the training.

A study that applied the CRM in the context of childhood obesity prevention also reported significant behavior change by adopting strategies aimed at increasing community readiness. As participants became more aware of the issue and the benefits of healthier behaviors, they actively shared their knowledge with others in the community [49]. In line with these findings, some participants in our study expressed interest in continuing the cognitive training after the program ended, a trend also observed in previous research on cognitive training interventions [47-50]. This suggests that as participants recognize the value of training, they are more likely to seek continued engagement, further reinforcing the importance of fostering sustained motivation and readiness.

The "community climate" dimension, which refers to people's attitudes toward change—in this case, engaging in digital cognitive training—progressed from the "vague awareness" phase to the "preparation" phase. This improvement can be attributed to the BCT implemented, which kept participants informed about the objectives of cognitive training throughout the process. A Canadian study similarly found that participants who underwent digital cognitive training reported cognitive improvements that were meaningful to their daily lives [51]. Moreover, the absence of adverse effects and the reported cognitive gains likely contributed to an increased sense of self-efficacy among participants, as observed in Edwards et al [52] study on the outcomes of digital cognitive training programs.

Before the training, the "resources" dimension was already at the "initiation" stage, as access to technological devices was a prerequisite for participation. Despite a significant improvement in this dimension, participants frequently reported difficulties in handling technology both at baseline and at postintervention. This barrier, also highlighted by stakeholders during baseline assessments, reflects the digital exclusion often faced by older adults, who are less likely to use computers and the internet [53-55]. Similarly, other studies have noted limited access to devices and a lack of proficiency in using digital tools among older adults, particularly computers [54]. To address this issue, we recommend that future studies incorporate digital inclusion initiatives prior to formal training sessions. Such initiatives could help familiarize older adults with technology, build their confidence, and improve their proficiency. Evidence from recent research, such as Djabelkhir et al [56], supports this approach, showing that pretraining technology familiarization enhances participants' skills, fosters positive attitudes toward training, and boosts self-efficacy, ultimately leading to more successful digital literacy outcomes.

In our study, stakeholders showed a low level of readiness, contrasting with research that highlights strong professional support for digital cognitive programs and their role in preserving cognitive function among older adults [56]. Stakeholders in our study faced challenges in helping participants consistently integrate training into their routines, primarily due to limited familiarity with technological resources. Addressing this issue calls for collaborative efforts to enhance stakeholders' understanding of digital cognitive training and to develop strategies that effectively promote older adults' engagement with these tools.

XSL•FO
RenderX

The main facilitators identified in the baseline were related to the interest in the training, getting results, searching for cognitive stimulation, and curiosity. The facilitators identified in the postintervention analysis for the group that did 10 hours of training were related to cognitive improvement, team support, and the desire to continue training. While, for the group that completed 20 hours, the facilitators were related to cognitive improvement, satisfaction, and training challenge. Notably, the "training challenge" reported by the 20-hour group highlights the importance of programs that adapt to an individual's progress throughout the intervention [10].

In this study, participants were not compensated for their involvement, reflecting real-world conditions where individuals typically participate without financial incentives. Their voluntary engagement created a naturalistic setting, enhancing the generalizability of the findings to everyday contexts. This study has some limitations. Although we tried to diversify the sample for data collection, the criteria for participation in this study included the participant having access to a computer and the internet. As a consequence, we have a selected sample, which does not allow the generalization of these findings to populations with different characteristics. Another limitation of this study concerns the fact that we did not interview stakeholders after the participants completed the digital cognitive training, which prevented us from determining whether there were similar changes in readiness. Finally, this feasibility study was designed to establish the viability, and not to demonstrate the effectiveness of the use of CRM to increase readiness in older adults [53]. Thus, the current findings of an increase in readiness should be interpreted as exploratory and need to be confirmed.

We propose several recommendations for future research aimed at increasing readiness for digital cognitive training among older adults: (1) disseminate psychoeducational content on digital cognitive training and its associated benefits; (2) incorporate digital literacy training before cognitive training, especially for older adults who have difficulty handling technology; (3) provide feedback on training performance and the cognitive domains being trained; and (4) disseminate study findings not only at scientific events but also through mainstream media.

In conclusion, CRM is a feasible tool for identifying facilitators and barriers that can inform strategies to enhance readiness for digital cognitive training in older adults. Future studies should assess the effectiveness of applying CRM to improve readiness and examine whether higher levels of readiness are associated with greater engagement and cognitive improvements following digital cognitive training in older adults. Notably, while CRM was selected in this study due to its comprehensive framework encompassing multiple dimensions, key stakeholders, and the population of older adults, alternative models addressing behavior change may also contribute to the successful implementation of digital cognitive training programs. Therefore, further research exploring diverse theoretical frameworks is encouraged to optimize cognitive training interventions for older adults.

Acknowledgments

We are grateful to the subjects who participated in this study. The cognitive training software used in this study and all technical support were provided to us free of charge by Posit Science, Inc.

Funding

This work was supported by grants from the Global Brain Health Institute, Alzheimer's Association, Atlantic Institutes, Fundac¸a˜o de Amparo a` Pesquisa do Estado do Rio de Janeiro (FAPERJ), and Conselho Nacional de Desenvolvimento Cienti´fico e Tecnolo´gico (CNPq). RP is an Atlantic Fellow of the Global Brain Health Institute. CMC was supported by fellowships from CNPq.

Conflicts of Interest

RP is the founder of NeuroForma LTDA, a company with a financial interest in cognitive training. The remaining authors have no conflict of interest to disclose.

References

1. Khaw KT. Epidemiological aspects of ageing. Philos Trans R Soc Lond B Biol Sci 1997;352(1363):1829-1835 [FREE Full text] [doi: 10.1098/rstb.1997.0168] [Medline: 9460067]
2. Merzenich MM, Jenkins WM, Johnston P, Schreiner C, Miller SL, Tallal P. Temporal processing deficits of language-learning impaired children ameliorated by training. Science 1996;271(5245):77-81. [doi: 10.1126/science.271.5245.77] [Medline: 8539603]
3. Mahncke H, Bronstone A, Merzenich M. Brain plasticity and functional losses in the aged: scientific bases for a novel intervention. Prog Brain Res 2006;157:81-109. [doi: 10.1016/S0079-6123(06)57006-2] [Medline: 17046669]
4. Mahncke HW, Connor BB, Appelman J, Ahsanuddin ON, Hardy JL, Wood RA, et al. Memory enhancement in healthy older adults using a brain plasticity-based training program: a randomized, controlled study. Proc Natl Acad Sci U S A 2006;103(33):12523-12528 [FREE Full text] [doi: 10.1073/pnas.0605194103] [Medline: 16888038]
5. Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature 2013;501(7465):97-101 [FREE Full text] [doi: 10.1038/nature12486] [Medline: 24005416]

XSL•FO
RenderX

6. Mishra J, de Villers-Sidani E, Merzenich M, Gazzaley A. Adaptive training diminishes distractibility in aging across species. Neuron 2014;84(5):1091-1103 [FREE Full text] [doi: 10.1016/j.neuron.2014.10.034] [Medline: 25467987]
7. Strobach T, Karbach J. Cognitive training. In: An Overview of Features and Applications. Cham, Switzerland: Springer International Publishing; 2021.
8. Hill NTM, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A. Computerized cognitive training in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis. Am J Psychiatry 2017;174(4):329-340. [doi: 10.1176/appi.ajp.2016.16030360] [Medline: 27838936]
9. Kelly ME, Loughrey D, Lawlor BA, Robertson IH, Walsh C, Brennan S. The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: a systematic review and meta-analysis. Ageing Res Rev 2014;15:28-43. [doi: 10.1016/j.arr.2014.02.004] [Medline: 24607830]
10. Lampit A, Hallock H, Valenzuela M. Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. PLoS Med 2014;11(11):e1001756 [FREE Full text] [doi: 10.1371/journal.pmed.1001756] [Medline: 25405755]
11. Peeters G, Black IL, Gomersall SR, Fritschi J, Sweeney A, Guedes de Oliveira Y, et al. Behaviour change techniques in computerized cognitive training for cognitively healthy older adults: a systematic review. Neuropsychol Rev 2023;33(1):238-254 [FREE Full text] [doi: 10.1007/s11065-022-09537-4] [Medline: 35157209]
12. Ware C, Damnee S, Djabelkhir L, Cristancho V, Wu Y, Benovici J, et al. Maintaining cognitive functioning in healthy seniors with a technology-based foreign language program: a pilot feasibility study. Front Aging Neurosci 2017;9:42 [FREE Full text] [doi: 10.3389/fnagi.2017.00042] [Medline: 28298892]
13. Wu YH, Damnée S, Kerhervé H, Ware C, Rigaud AS. Bridging the digital divide in older adults: a study from an initiative to inform older adults about new technologies. Clin Interv Aging 2015;10:193-200 [FREE Full text] [doi: 10.2147/CIA.S72399] [Medline: 25624752]
14. Haesner M, O'Sullivan JL, Gövercin M, Steinhagen-Thiessen E. Requirements of older adults for a daily use of an internet-based cognitive training platform. Inform Health Soc Care 2015;40(2):139-153. [doi: 10.3109/17538157.2013.879149] [Medline: 24725153]
15. Mororó LG, Guimarães AL, Costa AC, Genaro L, Cavalcanti MT, Scoriels L, et al. Association between motivation and engagement with changes in cognition and symptoms after digital cognitive training in schizophrenia. Schizophr Res 2023;251:1-9. [doi: 10.1016/j.schres.2022.12.002] [Medline: 36527953]
16. Money A, Hall A, Harris D, Eost-Telling C, McDermott J, Todd C. Barriers to and facilitators of older people's engagement with web-based services: qualitative study of adults aged >75 years. JMIR Aging 2024;7:e46522 [FREE Full text] [doi: 10.2196/46522] [Medline: 38416543]
17. Kelly KJ, Edwards RW, Comello MLG, Plested BA, Thurman PJ, Slater MD. The community readiness model: a complementary approach to social marketing. Mark Theory 2003;3(4):411-426. [doi: 10.1177/1470593103042006]
18. Plested BA, Smitham DM, Jumper-Thurman P, Oetting ER, Edwards RW. Community readiness: a tool for effective community-based prevention. Prev Res 1997;5:5-7.
19. Dunham L, Freeman RE, Liedtka J. Enhancing stakeholder practice: a particularized exploration of community. Bus Ethics Q 2015;16(1):23-42. [doi: 10.5840/beq20061611]
20. FitzGerald C. Multi-level meals on wheels insights: a qualitative study. Work Older People 2023;28(1):41-48 [FREE Full text] [doi: 10.1108/wwop-12-2022-0063]
21. Otto L, Schlieter H, Harst L, Whitehouse D, Maeder A. The telemedicine community readiness model-successful telemedicine implementation and scale-up. Front Digit Health 2023;5:1057347 [FREE Full text] [doi: 10.3389/fdgth.2023.1057347] [Medline: 36910573]
22. Edwards RW, Jumper-Thurman P, Plested BA, Oetting ER, Swanson L. Community readiness: research to practice. J Community Psychol 2000;28(3):291-307 [FREE Full text] [doi: 10.1002/(sici)1520-6629(200005)28:3<291::aid-jcop5>3.0.co;2-9]
23. FitzGerald C, Domegan C, Scharf T. Meals on wheels: a community-led approach and stakeholder analysis. In: Hastings G, Domegan C, editors. Social Marketing: Rebels With a Cause. London: Routledge; 2017:339-345.
24. Stallones L, Gibbs-Long J, Gabella B, Kakefuda I. Community readiness and prevention of traumatic brain injury. Brain Inj 2008;22(7-8):555-564. [doi: 10.1080/02699050802132487] [Medline: 18568708]
25. Kelly KJ, Stanley L. Identifying upstream factors using the community readiness model: the case of reducing alcohol use among college students. J Soc Mark 2014;4(2):176-191. [doi: 10.1108/jsocm-08-2013-0055]
26. Schröder M, Schnabel M, Hassel H, Babitsch B. Application of the community readiness model for childhood obesity prevention: a scoping review. Health Promot Int 2022;37(4):daac120. [doi: 10.1093/heapro/daac120] [Medline: 36047636]
27. Rütten A, Frahsa A, Abel T, Bergmann M, de Leeuw E, Hunter D, et al. Co-producing active lifestyles as whole-system-approach: theory, intervention and knowledge-to-action implications. Health Promot Int 2019;34(1):47-59 [FREE Full text] [doi: 10.1093/heapro/dax053] [Medline: 28973298]
28. Guimarães AL, Lin FV, Panizzutti R, Turnbull A. Effective engagement in computerized cognitive training for older adults. Ageing Res Rev 2025;104:102650. [doi: 10.1016/j.arr.2024.102650] [Medline: 39755175]

XSL•FO
RenderX

29. Grayling MJ, Wason JMS, Mander AP. Stepped wedge cluster randomized controlled trial designs: a review of reporting quality and design features. Trials 2017;18(1):33 [FREE Full text] [doi: 10.1186/s13063-017-1783-0] [Medline: 28109321]
30. Tandoh A, Laar A, Pradeilles R, Le Port A, Osei-Kwasi H, Amevinya GS, et al. Addressing the marketing and availability of unhealthy food and beverages in and around selected schools in Ghana: a community readiness appraisal. BMJ Open 2023;13(9):e075166 [FREE Full text] [doi: 10.1136/bmjopen-2023-075166] [Medline: 37770260]
31. Anderson-Carpenter KD, Watson-Thompson J, Jones MD, Chaney L. Improving community readiness for change through coalition capacity building: evidence from a multi-site intervention. J Community Psychol 2017;45(4):486-499 [FREE Full text] [doi: 10.1002/jcop.21860] [Medline: 28458405]
32. Stanley LR. Community Readiness for Community Change: Tri-Ethnic Center Community Readiness Handbook. 2nd ed. Fort Collins, CO: Tri-Ethnic Center for Prevention Research, Colorado State University; 2014.
33. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess 2015;19(99):1-188. [doi: 10.3310/hta19990] [Medline: 26616119]
34. Silveira SL, Huynh T, Kidwell A, Sadeghi-Bahmani D, Motl RW. Behavior change techniques in physical activity interventions for multiple sclerosis. Arch Phys Med Rehabil 2021;102(9):1788-1800 [FREE Full text] [doi: 10.1016/j.apmr.2021.01.071] [Medline: 33549545]
35. Carvalho CM, de Andrade KR, Poltronieri BC, de Oliveira YG, Ferreira RG, Woodruff E, et al. Translation and validation of the audiovisual version of the Montreal cognitive assessment in older adults in Brazil. BMC Geriatr 2024;24(1):10 [FREE Full text] [doi: 10.1186/s12877-023-04553-2] [Medline: 38172673]
36. Muñoz-Neira C, López OL, Riveros R, Núñez-Huasaf J, Flores P, Slachevsky A. The technology – activities of daily living questionnaire: a version with a technology-related subscale. Dement Geriatr Cogn Disord 2012;33(6):361-371 [FREE Full text] [doi: 10.1159/000338606] [Medline: 22797087]
37. Pachana NA, Byrne GJ, Siddle H, Koloski N, Harley E, Arnold E. Development and validation of the geriatric anxiety inventory. Int Psychogeriatr 2007;19(1):103-114 [FREE Full text] [doi: 10.1017/S1041610206003504] [Medline: 16805925]
38. Sheikh JI, YesavageJA. Geriatric depression scale (GDS): recent evidence and development of a shorter version. Clin Gerontol 2008;5(1-2):165-173. [doi: 10.1300/J018v05n01_09]
39. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
40. Fisher E, Venkatesan S, Benevides P, Bertrand E, Brum PS, El Baou C, et al. Online cognitive stimulation therapy for dementia in Brazil and India: acceptability, feasibility, and lessons for implementation. JMIR Aging 2024;7:e55557 [FREE Full text] [doi: 10.2196/55557] [Medline: 38861708]
41. Camargo BV, Justo AM. IRAMUTEQ: Um Software Gratuito para Análise de Dados Textuais [in Portuguese]. Temas Psicol 2013;21(2):513-518. [doi: 10.9788/tp2013.2-16]
42. Sousa Y. O Uso do Software Iramuteq: Fundamentos de Lexicometria para Pesquisas Qualitativas [in Portuguese]. Estud Pesqui Psicol 2021;21(4):1541-1560. [doi: 10.12957/epp.2021.64034]
43. Nigusie A, Azale T, Yitayal M, Derseh L. Effect of stage-matched educational intervention on use of institutional delivery in Northwest Ethiopia: using community readiness model. Pan Afr Med J 2024;48:16 [FREE Full text] [doi: 10.11604/pamj.2024.48.16.37504] [Medline: 39184843]
44. Kostadinov I, Daniel M, Stanley L, Gancia A, Cargo M. A systematic review of community readiness tool applications: implications for reporting. Int J Environ Res Public Health 2015;12(4):3453-3468 [FREE Full text] [doi: 10.3390/ijerph120403453] [Medline: 25811769]
45. Slater MD, Edwards RW, Plested BA, Thurman PJ, Kelly KJ, Comello MLG, et al. Using community readiness key informant assessments in a randomized group prevention trial: impact of a participatory community-media intervention. J Community Health 2005;30(1):39-53. [doi: 10.1007/s10900-004-6094-1] [Medline: 15751598]
46. Irazoki E, Sánchez-Gómez MC, Contreras-Somoza LM, Toribio-Guzmán JM, Martín-Cilleros MV, Verdugo-Castro S, et al. A qualitative study of the cognitive rehabilitation program GRADIOR for people with cognitive impairment: outcomes of the focus group methodology. J Clin Med 2021;10(4):859 [FREE Full text] [doi: 10.3390/jcm10040859] [Medline: 33669716]
47. González-Palau F, Franco M, Toribio J, Losada R, Parra E, Bamidis P. Designing a computer-based rehabilitation solution for older adults: the importance of testing usability. PsychNology J 2013;11(122):119-136. [doi: 10.3389/fpsyg.2021.636116]
48. Dieciuc M, Zhang S, Gray N, Lustria MLA, Carr DC, Charness N, et al. Terms of engagement: understanding the motivations, preferences, and attitudes of older adults toward mobile cognitive assessment and training. Gerontologist 2024;64(2):gnad048 [FREE Full text] [doi: 10.1093/geront/gnad048] [Medline: 37097773]
49. Whelan J, Love P, Millar L, Allender S, Morley C, Bell C. A rural community moves closer to sustainable obesity prevention – an exploration of community readiness pre and post a community-based participatory intervention. BMC Public Health 2019;19(1):1420 [FREE Full text] [doi: 10.1186/s12889-019-7644-x] [Medline: 31666042]

XSL•FO
RenderX

50. Manera V, Petit P, Derreumaux A, Orvieto I, Romagnoli M, Lyttle G, et al. 'Kitchen and cooking,' a serious game for mild cognitive impairment and Alzheimer's disease: a pilot study. Front Aging Neurosci 2015;7:24 [FREE Full text] [doi: 10.3389/fnagi.2015.00024] [Medline: 25852542]
51. Goghari VM, Lawlor-Savage L. Self-perceived benefits of cognitive training in healthy older adults. Front Aging Neurosci 2018;10:112 [FREE Full text] [doi: 10.3389/fnagi.2018.00112] [Medline: 29922146]
52. Edwards JD, Philllips CB, O'Connor ML, O'Brien JL, Hudak EM, Nicholson JS. Applying the health belief model to quantify and investigate expectations for computerized cognitive training. J Cogn Enhanc 2021;5(1):51-61 [FREE Full text] [doi: 10.1007/s41465-020-00183-3] [Medline: 33817548]
53. Ball C, Francis J, Huang K, Kadylak T, Cotten SR, Rikard RV. The physical-digital divide: exploring the social gap between digital natives and physical natives. J Appl Gerontol 2019;38(8):1167-1184. [doi: 10.1177/0733464817732518] [Medline: 29165009]
54. Tsai HS, Shillair R, Cotten SR, Winstead V, Yost E. Getting grandma online: are tablets the answer for increasing digital inclusion for older adults in the U.S.? Educ Gerontol 2015;41(10):695-709 [FREE Full text] [doi: 10.1080/03601277.2015.1048165] [Medline: 26877583]
55. Kunonga TP, Spiers GF, Beyer FR, Hanratty B, Boulton E, Hall A, et al. Effects of digital technologies on older people's access to health and social care: umbrella review. J Med Internet Res 2021;23(11):e25887 [FREE Full text] [doi: 10.2196/25887] [Medline: 34821564]
56. Djabelkhir L, Wu YH, Vidal JS, Cristancho-Lacroix V, Marlats F, Lenoir H, et al. Computerized cognitive stimulation and engagement programs in older adults with mild cognitive impairment: comparing feasibility, acceptability, and cognitive and psychosocial effects. Clin Interv Aging 2017;12:1967-1975 [FREE Full text] [doi: 10.2147/CIA.S145769] [Medline: 29200836]

Abbreviations

BCT: behavior change technique
CRM: Community Readiness Model
TS: text segment

Edited by R Yang; submitted 29.Nov.2024; peer-reviewed by M Lemonde, A Keleman; comments to author 17.Mar.2025; revised version received 10.May.2025; accepted 18.May.2025; published 19.Feb.2026.

Please cite as:
Carvalho CM, Poltronieri BC, Lima BPDS, Ferreira RG, Reis MEA, FitzGerald C, Lawlor B, Panizzutti R
Community Readiness Model Applied in Older Adults for the Promotion of Digital Cognitive Training: Mixed Methods Feasibility Study
JMIR Aging 2026;9:e69434
URL: https://aging.jmir.org/2026/1/e69434
doi:10.2196/69434
PMID:

©Cíntia Monteiro Carvalho, Bruno Costa Poltronieri, Beatriz Pereira da Silva Lima, Rafaela Guilherme Ferreira, Maria Eduarda Alves Reis, Christine FitzGerald, Brian Lawlor, Rogério Panizzutti. Originally published in JMIR Aging (https://aging.jmir.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Mahjong-Based Cognitive Engagement Intervention for Older Adults: Randomized Controlled Trial

Heng-Hsin Tung[1,2,3], RN, FNP, PhD; Chen-Yuan Kuo[4], PhD; Pei-Lin Lee[2,4], PhD; Chih-Wen Chang[2], MSc; Kun-Hsien Chou[4,5], PhD; Katelyn E Chen[6], BS, MS; Ching-Po Lin[4,5,7], PhD; Chih-Kuang Liang[8,9,10], MD, PhD; Liang-Kung Chen[2,11], MD, PhD

[1]Department of Nursing, National Yang Ming Chiao Tung University, Taipei City, Taiwan
[2]Center for Healthy Longevity and Aging Sciences, National Yang Ming Chiao Tung University, Taipei City, Taiwan
[3]Tungs' Taichung Metro Harbor Hospital, Taichung City, Taiwan
[4]Institute of Neuroscience, National Yang Ming Chiao Tung University, Taipei City, Taiwan
[5]Brain Research Center, National Yang Ming Chiao Tung University, Taipei City, Taiwan
[6]Department of Mechanical Engineering, Stanford University, Stanford, CA, United States
[7]Department of Education and Research, Taipei City Hospital, Taipei City, Taiwan
[8]Center for Geriatrics and Gerontology, Kaohsiung Veterans General Hospital, Kaohsiung City, Taiwan
[9]Department of Geriatric Medicine, School of Medicine, National Yang Ming Chiao Tung University, Taipei City, Taiwan
[10]Division of Neurology, Department of Internal Medicine, Kaohsiung Veterans General Hospital, Kaohsiung City, Taiwan
[11]Taipei Municipal Gan-Dau Hospital (Managed by Taipei Veterans General Hospital), Taipei City, Taiwan

Corresponding Author:
Chih-Kuang Liang, MD, PhD
Center for Geriatrics and Gerontology
Kaohsiung Veterans General Hospital
386 Dazhong 1st Road
Zuoying District
Kaohsiung City, 813414
Taiwan
Phone: 886 73422121 ext 2037
Email: ck.vghks@gmail.com

Abstract

Background: Gamified digital training may support long-term engagement in aging populations, yet randomized evidence linking behavioral effects to convergent neuroimaging outcomes remains limited.

Objective: This study tested whether a home-based digital mahjong-based cognitive engagement intervention enhanced cognition in community-dwelling older adults and explored brain changes using multimodal magnetic resonance imaging (MRI).

Methods: We conducted a single-blind, parallel-group randomized controlled trial in Taipei, Taiwan (July 2020-November 2021). Community-dwelling adults aged 55 years and older were randomized (1:1) to a 6-month digital mahjong-based intervention (two 30-minute sessions per week via an iPad app) with an incentivized walking component (Mi Band step tracking with in-game tokens) or no intervention. Primary MRI outcomes were gray matter volume, regional homogeneity (ReHo), fractional amplitude of low-frequency fluctuations (fALFF), and global brain connectivity. Secondary outcomes included cognition (Montreal Cognitive Assessment), physical activity (International Physical Activity Questionnaire), mental health (Brief Resilience Scale and Mandarin version of the Demoralization Scale), quality of life (EQ-5D with a visual analog scale), physical fitness and body composition, and blood biomarkers. Per-protocol analyses were performed. Voxel-wise imaging analyses used analysis of covariance in SPM12 with family-wise error (FWE) correction at the cluster level (P<.05). Sensitivity analyses controlling for MRI delay duration were conducted to assess robustness of findings.

Results: In total, 60 participants were enrolled (n=30, 50% per group); 96.7% (29/30) in the intervention group and 90% (27/30) in the control group completed primary assessments. Compared with controls, the intervention group showed a greater improvement in Montreal Cognitive Assessment delayed recall (between-group difference=−0.79 points, 95% CI −1.538 to −0.042; P=.04). Neuroimaging revealed increased ReHo in the right insula (FWE-corrected P<.05) and decreased gray matter volume in the left frontal pole and orbitofrontal cortex, with reduced fALFF in the left frontal pole. Changes in delayed recall were positively

XSL•FO
RenderX

associated with right insular ReHo and also showed significant associations with the right anterior cingulate gyrus (ReHo), left inferior frontal gyrus, and right paracingulate gyrus (fALFF; FWE-corrected $P<.05$ in all cases). In exploratory dose analyses, right insular ReHo correlated with total practice sessions ($r=0.48$; $P=.009$). Among physiological outcomes, the intervention group had a larger reduction in diastolic blood pressure (−3.50 mm Hg, 95% CI −6.911 to −0.091; $P=.04$), whereas the thyroid-stimulating hormone decreased but remained within the normal range. Sensitivity analyses confirmed that primary findings remained stable after controlling for MRI delay duration.

Conclusions: This 6-month home-based digital mahjong intervention improved delayed recall and increased right insular ReHo. Integrating culturally familiar cognitive challenges with lifestyle engagement may be a feasible strategy associated with salience network–related functional changes and memory performance. Future large-scale trials should use active control conditions and factorial designs to rigorously isolate component-specific mechanisms.

Trial Registration: ClinicalTrials.gov NCT05808426; https://clinicaltrials.gov/study/NCT05808426

(JMIR Aging 2026;9:e85339) doi:10.2196/85339

KEYWORDS

digital mahjong; cognitive engagement intervention; gamification; cognitive function; healthy aging; magnetic resonance imaging

Introduction

Background

Declines in physical and cognitive functions are major challenges for aging populations. Although multidomain interventions can help, maintaining long-term adherence remains difficult for many older adults. The World Health Organization healthy aging framework and the US National Academy of Medicine's healthy longevity framework [1,2] indicate that effective strategies should target physical and cognitive health simultaneously rather than separately. Therefore, combining exercise with cognitive training is a recommended approach [3,4].

As neurodegenerative disorders become more prevalent, preventing functional decline is a priority [5,6]. However, purely physical or cognitive interventions often fail to maintain long-term engagement [7,8]. Previous studies have shown that multidomain interventions can promote healthy longevity; however, their benefits often diminish if participants do not continue the activities after the intervention ends [4,9]. Therefore, sustainable approaches to healthy aging must involve scientifically grounded, multifaceted strategies capable of maintaining participant engagement over the long term.

The interactive nature of digital gaming offers distinct advantages over traditional training, particularly regarding immediate feedback and user engagement. This interactivity makes digital gaming a promising tool for improving adherence in older adults [10] while also providing a feasible and relatively safe approach to support cognitive and physical activity [11]. There is recent evidence further suggesting that digital games may benefit cognitive performance and motor-related outcomes in older adults, particularly when the interventions are structured and appropriately tailored [12]. Research by Ding et al [13] and Zhang et al [14] has shown that digital versions of games such as mahjong can improve cognitive function, support mental health, and reduce cognitive decline in older adults. Mahjong, a traditional Chinese tile-based game, requires players to perform coordinated visual, mental, and cognitive tasks to achieve winning combinations. Studies indicate that both traditional and digital forms of mahjong can positively impact early-stage dementia in older adults [15].

Mechanistically, digital game training engages multiple cognitive domains, such as executive function, attention, and memory, through multitasking activities [16-18]. This process requires the brain to integrate functions across different regions to generate appropriate responses. Research has shown that, after multitasking video game training, older adults can achieve cognitive performance levels comparable to those of control participants aged 20 years [19]. Furthermore, complex game training not only improves cognitive performance in older adults but also induces significant changes in brain function and structure [20-23]. These neural changes demonstrate substantial neuroplasticity in response to targeted cognitive training, suggesting that mahjong-based game training could be an effective approach to promote brain plasticity [24]. Despite the importance of structural and functional brain markers in understanding cognitive aging, there have been few randomized controlled trials investigating the effects of digital mahjong gameplay on brain measures, cognitive function, and overall health.

We selected mahjong as the intervention for 3 reasons. First, it requires working memory, executive function, and attention, aligning with multidomain cognitive training principles [19,24]. Second, observational and experimental evidence links mahjong practice to preserved cognition [13-15]. Third, unlike laboratory tasks, mahjong is culturally familiar, which may enhance intrinsic motivation and adherence [10]. In contrast to traditional computerized cognitive training that isolates specific domains, our cognitive engagement strategy embeds training within a familiar context to improve ecological validity. We combined the core digital mahjong game with incentivized walking and gamified progression to sustain engagement. This design balances efficacy with real-world acceptability. While targeted serious games can effectively engage specific processes, our complementary approach prioritizes cultural relevance and sustainable engagement.

Objectives

In this study, we aimed to evaluate whether a 6-month digital mahjong program could improve cognitive function and be

XSL•FO
RenderX

associated with magnetic resonance imaging (MRI)–derived brain changes. We also tested whether adding a walking incentive would increase physical activity. Our study compared the intervention to a no-intervention control group across multiple domains, including MRI-derived brain measures, global and domain-relevant cognitive performance, overall health status, and blood biomarkers. We hypothesized that this integrated intervention would improve cognitive function and be associated with measurable changes in brain structure and function, reflecting the combined effects of sustained cognitive engagement and lifestyle modification.

Methods

Study Design

This was a single-blind, parallel-group randomized controlled trial registered at ClinicalTrials.gov (NCT05808426). Using a computer-generated random number sequence, we randomly assigned eligible participants in a 1:1 ratio to either an intervention group or a control group without any stratification. The control group received no intervention.

Sample Size Estimation and Participants

We performed an a priori power analysis using G*Power (version 3.1.9.7) to determine the required sample size. The sample size was calculated based on Mini-Mental State Examination data reported in a previous study examining the effects of digital game interventions on cognitive function in older adults [20] in which the intervention and control groups showed mean scores of 28.87 (SD 1.64) and 27.72 (SD 1.60), respectively. We input these values into G*Power, and the resulting effect size was approximately 0.71. Assuming a 1-tailed *t* test with an α value of .05 and power $(1 - \beta)$ of 0.8, the analysis indicated that 26 participants per group would be needed. To account for an anticipated attrition of approximately 10%, we targeted a recruitment of 30 participants in each group. Ultimately, 30 participants were enrolled in each group; by the end of the study, 96.7% (29/30) participants in the intervention group and 90% (27/30) in the control group completed the trial.

Ethical Considerations

This study was approved by the institutional review board of National Yang Ming Chiao Tung University (YM108123F). All participants provided written informed consent in accordance with the Declaration of Helsinki. Participants were informed that they could withdraw from the study at any time, and all personal data were deidentified to ensure privacy and confidentiality. At the time of study design (2019-2020), it was genuinely unknown whether the intervention was efficacious. This study was approved as a minimal-risk proof-of-concept trial. A waitlist control design was not feasible due to resource constraints and COVID-19–related delays.

Procedure

The program design, trial protocol, and intervention completion rate are provided in Multimedia Appendix 1 and Table S1 in Multimedia Appendix 2. Before starting the mahjong intervention, all participants in both groups completed baseline assessments, including questionnaires, physical examinations, blood tests, and brain MRI scans. The intervention content (the digital mahjong game and related procedures) was beta tested by the sponsoring company and piloted with a group of older adults not enrolled in the study.

This trial recruited community-dwelling adults aged 55 years and above in Taipei, Taiwan, between July 2020 and November 2021, corresponding to a 6-month intervention period for each participant. Data were collected at baseline and immediately after the intervention. Participants were enrolled via convenience sampling responding to advertisements posted in local health centers. This study was interrupted for approximately 3 months due to the COVID-19 pandemic. Participants assigned to the intervention group engaged in the digital mahjong program for the full 6-month period as planned. However, pandemic-related restrictions on physical activity caused both groups to experience a 6- to 8-month delay in undergoing postintervention brain MRI scans and other functional assessments.

Because the intervention was delivered individually and remotely, participants' start and end dates varied. For example, the first participant underwent the baseline assessment on July 28, 2020; completed the 6-month digital mahjong game intervention by January 5, 2021; and then underwent the postintervention evaluation (including the MRI scan) on September 2, 2021 (approximately 8 months after intervention completion). The last participant finished all postintervention assessments by October 13, 2021. We paused the study for 3 months during Taiwan's level 4 COVID-19 alert as restrictions prevented the delivery of Wi-Fi SIM cards. Make-up sessions were provided once restrictions eased. To mitigate potential decay of training effects due to the delay between intervention completion and postintervention MRI, we administered a 1-month "booster" retraining session to the intervention group immediately prior to the scheduled MRI assessment based on perceptual learning principles suggesting that brief refresher training can help consolidate and maintain cognitive gains.

Randomization and Blinding

Participants were not informed of their group allocation, and they had no contact with the research team during the study. The research personnel were blinded to group assignments: all participant data were labeled with anonymous codes, so investigators could not discern which intervention a given participant received. Moreover, the staff members who delivered the intervention were separate from those who conducted outcome evaluations and data analysis. The research team also ensured that participants from different groups did not interact with one another throughout the trial.

Intervention Program

The digital mahjong game was delivered via an iPad app (the Bonus Winner app), in which a single participant plays against 3 computer-controlled players, replicating the traditional game in a virtual format. The app was a commercially available platform adapted for research purposes with the developer's cooperation. The key modification for this study was the implementation of a structured difficulty progression system based on response time constraints, allowing for parametric control over cognitive load. The research team provided an

XSL•FO
RenderX

orientation to the digital mahjong platform and supplied each intervention group participant with an iPad. As part of the gamified design, participants could earn virtual bonus points through walking activity; therefore, each participant was given a Xiaomi Mi Band wristband to track their steps. A research assistant set up each participant's device by installing a SIM card for Wi-Fi connectivity, instructed participants on using the mahjong app, and monitored their gameplay progress remotely. The assistant also demonstrated how the Mi Band's recorded walking distance would be converted into in-game bonuses.

The digital mahjong game protocol (see Multimedia Appendix 1 for details) used a progressive difficulty system with real-time interactive gameplay. The game consisted of 4 levels, each with increasingly stringent response time limits:

- Level 1: 20-second response window (familiarization phase; month 1)
- Level 2: 7-second response window (standard training; months 2-6)
- Level 3: 5-second response window (advanced; optional)
- Level 4: 3-second response window (highest difficulty; optional)

To minimize ceiling effects and baseline heterogeneity, we excluded participants with regular mahjong experience (playing >3 times per year). However, we permitted basic rule familiarity to avoid the confounding load of learning game mechanics from scratch. All participants spent the first month at level 1 to ensure mastery of the digital interface before progressing. This graduated structure facilitates adaptation to increasing cognitive loads. To sustain motivation, we implemented a coin-based progression system that incentivized advancement while mitigating the frustration of losses. Comprehensive intervention details are provided in Multimedia Appendix 1.

Outcomes

The primary outcome was defined as changes in brain structure and function as measured via MRI. MRI data acquisition and image processing for structural and functional analyses were conducted using standard protocols.

All MRI scans were performed on a 3T Siemens scanner at National Yang Ming Chiao Tung University. T1-weighted anatomical images were acquired using a 3D magnetization prepared rapid gradient echo sequence with the following parameters: repetition time=3500 ms; echo time=3.5 ms; inversion time=1100 ms; flip angle=7°; field of view=256 × 256 mm^2; matrix size=256 × 256; voxel size=1.0 × 1.0 × 1.0 mm^3; 192 sagittal slices; 1 excitation. Resting-state functional MRI (fMRI) data were collected using a T2*-weighted gradient-echo echo-planar imaging sequence using the following parameters: repetition time=2500 ms; echo time=30 ms; flip angle=90°; slice thickness=3.5 mm; image matrix size=64 × 64; field of view=222 × 222 mm; 43 interleaved slices (no gap); 200 volumes (acquisition of approximately 8 minutes and 27 seconds). All raw MRI scans were reviewed by an experienced researcher to ensure no anatomical abnormalities or substantial artifacts before further processing.

To characterize intervention-related changes in the brain, we evaluated four neuroimaging metrics: (1) gray matter volume (GMV; derived from structural MRI), which reflects regional GMV differences and was used as a structural MRI marker; (2) fractional amplitude of low-frequency fluctuations (fALFF; derived from resting-state fMRI), which reflects neural activity levels and their relationship to brain metabolic rate changes [25,26]; (3) regional homogeneity (ReHo; derived from resting-state fMRI), which measures the similarity of neural activity among neighboring brain regions [27]; and (4) global brain connectivity (GBC; derived from resting-state fMRI), which measures the interconnectivity of disparate brain regions, reflecting interactions among large-scale brain networks [28].

Detailed preprocessing methods for these MRI-derived features are provided in Multimedia Appendix 1. Whole-brain voxel-wise maps for GMV, ReHo, fALFF, and GBC were generated for statistical analysis.

Secondary outcomes were assessed using validated tools across cognitive, physical, mental, and quality of life domains. Cognitive function was measured using the Montreal Cognitive Assessment (MoCA), a 30-point cognitive screening tool covering multiple domains (orientation, memory, executive and visuospatial function, language, abstraction, naming, attention, and clock drawing) [29,30]. The MoCA takes approximately 10 to 12 minutes to administer, and scores of 26 or higher are considered normal. It has strong psychometric properties and is widely used as a reliable screening instrument for dementia.

Physical activity was measured using the short form of the International Physical Activity Questionnaire (IPAQ; 7-day recall) [31,32]. This questionnaire captures physical activity across work, household, transportation, and leisure domains at 3 intensity levels. Weekly physical activity was calculated in metabolic equivalent of task (MET) minutes by summing the time spent walking (3.3 METs), in moderate-intensity activities (4.0 METs), and in vigorous-intensity activities (8.0 METs). The short form of the IPAQ has excellent content validity (content validity index=0.994) and adequate test-retest reliability (r=0.67).

Mental health was evaluated using 2 self-report scales: the Brief Resilience Scale [33] and the Demoralization Scale (Mandarin version) [34-36]. The Brief Resilience Scale is a 6-item questionnaire measuring an individual's resilience, with items rated on a 5-point Likert scale (1="strongly disagree"; 5="strongly agree"). After reverse scoring the 3 negatively worded items, responses are averaged, with higher scores indicating greater resilience. The Mandarin version of the Demoralization Scale consists of 24 items rated from 0 ("never") to 4 ("all of the time"), with a total score of 30 or higher indicating demoralization syndrome (higher scores reflect greater severity of demoralization).

Quality of life was assessed using the EQ-5D questionnaire [37,38]. This instrument evaluates 5 domains: mobility, self-care, usual activities, pain and discomfort, and anxiety and depression, each on a 3-level scale ("no problems," "some or moderate problems," and "extreme problems"). Participants also rated their overall health on the EQ-5D's visual analog scale from 0 (worst imaginable health) to 100 (best imaginable

XSL•FO
RenderX

health). The EQ-5D (including the visual analog scale) is a widely used and validated measure of general health status in older adult populations.

Additionally, we administered standard physical fitness tests, measured body composition using an InBody S10 bioimpedance analyzer, and collected blood samples for a panel of biochemical markers. Other recorded information included participants' demographic characteristics, health-related behaviors (eg, tobacco and alcohol use), and any self-reported physician-diagnosed medical conditions.

Statistical Analysis

All analyses were conducted on a per-protocol basis comparing the intervention and control groups' outcomes from baseline to the postintervention time point. Continuous variables were summarized as means and SDs and compared between groups using appropriate significance tests (2-tailed Student *t* test for parametric data or Mann-Whitney *U* test for nonparametric data). Categorical variables were summarized as counts and percentages and compared using chi-square or Fisher exact tests. Statistical significance was set at a 2-sided *P* value of less than .05 (95% CIs that excluded the null value). Data were analyzed using SPSS Statistics (version 25.0; IBM Corp).

For primary outcomes derived from neuroimaging measures, we used a difference-in-differences approach to compare changes from baseline to the 6-month follow-up between groups. Specifically, after preprocessing, individual posttest-minus-pretest change maps were estimated for each MRI metric by subtracting each participant's baseline image map from the corresponding postintervention image map. These voxel-wise change maps were then entered into group-level analyses in SPM12 to test whether longitudinal changes differed between the intervention and control groups. As the dependent variables were posttest-minus-pretest change maps, baseline MRI maps were incorporated into the calculation of within-participant change rather than entered as separate covariates. We used analysis of covariance models to test for intervention effects on MRI outcomes adjusting for relevant covariates (baseline age, sex, and intracranial volume for structural measures and age and sex for functional measures). To correct for multiple voxel-wise comparisons, we applied a cluster-level family-wise error (FWE) correction at a *P* value of less than .05 using the AFNI program's 3dClustSim (version 20.3.03) with an initial voxel-wise threshold of a *P* value below .005. The minimum cluster size for significance was determined separately for each imaging metric (eg, 166 voxels for GMV, 57 voxels for ReHo, 19 voxels for fALFF, and 10 voxels for GBC).

To assess the robustness of our neuroimaging findings to variable delays between intervention completion and postintervention MRI assessment, we conducted sensitivity analyses including MRI delay duration as an additional covariate in all analysis of covariance models. This approach allowed us to evaluate whether observed brain changes were confounded by the timing of postintervention assessments. Details and results of these sensitivity analyses are provided in Multimedia Appendix 3.

To explore dose-response relationships in the intervention group, we examined whether the amount and type of mahjong practice correlated with brain changes. We first extracted the mean values from each participant's significant brain clusters (identified in the GMV, ReHo, fALFF, and GBC analyses). Then, using partial Pearson correlations (controlling for age, sex, and MRI delay duration in sensitivity analyses), we tested associations between these brain change metrics and the number of completed mahjong sessions at various difficulty levels (total sessions and sessions with 20-, 7-, 5-, and 3-second response time limits). The significance for these correlations was evaluated at a false discovery rate–corrected level of a *P* value below .05.

We also investigated whether improvements in cognitive scores were associated with neuroimaging changes. For each participant, we computed the change (baseline to posttest) in cognitive performance (eg, MoCA delayed recall) and extracted the corresponding change maps for brain outcomes. Using SPM12, we built multiple linear regression models to test the association between cognitive improvement and brain changes adjusting for age and sex (and also adjusting for baseline intracranial volume for structural MRI measures). These regression analyses did not yield any significant voxel-wise associations at the predefined threshold (cluster-level FWE $P<.05$). Further details of all statistical procedures are provided in Multimedia Appendix 1.

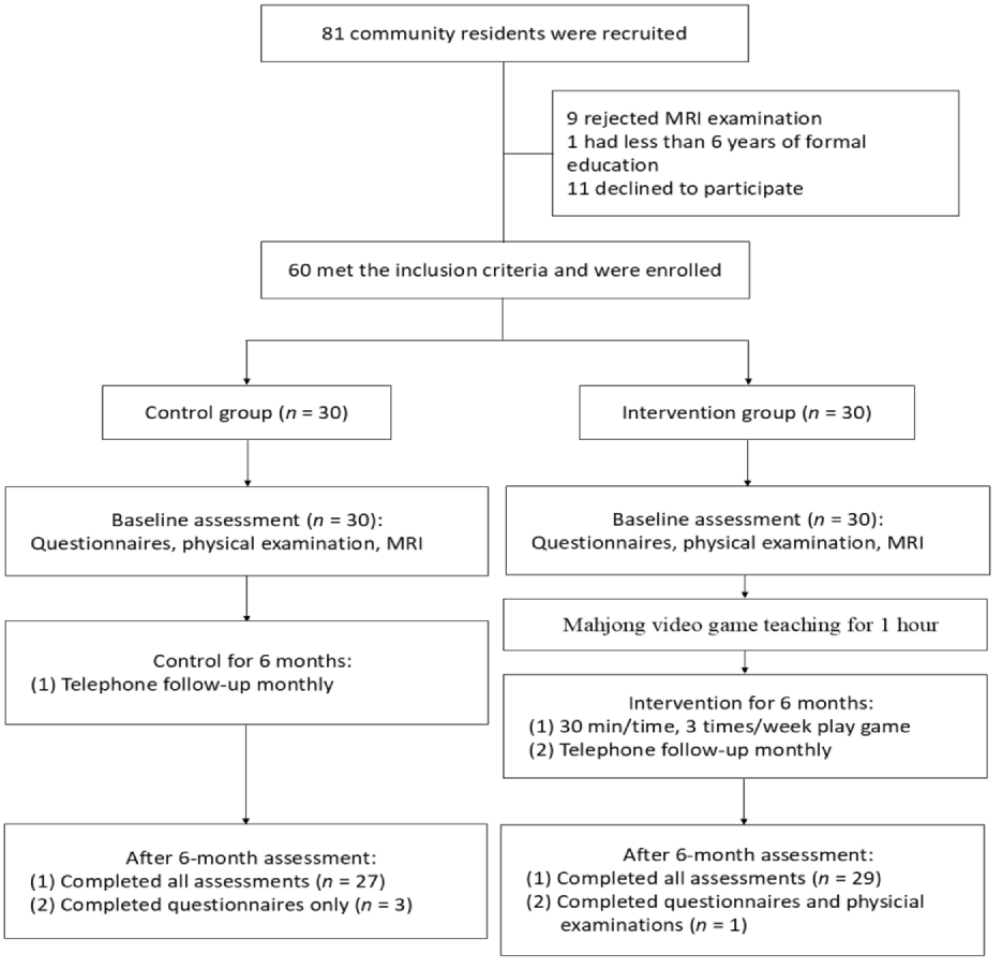
Results

Participants

Figure 1 shows the participant flow from initial screening and enrollment to random allocation (30 participants per group). Participant characteristics included a mean age of 64.45 (SD 5.23) years in the intervention group and 68.07 (SD 5.78) years in the control group ($P=.27$). The intervention group consisted of 34.5% (10/29) male participants, whereas the control group comprised 40.7% (11/27) male participants ($P=.78$). The groups did not differ significantly on baseline physical activity (IPAQ), physical fitness, or body composition measures, supporting comparability prior to randomization. Additional baseline demographic characteristics for each group are provided in Table S2 in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram indicating the number of participants enrolled, the randomization process, and the study flow. MRI: magnetic resonance imaging.



Intervention Engagement

Detailed engagement metrics are provided in Table S1 in Multimedia Appendix 2. Briefly, participants in the intervention group completed a mean total of 825.2 (SD 615.0; range 289-3055) sessions over the 6-month period. All participants in intervention group (30/30, 100%) completed the familiarization phase (level 1; 20-second response window), 100% (30/30) progressed to level 2 (7-second response window), 100% (30/30) advanced to level 3 (5-second response window), and 86.7% (26/30) attempted the highest difficulty (level 4; 3-second response window). The mean number of daily steps recorded via Mi Band was 8826 (SD 3205.4; range 85.0-19,543.5). Protocol adherence (completing ≥80% of scheduled sessions) was achieved by all participants, with a mean completion rate of 94.44% (SD 6.09%; range 83.33%-100%).

Primary Outcome

At the 6-month follow-up, the intervention group exhibited significant intervention-related changes in brain structure and function compared to controls. In voxel-wise analyses (Figure 2; Table S3 in Multimedia Appendix 2), the intervention group showed reduced GMV in the left frontal pole and left orbitofrontal cortex, as well as decreased fALFF in the left frontal pole. Conversely, ReHo in the right insula was significantly increased in the intervention group (cluster-level FWE P<.05). Notably, none of these brain changes were significantly associated with the frequency of digital mahjong gameplay during the intervention.

The regression analyses examining cognitive-brain relationships indicated that changes in GMV were not significantly related to improvements in cognitive performance (eg, MoCA delayed recall). In contrast, changes in MoCA delayed recall scores were significantly associated with functional connectivity changes: a greater improvement in delayed recall was associated with an increase in ReHo in the right insula and a decrease in ReHo in the right paracingulate gyrus and right frontal pole (cluster-level FWE P<.05 in all cases; Figure 3; Table S4 in Multimedia Appendix 2). Similarly, for fALFF, higher delayed recall gains were linked to increased fALFF in the left inferior frontal gyrus and decreased fALFF in the right paracingulate gyrus (FWE P<.05). Finally, intervention participants who completed more training sessions, especially those with faster response times (7- and 5-second limits), exhibited higher ReHo in the right insula. In exploratory dose analyses, right insular ReHo correlated with total practice sessions (r=0.48; P=.009; Figure 4).

XSL•FO
RenderX

Sensitivity analyses controlling for MRI delay duration confirmed the robustness of the primary findings. Detailed sensitivity analysis results are provided in Multimedia Appendix 3. Group-level results for GMV and fALFF remained consistent with those of the original analyses (Table S1 and Figure S1 in Multimedia Appendix 3). Although the ReHo activation pattern in the right insula did not survive strict cluster-level FWE correction in the adjusted model due to reduced statistical power, the effect remained visible at a relaxed threshold (uncorrected $P<.05$; cluster of >100 voxels). Importantly, the brain-behavior correlations and dose-response relationships remained significant after controlling for the delay, with peak coordinates identical to those of the original analysis (Table S2 and Figures S2 and S3 in Multimedia Appendix 3).

Figure 2. Brain measure changes between the control and intervention groups. Hot (red-yellow) regions reflect a significantly increased change in the intervention group. Cold (blue-green) regions reflect a significantly decreased change in the intervention group. fALFF: fractional amplitude of low-frequency fluctuations; FWE: family-wise error; GMV: gray matter volume; ReHo: regional homogeneity.

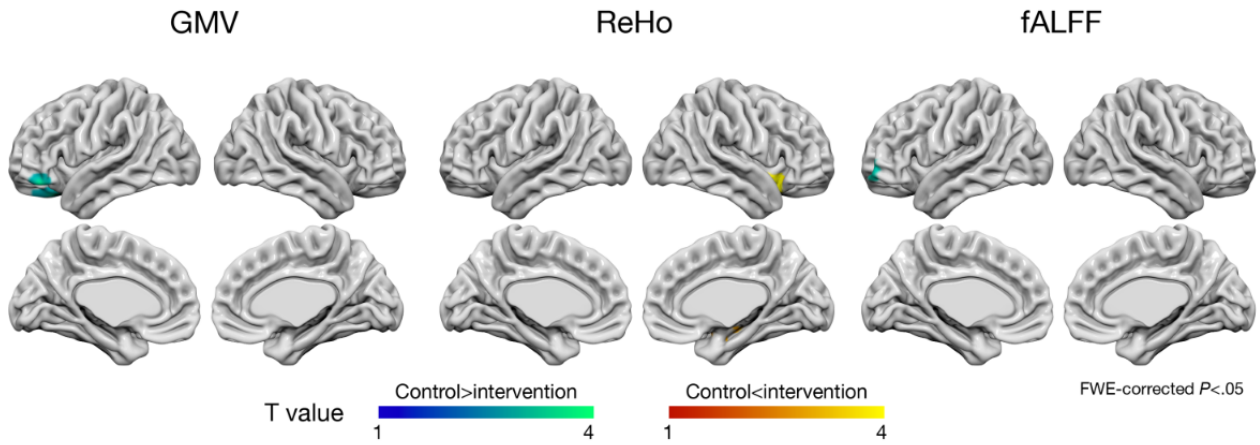
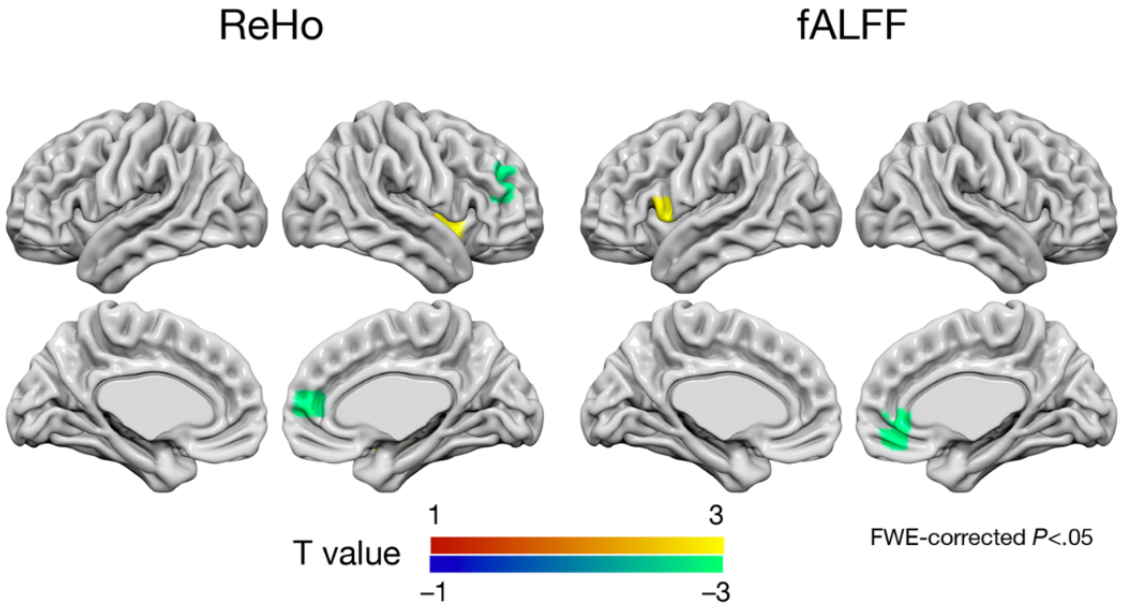
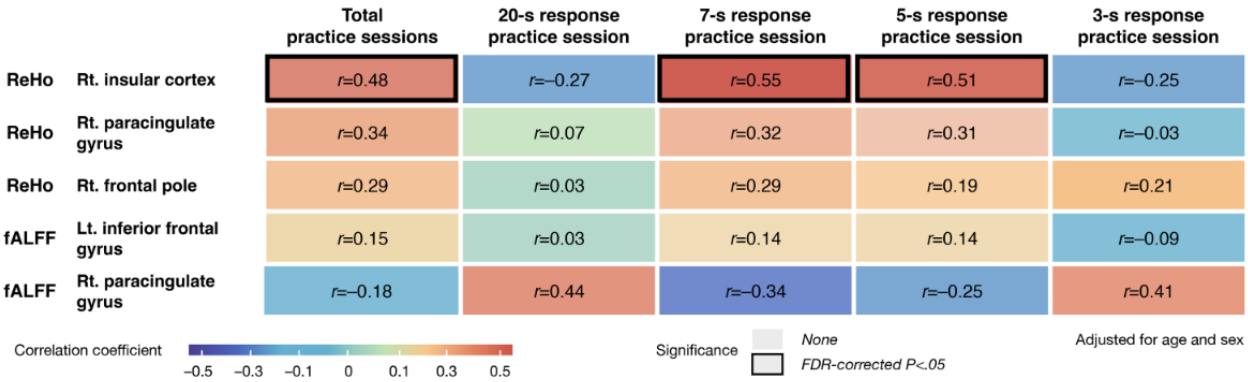


Figure 3. Associations between changes in intrinsic functional brain measures and improvements in delayed recall performance. Multiple linear regression analyses were performed to examine the associations between changes in intrinsic functional brain measures (regional homogeneity [ReHo] and fractional amplitude of low-frequency fluctuations [fALFF]) and improvements in Montreal Cognitive Assessment (MoCA) delayed recall scores. Analyses were adjusted for age and sex. Hot (red-yellow) regions reflect a significantly positive correlation with the cognition outcome. Cold (blue-green) regions reflect a significantly negative correlation with the cognition outcome. FWE: family-wise error.



XSL•FO
RenderX

Figure 4. Associations between brain signature changes and the number of mahjong training sessions completed at different response-time levels. Relationships were estimated using partial Pearson correlation analysis accounting for age and sex. Significant correlations are marked with a black box (false discovery rate [FDR]–corrected *P*<.05). fALFF: fractional amplitude of low-frequency fluctuations; Lt: left; ReHo: regional homogeneity; Rt: right.



Secondary Outcomes

Table S5 in Multimedia Appendix 2 summarizes the between-group comparisons for cognitive function, physical activity, psychological measures, quality of life, body composition, and blood biomarkers. The intervention group demonstrated a significantly greater improvement in MoCA delayed recall from baseline than the control group (between-group difference=−0.79, 95% CI −1.538 to −0.042; *P*=.04). Both groups completed identical MoCA assessments at baseline and follow-up under the same conditions. Importantly, the improvement was specific to delayed recall rather than domains more susceptible to practice effects (eg, orientation and naming), suggesting domain-relevant gains aligned with the memory demands of the intervention rather than generalized practice effects. The intervention group also reported a reduction in weekday sitting time relative to baseline, but this decrease was not significantly different from that in the control group (+71.7 minutes, 95% CI −7.8 to 141.1; *P*=.08). There were no significant between-group differences in resilience, demoralization, or quality of life outcomes.

Physical examination results showed that the intervention group had a significant reduction in diastolic blood pressure (3.50 mm Hg, 95% CI, 0.091-6.911; *P*=.04). Although the control group showed greater decreases in muscle mass measurements (right leg, left leg, total limb muscle mass, and relative appendicular skeletal muscle), these differences were not statistically significant between groups (Table S5 in Multimedia Appendix 2).

Biochemically, the intervention group showed a significant decrease in serum thyrotropin levels compared to the control group (−0.545 µIU/mL, 95% CI −1.018 to −0.072; *P*=.02). As all participants' thyrotropin level values remained within the normal range, the clinical significance of this drop is unclear. Total cholesterol levels changed slightly in the intervention group relative to controls, but there were no corresponding significant changes in high-density lipoprotein or low-density lipoprotein cholesterol.

Finally, exploratory hierarchical regression analyses (Table S6 in Multimedia Appendix 2) were conducted to investigate how the extent of engagement with the intervention related to the secondary outcomes. These analyses suggested that participants with longer average daily play time tended to have greater reductions in total cholesterol (*P*=.03) and those who frequently played with the 20-second response window tended to have higher high-density lipoprotein cholesterol (*P*=.047). Additionally, the total number of incentivized walking steps was positively correlated with thyroid-stimulating hormone levels at 7 months (*P*=.03).

Discussion

Principal Findings

Our study showed that a home-based digital mahjong-based cognitive engagement intervention significantly improved delayed recall in older adults. We also observed functional changes in the right insula, suggesting that the training involved specific brain networks. While many health interventions are effective in the short term, adherence often declines over time. Our study demonstrates that a culturally adapted digital tool can sustain engagement and improve delayed recall. Gamification has emerged as a promising approach to encourage healthy behaviors, but its long-term efficacy relative to more traditional interventions remains to be fully elucidated.

The findings from our trial indicate that the digital mahjong-based cognitive engagement intervention significantly enhanced delayed recall ability and showed accompanying changes in brain functional measures. The observed effects reflect the combined intervention package (digital mahjong, incentivized walking, and gamification) rather than any single component. In particular, the most challenging training sessions (with 7- and 5-second response time limits) showed the strongest associations with right insular ReHo. These findings align with those of previous research showing that cognitive training can enhance memory [39,40]. Our study adds exploratory neuroimaging findings to behavioral findings. Specifically, we observed increased ReHo in the right insula. The preferential engagement of this region during high-intensity (5- and

XSL•FO
RenderX

7-second) sessions may reflect increased demands on decision urgency and error monitoring. This pattern is consistent with the potential involvement of the salience network in mediating rapid switching between internal processing and external stimuli. However, because the right insular ReHo finding was attenuated in sensitivity analyses after controlling for MRI delay duration, this result should be interpreted with caution given its sensitivity to model specification.

We also noted significant GMV reduction in the frontal pole, a region critical for working memory [41,42]. Given the concurrent improvement in delayed recall performance, this decrease may be compatible with experience-related structural reorganization or enhanced neural efficiency rather than necessarily indicating maladaptive atrophy. This finding is consistent with those of recent randomized controlled trials, which similarly found no evidence of macroscopic structural growth (volume expansion) despite rigorous cognitive training. Together, these findings support the hypothesis that functional connectivity alterations (such as ReHo) may precede or accompany structural changes [43]. Behaviorally, we further noticed that participants relied heavily on in-game cues during the most challenging (7 and 5 seconds) response time sessions. This suggests that the high cognitive load necessitated external support, reinforcing the possibility that intensive cognitive engagement may be associated with changes in specific brain networks.

This trial demonstrated that a gamified cognitive engagement intervention can significantly improve cognitive performance (specifically delayed recall) in older adults, corroborating previous findings that cognitively stimulating games such as mahjong yield cognitive benefits [13,14]. Our study uniquely combined the digital game with an incentive program: participants earned virtual tokens for in-game rewards by walking, thereby motivating both cognitive engagement and physical activity. Unfortunately, the onset of the COVID-19 pandemic likely blunted the full benefits of this design. As infection concerns grew, participants became less willing to go outside, reducing their walking activity despite the token incentive. Indeed, pandemic-related restrictions and safety measures have been shown to adversely affect older adults' physical, cognitive, and mental well-being [44,45]. Although Taiwan experienced a relatively mild outbreak, these disruptions probably limited the effectiveness of our walking-based token system to simultaneously promote exercise and cognitive training.

Mahjong gameplay has also been linked to mental health benefits (eg, lower depression and better well-being), presumably through increased social interaction [46]. In our study, however, the digital format did not yield measurable gains in physical performance, psychological resilience, demoralization, or quality of life relative to controls. One explanation for the absence of these effects is a possible ceiling effect: our participants were relatively healthy at baseline, leaving little room for improvement on those measures. Moreover, both groups experienced comparable pandemic-related stress during the trial, which could have obscured subtle intervention benefits. Despite these null findings, our intervention did produce robust gains in delayed recall along with corresponding brain changes. The integrated walking incentive successfully promoted physical activity, with participants averaging 8826 (SD 3205.4) steps per day despite pandemic restrictions. This sustained activity level suggests that the gamified token economy remained effective in motivating physical movement even under constrained environmental conditions, which likely contributed to improvements in cardiovascular indicators. In fact, the intervention group showed significant reductions in diastolic blood pressure and total cholesterol, outcomes consistent with increased daily activity levels [47,48].

Unlike targeted serious games that isolate specific cognitive processes, our study prioritized ecological validity and cultural acceptability. Mahjong engages multiple domains simultaneously, mirroring real-world cognitive demands. We acknowledge the trade-off between the mechanistic precision of laboratory tasks and the implementability of lifestyle interventions. However, culturally meaningful activities such as digital mahjong often promote better long-term adherence than abstract computerized cognitive training programs. This sustained engagement may yield greater cumulative benefits over time even if the training target is less precise. Future studies should directly compare serious games with culturally adapted activities to guide recommendations for diverse populations.

Limitations

Our study has several limitations. First, the lack of an active control group limits our ability to exclude placebo effects. Participants in the intervention group received staff attention, technology, and incentives, all of which could trigger Hawthorne or expectancy effects. However, the observed gains were most evident in delayed recall, a core demand of mahjong, rather than generalized improvements. This pattern, combined with the neuroimaging correlations, suggests but does not establish intervention-specific effects. Future studies should use active controls, such as nonstrategic games or educational activities [43], to isolate the specific effects of mahjong training.

Second, COVID-19 restrictions necessitated a 6- to 8-month delay between intervention completion and MRI assessment. To mitigate potential signal decay, we administered a 1-month booster session prior to scanning and controlled for delay duration in sensitivity analyses [43]. Although these analyses confirmed the stability of brain-behavior correlations, we acknowledge that the observed neuroimaging differences may partially reflect consolidation processes or environmental factors during the hiatus rather than acute training effects alone.

Third, our intervention combined digital mahjong with incentivized walking, preventing us from isolating the specific contribution of each component. Consequently, the reported effects reflect the aggregate intervention. However, there are 2 lines of evidence suggesting that cognitive training was the primary driver of the observed neuroimaging differences. First, we found a specific dose-response relationship between mahjong practice intensity and right insular ReHo, supporting a link between cognitive training intensity and neural changes. Second, behavioral improvements were specific to delayed recall rather than physical function measures. While these findings support

XSL•FO
RenderX

the specificity of the cognitive training, we cannot rule out synergistic effects between the cognitive and physical elements.

Fourth, because the incentive (token) program was applied only to the intervention group, we could not evaluate whether control participants might also have increased their physical activity without incentives. That said, it is generally recognized that specific incentives are often required to induce meaningful behavior change in health promotion programs.

Fifth, the COVID-19 pandemic and resulting restrictions led to a temporary suspension of the trial, interrupting the link between walking and token rewards. Nevertheless, the core element of our intervention, the digital mahjong game, remained relatively resilient to pandemic disruptions. Notably, previous neuroplasticity trials have found that training gains can persist for 3 to 6 months after an intervention even if performance at follow-up is somewhat below the posttraining peak [19,49]. To help sustain the cognitive benefits in our study, we administered a 1-month "booster" retraining to the intervention group approximately 1 month before the final assessment following perceptual learning principles to reinforce memory traces [50]. Despite these efforts, the pandemic-related delays postponed our postintervention evaluations, which may have attenuated the observed effects. Therefore, our results could be underestimating the true efficacy of the intervention.

We were also concerned about this study not monitoring participants' use of the mahjong app after completion of the intervention; therefore, the sustainability of engagement, continued gameplay habits, and potential risks of excessive use could not be evaluated. Future studies should include longitudinal follow-up assessments to examine both the maintenance of intervention benefits and the promotion of appropriate, balanced use of gamified digital health apps.

Conclusions

This randomized controlled trial suggests that a digital mahjong-based cognitive engagement intervention improved delayed recall and was associated with MRI-derived functional changes (right insular ReHo) in middle-aged and older adults. We also observed secondary benefits for diastolic blood pressure and total cholesterol. These findings support the viability of integrating culturally meaningful activities into lifestyle interventions to promote healthy aging. Future large-scale trials should use active control conditions and factorial designs. Such rigor is necessary to distinguish the specific effects of cognitive training from those of physical activity and optimize these tools for routine geriatric care.

Acknowledgments

The authors would like to thank all the participants. They also would like to thank the Interdisciplinary Research Center for Healthy Longevity of National Yang Ming Chiao Tung University from the Featured Areas Research Center Program within the framework of the Higher Education Sustained Progress and Rise of Universities in Taiwan Project by the Ministry of Education in Taiwan for supporting this research. No generative artificial intelligence tools were used to generate scientific content, analyses, interpretations, or conclusions in this manuscript. All scientific content was developed and written by the named authors. Language editing was performed solely to improve clarity and readability.

Funding

This study was sponsored by Bonus Winner Mahjong, Taipei, Taiwan. The sponsor had no role in the design and conduct of the study; collection, analysis, and interpretation of the data; or preparation, review, and approval of the manuscript.

Data Availability

The data that support the findings of this study are available from the corresponding author on reasonable request.

Authors' Contributions

HHT contributed to conceptualization, data curation, formal analysis, methodology, project administration, validation, and writing—original draft. CYK contributed to data analysis, brain imagery analysis, validation, and visualization. PLL contributed to data analysis, brain imagery analysis, validation, and visualization. CWC contributed to data curation, investigation, and project administration. KHC contributed to validation and visualization. KEC contributed to data analysis and brain imagery analysis. CPL contributed to validation and visualization. CKL contributed to conceptualization, methodology, validation, writing—review and editing, and final approval of the manuscript. LKC contributed to conceptualization, writing—review and editing, and final approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Trial protocol and statistical analysis plan.
[DOCX File , 24 KB - aging_v9i1e85339_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Supplementary tables describing the digital mahjong intervention protocol, participant characteristics, magnetic resonance imaging findings, intervention outcomes, and exploratory analyses.
[DOCX File , 35 KB - aging_v9i1e85339_app2.docx]

Multimedia Appendix 3
Sensitivity analysis.
[DOCX File , 1155 KB - aging_v9i1e85339_app3.docx]

Multimedia Appendix 4
CONSORT checklist.
[PDF File (Adobe PDF File), 128 KB - aging_v9i1e85339_app4.pdf]

References

1. Beard JR, Officer A, de Carvalho IA, Sadana R, Pot AM, Michel JP, et al. The World report on ageing and health: a policy framework for healthy ageing. Lancet 2016 May 21;387(10033):2145-2154 [FREE Full text] [doi: 10.1016/S0140-6736(15)00516-4] [Medline: 26520231]
2. Dzau VJ, Inouye SK, Rowe JW, Finkelman E, Yamada T. Enabling healthful aging for all - the National Academy of Medicine grand challenge in healthy longevity. N Engl J Med 2019 Oct 31;381(18):1699-1701. [doi: 10.1056/NEJMp1912298] [Medline: 31633895]
3. Chen LK, Hwang AC, Lee WJ, Peng LN, Lin MH, Neil DL, et al. Efficacy of multidomain interventions to improve physical frailty, depression and cognition: data from cluster-randomized controlled trials. J Cachexia Sarcopenia Muscle 2020 Jun;11(3):650-662 [FREE Full text] [doi: 10.1002/jcsm.12534] [Medline: 32134208]
4. Lee WJ, Peng LN, Lin CH, Chen RC, Lin SZ, Loh CH, et al. Effects of incorporating multidomain interventions into integrated primary care on quality of life: a randomised controlled trial. Lancet Healthy Longev 2021 Nov;2(11):e712-e723 [FREE Full text] [doi: 10.1016/S2666-7568(21)00248-8] [Medline: 36098028]
5. Lee WJ, Peng LN, Lin MH, Loh CH, Chung CP, Wang PN, et al. Six-year transition of physio-cognitive decline syndrome: results from I-Lan Longitudinal Aging Study. Arch Gerontol Geriatr 2022;102:104743. [doi: 10.1016/j.archger.2022.104743] [Medline: 35687948]
6. Chung CP, Lee WJ, Peng LN, Shimada H, Tsai TF, Lin CP, et al. Physio-cognitive decline syndrome as the phenotype and treatment target of unhealthy aging. J Nutr Health Aging 2021;25(10):1179-1189 [FREE Full text] [doi: 10.1007/s12603-021-1693-4] [Medline: 34866145]
7. Razai MS, Oakeshott P, Kankam H, Galea S, Stokes-Lampard H. Mitigating the psychological effects of social isolation during the Covid-19 pandemic. BMJ 2020 May 21;369:m1904. [doi: 10.1136/bmj.m1904] [Medline: 32439691]
8. Santini ZI, Jose PE, York Cornwell E, Koyanagi A, Nielsen L, Hinrichsen C, et al. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. Lancet Public Health 2020 Jan;5(1):e62-e70 [FREE Full text] [doi: 10.1016/S2468-2667(19)30230-0] [Medline: 31910981]
9. Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. Lancet 2015 Jun 06;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5] [Medline: 25771249]
10. Valenzuela T, Okubo Y, Woodbury A, Lord SR, Delbaere K. Adherence to technology-based exercise programs in older adults: a systematic review. J Geriatr Phys Ther 2018;41(1):49-61. [doi: 10.1519/JPT.0000000000000095] [Medline: 27362526]
11. Zhao Y, Feng H, Wu X, Du Y, Yang X, Hu M, et al. Effectiveness of exergaming in improving cognitive and physical function in people with mild cognitive impairment or dementia: systematic review. JMIR Serious Games 2020 Jun 30;8(2):e16841 [FREE Full text] [doi: 10.2196/16841] [Medline: 32602841]
12. Chung SJ, An JY, Paik J, An SJ. Digital games for cognitive enhancement in healthy older adults: a scoping review. Games Health J 2026 Feb;15(1):1-13. [doi: 10.1177/2161783X251397928] [Medline: 41273126]
13. Ding M, Ouyang H, Zhang C, Jiang L, Zhuang R, Dong X, et al. Mahjong playing and leisure physical activity alleviate cognitive symptoms in older community residents. J Aging Phys Act 2022 Feb 01;30(1):89-97. [doi: 10.1123/japa.2020-0383] [Medline: 34388703]
14. Zhang H, Peng Y, Li C, Lan H, Xing G, Chen Z, et al. Playing mahjong for 12 weeks improved executive function in elderly people with mild cognitive impairment: a study of implications for TBI-induced cognitive deficits. Front Neurol 2020 Mar 27;11:178 [FREE Full text] [doi: 10.3389/fneur.2020.00178] [Medline: 32292385]
15. Cheng S, Chan AC, Yu EC. An exploratory study of the effect of mahjong on the cognitive functioning of persons with dementia. Int J Geriatr Psychiatry 2006 Jul;21(7):611-617. [doi: 10.1002/gps.1531] [Medline: 16779765]

XSL•FO
RenderX

16. Belchior P, Marsiske M, Sisco SM, Yam A, Bavelier D, Ball K, et al. Video game training to improve selective visual attention in older adults. Comput Human Behav 2013 Jul 01;29(4):1318-1324 [FREE Full text] [doi: 10.1016/j.chb.2013.01.034] [Medline: 24003265]
17. Nouchi R, Taki Y, Takeuchi H, Hashizume H, Nozawa T, Kambara T, et al. Brain training game boosts executive functions, working memory and processing speed in the young adults: a randomized controlled trial. PLoS One 2013;8(2):e55518 [FREE Full text] [doi: 10.1371/journal.pone.0055518] [Medline: 23405164]
18. Toril P, Reales JM, Mayas J, Ballesteros S. Video game training enhances visuospatial working memory and episodic memory in older adults. Front Hum Neurosci 2016 May 06;10:206 [FREE Full text] [doi: 10.3389/fnhum.2016.00206] [Medline: 27199723]
19. Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature 2013 Sep 05;501(7465):97-101 [FREE Full text] [doi: 10.1038/nature12486] [Medline: 24005416]
20. Hou HY, Jia XZ, Wang P, Zhang JX, Huang S, Li HJ. Intrinsic resting-state activity in older adults with video game experience. Front Aging Neurosci 2019 May 21;11:119 [FREE Full text] [doi: 10.3389/fnagi.2019.00119] [Medline: 31164816]
21. Farras-Permanyer L, Mancho-Fora N, Montalà-Flaquer M, Bartrés-Faz D, Vaqué-Alcázar L, Peró-Cebollero M, et al. Age-related changes in resting-state functional connectivity in older adults. Neural Regen Res 2019 Sep;14(9):1544-1555 [FREE Full text] [doi: 10.4103/1673-5374.255976] [Medline: 31089053]
22. Kawata NY, Nouchi R, Oba K, Matsuzaki Y, Kawashima R. Auditory cognitive training improves brain plasticity in healthy older adults: evidence from a randomized controlled trial. Front Aging Neurosci 2022 Mar 31;14:826672 [FREE Full text] [doi: 10.3389/fnagi.2022.826672] [Medline: 35431898]
23. Yin S, Xiong J, Zhu X, Li R, Li J. Cognitive training modified age-related brain changes in older adults with subjective memory decline. Aging Ment Health 2022 Oct;26(10):1997-2005. [doi: 10.1080/13607863.2021.1972931] [Medline: 34498987]
24. Ten Brinke LF, Davis JC, Barha CK, Liu-Ambrose T. Effects of computerized cognitive training on neuroimaging outcomes in older adults: a systematic review. BMC Geriatr 2017 Jul 10;17(1):139 [FREE Full text] [doi: 10.1186/s12877-017-0529-x] [Medline: 28693437]
25. Tomasi D, Wang GJ, Volkow ND. Energetic cost of brain functional connectivity. Proc Natl Acad Sci U S A 2013 Aug 13;110(33):13642-13647 [FREE Full text] [doi: 10.1073/pnas.1303346110] [Medline: 23898179]
26. Zang YF, He Y, Zhu CZ, Cao QJ, Sui MQ, Liang M, et al. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. Brain Dev 2007 Mar;29(2):83-91. [doi: 10.1016/j.braindev.2006.07.002] [Medline: 16919409]
27. Zang Y, Jiang T, Lu Y, He Y, Tian L. Regional homogeneity approach to fMRI data analysis. Neuroimage 2004 May;22(1):394-400. [doi: 10.1016/j.neuroimage.2003.12.030] [Medline: 15110032]
28. Cole MW, Pathak S, Schneider W. Identifying the brain's most globally connected regions. Neuroimage 2010 Feb 15;49(4):3132-3148. [doi: 10.1016/j.neuroimage.2009.11.001] [Medline: 19909818]
29. Coen RF, Robertson DA, Kenny RA, King-Kallimanis BL. Strengths and limitations of the MoCA for assessing cognitive functioning: findings from a large representative sample of Irish older adults. J Geriatr Psychiatry Neurol 2016 Jan 06;29(1):18-24. [doi: 10.1177/0891988715598236] [Medline: 26251108]
30. Kim H, Yu KH, Lee BC, Kim BC, Kang Y. Validity of the Montreal Cognitive Assessment (MoCA) index scores: a comparison with the cognitive domain scores of the Seoul Neuropsychological Screening Battery (SNSB). Dement Neurocogn Disord 2021 Jul;20(3):28-37 [FREE Full text] [doi: 10.12779/dnd.2021.20.3.28] [Medline: 34354755]
31. Liou YM, Jwo CJ, Yao KG, Chiang LC, Huang LH. Selection of appropriate Chinese terms to represent intensity and types of physical activity terms for use in the Taiwan version of IPAQ. J Nurs Res 2008 Dec;16(4):252-263. [doi: 10.1097/01.jnr.0000387313.20386.0a] [Medline: 19061172]
32. Zwolinsky S, McKenna J, Pringle A, Widdop P, Griffiths C. Physical activity assessment for public health: efficacious use of the single-item measure. Public Health 2015 Dec;129(12):1630-1636. [doi: 10.1016/j.puhe.2015.07.015] [Medline: 26296848]
33. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The Brief Resilience Scale: assessing the ability to bounce back. Int J Behav Med 2008;15(3):194-200. [doi: 10.1080/10705500802222972] [Medline: 18696313]
34. Hung HC, Chen HW, Chang YF, Yang YC, Liu CL, Hsieh RK, et al. Evaluation of the reliability and validity of the Mandarin version of Demoralization Scale for cancer patients [Article in Chinese]. J Intern Med Taiwan 2010;21(6):427-435. [doi: 10.6314/JIMT.2010.21(6).06]
35. Kissane DW, Wein S, Love A, Lee XQ, Kee PL, Clarke DM. The Demoralization Scale: a report of its development and preliminary validation. J Palliat Care 2004;20(4):269-276. [Medline: 15690829]
36. Mehnert A, Vehling S, Höcker A, Lehmann C, Koch U. Demoralization and depression in patients with advanced cancer: validation of the German version of the Demoralization Scale. J Pain Symptom Manage 2011 Nov;42(5):768-776 [FREE Full text] [doi: 10.1016/j.jpainsymman.2011.02.013] [Medline: 21592722]

XSL•FO
RenderX

37. Lin HW, Li CI, Lin FJ, Chang JY, Gau CS, Luo N, et al. Valuation of the EQ-5D-5L in Taiwan. PLoS One 2018 Dec 26;13(12):e0209344 [FREE Full text] [doi: 10.1371/journal.pone.0209344] [Medline: 30586400]
38. Schrag A, Selai C, Jahanshahi M, Quinn NP. The EQ-5D--a generic quality of life measure-is a useful instrument to measure quality of life in patients with Parkinson's disease. J Neurol Neurosurg Psychiatry 2000 Jul;69(1):67-73 [FREE Full text] [doi: 10.1136/jnnp.69.1.67] [Medline: 10864606]
39. Li BY, He NY, Qiao Y, Xu HM, Lu YZ, Cui PJ, et al. Computerized cognitive training for Chinese mild cognitive impairment patients: a neuropsychological and fMRI study. Neuroimage Clin 2019;22:101691 [FREE Full text] [doi: 10.1016/j.nicl.2019.101691] [Medline: 30708349]
40. Kim S, Park E, Cha H, Jung JC, Jung TD, Chang Y. Effects of cognitive training in mild cognitive impairment measured by resting state functional imaging. Behav Sci (Basel) 2020 Nov 17;10(11):175 [FREE Full text] [doi: 10.3390/bs10110175] [Medline: 33213021]
41. Fink A, Benedek M, Koschutnig K, Pirker E, Berger E, Meister S, et al. Training of verbal creativity modulates brain activity in regions associated with language- and memory-related demands. Hum Brain Mapp 2015 Oct;36(10):4104-4115 [FREE Full text] [doi: 10.1002/hbm.22901] [Medline: 26178653]
42. Bludau S, Eickhoff SB, Mohlberg H, Caspers S, Laird AR, Fox PT, et al. Cytoarchitecture, probability maps and functions of the human frontal pole. Neuroimage 2014 Jun;93 Pt 2(Pt 2):260-275 [FREE Full text] [doi: 10.1016/j.neuroimage.2013.05.052] [Medline: 23702412]
43. Brill E, Holfelder A, Falkner M, Krebs C, Brem AK, Klöppel S. Behavioural and neuronal substrates of serious game-based computerised cognitive training in cognitive decline: randomised controlled trial. BJPsych Open 2024 Nov 06;10(6):e200 [FREE Full text] [doi: 10.1192/bjo.2024.797] [Medline: 39501844]
44. Feter N, Caputo EL, Koth AP, Delpino FM, Tomaz GD, Doring IR, et al. Association between specific types of physical activity during the COVID-19 pandemic and the risk of subjective memory decline: findings from the PAMPA Cohort. Public Health 2021 Nov;200:49-55 [FREE Full text] [doi: 10.1016/j.puhe.2021.09.014] [Medline: 34673460]
45. Li C, Hua R, Gao D, Zheng F, Xie W. Cognitive decline before and during COVID-19 pandemic among older people with multimorbidity: a longitudinal study. J Am Med Dir Assoc 2023 Apr;24(4):419-25.e10 [FREE Full text] [doi: 10.1016/j.jamda.2023.01.001] [Medline: 36774966]
46. Wang R, Chen Z, Zhou Y, Shen L, Zhang Z, Wu X. Melancholy or mahjong? Diversity, frequency, type, and rural-urban divide of social participation and depression in middle- and old-aged Chinese: a fixed-effects analysis. Soc Sci Med 2019 Oct;238:112518. [doi: 10.1016/j.socscimed.2019.112518] [Medline: 31473574]
47. Hanson S, Jones A. Is there evidence that walking groups have health benefits? A systematic review and meta-analysis. Br J Sports Med 2015 Jun;49(11):710-715 [FREE Full text] [doi: 10.1136/bjsports-2014-094157] [Medline: 25601182]
48. Murtagh EM, Murphy MH, Boone-Heinonen J. Walking: the first steps in cardiovascular disease prevention. Curr Opin Cardiol 2010 Sep;25(5):490-496 [FREE Full text] [doi: 10.1097/HCO.0b013e32833ce972] [Medline: 20625280]
49. Draganski B, Gaser C, Busch V, Schuierer G, Bogdahn U, May A. Neuroplasticity: changes in grey matter induced by training. Nature 2004 Jan 22;427(6972):311-312. [doi: 10.1038/427311a] [Medline: 14737157]
50. Bang JW, Shibata K, Frank SM, Walsh EG, Greenlee MW, Watanabe T, et al. Consolidation and reconsolidation share behavioral and neurochemical mechanisms. Nat Hum Behav 2018 Jul;2(7):507-513 [FREE Full text] [doi: 10.1038/s41562-018-0366-8] [Medline: 30505952]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
fALFF: fractional amplitude of low-frequency fluctuations
fMRI: functional magnetic resonance imaging
FWE: family-wise error
GBC: global brain connectivity
GMV: gray matter volume
IPAQ: International Physical Activity Questionnaire
MET: metabolic equivalent of task
MoCA: Montreal Cognitive Assessment
MRI: magnetic resonance imaging
ReHo: regional homogeneity

XSL•FO
RenderX

Edited by Q Fan; submitted 06.Oct.2025; peer-reviewed by E Brill, O Shen; comments to author 15.Dec.2025; revised version received 25.Jun.2026; accepted 26.Jun.2026; published 23.Jul.2026.

Please cite as:
Tung HH, Kuo CY, Lee PL, Chang CW, Chou KH, Chen KE, Lin CP, Liang CK, Chen LK
Digital Mahjong-Based Cognitive Engagement Intervention for Older Adults: Randomized Controlled Trial
JMIR Aging 2026;9:e85339
URL: https://aging.jmir.org/2026/1/e85339
doi:10.2196/85339
PMID:

©Heng-Hsin Tung, Chen-Yuan Kuo, Pei-Lin Lee, Chih-Wen Chang, Kun-Hsien Chou, Katelyn E Chen, Ching-Po Lin, Chih-Kuang Liang, Liang-Kung Chen. Originally published in JMIR Aging (https://aging.jmir.org), 23.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Impact of Chatbot-Based Cognitive Training on Cognitive Function and White Matter Integrity in Cognitively Unimpaired Older Adults: Randomized Controlled Trial

Bori R Kim[1,2], PhD; Eunhee Kim[3], MD, PhD; Soo Mee Lim[4], MD, PhD; Jinwoo Kim[5,6], PhD; Jee Hyang Jeong[4], MD, PhD; Geon Ha Kim[1], MD, PhD

[1]Department of Neurology, College of Medicine, Ewha Womans University Mokdong Hospital, Ewha Womans University, Seoul, Republic of Korea
[2]Ewha Medical Research Institute, Ewha Womans University, Seoul, Republic of Korea
[3]Department of Radiology, College of Medicine, Ewha Womans University Mokdong Hospital, Ewha Womans University, Seoul, Republic of Korea
[4]Department of Neurology, College of Medicine, Ewha Womans University Seoul Hospital, Ewha Womans University, Seoul, Republic of Korea
[5]Department of Digital Therapeutics Research and Development, HAII Corp., Seoul, Republic of Korea
[6]Graduate School of Cognitive Science, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
Geon Ha Kim, MD, PhD
Department of Neurology
College of Medicine, Ewha Womans University Mokdong Hospital
Ewha Womans University
1071, Anyangcheon-ro, Yangcheon-gu
Seoul, 07985
Republic of Korea
Phone: 82 0226505847
Email: geonha@ewha.ac.kr

Abstract

Background: Cognitive training may enhance cognition and support brain health during aging, but few trials have evaluated readily accessible chatbot-based programs with neuroimaging outcomes.

Objective: This study aimed to evaluate the effects of a chatbot-based cognitive training program on white matter microstructural organization and cognition in cognitively unimpaired older adults.

Methods: In a single-blind randomized controlled trial, 72 cognitively unimpaired older adults were assigned to chatbot-based cognitive training (30 min/day, 5 days/week for 12 weeks; n=37) or a no-intervention control (n=35). Diffusion tensor imaging at baseline and 12 weeks quantified fractional anisotropy (FA) in predefined white matter tracts, and cognition was assessed with the Cambridge Neuropsychological Test Automated Battery.

Results: Groups were comparable at baseline. After 12 weeks, the training group showed significantly less FA reduction than controls in the bilateral tapetum and right posterior thalamic radiation, surviving false discovery rate control (P<.05). The training group also improved on Paired Associates Learning (PAL), an episodic memory measure (P=.04). PAL improvement was positively associated with FA change in the right posterior thalamic radiation and right sagittal stratum (false discovery rate–adjusted P=.04 for each).

Conclusions: These findings suggest that a 12-week chatbot-based cognitive training program may support episodic memory and the relative maintenance of white matter microstructural organization in selected pathways in cognitively unimpaired older adults, consistent with a brain-maintenance framework.

Trial Registration: CRIS KCT0004630; https://tinyurl.com/4xfesb68

(JMIR Aging 2026;9:e94127) doi:10.2196/94127

KEYWORDS

aging; chatbot; cognitive training; diffusion tensor imaging; randomized controlled trial; white matter

XSL•FO
RenderX

Introduction

Extensive research indicates that cognitive training can mitigate age-related cognitive decline, with improvements reported across multiple domains including memory, attention, and executive functions [1-3]. Large-scale randomized controlled trials, such as the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) trial, have demonstrated domain-specific gains in memory, processing speed, and reasoning, with benefits extending to everyday functioning over long-term follow-up [4,5]. However, many traditional interventions rely on in-person group delivery, which may be less accessible for older adults due to mobility or transportation barriers, scheduling constraints, or reluctance to participate in group-based settings. In addition, such programs often require trained facilitators, limiting scalability across communities [6]. These limitations underscore the need for accessible, scalable, home-based cognitive interventions, particularly when in-person delivery is constrained [7].

Computerized (digital) cognitive training offers a scalable alternative that can be delivered at home with fewer logistical barriers [8]. Particularly, chatbot-based cognitive training represents an evolution of mobile-delivered computerized cognitive training. By integrating conversational AI, chatbots simulate human-like interactions that provide social-cognitive stimulation, which is often lacking in passive digital platforms. This interactive nature may improve engagement and adherence—key challenges in home-based interventions for older adults—by reducing perceived technological burden and providing interactive guidance and feedback [9,10]. This is especially important for older adults living alone, where home-based assistive technologies can maintain accessibility without spatial or scheduling limitations [11,12]. Early studies suggest that chatbot interfaces reduce technological burden while supporting key adherence factors such as usability and consistent feedback [8,13].

Building on prior evidence, we developed Saemi, a chatbot-based cognitive training program delivered through KakaoTalk, a mobile messaging platform widely used in Korea [14,15]. Saemi provides structured cognitive training tasks targeting memory, attention, calculation, language, and executive function. Each daily session includes at least 2 different training modules: one consistently focused on memory and the other targeting another cognitive domain. This design emphasizes memory training as a recurring component, which is consistent with the evidence that memory decline is a key early feature of age-related cognitive changes [16-18]. By leveraging the interactive nature of chatbots, the program enhances personalization and user engagement, while its home-based format helps overcome spatial and logistical barriers, ultimately promoting sustained cognitive stimulation through accessible mobile technology.

While chatbot-based interventions represent a promising and scalable approach, their cognitive benefits and potential neurobiological correlates remain insufficiently characterized in cognitively unimpaired (CU) older adults. Relative maintenance of brain structural integrity—particularly in white matter (WM)—is considered a key mechanism supporting cognitive resilience during aging and is central to the brain maintenance framework [19-21]. WM integrity, particularly in tracts facilitating interhemispheric communication and sensory-cognitive integration, is essential for efficient neural communication and is strongly associated with memory, attention, and executive function [22-24]. Such WM microstructural changes may be sensitive to aging and experience, and may reflect plasticity-related processes over time [25-27].

Therefore, this study aimed to determine whether a 12-week chatbot-based cognitive training program could attenuate age-related decline in WM integrity and concurrently enhance cognitive performance, consistent with a brain maintenance framework. By addressing both neural mechanisms and cognitive outcomes, this work sought to evaluate the potential of a scalable, home-based intervention to support brain maintenance and cognitive resilience in aging.

Methods

Study Design

This was a prospective, single-center, rater-blind, randomized controlled trial, which was conducted from November 21, 2019, to July 13, 2020, at the Memory Disorder Clinic of Ewha Womans University Mokdong Hospital in Seoul, Republic of Korea. This study was reported in accordance with the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist (Multimedia Appendix 1). Patients and the public were not involved in the design, conduct, reporting, or dissemination plans of this trial.

Participants

A total of 81 CU community-dwelling individuals were recruited from the Ewha Womans University Mokdong Hospital in Seoul. All participants underwent a comprehensive neuropsychological assessment using the Seoul Neuropsychological Screening Battery, second edition (SNSB-II) [28] by trained clinical neuropsychologists. CU was defined as those aged 60 or above who scored more than –1 SD below the mean of age- and education-matched norms in tests evaluating memory, attention, language, visuospatial, and frontal executive functions in the SNSB-II. We excluded individuals with any of the following characteristics: (1) suspected or confirmed mild cognitive impairment or dementia; (2) suspected or confirmed major neurological or psychiatric illnesses, including major depressive disorders; (3) contraindications to MRI; (4) visual or hearing impairments severe enough to interfere with questionnaire response; (5) history of medications that could affect cognitive and emotional functions in the past 3 months; and (6) other major medical problems. Of the 81 participants, nine were excluded because of the withdrawal of consent before the baseline assessment. Therefore, a total of 72 participants were included in this study.

Sample Size

For the formal sample size calculation, we used the cortical thickness effect-size estimate specified in the study protocol:

XSL•FO
RenderX

an expected longitudinal cortical thickness change of 0.046 with an SD of 0.053, corresponding to a standardized effect size of Cohen *d*=0.87 [29]. The sample size calculation was based on a 2-tailed independent-samples *t* test comparing mean longitudinal change between the 2 groups. Assuming 2 parallel groups, a 2-tailed α of .05, and 90% statistical power to detect group differences, the required sample size was estimated to be 30 participants per group. To accommodate an anticipated dropout rate of approximately 20%—commonly observed in MRI-based intervention studies—we planned to enroll 36 participants per group, yielding a total target sample of 72 participants.

Ethical Considerations

Written informed consent was obtained from all participants. The study protocol was approved by the Institutional Review Board (IRB) of Ewha Womans University Mokdong Hospital (IRB number 2019-06-028). The study was conducted in accordance with the Declaration of Helsinki and relevant guidelines and regulations. Participant data were handled confidentially and deidentified for analysis in accordance with applicable institutional policies. Participants did not receive financial compensation for participation. This trial was registered with the Clinical Research Information Service (CRIS), Republic of Korea (KCT0004630).

Randomization

After the baseline assessment, 72 participants were randomly allocated in a 1:1 ratio to 2 groups: the intervention group (n=37) and the control group (n=35). A simple, computer-generated randomization sequence was created by an independent statistician who had no contact with study participants. To ensure allocation concealment, the sequence was stored in a password-protected file and revealed only after the baseline assessment was finalized. Trained research staff enrolled participants and had no access to the allocation sequence. Outcome assessors were blinded to group assignments throughout the study.

Blindness

All outcome evaluators were blinded to the group assignment of participants and were not involved in the intervention.

Intervention

All participants from both groups received 30 minutes of education about dementia prevention after undergoing the baseline assessment. Participants in the control group received education only and were instructed to continue their usual activities for 12 weeks without any additional cognitive training. This education included the definition, diagnosis, and current treatments of dementia and lifestyle modifications for dementia prevention, encompassing the importance of cognitive training, healthy foods for the brain, aerobic exercise, and stretching that could be done at home [30,31]. Beyond the study intervention, no other cointerventions were provided. To avoid confounding effects, participants continued their routine health care and agreed not to participate in any other cognitive enhancement programs throughout the trial.

All participants in the intervention group used Saemi (HAII Corporation), a chatbot-based cognitive training program delivered via a mobile messaging platform. The program provides representative brain-training tasks grouped by cognitive domains—language, memory, attention, calculation, and executive function—as illustrated in Figure 1. Participants were instructed to complete training for at least 30 minutes per day, 5 days per week, for 12 weeks. Each daily session comprised at least 2 tasks, including one consistently focused on memory and another targeting a different domain (language, visuospatial function, calculation, or executive function). Actual intervention exposure was quantified using app usage logs. Usage metrics included total training time, completed session counts, active days, and mean use time per active day. Time-based adherence rate was calculated as the percentage of the recommended minimum total training time completed according to app usage logs. The prescribed minimum dose was 30 minutes per day on 5 days per week for 12 weeks, corresponding to 30 total hours across the intervention period. Adverse events were monitored throughout the trial via participant self-reports and spontaneous reporting.

The figure shows example tasks targeting language, memory, attention, calculation, and executive function in a chat-based mobile format.

XSL•FO
RenderX

Figure 1. Representative interfaces of the Saemi chatbot-based cognitive training program across cognitive domains.



Outcome Measures

Changes in WM Microstructure

WM microstructural changes were assessed using diffusion tensor imaging (DTI). Fractional anisotropy (FA) was the principal WM metric due to its sensitivity to microstructural organization and its association with cognitive performance in aging [22,32]. Longitudinal FA changes were evaluated using predefined tract-level region-of-interest (ROI) measures extracted from tract-based spatial statistics (TBSS)–processed skeletonized diffusion maps, rather than as a single global FA average; therefore, false discovery rate (FDR) correction was applied across the predefined ROIs. Global baseline FA was used only to assess baseline comparability between groups. Additional metrics such as mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD) were analyzed as complementary measures to provide a broader characterization of WM changes. In line with conventional neuroimaging approaches in intervention studies, WM changes were evaluated as longitudinal alterations over the intervention period, with FA used as the principal imaging marker.

Cognitive Outcomes

Episodic memory performance was assessed using the Paired Associates Learning (PAL) task of the CANTAB [33,34]. PAL was selected because memory training was the core component of the intervention, and PAL is highly sensitive to early episodic memory decline and intervention-related improvements in CU older adults [35]. Raw scores were standardized to *z* scores based on the baseline mean and SD of the control group, with positive *z* scores reflecting better performance. PAL was the principal cognitive measure of episodic memory, reflecting its established validity as a sensitive measure of associative episodic memory.

In addition to PAL, 5 further tasks were included to assess additional cognitive domains: Delayed Matching to Sample (DMS), Rapid Visual Information Processing (RVP), Reaction Time (RTI), Spatial Working Memory (SWM), and One Touch Stockings of Cambridge (OTS). Raw scores were standardized to *z* scores as for PAL, with positive *z* scores representing better performance [36]. The models for cognitive outcomes were adjusted for age, sex, years of education, and baseline CANTAB scores; however, the reported *P* values were not corrected for multiple comparisons across cognitive outcomes.

MRI Acquisition and Processing

MRI Acquisition

All participants underwent brain MRI on the same scanner (Achieva, Philips 3.0T) using a standardized imaging protocol. The 3D T1 turbo field echo magnetic resonance images were acquired with the following imaging parameters: sagittal slice thickness=0.5 mm; no gap; repetition time (TR)=9.9 ms; echo time (TE)=4.6 ms; flip angle=8°; and a matrix size of 240×240 pixels. Fluid-attenuated inversion recovery magnetic resonance images were used to evaluate abnormal WM hyperintensities. These were assessed using the following imaging parameters: axial slice thickness=2 mm; no gap; TR=11000 ms; TE=125 ms; inversion time=2800 ms; matrix size=256×189 pixels. For TBSS, DTI were acquired using a diffusion-weighted single-shot echo-planar imaging sequence with the following parameters: TR=9.2-9.9 s; TE=91 ms; flip angle=90°; b-value=1000 s/mm^2; matrix size=128×126 pixels; 70 axial sections; slice thickness=2 mm; in-plane resolution=1.72×1.72 mm. With the baseline image without diffusion weighting (the reference volume), diffusion-weighted images were acquired from 32 different directions.

XSL•FO
RenderX

DTI Processing

DTI data were processed using the Functional Magnetic Resonance Imaging of the Brain (FMRIB) Software Library (FSL, v6.0.2). Because the JHU ICBM-DTI-81 white-matter labels atlas has been reported to contain left-right orientation and label inconsistencies in some distributed versions [37], we performed an explicit left-right orientation check after registration to standard space and verified atlas orientation before ROI masking and extraction [38,39]. Eddy current–induced distortions and head motion were corrected by affine registration of each diffusion-weighted volume to the non–diffusion-weighted (b0) image using eddy_correct (FSL), which relies on the FMRIB's Linear Image Registration Tool (FLIRT) for linear registration [40,41]. The b vectors were then rotated appropriately using the "fdt-rotate-bvecs" tool as part of FSL [38,42]. Diffusion tensors were then calculated at the level of an individual voxel to generate FA images using DTIfit, and the images of brain tissue were subsequently extracted using the Brain Extraction Tool [43] from the FSL.

The baseline and follow-up FA data of all subjects were realigned onto the standard FMRIB58_FA template provided by the FSL software using a nonlinear registration algorithm implemented in the TBSS package. To improve anatomical longitudinal alignment, we used FLIRT tools from FSL [40,41]. The registered FA images were averaged to generate a skeletonized mean FA image, which represents the main fiber tracts and the center of all tracts common to the group. A threshold of 0.2 was applied to exclude peripheral tracts where there was significant intersubject variability and/or partial volume effect due to gray matter. Each participant's aligned FA data was projected onto the mean skeleton to create a skeletonized FA map. A similar process was performed for the MD, AD, and RD. The skeletonized diffusion maps were subsequently used for atlas-based ROI extraction and ROI-level statistical analyses.

ROI Selection

Following TBSS-based preprocessing and skeletonization, ROI analyses were performed to obtain anatomically interpretable tract-level diffusion measures. ROIs were defined using the JHU ICBM-DTI-81 white-matter labels atlas (48 tracts) [44,45]. The final 17 ROIs were selected based on prior neuroanatomical and cognitive-aging literature, rather than on Shapley additive explanations (SHAP) results. As a supplementary exploratory analysis, SHAP feature-importance ranking across all 48 atlas-defined regions was examined in relation to PAL change (Figure S2 in Multimedia Appendix 2) and was not used to define the ROI set, to support confirmatory inference, or as a criterion for statistical significance [46]. The selected ROIs included the splenium of the corpus callosum, bilateral tapetum, bilateral anterior corona radiata (ACR), bilateral posterior thalamic radiation (PTR; including optic radiation), bilateral sagittal stratum (SS), bilateral cingulum (cingulate gyrus and hippocampal segments), bilateral superior longitudinal fasciculus (SLF), and bilateral uncinate fasciculus. The 17 selected ROIs and their corresponding JHU ICBM-DTI-81 atlas label numbers are provided in Table S3 in Multimedia Appendix 2. ROI-level diffusion metrics were extracted for each selected tract and used as longitudinal outcome measures.

Statistical Analyses

Statistical analyses were performed using STATA software (version 18; StataCorp). Baseline demographic and clinical characteristics were compared between groups using the chi-square test for categorical variables, independent-samples *t* tests for continuous variables, and the Mann-Whitney *U* test for nonnormally distributed variables. Baseline neuropsychological test performances were compared using analysis of covariance adjusted for age, sex, and years of education.

For cognitive outcomes, raw scores from each task were transformed into standardized *z* scores using the mean and SD of the control group at baseline; where necessary, scores were oriented such that higher *z* scores indicated better performance [47].

The main analytical approach for longitudinal changes in DTI metrics and CANTAB outcomes was a linear mixed-effects model, including fixed effects for group, visit (baseline vs follow-up), and the group-by-visit interaction, with a participant-specific random intercept to account for within-subject correlations. Models were adjusted for age, sex, years of education, and the baseline value of each outcome. Missing outcome data were handled implicitly using maximum likelihood estimation in the mixed-effects framework; inference relies on the standard missing-at-random assumption [48].

The main analysis followed the intention-to-treat (ITT) principle, including all randomized participants in the groups to which they were assigned, regardless of adherence. In addition, a supplementary per-protocol (PP) analysis included only those participants who completed the intervention and the follow-up assessment [49].

ROI diffusion metrics were analyzed using the same mixed-effects framework. To account for multiple comparisons across ROIs, FDR correction was applied, with Bonferroni-adjusted and uncorrected *P* values reported for descriptive purposes.

Associations between changes in WM microstructure (DTI metrics) and cognitive performance (CANTAB *z* scores) were examined using partial correlation analyses, adjusting for age, sex, years of education, and baseline imaging and cognitive scores. For these exploratory structure-function analyses, missing values were handled using multiple imputation by chained equations with predictive mean matching (m=20 imputations; 5 iterations) under a missing-at-random assumption, implemented in R (version 4.4.1; *mice* package) [50]. Imputed datasets were analyzed separately and combined using standard multiple-imputation pooling rules [51]. Imputation diagnostics were evaluated using density plots. *P* values were adjusted for multiple comparisons using the Benjamini-Hochberg FDR procedure.

Protocol Deviations

This trial was prospectively registered with the CRIS of the Republic of Korea (KCT0004630), with cortical thickness and

XSL•FO
RenderX

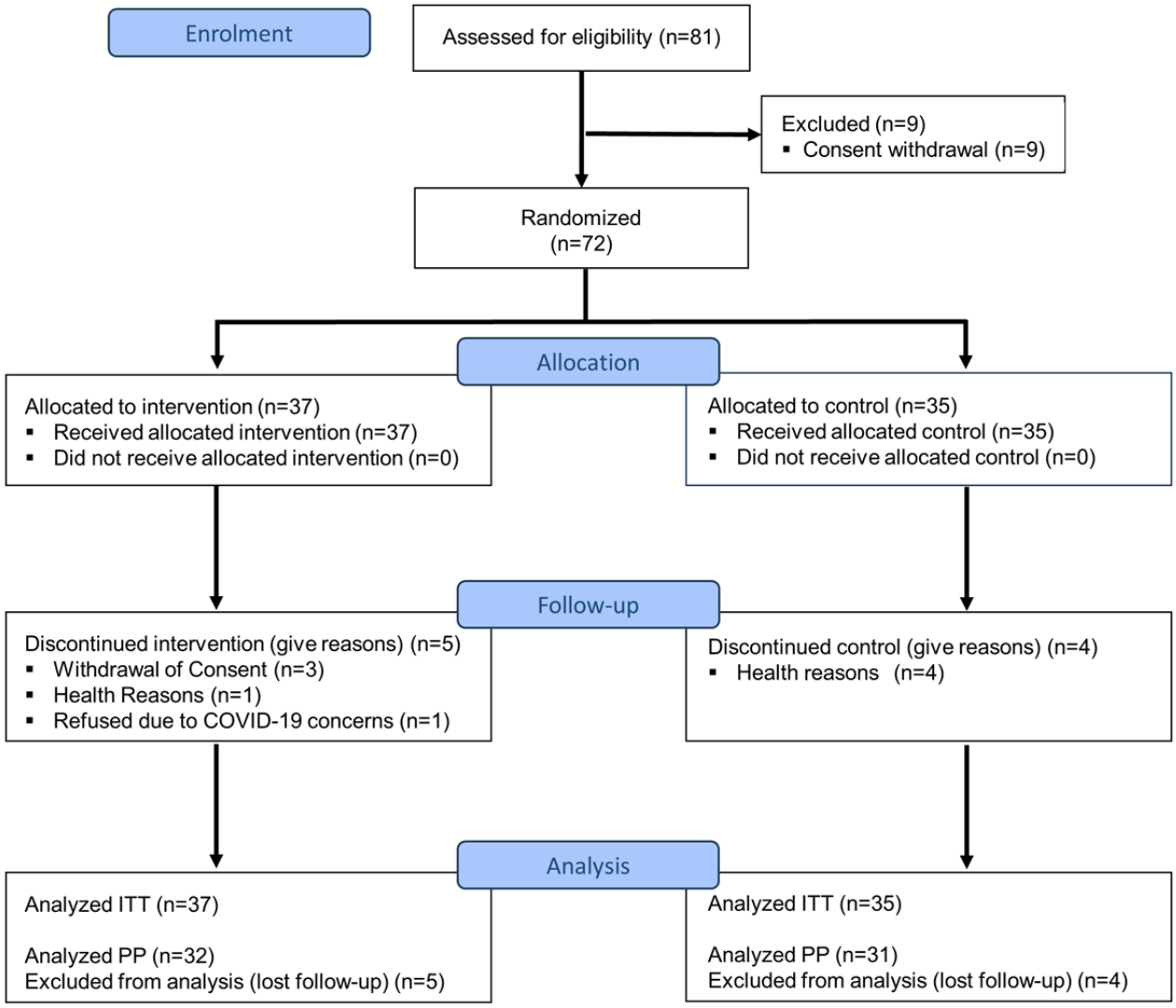
resting-state functional magnetic resonance imaging (fMRI) specified as the registered primary imaging outcomes. These modalities were acquired as planned; however, modality-specific longitudinal analyses of cortical thickness and resting-state fMRI have not yet been completed and are not included in the present report. DTI was included from the outset in the IRB-approved neuroimaging protocol but was not separately listed as an outcome in the CRIS registration. Because DTI processing and longitudinal analyses had been completed for the cohort, this study focuses on longitudinal changes in DTI-derived WM microstructure together with cognitive outcomes and should be interpreted as a focused analysis rather than as a report of the registered primary imaging outcomes. Separately, the registered protocol described a 60-minute daily cognitive training schedule, whereas the implemented intervention instructed participants to complete at least 30 minutes per day, 5 days per week for 12 weeks. This change was made before participant enrollment to reduce participant burden and improve the feasibility of sustained home-based training in older adults.

Results

Baseline Characteristics

Of the 72 participants who underwent baseline assessment, 9 (12.5%) participants withdrew from the study: 5 (13.5%) from the intervention group and 4 (11.4%) from the control group. Consequently, 63 (87.5%) participants completed the final assessment (Figure 2). No intervention-related adverse events were reported in either group. Baseline demographic characteristics, clinical variables, neuroimaging characteristics, and cognitive performance across attention, visuospatial function, memory, language, and frontal executive domains did not differ significantly between the intervention and control groups (Table 1).

Figure 2. Randomized trial flow diagram. A total of 72 cognitively unimpaired older adults aged 60 years or older were randomized into the intervention (n=37) and control (n=35) groups. During the trial, 5 participants from the intervention group and 4 from the control group withdrew from the study, and 63 participants completed the 12-week follow-up assessment. Baseline and follow-up assessments included Cambridge Neuropsychological Test Automated Battery cognitive testing and brain magnetic resonance imaging acquisition. ITT: intention-to-treat; PP: per-protocol.



XSL•FO
RenderX

Table 1. Demographic and clinical characteristics of participants.[a]

Characteristic	Intervention (n=37)	Control (n=35)	*P* value
Demographic and clinical characteristics			
Age (years), mean (SD)	73.49 (5.48)	73.60 (5.89)	.93
Education (years)[b], mean (SD)	9.20 (4.76)	9.84 (3.93)	.53
Female, n (%)	31 (84)	31 (89)	.56
K-MMSE[c], mean (SD)[b]	27.76 (2.24)	27.89 (1.78)	.90
Underlying disease, n (%)			
Diabetes mellitus[b]	10 (27)	6 (17)	.32
Hypertension	19 (51)	18 (51)	.99
Hyperlipidemia	17 (46)	13 (37)	.46
Dropout, n (%)[b]	5 (14)	4 (11)	.79
Neuroimaging characteristics, mean (SD)			
Intracranial volume (cm^3)	1448.76 (124.66)	1464.87 (101.53)	.55
Whole brain volume (% of ICV[d])	69.61 (2.55)	69.62 (2.81)	>.99
Global FA[e]	0.47 (0.02)	0.47 (0.02)	.22
Baseline cognitive performance (*z* scores)[f], mean (SD)			
Attention	0.03 (0.92)	0.13 (0.99)	.76
Visuospatial	–0.22 (0.88)	–0.02 (0.76)	.08
Memory	0.63 (0.82)	0.9 (0.94)	.21
Language	0.54 (0.79)	0.32 (0.8)	.07
Frontal executive	0.48 (0.51)	0.57 (0.6)	.94

[a]Data are presented as mean (SD) or number (%) for dichotomous variables. Between-group comparisons were performed using independent *t* tests unless otherwise indicated.

[b]*P* values were calculated using the Mann-Whitney *U* test for variables with nonnormal distributions.

[c]K-MMSE: Korean version of the Mini-Mental State Examination.

[d]ICV: intracranial volume.

[e]FA: fractional anisotropy.

[f]*P* values for baseline cognitive performance (*z* scores) were obtained using analysis of covariance, adjusted for age, sex, and years of education.

Changes in WM Microstructure

At week 12, the intervention group showed significantly less reduction in FA than the control group in the right tapetum, left tapetum, and right PTR after FDR correction (*P*=.03; *P*=.03; and *P*=.04, respectively; Table 2; Figure 3). Detailed longitudinal values for each group are provided in Table S1, and within-group pre-post comparisons based on paired *t* tests are provided in Table S4 in Multimedia Appendix 2.

Regarding MD, a nominal group difference was observed in the left cingulum (cingulate gyrus; *P*=.04), but this finding did not remain significant following correction for multiple comparisons (FDR-adjusted *P*=.37; Table S5 in Multimedia Appendix 2).

In the PP analyses (participants who completed the follow-up assessment), the results were consistent with the ITT findings. Participants who adhered to the intervention protocol demonstrated less reduction in FA compared with controls at week 12 in the right tapetum (–0.03 for intervention vs –0.06 for control group; FDR-adjusted *P*=.04) and right PTR (–0.02 vs –0.03; FDR-adjusted *P*=.04; Table S6 in Multimedia Appendix 2).

XSL•FO
RenderX

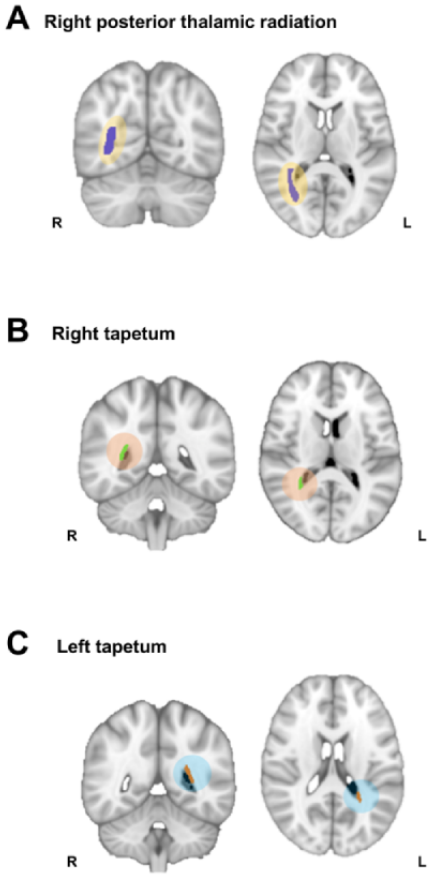
Table 2. Changes in white matter (WM) microstructure (fractional anisotropy) measured using diffusion tensor imaging in the intention-to-treat analysis after the 12-week intervention. False discovery rate–adjusted *P* values (q) <.05 were considered statistically significant.[a]

WM microstructure	Estimated between-group difference (95% CI)	*P* value[b]	*q* [c]
Commissural fibers			
SCC[d]	0.010 (–0.001 to 0.020)	.07	.20
Tapetum (R[e])	0.025 (0.008 to 0.043)	.004	.03
Tapetum (L[f])	0.023 (0.008 to 0.039)	.003	.03
Projection fibers			
ACR[g] (R)	0.003 (–0.011 to 0.017)	.66	.75
ACR (L)	0.001 (–0.008 to 0.010)	.82	.83
PTR[h] (R)	0.013 (0.003 to 0.022)	.007	.04
PTR (L)	0.008 (–0.001 to 0.016)	.07	.20
Association fibers			
SS[i] (R)	0.012 (–0.001 to 0.025)	.07	.20
SS (L)	0.005 (–0.004 to 0.013)	.27	.46
Cingulum-CG[j] (R)	0.003 (–0.009 to 0.015)	.67	.75
Cingulum-CG (L)	0.009 (–0.004 to 0.022)	.20	.37
Cingulum-HPC[k] (R)	0.013 (–0.017 to 0.042)	.40	.56
Cingulum-HPC (L)	0.018 (–0.009 to 0.044)	.19	.37
SLF[l] (R)	0.004 (–0.004 to 0.012)	.31	.47
SLF (L)	0.004 (–0.002 to 0.011)	.18	.37
UF[m] (R)	0.004 (–0.012 to 0.020)	.60	.75
UF (L)	0.002 (–0.014 to 0.017)	.83	.83

[a]Fractional anisotropy values are adjusted for age, sex, years of education, and baseline diffusion tensor imaging measures.
[b]Uncorrected *P* value from linear mixed models.
[c]False discovery rate–adjusted *P* value using the Benjamini-Hochberg method.
[d]SCC: splenium of the corpus callosum.
[e]R: right.
[f]L: left.
[g]ACR: anterior corona radiata.
[h]PTR: posterior thalamic radiation.
[i]SS: sagittal stratum.
[j]CG: cingulum (cingulate gyrus).
[k]HPC: cingulum (hippocampus).
[l]SLF: superior longitudinal fasciculus.
[m]UF: uncinate fasciculus.

XSL•FO
RenderX

Figure 3. Clusters showing between-group differences in longitudinal changes in fractional anisotropy. Projection and commissural fibers are shown: (A) right posterior thalamic radiation, (B) right tapetum, and (C) left tapetum. Colored regions indicate atlas-defined region of interest locations used for region of interest–based analyses.



Cognitive Outcomes

Episodic Memory (PAL)

In the ITT analysis, a linear mixed-effects model showed a significant between-group difference for PAL (estimated between-group difference 0.375, 95% CI 0.017 to 0.733; *P*=.04; Table 3 and Figure 4). Detailed longitudinal values for each group, including baseline, week 12, and within-group mean changes, are provided in Table S2 in Multimedia Appendix 2. A similar pattern for PAL was observed in the PP analysis (*P*=.049; Table S7 in Multimedia Appendix 2).

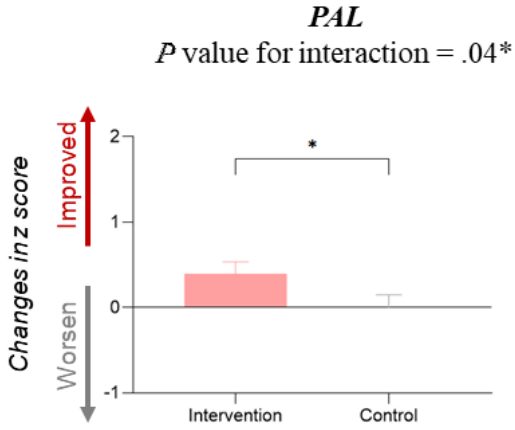
Table 3. Changes in cognitive outcomes measured by z scores after the 12-week intervention.[a]

Cognitive performance	Estimated between-group difference (95% CI)	*P* value
Memory		
Delayed Matching to Sample	0.164 (–0.421 to 0.748)	.58
Paired Associates Learning	0.375 (0.017 to 0.733)	.04
Attention		
Rapid Visual Information Processing	–0.723 (–3.006 to 1.561)	.54
Processing speed		
Reaction Time	–0.000 (–0.493 to 0.493)	>.99
Executive function		
Spatial Working Memory	–0.223 (–0.697 to 0.251)	.36
One Touch Stockings of Cambridge	0.014 (–0.464 to 0.492)	.96

[a]Adjusted for age, sex, years of education, and baseline Cambridge Neuropsychological Test Automated Battery scores. For all listed tasks, scores were transformed into *z* scores and, where necessary, oriented so that higher values indicate better cognitive performance. *P* values were not corrected for multiple comparisons across cognitive outcomes. Paired Associates Learning was the main cognitive measure of episodic memory; delayed matching to sample, Rapid Visual Information Processing, Reaction Time, Spatial Working Memory, and One Touch Stockings of Cambridge were additional cognitive outcomes.

XSL•FO
RenderX

Figure 4. Effects of the intervention on cognitive outcomes. Models were adjusted for age, sex, years of education, and baseline Cambridge Neuropsychological Test Automated Battery scores. PAL: Paired Associates Learning.



Other Cognitive Functions

No significant group × time interactions were observed for other cognitive domains including DMS (*P*=.58), RVP (*P*=.54), RTI (*P*>.99), SWM (*P*=.36), or OTS (*P*=.96; Table 3 and Figure S1 in Multimedia Appendix 2). These null results were consistent across both ITT and PP analyses. Thus, cognitive improvement was observed for PAL, whereas no detectable intervention-related changes were found in the other cognitive outcomes over the 12-week intervention period.

Structure-Function Correlations

In the ITT analysis, significant positive associations were observed between changes in FA and improvement in PAL performance in the right PTR and the right SS, and both remained significant after FDR correction (*P*=.04 for each; Table 4 and Figure 5). No other ROI demonstrated significant associations after correction for multiple comparisons. For MD, a nominal correlation was identified in the left ACR in the ITT analysis (*P*=.01), but this association did not survive FDR correction (Table S8 in Multimedia Appendix 2). In the PP analysis, nominal correlations were observed for FA in the right PTR (*P*=.03), right SS (*P*=.03), and right SLF (*P*=.03), and for MD in the right ACR (*P*=.02) and left ACR (*P*=.004); however, no additional ROI remained significant after multiple-comparison correction (Table S9 in Multimedia Appendix 2).

XSL•FO
RenderX

Table 4. Partial correlations between changes in Paired Associates Learning (PAL) performance and longitudinal changes in white matter microstructure (fractional anisotropy).[a]

Anatomical region	PAL			
	r	*P* value[b]	*P* value[c]	*q* [d]
Commissural				
SCC[e]	0.167	.23	>.99	.39
Tapetum (R[f])	0.179	.19	>.99	.38
Tapetum (L[g])	0.064	.64	>.99	.72
Projection				
ACR[h] (R)	0.219	.10	>.99	.29
ACR (L)	0.301	.03	.53	.18
PTR[i] (R)	0.417	.003	.048	.04
PTR (L)	0.080	.60	>.99	.72
Association				
SS[j] (R)	0.364	.004	.08	.04
SS (L)	0.162	.28	>.99	.39
Cingulum-CG[k] (R)	0.158	.27	>.99	.39
Cingulum-CG (L)	0.204	.18	>.99	.38
Cingulum-HPC[l] (R)	0.180	.20	>.99	.38
Cingulum-HPC (L)	0.032	.80	>.99	.80
SLF[m] (R)	0.234	.09	>.99	.29
SLF (L)	0.252	.07	>.99	.29
UF[n] (R)	0.054	.71	>.99	.76
UF (L)	–0.149	.30	>.99	.39

[a]All cognitive performance scores were oriented such that higher values indicate better performance; therefore, positive changes in PAL *z* scores reflect improvements in cognitive performance. Analyses were adjusted for age, sex, years of education, baseline cognitive scores, and diffusion tensor imaging metric at baseline. Change scores were calculated as follow-up minus baseline values.

[b]Uncorrected *P* value from partial correlation analysis.

[c]Bonferroni-corrected *P* value.

[d]False discovery rate–adjusted *P* value using the Benjamini-Hochberg method.

[e]SCC: splenium of the corpus callosum.

[f]R: right.

[g]L: left.

[h]ACR: anterior corona radiata.

[i]PTR: posterior thalamic radiation.

[j]SS: sagittal stratum.

[k]CG: cingulum (cingulate gyrus).

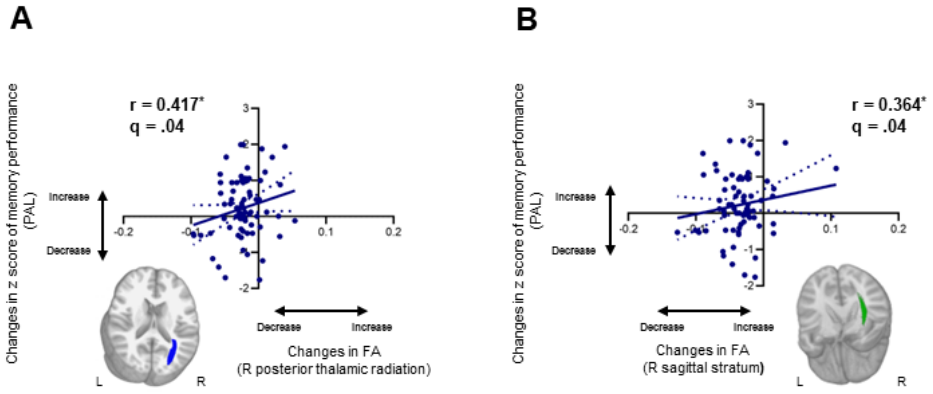
[l]HPC: cingulum (hippocampus).

[m]SLF: superior longitudinal fasciculus.

[n]UF: uncinate fasciculus.

XSL•FO
RenderX

Figure 5. Associations between longitudinal changes in fractional anisotropy (FA) and episodic memory performance (PAL). (A) Right (R) posterior thalamic radiation. (B) Right sagittal stratum. Scatterplots show changes from baseline to week 12 in FA and PAL z scores. Solid lines indicate fitted linear relationships. Reported r values denote correlation coefficients, and q values denote FDR-adjusted *P* values. Insets show the anatomical locations of each tract. Analyses were adjusted for age, sex, years of education, and baseline imaging and cognitive scores. *q<.05.



Intervention Use and Exploratory Dose-Response Analysis

Participants used the program for a mean of 30.73 (SD 12.84) minutes per active day and were active on a mean of 64.24 (SD 16.65) days over the 12-week intervention period, corresponding to approximately 5.35 (SD 1.39) active days per week. Participants completed a mean of 372.52 (SD 200.44) in-app training sessions. The mean total training time was 33.62 (SD 18.53) hours, and the mean time-based adherence rate was 112.1% (SD 61.8%), indicating that participants, on average, exceeded the prescribed minimum training dose of 30 total hours across 12 weeks. In unadjusted exploratory association analyses within the intervention group, the number of active days was positively correlated with PAL *z* score improvement (r=.374; P=.04), suggesting that more frequent engagement with the intervention was associated with greater memory improvement. No significant associations were observed between app-use metrics and FA changes.

Discussion

Principal Findings

This study demonstrated that a 12-week chatbot-based cognitive training intervention was associated with attenuated decline in FA within selected WM tracts and improved episodic memory in CU older adults, a pattern broadly consistent with a brain maintenance framework.

One of the major findings of this study was that DTI analysis revealed attenuated decline in FA in the intervention group compared with controls, most notably in the right PTR and the bilateral tapetum. Notably, significant group-by-time effects were observed primarily in FA, whereas complementary diffusivity metrics (MD, AD, and RD) showed largely null findings after multiple-comparison correction. FA is a composite measure influenced by fiber coherence, myelination, and axonal properties; thus, the present results are best interpreted as relative maintenance of microstructural organization rather than definitive structural restoration. Given the 12-week intervention duration and the modest effect sizes typically observed in diffusion metrics, detectable changes in MD/RD/AD may require longer follow-up and/or larger samples. These tracts are major long-range projection and commissural pathways that support interregional communication and have been implicated in cognitive aging [22,52]. While WM changes in older adults are often widespread, selective differences in projection and commissural fibers observed here are consistent with previous work suggesting activity-related modulation of myelinated pathways [53].

Projection fibers such as the PTR support posterior thalamo-cortical integration, which is relevant to visuospatial information processing and broader cognitive functions [52]. Given that PAL involves visuospatial associative binding [32,34], the observed attenuated FA decline in the right PTR may reflect relative maintenance of long-range pathways supporting visuospatial associative memory. Similarly, commissural fibers such as the tapetum showed attenuated FA decline, reinforcing their role in interhemispheric communication and integration. These commissural fibers support efficient interhemispheric information transfer, which is relevant to memory and higher-order cognition—functions that are particularly vulnerable to age-related decline [54-56].

Furthermore, the intervention group showed significant improvement in PAL performance, a task particularly sensitive to episodic memory. This domain-specific response aligns with the design of the Saemi program, in which memory training was a core, recurring element. The selective improvement in memory is consistent with previous studies reporting task-specific benefits of targeted cognitive interventions [8]. Because PAL is a visuospatial associative memory task, the observed improvement may reflect both a memory-related benefit and alignment between the intervention content and the specific cognitive demands of the task. Accordingly, these findings should not be interpreted as evidence of broad cognitive transfer across multiple domains. In contrast, no significant intervention-related changes were observed in the other cognitive outcomes, including attention, processing speed, executive function, and visuospatial recognition memory. This pattern suggests that the intervention effect on cognition was selective to episodic memory rather than generalized across all measured cognitive domains. The nonsignificant findings in the

XSL•FO
RenderX

other cognitive outcomes may reflect the memory-focused structure of the intervention, limited transfer to less directly trained domains, measurement sensitivity, the relatively short intervention duration, or limited statistical power to detect smaller effects.

These findings are more consistent with brain maintenance reflected by attenuated microstructural decline than with newly induced structural restoration. Because diffusion metrics provide an indirect index of WM microstructure and do not directly quantify myelin or synapse formation, these findings should not be interpreted as evidence of structural "restoration." Nevertheless, repeated engagement in cognitively demanding tasks has been linked to experience-dependent neuroplastic processes—including myelin-related changes and synaptic remodeling—in prior work [27,57,58], which posits that relative maintenance of brain integrity supports sustained cognitive function in older adults [21]. Although these findings are conceptually aligned with a brain-maintenance framework, the absence of longer-term follow-up assessments precludes definitive conclusions regarding the durability of this mechanism. In summary, the observed pattern is more consistent with relative maintenance of WM microstructural organization in pathways relevant to cognitive integration and memory processing, which may be related to intervention-associated improvements in episodic memory [46,53] and is conceptually aligned with a brain maintenance framework [21].

Building on the observed intervention effects on both cognitive performance and WM microstructure, we explored whether changes in these domains were interrelated. Using the imputed ITT dataset, exploratory partial correlation analyses were performed across all participants to examine whether changes in PAL performance were associated with changes in WM microstructure. Change scores were calculated as follow-up minus baseline values. Positive PAL *z* score changes indicate improvement in performance, whereas positive FA changes indicate relative preservation or increase in FA. Partial correlation analyses suggested an association between changes in episodic memory performance and WM microstructural integrity in task-relevant pathways. In addition, PAL improvement was also associated with FA change in the right SS, an associative pathway that includes major posterior cortico-cortical fibers (eg, inferior fronto-occipital and inferior longitudinal components) supporting visual information transfer and integration. Given that PAL relies on visuospatial associative binding, the concurrent PTR-SS associations may suggest that relative maintenance of posterior projection and associative pathways is related to memory-related performance during aging. However, these exploratory structure-function correlations cannot establish causal or mechanistic relationships and should be interpreted as hypothesis-generating findings that require confirmation in larger studies with longer-term follow-up. Nevertheless, this pattern is broadly consistent with the brain maintenance framework, which emphasizes relative preservation of neural structure and function in successful cognitive aging [21,22,52].

Taken together, these findings suggest a possible structure-function relationship in which improvements in episodic memory were associated with relative maintenance of white matter microstructural organization, consistent with a brain maintenance interpretation. Rather than demonstrating a definitive mechanism, our findings indicate that better episodic memory performance was associated with relatively preserved WM pathways implicated in memory-related processing. This interpretation is consistent with the integrated maintenance model, which emphasizes the contribution of both cognitive and structural resilience to successful cognitive aging [21,53,59,60].

Limitations and Future Directions

Several limitations should be acknowledged. First, the relatively small sample size may have limited the statistical power to detect smaller effect sizes and reduced the generalizability of the findings. Future studies with larger, more diverse samples are needed to confirm these results. Second, DTI-derived outcomes were not listed as outcomes in the trial registration, although DTI was included in the IRB-approved neuroimaging protocol. Therefore, these findings should be interpreted as a focused analysis rather than as the primary confirmatory findings of the registered trial. Third, although previous studies suggest that cognitive training effects can persist long-term [1,61], this study assessed only the immediate postintervention effects of a 12-week program. Therefore, the durability and longitudinal trajectory of the observed cognitive and WM changes could not be determined. In addition, without follow-up assessments beyond the intervention period, it remains unclear whether the observed WM differences reflect sustained neuroplastic change or more transient fluctuations. Longitudinal follow-up studies with repeated assessments over time are necessary to clarify the durability of these effects and the potential need for booster sessions. Fourth, potentially influential lifestyle factors such as sleep, physical activity, and social engagement were not assessed. These factors are known to affect WM integrity and may interact with cognitive training effects. Future research could consider multimodal interventions that integrate cognitive, behavioral, and lifestyle components [62]. Fifth, although the intervention emphasized memory training, likely contributing to the significant PAL improvements, no effects were observed in other cognitive domains. This domain-specific pattern suggests that the cognitive benefit was not broadly generalized across domains. Because PAL is a visuospatial associative memory task, the observed improvement may reflect both a memory-related benefit and alignment between the training content and the specific task demands. Future iterations of the program could adjust the training balance or incorporate more targeted modules to evaluate broader transfer effects. Another limitation is that the study did not include an active control condition. Accordingly, nonspecific effects such as expectancy, engagement, general mental stimulation, or digital interaction cannot be fully excluded. Therefore, the observed effects should be interpreted cautiously and should not be attributed solely to the specific cognitive-training content of the intervention. Future studies should include an active control condition matched for contact time, digital interaction, and engagement to better distinguish the specific effects of chatbot-based cognitive training from nonspecific intervention-related effects.

Nonetheless, this study demonstrated that a 12-week chatbot-based cognitive intervention was associated with

XSL•FO
RenderX

improved episodic memory and attenuated decline in WM microstructural integrity within key neural pathways. These findings provide early evidence that chatbot-based cognitive training was associated with improvements in episodic memory and attenuated decline in WM microstructural integrity, suggesting potential cognitive and neurostructural effects that warrant confirmation in larger trials.

Acknowledgments

ChatGPT (OpenAI) was used to support language editing during the preparation and revision of this manuscript. The authors reviewed and edited all AI-assisted content and take full responsibility for the accuracy, originality, and integrity of the manuscript.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT; RS-2025-23963523); the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2026-25576936); and the ANCHOR through the Seoul ANCHOR Center, funded by the Ministry of Education (MOE) and the Seoul Metropolitan Government (2026-ANCHOR-01-034-09).

Data Availability

Deidentified participant data (including the data dictionary) and statistical code are available from the corresponding author upon reasonable request, subject to institutional and ethical approvals.

Authors' Contributions

Conceptualization: BRK, GHK, JHJ, EK, SML
Methodology: BRK, GHK, JHJ, EK, SML
Formal analysis and investigation: BRK
Resources and Administrative support: GHK, JHJ, EK, SML
Writing - original draft preparation: BRK
Writing - review and editing: GHK, JHJ, EK, SML, JK
Supervision: GHK

Conflicts of Interest

Professor JK is the chief executive officer of HAII Corp., and Professor GHK holds shares in HAII Corp. However, the authors declare that the design, conduct, analysis, and reporting of this study were conducted independently of the company, and that HAII Corp. had no role in study conception, data collection, statistical analysis, interpretation, or manuscript preparation. All other authors declare no conflicts of interest.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1118 KB - aging_v9i1e94127_app1.pdf]

Multimedia Appendix 2
Supplementary tables and figures.
[DOCX File , 1145 KB - aging_v9i1e94127_app2.docx]

References

1. Chiu H, Chu H, Tsai J, Liu D, Chen Y, Yang H, et al. The effect of cognitive-based training for the healthy older people: a meta-analysis of randomized controlled trials. PLoS One 2017;12(5):e0176742 [FREE Full text] [doi: 10.1371/journal.pone.0176742] [Medline: 28459873]
2. Malatyali A, Reive C, Williams C. A systematic review of randomized controlled trials of cognitive training interventions in older adults. Innov Aging 2019;3(Suppl 1):S657 [FREE Full text] [doi: 10.1093/geroni/igz038.2434]
3. Smith GE, Housen P, Yaffe K, Ruff R, Kennison RF, Mahncke HW, et al. A cognitive training program based on principles of brain plasticity: results from the Improvement in Memory With Plasticity-Based Adaptive Cognitive Training (IMPACT) study. J Am Geriatr Soc 2009;57(4):594-603 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.02167.x] [Medline: 19220558]
4. Rebok GW, Ball K, Guey LT, Jones RN, Kim H, King JW, ACTIVE Study Group. Ten-year effects of the Advanced Cognitive Training for Independent and Vital Elderly cognitive training trial on cognition and everyday functioning in older adults. J Am Geriatr Soc 2014;62(1):16-24 [FREE Full text] [doi: 10.1111/jgs.12607] [Medline: 24417410]

XSL•FO
RenderX

5. Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, Advanced Cognitive Training for Independent Vital Elderly Study Group. Effects of cognitive training interventions with older adults: a randomized controlled trial. JAMA 2002;288(18):2271-2281 [FREE Full text] [doi: 10.1001/jama.288.18.2271] [Medline: 12425704]
6. Alnajjar F, Khalid S, Vogan AA, Shimoda S, Nouchi R, Kawashima R. Emerging cognitive intervention technologies to meet the needs of an aging population: a systematic review. Front Aging Neurosci 2019;11:291 [FREE Full text] [doi: 10.3389/fnagi.2019.00291] [Medline: 31798439]
7. Bernini S, Stasolla F, Panzarasa S, Quaglini S, Sinforiani E, Sandrini G, et al. Cognitive telerehabilitation for older adults with neurodegenerative diseases in the COVID-19 era: a perspective study. Front Neurol 2020;11:623933 [FREE Full text] [doi: 10.3389/fneur.2020.623933] [Medline: 33519704]
8. Lampit A, Hallock H, Valenzuela M. Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. PLoS Med 2014;11(11):e1001756 [FREE Full text] [doi: 10.1371/journal.pmed.1001756] [Medline: 25405755]
9. Kocaballi AB, Berkovsky S, Quiroz JC, Laranjo L, Tong HL, Rezazadegan D, et al. The personalization of conversational agents in health care: systematic review. J Med Internet Res 2019;21(11):e15360 [FREE Full text] [doi: 10.2196/15360] [Medline: 31697237]
10. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018;25(9):1248-1258 [FREE Full text] [doi: 10.1093/jamia/ocy072] [Medline: 30010941]
11. Kötteritzsch A, Weyers B. Assistive technologies for older adults in urban areas: a literature review. Cogn Comput 2015;8(2):299-317 [FREE Full text] [doi: 10.1007/s12559-015-9355-7]
12. Jang H, Yeo M, Cho J, Kim S, Chin J, Kim HJ, et al. Effects of smartphone application-based cognitive training at home on cognition in community-dwelling non-demented elderly individuals: a randomized controlled trial. Alzheimers Dement (NY) 2021;7(1):e12209 [FREE Full text] [doi: 10.1002/trc2.12209] [Medline: 35005202]
13. Kim Y, Kang Y, Kim B, Kim J, Kim G. Exploring the role of engagement and adherence in chatbot-based cognitive training for older adults: memory function and mental health outcomes. Behav Inf Technol 2024;44(10):2405-2417 [FREE Full text] [doi: 10.1080/0144929x.2024.2362406]
14. Kakao Corp. Earnings release. Kakao IR. URL: https://www.kakaocorp.com/ir/referenceRoom/earningsAnnouncement?lang=en [accessed 2026-01-06]
15. Digital 2024: South Korea. DataReportal. URL: https://datareportal.com/reports/digital-2024-south-korea [accessed 2026-01-06]
16. Wearn AR, Saunders-Jennings E, Nurdal V, Hadley E, Knight MJ, Newson M, et al. Accelerated long-term forgetting in healthy older adults predicts cognitive decline over 1 year. Alzheimers Res Ther 2020;12(1):119 [FREE Full text] [doi: 10.1186/s13195-020-00693-4] [Medline: 32988418]
17. Zelinski EM, Burnight KP. Sixteen-year longitudinal and time lag changes in memory and cognition in older adults. Psychol Aging 1997;12(3):503-513. [doi: 10.1037//0882-7974.12.3.503] [Medline: 9308097]
18. Brito DVC, Esteves F, Rajado AT, Silva N, ALFA score Consortium, Araújo I, et al. Assessing cognitive decline in the aging brain: lessons from rodent and human studies. NPJ Aging 2023;9(1):23 [FREE Full text] [doi: 10.1038/s41514-023-00120-6] [Medline: 37857723]
19. Cabeza R, Albert M, Belleville S, Craik FIM, Duarte A, Grady CL, et al. Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. Nat Rev Neurosci 2018;19(11):701-710 [FREE Full text] [doi: 10.1038/s41583-018-0068-2] [Medline: 30305711]
20. Stern Y, Barnes CA, Grady C, Jones RN, Raz N. Brain reserve, cognitive reserve, compensation, and maintenance: operationalization, validity, and mechanisms of cognitive resilience. Neurobiol Aging 2019;83:124-129 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2019.03.022] [Medline: 31732015]
21. Nyberg L, Lövdén M, Riklund K, Lindenberger U, Bäckman L. Memory aging and brain maintenance. Trends Cogn Sci 2012;16(5):292-305 [FREE Full text] [doi: 10.1016/j.tics.2012.04.005] [Medline: 22542563]
22. Bennett IJ, Madden DJ, Vaidya CJ, Howard DV, Howard JH. Age-related differences in multiple measures of white matter integrity: a diffusion tensor imaging study of healthy aging. Hum Brain Mapp 2010;31(3):378-390 [FREE Full text] [doi: 10.1002/hbm.20872] [Medline: 19662658]
23. Raz N, Gunning-Dixon FM, Head D, Dupuis JH, Acker JD. Neuroanatomical correlates of cognitive aging: evidence from structural magnetic resonance imaging. Neuropsychology 1998;12(1):95-114. [doi: 10.1037//0894-4105.12.1.95] [Medline: 9460738]
24. Schilling KG, Archer D, Yeh F, Rheault F, Cai LY, Hansen C, et al. Aging and white matter microstructure and macrostructure: a longitudinal multi-site diffusion MRI study of 1218 participants. Brain Struct Funct 2022;227(6):2111-2125 [FREE Full text] [doi: 10.1007/s00429-022-02503-z] [Medline: 35604444]
25. Lövdén M, Bodammer NC, Kühn S, Kaufmann J, Schütze H, Tempelmann C, et al. Experience-dependent plasticity of white-matter microstructure extends into old age. Neuropsychologia 2010;48(13):3878-3883 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2010.08.026] [Medline: 20816877]
26. Kolb B, Whishaw IQ. Brain plasticity and behavior. Annu Rev Psychol 1998;49:43-64. [doi: 10.1146/annurev.psych.49.1.43] [Medline: 9496621]

XSL•FO
RenderX

27. Zatorre RJ, Fields RD, Johansen-Berg H. Plasticity in gray and white: neuroimaging changes in brain structure during learning. Nat Neurosci 2012;15(4):528-536 [FREE Full text] [doi: 10.1038/nn.3045] [Medline: 22426254]
28. Ryu HJ, Yang DW. The Seoul Neuropsychological Screening Battery (SNSB) for comprehensive neuropsychological assessment. Dement Neurocogn Disord 2023;22(1):1-15 [FREE Full text] [doi: 10.12779/dnd.2023.22.1.1] [Medline: 36814700]
29. Kim GH, Jeon S, Im K, Kwon H, Lee BH, Kim GY, et al. Structural brain changes after traditional and robot-assisted multi-domain cognitive training in community-dwelling healthy elderly. PLoS One 2015;10(4):e0123251 [FREE Full text] [doi: 10.1371/journal.pone.0123251] [Medline: 25898367]
30. Barnard ND, Bush AI, Ceccarelli A, Cooper J, de Jager CA, Erickson KI, et al. Dietary and lifestyle guidelines for the prevention of Alzheimer's disease. Neurobiol Aging 2014;35 Suppl 2:S74-S78 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2014.03.033] [Medline: 24913896]
31. Lee Y, Back JH, Kim J, Kim S, Na DL, Cheong H, et al. Systematic review of health behavioral risks and cognitive health in older adults. Int Psychogeriatr 2010;22(2):174-187 [FREE Full text] [doi: 10.1017/S1041610209991189] [Medline: 19883522]
32. Kennedy KM, Raz N. Aging white matter and cognition: differential effects of regional variations in diffusion properties on memory, executive functions, and speed. Neuropsychologia 2009;47(3):916-927 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2009.01.001] [Medline: 19166865]
33. Robbins TW, James M, Owen AM, Sahakian BJ, McInnes L, Rabbitt P. Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. Dementia 1994;5(5):266-281. [doi: 10.1159/000106735] [Medline: 7951684]
34. Sabahi Z, Farhoudi M, Naseri A, Talebi M. Working memory assessment using Cambridge Neuropsychological Test Automated Battery can help in the diagnosis of mild cognitive impairment: a systematic review and meta-analysis. Dement Neuropsychol 2022;16(4):444-456 [FREE Full text] [doi: 10.1590/1980-5764-dn-2022-0006] [Medline: 36530766]
35. Lee A, Archer J, Wong CKY, Chen SA, Qiu A. Age-related decline in associative learning in healthy Chinese adults. PLoS One 2013;8(11):e80648 [FREE Full text] [doi: 10.1371/journal.pone.0080648] [Medline: 24265834]
36. Jardim NYV, Bento-Torres NVO, Tomás AM, da Costa VO, Bento-Torres J, Picanço-Diniz CW. Unexpected cognitive similarities between older adults and young people: scores variability and cognitive performances. Arch Gerontol Geriatr 2024;117:105206 [FREE Full text] [doi: 10.1016/j.archger.2023.105206] [Medline: 37742393]
37. Rohlfing T. Incorrect ICBM-DTI-81 atlas orientation and white matter labels. Front Neurosci 2013;7:4 [FREE Full text] [doi: 10.3389/fnins.2013.00004] [Medline: 23355801]
38. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 2004;23 Suppl 1:S208-S219. [doi: 10.1016/j.neuroimage.2004.07.051] [Medline: 15501092]
39. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 2006;31(4):1487-1505. [doi: 10.1016/j.neuroimage.2006.02.024] [Medline: 16624579]
40. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 2002;17(2):825-841. [doi: 10.1016/s1053-8119(02)91132-8] [Medline: 12377157]
41. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. Med Image Anal 2001;5(2):143-156. [doi: 10.1016/s1361-8415(01)00036-6] [Medline: 11516708]
42. Leemans A, Jones DK. The B-matrix must be rotated when correcting for subject motion in DTI data. Magn Reson Med 2009;61(6):1336-1349. [doi: 10.1002/mrm.21890] [Medline: 19319973]
43. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp 2002;17(3):143-155 [FREE Full text] [doi: 10.1002/hbm.10062] [Medline: 12391568]
44. Mori S, Wakana S, Van ZP, Nagae-Poetscher L. MRI Atlas of Human White Matter. Amsterdam, Netherlands: Elsevier; 2005.
45. Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich DS, et al. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. Neuroimage 2008;39(1):336-347 [FREE Full text] [doi: 10.1016/j.neuroimage.2007.07.053] [Medline: 17931890]
46. Bhattarai P, Thakuri DS, Nie Y, Chand GB. Explainable AI-based deep-SHAP for mapping the multivariate relationships between regional neuroimaging biomarkers and cognition. Eur J Radiol 2024;174:111403 [FREE Full text] [doi: 10.1016/j.ejrad.2024.111403] [Medline: 38452732]
47. Lee EH, Kim BR, Kim H, Kim SH, Chun MY, Park HK, et al. Four-week, home-based, robot cognitive intervention for patients with mild cognitive impairment: a pilot randomized controlled trial. Dement Neurocogn Disord 2020;19(3):96-107 [FREE Full text] [doi: 10.12779/dnd.2020.19.3.96] [Medline: 32985149]
48. Twisk J, de Boer M, de Vente W, Heymans M. Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. J Clin Epidemiol 2013;66(9):1022-1028. [doi: 10.1016/j.jclinepi.2013.03.017] [Medline: 23790725]

XSL•FO
RenderX

49. Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Jager KJ. Intention to treat and per protocol analysis in clinical trials. Nephrology (Carlton) 2020;25(7):513-517. [doi: 10.1111/nep.13709] [Medline: 32147926]
50. Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Soft 2011;45(3):1-67 [FREE Full text] [doi: 10.18637/jss.v045.i03]
51. Rubin DB. Multiple Imputation for Survey Nonresponse. New York: John Wiley & Sons; 1987.
52. Catani M, Thiebaut de Schotten M. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex 2008;44(8):1105-1132. [doi: 10.1016/j.cortex.2008.05.004] [Medline: 18619589]
53. Engvig A, Fjell AM, Westlye LT, Moberget T, Sundseth Ø, Larsen VA, et al. Memory training impacts short-term changes in aging white matter: a longitudinal diffusion tensor imaging study. Hum Brain Mapp 2012;33(10):2390-2406 [FREE Full text] [doi: 10.1002/hbm.21370] [Medline: 21823209]
54. Salat DH, Tuch DS, Greve DN, van der Kouwe AJW, Hevelone ND, Zaleta AK, et al. Age-related alterations in white matter microstructure measured by diffusion tensor imaging. Neurobiol Aging 2005;26(8):1215-1227. [doi: 10.1016/j.neurobiolaging.2004.09.017] [Medline: 15917106]
55. Schmahmann JD, Smith EE, Eichler FS, Filley CM. Cerebral white matter: neuroanatomy, clinical neurology, and neurobehavioral correlates. Ann NY Acad Sci 2008;1142:266-309 [FREE Full text] [doi: 10.1196/annals.1444.017] [Medline: 18990132]
56. Ryberg C, Rostrup E, Paulson OB, Barkhof F, Scheltens P, van Straaten ECW, LADIS study group. Corpus callosum atrophy as a predictor of age-related cognitive and motor impairment: a 3-year follow-up of the LADIS study cohort. J Neurol Sci 2011;307(1-2):100-105. [doi: 10.1016/j.jns.2011.05.002] [Medline: 21621224]
57. Park DC, Bischof GN. The aging mind: neuroplasticity in response to cognitive training. Dialogues Clin Neurosci 2013;15(1):109-119 [FREE Full text] [doi: 10.31887/DCNS.2013.15.1/dpark] [Medline: 23576894]
58. Fields RD. White matter in learning, cognition and psychiatric disorders. Trends Neurosci 2008;31(7):361-370 [FREE Full text] [doi: 10.1016/j.tins.2008.04.001] [Medline: 18538868]
59. Chapman SB, Aslan S, Spence JS, Hart JJ, Bartz EK, Didehbani N, et al. Neural mechanisms of brain plasticity with complex cognitive training in healthy seniors. Cereb Cortex 2015;25(2):396-405 [FREE Full text] [doi: 10.1093/cercor/bht234] [Medline: 23985135]
60. Lövdén M, Schaefer S, Noack H, Bodammer NC, Kühn S, Heinze H, et al. Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. Neurobiol Aging 2012;33(3):620.e9-620.e22 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2011.02.013] [Medline: 21497950]
61. Kelly ME, Loughrey D, Lawlor BA, Robertson IH, Walsh C, Brennan S. The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: a systematic review and meta-analysis. Ageing Res Rev 2014;15:28-43. [doi: 10.1016/j.arr.2014.02.004] [Medline: 24607830]
62. Wassenaar TM, Yaffe K, van der Werf YD, Sexton CE. Associations between modifiable risk factors and white matter of the aging brain: insights from diffusion tensor imaging studies. Neurobiol Aging 2019;80:56-70 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2019.04.006] [Medline: 31103633]

Abbreviations

ACR: anterior corona radiata
AD: axial diffusivity
CANTAB: Cambridge Neuropsychological Test Automated Battery
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
CRIS: Clinical Research Information Service
CU: cognitively unimpaired
DMS: delayed matching to sample
DTI: diffusion tensor imaging
FA: fractional anisotropy
FDR: false discovery rate
FLIRT: FMRIB's Linear Image Registration Tool
fMRI: functional magnetic resonance imaging
FMRIB: Functional Magnetic Resonance Imaging of the Brain
FSL: FMRIB Software Library
IRB: institutional review board
ITT: intention-to-treat
MD: mean diffusivity
MRI: magnetic resonance imaging
OTS: One Touch Stockings of Cambridge
PAL: Paired Associates Learning

XSL•FO
RenderX

PP: per-protocol
PTR: posterior thalamic radiation
RD: radial diffusivity
ROI: region of interest
RTI: Reaction Time
RVP: Rapid Visual Information Processing
SHAP: Shapley additive explanations
SLF: superior longitudinal fasciculus
SNSB-II: Seoul Neuropsychological Screening Battery, second edition
SS: sagittal stratum
SWM: Spatial Working Memory
TBSS: tract-based spatial statistics
TE: echo time
TR: repetition time
WM: white matter

Edited by Y Jiang, C Lee; submitted 29.Mar.2026; peer-reviewed by E Genc, CM Huang; comments to author 22.May.2026; revised version received 23.Sep.2026; accepted 24.Sep.2026; published 28.Oct.2026.

Please cite as:
Kim BR, Kim E, Lim SM, Kim J, Jeong JH, Kim GH
Impact of Chatbot-Based Cognitive Training on Cognitive Function and White Matter Integrity in Cognitively Unimpaired Older Adults: Randomized Controlled Trial
JMIR Aging 2026;9:e94127
URL: https://aging.jmir.org/2026/1/e94127
doi:10.2196/94127
PMID:

©Bori R Kim, Eunhee Kim, Soo Mee Lim, Jinwoo Kim, Jee Hyang Jeong, Geon Ha Kim. Originally published in JMIR Aging (https://aging.jmir.org), 28.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Facilitators and Barriers to Digital Self-Management in Older Adults With Depression: COM-B and Theoretical Domain Framework Qualitative Study

Ziping Zhu[1,2*], MS; Xinying Lin[1*], MS; Zhi Chen[1], MS; Hong Li[1], PhD; Rong Lin[1], PhD

[1]School of Nursing, Fujian Medical University, Shangjie Campus of Fujian Medical University, Minhou County, Fuzhou, Fujian, China

[2]Fudan Zhongshan Xiamen Hospital, Xiamen, Fujian, China

[*]these authors contributed equally

Corresponding Author:
Rong Lin, PhD
School of Nursing, Fujian Medical University, Shangjie Campus of Fujian Medical University, Minhou County, Fuzhou, Fujian, China

Abstract

Background: Depression in older adults presents unique challenges in self-management. Digital tools, such as mobile health (mHealth) apps, have the potential to support this population. This study explored the facilitators and barriers to digital self-management in older adults with depression to inform the design of effective mHealth apps.

Objective: This study aims to explore the facilitators and barriers to digital self-management in older patients with depression with the aim of informing the design and development of mHealth apps for older adults.

Methods: A purposive sampling method was used to recruit 25 older patients with depression from July to September 2024. Semistructured interviews were conducted to capture real-life experiences. Directed content analysis ensured objective and accurate data interpretation, and the Capability, Opportunity, Motivation, and Behavior (COM-B) model and the Theoretical Domains Framework were applied to identify facilitators and barriers related to behavior.

Results: Six themes were identified based on the COM-B model and the Theoretical Domains Framework: perception of illness and accumulation of personal experience; dual challenges of cognitive abilities and physical limitations; integration of digital technologies and acceptance differences; social influences and access to and utilization of support resources; environmental constraints and accommodations; and intertwined influences of beliefs, emotions, and motivation. A total of 13 barriers and 11 facilitators were identified.

Conclusions: Digital self-management in older adults with depression is complex and influenced by multiple interrelated factors. Effective mHealth apps must integrate the cognitive, emotional, and social contexts of patients to provide user-friendly, personalized solutions.

(JMIR Aging 2026;9:e79253) doi:10.2196/79253

KEYWORDS

geriatric depression; mHealth; Theoretical Domains Framework; TDF; Capability, Opportunity, Motivation, and Behavior model; COM-B model; qualitative descriptive study; mobile health

Introduction

Background

With the increasing integration of information and communication technologies into health care, mobile health (mHealth) apps have become an important modality within digital health and are increasingly used to support depression management [1-3]. In parallel, self-management has gained prominence in chronic disease care as a means of enabling individuals to actively manage symptoms, improve adherence to treatment, and enhance daily functioning while reducing dependence on formal medical services [4]. Self-management emphasizes personal agency, perceived control, and the translation of therapeutic principles into daily life practices [5,6]. In the context of depression, self-management commonly includes mental health literacy, psychoeducation, medication adherence support, behavioral activation, and cognitive behavioral strategies [7]. Compared with traditional face-to-face services, digital self-management interventions offer advantages, such as improved accessibility, reduced travel burden, scalability, lower costs, and enhanced protection of privacy [8-11]. Evidence from randomized controlled trials suggests that smartphone-based interventions can reduce depressive symptoms [12] and expand access to care, particularly for

XSL•FO
RenderX

individuals who face barriers to conventional mental health services [13-16].

The growing interest in digital self-management is particularly relevant given the global burden of depression, one of the most prevalent mental health disorders worldwide and a leading cause of disability. According to the World Health Organization, more than 300 million people globally are affected by depression, and its contribution to disability-adjusted life years continues to increase [17]. Depression is particularly prevalent among older adults and not only directly leads to individual disability but also imposes a substantial economic burden on society [18,19]. Studies project that between the years 2023 and 2032, cumulative all-cause medical expenditure for patients with depression will reach US $309 million, which is 2.6 times more than that for nondepressed individuals [20]. Among individuals aged 60 years and older, global prevalence estimates range from 28.4% to 35.1%, with mortality rates reported to be as high as 9.8% [21,22]. Late-life depression is characterized by high recurrence and mortality and yet remains markedly underdiagnosed and undertreated [23,24]. Beyond its clinical impact, depression imposes considerable economic costs through reduced workforce participation, impaired productivity, increased health care utilization, and the transfer of caregiving responsibilities to family members, placing sustained strain on intergenerational care systems [22,25,26]. Despite this burden, fewer than 10% of individuals with depression worldwide receive any form of treatment, and fewer than 1% receive minimally adequate care [27]. Older adults face additional barriers, including limited service accessibility, transportation difficulties, stigma, and reduced awareness of help-seeking pathways, further increasing the risk of untreated or poorly managed depression [28-31].

Effective self-management is therefore particularly important for older adults with depression [32], although its real-world implementation remains challenging. Depression frequently coexists with multimorbidity, functional decline, and age-related cognitive and sensory changes, increasing the complexity of daily symptom management. Core depressive symptoms, such as low motivation, fatigue, and hopelessness, directly undermine the ability of individuals to initiate and maintain self-management behaviors over time [33]. Although mHealth apps provide tools for symptom monitoring, psychoeducation, and remote support, existing evidence indicates a paradoxical pattern in their use among older adults with late-life depression. Specifically, while mHealth interventions have demonstrated efficacy in controlled trials, such as reducing depressive symptoms and improving sleep quality [34], these benefits have not translated into sustained engagement or effective self-management in the real-world setting, where low uptake, rapid disengagement, and poor long-term adherence are commonly reported [35].

This discrepancy has been attributed to multiple interacting barriers, including fear of mental illness–related stigma, suboptimal user interface designs, and age-related physical and cognitive limitations [36,37]. In the real-world context, additional constraints, such as inadequate digital literacy and privacy concerns, further restrict the adoption and continued use of mHealth tools among older adults [38,39]. In addition, existing apps are rarely tailored to the specific psychosocial needs of this population and often overlook themes central to late-life depression, such as grief, loss, loneliness, and social isolation [39]. However, most studies in this field have relied predominantly on quantitative methods, focusing on the acceptance, usability, or impact on symptom outcomes of technology. Although these approaches are valuable for identifying general influencing factors, such as the digital divide and technophobia, they provide limited insight into how older adults with depression perceive, interpret, and emotionally respond to mHealth interventions in everyday life [40]. In particular, they fail to capture how depressive symptoms dynamically interact with technology use to shape both the initiation and sustained engagement in digital self-management behaviors. Without such experiential and contextual understanding, mHealth interventions risk remaining technically effective in controlled trials but poorly aligned with the needs of users in the real world, perpetuating the persistent gap between effectiveness and participation. Therefore, a qualitative approach is essential to elucidate the mechanisms through which individual, technological, and contextual factors jointly influence engagement in digital self-management among older adults with depression.

Theoretical Framework

For older adults with depression, digital self-management is not merely a matter of technology adoption but rather a complex process of modifying health behavior. Therefore, understanding this process is crucial for designing effective intervention strategies. Evidence suggests that using theoretical frameworks to identify and analyze determinants of behavior can substantially increase the likelihood of intervention effectiveness [41]. Theoretical frameworks facilitate a comprehensive understanding and interpretation of the complexity of behavior change, providing a solid foundation for effective intervention design [42]. Interventions grounded in behavioral theory models have been shown to be more effective than interventions not based on theory [43]. Therefore, incorporating theoretical frameworks is essential for understanding behavioral changes. Among various theories, the Capability, Opportunity, Motivation, and Behavior (COM-B) model and its associated Theoretical Domains Framework (TDF) have been widely used to identify behavioral facilitators and barriers, providing strong support for designing targeted interventions [44,45]. To comprehensively understand, analyze, and explain the multidimensional factors influencing the adoption and sustained use of digital self-management by older adults, this study adopted the COM-B model and the TDF as its theoretical guiding frameworks.

The COM-B model posits that behavior arises from the interaction of capability, opportunity (social and physical), and motivation (automatic and reflective processes) [46]. Building on this model, the TDF translates these components into 14 operational domains, including knowledge, skills, social influences, and environmental resources, offering a comprehensive and structured tool for systematically identifying and analyzing behavioral determinants [46]. Together, the COM-B model and the TDF form the core of the behavior change wheel, serving as an integrated tool for describing

XSL•FO
RenderX

behavior and designing interventions through a concise yet comprehensive analytical structure. Although the COM-B model explains behavior formation through capability, opportunity, and motivation, the TDF further elaborates on these dimensions, enabling a more comprehensive exploration of behavior change possibilities through in-depth interviews and analysis of mHealth self-management among older adults with depression [46]. Compared with other commonly used behavior change theories, the COM-B model and TDF demonstrate advantages in analyzing complex health behaviors. For example, the health belief model emphasizes the perception of disease threat and behavioral benefits by the individual but does not adequately account for social and environmental factors [47]. The theory of planned behavior primarily focuses on individual cognitive factors, such as attitudes and subjective norms, resulting in a relatively narrow perspective [48]. Although the transtheoretical model emphasizes the dynamic nature of behavior change, it lacks systematic identification and classification of specific determinants, limiting its practical applicability [49]. Consequently, the integrated COM-B model and the TDF more comprehensively capture the complex interactions between individual internal factors and external environmental influences on health behaviors, making them particularly suitable for analyzing digital self-management, which is shaped by multiple interacting factors.

In this study, the TDF was used to guide the development of the semistructured interview guide to ensure comprehensive coverage of the theoretical domains that may influence digital self-management behaviors among older adults with depression. Subsequently, qualitative analysis of interview transcripts was conducted to categorize identified barriers and facilitators into relevant TDF domains and ultimately map them onto the three core COM-B components: capability, opportunity, and motivation. This analytical process allowed for a clear and structured presentation of the findings and provided direct theoretical justification and practical guidance to develop feasible and targeted interventions based on the behavior change wheel. Currently, qualitative studies that have applied the combined TDF and COM-B model to explore digital self-management behaviors among older adults with depression in China are scarce. Therefore, this framework provides a solid theoretical foundation for understanding digital self-management behaviors in older adults with depression to inform the design of tailored, context-sensitive digital interventions suitable for Chinese patients.

This study aims (1) to systematically identify the facilitators and barriers that influence the adoption and sustained participation in mHealth interventions among older adults with depression and (2) to map these factors onto the TDF and the COM-B model to elucidate key behavioral determinants, thus informing the development of future strategies to promote the use of mHealth in this population.

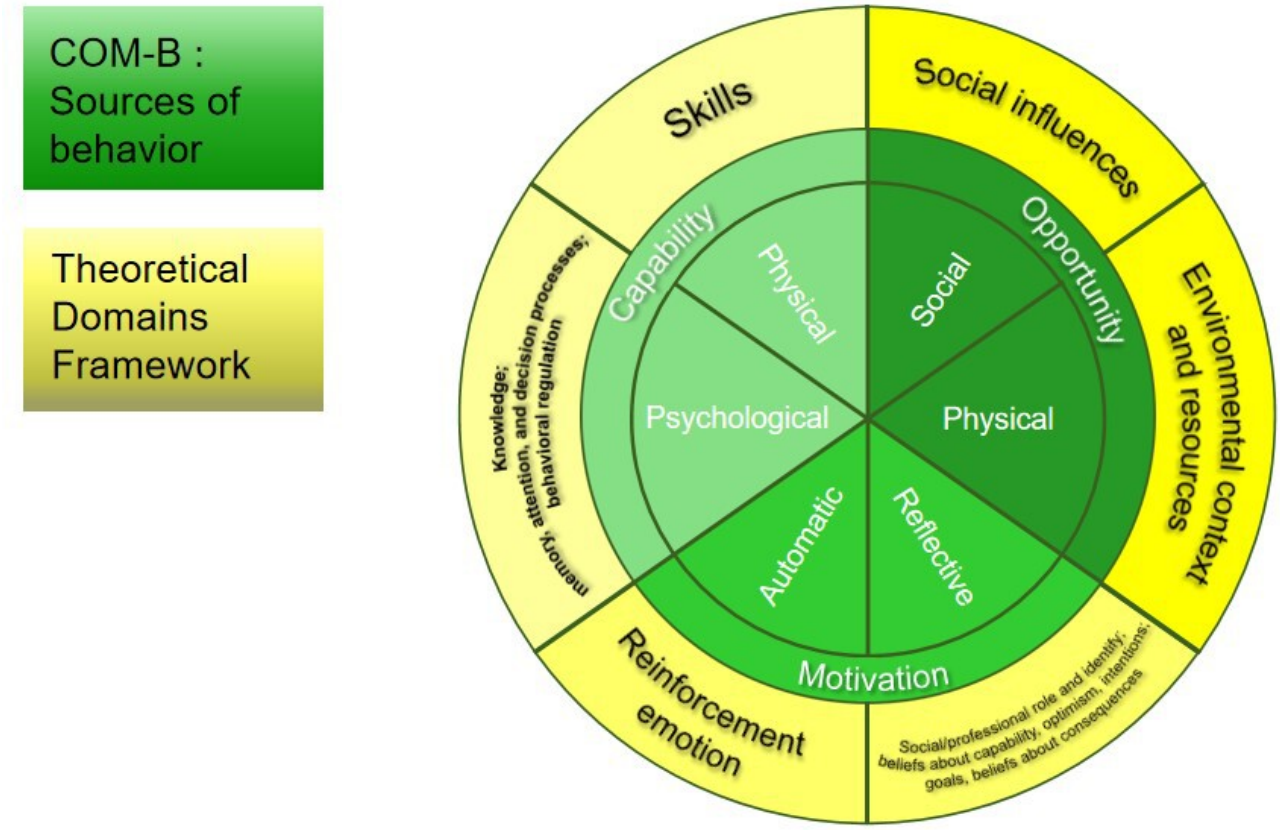
Methods

Study Design

This study used a descriptive qualitative research methodology that followed the philosophical basis of naturalistic inquiry, which emphasizes the interpretation of patient experience or the presentation of events in simple language [50]. Valuable information can be gathered from understanding an individual's reactions, thoughts, facilitators, or hindrances to events [51]. To better understand the factors influencing the use of digital tools for self-management in older people with depression, it is important to explore their feelings, experiences, and perceptions related to seeking help. Data were collected through semistructured interviews with open-ended questions. The study used the COM-B model and the TDF to design the outline of the interview and to provide a framework for organizing and analyzing the data (Figure 1). The study was presented in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Checklist 1 [52]).

XSL•FO
RenderX

Figure 1. Schematic diagram of the Capability, Opportunity, Motivation, and Behavior (COM-B) model and the Theoretical Domains Framework (TDF).



Study Setting and Recruitment

This study was conducted in 5-star nursing institutions and communities in Fuzhou City, Fujian Province, China. Purposive sampling was used to select interviewees, among those who met the above criteria, and the individuals most likely to provide rich information were selected. The maximum difference sampling strategy was used to account for differences in sex (male or female), age, literacy levels (ranging from primary school to college), economic situation (monthly income starting from <1000 to >5000 CNY [1 CNY=US $0.14]), and disease severity of older patients with depression (mild, moderate, or severe based on 15-item Geriatric Depression Scale scores). Additionally, to ensure diversity in digital experiences, we considered participants' smartphone ownership and their frequency of use. Based on participants' self-reported average daily usage duration during the interviews, usage frequency was categorized as low (≤1 hour/day), medium (1 - 4 hours/day), or high (≥4 hours/day). The sample size was determined based on the principle of meaning saturation (see Multimedia Appendix 1 for details). After each interview, the research team held regular meetings to review transcripts, identify emerging themes, and discuss whether new information was still being obtained. Saturation was considered achieved when 2 consecutive interviews did not yield new substantive themes related to the research questions, and additional data collection appeared redundant. The final decision on saturation was reached through consensus during these team discussions.

Inclusion and Exclusion Criteria

The inclusion criteria were as follows: (1) aged 60 years or older, (2) meeting the diagnostic criteria for depression in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, (3) a 15-item Geriatric Depression Scale score of 5 or higher, (4) ability to communicate; (5) voluntary participation and provision of informed consent, and (6) no requirement for smartphone ownership or prior experience using smartphones. Including participants with heterogeneous digital skills may increase variability, whereas restricting participation to digitally experienced older adults could limit the transferability of the findings [53,54]. Older adults are not a digitally homogeneous group, and excluding those without smartphone experience may overestimate the feasibility of digital self-management and exacerbate digital inequalities [55]. Including participants with diverse levels of digital engagement allowed for a more realistic assessment of barriers and facilitators in real-world settings. The exclusion criteria were as follows: (1) serious physical illnesses preventing cooperation with the interview, (2) severe mental illnesses preventing participation in the study, and (3) illiterate individuals or those unable to understand basic written information.

Data Collection

Data were collected between July and September 2024. Information from the literature [56,57] was used for the TDF and was mapped to the components of the COM-B model to develop a semistructured interview outline. Three patients who met the inclusion criteria for preinterviews through the convenience sampling method (preinterview information was

XSL•FO
RenderX

not included in the final analysis) were selected, and consultations with relevant experts on issues that arose during the interviews allowed for modifications to the outline. The final outline of the interview was confirmed after rediscussion by the research team and is provided in Multimedia Appendix 2.

Two methods were used for data collection: (1) a demographic survey (covering age, sex, education, and economic situation) and (2) in-depth, semistructured, one-on-one face-to-face interviews. Before the interview, the researcher (ZZ) introduced herself, explained the purpose and significance of the study, outlined the necessity of on-site audio recording and the principle of confidentiality, obtained written informed consent from the interviewees, and assisted the participants in filling out the basic information form. Face-to-face interviews were conducted when the patients were free and in good condition, in a separate room without interference. The interview was conducted with open-ended questions that referred to the outline of the interview but were not restricted to the language or order of the outline. During the interview, attention was paid to recording the patient's tone of voice, expressions, movements, and other nonverbal behaviors of the patient. The interviewer also listened carefully to the patient's statements, asked questions and responded to them at the right time, and avoided the use of guiding or suggestive language. The entire interview was recorded, and the key content was noted. The interview lasted 20 to 40 minutes. After the interview, the recorded information was listened to two or three times to ensure accuracy [1], and the information was entered into Microsoft Word.

To enhance the credibility of the data analyses, participants were invited to review and validate the content after the interviews to ensure an accurate understanding and reflection of their views. Detailed records and documentation, including interview outlines, data analysis processes, and rational decision-making, were maintained to facilitate the subsequent review of the data and validate the transparency of the research process.

Data Analysis

All audio recordings were transcribed into a Word document within 24 hours of the conclusion of the interview. The verbatim transcription of the audio-recorded material was accompanied by the labeling and integration of nonverbal information such as pauses, facial expressions, and body language recorded in the field notes of the interviews. The interview data were anonymized, and the transcripts were numbered with P1, P2, P3,..., Pn codes. Other basic demographic and disease information of the study participants was recorded in a separate document from the transcribed manuscript to ensure systematic and confidential data management.

The transcribed data were analyzed using a directed content analysis approach [58,59], guided by the TDF and the COM-B model. Directed content analysis is an appropriate method when existing theory or previous research can guide initial coding and categorization, while still allowing openness to new insights. The analytical process followed 3 main stages: preparation, organization, and reporting. First, the analysis unit was defined, with meaningful sentences reflecting the experiences, perceived barriers, and facilitators of participants related to mHealth-supported self-management used as the smallest analytical units. The transcripts were then read repeatedly to achieve immersion and gain an overall understanding of the data. In the organization stage, an initial coding framework was developed deductively based on the predefined TDF domains and corresponding COM-B components (capability, opportunity, and motivation). Line-by-line coding was conducted, during which relevant text segments were coded and mapped to the corresponding TDF domains. Similar codes were subsequently grouped into subthemes, which were further organized into higher-order themes aligned with the COM-B model. Although the analytic strategy allowed for inductive coding of data that did not fit the predefined theoretical framework, all the content identified in this study could be meaningfully mapped to the TDF domains and COM-B components. Therefore, the presentation of results reflects primarily findings derived from the directed content analysis. Two co–first authors independently performed the deductive coding and mapping process using established definitions of the COM-B model [46] and the TDF [60]. Any discrepancies were resolved through team discussion to ensure analytical rigor.

Rigor

The rigor of this study was ensured through systematic steps of directed content analysis and was elaborated upon using the criteria proposed by Colorafi et al [61], specifically across the 4 dimensions of credibility, transferability, confirmability, and consistency [62]. First, to establish credibility, we used in-depth interviews and member checking: all interviews were conducted by the same researcher (ZZ) to maintain consistency in interaction, and rapport was built during interviews to encourage participants to elaborate. Furthermore, we returned key interview summaries to participants for content verification, thereby validating the accuracy of understanding. Second, to ensure transferability, we comprehensively presented the research context, participant characteristics, purposive sampling strategy, and detailed data collection and analysis procedures in the *Methods* section, providing a sufficient basis for other researchers to judge the applicability of this study's context and findings. Regarding confirmability, we implemented cross-verification among coders: initial coding and the mapping of categories to the COM-B model were independently performed by two researchers (ZZ and XL), followed by review and approval by team members not involved in the initial coding. All analytical decisions and processes were documented for reference, forming a traceable audit trail. Finally, in terms of consistency, the entire study strictly followed the established stages of directed content analysis, and consensus was reached through discussions with the research team at each key analytical junction, thus ensuring the logical coherence and methodological stability of the analytical process.

Reflexivity

In qualitative research, ensuring the rigor and quality of the study is a core responsibility of the researcher. In this context, researcher reflexivity is regarded as a crucial standard to enhance the credibility of the study [63]. Qualitative researchers use reflexivity to explain how their own subjectivity influences the

XSL•FO
RenderX

research process and to make the entire process more transparent [64]. This study was led by 2 experts in the fields of geriatric depression and mental health. All interviews were conducted by ZZ, while the primary coding and mapping analysis to the COM-B model were performed independently by ZZ and XL, with ongoing mutual feedback throughout the process. As the researcher is the primary instrument for data collection and interpretation, throughout the research process, we continuously examined the researcher's professional background (eg, in clinical psychology and geriatric health) and performance considering preexisting knowledge of digital health technologies, as well as potential assumptions and emotional responses formed during interview interactions, through team discussions, research journals, and periodic reflections. This ongoing reflection aimed to clarify the influence of the researcher's position on the study. For example, building an empathetic relationship during interviews may enhance the openness of participants to sensitive topics (such as experiences of depression), whereas the researcher's academic training may also influence the initial interpretation of "self-management" behaviors in the data. By making such reflections explicit and incorporating them into the analytical considerations, we strived to achieve a deeper understanding of the participants' experiences while enhancing the transparency of the research process and the rigor of interpretation as much as possible.

Ethical Considerations

This study received approval from the Ethics Committee of Fujian Medical University (number 2024‐87). All participants were clearly informed of the research purpose, their rights, and the principles of data confidentiality. Written or oral informed consent was obtained from each participant prior to the interviews. Audio recordings were securely stored during the research period and were uniformly destroyed upon the completion of the study. Participants had the right to withdraw unconditionally at any stage of the research without any adverse consequences. All collected data were anonymized. Participants who completed all interviews received a gift as a token of appreciation.

Results

Study Participants

Twenty-five older adult patients with depression were enrolled in this study, with an age range of 60 to 86 years (mean 76.0, SD 6.1), including 15 female patients and 10 male patients. The education levels ranged from elementary school to undergraduate degree. In terms of severity of depression, 4 cases of severe depression, 12 cases of moderate depression, and 9 cases of mild depression were identified. Regarding smartphone ownership, 22 participants owned a smartphone, whereas 3 did not. Based on the content of the interview, the frequency of smartphone usage was categorized as follows: 14 participants had medium usage frequency, 8 had low usage frequency, and 3 had high usage frequency. This distribution across multiple dimensions supports the representativeness of the sample in capturing a broad spectrum of experiences related to digital self-management among older adults with depression. The basic characteristics of the participants are detailed in Table 1. The average duration of patient interviews in this study was approximately 30 minutes.

XSL•FO
RenderX

Table . General information on older patients with depression (n=25).

No.	Sex	Age (y)	Educational level	Economic situation (CNY)[a]	Smartphone	Smartphone use frequency	Depression severity
P1	Female	86	College	>5000	Yes	Medium	Moderate
P2	Female	85	Middle or high school	3000 - 5000	No	Low	Severe
P3	Female	75	College	>5000	Yes	Medium	Moderate
P4	Male	85	College	>5000	Yes	High	Mild
P5	Female	80	Middle or high school	3000 - 5000	Yes	Medium	Mild
P6	Male	80	Middle or high school	3000 - 5000	Yes	Medium	Moderate
P7	Female	70	Middle or high school	3000 - 5000	Yes	Medium	Mild
P8	Male	75	Primary school	<1000	Yes	Low	Moderate
P9	Female	77	Primary school	<1000	Yes	Medium	Severe
P10	Female	78	Middle or high school	1000 - 3000	Yes	Medium	Severe
P11	Female	82	College	>5000	Yes	Medium	Moderate
P12	Female	77	Primary school	>1000	No	Low	Severe
P13	Male	75	Middle or high school	3000 - 5000	Yes	Medium	Moderate
P14	Female	73	College	>5000	Yes	High	Mild
P15	Male	78	College	>5000	Yes	Medium	Moderate
P16	Female	82	Middle or high school	3000 - 5000	Yes	Low	Mild
P17	Female	78	College	>5000	Yes	Medium	Moderate
P18	Female	82	Middle or high school	>5000	No	Low	Mild
P19	Male	71	Middle or high school	3000 - 5000	Yes	High	Mild
P20	Male	60	Middle or high school	3000 - 5000	Yes	Low	Mild
P21	Male	74	Middle or high school	1000 - 3000	Yes	Low	Moderate
P22	Female	63	College	3000 - 5000	Yes	Medium	Moderate
P23	Male	73	College	>5000	Yes	Medium	Moderate
P24	Male	70	Primary school	1000 - 3000	Yes	Low	Moderate
P25	Female	72	Primary school	3000 - 5000	Yes	Medium	Mild

[a]An exchange rate of 1 CNY=US $0.14 is applicable.

Key Themes and Related Areas

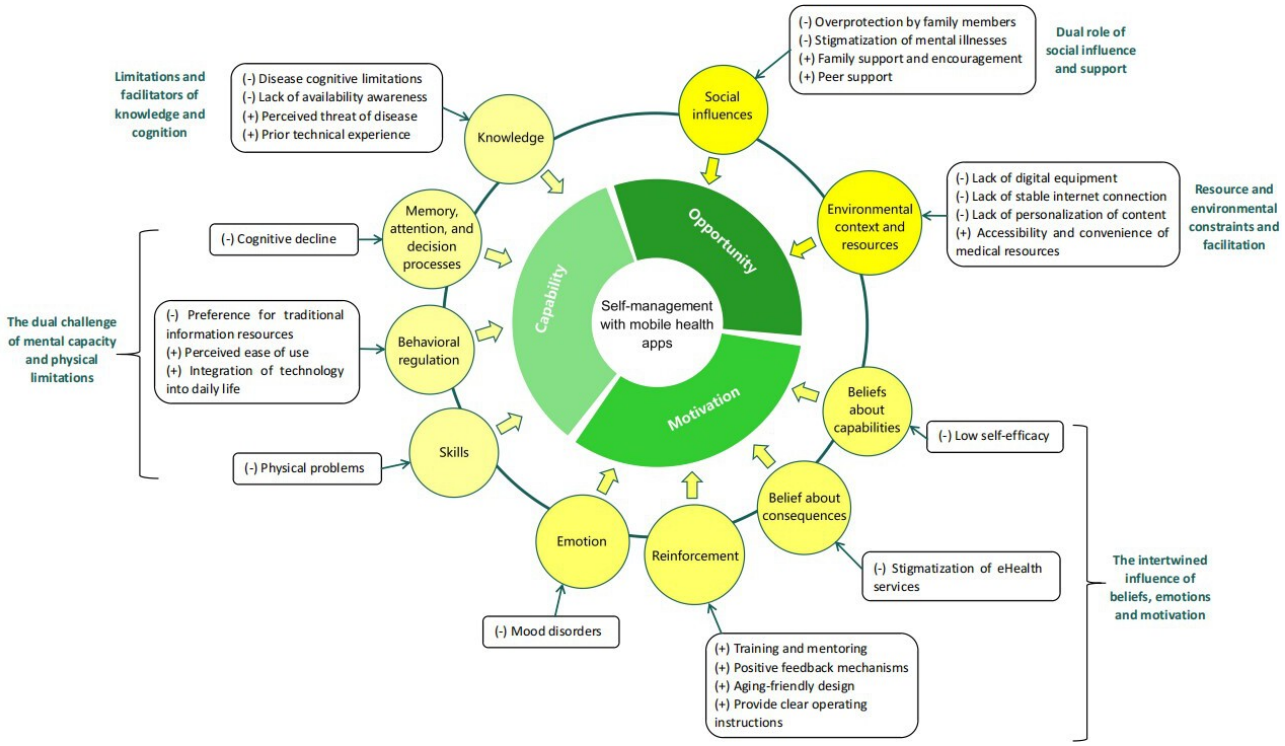
Overview

To improve theoretical clarity, the results are organized according to the COM-B model as the primary analytical hierarchy. Within each COM-B model component (capability, opportunity, motivation), relevant TDF domains and empirically derived themes are presented as subsections. Six themes were identified that were directly related to the 10 TDF domains and 3 COM-B components. The 10 TDF domains were "knowledge"; "memory, attention and decision processes"; "behavioral regulation"; "skills"; "social influences"; "environmental context and resources"; "belief about consequences"; "belief about capabilities"; "reinforcement"; and "emotion," whereas TDF domains not identified as relevant to the use of mHealth apps to support self-management behaviors were "optimism," "intentions," "goals," and "social or professional roles and identities." These findings are

XSL•FO
RenderX

described in more detail below using the TDF and the corresponding COM-B model (Figure 2). Table 2 provides an overview of the analytical framework, summarizing the facilitators and barriers identified across the TDF domains and COM-B components for participants, whereas detailed theme descriptions and illustrative quotes are presented in Multimedia Appendix 3. The complete data analysis process, from raw interview statements to open codes and subthemes, is provided in Multimedia Appendix 4.

Figure 2. Mapping of themes to the Capability, Opportunity, Motivation, and Behavior (COM-B) model and the Theoretical Domains Framework (TDF).



XSL•FO
RenderX

Table . Facilitators and barriers to digital self-management for older patients with depression, categorized by Theoretical Domains Framework (TDF) domains and Capability, Opportunity, Motivation, and Behavior (COM-B) components.

COM-B components and TDF	Themes	Barriers (number of participants who mentioned them)	Facilitators (number of participants who mentioned them)
Capability			
Psychological capability			
Knowledge	Disease perception and personal experience building	• Cognitive limitations of the disease (n=13) • Lack of availability awareness (n=8)	• Perceived threat of disease (n=7) • Prior technical experience (n=13)
Memory, attention, and decision processes	Dual challenges of cognitive function and physical limitations	Cognitive decline (n=16)	• Not identified
Behavioral regulation	Digital technology integration and life adaptation	• Preference for traditional information resources (n=7) • Not identified	• Perceived ease of use (n=9) • Integration of technology into daily life (n=9)
Physical capability			
Skills	Dual challenges of cognitive function and physical limitations	• Restrictions on physical functioning (n=10)	• Not identified
Opportunity			
Social opportunity			
Social influences	Access to and utilization of social impact and support resources	• Overprotection by family members (n=11) • Stigmatization of mental illnesses (n=4)	• Family support and encouragement (n=14) • Peer support (n=14)
Physical opportunity			
Environmental context and resources	Resource and environmental constraints and facilitation	• Lack of digital equipment (n=3) • Lack of stable internet connection (n=2) • Lack of personalization of content (n=2)	• Accessibility and convenience of medical resources (n=10) Not identified • Not identified
Motivation			
Reflective motivation			
Belief about capabilities	The intertwined influence of beliefs, emotions, and motivation	• Low self-efficacy (n=9)	• Not identified
Belief about consequences	Intertwined influence of beliefs, emotions, and motivation	• Stigmatization of eHealth services (n=6)	• Not identified
Automatic motivation			
Reinforcement	Intertwined influence of beliefs, emotions, and motivation	• Not identified	• Training and mentoring (n=10) • Positive feedback mechanisms (n=8) • Aging-friendly design (n=8) • Providing clear operating instructions (n=2)
Emotion	Intertwined influence of beliefs, emotions, and motivation	• Mood disorders (n=11)	• Not identified

COM-B: Psychological Capability; TDF: Knowledge (Disease Perception and Personal Experience Building)

Patients' perceptions about their disease and personal experiences constitute the starting point for their involvement in digital health apps. Due to a lack of accurate understanding of emotional symptoms, some patients (n=13/25, 52%) have a hard time understanding their symptoms and fail to connect their condition with a disease that requires management. This cognitive bias directly influences their motivation to seek help.

XSL•FO
RenderX

> *I know I sometimes feel down and bored, but I don't know if this mood is depression, I don't think I can be depressed.* [P1]
>
> *Most of the time when I am in a bad mood, it is due to my back pain, sometimes my back pain is so bad that my whole body wants to die.* [P3]

This tendency to attribute depression to physical discomfort or life events diminishes their motivation to actively use self-management tools. Meanwhile, 32% (n=8/25) of the patients demonstrated a lack of awareness or skepticism regarding the value of the app itself, which further dampened their willingness to give it a try.

> *I don't know what these apps can do for me, it feels like just one more thing to worry about.* [P2]

However, when health threats become tangible, such as a noticeable decrease in physical function, patients' attitudes often shift. For instance, 28% (n=7/25) of the patients reported that concerns about diminished quality of life could translate into a strong motivation to improve, thereby making them more willing to try new technological approaches.

> *I can't move my feet right now and I'm very limited in what I can do, that's my biggest struggle and I want to learn some ways to improve my life.* [P10]

In addition, 52% (n=13/25) of patients noted that prior experience with technology served as a valuable source of confidence, effectively enhancing their acceptance of new technological tools.

> *I have no difficulty using these apps, I use a variety of software daily.* [P4]
>
> *I am quite good at operating my phone, I already have several software programs; more than enough.* [P19]

Thus, the initial willingness to use such tools is deeply rooted in how individuals perceive their own illness and the technological experience they possess. However, willingness is only the first step. When patients actually begin to use the tools, they immediately face practical challenges shaped by their own capabilities.

COM-B: Psychological Capability and Physical Capability; TDF: Memory, Attention, and Decision Processes and Skills (Dual Challenge of Cognitive Abilities and Physical Limitations)

Even with the willingness to use such tools, the cognitive and physical limitations of older patients with depression often create a "mismatch" with the functional design of apps, which constitutes a major obstacle in the adoption process. Approximately 64% (n=16/25) of patients reported that cognitive impairments, such as memory decline and attention deficits, directly challenged the basic abilities required to learn and operate new digital tools.

> *I'm a poor reader … and when I use the app, the interface is very easy to accidentally exit, I've been taught by volunteers many times before, but I still don't know how to operate it.* [P8]
>
> *The brain is not very good, looking at the screen is blurred, memory is bad, easy to forget; the cell phone is used only to make and answer calls. WeChat is not used.* [P21]

Meanwhile, 40% (n=10/25) of the patients mentioned that the physical symptoms of the disease itself, such as dizziness, visual impairment, and extreme fatigue, made seemingly simple actions like "looking at the screen" and "tapping to operate" unusually strenuous and could even exacerbate their discomfort.

> *I'm not well at the moment, I come to the nursing home to get well, and I feel that using these will increase the burden of my illness.* [P6]
>
> *I mainly look at the computer because I think the phone is too small for my eyes and it's hard to look at it, it's hard to look at it twice with my eyes.* [P23]

In contrast, several participants (n=14/25, 56%) noted that support from family members or peers helped them continue using digital tools despite physical or cognitive difficulties, as assistance with operation and emotional encouragement reduced frustration during initial attempts.

> *My eyes are blurry and my fingers aren't flexible—I almost gave up several times. But my daughter often sits down and guides me through it step by step. In those moments, my heart feels warm and I find the courage to try again.* [P25]

The above challenges that stem from the patients' own capabilities inevitably affect their perception of the convenience offered by digital tools. Although the process of operation is fraught with frustration, their attitudes toward technology integration also tend to diverge.

COM-B: Psychological Capability; TDF: Behavioral Regulation (Digital Technology Integration and Life Adaptation)

Faced with their own ability limitations, patients show a clear divergence in their responses to the integration of digital technology into health management. For instance, 36% (n=9/25) of patients noted that convenient features of apps, such as medication reminders and simple logging, effectively compensate for cognitive deficits, help establish routines, and thereby lead to positive experiences.

> *I used to forget to take my medication all the time, but now I have reminders to know when I should take my medication.* [P9]
>
> *It's easy and I'd like to have more access to it.* [P14]

Conversely, for 28% (n=7/25) of patients, these limitations reinforced their reliance on and preference for traditional methods, which they perceived as more reliable and comfortable.

> *I think it's better for older people to read books than to look at their cell phones, there is gold in books.* [P2]
>
> *I like to read newspapers, I've been reading them for more than 40 years. I'm used to it and I seldom use a cell phone.* [P6]

XSL•FO
RenderX

This variation in acceptance, rooted in individual capabilities, habits, and comfort levels, does not exist in isolation. It is profoundly shaped by the social environment in which patients live, where the attitudes of family members and peers play a crucial role.

COM-B: Social Opportunities; TDF: Social Influences (Access to and Utilization of Social Impact and Support Resources)

Patient attitudes toward and attempts to use digital health tools are deeply embedded in their social networks. Positive social influences can serve as a key factor in overcoming self-doubt about their abilities and motivating them to learn. Approximately 56% (n=14/25) of the patients mentioned that encouragement from family members and the sharing of successful experiences by peers provided important external validation and boosted their confidence.

> *My wife tried it and said it was good and kept encouraging me to try it too.* [P6]
>
> *Lao Zhuang shares his experience using the app every day and encourages us to all try it, we all have to learn from him.* [P5]

However, this social network can also become a constraint. Several participants (n=11/25, 44%) described that family overprotection not only limited their opportunities for hands-on practice but also made them increasingly doubt their own ability to use digital technologies independently, as family members frequently emphasized the difficulty of these tools.

> *My kids won't let me use it, they always tell me to leave it alone, that it's too much trouble.* [P8]
>
> *They say the techniques are too complicated and I'm sure I can't use them, but how do I know if I can use them if I haven't even tried them yet?* [P15]

Meanwhile, 16% (n=4/25) of patients mentioned that broader societal prejudices, particularly the stigma surrounding mental illness, could cause them to feel anxious and ashamed even when using health apps in private.

> *Depression is an illness that a lot of people don't understand, especially within nursing homes, they talk about you behind your back all the time.* [P12]

The power of social support or hindrance ultimately interacts with the material resources and environmental conditions that patients actually possess, jointly determining whether a technology can truly be made "usable."

COM-B: Physical Opportunity; TDF: Environmental Context and Resources (Resource and Environmental Constraints and Facilitation)

Social support can improve willingness; however, if essential material resources and a supportive environment are lacking, actual usage remains unachievable. Approximately 40% (n=10/25) of participants noted that a key advantage of digital health apps lies in their ability to reduce the time and costs associated with traditional health care access.

> *I have found that seeing a doctor through my phone and not having to run to the hospital and stand in line is a real time and money saver.* [P4]
>
> *Now that seeing a doctor and making an appointment with a doctor works on my phone, …* [P14]

However, for some patients, the inability to own a smartphone (n=3/25, 12%) or the lack of a stable internet connection (n=2/25, 8%) still constitutes a fundamental barrier that is difficult to overcome.

> *I don't have a smartphone, and these apps don't work at all.* [P2]
>
> *Mainly because of the lack of traffic, my cell phone is used to take and make calls, and I don't use the other features until I have WiFi at home.* [P24]

Even after overcoming these hardware barriers, 8% (n=2/25) of patients indicated that if the app content lacked relevance and failed to meet personalized needs, the users' motivation for continued use diminished rapidly. This reveals a gap between technological offerings and the actual needs of users.

> *It's always the same content and I don't want to use it anymore.* [P7]
>
> *There are a lot of features, but I always feel like something is missing and the advice given to me is not tailored to my individual situation, it would be great if it could be adapted to my needs.* [P15]

The interplay of limited resources and design barriers created major obstacles to sustained use. The experiences of the 3 participants without smartphones underscore that engagement in digital self-management may be constrained not only by individual willingness but also by access to basic devices and connectivity. Rather than expressing resistance, these participants described relying on family members or institutional services for health-related support, suggesting that digital self-management interventions should consider providing foundational access support and proxy-assisted pathways for those lacking personal smartphone access.

COM-B: Reflective Motivation, Automatic Motivation; TDF: Beliefs About Capability, Belief About Consequences, Reinforcement, and Emotion (The Intertwined Influence of Beliefs, Emotions, and Motivation)

The entire process of using digital health apps is continuously influenced by a complex interplay of internal beliefs, emotional states, and motivational levels of patients. Approximately 32% (n=8/25) of patients mentioned that positive user experiences, such as receiving immediate feedback and encouragement, significantly enhanced their sense of self-efficacy and willingness to continue usage.

> *Even the app is complimenting me on a job well done, which gives me confidence to keep using it.* [P13]

However, low self-efficacy and the sense of "meaninglessness" inherent in depression can fundamentally erode the internal motivation to engage with such tools.

XSL•FO
RenderX

> *I'm so old, I haven't even figured out how to use an old person's phone, how can I possibly use a smartphone.* [P2]
>
> *When I am in a bad mood, I always feel that there is no point in doing anything.* [P3]

Furthermore, 24% (n=6/25) of patients expressed deep concerns about privacy breaches and online fraud, reflecting a more profound sense of insecurity in the digital age. Such concerns may be amplified by related negative reports in society.

> *I don't dare to use it, there are too many scammers online, in case I lose all my money.* [P7]
>
> *Worried about information leakage, have to pay attention to the privacy aspect.* [P22]

Approximately 8% (n=2/25) of patients also clearly expressed that reducing the difficulty of use through aging-friendly design, clear instructions, and continuous training support is crucial for building positive beliefs and motivation.

> *Provide training, guide us how to use it a bit, how to deal with the difficulties we encounter with cell phone things. For example, I don't want this advertisement to appear, but deleting it is troublesome. Sometimes I can't do it, so I go down and ask young people to click it off.* [P16]

Discussion

Principal Findings

Guided by the COM-B model and the TDF, this qualitative study synthesized how capability, opportunity, and motivation interact to shape the engagement in digital self-management among older adults with depression. Rather than treating barriers and facilitators as isolated determinants, the combined application of the COM-B model and the TDF enabled a mechanistic understanding of why digital self-management engagement is particularly difficult to initiate and sustain in this population. The findings demonstrate that digital self-management engagement extends beyond issues of technology acceptance and is instead embedded within a complex behavioral process influenced by the perception of the disease, cognitive and functional capacity, social context, environmental resources, and depression-related emotional and motivational vulnerability. Although participants generally recognized the potential value of digital tools for mood regulation and daily management, misalignment between individual capability, contextual opportunity, and motivational readiness frequently limited meaningful and sustained engagement.

Capability: Illness Perception, Cognitive Functioning, and Functional Constraints

At the capability level, digital self-management engagement was shaped less by isolated deficits in technical knowledge and more by how participants understood depression and evaluated their own functional readiness for self-management. Many participants did not conceptualize depressive symptoms as requiring ongoing management, instead attributing emotional distress to aging or physical discomfort, which diminished the perceived relevance of digital self-management tools. This finding refines prior evidence by highlighting that deficits in the TDF "knowledge" domain are not merely informational but also involve inaccurate symptom attribution, which undermines the perceived relevance of digital self-management tools [65-69]. Cognitive decline, fatigue, and attentional limitations further reduced psychological and physical ability, amplifying the effort required to engage with digital tools. However, some participants benefited from simplified functions such as reminders or basic tracking, consistent with evidence on digital prompts facilitating self-management behaviors [70,71]. In contrast, others expressed anxiety and avoidance of digital tools, often rooted in a long-standing reliance on traditional information sources and low confidence in their ability to use technology. This pattern reflects deficits across the TDF domains of "skills" and "beliefs about capabilities" [72-76]. Importantly, our findings suggest that even age-friendly interface principles may be insufficient when depressive symptoms exacerbate cognitive load and reduce tolerance for frustration [39,77-80]. Together, these results indicate that digital self-management capability among older adults with depression should be understood as a dynamic combination of illness perception, cognitive capacity, and functional energy, rather than as a static set of skills.

Opportunity: Social Influences and Environmental Resources

Social and physical opportunity conditions substantially shaped digital self-management engagement. Consistent with existing literature, supportive involvement from family members and peers enhanced confidence and facilitated behavioral initiation [81-86]. However, this study extends prior findings by demonstrating that social support is not uniformly beneficial. Overprotective family behavior, described by participants as discouraging use or emphasizing technical difficulty, limited hands-on practice and fostered self-doubt regarding their own capabilities, thereby constraining sustained engagement. This nuance refines the TDF "social influences" domain by emphasizing that only autonomy-supportive social environments promote long-term behavioral maintenance [87-89]. With regard to the level of physical opportunity, digital self-management engagement remained constrained by limited access to devices, unstable internet connectivity, insufficient technical support, and financial burden [90-92]. Beyond individual resource limitations, participants' experiences pointed to broader structural gaps within the digital health ecosystem, which have yet to adequately accommodate the intersecting needs of aging and depression. These findings echo concerns that poorly supported digital interventions may inadvertently widen health inequities [93-95], underscoring the need for system-level adaptations rather than reliance on individual-level solutions.

Motivation: Beliefs, Emotions, and Reinforcement Mechanisms

Motivation emerged as a central mechanism linking capability and opportunity with digital self-management behavior. Core depressive features, such as apathy, emotional numbness, and reduced reward sensitivity, attenuated intrinsic motivation even when participants cognitively acknowledged potential benefits

XSL•FO
RenderX

[96]. Reflective motivation was strongly shaped by beliefs about usefulness, controllability of symptoms, and personal competence, with early negative experiences exerting a disproportionately suppressive effect on continued engagement [97,98]. This finding extends traditional technology acceptance models by highlighting how depressive symptomatology amplifies the motivational consequences of usability failure. Conversely, positive reinforcement, including structured guidance, supportive feedback, and emotionally validating design, was critical in restoring confidence and facilitating habit formation by directly targeting the TDF domains of "reinforcement" and "emotion" [99]. Importantly, motivation did not operate independently: opportunity enhancements such as initial assistance or training partially compensated for limited capability, whereas persistent opportunity barriers nullified motivation regardless of willingness. These interaction patterns help explain why single-component interventions often fail and support the use of the COM-B model as a system-level explanatory model rather than a checklist of determinants [100,101].

Taken together, these findings suggest that older adults with depression may be willing to try digital self-management tools but often face substantial barriers to maintaining engagement over time. Interventions should move beyond simply increasing access to technology and instead address the understanding of the disease, the cognitive burden, and the emotional and motivational vulnerabilities associated with depression. Guided by the COM-B model and TDF, strategies may include targeted psychoeducation to strengthen illness recognition, a simplified and personalized design to reduce cognitive load, structured onboarding and hands-on training to support early success experiences, and supportive feedback mechanisms to maintain confidence during symptom fluctuations. At the system level, embedding digital self-management support within primary care and community-based mental health services, together with basic device and connectivity support, can improve accessibility and reduce avoidable disengagement. Overall, aligning the intervention design with the lived realities of aging and depression can offer a feasible pathway to improve sustained use and long-term self-management outcomes.

Limitations

As a qualitative study using purposive sampling, this research did not aim to achieve statistical generalizability but rather sought to generate in-depth, contextually grounded insights into the facilitators and barriers shaping digital self-management among older adults with depression. The findings should be considered within the Chinese sociocultural and healthcare context, and while they may not be directly generalizable to all settings, they offer analytical insights that could be transferable to other contexts where aging, depression, and digital health intersect. Second, although this study systematically collected the perspectives and experiences of older patients with depression regarding mHealth management, it did not incorporate professional viewpoints, such as those of health care professionals. These perspectives could have provided valuable insights for the clinical implementation of intervention programs. Third, the interview duration in this study was set at 20 to 40 minutes, considering the cognitive load and attention characteristics of older adults with chronic conditions. Longer interviews could have resulted in fatigue and reduced attention, potentially affecting response quality. Although previous qualitative studies with similar older populations have generally used relatively short interview durations (approximately 15-35 minutes) [102-105], the relatively brief interview time in this study may have somewhat limited the depth and richness of the data, particularly when exploring complex experiences, behavioral factors, and decision-making processes, and may have resulted in some subtle or latent themes not being fully captured.

Conclusions

This study provides a theory-informed explanation of participation in digital self-management among older adults with depression by applying the COM-B model and the TDF as an integrated analytical lens. The findings highlight that sustained engagement depends on the alignment of capability, opportunity, and motivation rather than on technological availability or willingness. In particular, depression-related cognitive and emotional vulnerability interacts with age-related functional constraints and contextual barriers, helping to explain why engagement is difficult to maintain in real-world settings. These insights emphasize the importance of designing and implementing interventions that are feasible for older adults with depression, including supports that reduce cognitive burden, enable early success, and promote autonomy. At the policy level, strengthening community-based delivery pathways and ensuring easy access to basic digital resources may help prevent the widening of mental health disparities as digital care strategies continue to expand.

Acknowledgments

We would like to acknowledge the staff who worked on the study for their assistance in this research. No generative artificial intelligence tools were used in the preparation of this manuscript. All authors take full responsibility for the accuracy, originality, and integrity of the work.

Funding

This work was supported by the National Natural Science Foundation of China (grant 72104050) and the Fujian Province Science and Technology-Economy Integration Service Platform (grant FJKX-2024XRH04).

Data Availability

All data generated or analyzed during this study are included in this published article.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: HL, RL, ZZ
Data curation: ZZ
Formal analysis: XL, ZZ
Funding acquisition: RL
Methodology: RL, HL
Study design: RL
Supervision: HL, RL, ZC
Visualization: XL
Writing – original draft: XL, ZZ
Writing – review & editing: XL, ZZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Meaning saturation grid.
[DOCX File, 16 KB - aging_v9i1e79253_app1.docx]

Multimedia Appendix 2
Interview questions posed to participants based on the Theoretical Domains Framework (TDF) and the Capability, Opportunity, Motivation, and Behavior (COM-B) model.
[DOCX File, 14 KB - aging_v9i1e79253_app2.docx]

Multimedia Appendix 3
Theme to Capability, Opportunity, Motivation, and Behavior (COM-B) mapping with Theoretical Domains Framework (TDF).
[DOCX File, 26 KB - aging_v9i1e79253_app3.docx]

Multimedia Appendix 4
From raw interview statements to open codes and subthemes.
[DOCX File, 36 KB - aging_v9i1e79253_app4.docx]

Checklist 1
COREQ checklist.
[PDF File, 605 KB - aging_v9i1e79253_app5.pdf]

References

1. Solomon DH, Rudin RS. Digital health technologies: opportunities and challenges in rheumatology. Nat Rev Rheumatol 2020 Sep;16(9):525-535. [doi: 10.1038/s41584-020-0461-x] [Medline: 32709998]
2. Kuwabara A, Su S, Krauss J. Utilizing digital health technologies for patient education in lifestyle medicine. Am J Lifestyle Med 2020;14(2):137-142. [doi: 10.1177/1559827619892547] [Medline: 32231478]
3. Ronquillo Y, Meyers A, Korvek SJ. Digital health. In: StatPearls [Internet]: StatPearls Publishing; 2026. URL: https://www.ncbi.nlm.nih.gov/books/NBK470260/ [accessed 2026-03-20] [Medline: 29262125]
4. Hagerty BM, Bathish MA. Testing the relationship between a self-management intervention for recurrent depression and health outcomes. Appl Nurs Res 2018 Dec;44:76-81. [doi: 10.1016/j.apnr.2018.10.001] [Medline: 30389064]
5. Help and information. Mental Health UK. URL: https://mentalhealth-uk.org/help-and-information/ [accessed 2026-01-10]
6. Davidson L. Recovery, self management and the expert patient—changing the culture of mental health from a UK perspective. J Ment Health 2005 Feb;14(1):25-35. [doi: 10.1080/09638230500047968]
7. Bilsker D, Goldner EM, Anderson E. Supported self-management: a simple, effective way to improve depression care. Can J Psychiatry 2012 Apr;57(4):203-209. [doi: 10.1177/070674371205700402] [Medline: 22480584]
8. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11. [doi: 10.1002/wps.20083] [Medline: 24497236]
9. Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P. Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. J Med Internet Res 2010 Dec 19;12(5):e53. [doi: 10.2196/jmir.1436] [Medline: 21169166]
10. Lal S, Adair CE. E-mental health: a rapid review of the literature. Psychiatr Serv 2014 Jan 1;65(1):24-32. [doi: 10.1176/appi.ps.201300009] [Medline: 24081188]

XSL•FO
RenderX

11. Richards D, Viganó N. Online counseling: a narrative and critical review of the literature. J Clin Psychol 2013 Sep;69(9):994-1011. [doi: 10.1002/jclp.21974] [Medline: 23630010]
12. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336. [doi: 10.1002/wps.20673] [Medline: 31496095]
13. Mohr DC, Ho J, Duffecy J, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409. [doi: 10.1002/jclp.20659] [Medline: 20127795]
14. Deady M, Johnston D, Milne D, et al. Preliminary effectiveness of a smartphone app to reduce depressive symptoms in the workplace: feasibility and acceptability study. JMIR mHealth uHealth 2018 Dec 4;6(12):e11661. [doi: 10.2196/11661] [Medline: 30514694]
15. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22. [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
16. Drissi N, Ouhbi S, Janati Idrissi MA, Ghogho M. An analysis on self-management and treatment-related functionality and characteristics of highly rated anxiety apps. Int J Med Inform 2020 Sep;141:104243. [doi: 10.1016/j.ijmedinf.2020.104243] [Medline: 32768994]
17. Depressive disorder (depression). World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/depression [accessed 2025-03-19]
18. Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011 Sep;21(9):655-679. [doi: 10.1016/j.euroneuro.2011.07.018] [Medline: 21896369]
19. Sobocki P, Jönsson B, Angst J, Rehnberg C. Cost of depression in Europe. J Ment Health Policy Econ 2006 Jun;9(2):87-98. [Medline: 17007486]
20. Chan VKY, Leung MYM, Chan SSM, et al. Projecting the 10-year costs of care and mortality burden of depression until 2032: a Markov modelling study developed from real-world data. Lancet Reg Health West Pac 2024;45:101026. [doi: 10.1016/j.lanwpc.2024.101026] [Medline: 38352243]
21. Hu T, Zhao X, Wu M, et al. Prevalence of depression in older adults: a systematic review and meta-analysis. Psychiatry Res 2022 May;311:114511. [doi: 10.1016/j.psychres.2022.114511] [Medline: 35316691]
22. Cai H, Jin Y, Liu R, et al. Global prevalence of depression in older adults: a systematic review and meta-analysis of epidemiological surveys. Asian J Psychiatr 2023 Feb;80:103417. [doi: 10.1016/j.ajp.2022.103417] [Medline: 36587492]
23. Vyas CM, Okereke OI. Late-life depression: a narrative review on risk factors and prevention. Harv Rev Psychiatry 2020;28(2):72-99. [doi: 10.1097/HRP.0000000000000240] [Medline: 31977599]
24. Park M, Unützer J. Geriatric depression in primary care. Psychiatr Clin North Am 2011 Jun;34(2):469-487. [doi: 10.1016/j.psc.2011.02.009] [Medline: 21536169]
25. Greenberg PE, Fournier AA, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). J Clin Psychiatry 2015 Feb;76(2):155-162. [doi: 10.4088/JCP.14m09298] [Medline: 25742202]
26. König H, König HH, Konnopka A. The excess costs of depression: a systematic review and meta-analysis. Epidemiol Psychiatr Sci 2019 Apr 5;29:e30. [doi: 10.1017/S2045796019000180] [Medline: 30947759]
27. Huang Y, Wang Y, Wang H, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. Lancet Psychiatry 2019 Mar;6(3):211-224. [doi: 10.1016/S2215-0366(18)30511-X] [Medline: 30792114]
28. van Beljouw IM, Verhaak PF, Cuijpers P, van Marwijk HW, Penninx BW. The course of untreated anxiety and depression, and determinants of poor one-year outcome: a one-year cohort study. BMC Psychiatry 2010 Oct 20;10(1):86. [doi: 10.1186/1471-244X-10-86] [Medline: 20961414]
29. Donohue JM, Pincus HA. Reducing the societal burden of depression: a review of economic costs, quality of care and effects of treatment. Pharmacoeconomics 2007;25(1):7-24. [doi: 10.2165/00019053-200725010-00003] [Medline: 17192115]
30. Collins KA, Westra HA, Dozois DJA, Burns DD. Gaps in accessing treatment for anxiety and depression: challenges for the delivery of care. Clin Psychol Rev 2004 Sep;24(5):583-616. [doi: 10.1016/j.cpr.2004.06.001] [Medline: 15325746]
31. Chekroud AM, Foster D, Zheutlin AB, et al. Predicting barriers to treatment for depression in a U.S. national sample: a cross-sectional, proof-of-concept study. Psychiatr Serv 2018 Aug 1;69(8):927-934. [doi: 10.1176/appi.ps.201800094] [Medline: 29962307]
32. Polacsek M, Boardman GH, McCann TV. A theory on the components of depression self-management in older adults. Qual Health Res 2021 Jan;31(1):160-171. [doi: 10.1177/1049732320961060] [Medline: 32996403]
33. Sherdell L, Waugh CE, Gotlib IH. Anticipatory pleasure predicts motivation for reward in major depression. J Abnorm Psychol 2012 Feb;121(1):51-60. [doi: 10.1037/a0024945] [Medline: 21842963]
34. Chou YH, Lin C, Lee SH, Chang Chien YW, Cheng LC. Potential mobile health applications for improving the mental health of the elderly: a systematic review. Clin Interv Aging 2023;18:1523-1534. [doi: 10.2147/CIA.S410396] [Medline: 37727447]
35. Gould CE, Loup J, Kuhn E, et al. Technology use and preferences for mental health self-management interventions among older veterans. Int J Geriatr Psychiatry 2020 Mar;35(3):321-330. [doi: 10.1002/gps.5252] [Medline: 31854029]

XSL•FO
RenderX

36. Almulhem JA. Factors, barriers, and recommendations related to mobile health acceptance among the elderly in Saudi Arabia: a qualitative study. Healthcare (Basel) 2023 Nov 23;11(23):3024. [doi: 10.3390/healthcare11233024] [Medline: 38063592]
37. Josephine K, Josefine L, Philipp D, David E, Harald B. Internet- and mobile-based depression interventions for people with diagnosed depression: a systematic review and meta-analysis. J Affect Disord 2017 Dec 1;223:28-40. [doi: 10.1016/j.jad.2017.07.021] [Medline: 28715726]
38. Ahmad NA, Mat Ludin AF, Shahar S, Mohd Noah SA, Mohd Tohit N. Willingness, perceived barriers and motivators in adopting mobile applications for health-related interventions among older adults: a scoping review protocol. BMJ Open 2020 Mar 16;10(3):e033870. [doi: 10.1136/bmjopen-2019-033870] [Medline: 32184309]
39. Yin R, Rajappan D, Martinengo L, et al. Depression self-care apps' characteristics and applicability to older adults: systematic assessment. J Med Internet Res 2025 Feb 21;27:e56418. [doi: 10.2196/56418] [Medline: 39983112]
40. Posselt J, Lander J, Dierks ML. Health literacy promotion and digital interventions for depressive disorders. Health Lit Res Pract 2024 Oct;8(4):e236-e245. [doi: 10.3928/24748307-20240730-01] [Medline: 39642046]
41. Herber OR, Atkins L, Störk S, Wilm S. Enhancing self-care adherence in patients with heart failure: a study protocol for developing a theory-based behaviour change intervention using the COM-B behaviour model (ACHIEVE study). BMJ Open 2018 Sep 11;8(9):e025907. [doi: 10.1136/bmjopen-2018-025907] [Medline: 30206096]
42. Noar SM, Zimmerman RS. Health Behavior Theory and cumulative knowledge regarding health behaviors: are we moving in the right direction? Health Educ Res 2005 Jun;20(3):275-290. [doi: 10.1093/her/cyg113] [Medline: 15632099]
43. Buchanan H, Newton JT, Baker SR, Asimakopoulou K. Adopting the COM-B model and TDF framework in oral and dental research: a narrative review. Community Dent Oral Epidemiol 2021 Oct;49(5):385-393. [doi: 10.1111/cdoe.12677] [Medline: 34270110]
44. Michie S, Johnston M, Abraham C, et al. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb;14(1):26-33. [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
45. Taylor N, Lawton R, Conner M. Development and initial validation of the determinants of physical activity questionnaire. Int J Behav Nutr Phys Act 2013 Jun 11;10(1):74. [doi: 10.1186/1479-5868-10-74] [Medline: 23758912]
46. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
47. Janz NK, Becker MH. The health belief model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
48. Armitage CJ, Conner M. Efficacy of the theory of planned behaviour: a meta-analytic review. Br J Soc Psychol 2001 Dec;40(Pt 4):471-499. [doi: 10.1348/014466601164939] [Medline: 11795063]
49. Brug J, Conner M, Harré N, Kremers S, McKellar S, Whitelaw S. The Transtheoretical Model and stages of change: a critique: observations by five commentators on the paper by Adams, J. and White, M. (2004) why don't stage-based activity promotion interventions work? Health Educ Res 2005 Apr;20(2):244-258. [doi: 10.1093/her/cyh005] [Medline: 15618324]
50. Dehlholm-Lambertsen B, Maunsbach M. Qualitative methods in empirical health research. III. The individual in-depth interview [Article in Danish]. Nord Med 1997 Mar;112(3):94-98. [Medline: 9190568]
51. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
52. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
53. Chang SJ, Yang E, Ryu H, Kim HJ, Yoon JY. Cross-cultural adaptation and validation of the eHealth literacy scale in Korea. Korean J Adult Nurs 2018;30(5):504. [doi: 10.7475/kjan.2018.30.5.504]
54. Clohessy S, Kempton C, Ryan K, Grinbergs P, Elliott MT. Exploring older adults' perceptions of using digital health platforms for self-managing musculoskeletal health conditions: focus group study. JMIR Aging 2024 Aug 1;7:e55693. [doi: 10.2196/55693] [Medline: 39088803]
55. Money A, Hall A, Harris D, Eost-Telling C, McDermott J, Todd C. Barriers to and facilitators of older people's engagement with web-based services: qualitative study of adults aged >75 years. JMIR Aging 2024 Feb 28;7:e46522. [doi: 10.2196/46522] [Medline: 38416543]
56. Flannery C, McHugh S, Anaba AE, et al. Enablers and barriers to physical activity in overweight and obese pregnant women: an analysis informed by the theoretical domains framework and COM-B model. BMC Pregnancy Childbirth 2018 May 21;18(1):178. [doi: 10.1186/s12884-018-1816-z] [Medline: 29783933]
57. Timlin D, McCormack JM, Simpson EE. Using the COM-B model to identify barriers and facilitators towards adoption of a diet associated with cognitive function (MIND diet). Public Health Nutr 2021 May;24(7):1657-1670. [doi: 10.1017/S1368980020001445] [Medline: 32799963]
58. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
59. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]

XSL•FO
RenderX

60. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37. [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
61. Colorafi KJ, Evans B. Qualitative descriptive methods in health science research. HERD 2016 Jul;9(4):16-25. [doi: 10.1177/1937586715614171] [Medline: 26791375]
62. Sullivan-Bolyai S, Bova C. Qualitative Description: a "How-To" guide. : Graduate School of Nursing, University of Massachusetts Medical School; 2021 URL: https://repository.escholarship.umassmed.edu/server/api/core/bitstreams/75fe4c77-3447-4214-9286-ead665fc97e6/content [accessed 2026-03-30]
63. Olmos-Vega FM, Stalmeijer RE, Varpio L, Kahlke R. A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. Med Teach 2022 Apr 7;45(3):1-11. [doi: 10.1080/0142159X.2022.2057287] [Medline: 35389310]
64. Dodgson JE. Reflexivity in qualitative research. J Hum Lact 2019 May;35(2):220-222. [doi: 10.1177/0890334419830990] [Medline: 30849272]
65. Kim B, Jeong CH, Blood E, et al. Multilevel factors influencing eHealth adoption among older adults during the pandemic. Front Public Health 2025;13:1531173. [doi: 10.3389/fpubh.2025.1531173] [Medline: 40356828]
66. Pocklington C. Depression in older adults. Br J Med Pract 2017;10(1):a1007 [FREE Full text]
67. Frost R, Nair P, Aw S, et al. Supporting frail older people with depression and anxiety: a qualitative study. Aging Ment Health 2020 Dec;24(12):1977-1984. [doi: 10.1080/13607863.2019.1647132] [Medline: 31619050]
68. Rong J, Cheng P, Li D, Wang X, Zhao D. Global, regional, and national temporal trends in prevalence for depressive disorders in older adults, 1990-2019: an age-period-cohort analysis based on the global burden of disease study 2019. Ageing Res Rev 2024 Sep;100:102443. [doi: 10.1016/j.arr.2024.102443] [Medline: 39097004]
69. Mosleh SM, Almalik MM. Illness perception and adherence to healthy behaviour in Jordanian coronary heart disease patients. Eur J Cardiovasc Nurs 2016 Jun;15(4):223-230. [doi: 10.1177/1474515114563885] [Medline: 25505161]
70. Orr JA, King RJ. Mobile phone SMS messages can enhance healthy behaviour: a meta-analysis of randomised controlled trials. Health Psychol Rev 2015;9(4):397-416. [doi: 10.1080/17437199.2015.1022847] [Medline: 25739668]
71. Arensman R, Kloek C, Pisters M, Koppenaal T, Ostelo R, Veenhof C. Patient perspectives on using a smartphone app to support home-based exercise during physical therapy treatment: qualitative study. JMIR Hum Factors 2022 Sep 13;9(3):e35316. [doi: 10.2196/35316] [Medline: 36098993]
72. Wu Y, Wen J, Wang X, et al. Associations between e-health literacy and chronic disease self-management in older Chinese patients with chronic non-communicable diseases: a mediation analysis. BMC Public Health 2022 Nov 29;22(1):2226. [doi: 10.1186/s12889-022-14695-4] [Medline: 36447176]
73. Kim KA, Kim YJ, Choi M. Association of electronic health literacy with health-promoting behaviors in patients with type 2 diabetes: a cross-sectional study. Comput Inform Nurs 2018 Sep;36(9):438-447. [doi: 10.1097/CIN.0000000000000438] [Medline: 29742548]
74. Zhou M, Zhao L, Kong N, Campy KS, Qu S, Wang S. Factors influencing behavior intentions to telehealth by Chinese elderly: an extended TAM model. Int J Med Inform 2019 Jun;126:118-127. [doi: 10.1016/j.ijmedinf.2019.04.001] [Medline: 31029253]
75. Lin CY, Ganji M, Griffiths MD, Bravell ME, Broström A, Pakpour AH. Mediated effects of insomnia, psychological distress and medication adherence in the association of eHealth literacy and cardiac events among Iranian older patients with heart failure: a longitudinal study. Eur J Cardiovasc Nurs 2020 Feb;19(2):155-164. [doi: 10.1177/1474515119873648] [Medline: 31516036]
76. Castarlenas E, Sánchez-Rodríguez E, Roy R, et al. Electronic health literacy in individuals with chronic pain and its association with psychological function. Int J Environ Res Public Health 2021 Nov 28;18(23):12528. [doi: 10.3390/ijerph182312528] [Medline: 34886254]
77. Cullen R. Addressing the digital divide. Online Inf Rev 2001 Oct 1;25(5):311-320. [doi: 10.1108/14684520110410517]
78. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
79. Chatsatrian M, Kunde K, Bosompem J, et al. Usability evaluation of digital health applications for older people with depressive disorders: prospective observational study in a mixed methods design. JMIR Hum Factors 2025 Nov 28;12:e66271. [doi: 10.2196/66271] [Medline: 41328453]
80. Guidance on applying WCAG 2 to non-web information and communications technologies (WCAG2ICT). World Wide Web Consortium (W3C). URL: https://www.w3.org/TR/wcag2ict/ [accessed 2026-01-10]
81. Quirke E, Klymchuk V, Suvalo O, Bakolis I, Thornicroft G. Mental health stigma in Ukraine: cross-sectional survey. Glob Ment Health (Camb) 2021;8:e11. [doi: 10.1017/gmh.2021.9] [Medline: 34026241]
82. Yu H, Gao Y, Tong T, et al. Self-management behavior, associated factors and its relationship with social support and health literacy in patients with obstructive sleep apnea-hypopnea syndrome. BMC Pulm Med 2022 Sep 18;22(1):352. [doi: 10.1186/s12890-022-02153-1] [Medline: 36115966]
83. Zhang XN, Qiu C, Zheng YZ, Zang XY, Zhao Y. Self-management among elderly patients with hypertension and its association with individual and social environmental factors in China. J Cardiovasc Nurs 2020;35(1):45-53. [doi: 10.1097/JCN.0000000000000608] [Medline: 31373957]

XSL•FO
RenderX

84. Tang R, Luo D, Li B, Wang J, Li M. The role of family support in diabetes self-management among rural adult patients. J Clin Nurs 2023 Oct;32(19-20):7238-7246. [doi: 10.1111/jocn.16786] [Medline: 37340624]
85. Jo A, Ji Seo E, Son YJ. The roles of health literacy and social support in improving adherence to self-care behaviours among older adults with heart failure. Nurs Open 2020 Nov;7(6):2039-2046. [doi: 10.1002/nop2.599] [Medline: 33072389]
86. Costa ALS, Heitkemper MM, Alencar GP, Damiani LP, Silva RMD, Jarrett ME. Social support is a predictor of lower stress and higher quality of life and resilience in Brazilian patients with colorectal cancer. Cancer Nurs 2017;40(5):352-360. [doi: 10.1097/NCC.0000000000000388] [Medline: 27171810]
87. Tovar E, Rayens MK, Gokun Y, Clark M. Mediators of adherence among adults with comorbid diabetes and depression: the role of self-efficacy and social support. J Health Psychol 2015 Nov;20(11):1405-1415. [doi: 10.1177/1359105313512514] [Medline: 24357864]
88. Lorig KR, Ritter P, Stewart AL, et al. Chronic disease self-management program: 2-year health status and health care utilization outcomes. Med Care 2001 Nov;39(11):1217-1223. [doi: 10.1097/00005650-200111000-00008] [Medline: 11606875]
89. Volz M, Möbus J, Letsch C, Werheid K. The influence of early depressive symptoms, social support and decreasing self-efficacy on depression 6 months post-stroke. J Affect Disord 2016 Dec;206:252-255. [doi: 10.1016/j.jad.2016.07.041] [Medline: 27513631]
90. Shi Z, Du X, Li J, Hou R, Sun J, Marohabutr T. Factors influencing digital health literacy among older adults: a scoping review. Front Public Health 2024;12:1447747. [doi: 10.3389/fpubh.2024.1447747] [Medline: 39555039]
91. Xie B, Charness N, Fingerman K, Kaye J, Kim MT, Khurshid A. When going digital becomes a necessity: ensuring older adults' needs for information, services, and social inclusion during COVID-19. J Aging Soc Policy 2020;32(4-5):460-470. [doi: 10.1080/08959420.2020.1771237] [Medline: 32507061]
92. Anderson M, Perrin A. Tech adoption climbs among older adults. : Pew Research Center; 2017 URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2017/05/PI_2017.05.17_Older-Americans-Tech_FINAL.pdf [accessed 2025-01-25]
93. Cornejo Müller A, Wachtler B, Lampert T. Digital Divide – Soziale Unterschiede in der Nutzung digitaler Gesundheitsangebote [Article in German]. Bundesgesundheitsbl 2020 Feb;63(2):185-191. [doi: 10.1007/s00103-019-03081-y]
94. Hale TM, Chou WY, Cotten SR. eHealth Current Evidence, Promises, Perils and Future Directions: Emerald Publishing Limited; 2018. [doi: 10.1108/S2050-2060201815]
95. Weiss D, Rydland HT, Øversveen E, Jensen MR, Solhaug S, Krokstad S. Innovative technologies and social inequalities in health: a scoping review of the literature. PLoS One 2018;13(4):e0195447. [doi: 10.1371/journal.pone.0195447] [Medline: 29614114]
96. Ma R, Wang Y, Wang XQ, Yu K, Zhang CC, Zhou YQ. Analysis of hindering and facilitating factors of help-seeking behavior in schizophrenia based on COM-B model: a descriptive qualitative study. BMC Psychiatry 2023 Oct 23;23(1):770. [doi: 10.1186/s12888-023-05226-5] [Medline: 37867190]
97. Tsai JM, Cheng MJ, Tsai HH, Hung SW, Chen YL. Acceptance and resistance of telehealth: the perspective of dual-factor concepts in technology adoption. Int J Inf Manage 2019 Dec;49:34-44. [doi: 10.1016/j.ijinfomgt.2019.03.003]
98. Knapova L, Klocek A, Elavsky S. The role of psychological factors in older adults' readiness to use eHealth technology: cross-sectional questionnaire study. J Med Internet Res 2020 May 28;22(5):e14670. [doi: 10.2196/14670] [Medline: 32348251]
99. Taylor SE, Stanton AL. Coping resources, coping processes, and mental health. Annu Rev Clin Psychol 2007;3:377-401. [doi: 10.1146/annurev.clinpsy.3.022806.091520] [Medline: 17716061]
100. Xia J, Merinder LB, Belgamwar MR. Psychoeducation for schizophrenia. Cochrane Database Syst Rev 2011 Jun 15;2011(6):CD002831. [doi: 10.1002/14651858.CD002831.pub2] [Medline: 21678337]
101. Lee WH, Lee WK. Cognitive rehabilitation for patients with schizophrenia in Korea. Asian J Psychiatr 2017 Feb;25:109-117. [doi: 10.1016/j.ajp.2016.10.010] [Medline: 28262129]
102. Qiu YF, Liu JL, Zeng LJ, et al. Older adults' perceptions and needs regarding mental health WeChat applet: a qualitative study. BMC Psychiatry 2025 Sep 30;25(1):889. [doi: 10.1186/s12888-025-07339-5] [Medline: 41029253]
103. Chen S, Yang K, Ko A, Giovannucci E, Stults-Kolehmainen M, Yang L. Facilitators and barriers of reducing sedentary behavior in sedentary and non-sedentary older adults: a descriptive qualitative study based on the COM-B model and TDF. BMC Public Health 2025 Jul 16;25(1):2472. [doi: 10.1186/s12889-025-23613-3] [Medline: 40670951]
104. Shen A, Wu P, Qiang W, et al. Breast cancer survivors' experiences of barriers and facilitators to lymphedema self-management behaviors: a theory-based qualitative study. J Cancer Surviv 2025 Apr;19(2):642-658. [doi: 10.1007/s11764-023-01497-9] [Medline: 37971555]
105. Schöne C, Fuchs TI, Kiselev J, et al. Facilitators and barriers to participation in prehabilitation prior to orthopaedic elective surgery—a qualitative study with elderly (pre-)frail patients. BMC Geriatr 2025 Nov 4;25(1):845. [doi: 10.1186/s12877-025-06592-3] [Medline: 41188708]

XSL•FO
RenderX

Abbreviations

COM-B: Capability, Opportunity, Motivation, and Behavior
COREQ: Consolidated Criteria for Reporting Qualitative Research
mHealth: mobile health
TDF: Theoretical Domains Framework

Edited by H LaMonica; submitted 21.Jun.2025; peer-reviewed by V Raj, X Zhang, X Dong; revised version received 03.Mar.2026; accepted 04.Mar.2026; published 10.Apr.2026.

Please cite as:
Zhu Z, Lin X, Chen Z, Li H, Lin R
Facilitators and Barriers to Digital Self-Management in Older Adults With Depression: COM-B and Theoretical Domain Framework Qualitative Study
JMIR Aging 2026;9:e79253
URL: https://aging.jmir.org/2026/1/e79253
doi:10.2196/79253

© Ziping Zhu, Xinying Lin, Zhi Chen, Hong Li, Rong Lin. Originally published in JMIR Aging (https://aging.jmir.org), 10.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Response Time Dynamics From Noncognitive Ordinal Ecological Momentary Assessment as a Proxy for Symptom Change in Geriatric Depression: Longitudinal Observational Study

Jooho Lee[1], MA; Jeehang Lee[1,2], PhD; Sehwan Park[1], MA; Gangho Do[1], BS; Jihye Noh[3], MDM; Sangjoon Moon[3], MS; Kyungmi Chung[3,4,5], PhD; Sang Joon Son[6*], MD, PhD; Jin Young Park[3,4,5*], MD, PhD

[1]Medical Research Team, Digital Medic Co., Ltd., Seoul, Republic of Korea

[2]Department of Human-Centered Artificial Intelligence, College of Intelligence Information Engineering, Sangmyung University, Seoul, Republic of Korea

[3]Center for Digital Health, Yongin Severance Hospital, Yonsei University Health System, Yongin, Republic of Korea

[4]Department of Psychiatry, Yongin Severance Hospital, Yonsei University College of Medicine, 363, Dongbaekjukjeon-daero, Giheung-gu, Yongin, Republic of Korea

[5]Institute of Behavioral Science in Medicine, Yonsei University College of Medicine, Yongin, Republic of Korea

[6]Department of Psychiatry, Ajou University School of Medicine, Suwon, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Jin Young Park, MD, PhD
Center for Digital Health, Yongin Severance Hospital, Yonsei University Health System, Yongin, Republic of Korea

Abstract

Background: Depressive symptoms in older adults are amplified by social isolation and limited access to clinic-based mental health care. Ecological momentary assessment (EMA) enables remote self-monitoring and unobtrusively captures response times (RTs), which may serve as indicators of psychomotor and cognitive functioning.

Objective: This study investigated the use of EMA-based RT dynamics for predicting symptom change and profiling potential responders for repeated self-monitoring in late-life depression.

Methods: Forty-nine community-dwelling adults aged 65 years or older (mean age 70.7, SD 5.8 years; female: 35; male: 14) with a history of major depressive disorder received case management incorporating daily EMA. Participants provided self-reports of mood, appetite, sleep quality, and general well-being. Preassessment and postassessment included the 15-item Short Geriatric Depression Scale (GDS-15), the Center for Epidemiologic Studies Depression Scale-Revised (CESD-R), the 9-item Patient Health Questionnaire, and the Beck Anxiety Inventory. RTs were cleaned with an asymmetric IQR rule, z standardized within-person × response level, and modeled with exponential decay curves over successive EMA trials. The efficacy of EMA-adjunctive care was evaluated using pre-post comparisons of symptom scales. We then examined associations between RT-derived features and symptom change using correlational analyses. Finally, Bayesian multilevel modeling was applied to assess the clinical relevance of RT dynamics, including group differences in adaptation patterns.

Results: Older adults at risk for depression showed significant symptom reductions over the 4-week EMA-adjunctive care period across all 4 psychological scales (CESD-R: mean Δ 11.5; rank-biserial r=0.78; GDS-15: mean Δ 2.14, Cohen d=0.76), alongside high EMA adherence (>90%). In correlational analyses, descriptive EMA score metrics and raw RTs showed modest, symptom-specific associations with symptom change (ΔCESD-R: $|r|\approx0.29$; Δ9-item Patient Health Questionnaire: $|r|\approx0.32$; ΔBeck Anxiety Inventory: $|r|\approx0.35$) but were not significantly related to change in geriatric depression (ΔGDS-15: $|r|\approx0.24$). In contrast, exponential-decay model parameters derived from standardized RT were significantly associated with geriatric depressive symptom change (Δ GDS-15), with the strongest effects observed for the feeling item (eg, decay rate θ_b: $r=-0.398$, asymptote θ_c: $r=-0.321$). Bayesian multilevel modeling further indicated that EMA-adjunctive care responders showed faster RT adaptation than nonresponders (median decay-rate ratio≈4.9, 95% credible interval 1.44-14.31), whereas differences in postadaptation RT levels were smaller and uncertain (median postadaptation RT ratio≈1.25, 95% credible interval 0.95-1.58). Sensitivity analyses showed consistent decay-rate effects across alternative specifications.

Conclusions: Dynamic characteristics of EMA-based RTs emerged as a sensitive proxy for monitoring changes in depressive symptoms among older adults at risk. These findings highlight the potential use of RTs as digital biomarkers derived from brief, low-burden EMA self-monitoring, supporting the development of scalable and personalized mental health interventions for geriatric populations.

XSL•FO
RenderX

(JMIR Aging 2026;9:e83891) doi:10.2196/83891

KEYWORDS

ecological momentary assessment; response time; digital biomarkers; geriatric depression; Bayesian multilevel modeling; self-monitoring; mobile health

Introduction

Modern Hazards of Late-Life Depression

Population aging has brought renewed attention to late-life depression. Growth in single-person households, shrinking family networks, and weaker community ties can intensify the risk of depression in older adults, particularly in highly urbanized settings, such as South Korea [1,2]. These shifts also strain clinic-centered care models, which are often inaccessible to older adults living alone and facing mobility barriers, underscoring the need for remote, daily-life–responsive geriatric mental health strategies.

Ecological Momentary Assessment: From Monitoring to Care

Ecological momentary assessment (EMA) involves the real-time, in situ measurement of behaviors, emotions, and thoughts in everyday contexts [3,4]. By capturing experiences as they occur, EMA reduces recall bias and is well suited to symptoms that fluctuate with context [4]. Advances in mobile and sensor technologies have enabled unobtrusive, scalable EMA deployment, positioning it as a potential infrastructure for practical routine mental health support methods [5,6]. Repeated self-monitoring via EMA may also confer therapeutic benefits, particularly when combined with standard treatments [7-11]; however, effects appear heterogeneous, with reduced responsiveness among individuals with more severe baseline symptoms, highlighting the need to identify who benefits most from EMA-based approaches [12].

EMA Response Times: Linking Cognition and Depression in Older Adults

Recent EMA research has moved beyond self-reported content to "side data" such as response time (RT), which may index underlying cognitive processes. Chung et al [13] reported a nonlinear (inverted U-shaped) association between EMA RT and depression severity, suggesting that both unusually fast and slow responses may reflect dysfunction (eg, impulsivity or cognitive slowing). Hernandez [14] further showed that EMA RTs correlate with Symbol Search performance, indicating that RTs capture general processing capacity even in noncognitive EMA tasks.

EMA-derived RTs may therefore help identify individuals vulnerable to depressive symptoms. Consistent with *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*) descriptions of major depressive disorder, which include observable psychomotor retardation [15], Hernandez et al [16] applied drift diffusion modeling to binarized EMA responses and found that RT-derived parameters (eg, drift rate and boundary separation) systematically tracked traits such as neuroticism and depressive symptoms [16]. Together, these findings position EMA RT as a candidate marker linking cognitive efficiency with affective symptomatology.

This approach may be particularly relevant in geriatric populations, where aging and depression jointly exacerbate psychomotor slowing [17-19]. Older adults with major depressive disorder (MDD) show prolonged initiation and movement times relative to controls, implicating combined executive and motor delays [17], and reviews report especially large effects for psychomotor slowing in late-life samples [20]. Depressed older adults also appear less adaptable under repeated cognitive demands: greater symptom severity has been associated with slower reaction-time learning curves and reduced improvement across training blocks, with sluggish adaptation predicting later functional decline [21,22]. Some interventions may even backfire when effort exceeds coping capacity; for instance, one randomized controlled trial reported worsening depressive symptoms following speed-of-processing training [23]. Collectively, these results motivate EMA-based RT metrics as scalable markers for monitoring cognitive-motor slowing and adaptation in late-life depression.

Challenges in Utilizing Ordinal-Scale, Noncognitive EMA RTs

Despite its potential usage, analyzing RTs from ordinal-scale EMA items presents distinct hurdles. While ordinal items, such as 5- or 7-point Likert scales, offer richer granularity than binary formats, their integration with RT data poses additional analytical considerations and requires further preprocessing steps.

One complication arises from emotional clarity. As shown in Hernandez et al [16], individuals with clearer emotional awareness due to intense emotions tend to respond more quickly to EMA questionnaires. Thus, trait-level factors, such as baseline emotional reactivity, as well as situational factors, including bereavement, acute illness, and exposure to daily stressors, may affect the RT of EMA reports. This can mask or confound the expected slowing associated with cognitive impairment, making RT a less straightforward marker when emotional clarity varies across individuals and between occasions.

A further complication lies in interface variability. Mobile EMA apps often use different visual formats (eg, sliders, buttons, or scales), which can unintentionally influence how quickly users respond. For example, a response option that is most proximal to a frequently used finger can lead to faster responses. These app-specific differences, such as user interface composition, introduce another layer of noise that must also be accounted for when analyzing ordinal-scaled EMA-based RT.

Research Purpose

In light of these issues, this study pursues the following objectives. First, we tested whether short-term EMA use (≤4 wk) alongside usual case management is associated with

XSL•FO
RenderX

symptom improvement among older adults at elevated risk for depression, evaluating EMA as a potential adjunctive intervention rather than a monitoring-only tool. Second, we examined how multiple EMA-derived features (eg, symptom ratings, response speed, and RT adaptation) relate to symptom change over time, with a focus on RT dynamics as a behavioral signature of improvement. We also introduce a scalable analytic approach that reduces key ecological confounds in RT (eg, momentary affect intensity and interface-related heterogeneity) and is broadly applicable across EMA apps and Likert-type scales. Third, we evaluated the clinical relevance of RT adaptation by using a Bayesian multilevel model to compare adaptive patterns between participants who benefited from EMA-adjunctive care and those who did not. With a modest sample size (n=50), partial pooling in a multilevel framework improves estimate stability while accommodating missing and unbalanced EMA observations [24]. Additionally, Bayesian posterior inference provides interpretable uncertainty for group differences in adaptation parameters, thereby quantifying the evidence for between-group differences in adaptive patterns [24-26]. This research ultimately supports the objective, remote stratification of individuals who are most likely to benefit from EMA-enabled care.

Methods

Eligibility Criteria

Participants eligible for this study were aged 65 years or older, regardless of gender, and had expressed their willingness to participate through recruitment notices posted at the community-based Suwon Geriatric Mental Health and Welfare Center in Gyeonggi-do, Republic of Korea. Individuals who met the following eligibility criteria were included: (1) participants must have had a prior diagnosis of depressive disorder by a psychiatrist, with no change in medication for the last 3 months and without symptom aggravation; (2) absence of medical or neurological disorders such as dementia, stroke, or Parkinson disease that may affect the research; and (3) current recipients of case management services at the Suwon Geriatric Mental Health and Welfare Center.

Ethical Considerations

All participants who met these criteria provided written informed consent prior to enrollment and completed demographic information and preassessment questionnaires. Those who completed these surveys were then given detailed guidance and support for installing and using the app, with additional assistance provided to those less familiar with digital tools. Given the specific population's limited experience with digital devices, only individuals who demonstrated the ability to follow the mobile app (BIG4+) usage guidelines independently were enrolled. Participation was entirely voluntary, and participants were free to withdraw at any time without any disadvantage. All procedures were approved by the Public Institutional Bioethics Committee designated by the Ministry of Health and Welfare of the Korean Government (MOHW; P01-202405-01-019) and were supported by Digital Medic Inc. Written informed consent was obtained from all participants. Participant privacy was protected by analyzing all data in deidentified form and restricting access to authorized research staff only. Electronic data were stored in password-protected encrypted files, and physical records were kept in locked storage. Participants received KRW 100,000 (approximately US $67.50) for completing the full study protocol. For early withdrawal, compensation was provided on a prorated basis according to the week of completion.

Research Design

The study used an observational, single-arm design. Participants were community-dwelling older adults enrolled through the Suwon Geriatric Mental Health and Welfare Center.

Case management was provided throughout the study, consisting of 1 monthly home visit and 1 to 2 interim telephone contacts, during which case managers monitored participants' daily living conditions and mental health status. In addition to this standard case management, participants were asked to use the Big4+ mobile app as an adjunctive care tool over a 4-week period. The Big4+ app, developed specifically for research purposes, was available on both Android (AOS) and iOS platforms and collected daily self-reports on mood, appetite, sleep quality, and overall well-being using a 7-point Likert scale (1=minimum to 7=maximum).

Each participant visited the center twice: once prior to initiating app usage (baseline assessment, V1) and once at the end of the 4-week period (postassessment, V2). During both visits, participants completed a battery of self-report questionnaires administered in paper format by case managers, with each session lasting approximately 30 minutes.

Standardized psychological scales administered to participants included the Korean version of the Center for Epidemiologic Studies Depression Scale-Revised (CESD-R) [27-29], the Korean version of the 15-item Short Geriatric Depression Scale (GDS-15) [30,31], the 9-item Patient Health Questionnaire-9 (PHQ-9) [32,33], and the Beck Anxiety Inventory (BAI) [34,35].

Data Preprocessing

Handling RT Outliers

To the best of our knowledge, there is no established gold standard for identifying outliers in RTs associated with noncognitive and ordinal EMA items. Furthermore, the population of this study (older adults at high risk for depression) was highly specific, making it necessary to define acceptable RTs empirically based on the distribution within the dataset. We adopted the IQR rule, where the IQR is calculated as the difference between the third quartile and the first quartile. Outlier boundaries were defined as specific multiples of the IQR from Q1 and Q3. Given that RT distributions were heavily right-skewed, we applied asymmetric boundaries across the dataset: a lower bound of $Q1 - 0.5 \times IQR$ and an upper bound of $Q3 + 1.5 \times IQR$. These boundaries were calculated separately for each EMA item (feeling, appetite, feeling of sleep, and general evaluation). Trials with RTs falling outside these bounds were excluded from the analysis. The details of the predefined outlier thresholds and the decision-making process are provided in Section I in Multimedia Appendix 1.

XSL•FO
RenderX

Standardization of RTs

EMA RTs are noisy and can be confounded by momentary affect (eg, faster responses during more intense or clearly defined states) and device or app characteristics [16]. To improve comparability, we standardized RTs within individuals and response options, expressing each latency relative to that participant's typical speed at a given Likert level.

For individual *i*, response option *j* (1-7), and trial *k*, standardized RT (Z-RT) was as follows:

$$Zijk=RTijk-\overline{RT}ij\sigma ij,$$

where RT_{ij} and σ_{ij} are the mean and SD of RT for participant *i* at option *j*. This z-transformation reduces between-context and between-person variability, supporting more reliable downstream analyses.

Extracting EMA-Based Digital Features

Descriptive features were extracted from 4 weeks of EMA data collected from each participant. For each EMA item, we computed the mean, SD, minimum, maximum, and interquartile values (25th, 50th, and 75th percentiles) of the reported scores to characterize overall mood patterns throughout the study period. In addition, identical descriptive statistics of EMA-based RTs were derived as indicators of cognitive processing speed and consistency.

Beyond overall response latency, we quantified how participants adapted to repeated EMA self-reflection over time. Such "practice effects" (ie, progressive reductions in RT with repeated exposure) are widely documented in cognitive tasks and are relevant to EMA given evidence that even noncognitive RTs reflect cognitive traits [16]. Although practice effects are often modeled with power-law decay [36], exponential decay may better capture adaptation in older adults, who typically show slower motor and cognitive adjustment [37,38]. Accordingly, we fit an exponential decay curve to each participant's longitudinal RT series (both raw RT and standardized RT):

$$ft=\theta a\cdot\exp-\theta b\cdot t+\theta c,$$

where *t* indexes time or trials. Parameters were estimated via nonlinear least squares (Levenberg-Marquardt; scipy.optimize.curve_fit; max 10,000 iterations). The parameter θ_a (amplitude) captures the initial elevation in RT above the asymptote (ie, $f(0)–\theta_c$), reflecting the potential magnitude of improvement across repeated trials. The parameter θ_b (decay rate) indexes the speed of adaptation, indicating how rapidly an individual becomes accustomed to repeated self-monitoring. The parameter θ_c (asymptote or postadaptation RT) represents the asymptotic minimum RT, interpreted as the stabilized latency after maximal adaptation within the observation window. The model was fit to both raw and standardized RTs, with parameter bounds set to match the scale and range of each outcome.

In this stage, exponential decay parameters were estimated separately for each participant (ie, without partial pooling) to serve as person-level features that can be computed in naturalistic settings where no external reference data are available. This individual-fit approach also facilitates a fair comparison with other EMA-derived features (eg, self-report scores and RT descriptives), which were likewise calculated at the individual level without borrowing strength from group-level information. Technical details regarding the fitting of the exponential-decay model are provided in Section II in Multimedia Appendix 1.

Defining Symptom Improvement

A key step prior to the analyses was to select an appropriate metric for symptom change. Although treatment response is often operationalized using simple change scores (post-pre), this approach has been criticized for its dependence on baseline severity and the resulting negative correlation with baseline values [39-41]. Percentage change scores adjust for baseline differences and are widely used in other clinical domains [42,43] but can be difficult to interpret and may reduce statistical power due to skewed distributions [40,41]. Regression-based alternatives (eg, analysis of covariance or residualized change) can control for baseline levels yet may yield biased estimates when baseline severity is related to treatment effects, which is common in observational settings [40,41]. Given these considerations and prior methodological work supporting the robustness of simple change scores in nonrandomized contexts [40,44,45], we used simple change score Δ(post-pre) as the primary symptom improvement outcome.

Responder Classification (Minimally Detectable Change Criteria)

Participants were classified as "responders" or "nonresponders" to EMA-adjunctive care based on minimally detectable change (MDC) thresholds derived from the GDS-15. A recent meta-analysis by González-Domínguez et al [46], which included 5876 older adults across 21 studies (mean age 76 y), reported that a score reduction of 3.81 points on the GDS-15 corresponds to the MDC threshold, defined as the level of change with only a 5% probability of being attributable to measurement error (95% CI 3.59-4.04) [47]. Based on these findings, an MDC threshold of –3.81 points was applied, reducing the arbitrariness of classification and accounting for measurement error.

Analysis

Explorative Effects of EMA-Based Self-Monitoring

Explorative analysis was conducted to assess the benefits of EMA-adjunctive care in older populations. Given the single-arm design, the effectiveness of EMA-based self-monitoring in geriatric depression was assessed by comparing symptom scores collected at baseline (V1) and postintervention (V2) using paired statistical tests. The choice of test was determined based on the normality of the difference scores, as assessed by the Shapiro-Wilk test. When the assumption of normality was satisfied, a paired-samples *t* test was applied. In cases where the normality assumption was violated, the nonparametric Wilcoxon signed-rank test was used instead. All paired tests in this analysis were 1-tailed.

Association Between EMA-Derived Features and Symptom Improvement

We examined whether EMA-derived features, particularly RT dynamics, track symptom improvement using correlational

XSL•FO
RenderX

analyses. We computed Spearman correlations between EMA-derived features and Δ on the GDS-15 and assessed specificity using parallel analyses for the CESD-R, PHQ-9, and BAI.

Clinical Relevance of EMA Trajectories for Symptom Improvement

To examine the clinical relevance of EMA-based RT dynamics, we fitted a moderated Bayesian multilevel exponential-decay model to test whether adaptive patterns of EMA-based RTs differed by the MDC group. The model equation is demonstrated as a fully randomized model. For detailed model specifications and rationales, see Section III in Multimedia Appendix 1.

Observation Model

For each observation n=1, ..., N_{obs},

$$RTn \sim LogNormal(\eta n, \sigma log) \Leftrightarrow log(RTn) \sim N(\eta n, \sigma log2),$$

we assumed each observation follows a log-normal distribution, ensuring positive predictions and accommodating the right-skew typical in RT data.

Mean Structure (RT Adaptation)

Expected log-RT followed a subject-specific exponential decay over normalized time $t \in [0, 1]$:

$$\eta n = \theta ai \cdot \exp(-\theta bi \cdot t + \theta ci,r),$$

where *i* indexes subjects and $r \in 1, ..., R$ indexes response options (Likert-7 scale).

Multilevel Effects and MDC Moderation

Amplitude and decay rate were modeled on the log scale (log-linked) for positive constraints and allowed to vary by subject, with MDC group moderation:

$$\log(\theta a) = \beta 0a + \beta 1a \cdot g(i) + \sigma a \cdot z(a) \quad z(a) \sim N(0,1)$$
$$\log(\theta b) = \beta 0b + \beta 1b \cdot g(i) + \sigma b \cdot z(b) \quad z(b) \sim N(0,1)$$

Baseline log-RT or postadaptation RT ($^{(\,,\,)}$) was decomposed into group, subject, and response-option components:

$$\theta c(i,r) = (\beta 0c + \beta 1c \cdot g(i)) \text{ Group-Moderated Baseline} + uc(i) \text{ Subject Intercept} + \alpha c(r) \text{ Option Main Effect} + c(i,r) \text{ Cell Residual}$$

with $uc(i) \sim N(0, \sigma subj2)$, $c(i,r) \sim N(0, \sigma cell2)$, and $\sum r=1R \alpha c(r)=0$ for identifiability. The decomposition of baseline log-RT or postadaptation RT aligns with the framework advocating for the within-person and within-response option standardization of RTs. By explicitly modeling these components, the model accounts for response-option biases in observed latency while simultaneously partitioning subject-specific baseline variance.

Priors and Inference

We used weakly informative priors $\beta 0* \sim N(\mu emp, 32)$, $\beta 1* \sim N(0,32)$, $\alpha c(r) \sim N(0,12)$, σ_a, σ_b, σ_{subj}, $\sigma_{cell} \sim HalfNormal(1.5)$, and $\sigma_{log} \sim HalfNormal(1)$. Models were fit in PyMC using NUTS (4 chains; 1000 warm-up + 1000 draws per chain; target acceptance = 0.95) [48].

To determine the appropriate degree of hierarchical structure in EMA-based RT trajectories, we compared a set of nested Bayesian exponential-decay models that differed only in their random-effects specification. Specifically, we evaluated (1) a nonhierarchical model with fixed parameters across participants, (2) partially hierarchical variants allowing subject-level variation in selected parameters (eg, random θ_b only; random θ_a and θ_b; random θ_b and θcir), and (3) a full hierarchical model with subject-level variation in key components. Models were compared using Pareto-smoothed importance sampling leave-one-out cross-validation (PSIS-LOO), summarizing out-of-sample predictive performance via the expected log pointwise predictive density leave-one-out cross-validation (ELPD-LOO). To isolate the contribution of random effects, MDC-group moderation terms were omitted during this model-structure comparison. Based on the best-performing random-effects structure (selected model; random θ_b and θ_c, with θ_a fixed across subjects), we then tested clinical relevance by extending the selected model to include MDC group moderation, allowing these parameters to vary systematically between responders and nonresponders.

Group differences were quantified using Bayesian posterior summaries rather than frequentist hypothesis tests. For log-linked parameters (eg, θ_a, θ_b), effects were expressed as responder to nonresponder multiplicative ratios exp (β1); for baseline parameters on the log-RT scale (eg, θ_c), group effects were additive on the log scale and additionally converted to RT ratios via exponentiation. We report the posterior median of the responder-to-nonresponder ratio, computed draw-by-draw from the posterior samples, along with the 95% credible interval (CrI) for this ratio. A ratio of 1 indicates no difference between the groups on the original scale; ratios greater than 1 indicate larger parameter values in responders (eg, a ratio of 1.30 corresponds to ~30% higher values), whereas ratios less than 1 indicate smaller values in responders relative to nonresponders. The 95% CrI represents the range containing 95% of the posterior probability for the quantity (given the specified model and priors). We also report the posterior probability of an increase, $P(\text{Ratio}_{\text{Responder/Nonresponder}} > 1|\text{data})$, which quantifies the evidence that responders have larger parameter values than nonresponders.

Models were fit separately for each EMA item and additionally for an averaged RT outcome across items; for the averaged outcome, continuous scores were discretized into 7 bins to align with the response-option structure used in item-level models. Finally, we conducted sensitivity analyses by refitting the model under alternative weakly informative priors (Student *t*) and a more complex plausible random-effects specification and verified that key group effect conclusions were robust across these reasonable modeling choices.

Results

Participants

A total of 50 older adults aged 65 years or older were recruited for the study (mean age 70.6, SD 5.8 y), of whom 72% (n=35) were female. Among the participants, 49 had a documented history of MDD, and 1 had a history of bipolar disorder. To ensure consistency and avoid confounding effects arising from the natural mood fluctuations associated with bipolar disorder, only participants with MDD were included in all statistical

XSL•FO
RenderX

analyses (n=49). Baseline psychological assessments indicated clinically elevated symptoms across multiple domains. Preassessment scores were as follows: GDS-15 mean=9.37 (SD 4.42), CESD-R mean=33.94 (SD 18.24), PHQ-9 mean=13.04 (SD 7.40), and BAI mean=22.76 (SD 16.46). Following the study period, scores were as follows: GDS-15 mean=7.22 (SD 4.44), CESD-R mean=22.65 (SD 16.15), PHQ-9 mean=8.59 (SD 6.30), and BAI mean=13.31 (SD 11.52). All participants successfully installed and used the BIG4+ app over the 4-week period, with a mean compliance rate above 93%. All descriptive details are summarized in Table 1.

XSL•FO
RenderX

Table . Descriptives of demographics, psychological assessments, and mobile app (BIG4+) compliance (N=49).

Descriptives	Values
Gender, n (%)	
Women	35 (71.4)
Men	14 (28.6)
Age (y), mean (SD)	70.7 (5.8)
Operating system, n (%)	
Android	49 (100)
iOS	0 (0)
Income level, n (%)	
Lowest	0 (0)
Low	6 (12.2)
Middle	15 (30.6)
High	8 (16.3)
Highest	18 (36.7)
No response	2 (4.1)
Residence type, n (%)	
Living with family	20 (40.8)
Living alone	29 (59.2)
Relationship with family, n (%)	
Very bad	5 (10.2)
Bad	8 (16.3)
Moderate	16 (32.7)
Good	15 (30.6)
Very good	4 (8.2)
No response	1 (2.0)
Intimate person present, n (%)	
Yes	28 (57.1)
No	19 (38.8)
No response	2 (4.1)
Belonging group present, n (%)	
None	15 (30.6)
1 group	19 (38.8)
2 groups	8 (16.3)
3 groups	5 (10.2)
Above 4 groups	2 (4.1)
Preassessment scale (V1), mean (SD)	
GDS-15[a]	9.4 (4.4)
CESD-R[b]	33.9 (18.2)
PHQ-9[c]	13 (7.4)
BAI[d]	22.8 (16.5)
Postassessment scale (V2), mean (SD)	
GDS-15	7.2 (4.4)

XSL•FO
RenderX

Descriptives	Values
CESD-R	22.7 (16.1)
PHQ-9	8.6 (6.3)
BAI	13.3 (11.5)
BIG4+ adherence (%), mean (SD)	93.8 (9.1)

[a]GDS: 15-item Geriatric Depression Scale.

[b]CESD-R: Center for Epidemiologic Studies Depression Scale-Revised.

[c]PHQ-9: 9-item Patient Health Questionnaire.

[d]BAI: Beck Anxiety Inventory.

EMA-Adjunctive Care and Short-Term Symptom Improvement in at-Risk Older Adults

To evaluate the benefit of EMA-based self-monitoring, preassessment and postassessment scores were compared across 4 psychological scales (GDS-15, CESD-R, PHQ-9, BAI). The results indicated that CESD-R, PHQ-9, and BAI scores violated the assumption of normality (*P* values =.01, .01, and .01, respectively), while the GDS-15 difference scores did not (*P*=.33). Consequently, Wilcoxon signed-rank tests were applied to the CESD-R, PHQ-9, and BAI, and a paired-samples *t* test was used for the GDS-15 (Table 2).

All 4 measures showed statistically significant improvements following the intervention. CESD-R scores decreased by a mean of 11.50 points (SD 15.8; SE 2.26; *W*=1045.00; *P*<.001), with a rank biserial correlation of 0.78. PHQ-9 scores decreased by a mean of 4.50 points (SD 6.9; SE 1.00; *W*=947.00; *P*<.001), with an effect size of 0.75. BAI scores were reduced by 9.00 points (SD 11.7; SE 1.67; *W*=940.00; *P*<.001), with a rank biserial correlation of 0.82. GDS-15 scores showed a significant reduction of 2.14 points (SD 2.8; SE 0.40; t_{48}= 5.30; *P*<.001), corresponding to a Cohen *d* of 0.76.

Table . Paired tests of psychological assessment scores between preintervention (V1) and postintervention (V2).

Scales	Test	Statistics	*P* value	Rank biserial correlation	Cohen *d*
GDS-15[a]	Student *t*	5.30	<.001	—[b]	0.76
CESD-R[c]	Wilcoxon *W*	1045.0	<.001	0.78	—
PHQ-9[d]	Wilcoxon *W*	947.0	<.001	0.75	—
BAI[e]	Wilcoxon *W*	940.0	<.001	0.82	—

[a]GDS-15: 15-item Geriatric Depression Scale.

[b]Not applicable.

[c]CESD-R: Center for Epidemiologic Studies Depression Scale-Revised.

[d]PHQ-9: 9-item Patient Health Questionnaire.

[e]BAI: Beck Anxiety Inventory.

Associations Between EMA-Derived Features and Symptom Improvement

Linear relationships between EMA-derived features and changes in depressive and anxiety symptoms were examined using correlational analyses. A full table of correlational analyses is presented in Multimedia Appendix 2.

Descriptive statistics of EMA scores showed generally small-to-modest associations with mental health change scores (Table 3), with no significant correlations for ΔGDS-15 across items. Significant associations emerged for other outcomes, suggesting item-specific links with symptom change: ΔCESD-R was correlated with the minimum of appetite (*r*=0.294; *P*=.04) and general evaluation (*r*=0.29; *P*=.04). ΔPHQ-9 was correlated with the median of feeling (*r*=0.291; *P*=.04), mean (*r*=0.325; *P*=.02), and minimum of appetite rating (*r*=0.291; *P*=.04), and the minimum of general evaluation (*r*=0.286; *P*=.046). For anxiety, ΔBAI was associated only with the minimum sleep quality rating (*r*=0.335; *P*=.02).

XSL•FO
RenderX

Table . Associations between mental health change scores (Δ), EMA score, and EMA response time descriptives.[a]

Type, EMA[b] item, and feature	ΔGDS-15[c] (r^p)	ΔCESD-R[d] (r^p)	ΔPHQ-9[e] (r^p)	ΔBAI[f] (r^p)
EMA score				
Feeling				
Mean	0.094	0.239	0.251	0.204
SD	0.074	–0.246	–0.193	–0.255
Minimum	–0.01	0.267	0.234	0.26
Median	0.039	0.252	0.291*	0.234
Maximum	0.233	0.001	0.062	0.025
Appetite				
Mean	0.241	0.273	0.325*	0.152
SD	0.145	–0.221	–0.136	–0.163
Minimum	0.013	0.294*	0.291*	0.14
Median	0.193	0.245	0.264	0.107
Maximum	0.263	0.006	0.086	–0.07
Sleep quality				
Mean	0.119	0.074	0.197	0.16
SD	–0.004	–0.231	–0.123	–0.192
Minimum	–0.06	0.245	0.277	0.335*
Median	0.116	0.06	0.141	0.085
Maximum	0.028	0.00	0.038	0.042
General evaluation				
Mean	0.121	0.168	0.198	0.149
SD	0.016	–0.26	–0.196	–0.203
Minimum	0.029	0.29*	0.286*	0.261
Median	0.106	0.156	0.175	0.162
Maximum	0.024	0.072	0.04	0.101
EMA response time				
Feeling				
Mean	0.091	–0.073	0.2	–0.108
SD	0.092	0.074	0.091	–0.146
Minimum	0.099	–0.028	0.17	–0.039
Median	0.101	–0.129	0.175	–0.164
Maximum	0.107	0.103	0.181	–0.05
Appetite				
Mean	0.195	–0.122	0.037	–0.301*
SD	0.008	–0.141	–0.077	–0.282
Minimum	0.105	–0.114	0.04	–0.158
Median	0.229	–0.053	0.077	–0.261
Maximum	0.044	–0.159	–0.049	–0.35*
Sleep quality				
Mean	0.251	–0.067	0.011	–0.224
SD	0.189	–0.003	–0.106	–0.023

XSL•FO
RenderX

Type, EMA[b] item, and feature	ΔGDS-15[c] (r^p)	ΔCESD-R[d] (r^p)	ΔPHQ-9[e] (r^p)	ΔBAI[f] (r^p)
Minimum	0.17	–0.104	0.053	–0.139
Median	0.257	–0.015	0.078	–0.189
Maximum	0.143	–0.079	–0.13	–0.13
General evaluation				
Mean	0.084	–0.071	0.129	–0.034
SD	0.201	–0.148	–0.13	–0.159
Minimum	–0.004	–0.09	0.097	–0.018
Median	0.083	–0.024	0.146	–0.034
Maximum	0.138	–0.193	–0.068	–0.182

[a]Significance thresholds: *P<.05, **P<.01, ***P<.001.

[b]EMA: ecological momentary assessment.

[c]GDS-15: 15-item Geriatric Depression Scale.

[d]CESD-R: Center for Epidemiologic Studies Depression Scale-Revised.

[e]PHQ-9: 9-item Patient Health Questionnaire.

[f]BAI: Beck Anxiety Inventory.

Descriptive statistics of raw EMA RTs showed limited associations with symptom change (Table 3). No RT features were significantly correlated with ΔGDS-15, ΔCESD-R, or ΔPHQ-9. The only significant RT associations were observed for anxiety change (ΔBAI), where mean (r=–0.301; P=.04) and maximum (r=–0.35; P=.01) of appetite RT were negatively correlated with ΔBAI, underscoring the challenge of using noisy, naturalistic raw RT descriptives to track depressive symptom change in at-risk older adults.

In contrast, exponential-decay parameters derived from Z-RT (within subject×response-option standardization) showed clearer associative patterns (Table 4). For the feeling item, both the decay rate and asymptote were significantly associated with ΔGDS-15 (θ_b: r=–0.398; P=.005, θ_c: r=–0.321; P=.03), representing the strongest ΔGDS-15 correlations among the RT-derived features and exceeding those observed for EMA score descriptives. Feeling-item parameters also showed significant associations with other outcomes: ΔCESD-R was related to model fit and dynamics (R^2: r=0.358; P=.01, θ_b: r=–0.302; P=.04, θ_c: r=–0.305; P=.04), while ΔBAI was associated only with goodness-of-fit (R^2: r=0.33; P=.02). Notably, no significant associations were observed for decay parameters fitted to appetite, sleep quality, or general evaluation across change scores (all P>.05), and no Z-RT decay features were significantly related to ΔPHQ-9.

XSL•FO
RenderX

Table . Associations between mental health change scores and EMA Z-RT exponential decay parameters.[a]

EMA[b] item and RT[c] parameter	ΔGDS-15[d] (r^p)	ΔCESD-R[e] (r^p)	ΔPHQ-9[f] (r^p)	ΔBAI[g] (r^p)
Feeling				
R^2	0.19	0.358*	0.271	0.330*
θ_a	0.209	0.144	0.017	0.244
θ_b	−0.398**	−0.302*	−0.263	−0.186
θ_c	−0.321*	−0.305*	−0.266	−0.166
Appetite				
R^2	0.034	0.198	0.206	0.266
θ_a	0.047	0.17	0.158	0.188
θ_b	0.23	0.072	−0.038	−0.12
θ_c	0.172	0.004	−0.084	−0.153
Sleep quality				
R^2	−0.194	−0.196	−0.28	−0.021
θ_a	0.107	0.165	0.042	0.206
θ_b	0.182	0.137	0.213	−0.059
θ_c	0.099	0.145	0.229	0.041
General evaluation				
R^2	0.206	−0.045	−0.05	−0.15
θ_a	0.242	0.032	−0.042	−0.054
θ_b	−0.104	0.227	0.181	0.127
θ_c	−0.057	0.119	0.155	0.213

[a]Significance thresholds: *P<.05, **P<.01, ***P<.001.

[b]EMA: ecological momentary assessment.

[c]RT: response time.

[d]GDS-15: 15-item Geriatric Depression Scale.

[e]CESD-R: Center for Epidemiologic Studies Depression Scale-Revised.

[f]PHQ-9: 9-item Patient Health Questionnaire.

[g]BAI: Beck Anxiety Inventory.

Group Differences in RT Trajectories by EMA-Adjunctive Care Response

Bayesian Multilevel Modeling

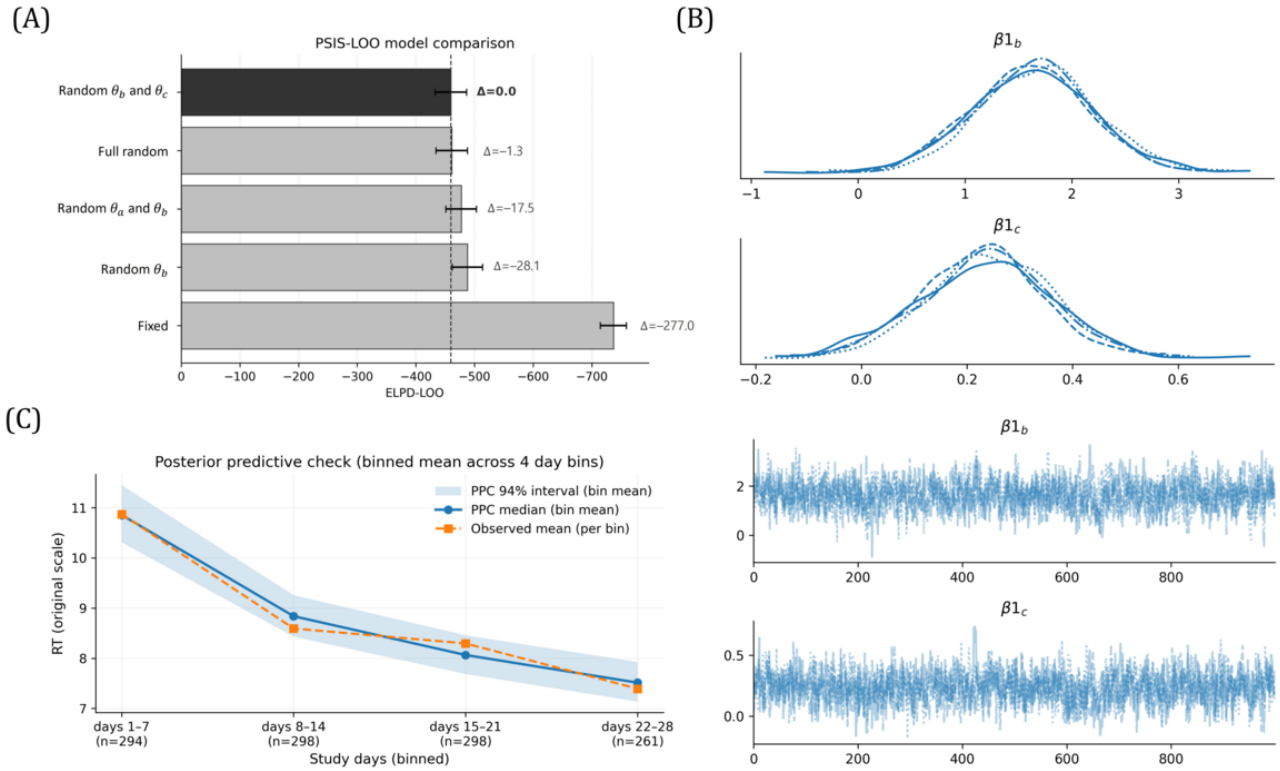
We next evaluated the clinical relevance of EMA-based RT trajectories in relation to depressive symptom improvement during EMA-adjunctive care. Responder-status moderation was tested via the posterior distribution of the group term (β1) in a Bayesian multilevel exponential-decay model, where responders were defined as participants whose GDS-15 change exceeded the MDC threshold (ΔGDS-15≤–3.81). Because prior correlational analyses indicated that only RT dynamic parameters from the feeling item were reliably associated with ΔGDS-15, the following results focus on models fit to feeling-item RT (including model selection, adequacy checks, and group effects). Group differences are summarized using posterior estimates, responder-to-nonresponder parameter ratios, and posterior probabilities of increase.

Model Selection (Random-Effects Structure)

PSIS-LOO model comparison favored a partial hierarchical exponential-decay model with subject-level random effects on the decay rate (θ_b) and postadaptation RT (θ_c) (Figure 1A; ELPD-LOO=−459.7). A full random-effects model performed similarly (ΔELPD-LOO=1.3) but showed higher effective complexity (ploo: 112.2 >109.5), indicating limited predictive benefit from additionally allowing subject-specific variability in amplitude (θ_a). Models with reduced or no random effects showed substantially poorer predictive performance, supporting the need to model between-person heterogeneity in RT adaptation. Accordingly, the model with random decay rate (θ_b) and postadaptation RT (θ_c) was selected for subsequent analyses.

XSL•FO
RenderX

Figure 1. Model selection and diagnostic summary. (A) Pareto-smoothed importance sampling leave-one-out (PSIS-LOO) expected log predictive density (ELPD) by random-effects structure (higher is better). (B) Markov chain Monte Carlo (MCMC) trace plots for responder-status moderation terms (β1) on key parameters (θ_b and θ_c). (C) Binned posterior predictive check comparing observed response time (RT) summaries with posterior predictive medians and 94% intervals. PPC: posterior predictive check.



Model Adequacy and Diagnostics

Building on the selected random-effects structure (random θ_b and θ_c), adding responder-status moderation produced a small but consistent gain in out-of-sample predictive performance relative to the nonmoderated model (ELPD-LOO: −457.2 vs −459.7; SE≈26.6 - 26.7). Although the absolute improvement was modest, it suggests that responder status explains some systematic variation in RT trajectories beyond subject-to-subject heterogeneity. LOO influence diagnostics were stable (99.7% of observations with Pareto k≤0.7), indicating that the model's predictive evaluation was not driven by a small set of highly influential observations and that PSIS-LOO approximations were reliable. Markov chain Monte Carlo diagnostics further supported reliable posterior inference. Trace plots showed good mixing with no visible chain separation (Figure 1B), and convergence metrics (R≈1.00–1.01 with generally high ESS) indicated that posterior summaries are numerically stable and unlikely to reflect sampling pathologies. Finally, posterior predictive checks demonstrated that the model captures the key empirical pattern of interest, simulated trajectories reproducing the overall decay trend, and binned observed means closely tracked posterior predictive check medians while remaining largely within the 94% predictive intervals on the original RT scale (Figure 1C). Together, these checks suggest that the moderated model provides an adequate and generalizable description of RT dynamics in this sample, supporting the downstream interpretation of responder-nonresponder differences in adaptation parameters.

Responder Versus Nonresponder Differences in RT Adaptation

Responders exhibited markedly faster RT adaptation for the feeling item than nonresponders (Table 5). The responder-to-nonresponder ratio for the decay rate was substantially greater than 1 (median θ_b ratio=4.86, 95% CrI 1.44-14.31; P (increase)=.99), indicating a steeper decline (4.86 times faster adaptation) in RT over repeated EMA administrations. Responders also exhibited higher postadaptation RT level than nonresponders (median exp (θ_c) ratio=1.25), indicating 1.25 times slower response after repeated administrations. However, group differences in the postadaptation RT level were more uncertain (95% CrI 0.95-1.58; P (increase)=.95), suggesting at most modest separation in late-phase baseline latency.

Group-averaged trajectories were consistent with the posterior estimates, showing a steeper RT decline among responders (Figure 2A). Differences in the postadaptation RT level became more apparent when model-implied trajectories were extrapolated to 100 days beyond the observed study window (shaded region in Figure 2A). Subject-specific estimates supported the group-level findings (Figure 2B). Although posterior estimates varied within each group, such that some responders exhibited slower adaptation or lower baseline RT (and vice versa), individual trajectories generally followed the overall group trend. Consistent with this, rank plots of participant-level parameters showed responders tending to

XSL•FO
RenderX

appear toward higher values of the decay rate and postadaptation RT level, indicating generally larger estimated values.

Table . Group moderation effects across EMA domains[a].

Quantity and EMA[b] item	Posterior median (IQR; width; 95% CrI[c])	Posterior *P* (increase)
Decay rate (θ_b) ratio		
Feeling	4.86 (3.524-7.609; 4.085; 1.438-14.31)	0.99
Appetite	2.22 (1.565-3.169; 1.604; 0.793-6.561)	0.93
Sleep quality	1.69 (1.186-2.373; 1.187; 0.608-4.707)	0.84
General evaluation	3.32 (1.828-6.229; 4.4; 0.593-19.26)	0.92
Average	3.11 (2.117-4.592; 2.475; 0.991-10.86)	0.97
Baseline RT (exp (θ_c)) ratio		
Feeling	*1.25* (1.164-1.38; 0.2164; 0.953-1.581)	0.95
Appetite	1.04 (0.975-1.102; 0.1267; 0.862-1.247)	0.66
Sleep quality	0.96 (0.9008-1.025; 0.1244; 0.792-1.161)	0.33
General evaluation	1.09 (0.9891-1.2; 0.2108; 0.813-1.433)	0.73
Average	1.08 (1.015-1.154; 0.1392; 0.886-1.315)	0.80

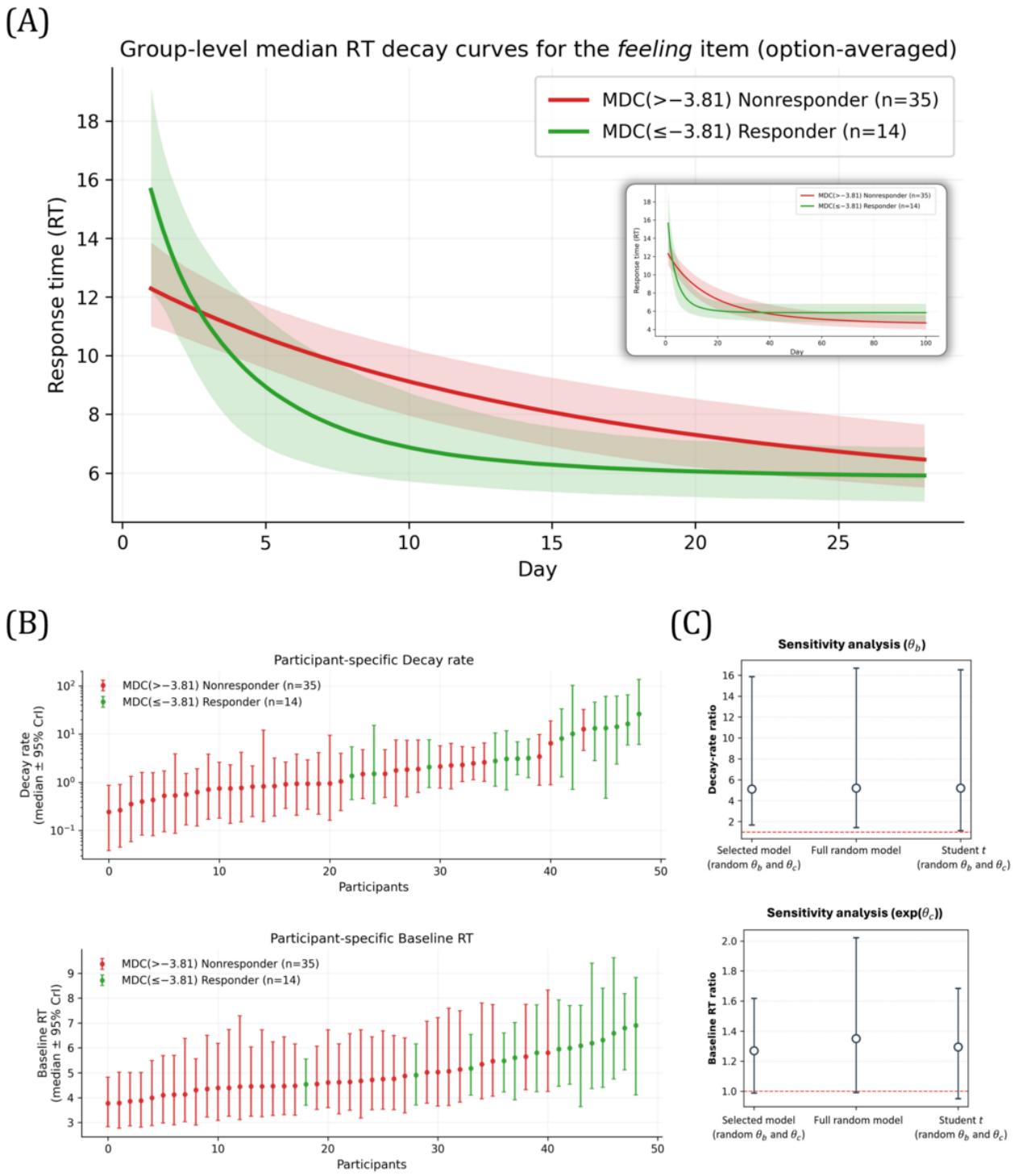
[a]Ratios >1 indicate larger values in responder than nonresponder.

[b]EMA: ecological momentary assessment.

[c]CrI: credible interval.

XSL•FO
RenderX

Figure 2. Responder-status (minimally detectable change [MDC]) moderation results for the feeling item. (A) Model-implied response time (RT) trajectories by group (posterior median with 95% credible interval [CrI]). The shaded region shows an illustrative extrapolation of the fitted trajectories up to 100 days. (B) Participant-level estimates for key parameters (θ_b and θ_c) by group. (C) Sensitivity analysis of group-effect ratios across alternative model specifications.



Item-Specific Moderation Across EMA Domains

Additionally, moderation effects were examined across individual EMA items and an averaged RT outcome. Evidence for group separation was strongest for the feeling item, which showed the largest and most certain decay-rate difference (θ_b ratio=4.86, 95% CrI 1.44-14.31; *P*=.99). Other items and the averaged outcome showed smaller, less certain decay-rate ratios with CrIs generally overlapping 1. A similar pattern was shown for the postadaptation RT level, the feeling item again showed the largest (but still uncertain) group ratio (exp (θc) ratio=1.25, 95% CrI 0.95-1.58), whereas the remaining items showed minimal differences with intervals spanning 1.

Sensitivity Analyses

Group-effect conclusions for the feeling item were robust across plausible alternative specifications, including a full

XSL•FO
RenderX

random-effects model and a Student *t* likelihood or alternative priors (Figure 2C). The responder-to-nonresponder decay-rate ratio remained consistently around 5 (~4.8), with 95% CrIs entirely above 1 across specifications, indicating a stable group difference in RT adaptation speed. By contrast, the postadaptation RT ratio remained modest (~1.3), and its CrIs overlapped 1 in all specifications, suggesting that late-phase RT differences are small and uncertain in this dataset.

Discussion

Summary of Results

This study provides preliminary evidence for the feasibility and potential efficacy of EMA-based self-monitoring in reducing psychological symptoms among older adults at elevated risk. Across 4 validated scales (CESD-R, PHQ-9, BAI, and GDS-15), participants showed statistically significant symptom reductions over the 4-week period, with medium-to-large effect sizes. These findings align with prior work suggesting that EMA-related benefits may be enhanced when used as an adjunct to treatment as usual, and they extend that evidence to a geriatric context. The high adherence to the EMA mobile app (>90%) further indicates that EMA-based self-monitoring is feasible and acceptable in older adults, including those with limited familiarity with mobile technologies, supporting its promise as a scalable adjunct to depression care in late life.

Correlational analyses also indicated that EMA-derived features can predict variation in symptom change, but with important limitations. Descriptive statistics of EMA scores and raw EMA RTs showed symptom-specific associations across measures, yet they were not significantly related to changes in the geriatric depression scale (ΔGDS-15), suggesting reduced sensitivity for tracking geriatric depressive symptom change using simple descriptives alone. In contrast, features derived from exponential-decay parameters fitted to standardized response times (Z-RT) showed consistent associations with depressive symptom change (ΔGDS-15 and ΔCESD-R). In particular, faster adaptation (higher decay rate), higher postadaptation RT level (asymptote/baseline), and better fit quality for the emotionally salient feeling item are associated with greater symptom improvement, suggesting that RT dynamics during EMA engagement may reflect clinically meaningful change during EMA-adjunctive care.

These patterns were further corroborated by the Bayesian multilevel modeling results. Compared with nonresponders, responders to EMA-adjunctive care exhibited markedly faster adaptation (higher decay rate) and modestly higher postadaptation RT levels. This suggests that potential responders show a distinct trajectory characterized by more rapid accommodation to repeated self-reflection, although the underlying mechanism cannot be determined from the current design (ie, whether EMA-adjunctive care enhances self-monitoring efficiency, or whether individuals with greater baseline capacity for such adaptation are more likely to benefit). Evidence for between-group differences in the postadaptation RT level was comparatively uncertain, and the direction (ie, slightly higher baseline RT in responders) may appear counterintuitive if faster responding is assumed to reflect greater efficiency. One plausible interpretation is that faster late-phase responding among nonresponders could reflect a tendency toward more superficial or less attentive responding rather than more efficient self-reflection, but this remains speculative and warrants direct validation. Additionally, sensitivity analyses across plausible alternative specifications yielded the same qualitative conclusions, supporting the robustness of the central group-difference inference. Collectively, these results highlight the potential clinical use of behavioral signatures embedded in EMA responding, particularly RT adaptation dynamics, for tracking depressive symptom improvement in older adults.

Interestingly, consistent with the correlational findings, group contrasts were strongest when the model was fitted to the feeling item among the 4 EMA domains examined. Moreover, although prior work suggests that averaging RT across multiple EMA items can increase measurement reliability [14,16], group differences in adaptive patterns were larger when using the feeling item alone. This finding suggests that EMA response dynamics may be item-specific rather than purely reflecting a global response-speed trait, raising the possibility that different symptom domains could exhibit distinct behavioral signatures in EMA trajectories.

Clinical Implications

These findings suggest that EMA-adjunctive care may be both feasible and clinically beneficial for older adults at risk of depression. Digital mental health interventions (DMHIs) are often considered difficult to implement in late-life populations because limited familiarity with digital tools can reduce sustained engagement. In contrast, adherence in this study remained high (on average >90%), indicating strong real-world acceptability. The EMA protocol used here, brief daily reporting on four items, appeared manageable even for participants who might be expected to show reduced willingness or capacity for sustained participation due to depressive burden. To support sustained engagement in real-world deployment, the intended implementation can incorporate an explicit feedback loop, such as clinician-facing dashboards, periodic review during EMA engagements, or automated triggers that prompt timely outreach when a missing response is detected. Importantly, symptom reductions observed across validated scales indicate that EMA use in this context may confer benefits beyond feasibility, supporting EMA-adjunctive care as a promising component of geriatric depression management.

A second implication concerns the objective monitoring of symptom change. Remote psychiatric care still relies heavily on self-reported symptoms, ranging from retrospective standardized questionnaires to daily EMA ratings, both of which remain vulnerable to recall bias and other reporting distortions (eg, mood-congruent evaluation, self-reflective bias). EMA-derived RT metrics are not fully “passive” because they require active responding; however, they provide a comparatively objective behavioral signal that reflects underlying cognitive and psychomotor processes and is less directly shaped by deliberative self-presentation or response framing. The observed association between RT adaptation dynamics and clinically meaningful symptom improvement, therefore, supports RT trajectories as a complementary marker

XSL•FO
RenderX

for tracking depressive symptom variation in naturalistic settings, potentially improving reliability when subjective reports are noisy or inconsistent.

Moreover, EMA-RT adaptation features may help make digital mental health interventions more actionable and personalized. As noted in the *Introduction* section, the effects of ecological momentary interventions are heterogeneous and depend on individual characteristics. Consequently, a practical barrier to scaling digital care is identifying early on who will benefit from a given protocol and who will require additional support. Our results suggest that RT adaptation patterns within an initial monitoring window (approximately 1 month) may provide an early indicator of EMA responsiveness. Such early stratification could guide stepped-care decisions (eg, prompting intensified clinician contact, alternative interventions, or safety monitoring for likely non-responders) while allowing responders to continue with lower-intensity, self-guided support, thereby supporting more efficient allocation of constrained clinical resources and optimizing just-in-time adaptive intervention approaches. We also note that this early stratification window may be shortened through systematic investigation of the optimal decision period. Specifically, building on the observed faster adaptation among responders, future work could quantify within-person day-to-day EMA-RT gradients (or slopes over rolling windows) and evaluate candidate "decision days" (eg, 7, 10, 14, or 21 days) to determine an earlier, reliable point for stepped-care decisions.

Finally, the item-specific links observed between EMA domains and symptom changes suggest a pathway to broaden EMA monitoring beyond depression. Because psychiatric symptoms frequently co-occur, identifying domain-specific behavioral signatures (eg, the feeling-depression linkage observed here) may enable EMA to track multiple symptom dimensions simultaneously, supporting more nuanced, individualized monitoring and intervention planning.

Limitations

First, regarding the observed effectiveness of the EMA-based self-monitoring, although significant reductions in depressive and anxiety symptoms were detected across multiple validated scales, the absence of a control group precludes definitive causal attribution. Participants were not randomized, and potential concurrent treatments, such as case management, pharmacotherapy, or psychotherapy, were not systematically recorded or controlled, introducing possible confounding influences. Additionally, postintervention assessments were conducted only once at the 4-week mark without follow-up measurements to evaluate the durability of treatment effects. As such, the long-term efficacy and sustainability of EMA-based self-monitoring in older adults with depression remain undetermined.

While several EMA-derived features were shown to be associated with symptom variation, some unexpected patterns emerged in the relationship between self-reported EMA scores and symptom change. Specifically, positive correlations were observed between mood-related EMA scores and symptom change in both depression and anxiety domains. These findings suggest that higher EMA ratings during the intervention period were paradoxically linked to less clinical improvement. One possible interpretation is that individuals with persistent depressive symptoms may have reduced self-reflective precision, leading to inflated or undifferentiated positive reports. Alternatively, healthier participants may have demonstrated more nuanced and accurate self-monitoring. However, this hypothesis requires further empirical validation.

Additionally, given the modest sample size (n<50), the correlational analyses were likely underpowered to detect small associations between symptom-change scores and EMA-derived features; thus, null or weak correlations should be interpreted cautiously. Larger cohorts will be needed to more precisely estimate small effects and confirm the robustness and generalizability of the observed associations, particularly those involving the exponential-decay parameters.

Finally, in examining dynamic EMA response trajectories, the most distinctive group-level differences between responders and non-responders emerged in the fitted parameters of the feeling item, whereas appetite, sleep quality, and general evaluation items yielded less pronounced contrasts. This discrepancy may reflect the unique psychological salience of the feeling item, which directly targets the immediate emotional state, in contrast to the more somatically anchored or abstract content of the other items. Alternatively, the fixed order of EMA item presentation, wherein the feeling item was always administered first, may have contributed to its stronger predictive utility. It is plausible that the first item elicits more cognitive effort or reflects the initiation speed of self-reflective responses, thus capturing more variance in related cognitive functioning. This potential order effect should be systematically examined in future studies to better delineate what EMA-based response latencies truly index in ecological settings.

Conclusions

This study demonstrates the feasibility of using EMA-based self-monitoring as both an adjunctive care and a behavioral phenotyping tool in older adults with a high risk of depression.

The implications for digital health care are manifold. First, the integration of response-time dynamics into mobile health platforms offers a nonobtrusive means of tracking cognitive and emotional engagement, even in populations with limited digital literacy. The method's compatibility with brief, low-frequency EMA schedules (eg, daily reports, about a month) makes it especially suitable for aging individuals, who may be less tolerant of intensive digital protocols. Second, the use of within-subject standardization and curve-fitting models supports individualized tracking, allowing for the detection of meaningful intra-individual changes. This has significant implications for the early identification of symptom relapse, monitoring of treatment responsiveness, and tailored interventions based on dynamic behavioral signatures.

In sum, the introduced EMA-based modeling offers a novel, sensitive, and pragmatic approach for advancing mental health care in aging populations. Future work should aim to validate these findings in randomized controlled settings, investigate the mechanistic underpinnings of RT adaptations, and explore broader applications in digital psychiatry and geriatric care.

XSL•FO
RenderX

Acknowledgments

We extend our sincere gratitude to the Suwon Geriatric Mental Health and Welfare Center for its invaluable role as an organizational partner and for its dedicated support in facilitating comprehensive data collection for this research. The authors used ChatGPT (OpenAI) to assist with language refinement. Sang Joon Son is the co-corresponding author (email: sjsonpsy@ajou.ac.kr; phone: 82 31-219-5180).

Funding

This research was supported by the ASTRA Project through the National Research Foundation funded by the Ministry of Science and Information and Communications Technology (number RS-2024-00470056).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical restrictions imposed by the institutional review board and the sensitive nature of clinical and behavioral data but may be obtained from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: Jeehang Lee (equal), Jooho Lee (lead)
Data curation: Jooho Lee
Formal analysis: Jeehang Lee (equal), Jooho Lee (lead)
Funding acquisition: GD (lead), JYP (equal)
Investigation: JN (supporting), KC (lead), SM (supporting), SP (equal)
Methodology: Jeehang Lee (equal), Jooho Lee
Project administration: KC (lead), JN (supporting), SM (supporting), SP (equal)
Resources: JN (supporting), KC (lead), SM (supporting), SP (equal)
Supervision: JYP (equal), SJS (lead)
Validation: JYP (equal), SJS (lead)
Visualization: JN (supporting), Jooho Lee (lead)
Writing – original draft: Jooho Lee
Writing – review & editing: Jeehang Lee (equal), Jooho Lee (lead), JYP (equal), SJS (equal)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Methodological details.
[DOCX File, 133 KB - aging_v9i1e83891_app1.docx]

Multimedia Appendix 2
Correlational analysis table.
[XLSX File, 28 KB - aging_v9i1e83891_app2.xlsx]

References

1. Future household projections 2022-2052 [Web Page in Korean]. Statistics Korea. 2024. URL: https://mods.go.kr/board.es?mid=a10301010000&bid=207&list_no=432696&act=view&mainXml=Y [accessed 2025-04-07]
2. Future population projections: 2022–2072 [Web Page in Korean]. Statistics Korea. 2024. URL: https://mods.go.kr/board.es?mid=a10301020600&bid=207&act=view&list_no=428476 [accessed 2025-04-07]
3. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavorial medicine. Ann Behav Med 1994 Sep 1;16(3):199-202. [doi: 10.1093/abm/16.3.199]
4. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
5. Saeb S, Zhang M, Karr CJ, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015 Jul 15;17(7):e175. [doi: 10.2196/jmir.4273] [Medline: 26180009]
6. Torous J, Kiang MV, Lorme J, Onnela JP. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health 2016 May 5;3(2):e16. [doi: 10.2196/mental.5165] [Medline: 27150677]

XSL•FO
RenderX

7. Burns MN, Begale M, Duffecy J, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011 Aug 12;13(3):e55. [doi: 10.2196/jmir.1838] [Medline: 21840837]
8. Kramer I, Simons CJP, Hartmann JA, et al. A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. World Psychiatry 2014 Feb;13(1):68-77. [doi: 10.1002/wps.20090] [Medline: 24497255]
9. Hur JW, Kim B, Park D, Choi SW. A scenario-based cognitive behavioral therapy mobile app to reduce dysfunctional beliefs in individuals with depression: a randomized controlled trial. Telemed J E Health 2018 Sep;24(9):710-716. [doi: 10.1089/tmj.2017.0214] [Medline: 29323626]
10. Marciniak MA, Shanahan L, Rohde J, et al. Standalone smartphone cognitive behavioral therapy-based ecological momentary interventions to increase mental health: narrative review. JMIR mHealth uHealth 2020 Nov 12;8(11):e19836. [doi: 10.2196/19836] [Medline: 33180027]
11. Myin-Germeys I, Kasanova Z, Vaessen T, et al. Experience sampling methodology in mental health research: new insights and technical developments. World Psychiatry 2018 Jun;17(2):123-132. [doi: 10.1002/wps.20513] [Medline: 29856567]
12. Bastiaansen JA, Ornée DA, Meurs M, Oldehinkel AJ. An evaluation of the efficacy of two add-on ecological momentary intervention modules for depression in a pragmatic randomized controlled trial (ZELF-i). Psychol Med 2020 Dec 14;52(13):1-10. [doi: 10.1017/S0033291720004845] [Medline: 33315003]
13. Chung K, Park JY, Joung D, Jhung K. Response time as an implicit self-schema indicator for depression among undergraduate students: preliminary findings from a mobile app-based depression assessment. JMIR mHealth uHealth 2019 Sep 13;7(9):e14657. [doi: 10.2196/14657] [Medline: 31586362]
14. Hernandez R, Hoogendoorn C, Gonzalez JS, et al. Reliability and validity of noncognitive ecological momentary assessment survey response times as an indicator of cognitive processing speed in people's natural environment: intensive longitudinal study. JMIR mHealth uHealth 2023 May 30;11:e45203. [doi: 10.2196/45203] [Medline: 37252787]
15. Diagnostic and Statistical Manual of Mental Disorders, 3rd edition: American Psychiatric Association; 1980.
16. Hernandez R, Hoogendoorn C, Gonzalez JS, Pyatak EA, Crespo-Ramos G, Schneider S. Reliability and validity of ecological momentary assessment response time-based measures of emotional clarity: secondary data analysis. JMIR Ment Health 2024 Jul 18;11:e58352. [doi: 10.2196/58352] [Medline: 39024004]
17. Pier M, Hulstijn W, Sabbe BGC. Psychomotor retardation in elderly depressed patients. J Affect Disord 2004 Jul;81(1):73-77. [doi: 10.1016/j.jad.2003.08.002] [Medline: 15183603]
18. Bonin P, Barry C, Méot A, Chalard M. The influence of age of acquisition in word reading and other tasks: a never ending story? J Mem Lang 2004 May;50(4):456-476. [doi: 10.1016/j.jml.2004.02.001]
19. Beheydt LL, Schrijvers D, Docx L, Bouckaert F, Hulstijn W, Sabbe B. Cognitive and psychomotor effects of three months of escitalopram treatment in elderly patients with major depressive disorder. J Affect Disord 2015 Dec 1;188:47-52. [doi: 10.1016/j.jad.2015.08.041] [Medline: 26342888]
20. Bennabi D, Vandel P, Papaxanthis C, Pozzo T, Haffen E. Psychomotor retardation in depression: a systematic review of diagnostic, pathophysiologic, and therapeutic implications. Biomed Res Int 2013;2013:158746. [doi: 10.1155/2013/158746] [Medline: 24286073]
21. Wolinsky FD, Vander Weg MW, Martin R, et al. The effect of speed-of-processing training on depressive symptoms in ACTIVE. J Gerontol A Biol Sci Med Sci 2009 Apr;64(4):468-472. [doi: 10.1093/gerona/gln044] [Medline: 19181719]
22. Brown PJ, Liu X, Sneed JR, Pimontel MA, Devanand DP, Roose SP. Speed of processing and depression affect function in older adults with mild cognitive impairment. Am J Geriatr Psychiatry 2013 Jul;21(7):675-684. [doi: 10.1016/j.jagp.2013.01.005] [Medline: 23567401]
23. Smith M, Jones MP, Dotson MM, Wolinsky FD. Speed of processing training and depression in assisted and independent living: a randomized controlled trial. PLoS One 2019;14(10):e0223841. [doi: 10.1371/journal.pone.0223841] [Medline: 31622386]
24. Veenman M, Stefan AM, Haaf JM. Bayesian hierarchical modeling: an introduction and reassessment. Behav Res Methods 2024 Aug;56(5):4600-4631. [doi: 10.3758/s13428-023-02204-3] [Medline: 37749423]
25. McGlothlin AE, Viele K. Bayesian hierarchical models. JAMA 2018 Dec 11;320(22):2365-2366. [doi: 10.1001/jama.2018.17977] [Medline: 30535206]
26. Ghahramani Z. Bayesian non-parametrics and the probabilistic approach to modelling. Philos Trans A Math Phys Eng Sci 2012;371(1984):20110553. [doi: 10.1098/rsta.2011.0553] [Medline: 23277609]
27. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):385-401. [doi: 10.1177/014662167700100306]
28. Eaton WW, Muntaner C, Smith C, Tien A, Ybarra M. Center for Epidemiologic Studies Depression Scale: Review and Revision (CESD and CESD-r). In: Maruish ME, editor. Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Instruments for Adults, 3rd edition: Lawrence Erlbaum Associates; 2004:363-377.
29. Lee S, Oh ST, Ryu SY, Jun JY, Lee K, Lee E, et al. Validation of the Korean version of Center for Epidemiologic Studies Depression Scale-Revised. Int. J. Neuropsychopharmacol 2016;19(1):81-82. [doi: 10.1093/ijnp/pyw043.223]
30. Yesavage JA, Sheikh JI. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. Clin Gerontol 1986;5:165-173. [doi: 10.1300/J018v05n01_09]

XSL•FO
RenderX

31. Bae JN, Cho MJ. Development of the Korean version of the Geriatric Depression Scale and its short form among elderly psychiatric patients. J Psychosom Res 2004 Sep;57(3):297-305. [doi: 10.1016/j.jpsychores.2004.01.004] [Medline: 15507257]
32. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
33. Han C, Jo SA, Kwak JH, et al. Validation of the Patient Health Questionnaire-9 Korean version in the elderly population: the Ansan Geriatric study. Compr Psychiatry 2008;49(2):218-223. [doi: 10.1016/j.comppsych.2007.08.006] [Medline: 18243897]
34. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988 Dec;56(6):893-897. [doi: 10.1037//0022-006x.56.6.893] [Medline: 3204199]
35. Lee HK, Lee EH, Hwang ST, et al. Psychometric properties of the Beck Anxiety Inventory in the community-dwelling sample of Korean adults. Korean J Clin Psychol 2016 Nov;35(4):822-830. [doi: 10.15842/kjcp.2016.35.4.010]
36. Newell A, Rosenbloom PS. Mechanisms of skill acquisition and the law of practice. In: Anderson JR, editor. Cognitive Skills and Their Acquisition: Lawrence Erlbaum Associates; 1981:1-55.
37. Heathcote A, Brown S, Mewhort DJ. The power law repealed: the case for an exponential law of practice. Psychon Bull Rev 2000 Jun;7(2):185-207. [doi: 10.3758/bf03212979] [Medline: 10909131]
38. Olivier GN, Walter CS, Paul SS, Dibble LE, Schaefer SY. How common is the exponential decay pattern of motor skill acquisition? A brief investigation. Motor Control 2021 May 13;25(3):451-461. [doi: 10.1123/mc.2020-0043] [Medline: 33992025]
39. Kaiser L. Adjusting for baseline: change or percentage change? Stat Med 1989 Oct;8(10):1183-1190. [doi: 10.1002/sim.4780081002] [Medline: 2682909]
40. Jennings MA, Cribbie RA. Comparing pre-post change across groups: guidelines for choosing between difference scores, ANCOVA, and residual change scores. J Data Sci 2016;14(2):205-230. [doi: 10.6339/JDS.201604_14(2).0002]
41. Patient-focused drug development: incorporating clinical outcome assessments into endpoints for regulatory decision-making. : U.S. Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER); 2023 URL: https://www.fda.gov/media/166830/download [accessed 2025-07-07]
42. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001 Nov;94(2):149-158. [doi: 10.1016/S0304-3959(01)00349-9] [Medline: 11690728]
43. Warkentin LM, Das D, Majumdar SR, Johnson JA, Padwal RS. The effect of weight loss on health-related quality of life: systematic review and meta-analysis of randomized trials. Obes Rev 2014 Mar;15(3):169-182. [doi: 10.1111/obr.12113] [Medline: 24118750]
44. Rogosa DR, Willett JB. Demonstrating the reliability the difference score in the measurement of change. J Educ Meas 1983 Dec;20(4):335-343. [doi: 10.1111/j.1745-3984.1983.tb00211.x]
45. Castro-Schilo L, Grimm KJ. Using residualized change versus difference scores for longitudinal research. J Soc Pers Relat 2018 Jan;35(1):32-58. [doi: 10.1177/0265407517718387]
46. González-Domínguez NP, Wu Y, Fan S, et al. Individual patient data meta-analysis estimates the minimal detectable change of the Geriatric Depression Scale-15. J Clin Epidemiol 2024 Sep;173:111443. [doi: 10.1016/j.jclinepi.2024.111443] [Medline: 38942179]
47. de Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006 Oct;59(10):1033-1039. [doi: 10.1016/j.jclinepi.2005.10.015] [Medline: 16980142]
48. Salvatier J, Wiecki TV, Fonnesbeck C. Probabilistic programming in Python using PyMC3. PeerJ Comput Sci 2016;2:e55. [doi: 10.7717/peerj-cs.55]

Abbreviations

BAI: Beck Anxiety Inventory
CESD-R: Center for Epidemiologic Studies Depression Scale–Revised
CrI: credible interval
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
ELPD-LOO: expected log pointwise predictive density leave-one-out cross-validation
EMA: ecological momentary assessment
GDS-15: 15-item Geriatric Depression Scale
MDC: minimal detectable change
MDD: major depressive disorder
PHQ-9: 9-item Patient Health Questionnaire
PSIS-LOO: Pareto-smoothed importance sampling leave-one-out cross-validation
RT: response time

XSL•FO
RenderX

Edited by J Wang; submitted 10.Sep.2025; peer-reviewed by JP Nunes, R Hernandez; revised version received 18.Feb.2026; accepted 17.Mar.2026; published 08.May.2026.

Please cite as:
Lee J, Lee J, Park S, Do G, Noh J, Moon S, Chung K, Son SJ, Park JY
Response Time Dynamics From Noncognitive Ordinal Ecological Momentary Assessment as a Proxy for Symptom Change in Geriatric Depression: Longitudinal Observational Study
JMIR Aging 2026;9:e83891
URL: https://aging.jmir.org/2026/1/e83891
doi:10.2196/83891

© Jooho Lee, Jeehang Lee, Sehwan Park, Gangho Do, Jihye Noh, Sangjoon Moon, Kyungmi Chung, Sang Joon Son, Jin Young Park. Originally published in JMIR Aging (https://aging.jmir.org), 8.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Emotional Regulation and Physiological Responses to a Cultural Heritage Virtual Reality Experience in Older Adults: Mixed Methods Study

Yafeng Lai[1,2], DPhil; Pohsun Wang[2], Prof Dr

[1]College of Fine Arts and Design, Guangdong Vocational Academy of Art, Guangzhou, China

[2]Faculty of Innovation and Design, City University of Macau, Avenida Padre Tomás Pereira, Taipa, Macau, China

Corresponding Author:
Pohsun Wang, Prof Dr
Faculty of Innovation and Design, City University of Macau, Avenida Padre Tomás Pereira, Taipa, Macau, China

Abstract

Background: Against the backdrop of global population aging, the development of low-burden and widely acceptable interventions to support emotional health in older adults has become a critical research priority. Virtual reality (VR) has been increasingly applied in medical rehabilitation and psychological interventions due to its immersive and interactive features. However, its potential as a form of digital therapy that integrates cultural content and artistic design remains underexplored, particularly with respect to the mechanisms through which such experiences facilitate emotional regulation in older populations. Empirical studies combining subjective emotional outcomes with objective physiological indicators are especially scarce.

Objective: This study aimed to examine the emotional regulation effects of an artistic cultural heritage–based VR experience among older adults and to explore its underlying mechanisms by integrating subjective psychological assessments with objective physiological responses.

Methods: A mixed methods experimental design was adopted using the United Nations Educational, Scientific and Cultural Organization Historic Centre of Macao as the study context. Two VR conditions were developed for comparison: (1) an artistic cultural heritage VR environment and (2) a conventional photorealistic replica VR environment. A total of 49 participants aged 60 years and older were recruited. During the VR experience, electrodermal activity (EDA) was continuously recorded to capture emotional arousal. Physiological data were triangulated with pre- and postintervention psychological questionnaires assessing emotional states, as well as semistructured interviews.

Results: Compared with the realistic replication VR condition, the cultural heritage VR condition was associated with significantly higher positive affect and lower negative affect (P<.05). At the physiological level, a significant between-group difference was observed in mean normal-to-normal interval (P=.01; Cohen d=0.91), whereas no significant differences were found in EDA measures (all P>.05). EDA results showed notable interindividual variability but did not demonstrate consistent group-level effects. These findings suggest that emotional changes may be more prominently reflected in subjective experience than in physiological arousal.

Conclusions: By integrating subjective and physiological measures, this study provides preliminary empirical evidence that cultural heritage VR is associated with improved emotional outcomes in older adults. The findings propose a conceptual interpretative pathway linking digital cultural experience, emotional engagement, and psychological well-being, while highlighting the importance of individual variability in physiological responses. This study contributes to the development of content- and experience-oriented approaches in digital therapeutics and provides an evidence-based framework for emotion-supportive design in digital heritage and aging-related contexts.

(JMIR Aging 2026;9:e91772) doi:10.2196/91772

KEYWORDS

digital therapy; cultural landscape; electrodermal activity; older adults; emotional health; Historic Centre of Macao

Introduction

The World Health Organization [1] projects that the global population aged 60 years and older will reach 2.1 billion by 2050. Mental health conditions such as depression and anxiety are expected to exert an increasingly significant impact on quality of life in later adulthood [2]. As population aging accelerates worldwide, enhancing emotional well-being and life satisfaction among older adults has become a central concern across public health, psychology, and design-related disciplines [3]. Declines in physical capacity and social participation increase vulnerability to loneliness, depression, and anxiety in

XSL•FO
RenderX

later life, which in turn affect daily functioning and overall physical and mental health [2,4].

Traditional approaches to emotional regulation in older adults—such as art therapy, group activities, and cultural visits—are often constrained by mobility limitations, health conditions, or external disruptions such as pandemics [5,6]. In response, digitally mediated therapeutic interventions have emerged as a promising alternative [7]. Identifying effective digital strategies to stimulate positive emotions and psychological resilience in older populations has thus become a key focus of interdisciplinary research [8]. In this context, virtual reality (VR), owing to its immersive, controllable, and spatially representational properties [9], offers new opportunities to enhance emotional experiences among older adults [10-12].

Presence, commonly defined as the subjective sensation of "being there" within a mediated environment, represents a central psychological construct in VR research. A strong sense of presence can enhance users' emotional engagement, attentional focus, and cognitive involvement in virtual environments. Previous studies have demonstrated that higher levels of presence are associated with stronger affective responses and more pronounced physiological reactions during immersive experiences. In particular, presence has been identified as an important psychological pathway through which immersive VR environments influence emotional and behavioral outcomes [13,14]. For example, the National Arts Council of Singapore, in collaboration with the National Gallery Singapore, has provided 1-hour virtual heritage art tours for older adults [15]. Although these programs have been widely welcomed and in high demand, empirical evaluations of their emotional impact remain limited.

Existing studies have demonstrated that VR-based interventions can alleviate depressive and anxiety symptoms and improve overall psychological well-being among older adults [16-18]. However, most prior research has focused on natural environments, cognitive training, or reminiscence-based content, with relatively little attention paid to immersive experiences that incorporate cultural depth and artistic expression [19-21]. Evidence suggests that older users—including individuals with dementia—prefer VR content related to travel, nature, art, social interaction, or familiar environments [22,23].

Unlike conventional cultural reproduction, cultural heritage–based VR emphasizes aesthetic transformation and contextual storytelling. It integrates historical spaces, cultural narratives, and artistic expression into immersive virtual experiences, rather than merely replicating physical environments. Such experiences may strengthen cultural identity and emotional immersion while eliciting deeper affective engagement. Despite this potential, heritage-based VR as a form of digital therapy remains underexplored at the intersection of VR research and aging studies.

Another limitation of this research lies in its heavy reliance on self-report measures, such as questionnaires and interviews, when examining the emotional effects of VR-based cultural experiences [8,24]. Objective physiological indicators are rarely incorporated. Electrodermal activity (EDA), a sensitive measure of autonomic nervous system activity, provides real-time information on emotional arousal and attentional engagement during immersive experiences [25]. Wearable physiological sensors have become an increasingly reliable method for collecting autonomic nervous system signals in naturalistic and experimental contexts [26]. When combined with self-report measures, EDA can reduce biases associated with retrospection and offer a dynamic representation of emotional fluctuations over time. Recent findings highlight the particular sensitivity of EDA responses to emotional arousal and presence in VR environments, underscoring its methodological value for emotional health research [27]. Incorporating EDA into VR-based cultural experience studies therefore represents an important methodological advancement.

To address these gaps, this study selected the Historic Centre of Macao as the experimental context, with a focus on the Ruins of St. Paul's and its surrounding urban spaces. Prior research suggests that historical heritage settings can foster meaning, identity, and subjective well-being among older adults [28,29]. As a United Nations Educational, Scientific and Cultural Organization World Heritage site, Macao embodies distinctive historical narratives and artistic symbolism. The virtual environments were modeled according to real spatial proportions and architectural characteristics to recreate the physical form, spatial scale, and ambient qualities of the historic district. Building on this foundation, an "artistic cultural experience" design approach was introduced, embedding cultural information through aesthetic and interactive elements.

Through heritage-oriented VR experiences, VR was positioned not merely as a tool for spatial replication, but as a medium for artistic translation and emotional regulation [30]. From the perspective of digital cultural heritage and art experience research, artistic interpretation can intensify emotional engagement and perceived meaning, enabling virtual environments to function as affective and interpretive media rather than purely technical simulations [31,32]. By integrating EDA with subjective assessments, this study further explores the relationship between virtual cultural-artistic experiences and physiological emotional responses.

The contributions of this study are 3-fold. First, it advances VR content design from literal replication toward artistic reinterpretation, encouraging emotional resonance and identity engagement among older participants. Second, it combines EDA with subjective measures to examine whether a coordinated pathway exists linking experience, emotional arousal, and emotional health. Third, by using the Historic Centre of Macao as a case study, it explores the social value of cultural heritage digitization in health promotion contexts. This interdisciplinary approach provides both a novel empirical pathway and a model for integrating design research, environmental psychology, and aging studies. Theoretically, it elucidates how design-led content and physiological feedback jointly shape emotional mechanisms in later life. Practically, it offers a replicable framework for digital cultural experiences applicable to cultural institutions, community care settings, and urban cultural policy.

Based on the theoretical framework and the identified research gaps, this study aims to examine whether cultural heritage VR experiences can enhance emotional health among older adults

XSL•FO
RenderX

and to explore the emotional mechanisms associated with such experiences. Specifically, this study addresses 3 core research questions (RQs): (RQ1) whether cultural heritage VR is more effective than realistic replication VR in improving emotional health among older adults, (RQ2) whether cultural heritage VR elicits stronger electrodermal responses reflecting emotional arousal, and (RQ3) whether individual psychological traits including emotion regulation difficulty and psychological resilience are associated with subjective emotional states and depression levels before and after VR intervention. Accordingly, 3 hypotheses (H1-H3) were formulated corresponding to these RQs (Table 1).

Table . Correspondence between research hypotheses (H) and research questions (RQs).

Research hypotheses	Corresponding RQs
H1. Older adults who experience cultural heritage VR[a] will exhibit higher levels of emotional health than those in the realistic replication VR condition.	RQ1. Can digital therapeutics based on cultural heritage VR significantly improve emotional health in older adults?
H2. Cultural heritage VR experiences will elicit stronger electrodermal responses, reflecting greater emotional arousal.	RQ2. Does cultural heritage VR induce significantly higher electrodermal activity, indicating stronger arousal, compared with the control condition?
H3. Individual psychological traits such as difficulty in emotion regulation and psychological resilience are significantly correlated with emotional states (PA[b]/NA[c]) before and after VR intervention.	RQ3. Do individual psychological traits such as emotion regulation ability and psychological resilience influence the intervention effect of VR cultural heritage experience on the emotional health of older people?

[a]VR: virtual reality.

[b]PA: positive affect.

[c]NA: negative affect.

Methods

Ethical Considerations

Older adult participants were recruited through local community committees and care institutions for older people in Zhuhai and Foshan, Guangdong Province. All participants met the study inclusion criteria and participated on a voluntary basis. Ethics approval for this study was granted by the Academic Committee of the School of Innovative Design, City University of Macao (approval Ref 2025 09291315). Written informed consent was obtained from all participants prior to study initiation. Participants' capacity to provide informed consent was independently assessed by a licensed clinical psychologist who was not involved in the study, in order to ensure adequate comprehension and voluntary participation. All collected data were anonymized. Data handling and storage complied with the Organic Law on the Protection of Personal Data and Guarantee of Digital Rights (Law 3/2018, enacted on December 5, 2018). The study was conducted in accordance with good research practice standards and the principles of the Declaration of Helsinki.

Participants

The sample size was estimated based on effect sizes reported in previous studies examining the impact of VR interventions on physiological outcomes, including heart rate variability (HRV) and EDA. Prior studies have demonstrated that VR-based interventions can significantly modulate autonomic and physiological responses, suggesting moderate to large effects [33,34]. Based on these estimates, a minimum of 30 participants per group was required to achieve 80% statistical power at a significance level of .05.

To further account for potential attrition during the VR intervention and follow-up assessments, an anticipated dropout rate of approximately 18% was incorporated, based on prior research involving older adults in VR settings [35]. Accordingly, the target sample size was increased, resulting in a final sample of 49 older adults; the study aimed to recruit at least 30 participants per group. A final sample of 49 older adults was included in the analysis.

Participants were eligible for inclusion if they met the following criteria: (1) aged 60 years or older and able to complete the experimental procedures independently or with assistance; (2) able to understand the study procedures and provide written informed consent, with additional auditory or visual explanations provided when necessary; (3) no uncontrolled cardiovascular disease or epilepsy (participants with stable conditions were included only after clinical approval); and (4) no severe cognitive impairment. Cognitive status was screened using the Mini-Mental State Examination, with a cutoff score of ≥24 indicating eligibility.

Exclusion criteria were (1) a history of severe motion sickness, recent ear surgery, or dermatological conditions that could interfere with EDA electrode attachment; and (2) recent use of medications known to strongly affect autonomic nervous system activity. When medication use could not be avoided, it was recorded and treated as a covariate in subsequent analyses.

This randomized experiment had a total duration of 3 months and concluded on November 30, 2025. Participant recruitment was conducted between September and November 2025. The baseline assessment took place on September 1, 2025, and the final assessment was completed on November 30, 2025.

Research Materials

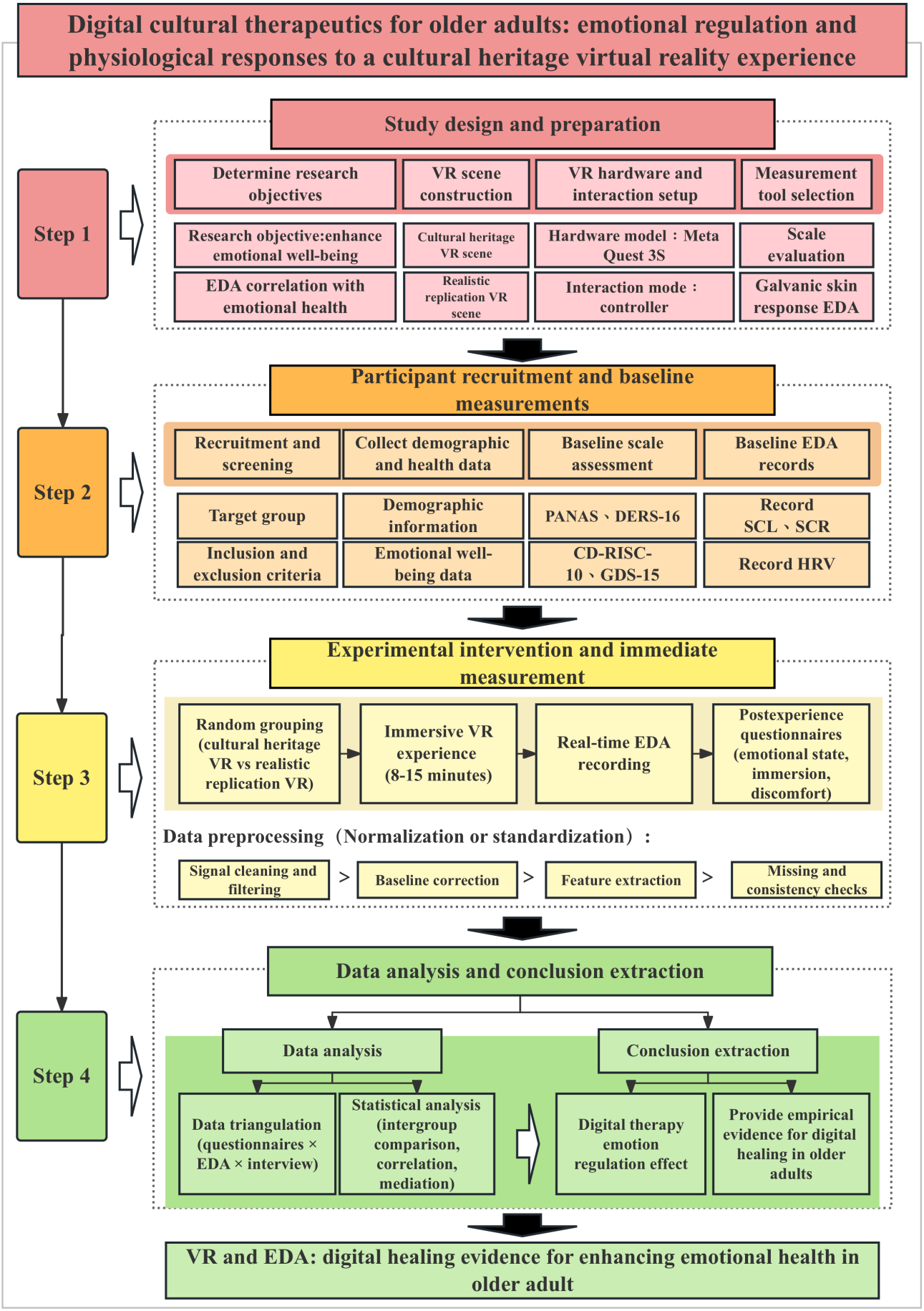
Overview

The study used VR equipment, standardized psychometric questionnaires, and multimodal physiological recording devices to establish an integrated assessment framework combining subjective and objective measures (Figure 1).

XSL•FO
RenderX

Figure 1. Experimental design of the cultural heritage VR digital therapeutics study. CD-RISC-10: 10-Item Connor-Davidson Resilience Scale; DERS-16: 16-Item Difficulties in Emotion Regulation Scale; EDA: electrodermal activity; GDS-15: 15-Item Geriatric Depression Scale; HRV: heart rate variability; PANAS: Positive and Negative Affect Schedule; SCL: skin conductance level; SCR: skin conductance response; VR: virtual reality.

XSL•FO
RenderX



XSL•FO
RenderX

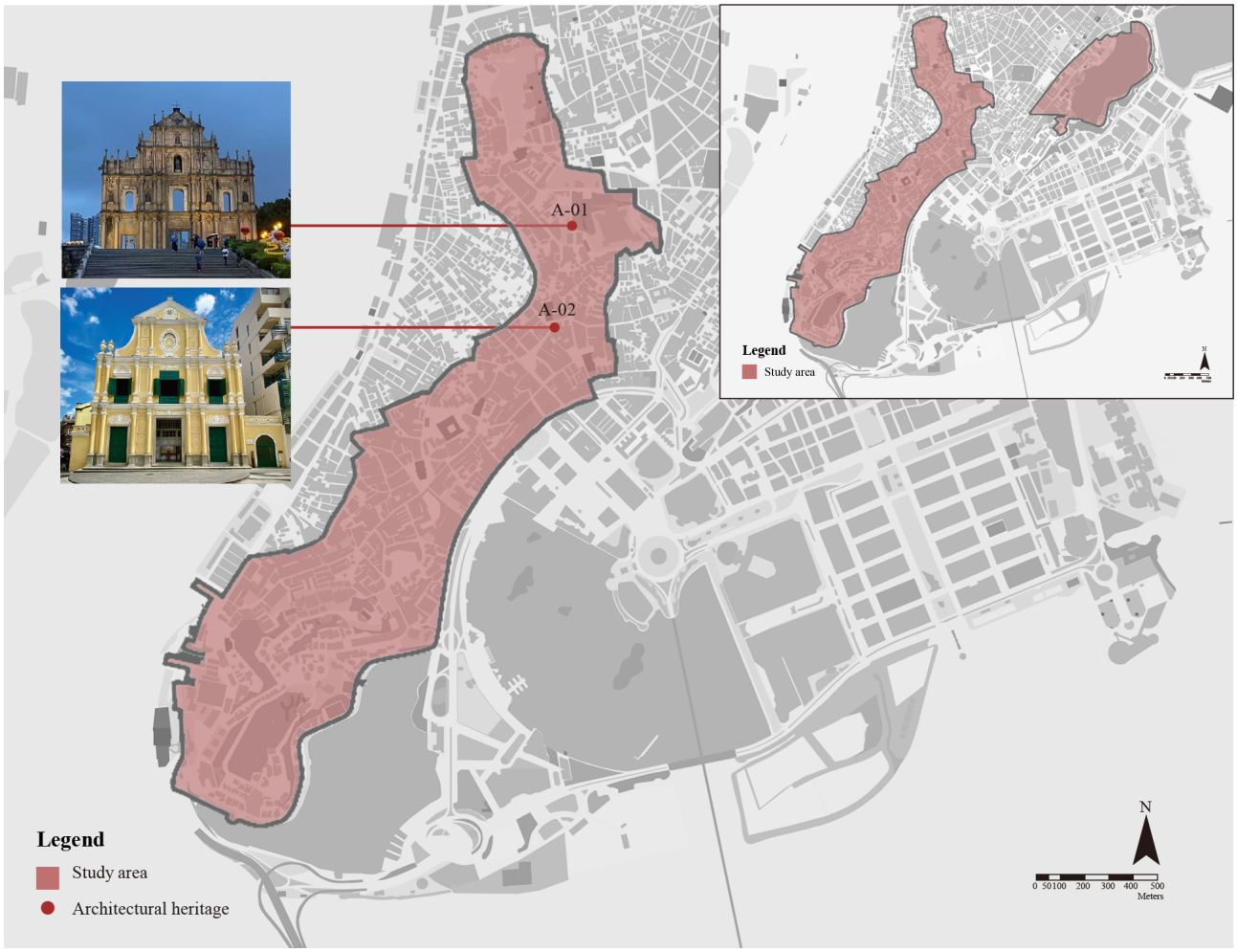
VR Scene Materials

This study adopted a mixed methods design that combined subjective self-report measures with objective physiological signals to examine the effects of cultural heritage–based VR experiences on emotional health in older adults.

The VR environment was developed based on the Historic Centre of Macao, a United Nations Educational, Scientific and Cultural Organization World Heritage Site, selected due to its rich cultural significance and well-preserved architectural heritage. The inclusion criteria for the VR environment were (1) cultural and historical relevance, (2) visual richness and environmental complexity, and (3) the potential to evoke emotional and cognitive engagement. These criteria were informed by prior research suggesting that immersive and meaningful environments can enhance user engagement and modulate emotional and physiological responses in VR settings [36,37].

The experimental framework comprised 2 parallel VR conditions: a cultural heritage VR scene and a realistic replica VR scene. Both conditions were modeled on the same landmarks—the Ruins of St. Paul's (A01) and St. Dominic's Church (A02)—to ensure consistency in spatial content across conditions (Figure 2).

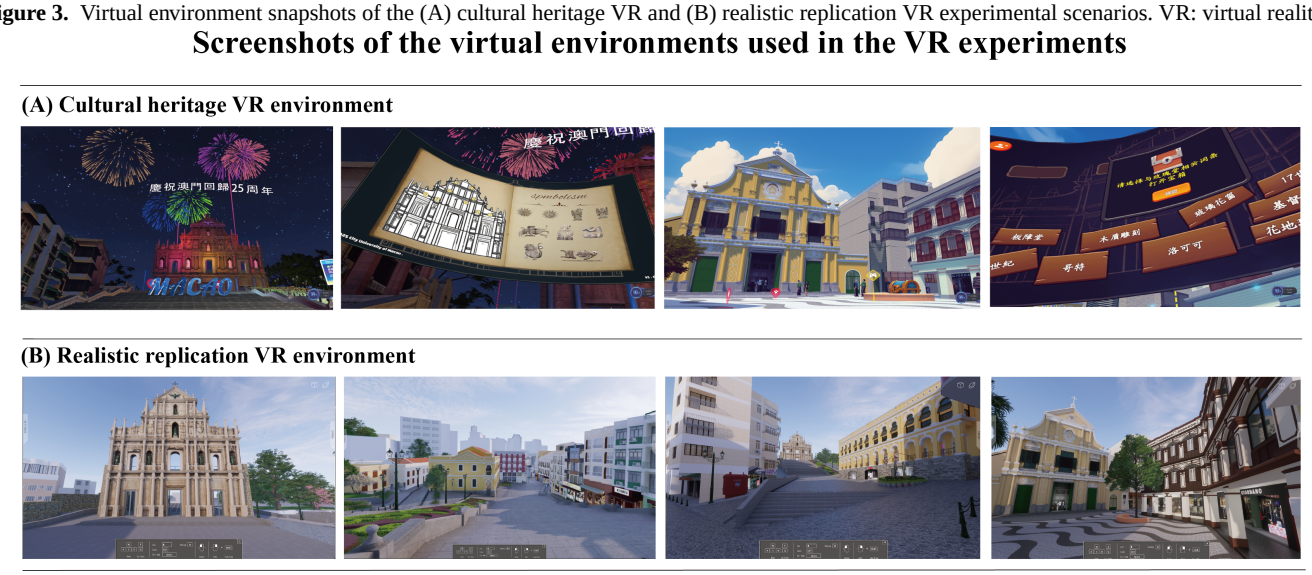
Figure 2. Study area.



Participants were randomly assigned to either the cultural heritage VR group or the realistic replica VR group. Prior to the experiment, participants were blinded to their group allocation. The cultural heritage VR scenes incorporated creative enhancements, including visual aesthetic augmentation, soundscape design, light interactive elements, gamified tasks, and embedded cultural information. These features were intended to increase immersion, cultural engagement, and the elicitation of positive emotional responses among older participants. In contrast, the realistic replica VR scenes served as the control condition and were designed to provide a neutral, realistic replication reproduction of the same spaces. These scenes included no background music, no gamified interactions, and no cultural prompts. Visual presentation, auditory input, and interactivity were deliberately kept minimal to control for artistic and experiential factors. Representative snapshots of the virtual environments used in the study are presented in Figure 3.

XSL•FO
RenderX

Figure 3. Virtual environment snapshots of the (A) cultural heritage VR and (B) realistic replication VR experimental scenarios. VR: virtual reality.



Equipment and Technical Setup

The experimental setup comprised immersive VR hardware and physiological signal acquisition devices. VR experiences were delivered using a head-mounted display (Meta Quest 3/3S), which offers 4K-level resolution, a refresh rate of 90 - 120 Hz, and 6 degrees of freedom tracking, ensuring stable and highly immersive visual performance. All VR content was developed by the research team led by PW at the School of Innovative Design, City University of Macau.

To ensure comparability across experimental conditions, the cultural heritage VR and realistic replica VR scenes were matched in spatial structure, viewpoint trajectories, and lighting parameters. The only differences between conditions concerned the presence or absence of artistic enhancements. Specifically, the cultural heritage VR condition incorporated environmental music, narrative voice prompts, cultural knowledge pop-ups, light interactive tasks, and gamified feedback, whereas these elements were deliberately excluded from the realistic replica condition.

Physiological data were collected using the EryLife S2 multimodal sensing device (Wecare Medical Co), which enables synchronized recording of EDA and HRV signals [38]. The sampling frequency for EDA was set at 4 Hz. Given that EDA signals primarily consist of slow-varying tonic and phasic components, relatively low sampling rates have been shown to be sufficient for capturing meaningful skin conductance responses [39]. Photoplethysmography signals were collected at 100 Hz for HRV analysis. This sampling rate is consistent with previous studies demonstrating that ≥100 Hz provides adequate temporal resolution for reliable interbeat interval detection and HRV estimation, in line with established HRV measurement principles [40]. EDA was measured using medical-grade dry electrodes attached to the inner side of the nondominant wrist, ensuring signal stability and wearing comfort for older participants. HRV data were obtained via a photoplethysmography sensor, providing complementary information on sympathetic and parasympathetic nervous system activity.

Psychological Measures

To systematically assess baseline emotional health prior to the VR intervention, psychological questionnaires were categorized into three functional domains:

1. Emotional state indicators: Positive and Negative Affect Schedule (PANAS),
2. Emotional burden indicators: Geriatric Depression Scale-15 and University of California, Los Angeles Loneliness Scale-8, and
3. Regulatory trait indicators: Difficulties in Emotion Regulation Scale-16 and Connor-Davidson Resilience Scale-10.

This classification facilitated the distinction between short-term emotional changes (state variables) and relatively stable emotional characteristics (trait variables) in subsequent analyses and supported the examination of potential mechanisms underlying VR intervention effects across different psychological dimensions.

The 20-Item Positive and Negative Affect Schedule [41] was used to assess changes in positive affect (PA) and negative affect (NA) before and after the VR experience. The 16-Item Difficulties in Emotion Regulation Scale (DERS-16) [42] measured individual differences in emotion regulation difficulties, allowing exploration of how regulatory capacity may moderate VR-related emotional benefits. The 10-Item Connor-Davidson Resilience Scale (CD-RISC-10) [43] was administered to evaluate psychological resilience and its potential influence on emotional improvement. The 15-Item Geriatric Depression Scale (GDS-15) [44] served as a screening tool for depressive symptoms in older adults and was used to control for baseline emotional burden. The short version of the University of California, Los Angeles Loneliness Scale [45] assessed perceived social loneliness, enabling further examination of whether immersive VR experiences could alleviate loneliness-related emotional distress.

In addition, demographic variables (eg, living arrangements) and behavioral information (eg, prior exposure to VR) were collected via questionnaire.

XSL•FO
RenderX

All psychological scales used in this study are well-established instruments with demonstrated validity in prior research. To assess internal consistency within the current sample, reliability analyses were conducted by the research team using Cronbach α coefficients in SPSS (IBM Corp). The results indicated satisfactory internal consistency across all measures, with Cronbach α values exceeding the recommended threshold of 0.70.

Experimental Procedure

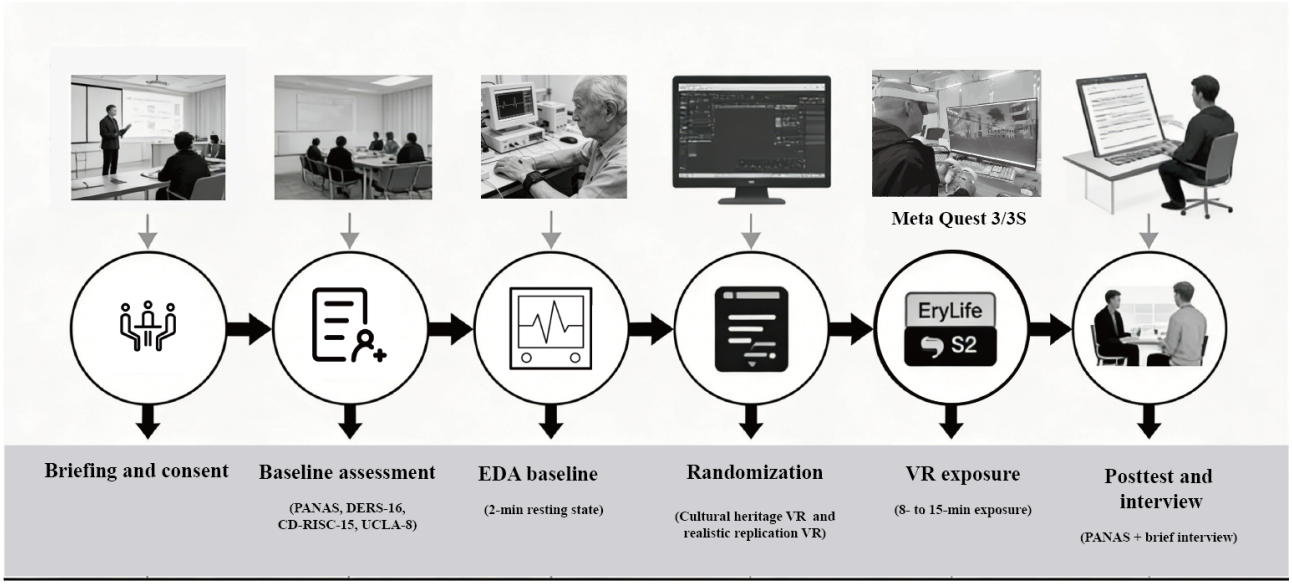
Overview of the Procedure

Upon arrival at the laboratory, all participants received a standardized explanation of the study procedures and provided written informed consent. Participants then completed baseline psychological assessments, including PANAS, DERS-16, CD-RISC-10, GDS-15, and 8-Item University of California, Los Angeles Loneliness Scale (UCLA-8), to establish initial emotional states, emotional burden, and regulatory and resilience-related traits.

Following the questionnaire assessment, participants underwent a 2-minute resting-state EDA recording in a quiet environment with constant lighting conditions. This baseline recording was used to obtain individual skin conductance level (SCL) and SCR measures. After baseline data collection, participants were randomly assigned by computer algorithm to either the cultural heritage VR group or the realistic replica VR control group.

Participants then engaged in an immersive VR experience lasting approximately 8‐15 minutes using the Meta Quest 3/3S head-mounted display. During the VR session, EDA and heart rate activity were continuously recorded using the EryLife S2 Watch device. Immediately after the VR experience, participants completed the postintervention PANAS assessment and, when appropriate, participated in a brief semistructured interview to supplement subjective experience data (Figure 4).

Figure 4. Experimental procedure (illustrative example). CD-RISC-15: 15-Item Geriatric Depression Scale; DERS-16: 16-Item Difficulties in Emotion Regulation Scale; PANAS: Positive and Negative Affect Schedule; UCLA-8: 8-Item University of California, Los Angeles Loneliness Scale; VR: virtual reality.



EDA Signal Preprocessing and Feature Extraction

EDA data were continuously collected during the VR experience using the EryLife S2 device and processed through a multistage standardized pipeline. First, the raw signals underwent automated artifact detection and low-pass filtering to remove motion-related noise and signal distortions caused by poor electrode contact. Second, baseline correction was performed using each participant's resting-state SCL, ensuring comparability across individuals. Subsequently, key EDA features were extracted from the cleaned signals, including SCR frequency, peak amplitude, temporal characteristics, and area under the curve. Finally, *z* score normalization and percentage transformation were applied to optimize data distribution properties and to provide a robust basis for subsequent statistical analyses.

RQs and Analytical Strategy

To address the RQs, a controlled experimental design was implemented. Specifically, RQ1 was examined by comparing emotional health outcomes between the cultural heritage VR condition and the realistic replication VR condition using self-report measures. RQ2 was investigated by analyzing physiological emotional responses based on EDA signals recorded during the VR experience. RQ3 was explored by examining the relationships between subjective emotional responses and physiological indicators, with the aim of identifying potential mechanisms underlying emotional health improvements associated with VR exposure.

Data Processing and Analysis

A multilevel analytical strategy integrating quantitative and qualitative approaches was used to examine the effects of artistic VR experiences on emotional health in older adults and to explore their underlying physiological mechanisms. First, statistical analyses were conducted on questionnaire data,

XSL•FO
RenderX

including PANAS, DERS-16, CD-RISC-10, GDS-15, and the 8-Item University of California, Los Angeles Loneliness Scale. Second, to mitigate potential interpretive bias associated with reliance on a single physiological indicator, EDA and HRV signals were jointly analyzed. This multimodal autonomic assessment enabled evaluation of VR-related emotional arousal across multiple dimensions of autonomic nervous system activity. Physiological measures were aggregated at the individual level prior to statistical testing to enhance the robustness and interpretability of the results. Finally, a data triangulation approach was applied to integrate findings from self-report measures, physiological indicators, and semistructured interviews, providing a comprehensive assessment of emotional responses to the VR experience.

To examine changes in emotional states associated with VR exposure, repeated-measures analysis of variance (RM-ANOVA) was conducted on PANAS PA and NA scores, with time (pre- vs postintervention) treated as the within-subject factor and group (cultural heritage VR vs realistic replication VR) as the between-subject factor. Because the within-subject factor consisted of only 2 levels, the assumption of sphericity for RM-ANOVA was inherently satisfied.

Results

Physiological Responses

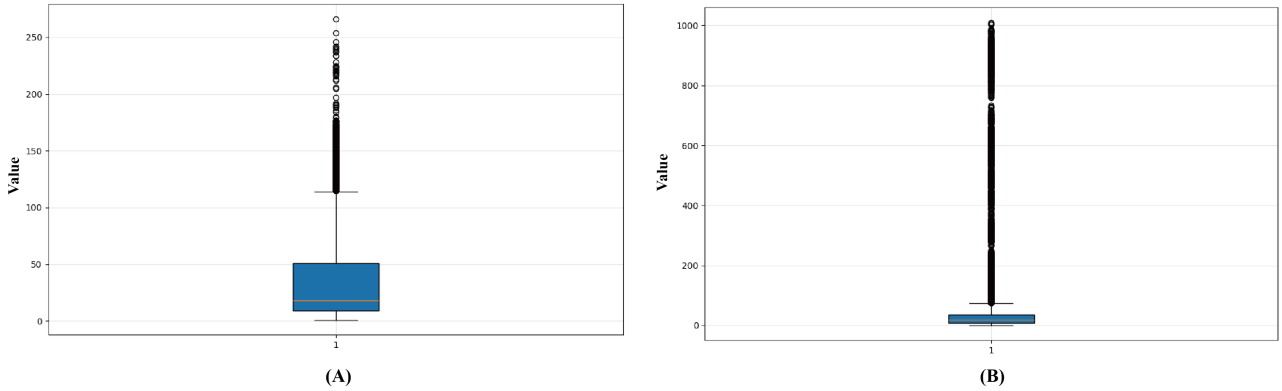
Overview

To objectively compare autonomic nervous system responses under different VR conditions in older adults, between-group analyses were conducted using HRV and EDA indicators. For all measures, change scores (Δ=post–pre) were calculated at the individual level and used for statistical comparisons between groups. Continuous variables were expressed as mean (SD). Descriptive analyses were first performed to examine the distributional characteristics of the physiological data. Boxplots were used to visualize the median, IQR, and dispersion of each indicator. Subsequently, nonparametric Mann-Whitney *U* tests were applied to both HRV and EDA measures to assess between-group differences.

EDA Result

Descriptive analyses revealed clear differences in the distributional properties of EDA between the cultural heritage VR and realistic replica VR conditions. Examination of the boxplots (Figure 5) allowed further differentiation of the sources of these distributional differences.

Figure 5. Distributional characteristics of electrodermal activity under different VR conditions: (A) cultural heritage VR and (B) realistic replication VR. VR: virtual reality.



Under the cultural heritage VR condition, the EDA box (representing the IQR, ie, the middle 50% of observations) was relatively compact, with the median located at a low level (approximately 25), indicating that baseline SCRs were clustered within a narrow range for most participants. However, the upper tail of the distribution was markedly extended and accompanied by several high-value outliers. This pattern indicates that a small subset of participants exhibited SCRs substantially higher than the group norm, thereby considerably increasing overall dispersion.

In the realistic replica VR condition, the EDA box was similarly concentrated in a low-value range, with a relatively low median, suggesting generally stable and homogeneous physiological responses across participants. Although a few high-value outliers were also observed, their number and influence on the overall distribution were limited, resulting in a lower degree of dispersion compared with the cultural heritage VR condition.

Overall, the primary differences between the 2 VR conditions lay in distributional dispersion and upper-tail structure rather than in a pronounced shift in central tendency. These findings suggest that cultural heritage VR experiences are more likely to elicit pronounced interindividual variability in autonomic responses, whereas realistic replica VR experiences are associated with more uniform and stable EDA response patterns.

The between-group comparisons of EDA indicators are presented in Table 2. No statistically significant differences were observed between the 2 VR conditions across all EDA measures, including mean EDA (P=.68; Cohen d=−0.16), SCL (P=.66; Cohen d=−0.16), SCR amplitude (P=.80; Cohen d=−0.09), and SCR peak rate (P=.87; Cohen d=0.06). All effect sizes were small ($|d|$<0.20), indicating limited practical differences in EDA responses between the 2 conditions.

XSL•FO
RenderX

Table . Group comparisons of electrodermal activity (EDA) indicators.

Signal source and feature	Cultural heritage VR[a] (n=25), mean (SD)	Realistic replication VR (n=24), mean (SD)	*P* value	Cohen *d*
EDA				
Mean EDA	1.58 (4.09)	2.61 (8.75)	.68	−0.16
SCL[b]	1.52 (4.12)	2.61 (8.77)	.66	−0.16
SCR[c] amplitude	−0.04 (0.44)	0.02 (0.81)	.80	−0.09
SCR peak rate	−0.29 (4.33)	−0.56 (4.67)	.87	0.06

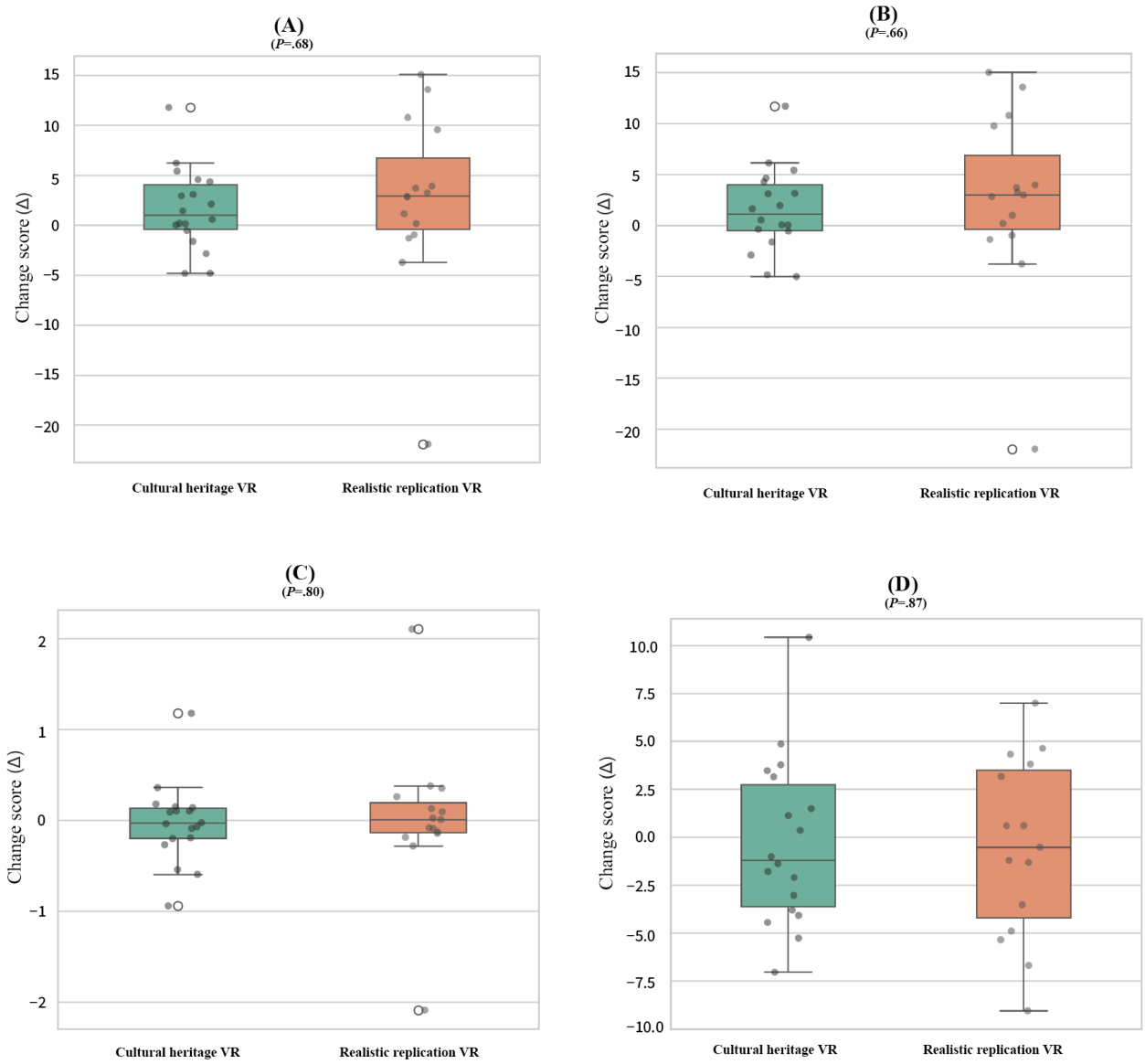
[a]VR: virtual reality.

[b]SCL: skin conductance level.

[c]SCR: skin conductance response.

Distributional analyses (Figure 6) further showed substantial overlap in medians and IQRs across groups, with no clear separation between conditions. Although several indicators exhibited occasional extreme values, these variations were primarily observed at the individual level and did not translate into consistent group-level effects.

Figure 6. Changes in EDA indicators across VR conditions. (A) Mean EDA, (B) SCL, (C) SCR amplitude, and (D) SCR peak rate. EDA: electrodermal activity; HRV: heart rate variability; SCL: skin conductance level; SCR: skin conductance response; VR: virtual reality.



XSL•FO
RenderX

From a physiological perspective, SCL reflects tonic sympathetic arousal, whereas SCR-related measures (amplitude and peak rate) capture phasic, event-related responses. These findings suggest that, under the present experimental conditions, EDA indicators did not demonstrate consistent group-level differentiation, potentially reflecting high interindividual variability and the complex nature of emotional arousal in immersive environments.

HRV Result

The between-group comparisons of HRV indicators are presented in Table 3. The results showed that a significant difference was observed for the mean HRV-derived value between the 2 VR conditions (*P*=.01; Cohen *d*=0.91), with higher change scores in the cultural heritage VR group (mean 21.48, SD 30.39) compared to the realistic replication VR group (mean −2.69, SD 21.03).

Table . Group comparisons of heart rate variability (HRV) indicators.

Signal source and feature	Cultural heritage VR[a] (n=25), mean (SD)	Realistic replication VR (n=24), mean (SD)	*P* value	Cohen *d*
HRV				
Mean NN[b] interval	21.48 (30.39)	−2.69 (21.03)	.01	0.91
SDNN[c]	0.71 (11.26)	12.40 (20.00)	.06	−0.74
RMSSD[d]	−0.96 (20.70)	3.52 (14.89)	.48	−0.24

[a]VR: virtual reality.

[b]NN: normal-to-normal.

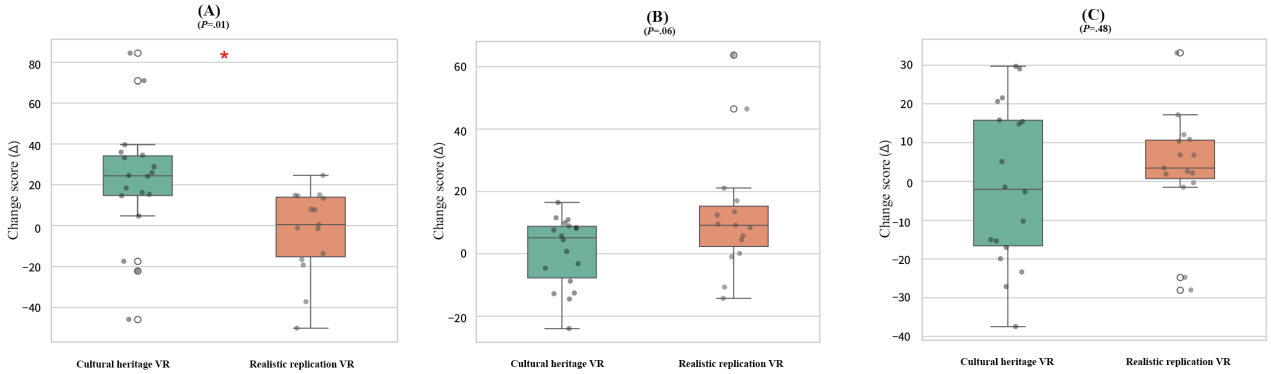
[c]SDNN: standard deviation of NN.

[d]RMSSD: root-mean-square of successive difference.

In contrast, conventional time-domain HRV indices, including the standard deviation of NN intervals (SDNN) and the root-mean-square of successive differences (RMSSD), did not show statistically significant differences between groups. Although SDNN exhibited a numerical difference (*P*=.06; Cohen *d*=−0.74), it did not reach statistical significance. Similarly, no significant difference was found for RMSSD (*P*=.48; Cohen *d*=−0.24).

Distributional analyses (Figure 7) indicated that the mean normal-to-normal (NN) interval (mean RR interval) tended to be higher under the cultural heritage VR condition, whereas substantial overlap was observed between groups for SDNN and RMSSD. Overall, the between-group difference was primarily reflected in the mean NN interval, while no consistent pattern was observed across standard time-domain HRV measures.

Figure 7. Changes in heart rate variability indicators across VR conditions. (A) Mean NN interval, (B) SDNN, and (C) RMSSD. NN: normal-to-normal; RMSSD: root-mean-square of successive differences; SDNN: standard deviation of NN; VR: virtual reality.



From a physiological perspective, RMSSD is generally considered to reflect parasympathetic activity, whereas SDNN represents overall HRV. However, as these conventional indices did not show significant differences, the physiological interpretation of the observed group difference should be approached with caution. The findings suggest that the effects of different VR conditions on autonomic regulation may be limited or influenced by substantial interindividual variability.

Subjective Emotional Outcomes and Psychological Measures

Baseline Sample Characteristics and Between-Group Comparability

To ensure comparability between the cultural heritage VR group and the realistic replica VR group prior to intervention, chi-square tests were conducted on participants' demographic and health-related background variables. A total of 12 indicators were examined, including sex, age group, educational attainment, marital status, living arrangement, occupational

XSL•FO
RenderX

background, economic status, visual and auditory conditions, sleep quality, experience with smart devices, and prior exposure to VR or digital therapeutic applications (Table 4).

No statistically significant differences were observed between the 2 groups for any of the variables examined (χ^2 tests; all P>.05). These results indicate a high degree of baseline equivalence between groups in terms of sociodemographic structure, health status, and technology-related experience, thereby minimizing potential confounding effects attributable to sample composition and providing a reliable baseline for subsequent analyses of emotional and physiological outcomes.

XSL•FO
RenderX

Table . Results of the chi-square analysis.

Item and response options	Group		Chi-square (*df*)	*P* value
	Cultural heritage VR[a], n (%)	Realistic replication VR, n (%)		
Sex			0.2 (1)	.67
Female	14 (56)	12 (50)		
Male	11 (44)	12 (50)		
Age group (years)			0.9 (3)	.84
60 - 64	11 (44)	8 (33)		
65 - 69	5 (20)	5 (21)		
70 - 74	5 (20)	5 (21)		
75 and above	4 (16)	6 (25)		
Educational attainment			1.7 (2)	.44
Junior college (associate degree)	5 (20)	4 (17)		
Bachelor's degree	5 (20)	2 (8)		
High school or below	15 (60)	18 (75)		
Marital status			2.0 (2)	.37
Widowed	1 (4)	0 (0)		
Married	23 (92)	24 (100)		
Unmarried	1 (4)	0 (0)		
Living arrangement			4.0 (3)	.26
Living with children	11 (44)	14 (58)		
Living with spouse	8 (32)	5 (21)		
Living alone	0 (0)	2 (8)		
Other	6 (24)	3 (13)		
Current occupational status			3.1 (3)	.38
Other	2 (8)	2 (8)		
Retired	20 (80)	22 (92)		
Self-employed or freelance	2 (8)	0 (0)		
Re-employed after retirement	1 (4)	0 (0)		
Economic status			5.3 (3)	.15
Generally sufficient	13 (52)	19 (79)		
Very difficult	1 (4)	0 (0)		
Somewhat difficult	2 (8)	0 (0)		
Relatively comfortable	9 (36)	5 (21)		
Visual condition			2.2 (2)	.33
Slightly blurred but adequate for daily activities without glasses	6 (24)	3 (12)		
Normal vision without glasses	18 (72)	21 (88)		
Requires glasses for normal	1 (4)	0 (0)		
Hearing condition			0.0 (1)	.98

XSL•FO
RenderX

Item and response options	Group		Chi-square (*df*)	*P* value
	Cultural heritage VR[a], n (%)	Realistic replication VR, n (%)		
Does not require a hearing aid but needs louder or repeated speech to hear clearly	1 (4)	1 (4)		
Normal hearing without a hearing aid	24 (96)	23 (96)		
Sleep quality			1.8 (3)	.62
Fair	7 (28)	5 (21)		
Very good	7 (28)	7 (29)		
Poor	3 (12)	1 (4)		
Average	8 (32)	11 (46)		
Use of smart devices (smartphone or tablet)			0.0 (1)	.98
No	1 (4)	1 (4)		
Yes	24 (96)	23 (96)		
Prior exposure to VR or digital therapeutic products			1.6 (1)	.20
No	23 (92)	19 (79)		
Yes	2 (8)	5 (21)		

[a]VR: virtual reality.

Between-Group Differences in Baseline Psychological Traits

Prior to the VR intervention, 2-tailed independent-samples *t* test were performed to compare baseline psychological trait measures between groups, including emotion regulation difficulties (DERS-16), psychological resilience (CD-RISC-10), and depressive symptoms (GDS-15). No significant between-group differences were detected for emotion regulation difficulties (t_{47}=−1.08; *P*=.29), psychological resilience ($t_{25.538}$=1.869; *P*=.07), or depressive symptoms ($t_{31.212}$=−0.688; *P*=.49; Table 5). These findings further confirm baseline homogeneity between the 2 groups at the psychological trait level, establishing an appropriate foundation for evaluating VR-induced emotional changes.

Table . Results of 2-tailed independent-samples *t* tests.

Variable	Group		*t* test (*df*)	*P* value
	Cultural heritage VR[a] experience (n=25), mean (SD)	Realistic replication VR experience (n=24), mean (SD)		
Emotion regulation difficulties	39.00 (8.48)	41.67 (8.81)	−1.08 (47)	.29
Psychological resilience	39.48 (1.19)	38.13 (3.35)	1.869 (25.538)	.07
Depression	23.44 (0.77)	23.71 (1.76)	−0.688 (31.212)	.49

[a]VR: virtual reality.

Effects of VR Intervention on PA and NA

To examine the effects of VR exposure on emotional state, RM-ANOVA was conducted on PANAS PA and NA scores, with time (pre- vs postintervention) as the within-subject factor and group (cultural heritage VR vs realistic replica VR) as the between-subject factor.

Descriptive statistics are presented in Table 6. At baseline, PA and NA scores were comparable between the 2 groups. Following the intervention, PA increased and NA decreased in both groups, with more pronounced changes observed in the cultural heritage VR group.

For PA, RM-ANOVA revealed a significant main effect of time ($F_{1,47}$=211.20; *P*<.001; η^2=0.818), indicating an overall increase following VR exposure. A significant main effect of group ($F_{1,47}$=5.67; *P*=.02; η^2=0.108) and a significant time×group interaction ($F_{1,47}$=12.47; *P*=.001; η^2=0.210) were also observed (Table 7).

XSL•FO
RenderX

Table . Descriptive statistics of positive affect (PA) and negative affect (NA) scores before and after the virtual reality (VR) intervention.

Measure and time	Cultural heritage VR experience (n=25), mean (SD)	Realistic replication VR experience (n=24), mean (SD)
PA		
Preintervention	30.96 (1.54)	30.79 (2.86)
Postintervention	43.48 (4.36)	38.42 (7.34)
NA		
Preintervention	19.00 (2.24)	21.38 (6.44)
Postintervention	10.32 (1.11)	17.46 (6.98)

Table . Repeated-measures ANOVA results.

Measure and effect	Type III sum of squares	*df*	Mean square	*F* test (*df*=1, 47)	*P* value	Effect size (η^2)
PA[a]						
Time	2484.619	1	2484.619	211.196	<.001	0.818
Group	167.573	1	167.573	5.674	.02	0.108
Time×group	146.7	1	146.7	12.47	.001	0.210
NA[b]						
Time	971.486	1	971.486	118.709	<.001	0.716
Group	554.103	1	554.103	14.151	<.001	0.231
Time×group	138.914	1	138.914	16.974	<.001	0.265

[a]PA: positive affect.

[b]NA: negative affect.

For NA, a similar pattern was observed. RM-ANOVA showed a significant main effect of time ($F_{1,47}$=118.71; P<.001; η^2=0.716), along with significant main effects of group ($F_{1,47}$=14.15; P<.001; η^2=0.231) and a significant time× group interaction ($F_{1,47}$=16.97; P<.001; η^2=0.265; Table 7).

Simple effects analyses (Tables 8 and 9) revealed a consistent pattern of significant time effects and differential group effects over time. At baseline, no significant between-group differences were observed in either PA or NA. Following the intervention, PA increased significantly in both the cultural heritage VR experience group and the realistic replication VR experience group; however, the increase was significantly greater in the cultural heritage VR group, resulting in higher postintervention PA compared with the control condition. Conversely, NA decreased significantly in both groups across time, with a significantly larger reduction observed in the cultural heritage VR group, which also exhibited significantly lower postintervention NA relative to the realistic replication VR group.

Table . Within-group simple effects analysis of positive affect (PA) and negative affect (NA) across time points[a].

Measure and group	Comparison	Mean difference (SE)	*P* value
PA			
Cultural heritage VR[b] experience	Pre versus post	−12.520 (0.970)	<.001
Realistic replication VR experience	Pre versus post	−7.625 (0.990)	<.001
NA			
Cultural heritage VR experience	Pre versus post	8.680 (0.809)	<.001
Realistic replication VR experience	Pre versus post	3.917 (0.826)	<.001

[a]Time 1=preintervention; time 2=postintervention.

[b]VR: virtual reality.

XSL•FO
RenderX

Table . Between-group simple effects analysis of positive affect (PA) and negative affect (NA) at each time point[a].

Measure and time	Comparison	Mean difference (SE)	*P* value
PA			
Pre	Cultural heritage VR[b] experience versus realistic replication VR experience	0.168 (0.652)	.80
Post	Cultural heritage VR experience versus realistic replication VR experience	5.063 (1.717)	.005
NA			
Pre	Cultural heritage VR experience versus realistic replication VR experience	−2.375 (1.366)	.09
Post	Cultural heritage VR experience versus realistic replication VR experience	−7.138 (1.414)	<.001

[a]Time 1=preintervention; time 2=postintervention.

[b]VR: virtual reality.

Associations Between Psychological Traits and Emotional Changes

To further explore individual differences in emotional responses to VR, Pearson correlation analyses were conducted to examine associations among emotion regulation difficulties, psychological resilience, depressive symptoms, and emotional state before and after the intervention (Table 10).

XSL•FO
RenderX

Table . Pearson correlation analysis between emotion regulation difficulty, depression, and emotional indicators before and after the intervention.

Variable and statistic	Difficulties in emotion regulation	Psychological resilience	Depression	Preintervention PA[a]	Preintervention NA[b]	Postintervention PA	Postintervention NA
Difficulties in emotion regulation							
r	1	−0.530	0.526	−0.525	0.638	−0.625	0.425
P value	—[c]	<.001	<.001	<.001	<.001	<.001	.002
Psychological resilience							
r	−0.530	1	−0.649	0.600	−0.754	0.770	−0.477
P value	<.001	—	<.001	<.001	<.001	<.001	<.001
Depression							
r	0.526	−0.649	1	−0.761	0.754	−0.661	0.441
P value	<.001	<.001	—	<.001	<.001	<.001	.002
Preintervention PA							
r	−0.525	0.600	−0.761	1	−0.699	0.610	−0.327
P value	<.001	<.001	<.001	—	<.001	<.001	.02
Preintervention NA							
r	0.638	−0.754	0.754	−0.699	1	−0.736	0.657
P value	<.001	<.001	<.001	<.001	—	<.001	<.001
Postintervention PA							
r	−0.625	0.770	−0.661	0.610	−0.736	1	−0.609
P value	<.001	<.001	<.001	<.001	<.001	—	<.001
Postintervention NA							
r	0.425	−0.477	0.441	−0.327	0.657	−0.609	1
P value	.002	<.001	.002	.02	<.001	<.001	—

[a]PA: positive affect.

[b]NA: negative affect.

[c]Not available.

Emotion regulation difficulties were significantly negatively correlated with psychological resilience (r=−0.530; P<.01) and positively correlated with depressive symptoms (r=0.526; P<.01). In addition, higher emotion regulation difficulties were associated with lower PA and higher NA both before and after the VR experience, indicating greater emotional burden among individuals with poorer regulation capacity.

Psychological resilience showed the opposite pattern, exhibiting significant positive correlations with PA and significant negative correlations with NA and depressive symptoms. These findings suggest that resilience may serve a buffering and facilitative role in VR-based emotional interventions.

Convergence Between Subjective Emotional Outcomes and EDA Responses

Comparative analysis of subjective emotional changes and EDA responses revealed consistent patterns across modalities. Relative to the realistic replica VR condition, participants exposed to cultural heritage VR exhibited higher and more sustained levels of skin conductance during the experience, alongside larger increases in PA and more pronounced reductions in NA on self-report measures.

This convergence between subjective emotional outcomes and autonomic physiological responses indicates coherent emotion-related effects across measurement modalities under different VR content conditions, providing an empirical basis for subsequent multimodal interpretations of VR-induced emotional regulation effects.

Discussion

Convergence of Subjective and Physiological Evidence and the Emotional Regulation Potential of Cultural Heritage VR

This study provides an exploratory evaluation of the role of cultural heritage VR in promoting emotional health among older adults by integrating subjective emotional measures with physiological indicators. The findings indicate that, compared with the visually neutral realistic replication VR condition, cultural heritage VR was associated with higher PA and lower NA. At the physiological level, only the mean NN interval showed a statistically significant between-group difference. In contrast, no significant differences were observed in EDA measures (all P>.05). However, EDA measures showed

XSL•FO
RenderX

substantial interindividual variability in their distribution patterns. It should be noted that the mean NN interval reflects overall cardiac cycle length rather than conventional HRV variability indices.

From the perspective of emotional arousal theory, changes in EDA are widely regarded as sensitive indicators of sympathetic nervous system activation and are closely linked to emotional engagement and attentional resource allocation [39,46,47]. The cultural heritage VR condition in this study incorporated enhanced visual aesthetics, soundscape design, narrative cues, and light interactive elements. These features may have facilitated affective engagement and sustained attention among older participants. However, this potential emotional arousal did not translate into stable group-level differences in EDA measures. Instead, it was reflected as variability at the individual level. Notably, although the physiological findings were limited, subjective emotional improvements showed a clearer pattern of differentiation between the 2 conditions. The discrepancy between subjective emotional improvement and physiological responses suggests a potential dissociation between perceived experience and autonomic activation.

This dissociation may be explained by several mechanisms. First, EDA primarily reflects sympathetic arousal. It may be less sensitive to changes in emotional valence when arousal levels remain moderate. In this study, the cultural heritage VR condition appeared to enhance PA without inducing strong physiological arousal, resulting in limited differentiation in EDA responses at the group level. Second, emotional engagement in VR environments may operate through cognitive-affective pathways, such as meaning-making and narrative immersion, which are more strongly reflected in subjective reports than in autonomic activation. Third, substantial interindividual variability in EDA responses may have obscured potential group-level effects. This variability may be particularly pronounced in older populations, where physiological responsiveness differs considerably.

Comparison With Previous Research: Cultural Heritage VR, Digital Therapeutics, and Emotional Health in Older Adults

Previous studies have widely reported that VR-based interventions can improve negative emotional states, enhance subjective well-being, and increase psychological engagement among older adults. These applications have primarily focused on rehabilitation training, cognitive stimulation, and social participation [11,48,49]. In such studies, emotional improvements are often attributed to immersion, novelty, and heightened attentional engagement afforded by VR technology [50,51].

However, much of the existing literature conceptualizes VR primarily as a technological medium, with limited systematic discussion of how artistic expression and cultural meaning embedded in virtual content contribute to emotional regulation [38]. In contrast, this study explicitly distinguishes between cultural heritage VR and realistic replica VR. Empirical evidence demonstrates that different types of virtual content do not elicit equivalent emotional responses. Emotional changes in older participants were not driven solely by VR technology itself but were more strongly influenced by artistic expression, cultural narrative structures, and experiential guidance embedded within the virtual environment. This finding extends existing VR intervention research by highlighting the emotional consequences of content differentiation.

Recent studies have begun incorporating music, narrative, or aesthetic elements into VR interventions and have reported positive outcomes at the subjective level. Nonetheless, most rely heavily on self-report measures and lack objective physiological validation [52-54]. By synchronously collecting EDA data, this study provides a complementary perspective to the existing literature. However, in contrast to prior studies reporting significant physiological changes, no significant between-group differences were observed in EDA indicators in this study. This discrepancy may be attributable to differences in experimental design and stimulus characteristics. Prior studies reporting significant physiological changes often used high-intensity or event-based stimuli, whereas this study used a more continuous and immersive experiential paradigm. Such differences may lead to distinct patterns of autonomic activation. Continuous experiences may elicit more subtle and distributed physiological responses. These responses may not be readily captured by conventional EDA metrics. This finding suggests that the emotional effects of VR interventions may not necessarily be accompanied by pronounced increases in physiological arousal, but may instead be more prominently reflected through subjective experiential pathways.

Nevertheless, distributional patterns indicate that a subset of participants in the cultural heritage VR condition exhibited relatively higher physiological responses, suggesting a greater potential for emotional activation at the individual level. Accordingly, the emotional effects of VR interventions appear to be partially individual-dependent, with physiological manifestations likely influenced by multiple factors, including individual differences and the degree of engagement with the virtual environment. These results support a shift in digital therapeutics research from a technology-centered effectiveness perspective toward a content- and experience-driven explanatory framework.

Mechanistic Pathways and Design Implications: A Conceptual Framework for Digital Therapeutics

Based on the integrated analysis of subjective emotional assessments and EDA data, this study proposes a conceptual pathway to explain the emotional effects of cultural heritage VR. The pathway emphasizes that culturally meaningful artistic experiences are associated with elevated emotional arousal, which in turn corresponds to changes in emotional states among older adults. It should be noted that this pathway is interpretative and hypothesis-generating rather than a fully validated causal model.

At the experiential level, cultural heritage VR may foster sustained emotional engagement and meaning construction through aesthetic visual language, soundscape design, narrative cues, and light interactive tasks. These features transform virtual environments from purely observable spaces into culturally situated emotional contexts, providing conditions conducive to

XSL•FO
RenderX

emotional activation. At the physiological level, higher levels of affective engagement and sustained attentional states may be reflected in autonomic activation, as indicated by electrodermal responses in a subset of individuals. However, this effect does not manifest as a consistent difference at the group level. This pattern suggests that emotional processing in VR may involve partially independent pathways for subjective experience and physiological activation. Affective engagement and meaning construction may enhance perceived emotional states. However, their translation into measurable autonomic responses may depend on factors such as stimulus intensity, individual sensitivity, and attentional dynamics. Unlike stress-induced or high-intensity stimulation responses, the observed EDA changes are more consistent with moderate, regulatable emotional arousal and co-occur with increased PA and reduced NA. This suggests that physiological arousal elicited by cultural heritage VR is more indicative of emotional participation and meaning experience rather than nonspecific sensory stimulation.

From the perspective of digital therapeutics and design practice, this conceptual pathway carries important implications. The findings underscore that content design in VR interventions is not ancillary but may constitute a primary condition for emotional effects to emerge. Moreover, incorporating physiological measures such as EDA into experience evaluation frameworks can address limitations inherent in exclusive reliance on self-report data and provide a more objective, verifiable basis for emotional health assessment in older adults. Future studies may further refine this pathway by introducing psychological constructs such as immersion and presence to test mediating mechanisms with greater precision. This pathway further suggests that emotional effects may depend on the dynamic coupling among contextual meaning, affective engagement, and physiological responses, rather than being driven by stimulus intensity alone.

Limitations and Future Research Directions

Despite its systematic design and methodological integration, this study has several limitations that warrant further investigation. First, the sample size was relatively modest, and participants were recruited from community settings within a limited geographic region. This may introduce homogeneity in cultural background and technology acceptance, thereby constraining the generalizability of the findings. Future research should incorporate multiregional and cross-cultural samples to assess the broader applicability of cultural heritage VR–based emotional interventions.

Second, the study primarily examined the immediate or short-term emotional effects of a single VR exposure. Long-term outcomes and the stability of emotional benefits remain unclear and should be explored through longitudinal designs with follow-up assessments. Third, although EDA was used as a sensitive indicator of emotional arousal, it cannot fully capture the complexity of autonomic nervous system responses. Future work may integrate additional physiological measures, such as electroencephalography or eye-tracking, to construct a more comprehensive multimodal analytical framework.

Furthermore, while the study emphasizes the emotional value of artistic design in cultural heritage VR, the experimental conditions incorporated multiple experiential elements simultaneously. The cultural heritage VR environment included background music, narrative voice prompts, interactive tasks, and cultural information pop-ups, whereas the photorealistic VR condition provided a comparatively minimal stimulus environment. As a result, the observed emotional differences may not be attributable solely to the cultural heritage content itself. Additional factors such as auditory stimulation, increased interactivity, task engagement, or novelty effects may also have contributed to the observed outcomes. This study therefore cannot fully disentangle the independent and interactive effects of these design components. Future research should adopt more controlled experimental paradigms—such as factorial or modular designs—in which specific elements (eg, auditory cues, interaction levels, narrative guidance, or cultural storytelling) are systematically manipulated in order to clarify their relative contributions to emotional responses.

In addition, this study did not directly assess immersion or sense of presence. Future research could incorporate validated presence scales or structural equation modeling approaches to more precisely examine the mediating role of emotional arousal within this process. Overall, this study provides preliminary but robust empirical support for the application of artistic VR in promoting emotional health among older adults. It also offers a methodological and theoretical foundation for future research at the intersection of design studies, environmental psychology, and digital therapeutics.

Conclusions

This study aimed to examine the effects of cultural heritage–based VR on emotional health in older adults, in comparison with a realistic replication VR condition. In addition, the study explored the relationships between subjective emotional responses, physiological indicators (HRV and EDA), and individual psychological characteristics.

The findings indicate that, compared with the realistic replication VR condition, cultural heritage VR significantly enhanced PA and reduced NA. These results support the hypothesis that culturally meaningful and artistically enriched VR environments can more effectively promote emotional well-being in older adults.

At the physiological level, a significant between-group difference was observed only for the mean NN interval, whereas no statistically significant differences were found in EDA measures. Furthermore, physiological indicators did not demonstrate fully consistent relationships with subjective emotional changes, suggesting that physiological responses to VR may be more variable and influenced by individual differences.

In addition, emotion regulation–related traits and psychological resilience were found to be associated with emotional outcomes, highlighting the importance of individual psychological characteristics in shaping responses to VR interventions. These findings provide preliminary support for a conceptual pathway

XSL•FO
RenderX

linking immersive experience to emotional health, although the underlying mechanisms require further empirical investigation.

This study contributes to the field in 3 main ways. First, it provides empirical evidence supporting the emotional benefits of cultural heritage VR as a potential digital intervention for older adults. Second, it demonstrates the value of integrating subjective measures with physiological indicators to achieve a more comprehensive assessment of emotional responses. Third, it offers a preliminary framework for understanding the mechanisms underlying VR-induced emotional change.

From a practical perspective, the findings inform the design of culturally meaningful and emotion-oriented VR applications for aging populations. Future research should examine the long-term effects of such interventions in larger and more diverse samples and further explore the generalizability of artistic VR approaches across different cultural contexts.

Acknowledgments

The authors thank the research team led by PW at City University of Macau for their support in conducting this study. The authors are also grateful to all older adults who participated in this research and to their families for their time, cooperation, and valuable contributions. The authors further acknowledge the China National Arts Fund for supporting the cultural heritage artistic creation project that provided the visual materials and creative foundation for the cultural heritage virtual reality environment used in this study. Special thanks are extended to Zhang Yangyang, Lu Yao, Dong Yuxuan, and Liu Gaolu for their contributions to the creation and development of the cultural heritage virtual reality materials. The authors used ChatGPT (OpenAI, GPT-5.5) to assist with language editing, grammar refinement, and manuscript organization. All scientific content, data analysis, interpretation of results, and final manuscript revisions were performed and verified by the authors.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated and analyzed during this study are not publicly available because they contain potentially identifiable participant information. Deidentified data may be available from the corresponding author upon reasonable request and subject to institutional ethics approval requirements.

Authors' Contributions

Conceptualization: PW
Methodology: PW
Software: YL
Investigation: YL
Data curation: YL
Validation: YL
Visualization: YL
Writing—original draft: YL
Writing—review and editing: PW
Supervision: PW

Conflicts of Interest

None declared.

References

1. Mental health of older adults. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2026-01-15]
2. Jalali A, Ziapour A, Karimi Z, et al. Global prevalence of depression, anxiety, and stress in the elderly population: a systematic review and meta-analysis. BMC Geriatr 2024 Oct 4;24(1):809. [doi: 10.1186/s12877-024-05311-8] [Medline: 39367305]
3. Zábó V, Csiszar A, Ungvari Z, Purebl G. Psychological resilience and competence: key promoters of successful aging and flourishing in late life. Geroscience 2023 Oct;45(5):3045-3058. [doi: 10.1007/s11357-023-00856-9] [Medline: 37418098]
4. Health effects of social isolation and loneliness. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/social-connectedness/risk-factors/index.html [accessed 2026-01-15]
5. Tuomi A, Moreira Kares E, Zainal Abidin H. Digital cultural tourism: older adults' acceptance and use of digital cultural tourism services. Scand J Hosp Tour 2023 May 27;23(2-3):226-247. [doi: 10.1080/15022250.2023.2256698]

XSL•FO
RenderX

6. McFarland J, Rice C, Changfoot N, et al. Graying arts access: crafting creative online programming to promote older adults' artistic engagement in and beyond pandemic time. Front Sociol 2024;9:1454143. [doi: 10.3389/fsoc.2024.1454143] [Medline: 39698027]
7. Cao Y, Yin H, Hua X, Bi S, Zhou D. Effects of artificial intelligence and virtual reality interventions in art therapy among older people with mild cognitive impairment. Australas J Ageing 2025 Mar;44(1):e70006. [doi: 10.1111/ajag.70006] [Medline: 39992794]
8. Ke Z, Wei M, Yang F, et al. The effectiveness of immersive virtual reality on the psychology of older adults: a systematic review and Meta-analysis of randomized controlled trials. Gen Hosp Psychiatry 2025;94:86-96. [doi: 10.1016/j.genhosppsych.2025.02.024] [Medline: 40036985]
9. Saha S, Dobbins C, Gupta A, Dey A. Differentiating presence in virtual reality using physiological signals. Pervasive Mob Comput 2025 Jun;111:102065. [doi: 10.1016/j.pmcj.2025.102065]
10. Baragash RS, Aldowah H, Ghazal S. Virtual and augmented reality applications to improve older adults' quality of life: a systematic mapping review and future directions. Digit Health 2022;8:20552076221132099. [doi: 10.1177/20552076221132099] [Medline: 36339904]
11. Carl E, Stein AT, Levihn-Coon A, et al. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. J Anxiety Disord 2019 Jan;61:27-36. [doi: 10.1016/j.janxdis.2018.08.003] [Medline: 30287083]
12. D'Cunha NM, Nguyen D, Naumovski N, et al. A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment. Gerontology 2019;65(4):430-440. [doi: 10.1159/000500040] [Medline: 31108489]
13. Ronca V, Ricci A, Capotorto R, et al. How immersed are you? State of the art of the neurophysiological characterization of embodiment in mixed reality for out-of-the-lab applications. Appl Sci 2024;14(18):8192. [doi: 10.3390/app14188192]
14. Ronca V, Babiloni F, Arico P. A novel mutual information-based approach for neurophysiological characterization of sense of presence in virtual reality. IEEE Trans Biomed Eng 2025 Aug;72(8):2313-2320. [doi: 10.1109/TBME.2025.3541438] [Medline: 40031843]
15. Tan MKB, Tan CM. Curating wellness during a pandemic in Singapore: COVID-19, museums, and digital imagination. Public Health 2021 Mar;192:68-71. [doi: 10.1016/j.puhe.2021.01.006] [Medline: 33647787]
16. Yang J, Li Y, Gao D, Xie X, Ji W, Gao J. Effects of virtual reality technology on anxiety and depression in older adults with chronic diseases: a systematic review and meta-analysis of randomized controlled trials. Worldviews Evid Based Nurs 2025 Feb;22(1):e12763. [doi: 10.1111/wvn.12763] [Medline: 39697092]
17. Cieślik B, Juszko K, Kiper P, Szczepańska-Gieracha J. Immersive virtual reality as support for the mental health of elderly women: a randomized controlled trial. Virtual Real 2023 May 7;27(3):1-9. [doi: 10.1007/s10055-023-00797-w] [Medline: 37360811]
18. Huang LC, Yang YH. The long-term effects of immersive virtual reality reminiscence in people with dementia: longitudinal observational study. JMIR Serious Games 2022 Jul 25;10(3):e36720. [doi: 10.2196/36720] [Medline: 35877169]
19. Mao Q, Zhao Z, Yu L, Zhao Y, Wang H. The effects of virtual reality–based reminiscence therapies for older adults with cognitive impairment: systematic review. J Med Internet Res 2024 Nov 12;26:e53348. [doi: 10.2196/53348] [Medline: 39531267]
20. Şansal KE, Şimşek AC, Aktan S, Özbey F, Paksoy A. Restorative effects of virtual nature on the emotional well-being of community-dwelling older adults. Eur J Geriatr Gerontol 2024 Apr 1;6(1):12-18. [doi: 10.4274/ejgg.galenos.2023.2023-3-2]
21. Khirallah Abd El Fatah N, Abdelwahab Khedr M, Alshammari M, Mabrouk Abdelaziz Elgarhy S. Effect of immersive virtual reality reminiscence versus traditional reminiscence therapy on cognitive function and psychological well-being among older adults in assisted living facilities: a randomized controlled trial. Geriatr Nurs 2024;55:191-203. [doi: 10.1016/j.gerinurse.2023.11.010] [Medline: 38007908]
22. Matsangidou M, Solomou T, Frangoudes F, Papayianni E, Pattichis CS. Offering outworld experiences to in-patients with dementia through virtual reality: mixed methods study. JMIR Aging 2023 Aug 31;6(1):e45799. [doi: 10.2196/45799] [Medline: 37656031]
23. Kim J, Lee J, Kim Y, Nuseibeh B, Han S. The effects of a nature-based virtual reality program on emotional health and quality of life among older adults with dementia. Am J Health Behav 2023 Feb 28;47(1):3-12. [doi: 10.5993/AJHB.47.1.1] [Medline: 36945101]
24. Wang D, Peng Y, Haddouk L, Vayatis N, Vidal PP. Assessing virtual reality presence through physiological measures: a comprehensive review. Front Virtual Real 2025;6:1530770. [doi: 10.3389/frvir.2025.1530770]
25. Saha S, Dobbins C, Gupta A, Dey A. Machine learning based classification of presence utilizing psychophysiological signals in immersive virtual environments. Sci Rep 2024;14(1):21667. [doi: 10.1038/s41598-024-72376-1]
26. Ronca V, Martinez-Levy AC, Vozzi A, et al. Wearable technologies for electrodermal and cardiac activity measurements: a comparison between Fitbit Sense, Empatica E4 and Shimmer GSR3+. Sensors (Basel) 2023 Jun 23;23(13):5847. [doi: 10.3390/s23135847] [Medline: 37447697]
27. Liberman L, Dubovi I. The effect of the modality principle to support learning with virtual reality: an eye‐tracking and electrodermal activity study. Computer Assist Learn 2023 Apr;39(2):547-557. [doi: 10.1111/jcal.12763]

XSL•FO
RenderX

28. Bowden J, Woolrych R, Kennedy C. Heritage, memory and well-being: exploring uses and perceptions of the historic environment amongst older adults in Nottinghamshire. Herit Soc 2025:1-34. [doi: 10.1080/2159032X.2024.2445449]
29. Bowden J, Woolrych R, Kennedy CJ. Situating place and wellbeing within heritage interactions for older adults. Heritage 2025;8(4):131. [doi: 10.3390/heritage8040131]
30. Tussyadiah IP, Wang D, Jung TH, tom Dieck MC. Virtual reality, presence, and attitude change: empirical evidence from tourism. Tour Manag 2018 Jun;66:140-154. [doi: 10.1016/j.tourman.2017.12.003]
31. Sylaiou S, Mania K, Karoulis A, White M. Exploring the relationship between presence and enjoyment in a virtual museum. Int J Hum Comput Stud 2010 May;68(5):243-253. [doi: 10.1016/j.ijhcs.2009.11.002]
32. Champion E. Critical Gaming: Interactive History and Virtual Heritage: Routledge; 2015. [doi: 10.4324/9781315574981]
33. O'Gara G, Murray L, Georgopoulou S, et al. SafeSpace: what is the feasibility and acceptability of a codesigned virtual reality intervention, incorporating compassionate mind training, to support people undergoing cancer treatment in a clinical setting? BMJ Open 2022 Feb 10;12(2):e047626. [doi: 10.1136/bmjopen-2020-047626] [Medline: 35144943]
34. Baldini A, Patron E, Gentili C, Scilingo EP, Greco A. Novel VR-based biofeedback systems: a comparison between heart rate variability- and electrodermal activity-driven approaches. IEEE Trans Affective Comput 2024;15(3):1580-1589. [doi: 10.1109/TAFFC.2024.3352424]
35. Lin C, Lee C, Lally D, Coughlin J. Impact of virtual reality (VR) experience on older adults' well-being. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population: Applications in Health, Assistance, and Entertainment ITAP 2018 Lecture Notes in Computer Science: Springer; 2018. [doi: 10.1007/978-3-319-92037-5_8]
36. Kerr JI, Weibel RP, Naegelin M, et al. The effectiveness and user experience of a biofeedback intervention program for stress management supported by virtual reality and mobile technology: a randomized controlled study. BMC Digit Health 2023;1(1):42. [doi: 10.1186/s44247-023-00042-z]
37. Wen Y, Shen X, Shen Y. Improving immersive experiences in virtual natural setting for public health and environmental design: a systematic review and meta-analysis of randomized controlled trials. PLoS One 2024;19(4):e0297986. [doi: 10.1371/journal.pone.0297986] [Medline: 38630668]
38. Dai Y, Ren Y, Li H, Wang M. Evaluating art exhibition spaces through space syntax and multimodal physiological data. Buildings 2025;15(11):1776. [doi: 10.3390/buildings15111776]
39. Boucsein W. Electrodermal Activity, 2nd edition: Springer; 2012. [doi: 10.1007/978-1-4614-1126-0]
40. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation 1996 Mar 1;93(5):1043-1065. [Medline: 8598068]
41. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54(6):1063-1070. [doi: 10.1037/0022-3514.54.6.1063]
42. Bjureberg J, Ljótsson B, Tull MT, et al. Development and validation of a brief version of the Difficulties in Emotion Regulation Scale: the DERS-16. J Psychopathol Behav Assess 2016 Jun;38(2):284-296. [doi: 10.1007/s10862-015-9514-x] [Medline: 27239096]
43. Campbell-Sills L, Stein MB. Psychometric analysis and refinement of the Connor-Davidson Resilience Scale (CD-RISC): validation of a 10-item measure of resilience. J Trauma Stress 2007 Dec;20(6):1019-1028. [doi: 10.1002/jts.20271] [Medline: 18157881]
44. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. Clin Gerontol 1986;5(1-2):165-173. [doi: 10.1300/J018v05n01_09]
45. Hays RD, DiMatteo MR. A short-form measure of loneliness. J Pers Assess 1987;51(1):69-81. [doi: 10.1207/s15327752jpa5101_6] [Medline: 3572711]
46. Mackersie CL, Calderon-Moultrie N. Autonomic nervous system reactivity during speech repetition tasks: heart rate variability and skin conductance. Ear Hear 2016;37(Suppl 1):118S-125S. [doi: 10.1097/AUD.0000000000000305] [Medline: 27355761]
47. Pop-Jordanova N, Pop-Jordanov J. Electrodermal activity and stress assessment. Pril (Makedon Akad Nauk Umet Odd Med Nauki) 2020 Sep 1;41(2):5-15. [doi: 10.2478/prilozi-2020-0028] [Medline: 33011695]
48. Kalantari S, Xu TB, Mostafavi A, et al. Using immersive virtual reality to enhance social interaction among older adults: a cross-site investigation. Innov Aging 2023;7(4):igad031. [doi: 10.1093/geroni/igad031] [Medline: 37213325]
49. Kalantari S, Bill Xu T, Mostafavi A, et al. Using a nature-based virtual reality environment for improving mood states and cognitive engagement in older adults: a mixed-method feasibility study. Innov Aging 2022;6(3):igac015. [doi: 10.1093/geroni/igac015] [Medline: 35592668]
50. Linares-Vargas BGP, Cieza-Mostacero SE. Interactive virtual reality environments and emotions: a systematic review. Virtual Real 2024;29(1):3. [doi: 10.1007/s10055-024-01049-1]
51. Syrigou K, Stoforou M, Kourtesis P. Time perception in virtual reality: effects of emotional valence and stress level. PsyArXiv. Preprint posted online on May 12, 2025. [doi: 10.31234/osf.io/63rek_v2]
52. Alexanian S, Foxman M, Pimentel D. Virtual reality and music's impact on psychological well-being. Front Rehabil Sci 2022;3:864990. [doi: 10.3389/fresc.2022.864990] [Medline: 36188907]

XSL•FO
RenderX

53. Tan J, Kannis-Dymand L, Jones C. Examining the potential of VR program Tilt Brush in reducing anxiety. Virtual Real 2022 Nov 10;27:1-13. [doi: 10.1007/s10055-022-00711-w] [Medline: 36405877]
54. Li Pira G, Aquilini B, Davoli A, Grandi S, Ruini C. The use of virtual reality interventions to promote positive mental health: systematic literature review. JMIR Ment Health 2023 Jul 6;10:e44998. [doi: 10.2196/44998] [Medline: 37410520]

Abbreviations

CD-RISC-10: 10-Item Connor-Davidson Resilience Scale
DERS-16: 16-Item Difficulties in Emotion Regulation Scale
EDA: electrodermal activity
GDS-15: 15-Item Geriatric Depression Scale
HRV: heart rate variability
NA: negative affect
NN: normal-to-normal
PA: positive affect
PANAS: Positive and Negative Affect Schedule
RM-ANOVA: repeated-measures analysis of variance
RMSSD: root-mean-square of successive difference
RQ: research question
SCL: skin conductance level
SCR: skin conductance response
SDNN: standard deviation of NN
VR: virtual reality

Edited by H Köttl; submitted 19.Jan.2026; peer-reviewed by K Yuenyongchaiwat, V Ronca; revised version received 15.Jun.2026; accepted 16.Jun.2026; published 14.Jul.2026.

Please cite as:
Lai Y, Wang P
Emotional Regulation and Physiological Responses to a Cultural Heritage Virtual Reality Experience in Older Adults: Mixed Methods Study
JMIR Aging 2026;9:e91772
URL: https://aging.jmir.org/2026/1/e91772
doi:10.2196/91772

© Yafeng Lai, Pohsun Wang. Originally published in JMIR Aging (https://aging.jmir.org), 14.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Demoralization and Mental Disorders in Community-Dwelling Older Adults: Cluster Analysis

Alice Onofrio[1*], MD; Maria Ferrara[1,2,3*], MD, PhD; Ilaria Domenicano[1], PhD; Martino Belvederi Murri[1,2], MD; Tommaso Toffanin[1,2], MD, PhD; Maria Giulia Nanni[1,2], MD; Alessandra Canuto[4], MD; Kerstin Weber[5], PhD; Jana Volkert[6], PhD; Holger Schulz[7], MD; Martin Härter[7], MD, PhD; Sylke Andreas[8], PhD; Luigi Grassi[1,2], MD

[1]Department of Neurosciences and Rehabilitation, University of Ferrara, via fossato di Mortara 64/a, Ferrara, Italy

[2]Integrated Department of Mental Health and Pathological Addiction, Local Health Trust (AUSL) Ferrara, Ferrara, Italy

[3]Department of Psychiatry, Yale School of Medicine, New Haven, CT, United States

[4]Valais Hospital - Hospital Centre of French-speaking Valais (CHVR), Sion, Switzerland

[5]Department of Psychiatry, University of Geneva, Geneva, Switzerland

[6]Clinic for Psychosomatic Medicine and Psychotherapy, University of Ulm, Ulm, Germany

[7]Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[8]Institute of Psychology, Alpen-Adria University Klagenfurt, Klagenfurt, Austria

[*]these authors contributed equally

Corresponding Author:
Alice Onofrio, MD
Department of Neurosciences and Rehabilitation, University of Ferrara, via fossato di Mortara 64/a, Ferrara, Italy

Abstract

Background: Demoralization negatively affects the quality of life, especially in people with mental disorders.

Objective: This study aims to assess demoralization in a cohort of community-dwelling older adults, to identify groups of individuals who share similar characteristics in terms of mental health conditions and the severity of demoralization domains, and to investigate the different characteristics across these groups.

Methods: Participants were enrolled in 3 centers as part of the cross-sectional MentDis_ICF65+ study and were assessed using the Demoralization Scale and the Composite International Diagnostic Interview at baseline. A cluster analysis was conducted using the partitioning around medoids algorithm with Gower's distance to classify individuals based on demoralization severity and the diagnosis of mental health disorders, allocating individuals to a distinct number of patterns by examining their similarities across these variables.

Results: Among the final sample of 1369 participants, 3 clusters were identified: the "low psychopathology" cluster, with very few mental health diagnoses; the "affective" cluster, characterized by affective disorders (100% of participants); and the "anxiety" cluster, characterized by anxiety disorders (100% of participants). Men were prevalent in the "low psychopathology" cluster (n=635, 59.1%; P<.001), while women were more prevalent in the "affective" cluster (n=133, 72.3%; P<.001) and the "anxiety" cluster (n=79, 71.8%; P<.001). The "affective" cluster had the highest mean scores for disheartenment (17.1, SD 6.7; P<.001), dysphoria (12.1, SD 3.5; P<.001), and loss of meaning (11.8, SD 4.9; P<.001) but the lowest for sense of failure (17.4, SD 3.2; P<.001).

Conclusions: The identification of different clusters of demoralization and mental health disorders in community-dwelling older adults, as well as sex-based differences, highlights the need for tailored mental health interventions in this vulnerable population.

(JMIR Aging 2026;9:e84657) doi:10.2196/84657

KEYWORDS

demoralization; depression; anxiety; cluster; older

Introduction

Demoralization was conceptualized more than 40 years ago as a state of existential distress denoting a persistent failure to cope with stress and occurs in patients with severe physical or mental illnesses, specifically those that threaten life or the integrity of being [1]. Subsequently, it has been described as a syndrome characterized by loss of purpose and meaning, a state of persistent feeling of subjective incompetence and existential distress, hopelessness, and helplessness, and an inability to cope

XSL•FO
RenderX

and solve problems [2-4]. As a construct, demoralization only partly overlaps with depression, with demoralized individuals not experiencing the lack of motivation typical of major depression, but rather an uncertainty in the appropriate course of action that results in an inability to cope with external sources of distress [5-7].

Demoralization has been extensively investigated in patients with medical comorbidities, including, but not limited to, those with cardiac, dermatological, and oncological conditions [8-10], as well as mental disorders. Additionally, demoralization has been shown to be associated with suicidal ideation, poorer quality of life, and worse outcomes in both mental and somatic disorders [7,11,12]. It has also been established that the prevalence of mental health diagnoses increases with demoralization [13,14] and that demoralization might play a role in the development or exacerbation of mental health conditions, such as anxiety disorders or depression [15].

In the older population, demoralization syndrome has not been examined in detail, apart from a few authors, such as Ramm et al [16], who found a higher prevalence of demoralization in people aged ≥65 years compared to younger people, as well as an increase in older adulthood (65-74 y, 75-84 y, and 85 yrs and older, 50% prevalence).

Since previous studies among the older individuals with mental disorders, especially depression, showed that they had a worse quality of life and higher levels of disability compared to their peers without a mental health diagnosis [17], it is imperative to understand the impact of demoralization in the aging population [18].

By using cluster analysis, or clustering, a machine learning technique that can automatically identify natural groupings in data by decomposing heterogeneous data [19], the impact of demoralization and mental health disorders in the older population could be examined in a proper way. While predictive modeling is used to predict the already-existing class, clustering algorithms play a key role when the classes are unknown, and the instances have to be classified into [20]. Cluster analysis offers a better understanding of the heterogeneity that exists in patient characteristics in clinical populations [21]. This method has already been applied to identify subtypes of medical patients suffering from demoralization. For instance, Rafanelli et al [22], by studying 1560 medically ill patients, found that 373 had demoralization, and 4 clusters were identified: in 3 out of 4 clusters, demoralization was comorbid with a mental health diagnosis (either major depression, anxiety, and somatoform disorders); in one cluster, it was not associated with any mental health condition (110/373, 29.5% of the sample). Cluster analysis has also been successfully applied to a large mental health registry to describe subgroups of patients in order to program resource allocations [23].

The aims of this study were to assess demoralization in a cohort of community-dwelling older adults and to identify whether groups of individuals cluster together by sharing similar characteristics in terms of demoralization scores as well as domains and mental health disorders. A secondary aim was to investigate clinical and sociodemographic differences across those groups.

Methods

Design

We examined data from the European project Mental Disorders in the Elderly (MentDis_ICF65+), a multicenter study aimed at investigating mental disorders in the community-dwelling older population, including incidence, prevalence, and correlations with the intensity of symptoms, levels of engagement in activities, participation, and use of services [24,25].

The study was conducted in 6 countries (Germany, Great Britain, Israel, Italy, Spain, and Switzerland) and included 3142 participants aged between 65 and 84 years of age. The exclusion criteria were severe cognitive impairment and insufficient proficiency in the corresponding language. It consisted of a community survey administered at baseline, with a follow-up at 12 months. The survey aimed to collect 12-month and 1-month prevalence and comorbidity of mental disorders, as well as information on the prior course of the disorders, quality of life, health care use, help-seeking behavior, impairments, and participation in community life.

Only 3 centers (ie, Ferrara, Geneva, and Hamburg) out of the 6 centers examined the construct of demoralization in addition to the standard protocol and were, therefore, included in the present analysis.

Measures

All participants in the 6 centers underwent individual clinical interviews, along with the completion of psychometric instruments. Additionally, aspects such as quality of life, health care use, help-seeking behavior, impairments, and participation in community life were assessed.

Mental health was assessed using the Composite International Diagnostic Interview for the Elderly (CIDI65+) [26], a version of the Composite International Diagnostic Interview (CIDI) specifically adapted to the needs and abilities of the older individuals. The CIDI65+ test-retest reliability has been examined and was considered satisfactory [26]. This interview assessed the 1-month and 12-month prevalence of somatic morbidity, somatoform disorders, anxiety disorders, obsessive-compulsive disorders, affective disorders, psychotic symptoms, substance abuse, adjustment and stress-related disorders, and cognitive impairment. For the purpose of this investigation, only the 1-month prevalence was considered.

The assessment of physical health included a checklist of primary medical conditions (such as cardiovascular disease, cancer, and musculoskeletal disease) self-reported by the participants, as well as a survey on the use of medications and access to health care facilities. Demographic and socioeconomic data were also gathered as part of the standard assessment.

The Demoralization Scale (DS), developed by Kissane et al [27], is a 24-item self-report tool measuring demoralization syndrome through a 5-point Likert scale (“never”=0; “all the time”=4) in 4 domains: disheartenment, loss of meaning or purpose, dysphoria, and sense of failure. The disheartenment domain includes feelings of discouragement and isolation, the

XSL•FO
RenderX

loss of meaning domain pertains to the loss of role, purpose, and sense of worth in life; the dysphoria domain encompasses emotions of distress and regret; and the sense of failure domain assesses the levels of accomplishment experienced by the individual [27]. The overall demoralization score (DS-Total) is calculated by summing item scores, ranging from 0 to 96. A total score equal to or exceeding 30 is indicative of a demoralized status, according to the criteria established by Kissane et al [27]. The DS has previously demonstrated robust levels of validity and reliability in several populations, such as patients with cancer [28], medically ill patients [5], and patients with nonpsychotic affective disorders [29].

Statistical Analysis

Descriptive statistics were computed for all study variables. Individuals were classified as demoralized in a categorical variable, according to Kissane [27].

In this study, we applied a cluster analysis method to allocate individuals to a distinct number of patterns by investigating individuals' similarities regarding variables that described several aspects of demoralization and the presence or absence of mental health disorders. Cluster analysis is an exploratory method for identifying groups of objects that share common characteristics. Generally, the groups are built such that objects belonging to the same groups share a similar response pattern on the observed variables, while such patterns are as different as possible between objects in different groups [30].

Eight variables were used as classifiers, namely (1) four continuous variables aimed at describing the severity of demoralization, that is, four domains of the DS as identified by Grassi et al [8]: disheartenment, sense of failure, dysphoria, and loss of meaning and purpose; (2) four binary variables aimed at identifying the presence or absence of mental health disorders: anxiety disorders (phobic disorder, panic attack, panic disorder, generalized anxiety disorder (GAD), obsessive-compulsive disorder), affective disorders, and nonaffective psychosis (major depressive disorder, dysthymic disorder with or without hierarchy, double depression, bipolar 1, bipolar 2, psychosis), alcohol dependence or abuse disorder, and somatoform pain disorder. Individuals with missing information regarding mental health diagnoses (5/1479) or demoralization (105/1479) were excluded from the cluster analysis.

Raw scores on demoralization domains were standardized to z-scale scores based on their mean and SD in the full sample pooled across centers to prevent variables with larger numerical ranges from dominating the clustering procedure [31]. Since the data contained both continuous and categorical variables, Gower's similarity measure was used to generate the matrix of similarities [32]. In computing the similarity measure, the same weight was assigned to each variable, and binary variables were treated as nominal variables. Based on this matrix, individuals were clustered by the partitioning around medoids (PAM) algorithm [33]. For a given number of groups, PAM allocates individuals so that the dissimilarity to the medoid is minimized within groups. We used several criteria to select the optimal number of clusters over 2‐7 cluster solutions. First, we constructed a total within sum of squares (WSS) plot and used the so-called elbow method to identify the optimal number of clusters so that adding another cluster would not improve the total WSS. Second, we computed the silhouette and identified the number of clusters for which the average silhouette width was maximized. Similarly, we employed the GAP statistics and found the solution cluster that maximized the statistics. Finally, we computed the average distance to cluster centroid to determine which solutions had the lowest value [33]. After identifying the optimal number of clusters, cluster-wise stability was assessed using bootstrap resampling with Jaccard similarity. The Jaccard index, computed as the mean over the Jaccard similarities between the original and the bootstrapped clusters, was used to evaluate the stability of a cluster. An index approaching 1 indicates a stable cluster, while lower values suggest instability [34]. The number of resampling was set to 1000.

Finally, we used t-distributed stochastic neighbor embedding (t-SNE) to visualize them. t-SNE is a nonlinear dimensionality reduction technique that allows the visualization of high-dimensional datasets in a lower-dimensional space. Data can, therefore, be represented as a scatter plot, where points that were close in the original space remain close, and points that were far away in the original space remain out to be far away [35]. The t-SNE representation was used only for visualization purposes, and not as evidence of separability.

After identifying the groups, we assessed the characteristics of the clusters by comparing the variables used for clustering (eg, the 4 subscales of the DS and the presence or absence of mental health disorders).

Furthermore, we assessed the characteristics of the clusters by comparing the external variables (eg, those not used for clustering, including sociodemographic variables such as sex, age, marital status, household, number of school years, working status, perceived financial situation, and physical illness). Differences in the aforementioned sample characteristics were compared according to cluster membership using independent sample ANOVA and χ^2 tests for continuous and categorical variables, respectively. Due to the exploratory nature of the study, analyses were not controlled for multiple testing.

The analyses were conducted with the R software (version 4.3.1; R Foundation for Statistical Computing) [36]. The packages *cluster* [37] and *fpc* [38] were used to perform and evaluate the cluster analysis. For all tests performed, the significance level was set at 0.05.

Ethical Considerations

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The study was approved by research ethics committees in all 6 centers (Germany: Hamburg Ethics Committee of the Medical Association No. 2895, Italy: Ferrara No. 0096637 5/11/2009, Israel: Jerusalem No. 0376‐09 -HMO, Spain: Madrid No. 22032010, Switzerland: University Hospitals of Geneva Ethics Committee, Protocol No. 09‐121 and UK: National Research Ethics Service No. 10/H0715/21).

XSL•FO
RenderX

Results

A total of 1479 participants were included in this study. However, 110 were excluded because the information regarding mental health diagnosis (n=5) or demoralization (n=105) was not available, leaving a final sample of 1369 individuals (Ferrara, n=466 participants; Geneva, n=509 participants; and Hamburg, n=394 participants; Figure S1 and Table S1 in Multimedia Appendix 1). Of the 110 missing outcomes, 50.9% (n=56) were female and 49.1% (n=54) were male; 26.4% (n=29) were aged 65 - 69 years, 29.1% (n=32) were aged 70 - 74 years, 29.1% (n=32) were aged 75 - 79 years, and 15.5% (n=17) were aged 80 years or older. Sociodemographic data are summarized in Table 1, while Figure 1 shows a visual representation of the 3 clusters.

XSL•FO
RenderX

Table . Sociodemographic characteristics of the 3 clusters.

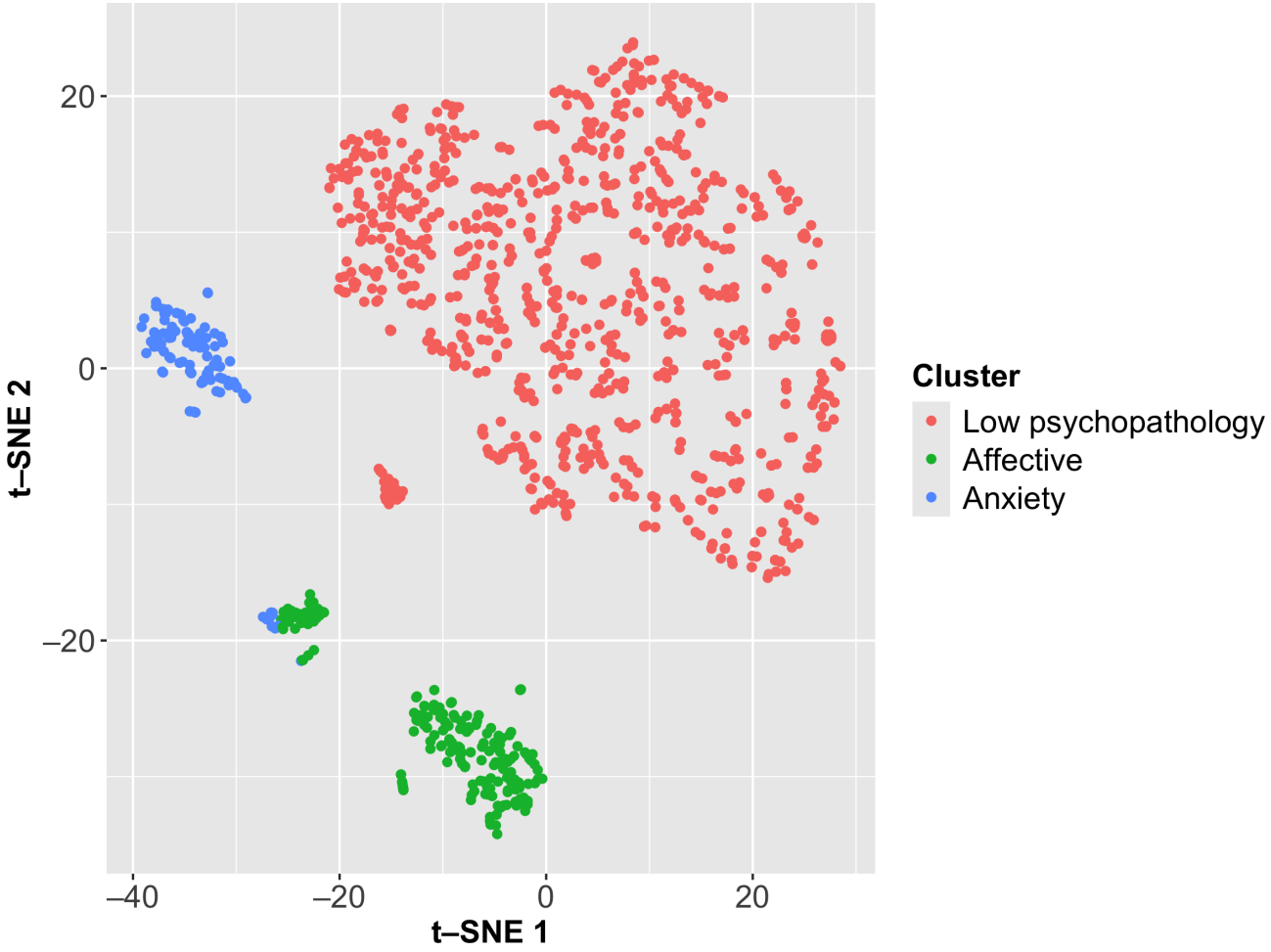
Characteristics	Total (N=1369)	"Low psychopathology" cluster (n=1075)	"Affective" cluster (n=184)	"Anxiety" cluster (n=110)	χ^2 (*df*)	*F* value (*df*)	Cohen[a] effect size	*P* value
Center, n (%)					34.69 (4)		0.16	<.001
Ferrara	466 (34.0)	383 (35.6)	46 (25.0)	37 (33.6)				
Geneva	509 (37.2)	368 (34.2)	104 (56.5)	37 (33.6)				
Hamburg	394 (28.8)	324 (30.2)	34 (18.5)	36 (32.8)				
Sex, male, n (%)	717 (52.4)	635 (59.1)	51 (27.7)	31 (28.2)	89.98 (2)		0.26	<.001
Age, mean (SD)	73.1 (5.3)	73.2 (5.3)	72.4 (5.3)	72.8 (5.0)		1.75 (2, 1366)	0.05	.17
Age group (y), n (%)					8.47 (6)		0.08	.20
65 - 69	435 (31.8)	331 (30.8)	72 (39.1)	32 (29.1)				
70 - 74	390 (28.5)	310 (28.8)	43 (23.4)	37 (33.6)				
75 - 79	343 (25.0)	268 (24.9)	47 (25.5)	28 (25.5)				
>80	201 (14.7)	166 (15.5)	22 (12.0)	13 (11.8)				
Marital status, n (%)					48.36 (4)		0.19	<.001
Married	865 (63.2)	725 (67.4)	78 (42.4)	62 (56.4)				
Separated/divorced/widowed	439 (32.1)	299 (27.8)	96 (52.2)	44 (40.0)				
Never been married/other	64 (4.7)	50 (4.7)	10 (5.4)	4 (3.6)				
Missing	1 (0.1)	1 (0.1)	0 (0)	0 (0)				
Household, n (%)					36.64 (2)		0.16	<.001
Lives alone	405 (29.6)	278 (25.9)	86 (46.7)	41 (37.3)				
Other	961 (70.2)	796 (74.0)	98 (53.3)	67 (60.9)				
Missing	3 (0.2)	1 (0.1)	0 (0.0)	2 (1.8)				
Number of school years, mean (SD)	11.6 (4.6)	11.7 (4.7)	11.9 (4.7)	10.6 (3.8)		3.36 (2, 1363)	0.07	.04
Number of school years group, n (%)					12.46 (4)		0.10	.01
0 - 8	356 (26.0)	278 (25.9)	45 (24.5)	33 (30.0)				
9 - 12	469 (34.3)	350 (32.7)	70 (38.0)	49 (44.5)				
13 and more	541 (39.5)	444 (41.4)	69 (37.5)	28 (25.5)				
Working status, n (%)					12.08 (8)		0.09	.15
Working or employed	72 (5.3)	58 (5.4)	8 (4.4)	6 (5.5)				
Homemaker or housewife	34 (2.5)	25 (2.3)	8 (4.4)	1 (0.9)				
Unemployed	1 (0.1)	0 (0)	1 (0.5)	0 (0)				
Retired	1242 (90.7)	975 (90.7)	166 (90.2)	101 (91.8)				

XSL•FO
RenderX

Characteristics	Total (N=1369)	"Low psychopathology" cluster (n=1075)	"Affective" cluster (n=184)	"Anxiety" cluster (n=110)	χ^2 (*df*)	*F* value (*df*)	Cohen[a] effect size	*P* value
Other	5 (0.4)	4 (0.4)	0 (0.0)	1 (0.9)				
Missing	15 (1.1)	13 (1.2)	1 (0.5)	1 (0.9)				
Financial situation perceived, n (%)					21.40 (8)		0.12	.006
Very good	180 (13.1)	144 (13.4)	20 (10.9)	16 (14.5)				
Good	661 (48.3)	507 (47.2)	105 (57.0)	49 (44.6)				
Just enough	450 (32.9)	371 (34.5)	41 (22.3)	38 (34.5)				
Poor	63 (4.6)	42 (3.9)	16 (8.7)	5 (4.6)				
Very poor	11 (0.8)	7 (0.6)	2 (1.1)	2 (1.8)				
Missing	4 (0.3)	4 (0.4)	0 (0)	0 (0)				

[a]Effect sizes are reported as Cohen d (continuous variables) or Cohen w (categorical variables), as appropriate

Figure 1. Clusters, as detected by the partitioning around medoids (PAM) algorithm, represented with t-distributed stochastic neighborhood embedding (t-SNE). T-SNE parameters were set as follows: perplexity=30, learning rate=200, iteration=1000, seed=123.



A total of 317/1369 (23.15%) patients were classified as demoralized. Amongst demoralized individuals, 147 (46%) had a comorbid psychiatric diagnosis. Amongst those who did not have any mental disorder (1041/1369), 16% (n=170) were classified as demoralized (Table 2).

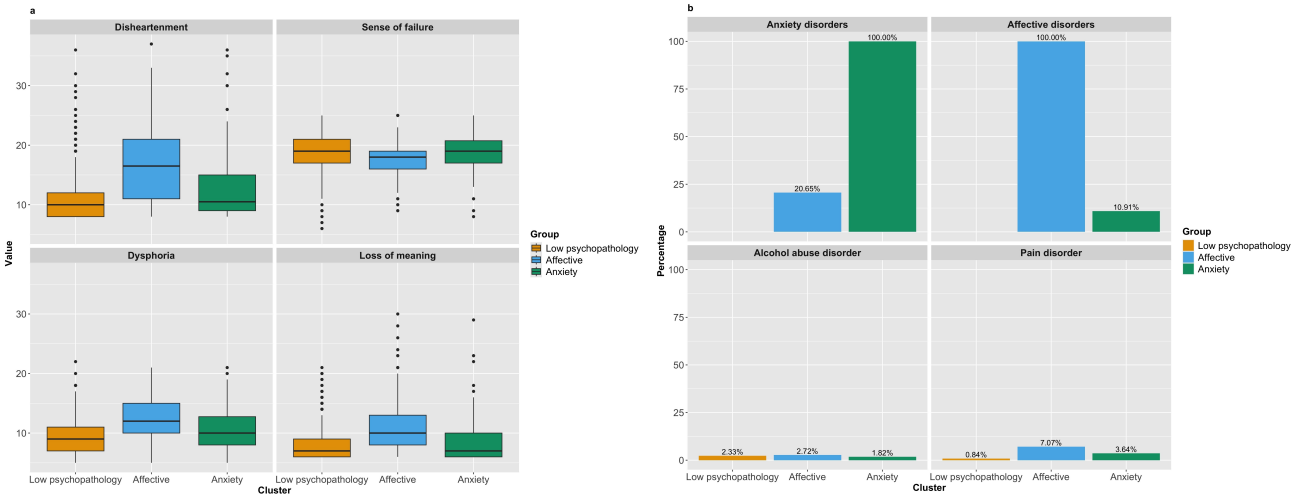
XSL•FO
RenderX

Table . Distribution of mental disorders and demoralization in the study sample for the cluster analysis (N=1369).

Characteristics	Demoralization score <30, not demoralized (n=1052), n (%)	Demoralization score ≥30, demoralized (n=317), n (%)	Total, n
Mental health disorders			
Anxiety disorders	83 (7.9)	65 (20.5)	148
Affective disorders	83 (7.8)	113 (35.6)	196
Alcohol abuse	28 (2.7)	4 (1.3)	32
Somatoform pain disorders	9 (0.9)	17 (5.4)	26
No mental health disorders	871 (82.8)	170 (53.6)	1041

The number of clusters that showed the best performance was 3, as it had an adequate average silhouette and elbow shape (Table S2, Table S3, and Figure S2 in Multimedia Appendix 1). The first cluster was characterized by "low psychopathology" (0% anxiety disorders, 0% affective disorders, 2.3% alcohol abuse disorder, and 0.8% somatoform pain disorder); the second was characterized by the presence of affective disorders (100% of participants); and the third by anxiety disorders (100% of participants; Figure 2B). Some individuals had both anxiety and affective disorders: 38 (20.7%) participants in the "affective" cluster also had an anxiety disorder, and 12 (10.9%) participants in the "anxiety" cluster also had an affective disorder (Table 3). All 3 clusters showed a robust structure, as they rarely separated or merged with the others upon resampling: specifically, the Jaccard bootstrap mean was 0.82 for the "low psychopathology" cluster, 0.88 for the "affective" cluster, and 0.83 for the "anxiety" cluster.

Figure 2. Characteristics of the 3 clusters across the variables used as classifiers (a): demoralization domains; (b): mental health disorders).



XSL•FO
RenderX

Table . Characteristics of the 3 clusters across the variables used as classifiers (demoralization domains and mental health disorders).

	"Low psychopathology" cluster (n=1075)	"Affective" cluster (n=184)	"Anxiety" cluster (n=110)	χ^2 (*df*)	*F* value (*df*)	Effect size Cohen[a] (95% CI)	*P* value
Demoralization, mean (SD)							
Total score	22.8 (7.4)	34.1 (11.8)	26.5 (10.9)		141 (2, 1366)	0.45 (0.0-0.50)	<.001
Disheartenment	11.1 (3.7)	17.1 (6.7)	13.0 (6.0)		146 (2, 1366)	0.46 (0.0-0.51)	<.001
Sense of failure	19.0 (3.0)	17.4 (3.2)	18.4 (3.2)		21.47 (2, 1366)	0.18 (0.0-0.22)	<.001
Dysphoria	8.9 (2.8)	12.1 (3.5)	10.5 (3.8)		97.71 (2, 1366)	0.38 (0.0-0.42)	<.001
Loss of meaning	7.9 (2.6)	11.8 (4.9)	8.6 (4.1)		95.14 (2, 1366)	0.37 (0.00.42)	<.001
DS ≥30, n (%)	175 (16.3)	112 (60.9)	30 (27.3)	176.7 (2)		0.36 (0.30-0.42)	<.001
Mental health disorders, yes, n (%)							
Anxiety disorders	0 (0.0)	38 (20.7)	110 (100)	1056.3 (2)		0.88 (0.85-0.90)	<.001
Affective disorders	0 (0.0)	184 (100)	12 (10.9)	1281.8 (2)		0.97 (0.95-0.98)	<.001
Alcohol abuse disorder	25 (2.3)	5 (2.7)	2 (1.8)	0.25 (2)		0.01 (0.00-0.07)	.88
Somatoform pain disorder	9 (0.8)	13 (7.1)	4 (3.6)	34.6 (2)		0.16 (0.09-0.23)	<.001
Anxiety disorders + affective disorders	0 (0.0)	38 (20.7)	12 (10.9)	208.3 (2)		0.39 (0.34-0.45)	<.001
Alcohol abuse disorder	0 (0.0)	0 (0.0)	0 (0.0)	—[b]	—	—	—
Somatoform pain disorder	—	4 (10.5)	1 (8.3)	—	—	—	—

[a]Effect sizes are reported as Cohen d (continuous variables) or Cohen w (categorical variables), as appropriate.

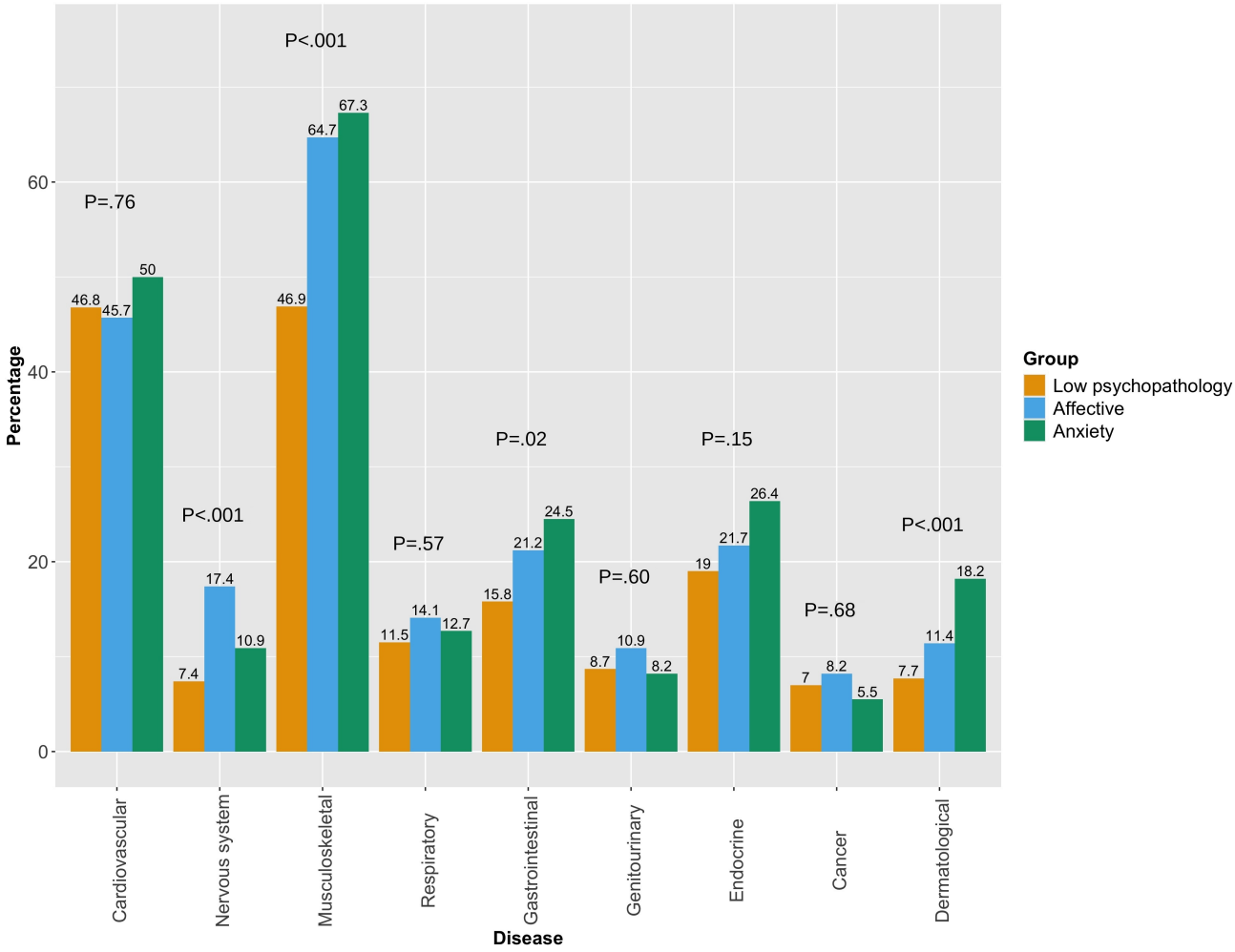
[b]Not available.

Men were more prevalent in the "low psychopathology" cluster (635/1075, 59.1%; *P*<.001), while women were largely prevalent in the "affective" cluster (133/184, 72.3%; *P*<.001) and "anxiety" cluster (79/110, 71.8%; *P*<.001). There was no significant difference in age between clusters (*P*=.20). A difference in marital status was observed between the 3 clusters: in the "low psychopathology" cluster and the "anxiety" cluster, married people were more prevalent (725/1075, 67.4% and 62/110, 56.4%, respectively; *P*<.001), while in the "affective" cluster, separated or widowed or divorced participants were more prevalent (96/184, 52.2%; *P*<.001; Table 1).

The proportion of people living alone was the highest in the "affective" cluster (n=86, 46.7%; *P*<.001), followed by the "anxiety" cluster (n=41, 37.3%; *P*<.001) and the "low psychopathology" cluster (n=278, 25.9%; *P*<.001). As shown in Figure 3 and Table S4 in Multimedia Appendix 1, the "affective" cluster had the highest prevalence of nervous system (*P*<.001) physical comorbidities, while the "anxiety" cluster had the highest prevalence of musculoskeletal (*P*<.001), gastrointestinal (*P*=.02), endocrine (*P*=.15), and dermatological (*P*<.001) comorbidities.

XSL•FO
RenderX

Figure 3. Physical comorbidities across the 3 clusters.



Regarding the different domains of demoralization, the "low psychopathology" cluster had the lowest scores for disheartenment (mean 11.1, SD 3.7; *P*<.001), dysphoria (mean 8.9, SD 2.8; *P*<.001) and loss of meaning (mean 7.9, SD 2.6; *P*<.001) subdomains, but the highest score for the sense of failure domain (mean 19.0, SD 3; *P*<.001). The "affective" cluster had the highest mean score for disheartenment (mean 17.1, SD 6.7; *P*<.001), dysphoria (mean 12.1, SD 3.5; *P*<.001) and loss of meaning (mean 11.8, SD 4.9; *P*<.001), but the lowest mean score for sense of failure (mean 17.4, SD 3.2; *P*<.001; Table 3).

Discussion

This study aimed to identify possible clusters based on mental health disorders and demoralization in a community-dwelling older population across 3 European centers. Using the criteria of the methods described above, we identified 3 distinct patterns. The first group was characterized by low psychopathology, the second by the presence of affective disorders, and the third by the presence of anxiety disorders.

About one fourth (23.2%) of the community-dwelling older individuals in our sample were found to be demoralized. This result is very close to the percentage of demoralized participants previously found in medically ill patients of any age [39,40] and in a more recent study of the older population [16]. This result underlines the need to take demoralization into careful consideration, since it is well known that with age, the burden of medical illnesses increases and that medical illnesses are often associated with demoralization [8,10,11,13].

Some participants had both anxiety and affective disorders. The cluster analysis assigned these comorbid patients either to the "anxiety" cluster or to the "affective" cluster based on the other classifiers included in the analysis, that is, the 4 domains of the DS identified by Grassi et al [7,28]. This finding suggests that the assessment of demoralization can help clarify the characteristics of comorbid mental disorders. This could allow the implementation of tailored interventions in otherwise mixed diagnostic categories.

The "affective" cluster showed higher scores in most of the demoralization domains, for instance, higher disheartenment, dysphoria, and loss of meaning. Affective disorders, especially depression, and demoralization syndrome have some overlapping symptoms, such as suicidal ideation, but are, in fact, separate concepts. Demoralization has been defined as a combination of distress and subjective incompetence, without the lack of motivation that characterizes Major Depressive Disorder [3]. Demoralization and Major Depression, however, are often correlated [5,7]. There might be a two-way relationship: depression, as with any other medical condition,

XSL•FO
RenderX

could result in demoralization, but at the same time demoralization could precede depression [22,41].

Women were more prevalent in the "affective" and "anxiety" clusters, which confirms gender differences in the prevalence of mental disorders. Mental disorders are more prevalent in females than in males, especially anxiety and mood disorders [42]. The underlying mechanisms of this difference remain unclear and may include society-mediated factors, genetics and hormonal fluctuations, the latter especially in the perimenopausal and postmenopausal period [43].

The "low psychopathology" cluster had the lowest levels of disheartenment, dysphoria and loss of meaning, but the highest sense of failure. The presence of a higher sense of failure in this group, especially compared to the relatively low prevalence in the "affective" cluster, has never been reported before. This finding, which could appear contradictory at first, aligns with the concept of *subjective incompetence* [44,45]. In fact, the main source of distress for older adults without an identifiable psychiatric diagnosis may arise from the comparison between the ever-present need for autonomy and age-related functional limitations, which in turn could trigger a more pronounced reactive perception of personal incompetence. This suggests that demoralization and mental health conditions do not necessarily overlap, corroborating the phenomenological difference between demoralization syndrome and mental health disorders such as depression. Furthermore, the elevation in the "Sense of Failure" domain suggests that, in this cluster, demoralization may act as a dimension of existential distress rather than a marker of a mental disorder ([5]M. Belvederi Murri et al., 2020 [3]. As individuals age, their temporal perspective shifts, often leading them to reflect more on past achievements rather than future goals. In this context, the 'Sense of Failure' may capture the appraisal of one's life trajectory rather than acute emotional distress. When this retrospective recollection collides with changes in social roles or functional decline, it can manifest as perceived inadequacy or unfulfilled potential.

The "affective" cluster had a higher prevalence of people living alone, as well as those who were separated, widowed, or divorced. This finding suggests that loneliness plays a role in affective disorders, particularly among older individuals. Previous literature studied this association, but further research may be needed to assess the impact of loneliness on demoralization [46], as also recently indicated in a study of older patients on hemodialysis [47] as well as in older patients with stroke and disability [48].

Nervous system diseases were most prevalent in the "affective" cluster, while musculoskeletal, gastric, endocrinological, and dermatological diseases were most prevalent in the "anxiety" cluster. These findings support the notion that psychopathology correlates with specific physical comorbidities. The correlation between medical conditions and a higher prevalence of mental health diagnoses is well documented in the literature. For example, patients with neurological conditions such as migraines or epilepsy have a higher prevalence of affective disorders compared to the general population [49,50], while anxiety disorders are often associated with gastrointestinal conditions (ie, peptic ulcers or IBS), endocrinological conditions (ie, Graves' disease), and dermatological diseases (ie, psoriasis) [51,52].

To our knowledge, this is the first study to apply a cluster analysis method to examine the relationship between demoralization and mental health disorders in a sample of community-dwelling older participants. Previous studies have mostly focused on demoralization in medical settings, whereas this study assesses on assessing demoralization among older people in the general population. In addition, the cluster analysis method offers a new strategy for understanding the heterogeneity of participants' characteristics within the sample. The identification of different clusters in the presentation of mental health conditions provides the opportunity to design more tailored psychotherapeutic, psychoeducational, and psychosocial interventions for the examined population. For the "affective" cluster, psychotherapeutic interventions tailored to older adults may reduce the psychological distress associated with demoralization. For instance, Managing Cancer and Living Meaningfully (CALM), a psychotherapeutic intervention designed to address end-of-life distress in cancer patients [53], could be adapted to accommodate the care needs of older adults, who share a similar burden of disability and loss of meaning. For the "anxiety" cluster, age-adapted psychotherapy and psychoeducation focusing on emotional regulation and coping strategies may be beneficial in reducing anxiety symptoms.

Some limitations need to be acknowledged. First, due to the cross-sectional design of this investigation, causal inferences are not possible. Second, the geographic limitations of the catchment area, as well as the exclusion of individuals with cognitive impairment and others living in residential facilities, those who are homeless, or those with a language barrier might limit the generalizability of the findings. Third, as in most studies on demoralization, we also used a psychometric tool, while more specific data might have emerged through a clinometric and interview-based approach [54,55], which represent the gold-standard for assessing demoralization. In relation to this, a fourth limitation is that we considered a clinical condition indicating demoralization syndrome by using a standard cut-off score of ≥30, which is, in fact, what we found in a global population of medically ill patients. Studies comparing the demoralization interview vs the DS in older individuals are necessary to define a cut-off that optimizes sensitivity and specificity for this construct. Finally, while the DS has been widely adopted in research and clinical contexts, however, its validation is not yet available in the French language.

As reported, 7% of participants were excluded from the analysis due to missing information on diagnosis or demoralization subscales. Since a relatively small fraction of the data was missing, and the dataset presented an adequate sample size, a complete-case analysis was used to handle missing data [56]. However, a sensitivity analysis to compare the current results (obtained through a complete-case analysis) to an imputed dataset would strengthen the study and ensure the robustness of the findings.

A further limitation that must be acknowledged is that psychiatric diagnosis variables were included as binary variables

XSL•FO
RenderX

within the clustering features; thus, the clustering algorithm might have been strongly incentivized to separate groups by such variables, potentially dominating the structure and reducing the added value of demoralization dimensions. To address this specific aspect, a sensitivity analysis clustering only on demoralization domains could be performed, subsequently comparing diagnoses across resulting clusters, or focusing on variable contribution by observing changes in silhouette scores when iteratively removing variables. Moreover, binary variables were automatically treated as symmetrical, even if potentially unbalanced. Finally, given the exploratory nature of this study, we did not systematically correct for multiple comparisons, which might increase the risk of Type I errors. To mitigate this issue, effect sizes and confidence intervals were reported.

This study identified three different clusters based on the presentation of demoralization and mental health disorders in community-dwelling older individuals, underlining sex-based differences in the presentation of demoralization and specific dimensions of demoralization, mostly associated with certain mental disorders. Further research is needed to guide tailored mental health interventions to address the vulnerabilities of these specific categories in the older population and the role of demoralization in influencing individuals' quality of life [57].

Acknowledgments

The authors would like to thank other collaborators of the Mentdis ICF65+ study: Chiara Da Ronch, Holger Schulz, Maria Dehoust, Susanne Sehner, Anna Suling, Karl Wegscheider, Mike J. Crawford, Yael Hershkovitz, Alan Quirk, Ora Rotenstein, Ana Belén Santos-Olmo, Arieh Shalev, and Jens Strehle. All participants to the study are gratefully acknowledged. We also express our gratitude to all interviewers, dedicated collaborators, and local institutions that made this study possible.

Generative AI tools were used solely for language editing purposes, including correction of grammar, syntax, and wording to improve readability and clarity of the manuscript. No generative AI tools were used for the generation, interpretation, analysis, or modification of scientific content, data, results, references, or conclusions. The authors reviewed and approved all revisions and remain fully responsible for the accuracy, originality, and integrity of the manuscript, including all references and citations.

Funding

This study was funded by a grant from the European Commission (grant 223,105) within the 7th Framework Research Programme of the European Commission. The funding body did not play any role in the design of the study, collection, analysis and interpretation of data nor in writing of this manuscript.

Authors' Contributions

Study concept and design: AO, MF, ID
Acquisition, analysis, or interpretation of data: All authors
Drafting of the manuscript: AO, MF, ID
Critical revision of the manuscript for important intellectual content: All authors
Statistical analysis: ID
Study supervision: MH, LG

Conflicts of Interest

None declared.

Multimedia Appendix 1
Participant flow diagram; exact missing counts; criteria used to determine the optimal number of clusters; performance of the k-prototypes algorithm; hierarchical clustering with Gower+Ward; physical comorbidities across the 3 clusters; and comparison of baseline characteristics across the 3 centers.
[DOCX File, 235 KB - aging_v9i1e84657_app1.docx]

References

1. de Figueiredo JM, Frank JD. Subjective incompetence, the clinical hallmark of demoralization. Compr Psychiatry 1982;23(4):353-363. [doi: 10.1016/0010-440x(82)90085-2] [Medline: 7116830]
2. Clarke DM, Kissane DW. Demoralization: its phenomenology and importance. Aust N Z J Psychiatry 2002 Dec;36(6):733-742. [doi: 10.1046/j.1440-1614.2002.01086.x] [Medline: 12406115]
3. de Figueiredo JM. Depression and demoralization: phenomenologic differences and research perspectives. Compr Psychiatry 1993;34(5):308-311. [doi: 10.1016/0010-440x(93)90016-w] [Medline: 8306640]
4. Hua R, Bhatarasakoon P. Demoralization: a concept analysis through a scoping review methodology. Int J Nurs Stud Adv 2025 Jun;8:100272. [doi: 10.1016/j.ijnsa.2024.100272] [Medline: 39737059]

XSL•FO
RenderX

5. Belvederi Murri M, Caruso R, Ounalli H, et al. The relationship between demoralization and depressive symptoms among patients from the general hospital: network and exploratory graph analysis. J Affect Disord 2020 Nov 1;276:137-146. [doi: 10.1016/j.jad.2020.06.074] [Medline: 32697691]
6. de Figueiredo JM. Distress, demoralization and psychopathology: diagnostic boundaries. Eur J Psychiat 2013;27(1):61-73. [doi: 10.4321/S0213-61632013000100008]
7. Grassi L, Pasquini M, Kissane D, et al. Exploring and assessing demoralization in patients with non-psychotic affective disorders. J Affect Disord 2020 Sep 1;274:568-575. [doi: 10.1016/j.jad.2020.05.043] [Medline: 32663989]
8. Grassi L, Nanni MG. Demoralization syndrome: new insights in psychosocial cancer care. Cancer 2016 Jul 15;122(14):2130-2133. [doi: 10.1002/cncr.30022] [Medline: 27171755]
9. Kissane DW, Clarke DM, Street AF. Demoralization syndrome--a relevant psychiatric diagnosis for palliative care. J Palliat Care 2001;17(1):12-21. [doi: 10.1177/082585970101700103] [Medline: 11324179]
10. Mangelli L, Fava GA, Grandi S, et al. Assessing demoralization and depression in the setting of medical disease. J Clin Psychiatry 2005 Mar;66(3):391-394. [doi: 10.4088/jcp.v66n0317] [Medline: 15766307]
11. Elfil M, Ahmed N, Alapati A, et al. Suicidal risk and demoralization in Parkinson disease. J Neurol 2020 Apr;267(4):966-974. [doi: 10.1007/s00415-019-09632-2] [Medline: 31802218]
12. Ko KT, Lin CJ, Pi SH, Li YC, Fang CK. Demoralization syndrome among elderly patients with cancer disease. Int J Gerontol 2018 Mar;12(1):12-16. [doi: 10.1016/j.ijge.2018.01.001]
13. Nanni MG, Caruso R, Travado L, et al. Relationship of demoralization with anxiety, depression, and quality of life: a Southern European study of Italian and Portuguese cancer patients. Psychooncology 2018 Nov;27(11):2616-2622. [doi: 10.1002/pon.4824] [Medline: 29943491]
14. Tang PL, Wang HH, Chou FH. A systematic review and meta-analysis of demoralization and depression in patients with cancer. Psychosomatics 2015;56(6):634-643. [doi: 10.1016/j.psym.2015.06.005] [Medline: 26411374]
15. Belvederi Murri M, Grassi L, Caruso R, et al. Depressive symptom complexes of community-dwelling older adults: a latent network model. Mol Psychiatry 2022 Feb;27(2):1075-1082. [doi: 10.1038/s41380-021-01310-y] [Medline: 34642459]
16. Ramm M, Jedamzik J, Lenz P, Jürgens L, Heuft G, Conrad R. Older adults coping with critical life events - results of the revised demoralization scale in a representative sample of older adulthood. Front Psychiatry 2024;15:1389021. [doi: 10.3389/fpsyt.2024.1389021] [Medline: 38800056]
17. Grassi L, Caruso R, Da Ronch C, et al. Quality of life, level of functioning, and its relationship with mental and physical disorders in the elderly: results from the MentDis_ICF65+ study. Health Qual Life Outcomes 2020 Mar 6;18(1):61. [doi: 10.1186/s12955-020-01310-6] [Medline: 32143635]
18. Kaneriya S, Mendonca C, Singh E, Wilkins KM. Distinguishing distress in the context of aging: Demoralization vs. Depression. Curr Geri Rep 2023;12(3):128-137. [doi: 10.1007/s13670-023-00390-8]
19. Gao CX, Dwyer D, Zhu Y, et al. An overview of clustering methods with guidelines for application in mental health research. Psychiatry Res 2023 Sep;327:115265. [doi: 10.1016/j.psychres.2023.115265] [Medline: 37348404]
20. Witten IH, Frank E, Hall MA, Pal CJ. Data Mining: Practical Machine Learning Tools and Techniques, 4th edition: Morgan Kaufmann Publishers; 2016.
21. Feczko E, Fair DA. Methods and challenges for assessing heterogeneity. Biol Psychiatry 2020 Jul 1;88(1):9-17. [doi: 10.1016/j.biopsych.2020.02.015] [Medline: 32386742]
22. Rafanelli C, Guidi J, Gostoli S, Tomba E, Porcelli P, Grandi S. Subtyping demoralization in the medically ill by cluster analysis. Eur J Psychiat 2013;27(1):7-17. [doi: 10.4321/S0213-61632013000100002]
23. Ferrara M, Gentili E, Belvederi Murri M, et al. Establishment of a public mental health database for research purposes in the ferrara province: development and preliminary evaluation study. JMIR Med Inform 2023 Aug 9;11:e45523. [doi: 10.2196/45523] [Medline: 37584563]
24. Andreas S, Härter M, Volkert J, et al. The MentDis_ICF65+ study protocol: prevalence, 1-year incidence and symptom severity of mental disorders in the elderly and their relationship to impairment, functioning (ICF) and service utilisation. BMC Psychiatry 2013 Feb 18;13(1):62. [doi: 10.1186/1471-244X-13-62] [Medline: 23418914]
25. Volkert J, Härter M, Dehoust MC, et al. Study approach and field work procedures of the MentDis_ICF65+ project on the prevalence of mental disorders in the older adult European population. BMC Psychiatry 2017 Nov 16;17(1):366. [doi: 10.1186/s12888-017-1534-5] [Medline: 29145800]
26. Wittchen HU, Strehle J, Gerschler A, et al. Measuring symptoms and diagnosing mental disorders in the elderly community: the test-retest reliability of the CIDI65. Int J Methods Psychiatr Res 2015 Jun;24(2):116-129. [doi: 10.1002/mpr.1455] [Medline: 25308743]
27. Kissane DW, Wein S, Love A, Lee XQ, Kee PL, Clarke DM. The Demoralization Scale: a report of its development and preliminary validation. J Palliat Care 2004;20(4):269-276. [Medline: 15690829]
28. Grassi L, Costantini A, Kissane D, et al. The factor structure and use of the Demoralization Scale (DS-IT) in Italian cancer patients. Psychooncology 2017 Nov;26(11):1965-1971. [doi: 10.1002/pon.4413] [Medline: 28263003]
29. Belvederi Murri M, Zerbinati L, Ounalli H, et al. Assessing demoralization in medically ill patients: factor structure of the Italian version of the demoralization scale and development of short versions with the item response theory framework. J Psychosom Res 2020 Jan;128:109889. [doi: 10.1016/j.jpsychores.2019.109889] [Medline: 31812103]

XSL•FO
RenderX

30. Everitt BS, Landau S, Leese M, Stahl D. Cluster Analysis: John Wiley & Sons; 2011.
31. Milligan GW, Cooper MC. A study of standardization of variables in cluster analysis. J Classif 1988 Sep;5(2):181-204. [doi: 10.1007/BF01897163]
32. Gower JC. A general coefficient of similarity and some of its properties. Biometrics 1971 Dec;27(4):857-871. [doi: 10.2307/2528823]
33. Kaufman L, Rousseeuw PJ. Finding Groups in Data: An Introduction To Cluster Analysis: John Wiley & Sons, Inc; 1990.
34. Hennig C. Cluster-wise assessment of cluster stability. Comput Stat Data Anal 2007 Sep;52(1):258-271. [doi: 10.1016/j.csda.2006.11.025]
35. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res 2008;9:2579-2605 [FREE Full text]
36. R: a language and environment for statistical computing. : R Foundation for Statistical Computing; 2021 URL: https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf [accessed 2026-06-19]
37. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. Cluster: cluster analysis basics and extensions. CRAN (The Comprehensive R Archive Network). 2023. URL: https://cran.r-project.org/web/packages/cluster/cluster.pdf [accessed 2026-06-19]
38. Hennig C. Fpc: flexible procedures for clustering. The Comprehensive R Archive Network (CRAN). 2026. URL: https://cran.r-project.org/web/packages/fpc/index.html [accessed 2026-07-13]
39. Fava GA, Guidi J, Porcelli P, et al. A cluster analysis-derived classification of psychological distress and illness behavior in the medically ill. Psychol Med 2012 Feb;42(2):401-407. [doi: 10.1017/S0033291711001231] [Medline: 24438853]
40. Grassi L, Rossi E, Sabato S, Cruciani G, Zambelli M. Diagnostic criteria for psychosomatic research and psychosocial variables in breast cancer patients. Psychosomatics 2004;45(6):483-491. [doi: 10.1176/appi.psy.45.6.483] [Medline: 15546825]
41. Klein DF. Endogenomorphic depression. a conceptual and terminological revision. Arch Gen Psychiatry 1974 Oct;31(4):447-454. [doi: 10.1001/archpsyc.1974.01760160005001] [Medline: 4420562]
42. World mental health report: transforming mental health for all. : World Health Organization; 2022 URL: https://www.who.int/publications/i/item/9789240049338 [accessed 2026-06-19]
43. Cohen LS, Soares CN, Vitonis AF, Otto MW, Harlow BL. Risk for new onset of depression during the menopausal transition: the Harvard study of moods and cycles. Arch Gen Psychiatry 2006 Apr;63(4):385-390. [doi: 10.1001/archpsyc.63.4.385] [Medline: 16585467]
44. Belvederi Murri M, Folesani F, Azzolina D, et al. A tale of two constructs: combined assessment of demoralization and subjective incompetence. Psychol Health Med 2024 Oct;29(9):1635-1651. [doi: 10.1080/13548506.2024.2405749] [Medline: 39302008]
45. Cockram CA, Doros G, de Figueiredo JM. Diagnosis and measurement of subjective incompetence: the clinical hallmark of demoralization. Psychother Psychosom 2009;78(6):342-345. [doi: 10.1159/000235737] [Medline: 19713728]
46. Van As BAL, Imbimbo E, Franceschi A, Menesini E, Nocentini A. The longitudinal association between loneliness and depressive symptoms in the elderly: a systematic review. Int Psychogeriatr 2022 Jul;34(7):657-669. [doi: 10.1017/S1041610221000399] [Medline: 33849675]
47. Zhang H, Yang L, Xing X, Huang Y, Lv X. Incidence and influencing factors of demoralization syndrome in elderly maintenance hemodialysis patients: a multi-center cross-sectional study. Int J Gen Med 2025;18:787-797. [doi: 10.2147/IJGM.S511458] [Medline: 39974563]
48. Wang Y, Zheng C, Zhao Y, et al. Latent profile analysis of demoralization syndrome in elderly stroke patients with disability. Geriatr Nurs 2025 Sep;65:103447. [doi: 10.1016/j.gerinurse.2025.103447] [Medline: 40618716]
49. Antonaci F, Nappi G, Galli F, Manzoni GC, Calabresi P, Costa A. Migraine and psychiatric comorbidity: a review of clinical findings. J Headache Pain 2011 Apr;12(2):115-125. [doi: 10.1007/s10194-010-0282-4] [Medline: 21210177]
50. Harden CL. The co-morbidity of depression and epilepsy: epidemiology, etiology, and treatment. Neurology 2002 Sep 24;59(6 Suppl 4):S48-S55. [doi: 10.1212/wnl.59.6_suppl_4.s48] [Medline: 12270969]
51. Culpepper L. Generalized anxiety disorder and medical illness. J Clin Psychiatry 2009;70 Suppl 2:20-24. [doi: 10.4088/jcp.s.7002.04] [Medline: 19371503]
52. Storer B, Kershaw KA, Braund TA, et al. The prevalence of anxiety disorders in dermatology outpatients: a systematic review and meta-analysis. J Eur Acad Dermatol Venereol 2023 Apr 28. [doi: 10.1111/jdv.19151] [Medline: 37118899]
53. Rodin G, Lo C, Rydall A, et al. Managing Cancer and Living Meaningfully (CALM): a randomized controlled trial of a psychological intervention for patients with advanced cancer. J Clin Oncol 2018 Aug 10;36(23):2422-2432. [doi: 10.1200/JCO.2017.77.1097] [Medline: 29958037]
54. Fava GA, Guidi J. Clinical characterization of demoralization. Psychother Psychosom 2023;92(3):139-147. [doi: 10.1159/000530760] [Medline: 37231894]
55. Romaniello C, Romanazzo S, Cosci F. Clinimetric properties of the diagnostic criteria for psychosomatic research among the elderly. Clin Psychol Psychother 2023 May;30(3):611-619. [doi: 10.1002/cpp.2822] [Medline: 36607260]
56. Kang H. The prevention and handling of the missing data. Korean J Anesthesiol 2013 May;64(5):402-406. [doi: 10.4097/kjae.2013.64.5.402] [Medline: 23741561]

XSL•FO
RenderX

57. Woźniewicz A, Cosci F. Clinical utility of demoralization: a systematic review of the literature. Clin Psychol Rev 2023 Feb;99:102227. [doi: 10.1016/j.cpr.2022.102227] [Medline: 36462221]

Edited by H LaMonica; submitted 23.Sep.2025; peer-reviewed by P Voukelatou, Y Chen; revised version received 23.Feb.2026; accepted 06.May.2026; published 19.Aug.2026.

Please cite as:
Onofrio A, Ferrara M, Domenicano I, Belvederi Murri M, Toffanin T, Nanni MG, Canuto A, Weber K, Volkert J, Schulz H, Härter M, Andreas S, Grassi L
Demoralization and Mental Disorders in Community-Dwelling Older Adults: Cluster Analysis
JMIR Aging 2026;9:e84657
URL: https://aging.jmir.org/2026/1/e84657
doi:10.2196/84657

© Alice Onofrio, Maria Ferrara, Ilaria Domenicano, Martino Belvederi Murri, Tommaso Toffanin, Maria Giulia Nanni, Alessandra Canuto, Kerstin Weber, Jana Volkert, Holger Schulz, Martin Härter, Sylke Andreas, Luigi Grassi. Originally published in JMIR Aging (https://aging.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Delirium Occurrence and Associated Factors Among Patients With Dementia and Mild Cognitive Impairment Admitted to Long-Term Care Hospitals: Cross-Sectional Study

Kyoung Ja Moon[1], PhD, RN; Yong-Hwan Hyeon[2], PhD, RN

[1]Department of Nursing, College of Nursing, Keimyung University, Daegu, Republic of Korea

[2]Department of Nursing, Ulsan College of Nursing, Ulsan College, Ulsan, Republic of Korea

Corresponding Author:
Kyoung Ja Moon, PhD, RN
Department of Nursing
College of Nursing
Keimyung University
1095 Dalgubeol-daero
Daegu, 42601
Republic of Korea
Phone: 82 53 258 7663
Fax: 82 53 258 7616
Email: kjmoon2150@kmu.ac.kr

Abstract

Background: Delirium is common among older adults, particularly those with cognitive impairment. However, delirium among patients with dementia or mild cognitive impairment (MCI) in long-term care (LTC) hospitals has received limited attention.

Objective: This study aimed to examine delirium occurrence and associated factors among patients with dementia or MCI in LTC hospitals.

Methods: This cross-sectional study included 142 patients aged 65 years or older with dementia (n=100, 70.4%) or MCI (n=42, 29.6%) at 2 LTC hospitals in South Korea between September 2021 and August 2022. Delirium occurrence was defined as at least 1 positive Short Confusion Assessment Method assessment during twice-daily assessments over 4 weeks, including the first assessment. Multivariable logistic regression examined associated factors overall and within each cognitive group.

Results: Delirium was identified in 55 (38.7%) of 142 patients. The proportion was 41% (41/100) in patients with dementia and 33.3% (14/42) in patients with MCI. Hypoactive delirium was the most common subtype (33/55, 60%). Overall, female sex (adjusted odds ratio [aOR] 4.344, 95% CI 1.794-10.515; *P*=.001), immobility (aOR 4.953, 95% CI 1.857-13.205; *P*=.001), and analgesic use (aOR 3.641, 95% CI 1.262-10.501; *P*=.02) were associated with delirium. Infection was inversely associated with delirium (aOR 0.251, 95% CI 0.101-0.623; *P*=.003). In patients with dementia, female sex, immobility, analgesic use, and pain were associated with delirium. Diaper use was associated with delirium in patients with MCI.

Conclusions: Delirium was common among patients with dementia or MCI in LTC hospitals. The findings support regular delirium assessment, with attention to mobility and pain-related care needs. Further studies are needed to clarify the factors associated with delirium in each cognitive group.

(JMIR Aging 2026;9:e90434) doi:10.2196/90434

KEYWORDS

delirium; dementia; mild cognitive impairment; long-term care; cross-sectional study

Introduction

Delirium is an acute disturbance in attention, awareness, and cognition that fluctuates over time [1]. Delirium occurs in 21% of patients after hip fracture surgery [2], 48% of patients during follow-up in long-term care (LTC) facilities [3], and 24.6% of patients on admission to acute hospitals [4]. Among hospitalized older adults with dementia, the prevalence of delirium is 48.9% [5]. Although these proportions differ by population and assessment period, delirium is common across care settings. Delirium is associated with cognitive decline [6], increased

XSL•FO
RenderX

mortality, longer hospital stays, institutionalization, higher health care costs, and increased caregiver burden [7-9]. Early recognition is particularly important for older patients with cognitive impairment, whose cognitive and functional difficulties may complicate delirium assessment and care.

Early identification of patients at risk is important for delirium prevention and management [5,7]. Delirium risk is related to both underlying vulnerability and acute clinical stressors [10]. Assessment therefore needs to consider cognitive status, physical illness, medication use, and daily care needs. Multicomponent nonpharmacological interventions can help prevent delirium in hospitalized patients outside intensive care units (ICUs) [11]. However, evidence for preventive interventions in institutional LTC remains limited [12]. Differences in patient characteristics and care needs require attention when applying findings from acute hospitals to LTC hospitals. Identifying factors associated with delirium in LTC hospitals can help nurses assess patients at risk and develop appropriate preventive care.

Most delirium research has focused on patients in ICUs, acute hospitals, and surgical settings. Delirium is also common among patients with dementia in LTC hospitals [3]. Patients with dementia who experience delirium have longer hospital stays and poorer outcomes than those without delirium [5,13]. Changes in attention, cognition, and behavior may be difficult to distinguish from existing dementia symptoms, making comparison with the patient's usual condition important [13]. In residential care, delirium prevalence ranges from 1% to 70.3%, with the highest proportion in a population with dementia. Cognitive impairment and dementia are consistently associated with delirium in these settings [14]. However, the wide variation in patient populations and assessment periods limits direct comparisons across studies.

Patients with mild cognitive impairment (MCI) also require attention in delirium care. MCI is associated with progression to dementia, although the rate varies by population, diagnostic criteria, and follow-up period [15]. Physical inactivity, smoking, diabetes, obesity, and hypertension are modifiable factors associated with dementia [16]. Dementia also places a substantial financial burden on care in South Korea [17]. These clinical and care burdens support the need to address cognitive health in patients with MCI as well as those with dementia. However, the clinical and care-related factors associated with delirium in patients with MCI in LTC hospitals remain unclear. Examining both groups may help identify their delirium care needs.

Older patients in LTC hospitals often have chronic illnesses and cognitive impairment that require ongoing nursing care. Nurses observe changes in patients' condition during daily care and have an important role in recognizing possible delirium. Information on delirium occurrence, motor subtypes, and associated factors can help nurses identify patients who need further assessment. Findings from patients with dementia and those with MCI may also help guide care according to cognitive status. However, evidence specific to these patients in LTC hospitals is needed to support delirium assessment and management.

This study examined delirium occurrence, motor subtypes, and associated factors among patients with dementia or MCI in LTC hospitals. Factors associated with delirium were examined in all participants and in the dementia and MCI groups. The findings may help guide nursing assessment and delirium care for patients with cognitive impairment in LTC hospitals.

Methods

Study Design

This observational study used a cross-sectional analysis of patient-level data collected through repeated assessments over 4 weeks. Delirium occurrence was defined as at least 1 positive Short Confusion Assessment Method (S-CAM) assessment during this period, including the first assessment. Each patient was classified once as having or not having delirium, regardless of the number of positive assessments. New episodes after a delirium-free baseline were not analyzed separately. The proportion of patients with delirium therefore represents the observed 4-week period prevalence, not incidence. The study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guideline for cross-sectional studies [18] (Multimedia Appendix 1).

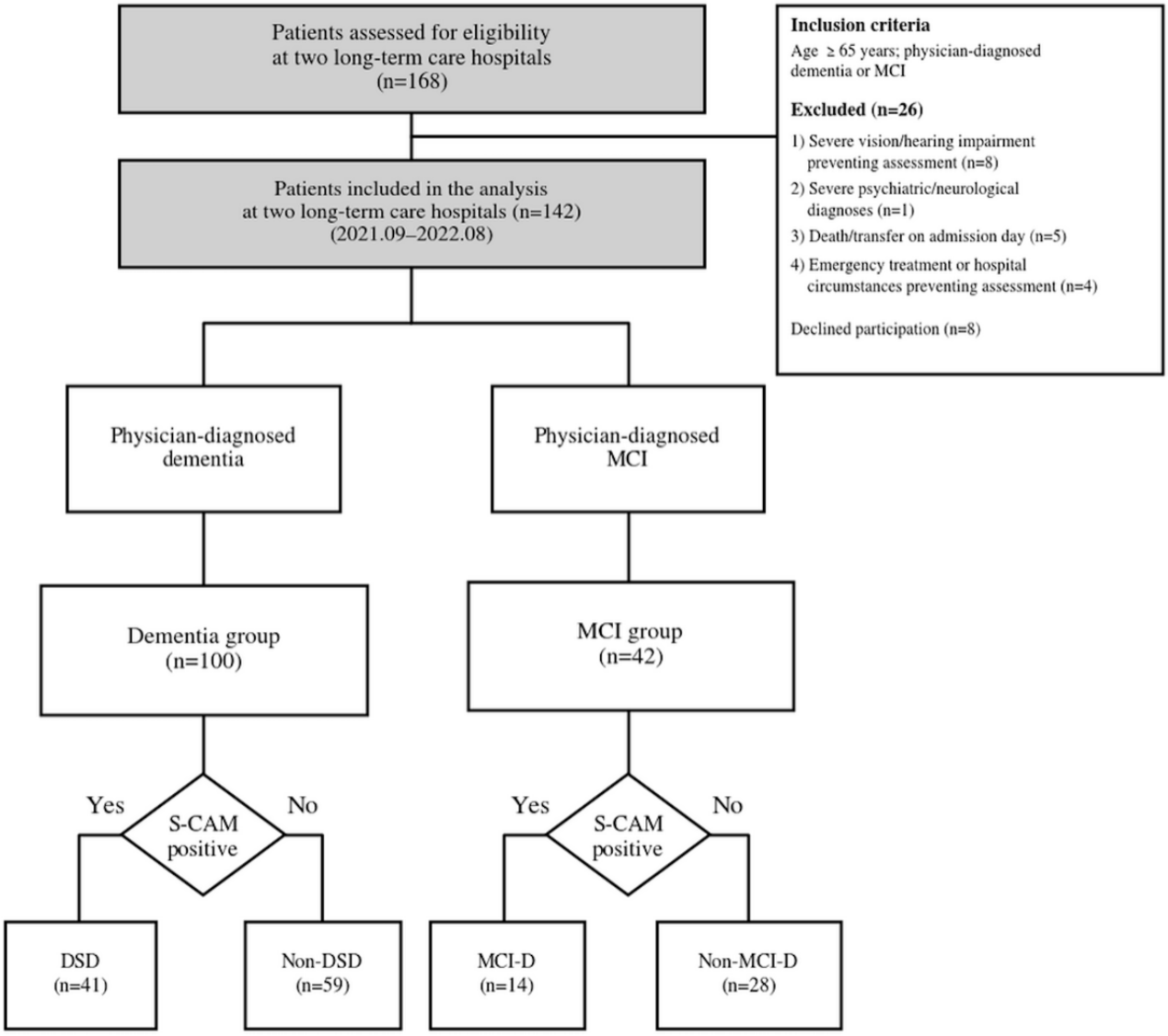
Participants and Recruitment

The minimum sample size was 136 patients, calculated using G*Power (version 3.1.9; Heinrich-Heine-University Düsseldorf) for multiple linear regression (f^2=0.15, α=.05, power=0.90, and 8 predictors).

Participants were patients aged 65 years or older at 2 LTC hospitals in cities D and K, South Korea, between September 2021 and August 2022. Each hospital had approximately 250 to 300 beds and mainly cared for older patients, including those with dementia. Patients were eligible if they had a physician-recorded diagnosis of dementia or MCI. After approval from the directors of nursing, recruitment notices were posted on ward notice boards. Head nurses received 2 training sessions and served as research assistants. They checked eligibility and explained the study individually using the recruitment notice and participant information sheet. Patients who expressed interest after reading the notice were also considered. Written consent was obtained from patients or their legal guardians. Of 168 patients assessed, 18 (10.7%) met the exclusion criteria, and 8 (4.8%) declined participation. The remaining 142 (84.5%) patients were included in the analysis (Figure 1).

XSL•FO
RenderX

Figure 1. Recruitment and classification in this cross-sectional study. DSD: delirium superimposed on dementia; MCI: mild cognitive impairment; MCI-D: MCI with delirium; non-DSD: dementia without delirium; non-MCI-D: MCI without delirium.



Patients were excluded because of severe visual or hearing impairment that prevented S-CAM assessment (n=8), severe psychiatric or neurological diagnoses (n=1), death or transfer on the day of admission (n=5), or inability to undergo assessment because of emergency treatment or hospital circumstances (n=4).

Patients were classified into the dementia group (n=100, 70.4%) or the MCI group (n=42, 29.6%) according to their physician-recorded diagnosis. The Korean version of the Mini-Mental State Examination (K-MMSE) cutoffs were not used for eligibility or group assignment. The groups were further divided according to whether patients had any positive S-CAM assessment during the 4 weeks (Figure 1): (1) patients with dementia and at least 1 positive S-CAM assessment (delirium superimposed on dementia [DSD]; n=41), (2) patients with dementia and no positive S-CAM assessment during the observation period (dementia without delirium [non-DSD]; n=59), (3) patients with MCI and at least 1 positive S-CAM assessment (MCI with delirium [MCI-D]; n=14), and (4) patients with MCI and no positive S-CAM assessment during the observation period (MCI without delirium [non-MCI-D]; n=28).

Measures

General and Clinical Characteristics

Baseline characteristics included age, sex, admission source, Charlson Comorbidity Index (CCI) score, BMI (kg/m²), K-MMSE score, number of medications, smoking, alcohol consumption, and comorbidities. The K-MMSE [19] score was used as a continuous measure of cognitive status rather than to classify patients as having dementia or MCI.

Clinical and care-related variables recorded during the 4 weeks included visual or hearing impairment, dehydration (blood urea nitrogen >18 mg/dL), malnutrition, fluid and electrolyte imbalance, infection, pain, sleep deprivation, physical restraint use, immobility, recent surgery, falls, pressure ulcers, analgesic use, blood transfusion, urinary catheterization, enteral feeding, diaper use, and sedative use. Haloperidol was used only in patients with delirium and was not included in the regression models. A clinical condition was coded as present if it was recorded at least once during the assessment period. Infection was identified through a physician diagnosis, laboratory evidence, or antibiotic use. Immobility was assessed through

XSL•FO
RenderX

routine activities-of-daily-living assessments and observation. Pain was assessed using recorded pain scores and pain medication use. Analgesic and sedative use were recorded as yes or no variables.

Delirium Assessment and Classification

Delirium was assessed using the S-CAM, based on the assessment guidelines [20] and previous research [3]. The S-CAM includes acute onset and fluctuating course, inattention, disorganized thinking, and altered level of consciousness. A positive result requires the presence of the first 2 features, along with either disorganized thinking or an altered level of consciousness. Nurses identified acute changes by comparing patients' current condition with their usual cognitive and behavioral status. Motor subtypes were classified as hypoactive or hyperactive according to the main observed pattern. Patients who alternated between hypoactive and hyperactive features were classified as having mixed delirium.

Delirium was assessed twice daily for 4 weeks, in the morning (8-10 AM) and afternoon (4-6 PM). Head nurses served as research assistants, and the nurse in charge of the participating unit performed the afternoon assessments. Before data collection, head nurses received 2 training sessions. A head nurse and a researcher then assessed the same 5 to 6 patients at the same time and compared their checklist ratings. Training continued until an agreement rate of approximately 98%-99% was achieved.

Statistical Analysis

Data were analyzed using SPSS (version 23.0; IBM Corp). Descriptive statistics summarized participant characteristics and delirium occurrence. The proportion with delirium was calculated as the number of patients with at least 1 positive S-CAM assessment during 4 weeks divided by the number analyzed (142 overall, 100 with dementia, and 42 with MCI), multiplied by 100. Group differences were examined using 1-way ANOVA, Pearson chi-square, or Fisher-Freeman-Halton exact tests. Fisher exact tests compared delirium occurrence and motor subtypes between cognitive groups. All tests were 2-sided (P<.05). No analyzed values were missing.

Multivariable logistic regression examined factors associated with delirium in the overall sample and separately within the dementia and MCI groups. The dependent variable was delirium occurrence, coded as at least 1 positive S-CAM assessment during 4 weeks vs none. The overall model included age, sex, CCI score, K-MMSE score, infection, immobility, and analgesic use. Subgroup variables were selected using univariable P<.05, previous research, and clinical relevance. Variables were entered simultaneously, with male sex and absence of binary characteristics as references. Adjusted odds ratios (aORs) and 95% CIs were reported. Each aOR describes the association between a factor and delirium within the analyzed sample, conditional on the other variables in that model. Multicollinearity and model fit were assessed using variance inflation factors and the Hosmer-Lemeshow test.

Ethical Considerations

The study was approved by the Institutional Review Board of Keimyung University (IRB 40525-202012-HR-073-06) and was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from patients or their legal guardians. Paper consent and data collection forms were stored in a locked cabinet and later destroyed. Anonymized electronic data are retained in encrypted form. The approved plan specified data retention for 3 years after study completion. Participants received a cotton T-shirt worth approximately ₩10,000 (₩1=US $0.000757 as of August 2022) as compensation.

Results

Participant Characteristics

Of 168 patients assessed for eligibility, 18 met exclusion criteria, and 8 declined participation. The analysis included 142 patients, whose mean age was 80.92 (SD 7.26) years; 58.5% (n=83) were female, and 41.5% (n=59) were male. Smokers accounted for 2.1% (n=3) of patients, and none reported alcohol consumption. Most patients (n=93, 65.5%) were admitted from home.

The mean BMI was 21.97 (SD 3.44) kg/m^2, the mean CCI score was 5.54 (SD 1.81), and the mean K-MMSE score was 10.89 (SD 7.35). Participants were taking an average of 7.50 (SD 4.81) medications. Of the 142 participants, hypertension, diabetes, and degenerative arthritis were present in 53.5% (n=76), 26.1% (n=37), and 2.1% (n=3), respectively. In addition, 54.2% (n=77) had cancer, and 1.4% (n=2) had a history of COVID-19. A history of stroke was recorded in 18.3% (n=26) of patients and a history of brain injury in 10.6% (n=15). Moreover, 1 (0.7%) patient had a visual or hearing impairment.

Among the 142 participants, 4.9% (n=7) had a blood urea nitrogen level above 18 mg/dL, 39.4% (n=56) had malnutrition, and 26.8% (n=38) had a fluid or electrolyte imbalance. A postoperative state was recorded in 7.7% (n=11) of patients, infection in 39.4% (n=56), and pain in 19.7% (n=28). Postoperative state referred to surgery performed at another hospital before admission to the LTC facility and before the delirium assessment period. The maximum interval since surgery and the type of procedure were not recorded.

Sleep deprivation was recorded in 8.5% (n=12) of patients and physical restraint use in 3.5% (n=5). Immobility was recorded in 25.4% (n=36) of patients, a history of falls in 18.3% (n=26), and pressure ulcers in 16.9% (n=24). Analgesics were used by 16.2% (n=23) of patients. Haloperidol was administered to 14 (9.9%) patients: 6 with DSD and 8 with MCI-D. Blood transfusions were recorded in 0.7% (n=1).

Of the 142 participants, urinary catheters were used by 1.4% (n=2), enteral feeding by 6.3% (n=9), and diapers by 60.6% (n=86). Sedative use was recorded in 19.7% (n=28) of participants. Using the full sample (N=142) as the denominator for dementia types, Alzheimer disease was the most common (n=81, 57%), followed by vascular dementia (n=15, 10.6%), Lewy body dementia (n=3, 2.1%), and mixed dementia (n=1, 0.7%; Table 1).

XSL•FO
RenderX

Table 1. Participants' general and clinical characteristics (N=142).

Variable	Participants
Age (years), mean (SD)	80.92 (7.26)
Sex, n (%)	
Male	59 (41.5)
Female	83 (58.5)
Smoking, n (%)	3 (2.1)
Alcohol, n (%)	0 (0)
Residence before admission, n (%)	
House	93 (65.5)
Hospital	41 (28.9)
Others	8 (5.6)
BMI (kg/m^2), mean (SD)	21.97 (3.44)
Charlson Comorbidity Index score, mean (SD)	5.54 (1.81)
K-MMSE[a], mean (SD)	10.89 (7.35)
Number of medications, mean (SD)	7.50 (4.81)
Present disease (multiple responses permitted), n (%)	
Hypertension	76 (53.5)
Diabetes	37 (26.1)
Degenerative arthritis	3 (2.1)
Cancer	77 (54.2)
COVID-19	2 (1.4)
Stroke	26 (18.3)
History of brain damage	15 (10.6)
Vision disturbances or hearing loss	1 (0.7)
Dehydration (BUN[b]>18 mg/dL)	7 (4.9)
Malnutrition	56 (39.4)
Electrolyte imbalance	38 (26.8)
Surgery	11 (7.7)
Infection	56 (39.4)
Pain	28 (19.7)
Insomnia	12 (8.5)
Restraints	5 (3.5)
Immobility	36 (25.4)
Falls	26 (18.3)
Pressure ulcer	24 (16.9)
Analgesic use	23 (16.2)
Delirium-related medication (haloperidol)	14 (9.9)
Transfusion	1 (0.7)
Indwelling urinary catheterization	2 (1.4)
Enteral feeding	9 (6.3)
Diaper use	86 (60.6)
Sedative use	28 (19.7)

XSL•FO
RenderX

Variable	Participants
Dementia	100 (70.4)
Type of dementia, n (%)	
Alzheimer	81 (57)
Vascular	15 (10.6)
Mixed	1 (0.7)
Lewy body	3 (2.1)

[a]K-MMSE: Korean version of the Mini-Mental State Examination.

[b]BUN: blood urea nitrogen.

Delirium Occurrence and Motor Subtypes

During the 4-week assessment period, delirium was identified in 38.7% (55/142) of patients, including 41% (41/100) of patients with dementia and 33.3% (14/42) of patients with MCI. Among patients with delirium, the hypoactive type was most common (33/55, 60%), followed by hyperactive (18/55, 32.7%) and mixed (4/55, 7.3%) types. In the dementia group, hypoactive, hyperactive, and mixed delirium accounted for 65.9% (27/41), 26.8% (11/41), and 7.3% (3/41) of patients, respectively. In the MCI group, the corresponding proportions were 42.9% (6/14), 50% (7/14), and 7.1% (1/14; Table 2).

Table 2. Delirium occurrence and motor subtypes.

Variable	Group, n (%)			*P* value
	Overall (N=142)	Dementia (n=100)	MCI[a] (n=42)	
Delirium				.45
Yes	55 (38.7)	41 (41)	14 (33.3)	
Motor subtype among patients with delirium				.25
Hypoactive	33 (60)	27 (65.9)	6 (42.9)	
Hyperactive	18 (32.7)	11 (26.8)	7 (50)	
Mixed	4 (7.3)	3 (7.3)	1 (7.1)	

[a]MCI: mild cognitive impairment.

Comparison of Variables by Cognitive Status and Delirium Occurrence

The 4 groups differed in sex (*P*=.006), admission source (*P*<.001), CCI score (*P*=.005), K-MMSE score (*P*<.001), number of medications (*P*<.001), cancer (*P*<.001), history of COVID-19 (*P*=.047), infection (*P*=.03), immobility (*P*<.001), analgesic use (*P*=.03), urinary catheterization (*P*=.009), and diaper use (*P*<.001; Table 3). Haloperidol use also differed across groups (*P*<.001) and was recorded only in patients with delirium.

XSL•FO
RenderX

Table 3. Comparison of variables by cognitive status and delirium occurrence (N=142).

Variable	Dementia		Mild cognitive impairment		*F* test (*df*)	Chi-square (*df*)	*P* value
	DSD[a] (n=41)	Non-DSD[b] (n=59)	MCI-D[c] (n=14)	Non-MCI-D[d] (n=28)			
Age (years), mean (SD)	81.37 (7.95)	80.59 (7.22)	80.79 (7.03)	81.04 (6.75)	0.093 (3, 138)	__[e]	.96
Sex, n (%)					—	12.5 (3)	.006
Male	11 (26.8)	31 (52.5)	2 (14.3)	15 (53.6)			
Female	30 (73.2)	28 (47.5)	12 (85.7)	13 (46.4)			
Smoking, n (%)					—	—	.85
Yes	1 (2.4)	1 (1.7)	0 (0)	1 (3.6)			
Alcohol consumption, n (%)					—	—	—
Yes	0 (0)	0 (0)	0 (0)	0 (0)			
Residence before admission, n (%)					—	—	<.001
Home	22 (53.7)	33 (55.9)	12 (85.7)	26 (92.9)			
Hospital	17 (41.5)	22 (37.3)	2 (14.3)	0 (0)			
Others	2 (4.9)	4 (6.8)	0 (0)	2 (7.1)			
BMI (kg/m^2), mean (SD)	21.89 (3.77)	21.43 (3.24)	23.29 (3.20)	22.56 (3.35)	1.468 (3, 138)	—	.23
CCI[f] score, mean (SD)	5.93 (2.00)	5.81 (1.58)	5.36 (2.02)	4.50 (1.53)	4.503 (3, 138)	—	.005
K-MMSE[g] score, mean (SD)	7.05 (4.63)	7.00 (5.23)	19.50 (1.70)	20.43 (1.32)	92.558 (3, 138)	—	<.001
Number of medications, mean (SD)	10.24 (3.20)	9.37 (2.97)	3.29 (4.66)	1.64 (3.51)	51.634 (3, 138)	—	<.001
Hypertension, n (%)					—	1.5 (3)	.68
Yes	24 (58.5)	28 (47.5)	8 (57.1)	16 (57.1)			
Diabetes, n (%)					—	2.9 (3)	.40
Yes	12 (29.3)	16 (27.1)	5 (35.7)	4 (14.3)			
Degenerative arthritis, n (%)					—	—	.50
Yes	1 (2.4)	1 (1.7)	1 (7.1)	0 (0)			
Cancer, n (%)					—	28.3 (3)	<.001
Yes	30 (73.2)	38 (64.4)	5 (35.7)	4 (14.3)			
COVID-19, n (%)					—	—	.047
Yes	0 (0)	0 (0)	0 (0)	2 (7.1)			
Stroke, n (%)					—	6.9 (3)	.08
Yes	6 (14.6)	16 (27.1)	0 (0)	4 (14.3)			
History of brain damage, n (%)					—	—	.60
Yes	5 (12.2)	7 (11.9)	2 (14.3)	1 (3.6)			
Vision or hearing impairment, n (%)					—	—	>.99
Yes	0 (0)	1 (1.7)	0 (0)	0 (0)			
Dehydration (BUN[h] >18 mg/dL), n (%)					—	—	.20
Yes	3 (7.3)	1 (1.7)	0 (0)	3 (10.7)			
Malnutrition, n (%)					—	2.0 (3)	.58
Yes	15 (36.6)	27 (45.8)	4 (28.6)	10 (35.7)			
Electrolyte imbalance, n (%)					—	3.5 (3)	.32
Yes	15 (36.6)	12 (20.3)	3 (21.4)	8 (28.6)			

XSL•FO
RenderX

Variable	Dementia		Mild cognitive impairment		*F* test (*df*)	Chi-square (*df*)	*P* value
	DSD[a] (n=41)	Non-DSD[b] (n=59)	MCI-D[c] (n=14)	Non-MCI-D[d] (n=28)			
Surgery, n (%)					—	—	.15
Yes	4 (9.8)	3 (5.1)	3 (21.4)	1 (3.6)			
Infection, n (%)					—	8.8 (3)	.03
Yes	10 (24.4)	26 (44.1)	4 (28.6)	16 (57.1)			
Pain, n (%)					—	5.9 (3)	.12
Yes	11 (26.8)	6 (10.2)	4 (28.6)	7 (25)			
Insomnia, n (%)					—	—	.18
Yes	3 (7.3)	7 (11.9)	2 (14.3)	0 (0)			
Restraints, n (%)					—	—	.56
Yes	1 (2.4)	4 (6.8)	0 (0)	0 (0)			
Immobility, n (%)					—	17.0 (3)	<.001
Yes	19 (46.3)	13 (22)	3 (21.4)	1 (3.6)			
Falls, n (%)					—	7.0 (3)	.07
Yes	12 (29.3)	7 (11.9)	4 (28.6)	3 (10.7)			
Pressure ulcer, n (%)					—	—	.30
Yes	11 (26.8)	8 (13.6)	2 (14.3)	3 (10.7)			
Analgesic use, n (%)					—	—	.03
Yes	10 (24.4)	4 (6.8)	4 (28.6)	5 (17.9)			
Delirium-related medication (haloperidol), n (%)					—	—	<.001
Yes	6 (14.6)	0 (0)	8 (57.1)	0 (0)			
Transfusion, n (%)					—	—	.10
Yes	0 (0)	0 (0)	1 (7.1)	0 (0)			
Indwelling urinary catheterization, n (%)					—	—	.009
Yes	0 (0)	0 (0)	2 (14.3)	0 (0)			
Enteral feeding, n (%)					—	—	.31
Yes	4 (9.8)	5 (8.5)	0 (0)	0 (0)			
Diaper use, n (%)					—	42.3 (3)	<.001
Yes	31 (75.6)	44 (74.6)	9 (64.3)	2 (7.1)			
Sedative use, n (%)					—	5.4 (3)	.14
Yes	12 (29.3)	12 (20.3)	2 (14.3)	2 (7.1)			

[a]DSD: delirium superimposed on dementia.

[b]Non-DSD: dementia without delirium.

[c]MCI-D: mild cognitive impairment with delirium.

[d]Non-MCI-D: mild cognitive impairment without delirium.

[e]—: not applicable. *F* statistics are reported for ANOVA and χ^2 statistics for Pearson chi-square tests. Neither statistic applies to Fisher-Freeman-Halton exact tests. No test was performed for alcohol consumption because all participants reported no alcohol use.

[f]CCI: Charlson Comorbidity Index.

[g]K-MMSE: Korean version of the Mini-Mental State Examination.

[h]BUN: blood urea nitrogen.

Factors Associated With Delirium

Overall, female sex (adjusted OR [aOR] 4.344, 95% CI 1.794-10.515; *P*=.001), immobility (aOR 4.953, 95% CI 1.857-13.205; *P*=.001), and analgesic use (aOR 3.641, 95% CI 1.262-10.501; *P*=.02) were positively associated with delirium, whereas infection was inversely associated with delirium (aOR 0.251, 95% CI 0.101-0.623; *P*=.003). Age (*P*=.69), CCI score (*P*=.69), and K-MMSE score (*P*=.48) were not significantly associated with delirium. The maximum variance inflation factor

XSL•FO
RenderX

was 1.22. The Hosmer-Lemeshow test showed no evidence of poor fit (χ^2 =7.1; *P*=.53). In the dementia group, female sex, immobility, analgesic use, and pain were positively associated with delirium, and infection was inversely associated with delirium. In the MCI group, diaper use was associated with delirium (aOR 20.327, 95% CI 3.005-137.500; *P*=.002; Table 4).

Table 4. Factors associated with deliriuma.

Variable	Overall (N=142)		Dementia (n=100)		MCI[b] (n=42)	
	aOR[c] (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value
Age	0.988 (0.931-1.048)	.69	—[d]	—	—	—
Female sex	4.344 (1.794-10.515)	.001	4.013 (1.410-11.423)	.009	5.552 (0.767-40.187)	.09
CCI[e] score	0.952 (0.749-1.211)	.69	—	—	—	—
K-MMSE[f] score	0.980 (0.925-1.037)	.48	—	—	—	—
Infection	0.251 (0.101-0.623)	.003	0.166 (0.052-0.537)	.003	—	—
Immobility	4.953 (1.857-13.205)	.001	6.302 (1.982-20.041)	.002	—	—
Analgesic use	3.641 (1.262-10.501)	.02	5.345 (1.106-25.818)	.04	—	—
Falls	—	—	2.252 (0.605-8.387)	.23	—	—
Pain	—	—	6.073 (1.326-27.818)	.02	—	—
Diaper use	—	—	—	—	20.327 (3.005-137.500)	.002

[a]Reference categories were male sex and absence of each binary characteristic.

[b]MCI: mild cognitive impairment.

[c]aOR: adjusted odds ratio.

[d]Not applicable; the variable was not included in the corresponding regression model.

[e]CCI: Charlson Comorbidity Index.

[f]K-MMSE: Korean version of the Mini-Mental State Examination.

Discussion

This study examined delirium occurrence and associated factors among patients with dementia or MCI in LTC hospitals. Delirium was identified in 38.7% (55/142) of patients, and hypoactive delirium was the most common subtype. Female sex, immobility, and analgesic use were positively associated with delirium overall, whereas infection was inversely associated with delirium.

Delirium was identified in 38.7% of patients, compared with 48% in the Korean LTC hospital study by Moon and Park [3], which used the same assessment instrument. Their study included patients with and without dementia and assessed delirium for 3 months, whereas our study focused on patients with dementia or MCI over 4 weeks. Differences in populations and assessment periods also contribute to the wide range of prevalence estimates in residential care (1%-70.3%) [14]. Despite these differences, delirium affected more than one-third of our participants, supporting the need for regular assessment in patients receiving care for cognitive impairment.

Hypoactive delirium was the most common subtype overall (33/55, 60%) and in patients with dementia (27/41, 65.9%), whereas hyperactive delirium was most common in patients with MCI (7/14, 50%). The predominance of hypoactive delirium is consistent with the ICU findings of la Cour et al [21], where this subtype accounted for approximately 50.3% of cases. However, mixed delirium predominated in the LTC hospital study by Moon and Park [3]. Ghezzi et al [22] found poorer cognition in patients with hypoactive than hyperactive delirium, although the effect size was small. These findings highlight the relevance of cognitive status when interpreting motor patterns. Hypoactive delirium may be overlooked when patients appear quiet [9], making observation of changes in activity and attention important [23]. In patients with dementia, nurses need to compare these changes with usual behavior. In patients with MCI, both increased and reduced activity require attention, as hypoactive delirium still accounted for 42.9% (6/14) of patients with MCI who developed delirium.

Female sex, immobility, and analgesic use remained associated with delirium after adjustment for age, CCI score, and K-MMSE score. Patients with dementia may have difficulty expressing pain [24,25], and pain has been associated with delirium in this population [26]. A meta-analysis also identified an association between pain and delirium among older hospital inpatients, although most included studies involved surgical patients [27]. Our finding for analgesic use therefore needs to be interpreted alongside pain and the reasons for treatment. Mobility also warrants attention, as physical dysfunction has been linked to delirium [28], and immobility was associated with delirium in the univariable analysis of older patients undergoing cholecystectomy [29]. Together, these findings support assessing pain in ways suited to patients' communication abilities and considering mobility, pain relief, and changes in attention during daily care.

XSL•FO
RenderX

Infection was inversely associated with delirium in our study. Associations between systemic infection and delirium in residential care have also been inconsistent or nonsignificant [14]. Our infection variable included physician diagnosis, laboratory evidence, or antibiotic use, and its timing relative to delirium was unknown. The inverse association should therefore not be interpreted as a protective effect of infection.

Diaper use was associated with delirium in the MCI group, with use recorded in 64.3% (9/14) of patients with delirium and 7.1% (2/28) without delirium. The association remained after adjustment for sex, but the wide CI indicates limited precision. Diaper use may reflect a need for assistance with continence, mobility, or toileting, although these needs were not established by our data. Clinically, the finding draws attention to the circumstances of diaper use and the patient's functional status. Documenting the reasons for diaper use and changes in toileting ability alongside cognitive and behavioral changes may help nurses assess these care needs. Future studies should examine these factors before delirium develops to clarify the observed association.

Delirium was identified in one-third of patients with MCI, indicating a need for assessment, as well as in patients with dementia. Sabbagh et al [30] emphasized linking MCI recognition with further assessment and appropriate care in primary care. In acute care hospitals, gaps remain in cognitive assessment, documentation, and communication with patients and caregivers [31]. These findings are relevant to continuity of cognitive care in LTC hospitals, where nurses can document usual cognitive and functional status, share changes during handover, and consult familiar caregivers when establishing the patient's baseline. Regular delirium assessment, together with attention to mobility, pain, and toileting needs, provides a focus for nursing care. The effectiveness of preventive care based on these findings requires evaluation in LTC hospitals, where intervention evidence remains limited [12].

This study has several limitations. First, data were collected during the COVID-19 pandemic, when masks and changes in the care environment and social contact may have affected delirium assessment and occurrence. Assessment completeness was not quantified, and formal interrater reliability was not evaluated during data collection. Symptoms of dementia and delirium can also overlap, complicating assessment in patients with moderate to severe dementia despite training [32]. Second, the timing of clinical conditions and care practices relative to delirium was unknown, precluding causal interpretation. Details of sedative use and the interval and type of prior surgery were unavailable. Third, the sample size was calculated for multiple linear regression rather than the logistic models used. The small cognitive groups, particularly the MCI group, limited estimate precision. Fourth, pain assessment included medication use, and infection assessment included antibiotic use, limiting the separation of symptoms from treatment-related associations. We did not adjust for multiple comparisons. Finally, recruitment from 2 LTC hospitals limits generalizability to other settings and patients without cognitive impairment. Larger longitudinal studies are needed to confirm these associations.

Delirium was common among patients with dementia or MCI in LTC hospitals, and hypoactive delirium was the most common subtype overall. Female sex, immobility, and analgesic use were positively associated with delirium, whereas infection was inversely associated. These findings support regular delirium assessment, with attention to mobility and pain-related care needs. Further longitudinal studies are needed to clarify these associations and the findings in each cognitive group.

Acknowledgments

The authors thank the staff and participants of the long-term care hospitals for their cooperation in this study.

Funding

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea, funded by the Ministry of Education (2019R1I1A3A01060561). The funder had no role in the study design; data collection, analysis, or interpretation; manuscript preparation; or the decision to submit the manuscript for publication.

Data Availability

The data are not publicly available because they contain sensitive patient information. Deidentified data may be requested from the corresponding author, subject to consent and institutional approval.

Generative AI Use

ChatGPT (Sol model; OpenAI) was used to check journal submission requirements, assist with editing tables and figures, and check consistency between in-text citations and the reference list.

Authors' Contributions

Conceptualization: KJM
Methodology: KJM
Funding acquisition: KJM
Investigation: YHH

XSL•FO
RenderX

Data curation: YHH
Formal analysis: KJM, YHH
Writing—original draft: KJM
Writing—review & editing: YHH

Conflicts of Interest

None declared.

Multimedia Appendix 1
STROBE checklist.
[DOCX File , 27 KB - aging_v9i1e90434_app1.docx]

References

1. Jaqua EE, Nguyen VT, Chin E. Delirium in older persons: prevention, evaluation, and management. Am Fam Physician 2023 Sep;108(3):278-287 [FREE Full text] [Medline: 37725462]
2. Lim EJ, Koh WU, Kim H, Kim HJ, Shon HC, Kim JW. Regional nerve block decreases the incidence of postoperative delirium in elderly hip fracture. J Clin Med 2021 Aug 15;10(16):3586 [FREE Full text] [doi: 10.3390/jcm10163586] [Medline: 34441882]
3. Moon KJ, Park H. Outcomes of patients with delirium in long-term care facilities: a prospective cohort study. J Gerontol Nurs 2018 Sep 01;44(9):41-50. [doi: 10.3928/00989134-20180808-08] [Medline: 30148531]
4. Reynish EL, Hapca SM, De Souza N, Cvoro V, Donnan PT, Guthrie B. Epidemiology and outcomes of people with dementia, delirium, and unspecified cognitive impairment in the general hospital: prospective cohort study of 10,014 admissions. BMC Med 2017 Jul 27;15(1):140 [FREE Full text] [doi: 10.1186/s12916-017-0899-0] [Medline: 28747225]
5. Han QY, Rodrigues NG, Klainin-Yobas P, Haugan G, Wu XV. Prevalence, risk factors, and impact of delirium on hospitalized older adults with dementia: a systematic review and meta-analysis. J Am Med Dir Assoc 2022 Jan;23(1):23-32.e27. [doi: 10.1016/j.jamda.2021.09.008] [Medline: 34648761]
6. Goldberg TE, Chen C, Wang Y, Jung E, Swanson A, Ing C, et al. Association of delirium with long-term cognitive decline: a meta-analysis. JAMA Neurol 2020 Nov 01;77(11):1373-1381 [FREE Full text] [doi: 10.1001/jamaneurol.2020.2273] [Medline: 32658246]
7. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. Lancet 2014 Mar 08;383(9920):911-922 [FREE Full text] [doi: 10.1016/S0140-6736(13)60688-1] [Medline: 23992774]
8. MacLullich AM, Shenkin SD, Goodacre S, Godfrey M, Hanley J, Stíobhairt A, et al. The 4 'A's test for detecting delirium in acute medical patients: a diagnostic accuracy study. Health Technol Assess 2019 Aug;23(40):1-194 [FREE Full text] [doi: 10.3310/hta23400] [Medline: 31397263]
9. Oh ES, Fong TG, Hshieh TT, Inouye SK. Delirium in older persons: advances in diagnosis and treatment. JAMA 2017 Sep 26;318(12):1161-1174 [FREE Full text] [doi: 10.1001/jama.2017.12067] [Medline: 28973626]
10. Ormseth CH, LaHue SC, Oldham MA, Josephson SA, Whitaker E, Douglas VC. Predisposing and precipitating factors associated with delirium: a systematic review. JAMA Netw Open 2023 Jan 03;6(1):e2249950 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.49950] [Medline: 36607634]
11. Burton JK, Craig L, Yong SQ, Siddiqi N, Teale EA, Woodhouse R, et al. Non-pharmacological interventions for preventing delirium in hospitalised non-ICU patients. Cochrane Database Syst Rev 2021 Nov 26;11(11):CD013307 [FREE Full text] [doi: 10.1002/14651858.CD013307.pub3] [Medline: 34826144]
12. Woodhouse R, Burton JK, Rana N, Pang Y, Lister JE, Siddiqi N. Interventions for preventing delirium in older people in institutional long-term care. Cochrane Database Syst Rev 2019 Apr 23;4(4):CD009537 [FREE Full text] [doi: 10.1002/14651858.CD009537.pub3] [Medline: 31012953]
13. Fong TG, Inouye SK. The inter-relationship between delirium and dementia: the importance of delirium prevention. Nat Rev Neurol 2022 Oct;18(10):579-596 [FREE Full text] [doi: 10.1038/s41582-022-00698-7] [Medline: 36028563]
14. Wouters I, Yang KL, Rex S, Milisen K. Prevalence, risk factors, and preventive interventions for delirium in older adults living at home and in residential care: a scoping review. BMC Geriatr 2026 Feb 06;26(1):371 [FREE Full text] [doi: 10.1186/s12877-026-07021-9] [Medline: 41652345]
15. McGrattan AM, Pakpahan E, Siervo M, Mohan D, Reidpath DD, Prina M, et al. Risk of conversion from mild cognitive impairment to dementia in low- and middle-income countries: a systematic review and meta-analysis. Alzheimers Dement (N Y) 2022 Mar 13;8(1):e12267 [FREE Full text] [doi: 10.1002/trc2.12267] [Medline: 35310524]
16. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec 16;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6] [Medline: 28735855]
17. Korean Dementia Observatory. National Institute of Dementia. URL: https://www.nid.or.kr/info/dataroom_view.aspx?BID=243 [accessed 2026-10-24]

XSL•FO
RenderX

18. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007 Oct 16;4(10):e296 [FREE Full text] [doi: 10.1371/journal.pmed.0040296] [Medline: 17941714]
19. Kang Y, Na DL, Hahn S. A validity study on the Korean Mini-Mental State Examination (K-MMSE) in dementia patients. J Korean Neurol Assoc 1997 Apr;15(2):300-308 [FREE Full text]
20. Inouye S. The Confusion Assessment Method (CAM): training manual and coding guide. Hospital Elder Life Program. 2014. URL: https://www.psychdb.com/_media/cl/cam_training_guide.pdf [accessed 2026-10-06]
21. la Cour KN, Andersen-Ranberg NC, Weihe S, Poulsen LM, Mortensen CB, Kjer CK, et al. Distribution of delirium motor subtypes in the intensive care unit: a systematic scoping review. Crit Care 2022 Mar 03;26(1):53 [FREE Full text] [doi: 10.1186/s13054-022-03931-3] [Medline: 35241132]
22. Ghezzi ES, Greaves D, Boord MS, Davis D, Knayfati S, Astley JM, et al. How do predisposing factors differ between delirium motor subtypes? A systematic review and meta-analysis. Age Ageing 2022 Sep 02;51(9):afac200 [FREE Full text] [doi: 10.1093/ageing/afac200] [Medline: 36153750]
23. Takeya Y, Arai H. Delirium in the long-term care: a challenge for clinicians and researchers-the negative outcomes related to delirium: a scoping review. Eur Geriatr Med 2025 Dec;16(6):1961-1970. [doi: 10.1007/s41999-025-01274-0] [Medline: 41076610]
24. Boltz M, Resnick B, Kuzmik A, Mogle J, Jones JR, Arendacs R, et al. Pain incidence, treatment, and associated symptoms in hospitalized persons with dementia. Pain Manag Nurs 2021 Apr;22(2):158-163 [FREE Full text] [doi: 10.1016/j.pmn.2020.08.002] [Medline: 32921569]
25. Sampson EL, West E, Fischer T. Pain and delirium: mechanisms, assessment, and management. Eur Geriatr Med 2020 Feb;11(1):45-52. [doi: 10.1007/s41999-019-00281-2] [Medline: 32297242]
26. Xiao Q, Zhang S, Li C, Zhu Y. Risk factors for delirium superimposed on dementia in elderly patients in comprehensive ward. Am J Alzheimers Dis Other Demen 2023;38:15333175231206023 [FREE Full text] [doi: 10.1177/15333175231206023] [Medline: 37812498]
27. White N, Bazo-Alvarez JC, Koopmans M, West E, Sampson EL. Understanding the association between pain and delirium in older hospital inpatients: systematic review and meta-analysis. Age Ageing 2024 Apr 01;53(4):afae073 [FREE Full text] [doi: 10.1093/ageing/afae073] [Medline: 38610062]
28. Gual N, García-Salmones M, Brítez L, Crespo N, Udina C, Pérez LM, et al. The role of physical exercise and rehabilitation in delirium. Eur Geriatr Med 2020 Feb;11(1):83-93 [FREE Full text] [doi: 10.1007/s41999-020-00290-6] [Medline: 32297245]
29. Park YM, Seo HI, Noh BG, Kim S, Hong SB, Lee NK, et al. Postoperative delirium after cholecystectomy in older patients: a retrospective study. Ann Hepatobiliary Pancreat Surg 2023 Aug 31;27(3):301-306 [FREE Full text] [doi: 10.14701/ahbps.23-012] [Medline: 37336783]
30. Sabbagh MN, Boada M, Borson S, Chilukuri M, Dubois B, Ingram J, et al. Early detection of mild cognitive impairment (MCI) in primary care. J Prev Alzheimers Dis 2020;7(3):165-170. [doi: 10.14283/jpad.2020.21] [Medline: 32463069]
31. Xiong B, Bailey DX, Prudon P, Pascoe EM, Gray LC, Graham F, et al. Identification and information management of cognitive impairment of patients in acute care hospitals: an integrative review. Int J Nurs Sci 2023 Dec 01;11(1):120-132 [FREE Full text] [doi: 10.1016/j.ijnss.2023.11.001] [Medline: 38352291]
32. Fong TG, Hshieh TT, Tabloski PA, Metzger ED, Arias F, Heintz HL, et al. Identifying delirium in persons with moderate or severe dementia: review of challenges and an illustrative approach. Am J Geriatr Psychiatry 2022 Oct;30(10):1067-1078 [FREE Full text] [doi: 10.1016/j.jagp.2022.04.003] [Medline: 35581117]

Abbreviations

aOR: adjusted odds ratio
CCI: Charlson Comorbidity Index
DSD: delirium superimposed on dementia
ICU: intensive care unit
K-MMSE: Korean version of the Mini-Mental State Examination
LTC: long-term care
MCI: mild cognitive impairment
MCI-D: mild cognitive impairment with delirium
Non-DSD: dementia without delirium
Non-MCI-D: mild cognitive impairment without delirium
S-CAM: Short Confusion Assessment Method
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by A Mavragani; submitted 29.Dec.2025; peer-reviewed by P Rani, E Wong; comments to author 25.Jun.2026; revised version received 09.Oct.2026; accepted 12.Oct.2026; published 27.Oct.2026.

Please cite as:
Moon KJ, Hyeon YH
Delirium Occurrence and Associated Factors Among Patients With Dementia and Mild Cognitive Impairment Admitted to Long-Term Care Hospitals: Cross-Sectional Study
JMIR Aging 2026;9:e90434
URL: https://aging.jmir.org/2026/1/e90434
doi:10.2196/90434
PMID:

©Kyoung Ja Moon, Yong-Hwan Hyeon. Originally published in JMIR Aging (https://aging.jmir.org), 27.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Prediction Models for Frailty in People Living with HIV: Systematic Review

Meilian Xie[1,2*], PhD; Yeye Hu[2*], BSN; Junjia Yuan[1], BSN; Zhiyun Zhang[2], MSc

[1]Capital Center for Children's Health, Capital Medical University, Capital Institute of Pediatrics, Beijing, China

[2]Beijing Ditan Hospital, Capital Medical University, No. 8 Jingshun East Street, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Zhiyun Zhang, MSc
Beijing Ditan Hospital, Capital Medical University, No. 8 Jingshun East Street, Beijing, China

Abstract

Background: Frailty is increasingly prevalent among people living with HIV, who experience earlier onset and a higher burden of frailty than the general population. Prediction models that identify people living with HIV at risk of future frailty could enable timely, targeted intervention. However, the evidence on prediction models developed specifically for frailty in people living with HIV has not yet been systematically evaluated.

Objective: This review aims to systematically identify and critically appraise multivariable prediction models for the future occurrence of frailty in people living with HIV and to assess and summarize their methodological quality and predictive performance.

Methods: We searched PubMed, Web of Science, Scopus, Cochrane Library, ProQuest, and Chinese databases (SinoMed, CNKI, Wanfang, VIP) from inception to July 24, 2026. Studies that developed or validated any multivariable statistical or machine-learning prediction model for frailty in people living with HIV were eligible. Title and abstract screening, full-text evaluation, data extraction (Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies [CHARMS]), and risk-of-bias and applicability assessment (Prediction Model Risk of Bias Assessment Tool [PROBAST]) were conducted independently in duplicate.

Results: The search identified 2523 records (initial search: 2205; updated search: 318); 137 full-text articles were assessed for eligibility, and only 1 study met all inclusion criteria. Frail RISC-HIV (Frailty Risk in the Short-Term for Care of People with HIV) was developed in the Centers for AIDS Research Network of Integrated Clinical Systems (CNICS) cohort (n=4680; training n=3170; validation n=1510) using Bayesian model averaging and LASSO (least absolute shrinkage and selection operator) regression combined with Cox proportional-hazards models to predict incident frailty within 2 years. The model comprised 8 predictors (baseline frailty score, depressive symptomatology [Patient Health Questionnaire-7], prescribed antidepressants, current illicit opioid use, current marijuana use, current antiretroviral therapy, age, and sex) and achieved acceptable discrimination in the validation set (Harrell C=0.76, 95% CI 0.73 - 0.79), with a sensitivity of 80% and specificity of 61% at a 5% risk threshold; calibration under-predicted the observed risk (predicted 7.9% vs observed 14.7% in the validation set). PROBAST assessment rated the study at high overall risk of bias, driven primarily by the analysis domain. To our knowledge, no external validation of the model was identified, and TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) reporting was incomplete. Quantitative synthesis was not feasible.

Conclusions: The evidence base for frailty prediction in people living with HIV currently comprises a single, internally validated model that has not yet been externally validated. This critical evidence gap underscores an urgent need for further model development and rigorous external validation in diverse settings, with standardized frailty definitions and reporting in accordance with TRIPOD.

(JMIR Aging 2026;9:e110499) doi:10.2196/110499

KEYWORDS

HIV; frailty; prediction model; systematic review; risk score

Introduction

With the widespread availability of antiretroviral therapy (ART), HIV infection has evolved from a fatal illness into a manageable chronic condition, and the life expectancy of people living with HIV has increased substantially [1,2]. The Joint United Nations Program on HIV and AIDS (UNAIDS) estimated that 40.8 million people living with HIV globally in 2024 [3]. As this population ages, the proportion of people living with HIV aged

XSL•FO
RenderX

50 years and older is projected to increase from 28% in 2010 to 73% by 2030 [4]. Clinicians and public health practitioners are therefore increasingly confronted not only with viral load management but also with the prevention and treatment of age-related comorbidities and geriatric syndromes among people living with HIV [5].

Frailty—a state of heightened vulnerability to stressors resulting from reduced physiological reserves—is among the most clinically relevant of these geriatric syndromes [6]. The 2 dominant conceptual frameworks are the frailty phenotype (FP), which defines frailty as the presence of 3 or more of 5 criteria (unintentional weight loss, exhaustion, weakness, slowness, and low physical activity) [7], and the frailty index (FI), which quantifies frailty as the accumulation of health deficits [8]. As a prognostic marker, frailty predicts a range of adverse outcomes, including falls, hospitalization, disability, and death [9,10]. Among people living with HIV, the burden is disproportionately high: a systematic review and meta-analysis reported pooled prevalences of frailty and prefrailty of 10.9% and 47.2%, respectively [11], and frailty has been shown to occur at younger ages and with greater severity in people living with HIV than in demographically matched HIV-negative peers [12,13]. These observations reflect persistent immune activation, chronic inflammation, accelerated immunosenescence, and ART-related metabolic disturbances [14].

A clinical prediction model is a statistical tool that combines multiple predictors to estimate an individual's risk of having a condition (diagnostic model) or of developing it in the future (prognostic model) [15]. Prognostic models for frailty could support the proactive identification of people living with HIV at highest risk, enabling early interventions to prevent or delay the progression of frailty. Numerous frailty prediction models exist in general geriatric medicine; however, applying standardized FP- or FI-based instruments directly to people living with HIV poses significant challenges because these instruments typically do not incorporate HIV-specific drivers of risk, such as nadir CD4 cell count, viral load, chronic immune activation, ART regimens, and common coinfections [16]. The comorbidity spectrum, patterns of functional decline, and psychosocial stressors among people living with HIV also differ from those of the general older population [16], collectively suggesting that general-population models may lack direct transferability to people living with HIV.

Although a growing body of research has developed frailty assessment or screening tools for HIV care [17,18], a comprehensive, systematic evaluation of the methodological rigor, predictive performance, and clinical applicability of models designed to predict future frailty in people living with HIV has been notably lacking. Whether generic frailty prediction models retain clinically useful performance in people living with HIV, or whether HIV-specific models are required, remains uncertain and warrants direct comparison. To address this gap, we published a protocol for a systematic review of prognostic models for frailty in people living with HIV and registered the review with PROSPERO (CRD420261332573) [19]. The objectives of this systematic review were to (1) identify all multivariable prediction models developed or validated to predict future frailty in people living with HIV, (2) summarize and compare their predictive performance, and (3) critically appraise their methodological quality, risk of bias, and applicability using recommended tools.

Methods

Protocol and Registration

This review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [20,21]. The protocol was developed following the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses–Protocols) checklist [22], registered with PROSPERO (CRD420261332573) in June 2025, and published in *JMIR Research Protocols* [19]. The review question was framed according to the CHARMS (Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies) guidance for prediction models [23]. After registration, two deviations from the protocol occurred: (1) the preliminary screening identified 5 potentially eligible studies, but strict reapplication of the prespecified eligibility criteria during full-text evaluation resulted in the exclusion of 4 of them (see the "Study Selection" section), and (2) meta-analysis was not performed because only 1 study was included; a narrative synthesis was conducted instead.

Eligibility Criteria

Eligibility criteria were defined using the PICOS (population, intervention, comparison, outcome, study design) framework and the CHARMS items to frame the review question [23] (Table 1). Studies were included if they (1) developed or validated a multivariable prediction model—whether a statistical model, machine-learning algorithm, or clinical risk score—with or without external validation; (2) enrolled adults (≥18 years) diagnosed with HIV/AIDS according to recognized diagnostic guidelines (eg, the Chinese guidelines for the diagnosis and treatment of HIV infection/ AIDS 2024 edition, WHO consolidated guidelines, US CDC laboratory testing guidance, European AIDS Clinical Society guidelines version 12, or UK testing guidelines) [24]; and (3) reported a primary outcome reflecting the future occurrence of frailty, with frailty defined using established multidimensional instruments (eg, FP, FI, Clinical Frailty Scale, Edmonton Frail Scale, FRAIL scale, or other validated frailty measures), regardless of the terminology used (eg, prediction model, risk score, prediction index, clinical prediction rule). Studies were excluded if they (1) evaluated the predictive value of more than 1 variable without reporting subgroups or evaluation outcomes; (2) were cross-sectional studies, because such designs do not assess future risk; (3) were nonprimary research (literature reviews, systematic reviews, meta-analyses, protocols, theses, editorials, comments, or letters); or (4) were unavailable in full text. No restrictions were applied regarding year or language. When multiple studies reported results from the same cohort for a given outcome, the study with the largest patient population was selected; if sample sizes were equal, the most recent publication was used.

XSL•FO
RenderX

Table . Eligibility criteria for study inclusion based on the PICOS (population, intervention, comparison, outcome, study design) framework and CHARMS (Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies) key items.

PICOS element	Criteria applied in this review
Population	Adults (≥18 years) diagnosed with HIV/AIDS according to recognized diagnostic guidelines
Index prediction model	Multivariable statistical models, machine-learning algorithms, or clinical risk scores developed or validated to predict the future occurrence of frailty, with or without external validation
Comparator	Not applicable (single-index-model review); comparator models described when reported
Outcome to be predicted	Future occurrence of frailty, defined by established multidimensional instruments (frailty phenotype, frailty index, Clinical Frailty Scale, Edmonton Frail Scale, FRAIL scale, or other validated measures)
Timing	Any prediction horizon for the future occurrence of frailty; moment of model use in clinical care
Setting	Any clinical or research setting in which adults with HIV receive care
Study design	Model development or validation studies. Excluded: cross-sectional studies, nonprimary research (reviews, protocols, theses, editorials, letters), studies unavailable in full text

Information Sources and Search Strategy

We searched PubMed, Web of Science, Scopus, Cochrane Library, ProQuest, and 4 Chinese databases (SinoMed, CNKI, Wanfang, and VIP) from database inception to July 20, 2025. The search was updated on July 24, 2026. The search strategy combined three thematic blocks: (1) population terms related to HIV/AIDS; (2) index-model terms related to prediction models, risk scores, and nomograms; and (3) outcome terms related to frailty (Table 2). A model-specific search filter was adapted where appropriate, and reference lists of relevant articles were screened for additional studies. All search strategies are provided in full in Multimedia Appendix 1.

Table . Summary of the search strategy used in each database.

Database	Key search concepts	Records identified (initial search, July 20, 2025), n	Records identified (updated search, July 24, 2026), n	Total, n
PubMed	#1 HIV/AIDS terms (MeSH + title/abstract) AND #2 prediction-model terms AND #3 frailty terms	282	44	326
Web of Science	TI/AB HIV/AIDS AND TI/AB prediction-model terms AND TI/AB frailty terms	1643	239	1882
Scopus	TITLE-ABS-KEY HIV/AIDS AND prediction AND frailty	11	1	12
Cochrane Library	MeSH and keyword combinations of HIV/AIDS, prediction-model, and frailty terms	112	11	123
ProQuest (gray literature)	summary (HIV AND frailty AND prediction terms)	157	23	180
SinoMed/CNKI/Wanfang/VIP	Title/keyword combinations of HIV, frailty, and prediction-model terms (Chinese)	0	0	0
Total	—[a]	2205	318	2523

[a]Not applicable

Study Selection

Records identified by the search strategy were imported into EndNote, and duplicates were removed. Two reviewers (YH and MX) independently screened titles and abstracts, followed by full-text evaluation. Disagreements were resolved by discussion or, if necessary, adjudicated by a third reviewer (JY). When published reports contained insufficient information, the

XSL•FO
RenderX

corresponding authors were contacted; if the information could not be obtained, the study was recorded as having missing data rather than automatically excluded. The study selection process is summarized in the PRISMA 2020 flow diagram in the Results section.

Data Extraction

Data were extracted independently by 2 reviewers (YH and MX) using a standardized, piloted extraction form based on the CHARMS checklist [23]. Extracted items included: source of data; participant eligibility, recruitment, and description; details of treatment received; study dates; outcome definition and measurement; candidate predictors and their handling; sample size and events per variable; missing data handling; model development methods (modeling technique, predictor-selection method, shrinkage); model performance (discrimination, calibration, classification); model evaluation (internal/external validation, model updating); and interpretation and discussion of results. Reporting completeness was additionally assessed against the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement [25]. Discrepancies between reviewers were resolved by discussion; if consensus could not be reached, a third reviewer (JY) adjudicated.

Risk of Bias and Applicability Assessment

The methodological quality of the included study was assessed using the Prediction Model Risk of Bias Assessment Tool (PROBAST) [26], which comprises 4 domains (participants, predictors, outcome, and analysis) and 20 signaling questions. Each domain was judged as being at low, high, or unclear risk of bias, and its applicability to the review question was rated as low, high, or unclear concern. Two reviewers (YH and MX) independently rated the study, with a target interrater reliability (Cohen kappa) of ≥0.8; disagreements were resolved by consensus or adjudication by a third reviewer (JY).

Data Synthesis

Consistent with the registered protocol, a meta-analysis of discrimination (C statistics) and calibration was planned a priori using random-effects models [27]; however, because only 1 study met the inclusion criteria, quantitative synthesis was not feasible. Instead, a narrative synthesis was conducted following the Synthesis Without Meta-Analysis (SWiM) reporting guidance [28], summarizing the characteristics, predictive performance, and methodological quality of the included study in tabular form. Between-study heterogeneity and subgroup or meta-regression analyses were not applicable.

Ethical Considerations

Ethical approval was not required because this study is a systematic review of published literature.

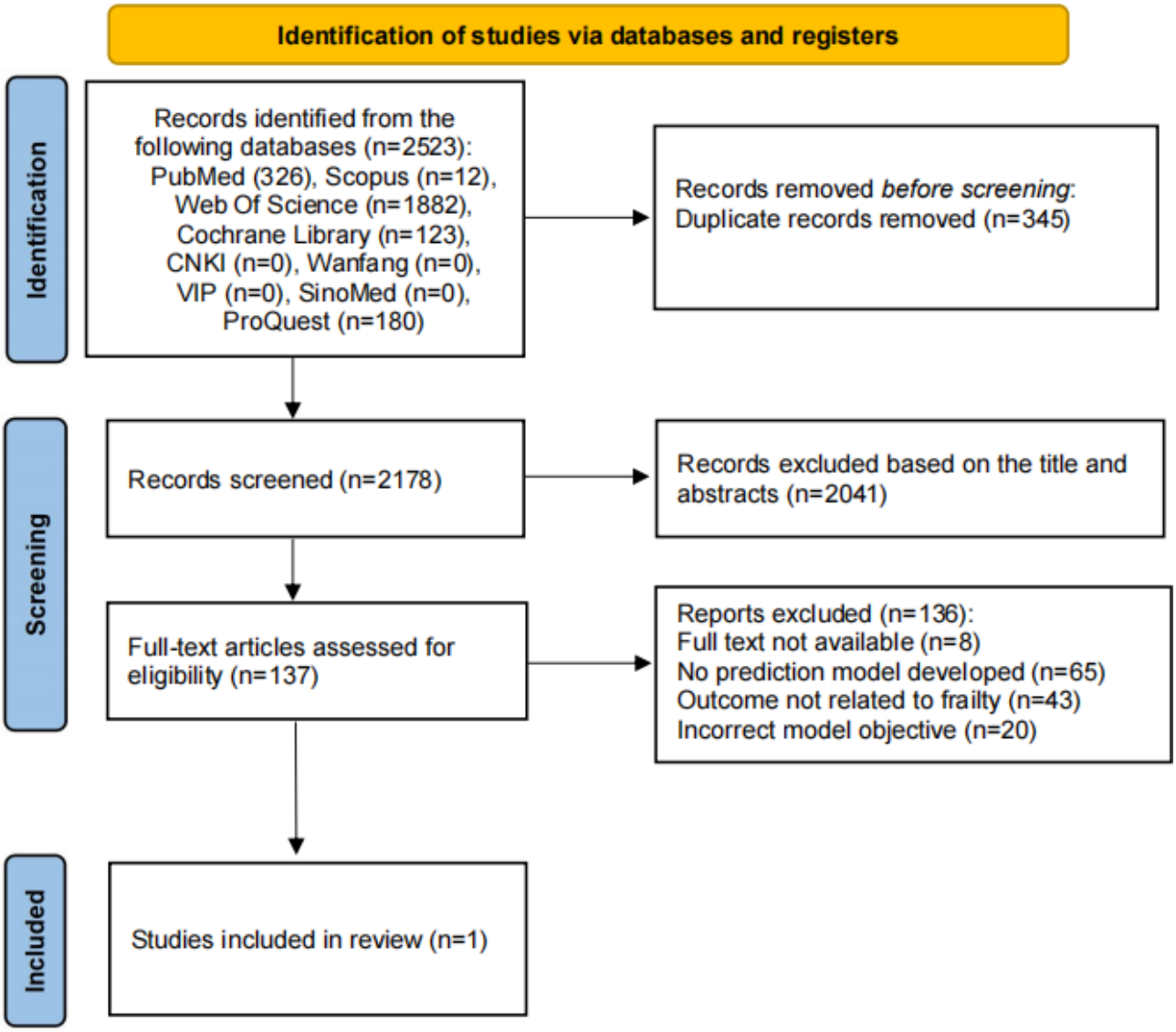
Results

Study Selection

The search identified 2523 records in total (initial search: 2205; updated search: 318). The initial search identified 2205 records (PubMed: 282; Web of Science: 1643; Scopus: 11; Cochrane Library: 112; ProQuest: 157; and none from the 4 Chinese databases). The updated search on July 24, 2026, identified 318 additional records (PubMed: 44; Web of Science: 239; Scopus: 1; Cochrane Library: 11; ProQuest: 23; and none from the 4 Chinese databases). After removal of 345 duplicates, 2178 records were screened by title and abstract, of which 2041 were excluded. The remaining 137 full-text articles were assessed for eligibility, and 136 were excluded for the following reasons: full text unavailable (n=8); no prediction model developed (n=65); outcome not related to frailty (n=43); and incorrect model objective, that is, studies that assessed frailty-related factors or cross-sectional correlates of current frailty status without developing or validating a multivariable model for the future occurrence of frailty (n=20). Ultimately, 1 study met all inclusion criteria (Figure 1). Following the published protocol [19], 5 studies were initially identified. However, full-text evaluation against predefined criteria—specifically the exclusion of cross-sectional designs and studies not predicting future frailty—led to the exclusion of 4 studies and left only 1 study that met all inclusion criteria. Excluded articles with reasons are detailed in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram of the study selection process.



Characteristics of the Included Study

The single eligible study was the development and internal validation of Frail RISC-HIV (Frailty Risk in the Short-Term for Care of People with HIV) by Ruderman et al [18], published in *AIDS* in 2023. The study was conducted within the Centers for AIDS Research Network of Integrated Clinical Systems (CNICS), a multicenter prospective clinical cohort of people living with HIV receiving care at 7 US academic HIV clinics (study period 2011 - 2021) [29]. The analytic cohort comprised 4680 nonfrail people living with HIV (training set: n=3170; validation set: n=1510). The validation set consisted of people living with HIV from the entire site, which was withheld during development to ensure independence of subpopulations (a split-site internal validation design). Detailed characteristics extracted according to the CHARMS domains are presented in Table 3.

XSL•FO
RenderX

Table . Characteristics of the included study (Frail RISC-HIV)[a], extracted according to the CHARMS[b] checklist.

CHARMS domain	Extracted details
Source of data	CNICS cohort (Centers for AIDS Research Network of Integrated Clinical Systems); multicenter prospective clinical cohort; 7 US academic HIV clinics; 2011 - 2021
Participants	4680 nonfrail people living with HIV (training: n=3170; validation: n=1510); aged ≥18 years; ≥2 patient-reported outcome (PRO) assessments; complete baseline candidate-predictor data; frailty measures at all follow-up assessments
Outcome to be predicted	Incident frailty within 2 years, defined as ≥3 of 4 components of a modified self-reported Fried phenotype (fatigue, unintentional weight loss, inactivity, poor mobility); correlated with the original Fried phenotype (ρ=0.81)
Candidate predictors	~40 variables: demographics (age, sex, race/ethnicity); HIV-related (nadir and current CD4 cell count, viral load, current ART[c] use, ever-prescribed DDI/d4T)[d]; comorbidities (HCV/HBV[e] coinfection, diabetes, dyslipidemia); anthropometric/metabolic (BMI, blood pressure, lipohypertrophy/lipoatrophy, eGFR[f], fibrosis-4); depression (PHQ-7[g] score, antidepressant prescription); substance use (marijuana, cocaine, methamphetamine, illicit opioids, tobacco, alcohol)
Sample size	Training: 3170 participants, 220 events (7%); validation: 1510 participants, 186 events (12%); events per variable adequate for the final model
Missing data	Complete-case analysis; extent and handling of missing data not fully described
Model development	Bayesian Model Averaging (BMA; >45% posterior inclusion probability) and LASSO[h] regression (5-fold cross-validation) for variable selection; Cox proportional-hazards models for effect estimation; age and sex forced into the model; no shrinkage of regression coefficients described
Final model (predictors and HRs)[i]	Baseline frailty score HR 2.9 (95% CI 2.4-3.6); current illicit opioid use HR 2.3 (1.3-4.2); female sex HR 1.5 (1.1-2.1); current marijuana use HR 1.4 (1.1-1.9); prescribed antidepressants HR 1.4 (1.1-2.0); age per decade HR 1.1 (0.96-1.2); depressive symptomatology (PHQ-7) HR 1.1 (1.0-1.1); not prescribed ART HR 0.6 (0.3-0.97)
Model performance	Discrimination: Harrell C 0.76 (95% CI 0.73-0.79) in the validation set. Classification: at 5% risk cutoff, sensitivity 80%, specificity 61%; at 10% cutoff, sensitivity 58%, specificity 81%. Calibration: under-prediction in the validation set (predicted 7.9% vs observed 14.7%). Lift plots: 5-fold higher risk in the top 3 deciles
Model evaluation	Internal validation only: split-site design (1 site withheld as the validation set); no external validation; no model updating or recalibration
Results and presentation	Risk score with ROC[j] curves and lift plots; recommended for administration during routine HIV care visits
Interpretation	Authors positioned the score as a simple tool to identify people living with HIV at risk of transitioning to frailty in the short term, supporting early targeted intervention

[a]Frail RISC-HIV: Frailty Risk in the Short-Term for Care of People with HIV.

[b]CHARMS: Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies

[c]ART: antiretroviral therapy

[d]DDI/d4T: didanosine and stavudine.

[e]HCV/HBV: hepatitis C and hepatitis B.

[f]eGFR: estimated glomerular filtration rate.

[g]PHQ-7: Patient Health Questionnaire-7.

[h]LASSO: least absolute shrinkage and selection operator.

[i]HR: hazard ratio.

[j]ROC: receiver operating characteristic.

XSL•FO
RenderX

Model Development and Predictors

Incident frailty within 2 years was defined using a modified, self-reported Fried phenotype (fatigue, unintentional weight loss, inactivity, and poor mobility; presence of ≥3 components), which has been shown to correlate highly with the original Fried phenotype (ρ=0.81) [30]. Approximately 40 candidate predictors spanning demographic, HIV-related clinical, comorbidity, depression, and substance-use variables were considered. Bayesian Model Averaging (inclusion threshold >45% posterior probability) and LASSO (least absolute shrinkage and selection operator) regression with 5-fold cross-validation were used for predictor selection, and the final score was estimated using Cox proportional hazards models; age and sex were forced into the model. The final 8-predictor model comprised baseline frailty score, Patient Health Questionnaire-7 depressive symptomatology, prescribed antidepressants, current illicit opioid use, current marijuana use, current use of ART, age, and sex (Table 3). The strongest predictor was the baseline frailty score (hazard ratio (HR) 2.9 per additional component, 95% CI 2.4 - 3.6), followed by current illicit opioid use (HR 2.3, 95% CI 1.3 - 4.2).

Model Performance

In the training set, 7% (220/3170) of people living with HIV became frail within 2 years, while in the validation set, 12% (186/1510) did. The model showed acceptable discrimination in the validation set (Harrell C=0.76; 95% CI 0.73 - 0.79). At a 5% predicted-risk cutoff, sensitivity was 80% and specificity was 61%; at 10%, they were 58% and 81%, respectively. Calibration was suboptimal; overall risk was under-predicted in validation (predicted 7.9% vs observed 14.7%), though training-set predictions aligned with observed events (8.6%). Lift plots revealed that most frailty risk was concentrated in the top 3 risk-score deciles (≈5-fold higher than the bottom 7 deciles). The model was formulated as a simple risk score with receiver operating characteristic curves and lift plots for use in routine HIV care. We identified no evidence of external validation in an independent population.

Risk of Bias and Applicability (PROBAST)

The PROBAST assessment is summarized in Table 4. Domains 1 - 3 (participants, predictors, and outcome) were judged at low risk of bias with low concern regarding applicability; the model was developed in a large, multicenter prospective cohort of people living with HIV using standardized, prospectively measured predictors and a validated outcome definition, consistent with the target population and outcome of this review. However, the baseline frailty score—the strongest predictor—was measured with the same modified Fried instrument used to define the outcome, which, although assessed at baseline and not part of the outcome event, creates partial overlap between a predictor and the outcome construct. Domain 4 (analysis) was judged at high risk of bias for the following reasons: handling of missing data was not described (a complete-case approach appears to have been used); candidate predictors were selected using data-driven methods without a fully pre-specified analysis plan; calibration was reported descriptively (predicted vs observed risk) without standard calibration metrics (eg, calibration slope or intercept); competing risks (eg, death) were not addressed; and model evaluation was limited to split-site internal validation, with no external validation. Overall, the study was judged to be at a high risk of bias, with low concern regarding applicability to the review question.

XSL•FO
RenderX

Table . PROBAST[a] risk-of-bias and applicability assessment of the included study (Frail RISC-HIV)[b].

Domain	Risk of bias	Rationale (key signaling questions)
Participants	Low	Large multicenter prospective clinical cohort of people living with HIV; inclusion/exclusion criteria appropriate for the target population; data source suitable for development of a frailty prediction model. Applicability: low concern.
Predictors	Low	Predictors defined and assessed in a standardized manner across sites; predictor data available at the time of intended model use; predictors measured before outcome ascertainment. Applicability: low concern.
Outcome	Low	Outcome (incident frailty within 2 years) defined with a validated instrument applied uniformly; time horizon explicit. Note: the baseline frailty score, a predictor, was measured with the same modified Fried instrument as the outcome, creating partial overlap between predictor and outcome constructs. Applicability: low concern.
Analysis	High	Handling of missing data not described (apparent complete-case analysis); data-driven variable selection without a fully prespecified analysis plan; calibration reported descriptively without standard calibration metrics; competing risks not addressed; evaluation limited to split-site internal validation with no external validation.
Overall	High	Overall high risk of bias driven by the analysis domain; applicability concerns were overall low.

[a]PROBAST: Prediction Model Risk of Bias Assessment Tool.

[b]Frail RISC-HIV: Frailty Risk in the Short-Term for Care of People with HIV.

Reporting Quality (TRIPOD)

Reporting completeness was partial when assessed against the TRIPOD statement. The study clearly reported participant selection, outcome definition, predictor definition, sample size, statistical analysis methods, model development and specification, discrimination with confidence intervals, and interpretation of results. Reporting was incomplete or absent for data-handling procedures, standard calibration metrics, model-updating procedures, study protocol or registration, and the electronic availability of the model.

Synthesis

Quantitative synthesis was not feasible because only 1 study met the inclusion criteria; a pooled estimate of predictive performance could therefore not be generated. In summary, the current body of evidence consists of a single internally validated model (Frail RISC-HIV) with acceptable but unreplicated discrimination, suboptimal calibration, high risk of bias in the analysis domain, and no external validation.

Discussion

Principal Findings

This systematic review, conducted according to a published, PROSPERO-registered protocol, found that the evidence base for predicting future frailty in people living with HIV comprises a single model: the Frail RISC-HIV, developed in the US CNICS cohort [18]. To our knowledge, no other study was identified that developed or validated a multivariable prediction model for incident frailty in people living with HIV, and we found no evidence of external validation for the existing model. This finding is itself the principal result of our review: despite the substantial and growing burden of frailty in the aging HIV population [5,11] and repeated calls to move from characterizing frailty toward prediction and intervention [31], the field remains in its infancy with respect to prognostic modeling.

The single available model demonstrated acceptable internal discrimination (C=0.76) but was rated at high risk of bias overall, driven primarily by methodological limitations in the analysis domain, and its calibration under-predicted the observed frailty risk in the validation set. Moreover, the model has not yet been externally validated, which limits any recommendation for its use in routine clinical practice. Its predictors are predominantly psychosocial and behavioral characteristics such as depression, antidepressant prescription, and substance use, which are clinically sensible and actionable, yet largely do not include the HIV-specific biomarkers (eg, nadir CD4 cell count, viral load, inflammatory markers) that many clinicians associate with frailty risk in people living with HIV [14,16]. Whether the omission of HIV-specific factors reflects true nonpredictiveness in the contemporary treatment era or reflects limitations of the candidate-predictor set remains unknown and warrants investigation.

XSL•FO
RenderX

Comparison With the Wider Literature

Our finding contrasts sharply with the general geriatric literature, in which dozens of frailty prediction and screening models exist and have been the subject of multiple systematic reviews. These generic models, however, have rarely been evaluated in people living with HIV. It does not necessarily follow that they are ineffective in this population; they may still capture shared aging-related mechanisms, such as multisystem physiological decline, comorbidity burden, and functional limitations, and could therefore retain some discriminative ability. However, their performance is likely to be suboptimal, particularly with respect to calibration, because they omit HIV-specific factors such as nadir CD4 cell count, viral load, chronic immune activation, and ART-related metabolic disturbances. To date, generic frailty prediction models have rarely been externally validated in people living with HIV, and we found no evidence of head-to-head comparisons with HIV-specific models. In addition, HIV-specific frailty prediction models themselves remain scarce, which is especially striking given that people living with HIV experience frailty earlier and more frequently than their HIV-negative peers [5,12]. Future studies should therefore both externally validate existing generic models in people living with HIV and develop HIV-specific models, comparing their performance before concluding that HIV-specific tools are universally necessary. The studies excluded from this review illustrate the nature of the gap: the large majority of retrieved articles characterized frailty or examined factors associated with current frailty status, including cross-sectional analyses of frailty correlates [32,33], frailty instrument construct validation [34], and machine-learning classification of prevalent frailty [35,36], rather than developing or validating models that predict future frailty. To our knowledge, only one other frailty-related tool for older people living with HIV, the FUNCFRAIL score, has been developed, but it was based on cross-sectional data and discriminates prevalent frailty rather than predicting its future occurrence [18]. Within the wider HIV literature, well-established prediction instruments such as the Veterans Aging Cohort Study Index predict mortality and other outcomes rather than frailty itself [37], and predictors of transitions in frailty severity or incidence have been reported in observational analyses [38,39], but neither constitutes a frailty prediction model of the type examined here.

Implications for Research

The critical evidence gap identified by this review has direct implications for future research. First, there is an urgent need for external validation and recalibration of Frail RISC-HIV in independent populations, particularly outside the United States and in low- and middle-income settings, where the HIV burden is greatest and where comorbidity and treatment contexts differ substantially. Such validation should include cohorts of people living with HIV who are confirmed to be nonfrail at baseline (0‐2 components of the modified Fried phenotype, ie, not frail or prefrail), because Frail RISC-HIV was developed with 4680 people living with HIV who were not frail at baseline. The baseline frailty score was a continuous count of 0‐2 components rather than a diagnosis of frailty (≥3 components); it therefore captured subclinical vulnerability or prefrailty rather than prevalent frailty. Although this score was the strongest predictor, its predictive performance should be confirmed in independently recruited nonfrail populations. Second, new prediction models should be developed that (1) incorporate HIV-specific predictors (eg, nadir and current CD4 cell count, viral load, markers of immune activation and inflammation) alongside established demographic, psychosocial, and behavioral factors; (2) use standardized and consensus, based frailty definitions; (3) are based on adequate sample sizes with sufficient events per variable; and (4) follow prespecified analysis plans with appropriate handling of missing data and competing risks. Third, all future model development and validation studies should adhere to the TRIPOD reporting statement [25], register their analysis plans, and make models electronically available. Finally, investigators should consider whether dynamic data from electronic health records can support the development of flexible, time-updated models that align with routine HIV care workflows.

Implications for Clinical Practice

From a clinical standpoint, our findings counsel caution. While Frail RISC-HIV is simple, brief, and potentially feasible for administration during HIV clinic visits, its use for risk stratification in routine care is premature in the absence of external validation and given its calibration under-prediction. Until such evidence is available, clinicians caring for aging people living with HIV should continue to rely on direct frailty assessment and on addressing the modifiable risk factors consistently associated with frailty, including depression, substance use, and polypharmacy, which were also identified as predictors in the sole available model.

Strengths and Limitations

This review has several strengths. It was conducted according to a peer-reviewed, published protocol and registered prospectively with PROSPERO. The search was comprehensive, covering 10 databases, including gray literature and 4 Chinese databases, with no language or date restrictions, and was updated through July 2026. Screening, data extraction, and quality appraisal were performed independently in duplicate, and methodological evaluation followed the internationally recommended CHARMS and PROBAST frameworks.

Several limitations should be acknowledged. First, the inclusion of a single study precluded quantitative synthesis and limited the generalizability of our conclusions; this scarcity of evidence is, however, itself the key finding of the review. Second, the review did not require prediction models to include HIV-specific health outcomes, such as viral load suppression or opportunistic infections, or HIV-specific biomarkers, such as nadir CD4 cell count and inflammatory markers. This inclusion criterion was intentionally broad to identify all available frailty prediction models for people living with HIV; however, it means that the only included model, Frail RISC-HIV, is predominantly based on psychosocial and behavioral predictors and lacks HIV-specific biomarkers. Had HIV-specific outcomes been required as an inclusion criterion, no eligible study would have been identified. This limitation reflects the current evidence base rather than the review question and highlights the need for future models that incorporate HIV-specific predictors and for comparative studies of frailty measurement between the general

XSL•FO
RenderX

population and people living with HIV. Third, despite the comprehensive search, it remains possible that eligible models were published in nonindexed journals or non-English sources and were missed, although no language restrictions were applied. Fourth, the validity of our appraisal is constrained by the completeness of reporting in the primary study; for example, several PROBAST signaling items could not be fully judged, and the corresponding authors were not contacted during this review. Finally, because only published information was assessed, unpublished or in-development models are not captured.

Conclusions

This systematic review identified a single prediction model for the future occurrence of frailty in people living with HIV, Frail RISC-HIV, which demonstrated acceptable internal discrimination but was rated at high risk of bias, showed suboptimal calibration, and has not been externally validated. The paucity of prognostic models for frailty in this vulnerable and rapidly aging population represents a critical evidence gap. Coordinated research efforts are urgently needed to develop robust models that incorporate HIV-specific factors, to externally validate and recalibrate the existing model in diverse settings, and to ensure that future studies are reported in accordance with TRIPOD guidelines. Until such evidence accumulates, model-based risk stratification for frailty in HIV care should be regarded as an emerging, rather than established, clinical tool.

Funding

This work was supported by the Special Fund for the Development of High-level Public Health Technical Talents (Discipline Leader-03-01) and the Chinese Nursing Association (ZHKY202422). The funders had no role in the design, conduct, or reporting of this review.

Data Availability

All data used in this systematic review were derived from published articles and the supplementary materials. The complete search strategies are provided in Multimedia Appendix 1.

Authors' Contributions

Conceptualization: MX, YH, ZZ, and JY
Data curation: YH, MX, ZZ, and JY
Formal analysis: YH and MX
Investigation: YH and MX
Methodology: YH, MX, and ZZ
Project administration: MX and ZZ
Supervision: MX and ZZ
Writing – original draft: MX, YH, JY, and ZZ
Writing – review and editing: MX and ZZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Complete search strategies for all databases.
[DOCX File, 26 KB - aging_v9i1e110499_app1.docx]

Multimedia Appendix 2
Full-text articles excluded after detailed evaluation, with reasons.
[DOCX File, 15 KB - aging_v9i1e110499_app2.docx]

Checklist 1
PRISMA checklist.
[DOCX File, 32 KB - aging_v9i1e110499_app3.docx]

References

1. Bekker LG, Beyrer C, Mgodi N, et al. HIV infection. Nat Rev Dis Primers 2023 Aug 17;9(1):42. [doi: 10.1038/s41572-023-00452-3] [Medline: 37591865]
2. Jones HT, Levett T, Barber TJ. Frailty in people living with HIV: an update. Curr Opin Infect Dis 2022 Feb 1;35(1):21-30. [doi: 10.1097/QCO.0000000000000798] [Medline: 34799510]

XSL•FO
RenderX

3. UNAIDS. 2024 Global AIDS Update. URL: https://crossroads.unaids.org [accessed 2024-10-27]
4. Kiplagat J, Tran DN, Barber T, et al. How health systems can adapt to a population ageing with HIV and comorbid disease. Lancet HIV 2022 Apr;9(4):e281-e292. [doi: 10.1016/S2352-3018(22)00009-1] [Medline: 35218734]
5. Falutz J, Brañas F, Erlandson KM. Frailty: the current challenge for aging people with HIV. Curr Opin HIV AIDS 2021 May 1;16(3):133-140. [doi: 10.1097/COH.0000000000000677] [Medline: 33833208]
6. Proietti M, Cesari M. Frailty: what is it? Adv Exp Med Biol 2020;1216:1-7. [doi: 10.1007/978-3-030-33330-0_1] [Medline: 31894541]
7. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
8. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. ScientificWorldJournal 2001 Aug 8;1:323-336. [doi: 10.1100/tsw.2001.58] [Medline: 12806071]
9. Kojima G. Frailty as a predictor of future falls among community-dwelling older people: a systematic review and meta-analysis. J Am Med Dir Assoc 2015 Dec;16(12):1027-1033. [doi: 10.1016/j.jamda.2015.06.018] [Medline: 26255098]
10. Hu K, Zhou Q, Jiang Y, et al. Association between frailty and mortality, falls, and hospitalization among patients with hypertension: a systematic review and meta-analysis. Biomed Res Int 2021;2021:2690296. [doi: 10.1155/2021/2690296] [Medline: 33575325]
11. Yamada Y, Kobayashi T, Condo A, et al. Prevalence of frailty and prefrailty in people with human immunodeficiency virus aged 50 or older: a systematic review and meta-analysis. Open Forum Infect Dis 2022 May;9(5):ofac129. [doi: 10.1093/ofid/ofac129] [Medline: 35415198]
12. Retornaz F, Petit N, Darque A, et al. Frailty phenotype in older people living with HIV: concepts, prevention and issues. Geriatr Psychol Neuropsychiatr Vieil 2019 Jun 1;17(2):123-128. [doi: 10.1684/pnv.2019.0795] [Medline: 31010801]
13. McMillan JM, Krentz HB, Gill MJ, Hogan DB. An emerging concern-high rates of frailty among middle-aged and older individuals living with HIV. Can Geriatr J 2019 Dec;22(4):190-198. [doi: 10.5770/cgj.22.387] [Medline: 31885759]
14. Leng SX, Margolick JB. Understanding frailty, aging, and inflammation in HIV infection. Curr HIV/AIDS Rep 2015 Mar;12(1):25-32. [doi: 10.1007/s11904-014-0247-3] [Medline: 25656346]
15. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ 2009 Feb 23;338(feb23 1):b375. [doi: 10.1136/bmj.b375] [Medline: 19237405]
16. Greene M, Justice AC, Covinsky KE. Assessment of geriatric syndromes and physical function in people living with HIV. Virulence 2017 Jul 4;8(5):586-598. [doi: 10.1080/21505594.2016.1245269] [Medline: 27715455]
17. Crane HM, Ruderman SA, Nance RM, et al. Approaches to optimally target frailty screening among people with HIV in clinical care: findings from the Centers for AIDS Research Network of Integrated Clinical Systems (CNICS). J Acquir Immune Defic Syndr 2025 Apr 17. [doi: 10.1097/QAI.0000000000003688] [Medline: 40243191]
18. Ruderman SA, Nance RM, Drumright LN, et al. Development of Frail RISC-HIV: a risk score for predicting frailty risk in the short-term for care of people with HIV. AIDS 2023 May 1;37(6):967-975. [doi: 10.1097/QAD.0000000000003501] [Medline: 36723488]
19. Hu Y, Xie M, Zhang Z. Prediction models for frailty in people living with HIV: protocol for a systematic review and meta-analysis of prognostic models. JMIR Res Protoc 2026 May 26;15:e89842. [doi: 10.2196/89842] [Medline: 42189797]
20. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
21. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160. [doi: 10.1136/bmj.n160] [Medline: 33781993]
22. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
23. Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014 Oct;11(10):e1001744. [doi: 10.1371/journal.pmed.1001744] [Medline: 25314315]
24. Acquired Immunodeficiency Syndrome Professional Group, Society of Infectious Diseases, Chinese Medical Association; Chinese Center for Disease Control and Prevention. Chinese guidelines for the diagnosis and treatment of human immunodeficiency virus infection/acquired immunodeficiency syndrome (2024 edition). Infect Dis Immun 2025 Jan;5(1):4-27. [doi: 10.1097/ID9.0000000000000152] [Medline: 42325861]
25. Collins GS, Reitsma JB, Altman DG, Moons KGM, members of the TRIPOD group. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD Statement. Eur Urol 2015 Jun;67(6):1142-1151. [doi: 10.1016/j.eururo.2014.11.025] [Medline: 25572824]
26. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. Ann Intern Med 2019 Jan 1;170(1):W1-W33. [doi: 10.7326/M18-1377] [Medline: 30596876]
27. Debray TPA, Damen J, Snell KIE, et al. A guide to systematic review and meta-analysis of prediction model performance. BMJ 2017 Jan 5;356:i6460. [doi: 10.1136/bmj.i6460] [Medline: 28057641]

XSL•FO
RenderX

28. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368:l6890. [doi: 10.1136/bmj.l6890] [Medline: 31948937]
29. Kitahata MM, Rodriguez B, Haubrich R, et al. Cohort profile: the Centers for AIDS Research Network of Integrated Clinical Systems. Int J Epidemiol 2008 Oct;37(5):948-955. [doi: 10.1093/ije/dym231] [Medline: 18263650]
30. Ruderman SA, Webel AR, Willig AL, et al. Validity properties of a self-reported modified frailty phenotype among people with HIV in clinical care in the United States. J Assoc Nurses AIDS Care 2023;34(2):158-170. [doi: 10.1097/JNC.0000000000000389] [Medline: 36652200]
31. Erlandson KM, Piggott DA. Frailty and HIV: moving from characterization to intervention. Curr HIV/AIDS Rep 2021 Jun;18(3):157-175. [doi: 10.1007/s11904-021-00554-1] [Medline: 33817767]
32. Guaraldi G, De Francesco D, Milic J, et al. The interplay between age and frailty in people living with HIV: results from an 11-year follow-up observational study. Open Forum Infect Dis 2019 May;6(5):ofz199. [doi: 10.1093/ofid/ofz199] [Medline: 31123697]
33. Onen NF, Agbebi A, Shacham E, Stamm KE, Onen AR, Overton ET. Frailty among HIV-infected persons in an urban outpatient care setting. J Infect 2009 Nov;59(5):346-352. [doi: 10.1016/j.jinf.2009.08.008] [Medline: 19706308]
34. Franconi I, Theou O, Wallace L, et al. Construct validation of a Frailty Index, an HIV Index and a Protective Index from a clinical HIV database. PLOS ONE 2018;13(10):e0201394. [doi: 10.1371/journal.pone.0201394] [Medline: 30332410]
35. Paul RH, Cho KS, Luckett P, et al. Machine learning analysis reveals novel neuroimaging and clinical signatures of frailty in HIV. J Acquir Immune Defic Syndr 2020 Aug 1;84(4):414-421. [doi: 10.1097/QAI.0000000000002360] [Medline: 32251142]
36. Inceer M, Brouillette MJ, Fellows LK, et al. Factors partitioning physical frailty in people aging with HIV: a classification and regression tree approach. HIV Med 2022 Aug;23(7):738-749. [doi: 10.1111/hiv.13236] [Medline: 35106895]
37. Escota GV, Patel P, Brooks JT, et al. Short communication: the Veterans Aging Cohort Study Index is an effective tool to assess baseline frailty status in a contemporary cohort of HIV-infected persons. AIDS Res Hum Retroviruses 2015 Mar;31(3):313-317. [doi: 10.1089/AID.2014.0225] [Medline: 25495766]
38. Brothers TD, Kirkland S, Theou O, et al. Predictors of transitions in frailty severity and mortality among people aging with HIV. PLOS ONE 2017;12(10):e0185352. [doi: 10.1371/journal.pone.0185352] [Medline: 28981535]
39. Lorenz DR, Mukerji SS, Misra V, et al. Predictors of transition to frailty in middle-aged and older people with HIV: a prospective cohort study. J Acquir Immune Defic Syndr 2021 Dec 15;88(5):518-527. [doi: 10.1097/QAI.0000000000002810] [Medline: 34757975]

Abbreviations

ART: antiretroviral therapy
CHARMS: Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies
CNICS: Centers for AIDS Research Network of Integrated Clinical Systems
FI: frailty index
FP: frailty phenotype
Frail RISC-HIV: Frailty Risk in the Short-Term for Care of People with HIV
HR: hazard ratio
LASSO: least absolute shrinkage and selection operator
PICOS: population, intervention, comparison, outcome, study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses–Protocols
PROBAST: Prediction Model Risk of Bias Assessment Tool
SWiM: Synthesis Without Meta-Analysis
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
UNAIDS: United Nations Program on HIV and AIDS

Edited by H LaMonica; submitted 26.Aug.2026; peer-reviewed by D Keitrys, T Lillie; comments to author 13.Oct.2026; revised version received 22.Oct.2026; accepted 23.Oct.2026; published 23.Nov.2026.

Please cite as:
Xie M, Hu Y, Yuan J, Zhang Z
Prediction Models for Frailty in People Living with HIV: Systematic Review
JMIR Aging 2026;9:e110499
URL: https://aging.jmir.org/2026/1/e110499
doi:10.2196/110499

XSL•FO
RenderX

© Meilian Xie, Yeye Hu, Junjia Yuan, Zhiyun Zhang. Originally published in JMIR Aging (https://aging.jmir.org), 23.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Global Frameworks Toward Effective Integration of Technology in National Dementia Strategies: Narrative Review and Case Study in Mexico

Lili Marlene Camacho-Bustamante[1], MSc; Claudia I Astudillo-García[2], PhD; Guillermo Velázquez López[3], MD; Luis Gerardo Moncayo-Samperio[3], MD; Alberto Agustín Palacios-García[4], MD; Sergio Alberto Navarro-Tuch[1], PhD

[1]Department of Mechatronics, School of Engineering and Sciences, Tecnologico de Monterrey, Campus Ciudad de México, Prol. Canal de Miramontes, Coapa, Mexico City, Mexico

[2]Dirección de Investigaciones Epidemiológicas y Psicosociales, Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Mexico City, Mexico

[3]Dirección de Servicios Clínicos, Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Mexico City, Mexico

[4]Department of Medical Sciences, School of Medicine and Health Sciences, Tecnologico de Monterrey, Campus Ciudad de México, Mexico City, Mexico

Corresponding Author:
Sergio Alberto Navarro-Tuch, PhD
Department of Mechatronics, School of Engineering and Sciences, Tecnologico de Monterrey, Campus Ciudad de México, Prol. Canal de Miramontes, Coapa, Mexico City, Mexico

Abstract

Background: Population aging is a global phenomenon that has led to a rise in age-related conditions, particularly dementia. This situation presents significant challenges for health care systems, caregivers, and society as a whole. Technologies have shown potential to address the needs of people living with dementia and their caregivers, offering support in daily activities, communication, and care monitoring. However, only a quarter of countries worldwide have developed national dementia plans, and even fewer have integrated technology-based solutions into these policies.

Objective: This study aimed to analyze existing gaps in dementia plans, with a particular emphasis on the limited integration of technological solutions as essential tools for dementia care and management. Based on this analysis, this study proposes a framework to facilitate context-appropriate technology integration.

Methods: This study presents a narrative review of public policy frameworks for dementia care. The review provides a global overview, examines the context in Latin America, and focuses on Mexico as a case study to analyze how technology is integrated into dementia strategies.

Results: Based on these findings, this study proposes a framework of general recommendations adaptable across multiple countries. Furthermore, it demonstrates the application of this framework in Mexico to strengthen its national dementia strategy through technological solutions tailored to people living with dementia and their informal caregivers.

Conclusions: Given that dementia is recognized as a global priority, countries with aging populations have developed numerous technological solutions. These international experiences serve as valuable references to guide the adoption or rejection of specific approaches, provided that local cultural and social contexts are considered. The applied technologies were merged into a framework that can be adapted to different countries, taking into account social aspects, governance, and region-specific barriers to generate a viable, customized option, as demonstrated in the case study focused on Mexico.

(JMIR Aging 2026;9:e83861) doi:10.2196/83861

KEYWORDS

dementia; technologies; policy analysis; literature review; health policy

Introduction

Dementia poses a major challenge for all countries. However, it cannot be addressed without first considering the broader context of aging. Aging is defined by the World Health Organization (WHO) as the gradual accumulation of molecular and cellular damage [1]. These modifications are natural, unavoidable, and start in early adulthood, generating physical and mental changes [2]. In addition to biological changes, aging often entails significant life variations, such as retirement or the loss of close friends and partners [3].

Projections estimate that by 2050, one in 5 individuals worldwide will be aged 60 years or older, totaling between 1.6 and 2.1 billion older adults [4,5]. Aging also increases the risk

XSL•FO
RenderX

of several chronic disorders like neurodegenerative, cardiovascular, and metabolic diseases, as well as other disorders [6].

Among these conditions, dementia is particularly significant. Dementia is a syndrome that may be caused by several diseases and causes cognitive decline [7]. The most common types of dementia include Alzheimer disease (AD), which accounts for approximately 60%-80% of cases, followed by vascular dementia (5%-10%), dementia with Lewy bodies (around 5%), and frontotemporal dementia (approximately 3%) [8]. The most common symptoms of this cognitive impairment may include memory loss, changes in behavior, issues with visuospatial perception, and language problems such as difficulty following conversations [9]. Dementia also impairs the ability to perform daily activities and eventually leads to a loss of individual autonomy [10].

Global costs associated with dementia are typically classified into 3 main categories, including direct medical costs (formal health care services), direct social costs (formal care outside hospital settings), and indirect costs (informal care provided by family members) [11,12].

Because of the importance of this issue and its high cost for governments around the world, as well as the growing number of people affected with between 152 million cases of dementia by 2050 [13], many countries have created public policies to address dementia. These policies aim to meet the needs of people living with dementia and their caregivers.

To effectively address the rising rates of dementia, public policies and action plans must adopt a holistic approach that accounts for the diverse causes and contributing factors of these conditions. In this context, the integration of technology into public policy represents a promising strategy to transform dementia care [14].

The integration of technology into dementia care is increasingly recognized as a key element to enhance early detection, optimize patient management, and improve caregivers' support. Advances in digital health tools, AI, and remote monitoring systems offer promising solutions to address the complex needs of people living with dementia. These technologies can bridge gaps in access to specialized services, particularly in underserved regions, and contribute to more personalized and effective interventions.

One approach to support people living with dementia is the use of assistive technology (AT), along with other technologies for early detection and prediagnosis. AT refers to any devices, tools, or systems that help improve or maintain a person's functional abilities [15,16]. Table 1 shows the types of devices that have been developed based on their specific functions.

Table . Types of assistive technologies. Authors' own creation using data adapted from Klimova et al [16].

Classification	Description
Devices for daily living	Technologies that provide aid with memory, orientation, communication, and other daily activities.
Safety devices	These devices aim to safeguard individuals in cases of disorientation or wandering, and to assist in the prevention of associated safety risks.
Telecare devices	These devices offer services and enable access to support systems, facilitating continuous monitoring and providing assistance when necessary.
Devices to support engagement, social participation, and leisure	Technologies that enhance quality of life by facilitating social interaction and engagement.

This study aimed to evaluate current dementia-related policies from a technological perspective and to propose strategies for the effective integration of technology into dementia care. The analysis progresses from global policies to more specific national policies. It culminates in a Mexican case study that explores potential approaches for integrating technology into national dementia policies.

The first section of this research presents a summary of various global policies and strategies, highlighting their evolution and the extent to which they incorporate technological solutions to address key challenges. Subsequently, an overview of dementia plans in selected Latin American countries and Mexico is provided, as Mexico will serve as the case study in subsequent sections. Thereafter, a framework for technology implementation is presented, taking into account the social context, governance, and regional public policies. Finally, the case study of Mexico is presented to demonstrate the application of this approach.

The main contributions of this paper are as follows:

- Policy-oriented review: a specialized review focused on the integration of emerging technologies within existing public dementia policies.
- Technology identification: identification of current technological tools and their formal inclusion in public health frameworks.
- Adaptive implementation framework: an adaptable decision-making framework designed to bridge the gap between public policies and technological implementation. While Mexico is presented as an illustrative case study to demonstrate its practical application, the framework is fully transferable to other national contexts.

Methods

Objective

The objective of this narrative review was to analyze the global landscape of public policies on dementia, with a specific focus on the application of technological solutions and an emphasis on Mexico, which is subsequently analyzed as a case study.

XSL•FO
RenderX

This analysis was conducted from the perspective of technological implementation, evaluating how official policy frameworks incorporated digital tools and innovation to address the dementia crisis at global, regional, and national levels.

Search Strategy

This review aimed to examine the evolution and current state of the field from the perspective of key stakeholders. Consequently, global analysis incorporated reports and guidelines from the WHO, Alzheimer's Disease International (ADI), and Pan American Health Organization (PAHO), given the review's focus on Latin America.

The search strategy used specific terminology across various databases. For WHO and ADI, the following Boolean query was applied: (public policy OR guide OR plan) AND dementia AND (technology OR tech OR digital OR assistive). The PAHO search was limited to (dementia AND policy) because the platform offered limited search customization. Furthermore, the search was supplemented by a direct review of regulatory documents to analyze the evolution of public policies on dementia, tracing their transition from a general aging context to their current global importance.

Regarding stakeholder websites, publications intended for general awareness or expressing opinions about dementia were excluded. Only reports and guidelines addressing global initiatives focusing on Latin America or discussing the use of technology in dementia care were considered. A screening process was conducted for studies that passed the initial filter, and the relevance of each was discussed with a supervisor. The use of gray literature was selected because it contains the documents related to public policies from different countries. However, the search was limited to the official websites of the main organizations involved in developing or promoting general public policies and dementia-specific policies.

Additional reports referenced within the studies identified during the initial search were also included in order to complement the review.

Data Extraction and Classification

Data extraction was focused on authorship, publication year, primary themes, and document type. For reports, all stakeholders involved in the drafting process were identified. The classification of entries was primarily divided into 3 categories, namely reports, news, and hubs. The latter category served to group related items of the same nature, such as entries that aggregated all dementia-related report results within a specific time period.

Related hubs and news sources were reviewed, and those that did not contain official documents were excluded to retain only official reports.

Results

Overview

The initial search of the WHO website yielded 52 results. Of these, 16 met one of the inclusion criteria, specifically addressing global public policies, focusing on the designated area of interest, or discussing the implementation of technology in dementia care. From these, only 3 official reports were ultimately included.

The search conducted in the ADI database yielded a total of 486 results. However, the most were excluded because they were not formal reports or focused on geographical regions outside the scope of this study.

Consequently, 65 results were selected, of which 25 were reports, primarily consisting of editions of the World Alzheimer Report and supporting documents from From Plan to Impact. Annual follow-up reports on the status of dementia policies were also identified.

The PAHO website search yielded 28 results, of which only one was considered relevant and classified as a technical document. Additionally, other reports cited within the analyzed literature were included. These documents were deemed essential for the development of this review to ensure an exhaustive approach and to provide a comprehensive overview of the current state of public policies on dementia. The review identified key reports with global, Latin American, and Mexican relevance; selected reports are presented in Table 2.

XSL•FO
RenderX

Table . Summary of the main reports identified. Authors' own creation using data adapted from [5,11,17-34].

Title and year	Issuing organization	Geographic scope	Core focus
Global status report on the public health response to dementia (2021) [19]	WHO[a]	Global	Evaluates progress on multisectoral efforts to address dementia globally; identifies implementation barriers of the global action plan and updates dementia burden and cost estimates
Decade of healthy aging: plan of action (2020) [5]	WHO	Global	Ten-year collaborative plan to improve lives of older adults, their families, and communities; placing older adults at the center
Global action plan on the public health response to dementia 2017-2025 (2017) [17]	WHO	Global	Seven priority areas of action to improve the lives of people with dementia and caregivers; include prevention, diagnosis, treatment, care, and caregiver support
The global impact of dementia 2013 - 2050 (2017) [20]	ADI[b]	Global	Policy brief updating prevalence figures, from 44 million in 2013 to 135 million by 2050; 71% will live in low- and middle-income countries by 2050.
Dementia: a public health priority (2012) [18]	WHO and ADI	Global	Joint report published in 2012 that raised awareness of dementia as a public health priority and addressed its global impact, costs, and the need for national and international action
World Alzheimer reports [11,26-34]	ADI	Global	Annual series of comprehensive reports on the global socioeconomic impact of dementia; each report addresses a different topic (2024: attitudes and stigma; 2025: rehabilitation; 2023: risk reduction; 2022: postdiagnostic support; 2021: diagnosis).
From plan to impact reports [21-25]	ADI	Global	Annual series monitoring member states' progress toward the WHO Global Action Plan; 7 chapters cover the 7 action areas of the plan

[a]WHO: World Health Organization.

[b]ADI: Alzheimer's Disease International.

Global Dementia Policies and Technology Integration

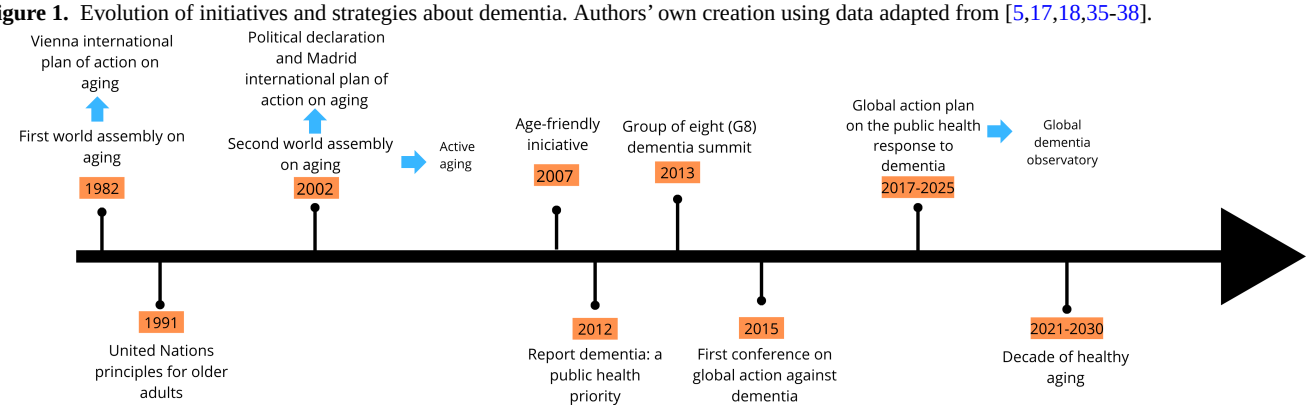
To propose a comprehensive approach to technology integration, it is necessary to understand the current state of global strategies and policies that prioritize dementia and the aging population. This understanding helps identify gaps and guide future policy development to better address the needs of individuals living with dementia and older adults.

The significance of this issue is underscored by the efforts of various organizations. In response to the global aging population and its resulting consequences, organizations such as the Alzheimer's Association, WHO, and the United Nations have emphasized the inclusion of this topic on the political agenda. They advocate developing strategies aimed at improving the quality of life for people living with dementia and their caregivers. However, according to the WHO, only a quarter of all countries have implemented a national plan or strategy to address dementia for 2021 [14]. Half of these countries are in Europe, and many of their plans had expired or were nearing expiration.

This global response is further illustrated in Figure 1, which demonstrates the chronological evolution of dementia policies and the primary events that have made dementia a visible and essential issue on the global stage. Because most people living with dementia are aged 60 years or older, early initiatives primarily focused on the needs of the older adults.

XSL•FO
RenderX

Figure 1. Evolution of initiatives and strategies about dementia. Authors' own creation using data adapted from [5,17,18,35-38].



The initial collaborative action addressing dementia was the First World Assembly on Aging in Vienna in 1982 [39]. This early initiative emphasized the imminent aging of the global population and focused on addressing humanitarian and developmental issues, including the needs of older adults and their role within society. As a result of this meeting, the International Plan of Action on Aging was developed. The first Action Plan on Aging sought to increase awareness of the challenges that population aging would present across all sectors. Its objectives included raising awareness of the economic, social, cultural, humanitarian, and socioeconomic implications of population aging [40].

Despite its intentions, the recommendations outlined in the first Action Plan on Aging lacked specificity regarding the actions needed to achieve its objectives. The plan also lacked specific timelines and measurable goals. Nevertheless, it served as a foundational reference for the development of more robust initiatives.

In response to the growing importance of aging, the United Nations established the Principles for Older Persons in 1991 [37]. The 18 principles address 5 areas: independence, participation, care, self-fulfillment, and dignity. By 2002, the primary focus remained on aging as a broad social issue. In response, the Second World Assembly on Aging was held in Madrid, resulting in the adoption of the International Plan of Action on Aging. This political declaration, composed of 19 articles, emphasized the need to incorporate aging into public policy. Its main priorities included older adults and development, promoting health and well-being in old age, and creating enabling and supportive environments [35]. Notably, the document emphasizes the role of technology as a critical tool in addressing the challenges associated with aging.

In 2012, the Global Report on Dementia was published, calling for urgent action on dementia worldwide [18]. Regarding technology, the report discussed ethical issues such as privacy and security, described existing tools, and highlighted the need for more research and evaluation to show the benefits of technology in dementia care.

In 2013, the G8, a group composed of the industrialized countries and representing approximately half of the world's economic activity, held the Dementia Summit in London [38]. At this event, ministers, researchers, and pharmaceutical companies gathered to discuss the importance of addressing dementia and the major challenges it presents, such as the lack of funding for dementia research.

In 2015, the first Conference on Global Action Against Dementia was held. The goal of this conference was to outline future action plans and to strengthen the call for all countries to recognize dementia as a public health priority. The conference also highlighted the importance of developing technological innovations to support people living with dementia, as well as the need for more information on how to effectively integrate these technologies into dementia care. It also stressed that national dementia policies must include a defined role for technology within coordinated care strategies [36].

These collaborative efforts culminated in the presentation of the Global Action Plan on the Public Health Response to Dementia in 2017 [17]. This plan specified objectives to be achieved during the period of 2017 - 2025 and the elements to be evaluated. In this document, technology was mentioned mainly in the seventh strategy action area, dementia research and innovation. The proposal included increasing investment in research focused on technological innovation in health care, particularly in the development of tools to address the needs of people with dementia and their caregivers. It also emphasized the importance of providing access to assistive technologies and encouraging the active participation of individuals living with dementia in the evaluation and design of these solutions.

To monitor progress on the goals of the Global Action Plan, the Global Dementia Observatory (GDO) was created [41]. It defines specific requirements for each area of action. In relation to technology use in care, only one of the ten indicators includes a subindicator that evaluates the availability of care products and assistive technologies. This subindicator considers 3 categories, including adult hygiene products, home modifications, and assistive devices. However, it only includes basic tools such as walking frames and hearing aids, without explicitly mentioning more advanced technologies like tracking systems or ambient assisted living.

Although some health policies address the use of AI and its ethical and legal implications, they often do so from a general perspective, without focusing specifically on dementia. For example, the WHO's Ethics and Governance of Artificial Intelligence for Health (2021) [42] acknowledges the role of digital tools and innovation but offers no specific guidance on their development, implementation, or evaluation in dementia care. This lack of direction limits the effectiveness of current

XSL•FO
RenderX

policies and highlights the need for frameworks that clearly define the role of technology in improving outcomes for people living with dementia.

As a significant milestone, The Decade of Healthy Aging (2021‐2030) [5] is a 10-year initiative aimed at changing perceptions about aging, promoting age-friendly environments, creating integrated and responsive health care systems, and ensuring access to long-term care for older adults. The initiative also emphasizes the role of technology in enhancing care for older adults. Its principal contributions to the technological framework encompass the development of standards for access to information and communication technologies and assistive devices, the promotion of affordable and effective digital tools, and the implementation of innovative long-term care services.

Latin American Dementia Policies: Regional Technology Landscape

Population aging is a phenomenon affecting the entire world, including Latin America, leading to a significant increase in dementia cases. In fact, in Latin America, dementia is the sixth leading cause of death and the primary cause of dependency [43]. Additionally, the prevalence of dementia in the region is 8.5% among individuals aged 60 years and older, the highest in the world; this figure is expected to reach 19.3% by 2050 [44].

The lack of formal studies in various Latin American countries makes it difficult to understand the existing issues and forces the generalization of incidence, prevalence, and cost figures, relying on countries that have conducted more studies, such as Cuba, the Dominican Republic, Mexico, Peru, and Venezuela [43].

The development and evolution of policies and plans to combat dementia have become crucial as countries acknowledge the growing proportion of their populations aged 60 years and older. This demographic shift emphasizes the prevalence of common illnesses and conditions associated with aging. Nations have adopted 2 main approaches: incorporating dementia initiatives within broader national plans that also address other neurodegenerative diseases or developing specific strategies targeting dementia or its most common type, AD. For instance, Mexico's initial plan focused on AD, while countries like Chile and the Dominican Republic have implemented general plans targeting dementia [45].

According to ADI, as of 2024, 8 countries in Latin America (Brazil, Chile, Cuba, Costa Rica, Dominican Republic, Mexico, Puerto Rico, and Uruguay) had a formal dementia plan, with 2 more countries (Panama and Peru) developing their plans. A review of policies from 6 Latin American countries indicates that all these plans align with the recommendations and action areas proposed by the WHO in the Global Action Plan on the Public Health Response to Dementia [17]. These areas include early diagnosis, staff training, public awareness campaigns, and improving the quality of life for people living with dementia. Additionally, each plan includes objectives tailored to their specific context and needs, such as enhancing monitoring and data collection related to the disease [45-49]. All these plans are acknowledged by ADI as the most recent for each country (Multimedia Appendix 1).

Some of these plans were established many years ago and have yet to be updated, such as those of Cuba and Chile [46,48]. Others were designed with a long-term perspective, like Costa Rica's plan, which was set to last 10 years until its deadline in 2024 [50]. Currently, many of these plans are undergoing updates. For example, Mexico's dementia plan, initially created in 2014, was updated in 2024.

In these plans, early diagnosis is prioritized through the training of frontline staff and the use of existing resources, as opposed to creating new services. Another important aspect is the inclusion of action lines aimed at informing people living with dementia about their legal rights.

Most of the dementia plans and strategies reviewed mention technology only as a tool for sharing information with caregivers and people living with dementia. Some also refer to promoting innovation in research. However, only Chile's plan explicitly emphasizes the use of technology to improve quality of life and the development of devices to support people with dementia [46].

Mexico's National Dementia Strategy

Overview

Dementia represents a growing concern in Mexico. Currently, an estimated 1.3 million people in Mexico are living with dementia, and projections indicate that this number could increase to 3.5 million by 2050 [47].

In response to this challenge, the first national plan specifically addressing dementia was introduced in 2014 under the title "Plan de Acción Alzheimer y Otras Demencias." This initiative aimed to enhance the well-being of individuals living with dementia by reinforcing the health care system through 7 key objectives. These included eliminating stigma, improving access to accurate information for patients and caregivers, and developing a comprehensive care system that integrates medical, social, and community support.

Despite the plan's commitment to periodic monitoring and evaluation, a formal update was not introduced until 2024 with the publication of the National Dementia Plan [47]. While this new strategy retains core objectives, such as eliminating stigma and supporting scientific research, it introduces additional measures aimed at strengthening early detection through a comprehensive care model. The primary goal of these enhancements is to improve patient outcomes and reduce the prevalence of dementia-related complications. A key addition to this plan is the strategic involvement of civil associations, which play a critical role in raising awareness and disseminating information about dementia. The plan also prioritizes the reinforcement of the primary health care system by integrating systematic screening processes and equipping health care professionals with improved diagnostic tools and training. These initiatives seek to establish a more effective and accessible support system for individuals living with dementia and their caregivers.

XSL•FO
RenderX

As observed, the application and development of technologies are largely absent from both new and existing plans. Therefore, it is essential to identify the key sectors where the use and development of technological solutions can be effectively integrated and leveraged for their potential impact.

The following points summarize the most pressing needs that must be addressed to build an effective and inclusive dementia care strategy in Mexico related to the application of technologies:

Strengthening Primary Health Care Networks

Primary health care providers, including those within the Mexican Social Security Institute (Instituto Mexicano del Seguro Social [IMSS]) and the Institute of Social Security and Services for State Workers (Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado [ISSSTE]), should receive training to recognize early signs of dementia, ensure timely referrals to specialized care, and use appropriate technologies for screening and early detection.

Support for Informal Caregivers

Given the cultural significance of family caregiving in Mexico, it is essential to implement policies that provide effective support to informal caregivers. This support may include financial assistance, respite care services, and access to training programs delivered through the Secretariat of Welfare or local community centers to strengthen caregiving capacity and enhance the quality of care provided.

Within this framework, the Mexican government has introduced the Salud Casa por Casa program, which delivers home-based health care services to individuals aged 65 years and older [51]. The program is structured across 4 levels of care that respond to varying degrees of clinical need, encompassing stable conditions as well as complex cases that require continuous medical attention and sustained caregiver involvement.

This initiative presents a valuable opportunity to integrate smart environments into long-term care and advance the development of large-scale technological solutions aimed at monitoring individuals with dementia. Such innovations can support early detection, ongoing cognitive assessment, and personalized care interventions within the home setting. However, this program remains in the process of structuring and implementation, and it is not yet operational.

Public Awareness and Education Campaigns

A nationwide public awareness campaign should be launched to reduce the stigma associated with dementia. The campaign could be spearheaded by the Ministry of Health, in collaboration with nongovernmental organizations (NGOs) like Federación Mexicana de Alzheimer (FEDMA), to educate the public about the importance of early diagnosis and the available resources. Notably, efforts are already underway through the Dementia Friends initiative, introduced in Mexico by FEDMA in 2018, with the primary objective of eliminating stigma [52].

Investment in Research and Innovation

To address the growing challenge of dementia in Mexico, it is crucial to increase funding for research through institutions like the Secretariat of Science, Humanities, Technology, and Innovation (SECIHTI). This funding should focus on studies investigating genetic and environmental factors contributing to dementia within the Mexican population and developing culturally appropriate care models. Using databases with cohorts on the health of the Mexican population offers significant opportunities to analyze data for better policies and informed decisions.

Since 2001, the Mexican Health and Aging Study (MHAS) has gathered essential data for policy development and health care strategies. Based on a nationally representative sample of 15,186 Mexicans aged 50 years and older, this longitudinal study collects information on health status, socioeconomic conditions, family structure, living environments, psychosocial factors, and major life events [53]. MHAS has expanded to include specialized assessments like cognitive evaluations and biological sample collection, providing valuable insights into aging and dementia.

Additional cohorts, such as the Cohorte de Obesidad, Sarcopenia y Fragilidad de Adultos Mayores Mexicanos, offer longitudinal data on sociodemographic information, cognition, nutrition, and other health aspects [54]. The Encuesta Nacional de Salud y Nutrición (ENSANUT) provides data on risk factors like smoking, diabetes, and cardiovascular disease [55]. Integrating these rich data sources can help Mexico advance research, develop effective policies, and create culturally tailored interventions to reduce the impact of dementia and other health challenges.

International Collaboration

Mexico should actively engage in international forums and collaborative initiatives focused on dementia care. This participation could involve partnerships with organizations such as the WHO and the PAHO, as well as the establishment of bilateral agreements with countries that have implemented advanced dementia care strategies. Through these collaborations, Mexico could both benefit from and contribute to the exchange of innovative technologies and best practices, particularly in the areas of early diagnosis and nonpharmacological treatment approaches.

Implementation of Community-Based Care Models

Community-based care models, such as those implemented by the Network of Municipalities for Health (Red de Municipios por la Salud) [56], should be expanded to include dementia-specific services. These models could provide localized support and resources for individuals living with dementia and their families. Additionally, they could facilitate the formation of dedicated teams to assess the acceptability, usability, and effectiveness of technological devices designed for people with dementia and their caregivers, thereby promoting more responsive and context-appropriate innovations in care.

Technology Implementation Framework

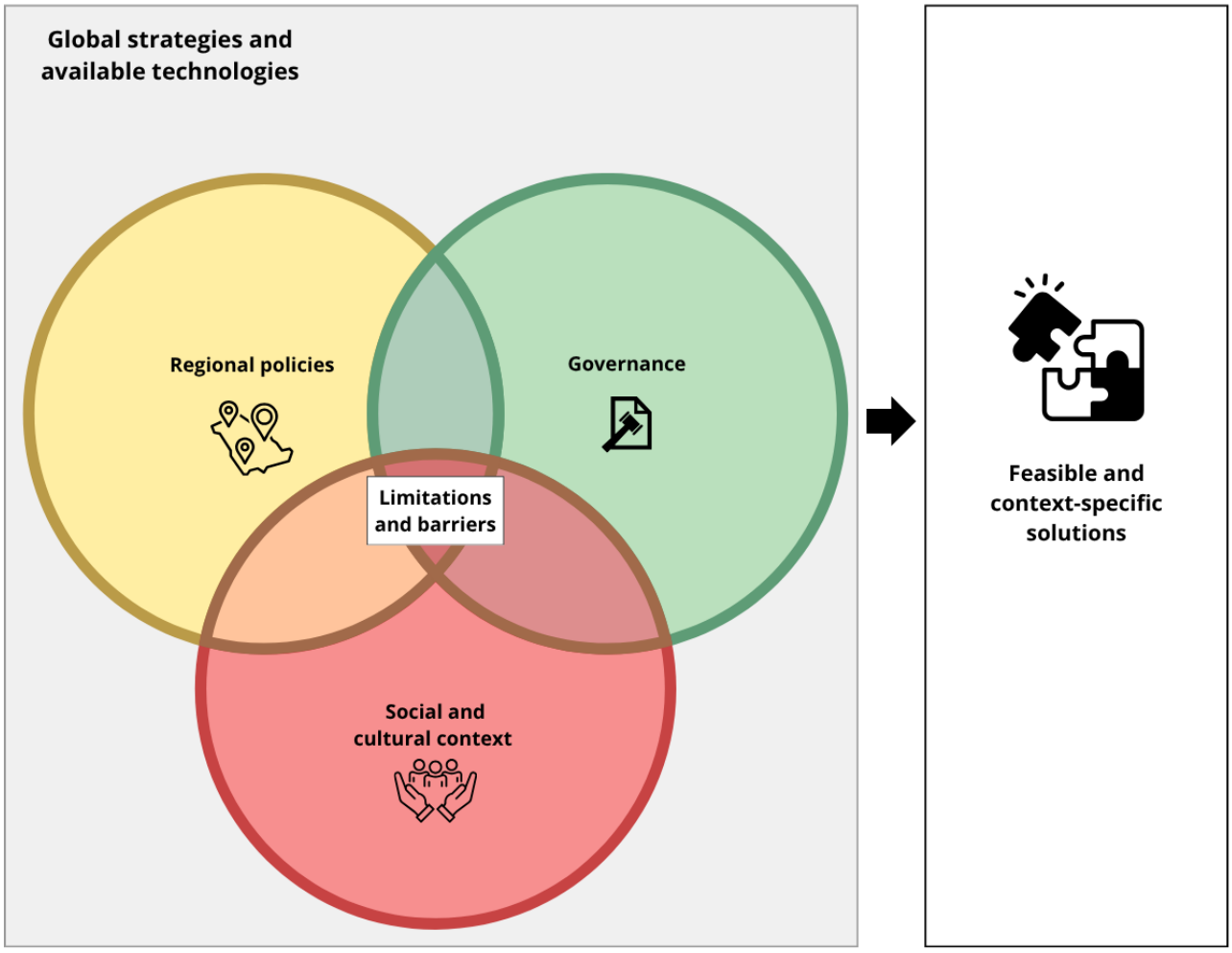
Technology implementation aligned with global and national strategic goals requires a structured procedure. This procedure considers overarching global expectations before evaluating country-specific conditions. This approach is presented first in

XSL•FO
RenderX

general terms and is subsequently illustrated through a case study focused on Mexico. The objective of this framework is to allow adaptation to the specific characteristics of different national contexts.

This section presents a framework, as depicted in Figure 2, which first outlines the global landscape with its corresponding policies, strategies, and the technologies to be implemented within each key area. Subsequently, a case study focused on Mexico is presented to illustrate the application of this framework in a specific environment. The intersection of regional policies, governance, and the social and cultural context is evaluated to identify constraints and barriers, thereby determining the viable options for this country.

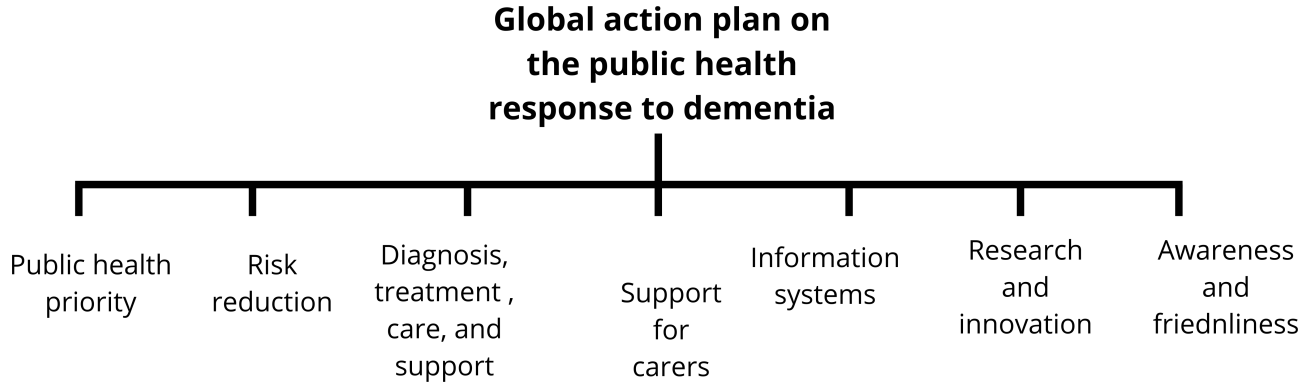
Figure 2. Conceptual framework for contextualizing global technology recommendations.



Following the analysis in the previous sections, the main areas where various policies converge were identified. Consequently, the action areas specified in the Global Action Plan on the Public Health Response to Dementia were selected as the foundational baseline, as shown in Figure 3. These areas serve to identify core priority approaches and determine which existing technologies can facilitate the fulfillment of their objectives.

XSL•FO
RenderX

Figure 3. Key strategic areas of the World Health Organization (WHO) global action plan on dementia. Authors' own creation using data adapted from [17].



Based on the analysis, the main issues addressed in global plans and proposals were identified and summarized in Table 3. Along with each specific strategy, the table outlines key problems within each domain, the technologies previously applied to address these issues, and specific examples from the literature. A more detailed examination of these applications follows.

XSL•FO
RenderX

Table . Technologies implemented based on the intervention strategy.

Strategy, issue, and technologies	Approach
Diagnosis, treatment, care, and support	
Underdiagnosis	
Machine learning	[57-62]
Data collection and tracking	[63,64]
Virtual reality	[65,66]
Automatized screening tools; apps for tablets and smartphones	[67,68]
Training	
Virtual reality	[69,70]
BPSD[a]	
Virtual reality	[71,72]
Cognitive training and rehabilitation	
Virtual reality	[69]
Support for carers	
Social interaction	
Virtual reality, wearables, social robots, and generative AI	[73,74]
Long-term care	
Smart environments and wearables	[75-77]
Safety and monitoring	
Wearables	[78,79]
Managing care	
Generative AI	[80]
Awareness and friendliness	
Removing stigma	
Virtual reality	[81]
Awareness	
App for tablets and smartphones	[82]
Risk reduction	
Risk index	
Machine learning	[83,84]
Cognitive training	
Virtual reality, apps for tablets and smartphones, and video games	[85-87]
Information systems	
Cohort construction	
Electronic health records	[88,89]

[a]BPSD: behavioral and psychological symptoms of dementia.

Integrating technology into dementia care improves effectiveness when each approach aligns with its intended support area. To facilitate this process, various technological solutions are identified for each strategy based on their relevance and potential applications. This framework serves as a practical decision-making guide to assist stakeholders in selecting suitable approaches based on their strategic priorities.

The areas of public health priority, as well as research and innovation, were excluded from this analysis, given that existing policies have already been evaluated. Furthermore, any technology integrated to support the fulfillment of these plans inherently drives innovation and, in many cases, requires research to ensure successful implementation.

Virtual reality is a type of technology that enables the creation of synthetic environments where users can engage emotionally,

XSL•FO
RenderX

cognitively, and behaviorally, resulting in immersive and meaningful experiences. In the context of dementia care, virtual reality offers several approaches that open new possibilities for early detection and more precise assessment in everyday situations. These include the use of eye-tracking technologies to support diagnostic processes [66], as well as systems that evaluate task performance in daily living [65]. Moreover, virtual reality has been used to train caregivers in the application of care techniques and in responding to behavioral and psychological symptoms [71,72], helping to create more supportive and calming environments. For people living with dementia, it has contributed to maintaining functionality in daily activities and encouraging social interaction [67,70]. It has also proven useful in promoting public understanding and empathy by simulating the sensory and cognitive experiences of a person with dementia [81].

Beyond immersive experiences, another key technological advancement is the use of machine learning algorithms, which offer the potential to identify early cognitive changes through pattern recognition across diverse data sources. Machine learning algorithms are useful tools in dementia research because they can detect subtle and complex patterns across different types of information. In the prediagnostic stage, they are often applied to analyze clinical histories, language use, and physiological indicators. These types of data contribute to identifying early cognitive changes that may not be evident through traditional clinical evaluations [57-59]. Machine learning has also been used to develop scoring instruments such as the Mini-Mental State Examination (MMSE) and the Clock Drawing Test [57-59], supporting the analysis of results and helping to identify individuals who may show signs of cognitive impairment.

Risk reduction is also an area to be addressed. To support these applications, it is necessary to develop longitudinal databases that organize and store diverse sources of information over time. This allows researchers to examine the progression of symptoms and identify traits shared by individuals at higher risk. Several countries have already established national databases with this purpose [63,64].

Smart environments offer a practical solution to provide this data by embedding sensors in the living spaces of people with dementia, enabling passive monitoring of daily activities and health indicators over time. These systems aim to monitor individuals living with dementia and enable them to remain in their homes for a longer period, preserving their autonomy and connection to familiar surroundings and their community. Smart environments can support daily living by tracking routines, providing reminders about medication, hydration, and other activities, and even issuing alerts to caregivers regarding the individual's condition or location.

As a more accessible and flexible alternative, wearable technologies provide real-time physiological and behavioral data, making them suitable for both urban and rural settings in dementia care. Wearable devices have become valuable tools for caregivers, largely due to their increasing accessibility. In countries such as Japan, these technologies have already been implemented with the primary goals of monitoring vital signs in people living with dementia, preventing wandering, and identifying behavioral patterns over time [82]. To complement this, mobile apps on smartphones and tablets offer a widely accessible means of raising public awareness and supporting early detection through interactive tools.

One of the most accessible ways to reach the general public is through apps available on smartphones and tablets. In some countries, these platforms have been used to raise awareness about the importance of brain health [82], while also providing ready-to-use apps that offer a general overview of an individual's cognitive status. Closely related to this are efforts to automate existing screening tools. These tools have the potential to expand the reach of detection efforts and support earlier identification of dementia. In addition, several tests have already been adapted into digital formats or modified to deliver fast and simple prediagnostic assessments to a larger segment of the population [67]. Despite their potential, these apps are often limited to predefined functions and lack the ability to adapt dynamically to users' specific needs.

Generative AI (GenAI) overcomes these limitations by enabling dynamic content creation and real-time conversational interactions, offering new ways to support both people with dementia and their caregivers. The emergence of GenAI has introduced new possibilities through technologies capable of producing text, images, and other multimedia formats using generative models [90]. These models are trained on large datasets to generate content based on learned patterns, with potential applications in dementia care. Their primary use has been to support social interaction and assist in managing the condition [80].

In particular, GenAI has been used to produce text designed to simulate social conversations, support reminiscence therapy, and assist caregivers by providing responses to common questions related to behavioral symptoms and everyday care scenarios. However, the development of these technologies requires advanced technical expertise, meaning that their application will likely come through ready-to-use apps or be driven by researchers and developers working in dementia care.

One particularly promising approach is the use of social robots to provide companionship for people living with dementia. Some of these robots are capable of verbal interaction or offer feedback based on users' responses. These features make them potentially valuable for supporting emotional well-being and engagement. However, this approach still presents important challenges. The high cost of these devices and the quantity required for use in clinical or institutional settings represent significant financial barriers. In addition, older adults may experience difficulties interacting with such technology, which can lead to resistance or rejection, especially given the digital divide often observed in this age group. Studying user interaction, acceptance, and functionality in local contexts would help identify improvements and inform future development.

Following the global identification of technologies aligned with broad needs, it is necessary to identify the specific strategies and requirements outlined in national or regional dementia plans. Based on these context-specific priorities, relevant technologies can be selected from Table 3. The opportunities for applying

XSL•FO
RenderX

these technologies must then be evaluated relative to the social context.

As a next step, current governance frameworks, laws, and regulations governing technology use in these specific environments must be identified. Subsequently, potential limitations and barriers arising from social, cultural, and broader regional characteristics should be determined. Finally, strategic actions can be proposed by synthesizing all contextual factors with available technologies.

The implementation of this framework is demonstrated below using the specific case study of Mexico.

Mexico as an Illustrative Case Study

Based on the public policies on Mexico analyzed in the previous sections, the identified areas of opportunity can be addressed using specific technologies, as outlined in Table 3.

Mexico, like many other countries, has low rates of dementia diagnosis, with cases often being detected only at advanced stages. Therefore, the development of technologies that enable early prediagnosis represents an essential area of focus. In this context, it is possible to build upon existing elements by integrating digital biomarkers into already established aging surveys, such as MHAS. Combined with the use of AI for pattern detection, this represents a significant opportunity for the development of AI models capable of identifying the main patterns associated with dementia specifically in the Mexican population, while considering the country's social and cultural context. These models could help inform public policies and support the development of risk indices with high specificity and sensitivity for the Mexican population.

The development of these risk indices could be used as a tool to identify populations with these characteristics or risk factors. In addition, the development of tools based on digital biomarkers, such as voice, gait, or facial expressions, could be used for screening at the primary care level and facilitate referral to specialist physicians. In particular, the creation of databases containing voice recordings from Mexican Spanish speakers is a major unmet need, since currently available databases of voices from people with dementia are predominantly in English, with only a limited number available in Spanish, but not specifically in Mexican Spanish. Consequently, potential biases may arise due to differences in linguistic and cultural characteristics, making localization and adaptation necessary.

MHAS, particularly the Cognitive Aging Ancillary Study (Mex-Cog) substudy, represents a valuable starting point [53]. Expanding the scope of collected data could further enhance its potential. Including additional variables commonly used in international research would strengthen early detection efforts and make it possible to generate risk profiles adapted to the characteristics of the Mexican population. However, the effectiveness of machine learning depends heavily on the availability of comprehensive, high-quality longitudinal data. Furthermore, the integration of digital biomarkers would contribute to the development of more complex and robust databases.

The integration of wearables into the health care system could represent an opportunity for digitalization through the incorporation of devices such as wristbands capable of monitoring vital signs and possible on/off states depending on the condition involved. The aim would be to integrate these data into electronic medical records and facilitate patient monitoring during treatment. These changes should be addressed in the short to medium term, since the health care system is still under development and regulations for the integration of such data into medical records are still required. Furthermore, the integration of digital biomarkers will require restructuring existing surveys, as well as clearly defining which data will be collected, how these data will be harmonized over time, and how they will be anonymized.

As discussed in previous sections, policies should not only focus on diagnosis but also address the context of caregiving. In Mexico, where informal care for older adults is predominantly provided by female family members [91], interventions should prioritize reducing caregiver burden and enhancing the quality of life of older adults. Care-related technologies could be implemented in 3 settings: the home environment, day-care centers, and caregiving systems. Their integration could also be considered the development of smart cities.

The application of technologies within public policies should be considered as part of a comprehensive strategy that also includes digital literacy. In Mexico, this could be promoted through free technology familiarization courses introducing care technologies in day-care centers, as well as through training programs directed toward primary users of these technologies. Furthermore, initiatives for technological familiarization could be delegated to institutions such as Instituto Nacional de las Personas Adultas Mayores (INAPAM), with the objective of improving digital literacy among populations at risk due to age while also promoting the acquisition of new knowledge.

Many caregivers have begun to show interest in implementing technologies in the care process; however, they often do not know how to use them, which represents a significant barrier. Some municipalities in Mexico have already implemented programs that provide wearable GPS devices free of charge to help ensure the safety of people with dementia and prevent them from getting lost [92-94]. Nevertheless, these initiatives have not yet been implemented throughout the entire country. Providing wearable devices that enable monitoring, together with the development of guidelines to support caregivers and disseminating these resources through day-care centers, Sistema Nacional para el Desarrollo Integral de la Familia (DIF) institutions, and health care centers, may represent a viable strategy to facilitate decision-making among caregivers. This approach could also help promote the appropriate use of technologies in ways that benefit both caregivers and people living with dementia.

With the increasing affordability and availability of virtual reality devices, the implementation of these tools in Mexico is becoming more feasible, with potential applications in home care, clinical settings, and day centers. Its implementation in adult day-care centers could be carried out in a phased and controlled manner. This approach would allow for initial user

XSL•FO
RenderX

familiarization, enabling subsequent implementation at home during visits from the proposed care system, specifically for rehabilitation purposes. Furthermore, this tool can facilitate personnel training in the early identification of symptoms and the administration of screening tests.

Beyond virtual reality, other technologies could support the care of people with dementia. In the Mexican context, this approach presents both opportunities and challenges. On the one hand, smart home technologies could help address the growing need for home-based care, especially in families where formal caregiving services are limited. On the other hand, the implementation of sensor-based environments may face obstacles such as high installation costs, technical complexity, and concerns about user privacy. These factors highlight the importance of considering culturally and economically appropriate adaptations to ensure the feasibility and acceptance of such technologies in Mexican households.

In fact, studies conducted in Mexico show that WhatsApp (Meta) is one of the most widely used technologies among older adults [95]. Therefore, older adults may be relatively familiar with mobile devices [96]. Consequently, the implementation of applications designed to support cognitive functions and adapted to the local context represents a promising alternative. This idea was reinforced by the Latin American Initiative for Lifestyle Intervention to Prevent Cognitive Decline (LatAm-FINGERS) study, in which cognitive interventions were provided through the use of mobile apps, and participants were instructed on how to use the devices [97].

Although the integration of GenAI into voice assistants represents an emerging area of technological innovation, empirical research regarding user acceptance of these systems in Mexico remains limited. However, existing studies indicate that the main difficulty is the accessibility of these devices, as only 3 of the 40 people interviewed reported having access to one [98]. Nevertheless, their potential for integration into homes and day-care centers was also identified.

This highlights the importance of supporting research and development in the field of health-related technologies. Furthermore, because these models are trained for specific cultural and linguistic contexts, their implementation in Mexico will require careful adaptation to the country's social and cultural environment.

Despite these limitations, acquiring a single unit for research purposes could offer a valuable opportunity to explore the necessary adaptations for implementation in Mexico. Studying user interaction, acceptance, and functionality in local contexts would help identify improvements and inform future development.

Given these challenges, the above approach offers a more accessible and scalable alternative that may be better suited for immediate integration into dementia care strategies in Mexico.

Data Management and Regulatory Governance

Although the use and management of data to support tools for treatment and diagnosis, as mentioned in the previous section, represent a significant opportunity for the digitalization of health care in Mexico, there are still no regulations capable of adequately governing these processes. Existing regulations are primarily limited to electronic medical records under NOM-024-SSA3-2012 and NOM-004-SSA3-2012 [99,100]. However, there are currently no specific regulations addressing the integration of wearable devices into these records or processes related to data collection, transfer from applications to clinical servers, and subsequent data use and management.

Furthermore, there are no clinical practice guidelines to support decisions about which devices should be integrated or recommended. This represents a major limitation for the implementation of these technologies in real-world clinical settings in Mexico.

Limitations and Barriers in Mexico

During the process of implementing technologies in Mexico, several factors arise that complicate their integration. When addressing dementia, the context of older adults cannot be overlooked. In Mexico, the living conditions of older adults vary considerably, resulting in socioeconomic, geographic, and educational differences that create multiple scenarios of difficulty for technology integration.

The integration of technologies often requires infrastructure capable of supporting their operation. For example, voice assistants require constant internet connectivity; however, in many rural areas of Mexico, this is not feasible, thereby creating additional gaps or inequalities in access to health care. Furthermore, the digital divide among older adults reflects differences in their circumstances and access to resources distributed across groups that reflect the limitations of their environment. While many older adults show interest in learning and using technology, they often lack the means to do so. Others may have little or no interest, possibly due to anxiety associated with interacting with technology or because of the lack of infrastructure necessary to sustain its usefulness [95,101,102].

Regarding the possible integration of precision medicine approaches supported by AI, this would require the centralization of data to enable the construction of well-founded individualized analyses. However, although electronic medical records have been mandatory for both public and private institutions in Mexico since 2026, their implementation is still ongoing [103]. In addition, the inclusion of digital biomarkers and data obtained from wearable devices has not yet been clearly defined or regulated.

Another important limitation is the fragmentation of the health care system, which prevents the establishment of a truly universal health care system applicable to the entire population [104]. This creates an additional barrier to the implementation of any public health policy.

Strategic Directions for Implementation

The integration of technologies into public policies will require a solid foundation for their implementation. At present, for Mexico, the short-term path appears to involve the development of these foundations, through which technologies may later be integrated more smoothly. In the short term, this includes the establishment of adequate infrastructure to guarantee internet

XSL•FO
RenderX

access for the entire population, as well as the promotion of digital literacy among older adults both with and without dementia. In addition, caregivers should also be trained through local programs on how to select appropriate devices or technologies according to their specific needs.

Furthermore, in the short term, regulations will also be required to guide health care professionals on how to integrate data obtained from wearable devices into electronic medical records in a precise and regulated manner.

With these foundations in place, and under the assumption of an already established care system, it would be possible to facilitate access for caregivers to basic monitoring devices that support daily life activities and enable integration into electronic medical records.

Once clear regulations regarding data storage and monitoring have been established, the next medium- to long-term step would be to develop dementia risk indices specific to the Mexican population, as well as standardizing metrics and criteria for decision-making based on these data. From this point onward, the pathway would involve the development and implementation of technologies specifically focused on the Mexican context.

Discussion

The prioritization of dementia as a critical issue has been increasing over time. However, the use of next-generation technologies is not yet fully considered. Therefore, it is necessary to leverage recent technological advances as an opportunity to reactivate the goals established in these international plans. Additionally, it is essential to include measures of the development and application in evaluations of new technologies, such as AI and smart environments.

Since dementia has been declared a global priority and aging has already impacted other countries, many technologies have been developed in those places. These technologies can guide the countries in deciding which paths to follow or avoid, always considering the cultural and social context. Although technology can bring many benefits, barriers such as cost and limited availability reduce its use. To move forward, it is necessary to create strategies that ensure access and promote the use of these tools in dementia care.

Moreover, the implementation of AI in the analysis of medical data can facilitate more accurate diagnoses and early identification of symptoms, which is crucial for appropriate treatment. Telemedicine platforms can also enable remote consultations with specialists, improving accessibility and reducing the burden on the health care system. These advancements can significantly contribute to enhancing the quality of life for individuals living with dementia and optimizing the time and effort of caregivers.

Leveraging these technological advancements requires increased investment in research and innovation. Increased funding for this field would not only improve existing tools but also drive the development of new solutions tailored to the specific needs of each population.

However, ethical issues must also be considered. Some technologies collect personal information, which raises questions about privacy and security. Other tools help people with daily tasks or social activities. In these cases, it is important that the technologies are reliable. There are also concerns about replacing human caregivers with machines, which could reduce the emotional connection in care.

Even if the technologies are available, their success will depend on how well they meet the needs of caregivers, people with dementia, and health care professionals. These tools must be useful, easy to use, and adapted to the reality of the people who need them.

Related to Mexico as a case study, integrating technology into dementia care can not only improve diagnostic and treatment processes but also help reduce geographical and economic barriers that limit access to specialized medical services. Digital tools, such as mobile apps for health monitoring, activity-tracking devices, and case management software, can provide effective solutions for continuous and personalized care, especially in rural areas or communities with limited resources. Although the proposed social programs have significant potential, Mexico’s health system is currently undergoing restructuring, and the effective integration of these technologies will depend on the final performance and organization of the new programs.

Acknowledgments

The CIMB Laboratory at Tecnológico de Monterrey, Mexico City Campus, is providing support for this research. The study has received financial assistance from Tecnológico de Monterrey through scholarship A01551004 and from the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) under CVU No. 11512001151200.

Generative AI Use

While preparing this manuscript, the authors used ChatGPT (OpenAI) in order to enhance the language and style of the text. Following this, the authors reviewed and made any necessary edits to the content and accept full responsibility for the content of the publication. This tool was not used to generate content, analyze data, or draw conclusions.

XSL•FO
RenderX

Funding

The CIMB Laboratory at Tecnológico de Monterrey, Mexico City Campus, is providing support for this research. The study has received financial assistance from Tecnologico de Monterrey through scholarship A01551004 and from the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) under CVU No. 1151200.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

Conceptualization: LMCB (lead), SANT (equal)
Data curation: LMCB
Formal analysis: LMCB (lead), CIAG (supporting), LGMS (supporting), GVL (supporting), AAPG (supporting)
Investigation: LMCB
Methodology: LMCB
Project administration: SANT (lead)
Supervision: SANT (lead), LMCB (equal)
Validation: AAPG (equal), CIAG (equal), GVL (equal), LGMS (equal)
Visualization: LMCB (lead)
Writing – original draft: LMCB (lead), AAPG (supporting)
Writing – review & editing: LMCB (lead), SANT (supporting), AAPG (supporting), CIAG (supporting), GVL (supporting), LGMS (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Review of national dementia action plans by country.
[DOCX File, 28 KB - aging_v9i1e83861_app1.docx]

References

1. World Health Organization. World Report on Ageing and Health: World Health Organization; 2015.
2. Gibson PC, Allen PH. Ageing: a biological and psychological perspective. In: Hall CN, editor. Introduction to Biological Psychology: University of Sussex Library; 2023:734-772. [doi: 10.20919/ZDGF9829/18]
3. Ageing and health. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2026-03-16]
4. United Nations, Department of Economic and Social Affairs. World Social Report 2023: Leaving No One Behind in an Ageing World: United Nations; 2023.
5. Decade of healthy ageing: plan of action. World Health Organization. 2020. URL: https://www.who.int/publications/m/item/decade-of-healthy-ageing-plan-of-action [accessed 2026-09-22]
6. Guo J, Huang X, Dou L, et al. Aging and aging-related diseases: from molecular mechanisms to interventions and treatments. Signal Transduct Target Ther 2022 Dec 16;7(1):391. [doi: 10.1038/s41392-022-01251-0] [Medline: 36522308]
7. Gale SA, Acar D, Daffner KR. Dementia. Am J Med 2018 Oct;131(10):1161-1169. [doi: 10.1016/j.amjmed.2018.01.022] [Medline: 29425707]
8. Alzheimer's Association. 2024 Alzheimer's disease facts and figures. Alzheimers Dement 2024 May;20(5):3708-3821. [doi: 10.1002/alz.13809]
9. Dementia. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2026-03-15]
10. Cipriani G, Danti S, Picchi L, Nuti A, Fiorino MD. Daily functioning and dementia. Dement Neuropsychol 2020;14(2):93-102. [doi: 10.1590/1980-57642020dn14-020001] [Medline: 32595877]
11. Prince M, Wimo A, Guerchet M, Ali GC, Wu YT, Prina M. World Alzheimer report 2015: the global impact of dementia – an analysis of prevalence, incidence, cost and trends. : Alzheimer's Disease International; 2015 URL: https://www.alzint.org/resource/world-alzheimer-report-2015/ [accessed 2026-09-22]
12. Kalochristianaki S, Vlotinou P, Bablekos G, et al. Cost estimation analysis of dementia: a scope review. Cureus 2025 May;17(5):e84547. [doi: 10.7759/cureus.84547] [Medline: 40546469]
13. Nichols E, Vos T. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Dec;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8]

XSL•FO
RenderX

14. World failing to address dementia challenge. World Health Organization. 2021. URL: https://www.who.int/news/item/02-09-2021-world-failing-to-address-dementia-challenge [accessed 2026-09-22]
15. Kiselica AM, Hermann GE, Scullin MK, Benge JF. Technology that CARES: enhancing dementia care through everyday technologies. Alzheimers Dement 2024 Dec;20(12):8969-8978. [doi: 10.1002/alz.14192] [Medline: 39508340]
16. Klimova B, Valis M, Kuca K. Exploring assistive technology as a potential beneficial intervention tool for people with Alzheimer's disease - a systematic review. Neuropsychiatr Dis Treat 2018;14:3151-3158. [doi: 10.2147/NDT.S181849] [Medline: 30532546]
17. Global action plan on the public health response to dementia 2017–2025. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/9789241513487 [accessed 2026-09-22]
18. World Health Organization and Alzheimer's Disease International. Dementia: A Public Health Priority: World Health Organization; 2012.
19. Global status report on the public health response to dementia. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240033245 [accessed 2026-09-22]
20. Prince M, Guerchet M, Prina M. The global impact of dementia 2013–2050: policy brief for heads of government. Alzimer's Disease International. 2023. URL: https://www.alzint.org/resource/policy-brief-the-global-impact-of-dementia-2013-2050/ [accessed 2026-09-22]
21. From Plan to Impact VII. Alzheimer's Disease International. 2024. URL: https://www.alzint.org/resource/from-plan-to-impact-vii/ [accessed 2026-10-13]
22. From plan to impact v. Alzheimer's disease international. 2022. URL: https://www.alzint.org/resource/from-plan-to-impact-v/ [accessed 2026-10-05]
23. From plan to impact IV: progress towards targets of the WHO global action plan on dementia. Alzheimer's Disease International. 2021. URL: https://www.alzint.org/resource/from-plan-to-impact-iv/ [accessed 2026-10-05]
24. From plan to impact VI: making every step count. Alzheimer's Disease International. 2023. URL: https://www.alzint.org/resource/from-plan-to-impact-vi/ [accessed 2026-10-05]
25. From plan to impact VIII. Alzheimer's disease international. 2025. URL: https://www.alzint.org/resource/from-plan-to-impact-viii/ [accessed 2026-10-05]
26. World Alzheimer report 2021: journey through the diagnosis of dementia. : Alzheimer's Disease International; 2021 URL: https://www.alzint.org/resource/world-alzheimer-report-2021/ [accessed 2026-09-22]
27. World Alzheimer report 2022: life after diagnosis – navigating treatment, care and support. : Alzheimer's Disease International; 2022 URL: https://www.alzint.org/resource/world-alzheimer-report-2022/ [accessed 2026-09-22]
28. World alzheimer report 2020: design, dignity, dementia – dementia-related design and the built environment. Alzheimer's Disease International. URL: https://www.alzint.org/resource/world-alzheimer-report-2020/ [accessed 2026-10-06]
29. World Alzheimer report 2024: global changes in attitudes to dementia. Alzheimer's Disease International. URL: https://www.alzint.org/resource/world-alzheimer-report-2024/ [accessed 2026-10-06]
30. World Alzheimer report 2025. : Alzheimer's Disease International; 2025 URL: https://www.alzint.org/resource/world-alzheimer-report-2025/ [accessed 2026-09-22]
31. World alzheimer report 2023: reducing dementia risk: never too early, never too late. Alzheimer's Disease International. URL: https://www.alzint.org/resource/world-alzheimer-report-2023/ [accessed 2026-10-06]
32. World alzheimer report 2018: the state of the art of dementia research. Alzheimer's Disease International. URL: https://www.alzint.org/resource/world-alzheimer-report-2018/ [accessed 2026-10-06]
33. World alzheimer report 2019: attitudes to dementia. Alzheimer's Disease International. URL: https://www.alzint.org/resource/world-alzheimer-report-2019/ [accessed 2026-10-06]
34. Jackson J, Prince M. World Alzheimer report 2009. Alzheimer's Disease International. 2009. URL: https://www.alzint.org/resource/world-alzheimer-report-2009/ [accessed 2026-10-10]
35. United Nations. Madrid political declaration and international plan of action on ageing, 2002. Int Soc Sci J 2006 Dec;58(190):633-665. [doi: 10.1111/j.1468-2451.2008.00660.x]
36. World Health Organization. First WHO ministerial conference on global action against dementia: meeting report. : World Health Organization; 2015.
37. Nusberg C. UN takes action on principles for older persons. Ageing Int 1991 Jun;18(1):3-7. [doi: 10.1007/BF03004295]
38. Wallace L, Walsh S, Brayne C. The legacy of the 2013 G8 Dementia Summit: successes, challenges, and potential ways forward. Lancet Healthy Longev 2021 Aug;2(8):e455-e457. [doi: 10.1016/S2666-7568(21)00145-8] [Medline: 36097994]
39. World Assembly on Ageing. World Assembly on Ageing. Aust J Ageing 1982 Sep;1(3):3-11. [doi: 10.1111/j.1741-6612.1982.tb01004.x]
40. United Nations. Report of the world assembly on aging, vienna, 26 july to 6 august 1982. United Nations Digital Library. 1982. URL: https://digitallibrary.un.org/record/36044 [accessed 2026-09-22]
41. World Health Organization. Global Dementia Observatory (GDO).: World Health Organization; 2018. URL: https://www.who.int/data/gho/data/themes/global-dementia-observatory-gdo [accessed 2026-10-06]
42. Ethics and governance of artificial intelligence for health. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240029200 [accessed 2026-09-22]

XSL•FO
RenderX

43. Pan American Health Organization. Dementia in Latin America and the Caribbean: Prevalence, Incidence, Impact, and Trends over Time: Pan American Health Organization; 2023. [doi: 10.37774/9789275126653]
44. Paradela RS, Calandri I, Castro NP, et al. Population attributable fractions for risk factors for dementia in seven Latin American countries: an analysis using cross-sectional survey data. Lancet Glob Health 2024 Oct;12(10):e1600-e1610. [doi: 10.1016/S2214-109X(24)00275-4] [Medline: 39304234]
45. Plan de respuesta a las demencias en la república dominicana 2020–2025. : Ministerio de Salud Pública; 2020 URL: https://www.alzint.org/u/Dominican-Republic-National-Dementia-Plan.pdf [accessed 2026-09-22]
46. Ministerio de Salud de Chile. Plan nacional de demencia 2017. : Ministerio de Salud, Gobierno de Chile; 2017 URL: https://www.minsal.cl/wp-content/uploads/2017/11/PLAN-DE-DEMENCIA.pdf [accessed 2026-09-22]
47. García Peña C, Ortega ML, Castro ST, et al. Plan Nacional de Demencia: Instituto Nacional de Geriatría; 2024.
48. Ministerio de Salud Pública de Cuba. Estrategia Nacional Para El Manejo de La Demencia En Cuba: Ministerio de Salud Pública de Cuba; 2013. URL: https://www.alzint.org/u/Cuba-National-Dementia-Plan.pdf [accessed 2026-10-06]
49. Departamento de Salud. Plan de acción para puerto rico: estrategias para mejorar el cuidado y la calidad de vida de las personas con la enfermedad de Alzheimer. : Departamento de Salud, Puerto Rico; 2015 URL: https://www.alzint.org/u/Puerto-Rico-Plan-de-Accion.pdf [accessed 2026-09-22]
50. Plan nacional para la enfermedad de alzheimer y demencias relacionadas esfuerzos compartidos 2014-2024. : Consejo Nacional de la Persona Adulta Mayor; 2014 URL: https://www.alzint.org/u/Costa-Rica-National-Dementia-Plan.pdf [accessed 2026-09-22]
51. Salud casa por casa. Gobierno de México. 2024. URL: https://programasparaelbienestar.gob.mx/salud-casa-por-casa/ [accessed 2026-09-22]
52. Dementia friends Mexico. World Health Organization. 2024. URL: https://globaldementia.org/en/dementia-friends-mexico [accessed 2026-09-22]
53. Wong R, Michaels-Obregon A, Palloni A. Cohort profile: the Mexican health and aging study (MHAS). Int J Epidemiol 2017 Apr 1;46(2):e2-e2. [doi: 10.1093/ije/dyu263] [Medline: 25626437]
54. Sánchez-García S, García-Peña C, Salvà A, et al. Frailty in community-dwelling older adults: association with adverse outcomes. Clin Interv Aging 2017;12:1003-1011. [doi: 10.2147/CIA.S139860] [Medline: 28721028]
55. Romero-Martínez M, Barrientos-Gutiérrez T, Cuevas-Nasu L, et al. Metodología de la Encuesta Nacional de Salud y Nutrición 2022 y Planeación y diseño de la Ensanut Continua 2020-2024. Salud Publica Mex 2022;64(5):522-529. [doi: 10.21149/14186]
56. ¿Conoces la red mexicana de municipios por la salud? Gobierno de México. 2025. URL: https://www.gob.mx/salud/articulos/crece-coordinacion-por-la-salud-de-los-mexicanos [accessed 2026-10-06]
57. Chen S, Stromer D, Alabdalrahim HA, Schwab S, Weih M, Maier A. Automatic dementia screening and scoring by applying deep learning on clock-drawing tests. Sci Rep 2020;10(1):20854. [doi: 10.1038/s41598-020-74710-9]
58. Lee B, Bang JU, Song HJ, Kang BO. Alzheimer's disease recognition using graph neural network by leveraging image-text similarity from vision language model. Sci Rep 2025;15(1):997. [doi: 10.1038/s41598-024-82597-z]
59. Xue C, Kowshik SS, Lteif D, et al. AI-based differential diagnosis of dementia etiologies on multimodal data. Nat Med 2024 Oct;30(10):2977-2989. [doi: 10.1038/s41591-024-03118-z] [Medline: 38965435]
60. Martin SA, Townend FJ, Barkhof F, Cole JH. Interpretable machine learning for dementia: a systematic review. Alzheimers Dement 2023 May;19(5):2135-2149. [doi: 10.1002/alz.12948]
61. Shao Y, Todd K, Shutes-David A, et al. Identifying probable dementia in undiagnosed Black and White Americans using machine learning in veterans health administration electronic health records. Big Data Cogn Comput 2023 Dec;7(4):167. [doi: 10.3390/bdcc7040167] [Medline: 40963735]
62. Kantayeva G, Lima J, Pereira AI. Application of machine learning in dementia diagnosis: a systematic literature review. Heliyon 2023 Nov;9(11):e21626. [doi: 10.1016/j.heliyon.2023.e21626] [Medline: 38027622]
63. Wang Y, Liu S, Spiteri AG, et al. Understanding machine learning applications in dementia research and clinical practice: a review for biomedical scientists and clinicians. Alzheimers Res Ther 2024 Aug 1;16(1):175. [doi: 10.1186/s13195-024-01540-6] [Medline: 39085973]
64. Butler PM, Yang J, Brown R, et al. Smartwatch- and smartphone-based remote assessment of brain health and detection of mild cognitive impairment. Nat Med 2025 Mar;31(3):829-839. [doi: 10.1038/s41591-024-03475-9] [Medline: 40038507]
65. Atkins AS, Khan A, Ulshen D, et al. Assessment of instrumental activities of daily living in older adults with subjective cognitive decline using the virtual reality functional capacity assessment tool (VRFCAT). J Prev Alzheimers Dis 2018;5(4):216-234. [doi: 10.14283/jpad.2018.28] [Medline: 30298179]
66. Mizukami K, Taguchi M, Kouketsu T, et al. A cognitive function test utilizing eye tracking technology in virtual reality. Nihon Ronen Igakkai Zasshi 2023;60(1):43-50. [doi: 10.3143/geriatrics.60.43] [Medline: 36889722]
67. Klil-Drori S, Bodenstein KC, Sun S, et al. Montreal Cognitive Assessment (MoCA) XpressO: validation of a digital self-administered cognitive prescreening tool. J Am Geriatr Soc 2024 Aug;72(8):2516-2522. [doi: 10.1111/jgs.18902] [Medline: 38558263]

XSL•FO
RenderX

68. Sum G, Lau LK, Jabbar KA, et al. The World Health Organization (WHO) integrated care for older people (ICOPE) framework: a narrative review on its adoption worldwide and lessons learnt. Int J Environ Res Public Health 2022 Dec 22;20(1):154. [doi: 10.3390/ijerph20010154] [Medline: 36612480]
69. Shin HT, Kim DY, Bae CR, et al. Fully-immersive virtual reality instrumental activities of daily living training for mild dementia: a feasibility study. Ann Palliat Med 2023 Mar;12(2):280-290. [doi: 10.21037/apm-22-847] [Medline: 37019637]
70. Hiramatsu T, Kamei M, Inoue D, Kawamura A, An Q, Kurazume R. Development of dementia care training system based on augmented reality and whole body wearable tactile sensor. Presented at: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); Oct 24, 2020 to Jan 24, 2021; Las Vegas, NV p. 4148-4154 URL: https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=9340668 [accessed 2026-09-22] [doi: 10.1109/IROS45743.2020.9341039]
71. Kim JH, Park S, Lim H. Developing a virtual reality for people with dementia in nursing homes based on their psychological needs: a feasibility study. BMC Geriatr 2021 Mar 7;21(1):167. [doi: 10.1186/s12877-021-02125-w] [Medline: 33678160]
72. Matsangidou M, Solomou T, Frangoudes F, Papayianni E, Kkeli N, Pattichi CS. "Transported to a better place": The influence of virtual reality on the behavioural and psychological symptoms of dementia. Int J Hum Comput Stud 2025 Feb;196:103413. [doi: 10.1016/j.ijhcs.2024.103413]
73. McCarron HR, Zmora R, Gaugler JE. A web-based mobile app with a smartwatch to support social engagement in persons with memory loss: pilot randomized controlled trial. JMIR Aging 2019 Jun 18;2(1):e13378. [doi: 10.2196/13378] [Medline: 31518270]
74. Iwabuchi Y, Sato I, Fujino Y, Yagi N. The communication supporting robot based on "humanitude" concept for dementia patients. Presented at: 2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech); Mar 12-14, 2019; Osaka, Japan p. 219-223. [doi: 10.1109/LifeTech.2019.8884049]
75. Sharma S, Dudeja RK, Aujla GS, Bali RS, Kumar N. DeTrAs: deep learning-based healthcare framework for IoT-based assistance of Alzheimer patients. Neural Comput Appl 2025 Oct;37(28):23017-23029. [doi: 10.1007/s00521-020-05327-2]
76. Goerss D, Köhler S, Rong E, et al. Smartwatch-based interventions for people with dementia: user-centered design approach. JMIR Aging 2024 Jun 7;7:e50107. [doi: 10.2196/50107] [Medline: 38848116]
77. Aujla GS, Jindal A, Chaudhary R, et al. DLRS: deep learning-based recommender system for smart healthcare ecosystem. 2020 Presented at: ICC 2019 - 2019 IEEE International Conference on Communications (ICC); May 20-24, 2019. [doi: 10.1109/ICC.2019.8761416]
78. Ebrahem AT, Mawfaq Mohamedsheet Al-Hatab M, Al-Jabbar EYA, Alkhaled WH, Al-Sawaff ZH. Using IoT technology for monitoring Alzheimer's and elderly patients. Indonesian J Electr Eng Comput Sci 2023;31(2). [doi: 10.11591/ijeecs.v31.i2.pp986-994]
79. Chávez A, Borrego G, Gutierrez-Garcia JO, Rodríguez LF. Design and evaluation of a mobile application for monitoring patients with Alzheimer's disease: a day center case study. Int J Med Inform 2019 Nov;131:103972. [doi: 10.1016/j.ijmedinf.2019.103972] [Medline: 31563010]
80. Löbe C, AboJabel H. Empowering people with dementia via using intelligent assistive technology: a scoping review. Arch Gerontol Geriatr 2022;101:104699. [doi: 10.1016/j.archger.2022.104699] [Medline: 35413610]
81. Ueno K, Tanaka H, Niki K, et al. Effects of real-time VR clinical practice on reducing the stigma toward dementia among students of occupational therapy: a randomized controlled trial. PCN Rep 2023 Dec;2(4):e160. [doi: 10.1002/pcn5.160] [Medline: 38868728]
82. Tochibayashi N, Ota M. World economic forum. Japan is leveraging digital solutions to tackle dementia. 2024. URL: https://www.weforum.org/stories/2024/11/japan-is-leveraging-digital-solutions-to-tackle-dementia/ [accessed 2026-09-22]
83. Ding H, Mandapati A, Hamel AP, et al. Multimodal machine learning for 10-year dementia risk prediction: the Framingham heart study. J Alzheimers Dis 2023;96(1):277-286. [doi: 10.3233/JAD-230496] [Medline: 37742648]
84. Sun H, Milton S, Fang Y, et al. Machine learning-based sleep electroencephalographic brain age index and dementia risk: an individual participant data meta-analysis. JAMA Netw Open 2026 Mar 2;9(3):e261521. [doi: 10.1001/jamanetworkopen.2026.1521] [Medline: 41854616]
85. Lim EH, Kim DS, Won YH, et al. Effects of home based serious game training (Brain TalkTM) in the elderly with mild cognitive impairment: randomized, a single-blind, controlled trial. Brain Neurorehabil 2023 Mar;16(1):e4. [doi: 10.12786/bn.2023.16.e4] [Medline: 37033004]
86. Kang S, Lim JI, Stenzel L, et al. Efficacy and safety of mobile app-based metamemory cognitive training for mild cognitive impairment: multicenter randomized clinical trial. JMIR Mhealth Uhealth 2026 Jan 19;14:e73464. [doi: 10.2196/73464] [Medline: 41554532]
87. Puig-Davi A, Moreu-Valls A, Aracil-Bolaños I, et al. Effects of immersive virtual reality and adaptive cognitive training on cognition and function in Parkinson's disease mild cognitive impairment: a randomized clinical trial. Sci Rep 2026 Jan 25;16(1):3499. [doi: 10.1038/s41598-025-31833-1] [Medline: 41582160]
88. Lusk JB, Choi S, Clark AG, et al. Dementia and Parkinson's disease diagnoses in electronic health records vs. Medicare claims data: a study of 101,980 linked patients. BMC Neurol 2023 Sep 12;23(1):325. [doi: 10.1186/s12883-023-03361-w] [Medline: 37700254]

XSL•FO
RenderX

89. Harding BN, Floyd JS, Scherrer JF, et al. Methods to identify dementia in the electronic health record: comparing cognitive test scores with dementia algorithms. Healthc (Amst) 2020 Jun;8(2):100430. [doi: 10.1016/j.hjdsi.2020.100430] [Medline: 32553526]
90. Sengar SS, Hasan AB, Kumar S, Carroll F. Generative artificial intelligence: a systematic review and applications. Multimed Tools Appl 2024;84(21):23661-23700. [doi: 10.1007/s11042-024-20016-1]
91. Nance DC, Rivero May MI, Flores Padilla L, Moreno Nava M, Deyta Pantoja AL. Faith, work, and reciprocity: listening to Mexican men caregivers of elderly family members. Am J Mens Health 2018 Nov;12(6):1985-1993. [doi: 10.1177/1557988316657049] [Medline: 27352926]
92. CUCBA ofrece pulseras de localización para personas con demencia o alzheimer de inicio temprano. Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. 2025. URL: https://cucba.udg.mx/noticia/cucba-ofrece-pulseras-de-localizacion-para-personas-con-demencia-o-alzheimer-de-inicio [accessed 2026-07-22]
93. Con brazalete, zapopan garantiza que adultos mayores vuelvan a casa. Gobierno Municipal de Zapopan. 2023. URL: https://www.zapopan.gob.mx/v3/noticias/con-brazalete-zapopan-garantiza-que-adultos-mayores-vuelvan-casa [accessed 2026-09-22]
94. Estudio de dispositivos de geolocalización 2024 [archived version]. : Secretaría de Igualdad Sustantiva entre Mujeres y Hombres (SISEMH), Gobierno del Estado de Jalisco; 2024 URL: https://web.archive.org/web/20250719183534/https://igualdad.jalisco.gob.mx/wp-content/uploads/2024/12/Estudio-Dispositivos-de-Geolocalizacion-2024.pdf [accessed 2026-09-22]
95. De la Peña-López IJ, Acosta-Gonzaga E. Adoption of technology in older adults in Mexico city: an approach from the technology acceptance model. Brain Sci 2025 Jun 12;15(6):632. [doi: 10.3390/brainsci15060632] [Medline: 40563802]
96. Castro JW, Madrigal GA. Uso de la tecnología por las personas mayores: un estudio exploratorio. 2024 Mar Presented at: Décima Cuarta Conferencia Iberoamericana de Complejidad, Informática y Cibernética; Conferencia Virtual p. 137-142. [doi: 10.54808/CICIC2024.01.137]
97. Crivelli L, Calandri IL, Suemoto CK, et al. Latin American initiative for lifestyle intervention to prevent cognitive decline (LatAm‐FINGERS): study design and harmonization. Alzheimers Dement 2023 Sep;19(9):4046-4060. [doi: 10.1002/alz.13101]
98. Loza Tello CE. Análisis de la accesibilidad al altavoz inteligente alexa en personas adultas mayores. Revista de estudios interdisciplinarios del arte, diseño y la cultura. 2023. URL: https://masam.cuautitlan.unam.mx/seminarioarteydiseno/revista/index.php/reiadyc/article/view/158 [accessed 2026-09-22]
99. Norma oficial mexicana NOM-004-SSA3-2012, del expediente clínico. Secretaría de Salud. 2012. URL: https://platiica.economia.gob.mx/normalizacion/nom-004-ssa3-2012/ [accessed 2026-10-06]
100. Sistemas de información de registro electrónico para la salud intercambio de información en salud. Secretaría de Salud. 2012. URL: https://platiica.economia.gob.mx/normalizacion/nom-024-ssa3-2012/ [accessed 2026-10-06]
101. Rocha FOD, López MDR, González M, Hernández CV. Impact of digital inclusion on older adults in Mexico. Rev Gestão Soc Ambient 2024;18(10):e08785. [doi: 10.24857/rgsa.v18n10-034]
102. Calderón EAA, Torre MDL. Design of eHealth systems and devices for monitoring older adults in Mexico. S F J of Dev 2025;6(10):e5851. [doi: 10.46932/sfjdv6n10-009]
103. Decreto por el que se reforman y adicionan diversas disposiciones de la ley general de salud en materia de salud digital. Diario Oficial de la Federación. 2026. URL: https://sidof.segob.gob.mx/notas/5778298 [accessed 2026-10-06]
104. Gómez-Dantés O, Flamand L, Cerecero-García D, Morales-Vazquez M, Serván-Mori E. Origin, impacts, and potential solutions to the fragmentation of the Mexican health system: a consultation with key actors. Health Res Policy Syst 2023 Jul 31;21(1):80. [doi: 10.1186/s12961-023-01025-2] [Medline: 37525130]

Abbreviations

AD: Alzheimer disease
ADI: Alzheimer's Disease International
AT: assistive technology
DIF: Sistema Nacional para el Desarrollo Integral de la Familia
ENSANUT: Encuesta Nacional de Salud y Nutrición
FEDMA: Federación Mexicana de Alzheimer
GDO: Global Dementia Observatory
GenAI : generative AI
IMSS: Instituto Mexicano del Seguro Social
INAPAM: Instituto Nacional de las Personas Adultas Mayores
ISSSTE: Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado
LatAm-FINGERS: Latin American Initiative for Lifestyle Intervention to Prevent Cognitive Decline
Mex-Cog: Cognitive Aging Ancillary Study
MHAS: Mexican Health and Aging Study
MMSE: Mini-Mental State Examination

XSL•FO
RenderX

NGO: nongovernmental organization
PAHO: Pan American Health Organization
SECIHTI: Secretariat of Science, Humanities, Technology, and Innovation
WHO: World Health Organization

Edited by M O'Connell; submitted 19.Sep.2025; peer-reviewed by G Gibson, H Wang; revised version received 03.Aug.2026; accepted 26.Aug.2026; published 22.Oct.2026.

Please cite as:
Camacho-Bustamante LM, Astudillo-García CI, Velázquez López G, Moncayo-Samperio LG, Palacios-García AA, Navarro-Tuch SA Global Frameworks Toward Effective Integration of Technology in National Dementia Strategies: Narrative Review and Case Study in Mexico
JMIR Aging 2026;9:e83861
URL: https://aging.jmir.org/2026/1/e83861
doi:10.2196/83861

© Lili Marlene Camacho-Bustamante, Claudia I Astudillo-García, Guillermo Velázquez López, Luis Gerardo Moncayo-Samperio, Alberto Agustín Palacios-García, Sergio Alberto Navarro-Tuch. Originally published in JMIR Aging (https://aging.jmir.org), 22.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparative Effectiveness of Digital Exercise-Based Interventions for Improving Balance in Older Adults: Systematic Review and Network Meta-Analysis With GRADE Certainty of Evidence

Fangbo Li[1*], PhD; Xuecheng Li[1*], PhD; Qingfang Li[2], MSc; Xiaoqin Gong[1], PhD; Kexin Shi[1], PhD; Yongyu Huang[1], PhD; Xiaowei Feng[3], PhD; Dong Zhang[4], PhD

[1]School of Physical Education and Sports Science, South China Normal University, Guangzhou, Guangdong, China

[2]Guangzhou Sport University, Guangzhou, Guangdong, China

[3]School of Football, Hainan Normal University, Haikou, Hainan, China

[4]College of Physical Education, Quanzhou Normal University, Main Campus: No. 398, Donghai Street, Fengze District, Quanzhou, Fujian, China

[*]these authors contributed equally

Corresponding Author:
Dong Zhang, PhD
College of Physical Education, Quanzhou Normal University, Main Campus: No. 398, Donghai Street, Fengze District, Quanzhou, Fujian, China

Abstract

Background: Technology-based rehabilitation is increasingly used for balance training in older adults. A 2023 network meta-analysis (NMA) found only exergaming with motion capture superior to no treatment; however, the evidence has since expanded, and no GRADE (Grading of Recommendations Assessment, Development, and Evaluation) appraisal has been applied.

Objective: This study aimed to update the comparative effectiveness of 5 technology categories—exergaming with balance platforms (EBP), exergaming with motion capture (EMC), other serious games, interventions with wearables, and telerehabilitation (Tele)—versus conventional exercise (CE), and no treatment (NT), applying GRADE adapted for NMA.

Methods: A systematic search of 4 databases from inception through April 2025, with an identical supplementary update search to January 31, 2026, identified 60 randomized controlled trials (N=3469; mean age 73.1, SD 5.2; range 63.3-86.2 y). Fifty-seven trials (2758 analyzed participants) contributed to 4 prespecified networks (all balance, functional scales, Timed Up and Go [TUG] test, and gait speed). A random-effects NMA was fitted using multivariate meta-regression. Treatment hierarchies were summarized by the surface under the cumulative ranking curve (SUCRA) and the probability of being best; certainty of evidence was rated with GRADE adapted for NMA.

Results: In the all-balance network, Tele (standardized mean difference [SMD] 0.69, 95% CI 0.08 - 1.29; *P*=.03), EBP (SMD 0.63, 95% CI 0.31 - 0.95; *P*<.001), EMC (SMD 0.58, 95% CI 0.22 - 0.95; *P*=.002), and CE (SMD 0.35, 95% CI 0.03 - 0.68; *P*=.03) were superior to NT, with consistent findings in the TUG network (EBP 0.81, 95% CI 0.42 - 1.20; Tele 0.82, 95% CI 0.16 - 1.49; EMC 0.79, 95% CI 0.36 - 1.21; all *P*<.05). Tele ranked first by SUCRA (76.3%), but its probability of being best was only approximately 41%, and the estimate rested on 5 trials (n=198). No technology-based intervention was superior to CE in any network. Certainty of evidence was moderate for EBP versus NT and EMC versus NT in the TUG network only; the 37 other comparisons were low or very low. Dropout did not differ between technology-based and control arms (odds ratio 1.10, 95% CI 0.87 - 1.40; *P*=.41).

Conclusions: Tele, EBP, and EMC were each associated with improved balance versus NT, although none exceeded CE. The first-place SUCRA ranking of Tele is preliminary. Certainty was moderate only for EBP and EMC versus NT on the TUG test; the remaining evidence is too uncertain for firm clinical recommendations. As no included trial reported falls, these findings are not direct evidence of fall prevention. Larger trials with standardized comparator intensity and posttraining follow-up are required.

Trial Registration: PROSPERO CRD420261380567; https://www.crd.york.ac.uk/PROSPERO/view/CRD420261380567

(JMIR Aging 2026;9:e101231) doi:10.2196/101231

KEYWORDS

network meta-analysis; technology-based rehabilitation; older adults; balance; accidental falls; telerehabilitation; Grading of Recommendations Assessment, Development, and Evaluation; GRADE approach

XSL•FO
RenderX

Introduction

Background

Population aging is a major global health challenge. The global population aged ≥60 years is projected to surpass 2.2 billion by the late 2070s [1]. Falls are among the most common threats to health and independence in later life. The World Health Organization estimates that between 28% and 35% of adults aged ≥65 years experience at least one fall per year, a rate that rises to 32% to 42% among those aged >70 years [2]. The Global Burden of Disease Study 2021 documents approximately 45.7 million new falls among older adults annually, with an estimated 37.3 million requiring medical attention [3]. Falls also increase the risk of mortality and reduce quality of life through fear of falling, functional decline, and institutionalization [4].

Exercise is the best-supported intervention for fall prevention in older adults. The 2022 World Guidelines for Falls Prevention and Management assign a strong recommendation to multicomponent exercise programs targeting balance, strength, and functional mobility [4]. However, adherence to conventional exercise (CE) programs is often poor, and motivation, access, and travel to clinical settings are common barriers [5]. Technology-based approaches have attracted interest because they may improve engagement and reach people who cannot attend in person.

The rationale for technology-based rehabilitation derives largely from augmented sensory feedback. Augmented feedback (ie, extrinsic information about movement outcomes) supports motor learning and postural adaptation [6]. In balance training, real-time feedback on postural sway promotes lasting changes in motor control by facilitating corrective movement strategies [7]. Digital technologies deliver this feedback through several distinct modalities.

Exergaming with balance platforms (EBP) provides real-time feedback on center-of-pressure displacement via devices such as the Nintendo Wii Balance Board. Exergaming with motion capture (EMC) captures whole-body kinematics to generate on-screen visual feedback, as implemented via the Microsoft Kinect. Other serious games (OSG) are purpose-built rehabilitation software increasingly delivered through head-mounted displays. Interventions with wearables (IW) provide haptic or auditory postural feedback via body-worn sensors. Telerehabilitation (Tele) encompasses remotely supervised exercise programs delivered via digital communication platforms, with uptake accelerating substantially after the COVID-19 pandemic [8].

Prior Work and Rationale for an Update

Most previous systematic reviews and meta-analyses have examined each technology category separately, so their relative effectiveness and their place within fall prevention programs remain unclear [9,10]. Network meta-analysis (NMA) combines direct and indirect evidence in a single model and can rank treatments even where head-to-head comparisons are absent from the primary literature [11]. Cieślik et al [12] published the first NMA in this field in 2023 and found that EMC was the only technology significantly superior to no treatment (Timed Up and Go [TUG]: standardized mean difference [SMD] −0.70, 95% CI −1.16 to −0.23), while Tele, serious games, and wearable interventions lacked sufficient evidence for definitive conclusions.

The evidence base has expanded substantially since that review's search cutoff in June 2023. Newer randomized controlled trials (RCTs) have tested Tele protocols in larger samples and over longer periods than the trials available to the original review [12], head-mounted virtual reality systems have been studied more widely as platforms for serious games, and wearable sensors now support home-based programs [13,14]. The original NMA also did not apply GRADE (Grading of Recommendations Assessment, Development, and Evaluation), so no formal certainty rating exists for any pairwise comparison [12]. An updated synthesis is therefore warranted to incorporate the newer evidence and to formally assess the certainty of evidence.

Objectives

We aimed to (1) quantify the comparative effectiveness of 5 categories of technology-based rehabilitation intervention (EBP, EMC, OSG, IW, and Tele) relative to CE and no treatment (NT) across 4 prespecified outcome networks (all balance outcomes, functional scales, TUG, and gait speed); (2) establish a probabilistic treatment hierarchy using the surface under the cumulative ranking curve (SUCRA) and the probability of being the best treatment; and (3) appraise the certainty of evidence for each comparison using the GRADE framework adapted for NMA [15].

Methods

Study Design and Registration

This study was designed as a systematic review and NMA, conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement and its NMA extension (PRISMA-NMA) [16,17]. The completed PRISMA 2020 and PRISMA-NMA checklists are provided in Checklists 1 and 2, respectively. The protocol was registered with PROSPERO (International Prospective Register of Systematic Reviews; CRD420261380567) and constitutes an updated and expanded synthesis of the original review by Cieślik et al [12].

As the PROSPERO record was completed after the primary searches had been executed, this review is reported as a retrospectively registered update rather than as a prospectively registered de novo review. No eligibility criterion, outcome definition, or analytical decision was altered after the authors had inspected the effect estimates. Two departures from the registered protocol occurred and are reported here in full. First, the protocol specified the *netmeta* R package with P-score ranking, whereas the primary analysis was implemented in Stata using SUCRA; SUCRA and the netmeta P-score are numerically near-equivalent frequentist ranking metrics, and the treatment hierarchy was unaffected by this change. Second, the meta-regressions of training dose and of CE intensity, the treatment-ranking probabilities, and the quantitative transitivity summary were added in response to peer review; these are

XSL•FO
RenderX

reported throughout as post hoc exploratory analyses and are interpreted accordingly.

Literature Search

A systematic search was conducted across 4 electronic databases (PubMed, Embase, Cochrane Central Register of Controlled Trials, and Web of Science) from inception through April 2025. Database-specific search strategies combined controlled vocabulary terms with free-text synonyms covering older adult populations (aged ≥60 y), technology-based interventions (virtual reality, exergaming, telerehabilitation, and wearables), and balance and functional mobility outcomes. A randomized-design filter was applied in PubMed, Embase, and Web of Science; the Cochrane Central Register of Controlled Trials contains only trial records and required none. Each strategy also carried an exclusion block to remove records indexed with the neurological, neurodegenerative, orthopedic, oncological, and other comorbid conditions listed under eligibility, as well as reviews, meta-analyses, and study protocols. The complete search strings for each database, with the date each was executed, are provided in Table S1 in Multimedia Appendix 1.

As peer review and revision extended into 2026, an identical supplementary update search was executed across all 4 databases on January 31, 2026, to capture records that were published, or made available online ahead of print, after the primary search date. Screening and data extraction for the supplementary search followed the same procedures as for the primary search. The record counts, screening numbers, and PRISMA flow diagram reported in the "Results" section represent the combined yield of the primary and supplementary searches. The supplementary search contributed 4 additional eligible trials [18-21], all of which were incorporated into the quantitative synthesis; no study identified after the primary search was used for contextual discussion only. One of these trials appears with a 2026 citation year because it was assigned to a 2026 print issue after online publication [18]; it was retrieved by the supplementary search.

Eligibility Criteria and Study Selection

Studies were eligible if they (1) enrolled community-dwelling or facility-dwelling adults aged ≥60 years without neurodegenerative, neurological, major orthopedic, or oncological comorbidities; (2) examined an eligible technology-based rehabilitation intervention with at least one control group; (3) reported at least one prespecified primary outcome; and (4) adopted a randomized design. Studies were excluded if the intervention involved robotic exoskeletons or if data were reported only as conference abstracts without peer review. Two independent reviewers (FL and XL) screened all records, and disagreements were resolved through discussion with a third reviewer (DZ). The number of full-text articles excluded at the eligibility stage, broken down by reason for exclusion, is provided in Table S2 in Multimedia Appendix 1.

Intervention Classification

Eligible technology-based interventions were classified into 5 mutually exclusive categories following the taxonomy of Cieślik et al [12]: (1) EBP, real-time center-of-pressure feedback devices (eg, Wii Balance Board); (2) EMC, systems capturing full-body kinematics for on-screen visual feedback (eg, Microsoft Kinect); (3) OSG, rehabilitation software not relying on balance platforms or motion capture; (4) IW, exercise programs using wearable sensor feedback; and (5) Tele, remotely supervised exercise delivered via digital communication platforms. Two control categories were designated: CE and NT.

Classification was performed independently by 2 reviewers for every arm of every included trial, with disagreements resolved by discussion and, where necessary, by a third reviewer. Where an intervention displayed features of more than 1 category, assignment followed the dominant source of augmented feedback: interventions delivering feedback derived primarily from center-of-pressure displacement were assigned to EBP even when delivered remotely, whereas interventions whose defining feature was synchronous remote supervision were assigned to Tele. Arms combining a technology with CE in a factorial manner (2 arms in 2 trials) could not be assigned to a single node and were excluded from the networks, while the remaining arms of those trials were retained. The as-extracted description and resulting node assignment for every arm of every included trial are reported in Table S3 in Multimedia Appendix 1, so that each classification decision can be checked.

Outcome Measures and Network Structure

Outcomes were prespecified into 4 analytical networks: (1) network 1 (all balance), aggregating results from all eligible balance-related instruments, including the TUG, Berg Balance Scale (BBS), Mini-Balance Evaluation Systems Test (Mini-BESTest), Short Physical Performance Battery (SPPB), Fullerton Advanced Balance Scale (FAB), Tinetti Performance-Oriented Mobility Assessment (POMA), Functional Reach Test, single-leg stance, center-of-pressure (COP) parameters, and sit-to-stand tests; (2) network 2 (functional scales), restricted to multidimensional clinical scales (BBS, Mini-BESTest, SPPB, FAB, and Tinetti/POMA); (3) network 3 (TUG), the TUG test and its equivalent, the 8-Foot Up-and-Go test; and (4) network 4 (gait speed), quantitative gait speed from 4- or 10-m walk tests. Where a trial reported more than 1 instrument eligible for network 1, a single instrument per trial was selected using a prespecified hierarchy (TUG, followed by BBS, Mini-BESTest, SPPB, and other instruments), so that no trial contributed more than 1 estimate to any network. The secondary outcome was participant acceptability, defined as the all-cause dropout rate.

Network 1 is a composite of instruments that measure different constructs (dynamic functional mobility, multidimensional clinical rating scales, static postural sway, and lower-limb function). It was retained because it maximizes the evidence contributing to each comparison, but it is reported alongside, and not in preference to, the construct-specific networks; network 3 (TUG) provides the most clinically interpretable single-construct estimates and is the basis for the principal clinical conclusions of this review.

Data Extraction

Data were extracted by 1 reviewer and independently verified by a second reviewer. For each eligible arm, we recorded the number of participants analyzed for the outcome of interest,

XSL•FO
RenderX

together with the number randomized, the postintervention mean, and the postintervention SD. Analyzed rather than randomized denominators were used in the effect-size calculations because postintervention means and SDs are reported for analyzed participants; the 2 denominators differ appreciably in a minority of trials (most notably in the largest included trial [22], in which 503 participants were randomized and 226 contributed postintervention balance data), and both are reported per trial in Table S4 in Multimedia Appendix 1. Postintervention means and SDs were used throughout, as combining posttest and change-score SMDs within a single analysis is not recommended when SMD is the effect measure [23]. Studies reporting only change-from-baseline scores without accompanying postintervention values were therefore excluded from network synthesis. For studies reporting only median and IQR, values were estimated using the formula of Wan et al [24]. Dropout counts were extracted per randomized arm.

We also sought to extract adverse events, technology-related difficulties (eg, device malfunction, connectivity failure, and cybersickness), and participant-reported satisfaction or acceptability to characterize the feasibility and user burden of each modality. These data were too sparsely reported to be synthesized and are summarized descriptively in the "Results" section.

Risk of Bias Assessment

The methodological quality of included trials was evaluated by 2 independent reviewers using version 2 of the Cochrane Risk of Bias tool (RoB 2) [25], covering 5 domains: randomization (D1), deviations from intended interventions (D2), missing outcome data (D3), outcome measurement (D4), and selective reporting (D5). Overall risk was classified as low risk, some concerns, or high risk. As blinding of participants is structurally impossible in exercise trials, an elevated D2 rating attributable solely to the absence of participant blinding was not by itself treated as a reason to rate down for risk of bias in the GRADE assessment. Instead, following the reasoning that the relevant question is whether the absence of blinding plausibly influenced the result [15], D2 was treated as a source of serious risk of bias only where the trial reports also indicated differential cointerventions, differential attention or contact time, or an intention-to-treat principle that was not followed; D4 was rated based on the presence of blinded outcome assessment; and D5 was rated based on the availability of a prespecified protocol or registration. Domain-level judgments for every trial are provided in Table S5 and Figure S1 in Multimedia Appendix 1.

Statistical Analysis

The SMD (Hedges *g*) and its SE were calculated for each pairwise arm comparison from postintervention means and pooled postintervention SDs. For outcomes where lower scores indicate better performance (TUG, COP, and sit-to-stand time), the SMD sign was reversed to ensure consistent positive directionality across all outcomes.

All analyses were performed in Stata/SE (version 15.1; StataCorp LLC) using the network suite [26]. Data were structured in the augmented format to accommodate multiarm trials [27], so that the correlation between effect estimates from the same trial was respected and no participant was counted more than once within a comparison; 8 trials contributed 3 arms each to network 1, and the remainder contributed 2. The NMA was implemented as a multivariate meta-analysis via the mvmeta command [28], fitting a random-effects consistency model with restricted maximum likelihood estimation. The between-study SD (τ) was assumed common across comparisons, and between-study covariances were modeled with an exchangeable correlation structure (ρ=0.5) [26,27]. NT was specified as the reference category.

Transitivity was evaluated by comparing the distribution across treatment nodes of mean age, intervention duration, training frequency, total sessions, session length, and study setting (community vs facility dwelling), summarized in Table S6 and Figure S2 in Multimedia Appendix 1. Modifiers reported too inconsistently to evaluate—baseline balance impairment, fall risk or history, frailty, sex distribution, supervision level, comparator intensity, prior technology familiarity, and outcome measurement protocol—are identified as a limitation of the transitivity appraisal. Between-study heterogeneity was quantified by τ and τ^2. Global statistical inconsistency was assessed using the design-by-treatment interaction model. Local inconsistency for each pairwise contrast was evaluated by node splitting [29-31], decomposing each estimate into its direct and indirect components and testing their agreement.

Treatment hierarchies were estimated by computing SUCRA from 5000 parametric bootstrap draws. SUCRA summarizes a rank distribution and carries no information about the magnitude or the certainty of the underlying effects, so the probability that each intervention was the best treatment was also estimated from the multivariate normal distribution of the network estimates [11] and is reported alongside SUCRA throughout. The post hoc analyses added during peer review—the random-effects meta-regressions of training dose and of CE intensity, the ranking probabilities, and the quantitative transitivity summary—were conducted in Python (version 3.12; Python Software Foundation) using numpy, scipy, and pandas, with figures produced in Matplotlib, and each was cross-checked for numerical agreement against the primary Stata NMA. In the dose meta-regressions, study-level SMDs were regressed on intervention duration, training frequency, total sessions, session length, weekly training minutes, and mean age; the direct CE versus NT effect was separately regressed on CE dose parameters.

Small-study effects were assessed visually using comparison-adjusted funnel plots [32] generated via netfunnel. Formal regression-based testing of funnel plot asymmetry was not undertaken because no comparison in any network was informed by 10 or more trials conventionally regarded as the minimum for such a test to be interpretable; the consequences of this limitation for the assessment of publication bias are considered in the "Discussion" section. As these methods have very low power when fewer than 10 trials contribute to a comparison, they were interpreted as exploratory in networks 2 and 4, and no comparison was rated down for publication bias on the basis of funnel plot appearance alone.

XSL•FO
RenderX

A post hoc sensitivity analysis examined the source of the global inconsistency in network 3 (TUG), where 4 trials provided direct NT versus CE evidence. As any exclusion threshold is arbitrary, 2 specifications were run, and both are reported: specification 1 removed the 2 trials whose direct effect exceeded SMD 2.0, and specification 2 removed the 3 trials exceeding SMD 1.0. All 4 are 3-arm trials, so removal also withdraws technology arms; the effect of each specification on every estimate, every SUCRA value, and the direct-indirect discrepancy is reported in full (Table S7, Multimedia Appendix 1), with ranking stability assessed by the Spearman correlation of SUCRA before and after removal. These analyses characterize the influence of the discrepant contrast; they were not used to select a model, and the primary analysis retains all trials.

GRADE certainty was rated for all pairwise comparisons using the framework adapted for NMA [15], starting at high certainty. The following decision rules were fixed before rating and applied uniformly. Imprecision was rated down 1 level where the 95% CI included the null or a trivial effect (|SMD|<0.20), and 2 levels where it additionally spanned appreciable benefit and harm (|SMD|>0.50) or fewer than 3 trials contributed. In network 1, imprecision was judged against the 95% prediction interval rather than the CI because the heterogeneity of a composite outcome makes the CI understate the uncertainty that applies to any single instrument; all network 1 prediction intervals included the null. Inconsistency was rated down where the design-by-treatment interaction test was significant and node splitting localized the discrepancy to that comparison. Indirectness was not rated down anywhere because all trials were randomized comparisons in the target population using the intended outcomes. Publication bias was rated down only where funnel plot asymmetry coincided with a plausible mechanism and at least 10 contributing trials. Two reviewers rated independently and reached a final rating by consensus.

Results

Study Selection and Characteristics

The primary and supplementary searches together yielded 2459 records (PubMed: n=408, 16.6%; Embase: n=775, 31.5%; Web of Science: n=491, 20%; and Cochrane: n=785, 31.9%), of which 1213 (49.3%) were duplicates. After title and abstract screening of 1246 unique records, 188 (15.1%) full-text articles were assessed for eligibility; 128 (68.1%) were excluded (incorrect study type, n=21, 16.4%; incorrect outcomes, n=62, 48.4%; incorrect intervention, n=26, 20.3%; incorrect participants, n=19, 14.8%). Sixty RCTs [18-22,33-87] involving 3469 randomized participants met all inclusion criteria (Figure 1). Four of these trials [18-21] were identified by the supplementary update search.

Fifty-seven (95%) of the 60 included trials (randomized: n=3301; and analyzed arms: n=2758) contributed data to at least one of the 4 networks. Three trials could not contribute quantitative data; they were retained in the systematic review but not included in any network: 1 reported outcomes as change scores only [37]; 1 reported balance solely as a composite *z* score that could not be mapped onto any single network instrument [67]; and 1 randomized participants to 2 arms that were both classified into the same network node, so that it contributed no between-node contrast [56]. All 3 are described in Table S4 in Multimedia Appendix 1.

Ten included trials [18-21,43,68,71,76,78,86] were published, or appeared online, after the June 2023 search cutoff of the original review [12]. Sample sizes ranged from 12 to 503 randomized participants (median 43, IQR 34-64), and the mean age was 73.1 (SD 5.2; range 63.3 - 86.2) years (unweighted mean of trial means). EBP (21 arms) and EMC (18 arms) were the most frequently studied categories in the all-balance network (Table 1). Tele trials were much larger on average (mean 124 randomized participants across the 6 Tele trials); although the mean number of participants contributing to each Tele arm in network 1 was 39.6, the largest Tele trial having supplied postintervention data for only a subset of those randomized [22]. Participant characteristics and intervention details for all included studies are summarized in Table S4 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of study identification, screening, and inclusion, showing the combined yield of the primary search (to April 2025) and the supplementary update search (to January 31, 2026).

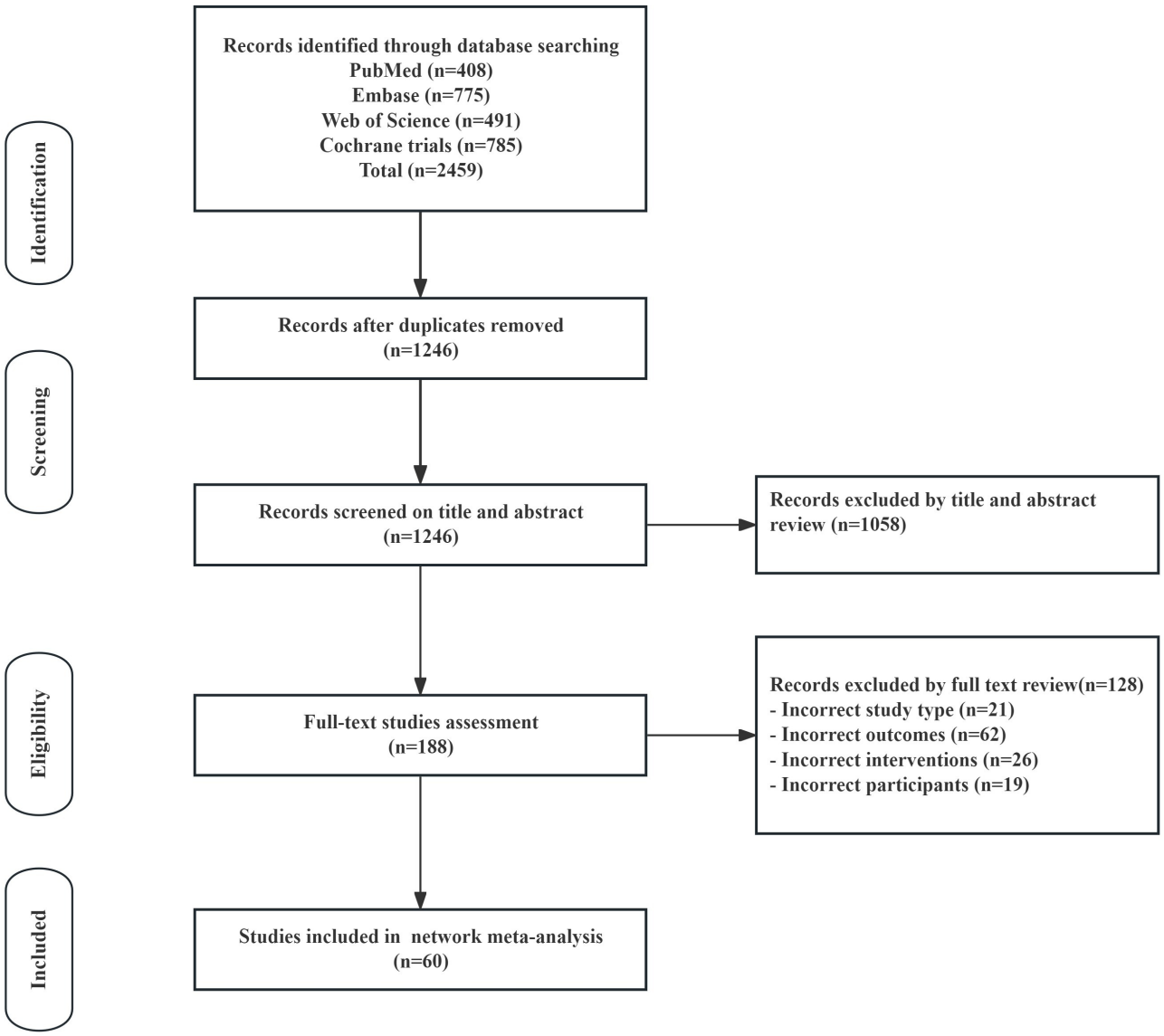


Table . Summary of trial arms and participants by intervention category in the primary (all-balance) network[a].

Intervention category	Trial arms (N=118), n (%)	Participants (N=2710), n (%)	Participants per arm, mean
Conventional exercise	29 (24.6)	581 (21.4)	20.0
Exergaming, balance platform	21 (17.8)	431 (15.9)	20.5
Exergaming, motion capture	18 (15.3)	360 (13.3)	20.0
Other serious games	10 (8.5)	318 (11.7)	31.8
Telerehabilitation	5 (4.2)	198 (7.3)	39.6
Interventions with wearables	3 (2.5)	29 (1.1)	9.7
No treatment	32 (27.1)	793 (29.3)	24.8

[a]Full per-study characteristics for all 60 included trials are provided in Table S4 in Multimedia Appendix 1. Values reflect arm-level contributions to network 1 (all balance outcomes). Network 1 included 55 (91.6%) of the 60 included trials; 2 further trials contributed only to the gait-speed network (network 4), and 3 trials contributed no quantitative data to any network. The 118 arms in the column header were contributed by 55 trials separately because 8 trials contributed 3 arms each and the remainder contributed 2.

Risk of Bias

RoB 2 assessment rated 2 (3%) of 60 studies at low overall risk of bias, 50 (83%) as having raised some concerns, and 8 (13%) at high risk. The predominant source of concern was D2 (deviations from intended interventions; 58/60, 97%, studies rated as having some concerns or high risk), attributable in large part to the structural impossibility of blinding participants to

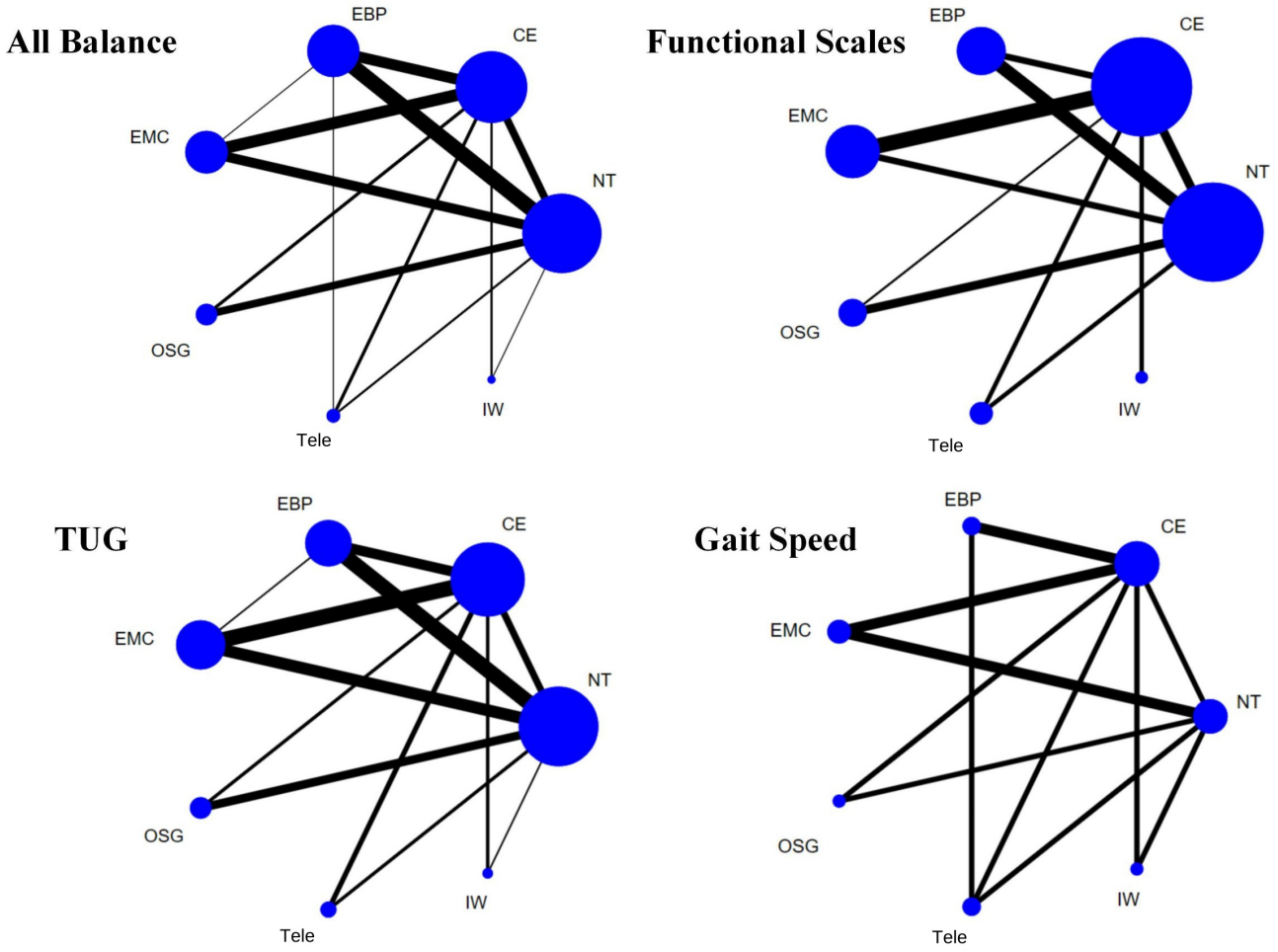
XSL•FO
RenderX

exercise interventions. D4 (outcome measurement) was rated as high risk in 6 (10%) studies, primarily reflecting the absence of blinded outcome assessors. Domain-level judgments for every trial are reported in Table S5 in Multimedia Appendix 1, and the RoB 2 traffic-light plot is presented as Figure S1 in Multimedia Appendix 1.

Network Structure

Network plots for all 4 analytical networks are presented in Figure 2. Network 1 (all balance) incorporated 55 trials, network 2 (functional scales) incorporated 26 trials, network 3 (TUG) incorporated 41 trials, and network 4 (gait speed) incorporated 12 trials. All 4 networks comprised a single fully connected component. The CE and NT nodes served as the principal hubs across all networks. Direct evidence for comparisons between active treatments was available for only a small proportion of dyads; the number of trials and participants contributing direct evidence to each contrast is tabulated in Table S8 in Multimedia Appendix 1.

Figure 2. Network plots presenting the direct evidence structure for each of the 4 analytical networks. CE: conventional exercise; EBP: exergaming with balance platforms; EMC: exergaming with motion capture; IW: interventions with wearables; NT: no treatment; OSG: other serious games; Tele: telerehabilitation; TUG: Timed Up and Go.



NMA Results

Main results for all 4 networks are presented in Table 2. Complete league tables are provided in Figure 3, and all-pairwise forest plots are provided in Figure 4.

XSL•FO
RenderX

Table . Network meta-analysis results: effect sizes (Hedges *g* SMD[a]), SUCRA[b] rankings, and GRADE[c] certainty of evidence for all treatment comparisons, sorted by SUCRA within each network.

Intervention	Studies, n	Participants, n[d]	SMD versus NT[e] (95% CI)	SMD versus CE[f] (95% CI)	SUCRA, %	GRADE certainty
Network 1: all balance outcomes (55 studies; global consistency test χ^2_{13} =16.1; *P*=.25)						
Tele[g]	5	198	0.69 (0.08 to 1.29)[h]	0.34 (−0.26 to 0.93)	76.3	⊕⊕◯◯ Low
EBP[i]	21	431	0.63 (0.31 to 0.95)[j]	0.28 (−0.07 to 0.63)	75.2	⊕⊕◯◯ Low
EMC[k]	18	360	0.58 (0.22 to 0.95)[j]	0.23 (−0.11 to 0.57)	68.7	⊕⊕◯◯ Low
OSG[l]	10	318	0.41 (−0.01 to 0.83)	0.06 (−0.41 to 0.52)	47.2	⊕⊕◯◯ Low
IW[m]	3	29	0.33 (−0.57 to 1.23)	−0.02 (−0.90 to 0.86)	41.2	⊕◯◯◯ Very low
CE	29	581	0.35 (0.03 to 0.68)[h]	—[n]	36.6	⊕⊕◯◯ Low
Network 2: functional scales (26 studies; global consistency test χ^2 =13.8; *P*=.13)						
IW	2	14	1.10 (−0.28 to 2.48)	1.00 (−0.25 to 2.25)	88.9	⊕◯◯◯ Very low
OSG	5	163	0.41 (−0.25 to 1.06)	0.31 (−0.48 to 1.10)	62.9	⊕⊕◯◯ Low
EMC	9	183	0.37 (−0.29 to 1.03)	0.27 (−0.24 to 0.79)	60.9	⊕⊕◯◯ Low
EBP	7	113	0.18 (−0.44 to 0.80)	0.08 (−0.61 to 0.78)	43.4	⊕⊕◯◯ Low
Tele	3	175	0.14 (−0.68 to 0.97)	0.05 (−0.79 to 0.88)	39.7	⊕⊕◯◯ Low
CE	15	286	0.10 (−0.47 to 0.67)	—	32.5	⊕⊕◯◯ Low
Network 3: Timed Up and Go test (41 studies; global consistency test χ^2_{12} =28.1; *P*=.005; inconsistency localized to NT vs CE)						
EBP	14	307	0.81 (0.42 to 1.20)[j]	0.21 (−0.21 to 0.64)	74.8	⊕⊕⊕◯ Moderate
Tele	4	190	0.82 (0.16 to 1.49)[h]	0.23 (−0.41 to 0.87)	72.5	⊕⊕◯◯ Low
EMC	14	272	0.79 (0.36 to 1.21)[j]	0.19 (−0.19 to 0.57)	71.7	⊕⊕⊕◯ Moderate
CE	22	442	0.60 (0.21 to 0.98)[j]	—	47.2	⊕⊕◯◯ Low[o]
IW	3	29	0.48 (−0.42 to 1.38)	−0.12 (−1.00 to 0.77)	41.3	⊕◯◯◯ Very low
OSG	7	245	0.49 (−0.005 to 0.978)	−0.11 (−0.66 to 0.44)	39.5	⊕⊕◯◯ Low
Network 4: gait speed (12 studies; supplementary and hypothesis-generating only; all comparisons nonsignificant)						
EMC	3	62	1.51 (−0.68 to 3.70)	0.36 (−1.78 to 2.49)	70.6	⊕◯◯◯ Very low
CE	7	144	1.15 (−0.82 to 3.13)	—	62.0	⊕◯◯◯ Very low
EBP	3	32	0.94 (−1.81 to 3.70)	−0.21 (−2.37 to 1.95)	53.6	⊕◯◯◯ Very low
Tele	3	165	0.91 (−1.53 to 3.35)	−0.24 (−2.49 to 2.00)	52.6	⊕◯◯◯ Very low
IW	2	21	0.85 (−1.83 to 3.52)	−0.31 (−3.00 to 2.39)	50.4	⊕◯◯◯ Very low
OSG	2	99	0.39 (−2.23 to 3.00)	−0.76 (−3.39 to 1.86)	38.8	⊕◯◯◯ Very low

[a]SMD: standardized mean difference (Hedges *g*; positive values favor the row treatment).

[b]SUCRA: surface under the cumulative ranking curve.

[c]GRADE: Grading of Recommendations Assessment, Development, and Evaluation (⊕⊕⊕⊕ high, ⊕⊕⊕◯ moderate, ⊕⊕◯◯ low, ⊕◯◯◯ very low).

[d]Number of participants in the respective intervention arm across contributing studies, summed at the arm level; the participant counts in Table 3 are the sum of both arms of each pairwise comparison and therefore differ from those reported here.

[e]NT: no treatment.

[f]CE: conventional exercise.

XSL•FO
RenderX

[g]Tele: telerehabilitation.

[h]*P*<.05 vs NT.

[i]EBP: exergaming with balance platforms.

[j]*P*<.01 vs NT. Exact *P* values for every comparison are reported in Table S9 in Multimedia Appendix 1.

[k]EMC: exergaming with motion capture.

[l]OSG: other serious games.

[m]IW: interventions with wearables.

[n]Not applicable.

[o]Certainty additionally rated down 1 level for inconsistency (significant direct-indirect discrepancy on node splitting, localized to the NT vs CE comparison).

Figure 3. League tables presenting all pairwise network meta-analysis effect estimates (standardized mean difference, 95% CI) for the 4 analytical networks. In the league tables, orange cells indicate statistically significant treatment comparisons, and white cells indicate nonsignificant comparisons. CE: conventional exercise; EBP: exergaming with balance platforms; EMC: exergaming with motion capture; IW: interventions with wearables; NT: no treatment; OSG: other serious games; Tele: telerehabilitation; TUG: Timed Up and Go.

All Balance

Tele						
0.05 (−0.58 to 0.69)	EBP					
0.10 (−0.55 to 0.76)	0.05 (−0.37 to 0.47)	EMC				
0.28 (−0.44 to 0.99)	0.22 (−0.28 to 0.72)	0.17 (−0.35 to 0.70)	OSG			
0.36 (−0.69 to 1.40)	0.30 (−0.62 to 1.22)	0.25 (−0.68 to 1.18)	0.08 (−0.89 to 1.05)	IW		
0.34 (−0.26 to 0.93)	0.28 (−0.07 to 0.63)	0.23 (−0.11 to 0.57)	0.06 (−0.41 to 0.52)	−0.02 (−0.90 to 0.86)	CE	
0.69 (0.08 to 1.29)	0.63 (0.31 to 0.95)	0.58 (0.22 to 0.95)	0.41 (−0.01 to 0.83)	0.33 (−0.57 to 1.23)	0.35 (0.03 to 0.68)	NT

Functional Scales

IW						
0.69 (−0.79 to 2.17)	OSG					
0.73 (−0.63 to 2.08)	0.04 (−0.84 to 0.91)	EMC				
0.92 (−0.52 to 2.35)	0.23 (−0.65 to 1.10)	0.19 (−0.61 to 0.99)	EBP			
0.95 (−0.55 to 2.46)	0.26 (−0.76 to 1.29)	0.23 (−0.71 to 1.16)	0.04 (−0.94 to 1.02)	Tele		
1.00 (−0.25 to 2.25)	0.31 (−0.48 to 1.10)	0.27 (−0.24 to 0.79)	0.08 (−0.61 to 0.78)	0.05 (−0.79 to 0.88)	NT	
1.10 (−0.28 to 2.48)	0.41 (−0.25 to 1.06)	0.37 (−0.29 to 1.03)	0.18 (−0.44 to 0.80)	0.14 (−0.68 to 0.97)	0.10 (−0.47 to 0.67)	CE

TUG

EBP						
−0.01 (−0.73 to 0.70)	Tele					
0.02 (−0.47 to 0.52)	0.04 (−0.68 to 0.75)	EMC				
0.21 (−0.21 to 0.64)	0.23 (−0.41 to 0.87)	0.19 (−0.19 to 0.57)	CE			
0.33 (−0.61 to 1.27)	0.35 (−0.72 to 1.41)	0.31 (−0.63 to 1.25)	0.12 (−0.77 to 1.00)	IW		
0.32 (−0.27 to 0.92)	0.34 (−0.46 to 1.13)	0.30 (−0.31 to 0.91)	0.11 (−0.44 to 0.66)	−0.01 (−1.01 to 0.99)	OSG	
0.81 (0.42 to 1.20)	0.82 (0.16 to 1.49)	0.79 (0.36 to 1.21)	0.60 (0.21 to 0.98)	0.48 (−0.42 to 1.38)	0.49 (−0.00 to 0.98)	NT

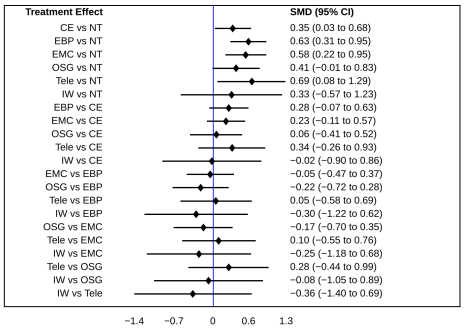
Gait Speed

EMC						
0.36 (−1.78 to 2.49)	CE					
0.56 (−2.40 to 3.52)	0.21 (−1.95 to 2.37)	EBP				
0.60 (−2.26 to 3.46)	0.24 (−2.00 to 2.49)	0.03 (−2.50 to 2.57)	Tele			
0.66 (−2.49 to 3.81)	0.31 (−2.39 to 3.00)	0.10 (−3.28 to 3.47)	0.06 (−3.22 to 3.34)	IW		
1.12 (−1.98 to 4.22)	0.76 (−1.86 to 3.39)	0.56 (−2.76 to 3.88)	0.52 (−2.70 to 3.74)	0.46 (−3.02 to 3.94)	OSG	
1.51 (−0.68 to 3.70)	1.15 (−0.82 to 3.13)	0.94 (−1.81 to 3.70)	0.91 (−1.53 to 3.35)	0.85 (−1.83 to 3.52)	0.39 (−2.23 to 3.00)	NT

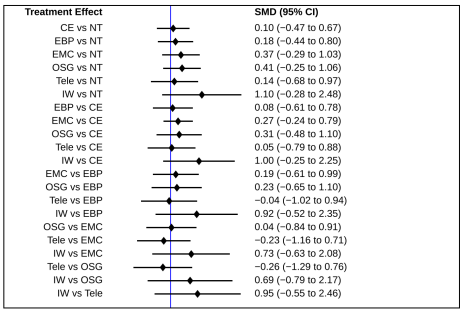
XSL•FO
RenderX

Figure 4. All-pairwise forest plots from the network meta-analysis consistency models for the 4 analytical networks. CE: conventional exercise; EBP: exergaming with balance platforms; EMC: exergaming with motion capture; IW: interventions with wearables; NT: no treatment; OSG: other serious games; Tele: telerehabilitation; TUG: Timed Up and Go.

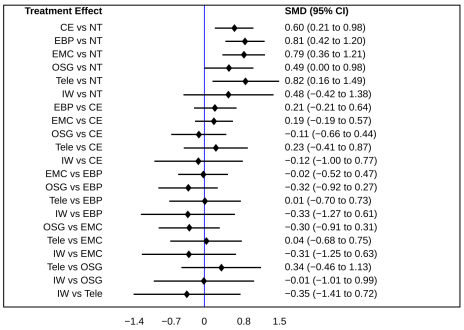
All Balance



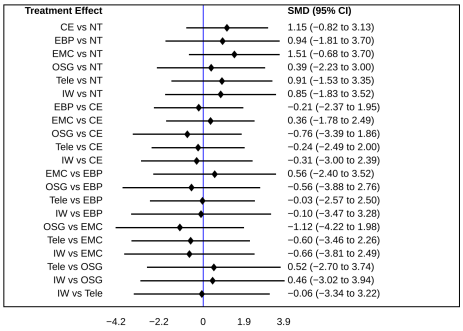
Functional Scales



TUG



Gait Speed



XSL•FO
RenderX

Table . GRADE[a] evidence profile for the network meta-analysis of technology-based rehabilitation interventions for balance and functional mobility in older adults (60 randomized controlled trials; 57 contributing data)[b].

Comparison	Studies, n	Participants, n[c]	SMD[d] (95% CI)	Reasons for rating down	Certainty
Network 1: all balance outcomes | 55 studies | global consistency test χ^2_{13} =16.1; *P*=.25					
EBP[e] versus NT[f]	21	1224	0.63 (0.31 to 0.95)	Risk of bias (serious); imprecision (serious)[g]	⊕⊕◯◯ Low
EMC[h] versus NT	18	1153	0.58 (0.22 to 0.95)	Risk of bias (serious); imprecision (serious)[g]	⊕⊕◯◯ Low
Tele[i] versus NT	5	991	0.69 (0.08 to 1.29)	Risk of bias (serious); imprecision (serious)[g]	⊕⊕◯◯ Low
OSG[j] versus NT	10	1111	0.41 (−0.01 to 0.83)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
IW[k] versus NT	3	822	0.33 (−0.57 to 1.23)	Risk of bias (serious); imprecision (very serious)	⊕◯◯◯ Very low
CE[l] versus NT	29	1374	0.35 (0.03 to 0.68)	Risk of bias (serious); imprecision (serious)[g]	⊕⊕◯◯ Low
EBP versus CE	21	1012	0.28 (−0.07 to 0.63)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
EMC versus CE	18	941	0.23 (−0.11 to 0.57)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
Tele versus CE	5	779	0.34 (−0.26 to 0.93)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
OSG versus CE	10	899	0.06 (−0.41 to 0.52)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
IW versus CE	3	610	−0.02 (−0.90 to 0.86)	Risk of bias (serious); imprecision (very serious)	⊕◯◯◯ Very low
Network 2: functional scales | 26 studies | global consistency test χ^2 =13.8, *P*=.13					
IW versus NT	2	399	1.10 (−0.28 to 2.48)	Risk of bias (serious); imprecision (very serious)	⊕◯◯◯ Very low
OSG versus NT	5	548	0.41 (−0.25 to 1.06)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
EMC versus NT	9	568	0.37 (−0.29 to 1.03)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
EBP versus NT	7	498	0.18 (−0.44 to 0.80)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
Tele versus NT	3	560	0.14 (−0.68 to 0.97)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
CE versus NT	15	671	0.10 (−0.47 to 0.67)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
IW versus CE	2	300	1.00 (−0.25 to 2.25)	Risk of bias (serious); imprecision (very serious)	⊕◯◯◯ Very low
OSG versus CE	5	449	0.31 (−0.48 to 1.10)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
EMC versus CE	9	469	0.27 (−0.24 to 0.79)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
EBP versus CE	7	399	0.08 (−0.61 to 0.78)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low

XSL•FO
RenderX

Comparison	Studies, n	Participants, n[c]	SMD[d] (95% CI)	Reasons for rating down	Certainty
Tele versus CE	3	461	0.05 (−0.79 to 0.88)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
Network 3: Timed Up and Go test | 41 studies | global consistency test $\chi^2_{1\ 2}$ =28.1; *P*=.005 (inconsistency localized to NT vs CE; all technology comparisons *P*>.08)					
EBP versus NT	14	940	0.81 (0.42 to 1.20)	Risk of bias (serious)	⊕⊕⊕◯ Moderate
EMC versus NT	14	905	0.79 (0.36 to 1.21)	Risk of bias (serious)	⊕⊕⊕◯ Moderate
Tele versus NT	4	823	0.82 (0.16 to 1.49)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
CE versus NT	22	1075	0.60 (0.21 to 0.98)	Risk of bias (serious); inconsistency (serious)[m]	⊕⊕◯◯ Low
OSG versus NT	7	878	0.49 (−0.005 to 0.978)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
IW versus NT	3	662	0.48 (−0.42 to 1.38)	Risk of bias (serious); imprecision (very serious)	⊕◯◯◯ Very low
EBP versus CE	14	749	0.21 (−0.21 to 0.64)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
EMC versus CE	14	714	0.19 (−0.19 to 0.57)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
Tele versus CE	4	632	0.23 (−0.41 to 0.87)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
OSG versus CE	7	687	−0.11 (−0.66 to 0.44)	Risk of bias (serious); imprecision (serious)	⊕⊕◯◯ Low
IW versus CE	3	471	−0.12 (−1.00 to 0.77)	Risk of bias (serious); imprecision (very serious)	⊕◯◯◯ Very low
Network 4: gait speed | 12 studies | global consistency test χ^2 =69.5; *P*<.001; supplementary and hypothesis-generating only					
EMC versus NT	3	285	1.51 (−0.68 to 3.70)	Risk of bias (serious); inconsistency (serious); imprecision (very serious)	⊕◯◯◯ Very low
CE versus NT	7	367	1.15 (−0.82 to 3.13)	Risk of bias (serious); inconsistency (serious); imprecision (very serious)	⊕◯◯◯ Very low
EBP versus NT	3	255	0.94 (−1.81 to 3.70)	Risk of bias (serious); inconsistency (serious); imprecision (very serious)	⊕◯◯◯ Very low
Tele versus NT	3	388	0.91 (−1.53 to 3.35)	Risk of bias (serious); inconsistency (serious); imprecision (very serious)	⊕◯◯◯ Very low
IW versus NT	2	244	0.85 (−1.83 to 3.52)	Risk of bias (serious); inconsistency (serious); imprecision (very serious)	⊕◯◯◯ Very low

XSL•FO
RenderX

Comparison	Studies, n	Participants, n[c]	SMD[d] (95% CI)	Reasons for rating down	Certainty
OSG versus NT	2	322	0.39 (−2.23 to 3.00)	Risk of bias (serious); inconsistency (serious); imprecision (very serious)	⊕◯◯◯ Very low

[a]GRADE: Grading of Recommendations Assessment, Development, and Evaluation (⊕⊕⊕⊕ high, ⊕⊕⊕◯ moderate, ⊕⊕◯◯ low, ⊕◯◯◯ very low).

[b]All evidence begins at high certainty (randomized trials) and is rated down 1 level (serious) or 2 levels (very serious) per domain. Prespecified decision rules: imprecision was rated down 1 level when the 95% CI included the null or a trivial effect (|SMD|<0.20) and 2 levels when the CI additionally spanned appreciable benefit and appreciable harm (|SMD|>0.50) or fewer than 3 trials contributed; inconsistency was rated down where the design-by-treatment interaction test was significant and node splitting localized the discrepancy to the comparison concerned; indirectness was not rated down for any comparison, as all trials were randomized comparisons in the target population using the intended outcomes; and publication bias was not rated down for any comparison, as no comparison was supported by the ≥10 trials required for a meaningful test of funnel plot asymmetry. GRADE ratings were performed independently by 2 reviewers and finalized by consensus, following the framework adapted for network meta-analysis [15]; risk of bias was assessed with RoB 2 [25].

[c]Sum of arm-level n across contributing studies for both arms of each comparison.

[d]SMD: standardized mean difference.

[e]EBP: exergaming with balance platforms.

[f]NT: no treatment.

[g]In network 1, imprecision was judged against the 95% prediction interval rather than the CI because the between-study heterogeneity of the composite outcome (τ=0.597) means the CI understates the uncertainty applying to any individual balance instrument; all network 1 prediction intervals included the null.

[h]EMC: exergaming with motion capture.

[i]Tele: telerehabilitation.

[j]OSG: other serious games.

[k]IW: interventions with wearables.

[l]CE: conventional exercise.

[m]Rating down for inconsistency applies to CE vs NT only: node-splitting showed a significant discrepancy between direct (SMD 1.61) and indirect (SMD 0.19) evidence (*P*<.001), whereas all technology-based comparisons showed nonsignificant direct-indirect differences (all *P*>.08).

Network 1: All Balance Outcomes

The consistency model (55 studies; τ=0.597; τ²=0.356) identified Tele (SMD 0.69, 95% CI 0.08 - 1.29; *P*=.03), EBP (SMD 0.63, 95% CI 0.31 - 0.95; *P*<.001), and EMC (SMD 0.58, 95% CI 0.22 - 0.95; *P*=.002) as significantly more effective than NT. CE was also significantly superior to NT (SMD 0.35, 95% CI 0.03 - 0.68; *P*=.03). OSG showed a numerically positive but nonsignificant effect versus NT (SMD 0.41, 95% CI −0.01 to 0.83; *P*=.06), as did IW (SMD 0.33, 95% CI −0.57 to 1.23; *P*=.47), the latter based on only 3 trials. No technology-based intervention was statistically superior to CE. SUCRA-based ranking placed Tele first (76.3%), followed by EBP (75.2%), EMC (68.7%), OSG (47.2%), IW (41.2%), CE (36.6%), and NT (4.8%). The narrow separation between the 3 leading nodes (7.6 SUCRA percentage points) should be read together with the ranking probabilities reported in the Ranking Probabilities section, which show that no single intervention emerged as clearly best.

Network 2: Functional Scales

The consistency model (26 studies) yielded no statistically significant treatment contrasts. IW ranked first by SUCRA (88.9%), although this estimate rested on only 2 trials contributing 14 participants and carried a very wide CI (SMD 1.10, 95% CI −0.28 to 2.48; *P*=.12), substantially limiting its interpretability. OSG (SUCRA 62.9%), EMC (60.9%), EBP (43.4%), Tele (39.7%), CE (32.5%), and NT (21.7%) occupied subsequent positions.

Network 3: Timed Up and Go Test

The consistency model (41 studies) demonstrated that EBP (SMD 0.81, 95% CI 0.42 - 1.20; *P*<.001), EMC (SMD 0.79, 95% CI 0.36 - 1.21; *P*<.001), Tele (SMD 0.82, 95% CI 0.16 - 1.49; *P*=.02), and CE (SMD 0.60, 95% CI 0.21 - 0.98; *P*=.002) were each significantly superior to NT. OSG did not reach significance (SMD 0.49, 95% CI −0.005 to 0.978; *P*=.05), and neither did IW (SMD 0.48, 95% CI −0.42 to 1.38; *P*=.30). None of the technology-based interventions was significantly superior to CE. SUCRA ranking identified EBP (74.8%), Tele (72.5%), and EMC (71.7%) as the 3 highest-ranked interventions, separated by 3.1 percentage points.

Network 4: Gait Speed

The gait speed network (12 trials) was characterized by extreme between-study heterogeneity and very wide CIs for all treatment contrasts, none of which attained statistical significance. Given the significant global inconsistency and the extreme heterogeneity, these estimates are unsuitable for treatment-effect inference. They are reported as supplementary and hypothesis generating only, are not used to support any conclusion of this review, and should not be cited as evidence that these interventions do or do not affect gait speed.

Inconsistency Testing

For network 1, the design-by-treatment interaction test was nonsignificant (χ^2_{13}=16.1; *P*=.25), providing no evidence of meaningful inconsistency. Network 2 similarly showed no

XSL•FO
RenderX

significant inconsistency (χ^2 =13.8; *P*=.13). Network 3 yielded a significant global statistic ($\chi^2_{1\ 2}$ =28.1; *P*=.005). Node splitting localized this to the NT versus CE comparison exclusively (direct SMD 1.612, SE 0.331 vs indirect SMD 0.189, SE 0.205; difference 1.423; *P*<.001); all 11 remaining contrasts showed nonsignificant direct-indirect differences (all *P*>.08). Network 4 also demonstrated significant global inconsistency (χ^2 =69.5; *P*<.001).

Two features of the direct evidence bear on how this discrepancy should be read. First, it does not arise from a small number of aberrant results: the 4 trials that carry this contrast, which span 3 study designs, each produced a large positive direct effect (Hedges *g*=0.90, 1.09, 2.36, and 2.53; Table S7, Multimedia Appendix 1), against an indirect estimate of 0.19. Second, the 4 trials share a distinctive structure rather than a distinctive dose—each is a 3-arm trial containing an NT arm, each randomized only 11 to 18 participants per arm, and each prescribed CE 3 times per week, which is the median frequency across the 29 CE arms in the review.

Both sensitivity specifications attenuated every treatment estimate, and neither eliminated the discrepancy (Table S7 and Figure S3, Multimedia Appendix 1). Under specification 1, the direct-indirect difference for CE versus NT remained significant, and under specification 2, only a single trial contributed direct evidence, so the contrast was no longer estimable. The treatment hierarchy was largely preserved under specification 1 (Spearman ρ=0.86) and less so under specification 2 (ρ=0.68), where removal of 1 of only 4 Tele arms substantially destabilized that node. Across both specifications, the 2 comparisons carrying moderate certainty in the primary analysis—EBP versus NT and EMC versus NT—remained statistically significant. We therefore do not claim that consistency was restored by removing trials; the primary analysis retains all trials, and the CE versus NT comparison is instead rated down for inconsistency in the GRADE assessment.

Transitivity Assessment

The distribution of the effect modifiers that could be extracted is summarized in Table S6 and Figure S2 in Multimedia Appendix 1. Across the 7 nodes of network 1, the mean of trial mean ages ranged from 70.7 years (Tele) to 78.6 years (IW), the mean intervention duration from 6.7 weeks (IW) to 30.0 weeks (Tele), and the mean session length from 33.8 minutes (Tele) to 50.0 minutes (IW). The 5 nodes supported by 10 or more trials (NT, CE, EBP, EMC, and OSG) were closely comparable to one another on every extracted modifier (mean age 71.7 - 73.8 years; mean duration 7.7 - 11.5 weeks; mean session length 38.8 - 48.9 minutes), and the transitivity assumption appears defensible for comparisons among them.

Two nodes depart from this pattern and require caution. The Tele node rests on 5 trials and has a much longer mean intervention duration (30.0 wk) than any other node; this is driven entirely by a single 104-week trial [22], and the median Tele duration is 8 weeks. The IW node rests on 3 trials whose participants were on average 5 to 7 years older than those in the other nodes. Comparisons involving Tele and IW are therefore more vulnerable to violation of transitivity than comparisons among the remaining nodes and are interpreted accordingly throughout.

A further limitation of this appraisal is that several effect modifiers of recognized importance for balance outcomes could not be evaluated because they were reported too inconsistently across trials: baseline balance impairment, baseline fall risk or fall history, frailty status, sex distribution, the level of therapist supervision, the intensity of the CE comparator, prior familiarity with the technology, and the outcome measurement protocol. Residual intransitivity attributable to these unmeasured modifiers cannot be excluded, and the CE-intensity analyses reported below indicate that at least one of them—comparator intensity—does materially influence effect estimates.

Meta-Regression of Training Dose

Post hoc random-effects meta-regressions found no statistically significant association between the technology versus control effect size and any training-dose parameter, in either the all-balance network or the TUG network (all *P*>.05; Table S10 and Figure S4, Multimedia Appendix 1). The inverse trend for session length and weekly training minutes approached but did not reach significance (*P*=.05-.10) and was opposite in direction to a simple “more-is-better” expectation, most plausibly reflecting confounding by baseline population severity: trials recruiting more impaired participants tended to prescribe longer sessions while achieving smaller standardized effects. As these analyses are based on study-level aggregate data, they are subject to ecological bias and are underpowered to detect modest dose-response relationships; the current evidence neither identifies an optimal training dose nor excludes one. What can be stated descriptively is the dose at which the observed effects were obtained: across the included trials, the median program ran for 8 weeks, with 3 sessions per week and 40 minutes per session. This describes the prescriptions that were tested rather than an established minimum effective dose, and the absence of a dose-response signal means it should not be read as a threshold below which benefit would be lost.

CE training frequency was associated with the magnitude of the direct CE versus NT effect (slope 1.46 per additional session per week; *P*=.04), while no other comparator-intensity parameter reached significance (Table S11 and Figure S5, Multimedia Appendix 1). This analysis rests on the 7 trials in network 1 that compared CE directly with NT and is hypothesis generating rather than confirmatory. It should also be noted that it does not, by itself, explain the direct-indirect discrepancy observed in the TUG network because all 4 trials contributing direct evidence to that contrast prescribed CE at the same frequency as the network median; comparator intensity is therefore a plausible general source of heterogeneity in the CE node, but not the specific explanation for the inconsistency detected in network 3.

Ranking Probabilities

Probabilities of being the best treatment temper the interpretation of the SUCRA hierarchy (Table S12 and Figure S6, Multimedia Appendix 1). In the all-balance network, Tele had a probability of approximately 41% of being the best treatment, followed by EBP (approximately 22%), EMC (approximately 16%), IW

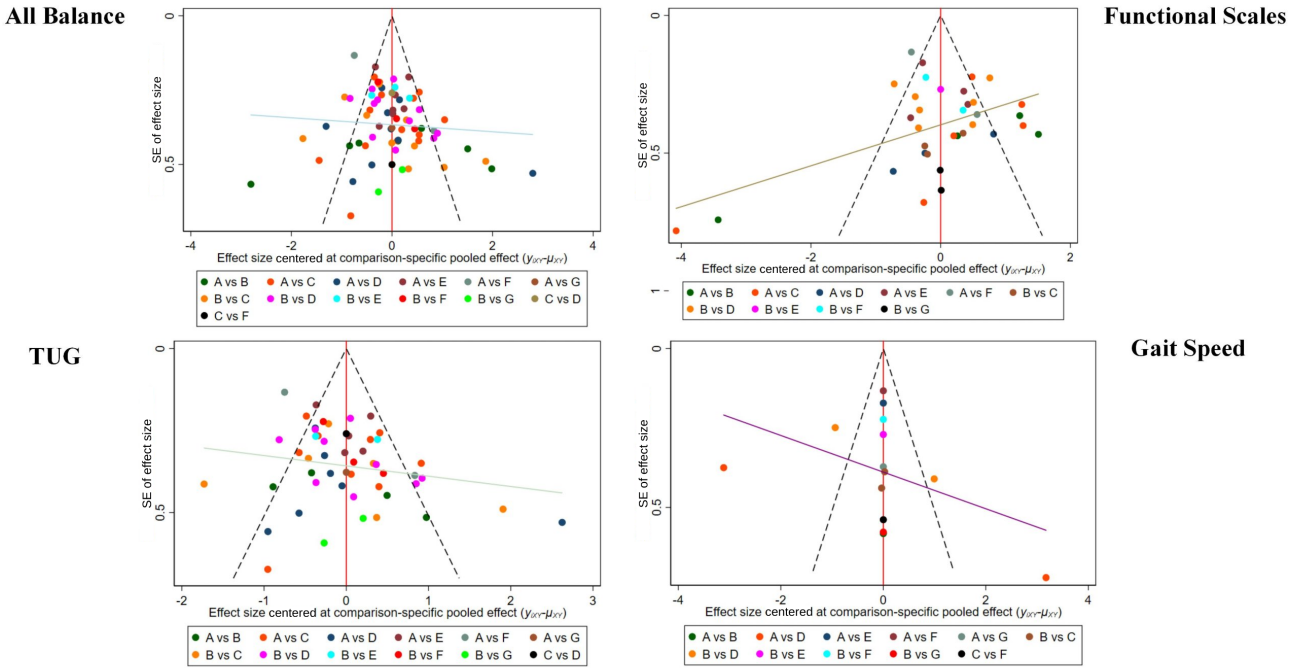
XSL•FO
RenderX

(approximately 15%), and OSG (approximately 5%); the probabilities for CE and NT were negligible. In the TUG network, the corresponding values were approximately 36% for Tele, 26% for EBP, 21% for EMC, and 13% for IW. The substantial probability attaching to IW despite its fifth-place SUCRA position arises from the extreme width of its CI (−0.57 to 1.23 in the all-balance network, 3 trials, 29 participants) rather than from any robust signal of benefit, and this illustrates why ranking metrics must not be read independently of effect size, precision, and certainty of evidence. No intervention emerged as clearly the best option, and the SUCRA hierarchy should not be read as a definitive ordering of clinical effectiveness.

Small-Study Effects and Publication Bias

For networks 1 and 3, visual inspection of comparison-adjusted funnel plots suggested broad symmetry around the comparison-specific pooled effects. Network 2 exhibited marginal asymmetry in the IW comparisons (Figure 5). Network 4 showed notable asymmetry. However, networks 2 and 4 contribute only 26 and 12 trials, respectively, across 7 nodes, and no comparison in either network is supported by the 10 trials conventionally regarded as the minimum for a meaningful test of funnel plot asymmetry. The apparent asymmetry in these 2 networks is therefore uninterpretable, and no comparison was rated down for publication bias on this basis; the possibility of unpublished negative trials, particularly for the sparsely studied IW and OSG categories, cannot be excluded and is treated as an unresolved limitation rather than as a graded judgment.

Figure 5. Comparison-adjusted funnel plots for assessment of small-study effects and publication bias in the 4 analytical networks. TUG: Timed Up and Go.



A: No Treatment; B: Conventional Exercises; C: Exergaming with Balance Platform; D: Exergaming with Motion Capture; E: Other Serious Games; F: Telerehabilitation; G: Interventions with Wearables

Certainty of Evidence

GRADE ratings are summarized in Table 3. Certainty was rated moderate for only 2 comparisons—EBP versus NT and EMC versus NT in the TUG network—both of which were rated down 1 level for risk of bias and met the prespecified precision criterion without further downgrading. Every other comparison in the review was rated low or very low. In network 1, all comparisons were rated down for imprecision because the 95% prediction interval included the null, reflecting the substantial between-study heterogeneity inherent to a composite balance outcome (τ=0.597). All comparisons involving IW were rated very low owing to very serious imprecision arising from 3 or fewer contributing trials. All network 4 comparisons were rated very low, being rated down for both very serious imprecision and serious inconsistency. No comparison was rated down for indirectness, and, for the reasons given above, none was rated down for publication bias. The explicit judgment reached for each of the 5 GRADE domains, for every comparison rated, is tabulated in Table S13 in Multimedia Appendix 1.

Acceptability, Adverse Events, and User Burden

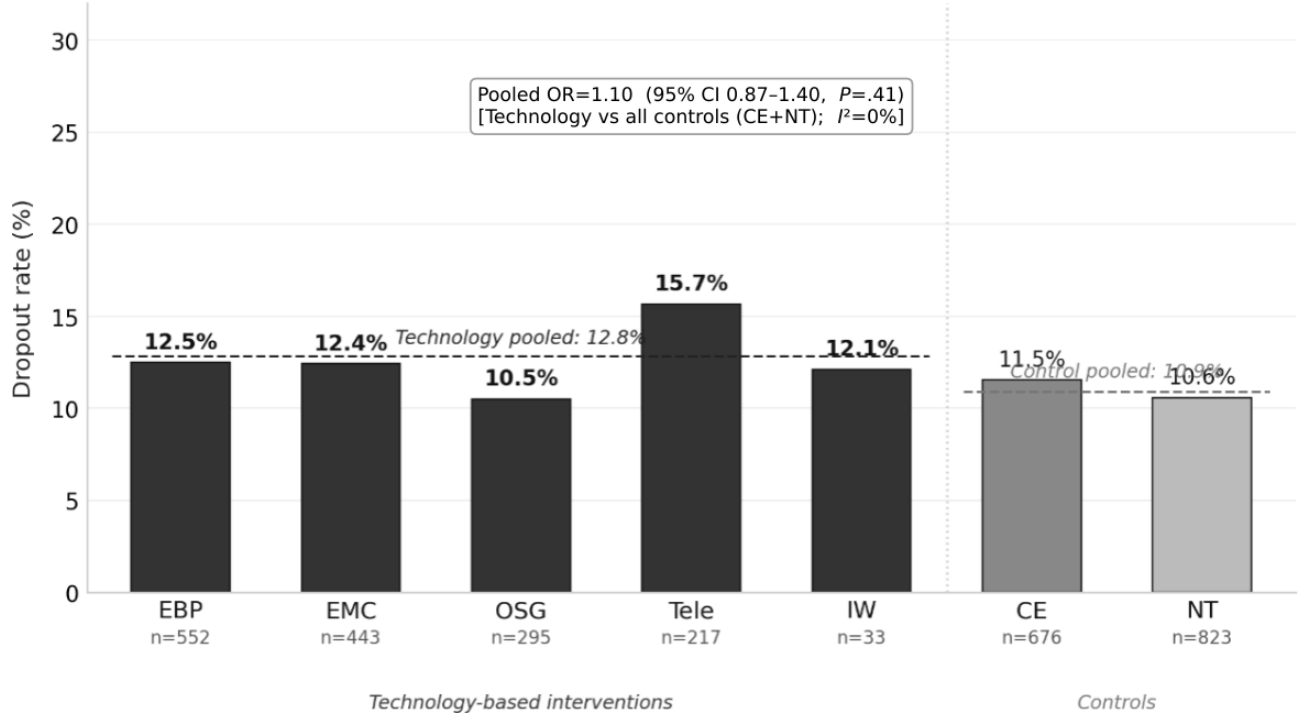
Dropout data were available for 59 (98.3%) of the 60 included trials (all except 1 [52], for which attrition was not reported). The crude dropout rates per category were 12.5% (69/552) for EBP, 12.4% (55/443) for EMC, 15.7% (34/217) for Tele, 12% (4/33) for IW, 10.5% (31/295) for OSG, 11.5% (78/676) for CE, and 10.6% (87/823) for NT. These denominators are the numbers of participants randomized to arms that both reported attrition and were assignable to a single network node; they therefore sum to 3039 rather than to the 3469 participants randomized overall. Pairwise odds ratios for all 5 technology categories versus CE were nonsignificant (all P>.05; Figure 6), with a pooled overall estimate of odds ratio 1.05 (95% CI 0.76 - 1.43; I^2=0%). When NT was included as an additional control condition, the pooled overall odds ratio for

XSL•FO
RenderX

technology-based interventions versus all control arms was 1.10 (95% CI 0.87 - 1.40; *P*=.41; I^2=0%), indicating no statistically significant difference in acceptability between technology-based and conventional rehabilitation (Figures 6 and 7).

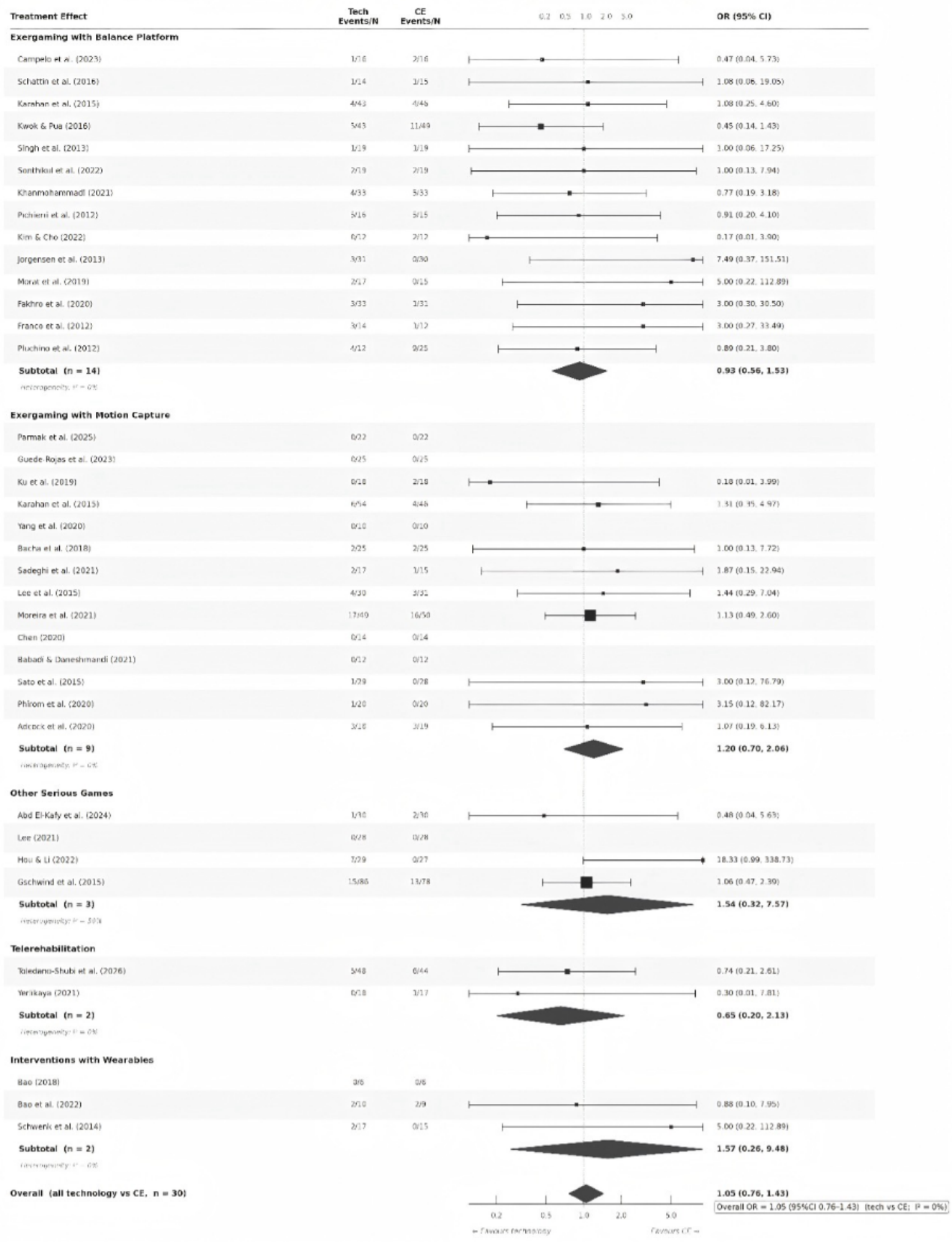
Reporting of the other feasibility outcomes was too sparse to permit synthesis. Only 3 (5%) of the 60 included trials reported a formal safety or adverse event outcome, and only 3 (5%) reported participant satisfaction or acceptability using a defined instrument; technology-related difficulties such as device malfunction, connectivity failure, or cybersickness were reported in a structured manner by a single trial [19], which used the Virtual Reality Sickness Questionnaire. Where adverse events were described, they were minor and were not attributed to the technology itself. The near-absence of standardized reporting of harms, usability, and satisfaction is an important evidence gap: for interventions whose principal claimed advantage is accessibility and engagement, these outcomes are as decision relevant as the effect on balance, and dropout alone is an inadequate proxy for them.

Figure 6. Pooled crude dropout rates by intervention category across all included trials with available dropout data (*N* below each bar denotes the total number of participants randomized per category). CE: conventional exercise; EBP: exergaming with balance platforms; EMC: exergaming with motion capture; IW: interventions with wearables; NT: no treatment; OR: odds ratio; OSG: other serious games; Tele: telerehabilitation.



XSL•FO
RenderX

Figure 7. Forest plot of all-cause dropout rates comparing each technology-based intervention category against conventional exercises [18,19,34-36,41,42,44-47,49-51,54,55,58,60-68,73,74,77,79,82-86]. CE: conventional exercise; EBP: exergaming with balance platforms; EMC: exergaming with motion capture; IW: interventions with wearables; OR: odds ratio; OSG: other serious games; Tele: telerehabilitation



XSL•FO
RenderX

Discussion

Principal Findings

This update includes 60 RCTs, of which 57 (95%) contributed data, and extends the evidence base by approximately 2 years beyond the original review [12]. Three findings stand out.

First, more technologies now show a benefit. The original review identified only EMC as statistically superior to no treatment, and only for TUG; here EBP, EMC, and Tele each improved outcomes significantly versus NT in both the all-balance composite and the TUG networks. This reflects the entry of larger, better-powered trials, particularly in the Tele and EBP categories, into nodes that previously lacked the precision to detect plausible effects.

Second, the treatment hierarchy has shifted. Tele rose from fifth of 7 ranked interventions in the original NMA (SUCRA 43%) to first in the all-balance network. However, the node contains only 5 trials and 198 participants, and Tele's probability of being the best treatment is approximately 41%, so this ranking is a hypothesis awaiting confirmation rather than an established hierarchy.

Third, GRADE appraisal, applied here for the first time in this technology category, shows that certainty is moderate at best and moderate for only 2 (5.1%) of the 39 pairwise comparisons rated. Risk of bias in the underlying literature is the main reason. The impossibility of blinding participants to an exercise intervention contributes to this, but it is not the whole explanation: blinded outcome assessment was absent or unclear in a substantial minority of trials, and fewer than half reported a prespecified protocol. Both deficiencies are remediable, and the moderate ratings for EBP and EMC versus NT in the TUG network show that the threshold is reachable in this field.

Telerehabilitation: Highest Ranked but Preliminary

Tele ranked highest across several outcome networks in this update, but the wide CI (SMD 0.69, 95% CI 0.08 - 1.29) and the small number of contributing trials call for caution. In the original review, only 3 Tele trials were available [12]. The addition of further large RCTs since 2023 contributed to statistically significant effects versus NT in both network 1 and network 3, consistent with a 2024 systematic review by Gamble et al [88], which found Tele to be noninferior to control interventions across 26 RCTs. The large randomized sample sizes of Tele trials (mean 124 participants per trial) confer greater statistical power per study, which may partly explain the elevation of the SUCRA ranking [22]. Three features of the evidence argue for restraint. The node rests on 5 trials, its 95% CI extends from a trivial effect (0.08) to a very large one (1.29), and its mean intervention duration is the longest and least typical of any node, driven by a single 104-week trial. The sensitivity analysis reported earlier, which removed 1 of the 4 Tele arms in the TUG network, widened the estimate further. Tele is therefore promising and highest ranked, but not demonstrably superior.

Tele works through a different mechanism from exergaming: expert supervision, error correction, and motivational support delivered in real time, which reproduces the cognitive behavioral elements of face-to-face physiotherapy without the access barriers [89,90]. The finding that Tele was not statistically superior to CE in any network is consistent with this and suggests a role as a scalable complement to in-person care rather than a replacement for it.

Convergence of Evidence for Exergaming Platforms and Motion Capture

EBP and EMC maintained significant effects versus NT across both the all-balance composite and TUG networks, with GRADE ratings of moderate certainty for EBP versus NT (SMD 0.81, 95% CI 0.42 - 1.20) and EMC versus NT (SMD 0.79, 95% CI 0.36 - 1.21) in network 3—the only moderate-certainty findings in this NMA. The comparable effect sizes of EBP and EMC are consistent with a recent meta-analysis [91] that found no significant difference between platforms, reflecting their shared therapeutic mechanism of real-time visual augmented feedback enabling immediate postural error correction [6]. Clinicians may reasonably select between them based on practical considerations, including equipment cost, installation requirements, and participant mobility constraints.

Serious Games and Wearable Interventions: Persistent Evidence Gaps

OSG fell just short of significance versus NT in network 3 (SMD 0.49, 95% CI −0.005 to 0.978; *P*=.05). This estimate should not be reported as a positive finding, and the accompanying certainty rating is low. Despite being purpose-built rehabilitation software, OSG did not outperform commercially available exergaming platforms in any network, suggesting that therapeutic benefits may depend more on movement engagement and practice volume than on software sophistication [92,93].

IW similarly failed to reach statistical significance in any comparison. The extreme width of the IW CIs reflects the small number of eligible trials (n=3), all involving custom-designed haptic or vibrotactile feedback systems. This gap has persisted over more than a decade of published work and is unlikely to reflect a lack of researcher interest. Three structural barriers are plausible. First, the wearable systems evaluated to date have been bespoke research prototypes requiring specialist calibration and supervision, which limits both trial size and clinical scalability; commercial devices with validated balance-feedback functions have only recently become available. Second, feedback modalities remain unstandardized—vibrotactile, auditory, and visual channels are used interchangeably across trials, so that the IW node aggregates mechanistically dissimilar interventions. Third, the cost of custom sensor arrays has kept sample sizes small: the 3 IW trials contributing to this review randomized 12, 16, and 33 participants, respectively. These findings indicate insufficient investigation rather than inefficacy. The priority for this category is agreement on a standardized feedback modality and outcome set that would make future trials poolable, not further small pilot trials [13].

The All-Balance Composite Network

Network 1 demonstrates that Tele, EBP, EMC, and CE each produced statistically significant improvements relative to NT (SMDs 0.35 - 0.69). The narrow span of effects across technically distinct platforms suggests a common minimum

XSL•FO
RenderX

benefit achievable through any adequately structured balance program, whatever the delivery mechanism. No technology exceeded CE in any network, consistent with technology acting as a delivery vehicle rather than a source of superior physiological stimulus.

How far this network can be interpreted is nevertheless limited by its construction. It pools instruments that measure genuinely different constructs—dynamic functional mobility (TUG), multidimensional clinical rating scales (BBS, Mini-BESTest, and SPPB), static postural sway (COP parameters and single-leg stance), and lower-limb function (sit-to-stand)—into a single standardized effect. A composite SMD of 0.63 does not tell a clinician which dimension of balance improved, and the substantial between-study heterogeneity (τ=0.597) is in part a direct expression of this structural pooling rather than of true clinical variability. For this reason, the prediction interval rather than the CI was used to rate imprecision in this network; every network 1 comparison is rated low or very low certainty, and the principal clinical conclusions of this review rest on network 3 (TUG), which is a single-construct outcome, is supported by the largest number of trials of any single instrument, and yields the only moderate-certainty estimates in the review. Network 1 therefore tests whether any balance benefit is detectable across the evidence base as a whole, while network 3 provides the estimate on which practice decisions should be based.

Outcome-Specific Results: Instrument Sensitivity and Task Alignment

The absence of significant effects in network 2 (functional scales) probably reflects the differential sensitivity of these instruments to the adaptations induced by short- to medium-duration technology-based training. The TUG mirrors the movement demands of many such programs, particularly sit-to-stand transitions, walking, and turning, giving it high construct validity and sensitivity to training-induced change [94]. Multidimensional scales such as the BBS, Mini-BESTest, and SPPB instead span a broader range of postural control domains, not all equally targeted by a typical exergaming program [95]. Future trials should adopt the TUG as the primary outcome for functional mobility, supplemented by domain-specific subscales where appropriate.

Network 3 (41 trials) yields the most internally consistent pattern of results and constitutes the primary basis for the moderate GRADE certainty ratings assigned to EBP and EMC. EBP (SMD 0.81), EMC (0.79), Tele (0.82), and CE (0.60) were each significantly superior to NT, while OSG (0.49) narrowly missed significance. The point estimates for EBP, EMC, and Tele were near-identical despite the mechanistically distinct pathways through which these platforms operate. The procedural demands of the TUG (rising from a chair, walking 3 m, turning, returning, and sitting) map directly onto the movement content targeted by all 3 platforms: balance board exergaming trains the postural transitions; motion capture training reinforces walking and turning kinematics; and Tele exercises typically incorporate functional mobility sequences supervised in real time. This task-training alignment plausibly explains why the TUG detects significant effects across a broader range of interventions than the functional scales network.

The gait speed network continued to yield no statistically significant treatment contrasts. This null result has now persisted across 2 successive NMAs, which points to a measurement problem: included trials used heterogeneous protocols (4-m vs 10-m tests, different starting conditions, single vs repeated trials), making meaningful pooling inherently difficult. Future trials should standardize gait speed assessment according to published consensus guidelines and should ideally supplement single-task gait speed with dual-task walking assessments, which may be more sensitive to the cognitive-motor integration benefits of technology-based training [96]. Until such standardization is achieved, gait-speed network estimates of this kind should not be used to draw treatment-effect conclusions.

Conventional Exercise as a Problematic Comparator

The significant global inconsistency in network 3, precisely localized by node splitting to the NT versus CE comparison (direct SMD 1.612 vs indirect SMD 0.189; P<.001), has important methodological implications. The aggregation of diverse exercise protocols into a single CE node implicitly assumes a uniform CE versus NT effect across study contexts. This assumption was violated.

The pattern of the underlying data constrains which explanations are tenable. The discrepancy is not driven by 1 or 2 aberrant trials: all 4 trials contributing direct CE versus NT evidence produced large effects (g=0.90 - 2.53; Table S7, Multimedia Appendix 1). Nor is it explained by comparator intensity, the hypothesis that the meta-regression reported earlier might be thought to support, because all 4 of these trials prescribed CE 3 times per week, the median frequency in the review. What the 4 trials do share is a structural feature: each is a small 3-arm trial (11 - 18 participants per arm) in which an NT arm ran alongside an active technology arm. Two mechanisms plausibly follow. Small-study effects would inflate all contrasts in such trials, and the direct CE versus NT evidence rests exclusively on them, whereas the indirect estimate is dominated by larger 2-arm technology versus CE trials. In addition, participants allocated to no treatment within a trial in which others visibly receive an engaging technology may disengage in a way that participants in a 2-arm design do not, depressing the NT arm and inflating every within-trial contrast against it.

The consequence is that removing the discrepant trials does not resolve the inconsistency: it attenuates every estimate, and the direct-indirect difference persists under the less aggressive threshold and becomes inestimable under the more aggressive one (Table S7, Multimedia Appendix 1). We therefore do not claim that the sensitivity analysis restored consistency, nor use it to justify the primary estimates. It supports only the narrower claim that the 2 comparisons on which our clinical conclusions rest, EBP versus NT and EMC versus NT in the TUG network, remain significant under every specification examined. The CE versus NT contrast itself is the least reliable estimate in the review, is rated down for inconsistency in GRADE accordingly, and should not be cited as this review's estimate of the effect of CE. The remedy is prospective. Future NMAs in this field should prespecify minimum comparator-intensity requirements or stratify the CE node, as emphasized in a 2024 Delphi process

XSL•FO
RenderX

[97], and trials comparing a technology with no treatment should report enough detail about the NT condition for its adequacy as a common comparator to be judged.

Acceptability

In contrast to the original review by Cieślik et al [12], which reported a marginally but significantly higher dropout probability in technology-based interventions versus controls (odds ratio 1.22, 95% CI 1.03 - 1.45; *P*=.02), the present updated analysis found no statistically significant difference in dropout rates between technology-based interventions and all control conditions (pooled odds ratio 1.10, 95% CI 0.87 - 1.40; *P*=.41; I^2=0%). This discrepancy likely reflects the expansion of the evidence base: the addition of 10 trials published since the original search, several of which reported zero or near-zero dropout, attenuated the modest signal detected in the smaller original dataset. Acceptability was comparable across all 5 technology categories, and the marginally elevated dropout risk identified in 2023 does not appear to be a stable feature of this intervention class as the literature matures.

The numerically highest crude dropout rate, observed in the Tele category (15.7%), did not differ significantly from that for CE (odds ratio 0.65, 95% CI 0.20 - 2.13) and reflects the composition of that node rather than poorer tolerability: it is driven by a single 104-week trial [22] whose comparator was no treatment, so that its attrition contributes to the crude Tele rate but not to the within-trial Tele versus CE comparison. In the trials that did compare Tele directly with CE, dropout was numerically lower in the Tele arms. A 2024 systematic review found high overall adherence (median >80%) to technology-based programs, with technology-related dropout accounting for only 3% of total withdrawals [98,99].

GRADE Certainty and Clinical Translation

This is the first formal certainty-of-evidence assessment for this technology category in older adults. The moderate ratings for EBP versus NT and EMC versus NT in the TUG network would support a conditional recommendation for EBP or EMC as options—not as proven superior alternatives—relative to no treatment for functional mobility. A recommendation of any strength, however, requires more than a certainty rating: a full GRADE Evidence-to-Decision appraisal [100] must weigh 3 further domains that this evidence base addresses only weakly.

On benefits and harms, the TUG effects for EBP and EMC (SMD approximately 0.8) are clinically appreciable, and attrition showed no excess; but harms were formally reported by only 3 (5%) of 60 trials, so that side of the balance rests on absence of data rather than evidence of safety. On values and preferences, no trial elicited older adults' preferences between technology-delivered and face-to-face exercise, and the choice is unlikely to be preference-neutral: some older adults value the social contact of supervised group exercise, whereas others value avoiding travel. As no technology exceeded CE, preference should carry substantial weight where both are available. On resource use, the technologies differ by more than an order of magnitude in acquisition, maintenance, and training costs, from consumer game consoles to bespoke sensor arrays, and no trial reported a cost outcome. In practical terms, where supervised CE is accessible and acceptable, this evidence provides no basis for substituting a technology-based program for it. These technologies are best placed to reach people for whom in-person supervised exercise is not accessible, the group for whom the comparison against no treatment, which carries the highest certainty, is the relevant one.

The predominance of low-certainty ratings reflects a structural feature of the current exercise rehabilitation literature rather than an absence of effect. As blinding of participants to an exercise intervention is not achievable, future GRADE assessments in this field would be better served by domain-specific rules that ask whether the absence of blinding plausibly affected adherence, cointerventions, outcome assessment, or selective reporting in the trial concerned, rather than by applying an automatic downgrade [25]. Applying that reasoning here, the remediable deficiencies are blinded outcome assessment and protocol prespecification, both of which are entirely feasible in this trial design.

Limitations

This review has several limitations. First, aggregation of heterogeneous exercise protocols within CE and within each technology node introduces within-node heterogeneity that the NMA model cannot fully accommodate; the CE-intensity meta-regression quantifies this for the comparator node, but analogous data were not available for every technology node. Second, IW and Tele remain underpowered relative to EBP and EMC, and the apparent elevation of Tele to first rank awaits confirmation. Third, restriction to postintervention means and SDs excluded change-score studies. Fourth, the very small number of head-to-head active-technology comparisons means that most technology versus technology estimates are wholly indirect and therefore depend on the transitivity assumption examined earlier. Fifth, several recognized effect modifiers—baseline balance impairment, fall history, frailty, sex distribution, supervision level, and technology familiarity—were too inconsistently reported to be assessed, so residual intransitivity cannot be excluded. Sixth, effect sizes were computed from analyzed rather than randomized denominators, which may introduce attrition bias in the minority of trials with substantial loss to follow-up. Finally, the gait speed network remained uninterpretable owing to extreme heterogeneity.

The absence of long-term follow-up data is a further limitation, and a clinically important one. All analyses used postintervention measurements obtained at the end of the training period; no study contributed follow-up data (eg, at 3, 6, or 12 mo after intervention cessation) to the quantitative synthesis. The present conclusions therefore apply to immediate, end-of-intervention effects and cannot be extrapolated to the durability of benefit or to long-term fall prevention outcomes. Durability is central to the clinical value of fall prevention interventions, and future trials should include follow-up of at least 6 months and report the trajectory of effects after training ends.

Related to this, the outcomes synthesized here are measures of balance, functional mobility, and gait speed, not of falls. No included trial contributed data on fall incidence, fall rate,

XSL•FO
RenderX

injurious falls, or time to first fall to any network. The interventions studied may plausibly reduce fall risk through the balance and mobility improvements demonstrated, and the instruments used are established predictors of falls, but this review provides no direct evidence that they prevent falls, and its findings should not be cited to that effect.

Implementation Considerations

Several practical considerations will determine whether these interventions can be deployed effectively in real-world aging care settings. Access to suitable devices and connectivity, digital literacy, usability for older users, the level of therapist or caregiver supervision required, data privacy, equipment and maintenance costs, and the digital divide among older adults may all influence uptake, adherence, and equity of access. These factors vary substantially across the technologies compared here and should be weighed alongside effectiveness when considering implementation, particularly for Tele and wearable systems whose real-world use may differ from the supervised conditions of the included trials.

Future Directions

The Tele evidence base most urgently requires expansion through large, rigorous RCTs that standardize the definition and intensity of remotely supervised exercise protocols. Standardizing the categorization of CE intensity across technology trial designs would substantially reduce the inconsistency problem identified in the TUG network. Trials should prioritize blinded outcome assessment, prospectively register their protocols, adopt at least 6 months of postintervention follow-up, report falls alongside balance outcomes, and report adverse events, technology-related difficulties, and participant satisfaction using standardized instruments. Two further design priorities follow from the present findings. First, instrumented or otherwise objective measurement of balance and functional mobility, applied by assessors blinded to allocation, would reduce the observer susceptibility that is inherent in timed and rated clinical tests and is the principal remediable source of risk of bias in this literature. Second, the near-absence of direct comparisons between technology categories means that almost every technology versus technology estimate reported here is wholly indirect; head-to-head trials, even of modest size, would contribute disproportionately to the precision and the certainty of the network.

Two further directions are worth noting. Individual participant data meta-analysis would overcome the ecological limitations of the study-level dose meta-regressions reported here and is the most promising route to identifying an optimal training dose. AI-adaptive exercise systems, in which machine learning continuously matches prescription to a dynamically assessed individual profile, form an emerging category that fell outside the present taxonomy because the available studies use outcome measures and designs that are not commensurable with network pooling [101-103]; incorporating them into future syntheses will require prior agreement on a core outcome set for digitally delivered exercise rather than further primary trials alone.

Conclusions

The evidence that technology-assisted rehabilitation improves balance and functional mobility in older adults has strengthened since the original 2023 review. EBP, EMC, and Tele each produced statistically significant improvements in functional mobility and overall balance versus no treatment. The emergence of Tele as the highest-ranked intervention by SUCRA is a preliminary observation resting on 5 trials with a wide CI and a probability of being best of only approximately 41%; it requires confirmation in larger trials. Moderate GRADE certainty for EBP and EMC versus no treatment on the TUG test may support their conditional use as options, rather than as proven superior alternatives, for improving functional mobility where supervised in-person exercise is not accessible; no technology exceeded CE in any network, and the low to very low certainty of the remaining 37 comparisons is insufficient to confirm effectiveness. As no included trial contributed fall outcomes, these findings should not be interpreted as direct evidence of fall prevention. Consolidating these findings will require rigorously designed, adequately powered RCTs, particularly for Tele, wearable interventions, and head-to-head technology comparisons, with standardized comparator intensity, blinded outcome assessment, and follow-up beyond the end of training.

Generative AI Use

Generative AI was used during the preparation and revision of this manuscript. Following the GAIDeT (Generative AI Delegation Taxonomy), the tasks delegated to the generative AI tool Claude were translation of text drafted by the authors into English and language editing (proofreading and polishing of style and readability). All AI-assisted text was critically reviewed, verified, and revised by the authors, who take full responsibility for the accuracy and integrity of the final manuscript. The generative AI tool is not listed as an author and bears no responsibility for the final content.

Funding

The authors declared no financial support was received for this work.

Data Availability

The risk-of-bias judgments for every included trial, the characteristics of every included trial, and the network geometry underlying all 4 networks are provided in Multimedia Appendix 1. The full arm-level dataset and the Stata and Python analysis scripts are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: FL (equal), XL (equal), DZ (equal)
Data curation: FL (equal), XL (equal), QL (supporting), KS (supporting)
Formal analysis: FL (equal), XL (equal), XG (supporting)
Investigation: FL (equal), XL (equal), QL (supporting), XG (supporting), KS (supporting)
Methodology: FL (equal), XL (equal), XF (supporting), DZ (supporting)
Project administration: DZ
Resources: DZ
Software: FL (equal), XL (equal)
Supervision: DZ (lead), XF (supporting)
Validation: QL (equal), XG (equal), YH (equal)
Visualization: FL (equal), XL (equal), KS (supporting)
Writing – original draft: FL (equal), XL (equal)
Writing – review & editing: DZ (lead), FL (supporting), XL (supporting), QL (supporting), XG (supporting), KS (supporting), YH (supporting), XF (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables and figures providing additional methodological details and results of the systematic review and network meta-analysis.
[DOCX File, 1667 KB - aging_v9i1e101231_app1.docx]

Checklist 1
PRISMA 2020 checklist.
[DOCX File, 41 KB - aging_v9i1e101231_app2.docx]

Checklist 2
PRISMA-NMA checklist.
[DOCX File, 26 KB - aging_v9i1e101231_app3.docx]

References

1. United Nations Department of Economic and Social Affairs. World population prospects 2024: summary of results. 2024 UN DESA/POP/2024/TR/no.9 URL: https://www.un.org/development/desa/pd/content/immigration-card4#1 [accessed 2026-10-12]
2. WHO global report on falls prevention in older age. WHO. 2008. URL: https://www.who.int/publications/i/item/9789241563536 [accessed 2026-10-12]
3. Chen Y, Dai F, Huang S, et al. Global, regional, and national burden of falls among older adults: findings from the Global Burden of Disease Study 2021 and projections to 2040. npj Aging 2025;11(1):85. [doi: 10.1038/s41514-025-00275-4]
4. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9). [doi: 10.1093/ageing/afac205]
5. Colón-Emeric CS, McDermott CL, Lee DS, Berry SD. Risk assessment and prevention of falls in older community-dwelling adults: a review. JAMA 2024 Apr 23;331(16):1397-1406. [doi: 10.1001/jama.2024.1416] [Medline: 38536167]
6. Moinuddin A, Goel A, Sethi Y. The role of augmented feedback on motor learning: a systematic review. Cureus 2021 Nov;13(11):e19695. [doi: 10.7759/cureus.19695] [Medline: 34976475]
7. Mansfield A, Aqui A, Fraser JE, Rajachandrakumar R, Lakhani B, Patterson KK. Can augmented feedback facilitate learning a reactive balance task among older adults? Exp Brain Res 2017 Jan;235(1):293-304. [doi: 10.1007/s00221-016-4790-6] [Medline: 27709269]
8. Park C, Lee BC. A systematic review of the effects of interactive telerehabilitation with remote monitoring and guidance on balance and gait performance in older adults and individuals with neurological conditions. Bioengineering (Basel) 2024 May 6;11(5):460. [doi: 10.3390/bioengineering11050460] [Medline: 38790328]
9. Chen Y, Zhang Y, Guo Z, Bao D, Zhou J. Comparison between the effects of exergame intervention and traditional physical training on improving balance and fall prevention in healthy older adults: a systematic review and meta-analysis. J Neuroeng Rehabil 2021 Nov 24;18(1):164. [doi: 10.1186/s12984-021-00917-0] [Medline: 34819097]

XSL•FO
RenderX

10. Ren Y, Lin C, Zhou Q, Yingyuan Z, Wang G, Lu A. Effectiveness of virtual reality games in improving physical function, balance and reducing falls in balance-impaired older adults: a systematic review and meta-analysis. Arch Gerontol Geriatr 2023 May;108:104924. [doi: 10.1016/j.archger.2023.104924] [Medline: 36680968]
11. Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol 2011 Feb;64(2):163-171. [doi: 10.1016/j.jclinepi.2010.03.016] [Medline: 20688472]
12. Cieślik B, Mazurek J, Wrzeciono A, et al. Examining technology-assisted rehabilitation for older adults' functional mobility: a network meta-analysis on efficacy and acceptability. NPJ Digit Med 2023 Aug 24;6(1):159. [doi: 10.1038/s41746-023-00907-7] [Medline: 37620411]
13. Nairn B, Tsakanikas V, Gordon B, et al. Smart wearable technologies for balance rehabilitation in older adults at risk of falls: scoping review and comparative analysis. JMIR Rehabil Assist Technol 2025 May 28;12:e69589. [doi: 10.2196/69589] [Medline: 40435383]
14. Alhasan H, Alandijani E, Bahamdan L, et al. Home-based virtual reality training for enhanced balance, strength, and mobility among older adults with frailty: systematic review and meta-analysis. JMIR Serious Games 2025 Jul 18;13:e67146. [doi: 10.2196/67146] [Medline: 40680186]
15. Brignardello-Petersen R, Murad MH, Walter SD, et al. GRADE approach to rate the certainty from a network meta-analysis: avoiding spurious judgments of imprecision in sparse networks. J Clin Epidemiol 2019 Jan;105:60-67. [doi: 10.1016/j.jclinepi.2018.08.022] [Medline: 30253217]
16. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. [doi: 10.1136/bmj.n71]
17. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med 2015 Jun 2;162(11):777-784. [doi: 10.7326/M14-2385]
18. Toledano-Shubi A, Livne D, Hel-Or H, Bahat HS. Remote fall prevention training for community-dwelling older adults: comparison with face-to-face and effect of delivery sequence-A randomized controlled trial. Geriatric Nursing 2026 Apr;69:103817. [doi: 10.1016/j.gerinurse.2026.103817]
19. Parmak D, Angın E, Iyigun G. Effects of immersive virtual reality on physical function, fall-related outcomes, fatigue, and quality of life in older adults: a randomized controlled trial. Health Care (Don Mills) 2025;13(15):1800. [doi: 10.3390/healthcare13151800]
20. Park TS, Lee JH, Jang MH, et al. Augmented reality-based exercise for older adults: a randomized controlled trial on physical function and mobility. J Aging Phys Act 2026 Oct 1;34(5):594-602. [doi: 10.1123/japa.2025-0076] [Medline: 41087006]
21. Gatica-Rojas V, Cartes-Velásquez R, Silva-Llanos J, et al. Telerehabilitation and face-to-face exergame delivery modalities to improve postural control in older adults: a randomised controlled trial. Med Sci (Basel) 2025 Nov 14;13(4):13. [doi: 10.3390/medsci13040270] [Medline: 41283271]
22. Delbaere K, Valenzuela T, Lord SR, et al. E-health StandingTall balance exercise for fall prevention in older people: results of a two year randomised controlled trial. BMJ 2021 Apr 6;373:n740. [doi: 10.1136/bmj.n740] [Medline: 33824131]
23. Deeks JJ, Higgins JPT, Altman DG, et al. Chapter 10: analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, editors. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane; 2023. [doi: 10.1002/9781119536604.ch10]
24. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol 2014 Dec 19;14:135. [doi: 10.1186/1471-2288-14-135] [Medline: 25524443]
25. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898. [doi: 10.1136/bmj.l4898]
26. White IR. Network Meta-analysis. The Stata Journal: Promoting communications on statistics and Stata 2015 Dec;15(4):951-985. [doi: 10.1177/1536867X1501500403]
27. White IR, Barrett JK, Jackson D, Higgins JPT. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. Res Synth Methods 2012 Jun;3(2):111-125. [doi: 10.1002/jrsm.1045] [Medline: 26062085]
28. White IR. Multivariate random-effects meta-regression: updates to Mvmeta. The Stata Journal: Promoting communications on statistics and Stata 2011 Jul;11(2):255-270. [doi: 10.1177/1536867X1101100206]
29. van Valkenhoef G, Lu G, de Brock B, Hillege H, Ades AE, Welton NJ. Automating network meta-analysis. Res Synth Methods 2012 Dec;3(4):285-299. [doi: 10.1002/jrsm.1054] [Medline: 26053422]
30. König J, Krahn U, Binder H. Visualizing the flow of evidence in network meta-analysis and characterizing mixed treatment comparisons. Stat Med 2013 Dec 30;32(30):5414-5429. [doi: 10.1002/sim.6001] [Medline: 24123165]
31. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med 2010 Mar 30;29(7-8):932-944. [doi: 10.1002/sim.3767] [Medline: 20213715]

XSL•FO
RenderX

32. Chaimani A, Salanti G. Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. Res Synth Methods 2012 Jun;3(2):161-176. [doi: 10.1002/jrsm.57] [Medline: 26062088]
33. Zhao C, Zhao C, Li Y, et al. The effects of active video game exercise based on self-determination theory on physical fitness and cognitive function in older adults. JCM 2022;11(14):3984. [doi: 10.3390/jcm11143984]
34. Yousefi Babadi S, Daneshmandi H. Effects of virtual reality versus conventional balance training on balance of the elderly. Exp Gerontol 2021 Oct 1;153:111498. [doi: 10.1016/j.exger.2021.111498] [Medline: 34311059]
35. Yerlikaya T, Öniz A, Özgùren M. The effect of an interactive tele rehabilitation program on balance in older individuals. Neurological Sciences and Neurophysiology 2021;38(3):180-186. [doi: 10.4103/nsn.nsn_91_21]
36. Yang CM, Chen Hsieh J(, Chen YC, Yang SY, Lin HCK. Effects of Kinect exergames on balance training among community older adults. Medicine (Baltimore) 2020;99(28):e21228. [doi: 10.1097/MD.0000000000021228]
37. Wu G, Keyes L, Callas P, Ren X, Bookchin B. Comparison of telecommunication, community, and home-based Tai Chi exercise programs on compliance and effectiveness in elders at risk for falls. Arch Phys Med Rehabil 2010 Jun;91(6):849-856. [doi: 10.1016/j.apmr.2010.01.024] [Medline: 20510973]
38. Whyatt C, Merriman NA, Young WR, Newell FN, Craig C. A Wii bit of fun: a novel platform to deliver effective balance training to older adults. Games Health J 2015 Dec;4(6):423-433. [doi: 10.1089/g4h.2015.0006] [Medline: 26469308]
39. Wang LT. Effectiveness of virtual reality exercise for functional fitness in community-dwelling older adults: a 12-week follow-up study. SAGE Open 2023 Oct;13(4):13. [doi: 10.1177/21582440231218515] [Medline: 37362769]
40. Toulotte C, Toursel C, Olivier N. Wii Fit® training vs. adapted physical activities: which one is the most appropriate to improve the balance of independent senior subjects? A randomized controlled study. Clin Rehabil 2012 Sep;26(9):827-835. [doi: 10.1177/0269215511434996] [Medline: 22324055]
41. Sonthikul C, Hadhoh N, Madeeyoh N, Ponlakarn A, Dolthamsiri N. Comparative effect of core stabilization exercise and active video gaming on dynamic balance in elderly female people: a singleblind randomized controlled clinical trial. J Health Sci Med Res 2021;40(1):53-65. [doi: 10.31584/jhsmr.2021820]
42. Singh DKA, Rajaratnam BS, Palaniswamy V, Raman VP, Bong PS, Pearson H. Effects of balance-focused interactive games compared to therapeutic balance classes for older women. Climacteric 2012 Jan;16(1):141-146. [doi: 10.3109/13697137.2012.664832]
43. Sheehy L, Bharadwaj L, Nissen KA, Estey JL. Non-immersive virtual reality exercise can increase exercise in older adults living in the community and in long-term care: a randomized controlled trial. Clin Interv Aging 2025;20:109-124. [doi: 10.2147/CIA.S498272] [Medline: 39931102]
44. Schwenk M, Grewal GS, Honarvar B, et al. Interactive balance training integrating sensor-based visual feedback of movement performance: a pilot study in older adults. J NeuroEngineering Rehabil 2014;11(1):164. [doi: 10.1186/1743-0003-11-164]
45. Schättin A, Arner R, Gennaro F, de Bruin ED. Adaptations of prefrontal brain activity, executive functions, and gait in healthy elderly following exergame and balance training: a randomized-controlled study. Front Aging Neurosci 2016;8:278. [doi: 10.3389/fnagi.2016.00278] [Medline: 27932975]
46. Sato K, Kuroki K, Saiki S, Nagatomi R. Improving walking, muscle strength, and balance in the elderly with an exergame using Kinect: a randomized controlled trial. Games Health J 2015 Jun;4(3):161-167. [doi: 10.1089/g4h.2014.0057] [Medline: 26182059]
47. Sadeghi H, Jehu DA, Daneshjoo A, et al. Effects of 8 weeks of balance training, virtual reality training, and combined exercise on lower limb muscle strength, balance, and functional mobility among older men: a randomized controlled trial. Sports Health 2021;13(6):606-612. [doi: 10.1177/1941738120986803] [Medline: 33583253]
48. Rendon AA, Lohman EB, Thorpe D, Johnson EG, Medina E, Bradley B. The effect of virtual reality gaming on dynamic balance in older adults. Age Ageing 2012 Jul;41(4):549-552. [doi: 10.1093/ageing/afs053] [Medline: 22672915]
49. Pluchino A, Lee SY, Asfour S, Roos BA, Signorile JF. Pilot study comparing changes in postural control after training using a video game balance board program and 2 standard activity-based balance intervention programs. Arch Phys Med Rehabil 2012 Jul;93(7):1138-1146. [doi: 10.1016/j.apmr.2012.01.023] [Medline: 22414490]
50. Pichierri G, Murer K, de Bruin ED. A cognitive-motor intervention using a dance video game to enhance foot placement accuracy and gait under dual task conditions in older adults: a randomized controlled trial. BMC Geriatr 2012 Dec;12(1):74. [doi: 10.1186/1471-2318-12-74]
51. Phirom K, Kamnardsiri T, Sungkarat S. Beneficial effects of interactive physical-cognitive game-based training on fall risk and cognitive performance of older adults. Int J Environ Res Public Health 2020 Aug 21;17(17):6079. [doi: 10.3390/ijerph17176079] [Medline: 32825555]
52. Park J, Yim J. A new approach to improve cognition, muscle strength, and postural balance in community-dwelling elderly with a 3-D virtual reality kayak program. Tohoku J Exp Med 2016;238(1):1-8. [doi: 10.1620/tjem.238.1]
53. Padala KP, Padala PR, Lensing SY, et al. Efficacy of Wii-Fit on static and dynamic balance in community dwelling older veterans: a randomized controlled pilot trial. J Aging Res 2017;2017:4653635. [doi: 10.1155/2017/4653635] [Medline: 28261500]
54. Moreira NB, Rodacki ALF, Costa SN, Pitta A, Bento PCB. Perceptive-cognitive and physical function in prefrail older adults: exergaming versus traditional multicomponent training. Rejuvenation Res 2021 Feb;24(1):28-36. [doi: 10.1089/rej.2020.2302] [Medline: 32443963]

XSL•FO
RenderX

55. Morat M, Bakker J, Hammes V, et al. Effects of stepping exergames under stable versus unstable conditions on balance and strength in healthy community-dwelling older adults: a three-armed randomized controlled trial. Exp Gerontol 2019 Nov;127:110719. [doi: 10.1016/j.exger.2019.110719] [Medline: 31513877]
56. Lim J, Cho JJ, Kim J, Kim Y, Yoon B. Design of virtual reality training program for prevention of falling in the elderly: a pilot study on complex versus balance exercises. Eur J Integr Med 2017 Oct;15:64-67. [doi: 10.1016/j.eujim.2017.09.008]
57. Liepa A, Tang J, Jaundaldere I, Dubinina E, Larins V. Feasibility randomized controlled trial of a virtual reality exergame to improve physical and cognitive functioning in older people. Acta Gymnica 2022;52:online. [doi: 10.5507/ag.2022.007]
58. Lee M, Son J, Kim J, Yoon B. Individualized feedback-based virtual reality exercise improves older women's self-perceived health: a randomized controlled trial. Arch Gerontol Geriatr 2015 Sep;61(2):154-160. [doi: 10.1016/j.archger.2015.06.010]
59. Lee K. Home-based exergame program to improve physical function, fall efficacy, depression and quality of life in community-dwelling older adults: a randomized controlled trial. Health Care (Don Mills) 2023;11(8):1109. [doi: 10.3390/healthcare11081109]
60. Lee K. Virtual reality gait training to promote balance and gait among older people: a randomized clinical trial. Geriatrics (Basel) 2020 Dec 22;6(1):1. [doi: 10.3390/geriatrics6010001] [Medline: 33375012]
61. Kwok BC, Pua YH. Effects of WiiActive exercises on fear of falling and functional outcomes in community-dwelling older adults: a randomised control trial. Age Ageing 2016 Sep;45(5):621-627. [doi: 10.1093/ageing/afw108] [Medline: 27496921]
62. Ku J, Kim YJ, Cho S, Lim T, Lee HS, Kang YJ. Three-dimensional augmented reality system for balance and mobility rehabilitation in the elderly: a randomized controlled trial. Cyberpsychol Behav Soc Netw 2019 Feb 1;22(2):132-141. [doi: 10.1089/cyber.2018.0261]
63. Kim SH, Cho SH. Benefits of virtual reality program and motor imagery training on balance and fall efficacy in isolated older adults: a randomized controlled trial. Medicina (B Aires) 2022;58(11):1545. [doi: 10.3390/medicina58111545]
64. Khanmohammadi R, Olyaei G, Talebian S, Hadian MR, Hossein B, Aliabadi S. The effect of video game-based training on postural control during gait initiation in community-dwelling older adults: a randomized controlled trial. Disabil Rehabil 2022 Sep;44(18):5109-5116. [doi: 10.1080/09638288.2021.1925360] [Medline: 34027754]
65. Karahan AY, Tok F, Taşkın H, Kuçuksaraç S, Başaran A, Yıldırım P. Effects of exergames on balance, functional mobility, and quality of life of geriatrics versus home exercise programme: randomized controlled study. Cent Eur J Public Health 2015 Nov;23 Suppl:S14-S18. [doi: 10.21101/cejph.a4081] [Medline: 26849537]
66. Jorgensen MG, Laessoe U, Hendriksen C, Nielsen OBF, Aagaard P. Efficacy of Nintendo Wii training on mechanical leg muscle function and postural balance in community-dwelling older adults: a randomized controlled trial. J Gerontol A Biol Sci Med Sci 2013 Jul;68(7):845-852. [doi: 10.1093/gerona/gls222] [Medline: 23114461]
67. Hou HY, Li HJ. Effects of exergame and video game training on cognitive and physical function in older adults: a randomized controlled trial. Appl Ergon 2022 May;101:103690. [doi: 10.1016/j.apergo.2022.103690] [Medline: 35066398]
68. Guede-Rojas F, Medel-Gutiérrez MJ, Cárcamo-Vargas M, et al. Effects of exergames and conventional physical therapy on functional physical performance in older adults: a randomized controlled trial. Games Health J 2023 Oct;12(5):341-349. [doi: 10.1089/g4h.2022.0194] [Medline: 37585611]
69. Gschwind YJ, Schoene D, Lord SR, et al. The effect of sensor-based exercise at home on functional performance associated with fall risk in older people – a comparison of two exergame interventions. Eur Rev Aging Phys Act 2015 Dec;12(1):11. [doi: 10.1186/s11556-015-0156-5]
70. Gschwind YJ, Eichberg S, Ejupi A, et al. ICT-based system to predict and prevent falls (iStoppFalls): results from an international multicenter randomized controlled trial. Eur Rev Aging Phys Act 2015 Dec;12(1):10. [doi: 10.1186/s11556-015-0155-6]
71. Ghous M, Masood Q, Nawaz Malik A, Afridi A, Mehmood Q. Comparison of nonimmersive virtual reality and task-oriented circuit training on gait, balance, and cognition among elderly population: a single-blind randomized control trial. Games Health J 2024 Jun;13(3):164-171. [doi: 10.1089/g4h.2022.0205] [Medline: 38512706]
72. Gallardo-Meza C, Simon K, Bustamante-Ara N, et al. Effects of 4 weeks of active exergames training on muscular fitness in elderly women. J Strength Cond Res 2022 Feb 1;36(2):427-432. [doi: 10.1519/JSC.0000000000003560] [Medline: 32483059]
73. Franco JR, Jacobs K, Inzerillo C, Kluzik J. The effect of the Nintendo Wii Fit and exercise in improving balance and quality of life in community dwelling elders. THC 2012;20(2):95-115. [doi: 10.3233/THC-2011-0661]
74. Fakhro MA, Hadchiti R, Awad B. Effects of Nintendo Wii fit game training on balance among Lebanese older adults. Aging Clin Exp Res 2020 Nov;32(11):2271-2278. [doi: 10.1007/s40520-019-01425-x] [Medline: 31802405]
75. Evangelista de Lima B, Passos GS, Youngstedt SD, Bandeira Santos Júnior LC, Gonçalves Santana M. Effects of Xbox Kinect exercise training on sleep quality, anxiety and functional capacity in older adults. J Bodyw Mov Ther 2021 Oct;28:271-275. [doi: 10.1016/j.jbmt.2021.07.029]
76. Elmas Ö, Cemali M, Livanelioğlu A. Comparison of the effects of video conference and video-based home exercise on physical performance and body composition in older adult individuals. Medicine (Baltimore) 2024;103(44):e40329. [doi: 10.1097/MD.0000000000040329]

XSL•FO
RenderX

77. Chen PJ, Penn IW, Wei SH, Chuang LR, Sung WH. Augmented reality-assisted training with selected Tai-Chi movements improves balance control and increases lower limb muscle strength in older adults: a prospective randomized trial. J Exerc Sci Fit 2020 Sep;18(3):142-147. [doi: 10.1016/j.jesf.2020.05.003] [Medline: 32514277]
78. Chan WLS, Chan CWL, Chan HHW, Chan KCK, Chan JSK, Chan OLW. A randomised controlled pilot study of a Nintendo Ring Fit Adventure™ balance and strengthening exercise program in community-dwelling older adults with a history of falls. Australas J Ageing 2024 Sep;43(3):533-544. [doi: 10.1111/ajag.13297] [Medline: 38404233]
79. Campelo AM, Weisberg A, Sheehan DP, Schneider K, Cossich VRA, Katz L. Physical and affective physical literacy domains improved after a six-week exergame exercise program in older adults: a randomized controlled clinical trial. Games Health J 2023 Oct;12(5):366-376. [doi: 10.1089/g4h.2022.0212] [Medline: 37311178]
80. Bieryla KA, Dold NM. Feasibility of Wii Fit training to improve clinical measures of balance in older adults. Clin Interv Aging 2013;8:775-781. [doi: 10.2147/CIA.S46164] [Medline: 23836967]
81. Benitez-Lugo ML, Suárez-Serrano C, Galvao-Carmona A, Vazquez-Marrufo M, Chamorro-Moriana G. Effectiveness of feedback-based technology on physical and cognitive abilities in the elderly. Front Aging Neurosci 2022;14:1050518. [doi: 10.3389/fnagi.2022.1050518] [Medline: 36438005]
82. Bao T, Noohi F, Kinnaird C, et al. Retention effects of long-term balance training with vibrotactile sensory augmentation in healthy older adults. Sensors (Basel) 2022 Apr 14;22(8):3014. [doi: 10.3390/s22083014] [Medline: 35459000]
83. Bao T, Carender WJ, Kinnaird C, et al. Effects of long-term balance training with vibrotactile sensory augmentation among community-dwelling healthy older adults: a randomized preliminary study. J NeuroEngineering Rehabil 2018 Dec;15(1):5. [doi: 10.1186/s12984-017-0339-6]
84. Bacha JMR, Gomes GCV, de Freitas TB, et al. Effects of Kinect adventures games versus conventional physical therapy on postural control in elderly people: a randomized controlled trial. Games Health J 2018 Feb;7(1):24-36. [doi: 10.1089/g4h.2017.0065] [Medline: 29239677]
85. Adcock M, Fankhauser M, Post J, et al. Effects of an in-home multicomponent exergame training on physical functions, cognition, and brain volume of older adults: a randomized controlled trial. Front Med (Lausanne) 2019;6:321. [doi: 10.3389/fmed.2019.00321] [Medline: 32047751]
86. Abd El-Kafy EM, Alayat MS, Subahi MS, Badghish MS. C-Mill virtual reality/augmented reality treadmill training for reducing risk of fall in the elderly: a randomized controlled trial. Games Health J 2024 Aug;13(4):258-267. [doi: 10.1089/g4h.2023.0162] [Medline: 38563663]
87. Bieryla KA. Xbox Kinect training to improve clinical measures of balance in older adults: a pilot study. Aging Clin Exp Res 2016 Jun;28(3):451-457. [doi: 10.1007/s40520-015-0452-y] [Medline: 26386865]
88. Gamble CJ, van Haastregt J, van Dam van Isselt EF, Zwakhalen S, Schols J. Effectiveness of guided telerehabilitation on functional performance in community-dwelling older adults: a systematic review. Clin Rehabil 2024 Apr;38(4):457-477. [doi: 10.1177/02692155231217411] [Medline: 38013415]
89. Cottrell MA, Galea OA, O'Leary SP, Hill AJ, Russell TG. Real-time telerehabilitation for the treatment of musculoskeletal conditions is effective and comparable to standard practice: a systematic review and meta-analysis. Clin Rehabil 2017 May;31(5):625-638. [doi: 10.1177/0269215516645148] [Medline: 27141087]
90. Jirasakulsuk N, Saengpromma P, Khruakhorn S. Real-time telerehabilitation in older adults with musculoskeletal conditions: systematic review and meta-analysis. JMIR Rehabil Assist Technol 2022 Sep 1;9(3):e36028. [doi: 10.2196/36028] [Medline: 36048520]
91. Cieślik B, Wrzeciono A, Mazurek J, Federico S, Szczepańska-Gieracha J, Kiper P. Balance board or motion capture? A meta-analysis exploring the effectiveness of commercially available virtual reality exergaming in enhancing balance and functional mobility among the elderly. Games Health J 2024 Dec 1;13(6):407-418. [doi: 10.1089/g4h.2023.0109]
92. Schmidt RA, Lee TD. Motor Control and Learning: A Behavioral Emphasis: Human Kinetics; 1982.
93. Carlier S, Naessens V, De Backere F, De Turck F. A software engineering framework for reusable design of personalized serious games for health: development study. JMIR Serious Games 2023;11:e40054. [doi: 10.2196/40054]
94. Podsiadlo D, Richardson S. The timed "up & go": a test of basic functional mobility for frail elderly persons. J American Geriatrics Society 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x]
95. Eusepi D, Pellicciari L, Ugolini A, et al. Reliability of the Short Physical Performance Battery (SPPB): a systematic review with meta-analysis. Eur Geriatr Med 2025 Dec;16(6):1993-2008. [doi: 10.1007/s41999-025-01277-x] [Medline: 40679769]
96. Wollesen B, Schulz S, Seydell L, Delbaere K. Does dual task training improve walking performance of older adults with concern of falling? BMC Geriatr 2017 Sep 11;17(1):213. [doi: 10.1186/s12877-017-0610-5] [Medline: 28893187]
97. Wollesen B, Herden M, Lamberti N, Giannaki CD. Defining and reporting exercise intensity in interventions for older adults: a modified Delphi process. Eur Rev Aging Phys Act 2024;21(1):3. [doi: 10.1186/s11556-024-00337-8]
98. Ashe MC, Dos Santos IK, Erome J, Grant J, Mollins J, Soh SE. Systematic review of adherence to technology-based falls prevention programs for community-dwelling older adults: reimagining future interventions. PLOS Digit Health 2024 Sep;3(9):e0000579. [doi: 10.1371/journal.pdig.0000579] [Medline: 39226315]
99. Dubbeldam R, Stemplewski R, Pavlova I, et al. Technology-assisted physical activity interventions for older people in their home-based environment: scoping review. JMIR Aging 2025 Sep 15;8:e65746. [doi: 10.2196/65746] [Medline: 40952779]

XSL•FO
RenderX

100. Morgano GP, Amato L, Parmelli E, Moja L, Davoli M, Schünemann H. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. Recenti Prog Med 2017 Oct;108(10):402-424. [doi: 10.1701/2802.28354] [Medline: 29105695]
101. Netz Y, Bar-Shalom S, Argov E, et al. A smartphone platform for remote motor fitness assessment and AI-generated personalized exercise programs for older adults: randomized controlled trial. J Med Internet Res 2025 Oct 15;27:e73145. [doi: 10.2196/73145] [Medline: 41092429]
102. Lin KC, Wai RJ, Chang Chien HY. AI-assisted dynamic postural control screening to improve functional mobility in older adult populations: quasi-experimental study. JMIR Aging 2025;8:e73290-e73290. [doi: 10.2196/73290]
103. Fan S, Tan X, Zheng H, et al. Comparing the effects of AI-assisted and traditional exercise on physical health outcomes in older adults: a systematic review and meta-analysis. Health Care (Don Mills) 2025;13(23):2999. [doi: 10.3390/healthcare13232999]

Abbreviations

BBS: Berg Balance Scale
CE: conventional exercise
COP: center of pressure
EBP: exergaming with balance platforms
EMC: exergaming with motion capture
FAB: Fullerton Advanced Balance Scale
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
IW: interventions with wearables
Mini-BESTest: Mini-Balance Evaluation Systems Test
NMA: network meta-analysis
NT: no treatment
OSG: other serious games
POMA: Performance-Oriented Mobility Assessment
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-NMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for network meta-analyses
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
RoB 2: Cochrane Risk of Bias tool, version 2
SMD: standardized mean difference
SPPB: Short Physical Performance Battery
SUCRA: surface under the cumulative ranking curve
Tele: telerehabilitation
TUG: Timed Up and Go

Edited by J Wang; submitted 13.May.2026; peer-reviewed by N Xiao, X Lyu; revised version received 02.Aug.2026; accepted 21.Sep.2026; published 28.Oct.2026.

Please cite as:
Li F, Li X, Li Q, Gong X, Shi K, Huang Y, Feng X, Zhang D
Comparative Effectiveness of Digital Exercise-Based Interventions for Improving Balance in Older Adults: Systematic Review and Network Meta-Analysis With GRADE Certainty of Evidence
JMIR Aging 2026;9:e101231
URL: https://aging.jmir.org/2026/1/e101231
doi:10.2196/101231

© Fangbo Li, Xuecheng Li, Qingfang Li, Xiaoqin Gong, Kexin Shi, Yongyu Huang, Xiaowei Feng, Dong Zhang. Originally published in JMIR Aging (https://aging.jmir.org), 28.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

GPS-Based Approaches to Assessing Psychosocial Health in Community-Dwelling Older Adults: Integrative Review

JaeWon Hyun[1,2*], PhD; Bomgyeol Kim[1*], PhD; Hun Kang[2,3], MPH; JiYeon Choi[1,2,4], PhD

[1]Mo-Im Kim Nursing Research Institute, Yonsei University College of Nursing, 50-1 Yonsei-ro, Seoul, Republic of Korea

[2]Yonsei University College of Nursing, Seoul, Republic of Korea

[3]Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA, United States

[4]Institute for Innovation in Digital Healthcare, Yonsei University, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
JiYeon Choi, PhD
Mo-Im Kim Nursing Research Institute, Yonsei University College of Nursing, 50-1 Yonsei-ro, Seoul, Republic of Korea

Abstract

Background: Psychosocial health in later life is closely related to how older adults move through and engage with their everyday environments. GPS technology enables repeated, individual-level assessment of real-world mobility, activity locations, and environmental contexts. However, GPS-based studies of psychosocial health among community-dwelling older adults have used heterogeneous devices, monitoring protocols, derived measures, and analytic approaches.

Objective: This integrative review aimed to synthesize associations between individual-level GPS-derived measures or GPS-anchored environmental exposures and psychosocial health, operationalized as psychological health, social health, and evaluative well-being, among community-dwelling older adults. It also aimed to characterize GPS devices, monitoring protocols, and measures. Finally, it sought to identify methodological limitations and evidence gaps.

Methods: Systematic searches were conducted in PubMed, CINAHL, Embase, Web of Science, and the Cochrane Library from database inception to July 10, 2026. Studies were eligible if they included community-dwelling adults aged 60 years or older; used repeated, individual-level GPS coordinates or trajectories to derive or anchor measures of mobility, life space, activity patterns, or environmental exposure; and analyzed these measures in relation to at least one psychosocial health indicator. For narrative synthesis, psychosocial indicators were grouped into psychological health, social health, and evaluative well-being.

Results: Of 1499 records identified, 15 reports representing 13 independent studies were included. GPS-based measures were classified into three nonmutually exclusive domains: (1) mobility and life space measures, (2) activity locations and patterns, and (3) GPS-anchored environmental exposure. Across studies, associations between GPS-based measures and psychosocial indicators were heterogeneous. Greater mobility, including more outdoor time, greater travel distance, more frequent trips, and larger activity spaces, was associated with more favorable indicators of psychological health, social health, or evaluative well-being in several reports, although null findings were also common. Some activity-location and movement-pattern measures showed associations with depressive symptoms, perceived social isolation, life satisfaction, and relaxation. GPS-anchored environmental exposure was examined in 3 independent studies. Substantial heterogeneity was observed in GPS devices, monitoring duration, sampling intervals, valid-data criteria, data-processing procedures, and missing-data handling.

Conclusions: Repeated, individual-level GPS data can characterize real-world mobility, activity patterns, and environmental exposure in relation to psychosocial health among community-dwelling older adults. However, associations were heterogeneous and were based on observational analyses. Future studies should prioritize adequately powered longitudinal designs, repeated and temporally aligned assessments, and standardized and transparent GPS data processing.

Trial Registration: PROSPERO CRD420251157640; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251157640

(JMIR Aging 2026;9:e103996) doi:10.2196/103996

KEYWORDS

digital technology; GPS; older adult; psychosocial health; life space; integrative review

Introduction

Healthy aging is increasingly understood as a multidimensional process shaped by physical, psychological, social, and environmental factors [1]. Psychological functioning, social relationships, and individuals' evaluations of their quality of life and well-being are important aspects of healthy aging and have been associated with functioning, social participation, and

XSL•FO
RenderX

life satisfaction in later life [2-4]. Understanding behavioral and contextual factors associated with these domains is therefore important for research on psychosocial health in later life.

Mobility in later life has been linked to health, independence, and social connectedness [5]. For example, restricted life space, referring to a limited geographic extent of movement in daily life, has been associated with depressive symptoms among community-dwelling older adults [6]. Mobility provides opportunities to access services, maintain social relationships, and participate in meaningful activities [5], and may therefore be associated with psychological, social, and evaluative well-being. Accordingly, mobility and life space provide important behavioral contexts for examining psychosocial health in older adults. However, repeated patterns of mobility in daily life can be challenging to characterize using retrospective self-report measures alone.

GPS technology provides an approach for characterizing daily-life mobility through repeated measurement of individual-level movement. GPS data can be used to derive measures such as travel distance, time spent outside the home, activity-space extent (the geographic area covered by an individual's recorded locations), the number of visited locations, and the regularity or variability of movement across observations [7,8]. By recording where and when movement occurs, GPS data can complement retrospective self-report measures by providing spatial and temporal information on daily mobility. GPS data can also be linked to environmental information to characterize exposures encountered along daily mobility paths, such as green or blue spaces, facilities, and road-network characteristics [9,10]. Thus, GPS-derived measures can characterize individual mobility and activity patterns, whereas GPS-anchored environmental measures can characterize environmental contexts linked to GPS-observed locations or mobility paths.

GPS-based studies have examined associations of these behavioral and contextual measures with psychological health, social health, and evaluative well-being among older adults [9-13]. However, GPS-based health studies have differed substantially in device type, monitoring duration, sampling frequency, valid-data criteria, preprocessing procedures, missing-data handling, and the operationalization of derived measures [7,14]. GPS data may also be incomplete because of signal loss, battery issues, or participant noncompliance [14]. Existing reviews have focused on GPS-based mobility measurement among community-dwelling older adults, mobility and well-being in later life, GPS data collection and processing, or associations between GPS mobility features and depressive symptoms across adult populations [5,7,14,15]. However, to our knowledge, no review has jointly synthesized repeated, individual-level GPS-derived and GPS-anchored environmental measures in relation to psychosocial health among community-dwelling older adults. In this review, psychosocial health was operationalized as encompassing 3 related domains: psychological health, social health, and evaluative well-being, reflecting psychological functioning, social relationships, and evaluations of life satisfaction and quality of life [16,17]. Therefore, this integrative review aimed to (1) synthesize evidence on associations between individual-level GPS-derived or GPS-anchored environmental measures and psychosocial health indicators, specifically, psychological health, social health, and evaluative well-being, among community-dwelling older adults; (2) characterize GPS devices and monitoring protocols, as well as the GPS-derived or GPS-anchored environmental measures used; and (3) identify methodological limitations and evidence gaps.

Methods

Study Design

This integrative review was conducted based on the methodological framework proposed by Whittemore and Knafl [18]. An integrative review approach was selected because the literature on GPS-based measures and psychosocial health in older adults encompasses heterogeneous study designs, monitoring protocols, derived measures, and outcome domains. The Whittemore and Knafl framework provides a systematic approach to synthesizing evidence across such diverse methodologies. The review protocol was registered in the PROSPERO (International Prospective Register of Systematic Reviews; CRD420251157640). This review was reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [19]. The PRISMA checklist is provided in Checklist 1.

Search Strategies

Systematic searches were conducted in PubMed, CINAHL, Embase, Web of Science, and the Cochrane Library from database inception to July 10, 2026. The search strategies were developed in consultation with a medical librarian, using database-specific subject headings and keywords related to GPS technology, psychosocial health, and older adults. In this review, psychosocial health was operationalized as comprising 3 domains: psychological health, social health, and evaluative well-being. Psychological health included mental health conditions and symptoms, psychological distress, and emotional or affective states; social health included social connectedness, social isolation, loneliness, and social engagement; and evaluative well-being included life satisfaction and perceived or health-related quality of life. Detailed search strategies for each database are available in Multimedia Appendix 1.

Eligibility Criteria

Studies were included if they (1) used repeated, individual-level GPS coordinates or trajectories to derive measures of real-world locations, mobility, life space, activity patterns, or to link environmental attributes to GPS-observed locations or mobility paths; (2) included community-dwelling older adults aged 60 years or older, or findings for this age group were separately extractable; (3) assessed at least one indicator of psychological health, social health, or evaluative well-being; (4) analyzed GPS-derived or GPS-anchored environmental measures in relation to at least one psychosocial health indicator; and (5) were published in peer-reviewed journals. In this review, older adults were defined as individuals aged 60 years or older, in accordance with the United Nations standard used by the World Health Organization [20].

XSL•FO
RenderX

Studies were excluded if they (1) used geospatial measures without repeated, individual-level GPS data; (2) used GPS solely for mapping, visualization, adherence monitoring, assessment triggering, diary validation, or retrospective reconstruction; (3) did not directly analyze GPS-derived or GPS-anchored environmental measures in relation to psychosocial health indicators; (4) were conducted exclusively in inpatient or institutional settings; (5) were conference abstracts, reviews, study protocols, editorials, case reports, or unpublished theses and dissertations; (6) were not published in English. After protocol registration, the review scope was further refined to include peer-reviewed journal articles only, and the outcome scope was broadened from psychological and emotional health to psychosocial health. For narrative synthesis, psychosocial outcomes were categorized into psychological health, social health, and evaluative well-being.

Study Screening and Selection

After completing database searches, all identified records were imported into EndNote X9 (Clarivate Analytics), where duplicate entries based on titles, authors, and publication year were removed. Two researchers (JH and HK) independently screened titles and abstracts according to the predefined eligibility criteria. Subsequently, full-text screening was conducted. Any discrepancies were reviewed by a third researcher (BK), who examined the full text and facilitated consensus through discussion among the review team. Final selection of reports was determined through this consensus-based process.

Quality Appraisal of the Selected Studies

The quality of the selected reports was evaluated using the Mixed Methods Appraisal Tool (MMAT [21]). The MMAT is designed to assess the quality of various study designs, including quantitative, qualitative, and mixed methods studies and includes 5 criteria tailored to each study design. Each criterion was rated as “yes,” “no,” or “can’t tell.” Two researchers (JH and HK) independently assessed the quality of each study, and disagreements were resolved through discussion with the third researcher (BK). We summarized overall study quality using asterisks (*), ranging from none (fulfilling none of the criteria) to ***** (fulfilling all 5 criteria).

Data Extraction and Synthesis

Data extraction was conducted using a standardized form in Microsoft Excel. Two research assistants independently extracted data from each included report, and 2 researchers (JH and HK) reviewed and validated all extracted information. Any discrepancies were resolved through discussion with the third researcher (BK). Extracted data included study characteristics (author, publication year, country, study aims, sample size, and participant age); GPS devices and monitoring protocols; GPS-derived or GPS-anchored environmental measures and their operational definitions; psychosocial health indicators; reported associations; and methodological and technological limitations. Companion reports originating from the same parent study were also identified during data extraction. GPS-derived measures refer to variables calculated directly from repeated GPS coordinates or trajectories, whereas GPS-anchored environmental measures refer to variables obtained by linking environmental attributes to GPS-observed locations or mobility paths.

Data synthesis followed the 5-step process described by Whittemore and Knafl [18]. First, extracted data were organized in Microsoft Excel (*data reduction*). Specifically, psychosocial indicators were initially extracted as reported in each report and subsequently categorized into 3 domains: psychological health, social health, and evaluative well-being. Second, the reduced data were summarized into tables describing study characteristics, GPS devices and monitoring protocols, GPS-derived or GPS-anchored environmental measures, psychosocial health indicators, and reported associations (*data display*). Third, findings were compared across reports to identify similarities, differences, and patterns (*data comparison*). Fourth, the findings were integrated and interpreted across the 3 psychosocial health domains (*conclusion drawing*). Fifth, synthesized findings were reviewed and verified through iterative discussions among the research team (*verification*).

The report was the unit of extraction, whereas companion reports from the same parent study were linked and counted as a single independent study when describing and interpreting the evidence base.

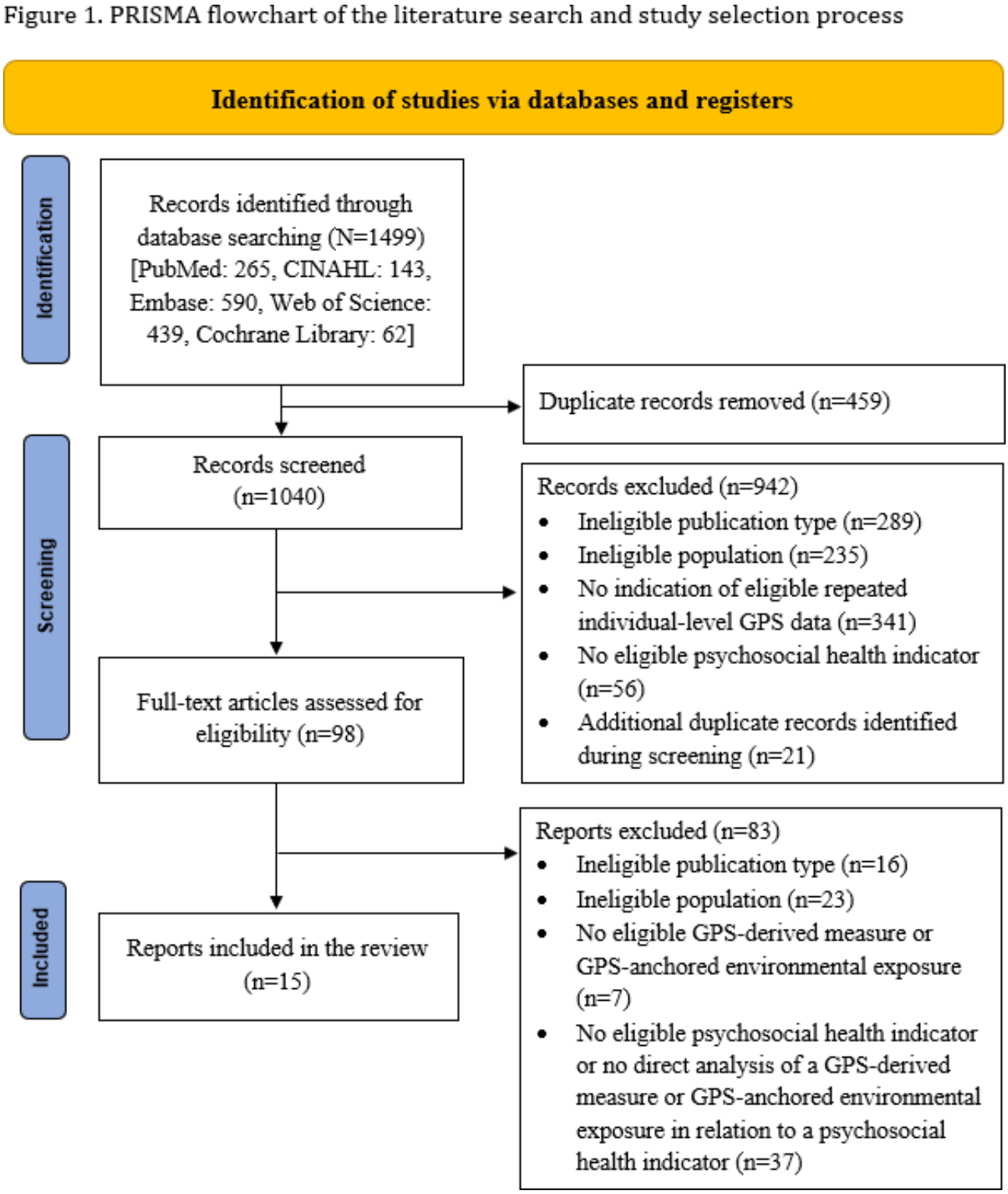
Results

Selection of Studies

Across the 5 databases, 1499 records were identified: 265 from PubMed, 143 from CINAHL, 590 from Embase, 439 from Web of Science, and 62 from the Cochrane Library. After removing 459 duplicates, 1040 records remained for title and abstract screening, during which 942 records were excluded. Subsequently, 98 reports were assessed through full-text review, and 83 were excluded. Consequently, 15 reports representing 13 independent studies were included in the review. Figure 1 summarizes the literature search and the study selection process.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the literature search and study selection process.



Characteristics of Included Studies

The 15 included reports represented 13 independent studies and were conducted across 4 geographic regions: Asia (Hong Kong, Japan, and China), Europe (France, Germany, and Switzerland), North America (Canada and the United States), and Oceania (Australia). Sample sizes varied substantially across studies, ranging from 8 to 279 participants, although analysis-specific sample sizes were not consistently reported.

Two companion reports from one parent study, focused on older adults with cognitive impairment [12,22], were included, and detailed characteristics and key findings of the included reports are provided in Multimedia Appendix 2.

All analyses relevant to this review were observational. Although 4 reports used secondary data from 2 parent randomized controlled trials, they did not compare randomized groups or estimate randomized intervention effects [12,22-24]. Accordingly, all 15 reports were appraised using the MMAT

XSL•FO
RenderX

criteria for quantitative nonrandomized studies. Based on the MMAT appraisal, 6 reports met 80% of the criteria (****), 5 reports met 60% (***), 3 reports met 40% (**), and 1 report met 20% (*). Detailed results of the quality appraisal for each report are presented in Multimedia Appendix 3.

GPS Technologies, Monitoring Protocols, and GPS-Derived or GPS-Anchored Environmental Measures

Across the 13 independent studies, GPS data were collected using dedicated portable GPS loggers or receivers (n=6 independent studies, represented by 8 reports) [9,11,12,22-26], smartphone-based GPS (n=5 independent studies) [10,27-30], or wearable GPS devices (n=2 independent studies) [13,31]. Ten reports collected complementary data using accelerometers, pedometers, inertial sensors, ecological momentary assessment, diaries, or other sensing devices alongside GPS data [9-13,23-25,27,28].

Monitoring protocols varied in assessment duration, sampling frequency, and criteria for valid GPS data. Eight reports collected GPS data over approximately 6 to 15 days [9-11,13,22-24,28], and 4 used monitoring periods of approximately 4 to 8 weeks [26,27,29,30]. One report used three daytime monitoring days [31], one assessed a single park visit [25], and one did not report the total GPS monitoring duration [12]. Where reported, sampling intervals ranged from 1 Hz [13,29] to 1 location fix every 2 minutes [10]. The raw GPS sampling interval was not reported in 2 reports [24,28]. Criteria for valid data included minimum daily monitoring time, a minimum number of valid days, and requirements for valid home and activity-location data [11-13,22-24]. Detailed GPS technologies and monitoring protocols are presented in Multimedia Appendix 4.

The measures were classified into 3 nonmutually exclusive domains: (1) GPS-derived mobility and life space, (2) GPS-derived activity locations and patterns, and (3) GPS-anchored environmental exposure. Mobility and life space measures were examined in 11 of the 13 independent studies, represented by 13 reports, including outdoor or out-of-home time, travel distance, activity-space area, maximum distance from home, and trip frequency and duration [10-13,22-24,26-31]. Eight independent studies examined activity locations and patterns, including visited destinations, location entropy (a measure of how evenly time is distributed across visited locations), significant locations, route diversity and familiarity, and park-use patterns [9-11,13,25,26,29,30]. Three independent studies examined GPS-anchored environmental exposures, including facilities and road intersections, green- and blue-space characteristics, microurban features encountered along mobility paths, and park-zone attributes [9,10,25].

Psychosocial Health Indicators

Psychosocial health indicators examined in relation to GPS-derived measures or GPS-anchored environmental exposures were organized into three domains: (1) psychological health, (2) social health, and (3) evaluative well-being. Studies could contribute to more than one domain.

Psychological health indicators were examined in 13 reports representing 11 independent studies [9,10,12,13,22-26,28-31]. Depressive symptoms were the most frequently assessed indicator and were examined in 11 reports representing 9 independent studies [9,12,22-26,28-31]. Other indicators included anxiety, relaxation, and contentment [25]; psychological well-being [12,22]; positive and negative affect [13]; happiness and calmness [10]; and fear of falling [23,24].

Social health indicators were examined across 4 reports representing 4 independent studies. These indicators included loneliness [11,28], perceived social isolation [27,29], social interactions [11], perceived help availability and social network [28], sense of belonging [27], and community participation [29].

Evaluative well-being indicators were examined across 4 reports representing 4 independent studies, including life satisfaction [13], daily satisfaction [10], perceived quality of life [23], and health-related quality of life [11].

Psychosocial Health in Relation to Mobility and Life Space Measures

Associations between mobility and life space measures and psychosocial health indicators were heterogeneous, with both significant and null findings across reports. Association directions, participant sample sizes, and covariate-adjustment status are summarized by GPS measure domain and psychosocial health domain in Multimedia Appendix 5.

For psychological health, longer outdoor or out-of-home time was associated with fewer depressive symptoms [12,26], whereas 12-month change in outdoor time was not associated with changes in depressive symptoms [22]. Outdoor time was not associated with psychological well-being [12], and 12-month change in outdoor time was not associated with changes in psychological well-being [22]. Spending at least 30 minutes outdoors per day was associated with fewer depressive symptoms and less fear of falling [23], and greater pedestrian trip frequency, distance, and duration were associated with fewer depressive symptoms and less fear of falling [24]. In the study by Giannouli et al [28], GPS-derived life space measures were not associated with depressive symptoms. In the study by Chung et al [31], no significant associations were observed in unadjusted analyses; however, after demographic adjustment, 95th-percentile movement speed was positively associated with depressive symptoms. Positive affect was not reliably associated with daily mobility measures [13]. In the study by Wang et al [10], activities occurring within the residential neighborhood were associated with lower momentary happiness.

For social health, a greater number of out-of-home trips was associated with more social interactions but not with loneliness [11]. Greater distance traveled from home was correlated with a stronger sense of belonging [27], and maximum action range showed a small association with perceived help availability but not with loneliness or social networks [28]. A relative-mobility component was also associated with community participation [29].

For evaluative well-being, outdoor time was not associated with perceived quality of life [23], and out-of-home trips were not associated with health-related quality of life [11]. In contrast,

XSL•FO
RenderX

a larger activity-space ellipse (an ellipse summarizing the spatial distribution of recorded locations) was associated with greater daily satisfaction [10], and larger daily convex-hull areas (the areas of the smallest convex polygons enclosing the recorded GPS locations) and more passive-transport time were associated with higher daily life satisfaction [13].

Psychosocial Health in Relation to Activity Locations and Patterns

Associations with activity-location and movement-pattern measures also varied across psychosocial health indicators. For psychological health, higher depressive-symptom scores were correlated with fewer unique driving routes [30], whereas no association between GPS-derived location-pattern measures and depressive symptoms remained after correction for multiple comparisons in the study by Hackett et al [29]. The between-person association between location entropy and negative affect was not robust after adjustment [13]. Among older park users, GPS-derived park-use classes differed in postvisit relaxation, with active park lingerers reporting greater relaxation than active walkers [25].

For social health, lower mean location diversity and lower day-to-day variability in location diversity were associated with greater COVID-19–related social isolation [29], whereas destination-specific trip counts were not associated with loneliness [11].

For evaluative well-being, greater daily location entropy was associated with higher life satisfaction at the within-person level [13], whereas destination-specific trip counts were not associated with health-related quality of life [11].

Psychosocial Health in Relation to GPS-Anchored Environmental Exposures

GPS-anchored environmental exposures were examined in 3 independent studies, whereas direct associations with psychological health indicators were reported in 2. Exposure to water and to commercial, leisure, or cultural attractors along daily mobility paths was associated with better momentary mental well-being in selected models, and selected attractors and walkable paths were also associated with better momentary mental well-being in within-person outdoor analyses [9]. In the study by Wang et al [10], selected GPS-anchored environmental attributes, including leisure facilities, road intersections, and the blue-view index, were associated with momentary happiness or calmness. Li et al [25] also characterized GPS-anchored environmental attributes within the park setting; however, the reported psychological health comparison was based on GPS-derived park-use classes rather than a direct environmental exposure-outcome association.

Methodological and Technological Limitations

Methodological and technological limitations were identified across the included reports. Several reports did not fully specify key GPS protocol elements, including total monitoring duration, raw sampling interval, valid-day criteria, participant-level data-completeness thresholds, home-geofence definitions, and trip definitions [9-12,23-25,27-29]. The application of valid-data criteria reduced analytic samples in 8 reports [11-13,22,23,25,30,31]. Three reports excluded, imputed, or reconstructed missing or incomplete GPS observations using different procedures [9,13,29].

GPS measurement and processing procedures introduced additional limitations. Data collection in 4 reports depended on participants manually activating GPS recording or carrying a study or personal smartphone, creating the potential for unrecorded observation periods [27,29-31]. Travel modes and trips were not consistently distinguished or fully defined, and classification was potentially subject to misclassification in 5 reports [11,12,22,24,28]. In the 3 studies of GPS-anchored environmental exposure, limitations included reconstructed location points and potential indoor-outdoor misclassification, incompletely reported interpolation or missing-data handling, and incomplete GPS coverage [9,10,25].

Three reports used small analytic samples or conducted exploratory or unadjusted analyses, limiting the precision and interpretability of the reported associations [27,30,31]. Feasibility and acceptability of GPS-based monitoring were explicitly evaluated in 3 reports using different measures [11,29,31]. Reported issues included variable device-use time, incomplete GPS records, additional charging requirements, and occasional technical problems, although Gough et al [11] and Hackett et al [29] generally reported favorable protocol completion and participant experiences.

Discussion

Principal Findings

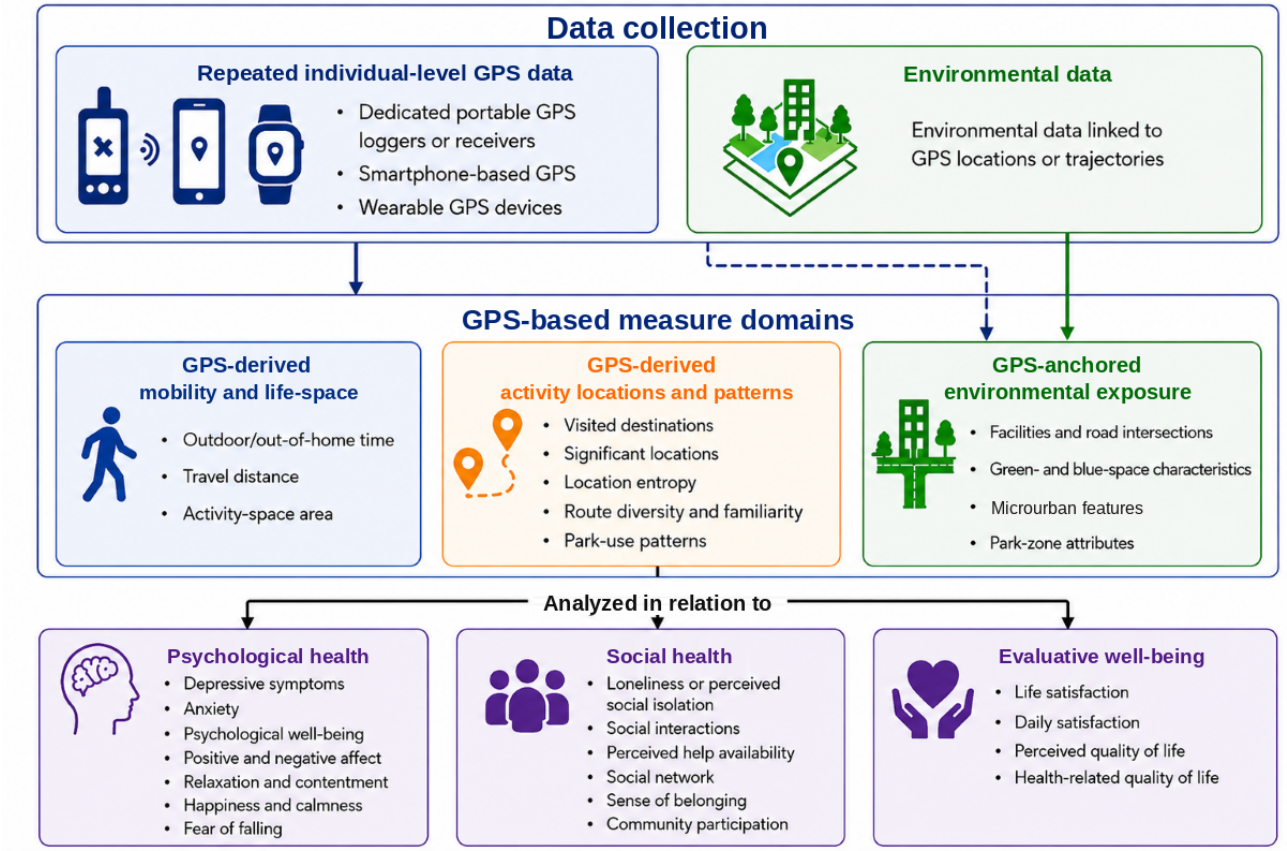
In this integrative review, we synthesized evidence from 15 reports representing 13 independent studies that used repeated, individual-level GPS data to examine psychosocial health among community-dwelling older adults. Overall, associations between GPS-based measures and psychosocial health indicators were heterogeneous, with both significant and nonsignificant findings reported across psychological health, social health, and evaluative well-being outcomes. GPS-derived mobility and life space measures were the most frequently examined indicators, whereas relatively few studies examined GPS-derived activity locations and patterns or GPS-anchored environmental exposures. Comparisons across studies were further limited by substantial variation in GPS devices, monitoring protocols, outcome measures, and analytic approaches, highlighting important methodological limitations and evidence gaps in the current literature [7,14,15].

To organize the evidence, GPS-based measures were classified into 3 nonmutually exclusive domains: GPS-derived mobility and life space, GPS-derived activity locations and patterns, and GPS-anchored environmental exposure. Consistent with our conceptual framework, psychosocial health outcomes were synthesized within 3 domains: psychological health, social health, and evaluative well-being. Figure 2 illustrates the relationships among repeated, individual-level GPS data, the 3 GPS measure domains, and the 3 psychosocial health domains examined in this review. This framework extends previous reviews focused primarily on GPS-based mobility measures and depressive symptoms by providing a broader synthesis of

XSL•FO
RenderX

psychosocial health outcomes among community-dwelling older adults.

Figure 2. Conceptual summary of data sources, GPS-based measure domains, and the 3 psychosocial health domains examined among community-dwelling older adults.



The mobility and life space domain was the most frequently examined GPS measure domain, assessed in 11 independent studies represented by 13 reports. Several reports found associations between greater mobility, including more time spent outdoors, greater travel distance, more frequent trips, or larger activity spaces, and fewer depressive symptoms, more frequent social interactions, a stronger sense of belonging, or greater life or daily satisfaction [10-13,23,24,26,27]. However, findings varied across GPS measures, psychosocial indicators, and assessment periods, and several reports also reported null associations. One report that examined change over time found that changes in outdoor time over 12 months were not associated with changes in depressive symptoms or psychological well-being [22]. Taken together, these findings do not support interpreting GPS-derived mobility measures as direct proxies for psychosocial health. The predominantly cross-sectional evidence also precludes conclusions about temporal direction. A recent meta-analysis similarly found small-to-moderate between-person associations between several GPS-derived mobility and activity features and depressive symptoms in adults but insufficient evidence to synthesize within-person associations [15]. Future studies should therefore combine repeated GPS monitoring with repeated psychosocial assessments and within-person analyses to better characterize changes over time.

Activity locations and patterns were examined in 8 independent studies [9-11,13,25,26,29,30]. Measures such as location entropy, location diversity, route diversity, and park-use patterns captured aspects of the spatial distribution, diversity, and organization of daily activity beyond mobility volume alone. Associations were reported with depressive symptoms, perceived social isolation, life satisfaction, and relaxation [13,25,29,30]. However, the evidence was limited, and some findings were based on small samples, used unadjusted analyses, or did not remain significant after correction for multiple testing. Evidence regarding social health was particularly limited, with only 2 of the 8 studies in this GPS measure domain examining social health indicators [11,29]. Thus, although GPS-derived activity-location and movement-pattern measures may provide information about daily behavior beyond mobility volume, current evidence is insufficient to support their use as validated indicators of psychosocial health.

GPS-anchored environmental exposure was the least frequently examined domain, addressed in only 3 independent studies. Fancello et al [9] and Wang et al [10] directly examined associations between environmental characteristics encountered during daily mobility and psychosocial health indicators, whereas Li et al [25] characterized environmental attributes within a park setting but evaluated psychosocial differences primarily across GPS-derived park-use classes. In the studies by Fancello et al [9] and Wang et al [10], time-stamped GPS locations or trajectories were linked to environmental data to characterize exposures encountered during daily mobility. Such approaches can incorporate the spatial and temporal contexts in which exposure occurs rather than relying solely on residential location [32]. The small evidence base and observational study

XSL•FO
RenderX

designs do not support causal interpretation. Future studies should more clearly define how GPS trajectories are linked to environmental data, including the spatial buffers and time windows used to define exposure.

GPS devices, monitoring periods, sampling intervals, valid-data criteria, preprocessing procedures, and missing-data handling varied substantially across reports. Several reports did not fully specify one or more key GPS protocol elements; valid-data requirements reduced the analytic sample in 8 reports; and incomplete observations were excluded, imputed, or reconstructed using different procedures in 3 reports. These differences limited direct comparisons across studies. Previous reviews have similarly identified substantial heterogeneity in GPS data collection, processing, and reporting practices [7,14]. Greater consistency in reporting GPS devices or applications, sampling intervals, monitoring schedules, valid-day and valid-participant criteria, preprocessing procedures, missing-data handling, and derivation of GPS measures would facilitate comparisons across studies. Reporting the number of participants and observation days retained at each stage of GPS processing and analysis would further improve transparency. Analyses of repeated GPS data should also account for within-person dependence and, where applicable, multiple testing.

Feasibility or acceptability of GPS-based monitoring was evaluated in only 3 reports using different measures, limiting comparisons across studies [11,29,31]. The limited reporting of these outcomes in the remaining reports should not be interpreted as evidence that GPS-based assessment was burden-free. Four reports depended on participants manually activating GPS recording or consistently carrying a study or personal smartphone, creating the possibility of unrecorded observation periods [27,29-31]. The extent to which these requirements affected participant burden, adherence, or retention was not directly evaluated. Future studies should report participant burden, device-related difficulties, technical support provided, and reasons for missing GPS data or study withdrawal.

The collection of detailed GPS trajectories raises privacy and data-governance concerns, including how location data are accessed, stored, retained, and shared. GPS trajectories can reveal home locations, frequently visited places, daily routines, and visits to potentially sensitive locations. Relevant ethical considerations include informed consent, data security, transparency regarding secondary use, equity, and data ownership [33,34]. GPS-based studies should therefore minimize the collection and retention of identifiable location information where feasible; clearly communicate how location data will be collected, processed, stored, used, and shared; and implement appropriate safeguards for data access, transfer, retention, and sharing.

Overall, the included studies used repeated, individual-level GPS data to characterize daily mobility, activity locations and patterns, and GPS-anchored environmental exposures in relation to psychosocial health among community-dwelling older adults. However, the current evidence is insufficient to establish these measures as validated indicators of psychological health, social health, or evaluative well-being. Further research should use adequately powered longitudinal designs with repeated and temporally aligned GPS and psychosocial assessments, adjust for relevant confounders, report GPS data collection and processing transparently, and replicate findings across diverse populations and geographic settings. These approaches may clarify whether and under what conditions GPS-based measures provide information about psychosocial health beyond conventional self-report measures and residentially defined environmental exposures.

Limitations

This review has several limitations. First, psychosocial health was operationalized as psychological health, social health, and evaluative well-being. Other conceptualizations and domains of psychosocial health were outside the scope of this review and were not systematically synthesized. Because studies without repeated, individual-level GPS data were excluded, evidence on environmental exposures assessed solely from residential location was not synthesized. Second, substantial heterogeneity in GPS devices, monitoring protocols, valid-data criteria, data-processing procedures, derived measures, and analytic approaches limited direct comparisons across studies. Incomplete reporting and differences in the handling of missing or invalid GPS data further limited assessment of measurement quality and comparability. Third, all analyses relevant to this review were observational, limiting causal inference. Several reports also relied on small samples or exploratory, unadjusted, or bivariate analyses, limiting the precision and robustness of the synthesized evidence. Some nonsignificant associations, particularly those estimated in small samples, may reflect limited statistical power rather than the absence of an association. Fourth, the evidence base was geographically concentrated and included limited representation of rural, socioeconomically disadvantaged, and mobility-limited older populations, restricting the generalizability of the findings. Finally, after protocol registration, the review scope was refined to include peer-reviewed journal articles only; therefore, relevant gray literature may not have been captured.

Conclusions

This integrative review synthesized evidence from studies using repeated, individual-level GPS data to examine psychological health, social health, and evaluative well-being among community-dwelling older adults. GPS-derived measures were used to characterize mobility and life space as well as activity locations and patterns, whereas GPS-anchored environmental measures were used to characterize exposures encountered during daily mobility; however, associations with psychosocial health indicators were heterogeneous and based on observational analyses. Future research should prioritize adequately powered longitudinal designs, repeated and temporally aligned GPS and psychosocial assessments, standardized and transparent GPS data processing, protection of identifiable location data, and inclusion of diverse older populations to determine the validity, interpretability, and potential utility of these measures in aging research.

XSL•FO
RenderX

Generative AI Use

The authors declare that generative AI (GenAI) was used in the creation of this manuscript. GenAI (ChatGPT; OpenAI) was used to assist with the visual rendering of a figure based on an author-developed draft, as well as proofreading, editing, or language polishing; the authors reviewed and revised the AI-assisted output and take full responsibility for the final figure. The authors verified all outputs of the tool(s) used in the preparation of this manuscript and take full responsibility for the final content of the manuscript.

Funding

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (numbers RS-2020-NR049581 and RS-2022-NR075577).

Data Availability

All data supporting the findings of this review are available within the article and its multimedia appendices.

Authors' Contributions

JH designed the study, searched and selected the articles, performed the quality assessment, analyzed and interpreted the articles, and drafted and revised the manuscript. BK designed the study, analyzed and interpreted the articles, and revised the manuscript. HK searched and selected the articles, performed the quality assessment, analyzed and interpreted the articles, and drafted the manuscript. JC designed the study, analyzed and interpreted the articles, and revised the manuscript for important intellectual content. All authors made substantial contributions to the study and approved the submitted version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies.
[DOCX File, 28 KB - aging_v9i1e103996_app1.docx]

Multimedia Appendix 2
Characteristics and key findings of the included reports (n=15).
[DOCX File, 26 KB - aging_v9i1e103996_app2.docx]

Multimedia Appendix 3
Results of quality appraisal of the included reports (n=15).
[DOCX File, 23 KB - aging_v9i1e103996_app3.docx]

Multimedia Appendix 4
Summary of GPS technologies, monitoring protocols, and GPS-derived or GPS-anchored environmental measures in the included reports (n=15).
[DOCX File, 26 KB - aging_v9i1e103996_app4.docx]

Multimedia Appendix 5
Associations between GPS measures and psychosocial health indicators (n=15).
[DOCX File, 37 KB - aging_v9i1e103996_app5.docx]

Checklist 1
PRISMA checklist.
[DOCX File, 25 KB - aging_v9i1e103996_app6.docx]

References

1. Piriu AA, Bufali MV, Cappellaro G, Compagni A, Torbica A. Conceptualisation and measurement of healthy ageing: insights from a systematic literature review. Soc Sci Med 2025 Jun;374:117933. [doi: 10.1016/j.socscimed.2025.117933] [Medline: 40198967]
2. Kim ES, Tkatch R, Martin D, MacLeod S, Sandy L, Yeh C. Resilient aging: psychological well-being and social well-being as targets for the promotion of healthy aging. Gerontol Geriatr Med 2021;7:23337214211002951. [doi: 10.1177/23337214211002951] [Medline: 33816707]

XSL•FO
RenderX

3. Reiner A, Steinhoff P. The association of social networks and depression in community-dwelling older adults: a systematic review. Syst Rev 2024 Jun 20;13(1):161. [doi: 10.1186/s13643-024-02581-6] [Medline: 38902787]
4. Sulandari S, Coats RO, Miller A, Hodkinson A, Johnson J. A systematic review and meta-analysis of the association between physical capability, social support, loneliness, depression, anxiety, and life satisfaction in older adults. Gerontologist 2024 Nov 1;64(11):gnae128. [doi: 10.1093/geront/gnae128] [Medline: 39233622]
5. Pantelaki E, Maggi E, Crotti D. Mobility impact and well-being in later life: a multidisciplinary systematic review. Res Transp Econ 2021 May;86:100975. [doi: 10.1016/j.retrec.2020.100975]
6. Polku H, Mikkola TM, Portegijs E, et al. Life-space mobility and dimensions of depressive symptoms among community-dwelling older adults. Aging Ment Health 2015;19(9):781-789. [doi: 10.1080/13607863.2014.977768] [Medline: 25376479]
7. Chung J, Sargent L, Brown R, Gendron T, Wheeler D. GPS tracking technologies to measure mobility-related behaviors in community-dwelling older adults: a systematic review. J Appl Gerontol 2021 May;40(5):547-557. [doi: 10.1177/0733464820979801] [Medline: 33356769]
8. Fillekes MP, Giannouli E, Kim EK, Zijlstra W, Weibel R. Towards a comprehensive set of GPS-based indicators reflecting the multidimensional nature of daily mobility for applications in health and aging research. Int J Health Geogr 2019 Jul 24;18(1):17. [doi: 10.1186/s12942-019-0181-0] [Medline: 31340812]
9. Fancello G, Vallée J, Sueur C, et al. Micro urban spaces and mental well-being: measuring the exposure to urban landscapes along daily mobility paths and their effects on momentary depressive symptomatology among older population. Environ Int 2023 Aug;178:108095. [doi: 10.1016/j.envint.2023.108095] [Medline: 37487375]
10. Wang B, Loo BPY, Zhang F. Aging in place: toward understanding momentary well-being and daily satisfaction of older people. Ann Am Assoc Geogr 2024 Oct 20;114(9):2044-2067. [doi: 10.1080/24694452.2024.2370919]
11. Gough C, Lewis LK, Barr C, Maeder A, George S. Community participation of community dwelling older adults: a cross-sectional study. BMC Public Health 2021 Mar 29;21(1):612. [doi: 10.1186/s12889-021-10592-4] [Medline: 33781223]
12. Harada K, Lee S, Lee S, et al. Objectively-measured outdoor time and physical and psychological function among older adults. Geriatr Gerontol Int 2017 Oct;17(10):1455-1462. [doi: 10.1111/ggi.12895] [Medline: 27633728]
13. Luo M, Kim EK, Weibel R, Martin M, Röcke C. GPS-derived daily mobility and daily well-being in community-dwelling older adults. Gerontology 2023;69(7):875-887. [doi: 10.1159/000527827] [Medline: 36657416]
14. Pearson AL, Tribby C, Brown CD, Yang JA, Pfeiffer K, Jankowska MM. Systematic review of best practices for GPS data usage, processing, and linkage in health, exposure science and environmental context research. BMJ Open 2024 Feb 1;14(2):e077036. [doi: 10.1136/bmjopen-2023-077036] [Medline: 38307539]
15. Terhorst Y, Knauer J, Philippi P, Baumeister H. The relation between passively collected GPS mobility metrics and depressive symptoms: systematic review and meta-analysis. J Med Internet Res 2024 Nov 1;26:e51875. [doi: 10.2196/51875] [Medline: 39486026]
16. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL group. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/s0033291798006667] [Medline: 9626712]
17. OECD guidelines on measuring subjective well-being. : OECD Publishing; 2013 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2013/03/oecd-guidelines-on-measuring-subjective-well-being_g1g28641/9789264191655-en.pdf [accessed 2026-10-31]
18. Whittemore R, Knafl K. The integrative review: updated methodology. J Adv Nurs 2005 Dec;52(5):546-553. [doi: 10.1111/j.1365-2648.2005.03621.x] [Medline: 16268861]
19. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
20. Active ageing: a policy framework. : World Health Organization; 2002 URL: https://iris.who.int/server/api/core/bitstreams/0418705f-1c82-4fe2-82d6-d4faee89dfa0/content [accessed 2026-08-07]
21. Hong QN, Fàbregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
22. Harada K, Lee S, Lee S, Bae S, Harada K, Shimada H. Changes in objectively measured outdoor time and physical, psychological, and cognitive function among older adults with cognitive impairments. Arch Gerontol Geriatr 2018;78:190-195. [doi: 10.1016/j.archger.2018.06.003] [Medline: 30006211]
23. Kerr J, Marshall S, Godbole S, et al. The relationship between outdoor activity and health in older adults using GPS. Int J Environ Res Public Health 2012 Dec;9(12):4615-4625. [doi: 10.3390/ijerph9124615] [Medline: 23330225]
24. Takemoto M, Carlson JA, Moran K, Godbole S, Crist K, Kerr J. Relationship between objectively measured transportation behaviors and health characteristics in older adults. Int J Environ Res Public Health 2015 Oct 30;12(11):13923-13937. [doi: 10.3390/ijerph121113923] [Medline: 26528999]
25. Li D, Zhai Y, Xiao Y, Newman G, Wang D. Subtypes of park use and self-reported psychological benefits among older adults: a multilevel latent class analysis approach. Landsc Urban Plan 2019 Oct;190:103605. [doi: 10.1016/j.landurbplan.2019.103605] [Medline: 31341342]

XSL•FO
RenderX

26. Wettstein M, Wahl HW, Diehl MK. A multidimensional view of out-of-home behaviors in cognitively unimpaired older adults: examining differential effects of socio-demographic, cognitive, and health-related predictors. Eur J Ageing 2013 Jun;11(2):141-153. [doi: 10.1007/s10433-013-0292-6] [Medline: 28804322]
27. Dayyani F, Chu CH, Abedi A, Khan SS. Correlations between social isolation and functional decline in older adults after lower limb fractures using multimodal sensors: a pilot study. Algorithms 2024;17(9):383. [doi: 10.3390/a17090383]
28. Giannouli E, Fillekes MP, Mellone S, Weibel R, Bock O, Zijlstra W. Predictors of real-life mobility in community-dwelling older adults: an exploration based on a comprehensive framework for analyzing mobility. Eur Rev Aging Phys Act 2019;16:19. [doi: 10.1186/s11556-019-0225-2] [Medline: 31700551]
29. Hackett K, Xu S, McKniff M, Paglia L, Barnett I, Giovannetti T. Mobility-based smartphone digital phenotypes for unobtrusively capturing everyday cognition, mood, and community life-space in older adults: feasibility, acceptability, and preliminary validity study. JMIR Hum Factors 2024 Nov 22;11:e59974. [doi: 10.2196/59974] [Medline: 39576984]
30. Zhu S, Chirles TJ, Keller JA, et al. Development of an algorithm for analysis of routes: case studies using novice and older drivers. J Safety Res 2024 Sep;90:319-332. [doi: 10.1016/j.jsr.2024.07.003] [Medline: 39251289]
31. Chung J, Boyle J, Wheeler DC. Relationship between life-space mobility and health characteristics in older adults using global positioning system watches. J Appl Gerontol 2022 Apr;41(4):1186-1195. [doi: 10.1177/07334648211054834] [Medline: 34719296]
32. Jankowska MM, Yang JA, Luo N, Spoon C, Benmarhnia T. Accounting for space, time, and behavior using GPS derived dynamic measures of environmental exposure. Health Place 2023 Jan;79:102706. [doi: 10.1016/j.healthplace.2021.102706] [Medline: 34801405]
33. Kilgallon JL, Tewarie IA, Broekman MLD, Rana A, Smith TR. Passive data use for ethical digital public health surveillance in a postpandemic world. J Med Internet Res 2022 Feb 15;24(2):e30524. [doi: 10.2196/30524] [Medline: 35166676]
34. Sun L, Yang B, Kindt E, Chu J. Privacy barriers in health monitoring: scoping review. JMIR Nurs 2024 May 9;7:e53592. [doi: 10.2196/53592] [Medline: 38723253]

Abbreviations

MMAT: Mixed Methods Appraisal Tool
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews

Edited by Y Jiang; submitted 08.Jun.2026; peer-reviewed by F Milella, N Stasinos, Z Liu; revised version received 19.Oct.2026; accepted 22.Oct.2026; published 19.Nov.2026.

Please cite as:
Hyun J, Kim B, Kang H, Choi J
GPS-Based Approaches to Assessing Psychosocial Health in Community-Dwelling Older Adults: Integrative Review
JMIR Aging 2026;9:e103996
URL: https://aging.jmir.org/2026/1/e103996
doi:10.2196/103996

© JaeWon Hyun, Bomgyeol Kim, Hun Kang, JiYeon Choi. Originally published in JMIR Aging (https://aging.jmir.org), 19.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI-Based Social Connection Interventions of Emerging Categories of Support for Older Adults: Scoping Review

Hana Kim[1], MPP; Cheng Ren[1], PhD; Heather Horton[1], PhD; Jeesoo Jung[2], PhD; Yanjun Dong[1,3], PhD; Victoria Rizzo[1], PhD

[1]School of Social Welfare, College of Integrated Health Sciences, University at Albany, State University of New York, 1400 Washington Avenue, Albany, NY, United States

[2]K-Inclusive Society Institute, Kyungpook National University, Daegu, Republic of Korea

[3]Center for Innovaton in Social Science, College of Art and Science, Boston University, Boston, MA, United States

Corresponding Author:
Cheng Ren, PhD
School of Social Welfare, College of Integrated Health Sciences, University at Albany, State University of New York, 1400 Washington Avenue, Albany, NY, United States

Abstract

Background: Social connection is a critical determinant of well-being in later life. However, older adults often experience social isolation and increased loneliness, which are associated with mental and physical health challenges. AI-based interventions have emerged as scalable approaches to support social connection. A growing body of research has examined the effects of individual AI technologies on improving social connection across diverse devices and care settings. Despite growing evidence of their effects, little attention has been paid to the common pathways through which these interventions support social connection.

Objective: This scoping review aimed to develop a typology of common categories of support across AI-based interventions and to map how these categories are configured in different intervention contexts.

Methods: Five databases (PubMed, MEDLINE, CINAHL, PsycINFO, and Web of Science) were searched for studies published from 2016 through 2025. Studies were included if they examined AI-based interventions designed to support social connection among older adults, conceptualized as a multidimensional construct encompassing loneliness reduction, social interaction, and behavioral engagement.

Results: A total of 1110 articles were retrieved, of which 47 met the inclusion criteria. Qualitative and mixed methods studies were the most prevalent (n=32, 68%), with quantitative studies accounting for the remainder (n=15, 32%). Four common support categories were identified: relational interaction (forming ties), emotional support (enhancing interaction quality), functional support (sustaining engagement), and risk detection (reducing constraints). Relational interaction and emotional support were present across all studies. Functional support and risk detection were observed in 70% (n=33) and 17% (n=8) of studies, respectively. The most prevalent configuration combined relational interaction, emotional support, and functional support, forming a companionship-plus-routine pathway that sustained engagement in everyday contexts. Functional support was associated with sustained participation in daily routines and activities, which may have contributed to opportunities for ongoing social interaction. Risk detection, while less prevalent, functioned as a constraint-reducing condition that supported continuity of engagement.

Conclusions: This review proposes a conceptual framework that shifts attention from specific technologies to the interactional configurations that enable AI to support social connection across care settings. This framework may inform the design and tailoring of AI-based interventions to address the multidimensional nature of social connection among older adults.

(JMIR Aging 2026;9:e99463) doi:10.2196/99463

KEYWORDS

AI; social connection; social isolation; AI interventions; loneliness; older adults; scoping review

Introduction

Background

Across many regions worldwide, rapid population aging, the rise of single-person older adult households, and the gradual erosion of traditional family–based support systems have intensified relational deficits among older adults [1]. These structural changes reduce opportunities for meaningful emotional exchanges and social interactions, thereby constraining older adults' access to consistent emotional support or social companionship [2,3].

Social connection is conceptualized as a multidimensional construct encompassing subjective emotional experiences (eg, loneliness and belonging), objective relational patterns (eg,

XSL•FO
RenderX

frequency and regularity of social interactions), and related social and emotional outcomes, such as engagement and companionship [3,4]. While multifaceted, social connection is fundamentally shaped through social interaction. In later life, however, opportunities for such interactions often decline due to structural and life-course changes. As these opportunities diminish, the interactional processes through which meaning and connection are constructed may be constrained, potentially contributing to relational deficits and weakened social connectedness among older adults.

Reduced social connection, in turn, increases vulnerability to adverse mental and physical health outcomes, including depression and anxiety [5], as well as dementia, cardiovascular disease, stroke, functional decline, and all-cause mortality [6,7]. These findings highlight the role of reduced social connection as a central determinant of well-being in later life and point to the need for effective strategies to address reduced opportunities for social interaction in aging populations.

Recent years have witnessed a rapid expansion of AI technologies in care for older adults, promoted as innovative strategies to respond to growing care demands and constraints on public and human resources [8-10]. AI refers to computational systems capable of performing tasks that traditionally require human intelligence, such as reasoning, perception, learning, and decision-making. AI encompasses a broad range of approaches, including rule-based symbolic AI, machine learning, neural networks, pattern recognition, and natural language processing, enabling systems to interpret information, learn from data, and generate predictions, recommendations, or decisions toward human-defined goals [11-14]. This conceptualization is broadly consistent with the European Union AI Act, which defines AI as a machine-based system capable of generating predictions, recommendations, or decisions with varying levels of autonomy [15].

AI has been widely applied in health-related contexts, including disease detection, diagnosis, health monitoring, predictive analytics, and patient management [11,13]. In addition, technologies such as social robots, conversational agents, and voice assistants have expanded into psychosocial domains, increasingly being used to support emotional well-being, social connection, and the reduction of loneliness among older adults [14,16,17], while also serving complementary functions in health monitoring and daily living support [18,19]. As traditional forms of human delivery are increasingly constrained by scalability and accessibility limitations [20], AI-based technologies have been introduced to complement existing care services and expand opportunities for social interactions.

Accordingly, a growing body of research, using diverse methods and AI systems across countries, has examined how AI can address psychological challenges faced by older adults. AI-based interventions for social connection can be conceptualized as intervention approaches that incorporate AI-based technologies to deliver, support, or evaluate activities intended to promote social connection [14,21,22]. These interventions encompass a wide range of technologies and service models designed to reduce loneliness and isolation and foster related emotional support. Emerging evidence includes socially assistive robots, conversational agents, personal voice assistants, digital companions, and other assistive technologies. For example, socially assistive robots such as Pepper (SoftBank Robotics), ElliQ (Intuition Robotics), and Hyodol (Hyodol Inc) have been designed to engage older adults through conversation, emotional interaction, reminders, games, and companionship, with evidence suggesting benefits for emotional well-being and psychological health [17]. Similarly, AI-powered personal voice assistants (eg, Amazon Alexa) have been associated with significant reductions in loneliness and increases in perceived social support among older adults, with greater social support linked to lower loneliness [23].

Existing systematic and scoping reviews have primarily examined intervention effectiveness and psychosocial outcomes, such as loneliness, social connectedness, and psychological well-being, or have organized the literature according to specific technology categories (eg, social robots and conversational AI) or phases of care delivery [14,16,17,21]. However, little attention has been paid to identifying common categories of support across diverse AI technologies and care contexts. Technologies that differ substantially in design and technical features may provide similar forms of support, while technologies within the same category may provide different forms of support. Identifying shared support categories may therefore offer a complementary way to understand the roles that AI technologies play across diverse intervention contexts. This study identifies and synthesizes common categories of support reported across AI-based interventions for older adults.

Conceptual Foundations

This study draws on 3 theoretical perspectives—the socioemotional selectivity theory (SST), the social convoy model (SCM), and the socio-ecological model (SEM)—to contextualize and interpret the support categories identified through the review. These perspectives highlight different dimensions of social connection across individual, relational, and structural contexts. Collectively, they provide a useful framework for understanding how AI-based interventions may support social connection in later life. SST posits that as individuals perceive their remaining time as limited, they increasingly prioritize emotionally meaningful goals and relationships [24,25]. In AI-assisted care, SST offers a useful lens for understanding how AI-based interventions may help older adults engage in emotionally meaningful interactions and relationships.

While SST highlights individual-level emotional processes, the SCM extends this understanding by situating social connection within dynamic relational networks. Within the SCM, AI-based interventions may augment existing convoys by maintaining interaction, facilitating communication, or providing companionship during life transitions such as bereavement, retirement, or declining mobility that may weaken older adults' broader networks of social ties [26]. Thus, the SCM provides a conceptual basis for understanding relational interaction as a key dimension through which AI-based interventions may support older adults' social ties.

Building on these relational dimensions, the SEM further situates social connection within broader environmental and structural

XSL•FO
RenderX

contexts. The SEM suggests that social connection is shaped by environmental, community, and structural conditions that enable or constrain sustained social engagement [27,28]. In the digital era, this ecological system has expanded to include the techno-subsystem, in which technological tools act as vital mediators of everyday engagement [29]. From this perspective, AI-based interventions can be viewed as an infrastructural support system that mitigates contextual barriers—such as functional decline or caregiver strain—thereby empowering older adults to sustain their social connections and agency.

Study Objectives

The specific objectives of this scoping review are to (1) identify and describe recurring categories of support that operate across AI-based interventions for social connection, and (2) map how combinations of these categories are configured within AI-based interventions and discussed alongside social connection–related outcomes in the literature. In this review, support categories are conceptualized as recurring functional and interactional processes through which AI systems engage with older adults and are discussed in relation to social connection outcomes. By explaining these categories and their combinations, this review develops a conceptual framework that may inform the interpretation and design of AI-based social connection interventions that are not tied to specific technologies or platforms.

Methods

Study Design

We used standard procedures for conducting scoping reviews. The review followed the 5 steps outlined by Arksey and O'Malley [30], including identifying the research question; identifying relevant studies; study selection; charting the data; and collating, summarizing, and reporting the results. The reporting of this review adhered to the Preferred Reporting Items for Systematic reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) guidelines [31]. The completed PRISMA-ScR checklist is available in Checklist 1.

Search Strategy

We searched 5 major databases spanning medicine, nursing, psychology, and interdisciplinary research: PubMed, MEDLINE (via ProQuest), CINAHL (via EBSCOhost), PsycINFO (via ProQuest), and Web of Science. These databases were selected to ensure comprehensive coverage of gerontology, psychosocial research, and AI and technology-related studies.

We included empirical studies that examined AI-based interventions designed to support social connection or related psychosocial outcomes among older adults. Consistent with prior research, social connection was operationalized to include both subjective (eg, loneliness and belonging) and objective (eg, social interaction patterns) dimensions, as well as related social and emotional outcomes [3,4]. Based on this operationalization, a broad set of keywords was used to reflect the multidimensional nature of social connection. All database searches were limited to studies published in English-language journals from 2016 through December 2025. This time frame was selected to reflect the rapid development of AI technologies over the past decade, including innovations designed for older adults.

The initial search was conducted in November 2025, and the updated search was conducted in January 2026. The search strategy is provided in Multimedia Appendix 1. The search algorithm included terms related to older adults (eg, "older adults," "elderly," and "seniors"), AI (eg, "AI," "social robot," and "chatbot"), and social connection–related terms (eg, "social isolation," "loneliness," "companionship," and "social connection"). All terms were searched within titles and abstracts to capture studies with an explicit conceptual and empirical focus on AI-based social connection interventions among older adults.

Eligibility Criteria

Studies were included in the review if they met all of the following criteria: (1) the intervention incorporated AI-based approaches as a core component of the intervention, with technologies considered AI-based when they used capabilities such as symbolic AI, machine learning, natural language processing, conversational agents, or adaptive decision-making functions; (2) the intervention facilitated or examined social connection, including social interaction, companionship, emotional comfort, perceived presence, loneliness, or related psychosocial experiences among older adults, regardless of whether these outcomes were the primary focus of the intervention; (3) study participants were older adults or individuals 60 years of age and older, or primarily older adults as defined by the original studies (eg, studies targeting "older adults," "elderly," "seniors," or "geriatric populations"); (4) technologies were eligible regardless of whether they were originally developed specifically for older adults or for broader populations, provided that the intervention was implemented and evaluated among older adults; (5) original, empirical, and published in peer-reviewed journals; (6) studies published within the past 10 years (2016 - 2025); and (7) papers written in English.

Studies were excluded from the review if they met any of these conditions: (1) AI techniques were mentioned only tangentially or were not implemented as an intervention through which social connection or related relational or emotional experiences were examined or observed; (2) reported outcomes did not reflect social connection–relevant experiences or processes (eg, companionship, emotional or relational engagement, social interaction, and loneliness), and instead focused exclusively on physical health, cognitive training, technical performance, or usability outcomes alone; and (3) the publication was nonempirical in nature, including letters, editorials, commentaries, conference abstracts, meeting materials, study or review protocols, or literature reviews.

Study Selection

Studies were selected using a 2-stage screening process. In the first stage, titles and abstracts were screened to assess relevance to the review aims. Duplicate records were identified and removed using Rayyan (Qatar Computing Research Institute). In the second stage, full-text studies were reviewed to determine

XSL•FO
RenderX

final eligibility. HK screened titles and abstracts and reviewed full-text studies using the predefined eligibility criteria.

Charting the Data

Data extraction was conducted to systematically capture relevant information from the included studies. Extracted variables included study characteristics (author, publication year, and country), intervention characteristics (AI systems, reported AI capabilities, care setting, intervention frequency, duration, and total study period), methodological features (study design and sample size), and outcome-related information (primary outcome focus and reported outcomes related to social connection). Data extraction was performed by HK and reviewed by coauthors (CR, HH, JJ, YD, and VR) to ensure accuracy and consistency. A detailed summary of extracted variables is provided in Multimedia Appendix 2.

Data Analysis

This review adopted an inductive approach to identify and organize key categories through systematic steps: (1) descriptive mapping of study characteristics, (2) analysis of AI system functions and reported outcomes, and (3) iterative cross-study comparison. All included studies were reviewed and summarized in a structured analytic matrix (Multimedia Appendix 2), providing a common framework for systematic comparison.

Data extraction and initial category development were conducted by the first author (HK). To enhance the transparency and consistency of category assignment, 2 coauthors(JJ, YD) independently reviewed the full-text studies and corresponding category assignments for all included studies. The operational criteria for assignment to the support categories are summarized in Multimedia Appendix 3. Intercoder agreement was assessed across all 4 support categories, with complete agreement for relational interaction, emotional support, and risk detection support. Functional support was the only category with coding discrepancies; Fleiss κ was 0.93. Discrepancies were resolved through discussion and consensus. Here are 2 borderline cases involving discrepancies in the functional support category [32,33]. In both cases, 2 coders initially assigned “yes,” while one assigned “no,” and the discrepancies were resolved through discussion and consensus. For Noguchi et al [33], the robot helped older adults communicate requests for support to family members or friends, rather than directly providing functional assistance itself. Therefore, this function was not coded as functional support. For Nagata et al [32], only functions actually provided by the intervention were coded; desired future functions were excluded.

First, a descriptive mapping of the included studies was used to characterize the overall structure of the literature. Studies were sorted according to key intervention characteristics (eg, AI systems, care settings, and intervention duration) to capture the breadth and diversity of the evidence base.

Second, a thematic analysis examined the operational features of AI-based interventions as described in the included studies. Full-text studies were analyzed to identify observable functions of AI systems and their links to reported social and emotional outcomes. To identify common categories of support across diverse AI interventions, recurrent intervention characteristics, AI system functions, and reported outcomes were compared across studies. Through iterative comparison, patterns reflecting similar forms of support were grouped into higher-order thematic categories.

Finally, through iterative cross-study comparison, recurrent patterns in AI system functions and interactional characteristics were inductively identified and grouped into higher-order categories. This bottom-up synthesis identified a set of common categories of support that characterize the ways in which AI-based interventions support social connection in later life. By grounding category identification in comparative analysis across the included studies, this approach ensured that the proposed category-based framework reflected recurring patterns observed across diverse AI technologies rather than effects attributable to any single technology. The identified categories were subsequently considered in relation to broader theoretical perspectives on social connection during the interpretation of the findings.

Ethical Considerations

This scoping review used publicly available literature and did not involve human participants or identifiable data. As a result, ethics approval was not required.

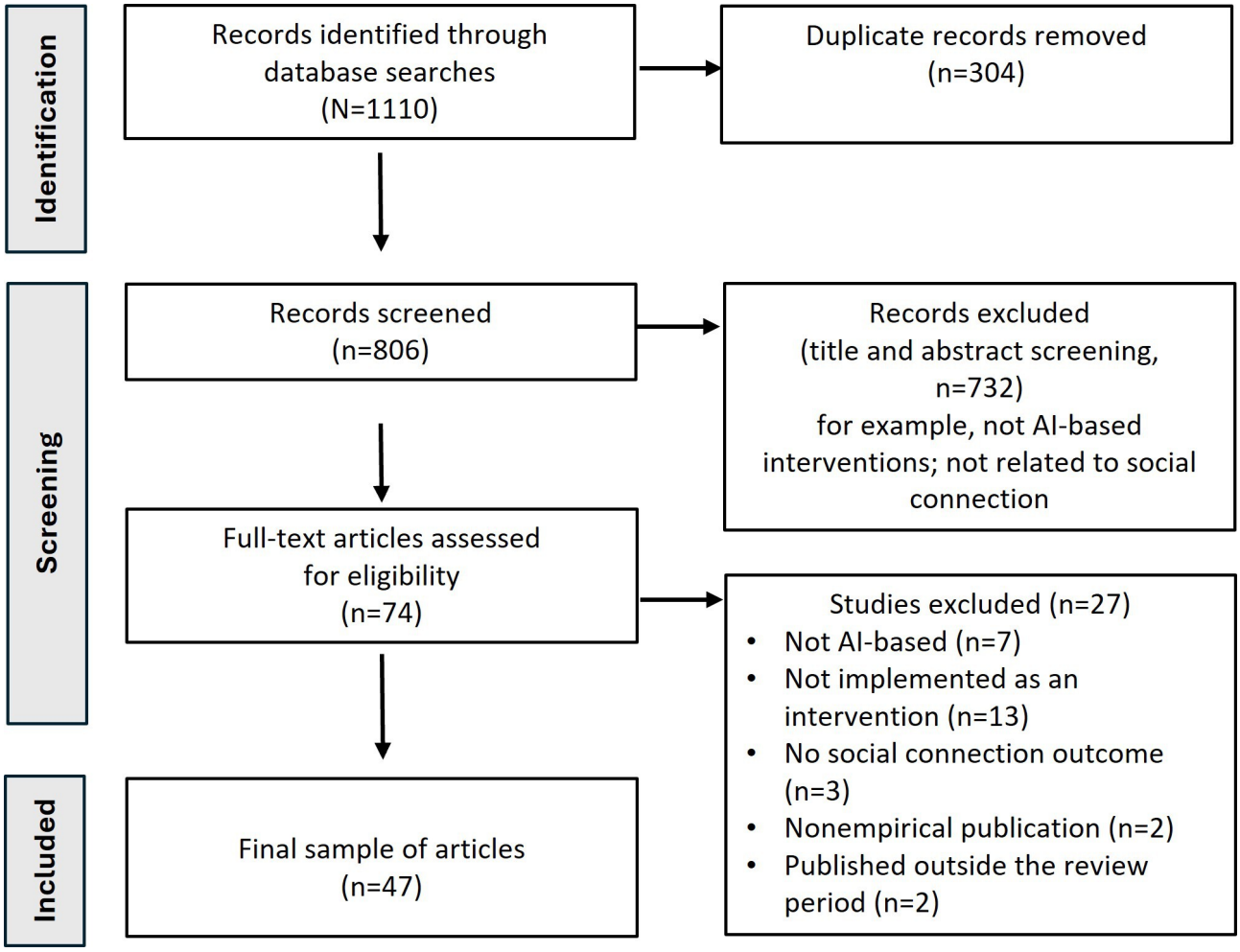
Results

Overview and Characteristics of Included Studies

Figure 1 provides a detailed outline of the search and selection process. A total of 47 studies met the inclusion criteria. Table 1 summarizes the key characteristics of the included studies. The literature spanned multiple publication years and geographic regions, indicating sustained and expanding interest in this area across national contexts. AI systems were classified according to the primary technological approach of each intervention (Table 2), which included socially assistive robots, pet-like social robots, voice-based AI assistants, telepresence robots, conversational agents, and humanoid service robots.

XSL•FO
RenderX

Figure 1. Flow diagram of the study identification and selection process.



XSL•FO
RenderX

Table . Characteristics of the studies included in the review (N=47).

Characteristics	Count, n (%)
Publication year	
2021 - 2025	39 (83)
2016 - 2020	8 (17)
Country/region where study was conducted	
United States	16 (34)
South Korea	6 (13)
Japan	5 (11)
Taiwan	5 (11)
Canada	3 (6)
Mainland China	2 (4)
Hong Kong	2 (4)
Netherlands	2 (4)
United Kingdom	2 (4)
Other countries/regions (ie, Australia, Belgium, Ireland, and Singapore)	4 (9)
Study design	
Quantitative study	15 (32)
Qualitative study	16 (34)
Mixed methods study	16 (34)
Care setting	
Community-dwelling	30 (64)
Residential care facilities (eg, care homes and nursing homes)	13 (28)
Mixed/hospital	4 (9)
Intervention duration (participant exposure), (weeks)	
<4	15 (32)
4-8	18 (38)
>8	14 (30)

XSL•FO
RenderX

Table . Types of AI systems in the included studies (N=47)[a].

AI system	Description	Count, n (%)
Socially assistive (companion) robots	Autonomous or semiautonomous robots that engage users through language-based interaction and provide social and emotional support, with varying levels of assistance for daily activities (eg, Hyodol, ElliQ, MARIO, PIO, and Maatje-pop)	23 (49)
Pet-like (therapeutic) social robots	Physically embodied, tactile robots that engage users mainly through nonverbal sensory interaction (eg, touch and movement) and are designed primarily to offer emotional comfort, reassurance, and companionship (eg, PARO, robotic pets, and LOVOT)	11 (23)
Voice-based AI assistants (smart speakers)	Voice-based AI systems without physical embodiment (eg, Alexa and Google Assistant)	7 (15)
Conversational agents (chatbots)	Text-, screen-, or voice-based conversational AI systems accessed via messaging platforms, web interfaces, or mobile apps, including both nonembodied systems and digitally embodied agents with visual or animated representations (eg, ChatGPT and PACO)	5 (11)
Humanoid service robots (nonsocially assistive robots)	Humanoid robots designed for general social interaction or service provision but not primarily framed as socially assistive or companion robots (eg, Pepper)	1 (2)

[a]AI systems were classified descriptively based on the terminology and descriptions reported in the included studies. AI systems were treated as mutually exclusive, with each study was assigned to the single category that best reflected the primary form of the AI intervention.

Among the 47 studies, 20 (43%) referenced the COVID-19 pandemic in relation to the intervention context, implementation, or interpretation of findings. Within this subset, 3 studies were explicitly framed as pandemic-responsive interventions to address heightened isolation under social distancing or visitation restrictions [18,34,35]. The remaining 17 studies referenced COVID-19 primarily as a contextual condition shaping implementation and interpretation (eg, recruitment constraints, modified data collection, or restricted social activity), while still demonstrating the continued use and contextual adaptability of AI technologies [23,36,37].

Regarding research methodology, qualitative (n=16) and mixed methods studies (n=16) were the most prevalent (n=32, 68%), while quantitative studies accounted for 32% (n=15) in the included studies. This distribution is consistent with an emphasis on capturing in-depth user experiences and contextual insights, while also incorporating quantitative evidence. Study designs also included longitudinal designs (n=2) and multigroup comparative designs (n=10). The sample sizes across the included studies varied substantially, ranging from small-scale studies (n=4) to large-scale surveys (n=589), with 5 studies involving over 100 participants. Participants included older adults, family caregivers, and care professionals.

Most interventions were implemented in real-world care settings, emphasizing routine care practices rather than experimental control, and were delivered to community-dwelling older adults, care home residents, and older adults in mixed-care or hospital settings. The duration of interventions varied widely, ranging from a single session to 45 months. Intervention duration was defined as the length of time participants were directly exposed to the intervention, rather than the total program duration. The majority of studies focused on midrange durations: 38% (n=18) described interventions lasting 4 to 8 weeks, 32% (n=15) described short-term interventions (<4 weeks), and the final 30% (n=14) examined longer-term effects lasting more than 8 weeks.

Social Connection–Related Outcomes

Across the 47 articles, outcomes related to social connection were grouped into four domains: (1) subjective social connection outcomes (subjective experiences), (2) interaction and engagement outcomes (observable behaviors), (3) emotional and quality-of-life outcomes (psychological consequences), and (4) broader relational outcomes (relational and contextual effects). These domains capture and summarize outcome measures reported in the included studies.

Subjective Social Connection Outcomes

The most frequently reported outcomes were measures reflecting social connection, including loneliness, perceived isolation, perceived connectedness, companionship, sense of presence, and friendship. Loneliness was the most commonly assessed outcome across studies, typically measured using validated scales such as the UCLA Loneliness Scale. Twelve studies reported reductions in loneliness and perceived isolation, whereas 8 studies reported nonsignificant or mixed findings [32,38,39]. Moreover, relational aspects of connection, including companionship, perceived presence, and feelings of being socially connected, were reported in many studies, particularly in those using socially assistive robots, pet-like social robots, and conversational agents [40-42]. In addition, several studies

XSL•FO
RenderX

reported related relational resources, such as perceived social support and stimulation of social networks [23,33], indicating that both subjective experiences of connection and broader relational perceptions were reported across studies.

Interaction and Engagement Outcomes

A second domain captured observable behavioral indicators of social connection, including interaction frequency, communication patterns, conversational exchanges, and participation in social or recreational activities. These outcomes involved both direct human-AI interaction and communication with family members, caregivers, or peers [43-45]. Several studies reported increases in interaction time, number of exchanges, and opportunities for conversation, as well as improvements in overall social engagement [44-46]. In some cases, AI-mediated interaction was also reported to stimulate human-human interaction, indicating that engagement outcomes extended beyond the immediate interaction with the technology.

Emotional and Quality-of-Life Outcomes

Emotional and quality-of-life outcomes included measures of emotional comfort, mood, depression, anxiety, and overall quality of life. These outcomes were reported alongside social connection outcomes in several studies. Many studies reported improvements in mood, reductions in depressive symptoms, and increased emotional comfort associated with AI-based interactions; however, findings varied across studies. While some studies showed statistically significant improvements in depression and quality of life [20,47], others reported limited or inconsistent effects across study settings and intervention types.

Broader Relational and Contextual Outcomes

Broader relational outcomes were reported in a small subset of studies. These included reductions in caregiver burden and increased reassurance among family members [48,49], as well as enhanced perceptions of safety and reduced vulnerability among older adults. In several studies, safety-related functions—such as monitoring, anomaly detection, and emergency alert systems—were described as contributing to a greater sense of security and reduced uncertainty in daily life [19,50]. In addition, some studies examined changes in digital self-efficacy, autonomy, and technology-mediated relational capacity [37,49]. These outcomes were reported at relational levels beyond the individual.

Together, these 4 domains illustrate the multidimensional nature of social connection, encompassing subjective, behavioral, psychological, and relational levels across AI-based interventions.

Support Categories for AI-Based Social Connection Interventions

Overview of Support Categories

This section delineates the common categories through which AI-based social connection interventions were implemented and evaluated. Inductive analysis identified 4 categories of support underlying AI-based interventions for social connection among older adults: relational interaction, emotional support, functional support, and risk detection. Each category reflects a distinct pathway through which AI technologies support social connection in later life.

Relational interaction and emotional support were present across all 47 included studies. These categories frequently co-occurred with functional components and, in some cases, risk-related components (Tables 3 and 4).

Table . Support categories in AI-based social connection interventions (N=47).

Support category	Operational examples	Count, n (%)
Relational interaction	Conversational exchange, reciprocal dialogue, companionship, and perceived social presence	47 (100)
Emotional support	Emotional comfort, reassurance, affective responsiveness, and mood stabilization	47 (100)
Functional support	Reminders, routine structuring, medication adherence, information access assistance, and structured activity–based programs (eg, cognitive, physical, or recreational activities)	33 (70)
Risk detection	Safety monitoring, emergency detection, fall, or health-risk alerts	8 (17)

Table . Category combinations and corresponding studies (N=47).

Category combination	Studies	Count, n (%)
Relational support, emotional support, and functional support	[18,23,35-39,45,49,51-66]	25 (53)
Relational support and emotional support	[32-34,40,42,46,47,67-73]	14 (30)
Relational support, emotional support, functional support, and risk detection	[19,20,41,50,74-77]	8 (17)

XSL•FO
RenderX

Identified Support Categories

The support categories were derived through an iterative analysis of observable operational behaviors of AI technologies and their reported links to social and emotional outcomes across the 47 studies. These categories represent distinct patterns through which AI-based interventions were operationalized and described. The 3 theoretical perspectives introduced in the conceptual foundations (SST, SCM, and SEM) were applied as interpretive lenses to contextualize these categories within broader understandings of social connection. The first category, relational interaction, emerged across AI-based systems that engaged in and facilitated ongoing social interactions. These systems fostered a sense of interpersonal presence and closeness by providing humanlike social interactions—such as conversation and companionship—or by enabling communication between older adults and others (eg, family members). Relational interaction was consistently observed across studies, regardless of the type of AI technology. It was associated with multiple social connection outcomes. For example, more frequent relational interactions with AI—such as greetings or initiating conversation—were associated with greater reductions in loneliness [63]. Similarly, AI-based communication features were linked to reduced self-disclosure anxiety and reengagement with family and friends [33].

Second, the emotional support category encompassed AI functions that responded to users' internal emotional states, including mood-aware dialogue, humor, reassurance, and other emotionally responsive interactions across the reviewed studies. The emotional support category was reflected in patterns such as emotional comfort, reduced loneliness and anxiety, improved mood, and enhanced psychological well-being [46,60,76]. These functions were described as alleviating negative emotional states, including loneliness, anxiety, and depression. For example, one participant described how the companion robot said "I love you" and "You are my best friend" during moments of distress, thereby contributing to reduced loneliness and emotional relief [41]. Emotional support was consistently observed across studies, regardless of the type of AI technology.

Third, the functional support category involved AI-enabled assistance with everyday activities that reduced practical barriers to social participation by supporting the planning, continuity, or execution of daily tasks. Within the reviewed studies, these functions included features such as medication or appointment reminders, prompts for daily routines, assistance with communication or information access, and structured activity-based programs (eg, cognitive, physical, or recreational programs) that supported daily functioning [35,37,52,62]. Functional support was coded based on implemented intervention functions rather than psychosocial outcomes alone. Therefore, companionship, meaningful engagement, reminiscence, memory recall, or social interaction that emerged through conversational exchange alone was not coded as functional support unless they were linked to a specific AI-delivered support function. Reported outcomes associated with functional support included improved daily routines, enhanced engagement in activities, increased autonomy, and sustained participation in everyday life.

Finally, risk detection and safety monitoring were reflected in AI-enabled monitoring and safety-related functions designed to identify potential health or environmental risks, such as falls, emergencies, or abnormal patterns in daily activity or conversation. Reported functions included monitoring systems integrated with socially assistive robots and nurse call systems. For example, Obayashi et al [76] showed that integrating socially assistive robots with bedside monitoring and nurse call systems enhanced older adults' sense of reassurance and continuity of care, particularly at night. Together, these 4 categories provided a structured way of organizing the diverse functions of AI-based interventions identified across the included studies. The following analysis examines how these categories co-occurred across studies and were configured in relation to social connection outcomes.

Configurations of Support Categories

We mapped how the categories were configured across the 47 articles. Across studies, AI-based interventions did not rely on a single category in isolation but instead reflected combinations of multiple categories. Accordingly, these combinations constituted the primary unit of analysis for understanding how AI-based interventions supported social connection. Three recurring configurations were identified.

First, the most prevalent configuration combined relational interaction, emotional support, and functional support, accounting for more than half of the 47 (n=25, 53%) studies. This configuration commonly supported social connection through a companionship-plus-routine pathway, in which conversational engagement fostered perceived companionship and emotional reassurance, while routine-oriented features—such as reminders, structured activities, and daily prompts—supported sustained engagement in everyday life [20,36]. In some studies, this pattern was further reinforced by personified or anthropomorphic interaction, through which older adults related to AI systems as social or relational partners [41,63].

Across studies, this configuration was frequently described in relation to reductions in loneliness and improvements in socioemotional well-being, particularly in contexts where relational and emotional exchanges were embedded in everyday routines [41,49,59].

Second, 14 (30%) studies focused primarily on relational and emotional categories but lacked functional support for organizing daily routines. Interventions in this group typically relied on physically embodied social robots, particularly pet-like therapeutic robots such as PARO, which provided tactile interaction, through which older adults were reported to relate to the robots as social or companion-like entities, commonly described as fostering emotional bonds and a sense of being with another presence [46,71,73]. These interventions were described as fostering a sense of presence, providing emotional comfort, and strengthening affective bonds.

However, there was limited evidence that these emotional gains were associated with longer-term behavior changes, such as increased social participation or changes in everyday routines. This pattern was often described alongside a lack of functional

XSL•FO
RenderX

features—such as reminders, scheduling, or activity coordination.

Third, a smaller subset of studies (n=8, 17%) incorporated all 4 categories, including risk detection alongside relational, emotional, and functional components. These interventions combined companionship and routine support with features such as anomaly detection and monitoring [50], emergency response and alerting [74], and caregiver-facing information sharing [75]. In this group, loneliness outcomes were reported together with a broader range of co-occurring outcomes, including emotional reassurance, support for daily living, and safety or peace of mind [50,74,76]. These studies often presented loneliness as one of multiple reported outcomes rather than as an analytically distinct end point or as a causally linked risk-detection component [19,75].

Overall, different configurations of support categories were associated with distinct patterns in reported social connection outcomes across the included studies.

Discussion

Principal Findings

This study proposes a category-based synthesis of AI-based social connection interventions, shifting attention from technology-centered classification to underlying support functions. We argue that social connection may be understood in relation to relational, emotional, functional, and risk-detection supports across different AI-based interventions. By mapping these support categories, our framework highlights how these supports work in coordination and may contribute to social connection in later life. Ultimately, this approach may inform the development of AI interventions that are responsive to the diverse social needs of older adults.

The 4 support categories identified in this review can also be interpreted through the theoretical perspectives introduced in the conceptual foundations. Relational interaction can be understood as an entry point through which older adults initiate and maintain social contact, whereas the emotional support category captures both emotion-sensitive responses and interactional features that provide emotional reassurance and comfort. Emotional support may function as a stabilizing mechanism that helps older adults manage emotional vulnerability and maintain psychological well-being. Functional support provides the practical stability necessary to sustain social participation in later life. Risk detection functions reduce uncertainty and perceived vulnerability, thereby lowering the psychological threshold for social participation.

SST helps explain why emotionally meaningful interactions and supportive relationships were consistently associated with perceived benefits, as older adults prioritize emotionally meaningful goals and relationships in later life [24,25]. The SCM situates AI-based interventions as supplementary elements within existing relational networks, supporting continuity of engagement rather than replacing human ties [26]. From a social-ecological perspective, functional support and risk detection operate as contextual enablers that reduce practical and safety-related constraints, thereby making sustained participation feasible [27,28].

The findings suggest that social connection outcomes varied with the configuration of relational, emotional, functional, and risk-detection support. Relational and emotional support categories were consistently associated with perceived companionship and affective comfort, although their presence across all included studies can be understood in relation to the review's eligibility criteria, which focused broadly on social connection, including social interaction, companionship, emotional comfort, perceived presence, loneliness, and related psychosocial experiences. Functional support was more commonly associated with sustained engagement in everyday life, whereas risk-detection support was primarily associated with a broader range of reported outcomes, including perceived safety and reassurance.

To illustrate how these common categories operate across contexts, empirical studies provide further insight. Evidence from multiple studies using the same technology underscores the value of a category-based perspective. A series of 5 studies by Lee et al [41,55-57,75] used the same companion robot, Hyodol, across different populations and implementation contexts but reported varying outcomes related to loneliness, depressive symptoms, and emotional well-being. These variations suggest that outcomes may be understood in relation to the configuration of support categories within specific contexts, as well as the technology used. Across these studies, Hyodol consistently demonstrated support in the relational, emotional, and functional categories. Complementing this pattern, Abdollahi et al [60] provided experimental evidence that outcomes can differ within the same robotic platform when specific categories are configured differently. Compared with a nonempathic condition, an empathic interaction mode elicited greater user engagement and improved mood indicators.

Beyond cross-context variation, evidence from recent studies highlights the fragility of AI-based social connection, particularly when underlying support processes fail to adapt over time. Findings from Lei et al [58] and Lee et al [75] show that when relational and emotional mechanisms are weak, misaligned, or insufficiently adaptive, interventions may fail to reduce loneliness or may even increase disengagement. In Lei et al [58], this misalignment is particularly evident when systems do not match users' communicative and cultural expectations (eg, local language recognition and culturally familiar forms of interaction). Relatedly, Lee et al [75] documented a novelty effect, in which short-term reductions in depression and improvements in health-related quality of life were observed at 3 months. However, these benefits were not sustained at 6 months, as the underlying categories plateaued and failed to adapt to users' changing emotional needs. These findings indicate that sustained psychosocial benefits depend on adaptive, evolving motivations rather than on static intervention features.

These insights suggest that AI-based interventions may be particularly valuable when integrated with existing human care and social relationships rather than used in isolation. Because AI alone has limited capacity to adapt over time, its effectiveness depends on how well it supports and complements ongoing

XSL•FO
RenderX

human interactions. Building on these observations, AI-based social connection interventions may benefit from tailoring the configuration of support categories to the specific needs and risk profiles of target populations. For example, interventions emphasizing relational and emotional supports may be appropriate for community-dwelling older adults with relatively low levels of social vulnerability, whereas interventions that integrate functional support and risk detection may be better suited for older adults with greater care needs or safety concerns. This category-based perspective may help guide the development of more tailored and context-sensitive AI interventions.

Future Directions

Future intervention research should focus on using longitudinal and mixed methods approaches that can capture the evolution of AI-based social connections over time. First, longitudinal observations are essential to determine whether AI-based interventions can meaningfully produce sustained improvements in multiple dimensions of social connection and function as an everyday form of relational infrastructure. While this review finds that AI interventions are associated with indicators of social connection, existing evidence also suggests a potential novelty effect, whereby perceived benefits may be confined to the initial phase of use and attenuate over time, resulting in limited sustained impact [75].

Second, mixed methods approaches are critical for examining how common support categories interact with cultural context, service systems, and individual life trajectories. Methodological approaches that hold technological artifacts constant—such as those demonstrated in the Hyodol studies—offer a particularly promising strategy for disentangling the effects of support categories from contextual influences and advancing theory-driven comparisons.

Third, in this type of review, it would be valuable to examine whether outcome patterns cluster within specific category configurations. This would help us observe the discriminative value of the proposed typology and further refine its application across different intervention contexts.

Limitations

Several limitations of this scoping review should be recognized. First, this review was limited to peer-reviewed journal articles published in English. As a result, relevant studies published in other languages, regions, or formats—including gray literature like policy reports, technical documents, and conference proceedings—may have been excluded. The broader AI research landscape, particularly studies published in computer science venues (eg, ACM Digital Library), may not have been fully captured. Moreover, given the rapid evolution of AI technologies, reliance on database searches alone may have resulted in the omission of some emerging interventions.

Second, the review is constrained by the characteristics of the existing evidence base, including the research questions, study designs, and analytic approaches used in prior studies. In addition, many included studies were short-term, individually delivered, and conducted within specific cultural or service contexts, limiting the conclusions about long-term sustainability and the broader applicability of AI-based social connection interventions.

Third, although full-text studies and category assignments were independently reviewed by coauthors, the initial stages of study selection and data extraction were conducted primarily by the first author (HK). Future reviews may benefit from implementing independent, multireviewer procedures throughout the study selection and coding process from the outset, to further enhance methodological rigor.

Conclusions

This scoping review analyzes 47 empirical studies and proposes an initial category-based conceptual framework for AI-driven social connection interventions. By identifying relational, emotional, functional, and risk-detection support categories as distinct yet interdependent pathways, this study shifts the analytical focus from technology-centered approaches to category-attentive, contextually grounded intervention design. In addition, the findings show that social connection outcomes span multiple dimensions, including subjective, behavioral, emotional, and relational domains, underscoring their multidimensional nature.

Across the reviewed studies, AI-based interventions rarely relied on a single support category in isolation; rather, they integrated multiple categories, underscoring the layered and interdependent nature of AI-based social connection. Future research may use this framework to examine how these categories evolve over time and across diverse care contexts, particularly through longitudinal and mixed methods approaches. This framework can provide a basis for designing AI interventions that are grounded in support categories, are context-sensitive, and are responsive to the social needs of older adults.

Generative AI Use

HK used ChatGPT (OpenAI) for English translation and English-language editing of text originally written by the author. All AI-assisted revisions were critically reviewed and edited by HK. All authors reviewed and approved the final manuscript.

Funding

JJ was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2026S1A5C4021468). No specific funding was received for this work.

Data Availability

Data sharing is not applicable to this article because no datasets were generated or analyzed during this study.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: HK
Data curation: HK
Formal analysis: HK (lead), CR (supporting)
Investigation: HK
Methodology: HK, CR, HH
Supervision: CR
Validation: HK, CR, JJ, YD
Writing – original draft: HK
Writing – review & editing: HK, CR, HH, JJ, YD, VR

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy across databases.
[DOCX File, 17 KB - aging_v9i1e99463_app1.docx]

Multimedia Appendix 2
Characteristics of AI intervention in the studies.
[XLSX File, 22 KB - aging_v9i1e99463_app2.xlsx]

Multimedia Appendix 3
Operational criteria for support-category coding.
[DOCX File, 16 KB - aging_v9i1e99463_app3.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 87 KB - aging_v9i1e99463_app4.docx]

References

1. United Nations Department of Economic and Social Affairs, Population Division. World population prospects 2022: summary of results. : United Nations; 2022 UN DESA/POP/2022/TR/NO. 3 URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf [accessed 2026-02-02]
2. Cacioppo JT, Cacioppo S. Social relationships and health: the toxic effects of perceived social isolation. Soc Personal Psychol Compass 2014 Feb 1;8(2):58-72. [doi: 10.1111/spc3.12087] [Medline: 24839458]
3. Holt-Lunstad J. Social connection as a public health issue: the evidence and a systemic framework for prioritizing the "social" in social determinants of health. Annu Rev Public Health 2022 Apr 5;43(1):193-213. [doi: 10.1146/annurev-publhealth-052020-110732] [Medline: 35021021]
4. Berkman LF, Glass T. Social integration, social networks, social support, and health. In: Berkman LF, Kawachi I, editors. Social Epidemiology: Oxford University Press; 2000:137-173. [doi: 10.1093/oso/9780195083316.003.0007]
5. Santini ZI, Jose PE, York Cornwell E, et al. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. Lancet Public Health 2020 Jan;5(1):e62-e70. [doi: 10.1016/S2468-2667(19)30230-0] [Medline: 31910981]
6. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
7. Finlay JM, Kobayashi LC. Social isolation and loneliness in later life: a parallel convergent mixed-methods case study of older adults and their residential contexts in the Minneapolis metropolitan area, USA. Soc Sci Med 2018 Jul;208:25-33. [doi: 10.1016/j.socscimed.2018.05.010] [Medline: 29758475]
8. Governing with artificial intelligence: the state of play and way forward in core government functions. : OECD Publishing; 2025 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/governing-with-artificial-intelligence_398fa287/795de142-en.pdf [accessed 2026-10-17]
9. National every-citizen AI everyday living implementation plan [Article in Korean]. Ministry of Science and ICT, Republic of Korea. 2023. URL: https://www.msit.go.kr/bbs/view.do?sCode=user&mId=307&mPid=208&bbsSeqNo=94&nttSeqNo=3183480 [accessed 2026-10-27]

XSL•FO
RenderX

10. Advancing AI-based welfare and care innovation accessible to citizens [Article in Korean]. Ministry of Health and Welfare, Republic of Korea. 2025. URL: https://www.mohw.go.kr/board.es?mid=a10503010100&bid=0027&act=view&list_no=1487393 [accessed 2026-10-17]
11. Esmaeilzadeh P, Mirzaei T, Dharanikota S. Patients' perceptions toward human-artificial intelligence interaction in health care: experimental study. J Med Internet Res 2021 Nov 25;23(11):e25856. [doi: 10.2196/25856] [Medline: 34842535]
12. Kaplan A, Haenlein M. Siri, Siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Bus Horiz 2019 Jan;62(1):15-25. [doi: 10.1016/j.bushor.2018.08.004]
13. Lukkien DRM, Nap HH, Buimer HP, et al. Toward responsible artificial intelligence in long-term care: a scoping review on practical approaches. Gerontologist 2023 Jan 24;63(1):155-168. [doi: 10.1093/geront/gnab180] [Medline: 34871399]
14. Yang Y, Wang C, Xiang X, An R. AI applications to reduce loneliness among older adults: a systematic review of effectiveness and technologies. Healthcare (Basel) 2025 Feb 20;13(5):446. [doi: 10.3390/healthcare13050446] [Medline: 40077009]
15. Regulation (EU) 2024/1689 of the European parliament and of the council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending regulations (EC) no 300/2008, (EU) no 167/2013, (EU) no 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (artificial intelligence act) (text with EEA relevance). : European Union; 2024 URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj [accessed 2024-06-13]
16. Oh KM, Hong SR, Beran K, Sanders L, Park JY, Lee JA. Utilizing conversational AI technology for social connectedness among older adults: a systematic review. J Appl Gerontol 2026 Mar;45(3):527-549. [doi: 10.1177/07334648251341629] [Medline: 40432591]
17. Park Y, Chang SJ, Kim HJ, Jeong HN. Effectiveness of artificial intelligence robot interventions on psychological health in community-dwelling older adults: a systematic review. J Korean Gerontol Nurs 2024 Aug;26(3):234-247. [doi: 10.17079/jkgn.2024.00353] [Medline: 41584876]
18. Chou YH, Lin C, Lee SH, Lee YF, Cheng LC. User-friendly chatbot to mitigate the psychological stress of older adults during the COVID-19 pandemic: development and usability study. JMIR Form Res 2024 Mar 13;8:e49462. [doi: 10.2196/49462] [Medline: 38477965]
19. Hofstede BM, Askari SI, van Hoesel TRC, et al. Huggable integrated socially assistive robots: exploring the potential and challenges for sustainable use in long-term care contexts. Front Robot AI 2025;12:1646353. [doi: 10.3389/frobt.2025.1646353] [Medline: 41170301]
20. Broadbent E, Loveys K, Ilan G, et al. ElliQ, an AI-driven social robot to alleviate loneliness: progress and lessons learned. JAR Life 2024;13:22-28. [doi: 10.14283/jarlife.2024.2] [Medline: 38449726]
21. Ni Y, Jia F. A scoping review of AI-driven digital interventions in mental health care: mapping applications across screening, support, monitoring, prevention, and clinical education. Healthcare (Basel) 2025 May 21;13(10):1205. [doi: 10.3390/healthcare13101205] [Medline: 40428041]
22. Murray E, Hekler EB, Andersson G, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851. [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
23. Jones VK, Yan C, Shade MY, et al. Reducing loneliness and improving social support among older adults through different modalities of personal voice assistants. Geriatrics (Basel) 2024 Feb 22;9(2):22. [doi: 10.3390/geriatrics9020022] [Medline: 38525739]
24. Carstensen LL. Social and emotional patterns in adulthood: support for socioemotional selectivity theory. Psychol Aging 1992;7(3):331-338. [doi: 10.1037/0882-7974.7.3.331] [Medline: 1388852]
25. Carstensen LL. The influence of a sense of time on human development. Science 2006 Jun 30;312(5782):1913-1915. [doi: 10.1126/science.1127488] [Medline: 16809530]
26. Antonucci TC, Ajrouch KJ, Birditt KS. The convoy model: explaining social relations from a multidisciplinary perspective. Gerontologist 2014 Feb;54(1):82-92. [doi: 10.1093/geront/gnt118] [Medline: 24142914]
27. Bronfenbrenner U. The Ecology of Human Development: Experiments by Nature and Design: Harvard university press; 1979. [doi: 10.2307/j.ctv26071r6]
28. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15(4):351-377. [doi: 10.1177/109019818801500401] [Medline: 3068205]
29. Johnson GM, Puplampu KP. Internet use during childhood and the ecological techno-subsystem. Can J Learn Technol 2008;34(1):1. [doi: 10.21432/T2CP4T]
30. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
31. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
32. Nagata Y, Satake Y, Yamazaki R, et al. A conversational robot for cognitively impaired older people who live alone: an exploratory feasibility study. Psychogeriatrics 2025 Sep;25(5):e70076. [doi: 10.1111/psyg.70076] [Medline: 40716073]
33. Noguchi Y, Kamide H, Tanaka F. How should a social mediator robot convey messages about the self-disclosures of elderly people to recipients? Int J Soc Robotics 2023 Jul;15(7):1079-1099. [doi: 10.1007/s12369-023-01016-x]

XSL•FO
RenderX

34. Fogelson DM, Rutledge C, Zimbro KS. The impact of robotic companion pets on depression and loneliness for older adults with dementia during the COVID-19 pandemic. J Holist Nurs 2022 Dec;40(4):397-409. [doi: 10.1177/08980101211064605] [Medline: 34889146]
35. Van Assche M, Moreels T, Petrovic M, Cambier D, Calders P, Van de Velde D. The role of a socially assistive robot in enabling older adults with mild cognitive impairment to cope with the measures of the COVID-19 lockdown: a qualitative study. Scand J Occup Ther 2023 Jan;30(1):42-52. [doi: 10.1080/11038128.2021.2009560] [Medline: 34871144]
36. Lim J. Effects of a cognitive-based intervention program using social robot PIO on cognitive function, depression, loneliness, and quality of life of older adults living alone. Front Public Health 2023;11:1097485. [doi: 10.3389/fpubh.2023.1097485] [Medline: 36815168]
37. Grigorovich A, Marcotte AA, Colobong R, et al. Using voice-activated technologies to enhance well-being of older adults in long-term care homes. Innov Aging 2024;8(12):igae102. [doi: 10.1093/geroni/igae102] [Medline: 39664605]
38. Kramer LL, van Velsen L, Clark JL, Mulder BC, de Vet E. Use and effect of embodied conversational agents for improving eating behavior and decreasing loneliness among community-dwelling older adults: randomized controlled trial. JMIR Form Res 2022 Apr 11;6(4):e33974. [doi: 10.2196/33974] [Medline: 35404255]
39. Bradwell H, Edwards KJ, Winnington R, Thill S, Allgar V, Jones RB. Implementing affordable socially assistive pet robots in care homes before and during the COVID-19 pandemic: stratified cluster randomized controlled trial and mixed methods study. JMIR Aging 2022 Aug 24;5(3):e38864. [doi: 10.2196/38864] [Medline: 35830959]
40. Chen SC, Davis BH, Kuo CY, Maclagan M, Chien CO, Lin MF. Can the Paro be my Buddy? Meaningful experiences from the perspectives of older adults. Geriatr Nurs 2022;43:130-137. [doi: 10.1016/j.gerinurse.2021.11.011] [Medline: 34883391]
41. Lee OE, Lee H, Park A, Choi NG. My precious friend: human-robot interactions in home care for socially isolated older adults. Clin Gerontol 2024;47(1):161-170. [doi: 10.1080/07317115.2022.2156829] [Medline: 36502295]
42. Tan CK, Lou VWQ, Cheng CYM, He PC, Khoo VEJ. Improving the social well-being of single older adults using the LOVOT social robot: qualitative phenomenological study. JMIR Hum Factors 2024 Aug 23;11:e56669. [doi: 10.2196/56669] [Medline: 39178408]
43. Austbø LBH, Testad I, Gjestsen MT. Using a robot to address the well-being, social isolation, and loneliness of care home residents via video calls: qualitative feasibility study. JMIR Form Res 2025 May 8;9:e59764. [doi: 10.2196/59764] [Medline: 40341128]
44. Follmann A, Schollemann F, Arnolds A, et al. Reducing loneliness in stationary geriatric care with robots and virtual encounters—a contribution to the COVID-19 pandemic. Int J Environ Res Public Health 2021 May 1;18(9):4846. [doi: 10.3390/ijerph18094846] [Medline: 34062776]
45. Lin YC, Fan J, Tate JA, Sarkar N, Mion LC. Use of robots to encourage social engagement between older adults. Geriatr Nurs 2022;43:97-103. [doi: 10.1016/j.gerinurse.2021.11.008] [Medline: 34847509]
46. Chen SC, Lin MF, Jones C, et al. Effect of a group-based personal assistive robot (PARO) robot intervention on cognitive function, autonomic nervous system function, and mental well-being in older adults with mild dementia: a randomized controlled trial. J Am Med Dir Assoc 2024 Nov;25(11):105228. [doi: 10.1016/j.jamda.2024.105228] [Medline: 39182512]
47. Chen SC, Jones C, Moyle W. The impact of engagement with the PARO therapeutic robot on the psychological benefits of older adults with dementia. Clin Gerontol 2024;47(5):909-921. [doi: 10.1080/07317115.2022.2117674] [Medline: 36062840]
48. Hung L, Wong JOY, Ren HL, et al. The impact of telepresence robots on family caregivers and residents in long-term care. Int J Environ Res Public Health 2025 May 1;22(5):713. [doi: 10.3390/ijerph22050713] [Medline: 40427829]
49. Zhao IY, Leung AYM, Huang Y, Liu Y. A social robot in home care: acceptability and utility among community-dwelling older adults. Innov Aging 2025;9(5):igaf019. [doi: 10.1093/geroni/igaf019] [Medline: 40386025]
50. Bott N, Wexler S, Drury L, et al. A protocol-driven, bedside digital conversational agent to support nurse teams and mitigate risks of hospitalization in older adults: case control pre-post study. J Med Internet Res 2019 Oct 17;21(10):e13440. [doi: 10.2196/13440] [Medline: 31625949]
51. Lee J, Joo C. Can social robots reduce the loneliness of older adults? -Interaction and communication between social robots and older adults. Geriatr Nurs 2025;64:103359. [doi: 10.1016/j.gerinurse.2025.04.032] [Medline: 40413816]
52. Murayama H, Takase M. Evaluating the effectiveness of digital social robots in reducing loneliness among community-dwelling older adults in Japan: randomized controlled trial and qualitative analysis. JMIR Aging 2025 Oct 22;8:e74422. [doi: 10.2196/74422] [Medline: 41124645]
53. Shade MY, Yan C, Jones VK, Boron JB. Interactive AI routines for pain symptoms and loneliness in older adults. Pain Manag Nurs 2025 Oct;26(5):530-537. [doi: 10.1016/j.pmn.2025.04.003] [Medline: 40328575]
54. Astell A, Clayton D. “Like another human being in the room”: a community case study of smart speakers to reduce loneliness in the oldest-old. Front Psychol 2024;15:1320555. [doi: 10.3389/fpsyg.2024.1320555] [Medline: 38711749]
55. Lee OE, Nah KO, Kim EM, Choi NG, Park DH. Exploring the use of socially assistive robots among socially isolated Korean American older adults. J Appl Gerontol 2024 Sep;43(9):1295-1304. [doi: 10.1177/07334648241236081] [Medline: 38410030]

XSL•FO
RenderX

56. Lee OE, Yun JC, Park DH. Perceptions and experiences of Korean American older adults with companion robots through long-term use: a comparative analysis of robot retention vs. return. Front Public Health 2024;12:1424123. [doi: 10.3389/fpubh.2024.1424123] [Medline: 39624414]
57. Lee OE, Shin H. Social workers' perspectives on socially isolated older adults living with a robot companion. J Gerontol Soc Work 2024 Jul;67(5):621-638. [doi: 10.1080/01634372.2024.2339974] [Medline: 38643400]
58. Lei J, Liang Y, Su Z, Dong P, Liang J, Lin L. Can socially assistive robots be accepted by older people living alone in the community? Empirical findings from a social work project in China. J Gerontol Soc Work 2024 Oct;67(7):995-1012. [doi: 10.1080/01634372.2024.2339992] [Medline: 38682357]
59. Yan C, Johnson K, Jones VK. The impact of interaction time and verbal engagement with personal voice assistants on alleviating loneliness among older adults: an exploratory study. Int J Environ Res Public Health 2024 Jan 16;21(1):100. [doi: 10.3390/ijerph21010100] [Medline: 38248563]
60. Abdollahi H, Mahoor MH, Zandie R, Siewierski J, Qualls SH. Artificial emotional intelligence in socially assistive robots for older adults: a pilot study. IEEE Trans Affect Comput 2023;14(3):2020-2032. [doi: 10.1109/taffc.2022.3143803] [Medline: 37840968]
61. Figueroa D, Yamazaki R, Nishio S, et al. Social robot for older adults with cognitive decline: a preliminary trial. Front Robot AI 2023;10:1213705. [doi: 10.3389/frobt.2023.1213705] [Medline: 38077459]
62. Leung AYM, Zhao IY, Lin S, Lau TK. Exploring the presence of humanoid social robots at home and capturing human-robot interactions with older adults: experiences from four case studies. Healthcare (Basel) 2022 Dec 22;11(1):39. [doi: 10.3390/healthcare11010039] [Medline: 36611499]
63. Jones VK, Hanus M, Yan C, Shade MY, Blaskewicz Boron J, Maschieri Bicudo R. Reducing loneliness among aging adults: the roles of personal voice assistants and anthropomorphic interactions. Front Public Health 2021;9:750736. [doi: 10.3389/fpubh.2021.750736] [Medline: 34957013]
64. Casey D, Barrett E, Kovacic T, et al. The perceptions of people with dementia and key stakeholders regarding the use and impact of the social robot MARIO. Int J Environ Res Public Health 2020 Nov 20;17(22):8621. [doi: 10.3390/ijerph17228621] [Medline: 33233605]
65. Chi NC, Sparks O, Lin SY, Lazar A, Thompson HJ, Demiris G. Pilot testing a digital pet avatar for older adults. Geriatr Nurs 2017;38(6):542-547. [doi: 10.1016/j.gerinurse.2017.04.002] [Medline: 28479082]
66. Chu MT, Khosla R, Khaksar SMS, Nguyen K. Service innovation through social robot engagement to improve dementia care quality. Assist Technol 2017;29(1):8-18. [doi: 10.1080/10400435.2016.1171807] [Medline: 27064692]
67. Hung L, Wong J, Wong KLY, Tan KCK, Lou VWQ. "It's always happy to see me": exploring LOVOT robots as companions for older adults. J Rehabil Assist Technol Eng 2025;12:20556683251320669. [doi: 10.1177/20556683251320669] [Medline: 39935956]
68. Wang Y, Li S. Tech vs. tradition: ChatGPT and mindfulness in enhancing older adults' emotional health. Behav Sci (Basel) 2024 Oct 10;14(10):923. [doi: 10.3390/bs14100923] [Medline: 39457795]
69. Hung L, Gregorio M, Mann J, et al. Exploring the perceptions of people with dementia about the social robot PARO in a hospital setting. Dementia (London) 2021 Feb;20(2):485-504. [doi: 10.1177/1471301219894141] [Medline: 31822130]
70. Chen SC, Moyle W, Jones C, Petsky H. A social robot intervention on depression, loneliness, and quality of life for Taiwanese older adults in long-term care. Int Psychogeriatr 2020 Aug;32(8):981-991. [doi: 10.1017/S1041610220000459] [Medline: 32284080]
71. Hudson J, Ungar R, Albright L, Tkatch R, Schaeffer J, Wicker ER. Robotic pet use among community-dwelling older adults. J Gerontol B Psychol Sci Soc Sci 2020 Oct 16;75(9):2018-2028. [doi: 10.1093/geronb/gbaa119] [Medline: 32789476]
72. Fields N, Xu L, Greer J, Murphy E. Shall I compare thee…to a robot? An exploratory pilot study using participatory arts and social robotics to improve psychological well-being in later life. Aging Ment Health 2021 Mar;25(3):575-584. [doi: 10.1080/13607863.2019.1699016] [Medline: 31851830]
73. Lane GW, Noronha D, Rivera A, et al. Effectiveness of a social robot, "Paro," in a VA long-term care setting. Psychol Serv 2016 Aug;13(3):292-299. [doi: 10.1037/ser0000080] [Medline: 27195530]
74. Cho M, Kim D, Jang M, et al. Evaluating human-care robot services for the elderly: an experimental study. Int J Soc Robotics 2024 Jul;16(7):1561-1587. [doi: 10.1007/s12369-024-01157-7]
75. Lee OEK, Nam I, Chon Y, Park A, Choi N. Socially assistive humanoid robots: effects on depression and health-related quality of life among low-income, socially isolated older adults in South Korea. J Appl Gerontol 2023 Mar;42(3):367-375. [doi: 10.1177/07334648221138283] [Medline: 36326599]
76. Obayashi K, Kodate N, Masuyama S. Assessing the impact of an original soft communicative robot in a nursing home in Japan: will softness or conversations bring more smiles to older people? Int J Soc Robotics 2022;14(3):645-656. [doi: 10.1007/s12369-021-00815-4] [Medline: 34394770]
77. O'Brien K, Light SW, Bradley S, Lindquist L. Optimizing voice‐controlled intelligent personal assistants for use by home‐bound older adults. J Am Geriatr Soc 2022 May;70(5):1504-1509. [doi: 10.1111/jgs.17625] [Medline: 35029296]

XSL•FO
RenderX

Abbreviations

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
SCM: social convoy model
SEM: socio-ecological model
SST: socioemotional selectivity theory

Edited by Y Luo; submitted 25.Apr.2026; peer-reviewed by A Staffini, L Zhou, X Zeng; comments to author 03.Jun.2026; revised version received 17.Sep.2026; accepted 05.Oct.2026; published 20.Nov.2026.

Please cite as:
Kim H, Ren C, Horton H, Jung J, Dong Y, Rizzo V
AI-Based Social Connection Interventions of Emerging Categories of Support for Older Adults: Scoping Review
JMIR Aging 2026;9:e99463
URL: https://aging.jmir.org/2026/1/e99463
doi:10.2196/99463

© Hana Kim, Cheng Ren, Heather Horton, Jeesoo Jung, Yanjun Dong, Victoria Rizzo. Originally published in JMIR Aging (https://aging.jmir.org), 20.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assigning Article-Level Themes in Bibliometric Analysis: Mode-Based Mapping Approach Using JMIR Aging Publications

Sam Yu-Chieh Ho[1,2], MD; Kang-Ting Tsai[3,4], MD

[1]Department of Emergency Medicine, Chi Mei Medical Center, No. 901, Chung Hwa Road, Yung Kung Dist, Tainan, Taiwan

[2]Center for Integrative Medicine, Chi Mei Medical Center, No. 901, Chung Hwa Road, Yung Kung Dist, Tainan, Taiwan

[3]Department of Family Medicine, Chi Mei Medical Center, 901 Chung Hwa Road, Yung Kung Dist, Tainan, Taiwan

[4]Department of Family Medicine, Well Clinic, No. 302, Yunong Rd, East Dist, Tainan, Taiwan

Corresponding Author:
Kang-Ting Tsai, MD
Department of Family Medicine, Chi Mei Medical Center, 901 Chung Hwa Road, Yung Kung Dist, Tainan, Taiwan

Abstract

Background: Although clustering techniques are commonly used in bibliometric analysis to identify research themes, few studies systematically assign these themes back to individual articles. This gap limits the interpretability of findings and hinders granular, article-level longitudinal analysis.

Objective: This study introduces the Theme Assignment Algorithm for Articles (TAAA), a data-driven framework designed to map clustered themes to individual publications. We demonstrate its utility by identifying dominant research patterns and thematic shifts within JMIR Aging.

Methods: TAAA was applied to 434 JMIR Aging articles published between 2020 and 2025. Keywords were harvested from 3 sources: Web of Science Core Collection (WoSCC) Keywords Plus, author-provided keywords, and abstract-derived terms. These were grouped into thematic clusters using the “following leader clustering algorithm”. The TAAA, implemented via R and a web-based application, determined each article’s primary theme using a statistical model to create a discrete article-level variable. Core themes were identified via h-index computation. Analytical visualization included Kano and Sankey diagrams, alongside volcano plots and heatmaps. The framework’s robustness was further tested by applying a “differentially expressed genes” analogy to map “unknown” core metadata across “known” pre/post publication stages using Cohen kappa as the extent of mapping power.

Results: Analysis across the 3 keyword sources yielded 9, 7, and 9 core themes, respectively. The most prominent themes identified were HEALTH (39.4%), OLDER ADULTS (43.1%), and DEMENTIA (25.3%). Notably, DEMENTIA emerged as a consistent core theme across all sources and visual layers, validating the TAAA’s ability to capture cross-source thematic coherence. The adaptation of dual heatmaps demonstrated the algorithm’s capacity for comparative bibliometric mapping in JMIR Aging. A mapping precision of 0.33 provided quantitative evidence of a 2-stage publication pattern, though the separation between stages was not strictly confined to predefined time intervals.

Conclusions: The TAAA framework provides a replicable, scalable, and interpretable method for article-level thematic assignment. Its ability to uncover consistent research patterns—specifically the dominance of dementia-related studies in JMIR Aging—demonstrates its value for bibliometricians and its potential adaptability to other domains, such as bioinformatics-inspired meta-analyses.

(JMIR Aging 2026;9:e79906) doi:10.2196/79906

KEYWORDS

bibliometric analysis; theme assignment; following leader clustering algorithm; FLCA clustering; R programming; visual analytics

Introduction

Overview

JMIR Aging is a peer-reviewed, open-access journal dedicated to advancing digital health, technology, and innovation for older adults [1]. It publishes research on eHealth interventions, gerontechnology, aging populations, and health services aimed at improving care and outcomes for aging societies [2]. Despite this clear mission, the underlying thematic structure of its publications remains underexplored, limiting the ability to systematically trace research foci, emerging trends, and evolving priorities [3]. Understanding the journal’s intellectual landscape requires an approach that integrates bibliometric theory, network clustering, and theme identification into a coherent framework.

Theoretical Background: Themes and Co-Word Structures in Bibliometrics

In bibliometric research, thematic classification plays a central role in mapping intellectual structures and identifying research

XSL•FO
RenderX

frontiers [4-6]. The theoretical foundation lies in co-word analysis, which assumes that the co-occurrence of keywords reflects conceptual proximity among research topics [4-9]. Tools such as Bibliometrix, VOSviewer, and CiteSpace operationalize this principle by constructing keyword networks and identifying clusters that represent thematic communities [7-9].

However, existing methods primarily operate at the cluster level, identifying thematic groupings among keywords rather than assigning a single dominant theme to each article [4-6,10,11]. This limits the interpretability of bibliometric maps and constrains microlevel analyses, such as tracing thematic shifts across journals, institutions, or regions.

Need for Article-Level Theme Assignment

Although advanced methods such as the following leader clustering algorithm (FLCA) [12] and latent Dirichlet allocation–based topic modeling can uncover latent thematic structures, a key gap remains: Cluster-level theme discovery does not automatically translate into clear, article-level theme representation—for example, mapping text-driven patterns (from item/factor structures) into labels that describe article or person attributes.

To date, few bibliometric models explicitly address this gap with a formal mechanism for assigning cluster-derived themes to individual publications. Without this linkage, thematic labeling becomes less consistent, and cross-sectional or longitudinal analyses are harder to interpret and compare [4,5].

To address this theoretical and methodological gap, an article-level theme assignment algorithm (Theme Assignment Algorithm in Articles [TAAA]) is required. TAAA formalizes the process of identifying each article's dominant theme from multiple candidate clusters, enhancing both interpretability and reproducibility in bibliometric studies.

An Application or R-Based Framework for Thematic Assignment and Visualization

Although some prior studies have conceptually mentioned "theme-by-article" analyses [4,5], no standardized nor open-source R-based (or application) implementation currently exists [13]. The proposed framework, TAAA, operationalizes the statistical mode of an article's keyword cluster assignments to determine its most representative theme (ie, from item/bifactor structures into labels that describe article or person attributes). This method enhances computational efficiency while maintaining theoretical consistency with co-word analysis principles.

Furthermore, the framework integrates visual analytic tools via an application (for the general public) and R-based scripts (for academic researchers), including Kano diagrams [14,15] and Sankey visualizations [11], to illustrate how the newly created theme variable propagates across articles, institutions, and nations. This integration advances existing bibliometric visualization paradigms by enabling article-level interpretability within cluster-based systems.

Study Aim and Contribution

This study aimed to develop and demonstrate an application and R-based TAAA for assigning dominant themes to publications in JMIR Aging. The contribution of this work is 2-fold:

1. Conceptual: It extends bibliometric theory by linking cluster-level co-word analysis to article-level thematic representation, addressing a critical gap in interpretive granularity.
2. Methodological: It provides a transparent, reproducible, and scalable application or R-based workflow for theme assignment and visualization.

We hypothesized that TAAA can accurately capture each article's primary thematic focus and improve interpretability in large-scale bibliometric analyses. We further expected the framework to be adaptable to other journals and research domains, as demonstrated through several practical examples.

Methods

Data Source

We retrieved metadata for 434 articles published in JMIR Aging from 2020 to 2025 using the Web of Science Core Collection (WoSCC) database. The dataset includes bibliographic information such as author names, affiliations, titles, abstracts, and keywords. Study data are provided in Multimedia Appendix 1. Since all records in WoSCC are publicly available and contain no identifiable patient information, institutional review board approval was not required.

Definition of Key Terms

To ensure conceptual clarity, the following key terms are defined within the context of this study: keyword source, item testlet, person theme, mapping power (MP), and theme variable.

Keyword source refers to the origin of the keywords used for analysis. This study distinguishes between 3 types: (1) author-provided keywords: terms explicitly selected by the authors; (2) Keywords Plus: terms automatically generated by the WoSCC; and (3) abstract-derived keywords: terms extracted from the article abstracts using a custom R script.

Item testlet draws from educational assessment terminology, where a "testlet" represents a cluster of keywords grouped by a common thematic thread. In this study, testlets were generated using FLCA [12], where each cluster reflects a cohesive research topic within the broader keyword network.

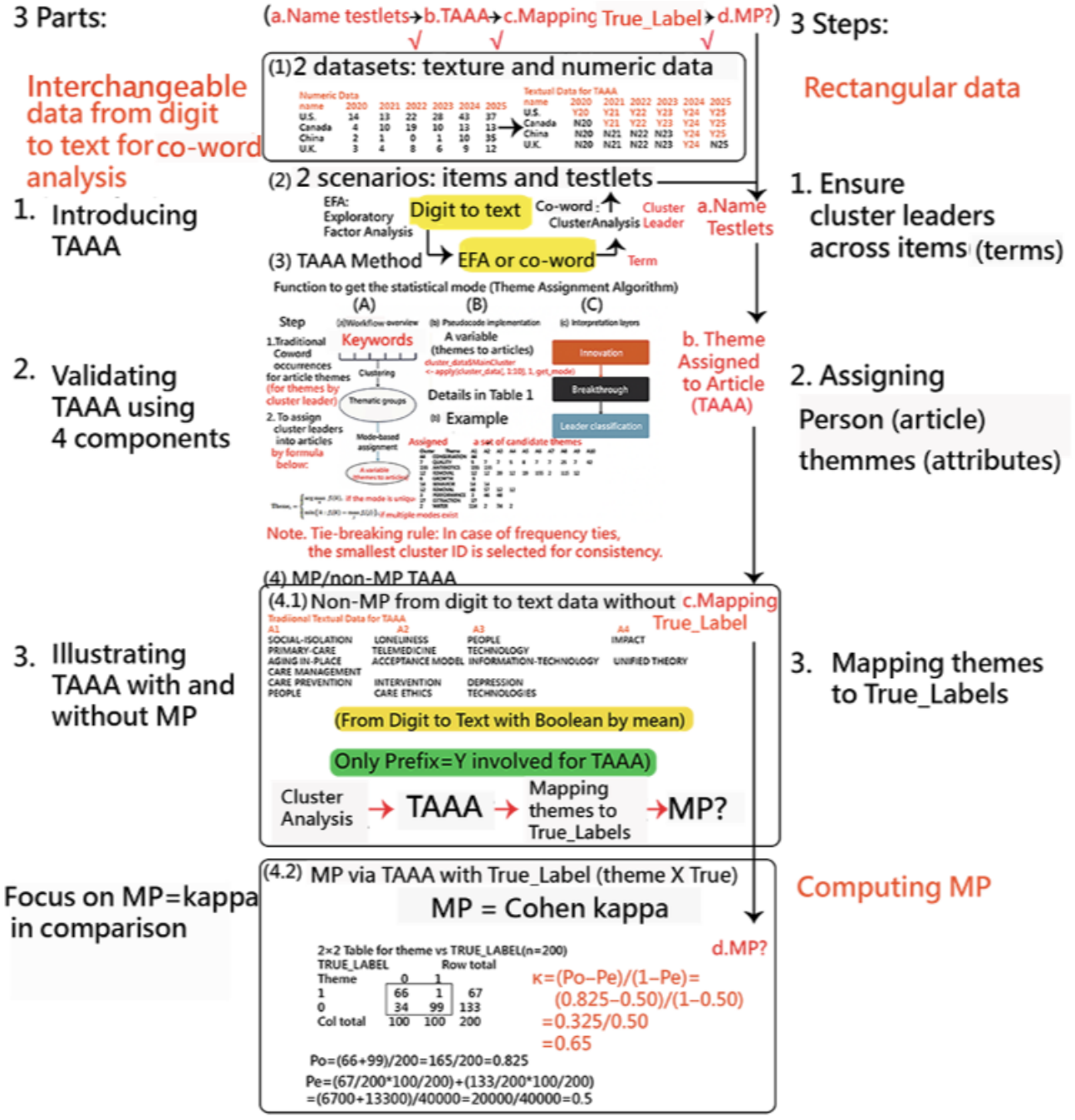
Person theme is the primary topical label assigned to an individual article. Using TAAA [4,5], the "person theme" is determined by identifying the statistical mode (the most frequent keyword cluster) associated with that specific article (or person).

MP is a metric used to validate the consistency of thematic classification. As illustrated in Figure 1, MP is computed by mapping keyword clusters (item testlets) to article-level classifications (person themes) and comparing them against established true labels.

XSL•FO
RenderX

For the theme variable, cluster leaders were first generated separately for each keyword source and then merged by mapping item-level terms to person-level themes. This design enables rigorous cross-source comparison and validation of mapping performance against external ground-truth labels (TRUE_LABEL), such as gender or cancer versus noncancer status.

Figure 1. Diagram showing the 3 parts and 3 steps for Theme Assignment Algorithm in Articles (TAAA) in 2 datasets and 2 scenarios (mapping power [MP] vs non-MP with and without True-Label for articles or persons).



Summary Performance Report

A clear and concise reporting format focusing on 10 key metadata elements, including top 10 themes, was used to improve readability and interpretation [11], offering an alternative to the extensive graphs and tables commonly found in traditional bibliometric analyses.

TAAA Applied to This Research

This research consists of 3 parts: (1) introducing TAAA, (2) validating TAAA, and (3) illustrating TAAA with and without MP.

XSL•FO
RenderX

Introducing TAAA

Creation of a Theme Variable for Co-Word Occurrence Analysis

This study used 3 scenarios to assign themes to articles using 3 keyword sources: Web of Science Keywords Plus, author keywords, and abstract keywords. These were processed via an R script [16] for the extraction of keywords in articles. Traditionally, article keywords are clustered into thematic groups using algorithms such as FLCA [12] or tools like CiteSpace [8], where the largest bubble typically represents the dominant theme in each cluster [10,11].

The TAAA was further implemented in R [17], leveraging the statistical mode to identify the most frequently occurring theme among a given article's clustered keyword assignments based on theme with the top one by count in each cluster. This newly created variable ("theme") was added to each article for downstream analysis as shown in part 3 in Figure 1 and Textbox 1.

Textbox 1. Conversion of multiple candidate clusters per article into a single dominant theme with R script.

```
# Function: get_mode()
# Purpose : Determine the dominant theme (mode) among candidate clusters
# Context : Used in the Theme Assignment Algorithm for Articles (TAAA) get_mode <- function(x) {
# Step 1. Clean input data
# Remove blank entries ("") and missing values (NA)
x <- x[x != "" & !is.na(x)]
# Step 2. Handle empty input case
# If no valid data remain after cleaning, return NA (no theme can be assigned)
if (length(x) == 0) return(NA)
# Step 3. Compute frequency table of cluster occurrences
# The 'table()' function counts how many times each cluster ID appears freq <- table(x)
# Step 4. Check for tie condition (all clusters occur equally often)
# If all frequencies are identical, select the smallest cluster number
# This ensures deterministic (non-random) and reproducible assignment
if (length(unique(freq)) == 1) {
return(min(as.numeric(names(freq)))) # smallest cluster ID (tie case)
# Step 5. Otherwise, return the most frequent cluster (the statistical mode)
} else {
return(names(freq)[which.max(freq)]) # dominant cluster (standard mode)
}
}
#Example usage:
# Apply the function across columns A1–A10 (keyword clusters per article)
# and create a new variable "MainCluster" representing the assigned theme cluster_data$MainCluster <- apply(cluster_data[, 1:10], 1, get_mode)
```

The approach improves interpretability and computational efficiency by resolving thematic ambiguity, as defined by Equation (1):

$$\text{Theme}_i = \begin{cases} \min\{k \mid f_i(k) = \max_j f_i(j)\}, & \text{if all } f_i(k) \text{ are equal} \\ \text{Mode}(\text{Cluster}_{\text{keywords}}), & \text{otherwise} \end{cases} \quad (1)$$

where $f_i(k)$ denotes the frequency of cluster k among the keywords of article i.

For instance, consider an article whose keywords are distributed across 3 clusters—Cluster 1, Cluster 2, and Cluster 3—with respective frequencies of 1, 3, and 2. The algorithm first computes the frequency of each cluster ($f_i(k)$) and identifies Cluster 2 as having the highest count. Therefore, the assigned theme for this article is Cluster 2, as it represents the statistical mode of the clustered keyword assignments.

If all clusters had occurred with equal frequency (eg, 2, 2, 2), the algorithm would select the smallest cluster identifier to ensure deterministic and reproducible assignment of the article's dominant theme, as shown in part 3 of Figure 1.

For each article, the assigned theme corresponds to the statistical mode of its keyword cluster assignments. In cases of ties (ie, multiple clusters with equal frequency), the cluster with the smallest numerical identifier is selected to ensure consistency and reduce thematic ambiguity.

XSL•FO
RenderX

TAAA in Bibliometric Analysis Based on Textual Data

The TAAA framework was applied using visual tools such as Kano [14,15] and Sankey [11] diagrams to illustrate thematic distributions across articles, countries, and journals based on co-word occurrence data (eg, "theme" and "country" variables in two columns). The R script [18] followed methods described in prior studies [11]. Additionally, leader classification was performed using the dimension coefficient (DC) [19,20], defining leadership categories as super (DC ≥0.8), quasisuper (DC≥0.7), strong (DC≥0.6), moderate (DC≥0.5), and weak (DC<0.5).

Venn Diagram Used to Identify the Common Themes From 3 Sources

A Venn diagram was used in R [21] to extract the overlaid themes from the 3 sources: Keywords Plus, author-defined keywords, and abstract. The screened themes denote the core elements in the studied journal.

Validating TAAA

We evaluated TAAA using 4 complementary validation components: interrater and intrarater validation, external validation and robustness across clustering methods, bootstrapping stability, and out-of-sample holdout robustness.

Interrater and Intrarater Validation (Core Theme Subsample)

A stratified subsample was drawn from dominant ("core") themes (86/431, 20% of core theme articles). Two independent reviewers (2 study authors) who were blinded to TAAA outputs assigned themes using the same predefined label set based on titles and abstracts. Agreement was assessed for TAAA vs Reviewer 1, TAAA vs Reviewer 2, and Reviewer 1 vs Reviewer 2 using confusion matrices and Cohen κ [22], where commonly cited cutoffs [23] are <0.00, poor; 0.00 - 0.20, slight; 0.21 - 0.40, fair; 0.41 - 0.60, moderate; 0.61 - 0.80, substantial; and 0.81 - 1.00, almost perfect.

External Validation and Robustness Across Clustering Methods

Because no universally accepted ground truth exists for topic clustering in keyword co-occurrence networks, we assessed robustness by comparing TAAA assignments under two clustering strategies applied to the same keyword set: (1) FLCA term-to-theme membership from our pipeline (terms_nodes.csv) and (2) Louvain communities from modularity optimization on the keyword co-occurrence graph.

Since cluster counts may differ, we evaluated concordance using (1) a majority-overlap mapping from Louvain communities to FLCA themes based on article-level cross-tabulation and (2) a top-*k* hierarchical criterion in which an article was counted as agreeing if its FLCA theme fell within the top-*k* most frequent FLCA themes inside its Louvain community (*k*=1, 2, 3, 4). Top-1 corresponds to strict one-to-one concordance; larger *k* quantifies hierarchical nesting where a coarse Louvain community may contain multiple FLCA subthemes.

Bootstrapping Stability

To assess sampling stability, we performed bootstrap resampling with replacement on the core-theme subsample. For each repetition, κ was recomputed and summarized by its mean and 2.5th-97.5th percentile interval. Jensen-Shannon (JS) distance and mean Spearman correlation across resamples were also computed to quantify stability [24-29].

Out-of-Sample Holdout Robustness

To evaluate robustness under dataset shift, we conducted repeated random holdout experiments with a prespecified holdout fraction. On holdout sets, we computed mean Spearman correlation and JS distance between training and holdout rankings to quantify generalization [30,31].

Illustrating TAAA With and Without MP

Textual Terms From Digital Data Assessed Using MP

This text integrates the technical logic of the TAAA with the validation process through MP, supporting both "non-MP" (digit-to-text without true labels) and "MP" (validation against true labels) scenarios to ensure robust bibliometric mapping.

The workflow illustrates the transition in Figure 1, from (1) multisource datasets to (2) the identification of item testlets. (3) The TAAA applies a mode-based function (with tie-breaking rules) to assign a primary person theme to each article (or person). (4) The process concludes with the computation of MP using Cohen kappa [22,23], validating the thematic assignment against true labels to ensure classification accuracy.

Steps for Implementing TAAA via R and a Web-Based Application

For the abstract keyword extraction, the input data came from a CSV file named methodsxabstract.csv containing columns for article title and abstract. In R [16] and the application [24], combined text is stored in df$CombinedText. Abstract keywords are extracted and saved to methodsxabstract_with_keywords_topics.csv.

TAAA Implementation via R and a Web-Based Application

Keywords across multiple columns are provided in top10keywords.csv. Themes are assigned per article using the TAAA model [17,25] and saved sequentially to methodsxabstract_with_keywords_topics.csv. The resulting "theme" variable is used to analyze co-word occurrences, visualizing thematic distributions across articles, countries, and journals (with variables like "theme" and "country" in two columns).

Applications of TAAA for Textual and Numeric Data With MP

The transition from numeric values to article-level (person-level) themes, as shown in part 1 in Figure 1, follows a structured 8-step process:

1. Dichotomization: Convert each numeric item into binary scores using the item-specific mean as a threshold.
2. Tokenization: Transform binary scores into textual tokens by prefixing item labels with "Y" (yes) or "N" (no).

XSL•FO
RenderX

3. Occurrence construction: Retain only "Y-type" tokens to simulate keyword occurrences within each row.
4. Co-word clustering: Apply clustering algorithms to these tokens to derive thematic "cluster leaders" representing coherent testlets across data columns.
5. Person-level assignment: Use the TAAA to map these term-based themes to individual rows (persons), establishing a dominant theme for each entry.
6. External validation (optional): If external reference labels are available, compute MP using Cohen kappa to assess classification accuracy.
7. Web implementation: To enhance user accessibility, the analytical framework [32-34] was deployed as a cloud-based application on Google App Engine (GAE) [35]. This provides a user-friendly interface for researchers to replicate the analysis without requiring local environment configuration, as done with R scripts.
8. Visual interpretation: Analyze the results through an automated HTML report [32,33] featuring Kano diagrams, Sankey plots, volcano plots, and heatmaps.

Data Sources for Application Examples

The application of TAAA to textual entities via the application was conducted with 434 articles published in JMIR Aging from 2020 to 2025. The application of TAAA to numerical entities via the application occurred via virtual data as a test on STEM and ARTS for men and women and as a questionnaire on 2 domains. In addition, the TAAA was applied to Gene Expression Omnibus (GEO) data via the application: A gene expression profile, GSE8401 [36], including 31 primary melanoma and 52 metastatic melanoma (MM) clinical samples, was downloaded from the GEO database [37]. The differentially expressed genes between MM and primary melanoma were assessed using MP.

R Packages and Application Used in This Study

All visualizations and data processing were conducted using the R platform [16,17,20,23] or GAE [35]. Kano diagrams were generated in the application or R scripts, with all statistical analyses performed in R version 4.1.3 [13]. Visualizations and results were produced within the RStudio integrated development environment [38]. The way to conduct this study using R scripts is described in Multimedia Appendix 2.

Results

Summary Performance Report for Top 10 Article Elements

Figure 2 presents a comprehensive summary of 434 publications in JMIR Aging (h-index=28 [39]) analyzed across 10 bibliometric dimensions—country, institution, department, year, article type, author, research area, keyword frequency, and citation count.

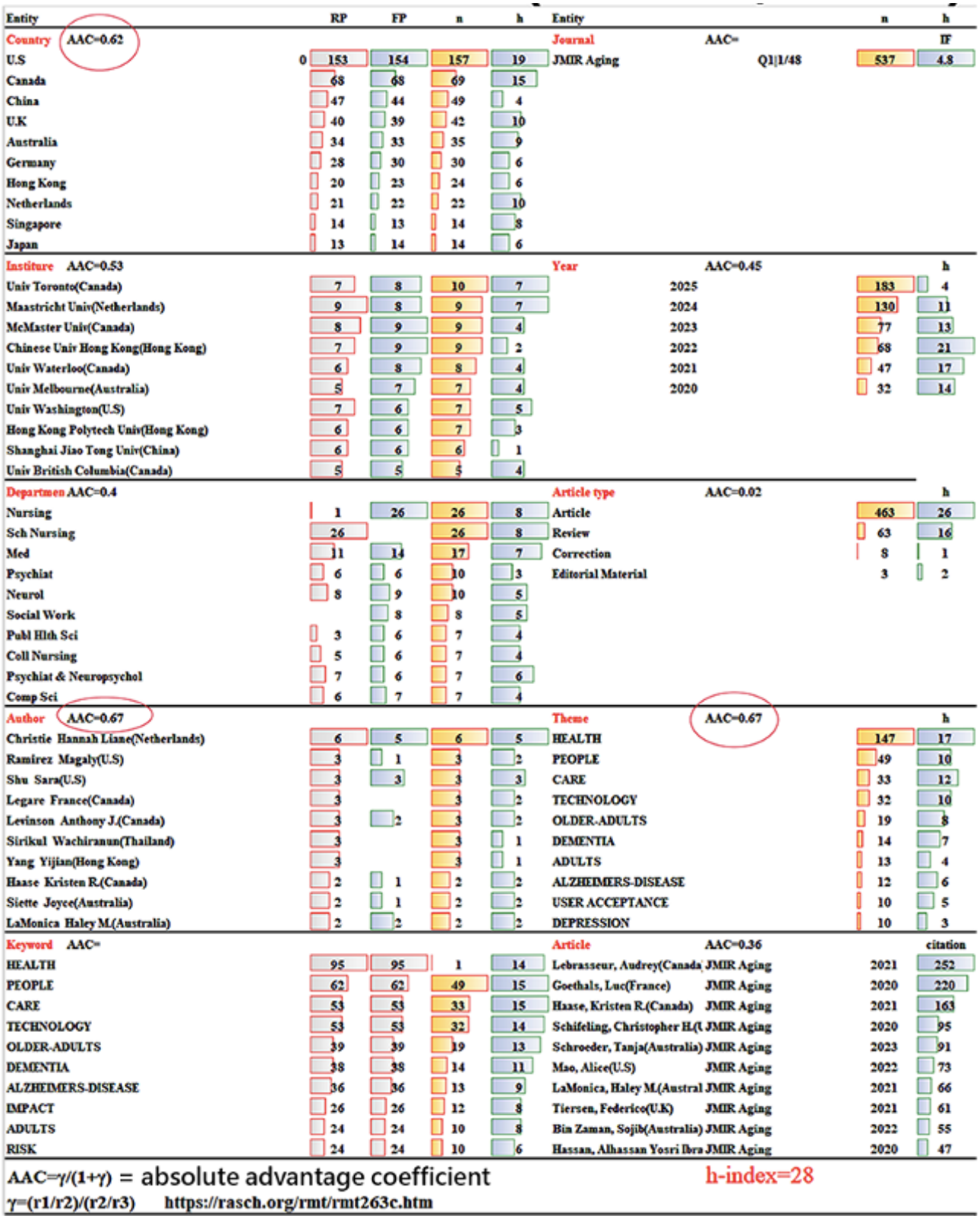
The United States, Canada, and the United Kingdom led in publication output, while the University of Toronto (Canada) and Maastricht University (the Netherlands) were among the most productive institutions. Nursing and medical departments contributed prominently, with the majority of papers classified as original research articles.

Notably, Hannah Liane Christie (the Netherlands) was the most prolific author (6 articles [40-45]), whereas Audrey Lebrasseur et al (Canada) produced the most-cited study [46], receiving 252 citations for a rapid review on COVID-19's impact on older adults. Core keywords such as HEALTH, PEOPLE, CARE, and TECHNOLOGY reflect the journal's central focus on digital health and aging innovation; see the keyword and theme panels in Figure 2.

Accordingly, this performance overview provides a structural snapshot of JMIR Aging's research landscape, enabling targeted insight into its geographic reach, institutional strengths, and topical emphasis.

XSL•FO
RenderX

Figure 2. Summary performance report for 10 article metadata categories in JMIR Aging (n=537). FP: first author; RP: first author.



TAAA for Theme Extraction and Implementation

Figure 3 illustrates how the TAAA identified dominant themes across 3 keyword sources: Keywords Plus (external indexing terms): 9 core themes (Figure 3A); author-supplied keywords: 7 core themes representing self-identified research emphases (Figure 3B); and abstract-derived keywords: 9 themes reflecting content-driven topic extraction (Figure 3C).

XSL•FO
RenderX

Figure 3. Core themes extracted from JMIR Aging from 3 keyword sources including (A) Keywords Plus, (B) author-defined keywords (top 7 themes contributing to the h-index), and (C) abstract-derived keywords (top 9 themes contributing to the h-index). DC: dimension coefficient.

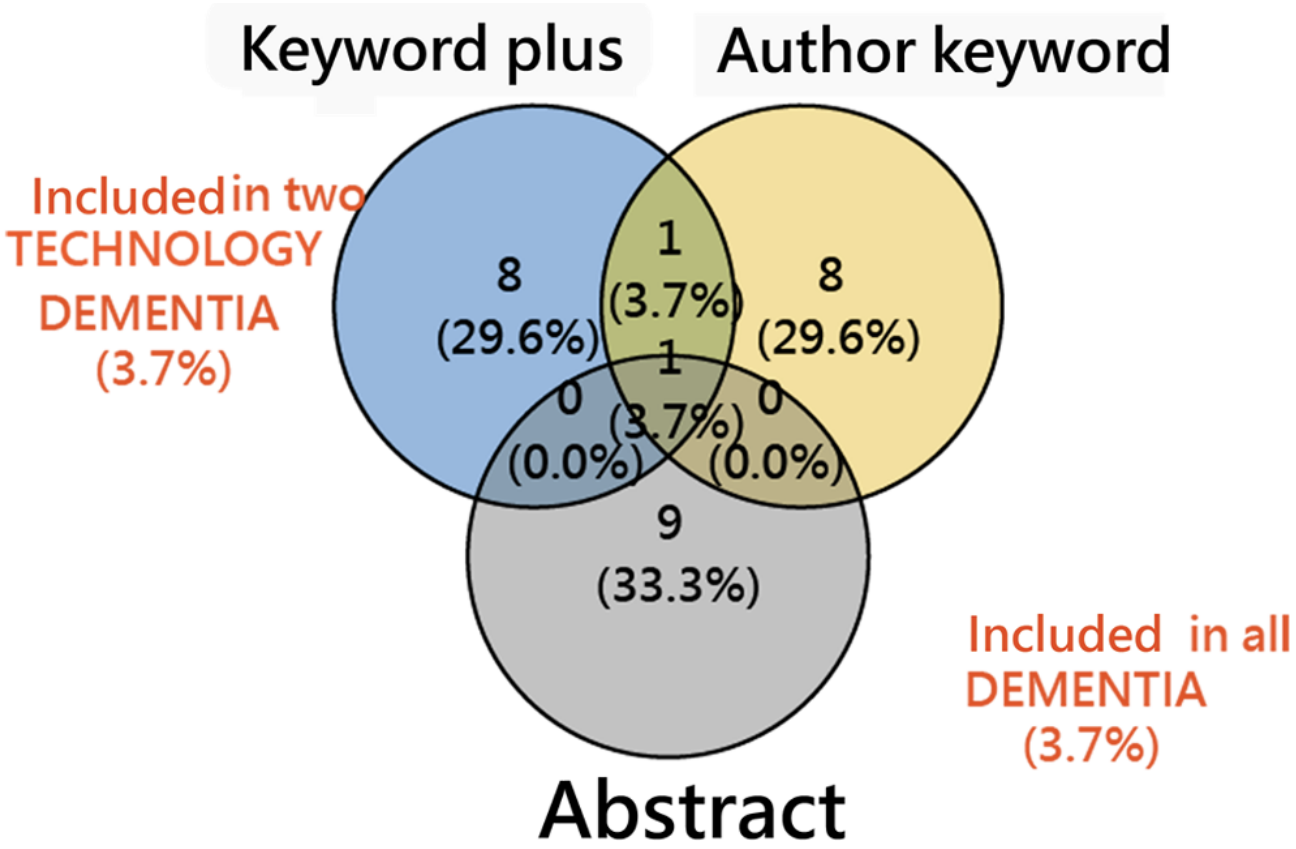
XSL•FO
RenderX

Across all 3 scenarios, the top themes were HEALTH (108/282, 39.4%), OLDER ADULTS (140/343, 43.1%), and DEMENTIA (38/158, 25.3%).

To illustrate theme overlap, Figure 4 (Venn diagram) shows that DEMENTIA was the only theme consistently identified in all 3 TAAA runs, underscoring its centrality in the journal's research portfolio.

Hence, the convergence of DEMENTIA across independent keyword sources validates TAAA's reliability in identifying dominant thematic foci and highlights the journal's sustained emphasis on cognitive health in aging research.

Figure 4. Venn diagram showing DEMENTIA as the only theme overlapping across all 3 TAAA approaches.



TAAA for Validation

Subsample Interrater Validation (Core Themes)

Subsample interrater validation on 57 core theme records (20% sampled from 431 articles) showed substantial agreement between TAAA and Reviewer 1 (κ=0.788; 84.2% exact agreement) and moderate to substantial agreement between TAAA and Reviewer 2 (κ=0.672; 75.4% exact agreement). Human-human agreement was moderate (κ=0.597; 68.4% exact agreement), indicating inherent ambiguity in theme boundary decisions.

Misclassifications were concentrated in a small number of adjacent theme pairs, particularly involving the dominant theme (T1) and conceptually related themes (eg, T13/T15), consistent with expected instability in low-prevalence categories (Table 1).

XSL•FO
RenderX

Table . Subsample interrater validation for core themes (n=57 core theme records).

	T1[a]	T3[b]	T2[c]	T5[d]	T4[e]	T13[f]	T9[g]	T15[h]	T10[i]
TAAA[j] versus Reviewer 1 (κ=0.788)									
T1	24	0	0	0	0	1	0	4	0
T3	0	5	0	0	0	0	2	0	0
T2	0	0	7	0	0	0	0	0	0
T5	0	0	0	3	0	0	0	0	0
T4	0	0	0	0	2	0	1	0	0
T13	0	0	0	1	0	1	0	0	0
T9	0	0	0	0	0	0	2	0	0
T15	0	0	0	0	0	0	0	2	0
T10	0	0	0	0	0	0	0	0	2
TAAA versus Reviewer 2 (κ=0.672)									
T1	22	1	0	0	0	2	2	0	2
T3	0	6	0	0	0	0	1	0	0
T2	1	0	6	0	0	0	0	0	0
T5	0	1	0	2	0	0	0	0	0
T4	0	0	0	0	2	0	0	0	1
T13	0	0	0	0	0	1	0	0	1
T9	0	0	0	0	0	1	1	0	0
T15	0	0	0	0	0	0	0	2	0
T10	0	1	0	0	0	0	0	0	1
Reviewer 1 versus Reviewer 2 (κ=0.597)									
T1	18	1	0	0	0	2	2	0	1
T3	0	5	0	0	0	0	0	0	0
T2	1	0	6	0	0	0	0	0	0
T5	0	1	0	2	0	0	0	0	1
T4	0	0	0	0	2	0	0	0	0
T13	1	0	0	0	0	1	0	0	0
T9	0	1	0	0	0	1	2	0	1
T15	3	0	0	0	0	0	0	2	1
T10	0	1	0	0	0	0	0	0	1

[a]T1: core theme 1.

[b]T3: core theme 3.

[c]T2: core theme 2.

[d]T5: core theme 5.

[e]T4: core theme 4.

[f]T13: core theme 13.

[g]T9: core theme 9.

[h]T15: core theme 15.

[i]T10: core theme 10.

[j]TAAA: Theme Assignment Algorithm in Articles.

When the human-human agreement (κ=0.597) is lower than the TAAA-human agreement (κ=0.788 and 0.672), it leads to several significant implications regarding the reliability and utility of the TAAA algorithm as a "stabilizing force." Although human reviewers may be influenced by subjective interpretation or fatigue, the TAAA applies a consistent, rule-based mode

XSL•FO
RenderX

logic that reduces the "noise" inherent in manual thematic assignment. In bibliometric studies involving hundreds or thousands of articles, manual review is prone to increasing error rates.

Agreement Between FLCA and Louvain

Among 431 records, 373 contained at least one keyword. FLCA yielded 53 themes in the term-membership file, whereas Louvain produced 15 communities on the same mapped keyword graph.

Because the two methods differ in granularity, we evaluated hierarchical concordance by allowing each Louvain community to correspond to its Top-*k* most frequent FLCA subthemes. Agreement increased from 49% (183/373; Top-1) to 65.5% (244/373; Top-2), 77.5% (289/373; Top-3), and 82.5% (308/373; Top-4), indicating that most Louvain communities are unions of a small number of FLCA subthemes. To report a chance-corrected measure at the Louvain granularity (*k*=15), we collapsed FLCA themes to Louvain communities via majority overlap at the term level and computed Cohen κ between the resulting labels, yielding κ=0.399, indicating an upper fair effect (0.21 - 0.40) [24].

Bootstrapping Stability (120 Repetitions)

Across 120 bootstrap resamples, the number of core themes was k=9, and the thematic structure showed high stability (Table 2). The mean JS distance between the bootstrap and reference distributions was 0.057 (SD 0.016), indicating only minor shifts in the theme proportion profile under resampling. The mean Spearman rank correlation was 0.896 (SD 0.071), suggesting that the relative ordering of theme importance was largely preserved across resamples.

Table . Bootstrapping stability and robustness.

Measurement	Repetitions, n	Core themes, k	Holdout fraction	JS[a] distance, mean (SD)	Spearman rank, mean (SD)
Bootstrapping	120	9	—[b]	0.06 (0.02)	0.90 (0.07)
OOL[c] holdout robustness	120	9	0.3	0.12 (0.03)[d]	0.72 (0.13)

[a]JS: Jensen-Shannon.

[b]Not applicabe.

[c]OOL: out-of-likelihood.

[d]Train, holdout.

Out-of-Distribution Holdout Robustness (120 Repetitions)

Under 120 repeated holdout (out-of-likelihood/out-of-sample) evaluations with a 30% holdout fraction, the model retained k=9 core themes (Table 2). The mean JS distance between the training and holdout theme distributions was 0.124 (SD 0.034), which is larger than the bootstrap JS distance, reflecting the expected distributional shift induced by data splitting. The mean Spearman correlation between training and holdout rankings was 0.724 (SD 0.127), indicating moderate agreement in theme ordering and suggesting that the primary thematic structure generalizes reasonably well, while lower-prevalence themes are more sensitive to sample partitioning.

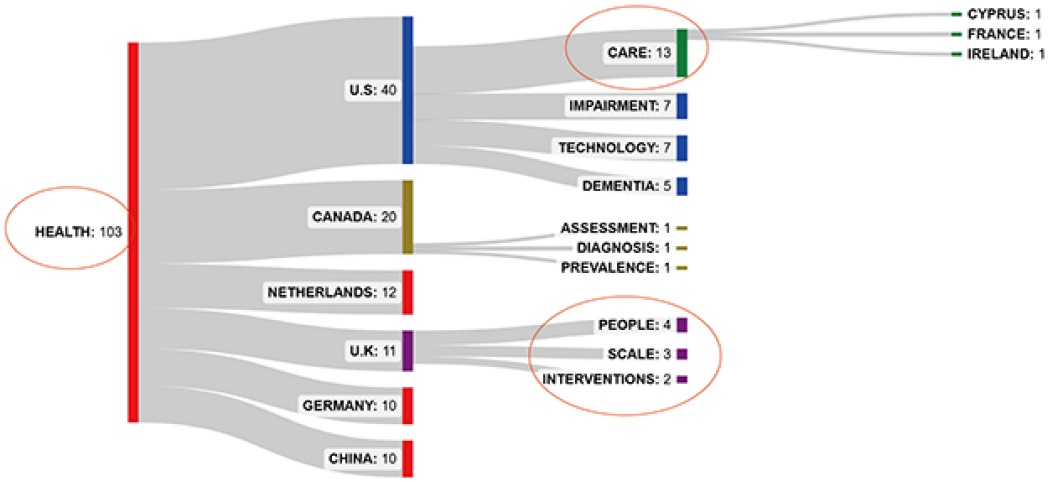
TAAA for Textual Entities via the Application

Figure 5 demonstrates how assigned themes distribute across countries, highlighting geographic research orientations. The United States remained a research leader, followed by Canada and the United Kingdom.

XSL•FO
RenderX

Figure 5. Textual co-word with a new variable of theme: (A) Sankey visualization by country-theme, (B) rectangular data, and (C) create network chart by theme and country via an application [25]. AAC=0.58 <0.7 without dominance role for health=[(99/62)/(62/53)]/(1+[99/62)]/(62/53)]). AAC: absolute advantage coefficient.

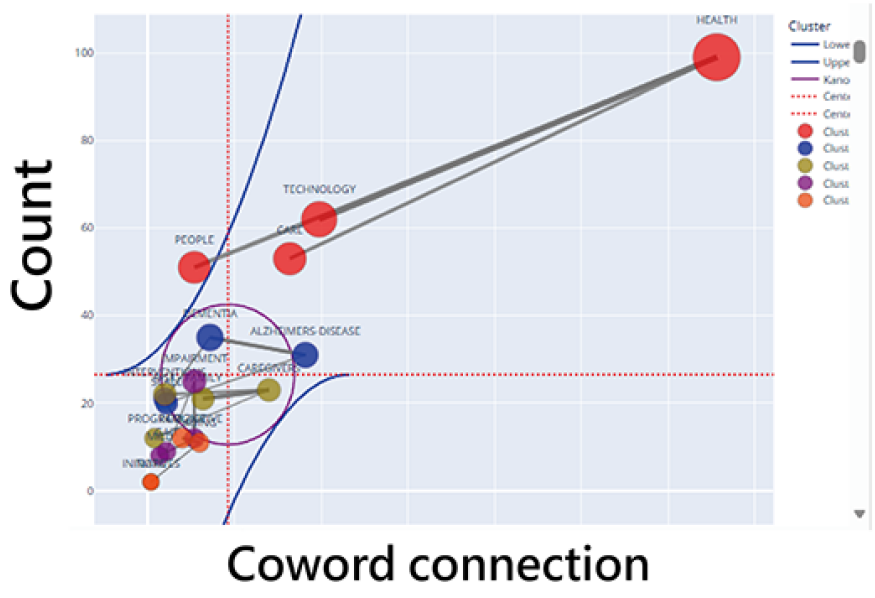
(A)



(B)

name	HEALTH	CARE	TECHNOLOGY	DEMENTIA	PEOPLE	RISK	IMPAIRMENT	CAREGIVERS	SCALE	ADULTS
U.S	41	13	7	5	2	2	7	6	2	3
Canada	20	8	4	1	3	0	0	1	0	0
U.K	12	1	2	3	4	0	1	1	3	1
Australia	13	2	3	0	0	1	0	0	0	1
China	11	2	1	0	1	1	1	0	0	0
Germany	10	1	0	2	3	0	1	0	1	1
Netherlands	12	1	4	1	0	0	0	0	0	0
Hong Kong	5	0	1	2	1	1	0	0	1	1
Japan	4	1	2	0	0	1	0	0	0	0
Singapore	4	0	4	0	0	0	0	0	0	0

(C)



DEMENTIA and CARE are closely associated with US authors, whereas ASSESSMENT is linked to Canadian research. SCALE and PEOPLE themes are common in UK-based publications. The Sankey diagram in Figure 5A visualizes these

XSL•FO
RenderX

relationships—flow width represents the proportional connection between countries and their dominant themes.

In Figure 5B, a network chart of country-theme was uploaded by the input form of rectangular data via an application created by our research team [25,34].

Together, these analyses reveal cross-level variation in thematic prominence—DEMENTIA dominates within specific institutions but shares space with broader aging-related themes at the national level, suggesting specialized research clusters driving global impact.

TAAA for Numeric Entities via the Application

Metadata for Publications in JMIR Aging

Dual heatmaps were constructed to examine metadata structure using complementary selection rules: statistical significance (*P* value) and effect magnitude (log10 fold change). Core metaterms were defined as the union of the top 10 up- and downregulated terms under each criterion. These terms were binarized using the mode of term presence > mean across persons and subjected to co-word analysis, followed by FLCA clustering (target_clusters=2).

Resulting term clusters were propagated to publications via TAAA and displayed beneath each heatmap alongside TRUE_LABEL, reflecting a 2-stage publication structure, with contrast themes denoted by social work versus ALZHEIMERS-DISEASE and editorial material versus China.

Although visual correspondence is evident, the modest mapping precision (MP=0.33) indicates residual misalignment between data-driven themes and external labels (Figure 6), suggesting that the thematic transition cannot be strictly partitioned into the periods 2020 - 2022 and 2023 - 2025.

XSL•FO
RenderX

Figure 6. Dual heatmaps for the 66 metaterms in JMIR Aging publications (mapping power [MP]=1.0): (A) heatmap (*P* value–selected) of the top 10 UP + top 10 DOWN metaterms, with increasing *P* values for ALZHEIMERS-DISEASE, Germany, United States, and RISK; (B) heatmap (log10FC-selected) of the top 10 UP + 1op 10 DOWN metaterms, wiith log10FC for China, HongKong, Japan, and RISK; and (C) the probability matrix. TAAA: Theme Assignment Algorithm in Articles.

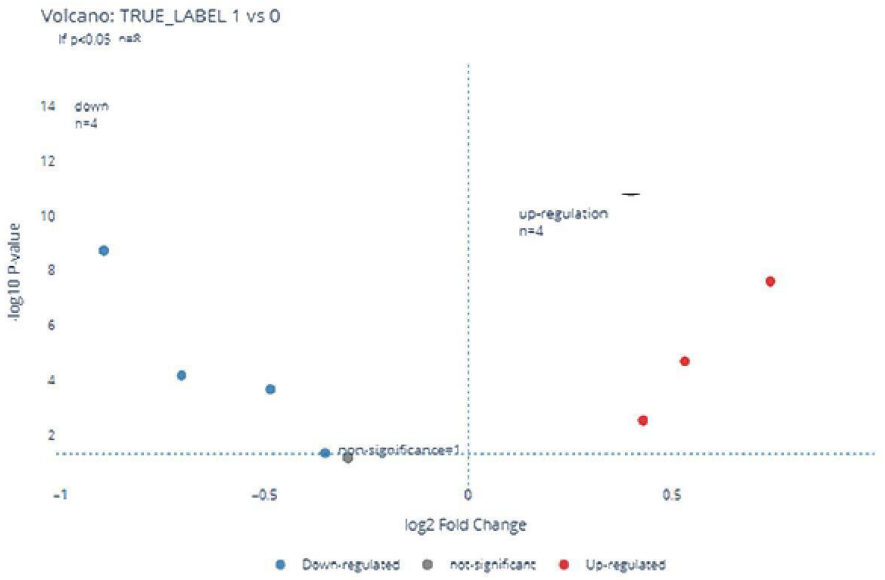
XSL•FO
RenderX

Test on STEM and ARTS for Men and Women via the Application

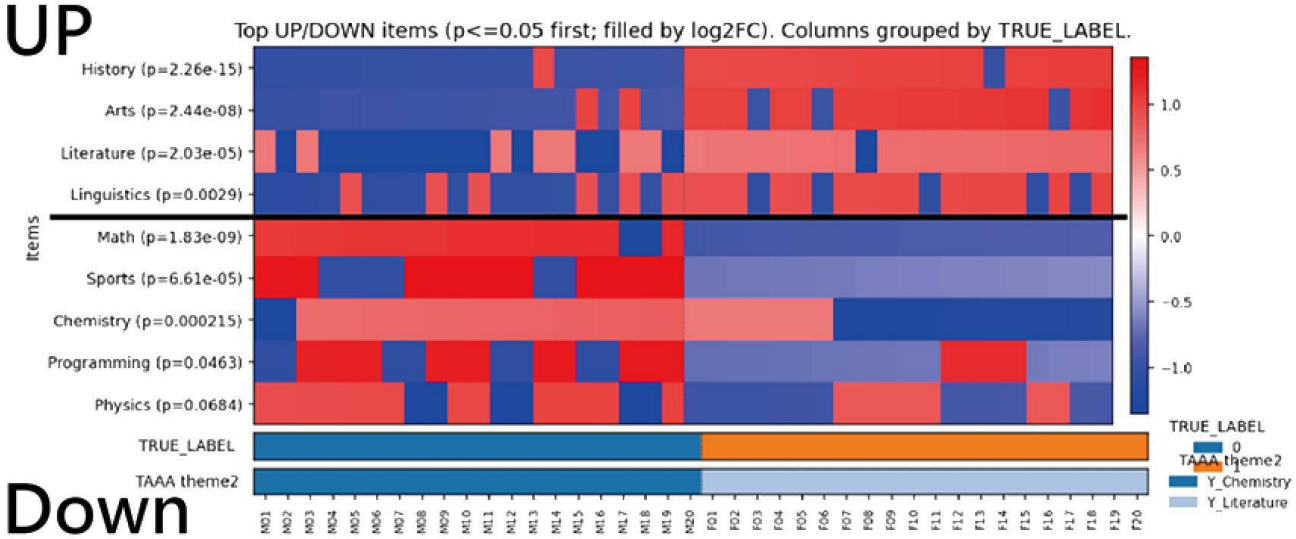
If the test demonstrated significant discriminatory power in identifying examinees' attributes [36], as assessed using TAAA against the external reference label (TRUE_LABEL; eg, men vs women), a corresponding volcano plot and heatmap can then be presented, as shown in Figure 7, with perfect agreement between TAAA-derived themes and TRUE_LABEL (MP Cohen κ=1.00).

Figure 7. Theme Assignment Algorithm in Articles (TAAA) applied to a test for examinees on STEM and ARTS against gender, shown via a (A) volcano plot, (B) heatmap (κ=1.0), and (C) paired *t* tests within STEM and ARTS.

(A)



(B)



STEM ARTS

(C)

Theme	n	mean_UP	mean_DOWN	*t* test	*P* value
STEM	20	48.71	69.45	-8.55	< .001
ARTS	20	75.91	47.817	17.30	< .001

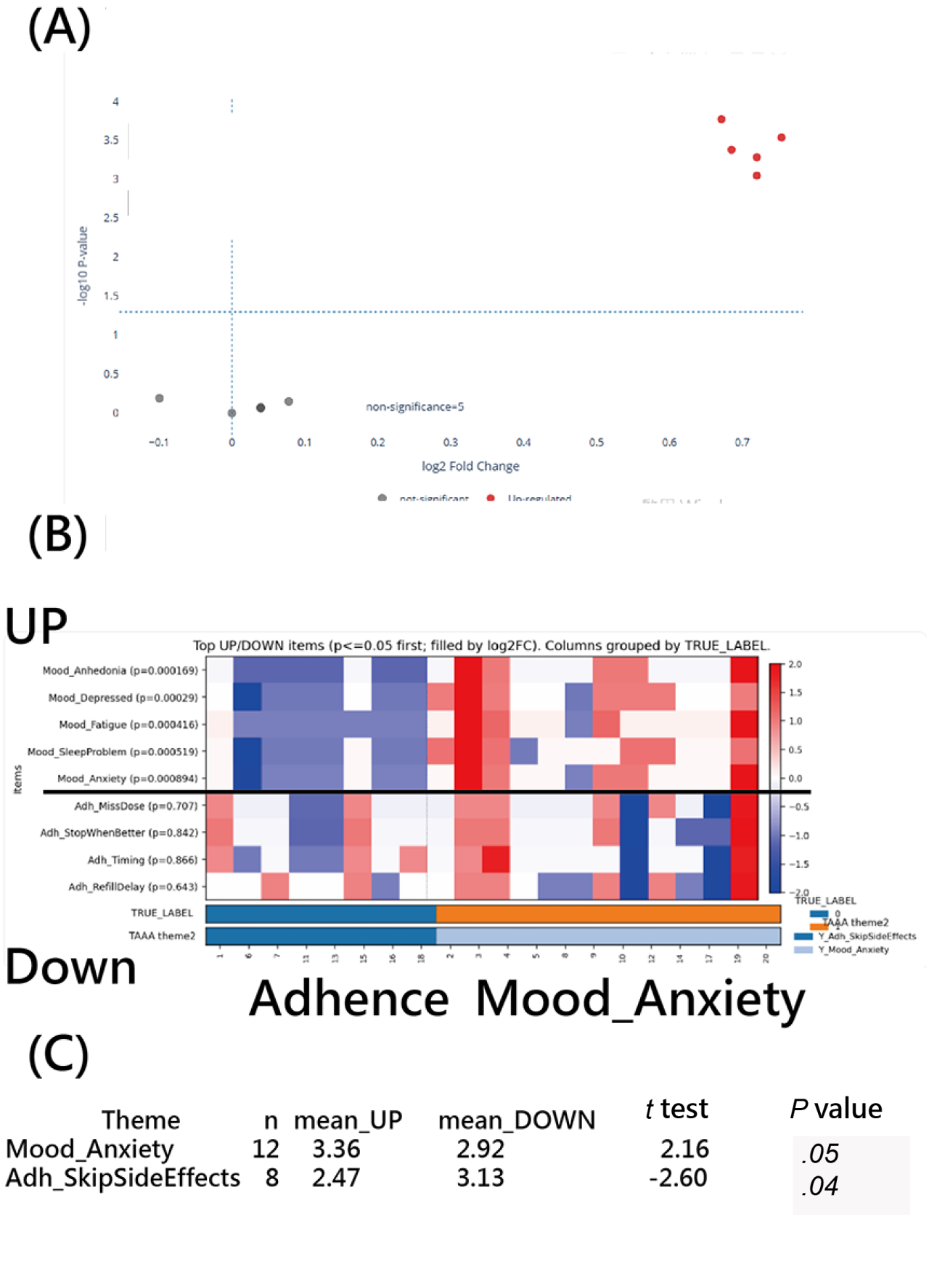
XSL•FO
RenderX

TAAA Applied to a Questionnaire on 2 Domains via the Application

If the questionnaire demonstrates significant discriminatory power in identifying patient features, as evaluated using TAAA against an external reference label (TRUE_LABEL; eg, depression/anxiety status and medication adherence) [34], the corresponding scree plot and network chart can be presented, as shown in Figure 8, indicating substantial agreement between TAAA-derived themes and TRUE_LABEL (MP: Cohen κ=1.0).

XSL•FO
RenderX

Figure 8. 2D scales of the questionnaire on differentially expressed genes of depression/anxiety and medication adherence using Theme Assignment Algorithm for Articles (TAAA): (A) volcano plot, (B) heatmap (κ=1.0), and (C) paired *t* tests within themes.



Theme	n	mean_UP	mean_DOWN	*t* test	*P* value
Mood_Anxiety	12	3.36	2.92	2.16	.05
Adh_SkipSideEffects	8	2.47	3.13	-2.60	.04

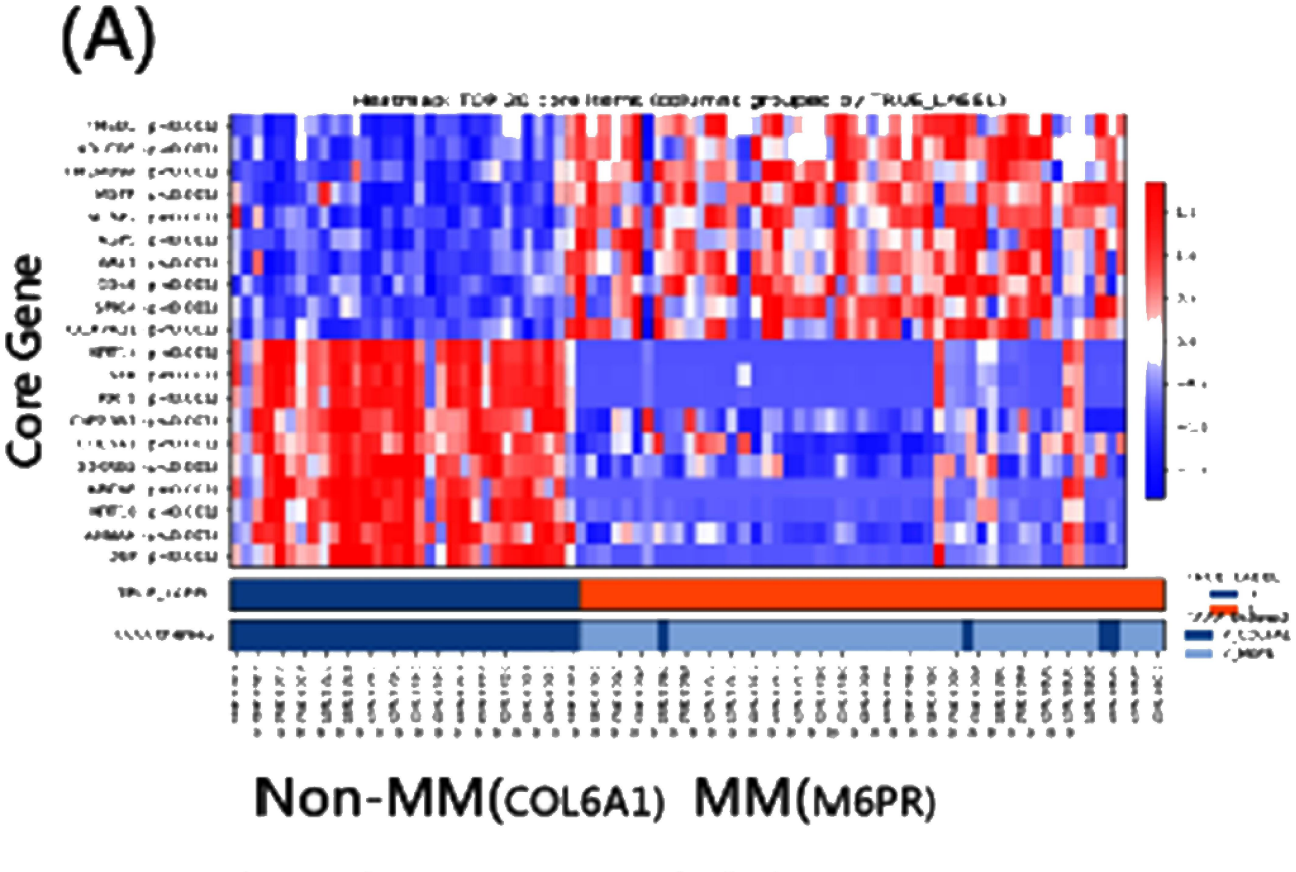
XSL•FO
RenderX

TAAA Applied to a Gene Expression Omnibus Data via the Application

Figure 9 demonstrates that applying TAAA to the GEO dataset GSE8401 robustly identifies metastasis-related gene patterns distinguishing MM from primary melanoma. The heatmap (MP: κ=0.90) shows a highly consistent separation of samples, with clear up- and downregulation blocks aligned with clinical status, indicating strong agreement between TAAA-derived themes and true labels.

Paired *t* tests within themes further confirmed the biological relevance: The MM and non-MM were statistically significant between gene testlets (*P*<.001 and *P*=.02).

Figure 9. The selection of significant genes for metastatic melanoma (MM) using differentially expressed genes with mapping power (MP)=0.90 based on a smaller *P* value: (A) heatmap (κ=0.90) and (B) paired *t* tests within True_Label.



(B)

True_Label	n	*F* test	*P* value	eta2	Adjused *P* value
0	31	229.58	<.001	0.88	<.001
1	52	5.88	.02	0.10	.02

Together, these results validate that, under a stringent MP threshold of 0.90, TAAA effectively selects significant differentially expressed genes and yields interpretable, metastasis-specific gene signatures rather than noisy individual genes—doing what a good method should: turning complexity into clarity with statistical poise.

Hypothesis Testing Results

The findings of this study confirmed that that the TAAA can reveal meaningful thematic alignments at the journal, country, and institute levels, supporting deeper insights into scholarly trends and research spots, as shown in Table 3.

XSL•FO
RenderX

Table . Summary of R scripts used in this study with the Theme Assignment Algorithm for Articles (TAAA) model.

Step	Visualization method	R script	Figure	Goal
1	Algorithm for keywords from abstracts	[16,24]	__[a]	Classification
2	Bar plot for core themes and TAAA	[17,25]	Figure 3	Assignment
3	Venn diagram	[22]	Figure 4	Screening
4	Kano and Sankey diagrams	[20,34]	Figures 6 and 7	Leadership
5	Sankey and network	[25]	Figure 5	Network
6	Volcano and heatmap	[34]	Figures 6–7,8,9	Mapping power

[a]Not applicable.

Table 4 summarizes key conceptual and methodological differences between TAAA and traditional co-word or cluster-based approaches, underscoring TAAA's capacity for article-level theme identification and reproducible bibliometric analysis.

Table . Comparison of Theme Assignment Algorithm for Articles (TAAA) with traditional co-word and theme-assignment methods.

Dimension	Traditional co-word/cluster-based methods	TAAA	Advancement/contribution
Analytical level	Operates at the keyword or cluster level; assigns themes to groups of co-occurring terms	Operates at the article level; assigns a single dominant theme to each article	Enables finer granularity and direct linkage between articles and their thematic focus
Assignment logic	Relies on cluster membership or keyword co-occurrence frequency without article-level aggregation	Uses the statistical mode of clustered keyword assignments within each article to determine its primary theme	Provides a reproducible and deterministic rule for article-level theme labeling
Interpretive output	Generates network maps or clusters without direct theme attribution to individual records	Produces a theme variable for every article, allowing downstream analyses (eg, country, institution, or temporal trends)	Enhances interpretability and enables quantitative comparisons across entities
Reproducibility and transparency	Results vary depending on clustering resolution, software, and visualization parameters.	Fully scripted in R, openly reproducible, with transparent equations and pseudocode (Equation 1)	Improves computational transparency and methodological reproducibility
Scalability and integration	Limited integration with other analytical tools; often visual only	Integrates seamlessly with visualization techniques such as Kano and Sankey diagrams for multidimensional mapping	Supports scalable, automated workflows for journal-, institution-, and country-level analyses

Discussion

Principal Findings

The results indicate that the 3 analytical scenarios identified 9, 7, and 9 core themes from their respective keyword sources, with HEALTH (39.4%), OLDER ADULTS (43.1%), and DEMENTIA (25.3%) emerging as the most prevalent themes. Among these, only DEMENTIA was consistently highlighted across all analyses. The application of dual heatmaps demonstrated the algorithm's capacity for comparative bibliometric mapping in JMIR Aging, with a mapping precision of 0.33 providing quantitative evidence of a 2-stage publication pattern.

These findings support the hypothesis that TAAA effectively assigns a dominant theme to each article using a mode-based approach that aligns with established thematic structures.

- Methodological innovation: Introduces the TAAA, a replicable R-based framework for precise article-level bibliometric mapping
- Multisource validation: Assigns primary themes to 434 JMIR Aging articles using mode-based clustering across 3 distinct keyword sources
- Thematic consistency: Identifies HEALTH, OLDER ADULTS, and DEMENTIA as the dominant research pillars, with dementia showing significant cross-source stability
- Advanced visualization: Uses Kano, Sankey, heatmap, and volcano plots to demonstrate the utility of treating bibliometric metadata with the same rigor as biological "differentially expressed" data

Adding Knowledge to Literature

This study advances bibliometric methodology by introducing the TAAA [17], which systematically assigns dominant themes at the article level using mode-based clustering, improving on

XSL•FO
RenderX

previous techniques [4,5] with R-based implementation [16,17]. Unlike traditional keyword co-occurrence analyses that remain at the cluster level (eg, a study [10] exploring trends and hotspots in research on exercise and Alzheimer's disease), TAAA enables article-level thematic assignment by adding a dedicated theme column to the dataset. This supports more granular analyses of journal focus, research trends, and associations between article entities.

The application and R-based TAAA framework [17,32-34] enhances reproducibility and computational efficiency while enabling seamless integration with visualization tools such as Kano and Sankey diagrams [11,14,15] for clearer thematic mapping with dual heatmaps. Applying TAAA [17,32-34] to JMIR Aging publications demonstrates its practical utility, consistently identifying core themes (eg, DEMENTIA) across multiple keyword sources. Notably, this approach can be applied to other journals to highlight their core themes for readers in future research.

TAAA thus offers a replicable, scalable method for bibliometric studies, supporting country-, institution-, and journal-level analyses that were previously limited by manual theme assignment.

Editorial Application of Abstract-Derived Keywords

In addition to analytical use, the extraction of keywords directly from article abstracts provides a potential editorial application. By comparing abstract-derived keywords with author-supplied keywords and Keywords Plus terms, editors can evaluate the thematic coherence of submitted or published manuscripts. Discrepancies between the automatically derived and author-assigned keywords may indicate whether an article aligns with the journal's primary thematic domains or falls outside its intended scope. Such comparison offers a data-driven mechanism for theme validation, assisting editorial boards with maintaining topical consistency, refining journal aims, and identifying emerging research areas within the publication's portfolio.

Cross-Level Interpretation of the DEMENTIA Theme

The greater prominence of DEMENTIA observed at the institutional level compared with the country level reflects the localized concentration of specialized research centers and collaborative networks focusing on dementia and cognitive aging. Institutions such as the University of Toronto and Harvard Medical School host long-standing dementia research programs that generate high keyword density and thematic coherence within institutional clusters.

At the national level, however, the thematic landscape becomes more diffuse as research activities span multiple subfields of aging and health technology. This pattern suggests that DEMENTIA functions as a core institutional specialization within certain centers of excellence rather than a uniform national research focus, highlighting TAAA's capacity to reveal hierarchical differences in thematic intensity across analytical levels.

Most Cited Article in JMIR Aging

Although Hannah Liane Christie from the Netherlands published the most articles, with 6 publications [24-29] in JMIR Aging, the most highly cited article [46] was authored by Audrey Lebrasseur and colleagues from Canada, with 252 citations. This study with a rapid review synthesizes research on COVID-19's impact on older adults, revealing increased psychological symptoms, ageism, and reduced social interactions. The findings highlight the need for tailored strategies to maintain social ties and address at-risk older adults' specific needs.

Strengths and Implications of This Study

The scalable framework of a reproducible application or R-based TAAA maps co-word clusters to one dominant theme per article using a deterministic mode rule.

Validated mapping via dual heatmaps and MP quantifies alignment with TRUE_LABEL, supporting a stage-like pattern while showing the boundaries are not strictly year-split.

Practical visuals integrating Kano, Sankey/SankeyMATIC, volcano, network, and receiver under the operating curve outputs provide interpretable, actionable bibliometric insights.

Potential Extensions to Other Domains

Beyond bibliometrics, TAAA can be adapted to clinical and other data-intensive settings where subject-level labels must be inferred from multi-item structures rather than directly observed. In clinical research, bifactor or multidomain measures are often reduced to composite scores and evaluated using logistic regression analysis to predict the patient status. TAAA offers a complementary structure-based alternative: It converts item-level patterns into a discrete subject style by assigning the dominant cluster (mode) across a subject's activated features, yielding an interpretable 2-style classification suitable for downstream prediction and stratification.

For example, TAAA can be applied to medication records to derive a patient's dominant treatment pattern (eg, drug class or regimen cluster), supporting analyses of treatment pathways and utilization without requiring explicit brand-company linkage. More broadly, the same framework can harmonize heterogeneous administrative datasets by inferring prevailing categories from fragmented records. These extensions can be implemented using the web-based application [32] provided in this study

Limitations and Suggestions

First, the analysis relied primarily on Keywords Plus from the WoSCC, which may omit domain-specific concepts or vary in granularity compared with controlled vocabularies such as MeSH used in PubMed. Future studies could integrate multiple bibliographic databases and standardized ontologies to improve thematic coverage and semantic precision.

Second, the mode-based theme assignment in TAAA simplifies articles to a single dominant theme, which may underrepresent genuinely multitopic publications. Extensions incorporating weighted frequencies or probabilistic theme assignments could

XSL•FO
RenderX

better capture thematic complexity and are available for further development, as outlined in our GitHub repository.

Third, although abstract- and title-based keyword extraction was automated using R [16] and the application [32], certain preprocessing steps (eg, keyword normalization and cluster interpretation) may still require expert oversight. Greater automation of preprocessing and quality control would further enhance scalability and reproducibility.

Fourth, interrater validation uses two study authors, acknowledges potential expectation bias, and describes blinding procedures more explicitly. Future studies could consider adding more experts to assign themes to articles in comparison withd TAAA.

Finally, validation was conducted using a single journal (JMIR Aging), limiting generalizability. Future work will apply TAAA across multiple journals and domains to more fully assess its robustness, scalability, and cross-disciplinary applicability.

Conclusions

This study introduces the application or R-based TAAA as a systematic approach for assigning dominant themes at the article level. Application to JMIR Aging publications demonstrates that TAAA enhances interpretability and reproducibility while enabling nuanced bibliometric analyses across journals, countries, and institutions. By bridging cluster-level co-word structures with article-level thematic representation, TAAA provides a transparent framework for large-scale thematic mapping. Future work may extend this approach by integrating semantic embeddings or applying the framework to multidisciplinary datasets.

Acknowledgments

We thank Coding Service for the English language review of this manuscript.

Funding

There are no sources of funding to be declared.

Data Availability

All data used in this study are available in Multimedia Appendices 1 and 2.

Authors' Contributions

SYCH initiated the research, collected the data, conducted the analysis, and wrote the manuscript. SYCH contributed to the design of the study and provided critical reviews of the manuscript, and SYCH and KTT contributed to the interpretation of the results. KTT monitored the research in process.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Dataset used in this study.
[XLSX File, 883 KB - aging_v9i1e79906_app1.xlsx]

Multimedia Appendix 2
How to conduct this study.
[PDF File, 1663 KB - aging_v9i1e79906_app2.pdf]

References

1. JMIR Aging. JMIR Publications. URL: https://aging.jmir.org [accessed 2025-06-29]
2. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
3. Aria M, Cuccurullo C. bibliometrix: an R-tool for comprehensive science mapping analysis. J Informetr 2017 Nov;11(4):959-975. [doi: 10.1016/j.joi.2017.08.007]
4. Chiang HY, Lee HF, Hung YH, Chien TW. Classification and citation analysis of the 100 top-cited articles on nurse resilience using chord diagrams: a bibliometric analysis. Medicine (Baltimore) 2023 Mar 17;102(11):e33191. [doi: 10.1097/MD.0000000000033191] [Medline: 36930064]
5. Huang YP, Pao JL, Chien TW, Lin JCJ, Chou PH. Thematic analysis of articles on artificial intelligence with spine trauma, vertebral metastasis, and osteoporosis using chord diagrams: a systematic review and meta-analysis. Medicine (Baltimore) 2022 Dec 30;101(52):e32369. [doi: 10.1097/MD.0000000000032369] [Medline: 36596060]

XSL•FO
RenderX

6. Havemann F, Gläser J, Heinz M, Struck A. Identifying overlapping and hierarchical thematic structures in networks of scholarly papers: a comparison of three approaches. PLoS ONE 2012;7(3):e33255. [doi: 10.1371/journal.pone.0033255] [Medline: 22479376]
7. de Lorenzo-Cáceres Ascanio A, Oter Puime AO. Publications on primary care evaluation in Spain after twenty years of reform (1984-2004). Thematic and bibliometric analysis. Rev Esp Salud Publica 2007;81(2):131-145. [doi: 10.1590/s1135-57272007000200004] [Medline: 17639681]
8. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010 Aug;84(2):523-538. [doi: 10.1007/s11192-009-0146-3] [Medline: 20585380]
9. Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. J Am Soc Inf Sci 2006 Feb;57(3):359-377 [FREE Full text] [doi: 10.1002/asi.20317]
10. Yang X, Li K, Zhang Y, Sun R, Yu J. Trends and hotspots in research on exercise and Alzheimer's disease: a decade of bibliometric review on prevention and molecular mechanisms. Psychiatry Clin Psychopharmacol 2025 Jun;35(2):185-197. [doi: 10.5152/pcp.2025.241016] [Medline: 40503464]
11. Lim SW, Chou W, Chen L. SankeyNetwork: a clear and concise visualization tool for bibliometric data. MethodsX 2025 Jun;14:103379. [doi: 10.1016/j.mex.2025.103379] [Medline: 40547885]
12. Cheng TY, Ho SYC, Chien TW, Chow JC, Chou W. A comprehensive approach for clustering analysis using follower-leading clustering algorithm (FLCA): bibliometric analysis. Medicine (Baltimore) 2023 Oct 20;102(42):e35156. [doi: 10.1097/MD.0000000000035156] [Medline: 37861508]
13. R Core Team. R: A language and environment for statistical computing.: R Foundation for Statistical Computing; 2025. URL: https://www.R-project.org [accessed 2025-06-12]
14. Kano N, Seraku N, Takahashi F, Tsuji S. Attractive quality and must-be quality. Journal of The Japanese Society for Quality Control 1984;14(2):39-48. [doi: 10.20684/quality.14.2_147]
15. Hsu SY, Chien TW, Yeh YT, Chou W. Using the Kano model to associate the number of confirmed cases of COVID-19 in a population of 100,000 with case fatality rates: an observational study. Medicine (Baltimore) 2022 Sep 16;101(37):e30648. [doi: 10.1097/MD.0000000000030648] [Medline: 36123944]
16. R Platform. Keyword extraction in R. URL: https://raschonline.com/raschonline/cbp.asp?cbp=NLPkeywordextractioninR [accessed 2025-06-14]
17. R Platform. Theme Assignment Algorithm for Articles (TAAA) in R. URL: https://raschonline.com/raschonline/cbp.asp?cbp=Top10keywordsfrequencytheme [accessed 2025-06-14]
18. R Platform. Kano and Sankey diagrams in R. URL: https://raschonline.com/raschonline/cbpgeo.asp?cbp=FLCAalgorithmexample5 [accessed 2025-06-14]
19. Chien TW. Cronbach's alpha with the dimension coefficient to jointly assess a scale's quality. Rasch Meas Trans 2012;26(3):1379 [FREE Full text]
20. Yang DH, Chien TW, Yeh YT, Yang TY, Chou W, Lin JK. Using the absolute advantage coefficient (AAC) to measure the strength of damage hit by COVID-19 in India on a growth-share matrix. Eur J Med Res 2021 Jun 24;26(1):61. [doi: 10.1186/s40001-021-00528-4] [Medline: 34167582]
21. R Platform. Venn diagram plotted in R. URL: https://raschonline.com/raschonline/cbp.asp?cbp=Venndiagramthen [accessed 2025-06-14]
22. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960 Apr;20(1):37-46. [doi: 10.1177/001316446002000104]
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [doi: 10.2307/2529310] [Medline: 843571]
24. Efron B. Bootstrap methods: another look at the jackknife. Ann Statist 1979;7(1):1-26. [doi: 10.1214/aos/1176344552]
25. Efron B, Tibshirani RJ. An Introduction to the Bootstrap: Chapman & Hall; 1993. [doi: 10.1007/978-1-4899-4541-9]
26. R Platform. Bootstrapping methods used in this study. URL: https://raschonline.com/raschonline/cbpchat.asp?cbp=Bootstrap_theme_stability [accessed 2026-03-21]
27. Lin J. Divergence measures based on the Shannon entropy. IEEE Trans Inform Theory 1991;37(1):145-151. [doi: 10.1109/18.61115]
28. Endres DM, Schindelin JE. A new metric for probability distributions. IEEE Trans Inform Theory 2003 Jul;49(7):1858-1860. [doi: 10.1109/TIT.2003.813506]
29. Xu QS, Liang YZ. Monte Carlo cross validation. Chemometr Intell Lab Syst 2001 Apr;56(1):1-11. [doi: 10.1016/S0169-7439(00)00122-2]
30. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Statist Surv 2010;4(none):40-79. [doi: 10.1214/09-SS054]
31. R Platform. Out-of-likelihood (OOL) procedures used in this study. URL: https://raschonline.com/raschonline/cbpchat.asp?cbp=Loo_cluster_theme_stability [accessed 2025-12-12]
32. OneClick keyword / mesh / co-word semantic tool. OpenAlex. URL: https://openalex2-dot-taaacoword2512.df.r.appspot.com/ [accessed 2026-05-06]

XSL•FO
RenderX

33. Chien TW. TAAA without mapping power(MP, i.e., no True_Label in the last column). URL: https://taaa-dot-taaacoword2512.df.r.appspot.com [accessed 2025-12-31]
34. Chien TW. TAAA from Digital → Text with mapping power(MP=kappa) vs True_Label. URL: https://taaa2factor-dot-taaacoword2512.df.r.appspot.com [accessed 2025-12-31]
35. App Engine standard environment documentation. Google Cloud. URL: https://cloud.google.com/appengine/docs/standard [accessed 2025-12-27]
36. National Center for Biotechnology Information Gene Expression Omnibus. Gene signature for aggression of melanoma metastases (GSE8401). URL: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8401 [accessed 2007-10-02]
37. Carbon S, Douglass E, Good BM. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res 2021 Jan 8;49(D1):D325-D334. [doi: 10.1093/nar/gkaa1113]
38. Posit. URL: https://posit.co/ [accessed 2026-03-21]
39. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A 2005 Nov 15;102(46):16569-16572. [doi: 10.1073/pnas.0507655102] [Medline: 16275915]
40. Van der Roest HG, Christie HL, Franco-Martin MA, Dröes RM, de Vugt ME, Meiland F. Determinants of successful implementation of assistive technologies for dementia: exploratory survey. JMIR Aging 2024 Sep 13;7:e53640. [doi: 10.2196/53640] [Medline: 39269371]
41. Christie HL, Dam AEH, van Boxtel M, Köhler S, Verhey F, de Vugt ME. Lessons learned from an effectiveness evaluation of Inlife, a web-based social support intervention for caregivers of people with dementia: randomized controlled trial. JMIR Aging 2022 Dec 7;5(4):e38656. [doi: 10.2196/38656] [Medline: 36476485]
42. Christie HL, Boots LMM, Hermans I, et al. Business models of eHealth interventions to support informal caregivers of people with dementia in the Netherlands: analysis of case studies. JMIR Aging 2021 Jun 3;4(2):e24724. [doi: 10.2196/24724] [Medline: 34081009]
43. Christie HL, Boots LMM, Tange HJ, Verhey FRJ, de Vugt ME. Implementations of evidence-based eHealth interventions for caregivers of people with dementia in municipality contexts (Myinlife and Partner in Balance): evaluation study. JMIR Aging 2021 Feb 5;4(1):e21629. [doi: 10.2196/21629] [Medline: 33544085]
44. Christie HL, Boots LMM, Peetoom K, et al. Developing a plan for the sustainable implementation of an electronic health intervention (Partner in Balance). JMIR Aging 2020;3(1):e18624. [doi: 10.2196/18624] [Medline: 32584261]
45. Christie HL, Schichel MCP, Tange HJ, et al. Perspectives from municipality officials on the adoption and implementation of electronic health interventions. JMIR Aging 2020;3(1):e17255. [doi: 10.2196/17255] [Medline: 32401217]
46. Lebrasseur A, Fortin-Bédard N, Lettre J, et al. Impact of the COVID-19 pandemic on older adults: rapid review. JMIR Aging 2021 Apr 12;4(2):e26474. [doi: 10.2196/26474] [Medline: 33720839]

Abbreviations

DC: dimension coefficient
FLCA: following leader clustering algorithm
GAE: Google App Engine
GEO : Gene Expression Omnibus
JS: Jensen-Shannon
MM: metastatic melanoma
MP: mapping power
TAAA: Theme Assignment Algorithm in Articles
WoSCC: Web of Science Core Collection

Edited by J Wang; submitted 01.Jul.2025; peer-reviewed by ChienSmile, M Cetin; revised version received 31.Dec.2025; accepted 16.Feb.2026; published 28.May.2026.

Please cite as:
Ho SYC, Tsai KT
Assigning Article-Level Themes in Bibliometric Analysis: Mode-Based Mapping Approach Using JMIR Aging Publications
JMIR Aging 2026;9:e79906
URL: https://aging.jmir.org/2026/1/e79906
doi:10.2196/79906

© Sam Yu-Chieh Ho, Kang-Ting Tsai. Originally published in JMIR Aging (https://aging.jmir.org), 28.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Application-Based Interventions for Family Caregivers of Older Adults: Scoping Review

Xin Yao Lin[1], PhD; Francesca Falzarano[2], PhD; E-Shien Chang[1], PhD; Kristin Hon[1], BA; Ruoying Zhang[3], MA; Andy Hickner[4], MSI; Sara Czaja[1], PHD

[1]Division of Geriatrics and Palliative Medicine, Weill Cornell Medicine, 420 East 70th Street, Office LH-303, New York, NY, United States

[2]Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, United States

[3]Department of Counseling, Developmental and Educational Psychology, Lynch School of Education and Human Development, Boston College, Chestnut Hill, NY, United States

[4]Samuel J. Wood Library, Weill Cornell Medicine, New York, NY, United States

Corresponding Author:
Xin Yao Lin, PhD
Division of Geriatrics and Palliative Medicine, Weill Cornell Medicine, 420 East 70th Street, Office LH-303, New York, NY, United States

Abstract

Background: As the global population ages, the increasing number of individuals with chronic conditions places a growing burden on family caregivers. Behavioral interventions delivered via application-based platforms, including those on mobile phones, tablets, or the web, have emerged as a powerful tool for enhancing caregiver support.

Objective: This study aims to identify and describe application-based interventions for family caregivers of older adults with chronic conditions, focusing on their designs, features, and impact on caregiver outcomes.

Methods: This review is reported according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. A search of publications from 2007 through December 20, 2022, was conducted across PsycINFO (EBSCO), MEDLINE ALL (Ovid), Embase (Ovid), Web of Science Core Collection (A&HCI, BKCI-SSH, BKCI-S, CCR-EXPANDED, ESCI, IC, CPCI-SSH, CPCI-S, SCI-EXPANDED, and SSCI; Clarivate), ACM Guide to Computing Literature (ACM Digital Library), and Engineering Village (Elsevier) using relevant keywords.

Results: The database search identified 9482 studies, resulting in 290 full-text papers being assessed, and a final 49 studies were included. Forty-four unique applications were identified, demonstrating different designs and features. These studies primarily examined the impact of applications on caregiver well-being, burden, and mental health. While the majority of interventions were perceived as beneficial, several design and research limitations were noted.

Conclusions: While many interventions demonstrate positive effects on caregiver outcomes, there are significant gaps in research design, evaluation, and reporting. Future studies should prioritize greater integration of self-care practices, clarity of intervention components, rigorous adherence to behavioral intervention frameworks, and greater inclusivity in participant samples.

(JMIR Aging 2026;9:e76115) doi:10.2196/76115

KEYWORDS

aging; mHealth; web application; mobile app; behavioral interventions; mobile health

Introduction

With the aging global population, the increasing number of individuals with chronic health conditions is placing a growing demand on family caregivers for assistance [1,2]. Family caregivers, typically family members or close friends, provide essential care to older adults with chronic conditions, managing activities of daily living and instrumental activities of daily living, including medication management, transportation, and meal preparation [3]. However, caregiver burden often results in negative outcomes such as stress, worse mental health, and reduced well-being [4-6]. Caregivers also face higher risks of social isolation and loneliness, issues that were further exacerbated during the COVID-19 pandemic [7-9]. Despite their central role, caregivers frequently neglect their own health and well-being, underscoring the need for interventions that support both the caregiver and care recipient [10-12].

In this context, behavioral application-based interventions delivered via mobile phones, tablets, or the web have emerged as a powerful tool for enhancing caregiver support. Applications allow caregivers to access support, resources, and health management tools, regardless of the intensity or demands of their caregiving role [13-17]. This flexibility is particularly valuable for caregivers who struggle to seek support due to time

XSL•FO
RenderX

constraints or geographical isolation [18,19]. Tablet, internet, and smartphone use has become prevalent across all age groups, including older adults, many of whom rely on smartphones as their primary or only digital device [20]. The widespread adoption of smartphones and use of web and mobile applications offer unique opportunities to address caregivers' needs in ways that traditional, nonvirtual services cannot [14,21,22]. Behavioral application-based interventions for caregivers could also broaden opportunities to improve psychosocial outcomes, such as reducing stress, improving social connectedness, and promoting well-being [14,23-25].

Despite growing interest in application-based interventions, evidence on their effectiveness and usability remains limited. Many caregivers are older adults themselves, and the usability of these applications is critical, especially for those experiencing age-related changes that affect their interaction with technology [26-28]. As the population of older adults grows, the need for usable, accessible applications to support caregivers in managing their own well-being while caring for others becomes increasingly urgent.

Previous systematic reviews have explored the use of digital health technologies for caregivers, with findings indicating that technology positively enhances self-efficacy, strain, and mastery [24,29,30]. However, many of these reviews focused broadly on mobile interventions [30,31] or specific chronic conditions [14,30], such as dementia or cancer, without thoroughly examining the features (eg, education, stress management, and peer support) that make mobile apps particularly useful or effective for caregivers. Furthermore, past reviews often evaluated existing apps available commercially in app stores (eg, Google Play Store and iOS App Store) rather than in the context of a behavioral intervention study without specificity regarding how these apps influence caregivers' well-being [30,31]. The importance of investigating mechanisms is underscored by the National Institutes of Health (NIH) [32].

This scoping review aims to address these gaps by focusing on application-based interventions for family caregivers of older adults with chronic conditions [33]. It will build upon existing reviews that have explored the broader role of technology in caregiving but have not fully examined application-based behavioral interventions. By summarizing current research on application designs, features, and their impact on caregiver outcomes, the results of this review will inform future application development and research, ultimately aiming to improve the well-being of family caregivers in an aging society.

The primary objective of this scoping review is to identify and describe application-based interventions (web, mobile, and tablet) designed to support family caregivers of older adults with chronic conditions. Specifically, this review will examine the designs (eg, user-centered approach and theory-based) and features (eg, education, stress management, and peer support) of application-based interventions (mobile, web, or tablet) for family caregivers of older adults, as well as their impact on key caregiver outcomes (eg, well-being, social support, and self-efficacy).

Methods

Design

This scoping review is reported using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines and is guided by the Joanna Briggs Institute methodology [34], which uses the population, concept, and context framework to define the scope of the review, eligibility criteria, and search strategy. The population is family caregivers of individuals with the following chronic diseases identified by the Centers for Disease Control and Prevention (CDC) to be prevalent among older adults in the United States [33]: heart disease, cancer, chronic lung disease, stroke, Alzheimer disease or related dementias (ADRD; including mild cognitive impairment), diabetes, and chronic kidney disease. The concept focuses on application-based interventions, including mobile, web, and tablet applications designed to support caregivers. The context encompasses any care environment or setting in which caregiving takes place, without limitations on geographic region or health care system.

A medical librarian (AH) searched on December 20, 2022, to identify studies focusing on application-based interventions for family caregivers of older adults. The search was conducted in the following databases: PsycINFO (EBSCO), MEDLINE ALL (Ovid), Embase (Ovid), Web of Science Core Collection (Editions=A&HCI, BKCI-SSH, BKCI-S, CCR-EXPANDED, ESCI, IC, CPCI-SSH, CPCI-S, SCI-EXPANDED, SSCI) (Clarivate), ACM Guide to Computing Literature (ACM Digital Library), and Engineering Village (Elsevier). Publication dates ranged from 2007, the year the first iPhone was released, to the date the search was conducted (2022). The search strategy consisted of selecting appropriate keywords such as "family caregiver" and "mhealth," combined with a "double-NOT" strategy to exclude records that were pediatric-only. Search term examples include Medical Subject Headings such as "mobile applications" and "caregivers," and keywords such as "Android," "iPhone," "care partner," and "carer." The full search strategy used for all databases is included in Multimedia Appendix 1. The current review protocol was registered and published on the Open Science Framework [35].

Selection Criteria

Included papers consisted of (1) published and peer-reviewed intervention studies, feasibility studies, pilot studies, randomized clinical trials (RCTs), or clinical trials, with English full texts available; (2) application-based interventions (mobile, tablet, or web) for family caregivers of older adults with one or more of the chronic diseases listed above; (3) applications designed and developed for family caregivers; and (4) papers reporting outcomes (eg, usability and well-being).

Papers were excluded based on the following criteria: (1) the paper is a review, perspective, opinion, conference abstract, commentary, book, book chapter, workshop, presentation, protocol paper, thesis, dissertation, or paper without outcome measures; (2) text-messaging or telephone-based interventions; (3) papers that focused on non-application–based interventions (eg, patient portals, virtual reality [VR], smart home, and wearables); (4) interventions targeting populations other than

XSL•FO
RenderX

caregivers (eg, care recipients and health care professionals); (5) application interventions that used general web platforms accessible by the general population (eg, social media and Zoom [Zoom Video Communications]); (6) studies that used an application prototype (eg, mock-ups) rather than a functional application; (7) interventions consisting solely of internet forums or discussion boards; and (8) studies that did not specify care recipients.

Selection of Studies

All retrieved studies were screened using the Covidence (Veritas Health Innovation Ltd) systematic review software developed by Veritas Health Innovation Ltd. First, duplicate titles and abstracts were identified and removed automatically using Covidence (AH). Then, four researchers (XYL, FF, ESC, and SC) applied the selection criteria and independently screened titles and abstracts. Full texts were then independently screened by 5 researchers (XYL, FF, RZ, KH, and SC) to ensure they met the eligibility criteria. Each title, abstract, and full text was reviewed by 2 researchers, and disagreements were resolved by a third researcher and through consensus meetings. Regular team meetings were conducted throughout the screening process to clarify questions or uncertainties.

Data Extraction and Analysis

Upon completion of full-text screening, data from included papers were extracted into a predesigned Microsoft Excel spreadsheet. The extracted fields included (1) study information (title, author, DOI, year of publication, and country of origin); (2) research questions and aims; (3) study characteristics (qualitative, quantitative, or mixed methods); (4) participant characteristics (number, condition, inclusion and exclusion criteria, relationship to the care recipient, and demographic characteristics of the caregiver and the care recipient); (5) recruitment strategy (where, how, attrition, and target recruitment goal); (6) intervention (type, duration, participants per condition, setting, treatment fidelity, and frequency of nonapplication components); (7) training (how and who trained participants); (8) application characteristics (type, features, application use frequency, recommended and actual application use, and satisfaction); (9) primary and secondary outcome measures; and (10) quantitative and qualitative results.

Five authors (XYL, FF, ESC, KH, and RZ) independently extracted data from included papers using predefined and standardized templates. The authors extracted and discussed 3 studies together before beginning the data extraction process to ensure proper and standardized extraction. Regular meetings were conducted throughout the extraction process to clarify questions or uncertainties. Moreover, 10% of papers were randomly selected for duplicate extraction to ensure reliability.

After extraction was completed, descriptive quantitative analyses were conducted to summarize the frequency and distribution of application interventions across various caregiver populations, study designs, and outcomes. Thematic analyses were also conducted, in which an initial inductive approach was used to identify themes that were validated and revisited until finalized. All discrepancies were discussed among the authors. The full review process is illustrated in the flowchart in Multimedia Appendix 2.

Ethical Considerations

Ethical approval was not required for this study as it is a scoping review of existing literature and does not involve human participants, animals, or the collection of original primary data.

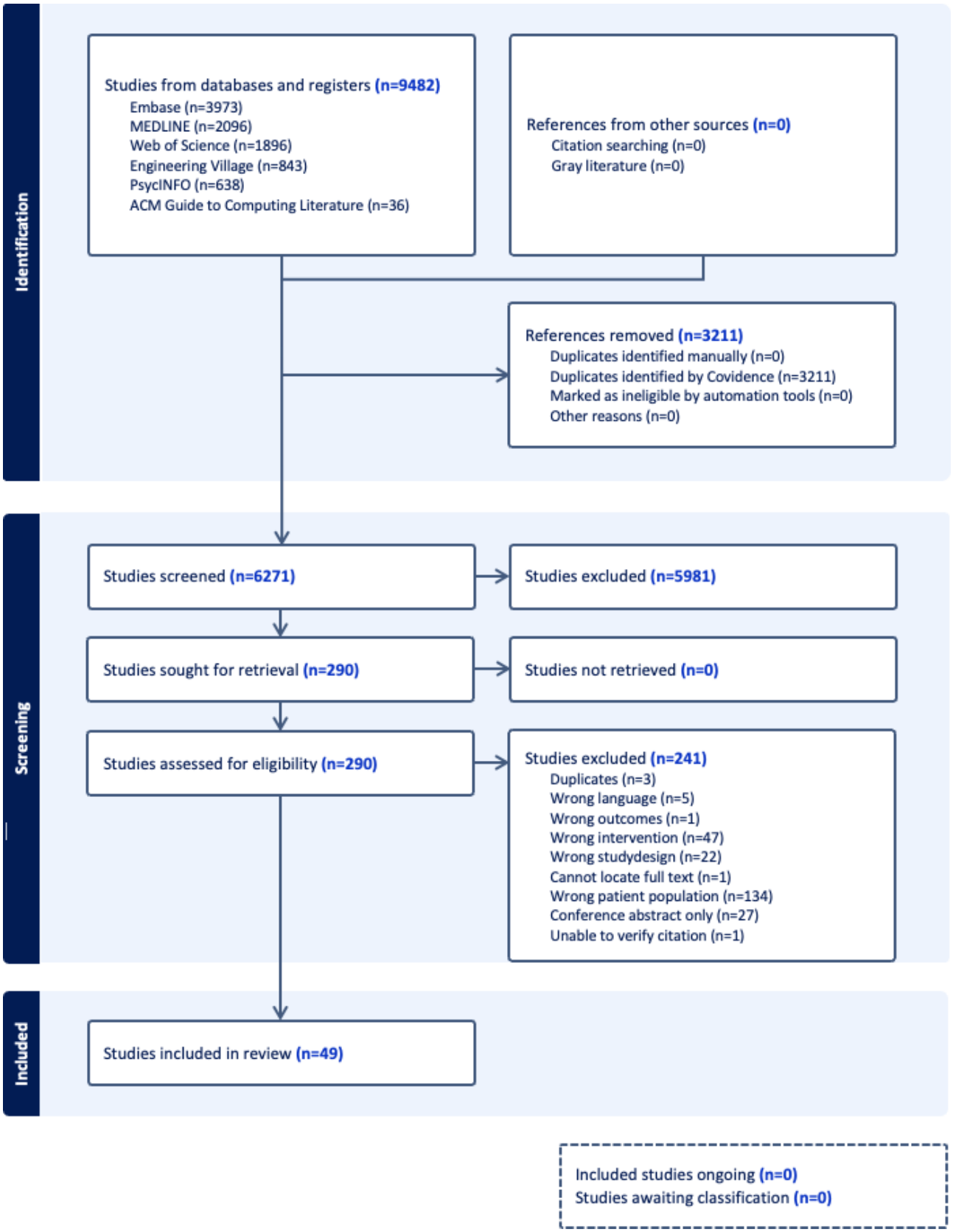
Results

Characteristics

Figure 1 shows the full PRISMA flow diagram outlining the study selection process. The database search yielded 9482 studies, in which 3211 duplicates were identified and removed, resulting in 6271 studies for screening. A total of 5981 studies were excluded during title and abstract screening, resulting in 290 full-text papers that were assessed for eligibility. Of those, 241 papers were excluded, which resulted in 49 studies included in this review (Figure 1). Studies were published between 2009 and 2022; 55% (27/49) studies were conducted in North America (United States and Canada), 39% (19/49) in Europe (Netherlands, Germany, Italy, Sweden, France, United Kingdom, Austria, Denmark, Poland, and Spain), 4% (2/49) in Asia (China and South Korea), and 2% (1/49) in Australia. The majority of studies (73%, 36/49) targeted family caregivers of people with ADRD, with 12% (6/49) and 8% (4/49) targeting cancer and stroke, respectively. A few studies (6%, 3/49) included family caregivers of older adults with various chronic conditions rather than a single condition (Table 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



XSL•FO
RenderX

Table . Paper and caregiver characteristics.

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Barbabella et al [36] (2016)	Italy, Sweden, and Germany	Quantitative	Baseline and post-3-month intervention	3 months	Home	Around 94 of 118 CG had access to web platform at least once.	NR[c]	Dropped out during intervention (n=5)	ADRD[d] or older adults (>60 years) with ADL[e] help.	Women: 61% and median age: 80 (IQR 74 - 85) years.	Spouse: 33%, children: 51%, and other: 16%.	Women: 68% and median age 58 (IQR 51-69) years.
Bartels et al [37] (2020)	Netherlands	Quantitative	RCT[f]	6-week intervention, 2-month follow-up, and 6-month follow-up.	Home	Experimental: ESM[g] self-monitoring and personalized feedback (n=17), pseudo-experimental: self-monitoring only (n=17), and control: regular care without ESM self-monitoring or feedback (n=16).	NR	Dropped out after the intervention (n=26)	Dementia	Reported CR dementia severity: very mild, 17%; mild, 37%; moderate, 33%; severe, 12%.	Spousal caregiver	Mean age 72 (SD 8.4) years; sex: 67% female; education: 51% low, 20% middle, 29% high; race and ethnicity: NR.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Blom et al [38] (2015)	Netherlands	Quantitative	RCT	5 - 6 months	Home	Intervention (n=151) and control (n=100)	NR	Seventy caregivers (28.6%) dropped out before the end of the intervention period: 59 in the experimental group (51 in the first 3 months) and 11 in the comparison group (all between 3 and 6 months).	Dementia	Most persons with dementia were female (60.4%), lived independently (82.9%), and had a mean age of 75.9 (SD 9.4; range 39–93) years. A formal dementia diagnosis was reported for 86.5% of CRs, with AD being the most common diagnosis (74.5%). Slightly more than half (52.4%) received the diagnosis within the previous 2 years. Mean IQ-CODE score was 58.9 (SD 5.75).	Almost all caregivers were spouses (58.4%) or children (in-law; 39.6%).	Predominantly female (69.4%) and living with the CR in the same household (60.4%). Mean age of CG: 61.2 (SD 12.37 years; range 26-87). Most CGs had children (78.8%; mean =2.54, (SD=1.10). All but 2 held Dutch nationality (99.2%). Education level varied from primary school to university, with 47.3% holding a bachelor's degree or higher.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Bodschwinna et al [39] (2022)	Germany	Quantitative	Feasibility	Within 2 months (average intervention duration was 7 weeks, and on average 9.79 days between 2 completed main sessions).	Home	Experimental (PartnerCARE; n=30): guided individually psycho-oncological online intervention delivered on the online platform minddistrict and control (n=30)	Since this was a feasibility study, a formal sample size calculation was not performed. A sample size of 60 participants (30 per group) was the predefined target, based on recommendations for sample sizes in feasibility studies.	The study dropout rate was low (T1: 17% and T2: 29%), and the high intervention completion rate (73.3%) indicates that once participants begin the intervention, they are likely to finish it. This treatment adherence is comparable to other online intervention studies with caregivers, but substantially higher than average intervention adherence in online intervention studies with patients.	Cancer		Supporting partner of patient with cancer. Relationship status: in partnership, 7 (11.9%), and married, 52 (88.1%). Mean relationship duration 18.35 (SD 11.32) years.	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
										Cancer diagnosis: breast cancer, 9 (15.3%); prostate cancer, 2 (3.4%); colon cancer, 6 (10.2%); lung cancer, 4 (6.8%); pancreatic cancer, 4 (6.8%); brain tumor, 6 (10.2%); hematological tumor, 18 (30.5%); other, 10 (16.9%). Time since diagnosis: ≤1 year, 24 (40.7%); 1-2 years, 14 (23.7%); >2 years, 21 (35.6%). Disease stage (UICC): I-II, 16 (27.1%); III-IV, 30 (50.8%); remission, 5 (8.5%); not known, 8 (13.6%). Treatment status: currently in treatment, 40 (67.8%); completed, 11 (18.6%);		Mean age: 47 (SD 9.62) years; range 28 - 71). Majority was female (69.5%), married (88.1%) and had children (69.5%) Mean age of 47.17 (SD 9.62). Gender: Male 18 (30.5%) and Female 41 (69.5%) Children: Yes 41 (69.5%) and No 18 (30.5%) Nationality: German 58 (98.3%) and Other 1 (1.7%) Level of Education: <10 years 3 (5.1%), 10 years 9 (15.3%), and >10 years 47 (79.7%) Employment Status: Employed (full time) 26 (44.1%), Employed (part time) 21 (35.6%), and Un-

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
										and pausing treatment, 8 (13.6%).		employed 12 (20.3%) Psychotherapy (current or last 8 wk): Yes 14 (23.7%) and No 45 (76.3%).
Bruinsma et al [40] (2021)	Netherlands	Mixed methods	Pre-post intervention	8 - 10 weeks of intervention but flexible	Home	Assessed (n=57), declined (n=17), version for spouses (n=30), and other family (n=25)	NR	NR	Young-onset dementia	Alzheimer dementia (AD; n=27) was the most prevalent cause, followed by frontotemporal dementia (FTD; n=9), vascular dementia (n=3), and Lewy body dementia (n=1).	Spouse: 45.8%, child: 33.3%, sibling: 12.5%, and siblings: 4.2%.	Men: 31% and age 18 - 74 years.
Bruinsma et al [41] (2021)	Netherlands	Mixed methods	Pre-post intervention	8 weeks	Home	Assessed for eligibility (n=33), excluded (n=6), allocated to intervention (n=27), lost to follow-up (n=7), and total analyzed (n=20)	NR	NR	Frontotemporal dementia	NR	Spouses	NR

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Carr et al [42] (2019)	United States	Mixed methods	RCT	NR	Home	Intervention group (n=36), treatment as usual group (n=36), final completion in postintervention assessment (n=26), and control group postintervention assessment (n=30)	NR	NR	Cancer	Fifty three percent of patients had lung disease.	Family caregivers, including 76% spousal caregivers.	Mean age=53.3; 73% female; 76% married; 76% spouse; 61% had college or higher education.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Caunca et al [43] (2020)	United States	Mixed methods	Phase 1: focus group; phase 2: semistructured interviews; and phase 3: usability study.	8 weeks	Home	Phase 3 usability study (n=9)	NR	NR	Stroke	NR	Family caregivers.	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												Phase 1: 7 stroke CG (n=3 Eng-speaking, n=4 Spa-speaking) participated in two focus groups Table 3. All female CGs ranged in age 49 - 72 (SD 60 years [9]). Of the 7 CGs, 4 identified as spouse, 5 reported > high school education, and 6 identified as Hispanic or Latino. Phase 2: 4 CGs participated in one-on-one, structured interviews to provide feedback (refined upon based on Phase I data). All 4 CGs were female with a mean age of 57 (SD 3) years and reported > high school educa-

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												tion. Of the 4 CGs, 2 were Eng-speaking, 3 were children of the stroke survivor, and 2 identified as Hispanic or Latino. Phase 3: 78% female; 78% Eng-speaking; 67% child CG; >78% high school education; mean age=46.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Chiu et al [44] (2009)	Canada	Mixed methods	Usability and intervention	6 months	Home	Experimental: ICSS[h] (n=28)	NR	NR	ADRD	Patient has ADRD.	Caring for a family member.	Typical participant was female, in her 40 - 50 s, worked full-time, college or higher education, not born in Canada, and immigrated to Canada 10 - 20 years previously. Majority of CRs were the parents who co-resided with the CG. CGs were in 2 groups (half provided care for <5 years and the other >10 y). Hours per week of providing care were <11 hours or >20 hrs.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Christie et al [45] (2022)	Netherlands	Quantitative	RCT	16 weeks	Home	Experimental, Inlife online social support platform for informal caregivers, people with dementia aimed at strengthening positive interactions, and social support (n=48) and control and waiting list (n=48)	Goal of 122 CGs	Two weeks after registration for the Inlife platform, participants were contacted by phone to reflect on their user experience to facilitate engagement, stave off attrition, and resolve any initial queries about the platform.	Dementia	Chart mean age=75 years (intervention); mean age=79.1 years (control); age range=47 - 91 years (intervention); age range=55 - 92 years (control); type of dementia: AD=25 intervention and 19 control, frontotemporal dementia=5 intervention and 2 control, vascular dementia=6 intervention and 9 control, dementia with Lewy bodies=1 intervention and 3 control, mixed dementia=2 intervention and 5 control, and dementia not otherwise specified=9 intervention and 10 control.	CG for patient with ADRD. Relationship with CR: spouse or partner, 21 (intervention) and 13 (control); son or daughter (-in-law), 24 (intervention) and 33 (control); brother or sister, 1 (intervention) and 1 (control); other, 2 (intervention) and 1 (control).	Chart mean age=58.1 years (intervention) and 55.7 years (control). Age range=26 - 84 years (intervention) and 22 - 82 years (control) Female=31, 64.6% (intervention) and 34, 70.8% (control). Mean years of education=13.1 (intervention) and 14.1 (control).

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Cristancho-Lacroix et al [46] (2015)	France	Mixed methods	Pilot RCT	3-month intervention, weekly sessions were 15 - 30 minutes on average, no time limit, and participants can access different website sections.	Home	Experimental, intervention of diapason (n=25) and control (n=25)	NR	Randomized (n=29); intervention: lost to 3 months (n=5) and lost to follow-up (n=3); control: lost to 3 months (n=4) and lost to follow-up (n=3); and total lost to follow-up (n=15)	AD	Chart mean onset of symptoms in years=4.62 (experimental) and 4.11 (control); mean MMSE[i]=18.5 (experimental) and 19 (control); and mean IADL[j] scale=0.6 (experimental) and 1.1 (control).	Adult-child caregivers: experimental group, 16 (64%); control group, 13 (54%).	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												Chart mean age=64.2 (experimental), 59 (control) Female=16 (experimental), 16 (control) Children of PWAD=16 (experimental), 13 (control) High level of education=19 (experimental), 18 (control) Middle level of education=6 (experimental), 3 (control) Living with the PWAD=12 (experimental), 10 (control) Visting PWAD daily=4 (experimental), 2 (control) Visiting PWAD at least once per week=9 (experimental), 9 (control) pscyhological or psychiatric treatment=3 (experi-

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												mental), 2 (control).
Dam et al [47] (2019)	Netherlands	Mixed methods	RCT	16 weeks	Home	Experimental, Inlife online social support platform for informal caregivers, people with dementia aimed at strengthening positive interactions, and social support (n=48) and control and waiting list (n=48)	Goal of 122 CGs and goal not met (n=96 recruited)	NR	Dementia	NR	Informal CG of person with dementia.	NR
Dam et al [48] (2017)	Netherlands	Quantitative	Feasibility and uncontrolled pilot study with repeated measures.	16 weeks	Home	Experimental group (n=25)	Goal: 20 participants (met)	NR	Dementia	CR diagnosis: AD 18 (75%), vascular dementia 1 (4.2%), FTD 1 (4.2%), Lewy body dementia 2 (8.3%), mixed dementia 2 (8.3%), and other 1 (4.2%). Mean years since diagnosis 2.1 (SD 1.6).	Spouse: 8 (32%), daughter: 10 (40%), and son: 6 (24%). No restrictions were placed on dementia type or CG relationship.	Table Males 13 (52%); Low-active group 54.4 (16.3%); High-active group 60.2 (7%); White 44 (86.2%) Education: High school 1 (4%), Lower vocational school 3 (12%), College 13 (52%), Graduate school 8 (32%), and NR 1 (2.0%).

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Davis et al [49] (2015)	United States	Quantitative	Feasibility	2 weeks	Home	Total sample (n=5)	NR	One withdrew	Dementia	NR	NR	All female, all non-Hispanic White, other info NR.
Duggleby et al [50] (2019)	Canada	Mixed methods	Pragmatic mixed methods RCT	6 months (data analysis was conducted for 3 months of data).	Home	Treatment (n=101) and control (n=98)	NR	1-month withdraw (n=9), 3-month withdraw (n=14), and 6-month withdraw (n=5).	AD and related dementia.	Female: 48% and age: 80.2 (SD 7.5) years.	Husband, wife, and life partner: 52%	Female: 80%; age: 63.5 years (SD: 12.0); education: 14.2 years (SD: 2.9); race: 96% White.
Duggleby et al [51] (2018)	Canada	Mixed methods	Pragmatic mixed methods RCT	6 months	Home	Treatment (n=101) and control (n=99)	NR	2-month withdraw (n=9), 3-month withdraw (n=14), and 6-month withdraw (n=5).	Dementia and multiple chronic conditions.	Female: 48% and age: 80.3 (SD 7.7) years.	Spouse or life partner: 49% and son or daughter: 46%.	Female: 81%; age: 63.6 years (SD: 11.6); education: 14.18 years (SD: 2.9); race: 93% Caucasian.
Duggleby et al [52] (2018)	Canada	Mixed methods	Feasibility	2 months	Home	n=37 (1 month: n=34 and 2 months: n=30)	Target (n=40, not met)	Baseline withdraw (n=3) and 1 month withdraw (n=4)	Dementia	Age: 84.7 (SD 7.4) years and female: 76%.	Son or daughter: 59.5% and spouse or life partner: 29.7%.	Mean age 63.2 (SD 11.7) years, 65% female, 82% Caucasian, education 15.5 (SD 3.5) years.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Ferre-Grau et al (2021) [53]	Spain	Quantitative	RCT	28 days	Online (can infer home) for the intervention group and the primary care center for the control group.	Intervention (n=71), control (n=60), analyzed intervention (n=56), and analyzed control (n=57)	Yes (target goal was 54/group)	Intervention group dropout (n=15) and control dropout (n=3); 56/71 of the sample in the intervention group finished the intervention program.	Multiple chronic conditions, Alzheimer, fragility, stroke, tetraplegia, neoplasia, Parkinson, and schizophrenia.	Mean age 83.27 (SD 9.66) years and female: n=73 (64.6); diagnosis of Alzheimer: n=24 (21.2%), Parkinson: n=1 (0.9%), stroke: n=7 (14.2%), other conditions included multiple chronic conditions: n=60 (53.1%), fragility: n=16 (6.2%), tetraplegic: n=2 (1.8%), neoplasia: n=2 (1.8%), and schizophrenia: n=1 (0.9%).	Parent: n=72 (63.7%), partner: n=25 (22.1%), other but family-related: n=10 (8.8%), other with no family relation: n=2 (1.8%), and nonprofessional CG contracted by family: n=4 (3.5%).	Total sample: mean age 60.65 (SD 12.37) years; n=104 (92%) female; n=108 Spanish nationality (95.6%); n=97 (85.8%) primary CGs.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Fossey et al [54] (2021)	United Kingdom	Quantitative	RCT	26 weeks	Online (can infer home)	cCBT[k] (n=53), cCBT+ telephone (n=101), and online psychoeducation (n=54)	NR	Randomized (n=638) and lost to follow-up across 3 arms (n=430); cCBT (n=213) allocated to group and discontinued or lost-to-follow-up (n=160); cCBT+ telephone (n=213) and discontinued or lost to follow up (n=112); and online psychoeducation allocated (n=212) and discontinued/lost-to-f/up (n=158).	Dementia	NR	Partner: n=278, child: n=235, and other: n=125.	*For total sample: mean age 59.9 (SD 12.2) years; female: n=545; race or ethnicity: NR.
Gaugler et al [55] (2015)	United States	Mixed methods	Feasibility	NR	Home (with face-to-face sessions)	Single group (n=30)	NA and NR	NR	Dementia	Mean age 80.75, (SD 9.52) years; female: 46.7%; White: 96.7%; and Alzheimer diagnosis: 53.3%.	Spouse: 60% and adult-child: 40%.	Mean age: 67.83 (SD 11.17) years; female=70%; and White=90%.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Goodridge et al [56] (2021)	Canada	Mixed methods	Participatory feasibility study	12 weeks	Virtual	Single group quantitative (n=29) and qualitative interview (n=21)	No (target sample size; n=40)	Enrolled (n=77), participants did not use the app (n=16), withdrew after several weeks (n=4); and final (complete) data were collected for 51% of participants (29/57).	Dementia	Mean age 78.9 (SD 10.1) years; CG rated health of PwD: fair (n=16, 55%).	Spouse: n=19 (36%), parents: n=17 (32%), other friends or relatives: n=17 (32%).	Mean age: 59.6 (SD 11.3) years, female: n=26 (90%), and race or ethnicity: NR.
Gustafson et al [57] (2019)	United States	Quantitative	RCT	6 months	Home	Intervention (n=14) and control (n=11)	Yes? At least before participant dropouts.	Randomized (n=31) and lost-to-follow-up or discontinued (n=5)	Dementia	Age: NR; gender: NR; race and ethnicity: NR; Clinical Dementia Rating Scale (CDR): intervention group (mean 1.63, SD 0.99) and control group (mean 15.40, SD 2.26).	Spouse or partner: n=28, adult child: n=1, and other relative: n=1.	Age (NR as continuous var): 55 - 64: n=6; 65 - 74; 65 - 74: n=16; 75+: n=9 Female: n=19; Race and ethnicity: 100% White or Caucasian.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Heynsbergh et al [58] (2019)	Australia	Mixed methods	Feasibility	30 days	Home	Single group (n=26)	NR	Lost-to-follow-up or discontinued (n=7)	Cancer	NR. Patients had diagnosis of colorectal cancer.	Spouse: n=19 (73%) and others (parent, adult-child, and other): n=7 (27%).	Mean age 57 (SD 12; range 30 - 79) years; gender: n=19 (73%) female; race and ethnicity: NR; education level: n=7 (27%) secondary; 19 (73%) tertiary.
Hughes et al [59] (2017)	United States	Qualitative	Feasibility	NR	NR	Single group (n=10)	NR	NR	Dementia	NR. Care recipients were patients diagnosed with dementia.	n=10	Mean age 60 (range 48 - 76) years; Gender: 70% female. Education: 100% HS graduates;

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Huisin Het Veld et al [60] (2020)	Netherlands	Mixed methods	3-arm RCT	12 weeks	Home	Major intervention arms (n=27), medium intervention arms (n=27), and minor intervention arms (n=27).	Goal: 24 participants per group were needed	NR	Dementia	NR	CG to relative with dementia and behavior changes.	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												Baseline: family CGs mean age 56.5 (SD 12.5) years (range 23 - 80 years), primarily female (71/81, 88%), and half of them had completed a professional or academic degree (40/81, 49%). Relatives with dementia they were caring for were mostly their mother or father (or a parent-in-law) (46/81, 57%) or their partner (32/81, 40%). Individuals with dementia had mean age of 75.1 (SD 9.9) years (range 49 - 96 years) and more often male (42/81, 52%), with AD being the most prevalent

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												form of dementia (47/81, 57%). In most cases, the first symptoms of dementia had appeared 4+ years previously (42/81, 52%). Behaviors that family CGs had the most difficulty dealing with were dependent (22/81, 27%) and masking behavior (19/81, 24%). At baseline, most family CGs stated that they were somewhat (35/81, 43%) or significantly (31/81, 38%) burdened by the care for their relative with dementia.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Jordan et al [61] (2022)	United States	Mixed methods	One-group pretest and posttest design	4 weeks	Medical center	Experimental group (n=72)	NR	NR	Stroke	NR	Caregives for someone with stroke diagnosis.	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												The majority of the CGs were women (n=69, 96%), White (n=49, 68%), married or living with partners (n=58, 81%), and college graduates or had some education after high school (n=55, 76%). Ages ranged from 34 - 84 years with a mean age of 62.63 (SD 10.51) years. CGs varied widely in the length of time they had been a stroke CG with a range from 2 months to 46.5 years, a mean of 6.05 (SD 8.39) years, and a median of 3 years. A majority (n=65, 90%) lived in the same house as the

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												stroke survivor. The most common CG problems reported on the problem checklist instrument were "feeling depressed or stressed a lot of the time" (n=46, 64%) and "relationship changes with the stroke survivor" (n=22, 31%). The least common CG problem was legal issues (n=9, 13%).

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Kajiyama et al [62] (2013)	United States	Quantitative	RCT	3 months	Home	Control and EOC[l] (n=75) and treatment and ICC[m] (n=75)	NR	NR	Dementia	The only significant difference between the 2 groups on any sociodemographic or baseline measures was that the mean age of the PWD in EOC was approximately 5 years greater than in ICC (*P*=.007).	CG to someone with dementia.	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												About 97% of the CGs had used the Internet in the past to search for information, close to half used it for prior online training of some kind, and over 80% watched online videos. Most CGs in the study (85%) were women in their late fifties with ages ranging from the midtwenties to the early eighties; their PWD tended to be about 20 years older. More than 80% of the CGs had at least some college, and about 35% had completed some graduate schoolwork. The majority (85% - 95%) were Caucasian;

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												there were also African Americans (2.9%), Asian Americans (4.8%), Hispanic American (2.9%), Native American (1.9%), and Hawaiian or Pacific Islanders (0.9%) enrolled. About two-thirds were caring for relatives with Alzheimer's disease, 10% were caring for patients with vascular dementia, and the remaining 23% were CGs for persons with another form of dementia (refer to Multimedia Appendix 3). Slightly more than half were caring for their spouse; 35% - 45%

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												were caring for a parent and the remainder (10% or so) were caring for another relative or a nonrelative (n=3;<1%). Mean number of hours per wk spent in caregiving activities was about 66 but the range was quite high 10 - 120 hours). Services provided ranged from transportation, shopping, and financial help to total personal care plus all other required services. On a 1 to 10 scale, both groups averaged approximately 7 discrete services being provided.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Kales et al [63] (2018)	United States	Quantitative	Two-site RCT	2 months	Home	Experimental and WeCareAdvisor (n=27) and control and waitlist (n=30)	NR	NR	Dementia	Regarding PLWD[n], their mean age was 80.4 (SD 10.2) years, most were female (63%), married (58%), and White (65%). The mean MMSE was 16.5 (SD 8.3), the mean number of NPI[o] behaviors was 7.3 (SD 3.1), and the mean level of functional dependence was 9.9 (SD 4.3).	Primary CG for a PLWD with clinical diagnosis or MMSE score <24 and residing with or close to the PLWD.	The mean age of CGs was 65.9 (SD 14) years and the majority (75%) were female. Most (84%) had more than a high school education and the majority (74%) were married and white (63%). Almost half (49%) were spouses. CGs had a mean of 4.1 (SD 3.1) medical and 0.4 (SD 0.6) mental health conditions.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Kovaleva et al [64] (2019)	United States	Qualitative	Feasibility and acceptability	7 weeks	Virtual	Single group (n=36)	NR	NR	Dementia	Diagnosed dementia	CG of a person with dementia.	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												CG's age 63.8 (SD 9.8 years; range 46-78). CG's gender (% female) 31 (86). PLWD gender (% male) 22 (61). Mean age of the PLWD 77.1 (SD 8.2; range 63-97). CG currently employed outside of home (%) 9 (25). CG's race (% African American) 7 (19). Education (% ≥high school) 36 (100). Duration of time CG has been personally providing care for the PLWD (years) mean 3.2 (SD 2.7; range 0.2 - 10). CGs who receive help in their caregiving responsibilities from family,

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												community, and/or paid help (%) 23 (64). CGs who are responsible for the provision of support (financial, emotional, logistical, legal, etc) to persons other than the PLWD (%) 14 (39). PLWD (% spouse) 21 (58), (% parent) 11 (31), (% other relative) 4 (11). Duration of time PLWD needed care due to dementia (years): mean 3.4 (SD 3.0; range 0.4 - 10).

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Leung et al [65] (2022)	China	Mixed methods	Pre- and post-quantitative surveys and postintervention semistructured interviews.	8 weeks	Virtual	Single group (n=28)	Goal (n=32)	20%	Dementia	Patients with dementia.	Family CG	The majority were female (71.4%), married (92.9%), and either retired or unemployed (78.6%). Mean age was 59.71 (SD 9.88) years. About one-third had obtained a Bachelor's degree or above (35.7%) and half had attained a secondary school education. Two-fifths (39.3%) had a monthly income of HK $12,001 (USD $1535.99) or more. Two-fifths of participants (39.2%) had been caring for their family member with dementia for 5 years or longer.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Lewis et al [66] (2010)	United States	Mixed methods	Feasibility and acceptability	NR	Virtual	Single group (n=47)	NR	Response rate: 74%	Dementia	Diagnosed dementia	Family caregiver	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												Mean age of the participants was 55 (SD ¼ 9) with a range from 32 to 87 years. The sample was predominantly Caucasian (85%), female (85%), and educated, with 65% college graduates. Participants came from 10 states with the majority coming from Colorado and Georgia, followed by North Carolina and Minnesota, with 19% living in rural areas. The participants were caregivers for their family members for an average of 3.8 years (SD ¼ 3.9 years) with a range from <1

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												to 21 years.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Linden et al [67] (2022)	United States	Qualitative	Feasibility	60 days	Home	Experimental (n=51)	NR	NR	Dementia	Females: 34 (66.7%) and mean age: 79.2 (SD 10.6) years; race and/or ethnicity: 2 Asian (3.9%), 1 Black or African American (2%), 2 Hispanic or Latinx (3.9%), 0 Native American or American Indian, 1 NR (2%), and 45 White (88.2%).	Family member or friend providing unpaid support for a person living with ADRD. Participants were allowed to define primary CG status. Child: 28 (54.9%), spouse or partner: 20 (39.2%), and other relative: 3 (5.9%).	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												Table where Females 38 (74.5%) and mean age is 60.3 (SD 9.8) years. Race and/or ethnicity: Asian 2 (3.9%), Black or African American 1 (2.0%), Hispanic or Latinx 2 (3.9%), Native American or American Indian 1 (2.0%), NR 1 (2.0%), White 44 (86.2%) Marital Status: Married or domestic partnership 37 (72.5%), Divorced 11 (21.6%), Single, never married 2 (3.9%), Widowed 1 (2.0%) Education: Postcollege 19 (37.2%), 4-year college 17 (33.3%), Technical school,

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												vocational training, community college 10 (19.6%), High school diploma or equivalent 5 (9.8%) Employment: Full-time 21 (41.2%), Retired 19 (37.3%), Part-time 7 (13.7%), Not working 4 (7.8%) Income: >$100 000 18 (35.3%), $40 000 - 60 000 8 (15.7%), Do not wish to answer 8 (15.7%), $80 000 - 100 000 6 (13.7%), $60 000 - 80 000 4 (11.8%), $20 000 - 40 000 2 (3.9%), <$20 000 1 (2.0%)

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Llaneza et al [68] (2022)	United States	Qualitative	Pre-post feasibility	15 - 30 minutes	Virtual	Single group (n=15)	Yes (goal: 12 - 15)	Selected: 25 and enrolled: 15.	Cognitive Impairment	Medical professional diagnosed cognitive impairment (n=67).	Unpaid support person	

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												Female 14 (93.3%) and Hours of care per week 75.5 (1.5 - 168) mo being a caregiver Age (years): 75.6 (6-180), 61.86 (34-80) Race: non-Hispanic White 11 (73.3%), African American 1 (6.6%), Multiracial 1 (6.6%), Hispanic 1 (6.6%), Asian or Pac. Islander 1 (6.6%) Education Status: High School graduate 1 (6.6%), Some college 5 (33.3%), College degree 5 (33.3%), Advanced degree 4 (26.3%) Living with person with cognitive impairment: Yes 9 (60.0%), No 6 (40.0%) Marital Status:

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
												Married 10 (66.6%), Single 4 (26.3%), Other 1 (6.6%) Relationship with person with cognitive impairment: Adult Child 8 (53.3%), Partner 5 (33.3%), Other 2 (13.3%)
Lundberg [69] (2014)	Sweden	Qualitative	Case study	2 years	Home	Single group (n=10)	NR	NR	Dementia and stroke.	Dementia: 9/10 and stroke: 1/10; age=74.5 years; and help with ADL=10/10.	Family caregiver	Age: 80 years; sex: 6/10 women.
Marziali and Garcia [70] (2011)	Canada	Mixed methods	Quasi-experimental	6 months	Home	Chat group (n=40) and video group (n=51)	NR	NR	Dementia	Sex: female (50%)	Spouses: 74% and adult-children (majority daughters): 26%.	Mean age: 65.51 years; female: 72%; education: 60% college or university; mean annual income: $40,000; average caregiving time per day: 14.68 hours; caregiving duration: 4.55 years.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Meichsner et al [71] (2019)	Germany	Quantitative	RCT	8-week intervention and 5-month follow-up from baseline.	Home	IG[p] (n=19) or WCG[q] (n=18)	NR	16.22% (n=6)	Dementia	Age: 74.92 years, sex: male (54.31%), disease: Alzheimer (54.1%), moderately severe (n=13, 35.1%) or severe (n=12, 32.4%) stage of dementia.	Seventy three percent of spousal caregivers, remaining caring for their parents.	Age: 62.11 years; sex: 78.4% female; caregiving duration: 4.5 years; residence: 81.1% living with CR; education: 46% tertiary.
Metcalfe et al [72] (2019)	England, France, and Germany	Mixed methods	Unblinded randomized wait‐list controlled trial	Week 6 midpoint and week 12 final evaluation	Home	Group A (immediate access; n=30) and group B (waitlist; n=31).	A sample size of 60 participants was considered adequate for a pilot study (30/group and 20 in each country).	Fifty-eight participants completed the protocol, with 1 at each intervention site dropping out during the study.	Young-onset dementia	Age: 61 years, sex: female (43.3%-54.8%), AD diagnosis: 54.7%-64.5%, and years since diagnosis: 3‐4 years (41.9%‐43.3%).	NR	Age: 57 years; caregiving duration: 1.8-2.6 years; sex: 60% female; education: 41.9%-47.7% higher education; employment: 50%-67.7% full- or part-time; retired: 22.6%-23.3%.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Núñez-Naveira et al [73] (2016)	Denmark, Poland, and Spain	Quantitative	Pilot RCT	3 months	Home	Experimental (n=30) and control (n=31)	NR	Dropout during intervention (n=16)	Dementia	GDS[r] stage 5: 36%	NR	Age: 25-88 years; female: 63.9%; occupation: intellectual labor, 41%; caregiving hours/week <20, 57.4%; received support: 96.7%; type of support: day-care center and relatives or friends, 44.3%; feelings of institutionalization: sometimes, 37.7%; self-perceived health: fair, 47.5%; no visits to general practitioner, 57.4%.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Park et al [74] (2020)	South Korea	Quantitative	Nonequivalent control group pretest-posttest design	4 weeks	Home	Experimental (n=12) and control (n=12)	Yes (aimed for 20)	Total participant drop-off rate of 7.7%.	Dementia	NR	Child: n=17, spouse: n=2, and daughter-in-law: n=5.	Age: 54-51 years; female: 14; religion: 14 Buddhist or atheist; marital status: 20 married; education: 15 high school or below; housing type: 12 households; employment: 12 employed; income: 20 below 2 million won; health status: 12 good; number of family members living together: 3-4; caregiving duration: 24-39 months; caregiving hours per day: 13 (6 hours); education on dementia: 15 no; use of long-term care facilities: 18 day and night care.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Pensak et al [75] (2021)	United States	Quantitative	Pilot RCT	12 weeks	Home	Pep - Pal (n=26) or treatment as usual (TAU; n=30)	NR	Noncompliant (n=8) and lost to follow-up during intervention (n=8).	Cancer	Patients are receiving hemopoietic stem cell transplant (HSCT).	Majority of CGs (82%) identified as a spouse or partner.	Participants ranged from 22 to 83 years of age (mean age 54.3 years) and were predominantly women (n=42). Most had received at least a college education (n=33) and were married (n=46), were employed at least part-time, and were White (n=49). Median caregiving hours ranged from 24 to 31.
Pierce and Steiner [76] (2013)	United States	Mixed methods	Secondary analyses of data from the RCT	1 year	Home	Intervention (n=36)	NR	About 36 of total 51 web users completed the yearlong RCT.	Stroke	First-time patients with stroke.	Spouse of the CR: 66.7%, child: 19.5%, and other: 13.8%.	Typical web user was female (69.4%), aged around 54 years, with college education (41.7%), and was White (86.1%).

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Pierce et al [77] (2009)	United States	Quantitative	Randomized 2-group, repeated measures design	1 year	Home	Web user (n=36) and nonweb user (n=37)	Met	NR	Stroke	NR	Wife: 41.7%, husband: 25%, daughter: 16.7%, and other relatives and friends: 13.8%.	Female: 69.4%; White or non-Hispanic: 86.1%. Age group: 51–60 years, 30.6%; 61–70 years, 22.2%; 31–40 years, 16.7%; and 41–50 years, 16.7%. Education: college (>12–16 years), 41.7%; high school (>8–12 years), 27.8%; and graduate school (>16 years), 16.7%.
Pot et al [78] (2015)	Netherland	Quantitative	Feasibility	8 weeks	Home	Completers (n=68) and noncompleters (n=81)	NR	Completers (n=68) and noncompleters (n=81).	Dementia	Female: 61.1% and age: 76.4 (39-93) years.	NR	Female: 69.8%; age mean=61.5 (range 33-87); at least bachelor's degree: 49.7%.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Reblin et al [79] (2018)	United States	Quantitative	Two-group randomized design	6 weeks	Home	eSNAP[s] (n=30) and control (n=10)	NR	NR	Neurological cancer	Female: 47.5%, White: 90%, and mean age: 52.2 (SD 16.5) years.	Spouse: 64.1%, child: 17.9%, and parent: 12.8%.	Female: 75%; White: 90%; mean age 52.2 (SD 16.5) years. Education level: some graduate or professional school, 31.6%; high school graduate or equivalent, 23.7%; some college or vocational school, 23.7%; and 4-year college graduate, 18.4%.
Reblin et al [80] (2018)	United States	Quantitative	Two-group randomized design	7 weeks	Home	eSNAP (n=30) and control (n=11)	NR	Retention rate: 80%	Neurological cancer	Age: 51.3 (22-76) years with 42.1% female.	Spouse: 55.2%, child: 20.7%, and parent: 17.2%.	Mean age 56.7 years (range 29-80); female: 75.9%; White or Caucasian: 92.9%.
Rettinger et al [81] (2020)	Austria	Mixed methods	RCT (feasibility, qualitative results, and RCT results published elsewhere)	3 months	Home	App use (n=15) and control (n=35)	NR	NR	Dementia	Female (60%): 75 - 80 years (33%), 81 - 85 years (20%), and >90 years (20%).	Child: 6 (40%), grandchild: 5 (33%), and partner: 3 (20%).	Female: 13 (87%) and mean age 49 (range 17-71) years.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Rodriguez et al [82] (2021)	United States	Quantitative	Feasibility	30 days	Home	72 hours=60, 14 days=57, and 30 days=55	NR	NR	Dementia	Moderately severe cognitive decline (42%), moderate cognitive decline (18%), severe cognitive decline (16%), and mild cognitive decline (15%).	Spouse: 50% (28) and child: 30% (17).	Mean age 64.96 (SD 10.9) years; female: 78% (42). Race: White, 76% (42); Black, 9% (5); and Hispanic, 10% (6).
Sikder et al [83] (2019)	United States	Mixed methods	Feasibility	4 weeks	Home	Single group (n=21)	NR	Not using the app or lost at follow-up (n=4)	Dementia	NR	NR	Mean age 66.52 (SD 8.61) years; women: 71%; race: all White.

XSL•FO
RenderX

Study and year	Country of origin	Study characteristics	Study design	Study duration	Study setting	Sample size	Met target goal	Attrition rate	Caregiving for what disease	CR[a] demographics	Relationship with CR	CG[b] demographic
Zimmerman et al [84] (2016)	United States	Mixed methods	Feasibility	12 weeks	Home	Single group (n=28)	NR	NR	Dementia	Mean age: 82 (SD 9.17) years	Mother: 8 (28.5%), father: 1 (3%), mother-in-law: 2 (7.1%), spouse: 13 (46.4%), and other: 2 (7.1%).	Mean age 64 (SD 12.77) years; female: 71.4%. Education: some high school, 7.1%; high school graduate or GED, 17.8%; vocational training or some college, 21.4%; associate degree, 10.7%; 4-year degree, 14.3%; and graduate degree, 32.1%. Race: Black, 10.7%; White, 75%; Asian, 7.1%; Hispanic, 3%; and Hawaiian, 3%.

[a]CR: care recipient.

[b]CG: caregiver.

[c]NR: not reported.

[d]ADRD: Alzheimer disease and related dementias.

[e]ADL: activities of daily living.

[f]RCT: randomized clinical trial.

[g]ESM: experience sampling method.

[h]ICSS: internet-based caregiver support service.

[i]MMSE: Mini-Mental State Examination.

[j]IADL: instrumental activities of daily living.

[k]cCBT: online cognitive-behavioral therapy.

[l]EOC: education/information-only condition.

[m]ICC: iCare condition.

[n]PLWD: person living with dementia.

[o]NPI: neuropsychiatric inventory.

XSL•FO
RenderX

[p]IG: intervention group.

[q]WCG: wait - list control group.

[r]GDS: Global Deterioration Scale.

[s]eSNAP: electronic Social Network Assessment Program.

Final studies included 21 quantitative, 5 qualitative, and 23 mixed methods studies with various designs: 25 RCTs, 15 feasibility studies, 5 pre-post intervention studies, and 4 usability studies. Intervention durations ranged from 1 session lasting 15 - 30 minutes to 2 years. The majority of interventions were home-based; 5 were delivered virtually, 2 in clinical settings (eg, medical center), and 1 was not reported. A few studies involved interventionists, often clinical psychologists, nurses, or social workers, who received specialized training to deliver the interventions along with an app. About half of all studies (n=27) reported that participants were trained on how to use the app by research assistants or staff via in-person sessions, phone calls, or online methods. Only 20% (10/49) of studies reported fidelity measures, including regular check-ins and adherence assessments to ensure protocol compliance.

Study Participants

Common eligibility criteria across studies included primary caregivers who were adults (typically aged >18 years) caring for a person with dementia or another chronic condition (eg, cancer). Most studies required caregivers to have access to the internet and a device (computer, smartphone, or tablet; n=30). Language proficiency was also a common criterion, with caregivers needing to read and write in the language of the study (n=22), such as English, German, or Chinese. Many studies excluded caregivers with mental health conditions or cognitive impairment (n=11). Recruitment methods varied, with caregivers recruited through health care settings such as memory clinics and hospitals (n=20), community organizations such as Alzheimer's Association and support groups (n=31), online platforms including websites and social media (n=17), and print media such as flyers and newspaper ads (n=18).

Demographic characteristics typically included middle-aged to older adults with a mean age of around 60 years. Caregivers were predominantly female (70% - 80%) and had varied educational backgrounds, with many having at least some college education. Employment status was mixed, with some caregivers employed full-time, part-time, or retired. Care recipients were generally older adults, aged >70 years, with a slight majority being female. The relationship between caregivers and care recipients was primarily spousal or adult children caring for their parents. Sample sizes varied widely, ranging from small pilot studies with about 10 - 30 participants (n=8) to larger trials with >100 participants (n=9). Many studies did not report their target recruitment goals (n=36) or attrition rate (n=22). Among those that reported these outcomes, most met their target recruitment (n=8), whereas some did not (n=5). Attrition rates also varied, with some studies reporting high retention (80% - 90%; n=6) and others reporting significant dropout rates (20% - 30%; n=8) (Table 1).

Application Characteristics

The dataset includes 44 unique applications designed to support caregivers of individuals with various chronic conditions. These applications are categorized as web-based (n=30), mobile (n=12), and tablet (n=2) platforms. The design of these applications often followed adaptations of in-person or previously developed intervention (n=9), a user-centered approach (n=9), and theory-based design (n=3). Other design approaches involve iterative development with user feedback and collaboration with health care professionals (n=3).

The primary functions of these applications include education (n=39), social support (n=19), and self-management tools (n=21). Common components across applications included informational resources (n=39), interactive services such as forums and chat (n=22), and personalized feedback or coaching from researchers or interventionists (eg, coaches and health care professionals; n=10).

Many applications provided educational content for caregivers, including a range of informational resources (n=37) to support chronic condition management and caregiving strategies. Structured modules further enhanced this learning experience by covering caregiving-related topics such as symptom management and communication skills. Practical tools were another important aspect of these applications, with features such as calendars and schedules helping caregivers organize tasks, appointments, and daily activities. Additionally, some applications included resource libraries designed for easy access to papers, videos, and external links for caregiving support.

Interactive services (n=22) were another key feature of these applications, allowing caregivers to connect with others through forums, chat rooms, and social networks. These platforms provided a space for caregivers to share experiences, offer support, and build community. In addition to forums, many applications include private messaging features for more personal communication with other caregivers or health care professionals. Customization and personalization options, including reminder notifications or email alerts, made the applications more responsive to individual caregiver needs.

Support tools in these applications (n=19) often incorporated interventionist-guided approaches to provide step-by-step assistance based on cognitive-behavioral principles or problem-solving techniques. Caregivers also received regular personalized feedback and coaching from interventionists to help monitor progress and adjust strategies as needed. Tracking and monitoring tools were also common (n=21), with applications offering progress tracking features such as daily assessments (eg, ecological momentary assessments) to regularly track mood, stress, and other key metrics.

Only about half of all studies (n=27) reported that participants were trained on how to use the app by research personnel (eg, in-person: n=10, phone calls: n=3, or online: n=6). Participants were typically trained through initial orientation sessions. App use was typically measured through log data (n=24) or self-reported usage (n=5), but many studies did not report app use (n=24). Actual app use during the intervention varied, with

XSL•FO
RenderX

5 studies showing consistent engagement and 11 studies reporting drop-offs in usage over time. Data on user satisfaction were reported in about half of all studies (n=24), with 16 studies reporting moderate to high satisfaction among caregivers.

Intervention Effects on Study Outcomes

Interventions with a comparison group (n=25) followed a waitlist model or received standard care. Control group participants typically had access to basic resources, such as printed materials or standard educational content, but did not receive the more comprehensive features or digital tools provided to the treatment group.

The primary and secondary outcomes across all studies predominantly focused on caregiver well-being and health (eg, depression and stress; n=33), burden (n=16), self-efficacy and competence (n=16), feasibility and acceptability (n=32), social support (n=9), quality of life (n=4), care recipient's symptoms (n=2), coping strategies (n=4), and intervention feedback (n=21).

Overall, the intervention effects on outcomes were generally positive. Many studies reported improvements in primary and secondary outcomes (eg, caregiver well-being, reduced caregiver burden, and enhanced mental health) across time or condition (n=39). For example, caregivers experienced lower levels of depression and anxiety and increased self-efficacy [41]. Usability and acceptability were commonly assessed (n=25), with caregivers often discussing ease of use and satisfaction. Qualitative results showed that a prevalent reason for app usage was the emotional impact of caregiving (n=8), where caregivers shared their experiences of stress, guilt, and isolation. Support and communication also emerged as significant themes regarding why and what caregivers need in terms of support, highlighting the importance of social support networks and effective communication with health care providers. Additional themes included practical caregiving challenges (n=8) and the need for personalized tools (n=2). However, 11 studies reported declining engagement over time, indicating the need for ongoing support and engagement strategies. Moreover, 10 studies showed null effects for outcomes across time or condition, which warrants further discussion (eg, possibility of app content or design failing to meet user needs, inadequate user training, or low engagement) [43,45,57,64,67-69,77,78,84].

Discussion

Principal Findings

This review synthesized 49 studies evaluating application-based interventions for family caregivers of chronically ill older adults (eg, ADRD, cancer, and stroke). These studies primarily examined the impact of applications on outcomes, including caregiver well-being, burden, and mental health. While most interventions were perceived as beneficial for enhancing well-being and social support, several design and research limitations were noted. In the following "Discussion" section, we discuss how these applications support caregivers and identify design and evaluation challenges to guide future research.

Application-Based Intervention Support for Caregivers

The majority of included studies aimed to provide caregivers with education, emotional support, and self-management tools. ADRD caregivers, in particular, benefited from applications that offered resources for managing caregiving tasks and emotional well-being. Caregivers appreciated interventions that enhanced their access to information and social support, especially during stressful caregiving periods. However, some caregivers reported dissatisfaction with applications that were too complex or poorly aligned with their needs, highlighting the importance of user-centered design. In line with broader caregiving research, these findings underscore how psychosocial factors—such as caregivers' perceived social support, resilience, and sense of coherence—shape their ability to benefit from digital tools and influence the extent to which applications can meaningfully support their emotional well-being [22,60,78,84].

For cancer and stroke caregivers, the feedback was similar. Applications focused on information delivery and psychosocial support were generally well received, but challenges arose when the app content was too general or failed to provide disease-specific advice. Across all caregiver groups, there was a demand for more personalized, adaptable interventions that could cater to the evolving needs of caregivers and care recipients as their conditions progressed. These needs are also influenced by well-established psychosocial influences on caregiver stress and coping, including role overload, work-related strain, and the extent to which caregivers have access to emotional and informational resources [39,43,58,76,77,80].

While many interventions aimed to reduce negative emotional states, they did not adequately address broader aspects of well-being, such as promoting caregivers' own self-care practices. Additionally, many interventions focused on reducing stress and burden without a parallel effort to enhance positive emotions or resilience, which could result in caregivers feeling only marginally supported. This gap is notable given evidence that resilience, emotional regulation, and prosocial coping behaviors are key determinants of caregiver health and may moderate the impact of app interventions [24,85]. There is a need for more holistic intervention approaches, promoting not just stress relief but also strategies to enhance self-care, emotional fulfillment, and resilience [25]. This highlights the need for application-based interventions to shift from solely focusing on minimizing negative outcomes to actively promoting well-being. Integrating established psychosocial constructs—and drawing on validated measures such as the Zarit Burden Interview, resilience scales, and sense-of-coherence instruments—may help future app designs more effectively target mechanisms that drive meaningful change for caregivers [5,6,36,38,46,86].

Design and Evaluation Considerations

A significant issue identified in many studies was the lack of clarity regarding which components of multifaceted interventions were responsible for specific outcomes. For instance, some applications aimed to increase both information access and social support, but studies often failed to specify which intervention component influenced outcomes such as

XSL•FO
RenderX

well-being or burden. This lack of specificity makes it challenging to identify the mechanisms through which the app intervention impacts outcomes, a key step in developing targeted behavioral interventions [32].

In some cases, negative or null findings were reported without clear explanations. Exploring potential reasons for these outcomes is crucial, including the possibility of app content or design failing to meet user needs, inadequate user training, or low engagement [29]. Additionally, negative or null findings were frequently underreported, making it challenging to understand why some interventions failed to produce the desired effects, which could be a result of publication bias. Unintended consequences, such as caregivers feeling overwhelmed by app content or difficulty integrating app use into their daily routines, were rarely explored but warrant attention in future studies. Future research should also explore potential unintended consequences or reasons for lack of effects, such as inadequate app design, caregiver disengagement, or issues with the intervention's delivery.

Another issue relates to studies that claimed to use mixed methods designs. While both qualitative and quantitative data were collected, these data were not integrated or triangulated to provide a comprehensive understanding of the intervention's effects, which may limit the study's implications.

Research Considerations

Many studies in this review lacked consistent reporting on key aspects of application-based interventions, such as usage, participant training procedures, and fidelity measures. For instance, of 49 studies, only half (n=27) reported whether participants received any training on using the applications. This omission is particularly problematic given that technology proficiency can vary widely among caregivers, especially older adults. Past research on technology training and proficiency highlights the importance of offering caregivers adequate guidance on digital tools [87], yet many studies failed to measure or report this. Studies that did not measure participants' baseline technology skills may have missed critical factors influencing application engagement and outcomes. Additionally, the lack of data on app satisfaction and usage makes it difficult to assess how well the applications were received and whether caregivers used them as intended. This lack of information limits the ability to assess intervention fidelity (only 20% of studies reported any fidelity measures) and their real-world applicability.

Moreover, adapting in-person behavioral interventions to app formats was common but not always well documented. When delivered via digital platforms, behavioral interventions must adhere to principles such as those outlined in the NIH Stage Model for Behavioral Intervention Development [32], which emphasizes the importance of specifying mechanisms, conducting feasibility studies, and ensuring intervention fidelity. Similarly, recruitment goals and attrition rates were often missing, particularly in feasibility studies and RCTs, which should adhere to standards such as CONSORT (Consolidated Standards of Reporting Trials) for transparent reporting [88]. However, many studies in this review did not clearly align their design with this model, resulting in unclear stages of intervention development and evaluation, where some studies attempted to combine intervention development and RCT phases without clearly defining how these stages were managed.

Limitations of Current Studies

Several limitations were identified. First, sample sizes were small and geographically limited, reducing the generalizability of findings. The studies predominantly involved caregivers from North America and Europe, with minimal representation from Asia and other regions. This limits the cultural generalizability of the findings, as caregiving practices, expectations of family responsibility, and help-seeking behaviors vary across cultural contexts. Additionally, the feasibility and effectiveness of application-based interventions may differ in settings where access to smartphones, internet connectivity, digital literacy, or technological infrastructure is limited [8]. Additionally, marginalized groups and long-distance caregivers were underrepresented, given limited participant diversity in race, ethnicity, and socioeconomic status. Further, many studies did not account for disease stage or care recipient etiologies, which could have provided more nuanced insights into the effectiveness of these applications for different caregiver populations.

Research design weaknesses were also prevalent. Many studies lacked robust control groups, relied on small sample sizes, or used short follow-up periods, limiting the strength and stability of the reported findings. In several studies, the intervention development process was not clearly described (eg, how the app was developed or adapted to digital formats), and reporting was inconsistent regarding levels of rigor, evaluation procedures, and adherence monitoring. These methodological limitations make it difficult to draw firm conclusions about the effectiveness or sustainability of application-based interventions for caregivers. Moreover, few studies examined how or why the interventions had any effects. Mediation or moderation analyses were rarely conducted, making it challenging to identify which components—such as educational content, stress-management tools, peer support features, or mood and symptom tracking—were most influential in improving caregiver outcomes. As a result, the mechanisms of change underlying application-based interventions remain unclear. To advance understanding of how app interventions influence caregiver outcomes, future studies should incorporate mediation and moderation analyses to clarify the mechanisms of change. Such work can help isolate whether specific features drive improvements in caregiver well-being. Finally, many interventions were prototypes or research-only platforms that were not publicly accessible, highlighting challenges related to scalability, implementation, and long-term integration into real-world caregiving contexts.

Limitations of the Current Review

This scoping review has several limitations. First, as a common limitation of scoping reviews [89], we did not include a critical appraisal of the methodological quality or risk of bias in the included studies, which may limit the validity of our findings. Additionally, the database search was conducted in December 2022; thus, any relevant studies published after this date were not included in the screening process. ChatGPT was released in November 2022 and significantly accelerated the development

XSL•FO
RenderX

of artificial intelligence (AI)-based applications. With the increasing interest in AI, future reviews can consider examining the application of AI to promote caregivers' health and well-being. Finally, we did not assess cost-effectiveness or broader health care system implications, which are important considerations for future research.

Conclusions

This review highlights the growing potential of application-based interventions to support family caregivers of older adults, particularly those caring for individuals with ADRD, cancer, and stroke. While many interventions are well-received and demonstrate positive effects on caregiver well-being, there are significant gaps in research design, evaluation, and reporting. Future studies should prioritize greater integration of self-care practices, clarity in intervention components, rigorous adherence to behavioral intervention frameworks, including the NIH Stage Model, and more diverse participant samples. Improved reporting on app usage, satisfaction, and recruitment outcomes is essential for advancing the field and ensuring these digital tools meet the complex needs of caregivers. Moreover, as technology evolves, ongoing research is needed to explore how emerging innovations can further enhance well-being. Addressing these gaps will be crucial for developing effective, scalable, and accessible digital interventions for caregivers across diverse populations.

Acknowledgments

Generative artificial intelligence tools were not used at any stage in the preparation of this manuscript.

Funding

XYL acknowledges support from a National Institute on Aging–funded T32 Training Program in Behavioral Geriatrics (NIA T32 AG049666). FF acknowledges support from the National Institute on Aging of the National Institutes of Health under award numbers K99AG073509 and R00AG073509. ESC acknowledges support from the National Institute on Aging under award number K01AG081540. The funding sources had no involvement other than providing financial support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search methods used for the review.
[DOCX File, 23 KB - aging_v9i1e76115_app1.docx]

Multimedia Appendix 2
Flowchart for the review process.
[PDF File, 106 KB - aging_v9i1e76115_app2.pdf]

Multimedia Appendix 3
Papers included in the study.
[DOCX File, 79 KB - aging_v9i1e76115_app3.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 192 KB - aging_v9i1e76115_app4.pdf]

References

1. Caregiving in the United States 2020. AARP. 2020 May 14. URL: https://www.aarp.org/ppi/info-2020/caregiving-in-the-united-states.html [accessed 2023-09-12]
2. Alzheimer's disease caregivers. : Alzheimer's Association; 2020 Mar URL: https://act.alz.org/site/DocServer/caregivers_fact_sheet.pdf?docID=3022 [accessed 2026-05-09]
3. Busch IM, Moretti F, Travaini G, Wu AW, Rimondini M. Humanization of care: key elements identified by patients, caregivers, and healthcare providers. a systematic review. Patient 2019 Oct;12(5):461-474. [doi: 10.1007/s40271-019-00370-1] [Medline: 31203515]
4. Alhasan DM, Hirsch JA, Jackson CL, Miller MC, Cai B, Lohman MC. Neighborhood characteristics and the mental health of caregivers cohabiting with care recipients diagnosed with Alzheimer's disease. Int J Environ Res Public Health 2021 Jan 21;18(3):913. [doi: 10.3390/ijerph18030913] [Medline: 33494425]

XSL•FO
RenderX

5. Ashare RL, Brewer B, Patterson F, Hubbard A, Longacre ML. Psychological and behavioral moderators of physical health among caregivers and non-caregivers. Geriatr Nurs 2022;43:77-84. [doi: 10.1016/j.gerinurse.2021.11.004] [Medline: 34844128]
6. Cohen SA, Cook S, Kelley L, Sando T, Bell AE. Psychosocial factors of caregiver burden in child caregivers: results from the new national study of caregiving. Health Qual Life Outcomes 2015 Aug 7;13:120. [doi: 10.1186/s12955-015-0317-2] [Medline: 26246132]
7. Hajek A, König HH. Social isolation and loneliness of older adults in times of the COVID-19 pandemic: can use of online social media sites and video chats assist in mitigating social isolation and loneliness? Gerontology 2021;67(1):121-124. [doi: 10.1159/000512793] [Medline: 33264778]
8. Lai J, Widmar NO. Revisiting the digital divide in the COVID-19 era. Appl Econ Perspect Policy 2021 Mar;43(1):458-464. [doi: 10.1002/aepp.13104] [Medline: 33230409]
9. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System: National Academies Press; 2020. [doi: 10.17226/25663]
10. Abdollahpour I, Nedjat S, Salimi Y. Positive aspects of caregiving and caregiver burden: a study of caregivers of patients with dementia. J Geriatr Psychiatry Neurol 2018 Jan;31(1):34-38. [doi: 10.1177/0891988717743590] [Medline: 29187025]
11. Anderson JG, Hundt E, Dean M, Rose KM. "A fine line that we walk every day": self-care approaches used by family caregivers of persons with dementia. Issues Ment Health Nurs 2019 Mar;40(3):252-259. [doi: 10.1080/01612840.2018.1499156] [Medline: 30422041]
12. Bressan V, Visintini C, Palese A. What do family caregivers of people with dementia need? A mixed-method systematic review. Health Soc Care Community 2020 Nov;28(6):1942-1960. [doi: 10.1111/hsc.13048] [Medline: 32542963]
13. Boots LMM, de Vugt ME, van Knippenberg RJM, Kempen GIJM, Verhey FRJ. A systematic review of Internet-based supportive interventions for caregivers of patients with dementia. Int J Geriatr Psychiatry 2014 Apr;29(4):331-344. [doi: 10.1002/gps.4016] [Medline: 23963684]
14. Castillo LIR, Tran V, Hadjistavropoulos T. Are mobile apps meeting the needs of caregivers of people living with dementia? An evaluation of existing apps for caregivers. Aging Ment Health 2024 Apr;28(4):577-586. [doi: 10.1080/13607863.2023.2177832] [Medline: 36775643]
15. Chi NC, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. J Telemed Telecare 2015 Jan;21(1):37-44. [doi: 10.1177/1357633X14562734] [Medline: 25475220]
16. Czaja SJ, Charness N, Rogers WA, Sharit J, Moxley JH, Boot WR. The benefits of technology for engaging aging adults: findings from the PRISM 2.0 trial. Innov Aging 2024;8(6):igae042. [doi: 10.1093/geroni/igae042] [Medline: 38854853]
17. Guisado-Fernández E, Giunti G, Mackey LM, Blake C, Caulfield BM. Factors influencing the adoption of smart health technologies for people with dementia and their informal caregivers: scoping review and design framework. JMIR Aging 2019 Apr 30;2(1):e12192. [doi: 10.2196/12192] [Medline: 31518262]
18. Schüz B, Czerniawski A, Davie N, et al. Leisure time activities and mental health in informal dementia caregivers. Appl Psychol Health Well Being 2015 Jul;7(2):230-248. [doi: 10.1111/aphw.12046] [Medline: 26097155]
19. Sheehan OC, Haley WE, Howard VJ, Huang J, Rhodes JD, Roth DL. Stress, burden, and well-being in dementia and nondementia caregivers: insights from the caregiving transitions study. Gerontologist 2021 Jul 13;61(5):670-679. [doi: 10.1093/geront/gnaa108] [Medline: 32816014]
20. Faverio M. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. 2022 Jul 13. URL: https://www.pewresearch.org/fact-tank/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2023-01-30]
21. Hill JR, Min EE, Abebe E, Holden RJ. Telecaregiving for dementia: a mapping review of technological and nontechnological interventions. Gerontologist 2024 Jan 1;64(1):gnad026. [doi: 10.1093/geront/gnad026] [Medline: 36919597]
22. Vaughan C, Trail TE, Mahmud A, Dellva S, Tanielian T, Friedman E. Informal caregivers' experiences and perceptions of a web-based peer support network: mixed-methods study. J Med Internet Res 2018 Aug 28;20(8):e257. [doi: 10.2196/jmir.9895] [Medline: 30154074]
23. Chelberg GR, Neuhaus M, Mothershaw A, Mahoney R, Caffery LJ. Mobile apps for dementia awareness, support, and prevention - review and evaluation. Disabil Rehabil 2022 Aug;44(17):4909-4920. [doi: 10.1080/09638288.2021.1914755] [Medline: 34034601]
24. Guessi Margarido M, Shah A, Seto E. Smartphone applications for informal caregivers of chronically ill patients: a scoping review. NPJ Digit Med 2022 Mar 21;5(1):33. [doi: 10.1038/s41746-022-00567-z] [Medline: 35314766]
25. Guo Y, Yang F, Hu F, Li W, Ruggiano N, Lee HY. Existing mobile phone apps for self-care management of people with Alzheimer disease and related dementias: systematic analysis. JMIR Aging 2020 Jan 24;3(1):e15290. [doi: 10.2196/15290] [Medline: 32012045]
26. Berkowsky RW, Sharit J, Czaja SJ. Factors predicting decisions about technology adoption among older adults. Innov Aging 2018 Jan;2(1):igy002. [doi: 10.1093/geroni/igy002] [Medline: 30480129]
27. Blok M, van Ingen E, de Boer AH, Slootman M. The use of information and communication technologies by older people with cognitive impairments: from barriers to benefits. Comput Human Behav 2020 Mar;104:106173. [doi: 10.1016/j.chb.2019.106173]

XSL•FO
RenderX

28. Chia BX, Teo WJ, Swa JX, et al. Design principles and considerations for a localisation app for caregivers of people with dementia. Presented at: 2022 International Conference on Platform Technology and Service (PlatCon); Aug 22-24, 2022; Jeju, Republic of Korea p. 19-24. [doi: 10.1109/PlatCon55845.2022.9932088]
29. Grossman MR, Zak DK, Zelinski EM. Mobile apps for caregivers of older adults: quantitative content analysis. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e162. [doi: 10.2196/mhealth.9345] [Medline: 30061093]
30. Sala-González M, Pérez-Jover V, Guilabert M, Mira JJ. Mobile apps for helping informal caregivers: a systematic review. Int J Environ Res Public Health 2021 Feb 10;18(4):1702. [doi: 10.3390/ijerph18041702] [Medline: 33578819]
31. Park JYE, Tracy CS, Gray CS. Mobile phone apps for family caregivers: a scoping review and qualitative content analysis. Digit Health 2022;8:20552076221076672. [doi: 10.1177/20552076221076672] [Medline: 35154806]
32. Nielsen L, Riddle M, King JW, et al. The NIH science of behavior change program: transforming the science through a focus on mechanisms of change. Behav Res Ther 2018 Feb;101:3-11. [doi: 10.1016/j.brat.2017.07.002] [Medline: 29110885]
33. Chronic diseases and cognitive decline – a public health issue. : Centers for Disease Control and Prevention (CDC); 2020 URL: https://www.cdc.gov/healthy-aging-data/media/pdfs/Chronic-Diseases-and-Cognitive-Decline-Pages-508.pdf [accessed 2023-01-30]
34. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews. JBI Manual for Evidence Synthesis. 2020. URL: https://jbi-global-wiki.refined.site/space/MANUAL/355862497/jbi-global-wiki.refined.site [accessed 2024-12-04]
35. Hickner A. Search methods for intervention studies for informal caregivers of older adults using mobile/web apps to promote social support and well-being: a scoping review. Open Science Framework (OSF). 2023. URL: https://osf.io/c9bmp [accessed 2025-12-15]
36. Barbabella F, Poli A, Andréasson F, et al. A web-based psychosocial intervention for family caregivers of older people: results from a mixed-methods study in three European countries. JMIR Res Protoc 2016 Oct 6;5(4):e196. [doi: 10.2196/resprot.5847] [Medline: 27713113]
37. Bartels SL, van Knippenberg RJM, Köhler S, et al. The necessity for sustainable intervention effects: lessons-learned from an experience sampling intervention for spousal carers of people with dementia. Aging Ment Health 2020 Dec 1;24(12):2082-2093. [doi: 10.1080/13607863.2019.1647130]
38. Blom MM, Zarit SH, Groot Zwaaftink RBM, Cuijpers P, Pot AM. Effectiveness of an internet intervention for family caregivers of people with dementia: results of a randomized controlled trial. PLoS One 2015;10(2):e0116622. [doi: 10.1371/journal.pone.0116622] [Medline: 25679228]
39. Bodschwinna D, Lorenz I, Bauereiß N, Gündel H, Baumeister H, Hönig K. A psycho-oncological online intervention supporting partners of patients with cancer (PartnerCARE): results from a randomized controlled feasibility trial. Psychooncology 2022 Jul;31(7):1230-1242. [doi: 10.1002/pon.5917] [Medline: 35233880]
40. Bruinsma J, Peetoom K, Bakker C, et al. Tailoring and evaluating the web-based “Partner in Balance” intervention for family caregivers of persons with young-onset dementia. Internet Interv 2021 Sep;25(Bruinsma):100390. [doi: 10.1016/j.invent.2021.100390] [Medline: 33996507]
41. Bruinsma J, Peetoom K, Boots L, et al. Tailoring the web-based “Partner in Balance” intervention to support spouses of persons with frontotemporal dementia. Internet Interv 2021 Dec;26(Bruinsma):100442. [doi: 10.1016/j.invent.2021.100442] [Medline: 34430221]
42. Carr AL, Jones J, Mikulich Gilbertson S, et al. Impact of a mobilized stress management program (Pep-Pal) for caregivers of oncology patients: mixed-methods study. JMIR Cancer 2019 May 3;5(1):e11406. [doi: 10.2196/11406] [Medline: 31066678]
43. Caunca MR, Simonetto M, Hartley G, Wright CB, Czaja SJ. Design and usability testing of the stroke caregiver support system: a mobile-friendly website to reduce stroke caregiver burden. Rehabil Nurs 2020;45(3):166-177. [doi: 10.1097/RNJ.0000000000000196] [Medline: 30418319]
44. Chiu T, Marziali E, Colantonio A, et al. Internet-based caregiver support for Chinese Canadians taking care of a family member with Alzheimer disease and related dementia. Can J Aging 2009 Dec;28(4):323-336. [doi: 10.1017/S0714980809990158] [Medline: 19925698]
45. Christie HL, Dam AEH, van Boxtel M, Köhler S, Verhey F, de Vugt ME. Lessons learned from an effectiveness evaluation of Inlife, a web-based social support intervention for caregivers of people with dementia: randomized controlled trial. JMIR Aging 2022 Dec 7;5(4):e38656. [doi: 10.2196/38656] [Medline: 36476485]
46. Cristancho-Lacroix V, Wrobel J, Cantegreil-Kallen I, Dub T, Rouquette A, Rigaud AS. A web-based psychoeducational program for informal caregivers of patients with Alzheimer’s disease: a pilot randomized controlled trial. J Med Internet Res 2015 May 12;17(5):e117. [doi: 10.2196/jmir.3717] [Medline: 25967983]
47. Dam AEH, Christie HL, Smeets CMJ, van Boxtel MPJ, Verhey FRJ, de Vugt ME. Process evaluation of a social support platform “Inlife” for caregivers of people with dementia. Internet Interv 2019 Mar;15:18-27. [doi: 10.1016/j.invent.2018.09.002] [Medline: 30510911]
48. Dam AEH, van Boxtel MPJ, Rozendaal N, Verhey FRJ, de Vugt ME. Development and feasibility of Inlife: a pilot study of an online social support intervention for informal caregivers of people with dementia. PLoS One 2017;12(9):e0183386. [doi: 10.1371/journal.pone.0183386] [Medline: 28886056]

XSL•FO
RenderX

49. Davis BH, Shehab M, Shenk D, Nies M. E-mobile pilot for community-based dementia caregivers identifies desire for security. Gerontechnology 2015;13(3):332-336. [doi: 10.4017/gt.2015.13.3.003.00]
50. Duggleby W, Ploeg J, McAiney C, et al. A comparison of users and nonusers of a web-based intervention for carers of older persons with Alzheimer disease and related dementias: mixed methods secondary analysis. J Med Internet Res 2019 Oct 17;21(10):e14254. [doi: 10.2196/14254] [Medline: 31625947]
51. Duggleby W, Ploeg J, McAiney C, et al. Web-based intervention for family carers of persons with dementia and multiple chronic conditions (My Tools 4 Care): pragmatic randomized controlled trial. J Med Internet Res 2018 Jun 29;20(6):e10484. [doi: 10.2196/10484] [Medline: 29959111]
52. Duggleby W, Jovel Ruiz K, Ploeg J, et al. Mixed-methods single-arm repeated measures study evaluating the feasibility of a web-based intervention to support family carers of persons with dementia in long-term care facilities. Pilot Feasibility Stud 2018;4(1):165. [doi: 10.1186/s40814-018-0356-7] [Medline: 30410783]
53. Ferré-Grau C, Raigal-Aran L, Lorca-Cabrera J, et al. A mobile app-based intervention program for nonprofessional caregivers to promote positive mental health: randomized controlled trial. JMIR Mhealth Uhealth 2021 Jan 22;9(1):e21708. [doi: 10.2196/21708] [Medline: 33480852]
54. Fossey J, Charlesworth G, Fowler JA, et al. Online education and cognitive behavior therapy improve dementia caregivers' mental health: a randomized trial. J Am Med Dir Assoc 2021 Jul;22(7):1403-1409. [doi: 10.1016/j.jamda.2020.10.009] [Medline: 33288467]
55. Gaugler JE, Reese M, Tanler R. Care to plan: an online tool that offers tailored support to dementia caregivers. Gerontologist 2016 Dec;56(6):1161-1174. [doi: 10.1093/geront/gnv150] [Medline: 26603183]
56. Goodridge D, Reis N, Neiser J, et al. An app-based mindfulness-based self-compassion program to support caregivers of people with dementia: participatory feasibility study. JMIR Aging 2021 Nov 26;4(4):e28652. [doi: 10.2196/28652] [Medline: 34842530]
57. Gustafson DH, Gustafson DH, Cody OJ, Chih MY, Johnston DC, Asthana S. Pilot test of a computer-based system to help family caregivers of dementia patients. J Alzheimers Dis 2019;70(2):541-552. [doi: 10.3233/JAD-190052] [Medline: 31256126]
58. Heynsbergh N, Heckel L, Botti M, Livingston PM. A smartphone app to support carers of people living with cancer: a feasibility and usability study. JMIR Cancer 2019 Jan 31;5(1):e11779. [doi: 10.2196/11779] [Medline: 30702432]
59. Hughes JC, Banerjee T, Goodman G, Lawhorne L. A preliminary qualitative analysis on the feasibility of using gaming technology in caregiver assessment. J Technol Hum Serv 2017 Jul 3;35(3):183-198. [doi: 10.1080/15228835.2017.1347553]
60. Huis In Het Veld JG, Willemse BM, van Asch IF, et al. Online self-management support for family caregivers dealing with behavior changes in relatives with dementia (part 2): randomized controlled trial. J Med Internet Res 2020 Feb 25;22(2):e13001. [doi: 10.2196/13001] [Medline: 32130142]
61. Jordan MM, Freytes IM, Orozco T, et al. The RESCUE problem solving intervention for stroke caregivers: a mixed-methods pilot study. Rehabil Psychol 2022 Nov;67(4):484-496. [doi: 10.1037/rep0000460] [Medline: 35925687]
62. Kajiyama B, Thompson LW, Eto-Iwase T, et al. Exploring the effectiveness of an internet-based program for reducing caregiver distress using the iCare Stress Management e-Training Program. Aging Ment Health 2013;17(5):544-554. [doi: 10.1080/13607863.2013.775641] [Medline: 23461355]
63. Kales HC, Gitlin LN, Stanislawski B, et al. Effect of the WeCareAdvisor TM on family caregiver outcomes in dementia: a pilot randomized controlled trial. BMC Geriatr 2018 May 10;18(1):113. [doi: 10.1186/s12877-018-0801-8] [Medline: 29747583]
64. Kovaleva M, Blevins L, Griffiths PC, Hepburn K. An online program for caregivers of persons living with dementia: lessons learned. J Appl Gerontol 2019 Feb;38(2):159-182. [doi: 10.1177/0733464817705958] [Medline: 28460557]
65. Leung AYM, Cheung T, Fong TKH, Zhao IY, Kabir ZN. The use of an electronic painting platform by family caregivers of persons with dementia: a feasibility and acceptability study. Healthcare (Basel) 2022 May 9;10(5):870. [doi: 10.3390/healthcare10050870] [Medline: 35628007]
66. Lewis ML, Hobday JV, Hepburn KW. Internet-based program for dementia caregivers. Am J Alzheimers Dis Other Demen 2010 Dec;25(8):674-679. [doi: 10.1177/1533317510385812] [Medline: 21131674]
67. Linden A, Jolliff A, Gonzalez D, et al. "It made me feel like I wasn't alone in the darkness": exploring dementia care network communication and coordination through a digital health platform. J Am Med Inform Assoc 2022 Nov 14;29(12):2003-2013. [doi: 10.1093/jamia/ocac172] [Medline: 36166717]
68. Llaneza DH, Llaneza AJ, Kozlov E. Perceived benefits and barriers of mHealth mindfulness use for caregivers of older adults with cognitive impairment: a qualitative exploration. Clin Gerontol 2024;47(1):66-77. [doi: 10.1080/07317115.2022.2154728] [Medline: 36502294]
69. Lundberg S. The results from a two-year case study of an information and communication technology support system for family caregivers. Disabil Rehabil Assist Technol 2014 Jul;9(4):353-358. [doi: 10.3109/17483107.2013.814170] [Medline: 23855387]
70. Marziali E, Garcia LJ. Dementia caregivers' responses to 2 internet-based intervention programs. Am J Alzheimers Dis Other Demen 2011 Feb;26(1):36-43. [doi: 10.1177/1533317510387586] [Medline: 21282276]

XSL•FO
RenderX

71. Meichsner F, Theurer C, Wilz G. Acceptance and treatment effects of an internet-delivered cognitive-behavioral intervention for family caregivers of people with dementia: a randomized-controlled trial. J Clin Psychol 2019 Apr;75(4):594-613. [doi: 10.1002/jclp.22739] [Medline: 30597537]
72. Metcalfe A, Jones B, Mayer J, et al. Online information and support for carers of people with young-onset dementia: a multi-site randomised controlled pilot study. Int J Geriatr Psychiatry 2019 Oct;34(10):1455-1464. [doi: 10.1002/gps.5154] [Medline: 31111516]
73. Núñez-Naveira L, Alonso-Búa B, de Labra C, et al. UnderstAID, an ICT platform to help informal caregivers of people with dementia: a pilot randomized controlled study. Biomed Res Int 2016;2016:5726465. [doi: 10.1155/2016/5726465] [Medline: 28116300]
74. Park E, Park H, Kim EK. The effect of a comprehensive mobile application program (CMAP) for family caregivers of home‐dwelling patients with dementia: a preliminary research. Japan Journal Nursing Sci 2020 Oct;17(4):e12343. [doi: 10.1111/jjns.12343] [Medline: 31397933]
75. Pensak NA, Carr AL, Jones J, et al. A pilot study of mobilized intervention to help caregivers of oncology patients manage distress. Psychooncology 2021 Apr;30(4):520-528. [doi: 10.1002/pon.5597] [Medline: 33217070]
76. Pierce LL, Steiner V. Usage and design evaluation by family caregivers of a stroke intervention web site. J Neurosci Nurs 2013 Oct;45(5):254-261. [doi: 10.1097/JNN.0b013e31829dba61] [Medline: 24025464]
77. Pierce LL, Steiner VL, Khuder SA, Govoni AL, Horn LJ. The effect of a Web-based stroke intervention on carers' well-being and survivors' use of healthcare services. Disabil Rehabil 2009;31(20):1676-1684. [doi: 10.1080/09638280902751972] [Medline: 19479528]
78. Pot AM, Blom MM, Willemse BM. Acceptability of a guided self-help Internet intervention for family caregivers: mastery over dementia. Int Psychogeriatr 2015 Aug;27(8):1343-1354. [doi: 10.1017/S1041610215000034] [Medline: 25648589]
79. Reblin M, Ketcher D, Forsyth P, et al. Outcomes of an electronic social network intervention with neuro-oncology patient family caregivers. J Neurooncol 2018 Sep;139(3):643-649. [doi: 10.1007/s11060-018-2909-2] [Medline: 29808340]
80. Reblin M, Ketcher D, Forsyth P, et al. Feasibility of implementing an electronic social support and resource visualization tool for caregivers in a neuro-oncology clinic. Support Care Cancer 2018 Dec;26(12):4199-4206. [doi: 10.1007/s00520-018-4293-z] [Medline: 29948397]
81. Rettinger L, Zeuner L, Werner K, et al. A mixed-methods evaluation of a supporting app for informal caregivers of people with dementia. 2020 Jun 30 Presented at: PETRA '20: Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments; Jun 30 to Jul 3, 2020; Corfu, Greece p. 337-344. [doi: 10.1145/3389189.3397981]
82. Rodriguez K, Fugard M, Amini S, et al. Caregiver response to an online dementia and caregiver wellness education platform. J Alzheimers Dis Rep 2021;5(1):433-442. [doi: 10.3233/ADR-200292] [Medline: 34368629]
83. Sikder AT, Yang FC, Schafer R, Dowling GA, Traeger L, Jain FA. Mentalizing imagery therapy mobile app to enhance the mood of family dementia caregivers: feasibility and limited efficacy testing. JMIR Aging 2019 Mar 21;2(1):e12850. [doi: 10.2196/12850] [Medline: 31518275]
84. Zimmerman S, Sloane PD, Ward K, et al. Helping dementia caregivers manage medical problems: benefits of an educational resource. Am J Alzheimers Dis Other Demen 2018 May;33(3):176-183. [doi: 10.1177/1533317517749466] [Medline: 29301414]
85. Wilks SE, Croom B. Perceived stress and resilience in Alzheimer's disease caregivers: testing moderation and mediation models of social support. Aging Ment Health 2008 May;12(3):357-365. [doi: 10.1080/13607860801933323] [Medline: 18728949]
86. Bédard M, Molloy DW, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. Gerontologist 2001 Oct;41(5):652-657. [doi: 10.1093/geront/41.5.652] [Medline: 11574710]
87. Czaja SJ, Sharit J. Designing Training and Instructional Programs for Older Adults: CRC Press; 2012. [doi: 10.1201/b13018]
88. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. EQUATOR Network. 2010. URL: https://www.equator-network.org/reporting-guidelines/consort/ [accessed 2023-06-06]
89. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]

Abbreviations

ADRD: Alzheimer disease or related dementias
AI: artificial intelligence
CDC: Centers for Disease Control and Prevention
CONSORT: Consolidated Standards of Reporting Trials
IADL: instrumental activities of daily living
NIH: National Institutes of Health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized clinical trial

XSL•FO
RenderX

Edited by Y Luo; submitted 16.Apr.2025; peer-reviewed by J Gehrmann, M Marti-Vilar; revised version received 17.Dec.2025; accepted 03.Feb.2026; published 11.Jun.2026.

Please cite as:
Lin XY, Falzarano F, Chang ES, Hon K, Zhang R, Hickner A, Czaja S
Application-Based Interventions for Family Caregivers of Older Adults: Scoping Review
JMIR Aging 2026;9:e76115
URL: https://aging.jmir.org/2026/1/e76115
doi:10.2196/76115

© Xin Yao Lin, Francesca Falzarano, E-Shien Chang, Kristin Hon, Ruoying Zhang, Andy Hickner, Sara Czaja. Originally published in JMIR Aging (https://aging.jmir.org), 11.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Multisensory Integration Training on Postural Stability Characteristics and Fall Risk in Older Adults: Systematic Review and Meta-Analysis

Chaoyu Guo[1], BSc; Lulu Yin[1], PhD; Peng Chen[1], MSc; Jianglong Zhan[1], MSc; Zhongqi Yu[1], MSc; Teck Cheng Tan[1], MSc; Yaru Wei[1], BSc; Yixue Gong[1], BSc; Menghan Xu[1], BSc; Van Minh Le[1], BSc; Lin Wang[2,3], PhD

[1]School of Exercise and Health, Shanghai University of Sport, Shanghai, China

[2]Sports Medicine and Rehabilitation Center, Shanghai University of Sport, Shanghai, China

[3]Shanghai Shangti Orthopaedic Hospital, 188 Hengren Road, Yangpu District, Shanghai, China

Corresponding Author:
Lin Wang, PhD
Sports Medicine and Rehabilitation Center, Shanghai University of Sport, Shanghai, China

Abstract

Background: The risk of falls escalates with advancing age, a consequence of the concomitant degeneration of multiple physiological systems, altered sensory processing capabilities, and reduced postural control. Multisensory integration (MSI) training has been demonstrated to enhance the brain's processing of multisensory information. However, existing studies show considerable variability in intervention duration and training modalities, limiting comparability across studies and contributing to inconsistent findings.

Objective: This study aimed to systematically evaluate the effectiveness of MSI training on postural stability and fall risk in healthy older adults and provide an evidence base for clinical practice.

Methods: Databases including PubMed, Embase, and Cochrane Library were searched by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The methodological quality of the included randomized controlled trials was assessed using the Cochrane risk of bias tool, and publication bias was evaluated through funnel plots. Meta-analyses using R packages quantified effects via standardized mean differences (SMDs) and 95% CIs, with fixed or random effects models selected based on heterogeneity (I^2). Subgroup analyses explored age, intervention duration, and type.

Results: A total of 14 randomized controlled trials were included in the meta-analysis. Results showed that MSI training significantly reduced center of pressure displacement in the anterior-to-posterior displacement (SMD −1.64, 95% CI −2.78 to −0.49, P<.001) and center of pressure displacement in the medio-to-lateral displacement (SMD −1.37, 95% CI −2.68 to −0.07, P<.001). In terms of postural stability, MSI training significantly improved Berg Balance Scale scores (SMD 3.42, 95% CI 2.41 to 4.44, P=.006). In terms of fall risk, MSI training significantly reduced the time to complete the timed up and go test, and intervention type significantly moderated this effect (SMD −1.43, 95% CI −2.36 to −0.50, P<.001). Additionally, MSI training reduced the risk of falls (SMD −1.27, 95% CI −2.03 to −0.52, P<.001).

Conclusions: In conclusion, MSI training is an effective intervention for enhancing static and dynamic postural control and reducing fall risk in healthy older adults, suggesting a beneficial effect on postural stability and fall-related outcomes.

Trial Registration: PROSPERO CRD420251023158; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251023158

(JMIR Aging 2026;9:e80345) doi:10.2196/80345

KEYWORDS

multisensory integration training; older adults; postural stability; fall risk; meta-analysis

Introduction

According to the World Health Organization, the risk of falls among older adults increases significantly with age [1], and falls impose substantial societal burdens, both directly and indirectly, including increased demands on family caregiving and health care resources [2]. These events are driven by a combination of age-related physiological decline and underlying pathological factors. It is worth noting that aging is not a single-system degeneration but a complex process of cognitive, musculoskeletal, sensory processing, and other multisystems synergistic decline [3]. On the one hand, sensory-motor system function decreases with aging; on the other hand, aging also leads to changes in the way of processing sensory information in the central nervous system [4-6]. The cognitive demands required to maintain balance during challenging postural tasks

XSL•FO
RenderX

increase in older adults [7-9]. In addition, decreased brain volume and structural changes in the brain, such as gray matter atrophy and decreased white matter integrity, affect the transmission and integration of sensory information in older adults [10,11]. At the same time, changes in neurotransmitter levels interfere with sensory information transmission and processing [12,13]. In particular, functional and connectivity declines in core multisensory brain regions, such as parietal and temporal cortices, reduce multisensory integration in older adults, which in turn affects postural stability and fall risk [8,14,15]. Deficits in postural control are the most important risk factor for falls in older adults, directly weakening the physiological basis for maintaining postural stability and significantly increasing fall risk. Understanding the effects of multisensory integration (MSI) training on postural control is essential to address fall risk and improve quality of life in older adults.

The maintenance of postural stability relies on the dynamic coordination of multisensory integration, in which the brain integrates and assigns weights to visual, vestibular, and somatosensory information in real time [16]. It is important to note that the decline in multisensory integration is not irreversible [17,18], and MSI training is a method to optimize the brain's ability to process multisensory information through targeted interventions. This type of training usually combines visual, vestibular, and somatosensory stimuli to design a variety of tasks [19-21]. Through repeated practice, the nervous system can readjust the weights of different sensory inputs and strengthen the efficiency of integrating sensory information, thus improving postural control and balance stability [22,23]. Studies have shown that MSI training can be beneficial in improving balance function in older adults and can restore sensory integration deficits in patients with neurological impairments [22,23]. Allison et al [24] found that MSI training significantly improved postural stability and fall risk in older adults; however, Dowd et al [25] did not find that MSI training could improve postural stability and fall risk in older adults. Existing studies have found that the length and type of interventions in training programs vary greatly, resulting in limited comparability of results, and their methodological differences may affect the consistency of validation of intervention effects [26-28].

Drawing on previous literature [29-33], when analyzing the effects of interventions targeting memory impairment or balance in older adults, researchers commonly use comparisons based on age stratification, intervention type classification, and intervention duration staging. Combined with regression analysis to examine the influence of these variables on effect sizes, this methodological framework provides a reliable reference for interpreting the heterogeneity observed in this study. Therefore, this study aimed to integrate the existing evidence to quantify the effects of MSI training on postural stability and fall risk in older adults through systematic evaluation and meta-analysis, and to reveal the sources of heterogeneity through subgroup analyses of age, duration of intervention, and intervention type, so as to provide evidence-based medical evidence for the application of MSI training in older adults.

Methods

This systematic review and meta-analysis was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist and registered on PROSPERO on April 1, 2025 (CRD420251023158).

Search Strategy

Two authors (CG and YW) individually conducted a literature retrieval in 5 electronic databases—PubMed, Embase, Cochrane Library, WoS, and Scopus—from the inception of each database up to April 5, 2025, without language restriction. Medical Subject Headings (MeSH) terms or keywords were used during the literature search. The search terms were ("sensory reweight*" OR "sensory recalibrat*" OR "multisensory integrat*" OR "sensory train*") AND ("elderly*" OR "older adult*") AND ("balance" OR "falls" OR "fall risk"), and the detailed search strategy is shown in Multimedia Appendix 1. After the search was completed, all search results were imported into EndNote reference management software (Clarivate) for deduplication, and the reference lists of full-text literature were manually searched to identify studies that were not retrieved through the database search.

Research Inclusion and Exclusion Criteria

The study eligibility criteria are shown in Textbox 1.

XSL•FO
RenderX

Textbox 1. Study eligibility criteria.

Inclusion criteria

- Participants: healthy older adults aged ≥60 years, with no musculoskeletal dysfunction within the past 6 months, no confirmed peripheral somatosensory or vestibular deficits, visual acuity sufficient for daily activities, and a Montreal Cognitive Assessment score ≥26 or a Mini-Mental State Examination score ≥24
- Interventions: the intervention was multisensory integration (MSI) training, defined as any protocol that simultaneously manipulates or stimulates ≥2 sensory modalities (visual, vestibular, and somatosensory) to provoke central sensory integration through conflict or reweighting
- Control group: presence of a healthy control group to establish baseline changes and accurately identify the effects of MSI training
- Outcome measures: the primary outcome was center of pressure (COP) displacement (including anterior-to-posterior and medio-to-lateral displacements). Secondary outcomes included (1) postural stability measures under different test conditions (single-leg standing, 2-legged tandem standing, and 2-legged standing), specifically COP displacement and Berg Balance Scale scores; and (2) fall risk prediction indicators, including direct risk measures (fall risk, times/year) and functional measures (timed up and go test), and total intervention duration (weeks)

Exclusion criteria

- Nonrandomized controlled trials
- Insufficient data for analysis
- Duplicate publications
- Non–peer-reviewed literature
- Serious methodological flaws, defined as a lack of adequate randomization or allocation concealment, intergroup cross-over contamination, use of invalid or noncomparable outcome measures, or dropout rates exceeding 20% without intention-to-treat analysis

The summary of outcome measures is shown in Table 1.

Table . Summary of outcome indicators.

Specific tests		Assessment of indicators
Postural stability test		
	Single-leg standing test	AP[a] displacement[b], ML[c] displacement[d], and BBS[e] score
	2-legged tandem standing test	AP displacement, ML displacement, and BBS score
	2-legged standing test	AP displacement, ML displacement, and BBS score
Fall risk prediction		
	Functional test metrics	TUG[f] time
	Risk prediction indicators	Fall risk (time/year)

[a]AP: anterior-to-posterior.

[b]Center of pressure displacement in the anterior-to-posterior.

[c]ML: medio-to-lateral.

[d]Center of pressure displacement in the medio-to-lateral.

[e]BBS: Berg Balance Scale.

[f]TUG: timed up and go.

Two reviewers independently screened the titles, abstracts, and full texts of the identified studies. Any disagreements were resolved through consultation with a third adjudicator, who made the final determination based on the predefined criteria. Relevant data, including study design, population characteristics, sample size, intervention details, and outcome measures, were extracted into a standardized Microsoft Excel 2024 (Microsoft Corporation) form.

Quality and Risk of Bias Assessment

Quality assessment was performed using the Cochrane risk-of-bias tool (Cochrane), which consists of seven domains: (1) random sequence generation, (2) allocation concealment, (3) blinding of experimental staff, (4) blinding of results staff, (5) data integrity, (6) selective reporting of results, and (7) other sources of bias. For each source of bias, studies were classified as low, high, or unclear (if reporting is insufficient to assess a specific domain) risk. Two independent reviewers (CG and YW) assessed the results. In case of disagreement, differences

XSL•FO
RenderX

of opinion were discussed and resolved with a third researcher (MX).

Data Extraction

Literature screening and data extraction were carried out using an independent double-blind method. Two researchers, CG and YW, first screened the literature according to the titles and abstracts of the literature, excluded obviously irrelevant studies, then read the full text of the potentially relevant literature, and determined the final selection of the literature according to the inclusion and exclusion criteria. In case of disputes during the screening process, a third researcher (MX) was involved in the discussion and decided. Data extracted included the first author, year of publication, sample size, age of subjects, intervention plan and specific protocol, outcome measures, and major findings. If the original text did not contain the target data or the information was ambiguous, an email was sent to contact the corresponding author for assistance or confirmation.

Statistical Analysis

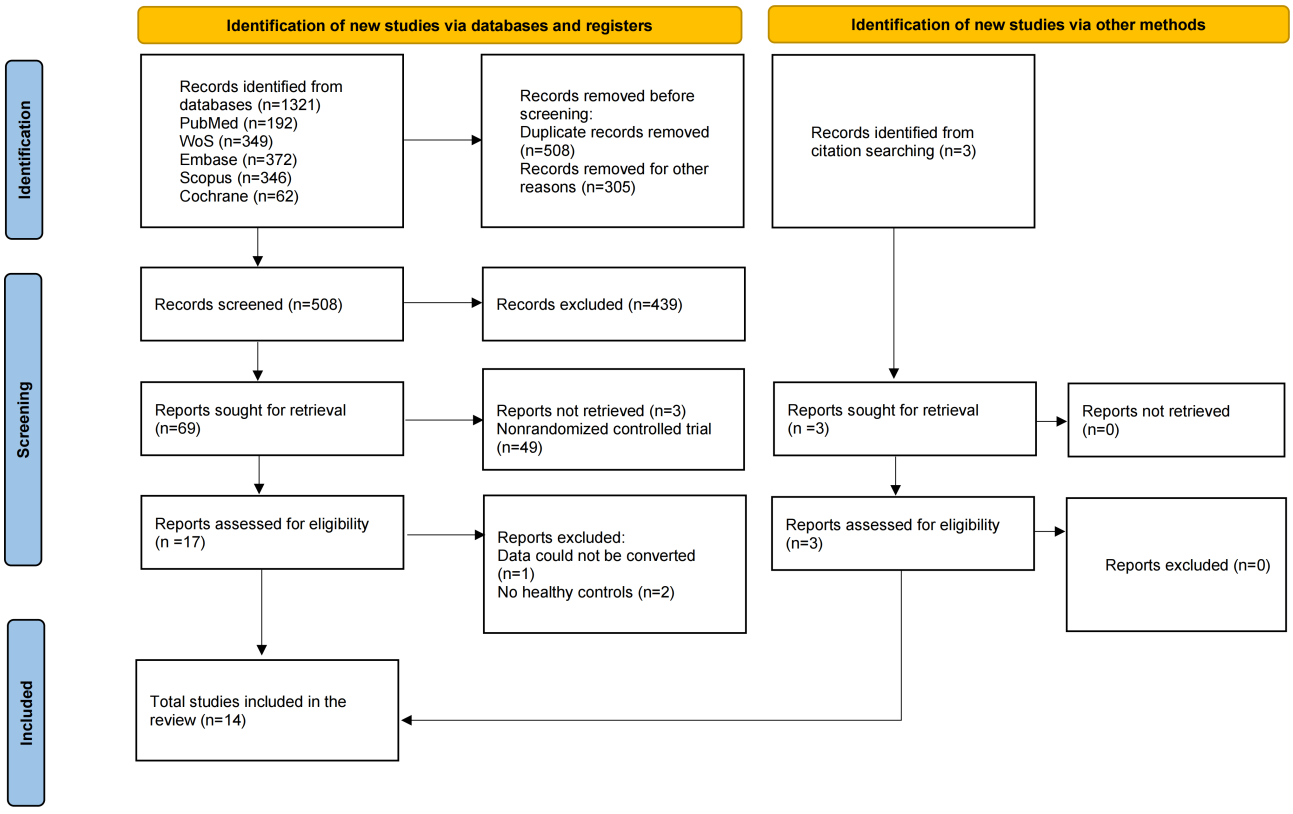
Meta-analyses were performed using the meta and metafor packages in the statistical software R (version 4.2.0; R Foundation for Statistical Computing). The effect size for continuous outcomes was expressed as the standardized mean difference (SMD), specifically Hedges g to correct for small-sample bias. Effect magnitudes were interpreted using Cohen conventions (0.2=small, 0.5=medium, and ≥0.8=large). This study adopted "95% CI not including zero" as the criterion for determining statistically significant differences between groups. The heterogeneity between studies was assessed by the I^2 statistic. When $I^2 \leq 50\%$, the heterogeneity between studies was considered low, and meta-analysis was conducted using the fixed-effects model; when $I^2 > 50\%$, the heterogeneity between studies was considered high, and meta-analysis was conducted using the random-effects model. Age stratification was based on the age distribution of the study population, dividing participants into 60 to 70 years and ≥70 years groups [34]. Intervention duration was categorized according to the median study period and the stabilization phase of multisensory training effects [35]. Intervention types were classified based on the core modalities (dynamic, static, or combined dynamic-static) used in the included studies, combined with clinical classification standards for balance training. Additionally, regression analysis was conducted to explore key findings of this study. Meta-regression models were fitted within a random-effects framework, and the significance of regression coefficients for each variable was tested to determine whether statistical differences existed in effect sizes between different intervention types. Sensitivity analyses were conducted by excluding the included studies one by one to assess the effect of individual studies on the overall results and to judge the reliability of the results. Funnel plots were used to assess publication bias, and the results were interpreted with caution if publication bias existed, and further unpublished studies were searched or statistically corrected if necessary.

Results

Study Selection and Characteristics

A total of 1321 records were obtained through systematic search of multiple databases. After deleting duplicates, screening titles and abstracts, and reading the full text, 14 studies were finally included. The specific screening process is shown in Figure 1.

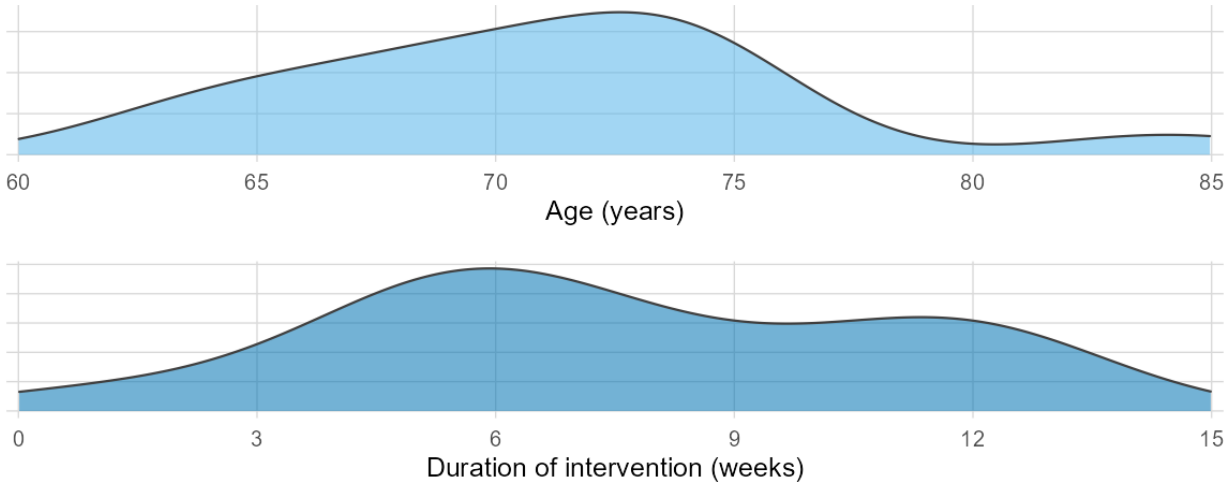
Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the study selection process.



XSL•FO
RenderX

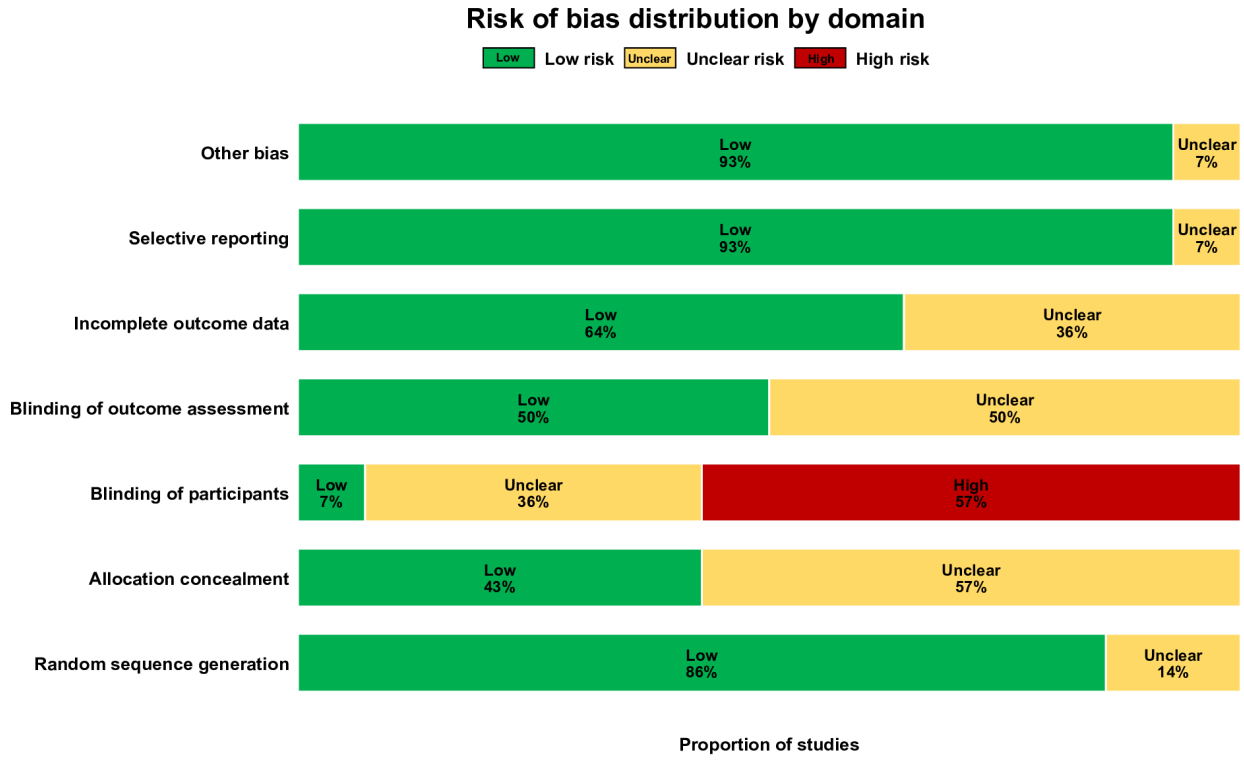
Included studies were published from 2010 to 2025, with a total sample size of 1321 cases, and the type of study was randomized controlled trial. The outcome metrics covered anterior-to-posterior (AP) displacement (5 items) [11,35-38], medio-to-lateral (ML) displacement (4 items) [11,35,36,38], BBS score (3 items) [39-41], TUG (7 items) [35,39,41-45], fall risk (4 items) [37,45-47], and the basic characteristics of each study are shown in Multimedia Appendix 1. Among the included studies, the age of the experimental group was mainly concentrated in 2 intervals, 60 to 70 years and >70 years, and the intervention time was mainly distributed in the period of <6 weeks and >6 weeks. The types of interventions were mainly dynamic balance training, static balance training, and combined dynamic and static balance training, and the distribution is shown in Figure 2.

Figure 2. Distribution maps of age and intervention duration.



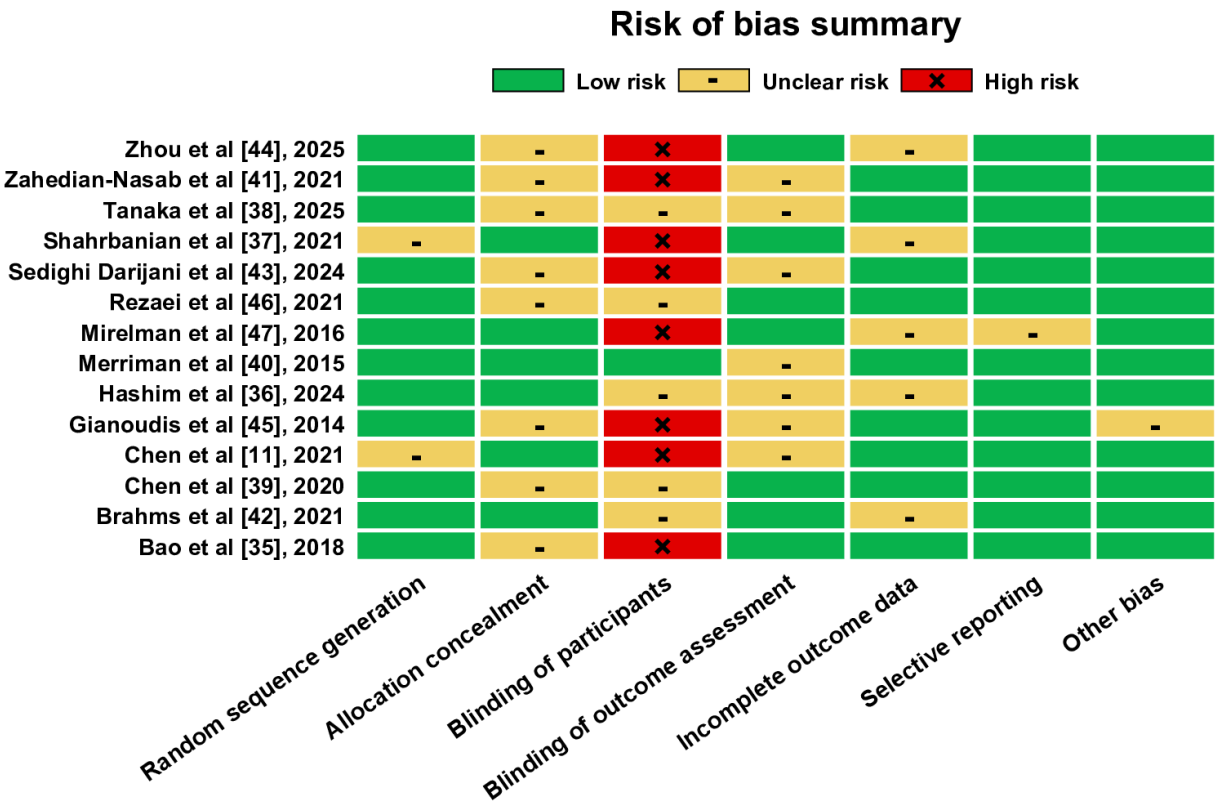
Of the 14 included studies, 12 [35,36,38-47] described specific methods of random sequence generation, 6 [11,36,37,40,42,47] described allocation concealment, and 7 [35,37,39,42,44,46,47] were single-blind studies. However, 9 studies [11,35,38-41,43,45,46] reported detachment without specifying the reason, as shown in Figures 3 and 4.

Figure 3. Distribution of risk of bias across assessment domains.



XSL•FO
RenderX

Figure 4. Results of the risk of bias assessment of the included studies.



Main Results

MSI training significantly improved AP displacement (SMD −1.64, 95% CI −2.78 to −0.49; I^2=83%, P<.001), and ML displacement showed a significant reduction (SMD −1.37, 95% CI −2.68 to −0.07; I^2=83%, P<.001). In subgroup analyses, age and intervention duration showed no significant moderating effects on AP or ML displacement. In contrast, the type of intervention significantly moderates AP displacement (P=.009) but not ML displacement. Additionally, we conducted regression analyses using study-level data points to validate these findings and to illustrate the observed distribution of effects across studies. Sensitivity analysis by excluding studies one by one revealed that the pooled effect size of AP displacement fluctuated between −3.41 and −1.45 and that of ML displacement fluctuated between −3.83 and −0.4. All effect sizes after exclusion iterations remained statistically significant (95% CIs did not include 0), indicating that the main results were robust, as shown in Figure 5, Figure 6, Multimedia Appendix 1, and Table 2.

XSL•FO
RenderX

Figure 5. Bubble chart of the regression analysis of anterior-to-posterior (AP) displacement based on Bao et al [35], Chen et al [11], Hashim et al [36], Shahrbanian et al [37], and Tanaka et al [38].

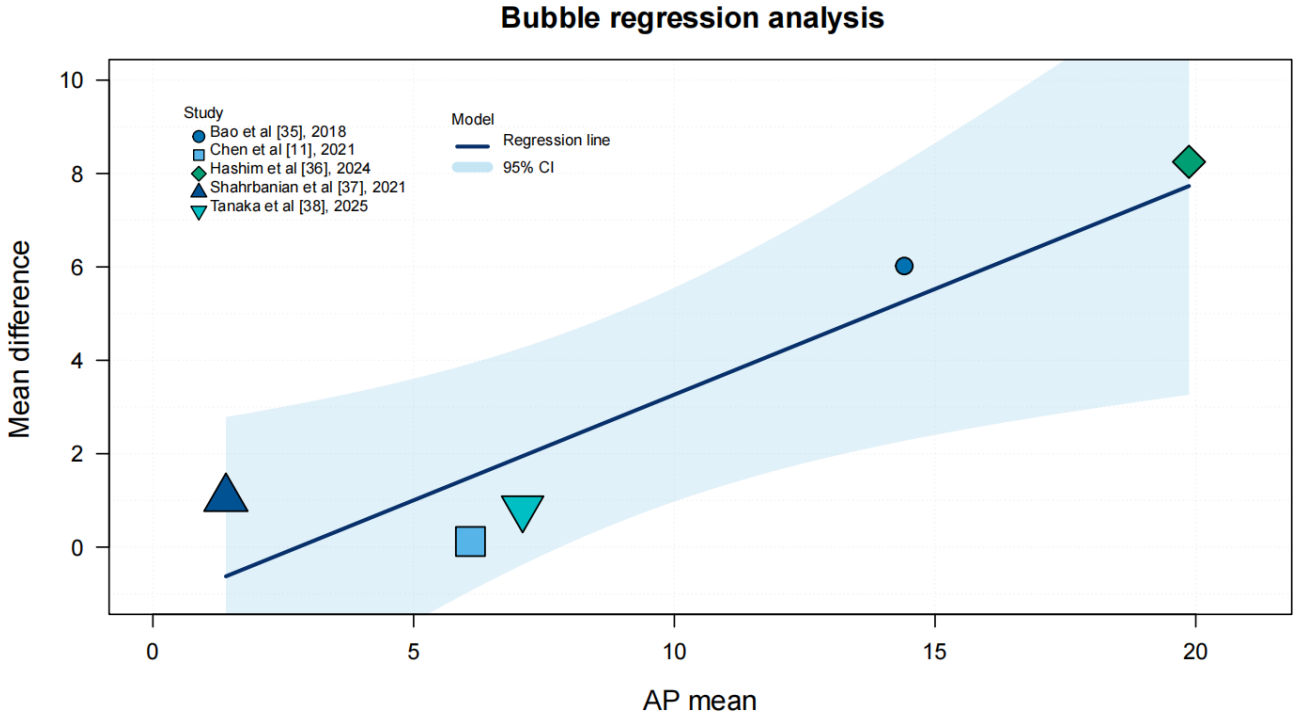
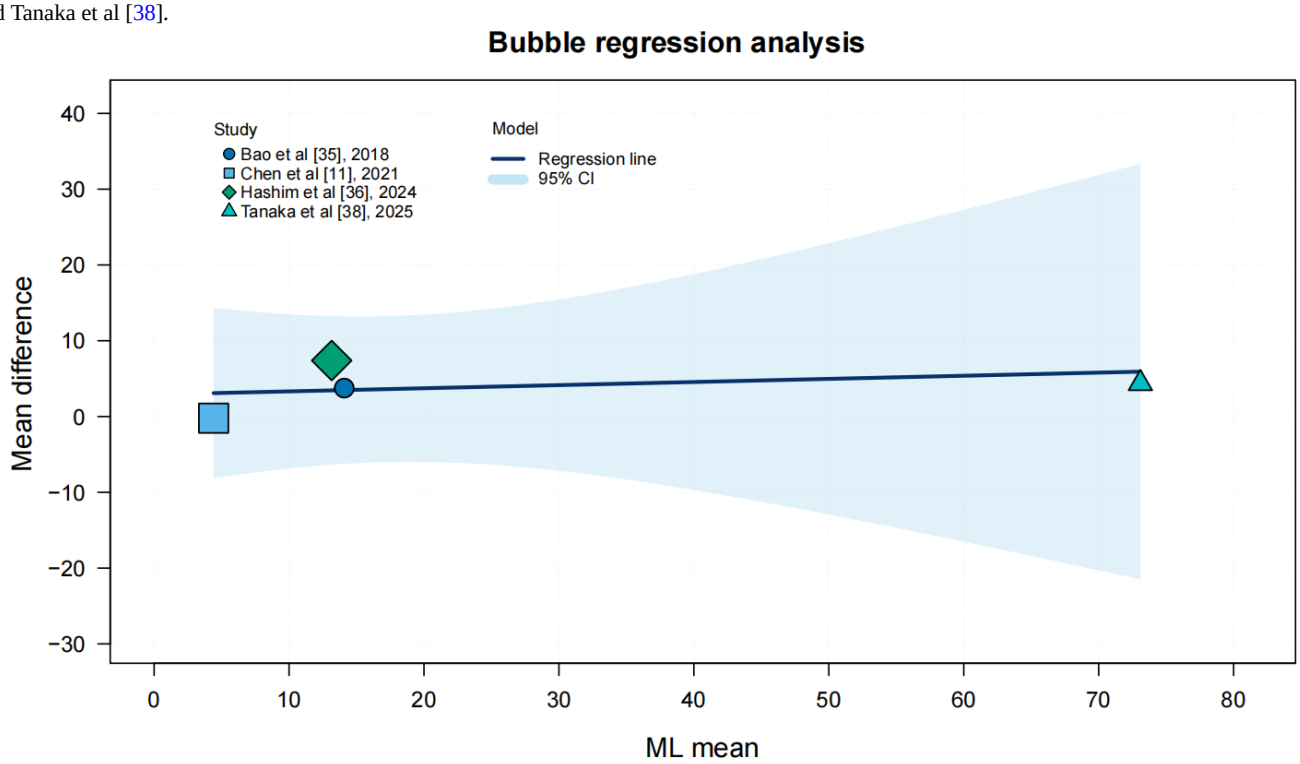


Figure 6. Bubble chart of the regression analysis of medio-to-lateral (ML) displacement based on Bao et al [35], Chen et al [11], Hashim et al [36], and Tanaka et al [38].



XSL•FO
RenderX

Table . Summary of the results of the included studies.

Outcomes	Studies		I^2 statistic, %	Subgroup, *P* value			Sensitivity analysis
	Reference (n)	SMD[a] (95% CI)		Age (years)	Duration of intervention	Type of intervention	
AP[b] displacement[c]	5 [11,35-38]	−1.64 (−2.78 to −0.49)	83	.13	.09	.009	Stable
ML[d] displacement[e]	4 [11,35,36,38]	−1.37 (−2.68 to −0.07)	83	.09	.58	.09	Stable
TUG[f] time	7 [35,39,41-45]	−1.43 (−2.36 to −0.50)	89	.26	.24	*<.001*[g]	Stable
BBS[h] score	3 [39-41]	3.42 (2.41 to 4.44)	0	N/A[i]	N/A	N/A	N/A
Fall risk	4 [37,45-47]	−1.27 (−2.03 to −0.52)	99	.07	.06	.001	Stable

[a]SMD: standardized mean difference.

[b]AP: anterior-to-posterior.

[c]Center of pressure displacement in the anterior-to-posterior.

[d]ML: medio-to-lateral.

[e]Center of pressure displacement in the medio-to-lateral.

[f]TUG: timed up and go.

[g]Statistically significant.

[h]BBS, Berg Balance Scale.

[i]N/A: no subgroup or sensitivity analysis was carried out due to insufficient data.

Secondary Results

Effect of MSI Training on Postural Stability in Older Adults

There was low heterogeneity in the studies of BBS scores (SMD 3.42, 95% CI 2.41 to 4.44; I^2=0%, *P*=.006), suggesting that the MSI training significantly improved BBS scores of the older adults and had reliable effects on the comprehensive quantitative static and dynamic balance enhancement, with a high degree of consistency among different studies. The results of the sensitivity analyses were not stable, and the funnel plot did not show obvious symmetry, suggesting that the results of the studies were influenced by individual studies, as shown in Multimedia Appendix 1 and Table 2.

Effect of MSI Training on Fall Risk in Older Adults

High heterogeneity was present in the studies of time to TUG completion (SMD −1.43, 95% CI −2.36 to −0.50; I^2=89%, *P*<.001). Subgroup analysis by intervention type was statistically significant (*P*<.001), and sensitivity analysis results were stable. The funnel plot analysis showed good symmetry, suggesting that the results were relatively reliable. Fall risk–related studies had extremely high heterogeneity (SMD −1.27, 95% CI −2.03 to −0.52; I^2=99%, *P*<.001). Subgroup analyses by intervention type was significant (P=.001), whereas age and intervention duration were not significant and the results of sensitivity analyses were stable, but the funnel plot showed asymmetry, as shown in Multimedia Appendix 1 and Table 2.

Discussion

This systematic review integrated data from 14 randomized controlled trials and demonstrated that MSI training significantly improved postural stability and reduced fall risk in older adults. Specifically, MSI training led to reductions in AP and ML displacement. However, high heterogeneity was observed in certain outcomes, which may be attributed to methodological variations, participant diversity, and differences in intervention protocols.

Effect of MSI Training on Postural Stability

Although MSI training significantly improved both AP and ML displacements, the high heterogeneity limits the generalizability of these findings. In terms of the intervention protocols, the training methods used in different studies varied greatly. Bao et al [35] used vibro-tactile sensory enhancement interventions, whereas Chen et al [11] used stroboscopic training in balance training; however, meta-regression and subgroup analyses did not identify significant moderating effects of age, intervention duration, or training type on these outcomes. For AP displacement, intervention type showed a significant moderating effect, whereas age and intervention duration did not. For ML displacement, no significant moderating effects were observed, suggesting that other moderating factors (eg, baseline balance ability or cognitive status) may be at play [13,22]. Future studies should standardize intervention protocols to better elucidate the mechanisms by which different training modalities influence balance performance in older adults.

XSL•FO
RenderX

Notably, BBS scores showed consistent improvement with low heterogeneity, indicating that MSI training effectively enhances overall balance control. This aligns with the concept that MSI training promotes central sensory recalibration and optimizes sensorimotor integration. This was attributed to the enhancement of the integration efficiency of the central nervous system by training, which optimized the distribution of sensory weights and thus improved postural control strategies. In a study by Eikema et al [48], it was found that older adults had a more pronounced decline in central sensory integration and a longer effective adaptive world than younger adults when faced with a challenging task. This finding provides an important reference for the development of stratified and individualized interventions in clinical practice based on baseline sensory function and age characteristics of older adults.

Effect of MSI Training on Fall Risk

There is a moderate heterogeneity in the study of the time of completion of the TUG test, and the results of the analysis of the subgroup of the intervention type are significant, which suggests that there is a significant difference in the effects of the different types of interventions on the ability of the older adults to rise and walk, which is clinically important for assessing the effect of training on functional mobility improvement. The results of the subgroup analyses were significant, indicating that the effects of different intervention types on the ability to get up and walk differed significantly, which is clinically important for evaluating the effects of training on the functional mobility of older adults. In addition, the overall heterogeneity of the studies on fall risk was extremely high, and most of the studies had small sample sizes, which resulted in insufficient statistical validity and limited the reliability and extrapolation of the results.

Mechanism of Action of MSI Training

With advancing age, balance and postural control become increasingly vision dependent, as neurodegenerative processes impair vestibular and somatosensory function. Additionally, concomitant neuronal loss and pathway damage further hinder afferent signaling [8,49,50]. MSI training is designed to mitigate these age-related deficits by jointly targeting peripheral sensory function, central integrative efficiency, and neuromuscular coordination [4,51-53]. At the peripheral level, structured visual exercises, such as object recognition under varied illumination and complex backgrounds, have been reported to improve visual acuity and the perception of shape, displacement, and motion [51,52]. Graded vestibular stimuli, including slow head rotations on stable surfaces and gentle tilting or rotation using dedicated apparatus, enhance responsiveness to head and trunk movements [54,55], whereas somatosensory tasks requiring precise limb positioning and force regulation strengthen proprioceptive feedback and movement accuracy [4,53]. These peripheral adaptations are accompanied by central reweighting of multimodal inputs: exposure to diverse sensory environments fosters more efficient integration of visual, vestibular, and somatosensory information [16,51], thereby accelerating postural adjustments and improving stability during dynamic tasks such as walking in simulated complex settings [56-58].

Concurrently, MSI training refines neuromuscular coordination by tightening the coupling between neural commands and muscular execution, manifested as smoother signal transmission, more accurate and rapid muscle responses, and reduced activation latency [56,59-61]. Bugnariu et al [62] reported that older adults exhibited 40 to 60 ms longer electromyographic latencies in the distal-to-proximal activation sequence compared to younger adults. Notably, a single 1-hour MSI session reduced ankle muscle latency by approximately 20 ms, suggesting improved neuromuscular timing [63]. Beyond cortical contributions, MSI training engages the cerebellum, basal ganglia, and other motor control centers: cerebellar metabolic activity increases during training, facilitating consolidation of motor skills, while the basal ganglia modulate initiation and termination of motor programs to enhance movement smoothness and precision [64]. Consistent with this systems-level account, Wang et al [17] reported less adaptive brain connectivity in older vs younger adults and demonstrated that MSI training can augment functional connectivity and neural adaptability, supporting structural and functional central adaptation and establishing a robust neuromodulator network for motor improvement [65]. Taken together, by strengthening sensory inputs, optimizing central integration, and improving neuromuscular coordination, MSI training constitutes a comprehensive and effective approach to enhancing postural stability in older adults, thereby reducing fall risk and improving quality of life [37].

Clinical Significance and Future Prospects of the Study

This study shows that MSI training is important for improving postural stability and reducing fall risk in older adults and provides a theoretical basis and practical guidance for clinical rehabilitation and geriatric care. To enhance the reliability and practicality of the results, future studies should expand the sample size and conduct multicenter studies to include older adults with different health conditions to verify the generalizability of the effect and continue to follow-up and evaluate various indicators to investigate the long-term effect, the best training mode, and the applicability to older adults with different health conditions, and the mechanism of action can be analyzed by combining neuroimaging and muscle electrophysiology techniques. We will continue to follow-up and evaluate various indicators to investigate the long-term effect, the optimal mode, and the applicability to different health conditions of older adults, to provide more accurate and effective interventions.

Comparison With Other Studies

Previous studies [22,66] have confirmed that although single balance or strength training can improve static postural stability, the lack of multisensory conflict simulation makes it difficult to transfer motor adaptations to dynamic multitasking scenarios and fails to effectively reduce the risk of falls. This issue was broken through in 2 systematic review studies [27,28] that used qualitative research to propose that MSI training significantly improves postural control, a finding that is highly compatible with the results of this study.

This study realizes dual innovations at the methodological and outcome levels. In terms of methodology, we broke through the

XSL•FO
RenderX

limitations of previous qualitative studies and adopted a combination of qualitative and quantitative analysis to construct a more rigorous research system by integrating multidimensional data. In contrast to the study by Sápi et al [67], which concluded that MSI training could not improve the COP displacement in the AP direction in older adults, this study, by incorporating the dynamic balance task and revamping the assessment scenarios, confirms that MSI training has a significant improvement in the COP displacement in the AP and ML directions, which is more in line with the real functional needs of older adults.

The innovative value of this study lies in the fact that it is the first quantitative study to systematically demonstrate the effectiveness of MSI training on the optimization of omnidirectional postural control and the reduction of fall risk, which not only provides a quantitative theoretical basis for the prevention of falls in older adults but also has a guiding significance for clinical practice, suggesting that individual functional shortcomings need to be based on individual design of differentiated MSI training programs integrating static and dynamic tasks, so as to accurately improve the motor function and quality of life of older adults. This finding not only provides a quantitative theoretical basis for fall prevention but also guides clinical practice.

Study Quality and Limitations

Several limitations should be considered when interpreting our findings. First, evidence on long-term effectiveness remains limited, as only 3 trials provided follow-up data, and their high heterogeneity precluded meta-analysis. Second, significant methodological and clinical heterogeneity, including variations in randomization, participant characteristics, intervention protocols, and outcome measures, may affect comparability across studies, though we addressed this through random-effects models and subgroup analyses. Third, most trials had small sample sizes, which could overestimate effect sizes and reduce statistical power; sensitivity analyses were performed to ensure robustness. Finally, potential publication bias, as indicated by funnel plot asymmetry, cannot be ruled out despite comprehensive search efforts. Future studies should adopt standardized interventions and assessments, include longer follow-up periods, and explore moderating effects of individual characteristics to strengthen the evidence base for clinical practice.

Conclusion

MSI training can improve postural stability and reduce fall risk in older adults, but we need to conduct long-term follow-up of multicenter and large-sample studies to identify the best intervention mode. Meanwhile, standardization of measurement tools, optimization of intervention protocols, attention to individual differences, and evaluation of long-term effects will help to improve the comparability of studies and provide more accurate guidance for clinical practice.

Funding

This work was supported by the Eastern Talent Plan Leading Project under grant BJJY2025064. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request. All relevant data are also provided in Multimedia Appendix 1 [11,35-47].

Authors' Contributions

Conceptualization: CG, LY, PC
Data curation: YW, CG, MX, JZ, ZY
Formal analysis: CG, LY, PC
Investigation: PC, MX, JZ
Methodology: CG, LY, PC
Project administration: LW
Supervision: LW
Validation: YW, CG, MX, JZ, ZY
Visualization: TCT, YW, YG, MX, VML
Writing—original draft: CG
Writing—review and editing: LY, PC, JZ, ZY, TCT, YW, YG, MX, VML, LW

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials, including search strategies, characteristics of included studies, additional analyses, figures, statistical code, dataset, and references.
[DOCX File, 2959 KB - aging_v9i1e80345_app1.docx]

XSL•FO
RenderX

Checklist 1
PRISMA checklist.
[DOCX File, 24 KB - aging_v9i1e80345_app2.docx]

References

1. WHO global report on falls prevention in older age. : World Health Organization; 2008 URL: https://www.who.int/publications/i/item/9789241563536 [accessed 2026-04-28]
2. Goetzl EJ, Feeley B. Musculoskeletal disorders of the elderly. Am J Med 2025 Sep;138(9):1197-1200. [doi: 10.1016/j.amjmed.2025.04.006] [Medline: 40228599]
3. Huang C, He X. Association between sensory impairment and sarcopenia in older Chinese adults: a 4-year longitudinal study. BMC Geriatr 2025 Feb 11;25(1):90. [doi: 10.1186/s12877-024-05642-6] [Medline: 39934655]
4. Mozolic JL, Hugenschmidt CE, Peiffer AM, Laurienti PJ. Multisensory integration and aging. In: Murray MM, Wallace MT, editors. The Neural Bases of Multisensory Processes: CRC Press; 2012.
5. Hirst RJ, Setti A, De Looze C, Kenny RA, Newell FN. Multisensory integration precision is associated with better cognitive performance over time in older adults: a large-scale exploratory study. Aging Brain 2022;2:100038. [doi: 10.1016/j.nbas.2022.100038] [Medline: 36908873]
6. Yin L, Xu X, Wang R, Li F, Wang Y, Wang L. Validity and reliability of inertial measurement units on gait, static balance and functional mobility performance among community-dwelling older adults: a systematic review and meta-analysis. EFORT Open Rev 2025 Apr 1;10(4):172-185. [doi: 10.1530/EOR-2024-0088] [Medline: 40167464]
7. Tsai YY, Chen YC, Zhao CG, Hwang IS. Adaptations of postural sway dynamics and cortical response to unstable stance with stroboscopic vision in older adults. Front Physiol 2022;13:919184. [doi: 10.3389/fphys.2022.919184] [Medline: 36105297]
8. Chen YC, Huang CC, Zhao CG, Hwang IS. Visual effect on brain connectome that scales feedforward and feedback processes of aged postural system during unstable stance. Front Aging Neurosci 2021;13:679412. [doi: 10.3389/fnagi.2021.679412] [Medline: 34366825]
9. Hu X, Zeng Z, Tang M, Wang L. Age-related changes in plantar sensation and ankle proprioception in adolescents to older adults. Motor Control 2023;27(3):596-615. [doi: 10.1123/mc.2022-0072] [Medline: 36990443]
10. Seidler RD, Bernard JA, Burutolu TB, et al. Motor control and aging: links to age-related brain structural, functional, and biochemical effects. Neurosci Biobehav Rev 2010 Apr;34(5):721-733. [doi: 10.1016/j.neubiorev.2009.10.005] [Medline: 19850077]
11. Chen YC, Chou YC, Hwang IS. Reliance on visual input for balance skill transfer in older adults: EEG connectome analysis using minimal spanning tree. Front Aging Neurosci 2021;13:632553. [doi: 10.3389/fnagi.2021.632553] [Medline: 33613272]
12. Semyanov A, Verkhratsky A. Neuroglia in aging. Handb Clin Neurol 2025;209:49-67. [doi: 10.1016/B978-0-443-19104-6.00002-4] [Medline: 40122631]
13. Lesinski M, Hortobágyi T, Muehlbauer T, Gollhofer A, Granacher U. Effects of balance training on balance performance in healthy older adults: a systematic review and meta-analysis. Sports Med 2015 Dec;45(12):1721-1738. [doi: 10.1007/s40279-015-0375-y] [Medline: 26325622]
14. Chen YC, Tsai YY, Huang WM, Zhao CG, Hwang IS. Cross-frequency modulation of postural fluctuations and scalp EEG in older adults: error amplification feedback for rapid balance adjustments. Geroscience 2024 Dec;46(6):5599-5613. [doi: 10.1007/s11357-024-01258-1] [Medline: 38910193]
15. Neville H, Bavelier D. Human brain plasticity: evidence from sensory deprivation and altered language experience. Prog Brain Res 2002;138:177-188. [doi: 10.1016/S0079-6123(02)38078-6] [Medline: 12432770]
16. Bolam J, Diaz JA, Andrews M, et al. A drift diffusion model analysis of age-related impact on multisensory decision-making processes. Sci Rep 2024 Jun 28;14(1):14895. [doi: 10.1038/s41598-024-65549-5] [Medline: 38942761]
17. Wang G, Yang Y, Liu X, et al. Neural connectivity and balance control in aging: insights from directed cortical networks during sensory conflict. Neuroimage 2025 May 15;312:121218. [doi: 10.1016/j.neuroimage.2025.121218] [Medline: 40239853]
18. Gerber ED, Huang CK, Moon S, Devos H, Luchies CW. Sensory reweighting of postural control requires distinct rambling and trembling sway adaptations. Gait Posture 2024 Jul;112:16-21. [doi: 10.1016/j.gaitpost.2024.04.028] [Medline: 38723391]
19. Cui J, Hao Z, Tian H, Yang Y, Wang J, Lin X. The effects of Tai Chi on standing balance control in older adults may be attributed to the improvement of sensory reweighting and complexity rather than reduced sway velocity or amplitude. Front Aging Neurosci 2024;16:1330063. [doi: 10.3389/fnagi.2024.1330063] [Medline: 38650868]
20. Sung P, Rowland P. Impact of sensory reweighting strategies on postural control using the sensory organization test in older adults with and without fall risks. Physiother Res Int 2024 Apr;29(2):e2075. [doi: 10.1002/pri.2075] [Medline: 38430540]
21. Pšeničnik Sluga S, Kozinc Z. Sensorimotor and proprioceptive exercise programs to improve balance in older adults: a systematic review with meta-analysis. Eur J Transl Myol 2024 Jan 11;34(1):12010. [doi: 10.4081/ejtm.2024.12010] [Medline: 38213185]

XSL•FO
RenderX

22. Lacroix A, Hortobágyi T, Beurskens R, Granacher U. Effects of supervised vs. unsupervised training programs on balance and muscle strength in older adults: a systematic review and meta-analysis. Sports Med 2017 Nov;47(11):2341-2361. [doi: 10.1007/s40279-017-0747-6] [Medline: 28573401]
23. Shih R, Shenvi C. Evaluation of older adults in the emergency department following a fall. Emerg Med Clin North Am 2025 May;43(2):189-198. [doi: 10.1016/j.emc.2024.08.003] [Medline: 40210340]
24. Allison LK, Kiemel T, Jeka JJ. Sensory-challenge balance exercises improve multisensory reweighting in fall-prone older adults. J Neurol Phys Ther 2018 Apr;42(2):84-93. [doi: 10.1097/NPT.0000000000000214] [Medline: 29547483]
25. O'Dowd A, Hirst RJ, Setti A, Donoghue OA, Kenny RA, Newell FN. The temporal precision of audiovisual integration is associated with longitudinal fall incidents but not sensorimotor fall risk in older adults. Sci Rep 2023 May 3;13(1):7167. [doi: 10.1038/s41598-023-32404-y] [Medline: 37137879]
26. Yin L, Nam H, Wei Y, et al. Gait and balance metrics comparison among different fall risk groups and principal component analysis for fall prediction in older people. Age Ageing 2025 Mar 28;54(4):afaf076. [doi: 10.1093/ageing/afaf076] [Medline: 40197782]
27. Zhang S, Xu W, Zhu Y, Tian E, Kong W. Impaired multisensory integration predisposes the elderly people to fall: a systematic review. Front Neurosci 2020;14:411. [doi: 10.3389/fnins.2020.00411] [Medline: 32410958]
28. de Dieuleveult AL, Siemonsma PC, van Erp JB, Brouwer AM. Effects of aging in multisensory integration: a systematic review. Front Aging Neurosci 2017;9:80. [doi: 10.3389/fnagi.2017.00080] [Medline: 28400727]
29. Cheng M, Ni J, Liu F, Wang S. Optimal type and dose of exercise to improve fall behavior in older adults: a systematic evaluation and network meta-analysis. Ageing Res Rev 2026 Jan;113:102924. [doi: 10.1016/j.arr.2025.102924] [Medline: 41110662]
30. Yu H, Zhong J, Li M, Chen S. Effects of exercise intervention on falls and balance function in older adults: a systematic review and meta-analysis. PeerJ 2025;13:e20190. [doi: 10.7717/peerj.20190] [Medline: 41122235]
31. Hernandez-Martinez J, Cid-Calfucura I, Vásquez-Carrasco E, Branco BH, Herrera-Valenzuela T, Valdés-Badilla P. Benefits of exergaming regarding to conventional physical therapies on balance and fall risk in prefrail and frail older people: a meta-analysis of randomized controlled trials. Inquiry 2025;62:469580251372362. [doi: 10.1177/00469580251372362] [Medline: 40956936]
32. Wei X, Huang C, Ding X, et al. Effect of virtual reality training on dual-task performance in older adults: a systematic review and meta-analysis. J Neuroeng Rehabil 2025 Jun 24;22(1):141. [doi: 10.1186/s12984-025-01675-z] [Medline: 40551196]
33. Wang D, Pang X, Shen P, Mao D, Song Q. Effectiveness of various exercise types in reducing fall risk among older adults with diabetic peripheral neuropathy: a systematic review and meta-analysis. J Exerc Sci Fit 2025 Jul;23(3):157-166. [doi: 10.1016/j.jesf.2025.03.005] [Medline: 40242133]
34. Shaffer SW, Harrison AL. Aging of the somatosensory system: a translational perspective. Phys Ther 2007 Feb;87(2):193-207. [doi: 10.2522/ptj.20060083] [Medline: 17244695]
35. Bao T, Carender WJ, Kinnaird C, et al. Effects of long-term balance training with vibrotactile sensory augmentation among community-dwelling healthy older adults: a randomized preliminary study. J Neuroeng Rehabil 2018 Jan 18;15(1):5. [doi: 10.1186/s12984-017-0339-6] [Medline: 29347946]
36. Hashim J, Campelo AM, Doyle-Baker PK, Cossich VRA, Katz L. Comparison between six-week exergaming, conventional balance and no exercise training program on older adults' balance and gait speed. Int J Comput Sci Sport 2024 Feb 1;23(1):93-109. [doi: 10.2478/ijcss-2024-0006]
37. Shahrbanian S, Hashemi A, Hemayattalab R. The comparison of the effects of physical activity and neurofeedback training on postural stability and risk of fall in elderly women: a single-blind randomized controlled trial. Physiother Theory Pract 2021 Feb;37(2):271-278. [doi: 10.1080/09593985.2019.1630877] [Medline: 31218913]
38. Tanaka T, Maeda Y, Miura T. Effects of tactile sensory stimulation training of the trunk and sole on standing balance ability in older adults: a randomized controlled trial. J Funct Morphol Kinesiol 2025 Mar 17;10(1):96. [doi: 10.3390/jfmk10010096] [Medline: 40137348]
39. Chen PJ, Penn IW, Wei SH, Chuang LR, Sung WH. Augmented reality-assisted training with selected Tai-Chi movements improves balance control and increases lower limb muscle strength in older adults: a prospective randomized trial. J Exerc Sci Fit 2020 Sep;18(3):142-147. [doi: 10.1016/j.jesf.2020.05.003] [Medline: 32514277]
40. Merriman NA, Whyatt C, Setti A, Craig C, Newell FN. Successful balance training is associated with improved multisensory function in fall-prone older adults. Comput Human Behav 2015 Apr;45:192-203. [doi: 10.1016/j.chb.2014.12.017]
41. Zahedian-Nasab N, Jaberi A, Shirazi F, Kavousipor S. Effect of virtual reality exercises on balance and fall in elderly people with fall risk: a randomized controlled trial. BMC Geriatr 2021 Sep 25;21(1):509. [doi: 10.1186/s12877-021-02462-w] [Medline: 34563120]
42. Brahms M, Heinzel S, Rapp M, et al. Cognitive-postural multitasking training in older adults - effects of input-output modality mappings on cognitive performance and postural control. J Cogn 2021 Mar 10;4(1):20. [doi: 10.5334/joc.146] [Medline: 33748665]

XSL•FO
RenderX

43. Sedighi Darijani S, Sahebozamani M, Eslami M, Babakhanian S, Alimoradi M, Iranmanesh M. The effect of neurofeedback and somatosensory exercises on balance and physical performance of older adults: a parallel single-blinded randomized controlled trial. Sci Rep 2024 Oct 15;14(1):24087. [doi: 10.1038/s41598-024-74980-7] [Medline: 39407027]
44. Zhou J, Liu B, Xu JF, et al. Home-based strength and balance exercises for fall prevention among older individuals of advanced age: a randomized controlled single-blind study. Ann Med 2025 Dec;57(1):2459818. [doi: 10.1080/07853890.2025.2459818] [Medline: 39918027]
45. Gianoudis J, Bailey CA, Ebeling PR, et al. Effects of a targeted multimodal exercise program incorporating high-speed power training on falls and fracture risk factors in older adults: a community-based randomized controlled trial. J Bone Miner Res 2014 Jan;29(1):182-191. [doi: 10.1002/jbmr.2014] [Medline: 23775701]
46. Rezaei K, Nami M, Sinaei E, Bagheri Z, Kordi Yoosefinejad A. A comparison between effects of neurofeedback and balance exercise on balance of healthy older adults. J Biomed Phys Eng 2021 Dec;11(6):713-722. [doi: 10.31661/jbpe.v0i0.1203] [Medline: 34904068]
47. Mirelman A, Rochester L, Maidan I, et al. Addition of a non-immersive virtual reality component to treadmill training to reduce fall risk in older adults (V-TIME): a randomised controlled trial. Lancet 2016 Sep 17;388(10050):1170-1182. [doi: 10.1016/S0140-6736(16)31325-3] [Medline: 27524393]
48. Eikema DJ, Hatzitaki V, Tzovaras D, Papaxanthis C. Age-dependent modulation of sensory reweighting for controlling posture in a dynamic virtual environment. Age (Dordr) 2012 Dec;34(6):1381-1392. [doi: 10.1007/s11357-011-9310-9] [Medline: 21894445]
49. Cassady K, Gagnon H, Lalwani P, et al. Sensorimotor network segregation declines with age and is linked to GABA and to sensorimotor performance. Neuroimage 2019 Feb 1;186:234-244. [doi: 10.1016/j.neuroimage.2018.11.008] [Medline: 30414983]
50. Yuasa-Kawada J, Kinoshita-Kawada M, Hiramoto M, et al. Neuronal guidance signaling in neurodegenerative diseases: key regulators that function at neuron-glia and neuroimmune interfaces. Neural Regen Res 2026 Feb 1;21(2):612-635. [doi: 10.4103/NRR.NRR-D-24-01330] [Medline: 39995079]
51. Legdeur N, Badissi M, Venkatraghavan V, et al. The temporal relation of physical function with cognition and the influence of brain health in the oldest-old. Gerontology 2025;71(1):13-27. [doi: 10.1159/000542395] [Medline: 39504937]
52. Liu Z, Wang Q, Sun W, Song Q. Balancing sensory inputs: somatosensory reweighting from proprioception to tactile sensation in maintaining postural stability among older adults with sensory deficits. Front Public Health 2023;11:1165010. [doi: 10.3389/fpubh.2023.1165010] [Medline: 37213635]
53. Schneeberger LC, Lynn A, Scarcelli V, Seif A, Stevenson RA. Enhanced multisensory gain in older adults may be a by-product of inverse effectiveness: evidence from a speeded response-time task. Psychol Aging 2024 Nov;39(7):770-780. [doi: 10.1037/pag0000850] [Medline: 39374148]
54. Reed CA, Chaudhari AM, Worthen-Chaudhari LC, Bigelow KE, Monfort SM. A new perspective on transient characteristics of quiet stance postural control. PLoS One 2020;15(8):e0237246. [doi: 10.1371/journal.pone.0237246] [Medline: 32776952]
55. Doumas M, Krampe RT. Adaptation and reintegration of proprioceptive information in young and older adults' postural control. J Neurophysiol 2010 Oct;104(4):1969-1977. [doi: 10.1152/jn.00345.2010] [Medline: 20702741]
56. Zou Z, Zhao B, Ting KH, Wong C, Hou X, Chan CC. Multisensory integration augmenting motor processes among older adults. Front Aging Neurosci 2023;15:1293479. [doi: 10.3389/fnagi.2023.1293479] [Medline: 38192281]
57. Pepper JL, Nuttall HE. Age-related changes to multisensory integration and audiovisual speech perception. Brain Sci 2023 Jul 25;13(8):1126. [doi: 10.3390/brainsci13081126] [Medline: 37626483]
58. Skirzewski M, Molotchnikoff S, Hernandez LF, Maya-Vetencourt JF. Multisensory integration: is medial prefrontal cortex signaling relevant for the treatment of higher-order visual dysfunctions? Front Mol Neurosci 2022;14:806376. [doi: 10.3389/fnmol.2021.806376] [Medline: 35110996]
59. Zhou C, Yang G, Wang Y, Zhu R, Zhu D. TaiChi-MSS protocol: enhancing cognitive and brain function in MCI patients through Tai Chi exercise combined with multisensory stimulation. Front Aging Neurosci 2025;17:1514127. [doi: 10.3389/fnagi.2025.1514127] [Medline: 40071122]
60. Chang JJ, Chen YC, Yeh SL, Tang PF, Tu CK. Impacts of audiovisual simultaneity perception on single-task and dual-task gaits in middle-aged and older adults. Gait Posture 2024 Sep;113:99-105. [doi: 10.1016/j.gaitpost.2024.06.003] [Medline: 38850854]
61. Roytman S, Paalanen N, Carli G, et al. Multisensory mechanisms of gait and balance in Parkinson's disease: an integrative review. Neural Regen Res 2025 Jan 1;20(1):82-92. [doi: 10.4103/NRR.NRR-D-23-01484] [Medline: 38767478]
62. Bugnariu N, Fung J. Aging and selective sensorimotor strategies in the regulation of upright balance. J Neuroeng Rehabil 2007 Jun 20;4:19. [doi: 10.1186/1743-0003-4-19] [Medline: 17584501]
63. Thompson AC, Mansfield T, Johnson E, et al. Relationship of multiple sensory impairments with physical performance in older adults in the study of muscle, mobility and aging. J Gerontol A Biol Sci Med Sci 2025 May 5;80(6):glaf065. [doi: 10.1093/gerona/glaf065] [Medline: 40151944]
64. Ma L, Marshall PJ, Wright WG. The impact of external and internal focus of attention on visual dependence and EEG alpha oscillations during postural control. J Neuroeng Rehabil 2022 Jul 26;19(1):81. [doi: 10.1186/s12984-022-01059-7] [Medline: 35883085]

XSL•FO
RenderX

65. Barbanchon C, Mouraux D, Baudry S. Repeated exposure to virtual reality decreases reliance on visual inputs for balance control in healthy adults. Hum Mov Sci 2024 Aug;96:103236. [doi: 10.1016/j.humov.2024.103236] [Medline: 38805764]
66. Liu H, Xing Y, Wu Y. Effect of Wii Fit exercise with balance and lower limb muscle strength in older adults: a meta-analysis. Front Med (Lausanne) 2022;9:812570. [doi: 10.3389/fmed.2022.812570] [Medline: 35602499]
67. Sápi M, Fehér-Kiss A, Csernák K, Domján A, Pintér S. The effects of exergaming on sensory reweighting and mediolateral stability of women aged over 60: usability study. JMIR Serious Games 2021 Jul 21;9(3):e27884. [doi: 10.2196/27884] [Medline: 34287215]

Abbreviations

AP: anterior-to-posterior
MeSH: Medical Subject Headings
ML: medio-to-lateral
MSI: multisensory integration training
TUG: timed up and go

Edited by R Bjarnadottir; submitted 09.Jul.2025; peer-reviewed by Q Song, S Liang, YJ Zhong; revised version received 03.Mar.2026; accepted 09.Apr.2026; published 07.May.2026.

Please cite as:
Guo C, Yin L, Chen P, Zhan J, Yu Z, Tan TC, Wei Y, Gong Y, Xu M, Le VM, Wang L
Effects of Multisensory Integration Training on Postural Stability Characteristics and Fall Risk in Older Adults: Systematic Review and Meta-Analysis
JMIR Aging 2026;9:e80345
URL: https://aging.jmir.org/2026/1/e80345
doi:10.2196/80345

© Chaoyu Guo, Lulu Yin, Peng Chen, Jianglong Zhan, Zhongqi Yu, Teck Cheng Tan, Yaru Wei, Yixue Gong, Menghan Xu, Van Minh Le, Lin Wang. Originally published in JMIR Aging (https://aging.jmir.org), 7.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital and Remote Health for Older Adults in Rural or Underserved Settings: Systematic Review of Clinical, Behavioral, and Implementation Outcomes

Víctor Beltrán[1,2*], PhD; Alain Chaple Gil[3*], PhD; Constanza Morales-Gómez[4*], MSc; Vanessa Campos-Bijit[5,6*], MSc; Rafael Contador[7*], DDS; Leonardo Díaz[8,9,10*], DS; Jaime Bustos[2,11*], DDS; Eduardo Fernandez[7,12*], PhD

[1]Clinical Investigation and Dental Innovation Center (CIDIC), Faculty of Dentistry and Center for Translational Medicine (CEMT-BIOREN), Universidad de La Frontera, Temuco, Chile

[10]Department of Prosthodontics, Faculty of Dentistry, University of Chile, Santiago, Chile

[11]Department of Industrial and System Engineering, Universidad de La Frontera, Temuco, Chile

[12]Instituto de Ciencias Biomédicas, Universidad Autónoma de Chile, Santiago, Chile

[2]Centro de Excelencia de Física e Ingeniería en Salud (CFIS), Universidad de La Frontera, Temuco, Chile

[3]Facultad de Ciencias de la Salud, Universidad Autónoma de Chile, Santiago, Chile

[4]Department of Pathology and Oral Medicine, Faculty of Dentistry, University of Chile, Santiago, Chile

[5]Laboratory of Nanobiomaterials, Research Institute of Dental Sciences, Faculty of Dentistry, University of Chile, Santiago, Chile

[6]Millennium Nucleus for Bioproducts, Genomics and Environmental Microbiology (BioGEM), Valparaíso, Chile

[7]Department of Restorative Dentistry, Faculty of Dentistry, University of Chile, Olivos 943, Santiago, Chile

[8]Department of Stomatology, Faculty of Dentistry, Universidad de Sevilla, Sevilla, Spain

[9]Perioplastic Institute, Santiago, Chile

[*]all authors contributed equally

Corresponding Author:
Eduardo Fernandez, PhD
Department of Restorative Dentistry, Faculty of Dentistry, University of Chile, Olivos 943, Santiago, Chile

Abstract

Background: Older adults in rural or underserved settings face persistent access barriers. Digital and remote health interventions may mitigate these gaps.

Objective: The aim of this study is to assess technologies, effectiveness, and implementation challenges of digital or remote health interventions for adults ≥60 years old in rural or underserved contexts.

Methods: A systematic search was conducted in 4 electronic databases, such as PubMed, Scopus, Web of Science, and Embase (inception: April 2025), under dual screening by 2 reviewers. The tool Cochrane Risk of Bias 2.0 was used to assess the risk of bias for randomized trials, and Risk Of Bias In Non-randomized Studies–of Interventions (ROBINS-I) was used for nonrandomized trials. For observational studies, we used the National Institutes of Health quality assessment tool, and a narrative synthesis was performed.

Results: Fourteen studies met the inclusion criteria (randomized and nonrandomized designs; median follow-up ~12 mo). Ten of 14 (71%) reported significant improvements in at least 1 clinical end point (eg, hemoglobin A_{1c}, blood pressure, and weight), 11 of 14 (79%) improved behavioral or psychosocial outcomes, and 5 of 9 (56%) reduced hospitalizations or acute episodes. Risk-of-bias concerns most frequently related to missing data and selective reporting. Digital literacy and broadband access were the most consistent barriers, while multicomponent, nurse-supported models achieved higher adherence. Equity-related dimensions (place, gender, and socioeconomic status) were variably reported, with limited attention to language, cultural tailoring, or social capital. No study included a formal cost-effectiveness analysis.

Conclusions: Digital and remote health interventions can benefit older adults in rural and underserved settings by improving selected clinical and behavioral outcomes. However, heterogeneity, small sample sizes, and the absence of economic evaluations temper certainty. Importantly, gaps in equity reporting and persistent barriers to digital inclusion highlight the need for future trials to integrate PROGRESS-Plus (place of residence, race/ethnicity, occupation, gender, religion, education, socioeconomic status, social capital, plus other context-specific determinants of health equity) dimensions, embed cost-effectiveness assessments, and design interventions that are both sustainable and accessible to the most underserved populations.

Trial Registration: PROSPERO CRD420251066174; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251066174

XSL•FO
RenderX

(JMIR Aging 2026;9:e78913) doi:10.2196/78913

KEYWORDS

older adults; rural health; digital health; telemedicine; remote care; underserved populations; health disparities

Introduction

Background

The aging of the global population poses unprecedented challenges for health systems, particularly in rural and underserved communities where access to care remains limited [1]. By 2050, it is projected that 1 in 6 people worldwide will be over the age of 65 years, with a significant proportion residing in rural regions [2]. These populations often face geographic, economic, and infrastructural barriers that impede access to timely and adequate health services [3]. Older adults in these regions experience a disproportionate burden of chronic diseases, such as diabetes, heart failure, and hypertension, as well as age-related cognitive decline and social isolation [4].

Digital and remote health interventions, such as telemedicine, mobile health (mHealth) applications, and remote patient monitoring, offer promising avenues to overcome these barriers [5]. These tools can facilitate timely communication with health care providers, enable continuous monitoring of vital signs, and support behavioral interventions aimed at chronic disease management [6]. The integration of such technologies into clinical practice has accelerated, particularly in response to the COVID-19 pandemic, which highlighted the need for resilient and accessible health systems [7].

In this context, telehealth platforms can strengthen provider-to-provider communication, facilitate referral pathways, and support coordination of care across rural and underserved settings.

Health Care Access Inequities in Rural Aging Populations

Access to health care in rural areas is often hindered by long travel distances, reduced mobility, and limited availability of specialists [8]. These issues are compounded for older adults, who may rely on caregivers or public transportation and experience difficulties navigating health care systems [9]. Studies such as the Informatics for Diabetes Education and Telemedicine (IDEATel) trial have demonstrated that telemedicine interventions can improve glycemic control and cardiovascular risk factors in low-income, underserved older adults [10]. Importantly, this study found that individuals from lower socioeconomic strata benefited as much or more than those from higher income groups, suggesting that tailored interventions can reduce health disparities [11].

Similarly, the Telemedical Interventional Monitoring in Heart Failure 2 (TIM-HF2) trial highlighted the effectiveness of remote patient monitoring in managing heart failure. Adherence to remote monitoring protocols was notably higher in rural residents compared to urban dwellers, especially when interventions were accompanied by patient education and regular contact with health care providers [12]. This suggests that rural patients may value the increased contact and support afforded by digital solutions, particularly when traditional in-person care is less accessible [13-15].

Barriers to Adoption and Usability Concerns

Despite the potential of digital health tools, several challenges persist. Digital literacy remains a significant barrier among older adults, many of whom have limited experience with smartphones, tablets, or web-based applications [16]. Moreover, infrastructural deficits—such as unreliable broadband access—disproportionately affect rural areas and can hinder the consistent use of telehealth platforms [17].

Usability and accessibility of devices also play a critical role in determining intervention success. The adoption of user-friendly interfaces, simplified monitoring equipment, and culturally tailored content can enhance engagement among older users. DiNapoli et al [18] emphasize the importance of addressing age-related cognitive and sensory impairments when designing digital interventions, noting that failure to do so can limit the reach and effectiveness of these programs.

Potential for Social and Behavioral Benefits

Beyond clinical outcomes, digital health interventions may also address broader social determinants of health. Perri et al [19] reported that remote interventions helped reduce social isolation and increased self-efficacy in older adults, fostering a sense of autonomy and connectedness. These psychological and behavioral dimensions are particularly important for rural seniors, who often report higher levels of loneliness and depression compared to their urban counterparts [20].

Rationale for a Systematic Review

Given the rapid expansion of digital health technologies and the pressing needs of aging rural populations, there is an urgent need to synthesize the existing literature to identify effective strategies, common pitfalls, and research gaps. Although some studies have reported promising results, the evidence remains fragmented, and many interventions lack rigorous evaluation. Understanding which approaches work best—and under what conditions—can inform policy decisions and guide the development of inclusive, scalable solutions.

Objective

This review aimed to synthesize evidence on digital and remote health interventions for adults aged ≥60 years in rural or otherwise underserved settings, describing (1) technologies and care models, (2) clinical, behavioral, and service-use outcomes, and (3) implementation barriers/facilitators through an equity lens. The primary analytic perspective was patient-centered, focusing on clinical, behavioral, and usability outcomes among older adults; provider-facing facilitators and system-level factors were considered as complementary.

XSL•FO
RenderX

Methods

Study Design

This systematic review adhered to the PRISMA 2020 [21] (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; CRD420251066174) guidelines and was framed using the PROPS [22] (Population, Remote Health Interventions, Outcomes, Procedures, and Study Design) approach. The PROPS framework provided a structured basis for defining the scope of the review as follows:

- Population: adults aged 60 years or older residing in rural or underserved communities
- Intervention: digital or remote health interventions, including telemedicine, mHealth, wearable devices, and remote monitoring tools
- Outcomes: clinical health parameters (eg, blood pressure and glucose levels), behavioral and psychosocial measures, implementation outcomes (feasibility, adherence, and satisfaction), and digital literacy-related outcomes
- Study design: randomized controlled trials, quasi-experimental studies, prospective and retrospective cohort studies, cross-sectional studies, and mixed methods designs

Eligibility Criteria

We included peer-reviewed papers published in all idioms until April 2025 that reported on digital or remote health interventions targeting adults aged ≥60 years living in rural or underserved settings. Eligible study designs included randomized controlled trials, quasi-experimental studies, cohort studies, cross-sectional studies, and mixed methods evaluations. Interventions had to utilize telemedicine, mHealth, remote monitoring, or other digital health modalities. Studies exclusively focused on urban populations, institutionalized participants, or individuals under the age of 60 years were excluded.

Search Strategy

A comprehensive search was conducted in 4 electronic databases: PubMed, Scopus, Web of Science, and Embase. The search strategy combined Medical Subject Headings and free-text terms related to “older adults,” “digital health,” “telemedicine,” “remote monitoring,” “mobile health,” “rural populations,” and “underserved populations.” Search strings were customized for each database. The last search was conducted in April 2025. Reference lists of included papers and relevant reviews were searched manually for additional eligible studies (Multimedia Appendix 1).

Study Selection

Two independent reviewers (EF and ACG) screened the titles and abstracts using the Rayyan QCRI (Qatar Computing Research Institute) [23] platform. Full texts were retrieved for papers that deemed potentially eligible. Disagreements were resolved through discussion or consultation with a third reviewer (LD). The selection process followed the PRISMA 2020 guidelines [20] (Checklist 1), and a flowchart depicting the number of studies screened, included, and excluded was prepared.

Data Extraction

A standardized extraction form was developed and piloted. Extracted data included author(s), year of publication, country, study design, sample size, mean participant age, intervention characteristics, technology used, health domain targeted, key outcomes, usability measures, digital literacy considerations, implementation strategies, and barriers and facilitators. Intervention types and target health outcomes were categorized following the framework derived from the included studies. One reviewer (ACG) conducted data extraction, and a second reviewer (EF) independently verified the entries.

Risk-of-Bias Assessment

Risk of bias was assessed using the Cochrane RoB 2.0 [24] tool for randomized trials and the ROBINS-I [25] tool for nonrandomized studies. For observational studies, we used the NIH quality assessment tool. Two reviewers conducted the assessments independently and discussed discrepancies to reach consensus. Risk-of-bias results were tabulated, and sensitivity analyses were performed when feasible.

Data Synthesis

Given the heterogeneity of intervention types, outcome measures, and study designs, a narrative synthesis was performed. Studies were grouped according to health conditions targeted (eg, cardiovascular and mental health), types of digital intervention (eg, teleconsultation, mobile app, and remote monitoring), and geographic regions. Themes from qualitative studies were synthesized using thematic analysis and included in cross-cutting narratives. Subgroup comparisons (eg, by rural vs semirural, by intervention duration, or technology type) were conducted where sufficient data were available.

Synthesis Under the PROPS Framework

Aligned with the PROPS framework, the population (P) included older adults in rural or underserved communities. Remote health interventions (R) varied in format but consistently included a digital or telehealth component. Outcomes (O) encompassed clinical parameters, behavioral change, patient-reported satisfaction, and health care utilization. Procedures (P) typically involved remote coaching, telemonitoring, or integrated care coordination. Study design (S) was predominantly randomized or quasi-experimental, ensuring a moderate to high level of evidence.

Ethical Considerations

As this study was based entirely on published literature, no ethical approval was required.

Results

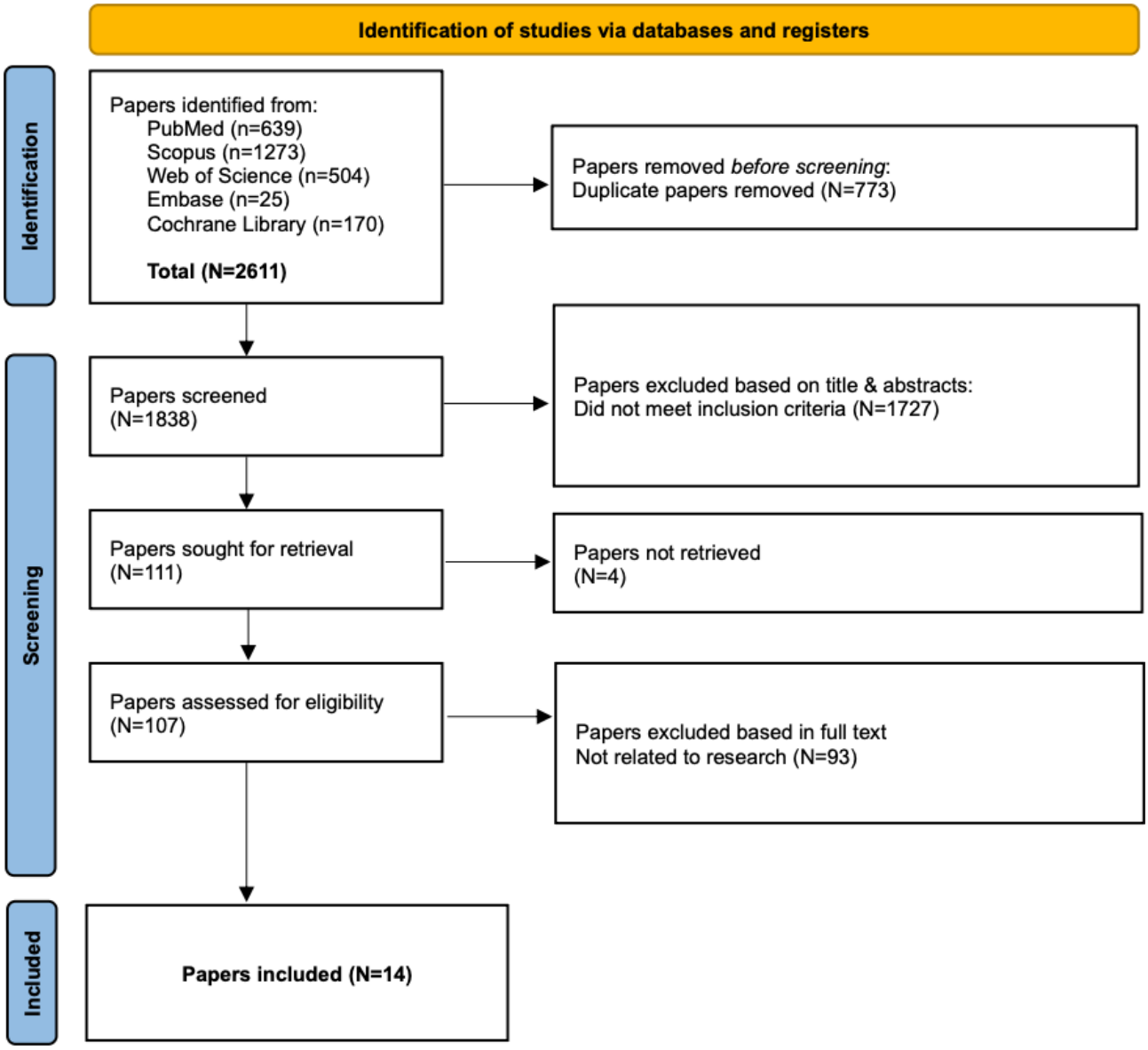
Characteristics of Included Studies

A total of 14 studies met the inclusion criteria and were included in this systematic review (Figure 1). These studies were conducted across diverse settings, including the United States, Canada, Germany, and Australia. The target populations comprised older adults residing in rural or underserved areas and patients with chronic conditions, such as diabetes mellitus, cardiovascular disease, obesity, heart failure, or postsurgical

XSL•FO
RenderX

recovery (Table 1). The majority of the interventions combined digital technologies—such as telemonitoring, phone-based coaching, and web-based rehabilitation—with behavioral, educational, or case management support. Table 1 summarizes the characteristics of included studies, Table 2 details intervention types and digital tools employed, and Table 3 presents the reported outcomes and implementation challenges.

Figure 1. PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020) flowchart of study selection process**.** This flow diagram illustrates the identification, screening, eligibility, and inclusion stages of the systematic review. A total of 14 studies met the eligibility criteria and were included in the final synthesis after full-text assessment.



XSL•FO
RenderX

Table . Characteristics of included studies.

Study	Country	Population	Telehealth intervention	Comparator	Main outcomes	Follow-up duration
Barnason et al [16] (2019)	United States	Cardiac patients (post-PCI[a] or CABS[b]), overweight/obese, rural	12-week phone coaching+nurse support	Usual care	Weight loss and self-efficacy	4 and 6 months
DiNapoli et al [18] (2017)	United States	Older adults, low QoL[c], mild cognitive impairment, rural	16 - 20 sessions of phone-delivered CBT[d]	Minimal support control (weekly calls)	Anxiety, depression, and QoL	6 months
Eberly et al [7] (2025)	United States	Heart failure, underserved	Phone-based GDMT[e] titration+blood pressure monitoring	Usual care	GDMT uptake	30 days
Krum et al [26] (2013)	Australia	Heart failure, rural	Phone support+nurse follow-up	Usual care	Hospitalizations and mortality	12 months
Lear et al [3] (2015)	Canada	Cardiac patients, rural and small urban	Web-based rehabilitation platform	Usual care	Exercise capacity and lipid profile	16 months
Perri et al [19] (2019)	United States	Obese adults (528), rural	Telecoaching+tracking	In-person	Weight loss and retention	4, 6, and 12 months
Perri et al [17] (2020)	United States	Obese adults (445), rural	Phone counseling (group/individual)+education	Usual care	Weight loss and maintenance	4, 10, 16, and 22 months
Prescher et al [12] (2023)	Germany	Elderly with HF[f] (1538), rural/urban	Remote monitoring+care coordination	Usual care	Function, QoL, BP[g], ECG[h]	12 months
Russell et al [27] (2011)	United States	Post-knee arthroplasty patients (65), rural	Telemonitoring	Usual care	Function and OA[i] Index	6 weeks
Shea et al [4](2009)	United States	Diabetes (1665), underserved	Telemonitoring+nurse case management	Usual care	HbA_{1c}[j], LDL[k], BP, and satisfaction	5 years
Shea et al [10,28] (2013)	United States	Diabetes, underserved	Telemonitoring+nurse care management	Usual care	HbA_{1c}, SBP[l], and LDL	12 months
Smith et al [29] (2000)	United States	Rural older adults	Telehealth counseling	Not reported	QoL and satisfaction	5 months
West et al [30] (2010)	United States	Diabetes, rural older adults	Telehealth counseling	Usual care	Behavior change	2 - 6 years
Wilson et al [20] (2016)	United States	Smokers, rural	Telemedicine advice	Usual care	Abstinence and self-efficacy	6 months

[a]PCI: percutaneous coronary intervention.

[b]CABS: coronary artery bypass grafting.

[c]QoL: quality of life.

[d]CBT: cognitive behavioral therapy.

[e]GDMT: guideline-directed medical therapy.

[f]HF: heart failure.

[g]BP: blood pressure.

[h]ECG: electrocardiogram.

[i]OA: osteoarthritis.

[j]HbA_{1c}: glycated hemoglobin.

[k]LDL: low-density lipoprotein.

XSL•FO
RenderX

[l]SBP: systolic blood pressure.

Table . Effectiveness outcomes by studies.

Study	Clinical outcome	Behavioral outcome	Patient satisfaction	Health care utilization
Barnason et al [16] (2019)	Sustained weight loss	Improved activity and self-efficacy	Moderate satisfaction	Not reported
DiNapoli et al [18] (2017)	No significant clinical change	Improved mood and activation	High satisfaction	Not reported
Eberly et al (2025)	Increased GDMT[a] use	Improved adherence	High protocol adherence	Fewer HF[b] hospitalizations
Krum et al [26] (2013)	Reduced hospitalizations	Not reported	Not reported	Reduced hospitalizations
Lear et al [3] (2015)	Better lipid profile	Not reported	Usability positive	Not reported
Perri et al [19] (2019)	Weight loss	High goal attainment	High satisfaction	Reduced episodes
Perri et al [17] (2020)	Weight loss	Behavior change	High satisfaction	Reduced episodes
Prescher et al [12] (2023)	Cardiac improvements	Improved motivation	High adherence	Lower hospitalization
Russell et al [27] (2011)	Improved OA[c] outcomes	Self-care adherence	High satisfaction	Reduced episodes
Shea et al [4] (2009)	HbA_{1c}[d], BP[e] reduction	Empowerment increased	Positive care perception	Lower rates
Shea et al [10,28] (2013)	HbA_{1c}, SBP[f] reduced	Health management	Not reported	Not reported
Smith et al [29] (2000)	No clinical change	Health awareness	Positive experience	Not reported
West et al [30] (2010)	Improved diabetes control	Goal execution improved	High satisfaction	Not reported
Wilson et al [20] (2016)	Verified smoking cessation	Health awareness	High acceptability	Not reported

[a]GDMT: guideline-directed medical therapy.

[b]HF: heart failure.

[c]OA: osteoarthritis.

[d]HbA_{1c}: glycated hemoglobin.

[e]BP: blood pressure.

[f]SBP: systolic blood pressure.

Table . Clinical recommendations.

Clinical domain	Recommended modality	Target population	Recommendation
Chronic conditions	Multicomponent (coaching+monitoring)	Rural older adults	Integrate digital tools with behavioral support
Diabetes	Telemonitoring+nurse coaching	Rural older adults	Combine remote feedback and coaching
Cardiovascular rehabilitation	Web-based rehabilitation+support	Postcardiac discharge, rural	Use online rehabilitation when in-person is not feasible
Weight management	Behavioral telecoaching	Obese rural adults	Support goals via telecoaching
Mental health	Phone-based behavioral therapy	Rural older adults	Implement CBT[a] remotely

[a]CBT: cognitive behavioral therapy.

Clinical Outcomes (Patient Perspective)

Clinical improvements were reported in most of the included studies. Glycemic control (hemoglobin A_{1c} [HbA_{1c}]), blood pressure, lipid profiles, and weight reduction were the most frequently evaluated parameters. For example, Shea et al [10,28] observed significant reductions in HbA1c and systolic blood pressure following nurse-led telemonitoring interventions in low-income and underserved diabetic populations. Similarly, Wilson et al [20] reported biochemical validation of smoking abstinence in rural participants receiving telemedicine counseling, whereas Perri et al [17,19] demonstrated sustained weight loss in older obese adults receiving telecoaching and tracking interventions for up to 22 months.

In the cardiovascular domain, Lear et al [3] found that a web-based rehabilitation platform improved lipid profiles and exercise capacity, whereas Krum et al [26], Barnason et al [16], and Prescher et al [12] reported fewer heart failure–related hospitalizations and improvements in weight and blood pressure through structured remote follow-up (Table 2). Notably, Eberly et al [7] demonstrated a significant increase in the uptake of

XSL•FO
RenderX

guideline-directed medical therapy for heart failure, supported by remote blood pressure monitoring and medication titration.

Behavioral and Psychosocial Outcomes

Fifteen studies evaluated behavioral and psychosocial variables, including self-efficacy, adherence, physical activity, and mood. DiNapoli et al [18] used cognitive behavioral therapy (CBT) delivered through telephone and observed improvements in mood, patient activation, and decreased depressive symptoms. Barnason et al [16] and Perri et al [17] reported increases in health-related behaviors, such as dietary control, physical activity, and treatment adherence, alongside weight loss outcomes.

Several studies emphasized the role of motivation, activation, and goal setting. For instance, West et al [30] and Smith et al [29] found that behavioral activation and health awareness were significantly improved among older rural adults participating in structured telehealth counseling sessions. Notably, these effects were more pronounced when interventions included personalized coaching and caregiver engagement.

Usability and Satisfaction (Patient-Reported)

Patient satisfaction and usability were assessed in 13 studies. Most interventions received favorable feedback from participants, who cited increased accessibility, reduced need for transportation, and personalized interactions as key benefits. Interventions led by nurses or care coordinators (eg, Shea et al [4], Prescher et al [12], and Smith et al [29]) had especially high ratings. Perri et al [19] reported high retention and satisfaction in group-based telecoaching models, emphasizing the value of social support and continuous feedback.

Although some participants noted challenges, such as intervention length or digital interface complexity, overall engagement and willingness to continue using digital health solutions remained strong across studies (Table 2).

Health Care Utilization and Efficiency

Nine studies reported health care utilization outcomes, including hospital readmissions, acute episode rates, and emergency department visits. Telehealth interventions consistently reduced hospitalization rates in older adults with heart failure or chronic disease, as evidenced in studies by Krum et al [26], Barnason et al [16], and Prescher et al [12]. Shea et al [4] and Perri et al [17] observed reduced acute episodes and improved care coordination through remote monitoring.

While Eberly et al [7] documented fewer cardiac hospitalizations through proactive titration of medications, the majority of studies lacked formal cost-effectiveness analyses. As shown in Table 4, future evaluations must address the economic implications of scaling telehealth models in rural and resource-constrained contexts.

Table . Limitations and research recommendations.

Limitation	Future research action	Justification
Short follow-up	Long-term RCTs[a] (≥12 mo)	Assess sustainability
Small samples	Larger trials with stratification	Avoid type II errors
Lack of cost analysis	Include economic evaluations	Guide policy and budgeting
Limited generalizability	Cultural tailoring across regions	Address population heterogeneity
Inconsistent outcomes	Standardized outcome sets	Enable synthesis
Digital barriers	User-centered design for low literacy	Improve accessibility

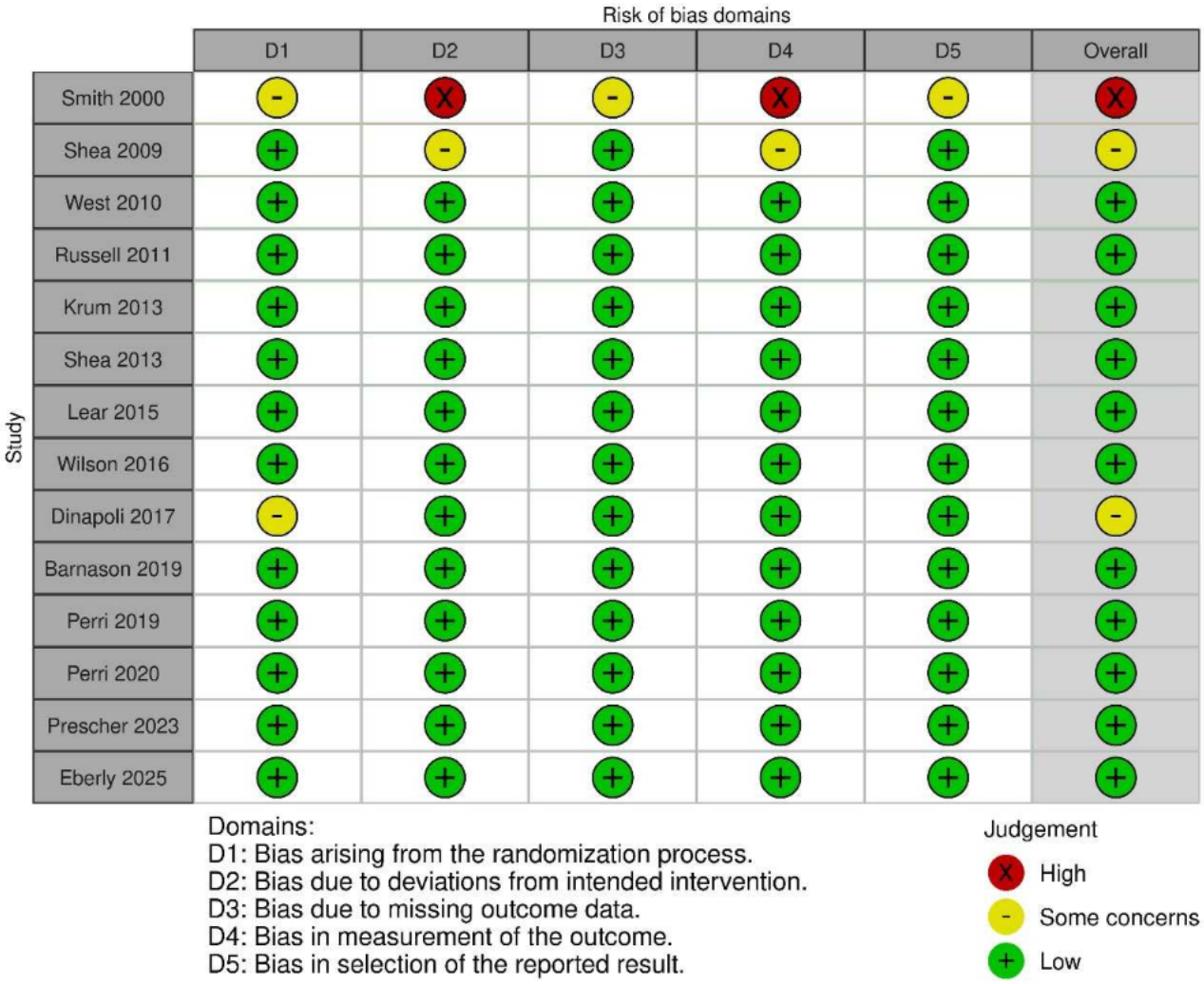
[a]RCTs: randomized controlled trials.

Risk of Bias and Study Quality

Figure 2 presents the risk of bias assessment. Using the RoB 2.0 tool, most studies were classified as low risk in randomization, outcome measurement, and reporting domains. However, Smith et al [29] was flagged as high risk due to deviation from protocols and unclear allocation procedures, and DiNapoli et al [18] presented some concerns in the randomization process. Despite these limitations, overall study quality was deemed adequate, with the majority of trials employing rigorous designs and clear outcome reporting.

XSL•FO
RenderX

Figure 2. Risk-of-bias assessment for the 14 included studies. Graphical summary of bias domains assessed using the Cochrane Risk of Bias 2.0 tool. Domains include D1 (randomization process), D2 (deviations from intended intervention), D3 (missing outcome data), D4 (measurement of the outcome), and D5 (selection of reported result). Green (+) indicates low risk, yellow (−) indicates some concerns, and red (×) indicates high risk. Overall risk of bias was low in 12 studies, with two studies (Smith et al [29] and DiNapoli et al [18]) rated as high or having some concerns [3,4,7,10,12,16-20,26,27,29,30].



Implementation Facilitators (Provider Perspective)

Key implementation barriers included limited digital literacy, lack of broadband connectivity, and insufficient cultural adaptation of interventions. Several studies highlighted the importance of preintervention training, simplified user interfaces, and language-appropriate materials. For example, Russell et al [27] and DiNapoli et al [18] integrated patient education modules and caregiver support, which improved both adherence and satisfaction.

Multicomponent interventions—combining synchronous coaching (eg, telephone and video) and asynchronous tracking (eg, mobile apps and remote monitoring)—were associated with greater behavioral change and health gains. Interventions with durations of 6 months or longer demonstrated more durable outcomes. Conversely, short follow-up periods (eg, 30 days in Eberly et al [7]) limited the understanding of long-term sustainability.

Heterogeneity and Feasibility of Meta-Analysis

The feasibility of conducting a meta-analysis was limited by marked heterogeneity across the included studies in terms of populations, intervention modalities, outcome measures, and follow-up durations. Clinical end points ranged from HbA_{1c} and blood pressure to functional capacity and psychological scales, with inconsistent definitions and reporting formats that precluded statistical pooling. Moreover, small sample sizes and diverse study designs further reduced comparability. For these reasons, a narrative synthesis was judged to be the most appropriate approach, allowing results to be interpreted within outcome domains while acknowledging variability in effect sizes and methodological rigor.

Synthesis Approach (PROPS)

Findings were organized according to the PROPS framework, grouping evidence by population characteristics, remote health intervention types, outcomes, and setting or contextual factors.

XSL•FO
RenderX

This approach allowed us to compare heterogeneous studies in a structured manner and highlight converging patterns. For example, in populations with cardiometabolic conditions, telemonitoring combined with nurse support consistently improved HbA_{1c} and blood pressure outcomes, whereas in mental health or behavioral domains, telephone-based CBT and lifestyle coaching showed greater impact on psychological well-being and adherence. Variability across settings, particularly in broadband access and digital literacy, was emphasized as a determinant of intervention success. Narrative weighting of evidence was guided by risk-of-bias judgments, giving higher emphasis to low-risk trials and cautious interpretation of studies with missing data or selective reporting

Equity (PROGRESS-Plus)

Equity-related variables, such as place, gender, and socioeconomic status, were variably reported across studies, while domains, such as religion and social capital, were largely absent. Detailed reporting of each dimension and study is provided in Table S4 in Multimedia Appendix 1.

Summary

Tables 1-4 synthesize the evidence on participants' characteristics, effectiveness, implementation strategies, and limitations. The findings support the potential of digital health interventions to improve health outcomes, reduce hospitalization, and enhance engagement among older adults in rural or underserved settings. However, sustainability, equity, and generalizability remain key areas for further research.

Comprehensive supplementary data provide additional granularity to the synthesized findings. Table S1 in Multimedia Appendix 1 details the complete database search strategy used across PubMed, Scopus, Web of Science, Embase, and Cochrane, confirming methodological transparency and reproducibility. Table S2 in Multimedia Appendix 1 summarizes the direction, magnitude, and certainty of effects for each included study by outcome domain, strengthening the interpretation of clinical and behavioral improvements reported in the main text. Table S3 in Multimedia Appendix 1 presents the frequency distribution of risk-of-bias domains across randomized, nonrandomized, and observational designs, supporting the overall quality assessment discussed earlier. Equity dimensions were systematically mapped in Table S4 in Multimedia Appendix 1, evidencing variable reporting across PROGRESS-Plus (place of residence, race/ethnicity, occupation, gender, religion, education, socioeconomic status, social capital, plus other context-specific determinants of health equity) categories. Table S5 in Multimedia Appendix 1 provides adherence and attrition data, illustrating high engagement rates (>75%) despite common technical barriers, and Table S6 in Multimedia Appendix 1 classifies intervention models by delivery type, clinical applicability, and observed outcomes, emphasizing the comparative advantages of multicomponent, hybrid telehealth formats. Collectively, these supplementary materials reinforce the robustness, contextual diversity, and implementation relevance of the reviewed evidence.

Discussion

Main Findings

This systematic review of 14 studies shows that remote and digital health interventions can significantly improve clinical, behavioral, and satisfaction-related outcomes in older adults living in rural or underserved settings. Telehealth modalities, such as telephone counseling, nurse-led monitoring, and web-based platforms, consistently enhanced chronic disease management, promoted healthier behaviors, and increased accessibility to care. Notable clinical benefits included reductions in HbA_{1c}, blood pressure, and low-density lipoprotein cholesterol in diabetic patients, improved uptake of guideline-directed medical therapy and fewer short-term hospitalizations in heart failure, and better weight loss and retention outcomes in obesity. Behavioral interventions also proved effective, with structured coaching and CBT programs improving patient activation, physical activity, and psychological well-being, highlighting the potential of remote care to address barriers faced by these populations.

Comparison With Prior Work

Our findings align with prior evidence indicating the efficacy of remote interventions in chronic disease management. For instance, Shea et al's [4] work on the IDEATel project has long emphasized the role of telehealth in improving glycemic and blood pressure control in underserved populations. Similarly, Lear et al [3] confirmed the value of web-based platforms in cardiac rehabilitation, echoing the current study's results. However, while prior meta-analyses (eg, Marcolino et al [31]) have generally combined urban and rural populations, this review narrows the focus to older adults in rural or resource-limited settings, highlighting unique challenges and opportunities in this demographic.

Several of the studies in this review support the role of multicomponent telehealth strategies that combine behavioral coaching, self-monitoring tools, and clinical oversight. For example, Prescher et al [12] demonstrated improvements in cardiac functions and reduced hospitalizations when remote monitoring was integrated with coordinated care. This convergence of digital tools and human support appears especially important in rural contexts, where health care infrastructure and workforce shortages often limit access to traditional services [14,32].

Interpretation by Outcome Domains

When interpreting the findings by outcome domains, the evidence suggests that digital and remote health interventions can generate benefits across clinical, behavioral, and health care utilization outcomes, though with varying levels of certainty. Clinical outcomes, such as HbA1c reduction, blood pressure control, and weight loss, were reported in a subset of studies, indicating that structured remote monitoring and nurse-supported programs are particularly effective for cardiometabolic management. Behavioral and psychosocial domains showed improvements in patient activation, medication adherence, and psychological well-being, highlighting the value of tailored telecoaching and CBT-based interventions. In contrast, health

XSL•FO
RenderX

service utilization outcomes, such as reduced hospitalizations or emergency visits, were less consistently observed, likely reflecting short follow-up periods and small sample sizes. Taken together, these results indicate that while the strongest evidence lies in clinical and behavioral benefits, sustained improvements in health care utilization may require longer interventions and broader system-level integration.

Equity Appraisal (PROGRESS-Plus)

Applying the PROGRESS-Plus framework revealed important gaps in how equity dimensions were reported and addressed across the included studies. Most trials provided some data on place of residence, education, or socioeconomic status, but few explicitly measured digital literacy, broadband availability, or cultural and linguistic adaptation of interventions. Gender and age were routinely reported, yet aspects such as occupation, religion, or social capital were rarely considered. Notably, none of the studies systematically evaluated how these factors moderated intervention outcomes, which limits the ability to determine whether digital tools reduced or inadvertently widened disparities. This underreporting underscores the need for future research to integrate equity indicators prospectively into study design and to tailor interventions to the structural barriers and lived experiences of rural and underserved older populations.

Limitations of Included Studies

Despite promising outcomes, limitations in study design, follow-up duration, and reporting hinder the generalizability of findings. As summarized in Table 4, many studies had short follow-up durations (eg, 30 d to 6 mo), which limit our understanding of long-term intervention sustainability. Only a few studies, such as West et al [30] and Shea et al [4], offered follow-up beyond 12 months.

Sample sizes were frequently small or underpowered (eg, Russell et al [27] and Smith et al [29]), increasing susceptibility to type II errors and reducing the robustness of subgroup analyses. The heterogeneity in intervention types and outcome measures also complicates cross-study comparisons. For example, while weight loss and HbA_{1c} were commonly measured, behavioral metrics varied widely and were often self-reported without standardized instruments.

Additionally, economic evaluations were notably absent. None of the included studies systematically reported on cost-effectiveness, despite the critical importance of economic data for policymaking and reimbursement frameworks in telehealth implementation [33,34]. All included studies were conducted in high-income countries; the absence of eligible studies from low- and middle-income countries limits generalizability. We explicitly searched without country restrictions; the lack of low- and middle-income country evidence appears to reflect publication scarcity rather than an a priori exclusion.

Limitations of This Review

This review also has several methodological limitations that should be acknowledged. First, the small and heterogeneous evidence base precluded quantitative pooling and limited the generalizability of conclusions. Second, despite a comprehensive search strategy, we restricted inclusion to peer-reviewed, indexed papers, which may have introduced publication and language bias by excluding gray literature and non-English studies. Third, although dual reviewers independently screened and extracted data, subjective judgments in the narrative synthesis may have influenced the interpretation of findings. Finally, all included studies originated from high-income countries, constraining the applicability of results to low- and middle-income settings where digital and infrastructural barriers may differ substantially.

Barriers to Implementation and Digital Equity

An important contextual finding across studies is the presence of digital and literacy-related barriers among older rural adults [35,36]. Digital literacy was rarely assessed systematically; however, studies like Wilson et al [20] and Smith and Weinert [29] hinted at limited technology familiarity and engagement. Moreover, intervention adherence and satisfaction appeared to be influenced by factors such as perceived complexity, trust in remote systems, and personal contact with health care providers.

These findings echo previous reports emphasizing the need for user-centered, accessible design tailored to the cognitive and sensory needs of older adults. As Barnason et al [16] and DiNapoli et al [18] noted, patient satisfaction and engagement were higher when interventions included regular human interaction, simplified instructions, and consistent technical support.

Implementation Barriers and Health Equity Considerations

The review highlights persistent digital and infrastructural barriers disproportionately affecting older adults in rural or underserved settings [37,38]. Limited broadband access, low digital literacy, and lack of culturally tailored content not only hinder adoption but reinforce existing health inequities [39-41]. These implementation challenges are structurally linked to social determinants of health, such as lower educational attainment and income levels in rural areas [42]. Without addressing these foundational disparities, digital health solutions risk exacerbating rather than alleviating access gaps. Therefore, policy interventions should include investments in digital infrastructure, community-based training, and participatory design processes that reflect local needs [43].

Innovation and Emerging Models

This review highlights not only the effectiveness of established telehealth modalities but also the emergence of hybrid care models and novel delivery platforms tailored to rural aging populations. Notably, interventions incorporating both synchronous (eg, phone-based coaching) and asynchronous (eg, mobile tracking or wearable monitoring) modalities were associated with higher patient engagement and adherence. For instance, the Rural Lifestyle, Exercise, Attitudes, and Relationships Program (LEAP) trial [19] compared group versus individual telecoaching, revealing that group-based formats may enhance social connectedness—a crucial factor for isolated older adults. Similarly, the use of wearable-enabled remote monitoring in heart failure management (eg, TIM-HF2 [12], Hózhó trial [7]) represents a shift toward continuous and passive

XSL•FO
RenderX

data collection, enabling earlier detection of clinical deterioration.

These findings align with a recent framework such as the National Academies of Sciences, Engineering, and Medicine 2022 Digital Health Equity Roadmap, which emphasizes human-centered design, infrastructure co-development, and inclusive implementation pathways to ensure equitable access to digital health. Compared to earlier reviews that emphasized general digital adoption [31], this study foregrounds context-specific innovations, such as culturally adapted content, caregiver-mediated interfaces, and localized broadband workarounds, which offer blueprints for scalable, equitable intervention design in marginalized aging populations.

Implications (Practice or Policy)

To translate the observed benefits of telehealth into routine care, integration into existing primary care models is essential [44]. This requires not only technological interoperability but also alignment with reimbursement policies, provider training, and patient education [45]. Interventions led by nurses or care coordinators demonstrated high satisfaction and adherence, suggesting that a team-based, hybrid care model may offer the most sustainable pathway [46]. Health systems must also establish clear protocols for digital triage, follow-up, and escalation of care to ensure clinical safety [47]. Policymakers should incentivize digital inclusion initiatives and consider telehealth reimbursement parity to support widespread adoption in rural and underserved populations [48]. Distinguishing patient-reported outcomes from provider-facing facilitators clarifies that improvements in usability or adherence reflect patient experiences, whereas professional networking and workflow integration represent implementation considerations at the provider/system level.

Economic Considerations

Although several interventions demonstrated clinical and behavioral benefits, none of the included studies conducted a formal cost-effectiveness analysis. This absence of economic evidence substantially limits the ability of health systems and policymakers to make informed, data-driven decisions regarding large-scale implementation. Without robust cost analyses, it is unclear whether the observed benefits can be achieved sustainably or equitably across diverse settings. Future trials should therefore embed cost-effectiveness and budget impact evaluations alongside clinical outcomes to determine value for money, support reimbursement frameworks, and guide resource allocation in rural and underserved populations.

Risk of Publication and Language Bias

The small number of included studies limited the feasibility of conducting formal statistical assessments of publication bias, such as funnel plots or regression-based tests. With only 14 trials, these methods would be underpowered and potentially misleading. Nevertheless, the predominance of positive findings raises the possibility of small-study effects. In addition, by restricting our search to peer-reviewed and indexed literature, we may have excluded gray literature or non-English studies that could provide null or negative results. These factors suggest that the overall effectiveness of digital health interventions in rural older populations may be somewhat overestimated, and future reviews should seek to mitigate this risk by incorporating broader data sources and multilingual evidence.

Recommendations for Future Research

To advance the field, future studies should prioritize longer follow-up durations to assess the sustainability of clinical and behavioral gains. Large, multicenter trials with stratified randomization would enhance generalizability and allow for meaningful subgroup analysis by age, comorbidity burden, or digital literacy level.

Moreover, standardized outcome reporting is critical. Adopting common core sets of metrics—particularly for behavioral and quality-of-life domains—would enable pooled synthesis and meta-analyses, strengthening the evidence base. Economic evaluations, including cost-effectiveness and budget impact analyses, should also be embedded into study designs.

The implementation of satellite-based internet in geographically remote areas with limited or no connectivity can improve the functionality and accessibility of telehealth platforms and applications, facilitating communication between health care providers and older adult patients. In this context, the limited availability of specialist care in underserved regions may be addressed through the progressive integration of telehealth technologies, expanding access to specialized services for aging populations.

Finally, cultural tailoring and language accessibility should be emphasized in designing telehealth platforms. Rural populations are not homogenous; successful interventions must adapt to the social, linguistic, and infrastructural realities of diverse communities.

Conclusions

Telehealth interventions represent a promising approach to enhancing the health and well-being of older adults in rural and underserved settings. Evidence from 14 studies indicates that remote strategies can contribute to improved management of chronic conditions, support behavioral change, and in some cases, reduce hospitalizations. Yet, important gaps remain regarding cost-effectiveness, outcome standardization, and digital inclusion. Integrating human support with technological solutions appears to enhance adherence and overall effectiveness. Equitable access and long-term sustainability will require greater emphasis on digital literacy, cultural adaptation, and economic feasibility. At present, the lack of formal cost-effectiveness analyses constrains data-driven decisions about large-scale adoption. Future studies should therefore embed economic evaluations and equity considerations to provide the evidence base needed for informed policy and resource allocation in underserved contexts.

XSL•FO
RenderX

Funding

This study would not have been possible without the support of the FRO2495 Ministerio de Educación de Chile (MINEDUC) project, "Integrated Technological Ecosystem of Tele-Dentistry to Provide General and Specialized Dental Care for Older Adults."

Conflicts of Interest

None declared.

Multimedia Appendix 1
Methodological materials and extended evidence synthesis.
[DOCX File, 31 KB - aging_v9i1e78913_app1.docx]

Checklist 1
PRISMA checklist.
[PDF File, 172 KB - aging_v9i1e78913_app2.pdf]

References

1. Asante D, McLachlan CS, Pickles D, Isaac V. Understanding unmet care needs of rural older adults with chronic health conditions: a qualitative study. Int J Environ Res Public Health 2023 Feb 13;20(4):3298. [doi: 10.3390/ijerph20043298] [Medline: 36833993]
2. Noto S. Perspectives on aging and quality of life. Healthcare (Basel) 2023 Jul 26;11(15):2131. [doi: 10.3390/healthcare11152131] [Medline: 37570372]
3. Lear SA, Singer J, Banner-Lukaris D, et al. Improving access to cardiac rehabilitation using the internet: a randomized trial. Stud Health Technol Inform 2015;209:58-66. [doi: 10.3233/978-1-61499-505-0-58] [Medline: 25980706]
4. Shea S, Weinstock RS, Teresi JA, et al. A randomized trial comparing telemedicine case management with usual care in older, ethnically diverse, medically underserved patients with diabetes mellitus: 5 year results of the IDEATel study. J Am Med Inform Assoc 2009;16(4):446-456. [doi: 10.1197/jamia.M3157] [Medline: 19390093]
5. Mumtaz H, Riaz MH, Wajid H, et al. Current challenges and potential solutions to the use of digital health technologies in evidence generation: a narrative review. Front Digit Health 2023;5:1203945. [doi: 10.3389/fdgth.2023.1203945] [Medline: 37840685]
6. Peyroteo M, Ferreira IA, Elvas LB, Ferreira JC, Lapão LV. Remote monitoring systems for patients with chronic diseases in primary health care: systematic review. JMIR Mhealth Uhealth 2021 Dec 21;9(12):e28285. [doi: 10.2196/28285] [Medline: 34932000]
7. Eberly LA, Tennison A, Mays D, et al. Telephone-based guideline-directed medical therapy optimization in Navajo nation: the Hózhó randomized clinical trial. JAMA Intern Med 2024 Jun 1;184(6):681-690. [doi: 10.1001/jamainternmed.2024.1523] [Medline: 38583185]
8. Coombs NC, Campbell DG, Caringi J. A qualitative study of rural healthcare providers' views of social, cultural, and programmatic barriers to healthcare access. BMC Health Serv Res 2022 Apr 2;22(1):438. [doi: 10.1186/s12913-022-07829-2] [Medline: 35366860]
9. Li H. Rural older adults' access barriers to in-home and community-based services. Soc Work Res 2006 Jun 1;30(2):109-118. [doi: 10.1093/swr/30.2.109]
10. Shea S, Kothari D, Teresi JA, et al. Social impact analysis of the effects of a telemedicine intervention to improve diabetes outcomes in an ethnically diverse, medically underserved population: findings from the IDEATel study. Am J Public Health 2013 Oct;103(10):1888-1894. [doi: 10.2105/AJPH.2012.300909] [Medline: 23488491]
11. Thornton RLJ, Glover CM, Cené CW, Glik DC, Henderson JA, Williams DR. Evaluating strategies for reducing health disparities by addressing the social determinants of health. Health Aff (Millwood) 2016 Aug 1;35(8):1416-1423. [doi: 10.1377/hlthaff.2015.1357] [Medline: 27503966]
12. Prescher S, Winkler S, Riehle L, et al. Patient reported experience and adherence to remote patient management in chronic heart failure patients: a posthoc analysis of the TIM-HF2 trial. Eur J Cardiovasc Nurs 2023 Apr 12;22(3):245-253. [doi: 10.1093/eurjcn/zvac080] [Medline: 36062451]
13. Klee D, Pyne D, Kroll J, James W, Hirko KA. Rural patient and provider perceptions of telehealth implemented during the COVID-19 pandemic. BMC Health Serv Res 2023 Sep 12;23(1):981. [doi: 10.1186/s12913-023-09994-4] [Medline: 37700286]
14. Maita KC, Maniaci MJ, Haider CR, et al. The impact of digital health solutions on bridging the health care gap in rural areas: a scoping review. Perm J 2024 Sep 16;28(3):130-143. [doi: 10.7812/TPP/23.134] [Medline: 39135461]
15. Gobburi RK, Olawade DB, Olatunji GD, Kokori E, Aderinto N, David-Olawade AC. Telemedicine use in rural areas of the United Kingdom to improve access to healthcare facilities: a review of current evidence. Inf Health 2025 Mar;2(1):41-48. [doi: 10.1016/j.infoh.2025.01.003]

XSL•FO
RenderX

16. Barnason S, Zimmerman L, Schulz P, Pullen C, Schuelke S. Weight management telehealth intervention for overweight and obese rural cardiac rehabilitation participants: a randomised trial. J Clin Nurs 2019 May;28:1808-1818. [doi: 10.1111/jocn.14784] [Medline: 30667588]
17. Perri MG, Shankar MN, Daniels MJ, et al. Effect of telehealth extended care for maintenance of weight loss in rural US communities: a randomized clinical trial. JAMA Netw Open 2020 Jun 1;3(6):e206764. [doi: 10.1001/jamanetworkopen.2020.6764] [Medline: 32539150]
18. DiNapoli EA, Pierpaoli CM, Shah A, Yang X, Scogin F. Effects of home-delivered cognitive behavioral therapy (CBT) for depression on anxiety symptoms among rural, ethnically diverse older adults. Clin Gerontol 2017;40(3):181-190. [doi: 10.1080/07317115.2017.1288670] [Medline: 28452665]
19. Perri MG, Ariel-Donges AH, Shankar MN, et al. Design of the Rural LEAP randomized trial: an evaluation of extended-care programs for weight management delivered via group or individual telephone counseling. Contemp Clin Trials 2019 Jan;76:55-63. [doi: 10.1016/j.cct.2018.11.006] [Medline: 30408606]
20. Wilson SM, Hair LP, Hertzberg JS, et al. Abstinence Reinforcement Therapy (ART) for rural veterans: methodology for an mHealth smoking cessation intervention. Contemp Clin Trials 2016 Sep;50:157-165. [doi: 10.1016/j.cct.2016.08.008] [Medline: 27521811]
21. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
22. Nasser M, Kumbargere S, Uhm S, Eachempati P, Bhaumik S. Do we need systematic reviews of research priority setting? A proposal for a new concept on conducting systematic reviews of research priority setting exercises. University of Southampton. Preprint posted online on Mar 1, 2025. [doi: 10.22541/au.174081962.27718420/v1]
23. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
24. Nejadghaderi SA, Balibegloo M, Rezaei N. The Cochrane risk of bias assessment tool 2 (RoB 2) versus the original RoB: a perspective on the pros and cons. Health Sci Rep 2024 Jun;7(6):e2165. [doi: 10.1002/hsr2.2165] [Medline: 38835932]
25. Igelström E, Campbell M, Craig P, Katikireddi SV. Cochrane's risk of bias tool for non-randomized studies (ROBINS-I) is frequently misapplied: a methodological systematic review. J Clin Epidemiol 2021 Dec;140:22-32. [doi: 10.1016/j.jclinepi.2021.08.022] [Medline: 34437948]
26. Krum H, Forbes A, Yallop J, et al. Telephone support to rural and remote patients with heart failure: the Chronic Heart Failure Assessment by Telephone (CHAT) study. Cardiovasc Ther 2013 Aug;31(4):230-237. [doi: 10.1111/1755-5922.12009] [Medline: 23061492]
27. Russell TG, Buttrum P, Wootton R, Jull GA. Internet-based outpatient telerehabilitation for patients following total knee arthroplasty: a randomized controlled trial. J Bone Joint Surg Am 2011 Jan 19;93(2):113-120. [doi: 10.2106/JBJS.I.01375] [Medline: 21248209]
28. Trief PM, Teresi JA, Eimicke JP, Shea S, Weinstock RS. Improvement in diabetes self-efficacy and glycaemic control using telemedicine in a sample of older, ethnically diverse individuals who have diabetes: the IDEATel project. Age Ageing 2009 Mar;38(2):219-225. [doi: 10.1093/ageing/afn299] [Medline: 19171951]
29. Smith L, Weinert C. Telecommunication support for rural women with diabetes. Diabetes Educ 2000;26(4):645-655. [doi: 10.1177/014572170002600412] [Medline: 11140073]
30. West SP, Lagua C, Trief PM, Izquierdo R, Weinstock RS. Goal setting using telemedicine in rural underserved older adults with diabetes: experiences from the informatics for diabetes education and telemedicine project. Telemed J E Health 2010 May;16(4):405-416. [doi: 10.1089/tmj.2009.0136] [Medline: 20507198]
31. Marcolino MS, Maia LM, Oliveira JAQ, et al. Impact of telemedicine interventions on mortality in patients with acute myocardial infarction: a systematic review and meta-analysis. Heart 2019 Oct;105(19):1479-1486. [doi: 10.1136/heartjnl-2018-314539] [Medline: 31253696]
32. Tagne JF, Burns K, O'Brein T, et al. Challenges for remote patient monitoring programs in rural and regional areas: a qualitative study. BMC Health Serv Res 2025 Mar 13;25(1):374. [doi: 10.1186/s12913-025-12427-z] [Medline: 40082928]
33. Uscher-Pines L, Sousa JL, Zachrison KS, Schwamm L, Mehrotra A. Financial impact of telehealth: rural chief financial officer perspectives. Am J Manag Care 2022 Dec 1;28(12):e436-e443. [doi: 10.37765/ajmc.2022.89279] [Medline: 36525663]
34. Schmeida M, McNeal R, Mossberger K. Policy determinants affect telehealth implementation. Telemed J E Health 2007 Apr;13(2):100-107. [doi: 10.1089/tmj.2006.0017] [Medline: 17489696]
35. Tomczyk Ł, Mascia ML, Gierszewski D, Walker C. Barriers to digital inclusion among older people: a intergenerational reflection on the need to develop digital competences for the group with the highest level of digital exclusion. Innoeduca Int J Technol Educ Innov 2023 Jun;9(1):5-26. [doi: 10.24310/innoeduca.2023.v9i1.16433]
36. Mohan R, Saleem F, Voderhobli K, Sheikh-Akbari A. Ensuring sustainable digital inclusion among the elderly: a comprehensive analysis. Sustainability 2024 Aug;16(17):7485. [doi: 10.3390/su16177485]
37. Panetta IC, Anjomrouz E, Paiardini P, Leo S. Digital by default? A critical review of age-driven inequalities in payment innovation. J Risk Financial Manag 2025;18(6):313. [doi: 10.3390/jrfm18060313]
38. Lund B. Well-being predicts digital literacy skills in rural older adults [Dissertation]. : Zenodo; 2022 URL: https://zenodo.org/records/18056171 [accessed 2026-03-02] [doi: 10.5281/zenodo.18056171]

XSL•FO
RenderX

39. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144. [doi: 10.2196/34144] [Medline: 35311682]
40. Badr J, Motulsky A, Denis JL. Digital health technologies and inequalities: a scoping review of potential impacts and policy recommendations. Health Policy 2024 Aug;146:105122. [doi: 10.1016/j.healthpol.2024.105122] [Medline: 38986333]
41. Graves JM, Abshire DA, Amiri S, Mackelprang JL. Disparities in technology and broadband internet access across rurality: implications for health and education. Fam Community Health 2021;44(4):257-265. [doi: 10.1097/FCH.0000000000000306] [Medline: 34269696]
42. Yu Y, Appiah D, Zulu B, Adu-Poku KA. Integrating rural development, education, and management: challenges and strategies. Sustainability 2024;16(15):6474. [doi: 10.3390/su16156474]
43. Koehle H, Kronk C, Lee YJ. Digital health equity: addressing power, usability, and trust to strengthen health systems. Yearb Med Inform 2022 Aug;31(1):20-32. [doi: 10.1055/s-0042-1742512] [Medline: 36463865]
44. Davidson R, Barrett DI, Rixon L, Newman S, ACT Program. How the integration of telehealth and coordinated care approaches impact health care service organization structure and ethos: mixed methods study. JMIR Nurs 2020;3(1):e20282. [doi: 10.2196/20282] [Medline: 34345792]
45. Torab-Miandoab A, Samad-Soltani T, Jodati A, Rezaei-Hachesu P. Interoperability of heterogeneous health information systems: a systematic literature review. BMC Med Inform Decis Mak 2023 Jan 24;23(1):18. [doi: 10.1186/s12911-023-02115-5] [Medline: 36694161]
46. Buljac-Samardzic M, Doekhie KD, van Wijngaarden JDH. Interventions to improve team effectiveness within health care: a systematic review of the past decade. Hum Resour Health 2020 Jan 8;18(1):2. [doi: 10.1186/s12960-019-0411-3] [Medline: 31915007]
47. Gellert GA, Rasławska-Socha J, Marcjasz N, et al. How virtual triage can improve patient experience and satisfaction: a narrative review and look forward. Telemed Rep 2023;4(1):292-306. [doi: 10.1089/tmr.2023.0037] [Medline: 37817871]
48. Zhang Z, Bundorf MK, Gong Q, Shea CM, Gilleskie D, Sylvia SY. Telehealth payment parity and outpatient service utilization: evidence from privately insured workers. Health Aff Sch 2025 Apr;3(4):qxaf068. [doi: 10.1093/haschl/qxaf068] [Medline: 40264702]

Abbreviations

CBT: cognitive behavioral therapy
IDEATel: Informatics for Diabetes Education and Telemedicine
LEAP: Lifestyle, Exercise, Attitudes, and Relationships Program
mHealth: mobile health
NIH: National Institutes of Health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROGRESS-Plus: place of residence; race/ethnicity, occupation, gender, religion, education, socioeconomic status, social capital, plus other context-specific determinants of health equity
PROPS: Population, Remote Health Interventions, Outcomes, Procedures, and Study Design
RoB: Risk of Bias
ROBINS-I: Risk Of Bias In Non-randomized Studies–of Interventions

Edited by Y Jiang; submitted 13.Jun.2025; peer-reviewed by I Adefolaju, M Peyroteo, ME Heidari, S Guan; revised version received 31.Oct.2025; accepted 09.Jan.2026; published 21.Apr.2026.

Please cite as:
Beltrán V, Chaple Gil A, Morales-Gómez C, Campos-Bijit V, Contador R, Díaz L, Bustos J, Fernandez E
Digital and Remote Health for Older Adults in Rural or Underserved Settings: Systematic Review of Clinical, Behavioral, and Implementation Outcomes
JMIR Aging 2026;9:e78913
URL: https://aging.jmir.org/2026/1/e78913
doi:10.2196/78913

© Víctor Beltrán, Alain Chaple Gil, Constanza Morales-Gómez, Vanessa Campos-Bijit, Rafael Contador, Leonardo Díaz, Jaime Bustos, Eduardo Fernandez. Originally published in JMIR Aging (https://aging.jmir.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effectiveness of Digital Interventions to Increase Preventive Care Uptake in Older Adults: Systematic Review

Lindsay Burton[1], MSc; Kathy L Rush[1], RN, PhD; Mindy A Smith[2], MD; Robert Janke[3], MLiS

[1]School of Nursing, University of British Columbia, 1147 Research Road, Kelowna, BC, Canada

[2]Department of Family Medicine, Michigan State University, East Lansing, MI, United States

[3]Library, University of British Columbia, Kelowna, BC, Canada

Corresponding Author:
Lindsay Burton, MSc
School of Nursing, University of British Columbia, 1147 Research Road, Kelowna, BC, Canada

Abstract

Background: Older adults face increasing health risks associated with aging and chronic disease; yet, uptake of recommended clinical preventive services remains low. Digital health interventions have the potential to enhance access and engagement, but their effectiveness in older adult populations remains unclear.

Objective: This systematic review aimed to examine the range and types of digital clinical preventive service interventions and assess their impact on preventive care uptake among older adults.

Methods: We conducted a systematic review of peer-reviewed research literature published since 2014. Eligible studies included experimental and quasi-experimental designs evaluating digital interventions targeting community-dwelling adults aged 60 years and older. Interventions focused on high-priority preventive services, including cancer screening and adult immunizations. Data were extracted using a standardized form and synthesized narratively due to heterogeneity in study designs and outcomes.

Results: In total, 24 studies involving over 1.3 million participants from 11 countries were included. Interventions used a range of digital tools, including telephone calls, SMS text messages, patient portals, and video-based education. While some digital and automated interventions demonstrated modest improvements in preventive services uptake, results were mixed. Interventions incorporating personalized elements (eg, tailored telephone counseling or in-person education) were generally more effective than generic, automated communications. Few studies reported on digital literacy support or intervention reach, and engagement with digital platforms was often low.

Conclusions: Digital interventions can support modest improvements in preventive services uptake among older adults, particularly when personalized or combined with human interaction. However, assumptions of digital fluency and limited reporting on engagement constrain generalizability. Future research should prioritize inclusive design, detailed reporting, and strategies that address digital equity to better support older adult populations.

(JMIR Aging 2026;9:e83446) doi:10.2196/83446

KEYWORDS

older adults; preventive care; digital health; systematic review; screening; vaccination

Introduction

Over the next decade, the percentage of Canadian people aged 65 years and older is projected to increase from 18.5% in 2021 to 22.9% by 2030. Similarly, the proportion of individuals aged 85 years and older is projected to increase from 2.3% in 2021 to 4.6% by 2050 [1]. Accelerated population aging, coupled with increasing chronic disease prevalence, has driven rising health care use and costs [2-5]. For instance, the average per-person health care spending for Canadian people aged 65+ years (US $9260) is 4 times higher than those <65 years (US $2083) [6]. Similarly, older adults in the United States represent 13.5% of the population but account for 45.2% of the top 10% of health care spenders [5]. The trend of population aging and estimates of increased costs highlight the urgent need for approaches that promote healthy aging, reduce chronic disease, and optimize quality of life for older adults.

Engaging older adults to stay current on recommended clinical preventive services is a proactive approach to promote healthy aging. Preventive services comprise activities aimed at disease prevention (eg, immunizations and alcohol or tobacco cessation) and early disease detection (eg, cancer screening, blood glucose, and blood pressure monitoring) [7]. Guidelines for recommended preventive services based on age, gender, and other risk factors have been developed using best evidence for cost-effectiveness and reduced morbidity and mortality [8,9]. However, despite national and provincial preventive services guidelines and recommendations, few older adults are up to

XSL•FO
RenderX

date on all recommended preventive activities. The number of recommended preventive services and their importance increases with age; yet, completion rates are alarmingly low for older adults. According to the British Columbia Cancer Agency, only 52% of women aged 50 - 74 years are screened for breast cancer compared to a target of 70% [10]. A US study estimated a completion rate for all relevant clinical preventive services of only 7.9% for older adults [11].

The rapid development and implementation of digital health services, including preventive services, during and after the COVID-19 pandemic have increased access and convenience [12] but have also had the untoward impact of excluding individuals with limited digital literacy. Although older adults are less frequent users of technology compared to younger age groups [13], evidence indicates that older adults' willingness to engage with digital technologies is growing steadily [14]. However, there remain challenges and barriers for older adults when accessing digital resources and health services (eg, lack of support and poor functionality) [15,16], which contribute to digital inequity and hinder older adult engagement in digital preventive services. Understanding the current state of the science regarding the effectiveness of digital interventions for preventive service uptake in older adults would provide valuable, evidence-based insights to better leverage technology in meeting their needs.

Although systematic reviews have synthesized evidence focused on individual preventive services, such as breast [17] and colon cancer [18] screening, no comprehensive review has been conducted looking specifically at older adults and digital aspects of preventive care services. As such, this review aims to answer this question: What are the range and types of digital clinical preventive services interventions (intervention), and what is their impact on older adults' (population) uptake of clinical preventive services (outcome)?

Methods

Design

This systematic review was designed following the JBI Manual for Evidence Synthesis [19,20] and the Cochrane Handbook for Systematic Reviews of Interventions [21].

Search Strategy

This systematic review captures peer-reviewed research papers addressing digital preventive services uptake targeting older adults. The digital tools considered included telephone (automated or synchronous), text messaging, web-based educational materials or materials delivered via video, or patient portal. The search strategy was developed in consultation with a health sciences librarian (RJ; see Multimedia Appendix 1 for the search strategy from Embase). Search terms were selected based on the review question's main content areas of older adults and clinical preventive services and included a combination of keywords and subject headings. Databases used in the search included MEDLINE, CINAHL, Embase, and PsycINFO. The search strategy was modified only slightly between the databases to incorporate their specific subject headings. Database searches were conducted on December 11, 2023, and updated on July 13, 2025.

Inclusion and Exclusion Criteria

Eligible literature included experimental (eg, randomized controlled trials) and quasi-experimental study designs. Qualitative and observational studies that did not test an intervention were excluded. Studies were included if they addressed preventive services uptake among older adults in response to single (eg, breast cancer screening, colorectal cancer, and blood pressure) or multiple services (eg, blood pressure, cancer screening, and seasonal vaccination). Interventions were limited to high-priority preventive care services based on high clinically preventable burden and established immunization schedules for older adults [8,22]. Preventive care services included 5 related to screening (cardiovascular disease, type 2 diabetes mellitus, hypertension, colorectal cancer, and breast cancer) and 6 related to vaccines (seasonal influenza, pneumococcal polysaccharide, shingles, COVID-19, and tetanus and or diphtheria).

Studies were eligible if they included samples that were primarily community-dwelling older adults. Because some studies did not report disaggregated age and used broad age categories (eg, "older adults" defined as 55+ years), we applied the following rules: studies in which all participants were aged 60 years or older were considered eligible; and if participants younger than 60 years were included, at least 50% of the sample had to be aged 65 years or older. This approach ensured consistency with our focus on older adults while allowing inclusion of studies with varied reporting practices. Studies were excluded if they focused on subpopulations (eg, familial risk and rheumatoid arthritis) or noncommunity-dwelling older adults (eg, long-term care residents) because they do not reflect the asymptomatic population for which preventive services guidelines are designed. Publications were limited to the last 10 years from when the review began to capture current strategies [9,22]. Studies were excluded if they did not report the results of primary research (ie, editorials, letters, search and research protocols, or systematic reviews).

Paper Selection

Two reviewers (LB and MAS) completed abstract screening and full-text screening, blinded to each other's decisions. Search results were added to the abstract management software Covidence for deduplication. A small sample of 10 - 15 records was tested against the stated inclusion and exclusion criteria [23]. The initial test of the screening process involved calibrating agreement, clarifying criteria, and making any necessary modifications before full screening. Following screening calibration, the remaining screening was conducted independently by the 2 reviewers (LB and MAS) [23]. Conflicts were assessed and resolved using consensus by 2 reviewers (LB and MAS). Interrater reliability was calculated using Cohen κ within Covidence software. The Cohen κ for abstract screening was 0.68, indicating substantial agreement [24].

Quality of Evidence

Quality of evidence was assessed using the Cochrane revised tool to assess risk of bias in randomized trials (Risk of Bias 2

XSL•FO
RenderX

[RoB 2]) [25] and the Risk of Bias in Non-randomized Studies—of Interventions (ROBINS-I) tool [26]. Both tools concentrate on the study's internal validity [25,26]. In both tools, bias refers to a tendency for the results to deviate systematically from what would be expected from a randomized trial, conducted on the same participant group with no flaws in its execution [25,26]. RoB 2 assesses 5 domains of bias: (1) randomization process, (2) deviations from intended interventions, (3) missing data, (4) measurement of outcomes, and (5) selection of reported results [25]. For the RoB 2 assessment, each domain is assigned a risk of bias assessment of low, some concerns, or high risk of bias [25]. ROBINS-I assesses six domains of bias: (1) confounding, (2) selection of participants, (3) classification of interventions, (4) missing data, (5) measurement of outcomes, and (6) selection of reported results [26]. For the ROBINS-I assessment, each domain is assigned a risk of bias assessment of low, moderate, serious, or critical [26]. An overall risk of bias assessment for each paper is then conducted for both RoB 2 and ROBINS-I tools based on the assessment of the domains [25,26]. Risk of bias assessments were conducted by 1 team member and verified by a second reviewer.

Data Extraction and Analysis

Data extraction was conducted using a standardized form. Elements for extraction included study characteristics, study population, intervention characteristics, type of preventive service targeted, single or multiple services, and change in preventive care uptake. Preventive services uptake was defined as the completion of a screening test or immunization, either by self-report, review of medical chart, or insurance claims. Data extraction was conducted by 1 research assistant and confirmed by 2 additional reviewers (MAS and LB).

Data were synthesized using a narrative approach to explore the effectiveness of digital preventive services interventions for older adults. This method was suitable, given the variability in designs and interventions among the studies analyzed [27]. Technology complexity was determined using a researcher-developed scale of low to high based on the degree of interactivity required of older adults to receive the preventive-services intervention [28]. Scale and rating criteria are detailed in Multimedia Appendix 2. The uptake of clinical preventive services was compared across different intervention strategies to identify potential patterns and insights. Data heterogeneity prevented meta-analysis.

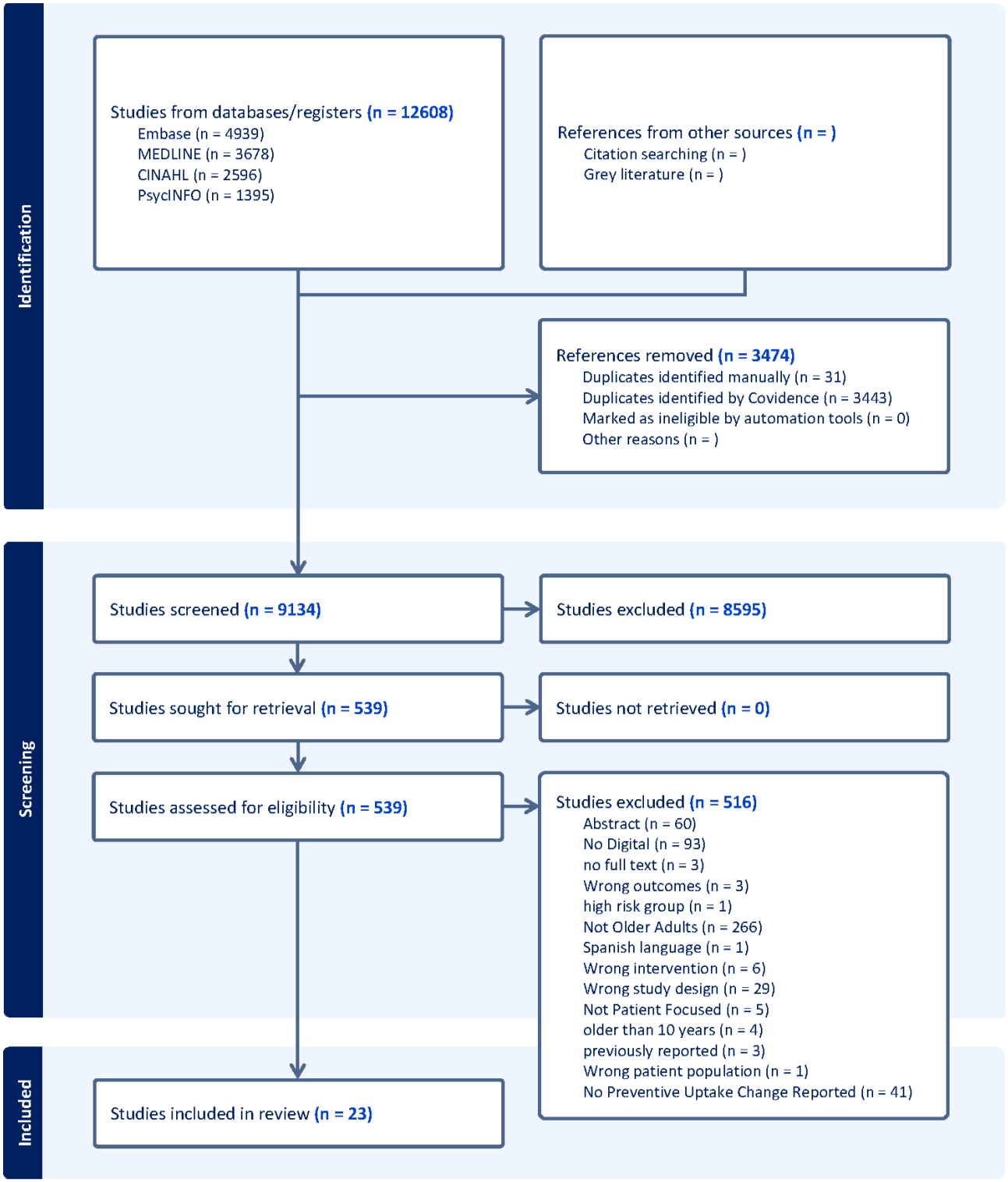
Results

Overview

The studies included in this review (n=23) represented 1,205,652 participants from 11 countries: the United States (n=11), Spain (n=3), Hong Kong (n=2), and 1 each from Australia, China, Denmark, France, Israel, Lebanon, and the United Kingdom. Studies were published between 2015 and 2025. A flowchart summarizing the search, screening, and study selection process is included in Figure 1 [29]. Designs included randomized controlled trials (n=20) and quasi-experimental studies (n=3). The majority (n=14) of included studies had a low risk of bias; the remaining had some or unknown risk of bias, but this was not a concern for the conclusions drawn by each study. The preventive services targeted in the interventions included colorectal cancer screening (n=10), influenza vaccination (n=6), breast cancer screening (n=3), pneumococcal vaccination (n=2), combined vaccinations (n=1), and COVID-19 vaccination (n=1; Table 1).

XSL•FO
RenderX

Figure 1. Summary of literature search and study selection process.



XSL•FO
RenderX

Table . Summary of included papers.

Study	Country	Design	Risk of bias	Sample	Targeted preventive care service	Simple intervention description	Change in preventive service uptake
Basch et al (2015) [30]	United States	RCT[a]	Low	564 participants, average age 57.6 (SD 4) years, 69.3% women, 17.6% White.	CRC[b] screening	• Comparison: CRC print education material • Intervention AD[c]: Comparison+AD • Intervention TTE[d]: Comparison, AD+targeted telephone education	Among those 60 years of age or older, AD plus telephone education group had a higher screening rate than the print group (27.3 vs 7.7%; *P*=.02).
Chan et al (2015) [31]	Hong Kong	RCT	Some concern—lack of detail of concealment	2517 participants aged >65 years, 57.3% women; ethnicity not reported.	Pneumococcal vaccination	• Comparison: Standard of care • Intervention: Education telephone intervention	1325 (53%) participants received the polysaccharide vaccine injection as recorded in the clinical management system—716 (57%) in the intervention group and 609 (48%) in the control group.
Denis et al (2017) [32]	France	RCT (3 arm)	Low	5691 participants, average age 61.2 (SD 6.8) years, 49.4% women; ethnicity not reported.	CRC screening	• Comparison: FOBT[e] kit • Intervention CATI[f]: CATI and FOBT kit • Intervention MI[g]: Motivational telephone interview and FOBT kit	Only 5691 (19.9%) people were actually counseled. There was no difference in completion between the intervention groups taken together and the control group (13.9% vs 13.9%) in intention-to-treat analysis.
Fernández et al (2015) [33]	United States	RCT (3 arm)	Some concern—outcome measure self-reported	656 participants, age 60 to 69 years: 31.8%, age 70+ years: 20.3%, 69.3% women, 100% Hispanic or Latino or Mexican.	CRC screening	• Comparison: No intervention • Intervention SMPI[h]: SMPI with DVD and flipchart • Intervention TIMI[i]: TIMI delivered on tablet computers	In the intention-to-treat analysis, 13.6% in the SMPI group, 10.2% in the TIMI group, and 10.8% in the comparison group completed CRC screening (not statistically significant).

XSL•FO
RenderX

Study	Country	Design	Risk of bias	Sample	Targeted preventive care service	Simple intervention description	Change in preventive service uptake
Fleming et al (2018) [34]	United States	Quasi-experimental design	Moderate concern	3415 participants, average age 60.0 (SD 6.3) years, 50.5% women, 32.8% White.	CRC screening	• Comparison: Mass-mailing FIT[j] kit • Intervention phone: Comparison+telephone or recorded message • Intervention in-person: Comparison+one-on-one conversations in clinic	One-on-one conversations either in person (OR[k] 24.63, 95% CI 19.28 - 31.46) or via telephone or recorded message (aOR 14.74, 95% CI 10.96 - 19.82) were more effective for completing the at-home CRC screening than mass mailing before receiving the FIT.
Ghadieh et al (2015) [35]	Lebanon	RCT (6 arm)	Low	350 participants aged 60+ years not receiving the vaccine, 54.9% women; ethnicity 100% Chinese.	Pneumococcal vaccination	• Intervention: Telephone reminder±education • Intervention: Email reminder±education • Intervention: SMS text message reminder±education	Vaccination rates increased to 16.5% in short phone call group, 7.2% in SMS text group, and 5.7% in the email group; from 17.2% to 20.4% in patients older than 65 years.
Hagoel et al (2016) [36]	Israel	RCT	Low	48,091 participants, average age 60.44 (SD=6.04) years, 51.1% women; ethnicity not reported.	CRC screening	• Comparison: No intervention • Intervention interrogative: Interrogative SMS text message reminders±social context • Intervention noninterrogative: Noninterrogative SMS text message reminders±social context	At 6 months, interrogative (9.8%; *P*=.002), interrogative-with-social-context (10.3%; *P*<.001), and noninterrogative-with-social-context (9.6%; *P*<.009) messages had greater uptake compared to no messages in control (8.5%).
Hirst et al (2017) [37]	United Kingdom	RCT	Low		CRC screening		

XSL•FO
RenderX

Study	Country	Design	Risk of bias	Sample	Targeted preventive care service	Simple intervention description	Change in preventive service uptake
				176,231 participants, 43.5% aged 60 - 70+ years, 53% women, 51% White.		• Comparison: Usual care • Intervention: SMS text message reminders	Uptake was 40.5% in the intervention group and 39.9% among controls. Uptake did not differ significantly between groups of older adults (aOR 1.03, 95% CI 0.94 - 1.12; *P*=.56) but did vary for first-time invitees (uptake 40.5% in intervention vs 34.9% in the control group; aOR 1.29, 95% CI 1.04 - 1.58; *P*=.02).
Hurley et al (2018) [38]	United States	Pragmatic RCT	Low	8269 participants, median age 66 years, 52% women; ethnicity not reported.	Influenza or tetanus, diphtheria, acellular pertussis vaccine	• Comparison: Usual care • Intervention: Up to 3 reminders or recalls	The intervention was associated with receipt of any needed vaccine in the adults aged ≥65 years (aOR 1.15, 95% CI 1.02 - 1.30).
Jiang et al (2022) [39]	China	RCT	Low	15,769 total participants, 616 aged 65+ years, 63.3% women; ethnicity not reported.	Influenza vaccination	• Comparison: No intervention • Intervention: Video-led educational intervention	Intervention group more willing to receive influenza vaccination than control group (64.6% vs 51.4%; *P*<.05). Vaccination uptake rate was higher in the intervention group (10.3% vs 3.4% among controls; OR 3.23, 95% CI 1.25 - 8.32; *P*<.001).
Lieu et al (2022) [40]	United States	RCT (3 arm)	Low	8287 participants not receiving the vaccine after prior outreach, average age 72.6 (SD 7) years, 56% women 100% Black or Latino.	COVID-19 vaccinations	• Comparison: Usual care • Intervention: Electronic and/or mail outreach from the provider with or without culturally tailored content	Higher COVID-19 vaccination rates for intervention with cultural content versus usual care (24% vs 21.7%; aHR[m] 1.22, 95% CI 1.09 - 1.37), and for primary care provider outreach without cultural content (23.1% vs 21.7%; aHR 1.17, 95% CI 1.04 - 1.31).

XSL•FO
RenderX

Study	Country	Design	Risk of bias	Sample	Targeted preventive care service	Simple intervention description	Change in preventive service uptake
López-Torres Hidalgo et al (2016) [41]	Spain	RCT (4 arm)	Some concern—self-reported outcome	1129 participants, average age 61.4 (SD 6.6) years, 60.1% women; ethnicity not reported.	CRC screening	• Comparison: No intervention • Intervention: Written, telephone, or face-to-face recommendation of CRC screening	Uptake was 15.4% (OR 2.32, 95% CI 1.23 - 4.37) in the written information group, 28.8% (OR 4.38, 95% CI 2.38 - 8.07) in the telephone information group, and 8.1% (OR 1.26, 95% CI 0.59 - 2.68) in the face-to-face group versus 5.9% in the control group.
Luckmann et al (2019) [42]	United States	RCT (3 arm)	Some concern—unclear description of randomization	30,160 participants, 52.3% aged 60 - 84 years, 100% women, 73% White.	Breast cancer screening	• Intervention letter: Reminder letter only • Intervention reminder: Letter plus reminder call • Intervention counseling: 2 letters and tailored education and MI	Over 4 years, mammography adherence was highest with a reminder call (83%) versus counseling (80.8%) or letter only (80.8%), *P*=.03.
Luque Mellado et al (2019) [43]	Spain	Quasi-experimental study	Moderate concern	2343 participants, 53.4% women; age and ethnicity not fully reported.	CRC screening	• Telephone interview	Participation increased with age for both genders (60 - to 64-year group: OR 1.55, 95% CI 1.35 - 1.78 and 64 - to 69-year group: OR 2.07, 95% CI 1.80 - 2.38).
Modin et al (2023) [44]	Denmark	RCT (9 arm; usual care or 9 electronic letters with designs based on behavioral concepts)	Low	964,870 participants, average age 73.1 (SD 6.2) years, 27.4% had CVD[n], 52% women; ethnicity not reported.	Influenza vaccination	• Comparison: Usual care • Intervention arms: 9 different electronic nudges emphasizing the potential cardiovascular benefits of influenza vaccination	

XSL•FO
RenderX

Study	Country	Design	Risk of bias	Sample	Targeted preventive care service	Simple intervention description	Change in preventive service uptake
							Influenza vaccinations were received by 83.1% of participants with CVD versus 79.2% of participants without CVD (*P*<.001). Compared with usual care, a letter intervention increased vaccination rates for those with (absolute difference,+0.60 percentage points; 99.55% CI –0.48 to 1.68) and without CVD (+0.98 percentage points; 99.55% CI 0.27 - 1.70; *P* for interaction .41). Only the repeated letter strategy with a reminder follow-up letter 14 days later was effective in increasing influenza vaccination, irrespective of CVD.
Regan et al (2017) [45]	Australia	RCT	Low	12,354 total participants, 3613 aged 65+ years, 52.6% women, 48.6% non-Indigenous.	Influenza vaccination	• Comparison: No reminder • Intervention: SMS text message reminders	For those ages 65+ years, 20.5% in the intervention group versus 15.8% in the control group were vaccinated (absolute difference +4.7; RR[o] 1.26, 95% CI 1.10 - 1.45; NNT[p] 21).
Sepucha et al (2023) [46]	United States	RCT	Some concern—not all intervention participants received decision coaching	800 participants, average age 60 (SD=8) years, 53% women, 73% White.	CRC screening	• Comparison: Usual care • Intervention: Decision aids for shared decision-making (SDM) and decision coaching	Intervention respondents reported higher SDM scores (mean difference 0.7; 95% CI 0.4 - 0.9; *P*<.001) and less decisional conflict than controls (−21%; 95% CI −35% to −7%; *P*=.003).
Szilagyi et al (2020) [47]	United States	RCT (4 arm)	Low		Influenza vaccination		

XSL•FO
RenderX

Study	Country	Design	Risk of bias	Sample	Targeted preventive care service	Simple intervention description	Change in preventive service uptake
				164,205 participants, 18.7% aged 65+ years, 58% women, 57% White.		• Comparison: No portal reminder • Intervention arms: Patient portal reminders provided one, two, or three times	The adjusted risk ratio for influenza vaccination among individuals aged 65 years and older was 0.97 (95% CI 0.92 - 1.02), with no dose response by number of reminders sent for all age groups.
Szilagyi et al (2022) [48]	United States	RCT (6 arm)	Low	196,486 total participants, 29,795 aged >65 years without diabetes, 58.5% women, 56.9% White.	Influenza vaccination	• Comparison: No messages • Intervention arms: Personalized messages (5 types around loss or gain frame) sent by a health system's patient portal or reminder alone	Influenza vaccination rates for older adults were 55.6% with no statistically different vaccination rates for any study group versus control.
Venishetty et al (2023) [49]	United States	Prospective, pragmatic, two-arm intervention study	Low	344 participants, average age 60.1 (SD 5.5) years, 68% women, 89% White.	CRC screening	• Comparison: Mailed FIT kit and follow-up call • Intervention: Mailed FIT kits with introductory phone calls by a patient navigator or community health worker; there was variability in FIT kits distributed by health clinics	Overall, 25.3% of participants returned the FIT kit. The return rate was higher for the intervention group versus the usual care group (27.8% vs 22.8%; *P*=.003). Participants who received a 1-day sampling kit had a higher return rate (31.3% vs 23.3%).
Vives et al (2024) [50]	Spain	RCT	Low	10,433 participants, average age 60.3 (SD 5.3) years.	Breast cancer screening		

XSL•FO
RenderX

Study	Country	Design	Risk of bias	Sample	Targeted preventive care service	Simple intervention description	Change in preventive service uptake
						• Comparison: Invitation letter+SMS text message reminder • Intervention: SMS text message invitation to mammography with a timed appointment+SMS text message reminder	No difference in the screening participation rate within 12 weeks of the invitation in the SMS text message group (86.6%) and in the letter group (87.3%).
Wang et al (2023) [51]	Hong Kong	RCT	Some concern—self-reported outcome	396 participants, average age 70.2 (SD=4.3) years, 63% women; ethnicity not reported.	Influenza vaccination	• Comparison: Standard online video • Intervention: Chatbot-delivered online intervention tailored to the stage of change once every 2 weeks for 4 sessions	Influenza vaccination rate was higher in the intervention group than the control group at month 6 (50.5% vs 35.3%; *P*=.002).
Wu and Lin (2015) [52]	United States	RCT	Some concern—self-reported outcome	193 participants, 100% women, 100% Chinese American.	Breast cancer screening	• Comparison: Mammography brochure • Intervention: Tailored telephone counseling to personal barriers	For women aged ≥65 years, there was a significant difference in uptake between intervention and control groups (51% vs 25%).

[a]RCT: randomized controlled trial.

[b]CRC: colorectal cancer.

[c]AD: academic detailing.

[d]TTE: tailored telephone education.

[e]FOBT: fecal occult blood test.

[f]CATI: computer-assisted tailored telephone interview.

[g]MI: motivational interview.

[h]SMPI: small media print intervention.

[i]TIMI: tailored interactive multimedia intervention.

[j]FIT: Fecal Immunochemical Test.

[k]OR: odds ratio.

XSL•FO
RenderX

[l]aOR: adjusted odds ratio.

[m]aHR: adjusted hazard ratio.

[n]CVD: cardiovascular disease.

[o]RR: relative risk.

[p]NNT: number needed to treat.

Digital tools used included the single or combined use of telephone calls, text messages, patient portal messages, email, and videoconferencing. No studies included a description of digital tool support or training for participants. Under half of the studies (n=9) included a description of the digital intervention's reach (ie, messages opened and calls answered; Table 1).

Range of Interventions

The analysis revealed a wide range of approaches aimed at improving patient engagement and adherence to preventive health screenings and vaccinations. Three main categories of interventions were classified: digital and automated communication, personalized and tailored communication, and in-person interventions supplemented with digital supports.

Digital and Automated Communication Interventions

To increase vaccination and screening rates, 12 studies tested interventions that used automated digital communication tools including electronic health record portal messages, interactive automated calls, automated text messaging, and 1 chatbot-integrated messaging intervention [32,35-38,40,44,45,47,48,50,51]. Across the 12 included trials, most interventions used low-complexity, system-integrated technologies, with a smaller subset relying on simple stand-alone channels (Table 2). Results were mixed, with only 4 interventions demonstrating significant improvements in vaccination or screening rates compared to usual care, as detailed below.

XSL•FO
RenderX

Table . Complexity of technology for digital and automated communication interventions.

Study	Targeted preventive care service	Simple intervention description	Technology	Digital intervention reach	Patient facing complexity
Denis et al (2017) [32]	Colorectal cancer screening	Tailored telephone—computer-assisted counseling or motivational interview	Telephone	33.6% of telephone counseling attempts were technically successful	Low
Ghadieh et al (2015) [35]	Pneumococcal vaccination	Patient reminder system	Telephone, email, or SMS text messages with or without education	2% (121/6,177) SMS transmission failures	Low
Hagoel et al (2016) [36]	Colorectal cancer screening	SMS text message reminders	SMS text messages	Not specified	Low
Hirst et al (2017) [37]	Colorectal cancer screening	Text message reminders	Text message	Mobile phone coverage varied by demographics: 42.7% - 53.3% had registered numbers; 73.4% (1,023/1,393) successfully received text reminders	Low
Hurley et al (2018) [38]	Influenza or tetanus, diphtheria, acellular pertussis vaccine	Centralized reminder or recall	Telephone	Not specified	Low
Lieu et al (2022) [40]	COVID-19 vaccinations	Electronic and/or mail outreach from provider with or without culturally tailored content or usual care	Electronic secure messages	Not specified	Moderate
Modin et al (2023) [44]	Influenza vaccination	Electronic nudges emphasizing potential cardiovascular benefits of influenza vaccination	Electronic letters	Not specified	Low
Regan et al (2017)[45]	Influenza vaccination	SMS text message reminders	SMS text messages	75.5% of eligible patients had mobile phone numbers in the EMR[a]	Low
Szilagyi et al (2020) [47]	Influenza vaccination	Patient portal reminders provided 1, 2, or 3 times or none (control)	Online patient portal messages	52.9% to 58.8% of intervention participants opened at least 1 portal reminder	Moderate
Szilagyi et al (2022) [48]	Influenza vaccination	Personalized messages (5 types around loss or gain frame) sent by a health system's patient portal or reminder alone	Online patient portal messages	Not specified	Moderate
Vives et al (2024) [50]	Breast cancer screening	SMS text message invitation to mammography with a timed appointment	SMS text message	Not specified	Low
Wang et al (2023) [51]	Influenza vaccination	Chatbot-delivered online intervention tailored to the stage of change once every 2 weeks for 4 sessions or standard online video	Chatbot-delivered online intervention	Text messages failed to be delivered to 92 (1.7%) women in intervention	Low

[a]EMR: electronic medical record.

Of the 4 studies aimed at cancer screening, only 1 significantly improved uptake compared to usual care. Hagoel et al [36] conducted a randomized experiment comparing 4 text messages with different theory-based content to no messages (control)

XSL•FO
RenderX

and found that 3 of the 4 messages significantly increased colorectal cancer screening rates compared to the control. Conversely, 2 interventions that used reminders, text messages [37], or phone calls [32], along with mailed Fecal Immunochemical Test (FIT) kits, found no difference in intervention groups compared to usual care (FIT kit plus 1 reminder letter). A breast cancer screening intervention found no difference between text message (intervention) and letter (control) invitations [50].

In 5 studies aimed at increasing seasonal influenza vaccination uptake rates, only marginal increases were observed. Even when increases were observed, absolute increases from baseline for different automated letters were small (0.73 and 0.89 percentage points, respectively) [44]. In the 2 studies examining generic reminders through a patient portal or via text messages, only marginal, nonsignificant increases in vaccination rates were found from baseline [45,47]. Another large-scale study used reminder messages sent through the health system patient portal, tailored to the health belief model loss or gain framework, to promote influenza vaccination uptake [48]. They similarly reported no significant increase in vaccination rates among older adults compared to generic digital reminders [48]. However, in 1 theory-based intervention that used a chatbot to deliver promotional videos corresponding to stages of change theory compared to generic messages, investigators found a significant between-group difference in vaccination rates at 6 months (50% of the intervention group vs 35% of the control group; P=.002) [51].

In total, 3 studies that examined interventions designed to increase the uptake of vaccinations showed small but statistically significant increases. One study compared 3 different reminder modalities to encourage pneumococcal vaccination: a standardized phone call reminder from a nurse, text messaging, or email [35]. Vaccination rates increased from 17.2% to 20.4%, with the largest uptake observed in the phone call group (16.5%), compared to text messaging (7.2%) and email (5.7%) [35]. Hurley et al [38] used a centralized vaccine reminder system for pneumococcal, tetanus, and influenza vaccinations, with the intervention group receiving up to 2 auto-dial phone calls followed by a postcard. Findings showed a significant increase in receiving any of the 3 targeted vaccines in the intervention arm (33.2%) compared with no centralized reminders (29.8%; odds ratio [OR] 1.20, 95% CI 1.06-1.36; P≤.01) [38]. Similarly, receiving either culturally tailored or standard electronic messages from primary care providers modestly increased COVID-19 vaccination uptake compared to no outreach from primary care providers (24% and 23.1% vs 21.7%; P<.01) [40].

Personalized and Tailored Communication Interventions

In total, 7 interventions examined tailored telephone calls to promote preventive services uptake [30,41-43,46,49,52]. Across the 7 interventions, most interventions used low-complexity approaches, primarily relying on manual telephone contact (Table 3). Only 4 of the tailored telephone interventions demonstrated statistically significant improvements in screening rates over usual care.

3 of the studies evaluating tailored or reminder-based interventions found largely nonsignificant improvements in screening uptake overall, though some age-related differences were reported. In one study, investigators used results from a baseline survey on knowledge and barriers to mammography to tailor messages to participants [52]. At the 4-month follow-up, although nonsignificant differences were found among younger participants, older women in the intervention group had significantly higher rates of obtaining a mammogram compared to controls (51% vs 25%) [52]. Luckmann et al [42] compared 2 interventions to a comparison group: a reminder letter with multiple reminder calls or reminder calls with tailored counseling, compared with only a reminder letter. Absolute difference in mammography adherence in the reminder call only intervention group compared to the other arms increased significantly in the 50 to 74 years age range by 3.3% but not in the 75 to 84 years age range [42]. However, the change in mammography adherence from baseline was not significant in either age group for either trial arm [42]. Another intervention to increase colorectal screening compared structured telephone calls after 2 reminder letters to no telephone call [43]. Although overall screening participation was 54.9%, and screening adherence was higher in older age groups than in younger ones (Table 1), there was no between-group difference [43].

XSL•FO
RenderX

Table . Complexity of technology for personalized and tailored communication interventions.

Study	Targeted preventive care service	Simple intervention description	Technology	Digital intervention reach	Patient-facing complexity
Basch et al (2015) [30]	Colorectal cancer (CRC) screening	CRC print education material, academic detailing to primary care providers, and targeted telephone education for patients	Printed brochures +telephone calls (staff-delivered)	Median 5 calls per patient	Low
López-Torres Hidalgo et al (2016) [41]	CRC screening	Written information, telephone call, or face-to-face meeting delivering CRC screening information; control group received no information	Telephone (voice call)	Not specified	Low
Luckmann et al (2019) [42]	Breast cancer screening (mammography)	Call with tailored education and motivational interviewing (MI)	Telephone (voice call)	Only 23.5% of women were reached and accepted MI	Low
Luque Mellado et al (2019) [43]	CRC screening	Structured telephone interview (scripted calls) to nonresponders after first invitation round	Telephone (voice call)	68.8% (1614/2343) contact rate	Low
Sepucha et al (2023) [46]	CRC screening	Telephone decision-coaching call	Telephone (voice call)	54% (214/399) of intervention arm reached by decision coaches	Low
Venishetty et al (2023) [49]	CRC screening	Mailed FIT[a] kits with introductory phone calls by a patient navigator or community health worker; there was variability in FIT kits distributed by health clinics	Telephone	Not specified	Low
Wu and Lin (2015) [52]	Breast cancer screening	Tailored messages to personal barriers	Telephone	Not specified	Low

[a]FIT: Fecal Immunochemical Test.

In contrast to the largely clinically nonsignificant results, 4 of the studies reported significant differences. Basch et al [30] supplemented print education materials with tailored telephone calls aimed at identifying and addressing colorectal screening barriers and encouraging follow-up with a primary care provider. Among those aged 60+ years who had visited a primary care provider during the intervention period, the intervention group was almost 5 times more likely to receive screening compared to print education material alone (36.4% vs 7.7%; χ^2_1=7.3; *P*=.01) [30]. Similarly, Venishetty et al [49] compared usual mailed FIT kits with an introductory phone call from a patient navigator, who introduced them to the program and its contents. The return rate for the usual care plus introductory call arm was significantly higher (27.8%) than that for the usual care arm (22.8%; *P*=.003) [49]. Another intervention on colorectal cancer screening found that a decision worksheet paired with a call from a decision coach significantly improved screening rates [46]. Despite only half (54%) of the intervention group having been reached by the decision coach, the intervention arm had higher screening rates than the control arm (35% vs 23%; *P*<.001) [46]. Another intervention compared written, telephone, and face-to-face tailored communication with a control group to improve colorectal cancer screening [41]. Interestingly, both written information and telephone communication, but not face-to-face communication, showed significantly greater screening uptake compared to the control group (Table 1) [41].

In-Person Interventions Supplemented With Digital Supports

In total, 4 interventions supplemented in-person preventive care education with digital tools [31,33,34,39]. Across these 4 trials, technology use was uniformly simple, with only 2 studies extending into low-moderate complexity (Table 4). In their study, Chan et al [31] found a 3-minute brief health education telephone intervention approximately 1 week before a 3-minute face-to-face health education intervention during participants' medical appointments increased vaccination uptake by 57% compared to the 9% increase for the control group that received standard of care (promotional leaflets and poster displays). A 3-modal intervention study found that compared to mass-mailing, an in-person conversation was the most effective (OR 24.63, 95% CI 19.28 - 31.46), followed by a telephone

XSL•FO
RenderX

one-on-one conversation (OR 14.74, 95% CI 10.96 - 19.82) [34]. A comparison study of a video-based intervention for influenza vaccination supplemented with a 15-minute small group in-person discussion with medical students, delivered in health care centers to older adults, improved vaccination uptake (10.3% increase in the intervention group and 3.4% in the control group; Table 1) [39]. Conversely, a tablet-delivered, tailored interactive multimedia intervention delivered in person by a lay health worker that provided individualized information based on participant responses did not increase colorectal cancer screening uptake compared with a small-media print intervention or no intervention [33].

Table . Complexity of technology for in-person interventions supplemented with digital supports.

Study	Targeted preventive care service	Simple intervention description	Technology	Digital intervention reach	Patient facing complexity
Chan et al (2015) [31]	Pneumococcal vaccination	Brief health-education telephone intervention (3 minutes) delivered by nurses before and during the clinic visit	Telephone (voice call)	Not specified	Low
Fernández et al (2015) [33]	CRC[a] screening	Tailored interactive multimedia intervention delivered on a tablet computer	Tablet-based interactive program	Not specified	Low-moderate
Fleming et al (2018) [34]	CRC screening	One-on-one conversations (in clinic, telephone, or recorded message) to promote FIT[b] completion	Telephone (voice call)	Not specified	Low
Jiang et al (2022) [39]	Influenza vaccination	Video-led educational intervention	Educational video	Not specified	Low

[a]CRC: colorectal cancer.

[b]FIT: Fecal Immunochemical Test.

Use of Theory

Of the 23 studies, 9 incorporated theoretical frameworks into their intervention development or design. These studies applied a variety of behavioral theories that focused on intentions, beliefs, motivations, readiness, economics, stages of change, and osteopathic principles. The primary purpose of using theory in the design of interventions was to inform the content and framing of messages (Table 5). Among these studies, 6 of 9 reported significant increases in uptake in the intervention groups, with 3 demonstrating clinically important differences. In contrast, only 4 of 13 atheoretical studies reported a statistically significant increase in preventive care uptake.

XSL•FO
RenderX

Table . Use of theory.

Author (year)	Theory	Theory use	Intervention development description
Fernandez et al (2015) [33]	Fishbein's integrated model of behavior change	Message content and framing	The study used behavioral theory to identify psychosocial factors influencing colorectal cancer (CRC) screening, which were then targeted in intervention, select behavior change methods, and develop tailored messages or content.
Fleming et al (2018) [34]	Osteopathic principles	Intervention delivery	The study applied osteopathic principles by designing the intervention to integrate patients as partners in their own care through one-on-one conversations, education, and motivational strategies.
Hagoel et al (2016) [36]	Question-behavior effect	Message framing	The 4 base intervention messages (reminders for CRC screening) were manipulated in terms of grammatical form (interrogative vs noninterrogative) and social context reference (yes or no).
Lieu et al (2022) [40]	Behavioral science	Message content	Outreach messages were crafted using behavioral science principles, including social proof, loss aversion, and authority figures.
Luckmann et al (2019) [42]	Precaution adoption process model	Message content and framing	The intervention counselors assessed which of the 5 stages of readiness each woman occupied: unaware, decided against, undecided, planning, or scheduled. Based on this, they delivered scripted and tailored education to specifically address her stage and encourage timely mammograms.
Szilagyi et al (2020) [47]	Health belief model	Message content	The development of portal reminder letter content drew on the health belief model to target key psychological drivers of vaccination, used health literacy principles to create accessible messages, applied behavioral economics to enhance motivation, and integrated patient or expert feedback to optimize content.
Szilagyi et al (2022) [48]	Health belief model	Message content and framing	The portal messages were structured following the health belief model, which focuses on individuals' perceptions about health risks (like susceptibility to influenza), perceived benefits of action (vaccination), perceived barriers, and cues to action. Messages were also modeled on precommitment and gain-loss. Precommitment messages asked patients if they planned to get the influenza vaccine, aiming to increase the likelihood of follow-through. Gain-loss messages were crafted to emphasize either the positive outcomes of vaccination (gain frame) or the negative outcomes of not vaccinating (loss frame).

XSL•FO
RenderX

Author (year)	Theory	Theory use	Intervention development description
Wang et al (2023) [51]	Transtheoretical model—stages of change (SOC)	Message content and framing	The intervention started with a chatbot assessment of each participant's stage of change regarding their intentions and plans to receive the vaccine. Participants then received a video tailored to their SOC every 2 weeks for 4 sessions. Their SOC was reassessed each time, allowing the intervention to stay relevant as their readiness evolved.
Wu and Lin (2015) [52]	Health belief model	Message content and framing	The telephone counseling protocol systematically began by assessing each woman's stage of mammography readiness using health belief model constructs (perceived susceptibility, perceived benefits, and perceived barriers), then delivered messages and action steps aligned with her specific needs and beliefs.

Discussion

Principal Findings

Overall, the findings from this review demonstrate that digital interventions for promoting preventive screenings and vaccinations in older adults yield mixed results, with effectiveness varying widely depending on the intervention. While certain strategies showed substantial improvements in uptake, the majority of digital or automated reminders (eg, generic text messages and portal alerts) resulted in only marginal or no increases in screening and vaccination rates.Among the interventions reviewed, those incorporating personalized elements, such as tailored telephone counseling or supplementary in-person education, were consistently more effective than impersonal, automated communications. This suggests that while automation can facilitate uptake, human interaction remains critical for promoting preventive health behaviors in older adults.

Personalization of Digital and Automated Interventions

Digital and automated communication tools, used in nearly half of the studies, can yield statistically significant improvement in uptake of preventive services among older adults. Older adults appear to respond well to generic approaches that notify and send reminders regarding their preventive activities and screenings. However, the effectiveness of simple, generic reminders is generally limited. These outcomes align with earlier systematic reviews and behavior change literature, which also report modest behavior change impacts from brief digital interventions, especially when applied generically to the older adult population [53].

Evidence from this work suggests that interventions incorporating personalized elements, such as tailored telephone counseling or supplementary in-person education, are generally more effective than impersonal, automated communications. This echoes findings from studies on preventive services and digital health adoption, which suggest that human interaction, particularly when delivered by a trusted source, such as a primary care provider, is a critical facilitator of successful behavioral change in older adults [54,55]. The preference for messages from known, trusted sources and for personalized content reflects broader trends in health communication and aligns with user preference studies, which show higher acceptability for content that has a connection to both their care and provider [56,57]. Further, studies have identified the personalization of messages to be an important condition for patient engagement [56,57].

Despite the importance of tailoring interventions to the user's context and to engage them effectively [58], there are challenges related to personnel cost, time, and effort with this approach. However, emerging evidence suggests that artificial intelligence (AI) may offer scalable solutions by automating personalization and tailoring at a scale for low cost [59]. One opportunity, such as AI chatbots, is able to automate interactions while also simulating interactive and empathic dialogue [60,61]. Chatbots trained with motivational interviewing techniques have shown promise in supporting smoking cessation [60,62], addressing vaccine hesitancy [61], and promoting behavior change [63]. However, AI-driven personalization carries important limitations and risks: models can amplify existing biases and worsen inequities if trained on unrepresentative data, produce clinically inaccurate or outdated recommendations without continuous validation, and generate "black box" outputs that are difficult to interpret or explain [64]. Additional research is needed to evaluate the comparative effectiveness, acceptability, and ethical implications of AI-based personalization in diverse populations.

Technology Complexity and Appropriateness

The appropriateness of digital health interventions for older adults is strongly influenced by technology complexity, digital literacy, and the availability of support [65]. While most studies included had low complexity, a consistent limitation across moderately complex digital interventions was the lack of attention to training needs, technical support, or onboarding required for older adults to effectively use digital health platforms such as patient portals, SMS text message reminders,

XSL•FO
RenderX

or email notifications. The assumption of a baseline digital fluency may inadvertently perpetuate ageism and disadvantage those with weaker digital skills, compounding barriers to access and potentially widening health disparities [15]. Studies that evaluated engagement metrics frequently reported low rates of message opening or system use, with only about half of older adult recipients successfully engaging with portal messages or receiving SMS text message reminders, highlighting a significant constraint on the intervention's reach. One recent survey has shown that while many older adults prefer patient portals and email, substantial proportions still rely on traditional methods (hospital websites and telephone calls), with significant numbers excluded by digital-only approaches, especially those aged 75+ years [55]. However, there is an increasing demand for patient portal communication in the population aged 65+ years [66]. This suggests that multimodal, accessible communication is critical for maximizing engagement in older populations.

Use of Theory

Notable in the digital health intervention studies included in this review was the use of theory in intervention design. Approximately half of the included studies explicitly referenced theory, most commonly behavior change, primarily to inform the content and framing of message delivery. While theory use was described in 9 studies, the depth of integration in intervention design is not explicit in all study descriptions. This is similar to critiques raised in a related cardiovascular preventive care research study, where the rigor and depth of theory application were often unclear or inadequately reported, making it difficult to determine whether theory guided the overall intervention architecture or was limited to message wording [67]. This risk of "surface-level" use of theory, where theoretical constructs inform only minor aspects, such as phrasing or framing, rather than the full intervention pathways, has been documented as an ongoing issue in behavioral health research [68,69]. Fewer than half of published behavior change interventions explicitly cite a guiding theory, making it difficult to determine how theory actually contributed to intervention outcomes. The use of theory to guide the studies would have afforded greater explanatory value to the findings, in particular, offering insights into the largely mixed findings that were observed.

Limitations

This review has several limitations that warrant consideration. The diversity of study designs, populations, and intervention types prevented meta-analysis, and findings were instead synthesized narratively. While this approach allowed for exploration of patterns across varied contexts, it limits the ability to quantify pooled effects. Data extraction was conducted by a single reviewer and verified by 2 others, which may introduce bias, as dual independent extraction is considered best practice [21]. Many studies lacked consistent control or comparison groups, and several used complex, multicomponent interventions, making it difficult to isolate the contribution of individual digital modalities. Although all included studies had a mean age of at least 60 years, many were not specifically designed for older adults and also targeted younger age groups. Finally, the review was limited to peer-reviewed, English-language publications, which may have excluded relevant findings from non-English–speaking regions.

Conclusions

While digital communication tools can support modest increases in preventive care uptake among older adults, their effectiveness appears to depend on the degree of personalization. This systematic review's findings highlight that automation alone may be insufficient and that interventions facilitating patient engagement or combining digital outreach with practical supports may be more likely to achieve meaningful improvements in screening and vaccination rates.

Acknowledgments

The authors are grateful for the screening and data extraction support from Sofia Knopf, Mayuran Thavaraj, and Victor Wang. The authors thank Dr. Kristen Haase and Dr. R. Colin Reid for proofreading and editorial review.

Funding

Funding support for this study was provided by the Canadian Institutes of Health Research Planning and Dissemination Grant (#GR027890).

Authors' Contributions

Conceptualization: LB, KLR
Data curation: LB
Formal analysis: LB (lead), KLR (supporting), MAS (supporting)
Funding acquisition: LB, KLR
Methodology: LB, RJ, KLR, MAS
Project administration: LB
Supervision: KLR
Writing—original draft: LB (lead), KLR (supporting)
Writing—review and editing: LB, KLR, MAS, RJ

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Embase search strategy.
[DOCX File, 66 KB - aging_v9i1e83446_app1.docx]

Multimedia Appendix 2
Technology complexity assessment criteria.
[DOCX File, 16 KB - aging_v9i1e83446_app2.docx]

Checklist 1
PRISMA checklist.
[PDF File, 141 KB - aging_v9i1e83446_app3.pdf]

References

1. A portrait of Canada's growing population aged 85 and older from the 2021 census. : Statistics Canada; 2022 URL: https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-x/2021004/98-200-x2021004-eng.cfm [accessed 2026-04-07]
2. Cantarero-Prieto D, Pascual-Sáez M, Lera J. Healthcare utilization and healthy lifestyles among elderly people living in Southern Europe: recent evidence from the SHARE. Atl Econ J 2020 Mar;48(1):53-66. [doi: 10.1007/s11293-020-09657-3]
3. Aging and chronic diseases: a profile of Canadian seniors. : Public Health Agency of Canada; 2020 URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/aging-chronic-diseases-profile-canadian-seniors-report.html [accessed 2026-04-07]
4. Terraneo M. Inequities in health care utilization by people aged 50+: evidence from 12 European countries. Soc Sci Med 2015 Feb;126:154-163. [doi: 10.1016/j.socscimed.2014.12.028] [Medline: 25562311]
5. Zayas CE, He Z, Yuan J, et al. Examining healthcare utilization patterns of elderly middle-aged adults in the United States. Proc Int Fla AI Res Soc Conf 2016 May;2016:361-366. [Medline: 27430035]
6. Meeting the care needs of Canada's aging population. : The Conference Board of Canada; 2018 URL: https://www.cma.ca/sites/default/files/2018-11/Conference%20Board%20of%20Canada%20-%20Meeting%20the%20Care%20Needs%20of%20Canada%27s%20Aging%20Population%20%281%29.PDF [accessed 2026-04-07]
7. Primary care networks' planning guide for preventive care. : BC Ministry of Health; 2019.
8. The lifetime prevention schedule: establishing priorities among effective clinical prevention services in British Columbia. : BC Ministry of Health; 2021 URL: https://www2.gov.bc.ca/gov/content/health/about-bc-s-health-care-system/health-priorities/lifetime-prevention [accessed 2026-04-07]
9. Krueger H, Robinson S, Hancock T, et al. Priorities among effective clinical preventive services in British Columbia, Canada. BMC Health Serv Res 2022 Apr 26;22(1):564. [doi: 10.1186/s12913-022-07871-0] [Medline: 35473549]
10. BC Cancer breast screening 2018 program results. : BC Cancer; 2019 URL: https://www.bccancer.bc.ca/screening/Documents/082019_Breast%20Annual%20Report2019_V20191025.pdf [accessed 2026-04-07]
11. Social determinants of health. Office of Disease Prevention and Health Promotion. 2023. URL: https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health [accessed 2026-04-07]
12. Expansion of virtual care in Canada: new data and information. : Canadian Institute for Health Information; 2023 URL: https://www.cihi.ca/sites/default/files/document/expansion-of-virtual-care-in-canada-report-en.pdf [accessed 2026-04-07]
13. Olson KE, O'Brien MA, Rogers WA, Charness N. Diffusion of technology: frequency of use for younger and older adults. Ageing Int 2011 Mar;36(1):123-145. [doi: 10.1007/s12126-010-9077-9] [Medline: 22685360]
14. Graham SA, Stein N, Shemaj F, Branch OH, Paruthi J, Kanick SC. Older adults engage with personalized digital coaching programs at rates that exceed those of younger adults. Front Digit Health 2021;3:642818. [doi: 10.3389/fdgth.2021.642818] [Medline: 34713112]
15. Kaihlanen AM, Virtanen L, Buchert U, et al. Towards digital health equity—a qualitative study of the challenges experienced by vulnerable groups in using digital health services in the COVID-19 era. BMC Health Serv Res 2022 Feb 12;22(1):188. [doi: 10.1186/s12913-022-07584-4] [Medline: 35151302]
16. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Dec;21(1). [doi: 10.1186/s12889-021-11623-w]
17. Vernon SW, McQueen A, Tiro JA, del Junco DJ. Interventions to promote repeat breast cancer screening with mammography: a systematic review and meta-analysis. J Natl Cancer Inst 2010 Jul 21;102(14):1023-1039. [doi: 10.1093/jnci/djq223] [Medline: 20587790]

XSL•FO
RenderX

18. Dougherty MK, Brenner AT, Crockett SD, et al. Evaluation of interventions intended to increase colorectal cancer screening rates in the United States: a systematic review and meta-analysis. JAMA Intern Med 2018 Dec 1;178(12):1645-1658. [doi: 10.1001/jamainternmed.2018.4637] [Medline: 30326005]
19. Aromataris E, Munn Z. JBI Manual for Evidence Synthesis. : JBI; 2020 URL: https://synthesismanual.jbi.global [accessed 2026-04-07]
20. Stern C, Lizarondo L, Carrier J, et al. Methodological guidance for the conduct of mixed methods systematic reviews. JBI Evid Synth 2020 Oct;18(10):2108-2118. [doi: 10.11124/JBISRIR-D-19-00169] [Medline: 32813460]
21. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions: John Wiley & Sons; 2011, Vol. 4.
22. Immunization: Schedules: Recommended vaccines for adults. HealthLink BC. 2023. URL: https://www.healthlinkbc.ca/health-library/immunizations/schedules/recommended-vaccines-adults [accessed 2026-04-07]
23. Garritty C, Gartlehner G, Nussbaumer-Streit B, et al. Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. J Clin Epidemiol 2021 Feb;130:13-22. [doi: 10.1016/j.jclinepi.2020.10.007] [Medline: 33068715]
24. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282. [Medline: 23092060]
25. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019:l4898. [doi: 10.1136/bmj.l4898]
26. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
27. Popay J, Roberts H, Sowden A, et al. Guidance on the conduct of narrative synthesis in systematic reviews. a product from the ESRC methods programme version. : Lancaster University; 2006. [doi: 10.13140/2.1.1018.4643]
28. Liu P, Li Z. Task complexity: a review and conceptualization framework. Int J Ind Ergon 2012 Nov;42(6):553-568. [doi: 10.1016/j.ergon.2012.09.001]
29. Moher D, Shamseer L, Clarke M, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
30. Basch CE, Zybert P, Wolf RL, et al. A randomized trial to compare alternative educational interventions to increase colorectal cancer screening in a hard-to-reach urban minority population with health insurance. J Community Health 2015 Oct;40(5):975-983. [doi: 10.1007/s10900-015-0021-5] [Medline: 25850386]
31. Chan SSC, Leung DYP, Leung AYM, et al. A nurse-delivered brief health education intervention to improve pneumococcal vaccination rate among older patients with chronic diseases: a cluster randomized controlled trial. Int J Nurs Stud 2015 Jan;52(1):317-324. [doi: 10.1016/j.ijnurstu.2014.06.008] [Medline: 25012957]
32. Denis B, Broc G, Sauleau EA, Gendre I, Gana K, Perrin P. Tailored telephone counselling to increase participation of underusers in a population-based colorectal cancer-screening programme with faecal occult blood test: a randomized controlled trial. Rev Epidemiol Sante Publique 2017 Feb;65(1):17-28. [doi: 10.1016/j.respe.2016.06.336] [Medline: 28089385]
33. Fernández ME, Savas LS, Carmack CC, et al. A randomized controlled trial of two interventions to increase colorectal cancer screening among Hispanics on the Texas–Mexico border. Cancer Causes Control 2015 Jan;26(1):1-10. [doi: 10.1007/s10552-014-0472-5]
34. Fleming TJ, Benitez MG, Weintraub MLR. Evaluating the effectiveness of one-on-one conversations to increase colorectal cancer screening in a community-based clinical setting. J Am Osteopath Assoc 2018 Jan 1;118(1):26-33. [doi: 10.7556/jaoa.2018.005] [Medline: 29309089]
35. Ghadieh AS, Hamadeh GN, Mahmassani DM, Lakkis NA. The effect of various types of patients' reminders on the uptake of pneumococcal vaccine in adults: a randomized controlled trial. Vaccine (Auckl) 2015 Oct 26;33(43):5868-5872. [doi: 10.1016/j.vaccine.2015.07.050] [Medline: 26232345]
36. Hagoel L, Neter E, Stein N, Rennert G. Harnessing the question-behavior effect to enhance colorectal cancer screening in an mHealth experiment. Am J Public Health 2016 Nov;106(11):1998-2004. [doi: 10.2105/AJPH.2016.303364] [Medline: 27631750]
37. Hirst Y, Skrobanski H, Kerrison RS, et al. Text-message Reminders in Colorectal Cancer Screening (TRICCS): a randomised controlled trial. Br J Cancer 2017 May 23;116(11):1408-1414. [doi: 10.1038/bjc.2017.117] [Medline: 28441381]
38. Hurley LP, Beaty B, Lockhart S, et al. RCT of centralized vaccine reminder/recall for adults. Am J Prev Med 2018 Aug;55(2):231-239. [doi: 10.1016/j.amepre.2018.04.022] [Medline: 29910118]
39. Jiang M, Yao X, Li P, et al. Impact of video-led educational intervention on uptake of influenza vaccine among the elderly in western China: a community-based randomized controlled trial. BMC Public Health 2022 Dec;22(1):1128. [doi: 10.1186/s12889-022-13536-8]
40. Lieu TA, Elkin EP, Escobar PR, et al. Effect of electronic and mail outreach from primary care physicians for COVID-19 vaccination of Black and Latino older adults: a randomized clinical trial. JAMA Netw Open 2022 Jun 1;5(6):e2217004. [doi: 10.1001/jamanetworkopen.2022.17004] [Medline: 35713906]
41. López-Torres Hidalgo J, Rabanales Sotos J, Simarro Herráez MJ, López-Torres López J, Campos Rosa M, López Verdejo M. Effectiveness of three interventions to improve participation in colorectal cancer screening. Rev Esp Enferm Dig 2016 Jun;108(6):315-322. [doi: 10.17235/reed.2016.4048/2015] [Medline: 27055722]

XSL•FO
RenderX

42. Luckmann R, Costanza ME, White MJ, et al. A 4-year randomized trial comparing three outreach interventions to promote screening mammograms. Transl Behav Med 2019 Mar 1;9(2):328-335. [doi: 10.1093/tbm/iby031] [Medline: 29796649]
43. Luque Mellado FJ, Paino Pardal L, Condomines Feliu I, et al. Impact of a Primary Care intervention on the colorectal cancer early detection programme. Gastroenterol Hepatol 2019;42(6):351-361. [doi: 10.1016/j.gastrohep.2019.01.007] [Medline: 30954319]
44. Modin D, Johansen ND, Vaduganathan M, et al. Effect of electronic nudges on influenza vaccination rate in older adults with cardiovascular disease: prespecified analysis of the NUDGE-FLU Trial. Circulation 2023 May 2;147(18):1345-1354. [doi: 10.1161/CIRCULATIONAHA.123.064270] [Medline: 36871213]
45. Regan AK, Bloomfield L, Peters I, Effler PV. Randomized controlled trial of text message reminders for increasing influenza vaccination. Ann Fam Med 2017 Nov;15(6):507-514. [doi: 10.1370/afm.2120] [Medline: 29133488]
46. Sepucha KR, Valentine KD, Atlas SJ, et al. Getting patients back for routine colorectal cancer screening: randomized controlled trial of a shared decision-making intervention. Cancer Med 2023 Feb;12(3):3555-3566. [doi: 10.1002/cam4.5172] [Medline: 36052811]
47. Szilagyi PG, Albertin C, Casillas A, et al. Effect of patient portal reminders sent by a health care system on influenza vaccination rates: a randomized clinical trial. JAMA Intern Med 2020 Jul 1;180(7):962-970. [doi: 10.1001/jamainternmed.2020.1602] [Medline: 32421168]
48. Szilagyi PG, Albertin CS, Casillas A, et al. Effect of personalized messages sent by a health system's patient portal on influenza vaccination rates: a randomized clinical trial. J Gen Intern Med 2022 Feb;37(3):615-623. [doi: 10.1007/s11606-021-07023-w] [Medline: 34472020]
49. Venishetty N, Calderon-Mora J, Shokar NK, et al. Implementing a mailed stool sample screening program in clinics providing care for an underserved Hispanic population. Cancer Treat Res Commun 2023;37:100756. [doi: 10.1016/j.ctarc.2023.100756] [Medline: 37659188]
50. Vives N, Vidal C, de Guzman EN, et al. The use of text messages as an alternative invitation method for breast cancer screening: a randomized controlled trial (M-TICS study). PLOS ONE 2024;19(8):e0306720. [doi: 10.1371/journal.pone.0306720] [Medline: 39208325]
51. Wang Z, Chan PSF, Fang Y, et al. Chatbot-delivered online intervention to promote seasonal influenza vaccination during the COVID-19 pandemic: a randomized clinical trial. JAMA Netw Open 2023 Sep 5;6(9):e2332568. [doi: 10.1001/jamanetworkopen.2023.32568] [Medline: 37695585]
52. Wu TY, Lin C. Developing and evaluating an individually tailored intervention to increase mammography adherence among Chinese American women. Cancer Nurs 2015;38(1):40-49. [doi: 10.1097/NCC.0000000000000126] [Medline: 24621965]
53. Sediva H, Cartwright T, Robertson C, Deb SK. Behavior change techniques in digital health interventions for midlife women: systematic review. JMIR Mhealth Uhealth 2022 Nov 9;10(11):e37234. [doi: 10.2196/37234] [Medline: 36350694]
54. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023;11:1281923. [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
55. Fridman I, Smalls A, Fleming P, Elston Lafata J. Preferences for electronic modes of communication among older primary care patients: cross-sectional survey. JMIR Form Res 2023 May 24;7:e40709. [doi: 10.2196/40709] [Medline: 37223979]
56. Lenoir KM, Sandberg JC, Miller DP, Wells BJ. Patient perspectives on a targeted text messaging campaign to encourage screening for diabetes: qualitative study. JMIR Form Res 2023 Jan 17;7(1):e41011. [doi: 10.2196/41011] [Medline: 36649056]
57. Weaver KE, Ellis SD, Denizard-Thompson N, Kronner D, Miller DP. Crafting appealing text messages to encourage colorectal cancer screening test completion: a qualitative study. JMIR Mhealth Uhealth 2015 Nov 4;3(4):e100. [doi: 10.2196/mhealth.4651] [Medline: 26537553]
58. Neuhauser L, Kreps GL, Morrison K, Athanasoulis M, Kirienko N, Van Brunt D. Using design science and artificial intelligence to improve health communication: ChronologyMD case example. Patient Educ Couns 2013 Aug;92(2):211-217. [doi: 10.1016/j.pec.2013.04.006]
59. Parsa S, Somani S, Dudum R, Jain SS, Rodriguez F. Artificial intelligence in cardiovascular disease prevention: is it ready for prime time? Curr Atheroscler Rep 2024 Jul;26(7):263-272. [doi: 10.1007/s11883-024-01210-w] [Medline: 38780665]
60. He L, Basar E, Wiers RW, Antheunis ML, Krahmer E. Can chatbots help to motivate smoking cessation? A study on the effectiveness of motivational interviewing on engagement and therapeutic alliance. BMC Public Health 2022 Apr 12;22(1):726. [doi: 10.1186/s12889-022-13115-x] [Medline: 35413887]
61. Li Y, Lee KC, Bressington D, et al. A theory and evidence-based artificial intelligence-driven motivational digital assistant to decrease vaccine hesitancy: intervention development and validation. Vaccines (Basel) 2024;12(7):708. [doi: 10.3390/vaccines12070708]
62. Brown A, Kumar AT, Melamed O, et al. A motivational interviewing chatbot with generative reflections for increasing readiness to quit smoking: iterative development study. JMIR Ment Health 2023 Oct 17;10:e49132. [doi: 10.2196/49132] [Medline: 37847539]
63. Meyer S, Elsweiler D. LLM-based conversational agents for behaviour change support: a randomised controlled trial examining efficacy, safety, and the role of user behaviour. Int J Hum Comput Stud 2025 May;200:103514. [doi: 10.1016/j.ijhcs.2025.103514]

XSL•FO
RenderX

64. Fisher S, Rosella LC. Priorities for successful use of artificial intelligence by public health organizations: a literature review. BMC Public Health 2022 Nov 22;22(1):2146. [doi: 10.1186/s12889-022-14422-z] [Medline: 36419010]
65. Jung SO, Son YH, Choi E. E-health literacy in older adults: an evolutionary concept analysis. BMC Med Inform Decis Mak 2022 Jan 31;22(1):28. [doi: 10.1186/s12911-022-01761-5] [Medline: 35101005]
66. Insights: 2024 Canadian Digital Health Survey. : Canada Health Infoway; 2025 URL: https://insights.infoway-inforoute.ca/2024-digital-health-survey [accessed 2026-04-07]
67. Ricci-Cabello I, Bobrow K, Islam SMS, et al. Examining development processes for text messaging interventions to prevent cardiovascular disease: systematic literature review. JMIR Mhealth Uhealth 2019 Mar 29;7(3):e12191. [doi: 10.2196/12191] [Medline: 30924790]
68. Cambon L, Terral P, Alla F. From intervention to interventional system: towards greater theorization in population health intervention research. BMC Public Health 2019 Dec;19(1):339. [doi: 10.1186/s12889-019-6663-y]
69. Prestwich A, Webb TL, Conner M. Using theory to develop and test interventions to promote changes in health behaviour: evidence, issues, and recommendations. Curr Opin Psychol 2015 Oct;5:1-5. [doi: 10.1016/j.copsyc.2015.02.011]

Abbreviations

AI: artificial intelligence
FIT: Fecal Immunochemical Test
OR: odds ratio
RoB 2: Risk of Bias 2
ROBINS-I: Risk of Bias in Non-randomized Studies—of Interventions

Edited by Y Luo; submitted 19.Sep.2025; peer-reviewed by J Lau, L Martinengo; revised version received 14.Feb.2026; accepted 27.Feb.2026; published 29.Apr.2026.

Please cite as:
Burton L, Rush KL, Smith MA, Janke R
The Effectiveness of Digital Interventions to Increase Preventive Care Uptake in Older Adults: Systematic Review
JMIR Aging 2026;9:e83446
URL: https://aging.jmir.org/2026/1/e83446
doi:10.2196/83446

© Lindsay Burton, Kathy L Rush, Mindy A Smith, Robert Janke. Originally published in JMIR Aging (https://aging.jmir.org), 29.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Digital-Based Exercise Interventions on Concerns About Falling, Falls Efficacy, and Physical Performance Among Older Adults: Systematic Review and Meta-Analysis

Zhaojun Wang[1*], MSc; Bochi Zhu[2*], MSc; Meng Zhou[1], MSc; Xiaojie Xie[1], MSc; Xueyan Zhang[1], PhD

[1]School of Nursing, Jilin University, NO. 965 Xinjiang Street, Changchun, Jilin, China

[2]Department of Neurology, The Second Hospital of Jilin University, Changchun, China

[*]these authors contributed equally

Corresponding Author:
Xueyan Zhang, PhD
School of Nursing, Jilin University, NO. 965 Xinjiang Street, Changchun, Jilin, China

Abstract

Background: Falls are prevalent and serious health problems among older adults. Concerns about falling and reduced falls efficacy are common fall-related psychological impairments, representing 2 distinct emotional and cognitive constructs, respectively. The impact of digital-based exercise interventions on these specific constructs remains unclear.

Objective: This systematic review and meta-analysis aimed to synthesize current evidence on digital-based exercise interventions for concerns about falling and falls efficacy among older adults, with a specific focus on determining their differential effects on emotional and cognitive constructs and evaluating their impact on physical performance.

Methods: The PubMed, Web of Science, Cochrane Library, Embase, PsycINFO, CINAHL, CNKI, SinoMed, VIP, and Wanfang databases were systematically searched from their inception dates to May 2025. We searched for published randomized controlled trials on the effects of digital-based interventions on the fear of falling, concerns about falling, and falls efficacy among older adults. The study followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines and was performed using Stata 17.0 software (StataCorp LLC).

Results: Eighteen studies involving 2435 participants were included. Meta-analyses revealed significant effects of digital-based exercise interventions on falls efficacy (standardized mean difference 0.70, 95% CI 0.51-0.90; *P*<.001), balance function (mean difference [MD] 4.03, 95% CI 2.57-5.49; *P*<.001), functional mobility (MD −1.65, 95% CI −2.52 to −0.77; *P*<.001), and physical function (MD 0.57, 95% CI 0.12-1.02; *P*=.006) among older adults. However, digital-based exercise interventions had no significant effect on concerns about falling, which is the emotional construct related to falls (standardized MD −0.12, 95% CI −0.28 to 0.05; *P*>.05).

Conclusions: The meta-analysis assessed the efficacy of digital-based exercise interventions on fall-related psychological impairments among older adults and revealed that the effects differed across the constructs. These findings suggest that digital-based exercise interventions have potential benefits for improving falls efficacy and physical performance among older adults compared with controls. However, the effect of digital-based exercise interventions on concerns about falling, which is the emotional construct related to falls, remains uncertain among older adults.

(JMIR Aging 2026;9:e87070) doi:10.2196/87070

KEYWORDS

digital-based; meta-analysis; older adults; concerns about falling; falls efficacy

Introduction

Falling is a prevalent and serious geriatric problem that affects approximately 52% of older adults worldwide [1,2]. The 2022 World Falls Guidelines recommend incorporating the assessment of concerns about falling into multifactorial risk assessments for fall prevention and management [3]. Such fall-related psychological impairments cannot simply be classified as post-fall syndrome, as they can occur in older adults who have never fallen [4]. Instead, they constitute the central mechanism driving the vicious cycle of falls [5,6]. These psychological impairments not only heighten the risk of physical injuries but also impose a significant psychological burden, manifesting as anxiety, depression, and frailty [7-10], and substantially compromise the overall quality of life of older adults. Moreover, the increased risk of falls and related complications can lead to higher mortality rates in older adults [11]. Effective intervention strategies are urgently needed given these profound implications.

XSL•FO
RenderX

To develop effective interventions, a precise understanding of fall-related psychological impairments is crucial. Common fall-related psychological factors include fear of falling, falls efficacy, concerns about falling, and balance confidence, among others [12-14]. Despite the frequent interchangeable use of these terms, they represent different psychological processes with different predictive values for falls [15]. Fear of falling and concerns about falling are both emotional constructs [16]. The former refers to an individual's emotional response to a real or perceived threat to balance, whereas the latter manifests as a persistent state of worry [17]. On the other hand, falls efficacy and balance confidence are classified as cognitive constructs, primarily referring to an individual's belief in their own ability or confidence in performing activities without falling. Furthermore, balance confidence is closely related to falls efficacy and is regarded as an important dimension of it [18]. Therefore, we use falls efficacy as an umbrella term to encompass the specific construct of balance confidence. This distinction is methodologically and clinically critical because concerns about falling predict future falls, whereas falls efficacy does not [15]. Combining them in meta-analyses dilutes their unique effects and may be a key source of heterogeneity, potentially obscuring the true efficacy of interventions. Distinguishing between them in research is crucial to avoid this bias and to reveal the specific effects of interventions on different constructs.

Owing to rapid advancements, digital technologies have been increasingly applied to exercise and health management for older individuals. Digital technologies are defined as the use of a series of computing platforms, software, internet-connected devices, and sensors for health care and related purposes, including wearable devices, virtual reality (VR), mobile health apps, and web-based communication platforms [19,20]. Digital-based exercise intervention refers to exercise interventions that use digital technologies to prevent, manage, and treat diseases or promote health. While traditional face-to-face interventions, such as group exercise and physical therapy, have demonstrated efficacy in improving physical function [21-23], they present challenges, including temporal and spatial constraints, health care resource disparities, and high costs [19]. Digital-based exercise interventions may offer novel solutions to these limitations through remote accessibility and automated personalization [24].

While several systematic reviews have investigated interventions for concerns about falling or falls efficacy [12,14], important distinctions set our work apart. First, previous reviews, while covering a broad range of interventions, have not provided a dedicated synthesis of digital-based exercise interventions [12-14]. Our review is distinguished by its exclusive focus on digital interventions and the inclusion of trials up to 2025. This is particularly necessary in the rapidly advancing field of digital interventions, such as those based on VR and smartphone apps, ensuring that our synthesis is current and comprehensive. Second, many syntheses of evidence share a methodological limitation in treating the fear of falling, concerns about falling, or falls efficacy as a single, homogeneous construct [25-27]. This approach contributes to significant heterogeneity and inconsistent findings across systematic reviews, as it fails to capture the differential effects of interventions on cognitive and emotional constructs. To address this, our study explicitly distinguishes cognitive constructs, such as falls efficacy, from emotional constructs, such as concerns about falling. This strategy allows us to reveal the specific intervention effects on each construct, thereby advancing the understanding of their distinct mechanisms in digital-based exercise interventions. Finally, while some reviews have focused on adults of broader ages [28-31], our precise targeting of older adults is designed to yield conclusions with greater specificity and clinical applicability for geriatric care.

This systematic review and meta-analysis aimed to (1) provide an updated synthesis of digital-based exercise interventions for concerns about falling and falls efficacy, (2) determine their specific effects on each distinct construct (cognitive and emotional), and (3) evaluate their impact on physical performance.

Methods

The study protocol was registered in PROSPERO (CRD42024567108), and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Checklist 1) were followed [32].

Search Strategy and Study Selection

The PubMed, Web of Science, Cochrane Library, Embase, PsycINFO, CINAHL, CNKI, SinoMed, VIP, and Wanfang databases were systematically searched for relevant publications. The search terms included Medical Subject Headings (MeSH) and free words related to digital-based interventions, older adults, fear of falling, concerns about falling, falls efficacy, and randomized controlled trials (RCTs). The search terms were combined using Boolean operations. There were no restrictions on the language of the publications. The search period ranged from database inception to May 2025 (the specific search strategies are presented in Multimedia Appendix 1).

Study selection and data extraction were performed independently by 2 researchers. All the retrieved studies were imported into EndNote X20 software (Clarivate) for screening, and duplicate studies were automatically removed by the software. The researchers read the titles, abstracts, and full texts of the studies and subsequently removed studies that did not meet the inclusion criteria. After the study selection was completed, two researchers cross-checked the screening results. Any disagreements were resolved through discussion with a third researcher.

Eligibility Criteria

Inclusion Criteria

The inclusion criteria are summarized in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion criteria

- Population: older adults (aged ≥60 y)
- Interventions: participants in the intervention group underwent an exercise intervention that was delivered via digital technologies (wearable devices, exergaming, virtual reality, mobile health apps, and web-based communication platforms).
- Comparison: no digital-based exercise intervention was used in the control group, which included active controls (motor exercise, balance exercise, rehabilitation exercise, and health education) and passive controls (no intervention or maintenance of normal daily living).
- Outcomes**:** studies reporting at least one of the following outcomes.
  - Primary outcome: concerns about falling or falls efficacy included as reported in the original studies, whether as a primary or secondary outcome. Concerns about falling was measured by validated scales (Falls Efficacy Scale–International, Icon-Falls Efficacy Scale, and Short Falls Efficacy Scale). Falls efficacy was measured by validated scales (Falls Efficacy Scale, Activities-Specific Balance Confidence Scale, and Modified Falls Efficacy Scale).
  - Secondary outcome: physical performance included measures of balance (Berg Balance Scale), functional mobility (Timed Up and Go Test), and physical function (Short Physical Performance Battery).
- Study design: randomized controlled trial

Exclusion Criteria

We excluded studies that focused primarily on individuals diagnosed with dementia, duplicate studies, studies with unavailable full texts or important data, meeting abstracts, study protocols, studies without treatment details or primary outcomes, pilot studies, and feasibility trials.

Data Extraction

Two researchers independently extracted data from the complete original studies using a uniform data extraction template. The data extracted included the author's name, country, year of publication, study population, age, sex, sample size, intervention, control, intervention period, duration of intervention, frequency of intervention, outcome indicators, end values in the control and intervention groups, and evaluation tools for the outcome indicators. Any discrepancies between the initial disagreements were first discussed between the 2 researchers until a consensus was reached. If an agreement could not be reached, the disagreement was referred to a third, senior researcher for adjudication and a final decision.

Assessment of the Risk of Bias and Certainty of the Evidence

An independent appraisal of the risk of bias for each included study was performed by 2 investigators who used the Cochrane Risk of Bias 2 (RoB2) instrument. The appraisal covered 5 predefined domains: randomization procedure, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result. On the basis of the assessments across these domains, the researchers subsequently assigned an overall risk judgment of "low risk," "some concerns," or "high risk" to each included study.

We evaluated the certainty of the evidence using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Profiler Guideline Development Tool. In accordance with the GRADE system, the certainty of evidence for each outcome was classified as high, moderate, low, or very low. Any disagreements were referred to a third researcher for adjudication.

Data Analysis

Statistical analysis was conducted using Stata 17.0 (StataCorp LLC). The outcome indicators were continuous variables, and effect sizes were expressed as the means and SDs combined with 95% CIs to report the results of the continuous variable data analysis. For continuous variables, the mean difference (MD) was selected when the same outcome indicator was measured on the same scale; otherwise, the standardized mean difference (SMD) was adopted. Between-group heterogeneity of synthetic effect sizes was tested using the I^2 statistic and P value. Random-effects models were used for the meta-analysis because of anticipated heterogeneity. To further explore the factors influencing the intervention, additional subgroup analyses were performed on the basis of the setting, intervention qualifications, intervention duration, exercise mode, digital platform, duration of the intervention, exposure dosage, and scale-scoring method. Reverse scoring scale–converted data (mean x−1, with the SD unchanged) were included in the analysis alongside positive scoring scale data to improve the accuracy and reliability of the results. For multiarm RCTs, this review included only comparisons that involved an eligible digital-based exercise intervention group and a control group. In the meta-analysis, each trial contributed only one comparison pair to prevent double-counting of control group data. To assess the robustness of the combined effect sizes, sensitivity analyses were performed using a remove-one-article method. For outcomes with data from more than 10 studies, publication bias was assessed via funnel plots [33]. The Egger test [34] and the Begg test [35] were used to further assess publication bias.

Results

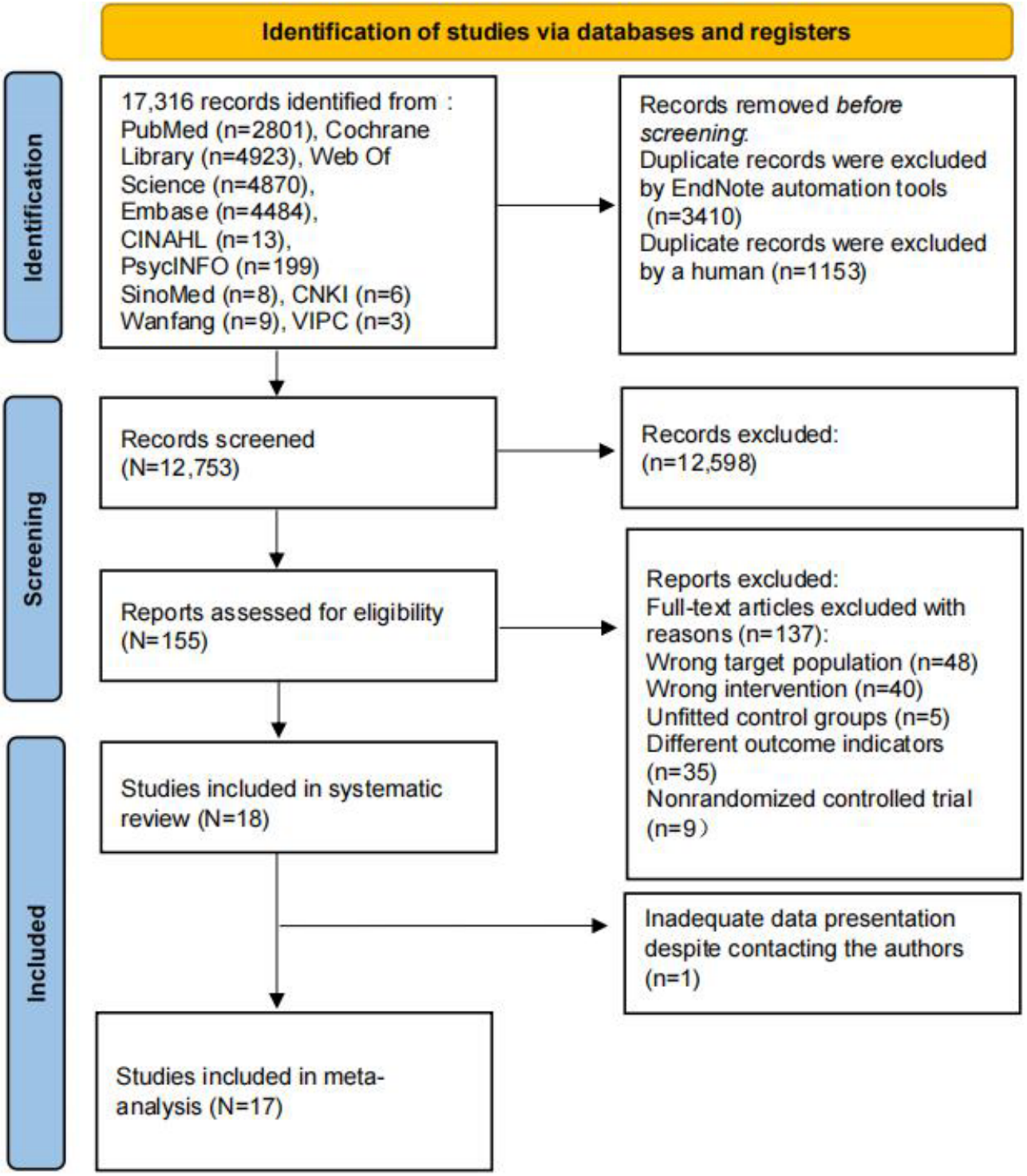
Search Results

In total, 17,316 articles were retrieved from 10 databases; 3410 duplicates were removed by EndNote X20 software, and 1153 duplicates were removed manually. A total of 12,598 articles were excluded because their titles and abstracts were not relevant to the aim of the study. After the full texts were reviewed, 137 studies were excluded. Ultimately, 18 studies [36-53] were included in this study. One study [44] was excluded from the quantitative synthesis because of inadequate data presentation

XSL•FO
RenderX

despite contacting the authors. Therefore, 17 studies were included in the meta-analysis (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the search process.



Characteristics of the Studies

The characteristics of the included studies are presented in Table 1. A total of 2435 older adults were included, with average ages ranging from 67.8 (SD 2.98) to 84.1 (SD 5.5) years. The proportion of women varied between 26.7% and 100%. The publication years ranged from 2013 to 2025. These studies were conducted in the following countries: Australia (n=2) [39,51], China (n=3) [40,42,47], Korea (n=2) [36,53], Switzerland (n=1) [49], Spain (n=2) [41,50], Iran (n=1) [45], Singapore (n=1) [37], Italy (n=1) [38], Türkiye (n=1) [46], the Netherlands (n=1) [48], Pakistan (n=1) [43], Chile (n=1) [52], and Brazil (n=1) [44]. The study settings included a range of environments, such as communities [37,41,46,48,51-53], hospitals [38,39,43,44,47,49,50], and nursing homes or welfare centers [36,40,42,45].

The intervention details of the included studies are presented in Table 2. The interventions were primarily delivered by qualified health care professionals [37-40,42-44,48-50], with

XSL•FO
RenderX

physiotherapists being the most common. In other cases, the delivery was managed by trained nonspecialists [36,41,51,52], including research assistants or physical education instructors. However, the specific qualifications of the interveners were not reported in several trials [45-47,53]. The digital-based exercise modes included balance-dominant balance training [38,41,44,45,48,51]; balance-inclusive multimodal training, which combines balance with one or more forms of resistance or aerobic exercise; and nonbalance training dominated by aerobic exercise [52]. Among these, balance-inclusive multimodal training was the most common. Furthermore, predesigned programs such as Baduanjin [42] and the Otago Exercise Program [47] incorporate various elements, including balance, strength, and flexibility, and are classified as balance-inclusive multimodal training. Exergaming [36-41,43-45,51-53] was the most frequently used digital delivery platform for exercise interventions, followed by computer-based applications [46,47,49,50], while only 2 studies [42,48] used VR to provide a virtual environment for exercise. The duration of individual interventions ranged from 30 to 120 minutes per session. The most frequently implemented duration for a single session was 60 minutes. The frequency of interventions ranged from 1 to 5 times per week, and the most commonly prescribed intervention frequency ranged from 2 to 3 times per week. The duration ranged from 2 weeks to 12 months, and the most common duration was 8 weeks.

The outcomes included concerns about falling, falls efficacy, balance, functional mobility, and physical function. A wide variety of measurement scales have been applied among studies. Nine studies [38,40,41,43,48-52] used the Falls Efficacy Scale–International (FES-I) scale, Short Falls Efficacy Scale, or Iconographical Falls Efficacy Scale, which reflect concerns about falling, and 9 studies [36,37,39,42,44-47,53] used the Activities-Specific Balance Confidence Scale, Falls Efficacy Scale, or Modified Falls Efficacy Scale (MFES), which reflect balance confidence or falls efficacy. Among the included studies, the MFES and Activities-Specific Balance Confidence Scale are positive scoring scales, where higher scores indicate greater falls efficacy. In contrast, the Falls Efficacy Scale is a reverse scoring scale, meaning that higher scores reflect lower falls efficacy. The FES-I, Short Falls Efficacy Scale, and Iconographical Falls Efficacy Scale also follow a positive scoring direction, but in their case, higher scores indicate a greater level of concerns about falling. The Berg Balance Scale [36,38,45,47,52,53], Timed Up and Go Test [36,37,40,43,45,47,49,51-53], and Short Physical Performance Battery [39,46,49-51] were used to evaluate balance, functional mobility, and physical function, respectively. Only 3 out of the 18 (16.7%) included studies [37,43,47] reported concerns about falling or falls efficacy as their primary outcome. Seven (38.9%) studies [39,40,44,48-51] explicitly defined it as a secondary outcome, whereas the remaining studies (44.4%) [36,38,41,42,45,46,52,53] did not clearly distinguish between primary and secondary outcomes in their original texts.

XSL•FO
RenderX

Table . Characteristics of the included studies (N=18).

Study (author, year)	Participant characteristics						Primary outcome (falls efficacy)
	Country	Participants	Setting	Age (y)	Total sample size	Female (%)	
Lee and Shin [36] (2013)	Korea	Older adults with diabetes mellitus	Welfare center	• CG[a]: 74.29 (5.20)[b] • IG[c]: 73.78 (4.77)[b]	55	70.9	MFES[d]
Kwok and Pua [37] (2016)	Singapore	Older adults with mild-to-moderate physical frailty	Community	• CG: 69.8 (7.5)[b] • IG: 70.5 (6.7)[b]	80	85	MFES
Morone et al [38] (2016)	Italy	Older women with bone loss	Hospital	• CG: 70.05 (4.93)[b] • IG: 67.8 (2.98)[b]	36	100	Short-FES-I[e]
van den Berg et al [39] (2016)	Australia	Older adults	Hospital	• CG: 82 (13)[b] • IG: 78 (10)[b]	56	62.1	FES[f]
Liao et al [40] (2019)	China	Frail older adults	Daycare centers	• CG: 84.1 (5.5)[b] • IG: 79.6 (8.5)[b]	52	69.2	FES-I[g]
Montero-Alía et al [41] (2019)	Spain	Older adults	Community	• CG: 75.4 (72.7 - 78.6)[h] • IG: 75.1 (72.6 - 78.7)[h]	608	59.1	Short-FES-I
Sun et al [42] (2020)	China	Older adults with osteoporosis	Nursing institution	• CG: 72.53 (7.04)[b] • IG: 73.96 (6.32)[b]	60	75.1	MFES
Zahedian-Nasab et al [45] (2021)	Iran	Older adults with fall risk	Nursing homes	• CG: 72 (7.81)[b] • IG: 69.67 (7.73)[b]	60	26.7	FES
Khushnood et al [43] (2021)	Pakistan	Older adults	Hospital	• N/A[i]	83	38.5	ABC[j]
Rebêlo et al [44] (2021)[k]	Brazil	Older adults with balance disorders	Hospital	• CG: 71.41 (5.94)[b] • IG: 69.25 (5.67)[b]	37	83.8	FES-I
Yang et al [47] (2022)	China	Older people with fear of falling	Hospital	• CG: 69.60 (3.21)[b] • IG: 69.92 (2.15)[b]	100	55	MFES

XSL•FO
RenderX

Study (author, year)	Participant characteristics						Primary outcome (falls efficacy)
	Country	Participants	Setting	Age (y)	Total sample size	Female (%)	
Tekin and Cetisli-Korkmaz [46] (2022)	Türkiye	Older adults	Community	• C: 70.34 (5.37)[b] • I: 68.34 (4.33)[b]	255	69.4	MFES
Gerards et al [48] (2023)	Netherlands	Older adults	Community	• CG: 73 (8)[b] • IG: 73 (10)[b]	82	79.3	FES-I
Lee [53] (2023)	Korea	Older adults	Community	• CG: 79.10 (3.90)[b] • IG: 80.39 (2.57)[b]	57	45.6	MFES
Sturnieks et al [51] (2024)	Australia	Older adults	Community	• CG: 72.5 (5.5)[b] • IG: 72.6 (5.7)[b]	507	71	Icon-FES[l]
Hager et al [49] (2024)	Switzerland	Older people at risk of falling	Community	• CG: 79 (6.6)[b] • IG: 79 (7.0)[b]	172	74	FES-I
Prieto-Moreno et al [50] (2024)	Spain	Older adults with hip fracture	Hospital	• CG: 80.07 (7.74)[b] • IG: 79.55 (7.11)[b]	105	71.4	Short-FES-I
Vásquez-Carrasco [52] (2025)	Chile	Older females	Community	• CG: 76.2 (2.03)[b] • IG: 72.2 (6.76)[b]	30	100	FES-I

[a]CG: control group.

[b]Mean (SD).

[c]IG: intervention group.

[d]MFES: Modified Fall Efficacy Scale.

[e]Short-FES-I: Short Falls Efficacy Scale International.

[f]FES: Falls Efficacy Scale.

[g]FES-I: Falls Efficacy Scale International.

[h]Medium (IQR).

[i]Not applicable.

[j]ABC: Activities-Specific Balance Confidence scale.

[k]Included in the qualitative synthesis but not in the quantitative synthesis.

[l]Icon-FES: Iconographical Falls Efficacy Scale.

XSL•FO
RenderX

Table . Intervention details of the included studies (N=18).

Study	Qualifications of interveners	Intervention	Control	Exercise mode	Digital platform	Duration for a single session	Exposure
Lee and Shin [36] (2013)	Research assistant	VR[a] exercise + health education on diabetes management	Health education on diabetes management	Balance + resistance + aerobic training	Exergaming	50 minutes	Twice a week, for 10 weeks
Kwok and Pua [37] (2016)	Physiotherapist and a therapist assistant	Nintendo Wii exercise program	Gym-based exercise intervention	Balance + resistance training	Exergaming	60 minutes	Once a week, for 12 weeks
Morone et al [38] (2016)	Physiotherapist	Balance training based on Nintendo WiiFit	Routine balance training	Balance training	Exergaming	60 minutes	Twice a week, for 8 weeks
van den Berg et al [39] (2016)	A physiotherapist and a therapist assistant	Computer-based interactive exercises + routine rehabilitation nursing	Routine rehabilitation nursing	Balance + resistance training	Exergaming	60 minutes	Five times a week, for 2 weeks
Liao et al [40] (2019)	Experienced physical therapist	Kinect-based exergaming exercise	Combined exercise training	Tai chi + balance + resistance + aerobic training	Exergaming	60 minutes	Three times a week, for 12 weeks
Montero-Alía et al [41] (2019)	Monitors who had received standardized training	Balance training based on Nintendo Wii game	Routine care	Balance training	Exergaming	30 minutes	Twice a week, for 3 months
Sun et al [42] (2020)	Therapist	VR-based Baduanjin training + conventional osteoporosis treatment	Conventional osteoporosis treatment	Baduanjin training	VR	50 minutes	Three times a week, for 12 months
Zahedian-Nasab et al [45] (2021)	N/A[b]	VR balancing exercise based on Xbox Kinect	Routine activities	Balance training	Exergaming	30 - 45 minutes	Twice a week, for 6 weeks
Khushnood et al [43] (2021)	Therapist	VR-based WiiFit exercise	Conventional balance training exercises	Balance + aerobic training	Exergaming	30 minutes	Twice a week, for 8 weeks
Rebêlo et al [44] (2021)[c]	Therapist	Balance training based on immersive VR	Conventional body balance training	Balance training	Exergaming	50 minutes	Twice a week, for 8 weeks
Yang et al [47] (2022)	N/A	Internet-based Otago Exercise Program workouts	Health education	Balance + resistance training	Application	30 minutes	Three times a week, for 6 months
Tekin and Cetisli-Korkmaz [46] (2022)	N/A	Internet-based calisthenic exercise	No intervention	Resistance + aerobic training	Application	Six different types of calisthenic exercises (repeated 10 times)	Five times a week, for 4 weeks
Gerards et al [48] (2023)	Physiotherapist	Perturbation-based balance training + usual care	Usual care	Balance training	VR	30 minutes	Once a week, for 3 weeks
Lee [53] (2023)	N/A	The home-based exergame program + health education	Health education	Balance + resistance training	Exergaming	50 minutes	Three times a week, for 8 weeks

XSL•FO
RenderX

Study	Qualifications of interveners	Intervention	Control	Exercise mode	Digital platform	Duration for a single session	Exposure
Sturnieks et al [51] (2024)	Exercise science graduates	Computerized game-based sports training	Health education	Balance training	Exergaming	120 minutes	Once a week, for 12 months
Hager et al [49] (2024)	Physiotherapist	Partially supervised exercise	Otago exercise program	Balance + resistance training	Application	30 - 45 minutes	Three times a week, for 12 months
Prieto-Moreno et al [50] (2024)	Physiotherapist or occupational therapist	ActiveHip intervention based on mHealth[d] + usual rehabilitation	Usual rehabilitation	Balance + resistance training	Application	One video-recorded session	Three times a week, for 3 months
Vásquez-Carrasco et al [52] (2025)	Trained physical education instructor	Xbox Kinect Sports	Usual daily routines	Aerobic training	Exergaming	60 minutes	Three times a week, for 24 weeks

[a]VR: virtual reality

[b]Not applicable.

[c]Included in the qualitative synthesis but not in the quantitative synthesis.

[d]mHealth: mobile health.

Risk of Bias in RCTs

The risk of bias was assessed for the 18 RCTs included. The majority of studies were judged to have a high risk of bias or to raise some concerns overall. For the domain of the randomization process, the majority of studies were judged to be at low risk of bias. Two studies were rated as having some concerns due to a lack of information on allocation concealment. One study was assessed as being at high risk because significant baseline differences suggested potential problems with the randomization procedure. For deviations from intended interventions, most studies were assessed as high risk or with some concerns, primarily attributed to a lack of blinding for participants and intervention providers. For the domain of outcome measurement, 3 studies were assessed as high risk because blinding of the outcome assessors was not implemented. The risk of bias was assessed as low across all remaining domains (Figure 2).

XSL•FO
RenderX

Figure 2. Risk of bias summary [36-53].



Effectiveness of Interventions

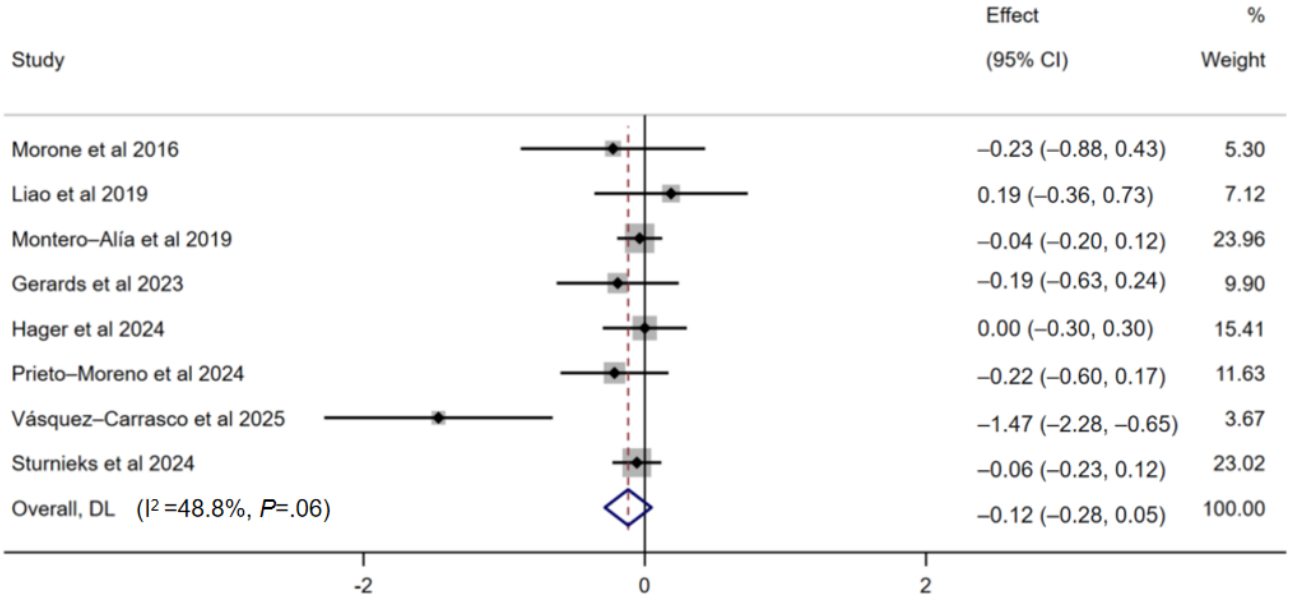
Concerns About Falling

A total of 8 studies evaluated the impact of digital-based exercise interventions on concerns about falling in older adults, involving 758 participants in the intervention groups and 834 in the control groups. The pooled results from 8 studies revealed that digital-based exercise interventions had a significant effect on concerns about falling (SMD −0.12, 95% CI −0.28 to 0.05; P>.05; I^2=48.8%; low certainty of the evidence; Figure 3). None of the subgroup analyses yielded statistically significant results. Although the point estimates varied, none of the tests for subgroup differences were significant (P>.05), with 1 exception: the exercise mode subgroup showed a nominally significant difference (P=.003). However, this finding was driven entirely by a single study in the nonbalance training subgroup, which reported a statistically significant reduction in concerns about falling (SMD −1.47, 95% CI −2.28 to −0.65). When this single

XSL•FO
RenderX

study was excluded, the difference between the remaining exercise mode subgroups was no longer significant (*P*=.86), indicating that the initial subgroup difference was not robust. Therefore, there is no clear evidence to identify which intervention characteristics are most effective for reducing concerns about falling (Table 3).

Figure 3. Forest plot showing the effects of digital-based exercise interventions on concerns about falling (n=8; lower scores indicate greater improvement) [38,40,41,48-52]. Weights are from a random-effects model.



XSL•FO
RenderX

Table . Subgroup analysis of digital-based exercise interventions on concerns about falling and falls efficacy.

Outcome and variable	Numbers of comparisons	Meta-analysis results, SMD[a] (95% CI)	Heterogeneity	Heterogeneity between groups (*P* value)
Concerns about falling				
Setting				.45
Community	5	−0.14 (−0.36 to 0.08)	I^2=66.9%, *P*=.02	
Senior care facilities	1	0.19 (−0.36 to 0.73)	—[b]	
Hospital	2	−0.22 (−0.55 to 0.11)	I^2=0%, *P*=.98	
Qualifications of interveners				.29
Nonprofessional physiotherapist	2	−0.05 (−0.16 to 0.07)	I^2=0%, *P*=.87	
Physiotherapist	6	−0.22 (−0.53 to 0.09)	I^2=60.7%, *P*=.03	
Exercise mode				*.003*
Balance-inclusive multimodal training	3	−0.04 (−0.26 to 0.18)	I^2=0%, *P*=.46	
Balance training	4	−0.06 (−0.17 to 0.05)	I^2=0%, *P*=.88	
Nonbalance training	1	−1.47 (−2.28 to −0.65)	—	
Digital platform				.88
Exergaming	5	−0.15 (−0.42 to 0.11)	I^2=68.3%, *P*=.01	
VR[c]	1	−0.19 (−0.63 to 0.24)	—	
Application	2	−0.08 (−0.32 to 0.15)	I^2=0%, *P*=.39	
Duration of the intervention (wk)				.67
<12	2	−0.20 (−0.57 to 0.16)	I^2=0%, *P*=.93	
≥12	6	−0.11 (−0.31 to 0.09)	I^2=62%, *P*=.02	
Exposure dosage				.46
<20	2	−0.20 (−0.57 to 0.16)	I^2=0%, *P*=.93	
20 - 50	3	−0.05 (−0.19 to 0.10)	I^2=0%, *P*=.48	
>50	3	−0.31 (−0.78 to 0.17)	I^2=82.7%, *P*=.003	
Falls efficacy				
Setting				.34
Community	3	0.55 (0.30 to 0.80)	I^2=24.1%, *P*=.27	
Senior care facilities	3	0.84 (0.53 to 1.15)	I^2=0%, *P*=.87	
Hospital	3	0.77 (0.27 to 1.27)	I^2=71.1%, *P*=.03	
Qualifications of interveners				.51
Nonprofessional physiotherapist	5	0.75 (0.57 to 0.93)	I^2=0%, *P*=.76	
Physiotherapist	4	0.59 (0.15 to 1.03)	I^2=69.8%, *P*=.02	
Exercise mode				.46
Balance-inclusive multimodal training	8	0.69 (0.48 to 0.89)	I^2=42.8%, *P*=.09	
Balance training	1	0.90 (0.37 to 1.43)	—	
Digital platform				.99
Exergaming	6	0.69 (0.37 to 1.00)	I^2=57.5%, *P*=.04	
VR	1	0.73 (0.21 to 1.25)	—	
Application	2	0.71 (0.48 to 0.93)	I^2=26.5%, *P*=.24	

XSL•FO
RenderX

Outcome and variable	Numbers of comparisons	Meta-analysis results, SMD[a] (95% CI)	Heterogeneity	Heterogeneity between groups (*P* value)
Duration of the intervention (wk)				.60
<12	6	0.75 (0.51 to 0.98)	I^2=37.2%, *P*=.16	
≥12	3	0.62 (0.23 to1.02)	I^2=55.9%, *P*=.10	
Exposure dosage				.72
<20	4	0.63 (0.16 to 1.10)	I^2=72.5%, *P*=.01	
20 - 50	3	0.69 (0.48 to 0.90)	I^2=0%, *P*=.68	
>50	2	0.83 (0.50 to 1.15)	I^2=0%, *P*=.64	
Scale-scoring method				.62
Positive scoring	7	0.73 (0.53 to 0.94)	I^2=34.3%, *P*=.17	
Reverse scoring	2	0.56 (−0.10 to 1.23)	I^2=68.3%, *P*=.08	

[a]SMD: standardized mean difference.

[b]Not applicable.

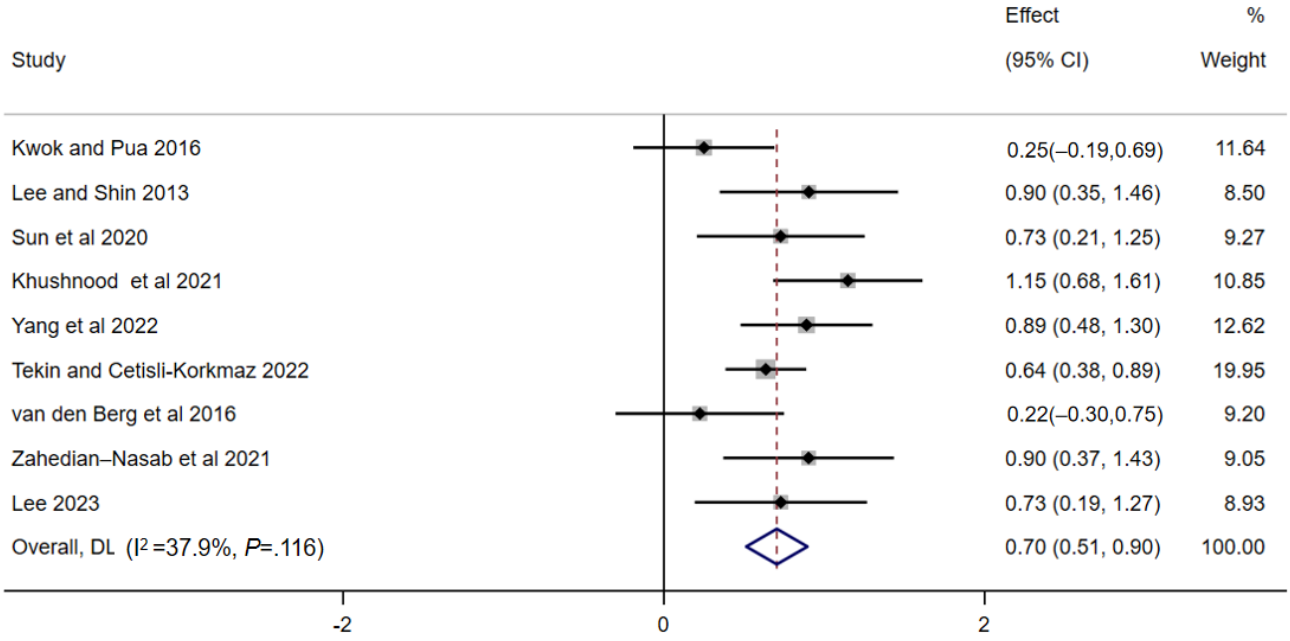
[c]VR: virtual reality.

The sensitivity analysis for the cognitive constructs demonstrated that the overall results were robust, as no single study significantly altered the pooled effect estimate. The sensitivity analysis of affect-based constructs revealed that a study by Vásquez-Carrasco et al [52] (SMD −1.47, 95% CI −2.28 to −0.65) was an outlier. After this study was excluded, the intervention effects remained nonsignificant (SMD −0.06, 95% CI −0.16 to 0.04; *P*>.05; I^2=0.0%).

Falls Efficacy

The meta-analysis of 9 studies revealed that digital-based exercise interventions significantly improved falls efficacy among older adults (SMD 0.70, 95% CI 0.51-0.90; *P*<.001; I^2=37.9%; moderate certainty of the evidence; Figure 4). With respect to intervention setting, the effect size was greatest for senior care facilities (SMD 0.84, 95% CI 0.53-1.15), followed by hospitals (SMD 0.77, 95% CI 0.27-1.27), and then community settings (SMD 0.55, 95% CI 0.30-0.80), although the difference between groups was not significant (*P*=.34). With respect to exercise mode, balance training had a greater effect (SMD 0.90, 95% CI 0.37-1.43) than multimodal training did (SMD 0.69, 95% CI 0.48-0.89), but the difference was not statistically significant (*P*=.46). Similarly, a higher exposure dosage was associated with a greater effect, with the >50 times subgroup (SMD 0.83, 95% CI 0.50-1.15) having a greater effect than the <20 times (SMD 0.63, 95% CI 0.16-1.10) and 20 to 50 times (SMD 0.69, 95% CI 0.48-0.90) subgroups did, and there was no significant between-group difference (*P*=.72). With respect to other subgroup variables, including the intervention qualifications, digital platform type, intervention duration, and scale-scoring method, the effect sizes were similar across categories (SMD range 0.56 - 0.90), with no significant differences between the subgroups (all *P*>.05; Table 3).

Figure 4. Forest plot showing the effects of digital-based exercise interventions on falls efficacy (n=9; higher scores indicate greater improvement) [36,37,39,42,43,45-47,53].



XSL•FO
RenderX

The sensitivity analysis for the falls efficacy demonstrated that the overall results were robust, as no single study significantly altered the pooled effect estimate.

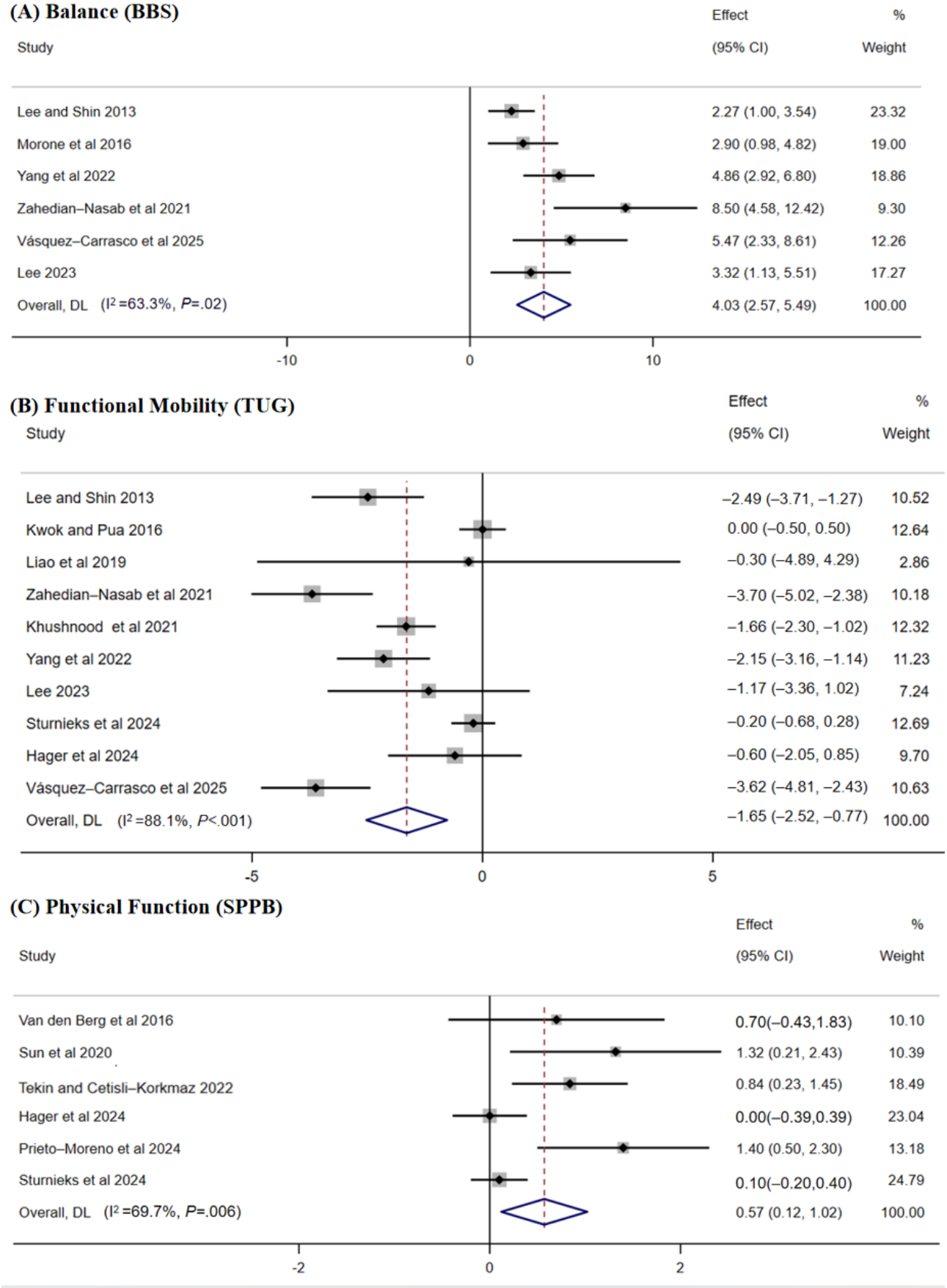
Balance (Berg Balance Scale)

The pooled results from 6 studies [36,38,45,47,52,53] revealed a significant effect of digital-based exercise interventions on balance (MD 4.03, 95% CI 2.57-5.49; $P<.001$; $I^2=63.3\%$; moderate certainty of the evidence; Figure 5A). With respect to balance, most subgroups, including the intervention setting, intervention qualifications, exercise mode, and intervention duration showed no significant differences between subgroups ($P>.05$). The only statistically significant subgroup differences were found between the exposure dosage subgroups ($P=.04$). An exposure dosage of less than 20 times (MD 5.43, 95% CI −0.03 to 10.89) had a greater effect than an exposure dosage of more than 50 times (MD 5.03, 95% CI 3.38-6.68) and 20 to 50 times (MD 2.53, 95% CI 1.43-3.63; Table S1 in Multimedia Appendix 1).

Sensitivity analysis revealed that no individual study was able to change the results, but the heterogeneity decreased from 63.3% to 42.6% when Zahedian-Nasab's [45] study was excluded (MD 3.46, 95% CI 2.30-4.61; $P<.001$; $I^2=42.6\%$).

XSL•FO
RenderX

Figure 5. Forest plot showing the effects of digital-based exercise interventions on (A) balance (Berg Balance Scale [BBS]), (B) functional mobility (Timed Up and Go Test [TUG]), and (C) physical function (Short Physical Performance Battery [SPPB]). A lower TUG score indicates a more significant improvement in functional mobility, whereas higher BBS and SPPB scores indicate more significant improvements in balance and physical function, respectively [36-40,42,43,45-47,49-53].



XSL•FO
RenderX

Functional Mobility (Timed Up and Go Test)

In ten studies [36,37,40,43,45,47,49,51-53], researchers reported functional mobility. The meta-analysis indicated a significant effect of digital-based exercise interventions on functional mobility compared with the control group; however, considerable heterogeneity was noted among the included studies (MD −1.65, 95% CI −2.52 to −0.77; *P*<.001; I^2=88.1%; moderate certainty of the evidence; Figure 5B). Subgroup analyses for functional mobility were performed across the intervention setting, exercise mode, the duration of intervention, digital platform, the qualifications of interveners, and exposure dosage. With respect to exercise mode, subgroup differences were initially observed (*P*=.01), with nonbalance training yielding the greatest effect size (MD −3.62, 95% CI −4.81 to −2.43) compared with balance-inclusive multimodal training (MD −1.29, 95% CI −2.22 to −0.37) and balance training (MD −1.89, 95% CI −5.32 to 1.53). However, the significant subgroup difference disappeared (*P*=.74) after the single study of nonbalance training was excluded. With respect to intervention setting, the digital-based exercise intervention delivered in senior care facilities (MD −2.89, 95% CI −4.11 to −1.68) showed a greater improvement effect than that in community settings (MD −1.03, 95% CI −1.20 to 0.07) and hospitals (MD −1.80, 95% CI −2.34 to −1.26), although the difference between groups was not statistically significant (*P*=.08). Interventions lasting less than 12 weeks (MD −2.30, 95% CI −3.33 to −1.28) improved functional mobility more strongly than interventions lasting 12 weeks or more did (MD −1.20, 95% CI −2.28 to −0.11; *P*=.15). For other subgroup variables, including the qualifications of the intervener, digital platform type, and exposure dosage, the effect sizes were similar across categories (MD range: −2.08 to −1.29), with no significant differences between the subgroups (all *P*>.05; Table S1 in Multimedia Appendix 1). Sensitivity analysis revealed that no individual study was able to change the results.

Physical Function (Short Physical Performance Battery)

In 6 studies [39,42,46,49-51], researchers reported the effects of digital-based exercise interventions on physical function. The meta-analysis revealed that digital-based exercise interventions significantly improved physical function among older adults (MD 0.57, 95% CI 0.12-1.02; *P*=.006; I^2=69.7%; moderate certainty of the evidence; Figure 5C). The hospital subgroup had the greatest effect size (MD 1.13, 95% CI 0.42-1.83), which was greater than that of the community subgroup (MD 0.24, 95% CI −0.16 to 0.64; I^2=64.5%; *P*=.060). After the senior care facility subgroup (only 1 study) was excluded, significant between-group differences were observed (*P*=.03). The effect size was greater for the application subgroup (MD 0.67, 95% CI −0.14 to 1.49) than for the exergaming subgroup (MD 0.14, 95% CI −0.16 to 0.44; I^2=81.5%; *P*=.005). However, after excluding the VR subgroup (only 1 study), the difference between groups was no longer statistically significant (*P*=.23). The 20-50 times subgroup yielded the strongest effect (MD 1.02, 95% CI 0.51-1.53), which was superior to that of the >50 times subgroup (MD 0.21, 95% CI −0.22 to 0.64; I^2=59.1%; *P*=.09). After the <20 times subgroup (only 1 study) was excluded, significant between-group differences were observed (*P*=.02). The intervention duration and qualifications of the interveners showed no significant between-group differences (*P*=.43 and *P*=.52, respectively; Table S1 in Multimedia Appendix 1). Sensitivity analysis revealed that no individual study was able to change the results.

Publication Bias

For the outcome of functional mobility, the funnel plot appeared asymmetrical; however, both the Egger test (*P*=.09) and the Begg test (*P*=.53) indicated no statistically significant publication bias (Figure S1 in Multimedia Appendix 1).

Narrative Synthesis

One study [44] was excluded from the quantitative synthesis because of inadequate data presentation, despite attempts to contact the authors. This study revealed that immersive VR-based balance training improved older adults' mobility (Timed Up and Go Test: MD −1.71, 95% CI −2.73 to −0.69) but did not alleviate concerns about falling (FES-I: MD 1.80, 95% CI −2.02 to 5.62), with no significant differences compared with the control group.

Discussion

Effect of Digital-Based Exercise Interventions on Falls Efficacy and Concerns About Falling

The meta-analysis assessed the efficacy of digital-based exercise interventions on concerns about falling and falls efficacy among older adults. Overall, the interventions were effective at improving falls efficacy but not concerns about falling, the emotional construct related to falls. This key distinction suggests that the cognitive and emotional constituents may represent 2 relatively independent intervention targets, requiring distinct strategies to address each effectively.

For falls efficacy, Gao et al [54] reported nonsignificant effects of VR interventions on falls efficacy measured by the MFES in nondisabled older adults. This difference may be because, compared with previous studies, this analysis pooled data from more studies, thus improving the statistical power and stability. The improvement in cognitive constructs through digital-based exercise interventions may be attributed to their ability to provide structured, repetitive exercises and feedback. These findings align with substantial evidence indicating that physical exercise directly enhances physical function and balance [55,56], thereby increasing self-efficacy. This process can be attributed to the underlying neurobiological mechanism: the multisensory stimulation (eg, visual, auditory, and vestibular) provided by digital exercise interventions promotes plasticity in the central nervous system, thereby enhancing the ability of the brain to process sensory information and maintain postural control [57-59]. Consequently, the experience of successfully completing motor challenges, reinforced directly through this neural reshaping, strengthens older adults' balance confidence or falls efficacy. Subgroup analyses explored the potential moderating effects of various intervention characteristics on falls efficacy, and no statistically significant between-group differences were found across all the subgroups, indicating that the improvement effect of digital-based exercise interventions on falls efficacy was consistent across different characteristics. For intervention setting, the effect size was the greatest in senior

XSL•FO
RenderX

care facilities, followed by hospitals and then community settings, which may be attributed to the more standardized implementation and higher participant compliance with interventions in institutional settings. For exercise mode, balance training yielded a slightly better effect than multimodal training, which was consistent with its core mechanism of directly improving postural control ability [60]. The absence of significant differences in the above dimensions might be due to the small sample size of each subgroup and the heterogeneity of intervention protocols across studies. The similar effect sizes across other subgroups, including interveners' qualifications, exposure dosage, and digital platform type, confirmed the robustness of such interventions, providing a reference for their flexible implementation in clinical and community settings.

For concerns about falling, our findings indicate that digital-based exercise interventions did not significantly improve concerns about falling, which is consistent with the findings of Lee et al [28], who used the FES-I. This lack of efficacy may be attributed to 2 primary reasons. First, concerns about falling, as an emotional construct, is fundamentally distinct from the cognitive construct. While a cognitive construct is closely tied to an individual's assessment of their balance capacity and can thus be improved through objective functional gains, the affective construct is characterized by concerns about falling, manifesting as a feeling of dread and apprehension toward situations perceived as threatening to balance [17]. This nonrational emotional response is resistant to standardized interventions that primarily target physical function alone. Second, there is a mismatch in the active components of the existing interventions. The digital interventions included in this analysis were predominantly based on physical training and lacked elements specifically designed to address worrying emotions, making them mechanistically inadequate for directly targeting the core of emotional constructs.

Effect of Digital-Based Exercise Interventions on Physical Performance

The pooled results demonstrated that digital-based exercise interventions significantly improved physical performance, including balance function, functional mobility, and physical function, in older adults. These findings are consistent with previous research. Most digital-based interventions implement structured exercise programs, particularly targeted balance training. These exercises have been proven to enhance lower limb strength, postural stability, and sensorimotor integration, which is the key physiological basis for evaluating balance function and activity ability [55,56]. Systematic exercise also improves the integrated physical capacity needed for daily activities, as reflected in physical function. In addition, digital platforms significantly amplify the effects of these exercises. Exergaming and immersive VR mitigate the monotony of traditional training, substantially increasing motivation, engagement, and adherence [61]. Critically, digital delivery overcomes spatial and temporal barriers, enhancing accessibility and convenience [62,63]. This allows for more frequent and consistent participation, enabling exercise benefits to accumulate fully. Furthermore, the improvement outcomes for both the moderate exposure dosage (20 - 50 times) and the lower dosage (<20 times) were significantly superior to those of the high dosage (>50 times), suggesting that beyond a certain threshold, further increases in exercise volume may not yield additional functional gains and could even diminish effects due to factors such as fatigue or reduced adherence [64].

Strengths and Limitations of the Study

This study has several strengths. First, this study provides the first meta-analytic evidence quantifying the effects of digital-based exercise interventions on both the cognitive and emotional constructs of fall-related psychological impairments among older adults, clearly demonstrating their differential effectiveness. Second, the study was rigorous, using comprehensive terminology and an extensive systematic search strategy. Only reliable evidence from RCTs was included, and registered protocols were adhered to, providing more precise effect estimates.

Moreover, our study has several limitations. First, this study included older adults with different baseline characteristics, and the heterogeneity of the population indicated that many subjects differed in terms of their pathophysiological behaviors and conditions, as well as their comorbidities, all of which may have affected the accuracy of the reported results. Second, most of the original studies included in this review were not designed primarily to improve "concerns about falling" or "falls efficacy," which could lead to an underestimation of the overall effect and an amplification of heterogeneity. The use of self-reported scales to measure fear of falling suggests the possibility of bias and may have led to inaccurate effect size estimates, especially for studies that did not have a blinded assignment to treatment versus control.

Suggestions for Further Research

While the study demonstrates limited efficacy in improving concerns about falling (an emotional construct), it reveals a more positive effect on falls efficacy (a cognitive construct). These differential outcomes underscore the need for more precisely targeted intervention designs in the future. Therefore, clinicians and researchers should select or design digital interventions on the basis of an individual's primary presentation.

To develop these more precisely targeted interventions, future research should prioritize the development of digital programs that integrate evidence-based psychological components, such as cognitive behavioral therapy and mindfulness training, to directly target the core mechanisms of emotional constructs. Future studies should employ more refined research designs to further clarify the specific pathways through which different intervention types (exercise, cognitive, and combined) influence distinct constructs. Furthermore, in terms of outcome assessment, in addition to self-report scales, future studies should explore the use of objective physiological biomarkers or behavioral tasks to more sensitively measure changes in emotional constructs.

Conclusions

The meta-analysis assessed the efficacy of digital-based exercise interventions on fall-related psychological impairments among older adults and revealed that the effects differed across the

XSL•FO
RenderX

constructs. These findings suggest that digital-based exercise interventions have potential benefits for improving falls efficacy and physical performance among older adults compared with controls. However, the effect of digital-based exercise interventions on concerns about falling, which is the emotional construct related to falls, remains uncertain among older adults.

Funding

The authors declared no financial support was received for this work.

Data Availability

The data supporting the findings of this study are available within the article and its online supplementary materials.

Authors' Contributions

Critical revision of the manuscript for important intellectual content: ZW, XZ, MZ, XX, BZ
Data acquisition: ZW, BZ
Data analysis and interpretation: ZW, BZ
Drafting of the manuscript: ZW, XX
Study concept and design: ZW, XZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy; subgroup analyses on balance (Berg Balance Scale), functional mobility (Timed Up and Go Test), and physical function (Short Physical Performance Battery); and funnel plots.
[DOCX File, 130 KB - aging_v9i1e87070_app1.docx]

Checklist 1
PRISMA checklist.
[PDF File, 98 KB - aging_v9i1e87070_app2.pdf]

References

1. Alawlah A, A'aqoulah A, Alamr R, El-Metwally A. Prevalence and risk factors of falls among the older adults in Gulf Cooperation Council countries: insights from a systematic review. BMC Public Health 2026 Jan 21;26(1):600. [doi: 10.1186/s12889-026-26348-x] [Medline: 41566329]
2. Salari N, Darvishi N, Ahmadipanah M, Shohaimi S, Mohammadi M. Global prevalence of falls in the older adults: a comprehensive systematic review and meta-analysis. J Orthop Surg Res 2022 Jun 28;17(1):334. [doi: 10.1186/s13018-022-03222-1] [Medline: 35765037]
3. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):36178003. [doi: 10.1093/ageing/afac205] [Medline: 36178003]
4. Friedman SM, Munoz B, West SK, Rubin GS, Fried LP. Falls and fear of falling: which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. J Am Geriatr Soc 2002 Aug;50(8):1329-1335. [doi: 10.1046/j.1532-5415.2002.50352.x] [Medline: 12164987]
5. Hadjistavropoulos T, Delbaere K, Fitzgerald TD. Reconceptualizing the role of fear of falling and balance confidence in fall risk. J Aging Health 2011 Feb;23(1):3-23. [doi: 10.1177/0898264310378039] [Medline: 20852012]
6. Peeters G, Bennett M, Donoghue OA, Kennelly S, Kenny RA. Understanding the aetiology of fear of falling from the perspective of a fear-avoidance model—a narrative review. Clin Psychol Rev 2020 Jul;79:101862. [doi: 10.1016/j.cpr.2020.101862] [Medline: 32442854]
7. Matsuda PN, Eagen T, Hreha KP, Finlayson ML, Molton IR. Relationship between fear of falling and physical activity in people aging with a disability. PM&R 2020 May;12(5):454-461. [doi: 10.1002/pmrj.12289] [Medline: 31736266]
8. Söylemez B, Çetİşlİ-Korkmaz N, TekİN S, BİR LS, Şenol H. The effect of balance, walking capacity, and fear of falling on the level of community integration in individuals with multiple sclerosis: a cross-sectional study. Physiother Theory Pract 2024 Sep;40(9):1974-1980. [doi: 10.1080/09593985.2023.2229903] [Medline: 37377114]
9. Gambaro E, Gramaglia C, Azzolina D, Campani D, Molin AD, Zeppegno P. The complex associations between late life depression, fear of falling and risk of falls. A systematic review and meta-analysis. Ageing Res Rev 2022 Jan;73:101532. [doi: 10.1016/j.arr.2021.101532] [Medline: 34844015]
10. Lee TH, Kim W. Is fear of falling and the associated restrictions in daily activity related to depressive symptoms in older adults? Psychogeriatrics 2021 May;21(3):304-310. [doi: 10.1111/psyg.12664] [Medline: 33576148]

XSL•FO
RenderX

11. Lee S, Hong GRS. The predictive relationship between factors related to fear of falling and mortality among community-dwelling older adults in Korea: analysis of the Korean Longitudinal Study of Aging from 2006 to 2014. Aging Ment Health 2020 Dec;24(12):1999-2005. [doi: 10.1080/13607863.2019.1663490] [Medline: 31512495]
12. Kruisbrink M, Crutzen R, Kempen GIJM, et al. Disentangling interventions to reduce fear of falling in community-dwelling older people: a systematic review and meta-analysis of intervention components. Disabil Rehabil 2022 Oct;44(21):6247-6257. [doi: 10.1080/09638288.2021.1969452] [Medline: 34511009]
13. Kruisbrink M, Delbaere K, Kempen GIJM, et al. Intervention characteristics associated with a reduction in fear of falling among community-dwelling older people: a systematic review and meta-analysis of randomized controlled trials. Gerontologist 2021 Aug 13;61(6):e269-e282. [doi: 10.1093/geront/gnaa021] [Medline: 32267498]
14. Lenouvel E, Ullrich P, Siemens W, et al. Cognitive behavioural therapy (CBT) with and without exercise to reduce fear of falling in older people living in the community. Cochrane Database Syst Rev 2023 Nov 15;11(11):CD014666. [doi: 10.1002/14651858.CD014666.pub2] [Medline: 37965937]
15. Ellmers TJ, Ventre JP, Freiberger E, et al. Does concern about falling predict future falls in older adults? A systematic review and meta-analysis. Age Ageing 2025 Mar 28;54(4):afaf089. [doi: 10.1093/ageing/afaf089] [Medline: 40197783]
16. Soh SLH, McCrum C, Okubo Y, et al. About falls efficacy: a commentary on “World guidelines for falls prevention and management for older adults: a global initiative”. J Frailty Sarcopenia Falls 2024 Dec;9(4):281-285. [doi: 10.22540/JFSF-09-281] [Medline: 39635562]
17. Ellmers TJ, Wilson MR, Kal EC, Young WR. The perceived control model of falling: developing a unified framework to understand and assess maladaptive fear of falling. Age Ageing 2023 Jul 1;52(7):afad093. [doi: 10.1093/ageing/afad093] [Medline: 37466642]
18. Soh SLH. Falls efficacy: The self-efficacy concept for falls prevention and management. Front Psychol 2022;13:1011285. [doi: 10.3389/fpsyg.2022.1011285] [Medline: 36438360]
19. Dong M, Huang Y, Zhang X, Wu L, Ghazal L, Cao F. Effects of psychological intervention based on digital health technology on psychological outcomes in cancer survivors: systematic review and meta-analysis. Psychooncology 2025 May;34(5):e70166. [doi: 10.1002/pon.70166] [Medline: 40325624]
20. Fatehi F, Samadbeik M, Kazemi A. What is digital health? Review of definitions. Stud Health Technol Inform 2020 Nov 23;275:67-71. [doi: 10.3233/SHTI200696] [Medline: 33227742]
21. Jimenez-Mazuelas MJ, de Villar NGP, De Casas-Albendea S, Martinez-Gimeno L, Jimenez-Gonzalez S, Angulo-Carrere MT. Somatosensory and dynamic balance improvement in older adults with diabetic peripheral neuropathy through sensorimotor exercise: a multisite randomized controlled trial. J Bodyw Mov Ther 2024 Oct;40:2062-2073. [doi: 10.1016/j.jbmt.2024.10.041] [Medline: 39593566]
22. Qiao X, Ji L, Jin Y, et al. Effectiveness of a theory-underpinning exercise intervention among community-dwelling (pre)frail older adults: a stepped-wedge cluster-randomized trial. Int J Nurs Stud 2025 Jan;161:104933. [doi: 10.1016/j.ijnurstu.2024.104933] [Medline: 39500251]
23. Bernard PL, de la Tribonniere X, Pellecchia A, et al. Effects on physical functioning and fear of falling of a 3-week balneotherapy program alone or associated with a physical activity and educational program in older adult fallers: a randomized-controlled trial. Clin Interv Aging 2024;19:1753-1763. [doi: 10.2147/CIA.S472816] [Medline: 39494368]
24. Niu XD, Chi JT, Guo JB, et al. Effects of nurse-led web-based interventions on people with type 2 diabetes mellitus: a systematic review and meta-analysis. J Telemed Telecare 2021 Jun;27(5):269-279. [doi: 10.1177/1357633X211010019] [Medline: 33966522]
25. He Z, Wu H, Zhao G, et al. The effectiveness of digital technology–based Otago Exercise Program on balance ability, muscle strength and fall efficacy in the elderly: a systematic review and meta-analysis. BMC Public Health 2025 Jan 7;25(1):71. [doi: 10.1186/s12889-024-21251-9] [Medline: 39773392]
26. Park BM, Choi H, Jeong H. Effects of virtual reality based on fall prevention intervention: a systematic review and meta-analysis. Healthcare (Basel) 2025 Jul 29;13(15):1845. [doi: 10.3390/healthcare13151845] [Medline: 40805878]
27. Saragih ID, Chen YM, Suarilah I, Susanto H, Lee BO. Virtual reality intervention for fall prevention in older adults: a meta-analysis. J Nurs Scholarsh 2025 Sep;57(5):759-775. [doi: 10.1111/jnu.70028] [Medline: 40533914]
28. Lee J, Phu S, Lord SR, Okubo Y. Effects of immersive virtual reality training on balance, gait and mobility in older adults: a systematic review and meta-analysis. Gait Posture 2024 May;110:129-137. [doi: 10.1016/j.gaitpost.2024.03.009] [Medline: 38581933]
29. Cortés-Pérez I, Osuna-Pérez MC, Montoro-Cárdenas D, Lomas-Vega R, Obrero-Gaitán E, Nieto-Escamez FA. Virtual reality–based therapy improves balance and reduces fear of falling in patients with multiple sclerosis. a systematic review and meta-analysis of randomized controlled trials. J Neuroeng Rehabil 2023 Apr 11;20(1):42. [doi: 10.1186/s12984-023-01174-z] [Medline: 37041557]
30. Ambrens M, Alley S, Oliveira JS, et al. Effect of eHealth-delivered exercise programmes on balance in people aged 65 years and over living in the community: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2022 Jun 10;12(6):e051377. [doi: 10.1136/bmjopen-2021-051377] [Medline: 35688586]

XSL•FO
RenderX

31. Thwaites C, Nayyar R, Blennerhassett J, Egerton T, Tan J, Bower K. Is telehealth an effective and feasible option for improving falls-related outcomes in community-dwelling adults with neurological conditions? A systematic review and meta-analysis. Clin Rehabil 2023 Jan;37(1):17-46. [doi: 10.1177/02692155221133468] [Medline: 36263524]
32. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
33. Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011 Jul 22;343:d4002. [doi: 10.1136/bmj.d4002] [Medline: 21784880]
34. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634. [doi: 10.1136/bmj.315.7109.629] [Medline: 21784880]
35. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994 Dec;50(4):1088-1101. [Medline: 7786990]
36. Lee S, Shin S. Effectiveness of virtual reality using video gaming technology in elderly adults with diabetes mellitus. Diabetes Technol Ther 2013 Jun;15(6):489-496. [doi: 10.1089/dia.2013.0050] [Medline: 23560480]
37. Kwok BC, Pua YH. Effects of WiiActive exercises on fear of falling and functional outcomes in community-dwelling older adults: a randomised control trial. Age Ageing 2016 Sep;45(5):621-627. [doi: 10.1093/ageing/afw108] [Medline: 27496921]
38. Morone G, Paolucci T, Luziatelli S, et al. Wii Fit is effective in women with bone loss condition associated with balance disorders: a randomized controlled trial. Aging Clin Exp Res 2016 Dec;28(6):1187-1193. [doi: 10.1007/s40520-016-0578-6] [Medline: 27154875]
39. van den Berg M, Sherrington C, Killington M, et al. Video and computer-based interactive exercises are safe and improve task-specific balance in geriatric and neurological rehabilitation: a randomised trial. J Physiother 2016 Jan;62(1):20-28. [doi: 10.1016/j.jphys.2015.11.005] [Medline: 26701163]
40. Liao YY, Chen IH, Wang RY. Effects of Kinect-based exergaming on frailty status and physical performance in prefrail and frail elderly: a randomized controlled trial. Sci Rep 2019 Jun 27;9(1):9353. [doi: 10.1038/s41598-019-45767-y] [Medline: 31249332]
41. Montero-Alía P, Miralles-Basseda R, López-Jiménez T, et al. Controlled trial of balance training using a video game console in community-dwelling older adults. Age Ageing 2019 Jul 1;48(4):506-512. [doi: 10.1093/ageing/afz047] [Medline: 31081504]
42. Sun Z, Gu X, Ouyang G, et al. Effects of Baduanjin based on virtual reality on fall risk and quality of life of elderly osteoporosis patients in nursing institution for the aged. China Med Herald 2020;17(1):90-97. [doi: 10.20047/j.issn1673-7210.2020.01.023]
43. Khushnood K, Altaf S, Sultan N, Ali Awan MM, Mehmood R, Qureshi S. Role Wii Fit exer-games in improving balance confidence and quality of life in elderly population. J Pak Med Assoc 2021 Sep;71(9):2130-2134. [doi: 10.47391/JPMA.319] [Medline: 34580500]
44. Rebêlo FL, de Souza Silva LF, Doná F, Sales Barreto A, de Souza Siqueira Quintans J. Immersive virtual reality is effective in the rehabilitation of older adults with balance disorders: a randomized clinical trial. Exp Gerontol 2021 Jul 1;149:111308. [doi: 10.1016/j.exger.2021.111308] [Medline: 33744393]
45. Zahedian-Nasab N, Jaberi A, Shirazi F, Kavousipor S. Effect of virtual reality exercises on balance and fall in elderly people with fall risk: a randomized controlled trial. BMC Geriatr 2021 Sep 25;21(1):509. [doi: 10.1186/s12877-021-02462-w] [Medline: 34563120]
46. Tekin F, Cetisli-Korkmaz N. Effectiveness of a telerehabilitative home exercise program on elder adults' physical performance, depression and fear of falling. Percept Mot Skills 2022 Jun;129(3):714-730. [doi: 10.1177/00315125221087026] [Medline: 35404709]
47. Yang X, Ma C, Jing S. The effect of internet-based OEP exercise on the psychology and mobility of older adults with fear of falling. J Clin Nurs 2022;21(4):53-55. [doi: 10.3969/j.issn.1671-8933.2022.04.018]
48. Gerards M, Marcellis R, Senden R, et al. The effect of perturbation-based balance training on balance control and fear of falling in older adults: a single-blind randomised controlled trial. BMC Geriatr 2023 May 17;23(1):305. [doi: 10.1186/s12877-023-03988-x] [Medline: 37198543]
49. Hager AGM, Mathieu N, Carrard S, Bridel A, Wapp C, Hilfiker R. Partially supervised exercise programmes for fall prevention improve physical performance of older people at risk of falling: a three-armed multi-centre randomised controlled trial. BMC Geriatr 2024 Apr 3;24(1):311. [doi: 10.1186/s12877-024-04927-0] [Medline: 38570773]
50. Prieto-Moreno R, Mora-Traverso M, Estévez-López F, et al. Effects of the ActiveHip+ mHealth intervention on the recovery of older adults with hip fracture and their family caregivers: a multicentre open-label randomised controlled trial. EClinicalMedicine 2024 Jul;73:102677. [doi: 10.1016/j.eclinm.2024.102677] [Medline: 38911836]
51. Sturnieks DL, Hicks C, Smith N, et al. Exergame and cognitive training for preventing falls in community-dwelling older people: a randomized controlled trial. Nat Med 2024 Jan;30(1):98-105. [doi: 10.1038/s41591-023-02739-0] [Medline: 38228913]
52. Vásquez-Carrasco E, Gómez CS, Valdés-Badilla P, et al. Xbox Kinect Sports effects on cognition status and physical performance in physically inactive older females: a randomized controlled trial. J Clin Med 2025 Mar 22;14(7):2165. [doi: 10.3390/jcm14072165] [Medline: 40217616]

XSL•FO
RenderX

53. Lee K. Home-based exergame program to improve physical function, fall efficacy, depression and quality of life in community-dwelling older adults: a randomized controlled trial. Healthcare (Basel) 2023 Apr 12;11(8):1109. [doi: 10.3390/healthcare11081109] [Medline: 37107943]
54. Gao Y, Wang N, Liu Y, Liu N. Effectiveness of virtual reality in preventing falls in non-disabled older adults: a meta-analysis and systematic review. Geriatr Nurs 2024;58:15-25. [doi: 10.1016/j.gerinurse.2024.04.010] [Medline: 38729063]
55. Duan Y, Yang M, Wang Y, et al. Effects of a blended face-to-face and eHealth lifestyle intervention on physical activity, diet, and health outcomes in Hong Kong community-dwelling older adults: a randomized controlled trial. BMC Public Health 2025 Jun 10;25(1):2145. [doi: 10.1186/s12889-025-23311-0] [Medline: 40495124]
56. Wan R, Huang J, Wang K, et al. Effectiveness of mind-body exercise in older adults with sarcopenia and frailty: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2025 Apr;16(2):e13806. [doi: 10.1002/jcsm.13806] [Medline: 40254030]
57. Li RO, Zhang Y, Jiang Y, Wang M, Ang WHD, Lau YO. Rehabilitation training based on virtual reality for patients with Parkinson's disease in improving balance, quality of life, activities of daily living, and depressive symptoms: a systematic review and meta-regression analysis. Clin Rehabil 2021 Aug;35(8):1089-1102. [doi: 10.1177/0269215521995179] [Medline: 33588583]
58. Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? Trends Cogn Sci 2006 Jun;10(6):278-285. [doi: 10.1016/j.tics.2006.04.008] [Medline: 16713325]
59. Liu D, Guo ZQ, Tian E, Zhu YT, Xu WC, Zhang SL. Efficacy of multisensory training and rehabilitation in patients with balance disorders. Zhonghua Yi Xue Za Zhi 2021 Jul 13;101(26):2050-2054. [doi: 10.3760/cma.j.cn112137-20210116-00143] [Medline: 34275237]
60. Talkowski JB, Brach JS, Studenski S, Newman AB. Impact of health perception, balance perception, fall history, balance performance, and gait speed on walking activity in older adults. Phys Ther 2008 Dec;88(12):1474-1481. [doi: 10.2522/ptj.20080036] [Medline: 18849479]
61. Xie H, Cong S, Wang R, et al. Effect of eHealth interventions on perinatal depression: a meta-analysis. J Affect Disord 2024 Jun 1;354:160-172. [doi: 10.1016/j.jad.2024.03.027] [Medline: 38490593]
62. Duque G, Boersma D, Loza-Diaz G, et al. Effects of balance training using a virtual-reality system in older fallers. Clin Interv Aging 2013;8:257-263. [doi: 10.2147/CIA.S41453] [Medline: 23467506]
63. Zak M, Sikorski T, Krupnik S, et al. Physiotherapy programmes aided by VR solutions applied to the seniors affected by functional capacity impairment: randomised controlled trial. Int J Environ Res Public Health 2022 May 15;19(10):6018. [doi: 10.3390/ijerph19106018] [Medline: 35627554]
64. Holtzer R, Shuman M, Mahoney JR, Lipton R, Verghese J. Cognitive fatigue defined in the context of attention networks. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2011 Jan;18(1):108-128. [doi: 10.1080/13825585.2010.517826] [Medline: 21128132]

Abbreviations

FES–I: Falls Efficacy Scale–International
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
MD: mean difference
MeSH: Medical Subject Headings
MFES: Modified Falls Efficacy Scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
RoB2: Risk of Bias 2
SMD: standardized mean difference
VR: virtual reality

Edited by M Gray; submitted 04.Nov.2025; peer-reviewed by F Li, T Ellmers; revised version received 10.Feb.2026; accepted 11.Mar.2026; published 30.Apr.2026.

Please cite as:
Wang Z, Zhu B, Zhou M, Xie X, Zhang X
Effects of Digital-Based Exercise Interventions on Concerns About Falling, Falls Efficacy, and Physical Performance Among Older Adults: Systematic Review and Meta-Analysis
JMIR Aging 2026;9:e87070
URL: https://aging.jmir.org/2026/1/e87070
doi:10.2196/87070

XSL•FO
RenderX

© Zhaojun Wang, Bochi Zhu, Meng Zhou, Xiaojie Xie, Xueyan Zhang. Originally published in JMIR Aging (https://aging.jmir.org), 30.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Associations Between Digital Exclusion and Physical or Cognitive Function in Middle-Aged and Older Adults: Systematic Review and Meta-Analysis

Jia Yu[1*], MSN; Jundan Huang[1*], PhD; Shuhan Zhou[1], PhD; Xiao Wan[1], PhD; Qi Xie[1], PhD; Yinan Zhao[1], PhD; Juan Yang[2], MSN; Hui Feng[1,3,4], PhD

[1]Xiangya School of Nursing, Central South University, No. 172, Tongzipo Road, Yuelu District, Changsha, Hunan, China

[2]Hunan Traditional Chinese Medical College, Zhuzhou, Hunan, China

[3]Oceanwide Health Management Institute, Central South University, Changsha, Hunan, China

[4]Hunan Engineering Research Center for Intelligent Medical Care, Central South University, Changsha, Hunan, China

[*]these authors contributed equally

Corresponding Author:
Hui Feng, PhD
Xiangya School of Nursing, Central South University, No. 172, Tongzipo Road, Yuelu District, Changsha, Hunan, China

Abstract

Background: Digital exclusion posed a significant challenge, especially in middle-aged and older adults, which affected their health outcomes. However, the evidence regarding the associations of digital exclusion on physical or cognitive function outcomes was controversial, and no systematic review had been performed to synthesize the pooled associations.

Objective: This study aimed to explore the relationship between digital exclusion and physical or cognitive function in middle-aged and older adults.

Methods: We conducted a systematic review and meta-analysis of cohort and cross-sectional studies, including Chinese or English publications retrieved from PubMed, Embase, Web of Science, PsycINFO, Scopus, CNKI, and Wanfang databases up to August 31, 2024. The risk of bias was evaluated using the Newcastle-Ottawa Scale (NOS). The pooled effect size was calculated based on odds ratios (ORs), hazard ratios, risk ratios, and 95% CIs. This study was registered on PROSPERO (CRD42024585459).

Results: Nineteen studies met the inclusion criteria, including 13 cohort studies and 6 cross-sectional studies, which had moderate-to-low risk of bias. The pooled analysis indicated that digital exclusion had prospective associations with decreased basic activities of daily living (incidence rate ratio 1.35, 95% CI 1.12 - 1.64, I^2=94.7%) and instrumental activities of daily living (incidence rate ratio 1.46, 95% CI 1.13 - 1.89, I^2=96.2%), or cross-sectional association with activities of daily living (OR 1.23, 95% CI 0.41 - 3.73, I^2=91%), with no statistical significance in the prospective association with frailty (OR 1.21, 95% CI 0.92 - 1.59, I^2=95.2%). There were prospective associations between digital exclusion and dementia (hazard ratio 1.78, 95% CI 1.43 - 2.22, I^2=0%), decreased Mini-Mental State Examination scores (OR 1.96, 95% CI 1.39 - 2.75, I^2=0%), as well as cross-sectional associations with Mini-Mental State Examination scores (OR 2.90, 95% CI 2.07 - 4.07, I^2=0%), and no statistical significance in the prospective association with cognitive impairment (risk ratio 2.08, 95% CI 0.98 - 4.44, I^2=78.2%).

Conclusions: Our findings indicated the negative associations of digital exclusion with physical and cognitive functions. Future research and clinical practice should consider designing digital interventions and services that match the physical and cognitive capacities and preferences of middle-aged and older adults, thereby improving digital engagement and reducing the harms associated with digital exclusion. Policies should focus on expanding access, reducing financial barriers, and improving digital literacy. However, due to the presence of heterogeneity and publication bias, the results should be interpreted cautiously.

(JMIR Aging 2026;9:e75920) doi:10.2196/75920

KEYWORDS

digital exclusion; internet nonuse; physical function; cognitive function; meta-analysis

Introduction

Population aging has become a global concern, posing significant challenges to public health systems and social development [1]. Although aging is an inevitable biological process, a growing body of research has focused on promoting healthy aging, which emphasizes maintaining and improving intrinsic capacity to enhance overall well-being among

XSL•FO
RenderX

middle-aged and older adults [2-4]. The World Health Organization defines healthy aging as "the process of developing and maintaining the functional ability that enables well-being in older age." [5] This functional ability is underpinned by 2 core components: physical and cognitive functions [6]. Physical function provides the strength, balance, and mobility necessary for action, while cognitive function supports the attention, memory, and decision-making needed to guide those actions effectively and safely [7-10]. Preserving both is essential for maintaining independence and preventing disability in later life, and accumulating evidence suggests that midlife might represent a valuable window for earlier monitoring or intervention [11,12].

Rapid digitalization has profoundly transformed the way individuals access information, communicate, and engage with health and social services, such as monitoring age-related diseases and promoting remote health management [13-15]. However, numerous middle-aged and older adults with substantial health care needs are excluded from digital technologies. They are often unwilling or unable to use digital technology critically, which exacerbates the vicious cycle of digital exclusion and health inequalities [16,17]. The prevalence of digital exclusion varies by country and is especially obvious in older adults, accounting for 21.69% in Denmark, 22.61% in the United Kingdom, 60.78% in Mexico, and 97.15% in China [18]. Digital exclusion currently lacks a unified definition and assessment tool. While some studies define it as the ability to send and receive online messages with difficulty [19], it is most commonly defined and measured as self-reported nonuse of the internet [18,20,21].

Growing evidence suggests that digital exclusion is related to adverse health outcomes, such as functional dependence [22], cognitive impairment [18], and psychological distress [23], ultimately elevating the risk of mortality and exacerbating socioeconomic burdens [24]. The evidence on the negative impact of digital exclusion on psychological well-being is particularly well-documented and has been proven to be associated with higher levels of depression [25,26], loneliness [21,27], and social isolation [28,29] in large cohort studies and systematic reviews.

However, evidence regarding the associations of digital exclusion on physical or cognitive function outcomes was controversial, and no systematic review had been performed to synthesize the pooled associations. Several studies analyzed data from 5 nationally representative cohorts conducted in both developed and developing countries. These studies reported that nonuse of the internet was associated with increased risks of cognitive impairment, functional dependence, and frailty [17,18,30]. However, Liu et al [31] found no statistically significant association between internet nonuse and functional disability in China, and García-Esquinas et al [32] observed a similar nonsignificant relationship between nonuse of computers and frailty in European populations. This inconsistency may arise from differences in study populations, definitions of digital exclusion, measurement methods, and sample sizes. To date, no systematic review has comprehensively synthesized the available evidence or assessed the overall magnitude and direction of these associations. A rigorous synthesis is therefore needed to clarify the impact of digital exclusion on functional health, identify potential sources of heterogeneity, and provide an evidence-based foundation for interventions aimed at promoting healthy aging in digitally marginalized populations.

Therefore, this study aims to systematically identify studies regarding the relationship between digital exclusion and physical or cognitive function in middle-aged and older adults, summarize their findings, and assess the certainty of their evidence to provide a comprehensive understanding of the associations.

Methods

Study Design

This review adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement [33] and was registered on PROSPERO (CRD42024585459; Checklist 1).

Search Strategies

We searched PubMed, Embase, Web of Science, PsycINFO, Scopus, CNKI, and Wanfang databases, from the inception of the databases to August 31, 2024. Only eligible full texts in English or Chinese were considered for review. Two thematic blocks of keywords were used: "Aged" and "Digital exclusion." The search terms were (aged or "middle aged" or "older adult*" or aging or "elder*") AND ("digital exclusion" or "digital inclusion" or "digital divide" or "Internet access" or "Internet use*" or "Internet non-use" or "web usage" or "web use" or "Internet usage" or "digital engagement"). In terms of developing search strategies, we first identified standardized terminology by consulting the Medical Subject Headings (MeSH) database in PubMed. This step ensured the inclusion of entry terms. Subsequently, we examined key published literature to incorporate nonstandardized terms that are conceptually related to digital exclusion, such as "digital divide," "digital inclusion," and "digital engagement" [27,34,35]. The detailed search strategies are shown in Multimedia Appendix 1.

Eligibility Criteria

The inclusion criteria for studies were as follows: (1) middle-aged and older adults as research participants, explicitly aged 45 years and older [36]; (2) digital exclusion was clearly specified as the exposure variable, without redefinition based on dose or frequency; and (3) reporting at least one measure of physical or cognitive function. Specifically, physical function was measured using standardized assessments of functional independence and mobility (eg, activities of daily living [ADLs] and Fried frailty phenotype). Cognitive function was assessed using standardized screening tools (eg, Mini-Mental State Examination [MMSE]) or based on a clinical diagnosis of cognitive impairment or dementia; (4) provision of the odds ratios (ORs), hazard ratios (HRs), risk ratios (RRs), and their corresponding 95% CIs for the associations between digital exclusion and health outcomes; and (5) observational studies.

This study excluded (1) editorials, letters, comments, meeting abstracts, and newspaper articles; (2) studies unable to provide the full text; (3) studies that only reported correlation metrics,

XSL•FO
RenderX

such as Pearson *r* or Spearman ρ, without providing data that could be used to calculate ORs, RRs, HRs, and their 95% CIs; and (4) studies that were not published in English or Chinese.

Study Selection and Data Extraction

Study selection was conducted independently in 2 stages by 2 researchers. A third researcher was invited to resolve the disagreement. Initially, the titles and abstracts of the studies were evaluated in terms of eligibility criteria. In the subsequent stage, full texts of preselected studies were reviewed to confirm their eligibility.

The selected studies were extracted by 2 researchers using a form to extract data on the setting, study design, sample size, basic participant information, digital exclusion assessment tools, categories of digital exclusion, outcome assessment tools, categories of assessment, covariates, and effect sizes of outcomes. Discrepancies between the researchers regarding data extraction were resolved through discussion and consensus with the assistance of a third researcher.

Methodological Quality Assessment

Two researchers independently assessed the methodological quality of each study using the Newcastle-Ottawa Scale (NOS), a quality assessment tool for both cohort and cross-sectional studies [37]. The tool comprises 8 items with a score range of 0‐9 for cohort studies, and 7 items from 0 to 8 for cross-sectional studies [38], as detailed in Multimedia Appendix 2 [17,18,30,32,39-47] and Multimedia Appendix 3 [19,24,31,48-50]. Initially, we conducted a pilot evaluation of 4 articles (2 cohort studies and 2 cross-sectional studies) to standardize the assessment criteria. Discrepancies between the researchers were discussed, and a third researcher was invited to resolve them collectively.

Synthesis and Data Analysis

Descriptive analysis was conducted to summarize the characteristics of the studies. For meta-analysis, the associations between digital exclusion and health outcomes were quantitatively synthesized, and effect sizes were measured using RR, OR, HR, and 95% CIs. Separate meta-analyses were conducted for ORs, RRs, and HRs. Effect sizes adjusted for potential confounders were extracted to ensure comparability across studies, allowing them to be regarded as standardized estimates [51]. When the original studies used "digital exclusion" as the reference group, the reported effect sizes were inverted to represent the impact of digital exclusion versus nonexclusion. This conversion was justified by the mathematical principle that inverting a ratio measure was equivalent to swapping the reference and exposure groups [52]. And all conversions were performed with double-checking for accuracy.

The random-effects model was used if significant heterogeneity (I^2>50% or P<.1) was found among the included studies; otherwise, the fixed-effect model was used. Due to the inclusion of fewer than 10 studies in the meta-analysis, publication bias analysis was not conducted. The meta-analysis and forest plots were performed using R software (R Core Team) with the main package of "*metafor*."

Sensitivity Analysis

Leave-one-out sensitivity analysis was conducted to evaluate the influence of each study on the pooled association estimate. In this procedure, the meta-analysis was repeated multiple times, each time omitting one study and recalculating the pooled association using the remaining studies. When between-study heterogeneity (I^2) exceeded 50%, a random-effects model was applied for the recalculated estimate [53]. Specifically, a study was considered to have a substantial influence if its exclusion caused the new point estimate to fall outside the 95% CI of the original pooled estimate. A study was also deemed influential if the effect size changed noticeably compared to the primary analysis. Noticeable changes were defined as either a reduction of 25‐50 percentage points or a shift from high heterogeneity (I^2>50%) to low or moderate heterogeneity (I^2≤50%) [54-56]. Due to the limited number of available studies, subgroup analysis based on different definitions, frequencies, or ages could not be performed. All results were presented visually using forest plots, along with corresponding CIs. A statistical significance threshold of P<.10 was adopted.

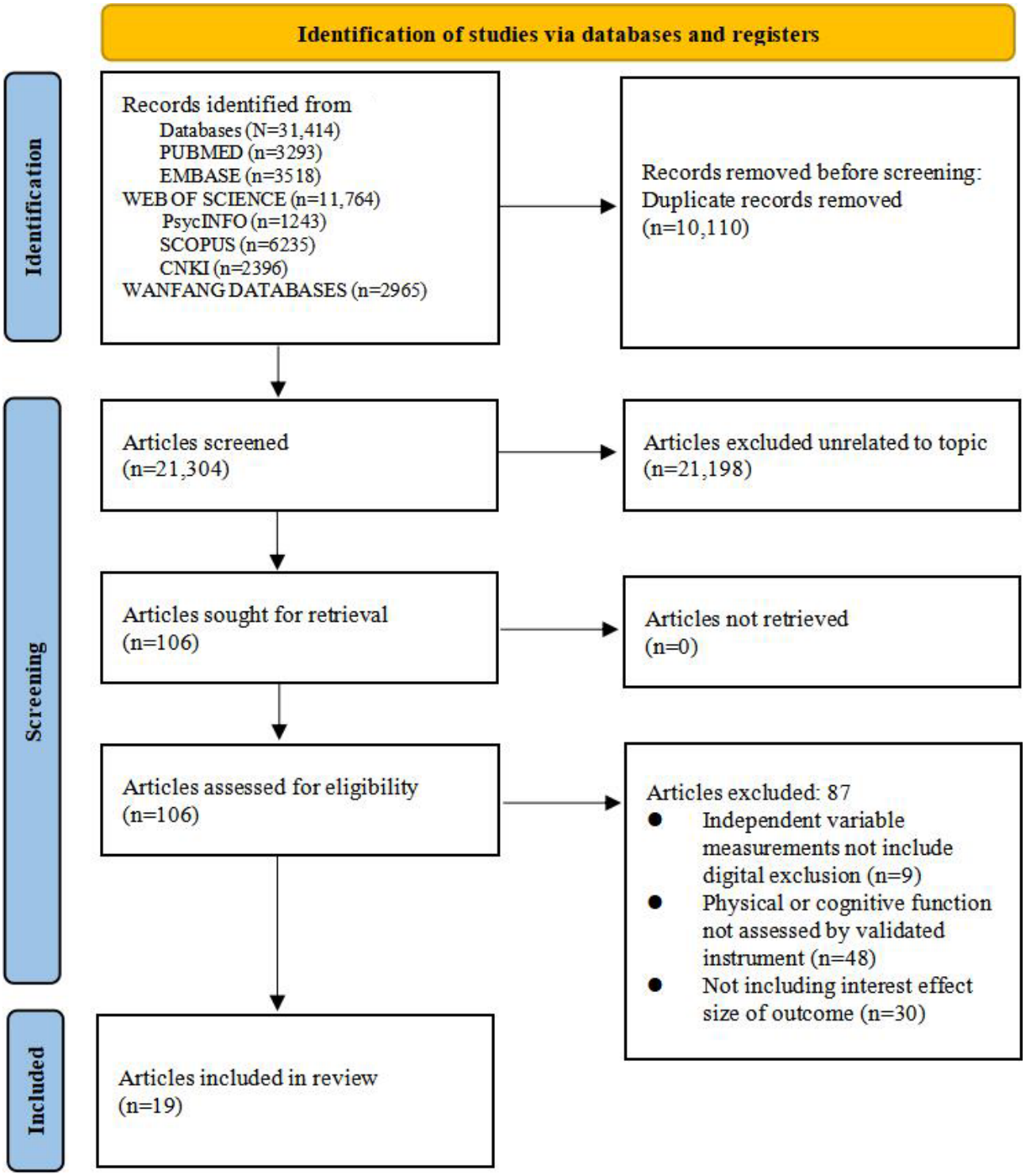
Results

Characteristics of the Studies

A total of 31,414 articles were searched from the databases. After excluding duplicate publications (n=10,110) and ineligible articles (n=21,198), 19 studies were included in this review (Multimedia Appendix 4) [17-19,24,30-32,39-50]. The selection process of the studies was summarized in the PRISMA flow diagram (Figure 1).

XSL•FO
RenderX

Figure 1. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Tables 1 and 2 provide the characteristics of the included cohort and cross-sectional studies, respectively. Three studies focused on middle-aged and older adults aged 45 years or older [24,42,45], while the remaining articles specifically concentrated on older adults aged 60 years or older [17-19,30-32,39-41,43,44,46-50]. Thirteen cohort studies [17,18,30,32,39-47] and 6 cross-sectional studies [19,24,31,48-50] were from different regions, including Spain, England, Japan, the United States, China, Europe, Mexico, Brazil, Sweden, and the Netherlands. These studies were conducted between 2012 and 2024, with the sample sizes ranging from 409 to 155,695. The follow-up duration of the cohort studies ranged from 3 to 17.1 years.

XSL•FO
RenderX

Table . Characteristics of the included cohort studies.

Study	Year	Country	Follow-up (year)	Sample size	Age (years), mean (SD)	Sex (male), n (%)	Digital exclusion definition	Measurements on outcome (outcome)	OR[a]/RR[b]/HR[c] (95%CI)
Physical function									
García-Esquinas et al [32]	2017	Spain	4	1882	≥60	912 (48.5)	First tertile of the usual number of hours or days spent on computers	Fried frailty phenotype (physical frailty)	• OR 0.909 (0.813-1)
García-Esquinas et al [32]	2017	England	4	3989	≥60	1893 (47.5)	First tertile of the usual number of hours or days spent on computers	Fried frailty phenotype (physical frailty)	• OR 1.000 (0.952-1.041)
García-Esquinas et al [32]	2017	Spain and England	4	5871	≥60	2805 (47.8)	First tertile of the usual number of hours or days spent on computers	Fried frailty phenotype (physical frailty)	• Random-effects meta-analysis OR 1.05 (0.95-1.13)
Li [17]	2024	United States	6	62,932	66.24 (10.28)	25,928 (41.2)	Nonuse of the internet	Frailty index (frailty)	• OR 1.299 (1.250-1.351)
Li [17]	2024	China	7	36,866	62.32 (8.37)	17,418 (47.25)	Nonuse of the internet in the past month	Frailty index (frailty)	• OR 1.613 (1.429-1.852)
Li [17]	2024	Europe	6	124,926	68.00 (9.75)	54,307 (43.47)	Use the internet less than once a week in the last 7 days	Frailty index (frailty)	• OR 2.128 (1.316-3.448)
Li [17]	2024	England	6	27,146	67.65 (9.25)	12,325 (45.40)	Nonuse of the internet	Frailty index (frailty)	• OR 1.408 (1.299-1.515)
Li [17]	2024	Mexico	6	34,273	65.15 (9.42)	14,404 (42.03)	Nonuse of the internet	Frailty index (frailty)	• OR 1.235 (1.176-1.316)

XSL•FO
RenderX

Study	Year	Country	Follow-up (year)	Sample size	Age (years), mean (SD)	Sex (male), n (%)	Digital exclusion definition	Measurements on outcome (outcome)	OR[a]/RR[b]/HR[c] (95%CI)
Li [17]	2024	32 countries	6 - 7	155,695	62 - 68	66,949 (43)	Nonuse of the internet	Frailty index (frailty)	• Random-effects meta-analysis OR 1.389 (1.493-1.266)
Tomioka et al [39]	2024	Japan	3	7913	≥65	3671 (46.4)	Nonuse of the internet	Public long-term care insurance (incident disability)	• Several times a year RR 1.205 (0.893-1.639) • Several times a month RR 1.031 (0.769-1.389) • Several times a week RR 1.205 (0.935-1.538) • Almost every day RR 1.493 (1.176-1.887)
Lu et al [30]	2022	United States	6	49,583	72	20,469 (41.3)	Nonuse of the internet	BADLs[d] and IADLs[e] (functional ability)	• BADLs IRR[f] 1.31 (1.24-1.37) • IADLs IRR 1.54 (1.45-1.64)
Lu et al [30]	2022	England	8	27,338	69	12,853 (47)	Nonuse of the internet	BADLs and IADLs (functional ability)	• BADLs IRR 1.21 (1.13-1.30) • IADLs IRR 1.25 (1.17-1.33)

XSL•FO
RenderX

Study	Year	Country	Follow-up (year)	Sample size	Age (years), mean (SD)	Sex (male), n (%)	Digital exclusion definition	Measurements on outcome (outcome)	OR[a]/RR[b]/HR[c] (95%CI)
Lu et al [30]	2022	Europe	4	96,184	70	42,160 (43.8)	Use the internet less than once a week in the last 7 days	BADLs and IADLs (functional ability)	• BADLs IRR 1.43 (1.36-1.51) • IADLs IRR 1.43 (1.37-1.49)
Lu et al [30]	2022	China	8	23,342	67	11,261 (48.2)	Nonuse of the internet in the past month	BADLs and IADLs (functional ability)	• BADLs IRR 2.04 (1.63-2.55) • IADLs IRR 2.45 (1.95-3.09)
Lu et al [30]	2022	Mexico	6	26,968	69	11,854 (44)	Nonuse of the internet	BADLs and IADLs (functional ability)	• BADLs IRR 1.07 (1.01-1.13) • IADLs IRR 1.07 (1.00-1.14)
Cognitive function									

XSL•FO
RenderX

Study	Year	Country	Follow-up (year)	Sample size	Age (years), mean (SD)	Sex (male), n (%)	Digital exclusion definition	Measurements on outcome (outcome)	OR[a]/RR[b]/HR[c] (95%CI)
Krug et al [40]	2019	Brazil	4	1197	≥60	NR[g]	Continued not using the internet	MMSE[h] (overall cognitive function)	• Cognitive loss OR 2.564 (1.136-5.882)
Krug et al [40]	2019	Brazil	4	1197	≥60	NR	Continued not using the internet	MMSE (overall cognitive function)	• Cognitive gain OR 0.301 (0.102-0.885)
Wang et al [18]	2024	China	8	7935	67.35 (5.97)	8872 (53.13)	Nonuse of the internet in the past month	Orientation, memory, and executive function tests	• OR 2.81 (1.84-4.28)
Wang et al [18]	2024	England	4	6824	70.91 (7.48)	7267 (47.59)	Nonuse of the internet	Orientation, memory, and executive function tests	• OR 1.92 (1.70-2.18)
Wang et al [18]	2024	United States	6	13,624	75.30 (7.36)	15,857 (41.04)	Nonuse of the internet	Orientation, memory, and executive function tests	• OR 2.48 (2.28-2.71)
Wang et al [18]	2024	Mexican	3	10,470	71.06 (7.59)	7546 (43.41)	Nonuse of the internet	Orientation, memory, and executive function tests	• OR 1.92 (1.74-2.12)
Wang et al [18]	2024	Europe	4	23,560	70.52 (7.56)	10,757 (45.53)	Use the internet less than once a week in the last 7 days	Orientation, memory, and executive function tests	• OR 2.60 (2.34-2.88)
Berner et al [41]	2019	Sweden	6	2872	66 - 96	1190 (41.4)	Nonuse of the internet	MMSE (overall cognitive function)	• OR 1.923 (1.25-2.941)
Berner et al [41]	2019	Netherlands	6	683	66 - 94	282 (41.3)	Nonuse of the internet	MMSE (overall cognitive function)	• OR 1.639 (0.763-3.571)
Cho et al [42]	2023	United States	17.1 (median 7.9)	18,154	55.17 [53.17 - 57.25]	7758 (47.36)	Nonuse of the internet	Modified telephone interview for cognitive status (Incident dementia)	• HR 1.852 (1.389-2.439)
Williams et al [43]	2020	England	10	3937	61.7(7.9)		Nonuse of the internet		

XSL•FO
RenderX

Study	Year	Country	Follow-up (year)	Sample size	Age (years), mean (SD)	Sex (male), n (%)	Digital exclusion definition	Measurements on outcome (outcome)	OR[a]/RR[b]/HR[c] (95%CI)
						1709 (43.4)		Modified telephone interview for cognitive status (Dementia or cognitive impairment)	• Wave 2 RR 1.351 (1.124-1.639) • Wave 3 RR 1.282 (1.064-1.538) • Wave 4 RR 1.389 (1.163-1.667) • Wave 5 RR 1.563 (1.299-1.852) • Wave 6 RR 1.515 (1.266-1.786)
Quialheiro et al [47]	2021	Brazil	10	594	≥60	482 (38.6)	Nonuse of the internet	MMSE (overall cognitive function)	• RR 3.333 (1.639-6.667)
Almeida et al [44]	2012	Australia	6	5506	75.5 (4.2)	5506 (100)	Nonuse of computers	Western Australian Data Linkage System (dementia)	• HR 1.613 (1.235-2.128)
d'Orsi et al [45]	2017	England	10	8238	≥50	3713 (45.07)	Nonuse of the internet	short-form IQCODE[i] questionnaire (dementia)	• HR 1.667 (1.176-2.381)
Nakagomi et al [46]	2021	Japan	3	4232	≥65	1575 (37.2)	Nonuse of the internet or email	Public long-term care insurance (dementia)	

XSL•FO
RenderX

Study	Year	Country	Follow-up (year)	Sample size	Age (years), mean (SD)	Sex (male), n (%)	Digital exclusion definition	Measurements on outcome (outcome)	OR[a]/RR[b]/HR[c] (95%CI)
									• Use a few times a month OR 1.190 (0.752-1.887) • Use a few times a week OR 1.450 (0.826-2.564) • Use almost every day OR 1.176 (0.758-1.852)

[a]OR: odds ratio.

[b]RR: risk ratio.

[c]HR: hazard ratio.

[d]BADL: basic activity of daily living.

[e]IADL: instrumental activity of daily living.

[f]IRR: incidence rate ratio.

[g]NR: not reported.

[h]MMSE: Mini-Mental State Examination.

[i]IQCODE: Informant Questionnaire on Cognitive Decline in the Elderly.

XSL•FO
RenderX

Table . Characteristics of the included cross-sectional studies.

Study	Year	Country	Sample size	Age (years), mean (SD)	Sex (male), n (%)	Digital exclusion definition	Measurements on outcome	OR[a]/RR[b]/HR[c] (95% CI)
Physical function								
Wen et al [24]	2023	China	13,474	61.50 (9.30)	6526 (48.43)	Nonuse of the internet	ADLs[d] (functional ability)	• OR 2.083 (1.667-2.564)
Medeiros et al [19]	2012	Brazil	1656	≥60	598 (36.1)	Unable to send or receive online messages	ADLs (functional ability)	• RR 1.639 (1.064-2.50)
García-Vigara et al [50]	2022	Spain	409	67.45 (7.81)	0	Nonuse of information and communication technology	Fried frailty phenotype (physical frailty)	• OR 10.62 (5.34-21.10)
Cognitive function								
Liu et al [49]	2023	China	10,325	60.32 (9.06)	5698 (55.19)	Nonuse of mobile payment	MMSE[e] (overall cognitive function)	• OR 3.48 (2.27-5.33)
Li et al [48]	2022	China	3020	≥60	1378 (45.6)	Nonuse of the internet	MMSE (overall cognitive function)	• MCI OR 2.092 (1.153-3.788) • Dementia OR 2.545 (0.601-10.753)
Physical function and cognitive function								
Liu et al [31]	2023	China	5868	68.07 (6.1)	3083 (52.54)	Nonuse of the internet in the last month	BADLs[f] and Brief Community Screening Instrument for Dementia (functional disability and dementia)	• Physical function OR 0.671 (0.357-1.260) • Cognitive function OR 19.231 (3.846-90.909)

[a]OR: odds ratio.

[b]RR: risk ratio.

[c]HR: hazard ratio.

[d]ADL: activity of daily living.

[e]MMSE: Mini-Mental State Examination.

[f]BADL: basic activity of daily living.

The assessment tools for digital exclusion showed diversity among the included studies. None of the studies used a validated scale; instead, they all used single-item measures to assess digital exclusion. Most studies (n=14) [17,18,24,30,31,39-43,45-48] defined digital exclusion as nonuse of the internet, while other studies defined digital exclusion as nonuse of computers (n=2) [32,44], nonuse of information and communication technology (n=1) [50], nonuse of mobile payment (n=1) [49], and inability to send or receive online messages (n=1) [19].

Seven studies focused exclusively on physical function [17,19,24,30,32,39,50], while 11 examined cognitive function [18,40-49]. Only 1 study assessed both physical and cognitive outcomes [31]. Physical function was assessed by ADLs (n=3) [19,24,30], Frailty index (n=1) [17], and Fried frailty phenotype (n=2) [32,50], with 1 study obtaining physical function outcomes

XSL•FO
RenderX

by linking to data from local government systems [39]. The cognitive function was assessed using numerous cognitive tests, including MMSE (n=5) [40,41,47-49], orientation, memory, and executive function tests (n=1) [18], Modified Telephone Interview for Cognitive Status (n=2) [42,43], short-form Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) questionnaire (n=1) [45], and Brief Community Screening Instrument for Dementia (n=1) [31], with some studies obtaining cognitive function outcomes by linking to data from local government systems.

Risk of Bias Assessment

The overall risk of bias assessment of cohort and cross-sectional studies is shown in Multimedia Appendices 2 and 3.

The overall risk of bias score of cohort studies had a low risk of bias, varying from 7 to 9 points. Regarding the selection bias, all studies included the selection of a nonexposed cohort and the ascertainment of exposure. Similarly, in the comparability and outcome domains, all studies ensured the comparability of cohorts based on design or analysis, assessment of outcomes, and an adequate follow-up period for outcomes to occur. However, 6 studies failed to demonstrate that the outcome of interest was not present at the start of the study [17,18,30,41,43,46], and 4 studies had inadequate cohort follow-up [18,40,41,47].

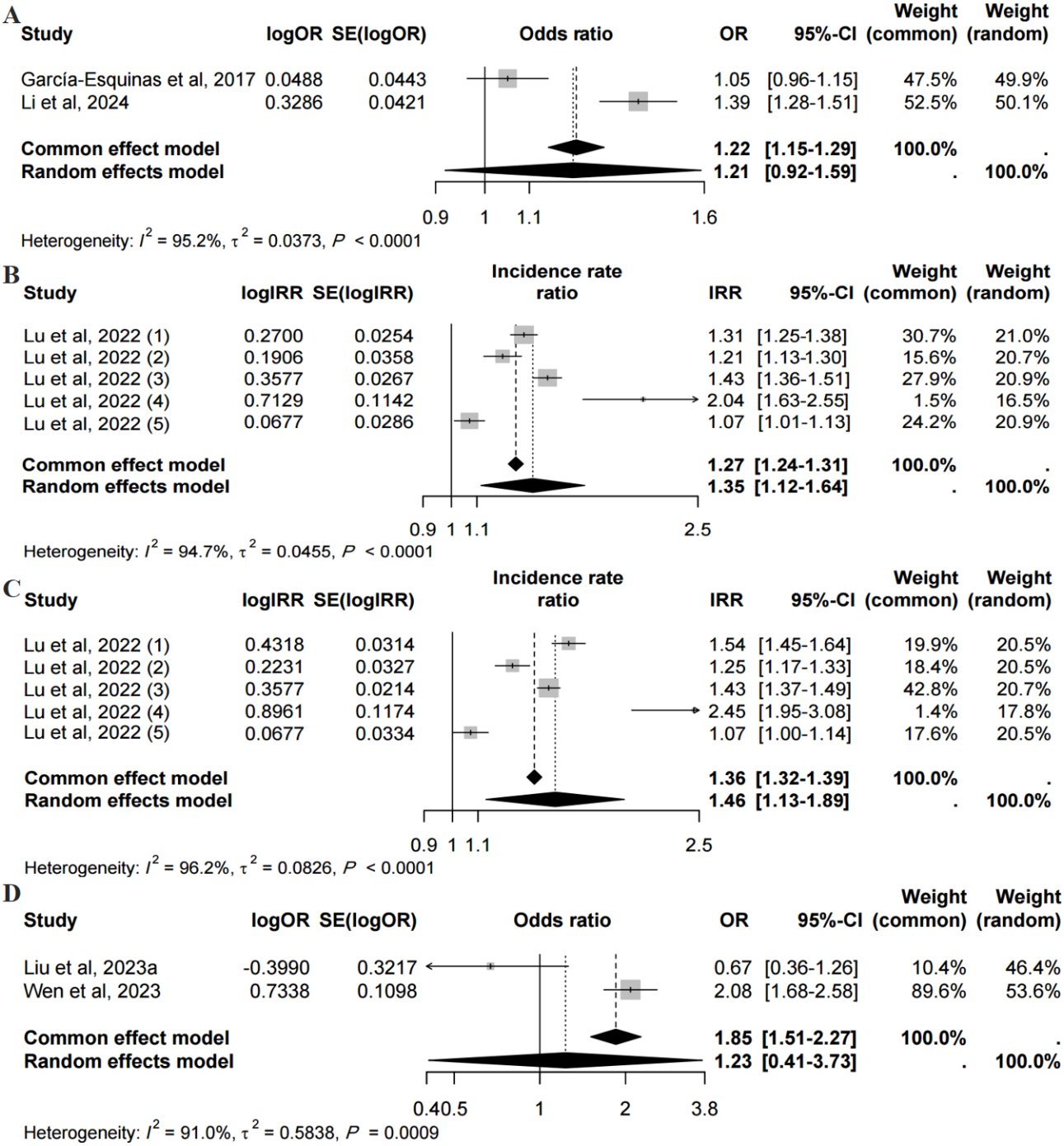
The overall risk of bias score of cross-sectional studies ranged from 3 to 7 points. All studies clearly stated the study population, had a sample size of more than 300, and assessed the outcome. However, none of the studies reported whether the studied sample had a participation rate of at least 70% (7/10) of the invited individuals.

Association Between Digital Exclusion and Physical Function

Four studies investigated the prospective association between digital exclusion and frailty, basic activities of daily living (BADLs), instrumental activities of daily living (IADLs), and incident disability [17,30,32,39]. García-Esquinas et al [32] and Li et al [17] investigated the OR value of the association between digital exclusion and frailty. The meta-analysis indicated that there was no statistically significant prospective association between them (OR 1.21, 95% CI 0.92 - 1.59, I^2=95.2%; Figure 2A). Lu et al [30] investigated the incidence rate ratio (IRR) value of the association between digital exclusion and BADLs or IADLs. The pooled associations indicated that digital exclusion was associated with decreased BADLs (IRR=1.35, 95% CI 1.12 - 1.64, I^2=94.7%; Figure 2B) and IADLs (IRR=1.46, 95% CI 1.13 - 1.89, I^2=96.2%; Figure 2C).

XSL•FO
RenderX

Figure 2. Forest plots of prospective and cross-sectional associations between digital exclusion and physical function [17,24,30-32]. (A) The odds ratio value of the prospective association between digital exclusion and frailty. (B) The incidence rate ratio value of the prospective association between digital exclusion and basic activities of daily living. (C) The incidence rate ratio value of the prospective association between digital exclusion and instrumental activities of daily living. (D) The odds ratio value of the cross-sectional association between digital exclusion and activities of daily living. IRR: incidence rate ratio; OR: odds ratio; RR: risk ratio.



Regarding the cross-sectional association, Liu et al [31] and Wen et al [24] analyzed the association between digital exclusion and ADLs. Meta-analysis indicated that there was no significant pooled association between digital exclusion and ADLs (OR 1.23, 95% CI 0.41 - 3.73, I^2=91%; Figure 2D). Tomioka et al [39] found an association between internet use and a lower risk of incident disability. While García-Vigara et al [50] revealed that digital exclusion was associated with a higher risk of frailty.

The pooled analysis indicated that digital exclusion had prospective associations with decreased BADLs and IADLs, no statistical significance in the prospective association with frailty, or cross-sectional association with ADLs.

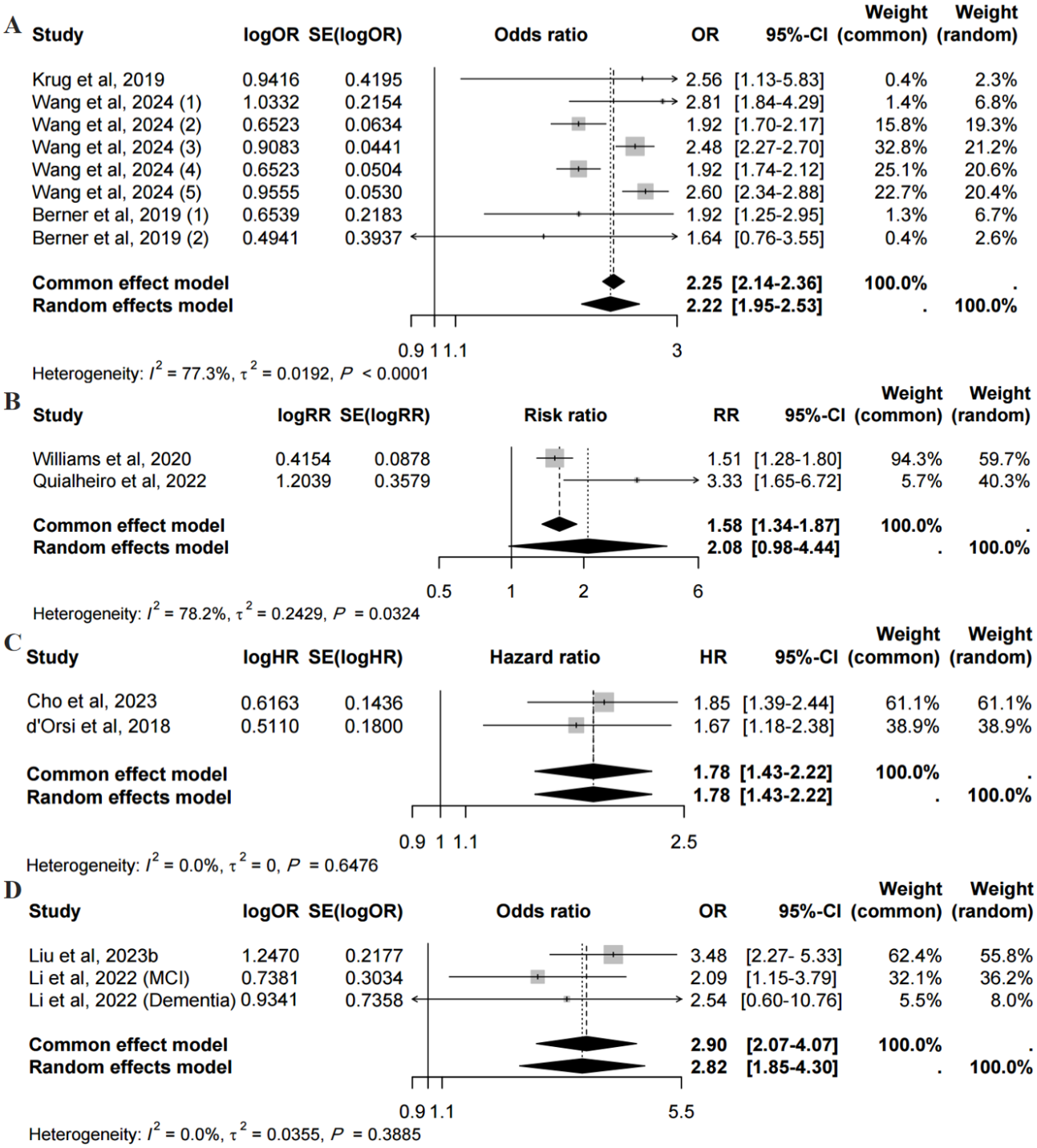
Association Between Digital Exclusion and Cognitive Function

Concerning the cohort studies, Krug et al [40], Berner et al [41], and Wang et al [18] explored the OR value of the association between digital exclusion and decreased cognitive function. Meta-analysis of these cohort studies demonstrated that digital

XSL•FO
RenderX

exclusion was significantly associated with higher odds of decreased cognitive function (OR 2.22, 95% CI 1.95 - 2.53, I^2=77.3%; Figure 3A). Williams et al [43] and Quialheiro et al [47] investigated the relationship between digital exclusion and cognitive impairment. However, meta-analysis of 2 studies [43,47] found no statistically significant association (RR 2.08, 95% CI 0.98 - 4.44, I^2=78.2%; Figure 3B). Cho et al [42] and d'Orsi et al [45] analyzed the HR value of the association between digital exclusion and dementia. Meta-analysis demonstrated a significant association between them (HR 1.78, 95% CI 1.43 - 2.22, I^2=0%; Figure 3C).

Figure 3. Forest plots of cohort and cross-sectional studies on digital exclusion and cognitive function [18,40-43,45,47-49]. (A) The odds ratio value of the prospective association between digital exclusion and cognitive function. (B) The risk ratio value of the prospective association between digital exclusion and cognitive impairment. (C) The hazard ratio value of the prospective association between digital exclusion and dementia. (D) The odds ratio value of the cross-sectional association between digital exclusion and Mini-Mental State Examination scores. HR: hazard ratio; OR: odds ratio; RR: risk ratio.



In terms of the cross-sectional studies, a meta-analysis of 2 studies [48,49] showed that digital exclusion was associated with a decline in MMSE scores (OR 2.90, 95% CI 2.07 - 4.07, I^2=0%; Figure 3D).

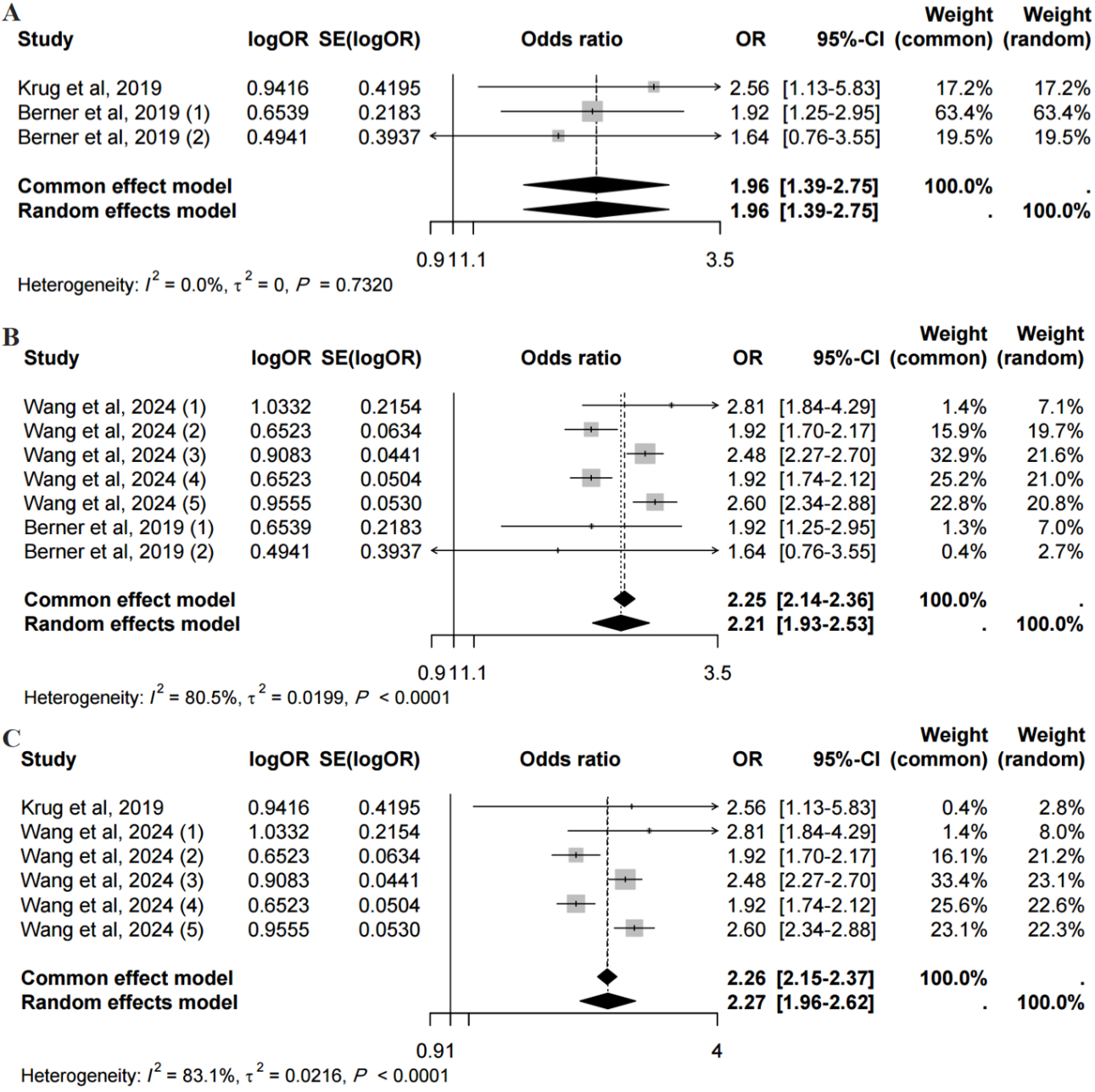
XSL•FO
RenderX

There were prospective associations between digital exclusion and dementia and decreased cognitive function, as well as cross-sectional associations with MMSE scores, but no statistically significant prospective association with cognitive impairment.

Sensitivity Analysis

Due to the limited number of studies, sensitivity analysis for the association between digital exclusion and physical function was not performed. In the sensitivity analysis of the prospective association between digital exclusion and cognitive function, a leave-one-out approach was applied, and only minor changes were observed in the results, indicating the robustness of our findings (Figure 4).

Figure 4. Forest plots of cohort studies on digital exclusion and cognitive functioning [18,40,41]. (A) The odds ratio value of the prospective association between digital exclusion and cognitive function, including studies of Krug et al [40] and Berner et al [41]. (B) The odds ratio value of the prospective association between digital exclusion and cognitive function, including studies of Wang et al [18] and Berner et al [41]. (C) The odds ratio value of the prospective association between digital exclusion and cognitive function, including studies of Krug et al [40] and Wang et al [18]. OR: odds ratio.



The sensitivity analysis revealed only minor variations in the results, confirming the stability of our findings, as evidenced by the prospective association between digital exclusion and cognitive function.

Discussion

Summary of Evidence

To the best of our knowledge, this study presents the first systematic review and meta-analysis to comprehensively

XSL•FO
RenderX

evaluate the associations of digital exclusion on both physical and cognitive functions across multiple countries. Synthesizing evidence from 13 cohorts [17,18,30,32,39-47] and 6 cross-sectional studies [19,24,31,48-50], we found that digital exclusion was prospectively associated with declines in IADLs and BADLs, as well as increased risks of incident disability, cognitive impairment, and dementia, and was cross-sectionally related to lower MMSE scores. However, no statistically significant prospective association was observed between digital exclusion and frailty.

Digital Exclusion and Physical Function

Consistent with previous studies [30,39], the meta-analysis identified a negative prospective association between digital exclusion and physical function. Several underlying explanations may account for the associations observed in this study. First, digital exclusion was associated with a minimal chance of using eHealth [57] and limited access to acquire digital preventive and therapeutic health services [58,59]. Second, it might cause a decrease in physical exercise behavior among middle-aged and older adults [60], which can lead to a reduction in skeletal muscle mass, strength, and physical mobility [61], ultimately accelerating the deterioration of physical functions and creating a vicious cycle [62-64]. Additionally, middle-aged and older adults experiencing digital exclusion were prone to dopamine imbalance and induced negative emotions such as loneliness and social isolation [65,66], increasing the risk of incident disability [67].

The associations of digital exclusion on frailty remained debated, and the meta-analysis found no statistically significant prospective association. Frailty is influenced by numerous factors such as age, gender, physiological changes, and other factors [68-70]. Furthermore, the influence of digital exclusion on frailty is likely indirect. It may operate through social isolation, limited access to health information, and reduced physical or cognitive engagement, which accumulate gradually over time [32]. These factors could mask a true association in short-term analyses. In addition, the limited number of studies included in the meta-analysis provided only a modest evidence base. The association between digital exclusion and frailty requires further validation through standardized longitudinal studies with longer follow-up periods and consistent assessments of both digital exclusion and frailty.

Meanwhile, digital exclusion did not demonstrate a cross-sectional association with physical function. One explanation for this inconsistency was the influence of confounding factors. It was also possible that socioeconomic and health-related factors masked the cross-sectional relationship between digital exclusion and physical function. Wen et al [24] reported a cross-sectional association between digital exclusion and physical function without adjusting for confounders [71], whereas Liu et al [31] found no such association after adjustment. In addition, the cross-sectional design was not well-suited to detect the slower impact of digital exclusion on physical function, which emerged gradually through changes in health behaviors and disease management. These findings underscore the need for longitudinal designs to clarify temporal relationships and potential causal pathways.

Digital Exclusion and Cognitive Function

Findings of this study regarding the relationship between digital exclusion and cognitive function broadly aligned with previous research [72-75]. Apart from the aforementioned impact of health care resources and negative psychological emotions, digital exclusion may influence cognitive function through several mechanisms. It not only reduces engagement in cognitively stimulating activities that build cognitive reserve but may also promote negative emotions, which can contribute to neurovascular dysfunction [73,76-78]. Additionally, it is linked to structural brain changes such as reduced volume of the globus pallidus [48].

However, it was noteworthy that no statistical prospective association was observed between digital exclusion and cognitive impairment. This discrepancy could be attributed to several methodological variations across studies. For instance, the studies by Williams et al [43] and Quialheiro et al [47] both investigated populations in the United Kingdom with similar follow-up durations. But the latter had a much smaller sample size (n=594 vs 3937), which could have limited its statistical power and affected the ability to detect significant associations. Furthermore, the 2 studies [43,47] used different assessment tools, which might contribute to the considerable heterogeneity in the meta-analysis.

Dose-Response Relationship and Types of Digital Exclusion

Digital exclusion or excessive internet use both had negative impacts on health outcomes [79-81]. These findings revealed “two sides of the same coin” and emphasized the importance of the frequency of internet use. In terms of the dose-response relationships, previous studies have found an inverted U-shaped association between internet use and cognitive function [42]. In the large-scale prospective cohort study of Cho et al [42], using the internet almost weekly was most beneficial to cognitive function. Additionally, Williams et al [43] and Liao et al [82] have indicated that different types of internet use exert varying impacts on physical and cognitive functions. Previous research showed diverse internet activities benefit cognition more than single ones [83]. However, the optimal types, combinations, and intensity of digital engagement for enhancing physical and cognitive functions in middle-aged and older adults remain unclear. Future research should investigate these factors across different age groups and geographic regions to tailor digital interventions to the needs and capacities of diverse populations.

Different Definitions and Assessment Tools of Digital Exclusion

The prevalence of digital exclusion varies substantially across studies due to differences in national populations, as well as inconsistencies in the definition and assessment tools of digital exclusion among middle-aged and older adults. Some studies had suggested that digital exclusion constituted a part of social exclusion [30], as digital technology marginalized older adults from the digital world [84-86]. It might be associated with a perceived lack of knowledge, inability to access necessary technology and services, and barriers posed by mental health

XSL•FO
RenderX

difficulties [85]. The absence of a unified definition and assessment tool not only reduced the reliability and generalizability of our findings due to high heterogeneity but also hindered the development of effective policies and interventions to address digital exclusion. Future research should focus on establishing a consensus-based conceptual framework and developing rigorously validated assessment instruments of digital exclusion. To ensure global applicability, these tools should be culturally and contextually adaptable, with careful attention to cross-cultural validation that ensures their relevance across diverse populations and settings.

Potential of Assessments Based on Digital Biomarkers

Current assessments of physical and cognitive function in the included studies rely largely on scales that capture activities of daily living, memory, orientation, and related abilities [18,30]. While these tools provide valuable information on functional status, they are largely subjective and may be influenced by self-report bias, assessor variability, or ceiling and floor effects [30,87]. Moreover, evidence directly linking these functional scales to objective biomarkers remains limited. Gotti et al [88] found that inflammatory and immune biomarkers changed significantly after internet-based psychological interventions, demonstrating that digital engagement can trigger measurable biological changes. This evidence presented a promising avenue for exploring underlying health mechanisms through objective biological markers. Advances in digital health further extended this possibility [89]. Chan et al [90] developed Watch Walk wrist-worn algorithms that accurately quantify gait speed [90]. This precision allowed large-scale use and turned routine movements into scalable digital biomarkers. Fan et al [70] demonstrated that integrating wearable gait data with machine learning models can substantially improve frailty prediction [91]. Together, these developments illustrate an emerging capacity to collect continuous or longitudinal data in real-world settings, enabling earlier detection of functional decline and a more nuanced understanding of how digital behaviors may shape health trajectories over time.

Mitigating Digital Exclusion for Middle-Aged and Older Adults

Although growing evidence supports the use of digital technologies to improve physical and cognitive function [92,93], reconciling digital exclusion and digital adoption remains a major challenge. To bridge this divide, it is essential to implement strategies that enhance digital access, strengthen digital literacy, and support digital assimilation [94]. Facilitating the introduction and adoption of user-friendly technologies, particularly those tailored to the needs of older adults, can play a critical role in ensuring equitable digital participation. A growing body of research highlighted co-design as a key approach [95-97]. Chan et al [98] used co-design workshops to involve older adults, health care providers, and community staff in refining a mobile health prototype, ensuring better usability and engagement. Digital navigators can offer literacy training, resolve technical issues, and guide app use, thereby improving technology adoption and integration into care [99]. Beyond design, social support also plays a crucial role. Hänninen et al [100] found that younger family members often act as “warm experts,” offering practical, personalized support that helps older adults manage digital tools with greater confidence.

Limitations and Implications

First, the generalizability of our findings is subject to certain constraints. Our meta-analysis was primarily restricted to studies published in English and Chinese, as these were the languages most accessible to the research team. Under the constraints of time and resources, conducting translations of studies in other languages was not feasible. This limitation may hinder the applicability of the results to regions where other languages are dominant or where cultural and socioeconomic contexts differ significantly. Additionally, the overall number of available studies focusing specifically on digital exclusion remains limited, particularly for middle-aged populations. Most included studies predominantly involved older adults, resulting in inadequate representation of middle-aged cohorts. Due to the limited number of included studies, it is hard to follow different age, sex, education level, income, or rural-urban residence groups to conduct subgroup analyses. This restricts the broader generalization of our conclusions across middle-aged and older adults. Future research should aim to incorporate more diverse geographical and linguistic sources, as well as prioritize well-defined studies involving middle-aged populations, to enhance the generalizability and age-specific understanding of digital exclusion.

Second, our meta-analysis is influenced by the substantial heterogeneity observed in some pooled estimates. This heterogeneity mainly arose from clinical diversity, including variations in population characteristics across different countries and income levels. Methodological diversity also contributed to heterogeneity, including differences in how digital exclusion was assessed across the large cohort studies, restriction to risk-based effect measures, variations in sample sizes, and inconsistent adjustment for covariates. By restricting quantitative synthesis to risk-based effect measures, our findings primarily reflect associations framed within an epidemiological risk perspective and may not fully capture evidence derived from correlation-based analyses using continuous measures. We performed sensitivity analyses to test the robustness of the findings and to identify key sources of heterogeneity. However, the limited number of studies in certain subgroups restricted more detailed investigations. If enough studies are available, conducting a subgroup analysis based on different definitions is informative. In the future, it is necessary to unify the definition of digital exclusion and develop recognized tools for assessment.

Third, the impact of the types and frequency of digital exclusion on physical and cognitive functions could not be explored. Future in-depth research could explore the dose-response relationships of various frequencies and types of digital exclusion on physical and cognitive functions.

Fourth, the exclusion of gray literature might lead to publication bias. It is advisable to search specialized gray literature databases when time, resources, and evidence permit.

Conclusions

Our findings emphasize the negative associations between digital exclusion and physical or cognitive functions. Although our

XSL•FO
RenderX

results do not directly address optimal patterns of digital engagement, it is noted that moderate and diverse digital use may be more beneficial than either nonuse or overuse for maintaining these functions. In clinical practice, integrating digital tools into interventions can support physical and cognitive health when tailored to users' abilities and needs. For policymakers, efforts should focus on expanding internet access, reducing financial barriers, and improving digital literacy through training and awareness programs. For future studies, the need for cautious interpretation and further validation are needed due to variations across studies, assessment tools, and populations.

Acknowledgments

During the preparation of this manuscript, no generative artificial intelligence tools were used at any stage. The authors assume full responsibility for the accuracy, originality, and integrity of all content.

Funding

This work is supported by the National Key Research & Development Program of China (2022YFC3601602), Project of the Zhuzhou Municipal Social Sciences Achievements Planning and Evaluation Committee, ZZSK2024085 the Hunan Provincial Natural Science Foundation (2023JJ30480), Fundamental Research Funds for the Central Universities of Central South University (grant number 2025ZZTS0171), the Hunan Provincial Innovation Foundation for Postgraduate (Project No. CX20250407), Central South University Research Programme of Advanced Interdisciplinary Studies (grant number 2023QYJC034).

Authors' Contributions

JY contributed to writing – original draft, methodology, formal analysis, data curation, and conceptualization. JH contributed to writing – original draft, methodology, formal analysis, and data curation. SZ contributed to data curation. XW contributed to data curation. QX contributed to data curation. JY contributed to writing – review & editing, supervision, and funding acquisition. HF contributed to writing – review & editing, supervision, and funding acquisition. All authors read and approved the final manuscript. HF and Juan Yang contributed equally to this work and share the co-corresponding authorship.

Conflicts of Interest

None declared.

Checklist 1
PRISMA checklist.
[DOCX File, 27 KB - aging_v9i1e75920_app1.docx]

Multimedia Appendix 1
Search strategy.
[DOCX File, 19 KB - aging_v9i1e75920_app2.docx]

Multimedia Appendix 2
The overall risk of bias assessment of cohort studies.
[DOCX File, 17 KB - aging_v9i1e75920_app3.docx]

Multimedia Appendix 3
The overall risk of bias assessment of cross-sectional studies.
[DOCX File, 15 KB - aging_v9i1e75920_app4.docx]

Multimedia Appendix 4
The detailed information of the included studies.
[DOCX File, 27 KB - aging_v9i1e75920_app5.docx]

References

1. Wang LY, Hu ZY, Chen HX, et al. Systematic review of longitudinal studies on daily health behavior and activity of daily living among older adults. Front Public Health 2025;13:1419279. [doi: 10.3389/fpubh.2025.1419279] [Medline: 40046125]
2. Leite MM, Silva ADO, Funghetto SS, et al. Home-based exercise and fall prevention in older adults: development, validation and usability of the Mais Equilíbrio mobile app. JMIR Aging 2025 Nov 4;8:e80724. [doi: 10.2196/80724] [Medline: 41187283]

XSL•FO
RenderX

3. Ho V, Merchant RA. The acceptability of digital technology and tele-exercise in the age of COVID-19: cross-sectional study. JMIR Aging 2022 Apr 13;5(2):e33165. [doi: 10.2196/33165] [Medline: 35294921]
4. Jeon S, Charles ST. Internet-based social activities and cognitive functioning 2 years later among middle-aged and older adults: prospective cohort study. JMIR Aging 2024 Dec 10;7:e63907. [doi: 10.2196/63907] [Medline: 39656544]
5. Graham SA, Lee EE, Jeste DV, et al. Artificial intelligence approaches to predicting and detecting cognitive decline in older adults: a conceptual review. Psychiatry Res 2020 Feb;284:112732. [doi: 10.1016/j.psychres.2019.112732] [Medline: 31978628]
6. Beard JR, Officer A, de Carvalho IA, et al. The World report on ageing and health: a policy framework for healthy ageing. Lancet 2016 May;387(10033):2145-2154. [doi: 10.1016/S0140-6736(15)00516-4]
7. Zhou Y, Li LD. Exercise training for cognitive and physical function in patients with mild cognitive impairment: a PRISMA-compliant systematic review and meta-analysis. Medicine (Baltimore) 2022;101(34):e30168. [doi: 10.1097/MD.0000000000030168]
8. Coelho-Junior HJ, Calvani R, Picca A, et al. Association between dietary habits and physical function in Brazilian and Italian older women. Nutrients 2020 Jun 1;12(6):1635. [doi: 10.3390/nu12061635] [Medline: 32492947]
9. Jing Y, Wang W, Peng W, et al. Social support and cognitive function in Chinese older adults who experienced depressive symptoms: is there an age difference? Front Aging Neurosci 2023;15:1175252. [doi: 10.3389/fnagi.2023.1175252] [Medline: 37251805]
10. Hameed M, Sani N, Sharif A, et al. Ecological factors influencing decision-making among middle-aged and older adults: a neurocognitive approach. Cureus 2025 Apr;17(4):e82598. [doi: 10.7759/cureus.82598] [Medline: 40395264]
11. Ritchie K, Carrière I, Su L, et al. The midlife cognitive profiles of adults at high risk of late-onset Alzheimer's disease: the PREVENT study. Alzheimers Dement 2017 Oct;13(10):1089-1097. [doi: 10.1016/j.jalz.2017.02.008] [Medline: 28365321]
12. Gates NJ, Rutjes AW, Di Nisio M, et al. Computerised cognitive training for maintaining cognitive function in cognitively healthy people in midlife. Cochrane Database Syst Rev 2019 Mar 13;3(3):CD012278. [doi: 10.1002/14651858.CD012278.pub2] [Medline: 30864746]
13. Alley SJ, Waters KM, Parker F, et al. The effectiveness of digital physical activity interventions in older adults: a systematic umbrella review and meta-meta-analysis. Int J Behav Nutr Phys Act 2024 Dec 18;21(1):144. [doi: 10.1186/s12966-024-01694-4] [Medline: 39696583]
14. Guan Z, Li H, Liu R, et al. Artificial intelligence in diabetes management: advancements, opportunities, and challenges. Cell Rep Med 2023 Oct 17;4(10):101213. [doi: 10.1016/j.xcrm.2023.101213] [Medline: 37788667]
15. Li W, Gui J, Luo X, Yang J, Zhang T, Tang Q. Determinants of intention with remote health management service among urban older adults: a unified theory of acceptance and use of technology perspective. Front Public Health 2023;11:1117518. [doi: 10.3389/fpubh.2023.1117518]
16. Wu M, Xue Y, Ma C. The association between the digital divide and health inequalities among older adults in China: nationally representative cross-sectional survey. J Med Internet Res 2025 Jan 15;27:e62645. [doi: 10.2196/62645] [Medline: 39813666]
17. Li L. Internet use and frailty in middle-aged and older adults: findings from developed and developing countries. Global Health 2024 Jul 3;20(1):53. [doi: 10.1186/s12992-024-01056-6] [Medline: 38961440]
18. Wang Y, Wu Z, Duan L, et al. Digital exclusion and cognitive impairment in older people: findings from five longitudinal studies. BMC Geriatr 2024 May 7;24(1):406. [doi: 10.1186/s12877-024-05026-w]
19. Medeiros FDL, Xavier AJ, Schneider IJC, Ramos LR, Sigulem D, d'Orsi E. Digital inclusion and functional capacity of older adults living in Florianópolis, Santa Catarina, Brazil (EpiFloripa 2009-2010). Rev Bras Epidemiol 2012 Mar;15(1):106-122. [doi: 10.1590/s1415-790x2012000100010] [Medline: 22450497]
20. Seifert A. The digital exclusion of older adults during the COVID-19 pandemic. J Gerontol Soc Work 2020;63(6-7):674-676. [doi: 10.1080/01634372.2020.1764687] [Medline: 32401181]
21. Zhong S, Wang Y. Digital exclusion and loneliness in older people: panel data analysis of three longitudinal cohort studies. BMC Geriatr 2025 Aug 26;25(1):662. [doi: 10.1186/s12877-025-06337-2] [Medline: 40859198]
22. Lu X, Yao Y, Jin Y. Corrigendum to "Digital exclusion and functional dependence in older people: findings from five longitudinal cohort studies" eClinical Medicine 54 (2022) 101708. EClinicalMedicine 2023 Oct;64:102239. [doi: 10.1016/j.eclinm.2023.102239] [Medline: 37799612]
23. Lu W. The impact of internet usage on the mental health of older adults in China: empirical analysis from a multi-dimensional perspective. Front Public Health 2024;12:1460558. [doi: 10.3389/fpubh.2024.1460558] [Medline: 39346596]
24. Wen W, Zhang Y, Shi W, Li J. Association between internet use and physical health, mental health, and subjective health in middle-aged and older adults: nationally representative cross-sectional survey in China. J Med Internet Res 2023 Mar 21;25:e40956. [doi: 10.2196/40956] [Medline: 36943368]
25. Cotten SR, Ford G, Ford S, Hale TM. Internet use and depression among retired older adults in the United States: a longitudinal analysis. J Gerontol B Psychol Sci Soc Sci 2014 Sep;69(5):763-771. [doi: 10.1093/geronb/gbu018] [Medline: 24671896]
26. Ge H, Li J, Hu H, Feng T, Wu X. Digital exclusion in older adults: a scoping review. Int J Nurs Stud 2025 Aug;168:105082. [doi: 10.1016/j.ijnurstu.2025.105082] [Medline: 40279791]

XSL•FO
RenderX

27. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. Inquiry 2022;59:469580221096272. [doi: 10.1177/00469580221096272] [Medline: 35471138]
28. Wiwatkunupakarn N, Pateekhum C, Aramrat C, Jirapornchaoren W, Pinyopornpanish K, Angkurawaranon C. Social networking site usage: a systematic review of its relationship with social isolation, loneliness, and depression among older adults. Aging Ment Health 2022 Jul 3;26(7):1318-1326. [doi: 10.1080/13607863.2021.1966745]
29. Latikka R, Rubio-Hernández R, Lohan ES, et al. Older adults' loneliness, social isolation, and physical information and communication technology in the era of ambient assisted living: a systematic literature review. J Med Internet Res 2021 Dec 30;23(12):e28022. [doi: 10.2196/28022] [Medline: 34967760]
30. Lu X, Yao Y, Jin Y. Digital exclusion and functional dependence in older people: findings from five longitudinal cohort studies. EClinicalMedicine 2022 Dec;54:101708. [doi: 10.1016/j.eclinm.2022.101708] [Medline: 36353265]
31. Liu J, Peng J, Chen M, Zhang T. Mediating and moderating effects of internet use on urban-rural disparities in health among older adults: nationally representative cross-sectional survey in China. J Med Internet Res 2023 Sep 28;25:e45343. [doi: 10.2196/45343] [Medline: 37768721]
32. García-Esquinas E, Andrade E, Martínez-Gómez D, Caballero FF, López-García E, Rodríguez-Artalejo F. Television viewing time as a risk factor for frailty and functional limitations in older adults: results from 2 European prospective cohorts. Int J Behav Nutr Phys Act 2017 Apr 26;14(1):54. [doi: 10.1186/s12966-017-0511-1] [Medline: 28446189]
33. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
34. Salzmann-Erikson M. Digital inclusion: a mixed-method study of user behavior and content on Twitter. Digit Health 2023;9:20552076231211277. [doi: 10.1177/20552076231211277] [Medline: 37928325]
35. Mannheim I, Köttl H. Ageism and (successful) digital engagement: a proposed theoretical model. Gerontologist 2024 Sep 1;64(9):gnae078. [doi: 10.1093/geront/gnae078] [Medline: 38874215]
36. Li X, Zhou X, Du H, et al. Longitudinal relationship between baseline Weight-Adjusted Waist Index and stroke risk over 8 years in Chinese adults aged 45 and older: a prospective cohort study. Front Public Health 2025;13:1505364. [doi: 10.3389/fpubh.2025.1505364]
37. Moskalewicz A, Oremus M. No clear choice between Newcastle-Ottawa Scale and Appraisal Tool for cross-sectional studies to assess methodological quality in cross-sectional studies of health-related quality of life and breast cancer. J Clin Epidemiol 2020 Apr;120:94-103. [doi: 10.1016/j.jclinepi.2019.12.013] [Medline: 31866469]
38. Ho PJ, Gernaat SAM, Hartman M, Verkooijen HM. Health-related quality of life in Asian patients with breast cancer: a systematic review. BMJ Open 2018 Apr 20;8(4):e020512. [doi: 10.1136/bmjopen-2017-020512] [Medline: 29678980]
39. Tomioka K, Shima M, Saeki K. Longitudinal association between frequency of Internet use and incident disability among community-dwelling older people during the COVID-19 pandemic. Environ Health Prev Med 2024;29:13. [doi: 10.1265/ehpm.23-00207] [Medline: 38447990]
40. Krug RR, d'Orsi E, Xavier AJ. Association between use of internet and the cognitive function in older adults, populational longitudinal study EpiFloripa Idoso. Rev Bras Epidemiol 2019 Mar 14;22:e190012. [doi: 10.1590/1980-549720190012] [Medline: 30892475]
41. Berner J, Comijs H, Elmståhl S, et al. Maintaining cognitive function with internet use: a two-country, six-year longitudinal study. Int Psychogeriatr 2019 Jul;31(7):929-936. [doi: 10.1017/S1041610219000668]
42. Cho G, Betensky RA, Chang VW. Internet usage and the prospective risk of dementia: a population-based cohort study. J Am Geriatr Soc 2023 Aug;71(8):2419-2429. [doi: 10.1111/jgs.18394] [Medline: 37132331]
43. Williams BD, Pendleton N, Chandola T. Cognitively stimulating activities and risk of probable dementia or cognitive impairment in the English longitudinal study of ageing. SSM Popul Health 2020 Dec;12:100656. [doi: 10.1016/j.ssmph.2020.100656] [Medline: 32984495]
44. Almeida OP, Yeap BB, Alfonso H, Hankey GJ, Flicker L, Norman PE. Older men who use computers have lower risk of dementia. PLoS One 2012;7(8):e44239. [doi: 10.1371/journal.pone.0044239] [Medline: 22937167]
45. d'Orsi E, Xavier AJ, Rafnsson SB, Steptoe A, Hogervorst E, Orrell M. Is use of the internet in midlife associated with lower dementia incidence? Results from the English longitudinal study of ageing. Aging Ment Health 2018 Nov 2;22(11):1525-1533. [doi: 10.1080/13607863.2017.1360840]
46. Nakagomi A, Shiba K, Kawachi I, et al. Internet use and subsequent health and well-being in older adults: an outcome-wide analysis. Comput Human Behav 2022 May;130:107156. [doi: 10.1016/j.chb.2021.107156]
47. Quialheiro A, Figueiró TH, Rech CR, et al. Can internet use reduce the incidence of cognitive impairment? Analysis of the EpiFloripa Aging Cohort Study (2009-2019). Prev Med 2022 Jan;154:106904. [doi: 10.1016/j.ypmed.2021.106904] [Medline: 34863810]
48. Li W, Yue L, Xiao S. Association between internet use, cognitive function, and globus pallidus volumes: a study among the elderly in Chinese communities. Front Public Health 2022;10:886974. [doi: 10.3389/fpubh.2022.886974]
49. Liu X, Wang X, Zhang H, Pei M, Li N. Relationship between digital exclusion and cognitive impairment in Chinese adults. Front Aging Neurosci 2023;15:1194348. [doi: 10.3389/fnagi.2023.1194348] [Medline: 37465320]

XSL•FO
RenderX

50. García-Vigara A, Cano A, Fernández-Garrido J, Carbonell-Asíns JA, Tarín JJ, Sánchez-Sánchez ML. Non-use of information and communication technology as a predictor of frailty in postmenopausal midlife and older women. Maturitas 2022 Feb;156:60-64. [doi: 10.1016/j.maturitas.2021.05.010] [Medline: 34275701]
51. Brumback BA, Bouldin ED, Zheng HW, Cannell MB, Andresen EM. Testing and estimating model-adjusted effect-measure modification using marginal structural models and complex survey data. Am J Epidemiol 2010 Nov 1;172(9):1085-1091. [doi: 10.1093/aje/kwq244] [Medline: 20801863]
52. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ. Cochrane Handbook for Systematic Reviews of Interventions Version 65: Cochrane; 2024. URL: https://www.cochrane.org/handbook [accessed 2026-04-03]
53. Chau JPC, Liu X, Lo SHS, Chien WT, Wan X. Effects of environmental cleaning bundles on reducing healthcare-associated Clostridioides difficile infection: a systematic review and meta-analysis. J Hosp Infect 2020 Dec;106(4):734-744. [doi: 10.1016/j.jhin.2020.08.019] [Medline: 32861741]
54. Dubale YM, Belachew T, Wondafrash B, Rameckers E, Verbecque E. Do physical fitness and motor skill performances in underweight children differ from normal weight peers? A meta-analysis. BMC Pediatr 2025 May 17;25(1):393. [doi: 10.1186/s12887-025-05738-x] [Medline: 40382586]
55. Zhang H, Donaldson J, Neves VCM, Scott J. Are dental magnetic resonance imaging and ultrasonography techniques reliable alternatives for treatment planning dental implants? A systematic review and meta-analysis. Int J Implant Dent 2025 Aug 11;11(1):52. [doi: 10.1186/s40729-025-00634-6] [Medline: 40788371]
56. Willems RPJ, van Dijk K, Vehreschild M, et al. Incidence of infection with multidrug-resistant Gram-negative bacteria and vancomycin-resistant enterococci in carriers: a systematic review and meta-regression analysis. Lancet Infect Dis 2023 Jun;23(6):719-731. [doi: 10.1016/S1473-3099(22)00811-8] [Medline: 36731484]
57. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
58. Clarke CS, Round J, Morris S, et al. Exploring the relationship between frequent internet use and health and social care resource use in a community-based cohort of older adults: an observational study in primary care. BMJ Open 2017 Jul 21;7(7):e015839. [doi: 10.1136/bmjopen-2017-015839] [Medline: 28733300]
59. Shim H, Ailshire J, Zelinski E, Crimmins E. The health and retirement study: analysis of associations between use of the internet for health information and use of health services at multiple time points. J Med Internet Res 2018 May 25;20(5):e200. [doi: 10.2196/jmir.8203] [Medline: 29802088]
60. Afshin A, Babalola D, Mclean M, et al. Information technology and lifestyle: a systematic evaluation of internet and mobile interventions for improving diet, physical activity, obesity, tobacco, and alcohol use. J Am Heart Assoc 2016 Aug 31;5(9):e003058. [doi: 10.1161/JAHA.115.003058] [Medline: 27581172]
61. Mcleod JC, Currier BS, Lowisz CV, Phillips SM. The influence of resistance exercise training prescription variables on skeletal muscle mass, strength, and physical function in healthy adults: an umbrella review. J Sport Health Sci 2024 Jan;13(1):47-60. [doi: 10.1016/j.jshs.2023.06.005] [Medline: 37385345]
62. de Magalhães JP. Cellular senescence in normal physiology. Science 2024 Jun 21;384(6702):1300-1301. [doi: 10.1126/science.adj7050] [Medline: 38900869]
63. Li X, Ma L. From biological aging to functional decline: insights into chronic inflammation and intrinsic capacity. Ageing Res Rev 2024 Jan;93:102175. [doi: 10.1016/j.arr.2023.102175] [Medline: 38145874]
64. Zeng P, Li M, Cao J, Zeng L, Jiang C, Lin F. Association of metabolic syndrome severity with frailty progression among Chinese middle and old-aged adults: a longitudinal study. Cardiovasc Diabetol 2024 Aug 16;23(1):302. [doi: 10.1186/s12933-024-02379-9] [Medline: 39152431]
65. Chen YRR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18. [doi: 10.2196/jmir.4596] [Medline: 26822073]
66. Li J, Zhou XC. Internet use and Chinese older adults' subjective well-being (SWB): the role of parent-child contact and relationship. Comput Human Behav 2021 Jun;119:106725. [doi: 10.1016/j.chb.2021.106725]
67. Tomata Y, Watanabe T, Tanji F, Zhang S, Sugawara Y, Tsuji I. The impact of psychological distress on incident functional disability in elderly Japanese: the Ohsaki Cohort 2006 Study. Int J Environ Res Public Health 2018 Nov 8;15(11):2502. [doi: 10.3390/ijerph15112502] [Medline: 30413102]
68. Picca A, Coelho-Junior HJ, Calvani R, Marzetti E, Vetrano DL. Biomarkers shared by frailty and sarcopenia in older adults: a systematic review and meta-analysis. Ageing Res Rev 2022 Jan;73:101530. [doi: 10.1016/j.arr.2021.101530] [Medline: 34839041]
69. Yang H, Li Z, Zhang Y, et al. Associations between frailty, genetic predisposition, and chronic kidney disease risk in middle-aged and older adults: a prospective cohort study. Maturitas 2024 Sep;187:108059. [doi: 10.1016/j.maturitas.2024.108059] [Medline: 38941958]
70. Fan J, Yu C, Guo Y, et al. Frailty index and all-cause and cause-specific mortality in Chinese adults: a prospective cohort study. Lancet Public Health 2020 Dec;5(12):e650-e660. [doi: 10.1016/S2468-2667(20)30113-4] [Medline: 33271078]
71. Valla LG, Rossi M, Gaia A, Guaita A, Rolandi E. The impact of the COVID-19 pandemic on oldest-old social capital and health and the role of digital inequalities: longitudinal cohort study. J Med Internet Res 2025 Jan 9;27:e62824. [doi: 10.2196/62824] [Medline: 39784108]

XSL•FO
RenderX

72. Jin Y, Jing M, Ma X. Effects of digital device ownership on cognitive decline in a middle-aged and elderly population: longitudinal observational study. J Med Internet Res 2019 Jul 29;21(7):e14210. [doi: 10.2196/14210] [Medline: 31359864]
73. Small GW, Lee J, Kaufman A, et al. Brain health consequences of digital technology use dialogues in clinical neuroscience. Dialogues Clin Neurosci 2020 Jun;22(2):179-187. [doi: 10.31887/DCNS.2020.22.2/gsmall] [Medline: 32699518]
74. Amini R, Chee KH, Mendieta M, Parker S. Online engagement and cognitive function among older adults. Geriatr Gerontol Int 2019 Sep;19(9):918-923. [doi: 10.1111/ggi.13749] [Medline: 31368165]
75. Xavier AJ, d'Orsi E, de Oliveira CM, et al. English longitudinal study of aging: can internet/e-mail use reduce cognitive decline? J Gerontol A Biol Sci Med Sci 2014 Sep;69(9):1117-1121. [doi: 10.1093/gerona/glu105] [Medline: 25116923]
76. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6]
77. Bhatti AB, Haq AU. The pathophysiology of perceived social isolation: effects on health and mortality. Cureus 2017 Jan 24;9(1):e994. [doi: 10.7759/cureus.994] [Medline: 28382237]
78. Schlomann A, Seifert A, Zank S, Woopen C, Rietz C. Use of Information and Communication Technology (ICT) devices among the oldest-old: loneliness, anomie, and autonomy. Innov Aging 2020;4(2):igz050. [doi: 10.1093/geroni/igz050] [Medline: 31911953]
79. Dresp-Langley B, Hutt A. Digital addiction and sleep. Int J Environ Res Public Health 2022 Jun 5;19(11):6910. [doi: 10.3390/ijerph19116910] [Medline: 35682491]
80. Bertocchi FM, De Oliveira AC, Lucchetti G, Lucchetti ALG. Smartphone use, digital addiction and physical and mental health in community-dwelling older adults: a population-based survey. J Med Syst 2022 Jun 18;46(8):53. [doi: 10.1007/s10916-022-01839-7] [Medline: 35716194]
81. Li P, Zhang C, Gao S, et al. Association between daily internet use and incidence of chronic diseases among older adults: prospective cohort study. J Med Internet Res 2023 Jul 17;25:e46298. [doi: 10.2196/46298] [Medline: 37459155]
82. Liao S, Zhou Y, Liu Y, Wang R. Variety, frequency, and type of Internet use and its association with risk of depression in middle- and older-aged Chinese: a cross-sectional study. J Affect Disord 2020 Aug 1;273:280-290. [doi: 10.1016/j.jad.2020.04.022] [Medline: 32421614]
83. Yu X, Mu A, Wu X, Zhou L. Impact of internet use on cognitive decline in middle-aged and older adults in China: longitudinal observational study. J Med Internet Res 2022 Jan 24;24(1):e25760. [doi: 10.2196/25760] [Medline: 35072642]
84. Helsper EJ. A corresponding fields model for the links between social and digital exclusion. Commun Theor 2012 Nov;22(4):403-426 [FREE Full text] [doi: 10.1111/j.1468-2885.2012.01416.x]
85. Greer B, Robotham D, Simblett S, Curtis H, Griffiths H, Wykes T. Digital exclusion among mental health service users: qualitative investigation. J Med Internet Res 2019 Jan 9;21(1):e11696. [doi: 10.2196/11696] [Medline: 30626564]
86. Helsper EJ, Reisdorf BC. The emergence of a "digital underclass" in Great Britain and Sweden: changing reasons for digital exclusion. New Media Soc 2017 Aug;19(8):1253-1270. [doi: 10.1177/1461444816634676]
87. Cao F, Vogel AP, Gharahkhani P, Renteria ME. Speech and language biomarkers for Parkinson's disease prediction, early diagnosis and progression. NPJ Parkinsons Dis 2025 Mar 24;11(1):57. [doi: 10.1038/s41531-025-00913-4] [Medline: 40128529]
88. Gotti G, Gabelli C, Russotto S, et al. Biomarkers of response to internet-based psychological interventions: systematic review. J Med Internet Res 2024 Nov 29;26:e55736. [doi: 10.2196/55736] [Medline: 39612489]
89. Huang J, Zhou S, Xie Q, Yu J, Zhao Y, Feng H. Digital biomarkers for real-life, home-based monitoring of frailty: a systematic review and meta-analysis. Age Ageing 2025 Mar 28;54(4):afaf108. [doi: 10.1093/ageing/afaf108] [Medline: 40251836]
90. Chan LLY, Choi TCM, Lord SR, Brodie MA. Development and large-scale validation of the Watch Walk wrist-worn digital gait biomarkers. Sci Rep 2022 Oct 10;12(1):16211. [doi: 10.1038/s41598-022-20327-z] [Medline: 36217013]
91. Fan S, Ye J, Xu Q, et al. Digital health technology combining wearable gait sensors and machine learning improve the accuracy in prediction of frailty. Front Public Health 2023;11:1169083. [doi: 10.3389/fpubh.2023.1169083]
92. Gao Y, Liu N. Effects of digital technology-based serious games interventions for older adults with mild cognitive impairment: a meta-analysis of randomised controlled trials. Age Ageing 2025 Mar 28;54(4):afaf080. [doi: 10.1093/ageing/afaf080] [Medline: 40192627]
93. Chen C, Huang N, Hu B, Zhang M, Yuan J, Guo J. The effectiveness of digital technology interventions for cognitive function in older adults: a systematic review and meta-analysis of randomized controlled trials. Geroscience 2025 Feb;47(1):653-683. [doi: 10.1007/s11357-024-01446-z] [Medline: 39688787]
94. Gong E, Wang H, Zhu W, et al. Bridging the digital divide to promote prevention and control of non-communicable diseases for all in China and beyond. BMJ 2024 Oct 18;387:e076768. [doi: 10.1136/bmj-2023-076768] [Medline: 39424328]
95. Connelly J, Swingler K, Rodriguez-Sanchez N, Whittaker AC. Identifying food preferences and malnutrition in older adults in care homes: co-design study of a digital nutrition assessment tool. JMIR Aging 2025 Mar 3;8:e64661. [doi: 10.2196/64661] [Medline: 40053797]
96. Hepburn J, Williams L, McCann L. Barriers to and facilitators of digital health technology adoption among older adults with chronic diseases: updated systematic review. JMIR Aging 2025 Sep 11;8:e80000. [doi: 10.2196/80000] [Medline: 40934502]

XSL•FO
RenderX

97. Timon CM, Heffernan E, Kilcullen S, et al. Developing independent living support for older adults using internet of things and AI-based systems: co-design study. JMIR Aging 2024 Oct 24;7:e54210. [doi: 10.2196/54210] [Medline: 39447196]
98. Chan KH, Oh LE, Bao H, Lee EWJ. Co-developing SingaporeWALK mobile health app: enhancing physical health and mental well-being among current and future older adults through inclusive technology. Digit Health 2025;11:20552076251390573. [doi: 10.1177/20552076251390573] [Medline: 41229926]
99. Perret S, Alon N, Carpenter-Song E, et al. Standardising the role of a digital navigator in behavioural health: a systematic review. Lancet Digit Health 2023 Dec;5(12):e925-e932. [doi: 10.1016/S2589-7500(23)00152-8] [Medline: 38000876]
100. Hänninen R, Taipale S, Luostari R. Exploring heterogeneous ICT use among older adults: the warm experts' perspective. New Media Soc 2021 Jun;23(6):1584-1601. [doi: 10.1177/1461444820917353]

Abbreviations

ADL: activity of daily living
BADL: basic activity of daily living
HR: hazard ratio
IADL: instrumental activity of daily living
IQCODE: Informant Questionnaire on Cognitive Decline in the Elderly
IRR: incidence rate ratio
MeSH: Medical Subject Headings
MMSE: Mini-Mental State Examination
NOS: Newcastle-Ottawa Scale
OR: odds ratio
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RR: risk ratio
WHO: World Health Organization

Edited by C Lee; submitted 13.Apr.2025; peer-reviewed by D Büyük, D Tang, W Zhu; revised version received 06.Mar.2026; accepted 06.Mar.2026; published 23.Apr.2026.

Please cite as:
Yu J, Huang J, Zhou S, Wan X, Xie Q, Zhao Y, Yang J, Feng H
The Associations Between Digital Exclusion and Physical or Cognitive Function in Middle-Aged and Older Adults: Systematic Review and Meta-Analysis
JMIR Aging 2026;9:e75920
URL: https://aging.jmir.org/2026/1/e75920
doi:10.2196/75920

© Jia Yu, Jundan Huang, Shuhan Zhou, Xiao Wan, Qi Xie, Yinan Zhao, Juan Yang, Hui Feng. Originally published in JMIR Aging (https://aging.jmir.org), 23.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Resources and Social Connectedness Among Ethnic Minority Older Adults: Systematic Review and Meta-Analysis

Padmore Adusei Amoah[1], PhD; Edward Kwabena Ameyaw[1], PhD; Miron Kumar Bhowmik[2], PhD; Prince Nyansah Adotey[3], BSc; Andrew Donkor[3,4], PhD

[1]School of Graduate Studies, Lingnan University, Hong Kong, 8 Castle Peak Rd, Tuen Mun, China (Hong Kong)

[2]Department of Education Policy and Leadership, The Education University of Hong Kong, Tai Po, China (Hong Kong)

[3]Department of Medical Imaging, Faculty of Allied Health Sciences, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

[4]IMPACCT (Improving Palliative, Aged and Chronic Care through Clinical Research and Translation), Faculty of Health, University of Technology Sydney, Sydney, Australia

Corresponding Author:
Padmore Adusei Amoah, PhD
School of Graduate Studies, Lingnan University, Hong Kong, 8 Castle Peak Rd, Tuen Mun, China (Hong Kong)

Abstract

Background: Being socially connected is essential for health and well-being. Nonetheless, many older adults face social isolation, especially in ethnically diverse societies. Digital technologies offer a pragmatic approach to addressing problems with social connectedness; however, a consolidated understanding of their association with social connectedness among ethnic minority older adults remains unaddressed.

Objective: We conducted a systematic review and meta-analysis to investigate the prevalence and utility of digital resources in promoting social connectedness among ethnic minority older adults.

Methods: This systematic review and meta-analysis was guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. Studies were retrieved from 4 electronic databases. A random-effects meta-analysis was used to estimate the pooled prevalence along with its 95% CI.

Results: Twelve studies were included. The research approaches used by the included studies were as follows: quantitative (n=5), qualitative (n=5), and mixed methods (n=2). Overall, the estimated prevalence of internet use, which was defined as engaging in online activities through digital devices, was 81.8%. Commonly used digital devices were computers (66.7%), smartphones (63.2%), and tablets (29.3%). Sociodemographic, economic, and health-related factors influenced the adoption and use of digital resources among ethnic minority older adults. These older adults relied on digital resources to build meaningful social connections through social participation, maintaining cultural and religious ties, promoting transnational social networks, and fostering their potential for social connectedness. However, digital resources were not always useful for social connectedness and adversely affected in-person contact.

Conclusions: Digital resources can promote social connectedness among ethnic minority older adults by enabling them to maintain their cultural and social values. Cities and regions aiming to promote the social connectedness and well-being of ethnic minority older adults should consider intersectional factors that affect access to digital resources and the sustainable adoption of digital resources.

(JMIR Aging 2026;9:e84962) doi:10.2196/84962

KEYWORDS

digital resources; ethnic minority; older adults; social connectedness

Introduction

There are approximately 5.6 billion internet users worldwide, which accounts for 67.9% of the global population, indicating the fundamental and crucial role of digital resources in the global socioeconomic and developmental milieu [1]. Digital resources encompass any electronic device, tool, system, or equipment that can connect to the internet and generate, store, and/or process data [2,3]. Digitalization and technological advancement are rapidly becoming the mainstay of contemporary communication, education, economic opportunities, and how people interact with their communities, even in low- and middle-income countries [4]. Digital engagements are increasingly underpinning most activities, and this situation also applies to ethnic minority older adults, especially given the myriad sociocultural barriers they face in accessing needed social services and maintaining healthy social connectedness [5]. Social connectedness, the primary focus of this study, is

XSL•FO
RenderX

considered the antonym of loneliness as it involves constructive relationships with others, including groups, communities, and individuals [6]. Thus, it is strongly related to social connections or networks, which represent "the structure, function, and quality of our relationships with others" [7], including family, friends, neighbors, and people within and outside our immediate and external sociophysical environments [6,8]. Such connections are fundamental to individual and social sustainability, resilience, safety, and prosperity, especially among vulnerable people [9].

Older adults, especially those in minority groups, are often vulnerable to social isolation (ie, an object deficit in adequate social connections) and loneliness (ie, psychological pain due to the gap between desired and actual social connections) due to limited opportunities to be socially connected—a situation often compounded by language barriers and economic challenges [7,10,11], which can have a negative impact on their health [12]. It is estimated that one-third of older adults experience some degree of social isolation as their social networks reduce with age [13,14]. This has been attributed to factors such as migration, death, and the age-related functional decline of their usual social connections [10]. Acculturation is the psychological, social, and cultural changes that occur when 2 cultures come in contact [15]. Many migrants, especially those of ethnic minority status, tend to struggle to effectively participate in the traditions, values, and practices of dominant cultures, making them feel a sense of exclusion [16]. A major challenge is often how people from minority backgrounds decide whether to prioritize their heritage or the host culture [16]. Such experience leads to acculturative stress and psychological and emotional strain (eg, language barriers, feelings of being an outsider, experiences of discrimination, and even racism) as people adapt to a different culture [17]. Difficulty in acculturating to a new culture means many ethnic minority people face significant challenges forming meaningful relationships with members of the majority group (ie, social connectedness), something that is likely more abundant among majority populations [18,19]. Likewise, from the perspective of the social identity theory [20], such challenging experiences due to acculturation stress leave questions about how ethnic minority people form or maintain their self-concept relative to the larger population and culture in terms of how they categorize themselves (eg, in terms of race, nationality, and social class), how they identify with the group (ie, the majority group), and how they compare themselves with majority groups and those of similar characteristics as themselves (in-group) [21]. This process of constructing and maintaining social identity can be inimical to the forms of social connections and connectedness available to ethnic minority people, as the formation and maintenance of social relationships are also part of social identity [22].

These challenges with acculturation and social identity make digital tools and resources quintessential to the successful adaptation and social connectedness of ethnic minority people, particularly older adults, as they can provide them with space to reinforce their ethnic and cultural identities and connect with others with similar sociocultural backgrounds and beliefs [3,23]. Digital resources are broadly considered to comprise devices, such as smartphones, computers, tablets, and remote monitoring devices, and the platforms they enable, such as video streaming, messaging app use (eg, WhatsApp and WeChat), online gaming, and social media use (eg, Facebook, TikTok, and Instagram), through the internet [3]. Such resources can promote social connectedness among older adults and reduce loneliness and social isolation [24,25]. For ethnic minority older adults, being in unfamiliar and racially diverse communities reduces their social and cultural capital, leading to limited opportunity for them to generate meaningful social connections, which often affects their well-being adversely [18,26,27]. Despite the benefits and widespread use of digital resources, ethnic minority populations continue to face significant barriers to adopting and utilizing these resources [28]. Identified barriers include limited access to digital technology (eg, the internet), negative attitudes and technology anxiety, data privacy and security concerns, low digital skills, low literacy skills, and unfavorable device and content design [29,30]. These factors align with the Technology Acceptance Model, which argues that factors, including perceived ease of use, perceived usefulness of a technology, attitudes, and intention toward technology use, determine how much a person is willing to accept and adopt different technologies [31,32]. Given these factors, emergent questions include how and to what extent digital tools/platforms serve as mediums for social connections among ethnic minority older adults.

To address such research gaps, some systematic literature reviews have summarized studies on the impact of digital resources on social connectedness among older adults [24,33-36]. However, these studies have mainly focused on older adult populations, leaving a significant knowledge gap regarding specific groups, which complicates efforts to develop inclusive digital strategies, particularly those focused on ethnic minority aging populations, many of whom have difficulty forming social connections [10]. Therefore, this systematic review and meta-analysis seeks to identify the prevalence and utility of digital resources for social connectedness among ethnic minority older adults. The findings from this study can pave the way for well-situated future research and a strengthened knowledge base for responsive policies aimed at promoting social connectedness among ethnic minority older adults to prevent and manage ill-consequences such as social isolation and loneliness.

Methods

Overview

This systematic review was conducted and reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [37]. The PRISMA checklist is provided in Checklist 1. It was registered with the International Prospective Register of Systematic Reviews (PROSPERO; CRD420251035114) [38].

Study Selection Criteria

Participants

This systematic review included studies in which some or all study participants were ethnic minority older adults, as defined

XSL•FO
RenderX

in various studies. An ethnic minority group is a group of people who share primarily racial and ethnic identities that differ from the dominant race in a given place. In most places, such people include immigrants (of different races/ethnicities and often different nationalities), indigenous people, people of color, and ethnic groups with characteristics such as different cultures/traditions and languages [27].

Intervention

Studies were included if they (1) explored the use of digital resources for communication or social function among ethnic minority older adults and (2) evaluated the association between the use of digital resources and outcomes related to social connectedness among ethnic minority older adults.

Studies

This systematic review and meta-analysis included primary studies of any design, as well as peer-reviewed journal articles and studies published in English. Editorials, opinion pieces, comments, letters, reviews, and any non–peer-reviewed materials were excluded.

Outcome Measures

The primary outcome of interest was the utility of digital resources for social connectedness. The secondary outcomes were the prevalence of digital resource usage and the factors that influence its use among ethnic minority older adults.

Data Source and Search Strategy

The following electronic databases were searched: PubMed, Web of Science (Core Collection), PsycINFO, and Scopus. Databases were searched between May 2, 2025, and May 4, 2025. Consistent with our consideration of digital resources as electronic devices and their platforms used for social connections, the search strategy included use of a wide range of terms related to the following concepts: electronic devices and platforms (eg, computers, smartphones, tablets, and social media), social connectedness (with related terms such as social network, social connection, social support, social capital, and social interactions), ethnic minority populations (with related terms such as people of color, minority group, human migration, racial minority, aboriginal, and others), and older adults (with related words such as elderly, seniors, and older persons). Detailed information is provided in Table S1 in Multimedia Appendix 1. The expansive search terms enabled the retrieval of comprehensive results of related research. Medical Subject Headings (MeSH), keywords, and free-text terms were combined using “AND” or “OR” Boolean operators. A hand search through Google Scholar and tracing of reference lists of relevant studies were performed to supplement the electronic database search. No geographical restrictions were applied to the search.

Study Selection

Duplicate articles were removed. Two reviewers independently screened the titles and abstracts of articles retrieved by the search to determine their eligibility. Agreement between reviewers during the title and abstract screening was almost perfect (Cohen κ=0.88). A third reviewer resolved any disagreement between the 2 reviewers. Full texts of any potentially relevant articles were obtained and assessed by 2 independent reviewers based on the defined inclusion criteria, with a Cohen κ score of 0.92. Disagreements were resolved by discussion with a third reviewer. The selection process was recorded in detail to produce a PRISMA flow diagram. EndNote (Clarivate) was used to manage the references.

Data Extraction

For studies that fulfilled the inclusion criteria, 2 reviewers independently extracted relevant data. A standard data extraction form was developed. The following data were extracted:

- Study characteristics: authors, year of publication, country, aim, study design, inclusion and exclusion criteria, sample size, gender distribution, and mean (SD) age
- Outcomes/results: type of digital resource, prevalence data, and determinants of digital resource use
- Conclusions: author discussion/conclusions, limitations, funding, and notable conflicts of interest

Study Quality Assessment

Two reviewers assessed the quality of the included studies. Any discrepancies were resolved by discussion. Interrater agreement between the reviewers was excellent, with a Cohen κ score of 0.90. Qualitative and quantitative studies were assessed by using the appropriate Joanna Briggs Institute (JBI) Critical Appraisal Checklist for qualitative research. Consistent with approaches reported in other previous systematic reviews and to facilitate comparison, each item in the appraisal checklist was rated as follows: yes (1 point), no (0 points), and unclear/not applicable (0 points) [39,40]. The sum was divided by the number of items in the appraisal checklist and multiplied by 100 to obtain a percentage. The risk of bias scores were categorized as follows: 80% or more (high quality), 50% - 79% (moderate quality), and below 50% (low quality), as suggested elsewhere [41].

Overall, most of the included studies fulfilled the relevant JBI criteria. Four quantitative studies were rated high quality [42-45]. Similarly, 4 qualitative studies were rated high quality [46-49]. Among the included studies, only 4 reported using a theoretical framework to guide their work [42-44,47].

Details of the quality assessment are provided in Tables S2 and S3 in Multimedia Appendix 1.

Data Analysis

Data were summarized statistically when available, sufficiently similar, and of sufficient quality. The included studies were characterized by the study design and the type of digital resource. Qualitative findings were synthesized thematically following the procedure described by Thomas and Harden [50]. This included (1) free line-by-line coding of the findings of primary studies (eg, identifying prevalence indicators of digital resources and different uses of digital resources for social connections such as religious, cultural, intranational, and cross-national connections), (2) organization of codes into related categories to develop descriptive themes that capture patterns across the included studies (eg, religious and cultural codes were pulled together), and (3) development of higher-order analytical themes (eg, defining and refining categories and names to represent the core idea, such as maintaining cultural and religious connections). The analysis

XSL•FO
RenderX

was mostly deductive, with a focus on addressing the research objectives guided by evidence from the literature on how digital resources affect social connectedness. Two authors led the data analysis, and the other members helped resolve contradictions during the process.

A meta-analysis was conducted on studies that reported quantitative data. A random-effects meta-analysis was used to calculate the pooled data with 95% CIs. The extracted data were used to compute the SE of the prevalence of digital resource usage for each included study, using the equation below, where n represents the sample size and p represents the sample proportion [51].

$$SE=p\times(1-p)n$$

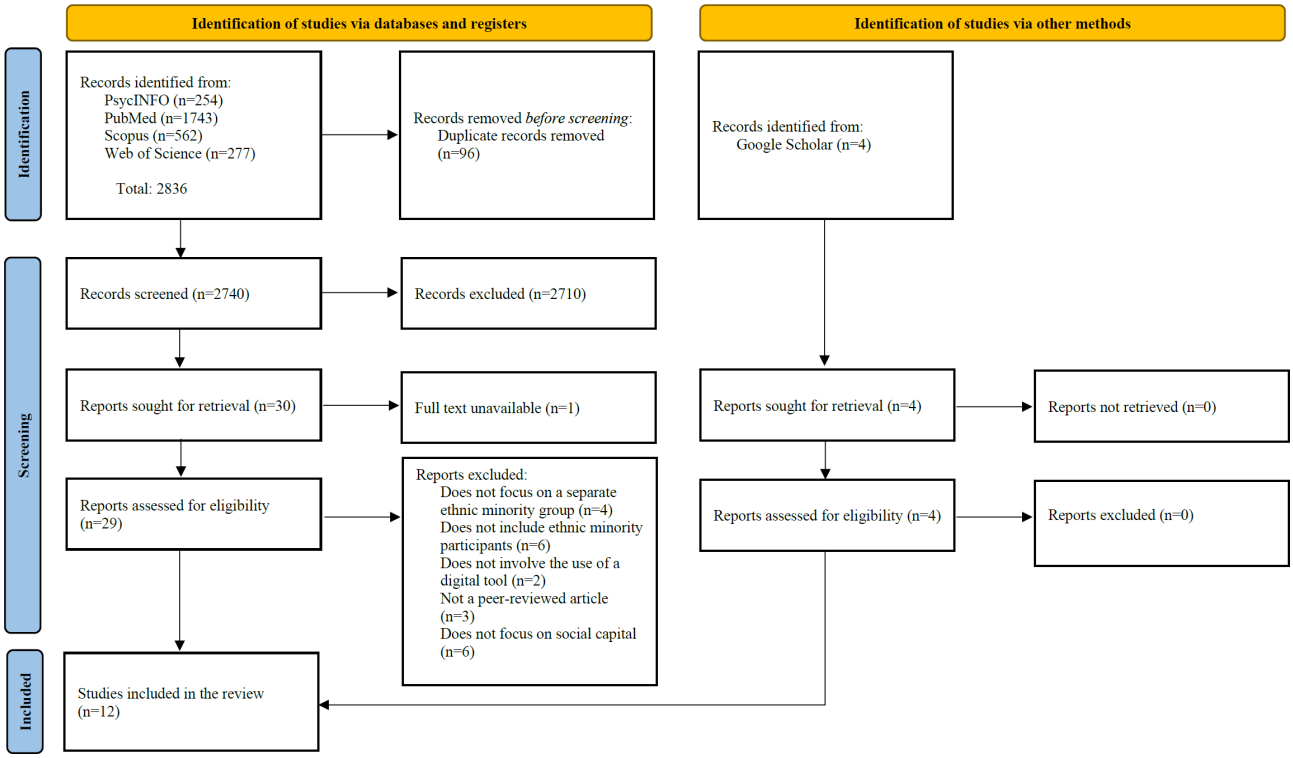
Heterogeneity among studies was estimated using the I^2 index [52], with values categorized as follows: 0%, no heterogeneity; 25% - 50%, low heterogeneity; 51% - 75%, moderate heterogeneity; and >75%, high heterogeneity. Heterogeneity was considered statistically significant at a P value of .05. Forest plots and tables were generated for the visualization of the results. Data were analyzed using Review Manager 5.4 (Cochrane).

Results

Search Results

An initial search of 4 electronic databases yielded 2836 articles, of which 96 duplicates were removed. The remaining articles were screened by title/abstract, and 2710 articles were excluded using the exclusion criteria. Of the remaining 30 articles, 8 met the inclusion criteria. Four additional articles were included through a Google Scholar hand search. Finally, 12 articles were included in this review (Figure 1).

Figure 1. Screening of studies based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [37].



Characteristics of the Included Studies

Table 1 presents the characteristics of the included studies. The research approaches used by the included studies were as follows: quantitative (n=5) [42-45,53], qualitative (n=5) [46-49,54], and mixed methods (n=2) [55,56]. The included studies were mainly conducted in 4 central countries (United States, n=6; Australia, n=3; Finland, n=1; Israel, n=1), and 1 study was conducted in multiple countries.

XSL•FO
RenderX

Table . Characteristics of the included studies.

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
DeLange Martinez et al [42], United States	To examine associations between PU[a], PEOU[b], ICT[c] use, social connectedness, and loneliness among low-income, older Asian Americans during the COVID-19 pandemic	Cross-sectional	Older Asian Americans in affordable senior housing	ICT	Social integration and cultural connection	401	79.07 (SD 7.0)	• PEOU was positively associated with PU (β=.15; *P*=.01). • PEOU and PU were positive predictors of ICT use (PEOU: β=.26; *P*<.001; PU: β=.18; *P*=.01). • ICT use increased social connectedness and reduced lonelines.	• The COVID-19 pandemic spurred an exponential increase in the use of ICT for social connection. When ICTs are perceived as more useful, their usage tends to increase. • Increased usage and familiarity make it easier, encouraging individuals to master new skills and diversify their usage.

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
Lee et al [56], United States	To explore whether a socially assistive robot (Hyodol) would have positive effects on Korean American immigrant older adults' health behaviors and emotional well-being and whether the older adults would be receptive to the robot	Mixed methods	Being aged 65 years or older; Having the ability to engage in verbal conversations and understand either Korean or English; Having the willingness to take Hyodol (in a boy or girl costume) home and use it	Hyodol socially assistive robot	Social integration, and cultural and emotional connections	30	82.6 (SD 6.36)	• Themes: (1) Bridging the gap in interpersonal relationships; (2) Serving individual health functions • Barriers to adoptability: (1) Technological challenges; (2) Health-related and other challenges	• Encouraging regular communication with robot companions in their mother tongue at home may help mitigate feelings of depression and loneliness among immigrant older adults by offering an avenue for social interaction.
Adeniji and Ashirifi [46], United States	To contribute to a broader understanding of the coping mechanisms employed by older African immigrants to promote social connections and interactions	Qualitative	Being aged 60 years or older; Having lived in the United States for 1 or more years; Being born in sub-Saharan Africa	Social media platform (Facebook and Instagram); Television (TV); Phone calls and video	Social integration, and intergenerational, religious, and transnational connections	11	72.18 (SD 5.23)		

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								• Positive self-talk and adaptation: "I have to cope with it" • Technology/social media: "If I cannot interact physically outside, I go through social media/watch TV" • Intergenerational engagement through caregiving: "They [grandchildren] are my immediate constituency" • Digging deep through faith: "Interactions have been mostly within the church"	• Older African immigrants' coping strategies can be enhanced through education that focuses on English language and technology classes. • Spending time with grandchildren brings them happiness and a sense of connection, and helps them learn English.
Nguyen et al [47], Australia		Qualitative	NR[d]	ICT	Cultural and transnational connections	22	72.8		

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
	To illustrate the ways ICTs facilitate the transnational life course transitions of Vietnamese migrant grandparents in Australia through life course digital learning							• Changes and constraints faced in transnational lives • Digital kinning: Maintenance of transnational care and social support networks • Digital homing: Development of migrant belongings and preservation of social and cultural identities	• Not all grandparents are proficient with digital skills and knowledge before their migration. • Digital citizenship, which is developed through the migration process as a part of life-course learning, can help older migrants cope with their increasingly transnational lives.
Pandya [45], United States	To examine the impact of an IMP[e] on sociocultural adaptation, coping, and quality of life of South Asian older widows emigrating in later life to live with their adult children in the United States	Quasiexperimental active control	Having an inclination to engage; Having immigrated to the United States to live with adult children in the recent 2-4 months (in the year 2017); Self-reported fair comfort levels with digital technology and smartphone usage	Graphical user interface software designed for synchronous, online multimedia sessions	Cultural, religious, intergenerational, and transnational connections	172	NR		

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
									• IMP is a meaningfully engaging intervention for late-life immigrant South Asian older widows, enabling adjustment in the new social and family contexts.

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								• Compared with the online informative games program, results indicated that the IMP group exhibited greater adjustment (Cohen *d* range 0.55-0.87; *P*≤.01). • Older widows with higher formal education, from Hindu and Buddhist backgrounds, who are financially independent, have other kith and kin in the destination country, are in fair health, and have moved in with their daughters or live	

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								with their adult children who are alone or only with a partner, reported higher scores on intended outcomes. • Immigrant adult child's gender and family configuration were stronger predictors, and IMP lessons attended and self-practice mediated the relationship between significant sociodemographics and outcomes.	
Pan et al [53], Belgium and Netherlands		Cross-sectional	Chinese migrants in Belgium and the Netherlands aged 50 years or older	Telephone or social media	Social participation and transnational connection	98	NR		

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
	To explore the consequences of the coronavirus pandemic for older Chinese migrants in Belgium and the Netherlands in terms of how increased low-person contact and decreased social participation affect loneliness and how they navigate those situations during the COVID-19 pandemic								• In-person social contact, for example, seems to have been replaced, or even supplemented, by increased non–in-person contact: 43.9% had increased non–in-person contact with their children and 81.6% with nonkin members.

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								• Reduced social participation increased loneliness (EXP β=7.028; *P*=.039) • Problem-focused coping strategies (measured as increased non–in-person contact via telephone or social media) and emotion-focused coping strategies (measured as increased distraction through increased participation in individual activities) were not found to protect against increased loneliness during the pan-	

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								demic.	
Kouvonen et al [43], Finland	To examine the associations among depressive symptoms and SRH[f] with different dimensions of DIT[g] use in older migrants	Cross-sectional	NR	DIT	Social interaction, social participation, social support, and emotional and intergenerational connections	1082	63.2 (SD 8.4)		

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								• After adjusting for sociodemographic and socioeconomic factors, depressive symptoms (OR[h] 2.68, 95% CI 1.37 - 5.24; *P*=.004) and poor SRH (OR 7.90, 95% CI 1.88 - 33.11; *P*=.005) were associated with a higher likelihood of not using the internet daily. • Depressive symptoms (OR 1.88, 95% CI 1.06 - 3.35; *P*=.03) and poor SRH (OR 5.05, 95% CI 1.58 - 16.19; *P*=.006) also increased the likelihood of smart-	• Social media can become an increasingly important source of feelings of connectedness in older age, particularly when retirement and declining health limit other forms of social engagement. • Technology usage can increase older migrants' ability to maintain and expand their dispersed support networks [41]. However, mental health problems, such as depression, may also result in decreased

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								phone nonuse. • Depressive symptoms were additionally associated with a lower likelihood of social media use, and poor SRH was associated with a lower likelihood of using the internet for messaging and calling.	initiative to engage in new activities.
Jun et al [44], United States	To understand the role of ICT use on social support and life satisfaction	Cross-sectional	NR	ICT: Emailing, smartphone usage, internet searching, and general social media usage	Social participation and social support	150	67.5 (SD 6.96)		

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
									• Many Korean immigrant elders are strongly attached to their home country's culture, which emphasises collectivism and interdependence. • Social service practitioners and Korean community agencies working with Korean immigrant elders must recognize their clients' capacity for technology use and encourage its use.

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								• Hierarchical multiple regression results indicated that a high level of smartphone use was significantly related to a high level of social support from families (β=0.05; *P*≤.05). • High levels of smartphone use (β=0.05; *P*≤0.05) and social media use (β=0.06; *P*≤.05) were significantly related to a high level of social support from friends. • A high level of social media use was significantly associated with higher life satisfaction	

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								among Korean immigrant elders (β=0.02; $P \leq .05$).	
Juul et al [48], Australia	To investigate the role of "touchscreen technology" in facilitating increased physical activity and stimulating social interaction in RACFs[i] in order to decrease social and physical inactivity	Qualitative	NR	Touchscreen technology (a large 165-cm interactive portable touch screen, with Windows software installed, on a mobile stand that could be rolled from room to room)	Social interaction, social participation, and cultural and emotional connections	In-depth face-to-face interviews=15; Observational session=around 16 per ward	NR	• New technologies can enhance meaningful physical and social engagement beyond cultural and linguistic barriers. • Technologies encourage social interactions and reduce boredom among residential care older adults. • Effective use of technologies in social engagement in residential care requires well-trained and motivated staff.	• The use of technology not only provides residents with instrumental and functional support, but it also has the potential to stimulate social interaction with each other and with staff, and can help transcend language barriers. • If key social dimsnions are not well-considered, technologies can cause reduced human contact.

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
Millard et al [49], Australia	To understand the increasingly important role of digital citizenship (the ability to participate in society online) in supporting the well-being of aging migrants	Qualitative	NR	Digital citizenship (the ability to participate in society online)	Social interaction, social support, and cultural, intergenerational, and transnational connections	15	NR		

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								• Older people's maintenance of support networks and social engagement, as well as their access to health care services, can be enhanced when they are motivated to increase their digital literacy (the ability to use the internet for information and communication) through appropriate educational, technological, and social support. • This support was more effective when developed	• Older migrants want to connect with members of their support networks who are both local and distant, and both online and offline. • Older migrants enjoy greater direct access to their social networks, information, services, and personal interests. • Older migrants report a greater sense of autonomy and social participation as their digital literacy increases. • Social learning settings functioning as communi-

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								through social learning systems that create communities of practice. Improving digital literacy has significant implications for the well-being of older migrants, as it can enhance their ability to exchange emotional support across distances.	ties of practice provide effective means for developing digital literacy and digital citizenship.
Andonian [55], United States	To explore the meanings, occupational engagement, and experiences associated with computer use	Mixed methods	Being aged 65 years or older; Living independently in the community; Having the ability to give informed consent; Having some exposure to computer use; Having no severe cognitive or mental health issues affecting the ability to participate	Computer use: Tablets; laptops; smartphones; and any computer-associated tasks, such as using the internet, downloading media, and creating documents	Social participation, social support, and cultural and emotional connections	9	NR		

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
									• Older adults discuss sharing their neighborhoods and architecture with friends from around the world through sharing photos. • Immigrants value computer experiences that allow them to exercise self-control, express themselves, and pursue their interests and values through participation in communities across the nation and the world.

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								• Engagement with digital resources led to: (1) Self-expression through self-directed engagement that reflects one's interests, personal history, values, identity, and roles; (2) Freedom and appreciation of unrestricted access to instant and multiple sources of information; (3) Personal growth and feeling open to new experiences; (4) Strengthened relationships that provide social and emo-	

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								tional support; (5) Shift in worldview and perception of the world as larger yet more reachable; (6) Membership in the digital world is active participation in the digital realm.	
Khvorostianov et al [54], Israel	To explore how using the internet may facilitate coping with the challenges of immigration in later life, based on the case of older Jewish immigrants from the Former Soviet Union in Israel	Qualitative	Immigrants who are active internet users	Internet use	Professional and social networks, and intergenerational and transnational connections	32	76		• Managing health online helps participants cope with aging despite lacking the host language skills. • Internet usage is not only vital for the participants' health but also provides significant personal empowerment.

XSL•FO
RenderX

Study and country	Aim	Study design	Inclusion criteria	Type of digital resource	Dimensions of social connectedness	Sample size	Mean age (years)	Key results	Key arguments
								• Results indicated that internet usage by the study participants involved (1) managing health; (2) nurturing professional interests; (3) maintaining and extending social networks; (4) appreciating the past; and (5) enjoying leisure. • Each usage seemed to preserve and even strengthen the participants' self-worth, improving their quality of life.	

[a]PU: perceived usefulness.

[b]PEOU: perceived ease of use.

[c]ICT: information and communication technology.

[d]NR: not reported.

[e]IMP: internet-based meditation program.

[f]SRH: self-rated health.

[g]DIT: digital information technology.

XSL•FO
RenderX

[h]OR: odds ratio.

[i]RACF: residential aged care facilities.

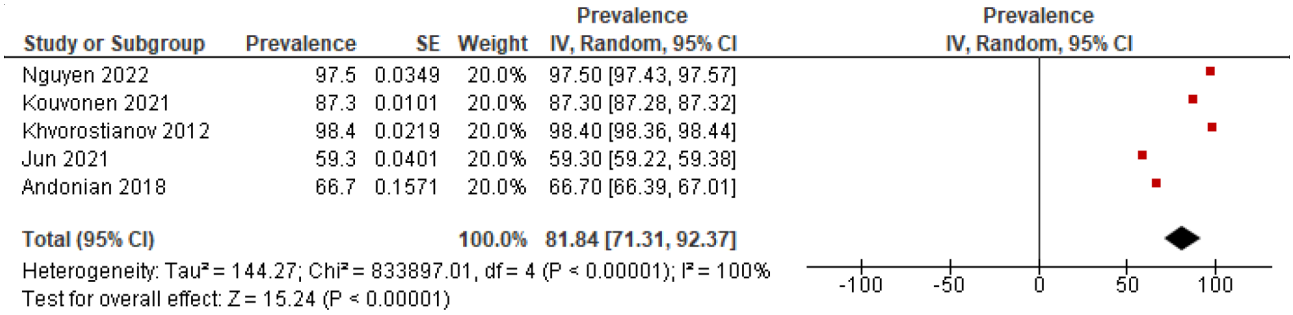
Meta-Analysis

Prevalence of Digital Resource Usage Among Ethnic Minority Older Adults

Prevalence of Internet Use Among Ethnic Minority Older Adults

Internet use was defined as engagement in online activities through digital devices. Five studies (1293 participants) reported prevalence data on internet use among ethnic minority older adults [43,44,47,54,55]. The estimated pooled prevalence of internet use was 81.84% (95% CI 71.31 - 92.37). However, the observed between-study heterogeneity was significantly high (I^2=100%; P<.001; Figure 2).

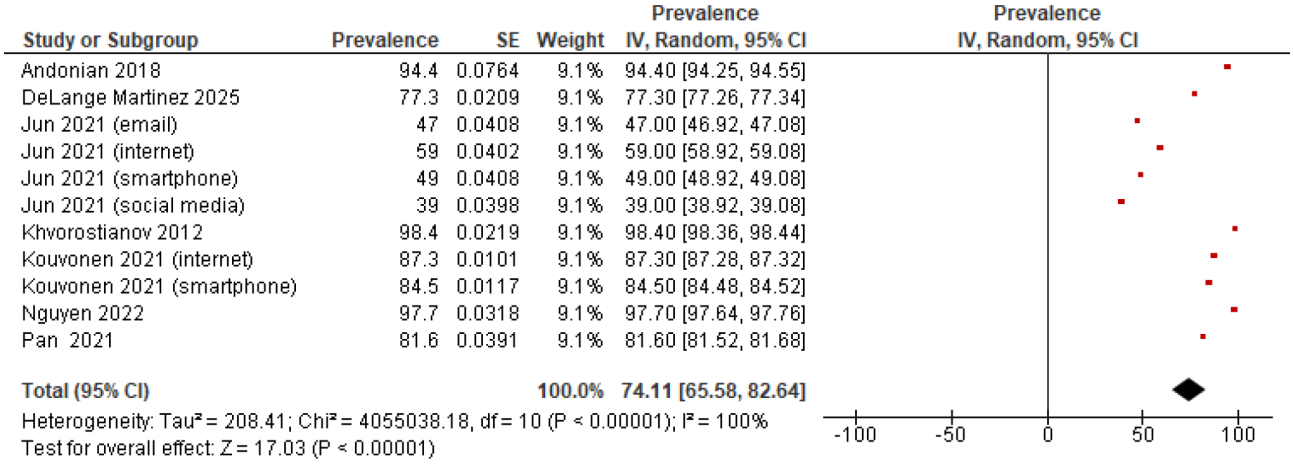
Figure 2. Forest plot of the prevalence of internet use [43,44,47,54,55].



Overall Use of Information and Communication Technology

The overall prevalence of information and communication technology resources that enabled digital engagements was calculated from 7 studies (1794 participants) [42-44,47,53-55]. The overall estimated pooled prevalence was 74.11% (95% CI 65.58 - 82.64). However, a significantly high between-study heterogeneity was observed (I^2=100%; P<.001; Figure 3).

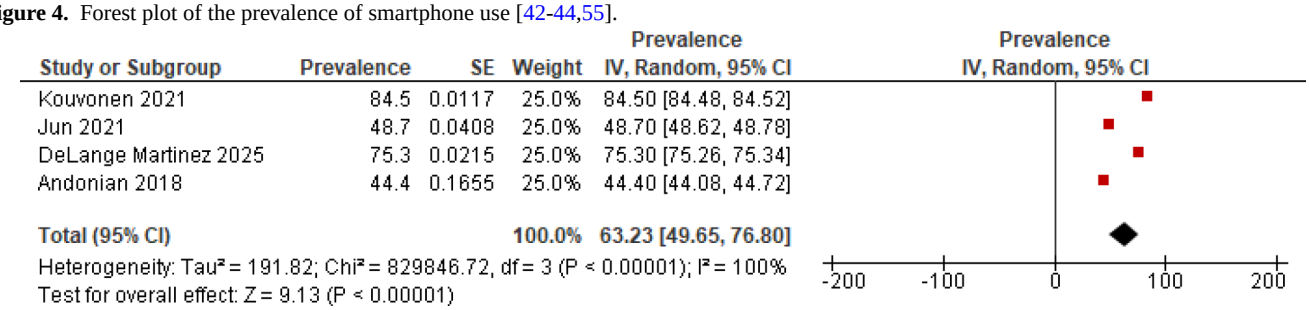
Figure 3. Forest plot of the prevalence of information and communication technology use [42-44,47,53-55].



Prevalence of Smartphone Use Among Ethnic Minority Older Adults

Four studies (1525 participants) reported prevalence data on smartphone use among ethnic minority older adults [42-44,55]. The estimated pooled prevalence of smartphone use was 63.23% (95% CI 49.65 - 76.80). However, a significantly high between-study heterogeneity was observed (I^2=100%; P<.001; Figure 4).

Figure 4. Forest plot of the prevalence of smartphone use [42-44,55].

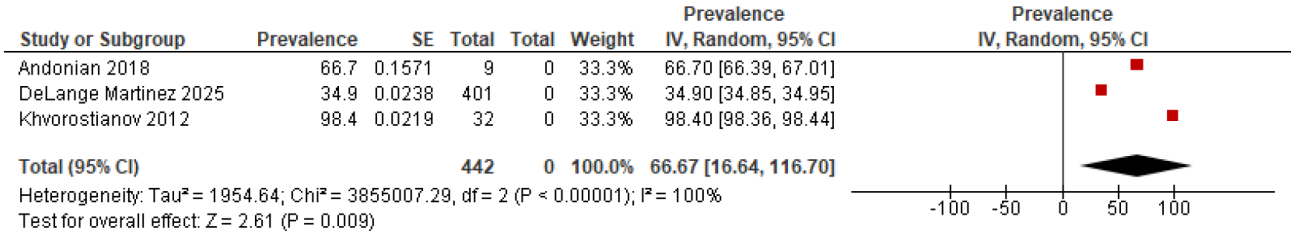


XSL•FO
RenderX

Prevalence of Computer Use Among Ethnic Minority Older Adults

Three studies (442 participants) reported prevalence data on computer use (ie, desktops and laptops) among ethnic minority older adults [42,54,55]. The estimated pooled prevalence of computer use was 66.67% (95% CI 16.64 - 116.70). However, a significantly high between-study heterogeneity was observed (I^2=100%; P<.001; Figure 5).

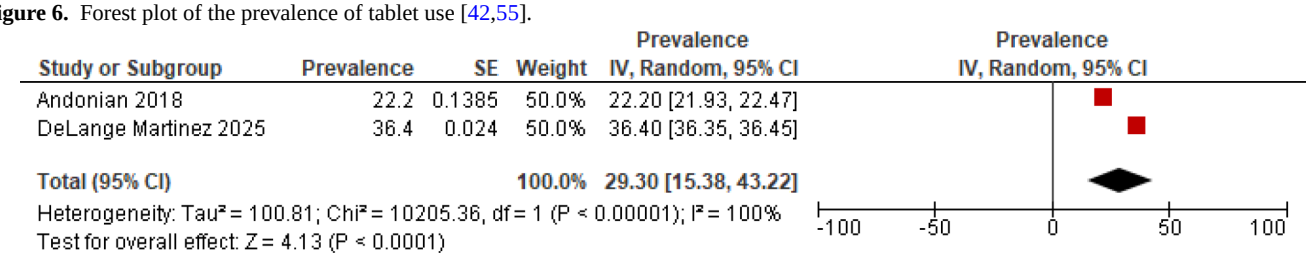
Figure 5. Forest plot of the prevalence of computer use [42,54,55].



Prevalence of Tablet Use Among Ethnic Minority Older Adults

Two studies (410 participants) reported prevalence data on tablet use among ethnic minority older adults [42,55]. The estimated pooled prevalence of tablet use was 29.30% (95% CI 15.38 - 43.22). A significantly high between-study heterogeneity was observed (I^2=100%; P<.001; Figure 6).

Figure 6. Forest plot of the prevalence of tablet use [42,55].



Thematic Synthesis

Thematic synthesis of the extracted data generated the following two major themes: (1) utility of digital resources for social connectedness, and (2) factors that influence the use of digital resources among ethnic minority older adults.

Utility of Digital Resources for Social Connectedness

Overview

The following six subthemes relating to the utility of digital resources for social connectedness were identified: (1) enhancing social integration and participation, (2) maintaining cultural and religious connections, (3) promoting transnational connectedness, (4) fostering intergenerational connectedness, (5) expanding social networks for social connectedness, and (6) downsides of digital resources for social connectedness.

Enhancing Social Integration and Participation

Findings related to digital resources enhancing social integration and participation emerged from quantitative (n=4) [42-44,53], qualitative (n=3) [46,48,49], and mixed methods (n=2) [55,56] studies. It was revealed that platforms, such as social media, enabled ethnic minority older adults to actively participate in social activities, access information and community resources, learn new skills, and interact with community members. For example, a cross-sectional survey of 98 Chinese immigrants in Belgium and the Netherlands found that 81.6% of respondents had more frequent non–in-person contact with kin and nonkin individuals via smartphone-enabled social media to support social participation during the COVID-19 pandemic [53]. Qualitative evidence highlighted the critical role of digital resources in promoting a sense of community belonging among ethnic minority older adults. A study among older adults from non–English-speaking backgrounds living in a residential aged care facility in Western Australia made the following argument [48]:

> *[Touchscreen technology] lifted the mood of the room, transforming it into a livelier space and facilitating a sense of community and shared fellowship among all participants… residents went from being individuals doing nothing to active members of a group, contributing to a sense of belonging.*

Maintaining Cultural and Religious Connections

Evidence on digital resources playing a critical role in maintaining cultural and religious connections among immigrant older adults was drawn from 4 qualitative studies [46-49], 2 quantitative studies [42,45], and 2 mixed methods studies [55,56]. A quasiexperimental study showed that participants who received an internet-based meditation program across various devices demonstrated significantly greater sociocultural adaptation compared with a control group, with effect sizes ranging from 0.55 to 0.87 (P≤.01) [45]. It was observed that digital resources created a convenient virtual space for religious practices, particularly during the COVID-19 pandemic, by connecting the older adults with people of similar faith. In a study of 11 older African immigrants, it was reported that avenues, such as Facebook, created opportunities for social, cultural, and spiritual engagement with others during the pandemic [45]. For instance, a study among older immigrants in Western Australia made the following statement regarding individuals' motivation for participating in the Internet Café community of practice [49]:

XSL•FO
RenderX

> *…desire to connect with the places and memories of their past…on one occasion, the cafe was filled with Irish folk music as a patron played the traditional sounds of his homeland from a website.*

Promoting Transnational Connectedness

Seven studies (4 qualitative studies and 3 quantitative studies) indicated that the availability of digital resources, such as the internet, smartphones, and media, such as video calls or group chats, allowed immigrant older adults to maintain connections with family members and friends across countries [42,45-47,49,53,54]. A cross-sectional survey of 401 older Asian Americans found that 89.2% of respondents agreed that digital resources were useful for engaging family and friends [42]. Khvorostianov et al [54] noted that transnational connections formed through digital resources often led to empowerment and better integration of immigrants into their host societies, including the development of lasting social relationships. It was observed that digital resources played a critical role in facilitating financial transactions, thereby enabling stronger social connectedness through the continuity of transnational caregiving and support. For example, a participant in a qualitative study involving 22 Vietnamese immigrant grandparents made the following statement [47]:

> *[She] maintained daily contact via video phone calls…with her son and grandchildren, as well as a group of Vietnamese close friends…For her mother, she provided transnational care by sending remittances to pay the domestic worker, cover her mother's daily expenses, and purchase nutritional supplements.*

Fostering Intergenerational Connectedness

Findings from 3 qualitative studies [46,49,54] and 2 quantitative studies [43,45] indicated that digital resources fostered intergenerational connections among ethnic minority older adults. Digital resources were important interventions for bonding with grandchildren, facilitating intergenerational language exchange, sharing family history, and mentoring the younger generation. According to Kouvonen et al [43], the frequent use of social media among 631 older immigrants enabled intergenerational interactions. Khvorostianov et al [54] identified that the internet served as a tool that connected grandchildren to their heritage:

> *[We] built a website where we publish articles, memories and documents telling how our relatives were murdered. My grandchildren browse this website too and learn about their family history.*

Expanding Social Networks for Social Connectedness

Evidence from 2 qualitative studies [49,54], 2 quantitative studies [43,44], and 1 mixed methods study [54] indicated that digital resources are critical for expanding the social networks of older adult immigrants. It was reported that internet-enabled computer usage helped ethnic minority older adults engage in social activities that expanded their social participation beyond their immediate environment [55]. Similarly, the use of social media platforms, facilitated by digital tools, supported development and reconnection with past social networks [43,44,49,54,55]. It was revealed that maintaining social networks online helped older adult immigrants cope with distance from family members and friends; thus, it provided physical and emotional support to remain socially connected and alleviate feelings of social isolation [49,54,55]. A study from Australia made the following statement [49]:

> *Improving digital literacy has special implications for the well-being of older migrants because it can enhance their ability to exchange emotional support across distance.*

Downsides of Digital Resources for Social Connectedness

Notwithstanding, having access to digital resources was not always impactful in reducing the outcomes of social connectedness, such as loneliness [53]. Other studies indicated the likelihood that digital resources will reduce human contact, which can worsen quality of life [48]. In the article by Andonian [55], some participants presented instances that can lead to poor social connectedness through physical isolation, such as the following:

> *When I came here, I didn't know anything. Everybody was like running all over the city with, you know [points to smartphone]. What are they doing?! You know, they're not saying "hello" to anybody around them…*

These observations align with others that indicate the limitations of digital resources in promoting social connectedness. For instance, it was found that having a poor health status and being socioeconomically disadvantaged led to less engagement with digital resources, making them less relevant to promoting low social connectedness, as presented in other findings [43,55].

Factors That Influence the Use of Digital Resources Among Ethnic Minority Older Adults

Overview

The systematic review identified sociodemographic, economic, and health-related factors that influence the adoption and use of digital resources among ethnic minority older adults.

Sociodemographic and Economic Factors

Age

Older age was found to be negatively associated with digital resource use [42,43,46,47,49,54,55]. For instance, Kouvonen et al [43] reported that nonuse of the internet and smartphones was prevalent among individuals aged 65 years or older compared to younger age groups (internet nonuse: 97/421, 23.0% vs 40/646, 6.2%; smartphone nonuse: 98/351, 27.9% vs 51/614, 8.3%). Additionally, Nguyen et al [47] found that age was associated with low digital literacy among ethnic minority older adults (Vietnamese Australians). However, they were motivated to learn how to use digital resources due to the social connection opportunities they offered, such as participation in religious events.

Gender

Gender was found to be a predictor of digital resource use among ethnic minority older adults [42,49,54]. DeLange

XSL•FO
RenderX

Martinez et al [42] revealed that female gender was significantly and negatively associated with the use of digital resources and platforms (r=–0.15; *P*=.01). Khvorostianov et al [54] found that older women were more likely to use the internet to develop new professional interests, whereas older men used it to maintain their original professional identity, which helped build their self-confidence as active members of their communities.

Educational Level

Seven studies found that ethnic minority older adults' adoption and use of digital resources were influenced by their levels of education [42-47,54]. DeLange Martinez et al [42] found that older adults with higher educational levels were more likely to utilize digital tools for social connections compared to those with lower educational levels (β=.21; *P*<.001). Nguyen et al [47] stated that Vietnamese migrant grandparents with lower education were less likely to engage with digital resources, with most using them for the first time after their migration. To address disparities in the adoption and use of digital resources, instructions for these resources should be clear and simplified for ethnic minority older adults with limited educational backgrounds to ensure ease of use [45].

Language Proficiency

Few studies stated that language proficiency was related to the adoption and perceived ease of use of digital resources among ethnic minority older adults [42,43]. Kouvonen et al [43] reported that older Russian-speaking immigrants in Finland who had effectively integrated into Finnish society and had strong local language skills were more likely to own a smartphone to access the internet. Linguistic bias in digital resource and application development, which favors English-speaking users, needs to be addressed to facilitate the use of technology among culturally and linguistically diverse older adults [42].

Religion

One study found religion to be associated with the use of digital resources among immigrant South Asian older widows [45]. Pandya [45] identified that an internet-based meditation program was more effective for older Hindu and Buddhist widows than Christian widows because of cultural familiarity with meditation practices.

Economic Status

Economic status was identified as an influential factor in the digital technology used by ethnic minority older adults [43-45,49]. For example, Jun et al [44] reported that Korean immigrant older adults with an annual income greater than US $20,000 were significantly more likely to use digital resources than those in the lower-income group (t_{148}=−4.53; *P*≤.001). The financial cost of digital resources, such as purchasing a smartphone or the monthly cost of internet services, was a significant barrier for low-income older adults, which increased the prevalence of nonuse of digital resources [43].

Health-Related Factors

Physical Health

The physical health of ethnic minority older adults has been identified as a critical factor in the adoption and use of digital resources [43-45,55,56]. Physical health challenges, including decreased vision and coordination, affected the use of digital resources. For example, Kouvonen et al [43] reported that older Russian-speaking immigrants in Finland with poor physical health were 5 times less likely to own and use smartphones compared to those with good physical health (odds ratio [OR] 0.20, 95% CI 0.06 - 0.69; *P*=.006).

Mental Well-Being

Mental health challenges, such as cognitive decline, anxiety, and depression, were identified as critical barriers to the adoption and use of digital resources among ethnic minority older adults [42,43,56]. As observed by Kouvonen et al [43], ethnic minority older adults experiencing depressive symptoms were almost 3 times more likely not to use the internet for messaging or calling compared to those without depressive symptoms (OR 2.68, 95% CI 1.37 - 5.24; *P*=.004).

COVID-19 Restrictions

One study investigated the impact of the COVID-19 pandemic on the adoption and use of digital resources among ethnic minority older adults [53]. Pan et al [53] stated that COVID-19 accelerated the adoption and use of digital resources among older Chinese migrants in Belgium and the Netherlands, with increases of 43.9% and 81.6% in non–in-person contact with family and friends, respectively.

Discussion

Principal Findings and Comparison With Other Studies

Maintaining active social connections is critical to the health and well-being of ethnic minority older adults, as it reduces the risk of social isolation and loneliness [10]. Thus, understanding and promoting the dynamics of social connections among this group of people are important for protecting their health and well-being. The study found significant variation in the prevalence estimates of digital resource use among ethnic minority older adults, with most preferring to use the internet for social connectivity. Overall, the prevalence of information communication technology use was 74.11%, and the estimated prevalence of internet use was 81.8%. Moreover, the values for computers, smartphones, and tablets were 66.7%, 63.2%, and 29.3%, respectively. Internet use ranged from 59.3% among older Korean immigrants in the United States [42] to 98.4% among older Jewish immigrants from the Former Soviet Union in Israel [54]. These findings indicate that while digital engagement for social connection is considerable among ethnic minority older adults, usage disparities exist across ethnic groups [42]. In comparison to studies about older adults generally, which also note a digital divide, these results emphasize the critical role of ethnic and cultural context in shaping digital resource utilization [25,32]. The high prevalence of internet use suggests that targeted digital inclusion strategies could be highly

XSL•FO
RenderX

effective. However, these must be culturally sensitive and address the unique needs and preferences of each ethnic minority group [28,42].

The qualitative synthesis identified two related themes: (1) the utility of digital resources for social connectedness, and (2) the factors that influence their use among ethnic minority older adults. These themes offer insights into the lived experiences of ethnic minority older adults in engaging with digital resources. It was revealed that digital resources function as essential tools for enhancing social integration and participation, maintaining cultural and religious connectedness, promoting transnational connections, and fostering intergenerational connectedness through expanded social networks. These findings suggest that digital resources facilitate access to information and opportunities for ethnic minority older adults to maintain meaningful social relationships with their families who have been left behind in their home countries or have been apart from them while also enabling these older adults to create new social connections to support their well-being [34,36]. From the perspective of social identity theory, digital resources can be instrumental in how ethnic minority populations construct and maintain their self-concepts [21,23]. In this review, such identity construction and maintenance are influenced by digital resources, which bridge geographical boundaries for social connectedness and enable ethnic minority older adults to meet cultural expectations, such as supporting younger generations.

Despite the overall positive aspects of digital resources for social connectedness, this review also points to instances or conditions that make them less relevant or even a causal mechanism for low social connectedness among aging ethnic minority people [53,55]. This finding underscores the importance of physical social connections in fostering social connectedness, as virtual interactions have significant limitations (eg, difficulty in conveying emotions effectively) [56]. These observations mean that the integration of digital resources in promoting social connectedness must be done from a balanced perspective—one that seeks to complement physical interactions with digital resources instead of attempting to entirely replace or overly prioritize the technologies over actual human contact [57]. For ethnic minority older adults, such a balance can be crucial to their acculturation, as found in studies showing that the high prevalence of smartphone use made entry into local cultural dynamics challenging for minority migrant individuals [55,56]. These downsides to digital resources were strongly related to intersectional factors, such as economic status, health status, and others, which require serious attention to enable the sustainable use of technologies in promoting social connectedness, as further elaborated in the following paragraphs.

The systematic review identified several factors that influenced the adoption and use of digital resources among ethnic minority older adults, which were categorized into sociodemographic, economic, and health-related factors. Identified sociodemographic factors included age, gender, educational level, language proficiency, and religion. These factors indicate the intersectional nature of digital engagement among older adults in general, with significant stratification across gender, race/ethnicity, education, and other areas [58]. Considering that many ethnic minority older adults are illiterate (at least in the context of host societies) and have poor language proficiency [26], their use of digital platforms/resources may be weaker than that of ethnic majority older adults in a given place, leading to fewer opportunities to build social connections. At the same time, limitations in language and other intersectional factors can also drive ethnic minority older adults to resort to digital resources [23]. In addition, the results across 7 studies were consistent on the issue that younger ethnic minority individuals were more likely to adopt and use digital resources, which further indicates the significant role of factors, such as educational attainment (mostly higher among young people), in the use of digital resources [42,43,46,47,49,54,55]. However, there was evidence that many ethnic minority older adults were highly motivated to utilize digital resources to stay connected with family and friends [46,47,55], which demonstrates the utility of digital resources for sociocultural adaptation and maintenance of social identity. A previous study has shown that the relationship between age and the adoption of digital resources is mediated by computer self-efficacy, cognitive ability, and computer anxiety [59]. Strong motivation toward digital inclusion can be crucial in cultivating and achieving desirable outcomes that promote social connection through digital means [5].

The review found that the income level of ethnic minority older adults significantly influenced their adoption and use of digital resources. A combination of factors, including high-cost digital resources, limited internet access, and a lack of digital skills, is a key barrier that excludes low-income older adults from accessing and using digital resources. This finding reinforces the theoretical perspectives of the fundamental cause theory, which argues that the health and well-being outcomes of populations are influenced by their access to critical socioeconomic resources [60]. It has been recommended that for ethnic minority older adults to experience the social connection benefits of digital resources, policies and programs cannot solely focus on the technological aspects of their digital inclusion, as economic resources are key to making such programs functional [49].

Furthermore, health-related factors significantly affected the adoption and use of digital resources among ethnic minority older adults. It was recognized that ethnic minority older adults experiencing physical or mental health challenges, such as cognitive decline, depression, anxiety, decreased vision, and coordination, were less likely to own and use digital resources. This is likely because physical and mental health issues can cause disinterest in activities of daily living, in addition to potential compromises to bodily functions needed for digital engagement, such as vision loss [25,28]. Going forward, a deeper understanding of these influential factors is essential for designing and implementing health-appropriate strategies that enable equitable access to digital resources among ethnic minority older adults. Moreover, as with other population groups [61,62], COVID-19 pandemic restrictions, including social distancing measures, led to an increased use of digital resources. Other studies among ethnic minority older adults have found similar results, showing that people used video calls to maintain social connections [63,64]. Thus, experiences during the pandemic can be instructive for incorporating digital tools to

XSL•FO
RenderX

promote social connectedness among minoritized older adults whose sociocultural and adaptation issues weaken their opportunities for lasting social connectedness. Considering the dangers of social isolation as an "epidemic," as observed during the COVID-19 pandemic, can hasten efforts toward integrating vulnerable older adults into social dynamics through digitalization.

Strengths and Limitations

To the best of our knowledge, this is one of the first systematic reviews and meta-analyses to explore the adoption of digital resources to address social connection challenges and the factors influencing their adoption and use among ethnic minority older adults globally. The study followed a rigorous approach to searching relevant articles and extracting and handling data. However, the study has limitations that must be acknowledged when interpreting the findings. First, we limited data extraction to full articles written in English and with full text available, which may have excluded relevant studies. Second, the included studies used different age thresholds to define older adults, including one that recruited participants aged 50 years or older [53]. Divergence in age criteria means that the experiences of older adults across studies are likely to differ, potentially affecting the compatibility of results across the included papers. Third, substantial heterogeneity was detected across the included studies; hence, a random-effects model was applied. Fourth, the lack of information regarding the use of validated instruments in the quantitative studies may have introduced measurement bias. Additionally, it was not possible to infer causal relationships between the use of digital resources and social connectedness because almost all the quantitative studies used cross-sectional designs. Notwithstanding these limitations, this study offers rare, collated evidence on the characteristics and influence of digital resource use among the understudied group of ethnic minority older adults. The evidence offers an opportunity to conduct further empirical studies to better understand ethnic minority older adults' engagement with digital resources and the causal pathways involved.

Conclusion

The findings of this review suggest that digital resources can enhance social connections among older adults from ethnic minority groups. Intersectional factors related to sociodemographic, economic, and health-related aspects are influential in the adoption and use of digital resources by ethnic minority older adults to maintain social connections. Ethnic minority older adults utilize a wide range of digital tools/platforms to enhance their social integration in new or foreign societies, maintain their cultural and religious heritage, and foster intergenerational connections. The findings suggest that access to culturally appropriate and suitable digital resources for ethnic minority older adults with diverse health statuses can enhance social connectedness. While digital resources appear to be positively associated with social connectedness among ethnic minority older adults, future research can extend existing knowledge by examining how these resources may have negative consequences for the extent and nature of the social connections that older ethnic minority people secure. This is in view of some observations that digital resources can discourage in-person contact and may not necessarily be useful for addressing concerns with low social connectedness.

Additionally, it will be instructive to critically examine the diverse social connections formed by these older adults. This is because the types of social networks one possesses determine the nature of the resources embedded in those networks and the impact they will have on various life areas [65]. Moreover, research is needed on how ethnic minority older adults sustain the social connections they form through digital resources compared to those made in person, to advance social sustainability in aging through digital resources. At the same time, in recognition of the limitations and other factors that affect the impact of digital resources on social connectedness in aging societies, research is needed to better understand how various social technologies can be engaged sustainably (eg, effective use over the long term) among aging populations in different places. Finally, longitudinal study designs will be an ideal step toward understanding the stability in the relations between digital resources and social connectedness among older ethnic minority populations.

Funding

The study received funding support from Lingnan University Research Committee, Hong Kong through the Direct Grant Scheme (reference number: DR25E1; principal investigator: PAA).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy and quality assessments.
[DOCX File, 25 KB - aging_v9i1e84962_app1.docx]

Checklist 1
PRISMA checklist.
[PDF File, 207 KB - aging_v9i1e84962_app2.pdf]

XSL•FO
RenderX

References

1. Kemp S. Digital 2025 global overview report. Data Reportal. 2025. URL: https://datareportal.com/reports/digital-2025-global-overview-report [accessed 2026-03-26]
2. Clohessy S, Arvanitis TN, Rashid U, et al. Using digital tools in clinical, health and social care research: a mixed-methods study of UK stakeholders. BMJ Open 2024 Apr 3;14(4):e076613. [doi: 10.1136/bmjopen-2023-076613] [Medline: 38569710]
3. Damant J, Knapp M, Freddolino P, Lombard D. Effects of digital engagement on the quality of life of older people. Health Soc Care Community 2017 Nov;25(6):1679-1703. [doi: 10.1111/hsc.12335] [Medline: 26919220]
4. Rao PB, Rammya B, Shamili SA, Inampudi P, Meegada VBR. Navigating globalization and digitalization: strategies for enhancing digital inclusion. In: Ahmed EM, editor. Developing Digital Inclusion Through Globalization and Digitalization: IGI Global Scientific Publishing; 2024:367-392. [doi: 10.4018/979-8-3693-4111-7.ch015]
5. Rasi-Heikkinen P, Doh M. Older adults and digital inclusion. Educ Gerontol 2023 May 4;49(5):345-347. [doi: 10.1080/03601277.2023.2205743]
6. O'Rourke HM, Sidani S. Definition, determinants, and outcomes of social connectedness for older adults: a scoping review. J Gerontol Nurs 2017 Jul 1;43(7):43-52. [doi: 10.3928/00989134-20170223-03] [Medline: 28399313]
7. Office of the Surgeon General. Our Epidemic of Loneliness and Isolation: The US Surgeon General'S Advisory on the Healing Effects of Social Connection and Community: US Department of Health and Human Services; 2023. URL: https://www.ncbi.nlm.nih.gov/books/NBK595227/ [accessed 2026-03-26]
8. Haslam C, Cruwys T, Haslam SA, Jetten J. Social connectedness and health. In: Pachana N, editor. Encyclopedia of Geropsychology: Springer; 2015:1-10. [doi: 10.1007/978-981-287-080-3_46-2]
9. Jaffe S. US Surgeon General: loneliness is a public health crisis. The Lancet 2023 May;401(10388):1560. [doi: 10.1016/S0140-6736(23)00957-1] [Medline: 37011643]
10. Joshi M, Finney N, Hale JM. Loneliness and social isolation of ethnic minority/immigrant older adults: a scoping review. Ageing Soc 2025 Jul;45(7):1395-1425. [doi: 10.1017/S0144686X24000205]
11. Social connection linked to improved health and reduced risk of early death. World Health Organization (WHO). 2025. URL: https://www.who.int/news/item/30-06-2025-social-connection-linked-to-improved-heath-and-reduced-risk-of-early-death [accessed 2026-03-26]
12. Fakoya OA, McCorry NK, Donnelly M. Loneliness and social isolation interventions for older adults: a scoping review of reviews. BMC Public Health 2020 Feb 14;20(1):129. [doi: 10.1186/s12889-020-8251-6] [Medline: 32054474]
13. Grenade L, Boldy D. Social isolation and loneliness among older people: issues and future challenges in community and residential settings. Aust Health Rev 2008 Aug;32(3):468-478. [doi: 10.1071/ah080468] [Medline: 18666874]
14. Su Y, Rao W, Li M, Caron G, D'Arcy C, Meng X. Prevalence of loneliness and social isolation among older adults during the COVID-19 pandemic: a systematic review and meta-analysis. Int Psychogeriatr 2023 May;35(5):229-241. [doi: 10.1017/S1041610222000199] [Medline: 35357280]
15. Berry JW, Hou F. Acculturation, discrimination and wellbeing among second generation of immigrants in Canada. Int J Intercult Relat 2017 Nov;61:29-39. [doi: 10.1016/j.ijintrel.2017.08.003]
16. Liu S, Dane S, Gallois C, Haslam C, Nghi Tran TL. The dynamics of acculturation among older immigrants in Australia. J Cross Cult Psychol 2020 Jul;51(6):424-441. [doi: 10.1177/0022022120927461]
17. Park C, Nam S, Redeker NS, et al. The effects of acculturation and environment on lifestyle behaviors in Korean immigrants: the mediating role of acculturative stress and body image discrepancy. Ethn Health 2022 Jan 2;27(1):157-173. [doi: 10.1080/13557858.2019.1658182]
18. Guo M, Wang Y, Liu J, Dong X. Ethnic enclaves, social capital, and psychological well-being of immigrants: the case of Chinese older immigrants in Chicago. Aging Ment Health 2023 Jun 3;27(6):1077-1085. [doi: 10.1080/13607863.2022.2084506]
19. Amoah PA, Koduah AO, Anaduaka US, Addae EA, Gwenzi GD, Amankwaa A. Psychological wellbeing in diaspora space: a study of African economic migrants in Hong Kong. Asian Ethnicity 2020 Oct 1;21(4):542-559. [doi: 10.1080/14631369.2020.1775485]
20. Tajfel H. Social identity and intergroup behaviour. Social Science Information 1974 Apr;13(2):65-93. [doi: 10.1177/053901847401300204]
21. Ward C, Szabó Á. Acculturation, cultural identity and well-being. Nat Rev Psychol 2023;2(5):267-282. [doi: 10.1038/s44159-023-00171-2]
22. Hu J, Cheung CKJ. Social identity and social integration: a meta-analysis exploring the relationship between social identity and social integration. Front Psychol 2024;15:1361163. [doi: 10.3389/fpsyg.2024.1361163] [Medline: 38638525]
23. Lin CC, Li NJ, Lee E. Exploring the intersections of immigrant seniors' digital literacies and social connectedness: a Canadian study. Front Educ 2023;8:2023. [doi: 10.3389/feduc.2023.1250240]
24. Wright PJ, Raynor PA, Bowers D, et al. Leveraging digital technology for social connectedness among adults with chronic conditions: a systematic review. Digit Health 2023;9:20552076231204746. [doi: 10.1177/20552076231204746] [Medline: 37799504]
25. Tappen RM, Cooley ME, Luckmann R, Panday S. Digital health information disparities in older adults: a mixed methods study. J Racial Ethn Health Disparities 2022 Feb;9(1):82-92. [doi: 10.1007/s40615-020-00931-3] [Medline: 33415705]

XSL•FO
RenderX

26. Shum TCT. Quality of life of South Asian older adults in Hong Kong: policy implications for a multicultural active aging framework. Gerontol Geriatr Med 2023;9:23337214231179837. [doi: 10.1177/23337214231179837] [Medline: 37324644]
27. Bamford J, Klabbers G, Curran E, Rosato M, Leavey G. Social capital and mental health among Black and minority ethnic groups in the UK. J Immigrant Minority Health 2021 Jun;23(3):502-510. [doi: 10.1007/s10903-020-01043-0]
28. Ang S, Lim E, Malhotra R. Digital independence: health-related difficulty in internet use and how it affects older adults' quality of life. Innovation in Aging 2019 Nov 8;3(Supplement_1):S442-S443. [doi: 10.1093/geroni/igz038.1661]
29. Frishammar J, Essén A, Bergström F, Ekman T. Digital health platforms for the elderly? Key adoption and usage barriers and ways to address them. Technol Forecast Soc Change 2023 Apr;189:122319. [doi: 10.1016/j.techfore.2023.122319]
30. Yazdani-Darki M, Rahemi Z, Adib-Hajbaghery M, Izadi-Avanji FS. Older adults' barriers to use technology in daily life: a qualitative study. Nurs Midwifery Stud 2020;9(4):229-236. [doi: 10.4103/nms.nms_91_19]
31. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
32. Yang HJ, Lee JH, Lee W. Factors influencing health care technology acceptance in older adults based on the technology acceptance model and the unified theory of acceptance and use of technology: meta-analysis. J Med Internet Res 2025 Mar 28;27:e65269. [doi: 10.2196/65269] [Medline: 40153796]
33. Chen YRR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18. [doi: 10.2196/jmir.4596] [Medline: 26822073]
34. Wei S, Kang B, Bailey DE, et al. Using technology to measure older adults' social networks for health and well-being: a scoping review. Gerontologist 2022 Aug 12;62(7):e418-e430. [doi: 10.1093/geront/gnab039] [Medline: 33754150]
35. Balki E, Hayes N, Holland C. Effectiveness of technology interventions in addressing social isolation, connectedness, and loneliness in older adults: systematic umbrella review. JMIR Aging 2022 Oct 24;5(4):e40125. [doi: 10.2196/40125] [Medline: 36279155]
36. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
37. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
38. Amoah PA, Ameyaw EK, Bhowmik MK, Adotey PN, Donkor A. Digital resources and social connectedness among ethnic minority older adults: a systematic review and meta-analysis. JMIR Aging (forthcoming) 2025. [doi: 10.2196/84962]
39. Donkor A, Atuwo-Ampoh VD, Yakanu F, et al. Financial toxicity of cancer care in low- and middle-income countries: a systematic review and meta-analysis. Support Care Cancer 2022 Sep;30(9):7159-7190. [doi: 10.1007/s00520-022-07044-z] [Medline: 35467118]
40. Donkor A, Adotey PN, Ofori EO, et al. Prevalence of preferences for end-of-life place of care and death among patients with cancer in low- and middle-income countries: a systematic review and meta-analysis. JCO Glob Oncol 2024 May;10:e2400014. [doi: 10.1200/GO.24.00014] [Medline: 38815191]
41. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol 2011 Apr;64(4):407-415. [doi: 10.1016/j.jclinepi.2010.07.017] [Medline: 21247734]
42. DeLange Martinez P, Tancredi D, Pavel M, Garcia L, Young HM. Adapting the technology acceptance model to examine the use of information communication technologies and loneliness among low-income, older Asian Americans: cross-sectional survey analysis. JMIR Aging 2025 Jan 8;8:e63856. [doi: 10.2196/63856] [Medline: 39778204]
43. Kouvonen A, Kemppainen L, Ketonen EL, Kemppainen T, Olakivi A, Wrede S. Digital information technology use, self-rated health, and depression: population-based analysis of a survey study on older migrants. J Med Internet Res 2021 Jun 14;23(6):e20988. [doi: 10.2196/20988] [Medline: 34125069]
44. Jun JS, Galambos C, Lee KH. Information and communication technology use, social support, and life satisfaction among Korean immigrant elders. J Soc Serv Res 2021 Jul 4;47(4):537-552. [doi: 10.1080/01488376.2020.1848969]
45. Pandya SP. The effect of internet-based programs on late-life immigrant older widows: socio-cultural adaptation, coping, and quality of life. J Technol Hum Serv 2021 Oct 2;39(4):349-378. [doi: 10.1080/15228835.2021.1978126]
46. Adeniji DO, Ashirifi GD. The voices of older African immigrants on how they cope with social isolation and loneliness in the United States. Clin Gerontol 2025 Mar 15;48(2):260-269. [doi: 10.1080/07317115.2024.2415949]
47. Nguyen HT, Baldassar L, Wilding R. Lifecourse transitions: how ICTS support older migrants' adaptation to transnational lives. Social Inclusion 2022;10(4):181-193. [doi: 10.17645/si.v10i4.5735]
48. Juul A, Wilding R, Baldassar L. The best day of the week: new technology enhancing quality of life in a care home. Int J Environ Res Public Health 2019 Mar 19;16(6):1000-1016. [doi: 10.3390/ijerph16061000] [Medline: 30893945]
49. Millard A, Baldassar L, Wilding R. The significance of digital citizenship in the well-being of older migrants. Public Health (Fairfax) 2018 May;158:144-148. [doi: 10.1016/j.puhe.2018.03.005]
50. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8(1):45-54. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
51. Collett D. Modelling Binary Data: Chapman & Hall; 2002. [doi: 10.1201/b16654]
52. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 6;327(7414):557-560. [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]

XSL•FO
RenderX

53. Pan H, Fokkema T, Switsers L, Dury S, Hoens S, De Donder L. Older Chinese migrants in coronavirus pandemic: exploring risk and protective factors to increased loneliness. Eur J Ageing 2021 Jun;18(2):207-215. [doi: 10.1007/s10433-021-00625-7] [Medline: 33967662]
54. Khvorostianov N, Elias N, Nimrod G. 'Without it I am nothing': the internet in the lives of older immigrants. New Media & Society 2012 Jun;14(4):583-599. [doi: 10.1177/1461444811421599]
55. Andonian LC. Meanings and experiences associated with computer use of older immigrant adults of lower socioeconomic status. Can J Occup Ther 2018 Apr;85(2):146-157. [doi: 10.1177/0008417418754395]
56. Lee OE, Nah KO, Kim EM, Choi NG, Park DH. Exploring the use of socially assistive robots among socially isolated Korean American older adults. J Appl Gerontol 2024 Sep;43(9):1295-1304. [doi: 10.1177/07334648241236081] [Medline: 38410030]
57. Cahill J, McLoughlin S, Wetherall S. The design of new technology supporting wellbeing, independence and social participation, for older adults domiciled in residential homes and/or assisted living communities. Technologies (Basel) 2018;6(1):18. [doi: 10.3390/technologies6010018]
58. Wang K, Chen XS, Gu D, Smith BD, Dong Y, Peet JZ. Examining first- and second-level digital divide at the intersection of race/ethnicity, gender, and socioeconomic status: an analysis of the National Health and Aging Trends Study. Gerontologist 2024 Sep 1;64(9):gnae079. [doi: 10.1093/geront/gnae079] [Medline: 38898816]
59. Czaja SJ, Charness N, Fisk AD, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352. [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
60. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. J Health Soc Behav 2010;51 Suppl(1_suppl):S28-S40. [doi: 10.1177/0022146510383498] [Medline: 20943581]
61. Wang X, Zhang R, Wang Z, Li T. How does digital competence preserve university students' psychological well-being during the pandemic? An investigation from self-determined theory. Front Psychol 2021;12:1252. [doi: 10.3389/fpsyg.2021.652594]
62. Kor PPK, Leung AYM, Parial LL, et al. Are people with chronic diseases satisfied with the online health information related to COVID-19 during the pandemic? J Nurs Scholarsh 2021 Jan;53(1):75-86. [doi: 10.1111/jnu.12616] [Medline: 33316121]
63. DesChâtelets JR, Khowaja AR, Mechelse K, Koning H, Ventresca D. Exploring the access and use of social technologies by older adults in support of their mental health during the COVID-19 pandemic: a rapid review. Can J Aging 2023 Dec;42(4):728-743. [doi: 10.1017/S0714980823000259] [Medline: 37727886]
64. Sixsmith A, Horst BR, Simeonov D, Mihailidis A. Older people's use of digital technology during the COVID-19 pandemic. Bull Sci Technol Soc 2022 Jun;42(1-2):19-24. [doi: 10.1177/02704676221094731]
65. Lu S, Chui C, Lum T. Promoting social capital for healthy aging: towards an integrative framework. Gerontologist 2023 May 9;63(4):628-636. [doi: 10.1093/geront/gnac062] [Medline: 35537375]

Abbreviations

JBI: Joanna Briggs Institute
OR: odds ratio
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by PH Chau; submitted 30.Sep.2025; peer-reviewed by A Hartanto, VQ Wang; revised version received 19.Feb.2026; accepted 19.Feb.2026; published 02.Apr.2026.

Please cite as:
Amoah PA, Ameyaw EK, Bhowmik MK, Adotey PN, Donkor A
Digital Resources and Social Connectedness Among Ethnic Minority Older Adults: Systematic Review and Meta-Analysis
JMIR Aging 2026;9:e84962
URL: https://aging.jmir.org/2026/1/e84962
doi:10.2196/84962

© Padmore Adusei Amoah, Edward Kwabena Ameyaw, Miron Kumar Bhowmik, Prince Nyansah Adotey, Andrew Donkor. Originally published in JMIR Aging (https://aging.jmir.org), 2.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Humanoid Robot–Assisted Support for Health Care in Older Adults: Systematic Scoping Review

Lei Cui[1,2], PhD; Yufei Li[1], PhD; Xinyao Yang[1], PhD; Xue Liu[1], MSc; Like Zhang[1], MSc; Lili Hou[3], PhD

[1]School of Nursing, Shanghai Jiao Tong University, Shanghai, China

[2]Pancreatic Center, The First Affiliated Hospital with Nanjing Medical University, Nanjing, China

[3]Department of Nursing, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, No. 639, Zhizaoju Road, Huangpu District, Shanghai, China

Corresponding Author:
Lili Hou, PhD
Department of Nursing, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, No. 639, Zhizaoju Road, Huangpu District, Shanghai, China

Abstract

Background: Humanoid robots offer a promising solution to the growing burden of care for older adults. However, existing evidence on their applications for general aging populations remains fragmented and lacks systematic synthesis.

Objective: This scoping review aimed to examine the literature on humanoid robot–assisted support for health care in older adults and identify gaps in the literature to guide future research.

Methods: The methodological framework by Arksey and O'Malley was used to conduct this scoping review. We conducted a comprehensive search in 8 databases, including IEEE Xplore Digital Library, CINAHL, Cochrane Library, EMBASE, PubMed, Scopus, Web of Science, and OpenGrey Repository, covering literature published up to April 30, 2025. The reference lists of key texts were examined, and citation chaining was conducted. Two independent reviewers examined all full articles for fitness with the eligibility criteria and extracted data elements. The study findings were then summarized, coded, and analyzed using the PAGER (Patterns, Advances, Gaps, Evidence for practice, and Research recommendations) framework.

Results: A total of 32,477 articles were retrieved, 59 of which were included in this review. The majority (49/59, 83%) were conducted in real-world settings. Methodologically, 34 studies (34/59, 58%) had small sample sizes (n≤25), with study designs comprising 26 quantitative (26/59, 44%), 22 mixed method (22/59, 37%), and 11 qualitative (11/59, 19%) approaches. Participant characteristics revealed female predominance (>50%) in 32 studies (32/59, 54%), while 27 studies (27/59, 46%) included participants with cognitive impairment. Through PAGER framework analysis, we identified 4 key patterns: (1) effects, perceptions, and experiences of humanoid robots; (2) preferences, expectations, and facilitators for humanoid robots; (3) implementation barriers and challenges; and (4) determinants of user experiences and outcomes.

Conclusions: This scoping review demonstrates the promising yet methodologically constrained potential of humanoid robots in health care for older adults while highlighting key challenges in their practical implementation. Successful integration will require addressing technical limitations, user acceptance barriers, and systemic adoption challenges.

(JMIR Aging 2026;9:e83849) doi:10.2196/83849

KEYWORDS

humanoid robots; older; older adults; robotics; scoping review; socially assistive robots

Introduction

Population aging has intensified worldwide, constituting a critical societal challenge. According to the United Nations World Population Prospects 2024 report, the proportion of individuals aged ≥65 years doubled from 5% in 1960 to 10% in 2024, and projections suggest this figure will surpass one-third of the population in the European Union, China, and Japan by 2050 [1]. A study tracking staffing trends since the onset of the COVID-19 pandemic found that 99% of nursing homes and 96% of assisted living facilities in the United States continued to face significant staffing shortages [2]. This study further indicated that 58% of nursing homes were limiting admissions due to lack of staff, with 78% expressing concern that workforce shortages might force them to close [2]. Despite measurable recovery observed in a 2024 follow-up study, employment levels in skilled nursing facilities remained 0.1% below projected benchmarks [3]. Concurrently, informal care systems face growing pressures, with Europe's care burden projected to increase 49.7% by 2050 [4], while caregivers endure physical, emotional, and financial strain [5]. The macroeconomic impact is equally concerning, with each 1% increase in care dependency reducing gross domestic product growth by 0.083%

XSL•FO
RenderX

through workforce attrition [6]. These intersecting crises in both professional and family-based care systems highlight the critical need for innovative support solutions in older adult care.

Humanoid robots present a promising avenue to address these systemic challenges. Since its inception in the 1970s, humanoid robotics has undergone a remarkable evolution, transitioning from rudimentary mechanical mimics to sophisticated platforms capable of complex social interactions [7]. Their anthropomorphic architecture—featuring human-like torsos, limbs, and movement patterns—confers unique advantages in care contexts for older adults. This biomimetic design enhances environmental navigation in human-centric spaces, fosters natural communication through familiar gestural cues, and improves social acceptance compared with nonanthropomorphic alternatives [7]. These attributes make humanoid robots a compelling platform for exploring the integration of physical and psychosocial support in older adult care. Empirical studies have begun to investigate these applications across multiple care domains. For example, the Pepper platform has been evaluated in the domains of cognitive training [8], object transfer [9], and social communication [10]. Similarly, the Nao robot has been studied in contexts such as health education delivery [11] and social interaction therapies [12] for older adults.

The field of robotics and gerontechnology has seen significant growth in research focused on developing care robots to support older adults and their caregivers [13]. Although existing reviews have examined robotic applications in health care, they present several important limitations that warrant addressing to advance the field. First, by conflating anthropomorphic and nonanthropomorphic robotic systems without distinguishing their distinct capabilities [14-16], prior syntheses obscure the unique potential of humanoid forms for fostering natural social engagement—a critical dimension in older adult care. Second, many reviews incorporate heterogeneous populations spanning multiple age groups and health conditions [15], thereby diluting findings that are specifically relevant to the physiological, cognitive, and psychosocial needs of the older adult population. Third, existing analyses often focus narrowly on isolated aspects, such as single-function applications, implementation challenges, or ethical issues in isolation [14,16-18], while neglecting comprehensive, multidimensional assessments of how multifunctional robotic systems operate within integrated older adult care environments. Although a previous scoping review did specifically investigate humanoid robots for older adult care, it was methodologically limited by its narrow search of only 2 databases and inclusion of literature only up to 2019 [19]. These limitations mean the review cannot account for the rapid technological advances and expanded evidence base in humanoid robotics over recent years. Consequently, the current evidence regarding humanoid robot applications for general populations of older adults remains fragmented and lacks an up-to-date, comprehensive synthesis to inform both research and practice.

Given this landscape, we used the scoping review framework of Arksey and O'Malley [20] to systematically map this emerging field. This approach optimally balances methodological rigor with the flexibility needed to (1) characterize current humanoid robot implementations and their care applications and (2) identify critical knowledge gaps requiring prioritized investigation. Our review thus provides both a comprehensive evidence synthesis and a roadmap for future research in this transformative domain.

Methods

Framework

This scoping review used a 5-stage methodological framework proposed by Arksey and O'Malley [20]. Prior to commencement, the study protocol was prospectively registered on the Open Science Framework on May 10, 2025, and is publicly accessible (registration DOI: 10.17605/OSF.IO/P2C8J) [21]. The review adhered strictly to this preregistered protocol without deviation. The conduct and reporting of this review strictly adhered to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines [22]. The completed PRISMA-ScR checklist is available in Checklist 1.

Identifying the Research Question

This scoping review systematically examined the current evidence regarding the development and implementation of humanoid robots in health care for older adults. To comprehensively map the existing literature, we formulated the following broad yet focused research questions: (1) What types of humanoid robots have been implemented and what specific interventions have been delivered to older adult populations? (2) What is the reported feasibility, preliminary efficacy, and user experience (from older adults and caregivers/staff) associated with these interventions? (3) What critical gaps exist in the current literature that warrant focused attention in future research?

Identifying Relevant Studies

The search strategy was designed to identify both published and unpublished sources, encompassing scientific and gray literature such as conference proceedings. This approach helped mitigate publication bias and capture emerging research, especially in the rapidly evolving field of robotics, where significant findings are often first presented outside traditional journals. An initial exploratory search was conducted in Web of Science to identify relevant articles on the topic. Key terms from the titles, abstracts, and index terms of these articles were extracted to inform the development of a comprehensive search strategy. The search combined robotics-related terms (eg, “humanoid robot,” “social robot,” “robot assist,” “human-robot interaction”) with terms referring to older adults (eg, “aged,” “older,” “elderly,” “geriatric”). A systematic 3-step search approach was used, covering literature published up to April 30, 2025. First, 7 major electronic databases including IEEE Xplore Digital Library, CINAHL, Cochrane Library, EMBASE, PubMed, Scopus, and Web of Science, were searched using predefined keywords and index terms (see Table S1 in Multimedia Appendix 1). Next, gray literature was sourced through the OpenGrey Repository. Finally, reference lists of key publications were examined, and citation chaining was performed to ensure thorough coverage of the relevant literature.

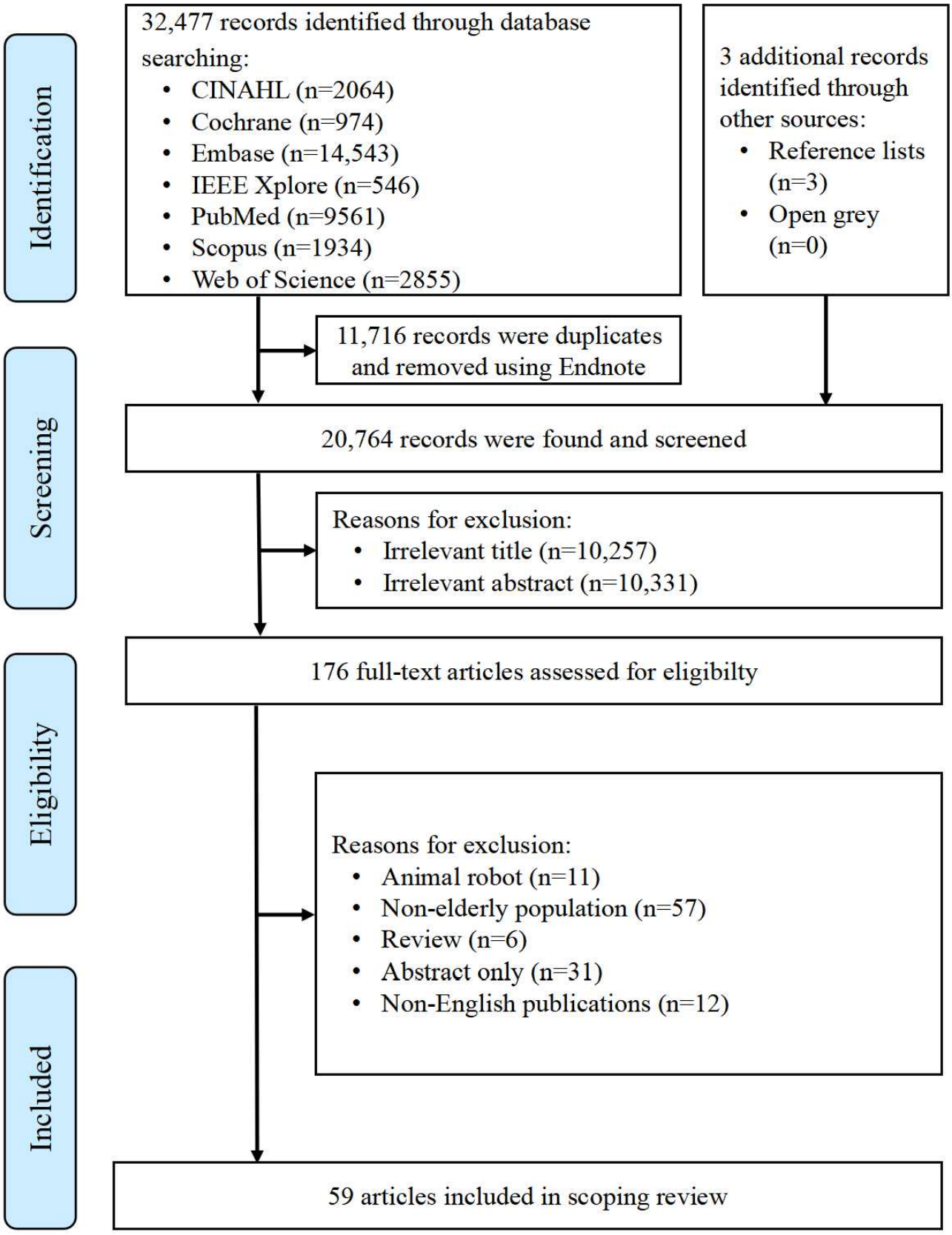
XSL•FO
RenderX

Study Selection

The review included original studies in which older adults (defined as individuals aged ≥60 years, aligning with the classification from the World Health Organization [WHO]) received any form of health care support through direct interaction with a physically embodied humanoid robot. We defined a humanoid robot as an autonomous or semi-autonomous machine with human-like morphology (eg, head, torso, arms; bipedal or wheeled mobility) and capable of performing social or physical tasks through human-like modalities such as speech, gesture, or facial expression. Studies that examined perspectives of other stakeholders (eg, care staff, family members, or health care professionals) within the same care context were also included, as these insights contribute to a holistic understanding of implementation and acceptability. For this review, health care is defined according to Merriam-Webster's Dictionary as "efforts made to maintain, restore, or promote someone's physical, mental, or emotional well-being especially when performed by trained and licensed professionals." Our focus within this broad domain is on applications in both clinical settings (eg, hospitals) and long-term social care settings (eg, residential institutions and private homes). All study designs (eg, experimental, observational, qualitative) were considered, and both published and unpublished English-language literature was included without date restrictions. Studies were excluded if they (1) primarily focused on robotic hardware or software development without evaluating a health care application; (2) used nonembodied or simulated representations of robots (eg, video-based or computer-simulated interactions); (3) focused solely on fundamental human-robot interaction mechanisms (eg, laboratory studies of perception or dialogue systems without a health care context or outcome), as these did not address our objective of mapping evidence related to health care support for older adults; or (4) lacked available full text. The study followed a rigorous 4-phase selection process in accordance with PRISMA-ScR guidelines (Figure 1). After initial duplicate removal using EndNote 21 software, 2 independent reviewers conducted title and abstract screening followed by full-text evaluation. Inter-rater agreement was assessed using Cohen kappa, with a value >0.6 considered acceptable. The reviewers then reconciled their findings to ensure comprehensive coverage, with any discrepancies resolved through arbitration by a third reviewer.

XSL•FO
RenderX

Figure 1. Flow diagram of the study retrieval and selection progress.



Charting the Data

We developed a standardized data extraction form aligned with the study objectives to systematically chart key information from included articles. Prior to full implementation, 2 independent reviewers piloted the data collection instrument using the first 5 eligible studies, refining the form through iterative discussion to enhance reliability. Extracted data elements encompassed (1) bibliographic details (authors, publication year, country of origin), (2) study characteristics (research setting, study design, study aims, sampling strategy,

XSL•FO
RenderX

participant demographics), (3) outcome measures, and (4) significant findings. For humanoid robot–assisted health care, we specifically documented (1) core intervention components, (2) program content, and (3) implementation parameters (frequency and duration).

Collating, Summarizing, and Reporting Results

Two independent reviewers conducted parallel analyses to ensure objectivity and minimize bias. Following thorough discussion and reconciliation of findings, the complete research team reviewed and validated the results. Using the PAGER (Patterns, Advances, Gaps, Evidence for practice, and Research recommendations) framework [23], we systematically identified 4 key patterns in the literature, along with their corresponding advances, gaps, practical implications, and future research directions. The PAGER framework is a structured methodological tool specifically designed for analyzing and reporting findings in scoping reviews, enhancing both the clarity and utility of the review output [23]. The acronym encompasses 5 interrelated analytical domains [23]: Patterns (key recurring themes emerging from the literature), Advances (notable developments within the recent evidence base), Gaps (critical limitations or underexplored areas), Evidence for practice (actionable insights for stakeholders), and Research recommendations (concrete suggestions for future research). A foundational step in applying the framework involves creating a patterning chart—a thematic table derived from inductive analysis of included studies—which visually organizes key findings and systematically guides the identification of advances, gaps, and resulting recommendations for research and practice [23].

Results

Study Selection and Characteristics

The study selection process, as depicted in the PRISMA-ScR flow diagram (Figure 1), commenced with the identification of 32,477 records from 8 electronic sources, with 3 additional records obtained from OpenGrey and manual reference searches. Following the removal of 11,716 duplicates using EndNote 21, we conducted initial screening of 20,764 records. Title and abstract review led to the exclusion of 20,588 records that did not satisfy the inclusion criteria. The analysis yielded a Cohen kappa value of 0.88 for inter-rater agreement. Subsequently, 176 full-text articles underwent comprehensive eligibility assessment, culminating in the inclusion of 59 studies for final analysis (see Table S2 in Multimedia Appendix 1).

Table 1 and Tables S3 and S4 in Multimedia Appendix 1 present the key characteristics of the included studies. Publication years spanned from 2012 onward, showing distinct temporal clustering with 10 studies (10/59, 17%) published before 2015, 18 studies (18/59, 31%) published between 2016 and 2020, and 31 studies (31/59, 53%) published after 2020. The publications originated from research teams across multiple regions: Europe (n=20), Asia (n=15), North America (n=7), and Oceania (n=4), with an additional 13 studies representing multinational collaborations. The majority (49/59, 83%) were conducted in real-world settings—defined for this review as environments where care is routinely provided with minimal researcher intervention, as opposed to controlled laboratory or research center environments—while laboratory environments accounted for 10% (6/59), and 5% (3/59) did not specify their location. Across the included studies, participants primarily comprised older adults (59/59, 100%), with a subset also involving informal caregivers (5/59, 9%), formal caregivers (14/59, 24%) and, in 3 studies, individuals categorized as “others.” Methodologically, 34 studies (34/59, 58%) had small sample sizes (n≤25). Among the 59 studies, 26 (44%) adopted a quantitative design, 22 (37%) used a mixed methods approach, and 11 (19%) used qualitative methodologies. Within the quantitative subset, designs varied considerably, encompassing 6 observational (cross-sectional) studies, 4 randomized controlled trials, 5 single-case/within-subjects experimental designs, 7 quasi-experimental (nonrandomized, group-based) studies, and 4 pre-post studies. Participant characteristics revealed female predominance (>50%) in 32 studies (32/59, 54%), while 16 studies (16/59, 27%) omitted gender data. Notably, 27 studies (27/59, 46%) incorporated participants with cognitive impairment.

XSL•FO
RenderX

Table . Characteristics of included studies (n=59).

Characteristics	Studies, n (%)	Study references
Publication year		
2021 - 2025	31 (53)	[8-12,24-49]
2016 - 2020	18 (31)	[50-67]
2012 - 2015	10 (17)	[68-77]
Country		
Europe		
Italy	6 (10)	[8,24,30,33,34,48]
France	5 (9)	[27,40,43,44,73]
Germany	2 (3)	[10,69]
Poland	5 (9)	[32,35,36,45,60]
Finland	1 (2)	[51]
Spain	1 (2)	[68]
Asia		
Japan	8 (14)	[28,42,56,57,63,64,67,77]
China	4 (7)	[37,52-54]
South Korea	2 (3)	[11,47]
Israel	1 (2)	[62]
North America		
United States	5 (9)	[12,39,58,65,75]
Canada	2 (3)	[9,72]
Oceania		
New Zealand	1 (2)	[70]
Australia	3 (5)	[66,74,76]
Multinational collaboration	13 (22)	[25,26,29,31,38,40,46,49,50,55,59,61,71]
Study design		
Quantitative study	26 (44)	[illegible]
Observational designs		
Cross-sectional study	6 (23)	[32,35,36,45,63,72]
Experimental designs		
Randomized controlled trial	4 (15)	[47,52,54,77]
Single-case experimental design/within-subjects design	5 (19)	[41,42,57,64,75]
Quasi-experimental studies		
Quasi-experimental study (group, nonrandomized)	7 (27)	[9,11,27,30,48,60,68]
Pre-post study	4 (15)	[12,28,33,59]
Qualitative study	11 (19)	[8,25,29,31,37,49,51,53,55,56,62]
Mixed method design	22 (37)	[illegible]
Setting		
Day care center	1 (2)	[8]
Clinic	3 (5)	[24,34,48]
Home	6 (10)	[25,31,37,69,70,77]
Community	5 (9)	[29,38,39,41,47]
Hospital	7 (12)	[33,43,44,50,53,60,75]

XSL•FO
RenderX

Characteristics	Studies, n (%)	Study references
Long-term care facility	15 (25)	[9,12,26-28,40,46,54,56-58,63,65,66,76]
University	1 (2)	[52]
Multiple sites	12 (20)	[32,35,36,42,45,49,51,55,59,64,68,74]
Laboratory	6 (10)	[10,11,61,62,71,73]
Not specified	3 (5)	[30,67,72]
Key robot functions in health care		
Social interaction	23 (39)	[12,26,28,32,34,39,45,51-53,55,58,59,61,63,65,66,70-74,76]
Cognitive training	15 (25)	[8,24,34,36,37,41,46-48,55,59,62,68,72,73]
Physical activity facilitation	11 (19)	[9,25,26,32,33,38,49,52,56,57,64]
Health care support	10 (17)	[29,32,37,38,60,65,70,74-76]
Daily life support	8 (14)	[37,38,43,45,60,61,66,69]
Companionship	8 (14)	[29,37-39,54,65,69,71]
Health monitoring	4 (7)	[32,59,69,71]
Communication	4 (7)	[10,49,63,77]
Memory training	3 (5)	[27,30,50]
Health education	3 (5)	[11,44,67]
Entertainment	3 (5)	[43,66,70]
Health coaching	2 (3)	[29,31]
Rehabilitation	2 (3)	[51,75]
Leisure facilitation	1 (2)	[42]
Handover	1 (2)	[40]
Emergency assistance	1 (2)	[71]

Description of Humanoid Robot–Assisted Health Care for Older Adults

See Table S5 in Multimedia Appendix 1 for details on humanoid robot–assisted health care interventions. Across the included studies, 25 distinct humanoid robots were used, with Pepper being the most frequently used (n=15), followed by Nao (n=9), TIAGo (n=4), Kabochan (n=3), and Matilda (n=3). The remaining robots were each used in 1 or 2 studies. As detailed in Table 1, the interventions covered 16 key functional domains, with the number of studies per domain ranging from 1 to 23. The most prevalent functions were social interaction (23/59, 39%), cognitive training (15/59, 25%), and physical activity facilitation (11/59, 19%). Among the 59 studies analyzed, 4 did not report intervention duration. Of the remaining 55 studies (55/59, 93%), the majority (30/55, 55%) lasted 1 week or less; 11% (6/55) lasted between 1 week and 4 weeks; and 35% (19/55) extended beyond 4 weeks, including 4 studies (4/55, 7%) that continued for more than 24 weeks.

Patterns of Research Findings

We identified 4 key PAGER patterns of research findings (see Table 2; also see Multimedia Appendix 2 for a detailed analysis of all included studies).

XSL•FO
RenderX

Table . PAGER (Patterns, Advances, Gaps, Evidence for practice, and Research recommendations) framework.

Pattern	Advances	Gaps	Evidence for practice	Research recommendations
Effects, perceptions, and experience of humanoid robots	Preliminary studies indicate the potential of humanoid robots to be effective in care for older adults, with many reporting positive user acceptance and experiences.	Lack of randomized controlled trials, short-duration humanoid robot applications, and small sample sizes and imbalanced population characteristics	Emerging studies indicate that humanoid robots may have the potential to improve cognitive, physical, and psychosocial health in aging populations.	Prioritize conducting rigorously designed randomized controlled trials with larger, more balanced samples to evaluate the clinical effectiveness of longer-term humanoid robot interventions while simultaneously assessing older users' perceptions and experiences.
Preference, expectations, and facilitators for humanoid robots	Heterogeneous and individualized preferences have been observed among older adults regarding humanoid robots' appearance, functionality, and interaction modalities.	Small sample sizes and imbalanced population characteristics	Available findings suggest that involving older adults in the design process to tailor a robot's appearance and functions to individual preferences could enhance its acceptance and utility in care settings.	Use representative sampling strategies and involve cross-disciplinary teams to develop and evaluate personalized robotic interventions, with specific attention to distinguishing between initial novelty effects and sustained engagement over time.
Implementation barriers and challenges	The adoption of humanoid robots in care for older adults may face multifaceted barriers across technological, user-specific, and contextual dimensions.	Lack of multistakeholder perspectives	To facilitate implementation, practice should focus on developing robots with senior-friendly interfaces and complementing their introduction with tailored digital literacy programs for older users and their caregivers.	Adopt a multistakeholder analytical framework, gathering and integrating perspectives from older adults, formal and informal caregivers, health care professionals, policymakers, and technology developers to identify comprehensive implementation strategies.
Determinants of user experiences and outcomes	Available evidence suggests that user experiences and outcomes with humanoid robots may be influenced by user-related factors (eg, demographics, health status, and psychological state) and robot-related characteristics (eg, appearance, personality, and interaction methods).	Small sample sizes and imbalanced population characteristics as well as narrow factor consideration (eg, omitting environmental and contextual variables)	Emerging studies point to the importance of considering user experiences and outcomes as key factors in shaping effective human-robot interactions.	Expand sample sizes and systematically examine a broader range of potential determinants, including environmental (eg, home vs institutional setting) and contextual factors, to better understand what shapes user experience and outcomes.

Effects, Perceptions, and Experience With Humanoid Robots

Research indicates that humanoid robot interventions can yield effects across psychosocial, cognitive, functional, and behavioral domains. A frequently reported psychosocial benefit is enhanced social interaction [8,11,12,25,26,29,37,38,42,43,46,48,51-53,55,58-60,63,71], with associated reductions in loneliness [12,29,37,45]. Improvements in mood and emotional states [8,12,24,25,28,39,42,46,55,59,65,77], alongside decreases in stress, depression, and anxiety [8,12,24,30,39,46,47,58,59,65], have also been observed, although not all studies report significant changes in these areas [58]. In cognitive and functional domains, evidence suggests potential for cognitive stimulation [8,24,29,30,41,46,47,50,55,68,74-77], enhanced task performance [24,41,48,64], and improved limb function [33,56,75]. Regarding health behaviors, some interventions show promise for promoting medication adherence [29,67,70], physical activity [29,31,33,51,56,76], better dietary habits [76], and sleep quality [77]. It is noteworthy that, in specific task-oriented contexts, robotic performance has been perceived as inferior to human care [27]. The sustainability of these positive effects is a critical concern, with evidence pointing to potential attenuation over time [38] or a lack of significant long-term improvement in outcomes such as core dementia symptoms, depression, and quality of life [38,54,58,59].

Perceptions of humanoid robots among older adults and formal oor informal caregivers are heterogeneous, encompassing both positive dimensions and reservations. A substantial body of literature reports favorable perceptions from older adults across multiple dimensions, including acceptability [8,25,26,29,30,38,44-46,49,52,53,60,61,64-66,69-74,76], usability [24,30,38,43,44,50,52,60,69,71], trust [31,64,71,73], satisfaction [11,33,40,43,61], and active or positive attitudes [8,10,24,35,37,43,46,51,55,57,61,65,66,69,72,75,76]. In practice, this manifests as increased user engagement and motivation [8,12,24-26,33,37,39,44,46,48,50,55,58,66,74,76,77] and strengthened perceived social support [59] in some contexts. However, user acceptance is not universal. Some older adults

XSL•FO
RenderX

perceive robots as "too technical" or "lacking a human touch" [51], a challenge exacerbated for those with sensory impairments [51]. Caregiver perceptions are similarly mixed. Although some studies indicated a potential reduction in caregiver workload [51,54,60,76], others report an increased burden due to the need to assist with the technology [46]. Formal caregivers frequently cite practical barriers, including the additional time required for operation, difficulties in workflow integration, and, in some cases, negative attitudes toward the technology's value or appropriateness [51]. Cost-effectiveness is another point of skepticism, raised both by older adults [25] and implicitly within the context of practical implementation challenges.

The lived experience of interacting with these robots underscores several critical considerations. Safety and adverse events, though infrequent, warrant attention. These range from user-expressed safety concerns [57] to documented incidents such as medication errors [38] and adverse psychological reactions leading to study withdrawal [54]. Ethical concerns prominently feature issues of autonomy and privacy, with a noted divergence in emphasis between older adults (prioritizing user control) and caregivers (exercising more caution) [35]. Furthermore, equity and access emerge as significant issues, exemplified by cases where robot design (eg, language support) inadvertently created barriers for participants, highlighting risks of exacerbating the digital divide [46].

Preferences, Expectations, and Facilitators for Humanoid Robots

This synthesis reveals a clear pattern in older adults' preferences for humanoid robots. Older adults show marked preference for humanoid robots' health management functions, especially medication reminders and daily routine notifications [25,43]. Assistive functions like safety monitoring and health reminders consistently receive higher utility ratings than purely social features [45]. At the same time, entertainment-oriented functions such as music playback and memory recall features exhibit high engagement frequency, with studies showing these applications significantly improve mood states and facilitate positive reminiscence [55]. These core functionalities are typically complemented by basic conversational interactions including weather inquiries and casual chats in real-world usage scenarios [37,65].

Beyond functionality, user expectations shape perception and interaction. Users perceive the empathetic version of Ryan robot as more attractive, friendly, and compassionate than its nonempathetic counterpart [39]. Notably, multisensory integration—combining visual, auditory, and tactile stimuli—elicited significantly stronger happiness responses than unimodal (visual-only) or bimodal (visual-auditory) stimulation approaches [28]. Some evidence suggests that older adults may prefer in-person human-robot interaction over remote formats [9]. In studies of walking assistance, participants have been observed to position the robot more as a "navigational guide" rather than a "walking companion" [64]. Regarding humanoid robotic appearance expectations, research indicates older adults expect larger-sized robots [25] but typically reject those taller than themselves [43]. Their appearance expectations encompass a "smiling" and "handsome" design with fluid movements [43], along with friendly aesthetics, strong interactivity, simple operation modes, and integrated verbal-physical interactions [44].

Research documents older adults' functional requests for humanoid robotic systems, which serve as key facilitators for adoption. These emphasize improved interaction quality (enhanced speech clarity, real-time responses, and personality adaptation) [25,31,37], personalized health features (tailored exercise programs and accurate reminders) [38], expanded utility (medication management and emergency services) [37,43], and technical enhancements (advanced voice recognition and adaptive mobility) [55,57]. Although users show interest in customizable cognitive activities and varied applications [65,72], they consistently express reservations about privacy-sensitive functions such as bathing assistance [61].

Implementation Barriers and Challenges

From the machine or technical perspective, limitations manifest across several operational domains. Hardware constraints significantly impair functional capacity, notably in fine motor manipulation [40] and ambulatory assistance [56]. System reliability issues are characterized by intermittent connectivity [43] and unanticipated software termination [26], while interaction protocols demonstrate inadequate natural language processing capabilities [27,49,62]. Sensory apparatuses exhibit suboptimal performance in acoustically challenging environments [55] and human motion detection [43]. User interface implementations frequently present usability barriers, including unintuitive volume modulation controls and instances of erroneous information dissemination (eg, inaccurate meteorological alerts) or unintended nocturnal activation [38]. Crucially, current systems predominantly operate under supervised autonomy paradigms [12], where robots execute predetermined tasks but remain reliant on human oversight and intervention for initiation, safety, and context management rather than achieving genuine operational independence [49].

From the human or user perspective, adoption barriers emerged across 3 principal dimensions. Physical limitations include difficulties engaging with activation protocols [31] and potential safety hazards from unpredicted robotic motions [49]. Psychological factors encompass both initial technophobia [46] and entrenched self-perceptions of digital incompetence [73], necessitating extended skill acquisition periods [47]. Sociocultural concerns involve professional displacement anxieties among care staff [51], diminished interpersonal engagement [73], and surveillance-related ethical dilemmas [49]. Most significantly, prolonged use of these systems may lead to a decline in users' own abilities due to lack of practice and could also undermine the human essence of caregiving [73].

Implementation challenges intersect both perspectives, as even technically proficient systems demand substantial care protocol modifications and extensive personnel training [51], with economic viability assessments challenging the utility of specific design implementations [25].

Determinants of User Experiences and Outcomes

Older adults' experiences and outcomes with humanoid robots are influenced by multiple factors, with significant individual

XSL•FO
RenderX

variations reported across studies. Available findings on demographic characteristics are mixed. For instance, one study suggested that men might be more inclined than women to perceive robots as companions [35], while another found that women exhibited greater cognitive improvement in a specific robot-assisted cognitive training program [47]; however, other research observed no significant gender differences in well-being outcomes [42]. Similarly, although age was not found to significantly affect older adults' attitudes toward robots or their well-being in one study [42], another reported more pronounced cognitive benefits from training among participants aged 75 years and older [47]. Regarding education level, findings present a dual pattern. One study reported that individuals with higher education levels exhibit lower acceptance of robots' social functions [36], whereas a different study found that those with lower education levels showed more significant cognitive improvement in a training context [47]. In terms of health status, some evidence indicates that older adults with less independence in daily living may display greater reliance on robotic assistance [35], and those with higher dependency levels have been associated with moderately reduced well-being scores in one study [42]. Psychological states appear to influence interaction tendencies. For example, one study observed that individuals exhibiting agitation or psychotic symptoms showed stronger engagement willingness than those with depressive or apathetic symptoms [54]; the same study also noted that hearing impairment and sleep deprivation were linked to significantly impaired interaction outcomes [54]. At the interaction design level, findings from a specific study indicated that multimodal (voice-gesture) interaction was most effective at enhancing emotional connection and familiarity [11]. In another study, introverted robot personalities were perceived as more natural in interaction, whereas extroverted robots were considered more entertaining but slightly intrusive [34]. Furthermore, key robot perception dimensions—such as anthropomorphism, animacy, likability, and perceived intelligence—have been found to show significant correlations with functional performance in one study [36].

Discussion

This scoping review provides a systematic examination of humanoid robot–assisted health care support for older adults, analyzing 59 studies that document 25 distinct humanoid robot models currently implemented in both real-world and laboratory settings. Using the PAGER framework as our analytical lens, we identified and organized the extant literature into 4 conceptually distinct patterns. These emergent patterns are detailed in the following sections.

Effects, Perceptions, and Experience With Humanoid Robots

This scoping review aligns with previous findings [19] by demonstrating the multifaceted applications of humanoid robots across various domains of older adult care, encompassing social interaction, cognitive stimulation, physical activity facilitation, and activities of daily living support. Evidence indicates these robotic solutions can successfully promote social interaction, deliver cognitive enrichment, and enhance motor function in older populations. Notably, older adults have demonstrated favorable reception of humanoid robots, expressing positive evaluations of their usability and satisfaction levels while exhibiting constructive attitudes toward this technology. Together, these findings underscore the considerable promise of humanoid robots for innovative implementations in older adult care. Nevertheless, the methodological heterogeneity and specific limitations of the included studies restrict the strength and generalizability of our inferences. In line with existing literature [14], a recurrent methodological limitation across the reviewed literature is the prevalent use of studies with small sample sizes. This issue is frequently compounded by significant gender imbalances, typically skewed toward female participants, and a predominant research focus on older adults with cognitive impairments, potentially limiting the generalizability of findings to the wider, more heterogeneous aging population. Methodologically, the field shows a striking lack of randomized controlled trials—the gold standard for clinical evidence—compounded by typically short intervention durations that fail to capture long-term effects. Additionally, approximately 10% of the included studies were conducted in laboratory settings, which may further restrict the ecological validity and real-world applicability of the reported outcomes. In terms of functionality, current humanoid robot applications remain narrowly focused on cognitive training and social interaction, as confirmed by both our findings and prior research [78]. This limited scope stands in stark contrast to the broader spectrum of care needs documented in the literature [79], which includes essential daily living supports like fall prevention, feeding assistance, toileting, bathing, and position changing. To advance the field, future work should prioritize the development of more versatile robotic systems while conducting rigorous, large-scale multicenter trials with extended observation periods to fully assess the potential of humanoid robots in older adult care settings.

Preference, Expectations, and Facilitators for Humanoid Robots

This scoping review elucidates older adults' preferences and expectations concerning humanoid robots, encompassing functionality, design features, and physical appearance, while also capturing their developmental aspirations for this technology. Our analysis reveals that seniors exhibit complex, multidimensional expectations that are not only highly personalized but also consistently oriented toward robotic systems with expanded capabilities, enhanced stability, sophisticated intelligence, and greater autonomy. These combined demands, which encompass both personalization and high-performance functionality, directly translate into a core technical requirement: Assistive robots must be highly adaptable to diverse home environments and individualized user needs. Achieving this necessary level of adaptability in a cost-effective manner represents a significant technical hurdle requiring advances in sensors, machine learning algorithms, and robotic hardware [80]. However, the current evidence base presents notable methodological constraints. Many included studies have similar methodological limitations as noted in the previous section, particularly regarding sample representativeness. Building upon nursing science's established tradition of

XSL•FO
RenderX

person-centered care [81,82], we propose an integrated development framework that synergizes technological innovation with individualized implementation strategies. This approach necessitates multidisciplinary collaboration among gerontologists, engineers, designers, and health care professionals to advance humanoid robot capabilities, particularly in modular customization, natural interaction algorithms, and reliable multicontext functionality [83]. Parallel to these technical advancements, comprehensive predeployment assessments incorporating detailed geriatric evaluations and user profiling will be essential to ensure optimal alignment between robotic solutions and individual user requirements, thereby maximizing both clinical efficacy and resource efficiency.

Implementation Barriers and Challenges

Similar to prior studies [84,85], we demonstrate that widespread adoption of humanoid robots faces substantial barriers from technical, human, systemic, and sociocultural perspectives. A critical gap in the included literature is its overemphasis on end user experiences while neglecting other essential stakeholders including family caregivers, health care providers, policymakers, and technology developers [18]. This narrow perspective creates an incomplete framework for implementation, which must account for seniors' needs alongside caregiver workflows, institutional policies, and financial sustainability. Furthermore, the implementation context is profoundly shaped by culture-related factors that are often underexamined. These include deeply held social norms regarding the appropriateness of delegating intimate care tasks to machines, varying ethical frameworks across societies for balancing technological efficiency with the preservation of human dignity in care, and differing levels of societal trust in automation. To address these limitations, future research should use a comprehensive approach that engages all relevant stakeholders and explicitly investigates the cultural contexts of deployment settings. For older adult users specifically, implementation strategies should use a phased approach that includes assessing initial perceptions, clarifying concepts, addressing concerns, and demonstrating benefits to enhance acceptance [86]. Structured communication channels and collaborative guideline development can help mitigate risks while fostering stakeholder engagement. From a technical perspective, developing adaptable systems for diverse environments and user needs remains paramount, though cost-effective solutions will require advances in sensors, machine learning, and hardware design. Ethical considerations must go beyond minimum standards, implementing robust data consent protocols that exceed basic International Organization for Standardization requirements [87,88]. Successful implementation further requires addressing affordability to ensure equitable access, developing comprehensive training programs to optimize use, and planning workforce transitions that include retraining opportunities. Culturally appropriate design remains essential for widespread adoption [87]. A successful integration strategy must therefore balance technological innovation with human-centered design and organizational preparedness.

Determinants of User Experiences and Outcomes

Our scoping review reveals that the effectiveness of humanoid robots is determined by both user-related factors (including demographics, health status, and psychological characteristics) and robot-related features (such as appearance, personality traits, and interaction modalities). The studies included in this review exhibit notable methodological constraints, particularly regarding the common issue of limited sample sizes with demographic imbalances and insufficient examination of contextual and environmental variables. Subsequent investigations should prioritize more representative participant cohorts and systematically evaluate the complete spectrum of influential factors to elucidate the key determinants of robot-assisted older adult care. Critically, as emphasized throughout this review, transitioning from standardized implementations to tailored, individualized approaches represents an essential paradigm shift for optimizing humanoid robot integration in older adult health care settings.

Strengths and Limitations

To our knowledge, this represents the most comprehensive and methodologically rigorous scoping review to date examining the applications of humanoid robots in older adult health care. Our study, which was prospectively registered, used the established 5-stage methodological framework by Arksey and O'Malley [20] and adhered to PRISMA-ScR guidelines for transparent reporting. Notwithstanding these strengths, several limitations warrant consideration. First, although our systematic 3-step search strategy identified 59 relevant articles across 7 databases and gray literature sources, we acknowledge the possibility of having missed pertinent studies, particularly those where humanoid robotics applications were not explicitly stated in titles or abstracts. The exclusion of internet searches may have further contributed to this limitation. Second, our English language restriction may have introduced selection bias by excluding potentially relevant non-English publications. Finally, the scope of this review did not encompass certain specialized domains such as robotic surgery or monitoring systems using humanoid platforms.

Implications

This scoping review underscores the considerable promise of humanoid robots in enhancing health care delivery for aging populations, positioning them as a viable solution to mitigate future older adult care challenges. The field has evolved from early locomotion-focused platforms to today's interactive, AI-enhanced systems capable of complex social and assistive functions. Future advancements are expected to encompass greater personalization, multimodal sensing, expanded physical assistance capabilities, and stronger ethical frameworks to support real-world integration [7]. However, the seamless integration of this technology into the daily lives of older adults remains a complex challenge that requires careful consideration. The findings reveal that the needs and expectations of older adults are multifaceted, demanding highly personalized approaches. Successful implementation of humanoid robots in older adult care must extend beyond technological functionality to incorporate the perspectives of all stakeholders—older adults themselves, family members, professional caregivers, health

XSL•FO
RenderX

care staff, institutional leaders, and policymakers. A one-size-fits-all model would be insufficient; instead, adaptable solutions tailored to individual needs are essential to enhance care quality and optimize resource efficiency. However, achieving this personalization at scale necessitates addressing its economic viability. Research and policy must concurrently focus on reducing costs through modular design, scalable software personalization, and innovative funding or reimbursement models to prevent the benefits of personalized robotic care from becoming a privilege of the few. Moreover, the effectiveness of humanoid robots in older adult care is shaped by a dynamic interplay of technological, human, environmental, and ethical factors. Advancing this field requires interdisciplinary collaboration across health sciences, computer science, and engineering, alongside the establishment of standardized frameworks to improve robotic functionality, safety, and autonomy while minimizing risks. Equally critical is the development of older adult–friendly interfaces and targeted training programs to improve digital literacy among older adults. Professional caregivers also need structured instruction to facilitate effective human-robot collaboration in care settings. Simultaneously, ethical concerns—including privacy protection, autonomy preservation, and the establishment of appropriate human-robot boundaries—must be systematically addressed to ensure responsible adoption and sustained trust in robotic care solutions. Although the promise of humanoid robots in older adult care is evident, their long-term viability hinges on addressing these multifaceted challenges through coordinated efforts across technological innovation, user-centered design, workforce training, and ethical governance. Finally, future rigorous randomized controlled trials should prioritize the inclusion of well-defined control groups to precisely isolate the unique effect of humanoid robot interventions. Appropriate control conditions may include, but are not limited to (1) standard care or treatment as usual to establish baseline effectiveness, (2) interventions delivered by human caregivers or companions to benchmark robotic performance against the human touch, and (3) other technological or robotic interfaces (eg, tablet apps, nonhumanoid robots) to distinguish the impact of humanoid form factors and interaction modalities from general technology effects. Such comparisons are essential for determining not just if humanoid robots are effective but also for whom, in what contexts, and relative to what alternatives.

Conclusions

This scoping review systematically mapped the literature on the implementation of humanoid robots in health care settings for older adults, identifying 25 distinct models currently in use. By applying the PAGER framework, the review delineated 4 key thematic patterns within the evidence: (1) effects, perceptions, and experiences; (2) preferences, expectations, and facilitators; (3) implementation barriers and challenges; and (4) determinants of user experiences and outcomes. The available evidence suggests that deployments to date have focused largely on social interaction and cognitive stimulation, with many studies reporting favorable initial user acceptance and experience. Findings also point to individualized user preferences and highlight that adoption is likely influenced by a complex interplay of technical, user-specific, and contextual factors. The analysis underscores persistent evidence gaps—most notably a lack of robust experimental designs, long-term outcome data, and studies conducted in ecologically valid settings with representative populations. Future research should, therefore, prioritize large-scale, real-world trials with diverse older adult populations to evaluate the long-term efficacy and implementation of multifunctional humanoid robots across the full spectrum of care needs. Ultimately, addressing these research gaps will be crucial for overcoming the technical, user-related, and systemic challenges to successful integration.

Funding

This work was supported by the National Natural Science Foundation of China (grant number 72574139).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Index, keywords, and list and characteristics of included studies.
[XLS File, 41 KB - aging_v9i1e83849_app1.xls]

Multimedia Appendix 2
Detailed PAGER (Patterns, Advances, Gaps, Evidence for practice, and Research recommendations) framework analysis of included studies.
[XLS File, 41 KB - aging_v9i1e83849_app2.xls]

Checklist 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.
[DOCX File, 68 KB - aging_v9i1e83849_app3.docx]

References

XSL•FO
RenderX

1. World Population Prospects 2024: Dataset. : United Nations; 2024. URL: https://www.un.org/development/desa/pd/content/World-Population-Prospects-2024 [accessed 2024-10-01]
2. Survey: nearly every US nursing home and assisted living community is facing a workforce crisis. The American Health Care Association and National Center for Assisted Living. 2021 Sep 22. URL: https://tinyurl.com/4ub6xfr3 [accessed 2022-05-23]
3. Roush K. Health care workforce rebounds post–COVID-19. Am J Nurs 2025;125(12):14-14. [doi: 10.1097/AJN.0000000000000208b]
4. Cattaneo A, Vitali A, Regazzoni D, Rizzi C. The burden of informal family caregiving in Europe, 2000-2050: a microsimulation modelling study. Lancet Reg Health Eur 2025 Jun;53:101295. [doi: 10.1016/j.lanepe.2025.101295] [Medline: 40255934]
5. Tu J, Li H, Ye B, Liao J. The trajectory of family caregiving for older adults with dementia: difficulties and challenges. Age Ageing 2022 Dec 5;51(12):afac254. [doi: 10.1093/ageing/afac254] [Medline: 36469090]
6. Tang B, Li Z, Hu S, Xiong J. Economic implications of health care burden for elderly population. INQUIRY 2022;59:469580221121511. [doi: 10.1177/00469580221121511] [Medline: 36062304]
7. Tong Y, Liu H, Zhang Z. Advancements in humanoid robots: a comprehensive review and future prospects. IEEE/CAA J Autom Sinica 2024 Feb;11(2):301-328. [doi: 10.1109/JAS.2023.124140]
8. Figliano G, Miraglia L, Manzi F, et al. “Tom and Pepper Lab”. Robotics for cognitive stimulation and social skills: a preliminary study. Asian J Psychiatr 2025 Feb;104:104375. [doi: 10.1016/j.ajp.2025.104375] [Medline: 39892296]
9. Liang N, Li Y, Nejat G. In-person vs remote HRI: a comparative study of robot facilitated dance with older adults in long-term care. Presented at: 2023 IEEE-RAS 22nd International Conference on Humanoid Robots (Humanoids); Dec 12-14, 2023. [doi: 10.1109/Humanoids57100.2023.10375155]
10. Strutz N, Perotti L, Heimann-Steinert A, Klebbe R. Older adults’ communication with an interactive humanoid robot: expectations and experiences of older adults in verbal and nonverbal communication with a socially interactive humanoid robot: a mixed methods design in Germany. Z Gerontol Geriatr 2024 Aug;57(5):371-375. [doi: 10.1007/s00391-023-02268-y] [Medline: 38180514]
11. Jung S, Ahn SH, Ha J, Bahn S. A study on the effectiveness of IT application education for older adults by interaction method of humanoid robots. Int J Environ Res Public Health 2022 Sep 2;19(17):10988. [doi: 10.3390/ijerph191710988] [Medline: 36078697]
12. Fields N, Xu L, Greer J, Murphy E. Shall I compare thee…to a robot? An exploratory pilot study using participatory arts and social robotics to improve psychological well-being in later life. Aging Ment Health 2021 Mar;25(3):575-584. [doi: 10.1080/13607863.2019.1699016] [Medline: 31851830]
13. Ohneberg C, Stöbich N, Warmbein A, et al. Assistive robotic systems in nursing care: a scoping review. BMC Nurs 2023 Mar 18;22(1):72. [doi: 10.1186/s12912-023-01230-y] [Medline: 36934280]
14. Trainum K, Tunis R, Xie B, Hauser E. Robots in assisted living facilities: scoping review. JMIR Aging 2023 Mar 6;6:e42652. [doi: 10.2196/42652] [Medline: 36877560]
15. Morgan AA, Abdi J, Syed MAQ, Kohen GE, Barlow P, Vizcaychipi MP. Robots in healthcare: a scoping review. Curr Robot Rep 2022;3(4):271-280. [doi: 10.1007/s43154-022-00095-4] [Medline: 36311256]
16. Vandemeulebroucke T, Dierckx de Casterlé B, Gastmans C. The use of care robots in aged care: a systematic review of argument-based ethics literature. Arch Gerontol Geriatr 2018 Jan;74:15-25. [doi: 10.1016/j.archger.2017.08.014] [Medline: 28926749]
17. Hung L, Liu C, Woldum E, et al. The benefits of and barriers to using a social robot PARO in care settings: a scoping review. BMC Geriatr 2019 Aug 23;19(1):232. [doi: 10.1186/s12877-019-1244-6] [Medline: 31443636]
18. Hung L, Wong J, Smith C, et al. Facilitators and barriers to using telepresence robots in aged care settings: a scoping review. J Rehabil Assist Technol Eng 2022;9:20556683211072385. [doi: 10.1177/20556683211072385] [Medline: 35083063]
19. Andtfolk M, Nyholm L, Eide H, Fagerström L. Humanoid robots in the care of older persons: a scoping review. Assist Technol 2022 Sep 3;34(5):518-526. [doi: 10.1080/10400435.2021.1880493] [Medline: 33481675]
20. Arksey H, O’Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
21. Humanoid robot-assisted support for health care in older adults: a systematic scoping review. Open Science Framework. URL: https://osf.io/zqx8y/overview [accessed 2026-03-05]
22. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
23. Bradbury-Jones C, Aveyard H, Herber OR, Isham L, Taylor J, O’Malley L. Scoping reviews: the PAGER framework for improving the quality of reporting. Int J Soc Res Methodol 2022 Jul 4;25(4):457-470. [doi: 10.1080/13645579.2021.1899596]
24. Catricalà B, Manca M, Paternò F, Sale A, Santoro C, Zedda E. Exploiting personal memories in humanoid robot serious games for mild cognitive impaired older adults. Behav Inf Technol 2025 Nov 8;44(18):4616-4641. [doi: 10.1080/0144929X.2025.2484647]

XSL•FO
RenderX

25. Andtfolk M, Fagerström L, Eide H, Nyholm L. Experiences of a humanoid robot-led physical training program for home-living older persons - a qualitative pilot study. Int J of Soc Robotics 2025 May;17(5):781-788. [doi: 10.1007/s12369-025-01212-x]
26. Wu D, Pu L, Jo J, Hexel R, Moyle W. Deploying robot-led activities for people with dementia at aged care facilities: a feasibility study. J Am Med Dir Assoc 2024 Jul;25(7):105028. [doi: 10.1016/j.jamda.2024.105028] [Medline: 38772526]
27. Tahan K, Cayrier A, Baratgin J, N'kaoua B. ZORA robot to assist a caregiver in prospective memory tasks: a preliminary study. Appl Neuropsychol Adult 2026;33(1):225-232. [doi: 10.1080/23279095.2024.2343766] [Medline: 38648448]
28. Otaka E, Osawa A, Kato K, et al. Positive emotional responses to socially assistive robots in people with dementia: pilot study. JMIR Aging 2024 Apr 11;7:e52443. [doi: 10.2196/52443] [Medline: 38623717]
29. Lee OE, Lee H, Park A, Choi NG. My precious friend: human-robot interactions in home care for socially isolated older adults. Clin Gerontol 2024;47(1):161-170. [doi: 10.1080/07317115.2022.2156829] [Medline: 36502295]
30. Gigliotta O, Padulo C, Chinzer E, Fairfield B. HUGO, a humanoid robot for training memory: a pilot study on healthy younger and older adults. Presented at: 2024 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE); Oct 23-25, 2024. [doi: 10.1109/MetroXRAINE62247.2024.10796359]
31. Browne R, Alam MM, Saleem Q, et al. Reflective dialogues with a humanoid robot integrated with an LLM and a curated NLU system for positive behavioral change in older adults. Electronics (Basel) 2024;13(22):4364. [doi: 10.3390/electronics13224364]
32. Tobis S, Piasek-Skupna J, Suwalska A. The Godspeed Questionnaire series in the assessment of the social robot TIAGo by older individuals. Sensors (Basel) 2023 Aug 18;23(16):7251. [doi: 10.3390/s23167251] [Medline: 37631787]
33. Bogliolo M, Burlando F, Germinario A, Vallone F, Pilotto A, Micheli E. Robot-assisted group exercise program for targeting sarcopenia in older adults: preliminary results. Presented at: 2024 IEEE-RAS 23rd International Conference on Humanoid Robots (Humanoids); Nov 19-21, 2024. [doi: 10.1109/Humanoids58906.2024.10769857]
34. Zedda E, Manca M, Paterno F, Santoro C. MCI older adults' user experience with introverted and extraverted humanoid robot personalities. 2023 Sep 20 Presented at: CHItaly 2023; Jun 12-15, 2023 URL: https://dl.acm.org/doi/proceedings/10.1145/3605390 [doi: 10.1145/3605390.3605405]
35. Tobis S, Piasek-Skupna J, Neumann-Podczaska A, Suwalska A, Wieczorowska-Tobis K. The effects of stakeholder perceptions on the use of humanoid robots in care for older adults: postinteraction cross-sectional study. J Med Internet Res 2023 Aug 4;25(1):e46617. [doi: 10.2196/46617] [Medline: 37540548]
36. Tobis S, Piasek-Skupna J, Neumann-Podczaska A, Religioni U, Suwalska A. Determinants of attitude to a humanoid social robot in care for older adults: a post-interaction study. Med Sci Monit 2023 Sep 8;29:e941205. [doi: 10.12659/MSM.941205] [Medline: 37679949]
37. Leung AYM, Zhao IY, Lin S, Lau TK. Exploring the presence of humanoid social robots at home and capturing human-robot interactions with older adults: experiences from four case studies. Healthcare (Basel) 2023;11(1):39. [doi: 10.3390/healthcare11010039]
38. Lee OEK, Nam I, Chon Y, Park A, Choi N. Socially assistive humanoid robots: effects on depression and health-related quality of life among low-income, socially isolated older adults in South Korea. J Appl Gerontol 2023 Mar;42(3):367-375. [doi: 10.1177/07334648221138283] [Medline: 36326599]
39. Abdollahi H, Mahoor MH, Zandie R, Siewierski J, Qualls SH. Artificial emotional intelligence in socially assistive robots for older adults: a pilot study. IEEE Trans Affect Comput 2023;14(3):2020-2032. [doi: 10.1109/taffc.2022.3143803] [Medline: 37840968]
40. Nowak J, Fraisse P, Cherubini A, Daures JP. Assistance to older adults with comfortable robot-to-human handovers. Presented at: 2022 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO); May 28-30, 2022. [doi: 10.1109/ARSO54254.2022.9802960]
41. Nault E, Baillie L, Broz F. Investigating the usability of a socially assistive robotic cognitive training task with augmented sensory feedback modalities for older adults. Presented at: 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); Aug 29 to Sep 2, 2022. [doi: 10.1109/RO-MAN53752.2022.9900825]
42. Inoue K, Yatsu C, Yao DPG, Kohno M, Wada K, Yamamoto S. Preliminary study on the benefits of using the robot PALRO® in facilitating leisure programs for older adults with dementia. Gerontechnology 2022 Jan 1;21(1):1-7 [FREE Full text] [doi: 10.4017/gt.2022.21.1.466.04]
43. Fattal C, Cossin I, Pain F, et al. Perspectives on usability and accessibility of an autonomous humanoid robot living with elderly people. Disabil Rehabil Assist Technol 2022 May;17(4):418-430. [doi: 10.1080/17483107.2020.1786732] [Medline: 32643466]
44. Blavette L, Rigaud AS, Anzalone SM, et al. A robot-mediated activity using the Nao robot to promote COVID-19 precautionary measures among older adults in geriatric facilities. Int J Environ Res Public Health 2022 Apr 25;19(9):5222. [doi: 10.3390/ijerph19095222] [Medline: 35564615]
45. Tobis S, Piasek J, Cylkowska-Nowak M, Suwalska A. Robots in eldercare: how does a real-world interaction with the machine influence the perceptions of older people? Sensors (Basel) 2022 Feb 22;22(5):1717. [doi: 10.3390/s22051717] [Medline: 35270866]

XSL•FO
RenderX

46. Tulsulkar G, Mishra N, Thalmann NM, Lim HE, Lee MP, Cheng SK. Can a humanoid social robot stimulate the interactivity of cognitively impaired elderly? A thorough study based on computer vision methods. Vis Comput 2021;37(12):3019-3038. [doi: 10.1007/s00371-021-02242-y] [Medline: 34345091]
47. Park EA, Jung AR, Lee KA. The humanoid robot Sil-Bot in a cognitive training program for community-dwelling elderly people with mild cognitive impairment during the COVID-19 pandemic: a randomized controlled trial. Int J Environ Res Public Health 2021 Aug 3;18(15):8198. [doi: 10.3390/ijerph18158198] [Medline: 34360490]
48. Manca M, Paternò F, Santoro C, et al. The impact of serious games with humanoid robots on mild cognitive impairment older adults. Int J Hum Comput Stud 2021 Jan;145:102509. [doi: 10.1016/j.ijhcs.2020.102509]
49. Betriana F, Tanioka T, Osaka K, Kawai C, Yasuhara Y, Locsin RC. Interactions between healthcare robots and older people in Japan: a qualitative descriptive analysis study. Japan Journal Nursing Sci 2021 Jul;18(3):e12409 [FREE Full text] [doi: 10.1111/jjns.12409]
50. Pino O, Palestra G, Trevino R, De Carolis B. The humanoid robot NAO as trainer in a memory program for elderly people with mild cognitive impairment. Int J of Soc Robotics 2020 Jan;12(1):21-33. [doi: 10.1007/s12369-019-00533-y]
51. Melkas H, Hennala L, Pekkarinen S, Kyrki V. Impacts of robot implementation on care personnel and clients in elderly-care institutions. Int J Med Inform 2020 Feb;134:104041. [doi: 10.1016/j.ijmedinf.2019.104041] [Medline: 31838291]
52. Ke C, Lou VWQ, Tan KCK, Wai MY, Chan LL. Changes in technology acceptance among older people with dementia: the role of social robot engagement. Int J Med Inform 2020 Sep;141:104241. [doi: 10.1016/j.ijmedinf.2020.104241] [Medline: 32739611]
53. Chen LY, Sumioka H, Ke LJ, et al. Effects of teleoperated humanoid robot application in older adults with neurocognitive disorders in Taiwan: a report of three cases. Aging Med Healthc 2020 Jun;11(2):67-71 [FREE Full text] [doi: 10.33879/AMH.2020.066-2001.003]
54. Chen K, Lou VWQ, Tan KCK, Wai MY, Chan LL. Effects of a humanoid companion robot on dementia symptoms and caregiver distress for residents in long-term care. J Am Med Dir Assoc 2020 Nov;21(11):1724-1728. [doi: 10.1016/j.jamda.2020.05.036] [Medline: 32713772]
55. Casey D, Barrett E, Kovacic T, et al. The perceptions of people with dementia and key stakeholders regarding the use and impact of the social robot MARIO. Int J Environ Res Public Health 2020 Nov 20;17(22):8621. [doi: 10.3390/ijerph17228621] [Medline: 33233605]
56. Tanioka R, Sugimoto H, Yasuhara Y, et al. Characteristics of transactive relationship phenomena among older adults, care workers as intermediaries, and the Pepper robot with care prevention gymnastics exercises. J Med Invest 2019;66(1.2):46-49. [doi: 10.2152/jmi.66.46] [Medline: 31064953]
57. Karunarathne D, Morales Y, Nomura T, Kanda T, Ishiguro H. Will older adults accept a humanoid robot as a walking partner? Int J of Soc Robotics 2019 Apr;11(2):343-358. [doi: 10.1007/s12369-018-0503-6]
58. Greer JA, Doelling K, Xu L, Fields N. Shakespeare and robots: participatory performance art for older adults. 2019 Presented at: 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); Oct 14-18, 2019. [doi: 10.1109/RO-MAN46459.2019.8956469]
59. D'Onofrio G, Sancarlo D, Raciti M, et al. MARIO Project: validation and evidence of service robots for older people with dementia. J Alzheimers Dis 2019;68(4):1587-1601. [doi: 10.3233/JAD-181165] [Medline: 30958360]
60. Gerłowska J, Skrobas U, Grabowska-Aleksandrowicz K, et al. Assessment of perceived attractiveness, usability, and societal impact of a multimodal robotic assistant for aging patients with memory impairments. Front Neurol 2018;9:392. [doi: 10.3389/fneur.2018.00392] [Medline: 29910769]
61. Bedaf S, Marti P, Amirabdollahian F, de Witte L. A multi-perspective evaluation of a service robot for seniors: the voice of different stakeholders. Disabil Rehabil Assist Technol 2018 Aug;13(6):592-599. [doi: 10.1080/17483107.2017.1358300] [Medline: 28758532]
62. Feingold Polak R, Elishay A, Shachar Y, Stein M, Edan Y, Levy Tzedek S. Differences between young and old users when interacting with a humanoid robot. 2018 Mar Presented at: HRI '18; Mar 5-8, 2018 URL: https://dl.acm.org/doi/proceedings/10.1145/3173386 [doi: 10.1145/3173386.3177046] [Medline: 33141717]
63. Ikeya K, Aoki JI, Ikeda N, et al. Use of a human-type communication robot to evaluate the categorized communicative ability of older adults with dementia. Geriatr Gerontol Int 2018 Jan;18(1):188-190. [doi: 10.1111/ggi.13185] [Medline: 29341483]
64. Piezzo C, Suzuki K. Feasibility study of a socially assistive humanoid robot for guiding elderly individuals during walking. Future Internet 2017 Jul;9(3):30. [doi: 10.3390/fi9030030]
65. Abdollahi H, Mollahosseini A, Lane JT, Mahoor MH. A pilot study on using an intelligent life-like robot as a companion for elderly individuals with dementia and depression. Presented at: 2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids); Nov 15-17, 2017. [doi: 10.1109/HUMANOIDS.2017.8246925]
66. Khosla R, Nguyen K, Chu MT. Human robot engagement and acceptability in residential aged care. International Journal of Human–Computer Interaction 2017 Jun 3;33(6):510-522. [doi: 10.1080/10447318.2016.1275435]
67. Sakata N, Majima Y, Ishiguro K. Deployment of ARCS model and utilization of communication robot in patient education. Presented at: Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies; Feb 21-23, 2016. [doi: 10.5220/0005774603710376]

XSL•FO
RenderX

68. Valentí Soler M, Agüera-Ortiz L, Olazarán Rodríguez J, et al. Social robots in advanced dementia. Front Aging Neurosci 2015;7:133. [doi: 10.3389/fnagi.2015.00133] [Medline: 26388764]
69. Doering N, Richter K, Gross HM, et al. Robotic companions for older people: a case study in the wild. Stud Health Technol Inform 2015;219:147-152. [Medline: 26799897]
70. Orejana JR, MacDonald BA, Ahn HS, et al. Healthcare robots in homes of rural older adults. 2015 Presented at: Proceedings of the International Conference on Social Robotics (ICSR 2015); Oct 26-30, 2015. [doi: 10.1007/978-3-319-25554-5_51]
71. Torta E, Werner F, Johnson DO, et al. Evaluation of a small socially-assistive humanoid robot in intelligent homes for the care of the elderly. J Intell Robot Syst 2014 Sep;76(1):57-71. [doi: 10.1007/s10846-013-0019-0]
72. Louie WYG, McColl D, Nejat G. Acceptance and attitudes toward a human-like socially assistive robot by older adults. Assist Technol 2014;26(3):140-150. [doi: 10.1080/10400435.2013.869703] [Medline: 26131794]
73. Wu YH, Wrobel J, Cornuet M, Kerhervé H, Damnée S, Rigaud AS. Acceptance of an assistive robot in older adults: a mixed-method study of human–robot interaction over a 1-month period in the Living Lab setting. CIA 2014;9:801. [doi: 10.2147/CIA.S56435]
74. Khosla R, Chu MT, Nguyen K. Enhancing Emotional Well Being of Elderly Using Assistive Social Robots in Australia. Presented at: 2013 International Conference on Biometrics and Kansei Engineering (ICBAKE); Jul 5-7, 2013. [doi: 10.1109/ICBAKE.2013.9]
75. Choe YK, Jung HT, Baird J, Grupen RA. Multidisciplinary stroke rehabilitation delivered by a humanoid robot: interaction between speech and physical therapies. Aphasiology 2013 Mar;27(3):252-270. [doi: 10.1080/02687038.2012.706798]
76. Khosla R, Chu MT, Nguyen K. Affective robot enabled capacity and quality improvement of nursing home aged care services in Australia. Presented at: Proceedings of the 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops (COMPSACW); Jul 26-27, 2013. [doi: 10.1109/COMPSACW.2013.89]
77. Tanaka M, Ishii A, Yamano E, et al. Effect of a human-type communication robot on cognitive function in elderly women living alone. Med Sci Monit 2012 Sep;18(9):CR550-CR557. [doi: 10.12659/msm.883350] [Medline: 22936190]
78. Maalouf N, Sidaoui A, Elhajj IH, Asmar D. Robotics in nursing: a scoping review. J Nurs Scholarsh 2018 Nov;50(6):590-600. [doi: 10.1111/jnu.12424] [Medline: 30260093]
79. Lee JY, Song YA, Jung JY, et al. Nurses' needs for care robots in integrated nursing care services. J Adv Nurs 2018 May 13;74(12):2894-2904. [doi: 10.1111/jan.13711] [Medline: 29754395]
80. Sørensen L, Johannesen DT, Johnsen HM. Humanoid robots for assisting people with physical disabilities in activities of daily living: a scoping review. Assist Technol 2025 May 4;37(3):203-219. [doi: 10.1080/10400435.2024.2337194] [Medline: 38709552]
81. Lauver DR, Gross J, Ruff C, Wells TJ. Patient-centered interventions: implications for incontinence. Nurs Res 2004;53(6 Suppl):S30-S35. [doi: 10.1097/00006199-200411006-00005] [Medline: 15586145]
82. Morgan S, Yoder LH. A concept analysis of person-centered care. J Holist Nurs 2012 Mar;30(1):6-15. [doi: 10.1177/0898010111412189] [Medline: 21772048]
83. Atia MGB, Mohammad A, Gameros A, Axinte D, Wright I. Reconfigurable soft robots by building blocks. Adv Sci (Weinh) 2022 Nov;9(33):e2203217. [doi: 10.1002/advs.202203217] [Medline: 36192162]
84. Koh WQ, Felding SA, Budak KB, Toomey E, Casey D. Barriers and facilitators to the implementation of social robots for older adults and people with dementia: a scoping review. BMC Geriatr 2021 Jun 9;21(1):351. [doi: 10.1186/s12877-021-02277-9] [Medline: 34107876]
85. Luo C, Yang C, Yuan R, Liu Q, Li P, He Y. Barriers and facilitators to technology acceptance of socially assistive robots in older adults - a qualitative study based on the capability, opportunity, and motivation behavior model (COM-B) and stakeholder perspectives. Geriatr Nurs (Lond) 2024 Jul;58:162-170. [doi: 10.1016/j.gerinurse.2024.05.025]
86. Xie B, Watkins I, Golbeck J, Huang M. Understanding and changing older adults' perceptions and learning of social media. Educ Gerontol 2012 Apr 1;38(4):282-296. [doi: 10.1080/03601277.2010.544580] [Medline: 22639483]
87. Battistuzzi L, Sgorbissa A, Papadopoulos C, Papadopoulos I, Koulouglioti C. Embedding ethics in the design of culturally competent socially assistive robots. Presented at: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); Oct 1-5, 2018. [doi: 10.1109/IROS.2018.8594361]
88. 11.180 aids for disabled or handicapped persons. International Organization for Standardization. 2022. URL: https://www.iso.org/ics/11.180/x/ [accessed 2026-03-06]

Abbreviations

PAGER: Patterns, Advances, Gaps, Evidence for practice, and Research recommendations
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
WHO: World Health Organization

XSL•FO
RenderX

Edited by C Lee; submitted 09.Sep.2025; peer-reviewed by K Trainum, L Nyholm, W Qiaofei, X Liu; revised version received 21.Dec.2025; accepted 16.Feb.2026; published 11.Mar.2026.

Please cite as:
Cui L, Li Y, Yang X, Liu X, Zhang L, Hou L
Humanoid Robot–Assisted Support for Health Care in Older Adults: Systematic Scoping Review
JMIR Aging 2026;9:e83849
URL: https://aging.jmir.org/2026/1/e83849
doi:10.2196/83849

© Lei Cui, Yufei Li, Xinyao Yang, Xue Liu, Like Zhang, Lili Hou. Originally published in JMIR Aging (https://aging.jmir.org), 11.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Facilitators and Obstacles in Piloting Dementia Initiatives Within a Living Lab Approach: Systematic Review

Sara Alves[1,2], PhD; Natália Duarte[1,2], PhD; Teodora Figueiredo[3,4], MSc; Luis Midao[3,4], PhD; Maria João Bessa[5], PhD; Olalla Rodriguéz[6], MSc; José Fidalgo[6], MBA; Irene Fernandez[6], MSc; Alba Felpete[7,8,9], PhD; Maxi Rodriguéz[10], PhD; Juan Carlos Bernardéz[10], BSc; David Facal[7], PhD; Joana Carrilho[3,4], PhD; Elísio Costa[3,4], PhD

[1]CIDIFAD – Centro de Investigação, Diagnóstico, Formação e Acompanhamento das Demências, Santa Casa de Misericórdia de Riba D'Ave, Travessa Conde de Riba de Ave, Braga, Portugal

[10]AFAGA Alzheimer - Asociación de Familiares de Enfermos de Alzheimer y otras Demencias de Galicia, Vigo, Spain

[2]ICBAS - School of Medicine and Biomedical Sciences, RISE-Health, Porto, Portugal

[3]Biochemistry Lab, Faculty of Pharmacy of the University of Porto, RISE-Health, Porto, Portugal

[4]Competences Centre on Active and Healthy Ageing (Porto4Ageing), Faculty of Pharmacy of the University of Porto, Porto, Portugal

[5]UPTEC-Science and Technology Park of University of Porto, Porto, Portugal

[6]Galician Health Service (SERGAS), ACIS - Galician Agency of Knowledge in Health, Santiago de Compostela, Spain

[7]Department of Developmental Psychology, University of Santiago de Compostela, Santiago de Compostela, Spain

[8]Instituto de Investigación Sanitaria de Santiago de Compostela (Health Research Institute of Santiago de Compostela, IDIS), Santiago de Compostela, Spain

[9]Institute of Psychology, University of Santiago de Compostela, Santiago de Compostela, Spain

Corresponding Author:
Sara Alves, PhD
CIDIFAD – Centro de Investigação, Diagnóstico, Formação e Acompanhamento das Demências, Santa Casa de Misericórdia de Riba D'Ave, Travessa Conde de Riba de Ave, Braga, Portugal

Abstract

Background: Dementia affects more than 55 million people worldwide, and its progressive cognitive decline creates substantial challenges for intervention testing and real-world implementation. Living Labs (LLs) have become increasingly relevant for piloting interventions in dementia care, offering real-world environments for cocreation and iterative testing. However, operational, ethical, and governance challenges can hinder the effective implementation of dementia-focused initiatives.

Objective: This study aimed to identify and synthesize the facilitators and challenges encountered when piloting dementia-related initiatives within an LL approach.

Methods: A systematic review was conducted in PubMed, Web of Science, Scopus, and EBSCOhost from database inception to January 31, 2024. Studies were eligible if they reported the piloting, implementation, or evaluation of dementia-related initiatives in an LL context. A narrative synthesis was performed using content analysis for initial coding and thematic analysis for theme development. The systematic review was conducted in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Results: The 15 included studies were mostly conducted in European countries. Two major themes emerged, namely (1) organizational and operational issues and (2) ethical and legal considerations. Within these, 4 key operational subthemes were identified: engaging people living with dementia (end users), research design and evaluation methods, co-design and testing phases, and LL governance. Ethical and legal considerations covered informed consent procedures and regulatory aspects. Across studies, the most frequently reported facilitators involved familiar environments, stakeholder collaboration, iterative evaluation, and flexible co-design processes. Barriers commonly included fluctuating cognitive symptoms, limited digital literacy, inconsistent stakeholder engagement, time and resource constraints, and fragmented regulatory guidance.

Conclusions: Piloting dementia initiatives in LLs requires methodological adaptability, strong governance structures, and ethical procedures tailored to cognitive impairment. These findings provide actionable guidance for researchers, practitioners, industry partners, and policymakers seeking to develop sustainable, user-centered LLs that support innovation in dementia care.

(JMIR Aging 2026;9:e77752) doi:10.2196/77752

KEYWORDS

cognitive impairment; digital health; innovation; quadruple helix approach; user-centered approach; real-world application

XSL•FO
RenderX

Introduction

Globally, it is estimated that 55 million people live with dementia, with this number projected to increase to 152 million by 2050 [1]. Dementia is a progressive neurodegenerative disorder that compromises memory, behavior, personality, judgment, attention, spatial awareness, language, abstract thought, and other executive functions [2]. These cognitive impairments significantly reduce individuals' ability to function socially, occupationally, and independently [3]. The growing impact of dementia presents a major public health challenge due to its extensive burden on people living with the disease, families, caregivers, health care professionals, and society as a whole [1,4].

Efforts to improve dementia care have led to the development of innovative approaches that extend beyond traditional clinical and technological interventions. While advances such as artificial intelligence–assisted diagnostics, smart home monitoring, and virtual reality therapies have shown potential, many fail to adequately address the lived experience of people living with dementia [5,6]. This highlights the need for more inclusive, user-centered strategies that embed people living with dementia and their caregivers throughout the innovation process.

In this context, Living Labs (LLs) have emerged as valuable frameworks for cocreating and testing solutions in real-life environments [7]. LLs are user-centered, real-world settings that integrate research and innovation through co-design, iterative testing, and multistakeholder collaboration, following a quadruple-helix approach that connects academia, industry, government, and civil society [8].

Their strength lies in their capacity to foster collaboration between researchers, health care providers, patients, and caregivers, bringing scientific knowledge into everyday practice and allowing continuous feedback loops for refinement and validation [9,10]. However, the application of LLs in dementia care is particularly complex. The multifaceted nature of dementia, ethical considerations related to vulnerable participants, and the logistical and governance demands of maintaining multiactor environments all present significant challenges [10-13].

Previous studies have explored some of these aspects, identifying both enabling and constraining factors in operating LLs in health care. Some of the reported barriers include challenges related to the disease's impact on end users' participation [14,15], such as difficulties in decision-making and evaluation processes [11,12]; securing adequate funding [10]; addressing ethical concerns related to vulnerable populations [16,17]; and managing logistical issues to create realistic yet controlled environments [18]. Conversely, facilitators of successful LLs include strong stakeholder collaboration, robust technological infrastructure, and well-established ethical frameworks [19,20], which enhance the effectiveness and impact of these initiatives. Nonetheless, existing evidence remains fragmented and largely focused on single-case experiences or technological pilots. A comprehensive synthesis of facilitators and barriers specific to dementia-related LLs is still lacking.

Given the increasing prevalence of dementia and the need to optimize innovation processes in real-world contexts, this systematic review aims to identify and synthesize the facilitators and barriers encountered when piloting dementia-related initiatives within an LL framework. By consolidating available evidence, this study seeks to inform the design, implementation, and governance of sustainable, ethically sound, and user-centered LLs that can enhance innovation in dementia care.

Therefore, this review aims to systematically identify and synthesize the facilitators and barriers encountered when piloting dementia-related initiatives within LL environments.

Methods

Search Strategy

A systematic search was conducted in 4 databases for relevant literature: PubMed, Web of Science, Scopus, and EBSCOhost, from database inception to January 31, 2024. Keywords included MeSH (Medical Subject Headings) terms and synonyms as follows: "living lab" AND "dementia" AND "barrier" OR "obstacles" OR "facilitators" OR "enablers." Details of the search strategy, including specific search strings and filters used, are available in Multimedia Appendix 1.

Inclusion and Exclusion Criteria

Studies were included if they met all of the following criteria, namely (1) reporting findings within an LL approach (with consideration for multimethod approaches, user engagement, multistakeholder participation, real-life settings, and cocreation); (2) piloting initiatives related to dementia or cognitive decline; (3) randomized controlled trials, observational studies, qualitative studies, mixed methods studies, intervention studies, implementation science studies, case studies, and systematic reviews or meta-analyses or scoping reviews; and (4) publications written in English, Spanish, or Portuguese.

Studies were excluded if they met any of the following criteria, namely (1) those not addressing the topic of barriers and facilitators of piloting initiatives; (2) those describing or analyzing the technical parameters of a digital solution; (3) those comprising samples with cognitive problems not directly related to dementia or degenerative cognitive impairment (eg, aphasia); and (4) study protocols, theoretical papers, conceptual papers, and position papers.

Study Selection and Data Extraction

After retrieving manuscripts, duplicates were removed using EndNote software (v21; Clarivate Analytics). Subsequently, 2 authors (SA and ND) independently assessed the titles and abstracts of all potentially relevant articles. The full texts of these articles were then meticulously reviewed to determine final eligibility based on predefined inclusion and exclusion criteria. Any discrepancies were resolved through consensus with a third reviewer (TF), who conducted an independent analysis of such manuscripts. Additionally, reference lists of the included articles were screened to identify supplementary relevant studies.

To ensure consistency in data extraction, an Excel spreadsheet (Office 21, v.2405; Microsoft Corporation) was created to

XSL•FO
RenderX

document key study information, including author details, publication year, country, study design, sample size, participant type (eg, people with dementia, professionals, caregivers, and family members), and a brief description of the piloted initiatives. Extracted data also included key facilitators and barriers reported in each study.

Data Synthesis

A narrative synthesis was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines. The synthesis followed two sequential analytical stages:

1. Content analysis was first applied to systematically code all reported facilitators and barriers across studies, identifying recurring concepts and relationships.
2. Thematic analysis was subsequently used to cluster and refine these codes into overarching themes and subthemes [21].

Two reviewers (SA and ND) independently performed the coding using NVivo software (v.14.23.3; Lumivero LLC). Regular meetings were held to discuss coding discrepancies and reach consensus, with a third reviewer (TF) available for arbitration when needed.

Following an iterative review of the extracted data and initial coding, two overarching themes were identified: (1) organizational and operational issues and (2) ethical and legal considerations. Each of these themes was further divided into subtopics that encompassed both barriers and facilitators [22]. Themes and subtopics were refined and organized into a final framework of facilitators and barriers.

A formal risk of bias assessment was not conducted, as the included studies primarily focused on methodological and participatory approaches for engaging people living with dementia under an LL approach rather than interventional studies. However, to ensure transparency, a completed PRISMA checklist is provided as supplementary material in Checklist 1. This review was registered in PROSPERO (CDR42024514481).

Results

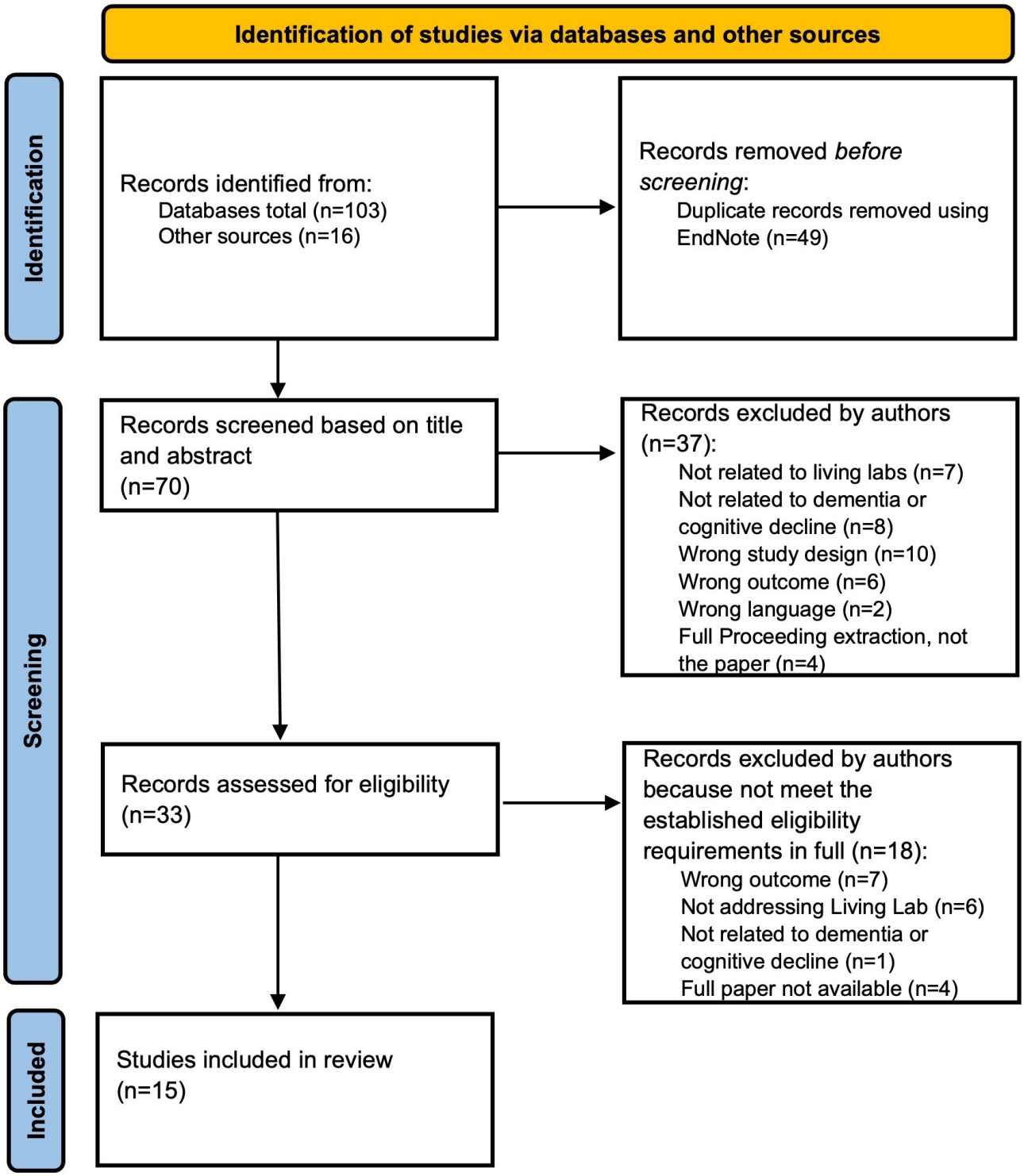
Overview

The initial search resulted in a total of 119 publications. After duplicate removal and full-text analysis, 15 met the inclusion criteria (Figure 1).

Findings were synthesized narratively through content and thematic analysis, according to the predefined analytical framework. Two overarching themes emerged, which were (1) organizational and operational issues and (2) ethical and legal considerations, each comprising several subthemes that captured both facilitators and barriers. Of the 15 studies [18,23-36], engaging people living with dementia was the most frequently reported subtheme (14 studies) [18,23-31,33-36], followed by research design and evaluation methods (12 studies) [23-31,33-36], co-design and testing phases (12 studies) [23-31,34-36], and LL governance (10 studies) [18,23,26,28-31,33-35].

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the included studies.



The studies were mainly qualitative (n=10) [18,23-25,28-32,35] and mixed methods (n=5) [26,27,33,34,36] (Multimedia Appendix 2). The reviewed articles were all from Europe, with the exception of 1 article from Canada, that is, Pigot and Giroux [23]. Facilitators and barriers were presented below by themes and subtopics. Multimedia Appendix 2 provides an overview of the included studies, while Table 1 summarizes facilitators and barriers by theme.

XSL•FO
RenderX

Table . Summary of facilitators and barriers by theme.

Theme	Subthemes	Most frequent facilitators	Most frequent barriers
Organizational and operational issues	• Engaging people living with dementia • Research design and evaluation methods • Co-design and testing phases • Living Lab governance	• Friendly atmosphere, familiar settings, caregiver involvement, and clear communication • Mixed methods approaches, visual aids, iterative feedback, and involvement of key informants • Real-life testing, pilot testing with healthy adults, training sessions, and flexible procedures • Quadruple-helix collaboration, shared vision, transparency, and regular meetings	• Digital literacy limitations, cognitive fluctuations, perceived stigma, and privacy concerns • Low adherence to questionnaires, positive-response bias, inconsistent data, and difficulty demonstrating real-world benefits • Time constraints, technical malfunctions, interruptions, and recruitment difficulties • Funding shortages, administrative burden, communication gaps, and divergent stakeholder priorities
Ethical and legal considerations	• Informed consent procedures • Regulatory aspects	• Rolling consent, adapted forms, and oral and written explanations • Alignment with ethical standards, involvement of gatekeepers, and certification	• Difficulty assessing capacity, limitations of one-time consent, and unclear evidence on audiovisual tools • Fragmented regulations, data-protection challenges, and lack of standardized contracts

Theme 1: Organizational and Operational Issues

Within the theme of organizational and operational aspects, 14 [18,23-31,33-36] out of 15 studies [18,23-36] identified various facilitators and barriers in piloting dementia initiatives using the LL approach. Four subthemes were identified within this theme: engaging people living with dementia, research design and evaluation methods, co-design and testing phases, and LL governance.

Engaging People Living With Dementia (End Users)

Facilitators

Fourteen studies emphasized strategies that promoted meaningful involvement of people living with dementia [18,23-31,33-36]. Common facilitators included a relaxed atmosphere and assisting with transportation [24].

The use of native language, careful pacing of information, and ensuring users were well-informed, using accessible language, were also mentioned [26,27]. Other studies highlighted the importance of a friendly atmosphere (eg, welcome participants with coffee or tea; coffee breaks; and not to dive into the product immediately, but starting by making informal contact first), familiar settings (eg, real-life conditions), and involving familiar people and care professionals to make end users feel safe and understood [18,23,28-31]. Recognizing the real-world benefits of initiatives and managing participants' expectations were described as crucial aspects for fostering engagement [18,27,34]. Another aspect that emerged was the possibility of offering rewards to increase users' willingness to participate in the initiatives [35].

Barriers

Several challenges related to engagement in LL activities were noted in 8 of the reviewed studies [26,28,29,31,34-36]. Some challenges include the participation of users with limited technical skills, which made it harder for them to fully participate. Additionally, not all user suggestions could be implemented due to technical or practical constraints (eg, development stage of the prototype), which sometimes caused frustration or a sense of exclusion because users felt their contributions were not fully considered [31,34]. Furthermore, negative perceptions of the solution, such as seeing it as stigmatizing or too complex to use, can lead to reduced confidence or reluctance to participate, ultimately hindering the engagement of people living with dementia [31,35,36]. Privacy concerns also limited participation, as participants reported reluctance to allow sensor monitoring in private home areas [36]. Other barriers include reluctance to express criticism or fear negative consequences, which can affect their willingness to provide feedback [28].

Research Design and Evaluation Methods

Facilitators

Mixed methods approaches to evaluate studies under an LL framework were endorsed in 12 studies [23-28,30,31,33-36], highlighting its potential to provide a comprehensive understanding of the LL approach, as well as real-world and user experiences. Qualitative methods, such as semistructured interviews and focus groups, effectively captured the experiences of people living with dementia. The use of visual aids, key informants, nonintrusive evaluation instruments [31,33], and postsession reflections [24,31] was recommended to enhance engagement. Additionally, a user-centered research design that emphasizes cocreation and continuous stakeholder

XSL•FO
RenderX

feedback was highlighted in 5 studies [26,28,30,31,33] as essential for developing relevant dementia interventions.

Barriers

Challenges in research design and evaluation methods were noted in 10 [23-25,27,28,30,31,33-35] of the 15 reviewed studies [18,23-36], and included low adherence to questionnaires, bias toward positive feedback, and inconsistent data collection, as well as issues like team changes and lack of diversity among participants that may limit study findings [35]. Other barriers included challenges in demonstrating the real-world benefits of initiatives [28,33,34], and ensuring sustained involvement throughout the project lifecycle [33].

Co-Design and Testing Phases

Facilitators

One widely recognized facilitator, mentioned in 12 [23-31,34-36] out of 15 studies [18,23-36], is the importance of integrating co-design and testing phases into daily activities and real-life settings. Other aspects included clear staff roles and environments tailored to the users' needs, using flexible facilities, nonintrusive methods, and adapting testing conditions for users' difficulties [23,25,26,28-31,34]. Conducting pilots with healthy older adults before involving people living with dementia was suggested as an effective strategy to identify potential issues early and adjust the testing process accordingly [24,30,34], helping to prevent confusion for end users. Implementing training sessions prior to deploying solutions [24,30] was also recognized as a valuable strategy.

Barriers

Challenges related to the co-design and testing phases were noted in 8 studies [23,26,28,30,31,34-36]. These included time constraints, budget limitations, and technical difficulties [26,34,35]. Additionally, the fluctuating symptoms of dementia require highly adaptable solutions, which can be difficult to find [31]. As an example, it was found that users' attention and motor fluctuations affected interaction with digital prototypes, and it was observed that caregivers had to adapt testing sessions daily to participants' cognitive variability [30,31].

Other studies pointed to the difficulties of tests in real-world settings [36], such as interruptions during home interventions, potential performance blur between user and relatives, and conflicts between user and researcher agendas [23].

LL Governance

Facilitators

Governance facilitators were highlighted by 10 studies [23,26,28-31,33-36]. Quadruple helix approach through the collaboration of academia, industry, government, and civil society, nurturing a culture of openness and team belonging, was described as an important aspect to ensure a smooth implementation of the innovation process [23,28,29]. Other facilitators highly mentioned included regular meetings and stakeholders' communication, clearly defined roles, and the development of roadmaps, which contribute to the LL transparency.

Barriers

Governance challenges were identified in 7 studies [18,26,28,31,33-35]. The main identified barriers included funding, time-consuming processes, defining a shared vision among stakeholders, balancing stakeholder interests, maintaining communication across all stakeholders, and managing complex administrative tasks.

Theme 2: Ethical and Legal Considerations

Eight [18,24,26,27,29-32] out of 15 studies [18,23-36] identified ethical and legal facilitators and barriers relevant to a dementia-focused LL approach. These findings were categorized into 2 main topics: informed consent procedures and regulatory aspects. While all studies discussed facilitators and/or barriers related to informed consent procedures, only 4 of them [18,27,29,31] addressed regulatory aspects (refer to Multimedia Appendix 2 for further details).

Informed Consent Procedures

Facilitators

"Process consent" (ie, ongoing seeking permission from participants throughout the study duration) or "rolling consent" (ie, seeking permission from participants as new phases or new activities are introduced) was advocated in 5 [18,26,27,29,32] studies with recommendations to monitor participant reactions and regularly remind them of their option to withdraw [31]. The importance of using both oral and written consent formats was also highlighted [29,32]. Additionally, sending consent forms in advance and assessing participants' capacity to consent were emphasized as key ethical practices [30,32]. Obtaining consent from both people living with dementia and their representatives, using adapted communication strategies, and considering innovative approaches, such as interactive consents (ie, a dynamic and ongoing process of seeking and reaffirming consent from participants) were also noted in some studies [31,32].

Barriers

Challenges related to the obtainment of informed consents were noted in 3 [26,27,32] studies. They include difficulties with one-time consent agreements [26,27] or the complexity of assessing consent capability due to fluctuating cognitive abilities [32]. It was also noted that there is still a lack of evidence that validates the benefits of using audio-visual aids in the consent process [32], which may limit their adoption.

Regulatory Aspects

Facilitators

Regulatory considerations, such as aligning prototypes and research methods with ethical standards and regulatory aspects, were discussed in 2 studies [27,31]. These studies emphasized the involvement of gatekeepers (ie, individuals or organizations who ensure the protection and well-being of people living with dementia, commonly including care institutions, relatives, and health professionals) or regulatory bodies and consistent regulations of recognized authorities (eg, European Network of Living Labs [ENoLL]) to foster collaboration while safeguarding participants' interests.

XSL•FO
RenderX

Barriers

Regulatory barriers were highlighted in 3 studies [18,29,31]. These included fragmented regulations across regions, lack of specific guidelines for cross-country research, absence of standardized contractual models, and challenges in managing sensitive data and intellectual property (mainly due to the collaborative nature and iterative process of LL, which often blurs the lines between roles).

Discussion

Principal Findings

This systematic review identified key facilitators and barriers that influence the successful piloting of dementia initiatives within an LL approach. Two overarching themes emerged, namely "Organizational and Operational Issues" and "Ethical and Legal Considerations," each encompassing a series of recurrent subthemes across the 15 included studies [18,23-36]. Overall, the findings show that LLs hold significant potential to support innovation in dementia care, but their effectiveness depends heavily on adapting methodological, ethical, and governance procedures to the specific needs and vulnerabilities of people living with dementia.

Consistent with previous literature, the studies highlighted that engaging people living with dementia requires approaches that accommodate fluctuating cognitive abilities, promote familiarity, and reduce cognitive burden [19,37,38]. Facilitators such as friendly atmospheres, familiar settings, accessible communication formats, and the involvement of caregivers and professionals emerged as crucial to fostering meaningful engagement. These findings illustrate the inherent complexity of piloting solutions with this population, reinforcing the need for tailored, adaptable, and iterative strategies. They also underscore the importance of ethical awareness and safeguarding the rights and well-being of people living with dementia while ensuring that their perspectives are integrated throughout the co-design process.

Based on the analysis of barriers and facilitators, 3 core aspects emerge as central to understanding the functioning of LLs in dementia care: the target population features, the LL ecosystem**,** and the LL framework gaps.

The Target Population Features

This dimension encompasses the unique needs and challenges faced by people living with dementia, including aspects that influence their participation, engagement, and adherence to the cocreation and testing processes.

The studies showed that users' heterogeneous and fluctuating cognitive symptoms create significant challenges from the earliest stages of the innovation process. These fluctuations affect communication, emotional stability, and the ability to engage with tasks, requiring highly flexible planning and individualized support [39-42].

Aspects related to communication processes (eg, accessible communication, native language, written and oral information, visual aids, regular reminders, and managing information load), flexible and adapted assessment methods and iteration phases, and familiar settings and people, and consistent team members are essential considerations to handle the unpredictable nature of dementia. Addressing these challenges often requires involving proxies (eg, caregivers and relatives), which extends the time needed for the process and may increase communication challenges and stakeholders' differing perspectives. These consequences ultimately influence project governance, sustainability, and the overall pace of development within LLs [38,43,44].

Another key aspect concerns the perceived benefits of the initiatives. Several studies emphasized that people living with dementia and their caregivers are more likely to accept and use a technology or intervention when they perceive clear, immediate value [18,28,31,33,36]. This aligns with previous research that documented that technology is either not accepted at all or is quickly abandoned when users do not perceive an immediate benefit [45-47]. Ensuring simple, intuitive interfaces and avoiding stigmatizing features (eg, making solutions too discreet or discouraging their use within social circles) is therefore essential. Failing to address these factors may lead to user frustration and a refusal to engage or discontinuation of use.

LL Ecosystem

At the ecosystem level, the findings highlight the importance of establishing strong communication flows and working on nurturing the strategic alliances. While the quadruple helix approach provides a solid basis for multiactor collaboration, many studies emphasize the need to expand beyond traditional stakeholders to include society, the media, and local communities [48-52]. This could create more opportunities for digital innovation, diversify funding sources, and contribute to the sustainability of the LL. Sustainability emerged as a recurring challenge, with several studies identifying resource instability and limited long-term planning as barriers to continuity [18,31,33].

Some authors have proposed a "mixed-economy" approach, where both state and private funding support the innovation process in social care [18,45,53]. This approach could be particularly useful in promoting the sustainability of LLs focused on innovation in dementia care.

Another key point is the establishment of a clear goal and shared vision for the LL. The studies reviewed strongly emphasize that LL success requires a clear goal that integrates reliability, real-world benefit, and the well-being of the LL team members. Communication is central to achieving this. It allows collaboration, builds trust, and fosters a sense of belonging. Attention must be given to all stages, from preimplementation, ongoing activities, and postimplementation. Notably, the postimplementation phase appears particularly neglected, with limited evidence on long-term monitoring, communication, and impact assessment (eg, reports, communication, and long-term impact assessment) [54-56]. This gap raises concerns about whether solutions remain relevant and functional after initial deployment, given the progressive nature of dementia [19,43,57].

XSL•FO
RenderX

In dementia care, especially with digital solutions, ongoing monitoring is crucial. Since dementia is a progressive neurodegenerative disorder, adherence to digital solutions will be affected by the disease's progression. Understanding these factors is vital to ensure that digital solutions effectively support the quality of life of people living with dementia and their caregivers, and therefore, should be a central concern of LL activity. Without this focus, LL risks introducing products that do not meet their intended purpose or do not provide real benefit.

LL Framework Gaps

As a relatively new approach, LLs still lack a well-defined regulatory structure. The absence of standardized guidelines leads to inconsistencies across countries and sectors, creating challenges in legal, ethical, and operational domains. This includes heterogeneity in data protection requirements, intellectual property procedures, and ethical oversight mechanisms. The limited dissemination of national dementia strategies, with less than a quarter of countries having them in place, further complicates the development of consistent standards, particularly in these populations [58].

Ethical challenges represent a core concern, particularly regarding informed consent and data protection [16,45,59,60]. Cognitive decline complicates the assessment of consent capacity and increases the need for sensitive, adaptive consent procedures. Innovative approaches such as "rolling consent" or interactive consent forms, ensuring that participants are continuously informed of their rights and can withdraw from studies at any time, fostering trust and ensuring participant safety. Privacy concerns, especially in data-rich environments such as LLs, where monitoring and data collection occur in real-life settings (including homes, hospitals, and nursing homes), further complicate the ethical management of these activities. This complexity underscores the need for robust ethical and legal regulations for dementia-focused LLs.

Mechanisms such as ethical review boards, certification authorities (eg, ENoLL), and standard contractual models can support the establishment of responsible LL operations. Strengthening these structures is critical to fostering trust, enhancing participation, and enabling LLs to function effectively in dementia-focused contexts.

Recommendations for Practice

Based on the results of this study, we have established some recommendations for piloting dementia initiatives (or on cognitive impairment) using an LL approach, providing a practical overview of the findings. The proposed guidelines are organized according to the topics identified in this systematic review (Table 2).

By considering these specific recommendations, LLs can improve their functioning and tackle common challenges in piloting dementia initiatives. Also, these guidelines can contribute to improving practices and to fostering sustainable and efficient innovation processes. By covering a broad range of aspects that include not only user and community involvement, but also research, governance, and ethical and regulatory considerations, these guidelines provide practical guidance for academia (research), policymakers, industry, and health care agents. For example, integrating iterative user feedback and ethical reviews can help tailor initiatives more effectively; furthermore, establishing clear metrics for evaluating the initiatives' impact will be crucial for assessing their effectiveness. At the ultimate end, these guidelines have the potential to enhance the impact of the LL approach in the dementia field, leading to improved health outcomes and quality of life of those affected, ensuring that innovation processes meet users' needs effectively.

XSL•FO
RenderX

Table . Recommendations for operating a dementia Living Lab (based on the categories of results).

Category	Recommendations
Engaging people living with dementia (end users)	• Develop engagement strategies tailored to the characteristics of people living with dementia by focusing on their needs, creating a relaxed atmosphere, using accessible language, and involving close persons, including both relatives and care professionals. • Consider technological difficulties of users, offering simplified digital solutions along with close technical support to enhance usability and acceptability. • Monitor frequently the users' feedback and experience to avoid feelings of frustration and to ensure users feel respected and understood. • Prioritize solutions that respect privacy and avoid being intrusive or stigmatizing. • Keep users well-informed by clearly presenting the potential real-world benefits of the solution, its limitations, and setting realistic expectations.
Evaluation methods	• Privilege mixed methods approaches, particularly semistructured interviews and focus groups, that help people living with dementia to better express their perspectives. • Use alternative communication methods, such as visual aids, and involve close family members or caregivers to enhance communication. • Consider cognitive limitations when selecting tools by using observation and cognitive-specific instruments. • Ensure staff consistency throughout the process and assure appropriate training on evaluation assessments. • Foster friendly and nonintrusive evaluation environments that are flexible and prevent positive biased feedback. • Incorporate a postreflection session to gather more in-depth feedback from users.
Co-design and testing phases	• Integrate co-design and testing phases in real-life settings to make the process more natural and less disruptive. Ensure that testing environments are familiar, avoiding the need for users to adapt to unfamiliar situations. • Conducting pilots with healthy older adults before involving people living with dementia to ensure the prototype functions well and is user-friendly. • Plan the co-design and testing phases with extended time and resources to accommodate potential fluctuations in dementia symptoms. • Consider the technical requirements of the solutions, including font size, format, battery life, menu design, icons, and audio-visual and written content. • Ensure staff consistency and provide appropriate training both for staff and users. • Tailor solutions to users' needs and conduct regular checkpoints to ensure they remain effective.
Research design	• Consider co-creation and exploratory approaches to develop effective solutions for dementia. • Prioritize study designs that use a user-centered approach, combining empirical and naturalistic methods. • Encourage collaboration among stakeholders to create a research design that meets the needs of all partners and maximizes the relevance and impact of the outcomes. • Clearly communicate research procedures to all stakeholders throughout the project lifecycle, especially those with varying levels of involvement, and establish defined research protocols specifically for dementia interventions.

XSL•FO
RenderX

Category	Recommendations
LL[a] governance	• Adopt a quadruple helix approach by involving academia, industry, government, and civil society. • Establish a shared vision early in the project and balance the interests of all stakeholders to ensure smooth implementation. • Foster a sense of team belonging through regular communication, clearly defined roles, and the development of roadmaps. • Implement efficient management and administrative strategies to streamline processes and enhance project execution, while avoiding unnecessary bureaucratic procedures. • Focus on building a broad network of local, national, and international partnerships and seek certification from recognized authorities (eg, ENoLL[b]). • Develop a sustainable business model to secure adequate funding and ensure the long-term sustainability of Living Lab. • Raise awareness of Living Lab initiatives by implementing a robust communication plan that encourages community involvement and engagement. • Consider establishing multidisciplinary boards, including financial, administrative, and scientific experts, to provide comprehensive oversight and guidance. • Integrate a long-term perspective in innovation processes, including postimplementation monitoring. • Regularly report on the impact of Living Lab initiatives, clearly identifying their benefits, with a particular focus on the real-world impact for users.
Informed consent procedures	• Prioritize “process consent” or “rolling consent,” regularly reminding participants of their rights and option to withdraw. Observe any reactions that may indicate reluctance to participate. • Adopt consent-friendly procedures, such as using adapted consent forms and sending informed consent documents in advance for review. • Use both oral and written consent formats, explore innovative methods such as interactive informed consent to enhance engagement. • Consider using tools to assess the capability to consent, and when necessary, obtain consent from both the individual and their representative.
Regulatory aspects	• Involve gatekeepers or regulatory bodies to ensure that research practices adhere to ethical standards and regulatory requirements. • Adopt regulatory procedures by standardizing contractual models and establishing clear guidelines for managing intellectual property and data privacy. • Advocate for enhanced regulatory procedures at both national and international levels to promote harmonized practices and facilitate cross-country activities.

[a]LL: Living Lab.

[b]ENoLL: European Network of Living Labs.

Limitations

This systematic review is subject to certain limitations. First, despite a comprehensive search strategy across 4 major databases, relevant studies may have been overlooked due to language restrictions (English, Spanish, and Portuguese), and the still-emerging nature of research on dementia-focused LLs. These factors may have contributed to publication bias and limited the diversity of included contexts.

Another limitation of this study pertains to the thematic analysis of the findings, which is susceptible to the influence of the authors’ backgrounds and understanding of the topic. To strengthen trustworthiness, 2 authors independently coded and discussed emerging categories, and interpretations were refined through consensus and alignment with existing literature and the review’s guiding question.

A formal risk of bias assessment was not conducted, as this review focused on synthesizing facilitators and barriers rather than evaluating intervention effects. While this may limit comparability across studies, a structured appraisal was not applicable given the descriptive and exploratory nature of the evidence. Future reviews should incorporate formal bias assessments when suitable tools for qualitative and mixed methods LL research become available.

Finally, the geographical distribution of the publications is predominantly from European Union countries. Out of 15 studies included [18,23-36], only 1 was outside Europe, which could limit the generalization of the findings and reduce the

XSL•FO
RenderX

applicability of the conclusions to other contexts. Further research is needed to explore how national dementia strategies, funding mechanisms, and societal attitudes influence LL implementation beyond European contexts.

Conclusion

In conclusion, this systematic review identifies key aspects for successfully piloting dementia initiatives within an LL framework. A comprehensive and strategically planned approach is essential to effectively address real-world needs. It encompasses a user-centered design that accounts for the cognitive limitations of people living with dementia and balances the interests of stakeholders. Equally important is the development of robust governance models that foster collaboration, ensure smooth communication, promote extensive networking, and secure consistent and long-term funding, all supported by clear ethical guidelines and regulatory standards. By establishing equilibrium among stakeholders' interests and ensuring adaptability, LLs can serve as effective platforms for driving innovation in dementia care.

Acknowledgments

Generative artificial intelligence (AI) tools were used only for language editing and text revision. No AI tools were used to generate scientific content, interpret results, or draw conclusions. The authors remain fully responsible for the accuracy, originality, and integrity of the manuscript.

Funding

This work was funded by the INTERREG V-A Spain-Portugal (POCTEP) program through the project 0088_INNOV4LIFE_1_P, in which all authors are members of the project team.This publication was also supported by national funds through the Portuguese Foundation for Science and Technology, I.P. (FCT), in the scope of the project RISE-Health - UID/06397/2025., in which SA, ND, TF, LM, JC, and EC are members; and by the Spanish AEI (DOI: 10.13039/501100011033; Refs. PID2020-114521RB-C21 and PID2023-151659OB-C21) and the Galician Government (Consellería de Cultura, Educación e Ordenación Universitaria; axudas para a consolidación e estruturación de unidades de investigación competitivas do Sistema Universitario de Galicia ED431C 2021/04; GI-1807-USC: Ref. 2021-PG011), which support the University of Santiago de Compostela, where DF and AF are collaborators. TF was also granted by national funds through the Portuguese Foundation for Science and Technology, I.P. (FCT) (2024.02137.BD).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search terms.
[DOCX File, 26 KB - aging_v9i1e77752_app1.docx]

Multimedia Appendix 2
Included studies.
[DOC File, 179 KB - aging_v9i1e77752_app2.doc]

Checklist 1
PRISMA checklist.
[DOCX File, 31 KB - aging_v9i1e77752_app3.docx]

References

1. Nichols E, Vos T. The estimation of the global prevalence of dementia from 1990‐2019 and forecasted prevalence through 2050: An analysis for the Global Burden of Disease (GBD) study 2019. Alzheimer's & Dementia 2021 Dec;17(S10):e105-e125 [FREE Full text] [doi: 10.1002/alz.051496]
2. Dementia. National Library of Medicine. 2024. URL: https://www.ncbi.nlm.nih.gov/mesh?Db=mesh&Cmd=DetailsSearch&Term=C10.228.140.380%5BAll+Fields%5D [accessed 2026-02-23]
3. Shi P, Chapman H, Liu L, et al. Social cognition in mild cognitive impairment and dementia: a systematic review and meta‐analysis. Alzheimer's & Dementia 2025 Dec;21(S3):21 [FREE Full text] [doi: 10.1002/alz70857_099687]
4. Alves S, Duarte N, Gomes B. Forecasted dementia prevalence in Portugal (2020-2080). J Geriatr Psychiatry Neurol 2024 Sep;37(5):403-412. [doi: 10.1177/08919887241237220] [Medline: 38436576]
5. Miguel Cruz A, Daum C, Comeau A, et al. Acceptance, adoption, and usability of information and communication technologies for people living with dementia and their care partners: a systematic review. Disabil Rehabil Assist Technol 2023 May;18(4):443-457. [doi: 10.1080/17483107.2020.1864671] [Medline: 33378627]

XSL•FO
RenderX

6. Sohn M, Yang J, Sohn J, Lee JH. Digital healthcare for dementia and cognitive impairment: a scoping review. Int J Nurs Stud 2023 Apr;140:104413. [doi: 10.1016/j.ijnurstu.2022.104413] [Medline: 36821951]
7. Higgins A, Klein S. Introduction to the living lab approach. In: Tan YH, editor. Accelerating Global Supply Chains with IT-Innovation: ITAIDE Tools and Methods: Springer; 2011:1-379. [doi: 10.1007/978-3-642-15669-4_2]
8. Living lab origins, developments, and future perspectives. : European Network of Living Labs; 2025 URL: https://living-in.eu/sites/default/files/files/enoll-2025.-living-lab-origins-developments-and-future-perspectives.pdf [accessed 2026-03-16]
9. da Penha ADL, Bonato SV, Passos JB, Fernandes EDS, Kulpa C, ten Caten CS. Navigating the urgency: an open innovation project of protective equipment development from a quadruple helix perspective. Sustainability 2024;16(4):1636. [doi: 10.3390/su16041636]
10. Miller K, McAdam R, McAdam M. A systematic literature review of university technology transfer from a quadruple helix perspective: toward a research agenda. R&D Manag 2018 Jan;48(1):7-24 [FREE Full text] [doi: 10.1111/radm.12228]
11. Fox S, Brown LJE, Antrobus S, et al. Co-design of a smartphone app for people living with dementia by applying agile, iterative co-design principles: development and usability study. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e24483. [doi: 10.2196/24483] [Medline: 35029539]
12. Hendriks N, Slegers K, Wilkinson A. Lessons learned from participatory design in dementia care: placing care partners at the centre. Stud Health Technol Inform 2017;233:63-77. [Medline: 28125414]
13. van den Kieboom RC, Bongers IM, Mark RE, Snaphaan LJ. User-driven Living Lab for assistive technology to support people with dementia living at home: protocol for developing co-creation-based innovations. JMIR Res Protoc 2019 Jan 28;8(1):e10952. [doi: 10.2196/10952] [Medline: 30688653]
14. Black BS, Wechsler M, Fogarty L. Decision making for participation in dementia research. Am J Geriatr Psychiatry 2013 Apr;21(4):355-363. [doi: 10.1016/j.jagp.2012.11.009] [Medline: 23498382]
15. Stecker M. A perspective: challenges in dementia research. Medicina (Kaunas) 2022 Sep 28;58(10):10. [doi: 10.3390/medicina58101368] [Medline: 36295529]
16. Chandra M, Harbishettar V, Sawhney H, Amanullah S. Ethical issues in dementia research. Indian J Psychol Med 2021 Sep;43(5 Suppl):S25-S30. [doi: 10.1177/02537176211022224] [Medline: 34732951]
17. van Geenhuizen M. A framework for the evaluation of living labs as boundary spanners in innovation. Environ Plann C: Politics Space 2018 Nov;36(7):1280-1298. [doi: 10.1177/2399654417753623]
18. Verloo H, Lorette A, Rosselet Amoussou J, et al. Using Living Labs to explore needs and solutions for older adults with dementia: scoping review. JMIR Aging 2021 Aug 19;4(3):e29031. [doi: 10.2196/29031] [Medline: 34420916]
19. Thordardottir B, Malmgren Fänge A, Lethin C, Rodriguez Gatta D, Chiatti C. Acceptance and use of innovative assistive technologies among people with cognitive impairment and their caregivers: a systematic review. Biomed Res Int 2019;2019:9196729. [doi: 10.1155/2019/9196729] [Medline: 30956989]
20. Zipfel N, Horreh B, Hulshof CTJ, de Boer A, van der Burg-Vermeulen SJ. The relationship between the living lab approach and successful implementation of healthcare innovations: an integrative review. BMJ Open 2022 Jun 28;12(6):e058630. [doi: 10.1136/bmjopen-2021-058630] [Medline: 35768105]
21. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
22. Creswell JW. Qualitative Inquiry and Research Design: Choosing among Five Approaches, 2nd edition: Sage Publications, Inc; 2007.
23. Pigot H, Giroux S. Living Labs for designing assistive technologies. Presented at: 2015 17th International Conference on E-health Networking, Application & Services (HealthCom); Oct 14-17, 2015. [doi: 10.1109/HealthCom.2015.7454493]
24. Boman IL, Lundberg S, Starkhammar S, Nygård L. Exploring the usability of a videophone mock-up for persons with dementia and their significant others. BMC Geriatr 2014 Apr 16;14(1):49. [doi: 10.1186/1471-2318-14-49] [Medline: 24739662]
25. Brankaert R, den Ouden E. The design-driven Living Lab: a new approach to exploring solutions to complex societal challenges. Technol Innov Manag Rev 2017;7(1):44-51. [doi: 10.22215/timreview/1049]
26. Brankaert R, Ouden ED, Brombacher A. Innovate dementia: the development of a Living Lab protocol to evaluate interventions in context. Info 2015 Jun 8;17(4):40-52. [doi: 10.1108/info-01-2015-0010]
27. Brankaert R, Snaphaan L. Stay in touch: an in context evaluation of a smartphone interface designed for people with dementia. In: Ambient Assisted Living and Daily Activities: Springer International Publishing; 2014. [doi: 10.1007/978-3-319-13105-4_42]
28. Bühler F, Leinen A, Schneider-Schelte H, et al. Exploring stakeholders' perspectives on the adaptation of the Living Lab approach to dementia care: a qualitative study. Geriatr Nurs (Lond) 2023 Nov;54:184-191. [doi: 10.1016/j.gerinurse.2023.09.008]
29. Callari TC, Moody L, Saunders J, Ward G, Woodley J. Stakeholder requirements for an ethical framework to sustain multiple research projects in an emerging Living Lab involving older adults. J Empir Res Hum Res Ethics 2020 Jul;15(3):111-127. [doi: 10.1177/1556264619873790] [Medline: 31530076]

XSL•FO
RenderX

30. Geerts I, Snaphaan L, Bongers IMB. Involving people with dementia and their informal caregivers in the development process of assistive technology: description and evaluation of a user-centered design (preprint). JMIR Formative Research. Preprint posted online on Feb 19, 2020. [doi: 10.2196/preprints.18319]
31. Pino M, et al. Contribution of the Living Lab approach to the development, assessment and provision of assistive technologies for supporting older adults with cognitive disorders. Studia informatica Universalis 2014;11:34-62.
32. Sacco G, Noublanche F, Blazek F, et al. How to deal with the consent of adults with cognitive impairment involved in European geriatric living labs? Philos Ethics Humanit Med 2021 Jun 16;16(1):3. [doi: 10.1186/s13010-021-00101-1] [Medline: 34130730]
33. Smith G, Dixon C, Neiva Ganga R, Greenop D. How do we know co-created solutions work effectively within the real world of people living with dementia? Learning methodological lessons from a co-creation-to-evaluation case study. Int J Environ Res Public Health 2022 Nov 2;19(21):14317. [doi: 10.3390/ijerph192114317] [Medline: 36361197]
34. Snaphaan L, Geerts I, Bongers IMB. Involving people with dementia in the development process of assistive technology: multi-stakeholder experiences of a user-driven living lab. Design for Health 2022 Jan 2;6(1):28-43. [doi: 10.1080/24735132.2022.2058827]
35. Toso F, Brankaert R, Hendriks N, Lenaerts L, Wilkinson A. Reflecting on Living Labs as multi-stakeholder collaborative networks to evaluate technological products for people living with dementia. Int J Environ Res Public Health 2023 Jan 17;20(3):1673. [doi: 10.3390/ijerph20031673] [Medline: 36767050]
36. Vasileiadis M, Giakoumis D, Votis K, Tzovaras D. A living lab infrastructure for investigating activity monitoring needs in service robot applications. 2016 Jun 29 Presented at: PETRA '16; Jun 29, 2016; Corfu Island Greece p. 1-7 URL: https://dl.acm.org/doi/proceedings/10.1145/2910674 [accessed 2026-03-16] [doi: 10.1145/2910674.2935830]
37. Dell'Era C, Landoni P. Living Lab: a methodology between user ‐ centred design and participatory design. Creat Innov Manage 2014 Jun;23(2):137-154 [FREE Full text] [doi: 10.1111/caim.12061]
38. Huang JH, Thomas E. A review of Living Lab research and methods for user involvement. Technol Innov Manag Rev 2021;11:88-107. [doi: 10.22215/timreview/1467]
39. Götzelmann TG, Strech D, Kahrass H. The full spectrum of ethical issues in dementia research: findings of a systematic qualitative review. BMC Med Ethics 2021 Mar 26;22(1):32. [doi: 10.1186/s12910-020-00572-5] [Medline: 33771131]
40. Griffiths S, Shepherd V, Volkmer A. Determining capacity of people with dementia to take part in research: an electronic survey study of researcher confidence, competence and training needs. BMC Med Ethics 2024 May 28;25(1):65. [doi: 10.1186/s12910-024-01056-6] [Medline: 38802779]
41. Pappadà A, Chattat R, Chirico I, Valente M, Ottoboni G. Assistive technologies in dementia care: an updated analysis of the literature. Front Psychol 2021;12:644587. [doi: 10.3389/fpsyg.2021.644587] [Medline: 33841281]
42. West E, Stuckelberger A, Pautex S, Staaks J, Gysels M. Operationalising ethical challenges in dementia research-a systematic review of current evidence. Age Ageing 2017 Jul 1;46(4):678-687. [doi: 10.1093/ageing/afw250] [Medline: 28104596]
43. Compagnucci L, Spigarelli F, Coelho J, Duarte C. Living Labs and user engagement for innovation and sustainability. J Clean Prod 2021 Mar;289:125721. [doi: 10.1016/j.jclepro.2020.125721]
44. Frank L, Jennings LA, Petersen RC, et al. Participation of persons with dementia and their caregivers in research. J Am Geriatr Soc 2021 Jul;69(7):1784-1792. [doi: 10.1111/jgs.17340] [Medline: 34245586]
45. Boyle LD, Husebo BS, Vislapuu M. Promotors and barriers to the implementation and adoption of assistive technology and telecare for people with dementia and their caregivers: a systematic review of the literature. BMC Health Serv Res 2022 Dec 23;22(1):1573. [doi: 10.1186/s12913-022-08968-2] [Medline: 36550456]
46. Frishammar J, Essén A, Bergström F, Ekman T. Digital health platforms for the elderly? Key adoption and usage barriers and ways to address them. Technol Forecast Soc Change 2023 Apr;189:122319. [doi: 10.1016/j.techfore.2023.122319]
47. Wilson M, Doyle J, Turner J, Nugent C, O'Sullivan D. Designing technology to support greater participation of people living with dementia in daily and meaningful activities. Digit Health 2024;10:20552076231222427. [doi: 10.1177/20552076231222427] [Medline: 38235415]
48. Arntzen C, Holthe T, Jentoft R. Tracing the successful incorporation of assistive technology into everyday life for younger people with dementia and family carers. Dementia (London) 2016 Jul;15(4):646-662. [doi: 10.1177/1471301214532263] [Medline: 24784941]
49. Dugstad J, Eide T, Nilsen ER, Eide H. Towards successful digital transformation through co-creation: a longitudinal study of a four-year implementation of digital monitoring technology in residential care for persons with dementia. BMC Health Serv Res 2019 Jun 10;19(1):366. [doi: 10.1186/s12913-019-4191-1] [Medline: 31182093]
50. Egan KJ, Pot AM. Encouraging innovation for assistive health technologies in dementia: barriers, enablers and next steps to be taken. J Am Med Dir Assoc 2016 Apr 1;17(4):357-363. [doi: 10.1016/j.jamda.2016.01.010] [Medline: 27016431]
51. Gibson G, Dickinson C, Brittain K, Robinson L. Personalisation, customisation and bricolage: how people with dementia and their families make assistive technology work for them. Ageing Soc 2019 Nov;39(11):2502-2519. [doi: 10.1017/S0144686X18000661]
52. Malmgren Fänge A, Carlsson G, Chiatti C, Lethin C. Using sensor-based technology for safety and independence - the experiences of people with dementia and their families. Scand J Caring Sci 2020 Sep;34(3):648-657. [doi: 10.1111/scs.12766] [Medline: 31614031]

XSL•FO
RenderX

53. Gibson G, Dickinson C, Brittain K, Robinson L. The everyday use of assistive technology by people with dementia and their family carers: a qualitative study. BMC Geriatr 2015 Jul 24;15:89. [doi: 10.1186/s12877-015-0091-3] [Medline: 26205957]
54. Ratna H. The importance of effective communication in healthcare practice. Harv Public Health Rev 2019;23(23) [FREE Full text] [doi: 10.54111/0001/W4]
55. Song Y, MacEachern L, Doupe MB, et al. Influences of post-implementation factors on the sustainability, sustainment, and intra-organizational spread of complex interventions. BMC Health Serv Res 2022 May 17;22(1):666. [doi: 10.1186/s12913-022-08026-x] [Medline: 35581651]
56. Verbeek H, Zwakhalen SMG, Schols J, Kempen G, Hamers JPH. The Living Lab In ageing and long-term care: a sustainable model for translational research improving quality of life, quality of care and quality of work. J Nutr Health Aging 2020;24(1):43-47. [doi: 10.1007/s12603-019-1288-5] [Medline: 31886807]
57. Whitney P. Design and the economy of choice. She Ji 2015;1(1):58-80. [doi: 10.1016/j.sheji.2015.09.001]
58. From plan to impact v: WHO global action plan: the time to act is now. : Alzheimer's Disease International; 2022 URL: https://www.alzint.org/resource/from-plan-to-impact-v/#:~:text=From%20Plan%20to%20Impact%20V%20(Full%20version%2C%20English) [accessed 2026-03-16]
59. David MCB, Del Giovane M, Wilson D, et al. Considerations for legal, ethical, and effective practice in dementia research. Brain Commun 2024;6(4):fcae211. [doi: 10.1093/braincomms/fcae211] [Medline: 38983620]
60. Soofi H. Ethical aspects of facilitating the recruitment of people with dementia for clinical trials: a call for further debate. Br J Clin Pharmacol 2022 Jan;88(1):22-26. [doi: 10.1111/bcp.14968] [Medline: 34192378]

Abbreviations

ENoLL: European Network of Living Labs
LL: Living Lab
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by M O'Connell; submitted 11.Jun.2025; peer-reviewed by C Talbot, S Mohanadas; revised version received 24.Nov.2025; accepted 17.Dec.2025; published 31.Mar.2026.

Please cite as:
Alves S, Duarte N, Figueiredo T, Midao L, Bessa MJ, Rodriguéz O, Fidalgo J, Fernandez I, Felpete A, Rodriguéz M, Bernardéz JC, Facal D, Carrilho J, Costa E
Identifying Facilitators and Obstacles in Piloting Dementia Initiatives Within a Living Lab Approach: Systematic Review
JMIR Aging 2026;9:e77752
URL: https://aging.jmir.org/2026/1/e77752
doi:10.2196/77752

© Sara Alves, Natália Duarte, Teodora Figueiredo, Luis Midao, Maria João Bessa, Olalla Rodriguéz, José Fidalgo, Irene Fernandez, Alba Felpete, Maxi Rodriguéz, Juan Carlos Bernardéz, David Facal, Joana Carrilho, Elísio Costa. Originally published in JMIR Aging (https://aging.jmir.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Key Challenges and Barriers to Digital Literacy for Older Adults: Scoping Review

Salman Khan[1], MPH; Sarah Webster[2], MHS; John Puxty[2], MD; Madison Robertson[2], PhD

[1]Public Health Sciences, Queen's University, Kingston, ON, Canada

[2]Centre for Studies in Aging and Health, Department of Medicine, Queen's University, Etherington Hall, Room 3046, 94 Stuart Street, Kingston, ON, Canada

Corresponding Author:
Madison Robertson, PhD
Centre for Studies in Aging and Health, Department of Medicine, Queen's University, Etherington Hall, Room 3046, 94 Stuart Street, Kingston, ON, Canada

Abstract

Background: Despite rising internet use in Canada, older adults continue to face significant barriers in adopting and using digital technologies. Digital literacy among older adults extends beyond technical proficiency, encompassing adaptation to new technologies, overcoming age-related limitations, and addressing socioeconomic disparities. Limited digital skills hinder social participation, access to essential services, and engagement with eHealth technologies, exacerbating disparities in health outcomes.

Objective: This scoping review aimed to identify key barriers to digital literacy among older adults, synthesizing evidence from existing literature to provide a comprehensive understanding of the challenges older adults face in adopting and utilizing digital technologies.

Methods: A systematic search was conducted across the Ovid MEDLINE, PubMed, Web of Science, and EBSCOhost databases to identify studies published in the past 10 years that examined digital literacy barriers among older adults. Data extraction captured population characteristics, study context, research methods, and outcomes, which were thematically analyzed to identify overarching categories and themes.

Results: A total of 22 studies met the inclusion criteria. Most studies used a qualitative design (n=16, 72.72%), with the remaining studies comprising 3 quantitative studies (13.63%) and 3 reviews (13.63%). Seven primary barriers to digital literacy among older adults were identified: (1) health barriers, (2) support networks, (3) convenience and ease of use, (4) knowledge and information, (5) perception barriers, (6) resource barriers, and (7) barriers for special populations. Within these themes, 18 associated categories provided a more detailed understanding of the challenges older adults face in adopting and using digital technologies.

Conclusions: These findings demonstrate that digital literacy challenges among older adults are shaped by interacting individual, social, and contextual factors rather than isolated skill deficits. Barriers often co-occur, with health-related limitations, access to support, technology design, and resource constraints jointly influencing digital engagement. This review highlights the need for integrated and inclusive digital literacy strategies that address compounded barriers and reflect the diverse circumstances of older adults. Future research should focus on establishing consistent definitions of digital literacy, evaluating the long-term impact of digital inclusion initiatives, and addressing the needs of subpopulations experiencing heightened or intersecting barriers.

(JMIR Aging 2026;9:e80647) doi:10.2196/80647

KEYWORDS

digital literacy; older adults; digital divide; eHealth; internet; technology; digital inclusion

Introduction

Background

Canada, with its diverse and aging population, faces unique challenges and opportunities in promoting digital inclusion for older adults. The older adult population in Canada is expected to grow by 68% between 2017 and 2037; the number of individuals aged 75 and older will double by 2037 [1]. These numbers highlight the urgency of addressing the unique needs and challenges of this population. Despite the high overall internet penetration in the country, a significant digital divide persists among older adults. The term “digital divide” refers to the various social disparities in access to and use of digital tools and services, such as personal computers and smartphones [2].

While digital literacy is essential for everyone in today’s technology-driven world, it is particularly crucial for older adults. The term “digital literacy” encompasses the skills required to use the internet and information and communication

XSL•FO
RenderX

technologies (ICT) effectively, such as computers, smartphones, and social media [2]. For older adults, digital literacy often extends beyond mere technical proficiency. This includes understanding and adapting to new technologies, as well as overcoming barriers related to age-specific physical and cognitive challenges and socioeconomic factors that can limit access to digital tools. Research findings have consistently identified that key barriers to internet use among older adults include age, education, income, and computer-related factors [2,3]. As the digital landscape continues to evolve—marked by the widespread adoption of smartphones, social media, and online services—understanding how digital literacy for older adults is defined and applied across various contexts has become increasingly critical.

In Canada, while internet use among older adults has increased from 32.2% in 2007 to 68.2% in 2016, substantial socioeconomic and access-related disparities persist [4]. Factors, such as preretirement exposure to technology, health status, education, living arrangements, and household income, significantly influence internet use among this demographic. Certain vulnerable populations are disproportionately affected by this digital divide. For example, older adults with higher education levels, better health, and those living with others are more likely to be digitally active, while those with less education, poorer health, and living alone are less likely to use the internet [4]. These disparities highlight the need for targeted digital literacy initiatives tailored to the needs of older adults. Addressing these unique barriers can enhance digital engagement, improve older adults' overall well-being, and enable greater participation in our increasingly digital world.

Benefits and Barriers to Digital Literacy

Achieving digital literacy is vital for social inclusion, accessing information, and managing health care through eHealth technologies [3]. However, older adults often face significant barriers to accessing and using digital technologies. Older adults represent a unique demographic in this regard due to factors such as age-related declines in physical and cognitive abilities, limited prior exposure to technology, and varying levels of education and income [2]. The lack of motivation and the belief that digital technologies are not relevant to their needs can further hinder older adults' engagement with these technologies, particularly in the context of digital literacy acquisition [4].

eHealth, which refers to health-related electronic technologies, particularly the internet, is one area where digital literacy is crucial for older adults [4]. eHealth literacy involves the ability to seek, understand, and apply health information from electronic sources. As more health care services move online, including booking and attending virtual appointments, accessing test results, and managing prescriptions, eHealth is increasingly essential for older adults to maintain autonomy and access necessary care [2]. However, the adoption of eHealth technologies remains low among older adults due to barriers such as design issues, cost, and lack of awareness [5]. A systematic review conducted identified a gap in the literature on eHealth literacy among underserved populations, including older adults [6]. While specific to the United States, this study emphasized the need for targeted research and interventions to improve digital health literacy in these groups.

The advantages of digital literacy for older adults are numerous. With improved digital skills, older adults gain better access to essential information, including health care resources, and can strengthen their social ties, positively influencing their mental and emotional well-being [3]. Additionally, engaging with ICT can stimulate cognition by offering access to games, news, and educational opportunities, which may help delay cognitive decline [3]. By improving digital literacy, age-related disparities can be reduced, fostering connections across diverse groups and promoting knowledge exchange [3]. However, despite these significant benefits, the substantial disparities in digital literacy among older adults in Canada, exacerbated by various demographic factors, highlight an urgent need to systematically explore and identify the most effective strategies and best practices for mitigating these barriers. By doing so, we can ensure that older adults are better equipped to obtain these critical benefits.

Although several reviews have explored older adults' engagement with digital technologies, most have focused on narrow subpopulations or specific digital domains. Prior syntheses have commonly examined digital health adoption, targeted specific older adult groups (eg, those with chronic conditions), or evaluated the impacts and barriers of particular digital literacy interventions, rather than conceptualizing digital literacy as a broader construct [7-10]. For example, Hepburn et al [11] reviewed barriers to digital health technology adoption among older adults with chronic diseases, emphasizing health-related contexts rather than general digital skill acquisition. Similarly, another review focused specifically on older adults' use of eHealth technologies, without addressing the broader range of skills, contexts, and supports that constitute digital literacy more generally [12]. Further research is needed to examine digital literacy as a multidimensional construct. To date, no review has comprehensively synthesized barriers and facilitators to digital literacy among the general older adult population across contexts, underscoring the need for an updated and inclusive scoping review.

Purpose

This scoping review synthesizes existing literature on barriers to digital literacy among older adults and addresses the question: *What are the key challenges and barriers to digital literacy among older adults across different settings?* Framed using the population, concept, and context framework, the population of interest is older adults, the concept is challenges and barriers to digital literacy, and the context includes any relevant setting (eg, community, health care, home). Challenges and barriers were conceptualized broadly to include individual, social, technological, and contextual factors, including health-related considerations where relevant. A scoping review was chosen to map the breadth of existing evidence, identify gaps, and include diverse study designs, which would not be feasible with a systematic or narrative review.

XSL•FO
RenderX

Methods

Overview

This scoping review was conducted and reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines [13]. The PRISMA-ScR guided the review process, including the development of the search strategy, study selection, data charting, and synthesis of results (Checklist 1).

Search Strategy

The search strategy for this scoping review involved a comprehensive and systematic approach to identify relevant studies using multiple databases. A librarian from Queen's University was consulted for this review. The databases used for this review included Ovid MEDLINE, PubMed, Web of Science, and EBSCOhost. The initial search was conducted using Ovid MEDLINE and PubMed databases in June 2024, with an updated search completed in December 2025 of all 4 databases. Search strategies combined controlled vocabulary terms (eg, Medical Subject Headings [MeSH]) and keyword searching, using Boolean operators (AND/OR) to capture concepts related to digital literacy (eg, "digital literacy," "digital divide," "digital inclusion," "computer literacy"), barriers and challenges (eg, "barriers," "challenges"), and the population of interest (eg, "older adults," "seniors"). Search terms were adapted for each database to align with database-specific indexing and syntax (Multimedia Appendix 1). The reference lists of all included studies were also hand-searched to identify additional relevant literature. The review was limited to studies published within the past 10 years (2014 to the date of the search) to ensure relevance to contemporary digital technologies and evolving conceptualizations of digital literacy.

Eligibility Criteria

This review included studies that (1) were conducted in any region, (2) included participants aged 50 years and older, (3) directly or indirectly examined digital literacy (including eHealth digital literacy), explicitly using this term or closely related synonyms, and (4) reported barriers to digital literacy. The age threshold was chosen to ensure that no relevant studies were missed and to align with previous studies that have similarly defined older adults with this age range to capture a broader range of experiences and challenges associated with aging and technology use [14,15].

This scoping review included peer-reviewed studies encompassing qualitative, quantitative, and mixed methods research studies, including systematic reviews, meta-analyses, and scoping reviews. Non–peer-reviewed studies, opinion pieces, and noncredible sources were not considered. Additionally, only studies and documents available in English were included to facilitate accessibility and comprehension.

Study or Source of Evidence Selection

The study selection process used a structured, multistep approach to ensure the inclusion of the most relevant and high-quality studies. All identified documents were imported into Covidence [16]. Duplicates were removed, and titles and abstracts were initially screened by the primary investigator against the predefined inclusion and exclusion criteria. To minimize selection bias, a secondary investigator independently reviewed all titles and abstracts, as well as the full texts of the included and excluded studies, to confirm consistency with the inclusion and exclusion criteria. Discrepancies were resolved through discussion or, if needed, consultation with a third independent investigator. Although formal interrater reliability statistics (eg, Cohen kappa) were not calculated, this double-review and resolution process provided a high degree of interrater reliability and ensured the accuracy and consistency of study selection. The results of this search were recorded using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [17].

Data Extraction

Data extraction was conducted to capture all relevant information from the selected studies. Key details, such as population characteristics (eg, age, gender, socioeconomic status), study context (eg, geographic location, setting), research methods, and outcomes (eg, levels of digital literacy, barriers identified, effectiveness of interventions), were systematically extracted by the primary investigator. A secondary investigator reviewed all extracted information for accuracy. The extracted information was organized into an Excel sheet (2019, version 2407) to facilitate easy comparison and analysis.

Data Analysis and Presentation

The synthesis of the extracted data followed a systematic, iterative process informed by thematic analysis as outlined by Braun and Clarke [18]. Initially, the findings were grouped into categories of statements that reflected similar meanings across studies. These categories were then further analyzed and synthesized into overarching themes representing different types of barriers to digital literacy among older adults. Throughout this process, all decisions regarding category formation, theme development, and coding were documented to maintain an audit trail, ensuring transparency and reproducibility. This approach allowed us to identify 7 key barriers and associated contextual factors, which are presented in a coherent narrative format highlighting the main findings and insights from the review. The methodological quality of the included studies was appraised using the Joanna Briggs Institute (JBI) critical appraisal tools, with the specific checklist applied according to each study's design (eg, qualitative, quantitative, or review). Quality appraisal was conducted to support the interpretation of the findings rather than to exclude studies.

Ethical Considerations

As this study is a scoping review of publicly available literature and does not involve human participants or identifiable data, ethics approval was not required in accordance with the Queen's University Research Ethics Board. An a priori internal protocol was established and used to guide the conduct of this scoping review. In accordance with current methodological guidance (eg, PRISMA-ScR, JBI Manual), prospective registration is not required for scoping reviews. The review was conducted as planned, and no substantive deviations from the internal protocol were made.

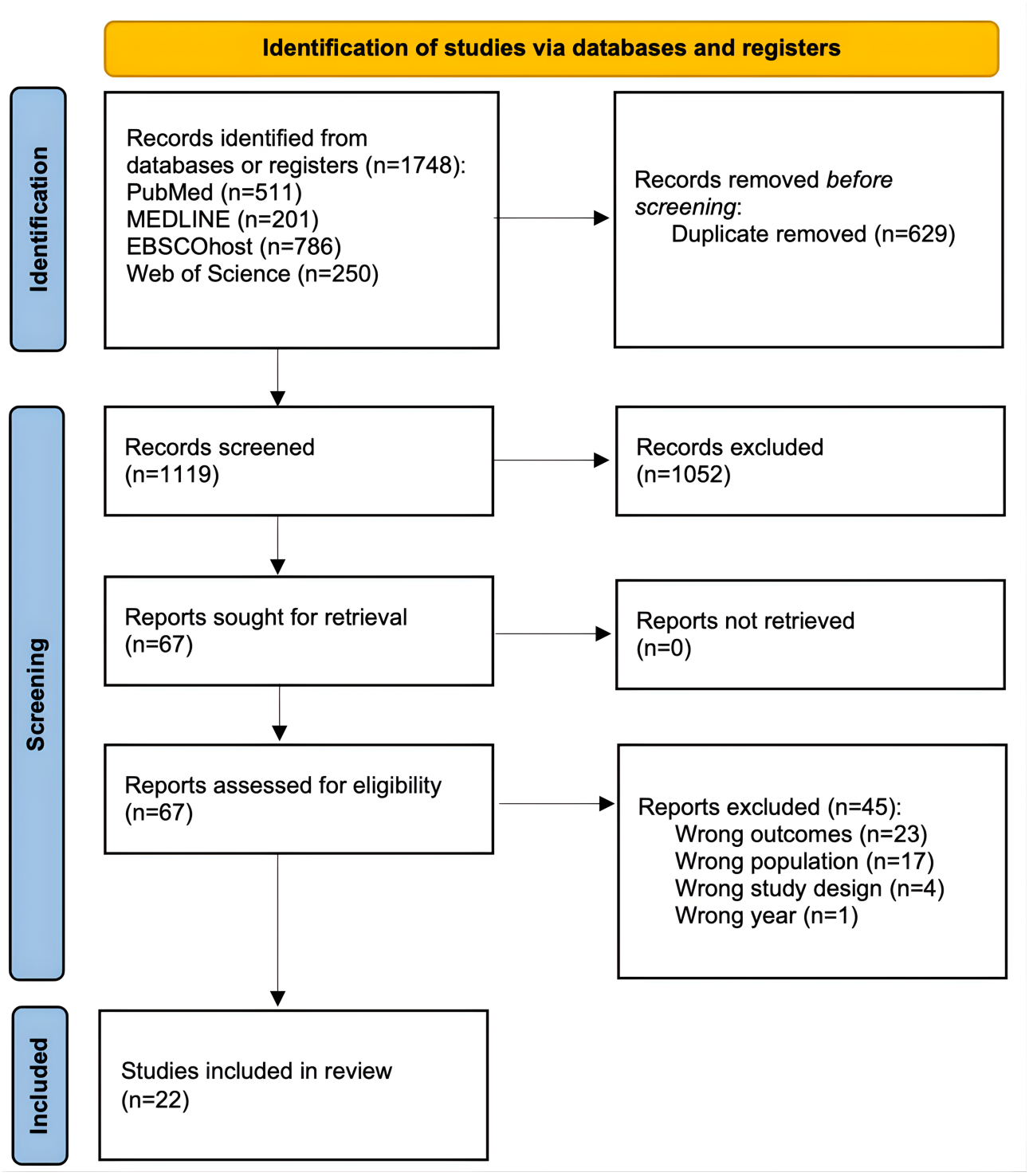
XSL•FO
RenderX

Results

Overview

This scoping review resulted in the identification of 1748 studies across 4 primary databases: PubMed (n=511), Ovid MEDLINE (n=201), Web of Science (n=786), and EBSCOhost (n=250). After electronically removing 629 duplicates, 1119 studies were screened against the inclusion and exclusion criteria at the title and abstract level. At this initial screening stage, 1052 studies were excluded. A total of 67 studies were assessed for eligibility through full-text review, leading to the exclusion of 45 studies for the following reasons: wrong outcomes (n=23), wrong population (n=17), wrong study design (n=4), and publication older than 10 years (n=1). Ultimately, 22 studies met the inclusion criteria and were included in the final review. A full outline of the search and selection process is provided in Figure 1.

Figure 1. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [17].



XSL•FO
RenderX

Study Demographics

Table 1 presents an overview of the characteristics of the studies included in this scoping review, highlighting the year of study publication, the region where the study was conducted, the study design, and the population age characteristics. Half (n=11, 50%) of the studies were recently published between 2022 and 2025. Geographically, the included studies were conducted across diverse global contexts. Four (18.18%) studies were conducted across multiple regions, while 7 (31.81%) studies were classified in the "Other" category, encompassing research conducted in South Korea, Malaysia, Norway, Turkey, Costa Rica, Spain, and China. A majority (n=16, 72.72%) of the studies used a qualitative study design, with 3 quantitative studies (13.63%) and 3 reviews (13.63%). A summary of each study's characteristics is provided in Multimedia Appendix 2 [19-40].

Table . Study demographic information of the included studies (n=22).

Category	Count, n (%)
Year of study publication	
2022 - 2025	11 (50)
2018 - 2021	6 (27.27)
2014 - 2017	5 (22.72)
Region where study was conducted	
Other (ie, South Korea, Malaysia, Norway, Turkey, Costa Rica, Spain, China)	7 (31.81)
Multiple regions or international	4 (18.18)
United States	3 (13.64)
United Kingdom	2 (9.09)
Canada	2 (9.09)
Singapore	2 (9.09)
Australia	2 (9.09)
Study design	
Qualitative	16 (72.72)
Review	3 (13.63)
Quantitative	3 (13.63)

Quality Appraisal

Among the 22 studies included in this review, the overall methodological quality ranged from high to moderate quality, with no studies being excluded on the basis of their quality. Using the appropriate JBI critical appraisal tools [41], 6 studies were appraised as high quality, 12 as moderate quality, and 3 were moderate-unclear quality due to insufficient methodological reporting. The high-quality studies (n=6) were primarily qualitative interview or focus group designs that demonstrated strong alignment between research questions, methodology, and data analysis and reported strategies to enhance rigor, such as multiple coders, audit trails, and triangulation. These studies provided robust, contextually grounded evidence on digital literacy barriers experienced by older adults. The majority (n=13) of studies were appraised as moderate quality, most commonly due to limited reporting of researcher reflexivity, small or context-specific samples, and restricted transferability to more digitally excluded populations. Several review and cross-sectional studies relied on self-reported measures and cross-sectional designs, limiting causal interpretation of identified barriers. Studies rated as moderate-unclear quality (n=3) were retained in the review because these studies still contributed valuable descriptive and contextual insights relevant to understanding barriers among older adults.

Primary Findings

Seven primary barriers of digital literacy among older adults were identified, including health barriers, support networks, convenience and ease of use, knowledge and information, perception barriers, resource barriers, and barriers for special populations. Within these 7 overarching themes, 18 associated categories were identified and provide a more detailed understanding of the specific challenges faced by older adults in adopting and using digital technologies. Consistent with scoping review methodology, the following sections thematically synthesize the 7 identified themes and categories, illustrating the diverse barriers to digital literacy among older adults.

Health Barriers

Health-related barriers are a significant factor impeding older adults' interaction with digital technologies. This theme encompasses 1 category that explores the physical and cognitive challenges faced by older adults when engaging with digital devices.

Evidence from the research shows older adults encounter significant health-related barriers that impede their interaction

XSL•FO
RenderX

with digital technologies. Across studies, age-related physical, sensory, and cognitive changes (eg, vision impairment, reduced dexterity, memory challenges, fatigue, or chronic disease burden) limited older adults' ability to learn, navigate, and sustain use of digital technologies [19-21,29,30,37,38]. In studies focused on digital health and self-management, chronic conditions compounded existing barriers by increasing cognitive load and task complexity, particularly where technologies required ongoing monitoring, interpretation of health information, and sustained engagement over time [24,35,40]. Cognitive impairments, such as memory loss and forgetfulness, were also frequently discussed barriers to digital literacy for older adults [34,35,38,39]. These barriers were frequently described as context-dependent, becoming more pronounced when digital tools were poorly aligned with aging-related needs or fluctuating health conditions, particularly in digital health and accessibility contexts [24,29,30,35,40].

Support Networks

Support networks are essential in shaping older adults' ability to engage with digital technologies. This theme explores 2 key categories: the impact of social support and social influence on digital literacy and technology adoption among older adults. While social support refers to the direct assistance and encouragement provided by family, friends, and community services, social influence encompasses the broader attitudes and behaviors within one's social network that can shape technology use.

The studies reviewed highlight the crucial role that support networks play in helping older adults navigate the digital landscape. However, many older adults lack sufficient social support from family, friends, or community services, which exacerbates their digital literacy challenges [20,21,26,28,32-35,37,39]. Dependence on younger family members for learning new digital skills is not always reliable, as negative interactions, such as lack of patience, can reinforce feelings of inadequacy and discourage use [22,26,33,34,38]. The absence of family or social networks often results in older adults being less likely to adopt internet use [20,21,26,33,34].

Evidence from the reviewed literature shows that social influence is pivotal in the adoption of digital technologies among older adults. Positive or negative attitudes within social networks can impact older adults' engagement with technology [19,24,29,34]. The encouragement or discouragement from family members and peers can either facilitate or hinder the adoption of digital tools [22,28,29,34,35,38].

Convenience and Ease of Use

The convenience and ease of use of digital technologies are crucial factors that determine their adoption and sustained use among older adults. This theme is explored through two key categories: (1) the impact of application interface and design and (2) the challenges posed by continuous technological changes.

The studies reviewed underscore that the design and interface of digital applications are critical determinants of their usability for older adults. Complex interfaces, small font sizes, poor sound quality, and multiple authentication requirements pose significant challenges because they fail to account for the physical and cognitive changes associated with aging, such as declining vision, hearing loss, and reduced memory or cognitive processing speed [22,29,30,35,37,38,40]. Several studies noted that the cumulative burden of multistep processes (ie, repeated logins, verification requirements, and navigating across multiple platforms) further increased cognitive load and reduced sustained engagement among older adults [20,21,25,35,40].

Continuous updates and changes in digital devices and websites create additional hurdles for older adults, causing confusion and a loss of confidence in their ability to use these technologies [29,30,34,37,38]. The constant evolution of technology can be overwhelming and discouraging for older users, leading to reduced engagement.

Knowledge and Information

The ability to effectively use digital technologies is heavily dependent on the knowledge and information available to older adults. This theme examines two key categories: (1) the impact of a lack of digital literacy and skills and (2) the consequences of limited awareness about digital technologies.

Research highlights that a pervasive lack of digital literacy among older adults is a significant barrier to the use of digital technology and services [19,20,22,25,26,33-40]. Many older adults did not grow up with advanced technology, making it challenging for them to acquire new digital skills [19,26,33,34,37,38]. This limited proficiency includes difficulty in navigating digital technologies, a lack of operational or technical knowledge, and minimal exposure to modern digital tools [20,22,27,33,34,37,38]. The general lack of computer proficiency hampers effective use of patient portals and other digital health platforms [25,28,31,35,40].

Many older adults are unaware of available digital technologies, services, and their benefits, which further limits their adoption of these technologies [19,20,27-29,33]. This lack of awareness hinders the effective use of digital health tools for managing health conditions. Beyond operational skills, several studies also reported difficulties appraising the credibility and relevance of online information, particularly in digital health contexts, which hindered confident use of digital services [20,21,25,31,35].

Perception Barriers

Overview

Perceptions and attitudes toward digital technologies play a crucial role in shaping older adults' willingness and ability to adopt and use these tools. This theme encompasses 4 key categories: concerns related to security, privacy, fear, and anxiety; challenges of self-efficacy and motivation; social and psychological barriers; and a preference for traditional methods over digital interactions.

Security, Privacy, Fear, and Anxiety

Concerns about online security and privacy are significant deterrents for older adults when considering the use of digital technology and services. Many older adults fear privacy breaches and distrust digital platforms, which prevent them

XSL•FO
RenderX

from engaging with these technologies [19,21,27,29,32]. This fear is particularly pronounced among rural residents, who express significant worries about breaches of privacy and confidentiality, especially with technologies that monitor daily activities and collect personal data [35]. Beyond privacy concerns, the fear and anxiety older adults experience around technology use often stem from broader security issues. Reports of online scams have heightened these concerns, leading many older adults to avoid digital engagement and digital health services altogether [21,27].

In addition to privacy concerns, older adults often experience anxiety about constant monitoring and the implications of being watched or recorded without consent, further increasing their reluctance to engage with technology [29]. The mistrust of technology, coupled with the complexity of digital tools, exacerbates their anxiety, making digital engagement even more daunting [21,35]. Several studies also reported fear of making irreversible mistakes or damaging devices, which contributed to the avoidance of independent technology use [21,27,32,36,38,39].

Self-Efficacy and Motivation

Low self-efficacy and motivation significantly impact older adults' engagement with digital technologies. Many older adults lack confidence in their ability to use telemedicine and other digital tools due to age-related issues, such as poor eyesight, memory loss, and unfamiliarity with technology [21,23,32,36]. Additionally, older adults' beliefs that they are too old to learn new technologies or that digital tools are unnecessary for daily tasks further hinder their willingness to engage with these tools [23,35]. Internalized age-related stereotypes about technology use further undermined confidence, reinforcing beliefs that digital technologies were better suited to younger generations [19,23,26,34].

Social and Psychological Barriers

The adoption of digital technologies among older adults is often hindered by concerns related to social isolation, stigmatization, and threats to independence. Some older adults fear that reliance on smart technologies could lead to reduced human interaction and increased loneliness, with worries that increased digital interaction might replace physical activity and social engagement [27,29,36]. Additionally, the use of assistive technologies can sometimes lead to stigmatization, as older adults may fear being labeled as dependent or frail, which further discourages their use [29]. Concerns about personal autonomy are also prevalent, with older adults expressing hesitation toward technologies that might impose rigid structures or routines, thereby undermining their sense of independence [29].

Preference for Traditional Methods

Many older adults prefer traditional methods of communication, such as letters and in-person consultations, instead of digital interactions. They value the personal connection and the comprehensiveness of physical examinations, which they believe digital tools cannot replicate [19,29]. The study found a low demand for telemedicine services among rural residents, often due to a preference for face-to-face visits and discomfort with online consultations [29,35].

Resource Barriers

Access to and the affordability of digital technologies are critical challenges that significantly hinder older adults' ability to engage with digital tools. This theme explores 3 key categories: limited access to technology; financial constraints; and the lack of personalized training, all of which contribute to the digital divide among older adults.

Limited access to digital devices and the internet is a significant barrier for many older adults [27,37,39]. This includes issues such as unreliable internet access and lack of necessary equipment, including smartphones or computers [19,20,27,39,40]. Many rural residents particularly struggle with a lack of reliable internet access, which is critical for telemedicine adoption [39]. Not possessing smartphones or having limited access to reliable internet and other digital resources further exacerbates these challenges [20,27]. Residential facilities often lack the necessary infrastructure, hindering older adults' ability to use digital health portals [27].

The affordability of digital devices and internet services is a substantial barrier. Many older adults have limited financial resources to purchase and maintain digital devices or pay for internet access [37,39,40]. The high cost of devices, maintenance, and indirect costs, such as electricity consumption, poses significant obstacles [39,40]. Lower income levels are associated with reduced internet use among older adults [20,21,24,26,40].

The absence of personalized training and ineffective teaching methods further hinder older adults' ability to learn and use digital technologies by failing to accommodate their specific learning needs. Rigid and nontailored instructional approaches can lead to frustration, disengagement, and a lack of confidence, making the learning process overwhelming and exacerbating the digital divide [24,33,38,40].

Barriers for Special Populations

Overview

In addition to the general barriers faced by older adults in adopting digital technologies, certain special populations encounter unique challenges that further complicate their ability to achieve digital literacy. Across studies, barriers affecting special populations were rarely experienced in isolation; instead, they reflected the interaction of health-related, perceptual, usability, and resource constraints, often compounding digital exclusion.

Visually Impaired Older Adults

Visually impaired older adults often lack awareness about the potential of technology to promote independence and reduce loneliness. This lack of awareness limits their engagement with digital tools that could otherwise enhance their quality of life. Additionally, municipalities frequently lack the expertise and capacity to provide adequate training, leaving visually impaired individuals without the necessary skills to use digital technologies effectively [30]. Another significant barrier for visually impaired older adults is the lack of long-term support after initial training. While some may receive training on how to use digital tools, the absence of continuous support,

XSL•FO
RenderX

particularly when faced with technical problems or software updates, can result in the loss of newly acquired skills, thereby hindering sustained technology use [30]. Technological changes present a unique challenge for visually impaired older adults. Learning to use gesture-based interactions and screen readers requires extensive training. Frequent updates and the presence of inaccessible content further complicate the use of digital devices, making it difficult for visually impaired individuals to keep up with technological advancements [30]. These challenges were frequently attributed to inconsistent accessibility standards and design practices, rather than individual limitations, and often resulted in frustration and the abandonment of digital technologies.

Older Adults With Chronic Conditions

Older adults with chronic conditions experienced increased digital literacy barriers due to the combined demands of health self-management and technology use. Studies reported that symptom burden, fatigue, and fluctuating health status increased cognitive load, making it difficult to learn, interpret, and sustain engagement with digital health tools. Technologies requiring frequent monitoring, data interpretation, or consistent interaction were particularly challenging, often leading to reduced confidence and disengagement [24,35,40].

Older Adults With Cancer and Their Caregivers

Both older adults with cancer and their caregivers often lack confidence in their ability to find relevant health information online. This low self-efficacy acts as a barrier to the effective use of digital health tools [34]. Limited familiarity with digital tools and apps is common among older adults with cancer and their caregivers. Additionally, lower levels of education are associated with lower eHealth literacy, making it more challenging for these individuals to engage with digital health resources effectively [31]. The rapid evolution of digital tools creates a barrier for older adults with cancer and their caregivers, who may struggle to keep up with technological changes. This unfamiliarity with new tools can prevent them from fully utilizing digital health resources [31]. Finally, limited access to necessary resources, such as broadband, computers, and other digital devices, particularly in rural areas, further exacerbates the digital divide for older adults with cancer and their caregivers. This geographic disparity highlights the need for targeted interventions to improve digital access and literacy in underserved areas [31].

Culturally and Linguistically Diverse Older Adults

Culturally and linguistically diverse older adults experience distinct digital literacy barriers related to language proficiency, cultural relevance, and reliance on informal support networks. Limited English proficiency and unfamiliarity with technical and health-related terminology restrict the understanding of digital platforms, online instructions, and digital health information, even when access to technology is available [26,28]. These barriers are often compounded by limited access to culturally appropriate training and support, resulting in greater dependence on family members or community intermediaries for navigation and interpretation of digital tools [28]. In some cases, this reliance reduced autonomy and confidence, further discouraging independent technology use. Cultural perceptions of technology, trust, and privacy also shaped engagement, highlighting the need for culturally responsive approaches to digital literacy support [26].

Defining Digital Literacy

In addition to the 7 themes identified in this scoping review, the included studies demonstrated heterogeneous definitions of digital literacy, underscoring the absence of a unified conceptualization in the field. Given this variability, we did not impose a single operational definition for the purposes of synthesis. Instead, we documented and compared the range of conceptualizations described in the included studies to accurately reflect the existing heterogeneity in the literature. Aslan et al [19] define digital literacy as the ability to find, evaluate, and communicate information using information and communication technologies, emphasizing the core skills necessary for effective information management. Kebede et al [20] offer a comprehensive framework, outlining digital literacy as comprising 5 key indicators: information and data literacy, communication and collaboration, digital content creation, safety, and problem-solving. This structured approach reflected the multidimensional nature of digital literacy. Chee [27] focuses on the practical application, defining digital literacy as the ability to use devices and online platforms, which highlights the functional aspect of digital literacy. Finally, Baek et al [22] define digital literacy as the ability to use and understand information obtained from various digital devices, stressing the significance of comprehension within the digital literacy framework.

Discussion

Principal Findings

This scoping review identified 7 interrelated domains of barriers to digital literacy among older adults: health barriers, support networks, convenience and ease of use, knowledge and information, perception barriers, resource barriers, and barriers for special populations. Together, these themes capture a broad range of challenges, such as physical and cognitive impairments, lack of social support, complex app interfaces, limited digital literacy, security concerns, financial constraints, and a preference for traditional methods of communication. These findings highlight the intricate and multifaceted nature of the digital divide among older adults.

While digital literacy and technology adoption barriers have often been reported separately, recognizing their complex interplay is crucial. Prior research has often examined digital literacy barriers in isolation, focusing separately on technology adoption, access, or specific older adult populations [7-12]. This synthesis suggests that older adults frequently encounter multiple barriers simultaneously, aligning with ecological and life-course perspectives on technology use in later life. This pattern is consistent with ecological and life-course models of aging and technology use, which emphasize how individual, social, and environmental factors interact over time to shape engagement with digital technologies in later life [42-44]. Many of these barriers do not exist in isolation; rather, they often work together to create compounded difficulties for older adults. For instance,

XSL•FO
RenderX

limited access to technology can exacerbate feelings of anxiety and insecurity, while financial constraints may prevent access to personalized training that could alleviate these concerns [20,21]. Similarly, social isolation may intensify the impact of low self-efficacy, leading to a reluctance to engage with digital tools [24,30,38]. An example of this interplay can be seen in how physical barriers, such as poor motor skills or visual impairments, can exacerbate challenges related to convenience and ease of use [30,31,42,43]. Complex interfaces or small font sizes can make digital devices particularly difficult for those with physical limitations, further discouraging use [23,30,31,43,45]. Therefore, it is important to view digital literacy barriers holistically, understanding that older adults often experience multiple, interconnected challenges that influence their ability to use and benefit from digital technologies. Addressing these barriers effectively requires an integrated approach that considers the full spectrum of factors influencing digital literacy.

These findings extend the literature on digital literacy among older adults by synthesizing a wide range of reported barriers and demonstrating how they cluster across multiple domains, including resource constraints, perceptions of technology, and social and psychological factors. Rather than occurring in isolation, barriers were frequently described as interconnected, with multiple challenges shaping older adults' experiences of digital engagement. This synthesis supports a more nuanced understanding of digital literacy that moves beyond single-factor explanations and highlights the cumulative and interacting nature of barriers encountered in later life.

In relation to eHealth technologies, the findings highlight how barriers commonly associated with eHealth literacy, such as design complexity, cost, and limited awareness, are frequently described alongside broader influences, including social support and perceptions of technology [5,19-21,39]. This pattern suggests that challenges related to eHealth engagement are embedded within wider digital literacy contexts rather than being solely attributable to individual skill deficits. Consistent with broader eHealth literacy literature [ie, 45,46], these findings suggest that engagement with digital health technologies is shaped by intersecting usability, access, and psychosocial factors rather than individual competence alone. Taken together, these observations point to the importance of considering multiple, interacting factors when examining eHealth literacy among older adults.

Defining Digital Literacy

Despite the increasing recognition of digital literacy as a critical competency for older adults, this review identified substantial variability in how digital literacy is defined and operationalized across studies. Definitions ranged from basic functional skills to more comprehensive frameworks encompassing information evaluation, communication, safety, and problem-solving. This heterogeneity complicates cross-study comparison and limits the development of targeted interventions.

Consistent with prior work, including the review by Oh et al [47], many commonly used measurement tools emphasize narrow aspects of digital literacy while overlooking broader competencies such as digital content creation and online safety. Although frameworks, such as the digital competence model, propose multidimensional conceptualizations, few studies apply these comprehensively in research with older adults. The lack of comprehensive assessment tools limits the ability to accurately evaluate digital literacy and develop targeted interventions.

The absence of a universally accepted definition of digital literacy complicates research efforts and impacts how digital literacy is assessed across studies. Varying definitions contribute to fragmented approaches, with some emphasizing technical proficiency, while others incorporate cognitive and social dimensions. This inconsistency leads to measurement gaps, as most widely used tools fail to capture the full scope of digital skills needed for meaningful engagement with technology. For instance, while some assessments, such as the Mobile Device Proficiency Questionnaire, evaluate all 5 digital competence competencies, many omit crucial aspects, such as digital content creation and online safety [47]. Without a standardized framework, it is difficult to compare findings across studies or develop interventions that effectively address older adults' digital literacy needs [47].

Future Directions

To address conceptual variability in the literature, future research would benefit from explicitly defining how digital literacy is conceptualized within the context of older adulthood or, at minimum, clearly acknowledging heterogeneity in existing definitions when designing studies. Greater conceptual clarity, supported by the development and adoption of standardized and validated measurement frameworks, would enhance the comparability and coherence of findings across studies. Establishing a shared foundation for defining and measuring digital literacy would also strengthen policy and practice efforts by ensuring that digital inclusion initiatives align with the real-world needs of older adults [47].

Future research should prioritize the development and evaluation of tailored interventions that address the specific barriers identified in this review [25,37]. Longitudinal studies are needed to assess the sustained impact of digital literacy programs on older adults' quality of life and health outcomes. Interdisciplinary research that integrates gerontology, technology design, and education can offer a more comprehensive approach to enhancing digital literacy among older adults. Collaboration with community organizations, policymakers, and technology developers will be essential in designing and implementing effective, scalable interventions.

Additionally, researchers should examine the barriers faced by specialized populations, such as older adults with visual impairments, chronic conditions, and those from rural or low-income backgrounds. These groups encounter unique challenges that often amplify difficulties in engaging with digital technologies. For instance, older adults with visual impairments struggle with the lack of accessible content and frequent technological changes, which can render screen readers and gesture-based interactions ineffective without extensive training [31]. Addressing these specialized needs through inclusive technology design and tailored interventions is crucial to ensure

XSL•FO
RenderX

that all older adults can achieve digital literacy and benefit from technological advancements.

This review highlights the need for multifaceted interventions that simultaneously address the interconnected barriers older adults face [20,21,25,30]. Future studies should explore approaches that integrate digital literacy training with accessible and affordable technology solutions. Additionally, research should evaluate the effectiveness of personalized and community-based approaches to digital literacy, particularly those that involve collaboration with educators, social workers, and family members. Investigating these strategies will provide evidence-based recommendations for scalable interventions that not only improve digital literacy among older adults but also enhance their overall quality of life.

Practical Implications

The findings of the scoping review have several practical implications for policymakers, practitioners, and technology developers. By highlighting the interconnected nature of digital literacy barriers, this review can support the development and funding of comprehensive interventions that address multiple challenges simultaneously, rather than focusing on isolated barriers. Policymakers should consider creating and funding programs that increase the accessibility and affordability of digital devices and internet services for older adults. Consistent with the findings of this review, affordability remains a significant barrier, with many older adults unable to purchase or maintain digital devices or pay for reliable internet access [20,21,25,27,45]. Addressing cost-related barriers is a necessary foundation for improving digital participation among older adults, particularly those living on fixed or limited incomes.

Practitioners can leverage insights from this review to design and implement community-based digital literacy programs that address the diverse and intersecting needs of older adults. The results highlight the importance of personalized, ongoing, and confidence-building approaches to training, particularly for individuals with limited prior exposure to digital technologies or reduced self-efficacy [25,37].

Technology developers play a critical role in reducing digital exclusion by prioritizing user-centered designs that accommodate age-related sensory, cognitive, and motor changes. The findings emphasize the need for intuitive interfaces, simplified navigation, and accessible design features, such as readable text, clear audio, and reduced cognitive load [23,30,31]. In addition, providing continuous education, guidance, and support related to technological updates can help older adults maintain confidence and sustained engagement, particularly as frequent changes were shown to undermine usability and self-efficacy [21,30,31,38]. By implementing these coordinated strategies, policymakers, practitioners, and technology developers can work together to address the multifaceted barriers identified in this review and support more equitable digital participation among older adults.

Limitations

Several limitations of the scoping review should be acknowledged. First, the exclusion of non-English studies and older publications may have limited the scope of insights into global digital literacy barriers among older adults. However, this decision was made to ensure consistency in data interpretation and to focus on more recent studies that reflect the rapidly evolving digital landscape. Additionally, this review focused solely on the experiences of older adults, excluding perspectives from individuals who play a key role in digital literacy education, such as tutors, librarians, and other educators. This approach was chosen to center firsthand accounts of the barriers faced by older adults. While the perspectives of educators and other stakeholders are undoubtedly valuable, the scope of this review was intentionally narrowed to prioritize firsthand accounts from the target population. As such, future research should consider incorporating educators' insights to provide a more comprehensive understanding of the barriers to digital literacy among older adults, ultimately informing more effective, tailored interventions.

Conclusion

The scoping review provides a comprehensive synthesis of the key barriers to digital literacy among older adults, highlighting the intricate and multifaceted nature of these challenges. The review identified 7 primary themes (ie, health barriers, support networks, convenience and ease of use, knowledge and information, perception barriers, resource barriers, and barriers for special populations) that collectively shape the digital experiences of older adults. These barriers often interact, potentially increasing the challenges older adults face in engaging with digital technologies. These findings suggest that engagement with digital technologies is shaped by a complex interplay of individual, social, and contextual factors. Future work could explore strategies that address these barriers while considering diverse abilities, levels of support, and contextual constraints. Additionally, research would benefit from greater consistency in defining digital literacy and examining how multifaceted approaches may support digital participation among subpopulations facing heightened barriers.

Funding

This work is funded in part by the Ministry for Seniors and Accessibility in Ontario, Canada.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Search strategy.
[DOCX File, 25 KB - aging_v9i1e80647_app1.docx]

Multimedia Appendix 2
Included study characteristics [19-40].
[DOCX File, 48 KB - aging_v9i1e80647_app2.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 673 KB - aging_v9i1e80647_app3.pdf]

References

1. Canada's seniors population outlook: uncharted territory [Infographic]. Canadian Institute for Health Information. 2017. URL: https://www.cihi.ca/en/infographic-canadas-seniors-population-outlook-uncharted-territory [accessed 2026-03-09]
2. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279]
3. Marcotte AA. Digital literacy interventions for older adults in Canada: visions for change policy challenge 2022–2023. : Agewell National Innovation Hub; 2024 URL: https://agewell-nih-appta.ca/wp-content/uploads/2023/04/Marcotte-Digital-Literacy-Final-Report.pdf [accessed 2026-03-02]
4. Study: evolving internet use among Canadian seniors. Statistics Canada. 2019. URL: https://www150.statcan.gc.ca/n1/daily-quotidien/190710/dq190710d-eng.htm [accessed 2026-03-02]
5. Arcury TA, Sandberg JC, Melius KP, et al. Older adult internet use and eHealth literacy. J Appl Gerontol 2020 Feb;39(2):141-150. [doi: 10.1177/0733464818807468] [Medline: 30353776]
6. Chesser A, Burke A, Reyes J, Rohrberg T. Navigating the digital divide: a systematic review of eHealth literacy in underserved populations in the United States. Inform Health Soc Care 2016;41(1):1-19. [doi: 10.3109/17538157.2014.948171] [Medline: 25710808]
7. Choudhary H, Bansal N. Barriers affecting the effectiveness of digital literacy training programs (DLTPs) for marginalised populations: a systematic literature review. J Tech Educ Train 2022 Jun 30;14(1):111-127. [doi: 10.30880/jtet.2022.14.01.010]
8. Zhang Y, Rosli Y, Mohd Saat NZ, Abd Talib R, Tang S. Scoping review of the e-health literature in older adults. Digit Health 2025 May;11:20552076251362545. [doi: 10.1177/20552076251362545]
9. Arora S, Huda RK, Verma S, Khetan M, Sangwan RK. Challenges, barriers, and facilitators in telemedicine implementation in India: a scoping review. Cureus 2024 Aug;16(8):e67388. [doi: 10.7759/cureus.67388] [Medline: 39310647]
10. Bobrowicz-Campos EM, da Mota Matos AP. Scoping review of barriers and facilitators to the successful implementation of education programs that promote digital literacy among older adults in the community. Sisyphus: J Educ 2025;13(2):139-167. [doi: 10.25749/sis.39536]
11. Hepburn J, Williams L, McCann L. Barriers to and facilitators of digital health technology adoption among older adults with chronic diseases: updated systematic review. JMIR Aging 2025 Sep 11;8:e80000. [doi: 10.2196/80000] [Medline: 40934502]
12. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
13. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
14. Kokorelias KM, Grigorovich A, Harris MT, et al. Coadaptation between smart technologies and older adults over time: protocol for a scoping review. JMIR Res Protoc 2023 Oct 9;12:e51129. [doi: 10.2196/51129] [Medline: 37812466]
15. Martinez L, Mirza RM, Austen A, et al. More than just a room: a scoping review of the impact of homesharing for older adults. Innov Aging 2020;4(2):igaa011. [doi: 10.1093/geroni/igaa011] [Medline: 32382662]
16. Covidence. URL: https://www.covidence.org [accessed 2026-03-02]
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
18. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
19. Aslan A, Mold F, van Marwijk H, Armes J. What are the determinants of older people adopting communicative e-health services: a meta-ethnography. BMC Health Serv Res 2024 Jan 11;24(1):60. [doi: 10.1186/s12913-023-10372-3] [Medline: 38212713]
20. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]

XSL•FO
RenderX

21. Money A, Hall A, Harris D, Eost-Telling C, McDermott J, Todd C. Barriers to and facilitators of older people's engagement with Web-Based services: qualitative study of adults aged >75 years. JMIR Aging 2024 Feb 28;7:e46522. [doi: 10.2196/46522] [Medline: 38416543]
22. Baek J, Kim H, Choi S, et al. Digital literacy and associated factors in older adults living in urban South Korea: a qualitative study. Comput Inform Nurs 2024 Mar 1;42(3):226-239. [doi: 10.1097/CIN.0000000000001109] [Medline: 38300124]
23. Köttl H, Gallistl V, Rohner R, Ayalon L. "But at the age of 85? Forget it!": internalized ageism, a barrier to technology use. J Aging Stud 2021 Dec;59:100971. [doi: 10.1016/j.jaging.2021.100971] [Medline: 34794716]
24. Chang J, McAllister C, McCaslin R. Correlates of, and barriers to, Internet use among older adults. J Gerontol Soc Work 2015;58(1):66-85. [doi: 10.1080/01634372.2014.913754] [Medline: 24941050]
25. Low STH, Sakhardande PG, Lai YF, Long ADS, Kaur-Gill S. Attitudes and perceptions toward healthcare technology adoption among older adults in Singapore: a qualitative study. Front Public Health 2021;9:588590. [doi: 10.3389/fpubh.2021.588590] [Medline: 33659232]
26. Tan KSY, Chan CML. Unequal access: applying Bourdieu's practice theory to illuminate the challenges of ICT use among senior citizens in Singapore. J Aging Stud 2018 Dec;47:123-131. [doi: 10.1016/j.jaging.2018.04.002] [Medline: 30447865]
27. Chee SY. Age-related digital disparities, functional limitations, and social isolation: unraveling the grey digital divide between baby boomers and the silent generation in senior living facilities. Aging Ment Health 2024 Apr;28(4):621-632. [doi: 10.1080/13607863.2023.2233454] [Medline: 37424361]
28. Hyman A, Stacy E, Mohsin H, et al. Barriers and facilitators to accessing digital health tools faced by South Asian Canadians in Surrey, British Columbia: community-based participatory action exploration using photovoice. J Med Internet Res 2022 Jan 13;24(1):e25863. [doi: 10.2196/25863] [Medline: 35023842]
29. Street J, Barrie H, Eliott J, et al. Older adults' perspectives of smart technologies to support aging at home: insights from five world café forums. Int J Environ Res Public Health 2022 Jun 25;19(13):7817. [doi: 10.3390/ijerph19137817] [Medline: 35805477]
30. Fuglerud KS, Tunold S, Kjæret K. Social contact for older people with visual impairment through mastery of smartphones: barriers and suggested solutions. Stud Health Technol Inform 2021 Jun 4;282:415-428. [doi: 10.3233/SHTI210417] [Medline: 34085985]
31. Verma R, Saldanha C, Ellis U, Sattar S, Haase KR. eHealth literacy among older adults living with cancer and their caregivers: a scoping review. J Geriatr Oncol 2022 Jun;13(5):555-562. [doi: 10.1016/j.jgo.2021.11.008] [Medline: 34810146]
32. Balcı E, Aslan GK. Older adults' experiences and perceptions of using digital technology for health purposes: a qualitative study. Geriatr Nurs 2025;65:103490. [doi: 10.1016/j.gerinurse.2025.103490] [Medline: 40633234]
33. Castro Rojas MD. Barriers and supportive factors for older adults learning about and using information and communication technologies for healthy aging in Costa Rica. Educ Gerontol 2021 Oct 3;47(10):433-446. [doi: 10.1080/03601277.2021.1989228]
34. Schreurs K, Quan-Haase A, Martin K. Problematizing the digital literacy paradox in the context of older adults' ICT use: aging, media discourse, and self-determination. Can J Commun 2017 May 23;42(2):359-377. [doi: 10.22230/cjc.2017v42n2a3130]
35. Shao Y, Yang X, Chen Q, et al. Determinants of digital health literacy among older adult patients with chronic diseases: a qualitative study. Front Public Health 2025;13:1568043. [doi: 10.3389/fpubh.2025.1568043] [Medline: 40231178]
36. Fernández-Piqueras R, González-García RJ, Sanz-Ponce R, Calero-Plaza J. Digital divides in older people: assessment of digital competencies and proposals for meaningful inclusion. Eur J Investig Health Psychol Educ 2025 Sep 26;15(10):196. [doi: 10.3390/ejihpe15100196] [Medline: 41149147]
37. Gitlow L. Technology use by older adults and barriers to using technology. Phys Occup Ther Geriatr 2014 Sep;32(3):271-280. [doi: 10.3109/02703181.2014.946640]
38. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 4;8:1687. [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
39. Baker S, Warburton J, Hodgkin S, Pascal J. The supportive network: rural disadvantaged older people and ICT. Ageing Soc 2017 Jul;37(6):1291-1309. [doi: 10.1017/S0144686X16000350]
40. Cajita MI, Hodgson NA, Lam KW, Yoo S, Han HR. Facilitators of and barriers to mHealth adoption in older adults with heart failure. Comput Inform Nurs 2018 Aug;36(8):376-382. [doi: 10.1097/CIN.0000000000000442] [Medline: 29742549]
41. JBI manual for evidence synthesis. Joanna Briggs Institute. 2020. URL: https://synthesismanual.jbi.global [accessed 2026-03-02]
42. Bronfenbrenner U. The Ecology of Human Development: Experiments by Nature and Design: Harvard University Press; 1979. [doi: 10.2307/j.ctv26071r6]
43. Bronfenbrenner U, Morris PA. The bioecological model of human development. In: Damon W, Lerner RM, editors. Handbook of Child Psychology, 6th edition: John Wiley & Sons; 2006, Vol. 1:793-828. [doi: 10.1002/9780470147658]
44. Czaja SJ, Charness N, Fisk AD, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352. [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]

XSL•FO
RenderX

45. van Dijk J. The Digital Divide: Polity Press; 2020. URL: https://openlibrary.org/books/OL29490934M/Digital_Divide [accessed 2026-03-02]
46. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9. [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
47. Oh SS, Kim KA, Kim M, Oh J, Chu SH, Choi J. Measurement of digital literacy among older adults: systematic review. J Med Internet Res 2021 Feb 3;23(2):e26145. [doi: 10.2196/26145] [Medline: 33533727]

Abbreviations

ICT: information and communication technologies
JBI: Joanna Briggs Institute
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by F Puga; submitted 14.Jul.2025; peer-reviewed by B Ong, K Butler-Henderson, M Rigby; revised version received 08.Jan.2026; accepted 19.Feb.2026; published 16.Mar.2026.

Please cite as:
Khan S, Webster S, Puxty J, Robertson M
Key Challenges and Barriers to Digital Literacy for Older Adults: Scoping Review
JMIR Aging 2026;9:e80647
URL: https://aging.jmir.org/2026/1/e80647
doi:10.2196/80647

© Salman Khan, Sarah Webster, John Puxty, Madison Robertson. Originally published in JMIR Aging (https://aging.jmir.org), 16.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying and Prioritizing Age-Friendly Design Principles and Guidelines for Developing Transportation Planning E-Tools: Scoping Review

Sara Bahrampoor Givi, PhD; Mireille Gagnon-Roy, OTR, PhD; Véronique Provencher, OTR, PhD

School of Rehabilitation, Faculty of Medicine and Health Sciences, Université de Sherbrooke, 3001, 12e Avenue Nord, Sherbrooke, QC, Canada

Corresponding Author:
Sara Bahrampoor Givi, PhD
School of Rehabilitation, Faculty of Medicine and Health Sciences, Université de Sherbrooke, 3001, 12e Avenue Nord, Sherbrooke, QC, Canada

Abstract

Background: Older adults often face mobility challenges and usability barriers when navigating transportation options due to age-related physical, cognitive, and sensory changes. While transportation planning e-tools can support their independence, most are not designed for their specific needs. There is a lack of comprehensive, age-friendly usability design principles tailored to this context.

Objective: This study aims to identify, synthesize, and prioritize the most relevant age-friendly usability design principles and guidelines for developing transportation planning e-tools that are tailored to the needs of older adults.

Methods: A scoping review was conducted following the Arksey and O'Malley methodological framework, enhanced by guidance from the Joanna Briggs Institute and reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) standards. The literature search was performed across the following six scientific databases: MEDLINE (PubMed and National Center for Biotechnology Information), AgeLine (EBSCOhost), Cochrane Library (Wiley Online Library), Scopus (Elsevier), IEEE Xplore, and TRID (Transportation Research Board), and was supplemented by gray literature identified through Google Scholar and Google Search, covering the period from January 2013 to April 2025. To guide the analysis, 4 foundational usability frameworks by Nielsen; Shneiderman and Plaisant; Gerhardt-Powals; and Weinschenk and Barker were used to inductively derive generic principles that structured the classification of age-friendly guidelines. This process resulted in the organization of extracted guidelines into 10 core usability principles. The analytic hierarchy process was applied by a single expert to rank the principles based on expert pairwise comparisons and their frequency of reference in literature.

Results: The study identified 31 relevant studies. From these, 500 age-friendly guidelines were refined into 68 actionable guidelines across 10 usability principles: visual clarity, structure and navigation, ease of use, information, minimizing memory load, feedback, accessibility, consistency, simplicity, and control. The analytic hierarchy process ranked visual clarity (36.4%), structure and navigation (22.1%), and ease of use (12.5%) as the top 3 age-friendly design priorities.

Conclusions: This study offers an evidence-based foundation for developing transportation planning e-tools that promote older adults' autonomy and digital inclusion, with prioritized guidelines applicable to transportation planning e-tools and to the broader field of age-friendly digital design. Ongoing updates and active user involvement are essential to ensure their sustained usability and long-term relevance.

(JMIR Aging 2026;9:e83668) doi:10.2196/83668

KEYWORDS

older adults; usability; guidelines; review; transportation; e-tools

Introduction

Background

As people age, they often encounter mobility-related difficulties, which can significantly impact their ability to engage in society [1]. These difficulties can make everyday activities, such as attending medical appointments or participating in social and recreational events, increasingly challenging [2]. Reduced physical capacity, changes in vision [3] and balance [4], as well as the onset of chronic illnesses [4], contribute to these mobility issues. As a result, many older adults find it difficult to maintain the independence and social connections that are crucial to their overall well-being [5]. Addressing these challenges requires a comprehensive approach, including accessible transportation options that help older adults remain active and engaged in their

XSL•FO
RenderX

daily lives, despite steep hills and heavy snowfall, which are common in Nordic countries like Canada.

In the province of Quebec (Canada), data from 2019 show that approximately 64% of older adults drive and 23% travel as passengers, while about 9% use taxis, adapted transport, or public transit, and only 3% rely on walking, cycling, or mobility devices. Men were more likely than women to drive (80% vs 50%), while women more often traveled as passengers (35% vs 10%) and used taxis, paratransit, or public transit more frequently (12% vs 7%) [6]. Although only a small proportion of older adults use public or shared transportation, studies indicate that this limited use does not necessarily reflect a lack of interest but rather barriers such as safety concerns, difficulties accessing clear travel information, and a lack of confidence or comfort when using available options like buses—often due to fears of becoming disoriented or encountering unwanted social interactions [7,8].

To mitigate these challenges, there is a growing need for transportation planning e-tools that can assist older adults in accessing reliable information about their travel options. E-tools (eg, route-planning systems) could help drivers feel safer by avoiding hazardous routes, such as steep hills or poorly lit roads, particularly in challenging conditions [9]. However, engaging with such digital tools often proves difficult for this population. Age-related changes such as reduced visual sensitivity [10] can make it hard to accurately select interactive elements, thereby disrupting the overall trip-planning experience [11]. In addition, cognitive factors, such as reduced attention, can make it more difficult to navigate digital interfaces [12]. Consequently, the design of transportation planning e-tools must consider both the functional changes associated with aging and the diversity in digital competencies among older adults [13]. It is therefore essential to adapt these tools, particularly their interfaces, to better meet the needs of older adults.

Given the notable lack of studies specifically focused on developing and evaluating transport planning interfaces tailored to the needs and preferences of older adults through a participatory design approach [14], it becomes increasingly important to identify comprehensive age-friendly design principles and usability guidelines. While a number of design frameworks exist for general usability [15], few explicitly address the cognitive, physical, and perceptual changes associated with aging [16,17]. Without clear, context-specific design guidance, there is a risk that digital solutions will continue to fall short of accessibility and usability standards for this population. A systematic synthesis of existing guidelines is relevant to ensure that future e-tool interfaces are not only inclusive but also practical, empowering older adults to confidently plan and manage their trips. It is therefore essential to fill the gap by identifying, consolidating, and prioritizing existing age-friendly usability design principles to inform the development of transportation planning e-tools that are both effective and attuned to the diverse capabilities of older users.

Objective

This study aims to identify, synthesize, and prioritize the most relevant age-friendly usability design principles and guidelines for developing transportation planning e-tools that are tailored to the needs of older adults.

Methods

Study Design

This scoping review was conducted using the methodological framework proposed by Arksey and O'Malley [18], which includes five stages: (1) identifying the research question; (2) identifying relevant sources; (3) selecting studies; (4) charting the data; and (5) collating, summarizing, and reporting the results. Methodological guidance from the Joanna Briggs Institute was also integrated to enhance the rigor of the review. The reporting process adheres to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines (Checklist 1) [19]. In the second phase, the analytic hierarchy process (AHP) [20] was used to prioritize the age-friendly usability design principles identified through the scoping review.

Scoping Review: Identifying the Research Question

The study is guided by the following research question: What are the most relevant age-friendly usability design principles and guidelines for developing transportation planning e-tools tailored to the needs of older adults? This question was structured using the population, concept, and context (PCC) framework, where the population is older adults, the concept focuses on age-friendly usability design principles and guidelines, and the context involves the development of transportation planning e-tools.

Identifying the Relevant Sources

This stage consisted of searching scientific and gray literature published in English from January 2013 to May 2023, with an updated search conducted twice, in April 2024 and April 2025. This timeframe was selected to capture the most recent decade of research on age-friendly usability guidelines and transportation planning e-tools, reflecting rapid advances in digital technologies. The search was performed across six (6) scientific databases: MEDLINE (PubMed and National Center for Biotechnology Information), AgeLine (EBSCOhost), Cochrane Library (Wiley Online Library), Scopus (Elsevier), IEEE Xplore, and TRID (Transportation Research Board). To supplement the scientific literature, gray literature was identified through Google Scholar and Google Search, with the first 50 results from each platform screened for inclusion to ensure relevance and manageability.

The development of the search strategy followed a systematic and collaborative approach. Initially, authors compiled a list of relevant databases and preliminary keywords based on the study objectives. This preliminary version was then critically reviewed by a professional librarian, who contributed methodological expertise throughout the process. His recommendations were informed by the PRESS (Peer Review of Electronic Search Strategies) 2015 guidelines [21], leading to substantial improvements in keyword formulation, the use of Boolean operators, and the integration of controlled vocabulary terms. To improve the effectiveness of the search strategy, the professional librarian suggested running preliminary tests to

XSL•FO
RenderX

evaluate how well the selected terms captured relevant studies. Based on the outcomes, he recommended several refinements. For example, the terms “Graphical User Interface” and “GUI” were added to ensure that studies using different terminology could be retrieved. These adjustments aimed to reduce the number of irrelevant results and to ensure that key studies were not overlooked.

Controlled vocabulary, such as Medical Subject Headings terms in PubMed, was mapped and complemented by a tailored list of free-text terms to reflect diverse indexing practices and author phrasing across databases. After executing the search, the results were imported into Zotero (Corporation for Digital Scholarship) for reference management. Initial deduplication was performed using Zotero’s automated function, followed by a thorough manual screening by the first author (SBG) to remove any remaining duplicate records. The final search strategy, including both controlled vocabulary and keywords, is presented in Table 1.

Table . Keywords and concepts.

Concepts	Keywords	Controlled vocabulary (by database)
Aging	“Older adult” OR “Old people” OR “Elderly” OR “Senior” OR “Aging” OR “Normal aging” OR “Frailty” OR “Aged” OR Old*	• MEDLINE or Cochrane Library: aged (MeSH[a]) • AgeLine: aged, older adults • TRID: elderly persons (thesaurus)
E-tool	“Website” OR “Application” OR “E-tool” OR “Web application” OR “App” OR “Web”	• MEDLINE or Cochrane Library: mobile applications (MeSH) • IEEE Xplore: software applications (IEEE thesaurus) • TRID: websites (TRID thesaurus)
Guideline	“Design principles” OR “Web design principle” OR “Design guideline” OR “Web design guideline” OR “Usability guideline” OR “User interface design guideline” OR “UI design guideline” OR “User experience design guideline” OR “UX design guideline” OR “GUI” OR “Graphical user interface”	• MEDLINE or Cochrane Library: user-computer interface (MeSH) and guidelines as topic (MeSH) • AgeLine: human-computer interaction and usability engineering • TRID: Human factors engineering
Transport	“transport* OR “public transport*” OR mobility* OR travel OR “public transit” OR “active transport*” OR “alternative transport” OR paratransit OR bus* OR carpool*	• MEDLINE or Cochrane Library: transportation (MeSH) • AgeLine: transportation • TRID: transit, mobility, and public transportation

[a]MeSH: medical subject headings.

Selecting the Studies

To be eligible for inclusion, sources needed to address usability guidelines specifically aimed at older adults.

Sources were excluded if they met any of the following criteria: (1) written in a language other than English and (2) focused on e-tools other than websites or apps.

The first author (SBG) conducted the initial screening of all identified sources based on titles and abstracts, followed by a full-text review to assess their eligibility. To mitigate potential bias and strengthen the reliability of the process, 20% of the study selection was independently reviewed and validated by a co-author (VP). Specifically, the co-author independently reviewed 20% of the records retained after the title and abstract screening stage, as well as 20% of the studies assessed at the full-text eligibility stage, to verify inclusion decisions and ensure consistency.

Charting the Data

To chart the data, 4 pioneering sets of general usability guidelines were selected: Nielsen [15], Shneiderman and Plaisant [22], Gerhardt-Powals [23], and Weinschenk and Barker [24]. These foundational works have stood the test of time, are widely cited in the fields of human-computer interaction and user experience, and have been extensively validated across diverse domains and user groups. Each set offers a distinct yet complementary perspective on usability: Nielsen outlined 10 usability heuristics, Shneiderman and Plaisant introduced 8 golden rules of interface design, Gerhardt-Powals proposed 10 cognitive principles for interface development, and Weinschenk and Barker organized their recommendations into 11 thematic categories. It is important to note that these frameworks were identified before screening relevant studies and were not part of the scoping review dataset. Rather, they were used to develop an initial analytical framework (charting form) to guide the systematic extraction and categorization of age-friendly usability guidelines from included studies. This approach ensured that the analysis was structured, consistent, and comprehensive, while preventing selection bias.

To develop the charting form, the first author (SBG) thematically analyzed the 4 pioneering sets of selected general usability guidelines using the Braun and Clarke [25] inductive thematic analysis approach. Guidelines from key authors were compiled into a cross-table and grouped into broader categories

XSL•FO
RenderX

based on emerging patterns. Redundant or overlapping items were removed or integrated to ensure clarity and uniqueness. These refined categories were then defined as 10 core usability principles associated with 19 generic usability guidelines (Table 2).

Table . Interface design principles with associated generic guidelines.

Principles and guidelines	Shneiderman	Nielsen	Weinschenk and Barker	Gerhardt‐Powals
Feedback				
Informative feedback	✓	✓	✓	
Confirmation of the actions that are done successfully	✓		✓	✓
Let users be aware of results			✓	
Consistency				
Present new information in familiar frameworks and structure				✓
Consistency in using familiar icons and colors	✓	✓		✓
Minimizing the memory load				
Memory load avoidance	✓	✓		✓
Avoid changes in familiar behavior to speed up the action of the experienced users	✓			
Simplicity				
Simple error message (not codes)	✓	✓	✓	
Simple and relevant information (eliminate unnecessary information to reduce the use of the resource in multitasks)		✓		✓
Simple layout			✓	✓
Accessibility				
Consider user profile in design (eg, digital literacy)	✓	✓		
Flexibility in providing information for a wide range of users with different cognitive abilities, in different tasks, for both experienced and novice users		✓		✓
Structure and navigation				
Error prevention by implying target limitation (eg, do not allow the user to enter numeric characters in alphabetic fields)	✓	✓	✓	
Data grouping				✓
Control				
Ability to control by undo, redo, or exit	✓	✓	✓	
Command and control model to let users know how they should answer the questions and enter the information			✓	
Ease of use				
Easy to access help and support		✓		✓
Visual clarity				
Use graphics as much as possible instead of text				✓
Information				
Use names that are conceptually related to function				✓

XSL•FO
RenderX

Subsequently, a structured Microsoft Excel grid was developed to chart the sources identified through this scoping review. In this grid, source characteristics (including publication year and study context), e-tool types, and age-friendly guidelines, categorized according to usability principles derived from the 4 pioneering general usability frameworks mentioned earlier, were captured. Similar to the generic usability design guidelines, the Braun and Clarke [25] thematic analysis framework was used. Since the generic usability guidelines served as the foundation for developing age-friendly usability guidelines, a deductive approach was chosen. In this context, all the identified guidelines were compiled into a cross-table. This process involved reading through the guidelines to understand their content and context, ensuring that each entry was accurately represented in a table. In the next step, each guideline was labeled with a principle that best represented its content, facilitating the organization of the guidelines into manageable pieces for further analysis. Subsequently, guidelines with the same labels were grouped together. The labeled guidelines were then reviewed and refined to ensure accurate clustering. Redundant guidelines within each group were identified and removed to simplify the dataset. Additionally, similar guidelines were integrated to reduce overlap and ensure that each guideline was distinct and relevant. Finally, the process concluded with final reporting. This process resulted in a clear and distinct set of guidelines.

Collecting, Summarizing, and Reporting Results

A combination of descriptive-analytical methods and thematic analysis [25] was applied to examine all data categories, including source characteristics, e-tool types, and age-friendly guidelines. The findings were then synthesized and interpreted to map the most relevant age-friendly usability design guidelines for creating transportation planning e-tools tailored to older adults' needs, categorized in the 10 core usability principles previously defined.

Prioritization Using the AHP

To prioritize the age-friendly design principles identified through the scoping review, the AHP [20] was used. This method enables prioritization through pairwise comparisons, relying on expert judgment to establish a structured hierarchy of priorities [20]. A key advantage of AHP is its systematic approach to comparing and ranking alternatives or criteria [26]. To guide the pairwise comparisons, the Saaty's [20] fundamental scale of relative importance was used. This scale assigns numeric values to verbal judgments: 1 for equally important, 3 for moderately important, 5 for strongly important, 7 for very strongly important, and 9 for extremely important. The even numbers 2, 4, 6, and 8 serve as intermediate values between these main levels.

To validate the results of the AHP, the consistency ratio (CR) is calculated using the formula CR=CI/RI, where the consistency index (CI), derived from pairwise comparisons, is measured using the following formula and then compared with the corresponding random index (RI).

$$CI=\lambda max-nn-1$$

In this study, the AHP Online System (AHP-OS), developed by Goepel [27], was used by the first author (SBG), who has extensive experience in human-centered design and usability evaluation, to conduct pairwise comparisons of each design principle. The software aggregated these comparisons and calculated normalized priority weights. Principles with higher frequencies in the scoping review were considered more influential and were assigned higher relative importance values on the Saaty fundamental scale (eg, "moderately" to "very strongly" more important) compared with less frequently cited principles. This approach ensured that the prioritization systematically reflected both the literature evidence and expert judgment. Using this method facilitated decision-making by quantifying expert input, producing a consistent ranking of the most critical age-friendly usability design principles, and resolving conflicts between design criteria.

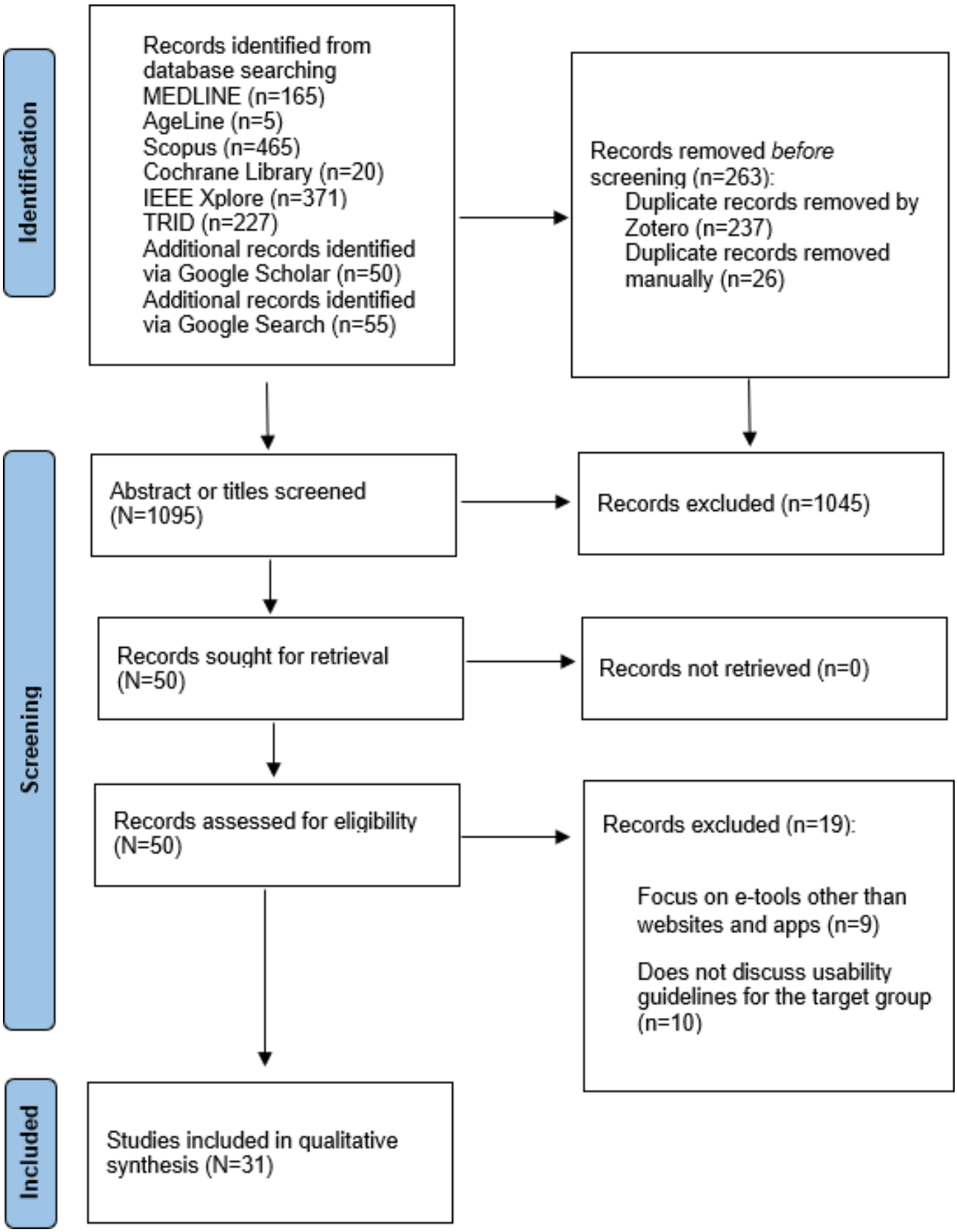
Results

Scoping Review

A total of 1358 sources were initially identified through database searches and gray literature. After removing duplicates, 1095 unique sources remained. Following the screening of titles and abstracts, 50 sources were retained. Of these, 19 did not meet the inclusion and exclusion criteria, leaving 31 sources for full-text analysis (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram detailing the process of literature search, screening, and selection.



Characteristic of Sources

Out of the 31 sources (Table 3), 26 (83.9%) clearly indicated the publication year, ranging from 2001 to 2024 [28-53]. Most (n=24, 77.4% publications) guidelines were published between 2015 and 2024 [28-30,32-44,46-53], with a peak in 2023 (n=5, 16.1% publications) [29,39,41,44,49].

XSL•FO
RenderX

Table . List of studies included in the scoping review.

Title	Year	Context of the study	E-tool type	Number of guidelines identified
Interface Design Features of Mobile Application for Senior Citizens [28]	2019	Religion	App	3
UsAge Guidelines: Toward Usable Saudi M-Government Applications for Elderly Users [29]	2023	Public services	App	22
Android OS [30]	2019	—[a]	App	5
Design and Evaluation of a Mobile User Interface for Older Adults: Navigation, Interaction and Visual Design Recommendations [31]	2014	Gaming	App	8
Designing Digital Games With & for Home-Dwelling Older Adults" Social Interaction Under Sheltering Measures [32]	2021	Gaming	App	2
UX[b]*/UI*[c] *Design for Elderly Users: Best Practices & Principles* [33]	2024	—	—	9
Guideline for Web Pages Design Inclusive for Senior Users [34]	2021	—	Website	56
Designing for an Aging Population: Toward Universal Design [35]	2016	—	Website	40
A Guide to Designing for Older Adults [36]	2024	—	—	9
Examining the Usability of Touch Screen Gestures for Older and Younger Adults [37]	2015	—	App	5
Towards Accessibility Guidelines of Interaction and User Interface Design for Alzheimer's Disease Patients [38]	2017	—	App	121
Design Guidelines of Mobile Apps for Older Adults: Systematic Review and Thematic Analysis [39]	2023	—	App	19
UX/UI Design for Elderly Users: A Comprehensive Guide [40]	2024	—	—	3
Age Inclusive Digital Platforms [41]	2023	—	—	17
Development of a Web-Based Insulin Decision Aid for the Elderly: Usability Barriers and Guidelines [42]	2017	Health	Website	25
Mobile Health Apps: Improving Usability for Older Adult Users [43]	2019	Health	App	19

XSL•FO
RenderX

Title	Year	Context of the study	E-tool type	Number of guidelines identified
Developing Usability Guidelines for mHealth Applications (UGmHA) [44]	2023	Health	App	1
Making Your Website Senior Friendly: A Checklist [45]	2001	—	Website	22
Increasing Usability of Homecare Applications for Older Adults: A Case Study [46]	2019	Health	App	5
Mobile Accessibility: How WCAG 2.0 and Other W3C/WAI Guidelines Apply to Mobile [47]	2015	—	App	16
Design of Mobile Phones for Older Adults: An Empirical Analysis of Design Guidelines and Checklists for Feature Phones and Smartphones [48]	2018	—	App	19
Evaluating User Interface and User Experience in Mobile Applications Designed for the Elderly [49]	2023	Health	App	7
Development of Universal Design Mobile Interface Guidelines (UDMIG) for Aging Population [50]	2016	—	App	5
User Interface on Smartphone for Elderly Users [51]	2017	—	App	5
A Preliminary Study on Designing Tour Website for Older People [52]	2015	Tourism	Website	14
Improving Messenger Accessibility for Elderly Users using User Centered Design (UCD) Methods (Study Case: WhatsApp) [53]	2021	Communication	App	5
Age Before Beauty—A Guide to Interface Design for Older Adults [54]	—	—	Website or app	9
A Guide to Interface Design for Older Adults [55]	—	—	Website or app	12
Factsheet On SS 618—Guidelines on User Interface Design for Older Adults [56]	—	—	Website or app	3
Designing for iOS [57]	—	—	—	9
UI Design for Elderly [58]	—	—	—	5

[a]Data not available.

[b]UX: user experience.

[c]UI: user interface.

Of the 31 sources, 11 (35.5%) specified the context of the study. Most of these contexts were related to health, representing 16.1% of the total sources [42-44,46,49]. Other identified contexts included tourism [52], communication [53], gaming [31,32], public service [29], and religion [28].

The majority of sources (n=17, 55%) focused on age-friendly usability guidelines related to developing apps, reflecting the

XSL•FO
RenderX

growing use of smartphones and mobile technology among older adults [28-32,37-39,43,44,46-51,53]. Five (16.1%) sources focused on guidelines for developing websites [34,35,42,45,52], and 3 (9.6%) sources covered both websites and apps [54-56]. Additionally, 6 (19.3%) sources did not indicate the type of e-tool [33,36,40,41,57,58].

A total of 500 guidelines were identified. Among them, visual clarity accounted for 26.6% (133 guidelines) and had the highest frequency of referral, while consistency, simplicity, and control each represented 5% (25 guidelines) and had the lowest frequency of referral. For more detailed information on the frequency of referral for each principle, see Table 4.

Table . Frequency of referral to guidelines associated with each principle.

Principles	Frequency of referral, n (%)
Visual clarity	133 (26.6)
Structure and navigation	82 (16.4)
Ease of use	63 (12.6)
Information	48 (9.6)
Minimizing the memory load	39 (7.8)
Feedback	31 (6.2)
Accessibility	29 (5.8)
Consistency	25 (5)
Simplicity	25 (5)
Control	25 (5)

After removing or integrating redundant or overlapping items to ensure clarity, 74 age-friendly usability design guidelines were identified and categorized under 10 core principles derived from general usability guidelines. The next step involved combining the guidelines identified through the generic usability design guidelines (19 guidelines) with the age-friendly usability design guidelines (74 guidelines), resulting in a total of 93 guidelines. In this process, the guidelines related to each principle were reviewed together, redundant guidelines were removed, and similar ones were combined to reduce overlap. The goal was to ensure that only the most relevant and unique guidelines were retained. As a result, from the initial 93 guidelines, 68 were finalized and categorized into the 10 design principles. Table 5 presents the final set of age-friendly usability design guidelines categorized into 10 core principles.

XSL•FO
RenderX

Table . Age-friendly usability design principles and guidelines.

Principles and guidelines	Relevant studies
Feedback	
Provide informative feedback by showing clicked links, task status, waiting periods, and error messages with resolution steps.	[15,22-24,29,34,35,38,39,41,48,49,54,57]
Request confirmation before performing actions that could cause significant changes or data loss, and confirm successful actions	[29,33,36,38,41,48]
Consistency	
	[15,22,23,33-35,38,40,42,43,45,47,57]
Maintain consistency in design elements across pages and screens such as by using familiar icons, input formats, colors, menu hierarchy, and so on.	
Present information clearly by using familiar frameworks and structures, and apply consistent navigation patterns and layouts.	[23,29,38,39,49,50,57]
Underline text only when it is a link.	[58]
Minimizing the memory load	
Present 1 task or process at a time.	[15,22,23,29,34-36,39,47]
Avoid complex dynamic input fields that require multiple gestures or the use of both hands.	[29,42,55]
Eliminate unnecessary or repetitive actions.	[29,38]
Limit the number of key functions on the home screen to 12 or fewer for tablets and 6 or fewer for smartphones.	[38,41,43,46,49,52]
Avoid changes in familiar behavior to speed up the actions of experienced users.	[22,34,35]
Use images, videos, or audio to complement text information.	[38,45]
Minimize time-sensitive content and provide options to adjust or extend time limits if necessary.	[34,35,38,39,41,42,56]
Simplicity	
Simplify data entry processes	[38,43]
Use simple and easy-to-understand graphics.	[43,48,52,55]
Ensure simplicity, clarity, and consistency in screen layout, navigation, and terminology.	[15,23,24,32,34,35,38,39,48,55]
Provide simple error messages (not codes).	[15,22,24,33,36,38,41]
Accessibility	
Consider user profiles in design such as their digital literacy, and visual and cognitive impairment.	[15,22-24,34,38,43,45,47,52]
Ensure button accessibility with a universal approach, accommodating left-handed users.	[34]
Use icons alongside text labels.	[34,38]
Include alternative data entry methods for users unable to use the primary method.	[34,35]
Ensure the app interface adapts to device rotation and user attempts to enlarge screen content.	[29]
Provide content in various multimedia formats, such as text, videos, pictures, and audio.	[30,31,34,35,38,47]
Structure and navigation	
Avoid requiring users to double-click or drag to interact with the page.	[34,36,41,45,50]
Offer both click-to-zoom and 2-finger zoom gestures.	[37,47,48]
Minimize data entry steps to ideally fewer than 3.	[43,52]
Provide help tutorials or brief animated instructions.	[29,33-35,38,46,47,52]
Clearly indicate touchable or clickable elements.	[34,35,39,47,56]
Position key elements (eg, buttons and information) at the center of the screen.	[31,34,35,38,39,41,47,48]
Guide users with systematic navigation to indicate their location and the next step.	[29,34,38,39,41-43,45,52]
Provide a virtual keyboard for numerical or alphabetical entry.	[15,22,24,38]
Use a short, simple, and static pull-down menu.	[34,38,45,48]
Include spell check for errors.	[38]

XSL•FO
RenderX

Principles and guidelines	Relevant studies
Maintain hierarchy in feedback, navigation, visual design, and data grouping.	[23,34,35,38,42,47,57]
Avoid using scroll bars, especially horizontally.	[34,35,38,45,48]
Position search tools at the top right.	[34]
Control	
Provide zoom-in ability up to 200%.	[33,38,47]
Consider the home menu as a safe return point.	[31,34,35,43]
Ability to control requests by undoing, redoing, and exiting.	[15,22,24,31,33,36,38,45,57]
Avoid disappearing messages; let users close them manually.	[36]
Offer adjustable text sizes, color schemes, and contrast to enable personalization of the user experience.	[35,38,42]
Ease of use	
Enable a function to remember usernames and passwords.	[34]
Make account creation optional for users.	[34]
Incorporate autosave features whenever possible.	[43,46]
Provide support for both orientations.	[47,57]
Include automatic data entry (eg, time and date) to reduce data entry time and keyboard use.	[29,31,35,38,39,47]
Ensure easy access to help and support.	[15,23,29,35,36,39,41,43,53,55]
Consider target touch areas of at least 9 x 9 mm, with a minimum of 8 dp of inactive space between them.	[30-34,37-39,41-43,47,48,50,54,55,58]
Visual clarity	
Avoid using fluorescent colors and colors with minimal contrast, such as green and red; yellow and blue; or green and blue.	[29,34,35,38,42,43,45,52,54]
Avoid using all capital letters words.	[34,38,45]
Consider 3:1 and 4:1 color contrast ratio for text and large text respectively.	[29,30,34,36,38-40,42,43,48,49,54,55]
Adjust the contrast range between 1.7 and 3.0.	[30,34]
Use blue, red, and magenta as the link colors on a white background due to their high visibility.	[34,36,38,43,45]
Use brightness levels of 75% and 50%.	[34,51]
Use font sizes of 12 point to 14 point for body text and 16 point for titles.	[28,33-35,38-40,43-45,47,48,50-53,55,58]
Use sans serif fonts like Arial, Helvetica, and Tahoma.	[29,34,38,45,51,54,55]
Avoid justified alignment, unfilled text, text shadows, large blocks of italic or bold text, and underlined text.	[38,42,45,54]
Maintain line spacing between 1.5 to 2.	[38,45]
Consider using bold typefaces for emphasis.	[28,38,45,54]
Avoid complex and distracting elements such as animations, blinking text, patterned backgrounds, pure white backgrounds, pure black text on white backgrounds, and transparent menus.	[23,34,38,39,42,44,45,51,57]
Information	
The website should not include pop-up messages or advertisements.	[52]
Avoid unclear or ambiguous language.	[31,33,35,38,39,41,42,45,49,52]
Do not use abbreviations or symbols.	[38]
Ensure each page has a clear and descriptive title.	[38]
Avoid using Roman numerals.	[38]
Use short sentences (less than 70 characters per line) and paragraphs (each focusing on a single idea) and avoid hyphenating words at the end of lines.	[38,39,45,54]
Highlight important information for better visibility.	[34,38]
Use active voice to enhance clarity.	[38,45]

XSL•FO
RenderX

Principles and guidelines	Relevant studies
Use names that are conceptually related to function.	[23,34,35,38,42,52,53]

Analytic Hierarchy Process

AHP was used to assess the relative importance of the 10 identified age-friendly usability design principles. This prioritization was based on expert judgment (SBG) and the frequency of references found in the scoping review. For example, visual clarity had the highest frequency of 133 (26.6%) references, whereas control had 25 (5%) references. During pairwise comparisons, this difference guided the assignment of the Saaty scale values: visual clarity was often judged as "strongly" or "very strongly" more important than principles with lower frequencies. The results indicated that visual clarity, structure and navigation, and ease of use were the most critical principles for the development of transportation planning e-tools tailored to older adults.

The CR, which measures the reliability of pairwise comparisons, was calculated at 7.5%, remaining below the acceptable threshold of 0.10 (10%) as recommended by Saaty [20]. The AHP decision matrix is presented in Multimedia Appendix 1.

According to the results, visual clarity was ranked as the highest priority, receiving 36.4% of the total weight. This indicates that guidelines associated with visual clarity should take precedence, particularly in cases where they may conflict with other guidelines or functional aspects of e-tool development. Structure and navigation and ease of use were ranked second and third, with priority weights of 22.1% and 12.5%, respectively. A complete ranking of the principles and their corresponding weights is presented in Table 6.

Table . List of principles sorted by their rank and the percentage of their priority.

Principles	Rank	Priority, %
Visual clarity	1	36.4
Structure and navigation	2	22.1
Ease of use	3	12.5
Information	4	10.5
Minimizing the memory load	5	6.3
Feedback	6	3.8
Accessibility	7	2.8
Consistency	8	2.1
Simplicity	9	1.9
Control	10	1.8

Discussion

Principal Findings

This study aimed to identify, synthesize, and prioritize age-friendly usability design principles and guidelines for transportation planning e-tools tailored to the needs of older adults. Through a comprehensive scoping review and the application of the AHP, the study provides a structured, evidence-based framework to guide the design of inclusive digital transport tools. The synthesis of existing usability guidelines identified 68 clear and actionable recommendations, which were categorized under 10 core design principles. Among these, visual clarity, structure and navigation, and ease of use emerged as the highest priorities for enhancing the usability and comfort of digital tools for this population.

In general, our results highlighted the effectiveness of these generic usability guidelines in categorizing design principles tailored to specific user groups, particularly older adults. As Ahmad et al [59] pointed out, these generic usability guidelines function as essential rules of thumb that all user interface designers should follow. Although general guidelines are widely applicable, there is significant value in customizing them to meet the specific needs of older adults when developing age-friendly e-tools. Through actionable recommendations, this customization enhances the usability and adaptability of such e-tools. This perspective aligns with Patsoule and Koutsabasis [60], who argued that identifying and validating a set of principles and guidelines for web design specifically for older adults improves the adaptability of an e-tool to their needs and preferences.

More specifically, our results revealed that the principle of visual clarity was ranked as the highest priority, followed by structure and navigation and ease of use, which were ranked second and third, respectively. This underscores the critical importance of visual clarity in designing e-tools for older adults, particularly for travel planning, which is inherently complex and involves managing multiple types of information such as destinations, schedules, and personal preferences. A clear and intuitive interface reduces cognitive load, making tasks more manageable and less stressful, thereby improving usability and encouraging long-term adoption among older adults [61]. In support of these findings, Patsoule and Koutsabasis [60] highlighted "Visibility" and "Ease of Understanding" as the top 2 principles in age-friendly heuristics for redesigning websites for older adults.

XSL•FO
RenderX

Their research aligns with the prioritization identified through the AHP methodology, further validating the emphasis on visual clarity, structure and navigation, and ease of use. These principles are vital in ensuring that older adults can effectively use e-tools without confusion or frustration.

Structure and navigation, ranked second (22.1%), further emphasizes the importance of logical content organization and user orientation within interfaces. Well-defined pathways, minimized data entry steps, and effective wayfinding features are essential to mitigate disorientation and cognitive overload [12,62]. Similarly, the third-ranked principle, ease of use (12.5%), points to the necessity of designing interfaces that are both intuitive and forgiving, offering support mechanisms such as autofill, orientation flexibility, and minimal setup requirements. At the other end of the spectrum, principles such as Control, Simplicity, and Consistency received the lowest prioritization. While their lower scores suggest they may not be as immediately critical as visual clarity or navigability, they remain essential for ensuring a robust and holistic user experience. To support this, Nahm et al [63] highlighted that "simple design with clear instructions" was a key preference among older adults, in contrast to design elements that rely on complex navigation or ambiguous features.

Given the dynamic nature of technology and the evolving needs of older adults, it is crucial to continuously refine and expand age-friendly usability design principles and guidelines. Nurgalieva et al [64] highlight that while many such guidelines exist, their adoption is often limited without ongoing efforts to make them more accessible, organized, and actionable. Furthermore, a recent systematic review of mobile app design for older adults emphasizes the importance of iterative, user-centered design, including feedback loops, adaptation to emerging devices, and integration of new accessibility features, to ensure that e-tools remain intuitive, usable, and relevant over time [65].

Limitations

Although the study highlights the most relevant age-friendly usability design principles and guidelines for transportation planning e-tools, certain limitations should be acknowledged. First, the protocol for the scoping review was not registered in advance. Although the review followed established methodological frameworks and incorporated rigorous processes for data collection and analysis, the absence of a registered protocol may limit the transparency and reproducibility of the review process. To mitigate this limitation, steps were taken to enhance transparency and minimize bias: a second reviewer validated 20% of the study selection and data extraction, the search strategy was fully documented, and a professional librarian reviewed and refined the search approach. Notably, the methodology did not include the consultation step recommended by Levac et al [66], which is often used to enhance the relevance and applicability of findings through stakeholder engagement. Nevertheless, the results of this study were extensively applied in the Mobilaînés project [67] to support the co-design and development of an age-friendly transportation planning e-tool. The implementation of this innovative e-tool helped validate the practical value and effectiveness of the guidelines identified through the scoping review in addressing the digital accessibility needs of older adults. Second, although the data selection and extraction processes were confirmed by co-authors, study screening was performed by the first author. The validation process, while helpful, might not have fully addressed all possible sources of subjective influence. Future research could benefit from involving multiple reviewers in the data extraction and study selection stages, thereby providing a more balanced perspective and reducing the risk of individual biases affecting the outcomes. Third, the current study relied on a single expert for the AHP analysis, which may limit the generalizability and robustness of the prioritization results. To reduce this limitation, frequency data from the scoping review were integrated into the pairwise comparisons to provide a more objective basis alongside expert judgment, thereby helping to reduce potential bias. Nevertheless, involving multiple experts in future AHP analyses could further strengthen the validity of the findings. Incorporating diverse viewpoints and ensuring thorough cross-checking of the data would enhance the robustness and credibility of the review results. Fourth, it is possible that some articles were not retrieved, due to the selected keywords and lack of consideration for gray literature. Although the database searches included a broad range of academic publication types (eg, journal articles, conference papers, and proceedings), many age-friendly usability guidelines and design frameworks are disseminated through gray literature such as institutional reports, professional guidelines, and practitioner-oriented documents, which are not consistently indexed in bibliographic databases. To ensure comprehensive coverage, Google Scholar and Google Search were therefore used to identify gray literature, with the first 50 results from each platform screened to balance systematic coverage with feasibility. This complementary approach strengthened the inclusiveness of the review while acknowledging the inherent limitations of gray literature searching. Finally, while the search strategy may not have captured every possible variation in spelling, pluralization, or word order, the integration of standardized indexing terms (eg, MeSH) and a wide range of keyword formats enhanced the comprehensiveness of the literature search (Multimedia Appendix 2). The review and input from the professional librarian also played a key role in ensuring methodological soundness and adherence to systematic review guidelines.

Conclusions

This study offers valuable insights into the design of transportation planning e-tools tailored to the needs of older adults, providing a robust foundation for the development of accessible and user-friendly digital interfaces. Indeed, the insights generated from this research are highly transferable, offering valuable contributions not only to transportation e-tools but also potentially to other apps and websites that support the autonomy and social participation of older adults. Continued refinement of these guidelines, through iterative updates and the active involvement of older users in the design process, will be crucial to ensure their long-term relevance and impact.

XSL•FO
RenderX

Acknowledgments

The authors gratefully acknowledge librarian Miguel Escobar for his invaluable contribution in reviewing and enhancing our search strategy to ensure its accuracy and comprehensiveness and Hélène Pigot for her insightful guidance and support throughout the study. The authors used the generative artificial intelligence tool ChatGPT (OpenAI) for copyediting and language refinement during manuscript preparation.

Funding

This study is financially supported by the Fonds de recherche du Québec–Santé (2020-VILL279242) and Bourse pour les nouveaux étudiants de 3e cycles du Centre de recherche sur le vieillissement, Sherbrooke.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary files.

Authors' Contributions

SBG and VP were responsible for the conceptualization of this study. VP secured funding and provided overall supervision of the research project. SBG conducted the database and web searches for this scoping review, under the guidance of VP and MG. Screening of publications, data extraction, and data analysis were performed by SBG and subsequently validated by VP. SBG drafted the initial manuscript, which was reviewed and revised by VP and MG. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Analytic heirarchy process decision matrix.
[DOCX File, 35 KB - aging_v9i1e83668_app1.docx]

Multimedia Appendix 2
Full search query example for MEDLINE.
[DOCX File, 18 KB - aging_v9i1e83668_app2.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 143 KB - aging_v9i1e83668_app3.docx]

References

1. Tomida K, Shimoda T, Nakajima C, Kawakami A, Shimada H. Social isolation/loneliness and mobility disability among older adults. Curr Geri Rep 2024;13(2):86-92. [doi: 10.1007/s13670-024-00414-x]
2. Che Had NH, Alavi K, Md Akhir N, Muhammad Nur IR, Shuhaimi MSZ, Foong HF. A scoping review of the factor associated with older adults' mobility barriers. Int J Environ Res Public Health 2023 Feb 27;20(5):4243. [doi: 10.3390/ijerph20054243] [Medline: 36901256]
3. Andersen GJ. Aging and vision: changes in function and performance from optics to perception. Wiley Interdiscip Rev Cogn Sci 2012 May;3(3):403-410. [doi: 10.1002/wcs.1167] [Medline: 22919436]
4. Robnett RH, Chop WC. Gerontology for the Health Care Professional, 3rd edition: Jones & Bartlett Publishers; 2013. URL: https://books.google.co.in/books?id=julfb3qx1h8C [accessed 2026-03-03]
5. Ring L, Barry B, Totzke K, Bickmore T. Addressing loneliness and isolation in older adults: proactive affective agents provide better support. Presented at: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction; Sep 2-5, 2013; Geneva, Switzerland p. 61-66. [doi: 10.1109/ACII.2013.17]
6. Moyen de transport le plus couramment utilisé [Article in French]. Institut de la statistique du Québec. URL: https://statistique.quebec.ca/vitrine/vieillissement/themes/mode-vie/moyen-transport-plus-couramment-utilise [accessed 2026-01-12]
7. Bayne A, Siegfried A, Beck LF, Freund K. Barriers and facilitators of older adults' use of ride share services. J Transp Health 2021 Jun;21:101055. [doi: 10.1016/j.jth.2021.101055] [Medline: 35572055]
8. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]

XSL•FO
RenderX

9. Xu R, Zhang S, Xiong P, Lin AY, Ma J. Towards better driver safety: empowering personal navigation technologies with road safety awareness. Presented at: 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC); Oct 8-12, 2022; Macau, China p. 3004-3009. [doi: 10.1109/ITSC55140.2022.9922545]
10. Whiteside MM, Wallhagen MI, Pettengill E. Sensory impairment in older adults: part 2: vision loss. Am J Nurs 2006 Nov;106(11):52-61. [doi: 10.1097/00000446-200611000-00019] [Medline: 17068436]
11. Dixon J, Crooks H, Henry K. Breaking the ice: supporting collaboration and the development of community online. Can J Learn Technol 2006;32(2). [doi: 10.21432/T25S3M]
12. Murman DL. The impact of age on cognition. Semin Hear 2015 Aug;36(3):111-121. [doi: 10.1055/s-0035-1555115] [Medline: 27516712]
13. Dodd C, Athauda R, Adam M. Designing user interfaces for the elderly: a systematic literature review. Presented at: Australasian Conference on Information Systems (ACIS 2017); Dec 4-6, 2017 URL: https://aisel.aisnet.org/acis2017/61 [accessed 2026-03-13]
14. Bahrampoor Givi S, Gagnon-Roy M, Pigot H, Provencher V. User experience of older adults with an age-friendly transportation planning e-tool: scoping review. JMIR Hum Factors 2025 Jul 28;12(1):e63273. [doi: 10.2196/63273] [Medline: 40720889]
15. Nielsen J. Usability Engineering: Morgan Kaufmann; 1994. URL: https://openlibrary.org/books/OL9283004M/Usability_Engineering_%28Interactive_Technologies%29 [accessed 2026-03-03]
16. Jiang Q, Deng L, Zhang J, Pengbo Y. User-centered design strategies for age-friendly mobile news apps. SAGE Open 2024 Oct;14(4). [doi: 10.1177/21582440241285393]
17. Zaphiris P, Ghiawadwala M, Mughal S. Age-centered research-based web design guidelines. Presented at: Proceeding CHI EA '05 CHI '05 extended abstracts on human factors in computing systems; Apr 2-7, 2005; Portland, Oregon, USA p. 1897-1900. [doi: 10.1145/1056808.1057050]
18. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
19. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
20. Saaty TL. Decision making with the analytic hierarchy process. Int J Serv Sci 2008;1(1):83. [doi: 10.1504/IJSSCI.2008.017590]
21. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. J Clin Epidemiol 2016 Jul;75:40-46. [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
22. Shneiderman B, Plaisant C. Designing the user interface: strategies for effective human-computer interaction, 4th edition: Pearson/Addison Wesley; 2004. URL: https://openlibrary.org/books/OL17216159M/Designing_the_user_interface [accessed 2026-03-03]
23. Gerhardt‐Powals J. Cognitive engineering principles for enhancing human‐computer performance. Int J Hum Comput Interact 1996 Apr;8(2):189-211. [doi: 10.1080/10447319609526147]
24. Weinschenk S, Barker DT. Designing Effective Speech Interfaces: Wiley; 2000. URL: https://www.oreilly.com/library/view/designing-effective-speech/9780471375456/ [accessed 2026-03-03]
25. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
26. Saaty TL. Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World: RWS Publications; 2012. URL: https://books.google.ca/books?id=c8KqSWPFwIUC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false [accessed 2026-03-03]
27. Goepel KD. Implementation of an online software tool for the Analytic Hierarchy Process (AHP-OS). Int J Analytic Hierarchy Process 2018;10(3). [doi: 10.13033/ijahp.v10i3.590]
28. Abdullah N, Abdul Hamid NFB. Interface design features of mobile application for senior citizens. Indones J Elect Eng Comput Sci 2019;14(1):436-442. [doi: 10.11591/ijeecs.v14.i1.pp436-442]
29. Alkhomsan MN, Alturayeif N, Alwadei S, Baslyman M. UsAge guidelines: toward usable Saudi M-Government applications for elderly users. J King Saud Univ Comput Inf Sci 2023 Jan;35(1):202-218. [doi: 10.1016/j.jksuci.2022.11.011]
30. Accessibility. Material Design. 2019. URL: https://m2.material.io/design/usability/accessibility.html [accessed 2025-04-10]
31. Barros AD, Leitão R, Ribeiro J. Design and evaluation of a mobile user interface for older adults: navigation, interaction and visual design recommendations. Presented at: 5th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (DSAI 2013); Oct 9-11, 2013; Vigo, Spain p. 369-378. [doi: 10.1016/j.procs.2014.02.041]
32. Bong WK, Bronshtein I. Designing digital games with & for home-dwelling older adults' social interaction under sheltering measures. Presented at: 7th International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE 2021); Apr 24-26, 2021; Online p. 57-68. [doi: 10.5220/0010467900570068]

XSL•FO
RenderX

33. UX/UI design for elderly users: best practices & principles. LinkedIn. 2024. URL: https://www.linkedin.com/pulse/uxui-design-elderly-users-best-practices-principles-7ow3c/ [accessed 2024-10-04]
34. Dombrovskaia L, Vilches A. Guideline for web pages design inclusive for senior users. Presented at: 2021 40th International Conference of the Chilean Computer Science Society (SCCC); Nov 15-19, 2021; La Serena, Chile p. 1-7. [doi: 10.1109/SCCC54552.2021.9650420]
35. Finn K, Johnson J. Designing for an aging population: toward universal design. Presented at: ACM CHI Conference on Human Factors in Computing Systems (CHI 2016); May 7-12, 2016; San Jose, California, USA p. 1011-1012. [doi: 10.1145/2851581.2856669]
36. Friedman V. A guide to designing for older adults. Smashing Magazine. 2024. URL: https://www.smashingmagazine.com/2024/02/guide-designing-older-adults/ [accessed 2024-10-04]
37. Gao Q, Sun Q. Examining the usability of touch screen gestures for older and younger adults. Hum Factors 2015 Aug;57(5):835-863. [doi: 10.1177/0018720815581293] [Medline: 25957042]
38. Ghorbel F, Métais E, Ellouze N, Hamdi F, Gargouri F. Towards accessibility guidelines of interaction and user interface design for alzheimer's disease patients. Presented at: Proceedings of the Tenth International Conference on Advances in Computer-Human Interactions (ACHI 2017); Feb 26 to Mar 2, 2017; Nice, France p. 143-149 URL: https://personales.upv.es/thinkmind/dl/conferences/achi/achi_2017/achi_2017_6_30_20152.pdf [accessed 2026-03-03]
39. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR mHealth uHealth 2023 Sep 21;11(1):e43186. [doi: 10.2196/43186] [Medline: 37733401]
40. Gruver J. UX/UI design for elderly users: a comprehensive guide. Bootcamp (Medium). 2024. URL: https://medium.com/design-bootcamp/ux-ui-design-for-elderly-users-a-comprehensive-guide-ee49d1870099 [accessed 2024-10-04]
41. Halperin Ben Zvi M, Briscoe G, Sidse C. Age inclusive digital platforms. Presented at: 11th International Conference on Inclusive Design (INCLUDE 2022); Jul 6-8, 2022; London, United Kingdom p. 9-18 URL: https://researchonline.rca.ac.uk/5207/ [accessed 2025-07-10]
42. Lum ASL, Chiew TK, Ng CJ, Lee YK, Lee PY, Teo CH. Development of a web-based insulin decision aid for the elderly: usability barriers and guidelines. Univ Access Inf Soc 2017 Aug;16(3):775-791. [doi: 10.1007/s10209-016-0503-y]
43. Morey SA, Stuck RE, Chong AW, Barg-Walkow LH, Mitzner TL, Rogers WA. Mobile health apps: improving usability for older adult users. Ergonom Des 2019 Oct;27(4):4-13. [doi: 10.1177/1064804619840731]
44. Nasr E, Alsaggaf W, Sinnari D. Developing Usability Guidelines for mHealth Applications (UGmHA). Multimodal Technol Interact 2023;7(3):26. [doi: 10.3390/mti7030026]
45. Making your web site senior friendly: a checklist. : National Institute on Aging and National Library of Medicine; 2001 URL: https://www.huzurevleri.org.tr/docs/MakingYourWebsiteSeniorFriendly.pdf [accessed 2025-04-10]
46. Panagopoulos C, Menychtas A, Tsanakas P, Maglogiannis I. Increasing usability of homecare applications for older adults: a case study. Designs 2019;3(2):23. [doi: 10.3390/designs3020023]
47. Patch K, Spellman J, Wahlbin K. Mobile accessibility: how WCAG 2.0 and other W3C/WAI guidelines apply to mobile. World Wide Web Consortium (W3C). 2015 Feb 12. URL: https://www.w3.org/TR/mobile-accessibility-mapping/ [accessed 2025-04-10]
48. Petrovčič A, Taipale S, Rogelj A, Dolničar V. Design of mobile phones for older adults : an empirical analysis of design guidelines and checklists for feature phones and smartphones. Int J Hum–Comput Interact 2018 Mar 4;34(3):251-264. [doi: 10.1080/10447318.2017.1345142]
49. Rosman S, Siau NZ, Bramantoro A. Evaluating user interface and user experience in mobile applications designed for the elderly. Presented at: 2023 6th International Conference on Applied Computational Intelligence in Information Systems (ACIIS); Oct 23-25, 2023; Bandar Seri Begawan, Brunei Darussalam p. 1-6. [doi: 10.1109/ACIIS59385.2023.10367221]
50. Ruzic L, Lee ST, Liu YE, Sanford JA. Development of universal design mobile interface guidelines (UDMIG) for aging population. In: Antona M, Stephanidis C, editors. Presented at: International Conference on Universal Access in Human-Computer Interaction (UAHCI 2016); Jul 17-22, 2016; Toronto, Ontario, Canada p. 98-108. [doi: 10.1007/978-3-319-40250-5_10]
51. Sakdulyatham R, Preeyanont S, Lipikorn R, Watakakosol R. User interface on smartphone for elderly users. Int J Autom Smart Technol 2017;7(4):147-155. [doi: 10.5875/ausmt.v7i4.1339]
52. Vasudavan H, Jayabalan M, Ramiah S. A preliminary study on designing tour website for older people. 2015 Presented at: 2015 IEEE Student Conference on Research and Development (SCOReD); Dec 13-14, 2015; Kuala Lumpur, Malaysia p. 670-675. [doi: 10.1109/SCORED.2015.7449423]
53. Zahida W, Effendy V, Hadikusuma A. Improving messenger accessibility for elderly users using user centered design (UCD) methods (study case: WhatsApp). Presented at: 2021 International Conference on Software Engineering & Computer Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM); Aug 24-26, 2021; Pekan, Malaysia p. 273-278. [doi: 10.1109/ICSECS52883.2021.00056]
54. A guide to interface design for older adults. Adchitects. URL: https://adchitects.co/blog/guide-to-interface-design-for-older-adults [accessed 2025-04-10]
55. Polyuk S. Age before beauty – a guide to interface design for older adults. Toptal. URL: https://www.toptal.com/designers/ui/ui-design-for-older-adults [accessed 2024-10-04]

XSL•FO
RenderX

56. FACTSHEET on SS 618 – guidelines on user interface design for older adults. : Singapore Standards Council / SPRING Singapore; 2016 URL: https://www.nas.gov.sg/archivesonline/data/pdfdoc/20170323006/Media%20Factsheet_%20Silver%20Industry%20Roadmap%20and%20SS%20618%20Launch_FINAL.pdf [accessed 2025-04-10]
57. Designing for iOS. Apple Developer Documentation. URL: https://developer.apple.com/design/human-interface-guidelines/designing-for-ios [accessed 2025-04-10]
58. UI design for elderly. UserPeek. 2022 Aug 12. URL: https://web.archive.org/web/20240920074455/https://userpeek.com/blog/ui-design-for-elderly/ [accessed 2024-10-04]
59. Ahmad N, Rextin A, Kulsoom UE. Perspectives on usability guidelines for smartphone applications: an empirical investigation and systematic literature review. Inf Softw Technol 2018 Feb;94:130-149. [doi: 10.1016/j.infsof.2017.10.005]
60. Patsoule E, Koutsabasis P. Redesigning websites for older adults: a case study. Behav Inf Technol 2014 Jun 3;33(6):561-573. [doi: 10.1080/0144929X.2013.810777]
61. Zhang X, Dolah J. A study of key factors influencing the usability of smartphone graphical user interfaces for older adults. PaperASIA 2024;40(5b):136-151. [doi: 10.59953/paperasia.v40i5b.232]
62. Zhou C, Yuan F, Huang T, Zhang Y, Kaner J. The impact of interface design element features on task performance in older adults: evidence from eye-tracking and EEG signals. Int J Environ Res Public Health 2022 Jul 28;19(15):9251. [doi: 10.3390/ijerph19159251] [Medline: 35954608]
63. Nahm ES, Preece J, Resnick B, Mills ME. Usability of health web sites for older adults: a preliminary study. Comput Inform Nurs 2004;22(6):326-334. [doi: 10.1097/00024665-200411000-00007] [Medline: 15602301]
64. Nurgalieva L, Jara Laconich JJ, Baez M, Casati F, Marchese M. A systematic literature review of research-derived touchscreen design guidelines for older adults. IEEE Access 2019;7:22035-22058. [doi: 10.1109/ACCESS.2019.2898467]
65. Amouzadeh E, Dianat I, Faradmal J, Babamiri M. Optimizing mobile app design for older adults: systematic review of age-friendly design. Aging Clin Exp Res 2025 Aug 14;37(1):248. [doi: 10.1007/s40520-025-03157-7] [Medline: 40804492]
66. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
67. Provencher V, Baillargeon D, Abdulrazak B, et al. Developing a one-stop platform transportation planning service to help older adults move around in their community where, when, and how they wish: protocol for a living lab study. JMIR Res Protoc 2022 Jun 9;11(6):e33894. [doi: 10.2196/33894] [Medline: 35679116]

Abbreviations

AHP: analytic hierarchy process
CI: consistency index
CR: consistency ratio
MeSH: Medical Subject Headings
PCC: population, concept, context
PRESS: Peer Review of Electronic Search Strategies
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
RI: random index

Edited by A Garcia; submitted 06.Sep.2025; peer-reviewed by A Sardeli, K Best; accepted 08.Feb.2026; published 19.Mar.2026.

Please cite as:
Bahrampoor Givi S, Gagnon-Roy M, Provencher V
Identifying and Prioritizing Age-Friendly Design Principles and Guidelines for Developing Transportation Planning E-Tools: Scoping Review
JMIR Aging 2026;9:e83668
URL: https://aging.jmir.org/2026/1/e83668
doi:10.2196/83668

© Sara Bahrampoor Givi, Mireille Gagnon-Roy, Véronique Provencher. Originally published in JMIR Aging (https://aging.jmir.org), 19.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI-Enabled Smart Glasses for Active Aging: Scoping Review

Claudio Delgado-Morales[1*], PhD; Juan Jesús García-Iglesias[2*], PhD; Ana Duarte-Hueros[1*], PhD

[1]Department of Education, Universidad de Huelva, Avenida de las Fuerzas Armadas, Huelva, Spain

[2]Department of Sociology, Social Work and Public Health, Universidad de Huelva, Huelva, Spain

[*]all authors contributed equally

Corresponding Author:
Claudio Delgado-Morales, PhD
Department of Education, Universidad de Huelva, Avenida de las Fuerzas Armadas, Huelva, Spain

Abstract

Background: The daily use of digital technologies is transforming the day-to-day lives of older adults. Among these technologies, artificial intelligence–enabled smart glasses have recently emerged, which allow constant interaction with the device itself and with the environment. They are designed to be used for multitasking, including options such as being able to take photographs and/or videos; record immersive audio; make calls; and share multimedia content through voice commands, touch, or blink detection.

Objective: The aim of this study was to map the existing evidence and gain insight into the effectiveness and potential benefits of artificial intelligence–enabled smart glasses in promoting active living in old age.

Methods: A scoping review was performed following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) statement by consulting the PubMed, Scopus, and Web of Science databases using a search strategy and syntax (“smart glasses” AND (“older adult” OR elderly OR aging) AND health). The review process was conducted through the Covidence online platform, and the final review protocol was prospectively registered in the Open Science Framework. Both the research question and eligibility criteria were based on the population, concept, and context framework for scoping reviews.

Results: From a total of 58 papers identified, 6 (10.3%) studies were finally included (2 pilot studies, 2 technological development studies with experimental validation of a prototype, 1 mixed methods feasibility trial, and 1 survey) published between 2015 and 2023 after eliminating duplications and screening titles, abstracts, and keywords. The results suggest that the use of artificial intelligence–enabled smart glasses may contribute to improving the quality of life, independence, autonomy, motor functions, and social interactions of older adults.

Conclusions: Because of the novelty of this type of digital technology, there is very little research on this topic at present. Moreover, the adoption and implementation of artificial intelligence–enabled smart glasses are conditioned by hindering factors such as data protection, the high price, and the lack of compatibility with conventional prescription glasses, as well as the lack of evaluation of their effectiveness, usability, and acceptance.

(JMIR Aging 2026;9:e81157) doi:10.2196/81157

KEYWORDS

smart glasses; artificial intelligence; AI; older adults; well-being; active aging

Introduction

The aging of populations worldwide raises important questions for health systems and social policies. In this context, the promotion of active aging, understood as older people’s ability to adapt and function in their daily environment [1-3], has become a priority to improve the quality of life of older adults in terms of health, participation, and safety. Given the decline in physical and cognitive functioning with advancing age, coupled with demographic shifts toward an increasingly aging society, it is essential to contribute to improving the well-being of this population [4]. Emerging digital technologies offer new opportunities to support these objectives [5]. These technologies, which include information and communications technologies, websites, cloud computing, big data, 3D printing, social networks, software, mobile devices, robots, and artificial intelligence (AI), enable immediate access to and transmission of information. Additionally, they enable communication in multiple formats over telecommunication networks anytime and anywhere, interconnection between devices, and automation [6-9].

AI-enabled smart glasses have recently stood out among these technologies [10]. These are wearable devices designed for multitasking use, integrating advanced technologies such as displays, cameras, internet connectivity, and AI to offer additional functions beyond those of traditional glasses (eg,

XSL•FO
RenderX

take photos, record audio and immersive videos, listen to music, make and receive calls, and share multimedia content) [11-13]. These devices can be controlled by users through voice commands, touch, or physical inputs such as blink detection.

The most advanced AI-enabled smart glasses use this technology to process and analyze data in real time, enhancing the user experience by interacting with the environment in the most efficient and enriching way [13]. Due to their technological features, AI-enabled smart glasses promise to help older adults gain higher daily functional autonomy for active aging by providing feedback, personalized assistance, and training, such as in taking medication, getting dressed, learning to eat more healthful food, or exercising [14-16]. Moreover, this technology takes advantage of the fact that older adults often wear traditional glasses [17].

While the potential of AI-enabled smart glasses has been studied in the area of health training [18,19], at the professional level [12,20-22], and as an assistive technology to support functional diversity [13,23,24], as well as in younger populations [4,25], research on their use in active living in old age, as a training resource, and/or in the promotion of the well-being of older adults is scarce. For this reason, a scoping review was carried out as it is appropriate for identifying key concepts that underpin a given field of research, as well as clarifying the definitions and/or conceptual boundaries of a topic [26-29].

The main objectives of this study were to map the available evidence and identify gaps in the existing research, as well as to examine the effectiveness and possible benefits of AI-enabled smart glasses in promoting active aging in older adults. We explore their potential impact and provide an overview of them, as well as clarify key concepts related to their use in fundamental areas such as autonomy, well-being, and social participation, while also identifying barriers that affect both the implementation and adoption of these devices.

Methods

Study Design

A scoping review was performed in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [30], consistent with the standards and the updated and extended checklist for scoping reviews, PRISMA-ScR (PRISMA extension for Scoping Reviews) [31]. Similarly, relevant recommendations to ensure the thoroughness and quality of the process were taken into account in the different phases and drafting of this scoping review [29].

The final review protocol was prospectively filed with the Open Science Framework on April 16, 2025 [32], and is publicly accessible.

Databases and Search Strategies

In March 2025, a total of 3 databases of high scientific impact, relevance, completeness, and functionality for research were consulted, with one of them specific to the health field given the nature of the study: PubMed, Scopus, and Web of Science. Two phases were determined to establish the search strategy. First, as phase 1, an iterative premapping was carried out in the 3 databases, combining terms such as “artificial intelligence glasses,” “well-being,” and “benefits” with those referring to older adults (“older adult,” “elderly,” and “aging”) to unify the keywords used in the first studies found and, in this way, be able to establish a subsequent search strategy that was as operational as possible. Thus, a network of co-occurring keywords was generated using the data visualization and analysis software VOSviewer (version 1.6.20; Centre for Science and Technology Studies, Leiden University) from a total of 10 preliminary documents: 2 (20%) from PubMed, 6 (60%) from Scopus, and 2 (20%) from Web of Science.

Then, as phase 2, based on the keywords co-occurring in the preliminary documents, it was decided to use and combine the following terms with the Boolean operators “AND” and “OR” using the following search strategy: (“smart glasses” AND (“older adult” OR elderly OR aging) AND health). In this way, where appropriate, words were combined to reflect the syntax and search parameters common to all 3 databases. In addition, it should be noted that the number of records generated did not vary when truncation was used in the term “health*.”

Selection Criteria and Participant Demographics

The research questions were formulated based on the stated objectives and according to the eligibility criteria, as follows: (1) what scientific evidence exists and what knowledge gaps can be identified on the use of AI-enabled smart glasses in promoting active aging in older adults? (2) What is the effectiveness of AI-enabled smart glasses and what potential benefits do they offer in the promotion of active aging in older adults? (3) What barriers or factors reported in the literature may hinder both the implementation and adoption of AI-enabled smart glasses in older adults for the promotion of active aging?

As shown in Table 1, the population, concept, and context framework was used in this scoping review. It is a structure that streamlines the approach and definition of inclusion criteria in scoping reviews [29].

XSL•FO
RenderX

Table . Details of the population, concept, and context (PCC) framework and eligibility criteria.

PCC domain	Description
Population	Older people, ie, those aged >65 years, following the definition of the World Health Organization [33], and older adults who live in nursing homes or residences
Concept	Use of AI[a]-enabled smart glasses by older adults and their effectiveness and possible benefits in the promotion of active aging; AI-enabled smart glasses similar to Ray-Ban Meta, Solos AirGo Vision, Halliday glasses, Microsoft HoloLens 2, Google Glass, or Epson Moverio BT-300, among others; and wearable assistive technology devices with built-in AI and/or AI-automated mechanisms for older adults
Context	Use of AI-enabled smart glasses in daily life; for self-management of health; for educational purposes; in improving cognitive aspects, autonomy, and quality of life; as a means of leisure for interactive entertainment and/or social connections in one's own home, in residences, or in health care settings; and consideration of barriers to implementation and/or adoption, such as technological availability, levels of digital literacy, technological accessibility, and factors inherent to socioeconomic and cultural conditions

[a]AI: artificial intelligence.

As this is a particularly recent topic and the use of this type of technology is on the rise, it was decided not to consider any exclusion criteria related to the type of document, year, area, and language, among others, so as not to discard related studies. The only exclusion criterion taken into account was that the studies should not be focused on conventional augmented or virtual reality glasses without built-in AI and/or capabilities that use AI through their software.

Data Collection Procedures

To approach the review process, the online platform Covidence (Veritas Health Innovation) was used, which facilitates and optimizes the conduct of this type of study (eg, import of references, screening of records, blind review, data extraction, and risk of bias assessment). Two reviewers independently and blindly carried out the entire review process on this platform based on the proposed eligibility criteria, justifying and recording the reasons for exclusion. Any discrepancies during the selection of studies were discussed and resolved through consensus.

Data extraction, focused on identifying, organizing, and compiling the most relevant information related to the main objective of this review, was carried out in an exhaustive manner, including all the findings reported in the selected studies, to guarantee a rigorous and coherent analysis in accordance with the available evidence. The extracted data included the following details: (1) author and year of publication, (2) title, (3) type of document, (4) country of origin, (5) objective, (6) methodology, (7) sample, (8) type of AI-enabled smart glasses used and/or analyzed, and (9) main results.

Data Analysis

After finalizing the data extraction process, the key findings obtained from the selected studies were summarized and presented in an evidence synthesis table (Multimedia Appendix 1 [14-16,23,34,35]).

In addition to examining the effectiveness and potential benefits of this type of technology in active living in old age, some barriers were detected that affect both the implementation and the adoption of AI-enabled smart glasses in this sector of the population. Furthermore, the gaps found in scientific literature regarding interventions based on these digital technologies were also synthesized, leading to the elaboration of a detailed summary of the objectives and findings of this study.

Results

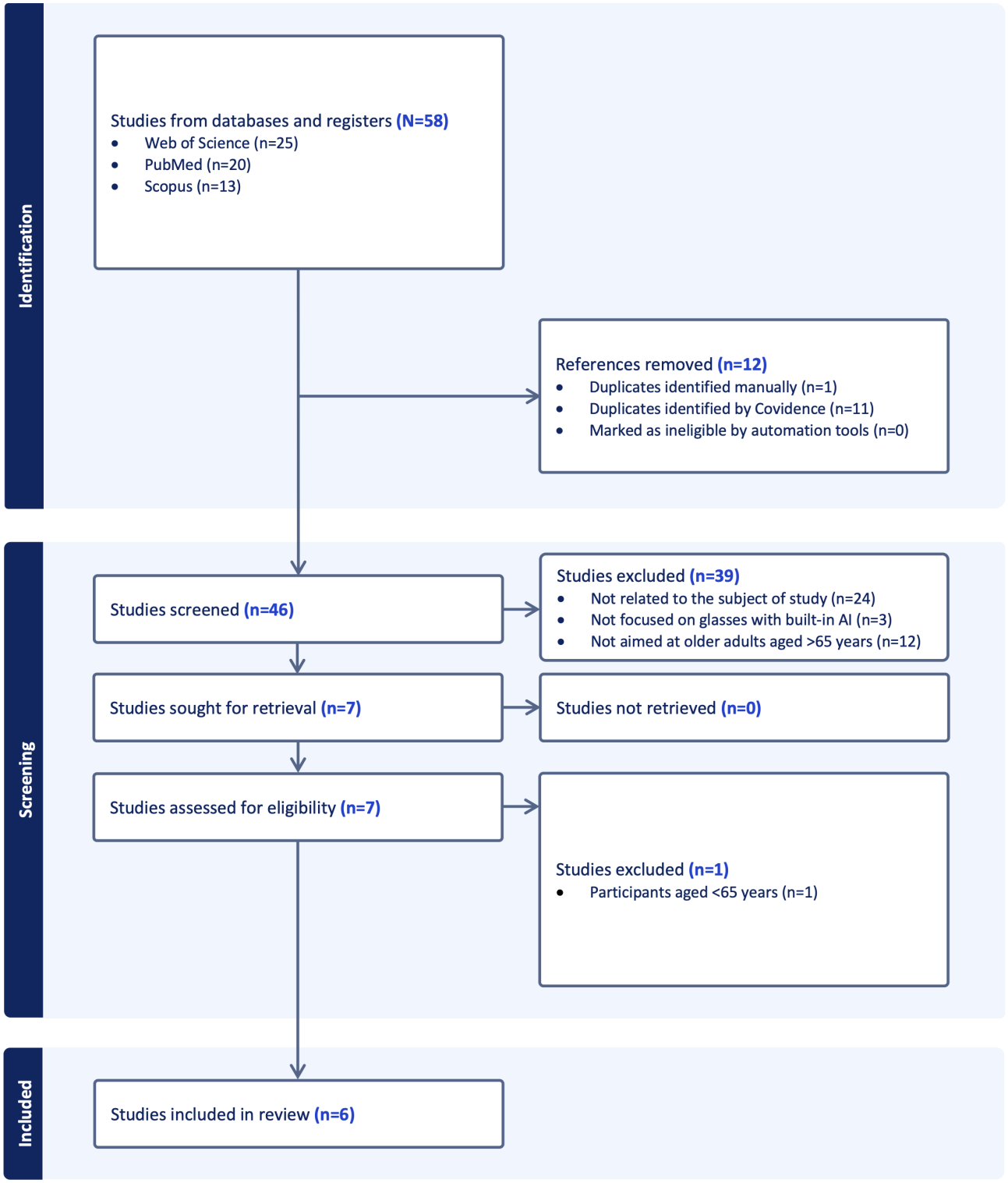
Inclusion of Studies

Of a total of 58 records, 12 (20.7%) duplicate documents were eliminated. Subsequently, the remaining 79.3% (46/58) of the results were screened by analyzing the titles, abstracts, and keywords. Of these 46 papers, 39 (84.8%) were excluded because they were not related to the topic under study, not focused on AI-enabled smart glasses, and/or not targeted at adults older than 65 years. Therefore, 7 studies were reviewed in full. Of the 7 studies analyzed in full, 1 paper was excluded because it was not aimed at older adults and did not consider age (>65 years) among the participants.

As shown in Figure 1 (flowchart adapted from the PRISMA-ScR guidelines) and Table 2, 6 studies were included in this scoping review.

XSL•FO
RenderX

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flowchart (extracted from Covidence). AI: artificial intelligence.



XSL•FO
RenderX

Table . Summary of the included studies on artificial intelligence (AI)–enabled smart glasses for older adults.

Study	Sample	Glasses or device	Study design	Outcomes
Muema et al [14]	Older adults with cognitive impairment (not specified)	Microsoft HoloLens 2	Experimental integration study	Improved independence and autonomy via companion robot–AI-enabled smart glasses integration and enhanced emotion recognition
Tunur et al [15]	7 older adults with Parkinson disease	Google Glass	Mixed methods feasibility study	Safe and acceptable, improvements in motor function, and positive participant satisfaction
Zhao et al [16]	62 older adults with Parkinson disease	AI-enabled smart glasses (general)	Online survey	Positive attitudes toward adoption and perceived usefulness for daily activity support and self-confidence
Moshtael et al [23]	23 older adults with macular disease	Epson Moverio BT-200	Pilot quantitative study	Increased reading speed and satisfaction with personalized text display
Wang et al [34]	50 older adults	Epson Moverio BT-300	Experimental prototype study	Accurate facial recognition, improved text readability, and voice command control validated
Chang et al [35]	5 older adults	Jorjin J7EF AR smart glasses	Pilot quantitative study	Improved exercise adherence and recovery rates and AI-enabled smart glasses feasible for home-based fitness promotion

Demographics of the Included Studies

Half of the papers were published in 2020 [15,23,34], with the publication period starting in 2015, when the first of the eligible studies was published [16], and including 2022 [14] and 2023 [35]. The studies were conducted in 5 different locations: Canada [14], the United States [15], the Netherlands [16], the United Kingdom [23], and Taiwan [34,35].

Designs of Included Studies

Five of the 6 studies used a quantitative methodology [14,16,23,34,35], and 1 study adopted a mixed methods approach [15]. This was a single-group feasibility pilot test where previously validated questionnaires and standardized scales were administered, as well as semistructured interviews. Three studies used an experimental-technological perspective [14,34,35]. The use of surveys [16] or the application of standardized tests [23] as instruments and techniques for data collection was also highlighted.

Selection Criteria and Participant Demographics

Sample size varied significantly in all the studies, ranging from 5 [35] to 62 [16]. Only 1 study did not specify the sample size, although it referred to older adults with reported cognitive impairment [14]. In 3 of the 6 studies, gender was specified, with predominantly female participants in 2 studies [15,23] and male participants in 1 study [16]. These 3 studies also reported the age of the participants, with an average of 69 years [15], 66 years [16], and 81 years [23]. Regarding place of residence, the home or residential centers were the norm in 3 studies [14-16], with residence in rural areas being relevant in 1 study [35]. Two studies omitted data on place of residence [23,34]. The diagnosis of an illness was an eligibility criterion in 3 studies, specifically Parkinson disease [15,16] and macular degeneration with oculomotor deficits [23]. No study detailed data on the occupation, educational level, or degree of digital literacy and technological training of the participants.

A wide variety of AI-enabled smart glasses were considered in the different studies, highlighting Microsoft HoloLens 2 [14], Google Glass [15], Epson Moverio BT-200 [23], Epson Moverio BT-300 [34], and Jorjin J7EF AR smart glasses [35]. Only 1 study addressed this type of technology in a generalized manner without specifying the model of glasses [16]. In addition to AI-enabled smart glasses, in the included studies, other technologies and devices were used, such as intelligent robots and virtual assistants [14], tablets and smartphones [16], cloud computing [34], and a portable Android device that performs real-time data processing (eg, image analysis or posture calculation) to send information to the glasses while the user moves or exercises [35].

The promotion of autonomy and the improvement of motor functions were the most noteworthy benefits as AI-enabled smart glasses can provide visual support or auditory cues by overlaying information in real time [14-16]. Similarly, benefits for physical health were evidenced through a more active lifestyle and increased mobility of the participants [15,35], as well as positive results in social interactions [14]. Another benefit identified was the improvement in reading performance and text readability due to the dynamization and personalization via AI-enabled smart glasses [23,34].

The main barriers or hindering factors that were identified in the results of the studies included sociocontextual, distinguishing

XSL•FO
RenderX

privacy issues and network vulnerabilities [14] or the high cost of the devices [15,16]. In terms of the inherent factors of the technology itself, such as its design and technical characteristics, some glasses may be inconvenient to use for this sector of the population due to the reduced size of the displays or a limited audio volume [15], as well as the possible absence of blue light filters in some displays [16]. An internal barrier was the lack of evaluations of the usability and acceptance of these technologies among older adults [14].

Discussion

Principal Findings

Preliminary findings suggest improvements in quality of life, independence, motor function, and social relationships among older adults resulting from the use of AI-enabled smart glasses. In terms of quality of life and overall well-being, aspects such as lack of social connections, independent living, or decreased mobility of older adults are some of the current concerns in this population. AI-enabled smart glasses have emerged as digital enabling technologies to alleviate these issues.

This scoping review contributes to scientific literature by mapping the available evidence on this type of technology and identifying indicators of effectiveness and potential benefits in promoting active aging. As this is an emerging technology with a marked innovative character, there are currently very few studies on it. Therefore, the contribution of this scoping review, from a holistic point of view, is even more significant. AI-enabled smart glasses facilitated independence, functional autonomy, and improvement of motor skills in the context of active and healthy aging [14-16]. These benefits were produced because these portable devices contribute to more conscious decision-making and greater confidence in the actions performed, as well as improvements in motor performance skills such as coordination, posture, and balance. For these reasons, AI-enabled smart glasses have the potential to empower older adults in the performance of their daily tasks and routines [16,34].

In addition, the implementation of AI-enabled smart glasses in this sector of the population produced an increase in daily physical activity levels due to the ubiquitous nature of this type of technology [15,35]. In other studies, links were found between the adoption of AI-enabled smart glasses and improvements in social connections through home automation [14]. It has also been demonstrated that AI-enabled smart glasses facilitate reading in real time [34]. In this sense, these technologies could offer alternatives for entertainment and personal enjoyment [15].

It was also shown that the use of AI-enabled smart glasses minimized symptoms of diseases such as Parkinson disease or macular degeneration [15,16,23]. All this translates into a significant increase in well-being and quality of life through daily assistive devices [14,16,35].

However, some hindering factors and/or barriers were identified regarding both the adoption and implementation of these wearable devices among older adults. First, more research is needed to guarantee aspects related to the safety, effectiveness, usability, and social acceptance of AI-enabled smart glasses regarding active living in old age [16].

Second, it is important that AI-enabled smart glasses are compatible with the prescription glasses worn by older adults, are easy to use, include a simple and natural interface, and recognize broad gestures rather than overly precise keystrokes. All these features would facilitate navigation [16].

The use of displays without blue light filters in this type of technology could alter quality of sleep and circadian rhythm. In fact, prolonged use of AI-enabled smart glasses may cause a feeling of visual fatigue or mild dizziness [15].

Furthermore, the efficient use of these digital technologies still needs to address the limited digital literacy of many older people [36,37]. In other words, the lack of technological skills and basic competencies in the use of electronic devices and digital tools is a limiting factor in the adoption of these devices and autonomous and independent management [38-40], showing the need to promote training strategies for digital learning and empowerment in older adults.

Limitations and Future Research

Although only 6 studies fulfilled the inclusion criteria following an exhaustive search that identified 58 documents, we consider that this scoping review is both relevant and valid, given that the subject of study (AI-enabled smart glasses) represents an emerging technology. Consequently, it is to be expected that scientific literature in this area remains limited or is still in the early stages of development. This review mapped the current state of knowledge and identified gaps in the literature, offering a meaningful and relevant preliminary overview that is consistent with the objectives of a scoping review. However, publication bias, methodological heterogeneity, sample size constraints, the instruments used, and the limited availability of rigorous conceptual frameworks, together with inherent language restrictions, made it difficult to establish relationships among the included studies. Therefore, further empirical research with more rigorous methods and periodic reviews is suggested as the scientific output grows and new AI-enabled smart glasses become commercially available.

Conclusions

The available studies included in this scoping review suggest that AI-enabled smart glasses enable improvements in the quality of life of older adults, especially in capabilities and skills such as independence, autonomy, and motor function. Preliminary evidence also indicates higher levels of physical activity and increased social interactions through the use of AI-enabled smart glasses. Although these portable devices could offer a more inclusive and accessible digital future by potentially delaying the effects of aging and contributing to active living in old age, these results should be interpreted cautiously given the limited number of studies and methodological heterogeneity.

This scoping review used a synthesis of evidence to clarify the state of the art and expand existing knowledge on the role of these digital technologies in promoting active aging, in addition to identifying gaps in research that serve as a basis for future studies. It is necessary to explore more deeply the potential

XSL•FO
RenderX

mechanisms underlying aging delay through this type of digital technology, which requires research with larger samples and longitudinal designs. Moreover, successful implementation and adoption require addressing digital literacy and ongoing training in technological skills among this population segment, as well as addressing challenges related to data protection, usability, acceptance, efficiency, and technical complexity.

Acknowledgments

The authors are grateful for the support received from the Agencia Estatal de Investigación; the Ministerio de Ciencia, Innovación y Universidades; the European Union NextGenerationEU and Recovery, Transformation and Resilience Plan and the European Regional Development Fund, European Union (projects TED2021-129253B-100 and PID2021-123552OB-I00); and the Consejería de Universidad, Investigación e Innovación of the Junta de Andalucía (Spain; project PROYEXCEL_00320).

Funding

This publication is part of the research, development, and innovation project TED2021-129253B-I00 (funded by grant MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU and Recovery, Transformation and Resilience Plan); project PROYEXCEL_00320 (funded by the Consejería de Universidad, Investigación e Innovación of the Junta de Andalucía, Spain); and project PID2021-123552OB-I00 (funded by grant MICIU/AEI/10.13039/501100011033 and by the European Regional Development Fund, European Union).

Data Availability

The datasets generated or analyzed during this study are available in the Open Science Framework (OSF) repository [32].

Authors' Contributions

CD-M contributed to conceptualization, database search, study selection, data extraction, data synthesis, and manuscript drafting. JJG-I contributed to conceptualization, database search, data synthesis, and manuscript drafting. AD-H contributed to conceptualization, study selection, data extraction, data synthesis, and manuscript drafting. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Organization of data as a synthesis of the final included studies and their key findings.
[DOCX File, 53 KB - aging_v9i1e81157_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 230 KB - aging_v9i1e81157_app2.pdf]

References

1. Keating N. A research framework for the United Nations Decade of Healthy Ageing (2021-2030). Eur J Ageing 2022 Sep;19(3):775-787. [doi: 10.1007/s10433-021-00679-7] [Medline: 35035341]
2. Xiao N, Xia B, Buys L, Susilawati C, Larbi M, Chen Q. Comparative analysis of 16 aging concepts and their influence on aging narratives: bibliometric and content analysis. JMIR Aging 2025 Sep 23;8:e72011. [doi: 10.2196/72011] [Medline: 40986788]
3. Zingmark M, Björk J, Granbom M, et al. Exploring associations of housing, relocation, and active and healthy aging in Sweden: protocol for a prospective longitudinal mixed methods study. JMIR Res Protoc 2021 Sep 21;10(9):e31137. [doi: 10.2196/31137] [Medline: 34546172]
4. Pan H, He X, Zeng H, Zhou J, Tang S. Pilot study of piano learning with AR smart glasses considering both single and paired play. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population Applications in Health, Assistance, and Entertainment: Springer; 2018:561-570. [doi: 10.1007/978-3-319-92037-5_39]
5. Sobrinho ADS, Gomes GDO, Bueno Júnior CR. Developing a multiprofessional mobile app to enhance health habits in older adults: user-centered approach. JMIR Form Res 2024 Apr 15;8:e54214. [doi: 10.2196/54214] [Medline: 38619865]
6. Fuentes S, Tongson E, Gonzalez Viejo C. Novel digital technologies implemented in sensory science and consumer perception. Curr Opin Food Sci 2021 Oct;41:99-106. [doi: 10.1016/j.cofs.2021.03.014]
7. Jayaraman PP, Forkan ARM, Morshed A, Haghighi PD, Kang Y. Healthcare 4.0: a review of frontiers in digital health. WIREs Data Min Knowl 2020 Mar;10(2) [FREE Full text] [doi: 10.1002/widm.1350]

XSL•FO
RenderX

8. Mihai F, Aleca OE, Gheorghe M. Digital transformation based on AI technologies in European Union organizations. Electronics (Basel) 2023 Apr;12(11):2386. [doi: 10.3390/electronics12112386]
9. Yeung AWK, Kulnik ST, Parvanov ED, et al. Research on digital technology use in cardiology: bibliometric analysis. J Med Internet Res 2022 May 11;24(5):e36086. [doi: 10.2196/36086] [Medline: 35544307]
10. Zhang Z, Bai E, Joy K, Ghelaa PN, Adelgais K, Ozkaynak M. Smart glasses for supporting distributed care work: systematic review. JMIR Med Inform 2023 Feb 28;11:e44161. [doi: 10.2196/44161] [Medline: 36853760]
11. Kaviani F, Lyall B, Koppel S. Exploring social perceptions of everyday smartglass use in Australia. PLoS One 2024;19(11):e0313001. [doi: 10.1371/journal.pone.0313001] [Medline: 39485765]
12. Romare C, Enlöf P, Anderberg P, Jildenstål P, Sanmartin Berglund J, Skär L. Nurse anesthetists' experiences using smart glasses to monitor patients' vital signs during anesthesia care: a qualitative study. PLoS ONE 2021;16(4):e0250122. [doi: 10.1371/journal.pone.0250122] [Medline: 33882100]
13. Waisberg E, Ong J, Masalkhi M, et al. Meta smart glasses—large language models and the future for assistive glasses for individuals with vision impairments. EYE (Lond) 2024 Apr;38(6):1036-1038. [doi: 10.1038/s41433-023-02842-z]
14. Muema C, Lawrence S, Anjum T. Social embodiment of companion robots in smart spaces: IoRT for independent living. In: Galambos P, Kayacan E, Madani K, editors. Robotics, Computer Vision and Intelligent Systems: Springer; 2022:147-171. [doi: 10.1007/978-3-031-19650-8_8]
15. Tunur T, DeBlois A, Yates-Horton E, Rickford K, Columna LA. Augmented reality-based dance intervention for individuals with Parkinson's disease: a pilot study. Disabil Health J 2020 Apr;13(2):100848. [doi: 10.1016/j.dhjo.2019.100848] [Medline: 31679951]
16. Zhao Y, Heida T, van Wegen EEH, Bloem BR, van Wezel RJA. E-health support in people with Parkinson's disease with smart glasses: a survey of user requirements and expectations in the Netherlands. J Parkinsons Dis 2015;5(2):369-378. [doi: 10.3233/JPD-150568] [Medline: 25855044]
17. Novac PE, Pegatoquet A, Miramond B, Caquineau C. UCA-EHAR: a dataset for human activity recognition with embedded AI on smart glasses. Appl Sci (Basel) 2022 Apr;12(8):3849. [doi: 10.3390/app12083849]
18. Kopetz JP, Wessel D, Jochems N. User-centered development of smart glasses support for skills training in nursing education. I-com 2019 Nov 18;18(3):287-299. [doi: 10.1515/icom-2018-0043]
19. Yoon H, Kim SK, Lee Y, Choi J. Google Glass-supported cooperative training for health professionals: a case study based on using remote desktop virtual support. J Multidiscip Healthc 2021;14:1451-1462. [doi: 10.2147/JMDH.S311766] [Medline: 34168458]
20. Sumner J, Lim HW, Bundele A, et al. Through the lens: a qualitative exploration of nurses' experiences of smart glasses in urgent care. J Clin Nurs 2025 Mar;34(3):948-958. [doi: 10.1111/jocn.17313] [Medline: 38837508]
21. Zuidhof N, Peters O, Verbeek PP, Ben Allouch S. Social acceptance of smart glasses in health care: model evaluation study of anticipated adoption and social interaction. JMIR Form Res 2025 Feb 11;9:e49610. [doi: 10.2196/49610] [Medline: 39933169]
22. Harari RE, Schulwolf SL, Borges P, et al. Applications of augmented reality for prehospital emergency care: systematic review of randomized controlled trials. JMIR XR Spatial Comput 2025 Feb;2:e66222-e66222. [doi: 10.2196/66222]
23. Moshtael H, Tooth C, Nuthmann A, Underwood I, Dhillon B. Dynamic text presentation on smart glasses: a pilot evaluation in age-related macular degeneration. Br J Vis Impair 2020 Jan;38(1):24-37. [doi: 10.1177/0264619619889998]
24. Seiple W, van der Aa HPA, Garcia-Piña F, Greco I, Roberts C, van Nispen R. Performance on activities of daily living and user experience when using artificial intelligence by individuals with vision impairment. Transl Vis Sci Technol 2025 Jan 2;14(1):3. [doi: 10.1167/tvst.14.1.3] [Medline: 39775799]
25. Nag A, Haber N, Voss C, et al. Toward continuous social phenotyping: analyzing gaze patterns in an emotion recognition task for children with autism through wearable smart glasses. J Med Internet Res 2020 Apr 22;22(4):e13810. [doi: 10.2196/13810] [Medline: 32319961]
26. Anderson S, Allen P, Peckham S, Goodwin N. Asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. Health Res Policy Syst 2008 Jul 9;6(1):7. [doi: 10.1186/1478-4505-6-7] [Medline: 18613961]
27. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
28. Ehrich K, Freeman GK, Richards SC, et al. How to do a scoping exercise: continuity of care. Res Policy Plan 2002 Jan;20(1):25-29 [FREE Full text]
29. Peters M, Godfrey C, Mcinerney P, et al. Scoping reviews. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: Joanna Briggs Institute; 2024. [doi: 10.46658/JBIMES-24-09]
30. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021 Mar 29;10(1):89. [doi: 10.1186/s13643-021-01626-4] [Medline: 33781348]
31. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
32. Delgado-Morales C, Duarte-Hueros A. AI-enabled smart glasses for active aging: scoping review protocol summary. OSF Registries. 2025 Apr 16. URL: https://osf.io/58ypw/overview [accessed 2026-02-18]

XSL•FO
RenderX

33. World report on ageing and health. World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789241565042 [accessed 2025-04-01]
34. Wang CS, Huang W, Chang YF, Yeh CM, Xu ZY. Development of an assistive device via smart glasses. Presented at: IEEE 2nd Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS); May 29-31, 2020; Tainan, Taiwan p. 39-42 URL: https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=9195411 [accessed 2025-05-25]
35. Chang WJ, Chen HW, Su JP, Liu SC, Ou YK, Chen MC. A 2 fitness: an artificial intelligent fitness assistance system using a augmented reality smart glasses for elderly health-promotion. Presented at: 2023 International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan); Jul 17-19, 2023. [doi: 10.1109/ICCE-Taiwan58799.2023.10227060]
36. Ngiam NHW, Yee WQ, Teo N, et al. Building digital literacy in older adults of low socioeconomic status in Singapore (project wire up): nonrandomized controlled trial. J Med Internet Res 2022 Dec 2;24(12):e40341. [doi: 10.2196/40341] [Medline: 36459398]
37. Rodríguez-Miranda FDP, Illanes-Segura R, Ceada-Garrido Y, Infante-Moro JC. Validation of a scale based on the DigComp framework on internet navigation and cybersecurity in older adults. Front Educ 2025 Jan;10:1520929. [doi: 10.3389/feduc.2025.1520929]
38. Cipolletta S, Gris F. Older people's lived perspectives of social isolation during the first wave of the COVID-19 pandemic in Italy. Int J Environ Res Public Health 2021 Nov 11;18(22):11832. [doi: 10.3390/ijerph182211832] [Medline: 34831586]
39. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018 May 8;58(3):467-477. [doi: 10.1093/geront/gnw249] [Medline: 28201730]
40. Mannheim I, Schwartz E, Xi W, et al. Inclusion of older adults in the research and design of digital technology. Int J Environ Res Public Health 2019 Oct 2;16(19):31581632. [doi: 10.3390/ijerph16193718] [Medline: 31581632]

Abbreviations

AI: artificial intelligence
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by C Lee; submitted 23.Jul.2025; peer-reviewed by A Gupta, E Oluokun, N Bievre, R Suri, S Mittal; accepted 16.Dec.2025; published 25.Feb.2026.

Please cite as:
Delgado-Morales C, García-Iglesias JJ, Duarte-Hueros A
AI-Enabled Smart Glasses for Active Aging: Scoping Review
JMIR Aging 2026;9:e81157
URL: https://aging.jmir.org/2026/1/e81157
doi:10.2196/81157

© Claudio Delgado-Morales, Juan Jesús García-Iglesias, Ana Duarte-Hueros. Originally published in JMIR Aging (https://aging.jmir.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of 4 Weeks or More Immersive Virtual Reality on Quality of Life and Physical Activity in Older Adults: Systematic Review and Meta-Analysis

Iria Trillo-Charlín[1,2], MSc; Javier Bravo-Aparicio[2,3], MSc; Juan Avendaño-Coy[1,2], PhD; Héctor Beltrán-Alacreu[2,3], PhD

[1]Toledo Physiotherapy Research Group (GIFTO), Faculty of Physical Therapy and Nursing, University of Castilla-La Mancha, Toledo, Spain

[2]Toledo Physiotherapy Research Group (GIFTO), Castilla-La Mancha Health Research Institute (IDISCAM), Toledo, Spain

[3]Pain, Mental Health, Exercise and Technology Research Group (PAIN+MET), Faculty of Physical Therapy and Nursing, University of Castilla-La Mancha, Avenue de Carlos III, 45004 Toledo, Toledo, Spain

Corresponding Author:
Javier Bravo-Aparicio, MSc
Toledo Physiotherapy Research Group (GIFTO), Castilla-La Mancha Health Research Institute (IDISCAM), Toledo, Spain

Abstract

Background: Population aging poses significant public health challenges. Older adults often face multimorbidity, functional decline, and diminished quality of life. While physical activity can mitigate these effects, adherence remains low. Immersive virtual reality (IVR) has emerged as a promising, engaging tool to promote physical and cognitive health in this population.

Objective: The review aims to evaluate the effectiveness of IVR interventions lasting 4 weeks or more on quality of life, physical activity, pain, perceived effort, and adverse events in older adults.

Methods: A systematic review and meta-analysis were conducted following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and Cochrane guidelines. Literature was searched across PubMed, Web of Science, PEDro, and Scopus, as well as sources of gray literature. Eligible studies included randomized controlled trials involving participants aged >60 years, using IVR via head-mounted display. Outcomes assessed included quality of life, physical activity, pain, perceived effort, and adverse events. Risk of bias and evidence certainty were assessed using Risk of Bias 2.0 and GRADE (Grades of Recommendation Assessment Development and Evaluation), respectively.

Results: A total of 14 studies with 839 participants were included in the qualitative synthesis, of which 8 were eligible for quantitative meta-analysis. The pooled analysis showed a statistically significant moderate effect of IVR on quality of life (standardized mean difference [SMD]=0.48, 95% CI 0.1-0.8; *P*=.007), particularly in interventions lasting 10 to 12 weeks or involving more than 600 minutes of exposure. For physical activity, no significant differences were found between IVR and control groups (SMD=–0.2, 95% CI –0.7 to 0.4; *P*=.50). Evidence for secondary outcomes (pain, perceived exertion, and adverse events) was limited and largely qualitative, with inconsistent findings. Pain outcomes, assessed in 2 studies, indicated reductions in the IVR group, especially when multimodal approaches were used. Perceived effort was not systematically measured. Adverse events were generally mild, with cybersickness being the most reported issue.

Conclusions: IVR interventions of 4 weeks or more appear to moderately improve quality of life in older adults, especially those with clinical vulnerabilities or in institutional settings. Although effects on physical activity were not significant, trends suggest potential with appropriate program design. Preliminary findings support IVR's use in pain reduction, particularly when incorporating emotional and multisensory elements. The low incidence of adverse events suggests good tolerability. Overall, IVR is a promising and safe tool to support healthy aging, though further high-quality studies are needed to confirm these findings and assess long-term outcomes.

(JMIR Aging 2026;9:e80820) doi:10.2196/80820

KEYWORDS

adverse events; immersive virtual reality; older adults; pain; perceived effort; physical activity; PRISMA; Preferred Reporting Items for Systematic Reviews and Meta-Analyses; quality of life

XSL•FO
RenderX

Introduction

By 2030, the population of adults aged ≥65 years is expected to reach 994 million (12% worldwide), intensifying pressures on public-health systems [1]. From a biological perspective, aging is defined as a complex and progressive process that affects multiple systems, leading to a decline in functional capacity and increased vulnerability to various pathologies [2]. Aging is often accompanied by impairments such as loss of muscle strength, cognitive decline, and a heightened risk of falls and disability, all of which negatively impact the autonomy and quality of life of older adults [2-4].

In this context, it is crucial to identify effective strategies that support health and well-being in later life, with physical activity being one of the fundamental pillars for promoting healthy aging [5]. In older adults, physical activity is associated with better physical function, preservation of cognition, and enhanced quality of life. It is also linked to lower disability in activities of daily living, fewer falls, and a reduced risk of neurodegenerative diseases such as dementia [6-9]. Moreover, it has been linked to decreased mortality rates, and it is a leading modifiable determinant of healthy aging with a dose–response association to all-cause and cardiovascular mortality in older adults [10,11]. However, despite the available evidence, the proportion of older adults who meet physical activity recommendations remains suboptimal [11]. Several factors contribute to this, including lack of awareness of the benefits, fear of pain or falling, low motivation, and environmental barriers [12].

In this regard, immersive virtual reality (IVR) emerges as a potential nonpharmacological therapeutic alternative. Through 3D-simulated environments experienced via devices known as head-mounted displays, IVR offers an immersive experience in which users can interact with the virtual environment [13]. IVR's sensorimotor immersion and playful elements may support engagement with physical activity, but current evidence is exploratory and based on a small feasibility trial [14]. Because IVR may enhance motivation and engagement, it is clinically and pragmatically important to test whether multiweek IVR programs can change physical activity behavior. Accordingly, our protocol designated physical activity as a coprimary outcome [15]. Additionally, its design allows for the adjustment of difficulty levels, making it a customizable alternative [16,17]—a feature considered essential in exercise prescription [18].

On the other hand, IVR has also proven useful in interventions aimed at cognitive stimulation and reminiscence, thereby expanding its range of applications in the context of active aging [19,20].

Currently, evidence suggests that IVR may benefit balance, mobility, cognition, and psychological well-being. However, some trials have reported mixed or modest effects [20-23]. These findings highlight the need for standardized outcome measures and adequately dosed programs in future research. These interventions have been implemented in both healthy individuals [21,24] and those with chronic medical conditions [22] or mild cognitive impairment [20]. However, existing reviews often involve short intervention periods [25]. This limits the ability to observe sustained long-term effects and may inflate immediate benefits [15,21]. Therefore, the focus is placed on programs delivered over periods that map onto the recall horizons of core quality-of-life instruments, a window that also coincides with the early emergence of neural and motor adaptations in response to training [26-30]. Trials in older adults have reported measurable gains over such intervals—cognition, balance, and functional outcomes [31-34]. Importantly, centering analyses on these exposure windows reduces the susceptibility of very brief interventions to novelty and Hawthorne effects [35]. In line with the World Health Organization's (WHO) healthy aging agenda, which prioritizes functional capacity and well-being as central goals in aging societies, we selected quality of life and physical activity as primary outcomes. These measures are patient-centered indicators of healthy aging and modifiable behavior, making them particularly relevant assessment criteria [36].

Prior reviews in older adults have primarily established that virtual reality (VR) is acceptable and feasible, while noting that evidence for effectiveness remains limited [15]. Other syntheses have focused on exergames and outcomes such as balance in long-term care facilities, aggregating heterogeneous exposure durations and without a prespecified minimum dose [37]. More recent randomized controlled trial (RCT)-only reviews of head-mounted display VR suggest benefits for physical activity and broader well-being, but do not isolate ≥4-week programs as an a priori inclusion criterion [38]. To our knowledge, no systematic review has focused exclusively on IVR interventions in older adults with a minimum duration of ≥4 weeks as an a priori inclusion criterion.

The objective of this systematic review and meta-analysis is to evaluate the effectiveness of IVR interventions lasting 4 weeks or more on the quality of life, physical activity, pain, perceived effort, and adverse events in older adults.

Methods

The protocol for this systematic review and meta-analysis was registered in the International Prospective Register of Systematic Reviews (CRD420251019170). It was conducted following the recommendations of the Cochrane Collaboration and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Checklist 1) [39]

Search Strategy

A bibliographic search was completed between March 15, 2025, and June 15, 2025, in all the following databases: PubMed, Web of Science, PEDro, and Scopus. Additionally, we systematically searched for gray literature. This included searches in ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform, ProQuest Dissertations & Theses Global, and medRxiv. Search strategies used are available in Multimedia Appendix 1 [14,18,24,29,40-53].

We also carried out a "snowball" search to identify additional studies by searching the reference lists of publications eligible for full-text review.

XSL•FO
RenderX

Inclusion and Exclusion Criteria

Inclusion criteria were defined according to the PICOS (population, intervention, comparison, outcome, study design) framework: (1) we included studies only if all participants were aged ≥60 years at baseline; (2) IVR intervention, alone or combined with other therapies, lasting 4 weeks or more; (3) the comparison group can be control, placebo, or another type of intervention; (4) outcome measures related to physical activity, quality of life, pain intensity, perceived effort, and adverse events; and (5) randomized controlled trials (RCTs), crossover clinical trials, and randomized mixed methods studies.

Exclusion criteria were: (1) studies with participants aged under 60 years, although the average age is 60 years or older; and (2) studies whose intervention is defined as immersive but does not use the head-mounted display.

No language restrictions were applied while searching.

Study Selection

Once the research question was defined, studies were identified and screened accordingly. Following the search strategy, all references were imported into Rayyan (Rayyan Systems Inc) to exclude duplicate studies [54].

Two independent researchers (IT-C and JB-A) conducted the study selection based on the predefined inclusion and exclusion criteria. In cases of disagreement, a third researcher (HB-A) was consulted to resolve discrepancies and reach a consensus.

Data Extraction

First, two researchers (IT-C and JB-A) independently extracted key information from the included studies. In case of discrepancies, a third researcher was consulted to resolve disagreements. The extracted data included: first author and year of publication, country, study design, sample size, age, diagnosis, setting, type of intervention of the groups, hardware and software of VR used, time of intervention (total amount of time in h, session time, frequency, and number of wk), outcomes, and follow-up period.

In the second phase, quantitative data for both primary and secondary outcomes were extracted. For the primary outcomes—physical activity and quality of life—as well as for the secondary outcomes of pain, perceived effort, and adverse events, mean and SD values were collected. When studies reported both change scores and final values, the final values were prioritized for analysis. When information regarding any of the above was unclear, we contacted authors of the reports to provide further details.

If data were only available in graphs, the graph digitization software GraphGrabber 2.0.2 (Quintessa Ltd) was used for extraction [55].

Risk of Bias

The methodological quality of the included studies was independently assessed by 2 researchers using the Cochrane Risk of Bias 2.0 tool for both parallel and crossover designs [56]. In instances where discrepancies arose between the two primary researchers, a third independent researcher was consulted to resolve disagreements. This tool evaluates the risk of bias across 5 domains: randomization process, missing outcome data, measurement of outcomes, selection of reported results, and deviations from intended interventions.

In addition, selective reporting will be judged by comparing published outcomes with registered protocols when available.

Main Outcomes

The primary outcomes of this study were quality of life and physical activity. Any objective or self-reported measure of these outcomes was considered eligible for inclusion, provided the instruments used demonstrated evidence of validity and reliability. No restrictions were placed on the number or timing of assessment time points (eg, baseline, postintervention, and follow-up).

The secondary outcomes included pain intensity, perceived effort, and adverse events. The same eligibility and selection criteria were applied, with preference given to the numeric rating scale (NRS) and visual analog scale (VAS) for pain intensity [57].

Statistical Analysis

A meta-analysis was conducted to estimate the overall effect of IVR interventions on quality of life and physical activity levels in older adults. As the included studies used different scales to assess these outcomes, the standardized mean difference (SMD) was used as the effect size measure. In all cases, higher scores indicated better outcomes, except for the study by Rodríguez-Fuentes et al [43], which used the Parkinson Disease Questionnaire (PDQ-39) scale, where lower scores reflect a better quality of life.

All hypothesis tests were 2-tailed, with statistical significance set at α=.05. Effect estimates are reported with 95% CIs. The inverse variance method was applied using a random-effects model, considering the expected clinical and methodological heterogeneity among studies. For each outcome, 95% CI were calculated, and a significance level of P<.05 was established. The magnitude of the effect was interpreted according to the Cohen criteria: small effect (SMD approximately 0.2), moderate (approximately 0.5), and large (≥0.8).

Statistical heterogeneity was assessed using the chi-square test, the ^{2} estimate, and the I^2 statistic. Heterogeneity was established as low for I^2=25%, moderate for I^2=50%, and high for I^2=75%. Subgroup analyses were conducted based on type of control group (active vs passive), intervention duration in weeks (6 - 8wk vs 10 - 12 wk), and total exposure time (180 - 480 min vs 600 - 1800 min). In addition, a sensitivity analysis was conducted by sequentially excluding each individual study to assess its impact on the overall effect size and heterogeneity.

All statistical analyses were performed using Review Manager (RevMan) version 5.4.1. The certainty of the evidence will be assessed using the GRADE (Grades of Recommendation Assessment Development and Evaluation) approach [58].

Deviations From Protocol

Several deviations from the International Prospective Register of Systematic Reviews protocol occurred. The mental component of the 12-item Short Form Survey (SF-12) was not

XSL•FO
RenderX

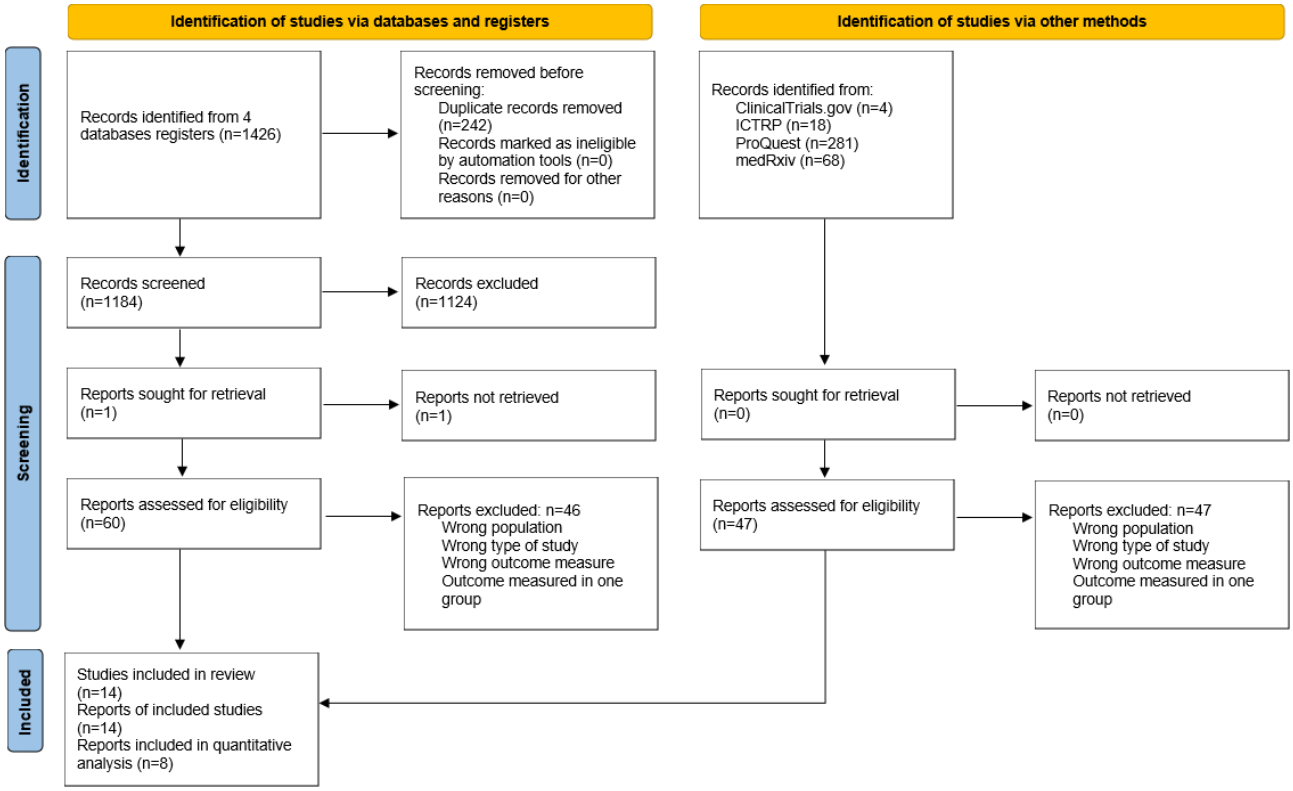
analyzed to avoid conceptual overlap (Multimedia Appendix 1 [14,18,24,29,40-53]). Publication bias assessment was planned but not performed due to <10 studies per outcome (Multimedia Appendix 1 [14,18,24,29,40-53]). When required, means and SDs were estimated from median (IQR) following validated procedures (Multimedia Appendix 1 [14,18,24,29,40-53]). Subgroup analyses were restricted to comparator type, intervention duration, and total exposure; definitions and limitations are detailed in Multimedia Appendix 1 [14,18,24,29,40-53].

Results

Study Selection

The selection process is illustrated in the PRISMA 2020 flow diagram (Figure 1) in this review.

Figure 1. Flow diagram in this review. ICTRP: International Clinical Trials Registry Platform.



We found 1426 records in database searching. After duplicate removal, we screened 1184 records, from which we reviewed 60 full-text documents. Later, we searched documents that cited any of the initially included studies as well as the references of the initially included studies. However, no extra studies that fulfilled the inclusion criteria were found in these searches.

Ultimately, 14 trials met the inclusion criteria and were included in the systematic review. Of these, 8 studies were included in the meta-analysis. The remaining 6 studies were excluded from the quantitative synthesis because they did not report outcome measures corresponding to the primary endpoints analyzed in the meta-analysis (ie, quality of life or physical activity).

Study Characteristics

This systematic review included 14 RCTs published between 2021 and 2025 [14,19,25,43-53], with a total of 839 older adults and sample sizes ranging from 9 to 293 participants. While most studies included older adults without specific diagnoses [14,19,25,44,45], others focused on populations with cognitive impairment or frailty [46-48,50], balance and mobility issues [51], Parkinson disease [43], knee osteoarthritis [53], chronic low back pain [52], or individuals with joint arthroplasty [49].

Regarding settings, 10 studies were conducted in community centers or nursing homes [19,43-48,50-52], 3 in laboratory conditions [44,52], and 2 at home [14,53].

Most studies focused on physical activity delivered via VR, using environments that required movement [14,44,51]. In some cases, VR-based physical activity was combined with usual care or conventional rehabilitation, such as occupational therapy or kinesiotherapy [45,49]. Other interventions emphasized cognitive training delivered via VR [46], sometimes including underdesk ergometers for simultaneous stimulation [47,48]. Some protocols included education sessions, followed by structured VR-based exercise programs [25,53].

A few studies explored more therapeutic applications of VR, such as immersive reminiscence therapy [19], or multimodal pain management programs that integrated psychoeducation and movement therapy [52]. One study used VR to simulate daily living environments, aiming to promote autonomy in participants with cognitive frailty [50].

Control conditions varied: some used usual care [19,45,47,50], others applied active comparators such as conventional rehabilitation or group-based exercise [14,43,48,49,51-53] and a few used no-intervention or educational controls [25,44,46].

XSL•FO
RenderX

Session durations typically ranged from 15 to 60 minutes, with frequencies varying from 1 to 5 times per week over periods from 4 to 12 weeks. The longest interventions, by Lo et al [53] and Rodríguez-Fuentes et al [43], lasted 12 weeks.

Detailed tables summarizing the characteristics of each included study are provided in Table 1 (study characteristics) and Table 2 (intervention characteristics). In addition, detailed specifications of software, hardware, and outcome instruments are provided in Multimedia Appendix 1 [14,18,24,29,40-53].

XSL•FO
RenderX

Table . Study characteristics.

Study ID	Study design	Sample size	Age (y)	Diagnosis	Setting
Barsasella et al (2021) [44]	RCT[a]	n=60; VRG[b] n=29; CG[c] n=31	NI[d]	NSD[e]	Laboratory-based
Campo-Prieto et al (2022) [45]	RCT	n=24; VRG n=13; CG n=11	VRG: mean 85.08 (SD 8.48) CG: mean 84.82 (SD 8.1)	NSD	Clinical-based
Chiu et al (2023) [46]	RCT	n=60; VRG n=30; CG n=30	VRG: mean 80.7 (SD 8.8) CG: mean 80 (SD 7.9)	Cognitive impairment	Clinical-based
Drazich et al (2023) [24]	Pilot RCT	n=20; VRG n=10; CG n=10	Mean 74.1 (SD 6.5)	NSD	Clinical-based
Kershner et al (2024) [14]	Pilot RCT	n=9; VRG n=5; CG n=4	Mean 66.8 (SD 4.8)	NSD	Home-based
Khirallah Abd el Fatah et al (2024) [18]	RCT	n=60. VRG n=20; RT group n=20; CG n=20	Mean 66.68 (SD 4.22)	NSD	Clinical-based
Kwan et al (2021) [48]	Pilot RCT	n=15; VRG n=9; CG n=8	Median: 74 (IQR 9,5)	Cognitive frailty	Clinical-based
Kwan et al (2024) [47]	RCT	n=293; VRG n=146; CG n=147	Mean 74.5 (SD 6.8)	Cognitive frailty	Clinical-based
Lo et al (2024) [53]	Mixed methods pilot RCT	n=30; VRG n=15; CG n=15	Median: 63.5 (IQR 61.8 - 66.3)	Knee osteoarthritis	Home-based
Mazurek et al (2023) [49]	RCT	n=68; VRG n=34; CG n=34	Mean 69.59 (SD 6.16)	Hip or knee joint arthroplasty surgeries	Laboratory-based
Rodríguez-Fuentes et al (2024) [43]	RCT	n=52; VRG n=30; CG n=22	Mean 70.79 (SD 6.59)	Parkinson	Parkinson Association
Sekar et al (2024) [51]	RCT	n=60; VRG=30; CG=30	NI	Balance and mobility issues	Clinical-based
Stamm et al (2022) [52]	Pilot RCT	n=22; VRG n=11; CG n=11	VRG: mean 75.0 (SD 5.8) CG: mean 75.5 (SD 4.39)	Chronic low back pain	Laboratory-based
Zheng et al (2025) [50]	RCT	n=66; VRG n=33; CG n=33	Mean 80.20 (SD 9.14)	Cognitive frailty	Clinical-based

[a]RCT: randomized controlled trial.

[b]VRG: virtual reality group.

[c]CG: control group.

XSL•FO
RenderX

[d]NI: no information.

[e]NSD: no specific diagnosis.

XSL•FO
RenderX

Table . Intervention characteristics.

Study ID	Experimental group (s)	Control group	Frequency	Follow-up
Barsasella et al (2021) [44]	PA[a] via VR[b]	No intervention	15’ 2 t/w[c] 6 weeks TT:[d] 180’	—[e]
Campo-Prieto et al (2022) [45]	Usual care+PA via VR	Usual care (occupational therapy and memory workshops)	6’ 3 t/w 10 weeks TT: 180’	—
Chiu et al (2023) [46]	Cognitive training intervention via VR	No intervention	60’ 1 t/w 8 weeks TT: 480’	—
Drazich et al (2023) [24]	PA education+VR	1 session of PA education	40’ 2 t/w 8 weeks TT: 640’	—
Kershner et al (2024) [14]	PA via VR	PA via group videoconference	45 - 60’ minimum/w 4 weeks TT: 180 - 240’ minimum	—
Khirallah Abd el Fatah et al (2024) [18]	EXP 1: IVR[f] reminiscence therapy EXP 2: Traditional reminiscence therapy	Usual care (daily personal care, primary nursing care, medical care)	30 - 45’ 2 t/w 6 weeks TT: 360 - 540’	3 months
Kwan et al (2021) [48]	Motor and cognitive training on VR+underdesk ergometer	Cognitive training on tablet computers and motor training cycling on ergometer	30’ 2 t/w 8 weeks TT: 480’	—
Kwan et al (2024) [47]	Motor and cognitive training on VR+underdesk ergometer	Usual care (activities provided by the community centers)	20 - 30’ 2 t/w 8 weeks TT: 320 - 480’	—
Lo et al (2024) [53]	Health talk+lower limb exercises via VR	Lower limb exercises	30’ 5 t/w 12 weeks TT: 1800’	—
Mazurek et al (2023) [49]	Relaxing VR+conventional rehabilitation	Conventional rehabilitation (kinesiotherapy, ergotherapy, laser therapy/magnetic therapy/electrotherapy)	VR sessions: 20’ 2 t/w 4 weeks TT: 160’	—
Rodríguez-Fuentes et al (2024) [43]	Cycloergometer+VR	Static Cycling using Smart Cycloergometers	25’ 2 t/w 12 weeks TT: 600’	—
Sekar et al (2024) [51]	Balance and mobility exercises with VR	Balance and mobility exercises	2 t/w 8 weeks	—

XSL•FO
RenderX

Study ID	Experimental group (s)	Control group	Frequency	Follow-up
Stamm et al (2022) [52]	Movement therapy and psychoeducation via VR	Conventional multimodal pain therapy (chair-based group exercises and psychoeducation units)	30’ 3 t/w 4 weeks TT: 360’	—
Zheng et al (2025) [50]	Scenarios with daily environments via VR	Usual care (nursing care, and routine activities like finger exercises and holiday paper cutting)	45’ 2 t/w 12 weeks TT: 1080’	—

[a]PA: physical activity.

[b]VR: virtual reality.

[c]t/w: times/week.

[d]TT: total time.

[e]—: not available.

[f]IVR: immersive virtual reality.

Main Outcomes

Quality of Life

Among all 6 studies assessed quality of life using various validated tools such as EuroQol VAS [44,53], SF-12 [45,50], World Health Organization Quality of Life Scale Brief Version (WHOQOL-BREF) [43,46] and PDQ-39 [46,50].

Overall, 5 studies reported improvements in quality of life following IVR interventions. Chiu et al [46] and Zheng et al [50] found the most notable effects, with WHOQOL-BREF scores nearly doubling and significant gains in SF-12 mental health, respectively. Only Barsasella et al [44] found no overall differences in EuroQol 5D-3L scores between groups, though improvements were noted in specific domains such as pain and anxiety.

Physical Activity

Among all 3 studies assessed physical activity using both self-report and objective measures (eg, Yale Physical Activity Survey, Garmin Vivosmart 4, and ActivPAL accelerometry) [14,25,53].

Drazich et al [24] observed modest gains in vigorous activity within the VR group, despite stable weekly activity levels. Kershner et al [14] reported greater gains in steps and vigorous activity for the video conference group, though the VR group improved more in functional capacity. Lo et al [53] found slightly higher metabolic equivalent of tasks in the VR group, but without significant differences.

Intensity of Pain

Out of all 2 studies addressed pain. Lo et al [53] used the NRS and Western Ontario and McMaster Universities Osteoarthritis Index subscale in patients with osteoarthritis, showing reductions in both scores within the VR group (NRS: 5.93-4.78; Western Ontario and McMaster Universities Osteoarthritis Index: 189.5-160.2), with minimal changes in controls. However, these differences did not reach statistical significance when comparing IVR to control, suggesting no conclusive effect attributable to the intervention. Mazurek et al [49] reported a significant drop in VAS scores in the VR group (5.27-0.88), exceeding improvements in the control group.

Perceived Effort

None of the included studies systematically assessed perceived effort using validated tools such as the Borg rating of perceived exertion scale or comparable measures.

Adverse Events

Adverse events were reported in 7 of the included studies, focusing on cybersickness [14,19,25,45,47,48,53]. Overall, most studies reported minimal or mild adverse events. In the pilot study of Kwan et al [47,48], 1 participant in the VR group withdrew early due to persistent symptoms of cybersickness, while the RCT reported low incidence rates (0.7%-3%) across 293 participants. Kershner et al [14] observed mild symptoms present in some participants. Lo et al [53] found that 5 out of 15 participants in the VR group reported mild adverse events such as dizziness or visual fatigue, though these did not result in discontinuation.

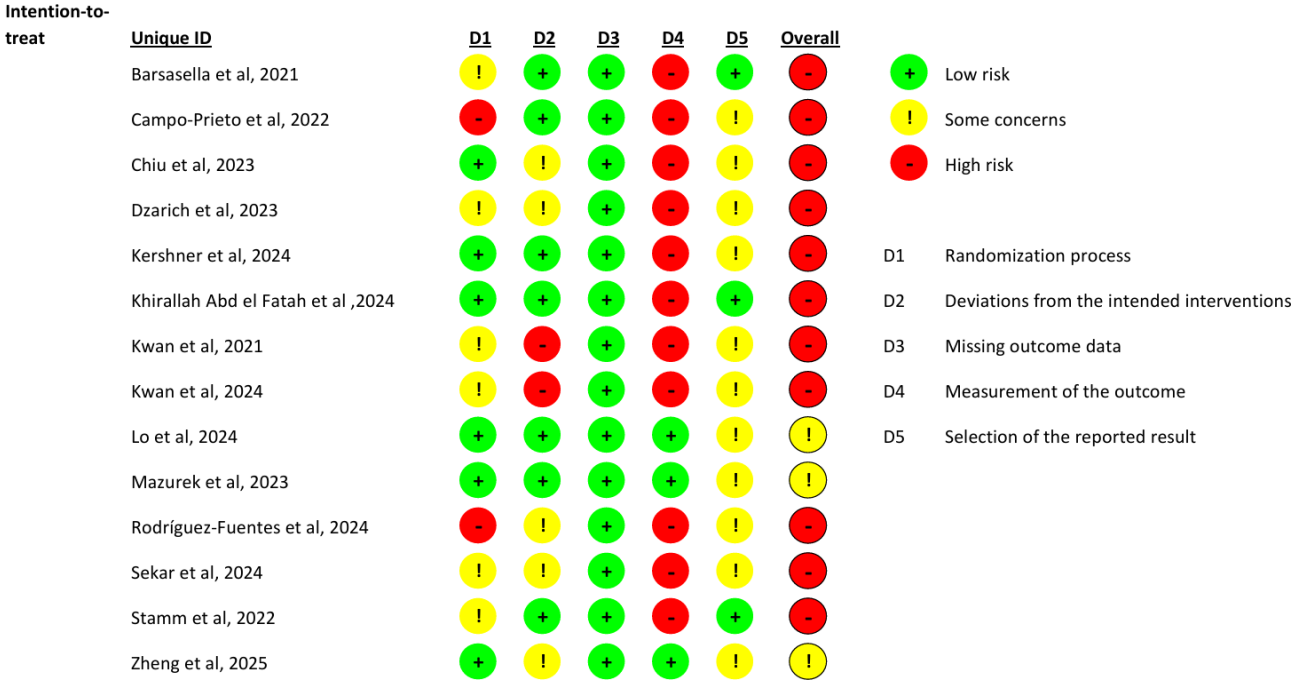
Risk of Bias in Studies

Most studies were judged to be at high risk of bias, except for 3 studies that showed a moderate risk [49,50,53]. All studies demonstrated low risk in domain 3 (missing outcome data). In domain 4 (measurement of the outcome), all but 3 studies were rated as high risk [49,50,53]. In domain 5 (selection of the reported result), 3 studies were rated as low risk, while the remaining studies were judged to have unclear risk.

As shown in Figure 2, the agreement rate achieved between the 2 researchers who completed risk of bias assessment was 81.43%. In case of disagreement, a third researcher resolved it.

XSL•FO
RenderX

Figure 2. Risk of bias [14,18,24,43-53].

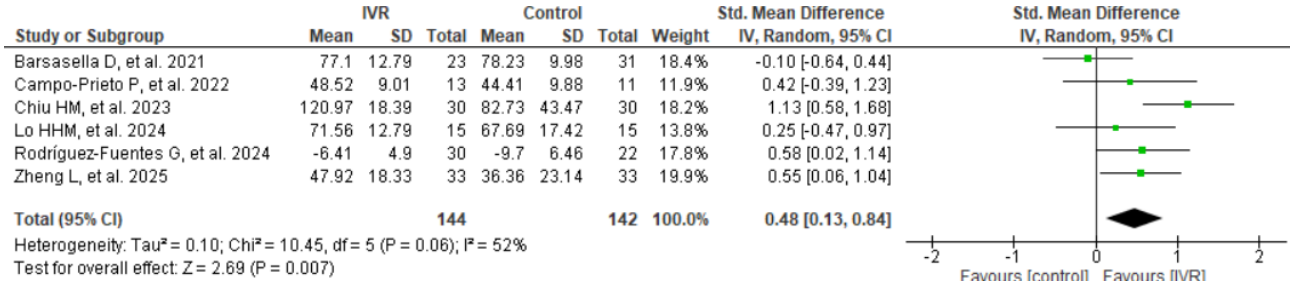


Results of Syntheses

Quality of Life

As shown in Figure 3 (effect of IVR on quality of life), a total of 6 RCTs including 286 participants (IVR group: n=144; control group: n=142) were synthesized to examine the effect of IVR interventions on quality of life in older adults. The meta-analysis yielded a statistically significant moderate effect in favor of IVR compared to control conditions (SMD=0.48, 95% CI 0.1-0.8; *P*=.007). Heterogeneity was moderate (I^2=52%; *P*=.06; ^{2}=0.10), indicating some variability across studies.

Figure 3. Effect of immersive virtual reality (IVR) on quality of life [43-46,50,53].



As shown in Table 3 subgroup analyses were performed to explore the potential influence of comparator type, intervention duration in weeks, and total intervention time in minutes.

XSL•FO
RenderX

Table . Subgroup analysis.

Subgroup	Studies, n	Participants, n	Random-effects model		Heterogeneity I^{2b} (%)	Between subgroups difference	
			SMD[a] (95% CI)	*P* value		Chi-square (*df*)	*P* value
Comparator type						0.02 (1)	.89
Passive control	4	110	0.50 (−0.04 to 1.04)	.07	71		
Active control	2	82	0.48 (0.12 to 0.84)	.05	0		
Weeks of intervention						0.00 (1)	.97
6 - 8	2	120	0.51 (−0.69 to 1.72)	.40	90		
10 - 12	4	172	0.48 (0.18 to 0.79)	.002	0		
Minutes of intervention						0.00 (1)	.99
180 - 480	3	138	0.49 (−0.31 to 1.28)	.23	80		
600 - 1800	3	286	0.49 (0.13 to 0.84)	.003	0		

[a]SMD: standardized mean difference.

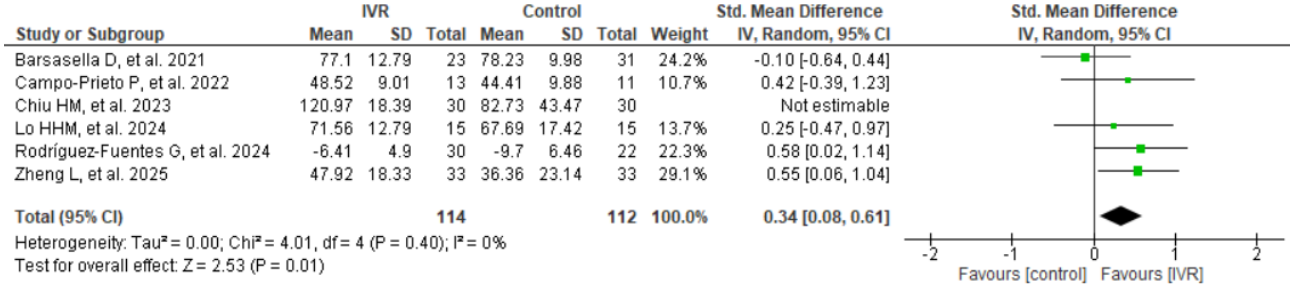
[b]I^2: inconsistency index.

Within-subgroup analyses showed statistically significant effects in some conditions. For instance, a significant effect was observed in the subgroup receiving 600 - 1800 minutes of intervention (SMD=0.49, 95% CI 0.13-0.84; *P*=.003), in contrast to the nonsignificant effect in the 180 - 480 minutes group. Similarly, the 10 - 12 week intervention subgroup showed a statistically significant moderate effect (SMD=0.48, 95% CI 0.18-0.79; *P*=.002), while the 6 - 8 week group did not (*P*=.40).

Regarding comparator type, significant effects were observed in the active control group (*P*=.05), but not in the passive control group (*P*=.07). However, despite these within-subgroup differences, the overall between-subgroup comparisons were not statistically significant for any of the 3 variables assessed: comparator type (*P*=.89), duration in weeks (*P*=.97), or total minutes of intervention (*P*=.99). These findings suggest that while effect sizes may vary descriptively across subgroups, such differences are not supported statistically.

As shown in Figure 4 (sensitivity analysis of IVR effects on quality of life), a sensitivity analysis was conducted excluding the study by Chiu et al [46], which presented the largest effect size (SMD=1.1). When this study was removed, the overall effect in favor of IVR remained statistically significant (SMD=0.3, 95% CI 0.08-0.6; *P*=.01), although the magnitude of the effect was reduced. Notably, heterogeneity was eliminated (I^2=0%; ^{2}=0.00), suggesting that this study contributed substantially to the observed heterogeneity in the main analysis.

Figure 4. Sensitivity analysis of IVR effects on quality of life. IVR: immersive virtual reality [43-46,50,53].



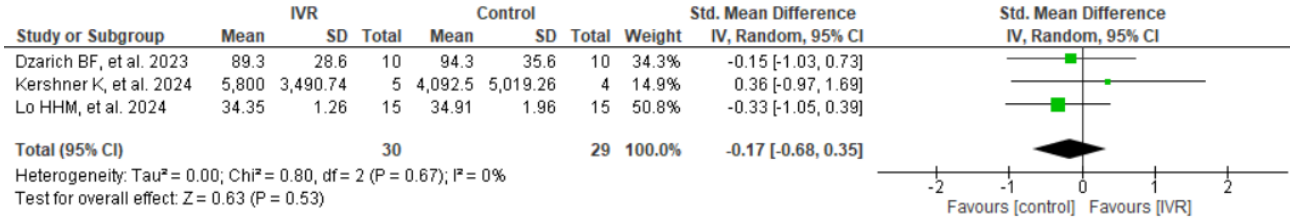
Physical Activity

Among all 3 RCTs involving a total of 59 participants (IVR group: n=30; control group: n=29), IVR was assessed on physical activity in older adults. The meta-analysis revealed no statistically significant difference between groups (SMD=−0.2, 95% CI −0.7 to 0.4; *P*=.50).

As shown in Figure 5 (effect of IVR on physical activity), heterogeneity among studies was minimal (I^2=0%; *P*=.70; ^{2}=0.00), indicating high consistency in the direction and magnitude of effects across studies. Due to the small number of studies, no subgroup or sensitivity analyses were performed for this outcome.

XSL•FO
RenderX

Figure 5. Effect of immersive virtual reality (IVR) on physical activity [14,24,53].



Certainty of Evidence

As shown in Figure 6 (assessment of evidence according to GRADE), the certainty of the evidence for the outcome quality of life was rated as low. Downgrading was applied due to very serious concerns related to the risk of bias. For the outcome of physical activity, the certainty of the evidence was rated as very low. Downgrading was applied due to serious risk of bias, very serious imprecision, and strong suspicion of publication bias.

Figure 6. Assessment of evidence according to GRADE (Grades of Recommendation Assessment Development and Evaluation). IVR: immersive virtual reality; SMD: standardized mean difference. *Out of 6 studies, 4 contributing to this outcome were judged to have high risk of bias, particularly due to issues in the randomization process, deviations from intended interventions, and selective reporting. The remaining 2 had some concerns. **Out of the 3 studies, 2 were judged to have high risk of bias, particularly in randomization and outcome assessment. The overall certainty was downgraded due to methodological limitations. ***The CI was wide and included both meaningful benefit and harm. The small sample size and lack of statistical significance led to downgrading by 2 levels. ****Only 3 studies were included in the meta-analysis for physical activity, which prevents formal assessment of publication bias. Due to the small number of trials and likelihood of selective reporting, strong suspicion of publication bias was considered.

Certainty assessment						N of patients		Effect		Certainty
N of studies	Study design	Risk of bias	Inconsistency	Imprecision	Other considerations	IVR	Control group	Relative (95% CI)	Absolute (95% CI)	
Quality of life										
6	Randomized trials	Very serious*	Not serious	Not serious	None	144	142	-	SMD 0.48 higher (0.13 higher to 0.84 higher)	Low*
Physical activity										
3	Randomized trials	Serious**	Not serious	Very serious***	Publication vias strongly suspected****	30	29	-	SMD 0.17 lower (0.68 lower to 0.35 higher)	Very low*,**,***,****

Discussion

Principal Findings

To our knowledge, this review is the first to focus exclusively on IVR interventions in older adults with a minimum duration of ≥4 weeks as a predefined inclusion criterion. The findings reveal a statistically significant moderate effect of IVR on quality of life, especially in individuals with clinical vulnerability or living in institutional settings. In contrast, no significant effects were observed for physical activity, and the direction of the results across included studies was inconsistent. Although some isolated findings indicated minor improvements in specific physical activity parameters, these were not replicated across trials and were derived from interventions of limited intensity or duration. As such, current evidence does not support definitive conclusions regarding the impact of IVR on physical activity, though it highlights important considerations for future program design. On the other hand, given the limited number of studies and heterogeneous measurement, secondary outcomes (pain, perceived exertion, and adverse events) could not be robustly quantified. We retained these outcomes because they were prespecified in our protocol and are clinically salient, but current evidence is insufficient to support firm conclusions. Additionally, preliminary results indicate potential benefits for pain reduction, and IVR was well tolerated, with low incidence of mild adverse events.

Quality of Life

Our findings suggest a moderate, statistically significant improvement in quality of life with IVR versus control; however, the certainty of evidence is low due to risk of bias. The observed effect size—approaching the 0.5 threshold commonly regarded as clinically meaningful in geriatric interventions—suggests a relevant improvement in this population. The psychological sense of presence elicited by IVR may diminish the perception of aversive stimuli such as pain or anxiety and promote emotional regulation, intrinsic motivation, and well-being [59,60].

Notably, the most pronounced benefits were evident among older adults with clinical diagnoses, indicating efficacy in individuals with functional vulnerabilities. This observation is consistent with previous research reporting cognitive gains in populations with mild cognitive impairment [61]. Moreover, interventions implemented in residential or institutional environments tended to yield more consistent and favorable outcomes. In support of this, Li et al [62] reported that IVR experiences in nursing homes enhanced not only well-being and social engagement but were also perceived as meaningful and motivating by participants.

Regarding intervention characteristics, programs incorporating cognitive or functional components generated more substantial improvements than those centered exclusively on physical exercise. This may reflect the inherently multidimensional nature

XSL•FO
RenderX

of quality of life, which integrates cognitive, emotional, and social dimensions in addition to physical health [63].

In terms of duration, longer interventions (10 - 12 wk or exceeding 600 min) could be associated with more reliable improvements in quality of life. This finding is aligned with results from Vasodi et al [64], who reported that extended IVR programs led to better outcomes in older adults' mood and well-being, potentially due to increased engagement, gradual adaptation, and the consolidation of behavioral changes over time.

Physical Activity

The effectiveness of IVR in promoting physical activity among older adults appears to be highly contingent on intervention design; however, in our review, the evidence is very uncertain. The meta-analysis included only 3 randomized trials and showed no statistically significant differences between IVR and control groups, with inconsistent directions of effect across studies. Coupled with the very low certainty of evidence, small total sample size, and wide confidence intervals, these findings should be interpreted as hypothesis-generating rather than decision-informing. These findings align with previous evidence suggesting that brief, lab-based programs are generally insufficient to generate sustained behavioral change. For instance, studies by Lo et al [53] and Drazich et al [24] reported only modest or transient increases in activity, likely due to seated, low-intensity exercises and short durations. In contrast, research in younger adults shows IVR can elicit greater physiological responses—such as increased oxygen consumption and enjoyment—when compared to traditional 2D formats, likely due to enhanced emotional engagement [65].

The success of longer, home-based interventions like those reported by Dinet and Nouchi [66] may reflect the critical importance of habit formation and environmental integration, which require extended exposure periods and real-world application contexts. However, such outcomes seem contingent on the intervention's ability to integrate into daily routines, adapt to user capacity, and sustain motivation over time.

A further limitation is the reliance on self-reported physical activity measures, which are subject to bias. Future research should use wearable devices with validated protocols for objective monitoring. Overall, IVR can support physical activity, but only if programs are engaging, adaptable, and promote long-term autonomous use.

Intensity of Pain

The effects of IVR on pain intensity in older adults show heterogeneous results, probably conditioned by the design of the interventions and the characteristics of the participants. In the pilot study by Lo et al [53] with older people with knee osteoarthritis, the differences in pain intensity did not reach statistical significance compared to the control group, which may be attributed to the limited format of the intervention, focusing exclusively on strength without aerobic elements or relevant visual distracters.

In contrast, the study by Mazurek et al [49] shows more robust results: after 8 sessions of IVR with psychotherapeutic approach and immersive relaxation, a significant reduction in pain was observed, with a significant difference compared to the control group. This effect could be explained by the inclusion of psychological components, such as attentional distraction, emotional reinforcement, and the use of therapeutic metaphors, which have been associated with a downward modulation of pain in previous neurobiological studies [67,68].

Additional studies reinforce this approach. Li et al [69] observed a significant reduction in chronic low back pain in older adults after 8 weeks of IVR combined with functional exercise, highlighting increased adherence and reduced analgesic use in the IVR group. Taken together, the evidence suggests that IVR may be effective in reducing pain in older adults, especially when applied with a multisensory, emotional, and adaptive approach.

Adverse Events

The results of this review suggest that IVR is generally well tolerated by older adults, with a low incidence of adverse events, mainly related to mild symptoms of cybersickness (dizziness, nausea, and visual fatigue). Studies such as those by Campo-Prieto et al [45], Drazich et al [24], or Khirallah et al [18] reported no adverse events, while others, such as Lo et al [53], did observe mild symptoms in one third of participants, with no related dropouts. In Kwan et al [47,48], one withdrawal due to persistent cybersickness was reported, but in their later study, with a larger sample size, the incidence was low (0.7% - 3%). These findings are consistent with previous reviews. Weech et al [70], in a systematic review, identified that symptoms of cybersickness are common in immersive environments, but their severity tends to be mild and dependent on factors such as content type, duration of exposure, and individual characteristics. Stanney et al [71] emphasized the importance of individualized visor fit, especially interpupillary distance, noting that poor fit significantly increases discomfort, especially in women.

Overall, the evidence suggests good overall tolerability, although not without some episodes.

Clinical Implications

The findings suggest that IVR may be a valuable clinical tool to enhance the quality of life in older adults, particularly those with functional limitations or in institutional settings. Given the predominance of high risk of bias across several trials and low to very low certainty by GRADE, the pooled effects should be interpreted as signals of possible benefit, not as precise estimates for clinical decision-making.

Interventions combining cognitive, physical, and motivational elements show greater effectiveness. Although physical activity outcomes were not significant, appropriate program design appears crucial for adherence. Preliminary evidence also supports IVR's use in pain management, especially with emotionally engaging approaches. Its low rate of adverse events indicates good tolerability. With proper device adjustment and supervision, IVR can be progressively integrated into geriatric rehabilitation across various care environments.

XSL•FO
RenderX

From a policy perspective, our findings align with the UN Decade of Healthy Aging, particularly its priority on maintaining functional ability and person-centered care. IVR could act as an engagement-enhancing tool when embedded within routine health and social care services [72]. Operationally, its integration should be guided by WHO's integrated care for older people pathways in primary and community care and adhere to the principles of the Global Strategy on Digital Health—namely interoperability, equity, and evidence-based implementation [72].

Limitations

This systematic review and meta-analysis present several limitations related to its design and execution. First, although the search strategy was comprehensive and included 4 major databases, relevant studies indexed in other sources may have been missed. However, no language restrictions were applied, which mitigates selection bias. Second, the number of studies eligible for inclusion in the meta-analysis was limited, particularly for some outcomes such as physical activity and pain, reducing the statistical power and precision of the effect estimates. Additionally, due to the low number of included trials per outcome (n<10), we could not perform funnel plots or the Egger test to formally assess publication bias. Third, the heterogeneity in outcome measurement instruments and reporting formats across studies made it necessary to apply specific criteria for data inclusion, such as prioritizing final values over change scores and estimating means and SD when only medians and interquartile ranges were reported. These decisions, although methodologically justified, may introduce some degree of imprecision. Finally, some planned subgroup analyses could not be conducted due to insufficient data availability. While sensitivity analyses were performed where possible, the overall ability to explore sources of heterogeneity was limited.

A wide range of instruments was used to assess quality of life and physical activity (SF-12, EuroQol-5 Dimensions, WHOQOL-BREF, PDQ-39, accelerometers, and self-report surveys). While we synthesized conceptually similar constructs, instrument heterogeneity reduces scale-specific interpretability and likely contributed to between-study variance.

Future RCTs should address current evidence gaps by prioritizing: (1) the use of standardized and psychometrically validated outcome measures, particularly for physical activity and perceived effort; (2) objective monitoring of physical activity levels through wearable devices to reduce reliance on self-report; (3) long-term follow-up assessments to determine the durability of IVR effects on quality of life and functional outcomes; and (4) cost-effectiveness analyses to evaluate the feasibility of implementing IVR programs in real-world geriatric care settings; and (5) future studies should focus on the analysis of contextual factors—such as delivery model (standalone IVR vs adjunct), supervision intensity, and care setting—as potential effect modifiers. In addition, the main bias in the review was due to the unblinding of assessors and nonregistered protocols; therefore, future studies should (1) blind outcome assessors, and (2) prospectively register protocols with a prespecified analysis plan to prevent selective reporting.

Conclusions

IVR interventions lasting 4 weeks or more appear to moderately improve quality of life in older adults, especially those with clinical vulnerability or living in institutional settings. In contrast, no significant effects were observed for physical activity, and available evidence does not support a consistent trend in favor of IVR. Further research is needed to determine whether specific program designs could enhance its impact in this domain. Pain, perceived effort, and adverse events are included as secondary outcomes, but the evidence is sparse and largely qualitative. Policies and previous research also support the potential of IVR in reducing pain, particularly when using multisensory and emotionally engaging approaches, although the certainty of evidence is low to very low according to GRADE; therefore, these findings should be interpreted with caution. Overall, IVR is well tolerated and shows promise as a safe, adaptable, and motivating tool to support healthy aging, warranting further research in diverse settings.

Funding

This research received no external funding. Part of JBA's and ITC's salaries were financed by the European Social Fund Plus.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study information.
[DOCX File, 45 KB - aging_v9i1e80820_app1.docx]

Checklist 1
PRISMA checklist.
[DOCX File, 274 KB - aging_v9i1e80820_app2.docx]

References

XSL•FO
RenderX

1. World population prospects 2022: summary of results. : United Nations; 2022 URL: https://www.un.org/development/desa/pd/content/World-Population-Prospects-2022 [accessed 2025-12-12]
2. McPhee JS, French DP, Jackson D, Nazroo J, Pendleton N, Degens H. Physical activity in older age: perspectives for healthy ageing and frailty. Biogerontology 2016 Jun;17(3):567-580. [doi: 10.1007/s10522-016-9641-0] [Medline: 26936444]
3. Bellantuono I. Find drugs that delay many diseases of old age. Nature New Biol 2018 Feb 15;554(7692):293-295. [doi: 10.1038/d41586-018-01668-0] [Medline: 29446384]
4. Aging and health [Web page in Spanish]. World Health Organization. URL: https://www.who.int/es/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-12-24]
5. Daskalopoulou C, Stubbs B, Kralj C, Koukounari A, Prince M, Prina AM. Physical activity and healthy ageing: a systematic review and meta-analysis of longitudinal cohort studies. Ageing Res Rev 2017 Sep;38(6–17):6-17. [doi: 10.1016/j.arr.2017.06.003] [Medline: 28648951]
6. Cunningham C, O' Sullivan R, Caserotti P, Tully MA. Consequences of physical inactivity in older adults: a systematic review of reviews and meta-analyses. Scand J Med Sci Sports 2020 May;30(5):816-827. [doi: 10.1111/sms.13616] [Medline: 32020713]
7. Roberts CE, Phillips LH, Cooper CL, Gray S, Allan JL. Effect of different types of physical activity on activities of daily living in older adults: systematic review and meta-analysis. J Aging Phys Act 2017 Oct 1;25(4):653-670. [doi: 10.1123/japa.2016-0201] [Medline: 28181837]
8. Alanazi MA. The role of physical activity in adjunctive nursing management of neuro-degenerative diseases among older adults: a systematic review of interventional studies. Life (Basel) 2024 May 7;14(5):597. [doi: 10.3390/life14050597] [Medline: 38792618]
9. Olanrewaju O, Kelly S, Cowan A, Brayne C, Lafortune L. Physical activity in community dwelling older people: a systematic review of reviews of interventions and context. PLoS ONE 2016;11(12):e0168614. [doi: 10.1371/journal.pone.0168614] [Medline: 27997604]
10. Fukushima N, Kikuchi H, Sato H, et al. Dose-response relationship of physical activity with all-cause mortality among older adults: an umbrella review. J Am Med Dir Assoc 2024 Mar;25(3):417-430. [doi: 10.1016/j.jamda.2023.09.028] [Medline: 37925162]
11. Picorelli AMA, Pereira LSM, Pereira DS, Felício D, Sherrington C. Adherence to exercise programs for older people is influenced by program characteristics and personal factors: a systematic review. J Physiother 2014 Sep;60(3):151-156. [doi: 10.1016/j.jphys.2014.06.012] [Medline: 25092418]
12. Schutzer KA, Graves BS. Barriers and motivations to exercise in older adults. Prev Med 2004 Nov;39(5):1056-1061. [doi: 10.1016/j.ypmed.2004.04.003] [Medline: 15475041]
13. Grassini S, Laumann K. Immersive visual technologies and human health. Presented at: ECCE '21: Proceedings of the 32nd European Conference on Cognitive Ergonomics; Apr 26-29, 2021. [doi: 10.1145/3452853.3452856]
14. Kershner K, Morton D, Robison J, N'dah KW, Fanning J. Assessing the feasibility and acceptability of virtual reality for remote group-mediated physical activity in older adults: pilot randomized controlled trial. JMIR Form Res 2024 Nov 8;8:e53156. [doi: 10.2196/53156] [Medline: 39514256]
15. Van Veelen N, Boonekamp R, Schoonderwoerd T, et al. Tailored immersion: implementing personalized components into virtual reality for veterans with post-traumatic stress disorder. Eur Psychiatr 2022 Jun;65(S1):S675-S675. [doi: 10.1192/j.eurpsy.2022.1737]
16. Cucinella SL, de Winter JCF, Grauwmeijer E, Evers M, Marchal-Crespo L. Towards personalized immersive virtual reality neurorehabilitation: a human-centered design. J Neuroeng Rehabil 2025 Jan 20;22(1):7. [doi: 10.1186/s12984-024-01489-5] [Medline: 39833912]
17. Wada T, Matsumoto H, Hagino H. Customized exercise programs implemented by physical therapists improve exercise-related self-efficacy and promote behavioral changes in elderly individuals without regular exercise: a randomized controlled trial. BMC Public Health 2019 Jul 9;19(1):917. [doi: 10.1186/s12889-019-7270-7] [Medline: 31288781]
18. Khirallah Abd El Fatah N, Abdelwahab Khedr M, Alshammari M, Mabrouk Abdelaziz Elgarhy S. Effect of immersive virtual reality reminiscence versus traditional reminiscence therapy on cognitive function and psychological well-being among older adults in assisted living facilities: a randomized controlled trial. Geriatr Nurs 2024;55:191-203. [doi: 10.1016/j.gerinurse.2023.11.010] [Medline: 38007908]
19. Porras-Garcia B, Rojas-Rincón J, Adams A, Garolera M, Chang R. Immersive virtual reality cognitive training for improving cognition and depressive symptoms among older adults. current evidence and future recommendations. a systematic review. Cyberpsychol Behav Soc Netw 2024 Oct;27(10):692-703. [doi: 10.1089/cyber.2024.0090] [Medline: 39180434]
20. Lee J, Phu S, Lord SR, Okubo Y. Effects of immersive virtual reality training on balance, gait and mobility in older adults: a systematic review and meta-analysis. Gait Posture 2024 May;110:129-137. [doi: 10.1016/j.gaitpost.2024.03.009] [Medline: 38581933]
21. Maheta B, Kraft A, Interrante N, et al. Using virtual reality to improve outcomes related to quality of life among older adults with serious illnesses: systematic review of randomized controlled trials. J Med Internet Res 2025 Feb 26;27:e54452. [doi: 10.2196/54452] [Medline: 40009834]

XSL•FO
RenderX

22. Dermody G, Whitehead L, Wilson G, Glass C. The role of virtual reality in improving health outcomes for community-dwelling older adults: systematic review. J Med Internet Res 2020 Jun 1;22(6):e17331. [doi: 10.2196/17331] [Medline: 32478662]
23. Ke Z, Wei M, Yang F, et al. The effectiveness of immersive virtual reality on the psychology of older adults: a systematic review and meta-analysis of randomized controlled trials. Gen Hosp Psychiatry 2025;94(86–96):86-96. [doi: 10.1016/j.genhosppsych.2025.02.024] [Medline: 40036985]
24. Drazich BF, Anokye D, Zhu S, et al. Motivating older adults through immersive virtual exercise (motive): a randomized pilot study. Geriatr Nurs 2023;54:229-236. [doi: 10.1016/j.gerinurse.2023.09.019] [Medline: 37844539]
25. Doré B, Gaudreault A, Everard G, et al. Acceptability, feasibility, and effectiveness of immersive virtual technologies to promote exercise in older adults: a systematic review and meta-analysis. Sensors (Basel) 2023 Feb 24;23(5):2506. [doi: 10.3390/s23052506] [Medline: 36904709]
26. Chen PJ, Hsu HF, Chen KM, Belcastro F. VR exergame interventions among older adults living in long-term care facilities: a systematic review with meta-analysis. Ann Phys Rehabil Med 2023 Apr;66(3):101702. [doi: 10.1016/j.rehab.2022.101702] [Medline: 36028201]
27. Kim KA, Ahn JA. Effectiveness of immersive virtual reality simulation programs using head-mounted displays in promoting physical activity in older adults: a systematic review. Clin Simul Nurs 2024 Sep;94:101593. [doi: 10.1016/j.ecns.2024.101593]
28. John W, Kosinski M, Gandek B. SF-36 Health Survey: Manual & Interpretation Guide 1993. URL: https://czresearch.com/info/SF36_healthsurvey_ch6.pdf [accessed 2025-12-12]
29. Ware JE, Kosinski M, Diane TB. How to Score SF-12 Physical and Mental Health Summary Scales, 2nd edition: The Health Institute, New England Medical Center; 1995.
30. Rong W, Geok SK, Samsudin S, Zhao Y, Ma H, Zhang X. Effects of strength training on neuromuscular adaptations in the development of maximal strength: a systematic review and meta-analysis. Sci Rep 2025 Jun 2;15(1):19315. [doi: 10.1038/s41598-025-03070-z] [Medline: 40456806]
31. Škarabot J, Brownstein CG, Casolo A, Del Vecchio A, Ansdell P. The knowns and unknowns of neural adaptations to resistance training. Eur J Appl Physiol 2021 Mar;121(3):675-685. [doi: 10.1007/s00421-020-04567-3] [Medline: 33355714]
32. Brown N, Bubeck D, Haeufle DFB, et al. Weekly time course of neuro-muscular adaptation to intensive strength training. Front Physiol 2017;8(JUN):329. [doi: 10.3389/fphys.2017.00329] [Medline: 28642711]
33. Nouchi R, Taki Y, Takeuchi H, et al. Four weeks of combination exercise training improved executive functions, episodic memory, and processing speed in healthy elderly people: evidence from a randomized controlled trial. Age (Dordr) 2014 Apr;36(2):787-799. [doi: 10.1007/s11357-013-9588-x] [Medline: 24065294]
34. Jagdhane S, Kanekar N, Aruin AS. The effect of a four-week balance training program on anticipatory postural adjustments in older adults: a pilot feasibility study. Curr Aging Sci 2016;9(4):295-300. [doi: 10.2174/1874609809666160413113443] [Medline: 27071477]
35. Yue T, Yan X, Lv Y, Ren X, Zhang S, Qi F. Effects of immersive virtual reality-based tennis training on balance in older adults. Gerontology 2025;71(10):861-871. [doi: 10.1159/000547548] [Medline: 40695260]
36. Sadowski I, Meilleur-Bédard M, Khoury B. A novel virtual reality-based nature meditation program for older adults' mental health: results from a pilot randomized controlled trial. Clin Gerontol 2025;48(4):663-683. [doi: 10.1080/07317115.2025.2482089] [Medline: 40152067]
37. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol 2014 Mar;67(3):267-277. [doi: 10.1016/j.jclinepi.2013.08.015] [Medline: 24275499]
38. WHO's work on the UN decade of healthy ageing (2021-2030). World Health Organization. URL: https://www.who.int/initiatives/decade-of-healthy-ageing [accessed 2025-12-12]
39. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
40. Shi J, Luo D, Weng H, et al. Optimally estimating the sample standard deviation from the five - number summary. Res Synth Methods 2020 Sep;11(5):641-654. [doi: 10.1002/jrsm.1429]
41. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. Stat Methods Med Res 2018 Jun;27(6):1785-1805. [doi: 10.1177/0962280216669183] [Medline: 27683581]
42. Cheng LJ, Tan RLY, Luo N. Measurement properties of the EQ VAS around the globe: a systematic review and meta-regression analysis. Value Health 2021 Aug;24(8):1223-1233. [doi: 10.1016/j.jval.2021.02.003] [Medline: 34372988]
43. Rodríguez-Fuentes G, Campo-Prieto P, Cancela-Carral JM. Immersive virtual reality high-intensity aerobic training to slow parkinson's disease: the ReViPark program. Appl Sci ;14(11):4708. [doi: 10.3390/app14114708]
44. Barsasella D, Liu MF, Malwade S, et al. Effects of virtual reality sessions on the quality of life, happiness, and functional fitness among the older people: a randomized controlled trial from taiwan. Comput Methods Programs Biomed 2021 Mar;200:105892. [doi: 10.1016/j.cmpb.2020.105892] [Medline: 33280934]

XSL•FO
RenderX

45. Campo-Prieto P, Cancela-Carral JM, Rodríguez-Fuentes G. Feasibility and effects of an immersive virtual reality exergame program on physical functions in institutionalized older adults: a randomized clinical trial. Sensors (Basel) 2022 Sep 6;22(18):6742. [doi: 10.3390/s22186742] [Medline: 36146092]
46. Chiu HM, Hsu MC, Ouyang WC. Effects of incorporating virtual reality training intervention into health care on cognitive function and wellbeing in older adults with cognitive impairment: a randomized controlled trial. Int J Hum Comput Stud 2023 Feb;170:102957. [doi: 10.1016/j.ijhcs.2022.102957]
47. Kwan RYC, Liu J, Sin OSK, et al. Effects of virtual reality motor-cognitive training for older people with cognitive frailty: multicentered randomized controlled trial. J Med Internet Res 2024 Sep 11;26:e57809. [doi: 10.2196/57809] [Medline: 39259959]
48. Kwan RYC, Liu JYW, Fong KNK, et al. Feasibility and effects of virtual reality motor-cognitive training in community-dwelling older people with cognitive frailty: pilot randomized controlled trial. JMIR Serious Games 2021 Aug 6;9(3):e28400. [doi: 10.2196/28400] [Medline: 34383662]
49. Mazurek J, Cieślik B, Wrzeciono A, Gajda R, Szczepańska-Gieracha J. Immersive virtual reality therapy is supportive for orthopedic rehabilitation among the elderly: a randomized controlled trial. J Clin Med 2023 Dec 14;12(24):7681. [doi: 10.3390/jcm12247681] [Medline: 38137750]
50. Zheng L, Li X, Xu Y, et al. Effects of virtual reality-based activities of daily living rehabilitation training in older adults with cognitive frailty and activities of daily living impairments: a randomized controlled trial. J Am Med Dir Assoc 2025 Feb;26(2):105397. [doi: 10.1016/j.jamda.2024.105397] [Medline: 39615543]
51. Sekar M, Suganthirababu P, Subramanian SS, et al. The effectiveness of virtual reality (VR) therapy on balance and mobility in elderly patients: a randomized controlled trial. Fiz Pol 2024 Dec 31;24(5):191-194. [doi: 10.56984/8ZG020C8UWP]
52. Stamm O, Dahms R, Reithinger N, Ruß A, Müller-Werdan U. Virtual reality exergame for supplementing multimodal pain therapy in older adults with chronic back pain: a randomized controlled pilot study. Virtual Real 2022;26(4):1291-1305. [doi: 10.1007/s10055-022-00629-3] [Medline: 35194374]
53. Lo HHM, Ng M, Fong PYH, et al. Examining the feasibility, acceptability, and preliminary efficacy of an immersive virtual reality-assisted lower limb strength training for knee osteoarthritis: mixed methods pilot randomized controlled trial. JMIR Serious Games 2024 Sep 27;12(2024):e52563. [doi: 10.2196/52563] [Medline: 39331525]
54. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
55. Graph Grabber 2.0.2. Quintessa. 2020. URL: https://www.quintessa.org/software/downloads-and-demos/graph-grabber-2.0.2 [accessed 2025-12-12]
56. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
57. Vicente Herrero MT, Delgado Bueno S, Bandrés Moyá F, Ramírez Iñiguez de la Torre MV, Capdevila García L. Pain assessment. Comparative review of scales and questionnaires [Title in Spanish]. Rev Soc Esp Dolor 2018;25(4):228-236. [doi: 10.20986/resed.2018.3632/2017]
58. McMaster University. GRADEpro. URL: https://www.gradepro.org/ [accessed 2025-12-12]
59. Riva G, Wiederhold BK, Mantovani F. Neuroscience of virtual reality: from virtual exposure to embodied medicine. Cyberpsychol Behav Soc Netw 2019 Jan;22(1):82-96. [doi: 10.1089/cyber.2017.29099.gri] [Medline: 30183347]
60. Chirico A, Lucidi F, De Laurentiis M, Milanese C, Napoli A, Giordano A. Virtual reality in health system: beyond entertainment. a mini-review on the efficacy of vr during cancer treatment. J Cell Physiol 2016 Feb;231(2):275-287. [doi: 10.1002/jcp.25117] [Medline: 26238976]
61. Yang Q, Zhang L, Chang F, Yang H, Chen B, Liu Z. Virtual reality interventions for older adults with mild cognitive impairment: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2025 Jan 10;27:e59195. [doi: 10.2196/59195] [Medline: 39793970]
62. Li Y, Wilke C, Shiyanov I, Muschalla B. Impact of virtual reality-based group activities on activity level and well-being among older adults in nursing homes: longitudinal exploratory study. JMIR Serious Games 2024 Mar 29;12:e50796. [doi: 10.2196/50796] [Medline: 38551635]
63. Liao Y, Hsu MHK, Xie FL, Dai HX, Liu M. The role of virtual reality on physical and psychological health among older adults: a systematic review. Edelweiss Appl Sci Technol 2024 Oct 25;8(6):2762-2777. [doi: 10.55214/25768484.v8i6.2556]
64. Vasodi E, Saatchian V, Dehghan Ghahfarokhi A. Virtual reality-based exercise interventions on quality of life, some balance factors and depression in older adults: a systematic review and meta-analysis of randomized controlled trials. Geriatr Nurs 2023;53:227-239. [doi: 10.1016/j.gerinurse.2023.07.019] [Medline: 37598426]
65. Barbour B, Sefton L, Bruce RM, Valmaggia L, Runswick OR. Acute psychological and physiological benefits of exercising with virtual reality. In: Sun F, editor. PLoS ONE 2024;19(12):e0314331. [doi: 10.1371/journal.pone.0314331] [Medline: 39693283]
66. Dinet J, Nouchi R. Promoting physical activity for elderly people with immersive virtual reality (IVR). Presented at: 13th International Conference on Applied Human Factors and Ergonomics (AHFE 2022); Jun 24-28, 2022. [doi: 10.54941/ahfe1001682]

XSL•FO
RenderX

67. Malloy KM, Milling LS. The effectiveness of virtual reality distraction for pain reduction: a systematic review. Clin Psychol Rev 2010 Dec;30(8):1011-1018. [doi: 10.1016/j.cpr.2010.07.001] [Medline: 20691523]
68. Hoffman HG, Chambers GT, Meyer WJ 3rd, et al. Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. Ann Behav Med 2011 Apr;41(2):183-191. [doi: 10.1007/s12160-010-9248-7] [Medline: 21264690]
69. Li A, Montaño Z, Chen VJ, Gold JI. Virtual reality and pain management: current trends and future directions. Pain Manag 2011 Mar;1(2):147-157. [doi: 10.2217/pmt.10.15] [Medline: 21779307]
70. Weech S, Kenny S, Barnett-Cowan M. Presence and cybersickness in virtual reality are negatively related: a review. Front Psychol 2019;10:158. [doi: 10.3389/fpsyg.2019.00158] [Medline: 30778320]
71. Stanney K, Fidopiastis C, Foster L. Virtual reality is sexist: but it does not have to be. Front Robot AI 2020;7:4. [doi: 10.3389/frobt.2020.00004] [Medline: 33501173]
72. Integrated care for older people (ICOPE) : guidance for person-centred assessment and pathways in primary care, 2nd ed. : World Health Organization; 2024 URL: https://www.who.int/publications/i/item/9789240103726 [accessed 2024-12-12]

Abbreviations

GRADE: Grades of Recommendation Assessment Development and Evaluation
IVR: immersive virtual reality
NRS: numeric rating scale
PDQ-39: Parkinson Disease Questionnaire-39
PICOS: population, intervention, comparison, outcome, study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
RevMan: Review Manager
SF-12: 12-item Short Form Survey
SMD: standardized mean difference
VAS: visual analog scale
VR: virtual reality
WHO: World Health Organization
WHOQOL-BREF: World Health Organization Quality of Life Scale Brief Version

Edited by H Köttl; submitted 17.Jul.2025; peer-reviewed by JF Costa, ME Heidari, P Adedigba; revised version received 03.Nov.2025; accepted 03.Nov.2025; published 12.Jan.2026.

Please cite as:
Trillo-Charlín I, Bravo-Aparicio J, Avendaño-Coy J, Beltrán-Alacreu H
Impact of 4 Weeks or More Immersive Virtual Reality on Quality of Life and Physical Activity in Older Adults: Systematic Review and Meta-Analysis
JMIR Aging 2026;9:e80820
URL: https://aging.jmir.org/2026/1/e80820
doi:10.2196/80820

© Iria Trillo-Charlín, Javier Bravo-Aparicio, Juan Avendaño-Coy, Héctor Beltrán-Alacreu. Originally published in JMIR Aging (https://aging.jmir.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Lessons Learned About Digital Health Tool Acceptability Among Rural Older Adults: Systematic Review Guided by the Technology Acceptance Model

Zachary M Siegel[1], BA; Ellie Quinkert[1], BS; Jiya Pai[2]; Corinne H Miller[3], MLiS; Marquita W Lewis[4], PhD, MPH, MS

[1]University of Kentucky, Lexington, KY, United States

[2]Case Western Reserve University, Cleveland, United States

[3]Northwestern University, Chicago, IL, United States

[4]Department of Medical Social Sciences, Northwestern University, Chicago, IL, United States

Corresponding Author:
Marquita W Lewis, PhD, MPH, MS
Department of Medical Social Sciences
Northwestern University
750 N. Lake Shore Drive
Chicago, IL, 60611
United States
Phone: 1 312 503 1986
Email: marquita.lewis-thames@northwestern.edu

Abstract

Background: Digital health tools are increasingly vital in rural health care due to widespread hospital closures and the rapid adoption of telehealth during the COVID-19 pandemic. Rural older adults, a uniquely vulnerable population, face barriers to accessing these tools due to rurality and usability challenges. Although a growing body of literature examines the acceptability and usability of digital tools among rural older adults, no study has synthesized this research to establish best practices.

Objective: This study aims to review existing literature on digital health tools for rural older adults, highlighting key lessons learned about their acceptability and identifying strategies to improve usability for this population.

Methods: Following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, this study reviewed literature that investigated the role of digital health tools on the health outcomes of rural older adults (ie, at least 60 years old). The literature was retrieved from 5 electronic databases through June 2023. This study and all reviewed literature were conducted in the United States. Guided by a systematic process, 2 reviewers assessed relevant articles for eligibility, analyzed data, and extracted relevant content. The extracted findings were organized according to the evidence-based technology acceptance model, which assesses the acceptability of a technology by its usefulness, ease of use, and intention to use.

Results: The preliminary title review produced 7728 results, and 38 eligible manuscripts were included in the final review. Studies included both rural older adults and providers of rural older adults as participants. Digital health tools included, but were not limited to, videoconferencing, phone calls, telehealth monitoring, telemedicine appointments, and computer-based interventions. Findings on the usefulness of digital health tools by rural older adults were mixed. While digital health tools were useful for overcoming barriers to accessing care, these tools were less useful for rural older adults with limited digital literacy. Additionally, some studies described that the technology was easy but difficult to use when faced with environmental barriers, equipment issues, and discomfort with the technology. Rural older adults often reported an intention to use the technology after the study. Yet, on a few occasions, participants who preferred in-person care visits or did not have buy-in on the technology reported no intention to use the technology again.

Conclusions: Our review highlights that rural older adults and their providers generally view digital health tools as acceptable for delivering care and, in some cases, as a viable alternative to in-person clinic visits. While certain barriers impacted the acceptance of these tools among rural older adults, many of these challenges were not directly linked to their age or rural location; thus, they are potentially applicable to urban older adults.

Trial Registration: PROSPERO CRD42021287924; https://www.crd.york.ac.uk/PROSPERO/view/CRD42021287924

XSL•FO
RenderX

(JMIR Aging 2026;9:e70012) doi:10.2196/70012

KEYWORDS

health services; older adult; rural health; digital health; social determinants of health

Introduction

Digital health tools facilitate communication between patients and health care providers and offer access to resources. These tools encompass a range of technologies, including mobile health apps, electronic health records, wearable devices, and telehealth services. Social distancing mandates related to COVID-19 facilitated increased funding to support improved access to broadband internet and the rapid uptake of digital health tools [1]. To increase digital tool access and use by rural residents, in the spring of 2022, the US Department of Health and Human Services announced a US $16.3 million expansion to telehealth care in the Title X Family Planning Program [2]. Thus, rural health care professionals and systems were able to integrate digital tool uptake in their care rapidly [3]. Telehealth uptake in health clinics and hospitals increased by 154% in March 2020 compared with March 2019 [4]. For many rural patients, digital health tools are an essential component of their health care management and will likely remain important for timely and continuous rural care coordination [1].

Rural older adults represent a vulnerable population at the intersection of aging and rural residency, facing well-documented yet preventable challenges in accessing health care [5]. Rural residents are rapidly aging in place. For instance, 25% of older adults live in a rural or small town, and this is expected to rise to 33% by 2030 [6]. Additionally, for many rural older adults, care management is complex, confusing, and further challenged by coordination between distant health care facilities [7]. Since 2010, over 160 rural hospitals have permanently closed their doors, reducing access to inpatient care, which is critical for improving rural community health [8]. Therefore, aging rural populations will increasingly experience limited access to specialty care and poorer health outcomes [9].

Digital health tools can potentially overcome care coordination challenges for rural older adults. Once rural older adults engage with digital health tools, they often find their experience satisfactory and, at times, comparable to in-person visits [10]. Rural older adults evaluated web-based consultations conducted by service providers with high efficiency and satisfaction scores [11]. Once older adults understand the technology, they often find it an acceptable mode of care when punctuated by in-person visits.

Despite high levels of satisfaction with digital tools by rural older adults, compared with urban older adults, this vulnerable population has reduced telehealth use [12,13]. Also, although rural residents are willing to adopt digital health tools [14,15], studies show that rural older adults report slower telehealth uptake than younger rural adults [16,17]. This is partially due to barriers that make using digital health tools difficult for rural older adults. Some of these barriers include technical literacy, lack of technical support, cost, ownership of technology, and visual acuity [13]. In a systematic review including rural adults aged 55 years and older who have used telehealth, older adults reported a willingness to learn how to use various digital tools, but 30% felt too inexperienced with technology to use them [18]. Similarly, in a sample of Medicare enrollees, rural cancer survivors had a significantly lower predicted probability of internet use for patient-provider communication when compared with urban cancer survivors with Medicare (28% vs 46%) [19]. Importantly, not all rural older adults will find digital health tools to be a favorable health care management tool. Yet, funding to increase broadband access and the threat of widening rural medical deserts will facilitate continued telehealth uptake of digital tools by health care systems, thereby reinforcing the increased uptake of digital health tools by rural older adults. Increasing the acceptability of digital health tools and reducing barriers to their uptake for rural older adults are essential for providing health care to rural older adults.

Given the rapid acceleration of digital health tools by rural health care providers, rural older adults find digital health tools helpful. Still, rural older adults have reduced uptake of these technologies compared with both urban older adults and younger rural adults. With the increasing use of digital health tools, understanding their acceptability among rural older adults is crucial for ensuring this vulnerable population stays engaged in their care management and coordination as reliance on these tools continues to grow. Details about the rural older adults' digital health tool acceptability and usage can inform tool intervention design, implementation, and evaluation. Existing research summarizes the effectiveness of services such as telehealth among older adults, but strategies to improve rural older adults' usage of digital health tools are limited [18]. Therefore, this study will systematically review the existing literature in the United States on rural adults' acceptability of digital health tools and assess lessons learned on digital health tool usage among rural older adults.

Methods

Study Design

The study was analyzed and reported in accordance with the Cochrane systematic review guidelines and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines (Multimedia Appendix 1) [20,21].

Search Strategy and Data Sources/Protocol Registration

A trained librarian conducted searches in MEDLINE (Ovid), Cochrane Database of Systematic Reviews (Wiley), Embase (Elsevier), CINAHL (EBSCO), and PsycINFO (EBSCO) databases. This search included articles published in English through June 5, 2023. Keywords and subject headings related to the following topics were used to identify possible articles: rural residents, older adults, the use of

XSL•FO
RenderX

technology-enhanced tools to navigate telehealth, and acceptability. See the supplementary materials for the complete search strategy. Following, we searched the references of eligible articles for additional relevant articles. The protocol was registered post hoc in the International Prospective Register of Systematic Reviews (PROSPERO CRD42021287924).

All articles eligible for data extraction underwent title and abstract review, and full-text review by a pair of reviewers. Reviewers independently assessed the articles based on the eligibility criteria (see below) using standardized procedures in the systematic review software Covidence (Covidence systematic review software, Melbourne, Victoria, Australia). A pair of reviewers discussed and resolved conflicts in weekly meetings. Each article was assessed for quality by 2 reviewers. Extracted content and Quality Assessments were reported using Microsoft Excel and Covidence.

Eligibility Criteria

We included research articles of investigations conducted in the United States that assessed a digital health tool's ability to connect patients with providers, where at least 25% of the sample population identified as rural, and at least 25% of the sample identified as at least 60 years old. Articles were excluded if they were a review (eg, systematic or scoping), withdrawn, a conference proceeding, an abstract, or a dissertation. We also excluded articles published before 1999, as we deemed that the technology or lessons learned from the technology over 25 years ago were antiquated.

Data Extraction

Paired reviewers conducted consensus meetings to agree upon the rationale for data extraction content and synthesize the results. Table 1 displays the extraction content. In short, paired reviewers reported each article's title, first author, and year of publication. The outcome variables collected from each study included the participants' age (average or mean), study design, and type of digital health tool technology (eg, videoconference or wearables). Following the technology acceptance model (TAM), we extracted data related to the core TAM domains: perceived usefulness, perceived ease of use, and intention to use [22]. The perceived usefulness domain describes how much technology improves a patient's performance. Perceived ease of use is the effort required to use the technology. Last, intention to use refers to a patient's willingness to use the technology. Given that we aim to synthesize acceptability and lessons learned, our analysis did not assess the effect of the outcomes. This systematic review did not need an exploration of heterogeneity or a sensitivity analysis.

XSL•FO
RenderX

Table 1. Summary of findings.

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
Anderson et al [23]	Mean age 61 years	Community-based outpatient clinics in rural Southeast Texas	Qualitative	Videoconferencing	• It was accessible. • Videoconferencing was convenient and reduced transportation issues.	• Participants reported that there was too much information covered in the 2 self-management classes and that 2.5 hours was too long for a single videoconferencing session. • Clinicians reported that they had no time to assist when technology problems occurred. • Clinicians reported that there was only limited clinic space to hold videoconferencing sessions.	• Not applicable.
Barton et al [24]	28.5% of the sample is 60 years or older	67.3% of the sample is from rural Colorado	Cross-sectional	Phone call, videoconference	• Providers reported that ease of access to patient records, scheduling follow-up visits, and timely follow-ups were all better accomplished during digital visits than office visits (2-4.5× higher, all P<.001).	• Not applicable.	• Patients who did not engage in telehealth listed "preference to seeing the provider in person" and "telehealth not being an option" as reasons.
Bernacchi et al [25]	Mean age 60.2 years	Rural dwelling residents from Southeast US	Mixed methods	Videoconferencing	• Access to an oncology nurse during COVID-19 increased rural cancer patients' access to care, information, and education. • Participants gave unfavorable scores to questions that asked about the use of the technology. For example, the lowest scoring items were "my health is better than it was before I used the technology" (—X = 3, SD = 0.89).	• Participants struggled with connecting to appointments due to a lack of equipment or discomfort with digital technology. • Three participants with insufficient broadband used nearby telehealth satellite sites at local clinics or hospitals. • Participants often relied on family members for connection due to limited internet experience or poor broadband signals.	• Participants were committed to overcoming barriers in order to speak with their nurse.

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
Bonsignore et al [26]	Mean age 72 years	Rural counties in Western North Carolina	Mixed methods	Telemonitoring	• TapCloud increased patient-provider relationship, accessibility to clinicians, increased response times, improved efficiency, and made medications more accessible. • Accessibility to a provider increased comfort, along with communication and preemptive management of problems.	• TapCloud's facilitation of direct, efficient contact with patients made it easy to use. • Patients and caregivers were particularly enthusiastic about how easy medication refills are with TapCloud. • Patients and caregivers find the TapCloud application intuitive, easy to use, and not time-intensive.	• Once coached on using the application, patients readily adopted the technology and often felt a sense of accomplishment in doing so.
Browning et al [27]	Mean age 81 years	Residents of rural Southwest Virginia	Retrospective quality improvement case series	Phone-based telehealth monitoring	• Participants felt nurses had better accessibility to vitals (88.9%). • 90% felt a better connection with their doctor.	• 77.8% found telehealth easy to use.	• Not applicable.
Collie et al [28]	Mean age 60.7 (SD 9.24) years	Resident of Intermountain region of Northeastern California	Pretest-posttest	Videoconferencing	• Not applicable.	• Participants adapted quickly to videoconferencing and reported no communication difficulties. • The facilitator faced challenges due to time lags and poor lighting.	• All said they would recommend it to other women with breast cancer.
Cummings et al [29]	Mean age 52.8 (SD 16.2) years	Rural counties in eastern North Carolina	Descriptive	Nonmydriatic retinal imaging telemedicine system	• 85% of images were "good or fair" when examined by a retinal specialist. • Retinal specialists reported "very certain or certain" in 84% of diagnoses.	• 96.3% of the participants were "very comfortable" or "comfortable" with the portable camera.	• Not applicable.
DeHart et al [30]	41.2% of the sample is 55 years or older	Rural resident from South-Eastern State	Mixed methods	Web and mobile-based telehealth with remote monitoring	• The application effects were seen through the increased coordination across agencies, technology fit with intraoffice demands, growing practice through engagement, flexible delivery, reduced provider travel, geographic access, reduced patient transport, saving time/visits for patients, reduced costs, and fewer challenges to sustainability.		• The most common patient-level challenge reported by providers was a lack of buy-in to use the service consistently.

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
						• Telehealth had organizational technology and space challenges. • Patient skills and comfort were a challenge. • Provider knowledge and skills were a challenge. • Rural patients struggled with technology access, often due to a lack of broadband technology in homes, but also due to technology skills for aging populations and those with less education.	
Demaerschalk et al [31]	Mean age 66.3 (SD 13.5) years	Patients at rural medical centers, more than 185 miles from a primary stroke center	Randomized control trial	Telestroke	• Results indicate effectiveness in the organization and structure of stroke telemedicine networks for extending stroke care in rural communities.	• Technical problems were noted in 20 of 27 (74%) of telemedicine consultations and in 0 of 27 telephone consultations. However, none of the issues kept the patient from being cared for.	• Not applicable.
Depatie and Bigbee [26]	Senior living facility serving adults aged 60 years or older	Rural northern California senior centers	Descriptive mixed methods	Mobile health technology		• Participants' comfort level with using email or the Internet to communicate with a health care provider was evenly split.	• Responses to interest in incorporating technology into daily life for health tracking and communication with a health care professional indicated that 33% had no interest, 23% were somewhat interested, and 13% indicated that they already used this technology.

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
					• 67% felt it was moderately important for the technology to monitor their health, but do not feel comfortable sharing with their provider. • 53% felt it was somewhat or extremely important to monitor their health and were comfortable sharing information with their provider. • 46% felt it was somewhat or extremely important to monitor their health as well as share information over the internet with their health care provider. • 44% indicated it was moderately, somewhat, or extremely important to use mobile health technology in combination with in-home nurse visits. • 40% felt it was somewhat or extremely important to connect patient education and support groups online. • 77% felt health technology had a clear benefit on their health.		
DeVido et al [32]	Mean age 54 (SD 19.4) years	Providers at a rural population-serving hospital	Case report	Telepsychiatry	• Telepsychiatry is capable of responding to many consultation questions in a hospital setting. • Cognitive test items were collected reliably using a camera function that interpreted physical images. • Telemedicine gave valuable clinical data regarding psychotic symptoms. • Telepsychiatry was applicable in accessing a diverse range of patients. • Communication between the resource nurse and referring physician was aided through telemedicine.	• Technological challenges included getting the cameras and software to operate reliably. • An ongoing challenge was the integrity of the image and audio. • Audio feed was compromised by Wi-Fi connectivity problems.	• Not applicable.

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
Donahue et al [33]	Mean age 57.9 (SD 12.4) years	Rural North Carolina county	Cohort	Phone-based digital health care	• 84% of participants set a goal and reached at least one goal due to digital health technology.	• Not applicable.	• Providers expressed interest in continuing the phone coaching program with their participants, if resources were available. • 96% would recommend phone coaching to others. • 58% of participants remained engaged with phone coaching over the 12-month period (missed fewer than 3 consecutive monthly calls), 17% were less engaged (received at least one call), and 25% were not engaged.
Finley et al [34]	82% of the sample is 65 years or older	Three rural outpatient telehealth clinics, or one urban outpatient clinic in Arizona	Mixed methods	Telecardiology	• Patients perceived service as competent and began to see benefits in accessibility, reduced transportation, and general timesaving.	• Not applicable.	• When compared with the telehealth group, in-person participants more frequently expressed concerns about technology, about care not being face-to-face, and made more negative statements about quality. • Rural participants held the most positive attitudes toward telehealth, while suburban participants held the least positive attitudes toward telehealth.
Geller et al [35]	54.6% of the sample is 61 years or older	Patients and their providers from rural practices in Vermont	Quasi-experimental	Computer-based interactive interventions			• Not applicable.

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
					• Preliminary evidence that the technology could educate patients about screening. • Technology promoted patient-provider discussion, provider recommendations, and positive patient intentions to get a screening.	• Digitally everyone found the program easy to use. • Older participants and less educated participants found the sound (reading the information and questions on each screen) helpful compared with the younger and more educated participants.	
Gutierrez et al [36]	Mean age 65.2 years	Patients from a VA[a] hospital in rural Wisconsin	Mixed methods	Telehospitalist	• Communication benefits • Telehospitalists reported confidence that the diagnosis accuracy and quality were that of in-person. • Patient satisfaction showed improvement in care coordination (18%; P=.02).	• Connectivity problems were prevalent, although most providers were able to resort to a backup plan. • Internet connectivity was inconsistent, leading to disruption in video communications. • It was easy to contact bedside providers/telehospitalist: 100% telehospitalist, 42.9% physicians, 10% nurses, and 22.2% other staff.	• Not applicable.
Hatch et al [37]	47.4% of the sample is 65 years or older	33.1% of the sample resided in a rural area	Descriptive cross-sectional cohort	Telehealth	• Not applicable.	• Older and low-education patients are less familiar with the required technology and experience some technological limitations.	• Not applicable.
Hicks et al [38]	Mean age 68.9 years	Residents in rural midwestern state	Experimental	Telemonitoring	• 78.3% indicated that the use of telehealth technology improved their care. • 78.3% indicated it was very convenient. • 95.7% of the respondents indicated that the telehealth technology affected their relationship with their nurses positively.	• 95.7% indicated it was very easy to communicate with the agency personnel using the equipment. • 95.7% of the respondents indicated that the telehealth technology was very easy to use. • 87.0% indicated that the telehealth technology worked very well.	• 91.3% of the respondents indicated definitely and 8.7% indicated maybe when asked if they would use the digital health technology again.
Holloway et al [39]	Mean age 61.3 (SD 11.6) years	Residents in rural Montana	Pretest-posttest	PRISM[b] digital health care videoconferencing			• Not applicable.

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
					• Patients learned to work with the technology, enabling them to interact effectively with patients at other sites. • Participants felt that the ability to participate in a team approach to diabetes management without traveling was beneficial. • Compared with baseline, patients reported improvements in diabetes care of 30-200% 1 year after the intervention.	• Scheduling was a difficult and time-consuming task. • Staff members easily learned to use telehealth technology. • In most cases, patients were able to learn manual skills using telehealth technology. • 99% felt the technology picture and sound were clear. • 100% of patients felt comfortable learning health information using this technology and said they understood the information as if it were imparted in person.	
Khairat et al [40]	Mean age 77.8 years	66.7% of the sample resided in rural North Carolina	Cross-sectional	Videoconferencing follow-up care	• Geriatric patients did not have trouble using digital health tools. • Telemedicine platforms improve primary care by allowing providers to follow-up with their geriatric patients at a time and place that is most convenient for both groups.	• Not applicable.	• Not applicable.
Kulcsar et al [41]	Mean age 60.3 (SD 16.1) years	The majority of patients resided in rural New Hampshire or rural Vermont	Cross-sectional quality improvement	Telerheumatology	• Providers reported that 19% of the patient visits seen via Telerheumatology were inappropriate for the telemedicine visit type due to poor understanding of symptoms, symptom complexity, and the limited ability to perform a physical examination. • 81% of patients rated that they were comfortable with the provider's ability to examine them, thought the provider spent an adequate amount of time with them, and made an accurate diagnosis of their condition.	• 94% of patients felt that each individual member of the staff made check-in easy, was friendly, and competent with the videoconferencing equipment. • An area of greatest dissatisfaction for patients stemmed from problems with scheduling appointments (usually follow-up).	• About half (53%) of the patients either agreed or strongly agreed that they would like to be seen via telerheumatology again if given the option.
			Qualitative				

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
Liu et al [42]	Mean age 67 years	Patients from a clinic in rural Wisconsin		Teleophthalmology	• Teleophthalmology was often preferred by patients because it was more comfortable than traditional eye examinations, since pharmacological pupil dilation was usually not needed. • Teleophthalmology is effective for increasing diabetic eye screening rates in rural populations.	• Tele-ophthalmology offered convenience through same-day scheduling, proximity to the patient's primary care provider, and short wait times. • Patients appreciated the quick, easy, and painless nature of tele-ophthalmology compared with traditional, in-person eye examinations.	• Participants reported being unaware of tele-ophthalmology prior to the study.
Locke et al [43]	Mean age 69.2 years	Residents of rural zip codes defined by the United States Census Bureau	Retrospective	Home computer video health technology	• 96% of participants preferred video telehealth rather than traditional training visits at the medical center. • 76% of participants would not have gotten additional inhaler training if not for the telehealth. • The telehealth inhaler training delivered via internet video telehealth demonstrated an improvement in technique overall. • The main benefits, as listed by participants of the program, were convenience, time-saving, and decreased travel expenses. • Inhaler training delivered via video telehealth by a pharmacist was well received. • The CHAT inhaler training program provided an alternative to in-person visits for rural patients with transportation burdens.	• Among 93 home telehealth program enrollees, 19 (20%) faced technical issues with the computer or video software that prevented participation. • A quarter of participants reported frequent technical problems, consistent with pharmacists noting issues in 149 (63%) of scheduled visits, and 19 visits (13%) were postponed or partially completed due to unresolved technical issues. • Common issues included patient errors or confusion (41% of visits) with the video telehealth program, particularly with logging into the Jabber program and basic computer skills, along with 11% experiencing computer/software issues and 25% having audio/video troubles. • Despite these challenges, over 90% of participants found the equipment easy to set up and appreciated the benefits of convenience, decreased travel time and expenses, and increased privacy.	• A majority (96%) preferred home video telehealth for inhaler training compared with going to the medical center for in-person training.
McIlhenny et al [44]		Patients of rural medical clinics	Quasi-experimental	Computer-based telemedicine	• Not applicable.		

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
	Mean age 63.8 (SD 12.46) years					• Nearly all participants found the information easy to access.	• Participants responded positively to the individualized education provided by the nurse educator and did not prefer receiving information from the internet.
Owolo et al [45]	Mean age 59.9 (SD 13.5) years	28.9% of the sample was considered rural	Retrospective cohort	Web-based telemedicine	• Patients who used the patient portal had 5.21 times higher odds of completing a video visit compared with patients who did not use the patient portal (95% CI: 1.28-21.23; P=.022). • The majority of telehealth visits were conducted over the phone; however, there was an increase in video visits in the post–initial surge period.	• 81.9% of physicians agreed that telemedicine was easy to use with a preference for imaging review, initial appointments, and postoperative care. • While telephone visits were still used at higher rates than video visits, the increased use of video visits potentially reflects a better organized infrastructure for performing this type of visit.	• Not applicable.
Robinson et al [46]	Mean age 71 (SD 6.8) years	Residents from rural New York	Cross-sectional	Telemedicine	• 72% found the in-home nurse visit "very helpful," in contrast to 55% for telephone tutoring and 46% for the 39-page user's manual. • 37.8% of the participants wanted additional telephone training to access the web.	• Twenty-two subjects (63%) rated the amount of training as "3" (about right) compared with their initial expectation. • Four responders (11%) rated the training lower than "3" (less than expected), while 9 participants (26%) rated it higher than "3" (more than expected).	• Not applicable.
Rodriguez et al [47]	Mean age 63 (SD 12) years	Patients and providers from rural medical centers and remote community-based outreach clinics in Pennsylvania	Mixed methods	Electronic consultations	• Improved communication, as it enabled effective information transfer and patient-centered care. • For PCPsc, time efficiency was the main reason for telehealth satisfaction. • E-consults improved access to specialty care, saving travel time, and enabling confident care.	• Not applicable.	• All patients indicated that they intended to use e-consults in the future. • The intent to use e-consults in the future focused primarily on quality of care and timeliness of care.
Schlittenhardt et al [48]	Mean age 54 years	78% of the sample from rural areas	Mixed methods	Tele-Continence care		• Tele-Continence Care implementation was uncomplicated.	

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
					• The option of telehealth as a follow-up appointment reduced the overall failure rate from 45% to 14.3%. • Patients appreciated the convenience and reduction in transportation costs.		• When asked if face-to-face visits would be preferred over telehealth visits, patients described a neutral opinion.
Schooley et al [49]	Mean age 63.5 years	Residents of rural Vermont	Mixed methods	Information technology mediums	• Telehealth programs, such as telephone triage, are accepted as an option to receive remote care and avoid travel. • The community care home telehealth enabled physiological data to be monitored remotely through landline phones amid staffing problems. • Email, computers, and assistive devices were imperative to aiding veterans with disabling conditions to communicate effectively while remote.	• Assessing mental health symptoms was challenging through telephone and email, as these methods do not allow for observing nonverbal cues, unlike advanced videoconferencing.	• 56.5% of respondents reported some level of likelihood of consulting with a doctor over the phone, compared with 13.2% who reported some level of likelihood of using the Internet to consult with a doctor. • 35% had an interest in communicating with their doctors and nurses via the Internet about their health care. • Younger veterans (age 41-55 years) were more likely than older veterans to report interest in the program.
Silvestrini et al [50]	Mean age 60 years	Patients referred from clinics in rural Washington, Oregon, or Alaska	Qualitative	TelePain	• TelePain reduced transportation barriers and travel costs. • Patients enjoy the convenience of its general accommodations.		

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
						• Patients generally did not have significant issues using telehealth technology, with many veterans rating TelePain video and audio quality as "good" or "fine." • TelePain staff effectively helped veterans navigate any technological problems, such as resolving audio issues by having the provider call the patient. • Some patients suggested improvements for telehealth, such as larger TVs, better monitors, or improved camera systems to enhance the video quality and overall experience.	• Many of these patients mentioned that they would like to continue to use TelePain because of its convenience, rather than having to travel far distances to receive pain care. • When asked if they would use TelePain again, one patient said, "Oh, of course. It's much easier than going clear to [the VA medical center]." Another patient replied, "Yeah, probably. It's easier for me than driving to [the VA] and back."
Strowd et al [51]	Mean age 44.5 (SD 24.1) years	26% of the sample resided in a rural zip code	Retrospective cohort	Video and phone-based digital technology	• 73% of respondents reported that clinical needs were met with the telehealth visit. • Patients completing the video visits were more likely to have needs met than those who did telephone-only visits (77% vs 71%, P=.34). This was further seen in both urban and rural communities (71% vs 80%, P=.27).	• The most common patient-reported barrier to scheduling a video visit was patient technology-related (44% of patients), which included lack or limited access to a smartphone or home computer (n=59), no camera for video (n=51), no internet availability (n=27), or other (n=12).	• 45% of the respondents would definitely consider a future telehealth visit, 28% might consider, 24% would only consider if required. • Patients who completed video visits were more likely to definitely consider a future telehealth visit compared with patients who completed a telephone-only visit (58% vs 38%, P=.02). • Patients from rural communities were more likely to definitely consider a future telehealth visit compared with those from urban communities (55% vs 42%, P=.05).
							• Not applicable.

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
Svistova et al [17]	A large portion of the sample is aged 65 years and older	Rural residents in Pennsylvania	Qualitative focus group	Zoom teleconferencing	• Technology has allowed service providers to adapt quickly to COVID-19. • The effects of telehealth eliminated the transportation barrier and reduced cancellations and missed appointments.	• Older adults faced significant barriers using telehealth due to unfamiliarity with technology, lack of recent devices, low comfort levels, and distrust leading providers to have fewer positive insights about serving them during COVID-19. • An insurance provider noted that older clients often lack the proper technology or skills for telehealth. • Using technology for health care communication is often uncomfortable for older adults.	
Switzer et al [52]	Mean age 70 years	Patients from one of 12 hospitals, 10 of which were in rural Georgia	Cohort	Telestroke	• Telestroke helped enroll patients into the acute stroke treatment program. • The initiation and treatment period for Telestroke patients was quicker compared with in-person patients.	• Not applicable.	• Not applicable.
Virmani et al [53]	Mean age 65.8 (SD 9.2) years	Residents in rural Arkansas	Cross-sectional	Web-based televideo	• 82% of participants reported televideo had a positive effect on their in-home visits. • 70% of participants reported that they liked that there were no travel arrangements. • 84% of participants reported that they liked the ability to stay in the comfort of their home. • In rural areas, the quality of the audio-video connectivity was enough to implement the routine clinical and research assessments.	• 28% preferred for in-person visits. • Audio-video quality was rated great for 60% of visits, with video slightly slow for 30%. • Only 14% of participants needed more than 5 minutes for setup; 42% needed additional time for understanding, locating, or completing research assessments.	• Most patients reported being more likely to participate in telemedicine research in the future. • 28% preferred for in-person visits.
Waymouth et al [54]	Providers of older adults	Providers in a rural area	Qualitative	Assistive and remote monitoring technology		• Not applicable.	• Not applicable.

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
					• Telehealth helped facilitate care for individuals with dementia through the use of assistive technology and remote monitoring technology such as fall sensors, webcams, alarm systems, GPS tracking, and smart home hubs.		
Weiner et al [55]	Mean age 69.7 (SD 12.8) years	Residents of Choctaw Nation of Oklahoma	Clinical trial	Videoconferencing	• Virtual care was used to overcome transportation burdens.	• Less than 3% no-show rate indicates the technology was easy to use.	• There were no refusals of videoconferencing for initial visits and 2 refusals of continued videoconferencing follow-up. • The no-show rate for all videoconferencing sessions in the past year was 3%.
West and Milio [56]	Mean age 68.3 years	Patients from a rural homecare organization	Mixed methods case study	Rural homecare organization telemedicine	• There were frequent issues with the telehealth device that caused disruptions. • Restrictions were seen in outdated telecommunications equipment, which impaired the transmission of audio and visual information. • An increase in productivity and lower costs needs to be implemented to meet the demand for these services.	• Informants reported that the complexity of the setup and the number of components confused and intimidated the nurses and patients. • Faulty equipment also caused frequent disruptions during telemedicine visits.	• Younger patients are more likely to receive telemedicine visits than those receiving traditional home care.
West et al [57]	55 years or older	Residents of rural upstate New York	Cohort clinical trial	Web-based telemedicine with videoconferencing	• The IDEATel project demonstrated the home televisits through videoconferencing were possible for rural underserved elderly adults with the right home education and behavior goal setting.	• Many participants initially had difficulty using the unit and required additional instruction.	• Not applicable.
Zulman et al [58]	Mean age 56 (SD 17) years	53% of the sample as rural	Mixed methods	Tablet-based health technology	• The VA's use of video telehealth tablets reached rural older adults effectively.	• Not applicable.	

XSL•FO
RenderX

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
							• 81% of tablet recipients used their tablets during the evaluation period. • Tablet recipients were more likely to use their tablets if they were 45-64 years or 65 years old (compared with < 45 years). • Patients who did not use the telehealth tablet were largely younger with more chronic conditions and a lack of social support.

[a]VA: Veterans Health Administration.

[b]PRISM: Promoting Realistic Individual Self Management.

[c]PCP: primary care provider.

Quality Appraisal

The quality assessments were achieved through critical appraisal tools used in Joanna Briggs Institute Systematic Reviews [59] and the McGill Mixed Methods Appraisal Tool Version 2018 [60]. Two reviewers assessed the quality of each included article represented in the extracted data. Weekly meetings were held to discuss conflicts. Studies with a quality score of less than 50% were not included in the analysis, indicating evidence of reporting bias.

Results

Overview of Reviewed Studies

Figure 1 provides a summary of the selection process and illustrates our systematic review of the literature. The preliminary title review produced 7728 results, out of which 7616 underwent the abstract review, and 511 completed the full-text review. Following the assessment, 16 studies were excluded because they scored less than 50% of the total possible score for each respective quality test. Our final review included 39 studies.

XSL•FO
RenderX

Figure 1. PRISMA flow diagram.

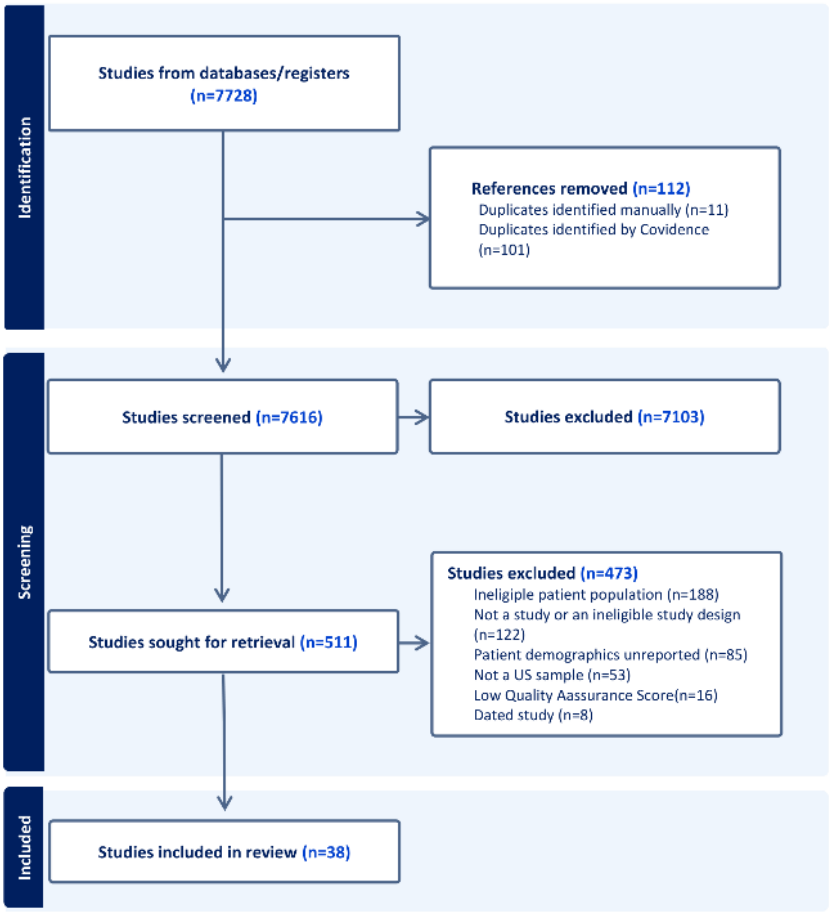


Table 1 details our data extraction, elucidating article information according to the first author and publication date, study population age, study design, technology used, and TAM domains. Since 25% of the population had to be at least 60 years old for eligibility, it is important to highlight that studies reported the mean age, included samples exclusively of older adults (eg, Medicare enrollees), and provided the percentage of participants within specific age ranges. The study design of eligible studies was cross-sectional (n=8), qualitative (n=7), and mixed methods (n=11). We will detail lessons learned from the eligible studies according to the TAM domains below.

Technology Usefulness

Digital health tools were useful for connecting patients to providers [25,26,35,38,54,61], assisting participants with improving health outcomes and care management [25,31,38-43,47-49,51-54,57,61], and reducing transportation burdens [17,23,30,34,43,47-50,53,55]. Uniquely, Anderson et al [23] reported on the role of telehealth videoconferencing in strengthening patients' social and peer connections. Providers also considered that digital health tools were useful for care management [24,29,32,36,40,49] and communicating with other providers [30,32,47]. Although the respective technology was considered useful, multiple studies reported that rural older adults needed additional supports to facilitate the usefulness of the technology, including behavioral goal setting [57] and assistive devices [49].

Despite the reported usefulness of digital health tools, some studies also reported that they were less than useful. Bernacchi et al [25] revealed that digital health technology was less useful for participants with limited experience with digital technology. Similarly, according to Kulcsar et al [41], nearly one in 5 patients scheduled for a telerheumatology visit was deemed unsuitable due to poor symptom understanding, symptom complexity, and the limitations of conducting a physical examination remotely.

Ease of Use

Articles that addressed the digital health tools' ease of use described the tools as uncomplicated [48], requiring an adequate amount of training to use [62], easy to access health information [39,44], easier than an in-person examination [42], easy to communicate with a health care professional [36,38,61], and intuitive [26]. Hybrid models that provided supplemental real-time instruction with the technology improved the ease of use [41,50]. Importantly, some digital health tools were easier to use than others. For example, Schooley et al [49] assessed that mental health evaluations via email and telephone were challenging because of barriers to observing nonverbal cues, yet were easier to accomplish via a videoconference. Additionally, Geller et al [35] reported that audio features improved the ease of using a computer-based intervention for older adults. Providers also reported that digital health tools improved the ease of managing appointments [24,45], accessing and reviewing patient records [24,45], and contacting patients [26]. Weiner et al [55] reported that a low "no-show" rate (3%)

XSL•FO
RenderX

evidenced the ease of using videoconferencing visits. Locke et al [43] reported that 90% of the providers reported that the technology was easy to use, despite the fact that more than half of the providers reported issues with scheduling, and 41% reported that patients were confused about using the technology.

Commonly, studies have reported on the challenges of using technology. Digital health tools were more complicated to use when faced with environmental barriers (eg, poor lighting) [23,28], technical or equipment issues [17,25,30-32,36,43,51,53,56], discomfort with the technology [17,25,30,56], reduced access to adequate internet [30,32,36,51], and distrust in the technology [17]. DeHart et al [30] and Hatch et al [37] observed a positive relationship between education levels and ease of use, which was particularly challenging for older adults with low levels of education. Although many studies reported the challenges of using digital tool technology, West et al [57] reported that supplemental instructions were critical for overcoming usage barriers.

Intention to Use the Technology

Most commonly, participants reported an intent to use the digital health tool technology through sustained use poststudy [34,38,41,47,51,53], future referrals or recommendations [28], high levels of digital tool uptake [43,50,55,58], and evidence of improved care management [25,43,47,50]. For example, Locke et al [43] reported that 96% of the participants preferred home video telehealth inhaler training rather than going to the clinic for in-person training. Bernacchi et al [25] detailed another example of intention to use videoconferencing technology through the patient's commitment to contact their health care provider despite challenges with equipment and broadband. Strowd et al [51] and Finley et al [34] reported that rural residents were more likely to consider telehealth in the future compared with urban residents. A study indicated that participants were more likely to continue with specific delivery modes of digital health tools, such as telephones, rather than web-based options [49].

When participants did not intend to use the digital health tool technology, it was often due to a preference for in-person visits [24,34,44,45,53,56] or a lack of buy-in about the technology [30,42]. Additionally, 2 studies identified age associations with intentions to use digital health tools. Older adults had reduced intentions of continuing their care digitally [49,51].

Discussion

Principal Findings

This review aimed to assess lessons learned about the acceptability of digital health tools among rural older adults. Following a systematic review approach, we organized our findings according to the TAM, focusing on the usefulness, ease of use, and intention to use digital health tools of rural older adults. The domains of the TAM aim to detail predictors of potential acceptance or rejection of the technology. Our findings revealed that digital health tools were, in most cases, useful for care management, reducing transportation burdens, and improving patient-provider communication. Two articles reported that digital health tools were less useful when the technology was misaligned with the participant's digital skills level or when the technology was unsuitable for the given care visit types. Several articles highlighted the ease of use of digital health tools for rural older adults, describing them as uncomplicated, intuitive, and effective for connecting with health care providers. Most articles that discussed the ease of using digital health tools focused on the ease of specific features of the tool (eg, audio capabilities) or the type of technology (eg, telephone vs digital). However, digital health tools were more difficult to use due to technical or equipment issues, discomfort with the technology, and limited access to broadband internet. Last, the findings on rural older adults' intention to use digital health tools were robust, as evidenced by participants' preference to continue using the technology after the study concluded and improved health outcomes. Comparatively, participants did not intend to use the digital health tools when they preferred an in-person visit or when they were not sold on the benefits of the digital health tool. Together, the TAM domains reveal that rural older adults and their providers largely consider digital health tools as acceptable modes of receiving care and, at times, a suitable alternative to in-person clinic visits. Despite barriers that reduced rural older adults' acceptance of digital health tools, many of these barriers were not associated with their age or rural residence.

Technology Usefulness: Lessons Learned

Digital health tools are useful for accessing health care and care management, but their effectiveness in improving health outcomes in older adults is mixed. Our review highlights that digital health tools were useful for mitigating burdens related to accessing health care and care management for both providers and patients. Articles reported that useful care management needs included scheduling, accessing health records, and patient-provider communication. The positive findings on remote care management and usefulness are specific to this review, and it is important to note that the effectiveness of remote care management and monitoring on health outcomes is mixed. In a review of remote care management of depression and anxiety in older adults, the findings on psychiatric outcomes were mixed, and no studies demonstrated a statistically significant effect of remote care management on health care use or cost [63]. Another review of mobile integrated health interventions for older adults revealed that these interventions reduced emergency department call volume and transports [64]. Thus signaling that digital health tools were useful for care management during emergency health events.

To build on this body of evidence, our review uniquely emphasizes the usefulness of digital/remote care management for rural older adults—a population that has complex care needs but often resides in a medically underserved area with reduced access to broadband internet and technology literacy programming. However, future systematic and meta-reviews are needed to assess the effectiveness of digital health tools with care management features on health outcomes and costs for rural older adults.

Ease of Use: Lessons Learned

Easy-to-use technology is associated with improved health outcomes; however, the design of technology may not be

XSL•FO
RenderX

sufficient, and external support (eg, timely technical assistance) may be necessary. According to the TAM, digital tools that are perceived to be easy to use are more likely to be accepted by the intended audience. Based on this review, rural older adults and their providers frequently highlighted user-friendly features of the tools that improved ease of use. However, there was strong evidence that external factors—such as technical issues, equipment limitations, and discomfort with the technology—hindered usability. There is a positive relationship between the health of older adults, their social connectedness, access to high-quality health care resources, and the perceived ease of use of digital health tools [46,65]. Notably, this evidence signals that technology design alone does not improve the ease of using digital health tools. Specifically for older adults, in-person synchronous technical assistance, access to remote technical assistance, and early interventions from hospital administrators are reported facilitators for increasing the ease of use of digital health tools [15,25,29]. Overall, this evidence underscores the need for additional resources and external support to enhance the perceived ease of using digital health tools for rural older adults.

Intention to Use the Technology: Lessons Learned

Despite design flaws or technical difficulties, rural older adults generally intended to continue using technology beyond the observed period. Factors such as the preference for in-person care and a lack of buy-in about the technology influenced the participants' intent to use digital health tools. Yet, our findings revealed that participants described an intention to use digital health tools despite also reporting that the technology was not always useful or easy to use [25,34,41,43]. In a similar study on patient portal use by older adults, challenges were noted with log-ins and the user interface design, such as color and font [62]. Despite these issues, older adults expressed an intention to continue using the portals due to their other beneficial features. In summary, in light of the barriers to ease of use and usefulness, rural older adults often overcame them to continue using digital health tools [25].

Additional Considerations for Digital Tool Acceptability

In synthesizing lessons learned, our review identified several phenomena, not salient enough to categorize as a lesson, but worthy of continued discussion. Namely, our review reveals both differences and commonalities in user behavior and preferences between older adults in rural and urban areas. As reduced access to broadband internet is a common barrier for rural residents, this was not the most reported impediment to digital tool acceptability in this review, as expected. According to the Federal Communications Commission's data reported in 2021, 23%-50% of rural residents had poor access to broadband internet [66]. The rather limited mention of challenges associated with rural internet connectivity, in this review, is inconsistent with the existing literature, which indicates that poor internet connection is a key barrier to digital tool use acceptability for rural residents [15]. Also notable is that multiple studies reported a measure of digital tool acceptance among rural residents compared with urban residents. For example, Finley et al [34] reported urban-rural differences in intention to use, indicating that rural patients, compared with urban and suburban patients, had more favorable attitudes toward telecardiology. This higher acceptance of digital health tools by rural patients likely punctuates the growing reliance on and acceptance of digital health tools in the wake of dwindling local health care resources and rising health care costs. While these findings suggest that rural and urban residents share common barriers to technology acceptability, our conclusions do not suggest that "one-size-fits-all" interventions should be considered. Rather, additional qualitative examinations on the rural-urban differences in attitudes toward digital health tools and the acceptability of these tools are warranted.

This review synthesized a diverse representation of technology modes (eg, videoconference or phone call), highlighting the robust intervention designs implemented in rural settings. These findings strengthen the evidence base on digital tool acceptability among rural patients, with all modes being reported as acceptable. Yet, important distinctions emerged across the reported technology modes, making salient conclusions about the most acceptable modes speculative. For example, both Anderson et al [23] and Svistova et al [17] used videoconferencing; however, Anderson et al [23] employed a Veterans Affairs-supported platform, and Svistova et al [17] used Zoom. The journal articles provided limited details regarding platform-distinctive features, though such distinctions could plausibly impact acceptability and usability very differently. In the current review, technology modes were extracted as they were identified in the original article to ensure transparency. A more detailed analysis of the technology's distinctive features and their impact on acceptability warrants further investigation.

The data collected for this study includes both pre– and post–COVID-19 pandemic publications, which provide key insights into how rural older adults' acceptance of digital health tools evolved, resulting from the rapid uptake of telehealth due to COVID-19 precautions. Many of the studies that were published prior to the COVID-19 pandemic often emphasize the same sentiment of the digital divide between younger and older generations. The existing literature evidences that the pandemic exacerbated this divide, as prepandemic older adults were less likely to benefit from technological innovations [67]. Specifically, pre- and early-pandemic trends indicated that age and rural zip codes were inversely related to continuous digital tool use [68]. Multiple studies from our review that were published prepandemic identified that younger participants were more likely to engage in digital tool technologies [49,56]. Yet, in a peripandemic investigation, Bernacchi et al [25] describe that videoconferencing with a nurse increased access to care for older rural cancer patients. Similarly, postpandemic, individuals older than 65 years used telemedicine over 3 times more when compared with prepandemic, further emphasizing this technological shift [19].

The findings of this review have practice and research implications. To improve practice, our findings suggest that health care providers adopt hybrid care models—combining both digital and in-person visits. It is essential for providers to emphasize the benefits of digital health tools, such as reducing travel burdens and offering greater convenience

XSL•FO
RenderX

[30,39,43,47,49,50,53]. Additionally, future research should include interventions aimed at increasing technology literacy among rural older adults and provide additional synchronous and asynchronous supports to help rural older adults better understand the technology [17,25,30-32,36,37,51,56,57]. For caregivers who support rural older adults, digital health tools have the potential to reduce caregiving burdens. Tools such as remote monitoring enable doctors and nurses to keep patient health under observation while the patient remains in the comfort of their home [30,38,54]. This technology could reduce the burden on caregivers and limit accidents. Last, it is our hope that these findings will influence policy that increases funding for the development of digital health tools that can decrease health disparities within rural older adult populations.

Limitations

This study has several strengths and limitations. The synthesis of our findings was guided by the core domains of the TAM [69]. Previous studies have reported that the Perceived Usefulness, Perceived Ease of Use, and Intention to Use domains are important for determining user acceptability. They are the core of all TAM models used across the literature, thus chosen for this study. Yet, additional domains and extensions have been added over time, including the "Attitudes Toward Using" domain [70,71]. Although our synthesis is limited to conclusions derived from the core domains, these domains provide sufficient information to inform our overall aim of assessing user acceptability. Additionally, a trained librarian conducted a data search of publicly available databases (eg, PubMed). Despite our comprehensive search conducted by a trained librarian, some relevant studies may have been missed due to factors such as publication dates postreview, non–English language restrictions, or potential oversight in keyword selection. It is important to note that participants in health outcomes research are often younger than 65, have higher household incomes, greater technological literacy, and are more likely to reside in urban areas [27,33,72]. Therefore, our review's focus on rural older adults and digital tool use may highlight a potential participant bias in our sample, limiting the generalizability of our findings to the broader rural older adult population. As with all scholarly reviews, researcher bias could have impacted the results. However, because multiple reviewers assessed each manuscript and attended consensus meetings for each manuscript, this bias was reduced.

Conclusion

This study aimed to systematically review articles that incorporated digital health tools used by rural older adults in order to assess their acceptability and usage of the tools. Following the TAM, we highlighted the usefulness, ease of use, and intention to use digital health tools. In summary, digital health tools were valuable for rural older adults with complex care needs, helping mitigate access barriers and support care management tasks like scheduling and patient-provider communication. While rural older adults and providers found the tools user-friendly, external factors such as technical issues and equipment limitations impeded usability, signaling the need for additional support and resources. Despite challenges with ease of use, rural older adults expressed an intention to continue using digital health tools, recognizing their overall benefits in managing care, especially in underserved areas. As medical deserts widen in rural communities, and in response to the rapid uptake of telemedicine due to COVID-19 precautions, digital health tool reliance is likely to grow for rural residents. Understanding what this uniquely vulnerable population views as acceptable and what facilitates their uptake of digital health tools is critical for addressing health disparities and bridging the digital divide in rural communities.

Funding

MWL and ZMS were supported by a grant from the National Cancer Institute (Grant Number: K01CA262342). MWL is also supported by a Northwestern University Clinical and Translational Sciences Institute grant (Grant Numbers: NUCATS; UL1TR001422, Principal Investigator: Richard D'Aquilla), an Institutional Research Grant (IRG-21-144-27) from the American Cancer Society (Principal Investigator: Leonidas Platanias), and funds from the Northwestern University Center for Community Health (grant number: not applicable). Opinions and comments expressed in this paper belong to the authors and do not necessarily reflect those of the National Institutes of Health.

Authors' Contributions

ZMS was responsible for formal analysis, methodology, project administration, visualization, writing the original draft, and reviewing and editing subsequent drafts. EQ was responsible for the formal analysis, writing the original draft, and reviewing and editing subsequent drafts. JP was responsible for writing the original draft and reviewing and editing subsequent drafts.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA checklist.
[PDF File (Adobe PDF File), 140 KB - aging_v9i1e70012_app1.pdf]

References

XSL•FO
RenderX

1. Shaver J. The state of telehealth before and after the COVID-19 pandemic. Prim Care 2022;49(4):517-530 [FREE Full text] [doi: 10.1016/j.pop.2022.04.002] [Medline: 36357058]
2. HHS announces $16.3 million to expand telehealth care in the title X family planning program. Office of Population Affairs. 2022. URL: https://opa.hhs.gov/about/news/grant-award-announcements/hhs-announces-16-million-expand-telehealth-care-title-x [accessed 2025-12-16]
3. Haque SN. Telehealth beyond COVID-19. Psychiatr Serv 2021;72(1):100-103. [doi: 10.1176/appi.ps.202000368] [Medline: 32811284]
4. Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly T, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020;69(43):1595-1599 [FREE Full text] [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
5. Asante D, McLachlan CS, Pickles D, Isaac V. Understanding unmet care needs of rural older adults with chronic health conditions: a qualitative study. Int J Environ Res Public Health 2023;20(4) [FREE Full text] [doi: 10.3390/ijerph20043298] [Medline: 36833993]
6. Skoufalos A, Clarke JL, Ellis DR, Shepard VL, Rula EY. Rural aging in America: proceedings of the 2017 connectivity summit. Popul Health Manag 2017;20(S2):S1-S10 [FREE Full text] [doi: 10.1089/pop.2017.0177] [Medline: 29251548]
7. Eastman MR, Kalesnikava VA, Mezuk B. Experiences of care coordination among older adults in the United States: evidence from the health and retirement study. Patient Educ Couns 2022;105(7):2429-2435 [FREE Full text] [doi: 10.1016/j.pec.2022.03.015] [Medline: 35331572]
8. Probst J, Eberth JM, Crouch E. Structural urbanism contributes to poorer health outcomes for rural America. Health Aff (Millwood) 2019;38(12):1976-1984. [doi: 10.1377/hlthaff.2019.00914] [Medline: 31794301]
9. Cyr ME, Etchin AG, Guthrie BJ, Benneyan JC. Access to specialty healthcare in urban versus rural US populations: a systematic literature review. BMC Health Serv Res 2019;19(1):974 [FREE Full text] [doi: 10.1186/s12913-019-4815-5] [Medline: 31852493]
10. Sekhon H, Sekhon K, Launay C, Afililo M, Innocente N, Vahia I, et al. Telemedicine and the rural dementia population: a systematic review. Maturitas 2021;143:105-114. [doi: 10.1016/j.maturitas.2020.09.001] [Medline: 33308615]
11. Hodge H, Carson D, Carson D, Newman L, Garrett J. Using internet technologies in rural communities to access services: the views of older people and service providers. J Rural Stud 2017;54:469-478 [FREE Full text] [doi: 10.1016/j.jrurstud.2016.06.016]
12. Call VRA, Erickson LD, Dailey NK, Hicken BL, Rupper R, Yorgason JB, et al. Attitudes toward telemedicine in urban, rural, and highly rural communities. Telemed J E Health 2015;21(8):644-651. [doi: 10.1089/tmj.2014.0125] [Medline: 25839334]
13. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. JMIR Med Inform 2020;8(8):e20359 [FREE Full text] [doi: 10.2196/20359] [Medline: 32784177]
14. Abdallah L, Stolee P, Lopez KJ, Whate A, Boger J, Tong C. The impact of COVID-19 on older adults' perceptions of virtual care: qualitative study. JMIR Aging 2022;5(4):e38546 [FREE Full text] [doi: 10.2196/38546] [Medline: 36054599]
15. Hunter I, Lockhart C, Rao V, Tootell B, Wong S. Enabling rural telehealth for older adults in underserved rural communities: focus group study. JMIR Form Res 2022;6(11):e35864 [FREE Full text] [doi: 10.2196/35864] [Medline: 36331533]
16. Cho Y, Yang R, Gong Y, Jiang Y. Use of electronic communication with clinicians among cancer survivors: health information national trend survey in 2019 and 2020. Telemed J E Health 2023;29(6):866-874 [FREE Full text] [doi: 10.1089/tmj.2022.0203] [Medline: 36355055]
17. Svistova J, Harris C, Fogarty B, Kulp C, Lee A. Use of telehealth amid the COVID-19 pandemic: experiences of mental health providers serving rural youth and elderly in Pennsylvania. Adm Policy Ment Health 2022;49(4):530-538 [FREE Full text] [doi: 10.1007/s10488-021-01181-z] [Medline: 34846613]
18. Rush KL, Singh S, Seaton CL, Burton L, Li E, Jones C, et al. Telehealth use for enhancing the health of rural older adults: a systematic mixed studies review. Gerontologist 2022;62(10):e564-e577. [doi: 10.1093/geront/gnab141] [Medline: 34661675]
19. Lama Y, Davidoff AJ, Vanderpool RC, Jensen RE. Telehealth availability and use of related technologies among medicare-enrolled cancer survivors: cross-sectional findings from the onset of the COVID-19 Pandemic. J Med Internet Res 2022;24(1):e34616 [FREE Full text] [doi: 10.2196/34616] [Medline: 34978531]
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
21. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
22. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-340. [doi: 10.2307/249008]

XSL•FO
RenderX

23. Anderson J, Godwin KM, Petersen NJ, Willson P, Kent TA. A pilot test of videoconferencing to improve access to a stroke risk-reduction programme for veterans. J Telemed Telecare 2013;19(3):153-159. [doi: 10.1177/1357633X13479703] [Medline: 23625913]
24. Barton AJ, Amura CR, Willems EL, Medina R, Centi S, Hernandez T, et al. Patient and provider perceptions of COVID-19-driven telehealth use from nurse-led care models in rural, frontier, and urban colorado communities. J Patient Exp 2023;10:23743735231151546 [FREE Full text] [doi: 10.1177/23743735231151546] [Medline: 36741820]
25. Bernacchi V, LeBaron VT, Hinton ID, DeGuzman PB. Rural cancer survivors' perceptions of a nurse-led telehealth intervention to manage cancer-related distress. Oncol Nurs Forum 2023;50(2):173-184. [doi: 10.1188/23.ONF.173-184] [Medline: 37677802]
26. Bonsignore L, Bloom N, Steinhauser K, Nichols R, Allen T, Twaddle M, et al. Evaluating the feasibility and acceptability of a telehealth program in a rural palliative care population: tapcloud for palliative care. J Pain Symptom Manage 2018;56(1):7-14 [FREE Full text] [doi: 10.1016/j.jpainsymman.2018.03.013] [Medline: 29551433]
27. Browning SV, Clark RC, Poff RM, Todd D. Telehealth monitoring: a smart investment for home care patients with heart failure? Home Healthc Nurse 2011;29(6):368-374. [doi: 10.1097/NHH.0b013e31821b7186] [Medline: 21633228]
28. Collie K, Kreshka MA, Ferrier S, Parsons R, Graddy K, Avram S, et al. Videoconferencing for delivery of breast cancer support groups to women living in rural communities: a pilot study. Psychooncology 2007;16(8):778-782. [doi: 10.1002/pon.1145] [Medline: 17253594]
29. Cummings DM, Morrissey S, Barondes MJ, Rogers L, Gustke S. Screening for diabetic retinopathy in rural areas: the potential of telemedicine. J Rural Health 2001;17(1):25-31. [doi: 10.1111/j.1748-0361.2001.tb00251.x] [Medline: 11354719]
30. DeHart D, King LB, Iachini AL, Browne T, Reitmeier M. Benefits and challenges of implementing telehealth in rural settings: a mixed-methods study of behavioral medicine providers. Health Soc Work 2022;47(1):7-18. [doi: 10.1093/hsw/hlab036] [Medline: 34910158]
31. Demaerschalk BM, Bobrow BJ, Raman R, Kiernan TJ, Aguilar MI, Ingall TJ, STRokE DOC AZ TIME Investigators. Stroke team remote evaluation using a digital observation camera in Arizona: the initial mayo clinic experience trial. Stroke 2010;41(6):1251-1258 [FREE Full text] [doi: 10.1161/STROKEAHA.109.574509] [Medline: 20431081]
32. DeVido J, Glezer A, Branagan L, Lau A, Bourgeois JA. Telepsychiatry for inpatient consultations at a separate campus of an academic medical center. Telemed J E Health 2016;22(7):572-576. [doi: 10.1089/tmj.2015.0125] [Medline: 26701608]
33. Donahue KE, Tillman J, Halladay JR, Cené CW, Hinderliter A, Cummings DM, et al. Lessons learned from implementing health coaching in the heart healthy lenoir hypertension study. Prog Community Health Partnersh 2016;10(4):559-567. [doi: 10.1353/cpr.2016.0064] [Medline: 28569681]
34. Finley BA, Palitsky R, Charteris E, Pacheco C, Kapoor D. Outpatient telecardiology perceptions among rural, suburban, and urban veterans utilizing in-person cardiology versus telecardiology services: a mixed methods analysis. J Rural Health 2021;37(4):812-820. [doi: 10.1111/jrh.12586] [Medline: 34002404]
35. Geller BM, Skelly JM, Dorwaldt AL, Howe KD, Dana GS, Flynn BS. Increasing patient/physician communications about colorectal cancer screening in rural primary care practices. Med Care 2008;46(9 Suppl 1):S36-S43 [FREE Full text] [doi: 10.1097/MLR.0b013e31817c60ea] [Medline: 18725831]
36. Gutierrez J, Moeckli J, Holcombe A, O'Shea AM, Bailey G, Rewerts K, et al. Implementing a telehospitalist program between veterans health administration hospitals: outcomes, acceptance, and barriers to implementation. J Hosp Med 2021;16(3):156-163 [FREE Full text] [doi: 10.12788/jhm.3570] [Medline: 33617436]
37. Hatch MN, Martinez RN, Etingen B, Cotner B, Hogan TP, Wickremasinghe IM, et al. Characterization of telehealth use in veterans with spinal cord injuries and disorders. PM R 2021;13(10):1094-1103. [doi: 10.1002/pmrj.12515] [Medline: 33098620]
38. Hicks LL, Fleming DA, Desaulnier A. The application of remote monitoring to improve health outcomes to a rural area. Telemed J E Health 2009;15(7):664-671. [doi: 10.1089/tmj.2009.0009] [Medline: 19694598]
39. Holloway B, Coon P, Kersten DW, Ciemins EL. Telehealth in rural Montana: promoting realistic independent self-management of diabetes. Diabetes Spectrum 2011;24(1):50-54. [doi: 10.2337/diaspect.24.1.50]
40. Khairat S, Haithcoat T, Liu S, Zaman T, Edson B, Gianforcaro R, et al. Advancing health equity and access using telemedicine: a geospatial assessment. J Am Med Inform Assoc 2019;26(8-9):796-805 [FREE Full text] [doi: 10.1093/jamia/ocz108] [Medline: 31340022]
41. Kulcsar Z, Albert D, Ercolano E, Mecchella JN. Telerheumatology: a technology appropriate for virtually all. Semin Arthritis Rheum 2016;46(3):380-385. [doi: 10.1016/j.semarthrit.2016.05.013] [Medline: 27395561]
42. Liu Y, Zupan NJ, Swearingen R, Jacobson N, Carlson JN, Mahoney JE, et al. Identification of barriers, facilitators and system-based implementation strategies to increase teleophthalmology use for diabetic eye screening in a rural US primary care clinic: a qualitative study. BMJ Open 2019;9(2):e022594 [FREE Full text] [doi: 10.1136/bmjopen-2018-022594] [Medline: 30782868]
43. Locke ER, Thomas RM, Woo DM, Nguyen EHK, Tamanaha BK, Press VG, et al. Using video telehealth to facilitate inhaler training in rural patients with obstructive lung disease. Telemed J E Health 2019;25(3):230-236 [FREE Full text] [doi: 10.1089/tmj.2017.0330] [Medline: 30016216]

XSL•FO
RenderX

44. McIlhenny CV, Guzic BL, Knee DR, Wendekier CM, Demuth BR, Roberts JB. Using technology to deliver healthcare education to rural patients. Rural Remote Health 2011;11(4):1798. [Medline: 21995854]
45. Owolo E, Petitt Z, Rowe D, Luo E, Bishop B, Poehlein E, et al. Sociodemographic trends in telemedicine visit completion in spine patients during the COVID-19 pandemic. Spine (Phila Pa 1976) 2023;48(21):1500-1507. [doi: 10.1097/BRS.0000000000004617] [Medline: 37235789]
46. Robinson KS, Morin PC, Shupe JAC, Izquierdo R, Ploutz-Snyder R, Meyer S, et al. Use of three computer training methods in elderly underserved rural patients enrolled in a diabetes telemedicine program. Comput Inform Nurs 2010;28(3):172-177. [doi: 10.1097/NCN.0b013e3181d785d5] [Medline: 20431360]
47. Rodriguez KL, Burkitt KH, Bayliss NK, Skoko JE, Switzer GE, Zickmund SL, et al. Veteran, primary care provider, and specialist satisfaction with electronic consultation. JMIR Med Inform 2015;3(1):e5 [FREE Full text] [doi: 10.2196/medinform.3725] [Medline: 25589233]
48. Schlittenhardt M, Smith SC, Ward-Smith P. Tele-continence care: a novel approach for providers. Urol Nurs 2016;36(5):217-223. [Medline: 29240333]
49. Schooley BL, Horan TA, Lee PW, West PA. Rural veteran access to healthcare services: investigating the role of information and communication technologies in overcoming spatial barriers. Perspect Health Inf Manag 2010;7(Spring):1f [FREE Full text] [Medline: 20697468]
50. Silvestrini M, Indresano J, Zeliadt SB, Chen JA. "There's a huge benefit just to know that someone cares:" a qualitative examination of rural veterans' experiences with TelePain. BMC Health Serv Res 2021;21(1):1111 [FREE Full text] [doi: 10.1186/s12913-021-07133-5] [Medline: 34656133]
51. Strowd RE, Strauss L, Graham R, Dodenhoff K, Schreiber A, Thomson S, et al. Rapid implementation of outpatient teleneurology in rural appalachia: barriers and disparities. Neurol Clin Pract 2021;11(3):232-241 [FREE Full text] [doi: 10.1212/CPJ.0000000000000906] [Medline: 34484890]
52. Switzer JA, Hall CE, Close B, Nichols FT, Gross H, Bruno A, et al. A telestroke network enhances recruitment into acute stroke clinical trials. Stroke 2010;41(3):566-569. [doi: 10.1161/STROKEAHA.109.566844] [Medline: 20056929]
53. Virmani T, Lotia M, Glover A, Pillai L, Kemp AS, Iyer A, et al. Feasibility of telemedicine research visits in people with Parkinson's disease residing in medically underserved areas. J Clin Transl Sci 2022;6(1):e133 [FREE Full text] [doi: 10.1017/cts.2022.459] [Medline: 36590358]
54. Waymouth M, Siconolfi D, Friedman EM, Saliba D, Ahluwalia SC, Shih RA. Barriers and facilitators to home- and community-based services access for persons with dementia and their caregivers. J Gerontol B Psychol Sci Soc Sci 2023;78(6):1085-1097 [FREE Full text] [doi: 10.1093/geronb/gbad039] [Medline: 36896936]
55. Weiner MF, Rossetti HC, Harrah K. Videoconference diagnosis and management of choctaw indian dementia patients. Alzheimers Dement 2011;7(6):562-566 [FREE Full text] [doi: 10.1016/j.jalz.2011.02.006] [Medline: 22055972]
56. West VL, Milio N. Organizational and environmental factors affecting the utilization of telemedicine in rural home healthcare. Home Health Care Serv Q 2004;23(4):49-67. [doi: 10.1300/J027v23n04_04] [Medline: 15778152]
57. West SP, Lagua C, Trief PM, Izquierdo R, Weinstock RS. Goal setting using telemedicine in rural underserved older adults with diabetes: experiences from the informatics for diabetes education and telemedicine project. Telemed J E Health 2010;16(4):405-416. [doi: 10.1089/tmj.2009.0136] [Medline: 20507198]
58. Zulman DM, Wong EP, Slightam C, Gregory A, Jacobs JC, Kimerling R, et al. Making connections: nationwide implementation of video telehealth tablets to address access barriers in veterans. JAMIA Open 2019;2(3):323-329 [FREE Full text] [doi: 10.1093/jamiaopen/ooz024] [Medline: 32766533]
59. Munn Z, Aromataris E, Tufanaru C, Stern C, Porritt K, Farrow J, et al. The development of software to support multiple systematic review types: the Joanna Briggs institute system for the unified management, assessment and review of information (JBI SUMARI). Int J Evid Based Healthc 2019;17(1):36-43. [doi: 10.1097/XEB.0000000000000152] [Medline: 30239357]
60. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The mixed methods appraisal tool (MMAT) version 2018 for information professionals and researchers. Educ Inf 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
61. Depatie A, Bigbee JL. Rural older adult readiness to adopt mobile health technology: a descriptive study. Online J Rural Nurs Health Care 2013;15(1):150-184. [doi: 10.14574/ojrnhc.v15i1.346]
62. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. J Med Internet Res 2019;21(4):e11604 [FREE Full text] [doi: 10.2196/11604] [Medline: 30958272]
63. Lim CT, Rosenfeld LC, Nissen NJ, Wang PS, Patel NC, Powers BW, et al. Remote care management for older adult populations with elevated prevalence of depression or anxiety and comorbid chronic medical illness: a systematic review. J Acad Consult Liaison Psychiatry 2022;63(3):198-212 [FREE Full text] [doi: 10.1016/j.jaclp.2022.02.005] [Medline: 35189427]
64. Louras N, Reading Turchioe M, Shafran Topaz L, Ellison M, Abudu-Solo J, Blutinger E, et al. Mobile integrated health interventions for older adults: a systematic review. Innov Aging 2023;7(3):igad017 [FREE Full text] [doi: 10.1093/geroni/igad017] [Medline: 37090165]

XSL•FO
RenderX

65. Kainiemi E, Saukkonen P, Virtanen L, Vehko T, Kyytsönen M, Aaltonen M, et al. Perceived benefits of digital health and social services among older adults: a population-based cross-sectional survey. Digit Health 2023;9:20552076231173559 [FREE Full text] [doi: 10.1177/20552076231173559] [Medline: 37312955]
66. Busby J, Tanberk J. FCC underestimates Americans unserved by broadband internet by 50%. BroadBand Now Team. 2020. URL: https://broadbandnow.com/research/fcc-underestimates-unserved-by-50-percent [accessed 2025-12-16]
67. Cheung K, Chau AKC, Woo J, Lai ET. The age-based digital divide in an increasingly digital world: a focus group investigation during the COVID-19 pandemic. Arch Gerontol Geriatr 2023;115:105225. [doi: 10.1016/j.archger.2023.105225] [Medline: 37837792]
68. Cousins MM, Van Til M, Steppe E, Ng S, Ellimoottil C, Sun Y, et al. Age, race, insurance type, and digital divide index are associated with video visit completion for patients seen for oncologic care in a large hospital system during the COVID-19 pandemic. PLoS One 2022;17(11):e0277617 [FREE Full text] [doi: 10.1371/journal.pone.0277617] [Medline: 36395112]
69. Davis FD, Venkatesh V. A critical assessment of potential measurement biases in the technology acceptance model: three experiments. Int J Hum Comput Stud 1996;45(1):19-45. [doi: 10.1006/ijhc.1996.0040]
70. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
71. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
72. Saphner T, Marek A, Homa J, Robinson L, Glandt N. Clinical trial participation assessed by age, sex, race, ethnicity, and socioeconomic status. Contemp Clin Trials 2021;103:106315 [FREE Full text] [doi: 10.1016/j.cct.2021.106315] [Medline: 33626412]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
TAM: technology acceptance model

Edited by M Mardini; submitted 30.Dec.2024; peer-reviewed by NA Puccinelli-Ortega, LR Guo; comments to author 22.Jul.2025; revised version received 09.Sep.2025; accepted 10.Nov.2025; published 04.Feb.2026.

Please cite as:
Siegel ZM, Quinkert E, Pai J, Miller CH, Lewis MW
Lessons Learned About Digital Health Tool Acceptability Among Rural Older Adults: Systematic Review Guided by the Technology Acceptance Model
JMIR Aging 2026;9:e70012
URL: https://aging.jmir.org/2026/1/e70012
doi:10.2196/70012
PMID:

©Zachary M Siegel, Ellie Quinkert, Jiya Pai, Corinne H Miller, Marquita W Lewis. Originally published in JMIR Aging (https://aging.jmir.org), 04.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Service Design Strategies to Enhance Exercise Adherence in Extended Reality Interventions for Older Adults: Systematic Review

Jiangpan Niu[1], BFA, MEng; Yuanyuan Yin[1], BEng, MA, PhD; Shan Wang[1], BA, MA, PhD

Department of Design, Winchester School of Art, University of Southampton, Winchester, United Kingdom

Corresponding Author:
Jiangpan Niu, BFA, MEng
Department of Design
Winchester School of Art
University of Southampton
Park Avenue Campus
Winchester, SO23 8DL
United Kingdom
Phone: 44 7726428590
Email: jn5n24@soton.ac.uk

Abstract

Background: Exercise has a positive impact on the health of older adults. However, due to physical conditions, psychological factors, and external environment constraints, older adults still face significant challenges in maintaining exercise. Exercise adherence is relatively low. Extended reality (XR) technology offers new ways for older adults to exercise and improve their adherence. Existing research mainly focuses on short-term effects, paying insufficient attention to maintaining long-term engagement and establishing effective incentive mechanisms. By introducing service design methods, user experience, stakeholder collaboration, and adherence support can be better integrated at different stages of exercise intervention, thereby enhancing the willingness and enthusiasm of older adults to continue to participate in exercise.

Objective: This review aims to evaluate how XR-based exercise interventions targeting populations with a mean age of older than 50 years integrate service design strategies to enhance exercise adherence and to identify critical gaps in their long-term application.

Methods: A systematic review was conducted across PubMed, Scopus, Web of Science, CINAHL, PsycINFO, and ACM Digital Library (January 2020-July 2025). Eligible studies (1) used XR to support exercise or rehabilitation, (2) included participants with a mean age of older than 50 years, (3) reported at least one service design activity, and (4) provided adherence-related outcomes. Dual independent screening and structured data extraction were performed.

Results: A total of 9 studies (242 participants) met the inclusion criteria. Most applied participatory, co-design, or user-centered design with iterative prototyping, but few advanced to full implementation or applied a complete service design cycle. The user experience was generally positive, but evaluations were primarily based on self-reports with limited objective tracking. Although exercise adherence was generally high in most studies (80%-100%), this assessment was primarily based on attendance-related indicators. There was a lack of consistency in how adherence was defined across studies, and no unified measure of exercise adherence was established. Exercise frequency, duration, and attendance were commonly reported, whereas exercise intensity and accuracy were often overlooked. Existing evaluations also lacked long-term tracking of exercise adherence. Regarding intervention delivery settings, most studies were conducted in laboratories, hospitals, and care facilities. Few studies investigated implementation in community settings, which made XR interventions difficult to adapt to the real-world conditions faced by older adults in their daily lives and hindered their promotion and application to a broader community population.

Conclusions: Current XR exercise interventions for older adults show promising short-term adherence but rarely embed service design continuity or comprehensive adherence monitoring. The combination of structured service design processes with standardized and multidimensional adherence indicators can provide strong support for participants to continue engaging in XR exercise projects. When implemented in community settings, these interventions can enhance scalability and better support an age-friendly XR exercise system.

XSL•FO
RenderX

(JMIR Aging 2026;9:e86595) doi:10.2196/86595

KEYWORDS

extended reality; older adults; exercise adherence; service design; systematic review

Introduction

As life expectancy increases, the older population has grown rapidly, and chronic disease prevalence has risen. Many older adults live with long-term conditions such as cardiovascular disease [1], type 2 diabetes [2], obesity [3], arthritis [4], and dementia [5]. The growing number of such cases is already putting pressure on health care systems globally [6,7]. For example, health care costs for frail older adults range from US $2540 to US $221,400 per person annually [8]. In Singapore, the societal cost associated with multimorbid older adults amounts to US $11,970 per person per year [9]. In Europe and North America, health care expenditures for older adults with multimorbidity are likewise considerable. For example, among older patients with heart failure and diabetes, annual hospital care spending was estimated at approximately US $10,956 per person in England and US $30,877 per person in the United States [10]. Regular exercise helps older adults maintain their health and reduces a variety of age-related problems, such as cardiovascular disease [11,12], decreased musculoskeletal strength [13], risk of falls [14], and psychological distress [15]. Despite the clear benefits of exercise, older adults generally lack the ability to continue exercising [16]. For example, the adherence rate in an aerobic exercise group of older adults averaged only 49.7% [17], while a 6-month intervention reported a 35% dropout rate among participants aged 65 to 74 years [18].

Older adults face numerous barriers to exercise, including physical fatigue and pain [19,20], psychological concerns and low confidence [21,22], and unsafe or unsuitable spaces in their communities [16,23]. Lack of sustained support and encouragement also makes it difficult for many older adults to maintain long-term exercise [19]. These issues suggest that more flexible interventions that can adapt to individual needs are needed to help older adults cope with physical, psychological, and environmental challenges.

To address these challenges, technologies such as virtual reality (VR), augmented reality (AR), and mixed reality (MR), collectively referred to as extended reality (XR) technology, have been proposed as a potential means of increasing the frequency and accessibility of physical activity among older adults. XR interventions also contain key features such as immersive environments, real-time interactive feedback, and gamification [24-26]. Some studies have reported that VR helps older adults better control blood sugar, improve motor function, and enhance emotional health [25,27,28]. In addition, studies have shown that group activities based on VR decrease loneliness among older adults and increase their enthusiasm and initiative for exercise [29,30]. AR and MR applications support mobility and balance training and are particularly suitable for older adults in institutional care [31,32]. Even though these results are encouraging, most XR-based exercise programs are short-term. Their main goal is to test the usability of the technology, rather than exploring how to sustain long-term behavioral changes [33,34]. Current XR device designs often overlook the needs of older users, leading many to experience difficulties and frustration. Interfaces may be too complex, visuals too intense, or the space required may exceed what their home can provide [35-38].

Since most XR interventions lack long-term planning, service design has gained attention as a way to build more sustainable solutions. It provides a flexible framework to connect user needs, coordinate stakeholders, and maintain consistent experiences over time [39,40]. User journey mapping [41,42], service blueprints [43], and co-creation methods [44] offer ways to identify behavioral touchpoints, track drop-off moments, and target interventions towards a dynamic user experience (UX). However, in many XR exercise studies, service design approaches are often applied in the early design or prototyping stages and rarely receive ongoing support during implementation and evaluation [45,46]. As a result, UXs become fragmented, and behavioral insights are not fully leveraged across the intervention process.

This review responds to existing gaps by taking a service design perspective to explore XR-based physical activity programs for older adults. Adherence is not treated as a fixed individual outcome, but as a dynamic process influenced by ongoing design efforts, emotional involvement, and multisystem collaboration [40,47]. This review adopts the TiSDD (This is Service Design Doing) framework [40] to analyze how service design methods contribute to each phase of XR interventions. The framework includes a 4-stage process covering research, ideation, prototyping, and implementation. Drawing on insights from participatory methods, user-centered design, and health service research, this review examines how sustained adherence is supported and identifies challenges arising from fragmented services. It further informs the development of XR interventions that are attuned to both the needs of older adults and the full exercise engagement cycle.

This review focuses on three core questions:

1. What kinds of service design strategies have been used in XR-based exercise programs for older adults?
2. In what ways are these tools applied to support continuity and improve the user experience throughout the intervention process?
3. What types of adherence outcomes have been observed in these interventions?

Methods

Study Design

This study conducted a systematic review to synthesize diverse research, focusing on how XR-based exercise interventions for older adults can address exercise adherence through service design strategies. This review followed the transparent, structured process of the PRISMA (Preferred Reporting Items

XSL•FO
RenderX

for Systematic Reviews and Meta-Analyses) guidelines (Multimedia Appendix 1) [48] to ensure reproducibility and methodological rigor.

Eligibility Criteria

Studies were included if they met all of the following criteria: (1) they enrolled participants with a mean age of older than 50 years—this standard is based on the World Health Organization's definition of older people in resource-limited settings, where life expectancy is generally shorter [49]; (2) they focused on XR-based interventions involving physical activity, exercise, or rehabilitation; (3) they reported the use of at least one service design method (eg, personas, cocreation, user journey mapping, user-centered design, or participatory design) during the development or implementation of the intervention; (4) they reported outcome data related to exercise adherence, including at least one of the following dimensions: frequency, duration, intensity, or accuracy of exercise performance; and (5) they used or proposed quantitative, qualitative, or mixed method evaluation strategies, with evidence of feasibility, usability, or preliminary outcomes.

Studies were excluded if they met any of the following conditions: (1) they focused solely on cognitive training or entertainment without an exercise component; (2) they did not include older adults as the primary target population; or (3) they were editorials, commentaries, opinion pieces, or lacked sufficient methodological detail or relevance to the research question.

Information Sources and Search Strategy

A structured search was conducted across 6 academic databases: PubMed, Scopus, Web of Science, CINAHL, PsycINFO, and the ACM Digital Library. The search strategy combined Boolean operators and thematic clusters covering: XR technologies ("virtual reality" OR "augmented reality" OR "mixed reality" OR "extended reality" OR "immersive technology"), service design methods ("service design" OR "user journey" OR "journey mapping" OR "experience map" OR "co-design" OR "participatory design" OR "user-centered design" OR "human-centered design"), and the target population and activity ("older adult" OR "elderly" OR "geriatric" AND "exercise" OR "physical activity" OR "rehabilitation"). Searches were limited to peer-reviewed papers published in English between January 2020 and July 2025. Full-text availability was required for eligibility assessment and data extraction; records for which the full text could not be accessed were excluded at the screening stage. Gray literature sources and trial registries were not systematically searched, as this review focused on peer-reviewed studies that provided sufficient methodological detail on service design activities and XR intervention implementation. The full database-specific search strategies are detailed in Multimedia Appendix 2.

Study Selection

All retrieved records were imported into a Zotero library (version 7.0.24; Corporation for Digital Scholarship), and duplicates were removed. Two independent reviewers (JN and SW) screened the titles, abstracts, and full texts of the remaining papers against the predefined inclusion and exclusion criteria. Discrepancies were resolved through discussion; if consensus could not be reached, a third reviewer (YY) was consulted.

Data Extraction and Management

Data extraction was performed using a standardized Microsoft Excel 2021 spreadsheet to ensure consistency across studies. Two reviewers independently extracted data on study characteristics, service design strategies, UX, and adherence outcomes. Discrepancies were resolved through discussion until consensus was reached. To classify the application stages of service design strategies, this study adopted the TiSDD framework [40]. Table 1 summarizes the definitions and typical activities of each stage, which served as the reference for coding service design applications in the included studies.

Table 1. Service Design Stages Defined by the TiSDD (This is Service Design Doing) framework.

Stage	Description	Typical activities
Research	Understanding users, behaviors, and context	Ethnography, interviews, observation, participatory workshops, and data analysis
Ideation	Generating, refining, and prioritizing concepts	Brainstorming, clustering, voting, and decision matrices
Prototyping	Developing and testing service concepts	Storyboarding, role-playing, mock-ups, and digital/physical prototypes
Implementation	Deploying and managing interventions	Change management, technical development, training, and monitoring

Data Synthesis

Data synthesis was organized around 3 research questions and divided into 4 sections, including study characteristics, service design outcomes, UX outcomes, and adherence outcomes. This thematic framework enabled a narrative synthesis that examined how service design tools were used, how they supported continuity and UX, and what types of adherence outcomes were reported.

Quality Assessment

To assess the quality of the studies, critical appraisal was conducted using the Mixed Methods Appraisal Tool (MMAT; version 2018; McGill University) [50]. The MMAT is a specially designed tool that can be used to assess the quality of different types of studies in the same review, including qualitative, quantitative, and mixed methods studies. The quality of all studies was assessed by the first reviewer (JN) and checked for accuracy by 2 other authors (YY and SW). In line with

XSL•FO
RenderX

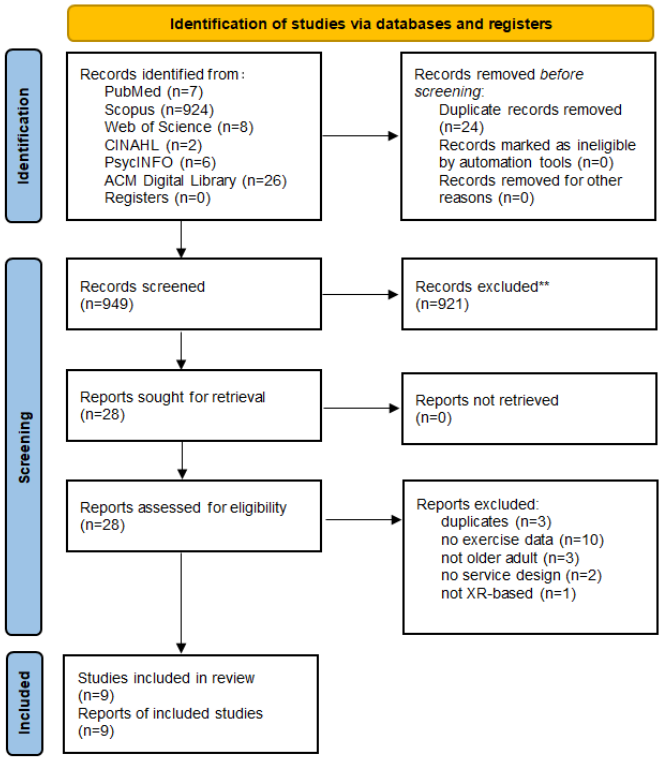
MMAT guidance, the tool was not used to score or exclude studies, but to support the interpretation of the findings [50].

Results

Sample

The total number of identified records was 973. After duplicates were removed, 949 records remained, and these were screened for title and abstract relevance. The number of excluded studies after title and abstract screening was 921. The full texts of the remaining 28 papers were read and analyzed, of which 19 were excluded due to factors such as the absence of exercise-related data, duplication, non–older adult populations, lack of service design elements, or lack of XR-based interventions. As a result, 9 studies were included in the final review (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram of the study selection process.



Study Characteristics

This review included 9 studies published between 2020 and 2025, conducted across 9 countries: Italy, Portugal, the Netherlands, the United Kingdom, Germany, Spain, Canada, Cyprus, and China. Of the 9 studies, 67% (6/9) were mixed methods designs [51-56], 11% (1/9) was a mixed methods study with an embedded randomized controlled trial [57], 11% (1/9) was a quasi-randomized controlled trial [58], and 11% (1/9) was a nonrandomized quantitative study [59]. A total of 242 participants were involved across all studies, with sample sizes ranging from 6 to 75. The mean age of participants ranged from 52.8 to 86.8 years. A total of 89% (8/9) included both male and female participants [51-54,56-59], while 11% (1/9) did not report gender information [55].

All 100% (9/9) studies used VR as the XR modality [51-59]. A total of 67% (6/9) of studies reported specific XR headset models [51-53,55,56,58], while 33% (3/9) did not specify the hardware used [54,57,59]. A total of 78% (7/9) of studies described the XR interaction modality [51,52,54-58], whereas 22% (2/9) did not report interaction details [53,59]. A total of 100% (9/9) of studies reported some form of XR supervision and support during intervention delivery [51-59]. A total of 33% (3/9) of studies reported information on XR cybersickness or tolerability [51,54,55], while 67% (6/9) did not report tolerability-related data [52,53,56-59].

Exercise types included 44% (4/9) upper limb rehabilitation interventions [51,52,55,56], 22% (2/9) gait or balance training [54,58], 11% (1/9) lower limb training [57], 11% (1/9) adapted physical activity [59], and 11% (1/9) multidomain physical-cognitive rehabilitation [53].

Intervention settings included 44% (4/9) hospital-based studies [51-53,57], 11% (1/9) university or research laboratory [58], 11% (1/9) biomedical research institute [54], 11% (1/9) long-term care residence [56], and 11% (1/9) home-based environment [55], 11% (1/9) study used a mixed-setting intervention across nursing homes, hospitals, and home environments [59] (Table 2).

XSL•FO
RenderX

Table 2. Study characteristics.

Author (year), country	Study design	Sample size	Age (years)	Sex	XR[a] type	XR hardware	XR interaction Modality	XR supervision and support	XR cybersickness / tolerability	Exercise type	Intervention setting
De Luca et al [59], (2025), Italy, Portugal, Netherlands	Nonrandomized quantitative	75	Mean age 79.8 (SD 9.96)	F[b] and M[c]	VR[d]	Not reported	Not reported	Trainer- or coach-supported use	Not reported	Adapted physical activity	Nursing homes, hospitals, and home settings
Sweeney et al [53], (2025), United Kingdom	Mixed methods	60	Mean age 69.6 (SD 14.0)	F and M	VR	Meta Quest headset	Not reported	Supervised group-based sessions with rehabilitation staff	Not reported	Multitechnology rehabilitation training: upper limb, lower limb, balance, cognitive, and whole-body functional training	Hospital-based intervention
Ciemer et al [58], (2025), Germany	Quasi-RCT[e]	32	Mean age 70.00 (SD 3.33)	F and M	VR	HTC Vive Pro 2	Walking-based interaction	On-site supervision by 2 trained supervisors, with real-time monitoring via an external screen	Not reported	Gait-based fall prevention exergame	University laboratory
Bosch-Barceló et al [54], (2024), Spain	Mixed methods	8	Mean age 61 (SD 9)	F and M	VR	Not reported	Treadmill-based gait training with foot tracking	Supervised by physiotherapists	Fatigue and sweating	Gait training (treadmill walking + dual-task intervention)	Biomedical Research Institute
Mehrabi et al [55], (2022), Canada	Mixed methods	12	≥60	None	VR	Oculus Quest 2	Handheld controllers	Remote support	Potential challenges related to nausea and motion sickness	upper-body physical activity	Home-based program
Matsangidou et al [51], (2022), Cyprus	Mixed methods	20	Mean age 77.85 (SD 11.35)	F and M	VR	VIVE Pro-Eye VR; Vive Trackers	Touch-based interaction without buttons, where objects in the VR environment are triggered by direct contact	On-site in-person supervision by physiotherapists, with real-time monitoring via an external screen	Motion sickness and anxiety noted as potential issues	Upper-limb training	Hospital setting
Matsangidou et al [52], (2023), Cyprus	Mixed methods	7	Mean age 81.14 (SD 3.84)	F and M	VR	VIVE Pro-Eye VR; Vive Trackers	Handheld controllers and Vive Trackers with gaze tracking	On-site in-person supervision by physiotherapists, with real-time monitoring via an external screen	Not reported	Upper limb rehabilitation exercises	Hospital setting

XSL•FO
RenderX

Author (year), country	Study design	Sample size	Age (years)	Sex	XR[a] type	XR hardware	XR interaction Modality	XR supervision and support	XR cybersickness / tolerability	Exercise type	Intervention setting
Eisapour et al [56], (2020), Canada	Mixed methods	6	Mean age 86.8 (SD 6.2)	F and M	VR	Oculus Rift CV1	Handheld touch controllers; audio (voice) instructions	Supervised by exercise therapists and researchers	Not reported	Upper limb exercises	Long-term care residence
Xu et al [57], (2021), China	Mixed methods study with an embedded RCT[f]	22	Mean age in the intervention group is 57.55, and in the control group is 52.82	F and M	VR	Not reported	Depth camera-based motion tracking	Supervised by physical therapists	Not reported	Lower limb functional training	Hospital setting

[a]XR: extended reality.

[b]F: female.

[c]M: male.

[d]VR: virtual reality.

[e]Quasi-RCT: quasi–randomized controlled trial.

[f]RCT: randomized controlled trial.

Quality Assessment of the Studies Included

Based on the MMAT assessment, the overall methodological quality of the included studies was acceptable.

Most studies demonstrated adequate quality across their respective methodological domains. The qualitative components generally met MMAT criteria, showing appropriate study design, data collection, and coherence between data, analysis, and interpretation.

Among quantitative studies, one randomized controlled trial met all MMAT criteria. Several nonrandomized studies showed limitations related to confounding, as confounders were not clearly accounted for in the study design or analysis.

Most mixed methods studies provided a clear rationale for using a mixed methods design and demonstrated integration between qualitative and quantitative components. However, some studies showed limited methodological transparency, reflected by “can’t tell” ratings on selected MMAT criteria.

No studies were excluded on the basis of quality appraisal. Detailed MMAT ratings for each study are provided in Multimedia Appendix 3 [51-59].

Service Design Outcomes

Service Design Tools and Strategies

Across the 9 included studies, service design strategies were predominantly grounded in participatory, co-design, and user-centered approaches. A total of 4 (44%) studies applied participatory design [52,56,57,59], 2 (22%) studies adopted co-design [51,53], 1 (11%) study used collaborative design [58], 1 (11%) study adopted user-centered design [54], and 1 (11%) study explicitly combined participatory and user-centered methods [55].

Stakeholder involvement demonstrated a multistakeholder pattern in most cases. All studies (9/9, 100%) involved older adults [51-59], 8 (89%) studies engaged rehabilitation professionals such as physical therapists, rehabilitation physicians, or kinesiologists [52-59]. Five (56%) studies involved technical and design experts [51,52,54,55,58], 4 (44%) studies included medical staff such as physicians, nurses, or paramedical professionals [51,53,54,59], and 1 (11%) studies engaged caregivers or community staff [59]. In total, 7 (78%) studies incorporated 3 or more stakeholder groups [51-55,58,59], reflecting a predominant emphasis on collaborative development across clinical, technical, and end user perspectives (Table 3).

XSL•FO
RenderX

Table 3. Service design tools and applications across service stages.

Author (year)	Service design strategy	Stakeholder groups involved	Application stage (TiSDD[a] model: research/ideation/prototyping/implementation)	Service design activities and functions
De Luca et al [59], (2025)	Participatory design	Older adults, caregivers, therapists, communities, and clinicians.	Research + Prototyping: Participatory design was used to engage older adults, caregivers, therapists, and other stakeholders to identify user needs and define goals (Research), iteratively refined the system prototype based on user feedback during development (Prototyping).	Service design activities structured adapted physical activity services for older adults in nursing homes and home settings by defining service workflows and delivery scenarios, positioning the VR[b] system as part of an integrated service rather than a standalone technology.
Sweeney et al [53], (2025)	Cocreation	Older adults with stroke, physicians, nurses, and therapists.	Implementation: The intervention was deployed, managed, and monitored in a real-world hospital ward environment.	Service design activities structured a multitechnology, group-based rehabilitation exercise service within an acute stroke unit by configuring service workflows, defining staff and patient roles, and integrating multiple rehabilitation technologies into a supervised group delivery model.
Ciemer et al [58], (2025)	Collaborative	Older adults, expert groups from the domains of Human Movement Science, Experience Design, and Game Design.	Research + Prototyping: User needs were identified through involvement of target users (Research), and the EXploVR prototype was developed and iteratively tested by a multidisciplinary team together with older adults (Prototyping).	Service design activities structured EXploVR as a supervised, community-oriented fall-prevention service by defining training workflows, staff roles, spatial and safety arrangements, and designing the intervention to support relevance to daily-life mobility and participation for older adults.
Bosch-Barceló et al [54], (2024)	User-centered design	Older adults with Parkinson disease, physiotherapists, nurses, and computing scientists.	Research + Prototyping + Implementation: Initial user needs were collected through feedback and expert interviews (Research), software was developed and validated using the Scrum method (Prototyping), and implementing the intervention in a real world setting and collecting feedback (Implementation).	Service design activities structured a supervised, technology-supported gait rehabilitation service by defining session workflows, physiotherapists and patient roles, safety arrangements, and progression rules, integrating a gamified VR environment into routine physiotherapy-led delivery.
Mehrabi et al [55], (2022)	Participatory and user-centered design	Older adults with no or mild cognitive impairment, exercise professionals, game designers, engineers from human factors and assistive technology fields, and a local VR studio.	Research + Prototyping: An extensive user-centered and participatory design process was applied, involving stakeholders including older adults, exercise professionals, game designers, and assistive technology experts (Research). The system was iteratively refined based on user feedback (Prototyping). The paper did not detail each phase separately.	Service design activities structured a remotely delivered home-based VR exercise service by defining session workflows, participant and staff roles, and remote onboarding and support procedures, enabling older adults to safely engage in a 6-week exergame intervention outside clinical settings.
Matsangidou et al [51], (2022)	Co-design	Older adults with dementia, medical and paramedical staff, Human-Computer Interaction (HCI) in health care experts.	Research + Ideation + Prototyping: Needs and barriers were identified through focus groups (Research), followed by concept generation and prioritization through workshops and brainstorming sessions (Ideation), and then rapid prototyping and testing with user feedback (Prototyping).	Service design activities structured a supervised VR-based physical training service for people living with dementia by defining session workflows, therapist-patient roles, safety and calibration procedures, and integrating VR into routine care delivery within a restricted hospital environment.
Matsangidou et al [52], (2023)	Participatory design	Older adults with dementia, physiotherapist, human-computer interaction (HCI) researcher.	Prototyping: The intervention applied the RITE rapid prototyping method, using observation notes and semistructured interviews to iteratively refine the system (Prototyping).	Service design activities structured a personalized, therapist-supported VR rehabilitation service for seated upper-body physical training for people living with dementia, by defining calibration-based exercise workflows, therapist monitoring roles, and adaptive feedback mechanisms to support safe and independent task execution in clinical settings.

XSL•FO
RenderX

Author (year)	Service design strategy	Stakeholder groups involved	Application stage (TiSDD[a] model: research/ideation/prototyping/implementation)	Service design activities and functions
Eisapour et al [56], (2020)	Participatory design	Older adults with dementia, exercise therapists, and kinesiologists.	Research + Prototyping: User needs were gathered through focus groups and observations (Research), and the prototypes were refined iteratively based on user feedback (Prototyping).	Service design activities structured dementia-friendly VR exercise delivery by co-designing with therapists and people living with dementia, iteratively refining 2 VR exercise environments, and embedding safety and accessibility supports (eg, audio guidance and individualized calibration) for seated upper-body training in long-term care.
Xu et al [57], (2021)	Participatory design	Older adults with stroke, physical therapists, and rehabilitation physicians.	Research + Prototyping + Implementation: Design inputs were collected from stroke patients, therapists, and physicians (Research); the game prototype was refined based on feedback to improve usability (Prototyping); and the final version was tested in a hospital rehabilitation ward with evaluations from multiple stakeholder perspectives (Implementation).	Service design activities structured a therapist-supervised VR service for poststroke lower-limb rehabilitation by translating stepping- and gait-training protocols into adjustable Stomp-Joy–based game tasks, and embedding role-based clinical workflows for session prescription, progression control, and safety supervision.

[a]TiSDD: This is Service Design Doing.

[b]VR: virtual reality.

Application Across Service Stages

The coverage of service design stages demonstrated apparent variation (Table 3). Seven of 9 (78%) studies applied the Research stage, identifying user needs and design elements through interviews, focus groups, or on-site observations [51,54-59]. One (11%) study addressed the Ideation stage, using workshops and brainstorming sessions to generate and prioritize concepts [51]. Eight (89%) studies engaged in the Prototyping stage, improving the usability and interaction experience of the system through rapid prototyping or RITE methods [51,52,54-59]. Three (33%) studies entered the implementation stage, deploying and managing the system in hospital, biomedical research, or rehabilitation ward settings [53,54,57].

Regarding stage coverage, 2 (22%) studies were single-stage [52,53] and focused exclusively on Implementation or Prototyping. A total of 7 (78%) studies were multistage, involving 2 or more phases [51,54-59]. Among these, 4 (44%) studies adopted a 2-stage process from the Research to Prototyping stage, enabling user needs to be identified and rapidly translated into iterative prototypes [55,56,58,59]. Two (22%) studies implemented a 3-stage combination of Research, Prototyping, and Implementation [54,57], 1 (11%) followed a 3-stage process covering the sequence of Research, Ideation, and Prototyping [51]. Notably, no study completed all 4 stages of the TiSDD model.

Across the included studies, service design activities were primarily used to configure how XR interventions functioned as services in real-world care contexts. Specifically, service design activities [51-59] were used to (1) define service workflows and delivery scenarios, (2) clarify roles and responsibilities of patients, therapists, and staff, and (3) integrate XR systems into existing rehabilitation, care, or home service structures (Table 3).

User Experience Outcomes

The included studies evaluated UX using 5 methods. Five of the 9 (56%) of studies used standardized usability questionnaires, such as System Usability Scale, The Single Ease Question, Simulator Sickness Questionnaire, Questionnaire for User Interaction Satisfaction, and NATU Quest (Assistive Technology Usability Questionnaire for people with Neurological diseases), to assess system usability and task difficulty [51,52,54,55,59]. Five of the 9 (56%) studies evaluated emotional and immersive experience through tools like Flow Short Scale, Physical Activity Enjoyment Scale, Exergame Enjoyment Questionnaire, Physical Activity Affect Scale, Visual Analogue Scale, Observed Emotion Rating Scale, and Slater-Usoh-Steed Questionnaire [51,52,55,56,58]. Two of the 9 (22%) studies captured motivation and engagement using the intrinsic motivation inventory and a 4-point engagement scale [56,57]. Two of the 9 (22%) of studies examined self-efficacy and safety perceptions using tools such as ABC-6/8, the Bandura scale, and custom safety items [55,58]. Six of the 9 (67%) studies used qualitative feedback methods, including interviews and open-ended forms [52-57].

Participants generally reported a positive UX. Five of the 9 (56%) studies reported that interventions were enjoyable and immersive, with participants frequently highlighting high levels of engagement and emotional benefits [51,52,54,56,58]. Three of the 9 (33%) of studies indicated perceived usefulness and value for rehabilitation, with self-report measures reflecting strong motivation and recognition of exercise benefits [55,57,59]. Two of the 9 (22%) studies included feedback from clinical staff who found the interventions both feasible and encouraging [53,57]. They highlighted outcomes such as increased patient confidence, improved mood, and reduced staff workload. However, 33% (3/9) of studies also pointed to negative aspects, including fatigue, discomfort from wearing head-mounted displays, and a decrease in enjoyment during cognitively demanding or extended sessions [51,54,57] (Table 4).

XSL•FO
RenderX

Table 4. User experience (UX) outcomes.

Author (year)	UX measurement method	UX feedback
De Luca et al [59], (2025)	QUIS7[a]-based questionnaire (5-point Likert)	76.7% (58/75) reported positive experience, 68.5% (51/75) were satisfied, 65.8% (49/75) found it useful.
Sweeney et al [53], (2025)	Staff interviews (qualitative)	Clinical staff considered the intervention acceptable, reporting that it improved patients' confidence, mood, and motivation for rehabilitation, while also alleviating staff workload.
Ciemer et al [58], (2025)	FKS[b], PACES-S[c], EEQ[d], ABC-6/8[e], custom questions on confidence	High levels of enjoyment (FKS, PACES, EEQ scores high and stable); participants reported the training as enjoyable, immersive, meaningful, and safe, with reduced fear of falling.
Bosch-Barceló et al [54], (2024)	SUS[f], NATU Quest[g], SSQ[h], user feedback	High satisfaction reported. participants preferred natural environments (countryside, forest), found the experience pleasant and immersive, and appreciated the dog design as reassuring and the environments as realistic. Some participants noted that fog and high speeds reduced enjoyment. Although they reported fatigue, the system was generally acceptable and motivating.
Mehrabi et al [55], (2022)	Ad hoc usability and game UX questionnaire, 5-point pictorial Likert enjoyment scale, PAAS[i], Bandura self-efficacy scale, semi-structured interviews	Positive UX: high satisfaction, usability, and immersion; high enjoyment (Likert); interviews show game is engaging and feasible; PAAS indicates positive emotions; self-efficacy supports future exercise intention.
Matsangidou et al [51], (2022)	OERS[j], VAS[k], VRSQ[l], SUS, SEQ[m]	FIVR[n] significantly increased positive affect (higher OERS scores), reduced sadness and anxiety; patients reported the training to be more enjoyable, immersive, and emotionally positive; SIVR[o] was better than conventional training but less effective than FIVR; some patients remained sensitive to wearing the HMD[p].
Matsangidou et al [52], (2023)	SUS, SEQ, OERS, VAS, task performance, and interviews	In the final prototype, significant improvements were observed in presence and usability (with SUS and SEQ scores significantly increased); participants showed increased positive emotions (pleasure) and reduced negative emotions (such as sadness, anxiety, anger); observation notes and interviews indicated that VR training elicited feelings of enjoyment, interest, and self-identification (eg, perceiving the training as more relevant to their personal context and natural to perform).
Eisapour et al [56], (2020)	Daily and weekly questionnaires (5-point Likert), 4-point engagement scale, interviews, open-ended feedback, and expert reviews.	Participants generally reported enjoying the VR experience, finding it relaxing and enjoyable, with high immersion, comfort, and usability; some expressed a preference for more challenging tasks or more variety rather than repeating the same movement.
Xu et al [57], (2021)	IMI[q] questionnaire (5 dimensions), semistructured interviews	Patients reported high scores in interest/enjoyment, perceived competence, and value/usefulness on the IMI questionnaire, indicating high engagement and perceived benefits of the game. However, it was noted that interest may decline with long-term use and that the system lacked a recording mechanism. Both patients and clinicians considered the experience pleasant and motivating.

[a]QUIS7: Questionnaire for User Interaction Satisfaction.

[b]FKS: Flow Short Scale.

[c]PACES-S: Physical Activity Enjoyment Scale.

[d]EEQ: Exergame Enjoyment Questionnaire.

[e]ABC-6/8: Activities-specific Balance Confidence Scales.

[f]SUS: System Usability Scale.

[g]NATU Quest: Assistive Technology Usability Questionnaire for people with Neurological diseases.

[h]SSQ: Simulator Sickness Questionnaire.

[i]PAAS: Physical Activity Affect Scale.

[j]OERS: Observed Emotion Rating Scale.

[k]VAS: Visual Analogue Scale.

[l]VRSQ: Slater-Usoh-Steed Questionnaire.

[m]SEQ: The Single Ease Question.

[n]FIVR: fully immersive virtual reality.

[o]SIVR: Semi Immersive VR.

[p]HMD: Head-Mounted Display.

[q]IMI: Intrinsic Motivation Inventory.

XSL•FO
RenderX

Exercise Adherence Outcomes

This review analyzed exercise adherence outcomes reported across the included studies. Given the heterogeneity in adherence definitions and measurement approaches, no conversion or standardization of adherence metrics was performed. Instead, adherence outcomes were synthesized by grouping indicators according to their underlying behavioral constructs, allowing comparison within similar categories while highlighting dimensions that remain underevaluated.

Attendance-based adherence indicators were the most consistently reported. Five of the 9 (56%) studies reported attendance-related indicators, including completion and dropout rates, with completion rates ranging from 80% to 100% and dropout rates varying between 2% and 13% [53,55-58]. In this review, references to “high adherence” were primarily derived from attendance-related indicators, as these were the most frequently and consistently reported measures across studies.

Performance-based adherence indicators were reported less consistently and captured different aspects of task execution and progression. Four of the 9 (44%) of studies monitored exercise accuracy, including task precision, movement quality, and success rates through in-game metrics or trajectory tracking systems [51,52,54,56]. Furthermore, 33% (3/9) of studies discussed performance progression across sessions, including observed improvements or declines in task success rates and response times [52-54].

Self-reported adherence indicators were the least frequently reported. One of the 9 (11%) studies assessed self-reported perceptions of adherence, with 65.8% (49/75) of participants indicating that the system supported their physical activity routines [59].

Across the 4 core adherence components synthesized in this review, including frequency, duration, intensity, and accuracy [60]. Seven of the 9 (78%) of studies clearly defined exercise frequency [51,52,54-58]. Furthermore, 89% (8/9) of studies reported the duration of exercise, including both session length and program length [51-58]. Three of the 9 (33%) studies addressed exercise intensity [55,56,59]. Four of the 9 (44%) of studies captured the accuracy of exercise performance [51,52,54,56]. None of the studies included long-term follow-up to assess sustained exercise adherence.

No studies were excluded from the review due to unclear adherence reporting; however, the variability and incompleteness of adherence measures limited cross-study comparability and constrained conclusions regarding long-term adherence (Table 5).

XSL•FO
RenderX

Table 5. Exercise adherence outcomes.

Author (year)	Exercise protocol (planned dose)	Exercise adherence indicators	Exercise adherence results	Follow-up
De Luca et al [59], (2025)	6-week program	IPAQ-SF MET[a] value (quantitative), IPAQ-SF[b] activity level classification (inactive/active), self-reported improvement in physical activity adherence.	No significant improvement in overall physical activity levels; 9% (7/75) of participants improved from inactive to sufficiently active; 65.8% (49/75) agreed and 9.6% (7/75) strongly agreed that VR2Care improved adherence to physical activity.	Not reported
Sweeney et al [53], (2025)	Duration not specified	Session attendance, training duration, task completion rate, performance progression.	On average, each patient completed 9.1 sessions (SD 7.4) during hospitalization, with a total rehabilitation time of 508.4 minutes (SD 521.2) and an average of 52.0 minutes per session (SD 15.7). Patients achieved an additional 52 minutes of rehabilitation per day prior to discharge. Dropout rate was low (1/60, 2%).	Not reported
Ciemer et al [58], (2025)	3-week program; 2 sessions per week; 60 minutes per session; total of 6 sessions	Adherence was tracked through an attendance log recorded at each session.	A total of 32 participants completed the study, with 16 in the intervention group and 16 in the control group. A total of 6 participants dropped out during the study (intervention group: n=4; control group: n=2), resulting in an overall dropout rate of only 13% (4/32).	Not reported
Bosch-Barceló et al [54], (2024)	3 sessions; 1 session per week; each session included treadmill warm-up and 4 bouts of gait training with rest intervals; total session duration approximately 20-21 minutes	Obstacle avoidance performance (success rate in each bout across 3 different difficulty environments).	Obstacle avoidance performance improved during the training phases, with better performance after familiarization on Day 1; performance slightly decreased on Days 2 and 3 as difficulty increased, but overall remained satisfactory.	Not reported
Mehrabi et al [55], (2022)	6 weeks; 3 sessions per week; 15-20 minutes per session	Session adherence rate; gameplay metrics.	The authors anticipated an adherence rate of ≥80%. Gameplay performance metrics were automatically recorded to assess participation.	Not reported
Matsangidou et al [51], (2022)	3 experimental sessions; each lasting approximately 30 minutes	Task accuracy (number of correctly performed repetitions out of 10), task independence (number of therapist interventions required), completion time (seconds), movement quality (match with baseline movements).	Under the FIVR[c] system, patients demonstrated higher task accuracy, greater independence, and better movement quality; the time required for simple movements (Side Arm Raise) was shorter, and although the time for complex movements (Overhead Press) was longer, accuracy improved.	Not reported
Matsangidou et al [52], (2023)	3 iterations of RITE[d] testing; each iteration approximately 30 minutes, including up to 20 minutes of VR[e] exposure	Task completion rate, task independence score, execution time, movement trajectory error.	Over the 3 prototype iterations, task completion and independence significantly improved, execution time decreased, and movement trajectories became closer to the target, indicating progressively improved exercise adherence.	Not reported
Eisapour et al [56], (2020)	3-week program; 5 sessions per week; ≈20 minutes per session	Completion rate (whether all exercises were completed), range of motion (reach), movement distance, velocity.	All participants completed all scheduled VR exercise sessions (6/6, 100% completion); exercise performance (range of motion, distance, velocity) was comparable to therapist-led exercise, with no significant differences; participants perceived VR exercise workload as adequate and comparable to therapist-led sessions.	Not reported
Xu et al [57], (2021)	2-week program; 5 sessions per week; VR training (30 minutes) combined with conventional physical therapy; total of 10 sessions	Intervention attendance/dropout rate; number of completed sessions during the intervention period.	Of the 24 participants, 2 withdrew—one due to a fall during training, without injury, and one for personal reasons.	Not reported

[a]IPAQ-SF MET: International Physical Activity Questionnaire Short Form - Metabolic Equivalent of Task.

[b]IPAQ-SF: International Physical Activity Questionnaire Short Form.

[c]FIVR: fully immersive virtual reality.

[d]RITE: Rapid Iterative Testing and Evaluation.

XSL•FO
RenderX

[e]VR: virtual reality.

Discussion

Principal Results

This review synthesized evidence from 9 XR-based exercise interventions targeting older adults and revealed a field that was promising yet constrained by structural and methodological limitations. Most studies grounded their service design strategies in participatory, co-design, and user-centered approaches [51-57,59], frequently combined with iterative prototyping [51,54-59]. However, service design processes were often incomplete, with most studies concentrating on early research and prototyping stages, showing limited engagement in ideation [51] and progression to implementation. Only 3 studies progressed to the Implementation stage [53,54,57]. Although multiple stakeholders were involved and participatory principles were widely adopted, their contributions were largely confined to early-stage design decisions and usability refinement, rather than extended to shaping long-term service integration or delivery models. This uneven coverage across service stages constrained the translation of interventions into sustained service offerings. When interpreted through behavior change frameworks such as Capability, Opportunity, Motivation–Behavior [61] and behavior change techniques [62], the limited continuity of service design across later stages helps explain why sustained exercise behavior was rarely examined. While early-stage co-design and prototyping supported initial engagement and motivation, fewer interventions addressed opportunity-related factors, such as integration into everyday contexts, reinforcement strategies, and long-term support mechanisms, which are essential for maintaining behavior over time.

UX outcomes were generally positive. Participants reported pleasant experiences, high levels of immersion, and increased motivation in questionnaires, usability evaluations, and interviews. However, some studies also noted negative adverse effects, such as fatigue, headset discomfort, and decreased enjoyment during challenging or long tasks [51,54,57]. Overall, most evaluations relied on subjective reports, with limited use of objective measures, such as physiological or behavioral data. This reliance on self-reported short-term experience indicators limits the extent to which positive UX can be interpreted as evidence of sustained engagement. While enjoyment and immersion appear effective in supporting initial participation, the lack of longitudinal and objective measures constrains understanding of how these experiences translate into continued exercise behavior over time.

Most studies reported high levels of exercise adherence, though inconsistently across dimensions, despite the use of a 4D adherence framework [60] in our synthesis; few studies comprehensively reported across all dimensions. Frequency and duration were the most reported measures of exercise adherence; however, few studies compared them with established exercise guidelines. Most studies conducted minimal monitoring of exercise intensity, and only 1 applied validated physiological measurement. While some VR systems included metrics related to movement accuracy, the studies did not integrate these metrics into the formal outcome framework for exercise adherence. This fragmentation of adherence dimensions impeded cross-study comparison and masked the contribution of specific design features to behavioral engagement. Most studies assessed exercise adherence using non-standardized, short-term measures. Attendance and task completion were the most commonly used surrogate measures of adherence; some studies also reported adherence through in-game performance metrics, performance progression over time, or self-reported perceptions of adherence, but none adopted a standardized long-term assessment framework. As a result, the consistently high adherence rates reported across studies should be interpreted cautiously, as they primarily reflect short-term participation rather than sustained exercise continuity. The absence of standardized, multidimensional, and longitudinal adherence assessments limits insight into whether engagement persisted beyond the intervention period.

Notably, 8 studies conducted interventions in hospital, laboratory, or long-term care settings [51-54,56-59], and 1 study conducted its intervention at home [55], with no studies implementing community-based deployment, highlighting a critical gap for real-world scalability. The traditional health management model mainly relied on hospital or home-centered interventions, and the increasing demand for population aging and chronic disease management posed greater challenges to this model [63-65]. In contrast, community-based health management was considered a sustainable strategy that could provide convenient, long-term services and promote healthy behaviors among older adults [66,67]. Community and group-based programs provide social interaction, emotional support, and a sense of belonging, which reduce isolation and strengthen psychological stability [16,22]. Community XR interventions also faced some challenges in resource sharing and management. Community-based XR exercise often requires multiple participants to share equipment, space, and network resources. However, most community spaces lacked the specialized infrastructure necessary to support immersive exercise programs, including adequate physical space, stable power supply, network connectivity, and safety facilities [68]. This created unresolved issues such as shared use time arrangements, equipment maintenance, and technical assistance training for community staff [68,69]. For community XR projects to be truly implemented, they also need the support of policies and management mechanisms [70]. For example, by optimizing community resource allocation and conducting technical personnel training, practical problems such as equipment sharing and insufficient infrastructure can be addressed [71].

The dominance of controlled environments in the reviewed studies, therefore, limits the interpretation of continuity and scalability outcomes. While such settings support feasibility testing and supervised delivery, they do not reflect the operational conditions under which older adults engage in exercise in everyday community contexts, making it difficult to assess whether XR-based interventions can be sustained beyond experimental or clinical settings.

XSL•FO
RenderX

Comparison With Prior Work

The findings of this study aligned with broader existing evidence. With the application of technology-assisted interventions in the field of rehabilitation and exercise for older adults, participatory methods and user-centered design principles were gradually integrated into the entire process of intervention development [72,73]. This trend also appeared in VR rehabilitation and digital health projects, where researchers generally emphasized the widespread application of participatory design and rapid prototyping design [35]. However, our synthesis identified a critical difference: early creative ideation and comprehensive multi-stage service design remained underused. Most studies adopted fragmented design processes without forming a closed loop between service design and adherence monitoring, which limited the ability to anticipate long-term engagement barriers and optimize real-world implementation.

Several studies confirmed that immersive VR interventions provided positive feedback to older adults, including heightened feelings of pleasure, increased willingness to participate in activities, and improved emotional well-being [74,75]. However, some studies reported side effects, such as fatigue, mild physical discomfort, or sensitivity to the headset after the intervention [76,77].

High short-term adherence appeared consistent with prior findings on gamified VR engagement [26,78]. However, this adherence pattern likely resulted from 2 limitations: first, the intervention duration was insufficient, and second, the adherence assessment did not form a standardized indicator, but rather participants achieved a real and sustained change in behavior. In interdisciplinary research practice, standardized compliance assessment played a crucial role. It bridged definitional gaps across disciplines, enabled unified evaluation criteria, and offered standardized measurement tools for longitudinal tracking [60,79]. Furthermore, due to the lack of follow-up data, researchers could not determine whether such adherence persisted beyond the short term [80].

However, a critical underexplored dimension in prior work and our synthesis concerned the setting of implementation. Existing research suggested that implementing XR intervention in community settings could support social interaction, emotional regulation, and behavioral improvement [81-83]. However, most studies focused on hospitals, laboratories, and care facilities. This narrow setting focus weakened the ecological validity of research findings and hindered a comprehensive evaluation of the long-term sustainability of interventions. Integrating XR into community programs could have improved accessibility, peer support, and emotional connection; however, this remained an overlooked area in current research.

Limitations and Future Research

This review had some limitations. First, the number of studies included in the analysis was relatively small, and the sample size of most individual studies was also small. More importantly, the studies showed pronounced heterogeneity in design, baseline characteristics of participants, and intervention implementation processes. These differences not only reduced the statistical power of the overall analysis but also limited the generalizability of the research conclusions. In addition, restricting inclusion to studies with accessible full texts may have excluded relevant work based on access rather than content, introducing potential selection bias; gray literature and trial registries were not systematically searched, which may have resulted in the omission of unpublished or ongoing studies. Second, none of the included studies established longitudinal follow-up links, which made it impossible to determine whether the short-term participation adherence translated into long-term behavioral change or functional improvement. Third, regarding intervention settings, most studies focused on hospitals, care facilities, or laboratory environments. In contrast, empirical evidence from community or home settings remained relatively limited, which also weakened the applicability of the findings to real-life scenarios. Finally, although this review adopted an XR-oriented scope, all included studies used VR interventions, reflecting the current concentration of evidence rather than an intentional focus on VR, and exploration of AR or MR modes remained relatively limited. Studies applying AR or MR technologies were identified in the broader literature and have demonstrated positive physical or cognitive outcomes among older adults [84,85]; however, these approaches have not yet been systematically examined within service design-oriented frameworks. The limitations of this technology exploration made it difficult to reveal the comparative advantages of different XR methods in terms of accessibility, usability, and situational adaptability, which affected the comprehensive understanding of the overall application value of XR technology.

To address the shortcomings of current research, future studies should adopt standardized, multidimensional adherence assessment metrics [60]. By establishing unified assessment dimensions and criteria, these metrics could resolve inconsistencies caused by variations in measurement and provide a stable foundation for long-term follow-up studies. Exploration of community-based interventions should be strengthened, with designs that closely mirror real-world exercise environments to address contextual barriers to adaptation effectively. Furthermore, service design tools such as user journey maps should be incorporated into the intervention process, leveraging their visualization capabilities to identify potential dropout risk points at each stage of the intervention, providing a concrete reference for iterative optimization of intervention plans. In parallel, multimodal tracking systems should be developed that integrate physiological data collected by wearable devices, participant behavioral logs, and in-game performance metrics. These systems should generate adaptive feedback to provide participants with more personalized support for interventions. Finally, future research should expand its technological scope by exploring AR and MR modalities and evaluating their potential for more inclusive and ecologically valid exercise interventions.

Conclusions

This review systematically evaluated how XR-based exercise interventions for older adults incorporated service design strategies to support exercise adherence. Although participatory design and gamified VR experiences have been widely used in related research, most studies have remained fragmented and

XSL•FO
RenderX

short-term. There were clear deficiencies in both the systematic integration of multistage service design processes and the effective incorporation of standardized compliance frameworks; a comprehensive research practice system had not yet been established. Research findings indicated generally high levels of exercise adherence; however, conclusions varied across studies, and few focused on the intensity of adherence, long-term maintenance, or the ecological validity of the findings. Further analysis from the perspectives of intervention setting and technology application revealed that most existing interventions focused on hospital or laboratory settings, whereas the exploration of community settings, as well as the development of other XR modalities such as AR and MR, was relatively limited. To address this gap, future XR intervention research should integrate structured service design tools throughout the intervention lifecycle, develop standardized adherence metrics, establish multidimensional adherence tracking mechanisms, and extend deployment to real-life settings. By promoting synergy among behavioral engagement, service continuity, and situational adaptability, the effectiveness of XR intervention in helping older adults build sustainable exercise habits, while also providing an empirical reference for the innovative development of inclusive digital health services.

Acknowledgments

The authors would like to thank the Library of the University of Southampton for assistance with scientific databases.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA checklist.
[DOCX File , 272 KB - aging_v9i1e86595_app1.docx]

Multimedia Appendix 2
Search strategy.
[DOCX File , 20 KB - aging_v9i1e86595_app2.docx]

Multimedia Appendix 3
Quality Assessment of the Studies Included.
[DOCX File , 34 KB - aging_v9i1e86595_app3.docx]

References

1. Gaziano TA, Bitton A, Anand S, Abrahams-Gessel S, Murphy A. Growing epidemic of coronary heart disease in low- and middle-income countries. Curr Probl Cardiol 2010;35(2):72-115 [FREE Full text] [doi: 10.1016/j.cpcardiol.2009.10.002] [Medline: 20109979]
2. Yang W, Lu J, Weng J, Jia W, Ji L, Xiao J, China National Diabetes Metabolic Disorders Study Group. Prevalence of diabetes among men and women in China. N Engl J Med 2010;362(12):1090-1101. [doi: 10.1056/NEJMoa0908292] [Medline: 20335585]
3. Doblhammer G, Hoffmann R, Muth E, Westphal C. A systematic literature review of studies analyzing the effects of sex, age, education, marital status, obesity, and smoking on health transitions. Acta Crystallogr B 2005;20:37-64 [FREE Full text] [doi: 10.1107/S010876810501726X] [Medline: 16041089]
4. Prince MJ, Wu F, Guo Y, Gutierrez Robledo LMG, O'Donnell M, Sullivan R, et al. The burden of disease in older people and implications for health policy and practice. Lancet 2015;385(9967):549-562. [doi: 10.1016/S0140-6736(14)61347-7]
5. van Alphen HJM, Volkers KM, Blankevoort CG, Scherder EJA, Hortobágyi T, van Heuvelen MJG. Older adults with dementia are sedentary for most of the day. PLoS One 2016;11(3):e0152457 [FREE Full text] [doi: 10.1371/journal.pone.0152457] [Medline: 27031509]
6. Zhou B, Lu Y, Hajifathalian K, Bentham J. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants. Lancet 2016;387(10027):1513-1530 [FREE Full text] [doi: 10.1016/S0140-6736(16)00618-8] [Medline: 27061677]
7. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract 2022;183:109119 [FREE Full text] [doi: 10.1016/j.diabres.2021.109119] [Medline: 34879977]
8. Alkhodary AA, Syed Mohamed A, Ismail A, Nur AM, Shahar S. The economic burden of frailty among elderly people: a review of the current literature. Malays J Public Health Med 2020;20(2):224-232. [doi: 10.37268/mjphm/vol.20/no.2/art.304]

XSL•FO
RenderX

9. Picco L, Achilla E, Abdin E, Chong SA, Vaingankar JA, McCrone P, et al. Economic burden of multimorbidity among older adults: impact on healthcare and societal costs. BMC Health Serv Res 2016;16:173 [FREE Full text] [doi: 10.1186/s12913-016-1421-7] [Medline: 27160080]
10. Figueroa JF, Papanicolas I, Riley K, Abiona O, Arvin M, Atsma F, et al. International comparison of health spending and utilization among people with complex multimorbidity. Health Serv Res 2021;56 Suppl 3(Suppl 3):1317-1334 [FREE Full text] [doi: 10.1111/1475-6773.13708] [Medline: 34350586]
11. Nystoriak MA, Bhatnagar A. Cardiovascular effects and benefits of exercise. Front Cardiovasc Med 2018;5:135 [FREE Full text] [doi: 10.3389/fcvm.2018.00135] [Medline: 30324108]
12. Fiuza-Luces C, Santos-Lozano A, Joyner M, Carrera-Bastos P, Picazo O, Zugaza JL, et al. Exercise benefits in cardiovascular disease: beyond attenuation of traditional risk factors. Nat Rev Cardiol 2018;15(12):731-743. [doi: 10.1038/s41569-018-0065-1] [Medline: 30115967]
13. D'Onofrio G, Kirschner J, Prather H, Goldman D, Rozanski A. Musculoskeletal exercise: its role in promoting health and longevity. Prog Cardiovasc Dis 2023;77:25-36. [doi: 10.1016/j.pcad.2023.02.006] [Medline: 36841491]
14. Papalia GF, Papalia R, Diaz Balzani LA, Torre G, Zampogna B, Vasta S, et al. The effects of physical exercise on balance and prevention of falls in older people: a systematic review and meta-analysis. J Clin Med 2020;9(8):2595 [FREE Full text] [doi: 10.3390/jcm9082595] [Medline: 32796528]
15. Mikkelsen K, Stojanovska L, Polenakovic M, Bosevski M, Apostolopoulos V. Exercise and mental health. Maturitas 2017;106:48-56. [doi: 10.1016/j.maturitas.2017.09.003] [Medline: 29150166]
16. Meredith S, Cox N, Ibrahim K, Higson J, McNiff J, Mitchell S, et al. Factors that influence older adults' participation in physical activity: a systematic review of qualitative studies. Age Ageing 2023;52(8) [FREE Full text] [doi: 10.1093/ageing/afad145] [Medline: 37595070]
17. Picorelli AMA, Pereira LSM, Pereira DS, Felício D, Sherrington C. Adherence to exercise programs for older people is influenced by program characteristics and personal factors: a systematic review. J Physiother 2014;60(3):151-156 [FREE Full text] [doi: 10.1016/j.jphys.2014.06.012] [Medline: 25092418]
18. Jancey J, Lee A, Howat P, Clarke A, Wang K, Shilton T. Reducing attrition in physical activity programs for older adults. J Aging Phys Act 2007;15(2):152-165. [doi: 10.1123/japa.15.2.152] [Medline: 17556782]
19. O'Neil-Pirozzi TM, Cattaneo G, Solana-Sánchez J, Gomes-Osman J, Pascual-Leone A. The importance of motivation to older adult physical and cognitive exercise program development, initiation, and adherence. Front Aging 2022;3:773944 [FREE Full text] [doi: 10.3389/fragi.2022.773944] [Medline: 35821853]
20. Herbert P. Promoting exercise in older people to support healthy ageing. Nurs Stand 2022;37(4):46-50. [doi: 10.7748/ns.2022.e11787] [Medline: 35156355]
21. Valenzuela T, Okubo Y, Woodbury A, Lord SR, Delbaere K. Adherence to technology-based exercise programs in older adults: a systematic review. J Geriatr Phys Ther 2018;41(1):49-61. [doi: 10.1519/JPT.0000000000000095] [Medline: 27362526]
22. Room J, Hannink E, Dawes H, Barker K. What interventions are used to improve exercise adherence in older people and what behavioural techniques are they based on? a systematic review. BMJ Open 2017;7(12):e019221 [FREE Full text] [doi: 10.1136/bmjopen-2017-019221] [Medline: 29247111]
23. Schutzer KA, Graves BS. Barriers and motivations to exercise in older adults. Prev Med 2004;39(5):1056-1061. [doi: 10.1016/j.ypmed.2004.04.003] [Medline: 15475041]
24. Baker S, Waycott J, Robertson E, Carrasco R, Neves BB, Hampson R, et al. Evaluating the use of interactive virtual reality technology with older adults living in residential aged care. Inf Process Manag 2020;57(3):102105. [doi: 10.1016/j.ipm.2019.102105]
25. Brimelow RE, Dawe B, Dissanayaka N. Preliminary research: virtual reality in residential aged care to reduce apathy and improve mood. Cyberpsychol Behav Soc Netw 2020;23(3):165-170. [doi: 10.1089/cyber.2019.0286]
26. Shah SHH, Karlsen AST, Solberg M, Hameed IA. A social VR-based collaborative exergame for rehabilitation: codesign, development and user study. Virtual Real 2022:1-18 [FREE Full text] [doi: 10.1007/s10055-022-00721-8] [Medline: 36465891]
27. Mirza RA, Yaqoob I. Effects of combined aerobic and virtual reality-based cognitive training on 76 years old diabetic male with mild cognitive impairment. J Coll Physicians Surg Pak 2018;28(9):S210-S212. [doi: 10.29271/jcpsp.2018.09.S210] [Medline: 30173701]
28. Lee Y, Hong J, Hur M, Seo E. Effects of virtual reality exercise program on blood glucose, body composition, and exercise immersion in patients with type 2 diabetes. Int J Environ Res Public Health 2023;20(5):4178 [FREE Full text] [doi: 10.3390/ijerph20054178] [Medline: 36901191]
29. Li Y, Wilke C, Shiyanov I, Muschalla B. Impact of virtual reality-based group activities on activity level and well-being among older adults in nursing homes: longitudinal exploratory study. JMIR Serious Games 2024;12:e50796 [FREE Full text] [doi: 10.2196/50796] [Medline: 38551635]
30. Lin C, Lee C, Lally D, Coughlin J. Impact of virtual reality (VR) experience on older adults' well-being. : Springer; 2018 Presented at: Human Aspects of IT for the Aged Population. Applications in Health, Assistance, and Entertainment; July 15–20, 2018; Cham p. 89-100. [doi: 10.1007/978-3-319-92037-5_8]

XSL•FO
RenderX

31. Bermúdez IBS, Avelino J, Bernardino A, Cameirão M, Muñoz J. Development and validation of a mixed reality exergaming platform for fitness training of older adults. In: Everyday Virtual and Augmented Reality. London (Global HQ), Berlin: Springer; 2023:119-145.
32. Blomqvist S, Seipel S, Engström M. Using augmented reality technology for balance training in the older adults: a feasibility pilot study. BMC Geriatr 2021;21(1):144 [FREE Full text] [doi: 10.1186/s12877-021-02061-9] [Medline: 33637043]
33. Chung OS, Robinson T, Johnson AM, Dowling NL, Ng CH, Yücel M, et al. Implementation of therapeutic virtual reality into psychiatric care: clinicians' and service managers' perspectives. Front Psychiatry 2021;12:791123 [FREE Full text] [doi: 10.3389/fpsyt.2021.791123] [Medline: 35058823]
34. Neira-Tovar L, Rodríguez IC, Salazar FG. A method to improve the design of virtual reality games in healthcare applied to increase physical activity in patients with type 2 diabetes mellitus. Appl Sci 2022;13(1):50. [doi: 10.3390/app13010050]
35. Lorenz EA, Bråten Støen A, Lie Fridheim M, Alsos OA. Design recommendations for xr-based motor rehabilitation exergames at home. Front Virtual Real 2024;5:1340072. [doi: 10.3389/frvir.2024.1340072]
36. Leonardi C, Mennecozzi C, Not E, Pianesi F, Zancanaro M. Designing a familiar technology for elderly people. Gerontechnology 2008;7(2):151. [doi: 10.4017/gt.2008.07.02.088.00]
37. Kouijzer MMTE, Kip H, Bouman YHA, Kelders SM. Implementation of virtual reality in healthcare: a scoping review on the implementation process of virtual reality in various healthcare settings. Implement Sci Commun 2023;4(1):67 [FREE Full text] [doi: 10.1186/s43058-023-00442-2] [Medline: 37328858]
38. Margrett JA, Ouverson KM, Gilbert SB, Phillips LA, Charness N. Older adults use of extended reality: a systematic review. Front Virtual Real 2022;2:760064. [doi: 10.3389/frvir.2021.760064]
39. Patrício L, Sangiorgi D, Mahr D, Čaić M, Kalantari S, Sundar S. Leveraging service design for healthcare transformation: toward people-centered, integrated, and technology-enabled healthcare systems. J Serv Manag 2020;31(5):889-909. [doi: 10.1108/josm-11-2019-0332]
40. Stickdorn M, Hormess ME, Lawrence A, Schneider J. This is service design doing. Santa Rosa, CA, USA: O'Reilly Media, Inc; 2018:569.
41. Simonse L, Albayrak A, Starre S. Patient journey method for integrated service design. Des Health 2019;3(1):82-97. [doi: 10.1080/24735132.2019.1582741]
42. Sijm-Eeken M, Zheng J, Peute L. Towards a lean process for patient journey mapping ? a case study in a large academic setting. Amsterdam, The Netherlands: IOS Press; 2020:1071-1075.
43. Koo Y, Kim E, Nam I. Development of a communication platform for patients with head and neck cancer for effective information delivery and improvement of doctor-patient relationship: application of treatment journey-based service blueprint. BMC Med Inform Decis Mak 2024;24(1):81 [FREE Full text] [doi: 10.1186/s12911-024-02477-4] [Medline: 38509511]
44. Mager B, Oertzen A, Vink J. Co-creation in health services through service design. Cham, Switzerland: Springer Cham; 2022:497-510.
45. James H, Buffel T. Co-research with older people: a systematic literature review. Ageing Soc 2023;43(12):2930-2956. [doi: 10.1017/s0144686x21002014]
46. Hewko M, Gagnon Shaigetz V, Smith MS, Kohlenberg E, Ahmadi P, Hernandez Hernandez ME, et al. Considering theory-based gamification in the co-design and development of a virtual reality cognitive remediation intervention for depression (bWell-D): mixed methods study. JMIR Serious Games 2025;13:e59514. [doi: 10.2196/59514]
47. Bitner MJ, Ostrom AL, Morgan FN. Service blueprinting: a practical technique for service innovation. Calif Manag Rev 2008;50(3):66-94. [doi: 10.2307/41166446]
48. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The prisma 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
49. Proposed working definition of an older person in Africa for the MDS project : definition of an older or elderly person. CiNii. 2013. URL: https://cir.nii.ac.jp/crid/1573668924062064128 [accessed 2026-09-08]
50. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The mixed methods appraisal tool (MMAT) version 2018 for information professionals and researchers. Educ Inform 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
51. Matsangidou M, Frangoudes F, Hadjiaros M, Schiza E, Neokleous KC, Papayianni E, et al. “Bring me sunshine, bring me (physical) strength”: the case of dementia. designing and implementing a virtual reality system for physical training during the COVID-19 pandemic. Int J Hum Comput Stud 2022;165:102840. [doi: 10.1016/j.ijhcs.2022.102840]
52. Matsangidou M, Frangoudes F, Schiza E, Neokleous KC, Papayianni E, Xenari K, et al. Participatory design and evaluation of virtual reality physical rehabilitation for people living with dementia. Virtual Real 2022;27(1):421-438. [doi: 10.1007/s10055-022-00639-1]
53. Sweeney G, Boyd F, Keogh M, Lyczba P, Forrest E, Rowe P, et al. A technology-enriched approach to increasing rehabilitation dose after stroke: clinical feasibility study. Clin Rehabil 2025;39(6):740-749 [FREE Full text] [doi: 10.1177/02692155251333542] [Medline: 40247700]
54. Bosch-Barceló P, Masbernat-Almenara M, Martínez-Navarro O, Tersa-Miralles C, Pakarinen A, Fernández-Lago H. A gamified virtual environment intervention for gait rehabilitation in parkinson's disease: co-creation and feasibility study. J Neuroeng Rehabil 2024;21(1):107 [FREE Full text] [doi: 10.1186/s12984-024-01399-6] [Medline: 38915103]

XSL•FO
RenderX

55. Mehrabi S, Muñoz JE, Basharat A, Boger J, Cao S, Barnett-Cowan M, et al. Immersive virtual reality exergames to promote the well-being of community-dwelling older adults: protocol for a mixed methods pilot study. JMIR Res Protoc 2022;11(6):e32955 [FREE Full text] [doi: 10.2196/32955] [Medline: 35700014]
56. Eisapour M, Cao S, Boger J. Participatory design and evaluation of virtual reality games to promote engagement in physical activity for people living with dementia. J Rehabil Assist Technol Eng 2020;7:2055668320913770 [FREE Full text] [doi: 10.1177/2055668320913770] [Medline: 32499921]
57. Xu Y, Tong M, Ming W, Lin Y, Mai W, Huang W, et al. A depth camera-based, task-specific virtual reality rehabilitation game for patients with stroke: pilot usability study. JMIR Serious Games 2021;9(1):e20916 [FREE Full text] [doi: 10.2196/20916] [Medline: 33759795]
58. Ciemer C, Schott N, Klotzbier TJ, Ghellal S. Effectiveness, enjoyment, and meaningfulness of a virtual reality gait-based fall prevention exergame in community-dwelling healthy older adults: an interdisciplinary pilot study. Front Psychol 2025;16:1610377 [FREE Full text] [doi: 10.3389/fpsyg.2025.1610377] [Medline: 40567879]
59. De Luca V, Qbilat M, Cuomo A, Bianco A, Cesaroni F, Lanari C, et al. Virtual reality solution to promote adapted physical activity in older adults: outcomes from VR2Care project exploratory study. Front Public Health 2025;13:1584406 [FREE Full text] [doi: 10.3389/fpubh.2025.1584406] [Medline: 40433492]
60. Frost R, Levati S, McClurg D, Brady M, Williams B. What adherence measures should be used in trials of home-based rehabilitation interventions? a systematic review of the validity, reliability, and acceptability of measures. Arch Phys Med Rehabil 2017;98(6):1241-1256.e1245 [FREE Full text] [doi: 10.1016/j.apmr.2016.08.482] [Medline: 27702555]
61. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
62. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
63. Anderson R. Patient empowerment and the traditional medical model: a case of irreconcilable differences? Diabetes Care 1995;18(3):412-415. [doi: 10.2337/diacare.18.3.412]
64. Morales-Asencio J, Gonzalo-Jiménez E, Martin-Santos F, Morilla-Herrera J, Celdráan-Mañas M, Carrasco AM, et al. Effectiveness of a nurse-led case management home care model in primary health care. a quasi-experimental, controlled, multi-centre study. BMC Health Serv Res 2008;8:193 [FREE Full text] [doi: 10.1186/1472-6963-8-193] [Medline: 18811927]
65. Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. mHealth 2017;3:38 [FREE Full text] [doi: 10.21037/mhealth.2017.08.07] [Medline: 29184890]
66. Buffel T. Social research and co-production with older people: developing age-friendly communities. J Aging Stud 2018;44:52-60. [doi: 10.1016/j.jaging.2018.01.012] [Medline: 29502790]
67. Farrance C, Tsofliou F, Clark C. Adherence to community based group exercise interventions for older people: a mixed-methods systematic review. Prev Med 2016;87:155-166 [FREE Full text] [doi: 10.1016/j.ypmed.2016.02.037] [Medline: 26921655]
68. Sarkar U, Lee JE, Nguyen KH, Lisker S, Lyles CR. Barriers and facilitators to the implementation of virtual reality as a pain management modality in academic, community, and safety-net settings: qualitative analysis. J Med Internet Res 2021;23(9):e26623. [doi: 10.2196/26623]
69. Levac DE, Miller PA. Integrating virtual reality video games into practice: clinicians' experiences. Physiother Theory Pract 2013;29(7):504-512. [doi: 10.3109/09593985.2012.762078] [Medline: 23362843]
70. Lewin S, Lehmann U, Perry HB. Community health workers at the dawn of a new era: 3. programme governance. Health Res Policy Sys 2021;19(S3):129. [doi: 10.1186/s12961-021-00749-3]
71. Chen AT, Smith DO, Ojikutu BO, Auerbach J. The community as a full partner: a new model for public health. Health Aff (Millwood) 2024;43(6):805-812. [doi: 10.1377/hlthaff.2024.00033] [Medline: 38830164]
72. Kerr A, Grealy M, Slachetka M, Wodu CO, Sweeney G, Boyd F, et al. A participatory model for cocreating accessible rehabilitation technology for stroke survivors: user-centered design approach. JMIR Rehabil Assist Technol 2024;11:e57227 [FREE Full text] [doi: 10.2196/57227] [Medline: 39177660]
73. Muñoz J, Mehrabi S, Li Y, Basharat A, Middleton LE, Cao S, et al. Immersive virtual reality exergames for persons living with dementia: user-centered design study as a multistakeholder team during the COVID-19 pandemic. JMIR Serious Games 2022;10(1):e29987 [FREE Full text] [doi: 10.2196/29987] [Medline: 35044320]
74. Moyle W, Jones C, Dwan T, Petrovich T. Effectiveness of a virtual reality forest on people with dementia: a mixed methods pilot study. Gerontologist 2018;58(3):478-487. [doi: 10.1093/geront/gnw270] [Medline: 28329867]
75. Villada Castillo JF, Montoya Vega MF, Muñoz Cardona JE, Lopez D, Quiñones L, Henao Gallo OA, et al. Design of virtual reality exergames for upper limb stroke rehabilitation following iterative design methods: usability study. JMIR Serious Games 2024;12:e48900 [FREE Full text] [doi: 10.2196/48900] [Medline: 38206670]
76. Goh L, Allen NE, Ahmadpour N, Ehgoetz Martens KA, Song J, Clemson L, et al. A video self-modeling intervention using virtual reality plus physical practice for freezing of gait in parkinson disease: feasibility and acceptability study. JMIR Form Res 2021;5(11):e28315 [FREE Full text] [doi: 10.2196/28315] [Medline: 34730537]

XSL•FO
RenderX

77. McTear M, Jokinen K, Alam MM, Saleem Q, Napolitano G, Szczepaniak F, et al. Interaction with a virtual coach for active and healthy ageing. Sensors (Basel) 2023;23(5):2748 [FREE Full text] [doi: 10.3390/s23052748] [Medline: 36904957]
78. Kruse L, Karaosmanoglu S, Rings S, Ellinger B, Steinicke F. Enabling immersive exercise activities for older adults: a comparison of virtual reality exergames and traditional video exercises. Societies 2021;11(4):134. [doi: 10.3390/soc11040134]
79. Mahmood A, Nayak P, Deshmukh A, English C, N M, Solomon M J, et al. Measurement, determinants, barriers, and interventions for exercise adherence: a scoping review. J Bodyw Mov Ther 2023;33:95-105. [doi: 10.1016/j.jbmt.2022.09.014]
80. Kaplan RM, Stone AA. Bringing the laboratory and clinic to the community: mobile technologies for health promotion and disease prevention. Annu Rev Psychol 2013;64:471-498. [doi: 10.1146/annurev-psych-113011-143736] [Medline: 22994919]
81. Zeng N, Pope Z, Lee J, Gao Z. Virtual reality exercise for anxiety and depression: a preliminary review of current research in an emerging field. J Clin Med 2018;7(3):42 [FREE Full text] [doi: 10.3390/jcm7030042] [Medline: 29510528]
82. Benham S, Kang M, Grampurohit N. Immersive virtual reality for the management of pain in community-dwelling older adults. OTJR 2019;39(2):90-96. [doi: 10.1177/1539449218817291] [Medline: 30595096]
83. Wang L. Effectiveness of virtual reality exercise for functional fitness in community-dwelling older adults: a 12-week follow-up study. SAGE Open 2023;13(4):21582440231218515. [doi: 10.1177/21582440231218515]
84. An Y, Min S, Park C. Mixed reality-based physical therapy in older adults with sarcopenia: preliminary randomized controlled trial. JMIR Serious Games 2025;13:e76357 [FREE Full text] [doi: 10.2196/76357] [Medline: 40749218]
85. Park TS, Shin M. Effectiveness of an exercise program for older adults using an augmented reality exercise platform: a pilot study. Ann Geriatr Med Res 2023;27(1):73-79 [FREE Full text] [doi: 10.4235/agmr.23.0016] [Medline: 36945873]

Abbreviations

AR: augmented reality
MMAT: Mixed Methods Appraisal Tool
MR: mixed reality
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
TiSDD: This is Service Design Doing
UX: user experience
VR: virtual reality
XR: extended reality

Edited by P Jansons; submitted 27.Oct.2025; peer-reviewed by J Ferrer Costa, SA Kristina; comments to author 14.Dec.2025; revised version received 02.Jan.2026; accepted 04.Jan.2026; published 01.Apr.2026.

Please cite as:
Niu J, Yin Y, Wang S
Service Design Strategies to Enhance Exercise Adherence in Extended Reality Interventions for Older Adults: Systematic Review
JMIR Aging 2026;9:e86595
URL: https://aging.jmir.org/2026/1/e86595
doi:10.2196/86595
PMID:41920595

©Jiangpan Niu, Yuanyuan Yin, Shan Wang. Originally published in JMIR Aging (https://aging.jmir.org), 01.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

“Living Inside the Walls”: Systematic Review and Qualitative Meta-Synthesis

Weitong Li[1*], PhD; Ziqi Mei[2*], PhD; WenJing Tu[1], Prof Dr; Yulei Song[1], Prof Dr; Yayi Zhao[1], Prof Dr; Yamei Bai[1], Prof Dr; Guihua Xu[1], Prof Dr

[1]School of Nursing, Nanjing University of Chinese Medicine, Nanjing, China

[2]Department of Psychology, Academy of Advanced Interdisciplinary Studies, Wuhan University, Wuhan, China

[*]these authors contributed equally

Corresponding Author:
Guihua Xu, Prof Dr
School of Nursing
Nanjing University of Chinese Medicine
Xianlin Road No 138
Nanjing, 210023
China
Phone: 86 1 395 204 9598
Email: 7115@njucm.edu.cn

Abstract

Background: Social isolation is a prevalent issue among older people in long-term care facilities (LTCFs), with profound negative impacts on their quality of life and mental health. However, the authentic experiences and underlying mechanisms of social isolation among older people in LTCFs remain understudied. A nuanced understanding of these experiences is essential for designing targeted nursing interventions.

Objective: This study aimed to systematically review and synthesize qualitative evidence to explore the experiences of social isolation among older adults in LTCFs and analyze them in terms of causes, mechanisms, outcomes, and strategies.

Methods: A systematic review and qualitative meta-synthesis were conducted. We searched PubMed, Web of Science, Embase, PsycInfo, Cochrane Library, CINAHL, Scopus, Chinese Biomedical Literature Database, CNKI, Wanfang Data, and Weipu (VIP) with no restriction on the start year, up to February 2025. Two reviewers (WL and ZM) independently screened studies, extracted data using a Microsoft Excel spreadsheet, and assessed study quality with the Joanna Briggs Institute Critical Appraisal Checklist for Qualitative Research. Qualitative data synthesis was performed following Thomas and Harden’s thematic and content analysis method, including line-by-line coding of participant quotes, organizing codes into descriptive themes, and developing analytical themes.

Results: A total of 4856 papers were retrieved through systematic searches, and 14 qualitative studies met the inclusion criteria. Four overarching analytical themes were identified: (1) root cause: person-environment mismatch (encompassing 4 interrelated dimensions: imbalance between declining individual capacity and insufficient environmental support, negative self-perception paired with inadequate positive environmental feedback, personal biases conflicting with homogeneous institutional social ecology, and disrupted social roles due to rigid institutional management); (2) overt behavioral patterns: proactive isolation, defensive isolation, and adaptive isolation; (3) inner emotional experiences: feelings of confinement, sorrow, and weariness; and (4) strategies for restoring person-environment match: active engagement in interaction, rational acceptance of coexistence, and passive avoidance and withdrawal. Notably, person-environment mismatch was confirmed as the core mechanism driving the onset and persistence of social isolation in this population.

Conclusions: Person-environment mismatch is the fundamental driver of social isolation among older adults in LTCFs. Interventions should prioritize restoring person-environment alignment—including optimizing physical environments, reconstructing supportive social environments, and enhancing psychological empowerment. Future studies should focus on developing tools to quantify person-environment matching degrees in LTCFs and conduct longitudinal evaluations of targeted interventions to further reduce social isolation and promote active social participation among older people.

(JMIR Aging 2026;9:e78142) doi:10.2196/78142

XSL•FO
RenderX

KEYWORDS

social isolation; meta-synthesis; LTCF; long-term care facility; older people; qualitative research; systematic review

Introduction

Population aging is an increasingly prominent global issue, with projections indicating that the number of adults aged 65 years and older will exceed 1.6 billion by 2050—more than double the current figure [1]. Against the backdrop of increasingly declining birth rates and an aging population, faced with practical challenges including family downsizing, regional separation, and a scarcity of community care resources, long-term care (LTC) is in high demand among older people. Given these circumstances, more families are choosing to place older relatives in long-term care facilities (LTCFs). The core mission of LTCFs is to support older adults in maintaining maximum independence, rather than providing services that replace their existing or potential abilities, undermine functional capacity, and increase care dependency [2]. Nevertheless, despite providing essential safety guarantees and professional care services, the inherent institutional attributes of LTCFs—marked by standardized structured environments, rigid administrative regulations, and centralized care protocols—may unintentionally exacerbate the risk of social isolation among older people [3]. To contextualize the focus of this study, a clear distinction is drawn between social isolation and loneliness: social isolation denotes an objective lack of meaningful social contacts and high-quality interpersonal relationships, while loneliness refers to a subjective emotional response to unmet interpersonal needs [4-6].

Social isolation represents a significant potential risk faced by older adults in LTCFs of older people care [7]. Evidence suggests that LTCFs do not adequately provide comprehensive care services suited to the diverse requirements of older people [8]. In LTCFs, tight space-time management standards restrict the timing, place, and manner of older individuals' daily activities. Due to understaffing and severe workloads, caregiving staff devotes the majority of their time and energy to satisfying older people's fundamental physical requirements [9], such as aid with daily living and medical care. In such settings, compared to those living in the community and at home, older people in LTCFs often lack meaningful social engagement, adequate sensory stimulation, and positive interpersonal relationships, and thus experience a far higher rate of social isolation [10,11]. A number of studies have shown that the social network of older people, as an important social resource, is closely related to their health and quality of life [12,13]. Older people in LTCFs who experience social isolation face an increased risk of mortality, higher morbidity, depression, and cognitive decline, and they are also more prone to physical pain and psychological distress [14]. Meanwhile, the deterioration in physical function makes older people in LTCFs more worried, worsens social isolation [15], and develops a vicious cycle, which may lead to a greater burden of institutional care and have negative societal impacts.

In actuality, for older people facing the risk of reduced social networks and increased social isolation, increasing social integration of older people is of vital importance, as it could alleviate the devastating sense of isolation and loneliness from LTC nursing and improve the quality of life of older people [16]. Despite decades of research on social isolation, most studies on social isolation of older people have focused primarily on community-dwelling older adults [17]. However, social isolation among older people in institutionalized care settings may differ significantly from that of older people living in communities in terms of mechanisms, manifestations, and effects [18]. Due to factors such as traditional cultural beliefs and weakened social functions of older people, with the aim of not burdening their children and reducing family burden, they may hide their emotional needs to some extent after entering in LTCFs, resulting in hidden and difficult to dynamically identify symptoms of social isolation. At the same time, due to management needs, LTCFs may encounter stricter institutional rules, limited environmental support, and reduced social network relationships [19]. Because of this group's unique physical functions, cognitive status, and social roles, the complexity of personal and environmental circumstances increases their risk of social isolation, making it more difficult to overcome through regular social support networks. However, older people living in the community could usually keep certain social roles and engage in specific social activities to interact with the outside world through their existing relatives, colleagues, and neighbors [20]. As a result, moving the study focus to older people in LTCFs could improve their quality of life, lower medical expenditures, and lessen the burden of social care for older adults [21]. This not only broadens the situational relevance of social isolation theory but also provides a practical foundation for creating more tailored intervention tactics.

Given the institutional characteristics of LTCFs and the objective nature of residents' social isolation, the person-environment fit (P-E fit) theory provides the core theoretical framework for explaining the underlying mechanisms of this phenomenon [22]. This theory focuses on the dynamic interaction between individuals and their environments, stressing that the alignment between personal needs or attributes and environmental resources or provisions determines adaptive outcomes [23]. Specifically, when older people's core needs (autonomy, social engagement, and dignity) match LTCFs' resources, norms, and support, they tend to maintain active social participation and well-being. Conversely, standardized care norms may suppress social initiative, and insufficient interaction opportunities fail to meet diverse needs. Such a mismatch may further induce psychological distress or physical health decline. From the P-E fit perspective, LTCF residents' social isolation can thus be viewed as a negative outcome of the imbalance between older people's personal resources and the environmental resources provided by the institution. However, this research offers a limited understanding of how older adults in LTCFs experience social isolation while neglecting older adults' subjective experiences of this mismatch. Key questions remain: How do LTCF environments reshape residents' social experiences? How do they make sense of, interpret, and respond to this sense of "isolation"? Answering

XSL•FO
RenderX

these requires targeted qualitative exploration—laying the foundation for this study.

Based on the results of systematic sorting and comparison, this study opted to use qualitative meta-synthesis as the research approach [24]: (1) social isolation among LTCF residents differs in its nature and manifestations from that among community-dwelling older people; it differs significantly in formation mechanisms, manifestations, and influencing variables. Through rigorous literature screening and quality assessment, qualitative meta-synthesis can not only systematically integrate scattered findings from existing qualitative research and connect fragmented evidence (environmental constraints, personal perceptions, and interaction patterns) but also reveal the unique formation mechanism of social isolation among LTCF-dwelling older adults via "reinterpretation" and "theoretical synthesis," thereby better addressing the uniqueness and innovation of this study [25]. (2) This study uses the P-E fit hypothesis to explain how environmental-person mismatch in LTCFs causes social isolation among older adults. Most current qualitative research focuses on 1-sided influencing factors, while qualitative meta-synthesis can identify qualitative evidence of "individual—environment—interaction" in existing studies, revise and refine the P-E fit theory framework for closed settings, and clarify the key mechanisms of social isolation in LTCFs and the specific forms of "mismatch" [26]. The need to deepen the P-E fit theory arises from the fundamental differences between LTCF residents and community-dwelling older adults [27]. Qualitative meta-synthesis could strengthen this theory to explain the formation mechanism of social isolation in LTCFs and potential solutions, thus promoting a systematic understanding of the phenomenon.

This review aimed to identify the experiences of social isolation among older people in LTCFs and analyze them in terms of causes, mechanisms, emotional experiences, and coping strategies. The findings of this study will inform the design and implementation of targeted interventions to alleviate social isolation and improve the well-being of this population.

Methods

Overview

We conducted a systematic review and qualitative meta-synthesis (PROSPERO CRD42024591240) and report the findings in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Multimedia Appendix 1) [28] and the ENTREQ (Enhancing Transparency in Reporting the Synthesis of Qualitative Research) statement [29] (Multimedia Appendix 2). We chose this study approach because qualitative meta-synthesis is useful for generating theoretical and conceptual knowledge, informing therapeutic interventions and research programs, and improving the clarity, comprehensiveness, and transparency of qualitative systematic reviews [30].

Search Strategy

We conducted a comprehensive electronic search of English and Chinese databases, including PubMed, Web of Science, Embase, PsycInfo, the Cochrane Library, CINAHL, Scopus, the Chinese Biomedical Literature Database, CNKI, Wanfang Data, and Weipu (VIP). The search included qualitative studies exploring social isolation experiences among older adults receiving LTC, with no start date restriction and a cutoff date of February 2025. According to the PICOS (population, intervention, context, outcome, study design) framework recommended by the Joanna Briggs Institute (JBI) Centre for Evidence-Based Healthcare in Australia, 4 key concepts were used: older adults, social isolation, LTC, and qualitative research. A comprehensive search was conducted for each concept using a combination of MeSH terms and free-text keywords. A backward citation search was also performed for all identified studies. For detailed examples of our search strategy, see Multimedia Appendix 3.

Inclusion and Exclusion Criteria

The PICOS model recommended by the JBI Evidence-Based Healthcare Center in Australia was used to determine the inclusion and exclusion criteria for the study. The inclusion criteria were as follows: (1) population (P): older people aged 60 years and older who are receiving LTC; (2) intervention (I): the causes and lived experiences related to social isolation among older adults receiving LTC; (3) context (C): residential facilities providing LTC for individuals who require assistance with daily living activities and have specific health and care needs, such as nursing homes or integrated medical and nursing care facilities; (4) outcome (O): the manifestations, impacts, and coping strategies of social isolation; and (5) study design (S): qualitative research or mixed method studies (only the qualitative component was included), using methodologies such as phenomenology, grounded theory, or ethnography.

The exclusion criteria were as follows: (1) studies where social isolation was not the primary research phenomenon but only a subtheme; (2) social isolation caused by lockdowns or isolation measures during the COVID-19 pandemic; (3) literature not available in English or Chinese; (4) full texts not accessible; (5) duplicate publications or studies with incomplete data; and (6) conference abstracts, commentaries, or reviews.

Study Selection and Data Extraction

Search results were entered into EndNote (version 8; Clarivate Analytics), then exported to Covidence after duplicates were removed. Five reviewers (WL, ZM, WT, YS, and YZ) independently screened titles and abstracts, with full-text reviews conducted in duplicate (2 reviewers [WL and ZM] per paper) to identify eligible studies. Eligible data were retrieved and entered into a Microsoft Excel spreadsheet, which included (1) general study and sample characteristics, (2) methodology, (3) participant quotes, and (4) themes and findings. Throughout, conflicts were settled through discussion and judgment by a third reviewer (YZ).

Appraisal of Methodological Quality

The JBI Critical Appraisal Checklist for Qualitative Research, with 10 items, guided the evaluation of the quality of the included publications [31]. Each publication was examined by 2 researchers (WT and YS) who worked separately. Disagreements were also resolved through discussion and further

XSL•FO
RenderX

evaluation. Inclusion needed a specified minimum of "yes" in 6 domains. All included studies clearly outlined their objectives and provided a thorough explanation of how the research was conducted in alignment with these objectives. Two studies reported all items recommended by the JBI-Qualitative Assessment and Review Instrument tool.

Data Synthesis

This study adopts a systematic review and thematic synthesis approach focusing on qualitative research. In the case of mixed method studies, only the qualitative data were included for analysis and integration to maintain methodological coherence. This study followed the three-stage thematic synthesis process proposed by Thomas and Harden [25]: (1) line-by-line coding of relevant texts: all descriptions related to the social isolation phenomenon among older people in LTCFs were compiled and coded; (2) organization of codes into descriptive themes: new descriptive themes were created by comparing the data from the first phase (eg, boredom and rational acceptance of coexistence), with connections clarified, similar themes merged, and grouped coding applied; and (3) development of analytical themes: the descriptive themes from the previous stage were re-examined, and similar descriptive themes were classified into a comprehensive outcome (the analytical themes included causes, manifestations, feelings, and coping strategies of social isolation). The synthesis analysis was conducted by 2 researchers (YS and YB), with a third experienced independent reviewer (GX) coordinating the thematic analysis. The entire research team carefully reviewed the data analysis process to ensure consistency in the interpretation of results and adequacy of the analytical themes.

Rigor, Trustworthiness, and Reflexivity

Rigor

This study used a comprehensive search strategy across 11 databases, rigorous quality appraisal using JBI criteria, and consistent data analysis following Thomas and Harden's [25] 3-stage thematic synthesis. Two researchers (WL and ZM) independently conducted screening, data extraction, and coding, with disagreements resolved through discussion and third-reviewer (YZ) adjudication, and the entire research team reviewed the coding and theme development process to ensure consistency and adequacy of the synthesized results.

Trustworthiness

Drawing on Lincoln and Guba's [32,33] criteria for establishing trustworthiness in qualitative research, we implemented four strategies: (1) credibility: prioritizing participants' direct quotes over authors' interpretations to ensure that findings authentically represent older adults' experiences. Member checking was indirectly achieved by verifying that the synthesized themes were consistent with the original findings of the included studies, with the research team discussing and confirming the alignment between coded data and thematic interpretations; (2) transferability: providing thick descriptions of study contexts (settings, participant characteristics, and care environments) to enable readers to assess applicability to other LTCF settings; (3) dependability: maintaining an audit trail documenting all analytical decisions, data extraction spreadsheets, coding frameworks, and theme development processes; this trail enables other researchers to replicate the study process and verify the consistency of the results; and (4) confirmability: 2 independent researchers (WL and WT) performed data extraction and coding, with a third reviewer (YS) resolving discrepancies to reduce bias. Thematic synthesis was guided by the P-E fit theory, and the research team repeatedly compared synthesized themes with original data to ensure that conclusions were data-driven rather than based on prior assumptions.

Reflexivity

The research team comprised nursing and psychology researchers with diverse backgrounds in geriatric care, institutional management, and qualitative methods. Acknowledging that team members' clinical experiences and cultural contexts (Chinese health care settings) might introduce bias, we implemented three reflexive practices [34]: (1) maintaining a reflexive journal to document preconceptions and analytical decisions, with explicit documentation during thematic synthesis of instances where personal experiences or disciplinary backgrounds could influence interpretation; (2) conducting peer debriefing with external qualitative experts not involved in data analysis—an independent researcher specializing in social isolation reviewed the final thematic framework, providing feedback on potential biases to ensure that findings were objective and data-grounded; and (3) explicitly bracketing prior assumptions during coding via iterative comparison of codes against raw data. While 2 cofirst authors (WL and ZM) led data analysis, multidisciplinary perspectives from psychology and the older people care management team members counterbalanced individual interpretations.

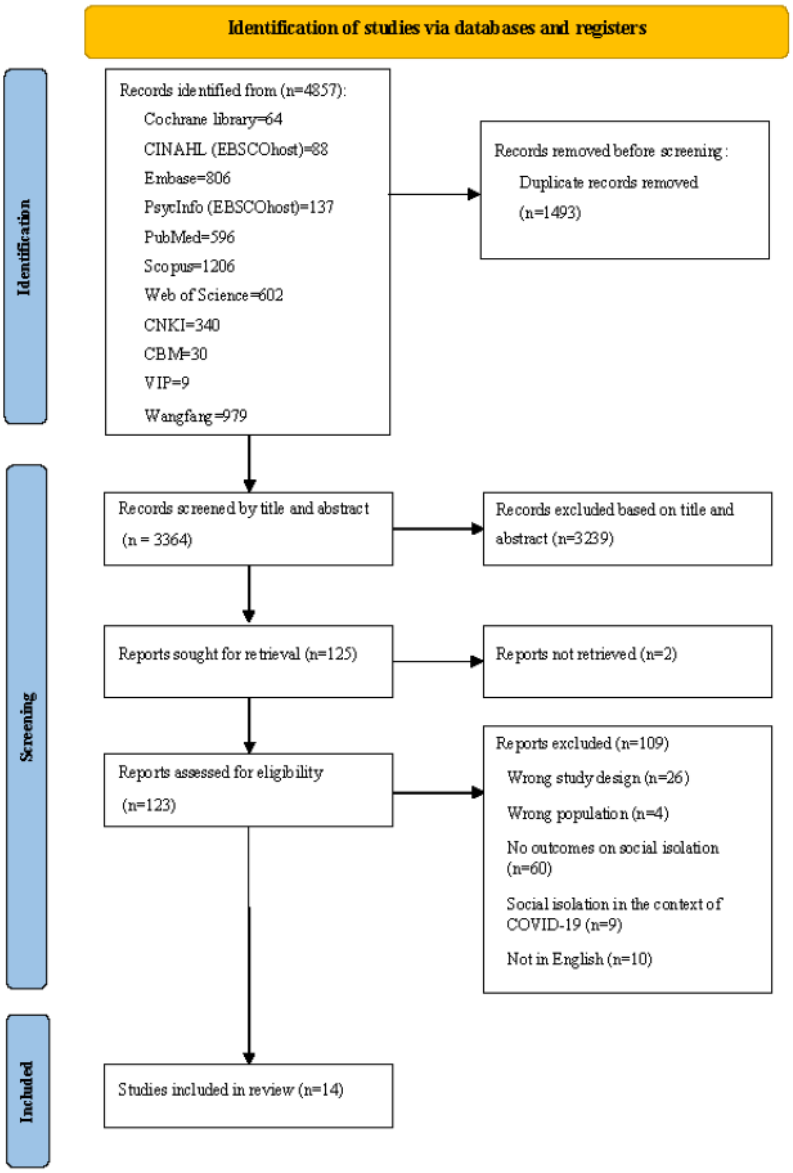
Results

Overview

A total of 4856 studies were identified via systematic database and hand searches. Of these, 123 underwent full-text review, and 14 qualitative studies involving 210 older people were ultimately included in the meta-synthesis. The screening process is shown in Figure 1. The studies (2006-2023) were conducted in various regions, including China (n=7), the United Kingdom (n=3), Ireland (n=1), New Zealand (n=1), Australia (n=1), and Finland (n=1). The research settings included nursing homes, social welfare institutions, and hospitals. In studies that clearly report the sex of participants, 94 female of 148 (64%) participants were included, with a relatively lower proportion of male participants. However, the study analyses were based on the entire study population of older people. In total, 6 studies reported the duration of LTC, with a minimum duration of 10 weeks, while the remaining studies did not provide information on the length of care. A variety of qualitative methodologies were used, with mixed methods designs being the most common, with data collection methods primarily relying on semistructured interviews and participant observation. The main findings of this review are presented in Table 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.



XSL•FO
RenderX

Table 1. Characteristics of included studies.

Authors	Location	Aim	Participants	Age (years) of older adults	Settings	Duration	Research design	Data collection methods	Main results or themes
Bartlett [35]	United Kingdom	To explore how men with dementia experience, and deal with, nursing home life	1 older adult with Alzheimer disease	65-84	Nursing home	≥5 months	Case study	Semistructured interview	• Experiences of social exclusion • Aligning self with other men • Aligning self with masculine behaviors
Buckley and McCarthy [36]	Ireland	To explore the lived experience of social connectedness as perceived by cognitively intact older adult residents	10 older people with scores greater than 6 on the Mini-Mental State Examination	71-99	Urban long-term care facility	≥1 year	Qualitative phenomenological design	Semistructured in-depth interviews	• Superficial relationships • Substitution • Outside world connection • Mental ability • Attitudes or actions of caregivers • Feelings of isolation
Cook et al [37]	United Kingdom	To understand the difficulties that residents with sensory impairments face when interacting with others	Study A: 8 older adults; study B: 18 older adults	Study A: 52-95; study B: 70-100	Nursing care home	Study A: 1.5-6 years; study B: not specified	Study A: hermeneutic inquiry; study B: constructivist study	Study A: interviews; study B: semistructured interviews, participant observation, and resident focus group interviews	• Being a member of the resident community • Getting to know other residents and supporting each other • Maintaining interaction and developing relationships with fellow residents
Goll et al [38]	United Kingdom	To explore the relationship between social participation and social identity among lonely older adults living independently in London	15 older adults, all of whom reported some form of illness or disability	62-100	Urban voluntary sector organizations situated	Not specified	Thematic analysis	Semistructured interview	• Overt barriers • Responses to barriers • Social fears • Fear of losing preferred identities
Li and Bai [39]	China	To explore the current situation of social isolation among disabled older adults in older people care institutions, the underlying mechanisms, and its impact on their physical and mental health	14 older adults with mild to moderate disabilities	≥65	Nursing home	Not specified	Qualitative research	In-depth interview method	• From spatial isolation to disconnection from external society • From spatial isolation to disconnection from internal society • From social disconnection to perceived isolation

XSL•FO
RenderX

Authors	Location	Aim	Participants	Age (years) of older adults	Settings	Duration	Research design	Data collection methods	Main results or themes
Li [40]	China	To explore the current situation, challenges, and causes of social isolation among partially disabled older adults in institutional care	22 older adults with partial disabilities	≥65	Nursing home	≥6 months	Mixed methodological research	Semistructured in-depth interviews; participant observation method	• Social isolation of semidisabled older people residents in X institution • Problems caused by social isolation among semidisabled older people residents in X institution • Causes of social isolation among semidisabled older people residents in X institution
Mu and Yin [41]	China	To understand the current situation and characteristics of the breakdown and reconstruction of social networks among older adults in nursing homes in southwestern Hubei	14 older people with clear thinking and articulate speech	65-87	Social welfare institute	≥6 months	Qualitative research	Semistructured in-depth interviews; participant observation method	• Breakdown and reconstruction of social networks among older people residents in H social welfare institution • Characteristics of social network breakdown and reconstruction among older people residents in H social welfare institution
Neves et al [42]	Australia	To better understand how frail older people experience social isolation and loneliness	22 residents experiencing (or at risk of) loneliness and/or social isolation, identified by staff	65-95	Care home	≥10 weeks	Qualitative research	Semistructured interviews; participant observation method	• Meanings and contexts of loneliness and social isolation • Loneliness as relational • Loneliness as oldering • Loneliness as a personal trouble • Loneliness as a sickness • Coping and protective strategies
Nielson et al [43]	New Zealand		12 cognitively competent older adults	70-85		2.5-6 years	Ethnographic approach		

XSL•FO
RenderX

Authors	Location	Aim	Participants	Age (years) of older adults	Settings	Duration	Research design	Data collection methods	Main results or themes
		To explore how residents of a commercially operated housing and care retirement complex experience inclusion and exclusion and their sense of community within a diverse environment of fit and frail residents			Commercially operated housing and care retirement complex			Semiformal interviews; participant observation method; artifact method	• Motivations and expectations • Building social connections through participation • Social group membership and belonging • Health decline and social change • Health decline and stigma
Pirhonen et al [44]	Finland	To shed light on residents' chances to reach affiliation in their new living surroundings, and thus adjust to that social environment	15 older adults	≥65	Sheltered home	Not specified	Ethnographic approach	Thematic interviews; participant observation method	• Detachment inside the assisted living facility • Separateness from the social world outside the assisted living facility
Wan [45]	China	To gain a comprehensive understanding of the real challenges of social isolation faced by institutionalized older adults	6 older adults experiencing both objective and subjective social isolation	74-84	Nursing institution	≥6 months	Mixed methodological research	Semistructured interviews; participant observation method	• Objective social isolation • Subjective social isolation
Wei [46]	China	To conduct a thorough analysis of the causes and factors influencing social isolation among older people with diabetes	11 older people with diabetes who experience social isolation	60-84	Hospital	Not specified	Mixed methodological research	Semistructured interview	• Changes in physical function • Insufficient inner strength • Weakening of social support • Ineffective social coping
Zhang [47]	China	To understand the levels of interpersonal alienation among semidisabled older people in care institutions and to explore the underlying causes of high interpersonal alienation	10 semidisabled older adults with a score exceeding 2.8 on the General Social Alienation Scale	79-85	Nursing institution	≥6 months	Mixed methodological research	Semistructured interviews; participant observation method	

XSL•FO
RenderX

Authors	Location	Aim	Participants	Age (years) of older adults	Settings	Duration	Research design	Data collection methods	Main results or themes
									• Manifestations of interpersonal alienation among institutionalized older Adults with partial disabilities • Causes of interpersonal alienation among institutionalized older adults with partial disabilities
Zhu [48]	China	To understand the manifestations and causes of social interaction barriers among partially disabled older adults in nursing homes	32 partially disabled older adults	≥65	Nursing institution	≥3 months	Mixed methodological research	Unstructured interview; participant observation method	• Manifestations of social interaction barriers among institutionalized older adults with partial disabilities • Causes of social interaction barriers among institutionalized older adults with partial disabilities

Quality Assessment

According to the JBI criteria, all studies, except for 1, described at least 7 of the key items. All studies followed strict qualitative research methodologies and designs. Except for 4 studies, the remaining studies provided detailed statements of participants' results. In total, 5 studies did not explicitly report whether appropriate institutional ethics approval had been obtained. Only 3 studies addressed the influence of the researcher on the research or the impact of the research on the researcher (Table 2), indicating a potential bias in the findings.

XSL•FO
RenderX

Table 2. Joanna Briggs Institute critical appraisal of included studies[a].

Authors	1	2	3	4	5	6	7	8	9	10	Quality rating
Bartlett [35]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	B
Buckley and McCarthy [36]	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	B
Cook et al [37]	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	B
Goll et al [38]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	B
Li and Bai [39]	Yes	Yes	Yes	Yes	Yes	No	No	No	Unclear	Yes	B
Li [40]	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Unclear	Yes	B
Mu and Yin [41]	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Unclear	Yes	B
Neves et al [42]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	B
Nielson et al [43]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	A
Pirhonen et al [44]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	A
Wan [45]	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Unclear	Yes	B
Wei [46]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	B
Zhang [47]	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Unclear	Yes	B
Zhu [48]	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	B

[a]Domains: (1) Is there congruity between the stated philosophical perspective and the research methodology? (2) Is there congruity between the research methodology and the research question or objectives? (3) Is there congruity between the research methodology and the methods used to collect data? (4) Is there congruity between the research methodology and the representation and analysis of data? (5) Is there congruity between the research methodology and the interpretation of results? (6) Is there a statement locating the researcher culturally or theoretically? (7) Is the influence of the researcher on the research, and vice versa, addressed? (8) Are participants, and their voices, adequately represented? (9) Is the research ethical according to current criteria or, for recent studies, is there evidence of ethics approval by an appropriate body? (10) Do the conclusions drawn in the research report flow from the analysis, or interpretation, of the data?

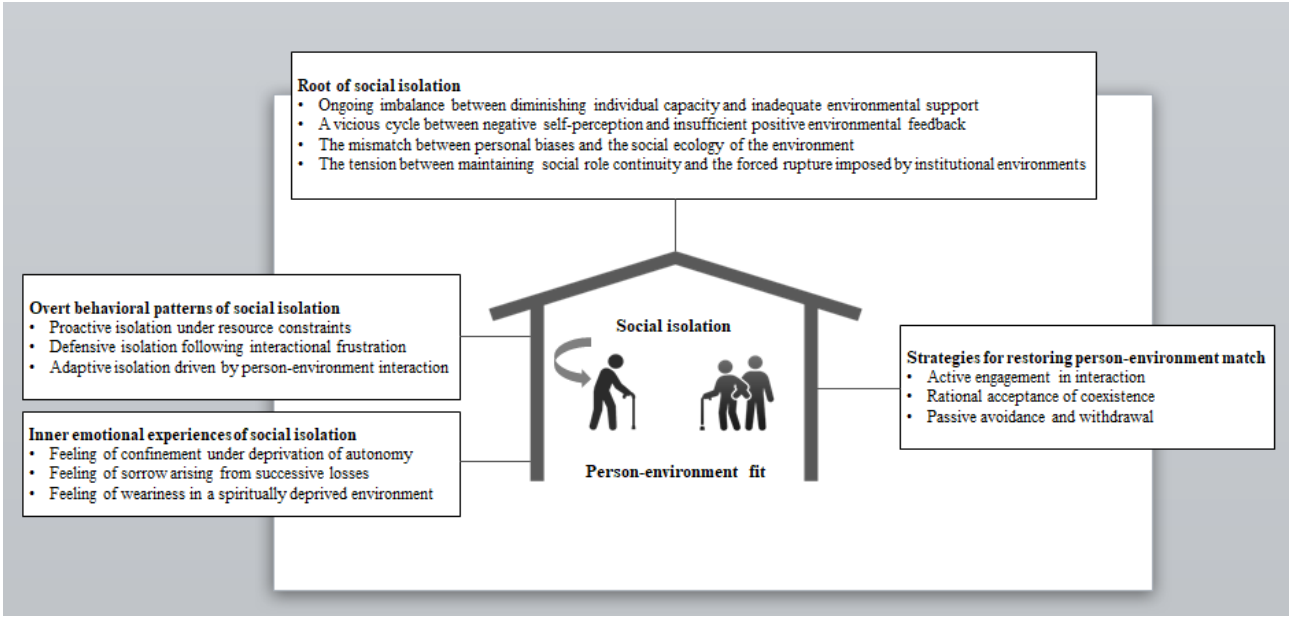
Main Findings of the Meta-Synthesis

Overview

This study generated 4 main themes (roots, manifestations, emotional experiences, and coping strategies of social isolation among older adults) and 13 subthemes.

The research team and qualified interpreters reviewed the identified themes to ensure alignment with the original authors' intended interpretations, thereby enhancing the accuracy of thematic analysis across all included studies. These themes effectively reflect the social isolation experiences and perceptions of older people in LTCFs. Figure 2 shows the examination of each theme.

Figure 2. Thematic overview diagram.



XSL•FO
RenderX

Theme 1: Person-Environment Mismatch as the Root of Social Isolation

Ongoing Imbalance Between Diminishing Individual Capacity and Inadequate Environmental Support

With advancing age, individuals experience a gradual decline in physiological functioning accompanied by a marked increase in the prevalence of chronic diseases. Older adults living in LTCFs frequently experience ongoing declines in mobility and capacity for independent daily living [49]. However, many LTC settings remain inadequate in providing supportive features—such as accessible infrastructure, mobility assistance, and functional spatial layout—thus failing to effectively offset the decline in residents' functional capacity. This insufficiency of environmental support diminishes older adults' autonomy and opportunities for external engagement, constraining their social participation and ultimately resulting in social isolation driven by the combined effects of functional decline and environmental misfit.

> *As I age and my health deteriorates, I need daily medication and injections. Over the past month, I have been receiving injections continuously and taking numerous pills each day. I truly have no desire to go out, nor is it physically feasible for me. Instead, I am just lying in bed* [48]
>
> *I used to be able to walk and visit friends, but now with this illness (diabetes) and my gout, along with foot sores, I can't walk anymore. I can only sit or lie down, and I can't go anywhere* [46]

A Vicious Cycle Between Negative Self-Perception and Insufficient Positive Environmental Feedback

Older adults living in institutional settings are susceptible to negative psychological states—including stigma, self-directed ageism, and feelings of guilt—triggered by aging and illness. In contrast, LTC environments generally provide far fewer affirming social interactions and psychological support than community contexts enriched by family and neighborhood ties, leaving older adults' negative self-perceptions insufficiently mitigated or corrected. Over time, these enduring negative self-perceptions lead older adults to withdraw from social engagement due to fear of being judged or misunderstood, thereby further limiting their access to positive environmental experiences. As a result, a self-perpetuating vicious cycle develops, in which negative self-cognition and insufficient positive feedback from the environment continuously reinforce each other.

> *If you go out by [accessible taxi] ... you wonder if they're gonna turn up* [38]
>
> *With this illness (diabetes), my savings and pension have been spent entirely on medication. Even going out to socialize incurs costs. Sometimes my children help out a little, but I feel embarrassed to ask for support; I don't want to be a burden to them* [46]

The Mismatch Between Personal Biases and the Social Ecology of the Environment

After entering LTCFs, older adults often form social circles based on individual differences such as personal values, behavioral patterns, and past experiences. For example, some may instinctively distance themselves from peers with highly stigmatized illnesses and choose to interact only with those who are moderately impaired or still self-sufficient. Institutional environments often fail to provide diverse and flexible social groups; thus, the avoidance strategies individuals adopt to maintain self-identity become barriers to rebuilding their social networks.

> *These severely disabled older people individuals should not have been placed on our floor. They are entirely dependent on the caregivers for mobility, and their temperaments are often volatile, leaving the area unclean. We do not want to live on the same floor with them* [39]
>
> *I don't want to go out. Those people lack manners. I went to college and used to be a Chinese teacher, so I can't communicate clearly with them. Take my neighbor, for example—she's practically crazy. She insults others and calls me names, saying I'm an introverted bore. She mocks my education, but her cultural level is far below mine. I refuse to share a room with her or interact with them* [47]
>
> *people don't want to socialise with frail people* [43]

The Tension Between Maintaining Social Role Continuity and the Forced Rupture Imposed by Institutional Environments

Moving into an institutional setting imposes an abrupt and involuntary disruption on older adults' established social functions and living contexts. Yet, the enclosed structure of LTC settings, the task-driven nature of care practices (focused on physical tasks rather than emotional connection), and the rigidity of institutional management collectively hinder the smooth transition and continuity of older adults' social roles. These conditions expedite the erosion of their social identity and emotional support, ultimately exacerbating social isolation.

> *There's nobody to talk to (...) no normal people, just the tea lady and the kitchen staff* [47]
>
> *I've been thinking about my friend who lives near Cuihu, but now that I'm living in this facility, I can't go out. Oh well, I'm fine here (telling myself that). It's probably best not to meet up anyway* [47]
>
> *I used to want to chat, so I would wait for the caregiver to come at mealtime just to say a few words. But she was always busy, rushing to feed me and then move on to the others. Gradually, I stopped trying to talk* [40]

Theme 2: Overt Behavioral Patterns of Social Isolation in the Context of Person-Environment Mismatch

Proactive Isolation Under Resource Constraints

Such isolation behaviors constitute a deliberate avoidance strategy adopted by older adults following an evaluation of their

XSL•FO
RenderX

functional restrictions and the physical or supportive barriers within institutional environments [50]. Although these older adults often desire social participation, when they perceive that environmental supports (such as hearing aids, barrier-free facilities, or accessible activity spaces) are inadequate to compensate for physical deterioration, they reasonably foresee possible frustration or risk during social engagement. Consequently, in order to prevent possible communication barriers, physical strain, or unsafe conditions, they opt to limit or withdraw from social engagement. This behavior reflects a rational, cost-benefit–driven decision aimed at minimizing the adverse outcomes that may result from social interaction.

> *I do want to talk to someone, but my health and other conditions don't allow it. Sometimes my spouse takes me out on a tricycle, but it still feels difficult to make friends or have a conversation* [46]

> *It's not that I don't want to go out and talk to people—it's just that I'm too deaf. Even if someone tries to speak to me, they'd have to shout, and I still might not catch what they're saying. So, I just stay in my room by myself* [40]

Defensive Isolation Following Interactional Frustration

This pattern of behavior arises from repeated negative interpersonal interactions experienced in institutional contexts. When older adults face frequent rejection or misunderstanding during social interactions, they are likely to develop a consistent pattern of negative attribution, placing the cause on others or on an unsupportive environment. As a defensive strategy, individuals deliberately retreat from social settings to shield themselves from additional interpersonal hurt and emotional pain. While this behavior offers short-term self-protection, in the long run, it depletes social capital and entrenches the individual's isolation.

> *Ask for help and you're turned down. That hurts ... It's not worth the aggravation ... I don't want to ask anybody for anything, nothing ... I don't want another knock-back* [38]

> *And they don't speak to me and I don't speak to them. We are like bits of kids, oh dear God. To think that I have come to this* [37]

Adaptive Isolation Driven by Person-Environment Interaction

Certain older adults demonstrate reduced intrinsic motivation for social interaction, influenced by personality characteristics or life stage factors. At the same time, the social activities offered by institutional environments tend to be limited in variety, depth, and attractiveness, thus insufficient to foster residents' motivation to engage. In this context, a passive equilibrium emerges between the individual's limited social needs and the environment's equally limited social offerings. Older adults' seeming acceptance of solitude is not only a manifestation of personal preference but also a consequence of the environment's insufficient provision of meaningful social engagement: "[Being alone] doesn't bother me anymore" [38].

Theme 3: Inner Emotional Experiences of Social Isolation in the Context of Person-Environment Mismatch

Feeling of Confinement Under Deprivation of Autonomy

The feeling of confinement arises from the tension between the individual's desire for autonomy and the physical as well as administrative constraints of institutional settings. As older adults' physical abilities deteriorate with age and illness, institutional environments often intensify restrictions through closed spatial layouts, rigid outing regulations, and standardized routines, resulting in a pronounced sense of being trapped within specific areas (eg, a room, a floor, or the entire facility). The confinement experience goes beyond physical restriction. It reflects a profound psychological response to the institutionalized deprivation of personal autonomy and indicates a severe loss of one's sense of control over the environment.

> *I just feel cut off from the outside world that's all ... I don't feel I have contact with the outside world at all* [36]

> *you can get in alright but you can't get out* [35]

Feeling of Sorrow Arising From Successive Losses

With the progression of aging and chronic illness, residents in LTCFs tend to endure a wider spectrum of losses than their peers, including diminished bodily autonomy, bereavement of close companions, separation from a spouse, and disruption of familiar environments. When faced with the loss of youth, vitality, and beloved relationships, certain older adults sink into a pervasive sorrow from ongoing losses, leading to an increasingly pessimistic outlook on life. During this journey marked by continual loss, some older people become absorbed in nostalgic recollections, rendering them more susceptible to distress and emotional isolation.

> *more of a torment I think ... That is the thing that gets me I start talking like this and then it goes out of my head. I have a hell of a job to get it back. It is awful really* [35]

> *All my family's gone. I had to accept that. You see other people coming in with families. It does hurt. But you move on. You just move on, say, "Well, they're lucky. Their family and that are still alive"* [42]

> *[My friend] who takes me shopping had been away, so I hadn't been out for quite some time* [38]

Feeling of Weariness in a Spiritually Deprived Environment

Some older persons in LTCFs find life tedious and uninspiring. On the one hand, their daily routines are dictated by established care processes, leaving little possibility for autonomy or personal choice, resulting in a lack of novelty and variation in their lives. On the other hand, the restricted range of activities available in LTCFs frequently fails to meet their different interests, leading people to believe that life is dull and lifeless. This lack of mental stimulation can quickly lead to boredom, exacerbating their feelings of loneliness and isolation. The absence of cognitive stimulation can readily induce fatigue and exacerbate their social isolation.

XSL•FO
RenderX

> *There's really nothing going on; I just can't do anything all day, and I feel a sense of emptiness inside* [48]
>
> *life is people; we need to belong; we need to talk, we old lonely men (...) we see each other every day, but we don't talk, we don't even know what to do* [42]

Theme 4: Strategies for Restoring Person-Environment Match to Alleviate Social Isolation

Active Engagement in Interaction

Some older people deal with social isolation by actively engaging in social activities and socializing with others. By expanding their exposure to the outside world, such as participating in group activities or attending institutions for older adults, they might form new relationships or pursue personal hobbies, re-establish social links with the outside world, and gain a sense of belonging and importance.

> *The nursing home has provided us with a sound system specifically for square dancing after dinner. Essentially, as long as one can walk, everyone comes together daily to enjoy dancing* [41]
>
> *if you don't participate you can be left out and feel isolated, so I participate in some of the things* [42]
>
> *It's nice to sit with the nurses and talk and laugh about things with them when they are not busy working* [44]

Rational Acceptance of Coexistence

Another group of older people copes by rationally accepting the institutional context and adjusting their mindsets to achieve inner balance, a strategy closely tied to the LTCF-specific experience of "relocating from family to institutional care." Over time, they seek inner peace by changing their mindset and progressively adjusting to live in harmony with their caregivers and cohabitants. This logical acceptance keeps individuals from experiencing more psychological distress, even though they still feel alienated from the outside world.

> *Now that the children have to work and take care of their own kids, they are under a lot of pressure and don't have time to look after us. That's why we've come here. We have to adapt to this environment, take care of ourselves, and avoid getting sick so we don't add to our children's burdens* [39]
>
> *Good child, I'm old now, and I'm stuck here. My spouse passed away early. My health isn't good; I have high blood pressure and have been on medication for a long time. I just don't have the energy to be around those lively people. Right now, I'm just taking it one day at a time and not thinking too much about it* [48]

Passive Avoidance and Withdrawal

Some older people adopt passive coping behaviors in response to LTCF-related social barriers, forming a maladaptive cycle of isolation. This coping strategy involves passive avoidance and withdrawal from social interactions among older people. Due to the fixed, inflexible schedules and routines in LTCF, residents' personalized social needs often cannot be met. In response, some older people choose to withdraw from institutional social activities, avoid social environments, and reduce interaction with others as a form of passive coping. Long-term avoidance may harm their psychological and physical health, making them feel even more alone. Over time, this creates a vicious cycle, in which older people retreat for extended periods of time, no longer seeking opportunities to contact with the outside world, and may even acquire suicidal ideation.

> *Well I just can't be bothered ... I just don't want to make the effort. I mean, I don't get up very early in the mornings ... When I feel lonely I don't want to do anything ... [I've] just lost interest* [38]

Discussion

Principal Findings

This synthesis identifies person-environment mismatch (encompassing 4 interrelated dimensions: functional decline vs insufficient environmental support, negative self-perception vs feedback deficit, personal bias vs homogeneous institutional ecology, and role disruption vs rigid management) as the core mechanism driving social isolation of older people in LTCFs. This mismatch manifests in 3 behavioral patterns (proactive, defensive, and adaptive isolation) and 3 emotional states (confinement, sorrow, and weariness), with residents adopting active engagement, rational acceptance, or passive withdrawal as coping strategies. Notably, LTCF residents' social isolation is a unique, institution-specific phenomenon (not an extension of community-dwelling), and interventions should prioritize restoring person-environment congruence—providing targeted guidance for optimizing institutional care, strengthening social support, and fostering active social participation among older people in LTCFs.

Core Mechanism Elaboration: Person-Environment Mismatch as the Root Cause of Social Isolation Among Older People in LTCFs

This study confirms that person-environment mismatch is the core driver of social isolation among older adults in LTCFs, validating the P-E fit theory and extending its utility to institutional older people care via empirical evidence. These 4 interrelated dimensions of this mismatch collectively contribute to social isolation, with synergistic rather than independent effects. First, the gap between declining physical capacity (eg, mobility loss from diabetes or gout) and insufficient environmental support (eg, lack of barrier-free facilities) forms a "functional restriction→participation barrier" loop. This finding is consistent with previous research linking environmental inadequacy to social withdrawal [51] and further highlights that this issue is persistent in LTCFs, where impaired mobility often confines residents to their beds. Second, negative self-perceptions—fueled by financial strain (eg, pensions spent on medication) or distrust in external support (eg, worries about accessible taxis)—are not mitigated by LTCFs' weak emotional feedback. Unlike community settings with family ties, this gap exacerbates self-directed ageism and active avoidance [46], and this study adds that such cycles are more intractable in closed

XSL•FO
RenderX

institutions. Third, personal biases (based on education or health status) collide with LTCFs' homogeneous social ecology. While Zhang et al [52] noted institutional homogeneity limits adaptation, this study further finds that the absence of flexible social groupings turns preferences (eg, avoiding severely disabled peers) into long-term isolation. Fourth, closed management and task-oriented care disrupt social roles (eg, losing "friend" or "teacher" identities), eroding emotional anchors. Consistent with the findings of Huang et al [53] on role disruption and isolation, this study highlights that LTCFs' lack of role-transition support accelerates this erosion. Together, these dimensions reveal that person-environment mismatch in LTCFs is a cumulative, interactive process—each dimension reinforces the others—making isolation more deeply rooted than in community or home settings.

Analysis of 3D Environmental Conflicts: Reflections on Person-Environment Mismatch

Overview

Drawing on the study's empirical findings, environmental conflicts in LTCFs can be categorized into 3 dimensions: physical, social, and psychological. Each dimension reflects specific manifestations of person-environment mismatch and deepens social isolation, consistent with the multidimensional environmental framework of Smith et al [54]. At the same time, this study found that the social isolation performance of older persons in LTCFs could be classified as "proactive," "defensive," and "adaptive," which is consistent with previous research findings [46,52].

Physical Environment Versus Physiological Needs: Aggravating "Sense of Confinement" and "Proactive Isolation"

Conflicts here stem from LTCFs' failure to align spatial-temporal arrangements with the declining physical capacity of older people. First, rigid time management (eg, unified meal or rest times) creates "flattened time," contradicting residents' need for autonomy. This exacerbates meaninglessness and passive isolation [51], and our results add that such rigidity leaves no room for adapting to individual energy levels (eg, residents too fatigued from medication to participate in scheduled activities). Second, insufficient barrier-free facilities (eg, lack of walking aids) trap mobility-impaired residents—our cases show that those with foot injuries or gout cannot reach activity areas, reinforcing physical confinement [52]. Third, closed spatial layouts sever external connections; residents unable to meet friends outside (as in our results) rely solely on internal resources, and when these are scarce, isolation deepens.

Social Environment Versus Social Needs: Triggering "Loneliness" and "Defensive Isolation"

This dimension's core issue is the mismatch between homogeneous social services and diverse individual needs. First, unified care (eg, passive waiting for meals or bathing) ignores social preferences—extroverts need group interactions, while introverts need one-on-one chats. We have a former Chinese teacher who is unable to communicate with peers with lower cultural levels, and this case illustrates this point [47]; this mismatch increases withdrawal. Second, caregivers prioritize tasks over emotional support; our results show that residents wanting to chat during meals are rushed, leading to "defensive isolation" (eg, no longer initiating interactions) [53-56]. Third, tense peer relationships—driven by health discrimination (eg, avoiding severely disabled peers)—undermine belonging. Our cases of residents refusing roommates who insult them reflect this, aligning with the observation of Nielson et al [43] of peer exclusion worsening isolation. Additionally, this study also believes that some older people may find it difficult to express their feelings of social isolation. This can make it difficult for caregivers to understand and sympathize with older people because of limitations and a lack of knowledge about the level of expertise of caregivers.

Psychological Environment Versus Psychological Resilience: Strengthening "Sense of Powerlessness" and "Adaptive Isolation"

Imbalances here lie in LTCFs' failure to support older people's mental coping with isolation. First, caregivers lack psychological skills to recognize loneliness or anxiety; uniformity of feeling among people remains impossible, and this study summarizes these themes as "active participation in interactions," "rational acceptance of coexistence," and "passive avoidance and withdrawal." Some older people stop expressing feelings when unheard, fostering powerlessness [57,58]. Second, monotonous activities (eg, repetitive handicrafts) fail to meet self-worth needs; older people in previous research have described "empty lives," leading to "adaptive isolation" (accepting loneliness) [59-62]. Third, static assessments ignore changing needs of older people (eg, hearing loss requiring hearing aids leads to a sense of powerlessness in expressing needs), and this static care deepens passive withdrawal.

Practical Implications: Person-Environment Matching-Based Interventions for Social Isolation

Overview

Guided by the study's core findings—effective coping strategies (eg, active participation and rational acceptance) and 3D environmental conflicts—interventions should focus on "repairing person-environment mismatch." Targeted, actionable measures are proposed across physical, social, and psychological dimensions to facilitate older adults' transition from "passive restriction" to "active adaptation," aligning with P-E fit theory's intervention orientation [23].

Optimizing the Physical Environment: Enhancing "Adaptability" to Reduce Physical Barriers to Socialization

Implement Flexible Time Management

Referring to older adults' need to "arrange activities independently" in the results, LTCFs should allow older adults to choose meal and rest times within a certain range (eg, "providing 2 meal periods in the morning and evening") [44]. Posters listing daily flexible slots should be displayed in residential wings. LTCFs could increase participation in 2 or more weekly timed activities (eg, 15:00 calligraphy and 17:30 group walks). This reduces the sense of meaninglessness caused

XSL•FO
RenderX

by "flattened time"—consistent with the "temporal autonomy" intervention [63].

Construct Semiopen Spatial Layouts

Within closed management, facilities can establish weekly outing schedules (eg, community park visits on Wednesday mornings and friend visits on Friday afternoons), with 1 staff member assigned to every 6 residents. On a monthly basis, facilities can invite 2 to 3 community volunteers (eg, retired teachers and craftsmen) to hold on-site interactive activities such as paper-cutting workshops [64]. These practices help older people reconnect with the outside world and reduce their sense of isolation from external society.

Reconstructing the Social Environment: Strengthening "Supportiveness" to Improve the Quality of Social Interaction

Transform Caregiver Role Positioning

LTCFs should strengthen training for caregivers in psychological support skills, transforming them from "task executors" to "social facilitators." For example, caregivers should "take the initiative to chat with older people during meals and listen to their feelings," integrating social support for older people into daily nursing work: nursing staff actively care about the older people's mood and feelings during meal delivery or daily patrols, communicate for about 5 minutes each time, and record their social and reaction situations in nursing records. This reduces the occurrence of "defensive isolation"—aligning with the view of Caspar and O'Rourke [65].

Foster an Inclusive Peer Atmosphere

Through "collective ice-breaking activities" and "peer mutual aid programs" [66] (eg, "pairing 1 high-functioning resident with 1 severely disabled resident for 30-minute daily interactions"), it breaks the prejudice of "unwillingness to live with severely disabled peers" and reconstructs positive peer relationships and eliminates "health discrimination."

Improving the Psychological Environment: Enhancing "Empowerment" to Boost Individual Coping Resilience

Strengthen Collaborative Support Between Families and LTCFs

Facilities can organize monthly "family activity days" (eg, parent-child tea parties on the second Sunday of each month) and provide free video call stations [67]. Staff can assist residents who are not familiar with digital devices, and each resident can make three 15-minute video calls per week. In addition, facilities can share "resident social progress reports" with family members quarterly to encourage family involvement in social goal-setting.

Establish a Dynamic Psychological Assessment Mechanism

LTCFs should regularly assess older adults' psychological states (eg, "loneliness level, social willingness") and changing needs (eg, "hearing loss requiring hearing aids"). Timely adjustments to intervention strategies can prevent "passive withdrawal due to unrecognized needs"—consistent with the proposal of Davies et al [68] for "dynamic monitoring to match psychological needs."

To ensure implementation, establish a three-tier oversight system: (1) micro: caregivers track daily intervention adherence via logs, (2) meso: LTCF managers conduct monthly spot checks (eg, verifying flexible meal times), and (3) macro: collaborate with local aging agencies to audit quarterly and provide funding for barrier-free renovations. This system ensures that interventions translate from policy to practice, driving residents' shift to active adaptation.

Strengths and Limitations

The strengths of this qualitative systematic study and meta-synthesis include a good grasp of older people's social isolation experiences. All relevant studies were identified using a careful search approach. However, unavoidable heterogeneity exists due to cross-national variations in LTCF resident characteristics and the care services provided. Factors including the sex and health status of the study population, the quality of care provided by caregivers, and the care environment quality may all influence how older people perceive their living environment. A key limitation is the uneven sex distribution among participants in the included studies, with most reporting a higher proportion of the female population. This sex imbalance may have influenced the salience of specific thematic findings. Future qualitative studies should aim for more balanced sex recruitment to capture sex-specific experiences of social isolation. At the same time, it is important to note that although many of the included studies were conducted from a Chinese perspective, several others originated from different countries. Cultural variations in older adults' perceptions of social isolation may additionally introduce bias into the interpretation of study findings.

Conclusions

This qualitative meta-synthesis explores social isolation experiences among older people in LTCFs and analyzes the underlying causes, behavioral manifestations, emotional experiences, and coping strategies associated with this phenomenon. Although this study describes social isolation experiences across both agentic and structural dimensions, most participants perceived social isolation as an individual challenge shaped by the interaction between their internal perceptions and the external institutional environment. As such, older people had to deal with these experiences using the social and individual strategies they had developed [69].

Additionally, there is a need to continue exploring the role of caregivers and other aged care staff in the prevention and management of social isolation of older people in LTCFs. Focusing on the needs of older people in LTCFs, through environmental transformation, relationship reconstruction, and dignity empowerment, the symbiotic relationship between humans and the environment can be reconstructed. Future research should further explore the quantification of person-environment matching in LTCFs—including the development of valid assessment tools—to enable more targeted interventions for social isolation.

XSL•FO
RenderX

Funding

This research is supported by the Research Project of Humanities and Social Sciences of the Ministry of Education (24YJC840018), 2024 University Philosophy and Social Science Research Project (2024SJYB0242), and National Natural Science Foundation of China (grants 72504149 and 72174095).

Data Availability

The datasets generated and/or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: WL, WT
Methodology: WL, ZM
Data curation: WL, ZM, YZ
Formal analysis: WL, ZM
Investigation: ZM, YZ
Project administration: WT, YS
Supervision: WT, YS
Writing—original draft: WL
Writing—review and editing: WT, YS, GX, YB

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA checklist.
[DOCX File , 655 KB - aging_v9i1e78142_app1.docx]

Multimedia Appendix 2
ENTREQ reporting guidelines.
[DOCX File , 20 KB - aging_v9i1e78142_app2.docx]

Multimedia Appendix 3
Literature search strategy and search terms.
[DOCX File , 21 KB - aging_v9i1e78142_app3.docx]

References

1. World social report 2023: leaving no one behind in an ageing world. United Nations | Department of Economic and Social Affairs. 2023. URL: https://desapublications.un.org/publications/world-social-report-2023-leaving-no-one-behind-ageing-world [accessed 2026-03-05]
2. Pot AM, Briggs AM, Beard JR. The sustainable development agenda needs to include long-term care. J Am Med Dir Assoc 2018;19(9):725-727. [doi: 10.1016/j.jamda.2018.04.009] [Medline: 30149838]
3. Plattner L, Brandstötter C, Paal P. Einsamkeit im pflegeheim? Erleben und maßnahmen zur verringerung: eine literaturübersicht [Loneliness in nursing homes-experience and measures for amelioration: a literature review]. Z Gerontol Geriatr 2022;55(1):5-10 [FREE Full text] [doi: 10.1007/s00391-021-01881-z] [Medline: 33856531]
4. Nicholson NR. Social isolation in older adults: an evolutionary concept analysis. J Adv Nurs 2009;65(6):1342-1352. [doi: 10.1111/j.1365-2648.2008.04959.x] [Medline: 19291185]
5. Wotherspoon N. Compelled loneliness and necessitated social isolation: "It's like being on the other side of a mirror, just looking in". Sociol Health Illn 2024;46(4):744-761. [doi: 10.1111/1467-9566.13732] [Medline: 38069691]
6. National Academies of Sciences, Engineering, and Medicine. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System. Washington, DC: The National Academies Press; 2020.
7. Boamah SA, Weldrick R, Lee TJ, Taylor N. Social isolation among older adults in long-term care: a scoping review. J Aging Health 2021;33(7-8):618-632 [FREE Full text] [doi: 10.1177/08982643211004174] [Medline: 33779366]
8. DeSalvo KB, Jones TM, Peabody J, McDonald J, Fihn S, Fan V, et al. Health care expenditure prediction with a single item, self-rated health measure. Med Care 2009;47(4):440-447. [doi: 10.1097/MLR.0b013e318190b716] [Medline: 19238099]
9. Diamond T. Making Gray Gold: Narratives of Nursing Home Care. Chicago, IL: University of Chicago Press; 1992.

XSL•FO
RenderX

10. Pomeroy ML, Cudjoe TKM, Cuellar AE, Ihara ES, Ornstein KA, Bollens-Lund E, et al. Association of social isolation with hospitalization and nursing home entry among community-dwelling older adults. JAMA Intern Med 2023;183(9):955-962. [doi: 10.1001/jamainternmed.2023.3064] [Medline: 37486647]
11. Abbott KM, Bettger JP, Hampton KN, Kohler H. The feasibility of measuring social networks among older adults in assisted living and dementia special care units. Dementia (London) 2015;14(2):199-219. [doi: 10.1177/1471301213494524] [Medline: 24339099]
12. Beridze G, Ayala A, Ribeiro O, Fernández-Mayoralas G, Rodríguez-Blázquez C, Rodríguez-Rodríguez V, et al. Are loneliness and social isolation associated with quality of life in older adults? Insights from northern and southern Europe. Int J Environ Res Public Health 2020;17(22):8637 [FREE Full text] [doi: 10.3390/ijerph17228637] [Medline: 33233793]
13. Durmuş Sarıkahya S, Terzi A, Kanbay Y, Çınar Özbay S. The mediating effect of social networks on the impact of health perceptions on the quality of life in older adults. Healthcare (Basel) 2025;13(2):122 [FREE Full text] [doi: 10.3390/healthcare13020122] [Medline: 39857150]
14. The Lancet. Reimagining long-term care. Lancet 2020;396(10259):1307 [FREE Full text] [doi: 10.1016/S0140-6736(20)32202-9] [Medline: 34338197]
15. Evans IEM, Martyr A, Collins R, Brayne C, Clare L. Social isolation and cognitive function in later life: a systematic review and meta-analysis. J Alzheimers Dis 2019;70(s1):S119-S144 [FREE Full text] [doi: 10.3233/JAD-180501] [Medline: 30372678]
16. Manjunath J, Manoj N, Alchalabi T. Interventions against social isolation of older adults: a systematic review of existing literature and interventions. Geriatrics (Basel) 2021;6(3):82 [FREE Full text] [doi: 10.3390/geriatrics6030082] [Medline: 34449636]
17. Phang JK, Kwan YH, Yoon S, Goh H, Yee WQ, Tan CS, et al. Digital intergenerational program to reduce loneliness and social isolation among older adults: realist review. JMIR Aging 2023;6:e39848 [FREE Full text] [doi: 10.2196/39848] [Medline: 36598801]
18. Hajek A, Volkmar A, König HH. Prevalence and correlates of loneliness and social isolation in the oldest old: a systematic review, meta-analysis and meta-regression. Soc Psychiatry Psychiatr Epidemiol 2025;60(5):993-1015. [doi: 10.1007/s00127-023-02602-0] [Medline: 38102477]
19. Chaudhury H, Cooke HA, Cowie H, Razaghi L. The influence of the physical environment on residents with dementia in long-term care settings: a review of the empirical literature. Gerontologist 2018;58(5):e325-e337. [doi: 10.1093/geront/gnw259] [Medline: 28329827]
20. Zhong S, Curran L, Yang H, Hahm Y, Lee C. Understanding intergenerational interactions between older adults and children: expert opinions about the roles of community environments. J Appl Gerontol 2025:7334648251347337 [FREE Full text] [doi: 10.1177/07334648251347337] [Medline: 40499168]
21. Dewan N, Sommerlad A, Chapman H, Banerjee S, Corazzini K, Edvardsson D, et al. Assessing social connection for long-term care home residents: systematic review using COnsensus-based Standards for the selection of health Measurement INstruments guidelines. Alzheimers Dement (N Y) 2024;10(3):e12492. [doi: 10.1002/trc2.12492] [Medline: 39104764]
22. Lawton M, Nahemow L. Ecology and the aging process. In: Eisdorfer C, Lawton MP, editors. The Psychology of Adult Development and Aging. Washington, DC: American Psychological Association; 1973:619-674.
23. van Vianen AEM. Person-environment fit: a review of its basic tenets. Annu Rev Organ Psychol Organ Behav 2018;5(1):75-101. [doi: 10.1146/annurev-orgpsych-032117-104702]
24. Barnett-Page E, Thomas J. Methods for the synthesis of qualitative research: a critical review. BMC Med Res Methodol 2009;9:59 [FREE Full text] [doi: 10.1186/1471-2288-9-59] [Medline: 19671152]
25. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008;8:45 [FREE Full text] [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
26. Chapman H, Bethell J, Dewan N, Liougas MP, Livingston G, McGilton KS, et al. Social connection in long-term care homes: a qualitative study of barriers and facilitators. BMC Geriatr 2024;24(1):857 [FREE Full text] [doi: 10.1186/s12877-024-05454-8] [Medline: 39438797]
27. Pomeroy SH, Scherer Y, Runkawatt V, Iamsumang W, Lindemann J, Resnick B. Person—environment fit and functioning among older adults in a long-term care setting. Geriatr Nurs 2011;32(5):368-378. [doi: 10.1016/j.gerinurse.2011.07.002] [Medline: 22059240]
28. Moher D, Liberati A, Tetzlaff J, Altman D, PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
29. Tong A, Flemming K, McInnes E, Oliver S, Craig J. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. BMC Med Res Methodol 2012;12:181 [FREE Full text] [doi: 10.1186/1471-2288-12-181] [Medline: 23185978]
30. Hannes K, Lockwood C. Synthesizing qualitative research: choosing the right approach. In: Synthesizing Qualitative Research. Hoboken, NJ: Wiley-Blackwell; 2011:1-14.
31. The Joanna Briggs Institute. The Joanna Briggs Institute Critical Appraisal tools for Use in JBI Systematic Reviews: Checklist for Qualitative Research. Australia: The Joanna Briggs Institute; 2017.

XSL•FO
RenderX

32. Lincoln YS, Guba EG. Naturalistic Inquiry. Newbury Park, CA: Sage Publications; 1985.
33. Korstjens I, Moser A. Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. Eur J Gen Pract 2018;24(1):120-124 [FREE Full text] [doi: 10.1080/13814788.2017.1375092] [Medline: 29202616]
34. Olmos-Vega FM, Stalmeijer RE, Varpio L, Kahlke R. A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. Med Teach 2022:1-11 [FREE Full text] [doi: 10.1080/0142159X.2022.2057287] [Medline: 35389310]
35. Bartlett R. "You can get in alright but you can't get out": social exclusion and men with dementia in nursing homes: insights from a single case study. Qual Ageing 2007;8(2):16-26. [doi: 10.1108/14717794200700009]
36. Buckley C, McCarthy G. An exploration of social connectedness as perceived by older adults in a long-term care setting in Ireland. Geriatr Nurs 2009;30(6):390-396. [doi: 10.1016/j.gerinurse.2009.09.001] [Medline: 19963148]
37. Cook G, Brown-Wilson C, Forte D. The impact of sensory impairment on social interaction between residents in care homes. Int J Older People Nurs 2006;1(4):216-224. [doi: 10.1111/j.1748-3743.2006.00034.x] [Medline: 20925766]
38. Goll JC, Charlesworth G, Scior K, Stott J. Barriers to social participation among lonely older adults: the influence of social fears and identity. PLoS One 2015;10(2):e0116664 [FREE Full text] [doi: 10.1371/journal.pone.0116664] [Medline: 25706933]
39. Li D, Bai G. Where is home: analysis of social isolation among disabled elderly in nursing home? Based on the survey of nursing home in city C. Acad J Zhongzhou 2020 Aug;08:67-72 [FREE Full text] [doi: 10.3969/j.issn.1003-0751.2020.08.010]
40. Li L. A study on group work intervention for social isolation among semi-disabled older adults in elderly care institutions: a case study of institution X in city L. Wanfang Data. 2022. URL: https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyNDAxMDkSCUQwMjg2NjA5NBoIdGcyc2kxeGs%3D [accessed 2026-03-05]
41. Mu G, Yin S. “Respectful care for the elderly”: challenges and solutions in building elderly-friendly nursing homes. Jianghuai Tribune 2023 Jun;06:94-102 [FREE Full text]
42. Neves BB, Sanders A, Kokanović R. "It's the worst bloody feeling in the world": experiences of loneliness and social isolation among older people living in care homes. J Aging Stud 2019;49:74-84. [doi: 10.1016/j.jaging.2019.100785] [Medline: 31229221]
43. Nielson L, Wiles J, Anderson A. Social exclusion and community in an urban retirement village. J Aging Stud 2019;49:25-30. [doi: 10.1016/j.jaging.2019.03.003] [Medline: 31229215]
44. Pirhonen J, Tiilikainen E, Pietilä I. Ruptures of affiliation: social isolation in assisted living for older people. Age Soc 2018;38(9):1868-1886. [doi: 10.1017/s0144686x17000289]
45. Wan K. An intervention study of behavioral activation therapy to improve social isolation of elderly people in nursing institutions. Master. 2023. URL: https://link.cnki.net/doi/10.27241/d.cnki.gnjgu.2023.000538 [accessed 2026-02-27]
46. Wei Z. A study on the status of social isolation and influencing factors among elderly diabetic patients. Wanfang Data. 2023. URL: https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyNDAxMDkSCUQwMzEyMzkyORoIdGcyc2kxeGs%3D [accessed 2026-02-27]
47. Zhang Y. Study on social work practice of reducing interpersonal alienation among the semi-disabled elderly in institutions. Master. 2021. URL: https://link.cnki.net/doi/10.27456/d.cnki.gyndu.2021.000044 [accessed 2026-02-27]
48. Zhu D. A practical study on overcoming interpersonal barriers for semi-disabled elderly people in nursing institutions-take the -walking together and enjoying silver age? Group as an example. Master. 2023. URL: https://link.cnki.net/doi/10.27036/d.cnki.ggxsu.2023.001980 [accessed 2026-02-27]
49. Buckinx F, Peyrusqué E, Kergoat M, Aubertin-Leheudre MJ. Reference standard for the measurement of loss of autonomy and functional capacities in long-term care facilities. J Frailty Aging 2023;12(3):236-243. [doi: 10.14283/jfa.2023.4] [Medline: 37493385]
50. Goll JC, Charlesworth G, Scior K, Stott J. Correction: Barriers to social participation among lonely older adults: the influence of social fears and identity. PLoS One 2018;13(7):e0201510 [FREE Full text] [doi: 10.1371/journal.pone.0201510] [Medline: 30044856]
51. Chew SY. Loneliness experience among cognitively-intact elderly nursing home residents in Singapore—an exploratory mixed methods study. Arch Gerontol Geriatr 2022;98:104572. [doi: 10.1016/j.archger.2021.104572] [Medline: 34808441]
52. Zhang D, Lu Q, Li L, Wang X, Yan H, Sun Z. Loneliness in nursing homes: a qualitative meta-synthesis of older people's experiences. J Clin Nurs 2023;32(19-20):7062-7075. [doi: 10.1111/jocn.16842] [Medline: 37605069]
53. Huang PH, Wang SY, Hu SH, Chuang YH. Older residents' perceptions of loneliness in long-term care facilities: a qualitative study. Int J Ment Health Nurs 2022;31(3):601-610. [doi: 10.1111/inm.12979] [Medline: 35118782]
54. Smith CB, Wong KLY, To-Miles F, Dunn S, Gregorio M, Wong L, et al. Exploring experiences of loneliness among Canadian long-term care residents during the COVID-19 pandemic: a qualitative study. Int J Older People Nurs 2023;18(1):e12509 [FREE Full text] [doi: 10.1111/opn.12509] [Medline: 36347829]
55. Wijesiri HSMSK, Samarasinghe K, Edberg AK. Loneliness among older people living in care homes in Sri Lanka. Int J Older People Nurs 2019;14(4):e12253. [doi: 10.1111/opn.12253] [Medline: 31274242]
56. Mansfield L, Victor C, Meads C, Daykin N, Tomlinson A, Lane J, et al. A conceptual review of loneliness in adults: qualitative evidence synthesis. Int J Environ Res Public Health 2021 Nov 02;18(21):11522 [FREE Full text] [doi: 10.3390/ijerph182111522] [Medline: 34770035]

XSL•FO
RenderX

57. Paque K, Bastiaens H, Van Bogaert P, Dilles T. Living in a nursing home: a phenomenological study exploring residents' loneliness and other feelings. Scand J Caring Sci 2018;32(4):1477-1484. [doi: 10.1111/scs.12599] [Medline: 30070385]
58. Maes M, Qualter P, Lodder GMA, Mund M. How (not) to measure loneliness: a review of the eight most commonly used scales. Int J Environ Res Public Health 2022;19(17):10816 [FREE Full text] [doi: 10.3390/ijerph191710816] [Medline: 36078540]
59. Sundström M, Edberg A, Rämgård M, Blomqvist K. Encountering existential loneliness among older people: perspectives of health care professionals. Int J Qual Stud Health Well-being 2018;13(1):1474673 [FREE Full text] [doi: 10.1080/17482631.2018.1474673] [Medline: 29869590]
60. Li Y, Li Z, Zhang Y. A qualitative study on nursing home care experience and needs of disabled elderly and visiting family members. Chinese Primary Health Care 2022;36(07):112-115. [doi: 10.3969/j.issn.1001-568X.2022.07.0033]
61. Tak SH, Kedia S, Tongumpun TM, Hong SH. Activity engagement: perspectives from nursing home residents with dementia. Educ Gerontol 2015;41(3):182-192 [FREE Full text] [doi: 10.1080/03601277.2014.937217] [Medline: 25489122]
62. Bedin MG, Droz-Mendelzweig M, Chappuis M. Caring for elders: the role of registered nurses in nursing homes. Nurs Inq 2013;20(2):111-120. [doi: 10.1111/j.1440-1800.2012.00598.x] [Medline: 22487323]
63. Bonifas RP, Simons K, Biel B, Kramer C. Aging and place in long-term care settings: influences on social relationships. J Aging Health 2014;26(8):1320-1339. [doi: 10.1177/0898264314535632] [Medline: 25502244]
64. Ndulue C, Oyebode O, Iyer RS, Ganesh A, Ahmed SI, Orji R. Personality-targeted persuasive gamified systems: exploring the impact of application domain on the effectiveness of behaviour change strategies. User Model User-adapt Interact 2022;32(1-2):165-214 [FREE Full text] [doi: 10.1007/s11257-022-09319-w] [Medline: 35281337]
65. Caspar S, O'Rourke N. The influence of care provider access to structural empowerment on individualized care in long-term-care facilities. J Gerontol B Psychol Sci Soc Sci 2008;63(4):S255-S265. [doi: 10.1093/geronb/63.4.s255] [Medline: 18689775]
66. Hedman M, Häggström E, Mamhidir A, Pöder U. Caring in nursing homes to promote autonomy and participation. Nurs Ethics 2019;26(1):280-292. [doi: 10.1177/0969733017703698] [Medline: 28425315]
67. Beogo I, Sia D, Collin S, Phaelle Gedeon A, Louismé MC, Ramdé J, et al. Strengthening social capital to address isolation and loneliness in long-term care facilities during the COVID-19 pandemic: systematic review of research on information and communication technologies. JMIR Aging 2023;6:e46753 [FREE Full text] [doi: 10.2196/46753] [Medline: 37578824]
68. Davies S, Laker S, Ellis L. Promoting autonomy and independence for older people within nursing practice: a literature review. J Adv Nurs 1997;26(2):408-417. [doi: 10.1046/j.1365-2648.1997.1997026408.x] [Medline: 9292377]
69. Nakou A, Dragioti E, Bastas NS, Zagorianakou N, Kakaidi V, Tsartsalis D, et al. Loneliness, social isolation, and living alone: a comprehensive systematic review, meta-analysis, and meta-regression of mortality risks in older adults. Aging Clin Exp Res 2025;37(1):29. [doi: 10.1007/s40520-024-02925-1] [Medline: 39836319]

Abbreviations

ENTREQ: Enhancing Transparency in Reporting the Synthesis of Qualitative Research
JBI: Joanna Briggs Institute
LTC: long-term care
LTCF: long-term care facility
P-E fit: person-environment fit
PICOS: population, intervention, context, outcome, study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by Y Jiang; submitted 27.May.2025; peer-reviewed by Q Lu, J Ji; comments to author 04.Sep.2025; revised version received 16.Feb.2026; accepted 16.Feb.2026; published 17.Apr.2026.

Please cite as:
Li W, Mei Z, Tu W, Song Y, Zhao Y, Bai Y, Xu G
"Living Inside the Walls": Systematic Review and Qualitative Meta-Synthesis
JMIR Aging 2026;9:e78142
URL: https://aging.jmir.org/2026/1/e78142
doi:10.2196/78142
PMID:25642774

©Weitong Li, Ziqi Mei, WenJing Tu, Yulei Song, Yayi Zhao, Yamei Bai, Guihua Xu. Originally published in JMIR Aging (https://aging.jmir.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any

XSL•FO
RenderX

medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

e-Counseling for Fall Prevention in Older Adults: Scoping Review

Lina Safarina[1*], M.Kep; Cecep Eli Kosasih[2*], S.Kp, MNS, PhD; Iqbal Pramukti[2*], BScN, MSc, PhD; Nucki Nursjamsi[2*], MSc, MD

[1]Department of Nursing, Faculty of Nursing, Universitas Achmad Yani, Cimahi, West Java, Indonesia

[2]Department of Nursing, Faculty of Nursing, Universitas Padjadjaran, Sumedang, Indonesia

[*]all authors contributed equally

Corresponding Author:
Lina Safarina, M.Kep
Department of Nursing, Faculty of Nursing
Universitas Achmad Yani
Jl. Terusan Jenderal Sudirman, Cimahi, Jawa Barat
Cimahi, West Java, 40525
Indonesia
Phone: 62 6656190
Email: lina22006@mail.unpad.ac.id

Abstract

Background: Falls are a major contributor to injury, disability, and mortality among older adults worldwide. Digital health–based e-counseling has emerged as a scalable strategy to support fall prevention by addressing physical, behavioral, and cognitive risk factors.

Objective: This scoping review aimed to map the scope, characteristics, and outcomes of e-counseling interventions for fall prevention among adults aged 60 years or older.

Methods: Guided by Arksey and O'Malley's framework and reported in accordance with PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews), we searched PubMed, ScienceDirect, CINAHL, and ProQuest for studies published between January 2019 and May 2025. Eligible studies evaluated digital or telehealth-based counseling interventions for fall prevention. Data were synthesized descriptively, and methodological quality was appraised using the Mixed Methods Appraisal Tool.

Results: A total of 13 studies were included, comprising randomized controlled trials, feasibility studies, qualitative research, and retrospective cohorts conducted across Asia, Europe, and North America. e-Counseling modalities included virtual reality, mobile health application, teleconsultations, and wearable-triggered feedback systems. Reported outcomes included reductions in fall incidence ranging from 15% to 90.9%, improvements in balance and mobility, increased fall risk awareness, and enhanced self-efficacy. Interventions incorporating cultural tailoring and digital literacy support demonstrated higher engagement and adherence.

Conclusions: e-Counseling appears to be a promising and scalable approach for fall prevention in older adults, with benefits across physical and behavioral domains. Future research should focus on long-term effectiveness, equitable access in low- and middle-income settings, and integration into routine health care systems.

(JMIR Aging 2026;9:e78444) doi:10.2196/78444

KEYWORDS

e-counseling; fall prevention; older adults; digital health; telehealth; scoping review

Introduction

The rapid aging of the global population is transforming societal structures and health care demands worldwide. The global population aged 60 years and older is expected to double, reaching 2.1 billion by 2050, with 16% of the world's population being 60 years or older by 2050 [1]. In Indonesia, 26 million people are aged 60 years and older, accounting for 9.8% of the total population [2]. The population of older adults is expected to reach 48 million by 2045, accounting for 15% of the total population [3]. This significant growth in the older adult population underscores the increasing demographic importance of this age group and the urgent need for attention to their needs and health challenges.

XSL•FO
RenderX

Falls are one of the leading causes of morbidity and mortality among older adults, often resulting in severe physical injuries, psychological distress, and increased health care costs [4]. Globally, it is estimated that 646,000 older adults die from falls each year, making it the second leading cause of accidental deaths worldwide [4]. In the United States, about 36 million falls are reported among older adults each year, leading to more than 32,000 fatalities [5]. In Indonesia, falls are also a major health issue, with approximately 23% of older adults experiencing at least 1 fall annually, contributing to significant morbidity and health care costs [6,7]. Contributing factors to the increased risk of falls include age-related physical decline, such as decreased muscle strength and balance, as well as chronic health conditions such as arthritis, visual impairments, and cognitive decline [8,9]. Furthermore, polypharmacy, or the use of multiple medications, can increase the likelihood of dizziness and instability, further elevating the risk of falls [10,11]. These alarming statistics underscore the urgent need for targeted interventions to prevent falls and promote the overall health and safety of the older adult population.

Falls among older adults have significant physical, emotional, and economic consequences, making them a critical public health concern. Physically, falls are the leading cause of injury and injury-related mortality in this demographic, accounting for approximately 646,000 deaths globally each year [4]. Survivors often experience serious injuries, including fractures and head trauma experience, leading to long-term disabilities and increased dependence on caregivers [12,13]. Emotionally, many older adults develop a fear of falling again, which can result in reduced physical activity, social isolation, and a diminished quality of life [14]. Economically, falls impose a substantial burden on health care systems, with US medical costs for fall-related injuries exceeding US $50 billion annually [5]. This financial strain affects not only individuals but also families and health care resources, highlighting the urgent need for effective fall prevention strategies to mitigate these impacts [15,16].

Various preventive measures have been implemented to address this critical issue. Traditional strategies, such as balance and strength training interventions, have shown effectiveness in reducing fall risk by improving physical capabilities [17]. Home safety assessments and modifications, including removing tripping hazards and improving lighting, are also essential strategies [18]. Additionally, community-based interventions that combine education and physical activity, such as Tai Chi classes, have been proven to enhance balance and reduce falls among older adults [19]. However, these interventions are not always accessible, particularly for older adults with mobility limitations or those living in rural and underserved areas [18]. In response, digital health innovations, particularly e-counseling, have emerged as promising tools for delivering tailored fall prevention strategies in a more scalable and accessible format.

In this context, e-counseling refers to digitally delivered, personalized counseling provided via telehealth platforms, mobile applications, or online systems to support behavior change e-counseling refers to digitally delivered, personalized counseling provided via telehealth platforms, mobile application, or online systems to support health behavior change. In fall prevention, e-counseling facilitates individualized goal setting, progress monitoring, and behavioral coaching grounded in self-efficacy theory [20]. Evidence suggests that e-counseling can improve fall risk awareness, adherence to preventive behaviors, and engagement in balance and mobility exercises [21,22]. By overcoming barriers related to mobility, distance, and resource constraints, e-counseling offers a scalable alternative to traditional in-person interventions. However, evidence remains fragmented regarding long-term sustainability, effectiveness across diverse populations, and integration into existing health care systems, highlighting the need for a comprehensive scoping review [23-26].

Despite its growing application, the evidence base on e-counseling for fall prevention in older adults remains fragmented [27]. Prior reviews have largely focused on in-person interventions or general eHealth approaches, with limited synthesis specifically addressing e-counseling modalities and outcomes. Additionally, there is a paucity of evidence on how these interventions are tailored for diverse populations, integrated into clinical practice, and sustained over time. Given these gaps, a scoping review is warranted to map the scope, characteristics, and outcomes of existing e-counseling interventions aimed at fall prevention among older adults. Therefore, this scoping review aims to (1) map the types and delivery modes of e-counseling interventions for fall prevention; (2) summarize reported physical, cognitive, and behavioral outcomes; and (3) identify research gaps to inform future intervention development and implementation.

Methods

Study Design

This study employed the methodological framework established by Arksey and O'Malley for conducting a scoping review [28] used, which consisted of 5 essential steps. First, the research question was clearly and objectively identified. Second, relevant articles were selected to ensure the study's focus. Third, related literature was extracted from the identified articles for further analysis. Fourth, the data collected from these sources were organized, summarized, and analyzed. Finally, the findings were reported, as outlined in the framework [28]. The review also adhered to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist to ensure comprehensive and replicable reporting [29].

Research Question and Framework

The review was guided by the population, concept, and context framework: population—older adults aged ≥60 years at risk of falls; concept—e-counseling interventions for fall prevention, including digital, telehealth, or app-based personalized counseling; and context—health care or community-based settings where digital interventions are implemented. The central research question was as follows: *What is the current scope of evidence regarding the use of e-counseling to prevent falls in older adults?*

XSL•FO
RenderX

Eligibility Criteria

Studies were included if they met the following criteria: (1) population—participants aged 60 years or older with a focus on fall prevention; (2) intervention—e-counseling interventions delivered through telehealth platforms, mobile application, video consultations, or other digital formats; (3) outcomes—reports on fall rates, balance, mobility, adherence, or related physical and cognitive outcomes; (4) study types—peer-reviewed empirical studies (quantitative, qualitative, or mixed methods), including randomized controlled trials (RCTs), quasi-experimental studies, longitudinal studies, and cohort studies; (5) timeframe—published between January 1, 2019, and May 31, 2025; and (6) language—English. Studies were excluded if they focused solely on in-person counseling, targeted populations other than older adults, or lacked full-text availability.

Information Sources and Search Strategy

Literature searches for this study were conducted using PubMed, ScienceDirect, CINAHL, and ProQuest in May 2025. The search strategy was designed to provide a comprehensive overview of e-counseling interventions for fall prevention among older adults. Keywords were aligned with Medical Subject Headings (MeSH) and incorporated Boolean operators “OR” and “AND” to refine the search results. The specific keywords used included “e-counseling,” “fall prevention,” “older adults” OR “elderly,” “telehealth” OR “digital counseling.” This systematic approach ensured that relevant literature was thoroughly explored to inform the study’s objectives.

Identification and Selection of Literature

The literature search was conducted independently through trusted databases. The retrieved articles were then analyzed, with differences and duplicates being removed. The process of searching for and selecting articles was illustrated using the PRISMA-ScR flowchart [30].

Data Extraction

Data were collected and organized in Microsoft Word in the form of a table containing several components: article title, authors, publication year, research objectives, study design, study location, type and delivery mode of e-counseling intervention, technologies or platforms used, fall-related outcomes (eg, fall rates, balance, and mobility), cognitive or behavioral outcomes (eg, adherence and self-efficacy), and key findings. Two reviewers performed data extraction independently to ensure consistency and accuracy.

Quality Appraisal

While scoping reviews do not typically require formal quality evaluation, this review incorporated a structured assessment to better understand the methodological soundness of the included studies. To achieve this, each study was appraised using the 2018 version of the Mixed Methods Appraisal Tool (MMAT), with consideration given to its specific research design. Two independent reviewers conducted the appraisal, examining elements such as the clarity of the research aim, the suitability of the chosen methods, the completeness of reported outcomes, and potential sources of bias. Importantly, this evaluation was not used to determine study inclusion or exclusion but served to enhance the interpretation of findings and highlight methodological gaps that could guide future investigations.

Data Synthesis and Analysis

A descriptive and thematic synthesis approach was applied. Data were first categorized based on study design, intervention type, and outcome domains. Themes were then identified around the delivery formats (eg, telehealth and mobile apps), effectiveness (eg, fall reduction and mobility improvement), and implementation factors (eg, adherence and accessibility). Findings were summarized narratively, supported by tabulated evidence to illustrate patterns, strengths, and gaps.

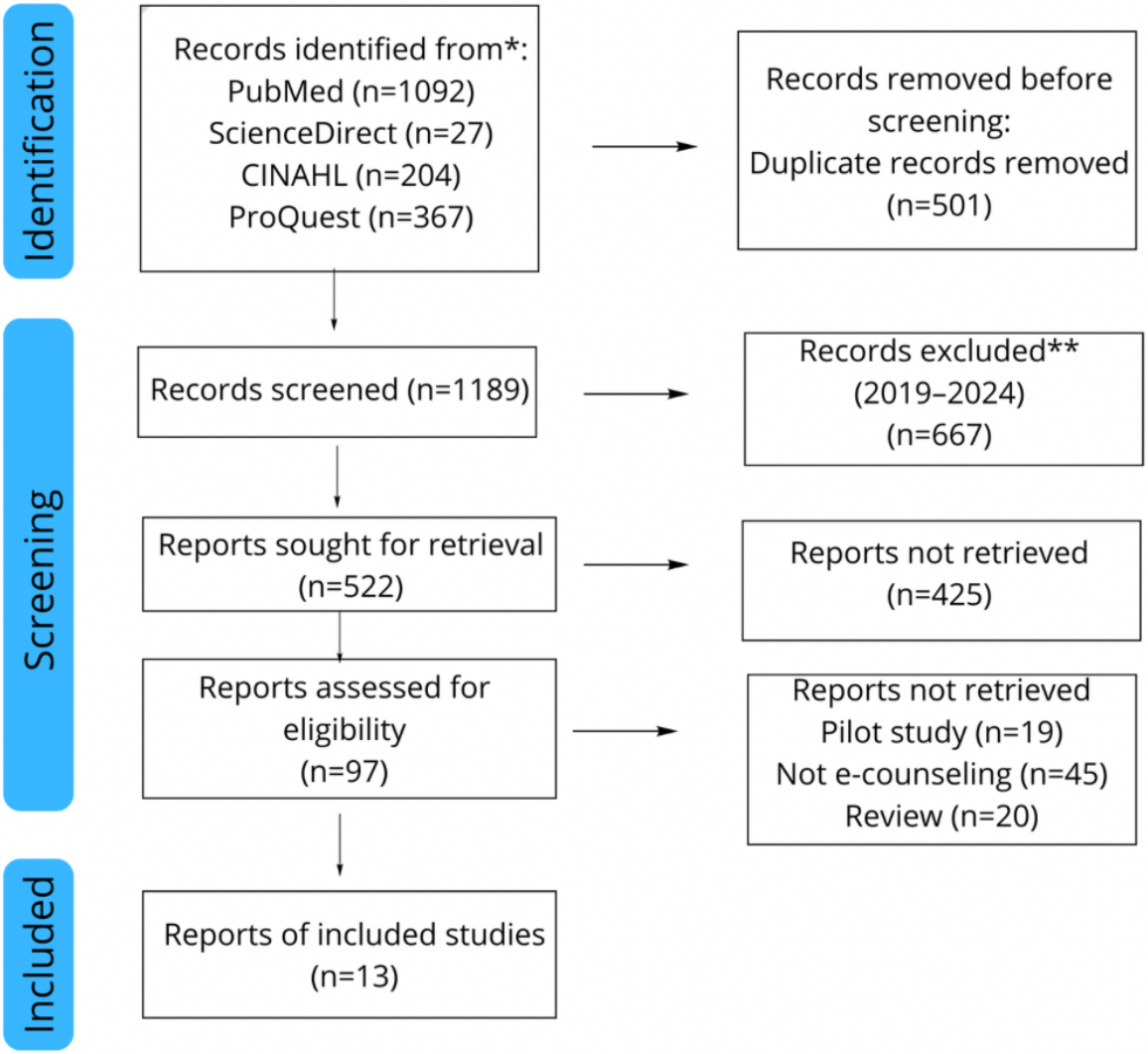
Results

Search Results

The flow of study selection is outlined through the PRISMA-ScR flow diagram, detailing each phase of the review process. An initial search across PubMed, ScienceDirect, CINAHL, and ProQuest yielded 1690 records. After eliminating 501 (29.6%) duplicates, 1189 (70.4%) titles and abstracts were screened. The remaining 97 articles underwent full-text review. During this phase, 84 studies (86.6%) were excluded: 19 (22.6%) were pilot trials, 45 (53.6%) did not involve e-counseling, and 20 (23.8%) were review papers. In the end, 13 studies (13.4%) were included in the final analysis. The study selection process is presented in the PRISMA-ScR flowchart (Figure 1; [30,31]).

XSL•FO
RenderX

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) flow diagram illustrating the study selection process for this scoping review. n indicates the number of records at each stage, and N indicates the total number of records identified. The figure is adapted from Moher et al [31]).



Findings

Characteristics of Included Studies

The included studies demonstrated a diverse range of research designs. RCTs [32,33] offered strong empirical evidence on intervention effectiveness. Pilot and feasibility studies [34,35] assessed initial usability and participant acceptance. Qualitative designs [36,37] provided insight into user perceptions, while retrospective cohort analyses [38,39] and observational designs [40] reflected real-world outcomes. Thiamwong et al [41] applied a cluster RCT to evaluate community-based peer-led interventions. The geographical distribution of the studies reflects global interest in digital fall prevention. Studies were conducted in Asia (Hong Kong, Thailand, and Indonesia), Europe (Germany and the United Kingdom), and North America (the United States). This variety shows the adaptability of e-counseling across different health care systems and cultural contexts.

Delivery modes ranged from immersive virtual reality (VR) [32], app-based learning modules [42], and artificial intelligence (AI)–assisted mobile health (mHealth) apps [39] to telehealth-facilitated interventions [33,43]. Peer-supported and community-led interventions [41] and hybrid models that transitioned from in-person to digital [34] highlight the flexibility of intervention formats. Wearable-triggered interventions, such as sensor belts [44], were also represented. The platforms used included tablet application [40], full-immersion VR systems [32], wearable smart belts [44], and wrist-worn fitness trackers [36]. mHealth application [39] and telehealth delivery models [43] were common. Digital literacy training [34] was crucial in improving accessibility among older adults.

Quality Appraisal Results

The pilot RCTs by Ip et al [32] and Arena et al [33] met most MMAT criteria, demonstrating well-defined aims, suitable randomization methods, and consistent outcome measures. Both studies had clear intervention protocols and provided sufficient follow-up data, although sample sizes were relatively small, limiting generalizability.

Dajpratham et al [38] and Alves et al [39], using retrospective designs, demonstrated strong documentation of real-world outcomes and substantial sample sizes. However, their reliance

XSL•FO
RenderX

on previously collected data posed risks of selection and reporting biases, particularly in the absence of control groups or standardized assessment timelines. Community-based interventions such as those conducted by Nazhira et al [42] and Thiamwong et al [41] provided robust insights into localized fall prevention strategies. While their designs were appropriate for evaluating feasibility and community relevance, some limitations were noted in participant blinding and long-term outcome measurement.

Feasibility and observational studies by Hladek et al [34], Jacobson et al [35], and Bajdek et al [40] generally adhered to MMAT standards, showing clearly defined objectives and consistent methods. However, several studies lacked comprehensive reporting on confounding variables and did not always include detailed rationales for sample size selection.

Qualitative studies, including those by Ojo et al [36] and Tang et al [37], scored highly on relevance and depth of data interpretation, effectively capturing user experiences and motivational factors. Nevertheless, transparency regarding coding frameworks and researcher reflexivity was limited in some cases.

The intervention led by Ritchey [43], though pilot in scope, aligned with MMAT criteria for interventional research, showcasing clear protocols and outcome metrics. Similarly, the real-world adaptive study by Kinosian and Stefanacci [44] demonstrated a high degree of ecological validity and intervention relevance, though it lacked some methodological transparency in data analysis.

Key Findings

Across the included studies, several consistent patterns emerged. First, e-counseling interventions were delivered through diverse digital modalities, including VR, mobile application, telehealth sessions, and wearable-based feedback systems [32]. Despite this variation, effectiveness did not appear to depend solely on technology type. Instead, interventions that incorporated personalized feedback, cultural tailoring, and usability support demonstrated higher engagement and adherence [43]. Second, most studies reported improvements in balance, mobility, and fall-related outcomes, with reported reductions in fall incidence ranging from 15% to 90.9% [39]. Third, behavioral and psychological outcomes—such as increased self-efficacy, confidence, and fall risk awareness—were commonly reported, particularly in interventions that emphasized goal tracking and continuous monitoring. Barriers to implementation included limited digital literacy, unequal access to technology, and resource constraints, especially in underserved populations [44].

Positive effects of these interventions were noted across a range of outcomes, particularly those related to fall risk. Improvements were frequently reported in participants' mobility, balance, and overall fall incidence. For example, the use of VR led to significant gains in both physical coordination and cognitive functioning [32]. Similarly, tools that used AI to assess fall risk enabled more accurate predictions and preventive action [39]. Interventions delivered through peer coaching and app-based platforms demonstrated measurable reductions in fall risk and improvements in physical capability [33,41]. Notably, the remarkable 90.9% reduction in hip fractures reported with the Tango Belt (ActiveProtective Technologies, Inc) highlighted the value of wearable, sensor-based protective devices.

Beyond physical outcomes, psychological and behavioral effects were also evident. Several studies recorded notable increases in participants' confidence in managing fall risk, better executive functioning, and improved knowledge of fall prevention techniques. Interventions that focused on building digital skills and promoting access, especially among vulnerable groups, also led to increased confidence and social connectedness [34]. Key to behavior change was the use of individualized feedback and goal tracking features, which fostered greater involvement and sustained commitment from users. The success of these interventions was also closely tied to how well they were implemented. User satisfaction and acceptance were higher in studies that emphasized intuitive design and meaningful interaction. At the same time, disparities in digital access, technological literacy, and economic resources emerged as ongoing barriers to equal adoption. These challenges point to the need for thoughtful planning when considering how such interventions might be scaled and integrated more broadly. The characteristics of the included studies are summarized in Table 1.

XSL•FO
RenderX

Table 1. Characteristics of the studies and summary of the findings (N=13).

Authors and publication year	Study design	Study location	Type and delivery mode of e-counseling	Technologies or platforms used	Fall-related outcomes	Cognitive or behavioral outcomes	Key findings
Ip et al [32], 2025	Pilot RCT[a]	Community setting, Hong Kong	VR[b] game-based interactive training	Full-immersive VR system	Mobility, balance, walk speed	Executive function, fall efficacy	Greater cognitive-motor improvement vs non-VR group
Dajpratham et al [38], 2025	Retrospective cohort study	Thailand	Personalized digital fall risk assessment	Clinical + digital tools	Fall rate, balance (Timed Up and Go, 30-s chair stand)	Medication adherence	Falls reduced from 80% to 27.4%; sedatives increased risk
Nazhira et al [42], 2025	Community intervention	Indonesia	Video-based balance training	Digital instructional videos	Balance improvement	Knowledge gain	Improved balance and daily confidence
Alves et al [39], 2024	Retrospective real-world analysis	Germany	AI[c]-assisted risk feedback	Longitudinal INdividualized Digital Evaluation of Risk and Aging mHealth[d] app	Fall incidence (6–24 months)	Risk awareness, engagement	Fall Risk Score predicts falls; supports stratified care
Bajdek et al [40], 2024	Observational study	United States	Digital exercise + cognitive assessment	Tablet-based app	Exercise adherence	Cognition, tech readiness	High adherence improves cognition
Ritchey et al [43], 2024	Interventional Veterans Affairs pilot)	United States	Virtual exercise + education	Telehealth platform	Balance, sit-to-stand	Fall self-efficacy	High feasibility and independence
Kinosian et al [44], 2024	Adaptive cohort	United States	Wearable protection system	Tango Belt	90.9% reduction in hip fractures	Adherence	Effective injury prevention
Ojo et al [36], 2024	Qualitative PEER[e] follow-up)	United States	Wearable engagement intervention	Fitbit	Increased physical activity	Motivation	Real-time feedback improves adherence
Tang et al [37], 2024	Qualitative study	United Kingdom	Digital instructional app	Keep On Keep Up (Chinese)	Balance awareness	Engagement, usability	Cultural tailoring improves uptake
Thiamwong et al [41], 2024	Cluster Randomized Controlled Trial	United States	Peer-led digital coaching	BTrackS	Reduced fall risk behavior	Perception, adherence	Effective in low-income groups
Hladek et al [34], 2023	Feasibility study	United States	Hybrid digital transition	Digital literacy + device training	Not directly measured	Confidence, inclusion	Improves access and digital skills
Arena et al [33], 2021	Randomized Controlled Trial	United States	HOP-UP-PT program[f]	In-person + phone	Reduced fall risk	Mobility confidence	Effective home-based intervention
Jacobson et al [35], 2021	Feasibility study	United States	Digital coaching	Online modules	Feasibility outcomes	Usability, acceptance	Scalable and acceptable

[a]RCT: randomized controlled trial.

[b]VR: virtual reality.

[c]AI: artificial intelligence.

[d]mHealth: mobile health.

[e]PEER: Participant Engagement and Empowerment Resources

[f]HOP-UP-PT program: Home-based Older Persons Upstreaming Prevention–Physical Therapy.

Discussion

Principal Findings

This scoping review demonstrates that e-counseling interventions offer meaningful benefits for fall prevention among older adults, with consistent improvements observed in physical outcomes, behavioral engagement, and fall risk awareness across diverse settings.

The study findings demonstrate that e-counseling offers substantial potential in addressing the multifaceted issue of fall prevention among older adults. The diverse modalities used, ranging from VR training to mHealth application and wearable technologies, demonstrate adaptability to various settings and

XSL•FO
RenderX

populations. Notably, the integration of culturally tailored content and user-friendly interfaces appears to enhance engagement and adherence, which are critical factors in the success of fall prevention interventions.

The observed benefits align with prior research indicating that digital interventions can effectively reduce fall incidence and improve balance and mobility. For instance, a previous study demonstrated that e-counseling interventions, whether delivered via telehealth, video consultations, or mobile apps, address many of these barriers by improving accessibility and allowing older adults to engage in fall prevention strategies from the comfort of their homes [45]. However, earlier reviews often focused on in-person interventions, such as group exercise interventions, home visits, or physical therapy, to address fall risk in older adults [46]. Moreover, earlier reviews highlighted issues with adherence, often related to the logistical difficulties of attending in-person sessions consistently [18]. Previous reviews had not emphasized the role of digital tools as strongly, as the technological infrastructure and adoption among older adults were not as widespread [47]. With the increased availability and acceptance of technology, recent studies have shown that digital platforms can provide continuous monitoring and personalized interventions, a feature that was less prominent in earlier studies that focused on more generalized group interventions.

Additionally, previous reviews noted a lack of sustained follow-up, with many interventions being short term and failing to assess the long-term sustainability of fall prevention strategies [48]. Recent findings show that e-counseling, through regular follow-ups and remote monitoring, enhances adherence and promotes long-term engagement in fall prevention activities [49]. This continuous interaction is a marked improvement over older models, which often relied on intermittent or in-person check-ins. Another critical advancement in recent studies is the inclusion of cognitive training and the exploration of how e-counseling can improve not only physical outcomes such as balance and mobility but also cognitive function [50]. Earlier reviews primarily focused on physical outcomes, leaving cognitive aspects underexplored. Recent studies have demonstrated that e-counseling can provide a holistic approach by incorporating both physical and cognitive training into fall prevention strategies, offering a more comprehensive intervention [25]. In summary, compared to previous reviews, the recent findings reflect significant advancements in accessibility, adherence, and the integration of personalized care through digital tools, positioning e-counseling as a scalable and effective intervention for fall prevention among older adults.

Despite the promising findings from existing studies, several research gaps remain in the field of e-counseling for fall prevention among older adults. First, although prior studies demonstrate improvements in knowledge, confidence, and adherence to safety measures, evidence on long-term effects on fall rates and overall health outcomes remains limited. Second, most studies focus on specific populations or geographic regions, limiting generalizability to more diverse groups of older adults with varying health statuses and cultural contexts. Third, the impact of different e-counseling modalities (eg, video calls vs SMS text messaging) on engagement and effectiveness remains underexplored. Finally, although some studies report reductions in fall-related anxiety and increases in physical activity, further research is needed to determine how e-counseling can be integrated into broader health care systems and coordinated with in-person care to support comprehensive fall prevention strategies. Addressing these gaps may enhance the effectiveness and applicability of e-counseling interventions for older adults [24,51].

A multidisciplinary home-based telehealth intervention has also been shown to be feasible and effective in reducing fall risk among older adults, highlighting the importance of integrated and technology-supported approaches in fall prevention programs [52].

Clinical Implications

The results underscore the potential of e-counseling as an integral component of fall prevention interventions in clinical practice. Health care providers should consider integrating e-counseling interventions into routine care for older adults, particularly those facing mobility challenges or residing in remote areas where access to in-person services is limited. By leveraging digital tools, clinicians can offer personalized care plans that adapt to each individual's needs, enhancing adherence to recommended exercises and preventive measures. Moreover, the observed improvements in balance, mobility, and cognitive function suggest that e-counseling not only reduces fall incidence but also contributes to the overall well-being of older adults. Clinicians should be proactive in promoting e-counseling options as part of comprehensive fall prevention strategies, particularly in light of the growing acceptance of telehealth among older populations. Additionally, the role of regular follow-ups and continuous monitoring emphasizes the need for a structured approach to e-counseling. Health care professionals should prioritize ongoing engagement with patients through scheduled virtual check-ins, ensuring that individuals remain motivated and accountable in their fall prevention efforts. This ongoing support can foster greater confidence and satisfaction, leading to more effective long-term outcomes.

Limitations

Several limitations must be acknowledged. First, the eligibility criteria restricted the inclusion of articles to those published in English and within the last 5 years (2019-2025), potentially excluding relevant studies in other languages or earlier research that could provide valuable insights into e-counseling for fall prevention among older adults. Additionally, the selection of articles was based on predefined criteria, which could introduce selection bias; focusing on specific databases may have led to the omission of studies published elsewhere or gray literature, affecting the review's comprehensiveness. Furthermore, the included studies varied in their designs, methodologies, and outcome measures, complicating the synthesis of findings and limiting the ability to draw definitive conclusions regarding the effectiveness of e-counseling interventions.

Conclusions

This scoping review shows that e-counseling represents a promising and scalable approach to fall prevention among older adults, particularly when interventions are personalized,

XSL•FO
RenderX

culturally appropriate, and supported by ongoing monitoring. These findings can inform the development of integrated digital fall prevention strategies within health care systems and support policy initiatives aimed at promoting healthy aging. Future studies should prioritize long-term effectiveness, equitable access in low- and middle-income settings, and sustainable integration into routine care.

Acknowledgments

This publication charge is funded by Unpad through the Indonesian Endowment Fund for Education (LPDP) on behalf of the Indonesian Ministry of Higher Education, Science and Technology and managed under the EQUITY Program (Contract No. 4303/ B3/DT.03.08/2025 and 3927/UN6. RKT/HK.07.00/2025).

Funding

This research and publication were supported by Universitas Padjadjaran.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR checklist.
[DOCX File , 16 KB - aging_v9i1e78444_app1.docx]

References

1. Progress report on the United Nations decade of healthy ageing, 2021-2023. World Health Organization. 2023 Nov 22. URL: https://www.who.int/publications/i/item/9789240079694 [accessed 2026-04-10]
2. Statistik Penduduk Lanjut Usia. Jakarta, Indonesia: Badan Pusat Statistik; 2005.
3. Ageing. World Health Organization. URL: https://www.who.int/health-topics/ageing#tab=tab_1 [accessed 2026-04-12]
4. Falls. World Health Organization. 2021 Apr 26. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2026-04-10]
5. Older adult falls data. Centers for Disease Control and Prevention. 2026 Feb 26. URL: https://www.cdc.gov/falls/data-research/index.html [accessed 2026-04-10]
6. Yeong UY, Tan SY, Yap JF, Choo WY. Prevalence of falls among community-dwelling elderly and its associated factors: a cross-sectional study in Perak, Malaysia. Malays Fam Physician 2016 Apr 30;11(1):7-14 [FREE Full text] [Medline: 28461842]
7. Susilowati IH, Nugraha S, Sabarinah S, Peltzer K, Pengpid S, Hasiholan BP. Prevalence and risk factors associated with falls among community-dwelling and institutionalized older adults in Indonesia. Malays Fam Physician 2020 Mar 18;15(1):30-38 [FREE Full text] [Medline: 32284802]
8. Tricco AC, Thomas SM, Veroniki AA, Hamid JS, Cogo E, Strifler L, et al. Comparisons of interventions for preventing falls in older adults: a systematic review and meta-analysis. JAMA 2017 Nov 07;318(17):1687-1699 [FREE Full text] [doi: 10.1001/jama.2017.15006] [Medline: 29114830]
9. Guirguis-Blake JM, Michael YL, Perdue LA, Coppola EL, Beil TL. Interventions to prevent falls in older adults: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2018 Apr 24;319(16):1705-1716. [doi: 10.1001/jama.2017.21962] [Medline: 29710140]
10. Laflamme L, Monárrez-Espino J, Johnell K, Elling B, Möller J. Type, number or both? A population-based matched case-control study on the risk of fall injuries among older people and number of medications beyond fall-inducing drugs. PLoS One 2015 Mar 27;10(3):e0123390 [FREE Full text] [doi: 10.1371/journal.pone.0123390] [Medline: 25815483]
11. Morin L, Calderon Larrañaga A, Welmer AK, Rizzuto D, Wastesson JW, Johnell K. Polypharmacy and injurious falls in older adults: a nationwide nested case-control study. Clin Epidemiol 2019 Jun 24;11:483-493 [FREE Full text] [doi: 10.2147/CLEP.S201614] [Medline: 31296999]
12. Dellinger A. Older adult falls: effective approaches to prevention. Curr Trauma Rep 2017 Jun;3(2):118-123 [FREE Full text] [doi: 10.1007/s40719-017-0087-x] [Medline: 28845383]
13. Appeadu MK, Bordoni B. Falls and fall prevention in older adults. In: StatPearls [Internet]. Treasure Island, FL: StatPearls Publishing; Jun 04, 2023.
14. Moncada LV, Mire LG. Preventing falls in older persons. Am Fam Physician 2017 Aug 15;96(4):240-247. [Medline: 28925664]
15. Bergland A. Fall risk factors in community-dwelling elderly people. Norsk Epidemiol 2012 Nov 29;22(2):151-164 [FREE Full text] [doi: 10.5324/nje.v22i2.1561]

XSL•FO
RenderX

16. Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults: a review of the literature. Maturitas 2013 May;75(1):51-61. [doi: 10.1016/j.maturitas.2013.02.009] [Medline: 23523272]
17. Sherrington C, Fairhall N, Wallbank G, Tiedemann A, Michaleff ZA, Howard K, et al. Exercise for preventing falls in older people living in the community: an abridged Cochrane systematic review. Br J Sports Med 2020 Aug;54(15):885-891. [doi: 10.1136/bjsports-2019-101512] [Medline: 31792067]
18. Cameron ID, Dyer SM, Panagoda CE, Murray GR, Hill KD, Cumming RG, et al. Interventions for preventing falls in older people in care facilities and hospitals. Cochrane Database Syst Rev 2018 Sep 07;9(9):CD005465 [FREE Full text] [doi: 10.1002/14651858.CD005465.pub4] [Medline: 30191554]
19. Li F, Harmer P, Fisher KJ, McAuley E, Chaumeton N, Eckstrom E, et al. Tai Chi and fall reductions in older adults: a randomized controlled trial. J Gerontol A Biol Sci Med Sci 2005 Feb;60(2):187-194. [doi: 10.1093/gerona/60.2.187] [Medline: 15814861]
20. Sibley KM, Thomas SM, Veroniki AA, Rodrigues M, Hamid JS, Lachance CC, et al. Comparative effectiveness of exercise interventions for preventing falls in older adults: a secondary analysis of a systematic review with network meta-analysis. Exp Gerontol 2021 Jan;143:111151. [doi: 10.1016/j.exger.2020.111151] [Medline: 33186739]
21. Ambrens M, Stanners M, Valenzuela T, Razee H, Chow J, van Schooten KS, et al. Exploring older adults' experiences of a home-based, technology-driven balance training exercise program designed to reduce fall risk: a qualitative research study within a randomized controlled trial. J Geriatr Phys Ther 2023;46(2):139-148. [doi: 10.1519/JPT.0000000000000321] [Medline: 34292258]
22. Pitre V, Sénéchal M, Bouchard DR. Comparing online and in-person delivery of a fall prevention exercise program for older adults. J Aging Phys Act 2023 Sep 23;32(1):55-61. [doi: 10.1123/japa.2023-0070] [Medline: 37741634]
23. Ilali M, Le Berre M, Vedel I, Khanassov V. Telemedicine in the primary care of older adults: a systematic mixed studies review. BMC Prim Care 2023 Jul 20;24(1):152 [FREE Full text] [doi: 10.1186/s12875-023-02085-7] [Medline: 37468871]
24. Duan Y, Shang B, Liang W, Du G, Yang M, Rhodes RE. Effects of eHealth-based multiple health behavior change interventions on physical activity, healthy diet, and weight in people with noncommunicable diseases: systematic review and meta-analysis. J Med Internet Res 2021 Feb 22;23(2):e23786 [FREE Full text] [doi: 10.2196/23786] [Medline: 33616534]
25. Robert C, Erdt M, Lee J, Cao Y, Naharudin NB, Theng YL. Effectiveness of eHealth nutritional interventions for middle-aged and older adults: systematic review and meta-analysis. J Med Internet Res 2021 May 17;23(5):e15649 [FREE Full text] [doi: 10.2196/15649] [Medline: 33999005]
26. Jung YH, Kang SH, Park EC, Jang SY. Impact of the acceptance of disability on self-esteem among adults with disabilities: a four-year follow-up study. Int J Environ Res Public Health 2022 Mar 24;19(7):3874 [FREE Full text] [doi: 10.3390/ijerph19073874] [Medline: 35409553]
27. Moon JY, Kim JH. Association between self-esteem and efficacy and mental health in people with disabilities. PLoS One 2021 Oct 6;16(10):e0257943 [FREE Full text] [doi: 10.1371/journal.pone.0257943] [Medline: 34613985]
28. O'Malley C, Vavoula G, Glew J, Taylor J, Sharples M, Lefrere P, et al. Guidelines for learning/teaching/tutoring in a mobile environment. HAL Authorization. 2005. URL: https://hal.science/hal-00696244 [accessed 2026-04-12]
29. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
30. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64 [FREE Full text] [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]
31. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. PRISMA 2009 flow diagram. Elsevier. 2009. URL: https://legacyfileshare.elsevier.com/promis_misc/prisma-2009-flow-diagram.pdf [accessed 2026-04-12]
32. Ip WK, Soar J, James C, Wang SY, Fong KN. Virtual reality game-based training for preventing falls among community-dwelling older adults with mild cognitive impairment: a pilot randomized control trial study. Virtual Real 2025 Jan 04;29:18. [doi: 10.1007/s10055-024-01084-y]
33. Arena SK, Wilson CM, Boright L, Peterson E. Impact of the HOP-UP-PT program on older adults at risk to fall: a randomized controlled trial. BMC Geriatr 2021 Oct 01;21(1):520 [FREE Full text] [doi: 10.1186/s12877-021-02450-0] [Medline: 34598692]
34. Hladek M, Rubio O, Curriero S, Crews D, DeMarco MM, Brennan D, et al. Empowering frail older adults through digital literacy: a theory and values-based intervention. Innov Aging 2024 Dec;8(Supplement_1):392-393. [doi: 10.1093/geroni/igae098.1276]
35. Jacobson CL, Foster LC, Arul H, Rees A, Stafford RS. A digital health fall prevention program for older adults: feasibility study. JMIR Form Res 2021 Dec 23;5(12):e30558 [FREE Full text] [doi: 10.2196/30558] [Medline: 34837492]
36. Ojo E, Loerzel V, Thiamwong L. Experience of older adults with objectively measured physical activity in fall prevention intervention. Innov Aging 2024 Dec;8(Supplement_1):837-838. [doi: 10.1093/geroni/igae098.2713]
37. Tang Y, Smith S, Burden S, Stanmore E. An exploration of a digital fall prevention program: “keep on keep up” for older adults in a Chinese cultural context. Innov Aging 2024 Dec;8(Supplement_1):330. [doi: 10.1093/geroni/igae098.1078]

XSL•FO
RenderX

38. Dajpratham P, Thitisakulchai P, Pongratanakul R, Prapavanond R, Haridravedh S, Muangpaisan W. Effectiveness of personalized multifactorial fall risk assessment and intervention in reducing fall rates among older adults: a retrospective study. Siriraj Med J 2025 Jan 01;77(1):64-72. [doi: 10.33192/smj.v77i1.271422]
39. Alves SA, Temme S, Motamedi S, Kura M, Weber S, Zeichen J, et al. Evaluating the prognostic and clinical validity of the fall risk score derived from an AI-based mHealth app for fall prevention: retrospective real-world data analysis. JMIR Aging 2024 Dec 04;7:e55681 [FREE Full text] [doi: 10.2196/55681] [Medline: 39631046]
40. Bajdek N, Latham NK, Dishaw M, Farrell S, Shang YV, Pencina KM, et al. Feasibility of a multicomponent digital fall prevention exercise intervention for at-risk older adults. J Frailty Aging 2024;13(4):349-358. [doi: 10.14283/jfa.2024.35] [Medline: 39574254]
41. Thiamwong L, Xie R, Lighthall N, Loerzel V, Park JH, Stout J. Technology-based physio-feedback exercise program (peer) to prevent falls and enhance health equity. Innov Aging 2024 Dec;8(Supplement_1):647-648. [doi: 10.1093/geroni/igae098.2119]
42. Nazhira F, Anggraeni RK, Faradillah KR, Rabia R, Ghani G, Puspadini D, et al. Implementation of digital technology to improve balance at elderly Posyandu. J Kreativitas Pengabdian Masy 2025;8(2). [doi: 10.33024/jkpm.v8i2.17746]
43. Ritchey K. Implementation of a virtual fall prevention program in older veterans. Innov Aging 2024 Dec;8(Supplement_1):219. [doi: 10.1093/geroni/igae098.0708]
44. Kinosian B, Stefanacci R. Reducing serious fall injuries in older adults: results of a multicenter wearable technology study. Innov Aging 2024 Dec;8(Supplement_1):684. [doi: 10.1093/geroni/igae098.2233]
45. Diener J, Rayling S, Bezold J, Krell-Roesch J, Woll A, Wunsch K. Effectiveness and acceptability of e- and m-health interventions to promote physical activity and prevent falls in nursing homes-a systematic review. Front Physiol 2022 May 20;13:894397 [FREE Full text] [doi: 10.3389/fphys.2022.894397] [Medline: 35669573]
46. Sherrington C, Michaleff ZA, Fairhall N, Paul SS, Tiedemann A, Whitney J, et al. Exercise to prevent falls in older adults: an updated systematic review and meta-analysis. Br J Sports Med 2017 Dec;51(24):1750-1758 [FREE Full text] [doi: 10.1136/bjsports-2016-096547] [Medline: 27707740]
47. Ambrens M, Alley S, Oliveira JS, To Q, Delbaere K, Vandelanotte C, et al. Effect of eHealth-delivered exercise programmes on balance in people aged 65 years and over living in the community: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2022 Jun 10;12(6):e051377 [FREE Full text] [doi: 10.1136/bmjopen-2021-051377] [Medline: 35688586]
48. Giordano A, Bonometti GP, Vanoglio F, Paneroni M, Bernocchi P, Comini L, et al. Feasibility and cost-effectiveness of a multidisciplinary home-telehealth intervention programme to reduce falls among elderly discharged from hospital: study protocol for a randomized controlled trial. BMC Geriatr 2016 Dec 07;16(1):209 [FREE Full text] [doi: 10.1186/s12877-016-0378-z] [Medline: 27923343]
49. Racey M, Markle-Reid M, Fitzpatrick-Lewis D, Ali MU, Gagne H, Hunter S, et al. Fall prevention in community-dwelling adults with mild to moderate cognitive impairment: a systematic review and meta-analysis. BMC Geriatr 2021 Dec 10;21(1):689 [FREE Full text] [doi: 10.1186/s12877-021-02641-9] [Medline: 34893027]
50. Shah TM, Weinborn M, Verdile G, Sohrabi HR, Martins RN. Enhancing cognitive functioning in healthly older adults: a systematic review of the clinical significance of commercially available computerized cognitive training in preventing cognitive decline. Neuropsychol Rev 2017 Mar;27(1):62-80. [doi: 10.1007/s11065-016-9338-9] [Medline: 28092015]
51. Chan JK, Klainin-Yobas P, Chi Y, Gan JK, Chow G, Wu XV. The effectiveness of e-interventions on fall, neuromuscular functions and quality of life in community-dwelling older adults: a systematic review and meta-analysis. Int J Nurs Stud 2021 Jan;113:103784. [doi: 10.1016/j.ijnurstu.2020.103784] [Medline: 33120138]
52. Bernocchi P, Giordano A, Pintavalle G, Galli T, Ballini Spoglia E, Baratti D, et al. Feasibility and clinical efficacy of a multidisciplinary home-telehealth program to prevent falls in older adults: a randomized controlled trial. J Am Med Dir Assoc 2019 Mar;20(3):340-346. [doi: 10.1016/j.jamda.2018.09.003] [Medline: 30366759]

Abbreviations

AI: artificial intelligence
MeSH: Medical Subject Headings
mHealth: mobile health
MMAT: Mixed Methods Appraisal Tool
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
RCT: randomized controlled trial
VR: virtual reality

XSL•FO
RenderX

Edited by P Jansons; submitted 02.Jun.2025; peer-reviewed by A Olanlesi-Aliu, C Weerth; comments to author 17.Oct.2025; revised version received 01.Jan.2026; accepted 27.Jan.2026; published 26.May.2026.

Please cite as:
Safarina L, Kosasih CE, Pramukti I, Nursjamsi N
e-Counseling for Fall Prevention in Older Adults: Scoping Review
JMIR Aging 2026;9:e78444
URL: https://aging.jmir.org/2026/1/e78444
doi:10.2196/78444
PMID:

©Lina Safarina, Cecep Eli Kosasih, Iqbal Pramukti, Nucki Nursjamsi. Originally published in JMIR Aging (https://aging.jmir.org), 26.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Use of Wearable Technology for Measuring and Characterizing Sedentary Behavior in People With Mild Cognitive Impairment and Dementia: Systematic Review

Jenny L Wales[1,2], MRES; Chloe Hinchliffe[1], PhD; Ryan Stanley Falck[3], PhD; Silvia Del Din[1,2], PhD; Alison J Yarnall[1,2], PhD; Ríona Mc Ardle[1,2], PhD

[1]Faculty of Medical Sciences, Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom

[2]NIHR Newcastle Biomedical Research Centre, Newcastle upon Tyne, United Kingdom

[3]School of Biomedical Engineering, Djavad Mowafaghian Centre for Brain Health, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Ríona Mc Ardle, PhD
Faculty of Medical Sciences
Translational and Clinical Research Institute
Newcastle University
Biomedical Research Building
Campus for Ageing and Vitality
Newcastle upon Tyne, NE4 5PL
United Kingdom
Phone: 44 1912081410
Email: riona.mcardle@newcastle.ac.uk

Abstract

Background: Sedentary behavior (SB) is a critical, modifiable risk factor for adverse health outcomes. Evidence suggests that SB is higher among individuals with cognitive impairment relative to their cognitively healthy peers. However, the nature and extent of SB across cognitive impairments remains unclear, largely due to the reliance on self-report data and the lack of synthesized evidence from more accurate methodology, such as wearable devices. Wearable device–based methodologies offer a reliable means of capturing SB in real-world settings, circumventing the recall bias inherent to self-report methods. Continuous remote monitoring of SB, via wearable devices, may provide nuanced insights important for understanding SB's contribution to cognitive impairment and health consequences.

Objective: This review aims to synthesize evidence on the volume, patterns, and variability of SB across cognitive impairment and critically appraise the wearable device–based methodology used to capture SB in this population.

Methods: Following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, we searched 5 databases (Embase, MEDLINE, PsycInfo, Scopus, and Web of Science) up to January 2025 for peer-reviewed English-language studies using wearable devices to measure SB in community-dwelling or aged residential care residents aged 50 years or older with cognitive impairment (PROSPERO: CRD42024616523). Study quality was assessed using an adapted version of the National Institute of Health Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies. Data were extracted on SB outcomes (eg, volume, pattern, and variability) and methodological characteristics (eg, device type, placement, SB classification/processing, and its corresponding validation).

Results: From 2824 screened records, 17 studies (2016-2025) were included. Most studies (n=11, 65%) were of "good" quality (scoring ≥5 on bias assessment). Synthesis revealed inconsistent evidence for differences in SB volume across cognitive impairment. However, individuals with dementia consistently exhibited a unique SB pattern, engaging in significantly fewer but longer sedentary bouts than other forms of cognitive impairment and cognitively intact controls. All (n=17, 100%) studies used volume metrics to describe SB, followed by pattern metrics (n=7, 41%); only 1 study reported on SB variability. Methodological appraisal found significant heterogeneity: 13 different device models across 6 body placements were used. Most studies quantified SB using count-based thresholds (counts per minute), which were largely unvalidated in cognitively impaired or older adult populations.

Conclusions: This review found that participants with dementia consistently exhibited a unique pattern of SB compared to other forms of cognitive impairment and healthy controls, while evidence for differences in SB volume was inconsistent. This may

XSL•FO
RenderX

indicate that differences in SB volume are not inherent to dementia pathology but may be mediated by other factors, such as neuropsychiatric symptoms or environmental influences. Furthermore, methodological heterogeneity and unvalidated thresholds were observed throughout the review, highlighting a need for standardized protocols to enhance the validity and clinical applicability of future research.

Trial Registration: PROSPERO CRD42024616523; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024616523

(JMIR Aging 2026;9:e85361) doi:10.2196/85361

KEYWORDS

cognitive impairment; dementia; mild cognitive impairment; sedentary behavior; systematic review; wearable technology

Introduction

Dementia is a leading global cause of disability and dependency, affecting 55 million people worldwide, a figure projected to double by 2050 [1,2]. Dementia is an umbrella term for multiple progressive neurological diseases that impact cognitive function, leading to significant impairments in functional abilities [2]. However, subtle changes in cognition often occur years before dementia diagnosis. Mild cognitive impairment (MCI) is a transition state between healthy cognitive aging and dementia, wherein cognitive performance is below the expected level for age and education level but does not interfere with independence [3,4]. Approximately 5%-15% of individuals with MCI progress to dementia each year, making it a significant risk factor [5]. Identifying modifiable lifestyle behaviors across cognitive impairment (including MCI and dementia) is therefore a critical public health priority, essential for prolonging cognitive health and reducing escalating health care expenditures [6-9].

Sedentary behavior (SB) is recognized as a distinct, modifiable risk factor for cognitive decline [7,10-12]. Although lifestyle interventions have historically targeted physical activity (PA) to promote healthy cognitive aging [13,14], emerging evidence clarifies that PA and SB are not merely opposite ends of a single continuum: individuals may meet PA guidelines yet still accumulate prolonged sedentary time that exacerbates cognitive decline [12,15-17]. This distinction carries particular significance for people with cognitive impairment (ie, MCI or dementia), who often face barriers such as executive dysfunction, apathy, and caregiver dependence that hinder active lifestyles and promote greater engagement in SB [18-21].

SB is defined as any sitting, reclining, or lying behavior, while awake, characterized by an energy expenditure ≤1.5 metabolic equivalent tasks (METs) [22-24], or any “nonupright” activity [24]. A growing body of evidence indicates that people with cognitive impairment engage in significantly more SB compared with their cognitively healthy counterparts [25]. This is quantitatively demonstrated by studies reporting that those with cognitive impairment spend up to 80% (9.7 hours) of their day sedentary [26-28], in contrast to 55% (7.7 hours) reported in the general older adult literature [29]. These quantitative findings are corroborated by qualitative evidence where 1 study found nursing home residents with cognitive impairment spend a large proportion of their day “doing nothing,” resulting in boredom, frustration, and hopelessness [30]. Collectively, these patterns underscore a critical target for intervention. First, excessive sedentary time is independently linked to exacerbated cognitive decline [15,25,27]. Second, the high volume of SB can worsen psychological symptoms commonly experienced in this population, including apathy, depression, and anxiety, thereby further diminishing overall quality of life [31-33]. This can also create an indirect pathway whereby increased SB exacerbates psychological symptoms, which in turn reduce PA engagement, forming a self-reinforcing cycle of SB [21]. Although differences in PA have been observed across cognitive impairment subtypes (eg, MCI and dementia) and severity levels (eg, mild, moderate, and severe) [34,35], it remains unclear whether SB varies similarly. The current evidence base is limited and methodologically heterogeneous [36,37]. Consequently, the precise nature and extent of SB across cognitive impairments remains inconclusive.

Accurately measuring SB is a critical first step toward developing interventions for individuals with cognitive impairment to reduce the associated health risks. The primary assessment methods offer a trade-off between context and precision. Self-report tools (eg, questionnaires) are popular for their practicality and ability to identify activity types but are undermined by recall bias and inaccuracy, which may be exacerbated by cognitive impairment [38-41]. In contrast, device-based measures are often regarded as the “gold standard” for assessing SB volume and pattern [40,42] and are recommended for assessing SB in people with cognitive impairment in free-living environments [40,41,43]. Unlike self-report tools, device-based methods enable detailed analyses of sedentary patterns, making them particularly valuable for detecting subtle behavioral changes over time and well-suited for research aiming to understand the nuanced relationship between SB and cognitive impairment [44,45]. As such, they offer a promising avenue to accurately capture SB in real-world settings with minimal participant burden [37].

Device-based methods can encompass a wide range of technologies, from smart-home systems (eg, indoor positioning systems [46]) to wearable devices (eg, accelerometers [47]). Smart-home technology can capture accurate contextual SB data in an unobtrusive manner [48]; however, it is often constrained to the home environment and requires complex infrastructure and data processing [46,48]. In contrast, wearable devices can capture SB data with high precision at a reduced cost [49]. Crucially, their ability to measure SB in free-living environments, beyond the home, provides the ecological validity necessary to understand SB patterns across all daily contexts in people with cognitive impairment [50].

Despite their advantages, wearable-based methodologies are marked by considerable heterogeneity in the SB literature

XSL•FO
RenderX

[15,51]. Wearable devices primarily assess SB through 2 approaches: by quantifying a lack of movement using accelerometer cut points or detecting body posture [52]. Although posture-based measurement is considered the most accurate method for identifying SB [22,43,52,53], much of the existing literature relies on cut-point thresholds (counts per minute [CPM]). These thresholds vary substantially across studies and can produce markedly different SB estimates [54]. Beyond threshold selection, a growing body of evidence indicates that SB estimates are also influenced by broader methodological factors, including device brand (eg, activPAL and Axivity), body placement (eg, thigh or wrist), and the metrics used to characterize SB (eg, sedentary time and bout length) [44,55-60]. Such methodological inconsistencies and lack of standardization may bias estimates of SB and complicate comparisons across studies. However, whether these methodological variations influence SB estimates in populations with cognitive impairment is yet to be synthesized.

This systematic review aims to synthesize the existing evidence on SB volume, patterns, and variability (fluctuation of SB) and to critically appraise the wearable technology methods and metrics used. By identifying and analyzing the key sources of methodological variation, this work seeks to promote more standardized and appropriate measurement approaches. Addressing these methodological challenges is a critical prerequisite for advancing SB research and translating findings into practical health benefits for populations with cognitive impairment.

Methods

Search Strategy

This review was preregistered on PROSPERO (CRD42024616523) and designed in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [61]. PRISMA 2020 checklist is available in Multimedia Appendix 1. Five databases were used for this search: Embase, MEDLINE, PsycInfo, Scopus, and Web of Science. The search criteria were restricted to studies conducted up to the date of the most recent search (January 2025), with no lower date limit. Keywords relevant to cognitive impairment, wearable digital technology, and SB formed a search string (Multimedia Appendix 2).

Selection Criteria

Articles were screened in accordance with the most generally adopted definitions of SB: any sitting, reclining, or lying behavior, while awake, characterized by an energy expenditure ≤1.5 METs [22-24] or any nonupright activity [24].

Inclusion and Exclusion Criteria

Articles were included if they were (1) peer reviewed; (2) published in the English language before 2025; (3) conducted within community-dwelling or aged residential care settings, including supportive living, assisted living, residential aged care, nursing homes, and care homes; (4) investigated individuals with MCI or dementia, inclusive of all subtypes to capture the full breadth of cognitive impairment presentations [62-65]; (5) involved participant samples with a mean age of ≥50 years to ensure the inclusion of MCI and young-onset dementia [62,63,66,67]; (6) reported SB metrics measured through quantitative wearable devices (eg, accelerometry); and (7) defined SB either in accordance with the established consensus definitions or through an alternative but clearly articulated and appropriate definition.

Articles were excluded if they were (1) published in a language other than English; (2) conducted in acute, in-patient hospital, or palliative care settings (as inactivity in these environments is often medically mandated and does not represent an individual's usual, volitional behavior [68,69]); (3) focused solely on PA or sleep-related metrics; (4) assessed SB through subjective self-report measures without the inclusion of quantitative wearable-derived data; (5) involved children, adolescents, or adults without cognitive impairments, including those with subjective cognitive impairment; (6) included mixed populations without disaggregated SB data for individuals with cognitive impairment; or (7) were not primary research articles such as conference abstracts, conference proceedings, posters, study protocols, letters to the editor, reviews, meta-analyses, or gray literature.

Data Extraction

All titles, abstracts, and full texts were independently screened by 2 reviewers (JLW and CH) using Rayyan software developed by Ouzzani et al [70]. Any disagreements were settled by a third reviewer (RMA).

Author JLW created data extraction forms using Excel (Microsoft Corp) and refined them in collaboration with 2 other authors (CH and RMA). One author (JLW) independently extracted data from all eligible articles, and the extracted data were verified by author CH.

The key measures of interest were: (1) study design, (2) participant characteristics and study setting, (3) stage of cognitive impairment (ie, MCI and mild, moderate, or severe dementia) and diagnostic criteria applied, (4) measure of outcome (SB), and (5) comparisons between groups (healthy controls or other forms of cognitive impairment). For outcome measures (SB), we extracted the (1) SB metrics and their values (eg, volume, pattern, and variability metrics); (2) method of assessment (eg, accelerometry); (3) definition of measurement (eg, body posture or cut-point thresholds); (4) device name and manufacturer; (5) data collection procedure (eg, device placement and length of assessment); (6) previously established validity and reliability of SB classification (eg, validation of cut point); and (7) data processing algorithms used to derive SB outcomes, including the classification approach (eg, cut-point–based or posture-based), algorithm source (eg, academic or proprietary), and evidence of algorithm validation.

Assessment of Study Quality

The quality of study articles was assessed using an adapted version of the National Institute of Health Quality Assessment Tool for Observational Cohort and Cross-sectional Studies [71]. This adapted version has previously been used in similar reviews [72]. Study quality was assessed based on 7 questions, which can be found in Multimedia Appendix 3.

XSL•FO
RenderX

Two raters (JLW and CH) independently assessed the quality of the studies and reached consensus through discussion; average scores determined the overall quality of each study. A binary system was used to rate the studies (1=yes and 0=no), which enabled a summed score out of 7 to aid in determining the overall quality of each study (rated good: scores 5-7, moderate: scores 3-4, or poor: scores 0-2).

Data Synthesis

Narrative data synthesis was applied to summarize and integrate findings across the studies. Reflecting this systematic review's aims, results were grouped by (1) synthesis of reported SB outcomes in people with cognitive impairment, considering disease severity and subtype; and (2) wearable technology methods and metrics used to measure SB, including device type and placement, characterization of SB, and processing techniques and validity. Due to heterogeneity in the data, meta-analysis was infeasible.

To support the interpretation of data and provide consistency across the literature, we categorized wearable technology metrics of SB into volumes, patterns, and variability of behavior. These categories were theoretically derived, drawing on established concepts in SB and PA literature [24,72-74], and were operationalized before the final data synthesis to ensure consistent classification of metrics across all included studies. Volume refers to the amount of time spent in SB during a specified period (eg, per day or full assessment period), expressed either as an absolute value (minutes or hours) or as a relative value (percentage of time) [22]. Where necessary, time-based data were converted to a common unit (eg, minutes) to ensure consistency in aggregation and comparison. Pattern explains the way sedentary time is accumulated, including the number and duration of SB bouts, the interruptions in SB, and the distribution of SB throughout the day [24]. Variability refers to the intraindividual fluctuations or day-to-day changes in sedentary time and captures how consistently (or inconsistently) a person engages in SB across different days or within various segments of the day (eg, morning vs evening) [75]. The categorization of metrics was performed by the primary researcher (JLW) and reviewed by all authors. Any amendments were discussed collaboratively, and all revisions were agreed upon by consensus.

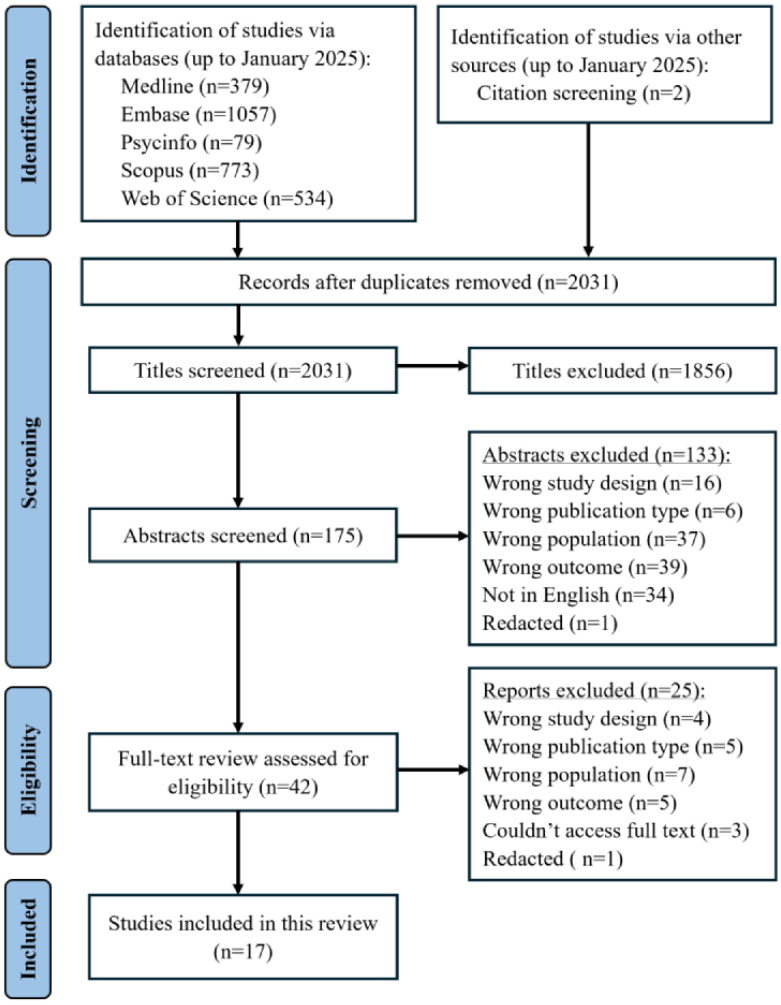
Results

Search Yield

Figure 1 describes the results of the search strategy for articles investigating SB using wearable digital technology in people with cognitive impairment. The search was conducted between November 2024 and January 2025 and initially yielded 2822 articles; 2 additional studies were identified through citation screening. Our final systematic review included 17 articles, all of which were published between 2016 and 2025.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram demonstrating the search yield for this systematic review.



Study Characteristics

Table 1 contains information relating to all the study characteristics. Studies took place in Germany (3/17, 18%), the Netherlands (3/17, 18%), the United States of America (3/17, 18%), Australia (2/17, 12%), Canada (2/17, 12%), Hong Kong (1/17, 6%), Japan (1/17, 6%), Norway (1/17, 6%), and Portugal (1/17, 6%). The sample size of participants with cognitive impairment ranged between 8 and 539 across all studies, with a mean age range of 68-91.7 years. Regarding participants with cognitive impairment, 59% (10/17) of studies reported ≥50% of participants as female. A total of 3 (18%) studies reported the ethnicity of participants with cognitive impairment, whereby more than 87% of the participants were White. Most studies were cross-sectional (14/17, 82%), the remaining studies (3/17, 18%) used baseline data from randomized controlled trials or intervention studies; there were no longitudinal observational studies. Of the 17 studies, 11 included only community-dwelling participants, 3 included only institutionalized participants, and 3 included a mix of both (Figure 2 [76-92]).

XSL•FO
RenderX

Table 1. Summary of included studies with main findings for SB[a] metrics reported in ≥2 studies.

Study; quality	Participant demographics	Data collection and processing	Main study findings	Comparisons
Abel et al [76]; good	• Dementia: n=53 (age: 82.3 years; 74% female; 72% community-dwelling)	• Device: PAMSys (chest) • Duration: 3 days • Classification: body posture • Algorithm: academic	• Lying: 9.1 hours/day • Sitting: 11.3 hours/day • Inactive: 20.4 hours/day	• No comparisons
Amagasa et al [77]; moderate	• Cognitive impairment: n=48 (MCI[b]/AD[c]; age: 77.6 years; 52% female; 100% community-dwelling) • Healthy controls: n=463 (age: 73, SD 5.4 years; 53% female; 100% community-dwelling)	• Device: HJA-750C (waist) • Duration: 7 days • Classification: ≤1.5 METs[d] • Algorithm: academic	• SB time: cognitive impairment, 476.2 minutes/day; healthy controls, 442.4 minutes/day	• SB time was not significant vs healthy controls
Balbim et al [78]; good	• MCI: n=253 (age: 73.69 years; 62% female; 100% community-dwelling)	• Device: Motionwatch8 (wrist) • Duration: 7 days • Classification: ≤178.5 CPM[e] • Algorithm: proprietary	• SB time: 626.9 minutes/day • SB percentage: 43.55% of the day	• No comparisons
Cerff et al [79]; good	• PDD[f]-MCI: n=22 (age: 68 years; 23% female; 100% community-dwelling) • PDD: n=9 (72 years; 0% female; 100% community-dwelling) • Healthy controls: n=17 (age: 71 years; 41% female; 100% community-dwelling)	• Device: MiniMod (lower back) • Duration: 3 days • Classification: body posture • Algorithm: proprietary	• SB percentage: PD-MCI, 78% of the day; PDD, 89% of the day; healthy controls, 75% of the day • SB bouts: PD-MCI, 129/day; PDD, 97/day; healthy controls: 134/day • SB bout length: PD-MCI, 515 seconds; PDD, 727 seconds; healthy controls, 506 seconds	• PDD: ↓ bouts and ↑ bout length vs PD-MCI and healthy controls. • SB percentage was not significant
Clina et al [80]; good	• AD: n=65 (age: 73.6 years; 37% female; 100% community-dwelling) • Healthy controls: n=65 (age: 69.4 years; 69% female; 100% community-dwelling)	• Device: wGT3x-BT (hip) • Duration: 7 days • Classification: ≤100 CPM • Algorithm: custom	• SB time: AD, 622.2 minutes/day; healthy controls: 591.8 minutes/day	• SB time was not significant vs healthy controls
Falck et al [81]; good	• MCI: n=81 (age: 72.5 years; 59% female; 100% community-dwelling) • Healthy controls: n=69 (age: 69.4 years; 78% female; 100% community-dwelling)	• Device: Motionwatch8 (wrist) • Duration: ≥4 days • Classification: ≤178.5 CPM • Algorithm: proprietary	• SB percentage: MCI, 61.65% of the day; healthy controls, 57.24% of the day • SB bouts (≥30 minutes): MCI, 4.07/day; healthy controls, 3.30/day	• MCI: ↑bouts ≥30 minutes vs Healthy controls. • SB percentage was not significant
Finnanger et al [82]; moderate	• Dementia: n=29 (FDC[g]; age: 74 years; 31% female; 100% community-dwelling) • Dementia: n=107 (RDC[h]; age: 84.3 years; 66% female; 100% community-dwelling)	• Device: Actisleep+ (wrist) • Duration: 7 days • Classification: ≤99 CPM • Algorithm: proprietary	• SB percentage: FDC, 39.7% of the week; RDC, 43.51% of the week	• Findings between care settings were not significant

XSL•FO
RenderX

Study; quality	Participant demographics	Data collection and processing	Main study findings	Comparisons
Hartman et al [83]; good	• Dementia: n=45 (age: 79.6 years; 49% female; 93% community-dwelling) • Healthy controls: n=49 (age: 80 years; 51% female; 98% community-dwelling)	• Device: Actiwatch 2 (wrist) • Duration: 7 days • Classification: ≤145 CPM • Algorithm: proprietary	• SB time: dementia, 8.5 hours/day; healthy controls, 8.2 hours/day • SB percentage: dementia, 57% of the day; healthy controls, 55% of the day • SB bouts (≥30 minutes): dementia, 2.3/day; healthy controls, 2.0/day • SB bout length: dementia, 18.3 minutes; healthy controls, 16.6 minutes	• Dementia: ↑ SB percentage and ↑ bout length vs healthy controls. • SB time and bouts ≥30 minutes was not significant
Hopkins et al [84]; good	• MCI: n=82 (age: 72 years; 56% female; 100% community-dwelling)	• Device: ActivPAL (thigh) • Duration: 7 days • Classification: body posture • Algorithm: proprietary	• SB time: 637 minutes/day • Sitting: 558 minutes/day • Lying: 21 minutes/day • SB bouts ≥30 minutes: 5/day	• No comparisons
Lu et al [85]; moderate	• MCI: n=105 (age: 83.6 years; 49% female; 100% community-dwelling) • Low MoCA[i]: n=252 (age: 83.4 years; 48% female; 98% community-dwelling) • AD: n=182 (age: 80.8 years; 66% female; 98% community-dwelling) • Healthy controls: n=271 (age: 81.9 years; 38% female; 100% community-dwelling)	• Device: wGT3x-BT (wrist) • Duration: 7 days • Classification: <1853 VM[j] CPM • Algorithm: proprietary	• SB percentage: MCI, 57.1% of the day; Low MoCA, 57.8% of the day; AD, 63.2% of the day; healthy controls, 58.4% of the day • SB bout length: MCI, 6.3 minutes; Low MoCA, 6.5 minutes; AD, 7.9 minutes; healthy controls, 6.6 minutes • SB bouts: MCI, 89.4/day; Low MoCA, 91.4/day; AD, 86.1/day; healthy controls, 91.5/day • SB bouts ≥30 minutes: MCI, 3.5/day; Low MoCA, 3.3/day; AD, 4.1/day; healthy controls 3.3/day	• AD: ↑ SB percentage, ↑ bout length, ↑ ≥30-minute bouts vs MCI, Low MoCA, healthy controls, and ↓ bouts vs Low MoCA and healthy controls
Marmeleira et al [86]; good	• Cognitive impairment: n=48 (age: 83.9 years; 73% female; 0% community-dwelling) • Healthy controls: n=22 (age: 82.2 years; 55% female; 0% community-dwelling)	• Device: GT1M (hip) • Duration: 7 days • Classification: ≤100 CPM • Algorithm: proprietary	• SB time: cognitive impairment, 603.7 minutes/day; healthy controls, 601.0 minutes/day • SB percentage: cognitive impairment, 87.2% of the day; healthy controls, 84% of the day	• All findings were not significant vs healthy controls
Muurling et al [87]; good	• Cognitive impairment: n=12 (MCI/AD; age: 91.7 years; 67% female; 83% community-dwelling) • Healthy controls: n=24 (age: 92.5 years; 45% female; 94% community-dwelling)	• Device: MoveMonitor (lower back) • Duration: 7 days • Classification: body posture • Algorithm: proprietary	• SB time: cognitive impairment, 142.7 hours/week; healthy controls, 140.7 hours/week • Sitting bouts: cognitive impairment, 681.5/week; healthy controls, 741.1/week	• All findings were not significant vs healthy controls
Parry et al [88]; moderate	• Dementia: n=8; healthy controls: n=20 (age: 83.1 years; 29% female; 0% community-dwelling)	• Device: GT3X (dementia, hip; healthy controls, thigh) • Duration: 5 days • Classification: ≤100 CPM • Algorithm: custom	• SB time: dementia, 565.6 minutes/day; healthy controls, 583.9 minutes/day • SB percentage: dementia, 86% of the day; healthy controls, 85% or the day	• No comparisons
Rackdoll et al [89]; moderate		• Device: GT3X+ (hip) • Duration: 7 days • Classification: ≤99 CPM • Algorithm: proprietary	• SB percentage: MCI, 72% of the day; healthy controls 74% of the day	• SB percentage was not significant vs healthy controls

XSL•FO
RenderX

Study; quality	Participant demographics	Data collection and processing	Main study findings	Comparisons
	• MCI: n=18 (age: 70 years; 56% female; 100% community-dwelling) • Healthy controls: n=48 (age: 65 years; 52% female; 100% community-dwelling)			
Resnick et al [90]; good	• Cognitive impairment: n=279 CI; healthy controls: n=101 (age: 89.5 years; 72% female; 0% community-dwelling)	• Device: Motionwatch8 (wrist) • Duration: 5 days • Classification: ≤178.5 CPM • Algorithm: proprietary	• SB time: cognitive impairment, 1203 minutes/day; healthy controls, 1138 minutes/day	• Cognitive impairment ↑ SB time vs healthy controls
van Alphen et al [91]; good	• Dementia (community-dwelling): n=37 (age: 77.3 years; 41% female; 100% community-dwelling) • Dementia (institutionalized): n=83 (age: 83 years; 80% female; 0% community-dwelling) • Healthy controls: n=26 (79.5 years; 50% female; 100% community-dwelling)	• Device: Actiwatch AW4 (wrist) • Duration: >6 days • Classification: ≤100 CPM • Algorithm: not specified	• SB time: dementia (institutionalized), 17.3 hours/day; dementia (community-dwelling), 15.83 hours/day; healthy controls, 14.52 hours/day	• Dementia ↑ SB time vs healthy controls • Dementia (institutionalized) ↑ SB time vs dementia (community-dwelling) and healthy controls
Varma et al [92]; moderate	• AD: n=39 (age: 73.5 years; 28% female; 100% community-dwelling) • Healthy controls: n=53 (age: 73.2 years; 70% female; 100% community-dwelling)	• Device: GT3X+ (hip) • Duration: 7 days • Classification: ≤149 CPM • Algorithm: proprietary	• SB percentage: AD, 60.94% of the day; healthy controls, 54.07% of the day	• SB percentage was not significant vs healthy controls

[a]SB: sedentary behavior.

[b]MCI: mild cognitive impairment.

[c]AD: Alzheimer disease.

[d]METs: metabolic equivalent tasks.

[e]CPM: counts per minute.

[f]PDD: Parkinson disease dementia.

[g]FDC: farm-based dementia care.

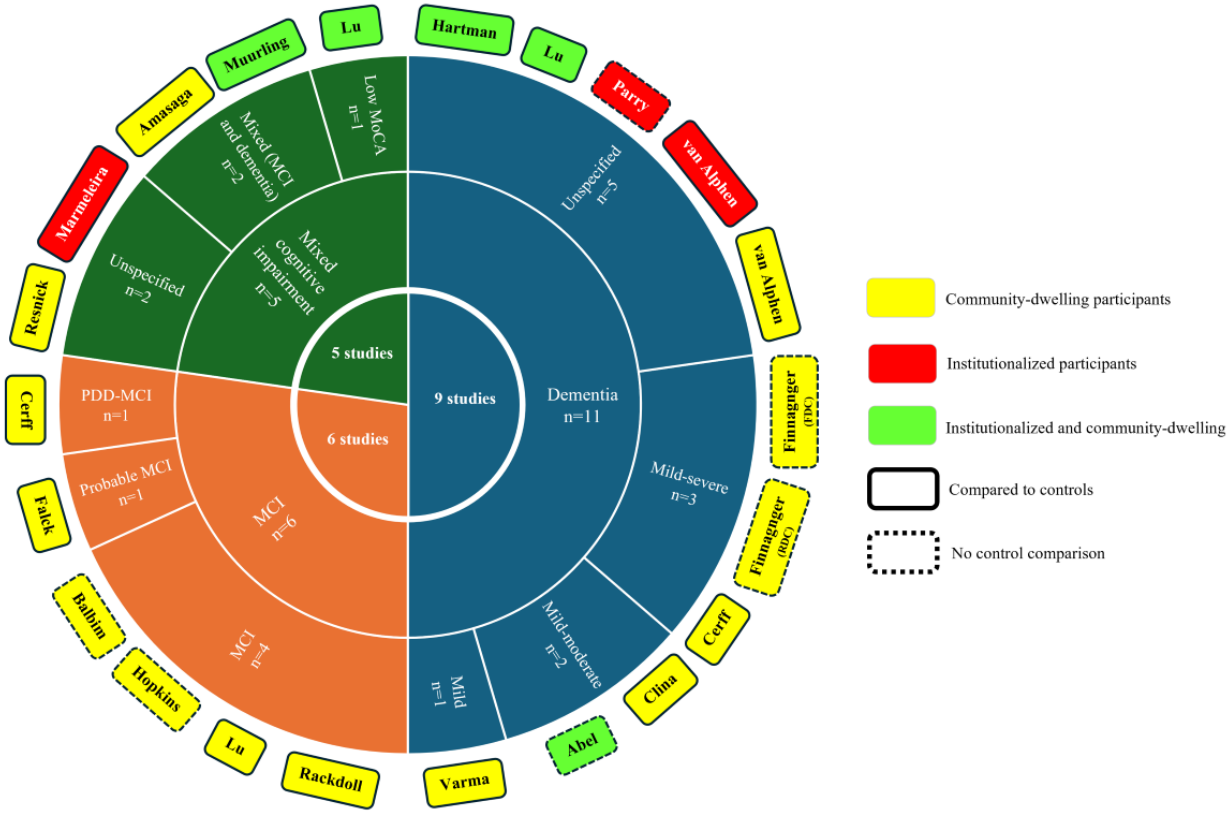
[h]RDC: regular day care.

[i]MoCA: Montreal Cognitive Assessment.

[j]VM: vector magnitude.

XSL•FO
RenderX

Figure 2. Overview of cognitive impairment categorization, study population, and control group comparisons for all studies included in this review. FDC: farm-based dementia care; MCI: mild cognitive impairment; MoCA: Montreal Cognitive Assessment; PDD: Parkinson disease dementia; RDC: regular day care.



Levels of cognitive impairment described across the 17 studies included MCI (n=5, 25%), unspecified severity of dementia (n=4, 20%), unspecified level of cognitive impairment (n=3, 15%), mild-moderate dementia (n=2, 10%), mild-severe dementia (n=2, 10%), mix of MCI and dementia (n=2, 10%), probable MCI (n=1, 5%), and mild dementia (n=1, 5%) (Figure 2).

For the purposes of this review, cognitive impairment was categorized into 3 broader cohorts: MCI (n=6), dementia (n=11), and mixed cognitive impairment (n=5), spanning the 17 studies. Mixed cognitive impairment refers to cohorts that have been identified as cognitively impaired with no specific diagnosis or mixed groups of dementia and MCI. Overall, 12 studies (71%) included comparisons to controls, resulting in 16 control cohorts corresponding to the dementia (n=7), MCI (n=4), and mixed cognitive impairment (n=5) cohorts (Figure 2).

A total of 8 (47%) studies specified dementia disease subtypes, of which 88% (n=7) included participants with Alzheimer disease (AD) [77,80,85,87,91-93]; 25% (n=2) reported participants with MCI/dementia due to Parkinson disease [79,91]; 25% (n=2) involved participants with vascular dementia or mixed dementia [91,93]; and 13% (n=1) included participants with frontotemporal dementia, dementia with Lewy bodies, or Korsakoff syndrome [91]. Additionally, 15 (88%) studies explicitly described procedures to characterize cognitive impairment (eg, clinician review and cognitive score thresholds), and 80% (n=12) of these used validated diagnostic criteria.

Quality Assessment

Mostly, studies were of adequate quality, with 65% (11/17) being rated as “good” and 35% (6/17) being rated as “moderate.” No studies were rated “poor.” The most common reasons for reduced study quality were a lack of sample size justification (n=15) (Multimedia Appendix 3).

Aim 1: Volume Patterns and Variability of SB in People With Cognitive Impairment

Overview

A summary of extracted outcomes (metrics reported by ≥2 studies) is provided in Table 1, with all metric definitions and prevalence detailed in Table 2. A comprehensive narrative synthesis of all extracted outcomes is provided in Multimedia Appendix 4.

XSL•FO
RenderX

Table 2. Prevalence and descriptions of SB[a] metrics captured in this review.

SB metric	Frequency, n (%)	Description
Volume		
Volume of SB	17 (100)	Time spent in SB for a specified time frame (eg, per day/per week) expressed as an absolute value (hours/minutes/seconds) or relative value (percentage of time)
SB time	10 (59)	Amount of time (hours/minutes/seconds) spent in SB over a specified period (eg, per day/week)
SB proportion	10 (59)	Percentage (%) of time spent in SB over a specified period (eg, per day/week)
Lying time	2 (12)	The length of time spent in sedentary lying (hours/minutes/seconds)
Sitting time	2 (12)	The length of time spent in sedentary sitting (hours/minutes/seconds)
Inactive time	1 (6)	The length of time spent inactive (sum of lying and sitting time) (hours/minutes/seconds)
Lying proportion	1 (6)	Relative percentage of time spent in sedentary lying
Sitting proportion	1 (6)	Relative percentage of time spent in sedentary sitting (%)
SB proportion (METs[b])	1 (6)	Relative percentage of time spent with intensity levels of ≤1.5 METs (%)
Pattern		
Pattern of SB	7 (41)	The number of sessions and distributions of sedentary activity over a specified period (eg, per day/week)
SB bouts (≥30 minutes)	4 (53)	The number of sedentary bouts lasting over 30 minutes, continuously
SB bouts	3 (18)	The number of sedentary bouts accumulated over a specified time (eg, per day/week)
Mean SB bout length	3 (18)	The average length of time spent in a singular sedentary bout
SB interruptions	1 (6)	The count of nonsedentary bouts that occur between two sedentary bouts
SB bouts (≥60 minutes)	1 (6)	The number of sedentary bouts lasting over 60 minutes, continuously
Hourly SB distribution	1 (6)	The extent to which an individual's SB differs across hours of the day
Variability		
Variability of SB	1 (6)	Fluctuations in sedentary time or patterns across a specified time (eg, day-to-day) within individuals or groups over time
Day-to-day variability	1 (6)	The extent to which an individual's SB differs across multiple days

[a]SB: sedentary behavior.

[b]METs: metabolic equivalent tasks.

Volume

All 17 studies used volume metrics to describe SB in individuals with cognitive impairment. Nine studies (53%; 10 cohorts) reported SB time per day: 4 in dementia (5 cohorts; ranging from 510 to 1038 minutes per day) [80,88,91,93], 2 in MCI (ranging from 626.9 to 637 minutes per day) [78,84], and 3 in mixed cognitive impairment (ranging from 476.2 to 1203 minutes per day) [77,86,90]. A total of 3 out of 4 dementia studies considered control group comparisons [80,91,93]; only 1 study [91] reported significantly higher SB time in institutionalized participants with dementia per day relative to both community-dwelling participants with dementia (P<.05) and healthy controls (P<.001), while the remaining 2 dementia studies reported no significant differences in SB time per day [80,93]. In MCI, neither of the 2 studies performed comparative analyses with control groups. For mixed cognitive impairment, all 3 studies considered control comparisons; only 1 study reported significantly higher SB in those with unspecified cognitive impairment (P<.01) [90], with the others reporting no significant differences [77,86]. Additionally, 1 study [87] reported SB time per week in mixed cognitive impairment (142.7 hours per week) and found no significant differences when compared with controls.

Nine studies (53%; 12 cohorts) reported SB as a percentage of the day: 5 in dementia (ranging from 57% to 89% per day) [79,85,88,92,93], 5 in MCI (ranging from 43.55% to 78% per day) [78,79,81,85,89], and 2 in mixed cognitive impairment (ranging from 57.8% to 87.2% per day) [85,86]. Among dementia studies, 4 included control comparisons [79,85,92,93], 2 reported statistically higher SB proportion (P<.05) [85,93], while the others found no significant differences [79,92]. Contrastingly, 4 of 5 MCI studies [79,81,85,89] and both mixed cognitive impairment studies considered group comparisons with controls, reporting no statistically significant differences. Additionally, 1 study [82] reported SB as a percentage of the week in 2 dementia cohorts (ranging from 39.7% to 43.51% per week), without comparison with healthy controls.

Two studies [76,84] assessed SB as sitting and lying minutes per day, respectively, 1 in dementia [76] (sitting: 678 minutes; lying: 546 minutes) and 1 in MCI [84] (sitting: 558 minutes;

XSL•FO
RenderX

lying: 21 minutes). Neither incorporated control comparisons, and no studies in mixed cognitive impairment reported this metric.

Finally, 1 study [79] assessed SB in people with dementia and MCI compared with cognitively intact controls using 3 volume metrics. First, it measured the proportion of time spent at ≤1.5 METs, finding that participants with dementia were sedentary for 89% of the day, compared with 84% for those with MCI. However, these differences were not statistically significant, either between MCI and dementia (P=.92) or between dementia and cognitively intact controls (P=.64). Second, the study reported the proportion of time spent lying and sitting. The MCI group spent the largest proportion of their day lying (42%), followed by sitting (34%), whereas the group with dementia spent more time sitting (42%) than lying (40%). However, no statistical comparisons were reported for this postural data, and these metrics were not assessed in mixed cognitive impairment groups.

Pattern

Seven studies (41%) used pattern metrics to describe SB in individuals with cognitive impairment. Four studies (57%; 6 cohorts) reported the number of prolonged (≥30 minutes) SB bouts per day: 2 in dementia (ranging from 2.3 to 4.1 bouts per day) [85,93], 3 in MCI (ranging from 3.5 to 5 bouts per day) [81,84,85], and 1 in mixed cognitive impairment (3.3 bouts per day) [85]. In dementia, both studies considered comparisons with controls [85,93]; 1 reported a significantly higher number of prolonged bouts compared with healthy controls, MCI, and mixed cognitive impairment groups (P<.05) [85], while the other found no significant differences [93]. In MCI, 2 of the 3 studies conducted comparative analyses with controls [81,85], with one reporting significantly more prolonged bouts than healthy controls (P<.05) [81] and the other finding no significant difference [85]. For mixed cognitive impairment, the single study found no statistically significant differences compared with controls [85].

Three studies (43%; 6 cohorts) reported average SB bout length: 3 in dementia (ranging from 474 to 1098 seconds) [79,85,93], 2 in MCI (ranging from 378 to 515 seconds) [79,85], and 1 in mixed cognitive impairment (390 seconds) [85]. All studies considered comparisons with control groups. All 3 dementia studies reported significantly longer bout durations compared with healthy controls (P<.05); 1 study further reported that this difference remained significant when compared to both MCI and mixed cognitive impairment groups [85]. In contrast, studies of MCI and mixed cognitive impairment found no significant differences in average bout length vs controls.

Two studies (43%; 5 cohorts) reported the number of SB bouts per day: 2 in dementia (ranging from 86.1 to 97.0 bouts per day) [79,85], 2 in MCI (ranging from 89.4 to 129.0 bouts per day) [79,85], and 1 in mixed cognitive impairment (91.4 bouts per day) [85]. All studies conducted comparative analyses with controls. For participants with dementia, both studies reported significantly fewer SB bouts compared with healthy controls (P<.05); 1 study further noted this finding remained significant when compared with a group with mixed cognitive impairment, but not MCI [85]. In contrast, studies of MCI and mixed cognitive impairment found no significant differences in this metric compared with controls. Additionally, 1 study reported the number of sitting bouts per week in mixed cognitive impairment (MCI and AD) [87], reporting 681.5 sitting bouts per week, with no significant differences compared with controls.

One study [93] compared daily sedentary interruptions between participants with dementia and healthy controls. Although the group with dementia averaged 27.2 interruptions per day, this difference was not statistically significant (P=.20). This metric was not assessed in participants with MCI or mixed cognitive impairment.

Another study [84] quantified the daily number of sedentary bouts exceeding 60 minutes in individuals with MCI, finding an average of 1 bout per day. However, this study did not consider control group comparisons, and this metric was not assessed in groups with dementia or mixed cognitive impairment.

Finally, 1 study [86] reported the hourly distribution of SB and found participants with unspecified cognitive impairment spent significantly more time sedentary, compared with healthy controls, during the hours of 7 AM to 11 AM, 1 PM, and 4 PM. This metric was not assessed in participants with dementia or MCI.

Variability

One study [76] reported day-to-day variability of participants with dementia, where they found SB parameters demonstrated consistently low day-to-day variability between Friday and Saturday and Saturday and Sunday (intraclass correlation coefficients ranging from 0.82 to 0.85), indicating minimal influence from day-specific or contextual factors. This metric was not assessed in participants with MCI or mixed cognitive impairment.

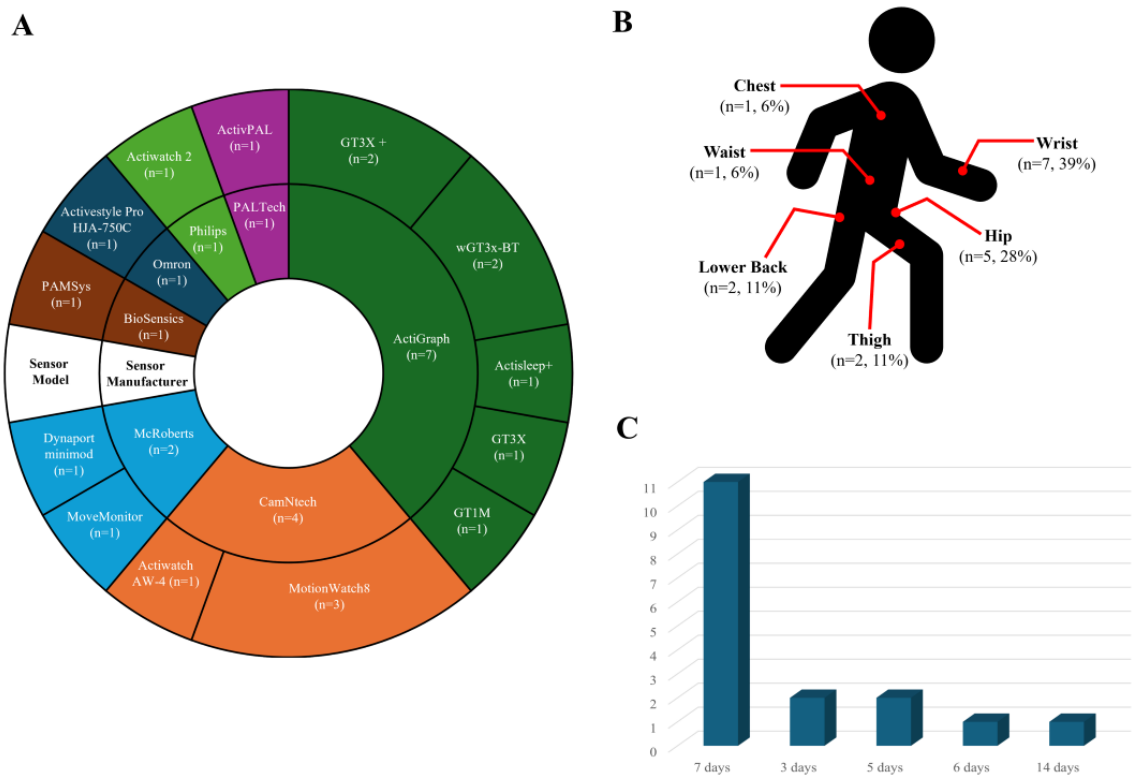
Aim 2: Wearable Digital Methods and Metrics Used to Measure SB in People With Cognitive Impairment

Overview

All studies (100%) used accelerometer-based wearable devices to measure SB in people with cognitive impairment. Across the 17 studies, 13 different device models were used across 7 different manufacturing brands (Figure 3A). The most common manufacturing brand reported was ActiGraph, with 41% (7/17) of studies using 5 different ActiGraph models to measure SB. However, the MotionWatch8 (CamNtech Ltd) was the most used wearable device across the studies (3/17, 18%). Six different body placements were used for devices, the most common placement being the wrist (7/17, 39%) (Figure 3B). Five different assessment periods were reported by studies; most study protocols (11/17, 65%) requested participants to wear the devices for 7 days (Figure 3C).

XSL•FO
RenderX

Figure 3. Methods and protocols used across studies included in this review. (A) Digital devices used to capture sedentary behavior (SB) metrics categorized by sensor type and manufacturer, (B) prevalence of placements of wearable devices on the body, and (C) prevalence of SB collection time periods.



Data Loss and Wear Compliance

Thirteen (76%) papers reported reasons for data loss, including insufficient data collection (n=11, 92%; <3 days [82,85,86]; <4 days [87,88]; <5 days [78]; <6 days [93]; <6 consecutive days [91]; <80% daily wear time [79]; <10 hours daily [84]; <10 hours on at least 3 days including 1 weekend day [80]), technical issues (n=6, 46% [87-91,93]), removal or refusal to wear devices (n=3, 23%; wrist [90,91]; lower-back [87]), lost devices (n=2, 17% [87,91]), forgetting to wear the device (n=1, 8% [87]), and organizational issues (n=1, 8% [91]).

Nonwear Identification and Wear Time Criteria

Eleven studies (65%) explicitly reported identifying nonwear in their analyses, either by specifying the criteria used (eg, zero accelerometer counts for a specified period, n=8 [77,80,81,85,86,88,90,92]) or by referencing nonwear algorithms (n=3 [79,82,87]). The remaining studies either did not report (n=4 [76,84,89,91]) or explicitly did not perform (n=2 [78,93]) nonwear identification. Fourteen studies (82%) established thresholds to determine data validity. The minimum required days for inclusion ranged from 3 to 6 days (n=11), and the minimum required wear per day ranged from 8 to 24 hours (n=11) (Multimedia Appendix 4).

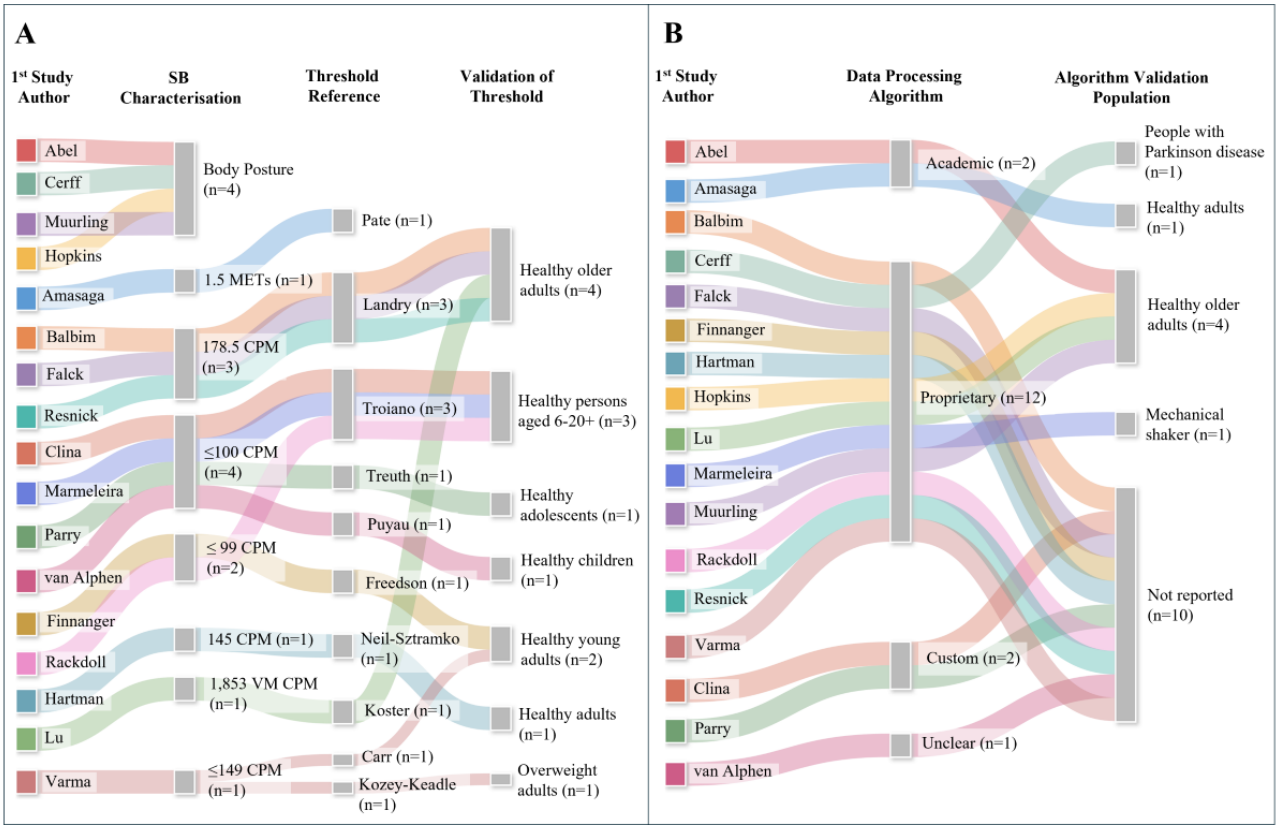
SB Quantification and Algorithm Validation

All studies (100%) reported their quantification of SB. A total of 4 (24%) classified SB using body posture [76,79,84,87], and 1 study (6%) classified SB as ≤1.5 METs [77]. A total of 71% (12/17) used CPM thresholds to classify SB.

In these cases, the thresholds were supported by references to validation studies conducted in 7 distinct populations (Figure 4A [76-92]). Furthermore, 16 (94%) studies described the data processing algorithms used to extract SB outcomes, 12 (75%) used proprietary algorithms provided by the manufacturer, 2 (13%) adopted academic algorithms provided from previous research, and 2 (13%) used custom-developed algorithms. Further, 7 (44%) studies explicitly reported the validation of the algorithms to extract SB outcomes (Figure 4B).

XSL•FO
RenderX

Figure 4. Variability in sedentary behavior (SB) characterization, thresholds, and processing across all studies (n=17). (A) Approaches to SB characterization, adopted thresholds, and validation of thresholds. (B) Data processing algorithms used to process SB with corresponding validation populations. CPM: counts per minute; METs: metabolic equivalent tasks; VM: vector magnitude.



Influence of Methodological Approaches on SB Estimates

The choice of classification approach appeared to influence the types of metrics reported across the included studies. Studies using posture-based classification were more likely to report pattern-related metrics (3/4, 75%) compared with those using CPM thresholds (4/12, 33%). The MET-based study did not report any pattern-related metrics. Day-to-day variability was the least frequently captured dimension, reported in only 1 study [76], which used a posture-based approach. The specific parameters used by each study are detailed in Multimedia Appendix 5, and definitions for all identified metrics are provided in Table 2. As SB volume was the only outcome reported by all studies, it was used to examine how methodological choices shaped the resulting estimates. Volume metrics reported by at least 2 studies are synthesized below and in Table 1, with a complete overview provided in Multimedia Appendix 4.

Device Placement

All CPM-based studies (n=12) applied devices on the wrist (n=7 [78,81,82,85,90,91,93]) or hip (n=5 [80,86,88,89,92]). Wrist-worn SB time ranged from 510 to 1203 minutes per day (n=4 [78,83,90,91]), and the proportion ranged from 43.55% to 63.2% per day (n=4 [78,81,83,85]). Hip-worn SB time ranged from 565.6 to 622.2 minutes per day (n=3 [80,86,88]), and proportion ranged from 60.94% to 87.2% per day (n=4 [86,88,89,92]). Posture-based studies (n=4) applied devices on the lower back (n=2 [79,87]), thigh (n=1 [84]), or chest (n=1 [76]). Lower-back placement produced sedentary proportions of 78% to 89% per day [79]. Thigh placement reported 637 minutes of SB time per day [84], while chest reported 1224 minutes per day [76]. The single MET-based study (n=1 [77]) used a waist-mounted device and reported 476.2 minutes of SB time per day.

Classification Approach

Among the studies adopting the lowest cut-point thresholds of ≤99-100 CPM, SB time per day ranged from 565.6 to 1038 minutes (n=4 [80,86,88,91]), while SB proportion per day ranged from 72% to 87.2% (n=3 [86,88,89]). In studies adopting mid-range thresholds of ≤145-149 CPM, SB proportion ranged from 57% to 60.94% (n=2 [92,93]). For those using a higher threshold of ≤178.5 CPM, SB time per day ranged from 626.9 to 1203 minutes (n=2 [78,90]), with SB proportion per day ranging between 43.55% and 61.65% (n=2 [78,81]).

Posture-based studies generally provided more granular data than threshold-based methods. Two such studies reported daily SB time (637-1224 minutes per day) and further distinguished between sitting (558-678 minutes per day) and lying time (21-546 minutes per day) [76,84]. Additionally, a study using body-posture classification estimated the proportion of SB spent at ≤1.5 METs to be between 84% and 89% per day [79]. The single MET-based study estimated SB time ≤1.5 METs (476.2 minutes per day) [77].

Denominator Choice in Proportion Calculations

Of the 10 studies that quantified SB proportion, a total of 6 [82,85,86,88,92,93] calculated proportion relative to wear or

XSL•FO
RenderX

wake time, and 4 [78,79,81,89] calculated proportion relative to a full day (24 hours). Specifically, within CPM-based studies, using wear or wake time as a denominator yielded higher SB proportions (range: 57%-87.2%) compared to those calculated relative to a 24-hour day (range: 43.55%-72%).

Algorithm Type

Of the 12 studies (16 cohorts) using proprietary algorithms, 8 reported SB proportion per day [78,79,81,85,86,89,92,93] (11 cohorts; range: 43.55%-89%), and 5 reported SB time [78,83,84,86,90] (range: 510-1203 minutes per day). Both studies using academic algorithms [76,77] reported SB time (range: 476.2-1224 minutes per day). The 2 studies using custom algorithms [80,88] also reported SB time (range: 565.6-622.2 minutes per day). No clear patterns emerged across algorithm types, as estimates overlapped considerably.

Discussion

Overview

This systematic review is the first to characterize SB across cognitive impairments (ie, MCI and dementia) and critically appraise the wearable digital methods used to measure it. A key finding is that while total SB volume did not consistently differ between those who are cognitively impaired and controls, their pattern appeared to. Specifically, individuals with dementia exhibited significantly fewer but longer sedentary bouts. However, this review also highlights significant methodological inconsistencies in how SB is measured. Variations in device type and placement, wear-time protocols, cut-point thresholds, proportion denominators, and reported outcomes limit cross-study comparisons and, consequently, our understanding of SB and its influencing factors.

Impact of Cognitive Impairment and Disease Subtype on SB

To reflect Aim 1, this review suggests that the relationship between cognitive impairment and SB may be characterized less by a simple increase in volume and more by a potential shift in pattern. The available evidence indicates that people with dementia tend to exhibit longer and fewer sedentary bouts compared with those with MCI and cognitively intact controls [79,85,93], a profile that could be indicative of more prolonged, uninterrupted SB, which is often considered the most hazardous type [94]. However, no longitudinal research was identified in the review, meaning we cannot ascertain how SB patterns change with disease progression or if it predicts worse outcomes. Furthermore, our synthesis points to a nuanced conclusion: a significantly greater volume of SB does not appear to be a consistent hallmark of cognitive impairment. Among the 12 studies (comprising 18 cohorts) that considered control comparisons, only 4 identified a significant increase in SB volume [85,90,91,93]. Of these, 3 (75%) were specific to dementia cohorts, while 1 pertained to a group with mixed cognitive impairment. The existing self-report literature, which focuses on physical inactivity, provides a relevant context for interpreting the current findings on SB. Studies have found that physical inactivity is not directly associated with lower global cognition in community-dwelling adults with dementia [95-100]. Instead, inactivity appears to be more closely linked to other factors, including depression, a history of falls, fewer waking hours, and environmental barriers [95,99-102]. This suggests that cognitive impairment may not inevitably lead to pervasive inactivity, which may also extend to SB, pointing to the influence of other mediating factors.

The apparent discrepancy between inconsistent volume and consistent pattern findings may be explained by neuropsychiatric symptoms and their interaction with the environment, rather than the disease pathology itself. For instance, apathy, a prevalent symptom in dementia, varies significantly between individuals and may manifest not as a greater total sedentary time, but as a reduced frequency of interruption, leading to the observed pattern of prolonged bouts [32,103]. This perspective is supported by Alonzo's [104] theory, which posits that the impact of an illness emerges from its interaction with the individual's environment. In this view, symptoms like apathy and disorientation interact with environmental contexts (eg, institutional settings and lack of stimulation) to promote sedentariness. Qualitative evidence corroborates this mechanism, identifying symptoms such as exhaustion and disorientation as common barriers to activity [103]. Therefore, future qualitative research is warranted to further elucidate the lived experience of sedentariness in this population.

It is important to note, however, that this interpretation is constrained by the methodological limitations identified in this review. The widespread use of wrist-worn devices and unvalidated cut points [37,105] is likely to impair the detection of short, nonpostural movements that break up sedentary time. Therefore, the current evidence base may underestimate the true extent of sedentary patterns in cognitive impairment, and the potential role of symptoms warrants further investigation with more precise and standardized measurement tools.

Furthermore, the generalizability of these findings across dementia subtypes remains uncertain. The available evidence is predominantly derived from cohorts with Alzheimer disease (88% of studies that reported dementia subtypes). Other dementia subtypes were included but not analyzed separately [91,93], preventing insights into how their unique symptomatic profiles might influence distinct SB patterns [106,107]. The current literature's focus on Alzheimer disease underscores a significant evidence gap regarding SB in non- Alzheimer disease dementias, which future research should devote attention to.

Current Approaches for Wearable Digital Methods and Metrics for Characterizing SB in People With Cognitive Impairment

Aligned with Aim 2, key findings in this review indicate that, in keeping with existing literature [108], studies quantified SB through 2 distinct methods: identifying body posture or applying cut-point thresholds. These approaches also shape the type of outcomes that can be derived; posture-based methods more readily capture pattern-related metrics (eg, bouts and breaks), whereas cut-point threshold approaches primarily quantify overall sedentary volume. This methodological divergence is compounded by differences in device placement and detection mechanisms. For posture-based measurements, thigh-worn

XSL•FO
RenderX

devices struggle to distinguish sitting from lying [43], while lower-back devices struggle to distinguish sitting from standing [109]. In contrast, cut-point thresholds do not distinguish between postures and are highly sensitive to nonambulatory arm movements. Consequently, wrist-worn devices may misclassify sedentary fidgeting as nonsedentary time [110,111].

This pattern is reflected in the findings of the present review: trunk placements (eg, hip, lower back, and waist) generally reported higher and more stable estimates of SB time and proportion, whereas wrist-worn devices produced the lowest SB proportions and the greatest variability in SB time. However, these differences may also be partially explained by how SB is characterized, as wrist-worn studies demonstrated greater variability in cut-point thresholds compared with more standardized placements such as the hip. Collectively, these findings suggest that differences in SB estimates may be driven less by device placement alone and more by underlying methodological choices (eg, CPM thresholds), which reflect differing conceptualizations of SB.

The CPM approach was the most common method adopted to measure SB in cognitive impairment, with 71% (12/17) of studies using this method. The CPM approach is widely used in SB research [112] and functions as a proxy measure for energy expenditure (METs) [40], aligning with the standard SB definition of ≤1.5 METs. However, its application is highly inconsistent; 6 different CPM thresholds were identified in this review, none of which were validated in cognitively impaired populations. Although 2 thresholds (178.5 CPM [78,81,90] and 1853 vector magnitude CPM [85]) were validated in healthy older adults, the remainder lacked validation in any older population, significantly undermining the validity and comparability of findings.

The consequences of this inconsistency were most apparent in estimates of sedentary proportion. Unlike the broader literature, where lower cut points typically yield lower sedentary estimates [43,113], studies in this review using lower thresholds (≤99–100 CPM) reported the highest sedentary proportions, whereas higher thresholds (≤178.5 CPM) produced lower estimates. This inverse pattern suggests that lower cut points validated in younger populations may overestimate sedentary time when applied to older adults with cognitive impairment, while thresholds derived from older populations may provide more accurate estimates. These trends were not observed for absolute sedentary time (minutes per day) or by algorithm type (proprietary, academic, and custom), indicating that proportional differences may instead reflect denominator choice. Studies using waking or wear-time denominators reported higher sedentary proportions than those using a 24-hour denominator, as excluding sleep and nonwear reduces the denominator and inflates the proportion of time classified as sedentary. This aligns with methodological evidence demonstrating that accelerometer processing decisions, including wear-time definitions, substantially influence sedentary estimates and complicate comparisons across studies [114-117]. Together, inconsistencies in cut-point selection and denominator choice limit comparability across studies and hinder evidence synthesis.

While less common, posture-based methods (used by 24% of studies) offer a more direct assessment of SB. This approach circumvents the need for population-specific energy expenditure validation and may offer more accurate detection of SB [24,43]. Matthews et al [29] have called for such objective definitions to incorporate posture, arguing that energy-based criteria alone may not fully capture the nature of SB. However, this method introduces a different set of limitations; crucially, it cannot estimate METs and is therefore limited to classifying posture rather than intensity [118]. This is a significant constraint, as it diverges from the consensus energy-based definition of SB (≤1.5 METs). Therefore, the definition provided by Chastin and Granat [24] (“any nonupright activity”) may offer a practical alternative. Thus, the choice of method to measure SB presents a fundamental trade-off: prioritizing conceptual alignment with the energy-based definition of SB, often via CPM, at the potential cost of measurement accuracy (due to the current lack of validated CPM thresholds), or opting for the practical accuracy of a posture-based definition (“any nonupright activity”) that may neglect the MET component of SB.

It is important to recognize that SB, like all behaviors, is multifaceted and challenging to fully capture with a single measurement [38]. Therefore, future research should carefully consider the limitations of existing measurement tools and select those best aligned with specific research objectives and target populations. Crucially, measures must adhere to core principles of good measurement: demonstrating reliability, validity, population specificity, and sensitivity to change [119,120]. By emphasizing such criteria, researchers can enhance the accuracy and relevance of SB assessment, particularly in populations with unique physiological profiles such as cognitive impairment.

Recommendations

To aid future research and advance the field, we propose a set of recommendations to guide more meaningful and standardized approaches to measuring SB in people with cognitive impairment (Textbox 1).

XSL•FO
RenderX

Textbox 1. Recommendations for standardizing sedentary behavior measurement in cognitive impairment research.

Recommendations

- Validate cut-point thresholds and/or body posture algorithms in older adults, or populations with cognitive impairment, where possible, ensuring they account for functional limitations and typical behavioral patterns for their application in people with cognitive impairment.
- Standardize the reporting of sedentary proportion. To support cross-study comparability, researchers should either:
  - Report sedentary proportion using a fixed 24-hour denominator.
  - If reporting proportion relative to a non–24-hour specified period (eg, wear or wake time), also report the absolute sedentary time (eg, total minutes per day) to allow readers and future analyses to calculate proportions using their preferred denominator.
- Establish minimum reporting standards for sedentary behavior methodology. Studies should explicitly report device model and placement, cut-point threshold, and validation source, denominator used for proportional estimates, wear-time criteria, and sleep inclusion/exclusion status.
- Develop and adopt a core set of outcomes for sedentary behavior research in cognitive impairment, including key metrics, such as total sedentary time, number of prolonged bouts, and mean bout length, to enable cross-study comparison.

Strengths and Limitations

This review has several strengths. It is the first to systematically examine the wearable digital technology methods and metrics used to assess SB in individuals with cognitive impairment, offering a comprehensive overview of current practices and highlighting key methodological inconsistencies. The study also extends its scope to consider the influence of cognitive impairment severity and disease subtype on SB, which has received limited attention in previous research. A rigorous search strategy, transparent inclusion criteria, and quality appraisal of included studies further strengthen the reliability of the findings.

However, this review's findings are subject to important limitations. The categorization of SB into volume, pattern, and variability, while necessary for synthesis and aligned with existing frameworks [72,121], may not capture all relevant dimensions of SB. Furthermore, the deliberate exclusion of home-based technologies, though methodologically justified by their current lack of standardized validation and restricted ecological scope [48,122], delimits the evidentiary basis exclusively to wearable-derived data. Consequently, the conclusions presented here should be interpreted within the context of these methodological boundaries.

Conclusions

This review establishes that while preliminary evidence suggests that SB patterns differ across cognitive impairment, differences in SB volume are not consistently observed. Furthermore, the increasing use of wearable technology to assess SB in cognitive impairment is hampered by significant methodological heterogeneity and a critical lack of population-specific validation. This heterogeneity underscores the critical need for a consistent methodology to advance the field and establish robust, reproducible evidence on the role of SB in cognitive impairment.

Acknowledgments

This paper presents independent research supported by the National Institute for Health and Care Research (NIHR) Newcastle Biomedical Research Centre (BRC). The NIHR Newcastle BRC is a partnership between Newcastle Hospitals National Health Service (NHS) Foundation Trust, Newcastle University, and Cumbria, Northumberland and Tyne and Wear NHS Foundation Trust, and is funded by the NIHR. Authors JLW, SDD, AJY, and RMA are supported by the NIHR Newcastle BRC.

The views expressed in this publication are those of the authors and not necessarily those of the NIHR, NHS, or the UK Department of Health and Social Care. The funders had no role in the design of the study or the preparation of this manuscript.

Data Availability

All data generated or analyzed during this study are included in this published article and its Multimedia Appendices 1-5. All articles included in the review are publicly available online. The complete search strategy, including the search strings used for each database, is provided in Multimedia Appendices 1-5 to facilitate reproducibility.

Funding

SDD and CH were supported by the Innovative Medicines Initiative 2 Joint Undertaking (IMI2 JU) project IDEA-FAST (grant agreement 853981). SDD and CH were supported by the UK Research and Innovation (UKRI) Engineering and Physical Sciences Research Council (EPSRC) (grant reference EP/X031012/1 and grant reference EP/X036146/1). RMA received funding from the NIHR (NIHR 301677) and the NIHR Three Schools Dementia Career Development Award (NIHR-SSCR-DP-CDA30). AJY has received research support from the NIHR Newcastle BRC, Dunhill Medical Trust, EU IMI, NIHR, Parkinson's UK, Michael J Fox Foundation, Weston Brain Institute, Lewy Body Society, Intercept Pharmaceuticals, and Electrocore, and has received

XSL•FO
RenderX

funding and/or honoraria from Britannia, UCB, AbbVie, GSK, Teva-Lundbeck, GE Healthcare, and Genus for attending or speaking at educational events.

Authors' Contributions

Conceptualization: JLW, RMA
Data curation: JLW
Formal analysis: JLW, CH
Funding acquisition: RMA
Investigation: JLW, CH, RMA
Methodology: JLW, RSF, RMA
Supervision: CH, SDD, AJY, RMA
Validation: JLW, CH, RMA
Visualization: JLW
Writing – original draft: JLW
Writing – review & editing: JLW, CH, RSF, SDD, AJY, RMA

Conflicts of Interest

SDD is a shareholder of Enoda Health Ltd. This shareholding did not have any relationship with the work in this paper.

Multimedia Appendix 1
PRISMA checklist.
[DOCX File , 269 KB - aging_v9i1e85361_app1.docx]

Multimedia Appendix 2
Search strategy: keywords and MeSH terms for systematic literature review.
[DOCX File , 16 KB - aging_v9i1e85361_app2.docx]

Multimedia Appendix 3
Quality assessment of all studies included in this systematic review.
[DOCX File , 351 KB - aging_v9i1e85361_app3.docx]

Multimedia Appendix 4
Comprehensive overview of study characteristics with main findings and comparisons.
[DOCX File , 75 KB - aging_v9i1e85361_app4.docx]

Multimedia Appendix 5
Matrix of all studies and corresponding parameters used to describe sedentary behaviour.
[DOCX File , 384 KB - aging_v9i1e85361_app5.docx]

References

1. Wimo A, Ali GC, Guerchet M, Prince M, Prina M, Wu YT. World Alzheimer Report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends. Alzheimer's Disease International. 2015 Sep 21. URL: https://www.alzint.org/resource/world-alzheimer-report-2015/ [accessed 2026-06-14]
2. Global action plan on the public health response to dementia 2017-2025. World Health Organization. 2017 Dec 8. URL: https://www.who.int/publications/i/item/9789241513487 [accessed 2026-06-14]
3. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 1999 Mar;56(3):303-308. [doi: 10.1001/archneur.56.3.303] [Medline: 10190820]
4. Feldman HH, Jacova C. Mild cognitive impairment. Am J Geriatr Psychiatry 2005;13(8):645-655. [doi: 10.1097/00019442-200508000-00003]
5. Dunne R, Aarsland D, O'Brien JT, Ballard C, Banerjee S, Fox NC, et al. Mild cognitive impairment: the Manchester consensus. Age Ageing 2021 Jan 08;50(1):72-80 [FREE Full text] [doi: 10.1093/ageing/afaa228] [Medline: 33197937]
6. Halloway S, Wagner M, Tangney C, Lange-Maia BS, Bennett DA, Arvanitakis Z, et al. Profiles of lifestyle health behaviors and cognitive decline in older adults. Alzheimers Dement 2024 Jan;20(1):472-482 [FREE Full text] [doi: 10.1002/alz.13459] [Medline: 37676928]
7. Maasakkers CM, Claassen JAHR, Gardiner PA, Olde Rikkert MGM, Lipnicki DM, Scarmeas N, COSMIC Collaborators, et al. The association of sedentary behaviour and cognitive function in people without dementia: a coordinated analysis

XSL•FO
RenderX

across five cohort studies from COSMIC. Sports Med 2020 Feb;50(2):403-413 [FREE Full text] [doi: 10.1007/s40279-019-01186-7] [Medline: 31529300]
8. Song S, Stern Y, Gu Y. Modifiable lifestyle factors and cognitive reserve: a systematic review of current evidence. Ageing Res Rev 2022 Feb;74:101551 [FREE Full text] [doi: 10.1016/j.arr.2021.101551] [Medline: 34952208]
9. Soaring dementia care costs reach £42 billion in UK - and families bear the brunt. Alzheimer's Society. 2024 May 13. URL: https://www.alzheimers.org.uk/news/2024-05-10/soaring-dementia-care-costs-uk-42-billion#:~:text=This%20Dementia%20Action%20Week%2C%20which,unless%20urgent%20action%20is%20taken [accessed 2026-06-04]
10. Thyfault J, Du M, Kraus WE, Levine JA, Booth FW. Physiology of sedentary behavior and its relationship to health outcomes. Med Sci Sports Exerc 2015 Jun;47(6):1301-1305 [FREE Full text] [doi: 10.1249/MSS.0000000000000518] [Medline: 25222820]
11. Carter S, Hartman Y, Holder S, Thijssen DH, Hopkins ND. Sedentary behavior and cardiovascular disease risk: mediating mechanisms. Exerc Sport Sci Rev 2017;45(2):80-86. [doi: 10.1249/JES.0000000000000106] [Medline: 28118158]
12. Gogniat MA, Khan OA, Li J, Park C, Hudson Robb W, Zhang P. Increased sedentary behavior is associated with neurodegeneration and worse cognition in older adults over a 7-year period despite high levels of physical activity. Alzheimers Dement 2025;21(5):e70157. [doi: 10.1002/alz.70157] [Medline: 40357887]
13. Fleiner T, Dauth H, Gersie M, Zijlstra W, Haussermann P. Structured physical exercise improves neuropsychiatric symptoms in acute dementia care: a hospital-based RCT. Alzheimers Res Ther 2017;9(1):68 [FREE Full text] [doi: 10.1186/s13195-017-0289-z] [Medline: 28851451]
14. Prick A, de Lange J, Scherder E, Twisk J, Pot AM. The effects of a multicomponent dyadic intervention with physical exercise on the cognitive functioning of people with dementia: a randomized controlled trial. J Aging Phys Act 2017;25(4):539-552. [doi: 10.1123/japa.2016-0038] [Medline: 28120631]
15. Falck RS, Davis JC, Liu-Ambrose T. What is the association between sedentary behaviour and cognitive function? A systematic review. Br J Sports Med 2017;51(10):800-811. [doi: 10.1136/bjsports-2015-095551] [Medline: 27153869]
16. Thivel D, Tremblay A, Genin PM, Panahi S, Rivière D, Duclos M. Physical activity, inactivity, and sedentary behaviors: definitions and implications in occupational health. Front Public Health 2018;6:288 [FREE Full text] [doi: 10.3389/fpubh.2018.00288] [Medline: 30345266]
17. Urroz Guerrero PD, Oliveira JM, Lewthwaite H, Gibson PG, McDonald VM. Key considerations when addressing physical inactivity and sedentary behaviour in people with asthma. J Clin Med 2023;12(18):5998. [doi: 10.3390/jcm12185998] [Medline: 37762938]
18. David R, Mulin E, Friedman L, Le Duff F, Cygankiewicz E, Deschaux O. Decreased daytime motor activity associated with apathy in alzheimer disease: an actigraphic study. Am J Geriatr Psychiatry 2012;20(9):806-814. [doi: 10.1097/JGP.0b013e31823038af] [Medline: 21997602]
19. Hobson N, Dupuis SL, Giangregorio LM, Middleton LE. Perceived facilitators and barriers to exercise among older adults with mild cognitive impairment and early dementia. J Aging Phys Act 2020;28(2):208-218. [doi: 10.1123/japa.2019-0010] [Medline: 31621645]
20. Kobayashi Y, Takahashi Y, Seki T, Kaneta T, Amarume K, Kasai M, et al. Decreased physical activity associated with executive dysfunction correlates with cognitive impairment among older adults in the community: a retrospective analysis from the Kurihara project. Dement Geriatr Cogn Dis Extra 2016;6(2):350-360. [doi: 10.1159/000448027] [Medline: 27703468]
21. Han J, Sun W, Yin X, Wang L, Zhou F, Wang S. Facilitators and barriers to physical activity participation in older adults with mild cognitive impairment or dementia: a systematic review based on the COM-B model. Am J Alzheimers Dis Other Demen 2026;41:15333175261420561. [doi: 10.1177/15333175261420561] [Medline: 41570228]
22. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE. Sedentary behavior research network (SBRN) - terminology consensus project process and outcome. Int J Behav Nutr Phys Act 2017;14(1):75. [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
23. Vancampfort D, Stubbs B, Lara E, Vandenbulcke M, Swinnen N, Smith L. Mild cognitive impairment and sedentary behavior: a multinational study. Exp Gerontol 2018;108:174-180. [doi: 10.1016/j.exger.2018.04.017] [Medline: 29704638]
24. Chastin SFM, Granat MH. Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity. Gait Posture 2010;31(1):82-86. [doi: 10.1016/j.gaitpost.2009.09.002] [Medline: 19854651]
25. Dillon K, Morava A, Prapavessis H, Grigsby-Duffy L, Novic A, Gardiner PA. Total sedentary time and cognitive function in middle-aged and older adults: a systematic review and meta-analysis. Sports Med Open 2022 Oct 12;8(1):127 [FREE Full text] [doi: 10.1186/s40798-022-00507-x] [Medline: 36224459]
26. Da Ronch C, Canuto A, Volkert J, Massarenti S, Weber K, Dehoust M, et al. Association of television viewing with mental health and mild cognitive impairment in the elderly in three European countries, data from the MentDis_ICF65+ project. Ment Health Phys Act 2015 Mar;8:8-14. [doi: 10.1016/j.mhpa.2014.11.002]
27. Raichlen DA, Aslan DH, Sayre MK, Bharadwaj PK, Ally M, Maltagliati S, et al. Sedentary behavior and incident dementia among older adults. JAMA 2023 Sep 12;330(10):934-940 [FREE Full text] [doi: 10.1001/jama.2023.15231] [Medline: 37698563]

XSL•FO
RenderX

28. Watts A, Garnier-Villarreal M, Gardiner P. Measurement of sitting time in older adults with and without Alzheimer's disease. J Meas Phys Behav 2018;1(2):70-78. [doi: 10.1123/jmpb.2018-0002] [Medline: 40027224]
29. Matthews CE, Chen KY, Freedson PS, Buchowski MS, Beech BM, Pate RR, et al. Amount of time spent in sedentary behaviors in the United States, 2003-2004. Am J Epidemiol 2008;167(7):875-881. [doi: 10.1093/aje/kwm390] [Medline: 18303006]
30. Gebhard D, Frank JI. Everyday life and boredom of people living with dementia in residential long-term care: a merged methods study. BMC Geriatr 2024 Dec 31;24(1):1049 [FREE Full text] [doi: 10.1186/s12877-024-05641-7] [Medline: 39736528]
31. Haque R, Alam K, Gow J, Neville C, Keramat SA. Staying active, staying sharp: the relationship between physical activity and health-related quality of life for people living with cognitive impairment. Qual Life Res 2025;34(5):1443-1456. [doi: 10.1007/s11136-025-03910-5] [Medline: 39921826]
32. Kuhlmei A, Walther B, Becker T, Müller U, Nikolaus T. Actigraphic daytime activity is reduced in patients with cognitive impairment and apathy. Eur Psychiatry 2013;28(2):94-97. [doi: 10.1016/j.eurpsy.2011.04.006] [Medline: 21696925]
33. Harding E, Sullivan MP, Camic PM, Yong KXX, Stott J, Crutch SJ. "I Want to Do Something" - exploring what makes activities meaningful for community-dwelling people living with dementia: a focused ethnographic study. Qual Health Res 2024 Nov;34(13):1286-1302 [FREE Full text] [doi: 10.1177/10497323241239487] [Medline: 38648467]
34. Batič K, Kozinc Ž, Rus Prelog P. Differences in physical function across dementia subtypes and cognitive decline: a cross-sectional study. Eur J Transl Myol 2025;35(3):13726. [doi: 10.4081/ejtm.2025.13726] [Medline: 40457961]
35. Montero-Odasso M, Oteng-Amoako A, Speechley M, Gopaul K, Beauchet O, Annweiler C, et al. The motor signature of mild cognitive impairment: results from the gait and brain study. J Gerontol A Biol Sci Med Sci 2014;69(11):1415-1421. [doi: 10.1093/gerona/glu155] [Medline: 25182601]
36. Cai XY, Qian GP, Wang F, Zhang MY, Da YJ, Liang JH. Association between sedentary behavior and risk of cognitive decline or mild cognitive impairment among the elderly: a systematic review and meta-analysis. Front Neurosci 2023;17:1221990 [FREE Full text] [doi: 10.3389/fnins.2023.1221990] [Medline: 37600015]
37. Quante M, Kaplan ER, Rueschman M, Cailler M, Buxton OM, Redline S. Practical considerations in using accelerometers to assess physical activity, sedentary behavior, and sleep. Sleep Health 2015;1(4):275-284. [doi: 10.1016/j.sleh.2015.09.002] [Medline: 29073403]
38. Ainsworth B, Rivière F, Florez-Pregonero A. Measurement of sedentary behaviour in population studies. In: Sedentary Behaviour Epidemiology. Cham: Springer International Publishing; 2023:81-106.
39. Ainsworth BE, Caspersen CJ, Matthews CE, Mâsse LC, Baranowski T, Zhu W. Recommendations to improve the accuracy of estimates of physical activity derived from self report. J Phys Act Health 2012 Jan;9 Suppl 1(0 1):S76-S84 [FREE Full text] [doi: 10.1123/jpah.9.s1.s76] [Medline: 22287451]
40. Prince SA, Cardilli L, Reed JL, Saunders TJ, Kite C, Douillette K, et al. A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2020 Mar 04;17(1):31 [FREE Full text] [doi: 10.1186/s12966-020-00938-3] [Medline: 32131845]
41. Watts AS, Vidoni ED, Loskutova N, Johnson DK, Burns JM. Measuring physical activity in older adults with and without early stage Alzheimer's disease. Clin Gerontol 2013;36(4):356-374 [FREE Full text] [doi: 10.1080/07317115.2013.788116] [Medline: 24062599]
42. Marconcin P, Júdice PB, Ferrari G, Werneck A, Marques A. Methods of assessing sedentary behaviour. In: Sedentary Behaviour: A Contemporary View. Rijeka: IntechOpen; 2021.
43. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. Med Sci Sports Exerc 2011;43(8):1561-1567. [doi: 10.1249/MSS.0b013e31820ce174] [Medline: 21233777]
44. Weizman Y, Tan AM, Fuss FK. The use of wearable devices to measure sedentary behavior during COVID-19: systematic review and future recommendations. Sensors (Basel) 2023;23(23):9449. [doi: 10.3390/s23239449] [Medline: 38067820]
45. Chastin SFM, Dontje ML, Skelton DA, Čukić I, Shaw RJ, Gill JMR. Systematic comparative validation of self-report measures of sedentary time against an objective measure of postural sitting (activPAL). Int J Behav Nutr Phys Act 2018;15(1):21. [doi: 10.1186/s12966-018-0652-x] [Medline: 29482617]
46. Rodrigues IB, Tariq S, Kouroukis A, Swance R, Adachi J, Bray S. Mapping sedentary behaviour (MAPS-B) in winter and spring using wearable sensors, indoor positioning systems, and diaries in older adults who are pre-frail and frail: a feasibility longitudinal study. PLoS One 2024;19(5):e0290197 [FREE Full text] [doi: 10.1371/journal.pone.0290197] [Medline: 38753692]
47. Yang CC, Hsu YL. A review of accelerometry-based wearable motion detectors for physical activity monitoring. Sensors (Basel) 2010;10(8):7772-7788 [FREE Full text] [doi: 10.3390/s100807772] [Medline: 22163626]
48. Ianovski A. A Smart Home Platform and Hybrid Indoor Positioning Systems for Enabling Aging in Place. Hamilton: McMaster University; 2018:72.
49. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med 2014;48(13):1019-1023 [FREE Full text] [doi: 10.1136/bjsports-2014-093546] [Medline: 24782483]

XSL•FO
RenderX

50. Dijkstra B, Kamsma YP, Zijlstra W. Detection of gait and postures using a miniaturized triaxial accelerometer-based system: accuracy in patients with mild to moderate Parkinson's disease. Arch Phys Med Rehabil 2010;91(8):1272-1277. [doi: 10.1016/j.apmr.2010.05.004] [Medline: 20684910]
51. Gogniat MA, Won J, Cruz C, Aranda A, Verma A, Gujral S. Sedentary behavior, cognition, and brain health in older adults: a systematic review. Front Aging Neurosci 2025;17:1622049. [doi: 10.3389/fnagi.2025.1622049] [Medline: 40697359]
52. Wullems JA, Verschueren SMP, Degens H, Morse CI, Onambélé GL. A review of the assessment and prevalence of sedentarism in older adults, its physiology/health impact and non-exercise mobility counter-measures. Biogerontology 2016;17(3):547-565. [doi: 10.1007/s10522-016-9640-1] [Medline: 26972899]
53. Aguilar-Farías N, Brown WJ, Peeters GGMEE. ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. J Sci Med Sport 2014;17(3):293-299. [doi: 10.1016/j.jsams.2013.07.002] [Medline: 23932934]
54. Phillips SM, Clevenger K, Bruijns B, Tucker T. Effect of accelerometer cut-points on preschoolers' physical activity and sedentary time: a systematic review and meta-analysis. J Meas Phys Behav 2024;7(1). [doi: 10.1123/jmpb.2023-0060]
55. Rosenberger M, Haskell WL, Albinali F, Mota S, Nawyn J, Intille S. Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. Med Sci Sports Exerc 2013;45(5):964-975. [doi: 10.1249/MSS.0b013e31827f0d9c] [Medline: 23247702]
56. Montoye A, Pivarnik JM, Mudd LM, Biswas S, Pfeiffer KA. Validation and comparison of accelerometers worn on the hip, thigh, and wrists for measuring physical activity and sedentary behavior. AIMS Public Health 2016;3(2):298-312. [doi: 10.3934/publichealth.2016.2.298] [Medline: 29546164]
57. Koster A, Shiroma EJ, Caserotti P, Matthews CE, Chen KY, Glynn NW, et al. Comparison of sedentary estimates between activPAL and hip- and wrist-worn actiGraph. Med Sci Sports Exerc 2016;48(8):1514-1522 [FREE Full text] [doi: 10.1249/MSS.0000000000000924] [Medline: 27031744]
58. Yano S, Koohsari MJ, Shibata A, Ishii K, Frehlich L, McCormack GR, et al. Physical activity and sedentary behavior assessment: a laboratory-based evaluation of agreement between commonly used actiGraph and omron accelerometers. Int J Environ Res Public Health 2019;16(17):3126. [doi: 10.3390/ijerph16173126] [Medline: 31466248]
59. Boerema ST, van Velsen L, Vollenbroek MM, Hermens HJ. Pattern measures of sedentary behaviour in adults: a literature review. Digit Health 2020;6:2055207620905418. [doi: 10.1177/2055207620905418] [Medline: 32095261]
60. Janssen X, Basterfield L, Parkinson KN, Pearce MS, Reilly JK, Adamson AJ, Gateshead Millennium Study Core Team. Objective measurement of sedentary behavior: impact of non-wear time rules on changes in sedentary time. BMC Public Health 2015 May 23;15:504 [FREE Full text] [doi: 10.1186/s12889-015-1847-6] [Medline: 26001579]
61. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
62. Bai W, Chen P, Cai H, Zhang Q, Su Z, Cheung T, et al. Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: a meta-analysis and systematic review of epidemiology studies. Age Ageing 2022 Aug 02;51(8):afac173 [FREE Full text] [doi: 10.1093/ageing/afac173] [Medline: 35977150]
63. Song WX, Wu WW, Zhao YY, Xu HL, Chen GC, Jin SY, et al. Evidence from a meta-analysis and systematic review reveals the global prevalence of mild cognitive impairment. Front Aging Neurosci 2023;15:1227112 [FREE Full text] [doi: 10.3389/fnagi.2023.1227112] [Medline: 37965493]
64. Yan S, Fu W, Wang C, Mao J, Liu B, Zou L, et al. Association between sedentary behavior and the risk of dementia: a systematic review and meta-analysis. Transl Psychiatry 2020;10(1):112. [doi: 10.1038/s41398-020-0799-5] [Medline: 32317627]
65. Gonçalves AC, Cruz J, Marques A, Demain S, Samuel D. Evaluating physical activity in dementia: a systematic review of outcomes to inform the development of a core outcome set. Age Ageing 2018;47(1):34-41. [doi: 10.1093/ageing/afx135] [Medline: 28985262]
66. Thomas KR, Edmonds EC. Objective criteria for subtle cognitive decline in aging and preclinical Alzheimer disease: a systematic review. Neurology 2025;104(9):e213536. [doi: 10.1212/WNL.0000000000213536] [Medline: 40198863]
67. Liu Y, Chen S, Li S, Pei Z, Fan S, Guo Y. Diagnostic and inclusion criteria in alzheimer's disease clinical trials: a systematic review of the past decade. J Alzheimers Dis Rep 2025;9:25424823251362444. [doi: 10.1177/25424823251362444] [Medline: 40755858]
68. Jawad BN, Petersen J, Andersen O, Pedersen MM. Variations in physical activity and sedentary behavior during and after hospitalization in acutely admitted older medical patients: a longitudinal study. BMC Geriatr 2022;22(1):209. [doi: 10.1186/s12877-022-02917-8] [Medline: 35291952]
69. Jasper U, Yadav L, Dollard J, Jadczak AD, Yu S, Visvanathan R. Sedentary behaviour in hospitalised older people: a scoping review. Int J Environ Res Public Health 2020 Dec 14;17(24):9359 [FREE Full text] [doi: 10.3390/ijerph17249359] [Medline: 33327552]
70. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]

XSL•FO
RenderX

71. Quality assessment tool for observational cohort and cross-sectional studies. National Heart, Lung, and Blood Institute study quality assessment tools website. 2022. URL: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools? [accessed 2026-06-09]
72. Mc Ardle R, Jabbar KA, Del Din S, Thomas AJ, Robinson L, Kerse N, et al. Using digital technology to quantify habitual physical activity in community dwellers with cognitive impairment: systematic review. J Med Internet Res 2023;25:e44352 [FREE Full text] [doi: 10.2196/44352] [Medline: 37200065]
73. Lord S, Chastin SFM, McInnes L, Little L, Briggs P, Rochester L. Exploring patterns of daily physical and sedentary behaviour in community-dwelling older adults. Age Ageing 2011;40(2):205-210. [doi: 10.1093/ageing/afq166] [Medline: 21239410]
74. Rochester L, Chastin SFM, Lord S, Baker K, Burn DJ. Understanding the impact of deep brain stimulation on ambulatory activity in advanced Parkinson's disease. J Neurol 2012;259(6):1081-1086. [doi: 10.1007/s00415-011-6301-9] [Medline: 22086738]
75. van der Berg JD, Willems PJB, van der Velde JHPM, Savelberg HHCM, Schaper NC, Schram MT, et al. Identifying waking time in 24-h accelerometry data in adults using an automated algorithm. J Sports Sci 2016 Oct;34(19):1867-1873 [FREE Full text] [doi: 10.1080/02640414.2016.1140908] [Medline: 26837855]
76. Abel B, Pomiersky R, Werner C, Lacroix A, Schäufele M, Hauer K. Day-to-day variability of multiple sensor-based physical activity parameters in older persons with dementia. Arch Gerontol Geriatr 2019;85:103911. [doi: 10.1016/j.archger.2019.103911] [Medline: 31400647]
77. Amagasa S, Inoue S, Murayama H, Fujiwara T, Kikuchi H, Fukushima N. Associations of sedentary and physically-active behaviors with cognitive-function decline in community-dwelling older adults: compositional data analysis from the NEIGE study. J Epidemiol 2020;30(11):503-508. [doi: 10.2188/jea.JE20190141] [Medline: 31656243]
78. Balbim G, Falck RS, Boa Sorte Silva NC, Kramer AF, Voss M, Liu-Ambrose T. The association of the 24-hour activity cycle profiles with cognition in older adults with mild cognitive impairment: a cross-sectional study. J Gerontol A Biol Sci Med Sci 2024 Jul 01;79(7):glae099 [FREE Full text] [doi: 10.1093/gerona/glae099] [Medline: 38642387]
79. Cerff B, Maetzler W, Sulzer P, Kampmeyer M, Prinzen J, Hobert MA. Home-based physical behavior in late stage parkinson disease dementia: differences between cognitive subtypes. Neurodegener Dis 2017;17(4-5):135-144. [doi: 10.1159/000460251] [Medline: 28441649]
80. Clina JG, Bodde AE, Chang J, Helsel BC, Sherman JR, Vidoni ED. Factors associated with physical activity in Alzheimer's disease: a cross-sectional study of individuals and their caregivers. J Aging Health 2026;38(3-4):148-158. [doi: 10.1177/08982643251318766] [Medline: 39894786]
81. Falck RS, Landry GJ, Best JR, Davis JC, Chiu BK, Liu-Ambrose T. Cross-sectional relationships of physical activity and sedentary behavior with cognitive function in older adults with probable mild cognitive impairment. Phys Ther 2017;97(10):975-984 [FREE Full text] [doi: 10.1093/ptj/pzx074] [Medline: 29029554]
82. Finnanger Garshol B, Ellingsen-Dalskau LH, Pedersen I. Physical activity in people with dementia attending farm-based dementia day care - a comparative actigraphy study. BMC Geriatr 2020;20(1):219. [doi: 10.1186/s12877-020-01618-4] [Medline: 32571238]
83. Hartman Y, Karssemeijer E, van Diepen LAM, Olde Rikkert MGM, Thijssen D. Dementia patients are more sedentary and less physically active than age- and sex-matched cognitively healthy older adults. Dement Geriatr Cogn Disord 2018;46(1-2):81-89. [doi: 10.1159/000491995] [Medline: 30145584]
84. Hopkins J, McVeigh J, Hill K, Ellis KA, Jacques A, Burton E. Associations between physical activity, sedentary behaviour and cognitive domain performance of people living with mild cognitive impairment in the community. Aust Occup Ther J 2024;71(4):527-539. [doi: 10.1111/1440-1630.12944] [Medline: 38616178]
85. Lu Z, Harris TB, Shiroma EJ, Leung J, Kwok T. Patterns of physical activity and sedentary behavior for older adults with Alzheimer's disease, mild cognitive impairment, and cognitively normal in Hong Kong. J Alzheimers Dis 2018;66(4):1453-1462. [doi: 10.3233/JAD-180805] [Medline: 30412502]
86. Marmeleira J, Ferreira S, Raimundo A. Physical activity and physical fitness of nursing home residents with cognitive impairment: a pilot study. Exp Gerontol 2017;100:63-69. [doi: 10.1016/j.exger.2017.10.025] [Medline: 29107061]
87. Muurling M, Badissi M, de Boer C, Legdeur N, Barkhof F, van Berckel BNM. Physical activity levels in cognitively normal and cognitively impaired oldest-old and the association with dementia risk factors: a pilot study. BMC Geriatr 2023;23(1):129. [doi: 10.1186/s12877-023-03814-4] [Medline: 36882690]
88. Parry S, Chow M, Batchelor F, Fary RE. Physical activity and sedentary behaviour in a residential aged care facility. Australas J Ageing 2019;38(1):E12-E18. [doi: 10.1111/ajag.12589] [Medline: 30281184]
89. Rackoll T, Neumann K, Passmann S, Grittner U, Külzow N, Ladenbauer J, et al. Applying time series analyses on continuous accelerometry data-a clinical example in older adults with and without cognitive impairment. PLoS One 2021;16(5):e0251544. [doi: 10.1371/journal.pone.0251544] [Medline: 33984029]
90. Resnick B, Boltz M, Galik E, Fix S, Zhu S. Feasibility, reliability, and validity of the MotionWatch 8 to evaluate physical activity among older adults with and without cognitive impairment in assisted living settings. J Aging Phys Act 2021;29(3):391-399 [FREE Full text] [doi: 10.1123/japa.2020-0198] [Medline: 33361499]

XSL•FO
RenderX

91. van Alphen HJM, Volkers KM, Blankevoort CG, Scherder EJA, Hortobágyi T, van Heuvelen MJG. Older adults with dementia are sedentary for most of the day. PLoS One 2016;11(3):e0152457. [doi: 10.1371/journal.pone.0152457] [Medline: 27031509]
92. Varma VR, Watts A. Daily physical activity patterns during the early stage of Alzheimer's disease. J Alzheimers Dis 2017;55(2):659-667. [doi: 10.3233/JAD-160582] [Medline: 27716669]
93. Hartman Y, Karssemeijer E, van Diepen LAM, Olde Rikkert MGM, Thijssen D. Dementia patients are more sedentary and less physically active than age- and sex-matched cognitively healthy older adults. Dement Geriatr Cogn Disord 2018;46(1-2):81-89. [doi: 10.1159/000491995] [Medline: 30145584]
94. Saunders TJ, Larouche R, Colley RC, Tremblay MS. Acute sedentary behaviour and markers of cardiometabolic risk: a systematic review of intervention studies. J Nutr Metab 2012;2012:712435 [FREE Full text] [doi: 10.1155/2012/712435] [Medline: 22754695]
95. McCurry SM, Pike KC, Logsdon RG, Vitiello MV, Larson EB, Teri L. Predictors of short- and long-term adherence to a daily walking program in persons with Alzheimer's disease. Am J Alzheimers Dis Other Demen 2010;25(6):505-512 [FREE Full text] [doi: 10.1177/1533317510376173] [Medline: 20660515]
96. Auyeung TW, Kwok T, Lee J, Leung PC, Leung J, Woo J. Functional decline in cognitive impairment--the relationship between physical and cognitive function. Neuroepidemiology 2008;31(3):167-173 [FREE Full text] [doi: 10.1159/000154929] [Medline: 18784415]
97. Vital TM, Hernandez SSS, Stein AM, Garuffi M, Corazza DI, de Andrade LP, et al. Depressive symptoms and level of physical activity in patients with Alzheimer's disease. Geriatr Gerontol Int 2012;12(4):637-642. [doi: 10.1111/j.1447-0594.2011.00830.x] [Medline: 22300241]
98. Christofoletti G, Oliani MM, Bucken-Gobbi LT, Gobbi S, Beinotti F, Stella F. Physical activity attenuates neuropsychiatric disturbances and caregiver burden in patients with dementia. Clinics (Sao Paulo) 2011;66(4):613-618. [doi: 10.1590/s1807-59322011000400015] [Medline: 21655755]
99. Allan L, McKeith I, Ballard C, Kenny RA. The prevalence of autonomic symptoms in dementia and their association with physical activity, activities of daily living and quality of life. Dement Geriatr Cogn Disord 2006;22(3):230-237. [doi: 10.1159/000094971] [Medline: 16902277]
100. Cedervall Y, Kilander L, Aberg AC. Declining physical capacity but maintained aerobic activity in early Alzheimer's disease. Am J Alzheimers Dis Other Demen 2012;27(3):180-187. [doi: 10.1177/1533317512442996] [Medline: 22573284]
101. Stubbs B, Eggermont L, Soundy A, Probst M, Vandenbulcke M, Vancampfort D. What are the factors associated with physical activity (PA) participation in community dwelling adults with dementia? A systematic review of PA correlates. Arch Gerontol Geriatr 2014;59(2):195-203. [doi: 10.1016/j.archger.2014.06.006] [Medline: 25034708]
102. van Alphen HJ, Hortobágyi T, van Heuvelen MJ. Barriers, motivators, and facilitators of physical activity in dementia patients: a systematic review. Arch Gerontol Geriatr 2016;66:109-118. [doi: 10.1016/j.archger.2016.05.008] [Medline: 27295140]
103. Tiemensma MD, Andersen PT, Meijering L, Sturge J. “If you just sit at home and look out of the window, then there is no life.” An ethnographic study of how home-dwelling people with dementia use the cityscape's life in practice. SSM Qual Res Health 2024;5:100448. [doi: 10.1016/j.ssmqr.2024.100448]
104. Alonzo AA. Everyday illness behavior: a situational approach to health status deviations. Soc Sci Med A 1979;13(4):397-404. [doi: 10.1016/0160-7979(79)90128-0]
105. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. Sports Med 2017;47(9):1821-1845. [doi: 10.1007/s40279-017-0716-0] [Medline: 28303543]
106. Boyle PA, Cohen RA, Paul R, Moser D, Gordon N. Cognitive and motor impairments predict functional declines in patients with vascular dementia. Int J Geriatr Psychiatry 2002;17(2):164-169. [doi: 10.1002/gps.539] [Medline: 11813280]
107. Matar E, Shine JM, Halliday GM, Lewis SJG. Cognitive fluctuations in Lewy body dementia: towards a pathophysiological framework. Brain 2020;143(1):31-46. [doi: 10.1093/brain/awz311] [Medline: 31612904]
108. Sasai H. Assessing sedentary behavior using wearable devices: an overview and future directions. J Phys Fit Sports Med 2017;6(3):135-143. [doi: 10.7600/jpfsm.6.135]
109. Abdul Jabbar K, Sarvestan J, Zia Ur Rehman R, Lord S, Kerse N, Teh R, et al. Validation of an algorithm for measurement of sedentary behaviour in community-dwelling older adults. Sensors (Basel) 2023;23(10):4605. [doi: 10.3390/s23104605] [Medline: 37430519]
110. Rowlands A, Olds TS, Hillsdon M, Pulsford R, Hurst TL, Eston RG, et al. Assessing sedentary behavior with the GENEActiv: introducing the sedentary sphere. Med Sci Sports Exerc 2014;46(6):1235-1247. [doi: 10.1249/MSS.0000000000000224] [Medline: 24263980]
111. Rowlands AV, Yates T, Olds TS, Davies M, Khunti K, Edwardson CL. Wrist-worn accelerometer-brand independent posture classification. Med Sci Sports Exerc 2016 Apr;48(4):748-754. [doi: 10.1249/MSS.0000000000000813] [Medline: 26559451]

XSL•FO
RenderX

112. Clarke-Cornwell AM, Farragher TM, Cook PA, Granat MH. Empirically derived cut-points for sedentary behaviour: are we sitting differently? Physiol Meas 2016 Oct;37(10):1669-1685 [FREE Full text] [doi: 10.1088/0967-3334/37/10/1669] [Medline: 27652920]
113. Gorman E, Hanson HM, Yang PH, Khan KM, Liu-Ambrose T, Ashe MC. Accelerometry analysis of physical activity and sedentary behavior in older adults: a systematic review and data analysis. Eur Rev Aging Phys Act 2014;11(1):35-49 [FREE Full text] [doi: 10.1007/s11556-013-0132-x] [Medline: 24765212]
114. Tanha T, Tornberg Å, Dencker M, Wollmer P. Accelerometer measured daily physical activity and sedentary pursuits--comparison between two models of the Actigraph and the importance of data reduction. BMC Res Notes 2013 Oct 31;6:439 [FREE Full text] [doi: 10.1186/1756-0500-6-439] [Medline: 24176143]
115. Winkler EAH, Gardiner PA, Clark BK, Matthews CE, Owen N, Healy GN. Identifying sedentary time using automated estimates of accelerometer wear time. Br J Sports Med 2012;46(6):436-442 [FREE Full text] [doi: 10.1136/bjsm.2010.079699] [Medline: 21504965]
116. Giné-Garriga M, Sansano-Nadal O, Tully MA, Caserotti P, Coll-Planas L, Rothenbacher D. Accelerometer-measured sedentary and physical activity time and their correlates in European older adults: the SITLESS study. J Gerontol A Biol Sci Med Sci 2020;75(9):1754-1762 [FREE Full text] [doi: 10.1093/gerona/glaa016] [Medline: 31943000]
117. Heesch KC, Hill RL, Aguilar-Farias N, van Uffelen JGZ, Pavey T. Validity of objective methods for measuring sedentary behaviour in older adults: a systematic review. Int J Behav Nutr Phys Act 2018;15(1):119 [FREE Full text] [doi: 10.1186/s12966-018-0749-2] [Medline: 30477509]
118. Kuster RP. Where to place which sensor to measure sedentary behavior? A method development and comparison among various sensor placements and signal types. J Meas Phys Behav 2020;3(4):274-284. [doi: 10.1123/jmpb.2019-0060]
119. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM modeling good research practices task force-7. Med Decis Making 2012;32(5):733-743. [doi: 10.1177/0272989X12454579] [Medline: 22990088]
120. Swan K, Speyer R, Scharitzer M, Farneti D, Brown T, Woisard V, et al. Measuring what matters in healthcare: a practical guide to psychometric principles and instrument development. Front Psychol 2023;14:1225850 [FREE Full text] [doi: 10.3389/fpsyg.2023.1225850] [Medline: 37790221]
121. Del Din S, Galna B, Godfrey A, Bekkers EMJ, Pelosin E, Nieuwhof F, et al. Analysis of free-living gait in older adults with and without Parkinson's disease and with and without a history of falls: identifying generic and disease-specific characteristics. J Gerontol A Biol Sci Med Sci 2019 Mar 14;74(4):500-506 [FREE Full text] [doi: 10.1093/gerona/glx254] [Medline: 29300849]
122. Hickey A, Del Din S, Rochester L, Godfrey A. Detecting free-living steps and walking bouts: validating an algorithm for macro gait analysis. Physiol Meas 2017;38(1):N1-N15. [doi: 10.1088/1361-6579/38/1/N1] [Medline: 27941238]

Abbreviations

AD: Alzheimer disease
CPM: counts per minute
MCI: mild cognitive impairment
MET: metabolic equivalent task
PA: physical activity
SB: sedentary behavior

Edited by H LaMonica; submitted 21.Oct.2025; peer-reviewed by R Stojan, W Peng; comments to author 03.Mar.2026; revised version received 07.Apr.2026; accepted 28.May.2026; published 25.Jun.2026.

Please cite as:
Wales JL, Hinchliffe C, Falck RS, Del Din S, Yarnall AJ, Mc Ardle R
Use of Wearable Technology for Measuring and Characterizing Sedentary Behavior in People With Mild Cognitive Impairment and Dementia: Systematic Review
JMIR Aging 2026;9:e85361
URL: https://aging.jmir.org/2026/1/e85361
doi:10.2196/85361
PMID:

©Jenny L Wales, Chloe Hinchliffe, Ryan Stanley Falck, Silvia Del Din, Alison J Yarnall, Ríona Mc Ardle. Originally published in JMIR Aging (https://aging.jmir.org), 25.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Effects of Pure Cognitive Training on Gait and Balance in Older Adults: Systematic Review and Meta-Analysis

Xiaohan Li[1,2], MSc, MPT; Yat Leung Chan[1], BScPT; Guoshi Liang[1], MSc; Cheuk Hin Lau[1], BScPT; Tsz Yau Yuki Lam[1], BScPT; Chak Hang Lam[1], BScPT; Tsz Wing Chan[1], BScPT; Hoi Kei Cheung[1], BSc; Wayne Lap Sun Chan[1], BScPT, PhD; Chun Liang Hsu[1], PhD; Freddy Man Hin Lam[1], BScPT, PhD

[1]Department of Rehabilitation Sciences, Faculty of Health and Social Sciences, The Hong Kong Polytechnic University, Kowloon, China (Hong Kong)
[2]Department of Allied Health, School of Health Sciences, Swinburne University of Technology, Victoria, Australia

Corresponding Author:
Freddy Man Hin Lam, BScPT, PhD
Department of Rehabilitation Sciences
Faculty of Health and Social Sciences
The Hong Kong Polytechnic University
11 Yuk Choi Road, Hung Hom
Kowloon
China (Hong Kong)
Phone: 852 27666720
Email: freddy-mh.lam@polyu.edu.hk

Abstract

Background: Physical and cognitive function, both of which decline with aging, are significantly interrelated. Although numerous studies have investigated the effect of physical exercise on cognitive function, relatively few have examined the impact of pure cognitive training on physical performance.

Objective: This review aimed to summarize the effects of pure cognitive training on balance and mobility in older adults.

Methods: Electronic databases (PubMed, Embase, CINAHL, and PsycInfo) were searched in February 2025. Randomized controlled trials that investigated the effect of pure cognitive training on balance and mobility in older adults were included. Pure cognitive training refers to a strictly nonphysical approach involving guided practice on a standardized set of cognitive tasks aimed at optimizing cognitive functioning. For the outcomes, balance performance focused on standing balance tests and comprehensive balance assessment scales. Mobility performance primarily focused on gait speed under a single-task condition (ie, walking only), gait speed under a dual-task condition (ie, conducting cognitive tasks while walking), and the Timed Up and Go test. The Physiotherapy Evidence Database (PEDro) scale was used to evaluate methodological quality. The level of evidence for the available outcomes was rated using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system. Meta-analyses and sensitivity analyses of studies with high methodological quality were performed if 3 or more studies were obtained.

Results: Fourteen studies were eventually included. The methodological qualities of 8 (57%) studies were rated as good, while those of 6 (43%) studies were rated as fair. A meta-analysis of studies with high methodological quality showed that pure cognitive training has a significant effect on cognitive-motor dual-task gait speed (standardized mean difference [SMD] 0.39, 95% CI 0.04-0.75; n=376 participants; moderate-quality evidence) but not single-task gait speed (SMD 0.12, 95% CI −0.05 to 0.30; n=506 participants; moderate-quality evidence). No significant improvements were found in functional mobility (SMD 0.29, 95% CI −0.60 to 1.18; n=566 participants; low-quality evidence) or balance (SMD 0.18, 95% CI −0.34 to 0.69; n=139 participants; very low-quality evidence).

Conclusions: Cognitive training does not improve balance or mobility under single-task conditions but does improve walking speed under cognitive-motor dual-task conditions in older adults. A cognitive training protocol targeting executive function, consisting of 40- to 60-minute sessions conducted 2 to 3 times per week over a period of 8 to 10 weeks, has been shown to be effective.

(JMIR Aging 2026;9:e85070) doi:10.2196/85070

XSL•FO
RenderX

KEYWORDS

computerized cognitive training; older adults; aging; physical performance; balance; gait

Introduction

Aging is associated with a decline in functional capacity, which could be attributed to decreased physical and cognitive functions. Existing evidence has demonstrated that physical and cognitive functions are closely interrelated. For instance, faster gait speed is positively correlated with better memory, executive function, and overall cognitive abilities [1]. Greater gait variability is independently linked to poorer cognitive performance among older adults regardless of age [2,3]. Higher cognitive function scores, as assessed by the Mini-Mental State Examination (MMSE), independently predicted better performance in functional mobility, independent of age and other clinical factors [4,5], while mild cognitive impairment (MCI) was found to affect specific gait parameters and functional mobility performance [6].

In terms of causality, substantial research has been conducted to evaluate the effect of physical exercise on cognitive functions. One systematic review found that physical exercise reduces global cognitive decline and improves working memory in individuals with MCI or dementia [7]. Another demonstrated that physical exercise enhances cognitive function across older adults with different levels of cognitive performance, ranging from intact cognition to MCI and dementia [8]. The improvement in cognitive function following physical exercise could be attributed to the increased cerebral perfusion and cardiac output during exercise [8]. Furthermore, physical exercise was found to stimulate the secretion of brain-derived neurotrophic factor (BDNF) in the hippocampi, increase the production of myokines such as irisin and insulin-like growth factor-1 (IGF-1), and reduce the accumulation and aggregation of β-amyloid in the brain [9-12]. These physiological changes provide a foundation for cognitive performance. Collectively, these findings underscore the pivotal role of regular physical activity in preserving cognitive functions in older adults.

On the other hand, research has shown that cognitive processing plays an important role in physical functions, particularly gait and balance. Walking requires motor planning and visuospatial perception, which involve executive function, perceptual-motor function, and visual attention [13]. A progressive decline in frontal cognitive functions, such as that seen in Alzheimer disease, may cause concomitant disorganization of the neural networks that control locomotion, leading to impaired gait performance [13,14]. Similarly, balance is closely related to executive function, mental flexibility, and response inhibition. It has been reported that older adults with balance issues performed worse on tasks requiring mental flexibility and response inhibition [15]. Therefore, there has been an increasing interest in the potential of cognitive training to improve physical function. A recent meta-analysis concluded that sequential programs combining cognitive training and exercise effectively enhance both cognitive and physical functions in healthy older adults. However, this review was unable to isolate the specific impact of cognitive training on physical function [16].

To date, a systematic review published in 2018 [17] evaluated the effect of cognitive training without a physical exercise component on physical function. It revealed that cognitive training can improve gait performance under complex task conditions but not under single-task conditions. On the basis of these findings, investigating the effect of pure cognitive training on mobility and balance, alongside gait, would offer a more comprehensive view of its impact on physical function, given that both domains are closely linked to cognition [1,18]. Furthermore, with new studies published in the past few years [19,20], it is timely to consolidate the latest evidence on the effects of cognitive training on physical function in older adults.

Therefore, this systematic review aims to isolate and evaluate the impacts of single-task cognitive training targeting any cognitive function, including computerized cognitive training (CCT) [21], video games, and cognitive remediation, on mobility and balance outcomes in older adults. The findings will provide an updated understanding of the relationship between cognitive training and physical outcomes.

Methods

Data Sources and Searches

A systematic review was conducted to investigate the effects of cognitive training on the mobility and balance of healthy older adults. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guideline was followed [22]. Searches were performed in 4 electronic databases (PubMed, Embase, CINAHL, and PsycInfo). The initial search was performed on September 8, 2022, and the final search was done on February 7, 2025. The population, intervention, comparison, outcome, and study design (PICOS) framework was followed for the search strategy. The keywords used in the different databases are shown in Multimedia Appendix 1. The protocol for this systematic review was registered with PROSPERO prior to data analysis (CRD42023401100) [23].

Study Selection

Search records were screened based on their titles and abstracts to exclude irrelevant articles. The full texts of the remaining articles were reviewed based on the PICOS framework, including the following components: participants (people aged ≥65 years), intervention (pure cognitive training), comparison (usual care or health education), outcome (physical performance), and study design (randomized controlled trial [RCT]). Therefore, this review included studies that (1) recruited older adults aged ≥65 years; (2) provided pure cognitive training as a strictly nonphysical approach involving guided practice on a standardized set of cognitive tasks aimed at optimizing cognitive functioning or slowing brain aging [24,25]; (3) included a comparison group receiving no cognitive training; (4) assessed outcomes related to balance, mobility, or gait under single-task conditions, which refer to the execution of a motor task in isolation, and cognitive-motor dual-task conditions, which involve the simultaneous execution of a primary motor

XSL•FO
RenderX

task alongside a concurrent cognitive task; and (5) used a RCT study design. Studies were excluded if they (1) included participants with cognitive impairments (eg, dementia or MCI) or neurological diseases known to impair gait (eg, Parkinson disease and stroke) or (2) applied only cognitive-motor dual-task interventions, preventing the specific effects of cognitive training from being isolated.

Methodological Quality Assessment and Quality of Evidence

The Physiotherapy Evidence Database (PEDro) scale was used to assess the methodological quality of each selected study [26]. The PEDro scale is a formally validated instrument with established reliability and validity for quantifying the methodological quality of RCTs. For articles that were not available on the PEDro website, they were rated independently by 2 of our team members using the PEDro criteria. A higher PEDro score indicates better methodological quality. The studies were then classified as excellent, good, or fair according to their PEDro score (9-10="excellent," 6-8="good," and 4-5="fair"). The quality of evidence for the outcomes was rated using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach [27-30]. The downgrade and upgrade criteria of GRADE are listed in Multimedia Appendix 2.

Data Analysis

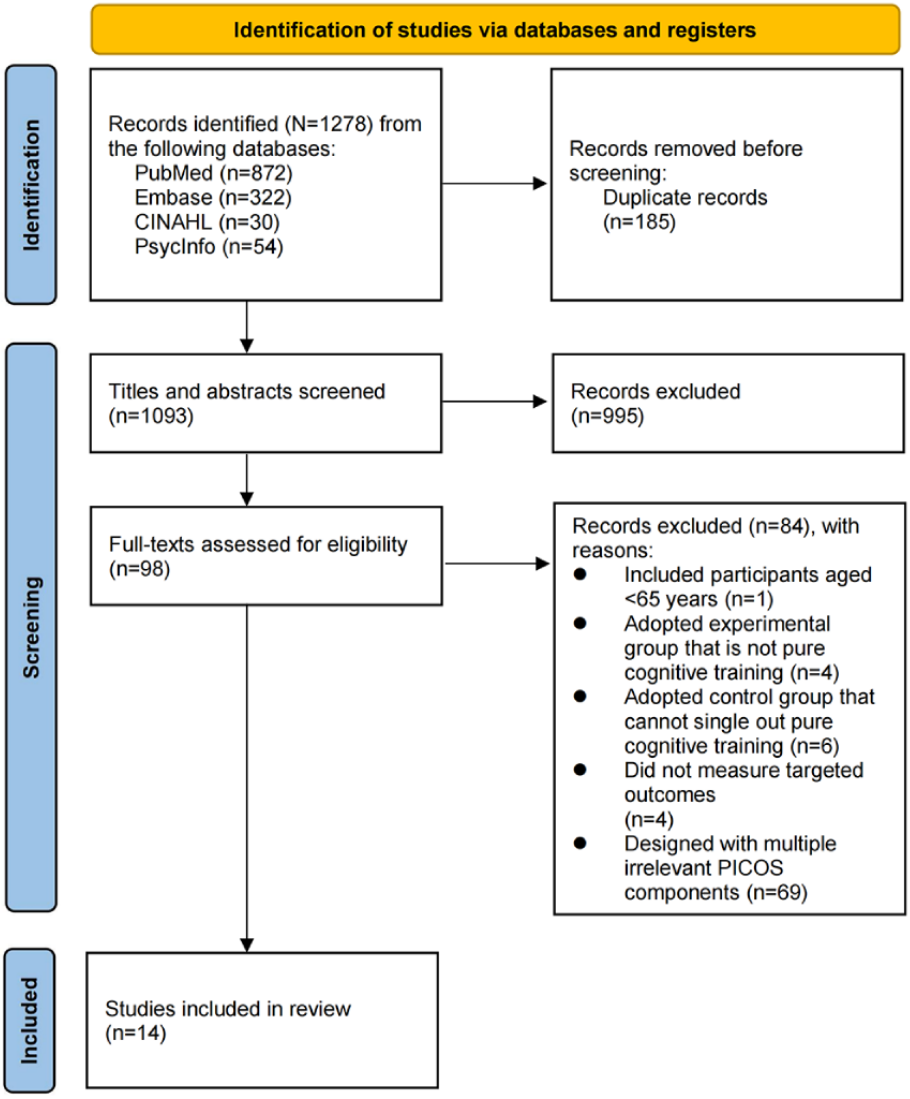
The effects of cognitive training on the outcomes of interest were categorized into mobility and balance. Meta-analysis was performed if ≥3 studies of similar outcomes were obtained. Random-effects models were used for all meta-analyses to address heterogeneity across studies. The standardized mean difference (SMD) and 95% CI were computed and presented in forest plots (Review Manager version 5.4.1; Cochrane Collaboration). A *P* value of ≤.05 was considered statistically significant. The I^2 index was used to test the heterogeneity of the included studies in each meta-analysis. A separate meta-analysis including studies with high methodological quality (PEDro score ≥6) was performed to avoid contamination of the results by low-quality studies. For meta-analyses with moderate to high heterogeneity (I^2>50%), subgroup analyses were performed if all subgroups could include at least 2 studies [31]. Publication bias was statistically assessed using the Egger regression asymmetry test to detect potential asymmetries in the evidence base if there were >10 studies included in the meta-analysis [32]. A *P* value <.10 on the Egger regression asymmetry test indicated the presence of publication bias [33]. The Egger regression method is a statistical tool used in meta-analyses to detect publication bias by assessing whether smaller studies with less significant findings are underrepresented in the published literature. When publication bias was detected, a trim and fill analysis was conducted to impute the findings of potentially unpublished studies and assess the robustness of the results. Narrative analyses were performed for outcomes for which meta-analyses could not be performed.

Results

Our search yielded a total of 1278 records. Ultimately, 14 (1%) trials [19,20,34-45] with a total of 898 participants met the eligibility criteria and were included in this systematic review (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flow diagram of the study selection process. PICOS: population, intervention, comparison, outcome, and study design; RCT: randomized controlled trial.



Study Characteristics

The characteristics of the included studies are summarized in Table 1. The sample size ranged from 20 to 238, with the mean age of participants ranging from 66.2 (SD 7.1) to 85.7 (SD 5.4) years. Most studies recruited community-dwelling individuals. One study included nursing home residents [42]. One study included only male participants [35]. The details of each study are listed in Multimedia Appendices 3 and 4 [1-14].

XSL•FO
RenderX

Table 1. Characteristics of the included studies (N=14).

Characteristics	Studies, n (%)	References
Total sample size		
<50	8 (57)	[34-38,41-43]
50-99	4 (29)	[39,40,44,45]
≥100	2 (14)	[19,20]
Mean age (years)		
60-64	0 (0)	N/A[a]
65-74	5 (36)	[20,37,39,44,45]
≥75	9 (64)	[19,34-36,38,40-43]
Sex		
Female participants only	0 (0)	N/A
Male participants only	1 (7)	[35]
Both sexes	13 (93)	[19,20,34,36-45]
Methodological quality (Physiotherapy Evidence Database scale)		
<4 (“poor”)	0 (0)	N/A
4-5 (“fair”)	6 (43)	[35-37,41,42,45]
6-8 (“good”)	8 (57)	[19,20,34,38-40,43,44]
Source of participants		
Community-dwelling older adults	12 (86)	[19,20,34,36-41,43-45]
Nursing home residents	1 (7)	[42]
Not mentioned	1 (7)	[35]
Types of cognitive training		
CCT[b]	10 (71)	[19,20,35-38,40,41,43,44]
Non-CCT	4 (29)	[34,39,42,45]
Targeted outcomes[c]		
Single-task gait speed	13 (93)	[19,34-36,38-44,46,47]
Cognitive-motor dual-task gait speed	7 (50)	[19,38,40-44]
Functional mobility	6 (43)	[19,20,34,36,40]
Static balance (one-leg stance test)	3 (21)	[20,34,37,45]
Overall balance ability (balance scales)	3 (21)	[34,41,46]

[a]N/A: not applicable.

[b]CCT: computerized cognitive training.

[c]The percentages sum to more than 100% as individual studies may report multiple outcomes.

Intervention Protocols

A total of 71% (10/14) studies used CCT as the intervention [19,20,35-38,40,41,43,44], while 4 (29%) used noncomputerized training [34,39,42,45]. Studies that adopted CCT used training programs designed to improve executive function, which include Mindfit (CogniFit Inc), CogniFit Program (CogniFit Ltd), and other self-designed programs. Noncomputerized training included therapist-led visual attention training [34], attentional capacity and working memory training [39,42], and board game–based cognitive training [45]. The duration of the training programs ranged from 3 weeks to 12 months. Most interventions were conducted 1 to 3 times per week, with each session lasting between 20 and 120 minutes. One study [19] involved self-directed home-based training, in which participants completed flexible sessions at their own pace, averaging 120 to 150 minutes of training per week.

Methodological Quality and Quality of Evidence

Among the 14 included studies, 8 (57%) [19,20,34,38-40,43,44] were rated as good (PEDro score 6-8) and 6 (43%) [35-37,41,42,45] were rated as fair (PEDro score 4-5). The PEDro score of each article is outlined in Multimedia Appendix 5 [1-14]. The quality of evidence for outcomes rated using GRADE is listed in Table 2. The quality of evidence for gait speed under single-task and cognitive-motor dual-task conditions

XSL•FO
RenderX

was rated as moderate. The quality of evidence for functional mobility and balance was rated as low and very low, respectively.

Table 2. Grade of Recommendations Assessment, Development, and Evaluation (GRADE) assessment of the quality of evidence[a].

Outcomes	Risk of bias[b]	Inconsistency	Indirectness	Imprecision	Publication bias	Large effect	Plausible confounding	Dose response gradient	GRADE quality	Effect size (95% CI)
Gait speed	0	0	0	−1[c]	0	0	0	0	Moderate	0.19 (−0.00 to 0.38)
Gait speed (high-quality studies)	0	0	0	−1[c]	0	0	0	0	Moderate	0.12 (−0.05 to 0.30)
Dual-task gait speed	0	0	0	−1[c]	0	0	0	0	Moderate	0.26 (0.03 to 0.49)
Dual-task gait speed (high-quality studies)	0	0	0	−1[c]	0	0	0	0	Moderate	0.39 (0.04 to 0.75)
Functional mobility	0	−1[d]	0	−1[c]	0	0	0	0	Low	0.29 (−0.60 to 1.18)
Balance scale	−1[e]	0	0	−2[b]	0	0	0	0	Very low	0.18 (−0.34 to 0.69)

[a]0 indicates that no serious limitations were identified in that domain; therefore, no downgrading was applied. Negative values (−1 or −2) indicate downgrading of the quality of evidence due to identified limitations (eg, imprecision or inconsistency).

[b]The *P* value of the Egger regression asymmetry test is less than 0.1.

[c]The number of participants included in the primary meta-analysis was <400, and the 95% CI spanned the appreciable harm or benefit thresholds.

[d]Heterogeneity I^2 larger than 50% in the meta-analysis.

[e]More than 50% of the participants included in the primary analysis from trials with a Physiotherapy Evidence Database score of <6.

Mobility

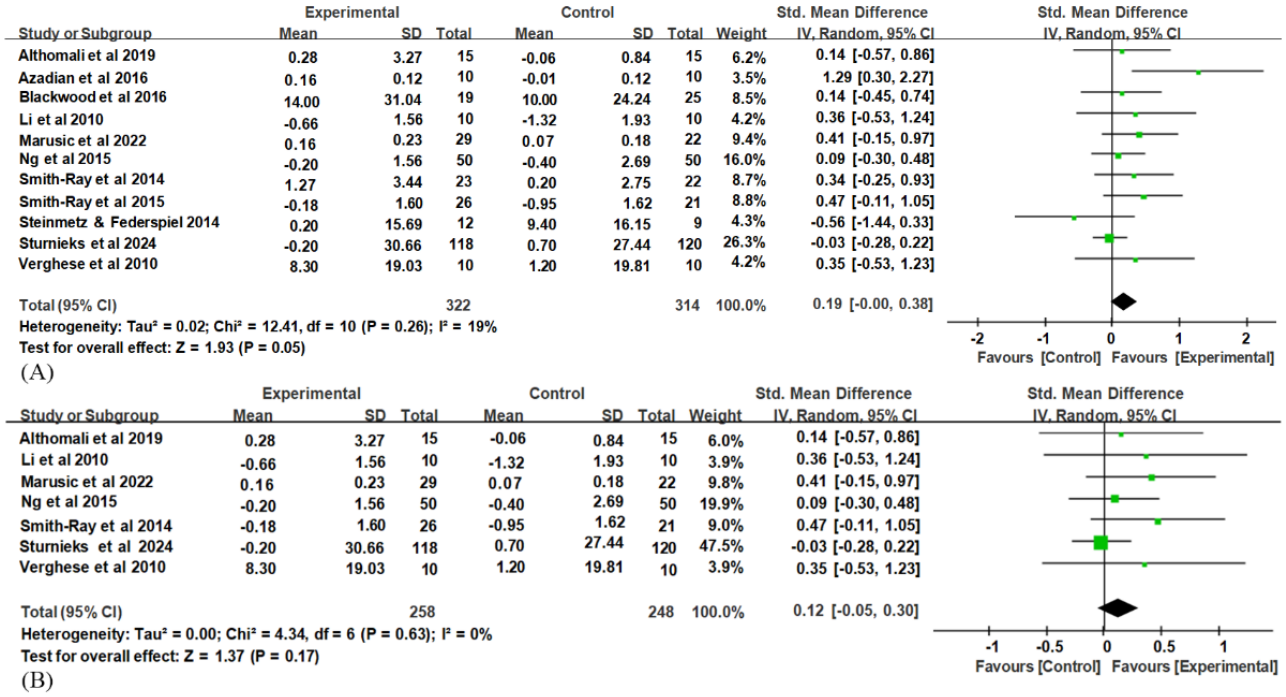
Fourteen studies [19,20,34-36,38-44,46,47] investigated the effect of cognitive training on mobility-related outcomes.

Single-Task Gait Speed

Of the 14 studies, 13 (93%) [19,34-36,38-44,46,47] investigated the effects of cognitive training on gait speed. Meta-analysis showed that cognitive training significantly increased single-task gait speed (SMD 0.19, 95% CI −0.00 to −0.38; Z=1.93; P=.05; I^2=19%; n=636 participants; moderate-quality evidence; Figure 2A). The Egger regression asymmetry test indicated no evidence of publication bias (P=.35). Yet, the result turned nonsignificant in the sensitivity analysis that included only studies with high methodological quality (SMD 0.12, 95% CI −0.05 to −0.30; Z=1.37; P=.17; I^2=0%; n=506 participants; moderate-quality evidence; Figure 2B) [19,34,38,39,41,43,44].

XSL•FO
RenderX

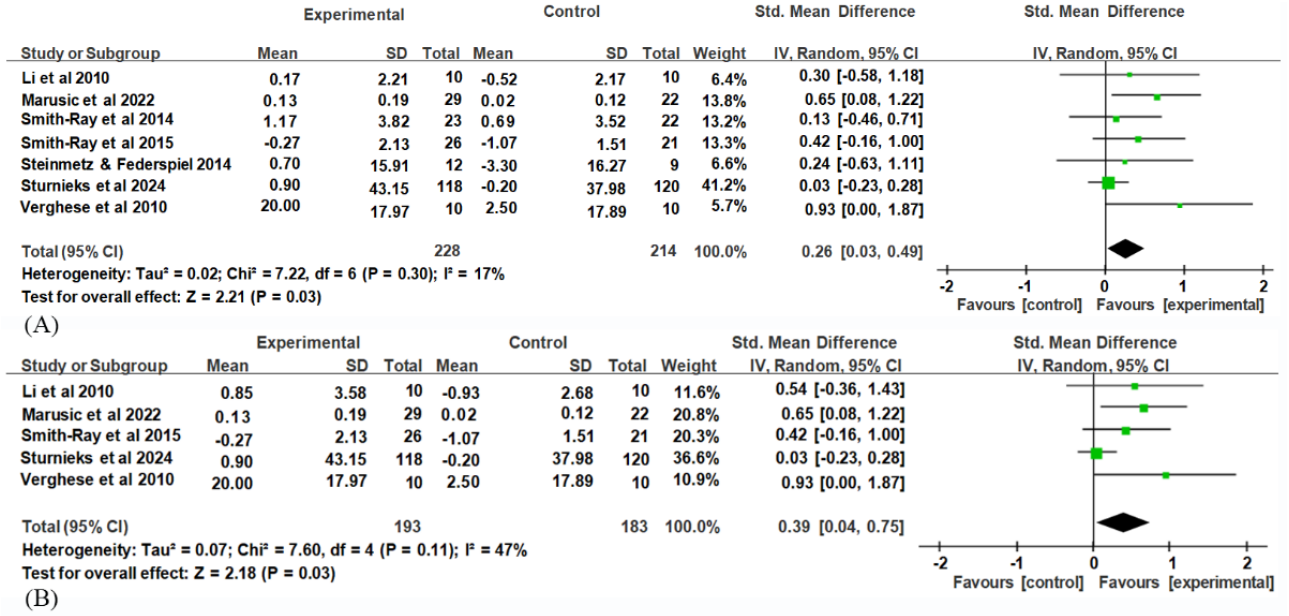
Figure 2. Gait speed. (A) Meta-analysis including all studies that reported gait speed outcomes and (B) meta-analysis including only studies rated as having high methodological quality.



Cognitive-Motor Dual-Task Gait Speed

A total of 50% (7/14) studies [19,38,40-44] investigated the effects of cognitive training on cognitive-motor dual-task gait speed. Meta-analysis showed that cognitive training significantly increased cognitive-motor dual-task gait speed (SMD 0.26, 95% CI 0.03 to −0.49; Z=2.21; P=.03; I^2=17%; n=442 participants; moderate-quality evidence; Figure 3A). The results remained significant in the sensitivity analysis that included only studies with high methodological quality (PEDro score ≥6; SMD 0.39, 95% CI 0.04 to −0.75; Z=2.18; P=.03; I^2=47%; n=376 participants; moderate-quality evidence; Figure 3B) [19,38,40,43,44].

Figure 3. Cognitive-motor dual-task gait speed. (A) Meta-analysis including all studies that reported cognitive-motor dual-task gait speed outcomes and (B) meta-analysis including only studies rated as having high methodological quality.



Functional Mobility

A total of 43% (6/14) studies [19,20,34,36,40,45] investigated functional mobility using the Timed Up and Go (TUG) test as an outcome measure. Meta-analysis showed that cognitive training had no significant effect in improving functional mobility (mean difference [MD] 0.29, 95% CI −0.60 to −1.18; Z=0.63; P=.53; I^2=77%; n=566 participants; low-quality evidence; Multimedia Appendix 6).

XSL•FO
RenderX

To delineate the contributing factor to the largest heterogeneity, subgroup analyses were conducted based on the methodological quality of the trials, the type of training (CCT vs non-CCT), total training hours of the program, individual training duration, and the baseline mean age of the participants. Nevertheless, the largest I^2 value ranged from 79% to 88% in all subgroup clusters. No clear contributing factor to the high heterogeneity could be identified. The results remained nonsignificant in all subgroup analyses (Multimedia Appendix 7).

Additionally, 7% (1/14) study reported no significant effect on dual-task TUG test (P=.53) [34].

Other Mobility Parameters

Only 7%(1/14) study [35] reported gait parameters other than gait speed, including stride and step length, stride duration, step, single-support duration, and double-support duration during single-task walking, and gait asymmetry during single-task and dual-task walking. Meta-analysis was not appropriate because of the insufficient number of studies. This study reported statistically significant improvements (P<.05) in some gait parameters and a significant decrease in gait asymmetry in the cognitive training group compared with the control group. Detailed results are presented in Multimedia Appendix 4.

Balance Performance

The effect of cognitive training on balance was evaluated in 5 studies [20,34,37,41,45].

Balance Scales

Of the 14 included studies, 2 (14%) used the Berg Balance Scale (BBS) and 1 (7%) used the Mini-Balance Evaluation Systems Test (Mini-BESTest) [34,41,45]. The 2 scales were combined in a meta-analysis. No significant effect was found in the cognitive training group compared with the control group (SMD 0.18, 95% CI −0.34 to −0.69; Z=0.67; P=.50; I^2=56%; n=139 participants; very-low–quality evidence; Multimedia Appendix 8). Despite the high heterogeneity, no subgroup analysis could be performed as there were only 3 studies in the meta-analysis. Any grouping would result in at least 1 subgroup containing only 1 study.

Static Balance

Only 21% (3/14) studies [20,34,37] reported static balance with contrasting results. The tests were carried out under either eyes-open [34] or eyes-closed [20,37] conditions. Althomali et al [34] reported that cognitive training had no significant effect on the duration of single-leg standing under eyes-open conditions relative to the control group (P=.56). Liu et al [20] reported that cognitive training led to a significant increase in the duration of single-leg standing under eyes-closed conditions when compared with the control group (P=.005). Hou et al [37] reported only a combined result of the one-leg stance test (OLST) and the TUG test, prohibiting the isolation of the balance outcome for analysis.

Discussion

Principal Findings

Our systematic review summarized the effects of cognitive training on mobility and balance in older adults. Overall, our meta-analysis showed that cognitive training significantly enhanced cognitive-motor dual-task gait speed (P=.03) and marginally improved single-task gait speed (P=.05). However, the results for single-task gait speed turned nonsignificant in the sensitivity analyses that included only studies with high methodological quality. No significant improvement was found in functional mobility or balance outcomes following cognitive training.

Effectiveness of Cognitive Training on Mobility and Balance Outcomes

Overview

Our meta-analysis revealed the effect of cognitive training on balance and mobility performance under different cognitive loads. According to the dual-task paradigm [48], single-task walking, functional mobility tests (ie, TUG test), and balance assessment tests (ie, BBS, Mini-BESTest, and OLST) were considered tasks with a low cognitive load, while cognitive-motor dual-task walking was considered a task with a high cognitive load [49]. Thus, results consolidated in this review suggest that cognitive training could improve only tasks with a high cognitive load and not those with a low cognitive load.

Effectiveness of Cognitive Training on Tasks With a High Cognitive Load

One of the possible explanations for this finding is the central capacity sharing model [50]. Cognitive-motor dual-task walking has been described as involving 2 simultaneous tasks (walking and a cognitive task) that interfere with each other and compete for cognitive resources. It has been stated that when attentional resources are limited in capacity, at least one of the tasks may deteriorate in performance. Among the 7 studies reporting cognitive-motor dual-task gait speed outcomes, 4 (57%) [38,40-42] did not assess changes in cognitive performance, 2 (29%) reported no significant change in cognitive function [19,43], and 1 (14%) demonstrated enhanced executive function [44]. Additionally, within the studies [43,44] with a significant improvement in dual-task gait speed, 1 (50%) study showed an improvement in cognitive function [44]. These findings suggest that the observed improvements in gait performance might stem from (1) enhanced cognitive capacity; (2) improved automaticity in performing cognitive tasks [51], allowing more cognitive resources to be allocated to walking, thereby enhancing physical performance without sacrificing cognitive performance; or (3) a mix of both. As we see no deterioration in cognitive performance from the available data, the improved dual-task walking speed would not be the result of a pure reallocation of existing cognitive resources, which would consequently reduce the resources available for concurrent cognitive tasks and potentially diminish cognitive performance [52]. More data are required to confirm our postulation.

XSL•FO
RenderX

Effectiveness of Tasks With a Low Cognitive Load

Our results suggest that cognitive training does not improve physical function directly, as seen in the lack of significant changes in outcomes for tasks that require a low cognitive load. Although our overall meta-analysis found a significant improvement in single-task gait speed (P=.05), the analysis of studies with high methodological quality did not show a significant improvement. This discrepancy may be attributed to a few factors: the stricter methodological standards of high-quality studies, which likely reduced the potential for overestimated effects, and the lower heterogeneity (I^2=0%) in the analysis of high-quality studies, suggesting more consistent and reliable findings. Meanwhile, studies in the high-quality group generally included larger sample sizes than those of lower quality (506 participants in the high-quality group vs 130 participants in the low-quality group). Notably, 1 large RCT (n=238 participants) reported no significant between-group improvements in single-task gait speed [19]. In contrast, the only study that reported a significant improvement in single-task gait speed had a small sample size, with only 10 participants in each group [35]. Additionally, this study included only male participants and focused on individuals with potential balance impairments, as indicated by a BBS score <52 and/or a self-selected gait speed of ≤1.1 m/s on the 10-meter walk test. These findings suggest that cognitive training may not significantly improve single-task gait speed in relatively healthy older adults.

Nevertheless, our review identified consistent results across outcomes involving low cognitive load, including functional mobility and balance measured using the TUG test, BBS, and Mini-BESTest, with no significant changes observed. We hypothesize that older adults, in general, have sufficient cognitive resources available for less cognitively challenging tasks. Consequently, the central capacity sharing model does not apply [50], which might explain the nonsignificant effect of cognitive training on tasks with a low cognitive load.

Conceptual and Mechanistic Interpretation

The pattern of benefit observed in this review can be interpreted within contemporary motor control theory. Healthy adult walking is partly automatic, meaning that steady-state gait can be regulated with limited executive attention, whereas aging and dual-task demands shift gait toward more deliberate, executive control [53,54]. From this perspective, pure cognitive training would not be expected to substantially alter well-practiced single-task walking or simple balance tasks in cognitively intact older adults, as these tasks may already be supported by sufficient automatic sensorimotor control. Its effect is more likely to appear during cognitive-motor dual-task walking, where cognitive resources are invoked, and training-related improvements in attention, inhibition, working memory, or task switching may reduce competition between cognitive and locomotor processes. This also aligns with the near- vs far-transfer framework [55,56]: improvement in dual-task gait speed reflects near transfer, given its shared executive-control demands with cognitive training, while the lack of effects on single-task gait speed, mobility, and balance is consistent with limited far transfer to outcomes with lower cognitive load.

Neurophysiologically, this interpretation is consistent with evidence that dual-task walking engages frontal cortical systems. Functional near-infrared spectroscopy (fNIRS) studies show that walking while talking in older adults is associated with prefrontal cortical activation, supporting the view that attentional and executive-control networks contribute to gait when task demands increase [57]. Systematic neurophysiological evidence also indicates that cognitive-motor interference involves distributed cortical resource allocation rather than a single gait-specific locus [58]. Therefore, the significant improvement in dual-task gait speed found in this review may reflect more efficient frontal compensation or cognitive resource allocation during cognitively demanding walking, although the included trials were not designed to directly test these neural mechanisms. Future trials should combine gait outcomes with neurophysiological measures, such as fNIRS or electroencephalogram (EEG), to determine whether behavioral gains correspond to reduced or more targeted prefrontal recruitment during dual-task gait.

Effective Training Protocol

As the meta-analysis revealed that pure cognitive training significantly improves only cognitive-motor dual-task gait speed, it is more pertinent to focus on the training protocols that could lead to an improvement in cognitive-motor dual-task gait speed.

Among the studies analyzed for dual-task gait speed with high methodological quality, all used CCT aimed at improving executive function. The training sessions lasted for 40 to 60 minutes, with 2 to 3 sessions per week, for a range of 8 to 10 weeks. After further analysis, we found no difference in training protocols between studies that reported significant improvements [43,44] and those that reported nonsignificant improvements [19,38,40].

Clinical Significance

This review demonstrated that pure cognitive training significantly improves cognitive-motor dual-task walking speed. Nevertheless, this review provides no direct evidence on whether the small-to-moderate effect size found on dual-task walking speed is clinically important. To contextualize the improvement in dual-task gait speed, a sensitivity analysis, including only studies using the same unit of measurement, was conducted to obtain the MD for comparison with the minimal detectable change (MDC) or minimal clinically important difference (MCID).

Pure cognitive training increased dual-task gait speed by 8.83 cm/s (MD 8.83, 95% CI 0.32-17.34; experimental n=157 participants: control n=152 participants; I^2=44%; P=.04; 3 trials of good methodological quality). This improvement falls within the ranges of MDC (5-10 cm/s) [59,60] and MCID (8-20 cm/s) [61,62] reported for various older populations. Therefore, a firm quantitative conclusion about its clinical significance cannot be drawn. Nevertheless, this improvement might be valuable as cognitive-motor dual-task walking speed is a critical predictor of falls, frailty, disability, and mortality [63].

XSL•FO
RenderX

In particular, the improvement in cognitive-motor dual-task gait speed may be relevant to fall prevention because walking in real-world situations often involves distractions, such as texting while walking [64]. Cognitive-motor dual-task gait speed reflects the ability to maintain walking performance under divided-attention conditions that commonly occur in daily life. A systematic review and meta-analysis reported that slower gait speed under both single- and dual-task conditions can differentiate fallers from nonfallers among older adults, supporting dual-task gait speed as a marker of fall risk [65]. More specifically, in a prospective cohort of 646 community-dwelling older adults, poorer walking-while-talking speed predicted incident falls after adjustment for normal walking velocity and other confounders. Conversely, participants in the fastest 14% of the walking-while-talking speed distribution had a significantly lower risk of incident falls than the remaining 86% of the cohort [66]. Therefore, the gain in cognitive-motor dual-task gait speed may indicate a greater capacity to walk safely while attention is divided between concurrent tasks.

However, pure cognitive training did not appear to improve physical performance directly in older adults. This could be attributed to the low cognitive requirement of simple balance and general mobility tasks among healthier older adults. It should be noted that the cognitive load of the same task could be different in older adults with different abilities. We observed a trend that studies that included older adults with poorer physical function [20,35,40,45], or included more difficult balance tasks (ie, single-leg standing with eyes closed) [20], tended to report significant improvements. Results might differ if more challenging balance tasks were involved or if participants were older adults with known physical or cognitive impairments.

Limitations

For the 5 trials that included balance outcomes, only 3 types of outcomes (BBS, Mini-BESTest, and OLST) were used. The insufficient diversity in balance outcomes prevents a comprehensive analysis of balance performances, thereby restricting the ability to evaluate the relationship between cognitive training and balance.

This meta-analysis primarily used gait speed and cognitive-motor dual-task gait speed as indicators of gait performance. However, gait performance should not be determined solely by gait velocity. Other gait parameters, such as spatiotemporal measures, can also be an important determining factor of gait performance. Yet, the effects of cognitive training on the spatiotemporal aspects of gait were not examined because of the limited number of studies available. Thus, future investigations are needed to explore the relationship between cognitive training and various spatiotemporal gait characteristics under single-task conditions among healthy older adults.

As this study included only cognitively intact older adults without any diseases affecting gait, the findings may not be generalizable to those with cognitive impairments or gait-related health conditions. Additionally, the cognitive training in most included studies focused primarily on enhancing executive function without targeting other cognitive domains.

Furthermore, for the meta-analysis with high heterogeneity (functional mobility), metaregression analysis was not performed as the meta-analysis included only 6 studies, falling short of the recommended minimum of 10 studies required for a metaregression model [32].

Conclusions

Cognitive training, aimed at improving executive function and lasting for 40 to 60 minutes per session, 2 to 3 sessions per week, over 8 to 10 weeks, improves cognitive-motor dual-task walking speed in older adults. However, cognitive training could not improve single-task mobility or balance performance in cognitively intact older adults.

Acknowledgments

The authors declare the use of generative AI (GAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing the manuscript and adapting and adjusting the emotional tone. The GAI tools used were ChatGPT (GPT-5.5; OpenAI) and Gemini (version 3; Alphabet Inc). Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit organizations for the research, authorship, or publication of this article.

Data Availability

Data are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: FMHL (lead), CLH (supporting), WLSC (supporting)

XSL•FO
RenderX

Formal analysis: XL (equal), YLC (equal), GL (equal), CH Lau (equal), TYYL (equal), CH Lam (equal), TWC (equal), HKC (equal)
Methodology: XL (equal), YLC (equal), GL (equal), CH Lau (equal), TYYL (equal), CH Lam (equal), TWC (equal), HKC (equal)
Validation: FMHL (lead), CLH (equal), WLSC (equal)
Writing—original draft: XL (equal), YLC (equal), GL (equal), CH Lau (equal), TYYL (equal), CH Lam (equal), TWC (equal), HKC (equal)
Writing—review and editing: FMHL (lead), CLH (equal), WLSC (equal)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details of search strategies.
[DOCX File , 24 KB - aging_v9i1e85070_app1.docx]

Multimedia Appendix 2
Downgrade and upgrade criteria of Grading of Recommendations Assessment, Development, and Evaluation (GRADE).
[DOCX File , 20 KB - aging_v9i1e85070_app2.docx]

Multimedia Appendix 3
Summary of participant characteristics and intervention protocols.
[DOCX File , 69 KB - aging_v9i1e85070_app3.docx]

Multimedia Appendix 4
Summary of intervention effects.
[DOCX File , 50 KB - aging_v9i1e85070_app4.docx]

Multimedia Appendix 5
Physiotherapy Evidence Database (PEDro) score of included articles.
[DOCX File , 42 KB - aging_v9i1e85070_app5.docx]

Multimedia Appendix 6
Meta-analysis comparing functional mobility between the intervention group and the control group.
[PNG File , 149 KB - aging_v9i1e85070_app6.png]

Multimedia Appendix 7
Subgroup meta-analyses for functional mobility to identify the source of heterogeneity.
[DOCX File , 113 KB - aging_v9i1e85070_app7.docx]

Multimedia Appendix 8
Meta-analysis comparing balance scales between the intervention group and the control group.
[PNG File , 31 KB - aging_v9i1e85070_app8.png]

Multimedia Appendix 9
PRISMA 2020 checklist.
[DOCX File , 33 KB - aging_v9i1e85070_app9.docx]

References

1. Divandari N, Bird ML, Vakili M, Jaberzadeh S. The association between cognitive domains and postural balance among healthy older adults: a systematic review of literature and meta-analysis. Curr Neurol Neurosci Rep 2023 Nov;23(11):681-693 [FREE Full text] [doi: 10.1007/s11910-023-01305-y] [Medline: 37856048]
2. Holtzer R, Wang C, Verghese J. The relationship between attention and gait in aging: facts and fallacies. Motor Control 2012 Jan;16(1):64-80 [FREE Full text] [doi: 10.1123/mcj.16.1.64] [Medline: 22402221]
3. Mc Ardle R, Taylor L, Cavadino A, Rochester L, Del Din S, Kerse N. Characterizing walking behaviors in aged residential care using accelerometry, with comparison across care levels, cognitive status, and physical function: cross-sectional study. JMIR Aging 2024 Jun 04;7:e53020 [FREE Full text] [doi: 10.2196/53020] [Medline: 38842168]

XSL•FO
RenderX

4. Yi D, Jang S, Yim J. Relationship between associated neuropsychological factors and fall risk factors in community-dwelling elderly. Healthcare (Basel) 2022 Apr 14;10(4):728 [FREE Full text] [doi: 10.3390/healthcare10040728] [Medline: 35455905]
5. Mori E, Aoyagi Y, Kono Y, Asai H, Tomita H, Izawa H. Exploring the factors associated with decreased dynamic balance ability in older patients with heart failure. Heart Lung 2023;58:139-143. [doi: 10.1016/j.hrtlng.2022.11.016] [Medline: 36512879]
6. Bahureksa L, Najafi B, Saleh A, Sabbagh M, Coon D, Mohler MJ, et al. The impact of mild cognitive impairment on gait and balance: a systematic review and meta-analysis of studies using instrumented assessment. Gerontology 2017;63(1):67-83 [FREE Full text] [doi: 10.1159/000445831] [Medline: 27172932]
7. Law CK, Lam FM, Chung RC, Pang MY. Physical exercise attenuates cognitive decline and reduces behavioural problems in people with mild cognitive impairment and dementia: a systematic review. J Physiother 2020 Jan;66(1):9-18 [FREE Full text] [doi: 10.1016/j.jphys.2019.11.014] [Medline: 31843427]
8. Xu L, Gu H, Cai X, Zhang Y, Hou X, Yu J, et al. The effects of exercise for cognitive function in older adults: a systematic review and meta-analysis of randomized controlled trials. Int J Environ Res Public Health 2023 Jan 07;20(2):1088 [FREE Full text] [doi: 10.3390/ijerph20021088] [Medline: 36673844]
9. Brown BM, Peiffer JJ, Taddei K, Lui JK, Laws SM, Gupta VB, et al. Physical activity and amyloid-β plasma and brain levels: results from the Australian Imaging, Biomarkers and Lifestyle Study of Ageing. Mol Psychiatry 2013 Aug;18(8):875-881. [doi: 10.1038/mp.2012.107] [Medline: 22889922]
10. Zhao R. Exercise mimetics: a novel strategy to combat neuroinflammation and Alzheimer's disease. J Neuroinflammation 2024 Feb 02;21(1):40 [FREE Full text] [doi: 10.1186/s12974-024-03031-9] [Medline: 38308368]
11. Gomez-Pinilla F, Zhuang Y, Feng J, Ying Z, Fan G. Exercise impacts brain-derived neurotrophic factor plasticity by engaging mechanisms of epigenetic regulation. Eur J Neurosci 2011 Feb;33(3):383-390 [FREE Full text] [doi: 10.1111/j.1460-9568.2010.07508.x] [Medline: 21198979]
12. Ben-Zeev T, Shoenfeld Y, Hoffman JR. The effect of exercise on neurogenesis in the brain. Isr Med Assoc J 2022 Aug;24(8):533-538 [FREE Full text] [Medline: 35971998]
13. Sheridan PL, Hausdorff JM. The role of higher-level cognitive function in gait: executive dysfunction contributes to fall risk in Alzheimer's disease. Dement Geriatr Cogn Disord 2007;24(2):125-137 [FREE Full text] [doi: 10.1159/000105126] [Medline: 17622760]
14. Coelho FG, Stella F, de Andrade LP, Barbieri FA, Santos-Galduróz RF, Gobbi S, et al. Gait and risk of falls associated with frontal cognitive functions at different stages of Alzheimer's disease. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2012 Sep;19(5):644-656. [doi: 10.1080/13825585.2012.661398] [Medline: 22360785]
15. Pieruccini-Faria F, Lord SR, Toson B, Kemmler W, Schoene D. Mental flexibility influences the association between poor balance and falls in older people - a secondary analysis. Front Aging Neurosci 2019 Jun 12;11:133 [FREE Full text] [doi: 10.3389/fnagi.2019.00133] [Medline: 31263408]
16. Gavelin HM, Dong C, Minkov R, Bahar-Fuchs A, Ellis KA, Lautenschlager NT, et al. Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. Ageing Res Rev 2021 Mar;66:101232. [doi: 10.1016/j.arr.2020.101232] [Medline: 33249177]
17. Marusic U, Verghese J, Mahoney JR. Cognitive-based interventions to improve mobility: a systematic review and meta-analysis. J Am Med Dir Assoc 2018 Jun;19(6):484-91.e3 [FREE Full text] [doi: 10.1016/j.jamda.2018.02.002] [Medline: 29680203]
18. Beauchet O, Allali G, Montero-Odasso M, Sejdić E, Fantino B, Annweiler C. Motor phenotype of decline in cognitive performance among community-dwellers without dementia: population-based study and meta-analysis. PLoS One 2014 Jun 09;9(6):e99318 [FREE Full text] [doi: 10.1371/journal.pone.0099318] [Medline: 24911155]
19. Sturnieks DL, Hicks C, Smith N, Ratanapongleka M, Menant J, Turner J, et al. Exergame and cognitive training for preventing falls in community-dwelling older people: a randomized controlled trial. Nat Med 2024 Jan;30(1):98-105. [doi: 10.1038/s41591-023-02739-0] [Medline: 38228913]
20. Liu Y, Zheng M, Liu J, Cao C, Ding Y. Effects of cognitive training in flanker task on mobility of older adults: a randomized controlled trial. Gerontology 2023;69(5):593-602. [doi: 10.1159/000529088] [Medline: 36646052]
21. Marusic U, Kavcic V, Giordani B, Gerževič M, Meeusen R, Pišot R. Computerized spatial navigation training during 14 days of bed rest in healthy older adult men: effect on gait performance. Psychol Aging 2015 Jun;30(2):334-340. [doi: 10.1037/pag0000021] [Medline: 25938245]
22. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
23. Lam F, Chan YL, Lau CH, Lam TY, Lam CH, Chan TW, et al. The effects of cognitive training on gait and balance of healthy elderly: systematic review and meta-analysis. PROSPERO. 2024. URL: https://www.crd.york.ac.uk/PROSPERO/view/CRD42023401100 [accessed 2026-07-08]
24. Marusic U, Grosprêtre S. Non-physical approaches to counteract age-related functional deterioration: applications for rehabilitation and neural mechanisms. Eur J Sport Sci 2018 Jun;18(5):639-649 [FREE Full text] [doi: 10.1080/17461391.2018.1447018] [Medline: 29557276]

XSL•FO
RenderX

25. Gobet F, Sala G. Cognitive training: a field in search of a phenomenon. Perspect Psychol Sci 2023 Jan;18(1):125-141 [FREE Full text] [doi: 10.1177/17456916221091830] [Medline: 35939827]
26. Cashin AG, McAuley JH. Clinimetrics: Physiotherapy Evidence Database (PEDro) scale. J Physiother 2020 Jan;66(1):59 [FREE Full text] [doi: 10.1016/j.jphys.2019.08.005] [Medline: 31521549]
27. Goldet G, Howick J. Understanding GRADE: an introduction. J Evid Based Med 2013 Feb;6(1):50-54. [doi: 10.1111/jebm.12018] [Medline: 23557528]
28. Liang SG, Chow JC, Leung NM, Mo YN, Ng TM, Woo CL, et al. The effects of ankle and foot exercises on ankle strength, balance, and falls in older people: a systematic review and meta-analysis. Phys Ther 2025 Jan 08;105(1):pzae157 [FREE Full text] [doi: 10.1093/ptj/pzae157] [Medline: 39657220]
29. Liang SG, Fan ES, Lam PK, Kwok WT, Ma CZ, Lam FM. The effect of adding real-time postural feedback in balance and mobility training in older adults: a systematic review and meta-analysis. Arch Gerontol Geriatr 2024 Aug;123:105439 [FREE Full text] [doi: 10.1016/j.archger.2024.105439] [Medline: 38643641]
30. Liang G, Ching HY, Fu CM, Lau SM, Chan WL, Huang M, et al. Does adding real-time postural feedback during training boost balance and mobility in individuals with Parkinson's disease? A systematic review and meta-analysis. J Appl Gerontol 2026 Jan 29:7334648261419637 [FREE Full text] [doi: 10.1177/07334648261419637] [Medline: 41609671]
31. Spineli LM, Pandis N. Problems and pitfalls in subgroup analysis and meta-regression. Am J Orthod Dentofacial Orthop 2020 Dec;158(6):901-904. [doi: 10.1016/j.ajodo.2020.09.001] [Medline: 33250104]
32. Geissbühler M, Hincapié CA, Aghlmandi S, Zwahlen M, Jüni P, da Costa BR. Most published meta-regression analyses based on aggregate data suffer from methodological pitfalls: a meta-epidemiological study. BMC Med Res Methodol 2021 Jun 15;21(1):123 [FREE Full text] [doi: 10.1186/s12874-021-01310-0] [Medline: 34130658]
33. Dalton JE, Bolen SD, Mascha EJ. Publication bias: the elephant in the review. Anesth Analg 2016 Oct;123(4):812-813 [FREE Full text] [doi: 10.1213/ANE.0000000000001596] [Medline: 27636569]
34. Althomali MM, Vallis LA, Leat SJ. Can older adults' balance and mobility improve with visual attention training? Eur J Appl Physiol 2019 Jul;119(7):1649-1661. [doi: 10.1007/s00421-019-04153-2] [Medline: 31055677]
35. Azadian E, Torbati HR, Kakhki AR, Farahpour N. The effect of dual task and executive training on pattern of gait in older adults with balance impairment: a randomized controlled trial. Arch Gerontol Geriatr 2016;62:83-89. [doi: 10.1016/j.archger.2015.10.001] [Medline: 26508098]
36. Blackwood J, Shubert T, Fogarty K, Chase C. The impact of a home-based computerized cognitive training intervention on fall risk measure performance in community dwelling older adults, a pilot study. J Nutr Health Aging 2016 Feb;20(2):138-145 [FREE Full text] [doi: 10.1007/s12603-015-0598-5] [Medline: 26812509]
37. Hou HY, Li HJ. Effects of exergame and video game training on cognitive and physical function in older adults: a randomized controlled trial. Appl Ergon 2022 May;101:103690. [doi: 10.1016/j.apergo.2022.103690] [Medline: 35066398]
38. Li KZ, Roudaia E, Lussier M, Bherer L, Leroux A, McKinley PA. Benefits of cognitive dual-task training on balance performance in healthy older adults. J Gerontol A Biol Sci Med Sci 2010 Dec;65(12):1344-1352. [doi: 10.1093/gerona/glq151] [Medline: 20837662]
39. Ng TP, Feng L, Nyunt MS, Feng L, Niti M, Tan BY, et al. Nutritional, physical, cognitive, and combination interventions and frailty reversal among older adults: a randomized controlled trial. Am J Med 2015 Nov;128(11):1225-36.e1 [FREE Full text] [doi: 10.1016/j.amjmed.2015.06.017] [Medline: 26159634]
40. Smith-Ray RL, Hughes SL, Prohaska TR, Little DM, Jurivich DA, Hedeker D. Impact of cognitive training on balance and gait in older adults. J Gerontol B Psychol Sci Soc Sci 2015 May;70(3):357-366 [FREE Full text] [doi: 10.1093/geronb/gbt097] [Medline: 24192586]
41. Smith-Ray RL, Makowski-Woidan B, Hughes SL. A randomized trial to measure the impact of a community-based cognitive training intervention on balance and gait in cognitively intact Black older adults. Health Educ Behav 2014 Oct;41(1 Suppl):62S-69S [FREE Full text] [doi: 10.1177/1090198114537068] [Medline: 25274713]
42. Steinmetz JP, Federspiel C. The effects of cognitive training on gait speed and stride variability in old adults: findings from a pilot study. Aging Clin Exp Res 2014 Dec;26(6):635-643 [FREE Full text] [doi: 10.1007/s40520-014-0228-9] [Medline: 24803285]
43. Verghese J, Mahoney J, Ambrose AF, Wang C, Holtzer R. Effect of cognitive remediation on gait in sedentary seniors. J Gerontol A Biol Sci Med Sci 2010 Dec;65(12):1338-1343. [doi: 10.1093/gerona/glq127] [Medline: 20643703]
44. Marusic U, Verghese J, Mahoney JR. Does cognitive training improve mobility, enhance cognition, and promote neural activation? Front Aging Neurosci 2022 May 23;14:845825 [FREE Full text] [doi: 10.3389/fnagi.2022.845825] [Medline: 35677205]
45. Ammar M, Amjad I, Nisar M, Ghoneim OS. Effects of board games on balance in association with cognition in community-dwelling older adults. J Musculoskelet Surg Res 2024 Aug 02;8(3):256-263. [doi: 10.25259/JMSR_122_2024]
46. Azadian E, Majlesi M, Jafarnezhadgero AA. The effect of working memory intervention on the gait patterns of the elderly. J Bodyw Mov Ther 2018 Oct;22(4):881-887. [doi: 10.1016/j.jbmt.2017.08.008] [Medline: 30368330]
47. Verghese J, Mahoney JR, Ayers E, Ambrose A, Wang C, Holtzer R. Computerised cognitive remediation to enhance mobility in older adults: a single-blind, single-centre, randomised trial. Lancet Healthy Longev 2021 Sep;2(9):e571-e579 [FREE Full text] [doi: 10.1016/s2666-7568(21)00173-2] [Medline: 34522910]

XSL•FO
RenderX

48. Esmaeili Bijarsari S. A current view on dual-task paradigms and their limitations to capture cognitive load. Front Psychol 2021 May 20;12:648586 [FREE Full text] [doi: 10.3389/fpsyg.2021.648586] [Medline: 34093335]
49. Ma D, Izzetoglu M, Holtzer R, Jiao X. Deep learning based walking tasks classification in older adults using fNIRS. IEEE Trans Neural Syst Rehabil Eng 2023;31:3437-3447 [FREE Full text] [doi: 10.1109/TNSRE.2023.3306365] [Medline: 37594868]
50. Tombu M, Jolicoeur P. A central capacity sharing model of dual-task performance. J Exp Psychol Hum Percept Perform 2003 Feb;29(1):3-18. [doi: 10.1037//0096-1523.29.1.3] [Medline: 12669744]
51. Haith AM, Krakauer JW. The multiple effects of practice: skill, habit and reduced cognitive load. Curr Opin Behav Sci 2018 Apr;20:196-201 [FREE Full text] [doi: 10.1016/j.cobeha.2018.01.015] [Medline: 30944847]
52. Riedel N, Herzog M, Stein T, Deml B. Cognitive-motor interference during walking with modified leg mechanics: a dual-task walking study. Front Psychol 2024 Apr 18;15:1375029 [FREE Full text] [doi: 10.3389/fpsyg.2024.1375029] [Medline: 38699569]
53. Clark DJ. Automaticity of walking: functional significance, mechanisms, measurement and rehabilitation strategies. Front Hum Neurosci 2015 May 05;9:246 [FREE Full text] [doi: 10.3389/fnhum.2015.00246] [Medline: 25999838]
54. Yogev-Seligmann G, Hausdorff JM, Giladi N. The role of executive function and attention in gait. Mov Disord 2008 Feb 15;23(3):329-42; quiz 472 [FREE Full text] [doi: 10.1002/mds.21720] [Medline: 18058946]
55. Barnett SM, Ceci SJ. When and where do we apply what we learn? A taxonomy for far transfer. Psychol Bull 2002 Jul;128(4):612-637. [doi: 10.1037/0033-2909.128.4.612] [Medline: 12081085]
56. Simons DJ, Boot WR, Charness N, Gathercole SE, Chabris CF, Hambrick DZ, et al. Do "brain-training" programs work? Psychol Sci Public Interest 2016 Oct;17(3):103-186. [doi: 10.1177/1529100616661983] [Medline: 27697851]
57. Holtzer R, Mahoney JR, Izzetoglu M, Izzetoglu K, Onaral B, Verghese J. fNIRS study of walking and walking while talking in young and old individuals. J Gerontol A Biol Sci Med Sci 2011 Aug;66(8):879-887 [FREE Full text] [doi: 10.1093/gerona/glr068] [Medline: 21593013]
58. Leone C, Feys P, Moumdjian L, D'Amico E, Zappia M, Patti F. Cognitive-motor dual-task interference: a systematic review of neural correlates. Neurosci Biobehav Rev 2017 Apr;75:348-360. [doi: 10.1016/j.neubiorev.2017.01.010] [Medline: 28104413]
59. Criss MG, Chui KK, Gallichio J, Centra J, Canbek J. Reliability, responsiveness, and validity of slow walking speed in community dwelling older adults. Gait Posture 2023 Jan;99:54-59. [doi: 10.1016/j.gaitpost.2022.10.016] [Medline: 36327539]
60. Goldberg A, Schepens S. Measurement error and minimum detectable change in 4-meter gait speed in older adults. Aging Clin Exp Res 2011;23(5-6):406-412. [doi: 10.1007/BF03325236] [Medline: 22526072]
61. Bohannon RW, Glenney SS. Minimal clinically important difference for change in comfortable gait speed of adults with pathology: a systematic review. J Eval Clin Pract 2014 Aug;20(4):295-300. [doi: 10.1111/jep.12158] [Medline: 24798823]
62. Thorman IB, Loyd BJ, Clendaniel RA, Dibble LE, Schubert MC. The minimal clinically important difference for gait speed in significant unilateral vestibular hypofunction after vestibular rehabilitation. J Otol 2023 Jan;18(1):15-20 [FREE Full text] [doi: 10.1016/j.joto.2022.11.001] [Medline: 36820158]
63. Verghese J, Holtzer R, Lipton RB, Wang C. Mobility stress test approach to predicting frailty, disability, and mortality in high-functioning older adults. J Am Geriatr Soc 2012 Oct;60(10):1901-1905 [FREE Full text] [doi: 10.1111/j.1532-5415.2012.04145.x] [Medline: 23002714]
64. Bruyneel AV, Reinmann A, Gafner SC, Sandoz JD, Duclos NC. Does texting while walking affect spatiotemporal gait parameters in healthy adults, older people, and persons with motor or cognitive disorders? A systematic review and meta-analysis. Gait Posture 2023 Feb;100:284-301 [FREE Full text] [doi: 10.1016/j.gaitpost.2023.01.009] [Medline: 36696854]
65. Menant JC, Schoene D, Sarofim M, Lord SR. Single and dual task tests of gait speed are equivalent in the prediction of falls in older people: a systematic review and meta-analysis. Ageing Res Rev 2014 Jul;16:83-104. [doi: 10.1016/j.arr.2014.06.001] [Medline: 24915643]
66. Ayers EI, Tow AC, Holtzer R, Verghese J. Walking while talking and falls in aging. Gerontology 2014;60(2):108-113 [FREE Full text] [doi: 10.1159/000355119] [Medline: 24192342]

Abbreviations

BBS: Berg Balance Scale
BDNF: brain-derived neurotrophic factor
CCT: computerized cognitive training
EEG: electroencephalogram
fNIRS: functional near-infrared spectroscopy
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
IGF-1: insulin-like growth factor-1
MCI: mild cognitive impairment
MCID: minimal clinically important difference

XSL•FO
RenderX

MD: mean difference
MDC: minimal detectable change
Mini-BESTest: Mini-Balance Evaluation Systems Test
MMSE: Mini-Mental State Examination
OLST: one-leg stance test
PEDro: Physiotherapy Evidence Database
PICOS: population, intervention, comparison, outcome, and study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SMD: standardized mean difference
TUG: Timed Up and Go

Edited by P Jansons; submitted 03.Oct.2025; peer-reviewed by D Vance, U Marusic; comments to author 29.Mar.2026; revised version received 26.Jun.2026; accepted 26.Jun.2026; published 29.Jul.2026.

Please cite as:
Li X, Chan YL, Liang G, Lau CH, Lam TYY, Lam CH, Chan TW, Cheung HK, Chan WLS, Hsu CL, Lam FMH
Effects of Pure Cognitive Training on Gait and Balance in Older Adults: Systematic Review and Meta-Analysis
JMIR Aging 2026;9:e85070
URL: https://aging.jmir.org/2026/1/e85070
doi:10.2196/85070
PMID:

©Xiaohan Li, Yat Leung Chan, Guoshi Liang, Cheuk Hin Lau, Tsz Yau Yuki Lam, Chak Hang Lam, Tsz Wing Chan, Hoi Kei Cheung, Wayne Lap Sun Chan, Chun Liang Hsu, Freddy Man Hin Lam. Originally published in JMIR Aging (https://aging.jmir.org), 29.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

AI-Based Monitoring Devices for the Identification of Circadian Parameters as Digital Predictive Biomarkers of Neurodegenerative Diseases in Older Adults: Systematic Review

J Manuel Herrera-Zamora[1*], PhD; Jesus Monroy-Rodriguez[2*], PhD; Jesus Garcia-Mancilla[3*], MSc; Pedro C Santana-Mancilla[4*], Prof Dr

[1]University Center for Biomedical Research, Facultad de Medicina, Universidad de Colima, Colima, Colima, Mexico

[2]Department of Basic Health Sciences, Centro Universitario del Sur, Universidad de Guadalajara, Ciudad Guzmán, Jalisco, Mexico

[3]Engineering, Sports Illustrated Tickets, Houston, TX, United States

[4]School of Telematics, Universidad de Colima, Colima, Colima, Mexico

[*]all authors contributed equally

Corresponding Author:
Pedro C Santana-Mancilla, Prof Dr
School of Telematics
Universidad de Colima
Av. Universidad 333
Colima, Colima, 28040
Mexico
Phone: 52 3123161075
Email: psantana@ucol.mx

Abstract

Background: Neurodegenerative diseases such as Alzheimer disease, Parkinson disease, and related dementias are a growing health burden among older adults. Circadian disruption is increasingly recognized as both a consequence and a potential early marker of neurodegenerative progression. Wearable sensing combined with AI enables continuous, noninvasive, and ecologically valid monitoring of circadian parameters as potential digital biomarkers.

Objective: This systematic review synthesizes evidence on AI-enhanced wearable and objective sensing devices used to monitor circadian parameters in older and cognitively vulnerable adults and evaluates their potential as digital predictive biomarkers for neurodegenerative diseases.

Methods: Following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines, an original search of PubMed, Scopus, and JMIR journals (June 2025) yielded 12 studies, 7 of which met updated eligibility criteria on rescreening. A complementary search of PubMed, IEEE Xplore, and the ACM Digital Library (March 24, 2026) added 3 eligible studies and 3 sensitivity references (excluded records retained to delimit the corpus boundary). Two reviewers (JMH-Z and JM-R) independently assessed risk of bias with the Mixed Methods Appraisal Tool. Heterogeneous end points precluded meta-analysis; the narrative synthesis covered 4 prespecified domains: device class, circadian construct, AI and machine learning (ML) method family, and population and neurocognitive context.

Results: The final corpus comprised 13 records (10 eligible studies and 3 sensitivity references) published from 2015 to 2025. Wrist actigraphy predominated, and samples ranged from community-dwelling older adults to nursing home residents with advanced dementia (n=16-7297). Circadian parameters included interdaily stability, intradaily variability, relative amplitude, midline estimating statistic of rhythm, acrophase, and circadian stimulus. AI/ML methods included supervised learning (random forests, gradient boosting, and regression models), unsupervised clustering, and algorithm development or validation; 7 studies applied deterministic algorithms without fitting a predictive model. Circadian fragmentation was associated with functional decline, dementia risk, and cognitive performance. Most studies were descriptive or observational with small, nondiverse samples; no AI model was externally validated, and no study followed an AI-specific reporting guideline.

Conclusions: AI-enhanced wearable devices show meaningful potential for noninvasive measurement of circadian disruption in older and cognitively vulnerable populations. Evidence certainty is moderate for actigraphy-derived fragmentation-function associations in institutional settings and low for predictive biomarker claims in community populations. From 6 recurring challenges

XSL•FO
RenderX

along the sensing-to-biomarker pipeline, we propose a 4-layer reference architecture (sensing and acquisition, circadian parameter extraction, AI/ML inference, and a clinical and validation interface) with a layered validation pathway. Longitudinal, multisite validation studies using standardized circadian metrics, diverse populations, and explainable AI remain necessary before clinical deployment as neurodegenerative biomarkers.

(JMIR Aging 2026;9:e83031) doi:10.2196/83031

KEYWORDS

circadian rhythms; wearable devices; AI; neurodegenerative diseases; older adults

Introduction

Background

Neurodegenerative diseases are a heterogeneous group of disorders characterized by the progressive loss of neuronal structure and function, significantly affecting memory, cognition, behavior, and sensorimotor function. These conditions, which include Alzheimer disease (AD), Parkinson disease (PD), and Lewy body dementia, represent a growing global public health challenge. The global number of people living with dementia is projected to reach 153 million by 2050 [1], underscoring the urgency of developing tools for earlier detection and monitoring.

Disruptions in circadian rhythms are increasingly recognized as a hallmark feature of neurodegeneration, appearing not only as a functional consequence of advancing disease but also as a potential early biological signal that may precede clinical onset [2,3]. The circadian system functions as the body's intrinsic biological clock, coordinating cellular processes, sleep-wake cycles, hormonal secretion, and behavioral rhythms with the 24-hour light-dark environment [4]. In older adults and those with neurodegenerative conditions, this system progressively deteriorates, resulting in measurable fragmentation of rest-activity rhythms, altered sleep architecture, and disrupted melatonin secretion [5]. Recent conceptual frameworks, developed in populations other than older adults, treat sleep and circadian disruption as the joint product of circadian, neuroendocrine, and behavioral stressors rather than as a single timing signal [6]; this multidimensional view is relevant to aging, where circadian decline coexists with endocrine and behavioral change.

Recent advances in wearable technologies and AI offer new approaches to continuously and ecologically monitoring circadian parameters outside clinical or laboratory settings. Wrist-worn actigraphs, electroencephalography (EEG) headbands, multisensor wearable patches, wearable light sensors, and ambient smart-home monitoring systems can collect physiological and behavioral data from individuals in their natural environments. When combined with AI and machine learning (ML) algorithms, these data streams can be processed to extract circadian parameters, such as interdaily stability (IS), intradaily variability (IV), relative amplitude (RA), midline estimating statistic of rhythm (MESOR), and acrophase, enabling continuous characterization of individual circadian profiles [7].

Despite the promise of this technology, the evidence base remains fragmented. Existing reviews have addressed wearable sensing and AI in aging broadly, without systematically evaluating the extent to which included studies operationalize genuine circadian constructs, apply AI methods rigorously, and target populations with neurodegenerative relevance. This review addresses that gap by synthesizing evidence from studies that explicitly combine objective sensing, AI, and circadian monitoring in older or cognitively vulnerable populations, with structured attention to the specificity of the circadian construct, the AI method family used, and the clinical relevance of reported outcomes. Drawing on this synthesis, we identify 6 structural challenges at the sensing-to-biomarker interface and propose a 4-layer reference architecture with a layered validation pathway intended to guide future system design and empirical evaluation in this field.

Contributions of This Review

This systematic review makes the following scholarly contributions:

- We synthesize evidence combining AI-enhanced wearable and objective sensing devices with circadian monitoring in older and cognitively vulnerable adults, organized across 4 prespecified analytical domains: device class, circadian construct, AI/ML method family, and population/neurocognitive context.
- We identify and characterize 6 structural challenges that currently limit the translation of circadian monitoring into validated predictive biomarkers for neurodegenerative disease, situating each challenge within the pipeline stage it most directly constrains.
- Drawing on the recurrence of these 4 analytical domains across the reviewed corpus, we propose a 4-layer reference pipeline architecture together with a layered validation pathway that maps each challenge to architectural responses and implementation propositions. This proposal represents a conceptual framework derived from the reviewed evidence; empirical validation through prospective field studies or multisite testbeds is outside the scope of this review but constitutes an explicit priority for future work.
- We articulate 4 implementation propositions grounded in the evidence base, including harmonized circadian end-point reporting, explainable AI (XAI) integration at model design time, population-specific device validation, and demographic representativeness as an operational validation criterion.

XSL•FO
RenderX

Methods

Overview

This systematic review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines [8] to ensure methodological transparency and rigor. The completed PRISMA 2020 checklist is available in Multimedia Appendix 1. The review protocol was prospectively deposited in Zenodo prior to data collection, in accordance with Findable, Accessible, Interoperable, and Reusable data principles and open science standards recommended by PRISMA 2020 [1]. The Zenodo record is available at [9].

During peer review, reviewers noted that the original search (PubMed, Scopus, and JMIR journals) underrepresented the engineering, biomedical informatics, and human-computer interaction literature, where AI-based sensing work is often published. To address this, we conducted a complementary systematic search across IEEE Xplore, the ACM Digital Library, and an updated PubMed query, applying the same prespecified eligibility criteria without modification. Because this search postdated the protocol deposited in Zenodo, we treat it as a documented amendment to the protocol rather than a change to the review question or eligibility rules; the deposited record reflects only the initial search.

The aim of this systematic review was to evaluate the current use of wearable devices that monitor circadian parameters using AI and to explore their potential as digital biomarkers for predicting the onset of neurodegenerative diseases in older adults, thereby contributing to improved health outcomes in aging populations. As a second, synthesis-driven aim, we sought to translate the analytical domains and structural challenges recurring across the reviewed evidence into a 4-layer reference pipeline architecture and an accompanying layered validation pathway, intended to guide the design and empirical evaluation of clinically deployable AI-circadian biomarker systems. This second aim is interpretive: the architecture and validation pathway were derived by the authors from the synthesized evidence and are not findings of the systematic review itself.

Data Sources and Search Strategy

Original Search

A primary literature search was conducted in June 2025 across PubMed, Scopus, and JMIR journals, targeting studies published between January 2015 and June 2025. Search terms were organized around three conceptual blocks: (1) wearable and objective sensing technologies, (2) circadian parameters and 24-hour behavioral constructs, and (3) aging or neurocognitive populations. This search identified 8441 records and, after duplicate removal and screening, yielded 12 included studies.

Complementary Search

Following peer-review recommendations, a complementary systematic search was conducted on March 24, 2026, across 3 additional bibliographic databases: PubMed (to capture publications beyond the original cutoff), IEEE Xplore, and the ACM Digital Library. These databases were selected to capture engineering, biomedical informatics, and human-computer interaction literature not adequately represented in the original search. Two parallel queries were run per database: a main search and a neurocognitive sensitivity search. Full database-specific Boolean search strings, including all field tags, filters (peer-reviewed journals only, 2015 onward), and date limits, are reported in Multimedia Appendix 2 in accordance with PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension) guidelines [10]. This search identified 162 records; 28 duplicates were removed, yielding 134 records for screening. Full-text assessment was conducted on 27 records. The publication window (2015-2025) was the same as in the original search; the later search date extended database coverage, not the period of eligible publications.

Eligibility Criteria

Studies were included if they met all the following criteria:

- Original peer-reviewed journal article.
- Study population comprising adults aged ≥45 years, or individuals of any age with a documented neurodegenerative condition relevant to this review (mild cognitive impairment [MCI], AD, PD, Lewy body dementia, or related dementia syndromes).
- Use of at least 1 wearable or objective sensing device for physiological or behavioral data collection.
- Application of at least 1 AI or ML method in data analysis, pattern detection, feature extraction, or predictive modeling.
- Report of at least 1 explicit circadian parameter (IS, IV, RA, MESOR, acrophase, cosinor-derived parameter, circadian stimulus, or circadian rhythm strength) or a 24-hour behavioral construct interpreted within a circadian or chronobiological framework.

Studies were excluded if they were reviews, meta-analyses, editorials, letters, conference abstracts, or protocols; used exclusively nonhuman participants; lacked full-text access; or did not incorporate an AI or ML component in the analysis.

Sensitivity references were defined as records excluded at full-text assessment that failed exactly 1 eligibility criterion, met all the others, and addressed a methodological element directly relevant to the review question: validation of a wearable sleep-wake algorithm against a reference standard, AI-based prediction from continuous sensor data, or extraction of circadian parameters in an institutional population. Sensitivity references are not counted as eligible studies. They are identified by a footnote in all tables, reported separately in the risk-of-bias assessment, and used to delimit the methodological boundary of the corpus; no conclusion of this review rests on them alone.

Study Selection Process

Records were imported into Zotero for deduplication and screening. Title and abstract screening were conducted using Rayyan Qatar Computing Research Institute (original search) and a standardized manual protocol (complementary search). Full-text eligibility assessment was performed by 2 independent reviewers (JMH-Z and JM-R for the original search; JMH-Z and PCS-M for the complementary search). Disagreements were resolved through consensus or by a third reviewer (JG-M).

XSL•FO
RenderX

The 12 studies from the original search were rescreened against the updated eligibility criteria during revision. In total, 5 were excluded: 3 because no AI or ML method was used in the analysis [11-13], and 2 because the enrolled population did not meet the population criterion or no explicit circadian parameter was analyzed [14,15]. A total of 7 studies were retained from the original corpus.

Data Extraction and Synthesis

Data were extracted from each included study using a standardized form covering authors, year and country, study design, population characteristics (sample size, age, condition), sensing device and modality, circadian construct and parameter names, AI/ML method and family, analytic approach, and main findings. Two reviewers (JMH-Z and JM-R) extracted data independently; discrepancies were resolved by discussion. For each study, we also extracted how computational methods were reported: validation scheme, unit of data partitioning, hyperparameter reporting, availability of code and data, explainability or variable-importance reporting, and adherence to AI-specific reporting guidelines (eg, TRIPOD+AI [Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis+Artificial Intelligence] [16], CLAIM [Checklist for Artificial Intelligence in Medical Imaging] [17], DECIDE-AI [Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence] [18], and CONSORT-AI [Consolidated Standards of Reporting Trials-Artificial Intelligence] [19]). A codebook with controlled vocabularies separated absent reporting (not reported) from inapplicability (not applicable and when no predictive model was fitted). A large language model–based agentic workflow (powered by Anthropic's Claude) produced a first-pass extraction from the full texts, recording a verbatim evidence locator for every value other than not reported; 2 reviewers (JMH-Z and PCS-M) then independently verified each value against the full text and supplementary materials of each source article, resolving disagreements by discussion or by a third reviewer (JG-M). Because AI assistance in scholarly evaluation can speed up the work but raises concerns about the depth and authenticity of the resulting judgments [20], the assistant's output was treated as a draft only, and no extracted value was retained without human verification. Consistent with the eligibility criterion, which counts feature extraction as a qualifying use of AI or ML, studies applying deterministic algorithms to sensor data were eligible even when they fitted no predictive model; this distinction is reported for every study in the Results section.

Narrative synthesis was organized around four prespecified analytical domains: (1) device class and sensing modality, (2) circadian construct and parametric specificity, (3) AI/ML method family, and (4) population and neurocognitive context. Studies were first cataloged descriptively within each domain, then analyzed comparatively to identify convergent evidence, methodological gaps, and key contributions. Three sensitivity-reference studies [21-23] were used to interpret the methodological boundary of the corpus and are indicated by a footnote in all tables.

Risk of Bias Assessment

Risk of bias was assessed using the Mixed Methods Appraisal Tool (MMAT) [24], which provides design-specific quality criteria for quantitative descriptive (type 4), quantitative nonrandomized (type 3), and quantitative randomized controlled trial (RCT; type 2) designs. Two reviewers (JMH-Z and JM-R) independently assessed each included study, applying the MMAT screening questions (S1 and S2) and 5 design-specific quality criteria (Q1-Q5). Disagreements were resolved by consensus. Domain-level judgments (Yes/No/Cannot tell) and overall appraisal (low, moderate, or high risk) are reported in the next section.

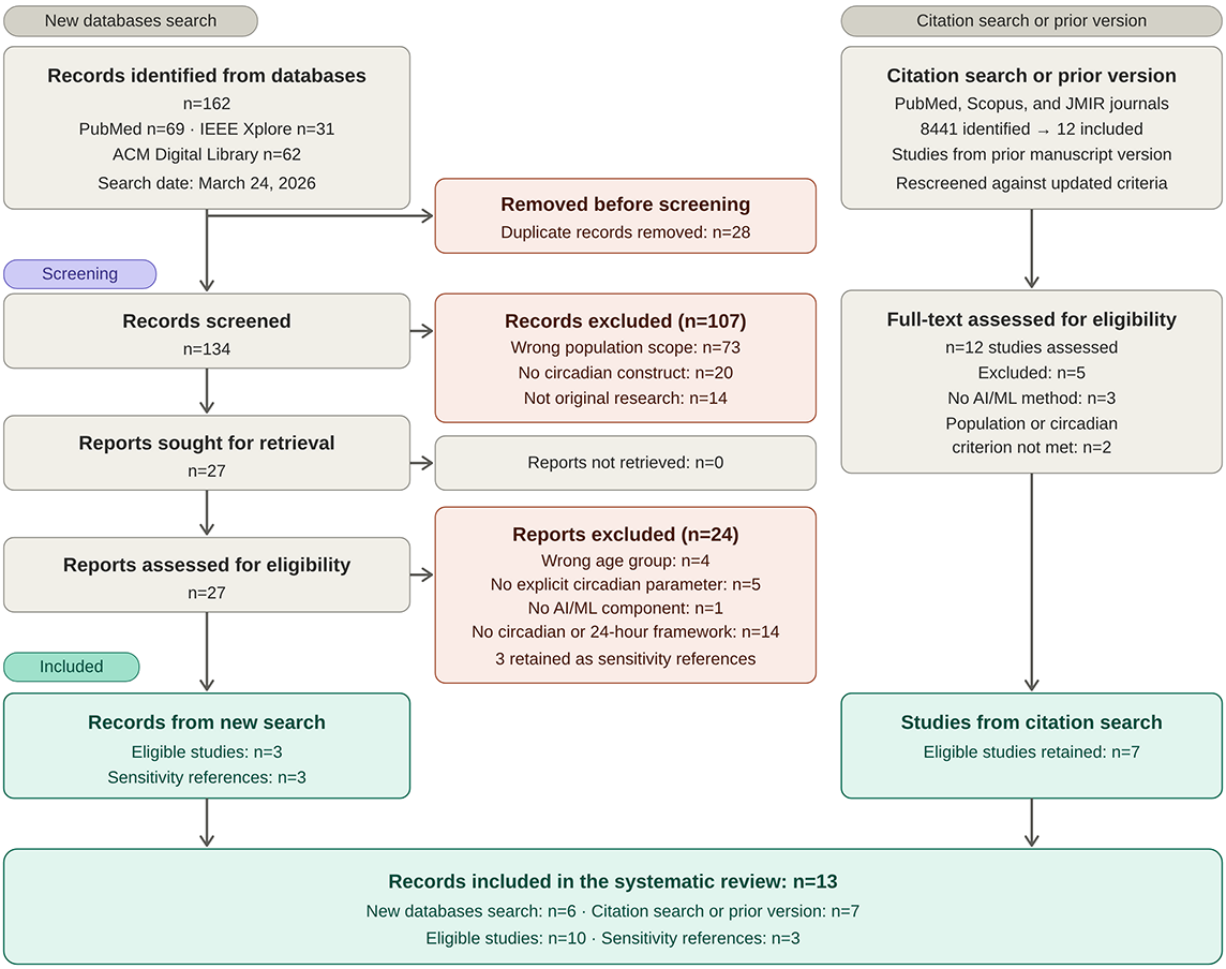
Results

Study Selection

The 2-stage search process yielded a final corpus of 13 records: 10 eligible studies and 3 sensitivity references (Figure 1). From the original search corpus, 7 of 12 previously included studies were retained after rescreening; 5 were excluded. From the complementary search, 162 records were identified across 3 databases (PubMed: n=69; IEEE Xplore: n=31; ACM Digital Library: n=62). After removing 28 duplicates, 134 records were screened at the title and abstract; 107 were excluded. In total, 27 records were assessed at full text; 24 were excluded (wrong age group: n=4; no explicit circadian parameter measured: n=5; no AI/ML component: n=1; no circadian or 24-hour framework: n=14). A total of 3 records met all eligibility criteria and were included; 3 of the 24 excluded records met the definition of sensitivity references and were retained as such. Figure 1 presents the PRISMA flow diagram with the full search strategy.

XSL•FO
RenderX

Figure 1. PRISMA flowchart for study selection (publications 2015-2025; searches run in June 2025 and March 2026). Of the 8603 records identified, 13 were included in the final corpus: 10 met all eligibility criteria, and 3 were retained as sensitivity references. ML: machine learning; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Study Characteristics

The 13 records (10 eligible studies and 3 sensitivity references) were published between 2015 and 2025. They were conducted in the United States (n=7), the United Kingdom (n=2), Taiwan (n=2), the Netherlands (n=1), and Finland (n=1). Sample sizes ranged from 16 to 7297 participants. In total, 6 studies enrolled populations with documented or high-risk neurodegenerative conditions; 4 enrolled community-dwelling or population-based samples without documented cognitive impairment; and 3 used middle-aged or mixed-age samples as methodological anchors. Full characteristics are presented in Table 1.

XSL•FO
RenderX

Table 1. Characteristics of included studies on AI-enhanced wearable devices for circadian monitoring in aging populations.

Author (year)	Population and setting	Device	Circadian parameters	AI/ML[a] method	Key findings
Chuang et al (2024) [25]	Taiwan; n=78 adults (mean 41.5 years; 59% women; 29 obese, 16 overweight, 33 controls)	Wrist actigraphy+Rhythm smartphone app	IS[b], social jet lag, sleep indicators	Actigraphy-based algorithm; multivariate regression	IS strongly correlated with adiposity. App-based measurement provided stronger associations than traditional actigraphy; IS from app correlated with obesity index, unlike social jet lag.
Razjouyan et al (2017) [26]	United States; n=21 adults (mean 50.8, SD 12.8 years; 47.8% women; poor sleep)	Chest sensor (BioPatch), wrist Actiwatch, PSG[c] (reference)	Sleep-wake epoch; implicit rest-activity rhythm	Posture+acceleration algorithm; logistic regression	Chest sensor achieved 85.8% accuracy versus 79.8% for Actiwatch. Combined posture and acceleration improved sleep/wake epoch detection.
Gabb et al (2025) [27]	United Kingdom; n=41 older adults (mean 70.9 years); 10 AD[d], 11 LBD[e], 20 controls	Axivity AX3, Dreem 2 EEG[f] headband, MyDignio app	Sleep continuity, circadian rhythm, cognitive task performance	GGIR[g] automated pipeline; Dreem EEG algorithm	Remote longitudinal monitoring feasible; 98% retention, 79%-97% data completeness. Multimodal approach viable in populations with MCI[h] and dementia.
Cavaillès et al (2025) [28]	United States (MrOS[i] cohort); n=2667 older men	SleepWatch-O wrist actigraph	24-hour activity profiles; sleep/circadian phenotypes	Unsupervised ML clustering (MCGHD[j]+PCA[k])	Three phenotypes: active healthy (64%), fragmented poor (14.1%), and long nappers (21.9%). Fragmented poor sleepers showed higher dementia and CVD[l] event risk.
Wang et al (2024) [29]	Taiwan; n=191 community adults ≥50 years	Actiwatch (3-axis accelerometer+light+temperature sensors), sleep diaries	IS, IV[m], RA[n], chronotype	Kuo actigraphy algorithm; ANOVA+regression	Older adults had lower activity, worse sleep efficiency, earlier chronotype. Women showed greater IS and RA. Age-gender interactions for life satisfaction and RA.
Shim et al (2023) [30]	United States (NHANES[o]); n=7297 adults	ActiGraph GT3X+wrist accelerometer	24-hour activity profiles as digital biomarkers (inflammation, biological age, mortality)	Hierarchical clustering (Ward method); recursive feature elimination	Five distinct 24-hour activity profiles identified. Circadian misalignment linked to systemic inflammation, accelerated biological age, and increased mortality risk.
Jones et al (2025) [31]	United States; n=40 older adults (mean 67.2 years; 72.5% women); 19 MCI, 21 controls	Apple Watch+myRhythmWatch platform	IS, IV, RA, cosinor (MESOR[p], acrophase, rhythm strength)	myRW[q] algorithm; cosinor analysis; linear regression	Later activity onset associated with worse memory. Fragmented rhythms linked to slower processing speed. Platform feasible for circadian assessment in MCI.

XSL•FO
RenderX

Author (year)	Population and setting	Device	Circadian parameters	AI/ML[a] method	Key findings
Beiglary et al (2025) [32]	United States; n=29 older adults with dementia; 2 assisted-living facilities	Wearable light sensor calibrated versus spectrophotometer	Circadian stimulus, personal circadian lighting exposure	Random forest (ML); laboratory calibration+field deployment	Random forest accurately predicted circadian stimulus. Substantial interindividual variability in circadian light exposure. In total, 40% device loss; validated against spectrophotometric ground truth.
Merilahti et al (2016) [23][r]	Finland; n=16 nursing home residents (mean 90.7 years; 7 with dementia)	Wrist-worn device with integrated activity sensor	MESOR, amplitude, acrophase, IS, IV, RA, CRS[s]	Single cosinor function (24 hours); nonparametric RAS[t] analysis	Physical functioning associated with rhythm strength and stability. Functional decline tracked with reduced activity, lower MESOR/amplitude, and increased IV.
Froggatt et al (2020) [33]	United Kingdom; n=32 nursing home residents with advanced dementia (mean 81.5 years)	ActiGraph wrist actigraphy	IS, IV (circadian fragmentation and synchronization)	Nonparametric actigraphy analysis (IS/IV computation)	Feasibility cluster RCT[u]. Highly inactive population with fragmented rhythms (IS: mean 0.08; IV: mean 0.7). Actigraphy feasible in advanced dementia.
Sprint et al (2022) [34]	United States; n=14 amnestic MCI (mean 74.4 years) in smart-home environment	Smart-home sensors+EMMA tablet app+text markers	Circadian rhythm strength, 24-hour circadian rhythm, regularity indices	Multimodal fusion; joint prediction (regression/ML)	Moderate-to-large prediction correlations (r=0.601-0.871) for 10 clinical assessments. Joint prediction outperformed independent modality models.
de Zambotti et al (2015) [21][r]	United States; n=28 midlife women (mean 50.1 years); sleep laboratory	Jawbone UP consumer tracker vs PSG (reference)	Sleep-wake rhythm (nightly sleep metrics)	Proprietary consumer algorithm; PSG comparison	High sleep sensitivity, poor wake specificity; overestimated TST[v], underestimated WASO[w]. Potentially useful for daily sleep-wake rhythm tracking over multiple days.

XSL•FO
RenderX

Author (year)	Population and setting	Device	Circadian parameters	AI/ML[a] method	Key findings
Robben et al (2017) [22][r]	Netherlands; n=13 older adults (mean 80, SD 7 years); independent living	Ambient smart-home sensors (PIR[x], pressure, switches); no on-body device	24-hour behavioral structure as temporal proxy	Delta-feature temporal modeling; random forest prediction	Delta features predicted functional health change, outperforming static baselines. Temporal behavioral regularity exploited as implicit circadian proxy.

[a]ML: machine learning.
[b]IS: interdaily stability.
[c]PSG: polysomnography.
[d]AD: Alzheimer disease.
[e]LBD: Lewy body dementia.
[f]EEG: electroencephalography.
[g]GGIR: R-package for accelerometry.
[h]MCI: mild cognitive impairment.
[i]MrOS: Osteoporotic Fractures in Men Study.
[j]MCGHD: multiple coalesced generalized hyperbolic distribution.
[k]PCA: principal component analysis
[l]CVD: cardiovascular disease.
[m]IV: intradaily variability.
[n]RA: relative amplitude.
[o]NHANES: National Health and Nutrition Examination Survey.
[p]MESOR: midline estimating statistic of rhythm.
[q]myRW: myRhythmWatch.
[r]Sensitivity-reference study.
[s]CRS: circadian rhythm strength.
[t]RAS: rest-activity rhythm analysis.
[u]RCT: randomized controlled trial.
[v]TST: total sleep time.
[w]WASO: wake after sleep onset.
[x]PIR: passive infrared.

Risk of Bias and Quality Assessment

Risk of bias was assessed with the MMAT (version 2018) [24], as described in the Methods section. The complete dataset to enhance transparency is available as Multimedia Appendix 3.

MMAT results are presented in Table 2. In total, 8 studies were classified as quantitative descriptive (type 4), 4 as quantitative nonrandomized (type 3), and 1 as a quantitative RCT (type 2). A total of 6 studies received an overall low-risk appraisal. Wang et al [29] was rated moderate risk due to a nonrepresentative sample (Q2: No). Beiglary et al [32] was rated moderate risk because sample representativeness could not be determined (Q2: Cannot tell), and Merilahti et al [23] because low reporting bias could not be confirmed (Q4: Cannot tell). Sprint et al [34] was rated as moderate-to-high risk, given a small sample size (n=14), no control of confounders, and uncertain completeness of outcome data. Froggatt et al [33] received a high-risk appraisal, as expected for a complex-care feasibility RCT with nonblinded outcomes; this does not diminish the descriptive validity of the actigraphy findings. All 3 sensitivity-reference studies were rated moderate risk.

XSL•FO
RenderX

Table 2. Mixed Methods Appraisal Tool (MMAT) results for included studies[a].

Author (year)	MMAT type	S1	S2	Q1	Q2	Q3	Q4	Q5	Appraisal
Chuang et al (2024) [25]	Type 4	Y[b]	Y	Y	Y	Y	Y	Y	Low risk
Razjouyan et al (2017) [26]	Type 4	Y	Y	Y	Y	Y	Y	Y	Low risk
Gabb et al (2025) [27]	Type 4	Y	Y	Y	Y	Y	Y	Y	Low risk
Cavaillès et al (2025) [28]	Type 3	Y	Y	Y	Y	Y	Y	Y	Low risk
Wang et al (2024) [29]	Type 4	Y	Y	Y	N[c]	Y	Y	Y	Moderate
Shim et al (2023) [30]	Type 3	Y	Y	Y	Y	Y	Y	Y	Low risk
Jones et al (2025) [31]	Type 4	Y	Y	Y	Y	Y	Y	Y	Low risk
Beiglary et al (2025) [32]	Type 4	Y	Y	Y	C[d]	Y	Y	Y	Moderate
Merilahti et al (2016) [23][e]	Type 4	Y	Y	Y	Y	Y	C	Y	Moderate
Froggatt et al (2020) [33]	Type 2	Y	Y	Y	N	C	N	C	High risk[f]
Sprint et al (2022) [34]	Type 3	Y	Y	C	Y	C	N	Y	Moderate-high
de Zambotti et al (2015) [21][e]	Type 4	Y	Y	Y	C	Y	Y	Y	Moderate
Robben et al (2017) [22][e]	Type 3	Y	Y	Y	Y	Y	C	Y	Moderate

[a]Type 2 (randomized controlled trial): Q1=randomization appropriately performed; Q2=groups comparable at baseline; Q3=complete outcome data; Q4=outcome assessors blinded; Q5=participants adhered to the assigned intervention. Type 3 (nonrandomized): Q1=participants representative of the target population; Q2=appropriate measurement of outcome and exposure; Q3=complete outcome data; Q4=confounders accounted for in design and analysis; Q5=exposure occurred as intended. Type 4 (descriptive): Q1=sampling strategy relevant to the research question; Q2=sample representative of the target population; Q3=appropriate measurements; Q4=low risk of nonresponse bias; Q5=appropriate statistical analysis.

[b]Y: yes.

[c]N: no.

[d]C: cannot tell.

[e]Sensitivity-reference study.

[f]High risk is expected and acceptable for a feasibility cluster randomized controlled trial.

Synthesis by Domain

Device Class and Sensing Modality

Wrist actigraphy was the most used modality, enabling derivation of nonparametric circadian metrics and cosinor parameters from continuous rest-activity data recorded over days to months. Devices deployed included ActiGraph GT3X+ [30], SleepWatch-O [28], Axivity AX3 [27], Actiwatch [29], Apple Watch with the myRhythmWatch platform [31], an unspecified integrated activity-sensing wrist device [23], and a spectrophotometer-calibrated wearable light sensor [32]. Device validation varied widely: Beiglary et al [32] performed explicit laboratory calibration against a spectrophotometric reference before field deployment; de Zambotti et al [21] compared consumer tracker output with polysomnography; and Razjouyan et al [26] validated a novel algorithm against polysomnography in a sleep laboratory setting. Most other studies used devices without concurrent within-study validation. Multisensor platforms that combine actigraphy with EEG headbands, apps, and pulse oximetry were used by Gabb et al [27]. Smart-home ambient sensor arrays without on-body devices were used in Sprint et al [34] and Robben et al [22]. A smartphone-actigraphy combination formed the sensing platform in Chuang et al [25].

Circadian Construct and Parametric Specificity

Studies differed substantially in how explicitly they operationalized circadian biology. Six studies derived or analyzed named chronobiological metrics as primary or clearly defined outcomes: IS, IV, RA, MESOR, acrophase, and circadian stimulus [23,25,29,31-33]. Four studies embedded temporal or 24-hour rhythm features within predictive or clustering frameworks without foregrounding them as primary end points [27,28,30,34]. Two studies used continuous sleep-wake rhythm monitoring as a functional proxy for circadian structure without computing formal circadian parameters [21,26]. One study treated ambient behavioral regularity as an implicit temporal proxy [22].

This gradient from explicit to implicit circadian operationalization has important implications for interpreting studies as evidence for circadian biomarkers. Only studies reporting formal chronobiological metrics provide direct evidence for the circadian construct. Studies that use sleep-wake rhythm or behavioral regularity as proxies provide contextual support but cannot serve as substitutes for direct circadian measurement.

AI/ML Method Family

Four AI/ML families were identified (Table 3). Supervised predictive modeling was applied in 3 studies: a random forest predicting circadian light exposure from a calibrated wearable

XSL•FO
RenderX

light sensor [32]; decision tree, random forest, and gradient-boosted regressors predicting clinical assessment scores from multimodal behavioral markers in MCI [34]; and regression forests predicting change in functional health from ambient sensor data [22]. Unsupervised phenotyping identified distinct 24-hour activity profiles in 2 large-sample studies, using hierarchical clustering after recursive feature elimination [30] and generalized hyperbolic mixture modeling on principal components of actigraphy variables [28]. Algorithm development or validation was the primary contribution of 2 studies: a chest-worn sleep-wake classifier developed against polysomnography [26] and the polysomnographic validation of a consumer tracker's proprietary scoring algorithm [21]. The remaining 6 studies applied deterministic algorithms (nonparametric rest-activity indices, cosinor fitting, and automated sleep-wake scoring) followed by conventional statistical inference to relate circadian metrics to functional, clinical, or feasibility outcomes [23,25,27,29,31,33].

Reporting of these methods was uneven (Table 3). No model was validated on an independent external cohort, and no study reported following an AI-specific reporting guideline; 3 cited general reporting guidelines (STROBE [Strengthening the Reporting of Observational Studies in Epidemiology] [27,28], TIDieR [Template for Intervention Description and Replication], and CONSORT [Consolidated Standards of Reporting Trials] [33]). Seven studies fitted no predictive model, so validation, partitioning, and hyperparameter criteria did not apply to them [21,23,25,27,29,31,33]. Of the 4 studies with a training-testing distinction, 2 stated that data were partitioned at the participant level [22,26]; 1 reported leave-one-out cross-validation over a participant-level feature matrix without stating the partition unit [34]; and 1 described a train-test split and 5-fold cross-validation for its linear models without stating the unit or proportion of the split, while its deposited random forest code fitted the complete training set without resampling [32]. The 2 unsupervised studies performed no held-out evaluation: one assessed cluster stability by repeated subsampling [28], and the other reported no validation of the clustering solution [30]. In total, 5 studies reported hyperparameter values partially [22,26,28,30,34], and 1 relied on package defaults that are visible only in its deposited code [32]; none reported every setting needed to reproduce its model. A total of 3 studies deposited analysis code, in each case covering only part of the analysis [28,32,34], and 3 made their data openly available [28,30,32].

Explainability was addressed only through model-level variable importance. One study applied a post-hoc attribution method, permutation importance for its random forest [32], and 3 reported model-internal importance measures: Gini importance [22], variables ranked by their contribution to cluster separation [28], and random forest importance within the feature-selection step that preceded clustering, rather than for the clustering itself [30]. No study used local attribution methods such as Shapley Additive Explanations (SHAP) or Local Interpretable Model-Agnostic Explanations (LIME), and none explained individual predictions.

XSL•FO
RenderX

Table 3. Reporting of computational methods in the included studies[a].

Author (year)	AI/ML[b] family	Validation	Partition unit	Hyperparameters	Code	Data	XAI[c]
Chuang et al (2024) [25]	Pattern association	Reference standard[d]	N/A[e]	N/A	NR[f]	Request	N/A
Razjouyan et al (2017) [26]	Algorithm validation	CV	Participant	Partial	NR	NR	NR
Gabb et al (2025) [27]	Pattern association	N/A	N/A	N/A	NR	Request	N/A
Cavaillès et al (2025) [28]	Unsupervised	Resubstitution[g]	N/A	Partial	Open[h]	Open	Importance
Wang et al (2024) [29]	Pattern association	N/A	N/A	N/A	NR	Request	N/A
Shim et al (2023) [30]	Unsupervised	NR	N/A	Partial	NR	Open	Importance[i]
Jones et al (2025) [31]	Pattern association	N/A	N/A	N/A	NR	NR	N/A
Beiglary et al (2025) [32]	Supervised	Unclear[j]	NR	Defaults[k]	Open[h]	Open	Post hoc
Merilahti et al (2016) [23][l]	Pattern association	N/A	N/A	N/A	NR[m]	NR	N/A
Froggatt et al (2020) [33]	Pattern association	N/A	N/A	N/A	NR	Request	N/A
Sprint et al (2022) [34]	Supervised	CV	Unclear[n]	Partial	Open[h]	NR	NR
de Zambotti et al (2015) [21][l]	Algorithm validation[o]	Reference standard	N/A	N/A	NR[m]	NR	N/A
Robben et al (2017) [22][l]	Supervised	CV	Participant	Partial	NR	NR	Importance[p]

[a]AI/ML family: supervised=supervised predictive modeling; unsupervised=unsupervised phenotyping or clustering; algorithm validation=algorithm development or validation; pattern association=pattern-association modeling with conventional statistical inference. Validation: CV=cross-validation; reference standard=comparison against a concurrent reference standard; resubstitution=performance reported on the fitting data. Partition unit: participant=data partitioned at the participant level. Hyperparameters: partial=some values reported; defaults=package defaults used. Code and data: open=publicly available; request=available on request. XAI: post hoc=post hoc attribution method; importance=variable importance without a named attribution framework; unclear=reporting insufficient to determine. No study reported external validation on an independent cohort or adherence to an AI-specific reporting guideline.

[b]ML: machine learning.

[c]XAI: explainable AI.

[d]Reference standard was research-grade actigraphy, not polysomnography.

[e]N/A: not applicable (no fitted predictive model).

[f]NR: not reported.

[g]Internal stability only (100 subsamples excluding 5% of participants; adjusted Rand index reported); no held-out performance evaluation.

[h]Deposited code covers only part of the analysis: cluster identification [28], the circadian light model [32], and digital marker extraction [34].

[i]Variable importance reported for the random forest used in recursive feature elimination, not for the clustering model.

[j]Five-fold cross-validation and a train-test split are described for the linear regression models; the deposited random forest code fits the full training set without resampling.

[k]Package defaults evident in the deposited code; not stated in the text; no random seed.

[l]Sensitivity-reference study.

[m]The device algorithm is stated to be proprietary; no statement on the authors' analysis code.

[n]Leave-one-out cross-validation over a participant-level feature matrix; partition unit not stated explicitly.

[o]Validation of a commercial algorithm against polysomnography; no algorithm was developed by the authors.

[p]The authors also state that the learned decision rules are interpretable.

Population and Neurocognitive Context

Six studies enrolled populations with documented or high-risk neurodegenerative conditions. Three targeted residents of nursing homes or assisted-living facilities with advanced dementia [23,32,33], where circadian fragmentation was most severe (IS mean ~0.08, IV mean ~0.7 in Froggatt et al [33]), and actigraphy feasibility was explicitly evaluated. Two enrolled participants with amnestic MCI [31,34], with circadian metrics linked to cognitive test scores and clinical assessment predictions. One enrolled a mixed MCI/dementia cohort alongside a cognitively normal control [27]. Four studies used

XSL•FO
RenderX

community-dwelling or population-based samples without documented cognitive impairment [22,28-30], examining 24-hour activity profiles in relation to dementia risk, biological age, or functional health as outcome proxies. Three studies provided methodological contributions using middle-aged or mixed-age samples [21,25,26].

Sex composition was notably imbalanced: Cavaillès et al [28] enrolled exclusively older men (Osteoporotic Fractures in Men Study [MrOS] cohort), and de Zambotti et al [21] enrolled exclusively women. Racial and ethnic diversity was rarely reported. No study was conducted in low- or middle-income country settings, despite the global burden of neurodegeneration in those populations.

The 4 analytical domains examined across this synthesis do not operate independently. Their recurrence across the included studies reveals a set of structural dependencies: sensing modality choices constrain the circadian constructs that can be extracted; construct operationalization specificity determines the inferential ceiling available to AI/ML methods; and the characteristics of the target population bound the generalizability of any biomarker claim. Taken together, these dependencies constitute the functional logic of a pipeline, not merely a descriptive taxonomy.

Across domains, the evidence shows a consistent pattern: continuous circadian-related data capture is technically feasible, but biomarker interpretation remains constrained by construct heterogeneity, limited external validation, and population-level generalizability gaps. The most robust evidence comes from actigraphy-derived rest-activity fragmentation in institutional or cognitively vulnerable populations, whereas AI-based prediction in community and preclinical settings remains exploratory. The Evidence-Derived Proposal: Toward Clinically Deployable AI-Circadian Biomarker Systems: Structural Challenges, a Reference Pipeline Architecture, and Validation Priorities section translates that logic into a proposed reference architecture and layered validation pathway grounded in the evidence synthesized earlier.

Evidence-Derived Proposal: Toward Clinically Deployable AI-Circadian Biomarker Systems: Structural Challenges, a Reference Pipeline Architecture, and Validation Priorities

The structural challenges, reference architecture, and validation pathway that follow were derived by the authors from the synthesized evidence; they are not outcomes of the screening and appraisal process described in the Methods section.

Structural Challenges at the Sensing-to-Biomarker Interface

The 13-study corpus reveals a field in which technical capacity has outpaced analytical standardization. Six structural challenges constrain the transition from ecologically valid continuous monitoring to validated predictive biomarkers for neurodegenerative disease, and each aligns with a specific stage of the pipeline that connects sensing hardware to clinical interpretation.

At the sensing layer, device heterogeneity is compounded by inconsistent within-study validation. Only 3 studies [21,26,32] performed explicit concurrent validation against a reference standard (polysomnography or laboratory spectrophotometry), while the remaining studies assumed device validity in populations for which independent performance evidence is often lacking. This limitation intensifies in advanced dementia, where motor disorganization, behavioral symptoms, and adherence constraints may invalidate calibration obtained from younger or healthier samples, as illustrated by the 40% device loss reported by Beiglary et al [32].

At the parameter extraction layer, circadian end-point heterogeneity prevents cross-study integration. Studies using nonparametric metrics (IS, IV, and RA), cosinor-derived parameters (MESOR, acrophase, and rhythm strength), circadian stimulus, and unsupervised activity phenotypes differ not only in end-point choice but also in epoch length, analysis software, and time window, precluding quantitative synthesis despite consistent directional findings. A parallel concern is the gradient from explicit chronobiological operationalization to implicit proxies based on sleep-wake rhythm or behavioral regularity; the distinction matters because only the former supports direct circadian biomarker inference; yet, it is frequently blurred in cross-study interpretation.

At the AI/ML inference layer, 2 interlocking challenges stand out. The first is the absence of external validation, compounded by incomplete reporting of how models were validated. Every model in the corpus was developed within a single dataset, and none was evaluated on an independent cohort; in 2 of the 4 studies with a training-testing distinction, the unit of data partitioning could not be determined from the report; no study reported every hyperparameter needed for reproduction; and none followed an AI-specific reporting guideline such as TRIPOD+AI [16] (Table 3). Even the large-sample clustering studies [28,30] relied on retrospective datasets without prospective biomarker validation designs. The same pattern has been documented for AI-based risk prediction in sports medicine, where only 5 of 39 injury prediction studies reported external validation [35], which points to a methodological gap shared across physiological monitoring domains rather than one specific to circadian research. The second challenge is the shallowness of explainability. Four studies reported variable importance [22,28,30,32], but only at the level of the whole model: no study used local attribution methods such as SHAP [36] or LIME [37], and none explained individual predictions, which is the level at which a clinician would act on a biomarker result. Converging regulatory expectations require transparency and human oversight before AI-based biomarkers are deployed clinically; for circadian models intended to inform neurodegeneration risk, this gap is a direct barrier to regulatory clearance and clinician trust.

At the population and deployment layer, 2 further challenges emerge. All included studies were conducted in high-income countries, with limited reporting of racial and ethnic diversity and pronounced sex imbalances, with exclusively male [28] or exclusively female [21] samples. Given that the global burden of neurodegeneration falls disproportionately on populations underrepresented in this corpus, biomarkers validated in this

XSL•FO
RenderX

sample frame cannot be assumed to transfer to the populations for whom early detection is most urgent. Compounding this, the descriptive-to-predictive gap limits clinical actionability: 7 of the 13 studies fitted no predictive model, applying deterministic algorithms for feature extraction followed by conventional statistical inference, and the 6 that did were small, single-site, or retrospective (Table 3).

These 6 challenges are not independent; they constrain one another. Device validation failures propagate into parameter extraction errors, which in turn exacerbate AI model overfitting and confound the association with clinical outcomes. Addressing them individually is insufficient; they must be addressed in structural relation to the pipeline that generates a circadian biomarker (Table 4).

Table 4. Structural challenges at the sensing-to-biomarker interface and their architectural responses.

Identified challenge/gap	Why it matters	Architectural/design response	Future validation implication
Device validation heterogeneity across populations	Calibration from younger or healthier samples may not hold in advanced dementia, where device loss and motion artifact are substantial	Layer 1: concurrent validation against reference standards in the target population as a precondition, not an add-on	Population-specific device validation required before biomarker claims
Circadian end-point heterogeneity (metrics, epoch lengths, and pipelines)	Prevents quantitative synthesis and direct effect-size comparison across studies	Layer 2: harmonized parameter set, standardized analysis pipelines, explicit distinction between direct circadian measurement and behavioral proxies	Minimum standard reporting of end points, epochs, and pipelines in biomarker studies
Absence of external validation and incomplete AI reporting	All included AI models are developed within a single dataset; validation procedures and hyperparameters are incompletely reported, and no study follows an AI-specific reporting guideline	Layer 3: external replication on independent cohorts as a design requirement	Cross-cohort generalization evidence and guideline-concordant reporting (eg, TRIPOD+AI[a]) before clinical translation claims
Explainable AI (XAI) gap	Regulatory deployment and clinical trust require transparency; importance reporting in the corpus is model-level only, with no explanation of individual predictions	Layer 3: explainability mechanisms integrated at design time, with model outputs mapped to chronobiologically interpretable parameters	Prespecified XAI reporting, including explanations of individual predictions (eg, SHAP[b] and LIME[c]), as a standard element of biomarker studies
Population diversity deficit	High-income, demographically narrow samples do not represent the global burden of neurodegeneration	Layer 4: demographic representativeness as an architectural concern, not a discussion caveat	Quantified transferability across sex, race, ethnicity, and country-income setting as a validation criterion
Descriptive-to-predictive gap	In total, 7 of 13 studies fit no predictive model; the 6 that do are small, single-site, or retrospective	Layer 3: principled distinction between algorithmic processing and learned representation, with implications for evidence strength	Prospective biomarker validation designs rather than retrospective cross-sectional association

[a]TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis + Artificial Intelligence

[b]SHAP: Shapley Additive Explanations.

[c]LIME: Local Interpretable Model-Agnostic Explanations.

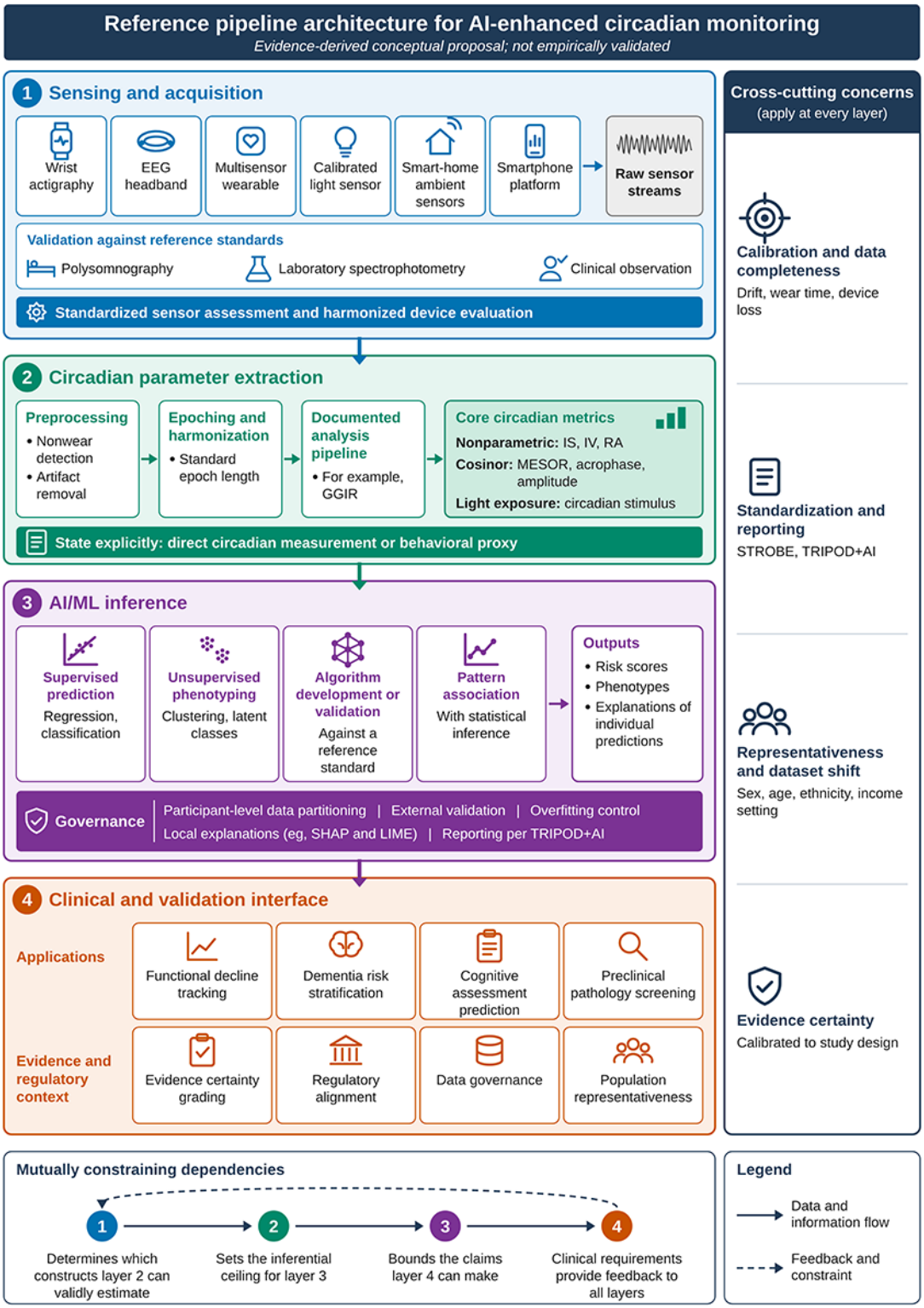
A Reference Pipeline Architecture for AI-Enhanced Circadian Monitoring

Overview

The recurrence of 4 analytical domains across the reviewed corpus supports a 4-layer reference architecture for AI-enhanced circadian biomarker pipelines (as Figure 2 shows). The architecture is conceptual: it does not prescribe a specific implementation but articulates the functional dependencies that any credible biomarker system in this field must respect.

XSL•FO
RenderX

Figure 2. Reference pipeline architecture for AI-enhanced circadian monitoring. EEG: electroencephalography; GGIR: R package for processing raw accelerometer data; IS: interdaily stability; IV: intradaily variability; LIME: Local Interpretable Model-Agnostic Explanations; MESOR: midline estimating statistic of rhythm; ML: machine learning; RA: relative amplitude; SHAP: Shapley Additive Explanations; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis+Artificial Intelligence.



Sensing and Acquisition Layer

This layer encompasses the modalities observed in the corpus: wrist actigraphy as the dominant channel, supplemented by EEG headbands, multisensor wearables, calibrated wearable light sensors, ambient smart-home sensor arrays, and smartphone-based platforms. Its function is to produce continuous, ecologically valid physiological and behavioral data streams with explicit, population-specific validation against reference measurements. The cross-cutting concerns at this layer are device calibration, data completeness under real-world conditions, and feasibility in the populations for whom the

XSL•FO
RenderX

biomarker is intended, particularly institutionalized individuals with advanced dementia, where adherence, motion artifact, and device loss are nontrivial [32,33]. Calls to standardize how sensor technologies and the algorithms built on them are assessed have also been made in applied fields outside medicine [38]; in circadian biomarker pipelines, this standardization matters most at layers 1 and 2, because device validation and parameter extraction set the ceiling for every inference downstream.

Circadian Parameter Extraction Layer

This layer converts raw sensor streams into chronobiologically interpretable metrics. The corpus supports a core parameter set comprising nonparametric metrics (IS, IV, and RA), cosinor-derived parameters (MESOR, acrophase, and rhythm strength), and specialized metrics such as circadian stimulus for light exposure studies. Its function is to produce reproducible, comparable circadian end points. The cross-cutting concerns are harmonized epoch length, standardized analysis pipelines (such as GGIR or equivalently documented alternatives), explicit operationalization of whether a study measures circadian biology directly or uses behavioral proxies, and transparent reporting that enables downstream meta-analytic integration.

AI/ML Inference Layer

This layer applies computational reasoning to extracted parameters or, in more advanced designs, directly to raw sensor streams. The corpus demonstrates 4 viable method families: supervised predictive modeling, unsupervised phenotyping, algorithm development and validation, and pattern-association modeling combined with conventional statistical inference. Its function is to extract clinically meaningful representations beyond the reach of single-metric description. The cross-cutting concerns are external validation across independent datasets, explainability mechanisms compatible with regulatory expectations, overfitting control under small-sample conditions, and the principled distinction between genuine learned representations and algorithmic feature engineering, a distinction that directly affects how results should be interpreted as biomarker evidence.

Clinical and Validation Interface Layer

This layer situates pipeline outputs within clinical, epidemiological, and regulatory contexts. Its function is to connect circadian biomarker outputs to actionable constructs: functional decline tracking, dementia risk stratification, cognitive assessment prediction, and preclinical pathology screening. The cross-cutting concerns are population representativeness across sex, race, ethnicity, socioeconomic status, and geography; regulatory alignment (including XAI requirements and medical device pathways); governance of longitudinal biometric data; and the evidence certainty calibration that distinguishes moderate-certainty institutional findings from lower-certainty community preclinical claims.

The layers are not merely sequential; they are mutually constraining. Layer 1 decisions determine the circadian constructs that layer 2 can produce. Layer 2 decisions determine the inferential ceiling available to layer 3. Layer 3 decisions determine the evidence strength with which layer 4 claims can be made. This dependency structure is the architectural translation of the cross-layer failure propagation described in the Structural Challenges at the Sensing-to-Biomarker Interface subsection.

Conceptual Validation Logic and Implementation Propositions

The proposed architecture suggests a layered validation logic that may guide subsequent empirical work. Rather than evaluating AI-circadian biomarker systems solely by end-point predictive accuracy, each layer should be assessed against criteria appropriate to its function, with the overall pipeline's evidence certainty bounded by its weakest layer.

At layer 1, validation should establish device performance in the specific population of interest, with concurrent comparison against an accepted reference (polysomnography, laboratory spectrophotometry, or supervised clinical observation, depending on modality) and transparent reporting of data completeness, device loss, and calibration drift over the monitoring period. At layer 2, validation should demonstrate that extracted circadian parameters are reproducible under standardized pipelines and interpretable within chronobiological theory, with an explicit statement of whether the study measures circadian biology directly or uses a behavioral proxy. At layer 3, validation should require external replication on independent cohorts, prespecified explainability reporting, and honest characterization of overfitting risk, particularly when sample sizes approach those observed in most included clinical studies. At layer 4, validation should evidence diagnostic or prognostic utility in prospectively followed, demographically representative cohorts, with certainty grading that respects the gradient already visible in the present corpus: moderate certainty for actigraphy-derived fragmentation-function associations in institutional settings and low certainty for AI-based predictive claims in community or preclinical populations.

From this logic, 4 implementation propositions follow directly from the evidence base. First, prospective longitudinal multisite cohort studies using harmonized circadian end points (nonparametric metrics and cosinor parameters computed with consistent epoch lengths and transparently reported pipelines) are required across cognitively normal aging, MCI, and dementia populations to move the field beyond retrospective and cross-sectional evidence. Second, AI models intended for clinical deployment should incorporate explainability mechanisms at design time rather than retrofit them, with explicit mapping between model outputs and chronobiologically interpretable parameters. Third, device and algorithm validation should be treated as population-specific rather than generic: performance in younger or healthier samples does not warrant extrapolation to institutionalized older adults or individuals with advanced dementia, where the biomarker is most clinically consequential. Fourth, demographic diversity should be operationalized as a validation criterion (not an ethical aspiration) with biomarker transferability quantified across sex, race, ethnicity, and country-income setting before generalizability is claimed.

XSL•FO
RenderX

Taken together, the challenge synthesis, reference architecture, and layered validation logic articulate what a clinically deployable AI–circadian biomarker system would require, and where the existing evidence in this review stops short of those requirements. Circadian disruption remains a biologically plausible, technically measurable, and increasingly computable candidate biomarker for neurodegeneration. Its maturation into a validated clinical tool depends less on additional sensing innovation than on structural coherence across the pipeline from device to clinical interface, on external and population-specific validation of AI components, and on explainability and governance frameworks commensurate with the clinical stakes.

Discussion

Principal Findings

Building on the cross-domain analysis, 3 convergent findings emerge from the 13 studies examining AI-enhanced wearable and objective sensing systems for circadian monitoring in older or cognitively vulnerable populations. First, circadian fragmentation (measurable as elevated IV, reduced IS, and attenuated RA from wrist actigraphy) is consistently documented in populations with advanced dementia and is associated with concurrent functional decline [23,33]. Second, AI and ML contribute meaningful value beyond descriptive actigraphy by enabling automated phenotyping of 24-hour activity profiles at a population scale [28,30], calibrated prediction of circadian light exposure [32], and multimodal fusion of behavioral markers into clinically interpretable predictions [34]. Third, the field's technical capacity to collect ecologically valid continuous data has advanced faster than its ability to standardize circadian end points, harmonize device-agnostic reporting, and independently validate AI models across diverse populations.

Contextual Interpretation and Recent Evidence

The association between actigraphy-derived circadian fragmentation and dementia risk is supported by prospective cohort evidence extending well beyond this review's corpus. Lim et al [39] and Musiek et al [40] demonstrated that rest-activity fragmentation predicts incident dementia and tracks preclinical Alzheimer pathology, supporting the biological plausibility of IS and IV as early biomarkers. The present review advances this framing by showing that AI approaches (particularly unsupervised clustering [28,30] and supervised learning [32]) can extract richer, phenotypically meaningful representations from these data streams that go beyond single-metric descriptors.

The multimodal sensing approach in Gabb et al [27] and Sprint et al [34] aligns with emerging frameworks for remote monitoring of cognitive decline that combine multiple behavioral data streams to improve predictive validity. However, neither study achieved external validation, and both were small. In Sprint et al [34], leave-one-out cross-validation on 14 participants leaves substantial overfitting risk; Gabb et al [27] fitted no predictive model, so its contribution is feasibility evidence rather than model performance.

Explainability was addressed only partially. One study applied permutation importance to a random forest [32], and 3 reported model-internal variable importance [22,28,30]; no study used local attribution methods such as SHAP [36] or LIME [37], and none explained individual predictions. Regulatory frameworks, including the US Food and Drug Administration's action plan for AI/ML-based software as a medical device [41], emphasize that clinical deployment of AI-based biomarkers requires transparency, interpretability, and human oversight, expectations that already apply to a growing number of approved AI/ML-enabled devices [42]. Knowing which hours of the day or which light-sensor channels drive a model helps during development, but a clinician deciding whether one patient's rest-activity rhythm signals elevated risk needs to know why the model flagged that patient. Circadian biomarker models that cannot give that account are unlikely to earn clinical trust or regulatory clearance in neurodegenerative contexts.

Population diversity remains a critical shortcoming. The Lancet Commission on Dementia has underscored that the global burden of neurodegenerative disease falls disproportionately on populations underrepresented in biomarker research [43]. All included studies were conducted in high-income countries, with limited reporting of racial and ethnic diversity. Wearable-based circadian biomarkers validated exclusively in these populations cannot be assumed to generalize to the populations most affected by neurodegeneration worldwide. In ML terms, this is a problem of dataset shift [44]: a model learns the relationship between sensor signals and outcomes in the population it was trained on, and that relationship may not hold when age, sex, ethnicity, comorbidity, or device-wearing behavior are distributed differently. Performance in the development sample is therefore not evidence of performance in underrepresented groups.

Limitations of the Evidence and of This Review

Sample sizes were predominantly small (n=16 to ~50 in clinical studies), with adequate statistical power only in the 2 large-scale epidemiological studies [28,30] that used retrospective data without prospective biomarker validation designs. Sex imbalances (male-only in MrOS [28]; female-only in de Zambotti et al [21]) limit generalizability. In total, 7 of the 13 studies fitted no predictive model, so conclusions about AI model performance rest on 6 studies [22,26,28,30,32,34], each developed within a single dataset and none externally validated. The assessment of methods reporting relied on the published articles and their supplementary materials; details available only in protocols or repositories not referenced by the articles were not captured. Circadian end-point heterogeneity across studies (different parametric conventions, epoch lengths, device positions, analysis software, and time windows) prevented quantitative synthesis and precluded direct cross-study comparison of effect estimates. Certainty in the evidence is moderate for actigraphy-derived fragmentation-function associations in institutional settings and low to very low for AI-based circadian biomarker claims in community or preclinical populations.

Regarding the review itself, the original search did not cover IEEE Xplore or the ACM Digital Library, which were added in the complementary search following peer-review recommendations. The complementary search is a documented

XSL•FO
RenderX

amendment to the deposited protocol, as described in the Methods. The reference architecture and validation pathway are an interpretive synthesis by the authors; they were not produced by the systematic review process and have not been empirically validated.

Implications and Future Directions

Four research priorities are indicated for the field to progress toward clinically deployable AI-based circadian biomarkers. First, prospective longitudinal cohort studies with standardized circadian metric reporting (including IS, IV, RA, and cosinor parameters calculated with harmonized epoch lengths and analysis pipelines) are needed across cognitively normal aging, MCI, and dementia populations. Second, AI models should incorporate explainability mechanisms to facilitate clinical interpretation and regulatory compliance. Third, device validation must be population-specific, accounting for the performance differences expected in very old, institutionalized, or demented populations. Fourth, population diversity in sex, race, ethnicity, and socioeconomic status must be actively prioritized to support generalizable biomarker development.

The structural interdependence of these 4 priorities (across sensing, parameter extraction, AI/ML inference, and clinical validation) motivates the reference architecture and layered validation pathway developed in the previous section, which maps each challenge to its corresponding pipeline layer and proposes implementation-oriented responses grounded in the reviewed evidence.

Conclusions

The 13-study corpus reviewed here documents a field in productive transition. AI-enhanced wearable and objective sensing devices can measure circadian fragmentation, derive 24-hour activity phenotypes, estimate circadian light exposure, and predict clinical assessments in older and cognitively vulnerable populations. Current evidence is strongest in institutional settings using actigraphy-derived nonparametric circadian metrics, where fragmentation is most pronounced and most consistently associated with functional outcomes. Circadian disruption, quantified through objective wearable sensing and AI-based analysis, constitutes a biologically plausible and technically measurable candidate biomarker for neurodegeneration. Transitioning this candidate into a validated predictive tool requires prospective longitudinal validation, cross-device harmonization, and transparent XAI approaches that the existing literature does not yet provide. The structural challenges, reference architecture, and layered validation pathway developed from this synthesis provide a conceptual map for researchers and system designers working toward that transition.

Acknowledgments

During the preparation of this manuscript, the authors used Claude (Anthropic) and Grammarly to refine language and improve clarity. An agentic workflow using Claude was also used to produce a first-pass draft of the extraction of computational-method reporting summarized in Table 3, as described in the Methods section; every extracted value was verified independently by 2 authors (JMH-Z and PCS-M) against the source articles and their supplementary materials. The authors reviewed and edited all AI-generated content and take full responsibility for the final version of the manuscript.

Funding

The authors declared no financial support was received for this work.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files (Multimedia Appendices 1-3). The title-abstract and full-text screening records are available from the corresponding author on reasonable request.

Authors' Contributions

All authors contributed equally to the conception and design of this review. The literature search and analysis were performed collaboratively. The manuscript was written by JMH-Z, JM-R, JG-M, and PCS-M, and critically reviewed and edited by all authors. All authors read and approved the final version of the manuscript.

Conflicts of Interest

PCS-M is an associate editor of *JMIR Human Factors* at the time of this publication. All other authors declare no conflicts of interest.

Multimedia Appendix 1
PRISMA checklist.
[PDF File (Adobe PDF File), 254 KB - aging_v9i1e83031_app1.pdf]

Multimedia Appendix 2

XSL•FO
RenderX

Search strategy (PRISMA-S).
[PDF File (Adobe PDF File), 112 KB - aging_v9i1e83031_app2.pdf]

Multimedia Appendix 3
Mixed Methods Appraisal Tool.
[XLSX File (Microsoft Excel File), 15 KB - aging_v9i1e83031_app3.xlsx]

References

1. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022;7(2):e105-e125 [FREE Full text] [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
2. Hood S, Amir S. Neurodegeneration and the circadian clock. Front Aging Neurosci 2017;9:170 [FREE Full text] [doi: 10.3389/fnagi.2017.00170] [Medline: 28611660]
3. Aubourg T, Demongeot J, Provost H, Vuillerme N. Exploitation of outgoing and incoming telephone calls in the context of circadian rhythms of social activity among elderly people: observational descriptive study. JMIR Mhealth Uhealth 2020;8(11):e13535 [FREE Full text] [doi: 10.2196/13535] [Medline: 33242018]
4. Dose B, Yalçin M, Dries SPM, Relógio A. TimeTeller for timing health: the potential of circadian medicine to improve performance, prevent disease and optimize treatment. Front Digit Health 2023;5:1157654 [FREE Full text] [doi: 10.3389/fdgth.2023.1157654] [Medline: 37153516]
5. Fifel K, Videnovic A. Circadian and sleep dysfunctions in neurodegenerative disorders—an update. Front Neurosci 2020;14:627330 [FREE Full text] [doi: 10.3389/fnins.2020.627330] [Medline: 33536872]
6. Dhahbi W, Chamari K, Dergaa I, Trabelsi K, BaHammam AS, Ben Saad H. Sleep under fire: towards a tripartite circadian-neuroendocrine framework for conflict-zone athletes. Br J Sports Med 2026;60(19):1453-1456. [doi: 10.1136/bjsports-2026-111904] [Medline: 42242882]
7. Reddy A, Reddy RP, Roghani AK, Garcia RI, Khemka S, Pattoor V, et al. Artificial intelligence in parkinson's disease: early detection and diagnostic advancements. Ageing Res Rev 2024;99:102410. [doi: 10.1016/j.arr.2024.102410] [Medline: 38972602]
8. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
9. Herrera JM, Monroy-Rodriguez J, Santana-Mancilla PC. Artificial intelligence-based monitoring devices for the identification of circadian parameters as digital predictive biomarkers of neurodegenerative diseases in older adults. Zenodo. 2025. URL: https://zenodo.org/records/16883416 [accessed 2025-11-10]
10. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, PRISMA-S Group. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev 2021;10(1):39 [FREE Full text] [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
11. Wei J, Boger J. Sleep detection for younger adults, healthy older adults, and older adults living with dementia using wrist temperature and actigraphy: prototype testing and case study analysis. JMIR Mhealth Uhealth 2021;9(6):e26462 [FREE Full text] [doi: 10.2196/26462] [Medline: 34061038]
12. Kuula L, Gradisar M, Martinmäki K, Richardson C, Bonnar D, Bartel K, et al. Using big data to explore worldwide trends in objective sleep in the transition to adulthood. Sleep Med 2019;62:69-76. [doi: 10.1016/j.sleep.2019.07.024] [Medline: 31563008]
13. Cai R, Chao J, Gao C, Gao L, Hu K, Li P. Association between sleep duration and cognitive frailty in older Chinese adults: prospective cohort study. JMIR Aging 2025;8:e65183 [FREE Full text] [doi: 10.2196/65183] [Medline: 40267503]
14. Tsanas A, Woodward E, Ehlers A. Objective characterization of activity, sleep, and circadian rhythm patterns using a wrist-worn actigraphy sensor: insights into posttraumatic stress disorder. JMIR Mhealth Uhealth 2020;8(4):e14306 [FREE Full text] [doi: 10.2196/14306] [Medline: 32310142]
15. Hu B, Dixon PC, Jacobs JV, Dennerlein JT, Schiffman JM. Machine learning algorithms based on signals from a single wearable inertial sensor can detect surface- and age-related differences in walking. J Biomech 2018;71:37-42. [doi: 10.1016/j.jbiomech.2018.01.005] [Medline: 29452755]
16. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024;385:e078378 [FREE Full text] [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
17. Tejani AS, Klontzas ME, Gatti AA, Mongan JT, Moy L, Park SH, et al. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 update. Radiol Artif Intell 2024;6(4):e240300 [FREE Full text] [doi: 10.1148/ryai.240300] [Medline: 38809149]
18. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med 2022;28(5):924-933. [doi: 10.1038/s41591-022-01772-9] [Medline: 35585198]

XSL•FO
RenderX

19. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AICONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020;26(9):1364-1374 [FREE Full text] [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]
20. Ben Saad H, Dergaa I, Ghouili H, Ceylan H, Chamari K, Dhahbi W. The assisted technology dilemma: a reflection on AI chatbots use and risks while reshaping the peer review process in scientific research. AI Soc 2025;40(7):5649-5656. [doi: 10.1007/s00146-025-02299-6]
21. de Zambotti M, Claudatos S, Inkelis S, Colrain IM, Baker FC. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chronobiol Int 2015;32(7):1024-1028 [FREE Full text] [doi: 10.3109/07420528.2015.1054395] [Medline: 26158542]
22. Robben S, Englebienne G, Krose B. Delta features from ambient sensor data are good predictors of change in functional health. IEEE J Biomed Health Inform 2017;21(4):986-993. [doi: 10.1109/JBHI.2016.2593980] [Medline: 27455530]
23. Merilahti J, Viramo P, Korhonen I. Wearable monitoring of physical functioning and disability changes, circadian rhythms and sleep patterns in nursing home residents. IEEE J Biomed Health Inform 2016;20(3):856-864. [doi: 10.1109/JBHI.2015.2420680] [Medline: 25861091]
24. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
25. Chuang H, Lin C, Lee L, Chang H, She G, Lin Y. Comparing human-smartphone interactions and actigraphy measurements for circadian rhythm stability and adiposity: algorithm development and validation study. J Med Internet Res 2024;26:e50149 [FREE Full text] [doi: 10.2196/50149] [Medline: 38838328]
26. Razjouyan J, Lee H, Parthasarathy S, Mohler J, Sharafkhaneh A, Najafi B. Improving sleep quality assessment using wearable sensors by including information from postural/sleep position changes and body acceleration: a comparison of chest-worn sensors, wrist actigraphy, and polysomnography. J Clin Sleep Med 2017;13(11):1301-1310 [FREE Full text] [doi: 10.5664/jcsm.6802] [Medline: 28992827]
27. Gabb VG, Blackman J, Morrison H, Li H, Kendrick A, Turner N, et al. Longitudinal remote sleep and cognitive research in older adults with mild cognitive impairment and dementia: prospective feasibility cohort study. JMIR Aging 2025;8:e72824 [FREE Full text] [doi: 10.2196/72824] [Medline: 40435500]
28. Cavaillès C, Wallace M, Leng Y, Stone KL, Ancoli-Israel S, Yaffe K. Multidimensional sleep profiles via machine learning and risk of dementia and cardiovascular disease. Commun Med (Lond) 2025;5(1):306 [FREE Full text] [doi: 10.1038/s43856-025-01019-x] [Medline: 40696147]
29. Wang C, Li C, Liang S. The lifestyle of new middle-aged and older adults in Taiwan described by wearable device: age and gender differences. Eur J Ageing 2024;21(1):27. [doi: 10.1007/s10433-024-00824-y] [Medline: 39340574]
30. Shim J, Fleisch E, Barata F. Wearable-based accelerometer activity profile as digital biomarker of inflammation, biological age, and mortality using hierarchical clustering analysis in NHANES 2011-2014. Sci Rep 2023;13(1):9326 [FREE Full text] [doi: 10.1038/s41598-023-36062-y] [Medline: 37291134]
31. Jones CD, Wasilko R, Zhang G, Stone KL, Gujral S, Rodakowski J, et al. Detecting sleep/wake rhythm disruption related to cognition in older adults with and without mild cognitive impairment using the myRhythmWatch platform: feasibility and correlation study. JMIR Aging 2025;8:e67294 [FREE Full text] [doi: 10.2196/67294] [Medline: 40194310]
32. Beiglary S, Feng Y, Wang N, Ghaeili N, Jao Y, Liao Y, et al. Using wearable sensors to measure and predict personal circadian lighting exposure in nursing home residents: model development and validation. JMIR Aging 2025;8:e72338 [FREE Full text] [doi: 10.2196/72338] [Medline: 41055091]
33. Froggatt K, Best A, Bunn F, Burnside G, Coast J, Dunleavy L, et al. A group intervention to improve quality of life for people with advanced dementia living in care homes: the namaste feasibility cluster RCT. Health Technol Assess 2020;24(6):1-140 [FREE Full text] [doi: 10.3310/hta24060] [Medline: 31971506]
34. Sprint G, Cook DJ, Schmitter-Edgecombe M, Holder LB. Multimodal fusion of smart home and text-based behavior markers for clinical assessment prediction. ACM Trans Comput Healthc 2022;3(4):41 [FREE Full text] [doi: 10.1145/3531231] [Medline: 36381500]
35. Dhahbi W, Jebabli N, Souaifi M, Ceylan H, Ben Saad H, Chamari K, et al. Application of artificial intelligence for predicting sports injuries and customizing personalized prevention strategies: a scoping review. Bioengineering (Basel) 2026;13(6):692 [FREE Full text] [doi: 10.3390/bioengineering13060692] [Medline: 42351936]
36. Lundberg SM, Lee SI. Unified approach to interpreting model predictions. ArXiv. Preprint posted online on November 25, 2017 2017. [doi: 10.48550/arXiv.1705.07874]
37. Ribeiro MT, Singh S, Guestrin C. "Why Should I Trust You?": explaining the predictions of any classifier. 2016 Presented at: KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, CA, United States p. 1135-1144. [doi: 10.1145/2939672.2939778]
38. Dhahbi W, Chamari K. The algorithmic athlete: a call to standardize assessment of sensor technologies and artificial intelligence. Int J Sports Physiol Perform 2026;21(4):505-506 [FREE Full text] [doi: 10.1123/ijspp.2025-0547] [Medline: 41493778]

XSL•FO
RenderX

39. Lim ASP, Yu L, Costa MD, Leurgans SE, Buchman AS, Bennett DA, et al. Increased fragmentation of rest-activity patterns is associated with a characteristic pattern of cognitive impairment in older individuals. Sleep 2012;35(5):633-640B [FREE Full text] [doi: 10.5665/sleep.1820] [Medline: 22547889]
40. Musiek ES, Bhimasani M, Zangrilli MA, Morris JC, Holtzman DM, Ju YS. Circadian rest-activity pattern changes in aging and preclinical alzheimer disease. JAMA Neurol 2018;75(5):582-590 [FREE Full text] [doi: 10.1001/jamaneurol.2017.4719] [Medline: 29379963]
41. Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan. Food & Drug Administration. 2021. URL: https://www.fda.gov/media/145022/download [accessed 2026-10-27]
42. Joshi G, Jain A, Araveeti SR, Adhikari S, Garg H, Bhandari M. FDA-approved artificial intelligence and machine learning (AI/ML)-enabled medical devices: an updated landscape. Electronics 2024 Jan 24;13(3):498. [doi: 10.3390/electronics13030498]
43. Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet 2024 Aug;404(10452):572-628. [doi: 10.1016/s0140-6736(24)01296-0]
44. Finlayson SG, Subbaswamy A, Singh K, Bowers J, Kupke A, Zittrain J, et al. The clinician and dataset shift in artificial intelligence. N Engl J Med 2021 Jul 15;385(3):283-286 [FREE Full text] [doi: 10.1056/NEJMc2104626] [Medline: 34260843]

Abbreviations

AD: Alzheimer disease
CLAIM: Checklist for Artificial Intelligence in Medical Imaging
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-AI: Consolidated Standards of Reporting Trials-Artificial Intelligence
DECIDE-AI: Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence
EEG: electroencephalography
IS: interdaily stability
IV: intradaily variability
LIME: Local Interpretable Model-Agnostic Explanations
MESOR: midline estimating statistic of rhythm
ML: machine learning
MMAT: Mixed Methods Appraisal Tool
MrOS: Osteoporotic Fractures in Men Study
PD: Parkinson disease
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension
RA: relative amplitude
RCT: randomized controlled trial
SHAP: Shapley Additive Explanations
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TIDieR: Template for Intervention Description and Replication
TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis+Artificial Intelligence
XAI: explainable AI

Edited by N Schütz; submitted 26.Aug.2025; peer-reviewed by S Mitra, W Dhahbi; comments to author 20.Nov.2025; revised version received 21.Sep.2026; accepted 19.Oct.2026; published 20.Nov.2026.

Please cite as:
Herrera-Zamora JM, Monroy-Rodriguez J, Garcia-Mancilla J, Santana-Mancilla PC
AI-Based Monitoring Devices for the Identification of Circadian Parameters as Digital Predictive Biomarkers of Neurodegenerative Diseases in Older Adults: Systematic Review
JMIR Aging 2026;9:e83031
URL: https://aging.jmir.org/2026/1/e83031
doi:10.2196/83031
PMID:

XSL•FO
RenderX

©J Manuel Herrera-Zamora, Jesus Monroy-Rodriguez, Jesus Garcia-Mancilla, Pedro C Santana-Mancilla. Originally published in JMIR Aging (https://aging.jmir.org), 20.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Rethinking the Prevention of Cognitive Decline in Healthy Older Adults: The Motor-Cognitive Active Gamified Immersive Cultural (MAGIC) Exergames

Jean-Jacques Temprado, PhD

Institute of Movement Science, Aix-Marseille Université, 163 Avenue de Luminy - BP 910, Marseille, Provence-Alpes-Côte d'Azur, France

Corresponding Author:
Jean-Jacques Temprado, PhD
Institute of Movement Science, Aix-Marseille Université, 163 Avenue de Luminy - BP 910, Marseille, Provence-Alpes-Côte d'Azur, France

Abstract

Though exergames attract considerable research interest as tools for preserving cognitive function in older adults, a decade of meta-analytic evidence reveals a persistent gap between theoretical promise and empirical demonstration: exergames rarely outperform conventional motor-cognitive training and do not generate the sustained engagement that preventive benefit requires. These results reflect limitations that are structural: the field focused predominantly on clinical populations rather than primary prevention; training concepts borrowed from rehabilitation protocols rather than grounded in the specific neuroplastic mechanisms of the exergaming medium; and virtual environments that are digitally sophisticated but culturally and esthetically neutral, systematically excluding the stimulation that converging evidence from neuroaesthetics, cultural epidemiology, and narrative neuroscience identifies as a core active ingredient of brain health promotion. To address these limitations simultaneously, this viewpoint introduces Motor-Cognitive Active Gamified Immersive Cultural (MAGIC) exergames, defined as "virtual reality solutions in which motor-cognitive interactions are embedded in artistic content, cultural heritage, or historically situated narrative to exploit optimal neurobiological substrates of the intervention." Drawing on an integrated theoretical framework, we develop mutually reinforcing mechanistic pillars: multimodal sensorimotor engagement within ecologically valid immersive environments; semantic and narrative activation of memory systems; emotional and autobiographical resonance as a dopaminergic driver of learning consolidation and sustained participation; and the re-engineering of physical effort evaluation from aversive to appetitive through aesthetic pleasure and narrative absorption. A fundamental implication is that in MAGIC exergames, motivation is not engineered through gamification mechanics—it is released by the intrinsic value of inhabiting a culturally meaningful world. We further argue MAGIC exergames correspond to a specific use that is, Cultural Snacking, which consists of the serialization of MAGIC content into brief, narratively open daily episodes of 5 to 10 minutes, designed for 3 to 5 daily consumptions across varied contexts. This dosing architecture is compatible with the motivational profile of healthy older adults who do not identify as patients or trainees. The paper opens perspectives with a design and research agenda structured around three priorities: characterizing MAGIC exergames by the type of motor-cognitive interactions they deliver rather than by the technology platform; conducting mechanistic neuroimaging investigations targeting the specific biomarkers of neuroplastic benefit that creative cultural engagement is expected to produce; and building the intersectoral innovation ecosystem that brings cultural institutions, digital health teams, and older adults together as co-designers. Without waiting for MAGIC-specific trials, we invite the digital medicine, public health, and silver economy communities to recognize that museums, heritage sites, and cultural institutions are potentially the most ecologically valid, intrinsically motivating, and cost-effective substrates for a new generation of brain health exergames that are as meaningful as they are effective. The scientific, institutional, and economic conditions for this convergence already exist.

(JMIR Aging 2026;9:e100013) doi:10.2196/100013

KEYWORDS

aging; exergames; cognition; neuroaesthetics; exercise snacking; immersive cultural experience

Exergames, Brain Health, and Cognition: A Promising Yet Unfulfilled Agenda

As populations age at unprecedented rates, primary prevention of cognitive decline and neurocognitive disorders has become a strategic priority of public health systems worldwide, with the aim of preserving the brain health of healthy older adults before clinical thresholds are reached [1]. Older adults are acutely aware of the prospect of cognitive decline, but evidence-based options available to them remain largely confined to aerobic and muscular exercise programs [2], that many people find insufficiently motivating to sustain. This

XSL•FO
RenderX

adherence problem is structural. Devices developed within a clinical or rehabilitation context benefit from externally supported adherence: medical prescription, professional supervision, a therapeutic framework, and an explicit recovery goal collectively sustain engagement, so that motivation does not need to be intrinsically generated by the intervention itself. This support structure is absent in primary prevention. Among healthy individuals who present no symptoms, face no institutional obligation, and answer to no supervisor, adherence can rest on nothing but intrinsic interest in the activity. This contextual shift is fundamental, and it constitutes, we argue, the central justification for the Motor-Cognitive Active Gamified Immersive Cultural (MAGIC) framework developed in this paper: designing a desirable, culturally meaningful environment that fits naturally into daily life is not a cosmetic optimization of an exergame intervention—it is the condition of its preventive effectiveness.

Over the past decade, exergames—that is, virtual reality solutions that combine physical activity with cognitive challenge within a gamified interface—have been considered as a promising solution for preventing, delaying, or compensating (at least partially) brain alterations and decrease in cognitive performance. By simultaneously delivering cardiovascular effort (endurance), complex goal-oriented movements, and cognitive stimulation in gamified virtual environments, they should, in theory, be more motivating and produce synergistic neuroplastic effects superior to single-domain interventions. Meta-analyses and systematic reviews have reported benefits in executive function, memory, and processing speed [3-5], yet a more rigorous reading of the cumulative evidence reveals a persistent gap between theoretical promise and empirical demonstration [6-9]. For instance, the first and most influential meta-analysis in the field, Stanmore et al [4], reported a moderate positive effect of exergames on global cognition, yet no evidence was found that exergames outperform cognitively demanding control conditions. Subsequent analyses have extended and deepened this assessment. Sokolov et al [7] found no significant advantage of exergames over conventional exercise for cognitive outcomes when head-to-head comparisons were made under controlled conditions, challenging the central claim that the combination of physical and cognitive stimulation in a gamified interface produces synergistic neuroplastic effects. Generally, the meta-analytic literature on exergames and cognition displays a pattern that is characteristic of confirmation bias: effect sizes consistently diminish as methodological rigor increases, and the findings most frequently cited as proof of concept contain, upon closer reading, the seeds of more uncertain conclusions. This conclusion is compounded by a well-documented publication bias, that is, selective reporting of positive findings; once publication bias is corrected for, the apparent cognitive benefits are substantially attenuated [6].

However, despite these conclusions, the field is maintaining a firm optimism about the cognitive benefits of exergaming [10,11], which is a direct consequence of the belief in the magical power of technology and the “product-first” approach, testing specific commercial platforms designed for entertainment as if the product itself were the intervention, rather than characterizing the type of combined training delivered. In a systematic review specifically designed to characterize the type of combined training delivered by exergames, Torre and Temprado [9] identified that the overwhelming majority of exergame studies do not characterize and control for the actual training content delivered, regardless of the technological interface. By misidentifying the exergame as the unit of analysis rather than the training program it supported, the field has systematically generated a literature too heterogeneous to be informative and too methodologically flawed to support causal inference. This failure has limited the understanding of the mechanisms engaged to enhance brain plasticity and cognitive performance in durable and transferable ways [12,13], so that whether and how exergames can be meaningfully improved remains an open question.

How Can Exergames Better Enhance Brain Health and Cognition in Older Adults?

The design logic of exergames required to better enhance brain health and cognition should overcome the limitations that have constrained exergame research in the discovery of effective neuroplastic principles. Four design recommendations follow directly from this diagnosis. First, new exergames should not limit themselves to a curative objective. Instead, the entire design logic—environments, task demands, and motivational architectures—should be structured for the primary prevention of cognitive decline in cognitively healthy older adults, that is, to preserve and enrich capacities rather than to remediate clinical deficits. This reorientation implies a fundamentally different user model, in which the older adult is not a patient to be treated but a capable individual whose neuroplastic resources are to be maintained and deepened through intrinsically motivating engagement. Throughout this paper, “cognitively healthy older adults” refers operationally to community-dwelling individuals, typically aged 65 to 85 years, without a diagnosed neurocognitive disorder and retaining full autonomy in activities of daily living; this population is distinct from, and considerably larger than, the clinical or prefrail populations that conventional exergame research has predominantly targeted. Specifying outcomes of interest for this population is less straightforward than in rehabilitation, where the therapeutic target and a success criterion can be defined a priori. In primary prevention, the relevant question is not whether performance improves in absolute terms but whether decline relative to an expected age-matched trajectory is slowed or arrested over a multiyear horizon, evaluated against domain-specific cognitive measures and, where available, neuroimaging-derived markers of functional brain aging. We return to this question of outcome specification in the research agenda below. Second, future exergames should be grounded in theoretically specified training concepts that exploit what is genuinely specific to the exergaming medium. Rather than developing motor-cognitive, physical-cognitive, or multidomain training concepts validated against neuroplastic mechanisms, the field has oscillated between 2 equally under-specified approaches: the gamification of existing rehabilitation protocols, principally to increase patient motivation [14], and the adoption of commercial

XSL•FO
RenderX

entertainment exergames [15,16]. However, neither approach deeply addresses the questions that a principled exergaming intervention must answer first: what type of movement, what cognitive demand, and what interaction between the two is neurobiologically productive? A related and underappreciated gap concerns the minimum intensity of physical effort, or alternatively the minimum complexity of motor movements, required to exploit facilitation-guidance mechanisms, which remains poorly specified [17,18]. Indeed, the field has largely imported conventional physical activity recommendation logic without questioning whether this dosing architecture is either optimal or realistic for cognitively healthy older adults whose motivational profile and daily routines are poorly served by the scheduled session model. Third, exergame research should capitalize more systematically on the video game literature, which has accumulated robust evidence that regular practice of action video games produces measurable and transferable benefits on attention, processing speed, executive function, and visuo-spatial cognition. This is a body of evidence that health-oriented exergame researchers have systematically ignored, though it provides both a source of theoretical insight into the mechanisms of game-based cognitive enhancement and a design benchmark against which the specific added value of physical and motor engagement in exergames could be rigorously evaluated. Ignoring it has left the field unable to answer a fundamental question: what does motor engagement specifically contribute, over and above the cognitive stimulation that commercial video games already provide? This question is not merely theoretical. Regularly engaging a cognitive process to maintain or strengthen it is also, at least implicitly, the logic underlying cognitive or "brain" training, whose promises and limitations are well-known regarding generalization of training-related benefits beyond the trained task [19,20]. In the present paper, we posit a new framework relative to this corpus: where brain training relies on the repeated, isolated engagement of a single trained process, MAGIC exergames hypothesize that benefit arises from the simultaneous, multisystem engagement of sensorimotor, semantic, and emotion-valuation networks within a single ecologically valid task. Whether this multisystem engagement yields more generalizable benefit than single-process training remains to be established and constitutes one of the central empirical questions this paper's research agenda is designed to address. Fourth, new exergames should actively exploit the neuroplastic potential of rich cultural, narrative, and autobiographically resonant content rather than designing virtual environments that mimic superficially senior-friendly aesthetics while remaining culturally neutral. The neuro-aesthetic evidence demonstrates that genuine aesthetic and cultural experiences simultaneously recruit sensorimotor, emotion-valuation, and meaning-knowledge neural systems in a configuration that no purposive cognitive task can replicate [21]. The brain clock framework shows that creative engagement delays brain aging in domain-independent ways [22], and the cultural epidemiology literature documents that regular cultural participation predicts slower cognitive decline across decades of longitudinal follow-up [23]. We hypothesize that embedding motor-cognitive training within culturally meaningful, esthetically rich, and narratively coherent environments is more than a cosmetic design choice; it constitutes one of the most promising design decisions available, pending dedicated empirical validation. Since no available exergame considers all these blind spots simultaneously, the principles that make exergaming effective to deliver on its neuroplastic promise remain poorly understood. Thus, the question arises of whether a fundamentally different approach to designing new exergames might succeed where most of the existing ones did not.

Recently, a recommended direction has been to move "beyond just-for-fun exergames" [24], to adopt rigorous step-by-step design processes [25,26], and a logic grounded in scientifically validated training concepts to target the specific mechanisms that underlie, combined training effects [8,13,24]. Though this approach represents a step forward relative to purely entertainment-driven exergames, it often results in libraries of motor-cognitive exercises that mimic therapeutic protocols, simply apply recommendations for conventional physical activity programs, and whose motivational appeal rests almost entirely on the thin scaffolding of gamification. Moreover, the "beyond just for fun approach" converges on the dominance of the facilitation-guidance model of brain-body interaction, either in its neurobiological formulation [27,28] or its more recent motor-cognitive variant that is, the Adaptive Capacity Model (ACM) [29]. Though very heuristic, this model is incomplete since it frames the body as a vehicle for stimulating brain plasticity rather than as the primary site of intelligent engagement with a meaningful world. Also, it considers cognitive training tasks as inputs to be optimized rather than as dimensions of a richer, culturally embedded experience. To address this issue, the present Viewpoint paper introduces a new category of preventive digital intervention and develops the theoretical and empirical foundations for a research and design agenda of these new digital brain-health solutions targeting healthy older adults.

A New Theoretical Foundation: MAGIC Exergames

The three structural blind spots identified above point toward a category of preventive digital interventions that: (1) target cognitively healthy older adults, (2) are grounded in a principled understanding of neuroplastic mechanisms rather than in the replication of existing exercise protocols, and (3) place culturally rich, esthetically meaningful, and narratively immersive content at the center of the experience. We propose to call this category MAGIC exergames and define it as: "a virtual reality solution in which brain-mind-body interactions are included in artistic content, cultural heritage, or historically situated narrative, to exploit the neurobiological, psychological and emotional mechanisms involved in problem-solving under changing conditions and constraints." The theoretical orientation underlying MAGIC exergames draws on two converging frameworks. The first is Bardy Human Movement Intelligence (HMI) [30], and the second is the theory of cognizant action developed by Araújo et al [31] within the Ecological Dynamics tradition [32]. Both frameworks share the common premise that movement is not a delivery mechanism for neuroplastic stimulation that ultimately enhances decomposed and isolated

XSL•FO
RenderX

cognitive processes, but the medium through which biological, psychological, emotional, and sensorimotor processes are integrated in a high-order dialogue between brain, body, and environment. In this view, cognition does not follow from movement; it is expressed through it. This perspective extends the ACM [29], which accounts for the neurocognitive and physiological bases of the relationships between physical activity, complex motor skills, brain plasticity, and cognition. Where the ACM grounds the benefits of motor-cognitive training in evolutionary neuroscience, HMI and Ecological Dynamics frameworks extend this foundation by incorporating the esthetic, cultural, social, and technological dimensions of the environments in which motor behavior takes place [30]. In doing so, a conceptual continuity is established between multidomain training [9], grounded in the evolutionary neuroscience model, and the richer theoretical architecture required to account for goal-directed movement in informationally dense, culturally and aesthetically meaningful environments. This reframing implies that cognition is inherently embodied, situationally grounded, culturally shaped, and inseparable from the emotional and aesthetic dimensions of the environment in which it unfolds—conditions that generic exergame environments, culturally neutral by design, systematically fail to provide. Applied to the exergame context, this suggests that the design of MAGIC exergames must immerse both the cognitive functions to train and the physical exercises to gamify within the richness and meaning of the environment itself: its affordances, its narrative coherence, its aesthetic resonance, and its capacity to invite the kind of fully embodied, intrinsically motivated engagement from which neuroplastic change most reliably emerges. This is precisely what MAGIC exergames are specifically designed to instantiate: the conservation of intentionality in informationally rich and culturally meaningful environments by coupled perception and action.

MAGIC Exergames: Design Considerations

The theoretical framework developed above has direct implications for the design of exergames that go beyond the standard recommendations of exercise and medical science to converge toward the capacity of arts, cultural, and creative engagement to prevent ill health, promote well-being, and support cognitive resilience across the lifespan [33,34]. MAGIC exergames instantiate this notion by making the cultural experience a neurobiological and motivational substrate of brain health prevention. Below, we present the different mechanistic pillars of the MAGIC exergames (Figure 1) and different design principles that make them genuinely distinct from both conventional exergames and existing immersive cultural experiences (Table 1). This distinction matters because cultural content could plausibly act through 2 nonequivalent pathways. The first is a neurobiological one: authentic aesthetic experience is hypothesized to jointly recruit sensorimotor, emotion-valuation, and meaning-knowledge systems in a configuration that purposive cognitive tasks and culturally neutral gamified environments do not reproduce. The second is purely motivational: cultural framing functions as a sophisticated wrapper that overrides the automatic cost evaluation described by the Theory of Effort Minimization in Physical Activity (TEMPA) without engaging any neural mechanism beyond those already recruited by the underlying motor-cognitive task. We hold that MAGIC exergames are designed to engage both pathways simultaneously and that their distinctiveness as a category, rather than a culturally themed variant of existing exergaming, rests on this dual and synergistic engagement. Thus, cultural content is not merely dressing applied over a validated cognitive-motor task to make it more palatable, but is hypothesized to codetermine, jointly with the motor-cognitive interaction itself, the neuroplastic outcome. This is a testable claim, and we return to it in the research agenda below.

XSL•FO
RenderX

Figure 1. The 4 mechanistic pillars of Motor-Cognitive Active Gamified Immersive Cultural (MAGIC) exergames and their convergence toward the preservation of cognitive function in healthy older adults. Pillar 1 (multimodal sensorimotor engagement) draws on the Adaptive Capacity Model and the Human Movement Intelligence framework to specify the types of motor-cognitive interactions most productive for neuroplasticity. Pillar 2 (semantic and narrative activation) engages episodic and semantic memory systems through culturally coherent narrative environments. Pillar 3 (emotional and autobiographical resonance) exploits dopaminergic reinforcement via personally meaningful cultural content. Pillar 4 (motivational architecture) operates through intrinsic motivation and the Theory of Effort Minimization in Physical Activity (TEMPA) bypass, operationalized as Cultural Snacking—brief, narratively open episodes designed for multiple daily consumptions. Together, these 4 pillars converge to support the preservation of cognitive function in healthy older adults through sustained, intrinsically motivated cultural immersion.

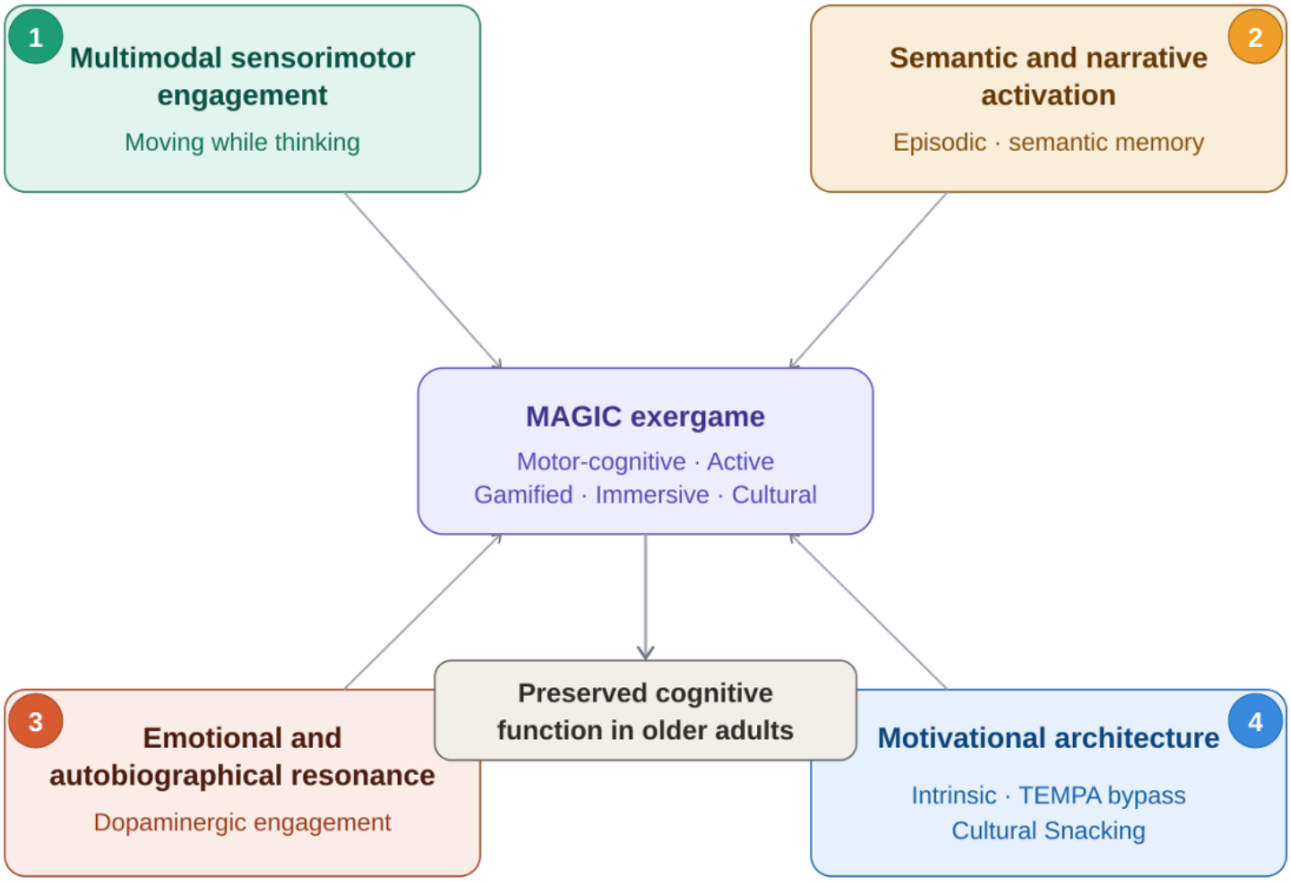


Table . Examples of Motor-Cognitive Active Gamified Immersive Cultural (MAGIC) interactions across the 3 engagement pathways, specified by the type of motor-cognitive interaction they deliver rather than by platform[a].

Engagement pathway	Example MAGIC interaction	Underlying motor-cognitive demand (HMI[b] framework)
Acti-gaming (curiosity-driven)	Wandering a reconstructed 1920s street market to locate a specific vendor's stall, identified only by a half-remembered visual cue from an earlier episode	Incidental locomotion and visual search under low cognitive load; episodic memory retrieval cued by environmental context
Exergaming proper (sustained return)	Restoring a damaged fresco panel by reproducing, with the hand controller, the gesture sequence of the original painter as revealed progressively across sessions	Bimanual coordination and motor sequence learning, with progressive task complexity invisibly tied to narrative deepening across sessions
Exergame-based training (explicit improvement goal)	Guiding a historical figure safely across a collapsing archive under time pressure, requiring rapid target pointing while withholding response to decoy objects	Speed-accuracy constrained pointing (Fitts-type) combined with response inhibition under time pressure—a moving-while-thinking architecture (stop-signal paradigm)

[a]Engagement pathways follow the motivational taxonomy proposed by Manser et al [24]. The HMI framework [31] specifies motor-cognitive demands in terms of perception-action coupling within a structured task environment rather than as isolated cognitive processes. All three examples can be delivered on the same platform (eg, a recumbent-bike exergame with a head-mounted display); the pathway distinction is motivational and design-level, not technological.

[b]HMI: Human Movement Intelligence.

XSL•FO
RenderX

Multimodal Sensorimotor Engagement and Immersive Presence

In the theoretical framework of MAGIC exergames, neuroplastic change emerges from the continuous, adaptive negotiation between a moving body and an informationally rich environment. High-fidelity immersive environments in three-dimensional virtual reality recruit a broader and more ecologically valid range of sensorimotor processes than two-dimensional screen-based exergames. Therefore, the sense of presence becomes a central mechanism that simultaneously engages proprioceptive, vestibular, and peripersonal space systems, creating the conditions for motor-cognitive interaction at a scale that flat-screen interfaces cannot achieve. Moreover, when cultural content structures the immersive environment, guiding participants through a virtual reconstruction of a heritage site, animating historical artifacts, and embedding movement cues within a narrative, the sensorimotor engagement acquires the additional quality of contextual relevance. Actions have meaning within the story: the player does not feel they move to fulfill a training prescription but to navigate a world through goal-directed behaviors whose significance is immediately legible. Such an engagement of motor-cognitive interactions moves beyond classical dual-task paradigms toward what Herold et al [35] term "Moving while Thinking" situations, in which cognitive tasks are incorporated into motor tasks rather than added alongside them. This configuration has been identified as particularly effective for enhancing adaptive behavior in informationally rich environments [17]. This pursuit of a high information rate must be reconciled with the fall risk inherent to physically engaged movement in immersive three-dimensional environments, a concern of particular relevance for healthy older adults. MAGIC design addresses this concern at the hardware and interaction level rather than by reducing informational richness: motor-cognitive bricks involving locomotion or high force production are preferentially delivered in seated or semireclined configurations (as in the recumbent-bike Silver Explorer prototype discussed below), and mixed reality passthrough, rather than fully occluded VR, is favored for daily domestic use because it preserves the user's perceptual awareness of the surrounding physical space throughout the experience. The cascade of formats discussed later in this paper, from supervised location-based installations to lightweight passthrough mixed reality for daily use, is itself a safety architecture as much as an accessibility one.

Semantic, Narrative, and Motor-Cognitive Engagement

In MAGIC exergames, participants navigating a virtual heritage site are not performing a motor task and a cognitive task in parallel; instead, they are inhabiting a world whose exploration is simultaneously sensorimotor, semantic, episodic, and aesthetic. These multimodal, semantically dense, and motor-cognitively demanding environments are lacking in most available exergames, which are focused on gamified exercise libraries rather than meaningful cultural worlds. Several studies have demonstrated that most brain structures vulnerable to aging are those recruited when older adults navigate a culturally rich, narratively coherent environment, which continuously engages episodic memory (recognizing historical objects), semantic memory (retrieving contextual knowledge), spatial navigation (orienting within an immersive scene), and prospective cognition (anticipating what comes next in the narrative) [36,37]. Recently, Coronel-Oliveros et al [22] published the first large-scale demonstration of the link between creative engagement and measurable protection of brain health through delaying functional aging of the involved cerebral networks. It should be noted that Coronel-Oliveros et al [22] documented a dose-dependent effect of active creative practice and expertise (dance, music, visual arts) rather than receptive exposure to a predesigned cultural environment. However, neurobiological evidence for receptive cultural engagement specifically comes from the epidemiological literature. For instance, Tymoszuk et al [38] showed that sustained, repeated engagement with theater, concerts, and museums was associated with improvements in well-being. Moreover, Fancourt et al [39], using epigenetic clocks in 3556 adults from a nationally representative UK cohort, provided the first biological evidence that arts and cultural engagement, including both participatory and receptive activities (attending performances, visiting heritage sites, reading, making music), is associated with a 2% to 4% slower pace of biological aging, in a dose-dependent fashion comparable in magnitude to physical activity. These studies directly support the hypothesis that receptive cultural immersion, of the kind delivered by MAGIC exergames, engages neurobiological mechanisms productive for healthy aging, independently of active creative practice. The ACM provides the complementary neurobiological rationale to the potential benefits of narrative, culturally, and esthetically rich environments for brain and cognition. It proposes that the simultaneous engagement of complex motor skills, physical activity, and higher-order cognitive functions in adaptive, goal-directed motor behavior activates neuroplastic mechanisms more effectively than either domain alone. Based on this model, motor-cognitive exergaming, in which complex motor sequences are simultaneously combined with incorporated cognitive demands (moving while thinking), is hypothesized to produce superior cognitive benefits to either physical-cognitive dual-task arrangements or single-domain training, even when the intensity of physical activity is low [9,17,35].

Emotional and Autobiographical Resonance

Cultural content anchored in the biographical and historical world of current older adults carries a neurobiological asset that purposive exergames entirely lack, that is, autobiographical resonance. Music from one's youth, heritage sites from one's region, artworks of personal significance—these stimuli engage the emotional-valuation system with an intensity that generic game content cannot match. According to the triad model proposed by Chatterjee and Vartanian [19], aesthetic experience simultaneously activates sensorimotor, emotion-valuation, and meaning-knowledge neural systems. Moreover, the physiological and affective "chills" triggered by music or visual art are presumably associated with dopaminergic [36], creating a

XSL•FO
RenderX

pleasure signal that the TEMPA framework identifies as the most powerful predictor of motivation and sustained participation [37]. This emotional engagement consolidates learning, potentiates synaptic plasticity, and biases future behavior toward re-engagement. A MAGIC exergame that generates genuine aesthetic emotion is assumed to create, with each session, an automatic motivational trace that draws older adults to a self-reinforcing dynamic that gamified training protocols have consistently failed to produce.

Pleasure, Effort Minimization, and Motivational Architecture

Designing for sustained engagement in MAGIC exergames therefore requires a motivational framework that operates simultaneously at 2 levels: the reflective level of conscious goal pursuit and the automatic, prereflective level that determines whether physical activity is initiated at all. The integration of these 2 frameworks yields the motivational architecture that MAGIC exergames require, in consistency with the classification of exergaming motivational dynamics proposed by Manser et al [24]. Self-Determination Theory [40], applied to MAGIC exergames, generates useful design principles—player choice to support autonomy, adaptive difficulty to support competence, cooperative or social modes to support relatedness—which are associated with higher deliberative engagement and internalization of motivation [24]. The relatedness dimension deserves further development than the present model currently affords, since the MAGIC framework as developed above focuses primarily on the individual's interaction with the cultural environment, while social engagement is independently established as protective against cognitive decline. Plausible avenues for incorporating a social dimension without displacing the individually paced, narratively open structure of Cultural Snacking include asynchronous shared exploration of the same cultural world, in which players leave discoverable traces or annotations for others who visit later, and noncompetitive collaborative narrative milestones reached jointly across a small group's cumulative individual sessions. These remain design proposals rather than validated mechanisms and are noted here as a direction for future elaboration of the model. However, this framework, which operates exclusively within the reflective motivational system, cannot account for the automatic barrier that precedes deliberate engagement. This limitation is addressed by the TEMPA [41,42], which demonstrates that movement-related cues are automatically processed as costly, that is, as an aversive signal that the brain evaluates and acts upon before any deliberate motivational reasoning is engaged. Exergames whose training logic is perceptually transparent are particularly vulnerable to TEMPA dynamics. Cultural immersion offers a fundamentally different architecture in this respect: a senior navigating a virtual heritage site is not experiencing effort; they are living an adventure. In the vocabulary of Csikszentmihalyi's flow theory [43], the challenge-skill balance is naturally calibrated by the cultural content itself, and the loss of self-consciousness that characterizes flow is precisely the condition under which automatic effort minimization is most effectively suspended. Converging evidence confirms that pleasure experienced during physical activity, not in anticipation, not afterward, is the strongest predictor of long-term adherence [41] and that individualized prescriptions targeting moment-to-moment affective valence generate higher adherence rates than standard prescriptions at 8-week follow-up [42]. In the language of TEMPA, cultural immersion does not merely motivate; it restructures the automatic cost-benefit calculation that determines whether the activity happens at all.

Thus, MAGIC exergames are distinct from conventional approaches to exergames in which motivation is engineered through gamification, considered as an extrinsic incentive architecture superimposed upon a training activity to make it more tolerable. In MAGIC exergames, gamification is not the solution to the motivational problem. Instead, the cultural experience, the narrative world, the aesthetic resonance, and the autobiographical significance of the content generate intrinsic motivation that requires no external scaffolding, because inhabiting a world that matters is already what the player desires.

The heterogeneity of motivational profiles among healthy older adults requires that this integrated framework be operationalized across three distinct engagement pathways, each corresponding to a different relationship between the player and exergaming, and each instantiating a different mode of engagement [22] (see Table 1). The first pathway concerns players motivated primarily by the intrinsic pleasure of cultural exploration and narrative discovery, rather than by any explicit improvement goal. Unlike in conventional exergames, where gamification mechanics (points, rewards, levels) provide the motivational scaffold, in MAGIC exergames the cultural experience constitutes the primary motivational substrate. Accordingly, engagement corresponds to *acti-gaming* [24], that is, occasional, curiosity-driven interaction in which physical and cognitive activity is a natural consequence of cultural immersion rather than a deliberate training investment. The design priority for this pathway is brief, narratively open episodes that generate voluntary return through incompleteness and aesthetic resonance rather than through reward accumulation. Attention is absorbed by the cultural world, and the physical activity embedded within it is experienced not as effort but as the natural condition of inhabiting an experience the player actively desires. The second pathway concerns players who return regularly not to improve but to sustain their capacities through the pleasure of cultural immersion, which could be termed "cultural discovering while exercising." Here the distinction from classical exergames is sharp: the player is not returning because a gamification system incentivizes them to do so, but because an unresolved narrative, an esthetically compelling environment, or an autobiographically resonant cultural world draws them back intrinsically. Their engagement corresponds to *exergaming proper*. Because these players are not motivated by explicit progress, failure must be rare and brief, and progression must remain invisible as training load, experienced instead as deepening cultural and aesthetic richness. The design priority is narrative coherence and continuity across sessions. The cultural world must persist, evolve subtly between episodes, and generate the sense that something awaits one's return. The third pathway concerns players who are explicitly motivated by improvement of

XSL•FO
RenderX

functional capacities and who seek in the MAGIC exergame an opportunity in this respect. Even for these players, the MAGIC framework proposes that the challenge structure be embedded within the cultural narrative rather than superimposed upon it as an explicit training layer. Challenges are presented as culturally situated opportunities that require precise motor sequences, demand spatial navigation under cognitive load, or call for sustained attentional engagement, so that the training logic remains legible to those who seek it while remaining transparent to those who do not. Their engagement corresponds to *exergame-based training* [22]. For these players, gamification mechanics may play a supporting role, but the primary motivational substrate remains the cultural experience and the narrative world—with gamification serving as a complementary layer rather than the foundational architecture.

Rethinking the Use of Exergames: Density, Intensity, and Cultural Snacking

To ensure their effectiveness in preserving brain health and cognition, a critical design challenge for MAGIC exergames concerns the translation of principles of exercise sciences into a format compatible with the lived reality of healthy older adults. Sustained neuroplastic benefit requires repetition, progressive challenge, and sufficient accumulated volume. Immersive cultural experiences are usually delivered as location-based events, typically designed as singular, unrepeatable events. Therefore, the question is how to make exergame formats more culturally appealing and how to reconceive the architecture of engagement itself. Two converging bodies of evidence may help to reframe this. First, Herold et al [44] demonstrated that the standard Frequency, Intensity, Time, Type framework is insufficient to characterize the dose-response relationship between physical activity and brain health, and that *density*—the temporal distribution of activity bouts across the day—must be treated as an independent and potentially decisive variable. Second, the exercise snacking literature shows that brief, intermittent bouts of physical and cognitive activity distributed across the day produce cognitive benefits at least comparable to, and in some respects superior to, equivalent continuous volumes, with an optimal accumulated daily duration of approximately 90 minutes best achieved through multiple short bouts rather than extended single sessions. These findings are directly transformative for MAGIC design. Rather than architecting sessions around conventional exercise prescriptions, MAGIC exergames should be conceived as high-density, narratively serialized micro-episodes whose temporal distribution and cultural continuity across the day become active design variables.

Regarding intensity, the conventional assumption that brain plasticity and cognitive benefits require moderate-to-vigorous aerobic effort is substantially qualified when movement occurs in informationally rich environments with complex motor-cognitive demands. Carrubba et al [17] demonstrated that even moderate- or low-intensity physical activity produces neuroplastic effects comparable to high-intensity training when embedded in rich environments, because it is the continuous adaptive negotiation between a moving body and a demanding environment, not cardiovascular intensity per se, that drives the neuroplastic response. What MAGIC exergames require is not a high heart rate but a high information rate. This opens MAGIC to the full spectrum of mobility and fitness levels characteristic of healthy older adults, without requiring the aerobic effort that purposive exergames demand and that TEMPA dynamics make so difficult to sustain. The design principle that synthesizes both insights is what we term *Cultural Snacking*, which consists of the serialization of MAGIC content into autonomous episodes of 5 to 10 minutes, each complete in itself yet narratively open, designed for 3 to 5 daily consumptions across varied contexts—home, residential facility, hotel, workplace—with or without full VR immersion. The mechanism of re-engagement in Cultural Snacking is categorically different from the gamification logic that underlies conventional exergame design. In gamified exergames, re-engagement is engineered through extrinsic incentive structures—point accumulation, streak rewards, leaderboard positioning, unlockable content—that create artificial obligations to return. Cultural Snacking operates on an entirely different motivational logic. Re-engagement is driven by narrative incompleteness and aesthetic absorption to continue a story that matters to players—a heritage site not yet fully explored, a historical figure whose motives remain unresolved, a cultural object whose significance has only been partially revealed. This is intrinsic motivation in its purest form; it requires no external scaffolding because the activity itself is the reward. It is precisely the motivation that cultural content is uniquely positioned to generate in an older adult population for whom personal history, aesthetic experience, and cultural resonance carry a depth of meaning that points, badges, and leaderboards cannot replicate. Critically, progression in Cultural Snacking does not occur through the explicit escalation of cognitive task difficulty. Instead, progression operates entirely through the gradual modification of environmental constraints, in the spirit of ecological dynamics: richer spatial configurations, more ambiguous cultural affordances, temporally more demanding motor-cognitive sequences, as the narrative world deepens across episodes. This progression is invisible as training load but experienced as increasing narrative and aesthetic richness. The player does not feel themselves becoming more capable—they feel the world becoming more meaningful. The neuroplastic benefit and the cultural experience are not merely adjacent: they are constitutive of each other. Between episodes, the cultural world persists and subtly evolves, generating the sense that something awaits one's return. In the framework of TEMPA, this mechanism is hypothesized to convert the automatic cost evaluation of physical effort into an automatic invitation. A legitimate concern is whether the repeated transition between daily activities and an immersive device, several times a day, would itself impose a cognitive and logistical load that conflicts with the effort-minimization principle the model otherwise seeks to exploit. We argue that this friction is directly addressed by the choice of delivery hardware rather than by the Cultural Snacking architecture itself: next-generation standalone mixed reality headsets require no external computer or cabling, start through a single gesture, and, in passthrough mode, do not require play-space recalibration between sessions. This remains a design hypothesis rather than an established finding, and we return to it in the

XSL•FO
RenderX

research agenda below, where repeated-engagement friction is identified as a measurable outcome in its own right [44]. The exercise snacking evidence [45] and the Cultural Snacking model establish that MAGIC exergames are not merely more engaging than conventional exergames—they are more neuroplastically effective and more motivationally coherent.

Conclusion and Perspectives

Based on underwhelming research on the hypothesized superiority over conventional alternatives. This Viewpoint paper recommends adopting a fundamentally different conceptual approach to designing MAGIC exergames: (1) targeting cognitively healthy older adults as the primary prevention population, (2) grounding their design in the neurobiological logic of HMI and Ecological Dynamics, and (3) placing culturally meaningful, aesthetically rich, and autobiographically resonant content at the center of the intervention. Moreover, the mechanistic pillars of MAGIC exergames developed in the present paper converge on a single design principle that inverts the classic logic of exergame design by assuming that in MAGIC exergames, pleasure is not a reward added to a training protocol: it is the training protocol. A practical challenge that this approach must confront is scalability: content that is culturally rich and personally meaningful risks becoming prohibitively costly to produce across different regions, languages, and population subgroups if each MAGIC environment is built as a bespoke production. Two factors are likely to mitigate this constraint. First, the motor-cognitive bricks specified by the combined-training taxonomy are largely reusable across narrative settings, so that production cost concentrates in environmental and narrative content rather than in the validated interaction mechanics themselves. Second, partnerships with heritage and cultural institutions, which already hold digitized or digitizable collections, offer a route to amortizing content production beyond what a single research or commercial actor could sustain alone. Whether this is sufficient to keep MAGIC content affordable at population scale remains an open empirical and economic question.

Practically, MAGIC exergames could be conceived as a system of complementary formats—from location-based event to the domestic sphere, organized along the reality-virtuality continuum, each optimal for a different use context, immersion depth, and accessibility level, yet all anchored in the same narrative and cultural world. Large-scale projection environments without headsets offer high aesthetic immersion with zero technological barrier, making them ideal for weekly location-based MAGIC sessions. Augmented reality on tablets and smartphones provides a zero-barrier entry point for the shortest micro-snacks integrated into daily life. Mixed reality passthrough devices represent the most promising format for daily domestic Cultural Snacking: immersive enough to sustain narrative engagement, safe enough to eliminate fall risk, and sufficiently lightweight for repeated daily use. This cascade architecture resolves simultaneously the cybersickness problem, the repetition problem by providing lightweight formats compatible with Cultural Snacking, and the accessibility problem by allowing older adults of very different technological comfort levels to enter the same cultural world through different doors. An advanced existing proof-of-concept for this convergence is *Submersive* (Austin, Texas; projected opening 2027), which is explicitly positioned at the intersection of immersive art, neuroaesthetics, and therapeutic wellness, while integrating wearable biometric sensors to measure in real time how immersive artistic environments affect the human body and brain, feeding data back into iterative space design. *Submersive* thus illustrates, as a project still under development rather than as established evidence, the convergence this paper advocates: a private operator from the entertainment world deliberately integrating the conceptual tools of neuroaesthetics and health research to transform a cultural experience into a wellness intervention. Should it open as planned and meet its stated design objectives, Submersive would provide further commercial validation of this convergence; at present, however, it remains a projected outcome rather than a demonstrated one, a distinction this paper is careful to maintain given its own critique, in earlier sections, of unrealistic optimism in the exergaming field.

What it still lacks to constitute the full MAGIC model is an explicit research agenda. This could be structured around three main recommendations. First, future studies must abandon the "technology-first approach" to characterize MAGIC exergames by the type of combined motor-cognitive interactions they deliver—specifying the physical intensity, motor challenge, cognitive load, and cultural content richness as independent and measurable variables. This requires the development of a dedicated taxonomy that incorporates dimensions of narrative coherence, aesthetic richness, and autobiographical resonance as primary design variables. Second, mechanistic investigations are needed, including neuroimaging to identify longitudinally sensitive biomarkers of preventive efficacy and neuroplastic change that creative cultural engagement is expected to produce. Third, comparative designs with active control groups—conventional exergames, standard combined training, and real-world cultural engagement—with adequate follow-up periods are necessary to establish both the superiority and the persistence of MAGIC-induced benefits. Concrete, measurable outcomes that would distinguish a MAGIC intervention from a standard multidomain exergame in a controlled trial should include functional brain-age indices derived from network-level connectivity measures, of the kind reported by Coronel-Oliveros et al [22]. For instance, electroencephalographic comparison between a MAGIC condition and a content-matched "library" condition presenting the identical cognitive-motor bricks without narrative or cultural framing could isolate the neurophysiological contribution of cultural embedding; spontaneous re-engagement rate could serve as a behavioral proxy for intrinsic adherence, which directly distinguishes the MAGIC motivational model from approaches reliant on extrinsic gamification incentives. These outcomes should also be tracked against the logistical-friction hypothesis raised above regarding repeated daily device use.

The future development of MAGIC exergames requires an intersectoral and multidisciplinary innovation ecosystem. The raw material of MAGIC resides in cultural institutions, while the neurobiological and design knowledge required to translate this content into effective interventions resides in digital health

XSL•FO
RenderX

research teams. Bringing these 2 communities into structured co-design programs, with older adults as co-designers, is the essential first institutional step. Indeed, location-based immersive cultural experiences, which already attract a predominantly senior audience, represent a natural entry point for proof-of-concept studies, combining high-quality cultural content with physical engagement in ecologically valid, nonclinical settings. Thus, without waiting for MAGIC-specific trials, we invite the digital medicine, public health, and silver economy communities to recognize that cultural and artistic content is potentially the most ecologically valid, intrinsically motivating, and cost-effective setting for a new generation of immersive exergaming interventions that are as meaningful as they are effective in enhancing brain health, cognition, and well-being. The scientific, institutional, and economic conditions for this convergence already exist.

Acknowledgments

AI was used to assist with language editing, formatting the references, and creating Figure 1.

Funding

The authors declared that no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Pozzi FE, Remoli G, Tremolizzo L, Appollonio I, Ferrarese C, Cuffaro L. Brain health and cognition in older adults: roadmap and milestones towards the implementation of preventive strategies. Brain Sci 2024 Jan 6;14(1):55. [doi: 10.3390/brainsci14010055] [Medline: 38248270]
2. Dhahbi W, Briki W, Heissel A, et al. Physical activity to counter age-related cognitive decline: benefits of aerobic, resistance, and combined training—a narrative review. Sports Med Open 2025 May 17;11(1):56. [doi: 10.1186/s40798-025-00857-2] [Medline: 40381170]
3. Manser P, Herold F, de Bruin ED. Components of effective exergame-based training to improve cognitive functioning in middle-aged to older adults - a systematic review and meta-analysis. Ageing Res Rev 2024 Aug;99:102385. [doi: 10.1016/j.arr.2024.102385] [Medline: 38914262]
4. Stanmore E, Stubbs B, Vancampfort D, de Bruin ED, Firth J. The effect of active video games on cognitive functioning in clinical and non-clinical populations: a meta-analysis of randomized controlled trials. Neurosci Biobehav Rev 2017 Jul;78:34-43. [doi: 10.1016/j.neubiorev.2017.04.011] [Medline: 28442405]
5. Stojan R, Voelcker-Rehage C. A systematic review on the cognitive benefits and neurophysiological correlates of exergaming in healthy older adults. J Clin Med 2019 May 23;8(5):734. [doi: 10.3390/jcm8050734] [Medline: 31126052]
6. Sala G, Tatlidil KS, Gobet F. Still no evidence that exergames improve cognitive ability: a commentary on Stanmore et al. (2017). Neurosci Biobehav Rev 2021 Apr;123:352-353. [doi: 10.1016/j.neubiorev.2019.11.015] [Medline: 31760047]
7. Sokolov AA, Collignon A, Bieler-Aeschlimann M. Serious video games and virtual reality for prevention and neurorehabilitation of cognitive decline because of aging and neurodegeneration. Curr Opin Neurol 2020 Apr;33(2):239-248. [doi: 10.1097/WCO.0000000000000791] [Medline: 32073439]
8. Temprado JJ. Can exergames be improved to better enhance behavioral adaptability in older adults? An ecological dynamics perspective. Front Aging Neurosci 2021;13:670166. [doi: 10.3389/fnagi.2021.670166] [Medline: 34122047]
9. Torre MM, Temprado JJ. Effects of exergames on brain and cognition in older adults: a review based on a new categorization of combined training intervention. Front Aging Neurosci 2022;14:859715. [doi: 10.3389/fnagi.2022.859715] [Medline: 35431905]
10. Béraud-Peigné N, Maillot P, Perrot A. Rethinking physical activity for active cognitive aging through exergames: an innovative approach. Geriatr Psychol Neuropsychiatr Vieil 2025 Jun 1;23(2):255-265. [doi: 10.1684/pnv.2025.1230] [Medline: 40586178]
11. Perrot A, Maillot P. Factors for optimizing intervention programs for cognition in older adults: the value of exergames. NPJ Aging 2023 Mar 29;9(1):4. [doi: 10.1038/s41514-023-00103-7] [Medline: 36991073]
12. Attoh-Mensah E, Bel JF, El Ayoubi K, Boujut A, Ramanoël S, Perrochon A. Neurobiological effects of exergame interventions in older adults with or without a neurocognitive disorder: a systematic review. Neurobiol Aging 2025 Oct;154:47-61. [doi: 10.1016/j.neurobiolaging.2025.06.009] [Medline: 40609365]
13. Monteiro-Junior RS, Vaghetti CAO, Nascimento OJM, Laks J, Deslandes AC. Exergames: neuroplastic hypothesis about cognitive improvement and biological effects on physical function of institutionalized older persons. Neural Regen Res 2016 Feb;11(2):201-204. [doi: 10.4103/1673-5374.177709] [Medline: 27073355]
14. Maggio MG, Baglio F, Maione R, et al. The overlooked role of exergames in cognitive-motor neurorehabilitation: a systematic review. NPJ Digit Med 2025 Jul 9;8(1):419. [doi: 10.1038/s41746-025-01843-4] [Medline: 40634592]

XSL•FO
RenderX

15. Béraud-Peigné N, Maillot P, Perrot A. The effects of a new immersive multidomain training on cognitive, dual-task and physical functions in older adults. Geroscience 2024 Apr;46(2):1825-1841. [doi: 10.1007/s11357-023-00952-w] [Medline: 37743415]
16. Moret B, Nucci M, Campana G. Effects of exergames on mood and cognition in healthy older adults: a randomized pilot study. Front Psychol 2022;13:1018601. [doi: 10.3389/fpsyg.2022.1018601] [Medline: 36420381]
17. Carrubba C, Torre MM, Langeard A, Temprado JJ. Enhancing cognition in older adults with interactive wall exergames. Sci Rep 2025 Oct 23;15(1):37104. [doi: 10.1038/s41598-025-21060-z] [Medline: 41131194]
18. Müller H, Baumeister J, Bardal EM, Vereijken B, Skjæret-Maroni N. Exergaming in older adults: the effects of game characteristics on brain activity and physical activity. Front Aging Neurosci 2023;15:1143859. [doi: 10.3389/fnagi.2023.1143859] [Medline: 37213536]
19. Simons DJ, Boot WR, Charness N, et al. Do "brain-training" programs work? Psychol Sci Public Interest 2016 Oct;17(3):103-186. [doi: 10.1177/1529100616661983] [Medline: 27697851]
20. Gobet F, Sala G. Cognitive training: a field in search of a phenomenon. Perspect Psychol Sci 2023 Jan;18(1):125-141. [doi: 10.1177/17456916221091830] [Medline: 35939827]
21. Chatterjee A, Vartanian O. Neuroaesthetics. Trends Cogn Sci 2014 Jul;18(7):370-375. [doi: 10.1016/j.tics.2014.03.003] [Medline: 24768244]
22. Coronel-Oliveros C, Migeot J, Lehue F, et al. Creative experiences and brain clocks. Nat Commun 2025 Oct 3;16(1):8336. [doi: 10.1038/s41467-025-64173-9] [Medline: 41044135]
23. Fancourt D, Steptoe A. Cultural engagement predicts changes in cognitive function in older adults over a 10 year period: findings from the English Longitudinal Study of Ageing. Sci Rep 2018 Jul 5;8(1):10226. [doi: 10.1038/s41598-018-28591-8] [Medline: 29977058]
24. Manser P, de Bruin ED, Temprado JJ, Bherer L, Herold F. Beyond "just" fun: the role of exergames in advancing health promotion and disease prevention. Neurosci Biobehav Rev 2025 Sep;176:106260. [doi: 10.1016/j.neubiorev.2025.106260] [Medline: 40518089]
25. Manser P, Hardeman LE, Wallin A, et al. Co-Develop-IT! Unifying methodological guideline for the co-design, development, and evaluation of individually tailored technology-enhanced training and rehabilitation concepts: consensus development study and tutorial. J Med Internet Res 2026 May 22;28:e84163. [doi: 10.2196/84163] [Medline: 42172569]
26. Manser P, de Bruin ED. Making the best out of IT: design and development of exergames for older adults with mild neurocognitive disorder–a methodological paper. Front Aging Neurosci 2021;13:734012. [doi: 10.3389/fnagi.2021.734012] [Medline: 34955806]
27. Kempermann G, Fabel K, Ehninger D, et al. Why and how physical activity promotes experience-induced brain plasticity. Front Neurosci 2010;4:189. [doi: 10.3389/fnins.2010.00189] [Medline: 21151782]
28. Kraft E. Cognitive function, physical activity, and aging: possible biological links and implications for multimodal interventions. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2012;19(1-2):248-263. [doi: 10.1080/13825585.2011.645010] [Medline: 22313174]
29. Raichlen DA, Alexander GE. Adaptive capacity: an evolutionary-neuroscience model linking exercise, cognition, and brain health. Trends Neurosci 2017 Jul;40(7):408-421. [doi: 10.1016/j.tins.2017.05.001] [Medline: 28610948]
30. Bardy BG. Human movement intelligence. Physical Activity 2026 [FREE Full text]
31. Araújo D, Lopes H, Farrokh D, Davids K. The ecological dynamics of cognizant action in sport. Psychol Sport Exerc 2025 Sep;80:102935. [doi: 10.1016/j.psychsport.2025.102935] [Medline: 40619115]
32. Seifert L, Araújo D, Button C, Chow JY, Davids K. Performance, learning, and expertise in sport: an ecological dynamics rationale. In: Nonaka T, Segundo-Ortin M, Romero V, Wagman JB, editors. The Routledge International Handbook of Ecological Psychology: Routledge; 2026:463-478. [doi: 10.4324/9781003529453-34]
33. Beauchet O, Cooper-Brown LA, Hayashi Y, Deveault M, Ho AHY, Launay CP. Health benefits of "Thursdays at the Montreal Museum of Fine Arts": results of a randomized clinical trial. Maturitas 2021 Nov;153:26-32. [doi: 10.1016/j.maturitas.2021.07.013] [Medline: 34654525]
34. Beauchet O, Cooper-Brown LA, Hayashi Y, Deveault M, Launay CP. Improving the mental and physical health of older community-dwellers with a museum participatory art-based activity: results of a multicentre randomized controlled trial. Aging Clin Exp Res 2022 Jul;34(7):1645-1654. [doi: 10.1007/s40520-022-02139-3] [Medline: 35578103]
35. Herold F, Hamacher D, Schega L, Müller NG. Thinking while moving or moving while thinking–concepts of motor-cognitive training for cognitive performance enhancement. Front Aging Neurosci 2018;10:228. [doi: 10.3389/fnagi.2018.00228] [Medline: 30127732]
36. Andrews-Hanna JR, Smallwood J, Spreng RN. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. Ann N Y Acad Sci 2014 May;1316(1):29-52. [doi: 10.1111/nyas.12360] [Medline: 24502540]
37. Smallwood J, Bernhardt BC, Leech R, Bzdok D, Jefferies E, Margulies DS. The default mode network in cognition: a topographical perspective. Nat Rev Neurosci 2021 Aug;22(8):503-513. [doi: 10.1038/s41583-021-00474-4] [Medline: 34226715]

XSL•FO
RenderX

38. Tymoszuk U, Perkins R, Spiro N, Williamon A, Fancourt D. Longitudinal associations between short-term, repeated, and sustained arts engagement and well-being outcomes in older adults. J Gerontol B Psychol Sci Soc Sci 2020 Aug 13;75(7):1609-1619. [doi: 10.1093/geronb/gbz085] [Medline: 31287550]
39. Fancourt D, Masebo L, Finn S, Mak HW, Bu F. Does leisure activity matter for epigenetic aging? Analyses of arts engagement and physical activity in the UK Household Longitudinal Study. Innov Aging 2026;10(6):igag038. [doi: 10.1093/geroni/igag038] [Medline: 42181697]
40. Ryan RM, Deci EL. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness: The Guilford Press; 2017. [doi: 10.1521/978.14625/28806]
41. Cheval B, Boisgontier MP. The theory of effort minimization in physical activity. Exerc Sport Sci Rev 2021 Jul 1;49(3):168-178. [doi: 10.1249/JES.0000000000000252] [Medline: 34112744]
42. Gerber M, Cheval B, Cody R, et al. Psychophysiological foundations of human physical activity behavior and motivation: theories, systems, mechanisms, evolution, and genetics. Physiol Rev 2025 Jul 1;105(3):1213-1290. [doi: 10.1152/physrev.00021.2024] [Medline: 39854639]
43. Csikszentmihalyi M. Flow: The Psychology of Optimal Experience: Harper & Row; 1990.
44. Herold F, Zou L, Theobald P, et al. Beyond FITT: addressing density in understanding the dose-response relationships of physical activity with health-an example based on brain health. Eur J Appl Physiol 2025 Oct;125(10):2679-2707. [doi: 10.1007/s00421-025-05858-3] [Medline: 40569418]
45. Liu X, Xia Y, Gao X, Wang R. Exercise snacks and cognitive health in older adults: mechanisms, implementation, and public health relevance. Front Public Health 2026;14:1850516. [doi: 10.3389/fpubh.2026.1850516]

Abbreviations

ACM: Adaptive Capacity Model
HMI: Human Movement Intelligence
MAGIC: Motor-Cognitive Active Gamified Immersive Cultural
TEMPA: Theory of Effort Minimization in Physical Activity

Edited by AF Abdulai; submitted 01.May.2026; peer-reviewed by D Emma, X Tao; revised version received 24.Jul.2026; accepted 24.Jul.2026; published 30.Sep.2026.

Please cite as:
Temprado JJ
Rethinking the Prevention of Cognitive Decline in Healthy Older Adults: The Motor-Cognitive Active Gamified Immersive Cultural (MAGIC) Exergames
JMIR Aging 2026;9:e100013
URL: https://aging.jmir.org/2026/1/e100013
doi:10.2196/100013

© Jean-Jacques Temprado. Originally published in JMIR Aging (https://aging.jmir.org), 30.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Why Frailty Should Be Assessed in Older Adults Affected by Leprosy and How the Use of Smartphones Could Enhance This Health Monitoring

Luzielma Macedo Gloria[1*], MSc; Josafá Gonçalves Barreto[1*], DSC; João Sérgio de Sousa Oliveira[2*], DSC; Areolino Pena Matos[1*], DSC; Bianca Callegari[1*], DSC; Givago Souza[1*], DSC

[1]Federal University of Pará, Generalíssimo Deodoro Avenue 92, Belém, Brazil

[2]University of Pará State, Belém, Brazil

[*]all authors contributed equally

Corresponding Author:
Givago Souza, DSC
Federal University of Pará, Generalíssimo Deodoro Avenue 92, Belém, Brazil

Abstract

Leprosy is endemic in many low-income countries, affecting mainly the poorest communities. The disease can lead to sensorimotor functional losses that can predispose patients to reduced independence. In older adults, these losses associated with leprosy are added to those already resulting from natural aging and associated comorbidities, which can contribute to the clinical picture of frailty syndrome in these older adults. Early detection of frailty can reduce several serious consequences for these individuals, preventing loss of quality of life and financial costs to patients and the health care system. Sensors built into smartphones have been proposed as low-cost, and highly sensitive tools for detecting motor loss and could help monitor the risk of some frailty features in older people affected by leprosy.

(JMIR Aging 2026;9:e55963) doi:10.2196/55963

KEYWORDS

leprosy; digital health; smartphones; balance control; mobility assessment; digital public health; frailty; sensors; mHealth; gerontology; older adults; aging; monitoring; motor function; technology use; health tools

Introduction

Population aging affects all countries today. While high-income nations have already plateaued, transitional and low-income countries are still undergoing the process [1]. Aging causes physiological and structural changes, resulting in some degree of loss of function and independence [2]. Additionally, various diseases can result in physical impairment in old people [3].

In certain circumstances, older adults may experience physical, psychological, and social challenges that reduce their ability to tolerate stressors and cause a cumulative decline in organic functions, making them susceptible to falls and potential negative consequences, including death [4]. The term used to describe this geriatric condition is frailty or frailty syndrome, which is characterized by three of five phenotypes: weakness, low level of physical activity, unintentional weight loss, slow gait speed, and exhaustion [5].

Transitional countries, such as Brazil and India have an impoverished population affected by numerous endemic diseases. Although frailty syndrome may be common in these populations, it is largely unrelated to these endemic diseases. However, leprosy can be a significant comorbidity factor, which, when combined with frailty, can greatly increase the risk of physical disabilities [6].

In regions in which leprosy is endemic, balance control and mobility are not adequately assessed in primary care services due to the high cost and limited availability of gold-standard assessment tools, which are primarily found in research or referral centers. Evaluation of balance or mobility is often conducted through the use of scales or low-cost clinical tests, which may lack the sensitivity necessary to identify early fall risks [7-9]. New options using portable digital technologies have been suggested and validated for conducting this kind of motor evaluation [10,11].

In this context, we suggest discussing the significance of evaluating frailty syndrome in older people with leprosy as a crucial public health policy. Additionally, we propose smartphones as a digital health intervention that could improve the monitoring of the risk of falls in the overlooked population of older adults with leprosy.

The Assessment of Frailty in Leprosy

Leprosy, caused by the bacterium *Mycobacterium leprae*, or less frequently by the *Mycobacterium lepromatosis*, is an

XSL•FO
RenderX

infectious disease that can result in neuropathies compromising proper somatosensory, visual, and vestibular system functioning, as well as muscle strength [12-14]. Although these functions are altered, balance assessment is not included in the standardized protocols established by the World Health Organization (WHO) [15], which serves as the foundation for national protocols in countries like Brazil.

The WHO disability grading system recommends assessing tactile sensitivity, muscle strength, and vision [15]. However, despite vision, somesthesia, and muscular activity contributing to balance control, individualized assessment of each of these functions does not provide a comprehensive understanding of balance control. Furthermore, evaluations of the vestibular system or other somatosensory modalities including temperature, pain, and proprioception, are not considered. It is crucial to note that there is no direct evaluation of mobility or balance in patients with leprosy.

Considering the widespread lack of assessment for both balance and mobility in patients with leprosy of all age groups, the issue becomes particularly concerning in older adults [6]. This population already experiences natural functional declines that are associated with physiological aging, as well as other comorbidities that are more prevalent as individuals age. These comorbidities, such as diabetes, hypertension, and osteoporosis, add to the functional impairment experienced by older adults.

Assessing frailty in older individuals with leprosy holds paramount importance for several reasons. The identification of frailty in older adults with leprosy would enable the timely detection of potential complications (increased risk of disability, delayed wound healing and ulcer development, increased falls, psychological complications, reduce quality of life, etc), facilitate physical and psychological interventions during milder clinical conditions, and provide health managers with valuable insights, thereby contributing to the formulation of new health policies and health care measures [16,17].

Since frailty specifically causes a decrease in muscle mass and strength, which can significantly affect balance control and mobility in older adults with leprosy, monitoring these functions may aid health managers in identifying those who may be at the risk of developing frailty syndrome.

Tracking Frailty Using Smartphones

There is a growing interest in using smartphones to track frailty in older adult populations, and several studies have explored their potential. Most of the apps used to track fragility used applications with forms to be filled with clinical descriptions of the patients following the criteria to classify the Fried's frailty phenotype [18,19]. Other apps have validated the use of sensors from smartphones (cameras and/or inertial sensors) to quantify functional activities that would be associated with the features of fragility [20,21].

Evaluation of Movement as a Potential Indicator of Functional Losses Related to Frailty

The high cost and limited accessibility of gold-standard movement analysis methods, like force platforms, electromyography, and video capture systems, pose significant barriers among poorer populations. These instruments require highly sensitive sensors and robust construction to ensure accuracy and regular, specialized calibration to maintain reliability, and are accompanied by advanced analytics software, thereby adding to the overall cost. These systems often require controlled environments, reliable power supplies, and ample physical space, which may be unavailable in resource-limited areas. Skilled professionals are necessary to operate and interpret the data from these systems, but such expertise is often lacking in underserved regions. These are often only found in specialized research laboratories, hospitals, or sports centers, limiting access for general populations.

However, smartphones with a multitude of built-in sensors have recently emerged, offering validated low-cost alternatives for a range of motor tests. These devices have the potential to provide high-precision assessments in primary care settings [22].

Specifically, for the evaluation of balance control and mobility, various proposals have demonstrated that the inertial sensors, namely the accelerometer and gyroscope, present in smartphones can extract numerous characteristics that describe these functions [23-25].

For static balance assessment, accelerations in the anteroposterior and mediolateral axes can be used to quantify body oscillations in the time domain and time frequencies. These measurements are comparable to those obtained with force platforms [10]. In the case of dynamic balance assessment, inertial sensors have been used during tests like the sit-to-stand test, which introduce various additional variables beyond the number of sit-and-stand cycles performed within a specific time, as is customary in clinical testing [26]. The assessment of mobility via the Timed Up and Go test solely relies on the time it takes to perform [27]. However, the addition of a smartphone to the test can provide measurements of acceleration during the process of getting up and down from a chair, as well as the time taken to walk to and from the chair, and the angular velocity of turns taken during the test [28].

In addition to balance and mobility tests, performing other motor tests on smartphones can also be of interest for evaluating patients with leprosy. The Finger Tapping Test and reaction time test, which indirectly indicate nerve conduction time, can provide valuable information, as nerve conduction time may be delayed in patients with leprosy [29,30].

Factors to Be Considered When Introducing a Smartphone to Monitor

XSL•FO
RenderX

Motor Functionality in Older People With Leprosy

Many experts discussing the implementation of smartphones in health care may initially assume that the proposal involves using the patient's smartphone to acquire all biological signals. However, for digital health interventions, such as motor assessment of patients with leprosy, it is recommended to use smartphones provided by the health care service. Some issues to consider would be the following: (1) the service can ensure that the device meets the necessary technical requirements and is compatible with the software; (2) few studies have evaluated the reliability of measurements using smartphones in unsupervised motor tests; (3) self-tests using smartphones may not be suitable for potentially fragile patients, as some tests (balance and mobility) require trained personnel to be present to prevent falls; and (4) accessing patient history data for evaluation purposes would be decentralized and challenging for managers. The latter is also a factor that does not indicate that the health care professional's smartphone should be used.

Few smartphone apps include more than one motor test. Most of the apps validated in the literature are still experimental and not available in virtual app stores. Some of the available apps provide files with time series of sensor recordings for offline analyses using computational routines. Furthermore, applications that report the values of the extracted characteristics do not provide normative reference values for interpreting those measurements. Therefore, before any intervention, it is necessary to have a plan for processing and storing the data so that the intervention is not frustrating.

While smartphone-based movement analysis apps offer great promise, these interventions are not without significant challenges. Addressing cost, privacy, digital literacy, regulatory, and systemic barriers will be critical for successful implementation, particularly in low- and middle-income countries (LMICs) and older populations. A well-rounded strategy involving stakeholder collaboration, robust funding mechanisms, and tailored solutions can help overcome these obstacles, ensuring that these tools fulfill their potential for equitable health care delivery.

While smartphones are more affordable than gold-standard equipment, they may still be expensive for underfunded health systems or patients in LMICs. Bulk purchases by governments or institutions might reduce costs, but initial investments could strain already limited budgets. Beyond initial procurement, ongoing costs for repairs, software updates, and network connectivity could be prohibitive. Health systems in LMICs often struggle with basic infrastructure issues (eg, lack of clean water and staff shortages). Allocating resources to digital health tools could be viewed as a lower priority compared to addressing these fundamental needs. Investing in smartphone-based interventions might divert funds from critical areas like vaccine programs, malnutrition initiatives, or essential drug supplies. Governments and stakeholders would need to weigh the benefits of early diagnosis and monitoring of movement disorders against these other pressing needs. Another possibility is looking for other mechanisms for funding such as partnerships with nonprofits or nongovernmental organizations focused on health equity, international aid or grants from organizations like the WHO or international foundations (such as the Gates Foundation), and public-private collaborations where technology companies subsidize costs for social impact.

Storing sensitive health data on cloud servers could expose patients to risks of data breaches or misuse. Many LMICs lack robust data protection laws, creating additional vulnerabilities. Some strategies can be applied to mitigate these vulnerabilities such as the use of encryption for data storage and transmission, storage of data locally on the device where feasible, minimization of cloud reliance, and development and enforcement of clear, transparent privacy policies, thereby ensuring compliance with international standards like General Data Protection Regulation where applicable. Patients need to understand how their data will be used and stored. Consent processes must be clear, culturally appropriate, and account for varying levels of digital literacy.

Many older adults, especially those in LMICs, may lack digital literacy (familiarity with smartphones), have physical barriers (poor vision, tremors, or limited dexterity), distrust in technology, or fear of making mistakes, which might discourage adoption. Developers should consider assistive technologies, such as voice commands or larger, high-contrast displays, and simplified user app interfaces; stakeholders should consider training and support programs for the patients, although they might add to implementation complexity.

The regulatory requirements for medical devices and health apps vary significantly among countries. In LMICs, regulatory frameworks for digital health tools may be underdeveloped or nonexistent. In contrast, high-income countries might impose rigorous standards, thus delaying deployment. Early engagement of local regulatory bodies is necessary to clarify requirements and ensure compliance. Some cultures may have reservations about adopting digital health tools due to mistrust of external technology or ethical concerns. Tailored community engagement and education are crucial.

Despite these challenges, we believe that the introduction of a smartphone for motor assessment as part of a digital health intervention holds potential benefits for all stakeholders involved in caring for patients with leprosy. Health professionals would be able to access objective data on various functional abilities for comparative assessment of therapy effectiveness. Health managers could offer personalized services based on functional characteristics of patient demand and plan health policies tailored to the population served by the service. Patients stand to benefit from enhanced health monitoring, as explained above.

XSL•FO
RenderX

Funding

This work was funded by National Council for Research Development (CNPq/Brazil, No. 431748/2016-0). GS is a CNPq productivity fellow (#309936/2022-5) and is supported by the Programa de Apoio a Produção Qualificada form Federal University of Pará.

Conflicts of Interest

None declared.

References

1. Shetty P. Grey matter: ageing in developing countries. Lancet 2012 Apr 7;379(9823):1285-1287. [doi: 10.1016/s0140-6736(12)60541-8] [Medline: 22489326]
2. Vasconcelos ABS, Marcos-Pardo PJ, Da Silva-Grigoletto ME. Editorial: Aging, personal autonomy and independence. Front Aging 2023;4:1326657. [doi: 10.3389/fragi.2023.1326657] [Medline: 38021399]
3. Karlamangla A, Tinetti M, Guralnik J, Studenski S, Wetle T, Reuben D. Comorbidity in older adults: nosology of impairment, diseases, and conditions. J Gerontol A Biol Sci Med Sci 2007 Mar;62(3):296-300. [doi: 10.1093/gerona/62.3.296] [Medline: 17389727]
4. Kolle AT, Lewis KB, Lalonde M, Backman C. Reversing frailty in older adults: a scoping review. BMC Geriatr 2023 Nov 17;23(1):751. [doi: 10.1186/s12877-023-04309-y] [Medline: 37978444]
5. Santos PHS, dos Santos L, Fernandes MH, Brito TA, Munaro HLR, Carneiro JAO. Factors associated with frailty syndrome in older adults with three- and four-criteria clustering. Geriatric Nursing 2023 May;51:121-128. [doi: 10.1016/j.gerinurse.2023.03.005]
6. Masresha BM, Yesuf KM, Moyehodie YA, Biresaw HB, Mulugeta SS, Addisia GD. Determinant factors of leprosy-related disability; comparison of acceleration failure time and parametric shared frailty models. PLoS ONE 2023;18(4):e0271883. [doi: 10.1371/journal.pone.0271883] [Medline: 37011051]
7. Menaldi SL, Harini M, Nelfidayani N, et al. Functional activity limitation of leprosy cases in an endemic area in Indonesia and recommendations for integrated participation program in society. PLoS Negl Trop Dis 2022 Aug;16(8):e0010646. [doi: 10.1371/journal.pntd.0010646] [Medline: 35984857]
8. Lima CA, Ricci NA, Nogueira EC, Perracini MR. The Berg Balance Scale as a clinical screening tool to predict fall risk in older adults: a systematic review. Physiotherapy 2018 Dec;104(4):383-394. [doi: 10.1016/j.physio.2018.02.002] [Medline: 29945726]
9. Beck Jepsen D, Robinson K, Ogliari G, et al. Predicting falls in older adults: an umbrella review of instruments assessing gait, balance, and functional mobility. BMC Geriatr 2022 Jul 25;22(1):615. [doi: 10.1186/s12877-022-03271-5] [Medline: 35879666]
10. Rodrigues LA, Santos EGR, Santos PSA, et al. Wearable devices and smartphone inertial sensors for static balance assessment: A concurrent validity study in young adult population. J Pers Med 2022 Jun 21;12(7):1019. [doi: 10.3390/jpm12071019] [Medline: 35887516]
11. Mellone S, Tacconi C, Chiari L. Validity of a Smartphone-based instrumented Timed Up and Go. Gait Posture 2012 May;36(1):163-165. [doi: 10.1016/j.gaitpost.2012.02.006] [Medline: 22421189]
12. Awasthi SK, Singh G, Dutta RK, Pahuja OP. Audiovestibular involvement in leprosy. Indian J Lepr 1990;62(4):429-434. [Medline: 2086677]
13. Daniel E, Appavoo R, Chacko S, Ragupathy A, Raju R. Impaired contrast sensitivity among leprosy patients with normal visual acuity. Lepr Rev 2005 Mar;76(1):55-64. [Medline: 15881036]
14. Frade MAC, Bernardes Filho F, Silva CML, et al. Evaluation of altered patterns of tactile sensation in the diagnosis and monitoring of leprosy using the Semmes-Weinstein monofilaments. PLoS ONE 2022;17(8):e0272151. [doi: 10.1371/journal.pone.0272151] [Medline: 35947601]
15. Brandsma JW, Van Brakel WH. WHO disability grading: operational definitions. Lepr Rev 2003 Dec;74(4):366-373. [Medline: 14750582]
16. Chen HJ, Chen PY, Kao CL, Sung WH. Smartphone frailty screening: Development of a quantitative early detection method for the frailty syndrome. J Chin Med Assoc 2020 Nov;83(11):1039-1047. [doi: 10.1097/JCMA.0000000000000409] [Medline: 32773591]
17. Deng Y, Zhang K, Zhu J, Hu X, Liao R. Healthy aging, early screening, and interventions for frailty in the elderly. Biosci Trends 2023 Sep 15;17(4):252-261. [doi: 10.5582/bst.2023.01204] [Medline: 37612123]
18. Aznar-Tortonda V, Palazón-Bru A, la Rosa DD, et al. Detection of frailty in older patients using a mobile app: cross-sectional observational study in primary care. Br J Gen Pract 2020 Jan;70(690):e29-e35. [doi: 10.3399/bjgp19X706577] [Medline: 31685541]
19. Zolnowski-Kolp V, Um Din N, Havreng-Théry C, et al. Assessment of frailty by the French Version of the Vulnerable Elders Survey-13 on Digital Tablet: Validation study. J Med Internet Res 2023 Aug 2;25:e42017. [doi: 10.2196/42017] [Medline: 37531175]

XSL•FO
RenderX

20. Montemurro A, Rodríguez-Juan JJ, Martínez-García MDM, Ruiz-Cárdenas JD. Validity of a video-analysis-based app to detect prefrailty or frailty plus sarcopenia syndromes in community-dwelling older adults: Diagnostic accuracy study. Digit Health 2024;10:20552076241232878. [doi: 10.1177/20552076241232878] [Medline: 38384370]
21. Montemurro A, Ruiz-Cárdenas JD, Martínez-García MDM, Rodríguez-Juan JJ. Validity of an iphone app to detect prefrailty and sarcopenia syndromes in community-dwelling older adults: The protocol for a diagnostic accuracy study. Sensors (Basel) 2022 Aug 11;22(16):6010. [doi: 10.3390/s22166010] [Medline: 36015771]
22. Paradis S, Roussel J, Bosson JL, Kern JB. Use of smartphone health apps among patients aged 18 to 69 years in primary care: Population-based cross-sectional survey. JMIR Form Res 2022 Jun 16;6(6):e34882. [doi: 10.2196/34882] [Medline: 35708744]
23. Rinderknecht MD, Zanon M, Boonstra TA, et al. An observational study to assess validity and reliability of smartphone sensor-based gait and balance assessments in multiple sclerosis: Floodlight GaitLab protocol. Digit Health 2023;9:20552076231205284. [doi: 10.1177/20552076231205284] [Medline: 37868156]
24. Pedrero-Sánchez JF, De-Rosario-Martínez H, Medina-Ripoll E, et al. The reliability and accuracy of a fall risk assessment procedure using mobile smartphone sensors compared with a physiological profile assessment. Sensors (Basel) 2023 Jul 20;23(14):6567. [doi: 10.3390/s23146567] [Medline: 37514860]
25. De Groote F, Vandevyvere S, Vanhevel F, Orban de Xivry JJ. Validation of a smartphone embedded inertial measurement unit for measuring postural stability in older adults. Gait Posture 2021 Feb;84:17-23. [doi: 10.1016/j.gaitpost.2020.11.017] [Medline: 33260077]
26. Ruiz-Cárdenas JD, Montemurro A, Del Mar Martínez-García M, Rodríguez-Juan JJ. Concurrent and discriminant validity and reliability of an Android App to assess time, velocity and power during sit-to-stand test in community-dwelling older adults. Aging Clin Exp Res 2023 Aug;35(8):1631-1640. [doi: 10.1007/s40520-023-02451-6] [Medline: 37306926]
27. Podsiadlo D, Richardson S. The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
28. Ponciano V, Pires IM, Ribeiro FR, Spinsante S. Sensors are capable to help in the measurement of the results of the Timed-Up and Go Test? A systematic review. J Med Syst 2020 Oct 17;44(12). [doi: 10.1007/s10916-020-01666-8] [Medline: 33070247]
29. Brito FAC, Monteiro LCP, Rocha Santos EG, et al. The role of sex and handedness in the performance of the smartphone-based Finger-Tapping Test. PLOS Digit Health 2023 Aug;2(8):e0000304. [doi: 10.1371/journal.pdig.0000304] [Medline: 37585430]
30. Serra-Añó P, Pedrero-Sánchez JF, Hurtado-Abellán J, Inglés M, Espí-López GV, López-Pascual J. Mobility assessment in people with Alzheimer disease using smartphone sensors. J Neuroeng Rehabil 2019 Aug 14;16(1):103. [doi: 10.1186/s12984-019-0576-y] [Medline: 31412893]

Abbreviations

LMIC: low- and middle-income country
WHO: World Health Organization

Edited by R Bjarnadottir; submitted 31.Dec.2023; peer-reviewed by KD Sabbata, MN Shalaby; revised version received 30.Nov.2024; accepted 23.Dec.2024; published 06.Aug.2026.

Please cite as:
Gloria LM, Barreto JG, Oliveira JSDS, Matos AP, Callegari B, Souza G
Why Frailty Should Be Assessed in Older Adults Affected by Leprosy and How the Use of Smartphones Could Enhance This Health Monitoring
JMIR Aging 2026;9:e55963
URL: https://aging.jmir.org/2026/1/e55963
doi:10.2196/55963

© Luzielma Macedo Gloria, Josafá Gonçalves Barreto, João Sérgio de Sousa Oliveira, Areolino Pena Matos, Bianca Callegari, Givago Souza. Originally published in JMIR Aging (https://aging.jmir.org), 6.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging AI to Advance Age-Friendly Care in the Veterans Health Administration

Elizabeth Fine Smilovich[1*], MD; Megha Kalsy[2,3*], PhD; Kimberly Wozneak[4], MS; Quratulain Syed[5], MD; Laurence M Solberg[6], MD, AGSF

[1]Geriatrics Research, Education, and Clinical Center (GRECC), Veterans Affairs Northeast Ohio Healthcare System, 10701 East Boulevard, Cleveland, OH, United States

[2]Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, OH, United States

[3]College of Public Health and Health Professions, University of Florida, Gainesville, FL, United States

[4]Geriatrics and Extended Care, Veterans Health Administration, Washington, DC, United States

[5]Geriatric Research Education and Clinical Center (GRECC), Atlanta VA Medical Center, Atlanta, GA, United States

[6]Geriatrics Research, Education, and Clinical Center (GRECC), NF/SG VHS, Malcom Randall VA Medical Center, Gainesville, FL, United States

[*]these authors contributed equally

Corresponding Author:
Elizabeth Fine Smilovich, MD
Geriatrics Research, Education, and Clinical Center (GRECC), Veterans Affairs Northeast Ohio Healthcare System, 10701 East Boulevard, Cleveland, OH, United States

Abstract

The aging population presents a pressing challenge for health care systems, necessitating effective strategies to address the complex needs of older adults. The Veterans Health Administration in the Department of Veterans Affairs (VA), the largest integrated health care system in the United States, has embraced the Age-Friendly Health Systems (AFHS) initiative from the Institute for Healthcare Improvement to ensure safe and high-quality care for older veterans. In AFHS, health care providers consistently use the evidence-based 4Ms framework (what matters, medication, mentation, and mobility) to deliver comprehensive care for older adults in all care settings. This viewpoint paper explores the potential of artificial intelligence (AI) to enhance the evidence-based implementation of the AFHS 4Ms framework in the VA to provide optimal care for older adults. By leveraging AI technologies, such as natural language processing, machine learning, large language models, clinical decision support, and data analytics, this viewpoint examines the opportunities and challenges of using AI to support the 4Ms domains in a large, integrated health care system. Furthermore, it discusses the potential benefits of integrating AI-driven decision support systems and predictive analytics to personalize care, reduce polypharmacy and potentially inappropriate medications, enhance cognitive and mood assessments, and better identify mobility issues and interventions. By examining the intersection of AI and age-friendly care in the VA, this viewpoint highlights the transformative potential of AI to expand 4Ms care and improve the experience of providers and older adults across diverse health care settings.

(JMIR Aging 2026;9:e75686) doi:10.2196/75686

KEYWORDS

age-friendly health systems; 4Ms framework; aging; artificial intelligence; technology; older adults; geriatrics; Veterans Health Administration

Introduction

The United States is experiencing a significant demographic shift, characterized by a rapidly aging population. By 2050, an estimated 82 million individuals will be aged 65 years or older [1]. The Veterans Health Administration in the Department of Veterans Affairs (VA) is the largest integrated health care system in the United States, and a significant number of older veterans are enrolled in health care provided by the VA [2]. An estimated 30% of adults aged 65 years or older could benefit from specialized care from geriatricians; however, there continues to be a scarcity of trained geriatricians despite innovative solutions to meet the rising demands within the geriatric workforce [3]. To meet the increasingly complex care needs of the growing older veteran population, the VA joined the Age-Friendly Health Systems (AFHS) movement in March 2020 [4]. The AFHS initiative within the VA continues to build on decades of evidence-based care and innovation in geriatrics, which aligns with the 4Ms framework (what matters, medication, mentation, and mobility) [5] and supports the VA's goal of becoming a high reliability organization [4,6-8].

As efforts to train health care professionals to optimize a geriatric-friendly workforce continue, artificial intelligence (AI) can be leveraged to supplement these efforts and support the

XSL•FO
RenderX

geriatric workforce in providing high-quality care to older adults with multiple chronic conditions.

The VA as an Age-Friendly Health System

The AFHS framework emphasizes developing and aligning care plans for older adults to meet the goals that matter most to them. AFHS was established in 2017 by the John A. Hartford Foundation and the Institute for Healthcare Improvement (IHI) in collaboration with the American Hospital Association and the Catholic Health Association of the United States. This initiative promotes a comprehensive framework known as the 4Ms, which focuses care on the "what matters" element to the patient and on assessing medication, mentation, and mobility to ensure evidence-based and reliable care for older adults [7,8].

The VA has made important strides in adopting the 4Ms on an enterprise scale. As of July 2025, the VA has implemented AFHS in 455 clinical sites, with all sites receiving Level 1 Participation in AFHS recognition from the IHI [9]. Level 1 recognition means that the clinical site has described how the 4Ms will be put into practice. Of these sites, 297 have achieved Level 2 Committed to Care Excellence recognition from the IHI. Level 2 recognition means that the clinical site has tracked and shared the number of older adults reached with the 4Ms over at least 3 months and has described how the 4Ms will be sustained in practice. The implementation and spread of AFHS in the VA has reached 164 Veteran Affairs Medical Centers nationwide.

The central focus of the AFHS movement is to prioritize "what matters" to older adults—the first "M"—placing their specific health outcome goals and care preferences at the forefront across all care settings. The VA has developed several ways, through different programs, to elicit this information, such as the Whole Health Mission, Aspiration, Purpose inquiry; the Patient Priorities Care survey; and the Well-Being Signs tool. These tools are accessible through an age-friendly 4Ms template and can be viewed, if previously populated, when the template is opened. Health care providers looking to readily view this information to position the visit in what matters to the patient must open the template. Currently, no function exists to easily elicit from the chart what a veteran's health priorities may be if the template wasn't previously used.

The second "M"—"what medication"—emphasizes limiting the use of high-risk medications and dose adjustments when a high-risk medication is essential to achieve what matters to older adults. Within the VA, there is active groundwork to support safer medication use: VIONE (Vital, Important, Optional, Not indicated), a nationwide deprescribing program with dashboards and risk tools embedded in the electronic health record (EHR), has been disseminated across more than 130 medical centers and supports pharmacist-led deprescribing [10]. VA informatics efforts have also applied natural language processing (NLP) to prescribing problems and are advancing AI- and machine learning (ML)-enabled decision support as part of a learning health system strategy [11]. Software built into the EHR notifies providers of medication interactions; however, currently, systems to review medications using Beer's criteria (which identify medications potentially inappropriate for older adults) or to alert health care providers to medication-disease interactions (eg, anticholinergic medications in patients with cognitive impairment) are not standard and require the expertise of geriatric-trained providers and pharmacists.

The third "M"—"what mentation"—concentrates on preventing, identifying, treating, and managing conditions such as dementia, depression, and delirium across care settings. Current use of the AFHS template with embedded cognitive and mood screening tools has improved access to these evidence-based assessments. Implementation, although progressing, faces challenges related to time and documentation constraints. Although there is a program for mental health clinicians to initiate and review mental health assessments (Mental Health Assistant), there is no repository within the EHR to easily view past screening results and actions; the AFHS template must be opened. Unless a veteran presents for cognitive evaluation to a mental health or geriatric specialist, the current system relies on individual providers to determine the best time to screen for cognitive impairment, often after warning signs for dementia are exhibited. Electronic frailty indices, based on predictive models such as the VA Frailty Index and the Care Assessment Needs score, exist within the VA system [12]; however, they are not used in clinical settings to identify patients at risk of cognitive decline.

The final "M"—"what mobility"—promotes mobility evaluation to encourage older adults to safely engage in daily physical activity to maintain function and pursue what matters to them. Current use of the AFHS template using embedded mobility screening tools prompts health care providers to perform either a history or physical examination-based screen. Real-time remote monitoring of gait or fall risk for ambulatory care patients is not yet part of routine care, and other potential predictors, such as the Care Assessment Needs score, do not correlate well with physical functioning [13].

Individually, each "M" has the potential to improve health outcomes, reduce low-quality services, and increase the use of cost-effective interventions. Introducing the 4Ms approach as part of AFHS improves care for older veterans [4]. To promote the AFHS 4Ms approach and offer better care to the older population, the VA must leverage all tools, including the potentially transformative use of appropriate AI techniques.

AI is a rapidly developing technology that encompasses various fields of knowledge and uses a wide range of techniques. Its primary goal is to simulate and extend human intelligence through machines [14]. Within the field of AI, research areas include expert systems, ML, robotics, decision support systems, and pattern recognition. There are many potential applications of AI in health care for older adults, such as clinical decision support (CDS) systems, robotics for physical rehab and physiotherapy exercises, NLP and speech analysis, medical imaging and video processing, wearable devices and signal processing from digital watches that track health data, and using ML for targeted drug therapy [15,16].

XSL•FO
RenderX

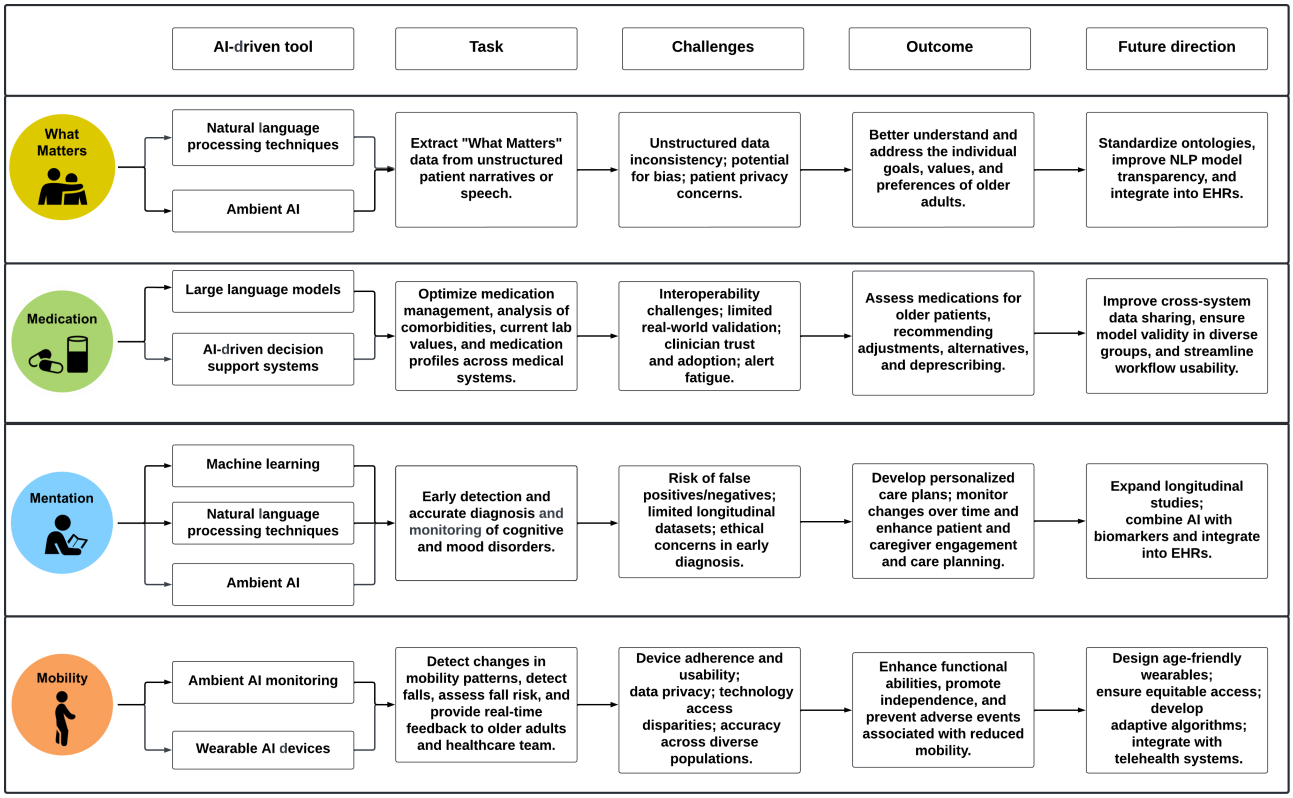
AI has already demonstrated its potential in geriatric care by addressing challenges such as caregiver shortages and resource disparities. Researchers increasingly believe that AI may further enhance the well-being of older adults by redistributing health care resources, including nursing and other professional support. Numerous studies have explored the application of AI in geriatric care, focusing on areas such as Alzheimer disease care, disease recognition, and medication reminders [17].

Our viewpoint is that appropriate use of AI could better use and integrate the 4Ms across the VA. The implementation and sustainment of 4Ms care require scalable and efficient strategies to deliver equitable, evidence-based care to reach all older adults. We propose leveraging AI technologies, such as NLP, ML, ambient AI, CDS, and data analytics to support the AFHS movement. This manuscript delves into the opportunities and challenges in using AI to support the 4Ms domains within the VA system.

Use of AI for 4M-Based Care at the VA

Although AI is rapidly advancing, its application to AFHS and the 4Ms framework has not been systematically characterized. To date, no research has comprehensively assessed the scope, trends, effectiveness, or gaps in AI-driven approaches to AFHS within the VA. This lack of synthesis represents a barrier to advancing implementation and scaling of age-friendly care. By leveraging AI methods such as NLP, ML, predictive modeling, and advanced data analytics, with emerging opportunities in deep learning and large language models (LLMs), there is significant potential to transform large volumes of unstructured and structured clinical data into actionable insights that enhance evidence-based, person-centered care for older veterans within the 4Ms framework. Figure 1 describes the proposed use of AI for age-friendly 4Ms-based care.

Figure 1. Proposed use of artificial intelligence (AI) for age-friendly 4Ms-based care. EHR: electronic health record; NLP: natural language processing.



What Matters

Using AI in health care for older adults can expand the available tools to improve care delivery and assist health care professionals in personalizing medicine for this population. Implementing AI within the context of AFHS, with a focus on what matters to the older adult, may enable more efficient development of care plans that are both personalized and optimized. The AFHS 4Ms framework represents a paradigm shift in the care of older adults by emphasizing alignment of clinical practice guidelines with patients' health goals and lifestyle preferences.

Extraction and compilation of EHR data across health care systems regarding patients' priorities and goals remain challenging. AI tools such as NLP and ML can support clinical decision-making by identifying and synthesizing these priorities, thereby assisting clinical teams in formulating care plans centered on what matters to older adults. NLP, a branch of AI, enables computers to understand, extract, and analyze information from human language in clinical text and documentation. These techniques can extract insights from unstructured patient narratives, including documentation of goals, life-sustaining treatment, and advance care planning, to provide clinicians with a clearer understanding of patients' values and preferences.

Recent work in this area has demonstrated promising applications. For example, Gleason et al [18] developed a rule-based NLP system to classify patient portal messages from

XSL•FO
RenderX

older adults by the 4Ms domains (including the "what matters" domain) and achieved an area under the curve of 0.90 in identifying messages sent by care partners [18]. Similarly, Min et al [19] developed an NLP algorithm to automatically capture daily care experience preferences from narrative data, demonstrating how AI can operationalize the "what matters" domain at scale [19].

Beyond text-based extraction, emerging technologies such as ambient AI (ambient scribing) can capture and synthesize clinical discussions in real time, including conversations around "what matters," to further support AFHS [20]. Integrating these AI-enabled insights into clinical workflows can promote more patient-centered care, facilitate shared decision-making, and improve patient satisfaction.

While some scholars caution that reliance on AI-driven clinical decision algorithms could risk a return to paternalistic models of care, AFHS emphasis on the "what matters" domain provides an important safeguard. By ensuring that AI tools are designed and implemented with patient priorities at the center, AFHS can help expand personalized medicine for older adults in a manner that enhances, rather than replaces, the patient voice.

Medications

In the "medication" domain, AI-enabled CDS can help optimize pharmacotherapy for older adults by screening for high-risk prescribing, drug-drug interactions, renal-dose mismatches, and deprescribing opportunities at the point of care. Contemporary reviews show that ML models can predict adverse drug events from EHR data and that CDS interventions in older populations improve medication safety and deprescribing outcomes [21]. Emerging LLM applications extend this capability. LLMs are advanced AI systems trained on large-scale text data to understand and generate human language. The term "large" refers to the size of the datasets that assist in learning patterns in language. The term "language" reflects the model's ability to understand, summarize, translate, and code text. The term "model" refers to the computational system that learns patterns in data to generate predictions or output. A 2025 scoping review mapped how generative AI and LLMs mitigate medication-related harm across the medication-use process [22] and a cohort study published in *JMIR Aging* demonstrates that an LLM-based pipeline can identify deprescribing opportunities for older adults, while also noting limits in complex clinical scenarios [23].

Modernization of medication reconciliation workflows highlights progress and remaining integration gaps that AI-driven CDS can address. These developments indicate that AI, including NLP, ML, and LLMs, may be potentially used to augment medication safety across the AFHS framework by surfacing risks, proposing safer alternatives and dose adjustments, and flagging candidates for deprescribing—particularly when embedded in workflows and paired with pharmacist and clinician review.

Mentation

AI also holds promise in the "mentation" domain, where timely identification and management of cognitive and mood disorders are critical. ML approaches trained on large datasets can aid early detection and differential diagnosis of cognitive impairment (eg, Alzheimer disease and mild cognitive impairment). NLP applied to clinical notes improves delirium ascertainment that is often undercoded in structured data [24]. In addition, prototype ambient and patient-facing systems such as the SeVA prototype system use artificial emotional intelligence with chatbot-based monitoring and 2 mobile apps to provide timely checks and health status recording; early deployments report feasibility for detecting delirium and supporting nursing workflows [25]. Similarly, a recent study demonstrated the feasibility of using ML to analyze Wi-Fi–based motion sensor data to predict depression among older adults [26]. Together, these AI methods can synthesize cognitive testing, narrative documentation, and real-time signals to inform personalized care plans and track changes over time.

Mobility

AI can strengthen the "mobility" domain by continuously detecting changes in gait and activity, assessing fall risk, and triggering timely interventions. Ambient and wearable sensing, ranging from step counters and actigraphy to contactless in-home systems, can monitor mobility patterns and surface deviations that matter clinically. In real-world living settings for older people, an AI video–based fall detection system reduced time on the ground and accelerated staff response after unwitnessed falls, demonstrating an actionable feedback loop to care teams [27].

In-home smart apartment platforms have similarly generated automated health alerts from passive sensors to notify clinicians of mobility and health changes, supporting earlier evaluation and intervention [28]. Contactless radio frequency sensing has enabled continuous home-based monitoring of gait speed and impairment (eg, in Parkinson disease), offering a path to track functional decline between visits [29].

Complementary evidence from reviews shows that ambient assisted-living and remote passive-sensing technologies can detect falls, changes in activities of daily living, and health status shifts, informing prevention and individualized mobility care plans [30].

The 4Ms as an Integrated Set

We propose implementing an integrated AI approach to the 4Ms. A prompted LLM extraction service will harvest 4Ms signals from structured and free-text EHR sources and patient messages ("what matters" domain preferences, high-risk medications, cognitive or behavioral changes, and falls or balance changes), building on evidence that LLMs can detect postoperative falls from clinical narratives [31]. A VA chat generative pretrained transformer chart reviewer will then compile the 4Ms into a provenance-linked 4Ms summary card, using structured data fields where permissible, and surface care gaps—such as missing goals-of-care documentation, high-risk prescribing, delirium language, or gait decline—as team-routable alerts. To ease documentation burden, ambient AI–assisted drafting and smart templates will prepopulate 4Ms documentation, supported by a previsit 4Ms gaps checklist and 2-click insertion into EHR notes. Success metrics could include

XSL•FO
RenderX

entity-level F_1-scores of at least 0.80 for key 4Ms concepts, note-level recall of at least 0.90 for any 4Ms signal, a 25% or greater reduction in documentation time, a 20% or greater increase in 4Ms completeness, improved usability scores, and demonstrable gains in goal-concordant care and time to action on fall risk and high-risk medications.

Barriers and Challenges to Integrating AI Into AFHS and EHRs

There are numerous challenges to integrating AI into health care for older adults. One consideration is the potential degradation of face-to-face human contact with clinicians and health care professionals resulting from the implementation of these technologies. This could lead to fewer routine caregiver interactions and increased social isolation [15]. Another challenge, as previous studies have demonstrated, is that wearable devices in long-term care environments can be uncomfortable and the sensors may disturb sleep [32]. Considerations such as these must be at the forefront of discussions regarding AI implementation in the care of older adults.

While AI has the potential to support health care professionals in the provision of high-quality health care services for older adults, numerous concerns have been raised in research studies, including bias in AI algorithms, which have the potential to precipitate existing disparities in access to health care resources, and concerns about patient privacy in the setting of possible cybersecurity threats [15].

Implementing AI within the VA to support the evidence-based AFHS 4Ms framework presents multiple challenges and barriers. One major challenge is integrating AI with existing EHRs. EHRs are often complex and fragmented across various platforms, necessitating sophisticated data harmonization strategies to enable AI-driven insights without compromising data integrity [33]. Furthermore, ethical considerations are paramount, as AI systems must ensure patient privacy and confidentiality, consistent with HIPAA (Health Insurance Portability and Accountability Act) regulations, while also being transparent in decision-making processes to gain stakeholders' trust [34]. The health system culture may pose another barrier, as clinicians may resist adopting AI tools due to perceived threats to clinical autonomy, potential disruptions to established workflows, and the learning curve associated with new technologies [35]. Addressing these issues requires a comprehensive strategy that aligns technological advancements with ethical standards and cultural readiness, involving stakeholders at all levels of the VA.

While the integration of AI into AFHS care within the VA presents exciting opportunities that may improve patient experience and allow older people to age in place and live in their homes longer, and decrease cognitive burden on clinicians and health care professionals, several challenges must be addressed. These include ensuring data privacy and security, data harmonization within the EHR, addressing potential biases in AI algorithms, and fostering user acceptance and trust in AI-powered systems. Overcoming these challenges requires interdisciplinary collaboration among health care professionals, data scientists, engineers, and policymakers supporting the AFHS to provide optimal personalized health care to older adults.

By applying a scientific lens to the intersection of AI and age-friendly care, this viewpoint paper has explored the transformative potential of AI in enhancing evidence-based implementation of the 4Ms framework. It highlights the opportunities and challenges in leveraging AI technologies to personalize care, improve medication management, enhance cognitive and mood assessments, and optimize mobility interventions for older adults across diverse health care settings. The findings of this study have implications for health care policymakers, administrators, and researchers seeking to maximize the benefits of AI in delivering high-quality, person-centered care for older adults.

Conclusion

This viewpoint paper addresses the urgent need to develop effective strategies for the use of AI to enhance evidence-based implementation of the AFHS framework within the VA as a model for AI integration in a large, integrated health care system. AI tools such as NLP, ML, and data analytics can be leveraged to support effective implementation of 4Ms-based care across the VA health care system. Integrating AI-driven decision support systems and predictive analytics can personalize care, reduce polypharmacy, improve cognitive and mood assessments, and identify mobility issues more effectively. This viewpoint highlights the transformative potential of AI in expanding care for older adults by creating easier and more effective ways to implement 4Ms care, potentially improving outcomes and experience for older adults across the VA. Embracing AI in AFHS holds promise for enhancing care quality, optimizing resources, and improving well-being. Further research and collaboration are needed to fully harness AI's potential in age-friendly care, leading to effective and person-centered health care solutions globally.

Acknowledgments

We thank the Veterans Health Administration Office of Geriatrics and Extended Care and the Health and Aging Policy Fellowship mentors. We thank the Geriatrics Research, Education, and Clinical Center (GRECC) and other facilities at the Veterans Affairs Northeast Ohio Healthcare System, and the Gainesville GRECC at the North Florida/South Georgia Veterans Health System.

XSL•FO
RenderX

Disclaimer

The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs or the United States government.

Conflicts of Interest

None declared.

References

1. Fact sheet: aging in the United States. Population Reference Bureau. URL: https://www.prb.org/resources/fact-sheet-aging-in-the-united-states/ [accessed 2025-10-14]
2. Augustine MR, Mason T, Baim-Lance A, Boockvar K. Reasons older veterans use the Veterans Health Administration and non-VHA care in an urban environment. J Am Board Fam Med 2021;34(2):291-300. [doi: 10.3122/jabfm.2021.02.200332] [Medline: 33832997]
3. Projected future need for geriatricians. The American Geriatrics Society. URL: https://www.americangeriatrics.org/sites/default/files/inline-files/Projected-Future-Need-for-Geriatricians_1.pdf [accessed 2025-10-10]
4. Church K, Munro S, Shaughnessy M, Clancy C. Age-friendly health systems: improving care for older adults in the Veterans Health Administration. Health Serv Res 2023 Feb;58 Suppl 1(Suppl 1):5-8. [doi: 10.1111/1475-6773.14110] [Medline: 36477634]
5. Age-friendly health systems: guide to using the 4ms in the care of older adults. Institute for Healthcare Improvement. 2020. URL: https://www.americangeriatrics.org/sites/default/files/inline-files/IHIAgeFriendlyHealthSystems_GuidetoUsing4MsCare.pdf [accessed 2026-02-09]
6. Farrell TW, Volden TA, Butler JM, et al. Age‐friendly care in the Veterans Health Administration: past, present, and future. J American Geriatrics Society 2023 Jan;71(1):18-25 [FREE Full text] [doi: 10.1111/jgs.18070]
7. Schwartz AW, Munro S, Echt KV, Mirk A, Solberg LM, Wozneak K. Scaling and spreading age‐friendly care: early lessons from the VA national age‐friendly action community. J American Geriatrics Society 2025 Feb;73(2):583-591 [FREE Full text] [doi: 10.1111/jgs.19321]
8. Church K, Munro S, Solberg LM, Gruber EA, Shaughnessy M. Age-friendly health systems and meeting the principles of high reliability organizations in the VHA. Fed Pract 2024 Jan;41(1):29-33. [doi: 10.12788/fp.0448] [Medline: 38835358]
9. Schwartz AW, Jindal SK, Wozneak KA, Burke RE. Implementation of the age-friendly health systems initiative in the Department of Veterans Affairs: 5 years of improving quality for older veterans. INQUIRY 2025;62:469580251366883. [doi: 10.1177/00469580251366883] [Medline: 40888158]
10. Koen S, Cavaletto M, Garris L, Hewitt J. Pharmacist-led deprescribing of aspirin for primary prevention of cardiovascular disease among geriatric veterans. Fed Pract 2025 Jan;42(1):22-28. [doi: 10.12788/fp.0537] [Medline: 40529601]
11. RePORTER: project 10862047 — a learning health system approach to using artificial intelligence–enabled decision support (AEDS) for medication optimization in veteran care. RePORTER. URL: https://reporter.nih.gov/search/fGCcdcc_lkuKIdrvCZXRIw/project-details/10862047 [accessed 2025-10-01]
12. Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. Med Care 2013 Apr;51(4):368-373. [doi: 10.1097/MLR.0b013e31827da95a] [Medline: 23269113]
13. Serra MC, Addison O, Giffuni J, Paden L, Morey MC, Katzel L. Physical function does not predict care assessment need score in older veterans. J Appl Gerontol 2019 Mar;38(3):412-423. [doi: 10.1177/0733464817690677] [Medline: 28380717]
14. Xu Y, Liu X, Cao X, et al. Artificial intelligence: a powerful paradigm for scientific research. Innovation (Camb) 2021 Nov 28;2(4):100179. [doi: 10.1016/j.xinn.2021.100179] [Medline: 34877560]
15. Shiwani T, Relton S, Evans R, et al. New horizons in artificial intelligence in the healthcare of older people. Age Ageing 2023 Dec 1;52(12):afad219. [doi: 10.1093/ageing/afad219] [Medline: 38124256]
16. Abadir P, Chellappa R. Artificial intelligence in geriatrics: riding the inevitable tide of promise, challenges, and considerations. J Gerontol A Biol Sci Med Sci 2024 Feb 1;79(2):glad279. [doi: 10.1093/gerona/glad279] [Medline: 38289911]
17. Wang J, Liang Y, Cao S, Cai P, Fan Y. Application of artificial intelligence in geriatric care: bibliometric analysis. J Med Internet Res 2023 Jun 23;25:e46014. [doi: 10.2196/46014] [Medline: 37351923]
18. Gleason KT, Powell D, DeGennaro AP, Wu MMJ, Zhang T, Wolff JL. Patient portal messages to support an age‐friendly health system for persons with dementia. J American Geriatrics Society 2024 Jul;72(7):2140-2147. [doi: 10.1111/jgs.18841]
19. Min SH, Woo K, Song J, et al. Understanding daily care experience preferences across the lifespan of older adults: application of natural language processing. West J Nurs Res 2025 Feb;47(2):71-81. [doi: 10.1177/01939459241306946] [Medline: 39707813]
20. Mess SA, Mackey AJ, Yarowsky DE. Artificial intelligence scribe and large language model technology in healthcare documentation: advantages, limitations, and recommendations. Plast Reconstr Surg Glob Open 2025 Jan;13(1):e6450. [doi: 10.1097/GOX.0000000000006450] [Medline: 39823022]

XSL•FO
RenderX

21. Hu Q, Li J, Li X, Zou D, Xu T, He Z. Machine learning to predict adverse drug events based on electronic health records: a systematic review and meta-analysis. J Int Med Res 2024 Dec;52(12):3000605241302304. [doi: 10.1177/03000605241302304] [Medline: 39668733]
22. Ong JCL, Chen MH, Ng N, et al. A scoping review on generative AI and large language models in mitigating medication related harm. NPJ Digit Med 2025 Mar 28;8(1):182. [doi: 10.1038/s41746-025-01565-7] [Medline: 40155703]
23. Socrates V, Wright DS, Huang T, et al. Identifying deprescribing opportunities with large language models in older adults: retrospective cohort study. JMIR Aging 2025 Apr 11;8:e69504. [doi: 10.2196/69504] [Medline: 40215480]
24. Wu C, Szep J, Hariri S, Agarwal NK, Agarwal SK, Nevarez C. SeVA: an AI solution for age friendly care of hospitalized older adults. In: Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2021) - Volume 5: HEALTHINF: SCITEPRESS; 2021. [doi: 10.5220/0010313605830591]
25. Gong KD, Lu R, Bergamaschi TS, et al. Predicting intensive care delirium with machine learning: model development and external validation. Anesthesiology 2023 Mar 1;138(3):299-311. [doi: 10.1097/ALN.0000000000004478] [Medline: 36538354]
26. Nejadshamsi S, Karami V, Ghourchian N, et al. Development and feasibility study of hope model for prediction of depression among older adults using Wi-Fi-based motion sensor data: machine learning study. JMIR Aging 2025 Mar 3;8(6):e67715. [doi: 10.2196/67715] [Medline: 40053734]
27. Bayen E, Nickels S, Xiong G, et al. Reduction of time on the ground related to real-time video detection of falls in memory care facilities: observational study. J Med Internet Res 2021 Jun 17;23(6):e17551. [doi: 10.2196/17551] [Medline: 34137723]
28. Janes WE, Marchal N, Song X, et al. Integrating ambient in-home sensor data and electronic health record data for the prediction of outcomes in amyotrophic lateral sclerosis: protocol for an exploratory feasibility study. JMIR Res Protoc 2025 Mar 12;14:e60437. [doi: 10.2196/60437] [Medline: 40073394]
29. Liu Y, Zhang G, Tarolli CG, et al. Monitoring gait at home with radio waves in Parkinson's disease: a marker of severity, progression, and medication response. Sci Transl Med 2022 Sep 21;14(663):eadc9669. [doi: 10.1126/scitranslmed.adc9669] [Medline: 36130014]
30. Orejel Bustos AS, Tramontano M, Morone G, et al. Ambient assisted living systems for falls monitoring at home. Expert Rev Med Devices 2023;20(10):821-828. [doi: 10.1080/17434440.2023.2245320] [Medline: 37610096]
31. Pillai M, Blumke TL, Studnia J, et al. Improving postsurgical fall detection for older Americans using LLM-driven analysis of clinical narratives. medRxiv. 2024 Jun 26. [doi: 10.1101/2024.06.25.24309480]
32. Loveys K, Prina M, Axford C, et al. Artificial intelligence for older people receiving long-term care: a systematic review of acceptability and effectiveness studies. Lancet Healthy Longev 2022 Apr;3(4):e286-e297. [doi: 10.1016/S2666-7568(22)00034-4] [Medline: 35515814]
33. Ambalavanan R, Snead RS, Marczika J, Towett G, Malioukis A, Mbogori-Kairichi M. Challenges and strategies in building a foundational digital health data integration ecosystem: a systematic review and thematic synthesis. Front Health Serv 2025;5:1600689. [doi: 10.3389/frhs.2025.1600689] [Medline: 40621432]
34. Liang M. Ethical AI in medical text generation: balancing innovation with privacy in public health. Front Public Health 2025;13:1583507. [doi: 10.3389/fpubh.2025.1583507] [Medline: 40756387]
35. Hoffman J, Wenke R, Angus RL, Shinners L, Richards B, Hattingh L. Overcoming barriers and enabling artificial intelligence adoption in allied health clinical practice: a qualitative study. Digit Health 2025;11:20552076241311144. [doi: 10.1177/20552076241311144] [Medline: 39906878]

Abbreviations

AFHS: Age-Friendly Health Systems
AI: artificial intelligence
CDS: clinical decision support
EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
IHI: Institute for Healthcare Improvement
LLM: large language model
ML: machine learning
NLP: natural language processing
VA: Department of Veterans Affairs

XSL•FO
RenderX

Edited by M Mardini; submitted 25.Apr.2025; peer-reviewed by MJ Kwak, RT Potla, S Keddem; accepted 08.Dec.2025; published 23.Feb.2026.

Please cite as:
Fine Smilovich E, Kalsy M, Wozneak K, Syed Q, Solberg LM
Leveraging AI to Advance Age-Friendly Care in the Veterans Health Administration
JMIR Aging 2026;9:e75686
URL: https://aging.jmir.org/2026/1/e75686
doi:10.2196/75686

© Elizabeth Fine Smilovich, Megha Kalsy, Kimberly Wozneak, Quratulain Syed, Laurence M Solberg. Originally published in JMIR Aging (https://aging.jmir.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Integrated High-Throughput Targeted Metabolomics and Machine Learning for Early Prediction and Prevention of Postoperative Delirium in Older Adult Surgical Patients: Prospective Multicenter Cohort Study

Gengrui Zhong[1*], MD; Xiaoli Huang[2*], MD; Congye Li[1], MD; Deqiang Wang[3], MD; Dingding Huang[4], MD; Menghan Sun[5], MD; Quanhong Zhou[1], MD, PhD; Yong Guo[1], MD, PhD

[1]Department of Critical Care Medicine, The Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, 600 Yishan Road, Shanghai, China

[2]Department of Critical Care Medicine, Jinshan District Central Hospital, Shanghai, China

[3]Department of Critical Care Medicine, Fengxian District Central Hospital, Shanghai, China

[4]Department of Anesthesiology, Fengxian District Central Hospital, Shanghai, China

[5]Department of Critical Care Medicine, Shanghai Eighth People's Hospital, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Yong Guo, MD, PhD
Department of Critical Care Medicine, The Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, 600 Yishan Road, Shanghai, China

Abstract

Background: Postoperative delirium (POD) is a common and severe complication in older adult patients with hip fracture, yet its pathogenesis remains unclear.

Objective: This study aimed to develop a predictive model for POD following hemiarthroplasty in older adult patients by integrating high-throughput targeted metabolomics and machine learning.

Methods: In this prospective multicenter cohort study, 260 older adult patients undergoing hemiarthroplasty for hip fracture were enrolled. Preoperative serum samples were analyzed via high-throughput targeted metabolomics. Differential metabolites were screened using random forest (RF) and least absolute shrinkage and selection operator regression. Predictive models were constructed using gradient boosting, logistic regression, and RF, with performance evaluated using receiver operating characteristic curves and the area under these curves (area under the receiver operating characteristic curve [AUC]).

Results: Absolute quantification of 201 metabolites revealed 41 (20.4%) significantly differentially expressed metabolites. RF and least absolute shrinkage and selection operator regression identified 16 candidate biomarkers. The logistic regression model demonstrated optimal performance, achieving an AUC of 0.855 (95% CI 0.800 - 0.910) in the overall cohort. Upon 7:3 partitioning into training and test sets, the model maintained robust predictive accuracy, with AUCs of 0.844 and 0.856, respectively.

Conclusions: Integration of preoperative metabolomics and machine learning enabled accurate prediction of POD in older adult patients with hip fracture, facilitating personalized risk stratification and tailored clinical management.

Trial Registration: Chinese Clinical Trial Registry ChiCTR-CPC-15006141; https://www.chictr.org.cn/showprojEN.html?proj=10640

(JMIR Aging 2026;9:e78495) doi:10.2196/78495

KEYWORDS

high-throughput targeted metabolomics; machine learning; prediction and prevention; postoperative delirium; older adult surgical patients

Introduction

Postoperative delirium (POD) is a common and severe neuropsychiatric syndrome affecting 10% to 50% of older adult surgical patients, particularly within the first 48 hours postoperatively. Characterized by acute fluctuations in attention, cognitive impairment, and altered consciousness, POD significantly prolongs intensive care unit stays, increases hospitalization costs, and elevates mortality rates [1]. Long-term consequences include persistent cognitive decline and an

XSL•FO
RenderX

elevated risk of dementia [2,3]. The pathogenesis of POD remains elusive, with hypotheses spanning neuroinflammation, neurotransmitter imbalances, metabolic dysregulation, and sleep-wake cycle disruptions [4]. Notably, preoperative cognitive impairment, advanced age, and perioperative stressors such as surgery duration and blood transfusions have been identified as independent risk factors [5,6]. Given the global aging population and rising surgical demand, there is an urgent need for early diagnostic tools and preventive strategies to mitigate POD's clinical and economic burdens.

Current diagnostic approaches rely heavily on clinical observation, which lacks sensitivity and specificity, particularly in the early stages of POD [7]. While metabolic dysregulation has emerged as a key feature of POD, traditional biomarkers lack predictive power. Metabolomics—a systems biology approach analyzing small-molecule metabolites—offers unprecedented opportunities to elucidate underlying pathophysiological pathways [8]. For instance, our preliminary studies revealed significant disturbances in polyunsaturated fatty acids, energy metabolism, oxidative stress, and amino acid imbalances in postoperative patients [9,10]. Furthermore, systemic inflammation markers (eg, interleukin-6 and C-reactive protein) and β-amyloid peptides correlated with POD onset in vulnerable populations [11]. However, existing research has yet to integrate high-throughput targeted metabolomics with artificial intelligence to develop predictive models capable of identifying at-risk individuals preoperatively [12].

Surgery-induced systemic inflammation is a core driver of POD, wherein metabolites act as dual-effect mediators—either amplifying inflammation or providing neuroprotection. Proinflammatory lipid mediators derived from arachidonic acid, such as prostaglandin E2, can directly damage hippocampal synapses [13,14], whereas preoperative deficiencies in gut microbiota–derived short-chain fatty acids such as propionate impair the suppression of neuroinflammation via the nuclear factor kappa-light-chain-enhancer of activated B cell pathway [15]. Concurrently, imbalances in neurotransmitter systems significantly contribute to cognitive dysfunction. Inadequate availability of choline, a precursor for acetylcholine synthesis, combined with metabolic alterations in γ-aminobutyric acid precursors such as L-serine, exacerbates postoperative neuropsychiatric symptoms [13]. Furthermore, the age-related decline in cerebral metabolic reserve predisposes older adult patients to neuronal energy failure under surgical stress. Disruptions in the tricarboxylic acid (TCA) cycle lead to lactate accumulation, and deficiencies in branched-chain amino acids—key substrates for neuronal energy metabolism—compromise neurotrophic support, establishing a pathological cycle of "energy substrate exhaustion–metabolic waste accumulation" [16]. These mechanisms underscore the value of metabolites as "molecular sensors" in POD pathophysiology, providing a theoretical foundation for our targeted metabolomics approach focused on inflammation, neurotransmission, and energy metabolism.

Building on prior work characterizing metabolic disturbances in geriatric perioperative patients [17-19], this study represents a paradigm shift by integrating high-throughput targeted metabolomics with machine learning to develop a clinically actionable prediction model for POD in older adult patients undergoing hip hemiarthroplasty. Unlike conventional approaches that rely on single biomarkers or broad genomic frameworks [19,20], our targeted metabolomics platform enables high quantitative accuracy and reproducibility, thereby facilitating clinical translation [21-23]. By moving beyond traditional physiological parameters and leveraging machine learning to integrate multiple metabolite markers, we aim to overcome the limitations of single-biomarker strategies and conventional statistical models.

Therefore, the primary objective of this multicenter prospective cohort study was to identify preoperative serum metabolite signatures associated with POD by using high-throughput targeted metabolomics. Furthermore, we sought to develop and validate a machine learning–based predictive model that integrates these metabolite biomarkers to enable early risk stratification and individualized prevention of POD in older adult surgical patients. This integrative approach not only aims to advance the mechanistic understanding of POD but also holds potential for clinical implementation in perioperative care.

Methods

Study Design

This prospective observational multicenter cohort study was conducted from June 2021 to January 2024, with patient recruitment from June 2021 to December 2023 and follow-up completion in January 2024. In this study, 586 older adult patients (aged ≥60 years) scheduled for hemiarthroplasty due to hip fractures were initially enrolled. Inclusion criteria were age of 60 years or above and undergoing total hip replacement under uniform general anesthesia. Exclusion criteria were American Society of Anesthesiologists physical status classification of grade III or higher, Mini-Mental State Examination (MMSE) score below 23, history of central nervous system (CNS) disorders or psychiatric diseases, current use of sedatives or antidepressants, preoperative biochemical abnormalities indicating renal dysfunction or active hepatic disease, preexisting diabetes mellitus, history of neurosurgical or cardiac surgery, refusal to complete study protocols or inability to comprehend the study language, severe auditory or visual impairment, illiteracy, Parkinson disease, and alcohol or substance dependence. All baseline data (demographics and clinical characteristics), preoperative serum samples, and postoperative outcomes (POD incidence and cognitive function) were collected prospectively according to a predefined protocol without retrospective data retrieval.

Ethical Considerations

All experimental procedures in this study were approved by the Ethics Committee of the Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (approval number 2021-YS-237). The study was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from each participant or their legally authorized representative prior to enrollment. To protect participant privacy and confidentiality, all personal identifiers were removed, and data were anonymized prior to analysis. Access to study data was restricted to authorized research personnel only. No

XSL•FO
RenderX

financial compensation was provided to participants. All study procedures in this research have been registered at the Chinese Clinical Trial Registry [24] with the registration number ChiCTR-CPC-15006141.

Dietary Management

All patients in the experimental group received a light diet preoperatively. Semiliquid feeding was initiated 6 hours after general anesthesia, followed by a gradual transition to a light regular diet.

Anesthetic Protocol

Patients underwent standardized general anesthesia for hemiarthroplasty. Premedication included intramuscular midazolam (3 mg). Anesthesia induction was performed with intravenous fentanyl (0.1 mg), propofol (1.5 mg/kg), and vecuronium (0.1 mg/kg). Maintenance anesthesia was administered via a semiclosed circuit system using sevoflurane (1‐1.5 maximum allowable concentration) and oxygen (1 L/min). Continuous intraoperative monitoring included pulse oximetry, invasive arterial blood pressure waveform, end-tidal partial pressure of carbon dioxide, and 3-lead electrocardiography.

Surgical Procedure

Hemiarthroplasty involved replacement of the fractured femoral head with a metallic prosthesis and insertion of a metal stem into the femoral shaft to enhance implant stability.

Biospecimen Collection

Preoperative blood samples were collected on the day of the surgery prior to entering the operating room. Postoperative samples were obtained immediately upon transfer from the postanesthesia care unit. Serum was cryopreserved for subsequent high-throughput targeted metabolomics profiling.

POD Diagnosis

Among the 260 enrolled patients undergoing total hip replacement, POD was assessed twice daily (8 AM to 8 PM on postoperative days 1‐3) using the Confusion Assessment Method (CAM) [25,26] as recommended by the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Diagnostic criteria were (1) acute onset with fluctuating symptoms, (2) inattention, (3) disorganized thinking, and (4) altered consciousness. POD was confirmed if criteria 1 and 2 were present along with either criterion 3 or 4. Patients were categorized into POD and non-POD groups.

The CAM (POD diagnosis) was conducted by research nurses certified in the CAM scale. Before study initiation, all involved nurses completed certified training for the Chinese version of the CAM scale (certificate numbers archived). The assessment results were reviewed, and the POD diagnosis was confirmed by the designated psychiatrists (see the Perioperative Cognitive Assessment section), thereby avoiding potential bias from nurses' subjective judgments.

Clinical Data Collection

Baseline demographic data, anesthesia- and surgery-related parameters, and other clinical variables were systematically documented. Clinical data collection was performed by research coordinators. Each center designated 1 coordinator (with a master's degree or higher in nursing or a related field) responsible for recording baseline characteristics (age, sex, and BMI) and surgical and anesthesia parameters (operation duration and anesthesia duration). Data entry was monitored weekly by a clinical research associate who spot-checked 20% of cases to ensure accuracy.

Observer Selection and Training for Clinical Assessments

CAM Assessors (Research Nurses)

All nurses completed standardized training on the Chinese CAM version led by psychiatrists and obtained certification through theoretical and practical examinations to ensure accurate identification of core delirium symptoms. Only certified nurses participated.

MMSE Assessors (Psychiatrists)

Each center designated 1 psychiatrist with 5 years or more of geriatric neuropsychological experience to conduct all MMSE evaluations, minimizing interassessor variability and ensuring consistent scoring practices.

High-Throughput Targeted Metabolomics Platform

Serum samples were analyzed using the Q300 metabolite assay kit (Metabo-Profile) [27]. Sample processing, including protein precipitation and derivatization, was automated on an Eppendorf epMotion workstation. Metabolite separation and detection were performed using ultrahigh-performance liquid chromatography (UPLC) coupled with tandem mass spectrometry (ACQUITY UPLC Xevo TQ-S system; Waters). The detailed sample preparation protocol and complete UPLC coupled with tandem mass spectrometry parameters are provided in Multimedia Appendix 1. Raw data processing, including peak integration and metabolite quantification, was conducted using the QuanMET software (version 2.0; Metabo-Profile). Quality control (QC) samples were analyzed intermittently to ensure analytical consistency.

Standardization Measures: Ensuring Data Consistency Through Training and Reliability Checks

To avoid multicenter data bias, the following standardization measures were implemented before and during the study.

Unified Training

Specific training sessions were conducted, covering (1) interpretation of the methods and inclusion and exclusion criteria (eg, American Society of Anesthesiologists grade determination standards and MMSE scoring details), (2) demonstration of scale administration procedures (eg, step-by-step demonstration of the Chinese MMSE and CAM), and (3) case exercises (all involved health care professionals [including research nurses and psychiatrists] practiced independent scoring, followed by group review of discrepancies to unify judgment standards).

Platform Standardization

Metabolomics analysis was performed by Metabo-Profile. All samples were preprocessed according to a unified workflow

XSL•FO
RenderX

(High-Throughput Targeted Metabolomics Platform section). One QC sample was inserted every 12 samples to ensure interbatch consistency of the assay results (coefficient of variation<30%).

Blinding Design: Assessors Were Double Blinded to Metabolomics Results to Avoid Diagnostic Bias

To prevent knowledge of metabolomics results from influencing cognitive assessments or POD diagnosis, this study used an assessor-analyst double-blinding design.

Assessor Blinding

The psychiatrists and research nurses responsible for MMSE and CAM assessments were only aware of the patients' clinical information (eg, age and surgery type) and were completely unaware of the patients' preoperative serum metabolomics results (eg, differential metabolite levels and whether they belonged to a potential high-risk POD group). Metabolomics testing was performed by an independent laboratory. The test results were provided only to the statistical analyst after all cognitive assessments were completed and POD diagnoses were finalized. Only then were the metabolomics data and clinical outcomes merged for model building and analysis.

Statistical Analyst Blinding

Before modeling, the statistical analyst received only anonymized metabolomics data (patient IDs replaced with random numbers) and anonymized clinical outcomes (1=POD; 0=non-POD). The analyst was unaware of the patients' specific clinical information to avoid subjective selection of feature variables.

Machine Learning Model Development

Hyperparameter Tuning

Model hyperparameters were optimized via 7-fold cross-validation with grid search. Least absolute shrinkage and selection operator (LASSO) regression (λ) was tuned using 10-fold cross-validation (log(λ): −10 to 2). Random forest (RF) used 500 trees with mtry tuned among {2,9,16}. Gradient boosting (Extreme Gradient Boosting) was optimized for tree number (50-150), depth (1-5), learning rate (0.01-0.1), and minimum node size (minimum number of observations in the trees' terminal nodes: 1-5). Logistic regression (LR) required no tuning.

Class Imbalance Handling

Overview

The POD-to–non-POD ratio was approximately 1:3.6. We used the area under the receiver operating characteristic (ROC) curve (AUC) as the primary metric due to its robustness to class imbalance. The test set achieved an AUC for the precision-recall (PR) curve of 0.966, confirming good minority-class identification.

Model Calibration

Calibration plots showed good agreement between predicted probabilities and observed outcomes for the final LR model.

Software

Analysis used R (version 4.3.2; R Foundation for Statistical Computing) with *caret* (version 6.0-94) for model training and tuning and *ggplot2* (version 3.5.0) for visualization.

Comprehensive Model Evaluation

The model's performance was comprehensively evaluated beyond discrimination. The overall model performance was quantified using the Brier score. Calibration was assessed visually using calibration plots and numerically by calculating the calibration in the large (intercept) and calibration slope. The clinical utility of the model was evaluated using decision curve analysis, which quantifies the net benefit across a range of probability thresholds. Given the class imbalance (POD prevalence of approximately 21.5%), PR curves and the average precision (AP) score were also analyzed alongside ROC curves to provide a more informative assessment of the model's performance in identifying the minority class (POD).

Statistical Analysis

Overview

Demographic and surgical data were analyzed using SPSS (version 26.0; IBM Corp). For continuous variables, normality was assessed via the Kolmogorov-Smirnov test; normally distributed data were analyzed via one-way ANOVA, and nonnormally distributed data were analyzed via the Kruskal-Wallis *H* test. For categorical variables, group differences were evaluated using the chi-square test (significance threshold: $P<.05$).

Metabolomics data underwent multivariate and univariate analyses on the iMAP platform (version 1.0; Metabo-Profile). For multivariate analysis, principal component analysis for dimensionality reduction and outlier detection was conducted, with partial least squares discriminant analysis and orthogonal partial least squares discriminant analysis to maximize intergroup discrimination (threshold: Variable Importance in Projection>1.0). For univariate analysis, differential metabolites were screened via 2-tailed *t* test and Mann-Whitney *U* test (threshold: $P<.05$).

RF and LASSO regression identified potential biomarkers [28]. The predictive performance of biomarkers was evaluated using gradient boosting, LR, and RF models, with ROC and PR curves generated to assess diagnostic accuracy.

A nested cross-validation approach was used to mitigate overfitting and robustly assess model generalizability. The cohort (N=260) was initially randomly split into a training set (70%; n=182) and a hold-out test set (30%; n=78). Model development and hyperparameter tuning were confined strictly to the training set using 7-fold cross-validation. In this inner loop, the training set was partitioned into 7 folds; the model was iteratively trained on 6 folds and validated on the remaining fold. The model with the best average performance across the 7 folds (LR) was retrained on the entire training set, and its final performance was evaluated on the independent test set. This strategy prevents information leakage and ensures an unbiased performance estimate.

XSL•FO
RenderX

The final analytical cohort comprised 260 patients, including 56 (21.5%) with POD. For the development of the machine learning model using the 16 selected metabolites, the event per variable (EPV) ratio was 56/16 = 3.5. While an EPV of 10 is often recommended for ideal stability in predictive modeling, an EPV greater than 3 is generally considered acceptable in exploratory biomarker discovery studies, particularly in fields such as metabolomics where prior data are limited [29]. Furthermore, the obtained AUC of 0.855 with a 95% CI of 0.800 to 0.910 indicates a reasonably precise estimate of the model's discriminative ability. This precision, evidenced by the CI width, supports the adequacy of the sample size for the primary objective of model development and initial internal validation within this cohort.

The 95% CIs for the AUC were calculated using the DeLong method [30], which accounts for the pairwise correlation between sensitivity and specificity across all possible thresholds.

Data Preprocessing and QC

Metabolites with more than 20% of missing values were excluded. In this study, all 201 targeted metabolites exhibited missing rates below 2%, and thus, no metabolites required exclusion. Given the minimal missingness, no imputation was performed to preserve data integrity. Batch effects were corrected using QC-based robust locally estimated scatterplot smoothing signal correction implemented in the QuanMET software. For class imbalance between the POD and non-POD groups (POD-to–non-POD ratio of approximately 1:3.6), class weight adjustment was applied to all machine learning models, with weights set inversely proportional to class frequencies.

Results

Summary Workflow Diagram to Illustrate the Complete Research Logic

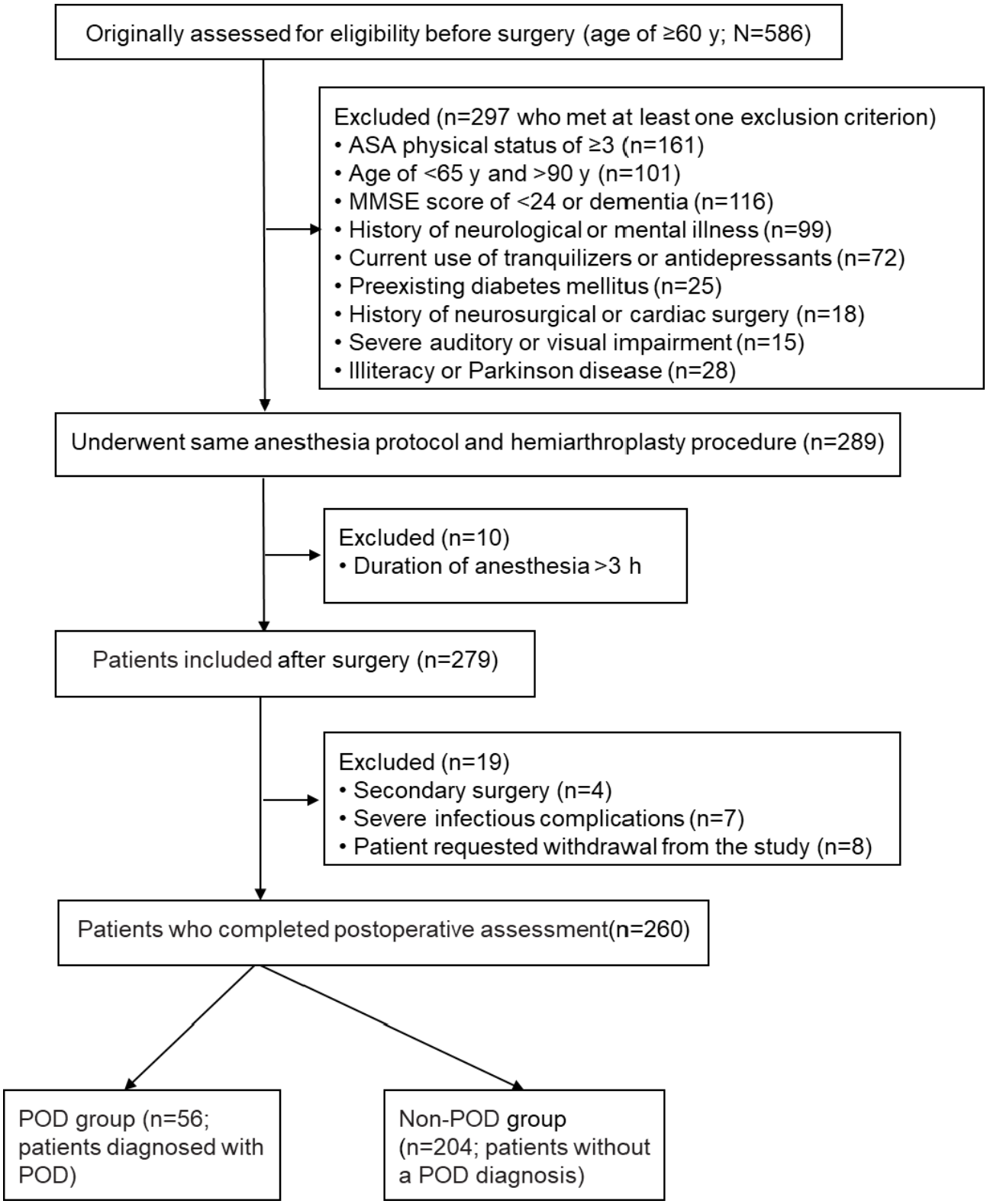
A schematic overview summarizing the complete study workflow is provided in Figure S1 in Multimedia Appendix 1. This diagram visually integrates the key stages of the research, from participant recruitment and data collection to metabolomics profiling, machine learning model development, and the final predictive outcome for POD.

Comparison of General Characteristics Between the POD and Non-POD Groups

The study flowchart (Figure 1) illustrates the patient enrollment process in this multicenter observational cohort study. Ultimately, 260 patients met the inclusion criteria and completed the study protocol. Of these 260 patients, 56 (21.5%) were diagnosed with POD, constituting the POD group, whereas 204 (78.5%) without a POD diagnosis were categorized as the non-POD group. As shown in Table 1, no significant differences were observed between the 2 groups in demographic characteristics, surgical duration, or anesthesia time.

XSL•FO
RenderX

Figure 1. Flowchart of patient screening (final analytical cohort: N=260 older adult patients with hip fracture undergoing hemiarthroplasty; postoperative delirium [POD] group: n=56; non-POD group: n=204). ASA: American Society of Anesthesiologists; MMSE: Mini-Mental State Examination.



XSL•FO
RenderX

Table . Demographic characteristics of the postoperative delirium (POD) and non-POD groups.

Characteristic	POD group (n=56)	Non-POD group (n=204)	*P* value
Age (y), mean (SD)	73.8 (6.8)	72.9 (7.1)	.39
Sex, n (%)			.86
Female	35 (62.5)	128 (62.7)	
Male	21 (37.5)	76 (37.3)	
BMI (kg/m^2), mean (SD)	20.33 (3.7)	20.87 (3.9)	.22
Education (y), mean (SD)	7.6 (4.1)	7.2 (3.9)	.40
ASA[a] grade, n (%)			.88
1	25 (44.6)	89 (43.6)	
2	31 (55.4)	115 (56.4)	
Preoperative MMSE[b] score (0-30), mean (SD)	27.49 (1.9)	27.83 (2.0)	.24
Preoperative CCI[c], mean (SD)	1.9 (0.7)	1.8 (0.9)	.21
Operation duration (min), mean (SD)	93 (38)	88 (39)	.30
Anesthesia duration (min), mean (SD)	119 (35)	117 (39)	.68

[a]ASA: American Society of Anesthesiologists.

[b]MMSE: Mini-Mental State Examination.

[c]CCI: Charlson Comorbidity Index.

Ethical Reporting of Patient Flow and Handling of Attrition

The patient flow throughout the study is detailed in Figure 1. All cases of attrition were meticulously documented in the study case report forms, including the specific reason and time point of withdrawal. This documentation was reviewed and signed by the principal investigator at each site to ensure adherence to ethical standards.

Data from patients who dropped out were handled as follows:

- Patients excluded for medical reasons (n=11): data collected from these patients prior to their exclusion (eg, preoperative blood samples and baseline demographics) were not included in the final statistical analysis or model development. This conservative approach was taken to prevent the introduction of confounding factors related to their severe postoperative medical complications, which are distinct from the pathophysiology under investigation and could distort the association between metabolic profiles and POD.
- Patients lost to follow-up (n=8): as these patients withdrew consent for continued participation in the study protocol postoperatively, no further data were collected. Any preliminary data obtained prior to withdrawal were not used in the analysis, respecting patient autonomy and the ethical principle of voluntary participation.

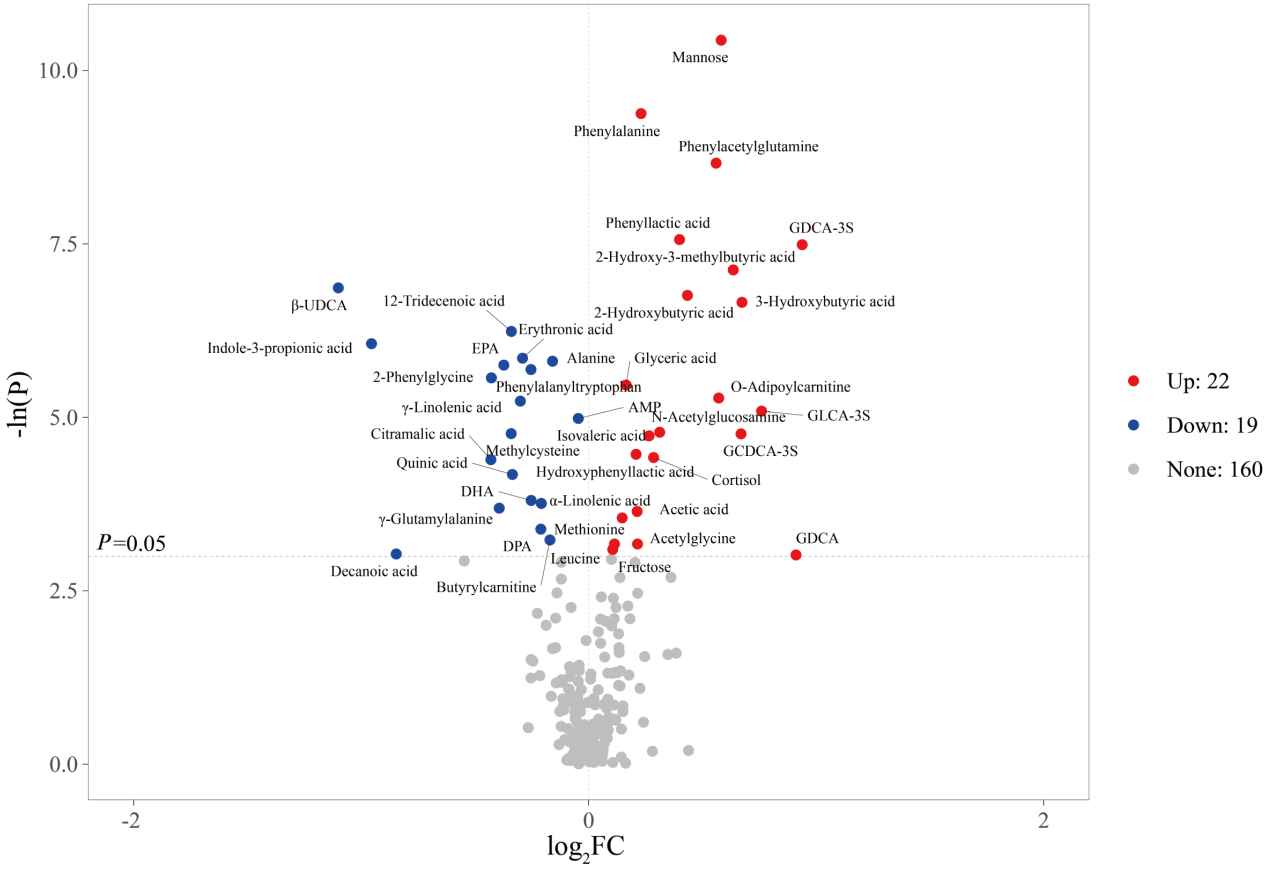
This rigorous approach to reporting patient flow and handling missing data ensures the transparency, reproducibility, and ethical integrity of our research findings. The final analytical cohort (N=260) represents a complete-case dataset for the planned metabolomics and machine learning analyses.

Differences in Preoperative Serum Metabolites Identified Using a High-Throughput Targeted Metabolomics Platform Between the POD and Non-POD Groups

To compare metabolite profiles, we used UPLC coupled with triple quadrupole mass spectrometry for absolute quantification of 201 metabolites. As shown in Figure 2, absolute quantification of 201 metabolites revealed 41 (20.4%) significantly differentially expressed metabolites between the POD and non-POD groups (P<.05). These differential metabolites were primarily enriched within 4 core metabolic domains (amino acid metabolism, energy metabolism, lipid metabolism, and bile acid metabolism), suggesting that POD development is closely associated with disruptions in these key biological pathways.

XSL•FO
RenderX

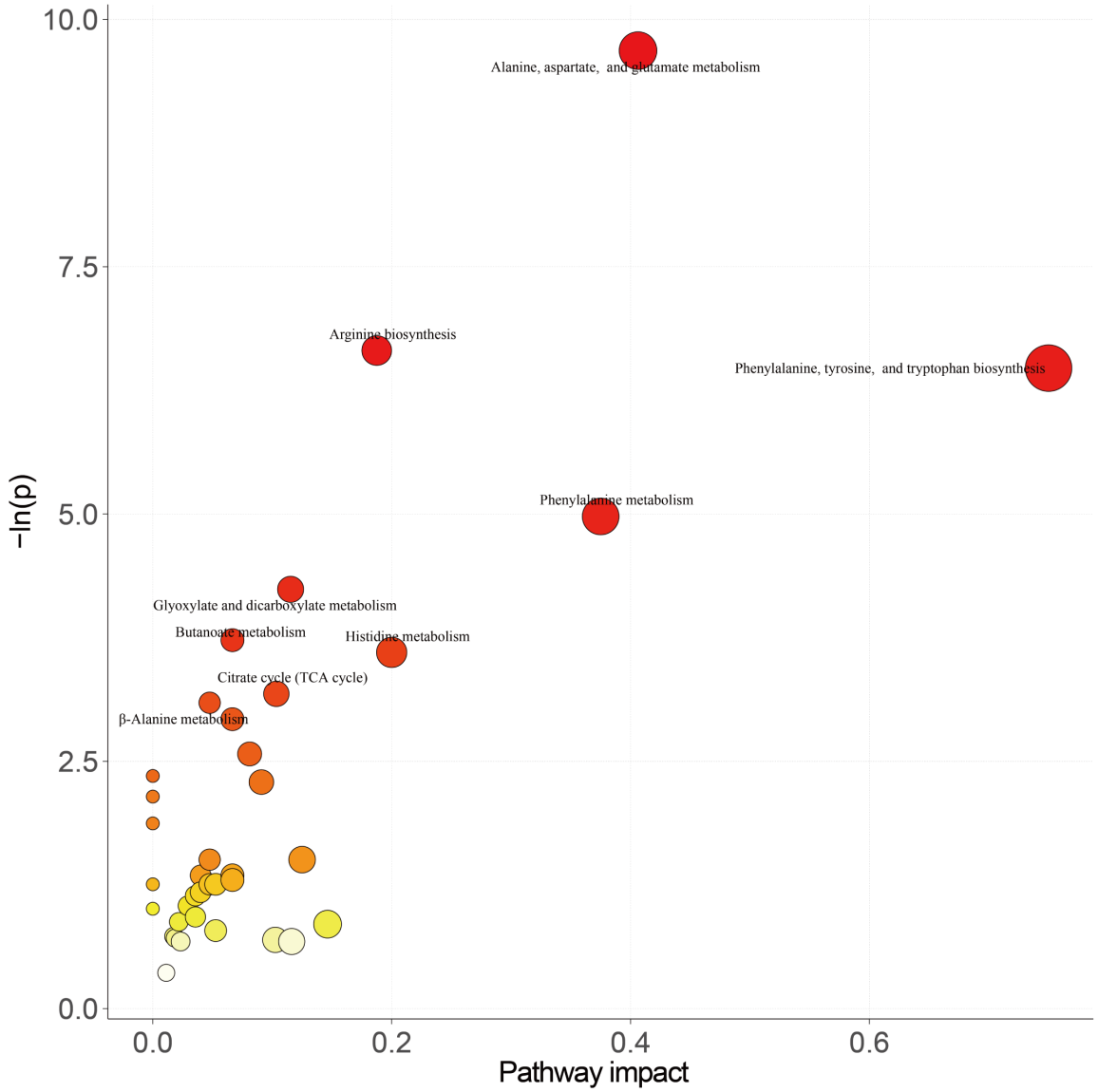
Figure 2. Volcano plot of differentially abundant preoperative serum metabolites between the postoperative delirium (POD) and non-POD groups (41 significant metabolites identified). β-UDCA: β-ursodeoxycholic acid; AMP: adenosine monophosphate; DHA: docosahexaenoic acid; DPA: docosapentaenoic acid; EPA: eicosapentaenoic acid; GCDCA-3S: glycochenodeoxycholic acid 3-sulfate; GDCA: glycodeoxycholic acid; GDCA-3S: glycodeoxycholic acid 3-sulfate; GLCA-3S: glycolithocholic acid 3-sulfate; log2FC: log 2 fold change.



Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis further confirmed that these differential metabolites were significantly enriched in key pathophysiological pathways, including alanine, aspartate, and glutamate metabolism (neurotransmitter synthesis); phenylalanine, tyrosine, and tryptophan biosynthesis (precursors for dopamine and serotonin); and the citrate cycle (TCA cycle and core energy metabolism; Figure 3).

XSL•FO
RenderX

Figure 3. Bubble plot of Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis for the differential metabolites (core enrichment in 4 metabolic pathways; $P<.05$). TCA: tricarboxylic acid.



Machine Learning–Powered Discovery of Specific Metabolic Biomarkers for Predictive Modeling of POD

By performing a union operation on the top 10 differentially abundant metabolites identified through RF model screening (Figure 4) and LASSO regression–based feature selection (Figure 5), we obtained 16 distinct metabolites (see Table S1 in Multimedia Appendix 1 for the complete list). These biomarkers span the aforementioned 4 core metabolic domains and collectively form the feature set for predictive modeling. Subsequently, predictive models were constructed using these 16 metabolites via gradient boosting, LR, and RF algorithms. Among these, the LR model demonstrated strong predictive performance, achieving an AUC of 0.855 (95% CI 0.800 - 0.910; Figure 6).

XSL•FO
RenderX

Figure 4. Top 10 core predictive metabolites screened using the random forest algorithm (ranked by Gini importance). β-UDCA: β-ursodeoxycholic acid; AMP: adenosine monophosphate.

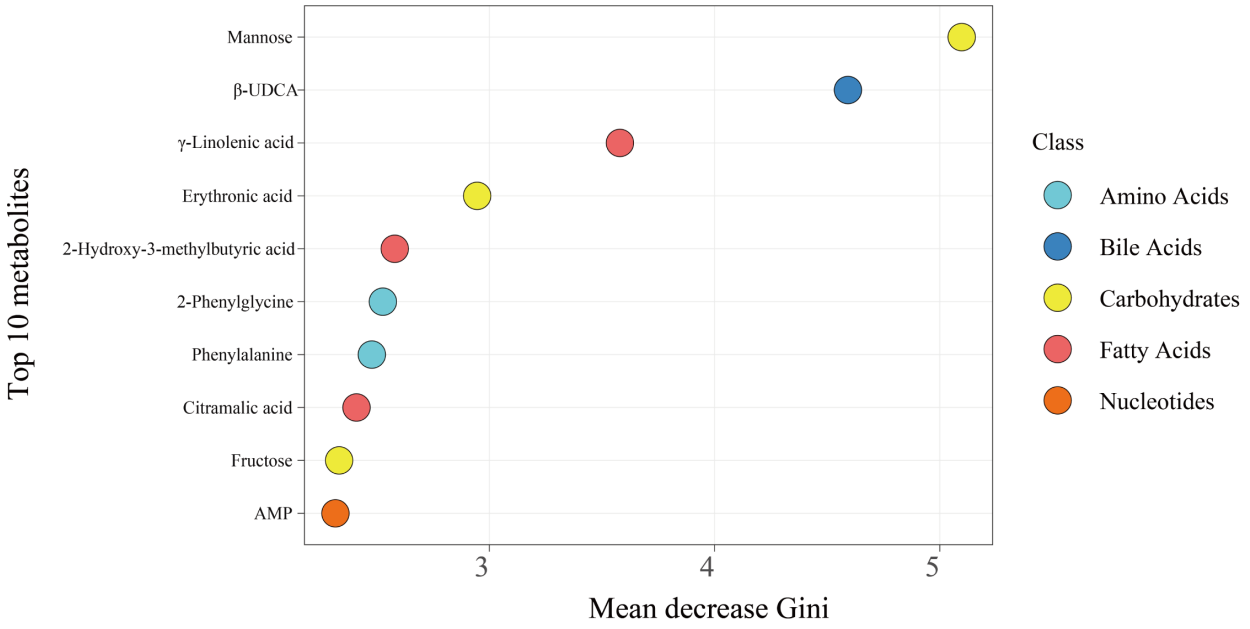
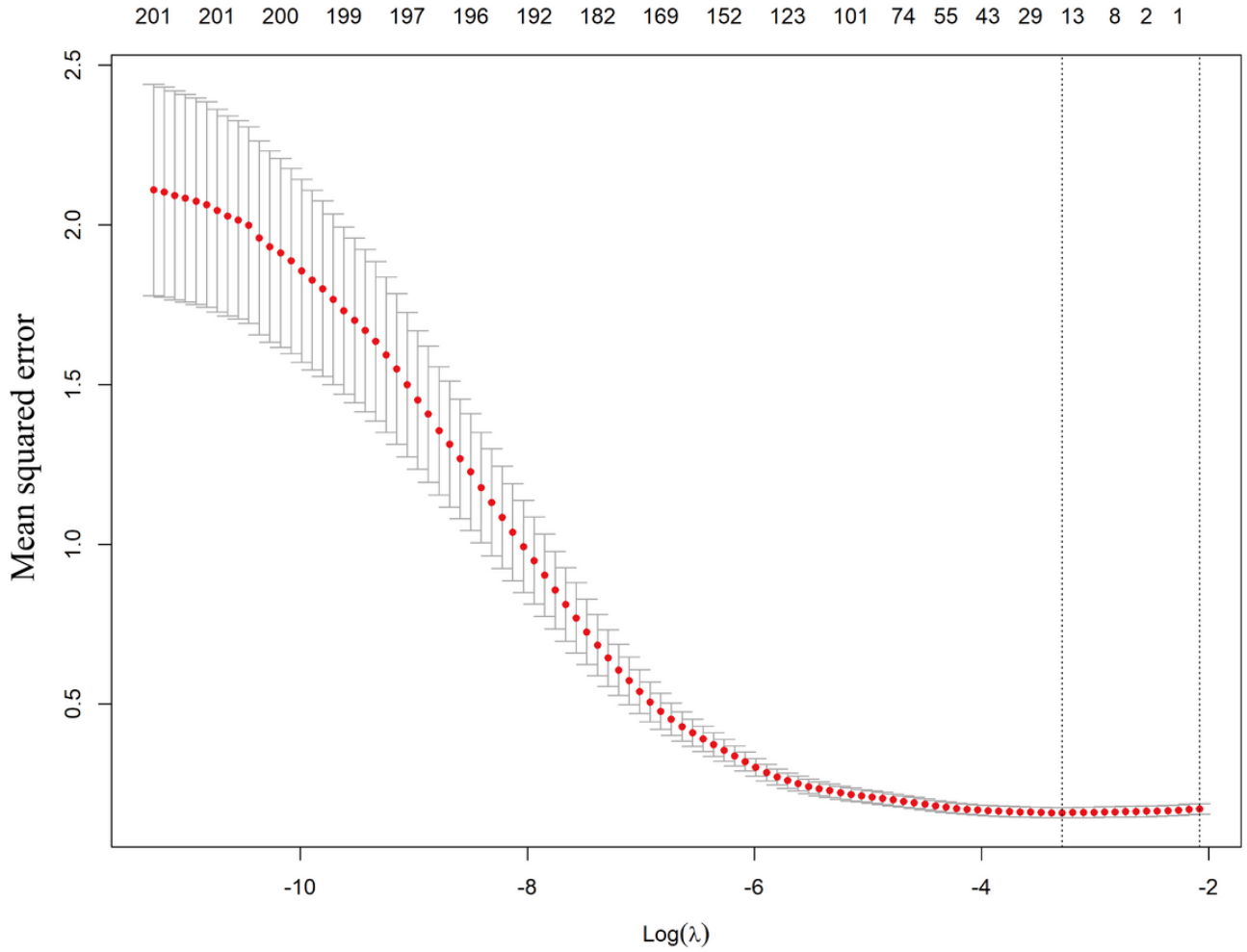
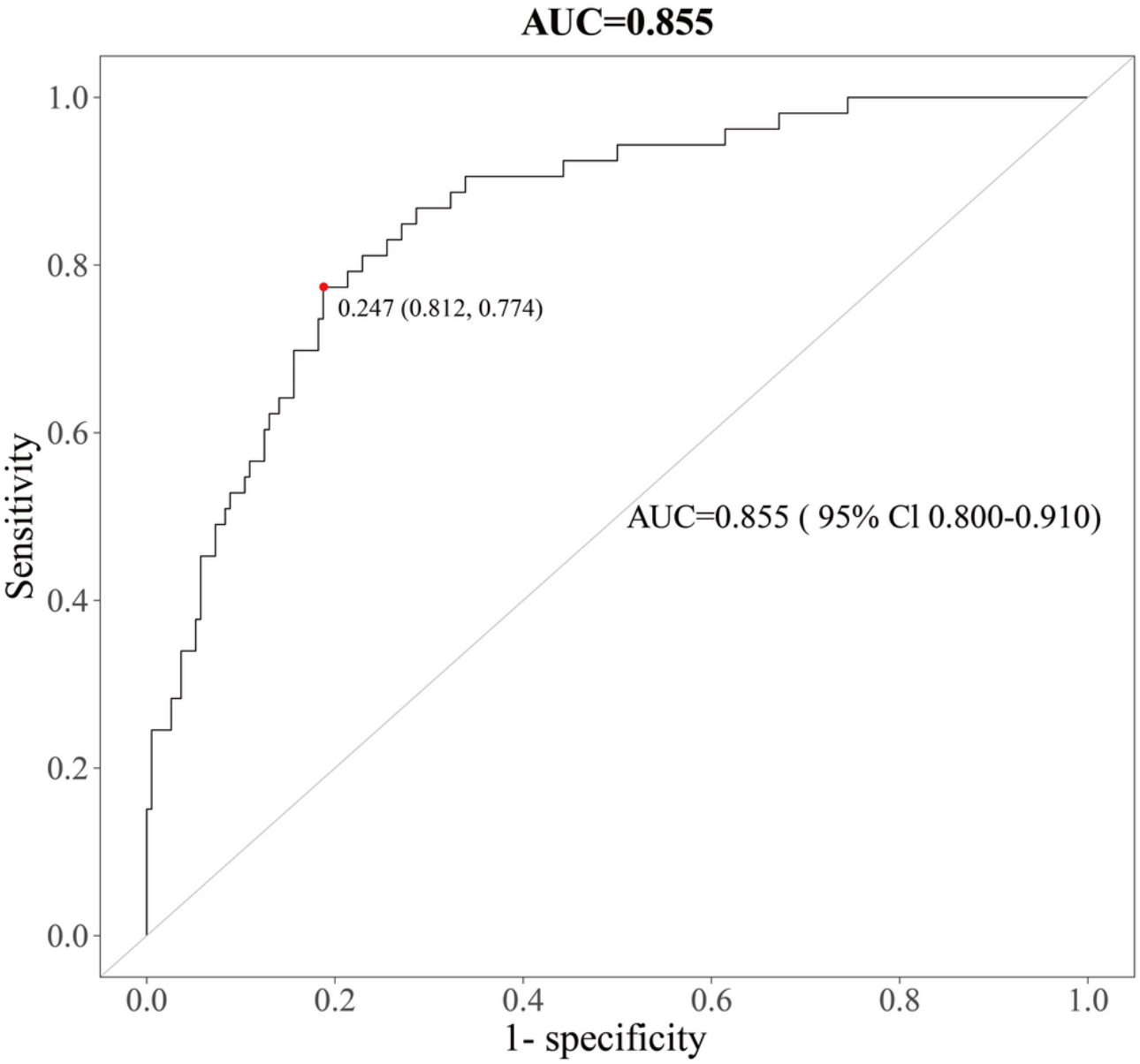


Figure 5. Feature selection plot for differential metabolites using least absolute shrinkage and selection operator regression (12 key metabolites selected at optimal λ).



XSL•FO
RenderX

Figure 6. Receiver operating characteristic (ROC) curve of the logistic regression model for predicting postoperative delirium based on 16 core metabolites in the overall cohort (area under the ROC curve [AUC]=0.855, 95% CI 0.800 - 0.910).

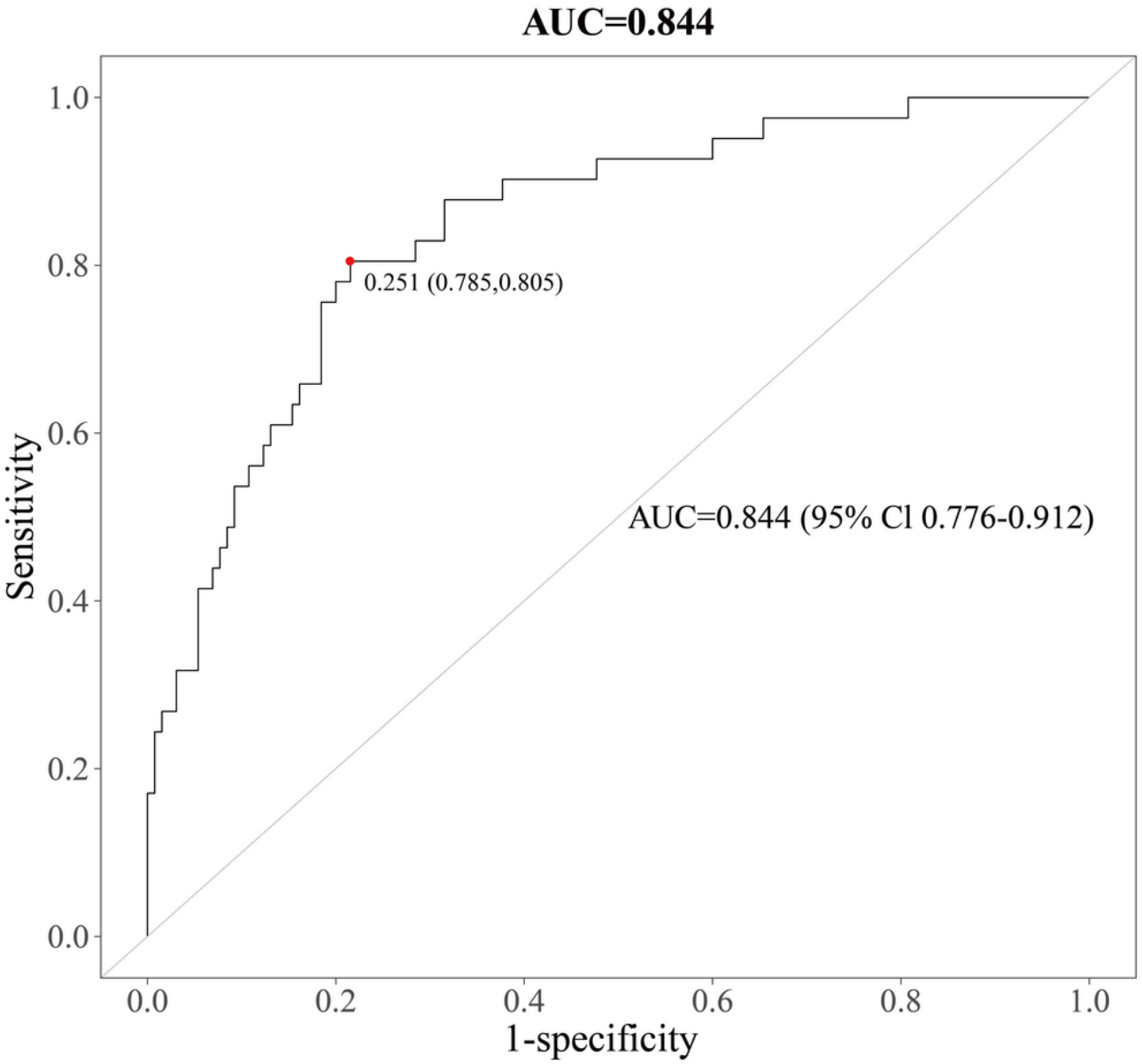


Randomized Cohort Partition Into Training and Test Sets for Machine Learning Validation of the POD Prediction Model

Total samples were randomly partitioned into training and testing sets (7:3 ratio) via computerized randomization algorithms. Differential metabolites underwent rigorous LR modeling, demonstrating robust predictive performance with AUC values of 0.844 (95% CI 0.776–0.912) in the training cohort (Figure 7) and 0.856 (95% CI 0.741–0.971) in the independent test cohort (Figure 8), meeting TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines for clinical prediction models.

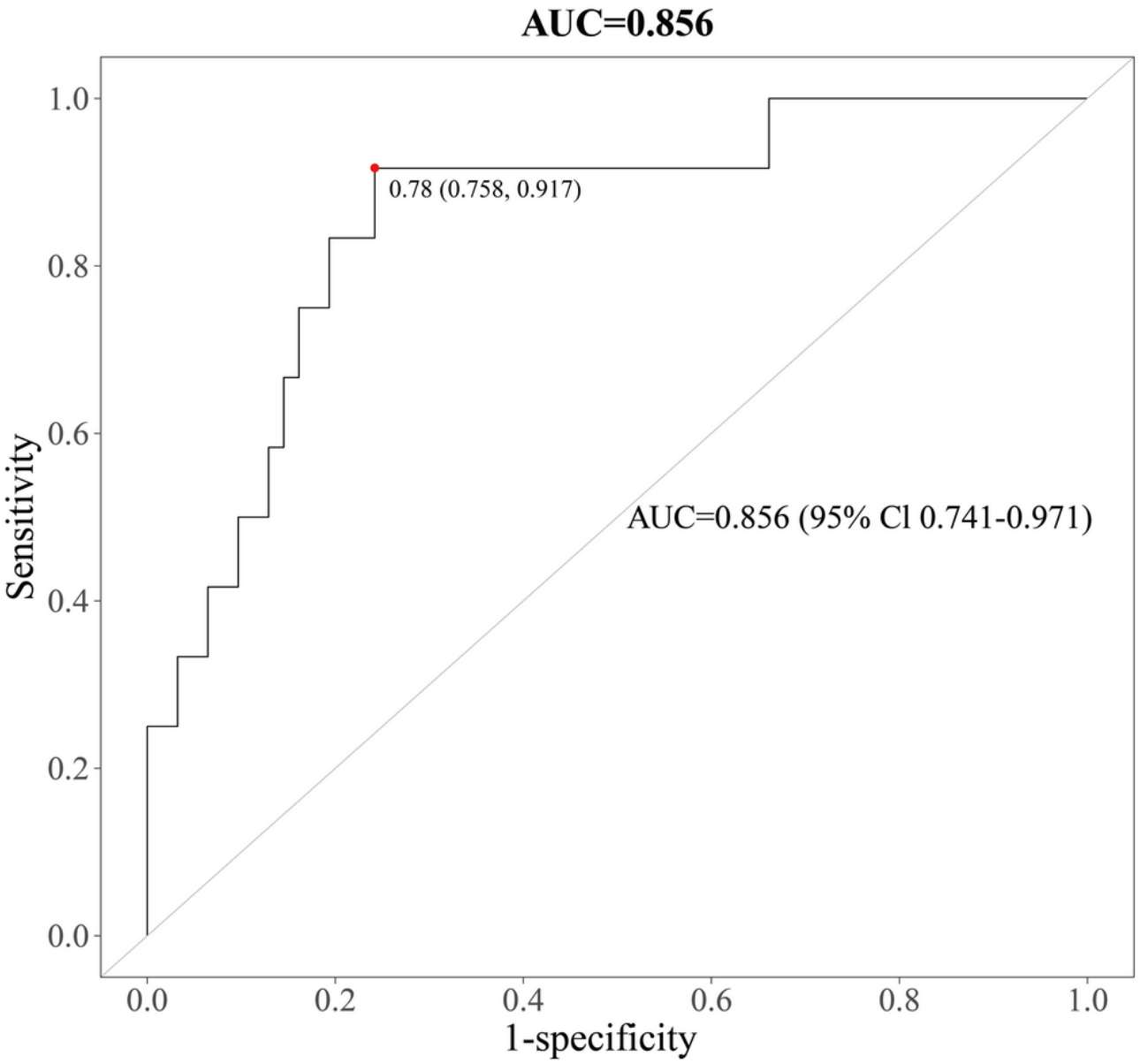
XSL•FO
RenderX

Figure 7. Receiver operating characteristic (ROC) curve of the logistic regression model based on 16 core metabolites in the training cohort (area under the ROC curve [AUC]=0.844, 95% CI 0.776 - 0.912).



XSL•FO
RenderX

Figure 8. Receiver operating characteristic (ROC) curve of the logistic regression model based on 16 core metabolites in the independent test cohort (area under the ROC curve [AUC]=0.856, 95% CI 0.741–0.971).

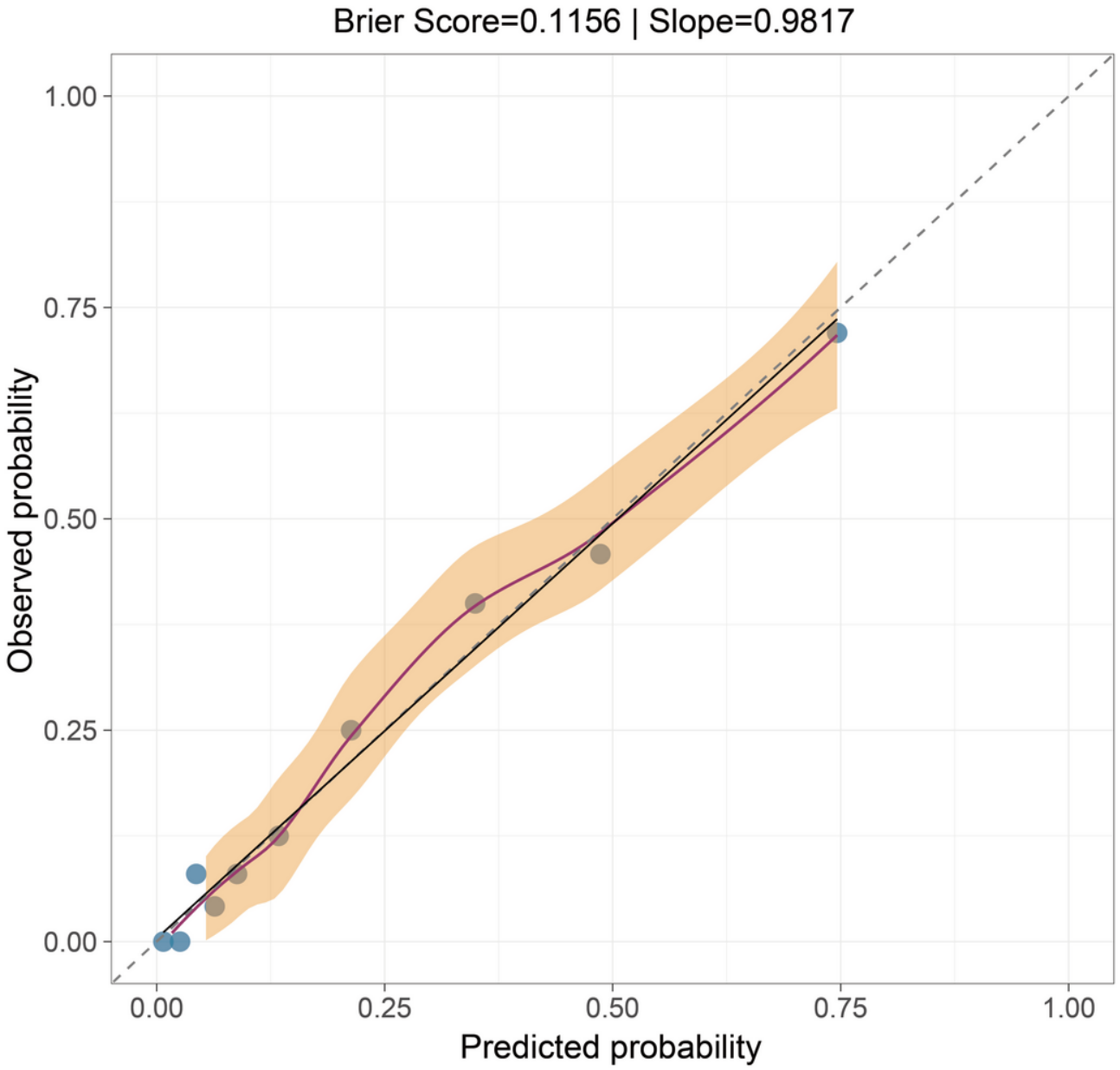


Comprehensive Evaluation of Model Performance: Discrimination, Calibration, and Clinical Utility

The final LR model was subjected to a comprehensive evaluation on the independent test set (n=78). The model demonstrated good overall accuracy, with a Brier score of 0.1156. The calibration plot (Figure 9) indicated excellent agreement between predicted probabilities and observed outcomes, with a calibration slope of 0.9817.

XSL•FO
RenderX

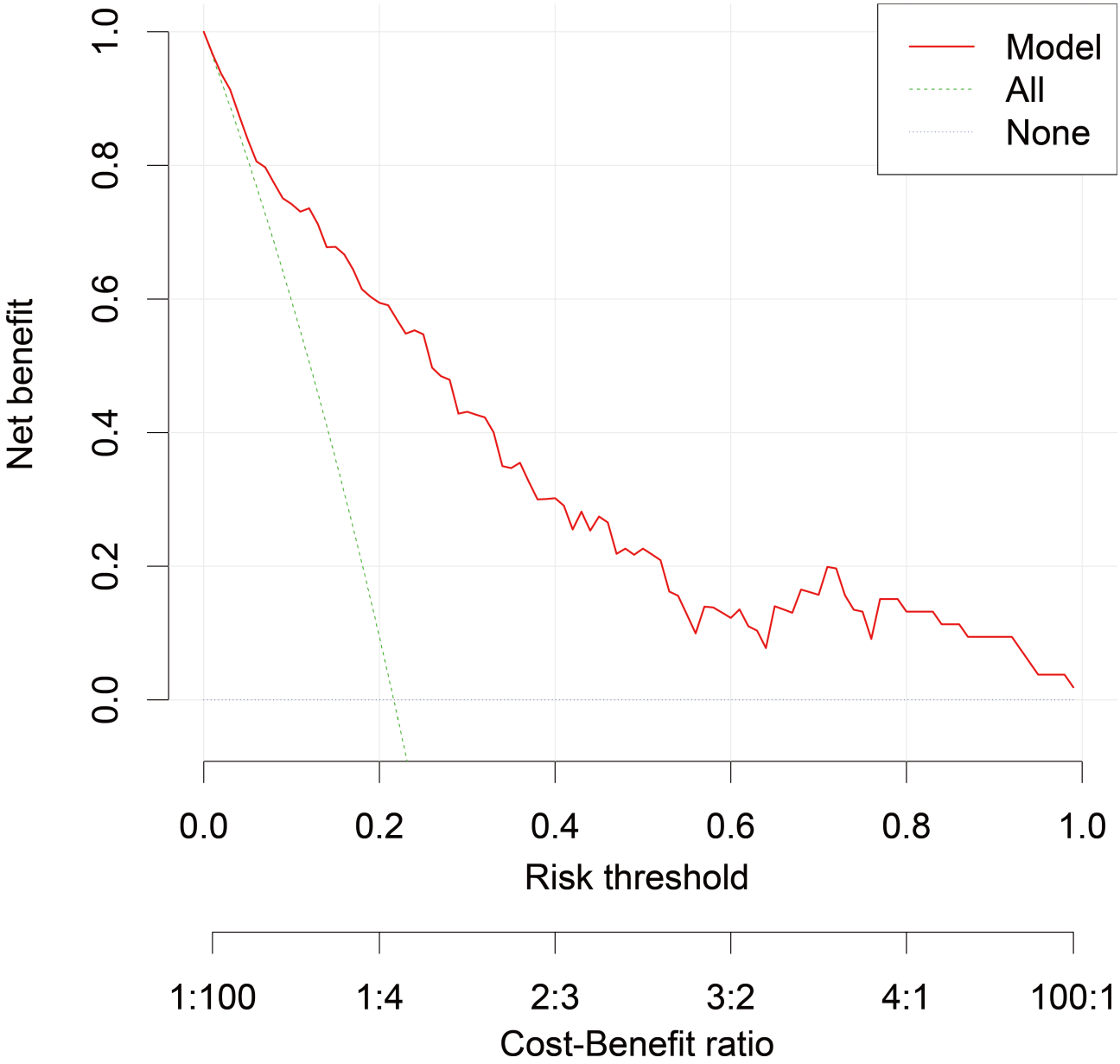
Figure 9. Calibration performance of the postoperative delirium prediction model, showing excellent fit (as indicated by the Brier score and calibration slope) between predictions and observations.



Decision curve analysis (Figure 10) revealed that using the metabolomics model for clinical decision-making provided a superior net benefit compared to strategies of intervening on all patients or intervening on no patients across a wide range of clinically reasonable probability thresholds (approximately 10% to 60%). This underscores the potential clinical utility of the model for targeting preventive interventions.

XSL•FO
RenderX

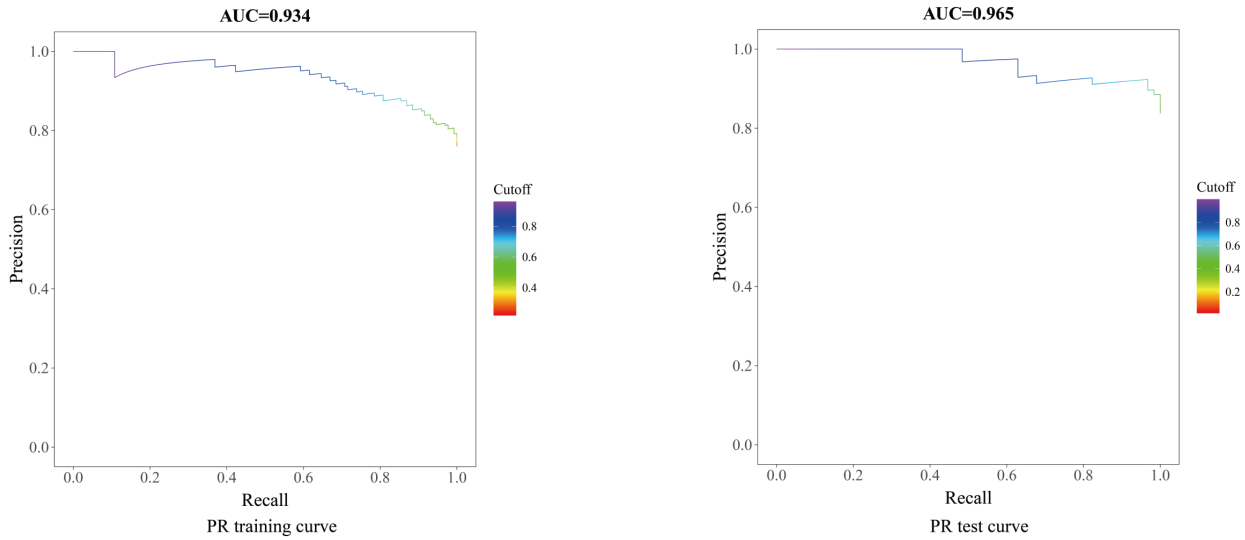
Figure 10. Decision curve analysis demonstrating the clinical utility of the metabolomics model, which provides a positive net benefit over a wide range of probability thresholds for postoperative delirium prediction.



In light of the class imbalance, PR curves were analyzed (Figure 11). The model achieved an AP of 0.965 on the test set (training set AP=0.934). This exceptionally high AP, in the context of a 21.5% (56/260) event rate, confirms the model's robust capability to identify patients with POD with high precision and recall. The PR curve is particularly informative for this imbalanced classification task as it focuses on the performance regarding the positive class (POD), unlike the ROC curve, which can be overly optimistic under class imbalance.

XSL•FO
RenderX

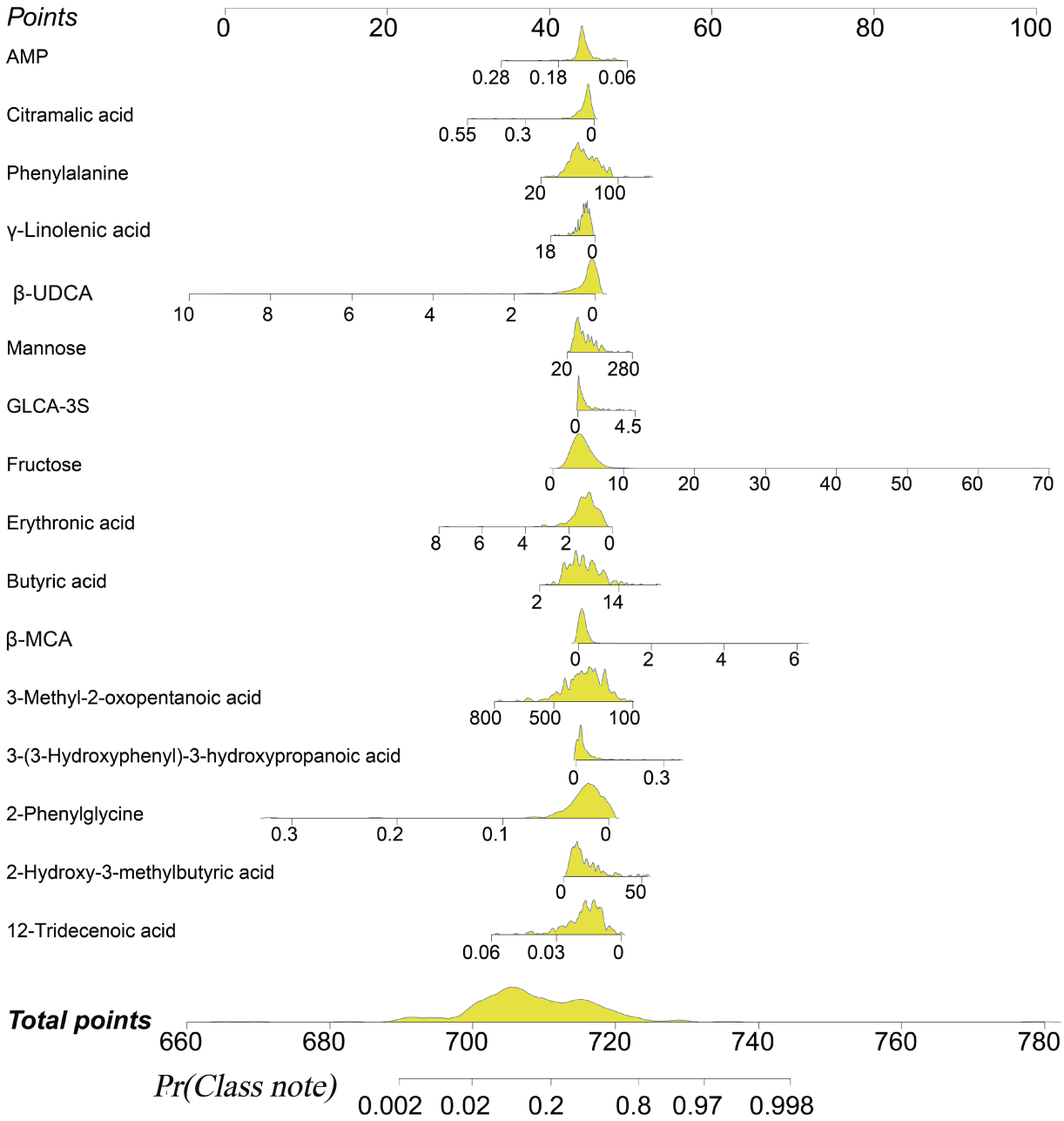
Figure 11. Precision-recall (PR) curves for the postoperative delirium prediction model. The high average precision scores on both the training and test sets confirm robust performance in identifying the minority class despite data imbalance. AUC: area under the receiver operating characteristic curve.



To facilitate clinical application, a nomogram was constructed based on the final LR model (Figure 12), allowing for the visual calculation of an individual patient's POD risk using the 16 preoperative metabolite levels.

XSL•FO
RenderX

Figure 12. Nomogram for individualized postoperative delirium (POD) risk prediction derived from the logistic regression model incorporating 16 preoperative metabolites. The total points from all metabolites correspond to an individual's predicted POD probability. β-MCA: β-muricholic acid; β-UDCA: β-ursodeoxycholic acid; AMP: adenosine monophosphate; GLCA-3S: glycolithocholic acid 3-sulfate.



Discussion

Principal Findings

POD, a severe neurological complication in older adult patients, is associated with neuroinflammation, oxidative stress, energy metabolism dysfunction, and cholinergic system imbalance [31]. Using a multicenter high-throughput targeted metabolomics approach combined with machine learning, we identified 16 differentially expressed metabolites as key biomarkers for constructing a POD prediction model.

Energy and Carbohydrate Metabolism

Altered levels of fructose and mannose indicate glucose metabolism dysregulation. Older adult patients often have insulin resistance, and postoperative stress exacerbates the imbalance between glycolysis and oxidative phosphorylation, leading to insufficient energy supply in the CNS [32]. Adenosine monophosphate, a key energy sensor, reflects mitochondrial dysfunction—overactivation of the adenosine monophosphate–activated protein kinase pathway under energy stress induces neuronal energy crisis linked to cognitive impairment in delirium [33].

Amino Acid Metabolism Dysregulation

Accumulation of phenylalanine, possibly due to reduced hepatic metabolism in older adults, disrupts the synthesis of catecholamine neurotransmitters (eg, dopamine and norepinephrine), exacerbating neuropsychiatric symptoms in

XSL•FO
RenderX

POD [34]. 2-Phenylglycine, a phenylalanine derivative, suggests dysregulation of aromatic amino acid metabolism associated with oxidative stress–induced neurotoxicity [35]. 3-Methyl-2-oxopentanoic acid, an intermediate of branched-chain amino acid metabolism, indicates dysfunction of the mitochondrial branched-chain ketoacid dehydrogenase complex, leading to CNS energy deficiency and neuroinflammatory cascades [35].

Lipid and Bile Acid Metabolism Disorders

Abnormal γ-linolenic acid levels, a polyunsaturated fatty acid, reflect membrane lipid homeostasis disruption, promoting proinflammatory mediator synthesis and neuroinflammation—a core mechanism of POD [36]. Butyric acid, a gut microbiota–derived short-chain fatty acid, regulates microglia anti-inflammatory function via the gut-brain axis; postoperative gut dysbiosis in older adult patients may reduce butyric acid, weakening its suppression of neuroinflammation [37]. β-Ursodeoxycholic acid and β-muricholic acid are associated with altered blood-brain barrier permeability; bile acids regulate neuroinflammation through the farnesoid X receptor and Takeda G protein–coupled receptor 5, and their imbalance activates astrocytes and impairs synaptic function [38]. Elevated glycolithocholic acid 3-sulfate suggests reduced hepatic detoxification, leading to toxic bile acid accumulation in the brain and neuronal injury [39].

Organic Acids and Oxidative Stress

Increased erythronic acid is linked to polyol pathway activation under oxidative stress, causing neuronal osmotic imbalance and reduced antioxidant capacity [40]. Citramalic acid, a TCA cycle intermediate, indicates mitochondrial citrate transporter dysfunction, leading to energy metabolism disorders and excessive reactive oxygen species production, intensifying neuronal oxidative damage [41]. 2-Hydroxy-3-methylbutyric acid, a leucine metabolite, is associated with enhanced catabolism under postoperative stress, indirectly reflecting the impact of systemic inflammatory load on the CNS [42]. Accumulation of 3-(3-hydroxyphenyl)-3-hydroxypropanoic acid, a phenylalanine metabolite, may result from abnormal phenylalanine hydroxylase activity, increasing neurotoxic metabolite production [43]. The role of 12-tridecenoic acid, a long-chain fatty acid, remains unclear but may involve membrane lipid regulation or inflammatory mediator synthesis, warranting further investigation [44].

Integration of Innovation Frameworks and Translational Implications

This study presents a distinctive research paradigm that integrates core principles from modern innovation frameworks. Our approach aligns with design thinking by centering on the clinical challenge of POD and iteratively developing an artificial intelligence–metabolomics model, with the resulting 16-metabolite panel representing a validated minimum viable product. This study follows the Lean Startup ethos through stringent RF- and LASSO-based feature selection to build a parsimonious yet powerful predictor (AUC>0.85), avoiding overengineering while enhancing clinical feasibility. Ultimately, this tool advances precision medicine by enabling individualized risk quantification based on objective metabolic profiles, facilitating a shift from reactive treatment to proactive prevention and demonstrating high translational potential in perioperative care.

This study also has limitations. Despite rigorous internal validation, the model requires validation in an external, geographically distinct cohort as our multicenter sample was drawn solely from Shanghai, potentially limiting generalizability. The model's scalability, while aided by its objective metabolic inputs, faces challenges including platform accessibility, ethnic variations in metabolic baselines requiring recalibration, and integration into diverse clinical workflows. Future work will focus on external validation across broader populations and surgical subtypes to advance clinical translation. Moreover, we agree that the clinical translation of our metabolomics-only model faces practical constraints regarding cost, turnaround time, and standardization. More fundamentally, as POD is multifactorial, a model relying solely on metabolites cannot incorporate key clinical risk factors such as frailty or comorbidity burden. To address these limitations, our primary future plan is to integrate this metabolomics panel with essential clinical variables to build a multimodal model. This integration is expected to enhance clinical practicality and robustness while maintaining high predictive performance. Furthermore, we have planned a prospective, external validation study to rigorously assess the model's generalizability.

Conclusions

This multicenter metabolomics study identified 16 differentially expressed metabolites involved in core pathological pathways of POD in older adult patients, including energy metabolism, amino acid homeostasis, bile acid regulation, and oxidative stress. These metabolites provide multidimensional biomarkers for early POD prediction and novel targets for interventions (eg, improving mitochondrial function and regulating gut microbiota). The machine learning–based prediction model has potential for integration into perioperative management to achieve precise POD prevention through dynamic metabolic profiling. However, to facilitate its clinical translation, future studies should focus on (1) integrating these metabolites with readily available key clinical variables to develop a more universally applicable and practical multimodal prediction tool; (2) rigorously validating the refined model in independent, multiregional external cohorts to assess its real-world performance; and (3) further elucidating the potential causal associations between these metabolites and the development of POD, and exploring their specific roles across different surgical subtypes.

Acknowledgments

The authors acknowledge Metabo-Profile Biotechnology (Shanghai) Co, Ltd, for providing the high-throughput targeted metabolomics platform.

XSL•FO
RenderX

Funding

This study was supported by the Hospital-Level Clinical Research Project of the Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (ynhg202131), the Medical Innovation Research Project of Shanghai Science and Technology Commission (21Y11904600), and the Young Scientific Research Project of the Anesthesia and Pain Special Committee of Shanghai Society of Integrated Traditional Chinese and Western Medicine (2021-469).

Data Availability

The original data of this study are not publicly available but are available from the corresponding author on reasonable request. Access to the data will be granted to researchers who provide a clear and justified purpose for their request in alignment with ethical guidelines and data protection regulations.

Authors' Contributions

YG, QZ, and XH conceived and designed the study. GZ, CL, DH, DW, XH, and MS performed sample collection. YG, GZ, and XH established the biobank. YG collaborated with Metabo-Profile [45] for data analysis and interpretation, whereas Metabo-Profile conducted the high-throughput targeted metabolomics sample analysis. YG, GZ, and XH drafted the manuscript. All authors critically reviewed the manuscript and approved its final version for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials including detailed ultrahigh-performance liquid chromatography-mass spectrometry/mass spectrometry parameters, study workflow diagram, and complete list of 16 candidate metabolic biomarkers.
[PDF File, 449 KB - aging_v9i1e78495_app1.pdf]

References

1. Zywiel MG, Hurley RT, Perruccio AV, Hancock-Howard RL, Coyte PC, Rampersaud YR. Health economic implications of perioperative delirium in older patients after surgery for a fragility hip fracture. J Bone Joint Surg Am 2015 May 20;97(10):829-836. [doi: 10.2106/JBJS.N.00724] [Medline: 25995494]
2. Mehta S, Cook D, Devlin JW, et al. Prevalence, risk factors, and outcomes of delirium in mechanically ventilated adults. Crit Care Med 2015 Mar;43(3):557-566. [doi: 10.1097/CCM.0000000000000727] [Medline: 25493968]
3. Oh ES, Fong TG, Hshieh TT, Inouye SK. Delirium in older persons: advances in diagnosis and treatment. JAMA 2017 Sep 26;318(12):1161-1174. [doi: 10.1001/jama.2017.12067] [Medline: 28973626]
4. Gleason LJ, Schmitt EM, Kosar CM, et al. Effect of delirium and other major complications on outcomes after elective surgery in older adults. JAMA Surg 2015 Dec;150(12):1134-1140. [doi: 10.1001/jamasurg.2015.2606] [Medline: 26352694]
5. Gottschalk A, Hubbs J, Vikani AR, Gottschalk LB, Sieber FE. The impact of incident postoperative delirium on survival of elderly patients after surgery for hip fracture repair. Anesth Analg 2015 Nov;121(5):1336-1343. [doi: 10.1213/ANE.0000000000000576] [Medline: 25590791]
6. Card E, Pandharipande P, Tomes C, et al. Emergence from general anaesthesia and evolution of delirium signs in the post-anaesthesia care unit. Br J Anaesth 2015 Sep;115(3):411-417. [doi: 10.1093/bja/aeu442] [Medline: 25540068]
7. Stukenberg S, Franck M, Spies CD, Neuner B, Myers I, Radtke FM. How can postoperative delirium be predicted in advance? A secondary analysis comparing three methods of early assessment in elderly patients. Minerva Anestesiol 2016 Jul;82(7):751-759. [Medline: 27453295]
8. Guenther U, Koegl F, Theuerkauf N, et al. Nursing workload indices TISS-10, TISS-28, and NEMS: higher workload with agitation and delirium is not reflected [Article in German]. Med Klin Intensivmed Notfmed 2016 Feb;111(1):57-64. [doi: 10.1007/s00063-015-0056-5] [Medline: 26253516]
9. Maldonado JR. Neuropathogenesis of delirium: review of current etiologic theories and common pathways. Am J Geriatr Psychiatry 2013 Dec;21(12):1190-1222. [doi: 10.1016/j.jagp.2013.09.005] [Medline: 24206937]
10. Patti GJ, Yanes O, Siuzdak G. Innovation: metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol 2012 Mar 22;13(4):263-269. [doi: 10.1038/nrm3314] [Medline: 22436749]
11. Nicholson JK, Lindon JC. Systems biology: metabonomics. Nature 2008 Oct 23;455(7216):1054-1056. [doi: 10.1038/4551054a] [Medline: 18948945]
12. Ivanisenko VA, Rogachev AD, Makarova AL, et al. AI-assisted identification of primary and secondary metabolomic markers for postoperative delirium. Int J Mol Sci 2024 Nov 4;25(21):11847. [doi: 10.3390/ijms252111847] [Medline: 39519398]
13. Maldonado JR. Delirium pathophysiology: an updated hypothesis of the etiology of acute brain failure. Int J Geriatr Psychiatry 2018 Nov;33(11):1428-1457. [doi: 10.1002/gps.4823] [Medline: 29278283]

XSL•FO
RenderX

14. Huang H, Han J, Li Y, et al. Early serum metabolism profile of post-operative delirium in elderly patients following cardiac surgery with cardiopulmonary bypass. Front Aging Neurosci 2022;14:857902. [doi: 10.3389/fnagi.2022.857902] [Medline: 35754961]
15. Huo J, Han S, Hao X, et al. Alterations in the gut microbiome and metabolome in elderly patients with postoperative delirium: a prospective nested case-control study. J Clin Anesth 2025 Apr;103:111833. [doi: 10.1016/j.jclinane.2025.111833] [Medline: 40228374]
16. Tripp BA, Dillon ST, Yuan M, et al. Targeted metabolomics analysis of postoperative delirium. Sci Rep 2021 Jan 15;11(1):1521. [doi: 10.1038/s41598-020-80412-z] [Medline: 33452279]
17. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015 Feb 26;372(9):793-795. [doi: 10.1056/NEJMp1500523] [Medline: 25635347]
18. Venter JC, Smith HO, Adams MD. The sequence of the human genome. Clin Chem 2015 Sep;61(9):1207-1208. [doi: 10.1373/clinchem.2014.237016] [Medline: 26185218]
19. Long T, Hicks M, Yu HC, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet 2017 Apr;49(4):568-578. [doi: 10.1038/ng.3809] [Medline: 28263315]
20. Hou YC, Yu HC, Martin R, et al. Precision medicine integrating whole-genome sequencing, comprehensive metabolomics, and advanced imaging. Proc Natl Acad Sci U S A 2020 Feb 11;117(6):3053-3062. [doi: 10.1073/pnas.1909378117] [Medline: 31980526]
21. Qing Y, Wang P, Cui G, et al. Targeted metabolomics reveals aberrant profiles of serum bile acids in patients with schizophrenia. Schizophrenia (Heidelb) 2022 Aug 18;8(1):65. [doi: 10.1038/s41537-022-00273-5] [Medline: 35982185]
22. Wang S, Wang J, Zhang R, et al. Association between serum haptoglobin and carotid arterial functions: usefulness of a targeted metabolomics approach. Cardiovasc Diabetol 2019 Jan 11;18(1):8. [doi: 10.1186/s12933-019-0808-2] [Medline: 30634984]
23. MahmoudianDehkordi S, Bhattacharyya S, Brydges CR, et al. Gut microbiome-linked metabolites in the pathobiology of major depression with or without anxiety-a role for bile acids. Front Neurosci 2022;16:937906. [doi: 10.3389/fnins.2022.937906] [Medline: 35937867]
24. Chinese Clinical Trial Registry (CHICTR). URL: https://www.chictr.org.cn/ [accessed 2026-04-20]
25. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med 1990 Dec 15;113(12):941-948. [doi: 10.7326/0003-4819-113-12-941] [Medline: 2240918]
26. Juan L, Yizhuang Z, Feng F, Yuying T. Revision of the delirium assessment method and testing of its reliability and validity [Article in Chinese]. J Clin Psychol Med 2003;13:147-149 [FREE Full text] [doi: 10.3969/j.issn.1005-3220.2003.03.009]
27. Xie G, Wang L, Chen T, et al. A metabolite array technology for precision medicine. Anal Chem 2021 Apr 13;93(14):5709-5717. [doi: 10.1021/acs.analchem.0c04686] [Medline: 33797874]
28. Sun Y, Zhang X, Hang D, et al. Integrative plasma and fecal metabolomics identify functional metabolites in adenoma-colorectal cancer progression and as early diagnostic biomarkers. Cancer Cell 2024 Aug 12;42(8):1386-1400. [doi: 10.1016/j.ccell.2024.07.005] [Medline: 39137727]
29. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996 Dec;49(12):1373-1379. [doi: 10.1016/s0895-4356(96)00236-3] [Medline: 8970487]
30. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988 Sep;44(3):837-845. [Medline: 3203132]
31. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. Lancet 2014 Mar 8;383(9920):911-922. [doi: 10.1016/S0140-6736(13)60688-1] [Medline: 23992774]
32. Titlestad I, Watne LO, Caplan GA, et al. Impaired glucose utilization in the brain of patients with delirium following hip fracture. Brain (Bacau) 2024 Jan 4;147(1):215-223. [doi: 10.1093/brain/awad296] [Medline: 37658825]
33. Lu Y, Chen L, Ye J, et al. Surgery/anesthesia disturbs mitochondrial fission/fusion dynamics in the brain of aged mice with postoperative delirium. Aging (Albany NY) 2020 Jan 12;12(1):844-865. [doi: 10.18632/aging.102659] [Medline: 31929114]
34. Watne LO, Idland AV, Fekkes D, et al. Increased CSF levels of aromatic amino acids in hip fracture patients with delirium suggests higher monoaminergic activity. BMC Geriatr 2016 Aug 2;16:149. [doi: 10.1186/s12877-016-0324-0] [Medline: 27484129]
35. Guo Y, Li Y, Zhang Y, et al. Post-operative delirium associated with metabolic alterations following hemi-arthroplasty in older patients. Age Ageing 2019 Dec 1;49(1):88-95. [doi: 10.1093/ageing/afz132] [Medline: 31711096]
36. Qiu Y, Mo C, Li J, et al. Acute changes in hippocampal metabolism after anesthesia and surgery: implications for perioperative neurocognitive disorder. Biochem Biophys Res Commun 2024 Dec 3;736:150492. [doi: 10.1016/j.bbrc.2024.150492] [Medline: 39116679]
37. Xu F, Chen H, Gao Y, Yang X, Zhang C, Ni X. Sodium butyrate ameliorates postoperative delirium by regulating gut microbiota dysbiosis to inhibit astrocyte activation in aged mice. Neurochem Res 2024 Dec;49(12):3342-3355. [doi: 10.1007/s11064-024-04245-2] [Medline: 39340594]

XSL•FO
RenderX

38. Fiaschini N, Mancuso M, Tanori M, et al. Liver steatosis and steatohepatitis alter bile acid receptors in brain and induce neuroinflammation: a contribution of circulating bile acids and blood-brain barrier. Int J Mol Sci 2022 Nov 17;23(22):14254. [doi: 10.3390/ijms232214254] [Medline: 36430732]
39. Williams E, Chu C, DeMorrow S. A critical review of bile acids and their receptors in hepatic encephalopathy. Anal Biochem 2022 Apr 15;643:114436. [doi: 10.1016/j.ab.2021.114436] [Medline: 34715070]
40. Ye S, Cheng Z, Zhuo D, Liu S. Different types of cell death in diabetic neuropathy: a focus on mechanisms and therapeutic strategies. Int J Mol Sci 2024 Jul 25;25(15):8126. [doi: 10.3390/ijms25158126] [Medline: 39125694]
41. Motori E, Atanassov I, Kochan SM, et al. Neuronal metabolic rewiring promotes resilience to neurodegeneration caused by mitochondrial dysfunction. Sci Adv 2020 Aug;6(35):eaba8271. [doi: 10.1126/sciadv.aba8271] [Medline: 32923630]
42. Bain CR, Myles PS, Corcoran T, Dieleman JM. Postoperative systemic inflammatory dysregulation and corticosteroids: a narrative review. Anaesthesia 2023 Mar;78(3):356-370. [doi: 10.1111/anae.15896] [Medline: 36308338]
43. Fitzpatrick PF. The aromatic amino acid hydroxylases: structures, catalysis, and regulation of phenylalanine hydroxylase, tyrosine hydroxylase, and tryptophan hydroxylase. Arch Biochem Biophys 2023 Feb;735:109518. [doi: 10.1016/j.abb.2023.109518] [Medline: 36639008]
44. Chung HL, Ye Q, Park YJ, et al. Very-long-chain fatty acids induce glial-derived sphingosine-1-phosphate synthesis, secretion, and neuroinflammation. Cell Metab 2023 May 2;35(5):855-874. [doi: 10.1016/j.cmet.2023.03.022] [Medline: 37084732]
45. Metabo-Profile. URL: https://www.metaboprofile.com/ [accessed 2026-04-20]

Abbreviations

AP: average precision
AUC: area under the receiver operating characteristic curve
CAM: Confusion Assessment Method
CNS: central nervous system
EPV: event per variable
LASSO: least absolute shrinkage and selection operator
LR: logistic regression
MMSE: Mini-Mental State Examination
POD: postoperative delirium
PR: precision-recall
QC: quality control
RF: random forest
ROC: receiver operating characteristic
TCA: tricarboxylic acid
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
UPLC: ultrahigh-performance liquid chromatography

Edited by R Bjarnadottir; submitted 03.Jun.2025; peer-reviewed by Q Wang, S Mohanadas; revised version received 25.Feb.2026; accepted 27.Feb.2026; published 30.Apr.2026.

Please cite as:
Zhong G, Huang X, Li C, Wang D, Huang D, Sun M, Zhou Q, Guo Y
Integrated High-Throughput Targeted Metabolomics and Machine Learning for Early Prediction and Prevention of Postoperative Delirium in Older Adult Surgical Patients: Prospective Multicenter Cohort Study
JMIR Aging 2026;9:e78495
URL: https://aging.jmir.org/2026/1/e78495
doi:10.2196/78495

© Gengrui Zhong, Xiaoli Huang, Congye Li, Deqiang Wang, Dingding Huang, Menghan Sun, Quanhong Zhou, Yong Guo. Originally published in JMIR Aging (https://aging.jmir.org), 30.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Web-Based Personalized Machine Learning Recommendations to Enhance Shared Decision-Making in Prostate-Specific Antigen Screening: Randomized Controlled Trial

Yi-Ting Lin[1,2*], MD, PhD; Yen Chun Huang[3*], PhD; Chih Kuang Liu[4], MD, PhD; Hsiao-Yun Cho[4,5], MD; Mingchih Chen[4], PhD

[1]Department of Urology, St. Joseph's Hospital, Yunlin, Taiwan

[2]College of Nursing, Meiho University, Pingtung, Taiwan

[3]Department of Artificial Intelligence, Tamkang University, New Taipei City, Taiwan

[4]Graduate Institute of Business Administration, College of Management, Fu Jen Catholic University, No.510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City, New Taipei City, Taiwan

[5]Department of Otorhinolaryngology—Head and Neck Surgery, Fu Jen Catholic University Hospital, Fu Jen Catholic University, New Taipei City, Taiwan

[*]these authors contributed equally

Corresponding Author:
Mingchih Chen, PhD
Graduate Institute of Business Administration, College of Management, Fu Jen Catholic University, No.510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City, New Taipei City, Taiwan

Abstract

Background: Prostate-specific antigen (PSA) screening involves complex trade-offs between early detection and the risks of overdiagnosis. For older adults (aged ≥50 years), shared decision-making (SDM) is often hindered by limited health literacy, sensory or cognitive impairments, and multimorbidity, which complicate risk comprehension. Traditional decision aids provide foundational knowledge but are often nonpersonalized. Machine learning (ML) may offer individualized recommendations, yet the psychological and behavioral effects of ML-assisted SDM in geriatric populations remain poorly characterized.

Objective: This study aimed to develop and evaluate a web-based, ML-driven decision aid integrated into an SDM workflow to provide personalized PSA screening recommendations and to assess its effects on decisional conflict (primary outcome), state anxiety, and decision satisfaction among middle-aged and older men.

Methods: The study followed a 2-stage design. First, a model establishment group (n=507) was used to train and evaluate 6 ML algorithms based on clinical and values-clarification data. A random forest model was selected for its superior performance (mean area under the curve 0.933, SD 0.350; 95% CI 0.902-0.963). Second, a randomized controlled trial was conducted with 367 participants (mean age 64.34, SD 10.30 years) randomly assigned 1:1 to the ML suggestion group (MLSG; n=185) or the control group (CG; n=182). Both groups received video-based education, counseling, and values clarification; only the MLSG received an ML-generated "second opinion" recommendation. Primary and secondary outcomes were assessed using the Decisional Conflict Scale (DCS), Spielberger State-Trait Anxiety Inventory (STAI), and Satisfaction with Decision scale.

Results: In the randomized controlled trial (n=367), the MLSG reported significantly lower decisional conflict than the CG (total DCS score: mean difference [MD] –3.77, 95% CI –5.55 to −1.99; Cohen *d*=–0.44; P<.001). The MLSG reported greater perceived support (DCS7: adjusted *P*=.03), more adequate advice (DCS9: adjusted *P*<.001), and higher decision confidence (DCS10: adjusted *P*=.03; DCS11: adjusted *P*<.001). Regarding psychological well-being, although total anxiety scores did not differ, the MLSG reported reduced worry (STAI item 6: MD –0.98, 95% CI –1.20 to −0.76; *d*=–0.89; adjusted *P*<.001) and increased calmness (STAI item 1: MD 0.30, 95% CI 0.06-0.54; *d*=0.25; adjusted *P*=.01). Decision satisfaction was higher in the MLSG across all items (total Satisfaction with Decision score: MD –7.38, 95% CI –8.54 to −6.18; *P*<.001). Behavioral choices were strongly influenced by the ML recommendation: participants in the MLSG who received an "accept" recommendation were more likely to select "accept" (34/67, 50.7%) than those in the CG (44/182, 24.2%; *P*<.001). When the system suggested "not now," only 17.8% (21/118) chose "accept," which was lower than in the CG.

Conclusions: Integrating personalized ML recommendations into SDM workflows provides emotional scaffolding for older men, reducing decisional distress and enhancing confidence without undermining autonomy. By addressing geriatric-specific vulnerabilities through a facilitated digital interface, this ML-driven approach complements traditional clinical consultations. These findings support the scalable integration of artificial intelligence–assisted decision support to foster patient-centered care in aging populations.

XSL•FO
RenderX

Trial Registration: Chinese Clinical Trial Registry ChiCTR2000034126; https://www.chictr.org.cn/showprojEN.html?proj=55682

(JMIR Aging 2026;9:e83238) doi:10.2196/83238

KEYWORDS

prostate cancer; PSA screening; prostate-specific antigen; shared decision-making; decision aids; machine learning; personalized recommendations; decisional conflict; patient-centered care

Introduction

Over the past 2 decades, shared decision-making (SDM) has become a cornerstone of patient-centered care, enabling individuals to actively participate in health care choices [1,2]. In this process, clinicians and patients exchange information, deliberate on alternatives, and align decisions with patient values [3-5]. Guidelines generally outline 5 key steps in SDM: engaging the patient, presenting options, clarifying values, making the decision, and reviewing outcomes [6]. This structured approach has been shown to not only enhance patient satisfaction but also reduce decisional conflict [7-9].

Prostate cancer remains one of the most common malignancies among men globally. However, the decision to undergo prostate-specific antigen (PSA) screening involves complex trade-offs between the benefits of early detection and the risks of overdiagnosis or overtreatment [10-12]. Among older adults, SDM is indispensable because health literacy limitations, sensory or cognitive impairments, and multimorbidity can hinder risk comprehension and value clarification, thereby increasing reliance on physician-directed choices [13-16]. In time-pressured encounters, lower educational attainment further diminishes active patient participation, shifting decisions toward clinician preferences unless patient values are explicitly elicited and supported through tailored explanations [3,15-17].

Consequently, strategies to strengthen SDM for older adults should prioritize cultivating strong physician-patient relationships, explicitly addressing psychological determinants of choice, and implementing patient-tailored decision aids. Collectively, these approaches have been shown to improve knowledge without increasing anxiety and can sustain reductions in decisional conflict over time, helping patients critically appraise the benefits, risks, and uncertainties of PSA testing [18-20].

Efforts to support SDM in PSA screening have primarily relied on traditional decision aids (DAs), such as educational brochures, interactive questionnaires, and electronic platforms or videos [3,6,21]. Although machine learning (ML) approaches have gained traction in clinical practice, their application has largely focused on risk prediction rather than personalized decision support. Notably, fully personalized electronic decision interfaces remain underdeveloped [22-24]. Furthermore, few studies have explicitly evaluated the impact of ML-assisted recommendations on psychological outcomes—such as anxiety, uncertainty, and decisional conflict—which are critical indicators of acceptability and long-term efficacy.

To address these gaps, we developed a novel ML-based decision aid designed as a complementary component within existing SDM workflows for PSA screening consultations. Distinct from traditional aids, this tool integrates individualized risk profiles with structured preference elicitation to align recommendations with both patient values and clinical outcomes observed in comparable cohorts, thereby facilitating preference-concordant choices in geriatric care.

The primary objective of this study was to evaluate whether this ML-integrated strategy reduces decisional conflict (primary outcome) among middle-aged and older men compared to standard care. Secondary objectives included assessing its impact on decision self-efficacy, state anxiety, and decision regret, as well as its potential to improve informed choice and downstream quality-of-life outcomes. These findings will inform the scalable integration of ML-based recommendations into routine PSA screening workflows, with the ultimate goal of reducing overtreatment while empowering informed patient choice.

Methods

Establishment of a Preference-Sensitive DA

Drawing on the frameworks proposed by Eila et al [18], Rychetnik et al [25], and Ferrante et al [26], we developed the "importance for physiological and psychological impact" (IPPI) questionnaire to systematically evaluate participants' values and preferences regarding concerns related to PSA screening. The instrument consisted of 10 items across two dimensions: 4 assessing physiological impacts and 6 evaluating psychological impacts. The full questionnaire is available in Multimedia Appendix 1.

Validation of Questionnaires

This 10-item questionnaire underwent expert validity assessment by 10 urologists, each with more than 7 years of clinical experience. Reliability analysis, based on data from the first 300 patients, demonstrated strong internal consistency, with Cronbach α values of 0.838 for the physiological impact dimension and 0.900 for the psychological impact dimension.

Factor analysis of the IPPI was conducted using the maximum likelihood method, followed by Varimax rotation. Maximum likelihood estimation was selected for its superior parameter estimation and statistical testing capabilities [27]. Varimax rotation was used to maximize the variance of squared factor loadings, creating a simpler and more interpretable factor structure that approaches the Thurstone principle of simple structure [28,29]. This orthogonal rotation preserves factor independence, providing greater statistical stability and replicability compared to oblique methods [29]. The analysis extracted two factors: personal values concerning physiological impacts and personal values concerning psychological impacts.

XSL•FO
RenderX

In addition, the Chinese versions of the Decisional Conflict Scale (DCS) and the International Prostate Symptom Score (IPSS) were validated and used in this study (Multimedia Appendix 2). The short form of the Spielberger State-Trait Anxiety Inventory (STAI) and the Decision Satisfaction Questionnaire were translated into Chinese by 2 independent bilingual experts. Comprehensive details of all research instruments are provided in Multimedia Appendices 3 and 4.

Participants Recruitment

This study used a closed, face-to-face recruitment strategy at Taiwan St. Joseph Hospital. We recruited men who were eligible if they (1) were aged 50 years or above, (2) presented for routine care between October 2017 and April 2018, and (3) were able to provide informed consent and communicate in Mandarin Chinese. This age threshold aligns with US Preventive Services Task Force and American Urological Association guidelines recommending PSA screening discussions starting at age 50, with SDM particularly emphasized for men aged 55 - 69 years [30,31]. Men were excluded if they had (1) any documented psychiatric disorder that might impair decision-making capacity, (2) prior PSA testing (to ensure treatment-naïve status), or (3) a history of prostate surgery or prostate cancer. Participant identity and eligibility were verified against electronic medical records by research staff to prevent multiple entries. Notably, computer and internet literacy were not required; although the intervention used a web-based interface, all assessments and digital components were administered face-to-face by trained research staff. This supervised approach ensured accessibility for older adults regardless of technological proficiency and guaranteed that the study team had direct contact with all participants.

Study Group Design Framework

A total of 520 participants were initially recruited for model development. After excluding 13 individuals due to low-quality responses—defined as ten consecutive identical scores or rhythmic scoring patterns—the remaining 507 participants formed the model establishment group (MEG) for comprehensive training and validation. To ensure progressive optimization, the ML model was iteratively retrained after every batch of 20 new participants. This development phase continued until the model achieved a predefined performance criterion (area under the curve [AUC] ≥0.80), indicating sufficient predictive accuracy for potential clinical application.

Subsequently, we conducted a parallel-group randomized controlled trial to evaluate the impact of ML-based decision aid recommendations. An additional 380 participants were recruited and randomly assigned with a 1:1 allocation ratio to either the ML suggestion group (MLSG, n=190) or the control group (CG, n=190). The randomization sequence was computer-generated and integrated directly into the web-based study platform to ensure allocation concealment. Group assignment was revealed to the research staff only after each participant's baseline data had been fully entered and locked, thereby preventing any foreknowledge of the allocation sequence. The MLSG received standard SDM support supplemented by ML-assisted recommendations, while the CG received standard SDM alone. After excluding participants with low-quality responses, the final analysis included 185 participants in the MLSG (n=5 excluded) and 182 in the CG (n=8 excluded). Prior to study commencement, all participating medical staff completed a standardized 10-minute training session to ensure proficiency in navigating the DA. The study was conducted in accordance with the original protocol, and there were no changes to the methods (including eligibility criteria and outcomes) after trial commencement.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and relevant ethical guidelines for research involving human subjects. The study protocol was reviewed and approved by the Institutional Review Board of Fu Jen Catholic University (approval C105142), and written informed consent was obtained from all participants before enrollment. The trial was registered in the Chinese Clinical Trial Registry (ChiCTR 2000034126). Only the patient medical record number was retained to protect privacy, and all study data were accessible exclusively to authorized research personnel; nonresearch staff were not permitted to access the dataset. To promote transparency and fairness, participants received a cash payment of NT $300 for their participation, which is equivalent to approximately US $9.38 based on an exchange rate of US $1=NT $31.97.

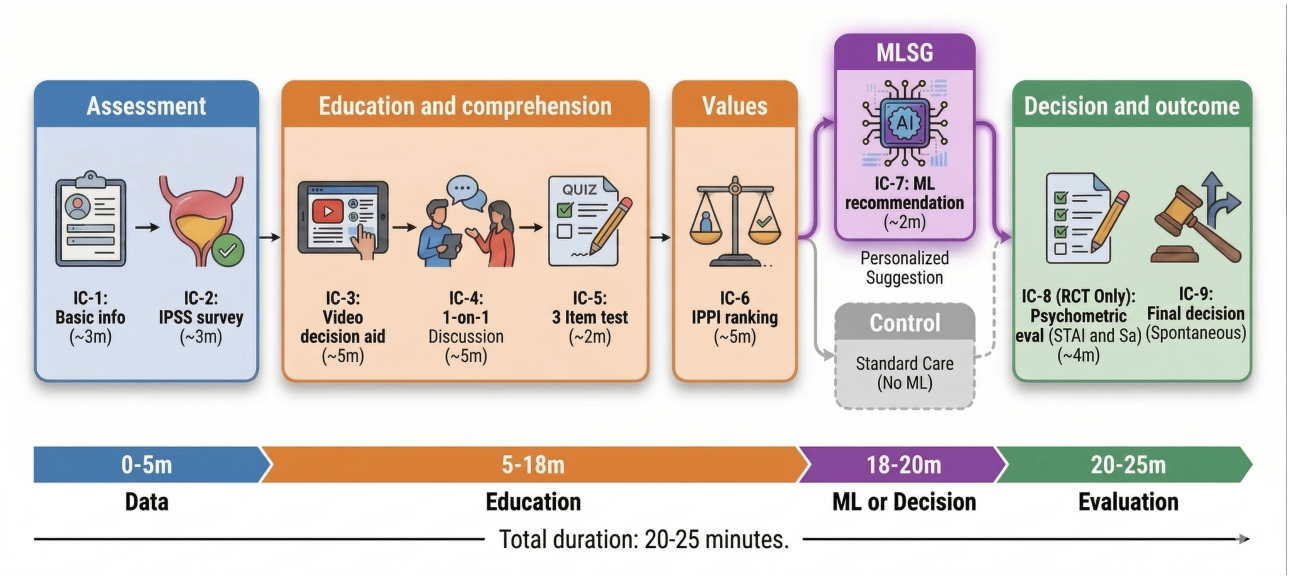
Intervention Specification

Intervention Flow and Core Components

The intervention comprised 9 standardized components delivered in approximately 20 - 25 minutes (Figure 1). Participants in the MEG completed components IC-1 to IC-6, whereas participants in the randomized controlled trial (RCT) completed the full sequence (IC-1 to IC-9). Initially, all participants provided baseline demographic and clinical data (IC-1; ~3 min), followed by a urinary symptom assessment using the IPSS (IC-2; ~2 min). The video-based decision aid (IC-3; ~5 min) then provided a structured introduction to PSA screening, sequentially covering (1) prostate physiology, (2) potential morbidity of prostate disease, and (3) benefits of early detection versus risks of overdiagnosis.

XSL•FO
RenderX

Figure 1. Prostate-specific antigen screening decision support intervention procedure. IC: informed consent; IPPI: importance for physiological and psychological impact; IPSS: International Prostate Symptom Score; ML: machine learning; MLSG: machine learning suggestion group; RCT: randomized controlled trial; Sa: Satisfaction with Decision scale; STAI: Spielberger State-Trait Anxiety Inventory.



At IC-4 (~5 min), trained medical personnel conducted individual counseling to ensure comprehension. Using neutral, nondirective language and the teach-back method, staff confirmed understanding and corrected misconceptions without soliciting patient preferences. Participants were explicitly reminded that the final screening decision remained entirely their prerogative. Subsequently, at IC-5, participants completed a brief 3-item knowledge assessment (true/false/unsure) covering screening accuracy, disease progression, and biopsy limitations (items detailed in Multimedia Appendix 5). At IC-6 (~3 min), participants engaged in the IPPI values-clarification exercise, rating the personal importance of screening attributes and ranking them by priority.

Facilitated ML Recommendation (IC-7; MLSG Only)

Data collected during IC-1 (baseline demographics and clinical history), IC-2 (IPSS urinary symptom assessment), and IC-6 (IPPI values clarification) were completed within the same web-based interface and were used as inputs to the deployed ML model. After finishing IC-1-IC-6, in the RCT (n=367), participants randomized to the MLSG received a personalized, on-screen recommendation at IC-7 (~2 min) via this dedicated interface. Figure 2 presents representative screenshots of the user interface (original traditional Chinese version), including the integrated data-entry and values-clarification screens (Figure 2) and the recommendation output screen (Figure 3). To support international readability and replicability, full English translations of all interface text and items are provided in Multimedia Appendix 6. The interface displayed the recommendation with the exact wording: “Your most likely choice is: [Immediately accept screening]” or “[Reconsider or refuse to undergo screening].” Clinicians conveyed this result using standardized, age-appropriate language, explicitly emphasizing its nature as an advisory “second opinion” based on the choices of similar patients, intended to support rather than replace the patient’s autonomy.

XSL•FO
RenderX

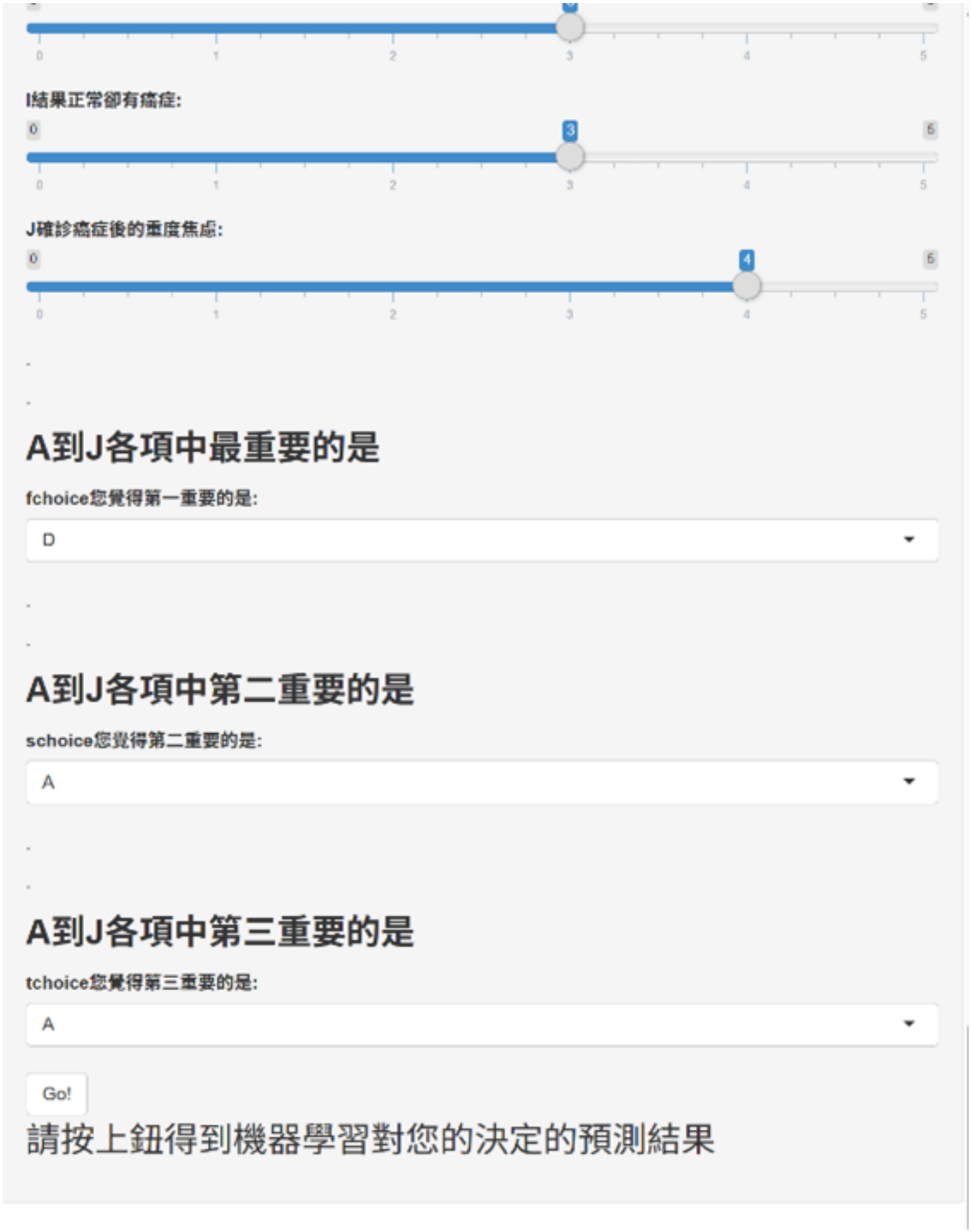
Figure 2. Screenshot of the web-based decision aid interface (original traditional Chinese version). Integrated participant interface for data entry. Full English translations are provided in Multimedia Appendix 6.



XSL•FO
RenderX

Figure 3. Screenshot of the web-based decision aid interface (original traditional Chinese version). Machine learning recommendation output screen (machine learning suggestion group only) displaying the personalized recommendation generated from participant inputs. Full English translations are provided in Multimedia Appendix 6.

XSL•FO
RenderX



XSL•FO
RenderX

Staff Training and Fidelity

To ensure digital equity and protocol adherence, the intervention was delivered in a facilitated setting. Prior to the study, all participating medical staff completed a comprehensive 10-minute training session focused on two key objectives: (1) technical proficiency in navigating the web-based decision aid to act as effective "digital navigators" and (2) adherence to a neutral facilitation protocol to avoid influencing participant decisions. Staff provided standardized, one-on-one technical assistance (eg, interface navigation and font enlargement) to any participant experiencing difficulty. This support was strictly operational, ensuring accessibility for older adults with varying technological proficiency while maintaining the neutrality of the SDM process. Periodic fidelity checks were conducted by the principal investigator to ensure standardized delivery.

Finally, all participants completed psychological measures (IC-8; ~4 min; STAI, SA, and DCS) and made their screening decision at IC-9 (self-paced).

Model Construction and Validation

Algorithm Selection and Optimization

To identify the optimal predictive model, we systematically benchmarked 5 baseline algorithms—logistic regression (LGR), random forest (RF), support vector machine, Extreme Gradient Boosting, and multilayer perceptron—alongside a deep neural network benchmark. All candidate models were evaluated using a rigorous 0.632 bootstrap protocol (n=507 iterations) to generate out-of-bag estimates of accuracy, sensitivity, specificity, and AUC. Technical specifications, including the hyperparameter optimization strategy (using design of experiments and ant lion optimizer), the full deep neural network architecture, and extended performance comparisons across all classifiers, are provided in Multimedia Appendix 7.

Model Deployment

Based on the comparative evaluation detailed in Multimedia Appendix 8, the RF model demonstrated the most robust performance, achieving the highest mean AUC of 0.933 (95% CI 0.902 - 0.963). Consequently, it was selected as the optimal algorithm for deployment. Using a standard probability threshold of 0.5 to generate binary recommendations ("accept" vs "not now"), the deployed model achieved a sensitivity of 0.69 and a specificity of 0.97. This high specificity was particularly valued to minimize false positive encouragement of screening, thereby reducing potential overdiagnosis. This finalized RF model was integrated into a secure web server (Shiny interface; Multimedia Appendix 6) to generate personalized recommendations for the intervention group.

Input Features and Outcome Definition

The deployed model used 17 input features derived from the MEG dataset (IC-1 to IC-6), including age, marital status, education, prior PSA knowledge, clinical history (eg, family history, prior biopsy), and IPPI-derived value priorities. Missing data were handled using complete case analysis, as the percentage of missing values was negligible (<1%) due to the forced-entry design of the digital interface. The prediction target was the participant's stated preference ("accept" vs "not now") recorded immediately after the values-clarification exercise but before viewing any ML recommendation, thus serving as an unbiased ground truth for supervised learning. Detailed definitions for all predictor variables are provided in Multimedia Appendix 9.

Statistical Analysis and Outcome Interpretation

Variable Scoring and Interpretation

To ensure consistent interpretation of group differences across psychological and decision-quality measures, scale directionality was defined as follows. For STAI, item direction varies: for negative affect items (STAI-2 tense, STAI-3 upset, STAI-6 worried), lower scores indicate lower anxiety; whereas for positive affect items (STAI-1 calm, STAI-4 relaxed, STAI-5 reassured), higher scores indicate greater calmness and reassurance. For the DCS, items were reverse scored (0 - 4; strongly agree=0, strongly disagree=4) such that higher scores indicate greater decisional conflict (worse decision quality). For Satisfaction with Decision (Sa), items were coded such that lower scores indicate greater satisfaction.

Statistical Methods

Continuous variables were compared using Student *t* tests, and categorical variables using χ^2 tests. To address multiplicity, we prioritized the total DCS score as the single primary end point; therefore, no adjustment was applied for this hypothesis. All other analyses, including subscales of the DCS, STAI, and Sa, were considered secondary and exploratory. For item-level comparisons within these scales, we controlled the false discovery rate using the Benjamini-Hochberg procedure to minimize type I errors. Adjusted *P* values are reported for these multiple comparisons.

Effect sizes were reported as Cohen *d* for continuous outcomes. Mean differences (MD) with 95% CIs were calculated. Statistical significance was defined as a 2-tailed $P<.05$. Analyses were performed using SPSS version 21.0 (IBM Corp) and R version 3.4.2 (R Core Team) (R package provided in Multimedia Appendix 10).

System Stability and Quality Control

The ML-based decision aid was deployed on a dedicated Shiny server. All intervention components (video content, questionnaires, IPPI interface, and the ML recommendation algorithm) were finalized and locked before randomization. No content modifications, algorithm updates, or bug fixes were implemented after the trial commenced. Study personnel remained consistent throughout the enrollment period, and no unexpected system failures affected data collection.

Results

Demographic Characteristics of the MEG Group

Baseline characteristics of the MEG cohort (n=507) are shown in Figure 4 and summarized in Table 1, with comprehensive results provided in Multimedia Appendix 11. Self-reported knowledge of PSA testing (KnowPSA score) was significantly higher in the accept group (mean 2.23, SD 0.81) than in the not now group (mean 2.07, SD 0.69; $P=.05$). Marital status

XSL•FO
RenderX

distribution also differed significantly between the 2 groups (P<.001), with a notably higher proportion of married individuals in the accept group. Regarding the IPSS, the accept group reported significantly higher scores for IPSS items 1‐5 (P values ≤.02) and item 7 (P<.001). Furthermore, a significant difference was observed in the IPSS Q3 (quality of life) score (P=.008), whereas no significant difference was found for IPSS item 6. For the IPPI questionnaire, the accept group scored significantly higher across all domains except for domain G (P=.34).

XSL•FO
RenderX

Figure 4. Flowchart of the SDM modeling research process. CG: control group; MEG: model establishment group; ML: machine learning; MLSG: machine learning suggestion group; PSA: prostate-specific antigen; RCT: randomized controlled trial; SDM: shared decision-making.

XSL•FO
RenderX

Table . Baseline characteristics in the MEG[a] group.

Variables	MEG group			*P* value
	All (n=507)	Accept (n=130)	Not now (n=377)	
Age (y), mean (SD)	62.89 (9.70)	63.63 (9.41)	62.63 (9.80)	.31
KnowPSA[b], mean (SD)	2.11 (0.73)	2.23 (0.81)	2.07 (0.69)	.05
Marriage, n (%)				<.001
Married	447 (88.16)	110 (84.62)	337 (89.39)	
Divorce	5 (0.98)	5 (3.85)	0 (0)	
Single	18 (3.55)	9 (6.92)	9 (2.39)	
Widow	37 (7.30)	6 (4.62)	31 (8.22)	
PCaFH[c], n (%)				.05
Yes	404 (79.68)	96 (73.85)	308 (81.70)	
No	103 (20.32)	34 (26.15)	69 (18.30)	
IPSS[d], mean (SD)				
IPSS 1	0.86 (1.21)	1.31 (1.48)	0.70 (1.06)	<.001
IPSS 2	0.89 (1.21)	1.33 (1.42)	0.74 (1.09)	<.001
IPSS 3	0.78 (1.25)	1.20 (1.53)	0.64 (1.10)	.008
IPSS 4	0.78 (1.21)	1.15 (1.46)	0.66 (1.08)	.02
IPSS 5	0.53 (1.02)	0.87 (1.26)	0.42 (0.89)	<.001
IPSS 6	0.64 (1.10)	0.81 (1.28)	0.58 (1.02)	.06
IPSS 7	1.71 (1.16)	2.02 (1.29)	1.60 (1.10)	<.001
IPSS Q3[e]	5.12 (1.33)	4.85 (1.45)	5.21 (1.28)	.008
IPPI[f], mean (SD)				
A	3.39 (1.27)	3.93 (1.07)	3.20 (1.29)	<.001
B	3.35 (1.35)	3.60 (1.26)	3.26 (1.37)	<.001
C	3.67 (1.17)	3.88 (1.15)	3.60 (1.17)	.006
D	3.14 (1.54)	3.42 (1.35)	3.05 (1.59)	.001
E	3.60 (1.26)	4.43 (0.85)	3.31 (1.26	<.001
F	3.27 (1.48)	4.16 (1.00)	2.96 (1.49)	<.001
G	2.26 (1.56)	2.29 (1.37)	2.25 (1.62)	.34
H	2.86 (1.57)	2.94 (1.26)	2.84 (1.67)	.004
I	3.53 (1.40)	3.71 (1.25)	3.47 (1.44)	.005
J	3.12 (1.56)	3.43 (1.29)	3.01 (1.63)	<.001
Top 3 ranked decision-making factors (A-J)[g], n (%)				
First				
A	88 (17.35)	32 (24.62)	56 (14.85)	.01
B	85 (16.76)	10 (7.69)	75 (19.89)	.001
D	42 (8.28)	3 (2.31)	39 (10.34)	.004
E	32 (6.31)	19 (14.62)	13 (3.45)	<.001
Second				
B	82 (16.17)	18 (13.85)	64 (16.98)	.40
E	43 (8.48)	27 (20.77)	16 (4.24)	<.001
F	29 (5.72)	17 (13.08)	12 (3.18)	<.001

XSL•FO
RenderX

Variables	MEG group			*P* value
	All (n=507)	Accept (n=130)	Not now (n=377)	
H	47 (9.27)	4 (3.08)	43 (11.41)	.004
Third				
C	112 (22.09)	28 (21.54)	84 (22.28)	.86
E	28 (5.52)	13 (10.00)	15 (3.98)	.009
F	28 (5.52)	19 (14.62)	9 (2.39)	<.001
J	65 (12.82)	9 (6.92)	56 (14.85)	.01

[a]MEG: model establishment group.

[b]KnowPSA: How would you rate your level of knowledge regarding prostate-specific antigen screening?

[c]PCaFF: prostate cancer among family or close friends.

[d]IPSS: International Prostate Symptom Score.

[e]Q3: quality of life.

[f]IPPI: importance for physiological and psychological impact.

[g]Decision factors derived from the Importance for IPPI questionnaire.

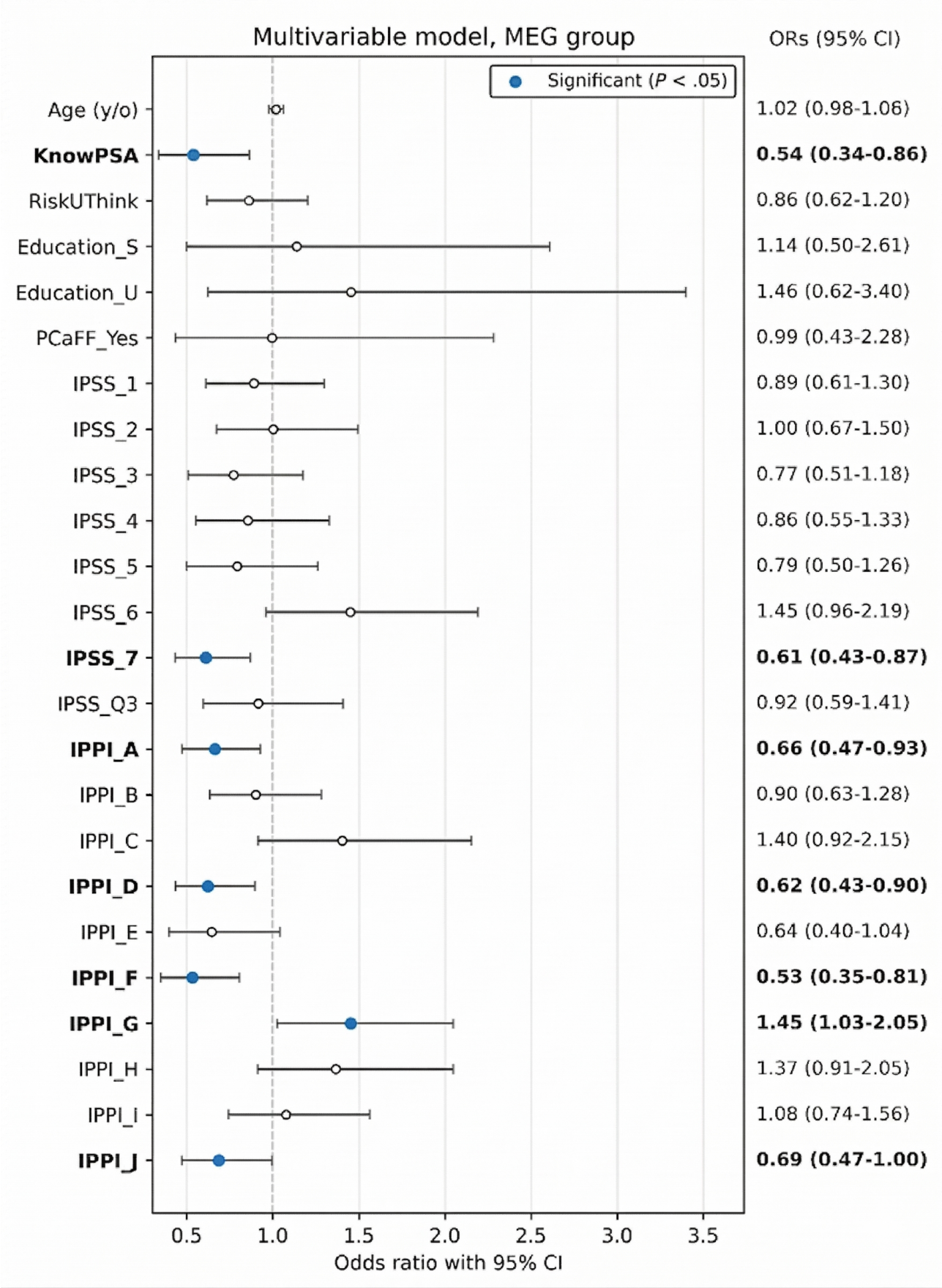
Regarding decision-making priorities, item A was most frequently identified as the primary factor in the accept group (n=32, 24.62%), which was significantly higher than in the not now group (14.85%; *P*=.01). In contrast, item B was more commonly prioritized as the first concern by the not now group (19.89%) compared to the accept group (7.69%; *P*=.001). Significant between-group differences were also noted for items D (*P*=.004) and E (*P*<.001). For the second most important factor, item E was significantly more endorsed by the accept group (20.77%) than the not now group (4.24%; *P*<.001), a trend also observed for item F (13.08% vs 3.18%; *P*<.001). Conversely, item H was less favored in the accept group (3.08%) compared to the not now group (11.41%; *P*=.004).

As for the third most important factor, item C was the leading choice overall (22.09%). However, item F was significantly emphasized in the accept group (14.62%) relative to the not now group (2.39%; *P*<.001). Additionally, significant between-group variations were observed for items E (*P*=.009) and J (*P*=.01). These findings highlight the differential influence of items A, B, C, E, F, and J on decision-making, suggesting nuanced variations in patient priorities as captured by the IPPI.

Figure 5 shows the multivariable LGR results for participants in the MEG group, presenting odds ratios (ORs) and 95% CIs for each covariate, with PSA screening uptake as the outcome. Statistically significant estimates (*P*<.05) are highlighted in blue, whereas nonsignificant estimates are shown as hollow markers. In this model, KnowPSA (OR 0.54, 95% CI 0.34 - 0.86; *P*=.01) and IPSS_7 (OR 0.61, 95% CI 0.43 - 0.87; *P*=.006) were statistically significant. Moreover, several IPPI subscales were also significantly associated with the outcome, including IPPI_A (OR 0.66, 95% CI 0.47 - 0.93; *P*=.02), IPPI_D (OR 0.62, 95% CI 0.43 - 0.90; *P*=.01), IPPI_F (OR 0.53, 95% CI 0.35 - 0.81; *P*=.003), IPPI_G (OR 0.53, 95% CI 0.35 - 0.81; *P*=.003), and IPPI_J (OR 0.69, 95% CI 0.47 - 1.00; *P*=.048). Together, these results suggest that both knowledge-related and psychological factors are associated with PSA screening decisions in this group. Among the significant predictors, KnowPSA, IPPI_A, IPPI_D, IPPI_F, and IPPI_G had the smallest ORs, and their 95% CIs were entirely below 1, supporting their independent associations with the outcome. Overall, these findings underscore the importance of addressing both informational and psychological dimensions when supporting patient decision-making in the context of ML-enhanced guidance.

XSL•FO
RenderX

Figure 5. Forest plot of multivariate analysis among MEG group. IPPI: importance for physiological and psychological impact; IPSS: International Prostate Symptom Score; MEG: model establishment group; OR: odds ratio; PSA: prostate-specific antigen.



XSL•FO
RenderX

Demographic Characteristics of the RCT Group

A total of 367 participants were included in the analysis, consisting of 185 in the MLSG group and 182 in the CG group (Table 2). Detailed results are provided in Multimedia Appendix 12. Regarding lower urinary tract symptoms measured by the IPSS, significant between-group differences were observed in specific domains. Specifically, scores for IPSS item 3 (0.93±1.28 vs 0.63±0.88; *P*=.008) and item 4 (0.93±1.33 vs 0.58±0.85; *P*=.003) were significantly higher in the MLSG group compared with the CG group.

Table . Baseline characteristics in the RCT[a] group.

Variables	RCT group			*P* value
	All (n=367)	MLSG[b] (n=185)	CG[c] (n=182)	
IPSS[d], mean (SD)				
IPSS 1	0.76 (1.16)	0.86 (1.28)	0.65 (1.02)	.08
IPSS 2	0.90 (1.32)	0.95 (1.38)	0.86 (1.26)	.54
IPSS 3	0.78 (1.11)	0.93 (1.28)	0.63 (0.88)	.008
IPSS 4	0.76 (1.13)	0.93 (1.33)	0.58 (0.85)	.003
Top 3 ranked decision-making factors (A-J)[e], n (%)				
First				
A	114 (31.06)	46 (25.27)	68 (36.76)	.01
B	48 (13.08)	30 (16.48)	18 (9.73)	.06
G	11 (3.00)	2 (1.10)	9 (4.86)	.03
Second				
E	73 (19.89)	27 (14.84)	46 (24.86)	.01
Third				
F	84 (22.89)	37 (20.33)	47 (25.41)	.24

[a]RCT: randomized controlled trial.

[b]MLSG: machine learning suggestion group.

[c]CG: control group.

[d]IPSS: International Prostate Symptom Score.

[e]Decision factors derived from the importance for physiological and psychological impact questionnaire.

When ranking primary decision-making concerns, item A was most frequently identified as the top priority overall (31.06%). Notably, a significantly larger proportion of the CG group selected item A as their first concern compared with the MLSG group (36.76% vs 25.27%; *P*=.01). Regarding subsequent priorities, item E was the most common second-ranked choice overall (19.89%), while item F was the most frequently selected third-ranked priority (22.89%).

Table 3 summarizes the impact of ML-based recommendations on participants' final decision-making outcomes. In the CG, 24.18% of participants selected "accept" as their final decision, while 75.82% chose "not now." In contrast, participants in the MLSG who received an "accept" recommendation were significantly more likely to select "accept" compared to those in the CG (50.75% vs 24.18%; *P*<.001).

Table . Effects of machine learning suggestions on final decision.

	Final decision		*P* value
	Accept	Not now	
MLSG[a] (decision aid suggestion)			<.001
Accept	34 (50.75)	33 (49.25)	
Not now	21 (17.80)	97 (82.20)	
CG[b]	44 (24.18)	138 (75.82)	—[c]

[a]MLSG: machine learning suggestion group.

[b]CG: control group.

[c]Not applicable.

XSL•FO
RenderX

Furthermore, when the ML system suggested "not now," only 17.8% of participants chose to "accept," which was lower than the baseline acceptance rate in the CG. These findings indicate that ML recommendations exerted a statistically significant influence on participants' decisions, substantially increasing the likelihood of selecting the recommended option.

Psychological Outcomes, Decision Satisfaction, and Decision Support Between MLSG and CG

For interpretability, lower scores on the DCS and Sa scale reflect better outcomes (ie, reduced conflict and enhanced satisfaction). Regarding the primary outcome, the total DCS score was significantly lower in the MLSG compared with the CG (MD –3.77, 95% CI –5.55 to −1.99; Cohen *d*=–0.44; *P*<.001), indicating superior overall decision quality. Item-level analysis revealed significant improvements in several key dimensions: participants in the MLSG reported greater perceived decision support (DCS7: adjusted *P*=.03), more adequate advice (DCS9: adjusted *P*<.001), higher decision confidence (DCS10: adjusted *P*=.03; DCS11: adjusted *P*<.001), and found the decision-making process significantly easier (DCS12: adjusted *P*<.001).

For decision satisfaction, the MLSG demonstrated a large and statistically significant improvement in the total Sa score (MD –7.38, 95% CI –8.54 to −6.18; *d*=–1.29; *P*<.001). Notably, the MLSG reported significantly higher satisfaction across all 5 items (adjusted *P* values <.001, Cohen $d \leq -1.10$). These gains were most pronounced in information adequacy (Sa1: MD −1.46, 95% CI –1.72 to −1.20), confidence that the decision was the best personal option (Sa2), alignment with personal values (Sa3), and a stronger intention to implement the chosen decision (Sa4).

Regarding anxiety, although the total STAI score did not differ significantly between groups (*P*=.37), item-level analysis identified significant improvements in specific psychological states. Participants in the MLSG reported feeling significantly less worried (STAI 6: MD –0.98, 95% CI –1.20 to −0.76; *d*=–0.89; adjusted *P*<.001) and significantly calmer (STAI 1: MD 0.30, 95% CI 0.06 to 0.54; *d*=0.25; adjusted *P*=.01) compared with the CG. Detailed item-level comparisons are presented in Table 4.

XSL•FO
RenderX

Table . Comparison of the STAI[a], Sa[b], and DSC[c] questionnaires between the MLSG[d] and CG[e].

Variables	RCT[f] group			*P* value	Cohen *d*	Adjusted *P*[g]
	MLSG (n=185), mean (SD)	CG (n=182), mean (SD)	MD[h] (95% CI)			
Anxiety						
STAI 1	2.28 (1.20)	1.98 (1.14)	0.30 (0.06 to 0.54)	.004	0.25	*.01*
STAI 2	2.22 (0.98)	2.08 (0.90)	0.14 (–0.05 to 0.33)	.14	0.15	.17
STAI 3	2.21 (1.09)	2.36 (1.05)	–0.15 (–0.37 to 0.07)	.17	–0.14	.17
STAI 4	2.15 (1.19)	1.97 (1.13)	0.18 (–0.06 to 0.42)	.06	0.15	.10
STAI 5	2.12 (1.21)	1.87 (1.10)	0.25 (0.01 to 0.49)	.03	0.21	.06
STAI 6	2.00 (1.02)	2.98 (1.16)	–0.98 (–1.20 to -0.76)	<.001	–0.89	*<.001*
Total	12.97 (2.39)	13.23 (3.13)	–0.26 (–0.83 to 0.31)	.37	–0.09	—[i]
Decision satisfaction						
Sa 1	1.75 (0.92)	3.21 (1.56)	–1.46 (–1.72 to -1.20)	<.001	–1.14	*<.001*
Sa 2	1.83 (0.88)	3.29 (1.44)	–1.46 (–1.71 to –1.21)	<.001	–1.22	*<.001*
Sa 3	1.77 (0.89)	3.19 (1.60)	–1.42 (–1.69 to –1.15)	<.001	–1.10	*<.001*
Sa 4	1.67 (0.82)	3.13 (1.69)	–1.46 (–1.73 to –1.19)	<.001	–1.10	*<.001*
Sa 5	1.57 (0.81)	3.12 (1.67)	–1.55 (–1.82 to –1.28)	<.001	–1.18	*<.001*
Total	8.58 (3.02)	15.94 (7.52)	–7.38 (–8.54 to –6.18)	<.001	–1.29	—
Decision conflicts						
DCS 1	2.09 (0.96)	2.29 (1.11)	–0.20 (–0.41 to 0.01)	.14	–0.19	.22
DCS 2	2.23 (1.02)	2.18 (1.15)	0.05 (–0.17 to 0.27)	.32	0.05	.40
DCS 3	2.24 (1.09)	2.19 (1.09)	0.05 (–0.17 to 0.27)	.65	0.05	.65
DCS 4	2.23 (1.09)	2.16 (1.18)	0.07 (–0.16 to 0.30)	.37	0.06	.43
DCS 5	2.17 (1.08)	2.33 (1.24)	–0.16 (–0.40 to 0.08)	.32	–0.14	.40
DCS 6	2.29 (1.08)	2.42 (1.25)	–0.13 (–0.37 to 0.11)	.43	–0.11	.46
DCS 7	1.85 (1.05)	2.19 (1.24)	–0.34 (–0.58 to –0.10)	.01	–0.30	*.03*
DCS 8	2.05 (1.17)	1.91 (1.06)	0.14 (–0.09 to 0.37)	.31	0.12	.40
DCS 9	1.73 (0.95)	2.30 (1.17)	–0.57 (–0.79 to –0.35)	<.001	–0.56	*<.001*
DCS 10	1.88 (0.99)	2.26 (1.26)	–0.38 (–0.61 to –0.15)	.008	–0.33	*.03*
DCS 11	1.81 (0.98)	2.38 (1.22)	–0.57 (–0.80 to -0.34)	<.001	–0.51	*<.001*
DCS 12	1.81 (0.93)	2.43 (1.23)	–0.62 (-0.84 to –0.40)	<.001	–0.57	*<.001*

XSL•FO
RenderX

Variables	RCT[f] group			*P* value	Cohen *d*	Adjusted *P*[g]
	MLSG (n=185), mean (SD)	CG (n=182), mean (SD)	MD[h] (95% CI)			
DCS 13	1.89 (1.05)	2.12 (1.17)	–0.23 (–0.46 to –0.00)	.07	–0.21	.15
DCS 14	1.96 (0.95)	2.15 (1.06)	–0.19 (–0.40 to 0.02)	.10	–0.19	.19
DCS 15	1.78 (0.87)	2.08 (1.16)	–0.30 (–0.51 to –0.09)	.04	–0.29	.09
DCS 16	1.60 (0.73)	1.98 (1.09)	–0.38 (–0.57 to –0.19)	.004	–0.41	*.01*
Total	31.60 (9.29)	35.37 (7.98)	–3.77 (–5.55 to –1.99)	<.001	–044	—

[a]STAI: Spielberger State-Trait Anxiety Inventory.

[b]Sa: Satisfaction with Decision.

[c]DCS: Decisional Conflict Scale.

[d]MLSG: machine learning suggestion group.

[e]CG: control group.

[f]RCT: randomized controlled trial.

[g]*P* values adjusted for multiplicity using the Benjamini-Hochberg false discovery rate method within each outcome (anxiety, decision satisfaction, decisional conflict). Italics values indicate statistical significance after adjustment (adjusted *P*<.05).

[h]MD: mean difference.

[i]Not applicable.

[j]Scale direction: For DCS and Sa, lower scores indicate better outcomes (ie, lower decisional conflict and higher decision satisfaction). For STAI, item direction varies: for STAI-6 (worry), lower scores indicate lower anxiety; whereas for STAI-1 (calm) and STAI-5 (reassured), higher scores indicate greater calmness and reassurance.

Discussion

Principal Findings

In this randomized controlled trial of 367 men aged 50 - 75 years, adding personalized ML recommendations to a structured SDM workflow was associated with significantly improved decision quality for PSA screening. Supported by a high-performing prediction model (AUC 0.93), the intervention group (MLSG) reported lower decisional conflict and reduced anxiety compared with standard care. Specifically, participants receiving ML-supported recommendations experienced greater calmness (STAI 1: adjusted *P*=.01) and less worry (STAI 6: adjusted *P*<.001). Beyond anxiety reduction, the intervention produced large, consistent improvements in decision satisfaction across all domains (Cohen *d* >1.10; adjusted *P* values <.001), including perceived adequacy of information and alignment with personal values.

Importantly, the intervention was designed to complement—rather than replace—clinician counseling. The ML output was presented as an advisory “second opinion” and reinforced by standardized clinician communication, ensuring that the final screening decision remained patient-directed. This approach is particularly relevant for older adults, who often face cognitive load and uncertainty when weighing the complex trade-offs of PSA screening [32].

Comparison With Prior Work

Although conventional prediction approaches (eg, LGR) can identify determinants associated with PSA screening intentions and uptake, they are generally not designed to deliver individualized, real-time recommendations at the point of decision-making [33]. Unlike prior decision-aid evaluations that focused primarily on presentation format or generic personalization, our intervention explicitly targeted decisional conflict and anxiety [34,35]—key barriers to informed choice in adults and aging populations [3,36]. Prior studies also suggest that education, age, symptom-related, and psychosocial factors influence screening behavior, and that incorporating constructs such as self-efficacy and emotional well-being may improve the use of decision support tools [37-39]. Watson et al [33] reported that men’s intention to undergo PSA testing was closely related to baseline knowledge, attitudes, and perceived risk, underscoring the central role of informational and psychological determinants.

Consistent with this literature, our multivariable LGR analysis in the MEG cohort confirmed that screening decisions were driven by a complex interplay of informational and preference-related factors. Specifically, higher PSA-related knowledge (KnowPSA) was independently associated with lower screening uptake (OR 0.54; *P*=.01). This inverse association suggests that participants with limited understanding may be more inclined to accept screening without fully weighing potential harms (eg, false positives, overdiagnosis), whereas greater awareness appears to foster a more cautious approach, reducing uninformed compliance.

In addition to knowledge, clinical symptoms and personal values were robust predictors of decision-making. Greater urinary symptom severity (IPSS item 7; OR 0.61; *P*=.006) was

XSL•FO
RenderX

associated with lower screening uptake, possibly reflecting a prioritization of symptom management over preventive screening in symptomatic individuals. Furthermore, several values-clarification subscales (IPPI items A, D, F, G, and J) significantly predicted screening behavior, indicating that participants who placed higher importance on specific physiological, psychological, or social impacts were more likely to decline or defer screening.

These results suggest that when patients are equipped with sufficient knowledge and an opportunity to clarify their values, they may become more discerning, potentially leading to informed decisions to decline screening based on their personal context. This aligns with the core tenet of SDM, which prioritizes value-concordant choices over passive adherence. The robust predictive power of these factors underscores the use of our ML model in capturing nuanced, individualized determinants, thereby moving beyond generic guidelines to support truly patient-centered care.

Clinical Implications

Recent evidence suggests that fewer than 25% of men aged 40 and older report discussing the benefits and harms of PSA testing with their providers, a gap even more pronounced among those with lower educational attainment [40]. The selection of the delivery mode for DAs is critical to addressing this disparity. Given that the target population for prostate cancer screening primarily consists of older adults, web-based DAs may present usability challenges due to age-related declines in fine motor control or limited digital literacy [41]. To mitigate these barriers, our intervention used a facilitated delivery approach. By providing personalized, accessible recommendations through a supported interface, our ML-driven decision aid ensures that screening decisions align with individual values rather than being driven by incomplete information or physician preference alone. Compared with conventional decision aids, ML recommendations significantly enhanced decisional confidence and yielded modest yet consistent increases in screening acceptance. This aligns with prior evidence that technology-supported tools can shape patient preferences and improve decision quality [24,42].

However, we emphasize that ML should be viewed as a decision-support tool, not a decision-maker. The modest increase in screening acceptance observed in the MLSG suggests that the tool facilitated decisive action, but ethical implementation requires ensuring that this influence does not coerce patients. Future integration into routine practice must prioritize transparency and physician training to maintain the delicate balance between algorithmic guidance and patient autonomy.

Taken together, these findings indicate that embedding personalized, ML-driven recommendations within SDM frameworks may not only improve decision quality but also enhance psychological outcomes. Before integrating ML-assisted tools into routine clinical practice, further research is warranted to evaluate their psychological and physiological safety. Such investigations are essential to determine not only the efficacy of ML recommendations in enhancing decision quality but also their broader implications for patient autonomy and well-being [43].

Limitations

Several limitations should be noted. First, while we provided uniform technical assistance, we did not formally assess baseline eHealth literacy or track the frequency of support required, limiting our ability to analyze effects based on digital proficiency. Future studies should incorporate validated measures like the eHealth Literacy Scale. Second, although our model demonstrated high discrimination (AUC >0.90), calibration metrics (eg, Brier score) were not assessed during the development phase; future work should include calibration plots to ensure risk estimates are accurate across subgroups. Third, unmeasured confounding remains a possibility, as factors such as physician influence or access barriers were not fully captured.

Finally, the current user interface provides binary recommendations (agree vs not now) without visualizing feature contributions. Integrating Explainable Artificial Intelligence techniques, such as Shapley Additive Explanations values, will be critical for future iterations to enhance transparency and trust.

Conclusion

This study demonstrates that integrating personalized ML recommendations into SDM frameworks can enhance decision quality while simultaneously reducing psychological burdens, such as anxiety and decisional conflict. Participants exposed to ML-supported recommendations reported greater confidence, reassurance, and satisfaction with their choices, underscoring the unique psychological value of individualized algorithmic guidance in prostate cancer screening.

Importantly, these findings suggest that ML-based decision support should be regarded as a complement to, rather than a replacement for, traditional decision aids and physician-patient discussions. By enriching the decision-making process with nuanced, personalized insights, ML has the potential to foster more supportive and less distressing decision environments.

Funding

The research is supported by the Fund by St. Joseph Hospital (project 10503) and The National Science and Technology Council, Taiwan (NSTC 112-2221-E-030 -009 -MY3).

Authors' Contributions

Y-TL and Y-CH contributed equally to this work and share first authorship.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Importance for physiological and psychological impact questionnaire.
[DOCX File, 26 KB - aging_v9i1e83238_app1.docx]

Multimedia Appendix 2
International Prostate Symptom Score questionnaire.
[DOCX File, 23 KB - aging_v9i1e83238_app2.docx]

Multimedia Appendix 3
Decisional Conflict Scale questionnaire.
[DOCX File, 22 KB - aging_v9i1e83238_app3.docx]

Multimedia Appendix 4
Expert validity review questionnaire.
[DOCX File, 34 KB - aging_v9i1e83238_app4.docx]

Multimedia Appendix 5
Knowledge test items.
[DOCX File, 16 KB - aging_v9i1e83238_app5.docx]

Multimedia Appendix 6
Website link and interface.
[DOCX File, 345 KB - aging_v9i1e83238_app6.docx]

Multimedia Appendix 7
Technical details of model development.
[DOCX File, 30 KB - aging_v9i1e83238_app7.docx]

Multimedia Appendix 8
Performance evaluation of machine learning models.
[DOCX File, 22 KB - aging_v9i1e83238_app8.docx]

Multimedia Appendix 9
Variables description.
[DOCX File, 22 KB - aging_v9i1e83238_app9.docx]

Multimedia Appendix 10
Detail baseline characteristics in the model establishment group.
[DOCX File, 31 KB - aging_v9i1e83238_app10.docx]

Multimedia Appendix 11
R packages.
[DOCX File, 18 KB - aging_v9i1e83238_app11.docx]

Multimedia Appendix 12
Detail baseline characteristics in the randomized controlled trial group.
[DOCX File, 30 KB - aging_v9i1e83238_app12.docx]

Checklist 1
CONSORT-EHEALTH checklist.
[PDF File, 1201 KB - aging_v9i1e83238_app13.pdf]

XSL•FO
RenderX

References

1. Makoul G, Clayman ML. An integrative model of shared decision making in medical encounters. Patient Educ Couns 2006 Mar;60(3):301-312. [doi: 10.1016/j.pec.2005.06.010] [Medline: 16051459]
2. Barry MJ, Edgman-Levitan S. Shared decision making: pinnacle of patient-centered care. N Engl J Med 2012 Mar 1;366(9):780-781. [doi: 10.1056/NEJMp1109283] [Medline: 22375967]
3. Pekala KR, Shill DK, Austria M, Langford AT, Loeb S, Carlsson SV. Shared decision-making before prostate cancer screening decisions. Nat Rev Urol 2024 Jun;21(6):329-338. [doi: 10.1038/s41585-023-00840-0] [Medline: 38168921]
4. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997 Mar;44(5):681-692. [doi: 10.1016/s0277-9536(96)00221-3] [Medline: 9032835]
5. Makarov DV, Chrouser K, Gore JL, et al. AUA white paper on implementation of shared decision making into urological practice. Urol Pract 2016 Sep;3(5):355-363. [doi: 10.1016/j.urpr.2015.10.006] [Medline: 37592546]
6. Makarov D, Fagerlin A, Finkelstein J, Forcino R, Gore J, Henning G, et al. Implementation of shared decision making into urological practice. : American Urological Association (AUA); 2022 URL: https://www.auanet.org/documents/Guidelines/Implementation%20of%20Shared%20Decision%20Making%20into%20Urological%20Practice%20unabridged%20version.pdf [accessed 2026-03-30]
7. Bagshaw HP, Martinez A, Heidari N, et al. A personalized decision aid for prostate cancer shared decision making. BMC Med Inform Decis Mak 2021 Dec 31;21(1):374. [doi: 10.1186/s12911-021-01732-2] [Medline: 34972513]
8. Chan EC, Sulmasy DP. What should men know about prostate-specific antigen screening before giving informed consent? Am J Med 1998 Oct;105(4):266-274. [doi: 10.1016/s0002-9343(98)00257-5] [Medline: 9809686]
9. Stiggelbout AM, Pieterse AH, De Haes JCJM. Shared decision making: concepts, evidence, and practice. Patient Educ Couns 2015 Oct;98(10):1172-1179. [doi: 10.1016/j.pec.2015.06.022] [Medline: 26215573]
10. Jayadevappa R, Chhatre S, Wong YN, et al. Comparative effectiveness of prostate cancer treatments for patient-centered outcomes: a systematic review and meta-analysis (PRISMA compliant). Medicine (Baltimore) 2017 May;96(18):e6790. [doi: 10.1097/MD.0000000000006790] [Medline: 28471976]
11. Cancer stat facts: prostate cancer. National Cancer Institute - SEER. 2020. URL: https://seer.cancer.gov/statfacts/html/prost.html [accessed 2026-03-30]
12. Chen L, Li Q, Wang Y, Zhang Y, Ma X. Comparison on efficacy of radical prostatectomy versus external beam radiotherapy for the treatment of localized prostate cancer. Oncotarget 2017;8(45):79854-79863. [doi: 10.18632/oncotarget.20078] [Medline: 29108367]
13. Davison BJ, Kirk P, Degner LF, Hassard TH. Information and patient participation in screening for prostate cancer. Patient Educ Couns 1999 Jul;37(3):255-263. [doi: 10.1016/s0738-3991(98)00123-2] [Medline: 14528551]
14. Riikonen JM, Guyatt GH, Kilpeläinen TP, et al. Decision aids for prostate cancer screening choice: a systematic review and meta-analysis. JAMA Intern Med 2019 Aug 1;179(8):1072-1082. [doi: 10.1001/jamainternmed.2019.0763] [Medline: 31233091]
15. Estevan-Vilar M, Parker LA, Caballero-Romeu JP, Ronda E, Hernández-Aguado I, Lumbreras B. Barriers and facilitators of shared decision-making in prostate cancer screening in primary care: a systematic review. Prev Med Rep 2024 Jan;37:102539. [doi: 10.1016/j.pmedr.2023.102539] [Medline: 38179441]
16. Gans EA, Pieterse AH, Klapwijk MS, et al. Shared decision-making with older adults with cancer: adaptation of a model through literature review and expert opinion. Psychooncology 2024 Jan;33(1):e6291. [doi: 10.1002/pon.6291] [Medline: 38282224]
17. Sheridan SL, Golin C, Bunton A, et al. Shared decision making for prostate cancer screening: the results of a combined analysis of two practice-based randomized controlled trials. BMC Med Inform Decis Mak 2012 Nov 13;12:130. [doi: 10.1186/1472-6947-12-130] [Medline: 23148458]
18. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367. [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
19. Wei JT, Barocas D, Carlsson S, et al. Early detection of prostate cancer: AUA/SUO guideline part I: prostate cancer screening. J Urol 2023 Jul;210(1):46-53. [doi: 10.1097/JU.0000000000003491] [Medline: 37096582]
20. Baggett ND, Schulz K, Buffington A, et al. Surgeon use of shared decision-making for older adults considering major surgery: a secondary analysis of a randomized clinical trial. JAMA Surg 2022 May 1;157(5):406-413. [doi: 10.1001/jamasurg.2022.0290] [Medline: 35319737]
21. Adeli M, Zayeri F, Rasekhi AA, Kiani Z, Amini P. Impact of decision aids on screening prostate cancer: a systematic review with a meta-analysis. Iran Red Crescent Med J 2024;26(1):1-9. [doi: 10.22034/ircmj.2024.203979]
22. Barocas DA, Alvarez J, Resnick MJ, et al. Association between radiation therapy, surgery, or observation for localized prostate cancer and patient-reported outcomes after 3 years. JAMA 2017 Mar 21;317(11):1126-1140. [doi: 10.1001/jama.2017.1704] [Medline: 28324093]
23. Tătaru OS, Vartolomei MD, Rassweiler JJ, et al. Artificial intelligence and machine learning in prostate cancer patient management-current trends and future perspectives. Diagnostics (Basel) 2021 Feb 20;11(2):354. [doi: 10.3390/diagnostics11020354] [Medline: 33672608]

XSL•FO
RenderX

24. Hassan N, Slight R, Bimpong K, et al. Systematic review to understand users perspectives on AI-enabled decision aids to inform shared decision making. NPJ Digit Med 2024 Nov 21;7(1):332. [doi: 10.1038/s41746-024-01326-y] [Medline: 39572838]
25. Rychetnik L, Doust J, Thomas R, Gardiner R, Mackenzie G, Glasziou P. A community jury on PSA screening: what do well-informed men want the government to do about prostate cancer screening: a qualitative analysis. BMJ Open 2014 Apr 30;4(4):e004682. [doi: 10.1136/bmjopen-2013-004682] [Medline: 24785399]
26. Ferrante JM, Shaw EK, Scott JG. Factors influencing men's decisions regarding prostate cancer screening: a qualitative study. J Community Health 2011 Oct;36(5):839-844. [doi: 10.1007/s10900-011-9383-5] [Medline: 21336996]
27. Fukasaku R, Hirose K, Kabata Y, Teramoto K. Algebraic approach to maximum likelihood factor analysis. Psychometrika 2025 Sep 15;90(5):1-33. [doi: 10.1017/psy.2025.10047] [Medline: 40947992]
28. Kaiser HF. The Varimax criterion for analytic rotation in factor analysis. Psychometrika 1958 Sep;23(3):187-200. [doi: 10.1007/BF02289233]
29. Jiang C, Fedewa SA, Wen Y, Jemal A, Han X. Shared decision making and prostate-specific antigen based prostate cancer screening following the 2018 update of USPSTF screening guideline. Prostate Cancer Prostatic Dis 2021 Mar;24(1):77-80. [doi: 10.1038/s41391-020-0227-1] [Medline: 32296126]
30. Final recommendation statement: prostate cancer: screening. United States Preventive Services Task Force. 2018 May 8. URL: https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/prostate-cancer-screening [accessed 2018-05-08]
31. American Cancer Society recommendations for prostate cancer early detection. American Cancer Society. URL: https://www.cancer.org/cancer/types/prostate-cancer/detection-diagnosis-staging/acs-recommendations.html [accessed 2023-11-22]
32. Kotwal AA, Walter LC. Cancer screening in older adults: individualized decision-making and communication strategies. Med Clin North Am 2020 Nov;104(6):989-1006. [doi: 10.1016/j.mcna.2020.08.002] [Medline: 33099456]
33. Watson E, Hewitson P, Brett J, et al. Informed decision making and prostate specific antigen (PSA) testing for prostate cancer: a randomised controlled trial exploring the impact of a brief patient decision aid on men's knowledge, attitudes and intention to be tested. Patient Educ Couns 2006 Nov;63(3):367-379. [doi: 10.1016/j.pec.2006.05.005] [Medline: 16875796]
34. Frosch DL, Kaplan RM, Felitti V. The evaluation of two methods to facilitate shared decision making for men considering the prostate-specific antigen test. J Gen Intern Med 2001 Jun;16(6):391-398. [doi: 10.1046/j.1525-1497.2001.016006391.x] [Medline: 11422636]
35. Salkeld G, Cunich M, Dowie J, et al. The role of personalised choice in decision support: a randomized controlled trial of an online decision aid for prostate cancer screening. PLoS ONE 2016;11(4):e0152999. [doi: 10.1371/journal.pone.0152999] [Medline: 27050101]
36. Taylor KL, Williams RM, Davis K, et al. Decision making in prostate cancer screening using decision aids vs usual care: a randomized clinical trial. JAMA Intern Med 2013 Oct 14;173(18):1704-1712. [doi: 10.1001/jamainternmed.2013.9253] [Medline: 23896732]
37. Robinson E, DiIorio C, DePadilla L, et al. Psychosocial predictors of lifestyle management in adults with epilepsy. Epilepsy Behav 2008 Oct;13(3):523-528. [doi: 10.1016/j.yebeh.2008.05.015] [Medline: 18595777]
38. Choi JHS, Jung DH. The role of psychological factors in predicting self-rated health: implications from machine learning models. Psychol Health Med 2025 Jul;30(6):1158-1170. [doi: 10.1080/13548506.2025.2450546] [Medline: 39778189]
39. Jarrar M, Al-Bsheish M, Alshahri BK, et al. Associations of self-management care and shared decision-making with glycemic control and psychosocial outcomes in type 2 diabetes mellitus. Patient Prefer Adherence 2025;19:2295-2307. [doi: 10.2147/PPA.S534066] [Medline: 40771654]
40. Bhojani N, Miller LE, Zorn KC, et al. Prevalence and determinants of shared decision-making for PSA testing in the United States. Prostate Cancer Prostatic Dis 2025 Jun;28(2):400-404. [doi: 10.1038/s41391-024-00843-x] [Medline: 38671083]
41. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
42. Brown SA, Chung BY, Doshi K, et al. Patient similarity and other artificial intelligence machine learning algorithms in clinical decision aid for shared decision-making in the prevention of cardiovascular toxicity (PACT): a feasibility trial design. Cardiooncology 2023 Jan 23;9(1):7. [doi: 10.1186/s40959-022-00151-0] [Medline: 36691060]
43. Cialdini RB. Influence (Rev) The Psychology of Persuasion: HarperCollins; 1993. URL: https://books.google.co.in/books?id=mTYj9XUlYvMC [accessed 2026-03-18]

Abbreviations

AUC: area under the curve
CG: control group
CONSORT: Consolidated Standards of Reporting Trials
DA: decision aid
DCS: Decisional Conflict Scale
IPPI: importance for physiological and psychological impact
IPSS: International Prostate Symptom Score

XSL•FO
RenderX

LGR: logistic regression
MD: mean difference
MLSG: machine learning suggestion group
OR: odds ratio
PSA: prostate-specific antigen
RCT: randomized controlled trial
RF: random forest
Sa: Satisfaction with Decision
SDM: shared decision-making
STAI: Spielberger State-Trait Anxiety Inventory

Edited by H Köttl; submitted 29.Aug.2025; peer-reviewed by AT Langford, M Pritchard; revised version received 20.Feb.2026; accepted 04.Mar.2026; published 13.Apr.2026.

Please cite as:
Lin YT, Huang YC, Liu CK, Cho HY, Chen M
Web-Based Personalized Machine Learning Recommendations to Enhance Shared Decision-Making in Prostate-Specific Antigen Screening: Randomized Controlled Trial
JMIR Aging 2026;9:e83238
URL: https://aging.jmir.org/2026/1/e83238
doi:10.2196/83238

© Yi-Ting Lin, Yen Chun Huang, Chih Kuang Liu, Hsiao-Yun Cho, Mingchih Chen. Originally published in JMIR Aging (https://aging.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Personalized Digital Care Program Allocation for Older Adults: Reinforcement Learning-Based Simulation Study

Hongsoo Kim[1,2,3], MPH, PhD; Kyungbok Lee[4], PhD; Jae Yoon Yi[1], MPH; Myunghee Cho Paik[5], PhD

[1]Department of Public Health Sciences, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea

[2]Artificial Intelligence Institute, Seoul National University, Seoul, Republic of Korea

[3]Institute of Health and Environment, Seoul National University, Seoul, Republic of Korea

[4]Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[5]Shepherd 23 Inc., Seoul, Republic of Korea

Corresponding Author:
Myunghee Cho Paik, PhD
Shepherd 23 Inc.
1 Gwanak-ro, Gwanak-gu
Seoul, 08826
Republic of Korea
Phone: 82 8802750
Email: myungheechopaik@gmail.com

Abstract

Background: As the demand for innovative older adult care grows alongside a shortage of care workers, personalization is key to optimizing services and enhancing long-term sustainability. This study proposes an adaptive reinforcement learning (RL)-based framework to promote precision digital care by dynamically assigning care programs based on individuals' unique characteristics and evolving needs. Its effectiveness was evaluated through simulation-based experiments comparing multiple allocation methods within an artificial intelligence (AI)-powered care call service for older adults.

Objective: This study aimed to develop and evaluate an RL-based model for personalizing digital care program allocation to optimize care engagement and health outcomes among low-income older adults living alone.

Methods: We developed the framework by using contextual bandits, specifically Thompson Sampling, to maximize user outcomes. Four program allocation strategies were tested using a synthetic dataset of user features and program attributes. Simulations were conducted over multiple iterations to evaluate how the model adapts over time and optimizes program assignments compared with static methods.

Results: Four program allocation methods were compared across 100 simulation runs (n=3000 assignments per run) using 2 datasets (AI Call: n=1196; Community Health Survey [CHS]: n=72,812): (1) systematic allocation (baseline), (2) single best program based on population average, (3) idealized personalized delivery (theoretical upper bound), and (4) precision digital care using Thompson Sampling. Precision digital care outperformed baseline and population-average approaches, achieving outcomes comparable to the theoretical upper bound. Compared to systematic allocation, call success rates increased by 84.2% (AI Call) and 54.4% (CHS), Patient Health Questionnaire-2 depression scores decreased by 32.1% (AI Call) and 41.4% (CHS), and self-reported health scores improved by 19% (AI Call) and 22% (CHS). It also showed improved learning efficiency, refining assignments dynamically as it learned from user responses.

Conclusions: Our findings emphasize the importance of personalization in digital care. We plan to refine and validate the model through a publicly funded AI care program for community-dwelling, low-income older adults living alone in the Republic of Korea. RL offers a scalable and effective approach to advance precision digital care delivery and support future innovations in aging services.

(JMIR Aging 2026;9:e86710) doi:10.2196/86710

KEYWORDS

precision digital care; adaptive program assignment; contextual bandits, AI-enabled care; artificial intelligence; aging population; simulation study

XSL•FO
RenderX

Introduction

The rapid aging of populations and increasing life expectancy demand integrated care and health services to support individuals throughout extended later life. Preventive care, aimed at promoting independent living and delaying institutionalization, has emerged as a key strategy in aging policies across developed countries [1]. Citizens increasingly expect robust social protection from their governments. Meanwhile, health and care workers are grappling with severe burnout and workforce shortages [2,3]. Despite numerous social policy reforms, traditional approaches to addressing the growing "care crisis" are proving inadequate. The failures of both policy and market-based strategies to protect vulnerable older adults during the COVID-19 pandemic underscore the urgent need for transformative changes to achieve global healthy aging [4].

Technological innovation is widely recognized as a powerful solution to the social risks posed by demographic shifts [5,6]. The COVID-19 pandemic catalyzed expansions in R&D and market investments, accelerating the transition from lab-based innovation to real-world applications, such as the widespread adoption of telemedicine [7]. In the postpandemic era, breakthroughs like GPT and large language models have ushered in a new age of digital health and care, opening unprecedented opportunities.

Within this digital health landscape, precision medicine has garnered significant attention [8]. The integration of big data from electronic medical records has propelled advancements in precision medicine, driving the development of personalized medical diagnostics and treatments [9,10]. These innovations hold immense promise, from curing previously untreatable diseases to rapidly discovering new drugs and materials. Consequently, many countries are actively investing in R&D and fostering global cooperation through public-private partnerships.

In contrast, the personalization of care services for older adults has received comparatively limited attention, despite significant progress in digital care for older adults. While precision medicine has primarily focused on genomics, biomarkers, and individualized treatment plans in clinical settings, precision care addresses broader contextual factors, including behavioral data, social determinants, and real-world service use. Precision medicine leverages structured biological data to optimize therapeutic decisions, often within highly controlled environments such as hospitals or labs. In contrast, precision care operates in complex, community-based settings and must contend with heterogeneous data types, evolving patient preferences, and ethical imperatives related to autonomy and accessibility.

We think precision care represents a paradigm shift that expands the individualization principle beyond disease treatment to service delivery and engagement support, especially in aging and chronic care domains. Therefore, the transition from precision medicine to precision care in this study reflects a shift in context, data sources, ethical priorities, and delivery models, rather than a simple semantic change. Several challenges impede this shift, including insufficient large-scale, high-quality care data, lack of data standardization, decentralized data governance, and fragmented health and care systems [11,12]. However, promising initiatives, such as the EU-funded iCare4Old project, demonstrate how these barriers can be overcome [13]. Such projects exemplify the potential of cross-national, standardized, comprehensive geriatric assessment databases linked to mortality and health care use data to support decision-making for older adults with complex chronic conditions.

Digital care is also a high policy priority in Korea, an East Asian country that is expected to face the most severe shortage of care workers among Organization for Economic Co-operation and Development countries by 2040 [2]. Various government-funded digital health and care programs have been developed and implemented to help delay cognitive decline and alleviate loneliness, with their delivery managed through diverse local government units. For example, the central agency for social services in Korea provides artificial intelligence (AI)-based care calls for older adults living alone for checking in and providing information [14]. Although still in the early stages of implementation both technologically and in terms of service design, many social welfare institutions have adopted the AI care call service to alleviate the overtime workload of direct care workers and to prevent unattended deaths among the older adults, an emerging social problem in Korea. The local government, in collaboration with interest groups from universities or industries, implements the Care Call program developed by these groups. While many programs are created through partnerships with local governments, users do not have a choice of programs; instead, the program they receive is determined by their location. As these programs expand to larger populations, they must be evaluated and delivered efficiently.

While AI-driven digital care is expanding, current approaches rely on rule-based or non-adaptive personalization, which may fail to optimize interventions dynamically at the individual level. In health care and digital interventions, the term personalization often refers to the adaptation of care strategies based on characteristics shared within predefined subgroups, such as age, gender, or geographic location. Individuals within these homogeneous groups receive the same intervention, under the assumption that group-level traits sufficiently capture variation in needs or responses. However, this approach can overlook meaningful individual differences. In contrast, hyper-personalization refers to a finer-grained strategy in which care is tailored at the individual level based on contextual signals such as prior responses, behavioral patterns, or dynamically changing health indicators. Rather than assigning treatments to groups, hyper-personalized care selects the best intervention for each instance, enabling adaptive learning and improved outcomes. Our work embraces this hyper-personalization framework by using contextual bandits to deliver care content that evolves with each person's unique context and feedback history.

Beyond technical optimization, hyper-personalization addresses a fundamental challenge in digital care for older persons: one-size-fits-all approaches often lead to disengagement, frustration, and abandonment among older adults [15]. Older people vary widely in cognitive capacity, health literacy, communication preferences, and daily routines [16]. Receiving

XSL•FO
RenderX

care content misaligned with these individual contexts can feel impersonal, patronizing, or irrelevant, undermining both engagement and dignity. For example, an older adult with mild cognitive impairment may benefit most from brief, frequent check-in calls, while someone experiencing loneliness may prefer longer, conversational interactions. By tailoring content to individual context and response patterns, precision digital care respects individual autonomy and maintains dignity through person-centered service delivery.

We model the care call assignment as a repeated decision-making problem, where the system observes individual-level context, selects a care message (action), receives feedback, and uses it to improve future decisions. This setting goes beyond supervised learning, which does not handle the need for sequential adaptation or balance between exploration and exploitation. While Q-learning and other reinforcement learning (RL) methods can technically be applied, they are designed for environments where actions influence future states—an unnecessary assumption in our context, where each care interaction is a standalone event with immediate feedback and does not affect the next person's reward. Contextual bandits with Thompson Sampling offer a natural, data-efficient, and interpretable approach for personalizing care at scale, without introducing unnecessary complexity.

This study proposes precision digital care as an RL-based framework using contextual bandits, specifically designed to personalize care services dynamically. This approach assigns care programs to individuals based on their unique characteristics and evolving care outcomes, continuously optimizing decisions over time. At its core, the framework leverages contextual bandits, a type of RL algorithm, specifically Thompson Sampling [17], to balance exploration and exploitation in digital health interventions. This computational approach enables the model to learn from user interactions, adapt in real time, and continuously refine program assignment decisions to maximize individual outcomes. Unlike traditional rule-based or static personalization methods, the proposed precision digital care approach dynamically optimizes care allocation, ensuring interventions remain responsive and effective.

By integrating Thompson Sampling, the proposed approach provides a scalable, data-driven solution for enhancing personalization, efficiency, and equity in digital health services, particularly in aging care. To evaluate its effectiveness, this study uses simulation-based experiments comparing multiple program allocation methods and analyzing their impact on key health and service-related outcomes. This study proposes precision digital care as an RL-based model designed to enhance quality of life in later years by delivering personalized, humane, and adaptive care. Rather than simply maximizing technical metrics, our approach uses RL as a tool to improve care engagement and health outcomes, supporting autonomy, reducing emotional distress, and ultimately enabling older adults to age in place with dignity. Importantly, by minimizing cumulative regret during the learning process, the approach reduces exposure to suboptimal care while enabling adaptive, sequential learning even in data-sparse or dynamically changing environments**.** By evaluating this approach through simulation, we aim to provide evidence for improving both efficiency and meaningful outcomes for vulnerable older adults before real-world implementation.

Methods

Aim of Simulations

In Korea, researchers are actively developing a range of government-funded digital care programs aimed at improving the well-being of older adults. These programs are administered by local authorities and are currently tailored to specific groups of older adults in distinct regions. As these initiatives progress, the accumulation of knowledge and data is expected to pave the way for large-scale implementation in the near future. This study focused on the critical role of personalization in enhancing the effectiveness of digital care. While personalization has not yet been fully integrated, demographic data on recipients of digital care programs is available. Leveraging this information, we conducted simulation studies that modeled the baseline characteristics of user populations, simulating the potential impact of personalized interventions.

Simulation Methods

To evaluate the potential of personalized digital care, we conducted a series of simulations mimicking the delivery of care programs to older adult residents in a community setting.

Four Allocation Methods

Our simulations were designed to evaluate the method that optimizes the outcome of interest. These 4 allocation methods were selected to represent commonly used approaches in digital health intervention deployment, ranging from traditional rule-based methods (systematic) to theoretically optimal but not feasible in practice (idealized personalized). By comparing these methods, we assessed the added value of the RL-based precision digital care approach in real-world care allocation settings.

Method 1: Systematic Delivery (Baseline)

This method reflects a common real-world practice, including the Korean local government model, where programs are assigned to users in a fixed order, without considering individual needs or program characteristics. This approach offers the simplest implementation, requiring no information about individual users or program effectiveness. However, it represents a naive baseline that can lead to suboptimal outcomes by ignoring potential benefits from personalized program allocation.

Method 2: Single Best Program (Theoretical Benchmark)

This method simulates a hypothetical situation where we know the single best program for the entire population based on average outcomes. This method mimics a scenario where the result of a large-scale randomized controlled trial, where individual characteristics are not used in program assignment, is known. This method provides a useful benchmark for comparison. It demonstrates the upper limit of effectiveness achievable when using a single program for everyone and highlights potential limitations of a one-size-fits-all approach.

XSL•FO
RenderX

If there were no individual variability, the one-size-fits-all approach would be optimal.

Method 3: Idealized Personalized Delivery (Theoretical Upper Bound)

This method assumes perfect knowledge of each individual's optimal program, representing an ideal but unattainable scenario in real-world settings. This method provides a theoretical upper bound on personalization benefits, demonstrating the best possible outcome achievable if we could perfectly predict the most effective program for each individual. Comparing this method to others highlights the potential gains from personalization.

Method 4: Precision Digital Care with Thompson Sampling (Personalized Approach)

This method addresses the challenge of sequentially assigning programs to individuals without prior knowledge of which program will be most effective for each person. This problem can be framed as a "sequential decision making under uncertainty" problem, where the goal is to maximize cumulative rewards over time. To tackle this, we used a contextual bandit framework, specifically using the linear Thompson Sampling algorithm [17].

In this setup, each care program is treated as an "arm" in the contextual bandit framework, where the reward function represents the expected benefit for a given user. The model dynamically updates its understanding of which arm (care program) is most effective based on observed user responses, allowing for personalized interventions over time. The relationship between the outcome (reward) and the context (individual characteristics and program features) is unknown and must be estimated as users experience the program. The challenge lies in balancing exploration (trying different programs to estimate or learn their effectiveness) and exploitation (assigning the program with the highest estimated reward based on current knowledge).

Among various contextual bandit algorithms, Thompson Sampling offers an effective approach for this setting. While upper confidence bound algorithms select arms by calculating an upper confidence bound on their estimated rewards, Thompson Sampling is motivated by a Bayesian approach. Specifically, it draws from the posterior distribution of each arm's reward, and the arm with the highest sampled reward is then selected.

This randomized approach effectively balances exploration and exploitation. In early stages, when uncertainty is high, the algorithm explores a wider range of programs due to the variability in sampled rewards. As more data is collected, the posterior distributions become more concentrated around the true reward values, and the algorithm increasingly exploits the program with the highest estimated reward for each individual. This approach enables precision digital care to adapt to individual differences and learn the optimal program assignment over time.

Data-Generating Mechanisms

To conduct simulation studies, we specified the distribution of the baseline user characteristics and the relationship between these characteristics and the outcomes. Our study involved a total of 6 distinct simulation scenarios, evaluating 3 different outcome measures (detailed in "Setup for Outcome Performance Measures") under 2 different assumptions about the user population and data availability, each informed by a specific Korean data source. The process for establishing these simulation environments was described below.

The first set of simulations aimed to closely mimic the characteristics and responses observed in a specific digital care intervention context. This was informed by data collected as part of a primary study involving approximately 1000 older adults actively engaging with an AI-based care call service (the "AI Call dataset"). This dataset provided detailed individual-level information on demographics, health status, intervention parameters (call success), and key outcomes (loneliness and self-rated health).

The second set of simulations aimed to represent a broader, more general population of older adults, reflecting demographic distributions found in large-scale surveys but potentially lacking detailed intervention-specific outcome data. This was informed by a representative sample from the national Community Health Survey (CHS) conducted in 2020 [18]. This survey provided robust population-level distributions for key demographic, health, and psychosocial variables. Table 1 describes the baseline distribution of key independent and dependent variables across the 2 datasets.

By comparing results across these 2 sets of simulations (AI Call dataset-based and CHS-based) for each outcome measure, we aimed to assess the robustness of our findings under different data conditions. The following sections detail how user profiles and the outcome generation process were specifically constructed for each simulation type.

XSL•FO
RenderX

Table 1. Baseline distribution of the key dependent and independent variables.

Variables	AI[a] Call dataset (n=1196)	CHS[b] dataset (n=72,812)
Independent variables		
Age, n (%)		
Young-old (65-74 years)	242 (20.2)	38,590 (53)
Old-old (≥75 years)	954 (79.8)	34,222 (470)
Sex, n (%)		
Male	273 (22.8)	30,326 (41.7)
Female	923 (77.2)	42,486 (58.4)
Region, n (%)		
Urban	342 (28.6)	29,846 (41)
Rural	854 (71.4)	42,966 (59)
Loneliness, n (%)		
Yes	136 (11.4)	4660 (6.4)
No	1060 (88.6)	68,152 (93.6)
Dependent variables		
Self-rated health, n (%)		
Very good	20 (1.7)	3073 (4.2)
Good	165 (13.8)	20,577 (28.3)
So-so	630 (52.7)	28,746 (39.5)
Bad	341 (28.5)	15,975 (21.9)
Very bad	39 (3.3)	4441 (6.1)
Depression, n (%)		
Nondepressed	923 (77.2)	70,264 (96.5)
Depressed	273 (22.8)	2548 (3.5)
Call success rate, mean (SD)	0.86 (0.35)	—[c]

[a]AI: artificial intelligence.

[b]CHS: community health survey.

[c]Not applicable.

Setup for Baseline Distribution

To investigate the potential of personalized program allocation under different scenarios, we constructed synthetic user profiles representative of the 2 populations described above. User profiles were defined by a combination of 7 binary demographic and health-related features: age (older or younger), sex (male or female), region (rural or urban), economic status (unstable or stable), presence of multiple chronic conditions (present or absent), functional difficulty (present or absent), social isolation (high or low), cognition (intact or not intact), and living arrangement (alone or with others). The generation process differed depending on the simulation set:

For simulations based on the AI Call dataset: Synthetic user profiles were generated by sampling directly from the characteristics of about 1000 users in the raw AI Call dataset. This ensured the simulated users closely mirrored the specific population engaging with that service. For simulations based on the CHS dataset: since only aggregated distributional information (eg, percentage breakdowns for age, sex, and region) was available from the CHS summary data, synthetic user profiles were generated differently. Each user characteristic was randomly generated independently according to the proportions reported in the CHS data. This method produced a synthetic population whose overall marginal distributions matched the CHS population structure, assuming independence between characteristics based on the available information. Similarly, we created a pool of programs with varying attributes (content type, frequency, duration, and voice type). Parameters defining individuals (based on the methods above) and the range of programs were set to represent typical variations observed in Korean older adult care contexts.

Setup for the Relationship Between Reward and Context

A critical component defining each simulation scenario is the relationship between a user's characteristics, the assigned program attributes, and the resulting outcome or reward. This relationship models the differential program effectiveness, which

XSL•FO
RenderX

personalization aims to exploit. The construction of this relationship also differed between the two simulation sets: (1) For simulations based on the AI Call dataset, the parameters (coefficients) governing the relationship between user and program features and outcomes were informed by observed patterns in the actual AI Call data. Specifically, the direction (sign) of main effect coefficients was set to match observed trends (eg, whether rural users in the data had higher or lower call success). For simplicity, the magnitude of these data-informed coefficients was held constant; and (2) for simulations based on the CHS, since the CHS data lacked information on how user characteristics relate to outcomes from specific digital care programs, the coefficients defining the relationship between user and program features and outcomes were assigned randomly. This simulated outcome variability without imposing specific assumptions based on the limited data.

In both simulation sets, to explicitly model the potential for personalization, several user-program interaction effects were included. These were based on domain knowledge and plausible scenarios. Importantly, these interaction effect scenarios were applied consistently across all 3 outcome measures. Specific examples of modeled interactions include scenarios where:

1. Women derived higher rewards from longer calls.
2. Users with functional limitations benefited more from longer calls.
3. Rural users preferred more frequent calls.
4. Those with older age preferred check-in calls.
5. Those with low cognitive function preferred check-in calls.
6. Lonely individuals preferred intervention calls less.
7. Those with chronic disease preferred health information calls.

Setup for Outcome (Performance Measures)

Within the simulation framework described above, we modeled 3 key outcomes to assess program effectiveness. The generation of these outcomes for each user-program interaction followed the relationships defined in "Setup for the Relationship Between Reward and Context" (differing based on whether it was an AI Call or CHS-based simulation). We compared the average value of each outcome measure over 100 repetitions, each involving 3000 user-item interactions (call selections) with a randomly shuffled sequence of users. The three specific outcome measures we used are:

1. Call success: modeled using a logistic linear model, where a higher success rate (completion of the call) indicated a more desirable outcome. The probability of success for a given user and call type was determined by the model described in "Setup for the Relationship Between Reward and Context."
2. Patient Health Questionnaire-2 (PHQ-2) Score [19]: modeled using a linear model, where lower scores on the PHQ-2, a measure of depressive symptoms, indicated better mental well-being. The predicted PHQ-2 score depended on user context and program assignment according to the relationships in "Setup for the Relationship Between Reward and Context." While the underlying relationship was modeled linearly, the final simulated PHQ-2 outcome for an interaction was generated as an integer using probabilistic rounding based on the linear prediction, reflecting the discrete nature of the scale.
3. Self-Reported Health (SRH) [18]: modeled using a linear model, with higher scores on the 5-point SRH scale (1 being "very bad" and 5 being "very good") representing better perceived health status. Similar to PHQ-2, the predicted SRH score was a function of user and program characteristics as defined in "Setup for the Relationship Between Reward and Context." The final simulated SRH outcome was generated as an integer on the 5-point scale using probabilistic rounding based on the linear model's prediction.

Evaluation Methods

Four allocation methods were evaluated based on 2 metrics. The first metric was the distribution of the average outcome across simulation runs. For each of the 3 performance measures, call success, PHQ-2 score, and SRH, the average value of the outcome for each of the 4 allocation methods was calculated. This represents the average scores of the subjects who received the care call according to each allocation rule.

The second metric was average cumulative regret over time, or with the increasing number of subjects. Cumulative regret represents the loss incurred from selecting a suboptimal arm over time. For the first 2 allocation methods, regret never decreases, and cumulative regret increases linearly over time. For Method 3, both regret and cumulative regret are zero. For the proposed method, RL-based precision digital care, if the method learns the optimal arm, cumulative regret tapers off after increasing initially. The tapering nature indicates that the method is learning the optimal arm.

Ethical Considerations

This study was reviewed and determined to be exempt from ethical approval by the Institutional Review Board of Seoul National University (E2408/004-007; August 14, 2024), as it used publicly available and deidentified data that did not involve direct interaction with human participants. The study analyzed anonymized AI care call data and aggregated data from the Community Health Survey, and no personally identifiable information was accessed or retained. Because the study used deidentified secondary data and simulation-based analyses, informed consent from participants was not required. All procedures were conducted in accordance with the Declaration of Helsinki.

Results

We compared 4 program allocation methods: systematic, single best, optimal, and our proposed precision digital care with Thompson Sampling, across 3 outcome measures: call success rate, PHQ-2 score (depressive symptoms), and SRH. Figures 1-3 show the outcome distributions for each method over 100 simulation runs, each involving 3000 program assignments. The summary statistics for each outcome, including mean values and standard deviations across 100 simulation runs, are presented in Table 2.

XSL•FO
RenderX

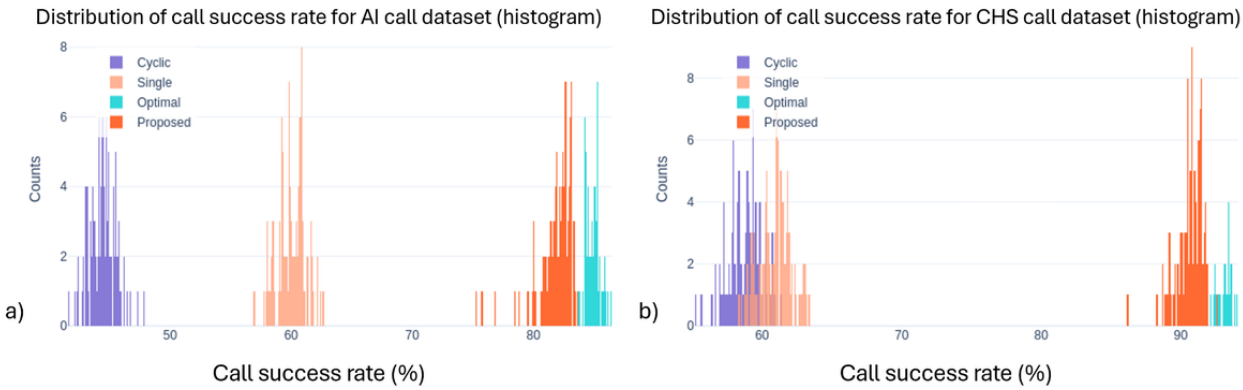
Values represent means (SDs) across 100 simulation runs (n=3000 assignments per run).

Our results demonstrate the significant potential of precision digital care with Thompson Sampling for personalized program allocation. Across all 3 measures, the proposed precision digital care consistently achieved outcomes comparable to the idealized optimal method, which assumes perfect knowledge of individual preferences and program effectiveness. Compared with systematic allocation, the proposed precision digital care approach substantially improved outcomes across all 3 measures. Call success rates increased from 44.45% to 81.86% in the AI Call dataset (84.2% relative improvement) and from 58.72% to 90.68% in CHS (54.4%). PHQ-2 depression scores decreased from 2.87 to 1.95 (32.1% reduction, AI Call) and from 2.73 to 1.60 (41.4%, CHS), while SRH scores improved from 2.94 to 3.50 (19%, AI Call) and from 3.18 to 3.88 (22%, CHS). To compare outcome distributions between systematic allocation and the proposed method, chi-square tests were used for call success rates, and Mann-Whitney *U* tests for PHQ-2 and SRH scores; all comparisons were statistically significant ($P<.001$). These findings indicate that RL-based personalization leads to substantial gains in care effectiveness compared to rule-based or static allocation strategies. These also suggest that precision digital care can effectively learn and adapt to individual needs, leading to near-optimal outcomes even without prior knowledge of individual preferences. As expected, the nonpersonalized baseline methods ("systematic" (cyclic) and "single best") yielded inferior performances compared to the personalized approaches. This highlights the limitations of assigning programs without considering individual needs and preferences.

Figure 4 illustrates cumulative regret over time or the number of users. Cumulative regret represents the loss incurred from selecting a suboptimal arm over time. It initially increases rapidly before tapering off, indicating that suboptimal arms are often selected in the beginning, followed by the selection of optimal arms. The figures also show that, depending on the outcome, learning occurs across a varying number of users. Since call success is binary, it takes more users to stabilize the learning process. This pattern suggests that precision digital care rapidly identifies optimal program assignments, reducing the frequency of ineffective interventions as it learns from accumulated user interactions. The stabilization of regret over time indicates that the RL model efficiently converges to near-optimal allocation decisions, ensuring that users receive the most suitable care programs based on their unique characteristics.

These findings underscore the potential of using contextual bandit algorithms like Thompson Sampling for personalized interventions. The ability to dynamically learn and adapt to individual needs can lead to significant improvements, especially when program effectiveness varies greatly across individuals. Overall, these results highlight the effectiveness of precision digital care in dynamically optimizing care program assignments, achieving outcomes close to the idealized optimal method. By leveraging RL, the proposed precision digital care has the potential to continuously improve care services based on user interactions, ultimately enhancing the quality and efficiency of digital health interventions.

Figure 1. Distribution of average call success rates across the four allocation methods for (a) AI call dataset and (b) community health survey dataset. The proposed precision digital care, a reinforcement learning-based approach, achieves superior performance in call success rates compared to systematic and single best allocation approaches. Note: "Cyclic" in the figure legend refers to the systematic (baseline) allocation method described in the text. AI: artificial intelligence; CHS: community health survey.



XSL•FO
RenderX

Figure 2. Distribution of average patient health questionnaire-2 scores across the four allocation methods for (a) AI call dataset and (b) community health survey dataset. The proposed precision digital care approach achieves superior performance in patient health questionnaire-2 scores compared to systematic and single best allocation approaches. Note: "Cyclic" in the figure legend refers to the systematic (baseline) allocation method described in the text. AI: artificial intelligence; CHS: community health survey; PHQ-2: Patient Health Questionnaire-2.

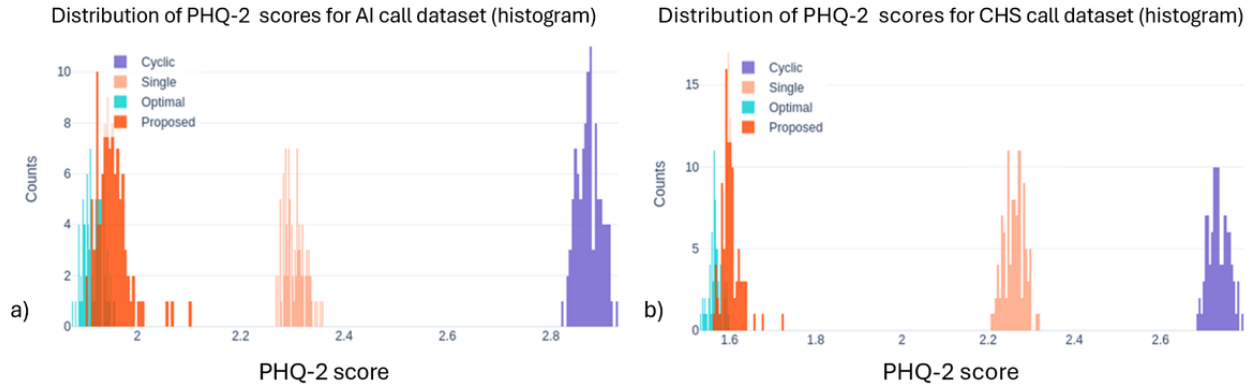


Figure 3. Distribution of average self-reported health scores across the four allocation methods for (a) AI call dataset and (b) community health survey dataset. The proposed approach achieves superior performance in self-reported health scores compared to systematic and single best allocation approaches. Note: "Cyclic" in the figure legend refers to the systematic (baseline) allocation method described in the text. AI: artificial intelligence; CHS: community health survey; SRH: self-reported health.

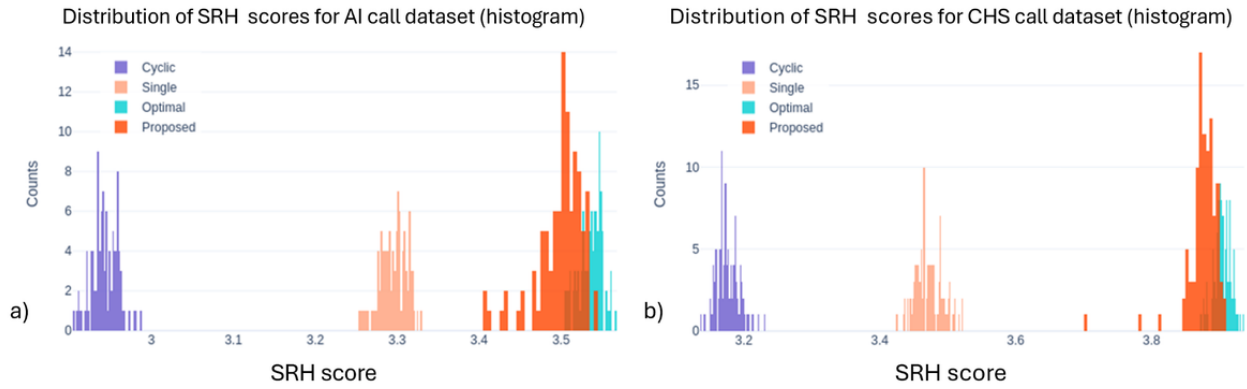


Table 2. Summary of the simulated care outcomes across allocation methods.

Outcomes and datasets	Systematic allocation (baseline)	Single best	Optimal	Precision digital care (proposed)
Call success rate[a]				
AI[b] Call, mean (SD)	44.45 (1.12)	59.98 (1.14)	84.89 (0.61)	81.86 (1.45)
CHS[c], mean (SD)	58.72 (1.20)	60.94 (1.13)	93.26 (0.48)	90.68 (0.94)
PHQ-2[d]				
AI Call, mean (SD)	2.87 (0.02)	2.3 (0.02)	1.92 (0.02)	1.95 (0.03)
CHS, mean (SD)	2.73 (0.02)	2.26 (0.02)	1.57 (0.01)	1.6 (0.02)
SRH[e]				
AI Call, mean (SD)	2.94 (0.02)	3.29 (0.02)	3.54 (0.01)	3.5 (0.03)
CHS, mean (SD)	3.18 (0.02)	3.47 (0.02)	3.9 (0.01)	3.88 (0.02)

[a]Call success rate: mean percentage of successful call completions (higher=better).

[b]AI: artificial intelligence.

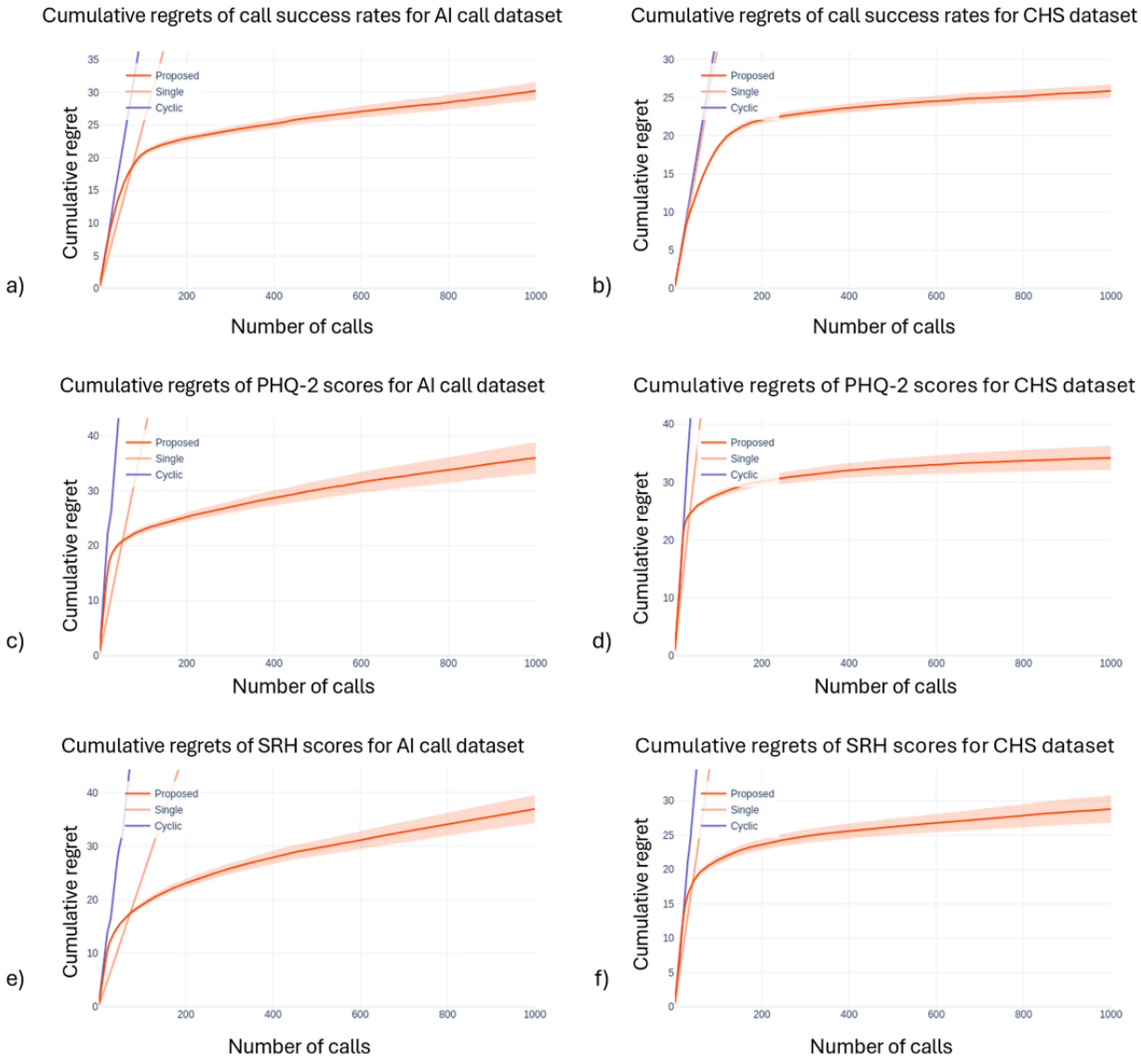
[c]CHS: Community Health Survey.

[d]PHQ-2: Patient Health Questionnaire-2; scores range from 0 to 6 (lower=fewer depressive symptoms).

[e]SRH: Self-Reported Health; scores range from 1 (very bad) to 5 (very good) (higher=better health).

XSL•FO
RenderX

Figure 4. Cumulative regret for (a) call success (AI call dataset), (b) call success (community health survey dataset), (c) Patient Health Questionnaire-2 (AI call dataset) (d) Patient Health Questionnaire-2 (CHS dataset), (e) self-reported health (AI call dataset), and (f) self-reported health (community health survey dataset) models. The rapid decline in regret for precision digital care indicates its ability to efficiently learn and optimize program allocation, significantly outperforming static allocation methods. Note: "Cyclic" in the figure legend refers to the systematic (baseline) allocation method described in the text. AI: artificial intelligence; CHS: community health survey; PHQ-2: Patient Health Questionnaire-2; SRH: self-reported health.



Discussion

Principal Findings

Our findings highlight the potential of precision digital care with Thompson Sampling in optimizing digital care service allocation for older adults. Through simulation studies, we demonstrated that personalized program allocation leads to improved care outcomes compared to nonpersonalized methods. Specifically, the precision digital care approach outperformed both systematic and single-best allocation approaches, achieving results close to the idealized theoretical upper bound. This underscores the importance of hyper-personalization in digital health care, particularly in aging populations facing growing care demands and workforce shortages.

The outcomes evaluated in this simulation study represent key dimensions of effective care delivery for older adults living alone. Call success indicates sustained engagement with care services and regular monitoring, both critical for timely detection of health or safety concerns and prevention of unattended deaths-an emerging social problem in Korea [20]. PHQ-2 scores represent depressive symptoms that affect daily functioning, social participation, and self-care capacity, while self-reported health reflects subjective well-being and perceived ability to age in place.

In real-world implementation, improvements in these outcomes would support the delivery of responsive, person-centered care. Importantly, precision digital care is designed to augment (not replace) human care relationships. In Korea's context, direct care workers often face overtime burdens and emotional exhaustion, contributing to workforce shortages [2,3]. The proposed approach optimizes routine care allocation through AI-driven monitoring, allowing human workers to focus on complex cases requiring empathy, clinical judgment, and relational continuity. This human-AI collaboration model

XSL•FO
RenderX

addresses workforce sustainability while preserving the irreplaceable value of human connection in care relationships. Rather than automating care relationships, the system automates administrative burden, the repetitive monitoring tasks that overwhelm workers, enabling them to provide higher-quality human care where it matters most. Together, these outcomes support the goal of enabling older adults to age with dignity and autonomy while ensuring appropriate care allocation.

These findings align with previous research emphasizing the importance of personalization in digital health interventions for older adults. A scoping review of digital health interventions targeting older populations found that personalized features improve self-management and psychosocial health, but also noted that experimental studies validating these effects are scarce [21]. Our study contributes to this gap by providing a systematic, AI-driven approach to personalization, demonstrating through simulation that RL can effectively optimize program allocation based on user characteristics. Unlike previous work, which primarily relies on static program designs, our approach adapts dynamically, ensuring that interventions are continually optimized over time.

Our study also relates to research exploring contextual bandit algorithms in mobile health settings, where RL has been used to optimize treatment timing and intervention strategies [22]. Existing studies have demonstrated the feasibility of such approaches in personalized patient engagement and behavior change interventions [23]. However, most prior applications focus on mobile health and preventive care rather than digital older adult care services. Our findings extend these applications by showing that contextual bandits can optimize not just treatment timing but also care program selection for older adults in government-supported digital health initiatives. This represents a novel application of RL in aging-related digital interventions, offering a scalable solution for optimizing care programs at the population level.

Beyond demonstrating the effectiveness of RL for digital care optimization, our findings have broader implications for the scalability of AI-driven decision models in government-funded care programs. As governments invest in digital health solutions to address aging populations and workforce shortages, integrating AI-driven program allocation systems can improve care efficiency without requiring extensive prelabeled datasets [24]. The Korean AI care call program exemplifies this trend, yet current models lack individualized program allocation strategies. By applying the proposed precision digital care approach for personalization of care in such settings, policymakers could transition from standardized digital interventions to truly personalized care, ensuring that digital services maximize impact and cost efficiency. More broadly, the integration of AI-driven decision-making into digital care services aligns with global trends toward precision health and personalized medicine, advocating for a shift away from static care delivery models.

Despite these promising findings, several limitations must be acknowledged. This study relied on synthetic data and simulations, rather than real-world implementation. Although the data was constructed using existing survey datasets and AI care call user characteristics, real-world trials are necessary to validate these findings in practice. Additionally, the current model was developed within the Korean digital care context, and further research is needed to assess the adaptability of precision digital care in different health care systems and cultural settings.

This study's focus on low-income, socially isolated older adults living alone reflects an intentional equity commitment. These populations face compounded vulnerabilities, limited access to family support, health care services, and social resources, placing them at increased risk of adverse outcomes [25]. Precision digital care has particular promise for reducing geographic and functional disparities. Our simulations incorporated domain-informed interaction effects in which rural users and those with functional limitations were modeled as benefiting from different care configurations. For example, based on observed patterns in Korean older adult care contexts, rural older adults were modeled as preferring more frequent calls (potentially compensating for limited in-person service access), while those with functional difficulties were assumed to benefit from longer interaction times (allowing more comprehensive needs assessment). Similarly, older adults with cognitive limitations were modeled as engaging better with simpler check-in style calls rather than complex health information content.

By identifying and responding to these differential needs, precision digital care can help bridge urban-rural care gaps, accommodate varying cognitive and functional capacities, and ensure that digital care innovations reduce rather than exacerbate existing inequities. However, realizing this potential requires careful attention to digital literacy barriers, technology access, and culturally appropriate implementation, critical considerations for future real-world validation studies.

Beyond technical and operational concerns, AI-driven personalization in digital health raises ethical considerations, such as potential bias in program allocation, equity in digital access, and privacy concerns related to AI-driven recommendations [26]. Future studies should explore fairness-aware RL approaches to ensure equitable care distribution, particularly for older adults who may have limited digital literacy or reduced access to technology. Furthermore, while RL-based precision digital care optimizes individualized program assignments, its integration into broader multilevel care frameworks, such as clinical care coordination, home-based services, and telemedicine ecosystems, remains an open challenge. Ensuring seamless interoperability with existing digital health platforms will be essential for widespread adoption.

To address these limitations, we plan to conduct real-world validation studies in a publicly funded AI care call program in Korea, assessing precision digital care's impact on user satisfaction, health outcomes, and long-term engagement, with an expanded set of objective measures (eg, health care use and device-measured activity) to complement self-reports. Although equity was not quantitatively evaluated in this study, the intervention setting, a publicly funded AI care call program for low-income, socially vulnerable older adults living alone,

XSL•FO
RenderX

reflects an inherent focus on underserved populations. Future research should also explore hybrid AI models, integrating deep learning and RL to further refine personalized digital interventions. Additionally, a critical next step is to examine how RL-based precision digital care performs across diverse populations, ensuring that AI-driven personalization remains both effective and equitable across different health care systems.

Conclusions

This study introduced precision digital care, an AI-driven RL model for optimizing personalized digital care service allocation. Our simulation-based evaluation demonstrated that precision digital care significantly improves care outcomes by dynamically adapting interventions to individual user characteristics. These findings highlight the potential of AI-driven personalization in enhancing the efficiency and effectiveness of digital health interventions, particularly in aging populations where scalable, individualized care is increasingly needed.

As digital care programs expand, ensuring personalized, data-driven decision-making will be critical in maximizing their impact. Precision digital care offers a scalable approach to deliver dignified, person-centered care at scale. By using RL to understand and respond to individual needs, the proposed framework can help older adults maintain autonomy, social connection, and well-being while supporting sustainable care systems. The ultimate goal is not simply algorithmic optimization but improving lived experiences and quality of life in later years through adaptive, humane digital care that honors individual dignity and preferences. However, further real-world validation is required to confirm the effectiveness of precision digital care beyond simulations and to address practical implementation challenges.

Future research should focus on scaling and adapting precision digital care to diverse health care settings, ensuring ethical AI deployment, and integrating hybrid AI models to enhance personalization. Establishing fair, transparent, and explainable AI systems will be key to promoting trust and equitable access in digital health services for aging populations worldwide.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (RS-2023-00276954) and the AI Institute at Seoul National University (AIIS) in 2025.

Data Availability

The datasets generated and/or analyzed during this study are not publicly available due to Institutional Review Board regulations, but are available from the corresponding author on reasonable request.

Authors' Contributions

HK contributed to conceptualization, funding acquisition, and led the writing of the original draft. MCP contributed to conceptualization, methodology, and writing of the original draft. KL was responsible for formal analysis and visualization, and JYY contributed to data curation. All authors participated in drafting sections of the manuscript and in reviewing and editing the final version.

Conflicts of Interest

None declared.

References

1. Beard JR, Officer AM, Cassels AK. The world report on ageing and health. Gerontologist 2016;56(Suppl 2):S163-S166. [doi: 10.1093/geront/gnw037] [Medline: 26994257]
2. OECD. Who cares? Attracting and retaining care workers for the elderly. In: OECD Health Policy Studies. Paris: OECD Publishing; 2020.
3. van Elk F, Robroek SJW, Burdorf A, Oude Hengel KM. Impact of the COVID-19 pandemic on psychosocial work factors and emotional exhaustion among workers in the healthcare sector: a longitudinal study among 1915 Dutch workers. Occup Environ Med 2023;80(1):27-33 [FREE Full text] [doi: 10.1136/oemed-2022-108478] [Medline: 36424169]
4. Gaskell J, Stoker G, Jennings W, Devine D. Covid‐19 and the blunders of our governments: long‐run system failings aggravated by political choices. Political Q. 2020;91(3):523-533 [FREE Full text] [doi: 10.1111/1467-923x.12894]
5. OECD. The Innovation Imperative: Contributing to Productivity, Growth and Well-Being. Paris: OECD Publishing; 2015.
6. Coughlin JF. Technology, innovation, and developing a NexGen aging services workforce. Publ Pol Aging Rep 2014;24(1):6-9 [FREE Full text] [doi: 10.1093/ppar/prt009]
7. Golinelli D, Boetto E, Carullo G, Nuzzolese AG, Landini MP, Fantini MP. Adoption of digital technologies in health care during the COVID-19 pandemic: systematic review of early scientific literature. J Med Internet Res 2020;22(11):e22280 [FREE Full text] [doi: 10.2196/22280] [Medline: 33079693]
8. Allen B. The promise of explainable AI in digital health for precision medicine: a systematic review. J Pers Med 2024;14(3):277 [FREE Full text] [doi: 10.3390/jpm14030277] [Medline: 38541019]

XSL•FO
RenderX

9. Topol EJ. A decade of digital medicine innovation. Sci Transl Med 2019;11(498):eaaw7610 [FREE Full text] [doi: 10.1126/scitranslmed.aaw7610] [Medline: 31243153]
10. Abul-Husn NS, Kenny EE. Personalized medicine and the power of electronic health records. Cell 2019;177(1):58-69 [FREE Full text] [doi: 10.1016/j.cell.2019.02.039] [Medline: 30901549]
11. Aljohani A. AI-driven decision-making for personalized elderly care: a fuzzy MCDM-based framework for enhancing treatment recommendations. BMC Med Inform Decis Mak 2025;25(1):119 [FREE Full text] [doi: 10.1186/s12911-025-02953-5] [Medline: 40055665]
12. Kim GS, Hong YS, Lee TH, Paik MC, Kim H. Bandit-supported care planning for older people with complex health and care needs. ArXiv Preprint posted online on March 13, 2023. [doi: 10.48550/arXiv.2303.07053]
13. Hoogendijk EO, Onder G, Smalbil L, Vetrano DL, Hirdes JP, Howard EP. Optimising the care for older persons with complex chronic conditions in home care and nursing homes: design and protocol of I-CARE4OLD, an observational study using real-world data. BMJ Open 2023;13(6):e072399 [FREE Full text] [doi: 10.1136/bmjopen-2023-072399] [Medline: 37385750]
14. AI checks on older adults living alone during the Chuseok holiday period [Web page in Korean]. Ministry of Health and Welfare (MOHW), Republic of Korea. URL: https://www.mohw.go.kr/board.es?mid=a10503010100&bid=0027&act=view&list_no=1483091&tag=&nPage=1 [accessed 2025-05-12]
15. Valkonen P, Kujala S. Older adults' emotional user experiences with digital health services. In: Digital Health and Wireless Solutions: NCDHWS 2024.: Springer; 2024 Presented at: Communications in Computer and Information Science; May 05, 2024; Cham. [doi: 10.1007/978-3-031-59080-1_10]
16. World Report on Ageing and Health. Geneva: World Health Organization; 2015.
17. Agrawal S, Goyal N. Thompson sampling for contextual bandits with linear payoffs. ArXiv Preprint posted online on February 3, 2014. [doi: 10.48550/arXiv.1209.3352]
18. Kang YW, Ko YS, Kim YJ, Sung KM, Kim HJ, Choi HY, et al. Korea community health survey data profiles. Osong Public Health Res Perspect 2015;6(3):211-217 [FREE Full text] [doi: 10.1016/j.phrp.2015.05.003] [Medline: 26430619]
19. Kroenke K, Spitzer RL, Williams JB. The patient health questionnaire-2: validity of a two-item depression screener. Med Care 2003;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
20. Lyu J, Kim YB, Park JS. AI services to prevent solitary death in South Korea. Int J Contents 2025;21(2):61-66. [doi: 10.5392/ijoc.2025.21.2.061]
21. Hwang M, Jiang Y. Personalization in digital health interventions for older adults with cancer: a scoping review. J Geriatr Oncol 2023;14(8):101652. [doi: 10.1016/j.jgo.2023.101652] [Medline: 37866009]
22. Lei H, Lu Y, Tewari A, Murphy S. An actor-critic contextual bandit algorithm for personalized mobile health interventions. arXiv 2017 Jun 28 [FREE Full text] [doi: 10.48550/arXiv.1706.09090]
23. Bi M, Klasnja P, Choudhury T, Tewari A, Murphy S. Optimizing mHealth interventions with a bandit. In: Digital Phenotyping and Mobile Sensing. Studies in Neuroscience, Psychology and Behavioral Economics. Cham: Springer; 2019:79-94.
24. Aerts A, Bogdan-Martin D. Leveraging data and AI to deliver on the promise of digital health. Int J Med Inform 2021;150:104456 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104456] [Medline: 33866232]
25. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System. Washington, DC: National Academies Press; 2020.
26. Maeckelberghe E, Zdunek K, Marceglia S, Farsides B, Rigby M. The ethical challenges of personalized digital health. Front Med 2023;10:1123863 [FREE Full text] [doi: 10.3389/fmed.2023.1123863] [Medline: 37404804]

Abbreviations

AI: artificial intelligence
CHS: community health survey
PHQ-2: Patient Health Questionnaire-2
RL: reinforcement learning
SRH: self-reported health

Edited by Y Jiang; submitted 29.Oct.2025; peer-reviewed by G Kim, R Singh; comments to author 09.Jan.2026; revised version received 28.Feb.2026; accepted 09.Mar.2026; published 13.Apr.2026.

Please cite as:
Kim H, Lee K, Yi JY, Paik MC
Personalized Digital Care Program Allocation for Older Adults: Reinforcement Learning-Based Simulation Study
JMIR Aging 2026;9:e86710
URL: https://aging.jmir.org/2026/1/e86710
doi:10.2196/86710
PMID:

XSL•FO
RenderX

©Hongsoo Kim, Kyungbok Lee, Jae Yoon Yi, Myunghee Cho Paik. Originally published in JMIR Aging (https://aging.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Speech Recognition–Based Dietary Assessment Tool for Older Adults: Validation and Usability Study

Yoonjee Sung[1], MSc; Soyoung Jung[2], MSc; Hae Jin Kang[2], PhD; So Young Moon[3], MD, PhD; Jee Hyang Jeong[4], MD, PhD; Seong Hye Choi[5], MD, PhD; Eun-Hye Lee[4], MD, PhD; Yoo Kyoung Park[2], PhD

[1]Department of Medical Nutrition, Graduate School of East-West Medical Science, Kyung Hee University, Gyeonggi, Republic of Korea

[2]Department of Medical Nutrition, Graduate School of East-West Medical Science AgeTech-Service Convergence, Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si,, Gyeonggi, Republic of Korea

[3]Department of Neurology, School of Medicine, Ajou University, Gyeonggi, Republic of Korea

[4]Department of Neurology, College of Medicine, Ewha Womans University, Seoul, Republic of Korea

[5]Department of Neurology, College of Medicine, Inha University, Incheon, Republic of Korea

Corresponding Author:
Yoo Kyoung Park, PhD
Department of Medical Nutrition, Graduate School of East-West Medical Science AgeTech-Service Convergence, Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si,, Gyeonggi, Republic of Korea

Related Article:

Correction of: https://aging.jmir.org/2026/1/e81336

(JMIR Aging 2026;9:e104170) doi:10.2196/104170

In “Speech Recognition–Based Dietary Assessment Tool for Older Adults: Validation and Usability Study” [1], the authors noted one error.

The corresponding author’s name appeared as:

Yoo Kyung Park

The name has been changed to:

Yoo Kyoung Park

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Sung Y, Jung S, Kang HJ, et al. Speech recognition–based dietary assessment tool for older adults: validation and usability study. JMIR Aging 2026;9:e81336. [doi: 10.2196/81336] [Medline: 13211599]

Submitted 09.Jun.2026; this is a non–peer-reviewed article;accepted 09.Jun.2026; published 16.Jul.2026.

Please cite as:
Sung Y, Jung S, Kang HJ, Moon SY, Jeong JH, Choi SH, Lee EH, Park YK
Correction: Speech Recognition–Based Dietary Assessment Tool for Older Adults: Validation and Usability Study
JMIR Aging 2026;9:e104170
URL: https://aging.jmir.org/2026/1/e104170
doi:10.2196/104170

© Yoonjee Sung, Soyoung Jung, Hae Jin Kang, So Young Moon, Jee Hyang Jeong, Seong Hye Choi, Eun-Hye Lee, Yoo Kyoung Park. Originally published in JMIR Aging (https://aging.jmir.org), 16.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Domain-Specific Associations of Physical Activity With Phase Angle and Extracellular Water–to–Total Body Water Ratio in Middle-Aged or Older Adults: Cross-Sectional Study

Seong-Uk Baek[1], MD; Jin-Ha Yoon[2,3,4], PhD

[1]Graduate School, Yonsei University College of Medicine, Seoul, Seoul, Republic of Korea

[2]Department of Preventive Medicine, Yonsei University College of Medicine, 50-1, Yonsei-ro, Seodaemungu, Seoul, Republic of Korea

[3]The Institute for Occupational Health, Yonsei University College of Medicine, Seoul, Republic of Korea

[4]Institute for Innovation in Digital Healthcare, Yonsei University Health System, Seoul, Republic of Korea

Corresponding Author:
Jin-Ha Yoon, PhD
Department of Preventive Medicine, Yonsei University College of Medicine, 50-1, Yonsei-ro, Seodaemungu, Seoul, Republic of Korea

Related Article:

Correction of: https://aging.jmir.org/2026/1/e90158

(JMIR Aging 2026;9:e114943) doi:10.2196/114943

In "Domain-Specific Associations of Physical Activity With Phase Angle and Extracellular Water–to–Total Body Water Ratio in Middle-Aged or Older Adults: Cross-Sectional Study" [1], the authors noted one error.

The Funding section was revised from the following:

> *This research was supported by a grant from the Korea Occupational Safety and Health Agency (KOSHA) under the project "Advancing Occupational National Safety and Health Data Management Systems and Establishing a Unified Integrated Platform" (grant RS-2026-25549947). This research was also supported by a grant from the National Bio Bigdata Project, funded by 4 Korean ministries (Ministry of Health and Welfare, Ministry of Science and ICT, Ministry of Trade, Industry and Resources, and Korea Disease Control and Prevention Agency) of Korea (grant HI20C1234).*

The Funding section now reads:

> *This research was supported by a grant from the National Bio Bigdata Project, funded by four ministries (Ministry of Health and Welfare; Ministry of Science and ICT; Ministry of Trade, Industry and Resources; and Korea Disease Control and Prevention Agency) of Korea (grant RS-2024 - 00462006).*

The correction will appear in the online version of the paper on the JMIR Publications website together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Baek SU, Yoon JH. Domain-specific associations of physical activity with phase angle and extracellular water–to–total body water ratio in middle-aged or older adults: cross-sectional study. JMIR Aging 2026 Sep 28;9:e90158. [doi: 10.2196/90158] [Medline: 42804760]

Submitted 21.Oct.2026; this is a non–peer-reviewed article;accepted 28.Oct.2026; published 18.Nov.2026.

Please cite as:
Baek SU, Yoon JH
Correction: Domain-Specific Associations of Physical Activity With Phase Angle and Extracellular Water–to–Total Body Water Ratio in Middle-Aged or Older Adults: Cross-Sectional Study
JMIR Aging 2026;9:e114943
URL: https://aging.jmir.org/2026/1/e114943
doi:10.2196/114943

XSL•FO
RenderX

© Seong-Uk Baek, Jin-Ha Yoon. Originally published in JMIR Aging (https://aging.jmir.org), 18.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Development and Validation of Machine Learning Models for Predicting Falls Among Hospitalized Older Adults: Retrospective Cross-Sectional Study

Xiyao Yang[1*], MSc; Juan Ren[1*], MSc; Dan Su[2,3*], PhD; Manzhen Bao[1], MSc; Miao Zhang[1], MSc; Xiaoming Chen[1], BSc; Yanhua Li[1], BSc; Zonggui Wang[1], BSc; Xiujing Dai[1], BSc; Zengzeng Wei[1], BSc; Shuiyu Zhang[2], BSc; Yuxin Zhang[1], MSc; Juan Li[1], BSc; Xiaolin Li[1], BSc; Junjin Xu[1], BSc; Nan Mo[1], MSc

[1]Second Hospital of Anhui Medical University, No.678 Furong Road, Shushan District, Hefei, China

[2]School of Nursing, Anhui Medical University, Hefei, China

[3]The Taikang Health and Wellness Industry Research Institute, Anhui Medical University, Hefei, China

[*]these authors contributed equally

Corresponding Author:
Nan Mo, MSc
Second Hospital of Anhui Medical University, No.678 Furong Road, Shushan District, Hefei, China

Related Article:

Correction of: https://aging.jmir.org/2026/1/e80602

(JMIR Aging 2026;9:e101284) doi:10.2196/101284

In "Development and Validation of Machine Learning Models for Predicting Falls Among Hospitalized Older Adults: Retrospective Cross-Sectional Study" [1], the authors noted one error.

In the Selection of Predictor Variables section, the following sentence was replaced:

> *Univariate LR identified 27 potential predictors (P<.05) in the training set, as detailed in Table S3 of Multimedia Appendix 1.*

The sentence now reads:

> *Univariate LR identified 26 potential predictors (P<.05) in the training set. The mBI did not reach statistical significance in the univariate analysis, but its P value was .054, which is very close to the threshold. Given that the mBI is an important clinical indicator associated with fall risk and to avoid omitting potentially influential factors due to a rigid threshold, we decided after expert discussion to include it as the 27th variable in the subsequent feature selection, as detailed in Table S3 of Multimedia Appendix 1.*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Yang X, Ren J, Su D, et al. Development and validation of machine learning models for predicting falls among hospitalized older adults: retrospective cross-sectional study. JMIR Aging 2026 Jan 5;9:e80602. [doi: 10.2196/80602] [Medline: 41489988]

Submitted 16.May.2026; this is a non–peer-reviewed article;accepted 29.May.2026; published 13.Aug.2026.

Please cite as:
Yang X, Ren J, Su D, Bao M, Zhang M, Chen X, Li Y, Wang Z, Dai X, Wei Z, Zhang S, Zhang Y, Li J, Li X, Xu J, Mo N
Correction: Development and Validation of Machine Learning Models for Predicting Falls Among Hospitalized Older Adults: Retrospective Cross-Sectional Study
JMIR Aging 2026;9:e101284
URL: https://aging.jmir.org/2026/1/e101284
doi:10.2196/101284

XSL•FO
RenderX

© Xiyao Yang, Juan Ren, Dan Su, Manzhen Bao, Miao Zhang, Xiaoming Chen, Yanhua Li, Zonggui Wang, Xiujing Dai, Zengzeng Wei, Shuiyu Zhang, Yuxin Zhang, Juan Li, Xiaolin Li, Junjin Xu, Nan Mo. Originally published in JMIR Aging (https://aging.jmir.org), 13.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Feasibility and Preliminary Effects of a Pressure Sensor–Based Exergame on Cognitive and Sensorimotor Frailty in Low-Income Older Adults Receiving Public Community Care Services: Pilot Randomized Controlled Trial

Sunmi Song[1,2*], MA, PhD; Suk-Chan Hahm[3*], PT, PhD; Seyun Chang[4], BS; Jinseung Lee[1,2], BS; Heewon Kang[5], MPH, PhD; Sanghun Lee[4], BS; Junesun Kim[1,6], PT, PhD

[1]Rehabilitation Science Program, Department of Health Science, Graduate School, Korea University, Seoul, Seoul, Republic of Korea

[2]Transdisciplinary Major in Learning Health Systems, Department of Healthcare Sciences, Graduate School, Korea University, Seoul, Seoul, Republic of Korea

[3]Graduate School of Integrative Medicine and Health, Cha University, Pocheon, Gyeonggi-do, Republic of Korea

[4]Geumcheon-gu, 205 Gasan digital 1-ro, MIDAS Health & Technology Co., Seoul, Seoul, Republic of Korea

[5]Department of Public Health Administration, Daejin University, Pocheon, Gyeonggi-do, Republic of Korea

[6]Department of Physical Therapy/Health and Environmental Science, Undergraduate School, College of Health Science, Korea University, Seoul, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Junesun Kim, PT, PhD
Department of Physical Therapy/Health and Environmental Science, Undergraduate School
College of Health Science
Korea University
665 Hana Health Science Building B
145 Anam-ro, Sungbuk-gu
Seoul, Seoul, 02841
Republic of Korea
Phone: 82 232905689
Email: junokim@korea.ac.kr

Related Article:

Correction of: https://aging.jmir.org/2026/1/e93467

(JMIR Aging 2026;9:e109510) doi:10.2196/109510

In “Feasibility and Preliminary Effects of a Pressure Sensor–Based Exergame on Cognitive and Sensorimotor Frailty in Low-Income Older Adults Receiving Public Community Care Services: Pilot Randomized Controlled Trial” [1], the authors wish to make the following correction.

The Funding section has been corrected from the following:

> *This study was supported by the ANCHOR through the Seoul ANCHOR Center, funded by the Ministry of Education (MOE) and the Seoul Metropolitan Government (2026-ANCHOR-01-003-01), the Basic Research Lab Program (BRL) of the National Research Foundation of Korea (NRF) (RS-2025-02303227), and the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science and ICT (MSIT) (grant number RS-2024-00348012).*

The Funding section now reads:

> *This study was supported by the ANCHOR through the Seoul ANCHOR Center, funded by the Ministry of Education (MOE) and the Seoul Metropolitan Government (2026-ANCHOR-01-003-01), the Pioneer Research Center Program (RS-2022-NR067566) through the National Research Foundation of Korea, and the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science and ICT (MSIT) (RS-2024-00348012).*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

XSL•FO
RenderX

Reference

1. Song S, Hahm SC, Chang S, Lee J, Kang H, Lee SH, et al. Feasibility and preliminary effects of a pressure sensor-based exergame on cognitive and sensorimotor frailty in low-income older adults receiving public community care services: pilot randomized controlled trial. JMIR Aging 2026 Jul 24;9:e93467 [FREE Full text] [doi: 10.2196/93467] [Medline: 42497405]

Submitted 13.Aug.2026; this is a non–peer-reviewed article;accepted 19.Aug.2026; published 09.Sep.2026.

Please cite as:
Song S, Hahm SC, Chang S, Lee J, Kang H, Lee S, Kim J
Correction: Feasibility and Preliminary Effects of a Pressure Sensor–Based Exergame on Cognitive and Sensorimotor Frailty in Low-Income Older Adults Receiving Public Community Care Services: Pilot Randomized Controlled Trial
JMIR Aging 2026;9:e109510
URL: https://aging.jmir.org/2026/1/e109510
doi:10.2196/109510
PMID:

©Sunmi Song, Suk-Chan Hahm, Seyun Chang, Jinseung Lee, Heewon Kang, Sanghun Lee, Junesun Kim. Originally published in JMIR Aging (https://aging.jmir.org), 09.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Temporal Trends in the Prevalence of Frailty in England From 2002 to 2022, With Forecasts to 2040: Longitudinal Secondary Analysis of a National Cohort Study

Dongfeng Tang[1], MSc; Katie Jane Sheehan[2], PhD; Simon Conroy[3,4,5], PhD; Aicha Goubar[1], PhD; Julie Whitney[1,6], PhD; Yanzhong Wang[1,7], PhD; Matthew D L O'Connell[1], PhD

[1]School of Life Course and Population Sciences, Faculty of Life Science and Medicine, King's College London, Great Maze Pond, Guy's Campus, London, England, United Kingdom

[2]Center for Bone & Joint Health, Blizard Institute, Queen Mary University of London, London, England, United Kingdom

[3]Medical Research Council Unit for Lifelong Health and Ageing, University College London, London, England, United Kingdom

[4]Royal London Hospital, London, England, United Kingdom

[5]Wolfson Institute of Population Health, Queen Mary University of London, London, England, United Kingdom

[6]Department of Clinical Gerontology, King's College Hospital NHS Foundation Trust, London, England, United Kingdom

[7]NIHR Applied Research Collaboration (ARC) South London, London, England, United Kingdom

Corresponding Author:
Dongfeng Tang, MSc
School of Life Course and Population Sciences, Faculty of Life Science and Medicine, King's College London, Great Maze Pond, Guy's Campus, London, England, United Kingdom

Abstract

Background: The increasing life expectancy provides an opportunity for healthy longevity, but frailty is a potential challenge. Existing studies of frailty prevalence in England are largely cross-sectional, use heterogeneous frailty definitions, and rarely break down population-level drivers or project future burden, limiting their use for health and social care planning.

Objective: This study aimed to explore temporal trends of frailty prevalence and project scenario-based future prevalence up to 2040 for the community-dwelling population aged 50 years and older in England.

Methods: Prevalence of frailty (measured using a frailty index) in the overall sample and subgroups of age, sex, educational level, and marital status from 2002 to 2022 was estimated among the community-dwelling population aged 50 years and older using 10 waves of nationally representative data from the English Longitudinal Study of Ageing (ELSA; n=6226 to 11,218 across waves). The number of individuals with frailty was also derived in each year. Joinpoint regression models were used to explore temporal trends of frailty prevalence. A linear probability regression model was performed using pooled data from 10 waves of the ELSA to explore driving factors and project future prevalence. Decomposition analysis was used to estimate the contribution of individual factors to the change in prevalence.

Results: A statistically nonsignificant (*P*=.61) increase of 2.3 percentage points from 15.5% (1738/11,218) in 2002 to 17.8% (1108/6226) in 2022 was observed, with the number of individuals with frailty growing from 2,560,122 to 3,791,038 among community-dwelling adults aged 50 years and older in England. Those with advanced age, who were female, had a lower educational level, and/or were single had a consistently higher prevalence of frailty. Opposing contributions netted to this overall change: improvements in sociodemographic conditions and educational attainment and healthier lifestyle behaviors offset 222.5%, 94.9%, and 51.6% of the observed increase, respectively. Under scenario-based projections, prevalence was estimated at 20.5% (95% CI 19.2%-21.9%) in 2040, corresponding to 5,213,689 people with frailty.

Conclusions: Despite a modest increase in prevalence among the community-dwelling population aged 50 years and older, the substantial growth in the absolute number of individuals with frailty indicates a likely rise in future care demands. The study contributes a 20-year trend, decomposition, and forecast framework for community-dwelling older adults. Findings support prioritizing broader sociodemographic development alongside educational attainment and healthy lifestyles, with potential policy approaches including income and social protection, housing improvements, access to health care, and prevention. Key limitations include self-reported measures, exclusion of care home residents, and potential influence of unmeasured effects.

(JMIR Aging 2026;9:e98665) doi:10.2196/98665

KEYWORDS

frailty; aging; prevalence; trend; projection

XSL•FO
RenderX

Introduction

The increasing life expectancy in high-income countries since at least 1900 represents one of the remarkable achievements of modern society, reflecting significant advances in social and economic development [1,2]. This longevity offers immense potential for societal contribution and enriched later life, and therefore, it is important to promote healthy life expectancy. In 2018, the UK government set a responsive goal to increase healthy life expectancy by at least 5 years by 2035 [3]. However, frailty is a potential challenge to healthy longevity. It is a geriatric syndrome associated with aging, characterized by multiple functional declines and increased vulnerability to stressors and associated with various adverse health outcomes, including disability, hospitalization, admission to nursing homes, and mortality [4,5]. Therefore, it is a critical target for innovative approaches to meet growing health and social care needs to ensure that those extended years are lived in good health. There is a need for health and care services to balance priorities, needs, and resources to support a response to population frailty [6]. Such informed policymaking and planning require a better understanding of the present and potential future prevalence of frailty and associated explanatory factors.

Frailty prevalence in the United Kingdom has been reported to range from 4.3% to 38.9%, largely due to differences in frailty measurement tools and study characteristics across the literature [7-10]. However, most of the studies are cross-sectional, limiting understanding of how this prevalence has changed over time [7-9]. In the study by Walsh et al [10] based on routine primary care data, the prevalence of frailty was found to increase from 26.5% to 38.9% (electronic frailty index [FI]) from 2006 to 2017 among adults aged 50 years and older. While this suggests an upward trend, understanding how this trend may develop into the future requires consideration of the evolving picture of population-level risk factors. Although risk factors for within-individual changes in frailty have been explored [10,11], the driving factors for population-level changes in frailty prevalence have received little attention. Studies describing temporal trends in the prevalence of age-related disorders, including musculoskeletal disorders, Parkinson disease, and age-related macular degeneration, have identified contributors such as demographic shifts, epidemiological transitions, and changes in exposures to risk factors [12-14]. These findings collectively suggest that a multidimensional lens is necessary to understand the factors driving the temporal trends in frailty prevalence. Moreover, an understanding of the evolution of frailty into the future is required to inform policy planning and target interventions.

Understanding how frailty prevalence has changed and will change is central to planning health and social care for an aging population. This study aimed to examine the temporal trends in frailty prevalence among the community-dwelling population aged 50 years and older in England from 2002 to 2022 and project the future prevalence up to 2040. It also sought to identify key factors driving changes in population frailty and estimate their effects on future projections, thereby providing a basis for targeting prevention and allocating future care resources.

Methods

Study Design and Participants

Ten waves of data (baseline commenced in 2002-2003, and follow-up surveys were conducted every 2 years up to 2018-2019, with another follow-up in 2021-2023) from the English Longitudinal Study of Ageing (ELSA) were used (n=6226 to 11,218 across waves). The ELSA is a representative survey of the population aged 50 years and older living in private households in England, involving the collection of economic, social, psychological, cognitive, health, biological, and genetic data [15]. Participants in the ELSA were initially recruited from households that participated in the Health Survey for England in 1998, 1999, and 2001, serving as the foundation for subsequent follow-ups. A multistage stratified probability sampling approach was used, starting with the selection of postcode sectors followed by the sampling of household addresses. In wave 1 of the ELSA, 12,099 participants from 11,578 households were included. To maintain representation of younger age groups as the original cohort aged, refreshment samples were introduced in waves 3, 4, 6, 7, 9, and 10 using younger participants from the Health Survey for England [16]. Inverse probability weighting was further used to reduce bias from nonresponse and ensure that the samples were representative of the target population. Weights were calculated for participants, and participants not representative received a weight of 0 or were coded as system missing. The sociodemographic characteristics of ELSA participants are similar to those of the national census, which suggests that the ELSA sample is broadly representative of the population of England [15]. More details about the study design and weighting strategy are provided in the publicly available online ELSA user guide [17].

Procedures

Frailty was assessed using a previously validated (in the ELSA) 31-item FI constructed using the health deficit accumulation approach [18,19]. Multiple dimensions, including self-reported health conditions, functional limitations, self-reported health status, mental health, and cognitive status, were assessed to capture the overall health status of individuals [18]. Details regarding specific items drawn to construct the FI are described in Section 1 in Multimedia Appendix 1. Participants with more than 20% data missingness for the FI were excluded [20]. The FI was computed as the ratio of present health deficits to the total potential deficits, yielding a continuous score from total fitness (0) to total frailty (1). An FI of 0.25 or higher was coded as frailty based on defined cutoff points from the study by Song et al [21]. Both continuous FI and categorical frailty were used for further analyses.

Explanatory variables from different dimensions selected based on the data availability in the ELSA and biological plausibility were used to explain the potential temporal changes in frailty and project its future prevalence. The following variables were included: age (50-64, 65-74, 75-84, and ≥85 years), sex (male and female), educational level (lower, indicating an educational level below the upper secondary level and/or having finished education before the age of 16 years; intermediate, indicating

XSL•FO
RenderX

upper secondary education or a vocational degree and/or having finished education at the age of 16-18 years; and higher, indicating a bachelor's degree or higher levels of tertiary education or having finished education at the age of 19 years or later), marital status (partnered or single), smoking (no or yes), metabolic equivalent of task (MET) score (MET hours per week were calculated to measure intensity and duration of physical activity [22]), and housing problems (the number of reported issues among 12 housing problems investigated in the ELSA, covering problems regarding space and lighting, noise, pollution, structure and maintenance, pest infestation, and temperature). Additionally, to depict the impacts of the socioeconomic development of England, the Socio-Demographic Development Index (SDI) was also used, which is based on 3 indicators associated with health outcomes: income per capita, educational attainment among individuals aged 15 years or younger, and total fertility rate among women aged under 25 years [23]. Finally, to account for effects of time and other unobserved factors, a year variable indicating studied year was also included. Detailed rationales and measurements for different variables are described in Section 2 in Multimedia Appendix 1.

Statistical Analysis

To obtain a nationally representative estimate, all statistics were weighted [24], and only representative individuals were included in the analyses.

Characteristics of participants were presented as means and SDs for continuous variables or frequencies and percentages for categorical variables by wave. Crude prevalences of frailty across waves from 2002 to 2022 were calculated as the proportion of participants with frailty. Nationally representative estimates were then obtained by weighting the crude prevalence to account for sampling probability and differential nonresponse. The prevalence of frailty as well as the mean value of the FI from 2002 to 2022 was estimated in the overall sample and subgroups stratified by age, sex, educational level, and marital status. Finally, the number of individuals affected by frailty among the population aged 50 years and older in England was estimated based on the population size of England each year estimated by the UK Office for National Statistics (ONS) [25].

To explore the temporal trend of frailty prevalence, joinpoint regression analysis, allowing for segmenting the temporal variations by using a piecewise regression model, was used. The annual percentage change (APC) between break points and average APC (AAPC) in the entire time horizon were then presented and regarded as statistically significant under the criteria of $P<.05$ [26]. The analysis was conducted in the overall sample and subgroups by age, sex, educational level, and marital status.

On the basis of a previous study [24], a linear probability model (LPM) was used to explore the driving factors associated with the change in frailty between 2002 and 2022 using cluster-robust SEs at the individual level to account for heteroscedasticity and within-individual correlation across repeated observations. The contribution of each explanatory variable, which was the proportion of the temporal changes in frailty prevalence that it explained, was denoted by the following equation [24]:

$$\beta j(\bar{x}j,2022-\bar{x}j,2002)/(\bar{y}2022-\bar{y}2002)$$

β_j denoted the coefficient of the explanatory variable estimated from the linear probability regression model, and $\bar{x}j2022$ and $\bar{x}j,2002$ were the wave averages of explanatory variables in 2022 and 2002, respectively, separately across individuals. $\bar{y}2022$ and $\bar{y}2002$ were the averages of frailty prevalence in 2022 and 2002, respectively, across individuals [24]. The future scenario-based prevalence of frailty in 2040 was projected using the LPM. Projected values of each explanatory variable and its corresponding model's coefficient were plugged into the model to project changes in frailty between 2002 and 2040. Specifically, the prevalence was estimated as the weighted sum of predicted probabilities using the mean or proportion of each explanatory variable in the future population. A summary of future covariate assumptions with sources can be found in Section 3 in Multimedia Appendix 1. Age and sex structures in 2040 in England were drawn from the UK ONS [27]. Educational attainment in 2040 was assumed to keep consistent with the situation in 2022 considering that the population of interest were aged 50 years or older, the youngest people in the target population would be aged over 30 years in 2022, and most people would have already completed their education. Assumed linear trends in marital status, MET, and housing problems were applied when forecasting corresponding values in 2040. Trends in the prevalence of smoking and SDI were obtained from the Global Burden of Disease study, and the annual change rate was applied to estimate the value of smoking prevalence and SDI in 2040. Calendar year was included as a continuous linear covariate and extended to the projection years assuming a linear continuation of the unobserved effects. This was justified by the broadly linear and nonsignificant overall trend observed between 2002 and 2022. Finally, the number of individuals affected by frailty among adults aged 50 years or older in England was calculated based on projected prevalence in 2040. To project the temporal trend, the frailty prevalence was also estimated in 2024, 2028, 2032, and 2036 using the same approach. Of note, all are scenario-based projections, not predictions.

Sensitivity analyses for testing the robustness of projection were conducted following probabilistic and scenario-based methodologies. First, a nonparametric bootstrap simulation was conducted with 1000 replications to obtain the 95% CI of projected frailty prevalence in 2040. The bootstrap 95% CI was interpreted as model-only uncertainty, conditional on the assumed future covariate values. Different values of explanatory variables with various assumptions were then imputed to obtain the projected prevalence of frailty in different scenarios. Assumptions included (1) projections of the age-sex structure of the population under different scenarios (low immigration and high immigration) [27]; (2) smoking prevalence increasing or decreasing by 20% relative to the original projected values; (3) 10% and 20% more individuals of an intermediate and higher educational level; and (4) use of the lower and upper 95% CI bounds of the likelihood of being single, MET hours per week, and number of housing problems. Additionally, because frailty was analyzed as a binary outcome, LPM may exhibit heteroscedasticity and generate fitted values outside the valid probability range. A logistic regression model using the same

XSL•FO
RenderX

covariates, sampling weights, and clustering specification as the primary analyses was fitted, and average marginal effects were estimated from the logistic regression model to facilitate comparison with the probability differences estimated by the LPM. To assess the potential overlap between the MET and the mobility-related items included in the FI, a sensitivity analysis in which the FI was reconstructed by excluding the mobility and physical performance deficits (items 23 to 29 in Section 1 in Multimedia Appendix 1) was performed. The LPM was then rerun using this modified FI, and the coefficient and contribution of the MET were compared with those from the primary analysis. Considering that the ELSA wave 10 coincided with the COVID-19 pandemic and the observed prevalence of frailty increased from 15.7% (1129/7195) in 2018 to 17.8% (1108/6226) in 2022, primary analyses were repeated after excluding data from wave 10. More details about the statistical analysis can be found in Section 2 in Multimedia Appendix 1. The joinpoint regression was conducted using the joinpoint trend analysis software developed by the National Cancer Institute (version 4.9.1.0). All other analyses were performed using Stata (version 18.0; StataCorp LLC). A 2-sided *P* value below .05 was considered significant for all statistical tests.

Ethical Considerations

Ethics approvals for the ELSA were obtained from different committees. Details can be found in Section 2 in Multimedia Appendix 1. All participants provided informed consent. All data used in this secondary analysis were deidentified and publicly available; thus, no additional institutional review was required. Projections based on publicly available ONS aggregate population data do not require additional ethical approval.

Results

After exclusions, a total of 85,190 records from 18,348 participants were included in the analysis, with a breakdown of the number of participants in each wave available in Section 4 in Multimedia Appendix 1. In the 2002 baseline survey, the mean age was 65.1 (SD 10.5) years, and 53.6% (6015/11,218) were female. A total of 59.5% (6678/11,218) had a lower educational level, and 31.4% (3522/11,218) were single. The characteristics of the participants across waves can be found in Section 5 in Multimedia Appendix 1. Of note, an improvement in educational attainment (proportion with a lower educational level from 6678/11,218, 59.5% to 1460/6226, 23.4%) was observed from 2002 to 2022. Section 6 in Multimedia Appendix 1 shows the characteristics of the included participants and those excluded for missing frailty data. Compared with included participants, excluded participants were older (mean age 70.9, SD 11.9 vs 61.9 years, SD 9.8) and more likely to report smoking (193/205, 94.1% vs 11,196/18,348, 61.0%) and had a lower mean MET score (4.8, SD 1.1 vs 5.8, SD 3.7).

Figure 1 shows the prevalence of frailty and the size of the frail population among community-dwelling adults aged 50 years and older in England from 2002 to 2022. The observed frailty prevalence in 2002 was 15.5% (1738/11,218) with a mean FI of 0.13 (SD 0.12). An increase of 2.3 percentage points in observed frailty prevalence was found in the previous 2 decades, reaching 17.8% (1108/6226) in 2022 with a corresponding mean FI of 0.14 (SD 0.13). The estimated number of individuals with frailty increased from 2,560,122 to 3,791,038 in 2022 among the community-dwelling population aged 50 years and older in England.

The prevalence of frailty across subgroups is shown in Figure 2. Those with advanced age, who were female, had a lower educational level, and/or were single had a consistently higher prevalence of frailty. In 2022, the prevalence of frailty was 42.6% (2652/6226) in the population aged 85 years and older compared to 12.9% (803/6226) in the age group of 50 to 64 years. The prevalence was 21.8% (1357/6226) among female individuals and was 13.5% (840/6226) among male individuals. Individuals with a lower educational level had a frailty prevalence of 32.1% (1998/6226), in contrast to 8.2% (510/6226) among those with higher education. Finally, frailty was more common among single individuals (1500/6226, 24.1%) than among those with partners (903/6226, 14.5%).

XSL•FO
RenderX

Figure 1. Prevalence of frailty and frail population from 2002 to 2022. FI: frailty index.

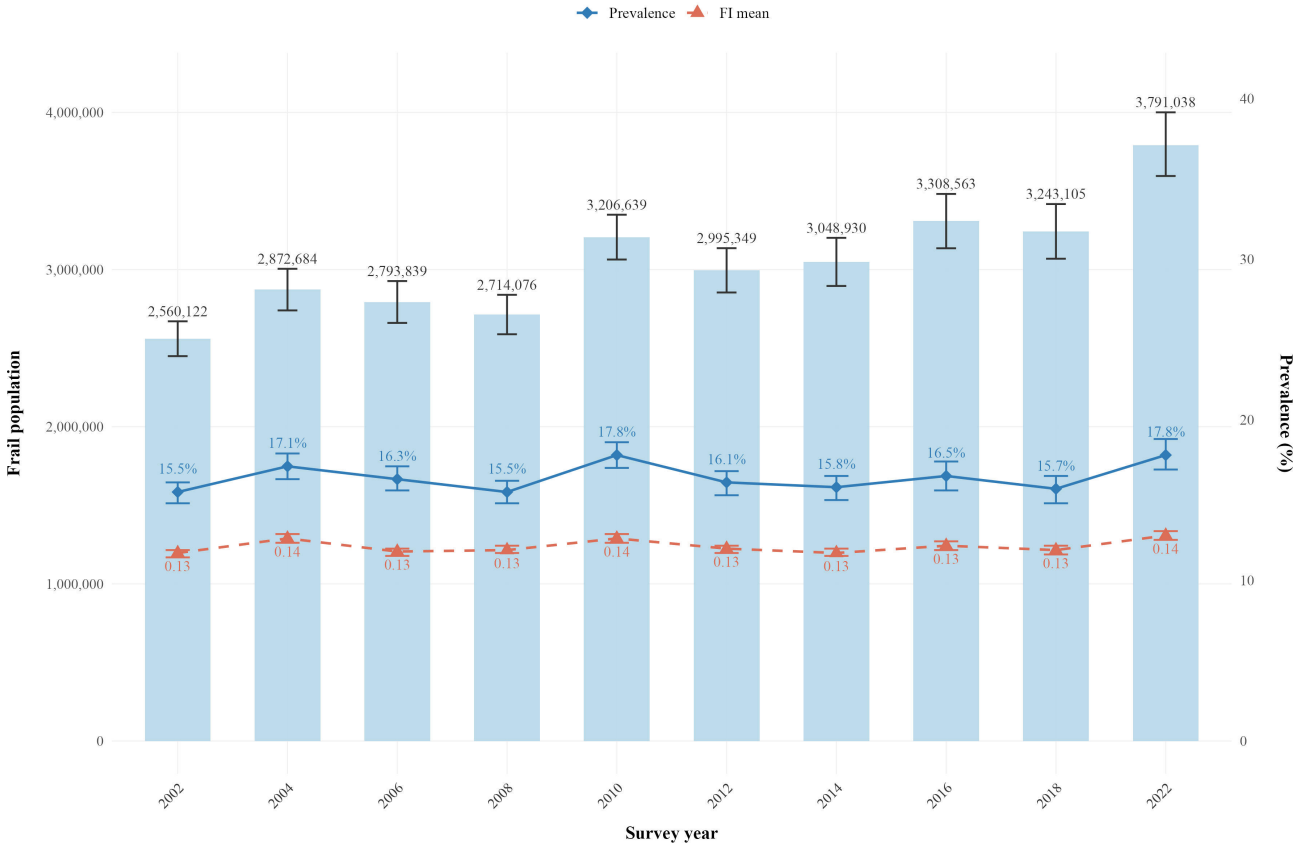


Figure 2. Prevalence of frailty by subgroup.

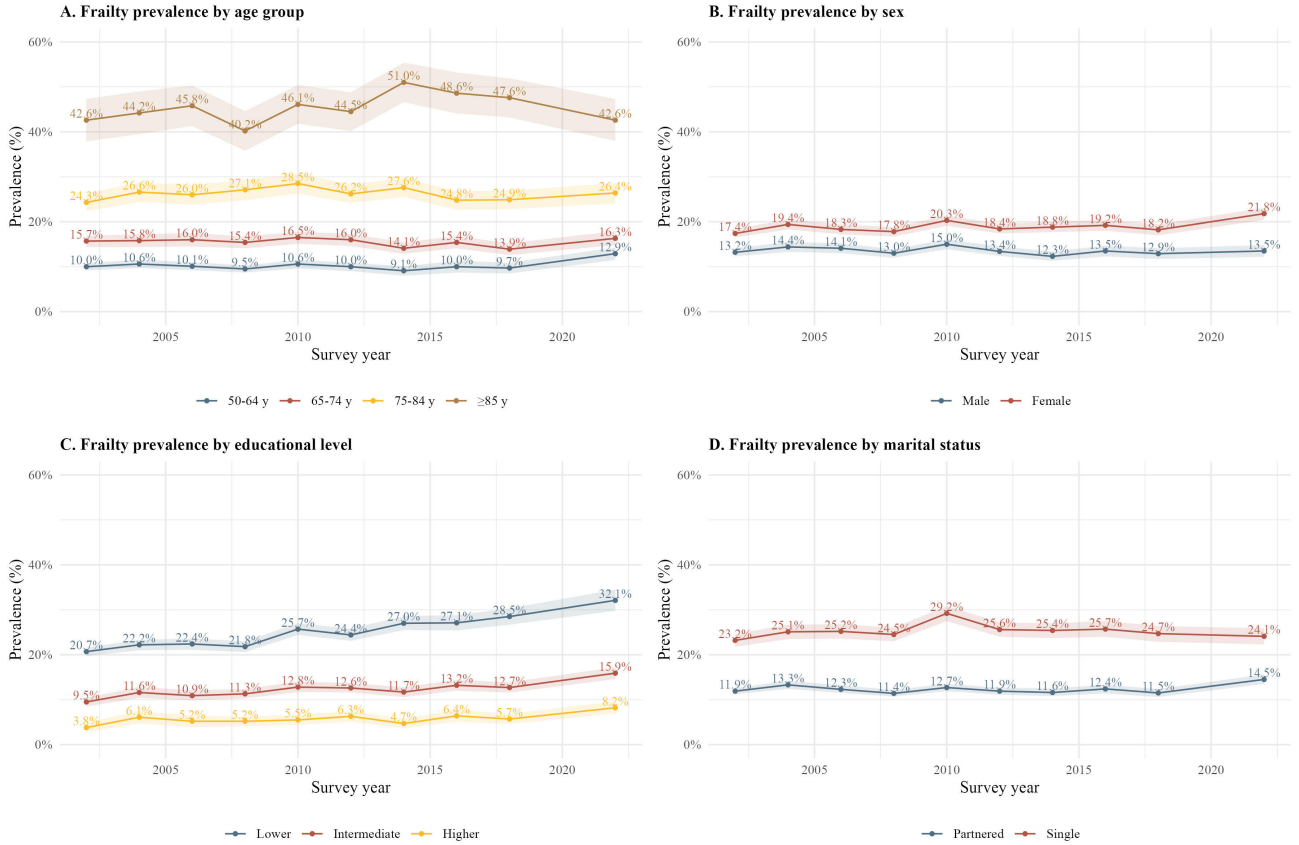


Table 1 summarizes the temporal trend of frailty prevalence. In the overall sample, the AAPC for frailty prevalence from 2002 to 2022 was 0.23% and not statistically significant, with no joinpoints identified. However, statistically significant upper trends were found among individuals with lower (AAPC=2.08%; *P*<.001) and intermediate (AAPC=1.91%; *P*=.03) educational

XSL•FO
RenderX

level. Among single individuals, a significant increasing trend (APC=1.92%; *P*<.01) in frailty prevalence was observed first from 2002 to 2010, followed by a significant decreasing trend (APC=–1.13%; *P*<.05) from 2010 to 2022. In the subgroup of individuals aged 85 years and older, frailty prevalence was found to increase at a relative rate of 1.11% per year from 2002 to 2016. No significant temporal trends were observed in the sex subgroup. The plots for visualization of different temporal trends can be found in Sections 7 to 18 in Multimedia Appendix 1.

The linear probability regression analysis showed that being older, being female, having a lower educational level, being single, having a lower MET score, smoking, and having more housing problems were positively associated with frailty from 2002 to 2022 (Section 19 in Multimedia Appendix 1). Several determinants associated with declining frailty showed improvements in England between 2002 and 2022. Educational attainment improved, with the proportion of individuals attaining an intermediate level increasing from 28.9% (3239/11,218) to 52.1% (3243/6226), while those with a higher educational level increased from 11.6% (1301/11,218) to 24.5% (1523/6226). The mean MET score increased from 5.37 (SD 3.60) to 5.67 (SD 3.74). Smoking prevalence declined, with the proportion of smokers falling from 64.4% (7222/11,218) to 58.4% (3637/6226). Finally, the mean SDI increased from 0.80 to 0.86. Meanwhile, some factors contributing to frailty showed increases to different extents from 2002 to 2022. A higher proportion of single individuals (3522/11,218, 31.4% vs 2149/6226, 34.5%) was observed in 2022. There was an increase of 1.2% in the proportion of the population aged 85 years and older (Section 20 in Multimedia Appendix 1).

Table . Temporal trend of frailty prevalence.

	Prevalence of frailty, n (%)		AAPC[a] (%; 95% CI)	*P* value	Joinpoint (prevalence in %)	APC[b] (%; 95% CI)			
	2002	2022				Slope 1	*P* value	Slope 2	*P* value
Overall	1738 (15.5)	1108 (17.8)	0.23 (–0.67 to 1.02)	.61	No joinpoint	—[c]	—	—	—
Age (y)									
50-64	1121 (10.0)	803 (12.9)	1.02 (–0.51 to 2.60)	.19	2016 (10.0)	–0.51 (–1.78 to 0.77)	.35	4.75 (–1.38 to 11.26)	.10
65-74	1761 (15.7)	1015 (16.3)	–0.26 (–0.95 to 0.43)	.41	No joinpoint	—	—	—	—
75-84	2726 (24.3)	1643 (26.4)	0.02 (–1.03 to 1.08)	.97	2016 (24.8)	1.25 (–1.34 to 3.93)	.27	–0.8 (–2.29 to 0.72)	.23
≥85	4779 (42.6)	2652 (42.6)	0.09 (–1.49 to 1.69)	.91	2016 (48.6)	1.11 (0.47 to 4.41)	.02	–2.25 (–4.63 to 0.05)	.06
Sex									
Male	1481 (13.2)	840 (13.5)	–0.25 (–1.18 to 0.54)	.44	No joinpoint	—	—	—	—
Female	1952 (17.4)	1357 (21.8)	0.6 (–0.11 to 1.28)	.10	No joinpoint	—	—	—	—
Educational level									
Lower	2322 (20.7)	1998 (32.1)	2.08 (1.49 to 2.56)	<.001	No joinpoint	—	—	—	—
Intermediate	1065 (9.5)	990 (15.9)	1.91 (0.22 to 3.53)	.03	No joinpoint	—	—	—	—
Higher	426 (3.8)	510 (8.2)	2.35 (–0.39 to 5.14)	.08	No joinpoint	—	—	—	—
Marital status									
Partnered	1335 (11.9)	903 (14.5)	0.27 (–1.08 to 1.39)	.66	No joinpoint	—	—	—	—
Single	2602 (23.2)	1500 (24.1)	0.08 (–0.59 to 0.64)	.87	2010 (29.2)	1.92 (0.60 to 5.04)	.003	–1.13 (–3.28 to –0.28)	.01

[a]AAPC: average annual percentage change.

[b]APC: annual percentage change.

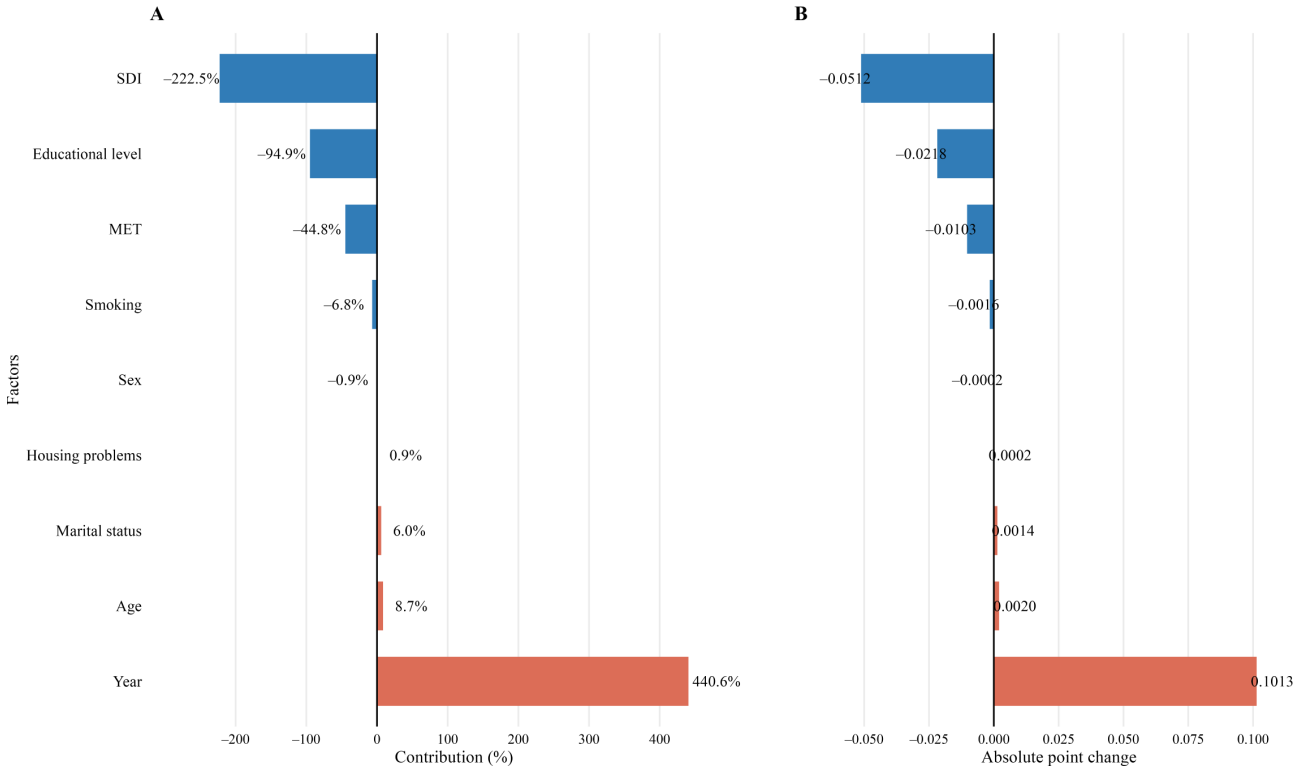
[c]Not applicable.

XSL•FO
RenderX

A total of 86.3% of the change in frailty prevalence was explained by this model. The overall change in frailty prevalence reflects the net result of opposing contributions from different factors, and a detailed contribution of each explanatory variable is shown in Figure 3. Aging population structure, a higher proportion of individuals being single, and more housing problems accounted for 8.7%, 6.0%, and 0.9% of the overall increase, respectively. Of note, year indicating other unobserved factors such as socioeconomic transitions, pandemic impacts, and hidden health exposures contributed to 440.6% of the increase. Conversely, improvements in sociodemographic conditions and lifestyle behaviors in England substantially mitigated the rise in frailty. The increase in SDI accounted for the largest reduction in frailty prevalence, offsetting up to 222.5% of the overall increase. Improvements in educational attainment explained 94.9% of the decrease in frailty, and more physical activity among the population aged 50 years and older contributed to a 44.8% reduction. Furthermore, a decline in smoking prevalence and fewer female individuals accounted for 6.8% and 0.9% of the decrease, respectively. Healthier lifestyle behaviors altogether offset 51.6% of the decrease. Detailed explained changes and corresponding proportions by subgroup can be found in Section 21 in Multimedia Appendix 1.

On the basis of the projected value of explanatory factors (Section 22 in Multimedia Appendix 1), the scenario-based projected prevalence of frailty among the English community-dwelling population aged 50 years and older was estimated as 20.5% in 2040, increasing from 17.1% in 2024. The corresponding projected number of individuals with frailty in England also increased from 3,829,786 in 2024 to 5,213,689 in 2040 based on extrapolating to the English population from ONS estimates on a national scale (Section 23 in Multimedia Appendix 1).

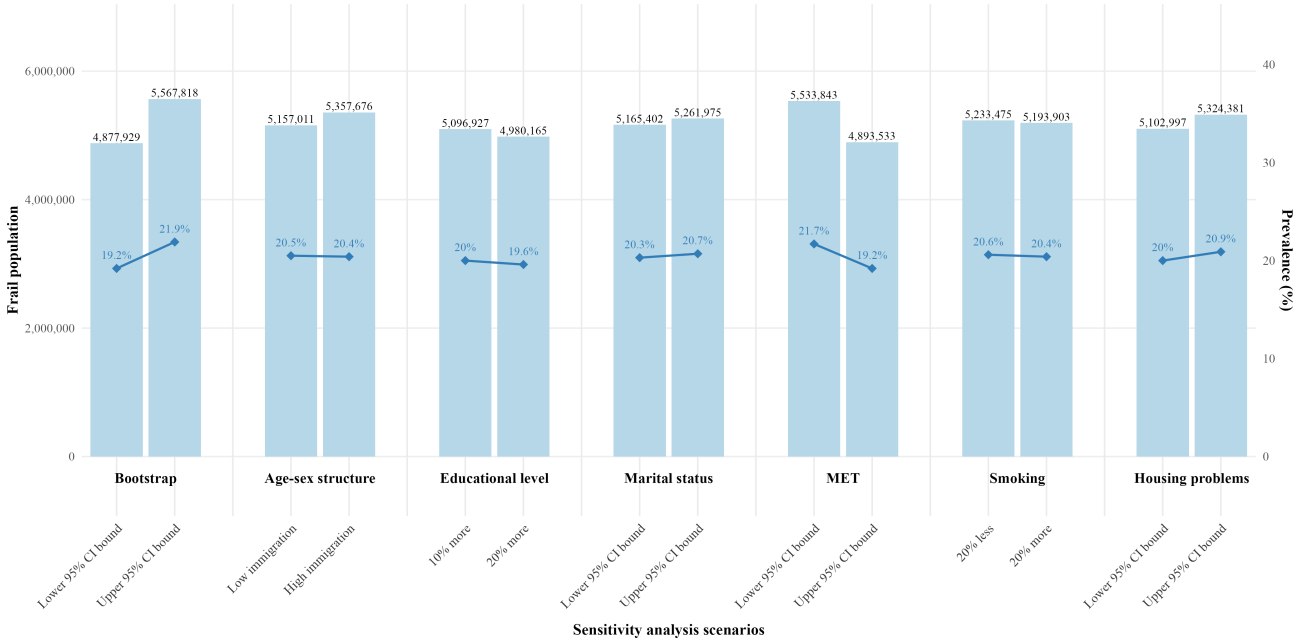
Figure 3. Contribution of different factors to changes in frailty prevalence in England between 2002 and 2022. Negative values indicate factors that offset the increase. Percentage contributions are relative to the net change and may exceed 100% or be negative; they do not add up to 100%. MET: metabolic equivalent of task; SDI: Socio-Demographic Development Index.



Sensitivity analyses for testing robustness of projection are shown in Figure 4. Bootstrap suggested a 95% CI of frailty prevalence in 2040 ranging from 19.2% to 21.9%, indicating a corresponding frail population size between 4,877,929 and 5,567,818. Immigration had limited effects on frailty prevalence, but the number of individuals with frailty could reach 5,357,676 under a scenario of high immigration. Assumptions of more educational improvement, fewer single individuals, more physical activity, and less smoking and housing problems were all associated with lower frailty prevalence and a lower frail population. Section 24 in Multimedia Appendix 1 shows the results estimated from the logistic regression model, showing that the direction and statistical significance of the associations were broadly consistent with those observed in the LPMs. Although some differences in the magnitude of the estimated effects were observed, particularly for the age group of 85 years and above (0.079 vs 0.176), the overall pattern of associations remained unchanged. In the sensitivity analysis excluding mobility and physical performance items from the FI, MET remained a statistically significant, independent protective factor (coefficient=–0.026; P<.001; Section 25 in Multimedia Appendix 1). Its contribution to offsetting the increase in frailty prevalence was 33.4% (Section 26 in Multimedia Appendix 1). In the sensitivity analysis excluding wave 10, the directions and statistical significance of all explanatory variables remained unchanged. The magnitudes of the LPM coefficients and factor contributions were also essentially unchanged (Sections 27 and 28 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 4. Sensitivity analyses of projected prevalence. The bootstrap 95% CI was interpreted as model-only uncertainty. MET: metabolic equivalent of task.



Discussion

This study found that the frailty prevalence in England among the community-dwelling population aged 50 years and older was broadly stable over 20 years, with a non–statistically significant increase from 15.5% (1738/11,218) in 2002 to 17.8% (1108/6226) in 2022, with the absolute number of individuals with frailty rising from 3,791,038 to 5,213,689. The aging population structure, a growing proportion of single individuals, and more housing problems contributed to frailty prevalence. However, these factors were largely offset by sociodemographic development in England, improvement in educational attainment, adoption of healthier lifestyles including more physical activity and less smoking, and a decreasing female proportion of the population, leading to the small overall change in frailty prevalence.

Over the past 2 decades, frailty prevalence in England has shown fluctuations but remained generally stable. This is consistent with the trends observed in healthy life expectancy within the same time frame [28]. An increase in frailty prevalence from 26.5% to 38.9% between 2006 and 2017 among the community-dwelling population aged 50 years and older was reported using primary care electronic health records [10]. This higher level and steeper rise could be attributed to several factors. First, the electronic FI was applied, with different items and thresholds from those used in the present study. Second, individuals in nursing homes were included, in contrast to the community-dwelling population considered in this study. Additionally, a tendency toward overselection of less healthy individuals inherent in retrospective electronic health record–based studies might have led to an overestimation of frailty prevalence [29]. In contrast, a modest increase from 13.5% to 16.3% over 9 years (2011-2020) was reported in China, while decreasing trends have been observed in high-income countries such as Finland and South Korea [30-32]. These findings provide evidence to guide public health approaches to prevention and for policymakers and health care providers to predict population health and care needs.

There is likely to be a substantial rise in care and health care demands because of the growing frail population size [5]. Resource input and innovation may be required in the future to support adults with frailty by integrating public health approaches with wider health, social, and long-term care systems [33]. Moreover, disparities in frailty prevalence and upward trends identified in specific subgroups highlighted the vulnerable subpopulations. On the basis of the results from this study, resources could be appropriately prioritized for older adults, female or single individuals, and those with lower educational attainment. Of note, this study excluded nursing home residents, among whom frailty is more common and more severe; our prevalence and projected absolute numbers are likely conservative estimates of the total population burden. Policy inference should therefore be interpreted as relevant to community-dwelling older adults, and separate estimates or projections for care home populations are needed for whole-population care planning.

This study provided a framework to explain changes in frailty prevalence and adopt targeted interventions accordingly from the perspectives of sociodemographic development, lifestyle behavior, environmental factors, and social development. Sociodemographic development and improvements in modifiable factors were key factors curbing the rapid increase in frailty prevalence. A higher SDI, reflecting broader social and economic development, is associated with better access to education in early life, improved living conditions, and stronger health care infrastructure—all of which contribute to more favorable health outcomes [34]. The benefits of education and healthier lifestyles were widely recognized by previous studies, suggesting the importance of shifting the care model further upstream by investing more in preventive measurements [35-38]. Among these, education appeared particularly influential in our study. This might reflect the cumulative advantages of higher

XSL•FO
RenderX

education, such as increased health literacy and better employment opportunities [38]. Therefore, a call on governments is needed to invest in early-life education and address the rising dropout rates in postsecondary enrollment to promote better long-term outcomes [38]. Although being single and housing problems accounted for a small proportion of the increase in prevalence, their significance in the regression model still offered insights for future interventions. Single individuals showed a higher risk of frailty, potentially due to increased psychological and emotional stress and reduced social support, highlighting the need for targeted interventions [39]. Additionally, housing is gradually gaining institutional recognition for its role as a social, economic, environmental, and cultural determinant of health, and its impacts on frailty have also been established [40,41]. Despite minor changes in housing problems during the study period, the cost of living crisis since 2020 might significantly worsen housing conditions and increase frailty risk, warranting closer attention [42].

This study has some limitations. First, the measurement of frailty relied on self-reported items in the ELSA, which might bring recall or response biases into the analysis. This was mitigated to some extent by the use of a standardized protocol for the ELSA, meaning that responses across waves should be reliable and comparable. Second, nursing home individuals were not included in this study; thus, the estimate of frailty prevalence should be generalized with caution to the full population. Third, although 86.3% of the changes in frailty prevalence were explained, the study year accounted for the largest proportion of the observed increase. However, as a residual proxy for unmeasured effects, study year may have masked the specific drivers of the rising frailty prevalence, such as socioeconomic transitions, pandemic impacts, and hidden health exposures. Consequently, key contributors to this trend may not have been fully identified. Additionally, to balance the simplicity and effectiveness of the projection model, and due to limitations in data availability or substantial missingness in self-completed questions in the ELSA, several variables known to influence frailty, such as nutrition, alcohol consumption, and social isolation, were not included in the analysis. Moreover, LPM may generate fitted values outside the valid probability range (0-1) and commonly exhibit heteroscedasticity. In our analysis, 20.7% of fitted values were below 0, although none exceeded 1. However, the use of cluster-robust SEs at the individual level partially addressed the issue of heteroscedasticity, and the sensitivity analysis using a logistic regression model confirmed the robustness of the results. Although analyses applied the ELSA survey weights to improve representativeness, weighting can only partially account for observed differences between eligible and noneligible respondents. Consequently, the estimates may not fully represent the entire community-dwelling population aged 50 years and older in England. Additionally, the bootstrap CI captures only model uncertainty, while scenario analyses vary assumptions individually and do not jointly account for uncertainty in future covariates, coefficients, baseline prevalence, and model structure. A fully joint uncertainty assessment is needed for future research. Finally, the mechanisms underlying the associations between these factors and frailty were not explored and may warrant future research to inform more targeted interventions. These limitations are balanced by key innovations, including 20 years of nationally representative data; a comprehensive set of explanatory variables; and an innovative analytical framework that combines temporal trend estimation, population-level decomposition, and scenario-based projection to 2040. Future studies are warranted to explore frailty prevalence in the entire population, incorporate objective frailty measurements, and identify unobserved driving factors for increasing frailty prevalence.

In conclusion, the past 20 years have seen a statistically nonsignificant increase in frailty prevalence among the community-dwelling population aged 50 years and older, and sociodemographic development, improvement in educational attainment, and adoption of healthier lifestyles served as the main curbing factors. By 2040, the scenario-based projected frailty prevalence is expected to increase modestly but with substantial growth in frail population size. Enhancing educational attainment and embedding healthy lifestyle behaviors across the population are suggested as key addressing strategies.

Generative AI Use

No generative AI tools were used in the preparation or revision of this manuscript.

Funding

This work was supported by the National Institute for Health and Care Research Barts Biomedical Research Centre (grant NIHR203330) and King’s College London–China Scholarship Council (grant 202306210057).

Data Availability

The English Longitudinal Study of Ageing is publicly available and free to access from the UK Data Service. Analysis code will be made available on request from the corresponding author.

Authors' Contributions

Conceptualization: DT (lead), KJS, SC, AG, JW, YW, MDLO (supporting)
Data curation: DT
Formal analysis: DT
Methodology: DT, MDLO, SC, KJS, YW

XSL•FO
RenderX

Supervision: MDLO
Validation: MDLO, AG, YW
Writing—original draft: DT
Writing—review and editing: DT, KJS, SC, AG, JW, YW, MDLO

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methodology details and results.
[PDF File, 2065 KB - aging_v9i1e98665_app1.pdf]

References

1. GBD 2021 Europe Life Expectancy Collaborators. Changing life expectancy in European countries 1990–2021: a subanalysis of causes and risk factors from the Global Burden of Disease Study 2021. Lancet Public Health 2025 Mar;10(3):e172-e188. [doi: 10.1016/S2468-2667(25)00009-X] [Medline: 39983748]
2. Stoodley I, Conroy S. An ageing population: the benefits and challenges. Medicine 2024 Nov;52(11):710-712. [doi: 10.1016/j.mpmed.2024.08.013]
3. Marteau TM, White M, Rutter H, et al. Increasing healthy life expectancy equitably in England by 5 years by 2035: could it be achieved? Lancet 2019 Jun 29;393(10191):2571-2573. [doi: 10.1016/S0140-6736(19)31510-7] [Medline: 31258113]
4. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
5. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
6. Cesari M, Prince M, Thiyagarajan JA, et al. Frailty: an emerging public health priority. J Am Med Dir Assoc 2016 Mar 1;17(3):188-192. [doi: 10.1016/j.jamda.2015.12.016] [Medline: 26805753]
7. Gale CR, Cooper C, Sayer AA. Prevalence of frailty and disability: findings from the English Longitudinal Study of Ageing. Age Ageing 2015 Jan;44(1):162-165. [doi: 10.1093/ageing/afu148] [Medline: 25313241]
8. Bai G, Wang Y, Mak JK, Ericsson M, Hägg S, Jylhävä J. Is frailty different in younger adults compared to old? Prevalence, characteristics, and risk factors of early-life and late-life frailty in samples from Sweden and UK. Gerontology 2023;69(12):1385-1393. [doi: 10.1159/000534131] [Medline: 37769628]
9. Kim J, Westra J, Tintle N, Harris WS, Park Y. Association of plasma n-3 polyunsaturated fatty acid levels and the prevalence of frailty in older adults: a cross-sectional analysis of UK Biobank. J Gerontol A Biol Sci Med Sci 2024 May 1;79(5):glae085. [doi: 10.1093/gerona/glae085] [Medline: 38518085]
10. Walsh B, Fogg C, Harris S, et al. Frailty transitions and prevalence in an ageing population: longitudinal analysis of primary care data from an open cohort of adults aged 50 and over in England, 2006-2017. Age Ageing 2023 May 1;52(5):afad058. [doi: 10.1093/ageing/afad058] [Medline: 37140052]
11. Welstead M, Jenkins ND, Russ TC, Luciano M, Muniz-Terrera G. A systematic review of frailty trajectories: their shape and influencing factors. Gerontologist 2021 Nov 15;61(8):e463-e475. [doi: 10.1093/geront/gnaa061] [Medline: 32485739]
12. Bouziri H, Descatha A, Roquelaure Y, Dab W, Jean K. Can we distinguish the roles of demographic and temporal changes in the incidence and prevalence of musculoskeletal disorders? A systematic review. Scand J Work Environ Health 2022 May 1;48(4):253-263. [doi: 10.5271/sjweh.4018] [Medline: 35260910]
13. Su D, Cui Y, He C, et al. Projections for prevalence of Parkinson's disease and its driving factors in 195 countries and territories to 2050: modelling study of Global Burden of Disease Study 2021. BMJ 2025 Mar 5;388:e080952. [doi: 10.1136/bmj-2024-080952] [Medline: 40044233]
14. Ni QY, Wu MY, Zha CK, et al. Trend and driving factors in burden of age-related macular degeneration in older adults aged 60-89 years: a global analysis over three decades. Age Ageing 2024 Sep 1;53(9):afae207. [doi: 10.1093/ageing/afae207] [Medline: 39324773]
15. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort profile: the English longitudinal study of ageing. Int J Epidemiol 2013 Dec;42(6):1640-1648. [doi: 10.1093/ije/dys168] [Medline: 23143611]
16. Dregan A, Armstrong D. Age, cohort and period effects in the prevalence of sleep disturbances among older people: the impact of economic downturn. Soc Sci Med 2009 Nov;69(10):1432-1438. [doi: 10.1016/j.socscimed.2009.08.041] [Medline: 19762138]
17. User guide to the main interview datasets waves 1 to 11. National Centre for Social Research. 2025. URL: https://11a183d6-a312-4f71-829a-79ff4e6fc618.filesusr.com/ugd/00de6d_96f357ea237748c297983250f5074790.pdf [accessed 2026-10-13]
18. He D, Wang Z, Li J, et al. Changes in frailty and incident cardiovascular disease in three prospective cohorts. Eur Heart J 2024 Mar 27;45(12):1058-1068. [doi: 10.1093/eurheartj/ehad885] [Medline: 38241094]

XSL•FO
RenderX

19. He D, Li J, Li Y, et al. Frailty is associated with the progression of prediabetes to diabetes and elevated risks of cardiovascular disease and all-cause mortality in individuals with prediabetes and diabetes: evidence from two prospective cohorts. Diabetes Res Clin Pract 2022 Dec;194:110145. [doi: 10.1016/j.diabres.2022.110145] [Medline: 36356844]
20. Cao X, Ma C, Zheng Z, et al. Contribution of life course circumstances to the acceleration of phenotypic and functional aging: a retrospective study. EClinicalMedicine 2022;51:101548. [doi: 10.1016/j.eclinm.2022.101548] [Medline: 35844770]
21. Song X, Mitnitski A, Rockwood K. Prevalence and 10-year outcomes of frailty in older adults in relation to deficit accumulation. J Am Geriatr Soc 2010 Apr;58(4):681-687. [doi: 10.1111/j.1532-5415.2010.02764.x] [Medline: 20345864]
22. Glasson L, Panter J, Ogilvie D, Patterson R. The physical activity implications of retirement across occupational activity groups. Prev Med 2023 Aug;173:107570. [doi: 10.1016/j.ypmed.2023.107570] [Medline: 37315902]
23. Wafa HA, Marshall I, Wolfe CD, et al. Burden of intracerebral haemorrhage in Europe: forecasting incidence and mortality between 2019 and 2050. Lancet Reg Health Eur 2024;38:100842. [doi: 10.1016/j.lanepe.2024.100842] [Medline: 38362494]
24. Gong J, Wang G, Wang Y, et al. Nowcasting and forecasting the care needs of the older population in China: analysis of data from the China Health and Retirement Longitudinal Study (CHARLS). Lancet Public Health 2022 Dec;7(12):e1005-e1013. [doi: 10.1016/S2468-2667(22)00203-1] [Medline: 36423656]
25. Estimates of the population for the UK, England, Wales, Scotland, and Northern Ireland. Office for National Statistics. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland [accessed 2026-10-13]
26. Meza R, Cao P, Jeon J, Warner KE, Levy DT. Trends in US adult smoking prevalence, 2011 to 2022. JAMA Health Forum 2023 Dec 1;4(12):e234213. [doi: 10.1001/jamahealthforum.2023.4213] [Medline: 38038988]
27. National population projections: 2022-based. Office for National Statistics. 2025. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/nationalpopulationprojections/2022based [accessed 2026-10-13]
28. Raleigh VS. Stalling life expectancy in the UK. BMJ 2018 Sep 27;362:k4050. [doi: 10.1136/bmj.k4050] [Medline: 30262572]
29. Williams BA. Constructing epidemiologic cohorts from electronic health record data. Int J Environ Res Public Health 2021 Dec 14;18(24):13193. [doi: 10.3390/ijerph182413193] [Medline: 34948800]
30. Tang D, Sheehan KJ, Goubar A, Whitney J, Dl O'Connell M. The temporal trend in frailty prevalence from 2011 to 2020 and disparities by equity factors among middle-aged and older people in China: a population-based study. Arch Gerontol Geriatr 2025 Jun;133:105822. [doi: 10.1016/j.archger.2025.105822] [Medline: 40068481]
31. Kang MG, Kim OS, Hoogendijk EO, Jung HW. Trends in frailty prevalence among older adults in Korea: a nationwide study from 2008 to 2020. J Korean Med Sci 2023 Jul 24;38(29):e157. [doi: 10.3346/jkms.2023.38.e157] [Medline: 37489714]
32. Pitkälä KH, Aalto UL, Roitto HM, et al. Secular trends in frailty among community-dwelling 75-95-year-old cohorts over three decades in Helsinki, Finland. Age Ageing 2024 Aug 6;53(8):afae172. [doi: 10.1093/ageing/afae172] [Medline: 39140371]
33. Dent E, Clegg A, Roller-Wirnsberger R, Vetrano DL, Hoogendijk EO. Reorienting frailty in clinical practice, public health, and policy: the Lancet Commission on Frailty. Lancet 2025 Jun 28;405(10497):2265-2266. [doi: 10.1016/S0140-6736(25)01101-8] [Medline: 40484017]
34. Lange S, Vollmer S. The effect of economic development on population health: a review of the empirical evidence. Br Med Bull 2017 Jan 1;121(1):47-60. [doi: 10.1093/bmb/ldw052] [Medline: 28069615]
35. Kojima G, Iliffe S, Walters K. Smoking as a predictor of frailty: a systematic review. BMC Geriatr 2015 Oct 22;15:131. [doi: 10.1186/s12877-015-0134-9] [Medline: 26489757]
36. Wang X, Hu J, Wu D. Risk factors for frailty in older adults. Medicine (Baltimore) 2022;101(34):e30169. [doi: 10.1097/MD.0000000000030169] [Medline: 36042657]
37. Ma LZ, Ge YJ, Zhang Y, et al. Identifying modifiable factors and their joint effect on frailty: a large population-based prospective cohort study. Geroscience 2025 Apr;47(2):2067-2078. [doi: 10.1007/s11357-024-01395-7] [Medline: 39441508]
38. Raghupathi V, Raghupathi W. The influence of education on health: an empirical assessment of OECD countries for the period 1995-2015. Arch Public Health 2020;78:20. [doi: 10.1186/s13690-020-00402-5] [Medline: 32280462]
39. Kojima G, Walters K, Iliffe S, Taniguchi Y, Tamiya N. Marital status and risk of physical frailty: a systematic review and meta-analysis. J Am Med Dir Assoc 2020 Mar;21(3):322-330. [doi: 10.1016/j.jamda.2019.09.017] [Medline: 31740150]
40. Mansour A, Bentley R, Baker E, et al. Housing and health: an updated glossary. J Epidemiol Community Health 2022 Sep;76(9):833-838. [doi: 10.1136/jech-2022-219085] [Medline: 35760516]
41. Gobbens RJ. Cross-sectional and longitudinal associations of environmental factors with frailty and disability in older people. Arch Gerontol Geriatr 2019;85:103901. [doi: 10.1016/j.archger.2019.103901] [Medline: 31352186]
42. Meadows J, Montano M, Alfar AJ, et al. The impact of the cost-of-living crisis on population health in the UK: rapid evidence review. BMC Public Health 2024 Feb 22;24(1):561. [doi: 10.1186/s12889-024-17940-0] [Medline: 38388342]

Abbreviations

AAPC: average annual percentage change
APC: annual percentage change

XSL•FO
RenderX

ELSA: English Longitudinal Study of Ageing
FI: frailty index
LPM: linear probability model
MET: metabolic equivalent of task
ONS: Office for National Statistics
SDI: Socio-Demographic Development Index

Edited by A Garcia; submitted 17.Apr.2026; peer-reviewed by B Cedeno-Veloz, L Ji; revised version received 21.Sep.2026; accepted 28.Sep.2026; published 03.Nov.2026.

Please cite as:
Tang D, Sheehan KJ, Conroy S, Goubar A, Whitney J, Wang Y, O'Connell MDL
Temporal Trends in the Prevalence of Frailty in England From 2002 to 2022, With Forecasts to 2040: Longitudinal Secondary Analysis of a National Cohort Study
JMIR Aging 2026;9:e98665
URL: https://aging.jmir.org/2026/1/e98665
doi:10.2196/98665

© Dongfeng Tang, Katie Jane Sheehan, Simon Conroy, Aicha Goubar, Julie Whitney, Yanzhong Wang, Matthew D L O'Connell. Originally published in JMIR Aging (https://aging.jmir.org), 3.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predicting Frailty Trajectories Using Interpretable Machine Learning Among Older Adults Following Hip Surgery: Prospective Longitudinal Study

Jingying Huang[1], MSN; Qinqin Fan[2], PhD; Huiqin Shi[3], BSN; Jiaojiao Chen[1], MSN; Hongbo Gao[2], BSc; Kailai Gu[3], BSN; Jin Yang[1], MSN

[1]Sir Run Run Shaw Hospital, No. 3, Qingchun East Road, Hangzhou, Zhejiang, China

[2]Shanghai Maritime University, Shanghai, China

[3]Zhejiang Chinese Medical University, Hangzhou, China

Corresponding Author:
Jingying Huang, MSN
Sir Run Run Shaw Hospital, No. 3, Qingchun East Road, Hangzhou, Zhejiang, China

Abstract

Background: Postoperative frailty is highly prevalent among older adults undergoing hip surgery and is closely linked to poor clinical outcomes. Despite growing interest in understanding its progression, the temporal patterns of frailty remain underexplored. Moreover, there is a lack of validated models that can predict frailty trajectories and stratify patients by risk in the early postoperative period.

Objective: This study aimed to identify distinct frailty trajectories within 6 months following hip surgery in older adults and to explore their associated predictors. An interpretable machine-learning model was developed and internally validated for individualized risk prediction and was implemented as a clinically accessible web-based calculator.

Methods: This prospective longitudinal observational study was conducted among older adults undergoing hip surgery at a tertiary hospital in China. Frailty assessments were performed preoperatively and at 1, 3, and 6 months postoperatively. A total of 209 participants who completed the 6-month follow-up were included in the analysis. Frailty was assessed using the Frailty Index, and group-based trajectory modeling was applied to identify distinct frailty progression patterns. Predictive variables were selected using the least absolute shrinkage and selection operator regression. An interpretable Extreme Gradient Boosting (XGBoost) model was developed using a 60:40 training-test data split. Model performance was evaluated in terms of discrimination, calibration, and clinical utility. Interpretability was assessed using SHAP (Shapley Additive Explanations) at both the global and individual levels.

Results: Three distinct frailty trajectories were identified: low-fluctuation frailty (55/209, 26%), high-improvement frailty (81/209, 39%), and high-deterioration frailty (73/209, 35%). Twelve predictors grounded in the Health Ecology Model were selected, spanning individual characteristics, interpersonal networks, and the living environment. The XGBoost model demonstrated excellent discrimination, with a microaverage area under the receiver operating characteristic curve of 0.98 (95% CI 0.96 - 0.99) in the training set and 0.93 (95% CI 0.90 - 0.96) in the test set. Calibration was acceptable, with a weighted Brier score of 0.0852. Decision curve analysis showed favorable clinical utility across a range of threshold probabilities. A web-based risk calculator was developed to facilitate personalized frailty trajectory prediction.

Conclusions: The XGBoost model demonstrated strong predictive performance and interpretability, enabling the early identification of older patients at risk for adverse frailty trajectories following hip surgery. This tool may support targeted interventions and improve perioperative care in geriatric populations.

(JMIR Aging 2026;9:e90705) doi:10.2196/90705

KEYWORDS

frailty trajectory; hip surgery; predictive model; Extreme Gradient Boosting; XGBoost; Shapley Additive Explanations; SHAP analysis; frailty; aging; older adults

Introduction

Population aging has emerged as a significant global public health challenge. With shifting demographic structures, the incidence of orthopedic conditions among older adults has risen substantially. In 2000, there were an estimated 1.6 million hip fracture cases worldwide, a number projected to increase to 6.3 million by 2050 [1]. Hip osteoarthritis is also common in older

XSL•FO
RenderX

adults, with an incidence of up to 10%, and its end-stage form leads to significant pain and disability, severely impacting quality of life [1]. For conditions such as hip fractures and advanced hip osteoarthritis, hip surgery remains the primary treatment, effectively alleviating pain, restoring joint function, and improving patient outcomes [2]. Notably, approximately 85% of patients undergoing hip surgery are aged 60 years or older and often present with multimorbidity and frailty [3].

Frailty is an age-related clinical syndrome characterized by a progressive decline in physiological function and multisystem reserve, which impairs the ability to withstand stressors, such as illness, trauma, or surgery, and increases vulnerability to adverse outcomes [4,5]. A growing body of evidence has identified frailty as an independent predictor of poor postoperative outcomes, including falls, functional decline, rehospitalization, and mortality [6-8]. Among older patients undergoing surgery, the prevalence of frailty ranges from 41.8% to 50.3% and may reach up to 65% in those undergoing hip surgery [9,10]. This elevated prevalence reflects the dual impact of both the underlying health status and surgical stress. On the one hand, hip fracture itself is often associated with frailty-related conditions such as sarcopenia, impaired balance, and malnutrition. On the other hand, perioperative factors such as prolonged bed rest, pain, and reduced mobility can accelerate muscle loss and disrupt homeostasis, further depleting physiological reserves through endocrine, metabolic, and immune dysregulation [11,12]. Common postoperative complications, including delirium and infection, may further delay recovery and exacerbate functional deterioration. Even 1 year after surgery, frailty remains present in 25.6% to 27.7% of patients [13].

Importantly, frailty is increasingly recognized as a dynamic and potentially reversible condition. Its core principle emphasizes that early identification and proactive intervention can delay or even reverse its progression [14]. The distinct phases of postoperative recovery, rehabilitation, and stabilization provide a critical window to observe changes in frailty status. These trajectories are shaped by factors such as baseline reserve, perioperative complications, and psychosocial conditions. While some patients recover steadily, others experience worsening frailty and increased dependency [15]. Studies suggest that frailty trajectory patterns better predict outcomes than static measurements, with rapid deterioration or high variability indicating increased risk [16]. These findings underscore the need for tools that can identify frailty trajectories early and guide intervention timing. While previous studies have examined frailty trajectories in various disease contexts, only a few have developed predictive models [17,18]. These models are typically restricted to specific populations, rely on a narrow set of variables, and employ conventional statistical approaches, limiting their robustness and generalizability. Furthermore, static frailty assessment tools, such as the Clinical Frailty Scale and the FRAIL Scale, fail to capture the multidimensional, nonlinear nature of frailty progression. Therefore, there is a pressing need for predictive models that integrate multilevel risk factors and offer strong interpretability.

The Health Ecology Model (HEM) offers a comprehensive framework for viewing health as the outcome of interactions among individual, behavioral, and environmental factors [19]. It emphasizes that health is determined by personal characteristics (eg, age, sex, body weight, psychological status, and lifestyle), as well as by social support, living conditions, economic status, and broader policy environments. This perspective aligns with the multifactorial etiology of frailty, which is shaped by both intrinsic and extrinsic factors [18,20]. Methodologically, Extreme Gradient Boosting (XGBoost) is a robust ensemble learning algorithm well-suited for modeling high-dimensional, nonlinear data and performs reliably even with relatively small sample sizes. SHAP (Shapley Additive Explanations), grounded in cooperative game theory, enhances the interpretability of machine-learning models by quantifying the global and individual-level contributions of each feature to prediction outcomes [21].

Therefore, guided by the HEM, this study integrates predictors across individual, behavioral, and environmental domains to construct an interpretable XGBoost-based predictive model for frailty trajectories within 6 months following hip surgery in older patients. SHAP is used to visualize and interpret the contribution of key predictors, and a web-based calculator is developed for clinical applications. This study aims to support the stratification of patients into distinct frailty trajectory subgroups, enabling more precise screening, targeted intervention, and optimized allocation of health care resources.

Methods

Study Design

This was a prospective, longitudinal cohort study, reported in accordance with the 2024 TRIPOD + AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis—Artificial Intelligence) guidelines for the development and validation of multivariable prediction models [22].

Population

This study employed convenience sampling to recruit older adults who underwent hip surgery under general anesthesia at a tertiary hospital in China between March 2024 and March 2025. Eligible patients were followed prospectively for 6 months postoperatively, with follow-up assessments conducted from March 2024 to September 2025.

Inclusion criteria were age 60 years or older; scheduled for elective hip surgery, including total hip arthroplasty, hemiarthroplasty, or internal fixation; and sufficient communication ability to complete study assessments. Exclusion criteria were severe psychiatric disorders that prevented adequate cooperation; substantial hearing or visual impairments that precluded completion of questionnaires; palliative status (ie, life expectancy <6 mo); or current enrollment in other studies likely to affect frailty outcomes. Participants were withdrawn if they were lost to follow-up due to relocation or hospital transfer, voluntarily withdrew consent, or died of causes unrelated to the study protocol.

Baseline data and preoperative frailty assessments were collected through face-to-face interviews conducted by trained researchers. Follow-up assessments at 1, 3, and 6 months

XSL•FO
RenderX

postoperatively were conducted via structured telephone interviews.

Sample Size

As there is currently no universally accepted sample size calculation method for trajectory-based multiclass machine-learning prediction models, the sample size was pragmatically estimated using a conventional epidemiological formula to ensure adequate baseline cohort representativeness. The required sample size was calculated as follows [23]: n=(Zδ)2φ^(1−φ^), where *Z* is the critical value at a 95% CI, φ∧ is the estimated incidence of frailty, and δ is the acceptable margin of error. Based on a pilot study indicating a 6-month postoperative frailty incidence of 14% and setting *d* at 5%, the calculated sample size was 185. After accounting for a 10% attrition rate, the final target sample size was set at no fewer than 206 cases to ensure the epidemiological representativeness of the baseline cohort.

Ethical Considerations

Ethical approval for the study was obtained from the Ethics Committee of Sir Run Run Shaw Hospital, Zhejiang University School of Medicine (number 2024-2687-01). Informed consent was obtained from all participants before the study commenced.

Outcomes Definition

Frailty was assessed preoperatively and at 1, 3, and 6 months postoperatively using a modified Frailty Index (FI) based on the cumulative deficit model developed by Rockwood and Mitnitski [24]. Tailored to the clinical characteristics of older adults undergoing hip surgery, we developed a 37-item population-specific FI covering 6 domains: chronic diseases, mental health, activities of daily living (ADLs), cognitive function, and symptoms/signs. Although several predictive variables share conceptual domains with components of the FI, none of the specific indicators used to construct the FI were directly included as predictors in the model. FI scores ranged from 0 to 1, with higher scores indicating greater frailty. A score <0.25 was classified as nonfrail and a score ≥0.25 was classified as frail [25]. The full list of FI items is provided in Multimedia Appendix 1.

Candidate Predictor Variables

Based on a review of the literature and expert consultation, potential risk factors were explored within the HEM, encompassing individual characteristics, behavioral and psychological factors, interpersonal networks, and the living environment. Key individual characteristics included demographic and socioeconomic factors (eg, gender, marital status, education level, and income), physiological indicators (high-sensitivity C-reactive protein [hsCRP]), health behaviors (smoking and drinking), and clinical history (number of comorbidities and surgery type). Functional status was assessed using the Barthel Index for ADLs, with scores indicating varying levels of independence [26]. Nutritional status was evaluated via the Mini Nutritional Assessment–Short Form [27], with scores categorizing participants as malnourished, at risk, or well-nourished.

Behavioral and psychological factors, including depressive symptoms, were measured using the 5-item Geriatric Depression Scale, while sleep quality was assessed using the Pittsburgh Sleep Quality Index. For interpersonal networks, family function was assessed using the Family APGAR (Adaptability, Partnership, Growth, Affection, Resolve) Index [28], and social support was measured using the Social Support Rating Scale. Finally, environmental factors were evaluated using a self-developed Living Environment Questionnaire (LEQ), which assesses home adaptability, perceived environmental burden, community resource accessibility, and residential satisfaction. The LEQ demonstrated strong content validity and was tailored to the needs of older adults. Detailed descriptions of these variables and the respective scales are provided in Multimedia Appendix 1.

Statistical Analysis

A complete-case analysis was conducted. Missing data primarily arose from loss to follow-up during the longitudinal assessments, and participants with incomplete follow-up data were excluded from the analysis. To assess the missing data mechanism, the Little test for missing completely at random (MCAR) was performed. In addition, baseline characteristics between participants included in the final analysis and those who were lost to follow-up were compared to evaluate the potential for selection bias. Categorical variables are expressed as frequencies and percentages. Differences between the training and test sets were assessed using the chi-square test. Continuous variables with a normal distribution are presented as mean (SD), while nonnormally distributed variables are shown as median (IQR). Depending on the data distribution, the *t* test or the Mann-Whitney *U* test was applied.

Group-based trajectory modeling was utilized to identify frailty trajectories in older patients with hip fractures. The optimal model was selected based on clinical relevance, fit indices, and classification quality, with a preference for simpler models with fewer groups. Model fit was evaluated using the following criteria [29]: (1) Bayesian Information Criterion (BIC): lower values indicate a better fit; (2) average posterior probability (AvePP): AvePP greater than 0.7 for all groups, with the smallest group proportion exceeding 5%; (3) odds of correct classification (OCC): OCC was greater than 5; (4) relative entropy (E_k): E_k was higher than 0.8.

The data were randomly split into training and test sets in a 60:40 ratio using stratified sampling to preserve the original distribution of the 3 frailty trajectory classes across the dataset. To prevent overfitting, LASSO (least absolute shrinkage and selection operator) regression was applied to the training set for high-dimensional feature selection. The L1 regularization term enabled simultaneous variable selection and coefficient shrinkage. Model performance was evaluated using 10-fold cross-validation, with the optimal penalty parameter (λ) selected according to the 1 SE rule to ensure model parsimony and generalizability. To evaluate potential overlap and multicollinearity among predictors, variance inflation factors (VIFs) were calculated for variables retained after LASSO selection. A VIF <5 was considered to indicate no significant multicollinearity [30].

XSL•FO
RenderX

The selected features were used to construct the XGBoost model. Because no trajectory class represented an extremely small proportion of the sample, no additional class imbalance handling strategies (eg, oversampling or SMOTE [Synthetic Minority Oversampling Technique]) were applied during model training. To improve model robustness and reduce overfitting, regularization and subsampling strategies were incorporated into the XGBoost model. During training, a 3-fold inner cross-validation was employed to estimate performance. For the test set, the area under the receiver operating characteristic curve, accuracy, F_1-score, sensitivity, and specificity were computed. The calibration curve was used to compare predicted probabilities with actual outcomes across frailty trajectories. Decision curve analysis was performed to assess net clinical benefit. Finally, SHAP values were used to interpret the XGBoost model, ranking feature importance based on the mean absolute SHAP values. The XGBoost model was deployed on a cloud platform [31] using the "Shiny" package, creating an online prediction calculator for frailty trajectories following hip surgery.

Statistical analysis and graphical representations were performed using IBM SPSS 26.0, R 4.5.1 (R Core Team), and Python 3.9 (Python Software Foundation), with the "traj," "glmnet," "xgboost," and "shap" packages. All statistical tests were 2-sided, and statistical significance was set at α=.05.

Results

Characteristics of Participants

A total of 300 patients were screened for the study, of whom 78 were excluded due to severe hearing impairment, significant visual decline, communication barriers, or other factors. Four patients with incomplete data were also excluded, leaving 218 patients enrolled. Multimedia Appendix 2 presents the flowchart of the participants included in this study. During the follow-up period, 9 participants were lost to follow-up at 6 months, yielding an attrition rate of 4.1%. Given that frailty trajectories were identified using group-based trajectory modeling based on 3 repeated follow-up assessments, participants with incomplete longitudinal data were excluded to ensure the stability and interpretability of trajectory classification. To evaluate the missing data mechanism, we conducted a missing data analysis (Multimedia Appendix 2). The Little MCAR test indicated that the missing data were consistent with the assumption of MCAR (χ^2_4=3.750, P=.44). In addition, no significant differences were observed in baseline characteristics between participants retained in the analysis and those lost to follow-up (Multimedia Appendix 2), suggesting a low risk of selection bias.

Frailty Prevalence and Trajectories

Frailty levels among older patients undergoing hip surgery exhibited a dynamic trend, with an initial increase followed by a gradual decline. The prevalence of frailty was 70.81% (148/209) before surgery, peaked at 84.21% (176/209) at 1 month postoperatively, and subsequently decreased to 63.16% (132/209) at 3 months and 45.45% (95/209) at 6 months.

A 3-class trajectory model demonstrated the best fit based on BIC, AvePP, and E_k values (Multimedia Appendix 3). It showed the lowest BIC change (59.36). All classes had AvePP values greater than 0.7 and OCC values exceeding 5. The E_k was 0.905, indicating high classification precision. Therefore, the 3-trajectory model was selected to characterize postoperative frailty trajectories among older adults undergoing hip surgery. Based on the estimated intercepts and slopes, trajectory 0 was labeled as low-fluctuation frailty, comprising 55 out of 209 individuals (26%), with a lower baseline frailty level, a modest increase at 1 month, and a subsequent decline. Trajectory 1 was labeled as high-improvement frailty, including 81 out of 209 (39%) individuals who had a higher baseline frailty level and experienced continuous functional recovery. Trajectory 2 was labeled as high-deterioration frailty, consisting of 73 out of 209 (35%) individuals who started with a high frailty level and demonstrated progressive worsening over time. Detailed parameter estimates and visual representations are provided in Multimedia Appendix 3. Given the distinct clinical patterns observed among the 3 classes, a multiclass modeling approach was employed to support the subsequent predictive analysis.

Predictor Variable Selection

A total of 209 participants were randomly assigned to a training set (n=126) and a testing set (n=83) in a 60:40 ratio. The training set was used for model development, while the testing set was reserved for independent performance evaluation. Baseline demographic and clinical characteristics were compared between the 2 sets, with statistically significant differences observed only in marital status and type of medical insurance (Table 1).

XSL•FO
RenderX

Table . Comparison of baseline characteristics between the training and test sets.

Variables	Total (N=209)	Training set (n=126)	Test set (n=83)	χ^2 (*df*) or *Z* value	*P* value
Age (y), median (IQR)	74 (68-83)	75 (68-83)	74 (68-82)	−3.343[a]	.73
Gender, n (%)				0.720 (1)[b]	.40
Male	66 (31.6)	37 (29.4)	29 (34.9)		
Female	143 (68.4)	89 (70.6)	54 (65.1)		
Marital status, n (%)				6.326 (1)[b]	.01
With spouse	58 (27.8)	27 (21.4)	31 (37.3)		
Without spouse	151 (72.2)	99 (78.6)	52 (62.7)		
Education level, n (%)				3.023 (3)[b]	.39
Primary or below	135 (64.6)	79 (62.7)	56 (67.5)		
Middle school	47 (22.5)	33 (26.2)	14 (16.9)		
High school	20 (9.6)	10 (7.9)	10 (12)		
College or above	7 (3.3)	4 (3.2)	3 (3.6)		
Type of work, n (%)				2.574 (2)[b]	.28
Manual labor	125 (59.8)	71 (56.3)	54 (65.1)		
Mental labor	64 (30.6)	40 (31.7)	24 (28.9)		
Unemployed or other	20 (9.6)	15 (11.9)	5 (6.0)		
Residence place, n (%)				1.669 (1)[b]	.20
Rural	117 (56.0)	66 (52.4)	51 (61.4)		
Urban	92 (44.0)	60 (47.6)	32 (38.6)		
Live arrangement, n (%)				1.752 (1)[b]	.19
Alone	41 (19.6)	21 (16.7)	20 (24.1)		
With family members	168 (80.4)	105 (83.3)	63 (75.9)		
Monthly income (￥), n (%)				0.182 (2)[b]	.91
<1000	68 (32.5)	42 (33.3)	26 (31.3)		
1000~5000	107 (51.2)	63 (50.0)	44 (53.0)		
>5000	34 (16.3)	21 (16.7)	13 (15.7)		
Medical insurance, n (%)				3.873 (1)[b]	.049
Yes	10 (4.8)	6 (4.3)	4 (1.2)		
No	199 (95.2)	117 (92.9)	82 (98.8)		
BMI (kg/m^2), n (%)				0.515 (2)[b]	.77
<18.5	24 (11.5)	15 (11.9)	9 (10.8)		
18.5~24	113 (54.1)	70 (55.6)	43 (51.8)		
>24	72 (34.4)	41 (32.5)	31 (37.3)		
Albumin (g/L), n (%)				0.430 (1)[b]	.51
<40	138 (66.0)	81 (64.3)	57 (68.7)		
≥40	71 (34.0)	45 (35.7)	26 (31.3)		
hsCRP[c] (mg/L), n (%)				0.160 (1)[b]	.69

XSL•FO
RenderX

Variables	Total (N=209)	Training set (n=126)	Test set (n=83)	χ^2 (*df*) or *Z* value	*P* value
≥6	130 (62.2)	77 (61.1)	53 (63.9)		
<6	79 (37.8)	49 (38.9)	30 (36.1)		
Smoking status, n (%)				4.117 (2)[b]	.13
Current smoker	14 (6.7)	5 (4.0)	9 (10.8)		
Never smoker	172 (82.3)	108 (85.7)	64 (77.1)		
Former smoker	23 (11.0)	13 (10.3)	10 (12.0)		
Drinking status, n (%)				2.194 (2)[b]	.33
Current drinking	36 (17.2)	18 (14.3)	18 (21.7)		
Never drinking	160 (76.6)	99 (78.6)	61 (73.5)		
Former drinking	13 (6.2)	9 (7.1)	4 (4.8)		
Number of teeth, median (IQR)	14 (8-18)	14 (8-18)	14 (8-18)	−0.079[a]	.94
Number of comorbidities, n (%)				1.352 (1)[b]	.25
≥2	101 (48.3)	65 (51.6)	36 (43.4)		
<2	108 (51.7)	61 (48.4)	47 (56.6)		
Surgery type, n (%)				0.947 (2)[b]	.62
Hip internal fixation	58 (27.8)	32 (25.4)	26 (31.1)		
Total hip arthroplasty	88 (42.1)	54 (42.9)	34 (41.0)		
Hemiarthroplasty	63 (30.1)	40 (31.7)	23 (27.7)		
Social activity frequency, n (%)				2.046 (1)[b]	.15
≥3 times per week	147 (70.3)	84 (66.7)	63 (75.9)		
<3 times per week	62 (29.7)	42 (33.3)	20 (24.1)		
Activities of daily living, n (%)				3.997 (3)[b]	.26
Complete dependence	6 (2.9)	5 (4.0)	1 (1.2)		
Severe dependence	46 (22.0)	25 (19.8)	21 (25.3)		
Moderate dependence	120 (57.4)	77 (61.1)	43 (51.8)		
Mild dependence	37 (17.7)	19 (15.1)	18 (21.7)		
Nutritional status, n (%)				1.565 (2)[b]	.46
Malnourished	36 (17.2)	25 (19.8)	11 (13.3)		
At risk	64 (30.6)	38 (30.2)	26 (31.3)		
Normal	109 (52.2)	63 (50.0)	46 (55.4)		
Depressive, n (%)				0.512 (1)[b]	.47
Yes	19 (9.1)	10 (7.9)	9 (10.8)		
No	190 (90.9)	116 (92.1)	74 (89.2)		
Sleep quality, n (%)				0.013 (1)[b]	.91
Poor	74 (35.4)	45 (35.7)	29 (34.9)		
Good	135 (64.6)	81 (64.3)	54 (65.1)		

XSL•FO
RenderX

Variables	Total (N=209)	Training set (n=126)	Test set (n=83)	χ^2 (*df*) or *Z* value	*P* value
Family function, n (%)				2.221 (2)[b]	.33
Severe dysfunction	62 (29.7)	40 (31.7)	22 (26.5)		
Moderate dysfunction	74 (35.4)	47 (37.3)	27 (32.5)		
Good family functioning	73 (34.9)	39 (31.0)	34 (41.0)		
Social support level, median (IQR)	30 (27-34)	30 (27-33)	30 (27-34)	−0.218[a]	.83
Living environment score, median (IQR)	12 (11-14)	12 (11-14)	12 (11-14)	−0.243[a]	.81

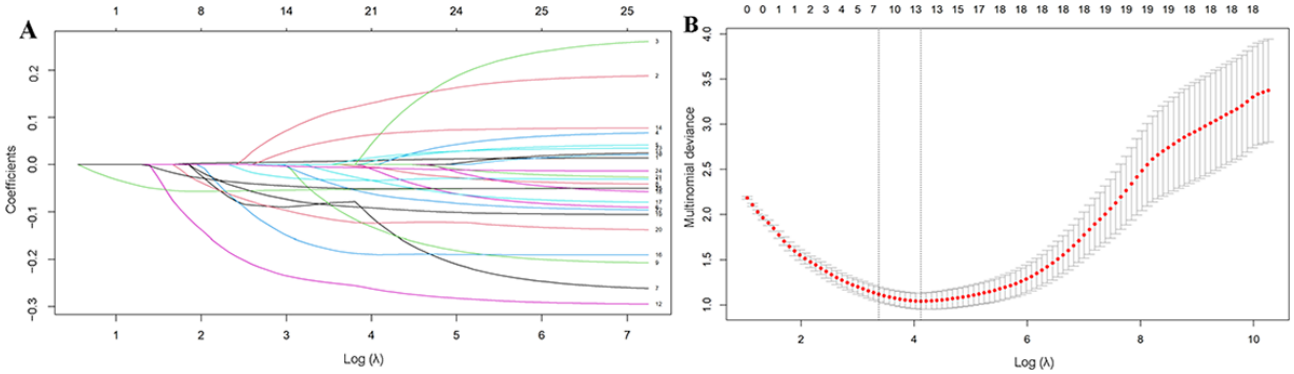
[a]Mann-Whitney *U* test.

[b]Chi-square test.

[c]hsCRP: high-sensitivity C-reactive protein.

Twenty-five candidate predictors were entered into the LASSO regression for feature selection. Using the 1 SE criterion (λ=0.0342), 12 variables were retained, including age, gender, living arrangement, albumin, hsCRP, drinking status, teeth number, comorbidities number, ADLs, nutritional status, social support level, and living environment. The feature selection process is illustrated in Figure 1, with additional details provided in Multimedia Appendix 4. Multicollinearity diagnostics showed that the VIF values of the retained predictors ranged from 1.07 to 1.67, with all values below 5, indicating no evidence of significant multicollinearity. Detailed results are presented in Multimedia Appendix 4. These variables were subsequently incorporated into the XGBoost model to predict postoperative frailty trajectories and to explore the relative contribution of each predictor to the classification outcomes.

Figure 1. Feature selection using the LASSO (least absolute shrinkage and selection operator) regression model. (A) LASSO coefficient profiles of the 25 candidate predictors. (B) Selection of the optimal tuning parameter (λ) using 10-fold cross-validation.



Model Development and Evaluation

To optimize XGBoost performance, key hyperparameters were tuned using grid search and cross-validation to improve model generalizability and accuracy. The model achieved excellent discriminative performance, with microaverage area under the receiver operating characteristic curve of 0.98 (95% CI 0.96 - 0.99) in the training set (Figure 2A) and 0.93 (95% CI 0.90 - 0.96) in the testing set (Figure 2B). Additional metrics, including accuracy, F_1-score, sensitivity, and specificity, are presented in Table 2. Calibration curves (Figure 2C) closely approximated the 45° line, indicating good agreement between predicted and observed probabilities, with a weighted Brier score of 0.085. Decision curve analysis demonstrated consistently greater net clinical benefit across a range of threshold probabilities compared to the "treat-all" and "treat-none" strategies, supporting the model's potential for clinical applications (Figure 2D).

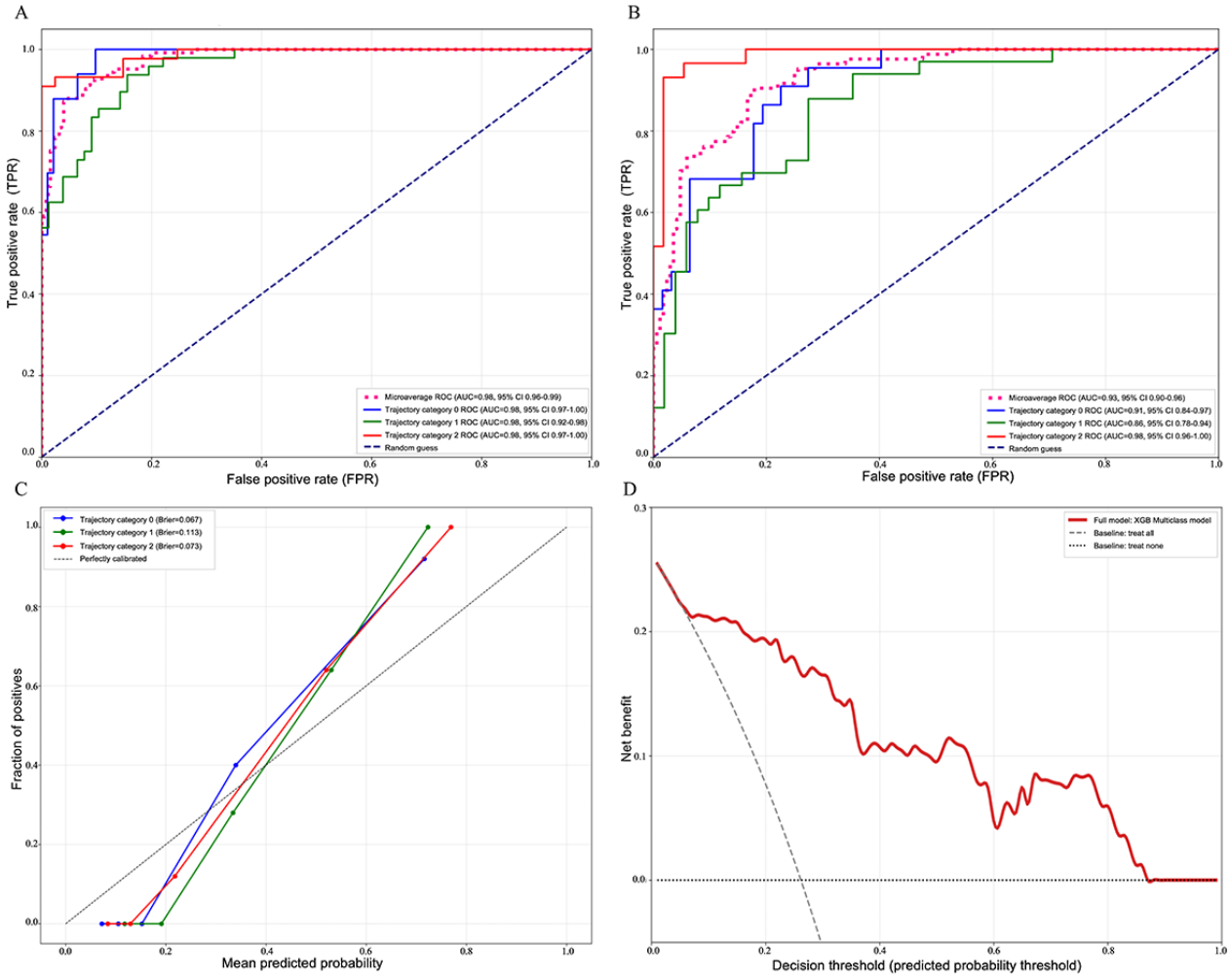
XSL•FO
RenderX

Table . The values of the evaluation metrics for the multiclass trajectory model in the test set.

Trajectory model	AUC[a] (95% CI)	Accuracy (95% CI)	F_1-score (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Frailty low fluctuation	0.91 (0.84-0.96)	0.88 (0.82-0.93)	0.87 (0.81-0.92)	0.88 (0.82-0.93)	0.88 (0.82-0.93)
Frailty high improvement	0.86 (0.77-0.94)	0.85 (0.78-0.91)	0.83 (0.76-0.89)	0.80 (0.72-0.87)	0.85 (0.78-0.91)
Frailty high deterioration	0.98 (0.96-1.00)	0.96 (0.93-0.99)	0.97 (0.94-1.00)	0.95 (0.92-0.98)	0.98 (0.95-1.00)
Overall	0.93 (0.90-0.96)	0.89 (0.84-0.94)	0.89 (0.83-0.94)	0.87 (0.80-0.93)	0.90 (0.85-0.95)

[a]AUC: area under the receiver operating characteristic curve.

Figure 2. Predictive performance and clinical utility of the frailty trajectory models. (A) Receiver operating characteristic (ROC) curves for the training set. (B) ROC curves for the test set. (C) Calibration plot of the Extreme Gradient Boosting (XGBoost) model. (D) Decision curve analysis (DCA) of the XGBoost model. AUC: area under the ROC curve. Dashed and dotted lines indicate the strategies of treating “all” or “no” patients, respectively.

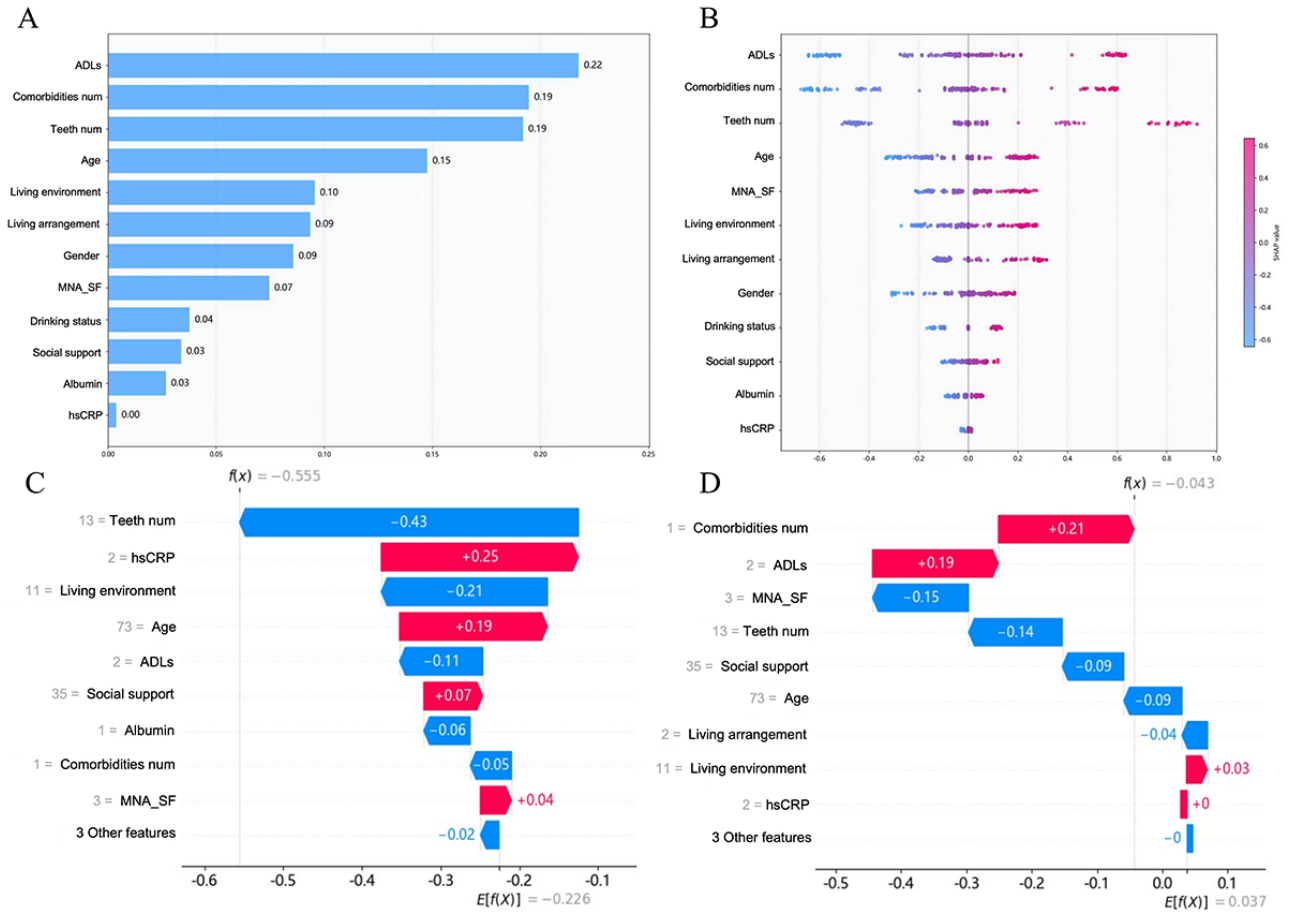


Model Interpretation

The SHAP algorithm was used to visualize and interpret the contribution of each predictor to the frailty trajectory model. As shown in Figure 3A, the bar plot displays the mean SHAP values of the 12 selected predictors, reflecting their average impact on the model output. ADLs had the greatest predictive importance, followed by the number of comorbidities and the number of teeth. Figure 3B presents the SHAP value distributions for each feature, with color gradients from blue (low values) to red (high values) representing the magnitude of each predictor and its influence on individual predictions.

XSL•FO
RenderX

Figure 3. SHAP (Shapley Additive Explanations)–based interpretation of the Extreme Gradient Boosting (XGBoost) frailty trajectory model. (A) Feature importance ranked by mean absolute SHAP values. (B) The beeswarm plot showing the distribution and direction of SHAP values for each feature. (C) The SHAP waterfall plot for an illustrative patient from the low-fluctuation frailty trajectory. (D) The SHAP waterfall plot for an illustrative patient from the high-deterioration frailty trajectory. ADL: activities of daily living; hsCRP: high-sensitivity C-reactive protein; MNA-SF: Mini Nutritional Assessment–Short Form.

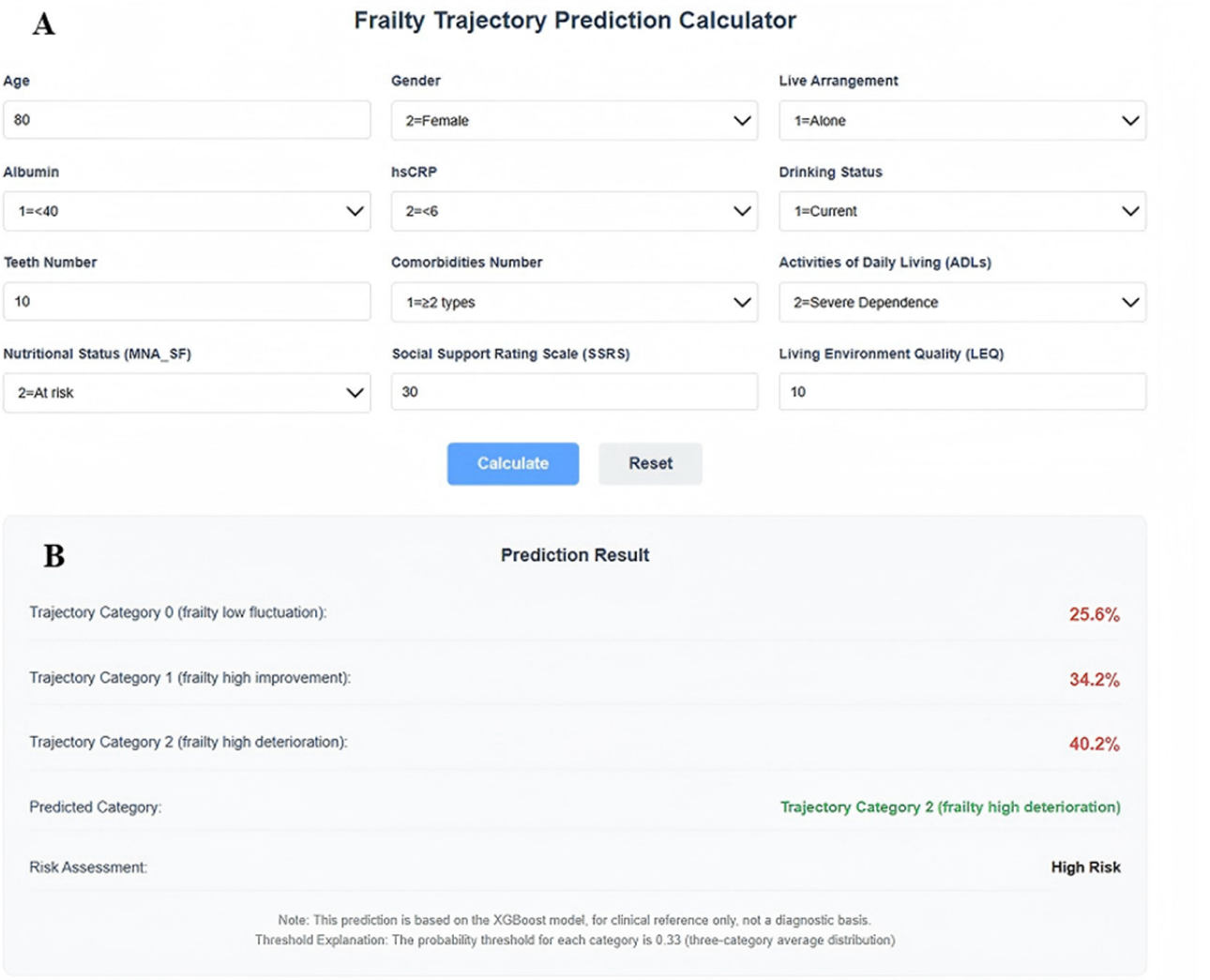


To visualize case-specific model explanations, Figure 3C and D presents SHAP waterfall plots for illustrative patients from the low-fluctuation and high-deterioration frailty trajectories, respectively. For these samples, the predicted model outputs were −0.555 and −0.043, compared with baseline values of −0.226 and 0.037. For the low-fluctuation trajectory, favorable living environments and higher teeth counts showed strong negative contributions to the model output, offsetting the positive contributions of age and hsCRP. In contrast, the high-deterioration trajectory was primarily characterized by positive contributions from comorbidities and ADL limitations, whereas nutritional status, social support, dental health, and age contributed negatively to the model output.

A web-based frailty trajectory prediction calculator was developed and is freely available online [32]. The user interface of the application is shown in Figure 4. This tool allows users to input key patient characteristics and obtain estimated probabilities for each frailty trajectory class, which may support frailty risk assessment and individualized management planning.

XSL•FO
RenderX

Figure 4. (A) Web-based calculator for predicting the probability of frailty trajectories following hip surgery. (B) Prediction result. hsCRP: high-sensitivity C-reactive protein; MNA-SF: Mini Nutritional Assessment–Short Form.



Discussion

Principal Findings

This study identified 3 distinct postoperative frailty trajectories among older adults following hip surgery: low fluctuation, high improvement, and high deterioration. These findings highlight that frailty after hip surgery is a dynamic process with substantial interindividual variability rather than a static condition. By moving beyond traditional binary outcomes [17,18], this trajectory-based approach provides a more nuanced understanding of recovery patterns that are often obscured in aggregate analyses. Furthermore, by integrating multidimensional predictors within the HEM, we developed a prediction model that demonstrated good discrimination and calibration performance. Collectively, these findings provide a basis for early risk stratification and the development of targeted, individualized intervention strategies in postoperative care.

Although the overall prevalence of frailty followed a rise-then-fall trajectory over the 6-month postoperative period, consistent with prior studies such as Huang et al [33], this general trend concealed the diversity of individual recovery paths. The early peak in frailty may be attributed to systemic inflammatory responses, catabolic stress, and functional limitations induced by postoperative pain, while the subsequent decline likely reflects recovery mechanisms, including tissue healing, pain relief, and rehabilitation engagement. Trajectory modeling in this study revealed 3 distinct frailty patterns: low fluctuation, high improvement, and high deterioration. These heterogeneous trajectories suggest that despite experiencing the same surgical stressor, older patients vary substantially in their recovery due to differences in baseline physiological reserve, comorbid conditions, social support, and environmental context.

A closer examination of these subgroups further highlights their clinical relevance. The low fluctuation group, despite having a relatively low baseline frailty level, experienced a notable increase at 1 month postoperatively. This transient deterioration may indicate reduced physiological resilience. Prior studies suggest that abrupt fluctuations in frailty may signal underlying instability and predict adverse events such as falls or rehospitalization more accurately than static frailty status [34]. For this group, interventions should aim to maintain functional stability by implementing prehabilitation, early mobilization, multimodal pain control, and prevention of iatrogenic complications [12]. In contrast, the high-improvement group demonstrated a steady recovery trajectory, likely reflecting

XSL•FO
RenderX

reversible functional impairments driven primarily by hip pathology. These patients may benefit from standard rehabilitation care without intensive resource input, allowing health care systems to prioritize those at greater risk. The high-deterioration group represents a clinically vulnerable subgroup with persistent decline, suggesting unresolved physiological or psychosocial deficits. Comprehensive geriatric assessments and tailored multidisciplinary interventions are warranted for this population, ideally guided by frameworks such as the World Health Organization (WHO)'s ICOPE (Integrated Care for Older People), which emphasizes multidomain support targeting mobility, cognition, nutrition, sensory function, and social engagement [35,36].

Based on the identified frailty trajectories, we developed a predictive model to support early risk identification and targeted intervention. Guided by the HEM, the model incorporated variables spanning individual characteristics, interpersonal networks, and living environment. Although some key predictors (eg, ADLs, comorbidity burden, and nutritional status) are conceptually related to domains of the FI, the multicollinearity analysis did not indicate significant statistical overlap among the predictors. This suggests that, despite conceptual relatedness, these variables do not represent redundant information from a statistical perspective. Nevertheless, these factors may partly reflect baseline vulnerability. Therefore, the model should not be interpreted as predicting frailty trajectory class membership entirely independent of baseline frailty severity. Instead, it is more appropriately viewed as an early risk stratification tool that integrates multidimensional information. Consistent with this interpretation, the SHAP analysis identified ADLs, number of comorbidities, number of teeth, age, and living environment as the top contributors to the model predictions.

In line with previous research [17,37], this study confirmed that impaired ADLs and a greater number of comorbidities were the strongest predictors of frailty deterioration. These variables represent fundamental aspects of functional dependence and chronic disease burden. Declining ADLs reflect a depletion of multisystem physiological reserves and may initiate a downward cycle in which the reduced function limits physical activity, weakens muscle strength, and accelerates frailty progression [37]. Similarly, multimorbidity contributes through persistent low-grade inflammation, polypharmacy, and cumulative organ damage, ultimately undermining mobility and physiological stability [38]. For individuals with impaired ADLs and multiple comorbidities, the early implementation of interventions such as individualized rehabilitation, optimized management of chronic conditions, and comprehensive medication review may be essential to interrupt this adverse trajectory [39,40]. In addition, age and teeth number were identified as important predictors. While age reflects irreversible biological decline, tooth loss serves as a modifiable marker of oral frailty. Reduced dentition in older adults impairs chewing function and nutritional intake, which in turn promotes sarcopenia and accelerates frailty [41]. These findings highlight the importance of incorporating oral health screening into perioperative management, along with timely interventions such as denture repair, chewing function training, and personalized nutrition support [14].

Among the identified predictors, living environment was the only environmental factor ranked among the top 5, highlighting its important role in frailty progression. Although the LEQ was a self-developed instrument, its core domains are conceptually aligned with the WHO age-friendly environment framework [42]. Previous studies have shown that environmental factors such as housing safety, community resources, social participation, and perceived environmental support are closely associated with frailty and its progression in older adults [43-45]. Notably, our findings demonstrated that living environment independently predicted frailty trajectories beyond individual-level characteristics. This finding is consistent with recent evidence showing that more positive perceptions of community environments are associated with lower frailty risk [46], as well as the findings of Mak et al [47], who reported sustained effects of environmental factors on frailty progression. For older adults undergoing hip surgery, environmental barriers after discharge may limit rehabilitation adherence and increase the risks of falls and readmission, thereby accelerating frailty progression. These findings emphasize the need to extend clinical care beyond hospital settings by incorporating environmental risk assessments into discharge planning and transitional care. Practical strategies may include predischarge home safety evaluations, caregiver education, and community linkage interventions to foster a more supportive recovery environment and reduce frailty-related risks [48].

In addition to the core predictors, the model identified several supplementary variables, including living arrangement, gender, nutritional status, drinking behavior, social support, albumin, and hsCRP. Although these variables contributed less individually, they enriched the understanding of frailty trajectories from physiological, behavioral, and demographic perspectives. These results indicate the importance of combining model-based risk identification with comprehensive biopsychosocial assessments to detect individualized frailty risk profiles and guide targeted interventions. Building on this, implementing a Fracture Liaison Service (FLS) may offer an effective integrative care pathway [49]. Through its structured components of assessment, education, intervention, and follow-up, Fracture Liaison Service enables continuous, coordinated management for older adults with hip fractures, supporting the transition from early risk prediction to precise clinical action.

Limitations

This study has several limitations. First, it was a prospective single-center study with a relatively small sample size, focusing exclusively on older patients undergoing elective hip surgery under general anesthesia, which may limit the generalizability of the findings to broader populations such as individuals receiving emergency surgery, regional anesthesia, or in other age groups. In addition, the relatively limited sample size for multiclass machine-learning modeling may increase the risk of overfitting and unstable feature selection despite the use of LASSO-based dimensionality reduction and internal cross-validation. Therefore, both the current prediction model and the associated web-based calculator should be considered exploratory and proof-of-concept tools rather than clinically deployable systems. Prospective multicenter external validation,

XSL•FO
RenderX

calibration assessment, and health-economic evaluation are warranted before routine clinical integration.

Second, a complete case analysis was used to handle missing data. Although the Little MCAR test suggested a completely-at-random mechanism and no baseline differences were observed between retained and excluded participants, departures from this assumption cannot be ruled out. If missingness were related to the severity of frailty or the presence of comorbidities, this could introduce bias and potentially affect trajectory classification and model performance. Therefore, the findings should be interpreted with caution. Future studies may consider more robust approaches for handling missing data, where appropriate, to further assess the robustness of the results.

Third, although the model incorporated multilevel predictors guided by the socioecological framework, some important but difficult-to-measure confounders, such as genetic predisposition or deeper physiological reserves, may have been missed. Future studies could leverage technologies such as wearable devices and multiomics approaches to improve measurement precision and expand variable inclusion. Finally, the 6-month follow-up period may capture only early or acute frailty changes and fail to reflect long-term trajectories. Longer follow-up and analytical methods such as competing risk models may help differentiate between temporary functional decline and persistent frailty, thus identifying critical intervention windows.

Conclusion

This study developed and internally validated a proof-of-concept machine-learning model to predict frailty trajectories within 6 months after hip surgery in older adults. Grounded in the HEM, the model integrated multidimensional predictors and showed good performance in discrimination, calibration, and clinical utility. Key contributing factors included functional status, comorbidity burden, oral health, age, and living environment. The model was translated into a web-based tool to support early identification and personalized frailty management; it currently serves as a foundational step. Comprehensive external validation across diverse clinical settings is strictly required before this tool can be deployed in routine clinical practice. These findings provide methodological and practical support for shifting perioperative care toward a proactive and individualized approach.

Acknowledgments

The authors thank Kaiqing Zhou (Shanghai Maritime University) for assistance with the data analysis and valuable discussion. They would like to express their gratitude for the funding provided, which enabled the completion of this study. During the preparation of this manuscript, generative artificial intelligence tools were used solely for English language editing and text polishing. The authors have thoroughly reviewed and edited the content and remain fully responsible for the accuracy, originality, and integrity of the final manuscript.

Funding

This work was supported by grants from the Department of Zhejiang Health Science and Technology Project Science and Technology (grant number 2025KY081).

Data Availability

The datasets generated and/or analyzed during this study are available from the corresponding author upon reasonable request. To ensure methodological transparency and reproducibility in compliance with the TRIPOD+AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis—Artificial Intelligence) guidelines, the machine-learning pipelines (LASSO [least absolute shrinkage and selection operator] and Extreme Gradient Boosting–SHAP [Shapley Additive Explanations]) are publicly available on GitHub [50].

Authors' Contributions

Conceptualization: JH
Data curation: JH, QF, HG
Formal analysis: JH, HG
Funding acquisition: JH
Investigation: JH, KG, JC
Methodology: QF, HG
Project administration: JH
Visualization: JH, HG, HS, JY
Writing – original draft: JH
Writing – review and editing: HS, YJ, KG, JC
All authors read, reviewed, and approved the final manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Variables and measurements.
[DOCX File, 26 KB - aging_v9i1e90705_app1.docx]

Multimedia Appendix 2
Participant characteristics and missing data.
[DOCX File, 66 KB - aging_v9i1e90705_app2.docx]

Multimedia Appendix 3
Frailty trajectory modeling (group-based trajectory modeling).
[DOCX File, 68 KB - aging_v9i1e90705_app3.docx]

Multimedia Appendix 4
Predictor selection and multicollinearity.
[DOCX File, 18 KB - aging_v9i1e90705_app4.docx]

References

1. Ridha M, Al-Jabri T, Stelzhammer T, Shah Z, Oragui E, Giannoudis PV. Osteosynthesis, hemiarthroplasty, total hip arthroplasty in hip fractures: all I need to know. Injury 2024 Mar;55(3):111377. [doi: 10.1016/j.injury.2024.111377] [Medline: 38324951]
2. Ong T, Vindlacheruvu M. A commentary update on NICE CG124. Hip fracture:management (2023). Age Ageing 2023 Jun 1;52(6):afad110. [doi: 10.1093/ageing/afad110] [Medline: 37389557]
3. Chen CL, Chen CM, Wang CY, et al. Frailty is associated with an increased risk of major adverse outcomes in elderly patients following surgical treatment of hip fracture. Sci Rep 2019 Dec 13;9(1):19135. [doi: 10.1038/s41598-019-55459-2] [Medline: 31836751]
4. Dent E, Hanlon P, Sim M, et al. Recent developments in frailty identification, management, risk factors and prevention: a narrative review of leading journals in geriatrics and gerontology. Ageing Res Rev 2023 Nov;91:102082. [doi: 10.1016/j.arr.2023.102082] [Medline: 37797723]
5. Wang SK, Wang QJ, Wang P, et al. The impact of frailty on clinical outcomes of older patients undergoing enhanced recovery after lumbar fusion surgery: a prospective cohort study. Int J Surg 2024 Aug 1;110(8):4785-4795. [doi: 10.1097/JS9.0000000000001594] [Medline: 38729123]
6. Eslami MH, Saadeddin Z, Rybin DV, Doros G, Siracuse JJ, Farber A. Association of frailty index with perioperative mortality and in-hospital morbidity after elective lower extremity bypass. J Vasc Surg 2019 Mar;69(3):863-874. [doi: 10.1016/j.jvs.2018.07.050] [Medline: 30527215]
7. Lunca S, Morarasu S, Rouet K, et al. Frailty increases morbidity and mortality in patients undergoing oncological liver resections: a systematic review and meta-analysis. Ann Surg Oncol 2024 Oct;31(10):6514-6525. [doi: 10.1245/s10434-024-15571-8] [Medline: 38856830]
8. Song Y, Wu Z, Huo H, Zhao P. The impact of frailty on adverse outcomes in geriatric hip fracture patients: a systematic review and meta-analysis. Front Public Health 2022;10:890652. [doi: 10.3389/fpubh.2022.890652] [Medline: 35844855]
9. Partridge JSL, Ryan J, Dhesi JK, CPOC-BGS perioperative frailty guideline group. New guidelines for the perioperative care of people living with frailty undergoing elective and emergency surgery-a commentary. Age Ageing 2022 Nov 2;51(11):afac237. [doi: 10.1093/ageing/afac237] [Medline: 36436009]
10. Liu Q. Construction of a predictive model for new-onset frailty in elderly patients after hip fracture surgery [Thesis in Chinese]. : Shanxi Medical University; 2023.
11. Wong BLL, Chan YH, O'Neill GK, Murphy D, Merchant RA. Frailty, length of stay and cost in hip fracture patients. Osteoporos Int 2023 Jan;34(1):59-68. [doi: 10.1007/s00198-022-06553-1] [Medline: 36197493]
12. Andaloro S, Cacciatore S, Risoli A, et al. Hip fracture as a systemic disease in older adults: a narrative review on multisystem implications and management. Med Sci (Basel) 2025 Jul 11;13(3):89. [doi: 10.3390/medsci13030089] [Medline: 40700118]
13. Pan Y, Li F. Analysis of factors associated with postoperative frailty in elderly patients with femoral neck fracture [Article in Chinese]. Pract Geriatr 2020;34(5):447-450. [doi: 10.3969/j.issn.1003-9198.2020.05.009]
14. Jia Y, Liang X, Huo Y, et al. Expert consensus on perioperative frailty management for elderly hip fracture patients [Article in Chinese]. J Orthop Clin Res 2025;10(2):65-71. [doi: 10.19548/j.2096-269x.2025.02.001]
15. Miao X, Guo Y, Chen Y, et al. Exploration of frailty trajectories and their associations with health outcomes in older gastric cancer survivors undergoing radical gastrectomy: a prospective longitudinal observation study. Eur J Surg Oncol 2024 Feb;50(2):107934. [doi: 10.1016/j.ejso.2023.107934] [Medline: 38160495]
16. Dent E, Martin FC, Bergman H, Woo J, Romero-Ortuno R, Walston JD. Management of frailty: opportunities, challenges, and future directions. Lancet 2019 Oct 12;394(10206):1376-1386. [doi: 10.1016/S0140-6736(19)31785-4] [Medline: 31609229]

XSL•FO
RenderX

17. Gou D, Wu H, Min C, et al. An interactive dynamic online nomogram for predicting the frailty trajectories among older people with chronic heart failure. J Affect Disord 2025 Aug 1;382:201-209. [doi: 10.1016/j.jad.2025.04.033] [Medline: 40246084]
18. Miao X, Guo Y, Ding L, et al. A dynamic online nomogram for predicting the heterogeneity trajectories of frailty among elderly gastric cancer survivors. Int J Nurs Stud 2024 May;153:104716. [doi: 10.1016/j.ijnurstu.2024.104716] [Medline: 38412776]
19. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15(4):351-377. [doi: 10.1177/109019818801500401] [Medline: 3068205]
20. Huang C, Sirikul W, Buawangpong N. Social frailty in community-dwelling older adults: a scoping review. BMC Geriatr 2025 May 13;25(1):329. [doi: 10.1186/s12877-025-05971-0] [Medline: 40360975]
21. Takefuji Y. Beyond XGBoost and SHAP: unveiling true feature importance. J Hazard Mater 2025 May 5;488:137382. [doi: 10.1016/j.jhazmat.2025.137382] [Medline: 39879771]
22. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
23. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020 Mar 18;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]
24. Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. J Gerontol A Biol Sci Med Sci 2007 Jul;62(7):722-727. [doi: 10.1093/gerona/62.7.722] [Medline: 17634318]
25. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6]
26. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
27. Soysal P, Isik AT, Arik F, Kalan U, Eyvaz A, Veronese N. Validity of the Mini-Nutritional Assessment Scale for evaluating frailty status in older adults. J Am Med Dir Assoc 2019 Feb;20(2):183-187. [doi: 10.1016/j.jamda.2018.07.016] [Medline: 30262439]
28. Smilkstein G, Ashworth C, Montano D. Validity and reliability of the family APGAR as a test of family function. J Fam Pract 1982 Aug;15(2):303-311. [Medline: 7097168]
29. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annu Rev Clin Psychol 2010;6(1):109-138. [doi: 10.1146/annurev.clinpsy.121208.131413] [Medline: 20192788]
30. Akinwande MO, Dikko HG, Samson A. Variance inflation factor: as a condition for the inclusion of suppressor variable(s) in regression analysis. Open J Stat 2015;05:754-767. [doi: 10.4236/ojs.2015.57075]
31. shinyapps.io. URL: https://www.shinyapps.io/ [accessed 2026-06-04]
32. Frailty Trajectory Prediction Calculator. URL: https://huangjingying-hub.github.io/Frailty-Calculator/ [accessed 2026-06-05]
33. Huang K, Ding S, Zhou H, Cai Y. Study on frailty trajectories and influencing factors in lung transplant patients [Article in Chinese]. Chin J Nurs 2023;58(9):1088-1095. [doi: 10.55111/j.issn2709-1961.202304134]
34. Stow D, Matthews FE, Hanratty B. Frailty trajectories to identify end of life: a longitudinal population-based study. BMC Med 2018 Sep 21;16(1):171. [doi: 10.1186/s12916-018-1148-x] [Medline: 30236103]
35. Duchesneau ED, Kim DH, Stürmer T, et al. Frailty trajectories following adjuvant chemotherapy and mortality in older women with breast cancer. JAMA Netw Open 2025 Mar 3;8(3):e250614. [doi: 10.1001/jamanetworkopen.2025.0614] [Medline: 40072432]
36. Integrated Care for Older People (ICOPE): guidance for person-centred assessment and pathways in primary care, second edition. : World Health Organization; 2024 URL: https://iris.who.int/server/api/core/bitstreams/7516b015-e205-43d1-883f-01f2c31af261/content [accessed 2026-06-03]
37. Du J, Zhang M, Zeng J, et al. Frailty trajectories and determinants in Chinese older adults: a longitudinal study. Geriatr Nurs 2024;59:131-138. [doi: 10.1016/j.gerinurse.2024.06.015] [Medline: 39002503]
38. Toh JJY, Zhang H, Soh YY, Zhang Z, Wu XV. Prevalence and health outcomes of polypharmacy and hyperpolypharmacy in older adults with frailty: a systematic review and meta-analysis. Ageing Res Rev 2023 Jan;83:101811. [doi: 10.1016/j.arr.2022.101811] [Medline: 36455791]
39. Izquierdo M, Ramírez-Vélez R, Fiatarone Singh MA. Integrating exercise and medication management in geriatric care: a holistic strategy to enhance health outcomes and reduce polypharmacy. Lancet Healthy Longev 2025 Sep;6(9):100763. [doi: 10.1016/j.lanhl.2025.100763] [Medline: 41015054]
40. Ruiz JG, Dent E, Morley JE, et al. Screening for and managing the person with frailty in primary care: ICFSR consensus guidelines. J Nutr Health Aging 2020;24(9):920-927. [doi: 10.1007/s12603-020-1492-3] [Medline: 33155616]
41. Dibello V, Zupo R, Sardone R, et al. Oral frailty and its determinants in older age: a systematic review. Lancet Healthy Longev 2021 Aug;2(8):e507-e520. [doi: 10.1016/S2666-7568(21)00143-4] [Medline: 36098000]
42. Age-friendly environments in Europe: indicators, monitoring and assessments. : World Health Organization, Regional Office for Europe; 2018 URL: https://iris.who.int/server/api/core/bitstreams/cf9f2ba0-85f5-45ed-a786-fe6b7e97f09f/content [accessed 2026-05-07]

XSL•FO
RenderX

43. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
44. Mori Y, Tsuji T, Watanabe R, Hanazato M, Chen YR, Kondo K. Built environments and frailty in older adults: the JAGES longitudinal study using mediation analysis. J Am Med Dir Assoc 2023 Nov;24(11):1677-1682. [doi: 10.1016/j.jamda.2023.06.023] [Medline: 37541649]
45. Gale CR, Westbury L, Cooper CJA. Social isolation and loneliness as risk factors for the progression of frailty: the English longitudinal study of ageing. Age Ageing 2018 May 1;47(3):392-397. [doi: 10.1093/ageing/afx188] [Medline: 29309502]
46. Ye B, Li Y, Chen L, Gong K, Bao Z, Gao J. Perceived environmental age friendliness and frailty among community-dwelling older adults. JAMA Netw Open 2026 Feb 2;9(2):e2558619. [doi: 10.1001/jamanetworkopen.2025.58619] [Medline: 41670998]
47. Mak JKL, Kuja-Halkola R, Bai G, et al. Genetic and environmental influences on longitudinal frailty trajectories from adulthood into old age. J Gerontol A Biol Sci Med Sci 2023 Feb 24;78(2):333-341. [doi: 10.1093/gerona/glac197] [Medline: 36124734]
48. Liu H, Zhang M, Chen B, Huang L, Zhao X. Links between life-course SES and frailty trajectory moderated by community environment resources: person-environment fit perspective. Adv Life Course Res 2023 Dec;58:100580. [doi: 10.1016/j.alcr.2023.100580] [Medline: 38054872]
49. Patel S, Daniels NF, Lim JA, Zhou AK, Thahir A, Krkovic M. The importance of fracture liaison services to the healthcare system: a review. Curr Rheumatol Rev 2023;19(1):36-41. [doi: 10.2174/1573397118666220510162127] [Medline: 35538826]
50. Huang J. Frailty-trajectory-prediction repository. GitHub. URL: https://github.com/HuangJingying-hub/Frailty-Trajectory-Prediction [accessed 2026-06-11]

Abbreviations

ADL: activity of daily living
APGAR: Adaptability, Partnership, Growth, Affection, Resolve
AvePP: average posterior probability
BIC: Bayesian Information Criterion
FI: Frailty Index
HEM: Health Ecology Model
ICOPE: Integrated Care for Older People
LASSO: least absolute shrinkage and selection operator
LEQ: Living Environment Questionnaire
MCAR: missing completely at random
OCC: odds of correct classification
SHAP: Shapley Additive Explanations
SMOTE: Synthetic Minority Oversampling Technique
TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis—Artificial Intelligence
WHO: World Health Organization
XGBoost: Extreme Gradient Boosting

Edited by M Gray; submitted 02.Jan.2026; peer-reviewed by B Cedeno-Veloz, KR Mun; revised version received 11.May.2026; accepted 28.May.2026; published 16.Jun.2026.

Please cite as:
Huang J, Fan Q, Shi H, Chen J, Gao H, Gu K, Yang J
Predicting Frailty Trajectories Using Interpretable Machine Learning Among Older Adults Following Hip Surgery: Prospective Longitudinal Study
JMIR Aging 2026;9:e90705
URL: https://aging.jmir.org/2026/1/e90705
doi:10.2196/90705

© Jingying Huang, Qinqin Fan, Huiqin Shi, Jiaojiao Chen, Hongbo Gao, Kailai Gu, Jin Yang. Originally published in JMIR Aging (https://aging.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Indoor Movement Complexity in Smart Homes to Detect Frailty in Older Adults: Multiple-Methods Case Series Study

Katherine Wuestney[1], PhD; Diane Cook[2], MSci, PhD; Catherine Van Son[1], MSN, PhD; Roschelle Fritz[3], MSN, PhD

[1]College of Nursing, Washington State University, Pullman, WA, United States

[2]School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA, United States

[3]Betty Irene Moore School of Nursing, UC Davis Health System, 2570 48th St, Sacramento, CA, United States

Corresponding Author:
Roschelle Fritz, MSN, PhD
Betty Irene Moore School of Nursing, UC Davis Health System, 2570 48th St, Sacramento, CA, United States

Abstract

Background: The theory of complexity in aging indicates that the complexity of sensor-derived physiological and behavioral signals reflects an older adult's adaptive capacity and, in turn, their frailty. Smart homes with ambient sensors offer a unique opportunity to longitudinally explore the complexity of older adults' indoor movement in a real-world setting. Here, we introduce a computational method to estimate behavior complexity from sensor data. We further conduct a multiple-methods case series to explore the relationship between entropy-measured smart home data complexity and older adult frailty.

Objective: This study aims to explore the relationship between entropy-measured ambient sensor data complexity and frailty in independent community-dwelling older adults.

Methods: The nature of older adults' indoor movement complexity is measured by quantifying the entropy of smart home data. Overall, 11 cases with persons aged 65 years and older were drawn from an ongoing smart home study to illustrate the method. We assessed weekly frailty for these cases using the Clinical Frailty Scale. For corresponding time ranges, we measured the complexity of smart home data using a fixed-width sliding window and an entropy-based complexity index (Rényi Complexity Index) built on a Universal Sequence Map (USM-Rényi). Descriptive statistics and graphical analysis were used to describe intraindividual frailty and sensor complexity change.

Results: The complexity of sensor-observed indoor movement does change over time in older adults as quantified by the computational method. In some individuals, these changes track with health transitions and frailty progression. The trends and monotonicity of complexity trajectories varied between cases. Overall, 3 of the cases demonstrated a negative association between frailty and complexity, while the association was not as clear for the other cases.

Conclusions: The complexity of older adults' smart home data is highly diverse. Changes in health and frailty influence indoor movement complexity. Although the findings suggest a relationship between frailty and complexity, confounding factors, such as home layout, visitors, external events, and technology disruptions, may influence sensor signals.

(JMIR Aging 2026;9:e77322) doi:10.2196/77322

KEYWORDS

entropy; complexity; smart home; movement; frailty

Introduction

Background

Frailty is a critical public health challenge among older adults globally. It is characterized as a clinically identifiable state of diminished physiologic reserve and heightened vulnerability to stressors and affects. An estimated 10% - 15% of community-dwelling individuals aged 65 years and older experience frailty, with the prevalence escalating to 51% among those aged 90 years and older [1]. This multifactorial syndrome, encompassing multiple impairments such as physical weakness, exhaustion, slow gait, low activity levels, and unintentional weight loss, elevates risks for adverse outcomes like falls, hospitalizations, functional decline, and mortality [1,2]. When unaddressed, frailty imposes substantial economic burdens, with frail older adult women incurring as much as 184% the health care cost of nonfrail older adult women [3,4]. Poor outcomes related to frailty strain health care systems and diminish quality of life [3-5].

Most older adults prefer to remain in their own homes and communities as they age. Despite this desire, age-related frailty and its sequelae remain a threat to their independence and quality of life. Older adults who are frail are more likely to present with atypical, nonspecific symptoms of acute illness, which include immobility, instability, incontinence, weakness, and delirium [6]. This can put them at risk for poorer outcomes

XSL•FO
RenderX

if such atypical signs are treated as the primary problem rather than merely the manifestation of underlying, seemingly unrelated illnesses. Thus, while frailty is a significant issue, it also functions as a gateway to a wide array of other salient health issues for older adults [7].

Two urgent challenges face older adults who wish to age in place: (1) needing validated methods to detect incipient frailty at home and (2) determining the best way to analyze these data for predicting frailty that focuses on the efforts of early intervention strategies [8]. This study addresses both challenges by exploring complexity as a feature of frailty in smart home sensor data. In the context of aging, the term "complexity" is often used to describe difficult problems that must be mitigated, making care more daunting [9]. While complex behavior is often seen as a challenge in caregiving, from a systems theory perspective, too little complexity may indicate diminished physiological adaptability. The theory of complexity in aging asserts that complexity is a direct indicator of the health of physiologic systems and aging reduces this complexity, resulting in frailty [10].

Sensors are ubiquitous in our world, and this reality is accompanied by an increasing interest in discovering indicators of human health using these sensors. These indicators serve a similar function as conventional biological and imaging biomarkers with less reliance on expensive lab equipment, visits to remote sites for time-consuming tests, or physically invasive procedures [11]. Digital biomarkers are valuable components of geriatric telehealth and precision medicine since these technologies support continuous, longitudinal, remotely delivered measurement of intraindividual changes in older adults' health [12]. Among this class of markers are behavior markers created from continuously collected sensor data, which open substantial opportunities to explore the complex dynamics of aging in an ecologically valid, real-world setting.

In this study, we enlist digital biomarkers to explore the relationships between behavior and frailty. This is increasingly important because the number of persons aged 80 years and older is expected to triple between 2020 and 2050 [13]. With the rise in age-related frailty and incidence of chronic conditions, meeting the health needs of older adults is increasingly burdensome. A review of unobtrusive frailty digital biomarkers concluded that passive infrared motion sensors, especially as part of a smart home, are the most promising type of embedded ambient sensor for detecting frailty [14]. Smart homes were promoted for their potential to uniquely inform individual responses to disease or treatment.

Older adults prefer digital biomarker technologies that minimally impose on their lifestyles [15]. From this perspective, digital behavior markers derived from completely passive monitoring (eg, ambient sensors embedded in residential environments) offer advantages over those measured via semipassive or active monitoring (eg, wearable sensors that must be routinely charged and positioned) [14]. Smart homes represent a passive biomarker technology that consists of ambient sensors to monitor movement and door interactions, combined with a computing infrastructure to collect, organize, and store the data [16]. The resulting time series data can be analyzed to understand the smart home resident's health status.

Prior Work

The theory of complexity in aging hypothesizes that measuring the complexity of a person's sensor-derived signals can indicate the underlying state of an older adult's adaptive capacity [17,18]. We analyze a person's behavioral signal complexity as an indicator of their adaptive capacity or functional reserve, referring to the capacity of their physiological and behavioral systems to maintain or regain function when perturbed. In a complex-systems view of aging, this reserve depends on the multiscale dynamics that are present between the system components [19]. These dynamics support homeostasis, the process that maintains internal stability while adapting to change.

In earlier work, researchers have investigated the use of multiscale entropy (MSE) to quantify complexity across time scales for physiological data. Bizovska et al [20] used MSE and Shannon entropy to analyze gait complexity as a mechanism for predicting fall risk in older adults. Castiglia et al [21] investigated the selection of MSE parameters that yield the best predictive probability in differentiating subjects with Parkinson disease from healthy subjects based on trunk acceleration patterns. Gao et al [22] use distribution entropy, which calculates the complexity of signal pattern distribution within a phase space representation, to determine whether pulse rate complexity is associated with corresponding cognitive decline in older adults.

Frailty is hypothesized to be an emergent state that arises from a critically dysregulated complex system [9]. In other words, the system dynamics may erode with aging and disease, causing complexity to decline and frailty vulnerability to increase. Evidence from cross-sectional studies suggests this process can be observed as a change in complexity in a diverse range of physiological and behavioral signals. For example, lower blood pressure interbeat interval complexity, when the beat-to-beat pattern becomes simpler and more uniform, is associated with greater frailty and dementia risk [23]. Similarly, lower moment-to-moment center-of-pressure complexity, such as simpler, more regular sway, during balance tasks is associated with increased future incidence of falls [24]. Reduced complexity of spontaneous brain activity, measured via the blood oxygenation signal, is associated with slower gait speed [25], and lower physical activity complexity and variance are associated with greater self-reported frailty [26] and mortality risk [27]. These examples suggest that reduced signal complexity may be a generalizable marker of physiologic decline.

Prior studies have measured the complexity of smart home data [28,29]. These earlier studies included complexity as one of a set of variables input to machine-learning models that were trained for specific tasks such as detecting visitors or predicting in-home movement [30,31]. Little is known regarding how within-person complexity, in isolation from other variables, evolves in relation to health outcomes over the long term. In a study by Schutz et al [28], Shannon entropy of refrigerator use was one of the strongest predictors of frailty (r=−0.25). However, this analysis did not explore the evolution of

XSL•FO
RenderX

complexity over time. A study by Takahashi et al [32] examined activities over a 2-year period and found that increased activity diversity manifested an inverse relationship with frailty. The findings support our hypothesis, but they are based on survey data rather than analysis of passively observed activity patterns.

Two prior studies applied complexity measures to data collected by the Center for Advanced Studies in Adaptive Systems (CASAS), the same data collection infrastructure used in this study. Specifically, Wang et al [30] estimated complexity using compression-based estimators to establish a theoretical limit on the predictability of indoor human mobility. In an earlier study by Gopalratnam and Cook [33], CASAS smart home data were analyzed with a Lempel-Ziv compression-based incremental parser to predict the resident's next interaction with the home. Although the smart home sites and analysis goals differed from this study, the prior work established the use of such behavior analyses from smart home data.

The common approach to predicting frailty leverages sources such as electronic health records and manually collected clinical data [34]. However, wearable sensors are increasingly accessible and offer a mechanism for passively sensing and detecting frailty [8]. Many frailty studies that analyze wearable data focus on predicting physical frailty components such as slowness and inactivity. These studies extract gait parameters such as cadence and indicators of time spent walking and standing [35,36]. One study instead analyzed Fitbit data that were collected while individuals performed an upper extremity function test [37]. While the primary component of these analyses is accelerometry, Merchant et al [38] combine these parameters with heart rate to analyze scripted movements such as sit-to-stand, walk, and climb stairs.

Wearable sensors have demonstrated the ability to sense and quantify changes in movement parameters that are associated with frailty. We focus here on monitoring activity and detecting frailty using ambient sensors in smart homes. Ambient sensors impose no user burden. Sensors collect data for multiple years on a charge, which results in continuous, uninterrupted monitoring of in-home behavior as a person's health status changes. Using wearable sensors, consistent multiday wear is challenging, and adherence varies with demographics and cognition [39]. While wearable sensors provide direct access to heart rate and gait parameters, the smart home sensors contribute context-rich information about location traces, sleep and wake routines, and activity patterns that are not easily modeled from wearable data [40]. Because we want to monitor uninterrupted longitudinal behavior patterns, we focus this analysis on data collected in smart home settings.

To address this knowledge gap, we present an exploratory case series investigating how the complexity of older adults' indoor movement patterns, as captured by the CASAS smart home, changes over time in relation to changes in their health status. Considering that this relationship between complexity and frailty has been observed across a diverse range of seemingly unrelated physiologic and behavioral signals, we hypothesize that changes in the complexity of time series obtained from smart home sensors are similarly associated with changes in health status and frailty of the older adult smart home occupant.

The case series design prioritizes investigation of intraindividual interpretation and allows us to integrate each participant's clinical narrative into the analysis. To promote replication of methods and application to new data, we make the analysis and visualization tools publicly available for the community to use in the calculation of sensor-derived behavior complexity.

Methods

Overview

We performed a multiple-methods exploratory case series, combining participant narrative and qualitative nursing data with complexity analysis of smart home sensor time series data to contextualize intraindividual changes in complexity of indoor movement. A case series analysis was chosen because the method is useful for exploring intricate, real-world issues in novel ways, especially when triangulating data from different sources to discover differences and similarities across similar cases. The method fosters a more nuanced, valid, and actionable understanding of the cases under study.

Participants

We used secondary data from sensors installed in the homes of community-dwelling older adults between October 2016 and December 2022 as part of the ongoing clinician-in-the-loop smart home research study [41]. To be included in the clinician-in-the-loop study, participants had to be aged 60 years and older, have at least 1 chronic condition, and had to be proficient in English. For this case study series, we applied the additional criteria of living alone without pets for the entire duration of the data collection and collected a minimum of 9 months of smart home data. Cases were further excluded if a majority of the days and sensors were missing. The resulting sample consisted of 11 cases, representing a balance between stable participants and those who exhibited frailty transitions. For this case series, each participant is considered as 1 case. Among these participants, cases 8 through 10 exhibited constant frailty scores, while the others experienced frailty that fluctuated throughout the data collection.

Participant cases included in the present analysis lived in independent living apartments in continuing care retirement communities. Most cases' ages were in the range of 80 to 89 years, although 2 were aged 70 - 79 years and 1 was aged 90 - 99 years. All included cases identified as non-Hispanic White, and 7 of the cases identified as women. Information summarizing participants, their chronic health conditions, and their home characteristics is provided in Table 1.

XSL•FO
RenderX

Table . Demographics and data characteristics for each participant case.

Case	Age[a] (years)	Sex	Home type	Sensors	Days	Window size[b]	Chronic conditions
1	80 - 89	Female	1-bedroom apartment	15	310	25,159	CV[c], NM[d], Pain[e]
2	70 - 79	Female	1-bedroom apartment	13	416	63,419	CV, Pulm[f], Pain
3	80 - 89	Female	3-bedroom duplex	20	366	19,714	CV, Pulm, NM
4	80 - 89	Female	1-bedroom apartment	13	349	33,358	CV, Pulm, NM
5	80 - 89	Male	1-bedroom apartment	15	629	60,800	NM, Pain
6	80 - 89	Female	1-bedroom apartment	12	571	28,951	CV, Pulm, NM, Pain
7	80 - 89	Female	1-bedroom apartment	14	385	28,182	Pain, CI[g]
8	70 - 79	Male	2-bedroom duplex	22	330	49,303	CV, NM, Pain
9	80 - 89	Female	Studio apartment	11	354	35,310	CV, Pain, CI
10	90 - 99	Male	1-bedroom apartment	12	264	18,951	CV, NM, Pain
11	80 - 89	Male	1-bedroom apartment	13	286	34,903	CV, Pain

[a]To preserve privacy, age is given as a range.

[b]Sliding window size was determined by the maximum biweekly count of sensor messages (excluding OFF and CLOSE) observed in the participant's data.

[c]CV: cardiovascular.

[d] NM: neuromuscular.

[e] Pain: chronic pain.

[f]Pulm: pulmonary.

[g]CI: cognitive impairment.

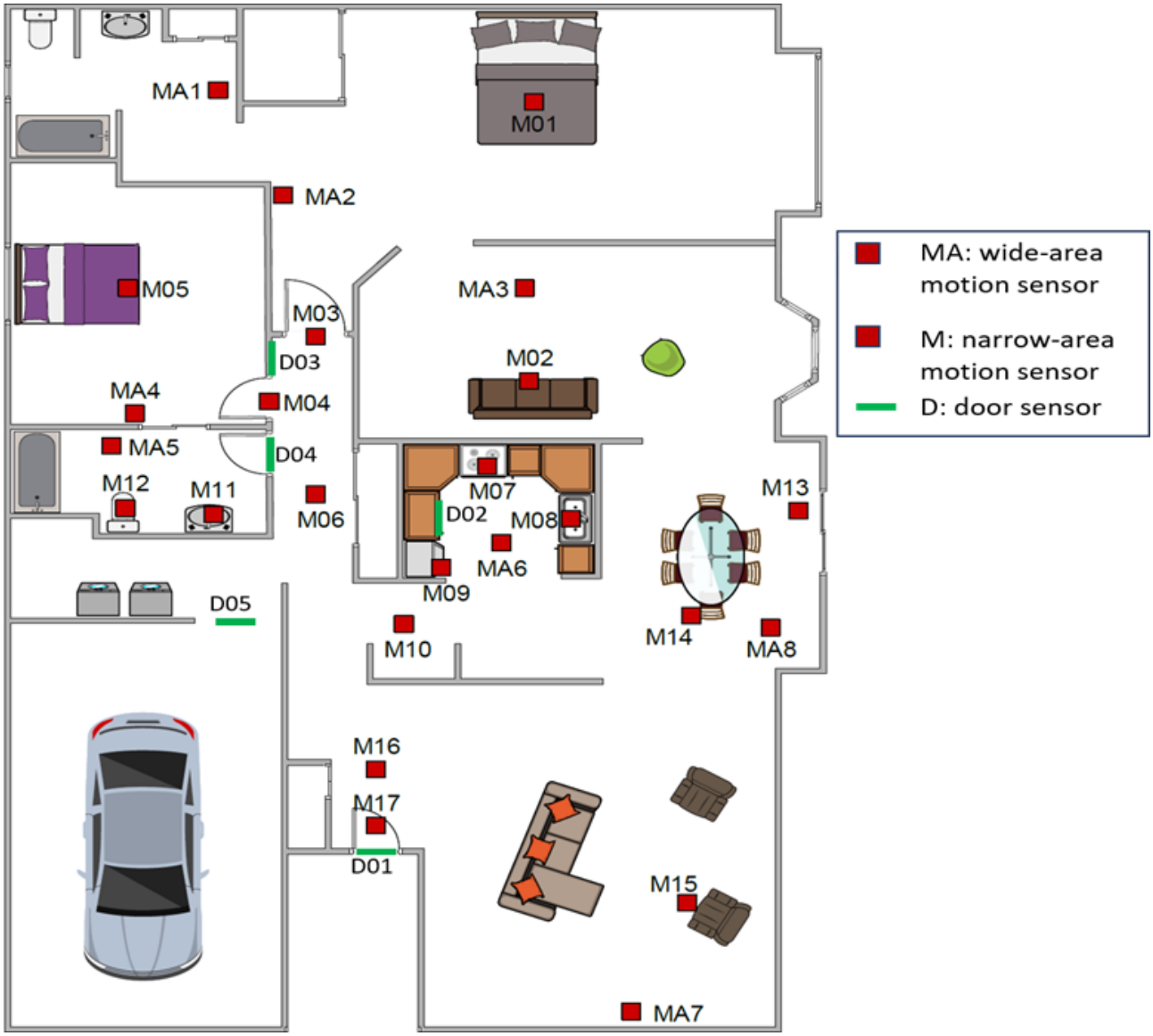
CASAS Smart Home

The CASAS smart home contains passive infrared motion detectors, light, magnetic door use, and temperature sensors. These sensors were installed strategically throughout each house to capture activity in critical locations (Figure 1). At least 1 motion detector with a 360° view was installed in each room. Additional motion detectors with a narrower field of view (approximately 1 m in diameter) were positioned in areas of high use, such as the bed, sinks, toilet, and frequented furniture (eg, preferred living room chair). Because floor plans, furniture layouts, and daily routines differed across each home, the number of sensors installed for included cases ranged from 11 to 22.

CASAS sensors send messages containing their readings to a middleware layer resident on a Raspberry Pi [16]. Architecture components communicate using a Zigbee wireless mesh. The middleware publish and subscribe manager allows hardware components to publish and receive messages. And annotates sensor readings with the corresponding sensor identifier and timestamp. All collected data are encrypted and securely transmitted to a password-protected server for storage and analysis.

XSL•FO
RenderX

Figure 1. Location of sensors in a Center for Advanced Studies in Adaptive Systems smart home.



We examined data collected from the passive infrared and door-use sensors. Each sensor samples the environment at 1.25 Hz. Rather than report the state at a constant frequency, the sensors record data when a change in state is sensed (eg, a door is opened, motion is detected). Once triggered, the sensor sends a message reflecting the new state to a central relay, which labels each message with the sensor identifier and timestamp, then transmits the data to a secured database. The resulting dataset is a timestamped series of binary messages (“ON” or “OFF” for motion sensors, “OPEN” or “CLOSE” for door sensors) indicating the time and location of the sensor reading in the home. Because an ON message from a motion sensor is followed by an OFF message (marking the end of movement within the sensor’s field of view or lack of activity for 1.25 s), both ON and OFF messages artificially inflated the regularity of the data sequence. Following previous literature measuring the entropy of smart home data [30,42], we excluded all OFF messages from motion sensors and all CLOSE messages from door sensors. Example deidentified CASAS datasets are available online [43].

Clinical Data

For each participant, nurse researchers conducted an initial comprehensive geriatric assessment, including functional status in activities of daily living (ADL) and instrumental activities of daily living (IADL), current health diagnoses, health history, medications, fall history, psychosocial supports and family presence, assistive device use, review of body systems, and personal demographic history. Participants then received weekly follow-up telehealth calls from a nurse researcher to assess for any changes in health or function from baseline. Weekly nursing data included, but was not limited to, vital signs, pain, sleep quality, psychosocial well-being (including the presence of visitors), changes in ADL and IADL status, and a brief review of physiologic systems and daily routines [41].

Although frailty was not measured as part of the primary data collection, the clinical data collected during the study provided information to retrospectively estimate weekly frailty using the Clinical Frailty Scale (CFS) [44]. The CFS is a 9-point scale (1=very fit to 9=terminal illness) designed to guide a clinician in assessing a holistic picture of a person’s frailty status using

XSL•FO
RenderX

elements of a comprehensive geriatric assessment, including overall activity level, functional dependence, and management and control of chronic condition symptoms [45]. Two CFS-trained nurse scientists reviewed the clinical data for each participant and assigned a frailty score for each week of data collection (Table 2). Changes in ADL and IADL independence, use of a new assistive device, and descriptions of increasing fatigue or "slowing down" were the most common health changes associated with an upward shift in the participant's CFS score.

Table . Example Clinical Frailty Scale codebook with scores for case 4.

Week[a]	Date (2017)	CFS[b] score	Rationale
45	July 24	5	No change
46	July 31	5	No change
47	August 7	5	Decreased activity, increased weeping lower legs
48	August 14	6	"I have to be careful not to fall"
49	August 21	6	Considering assisted living but hiring in-home help
50	August 28	6	No change
51	September 4	6	Losing weight, legs improving
52	September 11	6	Legs continue weeping due to heart failure
53	September 18	6	Began using pursed-lip breathing, moving less
54	September 25	6	Doctor's visit, medication change
55	October 2	6	Legs improving, taking diuretic

[a]Weeks 1 - 44 (CFS score: mean 4.9, SD 0.33; range 4 - 5); weeks 56 - 60 (CFS score: mean 6.4, SD 0.89; range 6 - 8).

[b]CFS: Clinical Frailty Scale.

Data Preprocessing

Because smart home data were collected in real-world settings over extended periods, we needed to address missing and noisy data. We screened each participant's sensor data for evidence of sensor malfunctioning, extended absences, and other issues. Periods associated with participant absence for more than one night (eg, vacation or hospitalization) were excluded from the analysis. Additionally, any periods where all sensors did not report readings, regardless of explanation, were excluded. Periods with no messages from a given sensor were cross-referenced with battery data from that sensor to confirm whether the absence was due to a change in behavior or a sensor malfunction. Sensors missing >50% of the observation time over one or more consecutive days were excluded. Sensors missing more than 50% of the observation period were either excluded or they were included, and the time associated with that sensor's absence was excluded.

The varying size of the homes and the corresponding number and density of sensors impact the scale of Rényi Complexity Index (RényiCI) values we observe in each home. A cross-sectional study would require that sensors be grouped into larger, consistently sized sets or that the values be normalized. For this study, we are interested in within-home RényiCI changes, so no adjustments are made to the per-home RényiCI scales. Because the sensors report binary state (motion ON or OFF, door OPEN or CLOSED), the raw sensor values are not normalized.

Some of the participants included in this study were enrolled during the onset of the COVID-19 pandemic, which had a dramatic global impact on daily activities. For those participants, if the majority of a participant's data were collected after the pandemic onset, data from before March 16, 2020, were excluded. Similarly, for participants with most data collected before the pandemic, we excluded data from March 16, 2020, onward.

Data cleaning included the removal of sensor data from analysis for sensors sending "error" signals, which can occur when low battery health or technical issues occur during installation. Only 2 homes were affected by this: case 1 (dining room area sensor, hallway, bathroom sink, and door for the primary bedroom) and case 5 (entry door, refrigerator, and bathroom area sensor). After data cleaning procedures were applied, the series of timestamped sensor messages was coded based on the sensor identifiers, resulting in a time series of discrete (categorical) sensor states. These discrete-valued series were then used to compute the complexity of sensor state transitions over time.

Complexity Measurement

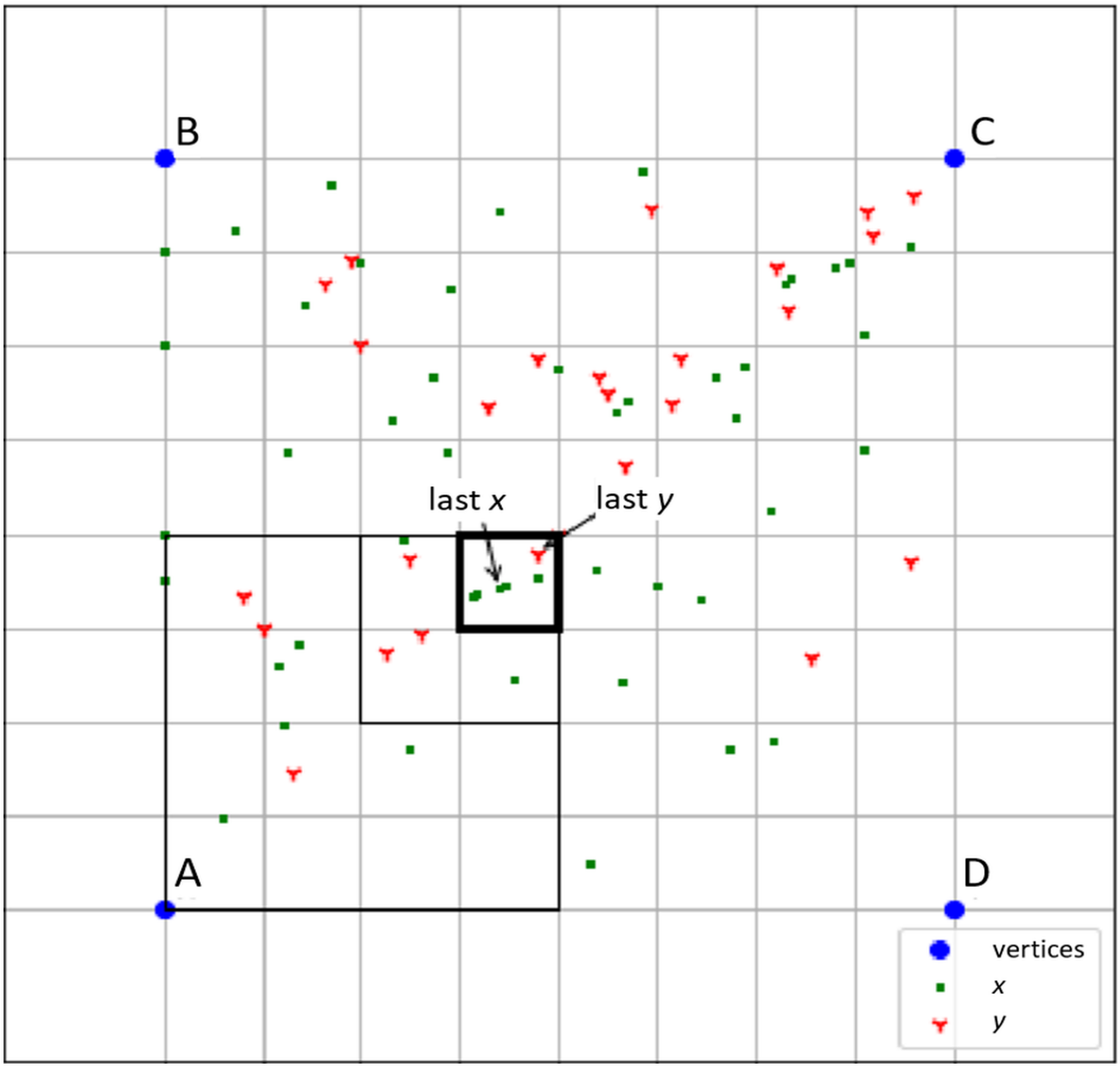
Understanding patterns in human behavior, especially those that signal changes in health or frailty, requires tools that quantify how predictable or irregular those behaviors are over time. One such method is based on entropy, a way of measuring complexity or unpredictability in a sequence of events. Almeida and Vinga [46] introduced a technique to calculate this complexity using a Universal Sequence Map (USM). This approach turns a sequence of events (eg, daily activities recorded

XSL•FO
RenderX

in a smart home) into a set of coordinates in a multidimensional space. These coordinates reflect how often and in what sequence specific symbols (eg, messages from home sensors) occur relative to one another. Once the sequence has been mapped to this space, the method estimates how densely packed these points are in space using the kernel density estimation statistical technique. The resulting density provides insight into whether the behavior is highly repetitive (low complexity) or highly varied (high complexity). Highly repetitive behavior may, for example, reflect a person moving primarily between the living room and bathroom each day. A more complex behavior will vary the daily sequence and perhaps more frequently introduce additional areas, including the guest room, the garage, and the front door to leave the home.

Figure 2 illustrates the process of creating a USM. Unique sensor readings are converted to symbols (A-D). USM coordinates are calculated by assigning each unit symbol in the sequence to a position in a multidimensional space. The positions are defined so that each symbol is equally distant from the others, ensuring that no symbol is biased in how the space is structured. The number of dimensions of the space is chosen so that each distinct symbol can be uniquely represented using binary digits. The sequence is processed forward (considering prior context) and backward (considering subsequent context), and the 2 resulting maps are merged to capture bidirectional structure in the behavior.

Figure 2. Plot of 2 sequences in a Universal Sequence Map. The last 4 symbols of sequence x are *ACCA*, and the last 4 of sequence y are *CCCA*. The highlighted subquadrant contains the coordinate of the sequences' last symbol, *A*.



The resulting space creates a unified framework to measure the complexity of sequences from the resulting coordinates. To quantify the complexity of these mapped sequences, Vinga and Almeida [47] introduced a method that computes Rényi entropy, a generalization of Shannon entropy, from the density of the USM coordinate distribution. This approach is particularly

XSL•FO
RenderX

effective for relatively short sequences, such as those representing daily behavior in smart homes.

Since the idea of entropy was introduced in information theory, many variations have been introduced to measure complexity in different contexts. These measures vary by the type and quantity of data they process, their sensitivity to noise, and their assumptions about the underlying state space distribution. An ideal measure of sensor-based time series complexity is one that tends toward a minimum value for both deterministic and random sequences while handling varying alphabet sizes and being sensitive to changes in complexity over short sample lengths. Rényi entropy of USMs was selected for our analysis as a method that meets these constraints.

A key strength of this method is its flexibility: it can emphasize either common or rare patterns, depending on how the parameters are configured. Importantly, the frequency of any subsequence of any length can be estimated by analyzing how dense different regions of the USM space are. The kernel size (ie, the size of the region considered) controls the length of the subsequences being emphasized. We use this principle to estimate Rényi entropy at multiple scales, where each scale corresponds to a different behavioral timespan or sequence length. This flexibility enables a multiscale view of behavioral complexity, which we refer to as the RényiCI. An in-depth tutorial and code are provided online [48].

Statistical Analysis

Because this case series investigates how the complexity of motion sensor transitions, representing indoor movement trajectories, evolves over time, we computed RényiCI for each participant using a sliding window approach with a fixed window size, *n*. The actual RényiCI values will shift with the number of sensors in the space and the window size; thus, the values should be examined for change within a single home across multiple time points. Higher RényiCI values indicate more complex behavior, while lower values suggest simpler, more predictable patterns. The sliding window method evaluates the time series in overlapping segments: starting with the first *n* data points, it computes summary statistics, shifts the window forward by a set number of steps, and repeats the process.

To ensure each window captured both routine cyclic behaviors (eg, weekly housekeeping) and short-term variations, we defined each participant's window size as the maximum number of sensor messages observed within any 2-week period (Table 3). The window was advanced using a step size equal to one-quarter of the window size.

XSL•FO
RenderX

Table . Sliding window statistics for each case. Runs test results were omitted as all resulted in *P* values <.001.

Sliding window			RényiCI statistics					ρ[a]	*P* value
Case	Count	Days, median (IQR; max)	Mean (SD)	CoV[b, c]	Median (IQR)	KS[d]	*P* value		
1	70	16.6 (15.5 to 17.2; 18.9)	−45.95 (0.26)	.006	−45.89 (−46.07 to −45.79)	0.10	.41	0.11	.48
2	66	24.7 (21.7 to 26.1; 30.8)	−38.85 (0.57)	.015	−38.65 (−39.09 to −38.47)	0.21	<.001	0.29	.06
3	80	16.3 (15.4 to 16.9; 18)	−61.97 (0.23)	.004	−61.94 (−62.18 to −61.78)	0.11	.29	0.12	.36
4	73	18.1 (16.8 to 19.2; 21)	−38.91 (0.27)	.007	−38.93 (−39.13 to −38.74)	0.09	.60	0.01	.96
5	121	18.1 (17.4 to 18.8; 30.7)	−44.66 (0.19)	.004	−44.68 (−44.81 to −44.51)	0.09	.23	−0.68	<.001
6	121	16 (15.5 to 16.6; 18.4)	−36.32 (0.10)	.003	−36.3 (−36.4 to −36.25)	0.09	.34	0.26	<.001
7	82	16.8 (15.6 to 17.7; 20.5)	−42.74 (0.14)	.003	−42.76 (−42.84 to −42.66)	0.07	.77	0.42	<.001
8	52	17.1 (16.1 to 18.1; 20.1)	−70.13 (0.20)	.003	−70.18 (−70.26 to −70.01)	0.12	.38	—[e]	—
9	73	18.3 (17.7 to 19; 20.6)	−33.07 (0.13)	.004	−33.07 (−33.15 to −32.99)	0.05	.98	—	—
10	49	18.9 (18 to 19.5; 22.6)	−36.64 (0.25)	.007	−36.64 (−36.84 to −36.43)	0.12	.47	—	—
11	65	15.5 (15 to 16.2; 18)	−39.58 (0.24)	.006	−39.58 (−39.78 to −39.35)	0.10	.53	0.24	.09

[a]Spearman rank correlation.

[b]CoV: coefficient of variance.

[c]Coefficient of variance was computed as the SD/mean.

[d]KS: Kolmogorov-Smirnov distance.

[e]Correlation is not provided because CFS is a constant.

To examine how the complexity of patterns relates to frailty status, we visualized RényiCI values using time series plots and categorical scatter (jitter) plots. Because RényiCI values can vary in scale depending on the number of sensors and the window size, we applied normalization within each case to enable comparison. To assess temporal fluctuations in complexity, we also computed the first-order difference of the normalized RényiCI sequence: ΔRenyiCI`t=RenyiCI`t-RenyiCI`t-1. Here, ΔRenyiCI`t represents the change in normalized complexity between consecutive windows.

To evaluate whether these complexity estimates varied systematically over time (in comparison to random changes in complexity), we applied Kolmogorov-Smirnov (KS) tests and runs tests to each participant's sequence of RényiCI values under the null hypothesis of randomness. The KS test checks whether the complexity values follow a normal distribution, as would be expected with random data. The runs test looks at the order of values in the sequence, rather than just the distribution, to determine if they appear in nonrandom patterns. Computation of USM-based RényiCI values was conducted in Python (version 3.9; Python Software Foundation) using our pyusm library [48]. This open-source package is publicly available and includes tools for computing USM, USM-Rényi, and generating 2D USM visualizations.

Finally, to resolve ambiguous quantitative results, sequential explanatory techniques were used. Quantitative results were reviewed alongside frailty scores assigned to each week of

XSL•FO
RenderX

nursing narrative documentation, which included written text about participants' physical and functional health recorded during weekly phone calls and monthly home visits. RényiCI complexity values were compared to recorded CFS scores. Lower complexity values combined with higher CFS scores meant the participant was frailer.

Ethical Considerations

The Washington State University Institutional Review Board approved the presented secondary analysis (protocol 18764) and parent study (protocol 15412). All participants provided informed consent, and their data were deidentified and securely managed for analysis. Participation was voluntary and without compensation.

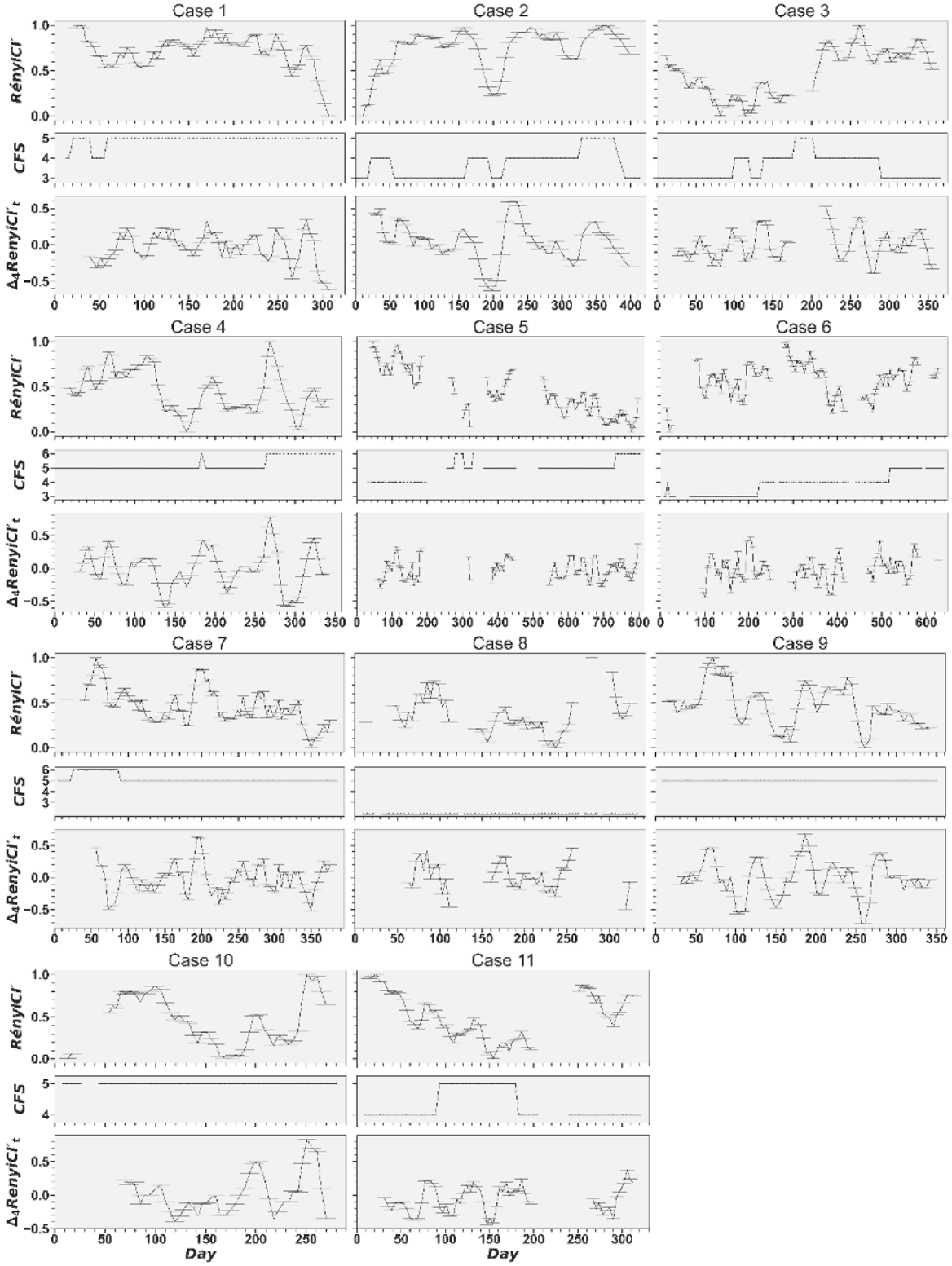
Results

Distributional Characteristics and Statistical Testing

Figure 3 presents time series plots of normalized RényiCI, frailty scores, and first-order differences in normalized RényiCI for each case. In the plots, time is measured in observation days. Summary statistics of overall RényiCI, KS, and Spearman rank correlation values are reported in Table 3. The shape of the RényiCI distributions varied notably across cases. Case 9 exhibited the only unimodal, symmetric distribution (Figure 4), while the remaining cases showed skewness or kurtosis. Cases 1 and 2 were strongly left-skewed, while cases 3, 5, 10, and 11 displayed low kurtosis. Cases 3 and 11 also showed bimodal distributions. Despite this heterogeneity, only case 2 showed a statistically significant deviation from a random normal distribution (P=.006) based on the KS test of normality.

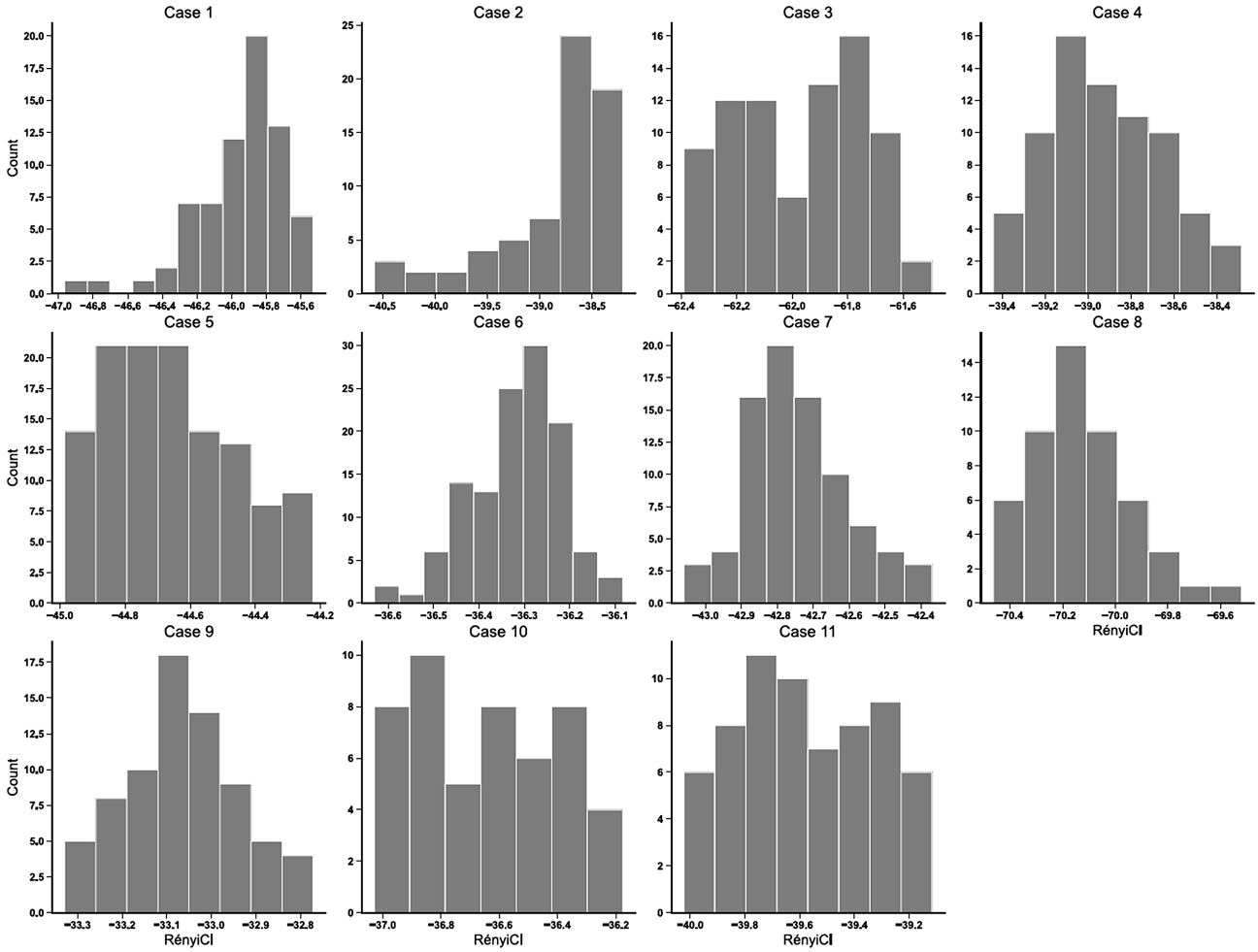
XSL•FO
RenderX

Figure 3. Plots of complexity, frailty, and complexity change as a function of time. Lower values reflect less complexity. Δ_4RényiCI' represents the value difference between sliding windows at times *t* and *t-4*. CFS: Clinical Frailty Scale; RényiCI: Rényi Complexity Index.



XSL•FO
RenderX

Figure 4. Histograms of RényiCI values.



Overall, case participants ranged from "fit" (CFS=2) to "living with moderate frailty" (CFS=6), although the trajectories of frailty within each participant varied considerably (Figure 3). For example, cases 2, 3, 7, and 11 all experienced periods of elevated frailty but all recovered and returned to baseline by the end of observation. Cases 4, 5, and 6 were the only cases that increased in frailty over time, ending frailer than their baseline. Most cases exhibited 2 - 3 transitions in frailty over time, with the extreme being case 2 with 7 frailty transitions.

While the goal of this study is not to directly infer CFS from RényiCI values, we note that Table 3 shows a significant correlation for all cases that have variable frailty scores. While they are significant, the correlations are mostly quite small. The overall correlation for all combined values is ρ=−.055 (P<.001). These results indicate that while a relationship between behavioral complexity and frailty can be observed, other factors must be considered when assessing a person's frailty from smart home sensor readings.

As Figure 3 demonstrates, trajectories of sensor complexity were similarly varied. RényiCI values for cases 1, 5, 7, and 9 exhibited downward trends over time, while cases 3 and 6 demonstrated an overall positive trend. Case 11 showed a concave shape with a general downward trend in sensor RényiCI for the first half of the data, followed by a general upward trend. The remaining cases exhibited nonmonotonic fluctuations. In each case, RényiCI values and frailty trajectories aligned with frailty scores assigned by the CFS-trained researchers during qualitative processing of clinical data.

Frailty-Complexity Associations

The relationship between frailty level and sensor data complexity also varied from case to case. Figure 5 shows jitter plots of RényiCI values by CFS score grouped by the number of sensors installed in the home. Cases 5 and 11 exhibited a negative trend between complexity and frailty, while cases 3 and 7 demonstrated a mostly positive trend. The range and mean of USM-Rényi values shift farther from 0 as the number of sensors in the home increases. The range of RényiCI for the home with the fewest sensors (11 sensors) spanned approximately −34 to −32, while in the home with the most sensors (22 sensors), the range extended from about −71 to −69.

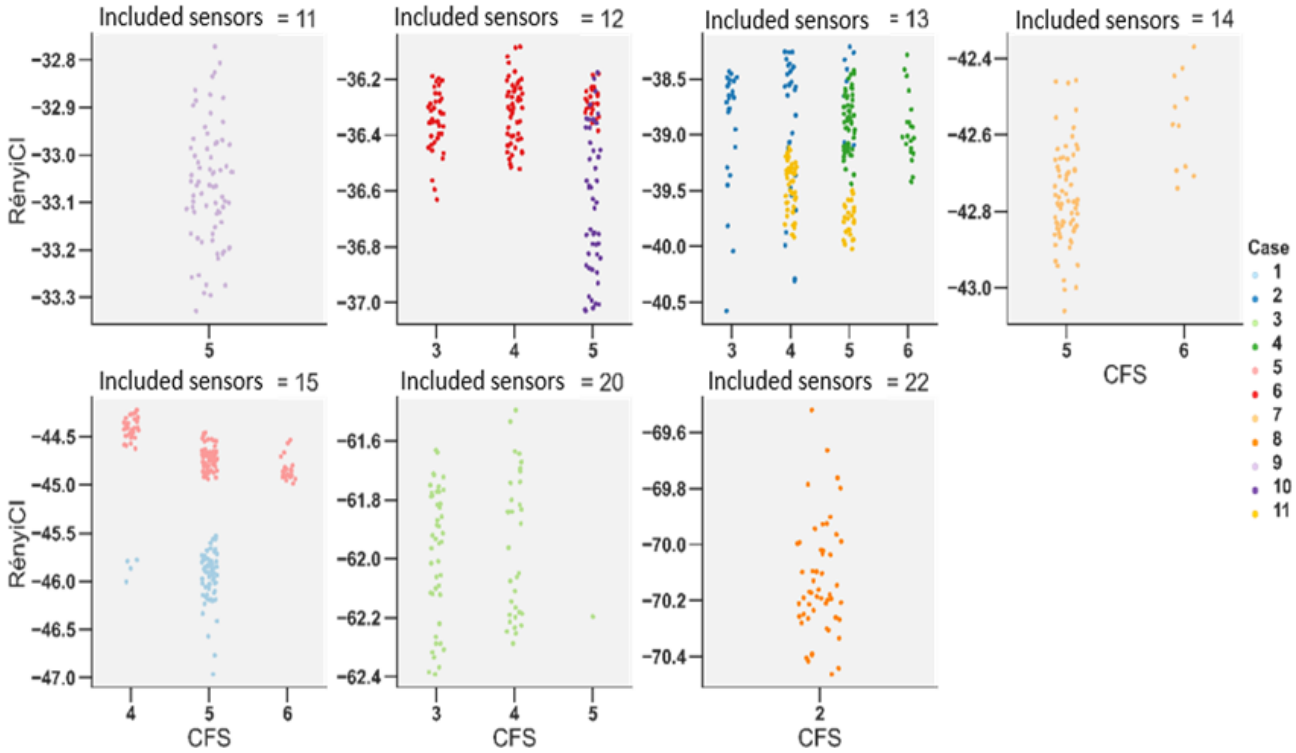
Initial runs tests applied to the full sequence of RényiCI values were statistically significant (P<.001) for all participants, suggesting nonrandom temporal ordering. To reduce potential autocorrelation introduced by overlapping windows, we repeated the runs tests on a downsampled sequence using every fourth window. Under this condition, only cases 3 (P<.001) and 5 (P=.002) remained statistically significant.

To account for these differences, Table 3 also includes the coefficient of variation (CoV) that normalizes the RényiCI SD by the mean for each case. Cases 6, 7, and 8 exhibited the least

XSL•FO
RenderX

amount of relative variability in RényiCI (CoV=0.003), while case 2 exhibited the highest (CoV=0.015).

Figure 5. Categorical scatter (jitter) plot of RényiCI values by frailty, grouped by sensor count and case. CFS: Clinical Frailty Scale.



The cases with statistically significant nonoverlapping runs tests, cases 3 and 5, also represented the clearest long-term monotonic trends. Case 3 demonstrated an overall increase in complexity between the start and end of her data, while case 5 demonstrated an overall decreasing trend. The only RényiCI distribution with a statistically significant KS test, case 2, had some of the most extreme variation among the cases, with an extremely left-skewed distribution and a coefficient of variation 5 times greater than the smallest coefficient of variation among the cases. To explore possible explanations for the diverse patterns of frailty and complexity trajectories observed, we compared the frailty and complexity trajectories of the cases with contextual information derived from the nursing assessment records.

Data selection for cases 2 and 3 started at the beginning of the COVID-19 pandemic shutdowns. The horizontal bars in Figure 3 represent sliding window durations (in days). For both cases, the shortest windows are at the beginning of the time series. The shorter sliding window durations during this period likely reflect increased in-home activity, with participants generating more sensor events due to spending more time indoors during COVID-19 lockdowns. However, where RényiCI increased steadily over the coming months for case 2, RényiCI decreased steadily for case 3 (Figure 3).

Cases 2 and 3 also showed pronounced shifts in RényiCI midway through the observation period. For both cases, this period roughly correlates to a time period of hazardous air quality caused by continued wildfire smoke over the course of about a month. However, the steep dip in RényiCI for case 2 occurring between days 186 and 212 is short, while for case 3, the sharp increase in RényiCI around day 231 appears to be a vertical shift in their average complexity that continues for the rest of her data.

Case Narratives

Case 2

A woman in her 70s with congestive heart failure and mild asthma. She was independent at baseline (CFS=3) and in stable health, which is consistent with the increasing RényiCI values plotted in Figure 3 at the beginning of data collection. This participant experienced 3 episodes of worsening fatigue and shortness of breath (days 20 - 54, 160 - 215, 258 - 397), which impacted her ADLs and IADLs and contributed to transient increases in frailty. During the first episode, the nurse's report indicated that "walking has been much more taxing on her this past week. She will walk around the building today but runs out of energy very quickly... her fatigue level has increased significantly over the past week." During the second episode, the nurse reported "the last 3 days she noticed … more soa [sic, shortness of air] and tired[ness]." During the third episode, the nurse recorded a direct quote from the participant indicating that she had "absolute fatigue beyond anything I've experienced." Two of these periods coincided with substantial troughs in RényiCI, suggesting alignment between behavioral simplification and functional decline. The primary movement patterns manifested in the CFS score and measured by RényiCI (ie, mechanisms of interest) were less in-home overall activity and less time spent out of the home on walks and social activities (frequency and duration). This case also had the highest variability in complexity and was the only one with a statistically significant KS test result. Possible gerontological clinical actions informed by these results include recommendations to follow up with the cardiologist, referral to a pulmonologist, and referral

XSL•FO
RenderX

to senior services to determine whether the patient qualifies for in-home care support services.

Case 3

A woman in her 80s with cardiovascular disease and allergy-induced asthma. Initially independent, she experienced progressive health decline, including 2 hospitalizations for acute hypertension and dyspnea. Her frailty peaked after the second hospitalization during a period of wildfire smoke, when she relied full-time on a walker for ambulation. Nursing records include statements during this timeframe like "no energy, has not left the house since Thursday [4 d]" and "overall health is declining." Following physical therapy, she recovered and reported no activity limitations by day 295. The nursing record indicated that she "went shopping" and had "several visitors over" across multiple days of the week. Her RényiCI trajectory reflected this pattern. As shown in Figure 3, the RényiCI values initially show a steady decrease aligned with the health issues. After she received treatment and improved her ambulation and functional independence, the RényiCI values showed a steady rise in complexity. Notably, this case showed a strong monotonic increase in RényiCI and passed the runs test even under downsampling. This result provides evidence that the pattern of increasing frailty, followed by improvement after treatment, is distinct and nonrandom. The mechanism of interest impacting her RényiCI trajectory was a renewed increase in time spent out of the home (frequency and duration) concurrent with an increase in the number of visitors. The case exemplifies how RényiCI trajectories could help clinical gerontologists understand treatment efficacy through novel remote patient monitoring tools that include sensor monitoring and associated behavior patterns.

Case 5

A man in his 80s with Parkinson disease. He began with mild frailty (CFS=4) and was independent but slowed by symptoms. Over time, he required increasing assistance with ADLs and IADLs. Nurses recorded that he began to require assistance "getting compression sock on in the morning and off at night" and "needing help with laundry and housekeeping" and that his daughter began assisting with bill paying. He experienced multiple hospitalizations and rehabilitation stays and ultimately progressed to moderate frailty (CFS=6). The moderate frailty score was based on the nurse reporting "unsteady gait" and that he "has cracked ribs from a fall last week" and his "symptoms of PD [are] increasing, [and] noticeable upon observation." His RényiCI trajectory followed a corresponding decline, with complexity peaking early and then falling across successive rehabilitation episodes. This case also exhibited a significant runs test and a clear downward trend in complexity. The mechanism of interest in this case was more overall time spent in his recliner chair, more nighttime sleeping in the recliner chair, and the decreased time spent out of the home (frequency and duration). This case illustrates how RényiCI trajectories may support automated smart home monitoring aimed at detecting increasing frailty upstream so interventions can be implemented.

Cases 4 and 10

Both cases involved sustained or increased caregiving over time. In case 4, RényiCI peaked just before caregiving began and declined thereafter. Case 10, who had consistent caregiving throughout, showed generally lower complexity than case 6, who lived alone with the same number of sensors. These comparisons suggest that increased caregiving frequency does not necessarily lead to increased behavioral complexity as measured by RényiCI. Older adults with consistent professional caregiving are likely to experience slower rates of decline due to the intentionality of caregiving, which aims to extend independence through building physical, functional, and cognitive strength. Findings could inform care planning and resource allocation.

Discussion

Principal Results

This study introduces and applies a novel entropy-based algorithm, the RényiCI, to quantify behavioral complexity from smart home sensor data in older adults. Using a USM framework with multiscale Rényi entropy, our method captures subtle temporal dynamics in sensor-derived movement sequences. In this exploratory case series, within-person indoor-movement complexity, as exhibited by RényiCI values, fluctuated over time. In several cases, these fluctuations coincided with frailty changes.

Across 11 participants, we observed diverse complexity trajectories, ranging from steady increases, steady declines, and concave patterns to nonmonotonic fluctuations. Case-level analysis revealed that greater fluctuations in complexity were frequently aligned with periods of functional decline or recovery. Notably, 2 cases (3 and 5) exhibited statistically significant nonrandom patterns in complexity over time, confirmed by runs tests on downsampled data, and showed clear monotonic trends in behavior complexity that matched health trajectories. Only one case (2) showed a RényiCI distribution that deviated significantly from normality, corresponding with extreme within-person variability and periods of worsening frailty. Scatter plots further revealed heterogeneous associations between complexity and frailty, with both positive and negative trends across cases. Importantly, increased caregiver presence was not associated with greater behavioral complexity, suggesting that RényiCI may reflect intrinsic changes in individual functional capacity rather than external support.

Changes in CFS scores fluctuated in alignment with changes in RényiCI values for some cases, like 2, 3, 6, and 7. These cases may suggest that changes in frailty do impact the person's behavioral routine and regularity. However, in cases 8 through 10, we observed changes in RényiCI values despite the lack of change in frailty scores. This observation highlights the fact that our findings provide 1 set of indicators of changes in frailty, but should not be analyzed in isolation. Other factors, such as visitors, seasonal effects, and external events, can also impact behavioral routines. These should be controlled for when examining frailty as a function of changes in RényiCI.

XSL•FO
RenderX

While observed repetitive behavior may correlate with frailty, the relationship is not one-to-one. Reduced complexity of movement often, but not always, aligns with frailty progression. Repetitive behavior can signal frailty because it reflects narrowed activity routines, reduced introduction of new routine elements, and corresponding reduced adaptability. At the same time, we note that complexity is multifactorial. Other influences, such as visitors in the home, home layout, and external events, also affected the entropy measures. The results showed nonmonotonic relationships in those cases. To interpret complexity, it is therefore best to consider an individual over time rather than compare cross-sectionally. Moreover, interventions aimed at slowing the impact of frailty on maintaining independence, like a smart home that projects RényiCI trajectories, would be more helpful for older adults living alone. Mechanisms of interest become difficult to automatically recognize in multiresident homes where ambient sensors detect movement from all residents.

Limitations

Several factors impacted the interpretation and generalizability of our findings. First, entropy-based measures like RényiCI are inherently sensitive to sample length and the number of sensors deployed in a participant's home. To prioritize intraindividual validity, we customized the sliding window size for each participant using a fixed number of sensor messages (n), rather than a fixed time duration. This approach allowed for consistent comparisons within individuals but introduced variability in the time span covered by each window, both within and across cases, limiting our ability to analyze complexity as a direct function of chronological time. Future work could develop correction factors for RényiCI to account for sample length, enabling the detection of periodic, seasonal, or event-driven patterns in indoor behavior.

Relatedly, interindividual comparisons were constrained by differences in sensor configurations across homes. Participants varied in the number and placement of sensors, affecting both the density of event data and the scale of RényiCI values. Standardizing sensor deployments in future studies would facilitate more robust cross-participant comparisons and support investigation into whether home-level sensor complexity systematically relates to frailty markers at the population level.

The impact of the COVID-19 pandemic further complicates interpretation. Several participants were enrolled during or shortly after the onset of pandemic-related lockdowns, which led to changes in daily routines, increased time spent indoors, and potentially long-term shifts in behavior and social support. These behavioral changes may have altered both the complexity of movement and its relationship to frailty. Additionally, one period of the study coincided with prolonged hazardous air quality due to regional wildfires, which may have further restricted participants' movement and contributed to abrupt changes in sensor complexity. Such exogenous events likely altered daily routines independent of health. We therefore interpret Rényi changes within homes and in the presence of annotated event periods. We also provide event-excluded sensitivities to reduce confounding.

Sensor noise and dropout also presented challenges. While preprocessing steps excluded known periods of sensor failure or participant absence, subtle forms of sensor drift or inconsistent message delivery could still introduce noise into the RényiCI estimates. Further improvements to sensor reliability and the integration of sensor health metrics into complexity analysis pipelines would strengthen future research.

In terms of statistical methods, the runs test was useful in identifying nonrandom patterns in behavioral complexity over time, but it is not well-suited to detecting more complex temporal structures such as oscillatory or nonlinear trends. Future research may benefit from time series models drawn from signal processing or machine learning that can more precisely characterize evolving behavioral dynamics.

Frailty measurement also posed a limitation. Because frailty was not a primary outcome in the parent study, we relied on retrospective CFS scoring based on weekly nursing reports. This limits temporal precision and may miss subtle fluctuations in functional status. Larger-scale studies using prospectively collected frailty data, including both clinician-reported and self-reported measures, could reveal more detailed associations between complexity and health.

Finally, this sample was racially and culturally homogeneous, limiting the generalizability of our findings. RényiCI analyses should be interpreted as a within-home monitoring signal rather than a cross-sectional diagnostic tool. As efforts to diversify smart home research populations expand, it will be essential to explore whether the relationships between sensor-derived behavioral complexity and frailty differ across racial, cultural, and socioeconomic groups. Inclusive, representative samples are critical to ensuring that digital biomarkers are both effective and equitable.

Conclusions

Detection of incipient frailty in community-dwelling older adults is a key component to supporting their independence. The findings in this study demonstrate that RényiCI, as a passive and unobtrusive complexity metric, offers a promising tool for monitoring functional health changes in aging populations and may help enable early detection of frailty in real-world settings. The PyUSM software package developed for this analysis is publicly available and supports future application of this method in diverse behavioral monitoring contexts. These findings support the potential of entropy-based digital behavior markers to unobtrusively monitor intraindividual health changes and capture early signs of frailty in aging-in-place.

Future enhancements of this analysis may reveal additional factors that influence change in indoor movement complexity and inform how the complexity of smart home data may inform clinical practice. For example, significant departures of RényiCI values from a person's complexity baseline may trigger a nurse call or follow-up when integrated into a remote monitoring or telemonitoring system. In routine care, weekly summaries of the analysis would support triage and help care providers select appropriate actions. Additionally, when performing a functional assessment of an individual, a summary of the complexity trend augments traditional frailty analysis to improve assessment and

XSL•FO
RenderX

treatment options. Future work will also emphasize analytical validity (repeatability and robustness across sensors and windowing), clinical validity (prospective prediction of frailty transitions), and clinical use (impact on downstream outcomes such as unplanned care, falls, and functional decline) for diverse homes and populations.

Additionally, future work should focus on integrating RényiCI in machine learning predictive modeling as a high-level feature to assist with identifying meaningful digital biomarkers [49]. Other temporal activities associated with frailty (eg, walking speed, ADL, and IADL behaviors) could also be integrated to optimize frailty classifications. Machine learning integration of features from RényiCI values that signal possible increasing frailty will support nurses and caregivers in providing timely interventions, thereby potentially extending independence and optimizing older adults' outcomes.

Acknowledgments

The authors would like to thank Jason Minor, Bryan Minor, and Brian Thomas for their assistance with the smart homes and Shandeigh Nikki Berry, Samantha Denison, and Ellen Hinderlie for their data collection contributions.

Funding

This work is supported in part by the National Institute of Nursing Research under grant R01NR016732, Touchmark Foundation, and the Washington State University College of Nursing's Linblad Scholarship funds.

Data Availability

Sensor data are available at CASAS DataDownloader [50]. Deidentified medical information about frailty cases included in this study may be made available for authorized research use upon request.

Conflicts of Interest

None declared.

References

1. Kim DH, Rockwood K. Frailty in older adults. N Engl J Med 2024 Aug 8;391(6):538-548. [doi: 10.1056/NEJMra2301292] [Medline: 39115063]
2. Kurnat-Thoma EL, Murray MT, Juneau P. Frailty and determinants of health among older adults in the United States 2011-2016. J Aging Health 2022 Mar;34(2):233-244. [doi: 10.1177/08982643211040706] [Medline: 34470533]
3. Jang J, Kim A, Choi M, et al. Association of Frailty Index at 66 years of age with health care costs and utilization over 10 years in Korea: retrospective cohort study. JMIR Public Health Surveill 2025 Jan 27;11:e50026. [doi: 10.2196/50026] [Medline: 39874179]
4. Ensrud KE, Schousboe JT, Kats AM, Taylor BC, Boyd CM, Langsetmo L. Incremental health care costs of self-reported functional impairments and phenotypic frailty in community-dwelling older adults: a prospective cohort study. Ann Intern Med 2023 Apr;176(4):463-471. [doi: 10.7326/M22-2626] [Medline: 37011386]
5. Crocker TF, Brown L, Clegg A, et al. Quality of life is substantially worse for community-dwelling older people living with frailty: systematic review and meta-analysis. Qual Life Res 2019 Aug;28(8):2041-2056. [doi: 10.1007/s11136-019-02149-1] [Medline: 30875008]
6. Lin YC, Yan HT. Frailty phenotypes and their association with health consequences: a comparison of different measures. Aging Clin Exp Res 2024 Dec 3;36(1):233. [doi: 10.1007/s40520-024-02887-4] [Medline: 39625598]
7. Fletcher J, Reid N, Hubbard RE, et al. Frailty Index, not age, predicts treatment outcomes and adverse events for older adults with cancer. J Frailty Aging 2024;13(4):487-494. [doi: 10.14283/jfa.2024.22] [Medline: 39574272]
8. Yixiao C, Hui S, Quhong S, Xiaoxi Z, Jirong Y. A review of utility of wearable sensor technologies for older person frailty assessment. Exp Gerontol 2025 Feb;200:112668. [doi: 10.1016/j.exger.2024.112668] [Medline: 39733783]
9. Woo J, Yu R, Tsoi K, Meng H. Variability in repeated blood pressure measurements as a marker of frailty. J Nutr Health Aging 2018;22(9):1122-1127. [doi: 10.1007/s12603-018-1082-9] [Medline: 30379313]
10. Fried LP, Cohen AA, Xue QL, Walston J, Bandeen-Roche K, Varadhan R. The physical frailty syndrome as a transition from homeostatic symphony to cacophony. Nat Aging 2021 Jan;1(1):36-46. [doi: 10.1038/s43587-020-00017-z] [Medline: 34476409]
11. Coravos A, Khozin S, Mandl KD. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. NPJ Digit Med 2019;2(1):1-5. [doi: 10.1038/s41746-019-0090-4] [Medline: 30868107]
12. Beattie Z, Miller LM, Almirola C, et al. The collaborative aging research using technology initiative: an open, sharable, technology-agnostic platform for the research community. Digit Biomark 2020;4(Suppl 1):100-118. [doi: 10.1159/000512208] [Medline: 33442584]
13. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-12-11]

XSL•FO
RenderX

14. Cobo A, Villalba-Mora E, Pérez-Rodríguez R, Ferre X, Rodríguez-Mañas L. Unobtrusive sensors for the assessment of older adult's frailty: a scoping review. Sensors (Basel) 2021 Apr 23;21(9):2983. [doi: 10.3390/s21092983] [Medline: 33922852]
15. Bian C, Ye B, Hoonakker A, Mihailidis A. Attitudes and perspectives of older adults on technologies for assessing frailty in home settings: a focus group study. BMC Geriatr 2021 May 8;21(1):298. [doi: 10.1186/s12877-021-02252-4] [Medline: 33964887]
16. Cook DJ, Crandall AS, Thomas BL, Krishnan NC. CASAS: a smart home in a box. Computer (Long Beach Calif) 2013 Jul;46(7):62-69. [doi: 10.1109/MC.2012.328] [Medline: 24415794]
17. Lipsitz LA. Dynamics of stability: the physiologic basis of functional health and frailty. J Gerontol A Biol Sci Med Sci 2002 Mar;57(3):B115-B125. [doi: 10.1093/gerona/57.3.b115] [Medline: 11867648]
18. Hao M, Zhang H, Li Y, et al. Using physiological system networks to elaborate resilience across frailty states. J Gerontol A Biol Sci Med Sci 2024 Oct 1;79(10):1-7. [doi: 10.1093/gerona/glad243] [Medline: 37824088]
19. Cohen AA, Ferrucci L, Fülöp T, et al. A complex systems approach to aging biology. Nat Aging 2022 Jul;2(7):580-591. [doi: 10.1038/s43587-022-00252-6] [Medline: 37117782]
20. Bizovska L, Svoboda Z, Vuillerme N, Janura M. Multiscale and Shannon entropies during gait as fall risk predictors-a prospective study. Gait Posture 2017 Feb;52:5-10. [doi: 10.1016/j.gaitpost.2016.11.009] [Medline: 27842283]
21. Castiglia SF, Trabassi D, Conte C, et al. Multiscale entropy algorithms to analyze complexity and variability of trunk accelerations time series in subjects with Parkinson's disease. Sensors (Basel) 2023 May 22;23(10):4983. [doi: 10.3390/s23104983] [Medline: 37430896]
22. Gao C, Lim ASP, Haghayegh S, et al. Reduced complexity of pulse rate is associated with faster cognitive decline in older adults. J Am Heart Assoc 2025 May 20;14(10):e041448. [doi: 10.1161/JAHA.125.041448] [Medline: 40331928]
23. Ma Y, Zhou J, Kavousi M, et al. Lower complexity and higher variability in beat‐to‐beat systolic blood pressure are associated with elevated long‐term risk of dementia: the Rotterdam study. Alzheimer's & Dementia 2021 Jul;17(7):1134-1144. [doi: 10.1002/alz.12288] [Medline: 33860609]
24. Zhou J, Habtemariam D, Iloputaife I, Lipsitz LA, Manor B. The complexity of standing postural sway associates with future falls in community-dwelling older adults: the MOBILIZE Boston study. Sci Rep 2017 Jun 7;7(1):2924. [doi: 10.1038/s41598-017-03422-4] [Medline: 28592844]
25. Zhou J, Poole V, Wooten T, et al. Multiscale dynamics of spontaneous brain activity is associated with walking speed in older adults. J Gerontol A Biol Sci Med Sci 2020 Jul 13;75(8):1566-1571. [doi: 10.1093/gerona/glz231] [Medline: 31585008]
26. Rector JL, Gijzel SMW, van de Leemput IA, van Meulen FB, Olde Rikkert MGM, Melis RJF. Dynamical indicators of resilience from physiological time series in geriatric inpatients: lessons learned. Exp Gerontol 2021 Jul 1;149:111341. [doi: 10.1016/j.exger.2021.111341] [Medline: 33838217]
27. Raichlen DA, Klimentidis YC, Hsu CH, Alexander GE. Fractal complexity of daily physical activity patterns differs with age over the life span and is associated with mortality in older adults. J Gerontol A Biol Sci Med Sci 2019 Aug 16;74(9):1461-1467. [doi: 10.1093/gerona/gly247] [Medline: 30371743]
28. Schütz N, Knobel SEJ, Botros A, et al. A systems approach towards remote health-monitoring in older adults: introducing a zero-interaction digital exhaust. NPJ Digit Med 2022 Aug 16;5(1):116. [doi: 10.1038/s41746-022-00657-y] [Medline: 35974156]
29. Cook DJ, Schmitter-Edgecombe M. Fusing ambient and mobile sensor features into a behaviorome for predicting clinical health scores. IEEE Access 2021;9:65033-65043. [doi: 10.1109/access.2021.3076362] [Medline: 34017671]
30. Wang T, Cook DJ, Fischer TR. The indoor predictability of human mobility: estimating mobility with smart home sensors. IEEE Trans Emerg Top Comput 2023;11(1):182-193. [doi: 10.1109/tetc.2022.3188939] [Medline: 37457914]
31. Howedi A, Lotfi A, Pourabdollah A. Employing entropy measures to identify visitors in multi-occupancy environments. J Ambient Intell Human Comput 2022 Feb;13(2):1093-1106. [doi: 10.1007/s12652-020-02824-z]
32. Takahashi J, Kawai H, Ejiri M, et al. Activity diversity is associated with the prevention of frailty in community-dwelling older adults: the Otassha study. Front Public Health 2023;11:1113255. [doi: 10.3389/fpubh.2023.1113255] [Medline: 37033071]
33. Gopalratnam K, Cook D. Online sequential prediction via incremental parsing: the active LeZi algorithm. IEEE Intell Syst 2007;22(1):52-58. [doi: 10.1109/MIS.2007.15]
34. Bai C, Mardini MT. Advances of artificial intelligence in predicting frailty using real-world data: a scoping review. Ageing Res Rev 2024 Nov;101:102529. [doi: 10.1016/j.arr.2024.102529] [Medline: 39369796]
35. Park C, Mishra R, Golledge J, Najafi B. Digital biomarkers of physical frailty and frailty phenotypes using sensor-based physical activity and machine learning. Sensors (Basel) 2021 Aug 5;21(16):5289. [doi: 10.3390/s21165289] [Medline: 34450734]
36. Osuka Y, Chan LLY, Brodie MA, Okubo Y, Lord SR. A wrist-worn wearable device can identify frailty in middle-aged and older adults: the UK Biobank study. J Am Med Dir Assoc 2024 Oct;25(10):105196. [doi: 10.1016/j.jamda.2024.105196] [Medline: 39128825]
37. Hosseinalizadeh M, Asghari M, Toosizadeh N. Sensor-based frailty assessment using fitbit. Sensors (Basel) 2024 Dec 7;24(23):7827. [doi: 10.3390/s24237827] [Medline: 39686364]

XSL•FO
RenderX

38. Merchant RA, Loke B, Chan YH. Ability of heart rate recovery and gait kinetics in a single wearable to predict frailty: quasiexperimental pilot study. JMIR Form Res 2024 Oct 3;8:e58110. [doi: 10.2196/58110] [Medline: 39361400]
39. Ding H, Ho K, Searls E, et al. Assessment of wearable device adherence for monitoring physical activity in older adults: pilot cohort study. JMIR Aging 2024 Oct 25;7:e60209. [doi: 10.2196/60209] [Medline: 39454101]
40. Tannou T, Lihoreau T, Gagnon-Roy M, Grondin M, Bier N. Effectiveness of smart living environments to support older adults to age in place in their community: an umbrella review protocol. BMJ Open 2022 Jan 25;12(1):e054235. [doi: 10.1136/bmjopen-2021-054235] [Medline: 35078843]
41. Fritz R, Wuestney K, Dermody G, Cook DJ. Nurse-in-the-loop smart home detection of health events associated with diagnosed chronic conditions: a case-event series. Int J Nurs Stud Adv 2022 Dec;4:100081. [doi: 10.1016/j.ijnsa.2022.100081] [Medline: 35642184]
42. Wang Y, Yalcin A, VandeWeerd C. An entropy-based approach to the study of human mobility and behavior in private homes. PLOS ONE 2020;15(12):e0243503. [doi: 10.1371/journal.pone.0243503] [Medline: 33301515]
43. CASAS datasets. Center for Advanced Studies in Adaptive Systems (CASAS). URL: https://casas.wsu.edu/datasets/ [accessed 2025-11-04]
44. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005 Aug 30;173(5):489-495. [doi: 10.1503/cmaj.050051] [Medline: 16129869]
45. Tamamura Y, Hachiuma C, Matsuura M, Shiba S, Nishikimi T. Frailty and energy intake deficiency reduce the efficiency of activities of daily living in patients with musculoskeletal disorders: a retrospective cohort study. Nutrients 2025 Apr 12;17(8):1334. [doi: 10.3390/nu17081334] [Medline: 40284199]
46. Almeida JS, Vinga S. Universal sequence map (USM) of arbitrary discrete sequences. BMC Bioinformatics 2002;3(1):6. [doi: 10.1186/1471-2105-3-6] [Medline: 11895567]
47. Vinga S, Almeida JS. Rényi continuous entropy of DNA sequences. J Theor Biol 2004 Dec 7;231(3):377-388. [doi: 10.1016/j.jtbi.2004.06.030] [Medline: 15501469]
48. Wuestney K. Demo of Universal Sequence Maps (USM) in Python. Documentation of pyusm Package. 2022. URL: https://katherine983.github.io/pyusm/demo_usm.html [accessed 2025-09-25]
49. Kong LN, Yang L, Lyu Q, Liu DX, Yang J. Risk prediction models for frailty in older adults: a systematic review and critical appraisal. Int J Nurs Stud 2025 Jul;167:105068. [doi: 10.1016/j.ijnurstu.2025.105068] [Medline: 40184783]
50. CASAS DataDownloader. URL: https://data.casas.wsu.edu/ [accessed 2025-12-23]

Abbreviations

ADL: activity of daily living
CASAS: Center for Advanced Studies in Adaptive Systems
CFS: Clinical Frailty Scale
CoV: coefficient of variance
IADL: instrumental activity of daily living
KS: Kolmogorov-Smirnov
MSE: multiscale entropy
RényiCI: Rényi Complexity Index
USM: Universal Sequence Map

Edited by R Yang; submitted 11.May.2025; peer-reviewed by J Wang, R Buenrostro; revised version received 22.Nov.2025; accepted 27.Nov.2025; published 02.Jan.2026.

Please cite as:
Wuestney K, Cook D, Van Son C, Fritz R
Using Indoor Movement Complexity in Smart Homes to Detect Frailty in Older Adults: Multiple-Methods Case Series Study
JMIR Aging 2026;9:e77322
URL: https://aging.jmir.org/2026/1/e77322
doi:10.2196/77322

© Katherine Wuestney, Diane Cook, Catherine Van Son, Roschelle Fritz. Originally published in JMIR Aging (https://aging.jmir.org), 2.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Early Identification of Mobility Limitations in Community-Dwelling Middle-Aged and Older Adults: Development of a Prediction Model Based on a Prospective Cohort

Alberto Conde Freniche[1*], MSc, MVetMed; Wei Hu[2*], MD; Mo Chen[1], MSc; Dantong Wang[1], PhD; Denis Breuillé[1], PhD; Yu-ming Chen[2], PhD

[1]Nestlé Institute of Health Sciences, Nestlé Research, Société des Produits Nestlé, Lausanne, Switzerland

[2]Department of Epidemiology, School of Public Health, Sun Yat-sen University, No. 135, Xingang Xi Road, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Yu-ming Chen, PhD
Department of Epidemiology, School of Public Health, Sun Yat-sen University, No. 135, Xingang Xi Road, Guangzhou, China

Abstract

Background: With the aging of the global population, preventing the onset of mobility limitations is considered a worldwide public health priority.

Objective: This study aimed to develop a predictive model for incident early mobility limitations (EMLs) in late middle-aged and older adults, based on a simple functional test and modifiable lifestyle factors to facilitate a home-based self-assessment of early mobility decline and promote lifestyle intervention strategies.

Methods: Our study population was community dwellers aged 45 years and above who participated in the second and fourth waves of the Guangzhou Nutrition and Health Study. The included participants were healthy, nonfrail adults reporting no limitations in activities of daily living at baseline. At the 6-year follow-up, participants with poor physical performance (walking speed <1 m/s or handgrip strength <28 kg for males and <18 kg for females) or reporting some difficulty in walking and/or climbing stairs were classified as experiencing EMLs. Least absolute shrinkage and selection operator (LASSO) was used to identify predictors from various factors, and 6 machine learning models were trained and evaluated for EML prediction, using bootstrap-based techniques to address class imbalance. Predictive ability was quantified using the area under the receiver operating characteristic curve (AUC). Variable importance analysis was used to identify key predictors.

Results: A total of 1344 participants were included in the analysis, of which 206 (15.33%) developed EMLs after a median follow-up of 6.67 (IQR 5.91 - 7.26) years. Those who developed EMLs were older, had a higher BMI, a lower alternate Mediterranean diet score, and poorer performance in the sit-to-stand (STS) test, as well as lower estimated muscle power from the STS at baseline. The final models included 6 out of 9 predictors: age, sex, BMI, alternate Mediterranean diet score, STS power, and dietary calcium intake. Four machine learning models (logistic regression, LASSO, support vector machine, and neural network) achieved an acceptable AUC value (≥0.70) or slightly below in the testing dataset, with the neural network performing the best (AUC 0.70, 95% CI 0.63 - 0.77). Curve analysis showed positive net benefits for LASSO and logistic regression. Bootstrapping did not improve classification performance. Advanced age, lower adherence to Mediterranean diet, lower muscle power estimated from STS test, and higher BMI at baseline emerged as the most important predictors of EMLs.

Conclusions: This study shows that EMLs in Chinese adults aged 45 years and older can be predicted using easy-to-obtain physical performance measures, age, sex, BMI, and specific nutritional factors. The combination of prediction models and variable importance analysis provides valuable insights for the early identification and intervention of EMLs. More efforts are needed to validate our findings in external cohorts.

Trial Registration: ClinicalTrials.gov NCT03179657; https://clinicaltrials.gov/study/NCT03179657

International Registered Report Identifier (IRRID): RR2-10.2188/jea.JE20230108

(JMIR Aging 2026;9:e77187) doi:10.2196/77187

KEYWORDS

prediction model; prognostic; China; self-assessment; mobility; mobility limitations; sarcopenia; frailty; diet; machine learning; healthy aging

XSL•FO
RenderX

Introduction

Maintaining optimal locomotive capacity—in this study referred to as "mobility"—is essential to successful aging. Optimal mobility determines the capacity to perform activities of daily living (ADLs) and engage in social and leisure activities. However, aging has a negative impact on mobility, underpinned by multiple mechanisms affecting the musculoskeletal, cardiorespiratory, and neurological systems. Along with the onset of chronic conditions, the aging process increases the likelihood of developing mobility limitations, such as difficulties in running, walking, or climbing stairs. These limitations often cause older adults to lose independence, reducing their quality of life and dramatically increasing health care costs. Therefore, with the aging of the global population, preventing the onset of mobility limitations is considered a worldwide public health priority [1,2].

Mobility limitations are often the first signs of further functional decline. The interplay between the emergence of mobility limitations, sarcopenia, and physical frailty involves intricate relationships. Sarcopenia, defined as "a progressive and generalized loss of skeletal muscle mass and strength" [3,4], requires the presence of at least 1 mobility limitation for its diagnosis, assessed through objective physical performance tests. In turn, sarcopenia is an important component of the physical frailty syndrome [5], where mobility limitations make up 2 out of 5 total diagnostic criteria, together with weight loss, low energy expenditure, and exhaustion [6]. The development of mobility limitations is therefore a critical component at the intersection of sarcopenia and physical frailty, serving as a primary target for interventions. In accordance, the World Health Organization has recently called health care professionals and gerontologists to focus on strategies for maximizing and preserving mobility across the entire life course independent of any specific disease.

The early stage of mobility decline is a window of opportunity for interventions, such as adequate nutrition intake and active exercise, to positively modify the trajectory in later life. Once physical frailty becomes manifest in older adults, it is increasingly challenging to reverse the development of mobility limitations [7]. Notably, the progressive age-related mobility decline may not clinically manifest at early stages because during the aging process, people unconsciously modify the way tasks are performed routinely to compensate for the accumulating physiological deficits [8]. For example, older adults typically increase the step width to preserve their balance while walking [9]. This strategy is maintained until the capacity reserve is exhausted and limitations become manifest. Therefore, to prevent mobility limitations, it is important to identify people who do not yet manifest them but who are at risk for their onset in the near future [10]. A recent scoping review and expert consensus supports the use of the term preclinical or early mobility limitation (EML), referring to such subtle and often unconscious changes in performance [11].

Models to predict EMLs in late middle-aged adults are scarce. Most of the models available were developed to predict mobility disability in older adults based on common physical performance measures, such as gait speed, balance tests, and handgrip strength [12-15]. However, these measures have poor sensitivity to detect mobility deficits in younger adults (under 65 y) [16,17]. Notably, measuring muscle power may increase the sensitivity of prognostic models compared to more commonly used muscle strength measures, such as handgrip strength. Muscle power declines at an earlier and faster rate than muscle strength and is likely to be a more discriminant predictor of mobility limitations [18,19]. Specifically, sit-to-stand (STS) transitions have been shown to decline about 2 decades earlier than handgrip strength and gait speed [16]. Therefore, it is possible that new predictive models built with power-based measures can improve their prognostic value among middle-aged adults. Recently, a formula to calculate muscle power from the STS test has been developed and validated, overcoming the difficulties of evaluating muscle power without any instrument [20,21].

Similarly, the importance of nutrition in preserving mobility has been studied [22,23]. Including nutrition elements in the prediction model would enable a holistic intervention approach to reduce the risk of mobility limitation. To date, few studies have investigated the association between muscle power-based physical performance measures or dietary factors and the development of EML among late middle-aged adults.

In the present study, we aimed to develop and internally validate a model to predict future preclinical mobility limitations in late middle-aged and older adults for patient self-assessment, based on simple–to-obtain measurements, including self-reported data and the STS test, using a cohort of healthy adults aged 45 years and above recruited in a community-based study in the south of China. This model could be used to identify which persons may benefit from preventive strategies at middle and older age to decelerate or delay the onset of mobility limitations in the Chinese population.

Methods

Guidelines for Model Development and Validation

We adhered to the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) guidelines, which provide transparent reporting standards for developing and validating prediction models for individual prognosis or diagnosis [24,25].

Study Design and Participants

This retrospective cohort study used data from the Guangzhou Nutrition and Health Study (GNHS), which is a longitudinal community-based cohort conducted by the School of Public Health, Sun Yat-sen University, Guangzhou, China. Between 2008 and 2013, a total of 4048 participants aged between 40 and 80 years, who lived in Guangzhou city for at least 5 years, were enrolled in the GNHS and have been followed up every 3 years approximately. Detailed information about the cohort protocol and baseline characteristics can be found elsewhere [26]. Briefly, data were collected from study visits from healthy participants, including physical performance tests, body composition measurements, dietary, and lifestyle factors. A validated semiquantitative Food Frequency Questionnaire was

XSL•FO
RenderX

used to collect dietary intake information [27]. Exclusion criteria such as myocardial infarction, stroke, diabetes, cancer, low-energy fractures, mental and physical disability, and chronic liver and kidney diseases were confirmed at the first enrollment.

Participants in the current analysis participated in 2 follow-ups of the GNHS (waves 2 and 4). At wave 2, from October 2013 to July 2017, 2970 participants were asked to take the 5-time STS test. Therefore, this study used wave 2 as the baseline. In addition, participants completed the ADL Scale questionnaire, grip strength test, and 4-minute walking test in wave 2 and subsequent follow-ups. The latest follow-up with available ADL and performance data was selected to allow EMLs to develop. At wave 4, from October 2020 to December 2023, 1952 participants had been followed-up.

Ethical Considerations

Written informed consent was obtained from all participants. The GNHS was approved by the Ethics Committee of the School of Public Health, Sun Yat-sen University (approval code: SYSU-GWYL2018048; approval date: July 17, 2018) and registered at ClinicalTrials.gov (NCT03179657). All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the World Medical Association Declaration of Helsinki. Data were anonymized and were accessible only to the research team. No financial compensation was provided.

Predicted Outcome

Mobility limitations were assessed using objective mobility tests and perceived difficulties by ADL questionnaire. Participants rated their ease at walking and climbing stairs as independent, have some difficulties, need help, and unable to mobilize. In the predictive model, EML was determined for those individuals meeting at least 1 of the following 3 criteria: (1) some difficulty in walking and/or climbing stairs; (2) poor handgrip strength (handgrip strength <28 kg for males and <18 kg for females); and (3) slow walking speed (<1 m/s).

Validated preclinical thresholds for middle-aged adults were lacking at the time of the study. Gait speed <1 m/s is a common threshold of slow walking speed among well-functioning older adults, and is often treated as a functional benchmark in middle-aged and older adults [28-31]. This threshold is recommended by the Asian Working Group on Sarcopenia (AWGS 2019) to screen for possible sarcopenia and has been applied to cohorts that include midlife participants across Asian countries, including China [32,33]. Similarly, for low handgrip strength, the threshold proposed by the AWGS 2019 was used in this study.

Baseline Candidate Predictors

The STS test is used to assess functional lower limb strength, transitional movements, balance, and fall risk in older adults. STS test score is based on the amount of time (to the nearest decimal in seconds) a participant is able to transfer from a seated to a standing position and back to sitting 5 times. The equipment needed to perform the STS test includes a stopwatch and standard height chair with straight back (42 cm high). Estimated STS power is calculated based on the results of the STS test, using the subject's height and BMI and the height of the chair.

In addition to STS power, other prefrailty phenotype-related predictors were considered, such as age, sex, BMI, alternate Mediterranean diet score (aMDS), energy-adjusted daily intake of vitamin D, protein, and calcium (Ca), and the use of vitamin supplements. Age, gender, and vitamin supplement use were self-reported by participants. BMI was measured by physical examination, the details of which are described in our previous study [26]. Nutrient intake levels were estimated based on a semiquantitative Food Frequency Questionnaire. The adherence to the alternate Mediterranean diet was measured by aMDS, calculated by summing the dichotomous points for the items of higher intakes of whole grain, vegetables, fruits, legumes, nuts, fish, and the ratio of monounsaturated fatty acids:saturated fatty acids, lower red meat, and moderate ethanol consumption, as described previously [34].

Finally, a least absolute shrinkage and selection operator (LASSO) regression model was built to screen the predictors for machine learning (ML) model. Feature selection was performed only in the training dataset. After feature selection, 6 variables (age, sex, BMI, aMDS, STS power, and dietary Ca intake) were selected from 9 predictors to construct a prediction model (Figure S1 in Multimedia Appendix 1).

Statistical Analysis

The continuous measurement data were represented as mean (SD). Student *t* test was used to compare differences between the EML and the normal (no EML) groups. The chi-squared test was used to compare categorical variables. Data were analyzed using R software (version 4.1.0; R Foundation for Statistical Computing).

ML Stages

The data were preprocessed before the models were trained. The predictors exhibit nonzero variance, and there is no significant correlation between them (Pearson correlation coefficient <0.75). Subsequently, the original dataset was randomly divided into a training dataset with 70% (940/1344) of the sample (796 normal and 144 EMLs) and a testing dataset with 30% (404/1344) of the sample (342 normal and 62 EMLs). Some predictor variables had a percentage of missing data of <10% (Figure S2 in Multimedia Appendix 1). For missing data, the training dataset was imputed with the random forest (RF) method by "mice" package [35]. Missing values of the testing dataset were imputed using median (for numerical) and mode (for categorical), with imputation parameters estimated from the training dataset to avoid leakage. This study performed variable (categorical and numerical) standardization on both the training and testing datasets.

Class imbalance is a common challenge encountered in medical datasets, and it can significantly impact the performance of classification models. The presence of imbalanced classes often leads to overfitting of the majority class data by learning algorithms, resulting in decreased classification performance [36]. Sampling is a widely used technique for adjusting the size of training datasets, and undersampling and oversampling are common approaches [37]. To deal with binary classification

XSL•FO
RenderX

problems in the presence of imbalanced classes, synthetic balanced samples were generated according to "ROSE (Random Over-Sampling Examples)" package [38]. ROSE is a bootstrap-based technique that aids the task of binary classification in the presence of rare classes. We combined oversampling and undersampling to rectify the imbalance in data classification. The distribution of the observations by category-wise in the balanced and imbalanced training datasets is shown in Figure S3 in Multimedia Appendix 1.

After the work of preprocessing, the imbalanced and balanced datasets were divided into 5 random parts in the cross-validation test, which was repeated 10 times to randomize the training dataset and improve model performance. In this stage, the models were manually adjusted during training and selected based on the best value of the receiver operating characteristic (ROC), specificity, and sensitivity. The values of these metrics and the parameters used are shown in Table S1 in Multimedia Appendix 1. Calibration curve and decision curve analyses (DCA) were used to evaluate the calibration and clinical utility of the models. The following 6 ML models were trained: logistics regression (LR), RF, support vector machine (SVM), neural network (NNET), lasso regression (LASSO), and Light Gradient-Boosting Machine (LightGBM).

Finally, the testing dataset was used to evaluate the performance of the models trained on the imbalanced and balanced datasets, to assess how well these models fitted the data that was not used during the training stage into the testing dataset. The ROC and precision-recall curves and the metrics area under the receiver operating characteristic curve (AUC), F_1-score, accuracy, balanced accuracy, specificity, sensitivity, positive predictive value (PPV), and negative predictive value were used to assess the performance of the 6 models. The Youden index was used to determine the optimal threshold for the prediction of EML by optimizing specificity and sensitivity. The "tidymodels" package was used in the ML stages.

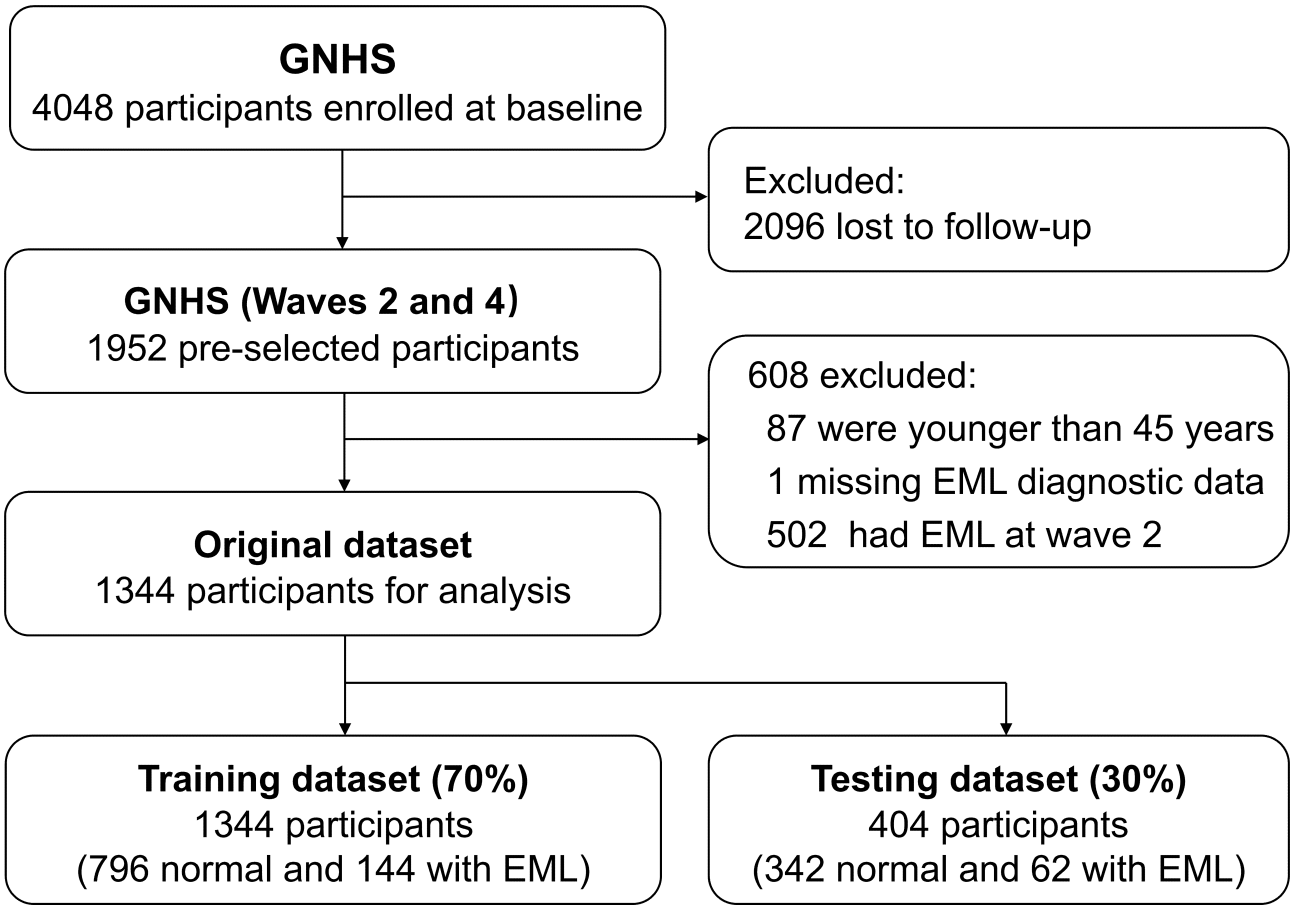
The Shapley Additive Explanations (SHAP) analysis was used to interpret the model and assess the marginal contribution of each feature to the model output. The waterfall of individualized SHAP was used to interpret single-sample data, illustrating the direction and strength of influence that an individual's variable values have on the predictive outcome.

Results

Population Characteristics

In the final analysis, 1344 participants attended the follow-up and were included. At baseline, the median age was 61.99 (IQR 59.03 - 65.68) years, and most participants were women (945/1344, 70.3%). After a median follow-up of 6.67 (IQR 5.91 - 7.26) years, 15.33% (206/1344) participants developed EMLs at 6-year follow-up (Figure 1). Those who developed EMLs were older, had a higher BMI, a lower aMDS, and poorer performance in the STS test, as well as lower estimated muscle power from the test (STS power) at baseline (Table 1).

Figure 1. Flowchart of the machine learning stages in this study. EML: early mobility limitation; GNHS: Guangzhou Nutrition and Health Study.



XSL•FO
RenderX

Table . Baseline candidate variables selected for machine learning by mobility limitations at follow-up in the dataset.[a,b,c,d]

Baseline variables	Preclinical mobility limitations (EML[e])			*P* value
	Total (N=1344)	No EML (n=1138)	EML (n=206)	
Age, y, mean (SD)	62.56 (5.14)	62.21 (5.03)	64.46 (5.33)	<.001
Sex, n (%)				.29
Female	945 (70.31)	807 (70.91)	138 (66.99)	
Male	399 (29.69)	331 (29.09)	68 (33.01)	
BMI (kg/m^2), mean (SD)	23.38 (3.17)	23.30 (2.99)	23.84 (4.02)	.02
Protein (g/d), mean (SD)	69.07 (10.50)	69.15 (10.54)	68.66 (10.31)	.54
Vitamin D (IU/d), mean (SD)	83.11 (57.96)	84.03 (59.89)	78.07 (45.62)	.17
Ca (mg/d), mean (SD)	547.46 (164.82)	548.81 (158.83)	540.05 (194.89)	.48
aMDS[f], mean (SD)	4.11 (1.71)	4.15 (1.71)	3.86 (1.69)	.02
Vitamin supplementation, n (%)				>.99
No	1117 (83.11)	946 (83.13)	171 (83.01)	
Yes	227 (16.89)	192 (16.87)	35 (16.99)	
STS[g] power (W), mean (SD)	268.41 (92.38)	271.66 (94.40)	250.49 (78.12)	.002
STS (s), mean (SD)	7.68 (1.88)	7.62 (1.92)	8.01 (1.67)	.003

[a]Continuous variables are presented as mean (SD), and categorical variables are presented as n (%).

[b]Student *t* tests were used to compare the mean differences between no EML and EML groups.

[c]Chi-square tests were used to count data.

[d]Statistics after imputation of missing data.

[e]EML: early mobility limitation.

[f]aMDS: alternate Mediterranean diet score.

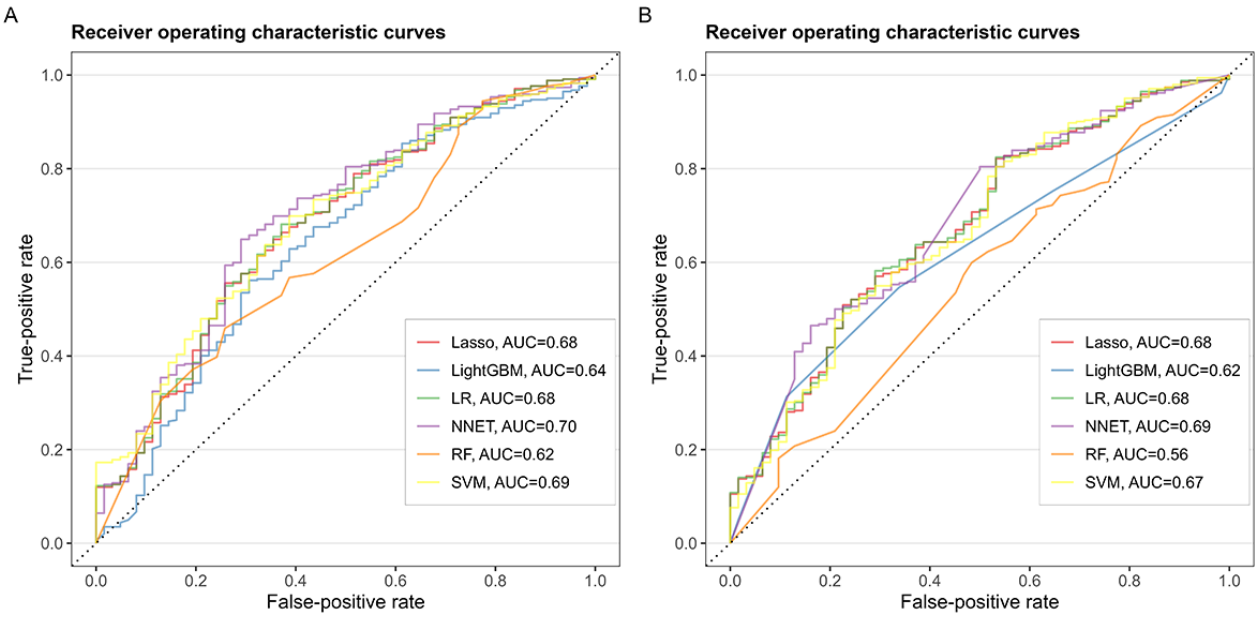
[g]STS: sit-to-stand test.

Models' Performance

Figure 2 shows the AUC values for the 6 ML models trained with imbalanced (AUC range 0.62-0.70; Figure 2A) and balanced (AUC range 0.56-0.69; Figure 2B) datasets. All 6 ML models selected all 6 variables (age, sex, BMI, aMDS, STS power, and dietary calcium intake) with different importance (Figure 3), except the LightGBM model for which age, BMI, and dietary calcium intake were not selected. The performance of most models on unseen data remained stable or decreased when the balanced dataset was used, and the LR model maintained the best performance across both datasets.

XSL•FO
RenderX

Figure 2. Receiver operating characteristic curves of the models trained with imbalanced (A) and balanced (B) datasets. AUC: area under the receiver operating characteristic curve; LASSO: Lasso regression; LightGBM: Light Gradient-Boosting Machine; LR: logistic regression; NNET: neural network; RF: random forest; SVM: support vector machine.



XSL•FO
RenderX

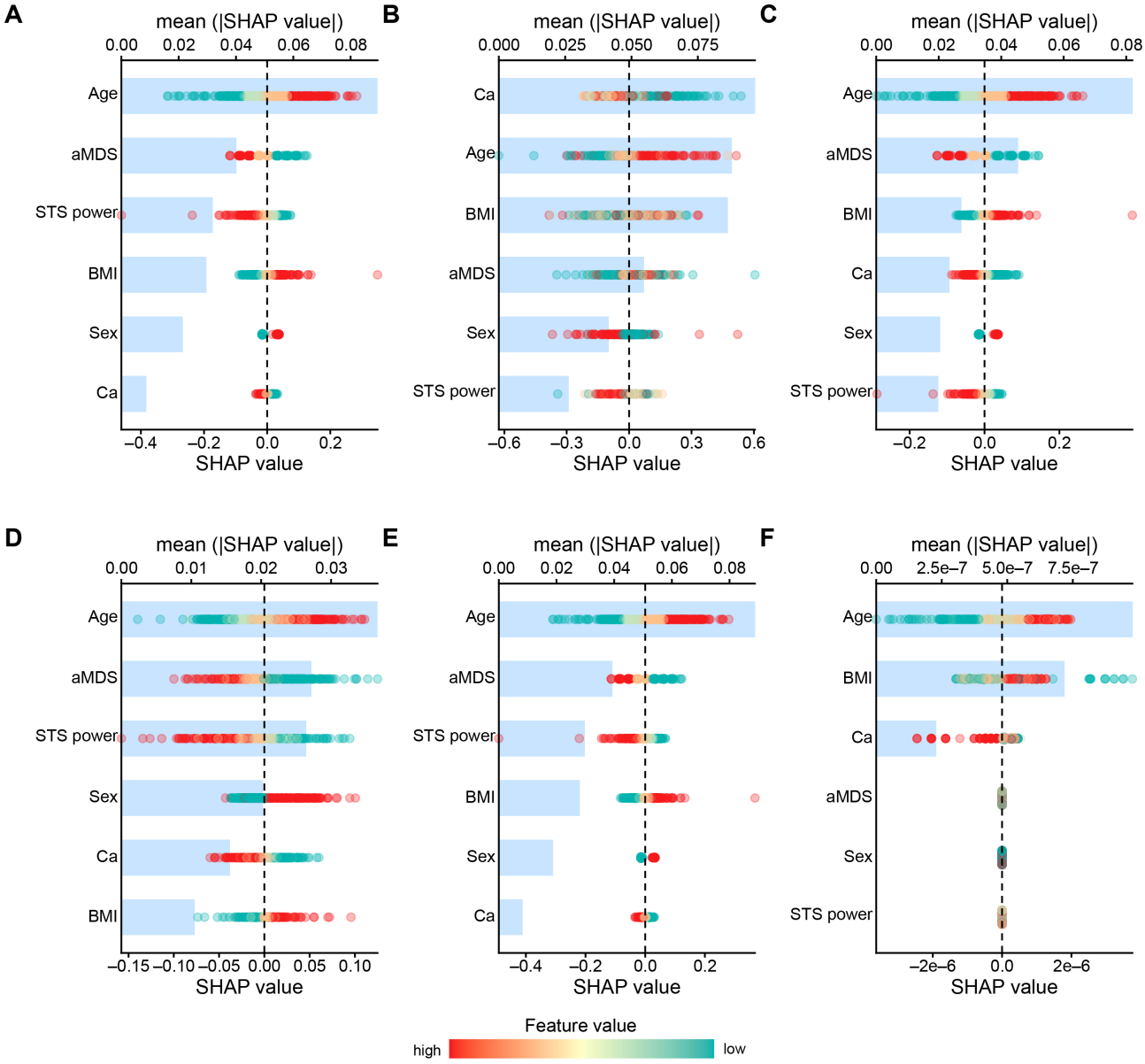


Figure 3. The SHAP for feature importance of the models trained with balanced datasets: (A) logistic regression; (B) random forest; (C) support vector machine; (D) neural network; (E) least absolute shrinkage and selection operator regression; and (F) Light Gradient-Boosting Machine. aMDS: alternate Mediterranean diet score; SHAP: Shapley Additive Explanations; STS: sit-to-stand test.

Table 2 shows the performance of the ML models on unseen data, trained with imbalanced and balanced datasets, based on different metrics. Some models (RF, SVM, and NNET) yielded better values for metric accuracy when they were trained with balanced data, whereas the unbalanced dataset showed better performance in all the other metrics and the rest of the models. LR, NNET, SVM, and LASSO had acceptable values for AUC (≥0.70) or slightly below when trained with imbalanced data, among which NNET had the best values for AUC (0.70; 95% CI 0.63 - 0.77) and balanced accuracy (0.68) compared with all the other models trained with imbalanced data. However, NNET models are typically more complex and prone to overfitting. Hence, in conditions of similar performance, simpler models like LR or LASSO may be preferred due to their interpretability and lower risk of overfitting. Specifically, the straightforward coefficients in LR can be easily understood in terms of odds ratios.

XSL•FO
RenderX

Table . Performance of the models trained with imbalanced and balanced datasets based on different metrics.[a]

Models	Accuracy	Balanced accuracy	F_1-score	PR[a]-AUC[b]	AUC (95% CI)	Threshold	Specificity	Sensitivity	PPV[c]	NPV[d]
Imbalanced dataset										
LR[e]	0.67	0.66	0.92	0.29	0.68 (0.61-0.76)	0.16	0.68	0.63	0.26	0.91
RF[f]	0.50	0.60	0.91	0.23	0.62 (0.55-0.70)	0.08	0.46	0.74	0.20	0.91
SVM[g]	0.69	0.66	0.92	0.27	0.69 (0.62-0.76)	0.15	0.70	0.61	0.27	0.91
NNET[h]	0.66	0.68	0.92	0.29	0.70 (0.63-0.77)	0.32	0.65	0.71	0.27	0.92
LASSO[i]	0.58	0.65	0.92	0.29	0.68 (0.61-0.76)	0.14	0.56	0.74	0.23	0.92
Light-GBM[j]	0.58	0.62	0.92	0.23	0.64 (0.56-0.72)	0.18	0.56	0.69	0.22	0.91
Balanced dataset										
LR	0.77	0.64	0.69	0.28	0.68 (0.60-0.75)	0.60	0.82	0.47	0.33	0.90
RF	0.59	0.56	0.74	0.16	0.56 (0.48-0.63)	0.43	0.60	0.52	0.19	0.87
SVM	0.75	0.64	0.71	0.29	0.67 (0.60-0.74)	0.59	0.80	0.47	0.30	0.89
NNET	0.76	0.65	0.66	0.24	0.69 (0.62-0.76)	0.52	0.80	0.50	0.32	0.90
LASSO	0.77	0.64	0.69	0.28	0.68 (0.60-0.75)	0.60	0.82	0.47	0.32	0.90
Light-GBM	0.56	0.60	0.92	0.17	0.62(0.55-0.69)	0.50	0.55	0.66	0.21	0.90

[a]PR: precision-recall.

[b]AUC: area under the receiver operating characteristic curves.

[c]PPV: positive predictive value.

[d]NPV: negative predictive value.

[e]LR: logistic regression.

[f]RF: random forest.

[g]SVM: support vector machine.

[h]NNET: neural network.

[i]LASSO: least absolute shrinkage and selection operator.

[j]LightGBM: Light Gradient-Boosting Machine.

DCA (Figure S8 in Multimedia Appendix 1) results showed that prediction model's effective range for test datasets are concentrated at low thresholds. All prediction model curves demonstrate positive net benefits exceeding the "no treatment" baseline only within the threshold range of 0% to 50%. When the threshold exceeds 50%, the model's net benefit drops below zero, losing clinical value.

In calibration curves (Figures S5 and S7 in Multimedia Appendix 1), LASSO regression and logistic regression exhibit relatively good calibration on imbalanced test datasets. However, in balanced datasets, most predictive models show that the predicted probabilities are higher than the actual probabilities, indicating that the models are overconfident in their predictions. For implementation in the real world, prediction models need to be further calibrated according to the specific application scenario.

In addition to discrimination metrics, we assessed model performance using the net reclassification improvement. Logistic regression was selected as reference because it represents a widely used baseline approach for binary risk prediction in preventive settings. Net reclassification improvement was decomposed into event and nonevent components to better understand potential gains and tradeoffs in risk stratification across modeling approaches, and the results are reported separately for models trained on imbalanced and balanced datasets (Table S1 in Multimedia Appendix 1). Only RF achieved a modest improvement in risk reclassification over

XSL•FO
RenderX

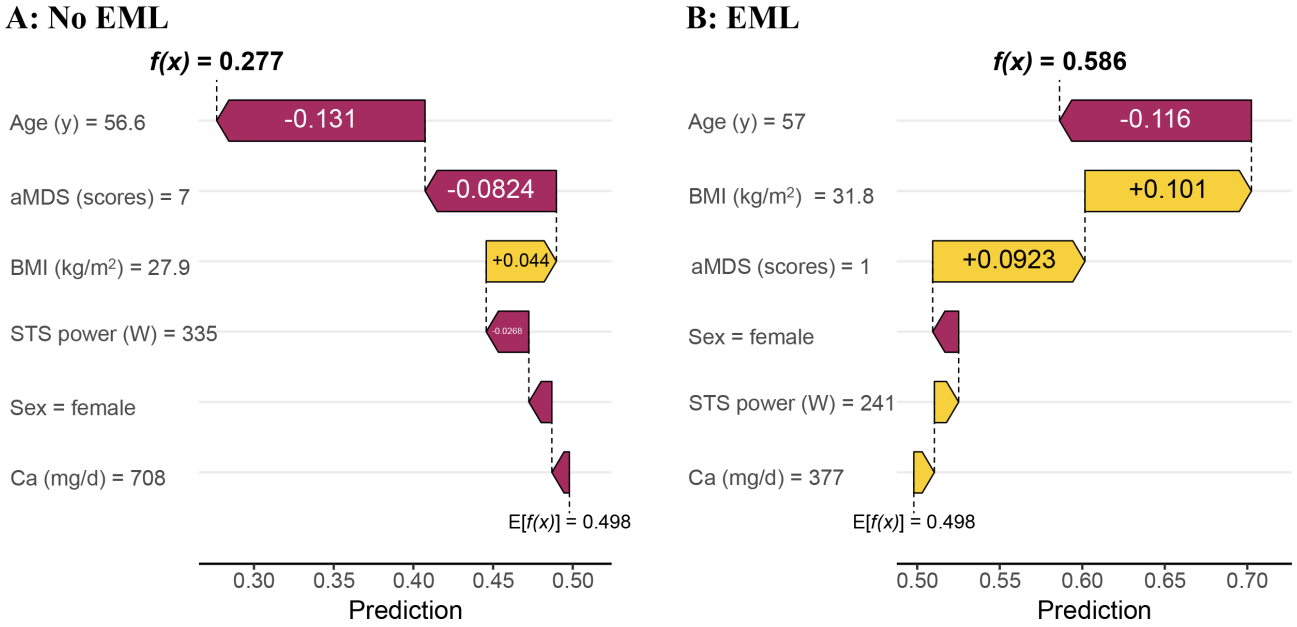
logistic regression, whereas other ML models did not demonstrate meaningful net reclassification gains.

Model Interpretability

The SHAP model was used to provide both global and local interpretations of the models. Globally, the more significant contributors to the models trained with the balanced dataset were age, aMDS, STS power, and BMI. Age was among the top 3 most important variables in all 6 models, followed by aMDS in 4 models (LR, SVM, NNET, and LASSO) and STS power (LR, NNET, and LASSO) and BMI (RF, SVM, and light GBM) in 3 models (Figure 3).

In the single-sample waterfall plot, baseline SHAP value was 0.498, model output values for a randomly selected individual with no EML was 0.277 (Figure 4A), and for an EML individual was 0.586 (Figure 4B). In no EML individuals, lower age as well as higher aMDS scores and STS power were identified as features that had a negative impact on EML, while higher BMI was considered to have a positive impact (Figure 4A). Among EML individuals, lower age was identified as a feature that negatively affected EML, while higher BMI and lower aMDS scores were considered to have a positive impact (Figure 4B).

Figure 4. Individualized SHAP waterfall plot for no EML (A) and EML participants (B), demonstrating the personalized contribution of each predictor to the risk score. aMDS: alternate Mediterranean diet score; EML: early mobility limitation; SHAP: Shapley Additive Explanations; STS: sit-to-stand test.



The SHAP bar plot ranks features by their overall importance, calculated as the mean absolute SHAP value across all participants.

The SHAP beeswarm plot illustrates the impact of each feature on the model's output for every participant. Each dot represents a participant; its position on the x - axis indicates the SHAP value (positive values increase the prediction of EML), and its color represents the feature's value (red for high, green for low).

Discussion

Principal Results

In this study, we developed 6 prediction models using ML for EML in healthy, mobility-intact middle-aged and older adults from a longitudinal Chinese population-based cohort who were followed-up over 6.67 years. A LASSO regression model was built to select the predictors. Finally, 6 variables (age, sex, BMI, aMDS, STS power, and dietary calcium intake) were identified and included in the model training process. The performance of the different ML models trained with balanced and imbalanced datasets was compared. Advanced age, lower adherence to Mediterranean diet, lower STS power, and higher BMI were identified as the most important predictors of EMLs.

Comparison to Prior Work

In this study, we have attempted to use commonly used ML models for binary variable prediction, including LR, RF, SVM, NNET, LASSO, and LightGBM. In a recent study with a large sample size (5466), ML models such as SVM, K-nearest neighbor, NNET, RF, and extreme gradient boosting demonstrated good results in predicting frailty among senior people in Canada [39]. Although our database may not have a comparable size, we were interested in exploring the feasibility of applying these ML models to the cohort in Guangzhou, China.

The discrimination power of 4 models (LR, NNET, SVM, and LASSO) trained with the imbalanced data was found acceptable (AUC≥0.70) or slightly below in the testing dataset. The performance moderately decreased when trained with the balanced dataset, with NNET (0.69), LR (0.68), SVM (0.67), and LASSO (0.68) achieving the highest AUC values and, therefore, showing more robust discrimination across datasets (Table 2). Royston et al [40] noted that the AUC of prognostic models is typically between 0.6 and 0.85. The commonly used prognostic tests include the ratio of total plasma cholesterol to high-density lipoprotein cholesterol level, with an AUC of 0.72 for predicting the risk of coronary heart disease [41], and the 6-minute walking test, with an AUC of 0.60 for predicting

XSL•FO
RenderX

increased risk of hospitalization among patients with chronic obstructive pulmonary disease [42]. Therefore, the predictive power of our models compares well with that of risk assessment tests widely used in clinical practice that are more invasive (eg, TC/HDL-C from a blood test) or burdensome (eg, 6-minute walking test requiring time and space). Additional DCA of the models trained with imbalanced datasets showed meaningful positive net benefits for LASSO and LR only, supporting the clinical utility of these 2 models within the threshold range of 0% to 35%, but no positive net benefits were found for the balanced trained models (Figure S8 in Multimedia Appendix 1). The incremental predictive benefit of other ML models compared to LR did not demonstrate meaningful net reclassification gains.

Similar ROC-AUC results were obtained by Chaves et al [43], who developed a screening tool from simple–to-obtain self-reported performance measures for predicting the probability of 18-month incident mobility difficulty in community-dwelling women aged 70 to 80 years, with an AUC of 0.73. However, a test is likely to be more sensitive when the prevalence of concomitant risk factors is higher, as it would be expected from older individuals such as those included in that study [44]. Likewise, as the time to event decreases, the predictive accuracy of a model tends to improve [45], leading to less uncertainty in short-term predictions like in the study by Chaves et al [43]. Taken together, considering the challenges of predicting the onset of the early stages of an unspecific and multifactorial condition such as EML in mobility-intact adults who are generally healthy, drawing mostly from modifiable predictors, this study yielded acceptable results for the LASSO and LR imbalanced trained models, with positive net benefits. Given the models' intended use for early identification of individuals who may benefit from low-risk lifestyle interventions, their limited discrimination remains potentially valuable.

The models developed in this study were designed to maximize clinical applicability. All 6 predictors (age, sex, BMI, STS power, aMDS, and dietary calcium intake) can be easily collected by individuals using a standard chair, a stopwatch, and a short diet questionnaire. These inputs could be incorporated into a digital self-assessment platform or mobile health tool to support interpretation and reporting. Because the predictors include modifiable factors such as diet quality, micronutrient intake, physical function, and body weight, such a tool could help identify individuals at higher risk of mobility decline and provide personalized lifestyle recommendations. The decision threshold can be implemented within preventive care workflows, with thresholds in the range of 10% to 20%—as supported by DCA—offering a balance between early intervention benefits and the likelihood of false positives. This approach may empower individuals to monitor and preserve their mobility health without requiring specialized training or extensive resources.

As expected, age was the most important predictor of incident EML across the different ML models. Advanced age is a risk factor in developing mobility limitations and has been consistently identified as a predictor of mobility outcomes in many predictive equations [46,47]. Therefore, the finding that advanced age was the most important variable across the different ML models as determined by SHAP values (Figure 3) served as a positive control variable in developing our models to predict EML. Other important variables were adherence to the Mediterranean diet, STS power, and BMI. Among the rest of the baseline candidate predictors selected by LASSO regression, other less important variables across the different ML models included being female and lower energy-adjusted calcium intake. Many of these variables confirm some of the risk factors most frequently found to be independently linked to mobility limitations in older people, which include advanced age, higher BMI, and reduced muscle function measured through performance tests [48]. However, our results cast a new light on the predictive ability of estimated muscle power from STS and diet quality.

This study found support for STS power as an important predictor of incident EML in older, as well as in late middle-aged adults. The STS test has been consistently identified as a predictor of subsequent falls, mobility limitations [49], disability [14], osteoporotic hip fractures, and mortality among older adults. Remarkably, we found a higher statistically significant difference and relative difference for STS power (7.79%, P=.002) than for STS time values (4.87%, P=.003) at baseline between the group developing EML and their counterparts (Table 1). Therefore, only STS power was included among the baseline candidate predictors. Whereas the STS test measures the minimum duration required to rise from a chair a specific number of times (eg, 5-time STS) or the maximum number of STS repetitions within a specified timeframe (eg, 30-s STS), the STS power formula also considers the chair height, and the subject's body mass and height. Previous evidence in older people suggests that STS power values are more clinically relevant in regard to physical function, cognitive function, and sarcopenia compared to traditional STS measures [20]. Consistent with these findings in older adults, our results extend the predictive value of STS power to identify prospective mobility limitations in late middle-aged adults.

Notably, although normative values for STS power in Asian populations are lacking, the 5-time STS values observed in our study suggest a rather optimal level of physical function in both the EML group (mean 8.01, SD 1.67 s) and the normal group (mean 7.62, SD 1.92 s) at baseline. Specifically, current guidelines from the AWGS 2019 recommend a cut-off of 12 seconds or more to screen for low physical performance and sarcopenia, based on a review of evidence from different Asian cohorts. The striking distance between that cut-off and the baseline functional performance observed in our cohort underlines the ability of our models to identify incident mobility limitations earlier in time, even in the absence of noticeable performance decline. Remarkably, a recent study redefined the cut-off value recommended by the AWGS to 9.2 seconds in men and 10.8 seconds in women aged 60 to 69 years, significantly improving the detection rate of sarcopenia among Chinese adults [50]. These lower thresholds tie well with the values observed in our study and further support the predictive value of subtle deficits in performance tests.

A further novel finding was the importance of diet quality, measured as adherence to the Mediterranean diet, to predict

XSL•FO
RenderX

incident EML. Studying the influence of dietary factors on mobility limitation has practical importance, and the findings could provide opportunities for dietary intervention to reduce the potential risk. In this study, we identified adherence to the Mediterranean diet as a predictor of mobility limitations, with a higher adherence score associated with a lower probability of having mobility limitations. This finding is consistent with the previous studies. The aMDS was found positively associated with skeletal muscle mass index, which is an important aspect of mobility, in the Chinese population aged 40 to 75 years [34]. Silva et al [51] conducted a meta-analysis in community-dwelling older people and found that higher adherence to the Mediterranean diet was inversely associated with frailty. The anti-inflammatory and antioxidant properties of some nutrients typically found in higher amounts in Mediterranean diets, such as vitamin C and vitamin E or carotenoids, could contribute to that benefit [52]. Future intervention studies are needed to provide direct evidence on the risk reduction of mobility limitation by adherence to aMDS and improvement of calcium intake among individuals aged 45 years and older.

Strengths and Limitations

Although many prediction models for mobility limitations or disability have been developed in older people with poorer baseline health and functional performance, to our knowledge, comparable models are lacking for mobility-intact adults under 65 years of age. Therefore, a major strength of the study is the sample selection of healthier and younger participants, allowing for the observation of the early transition from intact mobility to impairment. This design facilitates the identification of early predictors and risk factors of mobility limitations, which might be overlooked in studies restricted to older or already impaired individuals.

Another positive aspect of the study is the selected set of predictor variables, which can be easily self-assessed at home with a smartphone and comprises several modifiable risk factors such as diet, physical function, and weight. As a result, the information collected for the prognostic test could be used for guiding potential intervention strategies. Finally, despite the common practice among ML researchers of using extensive datasets and training models with numerous predictors, this research has shown the feasibility of applying ML in datasets with a reasonable number of predictors and in cases where the classes are initially unbalanced.

There are some limitations of the study. First, the composite outcome EML measured in this research integrated both patient-reported and performance-based measures, as recommended by Richardson et al [11]. The expert group conceptualized the term *early* or *preclinical* mobility limitations to capture subtle decrements in mobility; however, a standardized measure has not yet been established. Further research examining the reliability, validity, and responsiveness of outcome measures is necessary to strengthen the evidence base on preclinical mobility limitation.

Second, the PPV, which indicates the likelihood that a positive test result is a true positive, was not optimal across models. These results should be interpreted considering the relatively low prevalence of EML observed in our cohort (15.3%) compared to that reported in other studies with similar operationalizations of EML among Chinese (27.9%) [53]. Because predictive values are prevalence-dependent, this may have increased the proportion of false positives, lowering the PPVs of our models [54]. This prevalence discrepancy may stem from the high proportion of lost-to–follow-up participants in our study, over 50%. The cohort in this study was followed in a designated location, introducing attrition bias for participants who were more frail, developed disability, or moved to care facilities. To address this discrepancy, we recalculated the PPV using the prevalence in the target population (27.9% in China) for the LR, SVM, NNET, and LASSO models. The PPV increases to approximately 43% to 44% indicating a substantially higher probability that individuals classified as high risk will truly experience the outcome when the model is applied in the intended population. In addition, proper calibration of the models may further optimize the PPV [55]. Notably, in the context of rule-in tests, it is crucial to balance the benefits of correctly identifying at-risk individuals against the potential harm of false positives [56]. The prognostic models developed in this study are aimed at promoting healthy lifestyle advice, like increased physical activity and improved nutrition, which are particularly relevant for the intended at-risk population. Therefore, the limited PPVs observed in our sample shall not limit the applicability of some of our models to identify individuals who may benefit the most from low-risk preventive interventions in the target Chinese population. In particular, a consistent net benefit was shown in DCA for SVM and LR across threshold probabilities roughly in the 10% to 35% band.

Another major limitation of the study was the validation of the models with an internal testing dataset. Further studies on independent cohorts need to be undertaken to establish the validity of our findings. Lastly, we suffered from an imbalanced dataset with only 206 subjects developing EML out of the 1344 sample. Similar imbalance problems have been observed in related studies [37], and this imbalance can lead to biased estimates of training performance [36]. This is because algorithms always tend to predict negative outcomes to achieve higher accuracy, which can also bring implementation challenges, such as when the primary interest lies in discovering the predictive results for EML, and biased predictions render the algorithm meaningless. However, in this study, we used the ROSE bootstrap-based technique [38] to remedy the imbalanced data and generate synthetic balanced samples, leading to more reliable and consistent evaluation of our models across 2 datasets.

Conclusions

This study is the first, to the authors' knowledge, to develop a prediction model for identifying the onset of EMLs in middle-aged and older adults based exclusively on at-home, simple–to-obtain measurements. The evaluation of model performance on a testing dataset, considering both imbalanced and balanced data training, indicates the potential utility of the ML approach in identifying individuals who developed EMLs over time. LASSO and logistic regression yielded the most meaningful positive net benefits, supporting their potential clinical utility for early identification of individuals who may

XSL•FO
RenderX

benefit from low-risk lifestyle interventions, though external validation is still required. Across models, age, adherence to the Mediterranean diet, STS power, and BMI consistently emerged as the most important predictors. These models provide an initial foundation for developing a digital self-assessment tool to identify individuals at higher risk of mobility decline and support personalized preventive strategies in clinical and at-home settings.

Acknowledgments

The authors thank Fred (JZ) and Yu Kai for their work as scientific advisors.

Funding

The original cohort study was jointly supported by the Nutrition Research Foundation of the Chinese Nutrition Society (CNS-ZD2019049), National Natural Science Foundation of China (82574079, 82073533, and the 5010 Program for Clinical Researches (2007032) of Sun Yat-sen University (Guangzhou, China). No generative artificial intelligence tools were used at any stage in the preparation of this manuscript.

Data Availability

The data of Guangzhou Nutrition and Health Study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: ACF, DB, YC
Methodology: ACF, MC, WH
Data management and analysis: MC, WH
Writing – original draft: ACF
Writing – review & editing: ACF, DB, MC, DW, WH, YC

Conflicts of Interest

ACF, DW, and DB are employees of Société des Produits Nestlé SA. The authors declare no competing interests.

Multimedia Appendix 1
Supplementary figures and tables.
[DOCX File, 1091 KB - aging_v9i1e77187_app1.docx]

References

1. Inouye SK, Studenski S, Tinetti ME, Kuchel GA. Geriatric syndromes: clinical, research, and policy implications of a core geriatric concept. J Am Geriatr Soc 2007 May;55(5):780-791. [doi: 10.1111/j.1532-5415.2007.01156.x] [Medline: 17493201]
2. Webber SC, Porter MM, Menec VH. Mobility in older adults: a comprehensive framework. Gerontologist 2010 Aug;50(4):443-450. [doi: 10.1093/geront/gnq013] [Medline: 20145017]
3. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 1;48(1):16-31. [doi: 10.1093/ageing/afy169] [Medline: 30312372]
4. Cruz-Jentoft AJ, Sayer AA. Sarcopenia. Lancet 2019 Jun 29;393(10191):2636-2646. [doi: 10.1016/S0140-6736(19)31138-9] [Medline: 31171417]
5. Cesari M, Landi F, Vellas B, Bernabei R, Marzetti E. Sarcopenia and physical frailty: two sides of the same coin. Front Aging Neurosci 2014;6:192. [doi: 10.3389/fnagi.2014.00192] [Medline: 25120482]
6. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
7. Kojima G, Taniguchi Y, Iliffe S, Jivraj S, Walters K. Transitions between frailty states among community-dwelling older people: a systematic review and meta-analysis. Ageing Res Rev 2019 Mar;50:81-88. [doi: 10.1016/j.arr.2019.01.010] [Medline: 30659942]
8. Ferrucci L, Cooper R, Shardell M, Simonsick EM, Schrack JA, Kuh D. Age-related change in mobility: perspectives from life course epidemiology and geroscience. J Gerontol A Biol Sci Med Sci 2016 Sep;71(9):1184-1194. [doi: 10.1093/gerona/glw043] [Medline: 26975983]
9. Aboutorabi A, Arazpour M, Bahramizadeh M, Hutchins SW, Fadayevatan R. The effect of aging on gait parameters in able-bodied older subjects: a literature review. Aging Clin Exp Res 2016 Jun;28(3):393-405. [doi: 10.1007/s40520-015-0420-6] [Medline: 26210370]
10. Ferrucci L, Guralnik JM, Studenski S, et al. Designing randomized, controlled trials aimed at preventing or delaying functional decline and disability in frail, older persons: a consensus report. J Am Geriatr Soc 2004 Apr;52(4):625-634. [doi: 10.1111/j.1532-5415.2004.52174.x] [Medline: 15066083]

XSL•FO
RenderX

11. Richardson J, Beauchamp M, Bean J, et al. Defining and measuring preclinical mobility limitation: an expert consensus exercise informed by a scoping review. J Gerontol A Biol Sci Med Sci 2023 Aug 27;78(9):1641-1650. [doi: 10.1093/gerona/glad143] [Medline: 37300461]
12. Rantanen T, Guralnik JM, Foley D, et al. Midlife hand grip strength as a predictor of old age disability. JAMA 1999 Feb 10;281(6):558-560. [doi: 10.1001/jama.281.6.558] [Medline: 10022113]
13. Cavanaugh EJ, Richardson J, McCallum CA, Wilhelm M. The predictive validity of physical performance measures in determining markers of preclinical disability in community-dwelling middle-aged and older adults: a systematic review. Phys Ther 2018 Dec 1;98(12):1010-1021. [doi: 10.1093/ptj/pzy109] [Medline: 30256988]
14. Deshpande N, Metter EJ, Guralnik J, Bandinelli S, Ferrucci L. Predicting 3-year incident mobility disability in middle-aged and older adults using physical performance tests. Arch Phys Med Rehabil 2013 May;94(5):994-997. [doi: 10.1016/j.apmr.2012.10.032] [Medline: 23164980]
15. Perera S, Patel KV, Rosano C, et al. Gait speed predicts incident disability: a pooled analysis. J Gerontol A Biol Sci Med Sci 2016 Jan;71(1):63-71. [doi: 10.1093/gerona/glv126] [Medline: 26297942]
16. Suetta C, Haddock B, Alcazar J, et al. The Copenhagen Sarcopenia Study: lean mass, strength, power, and physical function in a Danish cohort aged 20-93 years. J Cachexia Sarcopenia Muscle 2019 Dec;10(6):1316-1329. [doi: 10.1002/jcsm.12477] [Medline: 31419087]
17. Rasmussen LJH, Caspi A, Ambler A, et al. Association of neurocognitive and physical function with gait speed in midlife. JAMA Netw Open 2019 Oct 2;2(10):e1913123. [doi: 10.1001/jamanetworkopen.2019.13123] [Medline: 31603488]
18. Reid KF, Fielding RA. Skeletal muscle power: a critical determinant of physical functioning in older adults. Exerc Sport Sci Rev 2012 Jan;40(1):4-12. [doi: 10.1097/JES.0b013e31823b5f13] [Medline: 22016147]
19. Lauretani F, Russo CR, Bandinelli S, et al. Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia. J Appl Physiol (1985) 2003 Nov;95(5):1851-1860. [doi: 10.1152/japplphysiol.00246.2003] [Medline: 14555665]
20. Alcazar J, Kamper RS, Aagaard P, et al. Relation between leg extension power and 30-s sit-to-stand muscle power in older adults: validation and translation to functional performance. Sci Rep 2020 Oct 1;10(1):16337. [doi: 10.1038/s41598-020-73395-4] [Medline: 33004970]
21. Alcazar J, Losa-Reyna J, Rodriguez-Lopez C, et al. The sit-to-stand muscle power test: an easy, inexpensive and portable procedure to assess muscle power in older people. Exp Gerontol 2018 Oct 2;112:38-43. [doi: 10.1016/j.exger.2018.08.006] [Medline: 30179662]
22. Mazza E, Ferro Y, Pujia R, et al. Mediterranean diet in healthy aging. J Nutr Health Aging 2021;25(9):1076-1083. [doi: 10.1007/s12603-021-1675-6] [Medline: 34725664]
23. Shahar DR, Houston DK, Hue TF, et al. Adherence to mediterranean diet and decline in walking speed over 8 years in community-dwelling older adults. J Am Geriatr Soc 2012 Oct;60(10):1881-1888. [doi: 10.1111/j.1532-5415.2012.04167.x] [Medline: 23035758]
24. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015 Jan 6;162(1):W1-W73. [doi: 10.7326/M14-0698] [Medline: 25560730]
25. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
26. Ling CW, Zhong H, Zeng FF, et al. Cohort profile: Guangzhou Nutrition and Health Study (GNHS): a population-based multi-omics study. J Epidemiol 2024 Jun 5;34(6):301-306. [doi: 10.2188/jea.JE20230108] [Medline: 37813622]
27. Zhang CX, Ho SC. Validity and reproducibility of a food frequency questionnaire among Chinese women in Guangdong province. Asia Pac J Clin Nutr 2009;18(2):240-250. [Medline: 19713184]
28. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. JAMA 2011 Jan 5;305(1):50-58. [doi: 10.1001/jama.2010.1923] [Medline: 21205966]
29. Cesari M, Kritchevsky SB, Penninx B, et al. Prognostic value of usual gait speed in well-functioning older people--results from the Health, Aging and Body Composition Study. J Am Geriatr Soc 2005 Oct;53(10):1675-1680. [doi: 10.1111/j.1532-5415.2005.53501.x] [Medline: 16181165]
30. Schrack JA, Zipunnikov V, Simonsick EM, Studenski S, Ferrucci L. Rising energetic cost of walking predicts gait speed decline with aging. J Gerontol A Biol Sci Med Sci 2016 Jul;71(7):947-953. [doi: 10.1093/gerona/glw002] [Medline: 26850913]
31. Makino K, Raina P, Griffith LE, et al. Lifetime physical activity and late-life mild cognitive impairment in community-dwelling older adults. J Am Med Dir Assoc 2024 Mar;25(3):488-493.e3. [doi: 10.1016/j.jamda.2023.12.006] [Medline: 38246592]
32. Wu X, Li X, Xu M, Zhang Z, He L, Li Y. Sarcopenia prevalence and associated factors among older Chinese population: findings from the China Health and Retirement Longitudinal Study. PLoS One 2021;16(3):e0247617. [doi: 10.1371/journal.pone.0247617] [Medline: 33661964]

XSL•FO
RenderX

33. Cheong CY, Yap P, Gwee X, et al. Physical and functional measures predicting long-term mortality in community-dwelling older adults: a comparative evaluation in the Singapore Longitudinal Ageing Study. Aging (Albany NY) 2021 Dec;13(23):25038-25054. [doi: 10.18632/aging.203756] [Medline: 34894397]
34. Tian HY, Qiu R, Jing LP, Chen ZY, Chen GD, Chen YM. Alternate Mediterranean diet score is positively associated with skeletal muscle mass index in middle-aged adults. Br J Nutr 2017 Apr;117(8):1181-1188. [doi: 10.1017/S0007114517001118] [Medline: 28514984]
35. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45(3):1-67. [doi: 10.18637/jss.v045.i03]
36. Kim MS. An effective under-sampling method for class imbalance data problem. Presented at: Eighth International Symposium on Intelligent Systems (ISIS 2007); Sep 5-8, 2007.
37. Li DC, Liu CW, Hu SC. A learning method for the class imbalance problem with medical data sets. Comput Biol Med 2010 May;40(5):509-518. [doi: 10.1016/j.compbiomed.2010.03.005] [Medline: 20347072]
38. Lunardon N, Menardi G, Torelli N. ROSE: a package for binary imbalanced learning. R J 2014;6(1):79. [doi: 10.32614/RJ-2014-008]
39. Aponte-Hao S, Wong ST, Thandi M, et al. Machine learning for identification of frailty in Canadian primary care practices. Int J Popul Data Sci 2021;6(1):1650. [doi: 10.23889/ijpds.v6i1.1650] [Medline: 34541337]
40. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. BMJ 2009 Mar 31;338:b604. [doi: 10.1136/bmj.b604] [Medline: 19336487]
41. Grover SA, Coupal L, Hu XP. Identifying adults at increased risk of coronary disease. How well do the current cholesterol guidelines work? JAMA 1995 Sep 13;274(10):801-806. [Medline: 7650803]
42. Spruit MA, Polkey MI, Celli B, et al. Predicting outcomes from 6-minute walk distance in chronic obstructive pulmonary disease. J Am Med Dir Assoc 2012 Mar;13(3):291-297. [doi: 10.1016/j.jamda.2011.06.009] [Medline: 21778120]
43. Chaves PH, Garrett ES, Fried LP. Predicting the risk of mobility difficulty in older women with screening nomograms: the Women's Health and Aging Study II. Arch Intern Med 2000 Sep 11;160(16):2525-2533. [doi: 10.1001/archinte.160.16.2525] [Medline: 10979066]
44. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007 Feb 20;115(7):928-935. [doi: 10.1161/CIRCULATIONAHA.106.672402] [Medline: 17309939]
45. Wang Z, Wang X. Evaluating the time-dependent predictive accuracy for event-to-time outcome with a cure fraction. Pharm Stat 2020 Nov;19(6):955-974. [doi: 10.1002/pst.2048] [Medline: 32776646]
46. Neumann JT, Thao LTP, Murray AM, et al. Prediction of disability-free survival in healthy older people. Geroscience 2022 Jun;44(3):1641-1655. [doi: 10.1007/s11357-022-00547-x] [Medline: 35420334]
47. Speiser JL, Callahan KE, Ip EH, et al. Predicting future mobility limitation in older adults: a machine learning analysis of Health ABC study data. J Gerontol A Biol Sci Med Sci 2022 May 5;77(5):1072-1078. [doi: 10.1093/gerona/glab269] [Medline: 34529794]
48. Brown CJ, Flood KL. Mobility limitation in the older patient: a clinical review. JAMA 2013 Sep 18;310(11):1168-1177. [doi: 10.1001/jama.2013.276566] [Medline: 24045741]
49. Pua YH, Tay L, Terluin B, et al. Estimating cutpoints of gait speed and sit-to-stand test values for self-reported mobility limitations in a cohort of community-dwelling older adults from Singapore: comparing receiver operating characteristic (ROC) analysis with adjusted predictive modelling. Arch Gerontol Geriatr 2023 Sep;112:105036. [doi: 10.1016/j.archger.2023.105036] [Medline: 37075584]
50. Li YH, Wang XH, Ya S, Jiaoling H, Hua N. The optimal cut-off value of five-time chair stand test for assessing sarcopenia among Chinese community-dwelling older adults. J Cachexia Sarcopenia Muscle 2024 Apr;15(2):756-764. [doi: 10.1002/jcsm.13441] [Medline: 38342700]
51. Silva R, Pizato N, da Mata F, Figueiredo A, Ito M, Pereira MG. Mediterranean diet and musculoskeletal-functional outcomes in community-dwelling older people: a systematic review and meta-analysis. J Nutr Health Aging 2018;22(6):655-663. [doi: 10.1007/s12603-017-0993-1] [Medline: 29806854]
52. Granic A, Sayer AA, Robinson SM. Dietary patterns, skeletal muscle health, and sarcopenia in older adults. Nutrients 2019 Mar 30;11(4):745. [doi: 10.3390/nu11040745] [Medline: 30935012]
53. Li L, Loo BP. Mobility impairment, social engagement, and life satisfaction among the older population in China: a structural equation modeling analysis. Qual Life Res 2017 May;26(5):1273-1282. [doi: 10.1007/s11136-016-1444-x] [Medline: 27796773]
54. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993 Apr;39(4):561-577. [Medline: 8472349]
55. Harrell FE, Lee KL, Mark DB. Prognostic/clinical prediction models: multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. In: Tutorials in Biostatistics, Statistical Methods in Clinical Studies: John Wiley & Sons, Ltd; 2004:223-249. [doi: 10.1002/0470023678.ch2b(i)]
56. Romero-Brufau S, Huddleston JM, Escobar GJ, Liebow M. Why the C-statistic is not informative to evaluate early warning scores and what metrics to use. Crit Care 2015 Aug 13;19(1):285. [doi: 10.1186/s13054-015-0999-1] [Medline: 26268570]

XSL•FO
RenderX

Abbreviations

ADL: activities of daily living
aMDS: alternate Mediterranean diet score
AUC: area under the receiver operating characteristic curve
DCA: decision curve analysis
EML: early mobility limitation
GNHS: Guangzhou Nutrition and Health Study
LASSO: least absolute shrinkage and selection operator
LightGBM: Light Gradient-Boosting Machine
LR: logistic regression
ML: machine learning
NNET: neural network
PPV: positive predictive value
RF: random forest
ROC: receiver operating characteristic
ROSE: Random Over-Sampling Examples
SHAP: Shapley Additive Explanations
STS: sit-to-stand
SVM: support vector machine
TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis

Edited by AF Abdulai; submitted 12.May.2025; peer-reviewed by S Chen, W Li; revised version received 02.Feb.2026; accepted 08.Mar.2026; published 11.May.2026.

Please cite as:
Freniche AC, Hu W, Chen M, Wang D, Breuillé D, Chen YM
Early Identification of Mobility Limitations in Community-Dwelling Middle-Aged and Older Adults: Development of a Prediction Model Based on a Prospective Cohort
JMIR Aging 2026;9:e77187
URL: https://aging.jmir.org/2026/1/e77187
doi:10.2196/77187

© Alberto Conde Freniche, Wei Hu, Mo Chen, Dantong Wang, Denis Breuillé, Yu-ming Chen. Originally published in JMIR Aging (https://aging.jmir.org), 11.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Novel Combined Audiovisual-Semantic Digital Tool for Early-Stage Cognitive Decline Detection: Development and Validation Study

Qiwei Wu[1,2], PhD; Binyu Zhao[1,2], PhD; Zihao Zheng[1,2], MS; Betsy Seah[3], PhD; Xueping Zhang[4], PhD; Jing Shao[1,2], PhD

[1]Department of Nursing, The Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, Hushan Road 200, Yiwu, Zhejiang Province, China

[2]Institute of Nursing Research, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

[3]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[4]Department of Geriatric Psychiatry, The Seventh People's Hospital of Hangzhou, Mental Health Center of Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

Corresponding Author:
Jing Shao, PhD
Department of Nursing, The Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, Hushan Road 200, Yiwu, Zhejiang Province, China

Abstract

Background: Dementia poses a significant public health challenge. Early detection is crucial for timely intervention but remains difficult in community settings, as many existing cognitive screening tools are either insufficiently sensitive to subtle decline or too burdensome for widespread use.

Objective: This study aimed to develop and validate a novel combined audiovisual-semantic digital tool for rapid and acceptable community-based screening of cognitive decline in older adults.

Methods: A total of 156 older adults completed 6 progressively complex task variants shifting from unimodal to multimodal stimuli and from basic to superordinate semantic categorization. Performance was measured using reaction time, accuracy, false alarms, and inverse efficiency. Diagnostic utility was assessed via logistic regression and receiver operating characteristic curve analysis, whereas qualitative interviews evaluated acceptability.

Results: Task outcomes showed significant declines across the cognitive continuum from no cognitive impairment to mild cognitive impairment to dementia (P<.001). The integrated audiovisual-semantic condition at the basic categorization level reliably differentiated cognitive groups and achieved higher area under the curve values for distinguishing no cognitive impairment from cognitive impairment and mild cognitive impairment from dementia compared to basic-level tasks. Incorporating semantic processing enhanced diagnostic discrimination. Participant feedback was highly positive, with 48.7% (76/156) describing the tasks as "fun/interesting."

Conclusions: The audiovisual-semantic integrated digital tool is a valid, well-accepted, and time-efficient instrument for cognitive screening in older adults. Its design, which increases cognitive load through multimodal and semantic integration, improves sensitivity to early decline, supporting its potential for practical community application.

(JMIR Aging 2026;9:e91165) doi:10.2196/91165

KEYWORDS

cognitive decline; working memory; audiovisual integration; semantic categorization; go/no-go task; community screening

Introduction

Dementia is one of the major causes of disability and dependency among older people, posing a significant public health challenge worldwide [1]. While dementia currently lacks effective treatment, it is understood to be a progressive continuum. There is cumulative evidence suggesting that early identification of and intervention for cognitive decline confer both clinical and health-economic benefits [2]. Notably, approximately 14.4% to 38% of individuals with mild cognitive impairment (MCI) who are identified through screening and receive interventions revert to normal cognitive functioning [3,4]. Consequently, screening tools with high specificity and sensitivity are critical. Current assessment methods, which primarily include physiological indicators, neuroimaging techniques, and questionnaires, are often invasive, costly, and time-consuming [5]. In contrast, some digital tools (eg, voice based or virtual reality [VR]) show substantial clinical promise

XSL•FO
RenderX

by addressing these shortcomings [6,7]. However, these new tools still have limitations (eg, cybersickness in VR and accent-related transcription errors) and require further exploration [5,6,8].

Impairment of working memory is a hallmark early symptom of dementia [9]. Working memory is a limited-capacity cognitive system that provides temporary storage and manipulation of information to support a wide range of complex activities such as reasoning, comprehension, and learning [10]. Neuroimaging studies have demonstrated altered activation in working memory–related brain regions during cognitive tasks in the early disease continuum [11,12]. However, translating these findings into early diagnostic tools remains a challenge as it hinges on deconstructing working memory into its measurable component processes. The multicomponent model of working memory by Baddeley [10] offers a simple yet robust framework that makes such deconstruction possible. This proposed framework comprises 4 core components: the central executive system, phonological loop, visuospatial sketchpad, and episodic buffer [13]. External sensory input is initially handled by the phonological loop and visuospatial sketchpad. The episodic buffer then binds this information with existing knowledge from long-term memory. Finally, the central executive system acts as the control system, achieving cognitive processes by regulating attention, updating information, and inhibiting irrelevant information [13]. Previous research has reported that inhibitory control represents the most frequently impaired executive domain in MCI [14]. Indeed, the central executive system has been confirmed as an interconnected network where inhibitory control relies on attention, inhibition, and other functions that are integrated across neural networks in an interactive manner [14]. Furthermore, the association between working memory capacity and higher-order cognitive abilities is more pronounced, with these differences largely attributed to attentional capacity [15]. Therefore, researchers believe that access to working memory is governed by inhibitory attention processes [16].

The go/no-go task is widely used as a measure of inhibition, attention, and impulsivity [17]. By designing tasks that require participants to respond to specific stimuli (“go”) and withhold responses to “no-go” stimuli, which assesses the ability of executive control, extensive behavioral and neuroimaging studies have demonstrated the neurobiological foundation of the go/no-go task [17-19]. Cognitive control–related tasks consistently elicit increased frontal midline θ power, which is a key neural marker for cognitive control, response inhibition, conflict monitoring, and attention allocation [18]. The increase in frontal midline θ power during go/no-go tasks is driven by both inhibitory and noninhibitory (attention-related) processing [19]. Furthermore, rapid response inhibition activates the prefrontal cortex (PFC), particularly the middle and inferior frontal gyri, as well as the parietal and temporal lobes and their functional networks [20]. Abnormalities in these regions are considered predictive of cognitive status [20]. Significantly, the PFC is particularly vulnerable to aging [21]. Some researchers have even proposed the frontal or executive decline hypothesis, suggesting that age-related cognitive decline may stem from damage to the prefrontal lobe and executive functions [22]. Therefore, the go/no-go task is an experimental paradigm that effectively reflects executive control and has been reported to be a potential practical early screening tool for cognitive decline [23,24].

According to the model by Baddeley [10], processing stimuli involves a multisensory operation that necessitates cross-modal integration within the episodic buffer. However, go/no-go tasks predominantly use visual stimuli [25]. Given that sensory function decline is a natural consequence of aging, relying on a single sensory modality may introduce measurement bias. Previous research has indicated that multisensory integration accelerates reaction time (RT) in older adults, potentially serving as a compensatory mechanism to enhance brain information processing efficiency [26]. Beyond facilitation effects, multisensory processing exhibits illusory effects such as the McGurk effect and the double-flash illusion [27]. These illusory effects are mediated by anatomical and functional connections between different brain regions, particularly the cross-modal associations involving neural resources in the PFC [28]. Cross-modal information is represented within the capacity-constrained episodic buffer, where more complex inputs entail the binding of a greater number of information sources, thereby elevating processing load [29]. Moreover, multisensory integration can be affected by top-down factors, also termed endogenous attention [30]. Endogenous attention requires purposeful and effort-intensive cognitive processing, whereas multisensory integration performance may exhibit instability or illusions under high–cognitive load conditions.

Semantic memory is a key component of long-term memory. Its integration with working memory occurs in the episodic buffer, followed by complex processing in the central executive system to support motor behavior. Prior research suggests that working memory is not related to basic sensory processing (eg, color or shape discrimination) but shows a reliable association with semantic tasks [31]. A meta-analysis revealed that semantic deficits constitute a crucial and early feature of MCI and should be incorporated into routine clinical assessments [32]. Furthermore, the complexity of conceptual semantic representations systematically influences inhibitory processing [33]. However, most published go/no-go studies measure perceptual discrimination of stimuli such as shapes or numbers, with little attention paid to changes in abstract concepts [34]. This neglect may result in insufficient sensitivity of go/no-go tasks in capturing early and subtle cognitive decline.

To sum up, based on the theoretical framework by Baddeley [10] and existing empirical evidence, we hypothesize that combining visual-auditory fusion and semantic categorization with the go/no-go task can heighten demand on central executive resources, which might lead to a better identification of cognitive decline. The aim of this study was to develop an integrated audiovisual-semantic go/no-go task and evaluate its performance in discriminating cognitive decline, as well as the task’s acceptability among older adults.

XSL•FO
RenderX

Methods

Study Design and Participants

This cross-sectional pilot study was conducted to preliminarily investigate and develop a digital tool integrating audiovisual-semantic stimuli for the detection of cognitive decline. From April 2025 to August 2025, a total of 156 participants aged 60 years and above were recruited from 7 communities in Hangzhou, China. Exclusion criteria were (1) individuals who had significant verbal and hearing impairments and were unable to see or hear the stimuli despite adjusting the volume or image size (by controlling the resolution in the E-Prime software [Psychology Software Tools]) and (2) those with major depression and other psychiatric disorders.

Data Collection

Overview

Data collection was conducted during a single session. All participants or their legally acceptable representatives signed a written consent form agreeing to take part in this study. After that, all participants underwent face-to-face interviews conducted by a trained staff member to collect the demographic information (eg, age and gender). We administered the Montreal Cognitive Assessment (MoCA) and the go/no-go tasks in an interleaved and alternating manner using a number-calling strategy, with at least a 30-minute interval between the 2 assessments. The MoCA assessment was conducted by trained research personnel.

Cognitive Assessment

The MoCA is a well-validated, multidomain, paper-and-pencil test that assesses a range of cognitive functions, including visuospatial and executive functions, naming, memory, attention, abstraction, and orientation [35]. We defined participants' cognitive status using the following MoCA cutoff scores (adjusted education; the MoCA should be scored based on educational level to reduce bias. According to the standard MoCA scoring rules, one point is added to the total score for participants with less than 6 years of education): a score of 23 or higher was categorized as no cognitive impairment (NCI), a score ranging from 14 to 22 was categorized as MCI, and a score of less than 14 was categorized as dementia [5].

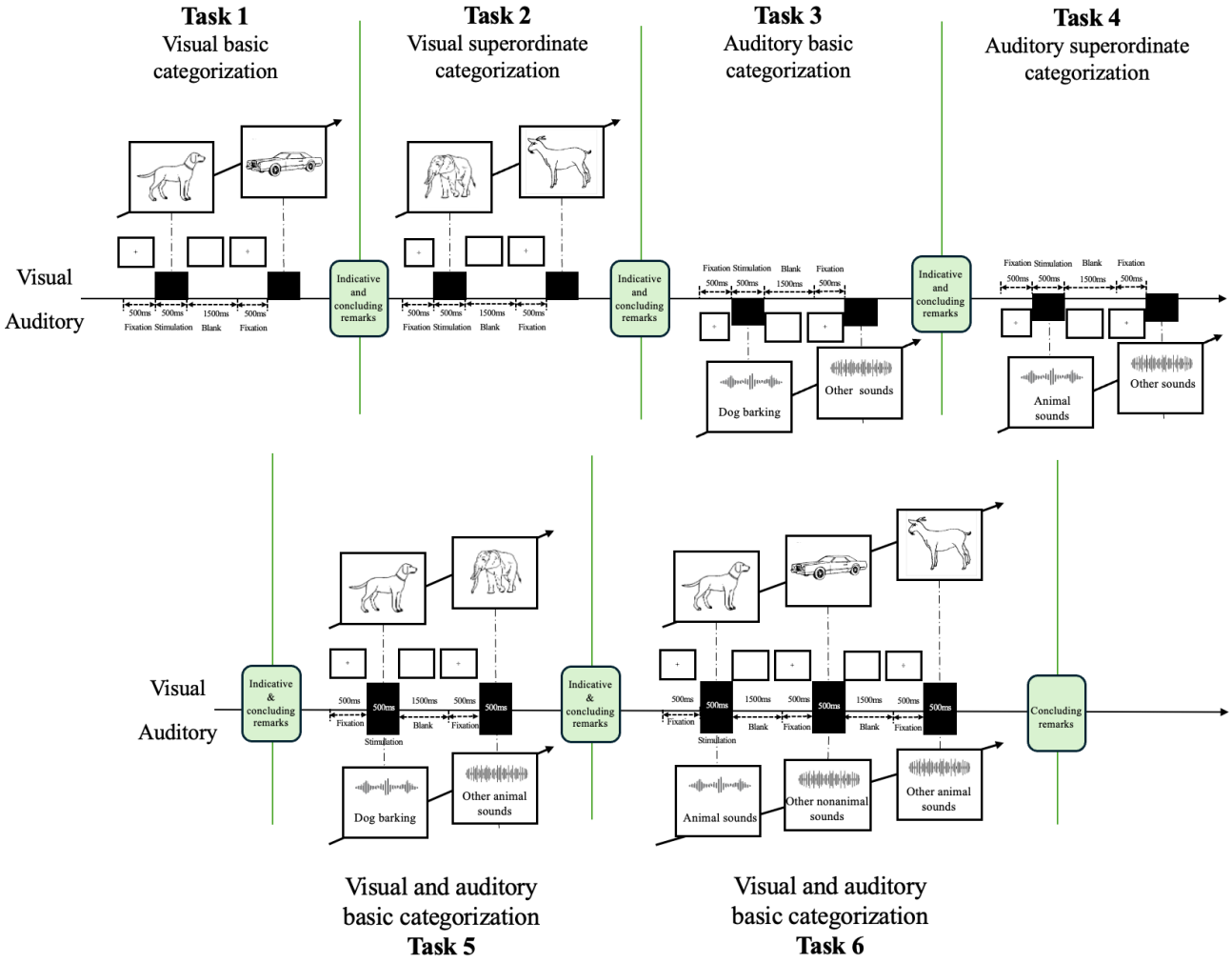
The Modified Go/No-Go Task

The task design was built on previous research and finalized through consensus among all authors [24]. This study comprised 6 go/no-go tasks (Figure 1), including visual, auditory, and audiovisual integration tasks. Every task consisted of 5 runs of 16 trials each (80 total trials). The total completion time for the 6 tasks was approximately 18 minutes (170 seconds per task). Participants were allowed a 1-minute break after every task. The tasks were conducted in a quiet environment (≤45 dB), verified using a decibel meter (model DLX-PS02303; Delixi Electric). Participants sat at approximately 58 cm from a 14-inch laptop computer (ThinkPad T480s-i7; Lenovo). Auditory stimuli were presented through wired headphones (model HD350BT; Sennheiser Electronic SE & Co). Volume was adjusted to a level clearly audible to participants before testing commenced. Visual materials were sourced from a canonical set of line drawings [36], and auditory materials were obtained from the Aigei website. As prior research has suggested that humans are more sensitive to biological sounds compared to nonbiological sounds, the auditory stimuli were designed as animal sound recognition tasks [37]. Visual and auditory stimulus materials underwent processing to standardize the image size and sound pressure level before being inserted into E-Prime (version 2.0). The sequence of stimuli in each task was fixed. A single-key keyboard (model K1-LY-WX; Yingyu IoT) was connected to the computer and positioned in front of the participant's dominant hand. Each stimulus was presented for 500 ms, with a 1500-ms intertrial interval during which a blank screen with a fixation cross was displayed. Each task was preceded by 12 practice trials at the same pace as the formal study trials (go/no-go ratio of 3:1) to ensure participant comprehension of task requirements. The go/no-go metrics evaluated were as follows:

- RT: mean response time for correct "go" trials
- "Go" trial accuracy: percentage of correct responses on "go" trials
- False alarm rate: percentage of incorrect responses on "no-go" trials
- Inverse efficiency score: RT divided by the overall proportion of correct responses

XSL•FO
RenderX

Figure 1. Go/no-go tasks.



A lower score reflects better inhibitory control by integrating speed and accuracy. All go/no-go metrics were standardized into *Z*-scores before performing regression analysis.

Informal Interview

Following completion of the MoCA and modified go/no-go tasks, we conducted informal interviews lasting approximately 3 to 5 minutes with participants. Semistructured questions included the following: “How did you feel about the overall testing process of Go/No-go?” “How was your experience with the Go/No-go task compared with the MoCA questionnaire?” “How do you feel about implementing this Go/No-go task for cognitive decline screening in a community setting?” Interviewers recorded key content words related to feelings, with each unique word recorded only once per participant.

Data Analysis

Statistical analyses were performed using SPSS (version 31; IBM Corp). As the data in this study did not pass normality tests, medians and IQRs were used for description. Kruskal-Wallis tests were used for demographic characteristics and go/no-go task–related metric analyses. Logistic regression analyses were conducted while controlling for age and educational level (P<.05 in both cases). We used backward likelihood ratio logistic regression analysis, with a P value of .05 for entry and .10 for removal. Backward likelihood ratio is considered more robust based on overall model fit changes [38].

The features extracted from the task were entered into a logistic regression model, where the diagnostic group (NCI vs cognitive impairment, MCI, or dementia and MCI vs dementia) was considered as the dependent variable and the task-extracted features were used as the independent variables.

To examine go/no-go diagnostic validity, receiver operating characteristic (ROC) curves were used to establish area under the curve (AUC) values for the following groups: NCI vs cognitive impairment (cognitive impairment including MCI and dementia), NCI vs MCI, NCI vs dementia, and MCI vs dementia. The optimal Youden index was used to calculate discrimination metrics, including sensitivity and specificity. AUC 95% CIs were calculated using bootstrap methods. To compare the discriminative validity of different tasks, additional ROC analyses were performed using the DeLong method. To analyze the qualitative data, the R package *wordcloud2* (R Foundation for Statistical Computing) was used to generate word clouds.

Ethical Considerations

All procedures were approved by the Ethics Committee for Human Studies of the Fourth Affiliated Hospital of Zhejiang University School of Medicine (2025-0326-1-F1). Written informed consent was obtained from all participants prior to measurements, in accordance with the Declaration of Helsinki. No identifying details are included in this paper or the

XSL•FO
RenderX

supplementary materials. Participants did not receive compensation. Access to raw data is restricted to research team members.

Results

Characteristics of the Participants

A total of 156 participants took part in this study, including those with NCI (n=75, 48.1%), MCI (n=66, 42.3%), and dementia (n=15, 9.6%). Baseline characteristics (Table 1) showed significant differences among the 3 groups in age, educational level, and MoCA scores (P<.001). No statistically significant differences were observed in gender or smoking or alcohol consumption habits across the 3 groups.

Table . Participant characteristics (N=156).

	NCI[a] (n=75)	MCI[b] (n=66)	Dementia (n=15)	Total	*P* value
Age (y), median (IQR)	70 (65-77)	76 (71-80)	82 (73-87)	74 (68-79)	<.001[c]
Sex (female), n (%)	23 (30.7)	13 (19.7)	1 (6.7)	37 (23.7)	.08[d]
Educational level, n (%)					<.001[d]
Illiterate	9 (12.0)	28 (42.4)	6 (40.0)	43 (27.6)	
Primary school	23 (30.7)	30 (45.5)	7 (46.7)	60 (38.5)	
Junior or senior high school	37 (49.3)	28 (42.4)	1 (6.7)	46 (29.5)	
College	6 (8.0)	0 (0.0)	1 (6.7)	7 (4.5)	
Smoking, n (%)					.57[d]
No	65 (86.7)	61 (92.4)	14 (93.3)	140 (89.7)	
Yes	7 (9.3)	3 (4.5)	10 (66.7)	10 (6.4)	
Quit	3 (4.0)	2 (3.0)	6 (40.0)	6 (3.8)	
Alcohol consumption, n (%)					.12[d]
No	69 (92.0)	65 (98.5)	15 (100.0)	149 (95.5)	
Yes	6 (8.0)	1 (1.5)	0 (0)	7 (4.5)	
MoCA[e] score (0-30), median (IQR)	26 (24-27)	19 (17-21)	9 (7-13)	22 (18-25)	<.001[c]

[a]NCI: no cognitive impairment.

[b]MCI: mild cognitive impairment.

[c]Kruskal-Wallis test.

[d]Fisher exact test.

[e]MoCA: Montreal Cognitive Assessment.

Task Performance and Logistic Regression

The results for task performance and logistic regression are shown in Table 2, Multimedia Appendix 1, and Figure 2.

XSL•FO
RenderX

Table . Logistic regression analysis of task variables for discriminating between cognitive groups (adjusted for educational level and age).

	NCI[a] vs cognitive impairment			NCI vs MCI[b]			NCI vs dementia			MCI vs dementia		
	Selected variable	β	OR[c] (95% CI)	Selected variable	β	OR (95% CI)	Selected variable	β	OR (95% CI)	Selected variable	β	OR (95% CI)
Task 1	NS[d]	NS	NS	NS	NS	NS	IES[e]_1[f]	0.940	2.561 (1.271-5.162)	IES_1[g]	0.563	1.757 (1.083-2.848)
Task 2	Go_RT[h]_2[f]	1.420	4.136 (1.876-9.119)	Go_RT_2[f]	1.323	3.755 (1.693-8.330)	Go_RT_2[f]	2.113	8.274 (2.207-31.016)	Go_RT_2[g]	0.570	1.769 (1.122-2.788)
Task 2	FAR[i]_2[g]	0.633	1.884 (1.112-3.192)	FAR_2[g]	0.602	1.825 (1.069-3.118)	—[j]	—	—	—	—	—
Task 3	IES_3[f]	2.159	8.666 (3.244-23.150)	IES_3[f]	2.185	8.891 (3.222-24.534)	IES_3[f]	2.923	16.665 (2.249-123.518)	IES_3[f]	2.278	9.761 (2.313-41.187)
Task 4	GTA[k]_4[f]	1.247	3.480 (1.630-7.433)	GTA_4[f]	1.123	3.074 (1.434-6.590)	GTA_4[f]	2.391	10.920 (1.817-65.629)	Go_RT_4[f]	2.746	15.578 (2.952-82.215)
Task 4	IES_4[f]	2.264	9.618 (3.632-25.469)	IES_4[f]	2.057	7.824 (2.886-21.209)	IES_4[f]	4.349	77.403 (8.050-744.251)	GTA_4[f]	1.721	5.589 (1.287-24.268)
Task 5	IES_5[f]	5.396	220.479 (17.180-2829.514)	IES_5[f]	5.242	189.081 (14.338-2493.478)	IES_5[g]	4.901	134.365 (7.004-2577.765)	Go_RT_5[f]	0.800	2.225 (1.346-3.678)
Task 5	FAR_5[g]	0.608	1.837 (1.010-3.341)	FAR_5[g]	0.691	1.996 (1.074-3.711)	—	—	—	—	—	—
Task 6	Go_RT_6[f]	1.914	6.781 (2.746-16.744)	Go_RT_6[f]	1.820	6.173 (2.483-15.348)	Go_RT_6[g]	1.914	6.781 (2.746-16.744)	FAR_6[f]	0.960	2.613 (1.331-5.131)

[a]NCI: no cognitive impairment.

[b]MCI: mild cognitive impairment.

[c]OR: odds ratio.

[d]NS: not significant.

[e]IES: inverse efficiency score.

[f]*P*<.01.

[g]*P*<.05.

[h]RT: reaction time.

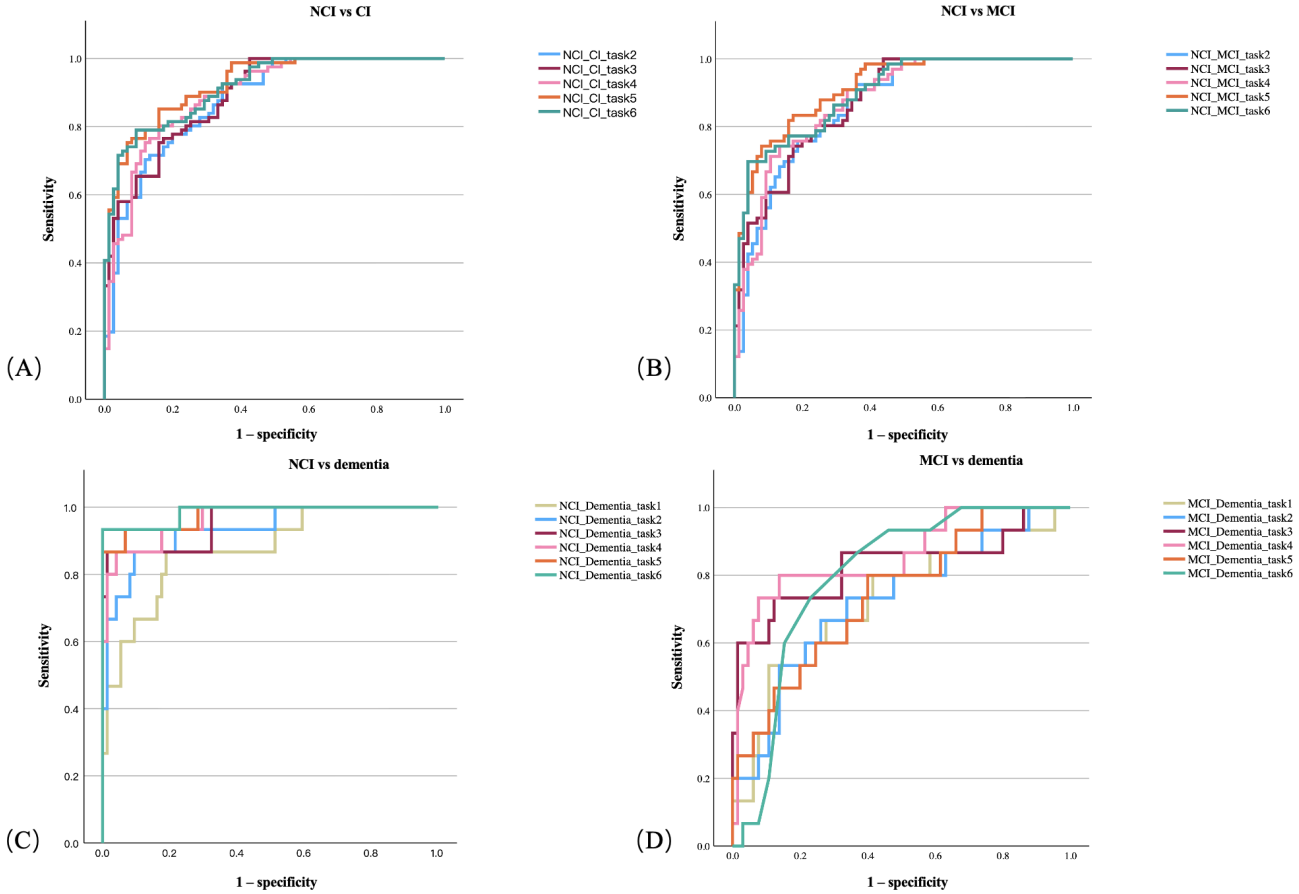
[i]FAR: false alarm rate.

[j]Not applicable.

[k]GTA: “go” trial accuracy.

XSL•FO
RenderX

Figure 2. Receiver operating characteristic curve performance of different tasks: (A) different task performance for identification of no cognitive impairment (NCI) vs cognitive impairment (CI), (B) different task performance for identification of NCI vs mild CI (MCI), (C) different task performance for identification of NCI vs dementia, and (D) different task performance for identification of MCI vs dementia.



Across all 6 go/no-go tasks, performance on the RT, “go” trial accuracy, inverse efficiency score, and false alarm rate metrics demonstrated a significant declining trend from NCI to dementia (Table 3), with no collinearity issues among metrics (variance inflation factor<10). Compared with participants with NCI, participants with both MCI and dementia exhibited slower RTs, lower accuracy, higher omission rates, and poorer inhibitory control (Table 3).

XSL•FO
RenderX

Table . Performance on go/no-go tasks across different cognitive groups.

	NCI[a], median (IQR)	MCI[b], median (IQR)	Dementia, median (IQR)	*P* value[c]
Task 1				
Go_RT[d]_1 (seconds)	510.28 (480.14-643.33)	515.22 (480.36-795.20)	601.63 (424.22-793.57)	.049
GTA[e]_1 (%)	93.33 (83.33-96.67)	90.00 (82.09-95.00)	86.67 (80.42-94.58)[f]	.03
IES[g]_1	553.86 (512.15-856.74)	606.93 (521.49-1010.24)	729.39 (529.29-1154.87)[f]	.003
FAR[h]_1 (%)	15.00 (10.00-25.00)	20.00 (10.00-30.00)	42.50 (20.00-52.50)[f,i]	.01
Task 2				
Go_RT_2 (seconds)	532.08 (470.87-619.95)	656.22 (562.44-836.66)[f]	705.88 (633.57-1239.25)[f,i]	<.001
GTA_2 (%)	91.67 (85.00-95.00)	85.00 (75.00-94.17)[f]	49.17 (31.25-89.17)[f,i]	<.001
IES_2	600.91 (523.75-761.38)	850.76 (617.57-1355.29)[f]	1751.80 (711.22-3061.99)[f,i]	<.001
FAR_2 (%)	10.00 (5.00-25.00)	20.00 (10.00-45.00)[f]	32.50 (7.50-53.75)[f]	<.001
Task 3				
Go_RT_3 (seconds)	734.56 (701.34-789.56)	880.23 (840.84-922.59)[f]	1026.96 (987.95-1105.29)[f,i]	<.001
GTA_3 (%)	90.00 (85.00-95.00)	71.67 (68.33-77.50)[f]	45.84 (42.92-50.83)[f,i]	<.001
IES_3	838.44 (787.68-910.63)	1318.70 (1208.50-1460.56)[f]	2221.76 (2027.79-2664.91)[f,i]	<.001
FAR_3 (%)	20.00 (7.50-30.00)	50.00 (40.00-62.50)[f]	57.50 (48.75-61.25)[f]	<.001
Task 4				
Go_RT_4 (seconds)	789.45 (745.56-812.73)	930.78 (896.55-975.40)[f]	1026.56 (980.53-1085.51)[f,i]	<.001
GTA_4 (%)	80.00 (72.50-84.17)	68.33 (63.33-71.67)[f]	66.67 (59.58-71.25)[f]	<.001
IES_4	1002.48 (937.43-1135.79)	1490.56 (1389.44-1691.61)[f]	1623.99 (1505.49-1958.68)[f,i]	<.001
FAR_4 (%)	30.00 (15.00-45.00)	55.00 (45.00-65.00)[f]	52.50 (45.00-56.25)[f]	<.001
Task 5				
Go_RT_5 (seconds)	593.53 (513.02-681.63)	705.52 (592.79-879.11)[f]	989.47 (852.59-1576.23)[f,i]	<.001
GTA_5 (%)	93.33 (85.00-96.67)	81.67 (71.67-95.00)[f]	53.33 (23.75-84.59)[f,i]	<.001
IES_5	677.63 (588.29-833.37)	1036.75 (729.19-1205.72)[f]	2244.61 (1116.53-4251.07)[f,i]	<.001
FAR_5 (%)	20.00 (10.00-35.00)	30.00 (10.00-50.00)[f]	42.50 (17.50-70.00)	<.001
Task 6				
Go_RT_6 (seconds)	767.34 (709.16-807.31)	920.67 (885.40-980.56)[f]	1121.00 (1087.95-1158.06)	<.001
GTA_6 (%)	81.67 (76.67-85.84)	68.33 (61.67-73.33)[f]	53.34 (47.92-58.75)[f]	<.001
IES_6	975.92 (912.57-1068.22)	1449.79 (1329.19-1716.65)[f]	2134.88 (2021.61-2437.72)[f]	<.001
FAR_6 (%)	25.00 (17.50-40.00)	45.00 (35.00-55.00)[f]	55.00 (50.00-56.25)[f,i]	<.001

XSL•FO
RenderX

[a]NCI: no cognitive impairment.
[b]MCI: mild cognitive impairment.
[c]Kruskal-Wallis test.
[d]RT: reaction time.
[e]GTA: "go" trial accuracy.
[f]Compared with NCI; *P*<.05.
[g]IES: inverse efficiency score.
[h]FAR: false alarm rate.
[i]Compared with MCI; *P*<.05.

ROC Results

As shown in Table 2 and Figure 2, we examined the discriminatory ability of different tasks for varying degrees of cognitive decline. The AUC ranges for the diagnostic comparisons were 0.877 to 0.926 for NCI vs cognitive impairment, 0.861 to 0.916 for NCI vs MCI, 0.909 to 0.968 for NCI vs dementia, and 0.724 to 0.855 for MCI vs dementia. Considering accuracy, sensitivity, specificity, and AUC simultaneously, tasks 5 and 6 demonstrated optimal performance in discriminating NCI from cognitive impairment compared to other discrimination models. Task 5 exhibited the most stable performance in discriminating NCI from cognitive impairment, MCI, and dementia (accuracy: 0.837-0.933; sensitivity: 0.742-0.933; specificity: 0.606-0.933; AUC: 0.744-0.977).

In the DeLong test for paired ROC curves, the AUC of task 5 was significantly higher than that of task 2 in discriminating NCI from cognitive impairment (Z=–2.129; ΔAUC=–0.048, 95% CI –0.093 to −0.004; P=.03) and from MCI (Z=–2.182; ΔAUC=–0.055, 95% CI –0.104 to –0.006; P=.03). The AUC of task 2 significantly outperformed task 1 when discriminating NCI from dementia (Z=–2.05; ΔAUC=–0.049, 95% CI –0.095 to –0.002; P=.04).

Acceptability

In informal interviews, "fun/interesting" emerged as the most frequently mentioned term (76/156, 48.7%), followed by "thinking" (65/156, 41.7%) and "high-tech vibe" (61/156, 39.1%). Participants also expressed perceptions of task difficulty, such as "easy" (57/156, 36.5%), "lagging" (43/156, 27.6%), and "difficult" (13/156, 8.3%). When comparing with standardized scales, participants mentioned the terms "engaging" (43/156, 27.6%) and "novelty" (30/156, 19.2%), with practical application perceptions including "convenient" and "roll out" (Multimedia Appendix 2).

Discussion

Principal Findings

Overview

This study provides empirical evidence that the increased cognitive load imposed by integrating visual-auditory stimuli and semantic categorization into a go/no-go task effectively amplifies performance differences across cognitive states, thereby supporting its utility as a discriminative tool. This novel approach indirectly assessed the episodic buffer's ability to form coherent episodic representations, enabling early detection of cognitive decline. Our findings showed that the task complexity incrementally increased from unimodal stimuli (visual or auditory) to multimodal stimuli (audiovisual integration) and from a basic level (concrete objects such as "dog") to a superordinate level (abstract categories such as "animal"), consistent with our hypothesis. Older adults with cognitive decline exhibited progressively worse performance in these incrementally difficult tasks.

Measuring Using Multimodal Stimuli

Unimodal tasks (task 1 and 3), which relied on basic sensory discrimination, did not demonstrate superior overall performance in discriminating between cognitive statuses. Such an observation is likely due to dependence on automatic processing and low-level perception, yielding similar performance across groups. In contrast, both cross-modal matching tasks (tasks 5 and 6) performed well; in particular, they outperformed the auditory-only MoCA-attention cognitive domain (AUC=88.7%, 95% CI 88.0% - 89.4%), suggesting that multisensory integration is a reliable biomarker for cognitive impairment [39]. This is supported by findings from The Irish Longitudinal Study on Ageing, which reported broad associations between multisensory integration and multiple cognitive domains rather than links to any specific subdomain [40]. This implies that measuring multisensory integration could serve as a window into global cognition. Furthermore, as noted previously, multisensory stimulation may serve as a compensatory strategy for older adults, enhancing behavioral performance in those with NCI [26]. However, this compensatory mechanism is limited as behavioral responses are governed by the central executive system, which may account for the more pronounced deficits observed in individuals with progressive cognitive decline. Bayesian integration theory further supports this premise, proposing that perceptual reliability weighting affects audiovisual integration. In other words, the brain optimally combines sensory inputs based on their relative reliability, but cognitive decline may change this weighting process [41]. Apart from that, some scholars regard temporal synchrony and functional connectivity as 2 primary mechanisms that facilitate audiovisual integration [42]. The integration time window posits 2 sequential, interdependent stages: initial parallel peripheral processing of stimuli followed by central processes involving audiovisual integration and motor response initiation [43]. This process may help explain impairments observed in cognitive decline. Previous research indicates that patients with MCI and Alzheimer disease exhibit delayed audiovisual integration compared to older adults with normal cognitive function, with MCI populations perceiving more sound-induced flash illusions [27]. However, patients with MCI and Alzheimer disease experience attentional deficits, and these patients become

XSL•FO
RenderX

distracted by irrelevant information, demonstrating impaired attention-switching speed [44]. This aligns with the theory by Baddeley [10]. Impaired performance of the episodic buffer may contribute to working memory deficit, particularly in cognitively declining populations [10]. Therefore, we propose that assessment methods involving multisensory processing can evaluate central executive performance and better identify cognitive decline.

Task Complexity

Task 5, rather than the more cognitively demanding task 6, exhibited the most stable discriminative performance across cognitive states. This suggests that excessive complexity may overwhelm the limitations of the episodic buffer and central executive system. A study indicated that older adults tend to adopt conservative action strategies when performing complex tasks, reflecting speed-accuracy trade-offs. Specifically, they adjust their decision boundaries to increase response caution, which results in prolonged RT as a means to enhance or maintain accuracy [45]. However, a study by Chen et al [46] showed different results, with MCI and dementia populations potentially exhibiting random performance without deliberation, resulting in shorter RTs when compared to normal groups. Nonetheless, the adaptive training trials at the beginning of each task in this study could, to some extent, strengthen participants' understanding of the tasks and avoid performance variability. However, such training for task 6, which placed the highest demands on central executive resources, may still have been insufficient, leading participants to adopt response strategies that influenced the outcomes. According to the Yerkes-Dodson law (an arousal-modulated mechanism) [47], task 5 likely reaches the peak of the inverted U-shaped curve, enhancing arousal and performance discrimination, whereas the increased cognitive load imposed by task 6 exceeds this optimum, resulting in overload and diminished utility. This may also explain why task 5 demonstrated greater stability than task 6. The findings further suggest that task difficulty must be carefully calibrated to avoid overloading the episodic buffer, which could undermine measurement consistency. Additionally, the fixed task order (task 1 to task 6) used in this study means that the decline in performance on task 6 may be partially attributable to time-on-task effects, manifested as decreased attention and less effective decision-making [48]. Previous research has shown that tasks shorter than 15 minutes are unlikely to induce significant fatigue, whereas subjective fatigue increases after 20 minutes of task engagement [49]. Although rest breaks were provided between tasks, this potential confounder should be considered when interpreting the results. Future studies should use task randomization or repeated measurements to rule out this alternative explanation.

Prospective Clinical Application

At present, the go/no-go task is typically used as an experimental paradigm in conjunction with other objective indicators and has not been widely adopted for screening purposes. In contrast, traditional instruments such as the Mini-Mental State Examination and MoCA require 5 to 15 minutes for administration [24,35]. These traditional tools require trained assessors and additional time and workload and could introduce potential bias from personnel differences [5]. Regarding diagnostic accuracy, our tool demonstrated a performance comparable to that of emerging digital screening instruments, including VR and voice-based tools (AUC=0.83-0.89; sensitivity=0.79-0.83; specificity=0.67-0.91) [7,8,50]. Nevertheless, in comparison with these tools, our novel tool effectively retains the diagnostic benefits of multisensory stimulation while eliminating adverse physical effects and high scene production costs [5-7].

Notably, after accounting for the influence of educational level, the tasks retained their independent predictive value. Within the tasks with independent predictive value, task 5, as the best-performing task, can be completed within 3 minutes. Thus, the use of task 5 as a stand-alone tool for efficient and rapid screening in community settings is highly promising. Task 5 can be integrated into various digital platforms (eg, smartphones or tablets) to quickly flag at-risk individuals for follow-up assessments and can also be deployed as a self-administered tool, thereby enabling timely interventions in resource-limited environments.

Results from the informal interview with community-dwelling older adults showed a positive perception of the modified go/no-go task, with nearly half of the participants describing it as interesting. This favorable experience is important for promoting community-based screening and aligns with previous reports of older adults' increasing willingness to use digital tools [51]. Additionally, participants' subjective evaluations of task difficulty were consistent with the study sample proportions of participants with NCI and dementia, indicating that the task's difficulty gradient was perceptible.

Limitations

The uneven distribution of participants with dementia, which accounted for 9.6% (15/156) of the total sample, might lead to wide CIs and present a risk of overfitting. This may be related to our community-based data collection. This potential bias calls for future validation studies expanding the sample size of older adults with dementia. Additionally, this study did not repeat measures for test-retest reliability. Therefore, we are currently conducting a follow-up study to fully assess the reliability of the modified go/no-go task. Furthermore, as a cross-sectional study, our findings demonstrate associations but cannot determine the tool's predictive validity for future dementia onset, which requires longitudinal investigation.

Conclusions

This study modified and validated a novel go/no-go task incorporating audiovisual stimuli and semantic categorization as an effective tool for early detection of cognitive decline. Positive participant feedback on engagement and novelty supports its feasibility. The applicability and reliability of the modified go/no-go task warrant further investigation in a larger-scale community-based setting.

XSL•FO
RenderX

Acknowledgments

The authors thank the participants for their time and willingness to share their experiences. The authors greatly appreciate the assistance with grammar provided by Google Gemini 3.

Funding

This work was supported by the "Pioneer" and "Leading Goose" Research and Development Programs of Zhejiang (grant 2026C02A1144), the Natural Science Foundation of Zhejiang Province (Y23G030061), the National Key Research and Development Program (grant 2022YFF0902000), the National Natural Science Foundation of China (grant 72004193), and the China Scholarship Council program (grant 202506320175).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Diagnostic performance of cognitive tasks for differentiating between cognitive groups.
[DOCX File, 19 KB - aging_v9i1e91165_app1.docx]

Multimedia Appendix 2
Task perception word cloud map.
[PNG File, 582 KB - aging_v9i1e91165_app2.png]

References

1. GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 May;18(5):459-480. [doi: 10.1016/S1474-4422(18)30499-X] [Medline: 30879893]
2. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet 2024 Aug 10;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
3. Zhao Y, Wei S, Liu Y, et al. The prevalence and influencing factors of reversion from mild cognitive impairment to normal cognition: a systemic review and meta-analysis. Geriatr Nurs 2025;63:379-387. [doi: 10.1016/j.gerinurse.2025.03.013] [Medline: 40245823]
4. Petersen RC, Lopez O, Armstrong MJ, et al. Practice guideline update summary: mild cognitive impairment [RETIRED]: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology 2018 Jan 16;90(3):126-135. [doi: 10.1212/WNL.0000000000004826] [Medline: 29282327]
5. Zhao X, Hu R, Wen H, et al. A voice recognition-based digital cognitive screener for dementia detection in the community: development and validation study. Front Psychiatry 2022;13:899729. [doi: 10.3389/fpsyt.2022.899729] [Medline: 35935417]
6. Huang L, Yang H, Che Y, Yang J. Automatic speech analysis for detecting cognitive decline of older adults. Front Public Health 2024;12:1417966. [doi: 10.3389/fpubh.2024.1417966] [Medline: 39175901]
7. Liu Q, Song H, Yan M, et al. Virtual reality technology in the detection of mild cognitive impairment: a systematic review and meta-analysis. Ageing Res Rev 2023 Jun;87:101889. [doi: 10.1016/j.arr.2023.101889] [Medline: 36806377]
8. Huang L, Li Y, Wu J, Chen N, Xia H, Guo Q. Shanghai Cognitive Screening: a mobile cognitive assessment tool using voice recognition to detect mild cognitive impairment and dementia in the community. J Alzheimers Dis 2023;95(1):227-236. [doi: 10.3233/JAD-230277] [Medline: 37482999]
9. Huntley JD, Howard RJ. Working memory in early Alzheimer's disease: a neuropsychological review. Int J Geriatr Psychiatry 2010 Feb;25(2):121-132. [doi: 10.1002/gps.2314] [Medline: 19672843]
10. Baddeley A. Working memory: theories, models, and controversies. Annu Rev Psychol 2012;63:1-29. [doi: 10.1146/annurev-psych-120710-100422] [Medline: 21961947]
11. Cheng CH, Hung CC, Chao YP, Nouchi R, Wang PN. Subjective cognitive decline exhibits alterations of resting-state phase-amplitude coupling in precuneus. Clin Neurophysiol 2023 Dec;156:281-289. [doi: 10.1016/j.clinph.2023.08.015] [Medline: 37722986]
12. Ben-Ami L, Ravona-Springer R, Tsarfaty G, et al. Neural correlates of subjective cognitive decline in adults at high risk for Alzheimer's disease. Front Aging Neurosci 2023;15:1067196. [doi: 10.3389/fnagi.2023.1067196] [Medline: 36819726]
13. Hitch GJ, Allen RJ, Baddeley AD. The multicomponent model of working memory fifty years on. Q J Exp Psychol (Hove) 2025 Feb;78(2):222-239. [doi: 10.1177/17470218241290909] [Medline: 39340360]
14. Happer JP, Wagner LC, Beaton LE, Rosen BQ, Marinkovic K. The "when" and "where" of the interplay between attentional capture and response inhibition during a Go/NoGo variant. Neuroimage 2021 May 1;231:117837. [doi: 10.1016/j.neuroimage.2021.117837] [Medline: 33577939]

XSL•FO
RenderX

15. Shipstead Z, Lindsey DR, Marshall RL, Engle RW. The mechanisms of working memory capacity: primary memory, secondary memory, and attention control. J Mem Lang 2014 Apr;72:116-141. [doi: 10.1016/j.jml.2014.01.004]
16. Tay D, McDonald JJ. Attentional enhancement predicts individual differences in visual working memory under go/no-go search conditions. PLoS Biol 2022 Nov;20(11):e3001917. [doi: 10.1371/journal.pbio.3001917] [Medline: 36441827]
17. Chikazoe J. Localizing performance of go/no-go tasks to prefrontal cortical subregions. Curr Opin Psychiatry 2010 May;23(3):267-272. [doi: 10.1097/YCO.0b013e3283387a9f] [Medline: 20308899]
18. Cavanagh JF, Frank MJ. Frontal theta as a mechanism for cognitive control. Trends Cogn Sci 2014 Aug;18(8):414-421. [doi: 10.1016/j.tics.2014.04.012]
19. Hong X, Sun J, Wang J, Li C, Tong S. Attention-related modulation of frontal midline theta oscillations in cingulate cortex during a spatial cueing Go/NoGo task. Int J Psychophysiol 2020 Feb;148:1-12. [doi: 10.1016/j.ijpsycho.2019.11.011] [Medline: 31857191]
20. Swick D, Ashley V, Turken U. Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. Neuroimage 2011 Jun 1;56(3):1655-1665. [doi: 10.1016/j.neuroimage.2011.02.070] [Medline: 21376819]
21. McEwen BS, Morrison JH. The brain on stress: vulnerability and plasticity of the prefrontal cortex over the life course. Neuron 2013 Jul 10;79(1):16-29. [doi: 10.1016/j.neuron.2013.06.028] [Medline: 23849196]
22. Luszcz M. Executive function and cognitive aging. In: Schaie KW, Willis SL, editors. Handbook of the Psychology of Aging: Elsevier Academic Press; 2011:59-72. [doi: 10.1016/B978-0-12-380882-0.00004-8]
23. Schmidt GJ, Boechat YE, van Duinkerken E, et al. Detection of cognitive dysfunction in elderly with a low educational level using a reaction-time attention task. J Alzheimers Dis 2020;78(3):1197-1205. [doi: 10.3233/JAD-200881] [Medline: 33136095]
24. Watanabe N, Kamijo M, Nishino T, et al. The feasibility of using the go/no-go task as a dementia screening test assessed with a cross-sectional design. Sci Rep 2024 Dec 1;14(1):29834. [doi: 10.1038/s41598-024-81301-5] [Medline: 39617763]
25. Simmonds DJ, Pekar JJ, Mostofsky SH. Meta-analysis of go/no-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. Neuropsychologia 2008 Jan 15;46(1):224-232. [doi: 10.1016/j.neuropsychologia.2007.07.015] [Medline: 17850833]
26. Freiherr J, Lundström JN, Habel U, Reetz K. Multisensory integration mechanisms during aging. Front Hum Neurosci 2013 Dec 13;7:863. [doi: 10.3389/fnhum.2013.00863] [Medline: 24379773]
27. Tang X, Wu J, Shen Y. The interactions of multisensory integration with endogenous and exogenous attention. Neurosci Biobehav Rev 2016 Feb;61:208-224. [doi: 10.1016/j.neubiorev.2015.11.002] [Medline: 26546734]
28. Fuster JM, Bodner M, Kroger JK. Cross-modal and cross-temporal association in neurons of frontal cortex. Nature 2000 May 18;405(6784):347-351. [doi: 10.1038/35012613] [Medline: 10830963]
29. Vandierendonck A. A working memory system with distributed executive control. Perspect Psychol Sci 2016 Jan;11(1):74-100. [doi: 10.1177/1745691615596790] [Medline: 26817727]
30. Macaluso E. Orienting of spatial attention and the interplay between the senses. Cortex 2010 Mar;46(3):282-297. [doi: 10.1016/j.cortex.2009.05.010] [Medline: 19540475]
31. Wiemers EA, Redick TS. Task manipulation effects on the relationship between working memory and go/no-go task performance. Conscious Cogn 2019 May;71:39-58. [doi: 10.1016/j.concog.2019.03.006] [Medline: 30928898]
32. Joubert S, Gardy L, Didic M, Rouleau I, Barbeau EJ. A meta-analysis of semantic memory in mild cognitive impairment. Neuropsychol Rev 2021 Jun;31(2):221-232. [doi: 10.1007/s11065-020-09453-5] [Medline: 32815030]
33. Maguire MJ, Brier MR, Moore PS, et al. The influence of perceptual and semantic categorization on inhibitory processing as measured by the N2-P3 response. Brain Cogn 2009 Dec;71(3):196-203. [doi: 10.1016/j.bandc.2009.08.018] [Medline: 19773108]
34. Mudar RA, Chiang HS, Maguire MJ, et al. Effects of age on cognitive control during semantic categorization. Behav Brain Res 2015;287:285-293. [doi: 10.1016/j.bbr.2015.03.042] [Medline: 25823764]
35. Cecato JF, Martinelli JE, Izbicki R, Yassuda MS, Aprahamian I. A subtest analysis of the Montreal cognitive assessment (MoCA): which subtests can best discriminate between healthy controls, mild cognitive impairment and Alzheimer's disease? Int Psychogeriatr 2016 May;28(5):825-832. [doi: 10.1017/S1041610215001982] [Medline: 26620850]
36. Snodgrass JG, Vanderwart M. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. J Exp Psychol Hum Learn 1980 Mar;6(2):174-215. [doi: 10.1037//0278-7393.6.2.174] [Medline: 7373248]
37. Tovar DA, Murray MM, Wallace MT. Selective enhancement of object representations through multisensory integration. J Neurosci 2020 Jul 15;40(29):5604-5615. [doi: 10.1523/JNEUROSCI.2139-19.2020] [Medline: 32499378]
38. Xu RF. Selection for independent variables and regression method in logistic regression: an example analysis. Chin J Evid Based Med 2016;16(11):1360-1364. [doi: 10.7507/1672-2531.20160205]
39. Liew TM. A normative calculator for MoCA domain scores: proxy for Z-scores of conventional neuropsychological tests. Alzheimers Res Ther 2025 Jul 21;17(1):167. [doi: 10.1186/s13195-025-01810-x] [Medline: 40696367]

XSL•FO
RenderX

40. Hirst RJ, Setti A, De Looze C, Kenny RA, Newell FN. Multisensory integration precision is associated with better cognitive performance over time in older adults: a large-scale exploratory study. Aging Brain 2022;2:100038. [doi: 10.1016/j.nbas.2022.100038] [Medline: 36908873]
41. Körding KP, Wolpert DM. Bayesian integration in sensorimotor learning. Nature 2004 Jan 15;427(6971):244-247. [doi: 10.1038/nature02169] [Medline: 14724638]
42. Bauer AK, Debener S, Nobre AC. Synchronisation of neural oscillations and cross-modal influences. Trends Cogn Sci 2020 Jun;24(6):481-495. [doi: 10.1016/j.tics.2020.03.003] [Medline: 32317142]
43. Diederich A, Colonius H, Schomburg A. Assessing age-related multisensory enhancement with the time-window-of-integration model. Neuropsychologia 2008 Aug;46(10):2556-2562. [doi: 10.1016/j.neuropsychologia.2008.03.026] [Medline: 18490033]
44. Perry RJ, Watson P, Hodges JR. The nature and staging of attention dysfunction in early (minimal and mild) Alzheimer's disease: relationship to episodic and semantic memory impairment. Neuropsychologia 2000;38(3):252-271. [doi: 10.1016/s0028-3932(99)00079-2] [Medline: 10678692]
45. Revie L, Metzler-Baddeley C. Age-related fornix decline predicts conservative response strategy-based slowing in perceptual decision-making. Aging Brain 2024;5:100106. [doi: 10.1016/j.nbas.2024.100106] [Medline: 38318456]
46. Chen YT, Hou CJ, Derek N, Huang SB, Huang MW, Wang YY. Evaluation of the reaction time and accuracy rate in normal subjects, MCI, and dementia using serious games. Appl Sci 2021;11(2):628. [doi: 10.3390/app11020628]
47. Tong C, Li W, Zou Y, et al. Norepinephrine-mediated arousal fluctuations drive inverted U-shaped functional connectivity dynamics. Nat Commun 2025 Dec 13;16(1):11318. [doi: 10.1038/s41467-025-66436-x] [Medline: 41390822]
48. Roose MM, Schampheleer E, Habay J, et al. Mental fatigue and brain activation using prolonged task-based fMRI: a systematic review on time-on-task and sequential task paradigms. Exp Brain Res 2026 Jan 11;244(2):24. [doi: 10.1007/s00221-025-07222-5] [Medline: 41521336]
49. Brahms M, Heinzel S, Rapp M, et al. The acute effects of mental fatigue on balance performance in healthy young and older adults - a systematic review and meta-analysis. Acta Psychol (Amst) 2022 May;225:103540. [doi: 10.1016/j.actpsy.2022.103540] [Medline: 35245722]
50. Zhao X, Wen H, Xu G, et al. Validity, feasibility, and effectiveness of a voice-recognition based digital cognitive screener for dementia and mild cognitive impairment in community-dwelling older Chinese adults: a large-scale implementation study. Alzheimers Dement 2024 Apr;20(4):2384-2396. [doi: 10.1002/alz.13668] [Medline: 38299756]
51. Ma T, Zhang S, Zhu S, Ni J, Wu Q, Liu M. The new role of nursing in digital inclusion: reflections on smartphone use and willingness to increase digital skills among Chinese older adults. Geriatr Nurs 2022;48:118-126. [doi: 10.1016/j.gerinurse.2022.09.004] [Medline: 36155310]

Abbreviations

AUC: area under the curve
MCI: mild cognitive impairment
MoCA: Montreal Cognitive Assessment
NCI: no cognitive impairment
PFC: prefrontal cortex
ROC: receiver operating characteristic
RT: reaction time
VR: virtual reality

Edited by R Lin; submitted 09.Jan.2026; peer-reviewed by D Zeng, Q Xie; revised version received 26.Feb.2026; accepted 17.Mar.2026; published 05.May.2026.

Please cite as:
Wu Q, Zhao B, Zheng Z, Seah B, Zhang X, Shao J
A Novel Combined Audiovisual-Semantic Digital Tool for Early-Stage Cognitive Decline Detection: Development and Validation Study
JMIR Aging 2026;9:e91165
URL: https://aging.jmir.org/2026/1/e91165
doi:10.2196/91165

© Qiwei Wu, Binyu Zhao, Zihao Zheng, Betsy Seah, Xueping Zhang, Jing Shao. Originally published in JMIR Aging (https://aging.jmir.org), 5.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any

XSL•FO
RenderX

medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Handwriting Kinematics and Physical Performance According to Pentagon-Copy Performance in Community-Dwelling Older Adults: Cross-Sectional Study

Pau Ferrer-Ramos[1,2,3], MSc; Marcos Faundez-Zanuy[4], PhD; Noemí Serra-Payá[5], PhD; Manuel Vicente Garnacho-Castaño[5,6], PhD; Montserrat Girabent-Farrés[5], PhD

[1]Faculty of Psychology, Education and Sport Sciences, Blanquerna, Universitat Ramon Llull, Barcelona, Catalonia, Spain

[2]VITAE Escola Universitària de l'Esport, Universitat Abat Oliba CEU, Barcelona, Catalonia, Spain

[3]Department of Health Sciences, Research Group in Technology Applied to High Performance and Health, TecnoCampus, Universitat Pompeu Fabra, Mataró, Catalonia, Spain

[4]TecnoCampus, Universitat Pompeu Fabra, Mataró, Catalonia, Spain

[5]DAFNiS Research Group (Pain, Physical Activity, Nutrition and Health), Campus Docent Sant Joan de Déu, Universitat de Vic - Universitat Central de Catalunya (UVic-UCC), Sant Boi de Llobregat, Catalonia, Spain

[6]Facultad de Ciencias de la Salud, Universidad Internacional de Valencia, Valencia, Spain

Corresponding Author:
Noemí Serra-Payá, PhD
DAFNiS Research Group (Pain, Physical Activity, Nutrition and Health)
Campus Docent Sant Joan de Déu
Universitat de Vic - Universitat Central de Catalunya (UVic-UCC)
C/ Sant Benito Menni, 18-20,
Sant Boi de Llobregat, Catalonia, 08830
Spain
Phone: 34 932800949
Email: noemi.serra@sjd.edu.es

Abstract

Background: Cognitive decline in older adults is often accompanied by subtle motor alterations. Digital handwriting analysis has emerged as a promising noninvasive approach for detecting these changes, but its usefulness in community-based settings remains unclear.

Objective: This study aims to examine the association of handwriting kinematics and physical performance measures with pentagon-copy performance classification in community-dwelling older adults.

Methods: This cross-sectional study included 174 community-dwelling adults aged 60 years or older (mean age 73.9, SD 6.1 years; 108/174, 62% women). Participants completed 10 digital handwriting tasks and a battery of physical performance tests assessing strength, balance, gait, and cardiorespiratory fitness. Group classification was based on pentagon-copy performance and categorized as normal (93/174, 53% participants) or altered (81/174, 46% participants). Adjusted linear regression models included group as the main predictor and age and sex as covariates. Multiple comparisons were controlled using the Benjamini-Hochberg false discovery rate.

Results: After adjustment for age, sex, and multiple comparisons, selected handwriting variables remained significantly associated with altered pentagon-copy performance, whereas no physical performance variables remained statistically significant. The most consistent differences were observed in cognitive effort and mechanical tasks, where participants with altered pentagon-copy performance showed longer contact time (β=526.8 ms; *P*<.001) and time on air (β=1111.5 ms; *P*<.001), together with lower mean writing pressure (β=–2058.8 au; *P*=.003). Overall, group differences were more consistently detected in handwriting-derived variables than in conventional physical performance outcomes after adjusted analyses.

Conclusions: Selected digital handwriting variables, particularly temporal measures, were more consistently associated with altered pentagon-copy performance than physical performance outcomes. These findings suggest that digital handwriting analysis may represent a sensitive complementary approach for exploring subtle functional differences in community-based settings.

(JMIR Aging 2026;9:e85074) doi:10.2196/85074

XSL•FO
RenderX

KEYWORDS

fine motor control; motor biomarkers; older adults; pentagon-copy test; visuoconstructional performance

Introduction

The global population is experiencing a marked demographic shift toward aging. According to the World Health Organization [1], the proportion of individuals aged 60 years or older is projected to nearly double, rising from 12% in 2015 to 22% by 2050. Furthermore, it is anticipated that by midcentury, approximately 80% of older adults will reside in low- and middle-income countries.

Aging is intrinsically associated with physiological changes that increase susceptibility to age-related conditions, including cardiovascular disease, musculoskeletal disorders, and neurodegeneration [2], thereby compromising individuals' functional capacity. To monitor changes in functionality, different tests have been developed to evaluate functional capacity in older adults, with particular emphasis on muscle strength, balance, and overall physical performance. Commonly used tests include handgrip strength, Timed Up and Go (TUG), 6-Minute Walk Test (6MWT), One-Leg Stand Test (OLST), Short Physical Performance Battery (SPPB), countermovement jump (CMJ), and midthigh pull test, all of which have demonstrated reliability and safety in older populations, including those with cognitive impairments [3-6]. Importantly, regular physical activity and structured exercise play a key role in maintaining functional capacity, delaying functional decline, and supporting healthy aging [7].

In parallel with these physical alterations, aging is also characterized by a progressive decline in cognitive function. Cognitive decline encompasses a broad continuum, ranging from normal age-associated cognitive deterioration to pathological dementia [8]. Numerous factors contribute to cognitive decline, with advancing age recognized as the primary independent risk factor, followed by genetic, socioeconomic, and environmental determinants, including nutrition and physical activity [9]. Cognitive functions comprise a range of mental processes integral to acquiring knowledge, processing information, and reasoning, which include perception, memory, learning, attention, decision-making, and language [10]. These domains are essential for the effective execution of activities of daily living.

It is well documented that functional decline is associated with cognitive deterioration in older adults [11,12], and there is evidence that targeted physical exercise interventions may alleviate or reverse these effects [13,14]. Consequently, early detection of functional and cognitive decline is vital for implementing appropriate medical and exercise interventions to slow disease progression. Traditional methods for assessing cognitive decline include verbal and written tests, such as the Mini-Mental State Examination [15], Clock Drawing Test [16,17], Trail Making Test (TMT) [18], and pentagon copying test [17]. Although failure on the intersecting-pentagons task strongly suggests underlying neurocognitive dysfunction, especially in visuospatial and constructional domains, the published literature indicates that the task functions best as a screening marker rather than an isolated diagnostic standard. Studies have shown its usefulness in older adult cognitive screening, in distinguishing Alzheimer disease from healthy aging, and in characterizing visuospatial impairment in dementia with Lewy bodies, while mechanistic work also shows that pentagon-copy errors reflect multifaceted constructional apraxia linked to parietal systems [19,20]. Accordingly, altered pentagon-copy performance can reasonably be interpreted as evidence of probable cognitive vulnerability, but not as sufficient on its own to establish clinical cognitive impairment. Additionally, physical assessments such as gait speed and balance tests have been shown to correlate with cognitive status [21,22].

Handwriting, a complex motor-cognitive task, is notably affected in individuals with neurodegenerative conditions, particularly in text production [23,24], and its kinematic characteristics may provide valuable insights into cognitive status [25,26]. For example, patients with Alzheimer disease and mild cognitive impairment (MCI) often exhibit slower handwriting speeds, increased in-air to on-surface time ratios, prolonged reaction times, decreased fluency, impaired coordination, and reduced consistency in their writing tasks [24,27]. Recent findings have also demonstrated that combining handwriting analysis with gait parameters enhances the classification accuracy of individuals with cognitive impairments [23].

While existing literature has predominantly focused on identifying behavioral markers and cognitive tests associated with cognitive decline, there remains a significant gap in research regarding the detection of cognitive deterioration through simple, noninvasive, and time-efficient motor assessments, such as handwriting analysis and physical performance tests.

Therefore, the aim of this study was to examine whether handwriting kinematic and physical performance variables differ by pentagon-copy performance classification in community-dwelling older adults. We hypothesized that altered pentagon-copy performance would be associated with poorer physical performance and less efficient handwriting kinematics.

Methods

Study Design and Population

We conducted a cross-sectional study of community-dwelling participants aged 60 years or older. Participants with a negative medical recommendation to exercise were excluded from the study. A total of 174 participants (age: 73.88, SD 6.05 years; weight: 70.93, SD 13.07 kg; height: 159.07, SD 9.47 cm) divided into men (n=66; age: 73.63, SD 5.96 years; weight: 71.47, SD 13.16 kg; height: 159.56, SD 9.52 cm) and women (n=108; age: 73.90, SD 6.06 years; weight: 70.85, SD 13.10 kg; height: 158.95, SD 9.44 cm), participated in this study.

XSL•FO
RenderX

Anthropometry and Body Composition

Anthropometric data, including height assessed with a wall-mounted stadiometer (Seca 220) and weight assessed with a scale (Tanita, MC 780-P MA), were collected. Fat-free mass and fat mass were assessed using a bioimpedance analysis system, the BIA 101 Biva Pro (Akern). This provided detailed body composition measurements for the study. Participants were assessed barefoot and in minimal clothing. They were also instructed to refrain from consuming alcohol for 48 hours before the assessments, avoid intense physical activity for at least 12 hours before the assessments, and not wear any metallic objects during the evaluation.

Handwriting Signal Acquisition

Overview

Handwriting signals were recorded using a digitizing tablet and stylus (Wacom Cintiq Pro 16, equipped with the Wacom Pro Pen 2), using an online data acquisition method. This approach represents a significant advancement over traditional offline acquisition, wherein handwriting is performed manually on paper and subsequently digitized through scanning. A principal advantage of online acquisition lies in its capacity to capture pen movements not only during direct contact with the writing surface but also during "in-air" phases, when the pen is lifted and moving above the surface [28]. The tablet operated at a sampling frequency of 100 Hz, recording spatial coordinates (x, y), pen pressure, pen tilt (altitude and azimuth angles), and pen status (pen-up/pen-down). These high-resolution temporal and spatial data allow for a comprehensive characterization of motor behavior during handwriting and drawing tasks.

Participants completed 10 handwriting tasks, each falling into 1 of 3 distinct categories as proposed by Garnacho-Castaño et al [29]: cognitive effort tasks, fine motor control tasks, and mechanical tasks.

Cognitive Effort Tasks

This category involves tasks that require substantial mental effort to replicate a complex figure or pattern. Successful completion necessitates working memory, spatial planning, and attentional control to remember the sequence of elements that have already been drawn and to anticipate those yet to follow. Such tasks are often particularly challenging for individuals with cognitive impairments [17], including MCI or early-stage dementia. Included in this category are the pentagon copying test (task 1) and the House Drawing Task (task 2).

Fine Motor Control Tasks

Fine motor control tasks, while not demanding in terms of cognitive processing, require precise motor execution. The visual patterns are generally simple and easily memorized after brief exposure; however, successful performance hinges on neuromotor control and sensorimotor integration. Tasks in this category include the Archimedean spiral (task 3), straight-line trace (task 4), spring trace (task 5), and concentric circles (task 6). These tasks are commonly used in neurological assessments and have demonstrated diagnostic value in conditions such as Parkinson disease, where subtle irregularities in amplitude, velocity, and timing of movement serve as early markers of dysfunction [30].

In certain tasks (eg, concentric circles), participants are instructed to trace a preprinted shape with minimal deviation, thereby assessing spatial fidelity and tremor control. Other tasks, such as the continuous line or spring patterns, emphasize rhythmicity, temporal consistency, and trajectory maintenance. The ability to sustain smooth and uniform movement represents an important marker of neuromotor function [31].

Mechanical Tasks

Mechanical tasks are characterized by low cognitive demands and involve automatic, overlearned motor patterns commonly used in daily routines. Examples include producing an individual's signature (tasks 7 and 10), writing capital letters (task 8), and copying a phrase (task 9). These tasks are generally straightforward for healthy individuals and are frequently used in biometric identification systems because of their high consistency and user specificity [32]. Given their automatic nature, deviations in task execution may be more easily attributed to motor abnormalities rather than cognitive deficits [33].

The sheet templates used for each of the 10 tasks are shown in Figures 1 and 2.

XSL•FO
RenderX

Figure 1. Drawing task sheet template, each identified by a number from 1 to 6.

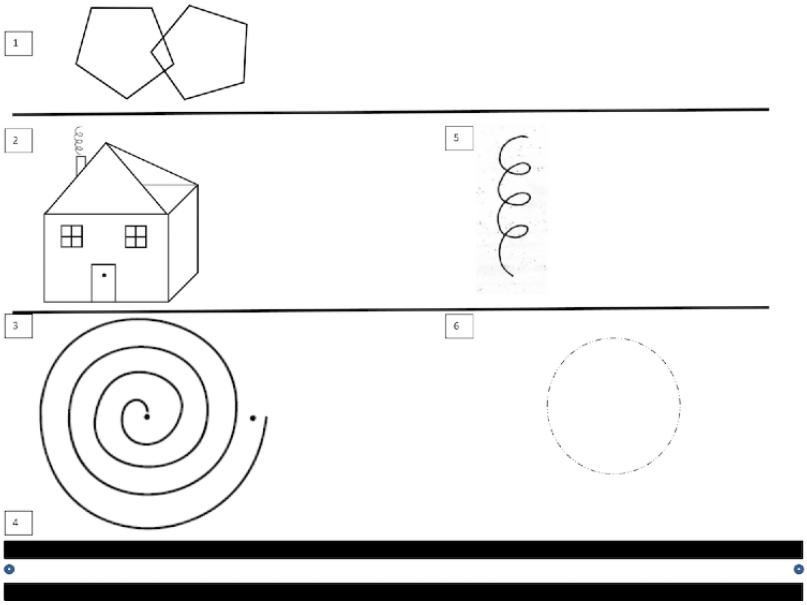
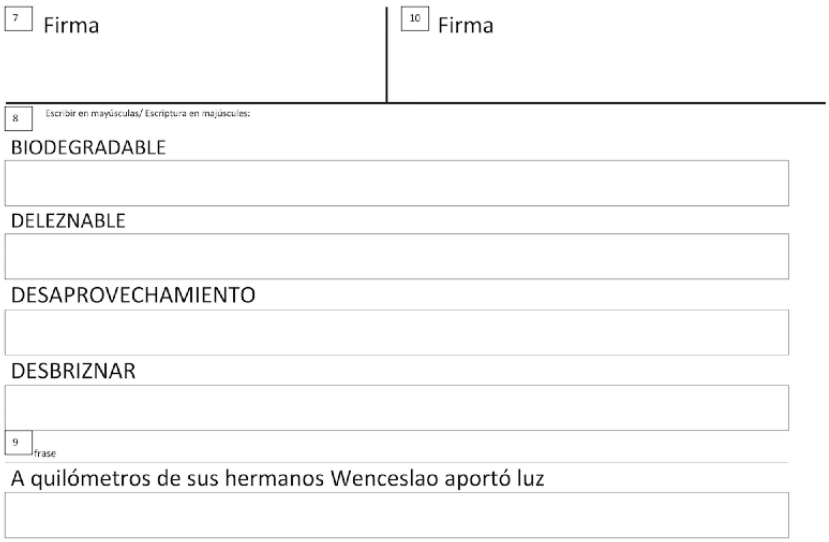


Figure 2. Writing task sheet template, each identified by a number from 7 to 10.



Handwriting Feature Extraction

The main variables obtained from each handwriting task were as follows:

Time on Air

Measured in milliseconds (ms) and refers to the total amount of time during which the pen is not in contact with the tablet surface, typically when the hand moves between letters or words. This measure reflects motor planning and writing fluency [34]. Higher values may suggest reduced motor efficiency or greater cognitive demand during the planning of subsequent strokes.

Time Down

Measured in milliseconds and refers to the total duration during which the pen remains in contact with the tablet surface while writing. This variable is closely associated with writing speed and the sustained pressure exerted during the task. Higher values may indicate a slower, more controlled writing process, often requiring greater cognitive or motor effort, whereas lower values may reflect a more automated and fluent execution.

Mean and Maximum Pressure

Measured in arbitrary units (au) and refer to the average and peak levels of pressure applied by the pen while in contact with the tablet surface. Mean pressure indicates the general force exerted throughout the writing task, whereas maximum pressure reflects the highest force recorded during a single stroke. These parameters are indicative of fine motor control and the physical effort involved in writing [35]. Deviations from typical values may be associated with motor rigidity, tremor, or impaired ability to modulate pressure during handwriting.

Mean and Maximum Speed

Measured in mm/s and refer to the average and highest velocities reached by the pen during the writing process, typically

XSL•FO
RenderX

measured as the distance covered per unit of time. Mean speed provides an overall indication of the writing pace, whereas maximum speed captures brief moments of rapid movement. These measures reflect the general rhythm and fluidity of handwriting. Alterations in speed may indicate motor slowing or, conversely, impulsive motor patterns characterized by sudden bursts of high velocity.

Mean and Maximum Acceleration

Measured in mm/s^2 and refer to the average and highest rates of change in the pen's velocity during the writing task. Mean acceleration reflects the general smoothness and consistency of pen movement, whereas maximum acceleration captures sudden, sharp changes in speed. These variables are closely linked to movement coordination and motor control. Elevated or irregular acceleration values may be indicative of tremors, spasmodic movements, or poor regulation of handwriting dynamics.

Group Classification Based on Pentagon-Copy Performance

To assess classification, the pentagon copying test from the Montreal Cognitive Assessment [36] was administered. In this task, participants were instructed to replicate a specific geometric figure, a pentagon. The test evaluates visuospatial abilities, executive functioning, and attention. Participants were provided with a digital sheet featuring a predrawn pentagon and instructed to replicate the figure by connecting its vertices using straight lines. The task aimed to ensure the accurate completion of the shape in accordance with the specified guidelines.

The test was scored based on the accuracy of the drawing and adherence to the instructions. Errors such as incorrect line placement, additional lines, or failure to follow instructions may indicate impairments in visuospatial processing and executive functioning, which are cognitive domains commonly affected in neurological conditions. Most experts assign a binary score of 0 or 1, with a correct response defined as producing 2 pentagons, each with 5 sides, and a correctly intersected region between them. If these criteria are not met, the score is 0. This scoring process is both straightforward and efficient.

In some contexts, a more granular scoring approach is used, particularly in computerized versions of the test. For example, Nagaratnam et al [19] proposed a scoring system ranging from 1 to 10 based on the quality of the drawing. In line with the original criteria, factors such as figure rotation and tremor were disregarded. In this study, a binary scoring system was adopted because the aim was to classify participants according to pentagon-copy performance rather than to perform a detailed visuoconstructive assessment. Importantly, pentagon-copy performance was used as a task-specific grouping variable and not as a diagnostic criterion or cognitive screening reference standard for cognitive impairment.

The same database used in this study was previously used to investigate cognitive impairment detection through a frontal camera during handwriting tasks, where we reported up to 83% accuracy based on head movement analysis, highlighting the potential of this approach as an automatic assessment method [37].

Strength Assessment

To evaluate muscular strength and power, a series of standardized tests targeting both upper- and lower-body performance were conducted. These tests included measures of isometric, dynamic, and explosive strength, as described below.

The handgrip strength test was used as a measure of muscular strength [38], particularly for assessing overall upper-body strength. Participants were instructed to grip a Jamar hand dynamometer (Patterson Medical) with maximal force for several seconds. The test was performed on both hands, and the highest value, measured in kilograms, was recorded as the result.

The 30-second arm curl test was used to evaluate upper-body strength and endurance [39]. Participants performed as many bicep curls as possible within 30 seconds using a 2.3 kg dumbbell for women and a 3.6 kg dumbbell for men, with the total number of repetitions recorded as the score.

The midthigh pull test was used to assess lower-body strength [40]. For this test, participants gripped a fixed bar connected to a strength sensor (Chronojump) positioned at midthigh height while maintaining a partial squat posture. They were instructed to pull upward with maximal effort for 3 to 5 seconds, with the peak force generated, measured in Newtons, recorded as a measure of lower-body strength.

The CMJ was used to dynamically assess lower-body explosive strength [41]. Participants started from an upright standing position, performed a rapid downward movement, and immediately jumped vertically as high as possible. Jump height (cm), flight time (ms), and power output (W/kg) were recorded using a force platform (MuscleLab). Each participant performed 3 trials with a 1-minute rest interval, and the highest value was used for analysis.

Cardiorespiratory Assessment

Cardiorespiratory fitness was evaluated using the 6MWT [42]. The test was conducted on a hardened surface along a 30-m straight-line circuit. Participants were instructed to walk at their fastest pace, without transitioning to running, to cover the greatest possible distance within 6 minutes. They wore comfortable clothing and footwear and were allowed to rest or stop if necessary. To objectively measure ventilatory parameters, participants were equipped with a portable gas analyzer (K5 COSMED). Peak oxygen (peak VO_2 , mL/kg/min) uptake was recorded as an indicator of cardiorespiratory fitness, alongside peak carbon dioxide (peak VCO_2 , mL/kg/min) production and pulmonary ventilation (VE; L/min) to assess ventilatory efficiency.

Functional Test and Balance Stability Assessment

To evaluate participants' functional mobility, balance, and strength, a series of standardized tests were used, including the TUG test, 400-m walking test, and SPPB. These assessments are designed to measure key aspects of physical performance, such as gait speed over short and medium distances, balance, and the ability to perform dynamic movements. The results from these tests help classify individuals based on their functional abilities and identify potential areas of decline in mobility and stability.

XSL•FO
RenderX

The TUG test involves measuring the time required for an individual to rise from a chair, walk 3 m as quickly as possible, turn around, and return to sit down. A completion time of less than 10 seconds is generally considered to be within normal limits. The onset of functional decline is indicated by times ranging from 10 to 20 seconds, although the individual may still be capable of performing basic transfers independently [43].

The 400-m walking test was incorporated within the 6MWT protocol, during which the total time taken by participants to complete a distance of 400 m was recorded. This measurement provides a practical indicator of functional walking capacity and endurance [44].

The SPPB comprises 3 components: a balance assessment, a 4-m gait speed test, and a sit-to-stand test involving 5 repetitions.

The balance assessment evaluates the participant's ability to maintain specific standing positions for 10 seconds each: feet side by side, semitandem, and tandem. For the feet side by side and semitandem positions, participants received scores of 0 or 1 based on their ability to hold the position (0 points for less than 10 seconds and 1 point for 10 seconds or more). For the tandem position, scoring ranges from 0 to 2 points (0 points for less than 3 seconds, 1 point for 3 to less than 10 seconds, and 2 points for 10 seconds or more) [45]. The total balance assessment score ranged from 0 to 4 points.

Gait speed was measured over an 8-m course, which included 2 m for acceleration, 4 m for timing, and 2 m for deceleration, thereby minimizing early braking bias. Participants were instructed to walk at their usual pace, as if walking on a street. Timing started on the command "Go," following a 3-second countdown, and concluded as the participant crossed the 4-m mark. Scoring was determined as follows: 0 points if the participant was unable to perform the test, 1 point for times exceeding 6.52 seconds, 2 points for times between 4.66 and 6.52 seconds, 3 points for times between 3.62 and 4.65 seconds, and 4 points for times below 3.62 seconds [45].

The sit-to-stand test involved a straight-backed chair stabilized against a wall. Participants were instructed to cross their arms over their chest and stand up and sit down 5 times as quickly as possible. Timing began from the initial seated position and ended when the participant reached the final standing position after the fifth repetition. The scoring for this component was based on the total time taken to complete the task: 0 points for more than 60 seconds or inability to perform the test, 1 point for between 16.7 and 59 seconds, 2 points for between 13.70 and 16.69 seconds, 3 points for between 11.20 and 13.69 seconds, and 4 points for less than 11.19 seconds [45].

A composite score ranging from 0 to 12 points was derived from all components, categorizing individuals into the following groups: severe limitations (0-3 points), moderate limitations (4-6 points), mild limitations (7-9 points), and minimal limitations (10-12 points) [46].

Static balance was assessed using a force platform (MuscleLab) [47]. Participants stood barefoot on both feet for up to 30 seconds across 3 trials, each separated by a 1-minute rest interval, without external support. The test was performed bilaterally to evaluate balance ability. Center of pressure parameters were recorded, including mean path velocity (mm/s), anteroposterior and mediolateral path velocities (mm/s), anteroposterior and mediolateral oscillation frequencies (Hz), and mean center of pressure displacement (mm).

Statistical Analysis

Continuous variables were expressed as mean (SD) and categorical variables as frequencies (%). Initial between-group comparisons were performed using 2-tailed Student *t* tests or Mann–Whitney U test, as appropriate, and chi-square tests for categorical variables.

Adjusted analyses were conducted using linear regression models, including group as the main predictor and age and sex as covariates. Regression coefficients (β) with 95% CIs were reported.

Variables with nonnormal distributions were log transformed before analysis.

Given the number of comparisons, the false discovery rate (FDR) was controlled using the Benjamini-Hochberg procedure, and both raw and adjusted *P* values are presented. Statistical significance was set at $P<.05$.

Ethical Considerations

All participants were informed about the aims and procedures of the study and provided written informed consent before participation. The study protocol was approved by the Ethics Committee of Universitat TecnoCampus Mataró-Maresme (CEI1/2022) and was conducted in accordance with the Declaration of Helsinki. To ensure privacy and confidentiality, all participant data were anonymized before analysis and stored securely in password-protected files accessible only to authorized members of the research team. No financial or other compensation was provided to participants for taking part in the study.

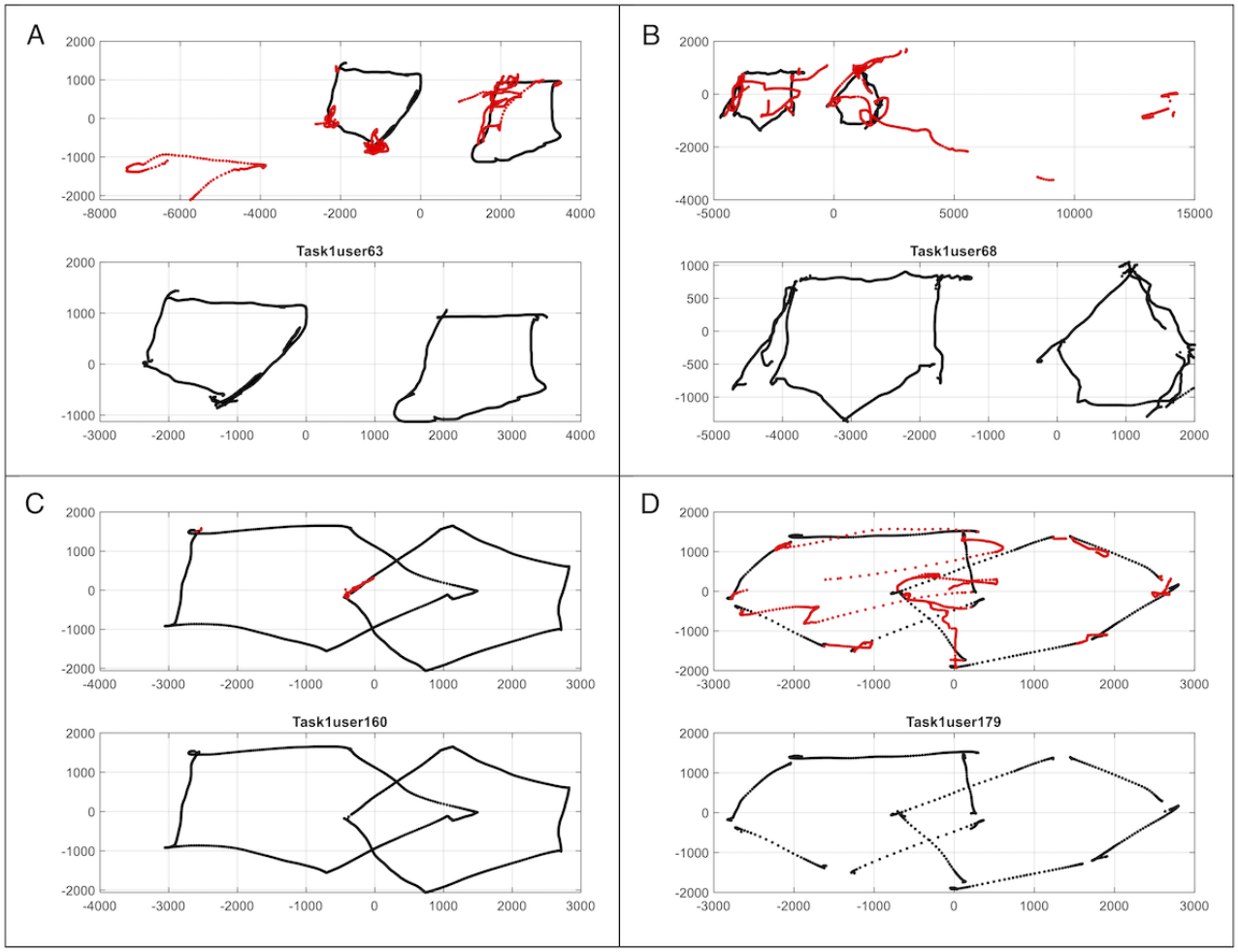
Results

Group Classification

Group classification was based on the results of the pentagon copying test. All drawings were manually reviewed, and participants were classified as either having normal pentagon-copy performance or altered pentagon-copy performance, depending on whether the drawing met the criteria: the presence of 2 pentagons, each with 5 sides, and a correctly formed intersection region between them. Participants who fulfilled these criteria received 1 point; those who did not were given a score of 0. Figure 3 presents the results of this task, illustrating the classification of 4 participants into 2 distinct groups: normal pentagon-copy performance and altered pentagon-copy performance.

XSL•FO
RenderX

Figure 3. Differences in pentagon-copying test performance between participants with altered pentagon-copy performance (A and B) and those with normal pentagon-copy performance (C and D).



Of the total sample, 93 participants were classified as having normal pentagon-copy performance (mean age 69.6, SD 5.9 years) and 81 as having altered pentagon-copy performance (mean age 72.9, SD 5.6 years). Among them, 56.1% (37/66) men and 51.9% (56/108) women were classified as having normal pentagon-copy performance, whereas 43.9% (29/66) men and 48.1% (52/108) women were classified as having altered pentagon-copy performance. Differences in pentagon-copy performance across age groups are presented in Table 1.

Table 1. Pentagon-copy performance by age group.

Age group (years)	Normal pentagon-copy performance, n (%)	Altered pentagon-copy performance, n (%)
60-69	57 (69.5)	25 (30.5)
70-79	31 (38.8)	49 (61.3)
≥80	5 (41.7)	7 (58.3)

Handwriting Performance

After adjustment for age and sex and correction for multiple comparisons using the FDR, several handwriting variables remained significantly associated with group classification based on pentagon-copy performance. The number of statistically significant differences, as well as mean values, regression coefficients, confidence intervals, and *P* values, are presented in Tables 2-5.

XSL•FO
RenderX

Table 2. Number of significant differences, after Benjamini–Hochberg correction applied, observed across task types in writing performance variables.

Task type (task numbers)	Total tasks, n	Time on air, n	Time down	Mean pressure, n	Max pressure, n	Mean speed, n	Max speed, n	Mean acceleration, n	Max acceleration, n
Cognitive (T1 and T2)	2	0	2	1	0	0	0	0	0
Fine motor control (T3, T4, T5, and T6)	4	0	0	0	0	0	0	1	0
Mechanical (T7, T8, T9, and T10)	4	2	2	1	0	0	1	0	1

Table 3. Statistical values for the variables obtained from the cognitive tasks.

Task and feature	Normal pentagon-copy performance, mean (SD)	Altered pentagon-copy performance, mean (SD)	β (regression coefficient)	95% CI	*P* value (row)	*P* value (FDR[a])
T1: pentagon copying test						
Time on air (ms)	1992.33 (722.49)	2305.01 (919.31)	153.146	–79.80 to 386.09	.20	.43
Time down (ms)	1044.43 (583.77)	1689.35 (1080.29)	526.799	279.44 to 774.16	<.001[b]	.001[b]
Mean pressure (au)	10919.99 (4511.14)	8719.79 (4253.78)	–2058.776	–3406.26 to –711.30	.003[b]	.03[b]
Max pressure (au)	21059.61 (3698.65)	19501.52 (4135.71)	–1326.480	–2524.12 to –128.84	.03[b]	.13
Mean speed (mm/s)	14.31 (5.82)	13.60 (4.82)	0.173	–1.35 to 1.70	.82	.94
Max speed (mm/s)	2254.08 (1877.30)	2585.35 (1890.03)	279.486	–272.80 to 831.78	.32	.55
Mean acceleration (mm/s^2)	5.51 (3.66)	5.92 (2.71)	0.591	–0.33 to 1.52	.21	.44
Max acceleration (mm/s^2)	2255.40 (1877.07)	2587.61 (1888.93)	281.083	–271.38 to 833.55	.32	.55
T2: house drawing						
Time on air (ms)	3806.12 (1574.59)	4192.51 (2126.70)	65.524	–423.79 to 554.84	.79	.93
Time down (ms)	3282.78 (1219.80)	4402.74 (2214.57)	825.700	316.08 to 1335.32	.002[b]	.02[b]
Mean pressure (au)	7294.31 (2404.32)	6186.05 (2223.39)	–1015.428	–1735.19 to –295.67	.006[b]	.054
Max pressure (au)	19921.83 (3515.07)	19152.14 (3546.15)	–526.070	–1637.40 to 585.26	.35	.58
Mean speed (mm/s)	18.99 (6.04)	17.55 (6.75)	–0.052	–1.82 to 1.72	.95	.97
Max speed (mm/s)	7105.57 (2395.51)	6524.47 (2450.45)	–556.572	–1269.58 to 156.43	.13	.32
Mean acceleration (mm/s^2)	9.47 (4.03)	9.00 (4.20)	0.282	–0.88 to 1.45	.63	.80
Max acceleration (mm/s^2)	7103.80 (2397.24)	6524.08 (2449.79)	–555.128	–1268.10 to 157.84	.13	.32

[a]FDR: false discovery rate.

[b]Statistically significant differences.

XSL•FO
RenderX

Table 4. Statistical values for the variables obtained from the fine motor control tasks.

Task and feature	Normal pentagon-copy performance, mean (SD)	Altered pentagon-copy performance, mean (SD)	β (regression coefficient)	95% CI	*P* value (row)	*P* value (FDR[a])
T3: Archimedean spiral						
Time on air (ms)	2570.31 (736.70)	2969.46 (1326.04)	294.475	–39.87 to 628.81	.08	.26
Time down (ms)	514.92 (458.77)	731.10 (552.95)	164.563	13.99 to 315.14	.03[b]	.14
Mean pressure (au)	15562.64 (4110.56)	13690.84 (3740.96)	–1418.389	–2637.34 to –199.44	.02[b]	.11
Max pressure (au)	21321.97 (4218.02)	20333.68 (3716.45)	–448.036	–1648.65 to 752.58	.47	.70
Mean speed (mm/s)	19.04 (5.90)	18.47 (6.37)	–0.062	–1.99 to 1.86	.95	.97
Max speed (mm/s)	8671.91 (2812.86)	9346.66 (2993.25)	707.225	–205.70 to 1620.15	.13	.32
Mean acceleration (mm/s^2)	6.18 (3.84)	6.27 (3.68)	0.099	–1.19 to 1.39	.89	.97
Max acceleration (mm/s^2)	8671.12 (2813.08)	9345.58 (2994.18)	706.648	–206.63 to 1619.92	.13	.32
T4: straight-line trace						
Time on air (ms)	1350.31 (468.86)	1336.15 (550.89)	–84.357	–236.68 to 67.96	.28	.49
Time down (ms)	270.80 (177.89)	356.42 (267.67)	49.522	–20.21 to 119.25	.16	.38
Mean pressure (au)	13804.48 (3479.02)	12186.11 (3655.28)	–1198.856	–2216.65 to –181.07	.02[b]	.11
Max pressure (au)	19500.58 (3921.30)	18338.06 (4592.50)	–685.930	–1985.92 to 614.06	.30	.52
Mean speed (mm/s)	26.47 (8.91)	28.37 (10.66)	3.541	0.50 to 6.58	.02[b]	.11
Max speed (mm/s)	8438.09 (3166.12)	9930.53 (4763.90)	1519.761	318.00 to 2721.52	.01[b]	.10
Mean acceleration (mm/s^2)	9.26 (6.86)	11.58 (9.69)	3.034	0.24 to 5.83	.03[b]	.14
Max acceleration (mm/s^2)	8434.92 (3166.60)	9935.58 (4757.37)	1525.121	323.41 to 2726.83	.01[b]	.10
T5: spring trace						
Time on air (ms)	767.43 (320.16)	910.09 (645.53)	73.060	–77.96 to 224.08	.34	.57
Time down (ms)	249.49 (160.02)	313.86 (246.21)	38.755	–28.03 to 105.54	.26	.48
Mean pressure (au)	12740.87 (3223.17)	11794.55 (3675.77)	–638.923	–1678.96 to 401.11	.23	.46
Max pressure (au)	20484.05 (4003.63)	19176.40 (4471.94)	–836.476	–2099.70 to 426.75	.19	.43
Mean speed (mm/s)	33.65 (12.74)	32.56 (16.19)	0.781	–3.99 to 5.55	.75	.91
Max speed (mm/s)	13403.85 (1955.36)	13569.41 (3590.10)	43.119	–837.30 to 923.54	.92	.97
Mean acceleration (mm/s^2)	18.53 (7.32)	18.32 (8.88)	0.522	–2.16 to 3.20	.70	.86
Max acceleration (mm/s^2)	13400.57 (1958.30)	13565.90 (3594.64)	42.238	–839.75 to 924.22	.92	.97
T6: concentric circles						
Time on air (ms)	5559.02 (3275.44)	5166.31 (3737.26)	–976.481	–1995.50 to 42.54	.06	.22
Time down (ms)	391.38 (329.59)	605.52 (771.61)	145.142	–15.45 to 305.74	.08	.25
Mean pressure (au)	18416.99 (3751.72)	16821.88 (4363.81)	–1263.018	–2526.12 to 0.09	.05	.19
Max pressure (au)	23032.83 (3731.70)	21890.59 (4058.39)	–686.131	–1891.81 to 519.55	.27	.49
Mean speed (mm/s)	31.59 (18.98)	36.71 (24.46)	8.652	1.96 to 15.34	.01[b]	.09
Max speed (mm/s)	3601.41 (1296.68)	4148.01 (1606.90)	414.529	–48.60 to 877.66	.08	.25
Mean acceleration (mm/s^2)	3.59 (2.09)	5.14 (4.26)	1.769	0.75 to 2.79	<.001[b]	.01[b]
Max acceleration (mm/s^2)	3600.78 (1298.13)	4147.52 (1607.63)	414.683	–48.66 to 878.02	.08	.25

[a]FDR: false discovery rate.

[b]Statistically significant differences.

XSL•FO
RenderX

Table 5. Statistical values for the variables obtained from the mechanical tasks.

Task and feature	Normal pentagon-copy performance, mean (SD)	Altered pentagon-copy performance, mean (SD)	β (regression coefficient)	95% CI	*P* value (row)	*P* value (FDR[a])
T7: signature						
Time on air (ms)	667.94 (364.30)	883.74 (551.08)	108.478	–22.64 to 239.59	.10	.29
Time down (ms)	191.84 (195.47)	337.99 (313.19)	88.176	12.91 to 163.44	.02[b]	.11
Mean pressure (au)	13161.94 (3818.82)	11722.73 (3735.97)	–741.676	–1898.19 to 414.84	.21	.44
Max pressure (au)	22306.63 (3949.26)	21314.91 (3355.68)	–353.550	–1460.81 to 753.71	.53	.73
Mean speed (mm/s)	60.19 (30.81)	49.65 (34.16)	–5.376	–15.12 to 4.37	.28	.49
Max speed (mm/s)	950.67 (2026.34)	857.06 (1464.04)	–162.247	–633.89 to 309.40	.50	.73
Mean acceleration (mm/s^2)	13.92 (10.27)	10.68 (6.47)	–2.059	–4.37 to 0.25	.08	.25
Max acceleration (mm/s^2)	859.66 (2056.96)	798.40 (1488.37)	–140.282	–620.54 to 339.98	.57	.77
T8: capital letters						
Time on air (ms)	5456.32 (1682.64)	7146.46 (2923.57)	1111.460	475.21 to 1747.71	<.001[b]	.01[b]
Time down (ms)	4839.24 (2191.81)	8749.05 (5224.63)	2921.053	1863.79 to 3978.31	<.001[b]	<.001[b]
Mean pressure (au)	7855.54 (1844.66)	6579.06 (2091.16)	–845.643	–1407.13 to –284.16	<.001[b]	.03[b]
Max pressure (au)	22000.49 (3758.36)	21329.35 (3435.94)	–115.567	–1179.81 to 948.67	.83	.94
Mean speed (mm/s)	20.98 (6.44)	17.69 (6.90)	–1.426	–3.28, to 0.43	.13	.32
Max speed (mm/s)	9060.52 (1802.76)	10252.66 (3050.16)	1166.011	450.70 to 1881.32	.001[b]	.02[b]
Mean acceleration (mm/s^2)	10.83 (3.60)	9.45 (3.89)	–.351	–1.41 to 0.71	.52	.73
Max acceleration (mm/s^2)	9059.39 (1803.40)	10252.81 (3049.91)	1167.717	452.31 to 1883.12	.001[b]	.02[b]
T9: phrase copying						
Time on air (ms)	3599.74 (939.12)	4840.90 (1885.47)	895.798	475.99 to 1315.61	<.001[b]	.001[b]
Time down (ms)	3005.53 (1463.13)	5611.05 (3955.06)	2086.101	1228.54 to 2943.66	<.001[b]	<.001[b]
Mean pressure (au)	8943.97 (2316.99)	7726.09 (2675.64)	–788.751	–1561.70 to –15.80	.05[b]	.18
Max pressure (au)	22723.55 (3763.68)	22099.78 (3845.61)	–47.616	–1198.17 to 1102.94	.94	.97
Mean speed (mm/s)	20.41 (5.86)	18.79 (12.41)	0.106	–3.01 to 3.22	.95	.97
Max speed (mm/s)	6871.78 (4078.34)	7552.58 (4185.74)	984.959	–309.76 to 2279.68	.14	.32
Mean acceleration (mm/s^2)	7.62 (3.08)	7.85 (7.17)	1.022	–0.82 to 2.86	.28	.49
Max acceleration (mm/s^2)	6869.59 (4078.49)	7550.90 (4184.92)	985.482	–309.06 to 2280.03	.14	.32
T10: signature						
Time on air (ms)	658.56 (356.86)	834.12 (521.77)	82.760	–43.61 to 209.13	.20	.43
Time down (ms)	280.35 (225.32)	430.60 (356.12)	77.437	1.10 to 153.77	.05[b]	.18
Mean pressure (au)	11924.60 (3236.74)	10739.14 (3427.03)	–612.123	–1618.27 to 394.03	.23	.45
Max pressure (au)	22437.45 (3398.14)	21973.26 (3891.51)	–72.172	–1171.64 to 1027.29	.90	.97
Mean speed (mm/s)	77.26 (43.62)	66.08 (43.97)	–3.623	–16.92 to 9.68	.59	.78
Max speed (mm/s)	14961.69 (2002.23)	15296.53 (1475.79)	386.022	–114.73 to 886.77	.13	.32
Mean acceleration (mm/s^2)	35.23 (22.95)	31.96 (24.13)	0.838	–6.36 to 8.04	.82	.94
Max acceleration (mm/s^2)	14958.35 (2013.05)	15292.96 (1476.56)	386.061	–116.01 to 888.13	.13	.32

[a]FDR: false discovery rate.

XSL•FO
RenderX

[b]Statistically significant differences.

The most consistent differences between groups were observed in temporal variables. In particular, time down was significantly higher in participants with altered pentagon-copy performance in tasks 1, 2, 8, and 9. In addition, time on air was significantly higher in this group in tasks 8 and 9.

Significant between-group differences were also found in pressure-related variables. Mean pressure was significantly lower in participants with altered pentagon-copy performance in tasks 1 and 8 after FDR correction.

For kinematic variables, significant differences after FDR adjustment were identified in selected tasks only. In task 6, participants with altered pentagon-copy performance showed significantly higher mean acceleration. In task 8, this group showed significantly higher maximum speed and maximum acceleration.

Overall, the greatest number of significant differences after FDR correction was observed in tasks 8 and 9. No other handwriting variables remained statistically significant after adjustment for multiple comparisons.

Physical Performance Results

In relation to physical performance tests, the unadjusted analyses showed statistically significant differences between participants with normal and altered pentagon-copy performance in selected variables, including CMJ height and flight time, anteroposterior axis balance stability, and 400-m walking time. However, after adjustment for age and sex, these differences were no longer statistically significant. Likewise, after controlling for multiple comparisons using the FDR, none of the physical performance variables remained statistically significant. No statistically significant differences were observed between the groups in other strength tests, such as handgrip strength, the 30-second arm curl test, or chair stand tests, nor in 4-m gait speed, total distance covered in the 6MWT, the TUG test, or metabolic variables such as peak VO_2 . The adjusted results for physical performance variables are presented in Table S1 in Multimedia Appendix 1. The number of participants differed across some tests because certain individuals were unable to complete them.

Discussion

Overview

The objective of this study was to examine whether handwriting kinematic and physical performance variables differ according to pentagon-copy performance classification in community-dwelling adults aged 60 years or older. After adjustment for age, sex, and multiple comparisons, only selected handwriting-derived variables remained significantly associated with altered pentagon-copy performance, whereas no physical performance variables remained significant.

The findings of this study should be interpreted within the context of a community-based sample without confirmed clinical diagnoses of cognitive impairment, in which group classification was based on pentagon-copy performance as a marker of task performance potentially related to cognitive vulnerability [20] rather than on a formal neuropsychological or clinical diagnostic assessment. Although failure on the intersecting-pentagons task is generally considered suggestive of underlying neurocognitive dysfunction, particularly in visuospatial and constructional domains, the literature supports its use primarily as a screening marker rather than as an isolated diagnostic standard [19,20]. Accordingly, altered pentagon-copy performance in this study should be interpreted as a marker of probable cognitive vulnerability within a community-based sample, but not as sufficient evidence on its own to establish clinical cognitive impairment. This study was not designed as a cognitive screening validation study; rather, it examined associations between handwriting and physical performance variables and group classification based on pentagon-copy performance in a community-based sample. Within this group-classification framework, differences were observed between individuals with normal and altered pentagon-copy performance in selected kinematic parameters recorded during digital handwriting tasks, whereas no corresponding differences were found in physical performance measures after adjusted analyses. These results are consistent with the hypothesis that early cognitive-related alterations may be reflected in fine motor execution [48], although not in overall physical performance in the present sample, in contrast to previous literature [49].

In the cognitively demanding handwriting tasks, which require mental planning, sustained attention, and working memory, the most consistent differences were observed between the 2 groups. Participants with altered pentagon-copy performance exhibited a significantly prolonged contact time with the screen and a markedly lower mean writing pressure. This pattern may reflect, among other possible explanations, greater disruption in movement planning during the writing process, a finding consistent with previous studies that have documented task execution slowing in complex sequential activities, as well as reductions in applied writing pressure among individuals with cognitive impairment [30,50]. Although velocity values in these tasks were not consistently lower in the altered pentagon-copy performance group, a nonsignificant tendency toward lower mean velocity was observed in some recordings, which may be indicative of less stable writing performance.

In fine motor control tasks, the altered pentagon-copy performance group exhibited higher values in mean acceleration. Contrary to what might be interpreted as an indicator of superior performance, this increase in acceleration likely reflects, among other possible explanations, reduced motor self-regulation and a more impulsive and inefficient execution. Previous evidence suggests that in contexts of cognitive impairment, individuals tend to adopt more erratic motor patterns, characterized by lower temporal and spatial consistency in stroke execution [30]. Higher acceleration in tasks demanding graphical precision may reflect less regulated motor execution.

The mechanically oriented tasks, designed to assess motor execution without additional cognitive demands, also revealed noteworthy differences. In tasks 8 and 9, the altered pentagon-copy performance group demonstrated a significantly longer total execution time (8749 ms vs 4839.2 ms and 5611

XSL•FO
RenderX

ms vs 3005.5, respectively) and higher time on air (7146.5 ms vs 5456.3 ms and 4840.9 ms vs 3599.7 ms, respectively). In addition, the altered pentagon-copy performance group showed lower mean writing pressure (6579.1 au vs 7855.5 au), higher maximum writing speed (10252.7 mm/s vs 9060.5 mm/s), and higher maximum acceleration (10252.8 mm/s^2 vs 9059.4 mm/s^2) in task 8. These results suggest reduced motor efficiency even in low-complexity tasks, consistent with previous evidence indicating that cognitive decline may compromise the basal function of motor networks involved in handwriting [27,51], even in the absence of overt clinical signs detectable through conventional neurological assessments.

The findings for physical performance should also be interpreted with caution. Although the unadjusted analyses initially suggested between-group differences in selected variables, including CMJ height and flight time, anteroposterior balance parameters, and 400-m walking time, these associations were no longer statistically significant after adjustment for age and sex. Furthermore, none of the physical performance variables remained significant after correction for multiple comparisons using the FDR, indicating that the initial differences were not robust after statistical adjustment. These findings suggest that, in this community-based sample, conventional physical performance measures may lack sensitivity to detect subtle functional differences associated with pentagon-copy performance classification. Previous studies have identified reduced explosive force production as a robust indicator of cognitively healthy aging [52], and associations have also been reported between MCI and other physical performance measures, including handgrip strength, TUG performance, gait speed, postural control, and cardiorespiratory fitness [53-59]. In this study, however, physical performance measures did not provide support as independent markers of group classification based on pentagon-copy performance. This discrepancy may reflect the relatively subtle nature of the classification approach used, the older age of participants with altered pentagon-copy performance, and other differences in study design, participant characteristics, cognitive reference standards, and severity of impairment across studies. It is also possible that such associations may become more evident in populations with more advanced cognitive impairment.

From a methodological perspective, this study provides a novel contribution by combining noninvasive digital tools for the precise quantification of handwriting kinematics with standardized assessments of physical performance in a community-based sample of older adults. This multimodal approach enabled the simultaneous evaluation of fine motor and functional measures in relation to group classification based on pentagon-copy performance. After adjustment for age, sex, and multiple comparisons, the most consistent associations were observed in selected handwriting-derived variables, particularly temporal measures, whereas physical performance outcomes did not remain statistically significant. In this context, digital handwriting analysis may represent a sensitive and scalable complementary approach for detecting subtle functional differences in community-based settings.

Several limitations should be acknowledged. First, the cross-sectional design does not allow conclusions regarding causality or predictive value. Second, group classification was based on pentagon-copy performance and does not constitute a clinical diagnosis of cognitive impairment. Third, participants with altered pentagon-copy performance were older on average, which may have contributed to the initial unadjusted differences despite statistical adjustment. Fourth, pentagon-copy performance may be influenced by factors other than cognitive status, such as visuoconstructional deficits, fine motor impairment, tremor, or visual difficulties. Finally, the study was conducted in a real-world community setting, where comprehensive neuropsychological evaluation was not feasible.

Future studies should examine these associations using more robust and clinically validated cognitive reference standards, such as the Montreal Cognitive Assessment total score, the Mini-Mental State Examination, or formal neuropsychological assessment. Longitudinal designs will also be needed to determine the predictive value of the handwriting variables identified here and to establish whether these measures may contribute to the development of digital screening tools or automated risk assessment models in primary care and community-based settings.

Conclusion

In this community-based sample of older adults, only a limited subset of digital handwriting variables remained significantly associated with group classification based on pentagon-copy performance after adjustment for age, sex, and multiple comparisons. The most consistent associations were observed in temporal handwriting measures, particularly time down and time on air, together with selected pressure and kinematic parameters, whereas none of the assessed physical performance variables remained statistically significant after adjusted analyses.

These findings suggest that digital handwriting analysis may be a more sensitive complementary approach than conventional physical performance tests for identifying subtle functional differences associated with altered pentagon-copy performance in older adults. However, because group classification was based on a single visuoconstructional task rather than on a formal clinical or neuropsychological diagnosis, the results should be interpreted with caution and should not be considered evidence of diagnostic screening performance for cognitive impairment.

Overall, this study supports the potential value of digital handwriting analysis as a noninvasive and scalable complementary approach for community-based assessment. Future research should use validated cognitive reference standards and longitudinal designs to determine the clinical relevance, predictive value, and potential application of these handwriting-derived measures in digital screening strategies for older adults.

XSL•FO
RenderX

Acknowledgments

The authors thank Francesco Di Luzio, from Sapienza Università di Roma, for his contribution to the segmentation of the handwriting tasks.

Funding

This research was partly supported by grant PID2023-146644OB-I00, funded by the Ministerio de Ciencia, Investigación y Universidades/Agencia Estatal de Investigación (MICIU/AEI; 10.13039/501100011033) and the European Union's Fondo Europeo de Desarrollo Regional (FEDER) program.

Data Availability

The datasets generated and analyzed during the current study are available from the authors upon reasonable request.

Authors' Contributions

All authors (PF-R, MF-Z, NS-P, MVG-C, and MG-F) contributed to the development of this study. PF-R contributed to conceptualization, investigation, data collection and extraction, and writing—original draft preparation; MF-Z contributed to conceptualization, data extraction, and writing—review and editing; NS-P contributed to conceptualization, investigation, data collection, and writing—review and editing; MVG-C contributed to conceptualization, investigation, and writing—review and editing; and MG-F contributed to investigation, methodology, data analysis, and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Physical performance variables according to pentagon-copy performance group. Values are presented as mean (SD), together with adjusted regression coefficients (β), 95% CIs, raw *P* values, and false discovery rate–adjusted *P* values.
[DOCX File , 20 KB - aging_v9i1e85074_app1.docx]

References

1. World population ageing 2019: highlights. United Nations. URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/aging_theme_world_population_ageing_2019_10_key_messages.pdf [accessed 2025-09-08]
2. Li Z, Zhang Z, Ren Y, Wang Y, Fang J, Yue H, et al. Aging and age-related diseases: from mechanisms to therapeutic strategies. Biogerontology 2021;22(2):165-187 [FREE Full text] [doi: 10.1007/s10522-021-09910-5] [Medline: 33502634]
3. Giannitsi S, Bougiakli M, Bechlioulis A, Kotsia A, Michalis LK, Naka KK. 6-minute walking test: a useful tool in the management of heart failure patients. Ther Adv Cardiovasc Dis 2019;13:1753944719870084 [FREE Full text] [doi: 10.1177/1753944719870084] [Medline: 31441375]
4. Steffen T, Seney M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. Phys Ther 2008;88(6):733-746. [doi: 10.2522/ptj.20070214] [Medline: 18356292]
5. Mehmet H, Robinson SR, Yang AWH. Assessment of gait speed in older adults. J Geriatr Phys Ther 2020;43(1):42-52. [doi: 10.1519/JPT.0000000000000224] [Medline: 30720555]
6. Espejo-Antúnez L, Pérez-Mármol JM, Cardero-Durán M, Toledo-Marhuenda JV, Albornoz-Cabello M. The effect of proprioceptive exercises on balance and physical function in institutionalized older adults: a randomized controlled trial. Arch Phys Med Rehabil 2020;101(10):1780-1788. [doi: 10.1016/j.apmr.2020.06.010] [Medline: 32663479]
7. Coelho-Júnior HJ, Uchida MC, Picca A, Bernabei R, Landi F, Calvani R, et al. Evidence-based recommendations for resistance and power training to prevent frailty in community-dwellers. Aging Clin Exp Res 2021;33(8):2069-2086. [doi: 10.1007/s40520-021-01802-5] [Medline: 33587271]
8. Chun CT, Seward K, Patterson A, Melton A, MacDonald-Wicks L. Evaluation of available cognitive tools used to measure mild cognitive decline: a scoping review. Nutrients 2021;13(11):3974 [FREE Full text] [doi: 10.3390/nu13113974] [Medline: 34836228]
9. Dominguez LJ, Veronese N, Vernuccio L, Catanese G, Inzerillo F, Salemi G, et al. Nutrition, physical activity, and other lifestyle factors in the prevention of cognitive decline and dementia. Nutrients 2021;13(11):4080 [FREE Full text] [doi: 10.3390/nu13114080] [Medline: 34836334]
10. Kiely KM. Cognitive function. In: Encyclopedia of Quality of Life and Well-Being Research. Dordrecht: Springer; 2014:974-978.

XSL•FO
RenderX

11. Biasin F, Ceolin C, Celli S, Terziotti C, Raffaelli C, Bontempi C. Interrelation between functional decline and dementia: the potential role of balance assessment. Hum Mov Sci 2023;89:103095 [FREE Full text] [doi: 10.1016/j.humov.2023.103095] [Medline: 37120906]
12. Auyeung TW, Lee JSW, Kwok T, Woo J. Physical frailty predicts future cognitive decline - a four-year prospective study in 2737 cognitively normal older adults. J Nutr Health Aging 2011;15(8):690-694 [FREE Full text] [doi: 10.1007/s12603-011-0110-9] [Medline: 21968866]
13. Rogge AK, Röder B, Zech A, Nagel V, Hollander K, Braumann KM, et al. Balance training improves memory and spatial cognition in healthy adults. Sci Rep 2017;7(1):5661 [FREE Full text] [doi: 10.1038/s41598-017-06071-9] [Medline: 28720898]
14. Li W, Luo Z, Jiang J, Li K, Wu C. The effects of exercise intervention on cognition and motor function in stroke survivors: a systematic review and meta-analysis. Neurol Sci 2023;44(6):1891-1903. [doi: 10.1007/s10072-023-06636-9] [Medline: 36781567]
15. Arevalo-Rodriguez I, Smailagic N, Roqué-Figuls M, Ciapponi A, Sanchez-Perez E, Giannakou A. Mini-mental state examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev 2021;7(7):CD010783 [FREE Full text] [doi: 10.1002/14651858.CD010783.pub3] [Medline: 34313331]
16. Cova I, Mele F, Zerini F, Maggiore L, Rosa S, Cucumo V, et al. The clock drawing test as a predictor of cognitive decline in non-demented stroke patients. J Neurol 2022;269(1):342-349 [FREE Full text] [doi: 10.1007/s00415-021-10637-z] [Medline: 34095964]
17. Vergouw L, Salomé M, Kerklaan A, Kies C, Roks G, van den Berg E, et al. The pentagon copying test and the clock drawing test as prognostic markers in dementia with lewy bodies. Dement Geriatr Cogn Disord 2018;45(5-6):308-317 [FREE Full text] [doi: 10.1159/000490045] [Medline: 29996144]
18. Bossers WJR, van der Woude LHV, Boersma F, Scherder EJA, van Heuvelen MJG. Recommended measures for the assessment of cognitive and physical performance in older patients with dementia: a systematic review. Dement Geriatr Cogn Dis Extra 2012;2(1):589-609 [FREE Full text] [doi: 10.1159/000345038] [Medline: 23341825]
19. Nagaratnam N, Nagaratnam K, O'Mara D. Intersecting pentagon copying and clock drawing test in mild and moderate Alzheimer's disease. Journal of Clinical Gerontology and Geriatrics 2014;5(2):47-52. [doi: 10.1016/j.jcgg.2013.11.001]
20. Martinelli JE, Cecato JF, Martinelli MO, de Melo B, Aprahamian I. Performance of the pentagon drawing test for the screening of older adults with Alzheimer's dementia. Dement Neuropsychol 2018;12(1):54-60 [FREE Full text] [doi: 10.1590/1980-57642018dn12-010008] [Medline: 29682234]
21. Gore S, Blackwood J, Ziccardi T. Associations between cognitive function, balance, and gait speed in community-dwelling older adults with COPD. J Geriatr Phys Ther 2023;46(1):46-52. [doi: 10.1519/JPT.0000000000000323] [Medline: 34334706]
22. Kuan YC, Huang LK, Wang YH, Hu CJ, Tseng IJ, Chen HC, et al. Balance and gait performance in older adults with early-stage cognitive impairment. Eur J Phys Rehabil Med 2021;57(4):560-567 [FREE Full text] [doi: 10.23736/S1973-9087.20.06550-8] [Medline: 33258361]
23. Qi H, Zhu X, Ren Y, Zhang X, Tang Q, Zhang C, et al. A study of assisted screening for alzheimer's disease based on handwriting and gait analysis. J Alzheimers Dis 2024;101(1):75-89. [doi: 10.3233/JAD-240362] [Medline: 39177597]
24. Fernandes CP, Montalvo G, Caligiuri M, Pertsinakis M, Guimarães J. Handwriting changes in Alzheimer's disease: a systematic review. J Alzheimers Dis 2023;96(1):1-11. [doi: 10.3233/JAD-230438] [Medline: 37718808]
25. Garre-Olmo J, Faúndez-Zanuy M, López-de-Ipiña K, Calvó-Perxas L, Turró-Garriga O. Kinematic and pressure features of handwriting and drawing: preliminary results between patients with mild cognitive impairment, alzheimer disease and healthy controls. Curr Alzheimer Res 2017;14(9):960-968 [FREE Full text] [doi: 10.2174/1567205014666170309120708] [Medline: 28290244]
26. Toffoli S, Abbate C, Lunardini F, Corno E, Diani N, Gallucci A, et al. Handwriting in mild cognitive impairment: reliability assessment and machine learning-based screening. JMIR Aging 2025;8:e73074 [FREE Full text] [doi: 10.2196/73074] [Medline: 40986851]
27. Yan JH, Rountree S, Massman P, Doody RS, Li H. Alzheimer's disease and mild cognitive impairment deteriorate fine movement control. J Psychiatr Res 2008;42(14):1203-1212. [doi: 10.1016/j.jpsychires.2008.01.006] [Medline: 18280503]
28. Sesa-Nogueras E, Faundez-Zanuy M, Mekyska J. An information analysis of in-air and on-surface trajectories in online handwriting. Cogn Comput 2011;4(2):195-205. [doi: 10.1007/s12559-011-9119-y]
29. Garnacho-Castaño M, Faundez-Zanuy M, Lopez-Xarbau J. On the handwriting tasks' analysis to detect fatigue. Applied Sciences 2020;10(21):7630. [doi: 10.3390/app10217630]
30. Drotár P, Mekyska J, Rektorová I, Masarová L, Smékal Z, Faundez-Zanuy M. Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. Artif Intell Med 2016;67:39-46. [doi: 10.1016/j.artmed.2016.01.004] [Medline: 26874552]
31. Schröter A, Mergl R, Bürger K, Hampel H, Möller HJ, Hegerl U. Kinematic analysis of handwriting movements in patients with Alzheimer's disease, mild cognitive impairment, depression and healthy subjects. Dement Geriatr Cogn Disord 2003;15(3):132-142 [FREE Full text] [doi: 10.1159/000068484] [Medline: 12584428]
32. Faundez-Zanuy M. Signature recognition state-of-the-art. IEEE Aerosp. Electron. Syst. Mag 2005;20(7):28-32. [doi: 10.1109/maes.2005.1499249]

XSL•FO
RenderX

33. Preti AN, Diana L, Castaldo R, Pischedda F, Difonzo T, Fumagalli G. Does cognitive decline influence signing? Aging Clin Exp Res 2023;35(11):2685-2691 [FREE Full text] [doi: 10.1007/s40520-023-02523-7] [Medline: 37661205]
34. Drotár P, Mekyska J, Rektorová I, Masarová L, Smékal Z, Faundez-Zanuy M. Analysis of in-air movement in handwriting: a novel marker for Parkinson's disease. Comput Methods Programs Biomed 2014;117(3):405-411. [doi: 10.1016/j.cmpb.2014.08.007] [Medline: 25261003]
35. Longstaff MG, Heath RA. Spiral drawing performance as an indicator of fine motor function in people with multiple sclerosis. Hum Mov Sci 2006;25(4-5):474-491. [doi: 10.1016/j.humov.2006.05.005] [Medline: 16859790]
36. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
37. Candela F, Romeo S, Faundez-Zanuy M, Ferrer-Ramos P. Cognitive impairment detection based on frontal camera scene while performing handwriting tasks. Cogn Comput 2024;16(3):1004-1021. [doi: 10.1007/s12559-024-10279-z]
38. Chai S, Zhao D, Gao T, Wang X, Wang X, Luo J, et al. The relationship between handgrip strength and cognitive function among older adults in China: functional limitation plays a mediating role. J Affect Disord 2024;347:144-149. [doi: 10.1016/j.jad.2023.11.056] [Medline: 37992778]
39. Rikli RE, Jones CJ. Development and validation of criterion-referenced clinically relevant fitness standards for maintaining physical independence in later years. Gerontologist 2013;53(2):255-267. [doi: 10.1093/geront/gns071] [Medline: 22613940]
40. Grgic J, Scapec B, Mikulic P, Pedisic Z. Test-retest reliability of isometric mid-thigh pull maximum strength assessment: a systematic review. Biol Sport 2022;39(2):407-414 [FREE Full text] [doi: 10.5114/biolsport.2022.106149] [Medline: 35309521]
41. Moreira-Reis A, Maté-Muñoz JL, Hernández-Lougedo J, Vilches-Sáez S, Benet M, García-Fernández P. Aerobic dance on an air dissipation platform improves cardiorespiratory, muscular and cellular fitness in the overweight and obese elderly. Biology (Basel) 2022;11(4):579 [FREE Full text] [doi: 10.3390/biology11040579] [Medline: 35453778]
42. Bautmans I, Lambert M, Mets T. The six-minute walk test in community dwelling elderly: influence of health status. BMC Geriatr 2004;4:6 [FREE Full text] [doi: 10.1186/1471-2318-4-6] [Medline: 15272934]
43. Podsiadlo D, Richardson S. The timed ‘Up & Go’: a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
44. Vestergaard S, Patel KV, Bandinelli S, Ferrucci L, Guralnik JM. Characteristics of 400-meter walk test performance and subsequent mortality in older adults. Rejuvenation Res 2009;12(3):177-184 [FREE Full text] [doi: 10.1089/rej.2009.0853] [Medline: 19594326]
45. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
46. Guralnik JM, Ferrucci L, Pieper CF, Leveille SG, Markides KS, Ostir GV. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci 2000;55(4):M221-M231. [doi: 10.1093/gerona/55.4.m221] [Medline: 10811152]
47. Pizzigalli L, Micheletti Cremasco M, Mulasso A, Rainoldi A. The contribution of postural balance analysis in older adult fallers: a narrative review. J Bodyw Mov Ther 2016;20(2):409-417 [FREE Full text] [doi: 10.1016/j.jbmt.2015.12.008] [Medline: 27210860]
48. Curreri C, Trevisan C, Carrer P, Facchini S, Giantin V, Maggi S. Difficulties with fine motor skills and cognitive impairment in an elderly population: the progetto veneto anziani. J Am Geriatr Soc 2018;66(2):350-356. [doi: 10.1111/jgs.15209] [Medline: 29322488]
49. Veronese N, Stubbs B, Trevisan C, Bolzetta F, De Rui M, Solmi M. What physical performance measures predict incident cognitive decline among intact older adults? A 4.4year follow up study. Exp Gerontol 2016;81:110-118. [doi: 10.1016/j.exger.2016.05.008] [Medline: 27235850]
50. Shi L, Ren Y, Wei Z, Qi H, Zhang R, Wang L. Evaluating handwriting characteristics as a novel assisted screening approach for mild cognitive impairment in community settings. Alzheimer's & Dementia 2025;20(S7). [doi: 10.1002/alz.092249]
51. Kawa J, Bednorz A, Stępień P, Derejczyk J, Bugdol M. Spatial and dynamical handwriting analysis in mild cognitive impairment. Comput Biol Med 2017;82:21-28. [doi: 10.1016/j.compbiomed.2017.01.004] [Medline: 28126631]
52. Steves CJ, Mehta MM, Jackson SH, Spector TD. Kicking back cognitive ageing: leg power predicts cognitive ageing after ten years in older female twins. Gerontology 2016;62(2):138-149 [FREE Full text] [doi: 10.1159/000441029] [Medline: 26551663]
53. Liu Y, Ma W, Li M, Han P, Cai M, Wang F. Relationship between physical performance and mild cognitive impairment in Chinese community-dwelling older adults. Clin Interv Aging 2021;16:119-127 [FREE Full text] [doi: 10.2147/CIA.S288164] [Medline: 33469279]
54. Shin BM, Han SJ, Jung JH, Kim JE, Fregni F. Effect of mild cognitive impairment on balance. J Neurol Sci 2011;305(1-2):121-125. [doi: 10.1016/j.jns.2011.02.031] [Medline: 21420690]

XSL•FO
RenderX

55. Zhou J, Manor B, McCarten JR, Wade MG, Jor'dan AJ. The effects of cognitive impairment on the multi-scale dynamics of standing postural control during visual-search in older men. Front Aging Neurosci 2023;15:1068316 [FREE Full text] [doi: 10.3389/fnagi.2023.1068316] [Medline: 36761178]
56. Deschamps T, Beauchet O, Annweiler C, Cornu C, Mignardot J. Postural control and cognitive decline in older adults: position versus velocity implicit motor strategy. Gait Posture 2014;39(1):628-630 [FREE Full text] [doi: 10.1016/j.gaitpost.2013.07.001] [Medline: 23891457]
57. Mignardot JB, Beauchet O, Annweiler C, Cornu C, Deschamps T. Postural sway, falls, and cognitive status: a cross-sectional study among older adults. J Alzheimers Dis 2014;41(2):431-439. [doi: 10.3233/JAD-132657] [Medline: 24625801]
58. Li H, Zhang J, Zou X, Jia X, Zheng D, Guo X, et al. The bidirectional association between cognitive function and gait speed in Chinese older adults: Longitudinal observational study. JMIR Public Health Surveill 2023;9:e44274 [FREE Full text] [doi: 10.2196/44274] [Medline: 36917163]
59. Pentikäinen H, Savonen K, Ngandu T, Solomon A, Komulainen P, Paajanen T. Cardiorespiratory fitness and cognition: longitudinal associations in the FINGER study. J Alzheimers Dis 2019;68(3):961-968. [doi: 10.3233/JAD-180897] [Medline: 30814346]

Abbreviations

6MWT: 6-Minute Walk Test
CMJ: countermovement jump
FDR: false discovery rate
MCI: mild cognitive impairment
OLST: One-Leg Stand Test
peak VCO_2 : peak carbon dioxide
peak VO_2 : peak oxygen
SPPB: Short Physical Performance Battery
TMT: Trail Making Test
TUG: Timed Up and Go
VE: pulmonary ventilation

Edited by M Antoniou; submitted 30.Sep.2025; peer-reviewed by H Lange, S Shimokihara; comments to author 14.Feb.2026; revised version received 31.Mar.2026; accepted 20.Apr.2026; published 18.May.2026.

Please cite as:
Ferrer-Ramos P, Faundez-Zanuy M, Serra-Payá N, Garnacho-Castaño MV, Girabent-Farrés M
Digital Handwriting Kinematics and Physical Performance According to Pentagon-Copy Performance in Community-Dwelling Older Adults: Cross-Sectional Study
JMIR Aging 2026;9:e85074
URL: https://aging.jmir.org/2026/1/e85074
doi:10.2196/85074
PMID:42149645

©Pau Ferrer-Ramos, Marcos Faundez-Zanuy, Noemí Serra-Payá, Manuel Vicente Garnacho-Castaño, Montserrat Girabent-Farrés. Originally published in JMIR Aging (https://aging.jmir.org), 18.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Machine Learning–Based Biomechanical and Clinical Assessment of the Five Times Chair Rise Test in Older Adults: Sensor-Based Functional Aging Study

Anna Schumacher[1], Dr rer nat; Ole Meyer[1,2], MSc; Vincent Quinten[3], MD, PhD; Andreas Hein[1], Prof Dr-Ing; Rebecca Diekmann[3], PD Dr oec troph

[1]Assistance Systems and Medical Device Technology, School VI - Medicine and Health Sciences, Carl von Ossietzky Universität Oldenburg, Ammerländer Heerstraße 140, Oldenburg, Lower Saxony, Germany

[2]OFFIS - Institute for Information Technology, Oldenburg, Lower Saxony, Germany

[3]Junior Research Group "Nutrition and Physical Function in Older Adults", School VI - Medicine and Health Sciences, Carl von Ossietzky Universität Oldenburg, Oldenburg, Lower Saxony, Germany

Corresponding Author:
Anna Schumacher, Dr rer nat
Assistance Systems and Medical Device Technology, School VI - Medicine and Health Sciences, Carl von Ossietzky Universität Oldenburg, Ammerländer Heerstraße 140, Oldenburg, Lower Saxony, Germany

Abstract

Background: Maintaining functional ability is essential for healthy aging. The Five Times Chair Rise Test (5CRT) is widely used to assess lower-extremity function, yet conventional stopwatch timing provides limited insight into movement quality and compensatory strategies. Sensor-based biomechanics combined with machine learning (ML) enables detailed analysis of movement dynamics and supports digital health approaches for early detection and monitoring of functional decline. However, the clinical relevance of ML-based 5CRT movement patterns and their relationship to established functional outcomes and aging trajectories remains insufficiently understood.

Objective: This study aimed (1) to classify 5CRT movement patterns using an established dynamic time warping (DTW)-based clustering workflow; (2) to examine associations between ML-based movement clusters and multidimensional indicators of strength, mobility, and functional performance across 24 months; and (3) to evaluate whether ML-based movement classification provides clinically meaningful differentiation beyond conventional stopwatch-based 5CRT timing within a digital health assessment context.

Methods: In this longitudinal observational study, community-dwelling adults aged ≥70 years completed standardized functional assessments at baseline, 6 months, and 24 months, including Handgrip Strength, Stair Climb Power Test, Timed "Up & Go" Test, Short Physical Performance Battery (including 5CRT), 4-meter Gait Speed, Balance Testing, 6-minute Walking Test, and Countermovement Jump. The assessments were recorded using a force plate and depth camera to capture biomechanical time-series data. DTW and k-means clustering were applied to ground reaction force signals to identify movement patterns.

Results: Among 172 participants included in the final analysis (mean age 76.9, SD 3.40 years; 96/172, 56% female), 2 distinct movement clusters were identified. Cluster 1 demonstrated faster chair rise performance and lower stabilization forces, whereas cluster 2 exhibited slower execution and increased stabilization demands. Across baseline, 6-month, and 24-month follow-up assessments, cluster 1 consistently showed higher muscle strength, greater leg power, faster mobility, higher gait speed, and longer walking distance than cluster 2 (all $P<.01$). At 24 months, large between-cluster differences were observed for the Timed "Up & Go" Test, 5CRT duration, gait speed, 6-minute walk distance, and Short Physical Performance Battery score. Conventional stopwatch scoring categorized most participants as high functioning, whereas ML-based clustering revealed substantial functional heterogeneity and clinically meaningful information not reflected by timing alone.

Conclusions: ML-based analysis of sensor-based chair rise performance identifies functional performance profiles beyond conventional timing metrics. Integrating biomechanical movement analytics into digital health assessment frameworks may improve early detection of mobility decline and support longitudinal monitoring and personalized intervention planning in older adults.

(JMIR Aging 2026;9:e94597) doi:10.2196/94597

KEYWORDS

healthy aging; digital health; machine learning; clustering; biomechanics; functional assessment; mobility; chair rise; older adults; frailty; movement analysis

XSL•FO
RenderX

Introduction

Striving for healthy aging and maintaining independence among older adults is a critical challenge in our rapidly aging global population [1,2]. The World Health Organization defines healthy aging as the process of developing and maintaining functional ability that enables well-being in older age [2,3]. However, functional ability tends to decline with age [4], impacting mobility, autonomy, and quality of life.

Functional ability is characterized by an individual's capacity to perform tasks essential for independent living, such as meeting basic needs, making decisions autonomously, maintaining social relationships, and moving independently [3,5-7]. Identifying the threshold at which age-related decline compromises mobility and independence, and understanding compensatory movement strategies individuals adopt to overcome these deficits, is a key focus of healthy aging research [4]. Early detection and quantification of these compensations can inform timely interventions, helping to prolong independence and reduce care needs.

The Five Times Chair Rise Test (5CRT) is a widely used clinical assessment to evaluate lower-extremity functional ability and strength in older adults [8]. Conventionally, performance is measured by timing the duration of 5 sit-to-stand cycles using a stopwatch, yielding a single temporal metric [8,9]. While practical, this approach does not fully capture the complexity and nuance of the movement patterns involved, nor does it exploit the rich biomechanical data potentially available from force and motion analyses [10,11]. Older adults with physical impairments, such as muscle weakness, often develop compensatory strategies—including altered center-of-mass trajectories or increased reliance on stabilization forces—that may enable task completion but mask underlying deficits [4]. Detailed biomechanical characterization of the 5CRT phases can reveal early functional decline and compensatory adaptations not reflected in timing metrics alone.

Meyer et al [10] introduced a machine learning (ML)–based workflow that combines dynamic time warping (DTW) and k-means clustering to classify 5CRT movement patterns from multidimensional force plate data in older adults. DTW was used to compensate for nonlinear temporal variability in sit-to-stand execution processes. Signal comparisons are not based on absolute duration. Rather, DTW aligns force trajectories according to their morphological characteristics. This allows k-means clustering to identify groups with similar movement patterns despite differences in execution speed. Meyer et al [10] demonstrated that unsupervised clustering of DTW-derived features can reveal distinct movement strategies that are invisible to stopwatch timing [10]. This established the methodological feasibility of the approach. However, Meyer et al [10] focused on discovering and describing biomechanical patterns without investigating how these movement clusters relate to established clinical indicators of physical function, mobility, and frailty.

This study uses the DTW clustering framework of Meyer et al [10] to ensure methodological comparability and reproducibility. The innovation lies in its integration with an extensive battery of validated clinical and functional assessments, including the Handgrip Strength (HGS) test, the Stair Climb Power Test (SCPT), the Timed "Up and Go" (TUG) test, the Short Physical Performance Battery (SPPB), including the 5CRT, the 4-meter Gait Speed (4mGS) test, and the Balance Test (BT), as well as the 6-minute Walking Test (6minWT) and the Countermovement Jump (CMJ). By directly correlating DTW-derived 5CRT movement clusters with these multidimensional performance measures, this study addresses a critical gap.

Preliminary analyses by Meyer et al [10] identified 2 distinct DTW-derived movement clusters with clear biomechanical and functional performance characteristics in a cohort of community-dwelling older adults. On average, cluster 1 was characterized by faster 5CRT completion times, lower vertical ground reaction forces (GRFs) during the stabilization phase, and higher mobility and strength assessment scores, indicating greater physical function and more efficient postural control. Conversely, cluster 2 exhibited slower chair rise times, higher stabilization forces, and poorer performance in mobility and functional tests. This suggests diminished lower-extremity power, potential balance limitations, and an increased risk of mobility-related decline. Contextualizing these biomechanically defined performance profiles within established geriatric assessments allows for a clinically meaningful classification, extending the applicability of ML-based motion analysis to real-world screening and intervention planning.

This work is not a methodological comparison but rather a clinical validation and transfer study. By applying the DTW and k-means clustering workflow by Meyer et al [10], methodological comparability is ensured, enabling validation of their findings in a longitudinal cohort while extending the analysis to its clinical implications across multiple functional domains.

The specific aims of this study are:

1. To classify 5CRT movement patterns using the established DTW and k-means workflow by Meyer et al [10], ensuring methodological comparability and enabling the investigation of aging trajectories in sensor-based movement characteristics.
2. To examine the association between DTW-defined 5CRT movement clusters and key indicators of lower extremity strength, mobility, exercise capacity, and functional performance as measured by HGS, SCPT, TUG, SPPB, 5CRT, 4mGS, BT, 6minWT, and CMJ, thereby exploring their relationship to interindividual variability and aging trajectories.
3. To evaluate whether ML-based 5CRT movement pattern classification provides clinically meaningful differentiation of functional status beyond conventional stopwatch-based 5CRT timing, and whether this approach contributes to capturing and monitoring functional aging trajectories over time.

This combined approach of traditional stopwatch timing and sensor-based, ML-driven movement pattern classification is used to create actionable functional performance profiles. These profiles could aid in the early identification of mobility decline, inform intervention planning, and ultimately support healthy

XSL•FO
RenderX

aging trajectories. Additionally, linking the analysis to established clinical outcomes connects biomechanical analytics to the decision-making requirements of geriatric care and digital health applications.

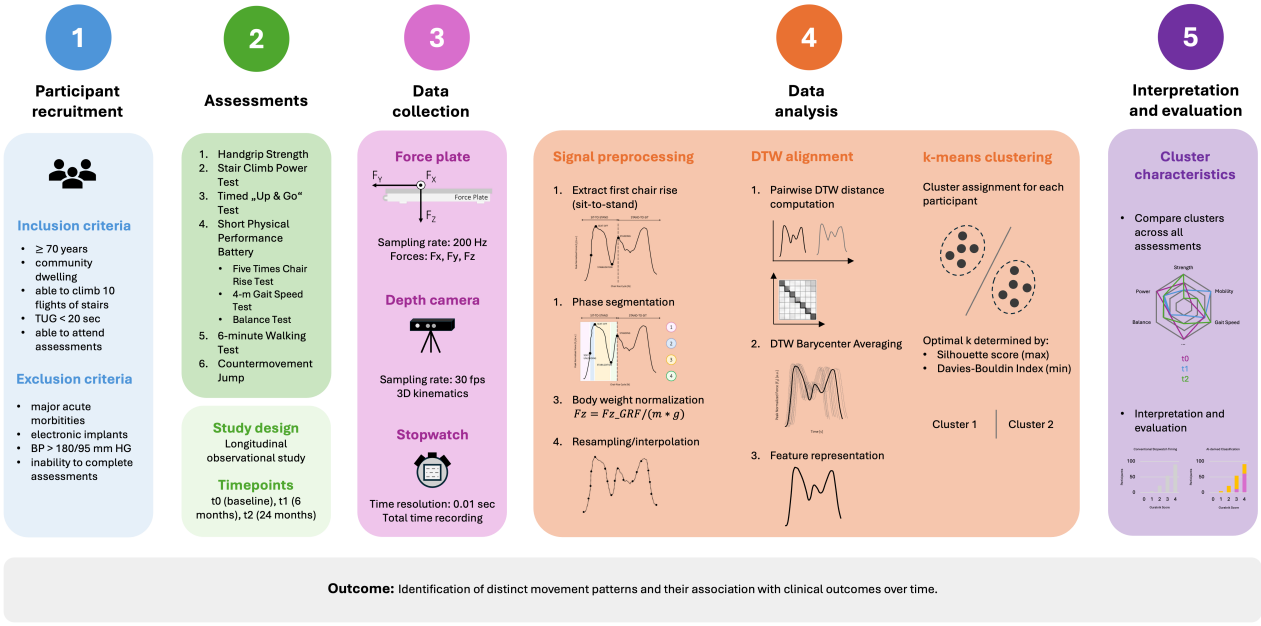
Methods

Study Participants

The study included a cohort of healthy, community-dwelling older adults [12,13]. Participants were recruited across diverse settings, including sports clubs, senior centers, music societies, rehabilitation sports facilities, and physiotherapy departments, as well as through newspaper advertisements. Functional geriatric assessments were systematically conducted under strict inclusion criteria. Participants had to be at least 70 years old and live in the community (Figure 1). They could not be included in the study if they had major acute morbidities such as pulmonary, renal, or cardiac diseases. They had to be able to climb 10 flights of stairs without difficulty and be self-sufficient enough to attend assessment sessions. They could not have pacemakers or other electronic implants. They had to be able to complete a TUG test in less than 20 seconds.

Prior to the assessment, participants underwent a telephone interview and received the written informed consent form at least one week in advance. Written informed consent was obtained from all participants prior to their participation. Additionally, a thorough medical history was collected, including a semistructured questionnaire on health status (eg, hypertension, history of stroke, chronic conditions such as diabetes mellitus and chronic obstructive pulmonary disease (COPD), falls, and general health) and a review of current medications. For safety reasons, blood pressure was measured, and participants with a reading exceeding 180/95 mm Hg were excluded from the study.

Figure 1. Graphical workflow of the study, summarizing participant recruitment, clinical assessment, biomechanical data acquisition, signal processing, dynamic time warping alignment, and machine learning–based clustering analysis. BP: blood pressure; DTW: dynamic time warping; TUG: Timed "Up & Go."



Study Design

This longitudinal observational study included 2 follow-up assessments (Figure 1). Eligible participants were evaluated at the Carl von Ossietzky Universität Oldenburg study center in Germany. Three assessment time points were conducted: (1) baseline (t0), (2) 6-month (t1) follow-up, and (3) 24-month (t2) follow-up.

Prior to each assessment, participants' height and weight were measured using a stadiometer. Participants then completed a standardized battery of the following physical performance tests (Table 1), administered consistently according to standardized protocols.

XSL•FO
RenderX

Table . Overview of physical performance tests applied in the study and corresponding test procedures.

Physical performance test	Test procedure
1. HGS[a]	
Reduced HGS = frailty criteria	• Measured with a JAMAR handheld dynamometer • Participant seated during the test • Both hands tested alternately, 3 trials per hand • Highest maximum value from 3 trials (left or right hand, whichever is stronger) is used • Strength values evaluated based on age and BMI according to [14]
2. SCPT[b]	
Leg power	• Participant climbs a flight of 10 stairs as quickly as possible without running or skipping steps • Time to climb the stairs is recorded • Body mass and total stair height are measured • Related power was calculated according to [15]: P=m x g x h/t (W), m=body mass (kg), g=9,81 m/s^{2,} h=height (m), t=time (s)
3. TUG[c]	
Functional mobility	• Participant sits with back against the chair • The signal “Go” starts the test • Participant stands up, walks 3 meters to a marked line, turns around, walks back, and sits down again • Time is measured from standing up to sitting down • Assistive devices allowed, but no physical help from others
4. SPPB[d]	
4.1. lower extremity physical function	• The test battery included the 5CRT[e], 4mGS,[f] and BT[g] • The time (in seconds) it takes to complete each test, as well as the cumulative score, are used to evaluate the results according to [8]
4.2. 5CRT	• Participant sits on a chair, and a signal starts the test • Participant stands up and sits down 5 times in a row as quickly as possible • Total time for all 5 repetitions is recorded with a stopwatch
4.3. 4mGS	• Participant walks a marked 4 meter distance at a normal pace • Time needed is recorded with a stopwatch
4.4. BT	• Semi-tandem and tandem stand • Participant holds specific positions (semi-tandem or full tandem stand) for a set of time • Ability to maintain position without support is assessed
5. 6minWT[h]	
Functional exercise capacity Endurance performance	• Participant walks up and down a 20 meter corridor for 6 minutes at the normal pace • Total distance covered is recorded in meters • Tester accompanies participant for safety reasons • Time needed is measured with a stopwatch
6. CMJ[i]	

XSL•FO
RenderX

Physical performance test	Test procedure
Functional performance	• Participant stands with feet at hip width and knees bent at about 90° at starting position on a ground reaction force platform (AMTI AccuPower, 200 Hz) • When jumping as high as possible, the knees are straightened (no leg tuck or arm instruction) • One preliminary jump: 3 test jumps with one-minute rest; weight is measured before first jump • Maximum jump power (W) was calculated from the 3 best jumps • The CMJ test was conducted at t0 and t1, but not at t2 for safety reasons

[a]HGS: Handgrip Strength.
[b]SCPT: Stair Climb Power Test.
[c]TUG: Timed "Up & Go" Test.
[d]SPPB: Short Physical Performance Battery.
[e]5CRT: Five Times Chair Rise Test.
[f]4mGS: 4-meter Gait Speed test.
[g]BT: Balance Test.
[h]6minWT: 6-minute Walking Test.
[i]CMJ: Countermovement Jump.

Assessments of Physical Function

To ensure methodological comparability with Meyer et al [10], the DTW and k-means clustering workflow was applied to the 5CRT at the 24-month follow-up (t2) assessment. Recent methodological advances have demonstrated the potential of using the biomechanical data collected during the 5CRT for time series analysis and unsupervised clustering. This allows for the identification of clinically relevant functional subgroups [10]. Thus, the 5CRT offers a chance to integrate quantitative biomechanical parameters with established clinical assessment frameworks. The detailed protocol for the 5CRT used in this study is described below.

Participants performed the 5CRT by rising from a chair to a fully stable standing position and sitting down again as quickly as possible, following the protocol described by Guralnik et al [8], with them being seated upright in the center of a height-adjustable chair without armrests. Participants started with their knees flexed at a 90-degree angle (verified by a goniometer) and their feet placed shoulder-width apart on the force plate positioned directly in front of the chair. Throughout the test, arms were crossed in front of the chest to prevent the use of arm support. A verbal command initiated the test. Participants were instructed to stand up completely (sit-to-stand) and then sit back down with their buttocks touching the chair (stand-to-sit). They repeated this movement cycle 5 times consecutively without using their arms. The total time to complete the 5 rises was recorded using a Schütt PC-90/60 Memory stopwatch with a time resolution of 1/100 second and a memory function for intermediate times (up to 60 splits). It was then scored according to Guralnik et al [8] (Table 2). Timing started with the command and stopped once the participant completed the fifth rise and made full contact with the back of the chair. All measurements were conducted by trained physiotherapists to ensure consistency.

Table . The 5CRT[a] is scored based on the measured test duration, as used in the SPPB[b] as a frailty measure according to Guralnik et al [8]. A higher score indicates better performance in the 5CRT.

5CRT[a] durations (in seconds)	Scores
≤11.19	4 points
11.20 - 13.69	3 points
13.70 - 16.69	2 points
16.70 - 60	1 point
>60 or unable to complete the 5CRT	0 points

[a]5CRT: Five Times Chair Rise Test.
[b]SPPB: Short Physical Performance Battery.

Data Collection

During the test, external GRFs were recorded at a sampling rate of 200 Hz using a calibrated AMTI AccuPower force plate (Figure 1). Participants stood on this plate with both feet. The plate was not moved during or between the assessments. Forces were captured in 3 directions during the assessment: medial-lateral (Fx), anterior-posterior (Fy), and vertical (Fz). Before each assessment, the force plate was zeroed and calibrated by inputting participant-specific data such as sex,

XSL•FO
RenderX

height, and weight. Additionally, a Microsoft Kinect V2 depth camera (30 fps) recorded movements in 3 dimensions, enabling correlation of kinematic data with kinetic force measurements for comprehensive biomechanical signal interpretation.

Data Analysis

Data analysis was performed using Python (Python Software Foundation) [16], including the *tslearn* package [17]. Raw force plate and motion capture data were synchronized, visually inspected for artifacts, and segmented into individual sit-to-stand trials [10] (Figure 1). Using the standardized, clinically established 5CRT as the basis for the clustering process ensures compatibility with widely used frailty and functional assessments and allows the use of detailed kinetic and kinematic data available from force plate and motion capture systems. This strategy allows to characterize the heterogeneity of functional ability in older adults, supporting a more individualized, data-driven approach to understanding and monitoring the decline of age-related mobility.

The sit-to-stand movement was analyzed by segmenting it into 4 distinct phases based on the model by Schenkman et al [18] (Figure 2):

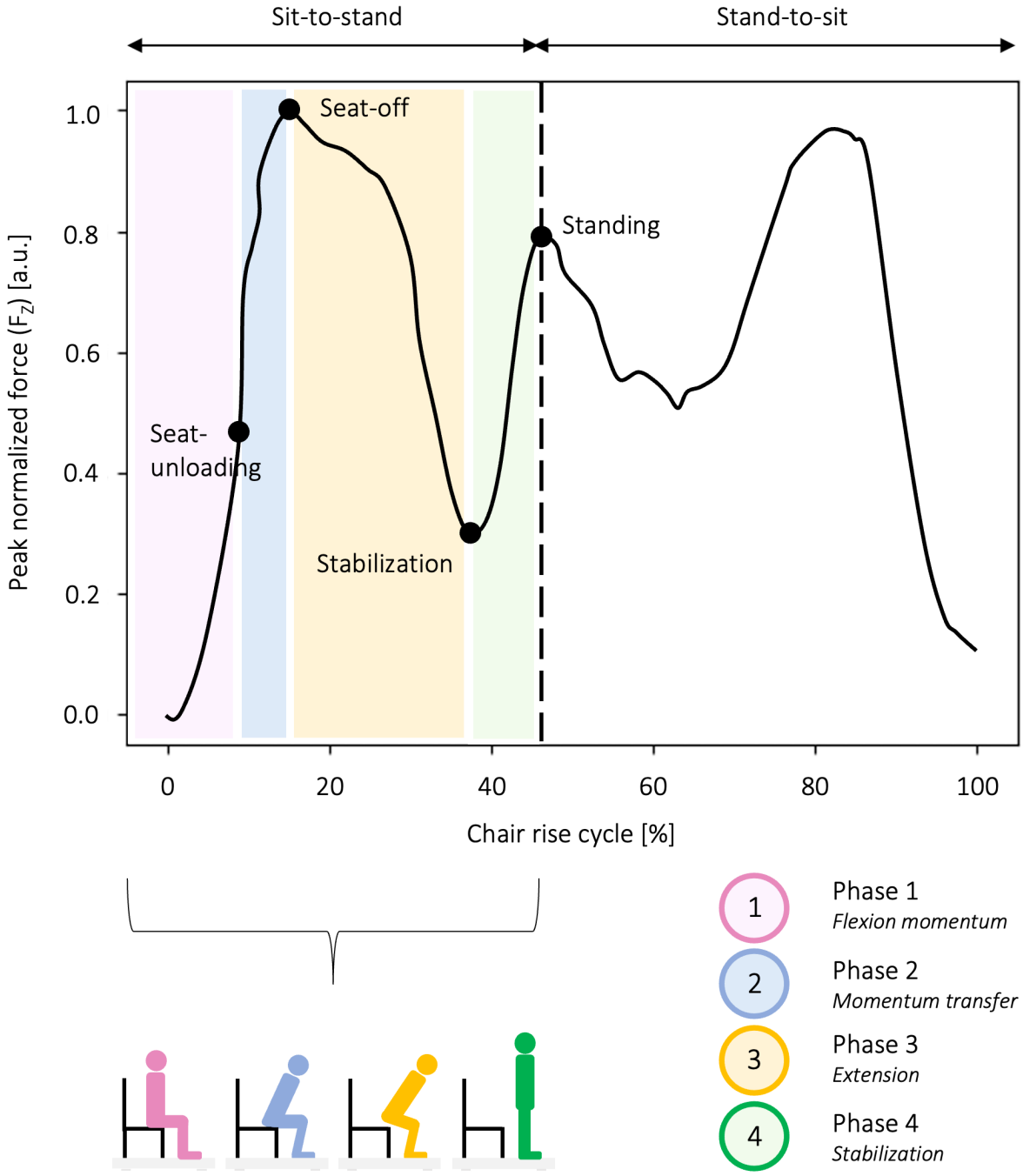
1. Phase 1: The "Flexion Momentum Phase" begins with movement initiation and continues until just before the participant's buttocks lift off the chair (seat-off). This phase is characterized by trunk and pelvic flexion, as well as anticipatory postural adjustments primarily involving the head, arms, and trunk.
2. Phase 2: The "Momentum Transfer Phase," spans from seat-off to maximum ankle dorsiflexion. During this phase, the body's center of mass shifts significantly forward and upward.
3. Phase 3: The "Extension Phase" starts immediately after maximum ankle dorsiflexion and concludes with full hip and knee extension.
4. Phase 4: The "Stabilization Phase" marks the final upright standing position, with the participant fully balanced.

To address the temporal variability and nonlinear time distortions inherent in human movement sequences, DTW was applied as a time series similarity measure [10]. For each participant, the vertical GRF signal (Fz) was peak-normalized to body weight, resampled to a common temporal length, and used as the input time series for DTW-based comparison. DTW aligns temporal force profiles to reveal latent spatiotemporal patterns within participant-specific movement cycles. This enables comparisons independent of the absolute duration of the chair rise.

The clustering model was based on previously derived centroids by Meyer et al [10] using k-means clustering with DTW Barycenter Averaging (DBA) as implemented in *tslearn* [17]. Clusters were assigned based on each participant's first chair rise performance. Therefore, for initial cluster assignment, only the first chair rise from the 5CRT sequence was used. This was done to minimize potential influences of fatigue, learning effects, or adaptation on movement strategy that could occur over successive repetitions. Subsequent chair rises were classified using the same model to evaluate stability and changes in movement patterns across repetitions. To reduce the risk of local minima, the clustering procedure was repeated 10 times with different centroid initializations, and the maximum number of iterations was set to 100 [10].

XSL•FO
RenderX

Figure 2. The different phases of the sit-to-stand part of the Five Times Chair Rise Test, as defined by Schenkman et al [18], can be seen in the Fz component of the ground reaction force data of a random participant. Phase 1 (flexion momentum): initiation of the movement. Phase 2 (momentum transfer): seat unloading until maximum dorsiflexion is reached. Phase 3 (extension): the buttocks leave the chair and the whole body moves upwards. Phase 4 (stabilization): The hip and knee joints are fully extended. The sit-to-stand transition is completed when the participant stands fully upright. The color-coded schematics illustrate the corresponding body posture during each phase of the sit-to-stand cycle according to Schenkman et al [18].



Evaluation

The evaluation of unsupervised clustering, which is inherently challenging due to its lack of predefined labels, required the application of multiple validation metrics (Figure 1). Silhouette analysis was initially used to evaluate the separation of clusters by measuring the mean intra versus intercluster distances for k=2 - 10 (*tslearn* implementation following Rousseeuw [17]). Higher values in this case indicate better separation, with a score near to 1 denoting well-defined clusters, and a score near to 0 indicating partial overlap, and a negative score indicates misclassification. The k-value that yielded the highest silhouette score was selected for clustering.

Subsequent assessment of cluster validity was conducted using the Davies-Bouldin Index [19], a quantitative metric that calculates the ratio of intracluster compactness to intercluster separation.

The Davies-Bouldin Index [19] is defined as follows:

$$\bar{R} \equiv 1N \sum i=1k \max j \neq i \, (Si+SjMij)$$

XSL•FO
RenderX

where k is the number of clusters, S_i and S_j denote the within-cluster scatter of clusters i and j, and M_{ij} represents the distance between the centroids of clusters i and j. Lower values indicate greater homogeneity and separation between clusters.

Statistical Analysis

The Shapiro-Wilk test was used to test the normal distribution of the force and duration values obtained from the chair rise dataset. Since the dataset contained nonnormally distributed data, the nonparametric Mann-Whitney *U* test was used to compare the clustered 5CRT groups. Statistical significance was defined as $P<.01$. The analysis was performed in Python [16] using the *SciPy* package.

Then, the 5CRT clusters were correlated with the conducted physical performance tests (Table 1) across the 3 time points (baseline: t0, 6 months: t1, and 24 months: t2). Independent samples *t* tests (2-tailed) were used to compare continuous outcomes between the 2 clusters at t0, t1, and t2. For categorical variables, the Pearson chi-square test was applied. Where normality assumptions were not met, the Mann-Whitney *U* test served as a nonparametric alternative. Normality and homogeneity of variance were assessed using the Shapiro-Wilk test and the Levene test, respectively. Statistical significance was set at $P<.05$, $P<.01$, and $P<.001$. Effect sizes were reported as Cohen *d* with 95% CIs. To compare effect sizes across assessments and time points, Cohen *d* values were visualized in a radar plot. All analyses were performed using IBM SPSS Statistics for Windows, version 30.0.0 (171).

Ethical Considerations

The study was approved by the Medical Ethics Committee of Hanover Medical School in Germany (approval number 6948) and was conducted in accordance with the Declaration of Helsinki [20], and all relevant guidelines and regulations. All participants provided written informed consent. Participant data were pseudonymized, and identifying information was stored separately from the study data. Access to the data was restricted to authorized study personnel, and electronic data were stored on password-protected or access-restricted systems. During transcription, identifying information was removed or modified. The participants identification list was deleted after the study was completed. Participants received reimbursement for travel expenses but no other compensation.

Results

Characteristics of Study Participants

A total of 251 participants were initially enrolled at baseline (t0). Four participants dropped out by the 6-month follow-up (t1), and an additional 19 participants were lost to follow-up by 24 months (t2; Table 3).

The following sections present the results in the order of the predefined study aims.

Table . Characteristics of study participants at baseline (t0), 6-month (t1), and 24-month (t2) follow-up assessments[a].

Assessments	Study participants (N)	Age (years)		Sex	
		Range	Mean (SD)	Female n/N (%)	Male n/N (%)
Baseline (t0)	251	70-87	75.10 (3.89)	148/251 (59)	103/251 (41)
6 months (t1)	247	70-88	75.87 (3.84)	146/247 (59)	101/247 (41)
24 months (t2)	228	71-89	77.23 (3.78)	132/228 (58)	96/228 (42)

[a] Values are presented as mean (SD), including minimum (Min), and maximum (Max) values for age (years) and sex (female and male).

Classification of 5CRT Movement Patterns Using ML-Based Clustering

At t2, 228 participants were included in the assessment, and 191 of these participants completed the 5CRT. Nine participants were then excluded from data analysis due to errors in the recorded measurement data, and 10 participants were excluded because they failed to complete the required 5 chair rises. Consequently, the final analytic 5CRT sample at t2 comprised 172 participants with a mean age of 76.90 (SD 3.40) years (Table 4). Among them, 96 (56%) were female and 76 (44%) were male.

A total of 860 chair rises from the t2 assessment were categorized into 2 clusters (clusters 1 and 2) using k-means clustering based on the DBA algorithm [10]. Silhouette analysis indicated that 2 clusters were the optimal solution (Davies-Bouldin index~0.92). The 2 resulting clusters were defined by the applied force (Fz) during the chair rise process. Figure 3 shows the soft-DTW barycenter to visualize the different clusters.

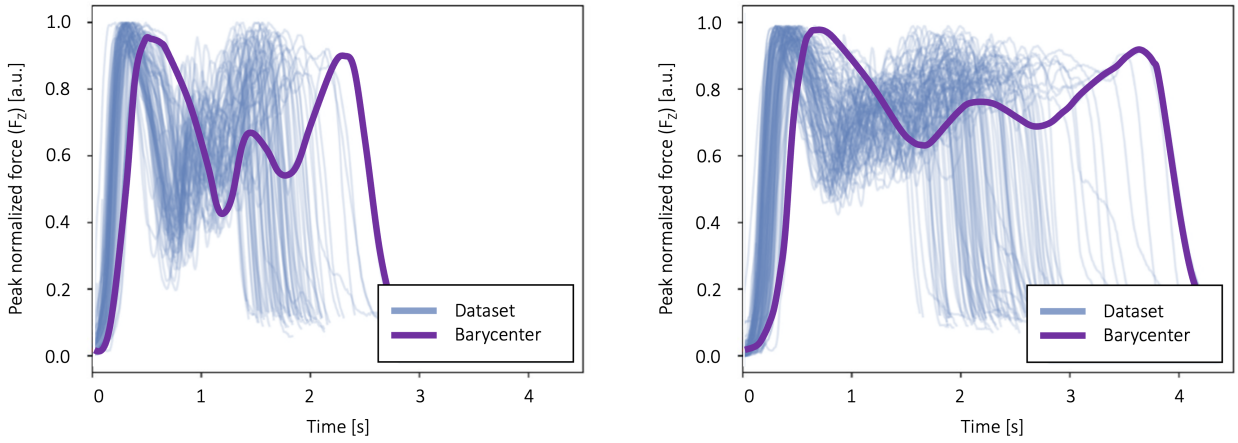
XSL•FO
RenderX

Table . Study participants' characteristics at 24-month follow-up (t2) measured via stopwatch.

Characteristics	Mean (SD)	Range
Age (years)	76.90 (3.40)	71-89
Weight (kg)	75.10 (13.40)	48.10-126
Height (m)	1.67 (0.09)	1.45-1.91
5CRT[a] stopwatch duration (seconds)	11.44 (2.49)	6.44-20.21

[a]5CRT: Five Times Chair Rise Test.

Figure 3. Dataset and barycenter of cluster 1 (left) and cluster 2 (right) after k-means clustering according to Meyer et al [10].



K-means clustering assigned 72 (42%) participants to cluster 1 and 100 (58%) participants to cluster 2 (Tables 5-7). Men and women were distributed unevenly across clusters: 47% (n=40) of male participants were assigned to cluster 1, while 53% (n=44) were categorized in cluster 2. Overall, 37% (n=32) of the female participants were assigned to cluster 1, while 63% (n=56) were assigned to cluster 2.

XSL•FO
RenderX

Table . Demographic and physical performance outcomes for participants in DTW[a]-defined 5CRT[b] movement clusters (cluster 1 and 2) according to Meyer et al [10].

Variables at t0 (baseline)	Cluster 1 (n=72)	Cluster 2 (n=100)	*P*[c] value	Effect size, Cohen *d* (95% CI)
Age (years), mean (SD)	74.40 (2.70)	75.30 (3.90)	.10	–0.259 (–0.563 to 0.045)
Sex, n (%)			.19[d]	
Male	36 (50)	40 (40)		
Female	36 (50)	60 (60)		
BMI (kg/m^2), mean (SD)	25.64 (2.90)	27.97 (4.09)	<.001	–0.640 (–0.949 to –0.328)
HGS[e] (kg), mean (SD)	32.41 (9.88)	28.07 (9.60)	.004	0.447 (0.140 to 0.753)
SCPT[f] (W), mean (SD)	236.15 (51.30)	214.66 (47)	.005	0.440 (0.133 to 0.746)
TUG[g] (s), mean (SD)	7.42 (0.95)	8.74 (1.44)	<.001	–1.047 (–1.368 to –0.723)
SPPB[h] total score, mean (SD)	11.35 (0.75)	10.93 (1.10)	.006	0.429 (0.122 to 0.735)
5CRT stopwatch duration (s), mean (SD)	11.15 (2.07)	12.80 (3.30)	<.001	–0.581 (–0.889 to –0.271)
4mGS[i] (s), mean (SD)	2.44 (0.30)	2.87±0.42)	<.001	–1.167 (–1.493 to –0,839)
4mGS (m/s), mean (SD)	1.66 (0.22)	1.42 (0.20)	<.001	1.190 (0.860 to 1.517)
6minWT[j] distance (m), mean (SD)	490 (60.27)	421.21 (60.34)	<.001	1.141 (0.813 to 1.465)
Semitandem stand (points), mean (SD)	1 (0)	0.99 (0.10)	.40[k]	
Tandem stand (points), mean (SD)	1.99 (0.12)	1.92 (0.32)	.06[l]	
CMJ[m] height (cm), mean (SD)	13.07 (4.71)[n]	9.55 (4.23)[o]	<.001	0.793 (0.463 to 1.121)
CMJ power (W), mean (SD)	2050.17 (577.12)[n]	1816.86[o] (482.32)	.007	0.445 (0.124 to 0.765)

[a]DTW: dynamic time warping.

[b]5CRT: Five Times Chair Rise Test.

[c]*P* values are derived from independent samples *t* tests (highly significant: *P*<.001, significant: *P*<.01, marginally significant: *P*<.05).

[d]Derived from Pearson chi-squared test for categorical variables: χ^2_1=1.697

[e]HGS: Handgrip Strength.

[f]SCPT: Stair Climb Power Test.

[g]TUG: Timed “Up & Go.”

[h]SPPB: Short Physical Performance Battery.

[i]4mGS: 4-meter Gait Speed test.

[j]6minWT: 6-minute Walking Test.

[k]Derived from Mann-Whitney *U* test for nonparametric data: *U*=3564.

[l]Derived from Mann-Whitney *U* test for nonparametric data: *U*=3361.5.

[m]CMJ: Countermovement Jump.

[n]n=66 for cluster 1.

[o]n=91 for cluster 2.

XSL•FO
RenderX

Table . Demographic and physical performance outcomes for participants in DTW[a]-defined 5CRT[b] movement clusters (cluster 1 and 2) according to Meyer et al [10].

Variables at t1 (6 months)	Cluster 1 (n=72)	Cluster 2 (n=100)	*P*[c] value	Effect size, Cohen *d* (95% CI)
Age (years), mean (SD)	74.94 (2.65)	75.84 (3.92)	.10	–0.260 (–0.564 to 0.045)
Sex n (%)			.19[d]	
Male	36 (50)	40 (40)		
Female	36 (50)	60 (60)		
BMI (kg/m^2), mean (SD)	25.55 (2.87)	27.90 (4.10)	<.001	–0.647 (–0.956 to –0.335)
HGS[e] (kg), mean (SD)	31.92 (9.70)	27.18 (9.41)	.002	0.497 (0.188 to 0.803)
SCPT[f] (W), mean (SD)	235.82 (47.59)	212.23 (44.72)	.001	0.513 (0.205 to 0.820)
TUG[g] (s), mean (SD)	7.40 (0.98)	8.80 (1.59)	<.001	–1.024 (–1.344 to –0.700)
SPPB[h] total score, mean (SD)	11.51 (0.71)	10.82 (1.16)	<.001	0.696 (0.384 to 1.007)
5CRT stopwatch duration (s), mean (SD)	10.38 (1.83)	12.49 (2.85)	<.001	–0.852 (–1.167 to –0.535)
4mGS[i] (s), mean (SD)	2.41 (0.31)	2.81 (0.46)	<.001	–1.008 (–1.328 to –0.686)
4mGS (m/s), mean (SD)	1.70 (0.23)	1.46 (0.22)	<.001	1.046 (0.722 to 1.367)
6minWT[j] distance (m), mean (SD)	499.73 (61.27)	430.25 (59.02)	<.001	1.159 (0.830 to 1.484)
Semitandem stand (points), mean (SD)	1 (0)	1 (0)	>.99[k]	
Tandem stand (points), mean (SD)	1.85 (0.36)	1.83 (0.43)	>.99[l]	
CMJ[m] height (cm), mean (SD)	12.86 (4.47)[n]	9.05 (3.84)[o]	<.001	0.926 (0.572 to 1.283)
CMJ power (W), mean (SD)	1952.28 (538.34)[n]	1766.56 (446.85)[o]	.03	0.383 (0.041 to 0.724)

[a]DTW: dynamic time warping.

[b]5CRT: Five Times Chair Rise Test.

[c]*P* values are derived from independent samples *t* tests (highly significant: *P*<.001, significant: *P*<.01, marginally significant: *P*<.05),

[d]Derived from Pearson chi-squared test for categorical variables: χ^2_1=1.697.

[e]HGS: Handgrip Strength.

[f]SCPT: Stair Climb Power Test.

[g]TUG: Timed “Up & Go.”

[h]SPPB: Short Physical Performance Battery.

[i]4mGS: 4-meter Gait Speed test.

[j]6minWT: 6-minute Walking Test.

[k]Derived from Mann-Whitney *U* test for nonparametric data: *U*=3600.

[l]Derived from Mann-Whitney *U* test for nonparametric data: *U*=3599.

[m]CMJ: Countermovement Jump.

[n]n=55 for cluster 1.

[o]n=86 for cluster 2.

XSL•FO
RenderX

Table . Demographic and physical performance outcomes for participants in DTW[a]-defined 5CRT[b] movement clusters (clusters 1 and 2) according to Meyer et al [10].

Variables at t2 (24 months)	Cluster 1 (n=72)	Cluster 2 (n=100)	*P*[c] value	Effect size, Cohen *d* (95% CI)
Age (years)	76.42 (2.70)	77.30 (3.85)	.09	–0.264 (–0.568 to 0.040)
Sex, n (%)			.19[d]	
Male	36 (50)	40 (40)		
Female	36 (50)	60 (60)		
BMI (kg/m^2), mean (SD)	25.60 (2.80)	27.850 (3.95)	<.001	–0.641 (–0.951to –0.330)
HGS[e] (kg), mean (SD)	27.64 (9.84)	23.21 (9.12)	.003	0.471 (0.163-0.777)
SCPT[f] (W), mean (SD)	222.67 (48.01)	193.05 (42.56)	<.001	0.659 (0.348-0.969)
TUG[g] (s), mean (SD)	7.55 (0.97)	9.52 (1.53)	<.001	–1.487 (–1.827 to –1.144)
SPPB[h] total score, mean (SD)	11.69 (0.57)	10.59 (1.08)	<.001	1.219 (0.888-1.547)
5CRT stopwatch duration (s), mean (SD)	9.84 (1.42)	12.60 (2.48)	<.001	-1.311 (-1.643, - 0.976)
4mGS[i] (s), mean (SD)	1.68 (0.19)	1.38 (0.20)	<.001	1.495 (1.152-1.836)
4mGS (m/s), mean (SD)	2.42 (0.28)	2.95 (0.44)	<.001	–1.411 (–1.747to –1.071)
6minWT[j] distance (m), mean (SD)	481.35 (65.99)[k]	411.74 (60.08)[l]	<.001	1.111 (0.777-1.443)
Semitandem stand (points), mean (SD)	1 (0)[m]	0.96 (0.20)[n]	.09[o]	
Tandem stand (points), mean (SD)	1.86 (0.39)[m]	1.62 (0.65)[n]	.008[p]	
CMJ[q] height (cm)	—[r]	—	—	—
CMJ power (W)	—	—	—	—

[a]DTW: dynamic time warping.

[b]5CRT: Five Times Chair Rise Test.

[c]*P* values are derived from independent samples *t* tests (highly significant: *P*<.001, significant: *P*<.01, marginally significant: *P*<.05).

[d]Derived from Pearson chi-squared test for categorical variables: χ^2_1=1.697.

[e]HGS: Handgrip Strength.

[f]SCPT: Stair Climb Power Test.

[g]TUG: Timed “Up & Go.”

[h]SPPB: Short Physical Performance Battery.

[i]4mGS: 4-meter Gait Speed test.

[j]6minWT: 6-minute Walking Test.

[k]n=69 for cluster 1.

[l]n=95 for cluster 2.

[m]n=72 for cluster 1.

[n]n=100 for cluster 2.

[o]Derived from Mann-Whitney *U* test for nonparametric data: *U*=3456.

[p]Derived from Mann-Whitney *U* test for nonparametric data: *U*=2980.

[q]CMJ: Countermovement Jump.

[r]Not available.

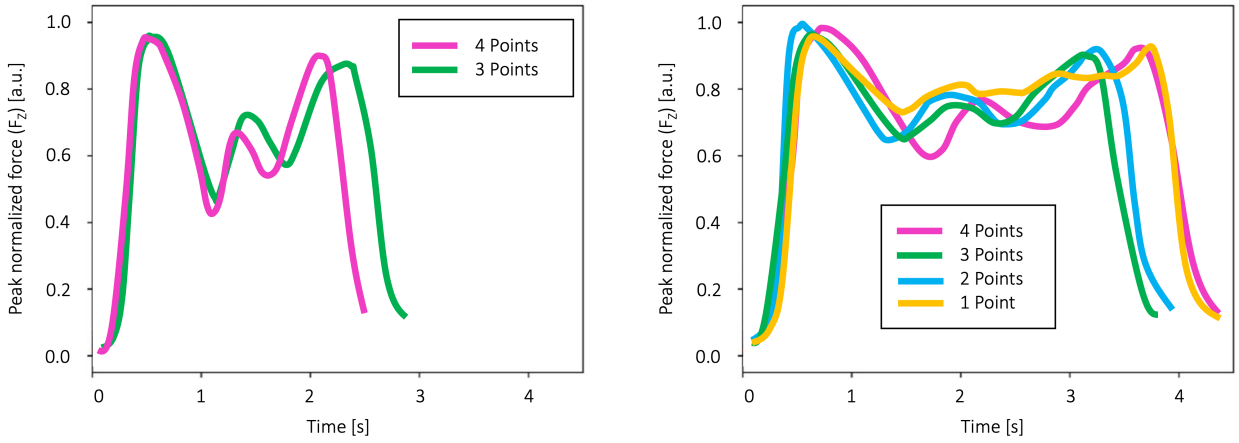
The t2 clusters differed significantly in several parameters (Table 7). To understand the effect of 5CRT duration, the clusters were divided by each score category, as presented by Guralnik et al [8]. Figure 4 shows the barycenter for each resulting dataset. Note that cluster 1 does not contain any participants with a 5CRT duration greater than 13 seconds, while cluster 2 contains participants in all score categories.

To sum up, 2 functional movement clusters (cluster 1 and 2) were identified among older adults performing the 5CRT

XSL•FO
RenderX

(Figures 3 and 4): cluster 1 was characterized by faster chair rise times and lower vertical GRFs during stabilization, reflecting better physical performance and postural control. Conversely, cluster 2 exhibited slower performance times and elevated forces during stabilization, suggesting compromised mobility and potential balance limitations. These clusters revealed underlying movement pattern differences that traditional stopwatch timing could not detect, demonstrating the enhanced sensitivity of ML-based time-series analysis.

Figure 4. Barycenter of subdatasets created by splitting cluster 1 (left) and cluster 2 (right) datasets according to Meyer et al [10] based on each scored category according to Guralnik et al [8].



Associations of ML-Based 5CRT Clusters With Clinical Physical Performance Tests

The demographic and physical performance characteristics of the 2 5CRT movement clusters identified by the ML-driven clustering approach proposed by Meyer et al [10] are presented in Tables 5-7 for each of the 3 assessment time points (t0, t1, and t2).

At baseline (t0; Table 5), age and sex distribution did not differ significantly between cluster 1 and cluster 2: participants in cluster 1 and cluster 2 were of comparable age (mean 74.40, SD 2.70 years vs mean 75.30, SD 3.90 years; Cohen *d*=–0.259; *P*=.09) and showed no significant difference in sex distribution (n=36, 50% male vs n=40, 40% male; *P*=.19). In contrast, both clusters differ in anthropometry and physical performance. Cluster 1 participants had a significantly lower BMI compared to cluster 2 participants (mean 25.64, SD 2.90 kg/m^2 vs mean 27.97, SD 4.09 kg/m^2; Cohen *d*=–0.640; *P*<.001). Measures of muscle strength and power were higher in cluster 1: participants in cluster 1 exhibited significantly higher HGS (mean 32.41, SD 9.88 kg vs mean 28.07, SD 9.60 kg; Cohen *d*=0.447; *P*=.004) and SCPT values (mean 236.15, SD 51.30 W vs mean 214.66, SD 47 W; Cohen *d*=0.440; *P*=.005) compared to cluster 2. Both clusters also differ in mobility and global physical function: compared to cluster 2, cluster 1 participants had faster TUG times (mean 7.42, SD 0.95 seconds vs mean 8.74, SD 1.44 seconds; Cohen *d*=–1.047; *P*<.001) and higher SPPB total scores (mean 11.35, SD 0.75 versus mean 10.93, SD 1.10, Cohen *d*=0.429; *P*=.006). The 5CRT performance showed differences between cluster 1 and cluster 2: the 5CRT stopwatch duration was significantly shorter in cluster 1 participants compared to cluster 2 participants (mean 11.15, SD 2.07 seconds vs mean 12.80, SD 3.30 seconds; Cohen *d*=–0.581; *P*<.001). Habitual gait speed, assessed via 4mGS, was faster in cluster 1 compared to cluster 2 (mean 2.44, SD 0.30 seconds vs mean 2.87, SD 0.42 seconds, Cohen *d*=–1.167; *P*<.001 and mean 1.66, SD 0.22 m/s vs mean 1.42, SD 0.20 m/s; Cohen *d*=1.190; *P*<.001). 6minWT distance was higher in cluster 1 (mean 490, SD 60.27 m vs mean 421.21, SD 60.34 m; Cohen *d*=1.141; *P*<.001). Balance measures showed near ceiling in both clusters: the semi-tandem stand revealed mean 1 (SD 0) points for cluster 1 participants vs mean 0.99 (SD 0.10) points for cluster 2 participants (*P*=.40). In contrast, tandem stance showed mean 1.99, (SD 0.12) points in cluster 1 and mean 1.92 (SD 0.32) points in cluster 2 (*P*=.05). Explosive lower-limb performance, assessed by CMJ, also differed between cluster 1 and cluster 2 at baseline: cluster 1 participants showed higher CMJ height (mean 13.07, SD 4.71 cm vs mean 9.55, SD 4.23 cm; Cohen *d*=0.793; *P*<.001) and power (mean 2050.17, SD 577.12 W versus mean 1816.86, SD 482.32 W, Cohen *d*=0.445; *P*<.007) than cluster 2 participants.

After 6 months (t1; Table 6), the patterns observed at baseline (t0) continue to show similar demographic profiles but differences in physical performance. Age and sex distribution did not differ significantly between cluster 1 and cluster 2: participants in cluster 1 and cluster 2 were of comparable age (mean 74.94, SD 2.65 years vs mean 75.84, SD 3.92 years, Cohen *d*=–.260; *P*=.09) and showed no significant difference in sex distribution (50% [36] male vs 40% [40] male; *P*=.19). Cluster 1 participants had a significantly lower BMI compared to cluster 2 participants (mean 25.55, SD 2.87 kg/m^2 vs mean 27.90, SD 4.10 kg/m^2, Cohen *d*=–0.647; *P*<.001). At t1, measures of muscle strength and power were still higher in cluster 1: participants in cluster 1 showed significantly higher HGS (mean 31.92, SD 9.70 kg vs mean 27.18, SD 9.41 kg, Cohen *d*=0.497; *P*=.002) and SCPT values (mean 235.82, SD 47.59 W vs mean 212.23, SD 44.72 W, Cohen *d*=0.513; *P*=.001) compared to cluster 2. Both clusters also differ in mobility and global physical function: compared to cluster 2, cluster 1 participants still had faster TUG times (mean 7.40, SD 0.98 s vs mean 8.80, SD 1.59 s, Cohen *d*=–1.024; *P*<.001) and higher

XSL•FO
RenderX

SPPB total scores (mean 11.51, SD 0.71 vs mean 10.82, SD 1.16, Cohen *d*=–0.696; *P*<.001). The 5CRT performance still showed differences between cluster 1 and cluster 2 at t1: the 5CRT stopwatch duration was significantly short in cluster 1 participants compared to cluster 2 participants (mean 10.38, SD 1.83 seconds vs mean 12.49, SD 2.85 seconds, Cohen *d*=–0.852; *P*<.001). 4mGS was faster in cluster 1 compared to cluster 2 (mean 2.41, SD 0.31 seconds vs mean 2.81, SD 0.46 s, Cohen *d*=–1.008; *P*<.001 and mean 1.70, SD 0.23 m/s vs mean 1.46, SD 0.22 m/s, Cohen *d*=1.046; *P*<.001). In addition, 6-minute walk test distance was higher in cluster 1 (mean 499.73, SD 61.27 m vs mean 430.25, SD 59.02 m, Cohen *d*=1.159; *P*<.001). Balance measures also showed near ceiling at t1 in both clusters: The semi-tandem stand revealed mean 1 (SD 0) points for cluster 1 and cluster 2 participants (*P*>.99). In contrast, tandem stance showed mean 1.85 (SD 0.36) points in cluster 1 and mean 1.83 (SD 0.43) points in cluster 2 (*P*>.99). Explosive lower-limb performance also differed cluster 1 and cluster 2 at t1: compared to cluster 2, cluster 1 participants showed higher CMJ height (mean 12.86, SD 4.47 cm vs mean 9.05, SD 3.84 cm, Cohen *d*=0.926; *P*<.001) and power (mean 1952.28, SD 538.34 W vs mean 1766.56, SD 446.85 W, Cohen *d*=0.383; *P*=.03).

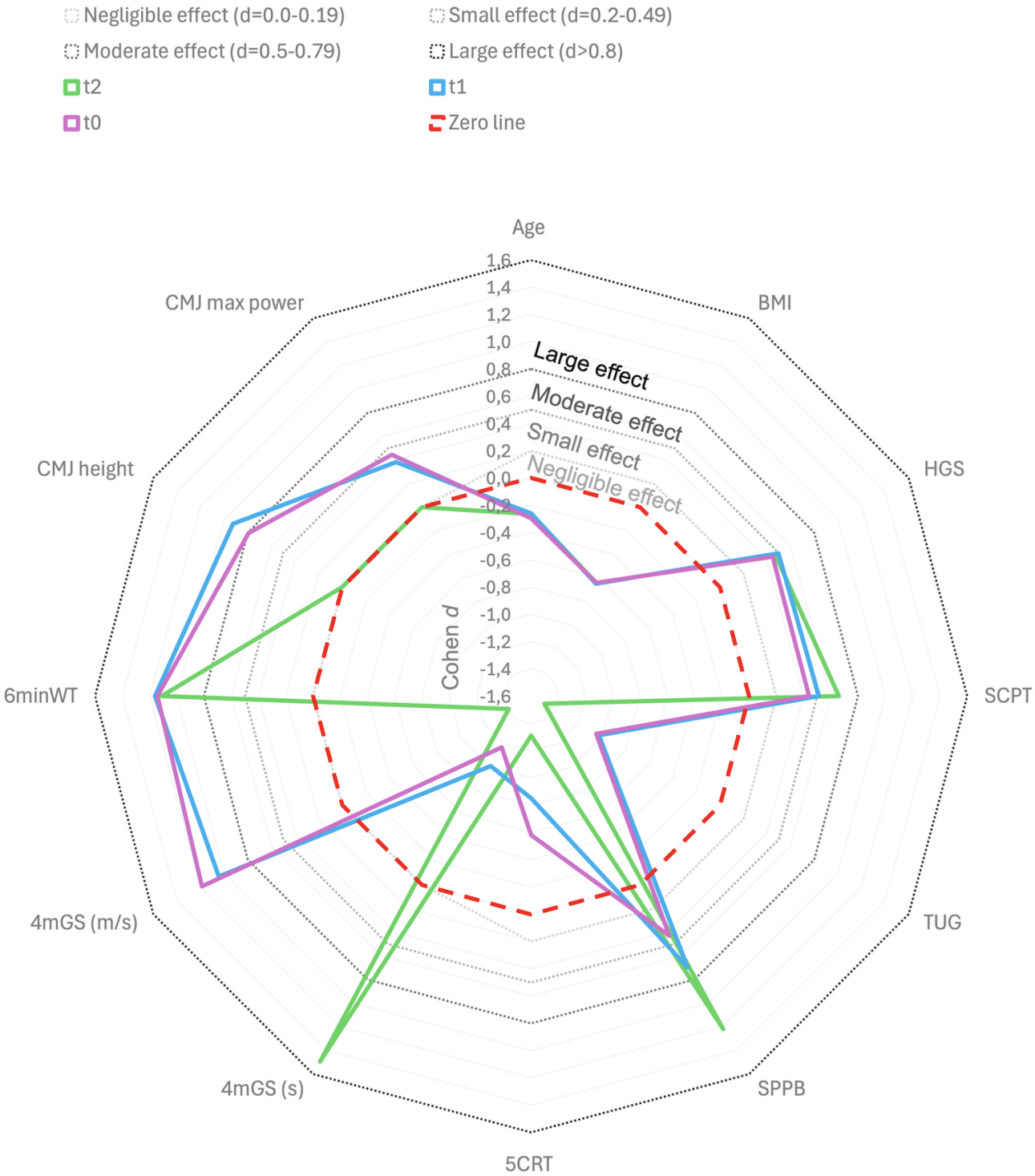
After 24 months (t2; Table 7), age and sex distributions remained similar across cluster 1 and cluster 2 (mean age 76.42, SD 2.70 years vs mean 77.30, SD 3.85 years, Cohen *d*=–0.264; *P*=.09; sex: n=36, 50% male vs n=40, 40% male; *P*=.19). The differences in BMI persisted as moderately higher in cluster 2 participants (mean 25.60, SD 2.80 kg/m² vs mean 27.85, SD 3.95 kg/m², Cohen *d*=–0.641; *P*<.001). At t2, several functional differences between cluster 1 and cluster 2 existed: cluster 1 participants maintained significantly higher HGS (mean 27.64, SD 9.84 kg vs mean 23.21, SD 9.12 kg, Cohen *d*=0.471; *P*=.003) and SCPT values (mean 222.67, SD 48.01 W vs mean 193.05, SD 42.56 W, Cohen *d*=0.659; *P*<.001). In addition, mobility and global performance differences further increased: TUG times remained faster in cluster 1 (mean 7.55, SD 0.97 seconds vs mean 9.52, SD 1.53 seconds, Cohen *d*=–1.487; *P*<.001), and SPPB total scores were higher (mean 11.69, SD 0.57 vs mean 10.59, SD 1.08, Cohen *d*=1.219; *P*<.001). The 5CRT performance showed differences between cluster 1 and cluster 2: the 5CRT stopwatch duration was significantly short in cluster 1 participants compared to cluster 2 participants (mean 9.84, SD 1.42 seconds vs mean 12.60, SD 2.48 seconds, Cohen *d*=–1.311; *P*<.001). Habitual gait speed, assessed via 4mGS, was faster in cluster 1 compared to cluster 2 (mean 1.68, SD 0.19 seconds vs mean 1.38, SD 0.20 seconds, Cohen *d*=1.495; *P*<.001 and mean 2.42, SD 0.28 m/s vs mean 2.95, SD 0.44 m/s, Cohen *d*=–1.411; *P*<.001). 6minWT distance remained higher in cluster 1 (mean 481.35, SD 65.99 m vs mean 411.74, SD 60.08 m, Cohen *d*=1.111; *P*<.001). Balance measures showed a more nuanced pattern. Semi-tandem stance remained near ceiling in both clusters, with only a nonsignificant trend toward a higher score in cluster 1 (mean 1, SD 0 vs mean 0.96, SD 0.20; *P*=.09). In contrast, tandem stance revealed a significant advantage for cluster 1 participants (mean 1.86, SD 0.39 vs mean 1.62, SD 0.65; *P*=.008).

CMJ height and CMJ power were not assessed at t2.

Figure 5 shows a radar plot of Cohen *d* effect sizes of each assessment and each variable at baseline (t0), 6-month (t1), and 24-month (t2) follow-up. The effect sizes are plotted relative to a central zero line, representing no difference between cluster 1 and cluster 2. Concentric rings correspond to conventional interpretation thresholds for negligible (0.00-0.19), small (0.20-0.49), moderate (0.50-0.79), and large (≥ 0.8) effects.

XSL•FO
RenderX

Figure 5. The radar plot shows Cohen *d* effect sizes for various assessments (age, BMI, Handgrip Strength test, Stair Climb Power Test, Timed "Up and Go" test, Short Physical Performance Battery, Five Times Chair Rise Test, 4-meter Gait Speed, 6-minute Walking Test, Countermovement Jump height and Countermovement Jump power) at three time points (t0=baseline, t1=6 months, and t2=24 months) in relation to the zero line (red). Concentric rings represent thresholds for negligible, small, moderate, and large effects. 4mGS: 4-meter Gait Speed; 5CRT: Five Times Chair Rise Test; 6minWT: 6-minute Walking Test; CMJ: Countermovement Jump; HGST: Handgrip Strength test; SCPT: Stair Climb Power Test; SPPB: Short Physical Performance Battery; TUG: Timed "Up & Go".



Across all 3 time points, age showed small negative effects, indicating that participants in cluster 1 were slightly younger. BMI consistently showed moderate negative effects, reflecting a higher BMI in cluster 2 at all time points. In contrast, functional and performance measures clearly fall outside the negligible range and generally favor cluster 1. At t0, the effect sizes were small for HGS (Cohen *d*=0.447), SCPT (Cohen *d*=0.440), SPPB (Cohen *d*=0.429), and CMJ height and power (Cohen *d*=0.445, Cohen *d*=0.793). Mobility and gait-related outcomes showed large between-cluster differences at baseline: TUG (Cohen *d*=-1.047), 4mGS in seconds (Cohen *d*=–1.167), 4mGS in m/s (Cohen *d*=1.190), and 6minWT (Cohen *d*=1.141). There was also a moderate-to-large effect for 5CRT duration (Cohen *d*=–0.581). These effects indicate that cluster 1 had higher strength, mobility, gait speed, endurance, and jump performance than cluster 2 at baseline.

XSL•FO
RenderX

After 6 months (t1), this pattern persisted and more pronounced for several measures. Effect sizes were small to moderate for HGS (Cohen *d*=0.497), SCPT (Cohen *d*=0.513), and CMJ power (Cohen *d*=0.383). They increased to moderate for the SPPB (Cohen *d*=0.696). CMJ height showed a large effect size (Cohen *d*=0.926), which further indicates higher explosive lower-limb power in cluster 1. Mobility and gait outcomes again exhibited large effects: TUG (Cohen *d*=–1.024), 4mGS in seconds (Cohen *d*=–1.008), 4mGS in m/s (Cohen *d*=1.046), 6minWT (Cohen *d*=1.159), and 5CRT duration (Cohen *d*=–0.852). Overall, the radar plot at t1 shows a widening separation between both clusters for key functional measures, particularly for mobility, gait speed, endurance, and CMJ height.

After 24 months (t2), the divergence in performance between cluster 1 and cluster 2 is even more pronounced for several outcomes. HGS (Cohen *d*=0.471) remained in the small-to-moderate range, while SCPT showed a moderate effect (Cohen *d*=0.659). SPPB (Cohen *d*=1.219), TUG (Cohen *d*=–1.487), 5CRT duration (Cohen *d*=–1.311), 4mGS in seconds (Cohen *d*=1.495), 4mGS in m/s (Cohen *d*=–1.411), and 6minWT (Cohen *d*=1.111) demonstrated large effect sizes. This indicates substantial, clinically meaningful differences in lower extremity performance, balance, gait speed, and endurance favoring cluster 1. CMJ height and CMJ power were not assessed at t2, as shown by the missing values on the radar plot at t2.

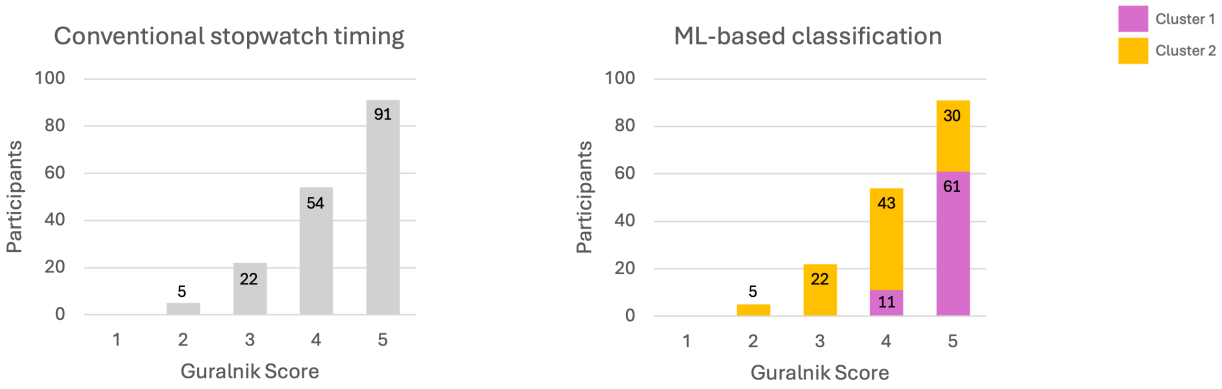
In summary, cluster 1 demonstrated higher functional performance at all 3 time points (t0, t1, and t2). Effect sizes indicated progressively larger differences compared to cluster 2. The largest differences were observed in mobility- and gait-related outcomes such as TUG, 4mGS, 6minWT, SPPB, and 5CRT. These outcomes showed the highest effect sizes and the greatest increases from baseline (t0) to the 24-month follow-up assessment (t2).

Added Value of ML-Based Classification Compared to Conventional Stopwatch Timing

To evaluate whether ML-based 5CRT movement pattern classification provides clinically meaningful differentiation of functional status beyond conventional stopwatch-based 5CRT timing, Figure 6 shows the distribution of participants classified by conventional, stopwatch-derived 5CRT scoring (left panel) and ML-based DTW cluster assignment (right panel). Using stopwatch-based thresholds according to Guralnik et al [8] (Table 2), most participants fell into the “good” performance category (3 - 4 points), while only a minority were in the impaired range (≤2 points). This suggests homogeneity and relatively preserved function in the cohort when interpreted by conventional timing alone.

However, the ML-based classification revealed substantial heterogeneity within the same stopwatch categories. While all cluster 1 participants were in the higher score ranges, cluster 2 participants spanned all performance levels, including many who scored 3 - 4 points by stopwatch criteria (Figure 6, right). Notably, these cluster 2 individuals exhibited elevated stabilization forces, slower mobility (TUG), and reduced endurance (6minWT), indicating clinically relevant impairments not reflected in their stopwatch score.

Figure 6. Comparison of conventional stopwatch-based Five Times Chair Rise Test scoring (left) and machine learning–based movement classification (right). While the conventional stopwatch-based Five Times Chair Rise Test scoring condensed most participants into high-performance categories (3-4 points), the machine learning–based clustering revealed hidden heterogeneity, separating participants into 2 clusters with clearly distinct functional profiles.



Discussion

Principal Findings

This study highlights the value of advanced, ML-based biomechanical analysis, specifically, DTW combined with k-means clustering, in identifying functional performance profiles among community-dwelling older adults performing the 5CRT. The results are discussed in the order of the predefined study aims.

Classification of 5CRT Movement Patterns Using ML-Based Clustering

The first study aim was to classify 5CRT movement patterns using DTW and k-means clustering, using the workflow established by Meyer et al [10] to ensure methodological comparability. The findings confirm that ML-based pattern recognition can differentiate between clusters of chair rise execution. Consistent with Meyer et al [10], 2 distinct movement clusters were identified at the 24-month follow-up assessment (t2). All participants lived in the community, were free from major acute morbidities, had to be able to climb 10 flights of

XSL•FO
RenderX

stairs without difficulty, and complete a TUG in less than 20 seconds. Despite this high baseline functional level, the ML-based pattern recognition detected differences in the shape and temporal dynamics based on 5CRT execution. Cluster 1 was characterized by faster mean 5CRT execution time of 9.84 (SD 1.42) seconds (Table 6) and lower stabilization forces (Figures 3 and 4). Cluster 2 exhibited 5CRT execution with a mean duration of 12.60 (SD 2.48) seconds (Table 6) and increased vertical GRFs during stabilization (Figures 3 and 4).

Although clustering was performed at t2, the functional differences between cluster 1 and cluster 2 participants were already evident at baseline and generally widened over the 24-month period (Tables 5-7 and Figure 5). This pattern suggests that the 5CRT movement patterns captured by the DTW-based approach may reflect individual differences in movement execution rather than purely transient variability. However, because cluster assignment was based on t2 data, these findings should be interpreted as descriptive stratification rather than evidence of distinct baseline trajectories or causal mechanisms. The identified clusters may represent a functional performance profile that is stable over time, but this remains hypothesis-generating and requires prospective validation.

Importantly, this multidimensional approach revealed functional heterogeneity within a population of community-dwelling older adults who were otherwise classified as healthy based on conventional criteria. Stopwatch timing produced a concentrated score range, with most participants falling into a high-functioning category (Figure 6), whereas DTW-based clustering identified subgroups with differing levels of mobility, strength, and postural control.

To sum up, the DTW-based clustering approach is a useful tool for identifying movement patterns in the 5CRT. However, the results may be influenced by preprocessing such as segmentation, resampling, and normalization. These steps can influence the input structure, thereby impacting the subsequent cluster solution. Additionally, different cluster sizes may lead to frequent movement patterns dominating the analysis, thereby underrepresenting less frequent patterns. Recent clinical movement studies [21] have demonstrated that data distribution and balancing decisions can substantially influence model-derived classifications, especially in small or heterogeneous datasets. Consequently, these findings should be interpreted as exploratory, and future work should assess the robustness of the results across a range of alternative preprocessing and sampling strategies.

Association Between Movement Pattern Clusters and Functional Performance Indicators

The second study aim focused on quantifying the associations between the ML-based movement clusters and established clinical outcomes. Across the baseline (t0), 6-month (t1), and 24-month (t2) follow-up assessments, participants in cluster 1 consistently outperformed those in cluster 2 on key performance indicators such as lower extremity strength (HGS), leg power (SCPT), mobility and gait (TUG and 4mGS), endurance (6minWT), global functional status (SPPB), and explosive capacity (CMJ power and height; Tables 5-7). These differences between both clusters were already present at baseline and generally increased over time, indicating diverging functional trajectories between cluster 1 and cluster 2.

The separation between cluster 1 and cluster 2 participants was most evident at t2, where effect sizes for 4mGS, TUG, and 6minWT reached the moderate-to-large range, pointing to disparities in mobility and endurance despite the cohort's initially high functional status. The pronounced separation at t2 is expected, as the clusters were assigned exclusively on t2 data using DTW-based k-means clustering, resulting in an inherent maximization of intercluster differences at this time point and strengthening the correlation between ML-based movement patterns and clinically meaningful functional performances.

In parallel, BMI and body weight were consistently higher in cluster 2 at all time points, with differences becoming particularly evident by t2. This pattern suggests that the ML-based movement clusters do not merely reflect transient performance fluctuations but align with broader anthropometric risk profiles that are known to coevolve with mobility decline and incident disability in older adults [22].

From a mechanistic point of view, these characteristics are plausible, but several interpretations remain possible. The 5CRT combines lower-limb power, postural control, and interlimb coordination [8,9]. The temporal dynamics of movement, as captured by DTW-based clustering, may encode compensatory strategies [10]. These strategies could include prolonged transition phases, altered sequencing, or pace variability that are not fully reflected by total task duration.

At the same time, the consistent superiority of cluster 1 across all functional domains suggests that the clusters may primarily reflect overall functional capacity or a latent fitness axis, rather than fundamentally distinct biomechanical strategies. In this view, DTW-based clustering may provide a more granular stratification of performance levels rather than identifying entirely novel movement phenotypes. This interpretation is consistent with the observation that cluster assignment is associated with stopwatch-based 5CRT duration.

The difference in BMI and body weight between cluster 1 and cluster 2 further supports a link between movement pattern quality, body composition, and functional aging. A higher BMI has been associated with adverse changes in body composition, reduced muscle quality, and increased risk of mobility disability and mortality [22].

In summary, the longitudinal association between ML-defined movement clusters and a broad spectrum of functional indicators suggests that temporal characteristics of chair-rise performance may be related to clinically relevant differences in functional aging. These results suggest that movement quality, as quantified by ML, may provide useful information about neuromuscular reserve and emerging frailty in otherwise high-functioning older adults.

Added Value of ML-Based Classification Compared to Conventional Stopwatch Timing

The third study aim was to evaluate the added clinical value offered by ML-based movement pattern classification beyond

XSL•FO
RenderX

conventional stopwatch-based timings. Distributional analysis revealed that stopwatch timing produced a more concentrated score range, with most participants falling within the high-functioning category (Figure 6). In contrast, DTW-based clustering revealed significant heterogeneity (Figure 6): cluster 2 participants displayed subclinical impairments in mobility, strength, and postural control despite achieving "good" stopwatch scores (Tables 5-7 and Figure 6). These discrepancies may highlight the value of sensor-based analytical methods in identifying latent deficits, particularly those related to compensatory strategies and stabilization, thus enabling more granular risk assessment and targeted early intervention.

From a clinical perspective, this approach may facilitate a transition toward personalized movement phenotyping for risk stratification, prognosis, and rehabilitation planning in older adults. Integrating ML-based biomechanical metrics with comprehensive geriatric assessments provides a more nuanced understanding of the aging process and functional decline trajectories that are not captured by traditional methods solely.

Our findings may support further development of precision geriatric assessment, in which multidimensional data and ML could contribute to the early identification of individuals at risk. This approach may help refine intervention strategies and support monitoring of mobility and independence in aging populations.

However, the incremental predictive value of cluster assignment beyond conventional 5CRT duration was not formally tested in this study. While the patterns suggest that sensor-based analytics may provide additional granularity, particularly for detecting subclinical deficits related to compensatory strategies and stabilization, this remains suggestive rather than proven. Formal analyses of incremental predictive value (eg, regression models with and without cluster assignment) are needed to establish whether ML-based clustering offers clinically meaningful added value beyond simple timing.

Limitations

Despite the strengths of a large cohort and robust biomechanical assessments, several limitations warrant consideration. The study sample consisted of relatively healthy community-dwelling older adults recruited from a specific geographical region in Germany, which may restrict the generalizability of our findings to more diverse or frailer populations. In addition, attrition over the 24-month period (including withdrawals and missing data) reduced the analytical sample size, potentially introducing bias toward healthier individuals who remained in the study.

Clusters were defined exclusively on the final time point (t2) data using k-means with DTW, which is sensitive to initial conditions, preprocessing choices, and cluster size imbalance. The pronounced separation observed at t2 is expected and should not be interpreted as independent evidence of longitudinal divergence. Instead, these discrepancies likely mirror the clustering procedure itself, which aimed to optimize inter-group separation at t2. Consequently, the observed variations at t0 and t1 may be indicative of post hoc stratification rather than genuinely distinct baseline trajectories. Future studies should use baseline clustering or temporal clustering approaches to evaluate prospective subgroup development. In addition, bootstrapping, stability analyses, or cluster consistency checks across time points as well as testing reproducibility in an independent cohort should be addressed in future work.

Furthermore, the strong alignment of cluster assignment with established functional variables raises the possibility that the clusters primarily reflect overall functional capacity rather than distinct movement strategies. In this study, it was not formally tested whether cluster assignment provides incremental predictive value beyond conventional metrics. These findings should therefore be interpreted as exploratory.

A limitation of this study is that the clustering approach did not explicitly model intraindividual changes across the 5 chair rises such as fatigue-, adaptation-, or motor control–related dynamics. In the data reported by Meyer et al [10], the first chair rise was generally consistent with the cluster assigned to the remaining repetitions, but subtle repetition-dependent effects may still not have been fully captured. Future studies should therefore consider repetition-aware or longitudinal approaches to investigate intraindividual changes in greater detail.

Although the use of a height-adjustable chair at t2 improved test standardization, slight deviations in chair height or leg position might still have influenced performance measures. Furthermore, the current setup using a force plate and depth camera offers high precision but may have limited scalability in everyday clinical environments. Wearable sensor technologies could serve as feasible alternatives for future large-scale or home-based assessments.

Finally, while ML-based clustering and DTW techniques enhanced sensitivity to subtle movement variations, broader validation across different sensor modalities, populations, and settings will be necessary to confirm their clinical utility and reproducibility. Future studies should also address psychosocial factors, environmental influences, and acceptability of such technology-driven assessments in routine geriatric care.

Claims regarding added clinical value should be considered preliminary.

Conclusions

This study demonstrates that ML-based analysis of dynamic chair rise movements using DTW and k-means clustering successfully identifies subgroups of older adults that cannot be discerned from conventional stopwatch timing metrics alone. These clusters exhibit significant differences in physical function, frailty, and mobility, as confirmed by comprehensive clinical assessments. The findings demonstrate that temporal characteristics of chair-rise performance provide information beyond total test duration. This study highlights the potential of advanced biomechanical analytics to improve geriatric screening and support the planning of personalized interventions. Further validation in larger, more diverse cohorts, and the integration of sensor-based assessments into routine practice, could optimize functional assessment and promote healthy aging at a population level.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank all participants and the entire project team for their dedication to collecting all AEQUIPA-Versa data.

Generative AI Use

Generative AI (GenAI) was used to support language editing and improve clarity in the manuscript. Specifically, the ChatAI service provided by Academic Cloud (Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen) was used with the Meta Llama 3.1 large language model. The manuscript is an original work of the authors, who reviewed and edited all AI-assisted content and take full responsibility for the final version.

Funding

The AEQUIPA project (Physical Activity and Health Equity: Primary Prevention for Healthy Aging) was funded by the Federal Ministry of Education and Research (grant number: 01EL1422D). Open access funding was enabled and organized by Projekt DEAL.

Data Availability

The datasets used and analyzed in this study are available from the authors on reasonable request.

Authors' Contributions

Conceptualization: AS, OM, RD
Data curation: AS, OM, RD
Formal analysis: AS, OM, RD
Funding acquisition: AH, RD
Investigation: AS, OM, RD
Methodology: AS, OM, RD
Project administration: AH, RD
Resources: AH, RD
Software: AS, OM, RD
Supervision: RD
Validation: AS, OM, RD
Visualization: AS, OM
Writing – original draft: AS
Writing – review & editing: AS, OM, VQ, AH, RD

Conflicts of Interest

None declared.

References

1. UN decade of healthy ageing: plan of action 2021-2030. : World Health Organization; 2020 URL: https://cdn.who.int/media/docs/default-source/decade-of-healthy-ageing/decade-proposal-final-apr2020-en.pdf [accessed 2025-08-12]
2. Rudnicka E, Napierała P, Podfigurna A, Męczekalski B, Smolarczyk R, Grymowicz M. The World Health Organization (WHO) approach to healthy ageing. Maturitas 2020 Sep;139:6-11. [doi: 10.1016/j.maturitas.2020.05.018] [Medline: 32747042]
3. Healthy ageing and functional ability. World Health Organization. URL: https://www.who.int/news-room/questions-and-answers/item/healthy-ageing-and-functional-ability [accessed 2025-08-19]
4. van der Kruk E, Silverman AK, Reilly P, Bull AMJ. Compensation due to age-related decline in sit-to-stand and sit-to-walk. J Biomech 2021 Jun 9;122:110411. [doi: 10.1016/j.jbiomech.2021.110411] [Medline: 33915476]
5. Michel JP, Sadana R. "Healthy Aging" concepts and measures. J Am Med Dir Assoc 2017 Jun 1;18(6):460-464. [doi: 10.1016/j.jamda.2017.03.008] [Medline: 28479271]
6. Fallon CK, Karlawish J. Is the WHO definition of health aging well? Frameworks for "Health" after three score and ten. Am J Public Health 2019 Aug;109(8):1104-1106. [doi: 10.2105/AJPH.2019.305177] [Medline: 31268759]
7. Beard JR, Officer A, de Carvalho IA, et al. The world report on ageing and health: a policy framework for healthy ageing. The Lancet 2016 May;387(10033):2145-2154. [doi: 10.1016/S0140-6736(15)00516-4]
8. Guralnik JM, Ferrucci L, Pieper CF, et al. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci 2000 Apr;55(4):M221-M231. [doi: 10.1093/gerona/55.4.m221] [Medline: 10811152]

XSL•FO
RenderX

9. Hardy R, Cooper R, Shah I, Harridge S, Guralnik J, Kuh D. Is chair rise performance a useful measure of leg power? Aging Clin Exp Res 2010;22(5-6):412-418. [doi: 10.1007/BF03324942] [Medline: 21422795]
10. Meyer O, Diekmann R, Hellmers S, Hein A, Schumacher A. Uncovering hidden insights in the chair rise performance of older adults using dynamic time warping and K-means clustering. Sci Rep 2025 Mar 5;15(1):7654. [doi: 10.1038/s41598-025-91015-x] [Medline: 40038392]
11. Hellmers S, Fudickar S, Lau S, et al. Measurement of the chair rise performance of older people based on force plates and IMUs. Sensors (Basel) 2019 Mar 19;19(6):1370. [doi: 10.3390/s19061370] [Medline: 30893819]
12. Diekmann R, Hellmers S, Elgert L, et al. Minimizing comprehensive geriatric assessment to identify deterioration of physical performance in a healthy community-dwelling older cohort: longitudinal data of the AEQUIPA Versa study. Aging Clin Exp Res 2021 Mar;33(3):563-572. [doi: 10.1007/s40520-020-01562-8] [Medline: 32358730]
13. Diekmann R, Hellmers S, Lau S, et al. Are vertical jumps able to predict 24-month follow-up functional geriatric assessment in a healthy community-dwelling older cohort? Aging Clin Exp Res 2022 Nov;34(11):2769-2778. [doi: 10.1007/s40520-022-02230-9] [Medline: 36053442]
14. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
15. Bean JF, Kiely DK, LaRose S, Alian J, Frontera WR. Is stair climb power a clinically relevant measure of leg power impairments in at-risk older adults? Arch Phys Med Rehabil 2007 May;88(5):604-609. [doi: 10.1016/j.apmr.2007.02.004] [Medline: 17466729]
16. Python reference manual. : Department of Computer Science [CS], Centrum Wiskunde & Informatica; 1995 URL: https://ir.cwi.nl/pub/5008 [accessed 2026-10-22]
17. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987 Nov;20:53-65. [doi: 10.1016/0377-0427(87)90125-7]
18. Schenkman M, Berger RA, Riley PO, Mann RW, Hodge WA. Whole-body movements during rising to standing from sitting. Phys Ther 1990 Oct;70(10):638-648. [doi: 10.1093/ptj/70.10.638] [Medline: 2217543]
19. Davies DL, Bouldin DW. A cluster separation measure. IEEE Trans Pattern Anal Mach Intell 1979 Feb;1(2):224-227. [doi: 10.1109/TPAMI.1979.4766909] [Medline: 21868852]
20. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
21. Trabassi D, Castiglia SF, Bini F, et al. Optimizing rare disease gait classification through data balancing and generative AI: insights from hereditary cerebellar ataxia. Sensors (Basel) 2024 Jun 3;24(11):3613. [doi: 10.3390/s24113613] [Medline: 38894404]
22. Murphy RA, Patel KV, Kritchevsky SB, et al. Weight change, body composition, and risk of mobility disability and mortality in older adults: a population-based cohort study. J Am Geriatr Soc 2014 Aug;62(8):1476-1483. [doi: 10.1111/jgs.12954] [Medline: 25039391]

Abbreviations

4mGS: 4-meter Gait Speed
5CRT : Five Times Chair Rise Test
6minWT: 6-minute Walking Test
BT: Balance Test
CMJ : Countermovement Jump
COPD: chronic obstructive pulmonary disease
DBA : DTW Barycenter Averaging
DTW: dynamic time warping
GRF : ground reaction force
HGS: Handgrip Strength
ML: machine learning
SCPT : Stair Climb Power Test
SPPB: Short Physical Performance Battery
TUG: Timed "Up & Go"

XSL•FO
RenderX

Edited by C Lee, J Wang; submitted 03.Mar.2026; peer-reviewed by D Trabassi, D Rivera; revised version received 05.Jul.2026; accepted 21.Sep.2026; published 09.Nov.2026.

Please cite as:
Schumacher A, Meyer O, Quinten V, Hein A, Diekmann R
Machine Learning–Based Biomechanical and Clinical Assessment of the Five Times Chair Rise Test in Older Adults: Sensor-Based Functional Aging Study
JMIR Aging 2026;9:e94597
URL: https://aging.jmir.org/2026/1/e94597
doi:10.2196/94597

© Anna Schumacher, Ole Meyer, Vincent Quinten, Andreas Hein, Rebecca Diekmann. Originally published in JMIR Aging (https://aging.jmir.org), 9.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multimodal Dementia Prediction With Large Language Models: Cross-Attention Over Text, Audio, and Image

Felix Agbavor[1], PhD; Hualou Liang[1,2,3,4], PhD

[1]School of Biomedical Engineering and Science, Drexel University, Philadelphia, PA, United States

[2]Division of Artificial Intelligence and the Humanities, The Hong Kong Polytechnic University, HHB717, 7/F, 8 Hung Lok Road, Hung Hom, Kowloon, China (Hong Kong)

[3]Department of Language Science and Technology, The Hong Kong Polytechnic University, Kowloon, China (Hong Kong)

[4]Departments of Data Science and Artificial Intelligence, The Hong Kong Polytechnic University, Kowloon, China (Hong Kong)

Corresponding Author:
Hualou Liang, PhD
School of Biomedical Engineering and Science, Drexel University, Philadelphia, PA, United States

Abstract

Background: Alzheimer disease (AD) is a leading cause of dementia, and there is growing interest in scalable approaches for early screening using speech-based tasks. While prior work has demonstrated promising results using either transcript-based language features or acoustic cues, most approaches remain unimodal or rely on simple fusion strategies that do not explicitly consider interactions across modalities.

Objective: In this study, we propose an attention-based trimodal fusion framework that integrates text, audio, and image representations of the Cookie Theft picture, which serves as the shared visual stimulus in the picture-description task.

Methods: Our method uses a new bidirectional cross-attention mechanism to achieve a unified multimodal embedding for downstream tasks. We evaluate the approach on 2 tasks: AD detection by classifying whether the participant has AD or not, and AD severity assessment by predicting Mini-Mental Status Examination cognitive scores.

Results: On the AD detection task, trimodal fusion achieves the best overall performance (F_1-score=0.8667, area under the receiver operating characteristic curve=0.9032), outperforming unimodal baselines, bimodal fusion, and conventional early or late fusion methods. For AD severity assessment, the proposed multimodal representation reduces prediction error of root mean squared error to about 4.20, improving over both unimodal and bimodal fusion settings. We further perform the ablation analysis to show that bidirectional cross-attention consistently outperforms conventional unidirectional cross-attention.

Conclusions: These results demonstrate that attention-based multimodal fusion can enhance dementia prediction from picture-description responses and provide a strong foundation for developing multimodal cognitive screening pipelines.

(JMIR Aging 2026;9:e93279) doi:10.2196/93279

KEYWORDS

Alzheimer disease; multimodal fusion; speech analysis; cognitive assessment; cross-attention

Introduction

Alzheimer disease (AD) is a progressive neurodegenerative disorder and a leading cause of dementia worldwide, with a growing public health impact due to aging populations [1,2]. Early screening and continuous monitoring are critical for timely intervention, planning, and patient support [3]. However, conventional clinical assessments remain resource-intensive, often requiring specialized expertise, controlled administration environments, expensive clinical facilities, and repeated in-person visits [4,5]. These limitations motivate scalable, accessible, noninvasive, and low-burden approaches for cognitive screening that can be deployed in realistic settings.

Spontaneous speech has emerged as a particularly promising biomarker for dementia because it reflects a wide range of cognitive-linguistic processes, including lexical retrieval, syntactic organization, semantic coherence, and executive control [6-8]. A substantial body of work demonstrates that spontaneous speech contains clinically meaningful signals for AD screening and diagnosis. However, most existing studies remain unimodal, relying solely on either textual features derived from transcripts or acoustic features extracted from audio, thereby neglecting complementary information that is critical for robust inference. More recent approaches have explored bimodal models (eg, audio + text), which generally outperform unimodal baselines by aligning linguistic and acoustic cues [8-10]. Despite these advances, bimodal analyses primarily focus on local speech characteristics such as pauses and lexical diversity and often fail to capture global discourse-level phenomena, including disrupted topic

XSL•FO
RenderX

maintenance and cross-modal inconsistencies between spoken narratives and visual image.

To address these limitations, we introduce image information into AD prediction by jointly modeling text, audio, and image representations in the Cookie Theft picture-description task. Because the Cookie Theft picture is identical across all participants, the visual modality is not intended to encode participant-specific variability; instead, it provides shared visual context that grounds the spoken narrative during multimodal fusion. This visual context helps relate the narrative to the scene being described and serves as a semantic anchor for fusion. Each modality contributes different information. Text captures lexical choice, semantic content, coherence, and informativeness of the spoken description. Audio contributes complementary paralinguistic cues, including pauses, hesitation, rhythm, fluency, and other prosodic features not fully preserved in transcripts. The Cookie Theft picture provides shared visual context that grounds the narrative and helps assess how well the spoken description aligns with the scene.

Multimodal integration is commonly approached through 3 fusion strategies: early (data-level), intermediate (joint), and late (decision-level) fusion. Early fusion projects features from different modalities into a single shared space, but it often fails to capture higher-order cross-modal interactions. Late fusion, by contrast, aggregates the outputs of modality-specific models, preserving individual modality strengths while overlooking deeper interdependencies among them [11-13]. Intermediate, or joint, fusion offers a principled compromise by explicitly modeling interactions across modalities during representation learning. In this context, Transformer-based cross-attention has emerged as a key mechanism for capturing fine-grained and long-range intermodal relationships [14]. Intuitively, cross-attention allows one modality to look at another and selectively use the parts that are most relevant for the task. For example, a text representation can attend to complementary information in the audio or visual representation, so the fused embedding reflects not only what was said, but also how it was spoken and how well it aligns with the Cookie Theft picture context.

Cross-attention has been widely adopted in language-vision representation learning [8,15], audio-text integration [11], and more recently in dementia prediction [16]. However, existing approaches typically rely on unidirectional cross-attention, in which one modality queries another, resulting in asymmetric information flow. To overcome this limitation, we propose a bidirectional cross-attention mechanism that enables symmetric and iterative information exchange across modalities. In this design, each modality both attends to and is updated by the other, allowing representations to be refined in both directions. This iterative bidirectional interaction facilitates richer cross-modal integration [8,15], promotes balanced modality contributions, and mitigates the risk of modality dominance. Building on this perspective, our trimodal fusion architecture leverages bidirectional cross-attention to produce unified embeddings that capture both complementary evidence and cross-modal agreement, rather than relying on fixed or heuristic fusion rules. Simply put, rather than allowing only one modality to query information from another, bidirectional cross-attention enables both modalities to interact and refine each other, resulting in a fused representation derived from 2-way information exchange rather than 1-way conditioning.

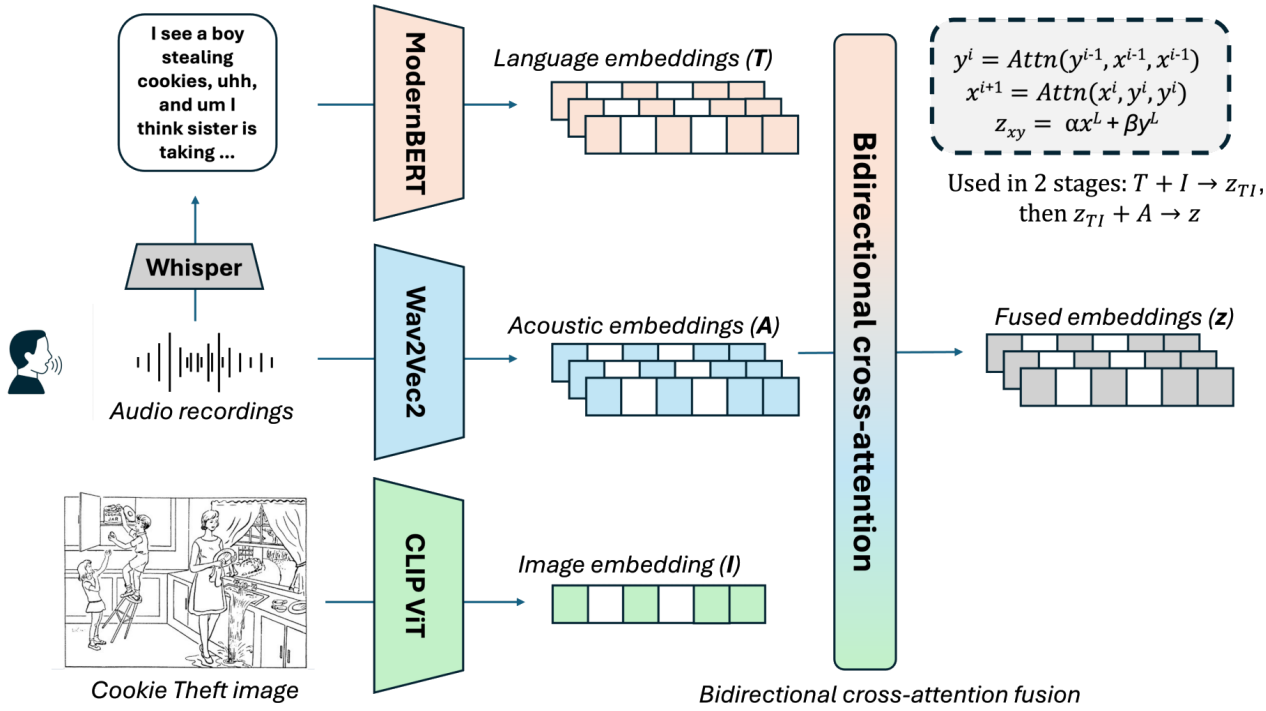
Taken together, we propose a trimodal embedding-level fusion architecture that integrates text, audio, and image through bidirectional cross-attention (Figure 1). The resulting fused embedding is used for downstream AD detection and cognitive score prediction. We evaluate this approach on the ADReSSo 2021 picture-description task against unimodal baselines, bimodal fusion models, and conventional early- and late-fusion strategies.

Our main contributions are as follows:

- We introduce a staged bidirectional cross-attention fusion framework for integrating text, audio, and image representations in the AD picture-description setting.
- On the ADReSSo 2021 picture-description task, we show that trimodal fusion improves over unimodal baselines and outperforms bimodal fusion for both AD detection and Mini-Mental Status Examination (MMSE) prediction.
- We show that the proposed fusion framework performs better than conventional fusion strategies, including early fusion and late fusion.
- We further show that bidirectional cross-attention yields more effective and stable performance than a standard 1-way cross-attention under the same task setting.

XSL•FO
RenderX

Figure 1. Proposed trimodal fusion architecture for dementia prediction. Pretrained 768-dimensional embeddings are extracted from text (ModernBERT), audio (Wav2Vec 2.0), and the Cookie Theft picture (contrastive language–image pretraining [CLIP] ViT-L/14). Fusion is performed in 2 stages for clarity. In stage 1, text and image are fused through a bidirectional bimodal cross-attention block to obtain an intermediate representation, *TI*. In stage 2, this intermediate *TI* representation is fused with audio through the same bidirectional fusion block to produce the final multimodal embedding, *z*, which is used for downstream classification and regression. The inset illustrates the generic bimodal fusion module used in each stage, where 2 modality embeddings are updated sequentially across layers by attending to one another, enabling reciprocal information exchange beyond standard 1-way cross-attention.



Methods

Dataset Description

The dataset used in this study is derived from the ADReSSo 2021 Challenge and consists of speech recordings from a picture-description task [8], where cognitively normal participants and individuals diagnosed with AD were asked to describe the Cookie Theft image from the Boston Diagnostic Aphasia Examination [17,18]. In total, the dataset contains 237 recordings, split into a 70/30 train–test partition with demographic balancing, resulting in 166 samples for training and 71 for testing. Within the training set, there are 87 AD and 79 non-AD (healthy control) recordings.

A key strength of ADReSSo is its focus on reducing common sources of bias in dementia screening benchmarks. The dataset was carefully curated to mitigate confounds such as repeated speech samples from the same individual, variation in recording quality, and imbalances in age and gender. Demographic matching was performed using a propensity-score procedure as described by Luz et al [8], yielding standardized mean differences below 0.001 for both age and gender, which helps ensure that downstream model performance reflects disease-related signal rather than demographic artifacts.

Dataset Preprocessing and Feature Representation

We preprocess the ADReSSo picture-description recordings to obtain aligned audio, text, and image representations, which serve as inputs to our multimodal fusion architecture. For the audio modality, we extract embeddings using Wav2Vec 2.0 [19], a self-supervised speech representation model trained to learn rich acoustic features directly from raw waveforms. Given an input recording, we obtain a sequence of frame-level hidden states and apply mean pooling over time to produce a fixed-dimensional 768-D audio embedding per sample.

For the text modality, we first transcribe each recording using OpenAI Whisper [20], which provides robust automatic speech recognition for spontaneous speech. To better preserve clinically relevant phenomena commonly observed in AD speech, we supply an initial decoding prompt (“Umm, Uhh, let me think like, hmm... Okay, here’s what I’m, like, thinking.") that encourages retention of fillers, disfluencies, repetitions, and incomplete phrases, rather than normalizing them away. We then encode the resulting transcripts using ModernBERT [21], a modern bidirectional Transformer encoder optimized for producing strong sentence-level representations. Similar to the audio pipeline, we derive a fixed-length 768-D text embedding for each transcript, providing a compact semantic representation that captures content, coherence, and lexical organization.

For the image modality, we represent the Cookie Theft stimulus using contrastive language–image pretraining (CLIP; ViT-L/14), a vision-language model trained via contrastive learning to align visual and textual representations in a shared embedding space. We extract the image encoder output and use a 768-D image embedding, which provides a high-level description of the visual scene structure and salient objects.

Attention-Based Multimodal Fusion

A central component of our fusion approach is bidirectional cross-attention, where each modality is updated by attending

XSL•FO
RenderX

to the others, allowing text, audio, and image embeddings to exchange information in both directions rather than being combined through simple concatenation. Unlike simple fusion methods (eg, concatenation), cross-attention allows the model to learn which parts of one representation are most relevant given the other representation, which is especially important in dementia screening where different modalities may carry complementary cues. Formally, we use the standard attention operation represented by Attention(Q,K,V)=softmax(QKTdk)V, where *Q* (queries) represents what we want to "look for," while *K* (keys) and *V* (values) represent the information we want to retrieve from.

Intuitively, bidirectional cross-attention can be understood as a 2-way exchange between modalities. In one direction, the text representation asks which parts of the audio or image are most relevant for interpreting what was said; in the other direction, the audio or image representation is also updated by attending back to the text. This allows the fused representation to reflect not only the content of the spoken description, but also how it was delivered and how well it aligns with the shared Cookie Theft picture context. In practice, the bidirectional design helps the model avoid treating one modality as fixed context for another and instead encourages reciprocal refinement of the modality embeddings across layers.

To combine information from text, image, and audio in a way that goes beyond simple concatenation, we propose an attention-based fusion module (Figure 1) that learns how modalities complement each other when predicting AD status and cognitive outcomes. The key idea is that each modality carries different but related information: text reflects linguistic content and coherence, audio captures acoustic patterns such as hesitation and fluency, and the image provides the visual context that grounds the picture-description narrative. Rather than assuming each modality contributes equally, our fusion approach learns to emphasize the most informative signals dynamically.

Our architecture performs multimodal fusion using bidirectional cross-attention, enabling each modality to iteratively incorporate information from the others. Rather than simply concatenating embeddings, the model allows the text representation to attend to complementary cues from the image and audio, while the non-text modalities also attend back to the evolving text representation. This bidirectional exchange supports richer alignment between what is said, how it is spoken, and the visual context of the picture-description stimulus. As a result, the fusion module produces a single multimodal embedding that reflects both modality-specific evidence and cross-modal dependencies, which is then used for downstream AD classification and its severity assessment.

For the implementation, we use a multilayer attention block that alternates between the 2 inputs, allowing one modality to "look at" and incorporate information from the other. Intuitively, this helps the model learn cross-modal interactions. For example, how acoustic hesitation patterns relate to lexical choice or how a visually grounded narrative differs between cognitively normal and impaired speakers. The final output is a single fused embedding per sample, which is passed to downstream classifiers or regressors for AD detection and its severity assessment.

Training Regimen

We train the proposed trimodal fusion model using the ADReSSo training split and further reserve 20% of the training data as a validation set for model selection and early stopping. This validation split is created randomly but reproducibly using a fixed seed of 42, ensuring consistent comparisons across experiments. To enforce reproducibility, the same seed is applied across Python, NumPy, and PyTorch (central processing unit and Compute Unified Device Architecture), with deterministic CUDA deep neural network behavior enabled. During training, samples are loaded in batches of 32. The fusion backbone uses a shared projection dimension of 768, with 4 attention heads and 3 cross-attention layer pairs per fusion stage, and models are trained for up to 50 epochs with a dropout rate of 0.1. To reduce scale mismatch across modalities, we apply L2 normalization to the input embeddings prior to fusion.

The model parameters are optimized using AdamW with a learning rate of 0.0001 and weight decay of 1×10^{-4}. For AD detection, we minimize cross-entropy loss, and for AD severity assessment, we use mean squared error loss. To stabilize learning, we use a ReduceLROnPlateau scheduler that monitors validation loss and reduces the learning rate by a factor of 0.5 after 2 consecutive epochs without improvement. We additionally apply gradient clipping with a maximum norm of 1.0. Early stopping is based on validation loss with a patience of 5 epochs, and the model weights from the epoch with the lowest validation loss are restored for final evaluation and embedding extraction.

AD Detection Task

The AD detection task is formulated as a binary classification problem, where the goal is to distinguish between participants with AD and cognitively normal (non-AD) participants using picture-description responses from the ADReSSo dataset. Because dementia-related signals can manifest across multiple modalities, we evaluate models using (1) audio-only representations, (2) text-only representations, and (3) multimodal fusion representations derived from combinations of text, audio, and image. Audio representations are extracted from Wav2Vec 2.0 embeddings, text representations are extracted from ModernBERT embeddings computed over Whisper transcripts, and image representations are extracted using CLIP ViT-L/14. We primarily use 3 standard machine-learning models for the AD classification task. These include support vector classifier (SVC), random forest (RF), and logistic regression (LR), all of which are implemented from the scikit-learn package [22-25].

AD Severity Assessment Task

In addition to AD classification, we evaluate our approach for AD severity assessment by predicting MMSE cognitive score, formulated as a regression problem. The objective is to predict a participant's cognitive score directly from their picture-description response, enabling a finer-grained estimate of cognitive status beyond binary diagnosis. Similar to the classification setting, we investigate the predictive use of multiple modalities by training models with audio-only,

XSL•FO
RenderX

text-only, and image-only representations, as well as multimodal fusion representations that combine complementary cues across modalities. Specifically, we use support vector regressor (SVR), random forest regressor (RFR), and ridge regression (Ridge) as our machine learning models [23-26]. For this task, we use the same feature representations described in the preprocessing pipeline: Wav2Vec 2.0 embeddings for audio, ModernBERT embeddings derived from Whisper transcripts for text, and CLIP ViT-L/14 embeddings for the image stimulus. We compare our proposed attention-based trimodal fusion model against unimodal baselines and bimodal fusion models.

Evaluation Metrics

We evaluate performance separately for the AD classification and AD severity assessment tasks. For the binary classification setting, we report accuracy, precision, recall, and F_1-score, providing a balanced view of overall correctness as well as performance on the positive AD class. Additionally, we also show the receiver operating characteristic (ROC) curve for modality comparisons and report the area under the ROC curve. For the regression task, we quantify prediction error using root mean squared error (RMSE), which measures the average magnitude of score prediction deviations. For all classical machine-learning baselines, optimal hyperparameters are selected using grid search with 5-fold cross-validation performed on the training split only. This ensures that model configuration is tuned in a statistically robust manner while keeping the held-out test set strictly reserved for final evaluation.

To better characterize uncertainty in model performance, we additionally report 95% CIs for the major classification and regression results using nonparametric bootstrap resampling of the held-out test set. For each metric, CIs were estimated from the empirical bootstrap distribution across 1000 repeated resamples. Note that the observed performance differences were not tested for statistical significance to avoid overreliance on arbitrary significance thresholds (eg, $P<.05$) and potential misinterpretations of P values.

Attention Variant Study (Bidirectional Cross-Attention Versus Cross-Attention Only)

To assess the impact of reciprocal information exchange during multimodal fusion, we compare 2 attention configurations under an otherwise identical fusion framework. The cross-attention-only, namely standard cross-attention, performs fusion using a single cross-attention direction, where 1 modality attends to another to incorporate complementary information. In contrast, the bidirectional cross-attention applies cross-attention in both directions, allowing each modality to be updated using information from the other in a sequential manner. Both variants use the same pretrained embeddings, projection dimensionality, optimization settings, and train or validation split, enabling a controlled comparison of the attention mechanisms.

Comparison with Conventional Fusion Methods

To assess whether the proposed attention-based multimodal fusion provides benefits beyond standard approaches, we compare it against 2 widely used conventional fusion strategies for the AD classification task: early fusion and late fusion. These baselines are commonly adopted in multimodal machine learning due to their simplicity, but they often struggle to capture complex interactions between modalities. For early fusion, we construct a single feature vector by concatenating the modality-specific embeddings (text, audio, and image) and train a classifier directly on the combined representation. While straightforward, early fusion does not explicitly model cross-modal relationships and instead relies on the downstream classifier to learn interactions implicitly from the concatenated embedding.

For late fusion, we train separate unimodal classifiers and combine their predictions at the decision level using majority voting. Because the visual stimulus is identical across all samples in the ADReSSo picture-description setting, an image-only classifier would not be meaningful for late fusion. Therefore, we report late-fusion performance using only the text and audio unimodal models. This comparison helps isolate the benefit of our attention-based fusion, which can jointly reason over modalities and learn cross-modal dependencies rather than combining independent unimodal decisions.

In addition to these conventional fusion strategies, we include a neural network baseline trained on the concatenated multimodal embeddings. The neural network architecture consisted of 2 hidden layers with 512 and 128 nodes, respectively, using the rectified linear unit activation function. This funnel-like bottleneck design by progressively stepping down from the high-dimensional input space to 512 and then 128 nodes was chosen to systematically compress the multimodal features. This encourages the network to learn robust, high-level representations while mitigating the risk of overfitting on the relatively small dataset [27]. The network was optimized using the Adam solver with an initial learning rate of 0.001. To further mitigate overfitting and ensure optimal generalization, we implemented early stopping. Specifically, 10% of the training data were set aside as a validation set, and training was terminated if the validation score did not improve by at least 1×10^{-4} for 10 consecutive epochs. This baseline provides a stronger learned comparison than linear or tree-based classifiers on early-fusion features, while still operating on the same underlying embedding inputs. In this way, we can assess whether any observed gains arise simply from using a more flexible classifier on concatenated features or from the proposed bidirectional fusion mechanism itself.

Ethical Considerations

The ADReSSo Challenge data used in this study are available through DementiaBank with approved credentials. The studies involving human participants were reviewed and approved by the DementiaBank consortium. All enrolled participants provided informed written consent to participate in this study. All data analyses in this work were conducted using deidentified data.

XSL•FO
RenderX

Results

AD Detection Task

Unimodal Baselines

In Table 1, we compare the performance of text and audio embeddings using the different classifiers. From the table, we can see that text provides stronger prediction for AD than the audio. With ModernBERT embeddings, SVC attains an F_1-score of 0.8334, outperforming LR which has an F_1-score of 0.7910 and RF which has an F_1-score of 0.7713. In general, audio (wav2vec2) is weaker across all the classifiers as the best audio-only result is achieved by RF with an F_1-score of 0.6736 and an accuracy of 0.7048. This performance gap between text and audio underscores that picture-description transcripts capture more discriminative information than the audio embeddings.

Table . Unimodal baselines on the ADReSSo 2021 unseen test set (n=71)[a].

Classifier	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
SVC^{b}_{text}	0.8448 (0.7606 - 0.9296)	0.8744 (0.7500 - 0.9714)	0.7999 (0.6571 - 0.9231)	0.8334 (0.7241 - 0.9231)
SVC_{audio}	0.7051 (0.5915 - 0.8028)	0.7069 (0.5484 - 0.8529)	0.6852 (0.5172 - 0.8294)	0.6928 (0.5588 - 0.8095)
LR^{c}_{text}	0.7876 (0.6901 - 0.8732)	0.7618 (0.6190 - 0.8919)	0.8275 (0.6969 - 0.9429)	0.7910 (0.6774 - 0.8889)
LR_{audio}	0.6617 (0.5493 - 0.7746)	0.6886 (0.5161 - 0.8519)	0.5712 (0.3953 - 0.7317)	0.6207 (0.4666 - 0.7500)
RF^{d}_{text}	0.8029 (0.7042 - 0.8873)	0.8886 (0.7500 - 1.0000)	0.6859 (0.5294 - 0.8333)	0.7713 (0.6429 - 0.8772)
RF_{audio}	0.7048 (0.5915 - 0.8169)	0.7338 (0.5714 - 0.8846)	0.6284 (0.4643 - 0.7858)	0.6736 (0.5283 - 0.7949)

[a]Performance of text-only (ModernBERT embeddings) and audio-only (wav2vec 2.0 embeddings) representations with 3 standard classifiers (support vector classifier, logistic regression, and random forest) is shown.

[b]SVC: support vector classifier.

[c]LR: logistic regression.

[d]RF: random forest.

Bimodal Fusion (2-Way)

Combining text and image embeddings yields consistently strong performance as shown in Table 2, reflecting the benefit of pairing transcript semantics with visual context from the Cookie Theft stimulus. Across classifiers, the results remain close to the text-only baselines, suggesting that the image modality provides incremental gain rather than a major boost in this setting. The top configuration is achieved by SVC, reaching an F_1-score of 0.8371 with an accuracy of 0.8437, slightly exceeding the text-only SVC F_1-score (0.8358). RF also performs competitively (F_1-score=0.8428), while LR attains an F_1-score of 0.8011. Overall, text continues to dominate prediction quality, while image representations appear to offer modest complementary signal.

Table . Text + image fusion for Alzheimer disease (AD) classification[a].

Classifier	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
SVC^{b}	0.8437 (0.7465 - 0.9296)	0.8509 (0.7241 - 0.9655)	0.8275 (0.6944 - 0.9429)	0.8371 (0.7333 - 0.9247)
LR^{c}	0.8011 (0.7042 - 0.8873)	0.7978 (0.6571 - 0.9231)	0.7987 (0.6571 - 0.9259)	0.7959 (0.6857 - 0.8919)
RF^{d}	0.8428 (0.7606 - 0.9296)	0.8731 (0.7500 - 0.9714)	0.7987 (0.6571 - 0.9286)	0.8321 (0.7241 - 0.9231)

[a]Performance of bimodal text + image representations on the ADReSSo 2021 unseen test set (n=71) using 3 classifiers (support vector classifier, logistic regression, and random forest) is shown. Text embeddings are extracted with ModernBERT, and image embeddings are extracted with contrastive language–image pretraining ViT-L/14.

[b]SVC: support vector classifier.

[c]LR: logistic regression.

[d]RF: random forest.

The text + audio pairing produces the best bimodal performance in Table 3 among all 2-way combinations. In particular, incorporating audio improves performance beyond transcript-only models, suggesting that features such as speech rhythm, hesitation patterns, and fluency irregularities provide additional discriminative value. The best overall results are obtained with SVC, which achieves an accuracy of 0.8584 and an F_1-score of 0.8503, outperforming both the text-only SVC model (F_1-score=0.8358) and the other bimodal settings. RF remains strong (F_1-score=0.8325), whereas LR performs slightly lower (F_1-score=0.7903). These results support the benefit of integrating acoustic information when strong text representations are available.

XSL•FO
RenderX

Table . Text + audio fusion for Alzheimer disease (AD) classification[a].

Classifier	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
SVC[b]	0.8584 (0.7746 - 0.9296)	0.8771 (0.7568 - 0.9722)	0.8286 (0.6944 - 0.9444)	0.8503 (0.7458 - 0.9315)
LR[c]	0.8016 (0.7042 - 0.8873)	0.8165 (0.6765 - 0.9394)	0.7704 (0.6249 - 0.9024)	0.7903 (0.6762 - 0.8889)
RF[d]	0.8440 (0.7606 - 0.9296)	0.8732 (0.7500 - 0.9714)	0.7993 (0.6571 - 0.9231)	0.8325 (0.7213 - 0.9206)

[a]Performance of bimodal text + audio representations on the ADReSSo 2021 unseen test set (n=71) using 3 classifiers (support vector classifier, logistic regression, and random forest) is shown. Text embeddings are extracted with ModernBERT, and audio embeddings are extracted with Wav2Vec 2.0.

[b]SVC: support vector classifier.

[c]LR: logistic regression.

[d]RF: random forest.

Table 4 shows results for audio + image fusion, a pairing that excludes transcript information and relies only on acoustic patterns and visual context. Performance is notably lower than text-based bimodal models, but it improves substantially over audio-only baselines, suggesting that image features contribute useful grounding signals when linguistic embeddings are not available. The top performance is achieved by SVC, with an accuracy of 0.7188 and an F_1-score of 0.6934, followed closely by RF (F_1-score=0.6830). LR produces a slightly weaker result (F_1-score=0.6530). Overall, these results reinforce that while audio + image alone can support AD prediction above chance, transcript-derived text embeddings remain critical for achieving top performance on picture-description based dementia screening.

Table . Audio + image fusion for Alzheimer disease (AD) classification[a].

Classifier	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
SVC[b]	0.7188 (0.6197 - 0.8169)	0.7427 (0.5806 - 0.8889)	0.6558 (0.4997 - 0.8065)	0.6934 (0.5574 - 0.8116)
LR[c]	0.6760 (0.5634 - 0.7746)	0.6869 (0.5185 - 0.8438)	0.6283 (0.4643 - 0.7857)	0.6530 (0.5098 - 0.7742)
RF[d]	0.7180 (0.6056 - 0.8169)	0.7582 (0.6000 - 0.9091)	0.6270 (0.4571 - 0.7812)	0.6830 (0.5385 - 0.8056)

[a]Performance of bimodal audio + image representations on the ADReSSo 2021 unseen test set (n=71) using 3 classifiers (support vector classifier, logistic regression, and random forest) is shown. Audio embeddings are extracted with Wav2Vec 2.0, and image embeddings are extracted with CLIP ViT-L/14.

[b]SVC: support vector classifier.

[c]LR: logistic regression.

[d]RF: random forest.

Multimodal Fusion (3-Way)

Using all 3 modalities (text, audio, and image) leads to the highest AD classification performance in our experiments. Among the 3 classifiers, SVC achieves the best overall results, with an accuracy of 0.8722 and an F_1-score of 0.8667. LR attains an F_1-score of 0.8542, while RF achieves an F_1-score of 0.8371. Overall, the 3-way fusion setting produces strong and stable performance across classifiers, with SVC providing the highest classification scores (Table 5).

Table . Multimodal fusion for Alzheimer disease (AD) classification[a].

Classifier	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score
SVC[b]	0.8722 (0.7887 - 0.9437)	0.8806 (0.7632 - 0.9730)	0.8564 (0.7317 - 0.9688)	0.8667 (0.7692 - 0.9474)
LR[c]	0.8582 (0.7746 - 0.9437)	0.8558 (0.7250 - 0.9677)	0.8563 (0.7273 - 0.9678)	0.8542 (0.7536 - 0.9367)
RF[d]	0.8437 (0.7465 - 0.9296)	0.8509 (0.7241 - 0.9655)	0.8275 (0.6944 - 0.9429)	0.8371 (0.7333 - 0.9247)

[a]Performance of text + audio + image fusion on the ADReSSo 2021 unseen test set (n=71) using 3 classifiers (support vector classifier, logistic regression, random forest). Text embeddings are extracted with ModernBERT, audio embeddings with Wav2Vec 2.0, and image embeddings with CLIP ViT-L/14.

[b]SVC: support vector classifier.

[c]LR: logistic regression.

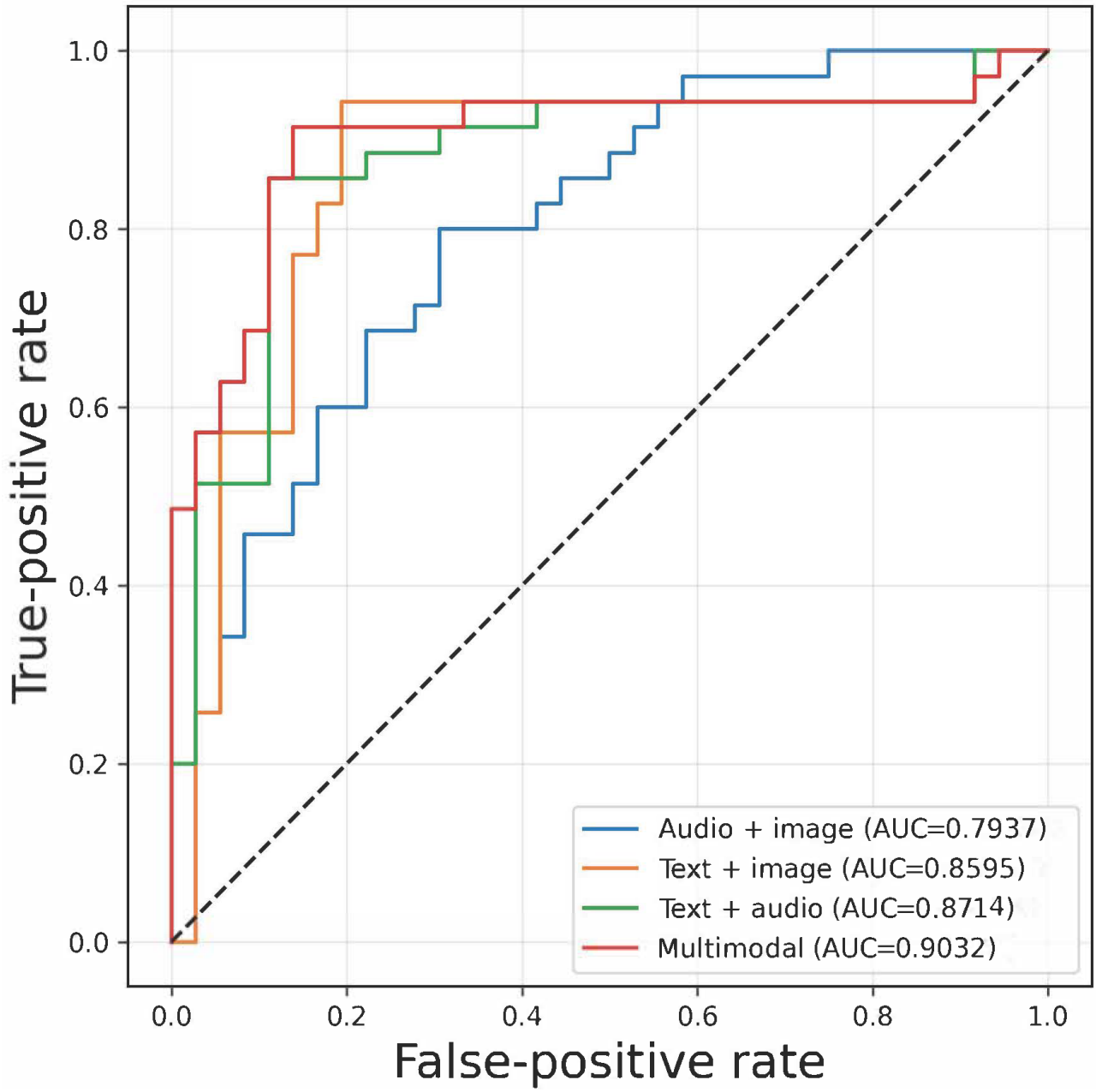
[d]RF: random forest.

XSL•FO
RenderX

ROC Curve Comparison Between Bimodal and Multimodal Fusion

Figure 2 compares ROC curves across bimodal fusion settings and the proposed trimodal multimodal fusion model for AD classification. Among the bimodal approaches, text + audio achieves the highest discrimination performance (area under the receiver operating characteristic curve [AUC]=0.8714), followed closely by text + image (AUC=0.8595), while audio + image performs notably lower (AUC=0.7937). The proposed multimodal fusion model yields the strong overall ROC performance with an AUC of 0.9032, indicating improved separability between participants with AD and without AD relative to all bimodal alternatives.

Figure 2. Receiver operating characteristic (ROC) comparison across bimodal and multimodal fusion (Alzheimer disease [AD] classification). ROC curves on the ADReSSo 2021 unseen test set (n=71) compare bimodal fusion models (audio + image, text + image, text + audio) against the proposed trimodal multimodal fusion (text + audio + image). The legend reports the area under the receiver operating characteristic curve (AUC) for each method.



Comparison to Other Fusion Methods

Table 6 compares the proposed multimodal fusion with conventional early fusion, late fusion, and a neural network baseline trained on concatenated embeddings. For SVC, both early fusion and late fusion achieve an accuracy of 0.8169, with F_1-scores of 0.7937 and 0.8000, respectively. The neural network baseline improves this comparison to an accuracy of 0.8311 and an F_1-score of 0.8099. In contrast, the proposed multimodal fusion achieves a higher accuracy of 0.8722 and a higher F_1-score of 0.8667. A similar pattern is observed across the other classifier families, indicating that the advantage of the proposed fusion is not limited to comparison against only simple conventional baselines.

XSL•FO
RenderX

Table . Comparison with conventional and learned fusion baselines for Alzheimer disease (AD) classification[a].

Fusion method	Accuracy	Precision	Recall	F_1-score
$SVC^{b}_{\text{early fusion}}$	0.8169	0.8929	0.7143	0.7937
$SVC_{\text{late fusion}}$	0.8169	0.8666	0.7428	0.8000
$SVC_{\text{multimodal fusion}}$	0.8722	0.8806	0.8564	0.8667
$LR^{c}_{\text{early fusion}}$	0.8028	0.8387	0.7429	0.7429
$LR_{\text{late fusion}}$	0.7887	0.8125	0.7428	0.7761
$LR_{\text{multimodal fusion}}$	0.8582	0.8558	0.8563	0.8542
$RF^{d}_{\text{early fusion}}$	0.8028	0.8889	0.6857	0.7742
$RF_{\text{late fusion}}$	0.7887	0.8333	0.7142	0.7692
Neural network	0.8311	0.8964	0.7427	0.8099
$RF_{\text{multimodal fusion}}$	0.8437	0.8509	0.8275	0.8371

[a]Performance comparison between early fusion (feature concatenation), late fusion (decision-level fusion of text and audio), a neural network baseline trained on concatenated embeddings, and the proposed attention-based multimodal fusion (text + audio + image) on the ADReSSo 2021 unseen test set (n=71) is shown.

[b]SVC: support vector classifier.

[c]LR: logistic regression.

[d]RF: random forest.

Bidirectional Cross-Attention Versus Cross-Attention Only

We evaluate 2 fusion variants to quantify the impact of the bidirectional cross attention. Across all 3 classifiers (Table 7), the bidirectional cross-attention configuration achieves stronger overall performance than cross-attention only. For SVC, cross-attention alone attains an F_1-score of 0.8571 with accuracy of 0.8732, while our bidirectional cross attention improves results to an F_1-score of 0.8667 and an accuracy of 0.8722. A similar trend is observed for LR. For RF, the cross-attention only baseline achieves an F_1-score of 0.8125 (accuracy=0.8309), and the bidirectional cross-attention variant improves to an F_1-score of 0.8371 (accuracy=0.8437). Overall, bidirectional cross-attention produces balanced AD classification performance metrics.

Table . Performance comparison between cross-attention–only fusion and the proposed bidirectional cross-attention fusion on the ADReSSo 2021 unseen test set (n=71) across 3 classifiers (support vector classifier [SVC], logistic regression [LR], random forest [RF]).

Fusion method	Accuracy	Precision	Recall	F_1-score
$SVC_{\text{cross\_attention}}$	0.8732	0.9642	0.7714	0.8571
$SVC_{\text{bidirectional}}$	0.8722	0.8806	0.8564	0.8667
$LR_{\text{cross\_attention}}$	0.8028	0.7837	0.8285	0.8055
$LR_{\text{bidirectional}}$	0.8582	0.8558	0.8563	0.8542
$RF_{\text{cross\_attention}}$	0.8309	0.8965	0.7428	0.8125
$RF_{\text{bidirectional}}$	0.8437	0.8509	0.8275	0.8371

AD Severity Assessment Task

Overview

We evaluated how well individual modalities support continuous MMSE score prediction and compare them to pairwise and their multimodal counterparts (Table 8). Across all 3 modalities for regression, text embeddings again outperform audio. With ModernBERT features, Ridge regression attains the lowest text-only error, followed closely by SVR and RFR. In contrast, the corresponding audio-only models with wav2vec2 embeddings are substantially less accurate.

XSL•FO
RenderX

Table . Mini-Mental Status Examination (MMSE) score prediction across unimodal, bimodal, and trimodal fusion settings on the ADReSSo 2021 unseen test set (n=71)[a].

Model and classifier	RMSE[b] (95% CI)
Unimodal	
SVR^{c}_{text}	4.3959 (3.7870 - 5.0445)
SVR_{audio}	6.0146 (4.6954 - 7.2511)
$Ridge^{d}_{text}$	4.5951 (3.7248 - 5.4198)
$Ridge_{audio}$	5.9914 (4.9110 - 7.1118)
RFR^{e}_{text}	4.5806 (3.8630 - 5.3138)
RFR_{audio}	8.0297 (6.8131 - 9.2356)
Pairwise	
$SVR_{text + image}$	4.3897 (3.7793 - 4.9916)
$Ridge_{text + image}$	4.3333 (3.5186 - 4.7430)
$RFR_{text + image}$	4.4310 (3.6517 - 5.2476)
$SVR_{audio + image}$	6.0201 (4.6987 - 7.2746)
$Ridge_{audio + image}$	5.8611 (4.9170 - 6.8695)
$RFR_{audio + image}$	6.1440 (5.0956 - 7.1654)
$SVR_{text + audio}$	4.6745 (3.9837 - 5.3725)
$Ridge_{text + audio}$	4.3122 (3.7134 - 4.9219)
$RFR_{text + audio}$	4.7432 (4.0348 - 5.4675)
Multimodal	
$SVR_{text + audio + image}$	4.2134 (3.4898 - 5.0544)
$Ridge_{text + audio + image}$	4.1440 (3.5511 - 4.7421)
$RFR_{text + audio + image}$	4.3856 (3.8126 - 5.0012)

[a]Root mean squared error with 95% CI is reported for unimodal models using text-only (ModernBERT embeddings) and audio-only (Wav2Vec 2.0 embeddings), bimodal fusion models using text + image, text + audio, and audio + image representations, and the proposed trimodal fusion model using text + audio + image. The results are shown for 3 regression heads: support vector regressor, ridge regression, and random forest regressor.

[b]RMSE: root mean squared error.

[c]SVR: support vector regressor.

[d]Ridge: ridge regression.

[e]RFR: random forest regressor.

For pairwise fusion, the lowest prediction error is achieved when text is included in the fusion pair, while combinations that exclude text remain substantially weaker. The best-performing bimodal configuration is Ridge with text + image, which attains the lowest RMSE of 4.3333, followed closely by SVR. These results indicate that integrating a second modality alongside text can reduce prediction error relative to text-only baselines. In contrast, pairings based on audio + image perform consistently worse across all regression heads.

Across the 3 regression heads, multimodal fusion achieves consistently low error, with Ridge obtaining the best overall RMSE of 4.1440, followed closely by SVR regression. RFR also performs competitively with an RMSE of 4.3856. Overall, the 3-way fusion setting yields strong and stable MMSE prediction accuracy across model families, with SVR and Ridge producing the lowest errors.

Comparison With Representative ADReSSo 2021 Multimodal and Transformer-Based Studies

To further position our method relative to recent benchmark-relevant approaches, we conducted a targeted scoping review of ADReSSo 2021 studies using multimodal, transformer, and attention-based methods. From this review, we selected e representative ADReSSo 2021 comparator studies for focused comparison: Wang et al [28], Zhu et al [29], Bang et al [30], and Shao and Fang [31]. These studies span attention-based multimodal fusion, wav2vec2 + Bidirectional Encoder Representations from Transformers (BERT) hybrid modeling, large language model–augmented multimodal modeling, and co-attention acoustic-text fusion, respectively.

XSL•FO
RenderX

Table 9 summarizes their reported AD classification performance alongside our proposed model. Details of the scoping-review procedure are provided in Multimedia Appendix 1.

Table . Comparison with representative modern ADReSSo 2021 multimodal and transformer-based comparator models for Alzheimer disease (AD) classification[a].

Study and year	Modalities	Model	Accuracy	F_1-score
Wang et al [28], 2021	Linguistic + IS10 + X-Vector	DNN[b]	0.80	0.83
Zhu et al [29], 2021	Audio + ASR[c] text	Wav2vec2 + BERT[d]	0.83	0.83
Bang et al [30], 2024	Audio + text + LLM[e]-generated "opinion" feature	ChatGPT-assisted multimodal model	0.87	0.87
Shao and Fang [31], 2025	Acoustic + ASR text	Co-attention multimodal model	0.83	0.84
Ours	Audio + text + image	Bidirectional attention + SVC	0.87	0.87

[a]The table reports a focused comparison between our proposed audio + text + image bidirectional attention framework and four representative ADReSSo 2021 studies selected from the targeted scoping review: Wang et al [28] (attention-based multimodal deep neural network), Zhu et al [29] (wav2vec2 + BERT), Bang et al [30] (large language model–augmented multimodal model), and Shao and Fang [31] (co-attention multimodal fusion). Accuracy and F_1-score are reported as provided in the respective studies.

[b]DNN: deep neural network.

[c]ASR: automatic speech recognition.

[d]BERT: Bidirectional Encoder Representations from Transformers.

[e]LLM: large language model.

Discussion

Our findings indicate that trimodal fusion of text, audio, and image provides the overall strong performance on ADReSSo for both AD classification and MMSE prediction. Relative to unimodal and bimodal alternatives, the trimodal setting yields the best overall classification performance and the lowest or near-lowest regression error across model families. Because ADReSSo was designed to reduce common confounds such as demographic imbalance, these gains are likely to reflect disease-relevant information rather than spurious correlations [8].

A consistent trend across the tasks is that text representations dominate unimodal performance, with transcript-derived embeddings outperforming audio-only models by a substantial margin. This aligns with longstanding clinical and computational evidence that connected speech contains measurable dementia-related changes in lexical selection, semantic content, and syntactic organization [32]. In our experiments, ModernBERT-based embeddings preserve much of this discriminative linguistic structure, yielding the strong unimodal baselines across both classification and MMSE prediction.

Although audio-only performance is weaker, audio contributes useful complementary information when paired with text, particularly for AD classification where text + audio is a rather competitive bimodal configuration and for MMSE prediction. This is consistent with prior work on paralinguistic and timing-related markers in dementia, such as pausing, hesitation, and fluency disruptions, which are not fully captured in transcripts alone [33]. The use of self-supervised acoustic representations such as wav2vec 2.0 provides a principled way to encode these cues from raw speech [19].

The image modality plays a more nuanced role in this benchmark. Because the Cookie Theft stimulus is fixed across participants, an image-only system is not expected to be discriminative; however, our results suggest that including image embeddings can still improve performance when used jointly with text and audio. One plausible explanation is that the image representation acts as a shared contextual anchor for the narrative. This interpretation is supported by several findings. First, drawing on the image-text matching approach [34], prior work [35] shows that healthy and dementia samples differ in their relevance to the target picture, with healthy samples yielding higher relevance scores. This finding suggests that visual information can contribute meaningfully to dementia detection. Second, we conducted a series of comparisons to evaluate the added value of the image modality (Tables 1-5), including text versus text + image, audio versus audio + image, and text + audio versus text + audio + image. Across all settings, incorporating the image leads to consistent, albeit modest, performance gains when combined with other modalities. Third, we performed a control analysis by randomly shuffling half of the image patches in the text + image fusion model. This manipulation resulted in a drop in performance relative to the original images (F_1-score: 0.84 vs 0.79), providing direct evidence that preserving the visual structure is important for model effectiveness. Taken together, these results indicate that the image serves as a semantic anchor to leverage the full contextual information available. Since the same image is used for every participant, the study demonstrates the usefulness of shared stimulus grounding rather than the predictive value of participant-specific image data.

A key methodological advance from our ablation study is that bidirectional cross-attention consistently outperforms standard

XSL•FO
RenderX

1-way cross-attention across classifier families. This suggests that reciprocal information exchange between modalities is more effective than a single one-way update for this task. In other words, allowing each modality to attend to and be updated by the other leads to stronger fusion than unidirectional conditioning alone [12]. The comparison with early and late fusion further shows that how modalities are integrated matters. Early fusion and late fusion are simple and widely used, but they do not model cross-modal relationships as explicitly as attention-based intermediate fusion. Their weaker performance in our experiments supports the value of learned cross-modal interaction over heuristic fusion strategies [35].

Across tasks, the regression results mirror classification trends: text-only is the strongest among unimodal baselines, bimodal improvements are most pronounced when text is included, and trimodal fusion produces the lowest overall error. This consistency suggests that the multimodal embedding learned through attention mechanisms supports both discrete diagnostic labeling and continuous cognitive assessment, a useful property for clinical decision support where disease manifestations often exist on a spectrum rather than a binary boundary. At the same time, RMSE values remain nontrivial, highlighting that cognitive score prediction remains a challenging objective likely influenced by heterogeneity in impairment profiles and the limited size of available labeled data.

Beyond the in-benchmark comparisons reported in the *Results* section, Table 9 provides additional context against representative modern ADReSSo 2021 multimodal and transformer-based studies. As shown in Table 9, our method remains competitive relative to these comparator models. At the same time, direct one-to-one comparison remains imperfect even within the ADReSSo-only subset because studies differ in modality definitions, auxiliary cues, and evaluation protocols. We therefore use Table 9 as a focused benchmark-context comparison rather than as a strict ranked state-of-the-art table.

Several limitations should be considered when interpreting these findings. First, the findings have not yet been validated across an independent dataset, a different elicitation image, or a different clinical population. ADReSSo is a relatively small benchmark dataset for training and evaluating multimodal deep learning models, with 237 total samples and a held-out test set of 71 recordings. Although the dataset is carefully curated and demographically balanced, the limited sample size may constrain generalizability and increase the possibility that strong performance partly reflects benchmark-specific characteristics rather than broader robustness across populations or recording conditions. Accordingly, this study should be interpreted primarily as a benchmarked methodological investigation on ADReSSo that evaluates the potential of bidirectional multimodal fusion, rather than as definitive evidence of clinical generalization. Second, picture-description tasks represent a structured elicitation paradigm; performance may differ for more naturalistic conversational speech, where topic drift and dialogue dynamics introduce additional complexity. Third, transcript-based modeling depends on automatic speech recognition quality; while Whisper-style transcription is robust in many conditions, transcription choices around disfluencies and fillers can influence downstream embedding behavior especially in tasks where such markers may carry clinical signal.

Finally, these results suggest several directions for future work. More fine-grained alignment strategies (eg, segment-level coupling between audio and text representations) may strengthen the ability of cross-attention to learn clinically meaningful correspondences. Additionally, evaluation across multiple elicitation tasks and datasets would help clarify whether the observed gains extend beyond Cookie Theft.

In this work, we introduced an attention-based trimodal fusion framework for dementia screening that integrates text, audio, and image through bidirectional attention. Evaluated on the ADReSSo Cookie Theft picture-description benchmark, the proposed approach achieved overall strong performance for both AD classification and MMSE prediction, outperforming unimodal baselines, bimodal fusion models, and conventional early- and late-fusion strategies. The ablation analysis further showed that bidirectional cross-attention performs better than standard 1-way cross-attention, highlighting the value of reciprocal cross-modal interaction for dementia-related prediction.

Acknowledgments

We are grateful for the ADReSSo Challenge data that were available via DementiaBank. This work was supported by computational resources provided by the Centre for Large AI Models (CLAIM) of The Hong Kong Polytechnic University. ChatGPT was used for basic error correction (grammar, typos, and editing) in the initial draft for the purpose of rephrasing or rewording.

Funding

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number P30AG073105. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Additional support is provided by PolyU Strategic Hiring Scheme and Faculty Reserve.

Data Availability

All the data are available online [36].

XSL•FO
RenderX

Authors' Contributions

Conceptualization: FA, HL
Formal analysis: FA
Investigation: FA
Methodology: FA, HL
Supervision: HL
Writing – original draft: FA, HL
Writing – review and editing: FA, HL

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details of the scoping review procedure.
[DOCX File, 23 KB - aging_v9i1e93279_app1.docx]

References

1. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):263-269. [doi: 10.1016/j.jalz.2011.03.005] [Medline: 21514250]
2. Taler V, Phillips NA. Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. J Clin Exp Neuropsychol 2008 Jul;30(5):501-556. [doi: 10.1080/13803390701550128] [Medline: 18569251]
3. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 8;396(10248):413-446. [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
4. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
5. Jack CR Jr, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. Alzheimers Dement 2024 Aug;20(8):5143-5169. [doi: 10.1002/alz.13859] [Medline: 38934362]
6. Agbavor F, Liang H. Predicting dementia from spontaneous speech using large language models. PLOS Digit Health 2022 Dec;1(12):e0000168. [doi: 10.1371/journal.pdig.0000168] [Medline: 36812634]
7. Eyigoz E, Mathur S, Santamaria M, Cecchi G, Naylor M. Linguistic markers predict onset of Alzheimer's disease. EClinicalMedicine 2020 Nov;28:100583. [doi: 10.1016/j.eclinm.2020.100583] [Medline: 33294808]
8. Luz S, Haider F, Fuente SDL, Fromm D, MacWhinney B. Detecting cognitive decline using speech only: the ADReSSo challenge. Presented at: INTERSPEECH 2021; Aug 30 to Sep 3, 2021. [doi: 10.21437/Interspeech.2021-1220]
9. Ksibi A, Walha A, Zakariah M, Ayadi M, Alshalali T, Almujally NA. Multimodal Siamese networks for dementia detection from speech in women. Sci Rep 2025 Aug 22;15(1):30938. [doi: 10.1038/s41598-025-13902-7] [Medline: 40847098]
10. Ilias L, Askounis D. Multimodal deep learning models for detecting dementia from speech and transcripts. Front Aging Neurosci 2022;14:830943. [doi: 10.3389/fnagi.2022.830943] [Medline: 35370608]
11. Tsai YHH, Bai S, Yamada M, Morency LP, Salakhutdinov R. Transformer dissection: an unified understanding for transformer's attention via the lens of kernel. In: Inui K, Jiang J, Ng V, Wan X, editors. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP: Association for Computational Linguistics; 2019:4344-4353. [doi: 10.18653/v1/D19-1443]
12. Atrey PK, Hossain MA, El Saddik A, Kankanhalli MS. Multimodal fusion for multimedia analysis: a survey. Multimedia Systems 2010 Nov;16(6):345-379. [doi: 10.1007/s00530-010-0182-0]
13. Arevalo J, Solorio T, Montes-y-Gómez M, González FA. Gated multimodal units for information fusion. arXiv. Preprint posted online on Feb 7, 2017. [doi: 10.48550/arXiv.1702.01992]
14. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. Presented at: 31st Conference on Neural Information Processing Systems (NIPS 2017); Dec 4-9, 2017 URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf [accessed 2026-09-10]
15. Tan H, Bansal M. LXMERT: learning cross-modality encoder representations from transformers. In: Inui K, Jiang J, Ng V, Wan X, editors. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): Association for Computational Linguistics; 2019:5100-5111. [doi: 10.18653/v1/D19-1514]
16. Agbavor F, Liang H. Dementia detection from spontaneous speech using cross-attention fusion. J Dement Alzheimers Dis 2026;3(1):12. [doi: 10.3390/jdad3010012]
17. Goodglass H, Kaplan E, Barresi B. The Assessment of Aphasia and Related Disorders, 3rd edition: Lippincott Williams & Wilkins; 2001.

XSL•FO
RenderX

18. Becker JT, Boller F, Lopez OL, Saxton J, McGonigle KL. The natural history of Alzheimer's disease. Description of study cohort and accuracy of diagnosis. Arch Neurol 1994 Jun;51(6):585-594. [doi: 10.1001/archneur.1994.00540180063015] [Medline: 8198470]
19. Baevski A, Zhou Y, Mohamed A, Auli M. Wav2vec 2.0: a framework for self-supervised learning of speech representations. Presented at: 34th Conference on Neural Information Processing Systems (NeurIPS 2020); Dec 6-12, 2020 URL: https://proceedings.neurips.cc/paper/2020/hash/92d1e1eb1cd6f9fba3227870bb6d7f07-Abstract.html [accessed 2022-07-14]
20. Radford A, Kim JW, Xu T, Brockman G, McLeavey C, Sutskever I. Robust speech recognition via large-scale weak supervision. Presented at: ICML'23: Proceedings of the 40th International Conference on Machine Learning; Jul 23-29, 2023 URL: https://dl.acm.org/doi/10.5555/3618408.3619590 [accessed 2026-09-10]
21. Warner B, Chaffin A, Clavié B, et al. Smarter, better, faster, longer: a modern bidirectional encoder for fast, memory efficient, and long context finetuning and inference. In: Che W, Nabende J, Shutova E, Pilehvar MT, editors. Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers): Association for Computational Linguistics; 2025:2526-2547. [doi: 10.18653/v1/2025.acl-long.127]
22. Alishiri GH, Bayat N, Fathi Ashtiani A, Tavallaii SA, Assari S, Moharamzad Y. Logistic regression models for predicting physical and mental health-related quality of life in rheumatoid arthritis patients. Mod Rheumatol 2008 Dec;18(6):601-608. [doi: 10.3109/s10165-008-0092-6]
23. Biau G. Analysis of a random forests model. J Mach Learn Res 2012;13:1063-1095 [FREE Full text]
24. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825-2830 [FREE Full text]
25. Schölkopf B, Smola AJ. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond: The MIT Press; 2001. [doi: 10.7551/mitpress/4175.001.0001]
26. Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 1970 Feb;12(1):55-67. [doi: 10.1080/00401706.1970.10488634]
27. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science 2006 Jul 28;313(5786):504-507. [doi: 10.1126/science.1127647] [Medline: 16873662]
28. Wang N, Cao Y, Hao S, Shao Z, Subbalakshmi KP. Modular multi-modal attention network for Alzheimer's disease detection using patient audio and language data. Presented at: Interspeech 2021; Aug 30 to Sep 3, 2021. [doi: 10.21437/Interspeech.2021-2024]
29. Zhu Y, Obyat A, Liang X, Batsis JA, Roth RM. WavBERT: exploiting semantic and non-semantic speech using wav2vec and BERT for dementia detection. Presented at: Interspeech 2021; Aug 30 to Sep 3, 2021. [doi: 10.21437/Interspeech.2021-332]
30. Bang JU, Han SH, Kang BO. Alzheimer's disease recognition from spontaneous speech using large language models. ETRI Journal 2024 Feb;46(1):96-105. [doi: 10.4218/etrij.2023-0356]
31. Shao Y, Fang T. Alzheimer's disease detection using co-attention mechanism for acoustic and ASR-transcribed text features. Presented at: Interspeech 2025; Aug 17-21, 2025. [doi: 10.21437/Interspeech.2025-219]
32. Ahmed S, Haigh AMF, de Jager CA, Garrard P. Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. Brain 2013 Dec;136(Pt 12):3727-3737. [doi: 10.1093/brain/awt269] [Medline: 24142144]
33. Lin H, Karjadi C, Ang TFA, et al. Identification of digital voice biomarkers for cognitive health. Explor Med 2020;1:406-417. [doi: 10.37349/emed.2020.00028] [Medline: 33665648]
34. Radford A, Kim JW, Hallacy C, et al. Learning transferable visual models from natural language supervision. Presented at: Proceedings of the 38th International Conference on Machine Learning; Jul 18-24, 2021 URL: https://proceedings.mlr.press/v139/radford21a/radford21a.pdf [accessed 2026-09-10]
35. Zhu Y, Lin N, Liang X, Batsis JA, Roth RM, MacWhinney B. Evaluating picture description speech for dementia detection using image-text alignment. ACM Trans Comput Healthcare 2025. [doi: 10.1145/3777482]
36. DementiaBank. TalkBank. URL: https://talkbank.org/dementia/ [accessed 2026-09-10]

Abbreviations

AD: Alzheimer disease
AUC: area under the receiver operating characteristic curve
BERT: Bidirectional Encoder Representations from Transformers
CLIP: contrastive language–image pretraining
LR: logistic regression
MMSE: Mini-Mental Status Examination
RF: random forest
RFR: random forest regressor
Ridge: ridge regression
RMSE: root mean squared error
ROC: receiver operating characteristic

XSL•FO
RenderX

SVC: support vector classifier
SVR: support vector regressor

Edited by M O'Connell; submitted 10.Feb.2026; peer-reviewed by BH Prasetio, J Fan; revised version received 07.Jun.2026; accepted 31.Aug.2026; published 21.Sep.2026.

Please cite as:
Agbavor F, Liang H
Multimodal Dementia Prediction With Large Language Models: Cross-Attention Over Text, Audio, and Image
JMIR Aging 2026;9:e93279
URL: https://aging.jmir.org/2026/1/e93279
doi:10.2196/93279

© Felix Agbavor, Hualou Liang. Originally published in JMIR Aging (https://aging.jmir.org), 21.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI-Enhanced Automatic Life Story Structuring for Reminiscence Therapy in Older Adults: Technical Feasibility Study

Fang Gui[1*], PhD; Mengchen Yang[1*], BE; Liuqi Jin[2*], PhD; Jing Qian[1*], PhD

[1]School of Artificial Intelligence, Chuzhou University, No. 1 Huifeng West Road, Chuzhou, Anhui Province, China

[2]School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China

[*]all authors contributed equally

Corresponding Author:
Jing Qian, PhD
School of Artificial Intelligence, Chuzhou University, No. 1 Huifeng West Road, Chuzhou, Anhui Province, China

Abstract

Background: Storytelling interventions have demonstrated substantial potential in improving emotional well-being, cognitive function, and quality of life for older adults. However, its effectiveness is often limited by the challenges of processing disorganized and redundant life stories, which impose substantial cognitive demands on caregivers. Although storytelling interventions are a well-established therapeutic approach, current practices depend heavily on manual narrative organization, restricting both the scalability and consistency of treatment delivery. Prior research has primarily focused on validating the clinical outcomes of storytelling interventions, with insufficient attention given to technological solutions that could enhance narrative processing while preserving therapeutic integrity. Digital approaches to life story structuring remain underexplored, despite their potential to amplify storytelling benefits by reducing cognitive load and improving recall accuracy.

Objective: This study aims to design an event timeline generation algorithm to optimize the prior work of the Story Mosaic system. The optimized system enables (1) the automatic extraction of event elements from life narratives, (2) the automatic organization of fragmented life stories into structured timelines, and (3) the preservation of clinically relevant contextual details during compression. The goal is to reduce manual intervention costs while increasing treatment efficacy through artificial intelligence–driven narrative structuring.

Methods: We have designed a novel method, CARE event timeline (CARE-ET), which combines a temporal attention mechanism with graph-based event relationship modeling. Furthermore, we used the CARE-ET algorithm to optimize existing story collage systems. The system uses multifeature extraction technology to capture event clues from oral histories, prioritizes the 6 elements of events through a hierarchical attention mechanism, and uses adaptive compression algorithms to reduce redundancy while maintaining narrative continuity. To verify the effectiveness of the CARE-ET method, this paper adopts a multidimensional evaluation framework, which encompasses event summary assessment, timeline quality evaluation, and usability testing of the optimized system.

Results: The proposed CARE-ET algorithm outperforms the baseline in both narrative flow and temporal accuracy. The Story Mosaic system, optimized by the CARE-ET algorithm, underwent usability evaluation by 10 caregivers recruited for this study. Based on standardized assessment metrics, the system received an A rating for usability. The comprehensive experimental results demonstrate that the CARE-ET method can effectively structure fragmented narratives from older adults, enhancing the usability of the Story Mosaic system.

Conclusions: The proposed method enables the structured extraction of representative event summaries, transforming disorganized life stories into an event timeline for caregiver-supported older adult well-being interventions. Future research should investigate longitudinal effects on cognitive preservation and explore integration with existing dementia care protocols. This work establishes a critical foundation for intelligent assistive technologies in geriatric mental health interventions.

(JMIR Aging 2026;9:e83122) doi:10.2196/83122

KEYWORDS

storytelling; life story organization; caregiver support; Story Mosaic system; older adult well-being interventions

Introduction

Cognitive impairment is a common condition affecting the health of older adults, and it is difficult to treat through medication [1]. People with mild cognitive impairment may encounter the following issues: (1) impaired short-term and long-term memory with reduced recall efficiency [2]; (2) deteriorated social functioning, exhibiting social withdrawal and diminished emotional interaction, which can lead to social

XSL•FO
RenderX

isolation [3, 4]; and (3) declined executive function in daily activities [5,6].

Through the analysis of over a hundred studies, Zhang et al [7] demonstrated that digital health technologies can effectively address the limitations of traditional medical models by delivering personalized and highly accessible intervention strategies, thereby facilitating timely intervention during the early stages of cognitive decline. Storytelling is an important application of digital health technology, which plays a significant role in the field of reminiscence [8]. Harnessing the power of storytelling, people with mild cognitive impairment can be encouraged to participate more proactively in their health care [9]. Additionally, life stories can serve as a database and decision support tool for understanding older adults and providing personalized services [10-13]. In practice, Zhu et al [14] have demonstrated that storytelling requires older adults to recall, organize, and narrate personal experiences—a process that naturally exercises cognitive functions such as memory, attention, and language abilities.

Interviews and daily interactions with older adults are a common way to access their life stories [15]. However, these life stories often have problems such as redundant content and disorganized structure, which increase the cognitive load on caregivers when learning them [16]. Moreover, as the number of exchanges grows and the volume of life stories continues to expand, the scattered information becomes increasingly difficult for caregivers to process [17,18]. In our prior work, our team developed the Story Mosaic system, designed to assist caregivers in integrating and organizing life stories for older adults. The findings were published in the international journal *JMIR Aging*. However, this version of the system has significant limitations in functional implementation: the extraction of story elements and the arrangement of narrative logic remain highly dependent on manual operations. This not only substantially increases the workload for caregivers but also necessitates the assistance of volunteers in practical applications to complete the recording and organization of older adults' life stories [17].

An effective solution is to extract the core events from older adults' life stories, rearrange the events in the order in which they occurred, and generate an event timeline structure of older adults' life stories. Then, based on the algorithm, develop a visualization system to present the timeline of older adults. An event timeline is a series of events, activities, or important moments recorded and displayed in chronological order, with a simple and clear presentation that allows users to quickly understand the temporal relationship and development of events. The structuring of life stories lies at the critical intersection point between the basic data infrastructure and artificial intelligence applications. It addresses the current problem of "fragmented data and lack of depth" in personalized intervention research, laying the foundation for the development of an intelligent intervention system that truly "understands you" [19].

Existing studies have made some progress in personal event timeline generation, but there are still limitations that do not meet the requirements of older adults' life stories. The Twitter (subsequently rebranded X) event extraction method proposed by Li et al [20] treats frequent mentions of events as general events, which contradicts the characteristics of older adults' narratives. Althoff et al's [21] timeline construction method relies on precise temporal information, which is difficult to apply to the temporally ambiguous life stories. Sun et al's [22] stage-by-stage story line generation method ensures plot integrity but lacks a generalized description of key events. The pretraining with extracted gap-sentences for abstractive summarization model learns to generate summaries by pretraining with masked key sentences, but its reliance on document structural integrity and sentence coherence makes it difficult to handle issues such as temporal confusion, information redundancy, or fragmentation in the life stories of older adults [23]. The long text-to-text transfer transformer model uses a local-global attention mechanism to efficiently process long texts and avoid excessive information repetition, but its limited ability to capture fine-grained temporal relationships constrains its performance in strictly reorganizing fragmented life events into chronological timelines [24]. The pyramid-based masked sentence pretraining for multidocument summarization model enhances multidocument information integration through pyramid-based masked pretraining; however, as the life stories of older adults are often fragmented and narrated individually, the lack of cross-document structural associations limits the effectiveness of its hierarchical modeling approach [25]. The life stories of older adults exhibit intrinsic particularities, such as ambiguous temporal information, conflicting event elements, and difficulty in determining event significance. These characteristics directly contribute to the limitations of existing methods in processing such narratives.

To address limitations in existing approaches to organizing life stories for older adults, this paper proposes a CARE event timeline (CARE-ET), an innovative event timeline generation method that systematically structures narratives through optimized event relevance analysis and implements the practical Story Mosaic system for caregiver application. Our research makes the following three key contributions. First, we developed a multimodal event fusion method using a 2-layer encoder-decoder architecture that integrates event triggers and timestamps. This structured representation model significantly improves accuracy in extracting key events from older adult narratives. Second, the theoretical novelty lies in the synergistic integration of graph attention networks (GATs) and deep belief networks (DBNs). In contrast to existing hybrids where DBNs serve merely as static feature extractors, the CARE-ET model uses a joint self-supervised module. Here, the latent clustering probability distribution generated by the DBN acts as a global before regularizing the GAT attention coefficients. This ensures that the learned values are optimized not only for local semantic alignment but also for global thematic coherence across the life story timeline. Finally, building upon the CARE-ET method, we have optimized and enhanced the previously developed Story Mosaic system. The improved system is designed to construct more logically coherent life story timelines for older adults, thereby providing reliable data support for personalized care.

XSL•FO
RenderX

Methods

This section first elaborates on the overall algorithmic framework of the CARE-ET method and concludes with an introduction to the optimized and enhanced functionalities of the Story Mosaic system.

Ethical Considerations

This study aims to propose a novel algorithm to optimize the Story Mosaic system, thereby better facilitating the automatic organization of fragmented life stories of older adults. This study has been approved by the research ethics committees of Hefei University of Technology (HFUT20220921001) and Chuzhou University (CZSC2025-021), respectively. All participants, including older adults and caregivers, signed a written informed consent form before the initiation of the study. The informed consent form explicitly included authorization for audio recording of the interviews, verbatim transcription of the recorded content, and the use of such materials for academic research purposes. Participation is entirely voluntary. Participants have the right to withdraw unconditionally at any stage of the study without assuming any liability or providing a reason. Given the nature and objectives of this study, there are no circumstances that necessitate the provision of financial or any other form of compensation to participants. All data collected during the research process (including interview recordings and transcribed texts) have undergone rigorous anonymization. All personally identifiable information has been removed or replaced, and comprehensive privacy protection measures have been implemented. The relevant information and images related to the older adults presented in this paper do not involve any personal privacy.

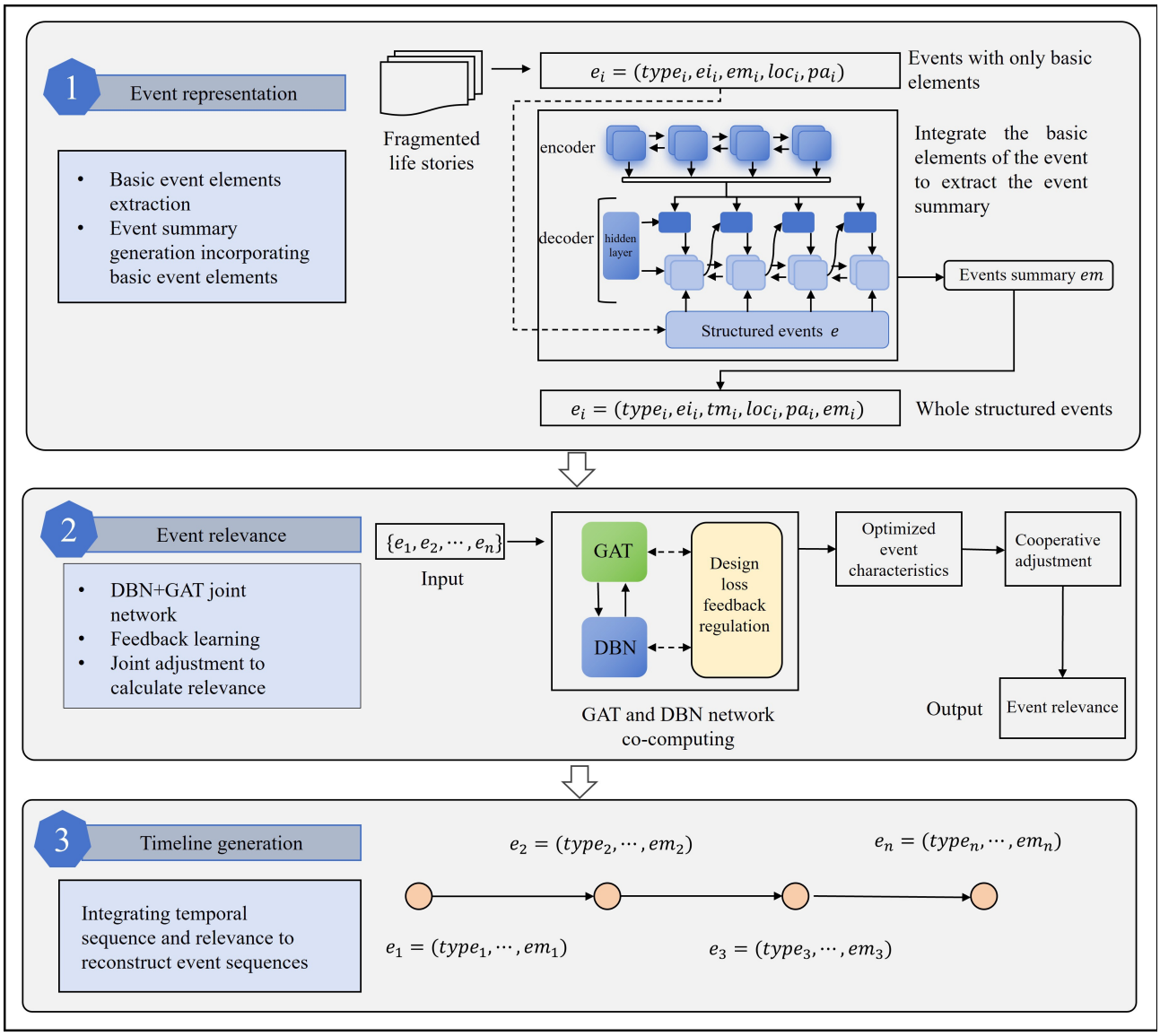
CARE-ET Method

Overview

The CARE-ET method automatically generates life story timelines through a three-stage computational process, as shown in Figure 1: (1) bidirectional long short-term memory (BiLSTM)-based event extraction and representation, which synthesizes triggers, time stamps, and semantic context into structured events; (2) hybrid DBN-GAT modeling for self-supervised event correlation computation using optimized loss functions and distance-density joint adjustment; and (3) integrating temporal features with semantic correlations to produce coherent event timelines.

XSL•FO
RenderX

Figure 1. The CARE event timeline (CARE-ET) algorithmic framework for automated life story structuring. The framework comprises three stages: (1) event extraction via bidirectional long short-term memory (BiLSTM), (2) relevance modeling using the joint network of graph attention network (GAT) and deep belief network (DBN), and (3) timeline generation. Rectangles represent computational modules, while arrows indicate the flow of feature vectors and processed narrative data.



Event Extraction and Presentation

Event extraction and representation consist of 2 main processes: the extraction of basic event elements and the generation of event summaries incorporating the basic event elements. Event information is extracted from older adults' life stories. The structured event representations include event type, event trigger word, time, location, participant, and event summary, and the formal expression is: ei=type,eii,tmi,loci,pai,emi(1)

Structured Event Extraction

The module achieves coreference resolution by using ALBERT (A Lite Bidirectional Encoder Representations from Transformers for Self-supervised Learning of Language Representations) embeddings with a 512-token sliding window for context management. Similarity between candidate expressions is measured via cosine distance, with a threshold of 0.8 established to determine coreferentiality [26,27]. We adopt a semisupervised open domain event extraction approach to expand event templates through manual annotation of triggers and types, enabling systematic extraction of basic event elements (type, trigger, time, location, and participants) for structured representation of older adults' life stories [28].

Context-Aware Summary Generation

This module builds a dual-level encoder-decoder architecture. It uses ALBERT-BiLSTM for encoding and combines BiLSTM decoding with dynamic attention mechanisms [16]. The model prioritizes event-centric content by jointly analyzing word or sentence-level features alongside spatiotemporal and participant context. For event-level summarization, the system selects the top three most salient sentences using maximal marginal relevance based on ALBERT's cross-attention scores. This ensures that the selected content maximizes coverage of the event trigger while minimizing information redundancy between sentences.

XSL•FO
RenderX

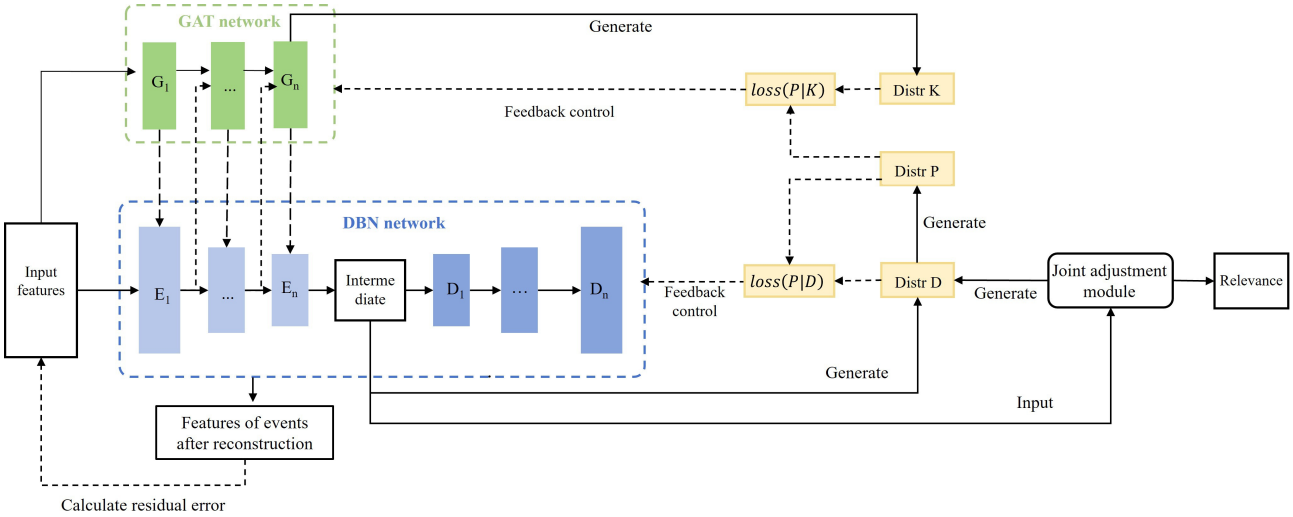
Event Relevance Calculation

Overview

After the event extraction module extracts the structured events for each life story, we need to calculate event correlations to arrange the event nodes on the timeline.

The CARE-ET method uses a novel GAT-DBN collaborative model to compute event relationships for timeline construction [29]. The architecture combines: (1) an enhanced GAT network that captures multihop event dependencies through dynamic attention weights, extending influence beyond immediate neighbors; (2) a DBN component that learns compressed event representations while preserving relational semantics; and (3) a hybrid k-means or density-based spatial clustering of applications with noise (DBSCAN) self-supervision module that optimizes feature clustering through distance-density joint optimization [30], as shown in Figure 2. This integrated approach automatically learns hierarchical event relationships while maintaining computational efficiency through backpropagation-based training.

Figure 2. Detailed architecture of the event relevance computing model. This diagram illustrates the collaborative workflow between the graph attention network (GAT; which captures global dependencies) and the deep belief network (DBN; which performs feature compression). The joint adjustment module integrates the results to output refined correlation coefficients used for timeline sequencing.



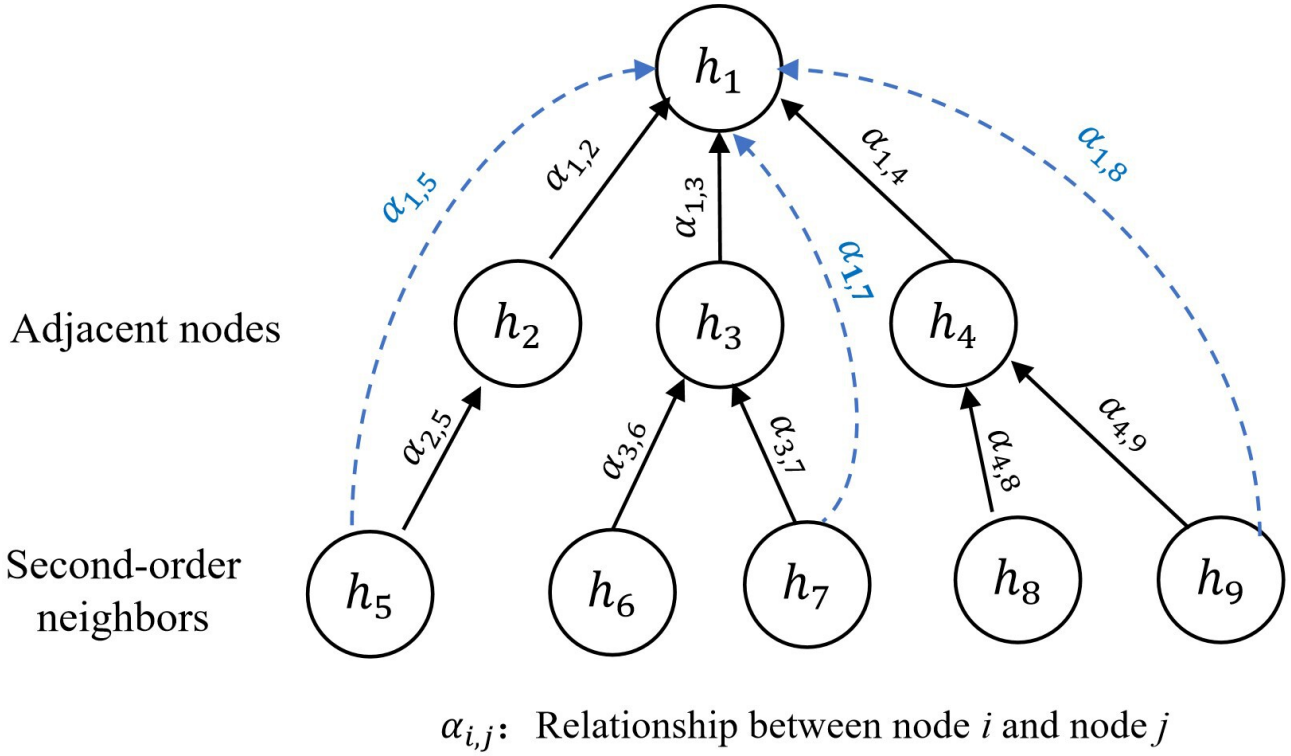
The Improved GAT Network

The proposed GAT extension structure advances traditional attention mechanisms by incorporating multihop node relationships during state updates. While standard GAT computes attention weights based solely on direct neighbors [31], our method dynamically integrates deeper network contexts through an adaptive depth strategy, enabling more comprehensive feature representation while maintaining computational efficiency. This enhancement allows each node to capture richer structural patterns by considering both immediate connections and semantically relevant indirect relationships within life story narratives.

Figure 3 illustrates the GAT network's enhanced state update mechanism, where each node's output (eg, $_1$) incorporates weighted influences from both direct neighbors ($_2$, $_3$, $_4$) and secondary neighbors ($_5$, $_7$, $_9$). The model automatically filters insignificant connections (eg, $_6$ and $_8$) through threshold-based attention weights, enabling efficient multihop relationship modeling while maintaining computational efficiency. This dynamic depth perception improves event relationship characterization without compromising network performance.

XSL•FO
RenderX

Figure 3. The mechanism of graph attention network node updates based on multihop relationships. Circles represent events $_i$. Solid arrows indicate direct influence from immediate neighbors, while dashed arrows denote filtered indirect connections identified by the adaptive depth strategy. Gray nodes (eg, $_6$ and $_8$) represent connections that have been discarded due to the similarity threshold.



Assuming that a total of events exist, denotes the event, and each event is characterized as an -dimensional vector. All event features can be represented as the set = { $_1$, $_2$, , }. Based on this, the features of all events can be integrated into a sample feature matrix of × :

(2)Φ=[e11e21 en1e12e22 en2 e1me2m enm]

The GAT network can be described as:

(3)O = G(Φ)

The improved GAT network facilitates node feature evolution through linear transformations and attention-weighted aggregation. The feature update for each node in a specific layer is defined as:

(4)hil=σ(∑j∈ (Ni∪Ii)aijlWlhjl→)

where is the network input, denotes the network function, and denotes the network output. Thus, the GAT network can be described as:

(5)O=[o11o21 on1o12o22 on2 o1m′o2m′ onm′]

where ′ is the new dimension of each event vector element, the actual output is a matrix consisting of new vectors describing each event.

The core of the GAT model lies in the graph attention layer, which maps the upper-layer graph node states to the lower-layer graph node states. The input to the graph attention layer, $^{l-1}$, is the set of graph node feature vectors output from the upper layer.

(6)Hl−1={e1Hl−1,e2Hl−1, ,enHl−1}

is the number of nodes in the network, and $^{-1}$ is the dimension of the output features of the previous layer. Thus, the size of the matrix $^{l-1}$ is × $^{-1}$:

(7)H=[h11h12h1nh21h22h2nhn1hn2hnn]

The output of this layer is a new set of node feature vectors M′, denoted as:

(8)M′={e1Hl,e2Hl, ,enHl}

The attention coefficient θij is computed using a single-layer feed-forward neural network, parameterized by a weight vector a, which operates on the concatenated features of the node pair:

(9)θij=a→(Wh→i,Wh→j)

(10)θij=LeakyReLu(a→T[Wh→i||Wh→j])

The proposed model extends traditional GAT attention computation through a dynamic coefficient function that incorporates both direct neighbors and validated indirect connections , as shown in Figure 3. By enhancing the masked attention mechanism with SoftMax regularization, the system captures multi-hop dependencies while maintaining stable gradient propagation, significantly improving relationship modeling for life event networks compared to standard single-hop approaches.

This formula represents the attention coefficient of node *i* to node *j*, where α is not a constant but a node-dependent function, and *W* denotes the trainable weight parameters. The key innovation of our GAT enhancement lies in its refined attention allocation mechanism.

Unlike traditional GAT models, which compute attention solely over the direct neighbors of node *i*, our improved masked

XSL•FO
RenderX

attention mechanism extends the scope of attention coefficient computation, enabling the exploration of indirect node relationships. Additionally, we introduce a SoftMax function to normalize attention scores across both direct neighbors and selected indirect nodes meeting specific criteria (equation 9). This modification enhances information propagation and dependency modeling between nodes:

(11)αij=softmaxj(θij)=exp (θij)∑(k∈ Ni∪Ii) exp (θik)

where denotes the set of adjacent nodes for node *i*, while represents the set of qualified indirect nodes for node *i*. According to the definition of and Ii, only events with a similarity exceeding a predefined threshold are established as related events.

Our GAT extension introduces an adaptive second-hop set Ii based on semantic relevance thresholds. While heterogeneous graph attention typically focuses on predefined meta-paths [32], our approach dynamically identifies high-influence indirect neighbors in the narrative graph. By redefining the training of attention coefficients α over Ii, the model captures long-range dependencies in disorganized oral histories while mitigating the oversmoothing issues common in deep GNNs.

The network function determines the node influence weights based on the correlation between the 2 events. The cosine similarity function is used to calculate the event similarity between the 2 events and :

(12)ϕ(ei,ej)=ei·ej|ei||ej|

To model the structural relevance of life events, we define the neighbor set Ni for a given event node ei as:

(13)Ni={ej ∣ cos_simhi,hj§amp;gt;τ1}

where τ1=0.6.

For global consistency adjustment, the second-order inference set Ii is constructed as:

(14)Ii={ek ∣ ∃ej∈ Ni,cos_simhj,hk§amp;gt;τ2}

where τ2=0.8.

Next, the LeakyReLU function is adopted as the single-layer feed-forward function of the attention mechanism, and the improved attention mechanism is shown below：

(15)[illegible]

The attention function can calculate the attention coefficients between different nodes and predict the output characteristics of each node. The output characteristics are：

(16)h→i=σ(∑j∈ (Ni∪Ii) aijWh→j)

Meanwhile, this method adopts the multihead attention mechanism, which can make the attention training more stable. The core idea is to execute *m* attention functions independently and use the average distribution of *m* attention outputs as the output of the overall attention mechanism：

(17)hi′→=σ(1m∑1m ∑j∈ (Ni∪Ii) aijmWmh→j)

GAT Synergy With DBN Networks

The model integrates GAT and DBN through a bidirectional encoding-decoding framework. GAT captures interevent correlations via attention-weighted node interactions, while DBN performs hierarchical feature compression through multilayer transformations. Their synergistic operation—alternating GAT's encoded outputs with DBN's hidden states during each encoding round—preserves individual event characteristics while enhancing semantic relationships. The complementary decoding process ensures reconstructed features maintain fidelity to original inputs while optimizing them for correlation computation, achieving both dimensionality reduction and relationship-aware representation learning.

In this synergistic interaction model, the inputs and outputs of each layer of the DBN coding part are shown in equation 16：

(18)hi+=sigmoid(∑j |Hj|wjvj+bi)

Also, the inputs and outputs of each GAT network can be adapted in the following form:

(19)Oi=G(Oi−1,hi−1+)

Self-Monitoring Module

To address the coordination challenge between GAT and DBN networks, we introduce a novel self-supervised joint adjustment module that serves as an integrative optimizer. This module dynamically aligns the feature representations from both networks during prediction by: (1) generating target distributions through hybrid k-means or DBSCAN clustering, (2) constructing a unified loss function for backpropagation, and (3) enabling simultaneous parameter optimization across both architectures. The resulting feature representations are specifically tuned for accurate event correlation computation while maintaining the distinct advantages of each network's processing paradigm.

If the 2 events are highly correlated, their features should also be highly similar. Multiple events describing similar contexts will form clusters in the feature space. In this paper, we use the output features of the DBN as inputs to the joint adjustment module to compute the interevent correlation and cluster structure. This is done by calculating the similarity between each event feature within a cluster and the center event feature of that cluster, which in turn generates a distribution *D* that contains the correlations of all the events, where $_i$ denotes the DBN feature of the ith event, is the center event of the jth cluster, and $_{cj}$ is the center event feature and is a fixed parameter. The construction of the correlation distribution is accomplished by traversing all clusters and their center events. The correlation distribution of the th event belonging to the th class is shown in equation 18:

(20)dij=(1+|hi−hcj|2vc)−(νc+1)∑n (1+|hi−hcn|2vc)−(νc+1)

The self-supervised module enhances event feature representations by optimizing correlations between individual events and their cluster centroids, thereby improving semantic similarity grouping among related events. The target distribution of event relevance is first computed ：

(21)pij=dij2/fi∑j′dij2/fi

where = ∑ is the sum of the frequencies of a category of event correlation, and ∑j`dij2/fi is the sum of the frequencies of all categories of event correlation, where ′ denotes the

XSL•FO
RenderX

individual categories in the computation. The following loss function is constructed from the distributions P and D for inverse training the DBN. The self-supervision module uses the Kullback-Leibler divergence to quantify the discrepancy between the predictive and target distributions. The joint loss function is formally defined as:

(22)L=KL(P|D)=∑i ∑j pijlog pijdij

GAT also requires backpropagation training. Since both distributions P and D depend on the features generated by the DBN, more noise may be introduced if the loss(P|D) is used directly to compute the loss of GAT. The GAT network itself can generate a distribution (which is also the output of the last GAT network), using K instead of the DBN-generated distribution D, using K and P to compute the loss. In this way, the GAT-generated distribution K is supervised by the distribution P, and the loss is designed to be used for the backpropagation of the GAT, where is in the last network output matrix .

(23)loss(P|K)=∑i ∑j pijlog pijkij

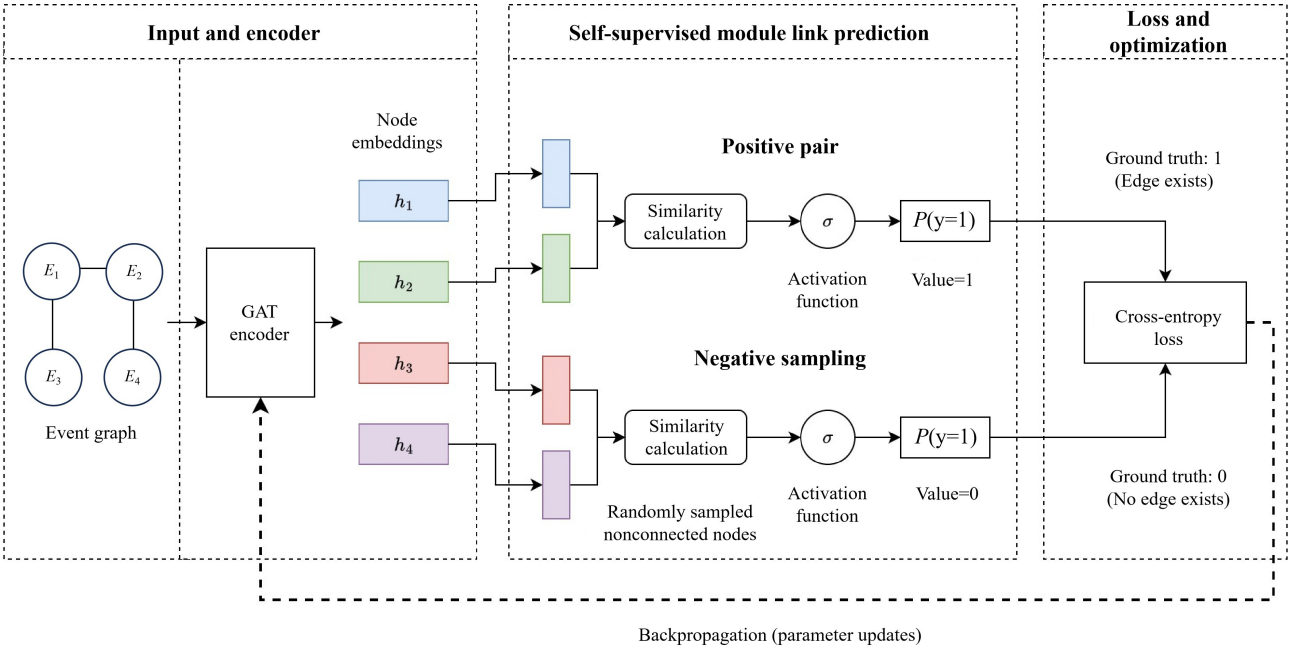
Intuitively, the self-monitoring module acts as an "internal auditor." While GAT makes initial predictions about event relevance based on local connections, the DBN provides a global thematic template. The module calculates the "surprise" or discrepancy between these two (via Kullback-Leibler divergence) and forces the model to self-correct, ensuring that local narrative links do not deviate from the overall life story logic.

The optimal number of clusters K in the joint adjustment module is determined using the silhouette coefficient and the elbow method applied to a validation subset of narratives. To prevent overfitting, we use L2 regularization and dropout (rate=0.5) within the GAT layers. Furthermore, the self-supervised loss function itself acts as a structural regularizer, penalizing the model when it focuses excessively on transient noise in individual narratives at the expense of global consistency.

The joint loss function is configured as Ltotal=λ1LGAT+λ2Ljoint, where λ1=0.7 and λ2=0.3 were empirically chosen to prioritize structural accuracy while maintaining global coherence. Optimization was performed using the AdamW routine with a weight decay of 0.01 and a linear learning rate scheduler that warmed up for the first 10% of total training steps.

Figure 4 illustrates the operational mechanism of the self-supervised module.

Figure 4. Self-supervised link prediction. A graph attention network (GAT) encoder generates node embeddings hi from the input graph. Positive and negative pairs are evaluated via dot products and sigmoid activation. The model minimizes cross-entropy loss through backpropagation to update encoder parameters and optimize latent representations.

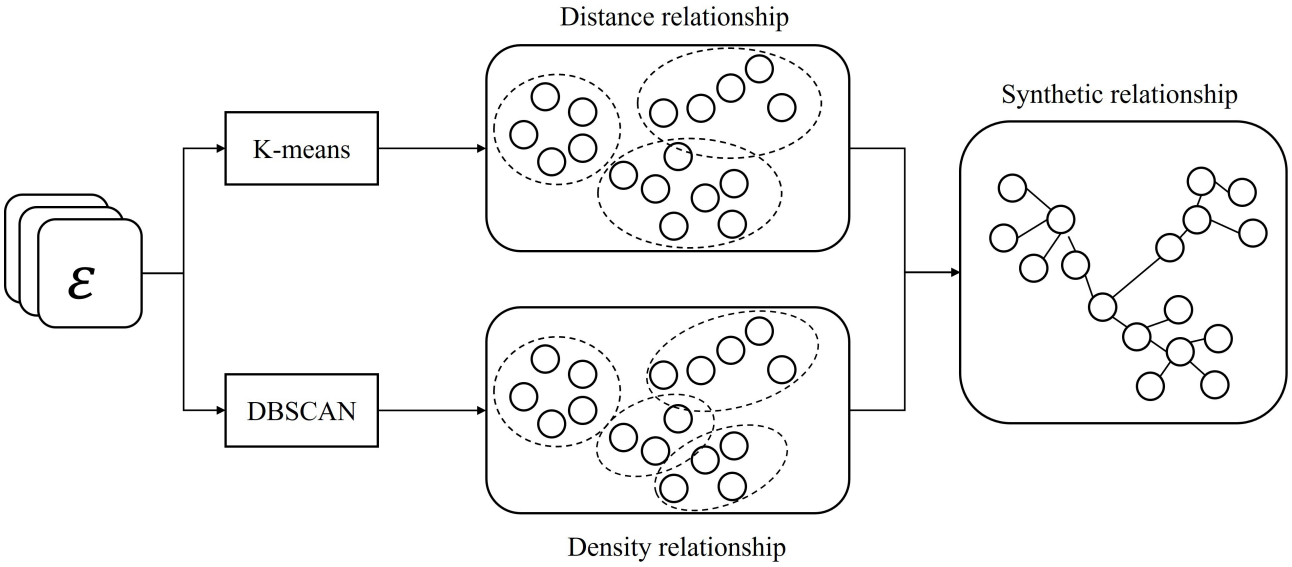


Joint Relevance Adjustment

The proposed method integrates k-means and DBSCAN clustering to compute event correlations through a dual-metric approach, as shown in Figure 5. K-means evaluates co-clustered events using relative centroid distances, while DBSCAN assesses connectivity through minimum path lengths between nodes. Both algorithms assign zero correlation to nonclustered event pairs.

XSL•FO
RenderX

Figure 5. Framework of the joint adjustment module combining geometric and density metrics. The process integrates k-means (based on Euclidean distance to centroids) and density-based spatial clustering of applications with noise (DBSCAN; based on local density connectivity). The final correlation score is a weighted fusion of these 2 metrics to ensure structural robustness in narrative clustering.



The final correlation coefficient synergistically combines both clustering results, leveraging k-means' geometric precision and DBSCAN's density awareness. This hybrid approach captures complementary aspects of event relationships while maintaining computational efficiency through conditional zero assignments for unrelated events.

Timeline Generation

To address temporal ambiguity in life stories, we first standardize diverse time expressions (eg, relative dates like "three years ago" or historical references like "Liberation War"). For events lacking absolute dates, we calculate cosine similarity between their text features and a curated historical event database, assigning standardized timestamps when similarity exceeds 0.85. This process ensures consistent temporal representation across both precise and ambiguous event references.

The generation algorithm first arranges temporally annotated events chronologically as the timeline backbone. For undated events, it calculates semantic similarity with existing timeline nodes to determine optimal insertion points between adjacent events. To integrate undated events Eu into the anchor timeline Tanchor={e1,e2,…,en}, we propose a relevance-based greedy insertion algorithm. For each eu∈ Eu, the optimal position i is determined by:

$$(21) i* = \arg\max_{0 \le i \le n} (\alpha \cdot \mathrm{Rel}(ei, eu) + (1-\alpha) \cdot \mathrm{Rel}(eu, ei+1))$$

where e0 and en+1 are virtual start and end nodes. If multiple positions yield the same maximum score, the tie is resolved by selecting the position that maintains the original narrative proximity in the transcript. This dual consideration of temporal sequence and semantic coherence produces timelines that maintain both chronological accuracy and narrative flow, essential for effective reminiscence therapy applications.

The cosine similarity threshold of 0.85 was empirically determined through a sensitivity analysis. In our pilot tests, thresholds above 0.95 resulted in a 36% loss of valid temporal anchors, while thresholds below 0.75 introduced significant "semantic drift," where personal anecdotes were incorrectly mapped to unrelated historical events. To mitigate the impact of potential misalignment on timeline validity, we introduced a "confidence flag" mechanism: events with similarity scores between 0.70 and 0.85 are flagged for human-in-the-loop verification by caregivers. Error analysis indicates that inaccuracies primarily occur in narratives with vague temporal adverbs (eg, "back then"), which the system addresses by placing them in a relative sequence based on narrative flow rather than fixed historical coordinates.

The Historical Event Database used for mapping undated narratives contains approximately 5000 major historical events spanning from 1900 to 2000, curated from open-source repositories including Wikipedia and Baidu Baike. It focuses on events highly relevant to the Chinese older adult cohort (eg, social movements, economic reforms, and significant natural disasters). The database is maintained in Chinese and is licensed under CC-BY-SA 4.0 for research purposes.

Story Mosaic System

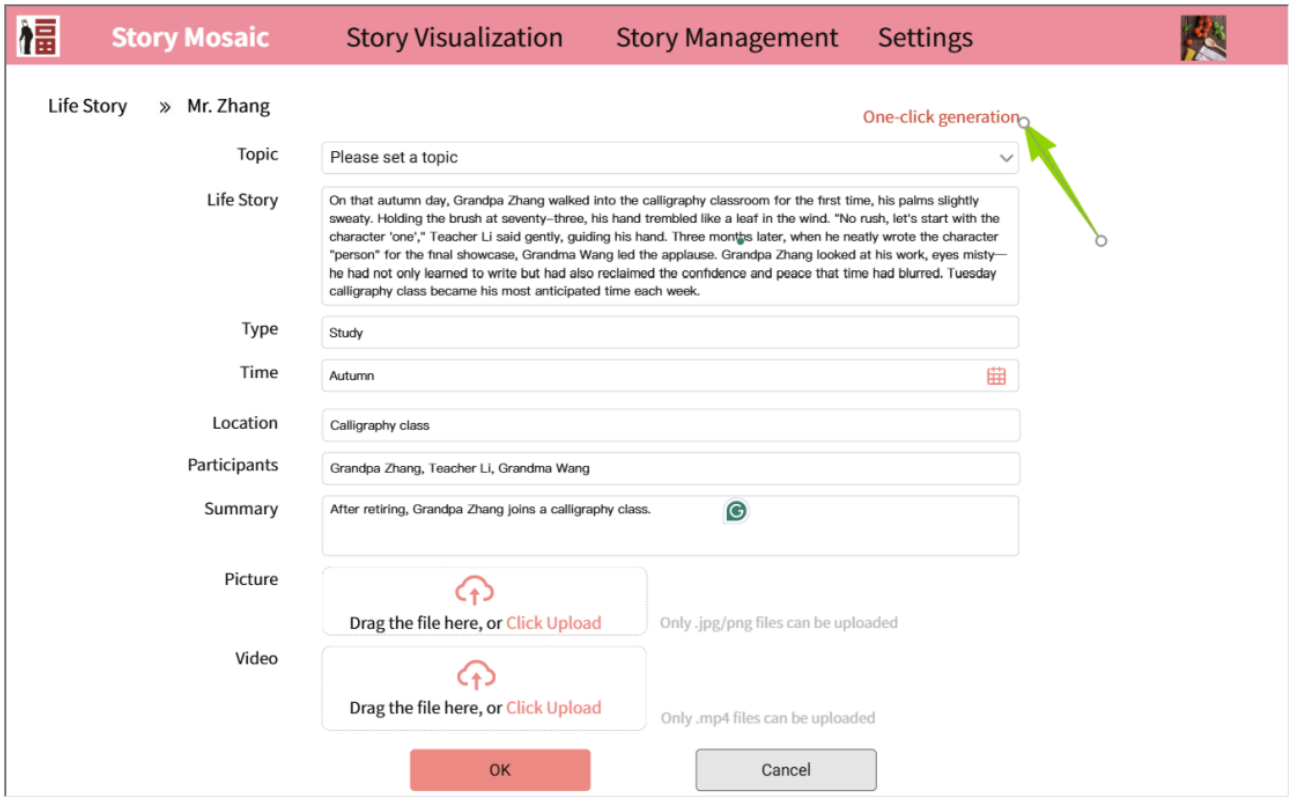
The core architecture of the Story Mosaic system is built around three principal functional modules: (1) a story input module supporting multiple formats, including text, images, and videos, with temporal annotations available for manual or automatic addition; (2) a theme-based content organization module that allows for customizable categorization (eg, "Career" and "Family Life"); and (3) a chronological timeline generation module for key life events [17].

In the preliminary work of the Story Mosaic system, the generation of life event timelines, while algorithmically automated, exhibited limited accuracy, with event elements largely reliant on manual annotation during story input. In this study, the timeline generation module has been specifically optimized based on the CARE-ET method. This enhanced module leverages the structured framework of the CARE-ET approach, integrating three core functions: event extraction, temporal sequencing, and narrative coherence, as shown in

XSL•FO
RenderX

Figure 6. It enables the automated generation of accurate event summaries and timelines, thereby alleviating the data entry burden on caregivers.

Figure 6. Schematic diagram of the automatic event feature generation module, supporting 1-click generation of event features.



Results

We conducted a comprehensive evaluation of the proposed system, organized into two key aspects: (1) quantitative assessment of event representation and timeline quality, and (2) systematic usability testing.

Datasets

We selected 2 datasets for the experiments to verify the effectiveness of the event timeline generation method. The first is the self-constructed OALS (Older Adults' Life Stories) dataset, which contains the event data TimID (timeline identity) that has been annotated with timelines. The second is the Chinese news summary dataset LCSTS (Large Scale Chinese Short Text Summarization), organized by the Harbin Institute of Technology, which is constructed based on Sina microblog text. For detailed information about the datasets, please refer to Multimedia Appendix 1.

Experimental Setup

Our model was implemented using PyTorch 1.12 (PyTorch Foundation) and trained on an NVIDIA Tesla T4 GPU (16GB VRAM). The ALBERT-base-v2 model was fine-tuned for 10 epochs with a learning rate of 2×10 , a batch size of 16, and a weight decay of 0.01 using the AdamW optimizer. Early stopping was implemented with a patience of 3 epochs based on validation loss. Random seeds were fixed at 42 for reproducibility. The more detailed parameters are listed in Table 1. For the CARE-ET framework, the GAT neighbor selection threshold was set to 0.6, and the joint adjustment threshold to 0.8. An 80/10/10 split was used for training, validation, and testing. To facilitate reproducibility, the code is provided as supplementary material for review and will be made public upon acceptance. The more detailed parameters are listed in Table 2.

XSL•FO
RenderX

Table . Training hyperparameters.

Parameter	Value
Framework	PyTorch 1.12 (PyTorch Foundation)
Hardware	NVIDIA Tesla T4 GPU (16GB VRAM)
Base model	ALBERT[a]-base-v2
Fine-tuning epochs	10
Batch size	16
Learning rate	2.00×10
Weight decay	0.01 (AdamW)
Optimizer	AdamW
Dropout rate	0.5
Early stopping	Patience=3 epochs
Random seed	42
Dataset split	80% / 10% / 10% (Train/Val/Test)

[a]ALBERT: A Lite Bidirectional Encoder Representations from Transformers for Self-supervised Learning of Language Representations.

Table . CARE-ET (CARE event timeline) and timeline generation parameters.

Parameter	Symbol and description	Value
GAT[a] neighbor selection threshold	Ni (selection)	0.6
Joint adjustment threshold	Ii (second-order neighbor)	0.8
Structural loss weight	λ1 (structural accuracy)	0.7
Global coherence weight	λ2 (consistency)	0.3
Temporal anchor similarity	Cosine threshold for history	0.85
Confidence flag range	For human-in-the-loop verification	[0.70, 0.85]
Coreference resolution window size	Sliding window for ALBERT[b]	512
Coreference resolution threshold	Cosine distance threshold	0.8

[a]GAT: graphic attention layer.

[b]ALBERT: A Lite Bidirectional Encoder Representations from Transformers for Self-supervised Learning of Language Representations.

The CARE-ET model was trained on a dataset of 3142 annotated narrative segments derived from 118 sessions. The annotation workflow involved 2 gerontology specialists who independently labeled event triggers and temporal relations; interannotator agreement reached a Cohen κ of 0.77. For model evaluation, we used a 5-fold cross-validation strategy to ensure the robustness of the results across different storytelling styles.

To ensure the statistical rigor of our results, we conducted bootstrap significance testing (n=1000) and used the paired 2-tailed *t* test to compare CARE-ET against the strongest baseline. Statistical significance is $P<.05$.

Event Summary Assessment

Life story event summarization enables caregivers to efficiently retrieve relevant narratives while providing the structural foundation for organizing fragmented life stories, with evaluation conducted using Recall-Oriented Understudy for Gisting Evaluation (ROUGE) metrics [33]. ROUGE is a commonly used evaluation method for automated text summarization, which measures the quality of summaries by calculating the overlap between automated summaries and human-referenced summaries for unigram words (ROUGE-1), bigram words (ROUGE-2), and the longest common subsequence (ROUGE-L). The method is simple, intuitive, highly consistent with manual evaluation results, and is one of the most effective evaluation metrics for current text summarization tasks.

For automated summarization evaluation, we report the F_1-scores of ROUGE-1, ROUGE-2, and ROUGE-L. All metrics were computed using the Python *rouge-score* package with Porter stemming enabled and standard English stopwords removed. To ensure the reliability of performance gains, we conducted bootstrap significance testing (n=1000) to report 95% CIs and *P* values against the baselines.

Table 3 and Table 4 present the experimental evaluations on both the OALS and LCSTS datasets [34], demonstrating CARE-ET's superior performance in event summarization tasks compared to several baseline methods (evolutionary timeline summarization, Timeline-sumy, temporally sensitive submodularity framework for timeline summarization, crisis

XSL•FO
RenderX

local timeline line-up summarization, pretraining with extracted gap-sentences for abstractive summarization, LongT5, and pyramid-based masked sentence pretraining for multidocument summarization) [23-25,35-38]. Using median and mean text lengths as input parameters (58/210 bytes for OALS; 234/248 bytes for LCSTS), CARE-ET consistently achieved higher accuracy across all test conditions, with particularly notable advantages in full-length input scenarios, showing 0.6% (SD 1.20%) and 1.6% (SD 0.69%) average improvements over the strongest baseline in OALS and LCSTS, respectively. These results validate the model's effectiveness in extracting structured event summaries from both domain-specific life story data (OALS) and general narrative texts (LCSTS), confirming its robust generalization capabilities while maintaining processing accuracy across variable input lengths.

Table . Performance on the OALS (Older Adults' Life Stories) dataset using different lengths.

Method	Length limited within 60 bytes (%), mean (SD)			Length limited within 220 bytes (%), mean (SD)			Full length (%), mean (SD)		
	ROUGE-1[a]	ROUGE-2[b]	ROUGE-L[c]	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-1	ROUGE-2	ROUGE-L
ETS[d]	26.4 (1.15)	13.9 (0.72)	19.5 (0.98)	31.1 (1.28)	14.3 (0.65)	26.3 (1.05)	34.9 (1.42)	15.1 (0.74)	30.1 (1.25)
Timeline-Sumy	26.4 (1.08)	13.9 (0.68)	19.5 (0.91)	31.1 (1.21)	14.3 (0.61)	26.3 (0.98)	34.9 (1.35)	15.1 (0.71)	30.1 (1.18)
TSSF-TLS[e]	26.9 (1.02)	15.5 (0.58)	19.9 (0.85)	31.3 (1.15)	15.9 (0.55)	26.9 (0.92)	35.8 (1.28)	15.9 (0.65)	30.8 (1.12)
CrisisLTL-Sum[f]	27.5 (0.95)	16.9 (0.62)	20.9 (0.78)	32.1 (1.08)	17.7 (0.52)	27.5 (0.88)	37.2 (1.15)	16.9 (0.68)	31.6 (1.05)
PEGASUS[g]	27.3 (0.91)	17.2 (0.55)	21.5 (0.82)	32.5 (1.02)	18.1 (0.48)	27.9 (0.85)	37.4 (1.12)	17.2 (0.58)	31.9 (0.98)
LongT5[h]	27.8 (0.88)	17.1 (0.52)	21.3 (0.75)	32.4 (0.98)	17.9 (0.45)	28.1 (0.82)	37.7 (1.08)	17.1 (0.55)	32.1 (0.95)
PRIMERA[i]	28.1 (0.82)	17.1 (0.48)	21.7 (0.68)	32.6 (0.92)	17.9 (0.42)	28.2 (0.78)	37.6 (0.98)	17.9 (0.52)	32.6 (0.88)
CARE-ET[j] (Word2Vec)	27.9 (0.85)	17.6 (0.51)	21.9 (0.72)	32.7 (0.95)	18.2 (0.45)	28.4 (0.81)	37.9 (1.02)	16.9 (0.55)	32.4 (0.91)
CARE-ET (LSTM[k])	28.5 (0.78)	17.7 (0.45)	21.7 (0.65)	33.1 (0.88)	18.6 (0.38)	28.7 (0.75)	38.1 (0.95)	17.4 (0.48)	32.8 (0.85)
CARE-ET	28.3 (0.72)[l]	17.5 (0.42)[l]	22.1 (0.61)[l]	33.2 (0.85)[l]	18.8 (0.35)[l]	28.9 (0.72)[l]	38.5 (0.92)[l]	17.7 (0.45)[l]	33.2 (0.82)[l]

[a]ROUGE-1: Recall-Oriented Understudy for Gisting Evaluation, variant 1.

[b]ROUGE-2: Recall-Oriented Understudy for Gisting Evaluation, variant 2.

[c]ROUGE-L: Recall-Oriented Understudy for Gisting Evaluation–Longest Common Subsequence.

[d]ETS: evolutionary timeline summarization.

[e]TSSF-TLS: temporally sensitive submodularity framework for timelines summarization.

[f]CrisisLTLSum: crisis local timeline line-up summarization.

[g]PEGASUS: pretraining with extracted gap-sentences for abstractive summarization.

[h]LongT5: long text-to-text transfer transformer.

[i]PRIMERA: pyramid-based masked sentence pretraining for multidocument summarization.

[j]CARE-ET: CARE event timeline.

[k]LSTM: long short-term memory.

[l]Statistically significant, $P<.05$.

XSL•FO
RenderX

Table . Performance on the LCSTS (Large Scale Chinese Short Text Summarization) dataset using different lengths.

Method	Length limited within 234 bytes (%), mean (SD)			Length limited within 248 bytes (%), mean (SD)			Full length (%), mean (SD)		
	ROUGE-1[a]	ROUGE-2[b]	ROUGE-L[c]	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-1	ROUGE-2	ROUGE-L
ETS[d]	25.4 (0.72)	12.8 (0.35)	22.1 (0.52)	27.2 (0.78)	14.1 (0.38)	24.5 (0.61)	29.1 (0.82)	16.3 (0.45)	26.2 (0.68)
Timeline-Sumy	25.5 (0.68)	12.9 (0.32)	22.2 (0.48)	27.3 (0.74)	14.2 (0.35)	24.6 (0.55)	29.2 (0.78)	16.4 (0.42)	26.3 (0.65)
TSSF-TLS[e]	26.4 (0.65)	14.1 (0.28)	23.3 (0.45)	28.5 (0.71)	15.6 (0.32)	25.7 (0.52)	30.2 (0.75)	17.5 (0.41)	27.4 (0.62)
CrisisLTL-Sum[f]	29.1 (0.58)	16.5 (0.25)	26.2 (0.38)	31.4 (0.62)	18.2 (0.30)	28.3 (0.48)	33.1 (0.68)	20.2 (0.38)	30.1 (0.58)
PEGASUS[g]	34.2 (0.55)	21.3 (0.32)	31.1 (0.42)	36.5 (0.65)	23.1 (0.38)	33.2 (0.51)	38.1 (0.71)	24.9 (0.42)	34.8 (0.61)
LongT5[h]	34.6 (0.51)	21.6 (0.28)	31.5 (0.38)	36.8 (0.58)	23.4 (0.32)	33.7 (0.48)	38.5 (0.65)	25.2 (0.38)	35.3 (0.55)
PRIMERA[i]	34.5 (0.48)	21.8 (0.25)	31.2 (0.35)	36.7 (0.55)	23.6 (0.28)	33.5 (0.45)	38.4 (0.62)	25.4 (0.35)	35.1 (0.52)
CARE-ET[j] (Word2Vec)	34.9 (0.45)	22.1 (0.22)	31.8 (0.32)	37.1 (0.52)	23.9 (0.25)	33.9 (0.42)	38.9 (0.58)	25.8 (0.32)	35.8 (0.48)
CARE-ET (LSTM[k])	35.3 (0.42)	22.4 (0.21)	32.2 (0.28)	37.4 (0.48)	24.2 (0.22)	34.2 (0.38)	39.2 (0.55)	26.1 (0.31)	36.2 (0.45)
CARE-ET	35.7 (0.38)[l]	22.7 (0.18)[l]	32.8 (0.25)[l]	37.8 (0.42)[l]	24.5 (0.21)[l]	34.8 (0.35)[l]	39.5 (0.52)[l]	26.4 (0.28)[l]	36.9 (0.42)[l]

[a]ROUGE-1: Recall-Oriented Understudy for Gisting Evaluation, variant 1.

[b]ROUGE-2: Recall-Oriented Understudy for Gisting Evaluation, variant 2.

[c]ROUGE-L: Recall-Oriented Understudy for Gisting Evaluation–Longest Common Subsequence (ROUGE-L).

[d]ETS: evolutionary timeline summarization.

[e]TSSF-TLS: temporally sensitive submodularity framework for timeline summarization.

[f]CrisisLTLSum: crisis local timeline line-up summarization.

[g]PEGASUS: pretraining with extracted gap-sentences for abstractive summarization.

[h]LongT5: long text-to-text transfer transformer.

[i]PRIMERA: pyramid-based masked sentence pretraining for multidocument summarization.

[j]CARE-ET: CARE event timeline.

[k]LSTM: long short-term memory.

[l]Statistically significant, $P<.05$.

To evaluate the specific contributions of each component, we conducted a rigorous ablation study. Table 5 indicates that replacing ALBERT with Word2Vec leads to a 1% drop in F_1-score, highlighting the necessity of contextual embeddings. Furthermore, the GAT-DBN synergy outperforms GAT-only and DBN-only configurations by 2.8% and 3.4%, respectively. The self-supervised module provides an additional 0.9% gain in timeline consistency. Our model contains 64M parameters, and the average training time per fold is approximately 4.5 (SD 0.34) hours on a Tesla T4 GPU, representing a balanced tradeoff between complexity and performance.

XSL•FO
RenderX

Table . Ablation experiment results.

Configuration	ROUGE-1[a]	ROUGE-2[b]	ROUGE-L[c]	SLEU[d]
CARE-ET[e] (Full)	*38.5* [f]	*17.7*	*33.2*	*0.72*
Without ALBERT[g] (Word2Vec)	37.4	16.8	32.2	0.69
Without BiLSTM[h] (LSTM[i])	37.3	16.5	32.0	0.68
Without DBN[j] (GAT[k] only)	35.7	15.2	30.4	0.61
Without GAT (DBN only)	35.1	14.8	29.8	0.58
Without self-supervised module	37.6	17.2	32.3	0.65
Without all (baseline Seq2Seq)	34.2	13.5	28.5	0.52

[a]ROUGE-1: Recall-Oriented Understudy for Gisting Evaluation, variant 1.

[b]ROUGE-2: Recall-Oriented Understudy for Gisting Evaluation, variant 2.

[c]ROUGE-L: Recall-Oriented Understudy for Gisting Evaluation–Longest Common Subsequence.

[d]SLEU: sequence link evaluation.

[e]CARE-ET: CARE event timeline.

[f]Italicized numbers indicate the best result among all experimental results.

[g]ALBERT: A Lite Bidirectional Encoder Representations from Transformers for Self-supervised Learning of Language Representations.

[h]BiLSTM: bidirectional long short-term memory.

[i]LSTM: long short-term memory.

[j]DBN: deep belief network.

[k]GAT: graphic attention layer.

Qualitative Analysis

To further understand the performance of CARE-ET, we present 2 anonymized case studies comparing the generated timelines against the therapeutic ground truth.

Case 1 (Success in Long-Range Dependencies)

During an interview, the participant mentioned family-related milestones (eg, marriage and grandchildren) and their prime years in a later segment. CARE-ET effectively captured the cross-segment dependencies, chronologically positioning these events after related life stages while maintaining the semantic links between the memories.

Case 2 (Failure in Near-Duplicate Events)

During a session discussing a wedding, the participant mentioned the date and event details in different segments. The model failed to merge these, creating two independent nodes for the same event (a redundancy error). This failure suggests that current GAT-based relevance thresholding remains insufficient for fine-grained deduplication in repetitive, fragmented oral testimonies.

Quality Assessment of Timelines

In this paper, sequence link evaluation (SLEU) metrics and TimID metrics are used to assess the quality of event timelines.

Sequence Link Evaluation

In this paper, the evaluation metrics of SLEU are used to verify the quality of the timeline [22]. The sentences in the system-generated storyline are ^{1}, ^{2},… , where *s*1 is the beginning of the story and is the end of the story. The sentences of the reference storyline are $'_{1}$, $'_{2}$, … $'$. n is the number of stages, and h-gram and h-gram′ denote a sequence of h sentences in a storyline, not necessarily consecutive, generated by the reference storyline and the generated storyline. The collocation of information between sentences in a storyline is utilized to compute the probability of the occurrence of the storyline and thus determine whether the storyline is complete or coherent.

The cosine similarity between (∈ [1,]) and (∈ [1,]) is first calculated. Then the probability of h-gram matching between the 2 storylines is calculated separately:

$$(22)\ ph=\min(\mathrm{Countmatch}(h\text{-}gram),\mathrm{Countmatch}(h\text{-}gram'))\mathrm{Count}(h\text{-}gram)$$

The possible values of the variable h depend on the specific case and are represented by the set H. The set H is a set of 1-grams. Storylines using 1-grams tend to satisfy the synthesis requirement more easily. Longer h-gram matches indicate better coherence. The count (h-gram) denotes the number of comparable h-grams (eg, if h=4, count(2-gram)=6). The count(hgram’) equals count(h-gram). The count_{match}(h-gram) denotes the number of h-grams that can be matched (ie, h-grams against h-grams). The count_{match}(h-gram) denotes the number of h-grams that can be matched (ie, the number of h-gram pairs for which the similarity of the corresponding sentences exceeds a specific threshold λ). In the experiments, we set λ=0.2. Then SLEU = (∑ ∈ log) is computed. The SLEU metric value ranges from 0 to 1 and is used to quantitatively assess the quality of the storyline.

XSL•FO
RenderX

We standardized the timeline consistency metric as SLEU. It evaluates the structural integrity of the generated timeline by calculating the weighted sum of correct temporal transitions:

$$(23)\ SLEU=\sum h=1Hwh\cdot \exp(-\lambda\cdot |pospred-posref|)$$

where H=3 denotes the maximum historical hop distance considered, wh represents the weight assigned to the h-th hop (set as 1/h), and λ=0.5 is the decay constant for distance penalties.

TimID

The study introduces TimID, a temporal annotation marker embedded within a novel life story dataset for older adults, enabling precise evaluation of event chronology through 2 complementary dimensions. Absolute position accuracy quantifies deviations between generated event nodes and their corresponding TimID reference points, where smaller positional differences indicate higher timeline fidelity. Simultaneously, relative order preservation assesses the correctness of pairwise event sequences by verifying their alignment with ground-truth temporal relationships. This dual-metric approach ensures comprehensive validation of both event placement precision and chronological logic in life story reconstruction.

TimID is defined to measure the precision of event placement. The "absolute position" of an event is defined as its normalized rank ri=i/N, where N is the total number of events. The relative order preservation is computed using the Kendall τ coefficient, which quantifies the ordinal correlation between the predicted sequence and the gold-standard therapeutic timeline.

Experimental Results Based on SLEU and TimID

The study uses 2 complementary assessment frameworks: SLEU evaluates storyline coherence through sentence-level probability analysis, while TimID assesses timeline quality via absolute position accuracy and relative order consistency. As shown in Table 6, CARE-ET outperforms baseline methods across both metrics, achieving superior SLEU scores (0.021 - 0.061 higher) and TimID performance (0.08% - 0.24%% better relative order accuracy, with 0.8 - 1.7 lower positional deviations). These results demonstrate significant improvements in both narrative flow and temporal precision.

Table . Performance on the dataset OALS (Older Adults' Life Stories) using sequence link evaluation (SLEU) and timeline identity (TimID).

Method	SLEU	TimID-ap[a]	TimID-rp[b]
ETS[c]	0.47 (0.02)	4.51 (0.38)	0.43 (0.03)
Timeline-Sumy	0.48 (0.02)	4.12 (0.25)	0.51 (0.03)
TSSF-TLS[d]	0.47 (0.02)	4.21 (0.31)	0.48 (0.03)
CrisisLTLSum[e]	0.51 (0.02)	3.64 (0.18)	0.59 (0.02)
CARE-ET[f]	0.53 (0.01)[g]	2.79 (0.15)[g]	0.67 (0.02)[g]

[a]TimID-ap: timeline identity–absolute position.

[b]TimID-rp: timeline identity–relatable position.

[c]ETS: evolutionary timeline summarization.

[d]TSSF-TLS: temporally sensitive submodularity framework for timeline summarization.

[e]CrisisLTLSum: crisis local timeline line-up summarization.

[f]CARE-ET: CARE event timeline.

[g]Statistically significant, $P<.05$.

The metrics validate CARE-ET's dual capacity to generate clinically meaningful timelines: SLEU confirms enhanced narrative coherence, which is crucial for therapeutic applications, while TimID verifies improved chronological accuracy that aligns with human cognition. This combination of strong textual fluency and precise event sequencing enhances both the interpretability of life stories for caregivers and their therapeutic use in reminiscence interventions, advancing digital health solutions for older adult care.

System Usability Assessment

This study has established cooperative relationships with the Hefei Tianyu Elderly Care Institution and the Chuzhou Lekang Elderly Care Institution. The system has been deployed within the daily care management platforms of both institutions, serving a total of 36 older adult users (21 at the Tianyu Institution and 15 at the Lekang Institution). For the optimized Story Mosaic system, we invited 4 caregivers and 1 manager from both the Tianyu and Lekang senior care facilities to complete the System Usability Scale (SUS) [39]. The statistical information for the 10 participants is shown in Table 7. Four participants (A1, C2, C3, and C4) had previously participated in the feasibility study of our published Story Mosaic system, while the other 6 were newly recruited for this study. The participants included 2 core roles—administrators and caregivers—with balanced gender representation. Aged 35 to 62 years, they represented young, middle-aged, and senior practitioner cohorts, while their elder care tenure (4 - 23 y) reflected experiential disparities among novice, proficient, and senior workers. Their educational qualifications showed a gradient distribution, ranging from no formal education to master's degrees. Participant selection ensured diversity across roles, sex, age, tenure, education, and affiliated institutions. This sampling strategy represented the target user groups and provided empirical support for the scientific rigor and objectivity of the system usability evaluation.

XSL•FO
RenderX

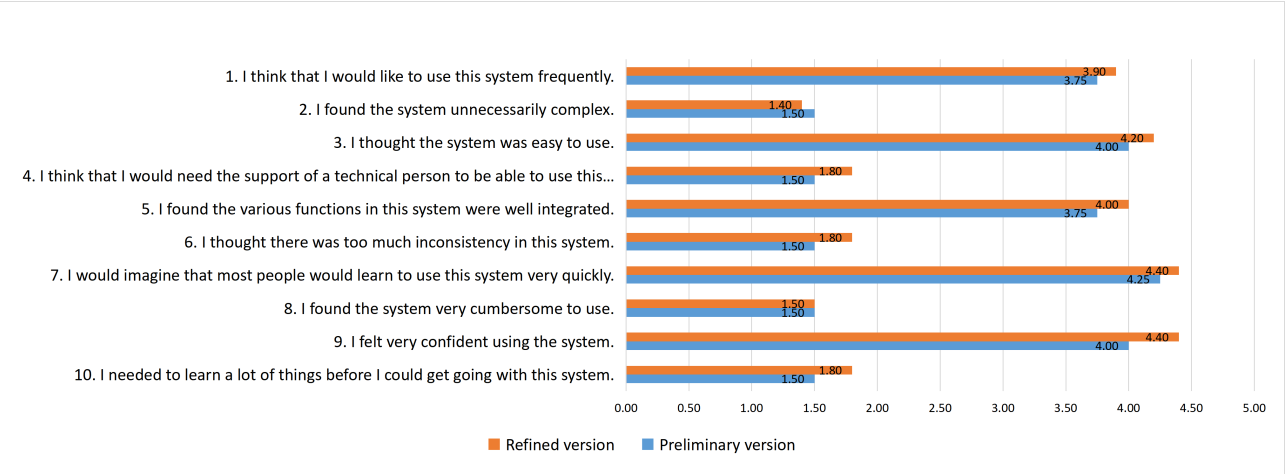
Table . Demographic characteristics of participants.

ID	Identity	Sex	Age (y)	Tenure	Education	Affiliated institution
A1	Administrator	Male	37	13	Master's degree	Tianyu
A2	Administrator	Female	52	23	Junior high school	Lekang
C1	Caregiver	Female	56	10	Primary school	Tianyu
C2	Caregiver	Female	48	6	Junior high school	Tianyu
C3	Caregiver	Male	36	5	Junior high school	Tianyu
C4	Caregiver	Female	57	8	Primary school	Tianyu
C5	Caregiver	Male	35	4	Bachelor's degree	Lekang
C6	Caregiver	Female	45	9	Junior high school	Lekang
C7	Caregiver	Female	62	14	Null[a]	Lekang
C8	Caregiver	Female	58	8	Primary school	Lekang

[a]Null: no education has been received.

In the preliminary work of the Story Mosaic system, the average SUS score was 82.63, achieving an A– grade. After enhancing the system based on the CARE-ET algorithm, a usability test was conducted again, and the evaluation results are shown in Figure 7 with SUS scores from both the preliminary and refined versions.

Figure 7. System Usability Scale scores from the preliminary and refined versions.



A comparison of the refined and preliminary versions of the SUS scale shows the following: for positively worded items (1, 3, 5, 7, 9), the refined version scored higher, reflecting more positive evaluations of the system's usability. For negatively worded items (2, 4, 6, 8, 10), the refined version still had relatively high scores, indicating that users perceived some remaining operational complexity. This is mainly because the refined version's evaluators included caregivers with primary education or no formal education. While these users gave high scores on item 7 ("Most people can learn to use this system quickly"), their actual onboarding speed was slower. Moving forward, it will be necessary to simplify processes and add basic guidance for this group to improve their efficiency.

In addition to the above questions, we also asked an open-ended question: "What other ways do you think the timelines can help you care for older adults"? One participant noted:

> *Timelines are critical in the care of older adults with cognitive impairment. These older adults' behavioral and psychological symptoms are closely related to their past experiences. Timelines help us understand their experiences and act appropriately to the situation.*

Objective Comparisons

To objectively evaluate the impact of the Story Mosaic system on cognitive load, we conducted a comparative study between the artificial intelligence–assisted workflow and a traditional manual workflow. We used the NASA Task Load Index, which assesses cognitive demand across 6 dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration [40]. The comparative evaluation results are shown in Table 8. Focusing on the 3 core dimensions of mental demand, temporal demand, and physical effort prioritized by participants, the Story Mosaic system achieved load reductions of 37.5%, 36.4%, and 42.9%, respectively, in these metrics. It precisely addressed users' key needs by effectively alleviating cognitive pressure, time constraints, and energy input. Meanwhile, the system also optimized loads across other

XSL•FO
RenderX

dimensions, cutting the overall workload by 37.4% and demonstrating excellent human-machine compatibility.

Table . NASA Task Load Index (NASA-TLX) evaluation of the manual workflow and the Story Mosaic system.

Metric	Manual workflow, mean (SD)	Story Mosaic system, mean (SD)	Relative reduction (%)
Mental demand	76.0 (13.1)	47.5 (23.6)	37.5
Physical demand	27.0 (7.50)	25.0 (8.5)	7.4
Temporal demand	77.0 (10.1)	49.0 (19.3)	36.4
Performance	43.0 (16.7)	31.0 (16.3)	27.9
Effort	73.5 (13.8)	42.0 (15.1)	42.9
Frustration	62.5 (14.4)	34.0 (14.3)	45.6
Overall workload	65.7 (13.6)	41.2 (16.9)	37.4

Discussion

Principal Results

The GAT framework incorporates dynamic node filtering using multilevel semantic similarity thresholds, automatically preserving clinically meaningful life event relationships. This adaptive graph optimization enhances event relationship modeling precision while maintaining computational efficiency for real-world care applications.

CARE-ET demonstrates robust performance across both specialized life story datasets and general narratives through its novel integration of event triggers, spatiotemporal markers, and contextual encoding. The system outperforms conventional methods in extracting key life events from older adult narratives, providing a solid foundation for timeline construction.

Rigorous evaluation confirms that the generated timelines meet clinical standards for both narrative coherence and temporal accuracy. Caregiver assessments highlight superior readability and content generalization, validating the system's ability to transform fragmented life stories into structured therapeutic materials for cognitive reminiscence therapy.

Limitations

While CARE-ET demonstrates promising results, several limitations warrant consideration. First, the method's performance relies heavily on the quality of temporal annotations (TimID) in the training data, which may limit generalizability to life stories with sparse or ambiguous time references. Second, the current implementation focuses primarily on textual narratives, potentially underutilizing multimodal cues (eg, emotional valence in vocal tones or visual context) that could enhance event interpretation. Third, the evaluation involved relatively small-scale caregiver assessments (n=10), necessitating broader clinical validation to confirm real-world utility across diverse older adult populations and care settings.

Furthermore, the collection and processing of older adult individuals' personal life stories involve highly sensitive data with significant privacy implications. Although this study followed ethical protocols including approval, informed consent, and data anonymization, several challenges persist. These include ensuring long-term data security, preserving participants' ongoing control over their narratives, and addressing potential emotional impacts from recalling personal memories—each representing key ethical considerations in this research domain.

Addressing these current limitations, future work will focus on advancing research in the following areas: First, the development of weakly supervised temporal annotation enhancement algorithms will leverage cross-modal temporal reasoning and the mining of implicit temporal cues to improve the model's robustness in parsing life stories with sparse temporal information. Second, the construction of a multimodal fusion framework that integrates speech emotion analysis, visual scene understanding, and textual semantic representation to establish a cross-modally aligned paradigm for life event comprehension. Third, collaboration with more elder care institutions to evaluate the long-term use of the system in real-world care scenarios.

Conclusions

This study proposes the CARE-ET method, which enables the structured extraction of core events from life stories through multimodal event element fusion–based summary generation techniques and improves the accuracy of event relevance calculation via a collaborative computation model of GAT and DBN. Experiments on the OALS and LCSTS datasets show that CARE-ET significantly reduces redundancy and disorder in narrative texts, generating event timelines that outperform baseline methods in readability and structural rationality. This study provides an effective technical solution for the digital presentation of older adult life stories, with important application value for enhancing the personalization and precision of older adult care services.

Funding

This work is supported by the Scientific Research Initiation Project (2024qd12 and 2025qd10) and the Anhui Provincial Department of Education Project (2025AHGXZK40376 and 2025AHGXZK40409).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Ethical approval of the dataset, data sources, construction basis, construction process, sample size, recruitment criteria, annotation specifications and processes for event elements and timelines, dataset public statement, etc.
[DOCX File, 56 KB - aging_v9i1e83122_app1.docx]

References

1. Morley JE. An overview of cognitive impairment. Clin Geriatr Med 2018 Nov;34(4):505-513. [doi: 10.1016/j.cger.2018.06.003] [Medline: 30336985]
2. Xu L, Fields NL, Daniel KM, Cipher DJ, Troutman BA. Reminiscence and digital storytelling to improve the social and emotional well-being of older adults with Alzheimer's disease and related dementias: protocol for a mixed methods study design and a randomized controlled trial. JMIR Res Protoc 2023 Sep 7;12:e49752. [doi: 10.2196/49752] [Medline: 37676706]
3. Piolatto M, Bianchi F, Rota M, Marengoni A, Akbaritabar A, Squazzoni F. The effect of social relationships on cognitive decline in older adults: an updated systematic review and meta-analysis of longitudinal cohort studies. BMC Public Health 2022 Feb 11;22(1):278. [doi: 10.1186/s12889-022-12567-5] [Medline: 35148704]
4. Barsuglia JP, Nedjat-Haiem FR, Shapira JS, et al. Observational themes of social behavioral disturbances in frontotemporal dementia. Int Psychogeriatr 2014 Sep;26(9):1475-1481. [doi: 10.1017/S104161021400091X] [Medline: 24846824]
5. Chehrehnegar N, Shati M, Esmaeili M, Foroughan M. Executive function deficits in mild cognitive impairment: evidence from saccade tasks. Aging Ment Health 2022 May;26(5):1001-1009. [doi: 10.1080/13607863.2021.1913471] [Medline: 33928806]
6. Che J, Cheng N, Jiang B, et al. Executive function measures of participants with mild cognitive impairment: systematic review and meta-analysis of event-related potential studies. Int J Psychophysiol 2024 Mar;197:112295. [doi: 10.1016/j.ijpsycho.2023.112295] [Medline: 38266685]
7. Zhang Y, Zhu X, Yang S, Wong AKC, Chen X. Digital health technologies applied in patients with early cognitive change: scoping review. J Med Internet Res 2025 Dec 29;27(1):e82881. [doi: 10.2196/82881] [Medline: 41461066]
8. Sommerlad A, Kivimäki M, Larson EB, et al. Social participation and risk of developing dementia. Nat Aging 2023 May;3(5):532-545. [doi: 10.1038/s43587-023-00387-0] [Medline: 37202513]
9. Gucciardi E, Jean-Pierre N, Karam G, Sidani S. Designing and delivering facilitated storytelling interventions for chronic disease self-management: a scoping review. BMC Health Serv Res 2016 Jul 11;16(1):249. [doi: 10.1186/s12913-016-1474-7] [Medline: 27401836]
10. Sljivic H, Sutherland I, Stannard C, Ioppolo C, Morrisby C. Changing attitudes towards older adults: eliciting empathy through digital storytelling. Gerontol Geriatr Educ 2022;43(3):360-373. [doi: 10.1080/02701960.2021.1900838] [Medline: 35837695]
11. Stargatt J, Bhar S, Bhowmik J, Al Mahmud A. Digital storytelling for health-related outcomes in older adults: systematic review. J Med Internet Res 2022 Jan 12;24(1):e28113. [doi: 10.2196/28113] [Medline: 35019845]
12. Rios Rincon AM, Miguel Cruz A, Daum C, Neubauer N, Comeau A, Liu L. Digital storytelling in older adults with typical aging, and with mild cognitive impairment or dementia: a systematic literature review. J Appl Gerontol 2022 Mar;41(3):867-880. [doi: 10.1177/07334648211015456] [Medline: 34009053]
13. Zhu D, Al Mahmud A, Liu W. Digital storytelling intervention for enhancing the social participation of people with mild cognitive impairment: co-design and usability study. JMIR Aging 2024 Jan 17;7(1):e54138. [doi: 10.2196/54138] [Medline: 38231541]
14. Zhu D, Al Mahmud A, Liu W. Effects of a mobile storytelling app (Huiyou) on social participation among people with mild cognitive impairment: pilot randomized controlled trial. JMIR Hum Factors 2025 Jun 18;12(1):e70177. [doi: 10.2196/70177] [Medline: 40532206]
15. Ferstad K, Rykkje L. Understanding the significance of listening to older people's life stories in whole person care-an interview study of nurses in gerontology. SAGE Open Nurs 2023;9:23779608231164077. [doi: 10.1177/23779608231164077] [Medline: 36949851]
16. An N, Gui F, Jin L, Ming H, Yang J. Toward better understanding older adults: a biography brief timeline extraction approach. Int J Hum Comput Interact 2023 Mar 16;39(5):1084-1095. [doi: 10.1080/10447318.2022.2077278]
17. Gui F, Yang J, Wu Q, Liu Y, Zhou J, An N. Enhancing caregiver empowerment through the Story Mosaic system: human-centered design approach for visualizing older adult life stories. JMIR Aging 2023 Nov 8;6:e50037. [doi: 10.2196/50037] [Medline: 37962517]
18. Valtolina S, Pugliese A. AI-assisted cognitive support for caregivers: a RAG and EUD framework for geriatric care. In: Santoro C, Schmidt A, Matera M, Bellucci A, editors. End-User Development: 10th International Symposium, IS-EUD 2025, Munich, Germany, June 16–18, 2025, Proceedings 2025:205-220. [doi: 10.1007/978-3-031-95452-8_13]

XSL•FO
RenderX

19. Almeida PV, Rocha RP. AI-powered social robot for cognitive stimulation in the elderly through personalized storytelling. 2025 Presented at: 2025 11th International Conference on Automation, Robotics, and Applications (ICARA); Feb 12-14, 2025; Zagreb, Croatia p. 122-126. [doi: 10.1109/ICARA64554.2025.10977621]
20. Li J, Cardie C. Timeline generation: tracking individuals on twitter. Presented at: WWW '14: 23rd International World Wide Web Conference; Apr 7-11, 2014; Seoul, Korea p. 643-652. [doi: 10.1145/2566486.2567969]
21. Althoff T, Dong XL, Murphy K, Alai S, Dang V, Zhang W. TimeMachine: timeline generation for knowledge-base entities. Presented at: KDD '15: The 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 10-13, 2015; Sydney, Australia p. 19-28. [doi: 10.1145/2783258.2783325]
22. Sun W, Wang Y, Gao Y, Li Z, Sang J, Yu J. Comprehensive event storyline generation from microblogs. 2019 Dec 15 Presented at: Proceedings of the 1st ACM International Conference on Multimedia in Asia; Dec 16-18, 2019. [doi: 10.1145/3338533.3366601]
23. Zhang J, Zhao Y, Saleh M, Liu P. PEGASUS: pre-training with extracted gap-sentences for abstractive summarization. 2020 Presented at: Proceedings of the 37th International Conference on Machine Learning PMLR; Jul 13-18, 2020 p. 11328-11339 URL: https://proceedings.mlr.press/v119/zhang20ae/zhang20ae.pdf [accessed 2026-03-10]
24. Guo M, Ainslie J, Uthus D, et al. LongT5: efficient text-to-text transformer for long sequences. 2022 Presented at: Findings of the Association for Computational Linguistics; Jul 10-15, 2022; Seattle, United States p. 724-736. [doi: 10.18653/v1/2022.findings-naacl.55]
25. Xiao W, Beltagy I, Carenini G, Cohan A. PRIMERA: pyramid-based masked sentence pre-training for multi-document summarization. 2022 Presented at: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers); May 22-27, 2022; Dublin, Ireland p. 5245-5263. [doi: 10.18653/v1/2022.acl-long.360]
26. Lan Z, Chen M, Goodman S, Gimpel K, Sharma P, Soricut R. ALBERT: a lite BERT for self-supervised learning of language representations. Presented at: 8th International Conference on Learning Representations; Apr 26-30, 2020 URL: https://openreview.net/pdf?id=H1eA7AEtvS [accessed 2026-03-10]
27. Che W, Feng Y, Qin L, Liu T. N-LTP: an open-source neural language technology platform for Chinese. 2021 Presented at: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing; Nov 7-11, 2021; Online and Punta Cana, Dominican Republic p. 42-49. [doi: 10.18653/v1/2021.emnlp-demo.6]
28. Liu X, Huang H, Zhang Y. Open domain event extraction using neural latent variable models. 2019 Presented at: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics; Jul 28 to Aug 2, 2019; Florence, Italy p. 2860-2871. [doi: 10.18653/v1/P19-1276]
29. Suo W, Muhammad BA, Zhang Z, et al. GAT-LS: a graph attention network for learning style detection in online learning environment. J Netw Netw Appl 2024;4(2):60-72. [doi: 10.33969/J-NaNA.2024.040202]
30. Singh NA, Kishore K, Deo RN, Lu Y, Kodikara J. Automated segmentation framework for asphalt layer thickness from GPR data using a cascaded k-means–DBSCAN algorithm. J Environ Eng Geophys 2022 Dec 1;27(4):179-189. [doi: 10.32389/JEEG22-019]
31. Velickovic P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y. Graph attention networks. 2017 Presented at: International Conference on Learning Representations (ICLR 2018); Apr 30 to May 3, 2017 URL: https://openreview.net/pdf?id=rJXMpikCZ [accessed 2026-03-10]
32. Li J, Sun Q, Zhang F, Yang B. Meta-structure-based graph attention networks. Neural Netw 2024 Mar;171:362-373. [doi: 10.1016/j.neunet.2023.12.025] [Medline: 38134599]
33. Lin CY. Rouge: a package for automatic evaluation of summaries text summarization branches out. 2004 Presented at: Text Summarization Branches Out: Proceedings of the ACL-04 Workshop; Jul 25-27, 2004; Barcelona, Spain p. 74-81 URL: https://aclanthology.org/W04-1013.pdf [accessed 2026-03-10]
34. Hu B, Chen Q, Zhu F. LCSTS: a large scale chinese short text summarization dataset. Presented at: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015); Sep 17-21, 2015; Lisbon, Portugal p. 1967-1972. [doi: 10.18653/v1/D15-1229]
35. Yan R, Wan X, Otterbacher J, Kong L, Li X, Zhang Y. Evolutionary timeline summarization: a balanced optimization framework via iterative substitution. Presented at: SIGIR '11: The 34th International ACM SIGIR Conference on Research and Development in Information Retrieval; Jul 24-28, 2011; Beijing, China p. 745-754. [doi: 10.1145/2009916.2010016]
36. Chang Y, Tang J, Yin D, Yamada M, Liu Y. Timeline summarization from social media with life cycle models. Presented at: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16); Jul 9-15, 2016; New York, NY p. 3698-3704 URL: https://www.ijcai.org/Proceedings/16/Papers/520.pdf [accessed 2026-03-30]
37. Martschat S, Markert K. A temporally sensitive submodularity framework for timeline summarization. 2018 Presented at: Proceedings of the 22nd Conference on Computational Natural Language Learning; Nov 1-2, 2018. [doi: 10.18653/v1/K18-1023]
38. Rajaby Faghihi H, Alhafni B, Zhang K, Ran S, Tetreault J, Jaimes A. CrisisLTLSum: a benchmark for local crisis event timeline extraction and summarization. Presented at: Findings of the Association for Computational Linguistics: EMNLP 2022; Dec 7-11, 2022; Abu Dhabi, United Arab Emirates p. 5455-5477 URL: https://aclanthology.org/2022.findings-emnlp.400.pdf [accessed 2026-03-10] [doi: 10.18653/v1/2022.findings-emnlp.400]

XSL•FO
RenderX

39. Vlachogianni P, Tselios N. Perceived usability evaluation of educational technology using the System Usability Scale (SUS): a systematic review. J Res Technol Educ 2022 May 27;54(3):392-409. [doi: 10.1080/15391523.2020.1867938]
40. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv Psychol 1988;52:139-183. [doi: 10.1016/S0166-4115(08)62386-9]

Abbreviations

ALBERT: A Lite Bidirectional Encoder Representations from Transformers for Self-supervised Learning of Language Representations
BiLSTM: bidirectional long short-term memory
CARE-ET: CARE event timeline
DBN: deep belief network
DBSCAN: density-based spatial clustering of applications with noise
GAT: graph attention network
LCSTS: Large Scale Chinese Short Text Summarization
OALS: Older Adults' Life Stories
ROUGE: Recall-Oriented Understudy for Gisting Evaluation
SLEU: sequence link evaluation
SUS: System Usability Scale

Edited by M Mardini; submitted 28.Aug.2025; peer-reviewed by AS Rigaud, J Yang, L Lyu, L Yan, Y Yang, Z Liu; revised version received 15.Jan.2026; accepted 10.Feb.2026; published 06.Apr.2026.

Please cite as:
Gui F, Yang M, Jin L, Qian J
AI-Enhanced Automatic Life Story Structuring for Reminiscence Therapy in Older Adults: Technical Feasibility Study
JMIR Aging 2026;9:e83122
URL: https://aging.jmir.org/2026/1/e83122
doi:10.2196/83122

© Fang Gui, Mengchen Yang, Liuqi Jin, Jing Qian. Originally published in JMIR Aging (https://aging.jmir.org), 6.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Machine Learning–Derived Cardiovascular Aging Phenotypes From Cardiac Function and Stroke Risk in the UK Biobank: Cohort Study

Kang Yuan[1], PhD; Deyan Kong[2], MD; Jinghui Zhong[3], MD; Mengdi Xie[1], MD; Rui Liu[1], PhD; Wen Sun[3], PhD; Xinfeng Liu[1], PhD

[1]Department of Neurology, Affiliated Jinling Hospital, Medical School of Nanjing University, 305 Zhongshan East Road, Xuanwu District, Nanjing, Jiangsu Province, China

[2]Department of Neurology, The Second Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, China

[3]Department of Neurology, Centre for Leading Medicine and Advanced Technologies of IHM, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui, China

Corresponding Author:
Xinfeng Liu, PhD
Department of Neurology, Affiliated Jinling Hospital, Medical School of Nanjing University, 305 Zhongshan East Road, Xuanwu District, Nanjing, Jiangsu Province, China

Abstract

Background: Cardiovascular magnetic resonance (CMR) is widely used across various cardiac conditions and systematically assesses cardiac anatomical structures and functional dynamics. Machine learning (ML) can accurately predict outcomes and understand the inherent features of clinical data.

Objective: This study aimed to derive CMR phenotypes related to cardiovascular aging, investigate the relationship between these phenotypes and stroke risk, and relearn these phenotypes using supervised ML.

Methods: We enrolled 36,467 participants without stroke and extracted CMR parameters from the UK Biobank, with follow-up data extending until September 30, 2023. Using the generative topographic mapping technique, we identified latent grid nodes among participants and then derived phenotypes through agglomerative hierarchical clustering. We used supervised ML models to predict cardiac function phenotypes and used Cox proportional hazards models to assess the association between these phenotypes and long-term stroke risk.

Results: We enrolled 36,467 participants in the study. The mean age was 54.9 (SD 7.5) years, with 17,442 (47.8%) male participants. During a mean follow-up time of 14.7 (SD 1.1) years, 500 (1.4%) participants developed stroke and 664 (1.8%) participants died, respectively. After generative topographic mapping modeling, we identified 2 distinct phenotypes: phenotype 1, characterized by adverse cardiac function and an accumulation of cardiovascular risk factors, reflecting cardiovascular aging; and phenotype 2, associated with a lower risk of stroke (hazard ratio 0.695, 95% CI 0.559-0.864; *P*=.001), which remained significant after accounting for competing mortality (hazard ratio 0.578, 95% CI 0.484-0.691; *P*<.001). We selected the random forest model as the optimal model for the phenotypes, demonstrating high accuracy (area under the curve 0.914, 95% CI 0.911-0.918 for training and 0.867, 95% CI 0.858-0.876 for validation) and calibration ability (Brier score 0.111, 95% CI 0.109-0.113 for training and 0.132, 95% CI 0.127-0.137 for validation).

Conclusions: By integrating unsupervised and supervised ML methods, we identified cardiovascular aging–related phenotypes that demonstrate robust predictive ability for incident stroke, which may have the potential to improve preventive and therapeutic strategies for high-risk populations.

(JMIR Aging 2026;9:e77017) doi:10.2196/77017

KEYWORDS

stroke; generative topographic mapping; UK Biobank; machine learning; cardiac function; aging; phenotypes

Introduction

Stroke is a major cause of mortality and morbidity worldwide [1]. Aging is associated with progressive structural and functional deterioration of the cardiovascular system [2]. Several conditions associated with cardiovascular aging [3,4], including atrial fibrillation, heart failure, and myocardial infarction, are recognized as important risk factors for stroke [5-8]. Cardiovascular magnetic resonance (CMR) enables systematic assessment of cardiac structural and functional remodeling and has been reported to be associated with the aging process [9].

XSL•FO
RenderX

Previous studies have shown that atrial mechanical dysfunction and left ventricular (LV) structural abnormalities assessed by CMR are associated with stroke events [10,11]. However, CMR metrics are relatively complex, and extracting phenotypes from these metrics may provide insight into their relationship with cardiovascular aging and help clarify their association with stroke risk.

The application of artificial intelligence in clinical practice for disease prediction and detection has grown tremendously in recent years [12,13]. Supervised machine learning (ML) can accurately predict clinical outcomes, while unsupervised ML can identify the inherent features of data without requiring labeled outcomes [14]. Previous studies have used unsupervised ML models to delineate prognostically distinct clusters using variables related to LV diastolic function derived from transthoracic echocardiography [15]. Additionally, the hierarchical k-means clustering algorithm has also been effective in distinguishing potential embolic source groups in cryptogenic stroke [16]. Generative topographic mapping (GTM) is a probabilistic model used for dimensionality reduction and data visualization, exhibiting superiority over the self-organizing map algorithm by explicitly modeling probability distributions [17,18]. GTM can reduce data dimensionality to a latent space and reveal the underlying structure with robustness across diverse sources [19]. However, previous studies applying GTM to derive phenotypes have been limited, with applications reported in general and critical care populations [20].

Therefore, the aim of this study was to use the GTM model to identify distinct cardiac function phenotypes related to cardiovascular aging, examine their association with long-term stroke risk, and develop supervised ML models to predict these phenotypes using data from a prospective longitudinal cohort.

Methods

Study Population

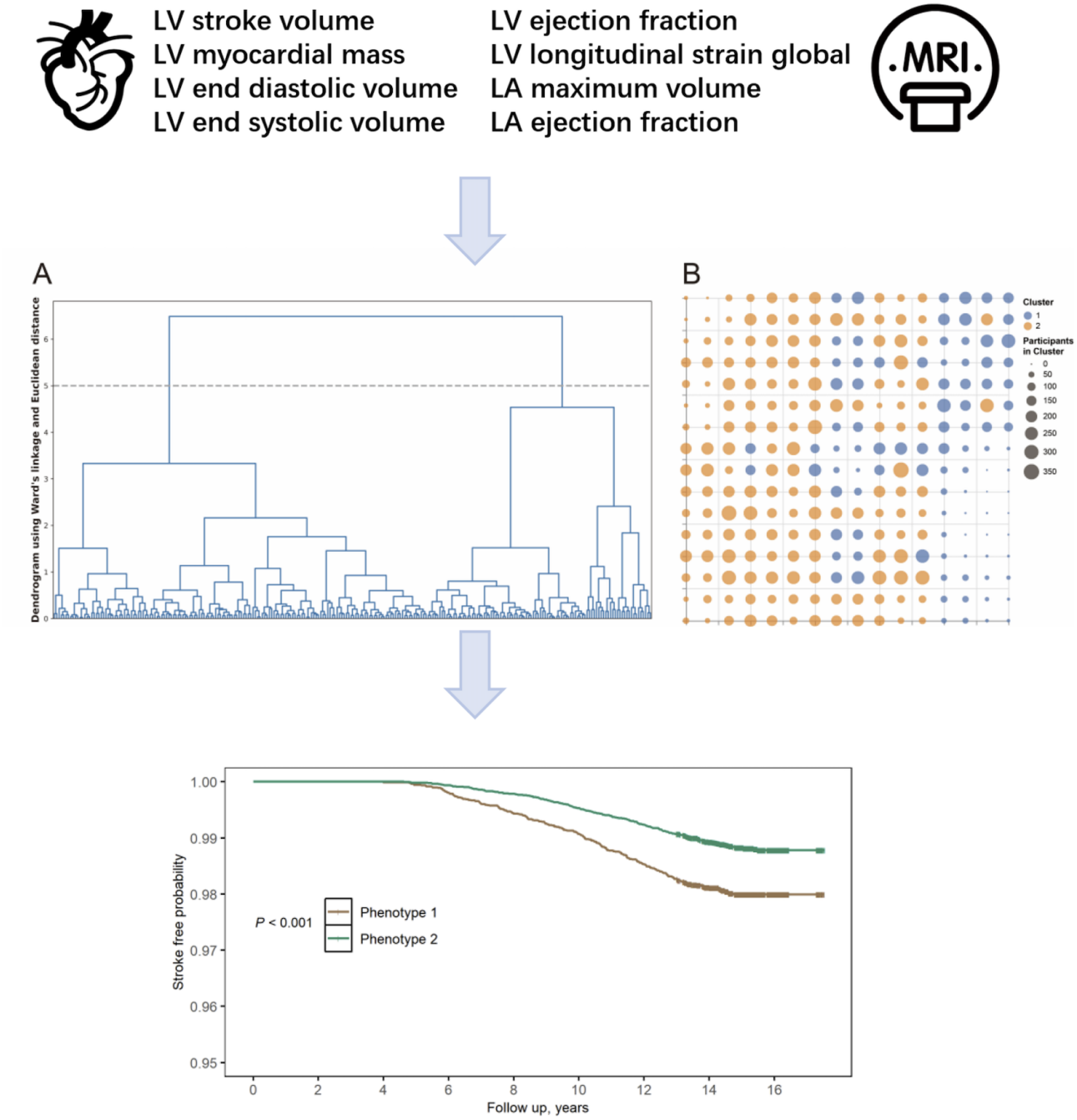
The UK Biobank is a large, population-based prospective cohort study comprising more than 500,000 participants aged 49 to 69 years, enrolled from 22 assessment centers in the United Kingdom between 2006 and 2010. Details of the UK Biobank protocols are accessible online [21]. Participants diagnosed with stroke at baseline (n=7780, 1.5%), lacking follow-up information (n=20,251, 4%), or without CMR data (n=437,713, 87.2%) were excluded, resulting in a final cohort of 36,467 (7.3%) participants.

We included the following CMR measures in the study: LV stroke volume (LVSV), LV myocardial mass (LVMM), LV end-diastolic volume (LVEDV), LV end-systolic volume (LVESV), LV ejection fraction (LVEF), LV global longitudinal strain (LVGLS), left atrial (LA) maximum volume (LAMV), and LA ejection fraction (LAEF) (Figure 1).

Stroke diagnosis was ascertained through linkage to self-reported medical conditions, hospital inpatient data, and death register records based on *International Classification of Diseases* codes, with the earliest recorded date of outcome provided in Table S1 in Multimedia Appendix 1. The follow-up time was defined as the interval from the date of recruitment to the occurrence of stroke, all-cause mortality, loss to follow-up, or September 30, 2023, whichever came first. Details of covariates were shown in Table S2 in Multimedia Appendix 1. This study was reported according to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines (Checklist 1).

XSL•FO
RenderX

Figure 1. Cardiovascular aging phenotypes and stroke. LA: left atrial; LV: left ventricular.



GTM Modeling and Phenotypes

Inspired by the algorithm introduced by Bellfield et al [20], we used a Gaussian mixture model to train the GTM model and visualize the distribution of data in the latent space. GTM uses a soft assignment method to estimate the probability of each participant being assigned to a cluster. The high-dimensional data were then visualized in a 2D latent space. After preprocessing the dataset, we used the *ugtm* Python (Python Software Foundation) package for GTM modeling and applied 10-fold cross-validation with the expectation-maximization algorithm to optimize the parameters. After GTM modeling, the generated latent space illustrated the distribution of participants in each cluster, with darker colors indicating higher densities. The modeling parameters (cardiac function variables) demonstrated the impact of cardiac function on each latent grid node in the GTM model. We aggregated the latent grid nodes by calculating the Euclidean distances between their modeling parameters and applied agglomerative hierarchical clustering using the Ward minimum variance method to derive phenotypes [22]. The characteristics of the phenotypes were compared to evaluate the distribution of risk factors for cardiovascular aging.

Phenotypes and Stroke Risk

We used Cox proportional hazards models to evaluate the association between cardiac function phenotypes and long-term stroke risk, reporting results as hazard ratios (HRs) with 95% CIs. In multivariable analyses, model 1 was an unadjusted model; model 2 was adjusted for age at recruitment, sex, Townsend deprivation index at recruitment, smoking status, alcohol intake frequency, systolic blood pressure, diastolic blood pressure, and BMI; and model 3 was additionally adjusted for

XSL•FO
RenderX

atrial fibrillation, type 2 diabetes, coronary heart disease, heart failure, antihypertensive drugs, and statins. In sensitivity analyses, we explored the association of cardiac function phenotypes with the risk of different stroke types. Additionally, we performed a competing risk analysis for stroke risk using the Fine and Gray method, accounting for mortality as the competing event.

Supervised ML Frameworks

Following the categorization of participants into different phenotypes through the GTM approach, we deployed several supervised ML models to predict cardiac function phenotypes. Hyperparameters for each model are detailed in Table S3 in Multimedia Appendix 1. Hyperparameter optimization was conducted using grid search and Bayesian optimization. Discriminatory ability was evaluated using the receiver operating characteristic curve. Calibration performance was assessed using the calibration curve. We selected the best-performing method as the final algorithm for developing a phenotype prediction model. We constructed a website using the Flask framework and the DeepSeek-R1 API (application programming interface). The predicted probabilities and individualized metrics were sent to the DeepSeek-R1 model to generate a detailed analysis of health status and provide professional recommendations. Detailed descriptions of the methods are provided in Multimedia Appendix 1.

Ethical Considerations

The ethical framework of the UK Biobank was approved by the North West Multi-centre Research Ethics Committee (11/NW/0382). All participants provided informed consent through electronic signature at enrollment.

Statistical Analysis

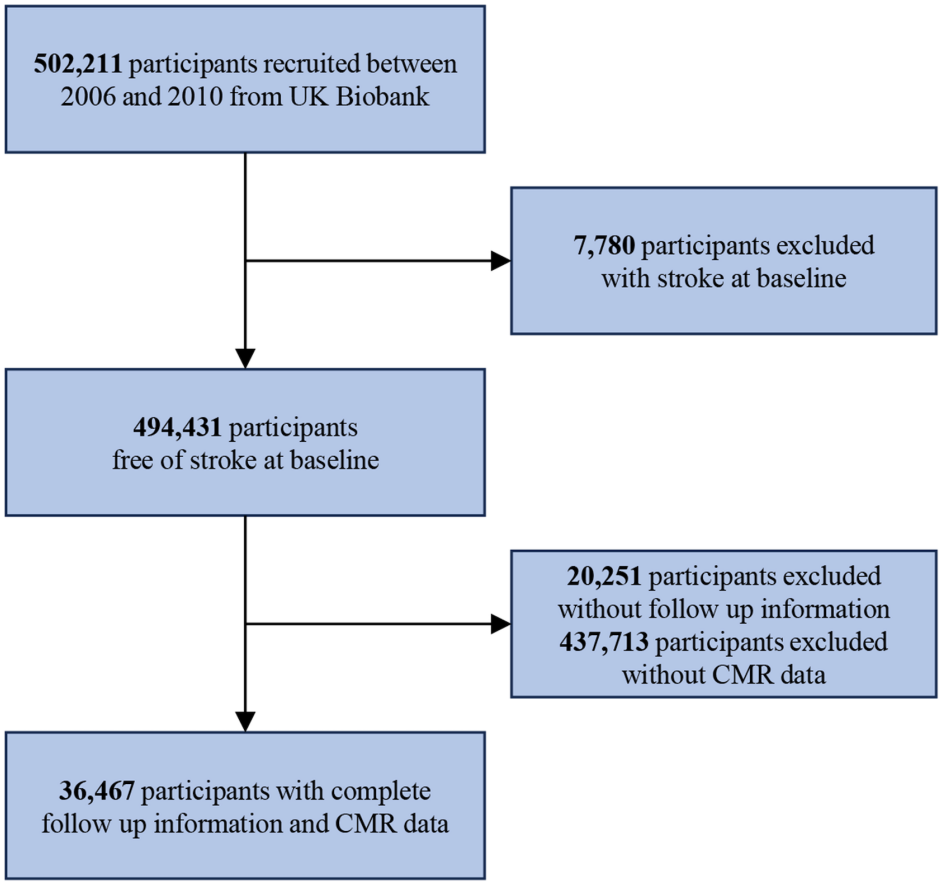
Normally distributed continuous data are presented as mean (SD), and group comparisons were performed using 2-tailed *t* tests. Categorical data were presented as frequency (percentage), and group comparisons were conducted using chi-square tests. To account for multiple comparisons, *P* values were adjusted using the false discovery rate method. Statistical tests were conducted with R software (version 4.2.1; R Foundation for Statistical Computing), and a 2-sided *P* value <.05 was considered statistically significant.

Results

Study Population

The study design is shown in Figure 2. Of the 36,467 participants included in the study (Figure 2), the mean age was 54.9 (SD 7.5) years, with 17,442 (47.8%) male participants. In total, 14,373 (39.5%) of participants were current or former smokers, 8180 (22.4%) had daily or almost daily alcohol intake, 1047 (2.9%) had type 2 diabetes, 345 (0.9%) had atrial fibrillation, and 48 (0.1%) had heart failure at baseline. During a mean follow-up time of 14.7 (SD 1.1) years, 500 (1.4%) and 664 (1.8%) participants developed stroke and died, respectively (Table 1).

Figure 2. Flowchart illustrating the study design. CMR: cardiovascular magnetic resonance.



XSL•FO
RenderX

Table . Characteristics of the participants included in the study.

Characteristics	Total (n=36,467)	Phenotype 1 (n=10,749)	Phenotype 2 (n=25,718)	Chi square (*df*) or t test (*df*)	*P* value	Adjusted *P* value[a]
Age (years), mean (SE)	54.9 (0.0)	54.4 (0.1)	55 (0.0)	−6.74 (19,563.5)[b]	<.001	<.001
Male, n (%)	17,442 (47.8)	9333 (86.8)	8109 (31.5)	9287 (1)[c]	<.001	<.001
Ethnicity, n (%)				78.9 (4)[c]	<.001	<.001
Asian	367 (1.0)	45 (0.4)	322 (1.3)			
Black	219 (0.6)	76 (0.7)	143 (0.6)			
Chinese	102 (0.3)	9 (0.1)	93 (0.4)			
Mixed	162 (0.4)	40 (0.4)	122 (0.5)			
White[d]	35,342 (97.7)	10,503 (98.4)	24,839 (97.3)			
Townsend deprivation index, mean (SE)	−1.9 (0.0)	−2 (0.0)	−1.9 (0.0)	−2.39 (20,325.8)[b]	.02	.02
Overall health rating, n (%)				35.6 (3)[c]	<.001	<.001
Poor	629 (1.7)	199 (1.9)	430 (1.7)			
Fair	5123 (14.1)	1683 (15.7)	3440 (13.4)			
Good	22,137 (60.8)	6359 (59.2)	15,778 (61.5)			
Excellent	8503 (23.4)	2492 (23.2)	6011 (23.4)			
Smoking status, n (%)				88.8 (2)[c]	<.001	<.001
Current	2297 (6.3)	771 (7.2)	1526 (5.9)			
Previous	12,076 (33.2)	3859 (36.0)	8217 (32.0)			
Never	22,007 (60.5)	6089 (56.8)	15,918 (62.0)			
Alcohol intake frequency, n (%)				415 (5)[c]	<.001	<.001
Daily or almost daily	8180 (22.4)	2800 (26.1)	5380 (20.9)			
3 or 4 times a week	10,190 (28.0)	3339 (31.1)	6851 (26.7)			
Once or twice a week	9389 (25.8)	2712 (25.2)	6677 (26.0)			
1-3 times a month	3963 (10.9)	946 (8.8)	3017 (11.7)			
Special occasions only	2997 (8.2)	554 (5.2)	2443 (9.5)			
Never	1728 (4.7)	393 (3.7)	1335 (5.2)			
Comorbidity, n (%)						
Dementia	8 (0.0)	2 (0.0)	6 (0.0)	0.001 (1)[c]	.99	.99
Major adverse cardiac events	857 (2.4)	432 (4.0)	425 (1.7)	184 (1)[c]	<.001	<.001
Type 2 diabetes	1047 (2.9)	368 (3.4)	679 (2.6)	16.4 (1)[c]	<.001	<.001
Liver disease	416 (1.1)	118 (1.1)	298 (1.2)	0.199 (1)[c]	.66	.73
Renal disease	446 (1.2)	132 (1.2)	314 (1.2)	0.001 (1)[c]	.99	.99
Atrial fibrillation	345 (0.9)	212 (2.0)	133 (0.5)	170 (1)[c]	<.001	<.001
Heart failure	48 (0.1)	35 (0.3)	13 (0.1)	41.6 (1)[c]	<.001	<.001
Coronary heart disease	1090 (3.0)	528 (4.9)	562 (2.2)	193 (1)[c]	<.001	<.001

XSL•FO
RenderX

Characteristics	Total (n=36,467)	Phenotype 1 (n=10,749)	Phenotype 2 (n=25,718)	Chi square (*df*) or t test (*df*)	*P* value	Adjusted *P* value[a]
Venous thrombosis	471 (1.3)	141 (1.3)	330 (1.3)	0.029 (1)[c]	.87	.91
Abdominal aortic aneurysm	5 (0.0)	3 (0.0)	2 (0.0)	1.01 (1)[c]	.31	.39
Peripheral arterial disease	255 (0.7)	68 (0.6)	187 (0.7)	0.844 (1)[c]	.36	.44
Cardiac function, mean (SE)[c]					<.001	<.001
Left ventricular stroke volume (mL)	87.5 (0.1)	104 (0.2)	80.6 (0.1)	113 (15,816.1)[b]		
Left ventricular myocardial mass (g)	86.1 (0.1)	109.2 (0.2)	76.4 (0.1)	161 (17,230.0)[b]		
Left ventricular end-diastolic volume (mL)	147.9 (0.2)	184.7 (0.3)	132.5 (0.1)	171 (16,189.1)[b]		
Left ventricular end-systolic volume (mL)	60.5 (0.1)	80.7 (0.2)	52 (0.1)	149 (14,438.1)[b]		
Left ventricular ejection fraction (%)	59.5 (0.0)	56.3 (0.1)	60.9 (0.0)	−64.2 (17,086.9)[b]		
Left ventricular longitudinal strain global (%)	−18.5 (0.0)	−17.3 (0.0)	−19 (0.0)	55 (19,909.5)[b]		
Left atrial maximum volume (mL)	72.8 (0.1)	92.6 (0.2)	64.4 (0.1)	109 (15,314.3)[b]		
Left atrial ejection fraction (%)	61.2 (0)	56.5 (0.1)	63.1 (0.1)	−59.8 (17,184.7)[b]		
Outcomes					<.001	
Follow-up time (y), mean (SE)	14.7 (0.0)	14.6 (0.0)	14.7 (0.0)	−3.05 (19,080.5)[b]		.003
Stroke, n (%)	500 (1.4)	209 (1.9)	291 (1.1)	36.4 (1)[c]		<.001
Mortality, n (%)	664 (1.8)	265 (2.5)	399 (1.6)	34.9 (1)[c]		<.001

[a]*P* values were adjusted by false discovery rate.

[b]*t* test.

[c]Chi-square test.

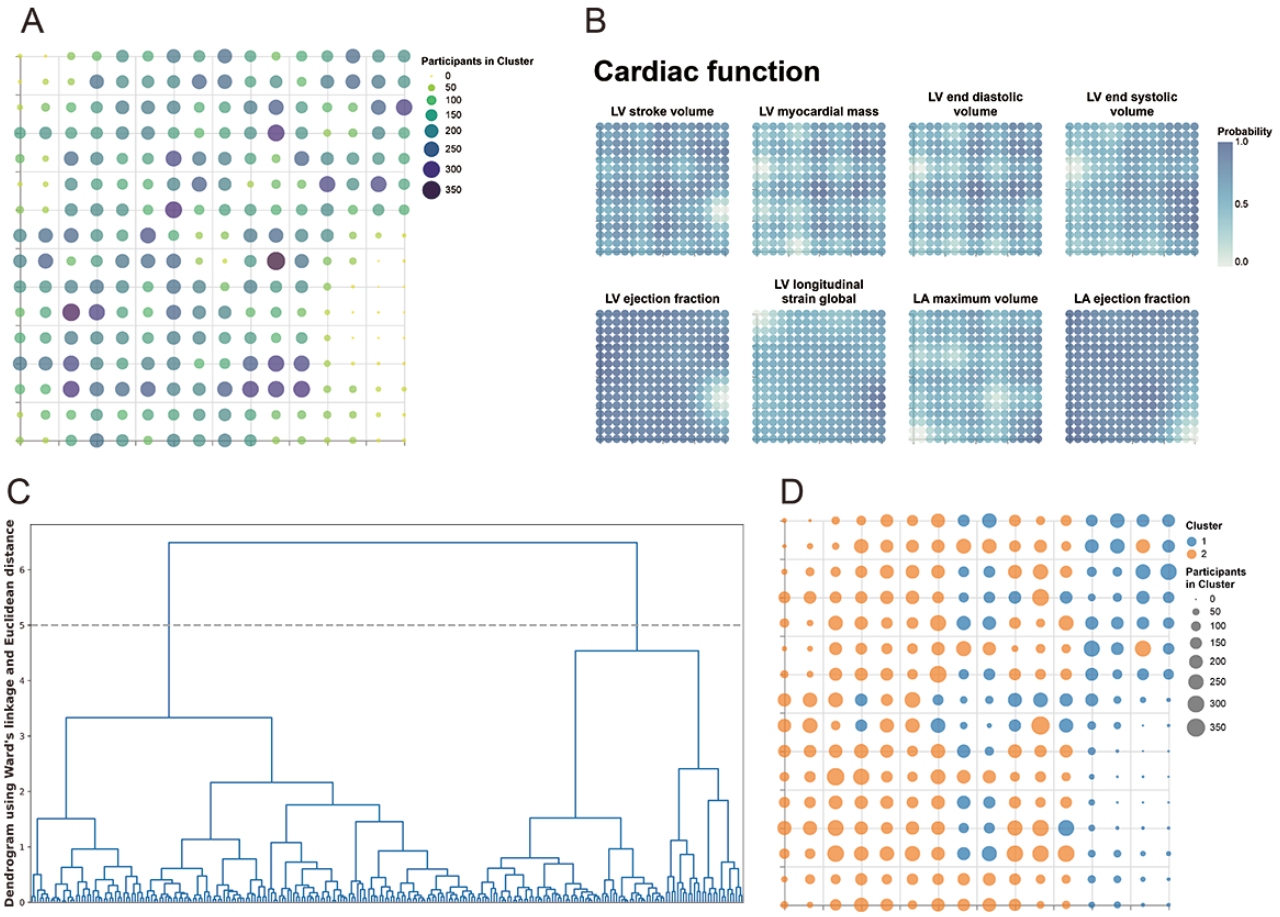
[d]Missing values for this row: n=275 missing for the Total column, n=76 missing for the Phenotype 1 column, and n=199 missing for the Phenotype 2 column.

GTM Modeling and Cardiac Aging Phenotypes

Figure 3 illustrates the latent space generated by the GTM model. The number of participants in clusters is represented using the “viridis” color scheme and visualized as circles, where a deeper blue hue and larger circle diameter indicate a higher number of participants. LVSV, LVMM, LVEDV, LVESV, LVEF, LVGLS, LAMV, and LAEF are aligned point-to-point with their latent space in a light teal blue scheme. In the dendrogram plot, 2 cardiac function phenotypes are identified, clearly separating the dataset into 2 major branches at a relatively large linkage distance. Figure S1 in Multimedia Appendix 1 illustrates that stroke diagnoses among participants were mainly distributed in the area of phenotype 1.

XSL•FO
RenderX

Figure 3. Derived phenotypes of cardiac function. (A) Latent space showing the distribution of participants; (B) latent space displaying the distribution of cardiac function markers derived from cardiovascular magnetic resonance; (C) dendrogram using Ward linkage and Euclidean distance to identify 2 cardiac function phenotypes; and (D) phenotypes and latent grid nodes of cardiac function. LA: left atrial; LV: left ventricular.



To further characterize the phenotype patterns, we compared the distributions of investigatory variables and visualized them on the latent space with the light orange color scheme (Figures S2-S7 in Multimedia Appendix 1). Significant differences were observed between the 2 cardiac function phenotypes across demographic characteristics, cardiovascular risk factors, and cardiac function parameters. Participants in phenotype 1 exhibited larger cardiac chamber volumes and greater myocardial mass, with higher LVEF and LAEF but lower LVSV, LVMM, LVEDV, LVESV, LVGLS, and LAMV compared with phenotype 2 (all $P<.001$). In addition, phenotype 1 demonstrated a higher burden of cardiovascular risk factors and comorbidities, including a greater prevalence of major adverse cardiac events, atrial fibrillation, heart failure, coronary heart disease, and type 2 diabetes ($P<.001$). Participants in phenotype 1 also had higher blood pressure, arterial stiffness index, and more adverse metabolic profiles. Taken together, these structural, functional, and clinical differences suggest that phenotype 1 represents a cardiovascular aging–related phenotype, characterized by cardiac remodeling, increased cardiovascular risk burden, and elevated incidence of stroke and mortality (Table 1 and Table S4 in Multimedia Appendix 1).

Phenotypes and Stroke Risk

Multivariable modeling identified that cardiovascular aging phenotypes were associated with the long-term risk of stroke and mortality. Compared with phenotype 1, phenotype 2 was significantly associated with a decreased risk of stroke in the unadjusted model (HR 0.579, 95% CI 0.484-0.691; $P<.001$); model 2 (HR 0.671, 95% CI 0.541-0.833; $P<.001$); and model 3 (HR 0.695, 95% CI 0.559-0.864; $P=.001$; Table 2).

XSL•FO
RenderX

Table . Association between cardiovascular aging phenotypes and long-term risk of stroke and mortality.

Outcome (phenotype 2 vs phenotype 1)	Stroke		Mortality	
	HR[a] (95% CI)	*P* value	HR (95% CI)	*P* value
Model 1[b]	0.579 (0.484-0.691)	<.001	0.625 (0.535-0.730)	<.001
Model 2[c]	0.671 (0.541-0.833)	<.001	0.760 (0.632-0.913)	.003
Model 3[d]	0.695 (0.559-0.864)	.001	0.772 (0.641-0.929)	.006

[a]HR: hazard ratio.

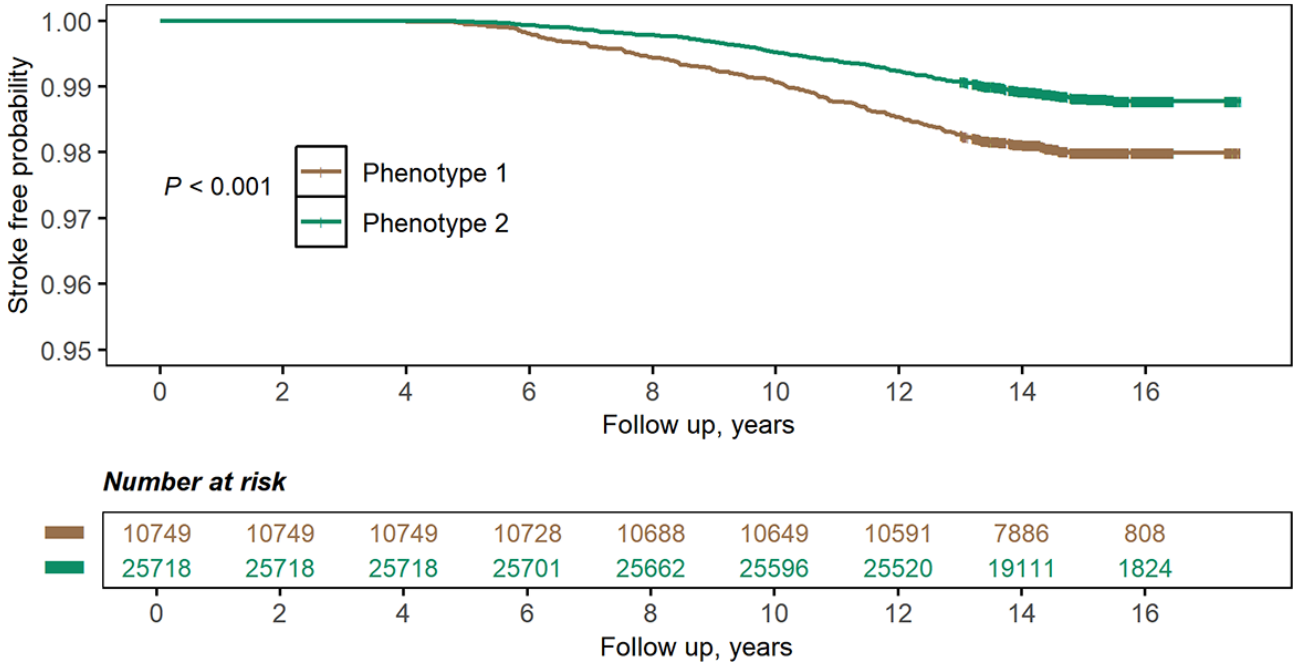
[b]Model 1 was unadjusted.

[c]Model 2 was adjusted for age at recruitment, sex, Townsend deprivation index at recruitment, smoking status, alcohol intake frequency, systolic blood pressure, diastolic blood pressure, and BMI.

[d]Model 3 was additionally adjusted for atrial fibrillation, type 2 diabetes, coronary heart disease, heart failure, antihypertensive drugs, and statins.

Kaplan-Meier survival curves revealed that the phenotype 1 group had a higher risk of stroke (*P*<.001; Figure 4).

Figure 4. Kaplan-Meier survival curves of stroke-free probability by phenotype.



Regarding stroke subtypes, cardiac function phenotypes were associated with cerebral infarction and other nontraumatic intracranial hemorrhage (all *P*<.05; Table S5 in Multimedia Appendix 1). When mortality was considered as a competing risk, phenotype 2 remained significantly associated with stroke risk (HR 0.578, 95% CI 0.484-0.691; *P*<.001; Figure S8 in Multimedia Appendix 1).

Supervised ML

Basal metabolic rate, sex, testosterone, standing height, weight, forced vital capacity, waist circumference, creatinine, forced expiratory volume in 1 second, and urate were included in the ML models (Table S6 in Multimedia Appendix 1). By integrating the accuracy in the training and validation sets, we selected the random forest model as the optimal model based on its excellent accuracy (area under the curve 0.914, 95% CI 0.911-0.918 for training and 0.867, 95% CI 0.858-0.876 for validation; and accuracy 0.837, 95% CI 0.833-0.841 for training and 0.805, 95% CI 0.804-0.806 for validation) without overfitting the training set (Figure 5). Calibration plots exhibited that the random forest model had good performance in the calibration quality (Brier score 0.111, 95% CI 0.109-0.113 for training and 0.132, 95% CI 0.127-0.137 for validation; Table 3 and Tables S7 and S8 in Multimedia Appendix 1). We drew a feature importance plot to better visualize the contributions of the predictors to the predictive power of the random forest model (Figure S9 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 5. Machine learning for phenotypes. ANN: artificial neural network; AUC: area under the curve; KNN: k-nearest neighbors; LGBM: light gradient boosting machine; LR: logistic regression; RF: random forest; SVM: support vector machine; XGB: extreme gradient boosting (XGBoost).

XSL•FO
RenderX

Table . Model evaluation and validation.

Models	Training set		Validation set	
	Accuracy (95% CI)	Brier score (95% CI)	Accuracy (95% CI)	Brier score (95% CI)
K-nearest neighbor	1.000 (1.000-1.000)	0.000 (0.000-0.000)	0.800 (0.800-0.800)	0.134 (0.130-0.139)
Logistic regression	0.797 (0.793-0.802)	0.136 (0.134-0.139)	0.802 (0.802-0.802)	0.132 (0.127-0.137)
Support vector machine	0.795 (0.791-0.800)	0.137 (0.135-0.140)	0.802 (0.802-0.802)	0.133 (0.128-0.138)
Random forest	0.837 (0.833-0.841)	0.111 (0.109-0.113)	0.805 (0.804-0.806)	0.132 (0.127-0.137)
Light gradient boosting machine	0.810 (0.805-0.814)	0.126 (0.124-0.128)	0.805 (0.805-0.805)	0.132 (0.128-0.137)
Extreme Gradient Boosting	0.799 (0.795-0.804)	0.134 (0.132-0.137)	0.804 (0.804-0.804)	0.132 (0.127-0.137)
Artificial neural networks	0.797 (0.793-0.802)	0.136 (0.133-0.138)	0.803 (0.803-0.804)	0.132 (0.127-0.136)

After entering the participant information, the local model calculated the corresponding phenotype, and the DeepSeek-R1 model provided an analysis of the participant's health status along with recommendations for potential clinical application (Figure S10 in Multimedia Appendix 1).

Discussion

Principal Findings

In this study, we identified distinct cardiac function phenotypes related to cardiovascular aging using the GTM model based on CMR parameters. These phenotypes were significantly associated with long-term stroke risk and could be reliably predicted by supervised ML models.

CMR is important for evaluating cardiovascular pathologies and for assessing the etiology and prognosis of patients with stroke. A previous study demonstrated that cardiovascular aging is associated with reduced LV volumes and increased LV concentricity on CMR, with notable sex-specific differences in systolic cardiac function [9]. CMR also exhibited high sensitivity in detecting intraventricular thrombi and thrombi within the LA appendage [23]. Additionally, CMR imaging could affect the management of acute stroke by detecting aortic plaques, cardiac structural abnormalities, and intracardiac thrombi in patients with stroke [24,25]. However, CMR metrics represented multidimensional data with complex interrelations, potentially involving collinearity and interaction effects. Using unsupervised ML may improve the ability to reveal the inherent structure of the data and achieve more accurate classification.

GTM incorporated probability distribution functions to refine the self-organizing map algorithms and mitigated their inherent limitations, such as convergence issues, insufficient neighborhood preservation, and the absence of a distinct objective function. Moreover, GTM was capable of visually representing phenotypes through latent spaces [17]. Sattarov et al [26] applied GTM to explore the latent space of Simplified Molecular Input Line Entry System–based autoencoders for generating targeted molecular libraries and visualized the autoencoder latent space on a 2D topographic map. These findings indicated a promising application of GTM in structure-based affinity assessments. However, unsupervised ML faces challenges related to limited interpretability and scalability, making the derived phenotypes and their clinical relevance difficult to translate into real-world medical practice. Therefore, we innovatively applied supervised ML to relearn the phenotype labels obtained from GTM, constructing a random forest model that accurately predicted the 2 distinct phenotypes. Given the high cost and inconvenience of CMR scans for patients with stroke [27], using clinical variables allowed for a reduction in model application requirements while preserving accuracy, making the classification model more suitable for clinical use.

The phenotypes derived from CMR metrics may reflect distinct patterns of cardiac structural and functional remodeling associated with cardiovascular aging. Key variables contributing to phenotype differentiation, including LVMM and concentric geometry, have been linked to vascular aging markers, supporting the connection between cardiovascular aging and cardiac structural changes [28]. LA enlargement and functional impairment are also recognized markers of cardiovascular aging and are related to chronic diastolic dysfunction and atrial fibrillation [29]. In addition, accelerated cardiovascular aging driven by cumulative risk factors such as hypertension, obesity, diabetes, and coronary artery disease may further promote cardiac remodeling, consistent with the differing comorbidity profiles observed between the phenotypes in our study [30]. Notably, the cardiovascular aging–related phenotype in our study was characterized by LV and LA abnormalities indicative of volume or pressure overload and myocardial dysfunction. These alterations have been associated with increased risks of cardioembolic stroke and subclinical cerebral infarction [31,32]. Therefore, the CMR-derived phenotypes may capture different stages of cardiovascular aging and help explain the heterogeneity of cardiac remodeling and its potential contribution to long-term stroke risk.

In our study, supervised ML models identified several clinical parameters as predictors for cardiac function phenotypes. Notably, basal metabolic rate is the energy required to maintain essential functions in a healthy state. Prior research indicated that basal metabolic rate was independently linked to heart failure, atrial fibrillation, and flutter and acted as a risk factor for cardiovascular mortality [33,34]. Furthermore, a pooled analysis of 8 cohorts showed that obstructive and restrictive physiology were associated with incident heart failure with reduced or preserved ejection fraction [35]. Moreover, a distinctive J-shaped curve was noted between BMI and waist

XSL•FO
RenderX

circumference and the risk of heart failure [36]. Additionally, renal impairment in chronic heart failure was typically linked to decreased cardiac output and subsequent renal hypoperfusion, and targeting volume overload may improve outcomes in cardiac dysfunction [37]. These findings suggested that the ML model could reflect systemic metabolic, cardiopulmonary, and renal alterations that may explain the clinical characteristics of the identified cardiac function phenotypes.

To the best of our knowledge, this was the first study to use GTM for clustering CMR parameters and to identify cardiovascular aging phenotypes associated with long-term stroke risk. However, there were several limitations to our study. First, the CMR parameters extracted from the UK Biobank were collected according to the cohort's standardized protocols. Because the CMR examination protocol may differ in other studies, this should be considered when applying our model to other cohorts. Second, while trained on the UK Biobank's large sample, the model's applicability to other cohorts may be limited by environmental, genetic, and cultural variations. Third, given the predominantly White and healthy population in the UK Biobank, selection and health biases may be inevitable. The generalizability of our findings to ancestrally diverse individuals will assist in determining whether more appropriate and ethnically relevant decisions are needed.

Conclusions

In conclusion, we used GTM to identify cardiac function phenotypes based on CMR parameters. The identified phenotypes were significantly associated with stroke risk, with the high-risk phenotype characterized by impaired cardiac function and features related to cardiovascular aging. Furthermore, we relearned phenotype labels using supervised ML models, enabling phenotype prediction from clinical variables. These models may facilitate the identification of individuals at higher risk and support the development of preventive and therapeutic strategies. Further studies are warranted to explore the potential mechanisms linking cardiovascular aging, cardiac phenotypic heterogeneity, and stroke risk.

Acknowledgments

The authors thank the UK Biobank participants. This study was conducted using the UK Biobank resource (application number 106487).

Funding

The project was supported by the Program for Innovative Research Team of the First Affiliated Hospital of University of Science and Technology of China (2023IHM01050), the National Natural Science Foundation of China (NSFC-82471341 and NSFC-U22A20341), the Guangxi Natural Science Foundation (2025GXNSFAA069341), and the Joint Project on Regional High-Incidence Diseases Research of the Guangxi Natural Science Foundation (2023GXNSFAA026197).

Data Availability

All data used in this study were accessed from the publicly available UK Biobank resource (application number 106487). Data from this study cannot be shared with other investigators. The source code for the analyses and the established machine learning models is available on GitHub [38].

Authors' Contributions

Conceptualization: KY, MX, WS
Data curation: JZ
Formal analysis: JZ
Funding acquisition: XL
Investigation: JZ
Project administration: WS
Software: JZ
Supervision: WS
Writing—original draft: KY, DK
Writing—review and editing: RL, WS, XL

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional methods and results detailing cardiovascular magnetic resonance analysis, generative topographic mapping for cardiac phenotyping, and machine learning frameworks for stroke risk prediction. Includes comprehensive methodology, participant characteristics, and model validation metrics using UK Biobank data.
[DOCX File, 8399 KB - aging_v9i1e77017_app1.docx]

XSL•FO
RenderX

Checklist 1
TRIPOD checklist.
[PDF File, 681 KB - aging_v9i1e77017_app2.pdf]

References

1. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820. [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
2. Vakka A, Warren JS, Drosatos K. Cardiovascular aging: from cellular and molecular changes to therapeutic interventions. J Cardiovasc Aging 2023;3(3):23. [doi: 10.20517/jca.2023.09] [Medline: 37274126]
3. Rim D, Henderson LA, Macefield VG. Brain and cardiovascular-related changes are associated with aging, hypertension, and atrial fibrillation. Clin Auton Res 2022 Dec;32(6):409-422. [doi: 10.1007/s10286-022-00907-9] [Medline: 36409380]
4. Subramanian V, Tucker WJ, Peters AE, Upadhya B, Kitzman DW, Pandey A. Cardiovascular aging and exercise: implications for heart failure prevention and management. Circ Res 2025 Jul 7;137(2):205-230. [doi: 10.1161/CIRCRESAHA.125.325531] [Medline: 40608862]
5. Polymeris AA, Meinel TR, Oehler H, et al. Aetiology, secondary prevention strategies and outcomes of ischaemic stroke despite oral anticoagulant therapy in patients with atrial fibrillation. J Neurol Neurosurg Psychiatry 2022 Jun;93(6):588-598. [doi: 10.1136/jnnp-2021-328391] [Medline: 35396339]
6. Doehner W, Böhm M, Boriani G, et al. Interaction of heart failure and stroke: a clinical consensus statement of the ESC Council on Stroke, the Heart Failure Association (HFA) and the ESC Working Group on Thrombosis. Eur J Heart Fail 2023 Dec;25(12):2107-2129. [doi: 10.1002/ejhf.3071] [Medline: 37905380]
7. Dutta P, Courties G, Wei Y, et al. Myocardial infarction accelerates atherosclerosis. Nature 2012 Jul 19;487(7407):325-329. [doi: 10.1038/nature11260] [Medline: 22763456]
8. Mayfield JJ, Otto CM. Stroke and noninfective native valvular disease. Curr Cardiol Rep 2023 May;25(5):333-348. [doi: 10.1007/s11886-023-01855-0] [Medline: 36971961]
9. Grassow L, Gröschel J, Saad H, et al. Sex-specific structural and functional cardiac remodeling during healthy aging assessed by cardiovascular magnetic resonance. Clin Res Cardiol 2025 Dec;114(12):1619-1630. [doi: 10.1007/s00392-024-02430-5] [Medline: 38466347]
10. Maheshwari A, Norby FL, Inciardi RM, et al. Left atrial mechanical dysfunction and the risk for ischemic stroke in people without prevalent atrial fibrillation or stroke: a prospective cohort study. Ann Intern Med 2023 Jan;176(1):39-48. [doi: 10.7326/M22-1638] [Medline: 36534978]
11. Tsao CW, Gona PN, Salton CJ, et al. Left ventricular structure and risk of cardiovascular events: a Framingham Heart Study cardiac magnetic resonance study. J Am Heart Assoc 2015 Sep 15;4(9):e002188. [doi: 10.1161/JAHA.115.002188] [Medline: 26374295]
12. Narang A, Bae R, Hong H, et al. Utility of a deep-learning algorithm to guide novices to acquire echocardiograms for limited diagnostic use. JAMA Cardiol 2021 Jun 1;6(6):624-632. [doi: 10.1001/jamacardio.2021.0185] [Medline: 33599681]
13. Kusunose K, Haga A, Abe T, Sata M. Utilization of artificial intelligence in echocardiography. Circ J 2019 Jul 25;83(8):1623-1629. [doi: 10.1253/circj.CJ-19-0420] [Medline: 31257314]
14. Verbeeck N, Caprioli RM, Van de Plas R. Unsupervised machine learning for exploratory data analysis in imaging mass spectrometry. Mass Spectrom Rev 2020 May;39(3):245-291. [doi: 10.1002/mas.21602] [Medline: 31602691]
15. Chao CJ, Kato N, Scott CG, et al. Unsupervised machine learning for assessment of left ventricular diastolic function and risk stratification. J Am Soc Echocardiogr 2022 Dec;35(12):1214-1225.e8. [doi: 10.1016/j.echo.2022.06.013] [Medline: 35840082]
16. Ntaios G, Weng SF, Perlepe K, et al. Data‐driven machine‐learning analysis of potential embolic sources in embolic stroke of undetermined source. Eur J Neurol 2021 Jan;28(1):192-201. [doi: 10.1111/ene.14524] [Medline: 32918305]
17. Kireeva N, Baskin II, Gaspar HA, Horvath D, Marcou G, Varnek A. Generative topographic mapping (GTM): universal tool for data visualization, structure-activity modeling and dataset comparison. Mol Inform 2012 Apr;31(3-4):301-312. [doi: 10.1002/minf.201100163] [Medline: 27477099]
18. Olier I, Vellido A. Advances in clustering and visualization of time series using GTM through time. Neural Netw 2008 Sep;21(7):904-913. [doi: 10.1016/j.neunet.2008.05.013] [Medline: 18653311]
19. Ezugwu AE, Shukla AK, Agbaje MB, Oyelade ON, José-García A, Agushaka JO. Automatic clustering algorithms: a systematic review and bibliometric analysis of relevant literature. Neural Comput Appl 2021 Jun;33:6247-6306. [doi: 10.1007/s00521-020-05395-4]
20. Bellfield RA, Olier I, Lotto R, et al. AI-based derivation of atrial fibrillation phenotypes in the general and critical care populations. EBioMedicine 2024 Sep;107:105280. [doi: 10.1016/j.ebiom.2024.105280] [Medline: 39153412]
21. Health research data for the world. UK Biobank. URL: https://www.ukbiobank.ac.uk [accessed 2026-04-04]
22. Vellido A, Lisboa P, Meehan K. The generative topographic mapping as a principal model for data visualization and market segmentation: an electronic commerce case study. Int J Comput Syst Signals 2000;1(2):119-138 [FREE Full text]

XSL•FO
RenderX

23. Takasugi J, Yamagami H, Noguchi T, et al. Detection of left ventricular thrombus by cardiac magnetic resonance in embolic stroke of undetermined source. Stroke 2017 Sep;48(9):2434-2440. [doi: 10.1161/STROKEAHA.117.018263] [Medline: 28818863]
24. Siebermair J, Kholmovski EG, Marrouche N. Assessment of left atrial fibrosis by late gadolinium enhancement magnetic resonance imaging: methodology and clinical implications. JACC Clin Electrophysiol 2017 Aug;3(8):791-802. [doi: 10.1016/j.jacep.2017.07.004] [Medline: 29759774]
25. Fonseca AC, Alves P, Inácio N, et al. Patients with undetermined stroke have increased atrial fibrosis: a cardiac magnetic resonance imaging study. Stroke 2018 Mar;49(3):734-737. [doi: 10.1161/STROKEAHA.117.019641] [Medline: 29371431]
26. Sattarov B, Baskin II, Horvath D, Marcou G, Bjerrum EJ, Varnek A. De novo molecular design by combining deep autoencoder recurrent neural networks with generative topographic mapping. J Chem Inf Model 2019 Mar 25;59(3):1182-1196. [doi: 10.1021/acs.jcim.8b00751] [Medline: 30785751]
27. Meinel TR, Eggimann A, Brignoli K, et al. Cardiovascular MRI compared to echocardiography to identify cardioaortic sources of ischemic stroke: a systematic review and meta-analysis. Front Neurol 2021;12:699838. [doi: 10.3389/fneur.2021.699838] [Medline: 34393979]
28. Liu T, Zeng J, Zhao X, et al. Relationship between vascular aging and left ventricular geometry in patients with obstructive sleep apnea hypopnea syndrome-related hypertension. Sci Rep 2025;15(1):6191. [doi: 10.1038/s41598-025-89964-4] [Medline: 39979427]
29. Fang Z, Raza U, Song J, et al. Systemic aging fuels heart failure: molecular mechanisms and therapeutic avenues. ESC Heart Fail 2025 Apr;12(2):1059-1080. [doi: 10.1002/ehf2.14947] [Medline: 39034866]
30. Triposkiadis F, Xanthopoulos A, Parissis J, Butler J, Farmakis D. Pathogenesis of chronic heart failure: cardiovascular aging, risk factors, comorbidities, and disease modifiers. Heart Fail Rev 2022 Jan;27(1):337-344. [doi: 10.1007/s10741-020-09987-z] [Medline: 32524327]
31. Kamel H, Okin PM, Merkler AE, et al. Relationship between left atrial volume and ischemic stroke subtype. Ann Clin Transl Neurol 2019 Aug;6(8):1480-1486. [doi: 10.1002/acn3.50841] [Medline: 31402612]
32. Shaikh Q, Ahmed B, Ahmed M, et al. Left atrial volumes and associated stroke subtypes. BMC Neurol 2013 Oct 18;13:149. [doi: 10.1186/1471-2377-13-149] [Medline: 24139054]
33. Zhao P, Han F, Liang X, et al. Causal effects of basal metabolic rate on cardiovascular disease: a bidirectional Mendelian randomization study. J Am Heart Assoc 2024 Jan 2;13(1):e031447. [doi: 10.1161/JAHA.123.031447] [Medline: 38156559]
34. Ruggiero C, Metter EJ, Melenovsky V, et al. High basal metabolic rate is a risk factor for mortality: the Baltimore Longitudinal Study of Aging. J Gerontol A Biol Sci Med Sci 2008 Jul;63(7):698-706. [doi: 10.1093/gerona/63.7.698] [Medline: 18693224]
35. Eckhardt CM, Balte PP, Barr RG, et al. Lung function impairment and risk of incident heart failure: the NHLBI Pooled Cohorts Study. Eur Heart J 2022 Jun 14;43(23):2196-2208. [doi: 10.1093/eurheartj/ehac205] [Medline: 35467708]
36. Zhou Y, Chai X, Yang G, Sun X, Xing Z. Changes in body mass index and waist circumference and heart failure in type 2 diabetes mellitus. Front Endocrinol (Lausanne) 2023;14:1305839. [doi: 10.3389/fendo.2023.1305839] [Medline: 38179309]
37. Sinkeler SJ, Damman K, van Veldhuisen DJ, Hillege H, Navis G. A re-appraisal of volume status and renal function impairment in chronic heart failure: combined effects of pre-renal failure and venous congestion on renal function. Heart Fail Rev 2012 Mar;17(2):263-270. [doi: 10.1007/s10741-011-9233-7] [Medline: 21365244]
38. Qianwg/ml_stroke. GitHub. URL: https://github.com/qianwg/ML_Stroke [accessed 2025-05-25]

Abbreviations

API: application programming interface
CMR: cardiovascular magnetic resonance
GTM: generative topographic mapping
HR: hazard ratio
LA: left atrial
LAEF: left atrial ejection fraction
LAMV: left atrial maximum volume
LV: left ventricular
LVEDV: left ventricular end-diastolic volume
LVEF: left ventricular ejection fraction
LVESV: left ventricular end-systolic volume
LVGLS: left ventricular global longitudinal strain
LVMM: left ventricular myocardial mass
LVSV: left ventricular stroke volume
ML: machine learning
TRIPOD: transparent reporting of a multivariable prediction model for individual prognosis or diagnosis

XSL•FO
RenderX

Edited by J Wang; submitted 06.May.2025; peer-reviewed by C Yang, L Dall'Olio; revised version received 09.Mar.2026; accepted 24.Mar.2026; published 27.Apr.2026.

Please cite as:
Yuan K, Kong D, Zhong J, Xie M, Liu R, Sun W, Liu X
Machine Learning–Derived Cardiovascular Aging Phenotypes From Cardiac Function and Stroke Risk in the UK Biobank: Cohort Study
JMIR Aging 2026;9:e77017
URL: https://aging.jmir.org/2026/1/e77017
doi:10.2196/77017

© Kang Yuan, Deyan Kong, Jinghui Zhong, Mengdi Xie, Rui Liu, Wen Sun, Xinfeng Liu. Originally published in JMIR Aging (https://aging.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Expectations, Ethical Dimensions, and Values for Voice-Enabled AI Assistants That Support Older Adults and Informal Family Caregivers: Participatory Speculative Design Study

Shweta Singh, MSc; Aqueasha Martin-Hammond, PhD

Human-Centered Computing Department, Luddy School of Informatics, Computing, and Engineering, Indiana University Indianapolis, 535 W Michigan Street, IT 585, Indianapolis, IN, United States

Corresponding Author:
Aqueasha Martin-Hammond, PhD
Human-Centered Computing Department, Luddy School of Informatics, Computing, and Engineering, Indiana University Indianapolis, 535 W Michigan Street, IT 585, Indianapolis, IN, United States

Abstract

Background: Continuous advancements in voice artificial intelligence technologies aim to assist older adults and caregivers, potentially improving quality of life and reducing caregiving burdens. Although research has explored the potential of voice-enabled artificial intelligence (VAI) assistants, such as Alexa (Amazon.com, Inc) and Google Home, to support older adults' health in informal care settings, there remains a significant gap in understanding the ethical dimensions and values that may influence their future adoption by caregivers and care recipients.

Objective: This research aims to explore older adult and informal family caregivers' perspectives of VAI assistants for supporting informal care, including the ethical dimensions and values that influence their decisions about future adoption for these purposes.

Methods: This research uses participatory speculative design to explore older adults' and informal family caregivers' perspectives of how VAI might support informal care in the future, and the ethical concerns they have about adopting VAI technologies. We conducted 8 workshop sessions with older adults and caregivers (n=9) over four months. Each phase focused on one of three goals: (1) to understand existing experiences, (2) to envision future VAI technologies, and (3) to reflect on ethical values that shape acceptance. In workshops, we aimed to gain insights into their experiences and challenges in managing informal care tasks and how future implementation of VAI might support the caregiving process to address their needs and concerns while emphasizing the ethical dimensions they value.

Results: The findings suggest older adults and informal family caregivers see potential opportunities for VAIs to support informal aging care by automating daily health tasks to improve efficiency, enhancing mental health and well-being, and offering companionship. However, participants felt that VAI alone might not be sufficient to address the complex needs of informal care. Additionally, they raised several ethical concerns related to transparency, privacy, inclusiveness, trust, affordability, and autonomy, which they felt needed to be addressed to encourage adoption of VAI technologies for informal care in the future.

Conclusions: Based on the findings, we offer insights and design implications for VAI systems that balance efficiency with ethical values to support diverse caregiving needs and potentially encourage future adoption in the informal care space.

(JMIR Aging 2026;9:e79740) doi:10.2196/79740

KEYWORDS

older adults; voice AI; voice assistants; ethics; design; human-computer interaction; participatory design; artificial intelligence

Introduction

Background

Often unpaid, informal family caregivers play an important role in helping older adults age in place by assisting with activities of daily living and other health and wellness needs [1,2]. As more caregiving tasks move into the home, older adult-informal care relationships become increasingly important in supporting older adults in maintaining and achieving positive health outcomes [3]. At the same time, voice-enabled artificial intelligence (VAI) assistants (eg, Alexa or Siri [Apple Inc]), which leverage voice input/output and artificial intelligence (AI), are becoming increasingly prevalent in consumer homes and show promise for supporting older adults and their caregivers [1,2,4]. VAI assistants are generally helpful for addressing various accessibility needs among older adults [5-7]. However, research also indicates that older adults and informal caregivers see VAIs as potentially useful for supporting health

XSL•FO
RenderX

care and other care needs at home [8-10]. For instance, VAIs have been found to be beneficial for informal caregivers and older adults with mild cognitive impairment [8], informal care dyads [10], and for scaffolding conversations between older adults and informal caregivers [9]. Thus, VAIs present a promising approach for enhancing the quality of life for older adults while helping to reduce some of the burdens faced by caregivers.

Significant prior research highlights the potential of VAIs to improve health care management and quality of life for older adults [10-12]. For older adult caregiving, research has also focused on understanding desired VAI features [4] and the benefits for tasks such as medication reminders, appointment scheduling [8,13], and improving communication between older adults and their informal caregivers [9,10]. Yet, ethical concerns such as privacy and transparency remain barriers to the acceptance and adoption of AI technologies such as VAIs generally [14,15], and these concerns are heightened within the health care context [16], where safety and confidentiality are paramount. Studies acknowledge that it is crucial to understand ethical considerations such as trust and transparency when considering AI-augmented health care interventions [17,18], calling for more emphasis on user values in health-related AI design and implementation [16]. However, little research has specifically explored the ethical values and dimensions that caregivers and older adults prioritize regarding VAI interactions in an informal caregiving context and how they might impact future adoption. Due to the diverse values individuals may have about AI use in health care [16], the growing ubiquity of VAIs for older adults' health [10-12], and the importance of exploring values in different contexts [19], it is important to explore ethical values specific to the older adult informal caregiving context.

In this study, we use participatory speculative design to explore older adults and informal family caregivers' ethical concerns about adopting VAI technologies. We aimed to gain insights into their experiences and challenges in managing informal care tasks, how future VAI features might support the caregiving process, and the ethical dimensions they value, including how those dimensions might impact potential acceptance. Based on our findings, we provide the following contributions: (1) We provide extended user perspectives on the role of VAIs in addressing informal caregiving needs, including the socioemotional aspects and burdens of informal care. (2) We highlight the ethical dimensions and values important to older adults and caregivers when considering accepting and adopting VAI assistants and the need to balance efficiency and ethical values in design to do good while doing no harm.

Prior Work

We provide background on the importance of informal, unpaid caregiving for supporting older adults' health and wellness needs, noting VAI assistants as emerging tools for supporting older adults and caregivers. We discuss open challenges of VAI acceptance and adoption for health and concerns raised in prior work that limit the adoption of VAI for health among older adults.

Informal Caregiving and VAI to Support Older Adults' Care

Informal caregiving usually involves unpaid assistance with activities of daily living and other health tasks [2] by someone, usually a family member or a close friend, who is outside of the formal health care system [20]. As informal caregivers, family or close friends take on responsibilities of looking after another's health and daily needs in addition to their own [2]. Informal caregiving tasks can include helping with feeding or personal hygiene, administering medication, accompanying care recipients to medical appointments, and fulfilling other essential needs in their daily lives [2]. For older adults, informal caregivers can play a significant role in maintaining their health when they are physically or mentally unable to perform daily tasks and self-care activities [3]. This form of caregiving can be physically and emotionally demanding, leading to exhaustion, neglecting one's health, and financial strain among those who provide care [2,3,20]. Due to the demanding nature of caregiving for older adults, assessing and reducing the workload for these individuals overseeing their care is crucial.

One technology, VAI assistants, has emerged in recent years as a potentially accessible support for older adults and their caregivers. Recent studies [10-12,21-23] have suggested that including VAI assistants such as Alexa or Google Home within households has the potential to improve health care management and quality of life for older adults. There has also been an emerging focus on the potential of physiologically expressive voice-based conversational agents to deliver valuable social support in health care contexts [24,25]. In addition to these individual care tasks, researchers have also explored how VAI might support informal caregiving environments, focusing on supporting caregiving relationships and collaboration. Research has indicated that despite some open concerns, both hired (eg, paid) and family (unpaid) informal caregivers felt that VAI systems could benefit home care settings, including supporting tasks related to documentation, care coordination, and person-centered care [13]. Researchers have found that VAI assistants are viewed as potentially useful for supporting older adults and their caregiver dyads generally [10], as well as for supporting the care of those with mild cognitive impairment [8]. Rudnik et al [9] also found VAIs useful for helping older adults and caregivers scaffold care conversations. Therefore, there is evidence that older adults and caregivers feel VAI might help support informal caregiving environments. Yet open challenges exist.

Ethical Concerns in Voice Activated-Augmented Care

Some of the challenges of VAIs in health relate to unmet needs for desired interactions that are tailored to specific caregiving needs and tasks. For example, Corbett et al [10] found that while older adults and caregivers found existing voice assistants useful, they wanted additional features to help them with health and medical information. Zubatiy et al [8] found preferences for more assistance with activities of daily living. Therefore, while virtual assistants for older adults represent a notable step forward in developing socially adept conversational agents, there is a consensus that further refinement is necessary, especially in the health care context. Part of this refinement aims to ensure that

XSL•FO
RenderX

these technologies are equipped to meet the demands of diverse real-world scenarios and align more closely with the user expectations and accessibility needs required to provide reliable interactions [11,22,24,26-28]. AI-enabled conversational health assistants such as VAIs have demonstrated their potential for accommodating individual preferences, particularly in situations where sensitivity plays a pivotal role, underscoring their relevance in health care and other contexts where tailored interactions are crucial [29]. However, while acknowledging these potential advantages, older adults and caregivers have also voiced ethical and trust concerns when leveraging AI-enabled health tools more generally [17], and these concerns have been found in VAI health research.

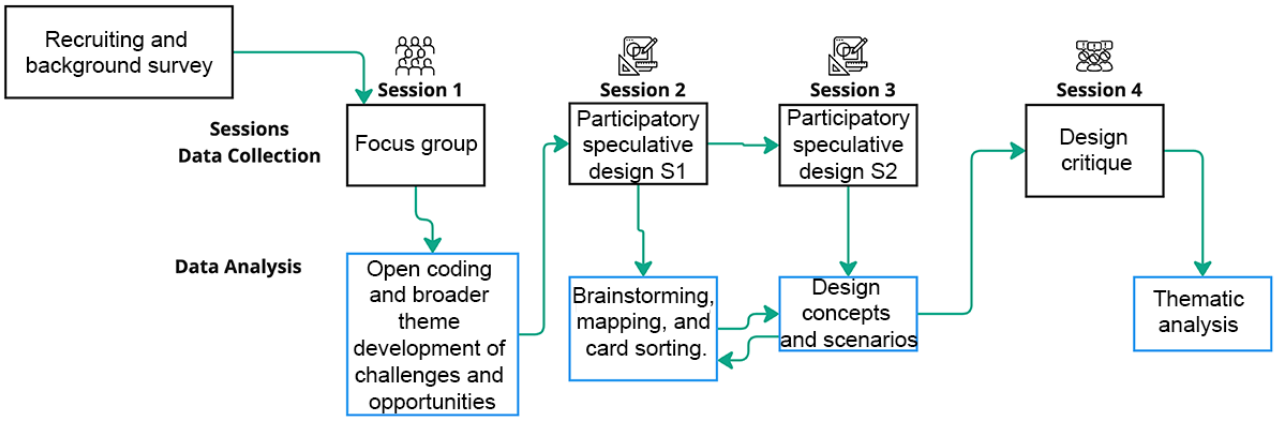
Ethics seeks to determine what is right, just, and morally acceptable within a given context or society [14]. Ethics in AI is a complex and evolving field, with convergence around 5 key principles: transparency, justice and fairness, trust, responsibility, and privacy [14]. Examining ethical components and their value when designing AI systems is thought to be the foundation of any successful interaction between humans and AI systems. VAI assistants have been found to raise ethical concerns with users [15,19,30-32]. The systematic examination of ethical concerns regarding voice assistants by Seymour et al [19] (eg, Amazon Alexa and Google Home) more broadly found that privacy was the most cited concern, yet other concerns related to accessibility, social interaction (eg, anthropomorphism and emotional connection to machines), and social order (eg, impact on parent-child relationships) were also discussed. Similar concerns apply to VAIs used within a health care context; however, examining ethical values is particularly important due to health care's safety-critical and privacy-sensitive nature [33]. For example, research has shown that the degree of health information sensitivity can impact users' preferences for interactions with VAI assistants such as Google Home [34]. Patients tend to prioritize trustworthiness and transparency when engaging with health-related assistants [21], and different aspects of trust, such as credibility and service quality, have also been found to impact perceptions of health-related VAIs and users' intentions to use them [33]. Therefore, though AI-driven assistive and voice-based interfaces offer substantial benefits, studies suggest an imperative need for tailored approaches that prioritize differences in context and individual and group experiences to establish trust and ethical standards [17], including exploring participatory methods to democratize design, expand who is included in the design, and prioritize usefulness [33,35].

Methods

Study Design

We conducted a 3-phase study (Figure 1) using participatory speculative design to explore how future VAI assistants might support informal caregiving practices among older adults and caregivers, as well as the values influencing the potential adoption of these technologies. Each phase addressed one of three goals: (1) understanding existing experiences, (2) envisioning future VAI technologies, and (3) reflecting on ethical values that shape acceptance. This study focused on the following research questions (RQs): (RQ1) How do participants perceive the role of VAIs in supporting caregiving tasks? (RQ2) What values do caregivers and older adults have around VAI assistants that support day-to-day care tasks? (RQ3) What ethical concerns do caregivers and older adults have about VAI assistants that support day-to-day care tasks?

Figure 1. Image of the participatory speculative design process. The process included 4 workshop sessions, including a focus group, two participatory speculative design sessions, and one design critique session.



Ethical Considerations

The Institutional Review Board at Indiana University reviewed and approved this study as exempt (Institutional Review Board application #21351) before recruitment or data collection began. In the first workshop session (session #1), we provided participants with a printed study information sheet that explained this study, the data we would collect, their rights as participants, and our procedures for minimizing risks to privacy and confidentiality. We allowed time for participants to read the sheet and also explained the information verbally. This study's information sheet included details on payment procedures (US $35 per session completed), principal investigator contact information, and how to contact our Institutional Review Board office. Participants were also informed of efforts to keep their personal information confidential, including restricting access to their data to only the research team and destroying identifiable data (such as audio or video recordings) at the end of this study. They were also told that their data might be shared with the University or the federal agency funding the research, if required

XSL•FO
RenderX

by law. Participants were informed of their right to withdraw from this study at any time. In accordance with our university's procedures for exempt research, participants provided verbal consent to continue after reviewing and asking questions about this study's information sheet.

Rationale for Participatory Speculative Design

While voice-activated technologies such as Siri and Alexa are currently available to consumers, they are often designed primarily for non–health-related uses. Therefore, while the current study of these tools for health is growing as an emerging technology, they are not commonly leveraged in health settings due to open concerns about safety [36] and privacy, and are typically not widely available outside of a research context. We chose speculative participatory design as an approach because it provided the flexibility for participants to imagine a future for these technologies in a health context, specifically for supporting informal, aging care, and to reflect on the ethical considerations around future adoption and integration within the home and informal care environment.

Speculative design offers users a way to speculate about future AI technologies and enables researchers to analyze those speculations to uncover values [37], which is a powerful approach for exploring ethics in the rapidly developing field of AI design [38]. The proposed method intersects with other human-computer interaction (HCI) methods, including participatory design [39], value-sensitive design [40], and design fiction [37]. While each of these approaches has been found useful for designing AI systems [21,41], speculative approaches have been useful for helping users envision future AI systems and the corresponding values that may not yet exist, or for imagining possible new futures [31,37,42-44]. Formally, speculative design often leverages the collective strengths of each method to use fictional stories to "collectively envision, design, and critique future AI agents," thereby allowing researchers to understand how people imagine future technologies and the values they consider important when designing them [37]. As an approach, it allows for the imaginative exploration of "future configurations of technology and society within the world" by adopting both creative and critical perspectives, including critiquing existing AI practices [38,45].

Recruitment and Participants

We recruited 9 participants, including 3 older adults and 6 caregivers, for a 4-month study (January-April 2024). The criteria for caregiver participation were that individuals be 18 years of age or older and serve as an informal, unpaid, family or friend caregiver to an individual 65 years of age or older. The criteria for older adult participation were that the individual be 65 years of age or older and be the recipient of informal (eg, family or friend), unpaid care. Additionally, because of our goal of speculative co-design of VAI technologies, we required that participants be currently using or have an interest in potentially adopting or using these technologies in an informal care setting in the future.

Our recruitment method was purposeful, aligning with our need for participants with specific caregiving experiences and interests in using and designing VAI assistants for that space. During recruitment, we found that 4 of 6 of our caregiving participants also self-identified as older adults. In those cases, we provided participants the flexibility to choose which role they would like to take on for the workshop sessions, letting them know that they could also share their perspectives as older adults receiving care. However, we did experience participant dropout over the 4 months. A total of 7 participants attended all 4 sessions. Further, 1 caregiver attended only 2 sessions in their series, and 1 caregiver attended only 3 sessions. Participants did not continue due to busy schedules and other time commitments (Table 1).

Table . Participant demographics and participation in study sessions.

Participant	Age group (years)	Sex	Caregiver or older adult	Sessions participated
P1	35 - 45	Female	Caregiver	S1, S2, S3, S4
P2	55 - 65	Female	Caregiver	S1, S2, S3, S4
P3	65 - 75	Male	Caregiver	S1, S2, S3, S4
P4	65 - 75	Female	Caregiver	S1, S2
P5	65 - 75	Female	Caregiver	S1, S2, S3
P6	65 - 75	Female	Caregiver	S1, S2, S3, S4
P7	75 - 85	Female	Older adult	S1, S2, S3, S4
P8	75 - 85	Male	Older adult	S1, S2, S3, S4
P9	75 - 85	Female	Older adult	S1, S2, S3, S4

Informal caregiver participants served primarily as family caregivers, including spouses, children, and siblings of the care recipient (Table 1). Their ages ranged from 35 to 75 years, with the majority also self-identifying as older adults (65-75 years). Most caregivers were primary caregivers (ie, the responsible party for providing care), although some occasionally received additional help from others. Caregiver participants supported individuals with a wide range of conditions, including dementia, chronic diseases, amputations, and cancer. Their caregiving durations spanned from 6 months to more than 10 years. The type of care tasks they assisted with involved primarily instrumental activities of daily living (eg, housework, transportation, and managing finances), companionship and emotional support, routine activities of daily living (eg, bathing,

XSL•FO
RenderX

toileting, and eating), and medical and nursing tasks, such as injections and colostomy care.

Older adult participants ranged from 75 to 85 years of age and had children or spouses as primary caregivers. They had from 2 to 5 doctors and nurses on their health care team in addition to 1 - 2 other caregivers who help them out with additional health tasks either daily or a few times a year. Older adults and caregivers were familiar to very familiar with voice assistants such as Siri and Alexa.

Participatory Design Sessions

Overview

Each group of participants (older adults and caregivers) engaged in 4 workshop sessions over the 4 months (Figure 1). Therefore, we conducted 8 group workshops in total, 4 with informal caregivers and 4 with older adults who receive care from family or friends. Topics and session progression were similar for both groups. Workshops were conducted online and in person. Initially, workshops were planned for in person; however, to accommodate participants with mobility and other challenges that prevented them from attending in person, we held some sessions online. During in-person workshops, participants engaged in activities such as card sorting, collaborative sketching, and crafting scenarios through role-play and discussion (Figure S2 in Multimedia Appendix 1). In-person and online workshops were conducted similarly, using online tools including videoconferencing software with breakout rooms and online collaboration tools (eg, Miro) to pivot and replicate activities for online participants. Each workshop session lasted approximately 1 - 1.5 hours (60 - 90 min), and participants were paid US $35 per session.

Session 1: Understanding Current Informal Caregiving Experiences

In the first workshop session, we aimed to create a common foundation for envisioning future technologies by understanding participants' experiences with caregiving, AI, and VAI technologies. We sought to gain a comprehensive understanding of the dynamics of informal family caregiving by engaging both caregivers and care recipients in an initial focus group to explore experiences shaping their caregiving values. We asked them to describe their daily routines, experiences, challenges, and key areas for improvement.

We also introduced AI and VAI assistants, sharing definitions of AI, its capabilities, and common VAI scenarios. We provided examples of consumer-facing VAI assistants, such as Siri, Alexa, and Google Assistant, and facilitated open discussions to address questions and assess participants' existing knowledge. We noted key points about experiences or challenges with caregiving or VAIs on sticky notes to prepare for session 2. The focus groups lasted approximately 60 minutes.

Sessions 2 and 3: Imagining VAI Informal Caregiving Futures

The second and third sessions aimed to envision VAI assistants that support informal caregiving. We used participatory speculative design to conceptualize AI voice applications that participants believed could help them in informal caregiving settings (Figure 1). In session 2, the group reviewed and validated notes from session 1 about participants' caregiving experiences and concerns. Participants used affinity mapping and brainstorming to identify and prioritize how VAI could support their informal caregiving experiences. To help prioritize initial ideas for further discussion, participants used card sorting to rank ideas and reflected as a group on (1) the potential feasibility of each idea and (2) how it might meet their needs.

After brainstorming and card-sorting, we used prompts, as shown in Figure S1 in Multimedia Appendix 1, to engage participants in co-design activities, such as scenario-based role-playing. Using these prompts, they brainstormed and created concepts for VAI interfaces suited for informal caregiving. During the third session, we continued these activities to complete the design concepts by creating a final set of scenarios. Consistent with the principles of speculative design [37], our goal was to explore values through the design of future technologies within the context of participants' narratives of use. While emphasizing VAI applications to keep discussions grounded in participants' stories of future use, sessions 2 - 3 also aimed to understand why participants made specific design choices, thereby enhancing our insight into their ethical values. Each workshop lasted about 90 minutes.

Session 4: Ethical Critiques of VAI for Informal Care

In the fourth session, we asked participants to revisit the concepts and scenarios they generated in sessions 2 - 3 to reflect more deeply on their ethical values. Some ethical values emerged naturally in sessions 2 - 3 as participants discussed ideas; however, in session 4, we asked participants to deliberately critique their scenarios through the lens of different ethical dimensions (eg, privacy, transparency, trust, and user autonomy) [14], consider potential challenges that could restrict their adoption, and discuss any needs related to various ethical dimensions, whether positive or negative. We also encouraged them to discuss additional issues that might not have been listed, using their own words, which we later categorized during analysis. The workshop lasted approximately 90 minutes.

Data Analysis

Each focus group and workshop session lasted 60 - 90 minutes. All sessions were recorded with the participants' permission. Data included transcripts generated from over 660 minutes of audio recordings from 8 workshops (4 with older adults and 4 with caregivers), researcher notes, and physical artifacts. While we did experience dropout during recruitment, resulting in a smaller sample, we observed recurring themes (thematic saturation) [46]. We used thematic analysis [47] to analyze the data. To conduct the analysis, the first author reviewed transcripts using memos to highlight interesting aspects of the data. Using a reflexive thematic process [47], we then used inductive and deductive coding to identify recurring themes from the data, focusing both within and across participant groups. Inductive coding was initially used to identify codes related to questions about caregiving experiences, opinions about VAIs, and discussions of ethical concerns. We later used deductive coding to further categorize ethical concerns into existing ethical dimensions: transparency, trust, privacy, justice and fairness, and freedom and autonomy (Table S1 in

XSL•FO
RenderX

Multimedia Appendix 1), aligning with known AI ethical dimensions identified in prior literature [14]. We met iteratively [48] to synthesize data, discussing and categorizing codes into themes. We used an axial coding process to organize the identified codes into broader categories. Analysis of data from participatory design activities, such as lists of caregiving challenges and VAI ideas, was taken directly from data generated during those activities and was refined by the research team using affinity diagramming to merge similar themes across workshops. Three high-level themes emerged: (1) caregiving experiences, (2) opinions about VAIs for caregiving, and (3) ethical dimensions and values aligning with study goals. A detailed summary of themes and subcodes is included in Table S2 in Multimedia Appendix 1.

Results

Overview

Findings from this study shed light on the complex interplay of emotional, physical, and logistical difficulties experienced by older adults and caregivers when engaging in caregiving tasks, as well as their needs and challenges, and their ideas for how VAI might support informal care in the future. Our findings also highlight tensions regarding the potential benefits of VAI and ethical concerns associated with its adoption in a care environment.

Envisioning VAI for Informal Care Support

Caregiving Experiences

From session 1, we learned that caregivers and care recipients represented varied caregiving experiences. However, some caregivers and care recipients have overlapping experiences of care. Aligned with HCI work exploring technology experiences of smaller numbers of participants [49], we present these experiences as personas. Our personas were based on real data and represent a grouping of common experiences across different participants. This method was used in the study by Waycott et al [49] exploring nontech use, where participant numbers were low (10 total) to balance sharing data about experiences while respecting participant confidentiality.

Our participants shared how they came to need care and the emotions surrounding that transition. Caregivers' experiences included (1) long-term caregivers beginning to care for an ill parent as a child into adulthood and older age, (2) partner caregivers who became primary caregivers of a partner due to a life-altering event, and (3) sandwich generation caregivers who were primary caregivers to both older parents and younger children (Figure 2).

Care recipients had similar varied experiences ranging from (1) needing care after an unexpected medical event, (2) needing care to support chronic illness management, and (3) needing shorter-term care after a serious medical procedure (Figure 3).

XSL•FO
RenderX

Figure 2. Caregiver personas representing experiences providing care for an older adult friend or family member.



The long-term caregiver

Margaret is in her mid-forties and is married with two children: a teenager who is 17 and a pre-teen who is 12. Margaret has been taking care of her parents since she was young. At 10, her mother was diagnosed with cancer, and as the eldest daughter, she began helping to care for her. Shortly after, her father was diagnosed with chronic heart failure and, due to his diabetes, went through multiple amputations. She dedicated her time to fully becoming a caregiver for her parents. Her responsibilities included assisting with daily tasks, managing medications and medical equipment, accompanying them to medical appointments, and ensuring their overall well-being. Being the only caregiver, she also took care of specific tasks like changing her mother's catheter, which was an unpleasant experience for her. She also, over time, faced the challenge of persuading her aging parents to accept necessary changes in their lifestyle, such as giving up driving or wearing diapers, without making them feel demeaned or dependent. Since she took the role of a caregiver from a very young age, it also took an emotional and physical toll on her. She went through emotional turmoil and stress and barely got time for herself. Moreover, she often suffers from back problems as well. Despite the discomfort and difficulty, Margaret perseveres, driven by her deep sense of duty and love for her family.



The partner caregiver

Jane is in her 50s and is married to her long-term partner, Tim. Recently, Tim had a stroke, which caused him to become paralyzed on his left side, making it difficult for him to get around. Before the stroke, Tim worked full-time and contributed to household expenses. However, after the stroke, Tim can no longer work and get around on his own. Jane has continued working at her job but has had to reduce her hours to become Tim's primary caregiver. With Tim unable to work, Jane faces the burden of being the sole breadwinner for the household. Reducing her work hours to care for Tim further impacts their financial stability, adding stress to an already difficult situation. She is also responsible for helping Tim get around to his appointments and managing his medication regimen. Balancing caregiving duties with her job responsibilities has become a daily struggle for her. Witnessing her partner's sudden disability and the loss of his independence is emotionally draining for her. Jane seeks moments of detachment, acknowledging the necessity of not fixating solely on one task or challenge as she finds clarity and balance amid the chaos of caregiving.



The sandwich generation caregiver

Betty is a 35-year-old woman who has a 5-year-old daughter. Recently, her mother was diagnosed with a leaky heart valve, and she feels her world has been turned upside down. This diagnosis not only brought a wave of worry and concern but also a significant shift in Betty's daily life. She finds herself sandwiched between the needs of her child and the increasing demands of caregiving for her ailing mother. She turned her daughter's room into a makeshift hospital room equipped with medical apparatus and necessities to accommodate her mother's condition. Betty encounters resistance from her mother, who struggles with relying on others for physical assistance. The shift from being cared for to becoming the primary caregiver strains their bond, as Betty longs for the comforting presence of her mother. She faces a relentless struggle to secure resources for caregiving tasks, from fixing wheelchairs to deciphering insurance policies and acquiring vital medical equipment for her mother's care. The demands of caregiving often leave Betty feeling isolated from friends, family, and her community.

XSL•FO
RenderX

Figure 3. Older adult personas representing experiences collaborating with a friend or family member for informal care.



Donna is 66, and her life took a sudden and unexpected turn when she was involved in a severe accident that left her with a metal rod in her leg, extending almost to her knee. This life-altering event added a new layer of challenges to Donna's already complex medical history. Ten years before the accident, Donna had faced another significant health battle when she underwent radiation treatment. The emotional toll of battling cancer was immense, and during this challenging time, Donna leaned heavily on the support of her husband, who served as her primary caregiver. Alongside him, Donna also received care from 3-4 professional caregivers. One of the most significant hurdles she faced was the loss of control and autonomy over her own body. Donna grappled with a deep-seated fear of the unknown, particularly regarding her recovery and what the future held in store. *What's gonna happen? Am I gonna get better?* - she thought to herself. During these times, she relied on her friends and connected with them often to chat, and it helped her with her anxiety.

Care after an unexpected medical event



Shawn, a man in his 60s, had always prided himself on his independence and self-sufficiency. He had lived a full life, working hard, raising a family, and cherishing every moment spent with his loved ones. However, when he was diagnosed with a chronic illness, cancer, his world was suddenly turned upside down. Shawn was fortunate to have the unwavering support of his daughter and granddaughter by his side. They became his primary caregivers, providing him with the physical and emotional support he needed to navigate his illness. Whether helping him with daily tasks, accompanying him to medical appointments, or simply offering a listening ear, their presence brought comfort and reassurance during tough moments. One of the most significant challenges Shawn faced was the loss of his independence. As someone who had always been fiercely independent, relying on others for assistance was a bitter pill to swallow. The feeling of helplessness and the loss of control over his own life weighed heavily on him, leaving him feeling frustrated.

Care after a chronic illness diagnosis



Annie's life took an unexpected turn when she underwent a major surgery—a hysterectomy—that left her physically weakened and dependent on the support of her three children and husband. Her three children and husband rallied around her, becoming her pillars of strength during this difficult time. From helping her with daily tasks like bathing and dressing to ensuring she took her medication on time, they spared no effort in ensuring her comfort. The stress of managing illness weighed heavily on Annie's shoulders as she navigated the complexities of her recovery. The most challenging part for her was adjusting to a new reality where she needed assistance with even the simplest tasks. The inability to do things quickly or independently left her feeling frustrated and restless. She missed driving and being able to do grocery shopping herself. Furthermore, Annie felt the pressure to maintain a positive outlook for her family and loved ones, even when she struggled internally.

Care after serious medical procedure

Both older adults and caregivers appreciated the caregiving relationship but noted challenges. For caregivers, most challenges were related to managing and coordinating care, including adjusting to their role, navigating the health care system, and accessing resources (Textbox 1). However, they also shared emotional and social challenges such as feelings of isolation, lack of support, and coping with feelings of being overwhelmed. They also shared that informal caregivers are often not trained; therefore, finding resources and training on different care-related tasks, such as navigating financial policies or communicating sensitive topics, made things more difficult. Textbox 1 lists the range of challenges caregivers listed during the session.

XSL•FO
RenderX

Textbox 1. Caregivers' caregiving needs and challenges.

- Lack of resources and information
  - Scarcity of external support networks
  - No access to information about available services
  - Potential loss of income due to caregiving responsibilities
  - Lack of understanding about financial policies
  - Shortcomings of training programs
  - Navigating health care systems and tasks
  - Managing medical appointments and tasks
  - Managing medical equipment and administering medication
  - Convincing older adults to accept help, medication, or medical devices
  - Advocating for the care recipient
- Emotional and psychological toll
  - Feelings of sadness, frustration, and guilt
  - Enduring emotional strain in prolonged caregiving
  - Witnessing decline in health and independence of loved ones
  - Lack of time for personal reflection
  - Dealing with the unpredictability of caregiving tasks
- Isolation and social dynamics
  - Transitioning roles within relationships or role reversal
  - Experiences of social isolation
  - Managing family dynamics
  - Balancing caregiving duties with other commitments
  - =Lack of self-care
- Communication and sensitive interactions
  - Communicating sensitive topics
  - Difficulty finding solutions to specific issues
  - Convincing older adults to accept help, medication, or medical devices

Older adults shared difficulties they had in transitioning to needing care. They discussed difficulties with losing independence, feeling like a burden, and needing to depend on others (Textbox 2). They shared that the transition to needing care came with a lot of emotional tolls, including adjusting to feelings of social isolation and loneliness, and frequently needing to balance positive and negative feelings that come along with needing care. Textbox 2 lists the range of challenges reported by older adult participants during the session.

XSL•FO
RenderX

Textbox 2. Older adults' caregiving needs and challenges.

- Loss of independence and autonomy
  - Loss of independence
  - Dependence on others
  - Sense of loss of control or autonomy
  - Asking for help with basic tasks such as picking up items from the floor
  - Not being able to drive
- Physical and emotional toll
  - Managing the stress of illness or injury
  - Physical limitations
  - Emotional toll
  - Feeling impatient because of not being able to do a task or do it quickly
- Sense of isolation and social disconnection
  - Sense of isolation
  - Fear of future and uncertainty about what lies ahead
  - Feelings of burdening or bothering others for assistance
- Feelings of burden and responsibility
  - Burden of trying to be better for others
  - Feelings of burdening or bothering others for assistance
- Balancing negative and positive emotions
  - Sense of kindness (despite struggles)
  - Sense of kindness (toward those providing help)

During the first session, participants reflected on their caregiving experiences and needs, prioritizing the challenges that VAIs might address. During sessions 2 and 3, caregiver and older adult participants brainstormed different ideas for how VAI assistants might augment and support them as caregivers and care recipients and address the challenges they experienced.

Many initial design ideas from caregivers and care recipients emphasized how they felt VAI could help make caregiving tasks more efficient, or how VAI might help shift caregiver time spent on physical or informational caregiving tasks toward more social and emotional ones. For example, during session 1, participants shared that informal caregivers often lack professional caregiving experience and frequently encounter unfamiliar situations, navigating uncharted territory. Therefore, when asked to prioritize how they felt VAI could address concerns listed from the initial focus group session, caregivers ranked their need for caregiving support as a top priority. They discussed how VAI might help them be more efficient by supporting them with (1) immediate access to relevant information and resources, (2) proactive assistance with caregiving tasks and resources, (3) persuading older adults to accept help, and (4) connecting for socioemotional support as top priorities. A full list of brainstormed ideas is included in Table S3 of Multimedia Appendix 1.

While informal caregivers generated more ideas in their sessions, some overlapped with those of older adults. As with caregivers, older adult participants saw opportunities for VAI to support health information needs with caregiving relationships and to provide socioemotional support for them and their caregivers. Older adults' ideas primarily focused on VAI providing support through (1) immediate access to relevant information and resources, and (2) connecting for socioemotional support as top priorities. A full list of brainstormed ideas is listed in Table S4 in Multimedia Appendix 1. The following subsections discuss the primary VAI concepts that caregivers and older adults prioritized for supporting informal care.

Providing Immediate Access to Information and Resources

Both older adults and caregivers felt that features of VAIs, such as voice input and automation, could help to make health information and other caregiving tasks more efficient. Caregiver participants shared that they manage different types of health information, ranging from supporting the care recipient's daily health care needs to navigating insurance coverage. For example, one participant gave an example of navigating insurance as a caregiver, C3, W2, "Some things get covered [by insurance]. Some things don't... You don't find out until somebody screams [at you] in the financial section of the hospital." They felt VAI

XSL•FO
RenderX

could help them find and navigate this and other information more easily and efficiently.

Older adult participants also expressed interest in using existing VAI to support health information needs, such as receiving reminders and alerts for appointments or events, and to provide efficient access to information for managing their health. They recognized the role VAI could play in helping them stay organized and on track with their schedules. They felt that voice-activated information requests, similar to voice-based Google searches, could be convenient for accessing information more quickly and easily. While these initial ideas align somewhat with what current VAI assistants might do, many of the systems on the market are not tailored to support health or care needs.

Proactive Assistance With Caregiving Tasks and Resources

Older adult participants tended to discuss ways they foresaw VAIs helping them better manage their health tasks, such as scheduling or quick access to health information. Caregiver participants also envisioned the potential for a more sophisticated VAI information assistant (beyond what currently exists) that could talk to or nudge them proactively, helping them streamline their schedules and manage caregiving tasks. Participants envision this system using their existing health information to intelligently set reminders, manage appointments, and optimize their time, acting as a proactive decision support assistant.

Caregiver participants discussed the potential of VAIs to reason to anticipate future caregiving duties by efficiently providing relevant information about the older adult patient's health. One participant described a VAI that could bring together relevant information around a caregiving task to help them complete it. C1, W2 described, "So basically, it's [the VAI] gonna say like you need a ramp, you need a hospital bed, you need to make sure they have the wheelchair you need. So, it's gonna give you a list of all the stuff you need and potentially where you get it and who pays for it." Another participant suggested a VAI that could help troubleshoot problems by understanding patterns and symptoms to identify potential causes (Figure 4). C2, W3 explained, "If it could troubleshoot and identify what their areas of need are, if this person is up all night, what can you do? Are they going to the bathroom too much? Are they irritable? And then suggest: These are the steps that might be able to help them." They felt this additional support could be valuable in reducing their feelings of being overwhelmed when managing various caregiving tasks and responsibilities.

Figure 4. Scenario demonstrating desired interactions between an informal caregiver and VAI that provides advice or tips to help them address caregiving concerns. UTI: urinary tract infection; VAI: vvoice-enabled artificial intelligence.



Persuading Older Adults to Accept Help

Caregiving recipients discussed VAI's potential to support their caregiving routines through virtual companionship; however, their ideas centered on ways virtual companions might leverage human-like characteristics to build trust and persuade older adults they care for to accept assistance, medication, and treatments prescribed for them. One idea presented was using VAI to mimic the voice of a trusted person, such as a celebrity, to encourage acceptance of essential care, such as transitioning to using adult diapers. C1, W2 described, "I'm imagining Harrison Ford making a video saying, Hey, it's ok to do these things [wear an adult diaper]." Challenging this notion, another caregiver, C3, W2, who also self-identified as an older adult, offered a contrasting perspective, suggesting that "with old age, you get fairly tough skin, so a diaper is a diaper." They explained that older adults may not find discussing topics such as diapers stigmatizing over time. However, participants generally agreed that leveraging VAI to imitate others to persuade or provide comfort might be helpful in specific scenarios. C1,W3 shared, "...so say the person's having trouble going to sleep if maybe a book or a scripture passage was read in the caregiver's voice. I do think something like that in the caregiver's voice, I think that would be real helpful."

Connecting for Socioemotional Support

Opportunities to Support Mental Health and Well-Being

Both older adult and caregiver participants saw opportunities for VAIs to support their mental health and well-being. Caregiver participants tended to focus on how VAIs could support their mental health through reducing caregiving management tasks and through direct mental health support.

Supporting Caregiving Management

One set of ideas from caregivers focused on tools that could ease some of the mental burden of managing care. For example, one participant saw promise in the VAI for recording older adult patient requests, ensuring care continuity while reducing the burden of constantly being in "caregiver mode." Reflecting on their conversations with their mother, one caregiving participant

XSL•FO
RenderX

shared that their relationship with their mother shifted over time. They initially had typical mother-daughter interactions; however, because of her dual roles as caregiver and daughter, preserving the natural relationship became more difficult. Therefore, C3,W3 suggested using voice-recorded messages could be helpful, stating "later on it [my role] became more caregiver. And so when we'd wake up in the morning, instead of being like, 'Hey mom, good morning,' it would be like, 'I need this, and I need this, and I need that.' So if she could leave a message saying I need all this, then it could open up the morning into being, 'Hello, good morning. I love you,' and then I can check the message. So, the request can be via the system so that you still have some of that relationship left." Participants suggested enhanced VAI systems with more reasoning capabilities could also provide more significant support in caregiving situations. For example, they discussed VAI systems that proactively provide a tailored review of patients' conditions to first responders in emergencies. C1,W2 described, "If there was an emergency and you called the fire department and they showed up, maybe they'd be able to connect to it [referring to the VAI], and that would immediately tell them what the patient's conditions are, what medications they're on."

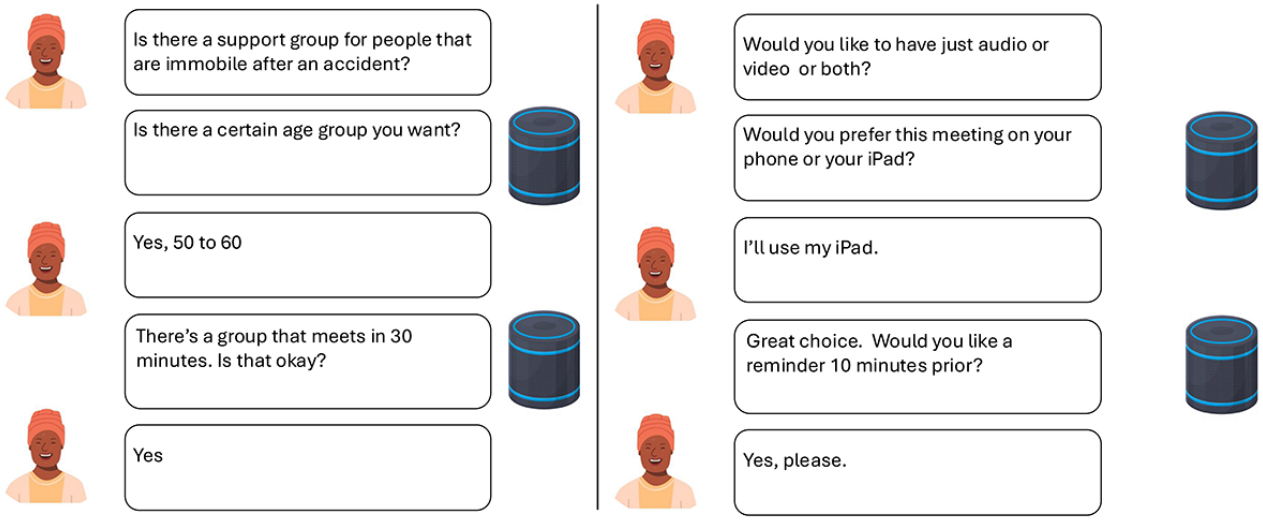
Supporting Caregiver's Mental Well-Being

Another set of ideas revolved around how VAI might directly support caregivers' mental health needs. Caregivers explained that being in a caregiving environment often comes with much stress, and therefore, they saw opportunities for VAI to support self-care and mental wellness. Participants suggested ideas for VAI to help them curate self-care plans, with occasional reminders to care for themselves or to complete self-care activities. C3, W2 shared, "It [the VAI] could give me a reminder to go do something for myself as self-care and maybe not for long periods of time because as a caregiver you don't have a lot of time ... also like have a self-care plan so that while it is reassuring, it can also walk you through yours." C3,W3 suggested using VAI as an emotional support, "I want this thing [VAI system] to comfort me in some way to. I want it to be like, look, I'm here to help you take a deep breath. It's gonna be OK." As shown in Figure S3 in Multimedia Appendix 1, participants brainstormed ways VAI systems that identify stress cues could offer positive affirmations or prompt caregivers to take breaks when needed. For example, the VAI could also converse with them to comfort them in difficult times. In addition, they explained that caregivers are often isolated, and the VAI could help connect them to support groups or people in similar situations.

Supporting Older Adult Self-Care

As with caregivers, older adults also felt VAI could help support self-care and well-being. They shared ideas for VAI systems to help them find leisure activities, viewing them as an accessible resource for quickly finding books, movies, or other media to enjoy as part of their well-being and self-care. Older adult participants, as with caregivers, also discussed generally using VAI for emergencies to establish instant connections to emergency services or loved ones, especially in times of crisis, when they expressed the need for an instant connection to emergency contacts. However, as with caregivers, most ideas centered around VAI as a facilitator for connecting with others (Figure 5). They shared ideas about the possibility of VAI helping them connect to their friends or loved ones through voice commands on a screen. O2,W2 explained, "I think [VAI] could be helpful in the emotional sense. If you're in bed and you don't wanna get up, you could always say call John." O3,W2 made similar suggestions, "It would be uh trained to be supportive of all their emotional needs and [when] you kind of shut down . she[the VAI] . on how much can we help this person who seems to have bouts of crying or they're really depressed and get to know them and see what the problems are and then [be] smart enough to try .and act out what would help that person get back to a stable emotional point." Participants also discussed the potential for VAI to connect them with people who have experienced similar situations automatically. O2,W2 later asked, "Is there a way that with AI, you could say connect me to the support group for people with recent car accidents, for instance."

Figure 5. Scenario representing a conversation between a person and a VAI assistant, discussing joining a support group for older adults experiencing mobility changes. VAI: voice-enabled artificial intelligence.



XSL•FO
RenderX

Supporting Older Adults Through Virtual Companionship

Older adult participants envisioned new ways in which VAI could enrich their lives as care recipients by providing companionship and social support. Participants envisioned new ways to engage in dialogue and interact with VAI as a virtual companion, which they felt could help alleviate feelings of loneliness during periods of solitude or immobility. Participants described a VAI system as an "AI girlfriend" (eg, a romance chatbot or robot). O3,W2 explained, "...they say [AI girlfriends] actually learn what they[the human partner] need and are looking for, and they're able to change how they respond to meet those needs. And if they can do that with AI girlfriends, I'm guessing they could absolutely do it with people that were depressed or down emotionally and actually learn what it is that that person is going through." However, participants also expressed apprehension about the possibility of humans developing emotional attachments to AI entities. O1,W2 posed, "The question is, can AI really meet someone's emotional needs?" At the same time, they expressed discomfort at conversing with a system rather than a human, stating, "Some of it [VAI] also scares me when they pretend, they're human or sound human."

Participants were also unsure whether a virtual companion should be embodied. One participant suggested embedding a VAI in robotic entities, such as robotic puppies equipped with AI to interact verbally and physically with individuals and to simulate lifelike behavior, such as mouth movements or eye blinking. In response, another participant shared that they preferred a discreet, compact device rather than a visually conspicuous one. They voiced concerns about potential discomfort associated with being constantly monitored by a visible robotic assistant, highlighting worries about their appearance or attire. 02,W2 mentioned, "I'm kind of stuck on something like a small device .I wouldn't necessarily want to see me . that'd be another thing to think about. Am I dressed or whatever? And I like to think that even though Alexa hears everything, she can't see anything." Older adult participants also considered that, while virtual assistants such as Alexa may capture audio, they lack visual capabilities. The idea of a small voice assistant device appealed to them because it is nonintrusive and occupies minimal space in a room.

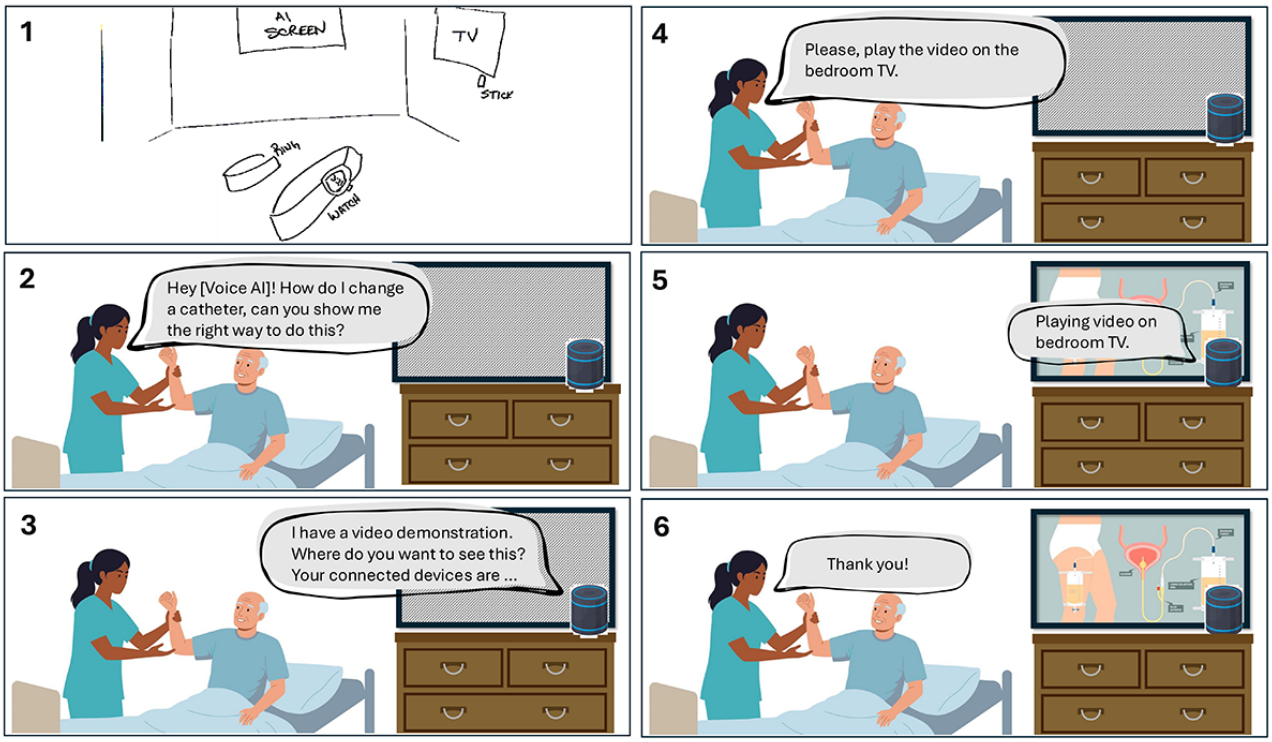
VAI Can Help, but it Is Not Enough

Older adult and caregiver participants viewed VAI as a potentially valuable tool to support their caregiving activities. However, participants also consistently pointed out that VAI would not be comprehensive enough to support the complex environment of caring for an older adult. Therefore, both groups also noted their desire for VAI to be considered broadly as one of several tools that might seamlessly integrate into a more extensive system. We learned that participants' reasoning for this perspective was related to several factors, including a desire to support various access and accessibility needs, to promote more integrated care, and to ensure fair access.

Participants suggested that VAI could be used as a tool to help them connect with multiple screens, noting that it may also be important for some users to have visual aids to enhance assistance and support varied accessibility needs. Additionally, participants noted that they sometimes preferred video demonstrations for tasks that might benefit from seeing them performed (eg, changing a catheter). They suggested that VAI-enhanced visual guidance might help them better perform the caregiving tasks (Figure 6). C3,W2 explained, "I would prefer having someone with a video showing me [for some tasks] every time [I was changing a catheter]. I like to see how things are done." These examples suggested that some participants preferred prioritizing the modality of information delivery that best supports the care task at hand, rather than defaulting to VAI alone.

XSL•FO
RenderX

Figure 6. Scenario representing a VAI caregiving ecosystem that connects other devices, such as smartwatches, televisions, and voice assistants, to coordinate various care tasks. On the wall, there is an "AI SCREEN" and a "TV" with an Amazon Fire TV Stick labeled "STICK." Next to it are a "RING" and a "WATCH" displaying the time. Followed is a storyboard of the VAI ecosystem scenarios participants discussed, showing a person interacting with a VAI to play a video demonstration of how to insert a catheter. AI: artificial intelligence; TV: television; VAI: voice-enabled artificial intelligence.



Caregivers also expressed the need for VAI to connect to multiple devices, enabling a more comprehensive approach to care support. C1,W3 discussed that it would be essential for the solution to have "different modalities because if something is also tracking their vital signs and everything, it has to be attached to the patient as well." They asked their group, "So what does that look like?" C2,W3 noted, "I personally would say again, things are gonna get smaller all the time, but like a ring, … or, you know, a watch. The necklace [fall necklace] got taken off much of the time. It got yanked. An Apple Watch?" C3, W3 explained, "The connectivity[has to be] built in with at least the phone." C1, W3 summarized, "Maybe this is like one program that can get attached to different equipment. Like your ring or watch or your TV, for that matter. And it can also attach to, like, you know, how Alexa can control your lights and everything. So, like something that can control different things differently, because, like it, when I was taking care of my mom, um, I'd have her watching TV, so I would want two separate units so that she could be watching TV while I was maybe watching something like a procedure." The participants drew a sketch illustrating their vision of a device ecosystem they would prefer to see integrated with VAI (Figure 6).

In addition to caregivers' perspectives, older adults emphasized the importance of VAI seamlessly integrating with the devices and technology they already use in their daily lives. For example, older adults wanted to give voice commands to the AI system to find information or pull up a book to read on their iPad (Apple Inc). They explained that the difference between their envisioned and existing VAI systems is that the primary agent assists in managing and connecting these different systems. They felt that integrating VAI could extend the flexibility and convenience of other devices, making care and wellness tasks easier for them overall. O2,W2 noted, "Maybe something that you can turn on if you want to talk, that could turn on through your television or through an iPad. So that if you want it to pull up something for you to read, for instance, Alexa, find me a good book on whatever. I mean, that's going pretty far out." In describing an interaction with the system, a participant demonstrated how they envision the VAI providing options when choosing a mode of interaction, for example, connecting them to a support group. Overall, participants view VAI as a tool within a broader ecosystem that supports care environments, and therefore proposed that a multimodal approach integrating VAI with other technologies would be ideal for addressing the complexities of caregiving support.

Exploring Ethical Values and Dimensions: Ethical Critiques of Envisioned VAI Concepts

Balancing Needs for Informal Care Support and Ethical Concerns

In the final workshop, participants provided ethical critiques of their VAI concepts, reflecting on and discussing potential ethical concerns about integrating solutions they designed into informal care. During workshops 2 and 3, we observed participants navigating various design tensions as they brainstormed ideas, revealing some contrasting perspectives and conflicting opinions regarding how VAI might support the care relationship (or not). For example, when suggesting ideas for VAI as emotional support, we found that participants questioned whether VAI should provide emotional support.

XSL•FO
RenderX

Further, while both older adult and caregiver participants remained positive about VAI's potential efficiency, in workshops 2‐3, we saw some different ideas of how VAIs might support their care. We expected some differences due to their experiences in their caregiving roles, although many of our caregivers were also older adults and provided perspectives from both roles. However, we found that when discussing ethical values, older adults and caregiver participants shared overlapping ethical concerns and expectations about how VAIs might handle health data and system transparency, as well as varied older adult needs, system trust and accountability, privacy and confidentiality, fairness and affordability, and shared autonomy. Therefore, participants highlighted a need to balance efficiency and ethics in designing and implementing such technologies. They discussed that if systems could address their ethical values, they would consider integrating VAI into their caregiving routines in the future.

To structure the discussion on ethical values, we report the ethical values and dimensions participants discussed to align with the 11 dimensions of AI ethics by Jobin et al [14], which include transparency, justice and fairness, nonmaleficence, responsibility, privacy, beneficence, freedom and autonomy, trust, sustainability, dignity, and solidarity. Similar to Jobin et al [14], we found that many of these dimensions overlap; however, of the eleven dimensions Jobin et al found in literature, our participants discussed seven dimensions: (1) transparency, (2) privacy, and (3) trust (including (4) nonmaleficence and (5) responsibility), (6) justice and fairness, and (7) freedom and autonomy.

Transparency, Privacy, and Trust

Overview

We found that participants considered the ethical dimensions of transparency, privacy, and trust to be essential yet complementary components in designing ethical VAI systems for informal care. While participants often discussed these components separately, the main takeaway was that the systems needed to provide avenues to build trust among users through transparent interactions, respect for privacy, and accurate information. Participants voiced a desire for clear, honest communication from VAI systems handling health data to ensure transparency about the systems and actions that help them better trust the system. At the same time, participants expected that systems would build trust by respecting their privacy, both by ensuring data confidentiality and by promoting privacy best practices. Relatedly, participants expected that VAI systems should proactively cultivate trust by ensuring the accuracy of information through personalization and contextually informed responses while limiting misinformation.

Ensure Sufficient Transparency of Health Data and System Actions

Discussions around transparency values related to data handling practices and knowledge of "how the system worked." Participants felt it was important to understand how the system handled their data, especially any health-related data or activity, including voice recordings of the surrounding environment or video capture, as they may contain sensitive data. Participants also wanted information about sources of responses and information from a VAI system. In a discussion about transparent communication of sources, a caregiver, C1,W4, stated that "I think there will be a lot of users that would not even care [about a VAI answer source], but there are people that would be like me, I really want to know where they got this information." However, while participants felt that knowing the source would be necessary, there was no consensus on how the system should provide it. The participant later noted that they would not want source information to be intrusive or prominently displayed, stating, "I think you want access to it, but don't put it in my face." Another participant, C2,W4, acknowledged that in critical or sensitive situations, the immediate need for assistance would likely take precedence over reviewing sources, stating, "I would absolutely wanna research so much of this information, but like in a situation that you're just thrown into, you just want the help, and you'll research it a little bit later."

Older adult participants shared that they often lacked understanding of the workings of AI systems and expressed that it would be important to have broader support for understanding the technology's capabilities and limitations. They emphasized the need to be informed about what they can and cannot do with it, noting that this has been an issue with past VAI technologies. O1, W4 referenced an incident in which Alexa altered its voice during the interaction, prompting them to question whether it was mimicking or mocking them. They shared, "this was just an experiment with my Echo, I just said, I love you. And she came back in a human-sounding voice. I thought that was kind of interesting in a weird way. Was it mocking?" Therefore, in addition to being transparent about answers and recommendations, participants also valued having a general understanding of "How an AI works." The participant mentioned, "What would make me trust it [VAI] is education, knowing more about it, reading more about it, learning how they do it."

Mitigate Privacy and Confidentiality Concerns Around Sensitive Data Handling

Caregivers and older adults voiced concerns about the substantial volume of sensitive data that VAI technology would need to effectively assist them with some caregiving tasks. They believed that accessing that volume of sensitive data carries the inherent risk of data breaches associated with storing and processing such confidential information. C2,W4 participant noted, "We're asking it to be kind of the keeper of all information. What if it were hacked?" Additionally, caregivers proposed implementing an authorization process for VAI technology, given its handling of health care information, noting that this process would ensure that only authorized individuals can issue commands and receive information. C1,W4 explained, "...if a new person is there [attempting to access personal health data], they should be authorized in some way to use this technology... the technology should only interact with the people it has registered as, like the usual voices." Here, the participant suggested that voice could be used as a unique identifier, similar to fingerprints or face recognition, to restrict or grant access to people using the system.

XSL•FO
RenderX

In one of the discussions, older adults contemplated whether asking sensitive questions or seeking personal information (eg, asking about "finding an economical nursing home") might pose additional risk compared to a traditional online search. They explained that they would be more reluctant to disclose certain types of sensitive health information to AI systems. O3,W3 noted, "I don't think I'd ever give my charge card numbers to AI. I guess one of the privacy things I thought about was just simply, do I want AI to know?" They indicated a reluctance to entrust AI with sensitive information, such as credit card numbers used to pay medical bills, even if it would be convenient, preferring not to "throw it out into the World Wide Web." They acknowledged that while some companies claim not to share information, they are often skeptical of such claims.

Cultivate Trust in the System and Be Accountable to Recommendations

Participants mainly discussed trust and accountability values through the lens of accuracy, highlighting their importance. Participants felt that, in a health care context, a VAI would need to be as accurate as possible. They felt this was true for any health care technology but discussed the unique benefits and challenges of accuracy in a VAI system.

Older adult participants discussed their worries about receiving incorrect information from a VAI system, especially for sensitive topics such as health. Caregivers and older adults, therefore, shared apprehensions about reliance on VAI technologies in caregiving situations. O2,W4 stated, "If I got a response that was totally crazy... I wouldn't use it anymore," explaining that if the system provided an answer that they felt was inaccurate, it would impact their use. O3,W4 discussed that in the news, they heard anecdotes about VAIs that might fabricate responses to appease the user. One older adult participant shared, "the little bits that I've really read about AI is that it tries real hard, and if it can't find what the legitimate answers are, it'll make stuff up to try and make sure that it satisfies."

Some participants also discussed the potential value of having a system that could personalize responses. They emphasized that they would value and trust personalized VAI systems that could provide contextually relevant responses tailored to individual needs and situations, rather than more abstract, generalized responses. A C1,W4 noted, "just having the technology to be able to individualize it to the environment, the person, and... their situation, I think. that'd be pretty accurate." In care-related contexts, participants felt that a degree of personalization was essential for providing more relevant responses to aid decision-making.

Older adult participants expressed concerns about the lack of fine-grained control they feel creators have over how AI might respond to queries, leading to inaccurate and unexpected responses. O3,W3 shared, "...I think AI, right now, I'm afraid they [big companies/AI developers] don't know how to control it." Similarly, they explained that they knew humans ultimately created AI systems, but felt humans could be limited in their ability to control how AI evolves because of how AI systems learn. O3,W3 said, "I know there's a programmer who writes the programs, but the way AI is evolving, it's really no longer under the control of a programmer; it's evolving on its own." As a result, participants mentioned that they would still consider AI-generated answers to some extent but would approach them with skepticism and exercise caution. Moreover, C3,W4 mentioned that they felt a trusted medical organization, such as the American Medical Association (ie, a trusted nonprofit organization in the United States consisting of medical professionals, students, and other stakeholders focused on issues related to patient care and policy), should validate and endorse any VAI caregiving system.

Justice and Fairness

Participants noted that they would expect ethically designed VAIs to consider the diversity of experiences in the design, including accessibility needs and inclusive design for older adults. While affordability was not identified as a component of current AI ethical dimensions in prior work [14], participants also noted that cost should be taken into account, as it could further widen gaps in access to health care resources.

Be Inclusive of Varying Older Adult Needs

Regarding accessibility, participants felt that while a VAI system could improve accessibility in some cases, it could also create additional accessibility challenges in others. A participant in the older adult session highlighted the potential benefits of voice commands in technology, noting that they may experience vision limitations as they age. O1,W4 explained, "...someday I won't be able to see very well, it enters my mind, and I would need help. ...It'll be helpful for you in a way if you can operate something with your voice, let's say you can just say, hey Alexa, connect me to this friend of mine, and they just do it." However, participants felt that if the VAI included the smaller screens found in commercial devices, access challenges could arise for people with visual disabilities. While they noted that the VAI could leverage voice as an alternative input method, larger screens could facilitate easier navigation and interaction and open possibilities for new functionalities, such as virtual doctor visits. O1,W3 explained, "So if I was developing a system and instead of Alexa, it was called nurse K or whatever that. Ok, that system and I think they're gonna move to screens on pretty much everything." O3,W3 agreed, noting access to a screen can be helpful, "Uh, I think Ring [the camera system] is one that comes with screens now. And so, it does the same thing Alexa does, it can search the internet or play songs or whatever, but it also has video and .I think you can get it on your phone or your laptop or iPad." Participants also emphasized the importance of adequate volume levels to address hearing loss among older adults. Therefore, participants raised concerns about balancing the introduction of a new interaction modality, VAI, with its inherent access challenges.

Participants also discussed values related to the diversity of speech when interacting with a VAI specifically designed to assist with care. Caregiver participants stressed the importance of representing diverse dialects to ensure the VAI system sounds natural and is easily understood, as overly mechanical voices can present challenges for users and make them less comfortable. C1,W3 explained, "because a lot of elderly people have hearing loss. So, I think it's important that we make sure the volume is good for them . and that the person's dialect is

XSL•FO
RenderX

represented. That's critical because if it sounds not too foreign but too mechanical, that could be a challenge too."

Make It Affordable, Promote Fair Access to Health Resources

Participants felt that affordability was an ethical concern for VAI systems that support caregiving. Participants felt that while the VAI system had the potential to improve efficiency, it would also incur higher costs. At the same time, cost is not always a factor formally considered in discussions of ethics. Participants noted that affordability could be a factor in determining who has access to these systems. Therefore, they related cost and affordability to justice and fairness in this context. They felt it was important that, if these VAI systems became available for caregiving purposes, they should be accessible to all who might elect to use them, so that all who wish to benefit might do so. Therefore, caregivers particularly emphasized a preference for integrating health-related VAI systems with insurance or existing health care systems, such as "MyChart," a patient-centered health record and messaging system, which already centralizes patients' health care information:

> *I go to Saint Francis, and they have something called MyChart … could it tie into something that's already existing like that …like a booster to that system. It would be good if it could tag on to what already exists out there that people are familiar with and be an enhancement. That would give you some legitimacy.* [C2,W4]

Moreover, older adults expressed apprehension about the future maintenance costs of such systems, especially because they felt their vigilance about maintaining subscriptions might decrease with age. A participant in the older adults' group, O2, W3, gave an example of how they were surprised to find that Alexa could facilitate games such as Jeopardy; however, they were surprised when it prompted them to start a free trial that later required a monthly payment of US $7.95. They shared, "I don't know much, and I asked my echo what she can do and, you know, it's funny. I said she could do games, and so she and I picked Jeopardy, and the music played, and I thought, well, this is really neat for people who can't get up and around, and it said you can start a free trial for a week. And then after that, it's $7.95 a month. Never occurred to me that there would be extra things on there that would cost money. I was kind of surprised by that. Another observation is dependency." Participants noted that in situations where they might rely on these types of VAI devices to support their care, they sometimes face a dilemma between continued care support (ie, from the system) and managing ongoing costs, which may limit access.

Freedom and Autonomy

Participants' discussions of freedom and autonomy mainly focused on the right to choose where and how VAIs are integrated into the care relationship, and on having sufficient autonomy to prioritize human relationships over functional support.

Participants voiced concerns about balancing technology-enabled caregiving assistance and the human touch. While acknowledging the potential benefits of technological support in caregiving, caregivers emphasized that it was essential to ensure that any technology, including VAI, does not overshadow or replace human interaction and compassion. They noted that the system must respect their values and rights to care autonomy (eg, rights to self-govern their care environment and experiences). O1,W2 shared, "I think the human touch and kindness is very important." C1,W4 shared, "It would be really important the individual is interacting with it [the VAI] instead of it taking over because it's real easy to do that."

Caregivers also emphasized that if technology aids caregiving activities, it should not be overly assertive, especially when empathy, compassion, and understanding of the situation and context might be necessary rather than a strict ultimatum. They gave the example of a VAI that might recommend an individual's water intake based on a predefined number of ounces, using established health guidelines. C1,W4 shared that it may not be productive for the VAI to recommend, "The person needs to drink a gallon of water a day when you're lucky to get eight ounces of water in them. I guess you wouldn't want it to be too aggressive." They felt that this was a situation where a human would understand the situation more holistically, and that an assertive AI might end up being discouraging rather than motivating.

Older adults expressed appreciation for the assistance and accountability that VAI technology could provide, including reminders to engage in beneficial activities such as stretching. However, they discussed that they felt it was important to maintain control over the technology's level of involvement in their care and have the option to override its actions if they feel it becomes too intrusive or overpowering. O3,W4 shared, "If it goes too far and I'm unhappy with it, I just want that kill switch to say no, you're done."

Discussion

Summary of Findings

In this study, we used a 3-phase research approach to explore the potential of VAI for informal caregiving and identify ethical dimensions that might impact adoption. We aimed to answer the following research questions: (RQ1) How do participants perceive the role of VAIs in supporting caregiving tasks? (RQ2) What values do caregivers and older adults have around VAI assistants that support day-to-day care tasks? (RQ3) What ethical concerns do caregivers and older adults have about VAI assistants that support day-to-day care tasks?

The findings illuminate a core tension between the perceived benefits and efficiency of VAI for supporting informal care and the ethical considerations surrounding its implementation. Participants identified several ways VAI could enhance caregiving efficiency, streamline caregiving tasks, and provide emotional, social, and training or educational support. Nevertheless, caregivers and older adults express apprehensions regarding the ethical implications of AI-driven caregiving solutions. Based on these findings, we also provide design considerations regarding the role of VAIs in assisting informal aging care and in addressing ethical concerns from a harm-reduction perspective.

XSL•FO
RenderX

Role of Voice-Activated Assistance in Informal Care: How Can Voice-Activated Systems Help?

Related to RQ1 and RQ2, we found that the primary appeal of VAI systems was the potential capacity to streamline caregiving routines, allowing caregivers and older adults to manage care tasks quickly and conveniently. Participants felt that this efficiency is particularly valuable when timely support and task management are critical. This finding aligns with other studies in the health care context, in which researchers have found that older adults desire VAI features that help with caregiving tasks, such as scheduling and appointment management, or medication reminders, for perceived convenience and efficiency [4,8,10,13]. We also found that VAI was viewed not only as a practical tool but also as a potential source of companionship to mitigate the feelings of loneliness and social isolation prevalent among caregivers and care recipients. Although not explicitly situated in the context of informal caregiving, prior work has explored the use of VAI as companions for older adults [4,5,50-53] and its potential to support social connections [54,55]. Rudnik et al [9] explored how a VAI system might benefit care conversations between older adults and caregivers. Yet our work extends prior work by finding that designing VAI for companionship and supporting social dimensions within a caregiving relationship will likely require careful additional exploration of existing relational values that form the foundation of the caregiving relationship. For example, participants envisioned VAI as a way to support socioemotional aspects of informal care, but also raised concerns about privacy, emotional attachment, and the erosion of human connection that VAI might bring. Therefore, our work suggests that, when designing, there needs to be a broader ethical discourse on integrating VAI and on the broader user values surrounding social relationships and human connectedness in caregiving contexts. Future work is likely needed to better understand those values and how they might shape the design of VAI technologies for informal care.

While some work has explored how VAI might be integrated to reduce caregiver burden [9,13], we also found that caregivers felt that a VAI system could help address the emotional toll of providing informal care. They saw opportunities for VAI to be a proactive companion in helping them remember to practice self-care and find related opportunities. Examples included proactively conducting self-care checks or guiding them through exercises when they felt overwhelmed. Additionally, we found that part of the emotional toll of caregiving lies in learning to provide care, as one must learn on the "job." VAI was also seen as potentially useful for helping caregivers adjust by providing an efficient way to obtain the training, resources, and information they need to function as informal caregivers. For example, VAIs might be integrated within a do-it-yourself smart home setup to provide caregivers with a hands-free way to interact with educational materials and resources. Researchers have found that older adults see potential in conversational agents to support access to personal health records and to help find and understand health data [27]. Voice-enabled systems might integrate with patient portals to create more accessible interactions with the health data included there. Future work should explore more ways VAI might assist with education, training, and support in informal care.

Our research findings also suggest that participants generally felt that relying solely on the voice modality is insufficient for fully addressing the complexities of a caregiving environment. While they felt that VAI provides a valuable interface for interaction, they felt its effectiveness would likely be enhanced if integrated into a broader care ecosystem that included multimodal features, such as visual displays, tactile feedback, and contextual cues. They discussed how often different caregiving tasks require different approaches and that different environments might have different needs. Therefore, in the future, it may be beneficial for designers to consider ways VAI might act as a facilitator of care tasks within informal care environments, for example, by exploring interoperable or multiplatform facilitation and connecting disparate devices into a connected ecosystem. For example, participants suggested that voice input and output be a way for them to seamlessly connect with different devices that support their self-care, health, and well-being, including wearable devices such as smartwatches and smart rings, instructional videos for caregivers, and smart televisions, to allow older adults and caregivers a more comprehensive picture of their health, while also supporting older adults' independence and reducing caregiver tasks.

Assist Me, but Do No Harm: What Ethical Values Should VAIs for Informal Care Consider?

Users' ethical concerns about AI technologies are not new and have been raised in other research [14]. For example, within health care, researchers have explored ethical concerns around using conversational agents for health-related areas such as mental health [56] and more broadly care providers [57] and public health [58], noting potential risks such as algorithmic bias, user harms, and privacy concerns due to users' unawareness that they are interacting with a virtual agent. Others have noted safety concerns, particularly in voice-only applications, due to a lack of transparency [36]. Privacy remains a core ethical concern among those using VAI systems, along with others related to personification and emotional connections to machines [19]. In parallel, general concerns about user personification of VAI technologies are well-known, with the HCI literature suggesting that the "human-like" aspect of voice user interfaces can sometimes be an enhancement or a detriment to ethical design, depending on context [59]. Therefore, while VAIs can have promising applications in health care, it is often not well known how ethical concerns vary among users and across contexts [19]. This has been found especially true within health-related contexts, where users tend to hold AI technologies to different standards depending on the context, given the sensitive data passed and the potential harms (eg, safety-related) that might arise from its use [33].

Related to RQ3, our findings highlight that VAIs are beneficial primarily because of perceptions of improved efficiency and social or emotional support for informal caregiving tasks. However, our research reveals a critical tension between the efficiency gains offered by VAI in caregiving and the ethical considerations that accompany its use. Therefore, the efficiency provided by VAI must be balanced against these ethical considerations, ensuring that the technology remains a tool for empowerment rather than a source of control or intrusion.

XSL•FO
RenderX

Within health care, the idea of interventional harm is typically associated with safety or injury [60]. However, in recent years, researchers have noted the significant potential harm of social-behavioral interventions that may not have immediate safety concerns, but due to the disruption they cause to complex social systems and needs for human agency, can lead to unintended effects [60]. Therefore, more emphasis has been placed on unearthing non–safety-related and unintentional harms to explore the potential impact of those concerns on intervention success (eg, acceptance and adoption) [60,61]. For example, Allen-Scott et al [61] found that for public health interventions, five high-level unintended harms are often associated with unsuccessful public health intervention deployment, including (1) physical, (2) psychosocial, (3) economic, (4) cultural, and (5) environmental which often occur for reasons including lack of community engagement when developing the intervention and ignoring root causes which lead to a lack of understanding of what users want. Examining harm is also a topic that has been explored among VAI technologies more generally [62,63]. For example, Reeves [62] examined the trade-off between framing VAI interactions as conversations and the potential for user harm. Aligned with these efforts, we summarize the ethical values highlighted by older adults and caregivers into broad categories of potential unintended harms that may ultimately lead to nonacceptance and nonuse of VAIs for informal care.

Physical Harms

Aligned with prior work on VAI systems, our findings suggest that assessing the potential safety-related harms of health-related VAI technologies remains paramount to users. Safety concerns associated with VAI use in health have been raised in studies where physicians note the potential harms to patients who rely on voice assistants for health information search [36]. However, beyond safety implications associated with receiving inaccurate health information or recommendations, our participants noted broader transparency issues that should be considered, such as helping users understand how decisions are made and/or the data privacy consequences that could affect their ability to receive care and resources (eg, breaches that impact insurance availability).

Financial Harms

Two categories related to finances emerged from our data: (1) the potential harm associated with data privacy breaches of sensitive information, and (2) increasing inequities in access to health resources due to limited financial resources to adopt health VAIs. Therefore, the implications of a financial data breach involving VAIs that could affect a user's financial wellness should be considered. Additionally, while not often listed as an ethical consideration for AI generally [14], others have found that cost is an aspect users consider in VAI adoption [11]. Therefore, cost and other financial factors should be considered in efforts to reduce inequities in access to caregiving resources. Mitigating the trade-offs between the efficiency gains and the ongoing costs of adoption would be essential to ensure that the benefits of VAI reach a wider population of users.

Relational or Social Harms

Our findings align with the informal caregiving literature, which notes that the social, relational, and emotional aspects of informal caregiving are just as important as physical caregiving tasks. However, we found that human autonomy concerns [64], such as shared governance of caregiving tasks, emerged when discussing the social or relational elements of VAI design for informal care, including those related to physical tasks that required social interaction. These findings suggest that users may value VAI in ways that can support social and relational engagement if it is designed in a way that does not overstep the boundaries of what makes certain aspects of an informal caregiving relationship a uniquely human endeavor (eg, human touch, social connectedness, human relationship building, and inclusiveness). Other researchers have found that the social aspects of VAI technologies, such as how they engage users, can impact perceptions, for example, among marginalized groups [65]. Therefore, while providing human-like characteristics to VAI systems is generally thought to have potential advantages for supporting health-related applications [66], our findings suggest that it may be necessary to further explore the impact of adding human-like characteristics on users' perceptions of autonomy when governing their caregiving relationships.

Design Considerations for Ethical VAI for Aging and Informal Care

Based on participants' suggestions, to further address RQ3, we also provide design implications for addressing their desired interactions and ethical values related to (1) transparency, privacy, and trust (including nonmaleficence and responsibility), (2) justice and fairness, and (3) freedom and autonomy when engaging with VAI, which supports informal caregiving practices.

We suggest that designers and developers encourage trust by incorporating features that help users assess the credibility of the information they receive (D1). Prior work has found that limited access to source information makes it difficult to assess the credibility of VAI systems for health [28,67], and the ability to view source information can reduce those concerns [27]. As VAIs - often inherently have less transparency due to the lack of visual guidance, VAIs designed for informal care should consider ways of integrating transparency within conversational design, such as allowing users to ask questions verbally about how their data is used or how the system works, or through multimodal approaches that allow users to better understand who has access to their data or fact-check information sources.

Older adults' privacy concerns related to VAI technologies have been well-documented in the literature [68,69]. Our findings suggest that older adults and caregivers have similar concerns about VAIs that are integrated into their care processes, including unauthorized access to personal health data. We suggest that designers and developers consider the dynamic privacy needs of potential users of these systems, including older adults, informal caregivers, and others involved in the informal care process (D2). Due to the sensitive nature of health data, VAI systems should implement industry-standard security protocols and help users build greater trust in the system through

XSL•FO
RenderX

transparent, clear communication about privacy policies. Systems should also emphasize user autonomy in data sharing, allowing users to control how their data is shared with others and third parties. Systems can also leverage VAI biometrics to prevent unauthorized access to sensitive information.

Prior literature also notes that older adults desire personalized experiences with intelligent voice technologies to support efficient task completion [11,21,22]. For example, Brewer et al [22] found in their study of older adults' use of current voice assistants for health information seeking that older adults asked for more personalized health information, feedback, and recommendations. In clinical settings, extending voice capabilities to personalize experiences has been found beneficial [70]. Our findings suggest that designers and developers should consider leveraging personalization to address diverse caregiving needs (D3). For example, VAIs might leverage contextual cues from users and environmental information to provide more personalized responses (eg, gauging mood to support caregiver wellness or older adult well-being).

Finally, balancing assistance and control is a longstanding open challenge in systems that automate tasks [71], leading to approaches such as mixed-initiative interaction [72] and, more recently, a focus on understanding human autonomy in AI systems [64]. Studying autonomy in AI-enabled systems has also been longstanding in technology designed for older adults [73-77], given frequent discussions of independence and agency as well as balancing AI and the human aspects of care [78]. Findings suggest that the successful integration of VAIs into informal caregiving relationships requires researchers to consider broader ethical values around the importance of the human aspects of caregiving relationships (D4). Further, participants noted that VAI alone is not enough to support the complex environment of informal care. Instead of viewing VAI as a standalone solution, designers should explore ways to leverage its capabilities to support more holistic care within the caregiving environment (D5). This can include exploring ways to leverage VAI to provide accessible information and resources to older adults and caregivers, support independence, or complement the broader care ecosystem through a set of connected devices. We summarize design implications for VAIs for informal aging care in Table 3.

XSL•FO
RenderX

Table . Design implications for informal VAI[a] care systems.

Design implication	Ethical dimensions addressed	Suggestions
D1: encourage trust through transparency	Transparency, trust	• Incorporate subtle opportunities for accessing information sources • Help users gauge when AI[b] is unsure by adding fact-checking features or the ability to flag information of uncertain origin
D2: prevent unauthorized access	Trust, privacy	• Implement industry-standard security protocols to protect user data • Be transparent about privacy policies at the start of the interaction • Let users control the kind of information they share with companies and caregivers • Leverage VAI capabilities (eg, voice biometrics) to automate system authorization processes • Suggest alternatives to voice commands, such as text input
D3: provide personalization through context	Justice and fairness	• Leverage contextual awareness such as environmental and user cues to deliver more relevant and dynamic responses (eg, gauging mood) and ensure more secure interactions (eg, verifying access)
D4: support autonomy through awareness	Freedom and autonomy	• Incorporate clear and user-friendly features that let users easily override or disable AI suggestions whenever they want • Enable personalization or automatically recognize and respond proactively to cues that indicate a need for human interaction
D5: provide flexibility through integrated support	Justice and fairness	• Leverage voice input and output to provide accessible information and resources • Leverage voice input to enhance older adults' independence in managing self-care and well-being • Leverage voice input and output for hands-free interaction with other devices that support care (eg, health education and demonstration materials, and wearables)

[a]VAI: voice-enabled artificial intelligence.

[b]AI: artificial intelligence.

Limitations

This research includes some limitations. While we conducted some in-person studies, much of the work took place online to make it more convenient for our participants. Moving the studies online might have prevented some participants from engaging in the work. For example, those who have lower digital literacy or limited access to technology. This study also had more caregiver participants than older adult participants. Although we did see recurring themes within and between groups, our sample was also smaller overall. We also acknowledge that, because of the age similarities between our older adult participants and caregiving participants (one caregiving participant was aged younger than 55 years) and the flexibility of caregivers to also comment on their experiences as older adults, older caregiver perspectives may be overrepresented. Additionally, while we attempt to provide a rich description [46] of the participants and their experiences through demographic data and personas, because individuals' care experiences vary, the participants' experiences in our study may not transfer to all populations and contexts. There may be additional ethical values that did not emerge from our data, or that may not apply in different contexts.

Conclusions

This study explores the experiences and challenges of older adults and informal family caregivers in managing informal care. Through participatory speculative design, we explore how future VAI technologies might support the caregiving process and address caregivers' needs and concerns. We intentionally explore the ethical dimensions that participants value, and that may affect the potential adoption of VAI technology in informal care. Our findings provide insight into the need for VAI systems

XSL•FO
RenderX

designed for the informal care context to balance efficiency and ethical values, supporting care processes while doing no harm.

Acknowledgments

We would like to thank the participants who took the time to share their experiences and opinions with us. No generative artificial intelligence tools were used at any stage in the preparation of this paper. The authors take full responsibility for the accuracy, originality, and integrity of all content, including all references and citations.

Funding

This research is based on work supported by the National Science Foundation (2130583). The National Science Foundation had no involvement in this study's design, data collection, analysis, interpretation, or the writing of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of additional tables and figures.
[DOCX File, 7129 KB - aging_v9i1e79740_app1.docx]

References

1. Gallant MP, Spitze GD, Prohaska TR. Help or hindrance? How family and friends influence chronic illness self-management among older adults. Res Aging 2007 Sep;29(5):375-409. [doi: 10.1177/0164027507303169]
2. Donelan K, Hill CA, Hoffman C, et al. Challenged to care: informal caregivers in a changing health system. Health Aff (Millwood) 2002;21(4):222-231. [doi: 10.1377/hlthaff.21.4.222] [Medline: 12117133]
3. Schulz R, Beach SR, Czaja SJ, Martire LM, Monin JK. Family caregiving for older adults. Annu Rev Psychol 2020 Jan 4;71(1):635-659. [doi: 10.1146/annurev-psych-010419-050754] [Medline: 31905111]
4. O'Brien K, Liggett A, Ramirez-Zohfeld V, Sunkara P, Lindquist LA. Voice-controlled intelligent personal assistants to support aging in place. J Am Geriatr Soc 2020 Jan;68(1):176-179. [doi: 10.1111/jgs.16217] [Medline: 31617581]
5. Pradhan A, Mehta K, Findlater L. "Accessibility came by accident": use of voice-controlled intelligent personal assistants by people with disabilities. CHI '18 2018:1-3. [doi: 10.1145/3173574.3174033]
6. Orlofsky S, Wozniak K. Older adults' experiences using Alexa. Geriatr Nurs 2022;48:247-257. [doi: 10.1016/j.gerinurse.2022.09.017] [Medline: 36332440]
7. Shinde PU, Martin-Hammond A. Designing to support blind and visually impaired older adults in managing the invisible labor of social participation: opportunities and challenges. CHI '24 2024 May 11:1-14. [doi: 10.1145/3613904.3642203]
8. Zubatiy T, Vickers KL, Mathur N, Mynatt ED. Empowering dyads of older adults with mild cognitive impairment and their care partners using conversational agents. CHI '21 2021 May 6:1-15. [doi: 10.1145/3411764.3445124]
9. Rudnik J, Raghuraj S, Li M, Brewer RN. CareJournal: a voice-based conversational agent for supporting care communications. CHI '24 2024 May 11:1-22. [doi: 10.1145/3613904.3642163]
10. Corbett CF, Combs EM, Wright PJ, et al. Virtual home assistant use and perceptions of usefulness by older adults and support person dyads. Int J Environ Res Public Health 2021 Jan 27;18(3):1113. [doi: 10.3390/ijerph18031113] [Medline: 33513798]
11. Nallam P, Bhandari S, Sanders J, Martin-Hammond A. A question of access: exploring the perceived benefits and barriers of intelligent voice assistants for improving access to consumer health resources among low-income older adults. Gerontol Geriatr Med 2020;6:2333721420985975. [doi: 10.1177/2333721420985975] [Medline: 33457459]
12. Brewer RN, Harrington C, Heldreth C. Envisioning equitable speech technologies for black older adults. FAccT '23 2023 Jun 12:379-388. [doi: 10.1145/3593013.3594005]
13. Tennant R, Allana S, Mercer K, Burns CM. Caregiver expectations of interfacing with voice assistants to support complex home care: mixed methods study. JMIR Hum Factors 2022 Jun 30;9(2):e37688. [doi: 10.2196/37688] [Medline: 35771594]
14. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell 2019;1(9):389-399. [doi: 10.1038/s42256-019-0088-2]
15. Seymour W, Cote M, Such J. Can you meaningfully consent in eight seconds? Identifying ethical issues with verbal consent for voice assistants. CUI 2022 2022 Jul 26:1-4. [doi: 10.1145/3543829.3544521]
16. McDougall RJ. Computer knows best? The need for value-flexibility in medical AI. J Med Ethics 2019 Mar;45(3):156-160. [doi: 10.1136/medethics-2018-105118] [Medline: 30467198]
17. Alam MAU. AI-fairness towards activity recognition of older adults. MobiQuitous '20 2020 Dec 7:108-117. [doi: 10.1145/3448891.3448943]

XSL•FO
RenderX

18. Lee MK, Rich K. Who is included in human perceptions of AI?: Trust and perceived fairness around healthcare AI and cultural mistrust. CHI '21 2021 May 6:1-14. [doi: 10.1145/3411764.3445570]
19. Seymour W, Zhan X, Coté M, Such J. A systematic review of ethical concerns with voice assistants. AIES '23 2023 Aug 8:131-145. [doi: 10.1145/3600211.3604679]
20. Wolff JL, Spillman BC, Freedman VA, Kasper JD. A national profile of family and unpaid caregivers who assist older adults with health care activities. JAMA Intern Med 2016 Mar;176(3):372-379. [doi: 10.1001/jamainternmed.2015.7664] [Medline: 26882031]
21. Martin-Hammond A, Vemireddy S, Rao K. Exploring older adults' beliefs about the use of intelligent assistants for consumer health information management: a participatory design study. JMIR Aging 2019 Dec 11;2(2):e15381. [doi: 10.2196/15381] [Medline: 31825322]
22. Brewer R, Pierce C, Upadhyay P, Park L. An empirical study of older adult's voice assistant use for health information seeking. ACM Trans Interact Intell Syst 2022 Jun 30;12(2):1-32. [doi: 10.1145/3484507]
23. Brewer RN. "If Alexa knew the state I was in, it would cry": older adults' perspectives of voice assistants for health. CHI '22 2022 Apr 27:1-8. [doi: 10.1145/3491101.3519642]
24. Jung I, Kim H, Lim YK. Understanding how users experience the physiological expression of non-humanoid voice-based conversational agent in healthcare services. DIS '21 2021 Jun 28:1433-1446. [doi: 10.1145/3461778.3462082]
25. Chen C, Johnson JG, Charles K, et al. Understanding barriers and design opportunities to improve healthcare and QOL for older adults through voice assistants. ASSETS '21 2021 Oct 17:1-16. [doi: 10.1145/3441852.3471218]
26. Karimi P, Bora P, Martin-Hammond A. Scribe: improving older adults' access to medical instructions from patient-physician conversations. W4A '21 2021 May 20:1-11. [doi: 10.1145/3430263.3452420]
27. Karimi P, Ballard K, Vazirani P. Designing conversational assistants to support older adults' personal health record access. In: Lewy H, Barkan R, editors. PH 2021 2022;431:253-271. [doi: 10.1007/978-3-030-99194-4_17]
28. Brewer R. Understanding voice‐based information uncertainty: a case study of health information seeking with voice assistants. J Assoc Inf Sci Tech 2023 Dec 5;75(10):1041-1057. [doi: 10.1002/asi.24854]
29. Kocielnik R, Langevin R, George JS, et al. Can I talk to you about your social needs? Understanding preference for conversational user interface in health. CUI '21 2021 Jul 27:1-10. [doi: 10.1145/3469595.3469599]
30. Seymour W, Binns R, Slovak P, Kleek M, Shadbolt N. Strangers in the room: unpacking perceptions of "smartness" and related ethical concerns in the home. DIS '20 2020 Jul 3:841-854. [doi: 10.1145/3357236.3395501]
31. Seymour W, Van Kleek M. Does Siri have a soul? Exploring voice assistants through shinto design fictions. CHI '20 2020 Apr 25:1-12. [doi: 10.1145/3334480.3381809]
32. Gupta A, Basu D, Ghantasala R, Qiu S, Gadiraju U. To trust or not to trust: how a conversational interface affects trust in a decision support system. WWW '22 2022 Apr 25:3531-3540. [doi: 10.1145/3485447.3512248]
33. Zhan X, Abdi N, Seymour W, Such J. Healthcare voice AI assistants: factors influencing trust and intention to use. Proc ACM Hum-Comput Interact 2024 Apr 26;8:1-37. [doi: 10.1145/3637339]
34. Cho E. Hey Google, can I ask you something in private? CHI '19 2019 May 2:1-9. [doi: 10.1145/3290605.3300488]
35. Martin-Hammond A, Vemireddy S, Rao K. Engaging older adults in the participatory design of intelligent health search tools. PervasiveHealth '18 2018 May 21:280-284. [doi: 10.1145/3240925.3240972]
36. Bickmore TW, Trinh H, Olafsson S, et al. Patient and consumer safety risks when using conversational assistants for medical information: an observational study of Siri, Alexa, and Google Assistant. J Med Internet Res 2018 Sep 4;20(9):e11510. [doi: 10.2196/11510] [Medline: 30181110]
37. Liao V, Muller M. Enabling value sensitive AI systems through participatory design fictions. arXiv. Preprint posted online on Dec 13, 2019. [doi: 10.48550/arXiv.1912.07381]
38. Muller M, Liao QV. Exploring AI ethics and values through participatory design fictions. : IBM Research, Cambridge and Yorktown Heights URL: https://www.qveraliao.com/hcic2017.pdf [accessed 2026-05-20]
39. Vines J, Clarke R, Wright P, McCarthy J, Olivier P. Configuring participation: on how we involve people in design. CHI '13 2013 Apr 27:429-438. [doi: 10.1145/2470654.2470716]
40. Friedman B, Kahn PH, Borning A. "Value sensitive design: theory and methods technical report 2.8. : University of Washington; 2002 Dec URL: https://research.cs.vt.edu/ns/cs5724papers/6.theoriesofuse.cwaandvsd.friedman.vsd.pdf [accessed 2026-06-09]
41. Šabanović S, Chang WL, Bennett CC. A robot of my own: participatory design of socially assistive robots for independently living older adults diagnosed with depression. In: Zhou J, Salvendy G, editors. ITAP 2015 2015 Jan 1:104-114. [doi: 10.1007/978-3-319-20892-3_11]
42. Ringfort-Felner R, Laschke M, Sadeghian S, Hassenzahl M. Kiro: a design fiction to explore social conversation with voice assistants. Proc ACM Hum-Comput Interact 2022;6(GROUP):1-21. [doi: 10.1145/3492852]
43. Søndergaard MLJ, Hansen LK. Intimate futures: staying with the trouble of digital personal assistants through design fiction. DIS '18 2018 Jun 8:869-880. [doi: 10.1145/3196709.3196766]
44. Wilcox L, Brewer R, Diaz F. AI consent futures: a case study on voice data collection with clinicians. Proc ACM Hum-Comput Interact 2023 Oct 4;7(CSCW2):1-30. [doi: 10.1145/3610107]

XSL•FO
RenderX

45. Auger JH. Living with robots: a speculative design approach. J Hum-Robot Interact 2014;3(1):20-42. [doi: 10.5898/JHRI.3.1.Auger]
46. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches: SAGE Publications; 2016.
47. Clarke V, Braun V, Hayfield N. Thematic analysis. In: Qualitative Psychology: A Practical Guide to Research Methods 2015:222-248.
48. McDonald N, Schoenebeck S, Forte A. Reliability and inter-rater reliability in qualitative research: norms and guidelines for CSCW and HCI practice. Proc ACM Hum-Comput Interact 2019;3(CSCW):1-23. [doi: 10.1145/3359174]
49. Waycott J, Vetere F, Pedell S, Morgans A, Ozanne E, Kulik L. Not for me: older adults choosing not to participate in a social isolation intervention. CHI '16 2016 May 7:745-757. [doi: 10.1145/2858036.2858458]
50. Corbett CF, Wright PJ, Jones K, Parmer M. Voice-activated virtual home assistant use and social isolation and loneliness among older adults: mini review. Front Public Health 2021;9:742012. [doi: 10.3389/fpubh.2021.742012] [Medline: 34708017]
51. Upadhyay P, Heung S, Azenkot S, Brewer RN. Studying exploration & long-term use of voice assistants by older adults. CHI '23 2023 Apr 19:1-11. [doi: 10.1145/3544548.3580925]
52. Jones VK, Hanus M, Yan C, Shade MY, Boron JB, Bicudo RM. Reducing loneliness among aging adults: the roles of personal voice assistants and anthropomorphic interactions. Front Public Health 2021;9:750736. [doi: 10.3389/fpubh.2021.750736] [Medline: 34957013]
53. Liu M, Wang C, Hu J. Older adults' intention to use voice assistants: usability and emotional needs. Heliyon 2023 Nov;9(11):e21932. [doi: 10.1016/j.heliyon.2023.e21932] [Medline: 38027966]
54. Joshi KR, Ulabhaje AA, Nataraj R, Martin-Hammond A. Skeptical yet curious: attitudes towards conversational assistants for supporting social engagement among older adults living alone. CHI EA '25 2025 Apr 26:1-7. [doi: 10.1145/3706599.3719917]
55. Karkera Y, Tandukar B, Chandra S, Martin-Hammond A. Building community capacity: exploring voice assistants to support older adults in an independent living community. CHI '23 2023 Apr 19:1-17. [doi: 10.1145/3544548.3581561]
56. Luxton DD, Hudlicka E. Intelligent virtual agents in behavioral and mental healthcare: ethics and application considerations. In: Advances in Neuroethics: Springer; 2021:41-55. [doi: 10.1007/978-3-030-74188-4_4]
57. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. Artif Intell Med 2014 Sep;62(1):1-10. [doi: 10.1016/j.artmed.2014.06.004] [Medline: 25059820]
58. Luxton DD. Ethical implications of conversational agents in global public health. Bull World Health Organ 2020 Apr 1;98(4):285-287. [doi: 10.2471/BLT.19.237636] [Medline: 32284654]
59. Luger E, Sellen A. "Like having a really bad PA": the gulf between user expectation and experience of conversational agents. CHI '16 2016 May 7:5286-5297. [doi: 10.1145/2858036.2858288]
60. Bonell C, Jamal F, Melendez-Torres GJ, Cummins S. "Dark logic": theorising the harmful consequences of public health interventions. J Epidemiol Community Health 2015 Jan;69(1):95-98. [doi: 10.1136/jech-2014-204671] [Medline: 25403381]
61. Allen-Scott LK, Hatfield JM, McIntyre L. A scoping review of unintended harm associated with public health interventions: towards a typology and an understanding of underlying factors. Int J Public Health 2014 Feb;59(1):3-14. [doi: 10.1007/s00038-013-0526-6] [Medline: 24381979]
62. Reeves S. Conversation considered harmful? CUI 2019 2019 Aug 22:1-3. [doi: 10.1145/3342775.3342796]
63. Fischer JE. Generative AI considered harmful. CUI '23 2023 Jul 19:1-5. [doi: 10.1145/3571884.3603756]
64. Calvo RA, Peters D, Vold K, Ryan RM. Supporting human autonomy in AI systems: a framework for ethical enquiry. In: Burr C, Floridi L, editors. Ethics of Digital Well-Being: A Multidisciplinary Approach Philosophical Studies Series: Springer; 2020:31-54. [doi: 10.1007/978-3-030-50585-1_2]
65. Harrington CN, Egede L. Trust, comfort and relatability: understanding black older adults' perceptions of chatbot design for health information seeking. CHI '23 2023 Apr 19:1-18. [doi: 10.1145/3544548.3580719]
66. Clark L, Pantidi N, Cooney O, et al. What makes a good conversation?: Challenges in designing truly conversational agents. CHI '19 2019 May 2:1-12. [doi: 10.1145/3290605.3300705]
67. Pradhan A, Lazar A, Findlater L. Use of intelligent voice assistants by older adults with low technology use. ACM Trans Comput-Hum Interact 2020 Aug 31;27(4):1-27. [doi: 10.1145/3373759]
68. Bonilla K, Gaitan B, Sanders J, Khenglawt N, Martin-Hammond A. Comparing older and younger adults perceptions of voice and text-based search for consumer health information tasks. AMIA Annu Symp Proc 2021;2021:227-236. [Medline: 35309005]
69. Bonilla K, Martin-Hammond A. Older adults' perceptions of intelligent voice assistant privacy, transparency, and online privacy guidelines. : Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020); 2020 URL: https://scholarworks.indianapolis.iu.edu/server/api/core/bitstreams/60e837c4-8292-4987-9559-3e7bf3dc892e/content [accessed 2026-05-20]
70. Pantazopoulos G, Bruyere J, Nikandrou M, et al. ViCA: combining visual, social, and task-oriented conversational AI in a healthcare setting. ICMI '21 2021 Oct 18:71-79. [doi: 10.1145/3462244.3479909]
71. Shneiderman B, Maes P. Direct manipulation vs. interface agents. interactions 1997 Nov;4(6):42-61. [doi: 10.1145/267505.267514]

XSL•FO
RenderX

72. Horvitz E. Principles of mixed-initiative user interfaces. CHI ’99 1999 May 1:159-166. [doi: 10.1145/302979.303030]
73. Karimi P, Martin-Hammond A. Finding the right balance: user control and automation in AI tools for supporting older adults’ health information tasks. CHI EA ’25 2025 Apr 26:1-10. [doi: 10.1145/3706599.3719773]
74. Sanders J, Martin-Hammond A. Exploring autonomy in the design of an intelligent health assistant for older adults. IUI ’19 2019 Mar 16:95-96. [doi: 10.1145/3308557.3308713]
75. Coghlan S, Waycott J, Lazar A, Neves BB. Dignity, autonomy, and style of company: dimensions older adults consider for robot companions. Proc ACM Hum Comput Interact 2021 Apr;5(CSCW1):1-25. [doi: 10.1145/3449178] [Medline: 34308262]
76. Caldeira C, Nurain N, Connelly K. “I hope I never need one”: unpacking stigma in aging in place technology. CHI ’22 2022 Apr 29:1-12. [doi: 10.1145/3491102.3517586]
77. Berridge C, Zhou Y, Lazar A, et al. Control matters in elder care technology: evidence and direction for designing it in. DIS ’22 2022 Jun 13:1831-1848. [doi: 10.1145/3532106.3533471]
78. Karimi P, Martin-Hammond A. Designing intelligent voice assistants for older adults’ collaborative care: exploring supportive and non-supportive interactions. Proc ACM Hum-Comput Interact 2025 Oct 18;9(7):1-35. [doi: 10.1145/3757465]

Abbreviations

AI: artificial intelligence
HCI: human-computer interaction
RQ: research question
VAI: voice-enabled artificial intelligence
VUI: VUI

Edited by M Mardini; submitted 17.Jul.2025; peer-reviewed by F Yang, P Voukelatou, R Boumans; revised version received 09.Mar.2026; accepted 03.May.2026; published 16.Jun.2026.

Please cite as:
Singh S, Martin-Hammond A
Exploring Expectations, Ethical Dimensions, and Values for Voice-Enabled AI Assistants That Support Older Adults and Informal Family Caregivers: Participatory Speculative Design Study
JMIR Aging 2026;9:e79740
URL: https://aging.jmir.org/2026/1/e79740
doi:10.2196/79740

© Shweta Singh, Aqueasha Martin-Hammond. Originally published in JMIR Aging (https://aging.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Potential Application Value and Needs of Using Generative AI Videos for Health Management for Older Adults: Qualitative Study

Ting Liu[1], MA; Yiming Taclis Luo[1], MSc; Zhenni He[1], MA; Patrick Pang[1], PhD; Kin Sun Chan[2], PhD; Ying Lau[3], PhD

[1]Faculty of Applied Sciences, Macao Polytechnic University, Rua de Luís Gonzaga Gomes, Macao, Macao

[2]Department of Government and Public Administration, Faculty of Social Sciences, University of Macau, Taipa, Macao

[3]The Nethersole School of Nursing, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong Special Administrative Region of China, China

Corresponding Author:
Patrick Pang, PhD
Faculty of Applied Sciences, Macao Polytechnic University, Rua de Luís Gonzaga Gomes, Macao, Macao

Abstract

Background: Population aging has emerged as a global concern, older adults' ability to access health knowledge and manage their well-being impacts their health outcomes. In the artificial intelligence era, generative artificial intelligence (GenAI) videos hold promise for enhancing geriatric health management. However, their potential and the needs of older adults in using GenAI videos for health-related purposes deserve a more in-depth investigation.

Objective: This study aims to explore the application potential and multifaceted needs of older adults in using GenAI videos for health information acquisition and management, while providing actionable recommendations for future aging-friendly GenAI video tools.

Methods: A qualitative approach was adopted. Twenty older adults (aged ≥60 y) with basic digital literacy were recruited from communities for participation in iterative GenAI video workshops. Semistructured interviews were conducted. Thematic analysis, following Braun and Clarke's 6-phase reflexive framework, was used to identify key themes from interview transcripts.

Results: Our results have identified a 3-layer hierarchical structure of needs when older adults interact with GenAI videos. The first layer is technical accessibility, which suggests the direct barriers preventing them from operating GenAI tools independently. The second layer is age-friendly design, which depicts the needs of age-friendly design adaptations, which can help older adults to better use GenAI features. The final layer is integration with professional resources, highlighting the need for integrating professional medical information resources with GenAI tools, so that the GenAI videos can be more appropriately used for health management purposes. By triangulating the core needs of older adults, the capabilities of GenAI, and their remaining gaps, we have found that the existing gaps limit the extent to which the health needs of older adults are met, and also inversely restrain the acceptance and use of GenAI videos by older adults.

Conclusions: This study explored the potential and gaps of using GenAI videos in older adults' health management. GenAI videos are prominent for self-management of health, but our triangular model reveals coexisting opportunities and challenges. While the core needs drive the development directions of GenAI video technologies, our study suggests that GenAI health videos can benefit from the improvements in technical capabilities, service innovation, localization, and integration of health resources.

(JMIR Aging 2026;9:e80661) doi:10.2196/80661

KEYWORDS

older adults; artificial intelligence; generative AI videos; health management; iterative workshop

Introduction

The global process of population aging is advancing at an unprecedented rate. United Nations' data indicate that the global population aged 60 years and older is projected to surge from approximately one billion in 2020 to nearly 2.1 billion by 2050 [1], with the proportion of this age group rising from 13% to 22% [2]. Effective acquisition of health knowledge and adoption of scientific health management strategies impact the quality of life and well-being of older adults [3]. From daily dietary planning and exercise regimen formulation to self-monitoring and management of chronic diseases [4], enhanced health knowledge and management capabilities empower older adults to better understand their health status [5], proactively prevent diseases, and thereby improve their autonomy and social participation [6]. Proper dietary combinations provide adequate nutrition to maintain physiological functions; moderate exercise enhances muscle strength, improves cardiopulmonary function, and optimizes sleep quality [7]. For older adults with chronic

XSL•FO
RenderX

conditions (eg, hypertension, diabetes, and cardiovascular diseases), effective health management controls disease progression, reduces complication risks, and lowers medical expenses [8]. Studies indicate that effective health management can decrease the incidence of chronic diseases and associated costs [9]. Older adults actively engaged in health management exhibit better mental health, with lower rates of depression and anxiety [10]. Thus, improving health management among older adults is essential for achieving active aging and intergenerational bonds [11].

In the context of rapid artificial intelligence (AI) advancements, leveraging generative artificial intelligence (GenAI) offers a new technology to assist older adults in health management by enhancing the efficient acquisition of health information and self-management [12]. GenAI, a cutting-edge technology integrating algorithms, such as deep learning and neural networks [13], excels not merely at content generation but, critically, at creating highly personalized, on-demand media. GenAI's advanced natural language processing [14], data analytics, and learning capacities allow it to transform complex medical knowledge into accessible content [15,16]. Traditional health videos often lack the personalization and real-time adaptability required for complex, individualized health needs [17]. GenAI directly addresses this gap. GenAI-based health videos have emerged as an innovative solution because they can dynamically tailor content based on users' demographics, specific health conditions, and preferences, for example, generating a short video on "low-sodium recipes for hypertension," improving relevance and engagement [18]. GenAI videos also enable interactive features, such as personalized recommendations and real-time feedback, which can further support older adults in navigating complex health topics [19]. GenAI videos can generate multilingual and culturally adapted content, reducing barriers for diverse older populations. Despite this immense potential to deliver granular, customized health information, the empirical study of GenAI video adoption and user experience in geriatric health management remains in its infancy.

Despite its potential, older adults face barriers in adopting AI technologies, including the GenAI-based ones. Age-related cognitive decline, manifesting as memory, attention, and reaction time impairments, hinders the understanding and operation of emerging technologies [20]. Neuroscientific research indicates that prefrontal cortex atrophy in older adults reduces multitasking capacity, complicating the execution of complex commands [21]. A historical lack of digital education leaves many unfamiliar with basic operations (eg, searching and installing apps), impeding independent use [22]. Information overload and the proliferation of misinformation further exacerbate challenges in accessing accurate health knowledge [23]. The internet's rapid, high-volume dissemination overwhelms older users, who struggle to discern reliable content [24]. Misleading health information, often disseminated via sensationalized headlines on social media, exploits these vulnerabilities, leading to misguided decisions [25]. Additionally, technology interfaces often neglect the physiological characteristics (eg, declining vision and hearing) and usage habits of older adults, resulting in poor usability [26]. Most devices and apps feature small fonts, complex icons, and convoluted workflows, rendering them inaccessible to those with sensory impairments [27]. Gesture-dependent interactions and ambiguous feedback mechanisms further increase operational difficulty [28]. Survey data reveal that fewer older adults proficiently use health management apps, with many expressing low trust in GenAI, preferring traditional consultations [29]. This digital divide not only limits GenAI videos' application in geriatric health management but also underscores urgent needs for age-friendly technological adaptations [30].

Currently, GenAI video applications in geriatric health management remain emergent, with designs and service models inadequately tailored to older adults' needs, necessitating in-depth research and improvement [31]. Existing tools exhibit unfriendly interaction modes, with content formats and delivery methods failing to accommodate cognitive traits and reading habits [32]. For instance, a study reports that older participants find GenAI videos' speech interaction too fast and their interfaces overly complex [33]. There is a gap regarding studies that focus specifically on older adults' interaction experiences, needs, and barriers when using GenAI for health. Against this backdrop, our research aims to investigate older adults' needs when using GenAI videos for health information acquisition and management, providing targeted recommendations for designing age-friendly GenAI health videos. This research addresses this identified gap in GenAI video applications for geriatric health management, offering evidence-based insights for developers, health care providers, and policy makers to construct human-centric, age-adapted intelligent health systems, ultimately advancing older adults' health and well-being. Based on these considerations, the following research questions are proposed:

1. What difficulties do older adults face, and what are their needs when using GenAI videos to access health information?
2. What functionalities and services can GenAI videos provide for geriatric health management, and where are the gaps?
3. How can technological design and interaction optimization enhance GenAI video's applicability and user experience in geriatric health management scenarios?

Addressing these questions will guide future research, offering theoretical and practical foundations for innovative GenAI video applications in geriatric health management. Resolving these issues will facilitate GenAI video's role in improving health management, narrowing the digital divide, and promoting social equity and harmony.

Methods

Participant Recruitment

With the assistance of community partners, we initially recruited 26 older adults residing in nearby communities to participate in an iterative GenAI health video workshop through offline posters. The following inclusion criteria were applied: Participants were aged 60 years or older, possessed basic operational skills with smart devices (eg, smartphones and tablet

XSL•FO
RenderX

PCs). Eventually, the results of 20 older adult participants were included in the study. As a qualitative study using thematic analysis, the final sample size of 20 participants was determined based on the principle of thematic saturation. Data collection and analysis proceeded iteratively, and after the 17th interview, minimal new codes or themes emerged, indicating that the core needs and challenges regarding GenAI videos had been sufficiently captured. The subsequent 3 interviews confirmed this saturation point, covering the hierarchical structure of needs. This sample selection was also designed to maximize information power through the deliberate inclusion of participants with diverse demographic factors (eg, varying educational levels and living arrangements) and heterogeneous health needs (eg, hypertension, joint health, and sleep disorders). This deliberate heterogeneity ensured a rich dataset and robust identification of core themes, justifying the sufficiency of the final sample size.

All participants expressed motivation and a need to leverage GenAI and health videos for better self-health management. Participants exhibited diverse health conditions (chronic diseases, healthy, and mild conditions) and living arrangements (eg, living alone, with family, or in care facilities). To clarify the health status definitions: Chronic Disease referred to participants who reported having a long-term medical condition requiring ongoing management, such as hypertension, hyperlipidemia, or diabetes. Mild condition referred to those who reported minor, nonchronic health issues or temporary discomforts that did not impair their daily function. Healthy referred to participants who did not report any diagnosed chronic disease or ongoing medical conditions.

In addition, community workers participated in the research. These professionals are engaged in community health-related work, possess a background in medical knowledge, and have experience assisting older adults with technological tools. Participants voluntarily joined the workshop and did not receive any additional compensation, though tea and snacks were offered. It is important to note that the participant sample, limited to 20 individuals from a single urban context and skewed heavily female (80%), limits the generalizability of the findings and may not fully capture the diverse experiences of the broader older adult population, particularly men and those in rural or low-income settings. We confirmed with the recruiting community workers that this gender imbalance is a common limitation in community-based research, reflecting the higher propensity of female older adults to actively participate in community health and social activities. This limitation is discussed further in the Discussion section.

Research Design

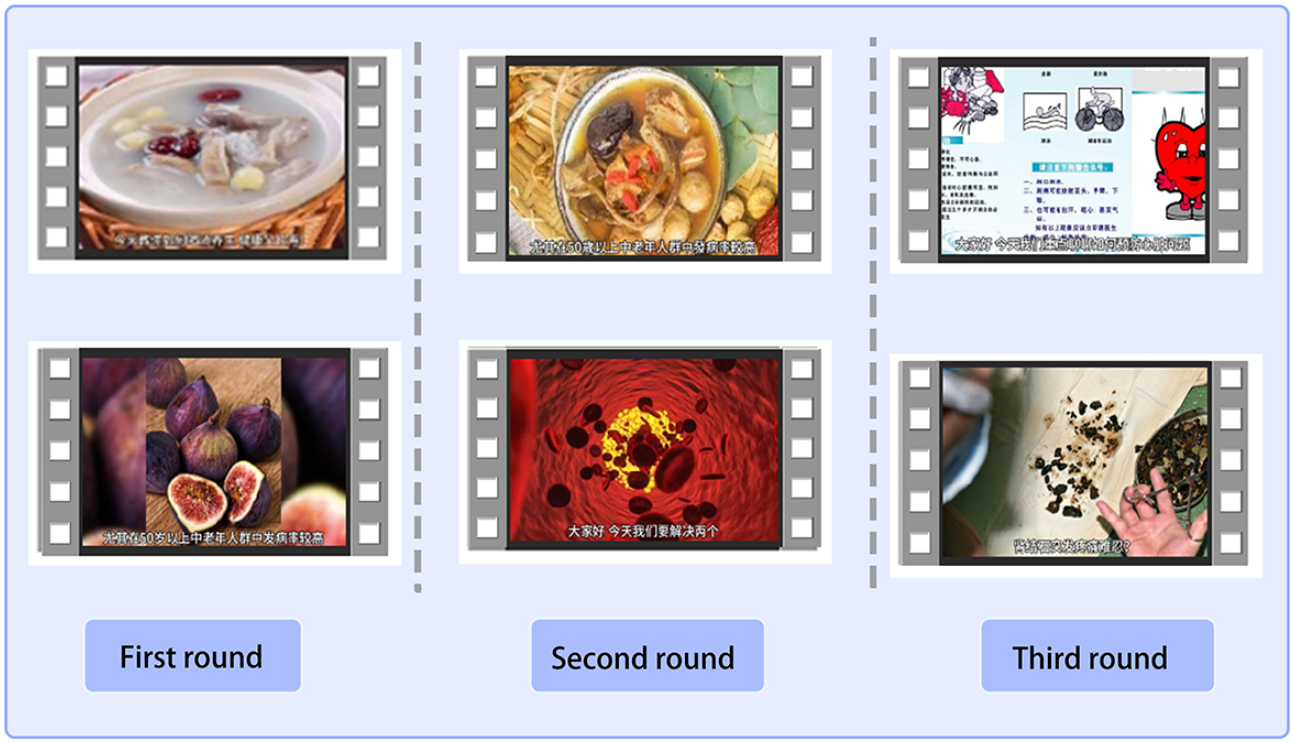
To reduce older adults' cognitive load, the study adopted a participatory iterative approach. Each participant received a 15-minute introduction to GenAI to ensure their basic understanding. Before the start of the workshop, a pretest was conducted to collect baseline data, such as participants' GenAI videos perceptions, technical skills, and self-evaluations of prior knowledge.

The methodology and approach were initially guided by established theoretical models, specifically the Technology Acceptance Model (TAM) [34] and the concept of self-efficacy [35]. We used TAM, which posits that perceived usefulness and perceived ease of use predict technology adoption, to structure the pre and posttest data collection and the iterative refinement process, focusing on understanding factors that influence the acceptance of GenAI-generated health videos among older adults.

The workshop implementation comprised four steps: (1) Participants proposed a health-related topic that suited their personal health needs. Then, they provided this health topic to GenAI, which converted it into a prompt for generating the GenAI video. They provided the video script generated by GenAI to the video production software CapCut to create the first video. Before generating the video scripts for GenAI, we conducted pretraining on them. We set the generated video scripts to introduce diseases or health issues, solutions and countermeasures, suggested structure generation, and recommended a duration of 3 - 5 minutes for short videos to prevent overly long video content from causing cognitive overload for older adults; (2) participants viewed the initial video, discussed their experiences, provided feedback, and proposed new requirements, which were refined into updated prompts for the second video generation; (3) this iterative process was repeated for a total of 3 cycles, culminating in a final optimized video; and (4) after each video generation, community health workers reviewed the content, flagged potential inaccuracies (eg, incorrect medical advice or misleading health claims), and explained the rationale to participants, ensuring discussions were grounded in reliable information. Finally, a posttest was performed to reassess satisfaction, acceptance, and needs after interacting with GenAI videos. The content and process of generating the videos are shown in Figure 1.

XSL•FO
RenderX

Figure 1. The content and process of generating the videos.



Data Collection and Analysis

Data sources included pre- and posttest interview transcripts. We conducted semistructured interviews in participants' native language (Cantonese) with audio recording to capture the challenges encountered by the participants and their preferences during the consumption of GenAI videos. The interview transcripts were analyzed with the thematic analysis approach following Braun and Clarke's 6-phase framework [36]. To ensure data reliability, 2 researchers [TL and PP] independently entered and cross-checked data and resolved discrepancies when disagreements occurred.

Ethical Considerations

This study was approved by the Human Ethics Committee of Macao Polytechnic University (Approval No.: HEA006-FCA-2025). All participants provided written informed consent before interviews, ensuring comprehension of the study's purpose, procedures, and potential impacts. All participants have signed an informed consent form before participation. Data were anonymized, with strict confidentiality maintained for personal identifiers. Before workshops, participants were informed that AI-generated content should supplement professional advice, with consent forms emphasizing technical limitations. To reduce the chance of adverse effects caused by misinformation from GenAI videos, community health workers guided the workshops, in which they confirmed the validity of information and stopped the video generation when problems occurred. No compensation was offered to the participants for their involvement in the study.

Results

Participant Characteristics

In total, 20 older adults participated (Table 1). Age ranged from 60 to 84 years (mean 71.4, SD 8.16): 8 out of 20 (40%) aged 60 - 69, 9 out of 20 (45%) aged 70 - 79, and 3 out of 20 (15%) aged ≥80. Gender skewed female (16/20, 80%) versus male (4/20, 20%), possibly reflecting higher female community engagement. Health status: 10 out of 20 (50%) had chronic diseases (hypertension, hyperlipidemia, etc), 8 out of 20 (40%) were healthy, and 2 out of 20 (10%) had mild conditions, highlighting prevalent health management needs. Living arrangements: 9 out of 20 (45%) lived alone, 11 out of 20 (55%) with family, suggesting varying reliance on external support. Education: 7 out of 20 (35%) primary school, 5 out of 20 (25%) junior high, 4 out of 20 (20%) high school, 4 out of 20 (20%) college or university, indicating diverse digital literacy levels. GenAI video experience: 17 out of 20 (85%) were naive users, 3 out of 20 (15%) had limited exposure, ensuring exploration of initial adoption needs. Health topics varied: joint health (n=5), diet (n=4), hypertension (n=3), sleep disorders (n=2), cardiovascular health (n=1), general health (n=2), exercise (n=1), obesity (n=1), and thyroid disease (n=1), reflecting diverse needs and opportunities for GenAI videos' applications.

XSL•FO
RenderX

Table . Participant characteristics.

Characteristics	Values (N=20), n (%)
Age (years)	
60 - 69	8 (40)
70 - 79	7 (35)
≥80	5 (25)
Sex	
Female	16 (80)
Male	4 (20)
Health status	
Chronic disease	12 (60)
Healthy	5 (25)
Mild disease	3 (15)
Living situation	
Living with family	13 (65)
Living alone	7 (35)
Education level	
Primary school or below	9 (45)
Junior high	5 (25)
High school	4 (20)
University or above	2 (10)
GenAI[a] experience	
No	16 (80)
Yes	4 (20)
Health topics of interest	
Hypertension or hyperlipidemia or hyperglycemia	3 (15)
Sleep disorders	3 (15)
Cardiovascular health	1 (5)
Healthy diet	4 (20)
Obesity	1 (5)
Joint health	4 (20)
Exercise health	1 (5)
Thyroid issues	1 (5)

[a]GenAI: generative artificial intelligence.

Thematic Analysis

Theme 1: Technical and Accessibility Barriers

Older adults faced foundational challenges in accessing and operating GenAI video tools, stemming from limited digital literacy and interaction difficulties. These barriers highlight a cognitive load required for adoption. Cognitive and digital literacy limitations were prevalent, with participants struggling with fundamental technological tasks. Participant #4 noted, “I have a poor memory and learn new things slowly,” a common reflection of age-related cognitive decline that hinders mastering complex workflows. Participant #12 admitted, “I am illiterate,” revealing the impact of basic educational background on technology adoption and independence. Operational and interface design flaws further compounded these issues, especially given physical challenges. Participant #16 reported, “Poor eyesight affects my typing, so I can only use voice or ask my granddaughter for help,” illustrating a hardware-software mismatch. Participant #14 criticized the “complex button functions” and convoluted workflows that were difficult to follow without prior experience. Consequently, participants often expressed feelings of exclusion, with #13 stating, “GenAI is currently more youth-friendly; it feels designed to make me make mistakes.” Furthermore, localization and support gaps

XSL•FO
RenderX

complicated accessibility, such as #4's explicit request for "Cantonese voice guidance" for better comprehension, and #8's frustration with having to rely on young family members "for downloading and setting up every new app." These findings underscore systemic shortcomings in age-friendly interface design and support structures.

Theme 2: Trust and Accuracy Expectations

Participants demonstrated a heightened level of caution and demanded rigorous verification regarding the accuracy and reliability of AI-generated health information, indicating a trust deficit. Concerns about misinformation were widespread, with participants often viewing GenAI as a source that requires external validation. Participant #3 expressed direct concerns about being misled, stating, "I worry about being misled by false content from GenAI, especially with medical advice that could harm me." This led to demands for professional verification and transparency. Participant #14 strongly emphasized the need for human input: "GenAI outputs must be validated by professionals, like a pharmacist or doctor, before I act on it." Participant #8 reiterated this cautious approach, insisting, "I will consult a doctor before acting on GenAI's advice; the video is just a starting point." To build trust, participants requested greater transparency, with #12 advocating for "source citations in AI-generated health information, showing where the medical facts came from," and #16 suggesting "alerts for abnormal data or warnings if the advice is outside standard guidelines." This cautious trust stems from their awareness of technological limitations, as #9 cautioned, "GenAI video might misdiagnose my health condition if I input the wrong symptoms." Despite these concerns, participants acknowledged the medium's potential, with #7 praising its format: "Videos are more intuitive than text; I can see how an exercise should be done."

Theme 3: Personalized Health Guidance

Participants expressed clear expectations for personalized functionalities that extended beyond generic content, spanning monitoring, risk assessment, and localized solutions. Demands for tailored monitoring and trend analysis were prominent, reflecting needs related to self-management of chronic diseases. Participant #1 requested the tool to provide "automated trend analysis of my blood glucose data over the last 2 weeks, not just today's number," while #4 needed specific, daily "calorie, intake calculations for my low-sodium diet." For risk assessment and nonpharmacological solutions, participants sought customized advice specific to their conditions. Participant #9 inquired about "joint health evaluations based on my walking data," and #17 sought "nonpharmacological hypertension solutions tailored to my limited mobility." The need for cultural and localized adaptation was also a key theme, demonstrating that personalization must consider the daily life context. Participant #4 requested "healthy recipe modifications for local Cantonese dishes" to make diet changes practical, and #10 desired "local hospital lookup features integrated into the video advice." Current GenAI tools fall short of these nuanced expectations. Participant #12 criticized the content as "too-fast, oversimplified video content that doesn't answer my specific questions," revealing a gap between generic output and true personalization.

Theme 4: Basic Health Management Needs

While GenAI videos demonstrated capability in offering foundational health functions, gaps persisted in terms of depth, integration, and systematic support. Foundational functions and clarity showed progress, such as #3 learning about heart disease basics and #9 understanding osteoporosis risks through visual content. However, limitations were evident in addressing multifaceted health scenarios. Participant #1's need for combined management of diabetes and cardiovascular concerns went unaddressed by single-topic videos, illustrating a lack of integrated care guidance. Similarly, #9's need for personalized solutions for concurrent osteoporosis and lung health issues lacked comprehensive guidance due to insufficient data integration. Gaps in professional support were clearly visible, with #14's request for "thyroid self-check guidance" and #19's need for a "local hospital locator" being unfulfilled. Consequently, participants often defaulted to traditional support channels, such as #8's preference for "doctor-confirmed advice" and #15's reliance on "family assistance" to navigate complex information, highlighting the tools' current inability to serve as a reliable, standalone health resource.

Theme 5: Diverse Health Concerns

The older adult population exhibited a highly diversified spectrum of health management needs, necessitating a flexible and broad application scope for GenAI video content. Chronic disease management was a central concern, with #1 explicitly highlighting the need for diabetes prevention, emphasizing blood glucose trend monitoring and personalized dietary recommendations. Health literacy through visualization was sought for complex topics, such as #3, focusing on cardiovascular health and stressing the importance of using GenAI video animations to understand electrocardiogram principles. Needs extended to highly specific daily life and preventative issues, including joint and sleep health, where #15 stated, "I want to understand sleep issues as I have had sleep apnea for over a decade," and #12 articulated comprehensive requirements for joint pain relief and vitamin supplementation. Furthermore, dietary and safety validation emerged as a distinct concern, with #12 asking, "I heard honeydew melon and eggs are harmful together—is this true?" reflecting a need for rapid validation of common health myths. This diversity confirms that GenAI videos must shift from single-disease treatment to comprehensive, highly specialized prevention and daily health information services in geriatric health management.

The abovementioned themes were summarized into three high layers of needs, representing the needs of (1) technical accessibility, (2) age-friendly design adaptations, and (3) integration with professional medical information resources. Table 2 lists the layers, themes, and subthemes described in this subsection.

XSL•FO
RenderX

Table . Summary of results of thematic analysis.

Layers, themes, and subthemes	Key insights from interview texts
Layer 1: Technical accessibility	
Technical and accessibility barriers: digital literacy limitations and language and localization gaps	• Inability to operate tools independently (eg, typing, navigating interfaces) • Insufficient Cantonese support and local specific health resources
Layer 2: Age-friendly design adaptations	
AI usage barriers: digital literacy challenges and interface simplicity needs	• Difficulty in operating complex interfaces due to poor memory or illiteracy • Demand for voice interaction, large fonts, and step-by-step guidance
Trust and accuracy expectations: cautious trust in AI[a] and localization and age-friendly features	• Openness to try but requiring human verification for critical decisions • Preference for Cantonese voice, localized services, and age-adapted design
Layer 3: Integration with professional medical information resources	
Personalized health guidance: dynamic health monitoring, visualized knowledge delivery, and multimodal interaction	• Autogenerated reports (eg, blood glucose trends and sleep reports) • Animated videos explaining medical concepts (eg, ECG[b] and osteoporosis) • Voice-guided operations and large-font interfaces
Basic health management needs: symptom identification, nonpharmacological solutions, professional medical diagnosis, and critical decision-making	• Early warning for osteoporosis, cardiovascular risks, and digestive issues • Sleep improvement tips and joint protection exercises • Inability to replace doctors for definitive diagnoses (eg, osteoporosis and sleep apnea) • Lack of human verification for health suggestions (eg, medication and surgery)
Diverse health concerns: chronic disease management, joint health and injury prevention, sleep health improvement, and digestive health and dietary safety	• Need for personalized monitoring and guidance on hypertension, diabetes, and cardiovascular diseases • Guidance on joint protection exercises and fall prevention during travel • Nonpharmacological solutions for insomnia and sleep apnea • Personalized dietary recommendations and food compatibility validation

[a]AI: artificial intelligence.

[b]ECG: electrocardiogram.

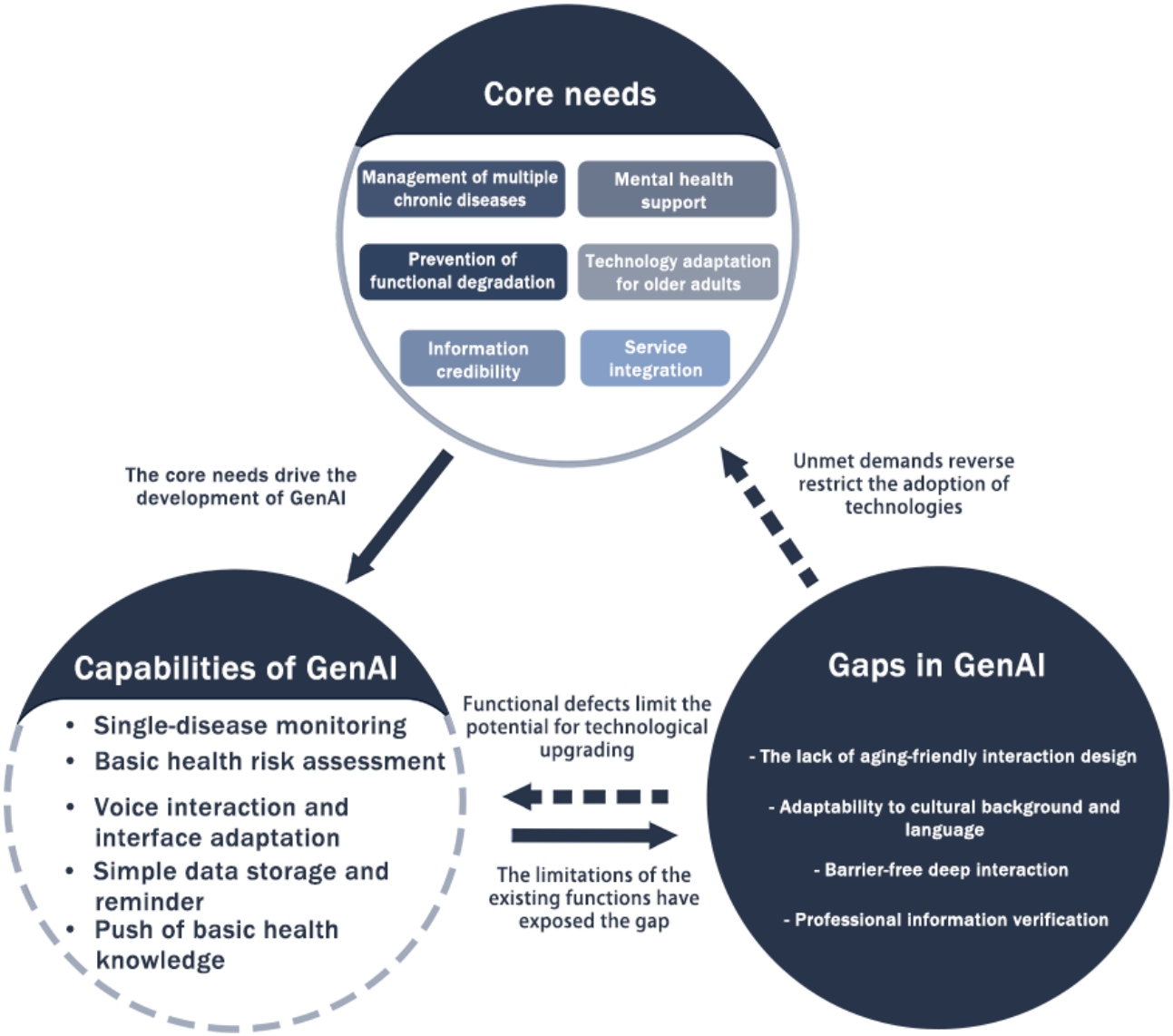
Discussion

Principal Findings

The results demonstrate the interactive relationship among the needs of older adults in using GenAI videos for health management, the capabilities of GenAI videos, and the existing gaps in GenAI video tools, as synthesized in Figure 2. The core needs of older adults drive the development and application direction of GenAI video technology, promoting the realization of its functions in areas such as health monitoring and information interaction. However, the current capabilities of GenAI video tools reveal unmet gaps in areas such as multidisease integration and professional medical diagnosis. On one hand, these gaps directly limit the extent to which the health needs of older adults are met; on the other hand, they also inversely restrain the further acceptance and use of GenAI video technology by older adults. As such, this forms a dynamic triangular relationship that requires more attention among user needs, GenAI capabilities, and the gaps to be filled.

XSL•FO
RenderX

Figure 2. Dynamic relation of the needs, capabilities, and gaps of generative artificial intelligence videos. GenAI: generative artificial intelligence.



This closed loop shows that the digital transformation in the field of geriatric health management is not a linear process of progress but rather a complex relationship where the potential of technological empowerment coexists with the challenges of age-friendly adaptation. Achieving sustainable alignment between technological development and user needs requires systematic design to balance efficiency and humanistic care. Table 3 also provides a detailed list of the specific context regarding the needs of older adults, the capabilities of GenAI video, and the gaps in GenAI video, which highlights the multidimensional characteristics and the mismatch of needs of current tools.

XSL•FO
RenderX

Table . The comparison of the needs, capabilities, and gaps of generative artificial intelligence videos.

Core dimension	Core needs	Capabilities of GenAI[a]	Gaps in GenAI
Health needs coverage	Comprehensive management of multiple chronic diseases, personalized health guidance, and mental health support	Single-disease monitoring (eg, blood glucose) and basic health risk assessment	Integrated management of multiple diseases, professional medical diagnosis, and solutions for complex health issues
Technical adaptability	Simple operations, cultural adaptation (dialect or localized content), and barrier-free interaction	Basic voice interaction and large-font interface	Fully text-free operation, automated complex tasks, and deep cultural customization
Information interaction	Authoritative and trustworthy information, personalized recommendations, and human verification channels	Basic information prompts and simple data recording	Medical-grade information verification, dynamic credibility rating, and real-time human validation
Service integration	Integration of medical, family, and community resources and end-to-end health management	Health data storage and basic consultation services	Real-time integration with medical systems and family-community resource network

[a]GenAI: generative artificial intelligence.

The health needs of older adults show a trend toward diversification and refinement, covering multiple areas, such as chronic disease management, prevention of functional decline, and mental health support, with highly heterogeneous combinations of needs among different individuals. At the level of technology use, insufficient digital literacy and interface interaction barriers constitute the main barriers to use, manifesting as fear of complex operations, dependence on clear guidance, and an urgent demand for culturally adapted features. In terms of trust, older adults exhibit a conservative attitude, recognizing the efficiency advantages of GenAI in information presentation and basic services while maintaining reservations about the reliability of professional medical advice and emphasizing the importance of human verification and professional support. Regarding functional supply, existing GenAI video tools have made initial progress in standardized service areas, such as basic health monitoring and visualized knowledge delivery, but they still show deficiencies in the depth of personalized adaptation, the ability to handle complex scenarios, and the level of integration with professional medical services.

Multidimensional Needs of Older Adults Using GenAI Video for Health Management

The core needs of older adults for GenAI-driven health management exhibit a hierarchical structure, which can be systematically deconstructed into 3 progressive dimensions. The first layer focuses on technological accessibility and operational inclusivity, manifesting as urgent demands for simplified interaction workflows [37], enhanced multimodal interactions [38], and interface designs adapted to physiological decline [39]. These needs stem from the structural deficiencies in digital literacy among older adults, characterized by barriers in comprehending complex technical logic and heightened sensitivity to operational memory load [40]. This cognitive load is not merely an individual failing but a manifestation of the digital divide, wherein technologies often neglect the specific cognitive and physical requirements of older users, reflecting ageism in design standards.

The second layer emphasizes personalized and precise health management adaptations, encompassing dynamic risk assessments based on longitudinal health data, culturally sensitive content delivery, and cross-scenario health guidance. This layer reflects a paradigm shift from passive acceptance to active participation in health management.

The last layer centers on professional medical support and system integration capabilities, including disease risk prediction models approaching clinical decision-support levels, interoperable health data exchange across medical institutions, and integrated emergency response systems. Such needs highlight a functional expectation shift from “information assistance” to “decision participation” for GenAI video tools. The corresponding challenges arise from multifaceted technical and social constraints. On the technical front, digital divides manifest as high learning costs, lack of age-friendly design, and complex functional logic [41]. Socially, these challenges are reflected in the lagging digital transformation of medical trust systems [42], fractured intergenerational digital support networks [43], and uneven regional health resource distribution [44]. The interplay between these needs and barriers underscores a structural mismatch in digital transformation. Specifically, the contradiction between standardized technological supply and personalized health demands, as well as the imbalance between digital access opportunities and capacity building. The persistent reliance on human verification is closely tied to low eHealth literacy, where users lack the critical skills to evaluate information quality, amplifying a lack of trust in nontraditional sources such as GenAI.

Capabilities of GenAI Video for Older Adults’ Health Management

GenAI video has achieved foundational breakthroughs in older adults’ health management across 3 interconnected dimensions. First, information delivery efficiency has improved [45], with natural language processing and multimodal interactions transforming medical knowledge into accessible formats, such as animated explanations and voice narrations, thereby lowering health information acquisition barriers [45]. Second, the level of service automation has increased, as big data analytics and

XSL•FO
RenderX

rule-based engines enable health monitoring, basic risk assessments, and standardized recommendations [46]. Third, accessibility has been enhanced through cloud computing and mobile internet technologies, transcending temporal and spatial constraints [47]. These advancements have facilitated a shift from experience-driven to data-driven health management. However, the operational scope of GenAI video tools remains generic, while video generation with complex information or expert knowledge still requires human guidance [48]. GenAI videos' role remains tool-oriented, and people tend to use them for enhancing efficiency rather than replacing professional medical opinions.

Gaps of GenAI Video for Older Adults' Health Management

Current applications of GenAI videos exhibit limitations that require breakthroughs in several areas. From the health care perspective, there is not enough integration of accurate health information in managing complex chronic diseases [49]. Regarding emotional interaction, the system shows a weak capacity to interpret unstructured health demands, such as ambiguous expressions, or provide empathetic responses [50]. Concerning system integration, barriers to interoperability within health care ecosystems persist [31]. Additionally, the cultural sensitivity of GenAI video tools remains insufficient, with inadequate adaptation to regional variations and traditional health beliefs, such as limited support of different spoken languages [51]. Technologically, it is essential to develop lightweight, age-friendly solutions, such as offline voice assistants and multimodal sensors, and enhance algorithm transparency. Regarding service models, exploring "AI+human expert" hybrid collaboration mechanisms will be crucial. Policy-wise, establishing cross-sector digital health standards, including data security protocols and service quality metrics, is necessary. Socially, activating family-community-health care collaboration networks will foster an inclusive ecosystem. The improvement strategies involve constructing a collaborative framework across technology, health care, and social domains. On the health care side, establishing clinical validation standards for AI-assisted diagnostics is crucial [52]. From the societal point of view, developing intergenerational digital support networks and community-based digital literacy programs will be necessary [53]. These gaps, particularly the lack of integrated, verified professional support, reinforce the existing digital divide by making the tool unreliable for those with lower health literacy, potentially deepening health inequities. These measures will facilitate the evolution of GenAI videos from a "basic service" to a "comprehensive health management partner."

Practical Implications

Based on the identified hierarchical needs and capability gaps, our findings offer concrete, actionable recommendations for key stakeholders to enhance the real-world applicability of GenAI health videos. For AI developers and designers, this means implementing mandatory Age-friendly interface design principles, prioritizing simplified, one-step workflows to reduce cognitive load, large-font visual interfaces, and enhanced multimodal interactions, especially high-accuracy dialect-supported voice control to overcome digital literacy barriers. Developers must also integrate transparent features showing the data source and reliability score for all medical claims. In parallel, health care organizations and policy makers should establish clinical vetting standards and certification processes for AI-generated health content to bolster trust. Policy makers should mandate the development of interoperable health systems that allow secure, real-time data exchange between GenAI tools and eHealth records, enabling true personalized, integrated chronic disease management. Community and social support networks have a crucial role in investing in community-based digital literacy programs that specifically teach older adults how to critically evaluate AI-generated health information, rather than just teaching basic operations. Implementation guidelines should focus on hybrid models where community health workers act as the "human verification" bridge between the AI tool and the end user.

Ethical Considerations in GenAI Video for Older Adults' Health Management

The application of GenAI videos in geriatric health management raises complex ethical issues that necessitate a comprehensive framework across technological design, service delivery, and policy regulation [54]. First, the balance between informed consent and autonomy presents a core challenge. Cognitive decline may hinder older adults' understanding of AI risks, while the over-reliance on proxies risks infringing their autonomy [55]. Solutions include developing visualized consent processes and clearly defining human-AI responsibility boundaries [56]. Second, algorithmic fairness and health inequities require urgent attention. Training data biases may exacerbate disparities, while hardware limitations could exclude economically disadvantaged groups [57]. Mitigation strategies involve privacy-preserving data integration and promoting affordable, age-friendly devices [58]. Third, the conflict between data privacy and generational security is prominent [59]. Finally, ambiguity in responsibility attribution challenges traditional medical ethics. Errors in AI-generated advice involve multiple stakeholders, including developers, clinicians, and users, necessitating decision-tracking systems and legislation for scenario-specific accountability [60]. These challenges underscore the need for a governance framework that balances efficiency, equity, and security through interdisciplinary collaboration.

Recommendations for Future Research

The sustainable development of GenAI videos for older adults' health management requires integrated strategies across multiple dimensions, which sheds light on the directions of future research. The ultimate goal is to create an inclusive health ecosystem where GenAI video empowers rather than replaces human care. Future research should focus on designing adaptive interfaces that accommodate the diverse digital literacy levels of older adults. This includes developing progressive disclosure mechanisms and context-aware guidance systems to reduce cognitive load, alongside iterative usability testing with age-diverse cohorts to refine interaction paradigms. From a gerontological standpoint, longitudinal studies are needed to map the dynamic evolution of older adults' digital competencies, identifying age-related changes in technology adoption patterns,

XSL•FO
RenderX

generational differences, culturally tailored content, and community-based digital literacy programs. For computer science researchers, the priority lies in advancing GenAI algorithms to enhance personalization accuracy. By integrating insights from these 3 domains, we can develop GenAI video systems that not only optimize technological efficiency but also embed humanistic values, ultimately fostering active aging through inclusive, empowering, and sustainable digital solutions.

Limitations

We acknowledge several limitations in this work. First, the single-point-in-time design limits our ability to capture the dynamic changes in older adults' digital literacy and GenAI acceptance over time; the barriers identified here may change rapidly as users gain experience, and a longitudinal study would be necessary to map the learning curve and long-term acceptance trajectory. Second, the sample primarily comprised urban, middle- to high-income older adults, with a notable skew toward female participants. While this skew reflects higher community engagement among this demographic, it may overestimate the average digital literacy level and underrepresent the unique health management needs and severe technological barriers faced by older adults in rural or low-income settings, particularly men. Finally, the qualitative nature of this study provides rich, deep insights into user needs but limits the statistical generalizability of the findings. While these limitations persist, they provide valuable directions for future research.

Conclusions

Based on a series of iterative workshops, this study systematically explored the opportunities and challenges of using GenAI videos for older adults' health management, yielding the following key conclusions. First, the needs of older adults for GenAI video exhibit a clear hierarchical structure: their fundamental needs focus on technological accessibility, then move to the emphasis of personalized adaptation, and the highest layer of needs prioritizes the integration with professional health information. The study confirms that GenAI videos have demonstrated potential in geriatric health management, notably enhancing information delivery efficiency, lowering the barriers of health knowledge acquisition, and transcending temporal and spatial constraints of health management. However, notable gaps remain between the features of GenAI video tools and older adults' needs, particularly in the depth of personalization, the capabilities of complicated responses, localization, as well as the accuracy and reliability of medical knowledge. The value of our research lies in recognizing the advantages and limitations of GenAI videos in this context, which can shed light on future research and solutions that can benefit older adults.

Acknowledgments

We are grateful to the Macao Federation of Trade Unions for their assistance in this research.

Funding

This research was funded by the Macao Science and Technology Development Fund (FDCT; funding ID 0032/2025/ITP1) and Macao Polytechnic University research grant (project code RP/FCA-24/2025).

Data Availability

The data used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

TL and YL (YTL) proposed the idea to conceive and design this study. TL analyzed the data. TL completed the manuscript with the assistance of YL (YTL) and ZH, PP, KC, and YL (YTL) supervised students and oversaw this work.

Conflicts of Interest

None declared.

References

1. Paul A, Nag P. Global trends of population dynamics oncoming 2050. 2022 URL: https://apgin.org/wp-content/uploads/2024/08/Journal_June_2024_Final-7-32.pdf?utm_source=chatgpt.com [accessed 2026-02-14]
2. Higo M, Khan HT. Global population aging: unequal distribution of risks in later life between developed and developing countries. Glob Soc Policy 2015 Aug;15(2):146-166. [doi: 10.1177/1468018114543157]
3. Rony MKK, Parvin M, Wahiduzzaman M, Akter K, Ullah M. Challenges and advancements in the health-related quality of life of older people. Adv Public Health 2024 Feb 1;2024:1-18. [doi: 10.1155/2024/8839631]
4. Burke LE, Warziski M, Starrett T, et al. Self-monitoring dietary intake: current and future practices. J Ren Nutr 2005 Jul;15(3):281-290. [doi: 10.1016/j.jrn.2005.04.002] [Medline: 16007557]
5. Kayser L, Karnoe A, Duminski E, Somekh D, Vera-Muñoz C. A new understanding of health related empowerment in the context of an active and healthy ageing. BMC Health Serv Res 2019 Apr 24;19(1):242. [doi: 10.1186/s12913-019-4082-5] [Medline: 31014350]

XSL•FO
RenderX

6. Evans JG. Prevention of age-associated loss of autonomy: epidemiological approaches. J Chronic Dis 1984 Jan;37(5):353-363. [doi: 10.1016/0021-9681(84)90101-2]
7. Shao T, Verma HK, Pande B, et al. Physical activity and nutritional influence on immune function: an important strategy to improve immunity and health status. Front Physiol 2021;12:751374. [doi: 10.3389/fphys.2021.751374] [Medline: 34690818]
8. Kastner M, Cardoso R, Lai Y, et al. Effectiveness of interventions for managing multiple high-burden chronic diseases in older adults: a systematic review and meta-analysis. CMAJ 2018 Aug 27;190(34):E1004-E1012. [doi: 10.1503/cmaj.171391] [Medline: 30150242]
9. Mbata AO, Soyege OS, Nwokedi CN, et al. Preventative medicine and chronic disease management: reducing healthcare costs and improving long-term public health. IJMRGE 2024;5(6):1584-1600. [doi: 10.54660/.IJMRGE.2024.5.6.1584-1600]
10. Reynolds CF 3rd, Jeste DV, Sachdev PS, Blazer DG. Mental health care for older adults: recent advances and new directions in clinical practice and research. World Psychiatry 2022 Oct;21(3):336-363. [doi: 10.1002/wps.20996] [Medline: 36073714]
11. Teater B. Intergenerational programs to promote active aging: the experiences and perspectives of older adults. Act Adapt Aging 2016 Jan 2;40(1):1-19. [doi: 10.1080/01924788.2016.1127041]
12. Fauziningtyas R. Empowering age: bridging the digital healthcare for older population. In: Digital Healthcare, Digital Transformation and Citizen Empowerment in Asia-Pacific and Europe for a Healthier Society: Elsevier; 2025:593-616. [doi: 10.1016/B978-0-443-30168-1.00015-3]
13. Buduma N, Buduma N, Papa J. Fundamentals of Deep Learning: Designing next-Generation Machine Intelligence Algorithms: " O'Reilly Media, Inc." 2022.
14. Luo Y, Pang PCI, Chang S. Enhancing exploratory learning through exploratory search with the emergence of large language models. Presented at: 58th Hawaii International Conference on System Sciences (HICSS-58).; Jan 7-10, 2025. [doi: 10.24251/HICSS.2025.007]
15. Pang PCI, Chang S, Verspoor K, Pearce J. Designing health websites based on users' web-based information-sdeeking behaviors: a mixed-method observational study. J Med Internet Res 2016 Jun 6;18(6):e145. [doi: 10.2196/jmir.5661] [Medline: 27267955]
16. Pang PI, Chang S, Pearce J, Verspoor K. Online health information seeking behaviour: understanding different search approaches. Presented at: Proceedings of the 18th Pacific Asia Conference on Information Systems (PACIS 2014); Jun 24-28, 2014 URL: https://aisel.aisnet.org/pacis2014/229/?utm_source=chatgpt.com [accessed 2026-02-14]
17. Singh P, Kushwaha AKS. Revolutionizing healthcare through optimized video retrieval. In: The Impact of Algorithmic Technologies on Healthcare: Wiley and Scrivener Publishing; 2025:177-188. [doi: 10.1002/9781394305490]
18. Rashid Z, Ahmed H, Nadeem N, Zafar SB, Yousaf MZ. The paradigm of digital health: AI applications and transformative trends. Neural Comput & Applic 2025 Jun;37(17):11039-11070. [doi: 10.1007/s00521-025-11081-0] [Medline: 35250181]
19. Crista V, Martinho D, Marreiros G. A multi-agent system approach with generative AI for improved elderly daily living. In: EPIA Conference on Artificial Intelligence: Springer; 2024:128-140. [doi: 10.1007/978-3-031-73497-7_11]
20. Schieber F. Human factors and aging: identifying and compensating for age-related deficits in sensory and cognitive function. In: Influences of Technological Change on Individual Aging: Springer; 2003:42-84 URL: https://scispace.com/pdf/human-factors-and-aging-identifying-and-compensating-for-age-472kla321m.pdf [accessed 2026-02-14]
21. Li KZH, Bherer L, Mirelman A, Maidan I, Hausdorff JM. Cognitive involvement in balance, gait and dual-tasking in aging: a focused review from a neuroscience of aging perspective. Front Neurol 2018;9:913. [doi: 10.3389/fneur.2018.00913] [Medline: 30425679]
22. Guri-Rosenblit S. Distance education in the digital age: common misconceptions and challenging tasks. International Journal of E-Learning & Distance Education/Revue internationale du E-learning et la formation à distance 2009;23(2):105-122 [FREE Full text]
23. Vivion M, Reid V, Dubé E, Coutant A, Benoit A, Tourigny A. How older adults manage misinformation and information overload - A qualitative study. BMC Public Health 2024 Mar 21;24(1):871. [doi: 10.1186/s12889-024-18335-x] [Medline: 38515081]
24. Demartini G, Mizzaro S, Spina D. Human-in-the-loop artificial intelligence for fighting online misinformation: challenges and opportunities. IEEE Data Eng Bull 2020;43(3):65-74 [FREE Full text]
25. Pittrow L. Medical misinformation on social media: assessing the ethical responsibilities of healthcare professionals, social media platforms, and users in combating the spread of false medical information. : Vilnius University; 2025 URL: https://epublications.vu.lt/object/elaba:238872406/ [accessed 2026-02-14]
26. Ramdowar H, Khedo KK, Chooramun N. A comprehensive review of mobile user interfaces in mHealth applications for elderly and the related ageing barriers. Univ Access Inf Soc 2024 Nov;23(4):1613-1629. [doi: 10.1007/s10209-023-01011-z]
27. Khamaj A, Ali AM. Examining the usability and accessibility challenges in mobile health applications for older adults. Alexandria Engineering Journal 2024 Sep;102:179-191. [doi: 10.1016/j.aej.2024.06.002]
28. Mihajlov M, Law ELC, Springett M. Intuitive learnability of touch gestures for technology-naive older adults. Interact Comput 2015 May 1;27(3):344-356. [doi: 10.1093/iwc/iwu044]
29. Wong AKC, Lee JHT, Zhao Y, Lu Q, Yang S, Hui VCC. Exploring older adults' perspectives and acceptance of ai-driven health technologies: qualitative study. JMIR Aging 2025 Feb 12;8:e66778. [doi: 10.2196/66778] [Medline: 39937162]

XSL•FO
RenderX

30. Yu C. Will AI be a Boon to an Aging Society. Thesis Commons: Center for Open Science Preprint posted online on 2024. [doi: 10.31237/osf.io/a8suh]
31. Anshari M, Pablos PO, Almunawar MN. AI-powered digital health: bridging the gap for elderly populations in the evolving healthcare ecosystem. In: Digital Healthcare in Asia and Gulf Region for Healthy Aging and More Inclusive Societies: Elsevier; 2024:265-281. [doi: 10.1016/B978-0-443-23637-2.00008-4]
32. McLaughlin A, Pak R. Designing Displays for Older Adults: CRC press; 2020. [doi: 10.1201/9780429439674]
33. Herrmann B. The perception of artificial-intelligence (AI) based synthesized speech in younger and older adults. Int J Speech Technol 2023 Jul;26(2):395-415. [doi: 10.1007/s10772-023-10027-y]
34. Portz JD, Bayliss EA, Bull S, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. J Med Internet Res 2019 Apr 8;21(4):e11604. [doi: 10.2196/11604] [Medline: 30958272]
35. Bandura A. Self-efficacy mechanism in human agency. American Psychologist 1982;37(2):122-147. [doi: 10.1037/0003-066X.37.2.122]
36. Braun V, Clarke V, Hayfield N, Davey L, Jenkinson E. Doing reflexive thematic analysis. In: Supporting Research in Counselling and Psychotherapy: Qualitative, Quantitative, and Mixed Methods Research: Springer; 2023:19-38. [doi: 10.1007/978-3-031-13942-0_2]
37. Ryu H, Kim S, Kim D, Han S, Lee K, Kang Y. Simple and steady interactions win the healthy mentality: designing a chatbot service for the elderly. 2020 Presented at: Proceedings of the ACM on human-computer interaction p. 1-25. [doi: 10.1145/3415223]
38. Schiavo G, Mich O, Ferron M, Mana N. Trade-offs in the design of multimodal interaction for older adults. Behav Inf Technol 2022 Apr 4;41(5):1035-1051. [doi: 10.1080/0144929X.2020.1851768]
39. He H, Raja Ghazilla RA, Abdul-Rashid SH. A systematic review of the usability of telemedicine interface design for older adults. Appl Sci (Basel) 2025;15(10):5458. [doi: 10.3390/app15105458]
40. Karkach A, Semigina T. Educational innovations in senior learning: the role of the university of the third age in addressing digital literacy gaps. SSRN. Preprint posted online on Apr 1, 2024. [doi: 10.2139/ssrn.4850037]
41. Chu CH, Nyrup R, Leslie K, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. Gerontologist 2022 Aug 12;62(7):947-955. [doi: 10.1093/geront/gnab167] [Medline: 35048111]
42. What explains the older users' continuance intention to use digital healthcare technologies in the UK. : Brunel University London; 2024 URL: http://bura.brunel.ac.uk/handle/2438/30673 [accessed 2026-02-04]
43. Azevedo C, Ponte C. Intergenerational solidarity or intergenerational gap? How elderly people experience ICT within their family context. OBS* 2020;14(3). [doi: 10.15847/obsOBS14320201587]
44. Xu J, Ma J. Urban-rural disparity in the relationship between geographic environment and the health of the elderly. Appl Spatial Analysis 2024 Sep;17(3):1335-1357. [doi: 10.1007/s12061-024-09586-7]
45. Khamaj A. AI-enhanced chatbot for improving healthcare usability and accessibility for older adults. Alexandria Eng J 2025 Mar;116:202-213. [doi: 10.1016/j.aej.2024.12.090]
46. Rahaman MM, Islam A. Automation and risk mitigation in healthcare claims: policy and compliance implications. RAST 2023;02(4):124-157 [FREE Full text] [doi: 10.63125/v73gyg14]
47. Liu Y, Shan Y, Zhang M, Ji B, Zhou Y. Digital divide in smart sports applications for the elderly and its mitigation. Qual Sport 2025;40:60013. [doi: 10.12775/QS.2025.40.60013]
48. Czaja SJ, Ceruso M. The promise of artificial intelligence in supporting an aging population. J Cogn Eng Decis Mak 2022 Dec;16(4):182-193. [doi: 10.1177/15553434221129914]
49. Zhang S, Wu L, Zhao Z, Massó JRF, Chen M. Artificial intelligence in gerontology: data-driven health management and precision medicine. Adv Gerontol 2024 Sep;14(3):97-110. [doi: 10.1134/S2079057024600691] [Medline: 28514535]
50. Abdollahi H, Mahoor MH, Zandie R, Siewierski J, Qualls SH. Artificial emotional intelligence in socially assistive robots for older adults: a pilot study. IEEE Trans Affect Comput 2023;14(3):2020-2032. [doi: 10.1109/taffc.2022.3143803] [Medline: 37840968]
51. Rao DV, Warburton J, Bartlett H. Health and social needs of older Australians from culturally and linguistically diverse backgrounds: issues and implications. Australas J Ageing 2006 Dec;25(4):174-179 [FREE Full text] [doi: 10.1111/j.1741-6612.2006.00181.x]
52. Gou F, Liu J, Xiao C, Wu J. Research on artificial-intelligence-assisted medicine: a survey on medical artificial intelligence. Diagnostics (Basel) 2024 Jul 9;14(14):1472. [doi: 10.3390/diagnostics14141472] [Medline: 39061610]
53. Aging well in a digital world: developing inclusive community-based technology programs for the mental wellbeing of older adults. : Simon Fraser University (SFU); 2024 URL: https://summit.sfu.ca/_flysystem/fedora/2024-09/etd23156.pdf [accessed 2026-02-04]
54. Mhlanga D. Artificial intelligence in elderly care: navigating ethical and responsible AI adoption for seniors. In: Fostering Long-Term Sustainable Development in Africa: Overcoming Poverty, Inequality, and Unemployment: Springer Nature; 2024:411-440. [doi: 10.1007/978-3-031-61321-0_18]
55. Finco MG, Mir N, Gresham G, Huisingh-Scheetz M. Ethical considerations of digital health technology in older adult care. Lancet Healthy Longev 2024 Jan;5(1):e12-e13. [doi: 10.1016/S2666-7568(23)00236-2] [Medline: 38183992]

XSL•FO
RenderX

56. Du Q, Wei X, Li J, et al. AI as a bridge across ages: exploring the opportunities of artificial intelligence in supporting inter-generational communication in virtual reality. Presented at: Proceedings of the ACM on Human-Computer Interaction; Oct 18-20, 2025; Bergen, Norway p. 1-29. [doi: 10.1145/3710924]
57. Alievska B. Ethical Challenges Arising from the Underrepresentation of the Elderly in the Development of Artificial Intelligence (AI) for Medical Applications: Exclusion, Bias, and the Limits of Accessibility. 2023.
58. Zhang H. AIoT for aging in place: from theory to practice. Rev Appl Sci Technol 2024 [FREE Full text]
59. Segkouli S, Giakoumis D, Votis K, Triantafyllidis A, Paliokas I, Tzovaras D. Smart Workplaces for older adults: coping "ethically" with technology pervasiveness. Univers Access Inf Soc 2023;22(1):37-49. [doi: 10.1007/s10209-021-00829-9] [Medline: 34305502]
60. Islam MR. Generative AI, Cybersecurity, and Ethics: John Wiley & Sons; 2024. URL: https://onlinelibrary.wiley.com/doi/book/10.1002/9781394279326?msockid=39c3084260a26d7d1e231c7061aa6cf4 [accessed 2026-02-14]

Abbreviations

AI: artificial intelligence
GenAI: generative artificial intelligence
TAM: Technology Acceptance Model

Edited by AF Abdulai; submitted 14.Jul.2025; peer-reviewed by H Zhang, K Malagas, K Mirkovski, M Yi; revised version received 26.Dec.2025; accepted 05.Jan.2026; published 09.Mar.2026.

Please cite as:
Liu T, Luo YT, He Z, Pang P, Chan KS, Lau Y
Potential Application Value and Needs of Using Generative AI Videos for Health Management for Older Adults: Qualitative Study
JMIR Aging 2026;9:e80661
URL: https://aging.jmir.org/2026/1/e80661
doi:10.2196/80661

© Ting Liu, Yiming Taclis Luo, Zhenni He, Patrick Pang, Kin Sun Chan, Ying Lau. Originally published in JMIR Aging (https://aging.jmir.org), 9.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Predicting Adherence to Computer-Based Cognitive Training Programs Among Older Adults Using Source-Free Domain Adaptation: Algorithm Development and Validation

Ronast Subedi[1], MS; Shayok Chakraborty[1], PhD; Zhe He[2,3], PhD; Yuanying Pang[2], MS; Shenghao Zhang[4], PhD; Mia Liza Lustria[2], PhD; Neil Charness[4], PhD; Walter Boot[4], PhD

[1]Department of Computer Science, Florida State University, Tallahassee, FL, United States

[2]School of Information, Florida State University, Tallahassee, FL, United States

[3]College of Medicine, Florida State University, Tallahassee, FL, United States

[4]Division of Geriatrics and Palliative Medicine, Weill Cornell Medicine, New York, NY, United States

Corresponding Author:
Shayok Chakraborty, PhD
Department of Computer Science
Florida State University
1017 Academic Way
Suite 253
Tallahassee, FL, 32304
United States
Phone: 1 850 645 8474
Email: schakraborty2@fsu.edu

Abstract

Background: Cognitive decline in the aging population presents an unprecedented challenge worldwide. Recent research has shown the potential of cognitive training programs to mitigate cognitive decline. However, these interventions require sustained adherence to be effective, which can be challenging.

Objective: In this study, we aim to enhance the accuracy of predicting adherence patterns in cognitive training programs for older adults, with the goal of developing personalized support systems that promote adherence and improve cognitive outcomes.

Methods: A major challenge in developing deep neural networks for predicting adherence patterns is the limited availability of individual participants' training data. Although domain adaptation techniques can address this issue by leveraging training data from other clinical studies, our research considers a more practical scenario where the use of such data from other studies is restricted due to privacy and confidentiality concerns. Therefore, we used source-free domain adaptation (SFDA), which uses models trained on other cognitive studies without requiring access to the corresponding datasets. To the best of our knowledge, this is the first effort to use SFDA to predict older adults' daily adherence to cognitive training programs.

Results: Using data from 3 previously conducted cognitive training intervention studies, our results demonstrated the efficacy of deep learning models combined with SFDA to accurately predict adherence lapses while addressing data privacy concerns.

Conclusions: Our findings indicate that deep learning and SFDA techniques can be useful in the development of adherence support systems for computerized cognitive training, aimed at improving the health and well-being of older adults.

(JMIR Aging 2026;9:e79123) doi:10.2196/79123

KEYWORDS

source-free domain adaptation; adherence prediction; cognitive training; data privacy; artificial intelligence; AI

Introduction

Background

Cognitive decline presents significant challenges for older people, impacting their well-being, decision-making, and overall quality of life. Although some cognitive decline is normal with age, severe cognitive impairments such as dementia and Alzheimer disease present greater risks [1]. Dementia currently ranks as the seventh leading cause of global deaths, and its prevalence could triple by 2050 [2]. The number of older adults

XSL•FO
RenderX

is increasing substantially, particularly among those aged 85 years and older, who represent the fastest-growing age category among older adults [3]. With the US population aged 65 years and older expected to nearly double in the next 40 years [4], addressing age-related cognitive changes is essential for promoting independence, improving quality of life, and reducing caregiving burdens on families and health care systems.

Cognitive training interventions have emerged as a promising approach to address cognitive decline. These interventions often involve technology-based training programs that target specific cognitive functions such as memory, attention, and problem-solving [5-8]. Recent studies have underscored the potential of digital cognitive training programs in improving cognitive performance and daily activities in older adults [9-11]. Lövdén et al [12] assert the following in their theoretical account of the necessary conditions for inducing cognitive plasticity: (1) mismatches between functional supply and experienced environmental demands drive manifestations of plasticity and (2) mismatches need to be maintained for extended periods to induce plasticity. Keeping people engaged and adherent for extended periods while performing tasks that force them to operate slightly above their cognitive limits is a major challenge. A necessary first step in promoting adherence is the ability to predict lapses and to find effective ways to intervene "just-in-time" to maintain engagement.

Recently, Li et al [13] conducted a meta-analysis that included 55 randomized controlled trials with 4455 participants (spanning normal cognition, mild cognitive impairment, and neurodegenerative dementia) to assess the effects of cognitive training in older adults. The study demonstrated that participants who maintained an adherence rate of 80% or higher experienced significant cognitive benefits. This systematic review further highlights the critical role of sustained adherence in augmenting the efficacy of cognitive training and underscores the importance of predicting adherence lapses. Studies also indicate that adherence to cognitive training is influenced by factors such as mental capacity, physical function, and program design [14-19]. Hence, understanding adherence to these interventions and developing strategies to enhance them are essential for improving the effectiveness and optimizing outcomes of computer-based cognitive training programs [20,21].

To this end, in our recent research, we have applied advanced machine learning (ML) and deep learning (DL) techniques to predict adherence lapses based on past behaviors among older adults participating in cognitive training programs [22,23]. Deep neural networks, together with techniques such as data augmentation, adaptive window size estimation, and temporal data clustering, have shown tremendous promise in predicting adherence patterns [24]. To capture the individual differences across participants, we trained a separate prediction model for each participant rather than a single model for all participants. However, the limited availability of data for each participant (only 30 days) made it difficult to train a reliable prediction model for each participant. In our recent work [25], we leveraged the potential of domain adaptation (DA) to address this challenge by using training data from participants in a similar, previously conducted clinical trial. Although the results were promising, this method required access to data from other participants in a different trial, which can raise privacy concerns, as health care data can contain personal or confidential information.

In this study, we used data from 3 technology-based cognitive intervention studies to predict adherence patterns among participants while also addressing data privacy issues. Although these datasets are deidentified, previous publications have shown that deidentification of health data, commonly used to protect privacy, is ineffective because modern techniques can easily reidentify individuals by linking deidentified datasets with external information [26,27]. Similarly, ML techniques can successfully reidentify individuals from incomplete datasets with high likelihood [28]. Despite these limitations, deidentified data are widely shared and sold without patient consent, highlighting the need for stronger protection and privacy practices in health care research.

In this work, we used state-of-the-art source-free domain adaptation (SFDA) techniques to address the challenge of limited training data for each participant and the privacy concerns surrounding clinical data. Our framework requires access only to a *deep neural network model trained on data from a different clinical study, and not the confidential data themselves*, thereby addressing data privacy concerns. Our extensive empirical analyses demonstrated that SFDA can efficiently leverage knowledge from other clinical studies to develop participant-specific deep neural network models to predict adherence patterns, while addressing data privacy issues at the same time.

Overall, our research highlights the potential of deep neural networks and SFDA techniques in predicting adherence lapses and developing personalized support systems for cognitive training programs. SFDA not only improves the prediction accuracy but also preserves the privacy of clinical datasets, which is crucial for confidentiality. By leveraging these technologies, we can enhance engagement and optimize the benefits of cognitive interventions, ultimately improving the quality of life for older people with cognitive impairments. To the best of our knowledge, this study represents the first effort to apply SFDA techniques to predict daily adherence among older adults participating in cognitive training programs.

Objectives

This research sought to address the following objectives:

1. Determine the potential of convolutional neural networks (CNNs) to predict adherence to cognitive training programs based on past behavior
2. Explore the use of SFDA techniques to address training data scarcity and maintain data privacy while predicting adherence patterns
3. Investigate how artificial intelligence (AI)–driven approaches can be implemented to predict engagement in cognitive training programs

Overall, by identifying adherence patterns, these approaches can inform the development of more effective interventions, benefiting individuals and society by improving conditions associated with aging.

XSL•FO
RenderX

Methods

Research Design

For this study, we investigated older adults' engagement with Mind Frontiers (Aptima Inc), an Android-based cognitive training game platform. The platform features a Wild West–themed interface with 7 interactive minigames designed to enhance key cognitive functions, including memory, attention, spatial navigation, multitasking, and logical reasoning. Cognitive training was conducted on Lenovo 10 tablets, with all participants receiving instructions on device operation and gameplay. The levels reached in the games ranged from 1 to 58, and the cognitive training program included 5 possible outcomes: *defeat*, *stalemate*, *victory*, *abort*, and *not yet finished*. After completing each game, participants received feedback, and the game's difficulty was adjusted based on their previous performance. This structured approach ensured consistent engagement and systematic data collection for assessing cognitive improvement over time.

Datasets

In this study, we used 3 distinct datasets, each derived from separate clinical research trials. The study 1 dataset [29] consisted of a 2-phase protocol. In the first phase of 12 weeks, participants were asked to follow a structured schedule, engaging in cognitive training for 45 minutes per day, 5 days per week. In contrast, the subsequent 6-week phase was unstructured, allowing participants to engage at their discretion. In our study, we analyzed data only from the structured phase, as adherence metrics rely on the prescribed game-playing schedule. The dataset is derived from 118 participants with a mean age of 72.6 (SD 5.5) years, including 78 (66.2%) female participants and 40 (33.8%) male participants. The dataset included multiple cognitive assessments, such as technical proficiency, self-efficacy, subjective cognition, perceived benefits, and objective cognitive measures such as processing speed and memory.

The study 2 dataset [30] came from a similar but independent cognitive training study that involved 120 adults aged 64 years and older from Leon County, Florida, and surrounding areas. Of these, 116 (96.7%) participants completed the structured phase, which included baseline and postintervention cognitive measures. The average age of participants was 70 (SD 4.42) years, with a range of 64 to 84 years; 77 (64.2%) participants were female, while 43 (35.8%) participants were male. Although participants were expected to complete both phases, the second phase did not require participants to engage in gaming interventions for specific time intervals.

The study 3 dataset also came from an independent cognitive training trial consisting of 92 participants. Among these, 90 (97.8%) participants provided complete demographic information, with ages ranging from 62 to 84 years and an average age of approximately 71 (SD 4.97) years. Of the participants, 67 (74.4%) were female and 20 (22.2%) were male, while 2 (2.2%) participants preferred not to disclose their sex, and 1 (1.1%) participant indicated that none of the listed sex options described them.

DL for Adherence Prediction

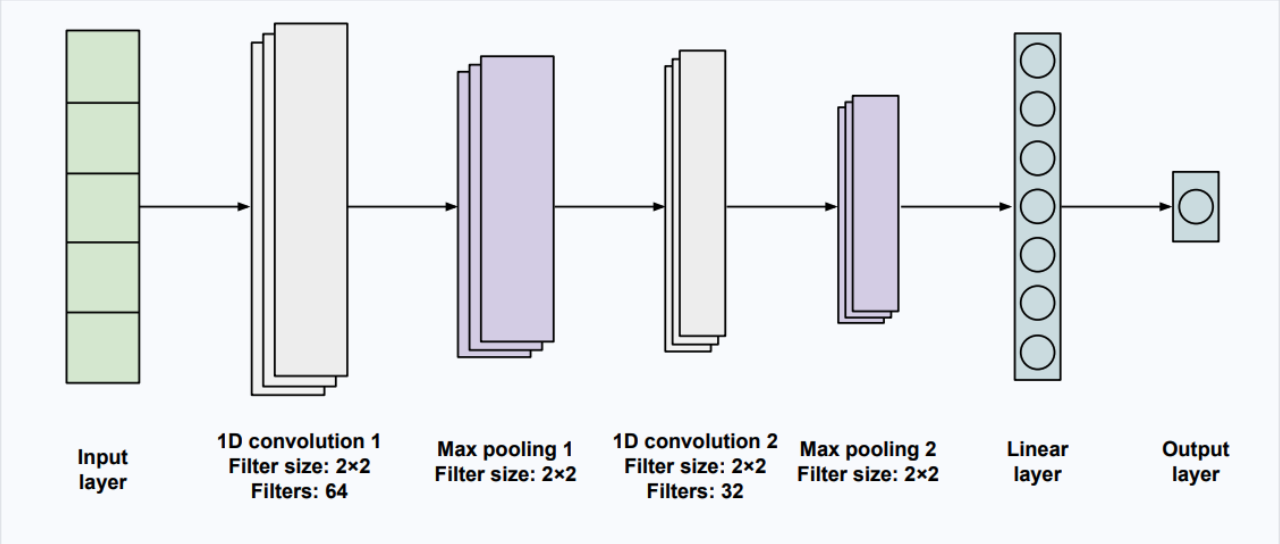
In this study, we used deep CNNs [31], which were initially designed for image data but are also widely used for time-series classification and forecasting tasks [32,33]. CNNs consist of convolutional layers comprising filters that slide over the input data to extract local patterns and features. For 1D time-series data, 1D convolution is applied, in which the kernel slides unidirectionally from the start of the time series to its end. For an input vector f of length n and a kernel g of length m, the convolution ($f \times g$) of f and g is defined as follows:



In this study, we used a DL architecture consisting of 2 convolutional blocks, each comprising a convolutional layer followed by a maximum pooling layer. The pooling layer down-samples the feature maps by selecting the maximum value within a predefined window, thus effectively retaining salient features. These blocks are succeeded by a linear layer and an output layer, as depicted in Figure 1.

XSL•FO
RenderX

Figure 1. Network architecture used in our experiments. It consists of 2 convolutional blocks, each comprising a convolutional layer followed by a max pooling layer. These blocks are succeeded by a linear layer and an output layer.



SFDA for Adherence Prediction

A key challenge in our study was the limited availability of labeled training data for each participant (only 30 days of training data). DA is a well-researched technique in the ML literature for addressing the challenge of training models with limited labeled data. DA algorithms use labeled data from a source domain to enhance the performance of DL models in a target domain of interest with limited labeled data, under the constraint that the data from the 2 domains are derived from different probability distributions [34,35]. Although DA is predominantly used in computer vision, recent studies have demonstrated its applicability in adapting temporal data patterns [36,37].

In our recent research [25], we studied the performance of adversarial DA [38] to enhance the accuracy of personalized adherence prediction models by using training data from participants in a different, previously conducted clinical study. For a given target participant, all participants from the other study, with similar play patterns (determined through a clustering algorithm), were considered source participants. Each participant has a unique play pattern; that is, the data for each participant are derived from a different probability distribution. Adversarial DA was used to address this disparity between the source and target domains, so that the data from the source participants can also be used to train a personalized model for the target participant. Our results demonstrated that DA consistently produced more accurate personalized adherence prediction models for each participant, compared with the case when only the data from the target participant was used to train the model for this participant, without using the data from the other clinical study.

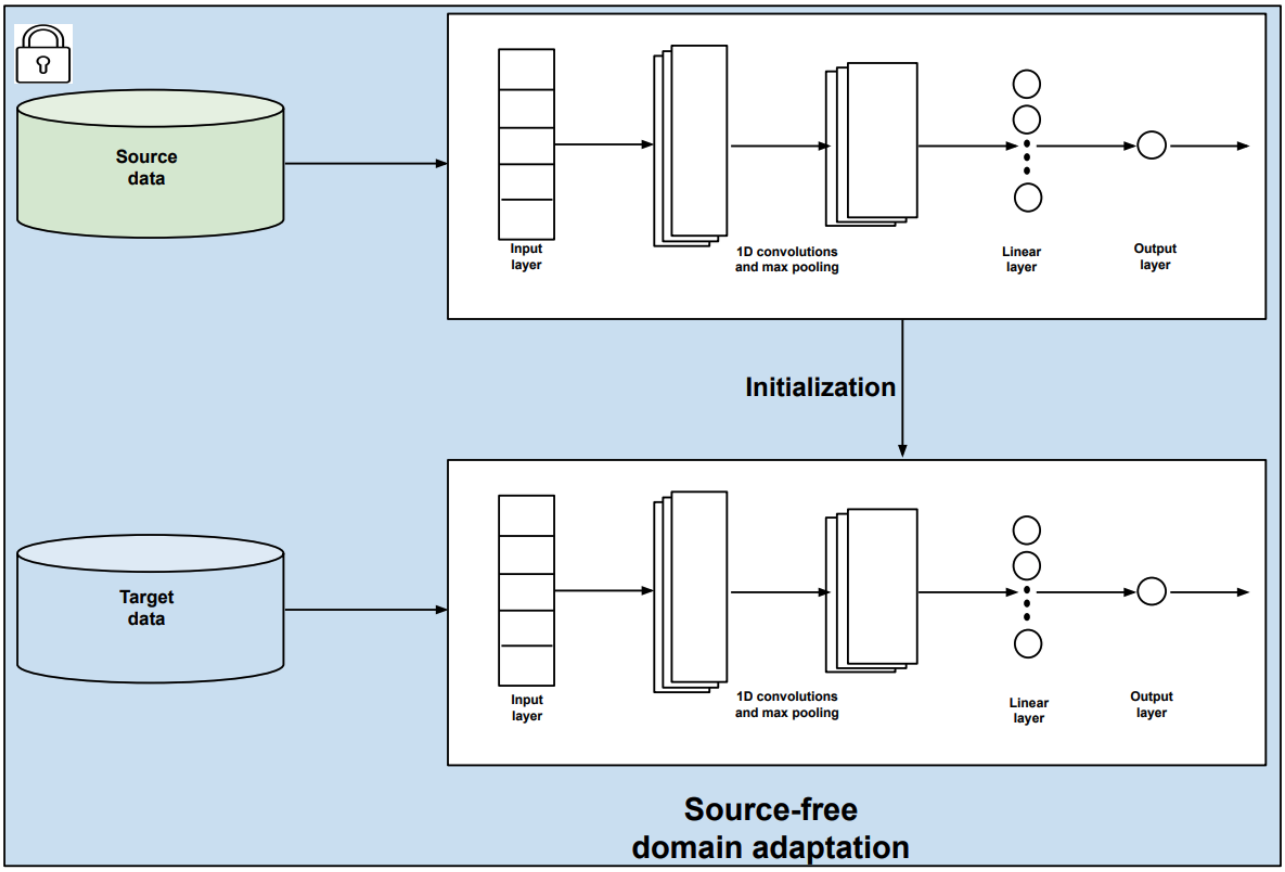
Although these results were promising, the DA algorithm required access to the source participants' data from the other clinical trial. This can pose privacy concerns as clinical data typically contain personal or confidential information. In this paper, we considered a more practical scenario, assuming we would only have access to a *deep neural network model trained on the source participants' data, rather than the raw data*. We attempted to leverage the source domain knowledge encoded in the trained source model to train an accurate adherence prediction model for a given target participant. This category of DA is referred to as SFDA [39-45].

Although much SFDA research focuses on computer vision, recent studies have extended SFDA approaches to medical and time-series domains. For instance, Ragab et al [39] introduced a temporal imputation strategy tailored for time-series SFDA, demonstrating improved performance on physiological datasets such as human activity recognition and sleep stage classification. Similarly, studies by Li et al [42] and Wu et al [43] applied SFDA techniques for medical image segmentation, including applications to fundus imaging. Moreover, Yang et al [44] explored the use of SFDA with Fourier-style feature mining for tasks such as polyp and prostate segmentation.

SFDA is effective in addressing data privacy concerns, as only a model trained on source domain data is used for knowledge transfer, rather than the raw source domain data. This is depicted in Figure 2. Our framework thus offers a safer mechanism for clinical trials to inform one another. In this work, we demonstrate the utility of our method using the 3 datasets mentioned as test cases. Our framework will be especially useful for datasets that contain even more sensitive information, given the ease of reidentifying the deidentified data as discussed previously.

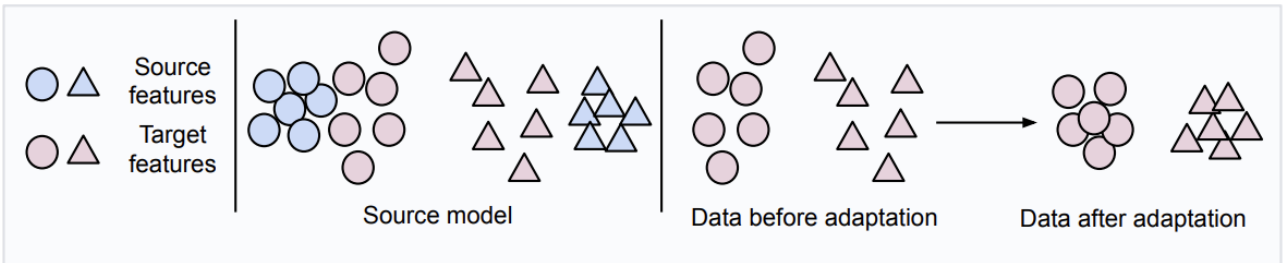
XSL•FO
RenderX

Figure 2. Illustration of source-free domain adaptation (SFDA). We only have access to a deep neural network model trained on the source participants' data, rather than the raw data. SFDA is effective in addressing data privacy concerns, as only a model trained on source-domain data is used for knowledge transfer to the target-domain participants, rather than the raw source-domain data.



We used in our study the SFDA method proposed by Yang et al [40]. This method performs adaptation to the target domain without access to source domain data by using local structure clustering (LSC). The main rationale of this method is that although the features of the target domain can be shifted with respect to the features of the source domain, the classes will still form clusters in the feature space. Thus, LSC attempts to move clusters of data points toward their most likely class prediction. Before adaptation, some target features given by the source model are expected to deviate from the corresponding dense source feature region due to domain shift. These samples will most likely be incorrectly predicted by the classifier. However, we assumed that the target features from the same class would be clustered together. Therefore, the nearest neighbors of target features would have a high probability of having the same class labels. To exploit this fact, LSC encourages samples that are close in the feature space to have similar predictions to those of their nearest neighbors. Consequently, clusters of points that are close in feature space will move jointly toward a common class. Thus, this process can correctly classify target features that would otherwise have been wrongly classified. Additional information about this method can be found in the study by Yang et al [40]. Figure 3 presents a visual illustration of the LSC method. To further validate the effectiveness of SFDA, we also conducted an additional experiment using the SFDA method proposed by Yang et al [45] in our setting.

Figure 3. Illustration of local structure clustering (LSC) for SFDA (as described in the studies by Yang et al [40,45]). Some target features from the source model will deviate from dense source feature regions due to domain shift. LSC aims to cluster target features based on their semantically close neighbors.



XSL•FO
RenderX

Experimental Setup

For all the experiments, we used the network architecture illustrated in Figure 1. The network was trained for up to 130 epochs, with early stopping applied when the validation loss ceased to decrease. We used the Adam optimizer with an initial learning rate of 1×10^{-3}, which is reduced by a factor of 0.5 every 20 epochs. All models are trained on an Ubuntu 20.04 system equipped with a 20 GB NVIDIA RTX A4500 graphics processing unit. The training of a single model takes approximately 45 minutes. We set aside 15% of the training data for validation and monitored the validation loss to ensure proper model convergence.

In our experiments, we divided each participant's data into 2 segments for training and testing. The first 30 days of data were considered for training, and the subsequent 30 days of data were used for testing. This division aligns with psychological research indicating that approximately 30 days is the median time required for habit formation to plateau [46].

The input to the CNN consisted of multivariate time-series data, including (1) play duration, (2) session count, (3) maximum level achieved, and (4) number of tasks completed. We established a threshold of 10 minutes of playing time per day to classify participants as adherent (*playing time ≥10 minutes*) or nonadherent, consistent with our previous work [22,24]. Our objective was to train deep neural networks to predict adherence on the (N + 1)th day based on data from the previous N days. We set the window size, N, to 7 in this study. A personalized model was trained for all participants to capture their unique playing patterns. We used accuracy, precision, recall, and F_1-scores as the metrics for evaluation. Precision, recall, and F_1-scores were used to assess the model's performance on positive predictions, whereas accuracy evaluated overall performance. The reported final scores represented the average performances across all participants. We conducted 2 different sets of experiments and performed a 1-tailed t test to assess the statistical significance of the obtained results. The sets of experiments were as follows: first, no source; in this setting, an individual DL model was trained for each target participant using only the training data of that participant. No information from any other dataset was used for model training. Second, SFDA; here, the SFDA method [40] was used to train personalized DL models for each target participant by transferring relevant knowledge from a source model trained on a source dataset.

For instance, when study 1 served as the source dataset, we first trained a deep model on all the participants' data from study 1. If the target participant belonged to study 2, the deep model was fine-tuned using data from all other participants in study 2, excluding the target participant. Subsequently, only the trained model, without direct access to the original study 1 dataset, was used to transfer relevant knowledge for developing a personalized adherence prediction model for the target participant.

Ethical Considerations

This study was approved by the Florida State University Institutional Review Board (study ID 2017.20622 and STUDY00000005). Informed consent was obtained from all participants, and participants were given the opportunity to opt out at any time. This consent process included institutional review board–approved language stating that the collected data could be made available for secondary analysis by other research teams and that any shared data would not include personal identifiers. Data were labeled with participant IDs, meaning that the data were deidentified in this study. Participants were compensated a total of US $200 (study 1) or US $75 (study 2) for their participation, with the difference driven by the shorter commitment involved in study 2.

Results

The results of our study are presented in Tables 1-3 for the study 1-3 datasets, respectively (using the SFDA method proposed in the study by Yang et al [40]). Each table reports the values of the evaluation metrics on the test sets of the respective datasets. The values reported are the average scores across all participants. The tables show that performance across all the metrics was lowest when no DA technique was applied.

Table 1. Average scores of source-free domain adaptation (SFDA) models on the study 1 dataset[a].

Experiment type	Accuracy (%), mean (SD)	Precision (%), mean (SD)	Recall (%), mean (SD)	F_1-score (%), mean (SD)
No source	63 (0.24)	58 (0.23)	57 (0.24)	54 (0.25)
SFDA with the study 2 dataset as the source	69 (0.20)	63 (0.20)	61 (0.19)	60 (0.21)
SFDA with the study 3 dataset as the source	65 (0.23)	58 (0.23)	56 (0.22)	56 (0.20)

[a]The scores are calculated by averaging the respective scores across all the participants.

XSL•FO
RenderX

Table 2. Average scores of source-free domain adaptation (SFDA) models on the study 2 dataset[a].

Experiment type	Accuracy (%), mean (SD)	Precision (%), mean (SD)	Recall (%), mean (SD)	F_1-score (%), mean (SD)
No source	57 (0.22)	54 (0.19)	52 (0.20)	49 (0.20)
SFDA with the study 1 dataset as the source	68 (0.20)	57 (0.21)	58 (0.19)	56 (0.21)
SFDA with the study 3 dataset as the source	64 (0.21)	55 (0.23)	57 (0.20)	54 (0.20)

[a]The scores are calculated by averaging the respective scores across all the participants.

Table 3. Average scores of source-free domain adaptation (SFDA) on the study 3 dataset[a].

Experiment type	Accuracy (%), mean (SD)	Precision (%), mean (SD)	Recall (%), mean (SD)	F_1-score (%), mean (SD)
No source	56 (0.20)	53 (0.19)	51 (0.22)	50 (0.23)
SFDA with the study 2 dataset as the source	62 (0.20)	61 (0.22)	58 (0.18)	56 (0.21)
SFDA with the study 1 dataset as the source	65 (0.21)	62 (0.25)	59 (0.19)	58 (0.23)

[a]The scores are calculated by averaging the respective scores across all the participants.

For instance, in the study 2 dataset (Table 2), the average accuracy was 57% when only the data for a given target participant was used to train a model for that participant, without access to any other source dataset. This shows that using only the limited participant-specific training data may result in prediction models with poor generalization capability. Using information from the other datasets through SFDA consistently produced better performance in terms of all the evaluation metrics.

For the study 2 dataset, using the study 1 dataset as the source yielded an accuracy of 68% (11% improvement), whereas using the study 3 dataset as the source produced an accuracy of 64% (7% improvement). The corresponding improvements on the F_1-score were 7% and 5%, respectively. A similar pattern was observed for the other 2 datasets.

For the study 1 dataset, using the other 2 datasets as the source datasets resulted in an improvement of 6% and 2% in accuracy and 6% and 2% in the F_1-scores, respectively. For the study 3 dataset, the improvements in accuracy and F_1-score were 6% and 9%, and 6% and 8%, respectively, when using the other 2 datasets as a source, compared with the "No source" setting.

Similarly, using the SFDA technique proposed in the study by Yang et al [45], the improvements in accuracy and F_1-score were 3% and 4%, and 6% and 7%, respectively, compared with the "No source" setting for the study 3 dataset, as depicted in Table 4. The greater performance gains observed when using the study 1 or study 2 datasets as the source dataset, compared with the study 3 dataset, are attributed to the larger sample size of the study 1 (118 participants) and study 2 (116 participants) datasets, compared with the study 3 dataset (92 participants).

Table 4. Average scores of source-free domain adaptation (SFDA) on the study 3 dataset (calculated by the method proposed in the study by Yang et al [45])[a].

Experiment type	Accuracy (%), mean (SD)	Precision (%), mean (SD)	Recall (%), mean (SD)	F_1-score (%), mean (SD)
No source	56 (0.20)	53 (0.19)	51 (0.22)	50 (0.23)
SFDA with the study 1 dataset as the source	59 (0.22)	56 (0.23)	55 (0.19)	56 (0.20)
SFDA with the study 2 dataset as the source	60 (0.23)	58 (0.24)	57 (0.20)	57 (0.25)

[a]The scores are calculated by averaging the respective scores across all the participants.

These results unanimously corroborate the potential of SFDA algorithms to not only develop more accurate personalized DL models for predicting adherence to cognitive training programs among older adults but also help address privacy and confidentiality concerns that are inherent in clinical datasets.

Additionally, we conducted statistical tests of significance using *t* tests (Table 5) and the Wilcoxon signed-rank test (Table 6) to assess whether the improvement in performance achieved by our SFDA method is statistically significant when compared with the *"no source"* method. We computed the *P* values on accuracy scores for all the experiments detailed in Tables 1-3. Tables 5 and 6 show that all *P* values were ≤.05 for all tests, signifying that the performance improvement achieved by the SFDA method over the *"no source"* method is statistically significant. These results further support the effectiveness of SFDA in improving the performance of DL models for predicting adherence to cognitive training programs among older adults while maintaining data privacy and confidentiality.

XSL•FO
RenderX

Table 5. *P* values for the accuracy scores obtained by performing a t test comparing the results of the source-free domain adaptation (SFDA) method with the "no source" method.

Datasets	*P* value
Study 1 dataset	
Source: study 2 dataset	.006
Source: study 3 dataset	.05
Study 2 dataset	
Source: study 1 dataset	.03
Source: study 3 dataset	.03
Study 3 dataset	
Source: study 1 dataset	.04
Source: study 2 dataset	.009

Table 6. *P* values for the accuracy scores obtained by performing the Wilcoxon signed-rank test comparing the results of the source-free domain adaptation (SFDA) method with the "no source" method.

Dataset	*P* value
Study 1 dataset	
Source: study 2 dataset	.005
Source: study 3 dataset	.03
Study 2 dataset	
Source: study 1 dataset	.02
Source: study 3 dataset	.01
Study 3 dataset	
Source: study 1 dataset	.02
Source: study 2 dataset	.008

Discussion

Principal Results

The findings of this study have several significant implications for adherence prediction in cognitive training programs for older adults. Our results demonstrate that CNNs can analyze behavioral patterns and predict adherence trends in cognitive training interventions. CNNs learn meaningful patterns and trends from past data and predict future adherence, which can be used to design more personalized and effective cognitive training programs. These findings contribute to the growing body of research supporting AI-driven approaches for optimizing health care interventions in the aging population.

Our work demonstrates that SFDA techniques can effectively leverage information from external datasets and substantially improve the predictive accuracy of DL models, particularly when limited training data are available for developing personalized models for individual participants. More importantly, these methods offer the critical advantage of preserving data privacy, as they only require access to a model trained on a different clinical dataset, without requiring access to the raw data. Given strict privacy concerns in clinical applications and the ability of modern ML techniques to reidentify the deidentified data, training effective predictive models without direct access to the source data is a significant advancement. To the best of our knowledge, this study is the first to apply SFDA techniques to the domain of adherence prediction in cognitive training programs. The ability to maintain data privacy while still achieving competitive predictive accuracy has significant potential benefits in health care settings.

Furthermore, our study's findings have important practical and theoretical implications for the design of future cognitive training interventions. By enabling accurate day-level adherence prediction from limited individual data, SFDA provides the foundation for just-in-time adaptive interventions that can detect early signs of disengagement. Early detection helps to deliver tailored support such as personalized reminders, motivational prompts, difficulty adjustments, and behavioral coaching. This predictive capability is critical for sustaining the prolonged engagement required for cognitive plasticity, as described in theoretical models of adult learning [12]. The results also indicate that adherence behaviors exhibit stable, transferable structural patterns, revealing a promising direction for using DA theory to model behavioral time-series and examine habit formation and individual differences in engagement. As SFDA enables cross-cohort knowledge transfer without accessing sensitive data, it supports the development of scalable, privacy-preserving personalized adherence support systems that can be deployed across sites and studies. From a practical

XSL•FO
RenderX

standpoint, improving adherence prediction has direct implications for clinical trials, where consistent participant engagement is essential for accurately evaluating the effectiveness of cognitive training interventions. By improving the predictive accuracy of CNNs, SFDA-enhanced systems could make cognitive training trials more reliable, efficient, and informative.

Limitations

Despite its advantages, applying SFDA in our study presents certain limitations. Deploying personalized models in real-world cognitive training applications raises several ethical and practical concerns, as users' experiences with AI-health systems depend heavily on trust, privacy, and output interpretability [47]. Although SFDA ensures data privacy and personalized models enhance user engagement, it remains crucial to address issues related to user trust, transparency, and interpretability.

In addition to SFDA, other privacy-preserving techniques such as federated learning (FL) and homomorphic encryption exist. However, our focus on SFDA is motivated by its practicality in small-scale clinical studies. FL requires distributed infrastructure, multiple rounds of communication, and active collaboration among participants, increasing both computational burden and vulnerability to privacy attacks [48,49]. In contrast, SFDA involves sharing only a single source model without accessing source data or maintaining continuous client interactions. Hence, it is more suitable for our setting, in which trials are completed, and no live infrastructure exists for FL deployment.

Furthermore, there is a risk of domain mismatch when transferring personalized models to new cognitive games or different user populations. Models trained on limited individual-level data may overfit to short-term noise, reducing reliability. Moreover, tailoring models using only 30 days of data may introduce bias and restrict external generalizability, as early playing patterns may not fully capture long-term patterns. Finally, all 3 datasets used in our study were derived from the same cognitive training platform (Mind Frontiers) and involve participants with similar demographic and geographic characteristics, which may further limit generalizability. In future work, we will focus on evaluating SFDA across different cognitive training platforms, more diverse populations, and varied intervention types to assess the robustness and broader applicability of our approach.

Conclusions

In conclusion, our research contributes to the growing body of research supporting AI-driven approaches to optimizing health care interventions for the aging population. Specifically, the findings support the use of DL and SFDA techniques for predicting adherence to cognitive training programs. With accurate adherence prediction models, AI-assisted personalized support systems can be developed to improve long-term engagement in cognitive training programs and maximize cognitive health benefits for older adults. We hope that our research will motivate further exploration of DA, particularly SFDA techniques in other contexts and populations, with the ultimate goal of improving the effectiveness of cognitive training interventions.

Funding

This work was supported by the National Institutes on Aging (R01AG064529; The Adherence Promotion with Person-centered Technology Project: Promoting Adherence to Enhance the Early Detection and Treatment of Cognitive Decline), and the University of Florida—-Florida State University Clinical and Translational Science Award, which is supported in part by the National Center for Advancing Translational Sciences of National Institutes of Health (UM1TR005128).

Data Availability

Data and analysis scripts will be made available upon request by contacting the corresponding author.

Authors' Contributions

Conceptualization: RS, SC, ZH, YP, NC, WB
Formal analysis: RS
Funding acquisition: WB
Methodology: RS, SC, NC, WB
Project administration: NC, WB
Software: RS
Supervision: SC
Validation: RS
Writing—original draft: RS
Writing—review and editing: SC, ZH, YP, SZ, ML, NC, WB
All authors contributed to the article and approved the submitted version.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Dementia. World Health Organization. 2025 Mar 31. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2026-02-22]
2. Fong TG, Inouye SK. The inter-relationship between delirium and dementia: the importance of delirium prevention. Nat Rev Neurol 2022 Oct 26;18(10):579-596 [FREE Full text] [doi: 10.1038/s41582-022-00698-7] [Medline: 36028563]
3. He W, Goodkind D, Kowal P. An aging world: 2015. United States Census Bureau. 2016 Mar. URL: https://www.census.gov/content/dam/Census/library/publications/2016/demo/p95-16-1.pdf [accessed 2026-02-22]
4. Mather M, Jacobsen LA, Pollard KM. Aging in the United States. Popul Bull 2015;70(2) [FREE Full text]
5. Ge S, Zhu Z, Wu B, McConnell ES. Technology-based cognitive training and rehabilitation interventions for individuals with mild cognitive impairment: a systematic review. BMC Geriatr 2018 Sep 15;18(1):213 [FREE Full text] [doi: 10.1186/s12877-018-0893-1] [Medline: 30219036]
6. Bahar-Fuchs A, Clare L, Woods B. Cognitive training and cognitive rehabilitation for persons with mild to moderate dementia of the Alzheimer's or vascular type: a review. Alzheimers Res Ther 2013;5(4):35 [FREE Full text] [doi: 10.1186/alzrt189] [Medline: 23924584]
7. Simons DJ, Boot WR, Charness N, Gathercole SE, Chabris CF, Hambrick DZ, et al. Do "brain-training" programs work? Psychol Sci Public Interest 2016 Oct;17(3):103-186. [doi: 10.1177/1529100616661983] [Medline: 27697851]
8. Chiu HL, Chu H, Tsai JC, Liu D, Chen YR, Yang HL, et al. The effect of cognitive-based training for the healthy older people: a meta-analysis of randomized controlled trials. PLoS One 2017;12(5):e0176742 [FREE Full text] [doi: 10.1371/journal.pone.0176742] [Medline: 28459873]
9. Chae HJ, Lee SH. Effectiveness of online-based cognitive intervention in community-dwelling older adults with cognitive dysfunction: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2023 Jan;38(1):e5853 [FREE Full text] [doi: 10.1002/gps.5853] [Medline: 36468299]
10. Tsai YI, Beh J, Ganderton C, Pranata A. Digital interventions for healthy ageing and cognitive health in older adults: a systematic review of mixed method studies and meta-analysis. BMC Geriatr 2024 Mar 04;24(1):217 [FREE Full text] [doi: 10.1186/s12877-023-04617-3] [Medline: 38438870]
11. Velloso V, Latgé-Tovar S, Bomilcar I, Mograbi DC. Cognitive interventions for healthy older adults: a systematic meta-review. Int J Clin Health Psychol 2025 Jan;25(1):100538 [FREE Full text] [doi: 10.1016/j.ijchp.2024.100538] [Medline: 39877892]
12. Lövdén M, Bäckman L, Lindenberger U, Schaefer S, Schmiedek F. A theoretical framework for the study of adult cognitive plasticity. Psychol Bull 2010 Jul;136(4):659-676. [doi: 10.1037/a0020080] [Medline: 20565172]
13. Li Z, He H, Chen Y, Guan Q. Effects of engagement, persistence and adherence on cognitive training outcomes in older adults with and without cognitive impairment: a systematic review and meta-analysis of randomised controlled trials. Age Ageing 2024 Jan 02;53(1):afad247. [doi: 10.1093/ageing/afad247] [Medline: 38266127]
14. Adherence to long-term therapies: evidence for action. World Health Organization. 2003. URL: https://iris.who.int/items/bf8058c0-03b2-4b47-838f-5534849927fb [accessed 2026-02-22]
15. Munro S, Lewin S, Swart T, Volmink J. A review of health behaviour theories: how useful are these for developing interventions to promote long-term medication adherence for TB and HIV/AIDS? BMC Public Health 2007 Jun 11;7(1):104 [FREE Full text] [doi: 10.1186/1471-2458-7-104] [Medline: 17561997]
16. Picorelli AM, Pereira LS, Pereira DS, Felício D, Sherrington C. Adherence to exercise programs for older people is influenced by program characteristics and personal factors: a systematic review. J Physiother 2014 Sep;60(3):151-156 [FREE Full text] [doi: 10.1016/j.jphys.2014.06.012] [Medline: 25092418]
17. Room J, Hannink E, Dawes H, Barker K. What interventions are used to improve exercise adherence in older people and what behavioural techniques are they based on? A systematic review. BMJ Open 2017 Dec 14;7(12):e019221 [FREE Full text] [doi: 10.1136/bmjopen-2017-019221] [Medline: 29247111]
18. Turunen M, Hokkanen L, Bäckman L, Stigsdotter-Neely A, Hänninen T, Paajanen T, et al. Computer-based cognitive training for older adults: determinants of adherence. PLoS One 2019 Jul 10;14(7):e0219541 [FREE Full text] [doi: 10.1371/journal.pone.0219541] [Medline: 31291337]
19. Rivera-Torres S, Fahey TD, Rivera MA. Adherence to exercise programs in older adults: informative report. Gerontol Geriatr Med 2019 Jan 22;5:2333721418823604 [FREE Full text] [doi: 10.1177/2333721418823604] [Medline: 30733977]
20. Tullo D, Feng Y, Pahor A, Cote JM, Seitz AR, Jaeggi SM. Investigating the role of individual differences in adherence to cognitive training. J Cogn 2023 Aug 22;6(1):48 [FREE Full text] [doi: 10.5334/joc.315] [Medline: 37636013]
21. Guimarães AL, Lin FV, Panizzutti R, Turnbull A. Effective engagement in computerized cognitive training for older adults. Ageing Res Rev 2025 Feb;104:102650. [doi: 10.1016/j.arr.2024.102650] [Medline: 39755175]
22. He Z, Tian S, Singh A, Chakraborty S, Zhang S, Lustria ML, et al. A machine-learning based approach for predicting older adults' adherence to technology-based cognitive training. Inf Process Manag 2022 Sep;59(5):103034 [FREE Full text] [doi: 10.1016/j.ipm.2022.103034] [Medline: 35909793]
23. Pang Y, Singh A, Chakraborty S, Charness N, Boot WR, He Z. Predicting adherence to gamified cognitive training using early phase game performance data: towards a just-in-time adherence promotion strategy. PLoS One 2024 Oct 2;19(10):e0311279 [FREE Full text] [doi: 10.1371/journal.pone.0311279] [Medline: 39356682]

XSL•FO
RenderX

24. Singh A, Chakraborty S, He Z, Tian S, Zhang S, Lustria ML, et al. Deep learning-based predictions of older adults' adherence to cognitive training to support training efficacy. Front Psychol 2022 Nov 17;13:980778 [FREE Full text] [doi: 10.3389/fpsyg.2022.980778] [Medline: 36467206]
25. Singh A, Chakraborty S, He Z, Pang Y, Zhang S, Subedi R, et al. Predicting adherence to computer-based cognitive training programs among older adults: study of domain adaptation and deep learning. JMIR Aging 2024 Sep 16;7:e53793 [FREE Full text] [doi: 10.2196/53793] [Medline: 39283346]
26. Benitez K, Malin B. Evaluating re-identification risks with respect to the HIPAA privacy rule. J Am Med Inform Assoc 2010;17(2):169-177 [FREE Full text] [doi: 10.1136/jamia.2009.000026] [Medline: 20190059]
27. Rothstein MA. Is deidentification sufficient to protect health privacy in research? Am J Bioeth 2010 Sep 09;10(9):3-11 [FREE Full text] [doi: 10.1080/15265161.2010.494215] [Medline: 20818545]
28. Rocher L, Hendrickx JM, de Montjoye YA. Estimating the success of re-identifications in incomplete datasets using generative models. Nat Commun 2019 Jul 23;10(1):3069 [FREE Full text] [doi: 10.1038/s41467-019-10933-3] [Medline: 31337762]
29. Harrell ER, Roque NA, Boot WR, Charness N. Investigating message framing to improve adherence to technology-based cognitive interventions. Psychol Aging 2021 Dec;36(8):974-982 [FREE Full text] [doi: 10.1037/pag0000629] [Medline: 34460281]
30. Harrell ER, Roque NA, Boot WR. Comparing the effectiveness of two theory-based strategies to promote cognitive training adherence. J Exp Psychol Appl 2023 Dec;29(4):782-792 [FREE Full text] [doi: 10.1037/xap0000485] [Medline: 37471032]
31. O'shea K, Nash R. An introduction to convolutional neural networks. ArXiv Preprint posted online on November 26, 2015 [FREE Full text]
32. Liu S, Zhang C, Ma J. CNN-LSTM neural network model for quantitative strategy analysis in stock markets. In: Proceedings of the 24th International Conference on Neural Information Processing. 2017 Presented at: ICONIP 2017; November 14-18, 2017; Guangzhou, China. [doi: 10.1007/978-3-319-70096-0_21]
33. Shah J, Vaidya D, Shah M. A comprehensive review on multiple hybrid deep learning approaches for stock prediction. Intell Syst Appl 2022 Nov;16:200111. [doi: 10.1016/j.iswa.2022.200111]
34. Farahani A, Voghoei S, Rasheed K, Arabnia H. A brief review of domain adaptation. In: Stahlbock R, Weiss GM, Abou-Nasr M, Yang CY, Arabnia HR, Deligiannidis L, editors. Advances in Data Science and Information Engineering. Cham, Switzerland: Springer; 2021:877-894.
35. Wang M, Deng W. Deep visual domain adaptation: a survey. Neurocomputing 2018 Oct;312:135-153. [doi: 10.1016/j.neucom.2018.05.083]
36. He H, Queen O, Koker T, Cuevas C, Tsiligkaridis T, Zitnik M. Domain adaptation for time series under feature and label shifts. In: Proceedings of the 40th International Conference on Machine Learning. 2023 Presented at: ICML'23; July 23-29, 2023; Honolulu, HI.
37. Wilson G, Doppa JR, Cook DJ, Wilson G. CALDA: improving multi-source time series domain adaptation with contrastive adversarial learning. IEEE Trans Pattern Anal Mach Intell 2023 Dec;45(12):14208-14221. [doi: 10.1109/tpami.2023.3298346]
38. Ganin Y, Ustinova E, Ajakan H, Germain P, Larochelle H, Laviolette F, et al. Domain-adversarial training of neural networks. J Mach Learn Res 2016;17:1-35 [FREE Full text] [doi: 10.1007/978-3-319-58347-1_10]
39. Ragab M, Eldele E, Wu M, Foo CS, Li X, Chen Z. Source-free domain adaptation with temporal imputation for time series data. In: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2023 Presented at: KDD '23; August 6-10, 2023; Long Beach, CA. [doi: 10.1145/3580305.3599507]
40. Yang S, Wang Y, van de Weijer J, Herranz L, Jui S. Generalized source-free domain adaptation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021 Presented at: ICCV 2021; October 10-17, 2021; Montreal, QC URL: https://ieeexplore.ieee.org/document/9710764 [doi: 10.1109/iccv48922.2021.00885]
41. Li H, Lin Z, Qiu Z, Li Z, Niu K, Guo N, et al. Enhancing and adapting in the clinic: source-free unsupervised domain adaptation for medical image enhancement. IEEE Trans Med Imaging 2024 Apr;43(4):1323-1336. [doi: 10.1109/tmi.2023.3335651]
42. Li L, Zhou Y, Yang G. Robust source-free domain adaptation for fundus image segmentation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2024 Presented at: WACV 2024; January 3-8, 2024; Waikoloa, HI. [doi: 10.1109/wacv57701.2024.00766]
43. Wu H, Zhou Y, Li X. Source-free domain adaptation for medical image segmentation via mutual information maximization and prediction bank. Electronics 2025 Sep 15;14(18):3656. [doi: 10.3390/electronics14183656]
44. Yang C, Guo X, Chen Z, Yuan Y. Source free domain adaptation for medical image segmentation with fourier style mining. Med Image Anal 2022 Jul;79:102457. [doi: 10.1016/j.media.2022.102457] [Medline: 35461016]
45. Yang S, Wang Y, van de Weijer J, Herranz L, Jui S. Exploiting the intrinsic neighborhood structure for source-free domain adaptation. In: Proceedings of the 35th International Conference on Neural Information Processing System. 2021 Presented at: NIPS'21; December 6-14, 2021; Virtual Event.
46. Lally P, van Jaarsveld CH, Potts HW, Wardle J. How are habits formed: modelling habit formation in the real world. Eur J Soc Psychol 2009 Jul 16;40(6):998-1009. [doi: 10.1002/ejsp.674]

XSL•FO
RenderX

47. Ryan K, Hogg J, Kasun M, Kim JP. Users' perceptions and trust in AI in direct-to-consumer mHealth: qualitative interview study. JMIR Mhealth Uhealth 2025 May 20;13:e64715 [FREE Full text] [doi: 10.2196/64715] [Medline: 40392584]
48. Kairouz P, Brendan McMahan H, Avent B, Bellet A, Bennis M, Bhagoji AN, et al. Advances and open problems in federated learning. Found Trends Mach Learn 2021;14(1-2):1-210. [doi: 10.1561/2200000083]
49. Li M, Xu P, Hu J, Tang Z, Yang G. From challenges and pitfalls to recommendations and opportunities: implementing federated learning in healthcare. Med Image Anal 2025 Apr;101:103497 [FREE Full text] [doi: 10.1016/j.media.2025.103497] [Medline: 39961211]

Abbreviations

AI: artificial intelligence
CNN: convolutional neural network
DA: domain adaptation
DL: deep learning
FL: federated learning
LSC: local structure clustering
ML: machine learning
SFDA: source-free domain adaptation

Edited by H Köttl; submitted 16.Jun.2025; peer-reviewed by L Nguyen, C Yang; comments to author 26.Aug.2025; revised version received 07.Dec.2025; accepted 18.Dec.2025; published 18.Mar.2026.

Please cite as:
Subedi R, Chakraborty S, He Z, Pang Y, Zhang S, Lustria ML, Charness N, Boot W
Predicting Adherence to Computer-Based Cognitive Training Programs Among Older Adults Using Source-Free Domain Adaptation: Algorithm Development and Validation
JMIR Aging 2026;9:e79123
URL: https://aging.jmir.org/2026/1/e79123
doi:10.2196/79123
PMID:

©Ronast Subedi, Shayok Chakraborty, Zhe He, Yuanying Pang, Shenghao Zhang, Mia Liza Lustria, Neil Charness, Walter Boot. Originally published in JMIR Aging (https://aging.jmir.org), 18.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Integrating Care Context With Skeleton and Depth Information for Older Adult Activity Recognition in a Care Facility Using Care-Assessment-Aware Spatiotemporal Transformer: Method and Validation Study

Nazmun Nahid[1], PhD; Iqbal Hassan[1], MEng; Md Atiqur Rahman Ahad[2], PhD; Sozo Inoue[1], PhD

[1]Kyushu Institute of Technology, Kitakyushu, Japan

[2]University of East London, London, United Kingdom

Corresponding Author:
Nazmun Nahid, PhD
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu Ward
Kitakyushu, 808-0135
Japan
Phone: 81 08069989264
Email: raian.nahid@gmail.com

Abstract

Background: Older adult activity recognition is a critical task in long-term care monitoring; yet, it remains challenging due to postural deformities and health-related variability. These factors cause different activities to appear visually similar, or the same activity to appear dissimilar, undermining the effectiveness of traditional human activity recognition models developed for the general population.

Objective: This study aims to develop an improved older adult activity recognition method that integrates care assessment information with motion data to capture and understand movement variability arising from different health conditions.

Methods: To achieve our objective, we propose a care-assessment-aware spatiotemporal transformer (CSTT) model that integrates body key points, heatmaps, and care level data for personalized and context-aware activity recognition. The model dynamically adjusts its attention mechanism based on care level context to improve recognition accuracy. CSTT was trained and validated on real-world older adult motion data. A total of 51 older adult participants (30 men and 21 women; age range of 64-95 years) were included in the study. Among them, 7 (13.7%) required high care assistance, 26 (51.0%) required medium care assistance, and 18 (35.3%) required low care assistance.

Results: Despite data imbalance and considerable intraclass variation due to differing care needs, the proposed CSTT model achieved an F_1-score and accuracy of 0.96 and area under the curve is 0.98. Analysis revealed that movement patterns differ significantly across care levels and that similar motions occur in distinct activities, highlighting the importance of care-aware modeling.

Conclusions: Incorporating care level information into activity recognition models significantly enhances performance in older adult care settings. The proposed CSTT framework demonstrates the value of personalized, context-sensitive approaches for accurate and ethical monitoring in long-term care environments.

(JMIR Aging 2026;9:e80102) doi:10.2196/80102

KEYWORDS

older adult activity recognition; activity recognition; care data; older adult dataset, transformer.

XSL•FO
RenderX

Introduction

Background

Advancements in medical science and technology have led to a global demographic shift, with the older adult population projected to reach 1.2 billion by 2025 and 2 billion by 2050 [1]. By 2050, nearly 20% of the world's population will be older adults [2]. As life expectancy increases, age-related physical and cognitive decline necessitates long-term care (LTC) [3], placing significant strain on health care systems and LTC facilities [4-8].

Caregivers in LTC facilities perform essential tasks, including hygiene assistance, feeding, dressing, and mobility support [7-10], while also providing emotional and psychological care [11]. However, their close interaction with residents exposes them to emotionally taxing experiences, such as witnessing chronic pain, cognitive decline, and end-of-life care [12], leading to compassion fatigue and emotional exhaustion. Caregivers also navigate ethical dilemmas, balancing residents' autonomy with safety, managing conflicts with families, and making complex decisions [13-15]. These challenges heighten stress and contribute to burnout [16-18], exacerbating the global caregiver shortage, particularly in Central Asia and Eastern Europe. High annual turnover rates (19%-55%) [19-21] worsen staff shortages, increase workloads, and reduce care quality [22]. To address shortages, LTC facilities increasingly rely on health care assistants and auxiliary nurses [23]. While vital, their limited training can hinder the care of residents with cognitive impairments or chronic conditions [24], affecting both caregiver well-being and overall health care sustainability. Ensuring caregiver well-being while maintaining high-quality older adult care is essential, as excessive strain can impact both mental health and service quality. Automatic monitoring systems offer a promising solution; however, their real-world implementation requires a robust human activity recognition (HAR) framework that accounts for older adult–specific mobility and health variations. Existing HAR methods mostly use RGB or RGB-D camera [25] and inertial measurement units [26]. Wearable inertial measurement units, although effective, can be intrusive and uncomfortable. In contrast, depth sensors and RGB cameras provide a nonintrusive, cost-effective alternative, advancing HAR through wireless sensor networks and the Internet of Things.

HAR approaches typically rely on RGB-based [27-29] or skeleton-based [30-32] methods, which perform well in general settings but degrade significantly for older adult individuals due to posture and mobility differences influenced by health conditions. A key limitation of traditional HAR models is the lack of care assessment scores (CASs), which are critical in older adult care. Caregivers assign CASs such as the activities of daily living score, Barthel Index, and care level (CL) to assess mobility, physical dependency, and assistance needs. These scores provide vital insights into an individual's functional abilities, yet state-of-the-art HAR methods fail to incorporate them, limiting their applicability to older adult populations. Since posture and motion patterns are directly affected by health conditions, neglecting these factors leads to poor generalization. Moreover, widely used HAR datasets such as RGBD-HuDaAct [33], 3D Action Pairs [34], MSR-DailyActivity [35], NTU RGB+D 120 [36], and UTD-MHAD [37] lack older adult–specific data, as they primarily feature younger or healthier participants. This demographic gap hinders HAR models from accurately recognizing older adult activities. To address this challenge, integrating CASs into HAR models is essential for personalized, context-aware monitoring, enabling more adaptive and accurate older adult activity recognition.

Related Works

Traditional HAR methods struggle with older adult individuals, leading to research on tailored models and datasets. However, challenges remain. This section reviews state-of-the-art HAR and older adult HAR methods using skeleton and video data, together with older adult–specific datasets.

Older Adult Activity Recognition

RGB-based HAR is widely used, particularly for monitoring applications. It primarily extracts motion information from video frames and can be classified into 2 categories: 2-stream networks and 3D convolutional networks. Two-stream networks leverage RGB data for spatial representation and optical flow for temporal dynamics [38-44]. However, computing optical flow is computationally expensive, creating bottlenecks in real-time applications. In contrast, 3D convolutional networks [45-48] aim to capture spatiotemporal features directly from video sequences but face challenges related to occlusions, camera motion, and environmental complexities [49]. These limitations are even more pronounced in older adult activity recognition, where subtle motion variations and fine-grained details play a crucial role. Skeleton-based approaches offer a compact and robust representation of human motion by focusing on skeletal joints and their temporal evolution. Early works used spatial graph-based models [50-53], while later studies introduced temporal relationships through recurrent and convolutional architectures [54-58]. However, skeleton-based methods lack environmental context, limiting their effectiveness in recognizing human-object interactions—an essential factor for assessing functional abilities in older adult individuals. To address the limitations of unimodal methods, researchers have explored multimodal fusion techniques that integrate RGB and skeleton data [59,60]. These methods typically extract features from each modality independently before performing fusion, with some incorporating contextual information such as human-object interactions and location data [61-63]. Despite improved recognition accuracy, existing multimodal approaches struggle with effective feature aggregation, as irrelevant modality-specific information can degrade overall performance. More importantly, these methods are predominantly trained on young, healthy individuals and lack adaptations tailored for older adult populations. A notable older adult HAR system was introduced in the study by Kim et al [64], using depth video and skeleton joint features. While effective in controlled environments, this approach failed to generalize well to real-world scenarios due to the exclusion of contextual and health-related mobility variations and the lack of real-world continuous activity patterns in the training data. Another older adult HAR method proposed a feature fusion model combining

XSL•FO
RenderX

handcrafted and deep-learned features using a dedicated dataset [65]. However, this approach remains limited by its reliance on a homogeneous environment and its inability to capture long-range dependencies.

To overcome the limitations of existing methods, we propose a cross-modal and personalized HAR approach that considers older adult health conditions. Our base model is built on transformers, which have demonstrated exceptional performance in HAR due to their ability to model long-range dependencies through self-attention mechanisms. Most transformer-based HAR methods process RGB frames as input tokens [66,67] or, less commonly, skeleton data [68,69]. However, these approaches suffer from high computational costs, restricting their applicability in real-time older adult care environments. In addition, existing transformer-based models do not efficiently integrate cross-modal information, limiting their ability to leverage multimodal dependencies for older adult HAR. To address these challenges, we propose a depth map-based approach instead of RGB to reduce computational costs. We have designed a care-aware attention mechanism (CAM) and incorporated it into the spatial layer, replacing the standard self-attention mechanism in transformers. This effectively facilitates cross-modal analysis by integrating care assessment information, skeleton data, and depth images, thereby enhancing older adult HAR performance.

Older Adult Dataset

Although most benchmark HAR datasets are designed for the general population, some datasets have been specifically collected to study older adult individuals. EGOFALLS [70] focuses on fall detection using egocentric camera data, containing 10,948 video samples from 14 participants, including 12 young adults and only 2 older adult individuals. Despite its large size, the dataset has a significantly small sample of older adult participants. Additionally, its controlled environment and lack of health condition–based variations make it unsuitable for real-world application training. The Toyota Smarthome Dataset [71] captures daily living activities using 7 Kinect sensors from 18 volunteers aged 60-80 years over 8 hours in a controlled apartment setting. While it includes interactions with household objects, its limitations stem from the controlled environment and the absence of health condition–based variations. IntelliRehabDS [72] was collected using a Kinect motion sensor and comprises 9 repetitive gestures performed by 29 individuals, including 15 patients and 14 healthy controls. It provides 3D body joint coordinates and depth maps, annotated for gesture type and position (sitting or standing). However, the dataset has a narrow older adult age range (20-60+ years), is collected in controlled conditions, and lacks health condition–specific data. ETRI-Activity3D [73] is a large-scale dataset that includes RGB videos, depth maps, and skeleton sequences from 100 participants—50 older adult individuals (aged 64-88 years, with an average age of 77 years) and 50 younger adults (average age of 23 years). Despite its scale, the dataset has limitations, including a lack of intervention data, health condition–based variations, and continuous activity recordings. Overall, existing older adult HAR datasets suffer from small and homogeneous older adult sample sizes, controlled environment constraints, and limited continuous activity data, often lacking health condition–specific variations. To address these issues, we collected our dataset in a real care facility without intervention or manipulation, capturing mealtime sessions of 28 older adult participants aged 62-95 years, representing 5 distinct CLs.

Objective and Contributions

In this work, our objective is to develop an improved older adult activity recognition method that integrates care assessment information with motion data to effectively capture and understand movement variability caused by different health conditions. To achieve this, we incorporated CL, one of the most widely used CASs in older adult care facilities, due to its availability and relevance in evaluating functional abilities, and introduced a care-assessment-aware spatiotemporal transformer (CSTT) that adapts its attention to key points and depth-based motion patterns based on an individual's CL, enabling personalized feature prioritization and improved activity prediction. We also addressed the issue of the lack of suitable data and collected a real-world older adult activity dataset incorporating CL information. The key contributions of our work can be summarized as follows:

1. We propose the first care-assessment-aware activity recognition approach by modeling the correlation between health conditions using CL and movement patterns to personalize and improve older adult activity recognition.
2. We propose CSTT, integrating skeleton, depth heatmaps, and CL information. Our proposed CAM dynamically adjusts focus based on care needs, ensuring personalized recognition while enhancing robustness and efficiency.
3. We present the first motion dataset with CL information, capturing real-world mealtime sessions to reflect aging and health impacts on mobility with high ecological validity.
4. This work improves older adult activity recognition by integrating care assessment information and capturing motion variations influenced by these conditions for improved accuracy.

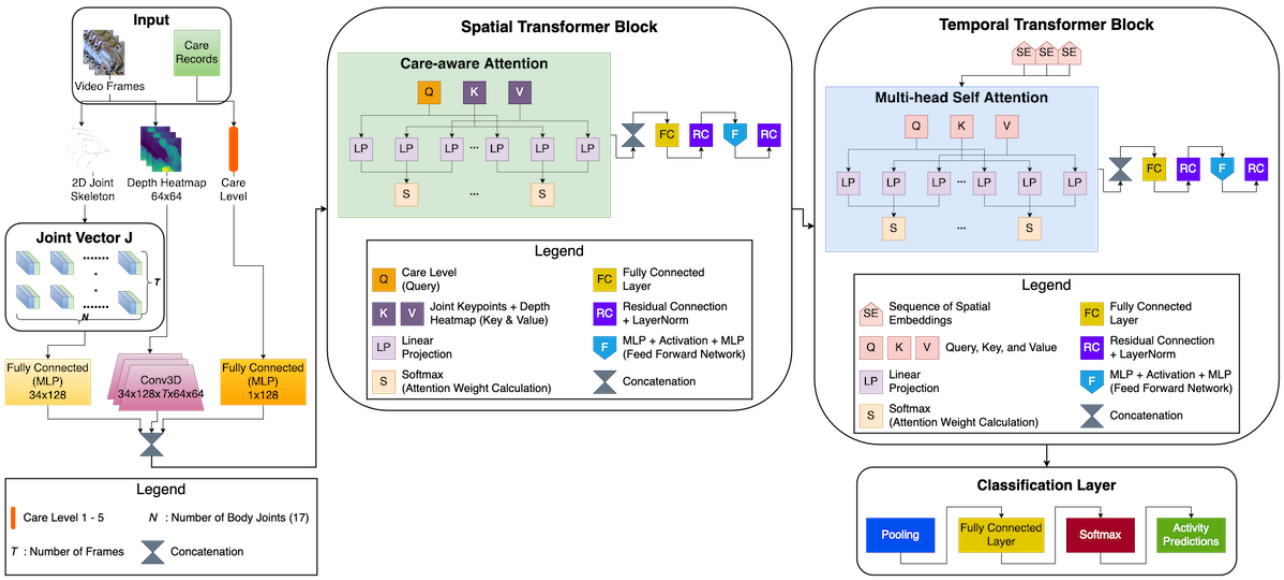
Methods

Study Design

In this work, we tackled the challenge of older adult activity recognition, which is complicated by pose deformities and motion limitations in older adults. To address these issues, we proposed a heterogeneous spatiotemporal motion transformer with CAM, specifically designed for recognizing older adult activities. For model training, we collected a dataset with informed consent and ethical permission,  where, X_i consists of body key points, depth images, and CL information, while y_i represents the corresponding activity label, and N denotes the number of training samples. The objective is to learn a function $f: X \rightarrow y$ that accurately classifies activities. In this section, we provide a detailed explanation of our data processing, followed by an in-depth discussion of the CSTT transformer. The overall architecture is illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Care-assessment-aware spatiotemporal transformer architecture. This transformer enhances older adult activity recognition by integrating joint skeleton data, depth heatmaps, and CL information. The Spatial Transformer Block uses a care-aware attention mechanism, where the care level acts as a query to capture spatial relationships. The Temporal Transformer Block uses multihead self-attention to model temporal dependencies across frames. Finally, the Classification Layer processes the learned representations to predict activity labels using a fully connected network and softmax activation.



Data Collection and Validation Method

Overview

The dataset was gathered in partnership with Global Care, a care facility in Japan, dedicated to supporting patients with dementia, which is equipped with an in-house video-monitoring system. The camera used for the monitoring is the AXIS M3048-P [74], a cost-effective fixed dome fish-eye camera featuring a 12-megapixel sensor. This camera is designed to provide a comprehensive 360° view of the surroundings, with distortion-corrected display options such as panoramic views, specific areas, corridors, corners, and quad displays, all offering exceptional sharpness. Additionally, the camera comes prefocused, eliminating the need for manual adjustments. To maintain the authenticity of the care environment, we chose not to incorporate any additional sensors. In collaboration with the care facilities, we accessed their video recordings, which specifically captured activities of 51 older adult people. For dataset 1, data was collected from 28 elderly during lunch mealtime from 3 different sites over a span of 15 days. Each recorded session lasted between 30 and 60 minutes. Older adult participants' ages are ranging from 64 to 95 (mean 79.5) years. The placement of each participant was predetermined and managed by the staff of the care facility. To maintain the facility's natural workflow, our team provided no instructions or interventions. In site 2, as shown in Figure 2, patient 7 was assigned multiple positions, as their location was occasionally adjusted by the staff during certain sessions. This variation in positioning was carefully considered during the data processing phase to ensure the accuracy of the labels. For dataset 2, the other 23 participants' data were collected from open-gathering space of dining and living area in site 4 from 2 different floor. Participants are aged between 71 and 92 (mean 81.5) years. The data were collected over 15 days period through continuous 24 hours monitoring. The site layouts are given in Multimedia Appendix 1.

Due to the limitations in posture and mobility, the activities of older adult individuals differ significantly from those of healthy individuals. Similar activities may present with different postures depending on CL, while different activities might appear with similar postures across different CL groups. As a result, CL may infer crucial insights for older adult activity recognition. To address this, we have collected each older adult person's CL alongside the video data. The score was provided by the medical professionals based on assistance requirements in meal, bath, excretion, movement, and dress-up. These levels, ranging from 1 (minimal assistance) to 5 (maximum assistance), determine the extent of supervision and support required [75]. Table 1 shows the explanation regarding the CL. For easier understanding of the readers, we have categorized the CLs based on assistance requirements into low, mid, and high.

From the recorded videos, skeleton data and depth images were extracted for analysis. For the purpose of this study, only skeleton data, depth images, and care-related information were used, while RGB images were excluded from the analysis. In Table 2, baseline comparison is shown between the 2 collected dataset.

Comparing the components between the 2 datasets in Table 2, we found that gender, care assistance requirement, and data collection time are significantly different. The age does not have much variation because the target group of our research is 60 above to 100.

To highlight the necessity of incorporating care context with motion data for activity recognition, we used 2 similarity comparison approaches. The first involves calculating pairwise similarity among different CL groups, while the second focuses on comparing the pairwise similarity between low and medium CL groups for various activities. To compute similarity scores,

XSL•FO
RenderX

we used mean per joint angle difference (MPJAD), cosine similarity (CS), and histogram of oriented gradients (HOG) similarity. Detailed explanations of the score calculation methods are given in the following sections:

Figure 2. Model performance evaluation: the receiver operating characteristic curve illustrates care-assessment-aware spatiotemporal transformer's prediction accuracy across different activities (class 0: sitting, class 1: eating, class 2: stand up, and class 3: trying to stand up). ROC: receiver operating characteristic curve.

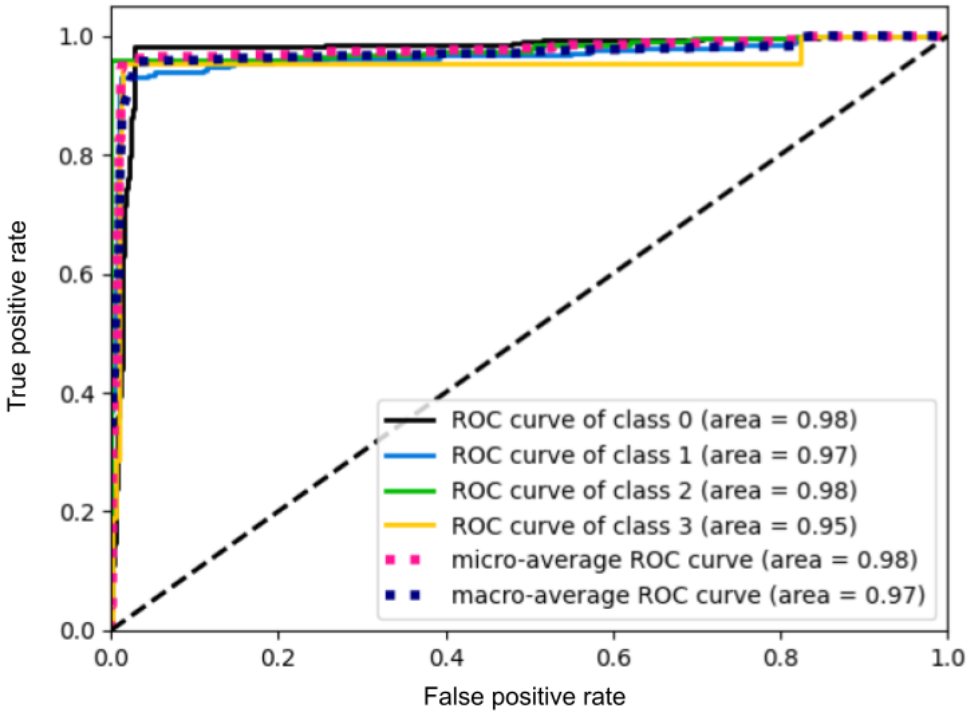


Table 1. Care-level interpretation for the older adult individuals.

AR[a]	CLs[b]	Older adult condition
Low	1 and 2	Older adult individuals in this group can move independently or with minimal aid. They sometimes require monitoring and reminders for meals, excretion, and baths. For dress-up, they sometimes need monitoring or partial assistance. They generally maintain an upright posture, with only minor stooping due to age-related spinal degeneration.
Mid	3	Older adult individuals at this stage require full support from caregivers for movement; however, they can stand up with the support of assistive devices. They require monitoring, reminders, and partial assistance for meals, excretion, and baths. For dress-up, they need full assistance. Posture is often characterized by forward leaning or hunching due to weakened core muscles and joint instability. Sitting may involve slumping as maintaining an upright position becomes difficult.
High	4 and 5	Older adult individuals at this stage require full support from caregivers for everything. They are predominantly immobile. Postural control is severely compromised. No standard posture can be seen.

[a]AR: Assistance requirement.

[b]CLs: care levels.

XSL•FO
RenderX

Table 2. Baseline comparison between the datasets.

Components	Dataset 1	Dataset 2	*P* value
Age (years), mean (SD)	79.5 (8.95)	81.5 (6.06)	.35
Gender			
Men	10	20	<.001
Women	18	3	
Care assistance requirement			
High	6	1	<.01
Medium	8	18	
Low	14	4	
Data collection time (minute)	750	21,600	<.001

Mean per Joint Angle Difference

MPJAD measures the average angular difference between 2 motion sequences across all frames and joints. Given 2 motion sequences and , the MPJAD is computed as:



(1)

Where P is the set of all motion sequence pairs, M is the number of motion sequence pairs, T_{ij} is the number of frames common between sequences S_i and S_j, N is the total number of joints (17 in this case),  is the joint angle at joint k and frame t in sequence S_i, and  is the absolute difference in joint angles. A lower MPJAD value indicates a higher similarity between motion sequences. The 2 motions follow similar joint movement patterns. A higher MPJAD value indicates greater dissimilarity between motion sequences. The 2 motions have significantly different joint movements.

Cosine Similarity

CS measures the angle between 2 vectors in a multidimensional space. Given 2 motion sequences S_i and S_j, the CS is computed by following the steps:

Derive the cosine dissimilarity (used to measure homogeneity):



(2)

Where S_i and S_j are the flattened motion sequences represented as vectors. $S_i \cdot S_j$ is the dot product of the 2 vectors. $|S_i|$ and $|S_j|$ are the Euclidean norms (magnitudes). Compute the mean pairwise cosine dissimilarity within the class:



(3)

Where N is the total number of unique pairs (i,j) within the class. Lower values of cosine dissimilarity indicate higher homogeneity (motion vectors within the class are closely aligned). Higher values of cosine dissimilarity indicate lower homogeneity (motion vectors within the class are more divergent).

HOG Similarity

HOG measures how structurally similar 2 images are based on their gradient information. Once the HOG feature vectors H_1 and H_2 are extracted for 2 images, their similarity is computed using the CS formula:



(4)

Where $H_1 \cdot H_2$ is the dot product of the 2 HOG feature vectors. $|H_1|$ and $|H_2|$ are the Euclidean norms of the vectors. If the HOG score is 1, it means the images have nearly identical edge and gradient structures, and if it is 0, it means the images have no similarity in gradient structure.

Ethical Considerations

This work involved human participants or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the research ethics board at Kyushu Institute of Technology (application no. 24-15). No participant compensation was provided. The data are not publicly available but they would be shared with upon the collaboration request only for research purpose to the applicants who adhere to the ethical and privacy policy of the ethical committee after careful consideration.

Skeleton Data Generation

To ensure robustness, we used skeleton data in our study. To get the skeleton data we used YOLOv7 [76], a state-of-the-art real-time object detection framework that builds upon the foundational YOLO (You Only Look Once) [77] family, offering enhanced performance in terms of accuracy. It maintains high precision in both detection and key point localization, and the architecture can handle multiple scales and dense environments effectively. The input is an RGB image, $I_{RGB} \in R^{H \times W \times 3}$ where H is the height of the image, W is the width of the image, and 3 represents the number of color channels (ie, red, green, and blue). YOLOv7 identifies bounding boxes for human figures in the input RGB image. Each bounding box B_k is represented as $B_k = (x_{min}, y_{min}, x_{max}, y_{max}, c_k)$, where (x_{min}, y_{min})

XSL•FO
RenderX

and (x_{max},y_{max}) define the corners of the box and c_k is the confidence score. Within each bounding box, the model then predicts the locations and confidence scores for the 17 key points $X = \{X_1,X_2,\ldots,X_{17}\}$, where each key point X_i is defined as $X_i = (x_i,y_i,c_i)$, i$\in$ {1,2,…,17}. Here, x_i,y_i: 2D coordinates of the ith key point in the image and c_i: confidence score indicating the likelihood that the ith key point is correctly detected ($c_i \in$ [2]). So, the whole body can be represented as:



(5)

where the total key point feature vector has 34 dimensions (17 key points × 2 coordinates).

Depth Image Generation

Skeleton data are highly robust but fail to capture the environmental context. In contrast, RGB data often include excessive information, which can lead to confusion due to varying lighting conditions. However, depth images offer valuable spatial insights, providing detailed information about the distance and positions of objects and body parts. Therefore, we generated depth maps from RGB images and incorporated them into our study. To achieve this, we used the "Depth Anything" model [78], an advanced monocular depth estimation technique that transforms RGB images into depth maps. This model leverages deep learning (DL) to predict pixelwise depth values based on visual features in the image. Trained on extensive datasets of RGB-depth pairs, the model knows how to estimate depth using visual information alone. Each pixel in the image has 3 values corresponding to its red, green, and blue intensities. The goal of the model is to convert this 3-channel image into a depth map, where each pixel value corresponds to the distance from the camera to the object in the scene. Let f_θ represent the trained model that maps an RGB image IRGB to its predicted depth map D. The model was learned from large datasets containing paired RGB images and their corresponding depth maps during training. The depth estimation function is given by $D = f_\theta(I_{RGB})$. Then, each pixel of the depth map represents that the estimated distance to the camera can be denoted by



(6)

where (x, y) are the coordinates of the pixel in the image. The value D (x, y) represents the depth of the corresponding point in the scene.

Care-Assessment-Aware Spatiotemporal Transformer

Our proposed CSTT analyzes body key points, heatmaps, and CL to predict human activities in a personalized manner. As a spatiotemporal heterogeneous transformer, it captures spatial features (key points and heatmaps) and temporal dynamics (motion sequences over time) while integrating multiple modalities with distinct representations—numerical key points, image-based heatmaps, and scalar CL—through CAM. In the Spatial Transformer, key points and heatmaps are processed to extract movement-related features, with the CL serving as an attention guide, determining which body movements and heatmap regions are most relevant. This enables the model to dynamically prioritize motion patterns based on an individual's care needs. The Temporal Transformer then analyzes the sequence of spatial features, capturing motion dynamics over time. Learning temporal dependencies helps recognize activities and transitions between postures. Finally, the extracted spatiotemporal features are passed to a classifier, which predicts the activity category.

Spatial Transformer

The Spatial Transformer extracts meaningful representations from the key points and heatmap while considering the CL as a crucial guiding factor. The CL attends to the key points and heatmap, allowing the model to emphasize relevant body parts or movements based on the older adult person's health condition.

Feature Embedding

Overview

To transform the raw input data into a meaningful and structured format suitable for further processing, we used an embedding mechanism. This step is crucial as it converts different types of input data—body key points, heatmaps, and CL—into a shared latent representation that can be effectively used by the transformer model. By mapping these diverse inputs into a common latent space of dimension d, the embedding process ensures that the model can seamlessly integrate and compare information from multiple modalities, facilitating efficient learning and cross-modal interactions. There are 3 types of feature embedding done here as mentioned in the following sections.

Key Point Embedding

The 2D pose key points (17 joints, each with (x,y) coordinates) are flattened and passed through a linear layer:



(7)

Where $X \in R^{B\times17\times2}$ is the input key point tensor and $W_k \in R^{(17\times2)\times d}$ is the weight matrix, which learns spatial dependencies. $K \in R^{B\times d}$ is the key point embedding.

Heatmap Embedding

The heatmap (grayscale image of size 64 × 64) is processed through a 3D convolution layer to capture local spatial dependencies and activation patterns:



(8)

where, $D \in R^{B\times1\times64\times64}$ is the heatmap input. $H \in R^{B\times T\times d}$ is the heatmap embedding after convolution, where T represents the number of spatial patches.

CL Embedding

The CL (a scalar value) is passed through a linear layer:

XSL•FO
RenderX



(9)

Where $c \in R^{B\times 1}$ is the CL tensor and $Wc \in R^{1\times d}$ is the weight matrix. $C \in R^{B\times d}$ is the CL embedding.

Care-Aware Attention Mechanism

Overview

Unlike conventional transformer self-attention mechanisms that treat all inputs equally, our proposed CAM dynamically adjusts attention based on CL. Using CL embeddings as queries in a multihead attention mechanism prioritizes key body parts and movements. Key points and heatmaps form key value pairs, allowing the model to focus on relevant features. The attended features are refined through a feedforward network, enhancing learning. By explicitly integrating CL into feature learning, CAM improves interpretability and efficiency, making activity recognition more personalized and accurate. The entire process is depicted as follows:

Attention Score Computation

The traditional attention mechanism follows the scaled dot product attention formulation:



(10)

Here, we made the modification in the embeddings of Q, K, and V values. CL embedding $C \in R^{B\times 1\times d}$ is assigned to Q (Query). Concatenated Key point + Heatmap Features $[K,H] \in R^{B\times(T+1)\times d}$ is assigned to both K (Key) and V (Value). So the modified attention weights computation formula is:



(11)

This results in a weighted sum of key points and heatmap features, emphasizing relevant information based on the CL.

Multihead Attention

To enhance model expressiveness, we use multihead attention, where different heads capture different aspects of the input:



(12)

where each head performs attention independently, and the outputs are concatenated and projected using W_o.

Feedforward Network

After attention, the output is passed through a feedforward network:



(13)

This enhances the feature representation before passing it to the temporal transformer.

Temporal Transformer

The temporal transformer is designed to capture and model the sequential dependencies that exist across frames. Since traditional transformers do not inherently account for the order of sequences, we incorporate positional encoding to inject information about the temporal order of the frames, allowing the model to distinguish between the different points in time:



(14)

Where ρ is a constant eprecally calculated and the value of ρ is 10,000. The Transformer Encoder then applies self-attention across the temporal dimension, processing a sequence of spatial features Z_t obtained from various timesteps. This process is mathematically represented as:



(15)

Where $Z_t \in R^{T\times d}$ is the temporal sequence of spatial representations and  is the transformed sequence after passing through the encoder. Throughout this process, the model allows the frames to “attend” to one another, meaning each frame is evaluated in relation to others. This enables the model to capture the dependencies of motion, such as recognizing how a body part moves or changes over time, effectively tracking the progression of movement across frames. After the temporal sequence is processed, the features are aggregated using mean pooling to summarize the sequence of attended frames into a single fixed-size representation:



(16)

This pooled feature vector encapsulates the entire movement sequence, providing a compact yet informative summary. Finally, this aggregated representation is passed into a classifier for further interpretation, enabling the model to make predictions based on the captured temporal dynamics.

Classification

The final feature vector encapsulates the entire activity, providing a contextual representation of the movement. After the pooled feature vector is obtained from the Temporal Transformer, it is passed through a softmax classifier for classification:



(17)

In this equation, W_{cis} and b_{cis} are learnable parameters that the model optimizes during training. Specifically, W_{cis} is responsible for projecting the final feature vector into activity categories,

XSL•FO
RenderX

effectively mapping the feature space to the set of possible classes. The softmax function then assigns probabilities to each of these activity categories, indicating the likelihood of each class being the correct one.

Results

In this work, we have proposed a care-assessment-aware approach for older adult activity recognition. This section presents the validation results for our dataset and evaluates the model's performance. In addition, we have conducted an ablation study and provided a detailed discussion of the obtained results.

Dataset Validation

To assess the dataset and demonstrate the importance of integrating care context, we applied 2 similarity-based evaluation strategies. The first measures pairwise similarity across all care-level groups, shown in Table 3 and the second examines pairwise similarity specifically between the low- and medium-care groups for each activity, shown in Table 4. Similarity was quantified using MPJAD, CS, and HOG similarity.

From Table 3, it is evident that the movement patterns of eating and sitting activities in the high care assistance group differ significantly from those in the other 2 groups. Since individuals in this category are almost immobile, activities such as trying to stand up and standing up are absent. For the low and medium care assistance groups, we conducted a pairwise similarity analysis of activities, as we observed that the motion patterns of trying to stand up and standing up share some similarities and yet exhibit notable differences. From Table 4, we can see that, despite being different activities, certain motion patterns show a high degree of similarity. In particular, the pairs eating–trying to stand up and standing up–trying to stand up demonstrate strong similarities, which is uncommon in standard activity recognition scenarios. These findings confirm that care context plays a crucial role in older adult activity motion patterns. To strengthen our claim, we also analyzed motion patterns using a comparative group of young adults. Since the study primarily focuses on older adult participants, these additional results are provided in Multimedia Appendix 1 for reference. All these comparison results validate our claim that care information plays a significant role in accurately distinguishing activities.

Table 3. Pairwise similarity calculation among different care level groups.

	Mean per joint angle difference			Cosine similarity			Histogram of oriented gradients similarity		
	[a]L-H	[b]H-M	M-L[c]	L-H	H-M	M-L	L-H	H-M	M-L
E[d]	0.95	0.92	0.18	0.95	0.92	0.16	0.07	0.14	0.91
S[e]	0.95	0.92	0.24	0.95	0.92	0.27	0.07	0.14	0.87
T[f]	N/A[g]	N/A	0.49	N/A	N/A	0.51	N/A	N/A	0.25
SU[h]	N/A	N/A	0.78	N/A	N/A	0.80	N/A	N/A	0.34

[a]L-H: low and high care–level pair.

[b]H-M: high and medium care–level pair.

[c]M-L: medium and low care–level pair.

[d]E: eating.

[e]S: sitting.

[f]T: trying to stand up.

[g]N/A: not applicable.

[h]SU: stand up.

XSL•FO
RenderX

Table 4. Activity pairwise similarity calculation between low and medium care–level groups. The first activity is from the medium group.

Activity pair	Mean per joint angle difference	Cosine similarity	Histogram of oriented gradients similarity
E[a]-S[b]	0.91	0.93	0.14
E-T[c]	0.21	0.23	0.91
E-SU[d]	0.56	0.60	0.72
S-E	0.92	0.95	0.12
S-T	0.55	0.59	0.71
S-SU	0.54	0.59	0.72
T-E	0.22	0.24	0.90
T-S	0.56	0.61	0.72
T-SU	0.21	0.23	0.88
SU-E	0.56	0.59	0.73
SU-S	0.51	0.53	0.71
SU-T	0.25	0.27	0.89

[a]E: eating.

[b]S: sitting.

[c]T: trying to stand up.

[d]SU: stand up.

Model Performance Evaluation

To evaluate the model's performance, we analyzed the attention weight matrix (AWM), the receiver operating characteristic curve, and the cumulative gain plot. The AWM in Figure 3 validates that the CSTT model efficiently integrates CL, body key points, and depth heatmaps through CAM. The model dynamically adjusts the importance assigned to different input features, demonstrating its ability to capture hierarchical dependencies. Notably, CL serves as a guiding factor, influencing how attention is distributed across other features. The receiver operating characteristic curve in Figure 2 evaluates the model's ability to distinguish activity classes by plotting the true-positive rate versus the false-positive rate. The area under the curve (AUC) values range from 0.95 to 0.98, indicating excellent classification performance. The microaverage AUC (0.98) reflects strong overall accuracy, while the macroaverage AUC (0.97) shows balanced performance across classes. Classes 0, 1, and 2 have the highest AUC values, ensuring clear separation, while class 3 (0.95) shows slight overlap but still performs well. The cumulative gain plot in Figure 4 highlights the model's ability to rank correct predictions early. The steep rise in curves shows effective prioritization. This confirms that the CSTT efficiently ranks true positives early, which is critical for applications needing confident and rapid classification.

We selected both machine learning and DL models as baselines to evaluate our approach. Table 5 presents the results with and without CL information, except for our model, where care context is essential when the data are benchmarked using a cross-day approach, applying an 80-20 train-test split. Machine learning models used only skeleton data, resulting in comparatively lower performance. However, adding just care context significantly improved all models' performance. Our model outperformed the traditional spatiotemporal transformer, achieving a 5% higher accuracy and a 9% increased F_1-score.

XSL•FO
RenderX

Figure 4. Model performance evaluation: the cumulative gain plot (one vs all) illustrates care-assessment-aware spatiotemporal transformer's prediction accuracy across different activities (class 0: sitting, class 1: eating, class 2: stand up, and class 3: trying to stand up). Here, red, orange, green, blue, and dotted lines represent class 3, class 1, class 2, class 0, and baseline accordingly.

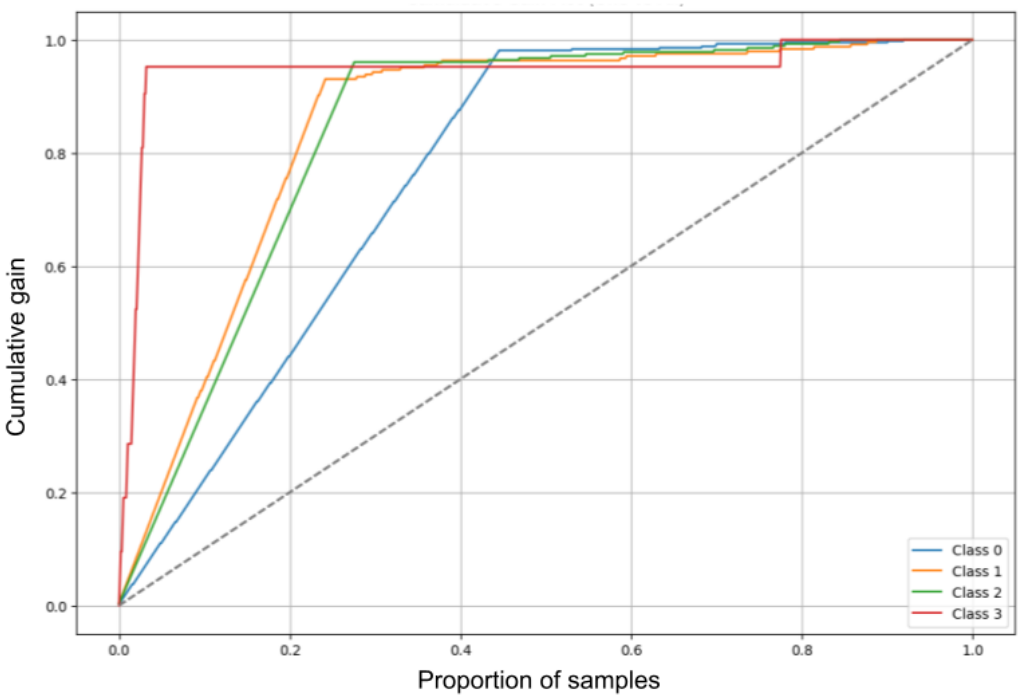


Table 5. Comparison with baseline methods.

Method	Without care (C) level				With care (C) level			
	P[a]	R[b]	F_1-score	A[c]	P	R	F_1-score	A
Random Forest (S[d])	0.58	0.65	0.56	0.65	0.69	0.69	0.69	0.69
XG Boost (S)	0.67	0.69	0.66	0.69	0.74	0.74	0.74	0.74
CNN[e] (S + D[f])	0.76	0.75	0.76	0.77	0.82	0.82	0.81	0.80
ResNet (S + D)	0.80	0.74	0.76	0.86	0.85	0.85	0.85	0.85
TSTT[g] (S + D)	0.82	0.83	0.82	0.83	0.83	0.91	0.87	0.91
Our model	N/A[h]	N/A	N/A	N/A	*0.97*[i]	*0.96*[i]	*0.96*[i]	*0.96*[i]

[a]P: precision.
[b]R: recall.
[c]A: accuracy.
[d]S: skeleton data.
[e]CNN: convolutional neural network.
[f]D: depth image.
[g]TSTT: traditional spatiotemporal transformer.
[h]N/A: not applicable.
[i]Best performance results values are in italics to highlight the importance.

Ablation Study

In the ablation study, we removed components from the key and value pairs to analyze their impact. Also, we have randomly removed some key points to create partial skeleton data. We compared the traditional spatiotemporal transformer without the CAM layer, a care-assessment-aware transformer using only skeleton data (CSTT [S + C]), a care-assessment-aware transformer using only partial skeleton data (CSTT [Sp + C]), one using only depth images (CSTT [D + C]), our full model integrating partial skeleton data (CSTT [Sp + D + C]), and our full model integrating all (CSTT [S + D + C]). As shown in Table 6, CSTT (S + C) outperforms CSTT (D + C), aligning with the AWM analysis in Figure 3. Also, CSTT (Sp + C) performs poorly compared with others, but CSTT (Sp + D + C) performance is in the acceptable range. This shows that the model is comparatively robust.

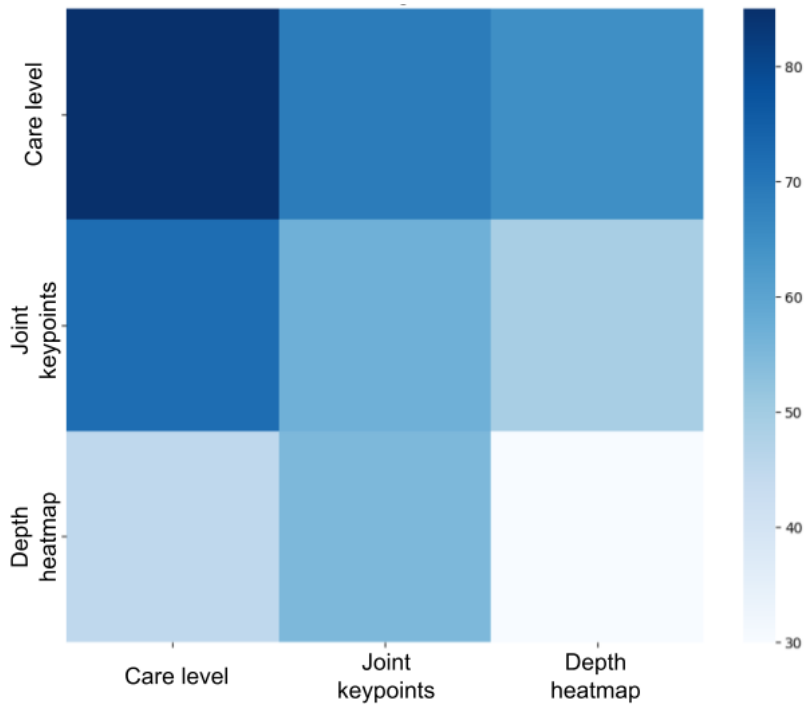
XSL•FO
RenderX

Table 6. Ablation study comparison.

Method	P[a]	R[b]	F_1-score	A[c]
TSTT[d] (S[e] + D[f])	0.82	0.83	0.82	0.83
TSTT (S + D + C[g])	0.83	0.91	0.87	0.91
CSTT[h] (S + C)	0.94	0.84	0.88	0.93
CSTT (Sp[i] + C)	0.81	0.88	0.84	0.85
CSTT (D + C)	0.91	0.82	0.84	0.90
CSTT (Sp + D + C)	0.92	0.90	0.91	0.91
Our model	*0.97*[j]	*0.96*[j]	*0.96*[j]	*0.96*[j]

[a]P: precision.
[b]R: recall.
[c]A: accuracy.
[d]TSTT: traditional spatiotemporal transformer.
[e]S: skeleton data.
[f]D: depth image.
[g]C: care context.
[h]CSTT: care-assessment-aware spatiotemporal transformer.
[i]Sp: partial skeleton data.
[j]Best performance results values are in italics to highlight the importance.

Figure 3. Model performance evaluation: attention weight matrix demonstrates the effective incorporation of care level through care-aware attention mechanism in care-assessment-aware spatiotemporal transformer.



Deployment

All experiments were implemented using the PyTorch DL framework and executed on an NVIDIA RTX 4080 GPU. The models were trained with a batch size of 8, an initial learning rate of 0.001, a weight decay of 0.0005, and for a total of 25 epochs. The trained model achieved an inference speed of 0.0043 seconds per sample, with approximately 1.2 M parameters and 1.09 Giga Floating-Point Operations per Second, indicating high computational efficiency.

Discussion

Principal Findings

This study demonstrated that older adult activity recognition can be significantly improved by integrating care-level

XSL•FO
RenderX

information into model design. The proposed CSTT effectively captured variations in motion patterns caused by differing health conditions and care requirements. Results revealed that even for the same activity, movements differed across CLs, while visually similar motions appeared in distinct activities (eg, eating vs trying to stand up). Despite class imbalance and naturally occurring variations, CSTT achieved high recognition performance. Misclassifications—such as confusion between eating and sitting—were primarily linked to overlapping movements during caregiver assistance. These findings validate that incorporating care assessment information enables more robust, context-aware, and personalized activity recognition, aligning closely with real-world monitoring needs in long-term older adult care environments.

Care-Assessment-Aware Older Adult Activity Recognition

Older adult activity recognition presents unique challenges due to the pose deformities, motion limitations, and variations in mobility caused by differing health conditions. Traditional HAR models often fail to capture these subtleties, as they are designed primarily for younger, healthy individuals and do not incorporate health-related variations. Furthermore, existing HAR approaches used for monitoring predominantly rely on RGB-based methods or, to a lesser extent, skeleton data, both of which struggle with computational efficiency and real-time applicability in older adult care environments. Our goal is to develop a context-aware HAR model that integrates CL assessments with motion data for personalized and accurate older adult activity recognition. To achieve this, we introduced CSTT, a heterogeneous spatiotemporal motion transformer incorporating skeleton data, depth-based heatmaps, and CL information, with CAM for improved adaptability. To ensure that the model captures real older adult activity patterns, we collected a motion dataset with CL annotations, recording uninterrupted mealtime sessions in a care facility for greater ecological validity. For model training, we focused on 4 key activities: eating, sitting, trying to stand up, and standing up. These were selected based on real nursing challenges—monitoring whether individuals are eating or sitting idle (requiring intervention if necessary) and ensuring continuous supervision during standing attempts to prevent falls. Unlike controlled datasets, our activity selection reflects real-world monitoring needs, capturing long-range dependencies in continuous time data, which aligns more closely with practical older adult care scenarios.

Performance of Care-Assessment-Aware Older Adult Activity Recognition

The results in Tables 3 and 4 support our initial hypothesis that older adult activity recognition differs from conventional methods, as their motion patterns are heavily influenced by their health conditions. Even for the same activity, movement can vary significantly across different CLs, while similar motion patterns may appear in completely different activities (eg, eating and trying to stand up). Our data collection was conducted without intervention to preserve real-world conditions. However, the dataset is highly imbalanced, particularly for the “trying to stand up” activity, which has significantly fewer samples than the other 3 classes. This reflects real-world scenarios but presents a challenge for recognition models. From Figure 5, the prediction accuracy for this class is 0.95, which is slightly lower than others. Among all classes, “eating” has the lowest accuracy at 0.93. Examining the confusion matrices provides insights into this misclassification. In Figure 6, eating is frequently confused with sitting, which is reasonable since caregivers often assist older adult individuals, reducing distinct motion patterns between these activities. In Figure 7, eating is often mistaken for trying to stand up, likely because older adult individuals extend their hands and lean forward for support when getting up—movements similar to hovering over a tray to pick up food. However, as seen in Figure 8, low care assistance individuals do not exhibit such overlapping movements, leading to perfect recognition of the eating activity. Based on overall performance, our model successfully outperforms the baselines.

XSL•FO
RenderX

Figure 5. Normalized confusion matrix for all groups together.

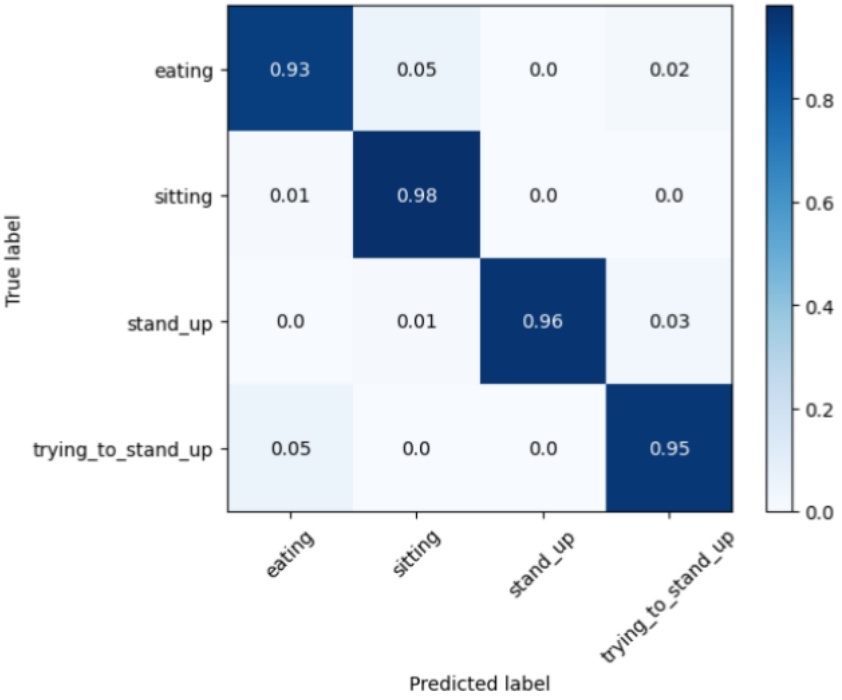
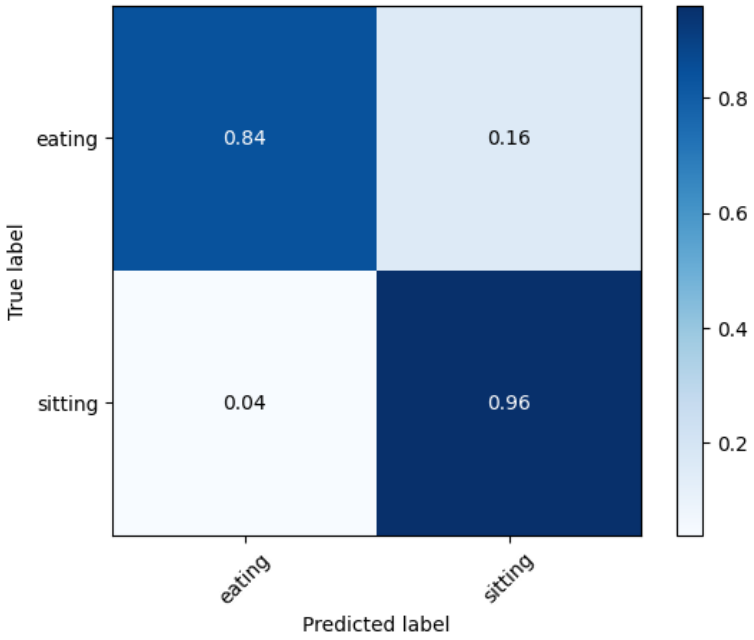


Figure 6. Normalized confusion matrix for high care assistance required group.



XSL•FO
RenderX

Figure 7. Normalized confusion matrix for medium care assistance required group.

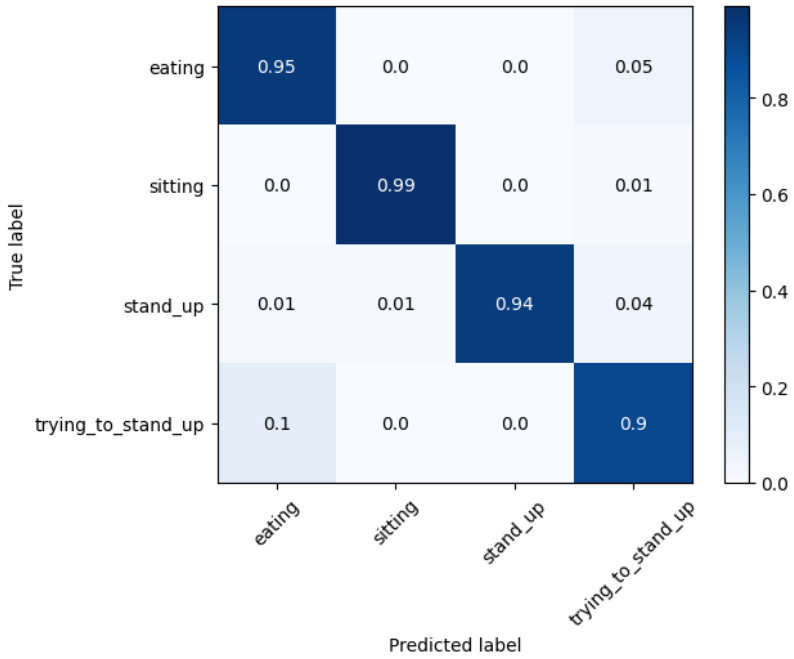
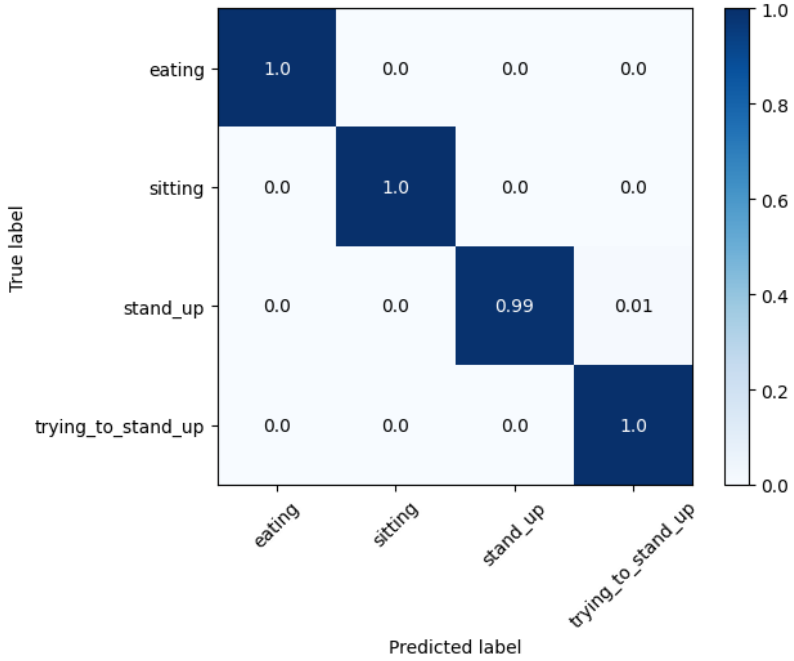


Figure 8. Normalized confusion matrix for low care assistance required group.



Current Limitations and Future Work

Although our dataset reflects real-world scenarios, it remains imbalanced, with activities such as “trying to stand up” having fewer samples. This affects classification accuracy (Figure 5), where “trying to stand up” achieves 0.95 and “eating” the lowest at 0.93 due to motion similarities, especially with caregiver assistance. We did not apply any data imbalance handling techniques, as “trying to stand up” naturally occurs less frequently in real-world settings. We did not apply data augmentation to address class imbalance because in older adult care monitoring, maintaining natural motion patterns and ecological validity is crucial. In cases of severe imbalance, augmentation risks producing unrealistic or biased samples, distorting class distributions, and leading to overfitting to synthetic data. The absence of CL annotations in existing datasets restricted direct comparisons with previous older adult HAR approaches. Although we manually annotated CLs for the ETRI Activity3D and Toyota Smarthome datasets to enable comparison, the potential inaccuracy of these inferred labels and the no representation of high-assistance older adult groups led us to include these results in Multimedia Appendix 2 rather than in the main manuscript. Future work will focus on

XSL•FO
RenderX

annotating CLs in public datasets with medical professionals for comparative analysis. Due to ethical restrictions, our dataset cannot be publicly shared but may be accessed through collaboration. To support replication, we provide references and guidelines for collecting similar datasets with integrated CL information. Since our model depends on accurate CL assessments, misclassification can impact recognition, emphasizing the importance of expert collaboration in dataset development.

Conclusions

In this work, we address the challenge of older adult activity recognition by considering the variability in movements influenced by CLs, which is often overlooked in existing datasets. To overcome this, we introduced a novel older adult motion dataset that includes CL information, collected from 51 older adult participants during real-world mealtime sessions in an ethical and privacy-preserving manner. We proposed CSTT, a spatiotemporal heterogeneous transformer model that integrates body key points, heatmaps, and CLs to predict activities in a personalized way by dynamically adjusting attention based on CLs. Our model surpassed the conventional spatiotemporal transformer, achieving a 5% higher accuracy and a 9% improvement in F_1-score, highlighting the significance of incorporating CL data. However, limitations include dataset imbalance and the inability to compare with similar works due to a lack of comparable datasets, as well as ethical restrictions on sharing the dataset. In conclusion, our work lays the foundation for more accurate, context-aware older adult activity recognition, with future research focusing on dataset expansion, model refinement, and real-world applications in care settings.

Acknowledgments

This research was supported by the JST Challenging Research Program for Next Generation Researchers (grant JPMJSP2154). Grammarly AI and Quilbot AI have been used for grammatical corrections and paraphrasing assistance. The depth images used in this study were generated using DepthAnything model. The details are provided in “Methods” (B. Depth Image Generation) section.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The data are not publicly available but they would be shared with upon the collaboration request only for research purpose to the applicants who adhere to the ethical and privacy policy of the ethical committee after careful consideration. Also, we have provided resources for reproducibility in Multimedia Appendix 1.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Validation of care aware data collection and dataset reproducibility resources.
[DOCX File , 242 KB - aging_v9i1e80102_app1.docx]

Multimedia Appendix 2
Generalization of care-assessment-aware spatiotemporal transformer by comparing performance on other data sets.
[DOCX File , 18 KB - aging_v9i1e80102_app2.docx]

References

1. World report on ageing and health. World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789241565042 [accessed 2025-12-24]
2. McNicoll G. World population ageing 1950-2050. 2002. URL: https://link.springer.com/rwe/10.1007/978-981-99-7842-7_154 [accessed 2025-12-24]
3. Gaudenz C, De Geest S, Schwendimann R, Zúñiga F. Factors associated with care workers' intention to leave employment in nursing homes: a secondary data analysis of the Swiss nursing homes human resources project. J Appl Gerontol 2019 Nov;38(11):1537-1563. [doi: 10.1177/0733464817721111] [Medline: 28715925]
4. Leiter M, Maslach C. Nurse turnover: the mediating role of burnout. J Nurs Manag 2009 Apr;17(3):331-339. [doi: 10.1111/j.1365-2834.2009.01004.x] [Medline: 19426369]
5. Yu M, Lee H. Impact of resilience and job involvement on turnover intention of new graduate nurses using structural equation modeling. Jpn J Nurs Sci 2018 Oct;15(4):351-362. [doi: 10.1111/jjns.12210] [Medline: 29508523]

XSL•FO
RenderX

6. Lu M, Zhang F, Tang X, Wang L, Zan J, Zhu Y, et al. Do type A personality and neuroticism moderate the relationships of occupational stressors, job satisfaction and burnout among Chinese older nurses? A cross-sectional survey. BMC Nurs 2022;21(1):88 [FREE Full text] [doi: 10.1186/s12912-022-00865-7] [Medline: 35428288]
7. Siciliano P, Burrage R. Chronic pain in the elderly: a continuing education program for certified nursing assistants. Geriatr Nurs 2005;26(4):252-258. [doi: 10.1016/j.gerinurse.2005.05.008] [Medline: 16109299]
8. Yang Q, Yi R, Wang N, Zhang P. Perception, behavior and experience of nursing assistants towards pain of older adults with dementia: a qualitative study. Geriatr Nurs 2024;56:100-107. [doi: 10.1016/j.gerinurse.2024.02.013] [Medline: 38340431]
9. Holloway K, McConigley R. Understanding nursing assistants' experiences of caring for older people in pain: the Australian experience. Pain Manag Nurs 2009 Jun;10(2):99-106. [doi: 10.1016/j.pmn.2008.10.001] [Medline: 19481049]
10. Dobbs D, Baker T, Carrion IV, Vongxaiburana E, Hyer K. Certified nursing assistants' perspectives of nursing home residents' pain experience: communication patterns, cultural context, and the role of empathy. Pain Manag Nurs 2014 Mar;15(1):87-96. [doi: 10.1016/j.pmn.2012.06.008] [Medline: 24602428]
11. Halifax E, Miaskowski C, Wallhagen M. Certified nursing assistants' understanding of nursing home residents' pain. J Gerontol Nurs 2018;44(4):29-36. [doi: 10.3928/00989134-20180131-01] [Medline: 29437185]
12. Jansen BDW, Brazil K, Passmore P, Buchanan H, Maxwell D, McIlfatrick SJ, et al. Exploring healthcare assistants' role and experience in pain assessment and management for people with advanced dementia towards the end of life: a qualitative study. BMC Palliat Care 2017;16(1):6 [FREE Full text] [doi: 10.1186/s12904-017-0184-1] [Medline: 28103847]
13. Lohne V, Høy B, Lillestø B, Sæteren B, Heggestad A, Aasgaard T, et al. Fostering dignity in the care of nursing home residents through slow caring. Nurs Ethics 2017 Nov;24(7):778-788. [doi: 10.1177/0969733015627297] [Medline: 26850071]
14. Moilanen T, Kangasniemi M, Papinaho O, Mynttinen M, Siipi H, Suominen S, et al. Older people's perceived autonomy in residential care: an integrative review. Nurs Ethics 2021;28(3):414-434 [FREE Full text] [doi: 10.1177/0969733020948115] [Medline: 33000683]
15. Preshaw DH, Brazil K, McLaughlin D, Frolic A. Ethical issues experienced by healthcare workers in nursing homes: literature review. Nurs Ethics 2016;23(5):490-506. [doi: 10.1177/0969733015576357] [Medline: 25870176]
16. Goethals S, Gastmans C, de Casterlé BD. Nurses' ethical reasoning and behaviour: a literature review. Int J Nurs Stud 2010;47(5):635-650. [doi: 10.1016/j.ijnurstu.2009.12.010] [Medline: 20096413]
17. Woods B, Pratt R. Awareness in dementia: ethical and legal issues in relation to people with dementia. Aging Ment Health 2005;9(5):423-429. [doi: 10.1080/13607860500143125] [Medline: 16024401]
18. Heggestad AKT, Nortvedt P, Slettebø Å. Dignity and care for people with dementia living in nursing homes. Dementia (London) 2015;14(6):825-841. [doi: 10.1177/1471301213512840] [Medline: 24381212]
19. Antwi YA, Bowblis JR. The impact of nurse turnover on quality of care and mortality in nursing homes: evidence from the great recession. Am J Health Econ 2018;4(2):131-163. [doi: 10.1162/ajhe_a_00096]
20. Castle N, Engberg J, Men A. Nursing home staff turnover: impact on nursing home compare quality measures. Gerontologist 2007;47(5):650-651. [doi: 10.1093/geront/47.5.650] [Medline: 17989407]
21. Zhang Z, Wang S, Miao C, Su M, Chen T, Yao N. Substantial differences in turnover intention between direct care workers in chinese hospitals and long-term care facilities. J Am Med Dir Assoc 2021;22(3):696-700. [doi: 10.1016/j.jamda.2020.09.006] [Medline: 33097399]
22. Krein SL, Turnwald M, Anderson B, Maust DT. "Sometimes it's not about the money... it's the way you treat people...": a qualitative study of nursing home staff turnover. J Am Med Dir Assoc 2022;23(7):1178-1184 [FREE Full text] [doi: 10.1016/j.jamda.2021.11.036] [Medline: 34990586]
23. Midje HH, Nyborg VN, Nordsteien A, Øvergård KI, Brembo EA, Torp S. Antecedents and outcomes of work engagement among nursing staff in long-term care facilities—a systematic review. J Adv Nurs 2024;80(1):42-59. [doi: 10.1111/jan.15804] [Medline: 37519065]
24. Heiks C, Sabine N. Long term care and skilled nursing facilities. Dela J Public Health 2022;8(5):144-149 [FREE Full text] [doi: 10.32481/djph.2022.12.032] [Medline: 36751604]
25. Lim MD, Connie T, Goh MKO, Saedon N. Model-Based Feature Extraction and Classification for Parkinson Disease Screening Using Gait Analysis: Development and Validation Study. JMIR Aging 2025;8:e65629. [doi: 10.2196/65629]
26. Zheng X, Zeng Z, S van Schooten K, Yang Y. Machine Learning Approach for Frailty Detection in Long-Term Care Using Accelerometer-Measured Gait and Daily Physical Activity: Model Development and Validation Study. JMIR Aging 2025;8:e77140. [doi: 10.2196/77140]
27. Shu X, Zhang L, Sun Y, Tang J. Host–parasite: Graph LSTM-in-LSTM for group activity recognition. IEEE Trans Neural Netw Learn Syst 2021;32(2):663-674. [doi: 10.1109/tnnls.2020.2978942]
28. Li C, Zhong Q, Xie D, Pu S. Collaborative spatiotemporal feature learning for video action recognition. 2019 Presented at: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); June 15-20, 2019; Long Beach, CA, USA. [doi: 10.1109/cvpr.2019.00806]
29. Shu X, Tang J, Qi G, Liu W, Yang J. Hierarchical long short-term concurrent memory for human interaction recognition. 2021 Presented at: IEEE Transactions on Pattern Analysis and Machine Intelligence; March 1, 2021; Washington, DC p. 1110-1118. [doi: 10.1109/tpami.2019.2942030]

XSL•FO
RenderX

30. Zhang P, Lan C, Xing J, Zeng W, Xue J, Zheng N. View adaptive neural networks for high performance skeleton-based human action recognition. IEEE Trans Pattern Anal Mach Intell 2019;41(8):1963-1978. [doi: 10.1109/tpami.2019.2896631]
31. Cheng K, Zhang Y, He X, Chen W, Cheng J, Lu H. Skeleton-based action recognition with shift graph convolutional network. 2025 Presented at: In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; June 13-19, 2020; Seattle, WA p. 183-192. [doi: 10.1109/cvpr42600.2020.00026]
32. Shu X, Zhang L, Qi G, Liu W, Tang J. Spatiotemporal co-attention recurrent neural networks for human-skeleton motion prediction. IEEE Trans Pattern Anal Mach Intell 2022;44(6):3300-3315. [doi: 10.1109/tpami.2021.3050918]
33. Ni B, Wang G, Moulin P. RGBD-HuDaAct: a color-depth video database for human daily activity recognition. : IEEE; 2011 Presented at: IEEE International Conference on Computer Vision Workshops (ICCV Workshops); November 6-13, 2011; Barcelona p. 1147-1153. [doi: 10.1109/iccvw.2011.6130379]
34. Oreifej O, Liu Z. HON4D: histogram of oriented 4D normals for activity recognition from depth sequences. : IEEE; 2013 Presented at: 2013 IEEE Conference on Computer Vision and Pattern Recognition; June 23-28, 2018; Portland, OR, USA p. 716-723. [doi: 10.1109/cvpr.2013.98]
35. Wang J, Liu Z, Wu Y, Yuan J. Mining actionlet ensemble for action recognition with depth cameras. : IEEE; 2012 Presented at: 2012 IEEE Conference on Computer Vision and Pattern Recognition; June 16-21, 2012; Providence, RI, USA p. 1290-1297. [doi: 10.1109/cvpr.2012.6247813]
36. Liu J, Shahroudy A, Perez M, Wang G, Duan L, Kot AC. NTU RGB+D 120: a large-scale benchmark for 3D human activity understanding. IEEE Trans Pattern Anal Machine Intell 2020;42(10):2684-2701. [doi: 10.1109/tpami.2019.2916873]
37. Chen C, Jafari R, Kehtarnavaz N. UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor. 2015 Dec 10 Presented at: 2015 IEEE International Conference on Image Processing (ICIP); September 27-30, 2015; Quebec City, QC, Canada p. 168-172. [doi: 10.1109/icip.2015.7350781]
38. Shu X, Tang J, Qi GJ, Song Y, Li Z, Zhang L. Concurrence-aware long short-term sub-memories for person-person action recognition. : IEEE; 2017 Aug 24 Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); July 21-26, 2017; Honolulu, HI, USA p. 1-8. [doi: 10.1109/cvprw.2017.270]
39. Yan R, Tang J, Shu X, Li Z, Tian Q. Participation-contributed temporal dynamic model for group activity recognition. 2018 Oct 15 Presented at: Proceedings of the 26th ACM International Conference on Multimedia; 2018; New York, NY, USA p. 1292-1300. [doi: 10.1145/3240508.3240572]
40. Yan R, Xie L, Tang J, Shu X, Tian Q. HiGCIN: hierarchical graph-based cross inference network for group activity recognition. IEEE Trans Pattern Anal Mach Intell 2023;45(6):6955-6968. [doi: 10.1109/TPAMI.2020.3034233] [Medline: 33108281]
41. Song K, Wei X, Shu X, Song R, Lu J. Bi-Modal progressive mask attention for fine-grained recognition. IEEE Trans Image Process 2020;29:7006-7018. [doi: 10.1109/tip.2020.2996736]
42. Jiang H, Song Y, He J, Shu X. Cross fusion for egocentric interactive action recognition. : Springer International Publishing; 2020 Presented at: MultiMedia Modeling: 26th International Conference, MMM 2020, Proceedings, Part I; January 5-8, 2020; Daejeon, South Korea p. 714-726. [doi: 10.1007/978-3-030-37731-1_58]
43. Simonyan K. Two-stream convolutional networks for action recognition in videos. In Proceedings of the 28th International Conference on Neural Information Processing Systems 2014;Volume 1:568-576.
44. Crasto N, Weinzaepfel P, Alahari K, Schmid C. MARS: Motion-augmented RGB stream for action recognition. : IEEE; 2019 Jan 09 Presented at: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; June 15-20, 2019; Long Beach, CA p. 7882-7891. [doi: 10.1109/cvpr.2019.00807]
45. Diba A, Fayyaz M, Sharma V, Karami AH, Arzani MM, Yousefzadeh R, et al. Temporal 3d convnets: new architecture and transfer learning for video classification. ArXiv. Preprint posted online on November 22, 2017 2017.
46. Tran D, Ray J, Shou Z, Chang SF, Paluri M. Convnet architecture search for spatiotemporal feature learning. arXiv. Preprint posted online on August 16, 2017 2017. [doi: 10.48550/arXiv.1708.05038]
47. Carreira J, Zisserman A. Quo vadis, action recognition? A new model and the kinetics dataset. 2017 Nov 09 Presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; July 21-26, 2017; Honolulu, Hawaii, USA p. 6299-6308. [doi: 10.1109/cvpr.2017.502]
48. Tran D, Bourdev L, Fergus R, Torresani L, Paluri M. Learning spatiotemporal features with 3D convolutional networks. : IEEE; 2015 Presented at: IEEE International Conference on Computer Vision (ICCV); February 18, 2016; Santiago, Chile p. 4489-4497. [doi: 10.1109/iccv.2015.510]
49. Jegham I, Ben Khalifa A, Alouani I, Mahjoub MA. Vision-based human action recognition: an overview and real world challenges. Forensic Sci Int Digit Invest 2020;32:200901. [doi: 10.1016/j.fsidi.2019.200901] [Medline: 38667983]
50. Kong Y, Fu Y. Human action recognition and prediction: a survey. Int J Comput Vis 2022;130(5):1366-1401. [doi: 10.1007/s11263-022-01594-9]
51. Yan S, Xiong Y, Lin D. Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition. In: AAAI. 2018 Apr 27 Presented at: Proceedings of the AAAI Conference on Artificial Intelligence; February 2-7, 2018; New Orleans, Louisiana. [doi: 10.1609/aaai.v32i1.12328]

XSL•FO
RenderX

52. Shi L, Zhang Y, Cheng J, Lu H. Two-stream adaptive graph convolutional networks for skeleton-based action recognition. : IEEE; 2020 Jan 09 Presented at: IEEE/CVF Conference on Computer Vision and Pattern Recognition; June 15-20, 2019; Long Beach, CA, USA. [doi: 10.1109/cvpr.2019.01230]
53. Qin Z, Liu Y, Ji P, Kim D, Wang L, McKay RI, et al. Fusing higher-order features in graph neural networks for skeleton-based action recognition. IEEE Trans Neural Netw Learning Syst 2024;35(4):4783-4797. [doi: 10.1109/tnnls.2022.3201518]
54. Li M, Chen S, Chen X, Zhang Y, Wang Y, Tian Q. Actional-structural graph convolutional networks for skeleton-based action recognition. : IEEE; 2020 Jan 09 Presented at: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; June, 15-20 2019; Long Beach, CA, USA p. 3595-3603. [doi: 10.1109/cvpr.2019.00371]
55. Song Y, Zhang Z, Shan C, Wang L. Constructing stronger and faster baselines for skeleton-based action recognition. IEEE Trans Pattern Anal Mach Intell 2023;45(2):1474-1488. [doi: 10.1109/tpami.2022.3157033]
56. Ke L, Peng K, Lyu S. Towards To-a-T spatio-temporal focus for skeleton-based action recognition. 2022 Feb 04 Presented at: Proceedings of the AAAI conference on artificial intelligence; 2022 April 29; NA p. 1131-1139. [doi: 10.1609/aaai.v36i1.19998]
57. Si C, Chen W, Wang W, Wang L, Tan T. An attention enhanced graph convolutional LSTM network for skeleton-based action recognition. : IEEE; 2020 Jan 09 Presented at: The IEEE/CVF Conference on Computer Vision and Pattern Recognition; June 15-20, 2019; Long Beach, CA, USA p. 1227-1236. [doi: 10.1109/cvpr.2019.00132]
58. Zhang J, Ye G, Tu Z, Qin Y, Qin Q, Zhang J, et al. A spatial attentive and temporal dilated (SATD) GCN for skeleton‐based action recognition. CAAI Trans on Intel Tech 2021;7(1):46-55. [doi: 10.1049/cit2.12012]
59. Liu G, Qian J, Wen F, Zhu X, Ying R, Liu P. Action recognition based on 3D skeleton and RGB frame fusion. : IEEE; 2020 Jan 28 Presented at: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); November 03-08, 2019; Macau, China p. 258-264. [doi: 10.1109/iros40897.2019.8967570]
60. Das S, Koperski M, Bremond F, Francesca G. Action recognition based on a mixture of RGB and depth based skeleton. : IEEE; 2017 Oct 23 Presented at: 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS); August 29, 2017 to September 01, 2017; Lecce, Italy p. 1-6. [doi: 10.1109/avss.2017.8078548]
61. Chan A. Feature fusion based human action recognition in still images. Intl J Comput Sci Netw Secur 2019;19(11):151.
62. Huan R, Jiang C, Ge L, Shu J, Zhan Z, Chen P, et al. Human complex activity recognition with sensor data using multiple features. IEEE Sensors J 2022;22(1):757-775. [doi: 10.1109/jsen.2021.3130913]
63. Guo J, Shi M, Zhu X, Huang W, He Y, Zhang W, et al. Improving human action recognition by jointly exploiting video and WiFi clues. Neurocomputing 2021;458:14-23. [doi: 10.1016/j.neucom.2020.11.074]
64. Kim K, Jalal A, Mahmood M. Vision-based human activity recognition system using depth silhouettes: a smart home system for monitoring the residents. J Electr Eng Technol 2019;14(6):2567-2573. [doi: 10.1007/s42835-019-00278-8]
65. Nagpal D, Gupta S, Kumar D, Illés Z, Verma C, Dey B. goldenAGER: a personalized feature fusion activity recognition model for elderly. IEEE Access 2023;11:56766-56784. [doi: 10.1109/access.2023.3282439]
66. Girdhar R, Carreira J, Doersch C, Zisserman A. Video action transformer network. : IEEE; 2020 Jan 09 Presented at: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; June 15-20, 2019; Long Beach, CA, USA p. 244-253. [doi: 10.1109/CVPR.2019.00033]
67. Wang X, Zhang S, Qing Z, Shao Y, Zuo Z, Gao C, et al. Oadtr: Online action detection with transformers. : IEEE; 2021 Presented at: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV); October 10-17, 2021; Montreal, QC, Canada p. 7565-7575. [doi: 10.1109/iccv48922.2021.00747]
68. Mazzia V, Angarano S, Salvetti F, Angelini F, Chiaberge M. Action transformer: a self-attention model for short-time pose-based human action recognition. Pattern Recognit 2022 Apr;124:108487. [doi: 10.1016/j.patcog.2021.108487]
69. Plizzari C, Cannici M, Matteucci M. Skeleton-based action recognition via spatial and temporal transformer networks. Comput Vision Image Understanding 2021 Jul;208-209:103219. [doi: 10.1016/j.cviu.2021.103219]
70. Wang X. EGOFALLS: a visual-audio dataset and benchmark for fall detection using egocentric cameras. Switzerland: Springer Nature; 2025 Presented at: Proceedings of the 27th International Conference on Pattern Recognition; December 1, 2024 to December 5, 2024; Kolkata, India p. 240-253. [doi: 10.1007/978-3-031-78166-7_16]
71. Dai R, Das S, Sharma S, Minciullo L, Garattoni L, Bremond F, et al. Toyota Smarthome Untrimmed: Real-World Untrimmed Videos for Activity Detection. IEEE Trans. Pattern Anal. Mach. Intell 2023;45(2):2533-2550. [doi: 10.1109/tpami.2022.3169976]
72. Miron A, Sadawi N, Ismail W, Hussain H, Grosan C. IntelliRehabDS (IRDS)—a dataset of physical rehabilitation movements. Data 2021;6(5):46. [doi: 10.3390/data6050046]
73. Jang J, Kim D, Park C, Jang M, Lee J, Kim J. ETRI-Activity3D: a large-scale RGB-D dataset for robots to recognize daily activities of the elderly. : IEEE; 2021 Feb 10 Presented at: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); October 24, 2020 to January 24, 2021; Las Vegas, NV, USA p. 10990-10997. [doi: 10.1109/iros45743.2020.9341160]
74. AXIS Communications. URL: https://www.axis.com/dam/public/7d/1f/55/datasheet-axis-m3048-p-network-camera-ja-JP-287941.pdf [accessed 2025-12-24]
75. Matsuda S, Muramatsu K, Hayashida K. Eligibility Classification Logic of the Japanese Long Term Care Insurance. Asian Pacific Journal of Disease Management 2011;5(3):65-74. [doi: 10.7223/apjdm.5.65]

XSL•FO
RenderX

76. Wang C. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. 2023 Presented at: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp.); 22 Aug 2023; Vancouver, BC, Canada p. 7464-7475. [doi: 10.1109/cvpr52729.2023.00721]
77. Redmon J. ou only look once: Unified, real-time object detection. 2016 Presented at: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 12 Dec 2016; Las Vegas, NV, USA p. 779-788. [doi: 10.1109/cvpr.2016.91]
78. Zhao H. Unleashing the Power of Large-Scale Unlabeled Data. Geneva: Springer; 2024:2024-2021.

Abbreviations

AUC: area under the curve
AWM: attention weight matrix
CAM: care-aware attention mechanism
CAS: care assessment score
CL: care level
CS: cosine similarity
CSTT: care-assessment-aware spatiotemporal transformer
DL: deep learning
HAR: human activity recognition
HOG: histogram of oriented gradients
LTC: long-term care
MPJAD: mean per joint angle difference
YOLO: You Only Look Once

Edited by Y Jiang; submitted 04.Jul.2025; peer-reviewed by H Sun; comments to author 01.Aug.2025; revised version received 10.Dec.2025; accepted 16.Dec.2025; published 02.Apr.2026.

Please cite as:
Nahid N, Hassan I, Ahad MAR, Inoue S
Integrating Care Context With Skeleton and Depth Information for Older Adult Activity Recognition in a Care Facility Using Care-Assessment-Aware Spatiotemporal Transformer: Method and Validation Study
JMIR Aging 2026;9:e80102
URL: https://aging.jmir.org/2026/1/e80102
doi:10.2196/80102
PMID:

©Nazmun Nahid, Iqbal Hassan, Md Atiqur Rahman Ahad, Sozo Inoue. Originally published in JMIR Aging (https://aging.jmir.org), 02.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Ways in Which Stakeholders Make Decisions About AI and Novel Technologies for the Health Care of Older Adults: Qualitative Interview Study

Zhang Zhang[1,2], MSc, PhD; Thomas KM Cudjoe[3], MA, MPH, MD; Sato Ashida[4], PhD; Jacqueline Massare[3], BS; Kacey Chae[5,6], MD; Phillip Phan[3,7,8], PhD; Peter Abadir[3], MD; Alicia I Arbaje[1,3], MPH, MD, PhD; Mathias Unberath[9], PhD; Nancy L Schoenborn[3], MHS, MD

[1]Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, NC, United States

[2]Bursky School of Public Health, Washington University in St. Louis, St. Louis, MO, United States

[3]Division of Geriatric Medicine and Gerontology, Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[4]Department of Community and Behavioral Health, College of Public Health, University of Iowa, Iowa City, IA, United States

[5]Division of General Internal Medicine, Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[6]Bariatric and Weight Management, Surgery, School of Medicine, Wake Forest University, Winston Salem, NC, United States

[7]Carey Business School, Johns Hopkins University, Baltimore, MD, United States

[8]Johns Hopkins Armstrong Institute for Patient Safety and Quality, Johns Hopkins University, Baltimore, MD, United States

[9]Department of Computer Science, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Nancy L Schoenborn, MHS, MD
Division of Geriatric Medicine and Gerontology, Department of Medicine
School of Medicine
Johns Hopkins University
5200 Eastern Avenue, MFL Center Tower Room 703
Baltimore, MD, 21224
United States
Phone: 1 410 550 7142
Email: nancyschoenborn@jhmi.edu

Abstract

Background: Artificial intelligence (AI) has the potential to improve health among older adults, yet how different stakeholders decide to develop, finance, and adopt AI innovations is not well understood.

Objective: This study aimed to understand the decision-making of different stakeholders regarding AI and technologies for the health care of older adults.

Methods: We conducted semistructured interviews with 15 older adults and care partners, 15 clinicians, 8 health system or insurance leaders, 5 investors, and 6 technology developers. Data were analyzed using thematic content analysis.

Results: All stakeholders considered cost, value, and usability important in adopting AI health technologies but emphasized different aspects of each concept. Older adults and care partners prioritized out-of-pocket costs and ease of use, whereas payers emphasized disease prevalence and implementation feasibility. Developers and investors focused on profitability and scalability, resulting in tension with end users' priorities. Participant suggestions included problem-driven design, greater stakeholder engagement, public-private partnerships, and educating older adults about AI.

Conclusions: As with prior health-related technology innovations, aligning decisional priorities across stakeholders is critical to motivate impactful AI health technologies for older adults.

(JMIR Aging 2026;9:e86148) doi:10.2196/86148

XSL•FO
RenderX

KEYWORDS

artificial intelligence; AI; novel technology; older adults; decision-making; artificial intelligence adoption; AI adoption; usability; technology acceptance model; perceived usefulness

Introduction

AI and novel technologies offer tremendous potential for improving the health of older adults. They are becoming an important focus for technologies that facilitate aging in place, reduce care partner burden, and improve health system efficiency [1-3]. AI, which uses various approaches to simulate human intelligence and behavior, has increasingly been used in health care, often paired with other digital health tools such as wearable sensors, smart home devices, virtual assistants, and robotics [4,5]. For example, wearable sensors combined with AI algorithms can monitor functional status, sleep patterns, and mobility to detect early signs of cognitive or physical decline in community-dwelling older adults [6-8]. AI-enabled robotics with autonomous navigation, multimodal sensing, and assistive manipulation capabilities improve health care efficiency by continuously monitoring patient function, supporting clinical workflows, and reducing preventable hospitalizations for older adults [9-11]. Despite these advancements, the broad adoption of such technologies remains limited.

The literature has examined facilitators of and barriers to general technological adoption by end users, including older adults, care partners, and clinicians [12-15]. These factors, such as perceived usefulness, ease of use, trust, cost, privacy concerns, and social support, are key determinants of technology acceptance in older adult populations [16-19]. The findings have been synthesized in conceptual frameworks that describe the key influences in technology adoption. A commonly used framework is the technology acceptance model (TAM) [20]. The TAM proposes that individuals' acceptance and use of a technology are primarily determined by 2 key factors: perceived usefulness and perceived ease of use [16]. Subsequent extensions of the TAM have also emphasized additional contextual factors, including trust, cost, privacy, and organizational support, that may influence technology adoption in health care settings [14,16]. Some recent studies have examined attitudes toward and acceptance of technology that can be paired with AI, such as wearable devices and robots [12,14,21,22]. However, we did not find studies that specifically examined how key stakeholders made decisions about AI health technologies designed to improve the care of older adults. We also did not find perspectives on how health system leaders and payers consider technology adoption and how investors and developers consider the decision of what technological products to develop and finance. As health system payers, investors, and developers are responsible for making novel health technologies available, understanding their decision-making in conjunction with that of the end users is critically important.

To address the above-mentioned knowledge gaps, we conducted a qualitative study to explore the perspectives of different stakeholder groups—older adults, care partners, clinicians, payers, investors, and technology developers—regarding their decision-making on the development, financing, and adoption of AI and novel technologies to improve the health of older adults. We aimed to identify areas of consensus and potential differences to help understand how to align stakeholders' key drivers of decision-making to accelerate the implementation of impactful technologies for the health and well-being of older adults.

Methods

Design, Study Setting, and Participant Recruitment

We chose to use qualitative semistructured individual interviews to explore stakeholder perspectives. This project is part of a larger study in which we also explored stakeholder perspectives on priority areas for AI applications, perceived benefits of and concerns about AI, and educational needs; these results have been reported previously [23,24]. In this paper, we report stakeholder perspectives on decision-making on the development, financing, and adoption of AI applications in the care of older adults. This study focused on stakeholders, including older adults (defined as those aged 60 years or older) and care partners for older adults (via self-report); clinicians; health system or health insurance leaders involved in technology adoption, such as the chief information officer (hereafter referred to as "payers"); and investors and technology developers. We combined older adults and care partners in the same group as many older adults were or had been care partners.

This project was supported by the Johns Hopkins Artificial Intelligence and Technology Collaboratory (JHU-AITC) for Aging Research [25]. The Stakeholder Engagement Core within the JHU-AITC engages longitudinally with a stakeholder council of older adults, care partners, and clinicians from Maryland and Iowa. We recruited older adults, care partners, and clinicians via referrals from the study team and council members; we recruited payers, investors, and technology developers via referrals from the study team, JHU-AITC colleagues, and the coordinating center that supports artificial intelligence and technology collaboratories. We also used snowball sampling to recruit contacts of existing participants. Recruitment continued until information saturation was reached.

All participants needed to be English speaking and be able to provide informed consent. Each participant was offered a US $50 gift card.

Interview Guide

The interview guide was developed by the study team, which included expertise from engineering, business development, patient engagement, geriatric medicine, and social and behavioral science. We defined AI as "technologies where machines try to do what human beings do" and described that there were many types of AI applications, such as alerts, sensors, prediction algorithms, or robots. For older adults, care partners, clinicians, and payers, we asked how they decided whether to use an AI product; for investors, we asked how they decided to invest in an AI product; for developers, we asked how they made decisions on what type of AI product to work on or move

XSL•FO
RenderX

forward with in development. The full interview guide is available in Multimedia Appendix 1.

Data Collection and Analysis

Three team members (NLS, KC, and JM) conducted interviews in person (at private workspaces) or virtually from May 2023 to January 2024. All 3 team members had prior qualitative research experience. One team member (NLS) had prior working relationships with some of the clinician participants.

Interviews were audio recorded and transcribed verbatim; the average duration of the interviews was 34 (SD 11) minutes. We collected demographic information from all participants. For older adults and care partners, we collected information about education, financial strain, health literacy, self-rated health, and duration of caregiving. For clinicians, we asked about their practice setting. For all other participants, we asked about their professional roles and the duration working in that role.

We analyzed interview transcripts using reflexive thematic analysis informed by the approach by Braun and Clarke [26] to identifying, analyzing, and reporting patterns of meaning across qualitative data. This method was appropriate because our study aimed to examine how different stakeholder groups described decision-making on the development, financing, and adoption of AI health technologies for older adults. Consistent with the guidance from Braun and Clarke [26] on good practice in thematic analysis, we used an iterative analytic process that emphasized methodological clarity; reflexive engagement with the data; and alignment among the research questions, analytic procedures, and resulting themes.

To code the transcripts, we used thematic analysis with an integrated approach that used both inductive coding and a deductive organizing framework that drew conceptually from the TAM [16,20]. Specifically, a codebook was developed based on the interview guide and revised iteratively to add new codes generated inductively based on the review of 9 randomly selected transcripts. We organized the codes, when applicable, by domains in the TAM and compared our findings to the TAM during the interpretation of the results. Each transcript was independently coded using the finalized codebook by at least 2 investigators (NLS, KC, and JM), and differences were reconciled through consensus. The investigators involved in the analysis had training in clinical general or geriatric medicine, health service research, and/or qualitative research. Results were summarized into major themes and illustrated via representative quotes. Analysis used the ATLAS.ti software (version 24; ATLAS.ti Scientific Software Development GmbH).

Reflexivity was addressed throughout data collection and analysis. The study team included members with expertise in engineering, business development, patient engagement, geriatric medicine, and social and behavioral science, which allowed for multiple disciplinary perspectives to inform interpretation. At the same time, our team recognized that these professional backgrounds, as well as prior relationships with some participants, could shape data collection and interpretation. To address this, the team discussed emerging interpretations, considered alternative explanations, and examined how researchers' assumptions and disciplinary perspectives may have influenced theme development.

Several strategies were used to enhance trustworthiness. Credibility was supported through independent coding by multiple investigators and consensus discussions using transcripts. Dependability was strengthened by maintaining documentation of analytic decisions, including codebook revisions, coding discussions, and theme refinement. Confirmability was supported through team-based reflexive discussion and by grounding interpretations in the data rather than a priori assumptions. Transferability was supported by describing the study sample, stakeholder groups, recruitment approach, interview procedures, and analytic process in sufficient detail to allow readers to assess the relevance of the findings to other settings and populations.

Ethical Considerations

This project was approved by the Johns Hopkins Medicine Institutional Review Board (IRB00356437). The reporting of this study followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) reporting guidelines [27]. Consent to participate was obtained for this project, and it adheres to human ethics guidelines. This study used deidentified data and posed minimal risk to participants.

Results

Participants

The 49 participants included 15 (30.6%) older adults or care partners, 15 (30.6%) clinicians, 8 (16.3%) payers, 5 (10.2%) investors, and 6 (12.2%) technology developers (Table 1). The older adult group had a mean age of 71.3 years; 60% (9/15) of the older adults reported caregiving experience. Clinicians comprised physicians, nurses, pharmacists, rehabilitation therapists, and a dentist. Most of the older adult or care partner participants (14/15, 93.3%) and all the clinician participants (15/15, 100%) were from Maryland and Iowa, whereas payer, investor, and technology developer participants were from 9 different states.

XSL•FO
RenderX

Table 1. Participant characteristics.

	Older adults and care partners (n=15)	Clinicians (n=15)	Payers (n=8)	Investors and technology developers (n=11)
Age (y), mean (range), SD	71.3 (8.6; 65-93)	50.3 (11.3; 33-69)	51.6 (10.9; 36-65)	42.3 (16.7; 27-62)
Women, n (%)	11 (73.3)	7 (46.7)	3 (37.5)	0 (0)
Race, n (%)				
Asian	1 (6.7)	3 (20)	0 (0)	6 (54.5)
Black	3 (20)	0 (0)	0 (0)	1 (9.1)
Hispanic	0 (0)	1 (6.7)	0 (0)	0 (0)
White	11 (73.3)	11 (73.3)	8 (100)	4 (36.4)
Location, n (%)				
Maryland	10 (66.7)	8 (53.3)	2 (25)	3 (27.3)
Iowa	4 (26.7)	7 (46.7)	2 (25)	0 (0)
Other states	1 (6.7)[a]	0 (0)	4 (50)[b]	8 (72.7)[c]
Past or current care partner, n (%)	9 (60)	—[d]	—	—
Duration as care partner (y), mean (range)	9.5 (0.5-25)	—	—	—
Educational level, n (%)				
College degree or higher	3 (20)	—	—	—
High school or <4-y college	11 (73.3)	—	—	—
Missing data	1 (6.7)	—	—	—
Self-reported financial strain[e], n (%)	3 (20)	—	—	—
Low health literacy[f], n (%)	3 (20)	—	—	—
Self-rated health, n (%)				
Excellent or very good	6 (40)	—	—	—
Good	4 (26.7)	—	—	—
Fair or poor	4 (26.7)	—	—	—
Missing	1 (6.7)	—	—	—
Clinician practice setting (could choose more than one), n (%)				
Inpatient	—	10 (66.7)	—	—
Outpatient	—	5 (33.3)	—	—
Nursing home	—	3 (20)	—	—
Home-based care	—	4 (26.7)	—	—
Duration in profession (y), mean (SD)	—	19.3 (11.1)	13.3 (9.5)	9.9 (8.5)

[a]Washington, DC.

[b]Colorado, Florida, Massachusetts, and Minnesota.

[c]California, Massachusetts, Virginia, and Washington, DC.

[d]Not asked for these participants.

[e]Financial strain was assessed using the following question: “During the last year, were there any times when you did not have enough money to pay rent or utilities or medical bills?” The response “yes” was counted as having self-reported financial strain [20].

[f]Low health literacy was assessed using the following question: “How confident are you in filling out medical forms by yourself?” Responses of “somewhat,” “a little bit,” and “not at all” were counted as low health literacy [14].

Themes

Three themes emerged from the interviews. First, stakeholders across groups identified cost, usability, and value as key considerations when making decisions about AI health technologies for older adults, although the meaning of these concepts differed by stakeholder group. Second, these differing

XSL•FO
RenderX

interpretations contributed to tensions between end users (older adults, care partners, and clinicians) and those responsible for developing and financing technologies (investors and developers), particularly regarding affordability and the design process. Third, participants described several potential strategies to better align stakeholder priorities and support the adoption of AI technologies for older adults. We describe each of the 3 themes below in detail.

Theme 1: The Value of AI Meant Different Things to Different Stakeholders

Overview

We found that all key stakeholders viewed cost, added value, and usability as important when considering AI health technologies for older adults, but each stakeholder group had different considerations within each concept. Although cost, value, and usability have been previously described as important to technology adoption in the literature, here, we contextualize each by stakeholder group and highlight that the same words encompass vastly different considerations (Table 2).

Table 2. Perspectives across stakeholder groups on cost, value, and usability when making decisions about AI health technologies for older adults.

	Consideration	Illustrative quote
Cost and value		
Older adults and care partners	Patient's out-of-pocket cost	"Price would be a barrier." [Older adult 5]
Clinicians	Patient's out-of-pocket cost	"Many older adults are on fixed incomes.... The group of older adults that probably need what AI might be able to deliver are just not going to have finance to do that." [Clinician 1]
Payers	Health system cost for purchasing, integrating, and maintaining the technology relative to health system savings from improvement in clinical outcomes	"...what are the things that improve quality outcomes and decrease cost." [Payer 2]
Investors and technology developers	Return on investment	"Bluntly, can you make money off of this?" [Investor 1]
Usability		
Older adults and care partners	Ease of use for older adults	"It has to be user-friendly. It has to be simple...accommodate slow fingers, eyes that don't see so well and ears that don't hear so well." [Older adult 4]
Clinicians	Ease of use for older adults and for clinicians	"If it's too complex then it makes it harder for our staff to use it and it becomes cumbersome." [Clinician 9]
Payers	Compatibility with existing health system	"It's the integration into the workflow and the health system that's important.... We were looking at a product that...I said, 'come back to us when [it'll] exist inside our electronic health record.' Because it just doesn't work well with the workflow." [Payer 8]
Investors and technology developers	Usability needs across other stakeholder groups	"You have to think like a doctor...about the healthcare need and impact of your project...then there are the technical and engineering type of things you have to think about...there's also the business related things...commercialization, finance, competition...and the actual implementation." [Technology developer 5]

Subtheme 1a: Cost and Value

Older adults, care partners, and clinicians made similar comments focused on the patients' out-of-pocket costs, and most were concerned about affordability. However, some commented on being willing to take on the cost if the technology offered sufficient value; one older adult said the following:

> *What is the return on investment? I'd rather go without some food if this type of technology could help me.* [Older adult 1]

Payers considered cost and value as a balance among the upfront cost of the technology to the health system or insurance plan, the expected improvement in clinical outcomes, and the potential savings from those outcome improvements. As such, payers commented that they were more likely to invest in technology that targeted high-cost or high-frequency health events. One payer said the following:

> *Rare event, but very detrimental and/or costly to the health system...something we really want to try to avoid, [we'd] make the investment there. Something that maybe doesn't have on an individual basis a lot of impact but is very common like hypertension, [we'd] make the investment there.* [Payer 8]

Payers also considered indirect costs required for supporting the technology:

> *If I have technology that we've acquired that I then need to hire an employee to interface with. Even though [the technology] is reimbursed, do I get enough reimbursement to hire the people? People*

XSL•FO
RenderX

> *are expensive, so I really need a massive return on investment to justify it.* [Payer 8]

Investors and technology developers considered cost and value as the magnitude of return on investment, which in turn included consideration of the potential market size for a given technology and whether the revenue stream for the technology would be directly from consumers or from insurance reimbursement. One investor said the following:

> *Who's going to pay for this and is the market large enough?* [Investor 1]

Subtheme 1b: Usability

Similarly, all stakeholder groups agreed on the importance of technology being easy to use, but they had distinct ideas on what that meant (Table 2). Older adults and care partners commented on technology being user-friendly and adapted to the needs of older adults who may have physical, sensory, and/or cognitive impairments. One older adult said the following:

> *Artificial intelligence can do many things, but it must be as simple as possible, and it must be user-friendly. If there's an awful lot that has to be learned by the person dealing with the artificial intelligence, it defeats its purpose.* [Older adult 4]

Clinicians considered the technology's user-friendliness to patients as well as to clinicians, especially in light of the existing clinical demands. One clinician said the following:

> *It has to be really user-friendly for me to be able to use it.* [Clinician 9]

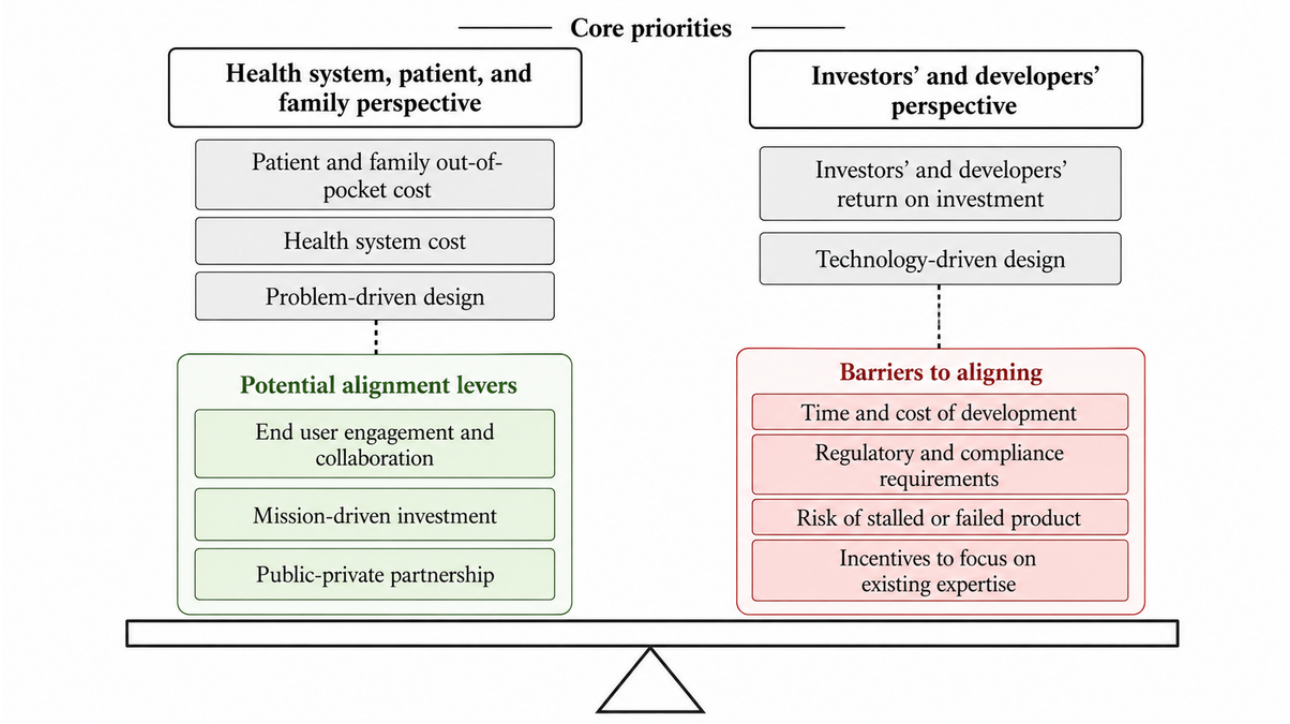
Another commented that "it has to be simple enough" (clinician 2), especially for clinicians who may already be burned out. Payers also commented on user-friendliness for clinicians but were more concerned about the ease of integration with existing systems. Finally, investors and developers commented on the complexity of trying to understand each stakeholder group's needs and expectations. One investor talked about an AI technology that was designed to help clinicians generate message responses to patients:

> *...it's like a quagmire of validation [with each group].... We talk to physicians...we talk to their staff...we talk to patients, do they like receiving automated AI response?... Then there's IT security issue.... We have to work with [the electronic health record vendor]..., and we are talking to payers to understand...if we were to bill for this, are you going to accept charges of an AI.* [Investor 1]

Theme 2: Tensions in Decision-Making Priorities

We observed tension between the decision-making priorities of the end users (eg, older adults, care partners, and clinicians) and those of the developers and investors. The decisional factors that were important to end users were at times in conflict with those that were important to investors and developers. Two areas of tension included the affordability of the technology and the design process (Figure 1).

Figure 1. Two identified tension areas between the decision-making priorities of end users and investors and developers were related to cost considerations and the design process. Suggested facilitators of aligning priorities are depicted in green, and existing barriers to aligning priorities are depicted in red.



As stated above, older adults, care partners, and clinicians valued low cost (ie, "costs should be kept as minimal as possible" [older adult 3]). On the other hand, investors and developers commented on multiple considerations in the development process that drive up cost due to the significant risk and the long time horizon involved in developing AI health technologies. Both investors and developers commented that direct-to-consumer products are much higher risk:

XSL•FO
RenderX

> *...consumer investments are very high risk, especially at early stages.* [Investor 4]

However, the alternative involves more regulatory cost and time:

> *...if you decide to go through the prescription route [where] you hope it becomes insured. The cost just balloons, it can be two, three, ten times more expensive...because the [medical device] industry is so conservative...you have to go through a lot of red tape to get it to market.... You are looking at 4 or 5 years to get from napkin sketch to through the FDA.* [Investor 4]

A developer commented the following on a recent publication looking at novel medical devices:

> *...after going through the FDA, which [is] very difficult, on average they had to wait 7 years for insurers to cover their technologies...and 44% didn't even get coverage at all.* [Technology developer 5]

The investment risks and the long wait, in turn, drive up cost:

> *...when you have odds like that for novel technologies...investors are going to ask for a big return because the risk is so high. You can't wait 10 years for a return on something without [the profit] being massive.... There are these huge pressures that drive cost up.* [Technology developer 5]

A second area of tension was regarding the design process for novel technologies. All stakeholder groups agreed that AI health applications should be designed after first understanding the target users and the health problem. However, many commented that, often, they saw the reverse—that existing technologies were offered to older adults even when they were not the most suitable solutions. One older adult said the following about health technology:

> *[It feels like] they are designing solutions in search of a problem.* [Older adult 1]

One engineer said the following:

> *I think the vast majority of engineers are coming from a place where we have a solution already and we think it might apply to this problem. Because that's by far easier.... It's very difficult to come from a purely problem-oriented perspective...even though I think that's ideal.* [Technology developer 5]

Another AI researcher commented on the incentive structure being a barrier to problem-driven design:

> *We have incentive structures to improve on what we already have in our areas of expertise, rather than pursuing new problems and things that may not work. For example, the geriatric population is not extremely well studied from an AI perspective.... It's much more risky for a researcher to start in the geriatric population and try to come up with a problem that fits AI than to take the AI and fit it to the geriatric population.* [Technology developer 2]

Theme 3: Promising Trends and Suggestions for Aligning Priorities

Subtheme 3a: Public Education About AI Health Technologies

Promising trends and suggestions can help enhance adoption and align decision-making priorities regarding AI health technology for older adults. Participants commented that older adults may not be sufficiently aware of AI health technology to recognize its value and that better promotion and public education are needed. One older adult said the following:

> *People need to do a much better job in educating the public on what AI really means and what AI can really do.* [Older adult 15]

An older adult suggested education through health care providers:

> *When they go into the doctor's office, they [can] have information about AI, how it can help.* [Older adult 10]

Another older adult talked about leveraging television shows to change cultural norms:

> *...a lot of people watch the soaps and regular TV, if those characters are using technology...it's going to make it more acceptable.* [Older adult 4]

One payer commented on more patient-friendly descriptions of AI:

> *The terminology of AI is not a natural, comfortable, or welcoming description for seniors. I think it's confusing, threatening, and it doesn't really get to the fact of helping you.... I would strongly recommend [improving] how we describe this new technology.* [Payer 4]

Subtheme 3b: Multi-Stakeholder Engagement and Problem-Driven Design Are Vital

Participants commented on the importance of engagement and collaboration to better understand the needs and expectations across stakeholder groups and approach technology design with a problem-driven mindset (Figure 1). All stakeholder groups commented on the importance of engaging older adults, care partners, and clinicians in the development of novel technologies. Several investors commented that they now used stakeholder engagement as a metric in evaluating potential projects:

> *...all good startup should be talking to their stakeholders before they even start developing a product...so we [ask]...how many people have you chatted with? What kind of feedback did they give?* [Investor 3]

Going beyond unilateral engagement from developers to end users, several participants also commented on the importance of collaboration and active participation from the public and clinicians. One payer who was also a clinician advocated for clinicians to be involved in technology development:

XSL•FO
RenderX

> *I wish more physicians realized the central place that software takes in our care, and realized that you can be just as much a physician designing a [health] application as you can on the wards.* [Payer 6]

Subtheme 3c: Public-Private Partnerships May Reduce Cost Barriers

Finally, participants commented on trends and suggestions to help align cost expectations for new technologies (Figure 1). One investor commented that some investors are willing to accept less profit because they are passionate about older adults and improving their health:

> *...there are...founders who don't care about the money part as much, and they are really passionate about a certain idea.... Maybe they had a family member that went through the same thing and they really want to solve this issue.* [Investor 3]

Another commented on public-private partnerships that can help align incentives:

> *...the economics may not be there to develop a solution for this...group, but if you can de-risk the research and development costs, [which] is the whole premise of the federal government funding, then it changes the dynamics for commercialization of those ideas.* [Investor 1]

Discussion

Principal Findings

This is the first study, to our knowledge, to examine decision-making related to AI technology development, financing, and adoption across diverse stakeholder groups. While previous research and frameworks such as the TAM have broadly acknowledged that cost, usability, and perceived value are significant determinants of technology acceptance by end users [28], this study adds depth by demonstrating how these concepts hold varied and nuanced meanings among stakeholder groups. In general, older adults and clinicians shared similar perspectives as end users, whereas investors and technology developers shared similar perspectives as those responsible for bringing about innovation; payers tended to have distinct perspectives from those of the other groups as they focused on broader, system-level concerns. The findings on the tension areas between decision-making priorities across stakeholder groups are also novel, drawing attention to factors in the existing technology development process that contribute to competing priorities.

How the different stakeholders viewed cost, value, and usability is consistent with what has been reported regarding other health technologies, but these have usually been reported *separately* and in different fields of study. For example, older adults, care partners, and clinicians primarily interpreted cost in terms of affordability, which has also been shown in studies on patient behavior regarding adoption of health care technologies [17,29]. Payers framed cost and value in terms of return on investment through reduced costs associated with high-frequency or high-severity clinical outcomes, a viewpoint common in the health care economics literature [30,31]. Investors and technology developers interpreted cost and value through the economic lens of market viability and profitability [32]. Our study design offering a juxtaposition of perspectives is a strength. A comparison of stakeholder perspectives has been described in the adoption of other technologies, such as patient portal use, but has focused on implementation without including the perspectives of developers or investors [18,33].

Recognizing that the same words and labels (eg, "cost" or "usability") represent different concepts to different stakeholders can help entrepreneurs and technology developers be more considerate and specific in the design process to meet stakeholder needs. More specific descriptions of these considerations may help highlight their importance. For example, instead of cost and usability, it may be more helpful for developers to recognize that older adults value products that are affordable and easy to use for individual users, whereas payers value technology that reduces health care costs and is easy to implement within an existing health care system. Similarly, funders of AI health technology research can evaluate proposals with metrics that are inclusive of different stakeholder viewpoints, for example, ensuring that a project addresses both individual usability and ease of health system–level implementation.

AI is unique in that the same underlying product can be used in multiple contexts. For example, large language models can be a direct-to-consumer tool (eg, ChatGPT) or a clinician-facing decision support system (eg, OpenEvidence). In contrast, prior health technologies such as patient portals have much more defined and stable target users. AI's adaptability means that the target user can shift after a product has already been developed and deployed. This has important implications for stakeholder alignment: the cost, usability requirements, accountability structures, and regulatory pathways may differ substantially depending on the target user. This flexibility creates both opportunity, allowing products to find their highest-value use case, and risk as it can obscure whose needs the technology was actually designed to serve.

Considering multiple perspectives together revealed critical tensions between end users (older adults, care partners, and clinicians) and technology creators (developers and investors). These tensions, particularly regarding affordability and design processes, signal significant barriers to wider adoption of patient-centered, impactful AI technologies for older adults. AI operates within a uniquely heterogeneous regulatory landscape where AI health applications can enter the market through multiple channels [34]. Some can be marketed directly to consumers as wellness products, which are subject to minimal regulatory oversight, whereas others may be considered medical devices that need to meet rigorous requirements before approval by the Food and Drug Administration [35]. Investors and developers highlighted factors in both pathways that contributed to tensions regarding affordability and design of AI health applications, including that direct consumer-facing technologies posed substantial financial risks whereas regulated products faced significant regulatory hurdles and lengthy approval timelines. The related incentive misalignments favor incremental enhancements to existing technologies over exploratory

XSL•FO
RenderX

problem-driven innovations addressing complex issues in older populations, which is a phenomenon that has also been reported in more traditional medical device development [35]. Indeed, the current health technology development infrastructure may very well inadvertently motivate high-cost products that are "solutions in search of a problem," and our results highlight the need for change.

Potential changes may include a revised or new regulatory and approval pathway for AI health technologies that is better tailored to the specific risks involved. Another change may involve altering the incentive structures, at least in research environments with a mission of discovery, to motivate problem-driven design and exploration of new ideas over honing existing expertise. We also identified actionable suggestions to facilitate adoption and align decision-making priorities from participants. Closer engagement across stakeholder groups throughout the development process was suggested by participants as one key solution to problem-driven design, which has also been reported in prior literature [12]. However, for older adults and clinicians to be interested and willing to engage, there first needs to be better public education and awareness campaigns to demystify AI and make clear its potential benefits. Participants suggested specific approaches for educating older adults about AI, such as through health care providers as well as media and cultural norms. Participants also suggested including stakeholder engagement as an explicit criterion for investment or funding, which can enhance accountability. Another suggestion by participants identified financial models such as public-private partnerships to balance cost considerations across stakeholders. The JHU-AITC is an example of one such partnership where public funding is awarded to promising pilot AI projects in very early development to help them advance to a stage where they can successfully attract private investment. Finally, it may be important to learn from other countries and health care models in terms of how they approach stakeholder engagement in AI health application development.

We acknowledge that our study has limitations. First, our participants were only restricted to English speakers, which may limit the generalizability of our findings to non-English speakers. Second, most older adult, caregiver, and clinician participants were based in Maryland and Iowa in the United States, which may limit geographic transferability. However, the inclusion of Iowa residents allowed us to obtain perspectives from rural residents and clinicians, where technology adoption is often delayed or limited. Third, several stakeholder groups lacked gender and racial diversity. For instance, all investor and technology developer participants were male. We did not assess prior experiences and familiarity with AI among our participants. Our future work should build on these findings through larger surveys with more diverse and representative samples across stakeholder groups.

Conclusions

In sum, we found that, across diverse stakeholder groups, the concepts of cost, usability, and value were considered universally important but had different meanings across different groups. This finding can help inform future research and technology development by highlighting that different elements of cost, usability, and value need to be separately considered and evaluated to meet the needs of diverse stakeholders. We also found competing decision-making priorities among different stakeholder groups when considering AI health applications for older adults, especially between those responsible for making new AI technologies available and the end users of such technologies. Implementing participant suggestions and other strategies to better align these priorities is essential for ensuring the widespread adoption of patient-centered, impactful AI solutions for older adults.

Funding

This study was supported by a grant from the Johns Hopkins Artificial Intelligence and Technology Collaboratory for Aging Research (P30AG073104). ZZ acknowledges funding from the National Institute on Aging (T32AG066576).

Authors' Contributions

ZZ was responsible for research design, data analysis, and writing of the manuscript. NLS and TKMC were responsible for study conception and design and data acquisition. KC and AIA conducted data preprocessing. SA, JM, PP, PA, and MU provided critical interpretation of the results. NLS supervised the overall project and secured funding. All authors contributed to the review and revision of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 31 KB - aging_v9i1e86148_app1.docx]

References

1. Loveys K, Prina M, Axford C, Domènec Ò, Weng W, Broadbent E, et al. Artificial intelligence for older people receiving long-term care: a systematic review of acceptability and effectiveness studies. Lancet Healthy Longev 2022 Apr;3(4):e286-e297 [FREE Full text] [doi: 10.1016/S2666-7568(22)00034-4] [Medline: 35515814]

XSL•FO
RenderX

2. Tian YJ, Felber NA, Pageau F, Schwab DR, Wangmo T. Benefits and barriers associated with the use of smart home health technologies in the care of older persons: a systematic review. BMC Geriatr 2024 Feb 14;24(1):152 [FREE Full text] [doi: 10.1186/s12877-024-04702-1] [Medline: 38355464]
3. Fakhimi MM, Hughes A, Gustavson AM. Human-centered design for smart home technologies: a framework for aging and mental health. Front Digit Health 2025 May 29;7:1555569 [FREE Full text] [doi: 10.3389/fdgth.2025.1555569] [Medline: 40510416]
4. Explanatory memorandum on the updated OECD definition of an AI system. Organisation for Economic Co-operation and Development. 2024 Mar 5. URL: https://www.oecd.org/en/publications/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_623da898-en.html [accessed 2026-06-27]
5. Recommendations on digital interventions for health system strengthening. World Health Organization. 2019 Jun 6. URL: https://www.who.int/publications/i/item/9789241550505 [accessed 2026-06-27]
6. Majumder S, Mondal T, Deen MJ. Wearable sensors for remote health monitoring. Sensors (Basel) 2017 Jan 12;17(1):130 [FREE Full text] [doi: 10.3390/s17010130] [Medline: 28085085]
7. Wu YH, Lewis M, Rigaud AS. Cognitive function and digital device use in older adults attending a memory clinic. Gerontol Geriatr Med 2019 May 2;5:2333721419844886 [FREE Full text] [doi: 10.1177/2333721419844886] [Medline: 31080848]
8. Lyons BE, Austin D, Seelye A, Petersen J, Yeargers J, Riley T, et al. Pervasive computing technologies to continuously assess Alzheimer's disease progression and intervention efficacy. Front Aging Neurosci 2015 Jun 10;7:102 [FREE Full text] [doi: 10.3389/fnagi.2015.00102] [Medline: 26113819]
9. Nicora G, Pe S, Santangelo G, Billeci L, Aprile IG, Germanotta M, et al. Systematic review of AI/ML applications in multi-domain robotic rehabilitation: trends, gaps, and future directions. J Neuroeng Rehabil 2025 Apr 09;22(1):79 [FREE Full text] [doi: 10.1186/s12984-025-01605-z] [Medline: 40205472]
10. Pookkuttath S, Rajesh Elara M, Sivanantham V, Ramalingam B. AI-enabled predictive maintenance framework for autonomous mobile cleaning robots. Sensors (Basel) 2021 Dec 21;22(1):13 [FREE Full text] [doi: 10.3390/s22010013] [Medline: 35009556]
11. Soori M, Arezoo B, Dastres R. Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cogn Robot 2023;3:54-70. [doi: 10.1016/j.cogr.2023.04.001]
12. Alzghaibi H. Adoption barriers and facilitators of wearable health devices with AI integration: a patient-centred perspective. Front Med (Lausanne) 2025 Apr 3;12:1557054 [FREE Full text] [doi: 10.3389/fmed.2025.1557054] [Medline: 40248075]
13. Hassan M, Kushniruk A, Borycki E. Barriers to and facilitators of artificial intelligence adoption in health care: scoping review. JMIR Hum Factors 2024 Aug 29;11:e48633 [FREE Full text] [doi: 10.2196/48633] [Medline: 39207831]
14. Felber NA, Lipworth W, Tian YJ, Roulet Schwab D, Wangmo T. Informing existing technology acceptance models: a qualitative study with older persons and caregivers. Eur J Ageing 2024 Mar 29;21(1):12 [FREE Full text] [doi: 10.1007/s10433-024-00801-5] [Medline: 38551677]
15. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506 [FREE Full text] [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
16. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
17. Chen K, Chan AH. Use or non-use of gerontechnology--a qualitative study. Int J Environ Res Public Health 2013 Sep 30;10(10):4645-4666 [FREE Full text] [doi: 10.3390/ijerph10104645] [Medline: 24084674]
18. Kooij L, Groen WG, van Harten WH. Barriers and facilitators affecting patient portal implementation from an organizational perspective: qualitative study. J Med Internet Res 2018 May 11;20(5):e183 [FREE Full text] [doi: 10.2196/jmir.8989] [Medline: 29752253]
19. Chan M, Estève D, Escriba C, Campo E. A review of smart homes- present state and future challenges. Comput Methods Programs Biomed 2008 Jul;91(1):55-81. [doi: 10.1016/j.cmpb.2008.02.001] [Medline: 18367286]
20. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug 1;35(8):982-1003 [FREE Full text] [doi: 10.1287/mnsc.35.8.982]
21. Hung L, Wong KL, Wong J, Park J, Mousavi H, Zhao H. Facilitators and barriers to using AI-enabled robots with older adults in long-term care from staff perspective: a scoping review protocol. BMJ Open 2023 Oct 30;13(10):e075278 [FREE Full text] [doi: 10.1136/bmjopen-2023-075278] [Medline: 37903609]
22. Elsheikh A, Al-Thani DA. Exploring the acceptance and adoption of humanoid robots among older adults: systematic review. In: Bramwell-Dicks A, Evans A, Winckler M, Petrie H, Abdelnour-Nocera J, editors. Design for Equality and Justice. Cham, Switzerland: Springer; Jul 9, 2024.
23. Chae K, Massare J, Ashida S, Cudjoe TK, Abadir P, Arbaje AI, et al. Stakeholders' perceived benefits and concerns regarding artificial intelligence in the care of older adults. J Am Geriatr Soc 2026 Feb;74(2):355-364. [doi: 10.1111/jgs.70228] [Medline: 41313148]
24. Schoenborn NL, Chae K, Massare J, Ashida S, Abadir P, Arbaje AI, et al. Perspectives on AI and novel technologies among older adults, clinicians, payers, investors, and developers. JAMA Netw Open 2025 Apr 01;8(4):e253316 [FREE Full text] [doi: 10.1001/jamanetworkopen.2025.3316] [Medline: 40184066]

XSL•FO
RenderX

25. Abadir P, Oh E, Chellappa R, Choudhry N, Demiris G, Ganesan D, et al. Artificial intelligence and technology collaboratories: innovating aging research and Alzheimer's care. Alzheimers Dement 2024 Apr;20(4):3074-3079 [FREE Full text] [doi: 10.1002/alz.13710] [Medline: 38324244]
26. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2022 Oct 25;24(1):1-6 [FREE Full text] [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
27. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
28. Peek ST, Wouters EJ, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJ. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014 Apr;83(4):235-248 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
29. Zhang X, Yu P, Yan J, Ton A M Spil I. Using diffusion of innovation theory to understand the factors impacting patient acceptance and use of consumer e-health innovations: a case study in a primary care clinic. BMC Health Serv Res 2015 Feb 21;15:71 [FREE Full text] [doi: 10.1186/s12913-015-0726-2] [Medline: 25885110]
30. Cutler D, Wikler E, Basch P. Reducing administrative costs and improving the health care system. N Engl J Med 2012 Nov 15;367(20):1875-1878. [doi: 10.1056/NEJMp1209711] [Medline: 23150956]
31. Neumann PJ, Cohen JT. QALYs in 2018-advantages and concerns. JAMA 2018 Jun 26;319(24):2473-2474. [doi: 10.1001/jama.2018.6072] [Medline: 29800152]
32. Herzlinger RE. Why innovation in health care is so hard. Harv Bus Rev 2006 May;84(5):58-66, 156. [Medline: 16649698]
33. Lyles CR, Nelson EC, Frampton S, Dykes PC, Cemballi AG, Sarkar U. Using electronic health record portals to improve patient engagement: research priorities and best practices. Ann Intern Med 2020 Jun 02;172(11 Suppl):S123-S129 [FREE Full text] [doi: 10.7326/M19-0876] [Medline: 32479176]
34. Warraich HJ, Tazbaz T, Califf RM. FDA perspective on the regulation of artificial intelligence in health care and biomedicine. JAMA 2025 Jan 21;333(3):241-247. [doi: 10.1001/jama.2024.21451] [Medline: 39405330]
35. Sandalow M, Adams K. FDA oversight: understanding the regulation of health AI tools. Bipartisan Policy Center. 2025 Nov 10. URL: https://bipartisanpolicy.org/issue-brief/fda-oversight-understanding-the-regulation-of-health-ai-tools/ [accessed 2026-03-08]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
JHU-AITC: Johns Hopkins Artificial Intelligence and Technology Collaboratory
TAM: technology acceptance model

Edited by M Mardini; submitted 19.Oct.2025; peer-reviewed by MM Shannon, V Dukhanin; comments to author 10.Feb.2026; revised version received 22.Jun.2026; accepted 23.Jun.2026; published 29.Jul.2026.

Please cite as:
Zhang Z, KM Cudjoe T, Ashida S, Massare J, Chae K, Phan P, Abadir P, I Arbaje A, Unberath M, Schoenborn NL
The Ways in Which Stakeholders Make Decisions About AI and Novel Technologies for the Health Care of Older Adults: Qualitative Interview Study
JMIR Aging 2026;9:e86148
URL: https://aging.jmir.org/2026/1/e86148
doi:10.2196/86148
PMID:

©Zhang Zhang, Thomas KM Cudjoe, Sato Ashida, Jacqueline Massare, Kacey Chae, Phillip Phan, Peter Abadir, Alicia I Arbaje, Mathias Unberath, Nancy L Schoenborn. Originally published in JMIR Aging (https://aging.jmir.org), 29.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging Multimodal Large Language Models for Fall Risk Reduction in Older Adults in the Home: Proposed Model Design

Justin Do[1], BS; Vivaswat Suresh[2], BS; Lily Zhang[1], BA; Bharvi M Chavre[1], BA; Jeremy Cha[1], BA; Robert Pugliese[1,2], PharmD

[1]Sidney Kimmel Medical College, Thomas Jefferson University, 925 Chestnut St., Basement Vault, Philadelphia, PA, United States

[2]Health Design Lab, Thomas Jefferson University, Philadelphia, PA, United States

Corresponding Author:
Justin Do, BS
Sidney Kimmel Medical College, Thomas Jefferson University, 925 Chestnut St., Basement Vault, Philadelphia, PA, United States

Abstract

This research letter proposes a novel model design leveraging natively multimodal large language models to identify fall risks and generate visualizations of recommended home environmental modifications, aiming to improve the accessibility and impact of personalized fall prevention advice for older adults. Through a pilot rating study, this work demonstrates that multimodal large language models can generate safe and actionable advice to reduce fall risk in lived spaces of older adults, and also generate realistic edits based on original images. While this concept needs further testing and clinical comparison, it highlights a promising avenue for further innovation of fall prevention tactics.

(JMIR Aging 2026;9:e77591) doi:10.2196/77591

KEYWORDS

LLM; multimodal; image generation; fall risk; older adults; large language model

Introduction

Falls among older adults cause significant mortality and increased health care costs [1]. Current literature has identified combined behavioral and exercise interventions as effective preventions for fall risks, improving balance performance, and reducing fear of falls [2], while limited evidence exists regarding medication-induced falls [3]. Home environmental intervention is also effective: safety assessments have been shown to reduce fall rate by 23% - 36% [1], with applied home modifications contributing a 7% risk reduction [4]. However, both external (insurance) and self-imposed (ie, the perception that safety assessments are invasive) barriers impede widespread implementation [5]. While research on frailty assessments is robust, gaps remain in technology-enabled interventions [6]. Prior studies have shown acceptability by older adults to embrace digital and electronic tools [5]. Existing remote home assessment protocols rely on caregiver camera operation, written instructions comprehension, and professional review of footage [7], while telehealth occupational therapy (OT) assessments may require insurance authorization, creating both obstacles and delays. Multimodal large language models (LLM) can fuse visual and text information, offering a scalable alternative while preventing encroachment on user values. This study aims to evaluate the ability of LLMs to produce safe, clinically useful, and actionable outputs that identify fall risks from user-provided home imagery and uniquely generate visualizations of the recommended environmental modifications.

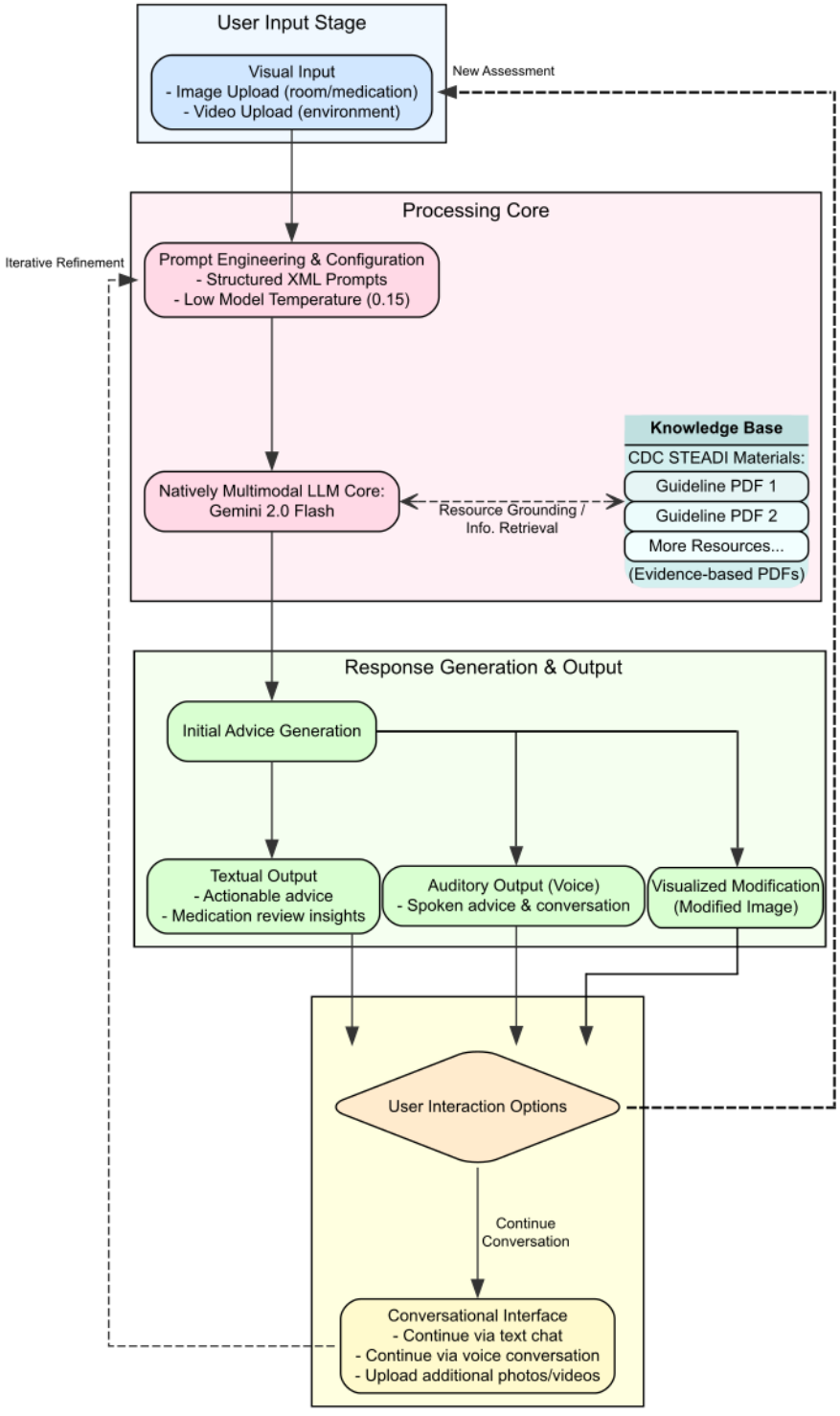
Methods

We selected Google's Gemini family due to its strong visual reasoning performance supported by validated benchmarks [8]. We designed our framework to focus on providing reliable output by employing a low model temperature (0.15), in-context learning through grounding responses in evidence-based CDC STEADI patient materials (Figure 1), and structured XML prompts iterated using artificial intelligence (AI)-driven prompt engineering. The core innovation of this study used the gemini-2.0-flash-exp-image-generation model to directly modify the uploaded images with the model's suggested changes. The model leverages a two-shot prompting system (Multimedia Appendix 1, Multimedia Appendix 2), in which the primary LLM generates textual recommendations based on an input image/video, then directs the image generation model to visually render these changes (eg, adding grab bars, removing hazards) onto the original image, iterating until the generated image reflects the proposed modifications. We conducted a formative, blinded, paired comparison of outputs generated from 27 publicly licensed "lived-in" home interior images. We compared a non-optimized baseline prompt with an enhanced multimodal pipeline ("Steadi"). Text and image outputs were compared based on clinical usefulness, safety, image fidelity/plausibility, and preference between baseline prompt output and our

XSL•FO
RenderX

enhanced multimodal pipeline output. Detailed methods and output can be found in Multimedia Appendices 3 and 4.

Figure 1. Enhanced model architecture for advice generation and multimodal interaction. CDC: Centers for Disease Control and Prevention; LLM: large language model; STEADI: Stopping Elderly Accidents, Deaths, & Injuries.



Results

Initial Advice Generation, Multimodal Communication, and Modification Visualization

The model takes an uploaded image or video and provides specific, actionable advice supported by evidence-based resources. The proposed architecture successfully applies both additive and subtractive modifications to images, providing users with a concrete visual representation of a safer environment.

XSL•FO
RenderX

Model Comparisons

We find that overall, our raters preferred the "Steadi" system output of both image and text (40/54 times (74.1%), Table 1). We demonstrate that contemporary LLMs produce relatively safe recommendations, regardless of the prompting system, with only one set of recommendations rated as unsafe in the baseline prompting system and none in the enhanced. We find that when given specific issues to visualize, image editing LLMs produce edits with good visualization fidelity (46/54 times, 85.2% for baseline and 43/54 times, 79.6% for enhanced; Table 1), and low rates of implausible/hazard-producing edits (6/54 times, 11.1% for both systems; Table 1). Text recommendations and visualized outputs ranged from generally "somewhat actionable" for the baseline system to highly actionable for our enhanced system.

Table . Results of rating study.

Outcome	Baseline	Enhanced	Comparison[a]
Q1 Overall clinical usefulness (preference)	Preferred: 10/54 (18.5%)	Preferred: 40/54 (74.1%)	Win rate: 40/50 (80.0%, 95% CI 67.0 - 88.8); sign test $P \leq .001$; ties=4
Q2 Unsafe/inappropriate recommendation rate (Yes)	1/54 (1.9%, 95% CI 0.3 - 9.8)	0/54 (0.0%, 95% CI 0.0 - 6.6)	Risk difference (Enhanced-Baseline): −1.9 pp
Q3 Visualization fidelity rate (Yes)	46/54 (85.2%, 95% CI 73.4 - 92.3)	43/54 (79.6%, 95% CI 67.1 - 88.2)	Risk difference (Enhanced-Baseline): −5.6 pp
Q4 Hazard-introducing/implausible edit rate (Yes)	6/54 (11.1%, 95% CI 5.2 - 22.2)	6/54 (11.1%, 95% CI 5.2 - 22.2)	Risk difference (Enhanced-Baseline):+0.0 pp
Q5 Actionability (1 - 5 Likert)	median 3.0 (IQR 3.0 - 4.0)	median 5.0 (IQR 4.0 - 5.0)	Δmedian (Enhanced-Baseline):+1.0 (IQR 0.0 - 2.0)

[a]n=rater case evaluations; outcomes are descriptive; Q1 sign test is exploratory.

Discussion

Principal Findings

This study introduces a novel application of multimodal LLMs, leveraging their image-generation capabilities for visualizing personalized home safety recommendations. We demonstrate that enhanced frameworks, such as structured prompting and grounding using trusted resources, produce safe, clinically useful, and actionable outputs that categorically rate better than outputs from baseline LLMs. The inherent flexibility of LLMs supports diverse interaction methods, uniquely enabling users to interact with their "consultant" in their preferred mode. LLMs may mitigate delays caused by insurance authorizations and restore autonomy to users.

The visual output capability is also key: generating suggestions directly onto uploaded images offers more intuitive, actionable guidance than abstract text instructions alone. The drive to protect the familiarity of their home from change was identified to be a major motive for older adults rejecting modification advice from OT [5]; direct visualization of user-fed images may help overcome this hurdle and increase acceptance. There are still limitations to the technology, namely outputs may be illogical such as the recommended soap placement, and movement of furniture and door in Multimedia Appendix 1. However, overall, this study demonstrates that LLMs generally produce visual outputs with high fidelity, low hazard introduction rates, and high actionability.

To adhere to HIPAA (Health Insurance Portability and Accountability Act) compliance, future work should consider working with LLM providers to sign a HIPAA Business Associate Amendment or other HIPAA-compliant program. Ethical considerations, such as disclosure of privacy and data protection, should be implemented in accordance with WHO guidance on AI in health [9].

Limitations

Further testing must be conducted against the current standard for in-home assessments to discover if the proposed model provides comparable advice to professionals. Implementation trials will be needed to mitigate concerns such as the digital divide and ensure accessibility among varying cognitive/visual functions. Implementations must comply with FDA digital health guidance, and characterization and limitation of unsafe output generation must be explored. This model is designed as a supplemental service to be integrated with OT rather than a replacement.

Conclusions

Multimodal LLMs that integrate image generation offer a novel, innovative approach to increasing end users' accessibility to personalized home environment recommendations for fall prevention. This capability represents a potential supplement to current care services that may enhance patient understanding, motivation, and adherence, serving as a valuable resource to patients who defer or cannot access in-home safety assessments. Rigorous validation of clinical efficacy and user acceptance is essential to translate this technological potential into improved patient outcomes.

XSL•FO
RenderX

Acknowledgments

We would like to extend our gratitude to Robert Pugliese and MaryEllen Daley for all of their generous support throughout this project. We would also like to thank Dr. Bracken Babula, MD, Dr. Zhe Chen, MD, Dr. Ryan Tomlinson, PhD, Dr. Deanna Gray-Miceli, PhD, CRNP, Dr. Christine Hsieh, MD, and Dr. Brooke Salzman, MD, for their feedback and guidance on this project. No generative AI was used in any portion of the manuscript text generation. We used the generative AI Tool "Gemini 1.5 Pro" made by Google to draft the system prompt found in Multimedia Appendix 2, with review and editing from the study group. Image portions of Multimedia Appendices 1, 4 and 5 were generated using "Gemini 2.0 Flash Image Preview" as described in the manuscript text as part of the model design. Image portions of Multimedia Appendices 1, 4 and 5 were generated using "Gemini 2.5 Flash Image" as described in the manuscript text as part of the model design.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Authors' Contributions

Conceptualization: JD, LZ, BC, JC. Methodology: JD, VS. Resources: JD, LZ, BC, JC.Supervision: RP. Writing – Original draft: JD. Writing – Revising and editing: JD, VS, LZ, BC, JC.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Two-shot image modification architecture.
[PNG File, 941 KB - aging_v9i1e77591_app1.png]

Multimedia Appendix 2
XML Prompt and model parameters.
[DOCX File, 10 KB - aging_v9i1e77591_app2.docx]

Multimedia Appendix 3
Blinded comparison methods.
[PDF File, 94 KB - aging_v9i1e77591_app3.pdf]

Multimedia Appendix 4
Rating packet.
[PDF File, 15079 KB - aging_v9i1e77591_app4.pdf]

Multimedia Appendix 5
Usage of model for recommendations for post-stroke patients.
[PNG File, 1063 KB - aging_v9i1e77591_app5.png]

References

1. Niedermann K, Meichtry A, Zindel B, et al. Effectiveness and cost-effectiveness of a single home-based fall prevention program: a prospective observational study based on questionnaires and claims data. BMC Geriatr 2024 Dec 28;24(1):1044. [doi: 10.1186/s12877-024-05586-x] [Medline: 39732675]
2. Azizan A, Justine M. Elders' exercise and behavioral program: effects on balance and fear of falls. Phys Occup Ther Geriatr 2015 Oct 2;33(4):346-362. [doi: 10.3109/02703181.2015.1093060]
3. Gillespie LD, Robertson MC, Gillespie WJ, et al. Interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev 2012 Sep 12;2012(9):CD007146. [doi: 10.1002/14651858.CD007146.pub3] [Medline: 22972103]
4. Lektip C, Chaovalit S, Wattanapisit A, Lapmanee S, Nawarat J, Yaemrattanakul W. Home hazard modification programs for reducing falls in older adults: a systematic review and meta-analysis. PeerJ 2023;11:e15699. [doi: 10.7717/peerj.15699] [Medline: 37489124]
5. Lee JJ, Patel D, Gadgil M, Langness S, von Hippel CD, Sammann A. Understanding barriers to home safety assessment adoption in older adults: qualitative human-centered design study. JMIR Hum Factors 2025 Jun 24;12:e66854. [doi: 10.2196/66854] [Medline: 40554799]

XSL•FO
RenderX

6. Azizan A. Exercise and frailty in later life: a systematic review and bibliometric analysis of research themes and scientific collaborations. IJPS 2024;11(1):1. [doi: 10.36922/ijps.3282]
7. Romero S, Lee MJ, Simic I, Levy C, Sanford J. Development and validation of a remote home safety protocol. Disabil Rehabil Assist Technol 2018 Feb;13(2):166-172. [doi: 10.1080/17483107.2017.1300345] [Medline: 28326967]
8. Yue X, Ni Y, Zheng T, et al. MMMU: a massive multi-discipline multimodal understanding and reasoning benchmark for expert AGI. arXiv. Preprint posted online on Jun 13, 2024. [doi: 10.48550/arXiv.2311.16502]
9. Regulatory considerations on artificial intelligence for health. : World Health Organization; 2023 URL: https://www.who.int/publications/i/item/9789240078871 [accessed 2025-11-01]

Abbreviations

AI: artificial intelligence
CDC: Centers for Disease Control and Prevention
HIPAA: Health Insurance Portability and Accountability Act
LLM: large language model
MMMU: Massive Multi-discipline Multimodal Understanding and Reasoning
NICE: National Institute for Health and Care Excellence
OT: occupational therapy
STEADI: Stopping Elderly Accidents, Deaths, & Injuries
XML: Extensible Markup Language

Edited by J Wang; submitted 21.May.2025; peer-reviewed by A Azizan, D Menychtas; revised version received 01.Mar.2026; accepted 16.Mar.2026; published 13.May.2026.

Please cite as:
Do J, Suresh V, Zhang L, Chavre BM, Cha J, Pugliese R
Leveraging Multimodal Large Language Models for Fall Risk Reduction in Older Adults in the Home: Proposed Model Design
JMIR Aging 2026;9:e77591
URL: https://aging.jmir.org/2026/1/e77591
doi:10.2196/77591

© Justin Do, Vivaswat Suresh, Lily Zhang, Bharvi M Chavre, Jeremy Cha, Robert Pugliese. Originally published in JMIR Aging (https://aging.jmir.org), 13.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Telehealth Barriers and Digital Ageism Experienced by Older Veterans: Formative Ethnographic Study to Inform a Telepharmacy Randomized Trial

Maria Venegas[1,2,3*], MPH, PhD; Megan B McCullough[2*], PhD; Chelsea E Hawley[3,4*], PharmD, MPH; Judith L Beizer[5*], PharmD; William W Hung[6,7], MD, MPH; Lauren R Moo[4,8,9*], MD

[1]Veterans Health Administration, 200 Springs Road Bld. 70 #128, Bedford, MA, United States

[2]Center for Healthcare Organization and Implementation Research, Veterans Health Administration, Bedford, MA, United States

[3]Department of Medicine, Division of Geriatrics, Boston University, Boston, MA, United States

[4]New England Geriatric Research Education and Clinical Center, Veterans Health Administration, Bedford, MA, United States

[5]College of Pharmacy and Health Sciences, St. John's University, New York, NY, United States

[6]Geriatric Research Education and Clinical Center, James J. Peters VA Medical Center, Veterans Health Administration, Bronx, NY, United States

[7]Department of Geriatrics and Palliative Medicine, Icahn School of Medicine, New York, NY, United States

[8]Department of Neurology, Harvard University, Cambridge, MA, United States

[9]Department of Neurology, Massachusetts General Hospital, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Maria Venegas, MPH, PhD
Veterans Health Administration, 200 Springs Road Bld. 70 #128, Bedford, MA, United States

Abstract

Background: As telehealth has become an increasingly common mode of care delivery, older adults may face structural, technological, and interactional barriers that limit their ability to engage with video-based care. Although digital ageism, defined as the presence of age-related stereotypes, lowered expectations, or assumptions about older adults' technology-related competence, has been described in prior literature, less is known about how such dynamics surface during real-time telehealth encounters and how they may shape participation in technology-focused clinical trials.

Objective: This formative ethnographic study aimed to (1) document real-world barriers encountered by older adults immediately before and during video telehealth visits and (2) inform recruitment and implementation procedures for a subsequent telepharmacy randomized controlled trial.

Methods: We conducted in-home, real-time ethnographic observation of 20 community-dwelling veterans aged ≥65 years participating in pharmacist-led video visits for medication management. Recruitment occurred over approximately 6 to 10 months using mailed invitation letters (>300 sent), supplemented with outbound telephone calls. Data sources included structured field notes completed independently by an in-house anthropologist and the remote clinical pharmacist, as well as observational documentation of previsit preparation, visit navigation, and postvisit reflections. Data were analyzed using qualitative rapid analysis, with iterative team review and triangulation across data sources.

Results: Participants had a mean age of 74 (SD 3.18) years; 19 of 20 (95%) were male, and 18 of 20 (90%) identified as White. All participants completed a video visit with technical support as needed. Structural barriers (eg, broadband access and device availability) and usability challenges (eg, audio-video setup and navigation) were common. Although digital ageism was not a predefined analytical category, age-related assumptions about technology emerged during observation, including participants attributing anticipated or experienced difficulties to age and expressing surprise or pride following successful completion of the visit. These age-related interpretations were analytically distinct from access and usability barriers and were interpreted as manifestations of digital ageism, particularly as internalized age-based expectations.

Conclusions: Formative ethnographic observation provided critical insights into how older adults experience telehealth encounters in real-world contexts and informed adaptations to recruitment and implementation procedures for a subsequent randomized controlled trial. Although digital ageism was not an original study aim, age-related assumptions about technology emerged as an important interpretive factor shaping engagement with video-based care. Incorporating ethnographic methods prior to trial implementation may help identify otherwise overlooked barriers and improve the inclusivity and feasibility of technology-focused clinical research involving older adults.

XSL•FO
RenderX

(JMIR Aging 2026;9:e79409) doi:10.2196/79409

KEYWORDS

digital ageism; telehealth; ethnographic research; older adults; telepharmacy; digital divide

Introduction

The COVID-19 pandemic accelerated the adoption of video telehealth as a primary mode of outpatient care delivery, including among older veterans [1-5]. Telehealth, defined by the National Institute on Aging as the use of communication technologies to exchange information and deliver health care services, holds the promise of increasing access to care and reducing logistical barriers while maintaining clinical effectiveness [6]. However, despite widespread interest in telemedicine among older adults, significant obstacles remain [7-12]. Limited broadband access and the high costs of digital devices and internet services—issues that are especially pronounced in rural and low-income communities—restrict many older adults' ability to engage with telehealth, potentially worsening health inequities [13-18]. Another critical challenge is digital ageism: the assumption that older individuals are less capable of using digital tools [19]. This bias can manifest through technological, structural, and attitudinal barriers [20,21]. Many virtual care platforms are not designed with older users in mind, often featuring small fonts, complex navigation, and controls that are not accommodating for those with motor or visual impairments [13]. In addition, the expectation of digital literacy can alienate individuals who have had limited exposure to technology, leading to frustration and disengagement [12,16,20]. Research within the Department of Veterans Affairs (VA), for example, has shown that while many older veterans are interested in video visits, they often require significant personalized assistance to participate effectively [1,17,22]. Furthermore, some health care providers may default to telephone visits or fail to offer adequate digital literacy support, reinforcing the notion that telemedicine is "not for them" and further discouraging engagement [21,23,24].

Although disparities in video telehealth use among older adults are well documented, a more nuanced understanding of the real-life barriers these individuals face is essential for designing digital tools and platforms that truly meet their needs. To address this gap, we conducted a formative ethnographic assessment of older adults' experiences during virtual visits with a clinician. This formative ethnographic study aimed to (1) document real-world barriers encountered by older adults immediately before and during a video telehealth visit and (2) inform recruitment and study procedures for a subsequent telepharmacy randomized controlled trial (RCT). Although digital ageism was not a predefined analytical category, age-related assumptions about technology were examined as an emergent interpretative theme.

Methods

Study Overview

This study draws on the same formative ethnographic dataset described in our prior publication examining feasibility and practical implementation solutions for pharmacist video visits among older veterans [25]. This manuscript extends that work by focusing on how ethnographic methods can identify and disentangle age-related assumptions from technology familiarity in the design of technology-based clinical trials. The analytical focus and conceptual framing in this manuscript are distinct from those in the prior implementation publication.

Ethical Considerations

This study was reviewed and approved by the VA Central Institutional Review Board (approval number 19-35). All participants provided informed consent prior to participation. Study staff explained the purpose of the study, procedures, potential risks, and voluntary nature of participation, and participants were informed of their right to withdraw at any time without consequences to their care. To protect participant privacy and confidentiality, all data were deidentified at the time of analysis. Field notes and observational data were stored on secure, VA-approved servers with access limited to authorized study personnel. No personally identifiable information is reported in this manuscript. Participants received a US $50 gift card as compensation for their time.

Recruitment

Potential participants were identified using an electronic health record–generated list of veterans aged ≥65 years who met the inclusion criteria (eg, ≥2 chronic conditions, no diagnosis of dementia, and at least 1 primary care visit in the prior year). Participants were recruited over approximately 10 months using a combination of mailed invitations and outbound telephone calls. In total, 316 invitation letters were mailed to veterans identified through the electronic health record as potentially eligible. Overall, 64 veterans responded to the mailed invitation; 41 indicated they were not interested in participating, and 23 initially agreed. Of these, 2 veterans subsequently declined enrollment due to personal reasons (ongoing cancer treatment and caregiving responsibilities) and 1 could not be scheduled within the recruitment window; recruitment stopped once the target sample size (n=20) was reached.

To supplement mailed invitations, the study staff conducted outbound telephone calls to veterans who did not respond to the letter. Across mail and telephone contact attempts, veterans either declined participation after learning more about the study or could not be reached (eg, disconnected phone numbers or no response after repeated attempts). In total, 17 veterans were determined to be ineligible due to no longer meeting the study criteria (eg, not currently taking ≥5 medications, recent diagnosis of cognitive decline, or enrollment in a VA satellite clinic). Recruitment ceased once the target sample size (N=20) was reached. All enrolled participants completed the observed telehealth visit.

Our ethnographic assessment used direct observation and structured field notes to capture the real-world experiences of

XSL•FO
RenderX

older adults as they navigated a clinical video visit using the VA's synchronous video platform VA Video Connect (VVC). An experienced anthropologist led this assessment by recruiting and enrolling participants and joining each participant at home before and during their virtual visit.

To participate, all veterans needed internet access at home and a device with a camera and microphone. For those lacking a video-capable device, the anthropologist arranged for the provision of a VA-loaned tablet. The VA's tablet distribution program was instrumental in bridging the digital divide for veterans with limited access to video care. Veterans receiving a tablet were scheduled for a test visit with a local VA employee, while others received a demonstration from the anthropologist prior to their pharmacist video visit. Before the video visit, the anthropologist provided a video demonstration when needed. Subsequently, the participant and pharmacist connected via VVC.

Ethnographic methods in anthropology involve qualitative techniques—such as semistructured interviews and focus groups—with a unique emphasis on situating specific events within their broader real-life contexts [22,26,27]. Unlike surveys or post hoc interviews, ethnographic methods also rely on in-situ, real-time observations to capture sociocultural, behavioral, and technical challenges that participants may not articulate in interviews [27]. In this assessment, field notes and direct observation notes were critical in documenting the technical, logistical, and emotional dimensions of older adults' virtual care experiences.

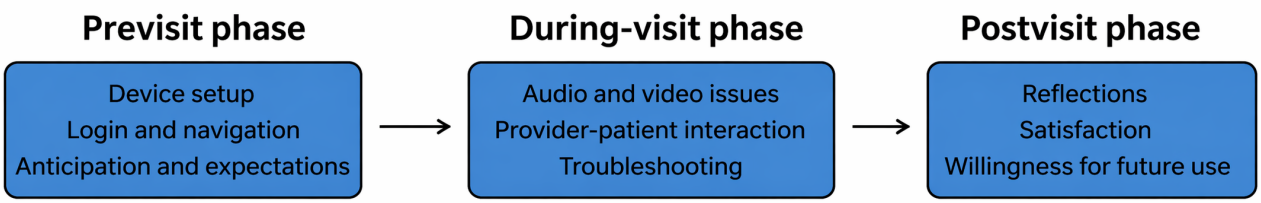
Field notes are detailed, descriptive records that researchers, particularly anthropologists and ethnographers, take during participant observation and interviews [28,29]. We designed a standardized field note form informed by a priori concepts from the Consolidated Framework for Implementation Research (CFIR) [30] and insights from experienced clinical pharmacists (Multimedia Appendix 1). Field notes are generally understood to be primary data collected in the field to document observations, interactions, and insights gained through immersive research methods [28]. Both the in-house anthropologist and the remote clinical pharmacist completed field notes for each participant. The anthropologist's (in-house) and the clinical pharmacist's (remote) field notes complemented each other by capturing a range of experiences and challenges, including the following:

- Technical challenges and troubleshooting; difficulties with video connections, microphone access, and screen navigation
- Participant engagement and reactions; variations in comfort, frustration, confidence, or hesitation in using virtual care platforms
- Health care provider interactions; the level of guidance provided, communication tone, and patience in assisting older adults
- Environmental and sociocultural factors; home setup, presence of caregivers, privacy concerns, and distractions

Direct observation is an ethnographic method ideal for studies on naturally occurring behaviors, actions, or events [29]. For example, observation is particularly useful for understanding patients' experiences because it provides an insider perspective and thus allows researchers to understand end users' experiences of a problem. Direct observation was conducted before, during, and after the virtual visit session to fully understand the participant's experience. Key components included the following (Figure 1):

- Previsit observations; how participants prepared (eg, setting up devices and logging in), their anxiety or reluctance related to technology use, and the role of caregivers
- During the virtual care visit; navigation challenges (eg, adjusting volume and enabling the camera), participant-provider interactions, and unanticipated disruptions such as device malfunctions or environmental distractions
- Postvisit reflections; participant feedback on the ease of use, overall satisfaction, willingness to use virtual care in the future, and any adaptive strategies used (eg, using written notes or seeking caregiver assistance)

Figure 1. Direct observation in telehealth interactions. This diagram outlines the key stages of direct observation in telehealth research with older adults, including previsit, during-visit, and postvisit phases.



XSL•FO
RenderX

Data Analysis

This formative ethnographic assessment was not originally designed to study digital ageism as a primary construct. Rather, the study focused on understanding how older adults navigated real-time telehealth encounters to inform the implementation of a subsequent RCT. Because the study population consisted exclusively of older adults, we anticipated that age-related assumptions about technology, whether expressed by participants, embedded in clinical workflows, or reflected in research procedures, might emerge alongside structural and usability barriers. Data sources included structured field notes completed independently by the in-house anthropologist and the remote clinical pharmacist for each observed visit, as well as observational documentation of previsit preparation and postvisit reflections. All field notes were compiled and organized by participant and visit stage.

We analyzed the ethnographic data (structured field notes and direct observation summaries) using a qualitative rapid analysis approach commonly applied in health services and implementation research, emphasizing matrix-based summarization, team-based interpretation, and iterative refinement of themes [31,32]. This approach is consistent with pragmatic ethnographic methods, including Rapid Assessment Procedure–Informed Clinical Ethnography, which integrates ethnographic observation with rapid, implementation-focused analysis [33].

The rapid qualitative analytic process proceeded in 4 iterative steps. First, field notes were summarized using structured analytic matrices organized by participant; visit phase (previsit, during visit, and postvisit); and focal domains (eg, technical barriers, participant responses, environmental context, and caregiver involvement). This step allowed for systematic comparison across cases while preserving contextual detail. Second, 2 analysts (the anthropologist and the clinical pharmacist) independently reviewed matrix summaries to identify recurring patterns and salient observations. Preliminary themes were generated inductively, with attention to both anticipated implementation challenges (eg, technology access and usability) and emergent phenomena observed during visits.

Third, the 2 primary analysts met weekly over a 6-month period to compare interpretations, refine thematic definitions, and resolve discrepancies through discussion and reference back to the original field notes. These meetings focused on clarifying analytic boundaries, distinguishing overlapping concepts, and ensuring consistency across cases. Fourth, biweekly analytic meetings with a senior qualitative methodologist and a clinical subject matter expert provided an additional layer of review. These meetings were used to challenge emerging interpretations, assess analytic coherence, and confirm that themes were grounded in observational evidence.

Several strategies were used to enhance analytic rigor and credibility [34]. First, triangulation across data sources (anthropologist field notes, pharmacist field notes, and direct observation) allowed for cross-validation of findings. Second, team-based analysis and consensus discussions reduced reliance on a single analyst's interpretation. Finally, analytic decisions were documented through iterative matrix revisions and meeting notes, creating an audit trail of theme development. This rapid, iterative analytic process enabled the identification of implementation-relevant insights while maintaining methodological transparency and rigor appropriate for ethnographic research embedded within clinical trial preparation.

We distinguished digital ageism from related but distinct phenomena, such as structural barriers to access (eg, broadband availability and device ownership) and usability or design limitations of telehealth platforms. We reserved the term "digital ageism" for age-related stereotypes, lowered expectations, and identity-based assumptions, whether externally imposed or internalized, that shaped participants' perceptions of their own technological competence and their engagement with virtual care. Consequently, digital ageism was used as an interpretive lens during analysis, rather than as a priori coding framework or outcome.

Results

Participant Characteristics

Participants had a mean age of 74 (SD 3.18) years; of 20 participants, 18 (90%) were male, and 18 (90%) identified as White. In total, 11 (55%) participants had completed high school or a General Educational Development test, and 13 (65%) participants reported some prior experience with digital technology (eg, owning or using a smartphone, tablet, or computer). However, experience with video-based platforms was limited for many participants, and assistance was frequently required; 7 (35%) participants reported neither regular access to a video-capable device nor prior experience using such technology and required assistance both before and during the virtual visit. Additionally, 3 (15%) participants reported having access to a device and some experience with technology but nonetheless encountered technical difficulties during the visit that required real-time support.

Lessons Learned to Refine the Implementation of a Subsequent Telepharmacy RCT

Overview

Direct in-home guidance was associated with successful visit completion, and participants frequently expressed comfort with specific tasks (eg, logging in and adjusting audio or video) during or after the visit. These observations informed mitigation strategies incorporated into the subsequent RCT protocol. Common challenges included difficulty navigating devices, audio-visual setup problems, and unstable internet connectivity. These barriers were identified prior to full trial implementation, allowing the study team to develop mitigation strategies such as providing VA-loaned tablets, offering previsit device and platform demonstrations, and delivering real-time troubleshooting support.

Identifying and Addressing Barriers Before the Trial

Pretrial qualitative and direct observation helped identify technological and structural barriers that would impact trial retention and data integrity. Direct in-home observations revealed challenges in device navigation, audio-visual accessibility, and broadband connectivity, allowing for proactive

XSL•FO
RenderX

solutions such as providing VA-loaned tablets and real-time technical assistance. These observations informed mitigation strategies for the subsequent RCT (eg, previsit platform demonstrations and caregiver involvement strategies).

Enhancing Participant Recruitment and Retention

Recruitment was the first opportunity to engage potential participants and build trust. We recruited 20 individuals over approximately 10 months during the COVID-19 pandemic. Despite the challenges of masks and the recommended 6-foot physical distance between parties, participants were accommodating and tolerant of the masking and other restrictions. For some veterans, the anthropologist was the only in-person visitor they had had in months and was greatly welcomed. By meeting patients at their home, we eliminated several barriers to research that homebound older adults face, such as a lack of transportation or mobility issues that limit their ability to participate in hospital-based research. The anthropologist was available via telephone for questions or concerns and was flexible, driving to participants' homes (sometimes 45 to 75 minutes away in rural locations) to conduct the enrollment and baseline screening when needed, and then returned to observe the in-home pharmacist video visit and conduct the postvisit interview. Often, this occurred immediately following the baseline screening to reduce the burden and infection-related safety risks of scheduling several in-person appointments with the participant. This unconventional approach proved to be successful in reaching out and engaging a population traditionally excluded from clinical and technology-based research and allowed our anthropologist to understand the physical constraints and isolation that many older adults experience on a day-to-day basis.

Improving Intervention Implementation

Observational field notes further captured deviations from the planned protocol that would not have been apparent in a traditional trial setting, including extended time needed for technical setup, the role of caregivers in facilitating participation, and the cumulative burden of scheduling multiple appointments. These insights informed refinements to streamline procedures and reduce participant burden. Our ethnographic assessment allowed for adaptation of the intervention before full-scale RCT implementation, such as:

- Adjusting session scheduling to minimize participant burden
- Integrating troubleshooting strategies for common technical difficulties
- Addressing digital literacy gaps before the RCT to improve participant readiness, reduce technical failures, and reduce user frustration during actual trial sessions

By iteratively refining procedures based on our direct observations in the participants' homes, our ethnographic approach ensured that the RCT was not only scientifically rigorous but also practical and participant-centered. The opportunity to be in the home of the participants ensured that older adults could access the technology and complete the visit and, more importantly, that the anthropologist was able to witness the process by which the video visit was delivered from the patient's perspective. This step was valuable in understanding how actual implementation differs from planned implementation and in identifying issues important to the transferability of an effective intervention outside experimental conditions [27].

Analytic Findings

Emergent Age-Related Assumptions Shaping Telehealth Engagement

Although digital ageism was not a predefined analytic category, age-related assumptions about technology emerged consistently during observation and participant interaction. We analytically distinguished age-related assumptions from participants' actual familiarity with technology or prior exposure to digital tools. Although technical challenges and levels of experience varied, age-related assumptions were evident when participants explicitly invoked age to explain anticipated reluctance to engage with technology, rather than when challenges reflected task-specific skills or access alone.

Multifaceted Barriers to Video Telehealth Engagement

Older adults faced technological barriers, such as difficulty navigating videoconferencing platforms, small touchscreens, and poor audio-visual accessibility. Structural barriers, including a lack of broadband access, reliance on outdated technology, and the need for assistance with device setup, limited participation in virtual visits. To overcome these hurdles, the anthropologist in our study brought a VA-owned tablet to the homes of those who did not have access to their own device. She also provided a quick demonstration of how to use the VA tablet before the scheduled video visit with the pharmacist. This was especially useful for those who were unfamiliar with touch screens and video call platforms. For participants who did not have access to a device or to reliable broadband internet, the study team took advantage of a national VA program to arrange for a 4G-enabled, VA-issued tablet to be sent to the veteran that included cellular service. Similar task-specific challenges (eg, navigating apps, managing passwords, and enabling audio permissions) were observed among participants with different levels of technology experience and skills, suggesting that difficulty alone did not map cleanly onto age-related assumptions. Furthermore, the anthropologist noted that participants voiced fewer concerns about technology when provided with access to a device and hands-on guided tutorials before and during the virtual visit, when needed.

Hands-On Support: Illuminated Barriers and Work-Arounds for Older Adults

Real-time troubleshooting, such as enabling microphone access, adjusting volume settings, or assisting with log-in procedures, was critical in ensuring the successful completion of the video visit with the clinician. We also learned where participants were encountering difficulties and needed support, allowing us to further adapt our protocols. For example, for the RCT, we began asking participants to bring their devices to the consent appointment so that study staff could assist them with tasks such as allowing the VVC app to use the camera and microphone.

XSL•FO
RenderX

Direct Engagement With Technology Challenges: Internalized Ageism

In addition to structural and usability barriers, ethnographic field notes documented age-related assumptions that shaped how participants engaged with video visits. Successful completion of the video visit was often accompanied by expressions of relief, appreciation, or surprise, particularly among participants with limited prior experience using video-based care independently. For example, a female participant in her late 70s was at first reluctant to participate in the study because she lacked prior experience with technology. However, when given the opportunity to learn a new skill through hands-on support, she agreed to participate. The participant did not own a smartphone or computer and was skeptical that a clinical video visit would provide the same "whole picture view" as an in-person visit. Prior to the scheduled video visit, the anthropologist provided a short demonstration of how to navigate the tablet, how to set it up, and how to handle it if asked to walk around during the virtual visit. The participant was a little intimidated by the touchscreen modality, as she only owned a flip phone and had never used such a device before. The anthropologist encouraged the participant to practice using the onscreen keyboard, test the video modality, and type her own information into the registration form. During the scheduled video visit, the participant was able to connect and navigate the tablet with no issues. She walked around and showed the clinician her surroundings; she was particularly proud of her home and garden. The participant was able to show the clinician the places where she stored her medications. After the home video visit, the participant was highly satisfied with the experience, indicating that it felt like an in-person visit.

Personalized Adaptations for Older Adults

For those participants with no digital technological skills, the anthropologist demonstrated how to use the video app before the actual clinical video visit and often had to assist the participant during the video visit by solving small issues such as enabling the video app, increasing the volume, or resolving Wi-Fi connection issues. However, even for those with more digital literacy, telemedicine and video telehealth policies often assume that older adults can seamlessly transition to digital platforms, without accounting for the need for additional training or technical support for practical issues.

For example, 1 participant had an older desktop without a camera that he used for email but had access to an iPad tablet with an integrated camera, which he had never used for email. The anthropologist sent instructions on how to add an email account to his iPad. However, as is common with many older adults, his internet service provider was also his email provider. The participant attempted to add his email to his iPad but gave up, stating, "I'm too old for this."

Adding an email account to a new device can be challenging for older adults who do not have much technical experience (even for those who are tech-savvy, as the instructions can be lengthy and difficult to follow). Similarly, email technology is not always up to date, which introduces potential problems: the interface may be difficult to use or may not work with web browsers. We used FaceTime (Apple Inc) briefly as a practice video call to familiarize the participant with camera and audio controls; the pharmacist visit itself occurred via VVC. The anthropologist assisted with enabling the video app and manually toggled the camera switch when needed. Participants described feeling respected and listened to during pharmacist interactions and emphasized the value of being able to speak "face to face" while accessing medications in their home. These responses reflected shifts in how participants evaluated their own capacity to engage with telehealth and their ability to speak openly with a pharmacist about medication-related questions.

Sociocultural and Environmental Factors Influence Virtual Care Success

Home setup challenges (eg, lack of private space, poor lighting, or technical constraints) impacted the quality and comfort of virtual care interactions. Participants with higher digital literacy or prior exposure to technology were more likely to engage diligently in virtual care, while those with minimal digital experience required more ongoing support. The anthropologist noted that some participants changed their environment by figuring out ways to prop up a device to be hands-free such as stacking books. Ethnographic observation further revealed that negative prior experiences with technology shaped participants' expectations and behaviors. Some participants expressed distrust of digital technologies due to concerns about privacy, hacking, and fraud, citing repeated scam calls or prior financial compromises. These concerns influenced device use practices, such as disabling or covering cameras, frequently changing passwords, or limiting technology use to specific tasks, such as email or banking. Thus, for these participants, reluctance to use video visits was rooted not only in technical skill but also in perceived risk and vulnerability. Family and caregiver involvement played a crucial role in supporting digital engagement. Some participants reported having their spouses, grandchildren, or children assisting with technical support when engaging with technology in general, as one participant noted, "You know who I learned that from? My granddaughter. She knows how to do this."

Discussion

Principal Findings

This formative ethnographic study examined how older adults engaged with video-based telehealth encounters in real-world settings to inform recruitment and implementation strategies for a subsequent telepharmacy randomized clinical trial. Through in-home observation conducted before, during, and after pharmacist video visits, we identified a range of structural, technological, and sociocultural factors that shaped telehealth engagement. Through direct observation, we gained insight into the disconnect between technological design and user needs, highlighting where telehealth platforms fail to accommodate older adults and how digital ageism manifests in real-world health care settings. These insights directly informed the development of recruitment and technology troubleshooting methods for the subsequent RCT focused on pharmacist video visits for medication management among older adults. In addition, although digital ageism was not a core focus of our study, age-related assumptions emerged as an important

XSL•FO
RenderX

interpretative factor that influenced how participants anticipated, experienced, and evaluated their own video visits.

Consistent with prior work, our findings underscore that telehealth engagement among older adults cannot be understood only in terms of individual technical skills or motivation [4-7,9-15]; our in-depth observational approach uniquely highlights both the facilitators and obstacles encountered before and during video visits. Structural barriers such as broadband access, device ownership, and email access intersect with sociocultural factors, including household layout, rurality, and caregiver presence. In-home observation allowed us to identify challenges that would not have been apparent through remote assessment alone, such as difficulty transporting devices within the home, concerns about privacy, and other physical limitations.

Family members frequently served as informal technology facilitators, assisting in device setup, account management, and troubleshooting. This finding highlights the importance of considering digital literacy programs tailored to older adults' needs as potentially dyadic, including community-based initiatives and caregiver-led tutorials [25,35].

Age-related assumptions about technology surfaced during observation and participant interaction. Participants' reluctance to engage with video visits was attributed to older age and lower technical experience. Although successful completion of video visits was frequently accompanied by expressions of surprise or pride, suggesting that telehealth encounters functioned as moments in which age-related self-perceptions were reinforced. These patterns can be further understood through a sociopsychological identity framework, which conceptualizes age and generational status as socially constructed identities rather than fixed deficits. As described by Lee and Hartanto [36], individuals' engagement with technology is shaped not only by skills or access but also by identity-relevant meanings, stereotypes, and expectations associated with belonging to an age group. From this perspective, observed changes in confidence reflect changes in how participants interpreted their own age-related technological competence rather than simple resolution of usability barriers.

It is important to note that not all age-related considerations observed in this study should be interpreted as manifestations of ageism. Some adaptations, such as preparing for visual, hearing, or mobility limitations, enabled and supported successful participation in the study. These practices reflect responsive design rather than ageist assumptions. The analytical challenge lies in distinguishing between accommodations that enhance inclusion and assumptions that lower expectations or discourage engagement. Although similar technology challenges might arise in younger populations with limited technology exposure, the tendency to attribute difficulty to age identity and the meaning attached to this appeared to be specific to older adults' experiences in this context.

Our findings suggest that incorporating ethnographic methods prior to trial implementation can surface otherwise overlooked barriers and facilitators to telehealth engagement among older adults. In-home observation provided actionable insights that informed protocol refinements, recruitment strategies, and support mechanisms for the subsequent randomized trial. Broadly, these methods may help ensure that technology-focused clinical research does not inadvertently exclude older adults by assuming access, skills, or comfort with digital tools.

Limitations

As with most qualitative research, our findings are not broadly generalizable to populations at large. The sample size was small and drawn from veterans willing to participate in a formative telehealth study and thus does not reflect the demographics of older adults in the United States. Given the technology acceptance and uptake during the COVID pandemic, the individuals enrolled in our study likely had additional motivation to learn to use video technology and thus participate in our study. Additionally, as this was a formative qualitative study, findings reflect observed experiences and reported perceptions rather than measured intervention effects. Finally, our hands-on ethnographic approach to recruitment and participant engagement with technology may not be feasible for all studies, especially those that are larger and time limited.

Conclusions

We conducted a focused ethnographic assessment of 20 older individuals engaging in a virtual clinical visit. We illustrate that the application of ethnographic methods to identify and reduce critical barriers to participation and retention in complex health interventions among older adults can yield critical information to improve an RCT. We offered critical insights into how older adults experience telehealth encounters in everyday contexts and revealed the interplay of structural barriers, social support, and age-related self-perceptions. Although digital ageism was not an initial focus of this study, age-related assumptions about technology emerged as a factor shaping engagement with virtual care. Integrating ethnographic approaches and methods into telehealth research and implementation efforts may help improve inclusivity, feasibility, and responsiveness for older adult populations.

Funding

This work was funded by a Veterans Health Administration Health Services Research and Development (IIR Award 18-286; co–principal investigators: WWH and LRM).

Authors' Contributions

MV conceptualized and designed the study, led data collection as the primary ethnographer, conducted the qualitative analysis, and drafted the manuscript. CH contributed to study design, participated in data interpretation, and provided clinical expertise

XSL•FO
RenderX

related to telepharmacy and patient care. MM contributed to the qualitative analytic approach, participated in regular consensus meetings, and supported interpretation of findings. LM provided subject matter expertise in telehealth and aging and contributed to interpretation of results. All authors reviewed, edited, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Structured field note template used to guide ethnographic observations during telehealth visits.
[DOCX File, 21 KB - aging_v9i1e79409_app1.docx]

References

1. Padala KP, Wilson KB, Gauss CH, Stovall JD, Padala PR. VA video connect for clinical care in older adults in a rural state during the COVID-19 pandemic: cross-sectional study. J Med Internet Res 2020 Sep 30;22(9):e21561. [doi: 10.2196/21561] [Medline: 32936773]
2. Weldon AL, Hagemann L. Telehealth use and COVID-19: assessing older veterans' perspectives. Psychol Serv 2023;20(Suppl 2):1-10. [doi: 10.1037/ser0000697] [Medline: 35951392]
3. Albritton J, Ortiz A, Wines R, et al. Video teleconferencing for disease prevention, diagnosis, and treatment: a rapid review. Ann Intern Med 2022 Feb;175(2):256-266. [doi: 10.7326/M21-3511] [Medline: 34871056]
4. Predmore ZS, Roth E, Breslau J, Fischer SH, Uscher-Pines L. Assessment of patient preferences for telehealth in post-COVID-19 pandemic health care. JAMA Netw Open 2021 Dec 1;4(12):e2136405. [doi: 10.1001/jamanetworkopen.2021.36405] [Medline: 34851400]
5. Hogan J, Amspoker AB, Walder A, Hamer J, Lindsay JA, Ecker AH. Differential impact of COVID-19 on the use of tele-mental health among veterans living in urban or rural areas. Psychiatr Serv 2022 Dec 1;73(12):1393-1396. [doi: 10.1176/appi.ps.202100363] [Medline: 35652189]
6. Telehealth: what is it, how to prepare, is it covered? National Institute on Aging, National Institutes of Health. URL: https://www.nia.nih.gov/health/medical-care-and-appointments/telehealth-what-it-how-prepare-it-covered [accessed 2026-04-15]
7. Frydman JL, Li W, Gelfman LP, Liu B. Telemedicine uptake among older adults during the COVID-19 pandemic. Ann Intern Med 2022 Jan;175(1):145-148. [doi: 10.7326/M21-2972] [Medline: 34748380]
8. Samson LW, Tarazi W, Turrini G, Sheingold S. Medicare beneficiaries' use of telehealth in 2020: trends by beneficiary characteristics and location. : Assistant Secretary for Planning and Evaluation; 2021 URL: https://aspe.hhs.gov/reports/medicare-beneficiaries-use-telehealth-2020 [accessed 2026-04-15]
9. Bhatia R, Gilliam E, Aliberti G, et al. Older adults' perspectives on primary care telemedicine during the COVID-19 pandemic. J Am Geriatr Soc 2022 Dec;70(12):3480-3492. [doi: 10.1111/jgs.18035] [Medline: 36169152]
10. Choi NG, DiNitto DM, Marti CN, Choi BY. Telehealth use among older adults during COVID-19: associations with sociodemographic and health characteristics, technology device ownership, and technology learning. J Appl Gerontol 2022 Mar;41(3):600-609. [doi: 10.1177/07334648211047347] [Medline: 34608821]
11. Donaghy E, Atherton H, Hammersley V, et al. Acceptability, benefits, and challenges of video consulting: a qualitative study in primary care. Br J Gen Pract 2019;69(686):e586-e594. [doi: 10.3399/bjgp19X704141] [Medline: 31160368]
12. Levy H, Janke AT, Langa KM. Health literacy and the digital divide among older Americans. J Gen Intern Med 2015 Mar;30(3):284-289. [doi: 10.1007/s11606-014-3069-5] [Medline: 25387437]
13. Mao A, Tam L, Xu A, et al. Barriers to telemedicine video visits for older adults in independent living facilities: mixed methods cross-sectional needs assessment. JMIR Aging 2022 Apr 19;5(2):e34326. [doi: 10.2196/34326] [Medline: 35438648]
14. Thomas-Jacques T, Jamieson T, Shaw J. Telephone, video, equity and access in virtual care. NPJ Digit Med 2021 Nov 18;4(1):159. [doi: 10.1038/s41746-021-00528-y] [Medline: 34795356]
15. Choi EY, Kanthawala S, Kim YS, Lee HY. Urban/rural digital divide exists in older adults: does it vary by racial/ethnic groups? J Appl Gerontol 2022 May;41(5):1348-1356. [doi: 10.1177/07334648211073605] [Medline: 35196918]
16. Moo LR, Schwartz AW. The urgent need for rigorous studies of telehealth for older adults who are homebound. JAMA Netw Open 2021 Sep 1;4(9):e2113451. [doi: 10.1001/jamanetworkopen.2021.13451] [Medline: 34499141]
17. Moo L, Nguyen A, Pugh K, McLaren J, Shirk S, O'Connor M. Technical supports to bridge the digital divide for older adults participating in telemedicine. Innov Aging 2022;6(Supplement_1):7. [doi: 10.1093/geroni/igac059.022]
18. Haimi M. The tragic paradoxical effect of telemedicine on healthcare disparities- a time for redemption: a narrative review. BMC Med Inform Decis Mak 2023 May 16;23(1):95. [doi: 10.1186/s12911-023-02194-4] [Medline: 37193960]
19. Chu CH, Nyrup R, Leslie K, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. Gerontologist 2022 Aug 12;62(7):947-955. [doi: 10.1093/geront/gnab167] [Medline: 35048111]
20. Chu C. Digital ageism and its implications for older adult inclusion. Innov Aging 2023 Dec 21;7(Supplement_1):312-312. [doi: 10.1093/geroni/igad104.1036]

XSL•FO
RenderX

21. Seifert A, Cotten SR, Xie B. A double burden of exclusion? Digital and social exclusion of older adults in times of COVID-19. J Gerontol B Psychol Sci Soc Sci 2021 Feb 17;76(3):e99-e103. [doi: 10.1093/geronb/gbaa098] [Medline: 32672332]
22. Smith-Morris C, Lopez G, Ottomanelli L, Goetz L, Dixon-Lawson K. Ethnography, fidelity, and the evidence that anthropology adds: supplementing the fidelity process in a clinical trial of supported employment. Med Anthropol Q 2014 Jun;28(2):141-161. [doi: 10.1111/maq.12093] [Medline: 24752942]
23. Leff B, Ritchie CS, Rising KL, Cannon K, Wardlow L. Addressing barriers to equitable telehealth for older adults. Front Med (Lausanne) 2025;12:1483366. [doi: 10.3389/fmed.2025.1483366] [Medline: 39917065]
24. Ladin K, Porteny T, Perugini JM, et al. Perceptions of telehealth vs in-person visits among older adults with advanced kidney disease, care partners, and clinicians. JAMA Netw Open 2021 Dec 1;4(12):e2137193. [doi: 10.1001/jamanetworkopen.2021.37193] [Medline: 34870680]
25. Hawley CE, Wagner C, Venegas MD, et al. Connecting the disconnected: Leveraging an in-home team member for video visits for older adults. J Am Geriatr Soc 2024 May;72(5):1408-1419. [doi: 10.1111/jgs.18663] [Medline: 37960887]
26. Morgan-Trimmer S, Wood F. Ethnographic methods for process evaluations of complex health behaviour interventions. Trials 2016 May 4;17(1):232. [doi: 10.1186/s13063-016-1340-2] [Medline: 27142662]
27. O'Cathain A, Thomas KJ, Drabble SJ, Rudolph A, Hewison J. What can qualitative research do for randomised controlled trials? A systematic mapping review. BMJ Open 2013 Jun 20;3(6):e002889. [doi: 10.1136/bmjopen-2013-002889] [Medline: 23794542]
28. Sanjek R. Fieldnotes: The Makings of Anthropology: Cornell University Press; 1990. URL: https://www.jstor.org/stable/10.7591/j.ctvv4124m [accessed 2026-04-16]
29. Fix GM, Kim B, Ruben MA, McCullough MB. Direct observation methods: a practical guide for health researchers. PEC Innov 2022 Dec;1. [doi: 10.1016/j.pecinn.2022.100036] [Medline: 36406296]
30. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
31. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
32. Hamilton AB, Fix GM, Finley EP. Pragmatic Healthcare Ethnography: Methods to Study and Improve Healthcare: Taylor & Francis; 2024.
33. Palinkas LA, Zatzick D. Rapid Assessment Procedure Informed Clinical Ethnography (RAPICE) in pragmatic clinical trials of mental health services implementation: methods and applied case study. Adm Policy Ment Health 2019 Mar;46(2):255-270. [doi: 10.1007/s10488-018-0909-3] [Medline: 30488143]
34. Abraham TH, Finley EP, Drummond KL, et al. A method for developing trustworthiness and preserving richness of qualitative data during team-based analysis of large data sets. Am J Eval 2021;42(1):139-156. [doi: 10.1177/1098214019893784]
35. Jacobs JC, Greene L, SooHoo S, Slightam C, Gujral K, Zulman DM. The role of social support in bridging the digital divide for older veterans. Med Care 2025 May 1;63(5):374-378. [doi: 10.1097/MLR.0000000000002131] [Medline: 39882998]
36. Lee D, Hartanto A. A cultural identity approach to the generational divide. New Ideas Psychol 2025 Dec;79:101183. [doi: 10.1016/j.newideapsych.2025.101183]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
RCT: randomized controlled trial
VA: Department of Veterans Affairs
VVC: VA Video Connect

Edited by F Puga; submitted 20.Jun.2025; peer-reviewed by A Hu, A Hartanto; revised version received 26.Feb.2026; accepted 13.Mar.2026; published 30.Apr.2026.

Please cite as:
Venegas M, McCullough MB, Hawley CE, Beizer JL, Hung WW, Moo LR
Telehealth Barriers and Digital Ageism Experienced by Older Veterans: Formative Ethnographic Study to Inform a Telepharmacy Randomized Trial
JMIR Aging 2026;9:e79409
URL: https://aging.jmir.org/2026/1/e79409
doi:10.2196/79409

XSL•FO
RenderX

© Maria Venegas, Megan B McCullough, Chelsea E Hawley, Judith L Beizer, William W Hung, Lauren R Moo. Originally published in JMIR Aging (https://aging.jmir.org), 30.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

It's Raining, It's Pouring, the Old Man Is Snoring: Content Analysis of Age Stereotypes in Nursery Rhymes

Reuben Ng, PhD; Nicole Indran, BSocSci (Hons)

Lee Kuan Yew School of Public Policy, National University of Singapore, 469C Bukit Timah Road, Singapore, Singapore

Corresponding Author:
Reuben Ng, PhD
Lee Kuan Yew School of Public Policy, National University of Singapore, 469C Bukit Timah Road, Singapore, Singapore

Abstract

Background: Ageist beliefs tend to take root in one's formative years and persist into adulthood, making it crucial to unpack the ways in which children are socialized to view old age. This study is the first to analyze portrayals of old age in nursery rhymes. Related literature has concentrated largely on depictions of older adults in books or movies targeted at children. As a staple of early childhood education, nursery rhymes merit examination as a vehicle through which age stereotypes are disseminated and reinforced.

Objective: Our content analysis of nursery rhymes is grounded in 3 research questions: to what extent is old age represented in nursery rhymes? What are the prevailing stereotypes associated with old age in nursery rhymes? Are these stereotypes primarily positive or negative?

Methods: To build a comprehensive dataset, we gathered material from the following websites, each of which houses an extensive collection of rhymes, including BBC Nursery Rhymes and Songs, the Nursery Rhyme Collections, All Nursery Rhymes, and NurseryRhymes.org. A web ingestion tool was used to compile the data. In total, 735 unique nursery rhymes were retrieved. To identify rhymes related to old age, we conducted a search using various terms related to old age (eg, old, aged, grandfather, and grandma), which yielded 85 nursery rhymes. After applying a rigorous set of exclusion criteria, 29 rhymes remained. Both deductive and inductive modes of reasoning guided our content analysis.

Results: Old age was a central theme in 4% (N=29) of the 735 nursery rhymes. Of the 29 rhymes analyzed, more than half contained negative age stereotypes (17/29, 59%), and a third (10/29, 34%) contained positive ones. A small proportion portrayed old age in a neutral manner (2/29, 7%). Examples of negative stereotypes include being physically debilitated, cognitively impaired, helpless, unhygienic, and incompetent. Examples of positive stereotypes include being wise, affectionate, and jovial. Neutral portrayals framed aging as a natural part of life.

Conclusions: In the context of an aging population, it is paramount that people embrace a less pessimistic outlook on aging. Although nursery rhymes may seem like mere tales not to be taken seriously, they are powerful cultural artifacts capable of molding thought processes. Our study highlights the need to give children access to more accurate and nuanced stories about older adults. By doing so, society can move beyond passively assimilating negative views of aging to actively fostering healthier ones, thus building a future where all are valued regardless of age.

(JMIR Aging 2026;9:e70841) doi:10.2196/70841

KEYWORDS

portrayals of old age; ageism; reframing aging; text as data; psychomics; age stereotypes

Introduction

Background

Ageism refers to stereotyping, discrimination, and prejudice on the grounds of age [1,2]. Globally, one in two individuals aged 16 years and older exhibits ageist attitudes toward older adults [1]. Such ageist beliefs tend to take root in one's formative years and persist into adulthood [3,4], making it crucial to unpack the ways in which children are socialized to view old age. In this study, we analyze depictions of old age in nursery rhymes.

Nursery rhymes are traditional poems, lullabies, or songs typically read or sung to children with the intention of instilling in them positive behaviors [5]. Most well-known nursery rhymes were penned decades or centuries ago but continue to be a staple of early childhood education [6]. On top of being entertaining, nursery rhymes support multiple aspects of a child's development [6], which could explain their enduring appeal. Today, nursery rhymes are enjoyed not only through books but also through television shows and videos, where they have amassed billions of views.

XSL•FO
RenderX

Theoretical Background

Central to the theory of cultivation is the notion that prolonged exposure to the media leads the individual to imbibe the views and values depicted [7,8]. Though originally formulated to elucidate the impact that television can have on viewers, the theory has since evolved to encompass different forms of media, including movies and books [9]. An important aspect of the theory of cultivation is the idea of mainstreaming, which posits that constant exposure to the media can result in a homogenization of perspectives. More specifically, through the process of mainstreaming, audiences who would otherwise hold very different views of the world come to assume a common view of reality [10]. There is evidence that the more people watch television programs that propagate negative messages about older adults, the more they perceive older adults unfavorably [11].

The communicative ecology model of successful aging postulates that perceptions of aging are derived from the routine ways in which aging is discursively constructed, whether through direct or mediated communication [12]. A key plank of this model is that uncertainty about aging influences one's sense of self-efficacy in navigating the aging process. As a form of early socialization, nursery rhymes contribute to this broader communicative ecology, molding perceptions of aging from a young age. When negative messages about aging are conveyed through these rhymes, they may induce feelings of uncertainty about later life. Meanwhile, positive depictions of aging can bolster one's sense of self-efficacy and promote more empowering attitudes toward later life [13].

There is evidence of a general prejudice against older adults among children [14-16]. Negative stereotypes of older persons have been observed in children as young as 3 years old [14,15]. By the time they reach elementary school, they may even verbalize these stereotypes [17]. The theory of stereotype embodiment submits that age stereotypes are internalized throughout one's life span and eventually become self-stereotypes that affect one's health. It is well documented that the assimilation of negative age stereotypes into one's self-concept correlates with poorer health outcomes such as a higher risk of depression and a shorter lifespan [3,18-21]. Meanwhile, positive age stereotypes are associated with improved functional health and overall well-being [3,18,19]. The types of stereotypes espoused early in one's life matter too. Research shows that younger individuals who hold negative age stereotypes are more likely to experience a cardiovascular event in later life than those with positive stereotypes [22]. In addition to being a threat to public health, ageism also constitutes a major economic burden. One study discovered that ageism costs the health care system in the United States approximately US $63 billion every year [23].

Existing Research and Gaps

Thus far, there have not been any studies exploring depictions of old age in nursery rhymes. Instead, related literature has concentrated largely on portrayals of older adults in books or movies targeted at children. The most foundational of these studies is by Ansello [24], whose content analysis of more than 600 children's books revealed that over 80% of them did not feature older characters. When older characters did appear, they were often typecast as sick, feeble, old-fashioned, out-of-touch, and unproductive. Younger characters were described with terms related to physical appearance or personality, while older characters were commonly reduced to their age. When additional descriptors were used, they usually had negative connotations, such as "poor" or "sad." Younger characters were also portrayed as multifaceted and integral to the plot, while older characters were presented as one-dimensional and peripheral to the storyline.

Findings from Ansello's [24] seminal study have since been corroborated by other scholars. In Hurst's analysis of children's picture books [25], 48% of older characters were explicitly called "old," with common descriptors including "nice," "funny," "small," "grumpy," "lonely," and "weak." Dodson and Hause [26] noted that older individuals were frequently depicted as "unhealthy," "ugly," "eccentric," and "passive" in books targeted at readers from kindergarten through adulthood. Janelli [27] argued that the diversity of the older demographic was not captured in picture books. In her review of over 73 storybooks, she noticed that older characters were commonly portrayed as bespectacled, having gray hair, and using walking sticks. Danowki and Robinson's [9] analysis of more than 700 literary works published between 2000 and 2010 in the United States observed that older adults were underrepresented, though there was a marked improvement in the overall portrayal of older characters in picture books over time. Robinson et al [28], conceded that commendable effort had been made to feature more older characters in Disney films. Nonetheless, these characters remained underrepresented and mainly inhabited roles of minimal significance to the plot [28].

This Study

From a conceptual standpoint, this study is the first to look at the ways in which old age is portrayed in nursery rhymes. Stereotypes and attitudes take hold early in life to become stable and influential guiding forces [29]. Besides affecting attitudes toward older people, age stereotypes exact a toll on one's self-concept [3]. From a practical standpoint, this study provides insights that can inform interventions to counteract negative portrayals of old age in nursery rhymes, thus paving the way for a more age-friendly society.

Our content analysis of nursery rhymes is grounded in three research questions: To what extent is old age represented in nursery rhymes? What are the prevailing stereotypes associated with old age in nursery rhymes? Are these stereotypes primarily positive or negative?

Methods

Dataset

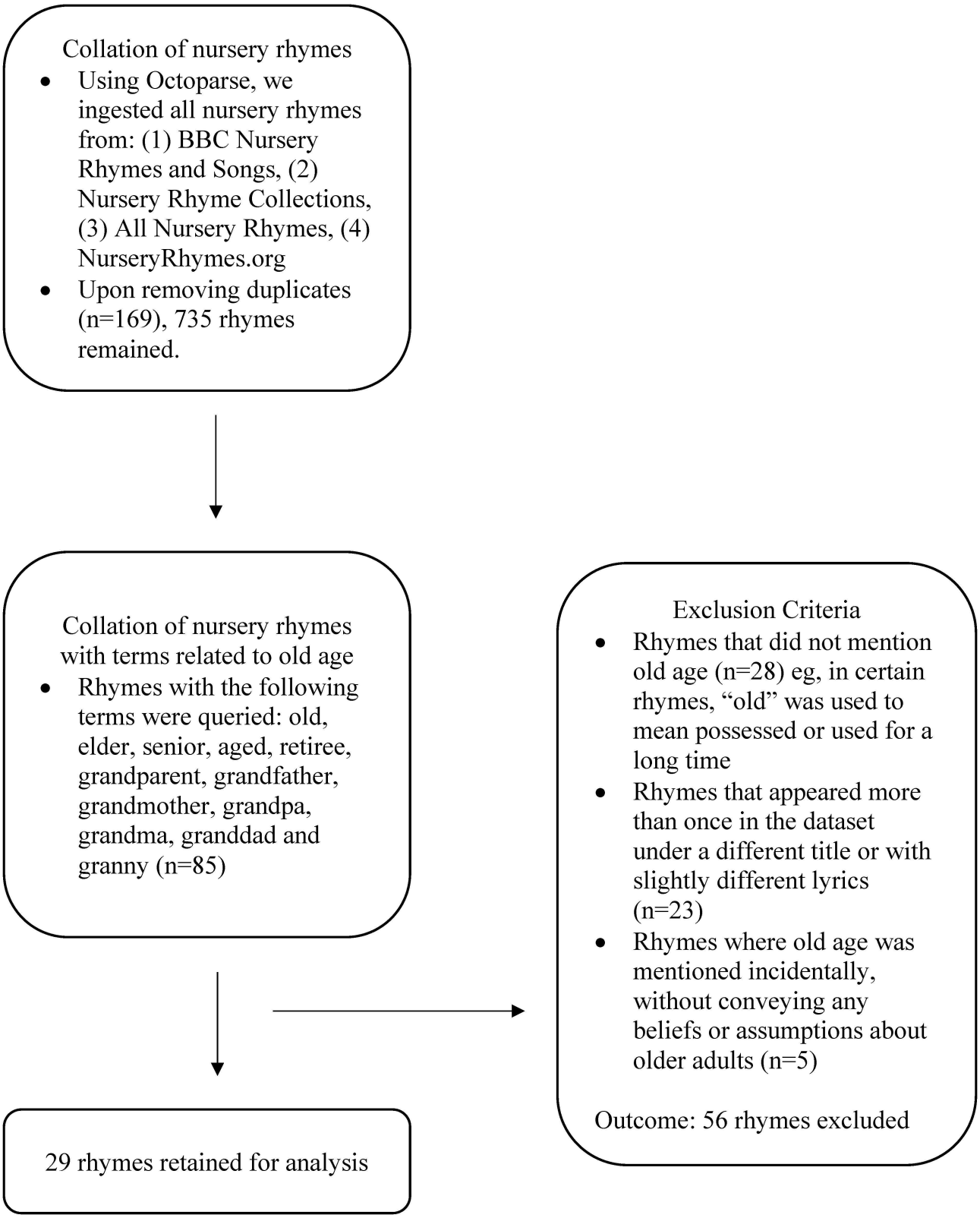
We gathered material from the following websites, each of which houses an extensive collection of rhymes: BBC Nursery Rhymes and Songs [30], the Nursery Rhyme Collections [31], All Nursery Rhymes [32], and NurseryRhymes.org [33]. Octoparse (Octopus Data Inc), a web-ingestion tool, was used to compile the dataset. In total, 904 nursery rhymes were retrieved. After filtering out duplicates (n=169), 735 rhymes

XSL•FO
RenderX

remained. Next, to identify rhymes related to old age, we conducted a search using the following terms: old, elder, senior, aged, retiree, grandparent, grandfather, grandmother, grandpa, grandma, granddad, and granny. This search yielded a total of 85 nursery rhymes. We subsequently removed rhymes where (1) the word “old” was used to mean possessed or used for a long time rather than having lived for a long time (n=28), (2) the rhyme appeared more than once in the dataset under a different title or with minor variations in lyrics (n=23), and (3) old age was mentioned incidentally, without conveying any beliefs or assumptions about the traits of older adults, whether positive, negative, or neutral (n=5). Any rhyme that referenced “old” in relation to old age was retained even if the term was used to describe animals instead of people. This left us with 29 nursery rhymes. Figure 1 depicts the process of collecting the data.

XSL•FO
RenderX

Figure 1. Process of collating nursery rhymes where old age was a central theme.



Content Coding of Nursery Rhymes

Following past scholarship [34-37], we conducted a thematic analysis. The coding rubric was designed using both deductive and inductive modes of reasoning [38]. We started by identifying a set of categories based on previous research [39-42] to develop a preliminary codebook. The analysis was then conducted in several stages, with the words of each rhyme read twice by the authors to ensure familiarity with the data. The aim of the first reading was to ascertain the validity of the initial set of categories, as well as to modify the codebook until all variables were refined and defined clearly. During this first reading, a new category was added whenever a rhyme featured a particular attribute that could not be suitably coded into any of the existing categories and appeared frequently enough to warrant its own category. The aim of the second reading was to make sure we

XSL•FO
RenderX

had a framework sufficiently representative of the different types of rhymes so as to finalize the coding rubric. To ensure rigor in the analysis, the 2 coders had regular discussions during which any discrepancies were reviewed and adjudicated. Areas of major overlap were identified and sectioned into broader themes. The percentage agreement between the 2 raters was 93.5% and the weighted Cohen κ was 0.91 (P<.001), indicating high interrater reliability.

Results

Summary of Insights From Content Analysis

Old age was a central theme in 4% (N=29) of the 735 nursery rhymes. Of the 29 rhymes, more than half contained negative age stereotypes (17/29, 59%) and a third contained positive ones (10/29, 34%). A small proportion portrayed old age in a neutral manner (2/29, 7%; Table 1).

Table . Themes and subthemes in portrayals of old age in nursery rhymes. Percentages may not sum to exactly 100 due to rounding.

Category and themes	Percentage
Negative portrayals of old age	59
Physically debilitated	14
Unlikable	14
Helpless	10
Cognitively impaired	7
Irresponsible	7
Unhygienic	3
Incompetent	3
Positive portrayals of old age	34
Wise	17
Affectionate	10
Jovial	7
Neutral portrayals of old age	7
Aging as natural	7

Negative Portrayals of Old Age (17/29, 59%)

Negative age stereotypes covered include being physically debilitated, cognitively impaired, unhygienic, helpless, unlikable, irresponsible, and incompetent.

Physically Debilitated

“Boom, Boom, Ain’t It Great to Be Crazy?” features a character named Peck, an older doctor who “(falls) down the well and (breaks) his neck.” The older man in “It’s Raining, It’s Pouring” is unable to “get up in the morning” after he “bump(s) his head.” “Michael Finnegan” revolves around a series of misadventures befalling an accident-prone older man. He “(catches) a fish and (drops) it in,” “(climbs) a tree and (barks) his shin,” and “(slides) back down and (scrapes) his skin.”

Cognitively Impaired

“I Know an Old Lady who Swallowed a Fly” depicts the titular character ingesting various creatures, including a “fly,” “spider,” “bird,” and “cat.” The repetitive and nonsensical nature of her actions suggests a possible decline in her mental faculties. In “Old Mother Hubbard,” old age is insinuated as being a time of forgetfulness. Old Mother Hubbard goes to the “fishmonger’s,” “grocer’s,” and “baker’s” but seemingly forgets to buy her dog food.

Unhygienic

Old age is equated with unkemptness in “Desperate Dan.” The eponymous character is classed as a “scruffy old man.” The portrayal of him “wash(ing) his face in a frying pan” and “comb(ing) his hair with a leg of a chair” adds to the depiction of someone who neglects personal hygiene. This eccentric behavior also suggests confusion and possibly cognitive decline or dementia.

Helpless

In “There was an Old Woman (As I’ve Heard Tell),” old age is positioned as a time of helplessness. After the older lady “(falls) asleep on the King’s highway” on her journey to the market, a peddler takes advantage of her by “(cutting) her petticoat up,” leaving her to “shiver and freeze.” She later wakes up disoriented and struggles to recognize herself. She says, “If it be I as I hope it may be,” her dog would be able to recognize her. When the dog barks at her, a sign it does not recognize her, she cries, “Oh dearie me.” In ”Christmas is Coming,” the line “put a penny in the old man’s hat” illustrates an impoverished older person dependent on the goodwill of others.

Unlikable

“The Old Woman Who Lived in a Shoe” describes a mother who provides her children “broth without any bread.” The phrase “[whipping] them all soundly” emphasizes her callous approach to parenting. “Goosey Goosey Gander” presents the older man

XSL•FO
RenderX

as uncooperative as he refuses to "say his prayers," leading the protagonist to take the former "by the left leg and [throw] him down the stairs."

Irresponsible

"Old Mother Leary" portrays the protagonist as behaving impetuously. She "[leaves] a lantern in the shed," which results in a fire when a "cow [kicks] it over." Her wink as the fire breaks out implies she took pleasure in the ensuing chaos.

Incompetent

In the sea shanty "Blow the Man Down," the "old skipper" is derided as incompetent. He is described as someone who "once played a remarkable game" but is now unable to navigate his ship, which is "becalmed in the tropical sea." The lyrics further underscore his nautical ineptitude by stating that he "whistled all day, but in vain for a breeze," hinting at his inability to get the ship moving.

Positive Portrayals of Old Age (10/29, 34%)

Positive age stereotypes covered include being wise, affectionate, and jovial.

Wise

In "A Wise Old Owl," the phrase "the more he saw the less he spoke, the less he spoke the more he heard" conveys that with age comes wisdom and the ability to listen and observe. The concluding line, "why can't we all be like that wise old bird?" demonstrates that the owl's qualities of being observant and thoughtful are worth emulating. Likewise, in "Dance Ter Yer Daddy," old age is celebrated as a stage of life where one bequeaths knowledge and experiences to the next generation. This is evident from the lines "when yer an old man, father to a son, sing to him the old songs, sing of all you've done, pass along the old ways, then let his song begin, dance to your daddy, ma bonnie lamb."

"Manx Lullaby" exalts old age as a time where the accrual of experience and knowledge offers guidance and insight. The lines "Through your life there are many, many, many tales, gems and stories you'll be told. Remember these songs and these rhymes as you grow old; I mean, the truth will never leave you" communicate the idea that the stories and experiences one gathers throughout life should be cherished. "If All the World was Apple Pie" opens with a hypothetical inquiry, "If all the world was apple pie, and all the sea was ink, and all the trees were bread and cheese, what should we have for drink?" The subsequent line, "it's enough to have an old man scratch his head and think," references the intellectual prowess of older individuals by suggesting that such a perplexing question would even stump someone of advanced age.

Affectionate

When a caterpillar crawls on the protagonist's grandmother in "Little Arabella Miller," she proclaims, "Arabella Miller, how I love your caterpillar!" In "Over the River and Through the Woods," the lines "when grandmother sees us come, she will say, "Oh, dear, the children are here, bring a pie for everyone" show a grandmother ever ready to indulge her grandchildren with treats.

Jovial

The central figure in "Old King Cole" is characterized as a "merry old soul." The verse notes that he "called for his pipe and he called for his bowl and he called for his fiddlers three," highlighting his zest for life. In "This Old Man," the refrain "played knick-knack" captures the character's spirited involvement in the world around him as he enthusiastically taps on different objects while engaging with others in a light-hearted manner.

Neutral Portrayals of Old Age (2/29, 7%)

Neutral portrayals depicted old age as a natural part of life. In "Grandfather's Clock," the clock symbolizes the passage of the life of the protagonist's grandfather. Significant life events are recounted, such as "when he entered at the door with a blooming and beautiful bride." The narrative concludes without judgment, stating that the clock "stopped short, never to go again, when the old man died." In "Do Do, Baby Do," the line "daddy old hen dozes over 'neath the roses, tiny chicks she'll have for you" situates old age within the larger context of family life. It suggests that aging is part of a natural process, with one generation giving way to the next.

Discussion

Principal Findings

This study delved into representations of old age in nursery rhymes. We found that only 4% of the nursery rhymes compiled mentioned old age. Of these, over half presented old age negatively, while about a third presented it positively. These findings comport with past literature about age stereotypes being multidimensional, albeit with a tendency toward negativity [14-16,43].

Considering the widespread tendency to overlook the experiences and contributions of older adults, it is unsurprising—though no less concerning—that just 4% of the nursery rhymes mentioned old age. This figure is especially stark given that by 2030, one in 6 people globally will be at least 60 years of age [44]. The issue of underrepresenting or, worse, excluding the older demographic has been flagged in other forms of children's media [9,24-28] as well as popular culture in general [45-48]. This erasure of older adults in the media in turn contributes to their invisibility in wider society.

The presence of negative age stereotypes in more than half of the rhymes raises important concerns about their potential impact on children's beliefs. It has been established that children tend to harbor certain negative stereotypes of older adults [39,40,42]. As per the theory of cultivation, repeated exposure to negative messages about older adults can influence children's perceptions of reality [7-11]. Besides perpetuating ageism and stymying the development of intergenerational solidarity, it also impacts them personally. From the perspective of the communicative ecology model of successful aging [12], negative portrayals of older adults can affect children's self-efficacy regarding their aging process. In addition, consistent with the theory of stereotype embodiment, these narratives can have detrimental effects on their mental and physical health [22].

XSL•FO
RenderX

Old age was cast in a positive light in 34% of the rhymes. These positive depictions offer a much-needed counterbalance to the ageist portrayals replete in popular culture [36,49-52]. In the endeavor to enhance representations of old age, it is crucial to bear in mind that the objective should not be for all rhymes to exclusively endorse positive age stereotypes. Overly romanticized portrayals of old age may wind up creating unrealistic expectations not in sync with the diverse realities of the aging experience [53-56]. Instead, the goal should be to portray old age in a balanced and authentic manner that acknowledges the heterogeneity of the older cohort, as well as the complexities inherent in the aging process [57].

Practical Implications

We propose a set of interventions to improve portrayals of old age in nursery rhymes.

First, new nursery rhymes could be composed to challenge negative age stereotypes. Jan Golden's creation of "Age-Friendly Vibes," a line of greeting cards developed to counter the ageism rampant in birthday cards, is an example of how new forms of expression are necessary to disrupt ageist thinking [58,59]. The creation of new rhymes could involve collaborations among writers, educators, creative professionals, gerontologists, and older adults themselves. These rhymes could be incorporated into educational materials, children's books, and digital platforms to ensure widespread accessibility and exposure. Efforts could also be made to revise existing nursery rhymes to depict the older population more positively, though it is essential to consider how such modifications might be received by educators and parents.

Second, curriculum developers should be mindful not to perpetuate negative attitudes toward older adults. There is evidence that stereotypes often take hold in one's childhood years and carry on into adulthood [29]. Educational materials should therefore be thoughtfully designed to foster understanding toward older individuals.

Third, intergenerational storytelling and performance sessions could be organized in schools, libraries, community centers, or senior living facilities. These sessions would create opportunities for older individuals to share their stories, experiences, and talents with their younger contemporaries. Crucially, these sessions encourage the reciprocal exchange of knowledge and skills while also enhancing the well-being of both groups [60,61].

Fourth, introducing opportunities for collaborative projects across age groups could bridge generational gaps. By encouraging joint artistic endeavors, both generations can contribute their perspectives, talents, and insights, thereby facilitating intergenerational learning [60,61].

Finally, there is a need to provide parents and caregivers with educational resources on ageism. As these individuals usually select the materials children are exposed to, it is crucial that they choose content that advances a more inclusive view of aging. Equipping them with the resources to challenge ageism will create a lasting impact on children's attitudes toward aging.

Limitations and Future Directions

This study has several limitations. First, since most of the rhymes in our dataset were written decades or even centuries ago, the sentiments expressed may be more representative of past attitudes than contemporary ones. Nonetheless, these rhymes warrant scrutiny for their continued relevance today. In addition, we did not manage to locate the country of origin and year of publication for all the rhymes, which limits our ability to fully contextualize them. Second, most, if not all, of the nursery rhymes in our dataset seem to have originated in the West, which means the depictions of old age likely reflect Western perspectives. Since norms of old age vary across cultures [62,63], future studies could investigate the depictions of old age in popular children's poems or songs from other countries. Third, as we were unable to determine authorial intent, our interpretation of each rhyme was inherently subjective. Nonetheless, the high interrater reliability supports the robustness of our conclusions.

To build on the findings of this study, subsequent investigations could dive into the illustrations that accompany nursery rhymes in books or videos. This will generate insights into the visual cues and symbols associated with old age in nursery rhymes. Exploring children's perceptions of depictions of older adults in nursery rhymes is another line of research worth pursuing. It would also be valuable to conduct a historical analysis of rhymes to determine whether newer rhymes contain more negative stereotypes than older ones, for instance. Finally, future research could compare nursery rhymes or other content targeted at children across different societies to examine whether there are cultural differences in the way aging is portrayed.

Conclusion

In the context of an aging population, it is paramount that people embrace a less pessimistic outlook on aging. Although nursery rhymes may seem like mere tales not to be taken seriously, they are powerful cultural artifacts capable of molding thought processes. Our study highlights the need to give children access to more accurate and nuanced stories about older adults. By doing so, society can move beyond passively assimilating negative views of aging to actively fostering healthier ones, thus building a future where all are valued regardless of age.

Acknowledgments

We appreciate LL's assistance in data collection.

XSL•FO
RenderX

Funding

We gratefully acknowledge the Commonwealth Fund's Harkness Fellowship in Health Care Policy and Practice awarded to RN. The funder had no role in study design, data collection, analysis, writing, and decision to publish.

Data Availability

Data are publicly available at BBC Nursery Rhymes and Songs, the Nursery Rhyme Collections, All Nursery Rhymes, and NurseryRhymes.org.

Authors' Contributions

Conceptualization: RN, NI
Formal analysis: RN, NI
Funding acquisition: RN
Resources: RN
Supervision: RN
Writing – original draft: RN, NI
Writing – review & editing: RN, NI

Conflicts of Interest

None declared.

References

1. Officer A, Thiyagarajan JA, Schneiders ML, Nash P, de la Fuente-Núñez V. Ageism, healthy life expectancy and population ageing: how are they related? Int J Environ Res Public Health 2020 May 1;17(9):3159. [doi: 10.3390/ijerph17093159] [Medline: 32370093]
2. Butler RN. Age-ism: another form of bigotry. Gerontologist 1969;9(4):243-246. [doi: 10.1093/geront/9.4_part_1.243] [Medline: 5366225]
3. Levy B. Stereotype embodiment: a psychosocial approach to aging. Curr Dir Psychol Sci 2009 Dec 1;18(6):332-336. [doi: 10.1111/j.1467-8721.2009.01662.x] [Medline: 20802838]
4. Bellingtier JA, Schenker LE, Weber AL. Judging a book by its older adult cover: age-related expectations and parental preference for children's books. Curr Psychol 2024 Jan;43(1):774-786. [doi: 10.1007/s12144-023-04298-6]
5. Patton DA. We all fall down: head injuries in nursery rhyme characters. BMJ 2021 Dec 15;375:e068256. [doi: 10.1136/bmj-2021-068256] [Medline: 34911678]
6. Mullen G. More than words: using nursery rhymes and songs to support domains of child development. J Child Stud 2017 Sep;42(2):42. [doi: 10.18357/jcs.v42i2.17841]
7. Gerbner G, Gross L, Morgan M, Signorielli N, Shanahan J. Growing up with television: cultivation processes. In: Bryant J, Zillmann D, editors. Media Effects: Advances in Theory and Research, 2nd edition: Lawrence Erlbaum Associates Publishers; 2002:43-67.
8. Gerbner G, Morgan M, Signorielli N. Living with television: the dynamics of the cu itivation process. In: Bryant J, Zillman D, editors. Perspectives on Media Effects Hilldale: Lawrence Erlbaum Associates; 1986:17-40 URL: https://www.semanticscholar.org/paper/Living-with-Television%3A-The-Dynamics-of-the-Cu/d376313acb16fcfa8160434ac01a23a8e44cb2dc [accessed 2025-03-29]
9. Danowski J, Robinson T. The portrayal of older characters in popular children's picture books in the US. J Child Media 2012 Aug;6(3):333-350. [doi: 10.1080/17482798.2011.630741]
10. Shrum J, Bischak D. Mainstreaming, resonance, and impersonal impact. Human Comm Res 2001 Apr;27(2):187-215 [FREE Full text] [doi: 10.1111/j.1468-2958.2001.tb00780.x]
11. Gerbner G, Gross L, Signorielli N, Morgan M. Aging with television: images on television drama and conceptions of social reality. J Commun 1980;30(1):37-47. [doi: 10.1111/j.1460-2466.1980.tb01766.x] [Medline: 7372841]
12. Fowler C, Gasiorek J, Giles H. The role of communication in aging well: introducing the communicative ecology model of successful aging. Commun Monogr 2015 Oct 2;82(4):431-457. [doi: 10.1080/03637751.2015.1024701]
13. Gettings PE, Kuang K. Extending the communicative ecology model of successful aging using talk about careers and retirement. Commun Monogr 2022 Oct 2;89(4):493-514. [doi: 10.1080/03637751.2021.2025410]
14. Aday RH, Aday KL, Arnold JL, Bendix SL. Changing children's perceptions of the elderly: the effects of intergenerational contact. Gerontol Geriatr Educ 1996;16:37-51. [doi: 10.1300/J021v16n03_04] [Medline: 23621412]
15. Rich PE, Myrick RD, Campbell C. Changing children's perceptions of the elderly. Educ Gerontol 1983 Sep;9(5-6):483-491. [doi: 10.1080/0380127830090512]
16. McTavish DG. Perceptions of old people: a review of research methodologies and findings. Gerontologist 1971;11(4 Part 2):Suppl. [doi: 10.1093/geront/11.4_part_2.90] [Medline: 4940690]

XSL•FO
RenderX

17. Isaacs LW, Bearison DJ. The development of children's prejudice against the aged. Int J Aging Hum Dev 1986;23(3):175-194. [doi: 10.2190/8GVR-XJQY-LFTH-E0A1] [Medline: 3557636]
18. Levy BR, Hausdorff JM, Hencke R, Wei JY. Reducing cardiovascular stress with positive self-stereotypes of aging. J Gerontol B Psychol Sci Soc Sci 2000 Jul;55(4):205-213. [doi: 10.1093/geronb/55.4.p205] [Medline: 11584877]
19. Levy BR, Slade MD, Kasl SV. Longitudinal benefit of positive self-perceptions of aging on functional health. J Gerontol B Psychol Sci Soc Sci 2002 Sep;57(5):409-417. [doi: 10.1093/geronb/57.5.p409] [Medline: 12198099]
20. Brothers A, Kornadt AE, Nehrkorn-Bailey A, Wahl HW, Diehl M. The effects of age stereotypes on physical and mental health are mediated by self-perceptions of aging. J Gerontol B Psychol Sci Soc Sci 2021 Apr 23;76(5):845-857. [doi: 10.1093/geronb/gbaa176] [Medline: 33057726]
21. Chang ES, Kannoth S, Levy S, Wang SY, Lee JE, Levy BR. Global reach of ageism on older persons' health: a systematic review. PLOS ONE 2020;15(1):e0220857. [doi: 10.1371/journal.pone.0220857] [Medline: 31940338]
22. Levy B. Stereotype embodiment. Curr Dir Psychol Sci 2009 Dec;18(6):332-336. [doi: 10.1111/j.1467-8721.2009.01662.x] [Medline: 20802838]
23. Levy BR, Slade MD, Chang ES, Kannoth S, Wang SY. Ageism amplifies cost and prevalence of health conditions. Gerontologist 2020 Jan 24;60(1):174-181. [doi: 10.1093/geront/gny131] [Medline: 30423119]
24. Ansello EF. Age and ageism in children's first literature. Educ Gerontol 1977 Jan;2(3):255-274. [doi: 10.1080/0360127770020305]
25. Hurst J. Images in Children's Picture Books 1981. URL: https://www.semanticscholar.org/paper/Images-in-Children%27s-Picture-Books.-Hurst/5c6ec5ed16d35ce13c9881535584797da20196e7 [accessed 2023-05-27]
26. Dodson AE, Hause JB. Ageism in literature. an analysis kit for teachers and librarians. In: Teaching and Learning about Aging Project, McCarthy-Towne School, Charter Road, Acton, MA 01720: Acton-Boxborough Regional School District; 1981. URL: https://eric.ed.gov/?id=ED211411 [accessed 2023-05-27]
27. Janelli LM. Depictions of grandparents in children's literature. Educ Gerontol 1988 Jan;14(3):193-202. [doi: 10.1080/0380127880140303]
28. Robinson T, Callister M, Magoffin D, Moore J. The portrayal of older characters in Disney animated films. J Aging Stud 2007 Aug;21(3):203-213. [doi: 10.1016/j.jaging.2006.10.001]
29. Klausmeier HJ, Goodwin W. Learning and Human Abilities: Educational Psychology: Harper; 1966. URL: https://eduq.info/xmlui/handle/11515/16557 [accessed 2023-06-09]
30. BBC nursery rhymes and songs - a to z BBC school radio. BBC School Radio. URL: https://www.bbc.co.uk/teach/school-radio/nursery-rhymes-a-to-z-index/z4ddgwx [accessed 2023-06-15]
31. All nursery rhymes in alphabetical order. The Nursery Rhyme Collections. URL: https://nurseryrhymescollections.com/lyrics/songlist.html [accessed 2023-06-15]
32. Nursery rhymes – popular nursery rhymes with lyrics. All Nursery Rhymes. URL: https://allnurseryrhymes.com/ [accessed 2023-06-15]
33. NurseryRhymes.org - nursery rhymes with lyrics and music. NurseryRhymes.org. URL: https://www.nurseryrhymes.org/ [accessed 2023-06-15]
34. Ng R, Indran N. Not too old for TikTok: how older adults are reframing aging. Gerontologist 2022 Sep 7;62(8):1207-1216. [doi: 10.1093/geront/gnac055] [Medline: 35512401]
35. Ng R, Indran N. Questions about aging and later life on Quora. Gerontologist 2024 Aug 1;64(8):gnae060. [doi: 10.1093/geront/gnae060] [Medline: 38813784]
36. Ng R, Indran N. Does age matter? Tweets about gerontocracy in the United States. J Gerontol 2023 Nov 14;78(11):1870-1878. [doi: 10.1093/geronb/gbad113]
37. Ng R, Indran N. Public interest in research on aging: analysis of altmetric attention scores over 5 years. J Appl Gerontol 2025 Jan;44(1):106-115. [doi: 10.1177/07334648241265191] [Medline: 39030665]
38. Armat MR, Assarroudi A, Rad M, Sharifi H, Heydari A. Inductive and deductive: ambiguous labels in qualitative content analysis. TQR . [doi: 10.46743/2160-3715/2018.2872]
39. Goldman RJ, Goldman JDG. How children view old people and ageing: a developmental study of children in four countries. Aust J Psychol 1981 Dec 1;33(3):405-418. [doi: 10.1080/00049538108254708]
40. Lynott PP, Merola PR. Improving the attitudes of 4th graders toward older people through a multidimensional intergenerational program. Educ Gerontol 2007 Jan;33(1):63-74. [doi: 10.1080/03601270600864041]
41. Seefeldt C, Jantz RK, Galper A, Serock K. Using pictures to explore children's attitudes toward the elderly. Gerontologist 1977 Dec;17(6):506-512. [doi: 10.1093/geront/17.6.506] [Medline: 924172]
42. Weinberger A. Stereotyping of the elderly: elementary school children's responses. Res Aging 1979:113-136. [doi: 10.1177/016402757911007]
43. Ng R, Indran N, Liu L. Social media discourse on ageism, sexism, and racism: analysis of 150 million tweets over 15 years. J Am Geriatr Soc 2024 Oct;72(10):3149-3155. [doi: 10.1111/jgs.19047] [Medline: 38958541]
44. Ageing and health. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2021-05-26]

XSL•FO
RenderX

45. Bishop JM, Krause DR. Depictions of aging and old age on Saturday morning television. Gerontologist 1984 Feb;24(1):91-94. [doi: 10.1093/geront/24.1.91] [Medline: 6698422]
46. Laughlin CA. Images of Aging and the Aesthetic of Actuality in Chinese Film: Reportage, Documentary, and the Art of the Real. Mod Chin Lit Cult 2019;31(2):207-248.
47. Lien SC, Zhang YB, Hummert ML. Older adults in prime-time television dramas in Taiwan: prevalence, portrayal, and communication interaction. J Cross Cult Gerontol 2009 Dec;24(4):355-372. [doi: 10.1007/s10823-009-9100-3] [Medline: 19757005]
48. Shary T, McVittie N. Fade to Gray: Aging in American Cinema: University of Texas Press; 2016.
49. Berger R. Aging in America: ageism and general attitudes toward growing old and the elderly. JSS 2017;05(8):183-198. [doi: 10.4236/jss.2017.58015] [Medline: 27928907]
50. Indran N, Ng R. Artistic provocation or ageist stereotyping? J Am Geriatr Soc 2025 Jul 15. [doi: 10.1111/jgs.70003] [Medline: 40660859]
51. Ng R, Indran N. Age advocacy on Twitter over 12 years. Gerontologist 2024 Jan 1;64(1):gnac183. [doi: 10.1093/geront/gnac183] [Medline: 36516463]
52. Ng R, Indran N. Videos about older adults on TikTok. PLOS ONE 2023;18(8):e0285987. [doi: 10.1371/journal.pone.0285987] [Medline: 37531317]
53. Ng R, Indran N. Reframing aging: foregrounding familial and occupational roles of older adults is linked to decreased ageism over two centuries. J Aging Soc Policy 2024 Sep 2;36(5):749-766. [doi: 10.1080/08959420.2023.2238538] [Medline: 37534573]
54. Ng R, Indran N. Impact of old age on an occupation's image over 210 years: an age premium for doctors, lawyers, and soldiers. J Appl Gerontol 2023 Jun;42(6):1345-1355. [doi: 10.1177/07334648231155025] [Medline: 37092180]
55. Ng R, Indran N. Youth is prized in medicine, old age is valued in law: analysis of media narratives over 200 years. J Med Internet Res 2024 Mar 26;26(1):e45855. [doi: 10.2196/45855] [Medline: 38530338]
56. Ng R, Indran N, Liu L. Advocating for older adults in the age of social media: strategies to achieve peak engagement on Twitter. JMIR Aging 2024 May 1;7:e49608. [doi: 10.2196/49608] [Medline: 38715315]
57. Ng R, Indran N. Role-based framing of older adults linked to decreased ageism over 210 years: evidence from a 600-million-word historical corpus. Gerontologist 2022 Apr 20;62(4):589-597. [doi: 10.1093/geront/gnab108] [Medline: 34323967]
58. Williams M. Denver designer unveils range of birthday cards with positive messages about getting older after claiming "damn, you're old" type rivals are driving the elderly into an early grave. Daily Mail Online. 2024 Feb 17. URL: https://www.dailymail.co.uk/news/article-13094913/birthday-ageism-woke-age-friendly-vibes-greeting-card.html [accessed 2025-03-29]
59. Lin SSH, Walden A. Ageism in birthday cards: a mixed-method content analysis. Gerontologist 2024 May 1;64(5):gnad115. [doi: 10.1093/geront/gnad115] [Medline: 37675965]
60. Boger J, Mercer K. Technology for fostering intergenerational connectivity: scoping review protocol. Syst Rev 2017 Dec 8;6(1):250. [doi: 10.1186/s13643-017-0652-y] [Medline: 29216904]
61. Burnes D, Sheppard C, Henderson CR Jr, et al. Interventions to reduce ageism against older adults: a systematic review and meta-analysis. Am J Public Health 2019 Aug;109(8):e1-e9. [doi: 10.2105/AJPH.2019.305123] [Medline: 31219720]
62. Ng R, Indran N. Societal perceptions of caregivers linked to culture across 20 countries: evidence from a 10-billion-word database. PLOS ONE 2021;16(7):e0251161. [doi: 10.1371/journal.pone.0251161] [Medline: 34197470]
63. Ng R, Indran N. #ProtectOurElders: analysis of Tweets about older Asian Americans and anti-Asian sentiments during the COVID-19 pandemic. J Med Internet Res 2024 Mar 29;26(1):e45864. [doi: 10.2196/45864] [Medline: 38551624]

Edited by M Gray; submitted 03.Jan.2025; peer-reviewed by KA Price, M Wettstein, N Dalmer, O Moreno-Fernandez; revised version received 01.Aug.2025; accepted 08.Aug.2025; published 17.Feb.2026.

Please cite as:
Ng R, Indran N
It's Raining, It's Pouring, the Old Man Is Snoring: Content Analysis of Age Stereotypes in Nursery Rhymes
JMIR Aging 2026;9:e70841
URL: https://aging.jmir.org/2026/1/e70841
doi:10.2196/70841

© Reuben Ng, Nicole Indran. Originally published in JMIR Aging (https://aging.jmir.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR

XSL•FO
RenderX

Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Integrating Community and Digital Support Through Social Prescribing to Improve Mental Health in Rural Older Adults in South Korea: Quasi-Experimental Study

Bo Zhao[1,2], PhD; Joo Aeh Lim[2], MPH; Subeen Lim[2], PhD; Hye Jin Yoon[3], MPH; Hae Kyung Kim[4], MS; Seok Mi Hong[5], MS; Xinyu Zhang[4], MPH; Eun Woo Nam[2,6], PhD

[1]Department of Health Services Management, School of Health Services Management, Anhui Medical University, Hefei, China

[2]Yonsei Global Health Center, Yonsei University, Wonju, Republic of Korea

[3]Gangwon Public Health Policy Institute, Wonju, Republic of Korea

[4]Graduate School of Health Science, Yonsei University, Wonju, Republic of Korea

[5]Saje Primary Health Care Post, Wonju, Republic of Korea

[6]Department of Health Administration, Graduate School, Yonsei University, 1 Yonseidae-gil, Wonju, Republic of Korea

Corresponding Author:
Eun Woo Nam, PhD
Yonsei Global Health Center, Yonsei University, Wonju, Republic of Korea

Abstract

Background: Mental health issues such as depression, loneliness, and cognitive decline are prevalent among older adults. They are particularly pronounced in rural areas due to socioeconomic disadvantages, limited health literacy, and social isolation. These challenges have been exacerbated by the COVID-19 pandemic, highlighting the urgent need for accessible, community-based mental health strategies. Social prescribing—linking individuals to nonmedical community resources—offers a promising model, especially when coupled with culturally relevant and digitally facilitated interventions.

Objective: This study aimed to evaluate the effectiveness of a pilot social prescribing project for individuals aged 65 years and older residing in rural South Korea. Specifically, it examined the changes in social support, loneliness, depression, cognitive function, and life satisfaction among the older adults before and after the intervention and the relationships among these factors.

Methods: A quasi-experimental, nonequivalent control group pretest-posttest design was conducted with 294 participants from Wonju City, Gangwon-do (n=148 in the experimental group, n=146 in the control group). The 8-week intervention featured a community-based music storytelling program incorporating local cultural elements and digital education. Statistical analyses, including *t* tests, difference-in-differences analysis, and structural equation modeling, were conducted to evaluate changes in scores and the effectiveness of the program intervention.

Results: The study confirmed that the social prescribing pilot project effectively reduced depression and increased social support and life satisfaction compared with the control group (P<.05, 95% CI). While cognitive function showed slight improvement, the change was not statistically significant. Social support emerged as a key mediating factor, positively influencing cognitive function and life satisfaction, and was negatively associated with depression. There was no statistically significant direct effect found between loneliness and cognitive function or social support.

Conclusions: This pilot study supports the feasibility and mental health benefits of integrating community and digital support through culturally embedded social prescribing for rural older adults. The findings highlight the importance of leveraging local resources and social networks to address mental health disparities in underserved aging populations, offering valuable insights for policymakers and practitioners developing inclusive aging and health promotion strategies.

International Registered Report Identifier (IRRID): RR2-10.2196/46371

(JMIR Aging 2026;9:e76201) doi:10.2196/76201

KEYWORDS

social prescribing; rural older adults; mental health; social support; cognitive function; community-based intervention; digital health; culturally tailored care.

XSL•FO
RenderX

Introduction

Developed countries like South Korea (hereafter "Korea") are now experiencing super-aging speed and an increasing number of noncommunicable diseases with long-term conditions [1]. Well-known noncommunicable diseases such as cancer, diabetes, heart disease, Alzheimer disease, dementia, and mental health ailments pose devastating health consequences for individuals, families, and communities and threaten to overwhelm health systems [2]. Disease-related risk factors include age, genes (inherited risks), long-term health conditions, lifestyle factors, sex and gender, cognitive function, and so on. Korea is a "super-aged society," in which over 20% of the total population is 65 years or older [3]. Mental health issues such as depression, loneliness, and suicide ideation have been reported to increase rapidly with soaring cognitive dysfunction among Korean elders. Meanwhile, the older adults have the highest poverty rate, and the suicide rate in South Korea is the highest among Organization for Economic Co-operation and Development countries [4]. This implies that the burden of psychological diseases in older Koreans is rapidly increasing, and the clinical approach alone cannot solve these problems. More innovative interventions should be included in the nonclinical approach to disease prevention and control for a healthier life [5].

Living alone or loneliness and a lack of social support are widely recognized risk factors for psychological disorders such as suicide, depression, and anxiety in later life [6]. These challenges are often exacerbated by life events like retirement or the loss of a loved one, which lead to smaller social networks and increased isolation [7]. The mental health of older individuals living alone is notably worse than that of those who cohabit, and those without spouses are at an even higher risk of depression [8]. During the COVID-19 pandemic, loneliness became more prevalent, with older adults being twice as likely to experience depressive symptoms compared to the prepandemic period, especially among those with lower socioeconomic status or existing health vulnerabilities [9-11]. These conditions not only increase the likelihood of depression and anxiety but also contribute to functional decline, disability, decreased life satisfaction, and even higher mortality rates [12-15]. The COVID-19 pandemic has highlighted the need for robust care support systems and high-quality life for the older adults, especially in addressing issues such as social isolation and loneliness.

Additionally, subjective health status, economic level, and literacy are crucial determinants of psychological health among the older adults. How older adults perceive their health often reflects their mental state, with positive perceptions being linked to better mental well-being and negative ones to poorer outcomes [16]. Economic conditions also play a pivotal role; older individuals with higher subjective economic levels and those receiving financial support from family members report higher levels of life satisfaction and lower risks of depression [17]. Notably, literacy significantly impacts mental health, particularly in rural older populations where low literacy or illiteracy is common. Poor literacy not only limits social participation but also reduces the effectiveness of interventions like social prescribing programs, hindering efforts to improve mental health outcomes in these communities [16]. Addressing these interconnected factors is essential for promoting better mental health among older adults.

Social prescribing originated in the United Kingdom in 2016 as an approach to connect patients or clients with nonclinical services in their local communities, addressing socioeconomic and psychosocial factors that influence health. The social prescribing model enables primary care professionals, particularly general doctors, to refer patients to various community-based activities and services, such as volunteering, gardening, and cooking, which voluntary and community organizations provide [18]. These services are designed to enhance the well-being of individuals, while simultaneously reducing the strain on conventional health care systems. The social prescribing structure has evolved over time. It generally consists of 3 key elements: the health care professional who makes the initial referral, community organizations offering support, and a link worker who facilitates the connection between the patient and the services [19]. The link worker works closely with the patient to identify their needs and goals, offering motivation and support throughout the care process. Recently, social prescribing has expanded internationally and is currently being developed in more than 30 countries and regions [20].

Social prescribing has been shown to benefit older adults [21,22]; however, given the diverse realities of aging, such as bereavement and cognitive decline, personalized interventions are essential [23]. Generic approaches often fail to meet individual needs, highlighting the importance of tailoring community care to specific situations and preferences. Integrating person-centered principles into social prescribing design can enhance intervention effectiveness by co-designing with individuals in ways that honor local culture and increase their sense of control, involvement, and engagement [24]. Around the world, social prescribing models vary but consistently reflect this emphasis on personalized care [25]. In the United Kingdom, social prescribing is integrated into the health care system, linking primary care professionals to community resources for social and psychological support that respect each person's cultural and social context [26]. In Japan, social prescribing initiatives focus on mental health and social isolation, as highlighted in a recent study that implemented social prescribing practices in local settings to improve community mental health and well-being [27]. Research on Canada's social prescribing in rural and underserved communities emphasizes the importance of local partnerships in providing social support and reducing isolation among older adults [28]. These diverse social prescribing approaches, though varying across countries, enhance older adults' well-being by prioritizing person-centered care, fostering greater acceptance and engagement [29].

Adopting the UK's concept at the beginning of 2019, Korea emerged as a notable example of a nongovernment-led program, with social prescribing being tested several times for older adults in a rural area [30]. This pilot project was led by a university and funded by the National School Research Foundation. Based on the special characteristics of the older adults in this rural area, the proportion of older people living alone who are

XSL•FO
RenderX

depressed is relatively high, and a 10-week social prescribing program that includes musical storytelling, community gardening activities [31], and support groups has been developed [32]. Previous qualitative studies reported positive results in reduced depression, improved self-esteem, increased vitality through intergenerational interactions, and expanded social relationships [33]. Reduced loneliness, increased social engagement attitudes and self-esteem [34], and more active social support and life satisfaction were also observed during the COVID-19 pandemic pilot period [35,36].

Despite the considerably high prevalence of chronic conditions, social isolation, loneliness, and mental illness among Korean older adults (prior to the pandemic and exacerbated during this time), social prescribing has not yet been implemented on a national scale. Although there is a growing body of qualitative social prescribing research on a global scale, significant gaps remain in the literature (particularly in quantitative studies). Most notably, how to evaluate the implemented project model with more samples is the priority problem in each region. To address this gap, the present study explores the significant differences in social support, loneliness, depression, cognition, and life satisfaction among rural Korean older adults before and after receiving the pilot social prescribing project intervention, and assesses the effectiveness of the pilot project in which they participated. The findings contribute to the development of social prescribing initiatives that target specific needs, preferences, and expectations of older adults, ultimately enhancing their health, social engagement, and well-being.

Methods

Project Objectives

The pilot project was organized by Yonsei University to achieve two goals: (1) to reduce depression and loneliness in the older population in rural settings in Korea and (2) to help establish city-rural social capital and relationships among older people.

Study Design

Participants were divided into experimental and control groups to evaluate the impact of a social prescribing program on depression, cognitive function, and loneliness among the older adults (April 2022 - December 2022). The design used a nonequivalent control group pretest-posttest model. This approach allowed researchers to compare outcomes before and after the intervention in both the experimental (who received the social prescribing intervention) and control groups (who did not participate in the intervention).

The following are the key steps in the study design:

1. Preliminary survey: Before each intervention, a preliminary survey was conducted to select eligible participants based on set inclusion criteria, such as age, absence of severe depression (as measured by the short form Geriatric Depression Scale-Korean version, GDS-K), and capability for everyday communication.
2. Assignment to groups: Participants were assigned to either the experimental or control group based on their willingness to participate in the social prescribing program.
3. Intervention implementation: The experimental group participated in a 2-month social prescribing program that involved activities designed to reduce isolation and improve mental health.
4. Postintervention assessment: After completing the program, both groups underwent a posttest to assess changes in depression, cognitive function, and loneliness on the final intervention day.

Intervention Programs

Music storytelling (Table 1), self-help groups, community gardening, Kalimba learning, flower tea classes, smartphone education, and COVID-19 prevention sessions were conducted. The related study protocol and research were also published in the previous studies [32,37].

XSL•FO
RenderX

Table . Intervention activities in the social prescribing program.

Intervention	Activities	Location	Time
Music storytelling	• Rhythm for life • Guide image music • Movement improvisation circle	Community library	1 per week (90 min)
Self-help group (School of Hunjang[a])	• Healthy 100-year lecture: fruit • Making natural soup for COVID-19 prevention • Making masks for COVID-19 prevention • Making natural diffuser • Songwriting for the community	Community library	1 per week (60 min)
Community gardening	• Gardening education in the spring and summer	Community library	Community garden 1 per week (30 min)
	• Nursery plant (cherry tomato, sweet potato, corn, and pepper) • Monitoring and managing community garden • Harvesting	Wonju City Agricultural Center	Community garden 1 per week (30 min)
Kalimba class (handheld musical instrument)	• Learning, practice, and public performance	Community library	1 per week (60 min)
Flower tea class (with nutrition facts)	• Learning and flower tea making • Private selling	Community library	1 per week (60 min)
Smartphone education	• Six-week smartphone education intervention for 65 and older people	Community library	1 per week (60 min)
COVID-19 prevention session	• Voluntary visit (per week)	—[b]	3 per week
	• COVID-19 phone call (per week)	—	4 per week
	• Education of healthy nutrition in COVID-19 pandemic	—	1 per week

[a]Hunjang: The teacher or headmaster of the Seodang was called the Hunjang.

[b]Not applicable.

- Music storytelling: Music therapy sessions were the main social program. The intervention was implemented as a multicomponent social prescribing package, with music storytelling as the core component and other activities serving complementary roles to support social engagement, peer support, and accessibility within the social prescribing

XSL•FO
RenderX

framework. Rhythms, group dances, and singing sessions are examples of music therapy. Sessions took place every Thursday for a duration of 10 weeks. Moreover, the team encouraged the participants to exercise while listening to music. Examples include stretching, practicing deep breathing, and joining an indoor walking program (first, the program instructor demonstrated appropriate movements and positions so that participants could easily follow along) [32].

- Self-help groups: Participants gathered at the community library every Friday to make various products using their own hands. The participants made COVID-19 masks, dolls, and soaps. While working on their personal projects, the participants told each other old stories and taught each other how to make the products.
- Community gardening activities, Kalimba class, and flower tea class: Participants and volunteers planted corn, peppers, sweet potatoes, cabbage, and tomatoes during the farming program. Kalimba and flower tea classes were also added to improve older people's hand activities and social contact.
- Smartphone education: Participants were divided into 6 groups to conduct the "smartphone education" session with education instructors and volunteer students. The education plan included basic smartphone operation methods to use apps most commonly used in Korea, such as Naver apps (information search platform) and KakaoTalk (mobile messaging app). Digital support in this study was intentionally designed as a low-threshold and supportive component rather than a technology-intensive intervention. Smartphone-based activities focused on basic functions and were embedded within face-to-face community programs to enhance accessibility and inclusivity among older rural adults. Additionally, every Friday, 2 instructors and a health major researcher shared feedback on the contents of the week's education, adjusting the overall flow of education to be consistent with the difficulty level.

Project Pilot Areas

The pilot areas for this research were Heung-eop-myeon and Munmak-eup in Wonju City, Gangwon Province, Korea. As of August 2022, Gangwon Province ranks fourth nationwide, with a high aging rate (Jeollanam-do 24.9%, Gyeongsangbuk-do 23.5%, Jeollabuk-do 22.9%, and Gangwon-do 22.4%). Both Heung-eop-myeon and Munmak-eup in Wonju City are experiencing high levels of aging with rates of 21.1% and 20.7%, respectively, indicating that both areas have become super-aged societies.

Participant Recruitment

The town office and welfare center of Heungup Town provided a potential participant list and communication support. As detailed below, participants, who were being monitored by the local welfare center, were older than 65 years of age, had mild depression, provided informed consent, and lived within walking distance of the community library were selected for the pilot project:

1. The social welfare officer of the Heungup-myeon Administrative Welfare Center recommended the older participants (age ≥65 y and living alone) living in rural areas, Wonju City, Gangwon-do.
2. First participant selection: The researchers conducted the GDS-K test with the candidates to identify the range of depression and selected those who had mild or moderate symptoms (GDS-K score: 9 - 18).
3. Second participant selection: The researchers selected older people who agreed to participate in the pilot and survey and had no difficulty with daily conversation and physical behavior.
4. Third participant selection (final): Public health doctors conducted an overall health diagnosis. Based on the results, the researchers selected the older adults who had no severe depression as the final participants in the pilot project.

Program Participants Selection

Selection criteria for program participants: The program participants, consisting of experimental and control groups, were residents aged 65 years and older from Heung-eop-myeon and Munmak-eup in Wonju City, Gangwon Province. The other criteria were as follows:

1. The first criterion was that the participants had to be recommended by at least 1 local social worker, primary health care provider (doctor), or community health nurse to have no issues participating in the pilot project.
2. The second criterion was that participants must be capable of everyday communication without hearing or visual impairments.
3. The third criterion was that during the preliminary survey, participants must not have severe depression (normal-0 - 9; mild depressives-10 - 19; and severe depressives-20 - 30), as determined by a score of 16 or lower on the GDS-K, a tool for detecting depression.

Sample Size

The sample size for participants was determined using G*Power 3.1 (Erdfelder, Faul, & Buchner, 1996), with an effect size of 0.5, a power of 0.95, and a significance level of .05, resulting in a minimum sample size of 54. Participants were conveniently sampled from community libraries and senior welfare centers in Wonju City, totaling 294 individuals. Among them, 148 who wanted to participate in the intervention program were designated as the experimental group, and 146 who did not wish to participate were designated as the control group.

Data Collection

Data were collected through one-on-one surveys by seven researchers, including 6 specializing in public health and 1 specializing in preventive medicine. The surveys were administered after explaining them to the participants, ensuring their understanding of the content, and obtaining their consent. The average survey time was approximately 15 - 20 minutes per participant. The surveys conducted before and after the intervention included questions designed to measure sociodemographic indicators, depression, cognitive function, loneliness, social support, and life satisfaction. Data collected from these assessments were analyzed to evaluate the impact of the social prescribing intervention by comparing the preintervention and postintervention results between the

XSL•FO
RenderX

experimental and control groups to determine the program's effectiveness.

Variables and Measurement

Selection of Study Variables

Building on previous research and the opinions of two local experts, this study selected age, sex, literacy, marital status, living alone status, religion, subjective economic level, and subjective health status as control variables. To measure the effectiveness of the pilot project, depression, social support, life satisfaction, cognitive function, and loneliness were selected as dependent variables.

Measurement of Study Variables

The short form GDS-K was used to measure depression among the participants. Higher scores indicated greater depression, with 15 items (15 scores) and a cutoff score of 8 [38]. The 8-item short version of the Perceived Social Support through Others Scale-8 estimated social support levels among these participants, and the highest score was 40 [39]. The Life Satisfaction Scale for Korean older adults was used to measure the index of their life satisfaction, which included 20 items and 3 scores [40]. Cognitive function was assessed using the Mini-Mental State Examination Korean Version (MMSE). For measuring loneliness, the shortened University of California, Los Angeles Loneliness Scale was used [41]. This scale consists of 8 items, with a total possible score of 32, where higher scores indicate greater levels of loneliness.

Statistical Analysis

The collected data were analyzed using SPSS Statistics version 26.0 (IBM Corp), using the following methods:

First, sociodemographic characteristics, depression levels, social support, life satisfaction, cognitive function, and loneliness levels of the participants were analyzed using percentage and frequency analyses.

Second, to verify the homogeneity between the experimental and control groups, *t* tests and Chi-square tests were conducted for all measured variables (including sociodemographic indicators).

Third, paired and independent *t*-tests were conducted to identify significant differences in social support, loneliness, depression, cognitive function, and life satisfaction scores before and after the intervention.

Fourth, to explore significant changes in social support, loneliness, depression, cognitive function, and life satisfaction between the experimental and control groups, simple and multiple difference-in-differences (DID) analyses were performed.

Finally, to examine the direct and indirect relationships between social support, loneliness, and life satisfaction with depression and cognitive function among older individuals, structural equation modeling (SEM) using SPSS Amos (version 23.0, IBM Corp) was conducted. Goodness-of-fit was evaluated using the following indicators: adjusted goodness-of-fit index (≥0.80), comparative fit index (≥0.80), incremental fit index (>0.80) [42,43], and root mean square error of approximation (≤0.80) [44,45].

All statistical significance levels were set at α=.05, and all statistical tests were 2-tailed.

DID analysis is a statistical method that can be applied when there is an experimental group and a control group, and the time point before and after the intervention is divided. This method is widely used to evaluate the effectiveness of health care policies [46-48] and community-based health projects [49,50]. A control group was selected from the same area where the experimental group lived, and presurvey and postsurvey data were collected and analyzed for the same subject in both groups through the same questionnaire. Through this, regional and individual characteristics were controlled, and the effectiveness of the intervention was verified. The amount of change in the effect variable value in the control group from the amount of change in the effect variable value before and after the intervention of the experimental group was used to confirm the effect of the social prescription program intervention. In the United Kingdom, a study using the double-difference analysis was used to assess the effectiveness of social prescription programs for patients with type 2 diabetes [51].

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki. Before conducting the survey, the purpose of data collection and use was explained to the participants, and consent was obtained with the promise of maintaining anonymity regarding the use of the research data. Each participant received a small gift (10 $) after each session. All data were deidentified, stored securely, and accessible only to authorized research team members to ensure privacy and confidentiality. This study also received institutional review board approval from the Yonsei University Future Campus Bioethics Committee (1041849 - 202206-SB-105-01).

Results

Participant Characteristics

Results in Table 2 demonstrate the key findings regarding the experimental and control groups' sociodemographic characteristics and changes in the measured variables. The mean age of the 2 groups did not have a statistically significant difference (*P*=.60). A notable difference in gender distribution was found, with a higher proportion of women in the experimental group than in the control group (*P*<.001). Additionally, there were no statistically significant differences in living alone status (*P*=.57) and partner status (*P*=.25), indicating similar living conditions in both groups. Education levels demonstrated that approximately 116 (78.4%) had education and literacy, and 97 (65.5%) to 103 (70.5%) had their own religion in both groups (*P*=.41 and *P*=.36, respectively). Economic levels and subjective health status showed statistically significant differences, with the experimental group having fewer individuals at a high economic level (*P*<.001) and the experimental group reporting worse health (*P*<.001). There were statistically significant differences in social support and life satisfaction between the 2 groups.

XSL•FO
RenderX

Table . Social demographic characteristics of experimental group and control group participants.

Variables and categories	Experimental group (n=148)	Control group (n=146)	*t* test/chi-square (*df*)	*P* value
Age (y), mean (SD)	79.31 (6.796)	78.25 (6.344)	1.396 (292)[a]	.60
Sex, n (%)			35.971 (1,294)[b]	<.001
Male	8 (5.4)	48 (32.9)		
Female	140 (96.6)	98 (67.1)		
Live alone, n (%)			0.330 (1,294)[b]	.57
No	64 (43.2)	68 (46.6)		
Yes	84 (56.8)	78 (53.4)		
Partner, n (%)			1.349 (1,294)[b]	.25
No	89 (60.1)	78 (53.4)		
Yes	59 (39.9)	68 (46.6)		
Education, n (%)			0.675 (1,294)[b]	.41
No	32 (21.6)	26 (17.8)		
Yes	116 (78.4)	120 (82.2)		
Have religion, n (%)			0.847 (1,294)[b]	.36
No	51 (34.5)	43 (29.5)		
Yes	97 (65.5)	103 (70.5)		
Subjective economic level, n (%)			16.384 (2,294)[b]	<.001
High	6 (4.1)	22 (15.1)		
Average	64 (43.2)	59 (40.4)		
Low	78 (52.7)	65 (44.5)		
Subjective health, n (%)			13.119 (1,294)[b]	<.001
Bad	87 (58.8)	55 (37.7)		
Good	61 (41.2)	91 (62.3)		
Social support, mean (SD)	24.96 (8.660)	23.75 (6.765)	1.160 (226)[a]	.009
Loneliness, mean (SD)	20.75 (13.404)	23.82 (14.054)	–1.587 (200)[a]	.43
Depression, mean (SD)	11.68 (7.279)	10.17 (7.811)	1.937 (292)[a]	.54
Cognitive function, mean (SD)	25.39 (4.125)	24.92 (4.447)	0.934 (292)[a]	.08
Life satisfaction, mean (SD)	23.01 (10.448)	25.09 (8.504)	–1.673 (248)[a]	.004

[a]*t* test.

[b]Chi-square test.

Results of Simple 2-Way ANOVA

Table 3 shows the results of a simple 2-way ANOVA of the experimental and control groups after the social prescription program intervention, including indicators of social support, loneliness, depression, cognitive function, and life satisfaction.

XSL•FO
RenderX

Table . Results of sample DID[a] analysis.

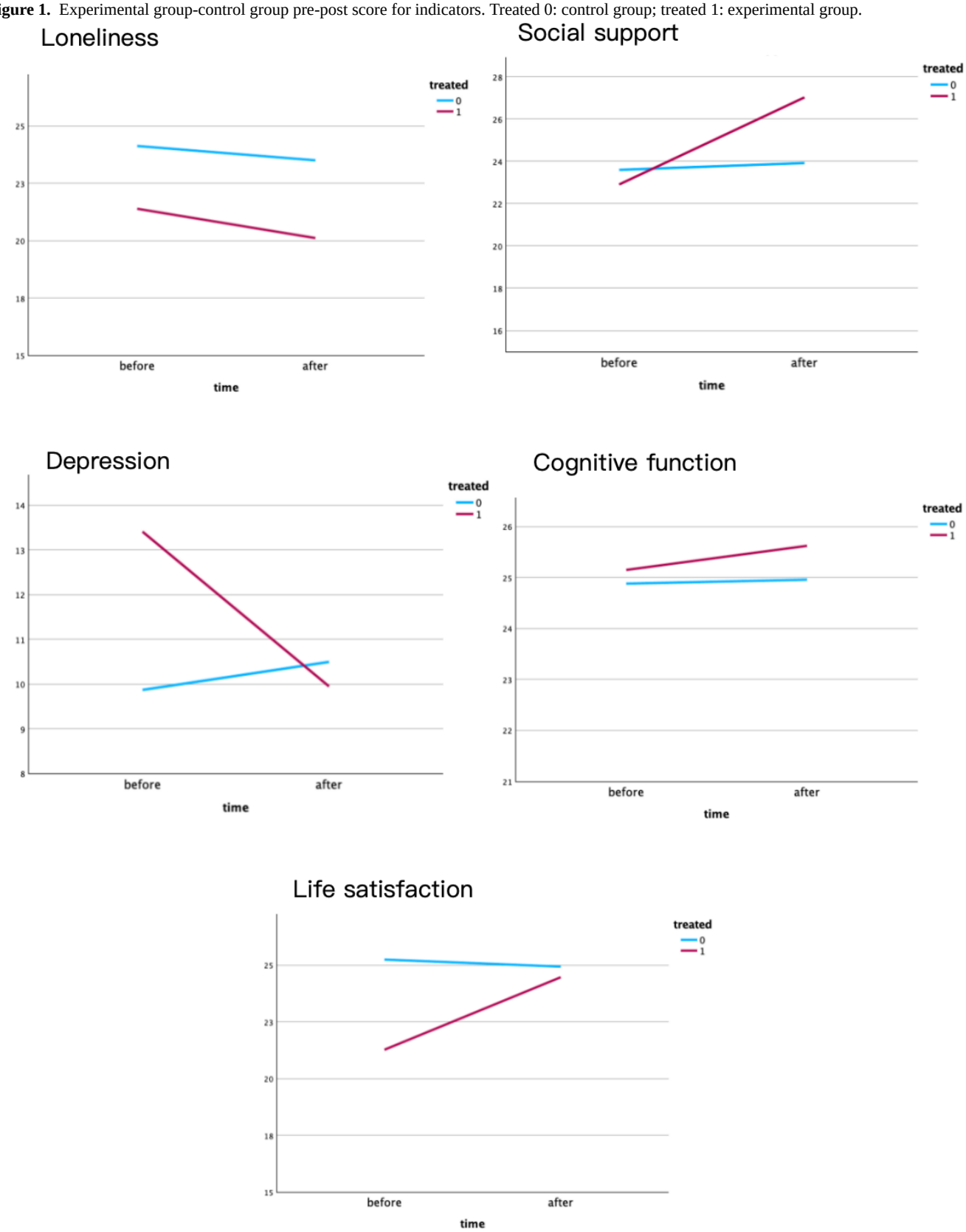
Variables and group (n)	Participants (n)	Pretest (A), mean (SD)	Posttest (B), mean (SD)	Difference, B–A (*P* values)	DID analyses
Social support					
Experimental	112	22.90 (8.40)	27.02 (8.51)	4.12 (.02)	2.79
Control	122	23.59 (6.77)	23.92 (6.82)	1.33 (.79)	
Loneliness					
Experimental	100	21.39 (13.21)	20.11 (13.69)	−1.28 (.62)	−0.91
Control	94	24.13 (13.90)	23.50 (14.36)	−0.37 (.83)	
Depression					
Experimental	146	13.41 (7.18)	10.34 (7.01)	−3.07 (.009)	−3.69
Control	144	9.87 (7.45)	10.49 (8.22)	0.62 (.63)	
Cognitive function					
Experimental	145	25.15 (3.72)	25.62 (4.51)	0.47 (.49)	0.39
Control	144	24.88 (4.41)	24.96 (4.52)	0.08 (.92)	
Life satisfaction					
Experimental	123	21.27 (9.64)	24.75 (11.01)	3.48 (.049)	4.14
Control	144	25.24 (8.56)	24.90 (8.49)	−0.66 (.83)	

[a]DID: difference-in-differences.

For social support, depression, and life satisfaction, the scores in the experimental group showed a statistically significant increase from a premeasure of 22.90, 13.41, and 21.27 to a post measure of 27.02 (*P*<.02), 10.34 (*P*=.009), and 24.75 (*P*=.049). Loneliness scores and cognitive function in both the experimental and control groups showed no statistically significant increase from the post test compared with the pretest. Depression in the experimental group demonstrated a statistically significant decrease, while the control group demonstrated increased depression scores without statistical significance (*P*=.63). Also, similar changes occurred in the control group in terms of social support (1.33) and life satisfaction (opposite, −0.66). Figure 1 shows the pre-post scores change graph in both groups.

XSL•FO
RenderX

Figure 1. Experimental group-control group pre-post score for indicators. Treated 0: control group; treated 1: experimental group.



Associations Between Socioeconomic Factors and Social Support, Loneliness, Depression, Cognitive Function, and Life Satisfaction

The results of multiple DID analyses for socioeconomic factors and other indicators are shown in Table 4.

XSL•FO
RenderX

Table . Results of multiple difference-in-difference analysis.

Variables and categories	β				
	Social support (n=456)	Loneliness (n=404)	Depression (n=588)	Cognitive function (n=588)	Life satisfaction (n=500)
Time (reference: pretest)					
Post test	0.149	−0.158	0.978	−0.138	−0.743
Group (reference: control)					
Experimental	−1.102	−3.195	−2.605[a]	0.477	−3.056[a]
Interaction (time×group)	4.136[a]	−1.255	−4.121[b]	0.553	4.292[a]
Age	0.023	−0.376[a]	0.068	−0.168[c]	0.059
Gender (reference: male)					
Female	3.201[a]	−0.361	0.979	−0.971	1.735
Living alone (reference: yes)					
No	−0.802	−6.182	−0.558	2.014[a]	1.870
Partner (reference: no)					
Yes	1.144	0.264	0.796	2.085[a]	0.210
Literacy (reference: no)					
Yes	2.586[a]	−2.336	−2.296[a]	2.847[c]	2.344
Have religion (reference: no)					
Yes	0.446	−1.668	−2.165[a]	0.727	2.820[a]
Subjective health status (reference: bad)					
Good	3.856[c]	−0.734[a]	−5.165[c]	1.176[a]	7.439[c]
Subjective economic level (reference: low level)					
Middle level	−1.682	10.803	−3.562[a]	4.537[c]	5.029[a]
Upper level	−3.074	10.417	−3.430[a]	4.895[c]	5.300[a]
R^2	0.155	0.155	0.221	0.323	0.196
F test (*df*)	3.289 (12,443)	2.892 (12,392)	6.652 (12,575)	11.180 (12,575)	4.802 (12,487)
P value	<.001	<.001	<.001	<.001	<.001

[a]*P*<.05.

[b]*P*<.01.

[c]*P*<.001.

In the analysis of the relationship between these factors and depression scores, group and interaction variables were statistically significant. Comparing the experimental group, the control group showed a higher level of depression (*P*<.05); statistical significance was also found in the negative direction in the interaction variable (β=−4.121, *P*<.01), which confirmed that the intervention effect was found on the change in depression in the experimental group. Moreover, literacy (vs no literacy), having religion (vs no religion), good subjective health status (vs bad subjective health status), and high subjective economic level (vs middle or low subjective economic level) demonstrated a statistically significant negative effect on the depressive condition, indicating that older people who have literacy, specific religion, better subjective health,

XSL•FO
RenderX

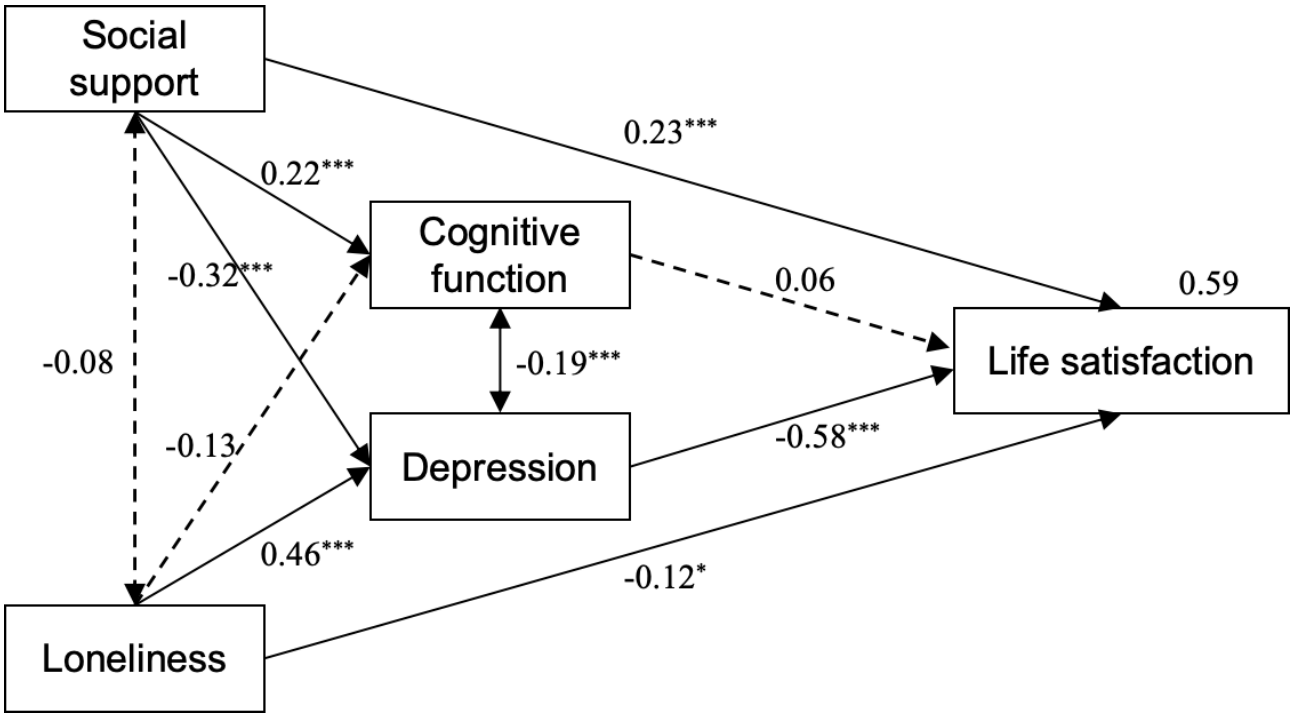
and better subjective economic level would have a lower level of depression. Similar results were found in the analysis of life satisfaction. However, no statistically significant relationship was found between literacy and life satisfaction.

Gender differences were apparent in social support, with female participants reporting higher perceived levels of social support than their male counterparts (β=3.201, *P*<.05) in the model of socioeconomic factors and social support. Additionally, social support scores seemed to have no relationship with subjective economic level. As a result of analyzing loneliness and cognitive function as an effect variable, neither the time variable, group variable, nor interaction variable were statistically significant. Surprisingly, statistically significant negative effects of age and subjective health status were reported in the loneliness analysis model; loneliness statistically significantly decreased as age increased (β=−.376, *P*<.05) with a better subjective health status (β=−.734, *P*<.05). However, for cognitive function, it was confirmed that younger persons who did not live alone and had a partner, literacy, better subjective health, and better subjective economic level had higher cognitive function scores.

Impacts of Social Support, Loneliness, Cognition, and Depression on Life Satisfaction Among the Participants

SEM analysis (Figure 2) illustrates the association and standardized effects of social support, loneliness, cognition, and depression on life satisfaction among the participants. The level of social support positively affected cognitive function (β=.22) and life satisfaction (β=.23) and had a negative relationship with depression levels (β=−.32). Additionally, all associations between loneliness, depression, and life satisfaction were statistically significant: loneliness had a positive effect on depression (β=.46), and life satisfaction was negatively impacted by depression (β=−.58) and loneliness (β=−.12), which showed the mediating effect of depression. Surprisingly, no statistically significant relationship was found between loneliness and social support or cognitive function in this study.

Figure 2. Structural equation modeling analysis of the association among social support, loneliness, cognitive function, depression, and life satisfaction. (χ^2=1.50; *P*=.22; comparative fit index=0.99; incremental fit index=0.99; Tucker-Lewis index=0.97; root mean square error of approximation=0.041). Solid lines: significant paths; dotted lines: nonsignificant paths; **P*<.05, ****P*<.001.



Regarding the relationships between cognitive function, depression, and life satisfaction, a statistically significant negative correlation was demonstrated between cognitive scores and depression (*P*<.001), whereas the opposite was observed for the relationship between social support and cognition. Moreover, although there were both negative effects between social support and depression, loneliness and cognitive function, statistical significance was only reported in social support and depression (β=−.32). This indicated that depression was negatively impacted by social support, and that a higher level of social support would result in a lower depressive condition. Fit indices of the SEM model showed an overall goodness of fit for the model: χ^2/df=1.50, root mean square error of approximation =0.041 (95% CI 0.000 - 0.050; *P*=.02), and comparative fit index =0.99.

Discussion

Principal Results

This study assessed the effects of a social prescribing pilot project on the older adults in rural South Korea. The intervention aimed to alleviate depression and loneliness, while fostering cognitive function, social support, and life satisfaction. Consistent with real-world social prescribing practice, this pilot study focused on evaluating the overall effectiveness of an integrated, multicomponent intervention model rather than isolating the relative contribution of individual components.

XSL•FO
RenderX

The results demonstrated statistically significant reductions in depression and improvements in social support and life satisfaction, alongside slight but nonsignificant improvements in cognitive function. SEM analysis further revealed a positive association between social support and life satisfaction, with depression and cognitive function acting as mediating factors, supporting the potential of social prescribing as a nonclinical, community-based approach to improving quality of life among rural older adults.

Research on social prescribing in the United Kingdom has highlighted its role as a valuable complement to clinical treatments for depression, with benefits for both physical and mental health [52,53]. A related study revealed that individuals who engaged more frequently with link workers showed greater improvements in anxiety and depression [54]. Consistent with this evidence, participants in our structured social prescribing program experienced significant reductions in depression and improvements in social support and life satisfaction. Similar positive effects have been observed in global social prescribing studies, particularly in reducing depression through community engagement and personalized support [55].

The differential outcomes observed in this study likely reflect varying responsiveness across outcome domains. Affective and evaluative indicators, such as depression and life satisfaction, appear more sensitive to short-term increases in social engagement and perceived support—core features of the intervention. In contrast, loneliness represents a more deeply rooted subjective experience that may be less responsive to change over a relatively short intervention period, particularly among older adults with long-standing social patterns in rural settings. Although cognitive function showed slight improvement, it did not reach a statistically significant association with the program intervention. This may be partly attributable to the relatively short intervention duration, which was sufficient to influence affective and psychosocial outcomes but may have been inadequate to produce detectable changes in global cognitive function [56]. In addition, cognitive function was assessed using the MMSE, a screening tool with limited sensitivity to short-term or domain-specific cognitive changes [57].

The study also highlighted the role of individual and contextual characteristics in shaping intervention outcomes. Factors such as age, gender, living arrangements, literacy, religion, subjective health, and economic status were associated with variations in social support, depression, cognitive function, and life satisfaction. In rural contexts characterized by long-term residence and stable social ties, older age was associated with lower reported loneliness, consistent with prior evidence suggesting greater emotional regulation, acceptance of social roles, and more selective social networks in later life [58-60]. Gender differences further indicated that women may benefit more readily from social support, aligning with research showing stronger community integration among older women in rural settings [61]. These findings suggest that future social prescribing programs may benefit from gender- and literacy-sensitive designs to enhance inclusivity and effectiveness [62,63].

Living arrangements and social resources also played an important role. Participants who did not live alone nor had a partner demonstrated higher cognitive function, underscoring the role of sustained social and emotional engagement in cognitive health [64], supporting previous studies linking partnership with cognitive resilience [65]. Literacy and religious affiliation were associated with lower depression and higher life satisfaction, highlighting the importance of culturally and cognitively accessible intervention materials and the protective role of spiritual and community support [63,66]. Better subjective health and economic conditions were consistently associated with lower depression and higher life satisfaction, likely reflecting reduced stress and enabling participation in meaningful social activities [25,67]. Conversely, reliance on basic pensions or combined financial and caregiving support from family members was linked to higher depressive symptoms, while informal family care was associated with greater life satisfaction compared with formal long-term or government-provided care [68-70].

The interaction effect between time and group demonstrated that the intervention's structured support and tailored activities effectively influenced mental health trajectories, suggesting that social prescribing interventions may be particularly effective in the older population by providing structured and continuous social engagement and support. These positive shifts emphasize the cumulative benefits of social prescribing, as repeated engagement in meaningful and supportive activities over time appears to foster mental well-being and social connectedness among older participants [71]. It is also possible that the cognitive effects of social prescribing operate indirectly through sustained improvements in social support and depression, suggesting that more pronounced cognitive benefits may emerge with longer engagement and follow-up [56]. This outcome may reflect older participants' established social skills, accumulated life experiences, or resilience, which may make them better equipped to benefit from social prescribing's community-based support, especially in terms of incremental mental health [53].

The results also indicated key relationships among social support, loneliness, depression, cognitive function, and life satisfaction in the older population. Social support had a positive impact on both cognitive function and life satisfaction, while showing a negative association with depression. This finding suggests that social support may serve as a buffer against depressive symptoms, facilitating a more satisfying life experience and supporting cognitive health in older adults. These outcomes are consistent with existing research demonstrating that social support is protective against mental health challenges, helping maintain cognitive function and emotional well-being in aging populations [72,73]. Additionally, loneliness was positively associated with depression, and depression negatively affected life satisfaction, indicating that depression may mediate the relationship between loneliness and life satisfaction [74,75]. This aligns with research indicating that depression is a critical factor in understanding the impact of loneliness on overall well-being, underscoring the importance of addressing both social support and mental health through social prescribing interventions [76].

XSL•FO
RenderX

Moreover, the absence of statistically significant direct associations between loneliness and social support or cognitive function warrants careful interpretation. Loneliness reflects a subjective emotional experience, whereas social support represents perceived or available resources embedded within social structures [77]. In rural communities with relatively stable and enduring social networks, older adults may continue to engage in routine social interactions despite experiencing emotional loneliness, which may attenuate direct associations in short-term analyses [78]. Consistent with this interpretation, our SEM findings indicated that loneliness may primarily influence well-being indirectly through depressive symptoms rather than exerting direct effects on cognitive function or perceived social support. Cultural norms emphasizing emotional restraint and endurance among older adults in rural Korea may further contribute to the underreporting of loneliness, partially explaining the observed pattern [79].

These results further suggest that the cognitive benefits observed in this study are more directly attributable to enhanced social support than to reductions in loneliness per se. Social support appears to play a central role in mental health resilience, as higher levels of social support were associated with lower depression and greater life satisfaction [80]. By strengthening social connections and perceived support, social prescribing interventions may therefore mitigate the negative impact of depressive symptoms on overall well-being, even in the absence of immediate changes in loneliness or cognitive function. This pattern aligns with existing evidence showing that structured social interactions and supportive relationships foster resilience against mental decline and contribute to improved quality of life in later life [81,82].

Taken together, these findings suggest that the mental health benefits of social prescribing in rural older adults are driven primarily by enhanced social support and its downstream effects on depression and life satisfaction, rather than by immediate reductions in loneliness. This highlights social support as a central and actionable mechanism through which community-based, culturally embedded interventions can promote mental health resilience and quality of life in later life. At the same time, the findings underscore the importance of considering contextual and affective pathways—particularly the indirect role of loneliness—when interpreting intervention mechanisms and designing future longitudinal studies to capture longer-term cognitive and psychosocial effects.

Limitations

This study had some limitations. First, the nonrandomized design and participation introduced potential selection bias. Although baseline equivalence was assessed and intervention effects were examined using DID analysis with covariate adjustment, residual confounding cannot be fully excluded. Participants in both the experimental and control groups may have had higher motivation for social engagement and more positive attitudes toward community activities, which could have influenced the observed outcomes. Second, as a real-world pilot study, it was challenging to fully implement all elements recommended by established intervention guidelines [83]. Third, the relatively small sample size and short intervention duration limit the generalizability of the findings and the ability to assess sustained or delayed effects, particularly with respect to cognitive outcomes. The use of the MMSE, while appropriate for screening, may have limited sensitivity to short-term cognitive changes in relatively high-functioning older adults. Future studies with larger samples, extended follow-up, and more sensitive or objective cognitive assessments are therefore warranted.

Conclusions

This pilot study examined the effectiveness of a community-supported and digitally supported social prescribing program in improving mental health and well-being among older adults in rural South Korea. The findings demonstrated that such interventions can significantly reduce depression and enhance social support and life satisfaction, reinforcing the value of integrating community-based resources with culturally relevant digital tools in elder care. These results underscore the incorporation of social prescribing into broader community care strategies to address mental health disparities in underserved aging populations.

The program's effectiveness was driven by culturally tailored activities—such as music sessions, flower tea making, and gardening—implemented through locally accessible and digital supports that promoted sustained engagement. By encouraging collaboration among health care professionals, local facilitators, and residents, social prescribing moves beyond traditional clinical models and toward inclusive, socially grounded care. The study underscores the importance of embedding social prescribing within broader community care and healthy aging strategies, with intermediary roles such as link workers serving as key connectors between digital and local community resources in sustaining and scaling such programs within national mental health and aging policy frameworks.

Implications for Practice and Policy

The findings of this study offer timely and policy-relevant implications for enhancing mental health promotion among older adults in rural and underserved communities. From a sustainability and public health systems perspective, the results suggest that social prescribing interventions can be effectively embedded within broader healthy aging strategies when they leverage existing community infrastructure, such as community libraries, welfare centers, and local volunteer networks. By emphasizing nonclinical, preventive, and person-centered approaches, social prescribing aligns with ongoing shifts toward community-based mental health promotion for aging populations.

In practice and policy, the integration of digitally supported components—implemented as low-threshold, supportive tools embedded within face-to-face community activities—can improve accessibility and continuity of care without exacerbating digital exclusion. At the same time, the institutionalization of intermediary roles, such as link workers or community coordinators, is critical for sustaining and scaling social prescribing frameworks by facilitating needs assessment, navigation of community resources, and coordination between medical and nonclinical services. At the national level, these

XSL•FO
RenderX

findings align with South Korea's priorities to reduce social isolation and promote healthy aging, underscoring the importance of embedding social prescribing within existing mental health and aging policy frameworks and establishing stable funding and governance mechanisms to support sustainable, community-driven mental health promotion.

Acknowledgments

The authors wish to acknowledge the contribution of participants and stakeholders to this program throughout the study process.

Funding

The study was funded by National Research Foundation of Korea, Grant/Award Number: NRF-2018S1A5A2A03039636; Ministry of Education of the Republic of Korea; This work was also supported by the Ph.D. research initiation project of Anhui Medical University (0501054201) and the Youth Social Science Research Program of the Education Department of Anhui Province (2025AHGXSK40223).

Data Availability

Due to ethical reasons, the raw data would remain confidential and would not be shared. The processed data that support the findings of this study are available from the corresponding author, upon reasonable request.

Authors' Contributions

EWN was involved in the design and conceptualization of the study. BZ, JAL, SL, HJY, HKK, SMH, XZ, and EWN were involved in program implementation and data collection. BZ finished data analysis, interpretation, and the initial draft of the manuscript and revised the manuscript. BZ and EWN contributed to funding acquisition, quality control, and reviewing of the final draft of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Soon GC, Bodeker G, Kariippanon K, editors. Healthy Ageing in Asia Culture, Prevention and Wellness: CRC Press; 2022. [doi: 10.1201/9781003043270]
2. Zablith N, Diaconu K, Naja F, et al. Dynamics of non-communicable disease prevention, diagnosis and control in Lebanon, a fragile setting. Confl Health 2021 Jan 11;15(1):4. [doi: 10.1186/s13031-020-00337-2] [Medline: 33430916]
3. The Lancet Regional Health – Western Pacific. South Korea's population shift: challenges and opportunities. Lancet Reg Health West Pac 2023 Jul;36:100865. [doi: 10.1016/j.lanwpc.2023.100865] [Medline: 37621310]
4. Jang H, Lee W, Kim YO, Kim H. Suicide rate and social environment characteristics in South Korea: the roles of socioeconomic, demographic, urbanicity, general health behaviors, and other environmental factors on suicide rate. BMC Public Health 2022 Mar 1;22(1):410. [doi: 10.1186/s12889-022-12843-4] [Medline: 35227243]
5. Fairbrass AJ, Chatterjee H, Jones KE, Osborn D. Human responses to nature- and culture-based non-clinical interventions: a systematised review. Perspect Public Health 2022 May;142(3):149-157. [doi: 10.1177/1757913920967036] [Medline: 33323042]
6. Wand APF, Zhong BL, Chiu HFK, Draper B, De Leo D. COVID-19: the implications for suicide in older adults. Int Psychogeriatr 2020 Oct;32(10):1225-1230. [doi: 10.1017/S1041610220000770] [Medline: 32349837]
7. Noguchi T, Hayashi T, Kubo Y, Tomiyama N, Ochi A, Hayashi H. Living alone and depressive symptoms among older adults in the COVID-19 pandemic: role of non-face-to-face social interactions. J Am Med Dir Assoc 2023 Jan;24(1):17-21. [doi: 10.1016/j.jamda.2022.10.014] [Medline: 36435273]
8. Courtin E, Knapp M. Social isolation, loneliness and health in old age: a scoping review. Health Soc Care Community 2017 May;25(3):799-812. [doi: 10.1111/hsc.12311] [Medline: 26712585]
9. Raina P, Wolfson C, Griffith L, et al. A longitudinal analysis of the impact of the COVID-19 pandemic on the mental health of middle-aged and older adults from the Canadian longitudinal study on aging. Nat Aging 2021 Dec;1(12):1137-1147. [doi: 10.1038/s43587-021-00128-1] [Medline: 37117519]
10. Robb CE, de Jager CA, Ahmadi-Abhari S, et al. Associations of social isolation with anxiety and depression during the early COVID-19 pandemic: a survey of older adults in London, UK. Front Psychiatry 2020;11:591120. [doi: 10.3389/fpsyt.2020.591120] [Medline: 33132942]
11. Krendl AC, Perry BL. The impact of sheltering in place during the COVID-19 pandemic on older adults' social and mental well-being. J Gerontol B Psychol Sci Soc Sci 2021 Jan 18;76(2):e53-e58. [doi: 10.1093/geronb/gbaa110] [Medline: 32778899]

XSL•FO
RenderX

12. Santini ZI, Jose PE, York Cornwell E, et al. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. Lancet Public Health 2020 Jan;5(1):e62-e70. [doi: 10.1016/S2468-2667(19)30230-0] [Medline: 31910981]
13. Sepúlveda-Loyola W, Rodríguez-Sánchez I, Pérez-Rodríguez P, et al. Impact of social isolation due to COVID-19 on health in older people: mental and physical effects and recommendations. J Nutr Health Aging 2020;24(9):938-947. [doi: 10.1007/s12603-020-1469-2] [Medline: 33155618]
14. Wong SYS, Zhang D, Sit RWS, et al. Impact of COVID-19 on loneliness, mental health, and health service utilisation: a prospective cohort study of older adults with multimorbidity in primary care. Br J Gen Pract 2020 Nov;70(700):e817-e824. [doi: 10.3399/bjgp20X713021] [Medline: 32988955]
15. Sayin Kasar K, Karaman E. Life in lockdown: social isolation, loneliness and quality of life in the elderly during the COVID-19 pandemic: a scoping review. Geriatr Nurs 2021;42(5):1222-1229. [doi: 10.1016/j.gerinurse.2021.03.010] [Medline: 33824008]
16. Huang C, Cui W, Yang Y, Huang X, Li C, Lin Z. Do health literate older people have fewer depressive symptoms? Evidence from the urban areas of Western China. Healthcare (Basel) 2024 Nov 26;12(23):2373. [doi: 10.3390/healthcare12232373] [Medline: 39684994]
17. Tan JJX, Kraus MW, Carpenter NC, Adler NE. The association between objective and subjective socioeconomic status and subjective well-being: a meta-analytic review. Psychol Bull 2020 Nov;146(11):970-1020. [doi: 10.1037/bul0000258] [Medline: 33090862]
18. Leavell MA, Leiferman JA, Gascon M, Braddick F, Gonzalez JC, Litt JS. Nature-based social prescribing in urban settings to improve social connectedness and mental well-being: a review. Curr Environ Health Rep 2019 Dec;6(4):297-308. [doi: 10.1007/s40572-019-00251-7] [Medline: 31713144]
19. Costa A, Henriques J, Alarcão V, et al. "You get out of the house, you talk to each other, you laugh…And that's fantastic" - a qualitative study about older people's perceptions of social prescribing in mainland Portugal. BMC Health Serv Res 2024 May 20;24(1):645. [doi: 10.1186/s12913-024-11086-w] [Medline: 38769571]
20. Khan H, Giurca BC, Sanderson J, Dixon M, Leitch A, Cook C. Social prescribing around the world: a world map of global developments in social prescribing across different health system contexts. : National Academy for Social Prescribing (NASP); 2023 URL: https://socialprescribingacademy.org.uk/media/thtjrirn/social-prescribing-around-the-world-2024.pdf [accessed 2026-03-24]
21. Bild E, Pachana NA. Social prescribing: a narrative review of how community engagement can improve wellbeing in later life. J Community Appl Soc Psychol 2022;32(6):1148-1215. [doi: 10.1002/casp.2631]
22. Liebmann M, Pitman A, Hsueh YC, Bertotti M, Pearce E. Do people perceive benefits in the use of social prescribing to address loneliness and/or social isolation? A qualitative meta-synthesis of the literature. BMC Health Serv Res 2022 Oct 19;22(1):1264. [doi: 10.1186/s12913-022-08656-1] [Medline: 36261835]
23. Savage RD, Stall NM, Rochon PA. Looking before we leap: building the evidence for social prescribing for lonely older adults. J Am Geriatr Soc 2020 Feb;68(2):429-431. [doi: 10.1111/jgs.16251] [Medline: 31840795]
24. Thomas G, Lynch M, Spencer LH. A systematic review to examine the evidence in developing social prescribing interventions that apply a co-productive, co-designed approach to improve well-being outcomes in a community setting. Int J Environ Res Public Health 2021 Apr 8;18(8):3896. [doi: 10.3390/ijerph18083896] [Medline: 33917681]
25. Bickerdike L, Booth A, Wilson PM, Farley K, Wright K. Social prescribing: less rhetoric and more reality. a systematic review of the evidence. BMJ Open 2017 Apr 7;7(4):e013384. [doi: 10.1136/bmjopen-2016-013384] [Medline: 28389486]
26. Lee H, Koh SB, Jo HS, et al. Global trends in social prescribing: web-based crawling approach. J Med Internet Res 2023 May 22;25:e46537. [doi: 10.2196/46537] [Medline: 37086427]
27. Uchihira T, Nakajima K. Limitations of social prescribing to help prevent isolation and loneliness. Jpn Archit Rev 2023 Jan;6(1):e12389. [doi: 10.1002/2475-8876.12389]
28. Saragosa M, Mulligan K, Hsiung S, et al. A qualitative study of national perspectives on advancing social prescribing using co-design in Canada. Health Expect 2024 Aug;27(4):e14144. [doi: 10.1111/hex.14144] [Medline: 38984442]
29. Paquet C, Whitehead J, Shah R, et al. Social prescription interventions addressing social isolation and loneliness in older adults: meta-review integrating on-the-ground resources. J Med Internet Res 2023 May 17;25:e40213. [doi: 10.2196/40213] [Medline: 37195738]
30. Nam EW. Social prescribing system to overcome social isolation and loneliness related to COVID-19. Korean J Health Educ Promot 2020;37(1):113-116. [doi: 10.14367/kjhep.2020.37.1.113]
31. Chung MA, Yoon HJ, Shin DE, Kim JE, Kim HK, Nam EW. A study on community gardening as a social prescribing program. Korean J Health Educ Promot 2021 Sep 30;38(3):39-51. [doi: 10.14367/kjhep.2021.38.3.39]
32. Lim JA, Yoon HJ, Nam HK, et al. Evaluating the social prescribing pilot project's effect on depression and loneliness among rural elderly in Korea. Korean J Health Educ Promot 2023 Mar 31;40(1):13-31. [doi: 10.14367/kjhep.2023.40.1.13]
33. Song GR, Cho HY, Nam EW. Effects of a social prescribing program on the depressive status of elderly people in a rural area of Gangwon-do. Korean Public Health Res 2019;45(4):77-87. [doi: 10.22900/kphr.2019.45.4.008]
34. Kim JE, Lee YL, Chung MA, et al. Effects of social prescribing pilot project for the elderly in rural area of South Korea during COVID-19 pandemic. Health Sci Rep 2021 Sep;4(3):e320. [doi: 10.1002/hsr2.320] [Medline: 34250272]

XSL•FO
RenderX

35. Jo HY, Lee H, Nam EW. Effectiveness of social prescribing for preventing depression and dementia of the elderly living in the rural area [Article in Korean]. J Korean Acad Rural Health Nurs 2023;18(1):1-10. [doi: 10.22715/jkarhn.2023.18.1.1]
36. Yoon H, Kim HK, Lim JA, et al. Evaluation of the effectiveness of a pilot project for digital platform-based social prescribing for the elderly amidst the COVID-19 pandemic [Article in Korean]. Korean J Public Health Res 2022;48(2):1-17. [doi: 10.22900/kphr.2022.48.2.001]
37. Lee H, Koh SB, Jo HS, et al. Evaluating the effectiveness of rural digital social prescribing in Korea: protocol for a cohort study. JMIR Res Protoc 2023 May 24;12:e46371. [doi: 10.2196/46371] [Medline: 37223972]
38. Park D, Kim HS. The factors affecting frailty among the elderly in Korea: a study using the frailty cohort. Int J Environ Res Public Health 2022 Dec 30;20(1):670. [doi: 10.3390/ijerph20010670] [Medline: 36613004]
39. Kim H, Kang J, Kim NR. Development of short perceived social support through others scale (PSO-8): a Rasch analysis. J Hum Underst Couns 2021 Jun 30;42(1):51-70. [doi: 10.30593/JHUC.42.1.3]
40. Choi HJ, Lee YB. Life satisfaction scale for elderly: revisited [Article in Korean]. Korean J Soc Welf 2006;58(3):27-49 [FREE Full text]
41. Xu S, Qiu D, Hahne J, Zhao M, Hu M. Psychometric properties of the short-form UCLA Loneliness Scale (ULS-8) among Chinese adolescents. Medicine (Baltimore) 2018 Sep;97(38):e12373. [doi: 10.1097/MD.0000000000012373] [Medline: 30235699]
42. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Testing Structural Equation Models: Sage Publications; 1993:136-162 URL: https://www.scirp.org/reference/ReferencesPapers?ReferenceID=213477 [accessed 2026-03-24]
43. Sharma S, Mukherjee S, Kumar A, Dillon WR. A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. J Bus Res 2005 Jul;58(7):935-943. [doi: 10.1016/j.jbusres.2003.10.007]
44. Markus KA. Principles and practice of structural equation modeling by Rex B. Kline. Struct Equ Modeling 2012 Jul 20;19(3):509-512. [doi: 10.1080/10705511.2012.687667]
45. Byrne BM. Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming: Routledge; 2013. [doi: 10.4324/9780203807644]
46. Farrar S, Yi D, Sutton M, Chalkley M, Sussex J, Scott A. Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis. BMJ 2009 Aug 27;339:b3047. [doi: 10.1136/bmj.b3047] [Medline: 19713233]
47. Kwon HJ, Cho YU, Ko JY. The effects of long-term care insurance on the life satisfaction and satisfaction in family relationships - the DD method combined with propensity score matching -. Korean J Soc Welf 2011 Nov;63(4):301-326. [doi: 10.20970/kasw.2011.63.4.013]
48. Zhou Z, Zhao Y, Shen C, Lai S, Nawaz R, Gao J. Evaluating the effect of hierarchical medical system on health seeking behavior: a difference-in-differences analysis in China. Soc Sci Med 2021 Jan;268:113372. [doi: 10.1016/j.socscimed.2020.113372] [Medline: 32979776]
49. Lee J, Sharma B, Jo H, Nam EW. Effectiveness of KOICA mental health project for adolescents in Peru using difference in differences analysis. Korean J Health Educ Promot 2019 Sep 30;36(3):51-59. [doi: 10.14367/kjhep.2019.36.3.51]
50. Lee HC, So AY, Lee TH, Nam EW. The effect of nutrition program for under-5 children by household's poverty status in the DR Congo: a difference-in-difference analysis [Article in Korean]. J Reg Dev Stud 2020;29(2):1-24. [doi: 10.22739/ipaid.2020.29.2.1]
51. Wildman J, Wildman JM. Evaluation of a community health worker social prescribing program among UK patients with type 2 diabetes. JAMA Netw Open 2021 Sep 1;4(9):e2126236. [doi: 10.1001/jamanetworkopen.2021.26236] [Medline: 34468751]
52. Fancourt D, Opher S, de Oliveira C. Fixed-effects analyses of time-varying associations between hobbies and depression in a longitudinal cohort study: support for social prescribing? Psychother Psychosom 2020;89(2):111-113. [doi: 10.1159/000503571] [Medline: 31658464]
53. Chatterjee HJ, Camic PM, Lockyer B, Thomson LJM. Non-clinical community interventions: a systematised review of social prescribing schemes. Arts Health 2018 May 4;10(2):97-123. [doi: 10.1080/17533015.2017.1334002]
54. Mercer SW, Fitzpatrick B, Grant L, et al. The Glasgow "Deep End" Links Worker Study Protocol: a quasi-experimental evaluation of a social prescribing intervention for patients with complex needs in areas of high socioeconomic deprivation. J Comorb 2017;7(1):1-10. [doi: 10.15256/joc.2017.7.102] [Medline: 29090184]
55. Drinkwater C, Wildman J, Moffatt S. Social prescribing. BMJ 2019 Mar 28;364:l1285. [doi: 10.1136/bmj.l1285] [Medline: 30923039]
56. Di Lorito C, Bosco A, Booth V, Goldberg S, Harwood RH, Van der Wardt V. Adherence to exercise interventions in older people with mild cognitive impairment and dementia: a systematic review and meta-analysis. Prev Med Rep 2020;19:101139. [doi: 10.1016/j.pmedr.2020.101139] [Medline: 32793408]
57. Stewart S, O'Riley A, Edelstein B, Gould C. A preliminary comparison of three cognitive screening instruments in long term care: the MMSE, SLUMS, and MoCA. Clin Gerontol 2012 Jan;35(1):57-75. [doi: 10.1080/07317115.2011.626515]
58. Graham EK, Beck ED, Jackson K, et al. Do we become more lonely with age? A coordinated data analysis of nine longitudinal studies. Psychol Sci 2024 Jun;35(6):579-596. [doi: 10.1177/09567976241242037] [Medline: 38687352]

XSL•FO
RenderX

59. Charles ST, Carstensen LL. Social and emotional aging. Annu Rev Psychol 2010;61:383-409. [doi: 10.1146/annurev.psych.093008.100448] [Medline: 19575618]
60. Jiang C, Chow JCC, Zhou L, Song H, Shi J. Community support, social isolation and older adults' life satisfaction: evidence from a national survey in China. Aging Ment Health 2024 Jun;28(6):849-857. [doi: 10.1080/13607863.2023.2277871] [Medline: 37921357]
61. Chung ACA, Rimal RN. Social norms: a review. Rev Commun Res 2016;4:1-28. [doi: 10.12840/issn.2255-4165.2016.04.01.008]
62. Ong CH, Pham BL, Levasseur M, Tan GR, Seah B. Sex and gender differences in social participation among community-dwelling older adults: a systematic review. Front Public Health 2024;12:1335692. [doi: 10.3389/fpubh.2024.1335692] [Medline: 38680931]
63. Rasi P, Vuojärvi H, Rivinen S. Promoting media literacy among older people: a systematic review. Adult Educ Q 2021;71(1):37-54. [doi: 10.1177/0741713620923755]
64. Chu HS, Oh J, Lee K. The relationship between living arrangements and sleep quality in older adults: gender differences. Int J Environ Res Public Health 2022 Mar 24;19(7):3893. [doi: 10.3390/ijerph19073893] [Medline: 35409576]
65. Chen C, Tian Y, Ni L, Xu Q, Hu Y, Peng B. The influence of social participation and depressive symptoms on cognition among middle-aged and older adults. Heliyon 2024 Jan;10(2):e24110. [doi: 10.1016/j.heliyon.2024.e24110] [Medline: 38293386]
66. Coelho-Júnior HJ, Calvani R, Panza F, et al. Religiosity/spirituality and mental health in older adults: a systematic review and meta-analysis of observational studies. Front Med (Lausanne) 2022;9:877213. [doi: 10.3389/fmed.2022.877213] [Medline: 35646998]
67. Subramanyam V, Haseena S, Sudha P, et al. Self-rated health, objective health, life satisfaction, and psychological wellbeing among middle adulthood and late adulthood elderly across different social classes: insights from a three-year perspective study. Libr Prog Int 2024;44(2s):1718-1734. [doi: 10.48165/bapas.2024.44.2.1]
68. Zhang D, Wang Y, Jiao Y. The impact of social pension schemes on the mental health of the chinese elderly: a mediating effect perspective of two-way intergenerational support. Int J Environ Res Public Health 2022 Jul 18;19(14):8721. [doi: 10.3390/ijerph19148721] [Medline: 35886572]
69. Park JI, Park TW, Yang JC, Chung SK. Factors associated with depression among elderly Koreans: the role of chronic illness, subjective health status, and cognitive impairment. Psychogeriatrics 2016 Jan;16(1):62-69. [doi: 10.1111/psyg.12160] [Medline: 26450373]
70. Lee SH. A comparative study of quality of life in the elderly care recipient between formal care and informal care [Article in Korean]. Korean J Soc Welf 2021 Nov 30;73(4):87-111. [doi: 10.20970/kasw.2021.73.4.004]
71. Tymoszuk U, Perkins R, Spiro N, Williamon A, Fancourt D. Longitudinal associations between short-term, repeated, and sustained arts engagement and well-being outcomes in older adults. J Gerontol B Psychol Sci Soc Sci 2020 Aug 13;75(7):1609-1619. [doi: 10.1093/geronb/gbz085] [Medline: 31287550]
72. Nicolini P, Abbate C, Inglese S, Rossi PD, Mari D, Cesari M. Different dimensions of social support differentially predict psychological well-being in late life: opposite effects of perceived emotional support and marital status on symptoms of anxiety and of depression in older outpatients in Italy. Psychogeriatrics 2021 Jan;21(1):42-53. [doi: 10.1111/psyg.12633] [Medline: 33230922]
73. Zanjari N, Momtaz YA, Kamal SHM, Basakha M, Ahmadi S. The influence of providing and receiving social support on older adults' well-being. Clin Pract Epidemiol Ment Health 2022;18:e174501792112241. [doi: 10.2174/17450179-v18-e2112241] [Medline: 37274857]
74. Kim C, Chun J. "Meaning in Life" mediates the relationship between loneliness and depressive symptoms among older adults during the COVID-19 pandemic. Healthcare (Basel) 2024 Mar 4;12(5):584. [doi: 10.3390/healthcare12050584] [Medline: 38470695]
75. Cooper M, Avery L, Scott J, et al. Effectiveness and active ingredients of social prescribing interventions targeting mental health: a systematic review. BMJ Open 2022 Jul 25;12(7):e060214. [doi: 10.1136/bmjopen-2021-060214] [Medline: 35879011]
76. Bandyopadhyay S. Loneliness and depression in older adults: living well in older age. Medicine 2024 Nov;52(11):719-724. [doi: 10.1016/j.mpmed.2024.08.009]
77. Rook KS. The functions of social bonds: perspectives from research on social support, loneliness and social isolation. In: Sarason IG, Sarason BR, editors. Social Support: Theory, Research and Applications: Springer; 1985:243-267. [doi: 10.1007/978-94-009-5115-0_14]
78. Freak-Poli R, Ryan J, Tran T, et al. Social isolation, social support and loneliness as independent concepts, and their relationship with health-related quality of life among older women. Aging Ment Health 2022 Jul;26(7):1335-1344. [doi: 10.1080/13607863.2021.1940097] [Medline: 34219569]
79. Yu G, Yuan T, Bu S, Zeng Y. Psychological distress and cultural role conflict among dementia family caregivers in China: a phenomenological study. BMC Nurs 2025 Oct 30;24(1):1352. [doi: 10.1186/s12912-025-03996-9] [Medline: 41168797]
80. Fang Y, Niu Y, Dong Y. Exploring the relationship between narcissism and depression: the mediating roles of perceived social support and life satisfaction. Pers Individ Dif 2021 Apr;173:110604. [doi: 10.1016/j.paid.2020.110604]

XSL•FO
RenderX

81. Liao Z, Zhou H, He Z. The mediating role of psychological resilience between social participation and life satisfaction among older adults in China. BMC Geriatr 2022 Dec 8;22(1):948. [doi: 10.1186/s12877-022-03635-x] [Medline: 36482364]
82. Bar-Tur L. Fostering well-being in the elderly: translating theories on positive aging to practical approaches. Front Med (Lausanne) 2021;8:517226. [doi: 10.3389/fmed.2021.517226] [Medline: 33898472]
83. Duncan E, O'Cathain A, Rousseau N, et al. Guidance for reporting intervention development studies in health research (GUIDED): an evidence-based consensus study. BMJ Open 2020 Apr 8;10(4):e033516. [doi: 10.1136/bmjopen-2019-033516] [Medline: 32273313]

Abbreviations

DID: difference-in-differences
GDS-K: Geriatric Depression Scale-Korean version
MMSE: Mini-Mental State Examination Korean version
SEM: Structural equation modeling

Edited by R Bjarnadottir; submitted 18.Apr.2025; peer-reviewed by HW Kuo, M Tusconi, Z Zandesh; revised version received 23.Dec.2025; accepted 27.Feb.2026; published 05.May.2026.

Please cite as:
Zhao B, Lim JA, Lim S, Yoon HJ, Kim HK, Hong SM, Zhang X, Nam EW
Integrating Community and Digital Support Through Social Prescribing to Improve Mental Health in Rural Older Adults in South Korea: Quasi-Experimental Study
JMIR Aging 2026;9:e76201
URL: https://aging.jmir.org/2026/1/e76201
doi:10.2196/76201

© Bo Zhao, Joo Aeh Lim, Subeen Lim, Hye Jin Yoon, Hae Kyung Kim, Seok Mi Hong, Xinyu Zhang, Eun Woo Nam. Originally published in JMIR Aging (https://aging.jmir.org), 5.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Online Communication Attitudes and Video Game Co-Play in Older Adults: Cross-Sectional Mediation Study

Jeffrey Tsifan Tseng, MSG

Department of Communication, University of California, Davis, 1 Shields Ave, Davis, CA, United States

Corresponding Author:
Jeffrey Tsifan Tseng, MSG
Department of Communication, University of California, Davis, 1 Shields Ave, Davis, CA, United States

Abstract

Background: Social isolation among older adults is a growing public health concern. While information and communication technologies offer opportunities for social engagement, few studies have examined how video game co-play, a form of interactive digital media, supports intergenerational connection and perceived social support among older adults.

Objective: Guided by uses and gratifications theory, this study investigates whether older adults' online communication attitudes predict video game co-play with younger family members and whether such co-play enhances perceived social support.

Methods: A total of 433 older adults (aged ≥60 years) were recruited using a snowball sampling approach facilitated by undergraduate students. Participants completed an online survey assessing online social connection; online self-disclosure (OSD); video game co-play with younger family members; and perceived informational, instrumental, and emotional support. Regression analyses predicting co-play behavior (H1) were conducted using the full sample (N=433), with participants reporting no co-play coded as 0 to indicate the absence of engagement. Hierarchical regression analyses examining associations between co-play and social support (H2) were restricted to participants reporting video game co-play (n=71). Mediation analyses testing indirect effects (H3-H4) were also conducted among co-players (n=71), controlling for age, gender, and race. Mediation models were estimated using PROCESS Model 4, with 5000 bootstrap samples to generate 95% CIs.

Results: OSD was positively associated with video game co-play behavior (b=0.12, SE=0.03; P<.001), whereas online social connection was not significantly related to co-play (b=0.06, SE=0.03; P=.08). Video game co-play was positively associated with perceived informational (b=0.11, SE=0.04; P=.01) and instrumental (b=0.16, SE=0.04; P<.001) social support but not emotional support (b=0.07, SE=0.04; P=.053). Among 71 participants reporting co-play, mediation analyses indicated that video game co-play mediated the association between OSD and informational support (indirect effect b=0.01, 95% CI .001-0.03) and instrumental support (indirect effect b=0.02, 95% CI 0.005-0.03), with partial mediation for informational support and full mediation for instrumental support.

Conclusions: While not interventional, these findings suggest that video game co-play may function as a culturally relevant, digitally mediated relational practice through which older adults engage in intergenerational connection and perceive informational and instrumental support. The results inform digital health efforts aimed at reducing social isolation and underscore the importance of integrating digital literacy and attitudinal readiness into age-inclusive policy and program design. However, the small subsample of older adults who reported co-play behavior warrants caution in generalizing the findings and highlights the need for broader inclusion in future research.

(JMIR Aging 2026;9:e80541) doi:10.2196/80541

KEYWORDS

older adults; video game co-play; digital health; social support; intergenerational communication; online communication attitudes; uses and gratifications theory

Introduction

Background

As the global population ages, addressing social isolation among older adults has become a pressing public health concern. Digital health tools, especially information and communication technologies (ICTs), have emerged as promising means to support emotional well-being and social connectivity [1,2]. Among these tools, video games are increasingly recognized not just as recreational platforms but as interactive environments capable of facilitating intergenerational engagement [3]. As digital connectivity among older adults continues to grow, there is a critical need to examine the psychological drivers that shape their engagement in online activities and evaluate the social and relational benefits that follow [4,5].

XSL•FO
RenderX

Older adults' online communication attitudes, specifically online self-disclosure (OSD) and online social connection (OSC), may influence how they interact with digital platforms [6,7]. OSD refers to the willingness to share personal thoughts and experiences online, while OSC reflects the perception that digital tools help sustain interpersonal bonds [7,8]. These attitudes influence participation in relationship-enhancing digital activities, such as video game co-play.

Video game co-play, which involves collaborative or competitive gaming between older adults and younger family members, enables shared interactive experiences that extend beyond entertainment [3]. Through gameplay, participants engage in storytelling, problem-solving, and coordinated activity, processes that have been shown to strengthen family closeness, shared identity, and intergenerational understanding [9-11]. Prior research also suggests that co-play can foster emotional support and a sense of inclusion in digital spaces, particularly for older adults navigating geographical distance from family members [12,13].

Unlike more passive or asynchronous ICTs, such as text messaging or social media browsing, video game co-play is characterized by real-time interaction, shared goals, and sustained mutual attention [3,11,14]. These affordances create structured opportunities for collaboration, reciprocity, and informal support exchange, making co-play particularly well suited for intergenerational connection [3,11,12]. Focusing on co-play allows this study to isolate a form of digital engagement that is both socially immersive and relationally consequential for older adults [9,12].

To explore these dynamics, this study draws on uses and gratifications theory (UGT) as a foundational framework for understanding older adults' agency in selecting and engaging with digital media [15,16]. While UGT has traditionally been applied to explain media choice and individual-level gratifications, recent scholarship highlights the need to extend the theory to interactive, relational, and socially embedded digital practices [16], particularly in aging populations [2,17]. Within this expanded perspective, older adults' online communication attitudes, such as openness to self-disclosure and perceived importance of online connection, may shape their engagement in interactive activities, such as video game co-play, which offer opportunities for synchronous communication and shared relational experiences across generations.

Social support, including informational, instrumental, and emotional forms, is a key determinant of psychological well-being in later life [18,19]. In digital intergenerational contexts, support may emerge through cooperative gameplay and emotionally resonant exchanges [20,21]. Understanding these relationships provides insight into how digital tools can sustain social bonds and promote resilience among older adults.

By conceptualizing video game co-play as a digitally mediated relational practice rather than a formal intervention, this study highlights a culturally situated pathway through which older adults may sustain intergenerational connection across geographic distance. In doing so, it extends UGT to later life by emphasizing how motivational dispositions translate into interactive relational behaviors, while also acknowledging that access, digital literacy, and cultural norms may shape engagement. These insights provide a theoretically grounded foundation for future research and age-inclusive digital engagement strategies.

Literature Review: Uses and Gratifications Theory

UGT provides a foundational framework for understanding how individuals actively select media to fulfill specific psychological and social needs [15]. Unlike passive models of media effects, UGT emphasizes users' agency in engaging with communication technologies based on anticipated gratifications such as entertainment, escapism, information-seeking, and social connection [22,23]. Although originally developed in the context of mass media, contemporary extensions of UGT emphasize its applicability to interactive, socially embedded, and relational digital practices, moving beyond individual media choice to account for co-use and participatory engagement [16,23]. This approach underscores the functional role of media in addressing users' everyday emotional and relational goals.

For older adults, UGT provides valuable insights into the adoption of digital tools designed to enhance well-being and maintain social relationships. Research has shown that motivations, such as companionship, leisure, and family connection, are central drivers of technology use in aging populations [2,16]. These motivations are especially salient in intergenerational contexts, where digitally mediated shared activities, such as video game co-play, can foster emotional reciprocity and familial closeness.

UGT also emphasizes the significance of individual differences, particularly in attitudes toward digital communication, in determining how and why technologies are utilized. Individuals actively choose media based on psychological and social needs, and these choices vary according to personal dispositions, such as motivation, past experiences, and social context [23]. For older adults, attitudinal orientations, such as openness to self-disclosure or valuing online social connection, may influence not only whether they engage with specific digital platforms but also how effectively those tools meet their evolving social needs. Understanding these motivations is crucial for informing age-appropriate digital engagement strategies that extend beyond usability to foster lasting and meaningful participation. As such, UGT provides a valuable framework for examining how older adults integrate video game co-play into their relational routines, particularly as a means of maintaining connections and reducing isolation in long-distance family relationships. Building on this perspective, this study advances UGT by applying it to intergenerational digital interaction in later life, emphasizing how attitudinal orientations toward online communication shape engagement in interactive, relational media practices rather than passive consumption.

Significance of Video Game Co-Play for Older Adults

Video game co-play has emerged as an innovative and increasingly meaningful form of intergenerational interaction, particularly for older adults seeking to sustain connections with younger family members [3,12]. More than a source of entertainment, co-playing video games provides a structured, immersive experience that fosters real-time engagement,

XSL•FO
RenderX

reciprocal communication, and collaborative problem-solving [3,14]. These interactions can serve as digital pathways for maintaining familial bonds and promoting intergenerational understanding.

For older adults, video game co-play can yield significant emotional and psychosocial benefits. Shared gameplay fosters mutual support and relational closeness, buffering against the health risks of isolation [14,24]. Studies found that older adults who engaged in digital gaming reported significantly higher levels of emotional well-being and reduced depressive symptoms compared to nonplayers [24]. Moreover, social gaming interventions have demonstrated feasibility and effectiveness in promoting intergenerational connection and alleviating loneliness. For example, past study has found that mobile gaming apps designed to facilitate intergenerational interaction can help to increase engagement and reduce social isolation among older participants [25]. These experiences enhance inclusion and self-worth, improving emotional health in later life.

Importantly, video game co-play supports synchronous engagement across physical distances, making it especially relevant for families separated by geography or mobility constraints [12,26]. Unlike more passive forms of communication, such as text messaging or phone calls, gameplay fosters dynamic interaction and shared accomplishment. Through these playful yet meaningful exchanges, older adults and younger relatives can maintain continuity in their relationships despite limited face-to-face contact [27].

In this sense, video game co-play can be understood as a low-barrier, home-based digitally mediated relational practice that holds relevance for digital health and social inclusion efforts. Its accessibility, emotional salience, and interactive format make it a promising tool for enhancing relational satisfaction and psychological well-being among older populations. As such, integrating co-play into broader digital health strategies and aging support programs may offer new avenues to promote social connectedness in aging communities.

Importantly, cross-cultural and comparative research suggests that engagement in intergenerational digital activities is deeply shaped by sociocultural norms related to family obligation, leisure, and emotional expression [28,29]. Prior research indicates that collectivist cultural contexts may place greater emphasis on shared activities and instrumental caregiving, potentially influencing both the adoption of intergenerational co-play and perceptions of its relational value [30,31]. In such contexts, shared digital practices may function as extensions of everyday family caregiving rather than as purely recreational activities [32]. Recognizing these cultural contingencies is essential for interpreting findings related to intergenerational digital engagement and for avoiding overgeneralization across diverse aging populations.

Online Communication Attitudes

Within the UGT framework, online communication attitudes serve as important predictors of media selection and engagement. Drawing on the social influence model of technology use, research shows that individuals' engagement with digital platforms is shaped by their preexisting attitudes toward mediated communication [33,34]. This highlights the central role of attitudinal dispositions, particularly among older adults, in influencing how and why digital tools are used. Attitudes are composed of both affective and cognitive dimensions, which guide individuals' behavioral tendencies [35]. The theory of reasoned action further posits that attitudes and behavioral intentions are key predictors of actual behavior, particularly in contexts that are socially salient, such as maintaining family relationships [36].

Two primary components of online communication attitudes include OSC and OSD [6,8]. OSC refers to the belief that digital technologies are essential for sustaining interpersonal relationships, while OSD reflects one's comfort with sharing personal information in online environments [6,7]. Individuals high in OSC may view digital engagement as a necessary aspect of their social well-being, whereas those high in OSD tend to interact more frequently and meaningfully across digital platforms [6-8]. These dimensions are especially relevant in older populations, as they help explain variability in how digital tools are adopted and used for relationship maintenance.

In the context of aging and family communication, OSC holds particular significance. For older adults with mobility limitations or who are geographically separated from loved ones, maintaining close contact often relies on digital channels [1,37]. Technologies, such as video games, offer new, interactive avenues for sustaining these relationships. Prior studies have shown that older adults who view online platforms as socially vital are more inclined to adopt tools that promote interactive engagement, including video game co-play [2,38]. From this perspective, co-play can be seen as a relational strategy grounded in OSC, a form of active, synchronous interaction that reinforces intergenerational connection.

Similarly, OSD may be a critical enabler of co-play behavior. The willingness to share personal experiences and emotions in digital contexts can deepen relationships and foster more meaningful interactions [7]. In video game environments, where gameplay often involves storytelling, collaboration, and real-time feedback [3], older adults who are comfortable with online self-disclosure may be more inclined to engage. Research has found that individuals high in OSD tend to be more active across various forms of digital media, including those that support emotionally rich exchanges [39,40]. This openness may enhance the depth of intergenerational connection within the gaming context.

Building on these insights, this study predicts that online communication attitudes, specifically OSC and OSD, will significantly influence the extent to which older adults participate in video game co-play with younger family members.

Hypothesis 1 (H1): (a) OSC and (b) OSD will be significantly associated with video game co-play behavior among older adults in the context of long-distance family relationships.

Video Game Co-Play and Perceived Social Support

Building on the understanding that attitudes influence technology adoption and use [39,41,42], the construct of perceived social support is central to evaluating how digital

XSL•FO
RenderX

interaction fosters relational well-being among older adults. Perceived social support encompasses several key dimensions: informational support (eg, sharing advice and guidance), instrumental or tangible support (eg, providing assistance or resources), emotional support (eg, offering empathy and care), social network support (eg, fostering connectedness and belonging), and esteem support (eg, affirming personal worth) [43]. These forms of support collectively contribute to emotional resilience and psychological health in later life.

In intergenerational contexts, family members are typically the primary providers of social support. Emotional support from close relatives is strongly associated with increased life satisfaction and psychological well-being in older adults [44], while instrumental support is essential for navigating health care systems, managing daily tasks, and responding to crises [45]. Informational support, in turn, plays a critical role in promoting autonomy and informed decision-making [46]. These resources support older adults' well-being and integration [47].

When geographic distance limits in-person interaction, digital communication technologies become vital tools for sustaining these support systems. Research indicates that ICTs enable older adults to maintain emotional intimacy and receive both practical and emotional assistance across distances [48,49]. Through interactive platforms, such as video games, older adults can share experiences, collaborate on tasks, and engage in emotionally resonant dialog with younger family members [38,50]. These forms of digital engagement can reinforce a sense of belonging, strengthen emotional ties, and affirm older adults' roles within their family networks.

Empirical studies have documented how ICT engagement improves perceptions of social support and reduces feelings of loneliness [17,51]. In senior living settings, technology use has been linked to enhanced social interaction, community integration, and emotional well-being [1]. Applying these findings to video game co-play, it is reasonable to hypothesize that engaging in shared digital gameplay with younger family members may lead to increased perceptions of informational, instrumental, and emotional support, especially in the context of long-distance relationships.

Hypothesis 2 (H2): Video game co-play will be positively associated with perceived (a) informational, (b) instrumental, and (c) emotional social support among older adults in long-distance relationships.

Mediating Role of Video Game Co-Play

While general ICT use has been shown to mediate the relationship between online communication attitudes and perceived social support [52], video game co-play represents a particularly interactive and emotionally rich form of digital engagement that may offer a stronger pathway to familial bonding [14,27]. Older adults who exhibit positive attitudes toward OSC and OSD are more likely to adopt collaborative digital platforms that enable synchronous, emotionally meaningful interaction [2,7]. Video game co-play exemplifies such a platform, providing not only entertainment but also opportunities for shared achievement, relational expression, and emotional reciprocity [3].

Co-playing video games encourages sustained communication, cooperation, and joint problem-solving, which are core components of emotionally supportive relationships and intergenerational understanding [3,27]. These dynamics are especially salient for older adults seeking to maintain connections with geographically distant family members. The interactive nature of co-play allows older users to both express care and receive encouragement through real-time engagement, reinforcing feelings of value and inclusion within family networks [17]. This process not only enhances emotional closeness but also strengthens perceptions of informational, instrumental, and emotional support [38,51].

Based on this theoretical and empirical foundation, this study examines a mediation model in which video game co-play serves as the mechanism through which online communication attitudes influence perceived social support among older adults. This framework positions co-play not merely as an outcome of digital attitudes but as a critical conduit for transforming psychological dispositions into relational benefits in digitally mediated family interactions.

Hypothesis 3 (H3): Video game co-play will mediate the relationship between OSC and perceived (a) informational, (b) instrumental, and (c) emotional social support among older adults in long-distance family relationships.

Hypothesis 4 (H4): Video game co-play will mediate the relationship between OSD and perceived (a) informational, (b) instrumental, and (c) emotional social support among older adults in long-distance family relationships.

Methods

Participants

Older adults aged 60 years and above were recruited using a snowball sampling approach through undergraduate students at a large West Coast university. Each college-aged participant was instructed to invite 1 older adult family member (aged 60 years or older) who resided in a different geographic location. The long-distance nature of the intergenerational dyads was determined based on participant self-report and supplemented by IP address filtering to reduce the likelihood that dyad members were responding from the same physical location. This recruitment method was incentivized through extra credit for the undergraduate students; no monetary or nonmonetary incentives were provided to older adult participants.

Initially, a total of 1306 older adult participants were recruited (mean age 70.14, SD 7.23 y). To ensure data quality, participants who had excessive missing data or spent insufficient time on the survey (ie, less than 300 s, based on pilot test standards) were excluded, resulting in the removal of 638 cases. An additional 235 participants were excluded due to missing data on the video game co-play measure, defined as incomplete or nonresponse to the co-play items administered to participants who indicated engagement in video game play. The final sample consisted of 433 older adults, ranging in age from 60 to 95 years. Of these 433 participants, 71 reported engaging in video game co-play with their younger family members, whereas 362 reported no co-play engagement. For analyses examining

XSL•FO
RenderX

predictors of co-play behavior (H1), participants who reported no co-play were coded as 0 to indicate the absence of engagement. However, mediation analyses requiring the continuous co-play composite were restricted to participants reporting co-play (n=71). Among these co-players, 41 were female and 30 were male, and the majority were Asian (53/71, 75%), followed by White and Hispanic participants. The gender composition of the full sample was 289 female participants and 144 male participants. In terms of ethnicity, the full sample included 252 Asian participants, 96 White participants, and 85 Hispanic participants.

Ethical Considerations

All procedures were approved by the university's institutional review board (1655269-1). College students were recruited through a study subject pool and received course extra credit for assisting with recruitment. A link to a survey on Qualtrics was provided after each college student gave informed consent; the survey requested the email address of 1 older adult family member aged 60 years or older. The system automatically sent the survey to the electronic address provided by the college student's family member. Older adult participation was voluntary, and no monetary or nonmonetary incentives were provided to older adult participants. The IP address was used as a filter to ensure that respondents were not located at the same physical address and to prevent multiple survey submissions from the same individual, providing supplemental verification of the reported long-distance nature of intergenerational dyads. The web-based survey was administered using a secure online platform between September 2020 and June 2021, and study procedures are reported in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) guidelines (Checklist 1). Because the survey was distributed via individualized referral emails rather than a public website, standard web-based response metrics, such as view rate and participation rate, could not be calculated.

Measures

Overview

The survey began with demographic information questions (ie, gender, age, and race). The rest of the survey had items measuring video game co-playing, perceived social support, and online communication attitudes. Participants were instructed to answer the items for their video game co-playing and perceived social support while considering their familial counterparts.

Online Communication Attitudes

This factor was measured with the OSD and OSC subscales [6]. OSD included 7 items, and OSC had 6 items. Both factors were framed as 7-point Likert-type scales (1=strongly disagree and 7=strongly agree). The OSD subscale included statements such as "I feel like I can sometimes be more personal during internet conversations" and "I feel like I can be more open when I am communicating online." The OSC subscale contained items such as "Without the internet, my social life would be drastically different" and "I would communicate less with my friends if I could not talk with them online." OSD (Cronbach α=0.91) and OSC (Cronbach α=0.83) for older adult participants showed good reliability.

Video Game Co-Play

Older adults' engagement in collaborative video game play with younger family members was assessed using 3 items adapted from Wang et al [3]. Participants reported (1) the typical duration of a video game session with a younger family member (eg, "When you play video games with your younger family member, about how long do you play at a time?"), (2) the percentage of their total gaming time spent co-playing with that family member, and (3) the estimated percentage of the younger family member's gaming time spent playing with them. These items captured complementary behavioral and perceptual indicators of co-play frequency and mutual engagement. Only participants who reported engaging in video game co-play completed these items. Participants who indicated no co-play engagement were not administered the follow-up items and were treated as missing for analyses requiring the continuous co-play composite. For analyses examining predictors of co-play behavior (H1), non–co-players were coded as 0 to reflect the absence of engagement. To construct the composite index, each item was standardized (*z*-score) to account for differences in scale and metric and then averaged to form a single co-play score, with higher values indicating greater engagement in intergenerational co-play. The composite demonstrated good internal consistency (Cronbach α=0.83).

Perceived Social Support

This factor was assessed using the UCLA social support inventory [53]. This scale measured informational (5 items), instrumental (5 items), and emotional support (10 items). The items were framed as a 5-point Likert-type scale (1=never to 5=very often). The items were reworded to target intergenerational relationships. Sample items included "how often did your younger/older family member provide assistance to you within the past three months?" and "how satisfied or dissatisfied have you been with the love and caring you've received from your younger/older family member within the past three months?" The scale showed good reliability for the older adult participants (Cronbach α=0.88).

Results

The mean scores for online communication attitudes, video game co-play frequency, and perceived social support were calculated. Descriptive statistics presented in Tables 1 and 2 reflect the full analytic sample (N=433). All regression and mediation analyses controlled for age, gender, and race (with White participants as the reference category) [54,55]. Regression analyses examining predictors of video game co-play behavior (H1) were conducted using the full sample (N=433), with participants who reported no co-play coded as 0 to indicate the absence of engagement. Hierarchical regression analyses were conducted for H2, with age, gender, and race entered in step 1 and video game co-play entered in step 2 to examine its incremental contribution to perceived social support. Mediation analyses (H3 and H4) were conducted using PROCESS Model 4 [56]. Because the continuous co-play composite was calculated only for participants reporting co-play engagement, cases with

XSL•FO
RenderX

missing mediator values were excluded by PROCESS, resulting in a mediation sample of 71 participants. Indirect effects were estimated using 5000 bootstrap samples to generate 95% CIs.

Table . Descriptive statistics among study variables (N=433).

Study variables	Range	Minimum	Maximum	Mean (SD)
Online social connection	6.00	1.00	7.00	2.88 (1.29)
Online self-disclosure	5.57	1.00	6.57	2.72 (1.27)
Video game co-play score[a,b,c]	5.54	−0.32	5.22	−0.00 (0.86)
Perceived informational social support	4.26	1.17	5.43	3.40 (0.74)
Perceived instrumental social support	4.50	1.00	5.50	3.31 (0.80)
Perceived emotional social support	3.75	2.00	5.75	4.05 (0.67)

[a]The video game co-play score is a composite variable derived from 3 standardized items: typical session duration, percentage of the older adult's gaming time spent co-playing, and percentage of the younger relative's gaming time spent with the older adult.

[b]The composite was calculated among participants reporting video game co-play engagement (n=71).

[c]Higher scores indicate greater frequency and intensity of video game co-play.

Table . Pearson correlations among study variables (N=433).

Study variables	1	2	3	4	5	6
Online social connection	—					
Online self-disclosure	0.36[a]	—				
Video game co-play score[b]	0.09	0.19[a]	—			
Perceived informational social support	0.03	0.14[c]	0.16[a]	—		
Perceived instrumental social support	−0.10	0.09	0.21[a]	0.62[a]	—	
Perceived emotional social support	0.001	0.07	0.09	0.59[a]	0.55[a]	—

[a]P<.001.

[b]Regression and mediation analyses involving the continuous co-play composite were restricted to participants reporting co-play (n=71).

[c]P<.01.

H1 posits that older adults' online communication attitudes of (a) OSC and (b) OSD would be associated with video game co-play behavior. Hierarchical regression analyses were conducted, with age, gender, and race entered at step 1 and online communication attitudes entered at step 2. For OSC, step 1 model was significant ($F_{3,429}$=3.53; P=.03) and explaining 2% of the variance (R^2=0.02). However, the addition of OSC in step 2 did not significantly improve the model (ΔR^2=0.01, F change$_{1,428}$=3.10; P=.08). The final model was significant ($F_{4,428}$=3.40; P=.02), explaining 3% of the variance (R^2=0.03), but OSC was not significantly associated with video game co-play behavior (b=0.06, SE=0.03, 95% CI −0.01 to 0.12; P=.08). For OSD, step 2 did significantly improve the model (ΔR^2=.04, F change$_{1,428}$=12.85; P<.001). The final model was significant ($F_{4,428}$=6.70; P<.001), explaining 6% of the variance (R^2=0.06). Moreover, OSD was significantly associated with video game co-play behavior (b=0.12, SE=0.03, 95% CI 0.05-0.18; P<.001). CIs were estimated using 5000 bootstrap samples with percentile-based intervals.

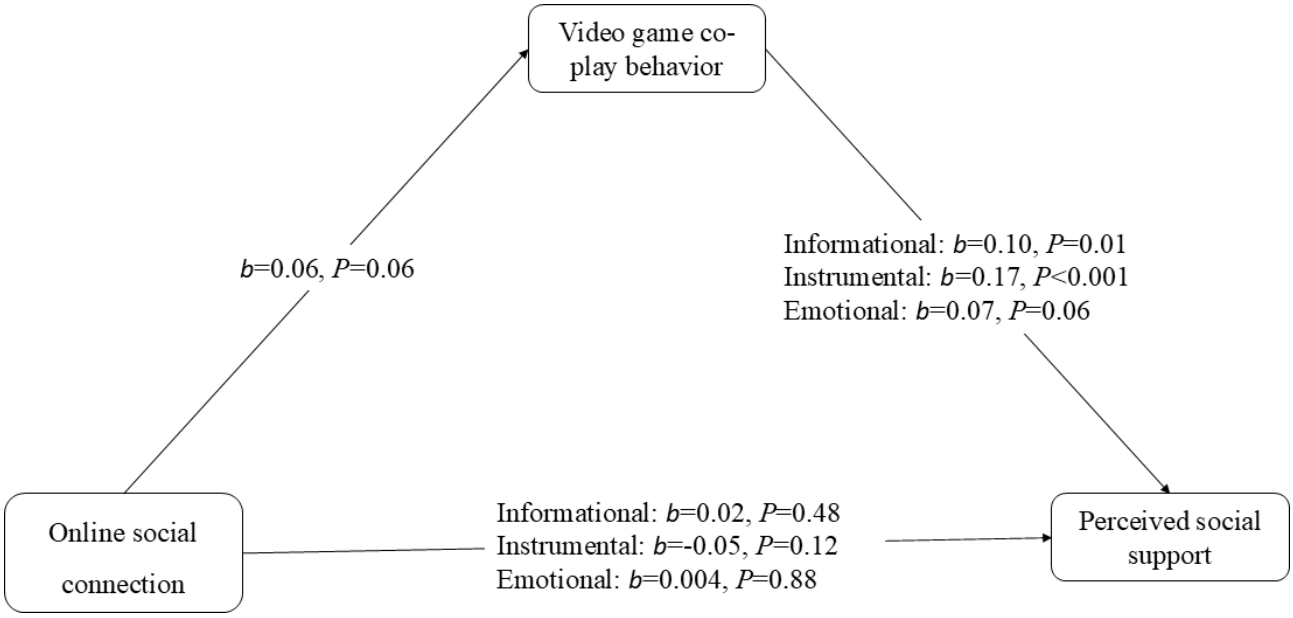
H2 posits that video game co-play would be positively associated with perceived (a) informational, (b) instrumental, and (c) emotional social support among older adults in long-distance relationships. Hierarchical regression analyses were conducted, with age, gender, and race entered in step 1 and video game co-play entered in step 2 (n=71). For informational support, the step 1 model was significant ($F_{3,67}$=7.01, R^2=0.06; P<.001). The addition of video game co-play in step 2 significantly improved the model (ΔR^2=0.02,

XSL•FO
RenderX

F change$_{1,66}$=7.10; *P*=.009). The final model was significant ($F_{4,66}$=6.90; *P*<.001), explaining 8% of the variance (R^2=0.08). Video game co-play was positively associated with informational support (*b*=0.11, SE=0.04, 95% CI 0.03-0.19; *P*=.01). For instrumental support, the step 1 model accounted for 11% of the variance (R^2=0.11, $F_{3,67}$=13.15; *P*<.001). Adding video game co-play significantly increased explained variance (ΔR^2=0.03, *F* change$_{1,66}$=13.92; *P*<.001). The final model was significant ($F_{4,66}$=13.62; *P*<.001), accounting for 14% of the variance (R^2=0.14), and video game co-play was positively associated with instrumental support (*b*=0.16, SE=.04, 95% CI 0.08-0.24; *P*<.001). For emotional support, the step 1 model including age, gender, and race was not statistically significant ($F_{3,67}$=1.65; *P*=.16), explaining 1% of the variance (R^2=0.01). Adding video game co-play in step 2 did not significantly improve the model (ΔR^2=0.01; *F* change$_{1,66}$=3.76; *P*=.07). The final model was not statistically significant ($F_{4,66}$=2.07; *P*=.07), explaining 2% of the variance (R^2=0.02). Consistent with the overall model, video game co-play was not significantly associated with emotional support (*b*=0.07, SE=.04, 95% CI −0.001 to 0.15; *P*=.053).

H3 posits that video game co-play will mediate the relationship between OSC and (a) informational, (b) instrumental, and (c) emotional social support among older adults in long-distance family relationships. Mediation analyses (PROCESS Model 4; 5000 bootstrap samples) indicated that the model predicting video game co-play was significant ($F_{4,66}$=4.18; *P*<.001), explaining 5% of the variance (R^2=0.05). However, because no significant association was found between OSC and video game co-play (*b*=0.06, SE=0.03, 95% CI −0.001 to 0.13; *P*=.06), the mediation hypothesis was not supported (Figure 1).

Figure 1. Mediation model of online social connection, video game co-play behavior, and perceived social support. This figure illustrates the mediation model examining the association between online social connection (OSC), video game co-play behavior, and perceived social support among older adults. Values represent unstandardized regression coefficients (*b*). The association between OSC and video game co-play was not statistically significant (*b*=0.06, SE=0.03; *P*=.06). Indirect effects were estimated using 5000 bootstrap samples. All paths were adjusted for age, gender, and race (White as the reference category).



H4 posits that video game co-play mediates the relationship between OSD and (a) informational, (b) instrumental, and (c) emotional social support among older adults in long-distance family relationships. For H4a, mediation analyses (PROCESS Model 4; 5000 bootstrap samples) indicated that the association between OSD and informational support was partially mediated by video game co-play. The model predicting video game co-play was significant ($F_{4,66}$=5.57; *P*<.001), explaining 6% of the variance (R^2=0.06). Higher OSD was associated with greater video game co-play (*b*=0.11, SE=0.03, 95% CI 0.04-0.17; *P*<.001). The overall model predicting informational support was also significant ($F_{5,65}$=6.61; *P*<.001), accounting for 9% of the variance (R^2=0.09). Video game co-play was positively associated with informational support (*b*=0.09, SE=0.04, 95% CI 0.01-0.17; *P*=.02). The direct effect of OSD on informational support remained significant after including video game co-play in the model (*b*=0.06, SE=0.03, 95% CI 0.001-0.11; *P*=.04), indicating partial mediation. The indirect effect of OSD on informational support through video game co-play was statistically significant (*b*=0.01, SE=0.01, 95% CI 0.001-.03).

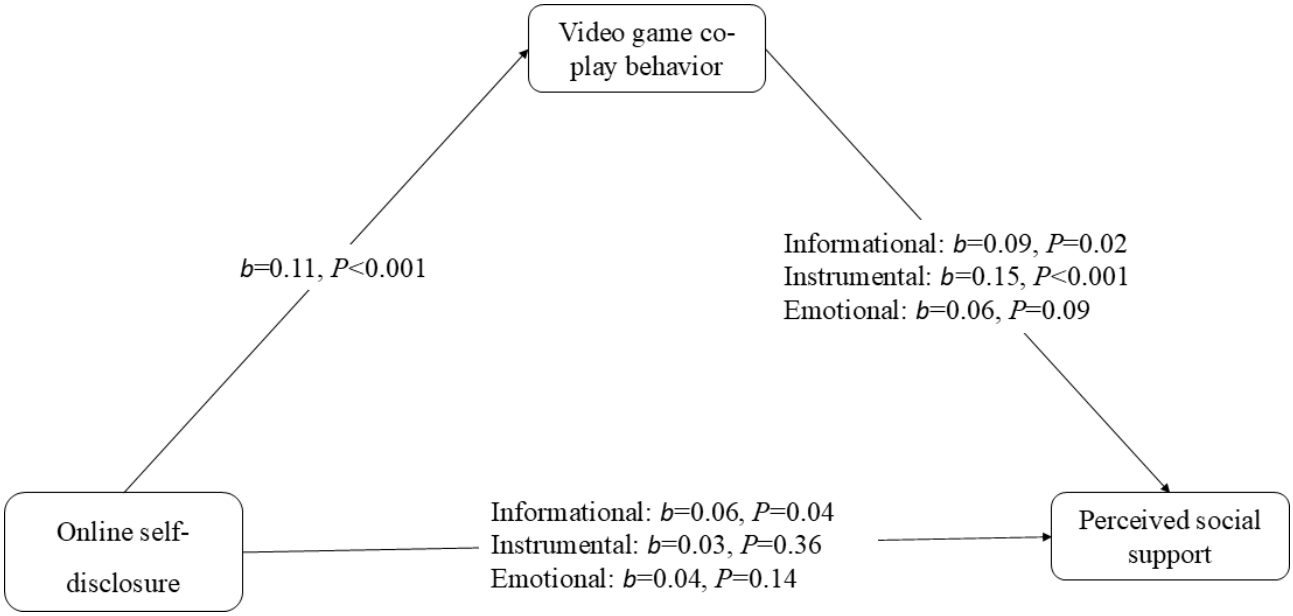
For H4b, mediation analyses (PROCESS Model 4; 5000 bootstrap samples) supported full mediation of the association between OSD and instrumental support through video game co-play. The model predicting video game co-play was significant ($F_{4,66}$=5.57; *P*<.001), explaining 6% of the variance (R^2=.06); OSD was positively associated with co-play (*b*=0.11, SE=.03, 95% CI 0.04-0.17; *P*<.001). The overall model predicting instrumental support was significant ($F_{5,65}$=11.5; *P*<.001), accounting for 14% of the variance (R^2=0.14), and co-play was positively associated with instrumental support (*b*=0.15, SE=0.04, 95% CI 0.07-0.24; *P*<.001). The indirect effect of OSD on instrumental support through co-play was statistically significant (*b*=0.02, SE=0.01, 95% CI 0.005-0.03), while the direct effect of OSD on instrumental support was not

XSL•FO
RenderX

significant after including co-play (b=0.03, SE=0.03, 95% CI −0.03 to 0.09; P=.36), consistent with full mediation.

For H4c, mediation analyses (PROCESS Model 4; 5000 bootstrap samples) did not support mediation of the association between OSD and emotional support through video game co-play. The model predicting video game co-play was significant ($F_{4,66}$=5.57; P<.001), explaining 6% of the variance (R^2=0.06). However, the overall model predicting emotional support was not statistically significant ($F_{5,65}$=2.11; P=.051), explaining 3% of the variance (R^2=0.03). Video game co-play was not significantly associated with emotional support (b=0.06, SE=0.04, 95% CI −0.01 to 0.13; P=.09). Consistent with this pattern, the indirect effect of OSD on emotional support through co-play was not statistically significant (b=0.01, SE=0.01, 95% CI −0.001 to 0.02; Figure 2).

Figure 2. Mediation model of online self-disclosure, video game co-play behavior, and perceived social support. Mediation model examining the association between online self-disclosure (OSD) and perceived social support through video game co-play among older adults. Values represent unstandardized regression coefficients (b). OSD was positively associated with video game co-play (b=.11, SE=.03; P<.001). Video game co-play was positively associated with informational support (b=.09, SE=.04; P=.02) and instrumental support (b=.15, SE=.04; P<.001) but not emotional support (b=.06, SE=.04; P=.09). Indirect effects were estimated using 5000 bootstrap samples. All paths were adjusted for age, gender, and race (White as the reference category).



Discussion

Principal Findings

This study advances understanding of how older adults engage with interactive digital platforms to sustain long-distance intergenerational bonds by extending UGT into the context of aging and digitally mediated family relationships. While UGT has traditionally been applied to explain media selection and individual-level gratifications [15], these findings highlight its relevance for understanding how older adults' online communication attitudes translate into interactive, relational digital practices. In particular, OSD emerged as a key motivational driver of video game co-play, which, in turn, was associated with enhanced informational and instrumental social support [7,39]. These findings position video game co-play not as passive media consumption but as an active, socially embedded digital health behavior through which older adults seek connection, reciprocity, and practical support within intergenerational family networks.

Although the findings indicate meaningful associations between video game co-play and perceived social support, only a small proportion of older adults in the sample reported engaging in co-play with younger family members (n=71). This limited prevalence suggests that intergenerational gaming is not yet a normative practice among older adults and may reflect broader barriers related to digital literacy, access, physical ability, or generational perceptions of gaming. The modest size of the co-player subsample necessarily limits statistical power and warrants cautious interpretation of the mediation findings, which should be viewed as exploratory rather than confirmatory. At the same time, the low uptake of co-play itself constitutes an important empirical insight, underscoring the need for targeted digital inclusion strategies that lower entry barriers and expand access to interactive digital activities for aging populations. These structural and attitudinal barriers warrant explicit consideration when interpreting these findings and when evaluating the broader applicability of video game co-play as a digitally mediated practice for aging populations.

Barriers to Access and the Digital Divide

Although the findings suggest that video game co-play may serve as a digitally mediated pathway to informational and instrumental social support among older adults, these benefits are not equally accessible to all aging populations. Structural barriers, such as access to gaming devices, reliable internet connectivity, and appropriate software, remain persistent challenges, particularly for older adults with limited socioeconomic resources [57,58]. In addition, individual-level barriers, including limited prior gaming experience, lower digital confidence, age-related sensory or cognitive limitations, and

XSL•FO
RenderX

the usability demands of commercial video games, may further constrain participation in video game co-play [39,59]. These barriers highlight the uneven distribution of opportunities to engage in interactive digital practices and underscore the importance of considering digital inequities when interpreting these findings [60-62].

Importantly, the current sample likely reflects older adults who already possess a baseline level of digital access and attitudinal readiness, which may partially explain the observed associations between online self-disclosure, co-play behavior, and perceived social support [39,58,59]. Engagement in video game co-play in this context often depended on younger family members initiating, facilitating, or sustaining participation, suggesting that intergenerational support may play a critical role in mitigating aspects of the digital divide [3,11,38]. From a digital health perspective, these findings emphasize the need for age-friendly design, intergenerational scaffolding, and community-based programs that reduce technological complexity and proactively address access barriers rather than assuming universal digital participation among older adults [62,63].

Health Literacy and Culturally Relevant Digital Engagement

Beyond access, effective engagement with digitally mediated practices also requires adequate digital health literacy, including confidence, perceived relevance, and comfort with online interaction [57,58,64]. Although digital health literacy was not directly measured in this study, online communication attitudes function as theoretically meaningful proxies for older adults' readiness to engage with interactive digital tools. In this study, OSD and OSC may be understood as psychosocial indicators of digital engagement readiness rather than measures of technical proficiency alone [8,39,59]. Older adults who feel comfortable sharing personal experiences online and who perceive digital platforms as socially meaningful may be better positioned to engage in interactive digital activities such as video game co-play and derive relational benefits from those interactions [16,41,51]. This interpretation aligns with broader digital health literature, emphasizing that motivation, trust, and perceived usefulness are central components of digital health literacy in later life [41,64].

Cultural context further shapes how older adults engage with digital technologies and how the benefits of such engagement are experienced [28,29,31]. The predominantly Asian composition of the co-playing subsample suggests that intergenerational digital engagement may be embedded within cultural norms, emphasizing family obligation, shared leisure, and instrumental support [30,32,65]. In such contexts, video game co-play may function less as recreational play and more as an extension of everyday family caregiving and relational maintenance [3,12,14]. These findings highlight the importance of culturally relevant digital health strategies that align with older adults' relational values and lived experiences, rather than relying on one-size-fits-all approaches to digital literacy or engagement. Designing inclusive digital programs that recognize cultural variation in communication norms and family roles may enhance both participation and perceived benefits among diverse aging populations [58,63].

Building on these findings, the study contributes theoretically by clarifying how online communication attitudes translate into relational outcomes through specific forms of media co-engagement. Although prior research has broadly documented the role of ICT use in maintaining familial connections [2,60], these findings identify video game co-play as a distinct, immersive activity that mediates the relationship between online self-disclosure and perceived social support. In particular, co-play fully mediated the association between OSD and instrumental support, suggesting that interactive gameplay creates conditions conducive to practical assistance from younger family members. This interpretation aligns with prior evidence that older adults derive both emotional and functional value from collaborative, task-oriented forms of communication [3,27].

This study extends prior research on intergenerational video game co-play, which primarily focused on younger populations and parent-child dynamics within coresiding families [3]. By contrast, this study uniquely focuses on older adults in geographically dispersed families, incorporating a broader age range and explicitly examining long-distance intergenerational ties. Furthermore, while a previous study emphasized family closeness and satisfaction [3], this study disaggregates the social support construct into informational, instrumental, and emotional components, providing a more nuanced understanding of how specific forms of support are cultivated through co-play. In doing so, it reframes video game co-play as a digitally mediated relational practice for aging populations rather than merely a recreational activity.

Drawing on Cutrona and Suhr's [43] social support typology, the results suggest that co-play enhances both informational and instrumental support through real-time problem-solving and skill-sharing. These findings align with prior evidence showing that active ICT engagement enhances the collaborative and cognitive capacities of older adults [17]. Within gameplay contexts, younger relatives may offer advice, guidance, or co-navigate challenges, reinforcing perceptions of competence and shared achievement among older players.

Implications for Policy and Community-Based Digital Programs

Practically, these findings inform the design of family-based digital health programs and future intervention development. They align with aging and digital inclusion frameworks, such as World Health Organization's Age-Friendly Digital Ecosystems and the Centers for Disease Control and Prevention's Healthy Aging initiatives, which emphasize digital equity, intergenerational connection, and inclusive design [62,63,66]. Integrating co-play into these frameworks could inform future health technology policy. Echoing prior research on shared media use and relational bonds, this study shows that video games can support both emotional and task-oriented intergenerational support [14], especially for families separated by distance or mobility constraints. Providers may integrate co-play into care programs to reduce isolation. Structured co-play modules in home or community services could enhance

XSL•FO
RenderX

engagement. Game features that encourage storytelling, collaboration, or personalization may enhance relational outcomes. Designers should prioritize emotional expressiveness and ease of use to maximize accessibility and meaningful interaction. By framing video game co-play as a relational digital health behavior rather than a recreational activity alone, this study aligns with aging-in-place and digital equity initiatives that emphasize socially meaningful technology use to reduce isolation among older adults [38,62,63].

Interestingly, although OSC was expected to predict co-play in video games, this relationship was not statistically significant. One explanation may be the difference between valuing online connections and engaging in specific behaviors, such as gaming. Older adults might appreciate digital connectivity but prefer more familiar tools (eg, email, video calls) over games, which may still seem youth-oriented [67,68]. Measurement limitations or ceiling effects may also have constrained variability in OSC scores, reducing predictive strength. Future research could further explore these dynamics, using qualitative approaches to examine older adults' preferences and comfort levels with various digital platforms.

Notably, video game co-play was more strongly associated with informational and instrumental support than with emotional support. This pattern suggests that structured, goal-oriented digital interactions may be particularly effective for facilitating task-focused assistance, such as problem-solving, guidance, and practical coordination, rather than explicit emotional expression. Emotional support may be communicated through more direct or verbally intimate modalities, such as video calls or messaging [48,49], whereas gameplay contexts emphasize shared activity and collaboration. In addition, cultural norms surrounding emotional expression, especially within collectivist family systems, may privilege instrumental forms of care over overt emotional disclosure [31,32,69]. Thus, the nonsignificant association with emotional support does not imply an absence of affective value but rather highlights the importance of considering how different digital modalities afford distinct types of social support.

These patterns are consistent with cross-cultural research on aging, family interaction, and digital engagement, which demonstrates that cultural norms strongly shape both technology use and expectations of social support in later life [31,58,59]. Comparative studies suggest that, in collectivist cultural contexts, family-based digital activities often emphasize practical assistance, shared responsibility, and instrumental caregiving over explicit emotional disclosure [29,30]. Emotional support may be conveyed implicitly through sustained participation and shared activities rather than through overt verbal expression [31,32]. Viewed through this cross-cultural lens, the observed association between video game co-play and informational and instrumental support, rather than emotional support, aligns with broader evidence on how culturally embedded family practices shape digitally mediated intergenerational relationships.

The predominantly Asian composition of the sample provides an important sociocultural context for interpreting the findings. Cultural values emphasizing familial obligation, interdependence, and instrumental caregiving may shape both engagement in intergenerational co-play and perceptions of support derived from such interactions. In collectivist cultural contexts, shared digital activities may function not only as leisure but also as relational maintenance practices that reinforce family roles and responsibilities [65,70]. While this cultural specificity limits the generalizability of the findings to more individualistic populations, it also offers valuable insight into how culturally embedded family norms intersect with digital media use in later life. Future cross-cultural research is needed to examine how gaming practices and support dynamics vary across cultural, national, and socioeconomic contexts. Although participants were recruited through a US-based university, the national context was not independently verified, and the findings should therefore be interpreted within this institutional and sociocultural setting.

These findings offer practical implications for digital game design and public policy. Developers should design intergenerationally friendly features, such as accessible interfaces, inclusive narratives, and cooperative challenges that foster caregiving dialogue. Policymakers and health professionals can support co-play integration into aging services, including telehealth, community centers, and digital literacy programs. Aging agencies and inclusion initiatives should embed co-play into broader digital engagement strategies, with funding for training and partnerships with libraries, senior centers, and telehealth platforms to pilot intergenerational gaming as a psychosocial support tool. The findings should be interpreted as observational evidence that informs future intervention design rather than as demonstrating causal intervention effects.

As synthesized above, while video game co-play shows promise, equitable implementation requires explicit attention to persistent socioeconomic disparities, digital literacy gaps, and physical or cognitive limitations that constrain older adults' access to interactive digital media [57]. Building on the barriers and literacy considerations discussed above, strategies such as device provision, subsidized internet access, and tailored digital skills training delivered through community-based settings are essential for inclusive adoption.

Intergenerational digital literacy programs, in which younger family members scaffold older adults' engagement in co-play, may simultaneously support skill acquisition and relationship building [71]. To ensure cultural relevance, interventions should incorporate design elements aligned with diverse familial values and communication norms. Multilingual interfaces, culturally familiar imagery, and collaborative mechanics suited to collectivist traditions can enhance engagement.

Age-inclusive policies must move beyond usability to support emotionally resonant, culturally congruent experiences. This study offers empirical grounding for future interventions that design and evaluate digital tools promoting relational well-being. Emphasizing co-play as a vehicle for relational support underscores the need for culturally and socially responsive technology–enabled aging innovation.

XSL•FO
RenderX

Limitations and Future Directions

Several limitations warrant consideration when interpreting this study's findings. First, the study used a cross-sectional design, which limits causal inference regarding the relationships among online communication attitudes, video game co-play, and perceived social support. Although mediation analyses were theoretically grounded, the results should be interpreted as exploratory and hypothesis-generating rather than as evidence of causal mechanisms. Longitudinal or experimental designs are needed to establish temporal ordering and causal direction.

Second, the sampling strategy relied on snowball recruitment facilitated by undergraduate students, which may limit representativeness. While this approach enabled access to geographically dispersed, long-distance intergenerational dyads, a key focus of the study, it may also have introduced selection bias toward families with stronger intergenerational ties or greater digital access. In addition, the sample was predominantly composed of Asian older adults. This ethnic composition likely reflects both institutional recruitment context and cultural norms emphasizing familial obligation and interdependence. Although this provides valuable insight into culturally embedded patterns of intergenerational digital engagement and contributes to theory-building in collectivist family contexts, it constrains the generalizability of the findings to other ethnic or cultural groups. This cultural context likely shapes both the forms of support emphasized and the meanings ascribed to co-play and thus should be understood as an interpretive lens rather than a confounding artifact.

Third, only a small proportion of participants (n=71) reported engaging in video game co-play with younger family members. This limited subsample reduces statistical power and increases the risk of type 2 error, particularly in mediation analyses. As such, null findings, especially regarding emotional support, should be interpreted cautiously. At the same time, the low prevalence of co-play itself represents a meaningful empirical finding, highlighting potential barriers related to digital literacy, access, generational norms, or physical ability that warrant further investigation.

Fourth, measures of video game co-play were based on self-reported estimates of shared gaming time and involvement, which may be subject to recall bias or social desirability effects. This concern may be especially salient among older adults with varying levels of cognitive load or digital familiarity. Future research should consider incorporating objective or behavioral measures, such as gameplay logs, platform-generated usage data, or ecological momentary assessment, to enhance measurement precision.

Fifth, the long-distance nature of intergenerational dyads was based on participant reports and supplemented by IP address filtering to ensure respondents were not located at the same physical address. However, geographic distance was not independently verified, and perceived distance may differ from objective distance. Future studies may benefit from more precise geographic measures to examine how physical separation interacts with digital engagement and support dynamics. Perceived geographic distance may differ from objective distance, and this distinction may influence how intergenerational support is experienced and interpreted, as perceived separation may be more salient for relational dynamics than physical distance alone.

Finally, the study did not directly assess socioeconomic status, national context, or broader structural factors that may shape access to digital technologies and gaming practices. Economic resources, educational background, and cultural exposure to gaming may act as unmeasured confounding variables influencing both co-play behavior and perceived social support. Future research should incorporate more diverse, cross-national samples and explicitly measure socioeconomic and structural determinants of digital engagement in later life.

Conclusion

In conclusion, this study highlights the potential of video game co-play as a meaningful, home-based, digitally mediated relational practice through which older adults may sustain intergenerational relationships across geographic distance. Rather than positioning co-play as an established intervention, the findings suggest that older adults' online communication attitudes, particularly online self-disclosure, are associated with engagement in interactive digital activities, which, in turn, relate to perceptions of informational and instrumental social support. These patterns are consistent with UGT and extend its application to aging populations by illustrating how motivational dispositions translate into relational digital practices in later life. By situating video game co-play within culturally embedded family contexts and acknowledging structural and access-related constraints, this study underscores the importance of age-inclusive, culturally responsive digital engagement strategies for addressing social isolation among older adults.

Acknowledgments

Generative artificial intelligence (ChatGPT, OpenAI) was used to assist with language refinement and editorial clarity during manuscript preparation. No AI tools were used for data analysis, the interpretation of results, or the writing of original scientific content.

Funding

This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

Data will be made available on request.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: JTT
Data curation: JTT
Editing: JTT
Formal analysis: JTT
Investigation: JTT
Methodology: JTT
Project administration: JTT
Review: JTT
Writing original draft: JTT

Conflicts of Interest

None declared.

Checklist 1
CHERRIES checklist.
[DOCX File, 20 KB - aging_v9i1e80541_app1.docx]

References

1. Cotten SR, Anderson WA, McCullough BM. Impact of internet use on loneliness and contact with others among older adults: cross-sectional analysis. J Med Internet Res 2013 Feb 28;15(2):e39. [doi: 10.2196/jmir.2306] [Medline: 23448864]
2. Quan-Haase A, Wang H, Wellman B, Zhang R. Weaving family connections on and offline: the turn to networked individualism. In: Neves BB, Casimiro C, editors. Connecting Families? Information & Communication Technologies, Generations, and the Life Course: Policy Press; 2018:59-80. [doi: 10.51952/9781447339953.ch004]
3. Wang B, Taylor L, Sun Q. Families that play together stay together: investigating family bonding through video games. New Media Soc 2018 Nov;20(11):4074-4094. [doi: 10.1177/1461444818767667]
4. Petersen B, Khalili-Mahani N, Murphy C, et al. The association between information and communication technologies, loneliness and social connectedness: a scoping review. Front Psychol 2023;14:1063146. [doi: 10.3389/fpsyg.2023.1063146] [Medline: 37034933]
5. Balki E, Hayes N, Holland C. Effectiveness of technology interventions in addressing social isolation, connectedness, and loneliness in older adults: systematic umbrella review. JMIR Aging 2022 Oct 24;5(4):e40125. [doi: 10.2196/40125] [Medline: 36279155]
6. Ledbetter AM. Measuring online communication attitude: instrument development and validation. Commun Monogr 2009 Dec;76(4):463-486. [doi: 10.1080/03637750903300262]
7. Ledbetter AM, Mazer JP. Do online communication attitudes mitigate the association between Facebook use and relational interdependence? An extension of media multiplexity theory. New Media Soc 2014 Aug;16(5):806-822. [doi: 10.1177/1461444813495159]
8. Ledbetter AM, Mazer JP, DeGroot JM, Meyer KR, Swafford B. Attitudes toward online social connection and self-disclosure as predictors of Facebook communication and relational closeness. Communic Res 2011 Feb;38(1):27-53. [doi: 10.1177/0093650210365537]
9. De Schutter B, Vanden Abeele V. Towards a gerontoludic manifesto. Anthropol aging 2015;36(2):112-120. [doi: 10.5195/aa.2015.104]
10. Peng W, Crouse J. Playing in parallel: the effects of multiplayer modes in active video game on motivation and physical exertion. Cyberpsychol Behav Soc Netw 2013 Jun;16(6):423-427. [doi: 10.1089/cyber.2012.0384] [Medline: 23509986]
11. Voida A, Greenberg S. Console gaming across generations: exploring intergenerational interactions in collocated console gaming. Univ Access Inf Soc 2012 Mar;11(1):45-56. [doi: 10.1007/s10209-011-0232-1]
12. Zhang F, Kaufman D. Older adults' social interactions in Massively Multiplayer Online Role-Playing Games (MMORPGs). Games Cult 2016 Jan;11(1-2):150-169. [doi: 10.1177/1555412015601757]
13. Zahn C, Leisner D, Niederhauser M, Roos AL, Iseli T, Soldati M. Effects of game mode in multiplayer video games on intergenerational social interaction: randomized field study. JMIR Form Res 2022 Feb 16;6(2):e29179. [doi: 10.2196/29179] [Medline: 35171097]
14. De la Hera T, Loos E, Simons M, Blom J. Benefits and factors influencing the design of intergenerational digital games: a systematic literature review. Societies (Basel) 2017;7(3):18. [doi: 10.3390/soc7030018]
15. Katz E, Blumler JG, Gurevitch M. Uses and gratifications research. Public Opin Q 1973;37(4):509-523. [doi: 10.1086/268109]
16. Sundar SS, Limperos AM. Uses and Grats 2.0: new gratifications for new media. J Broadcast Electron Media 2013 Oct;57(4):504-525. [doi: 10.1080/08838151.2013.845827]

XSL•FO
RenderX

17. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018 May 8;58(3):467-477. [doi: 10.1093/geront/gnw249] [Medline: 28201730]
18. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985 Sep;98(2):310-357. [doi: 10.1037/0033-2909.98.2.310] [Medline: 3901065]
19. Uchino BN. Social Support and Physical Health: Understanding the Health Consequences of Relationships: Yale University Press; 2004. URL: https://www.amazon.in/Social-Support-Physical-Health-Understanding/dp/0300102186 [accessed 2026-03-17]
20. Li C, Hu J, Hengeveld B, Hummels C. Supporting intergenerational memento storytelling for older adults through a tangible display: a case study. Pers Ubiquit Comput 2022 Jun;26(3):625-649. [doi: 10.1007/s00779-020-01364-9]
21. Zhang F, Schell R, Kaufman D, Salgado G. Social interaction between older adults (80+) and younger people during intergenerational digital gameplay. Presented at: Human Aspects of IT for the Aged Population Applications, Services and Contexts; Jul 9-14, 2017. [doi: 10.1007/978-3-319-58536-9_25]
22. Rubin AM. Uses-and-gratifications perspective on media effects. In: Bryant J, Oliver MB, editors. Media Effects: Routledge; 2009:181-200. [doi: 10.4324/9780203877111-14]
23. Ruggiero TE. Uses and gratifications theory in the 21st century. Mass Commun Soc 2000 Feb;3(1):3-37. [doi: 10.1207/S15327825MCS0301_02]
24. Allaire JC, McLaughlin AC, Trujillo A, Whitlock LA, LaPorte L, Gandy M. Successful aging through digital games: socioemotional differences between older adult gamers and non-gamers. Comput Human Behav 2013 Jul;29(4):1302-1306. [doi: 10.1016/j.chb.2013.01.014]
25. Châtel BDL, Janssen JHM, Peeters G, et al. Social gaming to decrease loneliness in older adults: recruitment challenges and attrition analysis in a digital mixed methods feasibility study. JMIR Serious Games 2024 Oct 16;12:e52640. [doi: 10.2196/52640] [Medline: 39412863]
26. Wen J, Kow YM, Chen Y. Online games and family ties: influences of social networking game on family relationship. Presented at: Human-Computer Interaction -- INTERACT 2011; Sep 5-9, 2011. [doi: 10.1007/978-3-642-23765-2_18]
27. Barr M, Copeland-Stewart A. Playing video games during the COVID-19 pandemic and effects on players' well-being. Games Cult 2022 Jan;17(1):122-139. [doi: 10.1177/15554120211017036]
28. Triandis HC. Individualism and Collectivism, 1st edition: Routledge; 1995. [doi: 10.4324/9780429499845]
29. Silverstein M, Gans D, Yang FM. Intergenerational support to aging parents: the role of norms and needs. J Fam Issues 2006;27(8):1068-1084. [doi: 10.1177/0192513X06288120]
30. Yeh KH, Bedford O. A test of the Dual Filial Piety model. Asian J of Social Psycho 2003 Dec;6(3):215-228. [doi: 10.1046/j.1467-839X.2003.00122.x]
31. Kim HS, Sherman DK, Taylor SE. Culture and social support. Am Psychol 2008 Sep;63(6):518-526. [doi: 10.1037/0003-066X] [Medline: 18793039]
32. Taylor SE, Sherman DK, Kim HS, Jarcho J, Takagi K, Dunagan MS. Culture and social support: who seeks it and why? J Pers Soc Psychol 2004 Sep;87(3):354-362. [doi: 10.1037/0022-3514.87.3.354] [Medline: 15382985]
33. Fulk J, Schmitz J, Steinfield CW. A social influence model of technology use. In: Fulk J, Steinfield C, editors. Organizations and Communication Technology: SAGE Publications, Inc; 1990:117-140 URL: https://sk.sagepub.com/book/edvol/organizations-and-communication-technology/chpt/social-influence-model-technology-use [accessed 2026-03-17] [doi: 10.4135/9781483325385.n6]
34. Schmitz J, Fulk J. Organizational colleagues, media richness, and electronic mail: a test of the social influence model of technology use. Communic Res 1991;18(4):487-523. [doi: 10.1177/009365091018004003]
35. Rokeach M. A theory of organization and change within value-attitude systems. J Soc Issues 1968;24(1):13. [doi: 10.1111/j.1540-4560.1968.tb01466.x]
36. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research: Addison-Wesley; 1975. URL: https://www.amazon.in/Belief-Attitude-Intention-Behaviour-Introduction/dp/0201020890 [accessed 2026-03-17]
37. Chen YRR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18. [doi: 10.2196/jmir.4596] [Medline: 26822073]
38. Barbosa Neves B, Franz R, Judges R, Beermann C, Baecker R. Can digital technology enhance social connectedness among older adults? A feasibility study. J Appl Gerontol 2019 Jan;38(1):49-72. [doi: 10.1177/0733464817741369] [Medline: 29166818]
39. Mitzner TL, Boron JB, Fausset CB, et al. Older adults talk technology: technology usage and attitudes. Comput Human Behav 2010 Nov 1;26(6):1710-1721. [doi: 10.1016/j.chb.2010.06.020] [Medline: 20967133]
40. Desjarlais M. The socially poor get richer, the rich get poorer: the effect of online self-disclosure on social connectedness and well-being is conditional on social anxiety and audience size. Cyberpsychology (Brno) 2022;16(4). [doi: 10.5817/CP2022-4-4]
41. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]

XSL•FO
RenderX

42. Wright K. The communication of social support within an on-line community for older adults: a qualitative analysis of the SeniorNet community. Commun Q 1999;47(4):S33-S43 [FREE Full text]
43. Cutrona CE, Suhr JA. Controllability of stressful events and satisfaction with spouse support behaviors. Communic Res 1992 Apr;19(2):154-174. [doi: 10.1177/009365092019002002]
44. Yeung GTY, Fung HH. Social support and life satisfaction among Hong Kong Chinese older adults: family first? Eur J Ageing 2007 Dec;4(4):219-227. [doi: 10.1007/s10433-007-0065-1] [Medline: 28794791]
45. Ashida S, Heaney CA. Differential associations of social support and social connectedness with structural features of social networks and the health status of older adults. J Aging Health 2008 Oct;20(7):872-893. [doi: 10.1177/0898264308324626] [Medline: 18815414]
46. Roberto KA, Jarrott SE. Family caregivers of older adults: a life span perspective. Fam Relat 2008 Jan;57(1):100-111. [doi: 10.1111/j.1741-3729.2007.00486.x]
47. Wang HC, Li CR, Lo C, et al. Effect of social support on changes in instrumental activities of daily living in older adults: a national population-based longitudinal study. Int J Gerontol 2018 Jul. [doi: 10.1016/j.ijge.2018.06.004]
48. Cornejo R, Tentori M, Favela J. Enriching in-person encounters through social media: a study on family connectedness for the elderly. Int J Hum Comput Stud 2013 Sep;71(9):889-899. [doi: 10.1016/j.ijhcs.2013.04.001]
49. Kacane I, Hernández-Serrano MJ. Social connection when physically isolated: family experiences in using video calls. Open Cult Stud 2023 Jan 14;7(1):20220165. [doi: 10.1515/culture-2022-0165]
50. Duggan M, Rainie L. Cell phone activities 2012: photo taking, texting, and accessing the internet are the most popular activities people pursue with their mobile phones. : Pew Research Center's Internet & American Life Project; 2012 URL: https://mediaimpact.issuelab.org/resources/15313/15313.pdf [accessed 2026-03-17]
51. Casanova G, Zaccaria D, Rolandi E, Guaita A. The effect of information and communication technology and social networking site use on older people's well-being in relation to loneliness: review of experimental studies. J Med Internet Res 2021 Mar 1;23(3):e23588. [doi: 10.2196/23588] [Medline: 33439127]
52. Kong H, Liu H. The relationship between ICT use and perceived life satisfaction among older people in Korea: the mediating effect of social capital. Sustainability 2023;15(12):9353. [doi: 10.3390/su15129353]
53. Dunkel-Schetter C, Feinstein L, Call J. UCLA social support inventory (UCLA-SSI). : University of California, Los Angeles; 1986 URL: https://cds.psych.ucla.edu/wp-content/uploads/sites/48/2021/03/UCLASocialSupportInventory.pdf [accessed 2026-03-17]
54. McAndrew FT, Jeong HS. Who does what on Facebook? Age, sex, and relationship status as predictors of Facebook use. Comput Human Behav 2012 Nov;28(6):2359-2365. [doi: 10.1016/j.chb.2012.07.007]
55. Jackson LA, Wang JL. Cultural differences in social networking site use: a comparative study of China and the United States. Comput Human Behav 2013 May;29(3):910-921. [doi: 10.1016/j.chb.2012.11.024]
56. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis, First Edition: A Regression-Based Approach: Guilford Press; 2013.
57. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/internet use. J Med Internet Res 2013 May 2;15(5):e93. [doi: 10.2196/jmir.2645] [Medline: 23639979]
58. Czaja SJ, Charness N, Fisk AD, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352. [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
59. Vroman KG, Arthanat S, Lysack C. "Who over 65 is online?" Older adults' dispositions toward information communication technology. Comput Human Behav 2015 Feb;43:156-166. [doi: 10.1016/j.chb.2014.10.018]
60. Barbosa Neves B, Fonseca JRS, Amaro F, Pasqualotti A. Social capital and internet use in an age-comparative perspective with a focus on later life. PLOS One 2018;13(2):e0192119. [doi: 10.1371/journal.pone.0192119] [Medline: 29481556]
61. Hargittai E, Dobransky K. Old dogs, new clicks: digital inequality in skills and uses among older adults. Can J Commun 2017 May 23;42(2):195-212. [doi: 10.22230/cjc.2017v42n2a3176]
62. Age-friendly environments. World Health Organization. 2025. URL: https://www.who.int/teams/social-determinants-of-health/demographic-change-and-healthy-ageing/age-friendly-environments [accessed 2025-07-08]
63. Sixsmith A, Horst BR, Simeonov D, Mihailidis A. Older people's use of digital technology during the COVID-19 pandemic. Bull Sci Technol Soc 2022 Jun;42(1-2):19-24. [doi: 10.1177/02704676221094731] [Medline: 38603230]
64. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
65. Lim AJ, Lau CYH, Cheng CY. Applying the Dual Filial Piety Model in the United States: a comparison of filial piety between Asian Americans and Caucasian Americans. Front Psychol 2021;12:786609. [doi: 10.3389/fpsyg.2021.786609] [Medline: 35185688]
66. Healthy aging. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/healthy-aging/index.html [accessed 2025-07-08]
67. Guo JW, Chow MC, Amici J, Yu PC, Kuo CT. Public opinions about an older adult playing a persuasive game: based on social media messages. Aging Health Res 2023 Jun;3(2):100145. [doi: 10.1016/j.ahr.2023.100145]

XSL•FO
RenderX

68. Wang YH. Understanding senior adults' needs, preferences, and experiences of commercial exergames for health: usability study. JMIR Serious Games 2024 Apr 5;12:e36154. [doi: 10.2196/36154] [Medline: 38578674]
69. Markus HR, Kitayama S. Culture and the self: implications for cognition, emotion, and motivation. Psychol Rev 1991;98(2):224-253 [FREE Full text]
70. Li WW, Singh S, Keerthigha C. A cross-cultural study of filial piety and palliative care knowledge: moderating effect of culture and universality of filial piety. Front Psychol 2021;12:787724. [doi: 10.3389/fpsyg.2021.787724] [Medline: 34925189]
71. Tseng JT, Hsiao H, Yu APL, Chen Y. Impact of a nonfamilial intergenerational program with a mobile application on college students' attitudes toward older adults in Taiwan. J Soc Social Work Res 2023 Jun 1;14(2):365-384. [doi: 10.1086/715817]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
ICT: information and communication technology
OSC: online social connection
OSD: online self-disclosure
UGT: uses and gratifications theory

Edited by H Köttl; submitted 12.Jul.2025; peer-reviewed by M Tusconi, S Xiong; revised version received 16.Feb.2026; accepted 04.Mar.2026; published 07.Apr.2026.

Please cite as:
Tseng JT
Online Communication Attitudes and Video Game Co-Play in Older Adults: Cross-Sectional Mediation Study
JMIR Aging 2026;9:e80541
URL: https://aging.jmir.org/2026/1/e80541
doi:10.2196/80541

© Jeffrey Tsifan Tseng. Originally published in JMIR Aging (https://aging.jmir.org), 7.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

CareMobi to Improve Communication Between Caregivers and Adult Day Centers for People With Dementia: Mixed Methods Feasibility Study

Tina Sadarangani[1], RN, PhD; Laura Peralta[1], BSN, MS, RN; Shahrzad Siamdoust[1], MPH; Jackie Finik[1], MPH

NYU Rory Meyers College of Nursing, New York, NY, United States

Corresponding Author:
Tina Sadarangani, RN, PhD
NYU Rory Meyers College of Nursing
433 First Avenue
New York, NY, 10010
United States
Phone: 1 2129927183
Email: tina.sadarangani@nyu.edu

Abstract

Background: Adult day centers (ADC) are well-positioned to address social isolation among the rapidly growing population of people living with dementia but are underused relative to other forms of long-term care. Mistrust of these centers among family caregivers remains a barrier to their use. Digital health tools offer a promising approach to enhance transparency, improve communication, and build trust between caregivers and ADCs. As such, researchers at New York University (NYU) developed CareMobi, a user-centered mobile app that supports care coordination between ADC, care providers, and caregivers.

Objective: This study aimed to evaluate the real-world usability, acceptability, and feasibility of CareMobi among staff in ADCs and the caregivers of people living with dementia who attend centers.

Methods: Guided by the Goal–Question–Metric (GQM) framework, we conducted a low-burden field usability evaluation of CareMobi in 2 ADCs. Data included baseline caregiver surveys, app usage logs, administration of the **Post-Study System Usability Questionnaire**, and participant interviews. Feasibility benchmarks were established a priori using STOP-AMEND-GO criteria for evaluable retention, usability, satisfaction, engagement, and acceptability. Quantitative analyses summarized demographics, caregiver confidence, and app usability (7-point scale and median scores). Qualitative analyses used thematic coding of interviews and written feedback to identify perceived benefits, barriers, and recommendations for implementation.

Results: Family caregivers (n=15) from 2 ADCs participated in a mixed-methods pilot. Participants were primarily female (87%), non-Hispanic White (93%), and middle-aged or older (67% aged 30-64 years). Most were sole caregivers (80%). Baseline confidence in managing dementia-related care was high, particularly in communication with providers. Of 15 enrolled caregivers, 14 (93%) logged into CareMobi at least once, and the majority “agreed/strongly agreed” that the app was easy to use (6/8, 75%) and easy to learn (7/8, 87.5%); 75% (6/8) of the participants liked the interface and found the information well organized. All staff (n=3) indicated the CareMobi app was easy to use, easy to learn, and well organized (3/3, 100%). Willingness to use the app beyond the study period was high among caregivers (6/8, 75%) and staff (3/3, 100%). Open-ended caregiver responses emphasized reassurance from daily updates, user-friendliness, and time savings. Staff reported reduced “phone tag” streamlining coordination, particularly for new ADC enrollees. However, established caregivers tended to use the app passively, defaulting to phone calls. Themes highlighted the need for structured logging guidance, role-specific templates, and integration with existing systems. Caregivers and staff noted that routine updates fostered trust, improved decision-making about attendance, strengthened communication with health care providers, and provided peace of mind.

Conclusions: CareMobi offers a promising approach to modernizing ADC in ways that respect resource limitations and caregiver needs, through combining digital transparency with low-burden communication. Future efforts should center on equity and focus on linguistic inclusion, infrastructure support, and engagement in underserved communities.

(JMIR Aging 2026;9:e85090) doi:10.2196/85090

KEYWORDS

adult day services; dementia; mobile health; community-based; feasibility; adult day care

XSL•FO
RenderX

Introduction

Social isolation and loneliness are widely recognized, modifiable social determinants of health in later life. In the United States, approximately one quarter of community-dwelling adults aged ≥65 years are socially isolated and large proportions report loneliness [1]. Social isolation and loneliness are associated with a 32% and 14% increased risk of all-cause mortality, respectively. They are also linked to increased morbidity, including a 29% higher risk of coronary heart disease and a 32% higher risk of stroke [1,2]. These risks are amplified for people living with dementia and their family care partners, for whom shrinking networks, mobility limitations, and caregiver burden compound barriers to engagement in community life [3].

Adult day centers (ADCs) are a cornerstone of community-based long-term care and a promising, equity-oriented platform to counter social disconnection at scale [4], especially for people living with dementia. They represent the most diverse sector of long-term care, serving more than 250,000 predominantly low-income (72%) and ethnically diverse (58%) individuals, who are able to remain in their own homes while receiving daytime care that fosters engagement, health monitoring, and meaningful socialization. By providing structured programming, culturally responsive services, and respite for care partners, ADCs buffer social isolation and loneliness for people living with dementia, and simultaneously reduce caregiver burden [4,5]. Yet, despite the fact that more than 80% of older adults report wanting to remain in their homes [5], and that social engagement is consistently linked to better health outcomes [6], ADCs remain vastly underused compared to other, more costly forms of long-term care that either don't offer socialization or allow people to remain at home. These include home health (serving more than 3 million) or nursing homes (with 1.2 million residents) [7].

There are several barriers that limit uptake of ADCs, including transportation challenges, cost, and limited awareness of services; but one of the most significant is trust. A key relational barrier to ADC use, particularly among socially isolated or underserved families, is mistrust of these facilities and difficulty communicating with center staff about day-to-day care [7]. Family caregivers often report uncertainty about whether staff truly understand their loved one's needs, whether cultural preferences will be respected, and whether information will be shared transparently, all of which cause apprehension and can deter sustained engagement.

Research has consistently highlighted improved information-sharing as crucial to overcoming barriers in the uptake of respite care [8]. While many older adults can articulate their own needs, for those who are unable to do so, like people living with dementia, effective communication often depends on carers and respite staff [8]. Digital health tools, when designed for usability and embedded into real-world workflows, offer a promising strategy to enhance transparency, strengthen bidirectional communication, and improve coordination between families and community providers [9,10]. By providing a lens into the day-to-day care of their loved ones, such tools can reduce caregiver apprehension, build trust in ADCs, and allow people living with dementia to be more socially engaged by attending these centers.

A growing number of mobile health (mHealth) interventions have been developed to support dementia caregiving. such as symptom-monitoring tools, secure messaging platforms, and caregiver education apps, but few are designed for or evaluated within ADC environments, which have distinct documentation practices, staffing structures, and relational communication needs. Existing tools often focus on clinical care coordination or remote monitoring rather than the day-to-day visibility that shapes caregiver trust in ADCs and influences continued attendance. CareMobi is a mHealth app designed to strengthen communication, coordination, and transparency between ADC staff, family caregivers, and health providers. The app allows for real-time updates on daily activities, health status, and logistical information, helping caregivers interpret changes in behavior at home and enhancing continuity of care across settings. CareMobi differs conceptually from prior interventions by centering a digital trust-building approach: it uses real-time, low-burden updates to reduce caregivers' uncertainty, strengthen relational continuity, and support socio-technical alignment between caregivers and ADC workflows. This focus on transparency, role-specific communication, and workflow feasibility represents a novel contribution to the mHealth landscape and addresses known barriers to ADC use [10-12].

We posit that by making ADC-based care more visible to families, through timely updates and low-burden, role-specific interactions, CareMobi can reduce uncertainty, foster trust, and simplify decisions about ADC attendance and sustained participation. Initial interactive prototype testing allowed us to make user-informed modifications that laid the foundation for real-world evaluation [10]. The purpose of this feasibility study was to: (1) evaluate acceptability (perceived usefulness, satisfaction, and burden for caregivers and staff), (2) assess demand (enrollment, uptake, and engagement with core features), and (3) examine implementation and practicality within ADC workflows (resources, role fit, and compatibility with existing communication practices). Our evaluation also sought to understand caregivers' baseline confidence levels and areas where ADCs could provide additional support, as well as the platform's potential impacts on caregivers' overall experience and confidence in the care provided by ADCs.

Methods

Overview and Rationale

CareMobi was developed using a user-centered design process that actively engaged ADS staff, family caregivers, and community stakeholders throughout its development. Iterative feedback was solicited through interviews, focus groups, and usability testing to ensure the tool was feasible to integrate into workflows, responsive to caregiver needs, and adaptable across diverse ADCs. This participatory process emphasized co-design to identify priority features (eg, streamlined messaging, daily activity summaries, and health-related observations) and reduce barriers such as duplicate documentation or technology burden. The full details of CareMobi's design and development process,

XSL•FO
RenderX

including prototype testing, have been reported elsewhere in prior research; here we focus on evaluation of its implementation, usability, and acceptability in practice [10].

Study Framework

We adopted a pragmatic, low-burden field usability evaluation of CareMobi in ADCs, guided by a Goal–Question–Metric (GQM/GAQM) [13] approach. Originally developed for software engineering and adapted for usability research, GQM ensures that data collection remains purposeful, relevant, and aligned with stakeholder needs. In practice, evaluators first articulate the overarching goals (eg, feasibility, usability, and acceptability), then derive key questions that operationalize those goals (eg, "Can caregivers complete core tasks without difficulty?"), and finally specify the metrics that can provide evidence (eg, task completion rates, satisfaction ratings, and interview themes). This hierarchical structure promotes rigor while keeping evaluations pragmatic and low burden, particularly in resource-constrained real-world settings such as ADCs. This allowed us to explicitly tie high-level goals (usability, acceptability, and feasibility) to specific questions and measurable metrics, while respecting constraints of under-resourced ADC settings. We aligned these questions and metrics with the GAQM framework's principle of first defining a goal (conceptual), then attributes (usability dimensions), then operational questions, and finally metrics (quantitative and qualitative). This approach is outlined in Table 1.

Table 1. Goal–questions–metrics framework.

Goal (Usability and implementation focus)	Usability or implementation attribute	Key questions	Metrics and indicators
Assess satisfaction and perceived usability	Perceived usability and satisfaction	How do users rate the app on ease, clarity, and burden?	Likert-scale items (eg, PSSUQ[a])
Understand barriers, user experience, and trust	Qualitative and contextual implementation	What aspects of the app or workflow did users find confusing, burdensome, or trust-enhancing? Where do they feel uncertain?	Thematic codes from semistructured interviews; free-text questionnaire comments
Engagement and uptake in real settings	Implementation and adoption	How many caregivers and staff enroll, how often do they use core features?	Enrollment counts, caregivers who use it at least once

[a]PSSUQ: Post-Study System Usability Questionnaire.

Overall Study Design

The evaluation combined field (in situ) testing, self-user usability (unmoderated use in real settings), post-use questionnaire, and qualitative interviews with staff and family caregivers in ADCs located in rural and underserved communities.

Setting

We partnered with 2 ADC sites primarily serving people living with dementia that were recruited at the Annual Meeting of the National Adult Day Services Association. Within each center, paid staff and family caregivers were recruited to pilot test CareMobi in their real daily workflows (as opposed to a lab setting), over a 12-week period. Our approach was consistent with "field studies" in mobile usability literature, where tasks are executed in the natural environment, and contextual factors (interruptions, real devices, ambient conditions) are preserved [10].

Recruitment and Eligibility

Recruitment occurred over 3 months through center leadership, who distributed study information to eligible caregivers and staff; interested individuals were then contacted by the research team to confirm eligibility and obtain informed consent. Eligible ADC staff included those directly involved in the day-to-day care of people living with dementia and responsible for communicating with family caregivers. Eligible family caregivers were those who identified as the primary caregiver of people living with dementia attending one of the participating ADCs, were aged 18 years or older, and owned a smartphone or tablet capable of running the CareMobi app. Both staff and caregivers were required to speak English and be willing to provide informed consent.

Training and Study Procedures

Prior to deployment, participants received minimal training, an intentional design choice to evaluate how intuitive the CareMobi platform was in real-world use. ADC staff participated in a 30-minute orientation covering the study overview and app functions. Family caregivers were assisted, as needed, in downloading the app using a QR code and were then encouraged to explore it independently. For ongoing support with the app, staff and caregivers were directed to a video library developed by an instructional technologist. Staff and caregivers were also advised to reach out to the study team with any questions or regarding technical issues. This pragmatic, low-burden approach mirrored real-world implementation conditions and aligned with the GQM framework, specifically addressing goals of assessing intuitiveness and efficiency of use with minimal external support.

Staff were instructed to use their discretion and share information through CareMobi that was tailored to each person living with dementia and caregivers' needs, while caregivers were encouraged to use the app as much or as little as they preferred to engage with the ADC. There was no minimum requirement for use in order to reflect real-world engagement patterns. CareMobi uses HIPAA (Health Insurance Portability and Accountability Act)–compliant security protocols, including encrypted data transmission and storage, authenticated user access, and restricted role-based permissions. No identifiable information was shared outside the secure platform, and all study data were deidentified prior to analysis." Following

XSL•FO
RenderX

consent, participants were instructed to download the CareMobi app from their preferred app store, and family caregivers completed a baseline survey. After the 12-week usage period, participants completed a post-use questionnaire assessing perceived usability, satisfaction, burden, and open-ended feedback. We also conducted semistructured interviews with a purposive subset of caregivers and staff, focusing on their experience, trust concerns, communication challenges, and suggestions for improvement. Staggered incentives, in the form of gift cards, were provided upon completion of the baseline survey, usability questionnaire, and interview (US $150 total).

Measures and Instruments

Demographics

We collected self-reported baseline demographic characteristics from family caregivers, including age, gender, role, and caregiving relationship, to contextualize study findings.

Caregiver Confidence

Two validated caregiver-reported instruments were used to assess caregiver experience and confidence. The Caregiver Confidence in Sign/Symptom Management (CCSM) Scale is a 4-item instrument that evaluates caregivers' perceived ability to recognize, monitor, and respond to changes in their care recipient's symptoms [14]. Items are rated on a 5-point Likert scale ranging from not at all confident (1) to extremely confident (5), with higher scores indicating greater confidence in symptom management. Prior research has demonstrated good internal consistency (Cronbach α=0.85) and construct validity, supporting its use in community-based caregiver populations [14].

Usability

To capture perceptions of usability and satisfaction, we administered the Post-Study System Usability Questionnaire (PSSUQ) [15,16], a validated 16-item scale that measures user perceptions across three domains: system usefulness, information quality, and interface quality. Items are rated on a 7-point Likert scale (strongly disagree [1] to strongly agree [7]), with lower scores indicating greater usability. The PSSUQ has been widely applied in digital health research and demonstrates high internal consistency (Cronbach α=0.83-0.96) [15,16].

Satisfaction

Finally, a **semistructured interview guide** was used to explore participants' experiences in greater depth. Interview domains included perceptions of transparency and trust, communication gaps, perceived benefits or frustrations, barriers to engagement, and suggestions for improving design or workflow alignment. Open-ended questions (eg, "Were there any parts of the app that felt confusing, frustrating, or difficult to use?"; "How, if at all, did it impact your experience with the ADC?") elicited participants' reactions to CareMobi, perceptions of usability, potential barriers, and impact on their relationship with the ADC as it pertained to the health of the people living with dementia.

Data Analysis

Quantitative

The pilot was conducted as a feasibility and usability study rather than an efficacy trial, and analyses were therefore descriptive and hypothesis-generating. Quantitative analyses included calculation of descriptive statistics (frequencies, medians, and ranges) for demographic and questionnaire results.

Qualitative

Participant interviews lasted approximately 30 minutes and were conducted by either the principal investigator (TS), who has formal training in mixed-methods research, or a research assistant (LP) with a master's in clinical research. Interviewers used member checking informally during data collection by pausing to confirm interpretations. All interviews were recorded, transcribed, and reviewed for accuracy. Field notes and reflexive journals were maintained to document interviewer observations, decision-making, and potential sources of bias.

Coding focused on domains relevant to the study aims, including trust and transparency, communication barriers, workflow fit, and caregiver confidence. Themes were then triangulated with quantitative findings (eg, task completion and usability scores) to provide a more comprehensive understanding of user experience. Two coders (one PhD, one undergraduate nursing student) independently coded transcripts in Dedoose, with the principal investigator reviewing a 20% subset to ensure reliability. The research team met regularly to discuss coding discrepancies, update the codebook, and validate emerging findings. Saturation was considered reached when no new categories were identified.

Feasibility Benchmarks

To evaluate feasibility, we applied a priori STOP-AMEND-GO [17] criteria across five domains: retention, usability, satisfaction, engagement, and acceptability. **Evaluable retention** was defined as completion of the 3-month follow-up survey, excluding non–study-related attrition (death or transition to a higher level of care); thresholds were STOP if <60%, AMEND if 60-69%, and GO if ≥70%. **Usability** was assessed using the PSSQ, with criteria based on the proportion of participants endorsing "agree" or "strongly agree" for items on ease of use and ease of learning; thresholds were STOP if <60% on any item, AMEND if 60-74%, and GO if >75% on each. **Satisfaction** was measured as the percentage of caregivers and staff endorsing "agree/strongly agree" on overall satisfaction with CareMobi, with thresholds STOP if <60%, AMEND if 60-74%, and GO if >75%. **Engagement** was defined as the proportion of participants who created an account and used the app at least once; thresholds were STOP if <50%, AMEND if 50-69%, and GO if ≥70%. Finally, **acceptability** was measured as the percentage of staff and caregivers expressing willingness to continue using CareMobi beyond the study period, with thresholds STOP if <50%, AMEND if 60-74%, and GO if >75% (Table 2).

XSL•FO
RenderX

Table 2. Study feasibility benchmarks (STOP-AMEND-GO criteria) and results.

Domains	Metric	STOP	AMEND	GO	Study result		
					Caregiver	Staff	Criterion
Evaluable retention	3-month completion of survey excluding nonstudy-related attrition (ie, death)	<60%	60%-69%	≥70%	100%	100%	GO
Usability	% "agree/strongly agree" on easy to use, easy to learn on PSSQ	<60% on any	60%-74% on any	75% on each	75%	100%	GO
Satisfaction	% "agree/strongly agree" on overall satisfaction	<60%	60%-74%	75% on each	>75%	100%	GO
Engagement	% of caregivers who created an account and used the app at least once	<50% any-use	50%-69% any-use	≥70% any-use	93%	100%	GO
Acceptability	% of staff and caregivers expressing that they will continue to use CareMobi beyond study period	<50%	60%-74%	75%	75%	100%	GO

Ethical Considerations

All study procedures were reviewed and approved by the Institutional Review Board at New York University (IRB-FY2023-7509). All participants provided informed consent prior to participation, and all data were deidentified to protect confidentiality. Compensation was US $150 for participation in all phases of the study.

Results

Caregiver Characteristics

In total, 15 caregivers completed baseline surveys (Table 3). Most participants identified as female (13/15, 86.7%), were aged 30-64 years (10/15, 66.7%), and reported some college or higher education (14/15, 93.3%). The sample was predominantly non-Hispanic White (14/15, 93.3% White; 14/15, 93.3% non-Hispanic). Most were married (11/15, 73.3%). Overall, 40% (6/15) were employed full-time and 20% (3/15) part-time; 26.7% (4/15) were retired. All respondents were family caregivers (15/15, 100%); 80% (12/15) identified as the sole caregiver. Relationship to the people living with dementia was most commonly adult child (7/15, 46.7%) or spouse or partner (5/15, 33.3%). Most had been caregiving for ≥1 year (11/15, 73.3%).

XSL•FO
RenderX

Table 3. Demographic characteristics of the family caregivers and staff at the adult day centers.

Characteristics	Values, n (%)
Family caregivers (n=15)	
Sex	
Male	2 (13.3)
Female	13 (86.7)
Age (years)	
30-49	4 (26.7)
50-64	6 (40)
65+	4 (26.7)
Unknown	1 (6.7)
Education	
High school graduate or less	1 (6.7)
Some college	6 (40)
College graduate or higher	8 (53.3)
Race (ethnicity)	
Non-Hispanic White	14 (93.3)
Other Non-Hispanic	1 (6.7)
Marital status	
Single	2 (13.3)
Married	11 (73.3)
Widowed	1 (6.7)
Divorced	1 (6.7)
Employment status	
Employed full-time	6 (40)
Employed part-time	3 (20)
Retired	4 (26.7)
Other	2 (13.3)
Annual household income (US $)	
<$25,000	2 (13.3)
$25,000–$49,999	4 (26.7)
$50,000–$99,999	6 (40)
≥$100,000	1 (6.7)
Prefer not to answer	2 (13.3)
Relationship to person with dementia	
Spouse or partner	5 (33.3)
Adult child	7 (46.7)
Grandchild	2 (13.3)
Other family member	1 (6.7)
Caregiving characteristics	
Sole caregiver	12 (80)
Duration of care	
6-12 months	2 (13.3)
≥1 year	11 (73.3)

XSL•FO
RenderX

Characteristics	Values, n (%)
Prefer not to answer	2 (13.3)
Staff demographics (n=3)	
Race (Ethnicity)	
Non-Hispanic White	2 (66.7)
Non-Hispanic Black or African American	1 (33.3)
Sex	
Female	3 (100)

Caregiver Confidence With Complex Care

Across 15 respondents, confidence was generally moderate to high for managing common dementia-related challenges (Table 4). For agitation, 10/15 (66.7%) reported being extremely or moderately confident; for mood changes, 10/15 (66.7%); for mental status changes, 12/15 (80%). Confidence in deciding whether to call a provider was high (13/15, 86.7% extremely or moderately). Communication-related competencies were strongest: 14/15 (93.3%) felt extremely or moderately confident talking to a medical provider, and 12/15 (80%) reported they could readily gather information a provider would want. Areas with more dispersion included understanding vital signs (10/15, 66.7% extremely or moderately; 3/15, 20% slightly) and assessing pain (9/15, 60% extremely or moderately).

Table 4. Caregiver confidence: frequency counts and median scores (n=15). Scores are based on a 5-point scale: 5=Extremely, 4=Moderately, 3=Somewhat, 2=Slightly, 1=Not at All. Median scores are reported alongside frequency counts.

Item	Extremely (5), n (%)	Moderately (4), n (%)	Somewhat (3), n (%)	Slightly (2), n (%)	Not at all (1),n (%)	Un-known	Median (IQR; 25th - 75th percentile)
Managing symptoms and behaviors							
Dementia-related behaviors (ie, agitation)	2 (13.3)	8 (53.3)	3 (20)	1 (6.6)	1 (6.6)	0 (0)	4 (3-4)
Mood-related changes (ie, depression)	1 (6.6)	9 (60)	4 (26.6)	0 (0)	1 (6.6)	0 (0)	4 (3-4)
Mental status changes (ie, confusion)	6 (40)	6 (40)	2 (13.3)	1 (6.6)	0 (0)	0 (0)	4 (4-5)
New medical problems or events	3 (20)	8 (53.3)	2 (13.3)	0 (0)	2 (13.3)	0 (0)	4 (3.5-4)
Ongoing chronic diseases	3 (20)	7 (46)	3 (20)	1 (6.6)	1 (6.6)	0 (0)	4 (3-4)
Decision-making ability							
Dementia-related behaviors	2 (13.3)	8 (53.3)	3 (20)	0	2 (13.3)	0	4 (3-4)
Mood changes	1 (6.6)	7 (46.6)	4 (26.6)	2 (13.3)	1 (6.6)	0 (0)	4 (3-4)
Mental status changes	2 (13.3)	8 (53.3)	2 (13.3)	3 (20)	0 (0)	0 (0)	4 (3-4)
New medical problems	3 (20)	5 (33.3)	5 (33.3)	2 (13.3)	0 (0)	0 (0)	4 (3-4)
Ongoing chronic diseases	3 (20)	3 (20)	5 (33.3)	2 (13.3)	1 (6.6)	1 (6.6)	3 (3-4)
Health care communication							
Deciding whether or not to call the provider	4 (26.6)	9 (60)	2 (13.3)	0 (0)	0 (0)	0 (0)	4 (4-4.5)
Gathering information that a medical provider would want	8 (53.3)	4 (26.6)	3 (20)	0 (0)	0 (0)	0 (0)	5 (4-5)
Talking to a medical provider	10 (66.6)	4 (26.6)	1 (6.6)	0 (0)	0 (0)	0 (0)	5 (4-5)
Clinical assessment skills							
Taking vital signs	8 (53.3)	1 (6.6)	1 (6.6)	3 (20)	2 (13.3)	0 (0)	5 (2-5)
Understanding vital signs	6 (40)	4 (26.6)	1 (6.6)	1 (6.6)	3 (20)	0 (0)	4 (2.5-5)
Assessing for pain	3 (20)	6 (40)	3 (20)	3 (20)	0 (0)	0 (0)	4 (3-4)
Taking action when someone is in pain	5 (33.3)	5 (33.3)	3 (20)	2 (13.3)	0 (0)	0 (0)	4 (3-5)
Assessing for dehydration	3 (20)	5 (33.3)	3 (20)	2 (13.3)	1 (6.6)	1 (6.6)	4 (3-4)
Managing dehydration	5 (33.3)	4 (26.6)	4 (26.6)	2 (13.3)	0 (0)	0 (0)	4 (3-5)
Managing medical problems at home	4 (26.6)	5 (33.3)	3 (20)	2 (13.3)	1 (6.6)	0 (0)	4 (3-4.5)

XSL•FO
RenderX

People Living With Dementia Clinical Stage and Health Care Use

Care recipients were primarily in moderate to later stages of dementia: 7/15 (46.7%) had moderately severe cognitive decline, 2/15 (13.3%) severe, and 2/15 (13.3%) very severe. Within the prior 6 months, 6/15 (40%) had at least one emergency department visit, and 4/15 (26.7%) had been hospitalized. Medication burden was substantial: 10/15 (66.7%) were taking ≥5 daily medications and 4/15 (26.7%) were taking 3-5 (Table 5).

Table 5. People living with dementia: disease stage and health care (n=15).

Characteristics	Values, n (%)
Stage of dementia	
Early-stage	4 (26.7)
Moderate-severe cognitive decline	7 (46.7)
Severe cognitive decline	2 (13.3)
Very severe cognitive decline	2 (13.3)
Health care use (past 12 months)	
Mean number of specialists seen	4 (26.7)
Health care use (Past 6 months)	
Emergency department visit	6 (40)
Hospitalization	4 (26.7)
Total daily medications	
1-2 medications	1 (6.7)
3-5 medications	4 (26.7)
>5 medications	10 (66.7)

Attrition and Engagement

Of the 15 caregivers who volunteered for the study, 14 logged into the app at least once, demonstrating high initial engagement. Four participants did not complete the study due to their loved one leaving the center, either due to death or requiring a higher level of care. Of the remaining caregivers, 8/10 completed the usability questionnaire and 6/10 participated in the follow-up interview. Among the 3 staff who participated, all (100%) logged into the app, completed the PSSQ, and completed the follow-up interview.

App Usability and Perceived Impact on Access and Communication

Among the 8 caregivers who completed the application usability questionnaire, perceived usability and satisfaction were high. At least three-quarters “agreed/strongly agreed” that the app was easy to use (6/8, 75%) and easy to learn (7/8, 87.5%); 6/8 (75%) liked the interface and found information well organized. Most felt comfortable using the app in social settings (6/8, 75%) and endorsed appropriate time demands (6/8, 75%). Willingness to use the app beyond the study period was high (6/8, 75%), and overall satisfaction was favorable (6/8, 75% “agree/strongly agree”).

Communication and access benefits were frequently endorsed: 7/8 (87.5%) reported that the app improved access to their loved one at the center, 7/8 (87.5%) said it made communication with the ADC convenient, and 6/8 (75%) felt comfortable communicating with the center via the app. Navigation consistency was also rated positively (7/8, 87.5%) (Table 6).

XSL•FO
RenderX

Table 6. Application usability among caregivers (n=8)[a].

Item, CareMobi app	Strongly Disagree (1), n (%)	Disagree (2), n (%)	Somewhat Disagree (3), n (%)	Neither Agree nor Disagree (4), n (%)	Somewhat Agree (5), n (%)	Agree (6), n (%)	Strongly Agree (7), n (%)	Median (IQR)
Easy to use	0 (0)	0 (0)	2 (25)	0 (0)	0 (0)	3 (37.5)	3 (37.5)	6 (5-7)
Easy to learn to use	0 (0)	0 (0)	1 (12.5)	0 (0)	2 (25)	2 (25)	3 (37.5)	6.5 (5.75-7)
Liked the interface	0 (0)	0 (0)	1 (12.5)	0 (0)	1 (12.5)	2 (25)	4 (50)	6.5 (5.75-7)
The information was well organized	0 (0)	0 (0)	1 (12.5)	0 (0)	1 (12.5)	2 (25)	4 (50)	6 (5.75-7)
Comfortable using in social settings	0 (0)	0 (0)	0 (0)	0 (0)	2 (25)	3 (37.5)	3 (37.5)	6 (5.75-6.25)
The amount of time involved in using it has been fitting for me	0 (0)	1 (12.5)	0 (0)	0 (0)	1 (12.5)	4 (50)	2 (25)	6 (5.75-6.25)
Would use it again	0 (0)	0 (0)	1 (12.5)	0 (0)	1 (12.5)	4 (50)	2 (25)	6 (5.75-7)
Overall, I am satisfied with it	0 (0)	0 (0)	0 (0)	1 (12.5)	1 (12.5)	3 (37.5)	3 (37.5)	4.5 (4-6)
Whenever I made a mistake using it, I could recover easily and quickly	0 (0)	0 (0)	1 (12.5)	3 (37.5)	1 (12.5)	2 (25)	1 (12.5)	6 (4.75-7)
It provides an acceptable way to send or receive health care communications	0 (0)	0 (0)	0 (0)	2 (25)	1 (12.5)	2 (25)	3 (37.5)	6 (5.75-7)
It adequately acknowledged and provided information to let me know the progress of my action	0 (0)	0 (0)	0 (0)	1 (12.5)	1 (12.5)	3 (37.5)	3 (37.5)	6 (6-7)
The navigation was consistent when moving between screens	0 (0)	0 (0)	0 (0)	1 (12.5)	0 (0)	4 (50)	3 (37.5)	6 (5.5-7)
The interface allowed me to use all the functions offered by the app (such as entering information, responding to notifications, viewing information)	0 (0)	0 (0)	0 (0)	2 (25)	0 (0)	3 (37.5)	3 (37.5)	6 (6-6)
It has all the functions and capabilities I expected it to have	0 (0)	0 (0)	0 (0)	1 (12.5)	0 (0)	6 (75)	1 (12.5)	6 (6-7)
It would be useful for my loved one's well-being	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	5 (62.5)	3 (37.5)	6.5 (5.75-7)
It improved my access to my loved one in the adult day center	0 (0)	0 (0)	0 (0)	1 (12.5)	1 (12.5)	2 (25)	4 (50)	6 (4.75-7)
It helped me manage my loved one's health effectively	0 (0)	0 (0)	0 (0)	2 (25)	1 (12.5)	2 (25)	3 (37.5)	6.5 (5.75-7)

XSL•FO
RenderX

Item, CareMobi app	Strongly Disagree (1), n (%)	Disagree (2), n (%)	Somewhat Disagree (3), n (%)	Neither Agree nor Disagree (4), n (%)	Somewhat Agree (5), n (%)	Agree (6), n (%)	Strongly Agree (7), n (%)	Median (IQR)
It made it convenient for me to communicate with my adult day center	0 (0)	1 (12.5)	0 (0)	0 (0)	1 (12.5)	2 (25)	4 (50)	6 (5.75-6.25)
I had more opportunities to interact with adult day center staff	0 (0)	0 (0)	1 (12.5)	0 (0)	1 (12.5)	4 (50)	2 (25)	6 (5.75-7)
I felt confident that any information I send to a caregiver would be received.	0 (0)	1 (12.5)	0 (0)	0 (0)	1 (12.5)	3 (37.5)	3 (37.5)	7 (5.75-7)
I felt comfortable communicating with my adult day center using it	0 (0)	1 (12.5)	0 (0)	0 (0)	1 (12.5)	1 (12.5)	5 (62.5)	__[b]

[a]Scores are based on a 7-point scale: 7=Strongly Agree, 6=Agree, 5=Somewhat Agree, 4=Neither Agree nor Disagree, 3=Somewhat Disagree, 2=Disagree, 1=Strongly Disagree. Median scores are reported alongside frequency counts.

[b]Not applicable.

Staff Usability of the CareMobi App

Staff reported highly favorable usability ratings, with median scores at or near the ceiling across nearly all items (median=7, “strongly agree”). All 3 participants indicated that the CareMobi app was easy to use, easy to learn, and well organized (median=7). Satisfaction was uniformly high (all median=7): staff strongly agreed they would use the app again, that it contained the functions and capabilities expected, and that it provided an acceptable way to communicate with caregivers.

Communication-related items also demonstrated the strongest endorsement. Staff unanimously reported that the app improved access to caregivers, made communication more convenient, and increased opportunities to interact with caregivers (median=7 across items). They also expressed confidence that any information sent would be reliably received (median=7).

Minor variability was observed on a small number of items. For example, one staff member reported difficulty recovering from mistakes, producing a median score of 6 (“agree”) on that item. Another staff member selected “neither agree nor disagree” on certain workflow-related items (eg, time fit and error recovery), though the median across sites still remained 7 (Table 7).

XSL•FO
RenderX

Table 7. Staff application usability—combined results (n=3).

Item, CareMobi app	Strongly Disagree (1), n (%)	Disagree (2), n (%)	Somewhat Disagree (3), n (%)	Neither Agree nor Disagree (4), n (%)	Somewhat Agree (5), n (%)	Agree (6), n (%)	Strongly Agree (7), n (%)	Median (IQR)
Easy to use	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (7-7)
Easy to learn to use	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (7-7)
Liked the interface	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (7-7)
The information was well organized	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (7-7)
Comfortable using in social settings	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (5.5-7)
The amount of time involved in using it has been fitting for me	0 (0)	0 (0)	0 (0)	1 (33.3)	0 (0)	0 (0)	2 (66.6)	7 (7-7)
Would use it again	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (7-7)
Overall, I am satisfied with it	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	4 (3-5.5)
Whenever I made a mistake using it, I could recover easily and quickly	0 (0)	1 (33.3)	0 (0)	1 (33.3)	0 (0)	0 (0)	1 (33.3)	7 (7-7)
It provides an acceptable way to send or receive health care communications	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (7-7)
It adequately acknowledged and provided information to let me know the progress of my action	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (6.5-7)
The navigation was consistent when moving between screens	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (33.3)	2 (66.6)	7 (5.5-7)
The interface allowed me to use all the functions offered by the app (such as entering information, responding to notifications, viewing information)	0 (0)	0 (0)	0 (0)	1 (33.3)	0 (0)	0 (0)	2 (66.6)	6 (5.5-6.5)
It has all the functions and capabilities I expected it to have	0 (0)	0 (0)	0 (0)	0 (0)	1 (33.3)	1 (33.3)	1 (33.3)	7 (7-7)
It would be useful for my loved one's well being	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (5.5-7)
It improved my access to my loved one in the adult day center	0 (0)	0 (0)	0 (0)	1 (33.3)	0 (0)	0 (0)	2 (66.6)	7 (7-7)
It helped me manage my loved one's health effectively	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (7-7)
It made it convenient for me to communicate with my adult day center	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)	7 (5.5-7)
I had more opportunities to interact with adult day center staff	0 (0)	0 (0)	0 (0)	1 (33.3)	0 (0)	0 (0)	2 (66.6)	7 (6-7)

XSL•FO
RenderX

Item, CareMobi app	Strongly Disagree (1), n (%)	Disagree (2), n (%)	Somewhat Disagree (3), n (%)	Neither Agree nor Disagree (4), n (%)	Somewhat Agree (5), n (%)	Agree (6), n (%)	Strongly Agree (7), n (%)	Median (IQR)
I felt confident that any information I send to a caregiver would be received.	0 (0)	0 (0)	0 (0)	0 (0)	1 (33.3)	0 (0)	2 (66.6)	7 (6.5-7)
I felt comfortable communicating with my adult day center using it	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (33.3)	2 (66.6)	—[a]

[a]Not applicable.

Explanatory Qualitative Findings (Caregivers and Staff)

Consistent with the high ratings for convenient communication and improved access (75-88%), interviews with caregivers and staff, and open-ended written comments, revealed eight major themes that described CareMobi's value and limitations in daily practice.

In general, caregivers expressed strong support for CareMobi in their open-ended survey responses: "CareMobi is very helpful to me." "It is great." "Keep it up :)" "I found the app to be user friendly and easy to communicate to others."

Transparency

Caregivers valued the app's ability to provide "visibility into the day" reducing uncertainty about what occurred at the center and allowing them to interpret behaviors at home. One staff member shared: "If they're acting different, then we may know it's because they didn't get a good night rest." Family caregivers described using updates to understand changes in appetite and weight: "He eats all the stuff at the center he won't touch at home. Seeing that helped me realize the weight loss wasn't from not eating at all." Another wrote, "This app gives caregivers the comfort of knowing they are a tap away from knowing the daily report of their loved ones."

Streamlined Communication and Time Savings

Staff and caregivers highlighted that CareMobi replaced "phone tag" with more efficient messaging. A staff member explained: "Better communication between our center and the families… We can just put it in CareMobi and communicate that way." Staff also emphasized the time saved: "Instead of spending half the morning calling families one by one, I can send the update once in the app." This freed staff to focus more on direct care tasks.

Guidance on What to Track

Both groups requested clearer guidance on what domains to log. Caregivers wanted consistent updates on hydration and bathroom use, saying "It's interesting that we don't track that in here." Staff asked for templates (eg, "Today at a Glance," "Transport Update") to reduce duplication and ensure the right staff member provided the right update. As one caregiver reflected:

Value Across the Care Trajectory

Staff reported the app was particularly helpful for new enrollees, giving families reassurance and setting expectations from the start. One staff member noted: "Really helpful with the new enrollees, because they don't know any different." While families of newly enrolled clients used it to communicate with center staff, established caregivers used it more to maintain daily stability and track ongoing concerns.

Technology Fit and Integration

While usability was generally positive, staff emphasized the need to minimize duplicate data entry and expressed interest in a web or desktop option. One explained: "[Caregivers] have their own portals and our own things you're using." This theme mirrored broader structural challenges, as many ADCs lack interoperable EHRs and rely on paper or billing software.

Operational Constraints

Participation was shaped by staffing capacity and caregiver demands. Rural centers with small teams were able to sustain use; however, caregivers cited competing responsibilities: "I'm juggling work and his appointments, some days I just can't get to it." One caregiver noted they still relied on phone calls for reassurance: "I have great communication with [the center] by phone. If any problems occur I know they will contact me. I also know I can contact them by phone."

Trust, Confidence, and Relationship-Building

Routine updates fostered trust and provided caregivers with peace of mind. One caregiver explained: "The better you know how things are going throughout the day, the better the day is for you both." Importantly, transparency also strengthened caregivers' confidence in communicating with physicians: "When I went to the doctor, I could show them what was happening at the center. I felt more prepared." Multiple caregivers described using CareMobi data to adjust medications: "We actually got him off one of his meds, because I could prove he was eating and drinking fine there." Staff echoed that regular updates improved caregiver satisfaction and contributed to steadier enrollment: "This app will give caregivers the comfort of knowing they are a tap away from knowing the daily report of their loved ones."

XSL•FO
RenderX

Discussion

Principal Findings

This study evaluated the acceptability, demand, and practicality of CareMobi, a mobile platform designed to strengthen communication between ADCs and family caregivers. We posit that by making ADC-based care more transparent to families (through timely updates and role-specific, low-burden interactions), CareMobi, if feasible and acceptable, can reduce uncertainty and build trust among caregivers. This, in turn, can support sustained participation in ADCs and bolster engagement and connectedness for people living with dementia who are at risk of loneliness and social isolation. Our findings provide preliminary support for this proposition.

Acceptability and Satisfaction

CareMobi demonstrated high acceptability among both caregivers and staff, meeting or exceeding all prespecified STOP-AMEND-GO thresholds [17]. Specifically, usability scores surpassed the >75% "agree/strongly agree" criterion across all items, satisfaction levels exceeded benchmarks, and more than 70% of caregivers engaged with the app at least once. Staff usability scores reached ceiling levels, with median ratings of 7 (strongly agree) across nearly all items. Acceptability was also met, with more than three-quarters of participants expressing interest in using CareMobi beyond the study. Caregivers emphasized the platform was "user friendly," "very helpful," and reassured them by being "a tap away from knowing the daily report of their loved ones." Together, these findings provide strong evidence of acceptability and support CareMobi's potential as a low-burden, scalable tool for use in ADCs, and provide robust evidence for moving forward with broader implementation trials.

Demand and Engagement

Technology adoption in dementia caregiving populations has historically been low, with many studies reporting attrition, underuse, or abandonment of digital tools due to steep learning curves, lack of personalization, or misalignment with caregivers' daily realities [18,19]. In contrast, CareMobi achieved high engagement among study volunteers: 93% of caregivers logged into the app at least once, and 75% of caregivers and 100% of staff expressed willingness to continue use beyond the study. This success may be attributed to its co-designed, user-centered development process, which emphasized simplicity. Staff specifically highlighted reduced "phone tag," noting that streamlined messaging saved time and allowed them to redirect effort toward direct care. This represents an important advantage over other health care technologies that inadvertently add to staff burden rather than reducing it [12]. **Staff observed that bidirectional communication via the app was more likely to be achieved with caregivers who were new to the ADC, as they relied on CareMobi as their primary means of communication with staff. In contrast, established caregivers, who were more accustomed to phone-based communication, tended to use the app more passively to track updates. This suggests that** caregivers who have developed routines with staff may perceive less added value in switching.

It is worth noting that our sample was composed of caregivers who reported relatively high baseline confidence in managing their loved ones' needs, which may have facilitated their willingness to engage with a new tool. Many were also sole caregivers, relying on the ADC as an extension of their kinship network. This close relational tie to the center may have further motivated consistent use of CareMobi, as research suggests that engagement is supported not only by the design of the technology but also by the social context in which it was deployed [20].

Implementation and Practicality

ADC staff were able to integrate CareMobi into their workflows, but several refinements were identified. Both caregivers and staff requested more structured guidance on which observations should be logged (eg, hydration, bathroom trips, and mood changes) and templates such as "Today at a Glance" or "Medication Updates" to enhance efficiency. Staff expressed demand for a desktop or web version since most entered data from a computer, as opposed to a smartphone or tablet. Research shows that community-based settings, especially those with limited resources, are often excluded from technological innovation and struggle because of a fragmented digital infrastructure [21,22]. By focusing on low-cost, high-utility design, and implementing requested refinements, CareMobi can help address these structural inequities, ensuring that centers with fewer resources are not left behind in digital health advancement.

Caregiver Confidence and Trust

Another critical finding was CareMobi's impact on caregiver confidence and trust. Families reported using real-time updates to interpret behaviors at home, guide conversations with medical providers, and even advocate for deprescribing when daily logs showed adequate eating and drinking. One caregiver shared: "We actually got him off one of his meds, because I could prove he was eating and drinking fine there." This echoes existing evidence that improved caregiver communication can enhance self-efficacy, reduce decisional conflict, and promote timely care coordination [23,24]. By strengthening transparency, CareMobi appears to foster not only trust in ADCs but also caregiver empowerment in broader health care decision-making. As another caregiver said, "The better you know how things are going throughout the day, the better the day is for you both."

Equity and Representation

Although findings were promising, the sample was predominantly Non-Hispanic White, female, and English-speaking. This does not reflect the broader demographics of ADC families nationally, where approximately 58% of clients identify as racial or ethnic minorities [4]. Language accessibility was a major limitation; CareMobi is not available in Spanish or other languages. This restricted participation for centers serving diverse populations. This represents both a limitation of the current evaluation and a critical direction for future work. Addressing language and cultural tailoring is essential for scaling CareMobi equitably and ensuring its benefits extend to the diverse populations that rely most on adult day services.

XSL•FO
RenderX

Broader Implications

Collectively, these findings suggest that CareMobi can address longstanding barriers to ADC use. By improving "day-visibility," simplifying communication, and providing structured yet flexible updates, the app has the potential to stabilize attendance and support sustained participation, barriers that have historically undermined ADCs' ability to demonstrate value in long-term care [25,26]. This is particularly timely given the underfunding of adult day services and their critical role in providing cost-effective, community-based care for people with dementia [7].

Limitations and Future Directions

While this pilot yielded encouraging results, this study had several limitations. First, although appropriate for a feasibility design, the sample size was small and drawn from 2 centers, limiting generalizability. The sample was also demographically homogenous (ie, primarily Non-Hispanic White, female, and English-speaking) and did not reflect the diversity of families served by ADCs nationally. As participation required English proficiency and smartphone ownership, findings may overrepresent caregivers who were more comfortable with technology and positively predisposed to digital tools. Second, the study duration was relatively short, constraining our ability to draw conclusions about long-term engagement or patient outcomes. Third, while we evaluated usability, satisfaction, and acceptability, we did not measure downstream effects like caregiver burden or health care use. These outcomes will be the focus of future trials.

The findings from this feasibility study will directly inform the next phase of research, which will involve a larger, multisite hybrid effectiveness–implementation trial. Future work will assess whether improved communication and transparency translate into measurable outcomes such as reductions in caregiver burden, improved people living with dementia engagement and quality of life, more consistent ADC attendance, and reduced acute-care use. Recruitment will intentionally prioritize racially, ethnically, and linguistically diverse caregivers and centers, including development of multilingual platform versions and community-engaged recruitment strategies."

Conclusions

By coupling digital transparency with low-burden communication, CareMobi offers a promising approach to modernizing ADCs in ways that respect resource constraints and caregiver realities. In doing so, it reduces caregivers' uncertainty about sending people living with dementia to the ADC, in turn, providing the people living with dementia an opportunity to socialize and engage. To fully realize this potential, future work must prioritize equity, especially linguistic inclusion, infrastructure support, and engagement in under-resourced settings. Ultimately, if scaled thoughtfully, tools like CareMobi may enrich the social fabric of dementia care, increase accessibility of ADCs, and deepen the connection between caregivers, centers, and people living with dementia.

Acknowledgments

The authors acknowledge their colleagues at the National Adult Day Services Association, as well as the adult day center staff and caregivers who devoted their time to support this research.

Data Availability

The data are unavailable to access publicly due to privacy and ethical restrictions.

Funding

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number: K23AG071948. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

Conceptualization: TS (lead)
Data curation: LP (lead) and JF (supporting)
Formal analysis: TS (lead) and JF (supporting)
Funding acquisition: TS (lead)
Investigation: TS (lead)
Methodology: TS (lead) and JF (supporting)
Project administration: LP (lead), SS (equal), and TS (supporting)
Supervision: TS (lead)
Validation: JF (lead)
Visualization: SS (lead) and JF (supporting)
Writing – original draft: TS (lead), SS (supporting), and LP (supporting)
Writing – review and editing: SS (lead), TS (supporting), and JF (supporting)

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. National Academies of Sciences, Engineering, and Medicine, Division of Behavioral and Social Sciences and Education, Health and Medicine Division, Board on Behavioral, Cognitive, and Sensory Sciences, Board on Health Sciences Policy, Committee on the Health and Medical Dimensions of Social Isolation and Loneliness in Older Adults. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System. Washington (DC): National Academies Press (US); Feb 2020:316.
2. Harris E. Meta-analysis: social isolation, loneliness tied to higher mortality. JAMA 2023 Jul 18;330(3):211. [doi: 10.1001/jama.2023.11958] [Medline: 37379015]
3. Xu L, Liu Y, He H, Fields NL, Ivey DL, Kan C. Caregiving intensity and caregiver burden among caregivers of people with dementia: The moderating roles of social support. Arch Gerontol Geriatr 2021;94:104334. [doi: 10.1016/j.archger.2020.104334] [Medline: 33516077]
4. Lendon JP, Singh P, Lu Z. Adult day services center participant characteristics: United States, 2022. NCHS Data Brief 2024(502):00. [doi: 10.15620/cdc/157498] [Medline: 39302253]
5. New AARP report: majority of adults 50-plus want to age in place; policies, communities must catch up. AARP Public Policy Institute. 2024. URL: https://press.aarp.org/2024-12-10-New-AARP-Report-Majority-Adults-50-plus-Age-Place-Policies-Communities-Catch-Up [accessed 2026-03-18]
6. Abugroun A, Shah SJ, Covinsky K, Hubbard C, Newman JC, Fang MC. Low social engagement and risk of death in older adults. J Am Geriatr Soc 2025;73(7):2166-2175. [doi: 10.1111/jgs.19511] [Medline: 40395045]
7. Sadarangani T, Fernandez Cajavilca M, Qi X, Zagorski W. Adult day services: a potential antidote to social isolation and loneliness in marginalized older adults. Front Public Health 2024;12:1427425 [FREE Full text] [doi: 10.3389/fpubh.2024.1427425] [Medline: 39310908]
8. Leocadie MC, Roy MH, Rothan-Tondeur M. Barriers and enablers in the use of respite interventions by caregivers of people with dementia: an integrative review. Arch Public Health 2018;76:72 [FREE Full text] [doi: 10.1186/s13690-018-0316-y] [Medline: 30479766]
9. Castro AR, Ould Brahim L, Chen Q, Arnaert A, Quesnel-Vallée A, Moffatt K, et al. Information and communication technologies to support the provision of respite care services: scoping review. JMIR Nurs 2023;6:e44750 [FREE Full text] [doi: 10.2196/44750] [Medline: 37252760]
10. Zheng A, Bergh M, Patel Murali K, Sadarangani T. Using mHealth to improve communication in adult day services around the needs of people with dementia: mixed methods assessment of acceptability and feasibility. JMIR Form Res 2024;8:e49492 [FREE Full text] [doi: 10.2196/49492] [Medline: 38427418]
11. Garnett A, Northwood M, Ting J, Sangrar R. mHealth interventions to support caregivers of older adults: equity-focused systematic review. JMIR Aging 2022;5(3):e33085 [FREE Full text] [doi: 10.2196/33085] [Medline: 35616514]
12. Zhong J, Boafo J, Brody A, Wu B, Sadarangani A. A qualitative analysis of communication workflows between adult day service centers and primary care providers. J Am Med Inform Assoc 2022;29(5):882-890 [FREE Full text] [doi: 10.1093/jamia/ocab284] [Medline: 34964467]
13. Wai SY, WaiShiang C, Phang P, Lam K, Phang E, Jali NB. Goal question metric as an interdisciplinary tool for assessing mobile learning application. International Journal of Advanced Computer Science and Applications (IJACSA) 2022;13(5):1-8. [doi: 10.14569/ijacsa.2022.0130518]
14. Piggott CA, Zimmerman S, Reed D, Sloane PD. Development and testing of a measure of caregiver confidence in medical sign/symptom management. Am J Alzheimers Dis Other Demen 2017;32(7):373-381 [FREE Full text] [doi: 10.1177/1533317517711247] [Medline: 28558474]
15. Vlachogianni P, Tselios N. Perceived usability evaluation of educational technology using the Post-Study System Usability Questionnaire (PSSUQ): a systematic review. Sustainability 2023;15(17):12954 [FREE Full text] [doi: 10.3390/su151712954]
16. Lewis JR. Psychometric evaluation of the Post-Study System Usability Questionnaire: the PSSUQ. Proceedings of the Human Factors Society Annual Meeting 1992;36(16):1259-1260 [FREE Full text] [doi: 10.1177/154193129203601617]
17. Herbert E, Julious SA, Goodacre S. Progression criteria in trials with an internal pilot: an audit of publicly funded randomised controlled trials. Trials 2019;20(1):493 [FREE Full text] [doi: 10.1186/s13063-019-3578-y] [Medline: 31399148]
18. Grewal KS, Gowda-Sookochoff R, Peacock S, Cammer A, McWilliams LA, Spiteri RJ, et al. Perspectives on technology use in the context of caregiving for persons with dementia: qualitative interview study. JMIR Form Res 2024;8:e63041 [FREE Full text] [doi: 10.2196/63041] [Medline: 39671589]
19. Panzavolta A, Arighi A, Guido E, Lavorgna L, Di Lorenzo F, Dodich A, et al. Patient-related barriers to digital technology adoption in Alzheimer disease: systematic review. JMIR Aging 2025;8:e64324 [FREE Full text] [doi: 10.2196/64324] [Medline: 40209218]

XSL•FO
RenderX

20. Hoel V, Ambugo EA, Wolf-Ostermann K. Sustaining our relationship: dyadic interactions supported by technology for people with dementia and their informal caregivers. Int J Environ Res Public Health 2022;19(17):10956 [FREE Full text] [doi: 10.3390/ijerph191710956] [Medline: 36078671]
21. Roy S, Lartey ST, Durneva P, Jha N, Ofori MA, Zeba Z, et al. Digital health technology infrastructure challenges to support health equity in the United States: scoping review. J Med Internet Res 2025;27:e70856 [FREE Full text] [doi: 10.2196/70856] [Medline: 40958727]
22. Lin CC, Dievler A, Robbins C, Sripipatana A, Quinn M, Nair S. Telehealth in health centers: key adoption factors, barriers, and opportunities. Health Aff (Millwood) 2018;37(12):1967-1974. [doi: 10.1377/hlthaff.2018.05125] [Medline: 30633683]
23. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. JAMA 2014;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]
24. Smith PD, Martin B, Chewning B, Hafez S, Leege E, Renken J, et al. Improving health care communication for caregivers: a pilot study. Gerontol Geriatr Educ 2018;39(4):433-444 [FREE Full text] [doi: 10.1080/02701960.2016.1188810] [Medline: 27379507]
25. Du Preez J, Millsteed J, Marquis R, Richmond J. The role of adult day services in supporting the occupational participation of people with dementia and their carers: an integrative review. Healthcare (Basel) 2018;6(2):43 [FREE Full text] [doi: 10.3390/healthcare6020043] [Medline: 29738489]
26. Savard J, Leduc N, Lebel P, Béland F, Bergman H. Determinants of adult day center attendance among older adults with functional limitations. J Aging Health 2009;21(7):985-1015. [doi: 10.1177/0898264309344311] [Medline: 19773597]

Abbreviations

ADC: adult day center
CCSM: Caregiver Confidence in Sign/Symptom Management Scale
HIPAA: Health Insurance Portability and Accountability Act
mHealth: mobile health
PSSUQ: Post-Study System Usability Questionnaire

Edited by A Quialheiro; submitted 30.Sep.2025; peer-reviewed by N Ogbonna, R Zhang; comments to author 03.Dec.2025; revised version received 09.Dec.2025; accepted 30.Jan.2026; published 24.Apr.2026.

Please cite as:
Sadarangani T, Peralta L, Siamdoust S, Finik J
CareMobi to Improve Communication Between Caregivers and Adult Day Centers for People With Dementia: Mixed Methods Feasibility Study
JMIR Aging 2026;9:e85090
URL: https://aging.jmir.org/2026/1/e85090
doi:10.2196/85090

©Tina Sadarangani, Laura Peralta, Shahrzad Siamdoust, Jackie Finik. Originally published in JMIR Aging (https://aging.jmir.org), 24.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Social Determinants and Oral Health Among Indigenous and Non-Indigenous Older Adults in La Araucanía, Chile: Cross-Sectional Study

Fernanda Muñoz-Sepúlveda[1], MSc; Karen Danke Hausdorf[2], PhD; Víctor Beltrán Varas[1], PhD; Claudia Acevedo Atala[1], PhD; Leonardo López[3], PhD; Alejandro Herrera[4], BA in Anthropology

[1]Clinical Investigation and Dental Innovation Center (CIDIC), Dental School and Center for Translational Medicine (CEMT-BIOREN), Universidad de La Frontera, Temuco, Chile

[2]Facultad de Odontología, Universidad San Sebastián, Santiago, Chile

[3]Instituto de Informática Educativa, Universidad de La Frontera, Temuco, Chile

[4]Facultad de Ciencias Agropecuarias y de Medioambiente, Universidad de La Frontera, Temuco, Chile

Corresponding Author:
Víctor Beltrán Varas, PhD
Clinical Investigation and Dental Innovation Center (CIDIC), Dental School and Center for Translational Medicine (CEMT-BIOREN)
Universidad de La Frontera
Francisco Salazar 1145
Temuco, 4811230
Chile
Phone: 56 45 2325000
Email: victor.beltran@ufrontera.cl

Abstract

Background: Social disparities in oral health persist, disproportionately affecting Indigenous populations.

Objective: This study aims to compare oral health status, dental care access, treatment needs, and social determinants between Mapuche and non-Mapuche older adults in the La Araucanía Region of Chile, within a hybrid dental care model, and to explore whether ethnicity is associated with these outcomes, guided by the World Health Organization (WHO) Commission on Social Determinants of Health (CSDH) framework.

Methods: We conducted a cross-sectional exploratory analytic study using data from 480 adults aged 60 years or older who received dental care through a hybrid teledentistry and mobile clinic model (Geriatric Dental Specialties Tele-Platform [TEGO]) in 4 municipalities of La Araucanía. The analytic sample was restricted to community-dwelling, functionally independent older adults. Outcomes included oral health status (nonfunctional dentition, edentulism, cavitated caries, and periodontitis), dental care access (fixed prosthetic rehabilitation and unmet denture need), and dental treatment needs (restorative and prosthetic). Structural (sex, age, education, rurality, and ethnicity) and intermediary determinants (behavioral, biological, psychosocial, and health system access) were analyzed following the WHO-CSDH framework. Bivariate analyses compared outcomes between Mapuche and non-Mapuche participants. Multivariable logistic regression models examined associations between determinants and outcomes.

Results: A total of 471 older adults (mean age 72.8, SD 7.7 years) were included; 76.9% (362/471) had nonfunctional dentition, 18% (85/471) were edentulous, 25.6% (98/383) had cavitated caries, and 46% (154/335) had periodontitis. In bivariate analyses, Mapuche participants showed lower educational attainment, higher rural residence, poorer oral health, more behavioral risk factors (eg, infrequent toothbrushing and higher sugar intake), greater restorative needs, and reduced access to complex dental care compared with non-Indigenous peers. In multivariable analyses, female sex was associated with increased odds of edentulism but lower odds of periodontitis. Lower education was associated with nonfunctional dentition (odds ratio [OR] 4.28, 95% CI 2.34-7.88; *P*<.001) and edentulism (OR 3.04, 95% CI 1.52-6.09; *P*=.002). Mapuche ethnicity was associated with higher prevalence of cavitated caries (OR 2.40, 95% CI 1.25-4.63; *P*=.009) and periodontitis (OR 2.43, 95% CI 1.21-4.88; *P*=.01). Dental care access was associated with behavioral factors (brushing, interproximal hygiene, and sugar intake) and structural determinants (age, sex, and rurality), while restorative and prosthetic needs were linked to sugar consumption and tobacco use, respectively.

XSL•FO
RenderX

Conclusions: Oral health inequalities among older adults in La Araucanía were associated with both structural and behavioral determinants. Although Mapuche participants exhibited poorer oral health, higher restorative needs, and reduced access to complex dental care in bivariate analyses, ethnicity remained associated only with cavitated caries and periodontitis in adjusted models. These findings underscore the importance of incorporating social determinants into oral health planning for underserved older adults and highlight the potential utility of data generated through digitally supported territorial care models for informing equity-oriented oral health strategies.

(JMIR Aging 2026;9:e85075) doi:10.2196/85075

KEYWORDS

delivery of health care; digital health; geriatric dentistry; health inequities; health services accessibility; human rights; Indigenous peoples; rural population

Introduction

Oral health disparities are a persistent global challenge, disproportionately affecting socially marginalized groups, including older adults and Indigenous populations [1-5]. These inequities are shaped by social determinants of health—the conditions in which people live, work, grow, and age [6]. The World Health Organization (WHO) Commission on Social Determinants of Health (CSDH) framework critically distinguishes between structural determinants (eg, socioeconomic position, education, ethnicity, age, and place of residence) and intermediary determinants (material circumstances, biological factors, psychosocial factors, and health system access), which jointly influence oral health outcomes [6-9]. Among older adults, multiple studies have documented robust social gradients in outcomes such as edentulism and oral health–related quality of life [10-15].

Despite global efforts to improve oral health among older adults, progress has been limited, highlighting the need for approaches that account for contextual complexities and structural inequities impacting health outcomes [16,17]. Indigenous populations worldwide experience some of the most persistent and systemic oral health inequities, deeply rooted in historical and structural injustices [2-5,18]. Older Indigenous populations may face compounded barriers to accessing oral health care due to the intersection of age, ethnicity, socioeconomic status, and geographic location [19].

Chile's Indigenous population is among those experiencing the highest levels of aging in the region, with a growing proportion of older adults, including a notably high share of people aged 80 years or older [20]. Approximately 8.8% of the national population self-identifies as Indigenous, and about 16% of this group is aged 60 years or older [21]. Nearly 80% of Indigenous people in Chile belong to the Mapuche, who have historically inhabited Wallmapu, the ancestral territory of the Mapuche people spanning present-day Chile and Argentina; its western portion, Ngulu Mapu, corresponds to what is now south-central Chile [21,22]. In recent decades, a considerable proportion of the Mapuche population has migrated to urban and semiurban areas nationwide, a process associated with longstanding structural inequities, loss of ancestral lands, and limited economic opportunities in rural communities [22,23]. Despite their demographic weight and historical relevance, Mapuche communities continue to experience pronounced health inequities compared with their non-Indigenous peers. At age 60, Mapuche individuals have a life expectancy 7.5 years shorter than their non-Indigenous counterparts, along with lower total and disability-free life expectancy and a greater number of years lived with disability [24,25]. These disparities are also reflected in documented differences in metabolic health, including higher levels of insulin resistance, as well as a greater prevalence of mental and communication impairments compared with non-Mapuche older adults [26,27]. Together, these findings illustrate broader inequalities in general health and aging conditions affecting Mapuche older adults.

Oral health status reflects similar social and health gradients observed among Indigenous populations worldwide. Several studies have documented a high prevalence of untreated dental caries, suboptimal oral hygiene practices, and limited access to preventive oral health services among Chilean Indigenous children and adults [28-31], as well as among Indigenous communities globally [2,3,32]. But despite growing awareness of oral health disparities, research on Indigenous populations in Latin America and Chile remains scarce.

In addressing persistent barriers to oral health care among vulnerable populations, innovative service delivery models, such as telehealth and mobile clinics, have gained traction for their ability to overcome geographic isolation and specialist shortages [33,34]. These approaches have shown considerable promise in expanding access to both general and oral health services, alleviating the burden of chronic diseases, and supporting culturally safe, community-centered models of care for Indigenous populations [34-36]. Such strategies are particularly relevant given the strong preference among many older Indigenous individuals to age in place, often in settings characterized by limited financial resources and scarce health infrastructure [37]. At the global level, digital health research targeting older Indigenous populations is gaining momentum—particularly in high-income countries—through the evaluation of innovations such as wearable technologies, videoconferencing, and mobile telemedicine clinics [38-42]. However, these advances remain less common in low- and middle-income countries, where disparities in technological infrastructure, digital literacy, cultural adaptation, and resource availability pose significant challenges to the implementation and sustainability of culturally appropriate telemedicine and eHealth solutions [20,43].

In response to these challenges, a pilot initiative was implemented in a Mapuche community, introducing a hybrid dental care model designed to reduce territorial barriers to oral

XSL•FO
RenderX

health services. The model integrates outreach clinical services delivered through a mobile unit with digitally supported coordination and remote specialist consultation, linking primary-level care with higher-level clinical expertise. Hybrid care approaches that combine digital and in-person services have been proposed as strategies to improve access and continuity of care in underserved populations [44]. The digital infrastructure supporting this model was provided by the Geriatric Dental Specialties Tele-Platform (TEGO), a web-based system that integrates electronic health records with structured medical-geriatric and odontogeriatric assessment modules and facilitates interdisciplinary collaboration [29,45-47].

Building on this initial experience, the care model was subsequently expanded to rural and urban municipalities across the La Araucanía Region. Although not exclusively targeted at Indigenous populations, the program engaged both Mapuche and non-Mapuche older adults across diverse settings. This broader implementation generated data that enabled the examination of associations between structural and intermediary social determinants of health and oral health outcomes in these populations.

Accordingly, this study aims to compare oral health status, dental care access, treatment needs, and social determinants of health between Mapuche and non-Mapuche older adults in the La Araucanía Region of Chile receiving care through a hybrid dental care model and to explore the association between ethnicity and these outcomes within the WHO-CSDH framework.

Methods

Ethical Considerations

The project was approved by the Ethics Committee of the Universidad de La Frontera (decision 109/22). All participants provided written informed consent, and strict confidentiality protocols were followed, including anonymization prior to analysis.

Overview

This cross-sectional exploratory analytic study is based on the integrated analysis of data collected at the conclusion of a community-based oral health intervention conducted in the La Araucanía Region of Chile. The intervention corresponded to the regional implementation of the previously described hybrid dental care model and was delivered through a mobile clinical unit serving rural and urban municipalities. Services included geriatric dental assessment and treatment, preventive and restorative care, prosthetic evaluation, and referral management according to treatment complexity. Clinical information, coordination of care, and specialist consultation were supported through the TEGO digital platform following standardized protocols described in previous publications [29,45,47,48].

Participants and Setting

The La Araucanía Region is distinguished by its high proportion of Indigenous Mapuche population (34.2%) [21]. The region continues to experience persistent socioeconomic disadvantages, including elevated poverty rates, a predominantly rural context, and pronounced health inequities, particularly reflected in limited access to preventive and routine oral health care services [49,50].

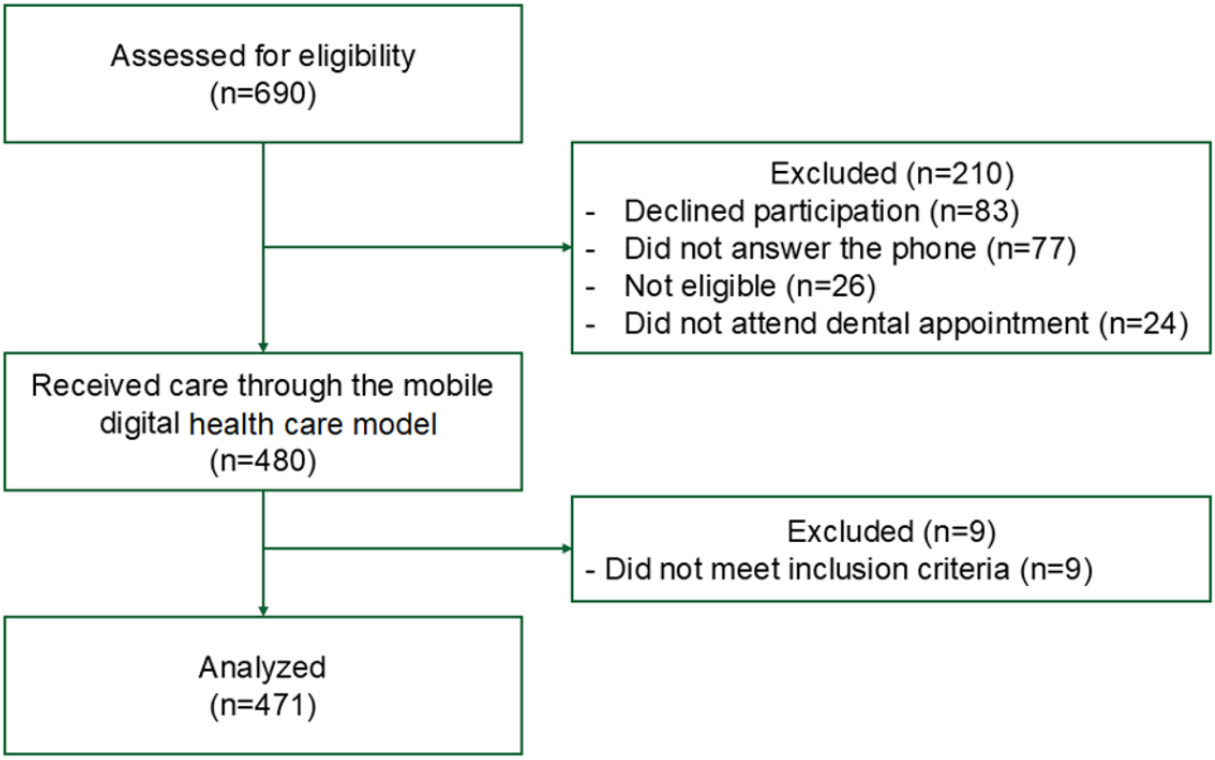
In response to these challenges, a community-based oral health intervention program was implemented with funding from a national Presidential Grant (exempt resolution number 472). Initial recruitment was conducted by telephone using databases provided by the regional branch of the National Senior Citizen Service, in collaboration with local community organizations. Municipalities were selected based on logistical feasibility and implementation criteria rather than probabilistic sampling. Program eligibility included adults aged 60 years or older, residing in the region, and physically and cognitively able to attend dental appointments.

Between January and September 2023, a total of 480 older adults received care through this program, which delivered services through scheduled outreach visits using a mobile dental clinic in 4 municipalities: 2 rural (Cholchol and Los Sauces) and 2 urban (Temuco and Angol). The intervention was implemented in coordination with municipalities and community organizations, including the Araucanía Regional Federation of Older Adults, which brings together municipal-level unions of older adult clubs, as well as local community groups such as church organizations.

This study was based on clinical and assessment data collected during program implementation and recorded through the TEGO digital platform. For the present study, eligibility criteria included being a community-dwelling, functionally independent older adult, defined as a Barthel Index score ≥90. Consequently, 9 institutionalized individuals were excluded, resulting in a final sample of 471 participants. Figure 1 presents the detailed participant flow, including eligibility assessment, application of inclusion and exclusion criteria, and the final analytic sample.

XSL•FO
RenderX

Figure 1. Participant flow diagram.



Conceptual Framework

The WHO-CSDH framework informed the research question, variable selection and classification, and interpretation of findings by conceptualizing oral health as shaped by processes of social stratification and the unequal distribution of power, resources, and opportunities [6]. This framework distinguishes between structural determinants—such as socioeconomic and political context and axes of social stratification, including education, ethnicity, and place of residence—and intermediary determinants, which encompass material conditions, health-related behaviors, psychosocial factors, and access to health systems. Within this framework, ethnicity was conceptualized as a structural determinant reflecting social position, guiding comparisons of oral health outcomes across groups.

The use of the WHO-CSDH framework is particularly relevant in this study context, as it allows oral health inequalities to be examined beyond individual-level risk factors by situating them within broader social and systemic conditions. This is especially important in Indigenous populations, where health disparities are shaped by historical, social, and territorial inequities that are not fully captured by conventional epidemiological approaches focused on proximal determinants [9]. By organizing variables into structural determinants, intermediary determinants, and oral health outcomes, this framework supports a more comprehensive assessment of how health inequalities are patterned across social groups and how different levels of determinants coexist.

Consistent with the WHO-CSDH model, structural and intermediary determinants are theorized to be hierarchically related through pathways linking social position to health outcomes. However, given the cross-sectional design, the present study examines the distribution of these determinants and their associations with oral health outcomes without inferring causal relationships.

Data Collection and Variable Structure

Clinical data were collected by 2 final-year dental interns under direct supervision of 2 experienced dentists (CAA and FM-S), using a standardized clinical protocol under uniform examination conditions. Examinations were conducted by the same team that implemented the initial pilot intervention in the Mapuche community, and the regional phase began 1 month later, maintaining the same examiners, instruments, and diagnostic criteria.

Prior to implementation, examiners underwent structured theoretical and supervised clinical training to apply the complete TEGO clinical record, which includes medical-geriatric, odontogeriatric, behavioral, and prosthodontic assessment modules. Training covered caries detection, periodontal examination, prosthetic evaluation, and standardized collection of self-reported variables (eg, oral hygiene practices and service use).

Formal calibration with assessment of interexaminer and intraexaminer reliability was conducted for dental caries using the International Caries Detection and Assessment System (ICDAS). Interexaminer and intraexaminer κ values were 0.66 and 0.78-0.85, respectively [29]. Periodontal and prosthetic assessments were recorded using standardized objective diagnostic criteria under direct supervision to ensure procedural consistency.

XSL•FO
RenderX

All findings were recorded directly into the TEGO [45,47], ensuring standardized data capture and quality control procedures.

Guided by the WHO-CSDH framework, variables were classified as structural determinants, intermediary determinants, or oral health outcomes.

Structural Determinants

Structural determinants were defined as variables reflecting the broader socioeconomic context that shapes health equity through social stratification. These included: sex (male or female, self-reported); age, recorded as a continuous variable in years and categorized as 60-69 years, 70-79 years, or ≥80 years; and ethnic self-identification, classified as Mapuche (based on self-reported Indigenous affiliation) or non-Mapuche. Educational attainment was dichotomized as no formal or only primary education (≤8 years) vs secondary or higher education (>8 years).

Socioeconomic status was approximated using Chile's Social Registry of Households (RSH), a national administrative classification used to target social programs. The RSH assigns households to 7 socioeconomic percentiles based on an algorithm that integrates income sources, household composition, and indicators of vulnerability (eg, age, disability, and access to goods and services) [51]. For this study, households were dichotomized into the most vulnerable group (0%-40%) vs higher percentiles (>40%).

Additional structural variables included place of residence, classified as urban or rural according to municipality-level administrative definitions [52], and health insurance coverage, categorized according to the Chilean health system as the National Health Fund (FONASA), the health systems for the armed forces and police (National Defense Social Security Fund [CAPREDENA] and Carabineros de Chile Social Security Directorate [DIPRECA]), the Comprehensive Health Care and Reparation Program for Victims of Human Rights Violations (PRAIS), private health insurance institutions (ISAPREs), or no health coverage.

Intermediary Determinants

Overview

These variables represent the mechanisms through which structural determinants influence health, including behavioral, biological, and psychosocial factors and access to care.

Behavioral Factors

Included oral hygiene practices, dietary habits, and smoking behavior, which are known to influence oral and general health [53-56]. Participants reported the frequency of tooth or denture brushing, interproximal hygiene (yes or no), and use of toothpaste (yes or no). Cariogenic dietary patterns were assessed through questions about the consumption of sugary foods and beverages between meals—such as candies, cookies, cakes, soda, sweetened juices, or sugar-added drinks (eg, coffee, tea, mate, or milk; yes or no). Daily water intake was also recorded and categorized as less than or more than 2 liters. Smoking status was documented through a yes or no question regarding current tobacco use.

Biological Factors

Several biological conditions—multimorbidity, polypharmacy, dry mouth, and swallowing disorders—were included as biological factors that may influence oral health, given their established association with impaired oral function, increased risk of caries, periodontal disease, and challenges in maintaining oral hygiene in older adults [57-60]. Multimorbidity and polypharmacy were obtained through self-report. Multimorbidity was defined as the presence of 2 or more chronic conditions, as reported by participants [61], while polypharmacy was defined as the regular use of 5 or more medications per day [62], based on participants' declarations of medication intake. The subjective sensation of dry mouth was recorded, and swallowing disorders were assessed using the Eating Assessment Tool-10 (EAT-10), a validated screening instrument; scores ≥3 indicate a positive screen for swallowing difficulties [63,64].

Psychosocial Factors

Psychosocial factors included variables related to social relationships, used as proxies for social isolation—a recognized risk factor for health in older adults [11]. Living arrangements were recorded and categorized as living alone or living with others. Perceived social support was assessed based on participants' subjective perception of receiving emotional or practical help when needed (yes or no).

Health System Access

Access to oral health services was assessed through various direct and indirect indicators. Self-reported time since last dental visit, categorized as ≤2 years or >2 years, was used as a direct indicator of access. The presence of fixed dental rehabilitation (eg, crowns and fixed bridges) and the presence of dental implants were used as indirect indicators of a history of better access to specialized dental treatment, whereas unmet denture need (absence of a removable dental prosthesis in individuals who were clinically identified as requiring a partial or complete denture) was used as a proxy for barriers to care.

In Chile, treatments such as restorations and removable dentures are commonly provided within primary care, although prosthetic programs do not cover all individuals [65]. In contrast, more advanced treatments, including dental implants and fixed prostheses, are typically available only through secondary care or private services. These indicators therefore reflect not only the availability of services but also the types of services they accessed, going beyond the traditional measure of time since the last dental visit. This approach aligns with previous studies that have used restorative, prosthetic, and visit-related measures as proxies for access to care and oral health inequalities [66,67].

Oral Health Status

Oral health status was assessed through clinical examination and included the number of natural teeth, functional dentition, edentulism, cavitated caries, and periodontal status. The number of natural teeth was recorded as the total count of teeth remaining in the oral cavity, while nonfunctional dentition was defined as the presence of fewer than 20 teeth, regardless of

XSL•FO
RenderX

their location. Edentulism was defined as the complete absence of natural teeth.

Cavitated caries were defined as lesions corresponding to the decayed (D) component of the decayed, missing, and filled teeth (DMFT) index, operationalized as ICDAS codes 4 to 6.

Periodontal status was assessed through a full-mouth periodontal examination, recording plaque levels, probing depth (PD), clinical attachment level (CAL), bleeding on probing, suppuration, and gingival recession at 6 sites per tooth, as well as furcation lesions and tooth mobility [68]. Based on these measures, participants were classified as having gingival health (<10% bleeding sites with PD ≤3 mm [69]), gingivitis (≥10% bleeding sites with PD ≤3 mm), or periodontitis (CAL ≥2 mm on nonadjacent teeth, or buccal/oral CAL ≥3 mm with PD >3 mm on ≥2 teeth [70]). For analytical purposes, periodontal status was dichotomized as presence vs absence of periodontitis.

Treatment Needs

Treatment needs were assessed through clinical examination and included both restorative and prosthetic requirements. Restorative treatment need was defined as the number of teeth requiring intervention due to caries, defective restorations, or other conditions requiring restorative care (eg, fractures or noncarious lesions). Including this measure provides a broader estimate of treatment requirements and supports planning and resource allocation.

Normative prosthetic treatment need was operationalized as a binary variable (presence or absence). Participants were classified as having prosthetic treatment need when clinical examination identified complete edentulism without prosthetic rehabilitation, partial tooth loss requiring prosthetic replacement, or existing removable prostheses in inadequate condition requiring repair or replacement. For participants with removable prostheses (complete or partial), functional adequacy was assessed based on stability, retention, and structural integrity. Stability and retention were evaluated following the Kapur index and the prosthodontic evaluation methods from phase 1 of the Third National Health and Nutrition Examination Survey (NHANES III) [71,72], assessing denture movement under manual lateral and vertical forces. Structural integrity was determined by identifying fractures, missing or chipped teeth, and broken clasps or framework components. Prostheses were classified as requiring repair or not requiring repair according to these criteria.

Statistical Analysis

Descriptive analyses were conducted to characterize the study population. Categorical variables were summarized as frequencies and percentages, and continuous variables as means with SDs or 95% CIs, as appropriate. Bivariate analyses compared structural determinants, intermediary determinants, and oral health indicators by ethnicity (Mapuche vs non-Mapuche). Associations between categorical variables were assessed using the Pearson chi-square test, with *P* values of <.05 considered statistically significant.

Guided by the WHO-CSDH framework, bivariate and multivariable logistic regression analyses were then conducted to examine the associations between structural and intermediary determinants and selected outcomes. Three groups of outcomes were evaluated: clinical oral health outcomes (edentulism, nonfunctional dentition, periodontitis, and caries), access-related outcomes (time since the last dental visit, presence of fixed dental rehabilitation, and clinically unmet denture need), and treatment needs (restorative and prosthetic). Separate multivariable logistic regression models were fitted for each binary outcome (9 models in total) to estimate adjusted associations between structural and intermediary determinants and the outcomes.

Missing data were present in a small proportion of observations for several variables (<10%), and complete-case analysis was therefore considered appropriate given the overall low level of missingness. The proportion of missing data for periodontitis was higher (51/386, 13%). For this outcome, missingness was not significantly associated with most evaluated covariates (all *P*>.05), although excluded participants were older, suggesting a potential age-related component. For all other variables, no substantial differences were observed between included and excluded participants across the evaluated sociodemographic characteristics. Missing periodontal data most likely reflects incomplete clinical examinations rather than item nonresponse, as periodontal assessment requires full-mouth probing in dentate participants. Such incomplete examinations may occur due to participant refusal, logistical constraints during fieldwork, or medical conditions preventing the assessment.

Because some outcomes were only applicable to specific subgroups (eg, dentate individuals or denture users) and some covariates had incomplete observations, the number of observations varied across models; therefore, logistic regression analyses were conducted using complete-case data for each model. Covariates were selected a priori based on the conceptual framework of social determinants of health, rather than through data-driven selection procedures. Indicators of health system access were conceptualized as intermediary determinants; however, in specific models they were also analyzed as outcomes to explore inequalities in access to care. In models examining oral health status outcomes, time since the last dental visit was included as a covariate to account for differential exposure to dental care. Models were also adjusted for the number of natural teeth present, when applicable, given its potential role as a confounder. Results are presented as odds ratios (ORs) with 95% CI, and all analyses were conducted using Stata (version 18; StataCorp LLC).

Results

Sociodemographic Characteristics and Structural Determinants

A total of 471 community-dwelling older adults were included in the analysis (Figure 1). Of these, 24.2% (114/471) self-identified as Mapuche. The mean age of the sample was 72.8 (SD 7.7) years, with Mapuche participants being significantly younger than non-Mapuche counterparts (*P*=.006).

Univariate and bivariate descriptive statistics for the study variables are presented in Table 1. Overall, 59.9% (282/471) of

XSL•FO
RenderX

participants were female, with no significant differences between Mapuche and non-Mapuche participants (*P*=.62). A total of 126 participants were not registered in RSH, as enrollment in this administrative system is voluntary in Chile. Among those registered (n=345 of 471), most (301/345, 87.2%) were classified within the ≤40% bracket, indicating high socioeconomic vulnerability. No significant differences were observed between Mapuche and non-Mapuche participants (*P*=.10). Regarding health coverage, most participants (405/471, 85.9%) were enrolled in FONASA, 9.8% (46/471) had no health insurance, and 4.2% (20/471) were covered by other schemes (ISAPRE, CAPREDENA/DIPRECA, and PRAIS). Coverage differed by ethnicity; nearly all Mapuche participants (111/114, 97.4%) had FONASA, while non-Mapuche participants showed more variation, including a higher proportion without insurance (45/357, 12.6%; χ^2_2=16.6; *P*<.001). Compared with their non-Indigenous counterparts, Mapuche individuals had lower educational attainment and resided predominantly in rural municipalities.

XSL•FO
RenderX

Table 1. Distribution of structural and intermediary determinants and oral health indicators according to ethnic group.

Variable and category	Total, n (%)	Non-Mapuche, n (%)	Mapuche, n (%)	*P* value
Structural determinants				
Sex				.62
Male	189 (40.1)	141 (39.5)	48 (42.1)	
Female	282 (59.9)	216 (60.4)	66 (57.9)	
Age group (years)				.003
60-69	184 (39.1)	124 (34.7)	60 (52.6)	
70-79	184 (39.1)	148 (41.5)	36 (31.6)	
≥80	103 (21.8)	85 (23.8)	18 (15.8)	
Educational attainment (years completed)				.003
≤8	252 (53.5)	177 (49.6)	75 (65.8)	
>8	219 (46.5)	180 (50.4)	39 (34.2)	
Rurality				<.001
Rural	188 (39.9)	99 (27.7)	89 (78.1)	
Urban	283 (60.1)	258 (72.3)	25 (21.9)	
Intermediary determinants				
Behavioral factors				
Brushing frequency (times per day; missing=12)				.001
<2	67 (14.6)	40 (11.5)	27 (24.1)	
≥2	392 (85.4)	307 (88.5)	85 (75.9)	
Interproximal oral hygiene[a]				.38
Yes	121 (31.3)	95 (32.5)	26 (27.7)	
No	265 (68.7)	197 (67.5)	68 (72.3)	
Sugary snacks consumption				<.001
Yes	240 (51)	161 (45.1)	79 (69.3)	
No	231(49)	196 (54.9)	35 (30.7)	
Daily water intake (≥2 liters)				.79
Yes	170 (36.1)	130 (36.4)	40 (35.1)	
No	301 (63.9)	227 (63.6)	74 (64.9)	
Smoking habit				.59
Smoker	43 (9.1)	34 (9.5)	9 (7.9)	
Nonsmoker	428 (90.9)	323 (90.5)	105 (92.1)	
Biological factors				
Polypharmacy (missing=38)				.15
Yes	72 (16.6)	59 (18.1)	13 (12.2)	
No	361 (83.4)	267 (81.9)	94 (87.8)	
Multimorbidity				.35
Yes	286 (60.7)	221 (61.9)	65 (57)	
No	185 (39.3)	136 (38.1)	49 (43)	
Dry mouth				.06
Yes	246 (52.2)	178 (49.9)	68 (59.7)	
No	225 (47.8)	179 (50.1)	46 (40.3)	

XSL•FO
RenderX

Variable and category	Total, n (%)	Non-Mapuche, n (%)	Mapuche, n (%)	*P* value
Swallowing disorders (EAT-10[b])				.007
Yes	38 (8.1)	22 (6.2)	16 (14)	
No	433 (91.9)	335 (93.8)	98 (86)	
Psychosocial factors				
Living arrangement				.008
Living alone	142 (30.2)	119 (33.3)	23 (20.2)	
Living with family	329 (69.8)	238 (66.7)	91 (79.8)	
Social support				.18
Yes	399 (84.7)	298 (83.5)	101 (88.6)	
No	72 (15.3)	59 (16.5)	13 (11.4)	
Health system factors				
Self-reported time since last dental visit (years; missing=4)				.16
≤2	229 (49)	180 (50.8)	49 (43.4)	
>2	238 (51)	174 (49.2)	64 (56.6)	
Unmet denture need				.44
Yes	70 (18.6)	55 (19.5)	15 (16)	
No	306 (81.4)	227 (80.5)	79 (84)	
Fixed rehabilitation				.003
Yes	75 (15.9)	67 (18.8)	8 (7)	
No	396 (84.1)	290 (81.2)	106 (93)	
Dental implants				.16
Yes	6 (1.3)	6 (1.7)	0 (0)	
No	465 (98.7)	351(98.3)	114 (100)	
Oral health status				
Nonfunctional dentition				.10
Yes	362 (76.9)	268 (75.1)	94 (82.5)	
No	109 (23.1)	89 (24.9)	20 (17.5)	
Periodontitis (missing=51)[a]				<.001
Yes	154 (46.0)	97 (38.6)	57 (67.9)	
No	181(54.0)	154 (61.3)	27 (32.1)	
Cavitated caries (missing=3)[a]				<.001
Yes	98 (25.6)	60 (20.7)	38 (40.9)	
No	285(74.4)	230(79.3)	55 (59.1)	
Complete edentulism (total tooth loss)				.87
Yes	85 (18)	65 (18.2)	20 (17.5)	
No	386 (82)	292 (81.8)	94 (82.5)	
Dental treatment need				
Restorative treatment needs (missing=3)[a]				.001
No need (0)	177(46.2)	143 (49.3)	34 (36.6)	
Moderate need (1-3)	174 (45.4)	131 (45.2)	43 (46.2)	
Severe need (≥4)	32 (8.4)	16 (5.5)	16 (17.2)	
Prosthetic treatment needs (missing=1)				.13

XSL•FO
RenderX

Variable and category	Total, n (%)	Non-Mapuche, n (%)	Mapuche, n (%)	*P* value
Yes	314 (66.8)	232 (64.9)	82 (72.6)	
No	156 (33.2)	125 (35.1)	31 (27.4)	

[a]Only dentate participants were considered (n=386).

[b]EAT-10: Eating Assessment Tool-10.

Intermediary Determinants

Behavioral Factors

Regarding behavioral determinants, overall, 85.4% (392/459) of participants reported brushing their teeth or dentures at least twice daily, whereas only 3.4% (16/459) reported never brushing. Among participants who were not completely edentulous, 31.9% (123/386) reported performing interproximal dental hygiene. Cariogenic dietary patterns were common, with 51% (240/471) consuming sugary foods or drinks between meals, with an average frequency of 2.3 (SD 0.9) times per day. Mapuche participants reported lower toothbrushing frequency and higher sugary consumption between meals (*P*<.001). Overall, 9.1% (43/471) of participants were current smokers. No statistically significant differences were observed between Mapuche and non-Mapuche participants in smoking status (*P*=.59) or daily water consumption (*P*=.79; Table 1).

Biological Factors

In terms of biological factors, the mean number of prescription medications per day was 2.63 (95% CI 2.43-2.83), and 16.6% (72/433) of participants met the criteria for polypharmacy. The mean number of chronic conditions was 2.36 (95% CI 2.20-2.51), with 60.7% (286/471) experiencing multimorbidity. Dry mouth was self-reported by 52.2% (246/471) of participants, while 8.1% (38/471) screened positive for swallowing disorders. Only swallowing disorders showed significant differences by ethnicity, being more prevalent among Mapuche participants (*P*=.007; Table 1).

Psychosocial Factors

Overall, perceived social support was high, with 84.7% (399/471) reporting access to emotional or practical help when needed, and 69.8% (329/471) living with family members. Mapuche participants were less likely to live alone compared with non-Mapuche participants (*P*=.008).

Health System Factors

Approximately half of participants (229/467, 49%) reported visiting a dentist within the last 2 years, with no significant differences by ethnicity (Table 1). Overall, 18.6% (70/376) of participants had unmet denture need, with no significant differences by ethnicity (*P*=.44).

More complex rehabilitations were less common. Fixed prosthetic rehabilitation was observed only in 15.9% (75/471) of participants (mean fixed rehabilitation per recipient: 3.63, 95% CI 2.88-4.38), while the presence of dental implants was rare (6/471, 1.3%), with a mean of 2.5 (95% CI 0.77-4.22) implants per recipient. Some participants received both types of treatments.

In the bivariate comparisons by ethnicity, no differences were observed in time since the last dental visit. However, Mapuche participants had a significantly lower prevalence of fixed prostheses (*P*=.003). A lower prevalence of implants was also observed among Mapuche participants, although this difference was not statistically significant (Table 1).

Oral Health Status

Only 4% (19/471) of participants had 28 or more natural teeth. The mean number of natural teeth present was 12.93 (95% CI 12.13-13.72), with no significant differences between Mapuche and non-Mapuche participants (*P*=.86). Nonfunctional dentition was present in 76.9% (362/471) of participants, while 18.1% (85/471) were edentulous; no significant differences were observed by ethnicity (Table 1).

All participants had a history of dental caries. The mean DMFT index was 22.82 (95% CI 22.23-23.42), ranging from 6 to 32. Among participants aged 65 to 74 years, the mean DMFT was 22.03 (95% CI 21.10-22.95). No significant differences were observed between Mapuche participants (mean DMFT 23.12, SD 5.4) and non-Mapuche participants (mean DMFT 21.6, SD 6.5) within this age group (*P*=.14) or overall (*P*=.09). Tooth loss represented the largest component of the index, accounting for 83.6% (mean 19.07), followed by restorations, which accounted for 15% (mean 3.42). Approximately one-quarter of dentate participants (98/383, 25.6%) had cavitated caries, with a significantly higher prevalence among Mapuche participants (*P*<.001).

Among dentate participants, 20.6% (69/335) had periodontal health, 33.4% (112/335) had gingivitis, and 45.9% (154/335) had periodontitis. Significant differences were observed between Mapuche and non-Mapuche participants (*P*<.001; Table 1).

Dental Treatment Needs

Restorative treatment needs were observed in 53.8% (206/383) of participants, and among dentate individuals, the mean number of teeth requiring restorative treatment was 1.25 (95% CI 1.23-2.24). Most participants used dentures (306/470, 65.1%). However, denture repair needs were high: 79.7% (244/306) of denture users required repair or replacement due to poor fit, retention, or structural defects. Overall, normative prosthetic treatment need was 66.8% (314/470). Mapuche participants had significantly greater restorative treatment needs, but prosthetic treatment need did not differ significantly between Mapuche and non-Mapuche participants (Table 1).

Association of Ethnicity and Other Social Determinants With Oral Health Status

After conducting bivariate analyses for all selected outcomes and covariates, multivariable logistic regression models were fitted to estimate the adjusted effect of each independent

XSL•FO
RenderX

variable. Statistically significant bivariate associations are reported in the corresponding table. Given the exploratory nature of the analysis, all predictor variables were included in the multivariable models regardless of their bivariate significance in order to control for potential confounding and obtain adjusted estimates.

In the multivariable logistic regression models, structural determinants exhibited the most consistent associations with oral health outcomes. Age and educational attainment were significantly associated with tooth loss: participants aged 80 years or older had higher odds of nonfunctional dentition (OR 5.55, 95% CI 2.34-13.08; *P*<.001) and complete edentulism (OR 3.75, 95% CI 1.62-8.68; *P*=.002) compared with those aged 60 to 69 years, while ≤8 years of education was linked to increased odds of nonfunctional dentition (OR 4.28, 95% CI 2.34-7.88; *P*<.001) and complete edentulism (OR 3.04, 95% CI 1.52-6.09; *P*=.002). Women had higher odds of edentulism (OR 2.38, 95% CI 1.24-4.58; *P*=.009) but lower odds of periodontitis (OR 0.53, 95% CI 0.31-0.90; *P*=.02). Mapuche ethnicity was associated with higher odds of cavitated caries (OR 2.40, 95% CI 1.25-4.63; *P*=.009) and periodontitis (OR 2.43, 95% CI 1.21-4.88; *P*=.01). Rurality showed no significant associations (Table 2).

XSL•FO
RenderX

Table 2. Adjusted associations between structural and intermediary determinants and oral health status outcomes from multivariable logistic regression models. For each binary outcome, a separate multivariable logistic regression model was fitted, including all structural and intermediary predictors. Categorical predictors with more than 2 levels were entered using dummy coding, and the number of natural teeth was included as a covariate.

Variable and category	Nonfunctional dentition		Complete edentulism		Cavitated caries		Periodontitis	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value
Structural determinants								
Sex[b-d]								
Male	1.00 (ref)	__[e]	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Female	1.32 (0.76-2.30)	.32	2.38 (1.24-4.58)	.009	0.81 (0.47-1.39)	.44	0.53 (0.31-0.90)	.02
Age groups (years)[c,d,f]								
60-69	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
70-79	2.01 (1.12-3.60)	.02	1.87 (0.89-3.92)	.09	0.58 (0.32-1.05)	.07	1.06 (0.59-1.91)	.84
≥80	5.55 (2.36-13.08)	<.001	3.75 (1.62-8.68)	.002	0.47 (0.21-1.09)	.07	1.49 (0.69-3.24)	.30
Educational attainment (years)[b,c]								
>8	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
≤8	4.28 (2.33-7.88)	<.001	3.04 (1.52-6.09)	.002	1.61 (0.88-2.95)	.12	0.76 (0.44-1.33)	.34
Rurality[b-d,f]								
Urban	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Rural	1.52 (0.78-2.97)	.22	1.59 (0.83-3.04)	.15	1.00(0.53-1.89)	.98	1.69 (0.91-3.13)	.09
Ethnicity[f]								
None	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Mapuche	1.30 (0.62-2.71)	.48	0.85 (0.40-1.82)	.68	2.40 (1.25-4.63)	.009	2.43 (1.21-4.88)	.01
Intermediary determinants								
Behavioral factors								
Brushing frequency (times per day)								
0	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
≥2	0.8 (0.34-1.89)	.62	0.99 (0.43-2.26)	.98	1.19 (0.55-2.57)	.65	1.16 (0.55-2.48)	.69
Interproximal oral hygiene[b,c,f]								
Yes	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
No	0.37 (0.21-0.65)	.001	0.97 (0.02-0.42)	.002	0.64 (0.33-1.22)	.17	0.8 (0.44-1.45)	.46
Sugary snacks consumption[d]								
No	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Yes	0.91 (0.52-1.58)	.72	0.67 (0.36-1.22)	.18	1.14 (0.67-1.96)	.62	1.86 (1.11-3.12)	.01
Daily water intake (≥2 liters)								
Yes	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
No	0.99 (0.56-1.74)	.96	1.71 (0.91-3.21)	.09	1.08 (0.62-1.90)	.78	0.9 (0.52-1.56)	.70

XSL•FO
RenderX

Variable and category	Nonfunctional dentition		Complete edentulism		Cavitated caries		Periodontitis	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value
Smoking habit[f]								
Nonsmoker	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Smoker	1.33 (0.53-3.36)	.54	2.31 (0.83-6.39)	.10	1.96 (0.86-4.51)	.11	1.25 (0.49-3.14)	.63
Biological factors								
Polypharmacy								
No	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Yes	0.65 (0.31-1.38)	.26	1.28 (0.59-2.78)	.52	0.93 (0.42-2.09)	.86	0.8 (0.38-1.67)	.55
Multimorbidity[f]								
No	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Yes	0.77 (0.42-1.41)	.40	1.17 (0.59-2.30)	.65	0.59(0.34-1.06)	.07	0.74 (0.42-1.30)	.30
Dry mouth[c]								
No	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Yes	2.08 (1.19-3.63)	.009	1.7 (0.93-3.10)	.08	1.78 (1.02-3.13)	.04	1.2 (0.71-2.03)	.49
Swallowing disorders								
No	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Yes	0.56 (0.22-1.43)	.22	0.93 (0.34-2.52)	.88	1.38 (0.58-3.28)	.47	1.05 (0.41-2.67)	.92
Psychosocial factors								
Living arrangement								
With family	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Alone	1.19 (0.65-2.18)	.58	1.28 (0.68-2.43)	.44	1.18 (0.64-2.17)	.59	0.83 (0.47-1.46)	.51
Social support								
Yes	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
No	0.51 (0.21-1.24)	.13	2.1 (0.79-5.59)	.13	1.3 (0.56-3.02)	.54	0.73 (0.35-1.55)	.41
Health System factors								
Self-reported time since last dental visit (years)[b,c]								
>2	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
≤2	1.25 (0.73-2.15)	.41	0.53 (0.28-0.99)	.048	0.75 (0.44-1.28)	.29	1.58 (0.94-2.66)	.08

[a]aOR: adjusted odds ratio.

[b]Indicates variables showing significant bivariate associations with edentulism.

[c]Indicates variables showing significant bivariate associations with nonfunctional dentition.

[d]Indicates variables showing significant bivariate associations with periodontitis.

[e]Not applicable.

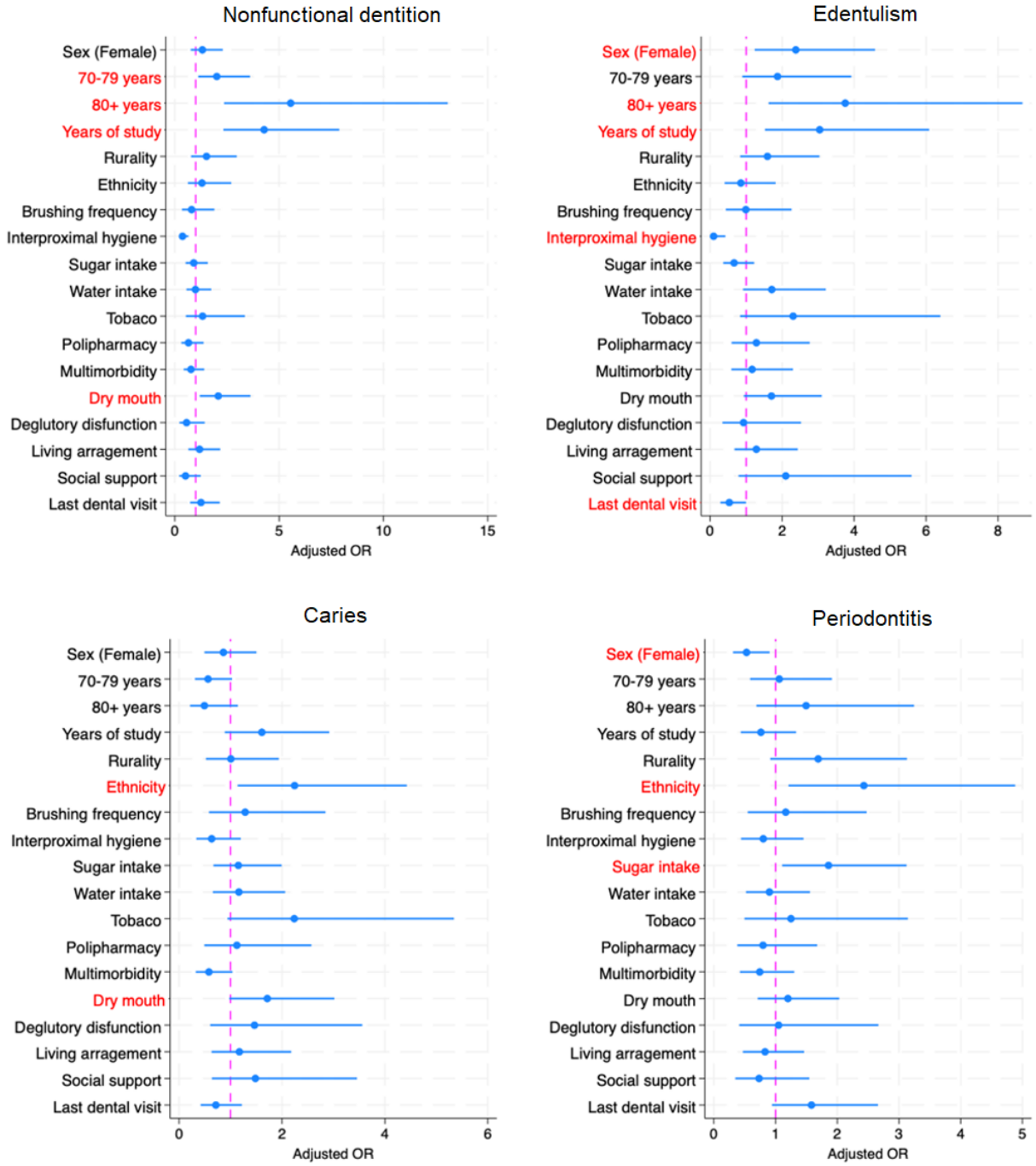
[f]Indicates variables showing significant bivariate associations with caries.

Among intermediary determinants, behavioral factors appeared to be the most relevant. Specifically, lack of interproximal hygiene was linked to tooth loss outcomes, while consumption of sugary snacks was associated with increased odds of periodontitis (OR 1.86, 95% CI 1.11-3.12; *P*=.01; Table 2). Psychosocial factors showed no significant associations, and biological factors had limited associations, with dry mouth linked to nonfunctional dentition (OR 2.08, 95% CI 1.20-3.63;

XSL•FO
RenderX

P=.009) and cavitated caries (OR 1.78, 95% CI 1.02-3.13; *P*=.04). Regarding dental care access, having received care in the past 2 years was associated with lower odds of total edentulism (OR 0.53, 95% CI 0.28-0.99; *P*=.048). These multivariable results, including the 4 oral health outcomes, are summarized in Figure 2.

Figure 2. Adjusted associations between predictor variables and oral health status outcomes. Forest plot showing adjusted odds ratios (aORs) and 95% CIs from separate multivariable logistic regression models for nonfunctional dentition, edentulism, cavitated caries, and periodontitis. Points represent aORs, and horizontal bars represent 95% CIs. The vertical line indicates an odds ratio (OR) of 1 (no association). Statistically significant associations are highlighted in red.



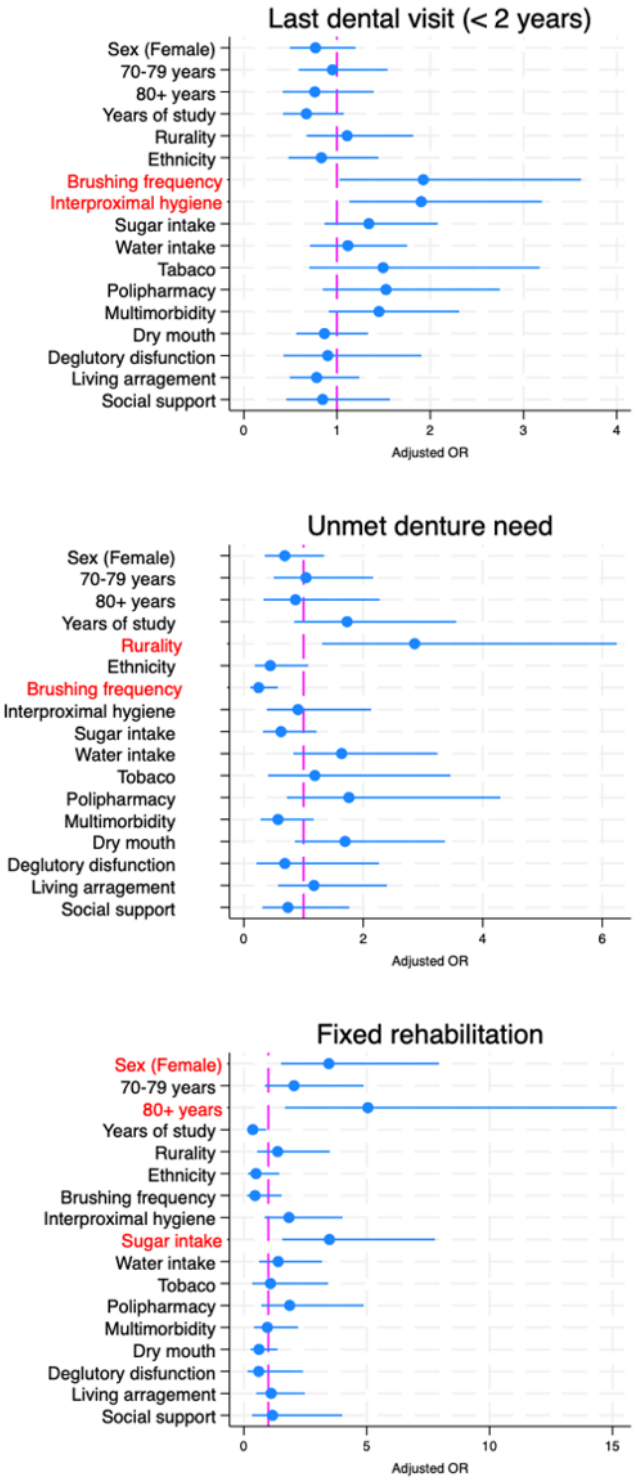
Association of Ethnicity and Other Social Determinants With Access to Dental Care

In the multivariable analysis, variables related to dental care were examined (full results are provided in Multimedia Appendix 1). Time since the last dental visit was associated with toothbrushing frequency (OR 1.92, 95% CI 1.02-3.62; *P*=.04) and interproximal hygiene practices (OR 1.90, 95% CI 1.13-3.19; *P*=.01). Rural residence was associated with higher odds of unmet denture need (OR 2.86, 95% CI 1.31-6.24; *P*=.008), whereas brushing frequency (OR 0.25, 95% CI 0.11-0.57; *P*<.001) was associated with lower odds.

XSL•FO
RenderX

In the adjusted model, women were more likely to have received fixed prostheses (OR 3.46, 95% CI 1.51-7.94; *P*=.003), as were participants aged 80 years or older (OR 5.05, 95% CI 1.68-15.17; *P*=.004). Conversely, participants with ≤8 years of education were less likely to have received fixed prostheses (OR 0.37, 95% CI 0.15-0.91; *P*=.03). Consumption of sugary snacks was also associated with higher odds of having received fixed prosthetic rehabilitation (OR 3.48, 95% CI 1.56-7.78; *P*=.002). Other determinants, including ethnicity, did not show significant associations with this outcome (Figure 3).

Figure 3. Adjusted associations between predictor variables and access to dental care outcomes. Forest plot showing adjusted odds ratios (aORs) and 95% CIs from separate multivariable logistic regression models for last dental visit, unmet denture need, and fixed rehabilitation. Points represent aORs, and horizontal bars represent 95% CIs. The vertical line indicates an odds ratio (OR) of 1 (no association). Statistically significant associations are highlighted in red.

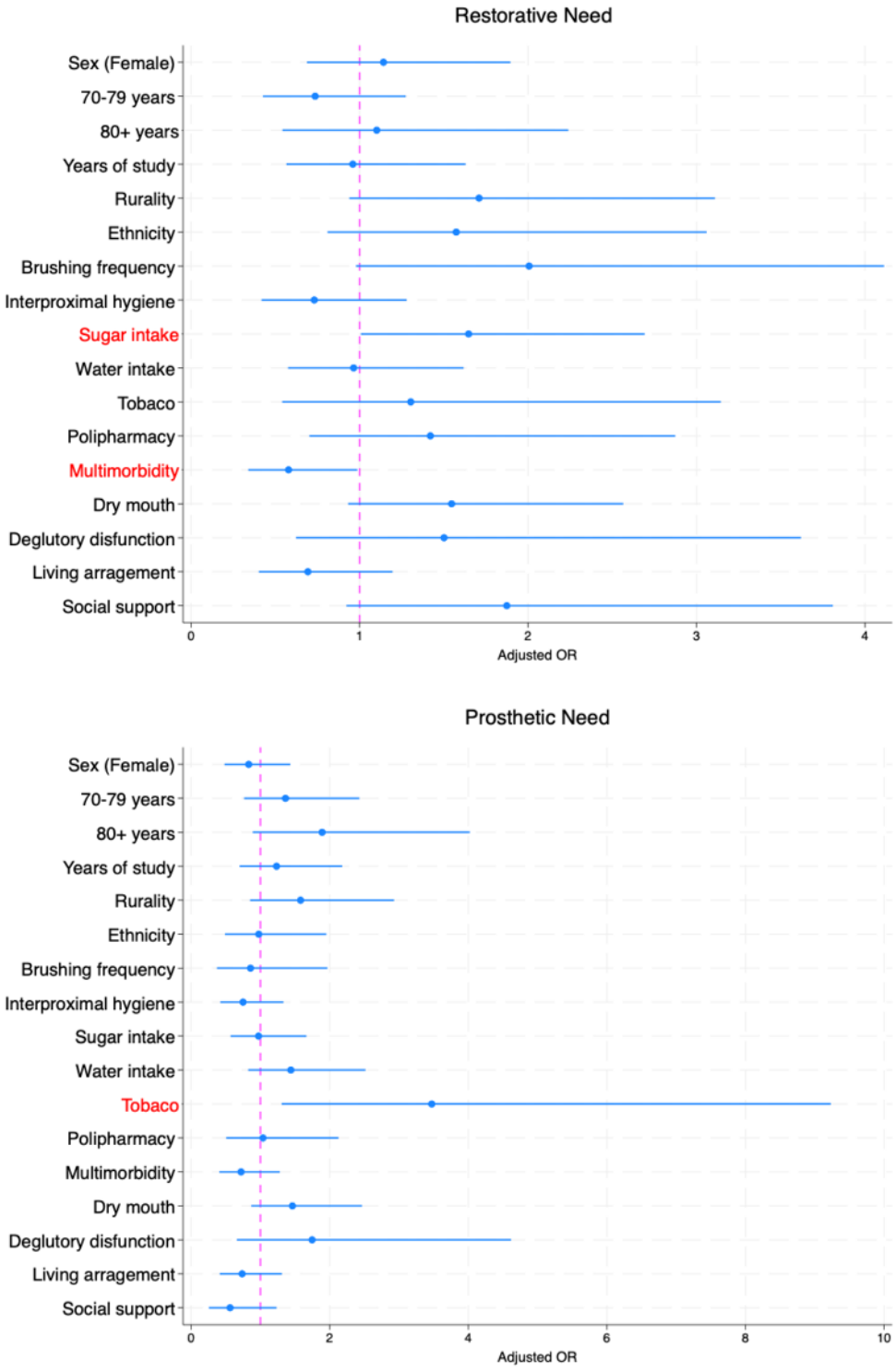


XSL•FO
RenderX

Association of Ethnicity and Other Social Determinants With Treatment Needs

When treatment needs were analyzed, no structural determinants were significantly associated with these outcomes. However, higher consumption of sugary snacks between meals was linked to increased odds of requiring restorative treatment (OR 1.69, 95% CI 1.03-2.77; *P*=.03). In addition, current smoking was associated with higher odds of normative prosthetic treatment need (OR 3.49, 95% CI 1.3-9.31; *P*=.01; Figure 4). Full adjusted model results are presented in Multimedia Appendix 2.

Figure 4. Adjusted associations between predictor variables and dental care needs. Forest plot showing adjusted odds ratios (aORs) and 95% CIs from separate multivariable logistic regression models for restorative treatment need and normative prosthetic treatment need. Points represent aORs, and horizontal bars represent 95% CIs. The vertical line indicates an odds ratio (OR) of 1 (no association). Statistically significant associations are highlighted in red.



XSL•FO
RenderX

Discussion

Principal Findings

This study highlights differences in both structural and intermediary social determinants of health between Indigenous (Mapuche) and non-Indigenous older adults in a sample from the La Araucanía Region of Chile. Compared with non-Indigenous participants, Mapuche individuals had lower educational attainment and were more likely to reside in rural areas. They also showed a higher prevalence of oral diseases, alongside a higher number of risk factors, including less frequent toothbrushing and higher sugar consumption. In addition, they had greater restorative treatment needs and reduced access to complex dental care services.

When examining the independent influence of structural and intermediary determinants on various oral health outcomes, structural factors showed more consistent associations than intermediary ones, in line with the WHO-CSDH framework. The associations observed for these variables may be consistent with the cumulative influence of structural disadvantage across the life course, whereby intermediary determinants in older age could reflect the long-term manifestation of social inequalities. Although our cross-sectional design does not allow direct assessment of life-course processes, longitudinal cohort research has demonstrated that early socioeconomic conditions shape oral health beliefs and behaviors, which subsequently influence adult clinical outcomes and quality of life, supporting the plausibility of cumulative pathways [73].

Notably, the pattern of associations varied depending on the outcome assessed. For instance, female sex was significantly associated with a higher prevalence of edentulism while showing lower odds of periodontitis. Similarly, lower educational attainment was associated with increased prevalence of nonfunctional dentition and edentulism; however, it showed no significant association with caries or periodontitis. For these latter outcomes, ethnicity showed the most consistent association. In particular, Mapuche participants had more than twice the odds of periodontitis compared with non-Mapuche counterparts (OR 2.43, 95% CI 1.21-4.88; *P*=.01). These findings are in line with international evidence reporting a 35% higher prevalence of periodontitis among Indigenous populations compared with non-Indigenous groups [2].

Regarding caries, Chilean adults aged 65 to 74 years exhibit one of the highest burdens worldwide (57,729 per 100,000) [74]. In our sample, all participants had a history of caries experience, and Mapuche participants showed a significantly higher prevalence of untreated (cavitated) caries compared with their non-Indigenous peers. This association remained significant in the multivariable model.

When examining variables related to access to dental care, no significant differences were observed in the frequency of the last dental visit between ethnic groups. However, the lower prevalence of complex treatments, such as dental implants and fixed prosthetic rehabilitations (eg, crowns and bridges), along with higher restorative treatment needs, including active caries and defective restorations, among Indigenous participants suggests a potential disparity in the quality of dental care received.

Ethnicity was significantly associated with 2 of the 9 outcomes in the multivariable models, a pattern similar to that observed for education, sex, and age. Within the WHO-CSDH framework, ethnicity is conceptualized as a structural determinant that interacts with socioeconomic, territorial, and cultural conditions, and is linked to cumulative disadvantages in health [6,9,75]. Historical and structural inequities have been documented as shaping access to services, use of dental prostheses, and the persistence of restorative treatment needs [4,5,76]. These findings highlight that ethnicity, alongside sex, age, and educational attainment, functions as a key axis of social stratification associated with differences in disease distribution and patterns of access and quality of treatment, rather than merely representing background characteristics.

Structural determinants are understood to shape differential exposure to intermediary determinants, including behavioral and biological risk factors [6,75]. In our study, unhealthy lifestyle behaviors, particularly the consumption of sugary snacks between meals, were more prevalent among Mapuche participants and were also linked to periodontitis and restorative treatment needs. These patterns are consistent with broader dietary changes reported in Indigenous populations, where colonization and globalization have contributed to the loss of traditional food systems, with a shift from subsistence diets toward store-bought and processed foods rich in refined sugars [77]. Structural factors, including corporate determinants of health such as the marketing of sugary products, may further reinforce this trend [78]. These dietary changes have been consistently linked with poorer oral health outcomes, including caries and periodontal disease [77]. The higher prevalence of periodontitis and caries in Mapuche participants may be related to a combination of factors, including ethnicity, dietary practices, and a lower brushing frequency, even though brushing frequency was not a significant predictor in the multivariable model. This finding supports the WHO-CSDH framework, which emphasizes the complex and not yet fully understood pathways through which structural determinants translate into biological outcomes [75].

Intermediary determinants played a more limited but still relevant role. Adverse outcomes were associated with dental visits occurring more than 2 years ago, frequent consumption of sugary snacks, and reduced toothbrushing frequency, while dry mouth, a biological factor linked with polypharmacy and aging [79], was also linked to both caries and nonfunctional dentition. Interestingly, better oral hygiene practices were associated with more recent dental visits. Addressing this requires expanding preventive interventions, particularly dietary counseling and management of xerostomia, as well as integrating tailored oral health education at the time of prosthetic delivery to improve long-term outcomes and avoid premature failures.

The Chilean health system, organized under a mixed public-private model, provides coverage to approximately 80% of the national population through its main government-funded scheme, FONASA, which primarily focuses on essential curative oral health care [3-5,80]. In this study, health insurance type

XSL•FO
RenderX

was not included as a covariate since most participants were enrolled in FONASA, resulting in limited variability in coverage. Nevertheless, despite formal insurance coverage, a high burden of oral health problems alongside important gaps in service provision was observed. Although dental visit rates within the past 2 years were comparable across groups, 54% (206/383) of participants required restorative treatment, and a substantial proportion of existing restorations were found to be defective. Access to prosthetic care also showed notable gaps. While the national health guarantee (Explicit Health Guarantees [GES]) provides removable dentures for individuals at age 60 years, and other programs extend this benefit to vulnerable adults aged older than 20 years, important unmet needs persist [65,81-84]. In this study, 2 prosthetic indicators were considered. Unmet prosthetic need referred to individuals requiring a prosthesis but not currently wearing one, serving as a proxy for limited previous access to prosthetic care. In contrast, prosthetic treatment needs included both individuals requiring a new prosthesis and those with existing prostheses requiring repair or replacement, reflecting the overall demand for prosthetic treatment and providing an indicator relevant for service planning. Nearly one-fifth of participants presented unmet denture needs, and among those who already had dentures, almost 80% required repair or replacement. These findings underscore the importance of ensuring the quality of removable dentures provided in both public and private settings, particularly within publicly financed dental care, where more than 300,000 dentures are delivered annually in Chile [65]. A considerable proportion of these prostheses remain unused by patients due to poor fit, discomfort, or insufficient follow-up and adaptation protocols, which are critical to achieving functional rehabilitation and patient adherence [65]. The limited provision of complex treatments, coupled with gender-, education-, and ethnicity-based disparities in the use of fixed prostheses, is indicative of persistent inequalities. Within our sample, access to oral rehabilitation and preventive care appears more limited among older adults characterized by structural disadvantages, including advanced age, lower educational attainment, rural residence, and Indigenous ethnicity. Taken together, these patterns suggest that insurance coverage alone does not guarantee equity in oral health.

An innovative aspect of this study was the implementation of a teledentistry ecosystem (TEGO) supported by a mobile clinic. Although TEGO was designed for the broader older population, its pilot implementation in a Mapuche community demonstrated feasibility, user satisfaction, and a culturally respectful approach through the incorporation of intercultural facilitators, suggesting potential for supporting access to care in underserved communities [29]. Within the WHO-CSDH framework, strategies to reduce health inequities must address both structural and intermediary determinants of health [6,7]. In this context, technology-supported models of care may help mitigate structural barriers related to geographic isolation and limited service availability through remote consultation, screening, and triage, particularly in rural and underserved populations [44,85-88]. At the same time, digital health interventions may also influence intermediary determinants by strengthening oral health literacy, preventive behaviors, and patient engagement [85,89]. In this regard, the TEGO model illustrates a potentially equity-oriented strategy that combines digital tools with mobile service delivery to address barriers across the continuum of care. By facilitating access to primary oral health services (eg, restorative care and repair of removable prostheses) and, in some cases, enabling referral to secondary care for complex treatments such as implants, this approach may help reduce gaps in service use among populations facing structural disadvantages. Such strategies are particularly relevant for older adults living in rural or historically underserved settings and align with broader calls to action aimed at reducing health inequities and expanding access to care, including the objectives of United Nations Sustainable Development Goals, particularly goal 3 (good health and well-being) and goal 10 (reduced inequalities), while also contributing to the growing agenda of digital health equity [35,90-94].

The educational and ethnic inequalities observed in oral disease are consistent with prior evidence and support the need to account for differential disease burden in the design of public health policies [30,67,95]. Ongoing monitoring of these indicators may help identify service gaps and inform more equitable interventions. Even though a “one-size-fits-all” approach focused on individual treatment has done little to improve oral health outcomes at the population level, priority actions should focus on reducing structural disadvantages that influence access to and the cost of dental services, with a strong emphasis on equity [78]. This may include prioritizing socially disadvantaged groups and ensuring that treatments, such as dentures, are provided appropriately to individuals whose anatomical and functional characteristics allow effective use, while simultaneously establishing a subsidized system to improve access to more complex dental treatments for those who require them. Efforts should focus on strengthening the quality and follow-up of dental care, as well as implementing culturally tailored health promotion strategies for Indigenous and rural populations. Additionally, community-based delivery models, including mobile and digitally supported services, may contribute to reducing geographic and social inequities in access to care.

Limitations and Generalizability

Certain limitations should be considered when interpreting these findings. First, the cross-sectional design precludes causal inference. The observed associations reflect patterns of co-occurrence rather than directional relationships, and reverse causation remains plausible. For example, poorer oral health may influence psychosocial domains such as depressive symptoms or social participation in older adults [96]. Accordingly, findings should be interpreted as evidence of social patterning rather than causal pathways.

Second, the sampling strategy and implementation context may have introduced selection bias and limited external validity. Participants were community-dwelling, functionally independent older adults recruited from a publicly funded, mobile, digitally supported care program implemented in municipalities selected for logistical feasibility. Participation required the capacity and willingness to attend scheduled dental visits, which likely underrepresents individuals with greater dependency, mobility limitations, or who are harder to reach—groups that may occupy

XSL•FO
RenderX

more disadvantaged positions within the social gradient. This may have resulted in a healthier or more health-engaged sample, potentially underestimating the magnitude of oral health inequalities. At the same time, as an access-oriented intervention, the program may have attracted individuals with heterogeneous levels of perceived need, including both those seeking care for existing conditions and those engaging in preventive checkups, which may further influence the composition of the analytic sample.

Third, measurement limitations may have affected internal validity. Several intermediary determinants were self-reported, including dental visits, oral hygiene practices, current tobacco use, and chronic conditions used to define multimorbidity. These measures are susceptible to recall and social desirability bias, and the absence of detailed exposure assessment (eg, pack-years of smoking) or verification through clinical records may have introduced measurement error. Clinical outcomes were assessed by trained examiners using standardized protocols and direct supervision; however, periodontal and prosthetic assessments were more susceptible to field variability than caries, which was formally calibrated. These factors may have affected the precision and consistency of associations, particularly for periodontitis and for behavioral and biological variables.

Fourth, residual confounding cannot be excluded. Although the multivariable models were adjusted for the number of natural teeth as a key clinical factor, other potentially relevant confounders may not have been fully captured. In particular, the dataset lacked more granular measures of socioeconomic position (eg, wealth and income) and detailed health system characteristics, limiting the full operationalization of the structural dimension of the WHO-CSDH framework.

Fifth, missing data may have introduced bias. Although missingness was generally low, it was higher for periodontitis and polypharmacy. Incomplete periodontal examinations may occur under field conditions when full assessment is not feasible (eg, due to pain, tooth mobility, or time constraints during clinical evaluation). Missingness was not consistently associated with most measured characteristics, including sex, education, ethnicity, and rurality. However, for periodontitis, excluded participants were older than those included in the analysis, suggesting that examination feasibility may have differed by age. Given the overall low level of missingness and the limited evidence of systematic differences across most covariates, complete-case analysis was considered appropriate. Nonetheless, some degree of bias cannot be entirely excluded.

Finally, the absence of an a priori sample size calculation and the low frequency of some outcomes (eg, implants or complex rehabilitations) resulted in limited statistical precision for certain estimates, as reflected in wide CIs. Also, the study may not have been specifically powered to detect all potential differences between Mapuche and non-Mapuche participants; therefore, subgroup analyses should be interpreted cautiously.

Taken together, these findings are primarily applicable to community-dwelling, functionally independent older adults in rural and urban municipalities of La Araucanía, including both Mapuche and non-Indigenous individuals. Extrapolation to populations with higher levels of dependency or barriers to access, other regions of Chile, larger metropolitan areas, or Indigenous populations with different social or health care contexts should be undertaken cautiously. Moreover, as data were collected within a mobile, digitally supported care model specific to this program, generalizability to older adults receiving standard primary care services may be limited. This program-based implementation reflects real-world conditions of service delivery within its operational context, which may support ecological validity in terms of care processes and service pathways.

Implications and Future Directions

Despite these limitations, this study contributes to the literature by examining oral health outcomes among older adults through a structured social determinants framework that simultaneously models structural determinants, including ethnicity, and intermediary factors. Ethnicity was associated only with cavitated caries and periodontitis. This analytical approach allows observed ethnic differences to be interpreted within broader social and territorial conditions, rather than attributing disparities to ethnic group membership alone. It does not eliminate ethnic gradients but situates them within patterned social conditions that may shape exposure, risk, and access over time [9]. In the Chilean context, national health surveys include self-identified Indigenous affiliation and have examined several social determinants, such as education, income, and insurance status. However, oral health outcomes have rarely been analyzed using multivariable frameworks that incorporate ethnicity alongside other determinants.

The findings further indicate that oral health inequalities in older adults may operate through multiple and outcome-specific pathways. Structural determinants—particularly age and educational attainment—were strongly associated with cumulative indicators of tooth loss, such as nonfunctional dentition and complete edentulism. In contrast, behavioral factors showed closer associations with selected access to dental treatment and treatment needs. This pattern suggests that different levels of determination may be more relevant depending on whether the outcome reflects long-term accumulation of disadvantage (eg, tooth loss) or more proximal conditions and service use patterns.

The inclusion of both rural and urban participants highlights the contextual dimension of these inequalities. While rural residence was not consistently associated with all clinical outcomes, it was linked to selected treatment-related indicators, suggesting that territorial factors may influence patterns of rehabilitation and unmet need even when not uniformly affecting disease prevalence.

Future longitudinal studies incorporating more detailed measures of lifetime exposures, including behavioral factors such as smoking history, as well as health service trajectories and contextual factors, would help clarify the pathways linking structural determinants to oral health outcomes in later life. Research may also benefit from intersectional approaches capable of examining how structural determinants interact to shape oral health disadvantages. Additionally, evaluating the long-term effectiveness and equity of innovative service delivery models, such as teledentistry and mobile dental clinics, may

XSL•FO
RenderX

provide valuable evidence to inform strategies aimed at reducing disparities. Mixed methods studies incorporating qualitative perspectives on social, cultural, and system-level barriers may further contribute to a more comprehensive understanding of the mechanisms underlying the observed patterns.

Conclusions

Oral health inequalities differed between Mapuche and non-Indigenous older adults in La Araucanía and were associated with both structural and behavioral determinants. Structural determinants, particularly lower education, female sex, older age, and ethnicity, showed consistent associations with oral health status, whereas dental care access reflected a combination of structural and behavioral factors. In contrast, current treatment needs were mainly associated with behavioral factors, including sugar consumption and tobacco use. Although Mapuche participants exhibited poorer oral health, higher restorative needs, and reduced access to complex dental care in bivariate analyses, ethnicity remained associated only with cavitated caries and periodontitis in adjusted models.

These findings underscore the importance of incorporating social determinants into oral health planning for underserved older adults, particularly in Indigenous and rural populations experiencing persistent social and territorial inequities. By applying the WHO-CSDH framework within a hybrid teledentistry and mobile clinic model, this study illustrates how digitally supported territorial care programs can generate clinically grounded population health data that help identify inequities in oral health status, access to care, and treatment needs. Integrating these approaches into oral health information systems and service planning may support more equity-oriented prioritization, resource allocation, and delivery of preventive, restorative, and prosthetic care for older adults in underserved territories.

Acknowledgments

The authors thank the Universidad de La Frontera for logistical support and the Inter-University Center of Healthy Aging (CIES) of the Council of Rectors of Chilean Universities (CRUCH) for their collaboration. This study would not have been possible without the support of the FRO2495 MINEDUC-Chile project, “Integrated Technological Ecosystem of Tele-Dentistry to Provide General and Specialized Dental Care for Older Adults,” funded by the Ministry of Education of Chile. The project was awarded to the Universidad de La Frontera under the Plan for Strengthening State Universities, Regional Higher Education (ESR) line.

Data Availability

The datasets analyzed during this study are not publicly available due to privacy restrictions regarding the Indigenous population but are available from the corresponding author on reasonable request, subject to ethical approval.

Funding

This research was funded by the Presidential Grant Project Clinical Center for Innovation in Oral Health of the Elderly in the La Araucanía Region (exempt resolution number 472).

Authors' Contributions

Conceptualization: FM-S, KDH
Data curation: FM-S, KDH
Formal analysis: KDH (lead), FM-S (supporting)
Funding acquisition: VBV (lead), CAA (supporting)
Investigation: CAA, AHA, FM-S
Methodology: FM-S, KDH, LL, VBV
Project administration: VBV (lead), CAA (equal), FM-S (supporting)
Resources: CAA, VBV
Supervision: LLN, AHA
Writing—original draft: FM-S (lead) and KDH (supporting)
Writing—review and editing: VBV (lead), CAA, LLN, AHA

Conflicts of Interest

None declared.

Multimedia Appendix 1
Multivariable logistic regression of structural and intermediary determinants on access to dental care–related outcomes.
[PDF File (Adobe PDF File), 189 KB - aging_v9i1e85075_app1.pdf]

Multimedia Appendix 2
Multivariable logistic regression of structural and intermediary determinants on dental treatment needs.

XSL•FO
RenderX

[PDF File (Adobe PDF File), 167 KB - aging_v9i1e85075_app2.pdf]

References

1. Global strategy and action plan on oral health 2023–2030. World Health Organization. 2024 May 26. URL: https://www.who.int/publications/i/item/9789240090538 [accessed 2025-08-27]
2. Nath S, Poirier B, Ju X, Kapellas K, Haag D, Jamieson L. Periodontal disease inequities among Indigenous populations: a systematic review and meta-analysis. J Periodontal Res 2022;57(1):11-29. [doi: 10.1111/jre.12942] [Medline: 34655251]
3. Nath S, Poirier BF, Ju X, Kapellas K, Haag DG, Ribeiro Santiago PH, et al. Dental health inequalities among Indigenous populations: a systematic review and meta-analysis. Caries Res 2021;55(4):268-287 [FREE Full text] [doi: 10.1159/000516137] [Medline: 34107490]
4. Schuch HS, Haag DG, Kapellas K, Arantes R, Peres MA, Thomson WM, et al. The magnitude of Indigenous and non-Indigenous oral health inequalities in Brazil, New Zealand and Australia. Community Dent Oral Epidemiol 2017;45(5):434-441. [doi: 10.1111/cdoe.12307] [Medline: 28509420]
5. Jamieson LM, Elani HW, Mejia GC, Ju X, Kawachi I, Harper S, et al. Inequalities in Indigenous oral health: Findings from Australia, New Zealand, and Canada. J Dent Res 2016;95(12):1375-1380. [doi: 10.1177/0022034516658233] [Medline: 27445131]
6. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. Paper 2 (Policy and Practice). World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/44489/9789241500852_eng.pdf?sequence=1 [accessed 2025-08-20]
7. Closing the gap in a generation: health equity through action on the social determinants of health: final report of the commission on social determinants of health. World Health Organization. 2008 Aug 27. URL: https://www.who.int/publications/i/item/WHO-IER-CSDH-08.1 [accessed 2025-08-20]
8. Watt RG, Daly B, Allison P, Macpherson LMD, Venturelli R, Listl S, et al. Ending the neglect of global oral health: time for radical action. Lancet 2019;394(10194):261-272. [doi: 10.1016/S0140-6736(19)31133-X] [Medline: 31327370]
9. Pulver LJ, Haswell M, Ring I, Waldon J, Clark W, Whetung V, et al. Indigenous health—Australia, Canada, Aotearoa New Zealand and the United States: laying claim to a future that embraces health for us all. World Health Report (2010). 2010. URL: https://cdn.who.int/media/docs/default-source/health-financing/technical-briefs-background-papers/whr-2010-background-paper-33.pdf?sfvrsn=eb0531c3_3&download=true [accessed 2026-05-05]
10. Aida J, Takeuchi K, Furuta M, Ito K, Kabasawa Y, Tsakos G. Burden of oral diseases and access to oral care in an ageing society. Int Dent J 2022;72(4S):S5-S11 [FREE Full text] [doi: 10.1016/j.identj.2022.06.012] [Medline: 36031325]
11. Hajek A, Kretzler B, König HH. Oral health, loneliness and social isolation. A systematic review and meta-analysis. J Nutr Health Aging 2022;26(7):675-680 [FREE Full text] [doi: 10.1007/s12603-022-1806-8] [Medline: 35842757]
12. Lyshol H, Grøtvedt L, Fagerhaug TN, Feuerherm AJ, Jakhelln G, Sen A. A study of socio-economic inequalities in self-reported oral and general health in South-East Norway. Sci Rep 2022;12(1):13721 [FREE Full text] [doi: 10.1038/s41598-022-18055-5] [Medline: 35962044]
13. Tsakos G, Demakakos P, Breeze E, Watt RG. Social gradients in oral health in older adults: findings from the English Longitudinal Survey of Aging. Am J Public Health 2011;101(10):1892-1899. [doi: 10.2105/AJPH.2011.300215] [Medline: 21852627]
14. Park H, Shin S, Ryu J. Edentulous disparities among geriatric population according to the sexual difference in South Korea: a nationwide population-based study. Sci Rep 2023;13(1):7854 [FREE Full text] [doi: 10.1038/s41598-023-35029-3] [Medline: 37188776]
15. Huang X, Kang L, Bi J. Epidemiology of oral health in older adults aged 65 or over: prevalence, risk factors and prevention. Aging Clin Exp Res 2025;37(1):193. [doi: 10.1007/s40520-025-03110-8] [Medline: 40549221]
16. Slack-Smith L, Ng T, Macdonald ME, Durey A. Rethinking oral health in aging: ecosocial theory and intersectionality. J Dent Res 2023;102(8):844-848 [FREE Full text] [doi: 10.1177/00220345231175061] [Medline: 37314086]
17. Fleming E, Bastos JL, Jamieson L, Celeste RK, Raskin SE, Gomaa N, et al. Conceptualizing inequities and oppression in oral health research. Community Dent Oral Epidemiol 2023;51(1):28-35. [doi: 10.1111/cdoe.12822] [Medline: 36749670]
18. Mejia GC, Parker EJ, Jamieson LM. An introduction to oral health inequalities among Indigenous and non-Indigenous populations. Int Dent J 2010;60(3 Suppl 2):212-215. [Medline: 20718305]
19. Slack-Smith L. Oral health in older Indigenous peoples. Gerodontology 2019;36(2):91. [doi: 10.1111/ger.12417] [Medline: 31124190]
20. The sociodemographic situation of Indigenous peoples in Latin America and the Caribbean: analysis in the context of aging and COVID-19. Pan American Health Organization. 2021. URL: https://www.paho.org/en/documents/sociodemographic-situation-indigenous-peoples-latin-america-and-caribbean-analysis [accessed 2025-08-20]
21. Resultados Censo 2024. Instituto Nacional de Estadísticas (INE), Gobierno de Chile. 2025. URL: https://censo2024.ine.gob.cl/resultados-dashboard/ [accessed 2025-07-29]
22. Iturriaga-Gutiérrez E, Rojo-Mendoza F, Escalona-Ulloa M. Permeability of the indigenous space: discourses of Mapuche landowners on urban expansion in periurbanian Temuco, Araucanía-Chile. Urbano 2020 Nov;23(42):124-134. [doi: 10.22320/07183607.2020.23.42.10]

XSL•FO
RenderX

23. Ancán Jara J. Urban Mapuches: reflections on a modern reality in Chile. Abya Yala News. 1997. URL: https://mapuche.info/archivo/mapuint/ancan00.htm [accessed 2025-08-20]
24. Sandoval MH, Portaccio MEA, Albala C. Ethnic differences in disability-free life expectancy and disabled life expectancy in older adults in Chile. BMC Geriatr 2024;24(1):116 [FREE Full text] [doi: 10.1186/s12877-024-04728-5] [Medline: 38297194]
25. Sandoval MH, Alvear Portaccio ME, Albala C. Life expectancy by ethnic origin in Chile. Front Public Health 2023;11:1147542 [FREE Full text] [doi: 10.3389/fpubh.2023.1147542] [Medline: 37397745]
26. Celis-Morales CA, Perez-Bravo F, Ibañes L, Sanzana R, Hormazabal E, Ulloa N, et al. Insulin resistance in Chileans of European and Indigenous descent: evidence for an ethnicity x environment interaction. PLoS One 2011;6(9):e24690 [FREE Full text] [doi: 10.1371/journal.pone.0024690] [Medline: 21931814]
27. Mella R, Alvear M, Carrillo B, Caire V. [Evaluation of mental and communication functions in Mapuche and non Mapuche elderly subjects in rural communities in Southern Chile]. Rev Med Chil 2003;131(11):1257-1265. [Medline: 14743686]
28. Angel P, Fresno M, Cisternas P, Lagos M, Moncada G. Prevalencia de caries, pérdida de dientes y necesidad de tratamiento en población adulta Mapuche-Huilliche de Isla Huapi. Rev Clin Periodoncia Implantol Rehabil Oral 2010;3(2):69-72. [doi: 10.4067/s0719-01072010000200002]
29. Beltrán V, Muñoz-Sepúlveda F, Acevedo C, Navarro P, Venegas B, Salgado C, et al. A rural teledentistry care experience: a geriatric approach to assessing oral health status and treatment needs in older adults from a Mapuche community in Chile. Front Public Health 2024;12:1356622 [FREE Full text] [doi: 10.3389/fpubh.2024.1356622] [Medline: 38903581]
30. Cantarutti C, Yévenes G, Muñoz-Del-Carpio-Toia A, Adorno-Farias D, Fernández-Ramires R, Santos-Silva AR, et al. Social determinants of oral health in an indigenous community of Chile: preliminary data of a mixed qualitative and quantitative study. BMC Oral Health 2025;25(1):87 [FREE Full text] [doi: 10.1186/s12903-025-05430-7] [Medline: 39825341]
31. de la Maza FJ, Cueto MV. Epidemiological study of oral health in a young adult Mapuche population. Odontol Chil 1989;37(1):183-185. [Medline: 2641955]
32. Soares GH, Pereira NF, Biazevic MGH, Braga MM, Michel-Crosato E. Dental caries in South American Indigenous peoples: a systematic review. Community Dent Oral Epidemiol 2019;47(2):142-152. [doi: 10.1111/cdoe.12436] [Medline: 30506750]
33. Estai M, Kanagasingam Y, Tennant M, Bunt S. A systematic review of the research evidence for the benefits of teledentistry. J Telemed Telecare 2018;24(3):147-156. [doi: 10.1177/1357633X16689433] [Medline: 28118778]
34. Shrivastava R, Couturier Y, Girard F, Papineau L, Emami E. Two-eyed seeing of the integration of oral health in primary health care in indigenous populations: a scoping review. Int J Equity Health 2020;19(1):107 [FREE Full text] [doi: 10.1186/s12939-020-01195-3] [Medline: 32605562]
35. Choukou M, Maddahi A, Polyvyana A, Monnin C. Digital health technology for Indigenous older adults: a scoping review. Int J Med Inform 2021;148:104408. [doi: 10.1016/j.ijmedinf.2021.104408] [Medline: 33609927]
36. Beks H, Ewing G, Charles JA, Mitchell F, Paradies Y, Clark RA, et al. Mobile primary health care clinics for Indigenous populations in Australia, Canada, New Zealand and the United States: a systematic scoping review. Int J Equity Health 2020;19(1):201 [FREE Full text] [doi: 10.1186/s12939-020-01306-0] [Medline: 33168029]
37. Jones L, Jacklin K, O'Connell ME. Development and use of health-related technologies in Indigenous communities: critical review. J Med Internet Res 2017;19(7):e256 [FREE Full text] [doi: 10.2196/jmir.7520] [Medline: 28729237]
38. McIlduff CD, Acharibasam J, Starr V, Chapados M. Engaging Indigenous older adults with technology use to respond to health and well-being concerns and needs. Healthc Manage Forum 2022;35(5):257-264 [FREE Full text] [doi: 10.1177/08404704221103521] [Medline: 35670368]
39. Moecke DP, Holyk T, Beckett M, Chopra S, Petlitsyna P, Girt M, et al. Scoping review of telehealth use by Indigenous populations from Australia, Canada, New Zealand, and the United States. J Telemed Telecare 2024;30(9):1398-1416 [FREE Full text] [doi: 10.1177/1357633X231158835] [Medline: 36911983]
40. Henson C, Freedman B, Rambaldini B, Carlson B, Parter C, Nalliah CJ, et al. Wearables are a viable digital health tool for older Indigenous adults living remotely in Australia (research). Digit Health 2024;10:20552076241277039 [FREE Full text] [doi: 10.1177/20552076241277039] [Medline: 39221087]
41. Henson C, Chapman F, Shepherd G, Carlson B, Rambaldini B, Gwynne K. Amplifying older Aboriginal and Torres Strait Islander women's perspectives to promote digital health equity: co-designed qualitative study. J Med Internet Res 2023;25:e50584 [FREE Full text] [doi: 10.2196/50584] [Medline: 37847550]
42. Henson C, Chapman F, Shepherd G, Carlson B, Rambaldini B, Gwynne K. How older Indigenous women living in high-income countries use digital health technology: systematic review. J Med Internet Res 2023;25:e41984 [FREE Full text] [doi: 10.2196/41984] [Medline: 37071466]
43. Global diffusion of eHealth: making universal health coverage achievable: report of the third global survey on eHealth. World Health Organization. 2016. URL: https://www.who.int/publications/i/item/9789241511780 [accessed 2025-07-29]
44. Drafta S, Macris A, Petre AE. Innovations on the Horizon: teledentistry, artificial intelligence, and hybrid models to improve oral health for vulnerable communities. Front Oral Health 2025;6:1649715 [FREE Full text] [doi: 10.3389/froh.2025.1649715] [Medline: 40860795]

XSL•FO
RenderX

45. Beltrán V, Acuña-Mardones P, Díaz J, Alvarado E, von Marttens A. TEGO: a new concept of teledentistry for the elderly through a web platform and mobile app in the context of the COVID-19 pandemic. J Oral Res 2022;S-1(Seminario Científico):1-8. [doi: 10.17126/joralres.2022.023]
46. Beltrán V, Acuña-Mardones. P, Fernández-Gil F, Acuña-Mardones D, Engelke W. Development of a teledentistry system for the elderly in the context of the COVID-19 pandemic: 3D diagnostic models integration in a web platform. J Oral Res 2022;S-1(Seminario Científico):1-7. [doi: 10.17126/joralres.2022.024]
47. Beltrán V, von Marttens A, Acuña-Mardones P, Sanzana-Luengo C, Rueda-Velásquez SJ, Alvarado E, et al. Implementation of a teledentistry platform for dental emergencies for the elderly in the context of the COVID-19 pandemic in Chile. Biomed Res Int 2022;2022:6889285 [FREE Full text] [doi: 10.1155/2022/6889285] [Medline: 35330690]
48. Sanzana-Luengo C, Díaz L, Abarza L, Alvarado E, Beltrán V. Teledentistry protocol for the elderly in the context of the COVID-19 pandemic through a web platform/mobile app: approach from the general dentist. J Oral Res 2022;1:1-8 [FREE Full text] [doi: 10.17126/joralres.2022.026]
49. National Health Survey 2016-2017 report: oral health. Ministry of Health, Government of Chile. 2017. URL: https://epi.minsal.cl/wp-content/uploads/2021/03/Informe_Salud_Bucal_ENS_2016_17.pdf [accessed 2025-08-27]
50. Regional Diagnoses with a Social Determinants of Health Approach: La Araucanía Region. Santiago, Chile: Ministry of Health, Government of Chile; 2017.
51. Guidelines for the Social Household Registry No. 8: calculation of the socioeconomic classification. Ministry of Social Development, Government of Chile. URL: http://www.registrosocial.gob.cl/docs/Orientaciones-complementarias-N8_calculo-CSE_VF.pdf [accessed 2025-08-27]
52. Araucanía regional profile. Office of Agricultural Studies and Policies (ODEPA). 2021. URL: https://www.odepa.gob.cl/estadisticas-del-sector/ficha-nacional-y-regionales [accessed 2025-08-27]
53. Jepsen S, Blanco J, Buchalla W, Carvalho JC, Dietrich T, Dörfer C, et al. Prevention and control of dental caries and periodontal diseases at individual and population level: consensus report of group 3 of joint EFP/ORCA workshop on the boundaries between caries and periodontal diseases. J Clin Periodontol 2017;44 Suppl 18:S85-S93. [doi: 10.1111/jcpe.12687] [Medline: 28266120]
54. Chang Y, Woo HG, Park J, Lee JS, Song T. Improved oral hygiene care is associated with decreased risk of occurrence for atrial fibrillation and heart failure: a nationwide population-based cohort study. Eur J Prev Cardiol 2020;27(17):1835-1845. [doi: 10.1177/2047487319886018] [Medline: 31786965]
55. Ford PJ, Rich AM. Tobacco use and oral health. Addiction 2021;116(12):3531-3540. [doi: 10.1111/add.15513] [Medline: 33822437]
56. Oral health. World Health Organization. 2025. URL: https://tinyurl.com/3a56yajt [accessed 2025-08-27]
57. van der Putten G, de Baat C. An overview of systemic health factors related to rapid oral health deterioration among older people. J Clin Med 2023;12(13):4306 [FREE Full text] [doi: 10.3390/jcm12134306] [Medline: 37445340]
58. Fernandes MS, Castelo PM, Chaves GN, Fernandes JPS, Fonseca FLA, Zanato LE, et al. Relationship between polypharmacy, xerostomia, gustatory sensitivity, and swallowing complaints in the elderly: a multidisciplinary approach. J Texture Stud 2021;52(2):187-196. [doi: 10.1111/jtxs.12573] [Medline: 33191538]
59. Kanungo S, Ghosal S, Kerketta S, Sinha A, Mercer SW, Lee JT, et al. Association of oral health with multimorbidity among older adults: findings from the longitudinal ageing study in India, wave-1, 2017-2019. Int J Environ Res Public Health 2021;18(23):12853 [FREE Full text] [doi: 10.3390/ijerph182312853] [Medline: 34886581]
60. Cannon I, Robinson-Barella A, McLellan G, Ramsay SE. From drugs to dry mouth: a systematic review exploring oral and psychological health conditions associated with dry mouth in older adults with polypharmacy. Drugs Aging 2023;40(4):307-316. [doi: 10.1007/s40266-023-01017-5] [Medline: 36943673]
61. Skou ST, Mair FS, Fortin M, Guthrie B, Nunes BP, Miranda JJ, et al. Multimorbidity. Nat Rev Dis Primers 2022;8(1):48 [FREE Full text] [doi: 10.1038/s41572-022-00376-4] [Medline: 35835758]
62. Masnoon N, Shakib S, Kalisch-Ellett L, Caughey GE. What is polypharmacy? A systematic review of definitions. BMC Geriatr 2017;17(1):230 [FREE Full text] [doi: 10.1186/s12877-017-0621-2] [Medline: 29017448]
63. Belafsky PC, Mouadeb DA, Rees CJ, Pryor JC, Postma GN, Allen J, et al. Validity and reliability of the Eating Assessment Tool (EAT-10). Ann Otol Rhinol Laryngol 2008;117(12):919-924. [doi: 10.1177/000348940811701210] [Medline: 19140539]
64. Fernández-Rosati J, Lera L, Fuentes-López E, Albala C. alidez y confiabilidad del cuestionario Eating Assessment Tool 10 (EAT-10) para detectar disfagia en adultos mayores chilenos. Rev Med Chil 2018;146(9):1008-1015 [FREE Full text] [doi: 10.4067/s0034-98872018000901008] [Medline: 30725021]
65. Danke K, Bevensee I, Silva AB, Morales A, Jara G, Werlinger F, et al. Rehabilitation with removable dentures in adults and older people with partial or total edentulism: Chilean Public Health System Experience, 2017–2019. Arch Gerontol Geriatr Plus 2024;1(4):100084. [doi: 10.1016/j.aggp.2024.100084]
66. Grytten J, Skau I. Educational inequalities in access to fixed prosthodontic treatment in Norway. Causal effects using the introduction of a school reform as an instrumental variable. Soc Sci Med 2020;260:113105 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113105] [Medline: 32653812]

XSL•FO
RenderX

67. Borgeat Meza M, Espinoza I, Carvajal P, Cuevas R. Changes in oral health inequalities in adults in Chile. Community Dent Oral Epidemiol 2022;50(6):506-512. [doi: 10.1111/cdoe.12701] [Medline: 34713473]
68. Herrera D, Tonetti MS, Chapple I, Kebschull M, Papapanou PN, Sculean A, et al. Consensus report of the 20th European Workshop on Periodontology: contemporary and emerging technologies in periodontal diagnosis. J Clin Periodontol 2025;52 Suppl 29(Suppl 29):4-33 [FREE Full text] [doi: 10.1111/jcpe.14152] [Medline: 40320757]
69. Chapple ILC, Mealey BL, Van Dyke TE, Bartold PM, Dommisch H, Eickholz P, et al. Periodontal health and gingival diseases and conditions on an intact and a reduced periodontium: consensus report of workgroup 1 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. J Periodontol 2018;89 Suppl 1:S74-S84. [doi: 10.1002/JPER.17-0719] [Medline: 29926944]
70. Botelho J, Machado V, Proença L, Mendes JJ. The 2018 periodontitis case definition improves accuracy performance of full-mouth partial diagnostic protocols. Sci Rep 2020;10(1):7093 [FREE Full text] [doi: 10.1038/s41598-020-63700-6] [Medline: 32341429]
71. Redford M, Drury TF, Kingman A, Brown LJ. Denture use and the technical quality of dental prostheses among persons 18-74 years of age: United States, 1988-1991. J Dent Res 1996;75 Spec No:714-725. [doi: 10.1177/00220345960750S11] [Medline: 8594095]
72. Kapur KK. A clinical evaluation of denture adhesives. J Prosthet Dent 1967;18(6):550-558. [doi: 10.1016/0022-3913(67)90221-1] [Medline: 4863771]
73. Broadbent JM, Zeng J, Foster Page LA, Baker SR, Ramrakha S, Thomson WM. Oral health-related beliefs, behaviors, and outcomes through the life course. J Dent Res 2016;95(7):808-813 [FREE Full text] [doi: 10.1177/0022034516634663] [Medline: 26936215]
74. GBD Compare. IHME, University of Washington. URL: https://vizhub.healthdata.org/gbd-compare/ [accessed 2025-09-01]
75. Bonnefoy J, Morgan A, Kelly MP, Butt J, Bergman V. Constructing the evidence base on the social determinants of health: a guide. World Health Organization. URL: https://cdn.who.int/media/docs/default-source/documents/social-determinants-of-health/guide--measurement-and-evidence-knowledge-network-2007.pdf?sfvrsn=e54891cd_3 [accessed 2025-08-27]
76. Lawrence HP, Cidro J, Isaac-Mann S, Peressini S, Maar M, Schroth RJ, et al. Racism and oral health outcomes among pregnant Canadian Aboriginal women. J Health Care Poor Underserved 2016;27(1 Suppl):178-206. [doi: 10.1353/hpu.2016.0030] [Medline: 26853210]
77. Poirier BF, Soares G, Neufeld HT, Hedges J, Sethi S, Jamieson L. Conceptualising the relationships between food sovereignty, food security and oral health among global Indigenous communities: a scoping review. Public Health Nutr 2024;27(1):e147. [doi: 10.1017/S1368980024001198] [Medline: 38804085]
78. Durey A, Naylor N, Slack-Smith L. Inequalities between Aboriginal and non-Aboriginal Australians seen through the lens of oral health: time to change focus. Philos Trans R Soc Lond B Biol Sci 2023;378(1883):20220294 [FREE Full text] [doi: 10.1098/rstb.2022.0294] [Medline: 37381845]
79. Tan ECK, Lexomboon D, Sandborgh-Englund G, Haasum Y, Johnell K. Medications that cause dry mouth as an adverse effect in older people: a systematic review and metaanalysis. J Am Geriatr Soc 2018;66(1):76-84. [doi: 10.1111/jgs.15151] [Medline: 29071719]
80. Oral health Chile 2022 country profile. World Health Organization. URL: https://www.who.int/publications/m/item/oral-health-chl-2022-country-profile [accessed 2025-08-27]
81. Final report: in-depth evaluation of the Primary Health Strengthening Program. Ministry of Health, Government of Chile. URL: https://www.dipres.gob.cl/597/articles-141038_r_ejecutivo_institucional.pdf [accessed 2026-03-08]
82. Final evaluation report on the evaluation of government programs: More Smiles for Chile Program. Ministry of Health, Government of Chile. URL: https://www.dipres.gob.cl/597/articles-163116_informe_final.pdf [accessed 2026-03-08]
83. Ministry of Health, Government of Chile. Technical-administrative guidelines for the implementation of the comprehensive dental program 2019. Ministry of Health, Government of Chile. Primary Care Division. URL: https://diprece.minsal.cl/wp-content/uploads/2019/02/Orientación-Técnica-Programa-Odontológico-Integral-2019.pdf [accessed 2026-03-08]
84. National oral health plan 2021-2030. Ministry of Health, Government of Chile. URL: https://saludbucal.minsal.cl/wp-content/uploads/2022/06/PLAN-NACIONAL-DE-SALUD-BUCAL-2021-2030.pdf [accessed 2026-03-08]
85. Nascimento da Silva Mulder J, Ramos Pinto M, Aníbal I, Dornellas AP, Garrido D, Huanca C, et al. Teledentistry applied to health and education outcomes: evidence gap map. J Med Internet Res 2024;26:e60590 [FREE Full text] [doi: 10.2196/60590] [Medline: 39602783]
86. Siegel ZM, Quinkert E, Pai J, Miller CH, Lewis MW. Lessons learned about digital health tool acceptability among rural older adults: systematic review guided by the technology acceptance model. JMIR Aging 2026;9:e70012 [FREE Full text] [doi: 10.2196/70012] [Medline: 41637737]
87. Beltrán V, López L, Acuña-Mardones P, Silva R, Acevedo C, Bustos J, et al. A teledentistry care model for older populations in remote settings in Chile using satellite communication technology. Front Oral Health 2025;6:1699401. [doi: 10.3389/froh.2025.1699401] [Medline: 41424661]

XSL•FO
RenderX

88. Kengne Talla P, Allison P, Bussières A, Rodrigues A, Bergeron F, Giraudeau N, et al. Teledentistry for improving access to, and quality of oral health care: overview of systematic reviews and meta-analyses. J Med Internet Res 2025;27:e65211 [FREE Full text] [doi: 10.2196/65211] [Medline: 40737673]
89. Muñoz-Sepúlveda F, Beltrán V, von Marttens A, Navarro P, Acevedo C, Mariño R, et al. WhatsApp-supported teledentistry to reinforce oral health promotion among older adults residing in rural and urban areas: randomized controlled trial. JMIR Mhealth Uhealth 2026;14:e71251 [FREE Full text] [doi: 10.2196/71251] [Medline: 42139689]
90. Labrique A, Agarwal S, Tamrat T, Mehl G. WHO Digital Health Guidelines: a milestone for global health. NPJ Digit Med 2020;3:120. [doi: 10.1038/s41746-020-00330-2] [Medline: 33015373]
91. Recommendations on digital interventions for health system strengthening. World Health Organization. 2019 Jun 06. URL: https://www.who.int/publications/i/item/9789241550505 [accessed 2026-03-08]
92. Zou C, Harvard A, Qian J, Fox BI. A systematic review of digital health technologies for the care of older adults during COVID-19 pandemic. Digit Health 2023;9:20552076231191050 [FREE Full text] [doi: 10.1177/20552076231191050] [Medline: 37529545]
93. Transforming our world: the 2030 agenda for sustainable development. United Nations. 2015. URL: https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf [accessed 2025-09-01]
94. Goldstein SP, Tovar A, Espel-Huynh HM, Stowers KC. Applying a social determinants of health framework to guide digital innovations that reduce disparities in chronic disease. Psychosom Med 2023;85(7):659-669. [doi: 10.1097/PSY.0000000000001176] [Medline: 36800264]
95. Aravena-Rivas Y, Venturelli R, Stennett M, Tsakos G. Inequalities in dental services use by older adults in Chile according to eligibility for a national dental programme. Community Dent Oral Epidemiol 2024;52(2):161-170. [doi: 10.1111/cdoe.12909] [Medline: 37691001]
96. Zwick L, Schmitz N, Shojaa M. Oral health-related quality of life and depressive symptoms in adults: longitudinal associations of the English Longitudinal Study of Ageing (ELSA). BMC Oral Health 2023;23(1):1029 [FREE Full text] [doi: 10.1186/s12903-023-03722-4] [Medline: 38124069]

Abbreviations

CAL: clinical attachment level
CAPREDENA: National Defense Social Security Fund
CSDH: Commission on Social Determinants of Health
DIPRECA: Carabineros de Chile Social Security Directorate
DMFT: decayed, missing, and filled teeth
EAT-10: Eating Assessment Tool-10
FONASA: National Health Fund
GES: Explicit Health Guarantees
ICDAS: International Caries Detection and Assessment System
ISAPRE: private health insurance institution
NHANES III: Third National Health and Nutrition Examination Survey
OR: odds ratio
PD: probing depth
PRAIS: Comprehensive Health Care and Reparation Program for Victims of Human Rights Violations
RSH: Social Registry of Households
TEGO: Geriatric Dental Specialties Tele-platform
WHO: World Health Organization

Edited by F Puga; submitted 30.Sep.2025; peer-reviewed by AM Meldrum, P Dhawan; comments to author 19.Feb.2026; revised version received 11.May.2026; accepted 04.Jun.2026; published 29.Sep.2026.

Please cite as:
Muñoz-Sepúlveda F, Danke Hausdorf K, Beltrán Varas V, Acevedo Atala C, López L, Herrera A
Social Determinants and Oral Health Among Indigenous and Non-Indigenous Older Adults in La Araucanía, Chile: Cross-Sectional Study
JMIR Aging 2026;9:e85075
URL: https://aging.jmir.org/2026/1/e85075
doi:10.2196/85075
PMID:

XSL•FO
RenderX

©Fernanda Muñoz-Sepúlveda, Karen Danke Hausdorf, Víctor Beltrán Varas, Claudia Acevedo Atala, Leonardo López, Alejandro Herrera. Originally published in JMIR Aging (https://aging.jmir.org), 29.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Design Preferences for a Digital Advance Care Planning Tool to Support People Living With Dementia: International Stakeholder Survey and Mixed Methods Analysis

Ana Milošič, MA; Urška Smrke, PhD; Izidor Mlakar, PhD

HUMADEX Research Group, Faculty of Electrical Engineering and Computer Science, University of Maribor, Koroška cesta 46, Maribor, Slovenia

Corresponding Author:
Izidor Mlakar, PhD
HUMADEX Research Group, Faculty of Electrical Engineering and Computer Science, University of Maribor, Koroška cesta 46, Maribor, Slovenia

Abstract

Background: Dementia is a progressive, life-limiting condition in which decision-making capacity diminishes over time, yet advance care planning (ACP) remains predominantly framed as a documentation-focused end-of-life task. This narrow framing fails to address the iterative, stage-specific clinical needs—such as behavioral change, care transitions, and caregiver support—that, in addition to clinical symptoms, characterize the full dementia trajectory. Although digital ACP (dACP) tools offer potential for integrating ACP with early palliative care principles, no study has systematically derived stakeholder-informed design priorities spanning ACP content, caregiver functionality, health system interoperability, and review frequency within a multigroup framework.

Objective: This study aims to identify and prioritize stakeholder-defined components and digital features for dACP and to examine whether priorities differ across people with dementia, informal caregivers, and health care professionals.

Methods: A cross-sectional mixed methods online survey was administered via Prolific (January-February 2025) to 232 participants: 13 people with dementia, 106 informal caregivers, and 113 health care professionals. Participants rated 12 dementia-tailored ACP content elements and 10 caregiver-support features on a 7-point Likert scale, indicated preferred care-plan review frequency, and rated the importance of health system integration. Open-ended responses were analyzed using large language model (LLM)–assisted thematic analysis (Anthropic Claude Sonnet 4) with subsequent dual-coder manual verification. Dimensionality reduction was performed via principal component analysis (PCA; direct oblimin rotation); intergroup differences were assessed by one-way ANOVA, and associations among central variables were assessed by Spearman correlations.

Results: The final sample included 232 participants. PCA identified 2 interrelated ACP domains—Comprehensive Care Planning and End-of-Life Preparation and Active Disease Symptom Management—and a caregiver domain capturing Caregiver Stress Management. Across items, strategies for maintaining quality of life as dementia progresses received the highest ratings. Tips for managing challenging behaviors were the highest-rated caregiver-support feature. Participants also rated integration with existing health care systems as highly important and were in favor of context-aware, regular plan review. Qualitatively, decision-making support and guidance, communication/information sharing, and documentation/record-keeping were the most frequently cited primary purposes of a dACP tool across all groups.

Conclusions: Across stakeholder groups, priorities converged on a dACP tool centered on quality-of-life support, active symptom management, caregiver guidance, and health system interoperability. The significant interdependence of symptom-management, caregiver-support, and integration domains supports the design of a unified adaptive platform rather than discrete modules. These empirically derived priorities provide actionable design targets for an integrated dACP tool extending early palliative care in dementia from diagnosis onward.

(JMIR Aging 2026;9:e93323) doi:10.2196/93323

KEYWORDS

dementia; advance care planning; digital health; co-design; mixed methods

Introduction

Dementia is characterized by a progressive decline in decision-making capacity accompanied by evolving care needs, which narrows the window during which individuals can meaningfully participate in future care planning. In contrast to other chronic conditions, the cognitive trajectory of dementia extends the scope of relevant care planning beyond conventional advance directives to encompass the broader spectrum of palliative care needs [1-5]. Advance care planning (ACP) is

XSL•FO
RenderX

intended to align care delivery with patients' values and preferences [6]. However, in the context of dementia, ACP must additionally accommodate evolving symptomatology, the progressive transfer of decisional responsibility to caregivers, and iterative review as circumstances change [7]. The conventional framing of ACP as an end-of-life documentation exercise is therefore particularly limiting in dementia, given that its nonlinear disease trajectory requires people with dementia, family members, and health care professionals to make recurring decisions regarding safety, behavioral change, symptom burden, and the appropriate timing of supportive or palliative interventions [8]. Consequently, a dementia-specific digital ACP (dACP) tool may need to integrate conventional anticipatory planning with continuous, stage-appropriate, person-centered guidance [9,10].

ACP is widely recognized as central to ensuring that health care delivery reflects the wishes of patients and their families [11], and its relevance extends to any individual with a life-limiting illness rather than being restricted to end-of-life decision-making [12-15]. For people living with dementia, the limitations inherent to conventional, documentation-oriented ACP are compounded by the prolonged and uncertain disease trajectory and the progressive loss of communicative and decisional capacity [16], necessitating an approach capable of supporting iterative, values-based decision-making across the disease course rather than end-of-life preparation alone [15,17,18]. Thus, the focus must shift from prolonging life at any cost to enhancing the quality and meaning of the time remaining [19,20]. This perspective is consistent with the principles of early palliative care, as it encompasses comprehensive symptom management, psychosocial support, and quality-of-life care started at diagnosis and in parallel with life-prolonging measures [21-23]. The approach is well established in oncology but remains comparatively underdeveloped in dementia, where palliative input is typically introduced only during the terminal stages of illness [24-29]. Existing evidence on ACP in dementia indicates consistent improvements in process-related outcomes, although effects on quality of care and quality of life remain mixed [12,30,31], and dementia continues to be frequently misclassified as a normative feature of aging rather than a terminal condition, contributing to the under-recognition of supportive care needs [32]. Most existing interventions consequently prioritize end-of-life directives over the comprehensive, longitudinal support model established in oncology [28,33,34]. To date, no dementia-specific digital tool has been designed to address this continuum by integrating anticipatory planning, ongoing symptom management, and caregiver support.

Digital health technologies offer a potential mechanism for addressing this gap by enabling structured elicitation of patient values and preferences, facilitating communication with care teams, and prompting timely review as clinical circumstances evolve [35-38], while also accommodating disease progression through ongoing symptom assessment, care coordination, and sustained caregiver support [39-41]. However, existing digital tools remain largely fragmented and have generally been developed without systematic, stakeholder-informed prioritization of features, raising the risk that development efforts emphasize technical feasibility over end user value and usability [35,42,43]. Given the individualized and evolving nature of dementia care needs, co-creation involving people with dementia, informal caregivers, and health care professionals is essential, as top-down design approaches are unlikely to reflect the lived and shifting experience of cognitive decline [44-48]. Lived experience is, however, fundamental to effective palliative care and meaningful ACP [49,50]. To date, no study has systematically elicited and prioritized the full range of content areas and digital features required for a trajectory-spanning, dementia-specific dACP tool, a gap that the present study directly addresses. Moreover, despite significant developments in the field, few dementia-specific digital interventions integrate anticipatory planning, ongoing symptom management, caregiver support, and cross-setting communication within a single platform [24-26,28,30-35].

This study was conducted within the AI4HOPE project and used stakeholder input to inform early-stage requirements for a dementia-specific dACP tool. The project is developing a dACP tool intended to extend the early palliative care model established in oncology to dementia care across the full disease trajectory. Building on a prior evidence review and gap analysis, the study aimed to identify and prioritize the components and features that stakeholders consider essential to such a tool. Five research questions (RQs) guided this investigation:

- (RQ1) Which components do ACP stakeholders, and people with dementia in particular, identify as most critical?
- (RQ2) Which digital caregiver-support features should be prioritized?
- (RQ3) How do priorities differ across people with dementia, informal caregivers, and health care professionals?
- (RQ4) How do ACP, symptom management, and caregiver support relate to one another as components of an integrated design?
- (RQ5) Which health-system integration capabilities and review-frequency preferences are considered most critical for implementation?

Methods

Study Design

This study was conducted as part of the EU-funded project AI4HOPE [51], which seeks to improve the quality of life and care for people with dementia through the use of AI and digital technologies. The present work focuses on establishing design priorities through stakeholder input before tool development and implementation. We adapted a cross-sectional mixed methods design implemented as an online survey to identify and prioritize design requirements for a dementia-specific dACP tool. Quantitative survey items were used to assess the perceived importance of ACP content elements, caregiver-support features, preferred frequency of reviewing and updating the care plan, and the importance of integrating the tool with existing health care systems. Qualitative open-ended questions were included to capture stakeholders' perspectives on the primary purpose of the dACP tool, how the tool should support balancing the person's expressed wishes with caregiver decision-making as dementia progresses, and anticipated implementation challenges.

XSL•FO
RenderX

Quantitative and qualitative data were analyzed in parallel and integrated at the interpretation stage to derive actionable design priorities for tool development.

Participants and Recruitment

Data were collected via an online survey administered on the Prolific platform in January and February 2025. Recruitment targeted 3 stakeholder groups relevant to dementia-oriented digital ACP: adults with a self-reported dementia diagnosis, informal caregivers, and health care professionals with experience working with people with dementia and/or providing palliative care. Eligible Prolific users were invited to participate through the platform, and interested individuals accessed the study link and reviewed study information prior to beginning the survey. The survey required approximately 15 minutes to complete. Eligibility was set using Prolific and was group-specific:

- Health care professionals were eligible if they currently resided in one of the following countries: Ireland, Germany, France, Spain, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Iceland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, or Sweden, and if their employment sector within medicine/health care was doctor, nurse, psychologist, or social worker.
- Persons with dementia were eligible if they currently resided in one of the following countries: Ireland, Germany, France, Spain, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Iceland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, or Sweden and indicated yes to the prescreener “mild cognitive impairment/dementia.”
- Caregivers were eligible if they currently resided in one of the following countries: Ireland, Germany, France, Spain, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Iceland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, or Sweden and indicated yes to “informal/unpaid caring responsibilities.”

The final sample consists of 232 participants (118 male, 111 female, 2 other, and 1 individual who preferred not to disclose their sex) from various European countries and the United States. The caregiver sample consists of 106 individuals (56 male, 48 female, 2 other), whose length of role ranged from 0 to 2 years (n=72, 67.9%) to over 10 years (n=4, 3.8%). Notably, 30.2% (n=32) are caregivers to a person with dementia. To examine whether the heterogeneity of the caregiver sample affected the results, we conducted a sensitivity analysis comparing ratings between caregivers of people with dementia (n=32) and caregivers of people with other conditions. Statistically significant differences were observed for only one item across the full Caregiver Stress Management domain, supporting the decision to report the caregiver group as a whole; full between-group comparisons are provided in Table S1 in Multimedia Appendix 1. The health care professionals sample was made up of 113 individuals (47 male, 66 female), whose experience ranged from 0 to 2 years (n=44, 38.9%) to over 10 years (n=15, 13.3%). This group included physicians (n=42, 37.2%), nurses (n=38, 33.6%), psychologists (n=10, 8.8%), and social workers (n=10, 8.8%). A majority (n=73, 64.6%) of health care professionals reported working with patients diagnosed with dementia or those receiving palliative care. The sample of individuals with dementia consisted of 13 individuals (8 male, 4 female, 1 did not wish to answer), whose mean age was 34.46 (SD 14.39; min=18, max=61) years. They reported different dementia diagnoses, including frontotemporal dementia (n=1, 7.69%), poststroke dementia (n=1, 7.69%), Lewy body dementia (n=1, 7.69%), Parkinson disease (n=1, 7.69%), mild dementia (n=2, 15.38%), slight memory loss (n=1, 7.69%), or other cognitive impairments (n=2, 15.38%).

Ethical Considerations

After consultation with the Commission for Ethics in Research at the Faculty of Arts, University of Maribor, formal ethics approval was deemed unnecessary because the study posed negligible risk to participants, in accordance with the commission’s guidelines [52]. Accordingly, no formal ethics approval was sought. The study was conducted in accordance with the guidelines of the Commission for Ethics in Research at the Faculty of Arts, University of Maribor. Before beginning the online survey, participants received information about the study and provided informed consent. Participation was voluntary, and participants were free to decide whether to take part. No directly identifying information was collected. Open-ended survey responses were deidentified before qualitative analysis, and data confidentiality was maintained throughout the research process. Only deidentified responses were used for the qualitative analysis, including the preliminary, LLM-assisted development of the codebooks. No foreseeable physical or psychological harms were identified. Participants received £4.50 (£1=US $1.32 as of November 16, 2026) as compensation for completing the online survey.

Measures

Overview

The survey included quantitative rating items and open-ended questions designed to elicit stakeholder priorities for a dementia-specific dACP tool. Survey items were developed specifically for this study because existing ACP instruments did not capture the combination of dementia-specific ACP content and digital tool functionality required for early, iterative planning across the dementia trajectory. An expert panel within the AI4HOPE consortium (including neurology, palliative care, medicine, AI, and social science expertise) generated an initial pool of candidate items based on the project’s design goals for a dementia-specific dACP tool. The panel iteratively refined items for clarity, relevance to dementia care, and suitability for an online survey format, resulting in two quantitative item sets: (1) elements of ACP tailored to dementia and (2) digital features intended to support caregiver stress management.

XSL•FO
RenderX

ACP Content Elements (Primary Quantitative Outcome)

Participants rated the importance of 12 elements of an advance care plan tailored to people with dementia (eg, decisions about future medical treatments and interventions, appointment of a health care proxy or power of attorney, strategies for maintaining quality of life as the disease progresses). Items were rated on a 7-point Likert scale ranging from 1 (“not important”) to 7 (“essential”).

Caregiver-Support Features (Primary Quantitative Outcome)

Participants rated the importance of 10 potential dACP features intended to support caregivers in managing dementia-related stress (eg, tips for managing challenging behaviors associated with dementia; self-assessment tools for caregiver stress and burnout; connection to professional mental health support services). Items were rated on the same 7-point Likert scale (1=not important to 7=essential).

Review/Update Preferences and Integration (Secondary Quantitative Outcomes)

Participants indicated how often the dACP tool should prompt review and potential updates of the care plan using a 5-option response scale (every 3 months, every 6 months, annually, only when prompted by the patient or caregiver, other). Participants also rated the importance of integrating the dACP tool with existing health care systems on a 7-point Likert scale (1=not important to 7=essential).

Open-Ended Questions (Qualitative Outcomes)

At the beginning of the survey, participants answered an open-ended question on what they believed should be the primary purpose of the dACP tool. Moreover, participants were also asked to provide open-ended responses regarding (1) how the dACP tool should balance respecting the person with dementia’s expressed wishes with the need for caregiver decision-making as the disease progresses and (2) the biggest anticipated challenges in implementing a dACP tool for people with dementia.

Quantitative Analysis

Quantitative analyses were conducted in SPSS (version 29; IBM Corp). After data cleaning, we included all respondents who met eligibility criteria and completed the survey, resulting in a final analytic sample with no missing data. We summarized item distributions using descriptive statistics and examined associations among central variables using correlation coefficients selected based on distributional assumptions.

For dimensionality reduction, we conducted principal component analyses (PCA) separately for (1) ACP content items and (2) caregiver-support feature items, using oblique rotation (direct oblimin) given the conceptual relatedness of domains. Sampling adequacy was evaluated using the Kaiser-Meyer-Olkin (KMO) statistic and the Bartlett test of sphericity. The number of components retained was guided by eigenvalues of greater than 1 and inspection of the scree plot, and components were interpreted based on rotated loadings. Component scores were computed as mean component scores based on item loadings and used for subsequent analyses.

Group differences between people with dementia, caregivers, and health care professionals were examined using ANOVA. For update frequency, we excluded responses that could not be ordered on the provided scale (eg, “only when prompted” and “other”) from ANOVA, as in the primary analysis plan.

Qualitative Analysis

We analyzed responses to the open-ended survey questions using a data-driven thematic analysis approach to identify recurring needs, expectations, and concerns regarding a dementia-specific dACP tool. Thematic analysis was selected because it is well suited to concise, heterogeneous survey responses and does not presuppose extended narratives required by other qualitative approaches [53]. An inductive orientation was adopted given the absence of a validated coding framework for stakeholder perspectives on dementia-specific dACP design, allowing priorities to emerge directly from the data. This approach also permits systematic categorization and descriptive frequency quantification necessary for cross-group comparison. LLM-assisted draft codebook generation was incorporated to manage response volume (N=232 across 3 open-ended questions), while dual-coder manual verification ensured that all interpretive decisions remained with human researchers, in line with emerging guidance on responsible LLM use in qualitative health research [54].

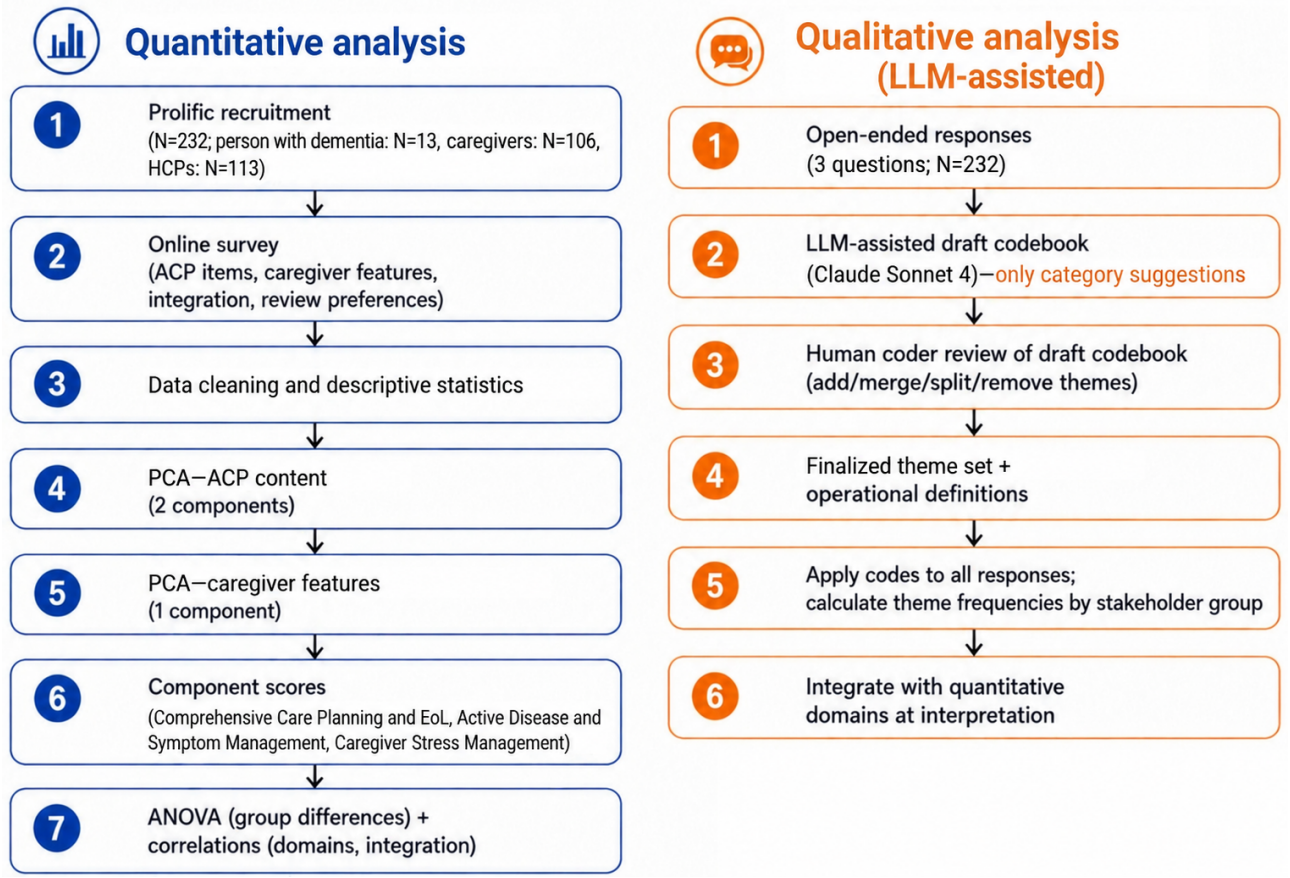
All open-ended responses were exported, deidentified, and cleaned (removal of empty responses and nonsubstantive entries). Responses to each of the 3 open-ended questions were submitted separately to Claude Sonnet 4 (Anthropic) [55] using a structured, constrained prompt. The prompt instructed the model to (1) read all responses to the given question, (2) identify recurring conceptual categories, (3) produce a concise label and brief operational definition for each candidate theme, and (4) provide illustrative groupings of similar response excerpts per theme. The LLM was explicitly instructed not to interpret, generalize beyond, or paraphrase the source responses. Its output was treated solely as a preliminary draft codebook and was used as a structured synthesis to accelerate initial organization of a large response set. No analytic claims were drawn from it directly [56].

Afterward, a researcher (AM or IM) reviewed the LLM-generated draft codebook against the full set of original participant responses for each question. They applied iterative edits to the codebook, including removing themes not sufficiently grounded in the data, merging conceptually overlapping themes, splitting overly broad themes into more specific categories, and adding themes absent from or underspecified in the draft. After independent review, coders met to reconcile disagreements through discussion and consensus, producing a finalized theme set with agreed operational definitions. The wording of theme labels was further refined for alignment with established terminology in ACP and palliative dementia care (eg, distinguishing documentation and legal functions from decision-support functions, distinguishing usability barriers from disease-progression-related adaptation requirements). The finalized codebook was then applied to all responses, and descriptive frequencies (proportion of respondents per stakeholder group mapping to each theme) were

XSL•FO
RenderX

calculated to support cross-group comparison. The full theme set is provided in Multimedia Appendix 1. Themes are reported in the *Results* section as the most frequently occurring categories per question and stakeholder group, with the full set of themes provided in Multimedia Appendix 1. Visual presentation of analysis workflow can be seen in Figure 1.

Figure 1. Study and mixed-methods analysis workflow. ACP: advance care planning; EoL: end of life; HCP: health care professional; LLM, large language model; PCA: principal component analysis.



Results

Reducing the Dimensionality of Questionnaire Items

Most Important Elements of an Advance Care Plan for People With Dementia

A PCA was conducted on 11 items with oblique rotation (direct oblimin) to reduce the number of responses regarding the most important elements of ACP for individuals with dementia into main components. The KMO measure confirmed the adequacy of the sample for the analysis (KMO=0.83). The Bartlett test of sphericity was significant (χ^2_{55}=596.11; P<.001), indicating that correlations between items were sufficient for component extraction.

An initial examination of data was done to determine the number of main components present. Using Kaiser's criterion (retaining components with eigenvalues >1), 2 components were extracted, together accounting for 46.83% of the total variance. The scree plot further supported a 2-component solution, which was also easily interpretable. Rotated component loadings are presented in Table 1. Component 1 includes items related to Comprehensive Care Planning and End-of-Life Preparation referring to long-term, anticipatory planning, such as values, legal arrangements, and end-of-life preferences, and component 2 includes items related to Active Disease and Symptom Management referring to active, ongoing, day-to-day clinical management of symptoms and crises.

XSL•FO
RenderX

Table . Rotated component loadings for most important elements of advance care planning (ACP).

Elements of ACP	Rotated component loadings	
	Comprehensive Care Planning and End-of-Life Preparation	Active Disease and Symptom Management
1. Decisions about future medical treatments and interventions	0.58[a]	0.07
2. Appointment of health care proxy or power of attorney	0.66[a]	–0.03
3. Strategies for maintaining quality of life as the disease progresses	0.06	0.56[a]
4. Monitoring and symptom management	–0.02	0.77[a]
5. Management of behavioral symptoms associated with dementia	0.003	0.78[a]
6. Crisis prevention and management strategies	–0.01	0.74[a]
7. Guidance on when to transition to palliative care	0.65[a]	0.07
8. Wishes for memory preservation and legacy (eg, recording life stories)	0.51[a]	–0.02
9. Preferences for end-of-life care	0.66[a]	0.16
10. Financial and legal planning considerations	0.70[a]	0.04
11. Cultural or spiritual preferences	0.75[a]	–0.15

[a]This value indicates the highest rotated component loading for each item, identifying the component with which that item is most strongly associated.

Features to Assist Caregivers in Managing Stress

A separate PCA was conducted on 10 items concerning features that may assist caregivers in managing stress, also using oblique rotation (direct oblimin). The KMO measure indicated sampling adequacy (KMO=0.89), and the Bartlett test of sphericity was significant (χ^2_{45}=823.08; P<.001), confirming the suitability of data for component extraction. According to Kaiser's criterion, one component, which accounted for 46.32% of the variance, was identified. Rotated component loadings are shown in Table 2. The 10 items loaded on a single component, labeled Caregiver Stress Management referring, which encompasses features designed to help caregivers manage stress, including access to professional mental health support, educational resources, and opportunities to build support networks.

Table . Rotated component loadings for assisting the management of caregiver stress.

Caregiver stress management features	Rotated component loadings for caregiver stress management
1. Self-assessment tools for caregiver stress and burnout	0.65
2. Personalized stress management techniques and resources	0.75
3. Information on respite care options and how to access them	0.70
4. Guided meditation or relaxation exercises	0.62
5. Tips for managing challenging behaviors associated with dementia	0.61
6. Resources for building a support network (eg, support groups, community services)	0.66
7. Time management and organization tools for caregiving tasks	0.64
8. Educational content on self-care and setting boundaries	0.71
9. Connection to professional mental health support services	0.70
10. Reminders for caregiver's own health check-ups and appointments	0.75

Comparisons of Central Variables Among Caregivers, Health Care Professionals, and People With Dementia

ANOVA was used to assess group differences across central variables, and the results are presented in Table 3.

XSL•FO
RenderX

Table . Mean and SD values for central variables, and ANOVA results.

Variable	Caregivers, mean (SD)	Health care professionals, mean (SD)	People with dementia, mean (SD)	ANOVA	
				F test (*df*)	η^2
Comprehensive Care Planning and End-of-Life Preparation[a]	3.45 (0.59)	3.48 (0.58)	3.60 (0.43)	0.40 (2, 229)	0.004
Active Disease and Symptom Management[b]	4.33 (0.46)	4.33 (0.58)	4.13 (0.56)	0.85 (2, 229)	0.007
Caregiver Stress Management[c]	3.74 (0.59)	3.76 (0.61)	3.80 (0.74)	0.07 (2, 229)	0.001
Importance of integration	6.12 (0.95)	6.18 (0.97)	6 (1.47)	0.22 (2, 229)	0.002
Frequency of review and updates[d]	1.55 (0.61)	1.55 (0.58)	1.5 (0.52)	0.04 (2, 229)	<0.001

[a]Min=1.45, max=4.52.

[b]Min=1.95, max=4.98.

[c]Min=1.07, max=4.75.

[d]Analysis was performed on a smaller sample (n=193), as values 4 (“only when prompted by patient or caregiver,” n=21) and 5 (“other,” n=18) were excluded from ANOVA. Mean component scores were calculated using the following equation: component mean score = ((item 1 × component loading 1) + (item 2 × component loading 2) + …(item n × component loading n))/n.

ANOVA indicated no statistically significant differences among group means; therefore, further analyses were done on the full sample. Mean and SD values for all items by group can be found in Table S2 (Multimedia Appendix 1).

Within the Comprehensive Care Planning and End-of-Life Preparation component, the item “Decisions about future medical treatments and interventions” was recognized as the most important (mean 5.84, SD 1.26), while the item “Cultural or spiritual preferences” was the least important (mean *M* 4.80, SD 1.49). In the Active Disease and Symptom Management component, the item rated as the most important (mean 6.32, SD 1.01) is “Strategies for maintaining quality of life as the disease progresses,” while “Crisis prevention and management strategies” was least important (mean 5.90, SD 1.10). In the Caregiver Stress Management component, the item “Tips for managing challenging behaviors associated with dementia” was recognized as the most important (mean 6.04, SD 1.13), while the item “Guided meditation or relaxation exercises” was the least important (mean 4.92, SD 1.53; see Table S1 for specific mean values in Multimedia Appendix 1).

Associations Between Central Variables

For the purpose of assessing the strength and direction of association between central variables, correlation coefficients were calculated (presented below in Table 4).

Table . Correlations between central variables.

Variable	1	2	3	4	5
Comprehensive Care Planning and End-of-Life Preparation	—[a]	—	—	—	—
Active Disease and Symptom Management	0.43[b]	—	—	—	—
Caregiver Stress Management	0.51[b]	0.52[b]	—	—	—
Importance of integration	0.06	0.24[b]	0.23[b]	—	—
Frequency of review and updates	–0.05	0.05	–0.05	0.06	—

[a]Not applicable.

[b]*P*<.001.

The Comprehensive Care Planning and End-of-Life Preparation component is statistically significantly positively correlated with both the Active Disease and Symptom Management component and the Caregiver Stress Management component. Additionally, the Active Disease and Symptom Management component is statistically significantly correlated with Caregiver

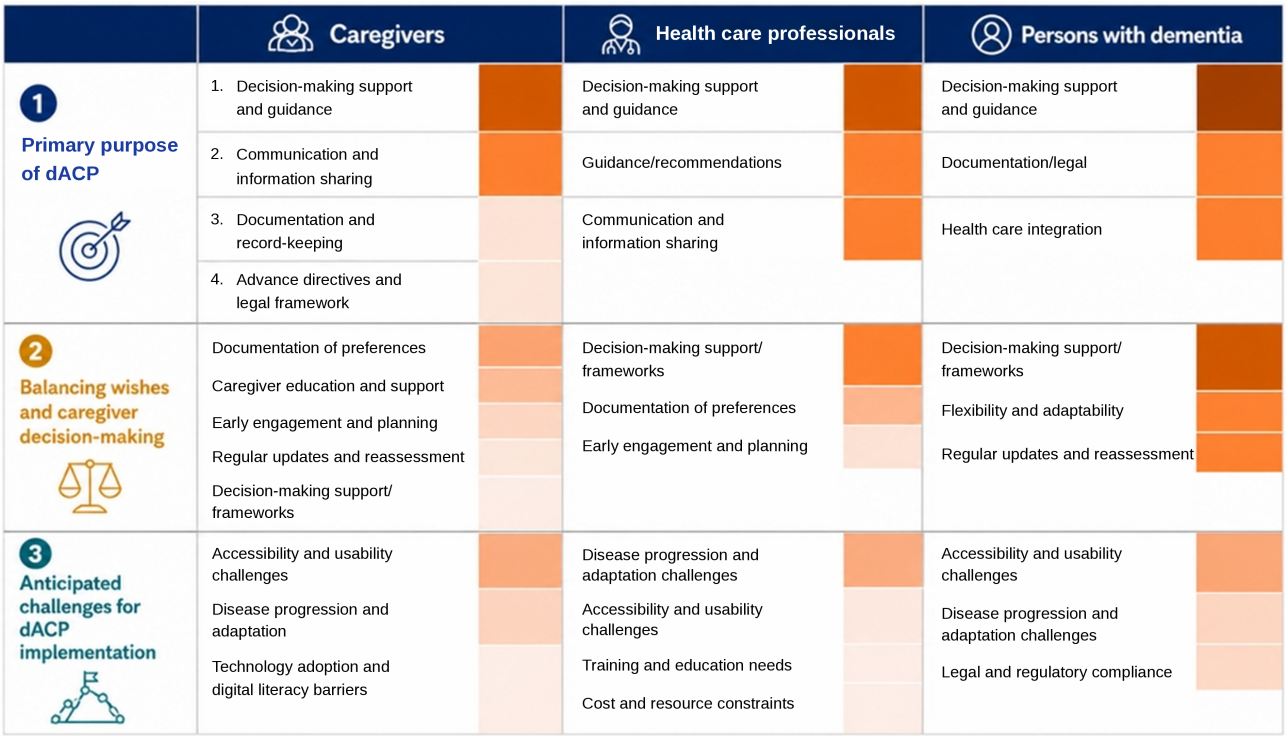
XSL•FO
RenderX

Stress Management and Importance of integration of dACP in the health care system. Finally, Caregiver Stress Management is also statistically significantly correlated with the Importance of integration variable.

Exploring Qualitative Responses on dACP

Answers of the 3 questions were categorized into 8 to 25 main themes for each of the questions. The most prominent 3 (or more, in cases of same frequency of mentions of the third) themes identified are reported in this subsection (all extracted themes can be found in Table S3 in Multimedia Appendix 1). For visual presentation of the main results, see Figure 2.

Figure 2. Stakeholder qualitative themes on purpose, balancing wishes, and anticipated challenges. Darker shading indicates higher mention frequency within each stakeholder group. dACP: digital advance care planning.



Primary Purpose of dACP

As the primary purpose of dACP, caregivers identified Decision-Making Support and Guidance (n=54, 50.94%), which refers to guiding care choices for both patients and caregivers based on the individual's preferences and needs; Communication and Information Sharing (n=38, 35.85%), referring to enabling sharing capabilities with family members and health care providers, as well as facilitating information exchange within the health care system; Documentation and Record-Keeping (n=21, 19.81%), referring to automated and intelligent filing of documents to maintain organized records of care preferences and plans; and Advanced Directives and Legal Framework (n=21, 19.81%), referring to creating and managing legally binding advanced directives that formalize the patient's wishes for future care.

Health care providers most prominently identified Decision-Making Support and Guidance (n=61, 54.46%), referring to guiding care choices for health care professionals and caregivers based on the patient's individual preferences; Guidance/Recommendations (n=52, 46.43%), which refers to providing care recommendations and guidance for managing daily activities, pain assessment, medication reminders, and symptom monitoring; and Communication and Information Sharing (n=42, 37.5%), referring to sharing capabilities including advanced directives, facilitating general communication between patients, health care professionals, and caregivers to ensure everyone is informed.

Finally, persons with dementia also identified Decision-Making Support and Guidance (n=10, 76.92%), referring to providing support and guidance to help individuals make informed decisions about care; Documentation/Legal (n=6, 46.15%), referring to managing advanced directives and legal documentation of care preferences in a formalized manner; and Health Care Integration (n=5, 38.46%), which refers to enabling information sharing within the health care system and coordinating with health care providers for comprehensive medical care management.

Balancing the Expressed Wishes of Patients With Dementia With the Need for Caregiver Decision-Making

When it comes to balancing patient wishes with caregiver decision-making, caregivers have identified the following themes as most prominent: Documentation of Preferences (n=29, 27.36%), referring to clear documentation of all relevant patient information and preferences within the tool to ensure nothing is forgotten or overlooked; Caregiver Education and Support (n=26, 23.21%), referring to providing caregivers with guidance through the tool's repository of patient preferences, reminding them of the patient's wishes, and offering prompts for consulting

XSL•FO
RenderX

with health care professionals; Early Engagement and Planning (n=21, 19.81%), referring to initiation of ACP discussions and documentation early in the disease process; Regular Updates and Reassessment (n=21, 19.81%), referring to continuous monitoring of the patient's condition, tracking disease progression, and prompting updates to care plans as needed through reminders and adjustments; and Decision-Making Support/Frameworks (n=21, 19.81%), referring to implementation of flexible frameworks such as gradual transfer of decision-making authority from patient to caregiver as the disease progresses, using weighted approaches that consider multiple factors.

Health care professionals most prominently identified Decision-Making Support/Frameworks (n=42, 37.5%), referring to balancing different perspectives with patient wishes as the priority while not limiting necessary care, providing alternatives and recommendations that weigh pros and cons using algorithms based on human and patient rights, and distinguishing between essential end-of-life needs vs preferences that can only be respected if they do not conflict with care needs; Documentation of Preferences (n=23, 20.54%), referring to sharing documented patient wishes and key information with caregivers and health care professionals to ensure everyone is aligned and informed; and Early Engagement and Planning (n=19, 16.96%), which refers to encouraging early conversations and relying primarily on wishes expressed before progression of the disease, documenting what the patient stated when still cognitively able to do so.

Finally, persons with dementia also identified Decision-Making Support/Frameworks (n=7, 53.85%), referring to providing frameworks that help both patients and caregivers understand their positions and roles at any given time as the disease progresses; Flexibility and Adaptability (n=4, 30.77%), which refers to maintaining adaptability in the tool's approach to accommodate changing circumstances and individual patient needs; and Regular Updates and Reassessment (n=4, 30.77%), referring to ensuring continuous monitoring and assessment of the patient's condition and care preferences over time.

Biggest Anticipated Challenges in Implementing a dACP Tool

As the biggest anticipated challenges in implementing a dACP tool, caregivers most prominently identified Accessibility and Usability Challenges (n=36, 33.96%), referring to difficulties with understanding the tool in later disease stages, adapting it to individual patients, and overcoming basic technological unfamiliarity, particularly for patients who may struggle with using technology; Disease Progression and Adaptation Challenges (n=26, 23.21%), referring to the tool having to adapt as cognitive status changes over time and need for quick implementation in order for the tool to be useful throughout the disease trajectory; and Technology Adoption and Digital Literacy Barriers (n=20, 18.87%), which refers to overall user adoption challenges, including the need for training and support, particularly given caregiver burden concerns.

Health care professionals identified Disease Progression and Adaptation Challenges (n=25, 22.32%), referring to adaptation to disease progression, keeping documentation updated, predicting future problems and decisions, and ensuring rapid implementation before cognitive abilities decline significantly; Accessibility and Usability Challenges (n=22, 19.64%), which refers to ensuring the tool works for diverse patient groups with varying needs, is easy to use, and doesn't create barriers where some people cannot use it effectively; Training and Education Needs (n=20, 17.86%), referring to providing adequate training and education for all users to effectively utilize the tool; and Cost and Resource Constraints (n=20, 17.86%), referring to managing time as a resource, implementation costs, and the human resources needed for development and ongoing support.

Lastly, persons with dementia also identified Accessibility and Usability Challenges (n=6, 46.15%), which refers to ensuring appropriate system design and clear communication that is suitable for individuals experiencing cognitive decline; Disease Progression and Adaptation Challenges (n=4, 30.77%), which addresses how the tool adapts as cognitive abilities decline over time; and Legal and Regulatory Compliance (n=4, 30.77%), referring to navigating legal and ethical concerns, along with regulatory requirements for such a health care tool.

Discussion

Principal Findings

The presented mixed methods study identified stakeholder-informed priorities for a dementia-specific dACP tool that supports ACP as a longitudinal, person-centered process. Across quantitative and qualitative findings, stakeholders prioritized maintenance of quality of life, decision-making support, caregiver guidance, and integration with health care systems. These priorities are consistent with concerns that conventional ACP can become overly focused on documentation and treatment decisions, rather than supporting values-based decisions about living with serious illness [12,14,57,58].

The findings also require cautious interpretation in relation to the group of people with dementia. This subgroup comprised 13 self-identified participants, with a mean age of 34.46 (SD 14.39; range 18 - 61) years and heterogeneous self-reported diagnoses, including both dementia subtypes and broader cognitive-concern labels. The age profile differs substantially from the epidemiology of even young-onset dementia, for which prevalence increases sharply across adulthood, from approximately 1.1 per 100,000 among adults aged 30 to 34 years to 77.4 per 100,000 among those aged 60 to 64 years [59]. Confirmed young-onset dementia cohorts also report a substantially later mean age at diagnosis and diagnostic delays of approximately 3.6 years [60]. Together with recruitment through a Prolific prescreener rather than clinical confirmation, these characteristics limit the subgroup's representativeness of clinically diagnosed dementia populations. The findings from this group should therefore be considered hypothesis-generating, rather than confirmatory.

ACP and Caregiver Priorities

With respect to RQ1, which examined stakeholders' priorities for comprehensive ACP in dementia, the ACP PCA identified 2 related domains, Comprehensive Care Planning and

XSL•FO
RenderX

End-of-Life Preparation and Active Disease and Symptom Management, which together explained 46.83% of variance. The highest-rated item was maintaining quality of life as dementia progressed (mean 6.32, SD 1.01), followed by decisions about future medical treatments and interventions (mean 5.84, SD 1.26). This pattern indicates that stakeholders considered ACP in dementia to encompass both prospective planning and support for day-to-day clinical needs, including symptoms, behavioral changes, and crises.

The qualitative findings further reinforced this interpretation. Decision-making support and guidance was the most frequently identified primary purpose of a dACP tool among caregivers (n=54, 50.94%), health care professionals (n = 61, 54.46%), and people with dementia (n=10, 76.92%). Communication and information sharing was also prominent among caregivers (n=38, 35.85%) and professionals (n=42, 37.50%), whereas people with dementia additionally identified documentation and legal functions (n=6, 46.15%) and health care integration (n=5, 38.46%) as central. These findings support a view of ACP as an iterative, capacity-sensitive communication process involving the person, family caregivers, and health care professionals, rather than a one-time advance-directive exercise [9,12,15-18,57,58].

Regarding RQ2, which examined priorities for digital caregiver support, caregiver-support functions formed one Caregiver Stress Management domain, accounting for 46.32% of variance. Guidance for managing challenging behaviors was the highest-rated caregiver-support feature (mean 6.04, SD 1.13). When considering how a dACP tool should balance the person's wishes with increasing caregiver responsibilities, caregivers emphasized documentation of preferences (n=29, 27.36%); education and support (n=26, 23.21%); and early engagement, regular reassessment, and decision-making frameworks (each n=21, 19.81%). These results indicate that caregivers need support not only for managing dementia-related symptoms but also for accessing and interpreting the person's values as responsibility for decisions evolves. This is aligned with palliative-care models that include caregiver well-being, psychosocial support, and shared decision-making as core elements of care [15,21,28,36,37].

Stakeholder Perspectives and Domains

For RQ3, which examined differences in priorities across people with dementia, caregivers, and health care professionals, ANOVA identified no statistically significant differences between groups on the central quantitative variables. However, this should not be interpreted as definitive evidence of stakeholder agreement. The small and demographically atypical dementia subgroup reduced statistical power to detect between-group differences and limits the generalizability of subgroup findings. The observed overlap in priorities should therefore be interpreted as preliminary, providing candidate shared priorities for subsequent co-design and validation in larger, clinically verified, and more balanced samples.

Addressing RQ4, which examined relationships among ACP, symptom management, caregiver support, and integration, the correlation analyses indicated that stakeholder priorities spanned connected care domains. Comprehensive Care Planning and End-of-Life Preparation was positively associated with Active Disease and Symptom Management (r=0.43) and Caregiver Stress Management (r=.51). Active Disease and Symptom Management was associated with Caregiver Stress Management (r=0.52) and perceived importance of health system integration (r=.24). Caregiver Stress Management was also associated with integration importance (r=0.23; all P<.01). However, these results do not establish causal relationships, but they do suggest that stakeholders regarded prospective planning, current symptom needs, caregiver capacity, and integration with services as interdependent elements of dementia care.

This interpretation is further supported by qualitative findings. When discussing changing decision-making responsibilities, people with dementia most often emphasized structured decision-making frameworks (n=7, 53.85%), flexibility and adaptability (n=4, 30.77%), and regular updates and reassessment (n=4, 30.77%). Health care professionals similarly prioritized decision-making frameworks (n=42, 37.50%), documentation of preferences (n=23, 20.54%), and early engagement and planning (n=19, 16.96%). These findings highlight that a clinically useful ACP process must preserve the person's expressed values while enabling transparent, context-sensitive involvement of caregivers and professionals as decisional capacity and care needs evolve [15,16,18,20,57,61,62].

Implementation and Review

Finally, addressing RQ5, which examined integration and care-plan review preferences, health system integration was rated highly important across all groups, with means ranging from 6.00 (SD 1.47) among people with dementia to 6.18 (SD 0.97) among health care professionals. Integration importance was associated with both symptom-management and caregiver-support domains, suggesting that stakeholders considered an isolated tool insufficient for supporting dementia care. Participants also preferred regular review, with mean scores of 1.50 to 1.55 (SD 0.52-0.61) on the ordered review-frequency scale, and regular updates and reassessment was a prominent qualitative theme.

Implementation barriers reflected the requirements of an aging population with progressive cognitive impairment. Accessibility and usability concerns were raised by 33.96% (n=36) of the caregivers, 19.64% (n=22) of the health care professionals, and 46.15% (n=6) of the people with dementia. Disease-progression and adaptation concerns were identified by 23.21% (n=26), 22.32% (n=25), and 30.77% (n=4), respectively. These findings emphasize that digital ACP requires more than electronic preference documentation: it must remain usable as cognitive abilities change; be understandable to caregivers; and fit care pathways that span home, community, primary, specialist, and long-term care settings [15,20,24-26,35,61-63].

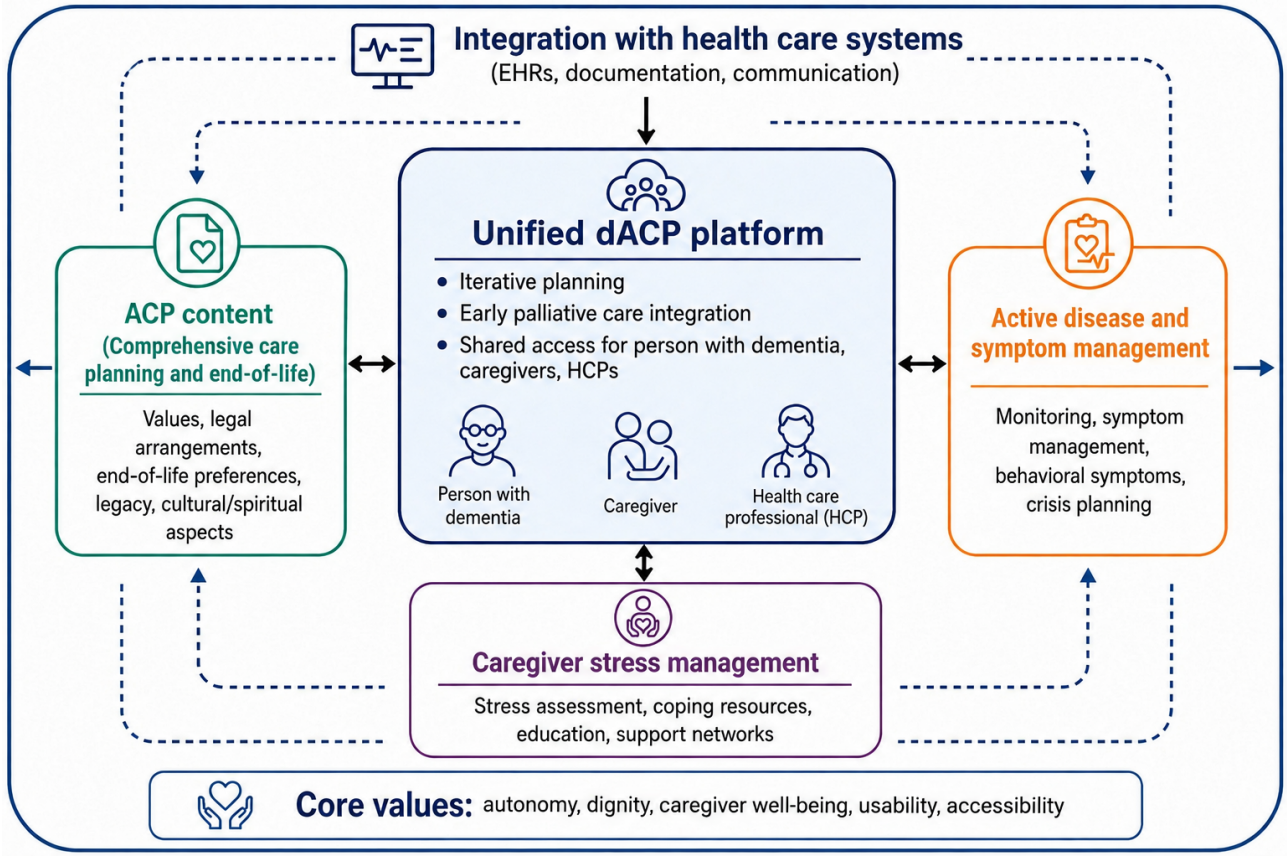
Interoperability should be treated as a staged implementation objective, initially prioritizing secure, role-appropriate access to current preferences and later addressing structured exchange and governance within local clinical workflows. Although HL7 FHIR offers a relevant technical standard, successful integration also depends on clear responsibility for reviewing preferences at clinically meaningful transitions [61,63-68].

XSL•FO
RenderX

Implications for Design

The results support a value-sensitive approach to dACP development, as value-sensitive design combines conceptual investigation of values and stakeholders, empirical investigation of user experiences, and technical translation into system features [69,70]. Based on the facilitated methodology, this study primarily focuses on the conceptual and empirical phases by identifying values and needs that should guide later co-design, prototyping, usability testing, and implementation research. Figure 3 summarizes the resulting conceptual design priorities.

Figure 3. Conceptual design priorities for unified dementia-specific digital advance care planning (dACP) platform. ACP: advance care planning; EHR: electronic health record; HCP: health care professional.



As outlined in Figure 3, the main identified conceptual design priorities include the following:

- Autonomy and self-determination: Enable values clarification and iterative decision support, not only static preference recording [9,12,14-16].
- Dignity and continuity: Link current symptom and behavioral needs with care preferences across disease progression [15,16,18,27,57,62].
- Caregiver well-being: Include practical behavioral guidance, caregiver stress support, and access to professional resources as core functions [15,21,28,36,37].
- Accessibility: Use adaptive, low-burden interactions that accommodate cognitive decline and differing digital literacy [35,40,44,45].
- Proxy accountability: Support traceable preference updates, decision rationales, and transparent transitions in decision-making responsibility [14,16,63].
- Care continuity: Enable structured sharing of current preferences across care settings and professionals [24-26,35,61,63].
- Iterative review: Trigger review after clinically or socially meaningful change rather than relying only on fixed schedules [15,20,61].

Limitations

Several limitations require consideration when interpreting our findings. The people-with-dementia subgroup was small (n=13), younger than expected for a clinically representative dementia population, and recruited on the basis of a self-reported Prolific prescreener. Therefore, it may include individuals with self-identified cognitive concerns rather than clinically verified dementia, and the findings from this subgroup should not be generalized to the broader dementia population [59,60,71-73]. The caregiver sample was also heterogeneous; however, only one of 10 Caregiver Stress Management items differed significantly between caregivers of people with dementia and caregivers of people with other conditions, suggesting that pooling caregivers did not materially alter the central design priorities.

The international sample did not ensure broad cultural representation, and the cross-sectional design captured preferences at one time point despite the changing needs associated with dementia. Future studies should use longitudinal, culturally diverse, and clinically based recruitment pathways, with larger and more balanced stakeholder samples [15,20,47,62]. The study-specific measures require further

XSL•FO
RenderX

psychometric evaluation, and their relation to clinical, caregiver, and implementation outcomes remains uncertain [31].

Finally, exclusive online recruitment may have underrepresented people with advanced dementia, limited digital access, lower digital literacy, or substantial caregiver burden. Online self-selection may also have favored people with prior interest in ACP, potentially increasing importance ratings. Although Prolific can support high-quality data collection, data quality depends on validation procedures and study design [71-73]. The LLM was restricted to draft-codebook generation, with all coding, theme refinement, and interpretation performed by human researchers; nevertheless, LLM-assisted qualitative methods can reproduce biases or introduce data-distant conceptual framings [54]. Future studies should assess real-world acceptability, equity, clinical workflow impact, and cost-effectiveness once dACP is implemented in routine care [33,34].

Conclusions

This mixed methods study identified stakeholder-informed priorities for a dementia-specific dACP tool centered on quality-of-life maintenance, decision-making support, caregiver guidance, health-system integration, and context-sensitive review. The findings support a longitudinal, person-centered approach that complements communication among people with dementia, caregivers, and health care professionals. However, because the dementia subgroup was small, demographically atypical, and based on self-reported eligibility, apparent cross-group convergence should be interpreted as preliminary. Future work should use clinically verified and more representative samples to co-design, test, and evaluate the feasibility, equity, implementation, and clinical impact of dACP in routine dementia care.

Generative AI Use

A large language model was used to assist with rapid synthesis and to propose candidate themes, but final coding decisions, theme boundaries, and interpretation were determined by the research team through manual review of the original responses. In addition, the GPT-5.2 (OpenAI) and Gemini 3 (Google) were used for language editing (eg, grammar and clarity improvements) of manuscript text. All information generated by the models was thoroughly reviewed by the authors.

Funding

This work was funded by (1) the European Union Horizon Europe Research and Innovation Program project AI4HOPE (grant number 101136769) and (2) the Slovenian Research and Innovation Agency, Advanced methods of interaction in telecommunication research program (grant number P2-0069). The content of this work does not reflect the official position of the European Union or any other institution. The information and views expressed are the sole responsibility of the authors.

Data Availability

The datasets generated and/or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: IM
Data curation: AM, IM
Formal analysis: AM, US
Funding acquisition: IM
Methodology: IM, US, AM
Supervision: US
Writing – original draft: IM, US, AM
Writing – review & editing: AM, US, IM

Conflicts of Interest

None declared.

Multimedia Appendix 1
Caregiver subgroup comparisons, descriptive statistics by stakeholder group and for the full sample, correlations among study variables, and the complete list of qualitative themes with frequencies by stakeholder group and question.
[DOCX File, 64 KB - aging_v9i1e93323_app1.docx]

References

1. Volicer L, Simard J. Palliative care and quality of life for people with dementia: medical and psychosocial interventions. Int Psychogeriatr 2015 Oct;27(10):1623-1634. [doi: 10.1017/S1041610214002713] [Medline: 25573531]
2. Fetherston AA, Rowley G, Allan CL. Challenges in end-of-life dementia care. Evid Based Ment Health 2018 Aug;21(3):107-111. [doi: 10.1136/eb-2018-102889] [Medline: 29776973]

XSL•FO
RenderX

3. Clayton JL, Supiano KP, Aruscavage N, et al. Advance care planning in the context of dementia: defining concordance. Gerontologist 2024 Jun 1;64(6):gnae029. [doi: 10.1093/geront/gnae029] [Medline: 38537649]
4. Yorganci E, Bone AE, Evans CJ, Sampson EL, Stewart R, Sleeman KE. Estimating the escalating future need for palliative care among people living with dementia. Palliat Med 2024 Oct;38(9):1069-1071. [doi: 10.1177/02692163241269773] [Medline: 39132703]
5. Gupta E, Patel P. Palliative care in dementia. Ann Palliat Med 2024 Jul;13(4):791-807. [doi: 10.21037/apm-23-503] [Medline: 37953222]
6. Malhotra C, Yee A, Ramakrishnan C, et al. Development and usability of an advance care planning website (My Voice) to empower patients with heart failure and their caregivers: mixed methods study. JMIR Aging 2024 Dec 18;7:e60117. [doi: 10.2196/60117] [Medline: 39693573]
7. Monnet F, Dupont C, Pivodic L, et al. A user-centred website to support advance care planning for people with dementia and their family caregivers: development and usability study. Digit Health 2025;11:20552076241304018. [doi: 10.1177/20552076241304018] [Medline: 39991741]
8. Dupont C, Smets T, Potts C, et al. Uncovering specific navigation patterns by assessing user engagement of people with dementia and family caregivers with an advance care planning website: quantitative analysis of web log data. JMIR Aging 2025 Feb 11;8:e60652. [doi: 10.2196/60652] [Medline: 39935007]
9. Bosisio F, Sterie AC, Rubli Truchard E, Jox RJ. Implementing advance care planning in early dementia care: results and insights from a pilot interventional trial. BMC Geriatr 2021 Oct 19;21(1):573. [doi: 10.1186/s12877-021-02529-8] [Medline: 34666711]
10. Giotas D, Dening KH. Incorporating advance care planning in dementia care. Nurs Older People 2024 Jul 29;36(4):20-25. [doi: 10.7748/nop.2024.e1463] [Medline: 38715540]
11. Mullick A, Martin J, Sallnow L. An introduction to advance care planning in practice. BMJ 2013 Oct 21;347:f6064. [doi: 10.1136/bmj.f6064] [Medline: 24144870]
12. Morrison RS, Meier DE, Arnold RM. What's wrong with advance care planning? JAMA 2021 Oct 26;326(16):1575-1576. [doi: 10.1001/jama.2021.16430] [Medline: 34623373]
13. Goswami P. Advance care planning and end-of-life communications: practical tips for oncology advanced practitioners. J Adv Pract Oncol 2021;12(1):89-95. [doi: 10.6004/jadpro.2021.12.1.7] [Medline: 33552664]
14. Auriemma CL, O'Donnell H, Klaiman T, et al. How traditional advance directives undermine advance care planning: if you have it in writing, you do not have to worry about it. JAMA Intern Med 2022 Jun 1;182(6):682-684. [doi: 10.1001/jamainternmed.2022.1180] [Medline: 35467697]
15. Piers R, Albers G, Gilissen J, et al. Advance care planning in dementia: recommendations for healthcare professionals. BMC Palliat Care 2018 Jun 21;17(1):88. [doi: 10.1186/s12904-018-0332-2] [Medline: 29933758]
16. van der Steen JT, Nakanishi M, Van den Block L, et al. Consensus definition of advance care planning in dementia: a 33‐country Delphi study. Alzheimers Dement 2024 Feb;20(2):1309-1320. [doi: 10.1002/alz.13526] [Medline: 37985444]
17. Murali KP, Carpenter JG, Kolanowski A, Bykovskyi AG. Comprehensive dementia care models: state of the science and future directions. Res Gerontol Nurs 2025;18(1):7-16. [doi: 10.3928/19404921-20241211-02] [Medline: 39836766]
18. van der Geugten W, Goossensen A. Dignifying and undignifying aspects of care for people with dementia: a narrative review. Scand J Caring Sci 2020 Dec;34(4):818-838. [doi: 10.1111/scs.12791] [Medline: 31750569]
19. Miranda R, Bunn F, Lynch J, Van den Block L, Goodman C. Palliative care for people with dementia living at home: a systematic review of interventions. Palliat Med 2019 Jul;33(7):726-742. [doi: 10.1177/0269216319847092] [Medline: 31057088]
20. Wendrich-van Dael A, Bunn F, Lynch J, Pivodic L, Van den Block L, Goodman C. Advance care planning for people living with dementia: an umbrella review of effectiveness and experiences. Int J Nurs Stud 2020 Jul;107:103576. [doi: 10.1016/j.ijnurstu.2020.103576] [Medline: 32380259]
21. Quinn KL, Shurrab M, Gitau K, et al. Association of receipt of palliative care interventions with health care use, quality of life, and symptom burden among adults with chronic noncancer illness: a systematic review and meta-analysis. JAMA 2020 Oct 13;324(14):1439-1450. [doi: 10.1001/jama.2020.14205] [Medline: 33048152]
22. Kassianos AP, Ioannou M, Koutsantoni M, Charalambous H. The impact of specialized palliative care on cancer patients' health-related quality of life: a systematic review and meta-analysis. Support Care Cancer 2018 Jan;26(1):61-79. [doi: 10.1007/s00520-017-3895-1] [Medline: 28932908]
23. de Sola-Smith K, Gilissen J, van der Steen JT, et al. Palliative care in early dementia. J Pain Symptom Manage 2024 Sep;68(3):e206-e227. [doi: 10.1016/j.jpainsymman.2024.05.028] [Medline: 38848792]
24. Martinsson L, Lundström S, Sundelöf J. Quality of end-of-life care in patients with dementia compared to patients with cancer: a population-based register study. PLoS One 2018;13(7):e0201051. [doi: 10.1371/journal.pone.0201051] [Medline: 30059515]
25. Quinn KL, Wegier P, Stukel TA, Huang A, Bell CM, Tanuseputro P. Comparison of palliative care delivery in the last year of life between adults with terminal noncancer illness or cancer. JAMA Netw Open 2021 Mar 1;4(3):e210677. [doi: 10.1001/jamanetworkopen.2021.0677] [Medline: 33662135]

XSL•FO
RenderX

26. dos Santos EAA. Barriers associated with palliative care in dementia: a review of the literature [Article in Portuguese]. Geriatr Gerontol Aging 2018;12(2):105-112. [doi: 10.5327/Z2447-211520181800014]
27. Nishimura M, Harrison Dening K, Sampson EL, et al. A palliative care goals model for people with dementia and their family: consensus achieved in an international Delphi study. Palliat Med 2024 Apr;38(4):457-470. [doi: 10.1177/02692163241234579] [Medline: 38634232]
28. Timmons S, Fox S, Drennan J, Guerin S, Kernohan WG. Palliative care for older people with dementia-we need a paradigm shift in our approach. Age Ageing 2022 Mar 1;51(3):afac066. [doi: 10.1093/ageing/afac066] [Medline: 35333919]
29. Nakanishi M, Martins Pereira S, Van den Block L, et al. Future policy and research for advance care planning in dementia: consensus recommendations from an international Delphi panel of the European Association for Palliative Care. Lancet Healthy Longev 2024 May;5(5):e370-e378. [doi: 10.1016/S2666-7568(24)00043-6] [Medline: 38608695]
30. Jimenez G, Tan WS, Virk AK, Low CK, Car J, Ho AHY. Overview of systematic reviews of advance care planning: summary of evidence and global lessons. J Pain Symptom Manage 2018 Sep;56(3):436-459.e25. [doi: 10.1016/j.jpainsymman.2018.05.016] [Medline: 29807158]
31. McMahan RD, Tellez I, Sudore RL. Deconstructing the complexities of advance care planning outcomes: what do we know and where do we go? A scoping review. J Am Geriatr Soc 2021 Jan;69(1):234-244. [doi: 10.1111/jgs.16801] [Medline: 32894787]
32. Harris D. Forget me not: palliative care for people with dementia. Postgrad Med J 2007 Jun;83(980):362-366. [doi: 10.1136/pgmj.2006.052936] [Medline: 17551065]
33. Eisenmann Y, Golla H, Schmidt H, Voltz R, Perrar KM. Palliative care in advanced dementia. Front Psychiatry 2020;11:699. [doi: 10.3389/fpsyt.2020.00699] [Medline: 32792997]
34. Vellani S, Puts M, Iaboni A, Degan C, McGilton KS. Integration of a palliative approach in the care of older adults with dementia in primary care settings: a scoping review. Can J Aging 2022 Sep;41(3):404-420. [doi: 10.1017/S0714980821000349] [Medline: 34743774]
35. Arioz U, Allsop MJ, Goodman WD, et al. Artificial intelligence-based approaches for advance care planning: a scoping review. BMC Palliat Care 2025 Oct 23;24(1):268. [doi: 10.1186/s12904-025-01827-x] [Medline: 41131546]
36. DuBose L, Fan Q, Fisher L, et al. Development and evaluation of a web-based platform for personalized educational and professional assistance for dementia caregivers: proposal for a mixed methods study. JMIR Res Protoc 2024 Aug 7;13:e64127. [doi: 10.2196/64127] [Medline: 39110962]
37. Fan Q, Hoang MN, DuBose L, et al. The Olera.care digital caregiving assistance platform for dementia caregivers: preliminary evaluation study. JMIR Aging 2024 Apr 17;7:e55132. [doi: 10.2196/55132] [Medline: 38630527]
38. Gerbaudo-González N, Rodríguez-González R, Facal-Mayo D, Gandoy-Crego M. Personalized interactive multimedia systems to support meaningful activities in dementia care: a systematic review. Arch Gerontol Geriatr 2024 Dec;127:105575. [doi: 10.1016/j.archger.2024.105575]
39. Lott SA, Streel E, Bachman SL, et al. Digital health technologies for Alzheimer's disease and related dementias: initial results from a landscape analysis and community collaborative effort. J Prev Alzheimers Dis 2024;11(5):1480-1489. [doi: 10.14283/jpad.2024.103] [Medline: 39350395]
40. Lyall DM, Kormilitzin A, Lancaster C, et al. Artificial intelligence for dementia—applied models and digital health. Alzheimers Dement 2023 Dec;19(12):5872-5884. [doi: 10.1002/alz.13391] [Medline: 37496259]
41. Mc Carthy M, Schueler P. Editorial: can digital technology advance the development of treatments for Alzheimer's disease? J Prev Alzheimers Dis 2019;6(4):217-220. [doi: 10.14283/jpad.2019.32] [Medline: 31686090]
42. Hooper S, Sabatino CP, Sudore RL. Improving medical-legal advance care planning. J Pain Symptom Manage 2020 Aug;60(2):487-494. [doi: 10.1016/j.jpainsymman.2020.03.004] [Medline: 32240757]
43. Allsop MJ, Birtwistle J, Bennett MI, et al. Optimising digital advance care planning implementation in palliative and end-of-life care: a multi-phase mixed-methods national research programme and recommendations. BMC Med 2025 May 20;23(1):291. [doi: 10.1186/s12916-025-04114-x] [Medline: 40394659]
44. Kilfoy A, Hsu TCC, Stockton-Powdrell C, Whelan P, Chu CH, Jibb L. An umbrella review on how digital health intervention co-design is conducted and described. NPJ Digit Med 2024 Dec 23;7(1):374. [doi: 10.1038/s41746-024-01385-1] [Medline: 39715947]
45. Moyle W, Sriram D, Murfield J, Pu L, Lion K. Co-designed technologies and people living with dementia: a qualitative systematic review. Contemp Nurse 2026 Apr;62(2):216-233. [doi: 10.1080/10376178.2025.2500377] [Medline: 40323305]
46. Bentur N, Sternberg SA. Dementia care in Israel: top down and bottom up processes. Isr J Health Policy Res 2019 Feb 20;8(1):22. [doi: 10.1186/s13584-019-0290-z] [Medline: 30782215]
47. Shatnawi E, Steiner-Lim GZ, Karamacoska D. Cultural inclusivity and diversity in dementia friendly communities: an integrative review. Dementia (London) 2023 Nov;22(8):2024-2046. [doi: 10.1177/14713012231206292] [Medline: 37871120]
48. Serbser-Koal J, Rommerskirch-Manietta M, Purwins D, Roes M. Person-centredness in dementia care: an integrative review of theoretical approaches. BMJ Open 2024 Jul 1;14(6):e085051. [doi: 10.1136/bmjopen-2024-085051] [Medline: 38951009]
49. Monnet F, Diaz A, Gove D, Dupont C, Pivodic L, Van den Block L. The perspectives of people with dementia and their supporters on advance care planning: a qualitative study with the European Working Group of People with Dementia. Palliat Med 2024 Feb;38(2):251-263. [doi: 10.1177/02692163231219915] [Medline: 38293945]

XSL•FO
RenderX

50. Sedini C, Biotto M, Crespi Bel'skij LM, Moroni Grandini RE, Cesari M. Advance care planning and advance directives: an overview of the main critical issues. Aging Clin Exp Res 2022 Feb;34(2):325-330. [doi: 10.1007/s40520-021-02001-y] [Medline: 34655048]
51. AI4HOPE: Transforming Dementia Care. URL: https://www.ai4hope.eu/ [accessed 2026-11-04]
52. Pravilnik o delovanju komisije za etičnost raziskovanja na filozofski fakulteti (KER FF) univerze v mariboru [Report in Slovenian]. : Univerza v Mariboru; 2020 URL: https://ff.um.si/wp-content/uploads/2020/08/Pravilnik-o-delovanju-Komisije-za-eti%C4%8Dnost-raziskovanja-na-Filozofski-fakulteti.pdf [accessed 2026-10-12]
53. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
54. Tai RH, Bentley LR, Xia X, et al. An examination of the use of large language models to aid analysis of textual data. Int J Qual Methods 2024;23. [doi: 10.1177/16094069241231168]
55. Introducing Claude 4. Anthropic. 2025. URL: https://www.anthropic.com/news/claude-4 [accessed 2025-05-25]
56. Castellanos A, Jiang H, Gomes P, Vander Meer D, Castillo A. Large language models for thematic summarization in qualitative health care research: comparative analysis of model and human performance. JMIR AI 2025 Apr 4;4:e64447. [doi: 10.2196/64447] [Medline: 40611510]
57. van der Steen JT, Van den Block L, Nakanishi M, et al. Optimizing advance care planning in dementia: recommendations from a 33-country Delphi study. J Pain Symptom Manage 2025 Jun;69(6):e755-e772. [doi: 10.1016/j.jpainsymman.2025.02.471] [Medline: 40032035]
58. Burghout C, Bolt SR, Nahar-van Venrooij LMW, Smilde TJ, van der Rijt CCD, Wouters EJM. What is the appropriate timing for advance care planning according to patients and their relatives? A scoping review. PLoS One 2026;21(3):e0345093. [doi: 10.1371/journal.pone.0345093] [Medline: 41860815]
59. Hendriks S, Peetoom K, Bakker C, et al. Global prevalence of young-onset dementia: a systematic review and meta-analysis. JAMA Neurol 2021 Sep 1;78(9):1080-1090. [doi: 10.1001/jamaneurol.2021.2161] [Medline: 34279544]
60. Sharma S, Ilse C, Brickell K, et al. Determinants of time to diagnosis in young-onset dementia. Am J Alzheimers Dis Other Demen 2024;39:15333175241309525. [doi: 10.1177/15333175241309525] [Medline: 39689290]
61. Bradshaw A, Birtwistle J, Evans CJ, et al. Factors influencing the implementation of digital advance care planning: qualitative interview study. J Med Internet Res 2024 Aug 16;26:e50217. [doi: 10.2196/50217] [Medline: 39151167]
62. Li W, Wu Y, Song D, et al. How to implement advance care planning for people with dementia: a scoping review. J Adv Nurs 2026 May;82(5):4489-4505. [doi: 10.1111/jan.70083] [Medline: 40693770]
63. Çevik HS, Muente C, Muehlensiepen F, et al. Systems for electronic documentation and sharing of advance care planning preferences: a scoping review. Prog Palliat Care 2024;32(3):149-159. [doi: 10.1080/09699260.2024.2339106] [Medline: 38831759]
64. Torab-Miandoab A, Samad-Soltani T, Jodati A, Rezaei-Hachesu P. Interoperability of heterogeneous health information systems: a systematic literature review. BMC Med Inform Decis Mak 2023 Jan 24;23(1):18. [doi: 10.1186/s12911-023-02115-5] [Medline: 36694161]
65. Pournik O, Ahmad B, Lim Choi Keung SN, et al. Interoperable e-health system using structural and semantic interoperability approaches in CAREPATH. Stud Health Technol Inform 2023 Jun 29;305:608-611. [doi: 10.3233/SHTI230571] [Medline: 37387105]
66. Saripalle R, Runyan C, Russell M. Using HL7 FHIR to achieve interoperability in patient health record. J Biomed Inform 2019 Jun;94:103188. [doi: 10.1016/j.jbi.2019.103188] [Medline: 31063828]
67. Three challenges FHIR 4 faces. Rhapsody. 2019. URL: https://rhapsody.health/blog/three-challenges-fhir-4-faces/ [accessed 2026-11-04]
68. Molnár-Gábor F, Sellner J, Pagil S, Slokenberga S, Tzortzatou-Nanopoulou O, Nyström K. Harmonization after the GDPR? Divergences in the rules for genetic and health data sharing in four member states and ways to overcome them by EU measures: insights from Germany, Greece, Latvia and Sweden. Semin Cancer Biol 2022 Sep;84:271-283. [doi: 10.1016/j.semcancer.2021.12.001] [Medline: 34896635]
69. Friedman B, Hendry DG. Value Sensitive Design: Shaping Technology with Moral Imagination: The MIT Press; 2019. [doi: 10.7551/mitpress/7585.001.0001]
70. Friedman B, Kahn PH, Borning A. Value sensitive design: theory and methods. : University of Washington; 2002 URL: https://dada.cs.washington.edu/research/tr/2002/12/UW-CSE-02-12-01.pdf [accessed 2026-11-04]
71. Alami A, Zahedi M, Ernst N. Are you a real software engineer? Best practices in online recruitment for software engineering studies. Proc IEEE/ACM Int Workshop Methodol Issues Empir Stud Softw Eng 2024:52-57. [doi: 10.1145/3643664.3648207]
72. Reid B, Wagner M, d'Amorim M, Treude C. Software engineering user study recruitment on Prolific: an experience report. In: Proceedings of the 1st International Workshop on Recruiting Participants for Empirical Software Engineering (RoPES '22): Singapore Management University; 2022:1-3. [doi: 10.1145/1122445.1122456]
73. Douglas BD, Ewell PJ, Brauer M. Data quality in online human-subjects research: comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. PLoS One 2023;18(3):e0279720. [doi: 10.1371/journal.pone.0279720] [Medline: 36917576]

XSL•FO
RenderX

Abbreviations

ACP: advance care planning
dACP: digital advance care planning
KMO: Kaiser-Meyer-Olkin
LLM: large language model
PCA: principal component analysis
RQ: research question

Edited by M O'Connell; submitted 11.Feb.2026; peer-reviewed by M Chakit, O Tutsoy, WU Hasan; revised version received 25.Aug.2026; accepted 14.Oct.2026; published 24.Nov.2026.

Please cite as:
Milošič A, Smrke U, Mlakar I
Design Preferences for a Digital Advance Care Planning Tool to Support People Living With Dementia: International Stakeholder Survey and Mixed Methods Analysis
JMIR Aging 2026;9:e93323
URL: https://aging.jmir.org/2026/1/e93323
doi:10.2196/93323

© Ana Milošič, Urška Smrke, Izidor Mlakar. Originally published in JMIR Aging (https://aging.jmir.org), 24.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX